

Received April 15, 2020, accepted April 26, 2020, date of publication May 4, 2020, date of current version May 18, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2992070

Lane Work-Schedule of Toll Station Based on Queuing Theory and PSO-LSTM Model

PENG WANG¹, JIANDONG ZHAO^{1,2,3}, YUAN GAO¹, MIGUEL ANGEL SOTELO⁴, AND ZHIXIONG LI⁵

¹School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, Beijing 100044, China

²School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China

³Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport, Ministry of Transport, Beijing Jiaotong University, Beijing 100044, China

⁴Department of Computer Engineering, University of Alcalá, 28801 Alcalá de Henares, Spain

⁵School of Mechanical, Materials, Mechatronic and Biomedical Engineering, University of Wollongong, Wollongong, NSW 2522, Australia

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

This work was supported by the National Natural Science Foundation of China under Grant 71871011, Grant 71890970/71890972, and Grant 71621001.

ABSTRACT A reasonable lane work-schedule in each time period can not only guarantee the traffic efficiency of toll stations, but also reduce the operating cost of toll stations. This paper proposes a comprehensive solution for lane work plan. Firstly, the average queue length is selected as a good index for measuring the congestion of toll station. And then, based on the queuing theory, the service level of toll station is divided into four levels according to the relationship between the average queue length and traffic capacity. Secondly, based on the toll data, a toll station congestion prediction model is established with the Long Short-Term Memory model (LSTM) and the particle swarm optimization (PSO) algorithm. In this model, the average queue length, service time and traffic volume are selected as three inputs, the average queue length value of the next hour is the output. Thirdly, on the basis of meeting the secondary level service level of toll stations, the lane work-schedule model is established. Then, the number of lanes opened in each time period can be calculated by using this model and congestion prediction results. Fourthly, considering the two scenarios of weekday and weekend, the effectiveness of the methods proposed in this paper is analyzed and verified with the toll data of the Dongshe, Changfeng, and Linfen toll stations in Shanxi Province. Finally, based on operating costs analysis, the results show that the proposed solution could effectively realize the reasonable work-schedule of the toll station.

INDEX TERMS Lane work-schedule, LSTM model, queuing theory, toll data, traffic management.

I. INTRODUCTION

According to the “Technical Requirements for Toll Road Networking Toll”, in China, toll stations have designed two types of lanes, MTC (Manual Toll Collection) and ETC (Electronic Toll Collection). Based on two important parameters in the queuing theory method, traffic volume and service time, this requirement suggests setting up 1-2 ETC lanes and 2-25 MTC lanes [1].

In recent years, with the increase in transportation, different demands have emerged at different times. In addition, compared to the previous single cash payment method, MTC lane has adopted a variety of payment methods including credit card payment, mobile payment and so on [2].

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

Thus, the combination of different models and different payment models has caused changes in service time. Obviously, according to the previous traffic volume and service time, the management of the MTC lane work-schedule is no longer sufficient to meet the current operating needs of toll stations. Unreasonable opening of MTC lanes will cause two situations in toll station. One is that MTC lanes are often congested, and another is that MTC lanes are idle for a long time. This will cause environmental pollution or waste of operating costs.

Therefore, it is a very important issue to arrange the number of MTC lanes to be opened according to the current traffic demand and the actual service time in different time periods. However, there is currently a lack of specifications and standards for the number of MTC lanes opened in China. Therefore, how to scientifically obtain the work-schedule of MTC lanes is very important.

Toll station is a typical queuing service system [3], and the queuing theory model in operations research can well simulate the process of vehicles passing through toll station [4]. Guo [5] used the queuing system model to construct the optimization model of toll station. Komada and Nagatani [6] studied the relationship between traffic volume, number of lanes and queue length. Astarita *et al.* [7] used queue length to measure status of the toll station in a microscopic simulation model. Mahdi *et al.* [8] used the average queue length of the simulation output as an indicator to measure the traffic condition of the toll station. Al-Deek *et al.* [9] used the average queue length of the simulation output as an indicator to measure the traffic condition. Mahdi *et al.* [10] aimed at the effect of the percentage of heavy vehicles on the performance of toll station in terms of the queue length. From the literature, the queuing system consists of three parts: system input, system service and system output. For toll station, the system input is the traffic volume, and the service time reflects the system service. Also, we found the average queue length of lane is an important index and has a good advantage to measure the congestion status of toll station.

At present, many roads rely on predicting congestion to manage traffic. So, can we plan the number of MTC lanes by predicting the congestion of toll stations in advance? This may be a good attempt. Tian *et al.* [11] proposed a LSTM traffic flow prediction model. Hua *et al.* [12] proposed an improved RC-LSTM to achieve accurate prediction of time series. Wang *et al.* [13] proposed a Data-Driven Short-Term Forecasting model for urban road network traffic. Yang *et al.* [14] proposed a method for traffic flow prediction using LSTM. Niu *et al.* [15] proposed a real-time taxi-passenger prediction model based on L-CNN. Wang *et al.* [16] proposed a prediction method for predicting urban network traffic speed based on Bi-LSTM. Liu *et al.* [17] proposed a traffic flow prediction method for using toll data based on ARIMA model. Yang *et al.* [18] proposed a method for judging highway congestion based on toll data. We found that the LSTM model has a good effect on time series prediction, also, PSO algorithm is a commonly used optimization algorithm, which has the advantage of simple and easy to implement the model [19]. In order to improve the prediction accuracy, using PSO algorithm to optimize LSTM model is a feasible strategy. Therefore, we want to use PSO-LSTM to establish a prediction model of the average queue length of to predict whether the toll station is congested in advance.

Next, how can we determine the opening quantity of each periods according to the prediction results? We know that the traffic capacity of the expressway is divided into four levels according to the speed of the vehicle. Can the traffic capacity of the toll station be divided into levels? Then what is it based on? In the reference, the service level of the toll station is a quality standard for the satisfaction of the passengers on the traffic flow status [20]. It should be a good method to classify the service level of toll stations. So, we try to use the average queue length to determine the appropriate number of open lanes based on the predicted result under certain service level.

Then there is another aspect, so how to measure the number of toll station lanes? Jiyang and Zhou [21] proposed a toll station lane layout scheme based on cost analysis. Liu [22] proposed an optimization model for the number of toll lanes based on cost analysis. Wang [23] studied a toll lane optimization model for the number of toll lanes. Park *et al.* [24] proposed an optimal operation strategy for highway toll stations based on the benefit cost analysis. Yang *et al.* [25] used the toll station operating and vehicle queuing cost to propose the optimization model. Lin *et al.* [26] proposed a lane configuration model based on queuing theory with the operating cost and delay. As can be seen from the literatures, many studies have focused on using construction costs to set the total number of toll lanes that need to be charged, rather than operating costs [27]. However, the lane work-schedule directly determines the operating costs.

Hence, it is very meaningful to study the congestion prediction and lane work-schedule of toll station based on the toll data.

1) The average queue length of the lane is selected as a measure of the congestion of the toll station. And then, based on queuing theory, the average queue length and the service time are calculated for using to divide the level of service.

