

Received April 10, 2020, accepted April 19, 2020, date of publication April 27, 2020, date of current version May 15, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2990738

Traffic Flow Forecast Through Time Series Analysis Based on Deep Learning

JIANHU ZHENG^{1,2} AND MINGFANG HUANG¹

¹School of Economics and Management, Minjiang University, Fuzhou 350108, China

²Internet Innovation Research Center, Fujian Provincial Research Base for Humanities and Social Sciences of Colleges and Universities, Minjiang University, Fuzhou 350108, China

Corresponding author: Mingfang Huang (1884@mju.edu.cn)

This work was supported in part by the Natural Science Foundation of Fujian Province, China, under Grant 2016J01757, in part by the Science and Technology Project of Minjiang University under Grant MYK18039, and in part by the Internet Innovation Research Center, Fujian Provincial Research Base for Humanities and Social Sciences of Colleges and Universities, Minjiang University, under Grant IIRC20180103.

ABSTRACT Traffic congestion is a thorny issue to many large and medium-sized cities, posing a serious threat to sustainable urban development. Recently, intelligent traffic system (ITS) has emerged as an effective tool to mitigate urban congestion. The key to the ITS lies in the accurate forecast of traffic flow. However, the existing forecast methods of traffic flow cannot adapt to the stochasticity and sheer length of traffic flow time series. To solve the problem, this paper relies on deep learning (DL) to forecast traffic flow through time series analysis. The authors developed a traffic flow forecast model based on the long short-term memory (LSTM) network. The proposed model was compared with two classic forecast models, namely, the autoregressive integrated moving average (ARIMA) model and the backpropagation neural network (BPNN) model, through long-term traffic flow forecast experiments, using an actual traffic flow time series from OpenITS. The experimental results show that the proposed LSTM network outperformed the classic models in prediction accuracy. Our research discloses the dynamic evolution law of traffic flow, and facilitates the decision-making of traffic management.

INDEX TERMS Traffic flow forecast, time series analysis, deep learning (DL), long short-term memory (LSTM).

I. INTRODUCTION

Owing to economic growth and urbanization, many large and mid-sized cities are increasingly troubled by traffic congestion, which brings a series of social problems (e.g. long travel time, frequent traffic accidents, and severe environmental pollution) [1]. Many measures have been developed to mitigate urban congestion, namely, improving transport infrastructure, charging congestion fee, providing route guidance, promoting public transit, and implementing traffic control [2], [3]. However, it is difficult to ease traffic congestion by adding road facilities, if the transport infrastructure in a city is already complete. In this case, the traffic congestion could be effectively reduced by improving the traffic management in a digitized and intelligent manner, i.e. building an intelligent traffic system (ITS) [4]. Being a key aspect of the ITS,

the short-term traffic flow forecast is the basis and premise of traffic management measures (e.g. traffic planning, route guidance, and traffic control). The effectiveness of the ITS relies on the accurate and reliable forecast of short-term traffic flow.

The short-term traffic flow forecast aims to predict the traffic state of a road section or intersection in the near future based on the historical traffic data and travel experience. It has long been a research hotspot in the field of the ITS. The early methods for traffic flow forecast are model-driven, which work only if the data and model parameters satisfy specific assumptions. As a result, the model-driven methods cannot describe the complex nonlinearity of the traffic system, and have not been widely applied [5]. In this era of big data, many scholars have attempted to predict traffic flow based on the massive traffic data, eliminating the need for multiple assumptions, creating and implementing lots of data-driven forecast methods [6], [7]. The typical data-driven

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

forecast methods are machine learning (ML) tools like neural networks (NNs). Currently, novel traffic forecast methods are emerging constantly. Among them, deep learning (DL) strategies stand out for its effective use of massive raw data.

Despite the extensive studies and fruitful results on traffic flow, the complex evolution of traffic flow on surface roads is not yet clear. There is no definite answer to the following questions: Whether the traffic flow time series is continuous? Does the traffic flow time series share the same statistical features with other time series? Whether the traffic flow is predictable? How far in advance could the traffic flow be predicted? The time series analysis provides a possible solution to these questions. Over the years, much attention has been paid to the analysis of classic and nonlinear time series, and the analysis results have been successfully applied to traffic management like traffic control and route guidance.

Considering the superiority of the DL in processing the big data on traffic, this paper identifies the periodicity and stationarity of traffic flow time series measured on roads in Changsha, Central China's Hunan Province, and then constructs a traffic flow forecast model based on the long short-term memory (LSTM) network. The proposed model was compared with two classic forecast models at different prediction periods, in rush hours, and in non-rush hours. Our research discloses the dynamic evolution law of traffic flow, facilitating the decision-making of traffic management.

The remainder of this paper is organized as follows: Section 2 reviews the literature on traffic flow forecast; Section 3 introduces the proposed model and two contrastive models; Section 4 evaluates the performance of our model with real-world data; Section 5 puts forward conclusions and looks forward to the future research.

II. LITERATURE REVIEW

In recent years, the limited road resources in many cities can no longer accommodate the surging number of vehicles. The imbalance between travel demand and transport capability has worsened urban congestion. The ensuring problems severely bottleneck the sustainable development of urban transport, including the rising traffic accidents, growing environmental pollution and reduced travel efficiency.

As mentioned before, traffic flow forecast is the precondition for traffic management measures like traffic planning, route guidance, and traffic control. The accurate forecast of future traffic flow makes it possible to evaluate the economy and effectiveness of the ITS. To ensure the forecast accuracy, the daily mean traffic flow must be sufficiently smooth. The smoother the daily mean traffic flow, the lower the forecast error.

The application of time series analysis in traffic flow forecast can be traced back to the 1970s [8]. Time series analysis enjoys immense popularity in the real world, because its model is simple and only relies on traffic flow time series. There are currently three kinds of traffic flow methods based on time series analysis: model-driven methods, data-driven methods, and combinatory models.

