Received December 16, 2018, accepted December 26, 2018, date of publication January 1, 2019, date of current version January 29, 2019.

Digital Object Identifier 10.1109/ACCESS.2018.2890414

Truck Traffic Speed Prediction Under Non-Recurrent Congestion: Based on Optimized Deep Learning Algorithms and GPS Data

JIANDONG ZHAO[1,2], YUAN GAO[3], ZHENZHEN YANG[1], JIANGTAO LI[1], YINGZI FENG[1], ZIYAN QIN[1], AND ZHIMING BAI[3]

[1]School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China
[2]Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport, Ministry of Transport, Beijing Jiaotong University, Beijing 100044, China
[3]School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, Beijing 100044, China

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

This work was supported in part by the Fundamental Research Funds for the Central Universities under Grant 2018YJS138, and in part by the National Natural Science Foundation of China under Grant 71871011, Grant 71871010, Grant 71621001, and Grant 71671014.

ABSTRACT Due to the restriction of traffic management measure in large cities, large heavy-haul trucks can only travel on the circuits and expressways around the city, which often causes congestion in these areas. It is necessary to study the travel speed prediction of trucks on the urban ring road and provide special information services for trucks. Based on the data generated by the trucks driving on the Sixth Ring Road in Beijing, an optimized GRU algorithm is proposed to predict the travel speed of trucks driving on urban express roads under non-recurrent congested conditions. First, a GPS map-matching algorithm that can simultaneously meet the accuracy and efficiency requirements of matching is proposed. Then, the trucks' data traveling on the Sixth Ring Road in Beijing are extracted from the original data. Aiming at getting rid of the abnormal data in GPS data, the screening and processing rules of the abnormal data are made, and then, the traffic speed sequence is extracted. Aiming at the problem that the commonly used weight optimization algorithm SGD cannot adaptively adjust the learning rate, Adam, Adadelta, and Rmsprop are used to optimize the weights in the GRU model in this paper. Considering the four scenarios, including workday, weekend, rainy, and accident, the accuracies of the proposed methods are verified.

INDEX TERMS GPS data, GRU model, optimization algorithm, traffic congestion, traffic speed prediction.

I. INTRODUCTION

Nowadays, freight is an important supply guarantee for urban economic development and public life. However, with the restriction of city traffic management, large heavy-haul trucks can only travel on the circuits and expressways around the city, and they cannot enter the urban area at the fixed time. As a result, the trucks can only gather in the urban peripheral expressways and the areas where the highways and urban roads are connected. The aggregation of trucks and the impact of bad weather often lead to low average speed in these city's entrances and exits areas. In other words, non-recurrent traffic congestion often occurs in these areas, resulting in a decline in the traffic capacity of the urban entrances and exits areas. At the same time, researchers pay more attention to the information service of small passenger cars [1], and there are less specific information services for trucks driver.

As shown in Fig. 1, this paper takes Beijing as an example. The trucks mainly travel on the area of the Sixth Ring Road, highways and ordinary urban roads. According to statistics, in May 2018, the total traffic flow of Sixth Ring Road in Beijing was about 7.59 million, of which trucks accounted for about 17%. From January 1 to August 9, there were 1,442 traffic accidents on the Sixth Ring Road in Beijing in 2018. Trucks caused 404 traffic accidents. The occurrence of these accidents caused non-recurrent congestion on the Sixth Ring Road in Beijing. Therefore, it is necessary to study the speed prediction of trucks to provide information services for the drivers of truck.

In recent years, as China has strengthened the supervision of freight vehicles, it is required that the GPS devices be installed on the large heavy-haul trucks to monitor the running status of trucks. This produces a large amount of trajectory data. Trajectory data, which records locations of moving objects at certain moments, have long been an important means of studying human behavior and solving traffic problems [2]. The convergence of a large number of truck

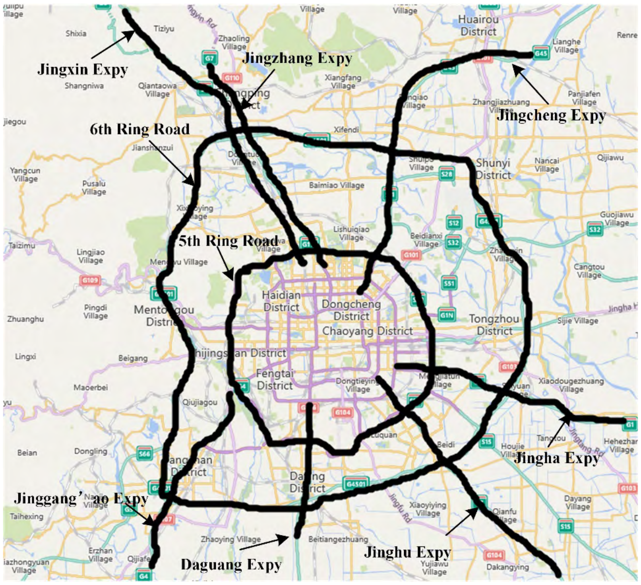

FIGURE 1. **Distribution of urban expressways and highways in Beijing.**

GPS data provides a very good data foundation for analyzing the traffic operation status of trucks and providing effective traffic information services for trucks.

Traffic congestion caused by accidents or bad weather can make a sudden change in the evolution of road traffic flow. It is very important to ensure the robustness of prediction algorithms under traffic congestion conditions.

Therefore, to provide effective information services for truck operations, this paper carries out the identification and prediction of the traffic status of the Sixth Ring Road based on the GPS data of trucks running on the Sixth Ring Road in Beijing. The accuracy of the proposed algorithm is verified in four scenes. It is expected to reduce non-recurrent traffic congestion and the probability of traffic accidents.

The organizational structure of the rest of the paper is as follows: The second part summarizes the existing research, and the third part introduces the preprocessing process of GPS data and related data. The fourth part introduces the GRU algorithm and the related optimization algorithms, the fifth part carries out case verification, and the sixth part is the conclusion.

The research framework of this paper is shown as Fig. 2.

II. RELATED WORKS

A. DATA SOURCE ACQUISITION

The prediction of the speed of the truck is essentially a study of traffic flow parameters prediction. Nowadays, multi-source traffic data are more easily collected than before [3], and many scholars have used different data sources to study prediction of traffic flow parameters. Baek *et al.* [4] used bluetooth beacon system and reverse RFID system to measure the performance travel time of freight vehicle. Wang [5] and Zhao [6] used GPS points speed data to estimate freight car speed. Moniruzzaman *et al.* [7] used volume data from remote traffic microwave sensors combined with GPS data to determine crossover time and establish the Artificial Neural Network Prediction Model. Wang *et al.* [8] used the probing GPS data and loop detector data to predict the travel time of highway freight cars. Spoel *et al.* [9] built prediction models of truck arrival time using actual traffic and weather data. Figliozzi *et al.* [10], Wang and Goodchild [11], Flaskou *et al.* [12], Greaves and Figliozzi [13], and Zhao and Goodchild [14] all used GPS data to study the travel time of trucks.

B. MAP MATCHING

Bierlaire *et al.* [15] proposed a probability mapping matching method based on GPS data obtained from smart phones. Quddus and Washington [16] proposed a new weight-based shortest path and vehicle trajectory aided map-matching (STMM) algorithm based on low frequency GPS data. Li *et al.* [17] developed a new tightly coupled integrity monitoring method for map matching by properly treating the uncertainties from all sources. Zhao *et al.* [18] proposed a local map-matching algorithm that combines online map matching with offline computation. Benny *et al.* [19] tried to find a new way to match 2D local maps with actual GPS trajectories from mobile phones. Chen *et al.* [20] proposed a tightly integrated vehicle navigation system to improve the reliability of the map matching method.