2) The prediction model based on LSTM and PSO algorithm is established. Meanwhile, the historical time series of lane average queue length, service time mean and traffic volume are selected as three inputs of the model to predict the average queue length.

3) Based on toll station service level, a lane work-schedule model is established. Then, the number of lanes opened in each time period can be determined by using the congestion prediction results and the lane work-schedule model.

4) The operating costs are used to verify the accuracy of the algorithm model through the toll data of Dongshe, Changfeng, and Linfen toll stations of Shanxi province.

II. PROCESSING ANALYSIS OF TOLL DATA AND TOLL STATION SERVICE LEVEL

A. TOLL DATA

The toll network collection system can collect real-time traffic information. The collected information is stored in the database, which includes static information and dynamic information. Static information mainly includes toll station name number, lane number, vehicle pit stop time, toll amount and toll method, etc. The dynamic information mainly includes information such as toll service time of each vehicle type. The main field information used in this paper are shown in the Table 1.

Due to equipment failure, transmission, environmental factors and other reasons, we usually get abnormal data. Through the statistical analysis of the historical data, we find that the abnormal data can be divided into three categories:

1) Completely duplicated data: Each field of two or more data are identical. We only keep one valid data for this kind of data.

TABLE 1. Main information of toll data.

Field Name	Directions
Date	YYYYMMDD
Time	HHMMSS
Toll Station	The Name of Toll Station
Mode of Payment	Cash
	Mobile
	Credit Card
Service Time	Seconds
	Small Vehicle
Vehicle Type	Middle-Sized Vehicle
	Large Vehicle

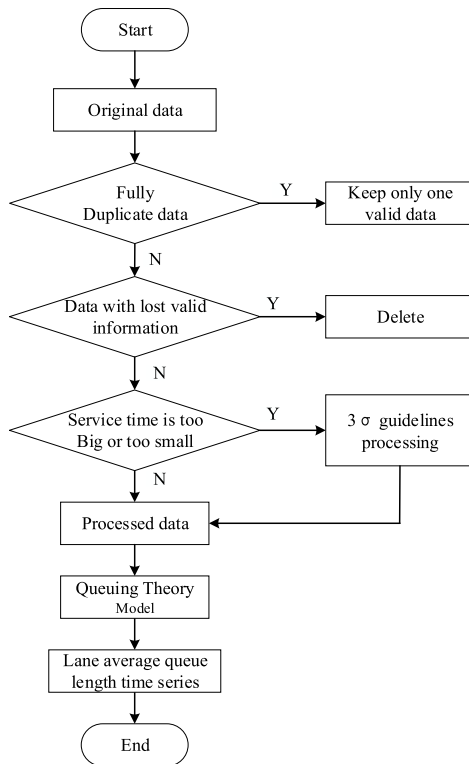


FIGURE 1. Data cleaning flow chart.

2) Missing data: Mainly including service time missing items, we delete such data.

3) There are too large or too small data in the field of service time. We use 3σ rule to deal with abnormal data.

The abnormal data identification and cleaning flow is shown in Fig. 1.

B. AVERAGE QUEUE LENGTH AND SERVICE TIME

The average queue length of the toll station is the average number of vehicles waiting to charge in each lane of the toll station [7]. The average queue length has the advantage of being easy to observe and operability. The more vehicles waiting in line, the lower the service level of toll stations. Therefore, the average queue length of toll lane is taken as an indicator to evaluate toll station service. The M/G/K model in queuing theory can well describe the state of vehicle passing at the toll station. The average queue length can be calculated

by the M/G/K model [28]:

$$W_q = \frac{D + E^2}{2E \times [\gamma - \lambda E]} \cdot \left[1 + \sum_{i=0}^{\gamma-1} \frac{(\gamma - 1)! (\gamma - \lambda E)}{i! (\lambda E)^{\gamma-i}} \right]^{-1} \quad (1)$$

$$L_q = \frac{\lambda D + \lambda E^2}{2E \times [\gamma - \lambda E]} \cdot \left[1 + \sum_{i=0}^{\gamma-1} \frac{(\gamma - 1)! (\gamma - \lambda E)}{i! (\lambda E)^{\gamma-i}} \right]^{-1} \quad (2)$$

$$W = E + W_q \quad (3)$$

W is the average residence time; Wq is the average queue time; Lq is average lane queue length; λ is average incoming strength (pcu/s), it can be calculated from traffic volume; γ is the number of toll lanes.

D is the variance of service time, E is the mean of service time. These two parameters can be calculated as follows:

$$E = \sum_{i=1}^n l_i \cdot E_{i1} \cdot \sum_{j=1}^m \beta_{ij} \cdot \lambda_{ij} \quad (4)$$

$$D = \sum_{i=1}^n l_i \cdot D_{i1} \cdot \sum_{j=1}^m \beta_{ij} \cdot \lambda_{ij} \quad (5)$$

$$\sum_{i=1}^n l_i = 1 \quad (6)$$

E_{i1} is the mean of service time of small vehicle using the type i payment; D_{i1} is the variance of service time of small vehicle using the type i payment; n is the number of payment methods; m is the number of vehicle types; l_i is the proportion of each payment; λ_{ij} is the proportion of vehicle type j with payment i; β_{ij} is the service time conversion coefficient of vehicle type j with payment i;

$$\beta_{ij} = t_{ij} / t_{i1} \quad (7)$$

t_{i1} is the service time of small vehicle with payment i; t_{ij} is the service time of vehicle type j with payment i.

Based on the processed toll data, we firstly calculate the mean and variance of the service time according to the (4) ~ (7), then calculate the traffic volume of the toll station in each period. Finally, we calculate the average queue length according to the (1) ~ (3).

C. TOLL STATION SERVICE LEVEL

Based on the toll data of 43 toll stations in Shanxi province from January to March 2019, we calculate the toll station capacity under different lane number and queue length by using the above methods. Taking Dongshe Toll Station as an example, the results are shown as Fig. 2.

When the average number of queued vehicles increases from 1 to 4, the traffic capacity of the toll stations increases significantly.

When the average number of queued vehicles is from 4 to 8, the toll traffic capacity of the toll stations will increase slowly.

When the average number of queued vehicles is from 8 to 10, the increase in the traffic capacity of the toll stations is slowed down more significantly.

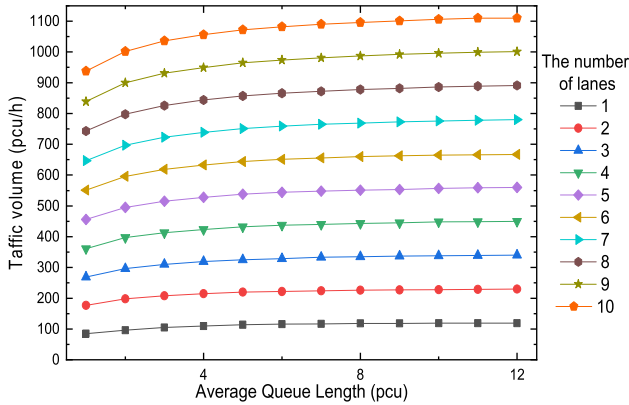


FIGURE 2. Traffic capacity of toll lane.