The model-driven methods mainly include the autoregressive integrated moving average (ARIMA) model [9]–[11], seasonal ARIMA (SARIMA) model [12], [13], Markov chain (MC) [14]–[17], Bayesian network (BN) [18]–[20], and Kalman filter (KF) [21]–[23]. These methods cannot perform normally without several preconditions, e.g. the model structure is predefined, the residuals obey the normal distribution, and the time series is stationary. These preconditions fail to demonstrate the nonlinearity and randomness of actual traffic flows. Hence, the model-driven methods have a limited applicable scope in the transport field.

The emerging technique of artificial intelligence (AI) has made it possible to manage traffic based on big data. Thus, the data-driven methods have become the mainstream strategy for traffic flow prediction. The most famous data-driven method is the NNs, which consist of fully-connected layers and radial basis functions (RBFs) [24]–[27]. Nonetheless, the traditional NNs are too shallow to effectively handle a massive amount of long time series.

Recently, more and more scholars have implemented the DL in traffic flow forecast. The DL can effectively utilize the information in the videos, images, time series, and spatial series of traffic data, which greatly enhances the forecast accuracy [28]–[36]. Lv *et al.* [37] designed a stacked auto encoder (SAE) model for traffic flow forecast, and proved the superiority of the SAE model over multilayer perceptron (MLP). Xiang *et al.* [28], [38] developed stacked denoising auto encoder (SDAE) to predict the traffic flow with missing data. Xie *et al.* [39] and other scholars analyzed the transport problem by a deep belief network (DBN) [40]–[43]. Sun *et al.* [44] and other scholars applied convolutional neural network (CNN) to process time series data for traffic forecast [45]–[47]. Hu *et al.* [48] and other scholars demonstrated the excellence of recurrent neural network (RNN) in traffic flow prediction based on time series analysis [49]–[55].

Adding a feedback loop to the hidden layer, the RNN can handle the correlations between data at different moments well. If the time series is very long, however, vanishing gradient and exploding gradient may occur in the RNN training. To solve the problems, Hochreiter and Schmidhuber [56] proposed the LSTM network, in which the traditional hidden layer neurons are replaced with a memory block. The block prevents vanishing and exploding gradients in prolonged trainings. Hence, the LSTM enables the RNN to process long time series. Many LSTM-based DL models have been successfully applied to forecast traffic flow based on time series [57]–[63]. However, it remains a challenge to forecast traffic flow accurately and reliably in the complex, nonlinear, time-varying traffic system.

Each forecast model has its merits and defects, and specific applicable scope. To combine the merits and resolve the defects, single models have been integrated into various combinatory forecast models. For example, Luo *et al.* [64] proposed a short-term traffic flow forecast model based on the k-nearest neighbors (k-NN) algorithm and the LSTM. Guo *et al.* [65] combined support vector regression (SVR)

and the LSTM into a hybrid forecast model for traffic flow. Based on the RNN and the LSTM, Wang *et al.* [50] presented a combinatory forecast model for traffic flow on ground-level roads, and verified its forecast performance of the proposed model. Despite their high forecast accuracy, the combinatory models are not widely applied in traffic engineering, because of their complexity and poor real time performance.

To sum up, the model-driven methods predict the future traffic flow by setting up accurate mathematical models based on the statistical features of the raw data. These models are easy to interpret, but many parameters and assumptions need to be determined in advance. Thus, few model-driven methods can adapt to the stochastic traffic flows, or make highly accurate predictions. Combinatory models are more accurate than model-driven methods. However, these models are not widely applied in traffic engineering, due to their complexity and poor real time performance. By contrast, the ML-based forecast methods are not constrained by preconditions and strong in nonlinear approximation and self-learning. Among them, artificial neural networks (ANNs) are well received, owing to their mature theories and good predictive performance. The LSTM, a superior DL network for time series analysis, has attracted a growing attention from researchers engaging in traffic flow forecast.

III. PREDICTION MODELS

The time series of traffic flow is a set of observations arranged in chronological order. If the observations only contain one variable, the series is a univariate time series. To analyze the time series, a basic assumption is required: the future traffic flow is affected by the past values. Unlike other models, time series models take the target variable as predictor variable.

In this paper, three representative time series analysis methods, namely, ARIMA model, backpropagation neural network (BPNN) and LSTM, are compared in the efficiency of traffic flow forecast. The ARIMA model has been broadly adopted to forecast traffic flow, for its strong potential for online implementation. The BPNN is the most popular method to process traffic data without an accurate mathematical model. However, the shallow structure of BPNN has a low efficiency in handling massive data. Many DL models have been developed to deal with massive data in an effective manner. Among them, the LSTM is the most promising one for time series analysis.

A. ARIMA MODEL

The ARIMA model provides a regression method for time series: First, the model judges whether the target time series is stationary; if the time series is non-stationary, it will be subjected to differential treatment and modified into a stationary one for modelling. For the traffic flow on roads, the continuous observations are correlated in the time series. The traffic flow data might be non-stationary, due to the stochasticity and complexity of the traffic system. Hence, the ARIMA model has often been used to predict the traffic flow based on the time series [66].

If the traffic flow time series X_t is stationary, the time series can be described as a linear combination of the previous traffic flows by autoregressive moving average (ARMA) model:

$$X_t = \varphi_1 X_{t-1} + \varphi_2 X_{t-2} + \dots + \varphi_p X_{t-p} + \mu_t - \theta_1 \mu_{t-1} - \theta_2 \mu_{t-2} - \dots - \theta_q \mu_{t-q} \quad (1)$$

where, p and q are the orders of the model; $\varphi_1, \varphi_2, \dots, \varphi_p$ are autoregressive coefficients; $\theta_1, \theta_2, \dots, \theta_q$ are moving average coefficients; μ_t is the residual at time t . By introducing the backshift operator B and $B^j X_t = X_{t-j}$, formula (1) can be simplified as:

$$\phi(B) X_t = \theta(B) \mu_t \quad (2)$$

where, $\phi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p$; $\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$.

Since most traffic flow time series are non-stationary, the differential treatment is implemented by replacing X_t with $(1 - B)^d X_t$, where $d \leq 2$ is the difference coefficient of X_t . Thus, the ARIMA model can be formulated as:

$$(1 - B)^d \phi(B) X_t = \theta(B) \mu_t \quad (3)$$

The key to ARIMA modelling is to determine the values of p and q . The values of parameters like ϕ_i and θ_i can be derived through least squares estimation, moment estimation, maximum likelihood estimation, etc. Once the parameter values are identified, the ARIMA model can be employed to predict the future traffic flow.