C. PREDICTION METHODS

In terms of prediction and estimation methods of traffic flow parameter, Wang *et al.* [8] proposed a pragmatic multi-regime speed-density relationship-based approach to improve travel time prediction. Figliozzi *et al.* [10] applied statistical techniques to calculate the travel time and reliability of freight vehicles. Wang and Goodchild [11] developed a set of logit models to determine the influencing factors of freight car routes and estimate travel time. Zhao *et al.* [21] proposed a tensor completion method to recover lost RTMS speed and volume data and the optimal K-nearest neighbor algorithm is proposed for travel time prediction. Sun *et al.* [22] used k-nearest neighbor, SVM and random forest to predict the travel time of open-pit tramcars. Zhao *et al.* [23] proposed a multi-dimensional travel time prediction method based on toll collection data and meteorological data of highway. Liu *et al.* [24] proposed Elman neural networks to predict the truck's 85th percentile operating speed. Bederina and Hifi [25] optimized total vehicle travel cost based on the overall method. Zhao *et al.* [26] combined charging data with microwave detection data for travel time prediction. Quiroga and Bullock [27] used GPS and GIS techniques for travel time studies. Zhao *et al.* [28] proposed a gated recurrent unit based method to predicted travel time. Jula *et al.* [29] developed a hybrid approach consisting of DP and genetic algorithms to find the shortest path to the freight car. Yan *et al.* [30] used network flow technology to build a system model to predict freight car travel time, and proposed a simulation-based evaluation method model. Le *et al.* [31] used an Inductive Loop Detector (ILD) network to estimate travel time, using

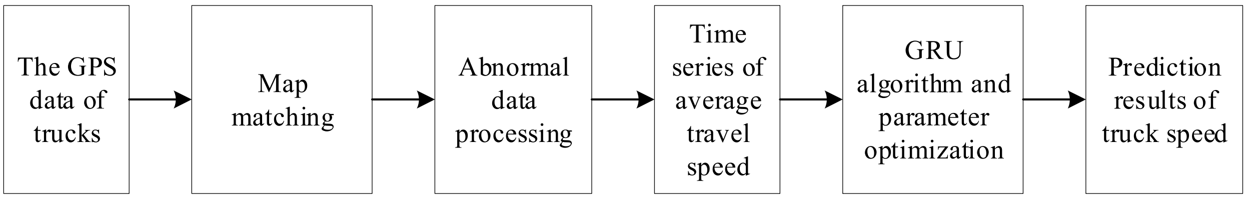

FIGURE 2. Flow chart of GRU model.

Bayesian-based learning methods, fuzzy logic methods, and SVM methods to identify freight cars.

In summary, the data sources used for traffic flow parameter prediction mainly include GPS, remote traffic microwave sensors, Bluetooth system and reverse RFID system. GPS is the mainly used data source for truck speed prediction.

The prediction methods mainly include shallow machine learning methods and deep learning methods. Shallow machine learning methods include neural networks (NNs), statistical modeling, SVM, and K-nearest neighbor algorithm.

The neural networks method has a strong ability to recognize nonlinearities, but the training convergence rate of the model is slow. Statistical modeling has a stable prediction accuracy but poor real-time performance. The SVM algorithm does not need a large number of samples, but the prediction accuracy still needs to be improved under unstable traffic flow. The K-nearest neighbor algorithm does not require modeling and a large number of parameter identification, but requires a large amount of historical data.

At present, with the development of the city size, urban traffic big data [32] has been produced. On this background, for shallow machine learning methods, model training is computationally expensive with frequent updating being prohibitive and the computed forecasting results need to be more accurate. Deep learning method can find a sparse model, which can be frequently updated in real time [33].

Deep learning involves a variety of methods, each with its own advantages in terms of prediction. RNN algorithm is one kind of deep learning method that is more suitable for time series prediction. The GRU algorithm is a relatively variant of the RNN (Recurrent Neural Network).

In the training process of the deep learning model, the traditional SGD algorithm has the problem that individual parameters cannot be adaptively adjusted, which limits its prediction accuracy. With the advent of various optimization methods, such as Adam, Adadelta, Rmsprop, etc., it is possible to adjust the above parameters adaptively.

III. DATA PREPROCESSING

This paper collected GPS data for trucks driving on Beijing's Sixth Ring Road from March 1 to May 31, 2018. The reception cycle of GPS data is about 10 s. As shown in Table 1, the main data fields include Date, Time, CarID, Longitude, Latitude, Speed, Azimuth and Receive Time.

TABLE 1. Field definition of GPS data.

No.	Field Name	Directions
1	Date	YYYYMMDD
2	Time	HHMMSS
3	Car ID	Identification of car
4	Longitude	Correct to six decimal places
5	Latitude	Correct to six decimal places
6	Speed	km/h
7	Azimuth	The north direction is 0 degrees, and the value ranges from 0 to 360 degrees.
8	Receive Time	YYYY-MM-DD HH:MM:SS

A. MAP MATCHING

Analysis of mobile big data has sparked high attentions of academia and industry [34]. Map matching technology is one kind of analysis technique of mobile big data that locates GPS track points on specific road segments in a digital map and then analyzes them using other methods. This paper uses the following algorithm to map matching, and finally obtains the trajectory data of the vehicle traveling on the Sixth Ring Road.

Firstly, as shown in Fig. 3, the original trajectory data are displayed in the GIS, it can be seen that the trajectory data are in the shape of the road network in Beijing. However, the noise data account for about 98%, which is difficult to process directly. At the same time, due to the large amount of data (about 1.5T text file), it is difficult to use GIS and other software for processing. Therefore, this paper uses the corresponding trajectory-matching algorithm to screen out the track points on the Sixth Ring Road section and then analyzes them.

FIGURE 3. Track point before map matching.

FIGURE 4. Track point after map matching. (a) After trajectory matching, (b) Direction of the vehicle.

The specific rules are as follows:

Step1: The map data of the Sixth Ring Road in Beijing is analyzed. Moreover, the map data is divided into square areas with a side length of about 60m, and each area is numbered. At the same time, the sections in each area are counted.

Step2: For a single-track point, the grid area where it is located is determined firstly. If there are road sections contained in the grid or the surrounding eight grids, Step 3 will be carried out; otherwise, the next track point is judged.

Step3: Determine whether the track point is on the Sixth Ring Road. Firstly, determine whether the track point is in the buffer of the Sixth Ring Road. If so, Step 4 will be carried out; otherwise, go back to Step 2.

Step4: If the track point is in the buffer of the Sixth Ring Road, and the driving direction of vehicle meets certain rules, the driving direction of the vehicle can be determined as same as the direction of the road segment; otherwise, multiple track points are used to judge together, as shown in Fig. 4.

Step5: If all the track points are matched, the process of map matching will be ended; otherwise return to Step 2.

The trajectory data after map matching is shown in Fig. 4. It can be seen that almost all the noise data is removed, and the track points obtained are highly matched with the road. At the same time, the direction of the track points can be determined by searching the road where the track points belong. Finally, the distribution of the track points on the Sixth Ring Road can be got, and the next analysis and research can be carried out.

B. PROCESSING OF ABNORMAL DATA

Some abnormal data are inevitable existed in the original data. Abnormal data cannot reflect the normal operating rules of the vehicle. Therefore, it is necessary to carry out the abnormal data processing firstly, so that the data can reflect the real operating state of the traffic flow. What's more, it can improve the prediction accuracy of traffic flow parameter. Therefore, this paper summarizes the expression forms of the abnormal data. Then, according to the expression form of the abnormal data, the corresponding abnormal data processing rules are formulated as follows:

Step1: Repeated data. One valid data is reserved and the other duplicate data are eliminated.

Step2: Different data are collected at the same time of the same car. The speed of multiple data is approximately the same. Therefore, only one valid data is reserved to reflect the status of the vehicle.

Step3: No accident occurs on the road, the latitude and longitude of the vehicle remains unchanged, and the vehicle speed is all zero. This kind of abnormal data needs to be eliminated.

Step4: The latitude and longitude of the vehicle does not change, but the speed value of the vehicle is not 0km/h. The latitude and longitude of the vehicle does not change, indicating that the vehicle did not move. The speed of the vehicle should be zero. Therefore, this kind of abnormal data needs to be eliminated.

Step5: The vehicle's latitude and longitude changes normally, but the vehicle speed is 0km/h. For this case, this paper uses the distance between two adjacent points and time difference to get the speed.

Step6: The speed of the vehicle far exceeds the speed limit. This type of data needs to be eliminated.

