
Jun Niu, Jiami He, Yurong Li, and Saifei Zhang

1 School of Electronics and Control Engineering, Changan University, Xi’an 710061, China
2 Airport Branch, Henan Transportation Development Group Co., Ltd, Zhengzhou 450016, China
3 School of Transportation Engineering, East China Jiaotong University, Nanchang 330013, China

Correspondence should be addressed to Jiami He; 2021138086100032@ecjtu.edu.cn

Received 28 April 2022; Accepted 18 July 2022; Published 31 August 2022

In order to accurately predict the short-time traffic flow of highways, to achieve the purpose of alleviating highway traffic congestion, saving travel time, and reducing energy waste and pollution, this paper considers the spatiotemporal characteristics of the road network, uses the advantages of long short-term memory network (LSTM) to analyze time series data, divides time intervals to collect traffic flow, and substitutes it into the model. Make traffic flow predictions. Three evaluation indicators, mean square error (MSE), root mean square error (RMSE), and mean absolute error (MAE), were used for training and testing on the section k602 + 630-k636 + 090 of Chang-Jiu Expressway. It is divided into average data with intervals of 1, 3, 5, and 10 minutes for prediction. The results show that each index obtained by dividing the scale of 1 minute is the smallest, which proves that the suitable prediction scale of road traffic flow is 1 minute.

1. Introduction

Accurate and timely traffic flow information is currently strongly needed for individual travelers, business sectors, and government agencies. It has the potential to help road users make better travel decisions, alleviate traffic congestion, reduce carbon emissions, and improve traffic operation efficiency. The objective of traffic flow prediction is to provide such traffic flow information. Generally, the prediction of long-time traffic flow means to calculate the value of traffic flow in a certain location by monitoring it for a month or a year for that period of time. However, the time-varying nature of traffic makes it nonprominent in real-life applications for predicting traffic flow for such a long period of time, while it has a very important value for predicting traffic flow for a shorter period of time. Timely adjustment of major roads through traffic flow short time prediction is one of the important ways to reduce road traffic congestion. Traffic flow prediction has gained more and more attention with the rapid development and deployment of intelligent transportation systems (ITs). It is regarded as a critical element for the successful deployment of ITS subsystems, particularly advanced traveler information systems, advanced traffic management systems, advanced public transportation systems, and commercial vehicle operations.

Traffic flow prediction heavily depends on historical and real-time traffic data collected from various sensor sources, including inductive loops, radars, cameras, mobile global positioning system, crowdsourcing, and social media. With the widespread traditional traffic sensors and new emerging traffic sensor technologies, traffic data are exploding, and we have entered the era of big data transportation. Transportation management and control are now becoming more data-driven. Although there have been already many traffic flow prediction systems and models, most of them use shallow traffic models and are still somewhat unsatisfying. This inspires us to rethink the traffic flow prediction problem.
based on deep architecture models with such a rich amount of traffic data [1].

Recently, deep learning, which is a type of machine learning method, has drawn a lot of academic and industrial interest. It has been applied with success in classification tasks, natural language processing, dimensionality reduction, object detection, motion modeling, and so on. Deep learning algorithms use multiple-layer architectures or deep architectures to extract inherent features in data from the lowest level to the highest level, and they can discover huge amounts of structure in the data. As a traffic flow process is complicated in nature, deep learning algorithms can represent traffic features without prior knowledge, which has good performance for traffic flow prediction.

In this paper, we propose a model for traffic flow prediction based on samples from the section k602 + 630- k636 + 090 of Chang-Jiu Expressway. By considering the spatiotemporal distribution characteristics of the road network, we use the advantages of a long short-term memory network (LSTM) to analyze the time series data, collect the data and predict the traffic flow. The contribution of this paper is to analyze the time series data through LSTM and divide the time interval to collect the traffic flow to estimate the traffic flow information of the road section. Our research also supports highway gantry systems to collect traffic flow data at a minimal cost.