The number of queued vehicles is more than 10, and the traffic capacity of the system queuing system is saturated.

We found that traffic congestion occurred under the condition of the Tertiary service level and Fourth service level of the toll station. Under the condition of the Secondary service level, although there are vehicles queuing in the lane, it can guarantee the fast passage of vehicles, the queue length is not so long, and the drivers and passengers have a good service experience. In order to ensure the traffic of toll stations, the work-schedule of the toll lanes should meet at least the secondary service level [5].

III. CONGESTION PREDICTION AND LANE WORK-SCHEDULE MODEL

A. CONGESTION PREDICTION BASED ON LSTM

Based on LSTM method, we establish a toll station congestion prediction model, which algorithm flow is shown as Fig. 3.

It can be seen that a LSTM memory block contains a single cell. The three gates are nonlinear and aggregate units, which gather all the internal and external excitation of the block, and control the excitation of the unit through the proliferous node.

LSTM contains a set of interconnected recurrent networks, which is also known as memory block. Each network block contains one or more auto correlative memory cells and three accrued units: input gate, output gate and forget gate.

The excitation function of the forget gate is usually a logarithmic logic curve, so gate excitation is evaluated between 0 and 1. The main model formulas of LSTM are as follows:

$$a_i^t = \sum_{i=1}^I w_{il}x_i^t + \sum_{h=1}^H w_{hl}b_h^{t-1} + \sum_{c=1}^C w_{cl}s_c^{t-1} \quad (8)$$

$$a_\phi^t = \sum_{i=1}^I w_{i\phi}x_i^t + \sum_{h=1}^H w_{h\phi}b_h^{t-1} + \sum_{c=1}^C w_{c\phi}s_c^{t-1} \quad (9)$$

$$a_c^t = \sum_{i=1}^I w_{ic}x_i^t + \sum_{h=1}^H w_{hc}b_h^{t-1} \quad (10)$$

$$s_c^t = b_\phi^t \odot s_c^{t-1} + b_i^t \odot g(a_c^t) \quad (11)$$

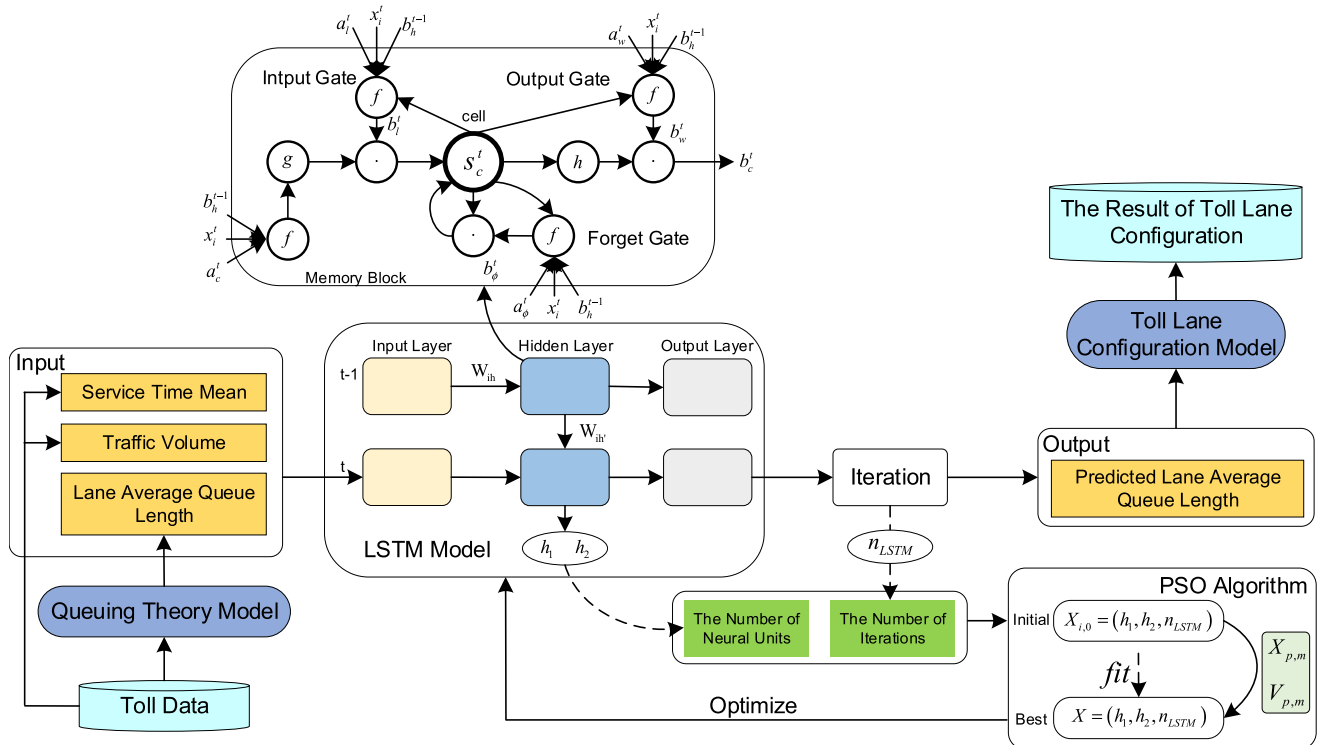


FIGURE 3. Congestion prediction algorithm flow.

$$a_w^t = \sum_{i=1}^I w_{iw}x_i^t + \sum_{h=1}^H w_{hw}b_h^{t-1} + \sum_{c=1}^C w_{cw}s_c^t \quad (12)$$

$$b_c^t = b_w^t \odot h(s_c^t) \quad (13)$$

where, x_i^t is the input of average queue length at time t , b_h^{t-1} is the output of hidden layer at time $t - 1$, s_c^{t-1} is the state value of the memory Cell. b_j^t is the output expression of the input gate at time t ; a_ϕ^t is the input of forget gate at time t ; b_ϕ^t is the output expression of the forget gate at time t ; a_l^t is the input of input gate at time t ; a_c^t is the input of memory cell at time t ; s_c^t is the state value of the memory Cell at time t ; a_w^t is the input of output gate at time t ; b_w^t is the output expression of the output gate at time t ; b_c^t is the output expression of memory block at time t ; i is a layer of input neurons, j is the next layer of neurons, W_{ij} is the weight of units i to j , a_j^t is the input of neuron j at time t , b_j^t is the output of control gate, f is activation function, g and h represent the input and output activation functions of Cell, I is the number of input layer neural units, K is the number of neural units in the output layer, H is the number of memory modules, C is the memory cell, S is the status of the memory unit.

The composition of the time series of the average queue length of the lane is related to the traffic volume of the toll station and the service time per period. In order to ensure the accuracy of the LSTM model prediction, we use the time series of lane average queue length, service time mean and traffic volume as the input of the model to predict the average queue length. The input vector expression is as follows:

$$x(t) = [L_{qu}, E_u, Q_u] \quad (14)$$

where, L_{qu} is the average queue length at time u , E_u is the service time average value at time u , and Q_u is the traffic volume at time u .