In order to verify the processing effect of abnormal data, the two days' data between Xuzhuang Bridge and Shiyuan Bridge on the Sixth Ring Road in Beijing are analyzed. Based on the original data, the average speeds of all vehicles in the road section collected in a unit time are extracted in time sequence to form a time series of average speed. Take the time series of average speed on May 18, 2018, and May 19, 2018, as an example to analyze the processing effect of abnormal data. Fig. 5 shows the traffic speed sequence before and after the process of abnormal data in the non-accident situation on May 18, 2018, with a 5-minute cycle. Fig. 6 shows the sequence of traffic speed before and after the process of abnormal data on May 19, 2018, with a 5-minute cycle. It can be seen in Fig. 5 and Fig. 6 that the processed data are smoother than the original data, and it is proving

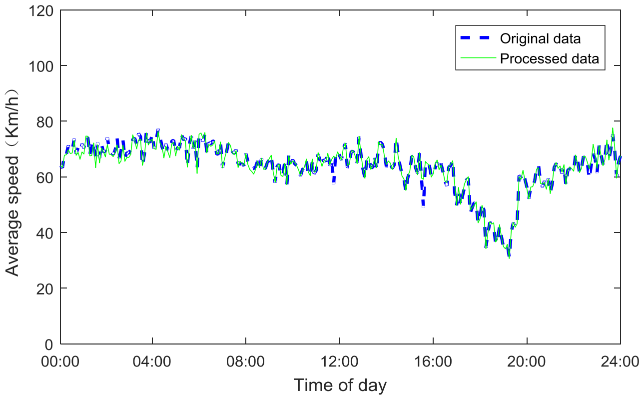

FIGURE 5. The sequence of traffic speed before and after the process of abnormal data on May 18, 2018.

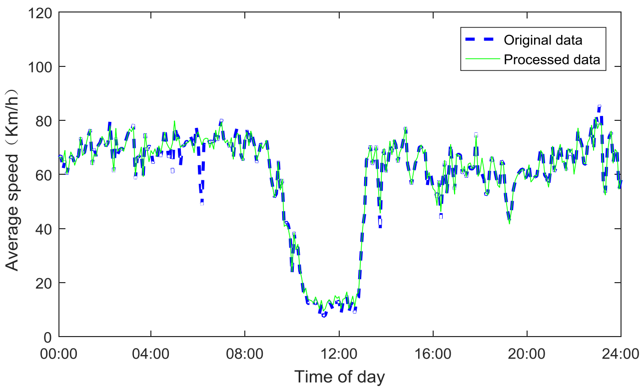

FIGURE 6. The sequence of traffic speed before and after the process of abnormal data on May 19, 2018.

that the data processing reduces the interference of abnormal data on traffic condition identification, and therefore the data cleaning work in this paper is effective.

IV. GRU MODEL WITH OPTIMIZATION ALGORITHM OF PARAMETER

The GRU model is an improvement of Recurrent Neural Network. The hidden layer nodes are connected to each other. Therefore, the structure of the GRU model exhibits the characteristics of circulation. Different from the traditional Recurrent Neural Network, the internal structure of the GRU's hidden layer nodes does not use a single activation function. The principle of the GRU model is described by the following formulas:

$$\tilde{s}_t = \phi(W_S(r_t \odot s_{t-1}) + U_S x_t + b_S) \quad (1)$$

$$s_t = (1 - z_t) \odot s_{t-1} + z_t \odot \tilde{s}_t \quad (2)$$

$$z_t = \sigma(W_z s_{t-1} + U_z x_t + b_z) \quad (3)$$

$$\mathrm{r}_t = \sigma(W_r s_{t-1} + U_r x_t + b_r) \quad (4)$$

where $\odot$ denotes elementwise production; z_t and r_t are update gate and reset gate, respectively. W_z and W_r are weight matrix of update gate and reset gate. W_s is the weight matrix of the output state. x_t is the input data of the moment t. $\tilde{s}_t$ and s_t are the candidate inner state value and output state of the moment t. b_s, b_r, and b_z are constant. σ and ϕ are activation functions that are used to activate the control gate and candidate states, respectively.

The GRU model used in this paper consists of input layer, hidden layer, fully connected layer, and output layer. Its structure is shown in Fig. 7.

The GRU model is essentially a neural network model. Each connection of the neural network model has a specific weight parameter. The purpose of neural network training is to optimize the weight parameter in the model structure based on the initial weight value and find the most suitable weight setting. After knowing the most suitable weight, the entire model can be built. The SGD algorithm is a commonly used optimization algorithm of weighting parameter, which updates each parameter with the same learning rate. However, the deep neural network often contains a large number of parameters, which are not always used. For the parameters that are frequently updated, a lot of knowledge about them have been accumulated. They are not expected to be affected by a single sample seriously. Their learning rate is expected to be slower. For the parameters that are updated occasionally, less information is known about them. It is hoped to get more information about them from every samples that appear occasionally. Therefore, an algorithm that can adaptively set the learning rate is needed. In response to this problem, many variants based on the gradient descent algorithm have been proposed, including Adam, Adadelta, Rmsprop, etc.

This paper focuses on the introduction of Adadelta and Adam algorithms.

The Adadelta algorithm uses mini-batch stochastic gradients g_t to weight the moving average variable s_t by the exponent of the square of the element. At the time step 0, all elements are initialized to zero. Preset hyper-parameter $0 \leq \rho < 1$, at the time step $t > 0$,

$$s_t \leftarrow \rho s_{t-1} + (1 - \rho) g_t \odot g_t \quad (5)$$

In addition, Adadelta maintains an additional state variable Δx_t, whose elements are also initialized to zero at the time step 0. The variation of the independent variable can be calculated by using Δx_{t-1}:

$$g_t' \leftarrow \sqrt{\frac{\Delta x_{t-1} + \epsilon}{s_t + \epsilon}} \odot g_t \quad (6)$$

where ϵ is a constant that is added to maintain numerical stability, for example 10^{-5}. Then, the independent variable is updated:

$$\mathrm{x}_t \leftarrow x_{t-1} - g_t' \quad (7)$$

Finally, we use Δx to record the variation of the independent variable g' by the exponent of the square of the element:

$$\Delta x_t \leftarrow \rho \Delta x_{t-1} + (1 - \rho) g_t' \odot g_t' \quad (8)$$

The Adam (Adaptive Moment Estimation) algorithm can calculate the adaptive learning rate for each parameter. This method not only stores the average exponential attenuation

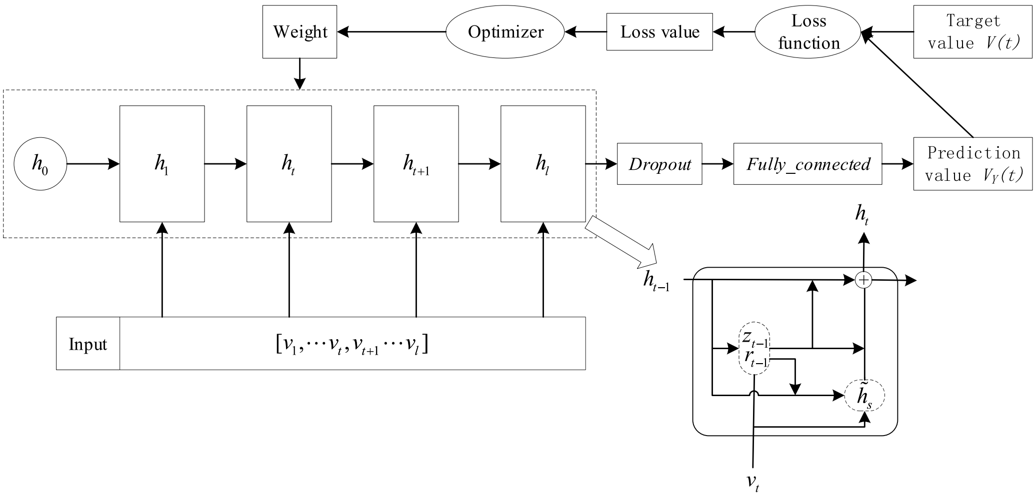

FIGURE 7. Flow chart of GRU model.

value of the previously squared gradient in AdaDelta algorithm, but also maintains the exponential decay average of the previous gradient.