The rest of this paper is organized as follows. Section 2 reviews the studies on traffic flow prediction. Section 3 introduces the long short-term memory (LSTM) method. Section 4 discusses the data sources and experimental results. Concluding remarks are described in Section 5.

2. Related Works

Traffic forecasting is a challenging problem. Key traffic metrics, such as flow and speed, exhibit complex spatial and temporal correlations that are difficult to model with classical forecasting approaches [2]. From the spatial perspective, locations that are close geographically in the Euclidean sense (e.g., two locations located in opposite directions of the same highway) often do not exhibit a similar traffic pattern, whereas locations in the highway network that are relatively far apart (e.g., two locations separated by a mile in the same direction of the same highway) can show strong correlations. Many traditional predictive modeling approaches cannot handle these types of correlations. From the temporal perspective, because traffic conditions vary across different locations (e.g., diverse peak hour patterns, varying traffic flow and volume, highway capacity, incidents, and interdependencies), the time series data becomes nonlinear and nonstationary, rendering many statistical time series modeling approaches ineffective.

As a research hotspot, highway traffic flow prediction has achieved fruitful results in different development periods, and the research trend has shifted from a single parametric model to a nonparametric model and a mixed model. Typical parametric models include autoregressive integrated moving average (ARIMA) model and the Kalman filter model. Levin and Tsao applied Box–Jenkins time series analyses to predict expressway traffic flow and found that the ARIMA (0, 1, 1) model was the most statistically significant for all forecasting [3]. Hamed et al. applied an ARIMA model for traffic volume prediction in urban arterial roads [4]. Many variants of ARIMA were proposed to improve prediction accuracies, such as Kohonen-ARIMA (KARIMA) [5], subset ARIMA [6], ARIMA with explanatory variables (ARIMAX) [7], vector autoregressive moving average (ARMA) and space-time ARIMA [8], and seasonal ARIMA (SARIMA) [9]. Compared with nonparametric models, parametric models rely on the assumption of stationarity and cannot reflect the nonlinear and uncertain characteristics of traffic flow. Therefore, nonparametric models become an effective method for traffic flow prediction. Nonparametric models include support vector machines (SVMs), artificial neural networks (ANNs), deep learning models, etc. However, SVMs are computationally expensive for large networks, and ANNs cannot capture the spatial dependencies of the traffic network. Furthermore, the shallow architecture of ANNs makes the network less efficient compared with a deep learning architecture. Moreover, deep learning models can be more accurately express the complex structure of traffic flow data. For example, Zhao et al. [10] applied long-short term memory (LSTM) neural network to traffic flow prediction to improve the prediction accuracy of the model. CNN-LSTM [11] is a neural network-based architecture for multi-lane short-term traffic forecasts. It highlights the importance of applying multiple features to characterize traffic conditions, explicitly considering the route between neighboring lanes and downstream/upstream traffics, and predicting multiple timestep traffic in a rolling-prediction manner.

Although many previous studies clearly predict highway traffic flow, there are three issues remaining unsolved in existing methods. First, parametric models rely on stationarity assumptions and cannot reflect the nonlinear and uncertain characteristics of traffic flow. Second, SVMs are too computationally expensive for large networks, and ANNs cannot capture the spatial dependencies of traffic networks. Third, the shallow architecture of ANN makes the network less efficient compared to deep learning architectures.

3. Methodology

In this section, we introduce a method for predicting traffic flow based on data acquired from highway gantry using a long short-term memory network (LSTM), the architecture of the short-term traffic flow prediction algorithm based on the LSTM model is shown in Figure 1, which includes the following steps:

Step 1: The highway gantry separately collects the traffic flow data sets on adjacent road sections, uploads the data sets to the Traffic Management Center (TMC) uniformly, processes the historical traffic flow data, eliminates the interference data with excessive deviation, and then processes the data. The time series is analyzed and divided into a test set and training set, and \((x_1, x_2, \ldots, x_t)\) is obtained as the input of the model.
Step 2: Use the LSTM model to predict the short-term traffic flow for the segment.