B. PSO OPTIMIZING ALGORITHM

According to the preliminary research, we found that the number of neurons and model iterations in the hidden layer of the LSTM are difficult to determine. The number of neural units in the hidden layer directly determines the fitting ability of the model, and the number of iterations determines the training effect. Usually the above parameters are manually adjusted that would cause greater randomness. So, we use the

PSO algorithm to optimize the LSTM parameters for improving the prediction accuracy.

$$V_{p,m+1} = wl \cdot V_{p,m} + c_1 \cdot rand \cdot (pbesm_p - X_{p,m}) + c_2 \cdot rand \cdot (gbesm_m - X_{p,m}) \quad (15)$$

$$X_{p,m+1} = X_{p,m} + V_{p,m+1} \quad (16)$$

where: wl is the weight; c_1 and c_2 are the learning factors; $rand$ is the random number between $[0, 1]$. $pbesm_p$ is the particle's individual optimal value, and $gbesm_m$ is the global optimal value. In this paper, the solution of the optimization problem is treated as a particle.

The particles are constantly flying in space, searching for the best position, this position is the optimal solution.

TABLE 2. Level of service.

Service Level	Description
Primary	Smooth
Secondary	Basically Smooth
Tertiary	Congested
Fourth	Severely Congested

The core idea of the PSO algorithm is to first initialize a set of random solutions and then find the optimal solution by iteratively. In the m th iteration, the position and velocity of the particle p are $X_{p,m}$ and $V_{p,m}$, respectively. Particles find the optimal value by updating their speed and position. Particles update their speed and new position according to (15) and (16). Due to the limited global optimization ability and convergence speed of the basic PSO algorithm. In this paper, we use the nonlinear variable weight method to improve the PSO algorithm.

$$wl = wl_{\max} - (wl_{\max} - wl_{\min}) (m/m_{\max})^2 \quad (17)$$

wl_{\max} and wl_{\min} are the maximum and minimum values of wl , respectively; m_{\max} is the maximum number of iterations.

Through the above method, the problem of insufficient optimization ability due to inertia weight is avoided. The initial particle expression is as follows:

$$X_{i,0} = (h_1, h_2, n_{LSTM}) \quad (18)$$

h_1 is the number of hidden layer neurons in the first layer LSTM, h_2 is the number of hidden layer neurons in the second layer LSTM, and n_{LSTM} is the number of iterations of the LSTM. Assign each particle to the LSTM and enter the data into the LSTM for training. The fitness function of the algorithm is:

$$fit = 0.5 \cdot \frac{1}{Z} \sum_z \frac{|yy^z - y^z|}{y^z} + 0.5 \cdot \frac{1}{L} \sum_l \frac{|yy^l - y^l|}{y^l} \quad (19)$$

yy^z is the output of the LSTM test sample obtained after reaching the iteration limit; y^z is the expected output of the LSTM test sample, Z is the number of test samples, yy^l is the output of the training sample, y^l is the expected output of the training sample, and L is the training sample number. In order to prevent the model from overfitting, the fitness function includes the error of the training sample and the test sample, and gives the same weight 0.5. The optimal LSTM network model is finally determined by iterative iteration.

C. TOLL LANE WORK-SCHEDULE MODEL

In order to ensure the normal traffic of the toll station, it is necessary to guarantee the service level above the secondary level. Considering the working time of the staff and the ease of operation, we take one hour as the work-schedule time period. The toll lane work-schedule model is set as the (20).

$$\gamma_{u-mtc} = \left\lceil \frac{\gamma_{mtc} \cdot L_{q-predict-u}}{4} \right\rceil (\gamma_{u-open} \leq \gamma_{mtc}) \quad (20)$$

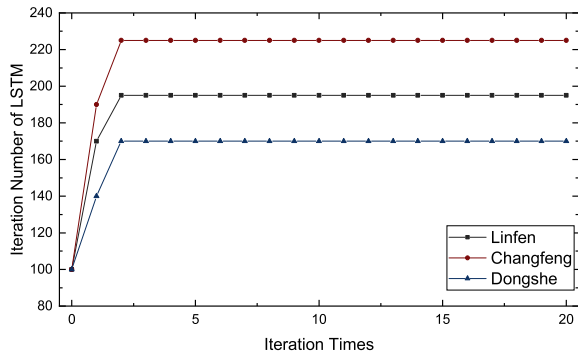


FIGURE 4. Iteration number of LSTM.

γ_{u-mtc} is the number of MTC open lanes at time u period, $L_{q-predict-u}$ is the prediction of average Lane queue length at time u , γ_{mtc} is the number of MTC toll lanes.

IV. CASE ANALYSIS

In this paper, we select the toll data of the three toll stations of Dongshe, Linfen and Changfeng in Shanxi Province from February to March 2019. Dongshe toll station is set up with 7 MTC lanes and 2 ETC lanes, Linfen toll station has 7 MTC lanes and 1 ETC lane, Changfeng station has 8 MTC lanes and 2 ETC lanes.

A. CONGESTION PREDICTION ANALYSIS

According to the parameter value range in the reference [29], the values of average queue length, service time and traffic volume of the first five hours are used as the input to predict the average queue length for the next hour. And then, a rolling horizon method is used to predict. The 80% of the data are used for training and 20% of the data are used for prediction. The input vector is three, so the number of neurons in the input layer of the LSTM model is three. We use Adam algorithm to optimize the target function in the back propagation of the LSTM model, and use the PSO algorithm to find the optimal LSTM structural parameters.

The number of particles in the PSO algorithm is set to 5, the number of iterations is 20, the learning factors c_1 and c_2 are 2, and the internal parameters of particle $X_{i,0} = (h_1, h_2, n_{LSTM})$ are in the range of [1, 50], [1, 50], and [50-350]. $w_{l_{max}}$ and $w_{l_{min}}$ are 0.9 and 0.4, respectively.

With the iteration of the PSO algorithm, the number of LSTM hidden layer neurons and the number of iterations are shown as Fig. 4 and Fig. 5. Taking the Dongshe toll station as an example, the number of iterations is 170, the hidden layer neuron h1 is 11, and h2 is 9. We found that through the iteration of the PSO algorithm, the hidden layers h1 and h2 of the three toll stations eventually stabilize, and the optimal number of iterations of the LSTM model is also obtained.

In order to evaluate the prediction accuracy, this paper mainly uses two types of indicators, Mean Absolute Percentage Error (MAPE) and Root Mean Squared Error (RMSE), to quantify the prediction error. The errors are calculated

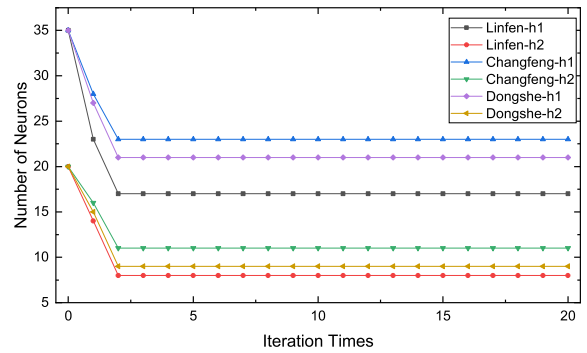


FIGURE 5. The number of LSTM hidden layer neurons.