Preset hyper-parameter $0 \leq \beta_1 < 1$. The momentum variable v_t of the time step t is the exponential weighted moving average of the mini-batch random gradient g_t:

$$v_t \leftarrow \beta_1 v_{t-1} + (1-\beta_1) g_t \tag{9}$$

Preset hyper-parameter $0 \leq \beta_2 < 1$. Square the mini-batch random gradient by the element to get $g_t \odot g_t$. s_t is the exponential weighted moving average of $g_t \odot g_t$:

$$s_t \leftarrow \beta_2 s_{t-1} + (1-\beta_2) g_t \odot g_t \tag{10}$$

Since the elements in both v_0 and s_0 is initialized to zero, in time step t, $v_t = (1-\beta_1)\sum_{i=1}^{t} \beta_1^{t-i} g_i$ is got. Adding the weights of mini-batch random gradients in the past time steps together, $(1-\beta_1)\sum_{i=1}^{t} \beta_1^{t-i} = 1-\beta_1^t$ is got. It should be noted that when t is small, the sum of mini-batch random gradient weights in each time step will be small. For example, when $\beta_1 = 0.9$, $v_1 = 0.1 g_1$. In order to eliminate such an effect, for any time step t, v_t is divided by $1-\beta_1^t$, so that the sum of the mini-batch random gradient weights in each time step is 1. This is also called deviation correction. In the Adam algorithm, deviation corrections are made for both variables v_t and s_t:

$$\hat{v}_t \leftarrow \frac{v_t}{1-\beta_1^t} \tag{11}$$

$$\hat{s}_t \leftarrow \frac{s_t}{1-\beta_2^t} \tag{12}$$

Next, the Adam algorithm uses the $\hat{v}_t$ and $\hat{s}_t$ to re-adjust the learning rate of each model parameters by elemental operations:

$$g'_t \leftarrow \frac{\eta \hat{v}_t}{\sqrt{\hat{s}+\epsilon}} \tag{13}$$

where η is the learning rate, ϵ is a constant added to maintain numerical stability. For example, it can be 10^{-8}. Like Adadelta, each element in the objective function argument has its own learning rate. Finally, use g'_t to iterate independent variable:

$$x_t \leftarrow x_{t-1} - g'_t \tag{14}$$

V. CASE STUDY

The GPS data used in this paper is derived from trucks driving on the Sixth Ring Road in Beijing.

It is a circular highway with a total length of 187.6 kilometers. The Sixth Ring Road is a bi directional and four lanes highway. It is connected to seven radial highway and many national highways in Beijing, 15-20 km from the downtown. The traffic conditions on the Sixth Ring Road are good, and the traffic flow is always smooth. However, sometimes the large trucks overtake each other, which may result in traffic congestion in a small area. Trucks account for 17% of the total traffic on the Sixth Ring Road. Fig. 8 shows the trend curves of the total flow of vehicles on the Sixth Ring Road from May 1, 2018 (Tuesday) to May 31. It can be seen that the traffic flow is obviously high on the festival (International Workers' Day). The traffic flow on the Sixth Ring Road is lower from Monday to Thursday, and higher from Friday to Sunday.

Fig. 9 shows the statistics of truck accidents. 93 accidents occurred due to weather, 79 accidents occurred on weekends, 26 accidents occurred on holidays, and 325 accidents occurred on working days. According to the analysis, there were 21 accidents caused by weather on the weekend, and 8 accidents caused by weather during the holidays.

The time range for the selected data is from March 1, 2018, to May 31, 2018. In this paper, the road section between two neighboring interchanges is taken as the basic research sections. Due to the dense interchanges at the Southeast Sixth Ring Road and the high frequency of accidents, as shown in Fig. 10, two sections of the Southeast Sixth Ring Road are selected as experimental verification segments, respectively.

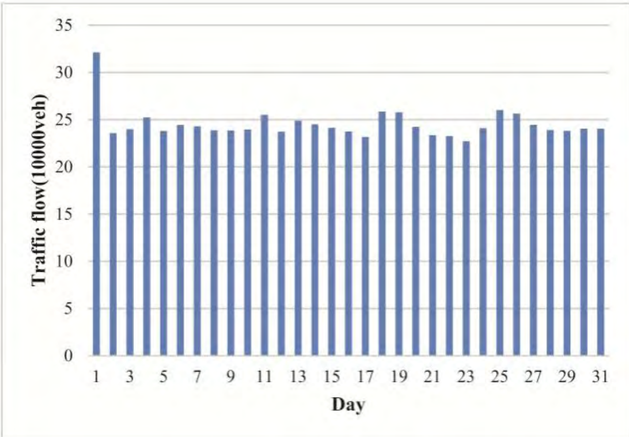

FIGURE 8. Traffic Statistics from May 1 to May 31, 2018.

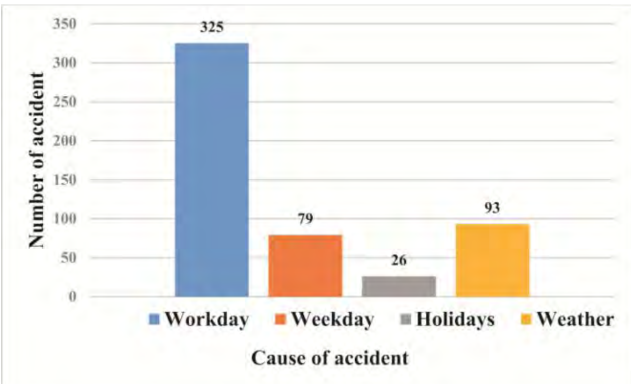

FIGURE 9. Accident statistics of trucks running on the sixth ring road.

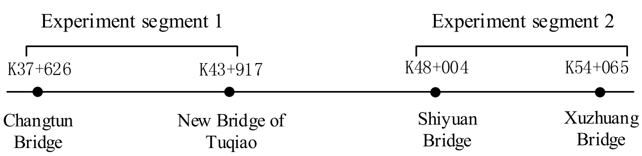

FIGURE 10. Schematic diagram of the experiment segment.

Tuqiao New Bridge to Changtun Bridge is selected as experiment segment 1 and Xuzhuang Bridge to Shiyuan Bridge is selected as experiment segment 2.

The length of the experiment segment 1 is 6.291 km. The length of the experiment segment 2 is 6.061 km. The two experiment segments are bi directional and four lanes with a speed limit of 80km/h. In the experimental data, the first 80% of the data are used as training data, and the last 20% of the data are used as test data.

The indexes of MAPE and RMSE are selected to evaluate the prediction accuracy of the model. The calculation methods of MAPE and RMSE are shown as equation (15) and (16).

$$MAPE = \frac{1}{L}\sum_{L}\frac{|V_Y(t) - V(t)|}{V(t)} \tag{15}$$

$$RMSE = \sqrt{\frac{1}{L}\sum_{L}(V_Y(t) - V(t))^2} \tag{16}$$

TABLE 2. Standard for the classification of congestion levels on highway sections.

Congestion level	Mark	Design speed (km/h) 120 Speed (km/h)	100 Speed (km/h)	80 Speed (km/h)
Unblocked	5	⩾90	⩾80	⩾60
Basically unblocked	4	[70, 90)	[60, 80)	[50, 60)
Slight congestion	3	[50, 70)	[40, 60)	[35, 50)
Moderate congestion	2	[30, 50)	[20, 40)	[20, 35)
Severe congestion	1	[0, 30)	[0, 20)	[0, 20)

where $V_Y(t)$ is the predicted speed at time t, $V(t)$ is the actual speed at time t, and L is the total number of prediction periods.

This paper focuses on the prediction of the travel speed of trucks under non-recurrent congestion. Table 2 is the division standard of the Ministry of Transport for the congestion level of expressway segments, in the "Temporary Technical Requirements for Operation and Monitoring of Highway Network Operation". In this paper, the design speed of the road section is 80km/h, and the congestion degree is divided by the standard.