Step 3: During the training phase, the function is continuously adjusted based on the predicted performance.

Step 4: Output traffic flow prediction, return to Step 1.

When learning data with a very large time interval, the Recurrent Neural Network (RNN) will forget the information that is older than the current data interval, so the RNN will fail for data with a very large time interval. The disadvantage of RNN is that it fully accepts the information transmitted by the network layer, and there is no corresponding selection process, which causes it to accept a lot of unnecessary information and reduces its learning ability. To solve this problem, the long short-term memory network (LSTM) is proposed. LSTM is a special form of RNN that has the ability to deal with long-distance problems. It can use its gating unit to screen information and selectively forget previous information, which can effectively solve problems such as gradient disappearance and gradient explosion. The traffic flow data of road sections 24 hours a day has very obvious time series characteristics, and LSTM can carry out long-term memory of historical information and analyze the characteristics of long-term series data through its own algorithm setting time interval, effectively mining its inherent regularity, so LSTM is very suitable as a tool for processing traffic flow prediction.

The basic unit structure of LSTM is shown in Figure 2. It is the same as RNN. The input layer of the model will receive the input $x_t$ for each operation. After processing the output $h_t$, the model will also add the current state $C_t$ of the input unit according to the training and parameter updates, the current output state $C_t$, and consider the output state $C_{t-1}$ at the last moment.

The basic structure of the LSTM hidden layer is a storage cell, which has a linear structure in the village and is not affected by other cells. Activation of cells to obtain the state of existing cells is a key part of LSTM. LSTM adds and deletes the error information mentioned above through the gating unit, and has completed the screening of data information in the process of transmission.

Through the abovementioned structural system, an output value between 0 and 1 can be obtained, and a selection of input data can be completed by compressing the data into a value between 0 and 1. If it is 0 then no information can pass through the structure, if it is 1 then pass all information. To sum up, a basic LSTM unit can be summarized into three gating units: input gate, forget gate, and output gate. Among them, the forget gate can effectively analyze the data, filter out the key information and eliminate the redundant information to help LSTM complete the sequence processing problem. For the $n$th iteration, the abovementioned three gates can be represented by $f_t$, $i_t$, and $o_t$, respectively.

LSTM determines which information needs to be remembered and which information should be eliminated through the forget gate. The forget gate can learn the input $x_t$ received by the current unit and the output $h_{t-1}$ of the hidden layer output of the previous unit as the input of this unit.

The next step for the cell is to decide how much new information to admit to adding to the current cell state. On
(4) The combined effect of the forget gate and the input gate solves the current network state value. In formula (4), the current memory state value and the state value of the previous moment are used to update the current state of the network.

\[ c_t = f_t c_{t-1} + i_t \bar{c}_t. \]  

(5) Solve the value of the output gate of the system. Formula (5) represents the weight between the input data and the output gate, represents the weight between the hidden layer and the output gate, represents the weight connecting the current memory state and the output gate, and represents the output gate bias parameter.

\[ o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + W_{co}c_{t-1} + b_o). \]  

(6) Calculate the output value of the network at time \( t \) by activating the current state value through the activation function by formula (6).

\[ h_t = \sigma_c \tan h(c_t). \]  

By comparing the error between the predicted value and the actual value, the system backpropagation to update the weights to iterate the network. Finally, when the prediction error obtained by the model after training meets certain requirements, the predicted value is output, and the predicted value of traffic flow is obtained by inverse normalization. The model prediction performance is judged. Mean square error, root mean square error, and mean absolute error are shown by formulas as follows:

\[ \text{MSE} = \frac{1}{n} \sum_{i=1}^{n} (x_{\text{true}} - x_{\text{predicted}})^2, \]

\[ \text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_{\text{true}} - x_{\text{predicted}})^2}, \]

\[ \text{MAE} = \frac{1}{n} \sum_{i=1}^{n} |x_{\text{true}} - x_{\text{predicted}}|. \]
4. Experiment