B. BPNN

The BPNN is a multi-layer feedforward NN, in which the signal propagates forward and the error propagate backwards through each layer. In the BPNN, the neurons on the current layer only affect the state of those on the next layer. The signal passes from the input layer through the hidden layer to the output layer. Meanwhile, the error passes through the network in the opposite direction, and adjusts the weight and bias of each layer, thereby optimizing the model continuously. In this way, the prediction of the BPNN gradually approaches the desired value [24], [26].

To describe the iterative learning of the BPNN, the following parameters were defined: the input x ; the input of the hidden layer h_i ; the output of the hidden layer h_o ; the input of the output layer y_i ; the output of the output layer y_o ; the desired output d_o ; the connection weight between the input layer and the hidden layer w_{ih} ; the connection weight between the hidden layer and the output layer w_{ho} ; the bias of each hidden layer neuron b_h ; the bias of each output layer neuron b_o . Then, the learning process of the BPNN can be explained as follows:

Step 1. Network initialization

According to the prediction target and the known conditions, determine the number of neurons on each layer, define the biases and initial weights of the hidden layer and output layer, and specify the calculation precision ε and maximum number of iterations M .

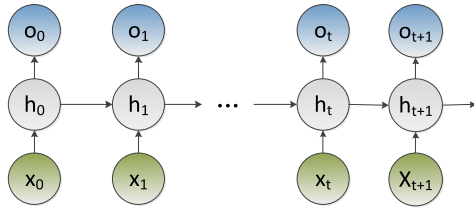


FIGURE 1. Structure of the RNN.

- Step 2. Hidden layer calculation
Derive h_o from x , w_{ih} and b_h .
- Step 3. Output layer calculation
Derive y_o from y_i , w_{ho} and b_o .
- Step 4. Error calculation
Derive error of the BPNN e from y_o and d_o .
- Step 5. Weight update
Update the weight and bias of each layer based on e .
- Step 6. Termination

Terminate the iterative learning if the error satisfies the calculation precision or the maximum number of iterations is reached. The trained BPNN is ready to predict new samples.

C. LSTM NETWORK

As mentioned before, the RNN (Figure 1) relies on the feedback loop in the hidden layer to clarify the correlation between data at different moments, where x_t , h_t and o_t are the values of input, hidden layer and output. But the standard RNN may suffer from vanishing and exploding gradients in the training process, if the data form a long time series. Hence, the RNN was extended by Hochreiter and Schmidhuber [56] into the LSTM network. Compared with the RNN, the LSTM network has a memory block instead of hidden layer neurons, which effectively prevents vanishing and exploding gradients in prolonged trainings.

In the LSTM network, several gates are introduced to control the memory of the RNN. During the training, the weight and bias of each gate are learned from the historical time series, and the features of historical states are identified and memorized. On this basis, the trained network can estimate the future state from new input samples. As a result, the LSTM network [28] can fully consider the long-term correlations between traffic flows, and make effective forecast of future traffic flow.

As shown in Figure 2, the memory block, the core of the LSTM network, consists of a memory cell C_t , an input gate i_t , a forget gate f_t , and an output gate o_t . The operations of the memory block are explained as follows:

1) FORGETTING

The redundant information in long-term cell states is discarded by the forget gate. Based on the input x_t , the forget gate f_t output a matrix x_t of elements $\in(0, 1)$. Each element is multiplied with the corresponding element in the cell state matrix C_{t-1} :

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (4)$$

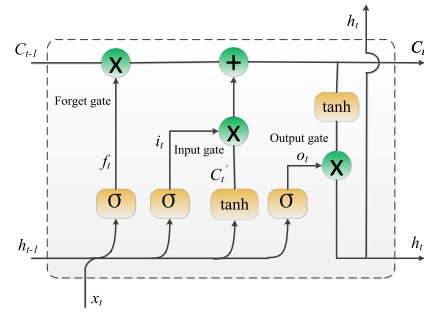


FIGURE 2. Structure of the memory block in the LSTM.

2) MEMORIZATION

The new information is stored as long-term cell states in three steps: first, each tanh function creates a new candidate vector; next, the sigmoid function of the input gate i_t updates some elements of each candidate vector; after that, the new information is added to long-term cell states:

$$C'_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (5)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (6)$$

$$C_t = f_t \otimes C_{t-1} + i_t \otimes C'_t \quad (7)$$

3) INFORMATION OUTPUT

The sigmoid function determines the output information h_t : the long-term cell states are processed by the tanh function, and then multiplied with the information filtered by the output gate o_t :

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (8)$$

$$h_t = o_t \otimes \tanh(C_t) \quad (9)$$

where, C'_t is the updated cell state at time t ; C_t and C_{t-1} are the outputs at time t and time $t - 1$, respectively; W_c and b_c are the weight and bias of the cell, respectively; i_t , f_t , o_t and h_t are the outputs of the input gate, the forget gate, the output gate and the hidden layer at time t , respectively; x_t is the input at time t ; h_{t-1} is the output of the hidden layer at time $t - 1$; W_i , W_f and W_o are the weights of the input gate, forget gate and output gate, respectively; b_i , b_f and b_o are the biases of the input gate, forget gate and output gate, respectively.

The activation functions sigmoid and tanh can be respectively expressed as:

$$f(x) = \frac{1}{(1 + e^{-x})} \quad (10)$$

$$f(x) = \frac{(e^x - e^{-x})}{(e^x + e^{-x})} \quad (11)$$

IV. EXPERIMENTS AND RESULTS ANALYSIS

A. DATA PREPROCESSING

Downloaded from OpenITS, the original data were collected at an interval of 5min in 10 days (September 17-26, 2013) by an inductive loop deployed beneath the road surface at No.1 Yuanda Road, Changsha, and the SCATS Traffic Reporter.