The structure of the GRU model in this paper includes the input layer (12 neurons), the GRU layer (12 neurons), the Dropout layer (64 neurons), and the output layer (1 neuron). The activation function uses a linear activation function.

In this paper, the accuracy of the proposed algorithm is verified in four scenes: the weekday, the weekend, the rainy day and the day of accident. Use the data of the experiment segment 1 on May 24, May 18, May 19, and May 20 for verification. Use the data of the experiment segment 2 on May 18, May 23, May 20, and May 21 for verification.

A. EXPERIMENT SEGMENT 1

1) THE WEEKDAY

The predicted results of the experiment segment 1 on May 24 (the weekday) are shown in Fig. 11. The prediction errors are shown in Fig. 12. Combined with the congestion evaluation standard, it can be seen that the section is in a mild congestion and moderate congestion state at the morning peak on May 24 and the section is in an unblocked and basic unblocked state during the rest of the period. Comparing various optimization algorithms, the GRU algorithm can get good prediction results. The results of the error evaluation indicator MAPE, RMSE are shown in Table 3. MAPE using "Adadelta", "Rmsprop", "Adam", and "Sgd" were 8.26%, 8.48%, 8.20%, and 8.90%, respectively. RMSE using "Adadelta", "Rmsprop", "Adam", and "Sgd" were 5.63, 5.73, 5.51, and 5.77, respectively.

TABLE 3. Error of experiment segment 1.

Evaluation index	Workday 5.24		Weekend 5.19		Accident 5.18		Rainy 5.20	
	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
Adadelta	8.26%	5.63	8.66%	6.06	17.76%	6.93	7.61%	6.05
Rmsprop	8.48%	5.73	9.07%	6.20	16.91%	6.96	7.92%	6.21
Adam	8.20%	5.51	8.47%	5.94	18.19%	6.95	7.40%	5.87
Sgd	8.90%	5.77	8.64%	6.14	21.49%	7.65	8.13%	6.19

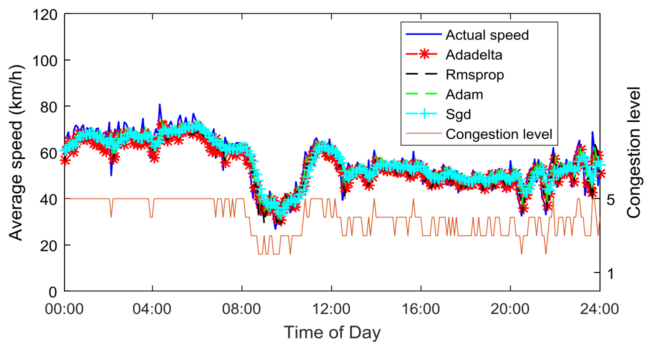

FIGURE 11. Prediction results on May 24, 2018 (workday), experiment segment 1.

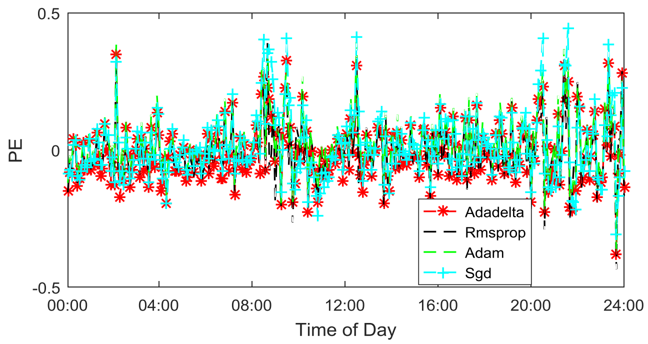

FIGURE 12. Errors on May 24, 2018 (workday), experiment segment 1.

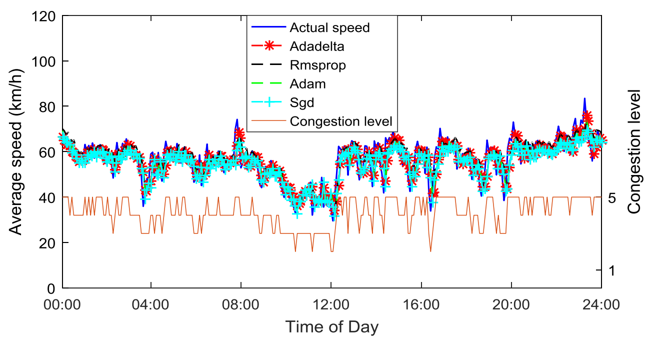

FIGURE 13. Prediction results on May 19, 2018 (weekend), experiment segment 1.

2) THE WEEKEND

The predicted results of the experiment segment 1 on May 19 (the weekend) are shown in Fig. 13. The errors are shown in Fig. 14. Combined with the congestion evaluation standard, it can be seen that the section is in an unblocked and basic unblocked state during most of the time on May 19,

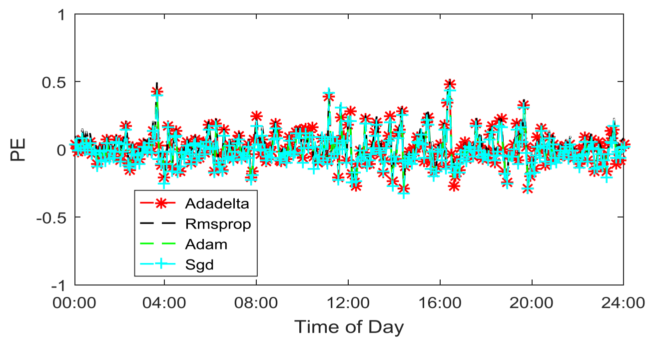

FIGURE 14. Errors on May 19, 2018 (weekend), experiment segment 1.

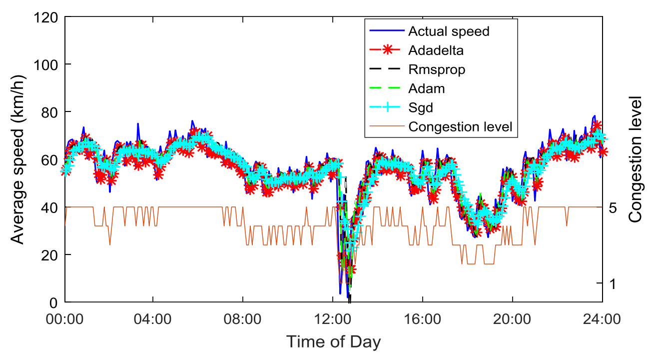

FIGURE 15. Prediction results on May 18, 2018 (accident), experiment segment 1.

and the section is in a mild congestion and moderate congestion state at noon. The results of the error evaluation indicator MAPE, RMSE are shown in Table 1. MAPE using "Adadelta", "Rmsprop", "Adam", and "Sgd" were 8.66%, 9.07%, 8.47%, and 8.64%, respectively. RMSE using "Adadelta", "Rmsprop", "Adam", and "Sgd" were 6.06, 6.20, 5.94, and 6.14, respectively.

3) THE DAY OF ACCIDENT

At 12 o'clock on May 18, a traffic accident on experiment segment 1 caused heavy congestion. The predicted results of the section 1 on May 18 (the day of accident) are shown in Fig. 15. The errors are shown in Fig. 16. The results of the error evaluation indicator MAPE, RMSE are shown in Table 1. MAPE using "Adadelta", "Rmsprop", "Adam", and "Sgd" were 17.76, 16.91%, 18.19%, and 21.49%, respectively. RMSE using "Adadelta", "Rmsprop", "Adam", and "Sgd" were 6.93, 6.96, 6.95, and 7.65, respectively.

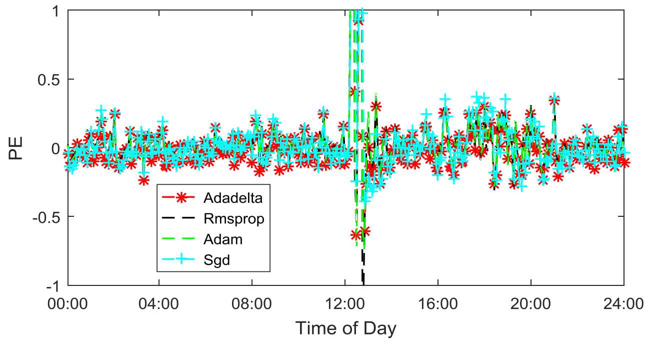

FIGURE 16. Errors on May 18, 2018 (accident), experiment segment 1.