4.1. Data Description

4.1.1. Data Source and Processing. Based on the highway gantry data, this paper takes the k602 + 630-k636 + 090 section of Chang-Jiu Expressway as the research object, with a total length of 40.89 km and three gantry detection devices. The regional situation is shown in Figure 3:

Adjacent gantry records the entry and exit times of a single vehicle. For example, the time when the vehicle enters the road section is \( t_p \), and the time when the vehicle exits is \( t_{p+1} \), then the travel time \( t = t_{p+1} - t_p \), divided by the travel distance to get the vehicle speed. It is worth mentioning that as two adjacent highway gantries can get how many vehicles enter and exit the section, the vehicles going out from the downstream gantries minus the vehicles coming in from the upstream gantries is the number of vehicles in the section, the distance among the gantries divided by the number of vehicles can get the density, through traffic volume = traffic density * travel speed can get the traffic volume, and the traffic volume can reflect how many vehicles have passed a certain point.

The quality of road traffic data should be guaranteed, and the problem data will have an impact on the analysis of the running state and the construction of the model, so scientific methods should be used to eliminate abnormal data. The generation of problem data mainly comes from the detector and the vehicle. For the vehicle, due to the complex road conditions, some vehicles may break down and stay on the road for too long, and some may have excessive driving. Behaviors lead to abnormal speeds, and there are cases where traffic collisions lead to abnormal data. The problem considered in this paper is the analysis of the macroscopic operating state of the road network. A small number of data generated by special traffic accidents and vehicle failures may cause errors in the accuracy of the model, so this type of data is treated as abnormal data. In addition, excessive driving behavior should be regarded as the driving data under abnormal conditions, which is mainly manifested in that the staying time in the road section is too long or too short, which seriously deviates from the normal range, is not representative, and cannot reflect the real operation of the current road network status, it should also be eliminated.

Common data cleaning methods include the quartile method and statistical principle, both of which belong to symmetry cleaning. The operation rules of quartile method and statistical principle are as follows:

1. **Quartile Method.** Calculate the 25% and 75% quantiles of the data, and use the upper and lower values of these two quantiles as the upper and lower limits of the interval. Data beyond this range are regarded as abnormal data. Mathematical expressions are shown in formulas (8) and (9).

   \[
   t_{\text{limit--down}} = t_{25\%} - 1.5 \times (t_{75\%} - t_{25\%}), \tag{8}
   \]

   \[
   t_{\text{limit--up}} = t_{25\%} + 1.5 \times (t_{75\%} - t_{25\%}). \tag{9}
   \]

In the formula, \( t_{\text{limit--down}} \) is the lower limit of the valid data interval, \( t_{\text{limit--up}} \) is the upper limit of the valid data interval, \( t_{25\%} \) and \( t_{75\%} \) represent the sample size, respectively, 25th and 75th percentiles.

2. **Statistical Principal Method.** According to the 2\( \sigma \) principles in statistics, 95.4% (the vast majority) of the samples are included in the following interval, as shown in equations (10) and (11).

   \[
   t_{\text{limit--down}} = t_{25\%} - 1.5 \times (t_{75\%} - t_{25\%}), \tag{10}
   \]

   \[
   t_{\text{limit--up}} = t_{\text{mean}} + 2\sigma. \tag{11}
   \]

   In the formula, \( t_{\text{mean}} \) is the sample mean, and \( \sigma \) is the sample variance.

4.1.2. Data Analysis

1. **Spatial Similarity.** Since the road network is an overall structure, the internal structure is complex and the influences are transmitted dynamically, so the traffic flow on the road will be affected by the traffic flow of the adjacent road sections, showing certain characteristics. The traffic flow states of several adjacent road segments at a certain time will transfer to each other to produce similar characteristics. Taking the high-speed gantry data in this paper as an example, select the adjacent road sections upstream and downstream of k602 + 630-k636 + 090, name them as road Sections 1, 2, and 3, and calculate the traffic flow and vehicle average of the three adjacent road sections in one day. The velocity diagram, as shown in Figures 4 and 5.