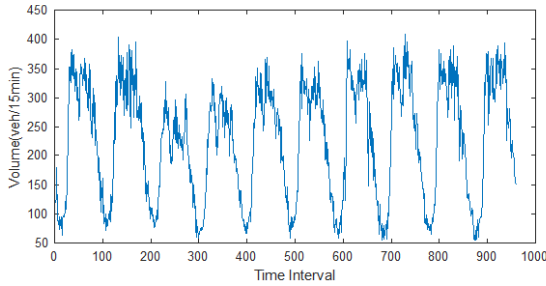


FIGURE 3. Our traffic flow time series.

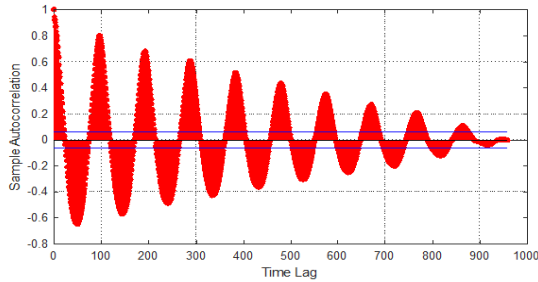


FIGURE 4. Autocorrelation curve of our traffic flow time series.

The original data were transformed into 960 samples (Figure 3) with 15min intervals, i.e. 96 sampling points per day. Then, the data of the first eight days were taken as the training set to build up forecast models, and the data of the other two days were taken as the test set to verify the effectiveness of each model. Hence, the training set and the test set respectively contain 768 samples and 192 samples.

As most traffic flow time series, the preprocessed data change periodically between days, and have similarities between some periods on the same day [67]. Morning and evening rush hours are clearly seen from the daily variation in traffic flow. The periodicity of traffic flow is a feature of the traffic system that should be considered to achieve accurate prediction of future traffic flow.

To judge if our traffic flow time series is periodic, autocorrelation analysis was carried out. The autocorrelation coefficient r_k can be computed by:

$$r_k = \frac{\sum_{t=1}^{n-k} (X_t - \bar{X})(X_{t-k} - \bar{X})}{\sum_{t=1}^n (X_t - \bar{X})^2} \quad (12)$$

where, X_{t-k} is the traffic flow time series that lags X_t by k intervals; \bar{X} is the mean traffic flow; n is the length of time series.

According to the autocorrelation curve (Figure 4), our traffic flow time series has significant periodicity. Despite the fluctuations, the data in the series tend to be stable.

B. PARAMETER SETTINGS

The LSTM network for our experiments contains five layers, including an input layer, three hidden layers and an output layer. Every layer contains 96 neurons. The preprocessed data was normalized as a time series with zero mean and unit variance. During training, there is one interval difference between each input and each output: if the input is x_t , then the

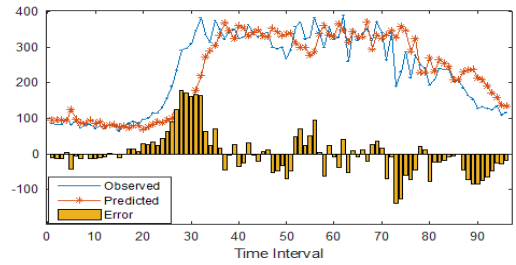


FIGURE 5. The predicted results of the ARIMA model on the samples of September 25th, 2013.

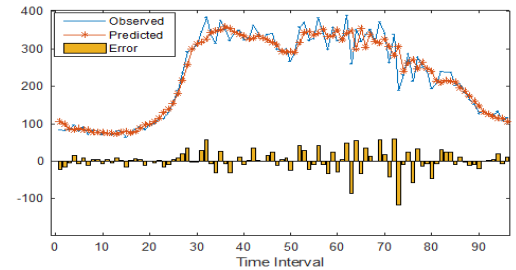


FIGURE 6. The predicted results of the BPNN on the samples of September 25th, 2013.

output will be x_{t+1} . The maximum number of iterations was set to 200, the gradient threshold was set to 1, and the learning rate was initialized as 0.005. After 100 iterations, the learning rate was reduced to the product between the initial value and a drop factor of 0.2. The dropout on each layer was set to 0.2.

For the ARIMA model, p , d and q were all set to 1. Meanwhile, the BPNN was configured as a three-layer network with one input layer neuron, ten hidden layer neurons and one output layer neuron.

C. RESULTS ANALYSIS

Based on the above parameters, our experiments were conducted on the MATLAB. Each forecast model, namely, the ARIMA model, the BPNN and the LSTM network, was trained by the training set and verified by the test set.

Two plans were designed to compare the forecast effects of the three models: (1) The daily predictions of the three models were compared to reveal how the prediction performance changes with the elapse of time; (2) The 2h-interval predictions of the three models were compared to disclose the difference in prediction performance between rush hours and non-rush hours.

Firstly, the ARIMA model, BPNN and LSTM network were tested by the 96 samples on September 25th, 2013. The predicted values, actual values and residual errors of the three models are displayed in Figures 5-7, respectively. It can be seen that the LSTM network achieved the best prediction performance (i.e. the smallest deviation between the predicted value and the actual value), followed in turn by the BPNN and the ARIMA model.

Secondly, the ARIMA model, BPNN and LSTM network were tested by the 96 samples on September 26th, 2019. The predicted values, actual values and residual errors of

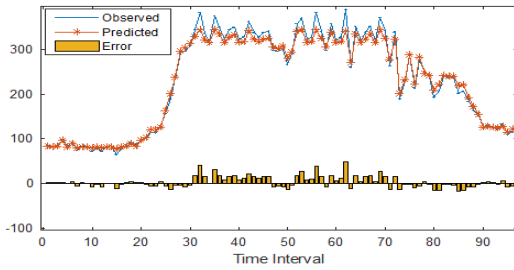


FIGURE 7. The predicted results of the LSTM network on the samples of September 25th, 2013.

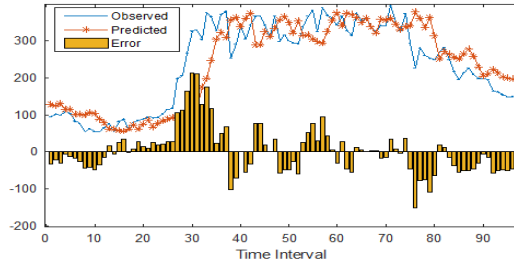


FIGURE 8. The predicted results of the ARIMA model on the samples of September 26th, 2013.