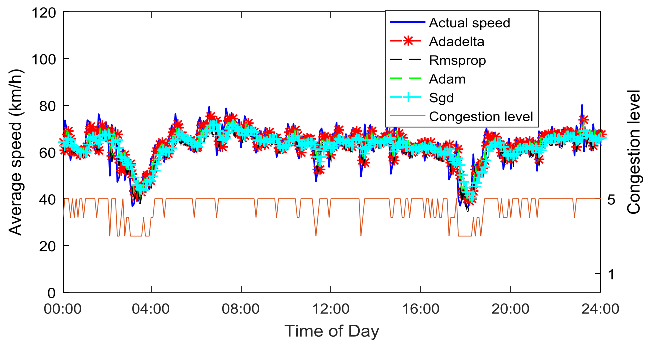

FIGURE 17. Prediction results on May 20, 2018 (rainy), experiment segment 1.

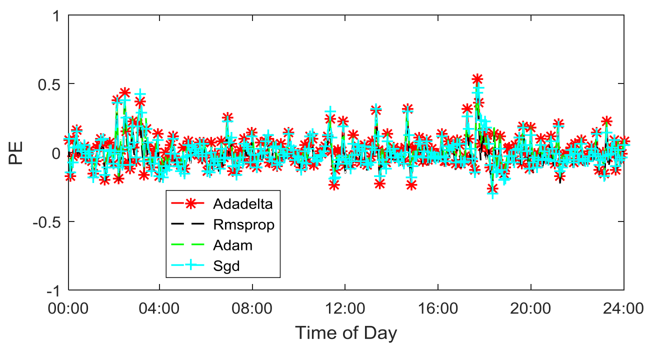

FIGURE 18. Errors on May 20, 2018 (rainy), experiment segment 1.

4) THE RAINY DAY

On May 20, the weather in the area near the experiment segment 1 was light rain. The predicted results of the section 1 on May 20 (the rainy day) are shown in Fig. 17. The errors are shown in Fig. 18. The results of the error evaluation indicator MAPE, RMSE are shown in Table 1. MAPE using "Adadelta", "Rmsprop", "Adam", and "Sgd" were 7.61%, 7.92%, 7.40%, and 8.13%, respectively. RMSE using "Adadelta", "Rmsprop", "Adam", and "Sgd" were 6.05, 6.21, 5.87, and 6.19, respectively.

B. EXPERIMENT SEGMENT 2

1) THE WEEKDAY

The predicted results of the experiment segment 2 on May 18 (the weekday) are shown in Fig. 19. The errors are shown in Fig. 20. Combined with the congestion evaluation standard,

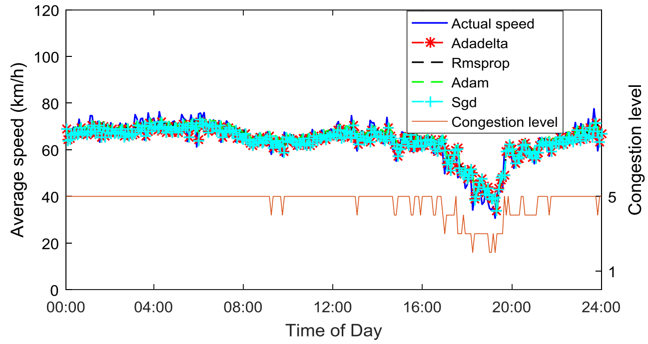

FIGURE 19. Prediction results on May 18, 2018 (workday), experiment segment 2.

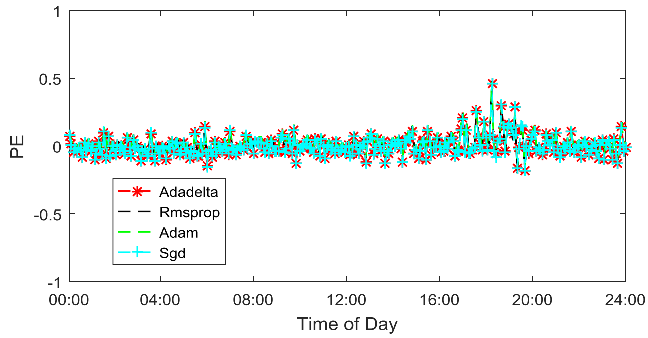

FIGURE 20. Errors on May 18, 2018 (workday), experiment segment 2.

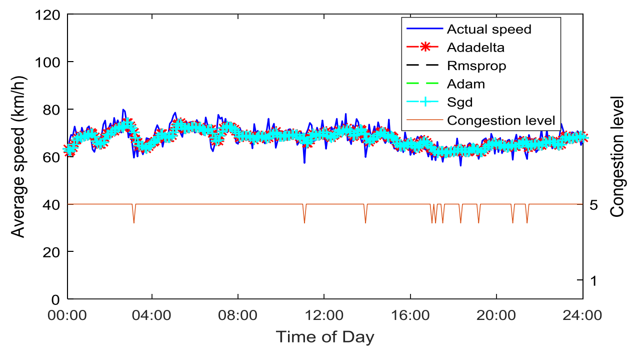

FIGURE 21. Prediction results on May 20, 2018 (weekend), experiment segment 2.

it can be seen that the section is in a mild congestion and moderate congestion state at the evening peak on May 18, and the section is in an unblocked and basic unblocked state during the rest of the period. The results of the error evaluation indicator MAPE, RMSE are shown in Table 1. MAPE using "Adadelta", "Rmsprop", "Adam", and "Sgd" were 5.12%, 5.34%, 5.52%, and 5.44%, respectively. RMSE using "Adadelta", "Rmsprop", "Adam", and "Sgd" were 3.97, 4.12, 4.23, and 4.19, respectively.

2) THE WEEKEND

The predicted results of the experiment segment 2 on May 20 (the weekend) are shown in Fig. 21. The errors are shown in Fig. 22. Combined with the congestion evaluation standard, it can be seen that the section is in an unblocked and basic unblocked state during the entire day on May 20. The results

TABLE 4. Error of experiment segment 2.

Evaluation index	Workday 5.18		Weekend 5.20		Accident 5.23		Rainy 5.21	
	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE	MAPE	RMSE
Adadelta	5.12%	3.97	4.63%	3.84	5.90%	4.52	4.62%	3.86
Rmsprop	5.34%	4.12	4.66%	3.88	5.85%	4.58	4.63%	3.93
Adam	5.52%	4.23	4.66%	3.88	5.85%	4.58	4.41%	3.75
Sgd	5.44%	4.19	4.64%	3.85	6.51%	5.28	5.02%	4.06

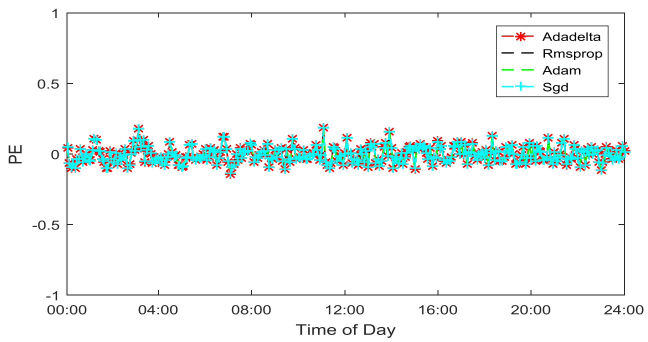

FIGURE 22. Errors on May 20, 2018 (weekend), experiment segment 2.