   It can be seen from the figure that the traffic flow and the average speed of the three road sections show roughly the same trend of change. The traffic flow is almost the same. The average speed of Section 1 and Section 2 is roughly the same, while Section 3 is quite different from the first two. Considering the possible cause is external factors such as terrain or structures. According to the change law of traffic flow and speed of each road section, further, explore its internal relationship and similarity, and introduce the Pearson correlation coefficient to analyze the similarity of traffic flow in time series. Its mathematical expression is shown in (8):

   \[
   \rho_{X,Y} = \frac{\text{Cov}(X,Y)}{\sigma_X \sigma_Y}. \tag{12}
   \]

   Among them: \( X, Y \) are two random quantities, \( \text{cov}(X,Y) \) is the sample covariance, and \( \sigma_X, \sigma_Y \) is the two-sample standard deviation.

   The value of \( \rho_{X,Y} \) is \([-1, 1]\). When it is 1, it proves that the two random samples have a complete positive correlation. When it is -1, it proves that the two random samples have a complete negative correlation, when it is 0, it proves that there is no correlation between the two samples. When
\( \rho_{XY} > 0.8 \), it proves that the two samples are highly correlated.

From this, the similar results of the traffic flow and the average driving speed of the three road sections are calculated, respectively, and the respective correlation coefficient tables are drawn as shown in Tables 1 and 2.

It can be seen from Table 1 that the traffic flow of the three sections is highly correlated, and the traffic flow of the three sections basically presents a saddle shape, that is, there are two peak periods in the morning and in the evening, and the growth trend of the morning peak is more obvious. Section 2 is at 8: The fluctuation of traffic flow from 00:00 to 10:00:00 is larger than that of Section 1 and Section 3. However, the correlation between Section 3 and Section 2 during this period is relatively low. The possible reason for consideration is that the traffic flow of Section 2 is concentrated, there may be some traffic accidents or excessive driving behaviors that cause the vehicle to stay in the road section, while the traffic flow of road Sections 1 and 3 are relatively stable, so they show high similarity.

It can be seen from Table 2 that the correlation of the average vehicle speed of the three sections is low, and the
The average speed of Section 3 is significantly lower than that of Section 1 and Section 2. Considering that the speed is reduced because the traffic flow is generally concentrated in downstream, another observation is made. The deviation of the driving speed to the road in Section 2 is large, and there are some high or low speeds, so the driving behavior of the vehicle on the road in Section 2 is more complicated. As the key research object. To sum up, the traffic flow in space has both a certain similarity and a certain influence and fluctuation, which provides a certain rule for the subsequent analysis of the traffic operation state.

(2) Space Stationarity. For the traffic flow on the road section, the traffic parameters will have a certain fluctuation in a certain time but will maintain certain stability with the expected value, that is, there will be individual fluctuations. The spatial stability of traffic flow will affect the traffic operation state of the road section. Different roads will generate traffic flow in different states due to differences in driving conditions and spatial structures. For the three road sections selected abovementioned, analyze the traffic flow and the average vehicle speed, respectively, and draw a box diagram to check their distribution characteristics and degree of dispersion, as shown in Figures 6 and 7.

It can be seen from the results that there is a certain difference between the traffic flow and the average vehicle speed of different road sections. This is because the geographical locations of different road sections are different, and they are affected by the upstream and downstream traffic flow to different degrees. The speed fluctuation is obvious, the vehicle speed is high all day, there are many outliers, and the spatial fluctuation of the traffic flow is more obvious.