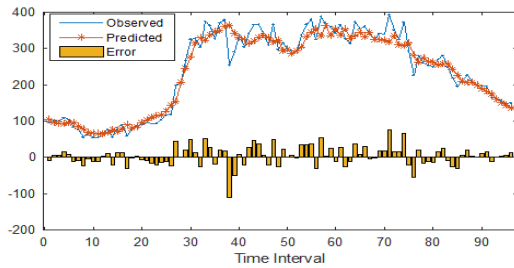


FIGURE 9. The predicted results of the BPNN on the samples of September 26th, 2013.

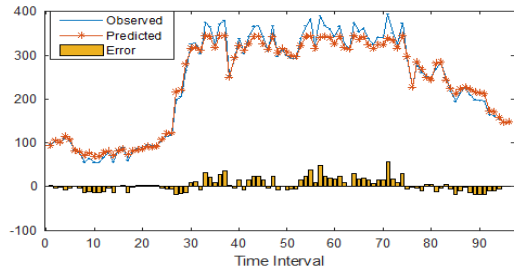


FIGURE 10. The predicted results of the LSTM network on the samples of September 26th, 2013.

the three models are displayed in Figures 8-10, respectively. It can be seen that the LSTM network achieved the best prediction performance (i.e. the smallest deviation between the predicted value and the actual value), followed in turn by the BPNN and the ARIMA model.

The prediction performance of each model was measured by three metrics: mean absolute error (MAE), mean absolute percentage error (MAPE) and root mean square error (RMSE):

$$MAE = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t| \quad (13)$$

TABLE 1. The metrics of the three models.

Performance		ARIMA	BPNN	LSTM
Test samples of September 25th, 2013	RMSE	58.9666	26.7399	12.9668
	MAE	45.5681	18.4108	9.4797
	MAPE (%)	0.1951	0.0840	0.0398
Test samples of September 26th, 2013	RMSE	63.2966	27.0140	15.7832
	MAE	45.9916	20.2150	11.9351
	MAPE (%)	0.2243	0.0973	0.0565
All samples	RMSE	61.1699	26.8773	14.4438
	MAE	43.7798	19.3129	10.7074
	MAPE (%)	20.97	9.06	4.82

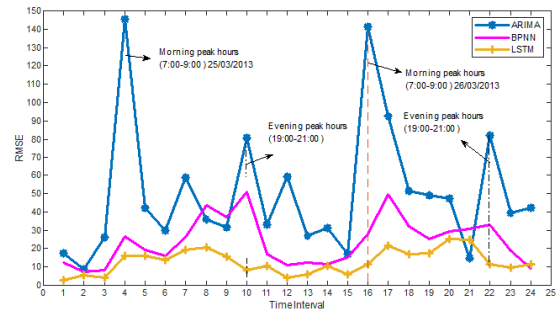


FIGURE 11. The RMSE curves of three models with 2h intervals.

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right| \times 100\% \quad (14)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2} \quad (15)$$

where, y_t is the actual traffic flow; \hat{y}_t is the predicted traffic flow; n is the number of samples. The three metrics are negatively correlated with the goodness-of-fit and prediction

Table 1 lists the metrics of each model on daily predictions. It can be seen that the LSTM network achieved the best RMSE, MAE and MAPE, followed in turn by the BPNN model and the ARIMA model. Hence, the LSTM network boasts the best prediction performance. For different time steps, the LSTM predictions on September 26th, 2013 were less accurate than those on September 25th, 2013, indicating that a long time step suppress the prediction accuracy. On the samples of both days, the LSTM network outperformed the two contrastive models, showcasing the superiority of the DL.

Figure 11 presents the RMSE curve of each model with 2h intervals. The 4th and 16th intervals (7-9am) are the morning rush hours of September 25 and September 26, 2013, respectively; the 10th and 22nd intervals (7-9pm) are the evening rush hours of September 25 and September 26, 2013, respectively. Obviously, the ARIMA model had a far greater RMSE in the morning and evening peak hours than that in the non-peak hours. Thus, this model cannot accurately forecast the traffic flow in peak hours. The BPNN model also showed a large RMSE in in the morning and evening peak hours, but not as high as that of the ARIMA model. The LSTM network controlled the RMSE at a low level in the morning and evening peak hours, an evidence of its excellent forecast effect on long time series.

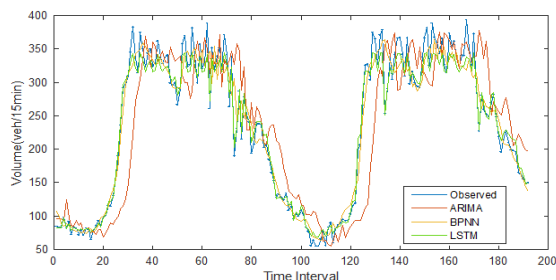


FIGURE 12. Comparison between predicted values and observed value.

V. CONCLUSION

Traffic flow forecast is critical to measures of traffic management, such as route guidance and traffic control. The significance of traffic flow forecast is growing with the proliferation of the ITS and autopilot. Currently, lots of traffic flow data are available thanks to the development of information technology (IT) and big data. Thus, it is very meaningful to mine out the evolution law of traffic flow through time series analysis, and make accurate forecast of future traffic flow. Therefore, this paper compares the traffic flow forecast effects of the LSTM network, BPNN model and ARIMA model on time series captured at a single point. The main conclusions are as follows:

(1) The contrastive experiments show that the LSTM network outperformed the ARIMA model and the BPNN model in prediction accuracy: the mean RMSEs of the ARIMA, BPNN and LSTM were 61.1699, 26.8773 and 14.4438, respectively.

(2) For different time steps, the RMSE of the LSTM network was 12.9668 on September 25, 2013 and 15.7832 on September 26, 2013. Hence, the prediction accuracy of the LSTM network was higher on the first day, which reflects the high correlation between adjacent data in time series.