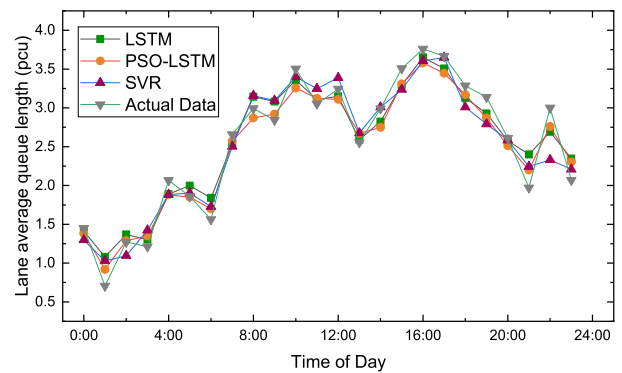


FIGURE 6. Prediction results on March 28 (weekday), Dongshe toll station.

as follows:

$$MAPE = \frac{1}{T_{sum}} \sum_{T_{sum}} \frac{|L_{q-predict}(u) - L_q(u)|}{L_q(u)} \quad (21)$$

$$RMSE = \sqrt{\frac{1}{T_{sum}} \sum_{T_{sum}} (L_{q-predict}(u) - L_q(u))^2} \quad (22)$$

where, $L_{q-predict}(u)$ is the prediction Lane average queue length at time u , $L_q(u)$ is actual prediction Lane average queue length at time u , T_{sum} is the total number of prediction cycles.

We selected the weekday and weekend scenarios to verify the average queue length of the three toll stations, and used Support Vector Regression model (SVR) [30] to compare the algorithm precision.

1) THE WEEKDAY

The prediction results of Dongshe Toll Station, Changfeng Toll Station and Linfen Toll Station on March 28 (weekday) are shown as Fig. 6, Fig. 7, and Fig. 8. The percent errors (PE) of Dongshe toll station, Changfeng toll station and Linfen toll station are shown as Fig. 9. The results of the error evaluation indicator are shown in Table 3.

The mean absolute percentage error of LSTM, PSO-LSTM, and SVR in Dongshe toll station are 8.3%, 6.1%, and 9.0%, respectively. The root mean squared error of LSTM, PSO-LSTM, and SVR in Dongshe toll station are 0.20, 0.16, and 0.23, respectively.

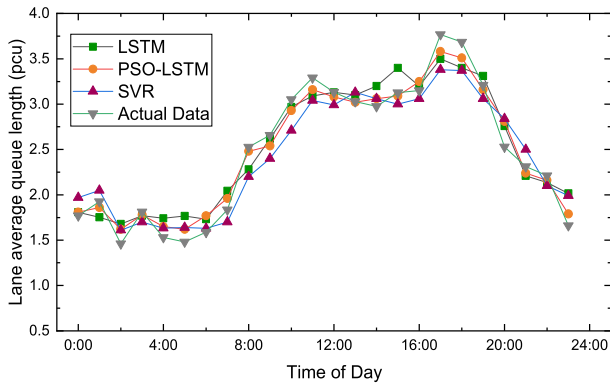


FIGURE 7. Prediction results on March 28 (weekday), Changfeng toll station.

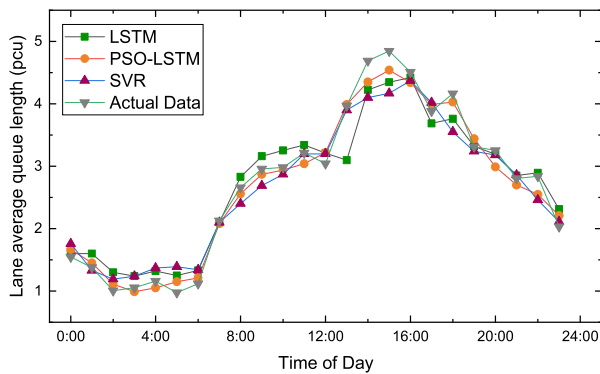


FIGURE 8. Prediction results on March 28 (weekday), Linfen toll station.

The mean absolute percentage error of LSTM, PSO-LSTM, and SVR in Changfeng toll station are 7.6%, 5.0%, and 8.1%, respectively. The root mean squared error of LSTM, PSO-LSTM, and SVR in Changfeng toll station are 0.19, 0.12, and 0.21, respectively.

The mean absolute percentage error of LSTM, PSO-LSTM, and SVR in Linfen toll station are 9.7%, 6.0%, and 9.6%, respectively. The root mean squared error of LSTM, PSO-LSTM, and SVR in Linfen toll station are 0.26, 0.16, and 0.25, respectively.

2) THE WEEKEND

The prediction results of Dongshe Toll Station, Changfeng Toll Station and Linfen Toll Station on March 30 (weekend) are shown in Fig. 10, Fig. 11, and Fig. 12. The percent errors (PE) of Dongshe toll station, Changfeng toll station and Linfen toll station on March 30 (weekend) are shown as Fig 13. The results of the error evaluation indicator MAPE, RMSE are shown in Table 4.

The mean absolute percentage error of LSTM, PSO-LSTM, and SVR in Dongshe toll station are 7.1%, 5.1%, and 9.8%, respectively. The root mean squared error of LSTM, PSO-LSTM, and SVR in Dongshe toll station are 0.23, 0.14, and 0.33, respectively.

TABLE 3. Prediction error.

Station	MAPE			RMSE		
	LSTM	PSO-LSTM	SVR	LSTM	PSO-LSTM	SVR
Dongshe toll station	8.3%	6.1%	9.0%	0.20	0.16	0.23
Changfeng toll station	7.6%	5.0%	8.1%	0.19	0.12	0.21
Linfen toll station	9.7%	6.0%	9.6%	0.26	0.16	0.25

TABLE 4. Prediction error.

Station	MAPE			RMSE		
	LSTM	PSO-LSTM	SVR	LSTM	PSO-LSTM	SVR
Dongshe toll station	7.1%	5.1%	9.8%	0.23	0.14	0.33
Changfeng toll station	7.8%	5.3%	7.7%	0.21	0.13	0.19
Linfen toll station	8.6%	6.4%	8.4%	0.21	0.15	0.20

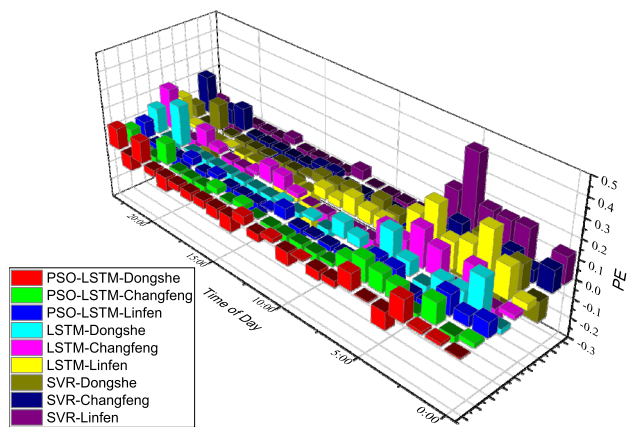


FIGURE 9. Errors on March 28 (weekday).