4.2. Result Analysis. The LSTM algorithm used in this article is based on the python3.8 environment and uses the Keras deep learning library to build a neural network structure. Through the actual situation and parameter adjustment, it is determined that the actual number of layers of the model is 2, and the number of nodes is 80 and 80, respectively; used in the model Adam is used as the optimization function and MSE is used as the objective function. It should be noted that we split the sample data into two parts, the traffic flow history data of the first two road sections as the training set, and the traffic flow data of the last road section as the test set. We use the data from the training set to iteratively train the model, and later, to verify the final effect of the model, we calculate the error of the trained model on the test set so that the error of the trained model on the test set is minimized to obtain a model that fits the data set reasonably well. For predicting the traffic flow on other links/locations, it should collect consecutive link traffic flow samples in time interval $t_0-t_{-1}$, which in turn, is used as input of LSTM model to predict the traffic flow of $t$ slot on the designated link.

It can be seen from the previous analysis that the nature of traffic flow changes with time, and the traffic flow collected in different time intervals also reflects different laws. Using LSTM to analyze the advantages of time series data, dividing the time interval to collect traffic flow and substitute it into the model for prediction will have different evolution laws and accuracy. In this experiment, in order to determine the influence of different time intervals on the model prediction, the historical traffic flow data were divided into average data with 1, 3, 5, and 10 min intervals for prediction.

Visualize the experimental results to obtain the prediction result graphs at 4-time scales, as shown in Figures 8–11.

The prediction performance of the four-time intervals is evaluated by indicators and organized into a table, as shown in Table 3.

From Table 3, the performance results of the model's prediction performance on traffic flow under different prediction time scales can be obtained. When the scale increases, the values of the three indicators all increase, indicating that the prediction performance of the model is declining, and the model has a long statistical interval. The prediction performance of the data continues to decline, indicating that the prediction time scale affects the prediction performance of the model. Among them, the indicators obtained by dividing the scale by 1 min are the smallest, and with the increase of the prediction time scale, the errors also increase immediately, and the prediction performance of the model for traffic flow is also weaker, which proves that the traffic flow of the road section is suitable. The prediction scale of 1 min is better. The possible reason is that the road section is unblocked, and the vehicles with high-speed are concentrated, resulting in a short time for the vehicle to drive through a single road segment, and the vehicle has already driven through three road segments at the interval of the larger prediction scale, so the effect is poor.
Figure 8: Traffic flow prediction results with 1 min as the statistical scale.

Figure 9: Traffic flow prediction results with 3 min as the statistical scale.

Figure 10: Traffic flow prediction results with 5 min as the statistical scale.
5. Conclusion

Accurately predicting expressway traffic flow can not only provide decision-making assistance for expressway managers but also provide a reference for the selection of public travel routes, facilitate vehicle diversion, and thus alleviate expressway traffic congestion. At the same time, short-term traffic forecasting has a positive impact on reducing accident rates and providing a convenient and comfortable travel environment for the public. In order to predict highway traffic flow more accurately, this paper considers the spatiotemporal characteristics of the road network, uses the advantages of long short-term memory network (LSTM) to analyze time series data, divides time intervals to collect traffic flow, and substitutes it into the model to predict traffic flow. The model is verified by the data of section k602+630-k636+090 of Chang-Jiu Expressway. Firstly, the spatial similarity and spatial stationarity of the obtained traffic flow data are analyzed.

And the historical traffic flow data of the first two sections are used as the training set, and the traffic flow data of the last section is used as the test set. LSTM is used to predict the traffic flow of the section, using mean square error (MSE), root mean square error (RMSE). The three evaluation indicators of mean absolute error (MAE) divide the historical traffic flow data into average data with intervals of 1, 3, 5, and 10 minutes, respectively, for prediction. The suitable prediction scale of road traffic flow is 1 min.

Data Availability

The data applied in this paper are privately restored on a private server by the local transport management department, which is not open to the public without their authorization.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This research was supported by Jiangxi Provincial Major Science and Technology Project, 5G Research Project (Grant no. 20212ABC03A07).

References