(3) For 2h intervals, the ARIMA model had relatively large prediction errors in morning and evening peak hours. This is because the traffic flow tends to fluctuate in peak hours under various disturbances, while the ARIMA is not good at predicting highly volatile data.

(4) As shown in Figure 12, the predicted values of the three models basically agreed with the observed value, due to the relative stability of traffic flow time series under normal conditions. The LSTM network made the most realistic prediction. Besides, the superiority of the DL in traffic flow prediction was confirmed by the results in Table 1: the MAPEs of the ARIMA, BPNN and LSTM were 20.97%, 9.06% and 4.82%, respectively.

The proposed LSTM network can accurately predict the traffic flow based on the relatively stable time series under normal conditions. However, the traffic system on roads is stochastic and complex, and often affected by abnormal factors like bad weather, traffic accident and large events. Therefore, the future research will fully utilize the abilities of the ML (e.g. big data processing and self-learning) to predict the traffic flow under various influencing factors.

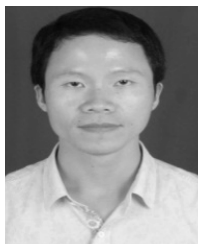
ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers for their worthwhile comments and the OpenITS for providing the experimental data.

REFERENCES

- [1] I. Spyropoulou, "Impact of public transport strikes on the road network: The case of Athens," *Transp. Res. A, Policy Pract.*, vol. 132, pp. 651–665, Feb. 2020.
- [2] J. Ma, M. Xu, Q. Meng, and L. Cheng, "Ridesharing user equilibrium problem under OD-based surge pricing strategy," *Transp. Res. B, Methodol.*, vol. 134, pp. 1–24, Apr. 2020.
- [3] J. Sun, J. Wu, F. Xiao, Y. Tian, and X. Xu, "Managing bottleneck congestion with incentives," *Transp. Res. B, Methodol.*, vol. 134, pp. 143–166, Apr. 2020.
- [4] D. Nallaperuma, R. Nawaratne, T. Bandaragoda, A. Adikari, S. Nguyen, T. Kempitiya, D. De Silva, D. Alahakoon, and D. Pothuhera, "Online incremental machine learning platform for big data-driven smart traffic management," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 12, pp. 4679–4690, Dec. 2019.
- [5] M. Ma, S. Liang, H. Guo, and J. Yang, "Short-term traffic flow prediction using a self-adaptive two-dimensional forecasting method," *Adv. Mech. Eng.*, vol. 9, no. 8, pp. 1–12, Aug. 2017.
- [6] H.-H. Chang and B.-J. Yoon, "High-speed data-driven methodology for real-time traffic flow predictions: Practical applications of ITS," *J. Adv. Transp.*, vol. 2018, pp. 1–11, 2018.
- [7] M. Tang, Z. Li, and G. Tian, "A data-driven-based wavelet support vector approach for passenger flow forecasting of the metropolitan hub," *IEEE Access*, vol. 7, pp. 7176–7183, 2019.
- [8] D. C. Gazis and C. H. Knapp, "On-line estimation of traffic densities from time-series of flow and speed data," *Transp. Sci.*, vol. 5, no. 3, pp. 283–301, Aug. 1971.
- [9] C. Ding, J. Duan, Y. Zhang, X. Wu, and G. Yu, "Using an ARIMA-GARCH modeling approach to improve subway short-term ridership forecasting accounting for dynamic volatility," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 4, pp. 1054–1064, Apr. 2018.
- [10] B. Moghimi, A. Safikhani, C. Kanga, and W. Hao, "Cycle-length prediction in actuated traffic-signal control using ARIMA model," *J. Comput. Civil Eng.*, vol. 32, no. 2, Mar. 2018, Art. no. 04017083, doi: 10.1061/(ASCE)CP.1943-5487.0000730.
- [11] H. Wang, L. Liu, S. Dong, Z. Qian, and H. Wei, "A novel work zone short-term vehicle-type specific traffic speed prediction model through the hybrid EMD-ARIMA framework," *Transportmetrica B, Transp. Dyn.*, vol. 4, no. 3, pp. 159–186, Sep. 2016.
- [12] X. Luo, L. Niu, and S. Zhang, "An algorithm for traffic flow prediction based on improved SARIMA and GA," *KSCE J. Civil Eng.*, vol. 22, no. 10, pp. 4107–4115, Oct. 2018.
- [13] M. Milenković, L. Švadlenka, V. Melicher, N. Bojović, and Z. Avramović, "SARIMA modelling approach for railway passenger flow forecasting," *Transport*, vol. 33, no. 5, pp. 1113–1120, 2018.
- [14] S. Zhang, Z. Kang, Z. Zhang, C. Lin, C. Wang, and J. Li, "A hybrid model for forecasting traffic flow: Using layerwise structure and Markov transition matrix," *IEEE Access*, vol. 7, pp. 26002–26012, 2019.
- [15] H. Qi and X. Hu, "Real-time headway state identification and saturation flow rate estimation: A hidden Markov chain model," *Transportmetrica A, Transp. Sci.*, vol. 16, no. 3, pp. 840–864, Jan. 2020.
- [16] M. Yun, W. Qin, X. Yang, and F. Liang, "Estimation of urban route travel time distribution using Markov chains and pair-copula construction," *Transportmetrica B, Transp. Dyn.*, vol. 7, no. 1, pp. 1521–1552, Dec. 2019.
- [17] M. Abareshi, M. Zaferanieh, and M. R. Safi, "Origin-destination matrix estimation problem in a Markov chain approach," *Netw. Spatial Econ.*, vol. 19, no. 4, pp. 1069–1096, Dec. 2019.
- [18] Z. Li, S. Jiang, L. Li, and Y. Li, "Building sparse models for traffic flow prediction: An empirical comparison between statistical heuristics and geometric heuristics for Bayesian network approaches," *Transportmetrica B, Transp. Dyn.*, vol. 7, no. 1, pp. 107–123, Dec. 2019.
- [19] Z. Zhu, S. Zhu, Z. Zheng, and H. Yang, "A generalized Bayesian traffic model," *Transp. Res. C, Emerg. Technol.*, vol. 108, pp. 182–206, Nov. 2019.
- [20] S. Wang, W. Huang, and H. K. Lo, "Traffic parameters estimation for signalized intersections based on combined shockwave analysis and Bayesian network," *Transp. Res. C, Emerg. Technol.*, vol. 104, pp. 22–37, Jul. 2019.