The mean absolute percentage error of LSTM, PSO-LSTM, and SVR in Changfeng toll station are 7.8%, 5.3%, and 7.7%, respectively. The root mean squared error of LSTM, PSO-LSTM, and SVR in Changfeng toll station are 0.21, 0.13, and 0.19, respectively.

The mean absolute percentage error of LSTM, PSO-LSTM, and SVR in Linfen toll station are 8.6%, 6.4%, and 8.4%, respectively. The root mean squared error of LSTM, PSO-LSTM, and SVR in Linfen toll station are 0.21, 0.15, and 0.20, respectively.

B. TOLL LANE WORK-SCHEDULE ANALYSIS

According to the prediction model, the predicted average queue length of lanes in each time period is obtained, and then the number of lanes opening in each time period is calculated by using the toll work-schedule model. We divided the weekday and weekend into two scenarios. The results are shown in Fig. 14 and Fig. 15.

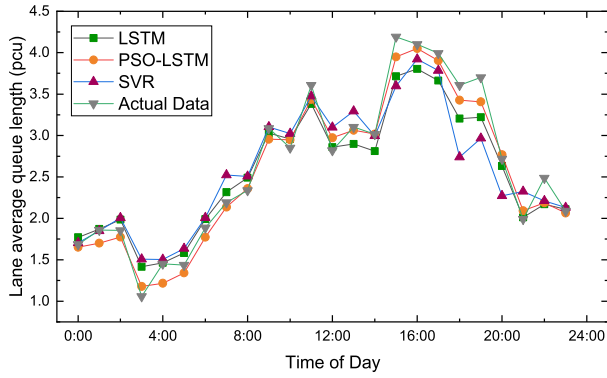


FIGURE 10. Prediction results on March 30 (weekend), Dongshe toll station.

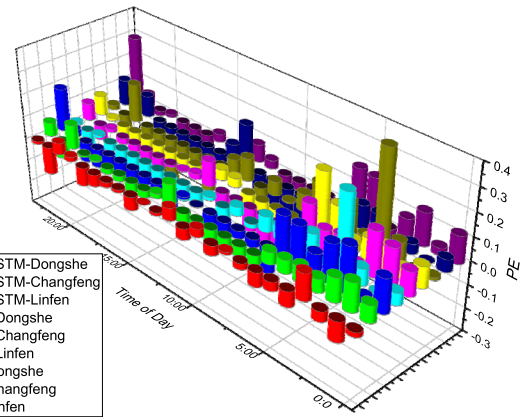


FIGURE 13. Errors on March 30 (weekend).

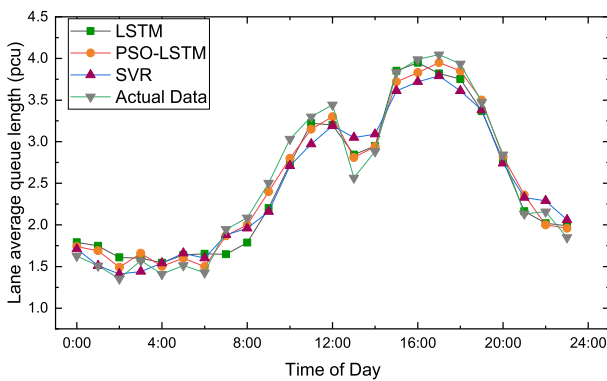


FIGURE 11. Prediction results on March 30 (weekend), Changfeng toll station.

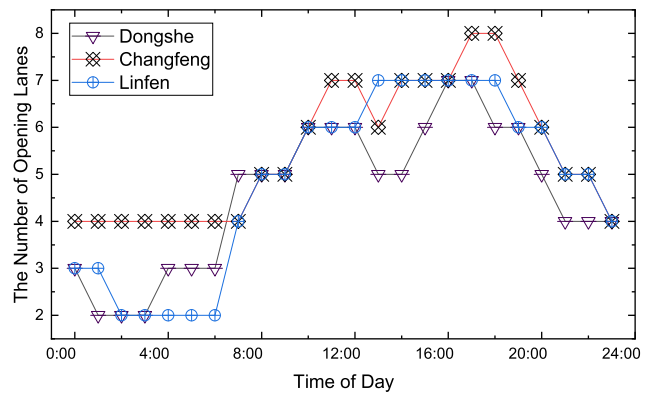


FIGURE 14. Toll lane work-schedule on March 28 (weekday).

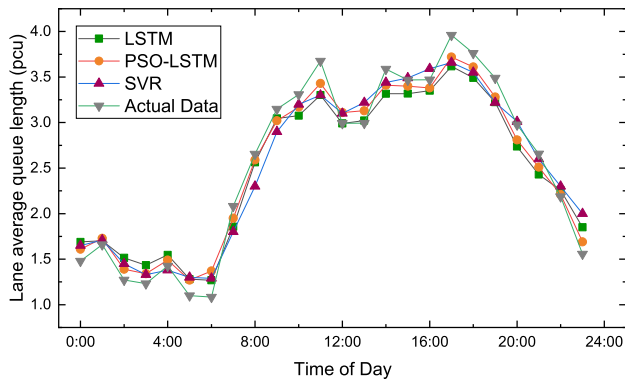


FIGURE 12. Prediction results on March 30 (weekend), Linfen toll station.

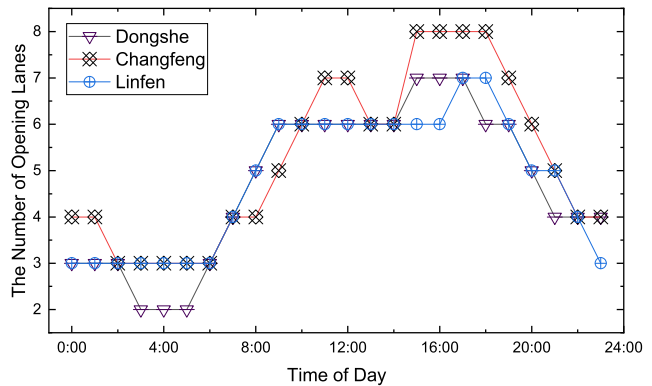


FIGURE 15. Toll lane work-schedule on March 30 (weekend).

Taking Shanxi Dongshe Toll Station on March 28 as an example. We compare the results of the proposed method with the results of the reference [27] as shown as Fig. 16.

The cooperation details are analyzed as follows:

1) According to the reference method, 2 lanes should be opened at 0:00, but according to the prediction queue length and lane work-schedule model, 3 lanes should be opened for meeting the secondary service level.