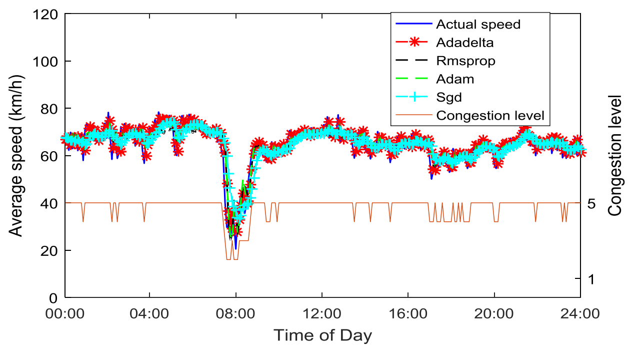

FIGURE 23. Prediction results on May 23, 2018 (accident), experiment segment 2.

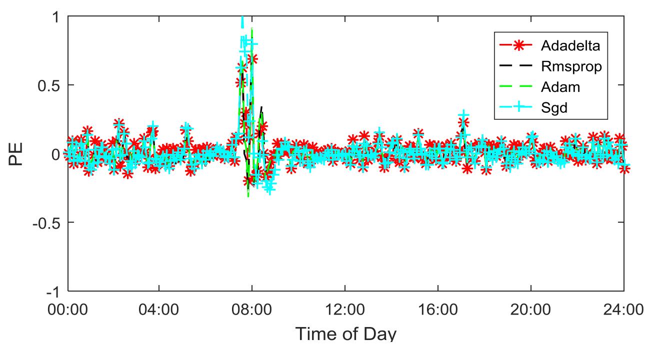

FIGURE 24. Errors on May 23, 2018 (accident), experiment segment 2.

of the error evaluation indicator MAPE, RMSE are shown in Table 1. MAPE using "Adadelta", "Rmsprop", "Adam", and "Sgd" were 4.63%, 4.66%, 4.66%, and 4.64%, respectively. RMSE using "Adadelta", "Rmsprop", "Adam", and "Sgd" were 3.84, 3.88, 3.88, and 3.85, respectively.

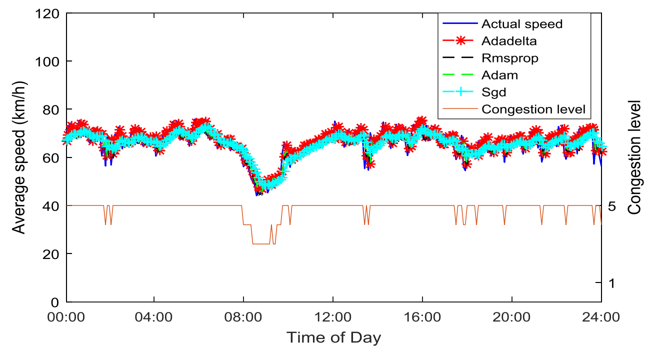

FIGURE 25. Prediction results on May 21, 2018 (rainy), experiment segment 2.

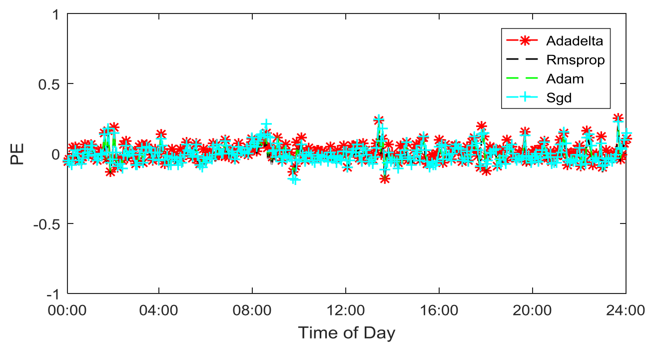

FIGURE 26. Errors on May 21, 2018 (rainy), experiment segment 2.

3) THE DAY OF ACCIDENT

At 10 o'clock on May 23, a traffic accident on experiment segment 2 caused heavy congestion. The predicted results of the section 2 on May 23 (the day of accident) are shown in Fig. 23. The errors are shown in Fig. 24. The results of the error evaluation indicator MAPE, RMSE are shown in Table 1. MAPE using "Adadelta", "Rmsprop", "Adam", and "Sgd" were 5.90%, 5.85%, 5.85%, 6.51%, respectively. RMSE using "Adadelta", "Rmsprop", "Adam", and "Sgd" were 4.52, 4.58, 4.58, and 5.28 respectively.

4) THE RAINY DAY

On May 21, the weather in the area near the experiment segment 2 was light rain. The predicted results of the section 2 on May 21 (the rainy day) are shown in Fig. 25. The errors are shown in Fig. 26. The results of the error evaluation indicator MAPE, RMSE are shown in Table 1. MAPE using "Adadelta", "Rmsprop", "Adam", and "Sgd" were 4.62%, 4.63%, 4.41%, 5.02%, respectively. RMSE using

"Adadelta", "Rmsprop", "Adam", and "Sgd" were 3.86, 3.93, 3.75, and 4.06, respectively.

As can be seen from the above chart, various optimization methods can get good prediction results on weekdays and weekends. Among them, the algorithms of "Adadelta" and "Adam" have higher precision. In this paper, the selected rainy days are light rain, which has less impact on actual road traffic. Therefore, in this scene, the prediction effect of the algorithm is similar to that on weekdays and weekends. The algorithms of "Adadelta" and "Adam" have higher precision. It can be seen from the prediction results of the accident data of the two sections that the accuracy of the SGD is significantly lower than other optimization methods under non-recurrent congestion.

VI. CONCLUSION

Large heavy-haul trucks can only travel on the circuits and expressways around the city due to the restriction of traffic management measure in large cities, which often causes non-recurrent congestion in these areas. At the same time, researchers are paying more attention to the information service of small passenger cars. The special information services for trucks is relatively less. Therefore, the optimized GRU algorithm is proposed to predict the travel speed of trucks driving on urban express roads under non-recurrent congested conditions. This paper aims to provide an effective information service for truck drivers.

This paper verifies the algorithm based on the data generated by the trucks driving on the Sixth Ring Road in Beijing. The truck driving in Beijing generates the original data used in this paper. Aiming at the problem of a large amount of original data, this paper proposes a GPS map-matching algorithm that can simultaneously meet the accuracy and efficiency of matching. Then, the truck data traveling in Beijing Sixth Ring Road are extracted from the original data. Aiming at getting rid of the abnormal data in GPS data, the screening and processing rules of abnormal data are made, and then the traffic speed sequence is extracted.

The obtained historical time series of average travel speed is input into the GRU model. Aiming at the problem that the commonly used weight optimization algorithm SGD cannot adaptively adjust the learning rate, Adam, Adaddala, and Rmsprop are used to optimize the weights in the GRU model in this paper. Considering the four scenarios including workday, weekend, rainy, and accident, the accuracy of the proposed algorithm is verified. The GRU model with "Adadelta" and "Adam" algorithms have higher precision on weekdays, weekends and light rain scenes. On the accident scenes, although the accuracy of all the algorithms is reduced, the accuracy of the GRU model with SGD algorithms is obviously the lowest than other algorithms. The algorithm with the best optimization result is "Rmsprop".

REFERENCES

[1] 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. Sci. Eng.*, vol. 15, no. 4, pp. 1507–1520, Oct. 2018.

[2] X. Kong *et al.*, "Big trajectory data: A survey of applications and services," *IEEE Access*, vol. 6, pp. 58295–58306, 2018.

[3] F. Xia, J. Wang, X. Kong, Z. Wang, J. Li, and C. Liu, "Exploring human mobility patterns in urban scenarios: A trajectory data perspective," *IEEE Commun. Mag.*, vol. 56, no. 3, pp. 142–149, Mar. 2018.

[4] J. Baek, Y. Choi, J. Jung, and C. Lee, "Performance comparison of Bluetooth beacon and reverse RFID systems as potential tools for measuring truck travel time in open-pit mines: A simulation experiment," *Geosyst. Eng.*, vol. 21, no. 1, pp. 43–52, 2018.

[5] Z. Wang, A. Goodchild, and E. McCormack, "Measuring truck travel time reliability using truck probe GPS data," *J. Intell. Transp. Syst.*, vol. 20, no. 2, pp. 103–112, Mar. 2016.