- [21] L. Cai, Z. Zhang, J. Yang, Y. Yu, T. Zhou, and J. Qin, "A noise-immune Kalman filter for short-term traffic flow forecasting," *Phys. A, Stat. Mech. Appl.*, vol. 536, Dec. 2019, Art. no. 122601.
- [22] T. Zhou, D. Jiang, Z. Lin, G. Han, X. Xu, and J. Qin, "Hybrid dual Kalman filtering model for short-term traffic flow forecasting," *IET Intell. Transp. Syst.*, vol. 13, no. 6, pp. 1023–1032, Jun. 2019.
- [23] S. Liang, M. Ma, S. He, and H. Zhang, "Short-term passenger flow prediction in urban public transport: Kalman filtering combined K-Nearest neighbor approach," *IEEE Access*, vol. 7, pp. 120937–120949, 2019.
- [24] M. S. Ghanim and G. Abu-Lebdeh, "Projected state-wide traffic forecast parameters using artificial neural networks," *IET Intell. Transp. Syst.*, vol. 13, no. 4, pp. 661–669, Apr. 2019.
- [25] Z. Liu, J. Guo, J. Cao, Y. Wei, and W. Huang, "A hybrid short-term traffic flow forecasting method based on neural networks combined with K-nearest neighbor," *PROMET-Traffic&Transp.*, vol. 30, no. 4, pp. 445–456, 2018.
- [26] T. Pamula, "Impact of data loss for prediction of traffic flow on an urban road using neural networks," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 3, pp. 1000–1009, Mar. 2019.
- [27] M. X. Liu, L. Xia, J. Zhong, N. N. Dou, and B. Li, "Is it necessary to approach the compressed vertebra bilaterally during the process of PKP?" *J. Spinal Cord Med.*, vol. 43, no. 2, pp. 201–205, Mar. 2020.
- [28] Y. Duan, Y. Lv, Y.-L. Liu, and F.-Y. Wang, "An efficient realization of deep learning for traffic data imputation," *Transp. Res. C, Emerg. Technol.*, vol. 72, pp. 168–181, Nov. 2016.
- [29] W. Meng, C. Mao, J. Zhang, J. Wen, and D. Wu, "A fast recognition algorithm of online social network images based on deep learning," *Traitement Du Signal*, vol. 36, no. 6, pp. 575–580, Dec. 2019.
- [30] Y. Gu, W. Lu, L. Qin, M. Li, and Z. Shao, "Short-term prediction of lane-level traffic speeds: A fusion deep learning model," *Transp. Res. C, Emerg. Technol.*, vol. 106, pp. 1–16, Sep. 2019.
- [31] Y. Liu, Z. Liu, and R. Jia, "DeepPF: A deep learning based architecture for metro passenger flow prediction," *Transp. Res. C, Emerg. Technol.*, vol. 101, pp. 18–34, Apr. 2019.
- [32] 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.
- [33] Y. Wu, H. Tan, L. Qin, B. Ran, and Z. Jiang, "A hybrid deep learning based traffic flow prediction method and its understanding," *Transp. Res. C, Emerg. Technol.*, vol. 90, pp. 166–180, May 2018.
- [34] C. Xu, J. Ji, and P. Liu, "The station-free sharing bike demand forecasting with a deep learning approach and large-scale datasets," *Transp. Res. C, Emerg. Technol.*, vol. 95, pp. 47–60, Oct. 2018.
- [35] R. Neelapu, G. Lavanya Devi, and K. Srinivasa Rao, "Deep learning based conventional neural network architecture for medical image classification," *Traitement Du signal*, vol. 35, no. 2, pp. 169–182, Jun. 2018.
- [36] Z. Zhang, M. Li, X. Lin, Y. Wang, and F. He, "Multistep speed prediction on traffic networks: A deep learning approach considering spatio-temporal dependencies," *Transp. Res. C, Emerg. Technol.*, vol. 105, pp. 297–322, Aug. 2019.
- [37] Y. Lv, Y. Duan, W. Kang, Z. Li, and F.-Y. Wang, "Traffic flow prediction with big data: A deep learning approach," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 2, pp. 865–873, Sep. 2014.
- [38] J. Xiang and Z. Chen, "Traffic state estimation of signalized intersections based on stacked denoising auto-encoder model," *Wireless Pers. Commun.*, vol. 103, no. 1, pp. 625–638, Nov. 2018.
- [39] D.-F. Xie, Z.-Z. Fang, B. Jia, and Z. He, "A data-driven lane-changing model based on deep learning," *Transp. Res. C, Emerg. Technol.*, vol. 106, pp. 41–60, Sep. 2019.
- [40] L. Li, L. Qin, X. Qu, J. Zhang, Y. Wang, and B. Ran, "Day-ahead traffic flow forecasting based on a deep belief network optimized by the multi-objective particle swarm algorithm," *Knowl.-Based Syst.*, vol. 172, pp. 1–14, May 2019.
- [41] F. Kong, J. Li, B. Jiang, and H. Song, "Short-term traffic flow prediction in smart multimedia system for Internet of vehicles based on deep belief network," *Future Gener. Comput. Syst.*, vol. 93, pp. 460–472, Apr. 2019.
- [42] L. Zhao, Y. Zhou, H. Lu, and H. Fujita, "Parallel computing method of deep belief networks and its application to traffic flow prediction," *Knowl.-Based Syst.*, vol. 163, pp. 972–987, Jan. 2019.
- [43] Y. Zhang and G. Huang, "Traffic flow prediction model based on deep belief network and genetic algorithm," *IET Intell. Transp. Syst.*, vol. 12, no. 6, pp. 533–541, Aug. 2018.
- [44] S. Sun, H. Wu, and L. Xiang, "City-wide traffic flow forecasting using a deep convolutional neural network," *Sensors*, vol. 20, no. 2, p. 421, 2020.
- [45] S. Guo, Y. Lin, S. Li, Z. Chen, and H. Wan, "Deep spatial-temporal 3D convolutional neural networks for traffic data forecasting," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 10, pp. 3913–3926, Oct. 2019.
- [46] Z. Ye, Q. Chen, Y. Zhang, J. Zou, and Y. Zheng, "Identification of vortex structures in flow field images based on convolutional neural network and dynamic mode decomposition," *Traitement Du Signal*, vol. 36, no. 6, pp. 501–506, Dec. 2019.