2) From 1:00 to 3:00, the calculation results of the two methods are the same.

3) The average queue length of lanes increased from about 1.5 pcu to 2 pcu at 4:00, which also indicated that the average queue length of lanes increased. Although the length of the queue is small, in order to meet the secondary service level of the toll station, according to lane work-schedule model, we found that additional lane needs to be opened at 4:00 and 5:00.

4) At 6:00, four lanes were opened according to the reference method, but according to the proposed method, only

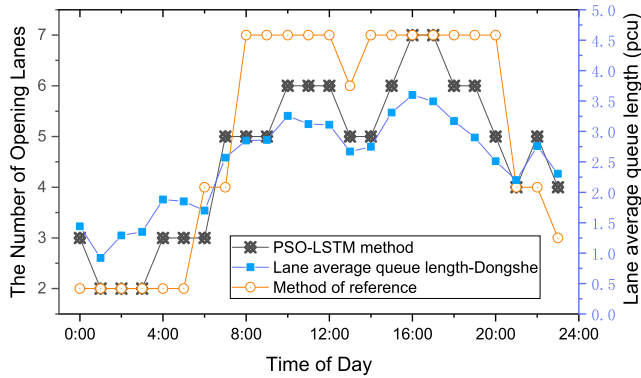


FIGURE 16. Lane work-schedule result comparison.

three lanes need to be opened to meet the secondary service level.

5) At 7:00, four lanes were also opened according to the reference method, but according to the proposed method, five lanes need to be opened to meet the secondary service level.

6) According to the method in reference method, in the time period from 8:00 to 20:00, all lanes were opened except for six lanes at 13:00. However, according to proposed method, all lanes need to be opened only from 16:00 to 17:00, and 1-2 lanes can be closed in other periods to meet the secondary service level of toll station.

7) At 21:00, the calculation results of the two methods are the same.

8) At 22:00 and 23:00, compared to reference method, we found that we need to increase the number of lanes opened in order to meet the secondary service level.

C. OPERATING COST ANALYSIS

The next, we use the toll station operating cost to evaluate the toll lane work-schedule. The toll station operating cost $\varepsilon_{station}$ in an hour includes: labor hour cost ε_{pe} , single lane operation hour electricity cost ε_{el} and single lane operation hour maintenance cost ε_{re} .

$$\varepsilon_{station} = (\varepsilon_{pe} + \varepsilon_{el} + \varepsilon_{re}) \cdot \gamma_u \quad (23)$$

$$\varepsilon_{pe} = \frac{2 \cdot \varepsilon_{pe-m}}{30 \cdot 8} \quad (24)$$

$$\varepsilon_{el} = \frac{\varepsilon_{el-m}}{30 \cdot 24} \quad (25)$$

$$\varepsilon_{re} = \frac{\varepsilon_{re-m}}{30 \cdot 24} \quad (26)$$

where, ε_{pe-m} is the monthly salary of staff, ε_{el-m} is the monthly electricity cost for a single toll lane, ε_{re-m} is the monthly maintenance cost for a single toll lane.

We conducted a survey of the toll stations, the monthly salary of toll station staff is 4,500 RMB, two toll collectors are usually required for a single lane, and the monthly electricity cost is 1,200 RMB and the monthly maintenance cost is 1,000 RMB. According to the lane work-schedule of reference [27], the operating cost of the toll station is 4,907RMB

per day. According to the method proposed in this paper, the toll station operating cost is 4,502RMB.

So, we could save 405RMB operating costs in one day. Using the cost of the above calculation to estimate the cost of one year, we found that the operating cost of 147,825RMB could be saved in one year. The analysis of the other two toll stations show that this method is also effective.

Hence, the lane work-schedule proposed in this paper could not only meet the secondary service level of the toll station, but also reduce the operating cost of the toll station

V. CONCLUSION

In order to determine the number of MTC lanes open in each time period of toll station, in this paper, firstly, the average queue length is used to measure the congestion of toll station. And then by queuing theory, we divide the service level of toll station into four levels according to the relationship between the average queue length and traffic capacity. Secondly, based on the LSTM and PSO algorithm, a prediction model is established to predict the average queue length of toll station. Thirdly, on the basis of meeting the secondary service level of toll stations, that is the average queue length is 4 pcu, a lane work-scheduling model is presented. According to this model, the number of MTC lanes opened in each period could be determined. Fourthly, taking Dongshe, Changfeng and Linfen toll stations in Shanxi Province as examples, the predictive performance of the algorithm and the accuracy of the toll work-schedule model are verified. The prediction results are compared with the SVR model. The MAPE of the PSO-LSTM model improves the prediction accuracy by about 2% and 3% compared to the LSTM and SVR models. Finally, we use the toll station operating cost to evaluate the toll lane work-schedule. Taking Dongshe Toll Station as an example, the proposed method of this paper is compared with the method of reference [27], the operating cost of 147,825RMB could be saved in one year. The results show that the congestion prediction model and lane work-schedule model proposed in this paper can determine the number of lanes opened in each period of the toll station effectively.

REFERENCES

- [1] *Technical Requirements for Toll Road Networking Toll, Announcement No. 35 of 2007*, Ministry Transp. People's Republic China, Beijing, China, Oct. 2007.
- [2] T. Yang, P. Wang, J. J. Zhu, and J. D. Zhao, "Optimization of lane number of highway toll station based on multiple payments," *J. Transp. Syst. Eng. Inf. Technol.*, vol. 19, no. 5, pp. 212–218, Oct. 2019.
- [3] Z. W. Dong and Y. M. Ni, "Analysis and simulation of the features of queues in highway toll station," *Sci. Mosaic*, vol. 10, pp. 47–49, Dec. 2016.
- [4] L. Y. Wang and Y. Feng, "A study to setting function of queuing discipline in highway tollgate," *Adv. Mater. Res.*, vols. 734–737, pp. 1594–1597, Aug. 2013.
- [5] Y. M. Guo, "Optimization of expressway toll-gate resource allocation based on lane queuing system model," *Value Eng.*, vol. 11, pp. 198–201, Apr. 2015.
- [6] K. Komada and T. Nagatani, "Traffic flow through multi-lane toll-booths on a toll highway," *Phys. A, Stat. Mech. Appl.*, vol. 389, no. 11, pp. 2268–2279, Jun. 2010.
- [7] V. Astarita, M. Florian, and G. Musolino, "A microscopic traffic simulation model for the evaluation of toll station systems," in *Proc. ITSC IEEE Intell. Transp. Syst.*, Aug. 2001, pp. 692–697.