[6] W. J. Zhao, A. V. Goodchild, and E. D. McCormack, "Evaluating the accuracy of spot speed data from global positioning systems for estimating truck travel speed," *Trans. Res. Rec., J. Transp. Res. Board*, vol. 2246, pp. 101–110, 2011.

[7] M. Moniruzzaman, H. Maoh, and W. Anderson, "Short-term prediction of border crossing time and traffic volume for commercial trucks: A case study for the Ambassador Bridge," *Transp. Res. C, Emerg. Technol.*, vol. 63, pp. 182–194, Feb. 2016.

[8] Z. Wang, A. V. Goodchild, and E. McCormack, "Freeway truck travel time prediction for freight planning using truck probe GPS data," *Eur. J. Transp. Infrastruct. Res.*, vol. 16, no. 1, pp. 76–94, Jan. 2016.

[9] V. der Spoel, C. Amrit, and J. van Hillegersberg, "Predictive analytics for truck arrival time estimation: A field study at a European distribution centre," *Int. J. Prod. Res.*, vol. 55, no. 17, pp. 5062–5078, Sep. 2017.

[10] M. Figliozzi, N. Wheeler, E. Albright, L. Walker, S. Sarkar, and D. Rice, "Algorithms for studying the impact of travel time reliability along multisegment trucking freight corridors," *Trans. Res. Rec., J. Transp. Res. Board*, vol. 2224, pp. 26–34, Oct. 2011.

[11] Z. Wang and A. V. Goodchild, "GPS data analysis of the impact of tolling on truck speed and routing: Case study in Seattle, Washington," *Trans. Res. Rec., J. Transp. Res. Board*, vol. 2411, no. 1, pp. 112–119, 2014.

[12] M. Flaskou, M. A. Dulebenets, M. M. Golias, S. Mishra, and R. M. Rock, "Analysis of freight corridors using GPS data on trucks," *Trans. Res. Rec., J. Transp. Res. Board*, vol. 2478, pp. 113–122, Sep. 2015.

[13] S. P. Greaves and M. A. Figliozzi, "Collecting commercial vehicle tour data with passive global positioning system technology: Issues and potential applications," *Trans. Res. Rec., J. Transp. Res. Board*, vol. 2049, pp. 158–166, Sep. 2008.

[14] W. Zhao and A. V. Goodchild, "Truck travel time reliability and prediction in a port drayage network," *Maritime Econ. Logistics*, vol. 13, no. 4, pp. 387–418, 2011.

[15] M. Bierlaire, J. Chen, and J. Newman, "A probabilistic map matching method for smartphone GPS data," *Transp. Res., C, Emerg. Technol.*, vol. 26, pp. 78–98, Jan. 2013.

[16] M. Quddus and S. Washington, "Shortest path and vehicle trajectory aided map-matching for low frequency GPS data," *Transp. Res., C, Emerg. Technol.*, vol. 55, pp. 328–339, Jun. 2015.

[17] L. Li, M. Quddus, and L. Zhao, "High accuracy tightly-coupled integrity monitoring algorithm for map-matching," *Transp. Res. C, Emerg. Technol.*, vol. 36, pp. 13–26, Nov. 2013.

[18] D. Zhao, X. Liu, and L. Guo, "Real time map matching algorithm of floating car in support of spatial grid index," *J. Comput.-Aided Des. Comput. Graph.*, vol. 26, no. 9, pp. 1550–1556, Sep. 2014.

[19] B. Hardjono, A. Wibowo, M. F. Rachmadi, and W. Jatmiko, "Mobile phones as traffic sensors with map matching and privacy considerations," in *Proc. Int. Symp. Micro-NanoMechatronics Hum. Sci. (MHS)*, Nagoya, Japan, Nov. 2012, pp. 450–455.

[20] W. Chen, Z. Li, M. Yu, and Y. Chen, "An integrated map-match algorithm with position feedback and shape-based mismatch detection and correction," *J. Intell. Transp. Syst.*, vol. 12, no. 4, pp. 168–175, Oct. 2008.

[21] J. Zhao, Y. Gao, J. Tang, L. Zhu, and J. Ma, "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.

[22] X. Sun, H. Zhang, F. Tian, and L. Yang, "The use of a machine learning method to predict the real-time link travel time of open-pit trucks," *Math. Problems Eng.*, vol. 2018, Apr. 2018, Art. no. 4368045.

[23] J. Zhao *et al.*, "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.

[24] X. Liu, K. Zhang, and M. Li, "Application of Elman neural networks to predict truck's operating speed," *Adv. Mater. Res.*, vols. 694–697, pp. 2846–2849, May 2013.

[25] H. Bederina and M. Hifi, "A hybrid multi-objective evolutionary optimization approach for the robust vehicle routing problem," *Appl. Soft Comput.*, vol. 71, pp. 980–993, Oct. 2018.

[26] 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.

[27] C. A. Quiroga and D. Bullock, "Travel time studies with global positioning and geographic information systems: An integrated methodology," *Transp. Res. C, Emerg. Technol.*, vol. 6, nos. 1–2, pp. 101–127, Feb. 1998.

[28] J. Zhao, Y. Gao, Y. Qu, H. Yin, L. Liu, and H. Sun, "Travel time prediction: Based on gated recurrent unit method and data fusion," *IEEE Access*, vol. 6, pp. 70463–70472, 2018.

[29] H. Jula, M. Dessouky, P. Ioannou, and A. Chassiakos, "Container movement by trucks in metropolitan networks: Modeling and optimization," *Transp. Res. E, Logistics Transp. Rev.*, vol. 41, no. 3, pp. 235–259, May 2005.

[30] S. Yan, H. C. Lin, and X. Y. Jiang, "A planning model with a solution algorithm for ready mixed concrete production and truck dispatching under stochastic travel times," *Eng. Optim.*, vol. 44, no. 4, pp. 427–447, 2012.

[31] C. Le Bastard, D. Guilbert, A. Boubezoul, S.-S. Ieng, Y. Wang, and A. Delepoulle, "Vehicule identification from inductive loops application: Travel time estimation for a mixed population of cars and trucks," in *Proc. 14th IEEE Int. Intell. Transp. Syst. Conf.*, Paris, France, Oct. 2011, pp. 507–512.

[32] F. Xia, A. Rahim, X. Kong, M. Wang, Y. Cai, and J. Wang, "Modeling and analysis of large-scale urban mobility for green transportation," *IEEE Trans. Ind. Informat.*, vol. 14, no. 4, pp. 1469–1481, Apr. 2018.

[33] L. Lin, Z. He, and S. Peeta, "Predicting station-level hourly demand in a large-scale bike-sharing network: A graph convolutional neural network approach," *Transp. Res. C, Emerg. Technol.*, vol. 97, pp. 258–276, Dec. 2018.

[34] 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.

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 in vehicle operation engineering with the School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, China. His research interests include traffic control and traffic flow prediction.

ZHENZHEN YANG received the B.S. degree from the School of Civil Engineering and Architecture, Hubei University of Technology, in 2009, and the M.S. degree from the School of Management, University of Shanghai for Science and Technology, in 2012. She is currently pursuing the Ph.D. degree with the School of Traffic and Transportation, Beijing Jiaotong University, Beijing, China. Her research interests include intelligent transportation systems, traffic big data analysis, and machine learning and its application.

JIANGTAO LI received the B.E. degree from Fuzhou University, in 2017. He is currently pursuing the master's degree in traffic engineering with the School of Traffic and Transportation, Beijing Jiaotong University, China. His research interest includes data mining.

YINGZI FENG received the B.S. degree from the Hebei University of Economics and Business, in 2018. She is currently pursuing the master's degree in transportation planning and management with the School of Traffic and Transportation, Beijing Jiaotong University, China. Her research interest includes traffic big data analysis.

ZIYAN QIN received the B.S. degree from the Xi'an University of Architecture and Technology, in 2018. She is currently pursuing the master's degree in transportation engineering with the School of Traffic and Transportation, Beijing Jiaotong University, China. Her research interest includes traffic data analysis.

ZHIMING BAI received the B.S. degree from Lanzhou Jiaotong University, in 2016. He is currently pursuing the master's degree in mechanical engineering with the School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, China. His research interests include traffic control and traffic flow prediction.