- [47] M. Khajeh Hosseini and A. Talebpour, "Traffic prediction using time-space diagram: A convolutional neural network approach," *Transp. Res. Rec., J. Transp. Res. Board*, vol. 2673, no. 7, pp. 425–435, Jul. 2019.
- [48] R. Hu, Y.-C. Chiu, and C.-W. Hsieh, "Crowding prediction on mass rapid transit systems using a weighted bidirectional recurrent neural network," *IET Intell. Transp. Syst.*, vol. 14, no. 3, pp. 196–203, Mar. 2020.
- [49] D. Zang, J. Ling, Z. Wei, K. Tang, and J. Cheng, "Long-term traffic speed prediction based on multiscale spatio-temporal feature learning network," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 10, pp. 3700–3709, Oct. 2019.
- [50] 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.
- [51] D. Duives, G. Wang, and J. Kim, "Forecasting pedestrian movements using recurrent neural networks: An application of crowd monitoring data," *Sensors*, vol. 19, no. 2, p. 382, 2019.
- [52] C. Li, J. Wang, and X. Ye, "Using a recurrent neural network and restricted Boltzmann machines for malicious traffic detection," *NeuroQuantology*, vol. 16, no. 5, pp. 823–831, 2018.
- [53] A. Fandango and R. P. Wiegand, "Towards investigation of iterative strategy for data mining of short-term traffic flow with recurrent neural networks," in *Proc. 2nd Int. Conf. Inf. Syst. Data Mining (ICISDM)*, 2018, pp. 65–69.
- [54] Y. Song and J. Lu, "RNN-based traffic flow prediction for dynamic reversible lane control decision," *Data Sci. Knowl. Eng. Sens. Decis. Support*, vol. 1, pp. 323–330, Sep. 2018.
- [55] Y.-J. Lee and O. Min, "Long short-term memory recurrent neural network for urban traffic prediction: A case study of seoul," in *Proc. 21st Int. Conf. Intell. Transp. Syst. (ITSC)*, Nov. 2018, pp. 1279–1284.
- [56] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [57] B. Du, H. Peng, S. Wang, M. Z. A. Bhuiyan, L. Wang, Q. Gong, L. Liu, and J. Li, "Deep irregular convolutional residual LSTM for urban traffic passenger flows prediction," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 3, pp. 972–985, Mar. 2020.
- [58] L. Mou, P. Zhao, H. Xie, and Y. Chen, "T-LSTM: A long short-term memory neural network enhanced by temporal information for traffic flow prediction," *IEEE Access*, vol. 7, pp. 98053–98060, 2019.
- [59] M. Feng and Y. Cheng, "Optimization of Drop-and-Pull transport network based on shared freight station and Hub-and-Spoke network," *J. Européen des Systèmes Automatisés*, vol. 52, no. 5, pp. 457–464, Nov. 2019.
- [60] 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.
- [61] Y. Lu, X. Pei, C. Zhang, H. Luo, B. Liu, and Z. Ma, "Design of multimodal transport path optimization model and dual pheromone hybrid algorithm," *J. Européen des Systèmes Automatisés*, vol. 52, no. 5, pp. 477–484, Nov. 2019.
- [62] D. Yang, K. Chen, M. Yang, and X. Zhao, "Urban rail transit passenger flow forecast based on LSTM with enhanced long-term features," *IET Intell. Transp. Syst.*, vol. 13, no. 10, pp. 1475–1482, Oct. 2019.
- [63] J. Zhang, F. Chen, and Q. Shen, "Cluster-based LSTM network for short-term passenger flow forecasting in urban rail transit," *IEEE Access*, vol. 7, pp. 147653–147671, 2019.
- [64] X. Luo, D. Li, Y. Yang, and S. Zhang, "Spatiotemporal traffic flow prediction with KNN and LSTM," *J. Adv. Transp.*, vol. 2019, pp. 1–10, Feb. 2019.
- [65] J. Guo, Z. Xie, Y. Qin, L. Jia, and Y. Wang, "Short-term abnormal passenger flow prediction based on the fusion of SVR and LSTM," *IEEE Access*, vol. 7, pp. 42946–42955, 2019.
- [66] Q. Cui and J. Xia, "Time-varying confidence interval forecasting of travel time for urban arterials using ARIMA-GARCH model," *J. Southeast Univ. (English Ed.)*, vol. 2014, vol. 30, no. 3, pp. 358–362.
- [67] J. Tang, F. Liu, Y. Zou, W. Zhang, and Y. Wang, "An improved fuzzy neural network for traffic speed prediction considering periodic characteristic," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 9, pp. 2340–2350, Sep. 2017.



JIANHU ZHENG was born in Fuzhou, Fujian, China, in 1975. He received the M.S. degree in environment engineering from Fuzhou University, in 2004, and the Ph.D. degree in transportation information engineering and control from Tongji University, in 2007. From 2007 to 2010, he was a Research Assistant with the Department of Transportation, Minjiang University. Since 2011, he has been a Vice Professor with the School of Economics and Management, Minjiang University, where he has also been a Professor, since 2019. His research interests include machine learning, traffic safety, and intelligent transportation systems.



MINGFANG HUANG was born in Fuzhou, Fujian, China, in 1981. She received the B.S. degree in transportation engineering from Fujian Agriculture and Forestry University, in 2005, and the M.S. degree in transportation planning and management from Fujian Agriculture and Forestry University, in 2008. From 2008 to 2010, she was a Research Assistant with the Department of Transportation, Minjiang University, where she has been a Research Assistant with the School of Economics and Management, since 2011. Her current research interests include road traffic safety, transportation planning, and intelligent transportation systems.

• • •