- [8] M. B. Mahdi and L. V. Leong, "Assessment of queue length and delay at toll plaza using microscopic traffic simulation," *Appl. Mech. Mater.*, vol. 802, pp. 387–392, Oct. 2015.
- [9] H. M. Al-Deek, A. A. Mohamed, and E. A. Radwan, "New model for evaluation of traffic operations at electronic toll collection plazas," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 1710, no. 1, pp. 1–10, Jan. 2000.
- [10] M. B. Mahdi, L. V. Leong, and A. F. M. Sadullah, "Use of microscopic traffic simulation software to determine heavy-vehicle influence on queue lengths at toll plazas," *Arabian J. Sci. Eng.*, vol. 44, no. 8, pp. 7297–7311, Aug. 2019.
- [11] Y. Tian, K. Zhang, J. Li, X. Lin, and B. Yang, "LSTM-based traffic flow prediction with missing data," *Neurocomputing*, vol. 318, pp. 297–305, Nov. 2018.
- [12] Y. Hua, Z. Zhao, R. Li, X. Chen, Z. Liu, and H. Zhang, "Deep learning with long short-term memory for time series prediction," *IEEE Commun. Mag.*, vol. 57, no. 6, pp. 114–119, Jun. 2019.
- [13] W. Xiangxue, X. Lunhui, and C. Kaixun, "Data-driven short-term forecasting for urban road network traffic based on data processing and LSTM-RNN," *Arabian J. Sci. Eng.*, vol. 44, no. 4, pp. 3043–3060, Apr. 2019.
- [14] B. Yang, S. Sun, J. Li, X. Lin, and Y. Tian, "Traffic flow prediction using LSTM with feature enhancement," *Neurocomputing*, vol. 332, pp. 320–327, Mar. 2019.
- [15] K. Niu, C. Cheng, J. Chang, H. Zhang, and T. Zhou, "Real-time taxi-passenger prediction with L-CNN," *IEEE Trans. Veh. Technol.*, vol. 68, no. 5, pp. 4122–4129, May 2019.
- [16] J. Wang, R. Chen, and Z. He, "Traffic speed prediction for urban transportation network: A path based deep learning approach," *Transp. Res. C, Emerg. Technol.*, vol. 100, pp. 372–385, Mar. 2019.
- [17] Y. L. Liu, "Short-term traffic flow forecasting method based on big charging data," *Comput. Digit. Eng.*, vol. 47, no. 5, pp. 1164–1169 and 1188, May 2019.
- [18] J. Yang, G. Jiang, and Q. Li, "The automatic traffic congestion identification of freeway based on charging date," *J. Harbin Inst. Technol.*, vol. 46, no. 12, pp. 108–113, Dec. 2014.
- [19] B. Yang and W. Y. Qian, "Summary on improved inertia weight strategies for particle swarm optimization algorithm," *J. Bohai Univ.*, vol. 40, no. 3, pp. 274–288, Sep. 2019.
- [20] W. T. Hugh and L. A. Hoel, "Toll plaza capacity and level of service," *Transp. Res. Board*, pp. 119–127, Apr. 1991.
- [21] B. B. Jiyang and J. F. Zhou, "Lane allocation of highway toll gate based on cost analysis," *J. Chongqing Jiaotong Univ. (Natural Sci.)*, vol. 28, no. 8, pp. 85–91, Sep. 2018.
- [22] Y. Liu, *The Research on Regional Traffic Characteristics and Lane Setting of Toll Station*, Nanjing, China: Southeast Univ., 2018.
- [23] H. Y. Wang, *Study on the Capacity and Lane Configuration Strategy of Expressway Toll Station*, Chengdu, China: Southwest Jiaotong Univ., 2017.
- [24] K.-C. Park, S.-M. Rhee, and H.-S. Shin, "Optimal integrated operation strategy for highway tollgate using micro simulation," in *Proc. Int. Conf. Multimedia Ubiquitous Eng. (MUE)*, Apr. 2007, pp. 429–445.
- [25] Z. W. Yang, "Research on charging station ETC lane opening amount based on cost analysis," *Shanxi Archit.*, vol. 39, no. 24, pp. 172–173, Aug. 2013.
- [26] P. Q. Lin and Y. Q. Liang, "Modeling and solving for lane type setting problem in highway toll station," *J. Transp. Sys. Eng. Inf. Technol.*, vol. 20, no. 1, pp. 152–159, Feb. 2020.
- [27] *Highway Toll Station Personnel Arrangement Plan*. Baidu Wenku, Beijing, China. Accessed: Aug. 23, 2016. [Online]. Available: <http://wenku.baidu.com/view/48cfe84af1ffc4ff47ace3.html?from=search>
- [28] M. G. Qu, "Study on the number of toll lanes at the entrance and exit of expressway," *Highway*, no. 5, pp. 262–267, May 2012.
- [29] Q. Wang, T. Li, and C. Zhu, "Short-time prediction for traffic flow based on wavelet de-noising and LSTM model," *J. Meas. Sci. Instrum.*, vol. 11, pp. 1–14, Mar. 2020.
- [30] J. D. Zhao, Y. Gao, Z. M. Bai, H. Wang, and S. H. Lu, "Traffic speed prediction under non-recurrent congestion: Based on LSTM method and BeiDou navigation satellite system data," *IEEE Intell. Transp. Syst. Mag.*, vol. 11, no. 2, pp. 70–81, Mar. 2019.



PENG WANG was born in 1994. He received the B.S. degree from Inner Mongolia Agricultural University, in 2017. He is currently pursuing the master's degree with the School of Mechanical, Electronic and Control Engineering, Beijing Jiaotong University, China.

He is also a Professional in vehicle operation engineering. His research interests include traffic big data, traffic control, and intelligent transportation technology.



JIANDONG ZHAO received the Ph.D. degree in mechatronic engineering from Tsinghua University, China, in 2004. He is currently a Professor with the School of Traffic and Transportation, Beijing Jiaotong University, China. He is also a member of the Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport, Ministry of Transport, China. His research interests include traffic flow forecast, 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. He is also a Professional in vehicle operation engineering. His research interests include traffic control and traffic flow prediction.



MIGUEL ANGEL SOTELO received the Ph.D. degree in electrical engineering from the University of Alcalá (UAH), Madrid, Spain, in 2001. He is currently a Full Professor with the Department of Computer Engineering, University of Alcalá (UAH). His research interests include autonomous vehicles and prediction of intentions. He is member of the IEEE ITSS Board of Governors and an Executive Committee. He has served as a Project Evaluator, a Rapporteur, and a Reviewer with the European Commission in the field of ICT for Intelligent Vehicles and Cooperative Systems in FP6 and FP7. He is also the President of the IEEE Intelligent Transportation Systems Society. He served as the Editor-in-Chief for the *IEEE Intelligent Transportation Systems Magazine* and ITSS Newsletter. He served as an Associate Editor for the *IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS*.



ZHIXIONG LI received the Ph.D. degree in transportation engineering from the Wuhan University of Technology, China. He is currently an ARC DECRA Fellow with the School of Mechanical, Materials, Mechatronic and Biomedical Engineering, University of Wollongong, Australia. He is also an Adjunct Professor with the Faculty of Engineering, Ocean University of China. His research interests include intelligent vehicles and control, loop closure detection, and mechanical system modeling and control. He is also an Associate Editor of *Measurement* (Elsevier) and *Measurement: Sensors* (Elsevier). He is also a Column Editor of the *IEEE Intelligent Transportation Systems Magazine*.

...