Research on Short-Term Traffic Flow Forecast Based on Improved Cuckoo Search Algorithm

Ruoxi Wang[1(✉)], Jingjing Zheng[1], and Zaoning Wang[2]

[1] Wuhan Fiberhome Technical Services Co., Ltd., Wuhan 430205, China
{rxwang,jjzheng}@fiberhome.com
[2] School of Computer Science, Hubei University of Technology, Wuhan 430205, China
101910734@hbut.edu.cn

Abstract. In recent years, the traffic flow in big cities has increased significantly, and the development of intelligent transportation systems (ITS) has become the general trend. Most of the traditional traffic flow prediction models are highly dependent on experienced experts and lack the ability to learn independently. Since the traffic flow depends on various factors such as weather, road conditions, and whether there are major events, it's influenced by multi-factors and huge amounts of data. It is difficult to fit and process traffic flow data well, so traditional traffic forecasting models are no longer suitable for the analysis and prediction of current and even next-generation traffic flow. Now artificial neural networks are widely used in traffic flow forecasting due to their strong robustness and fault tolerance, high operating efficiency, ability to process massive data, strong nonlinear mapping capabilities, and strong learning and adaptive capabilities. To tackle the problem of random initial weights and thresholds of the traditional model for the neural network, we proposes an improved cuckoo search algorithm, and uses the improved algorithm to optimize the initial threshold in the long short-term memory neural network, and applies the improved model to the research on traffic flow prediction, in order to improve traffic planning and save people's travel time and fuel costs laid the groundwork.

Keywords: Traffic Flow Prediction · CS Search Algorithm · Weighted Traction Component · Long Short-Term Memory Neural Network

1 Introduction

In recent years, people's economic conditions improved significantly, and speed up the pace of urban life, more and more people choose to drive or take a taxi, thus significantly increased traffic in big cities. According to the statistics of relevant departments, by 2021, the national motor vehicle population is up to 395 million. While private cars bring convenience to people's travel, urban traffic problems are becoming more and more serious. Therefore, the construction of smart city has attracted more and more attention [1].

Since 1980s, there have been researches on short-time traffic flow prediction technology for Intelligent Transportation System (ITS) [2]. Up to now, researchers have put

forward about dozens of different short-time traffic flow prediction methods. The famous short-time traffic flow prediction methods can be divided into non-parametric technology and parametric technology. Among non-parametric techniques, the well-known Neural Network methods include Back Propagation (BP) Neural Network [3], Generalized Regression Neural Network, GRNN) [4] and Wavelet Neural Network (WNN) [5]. Technology with parameters includes linear regression, nonlinear regression, historical average algorithm, smoothing technology, etc. According to the prediction method, it can be divided into traditional traffic flow prediction and traffic flow prediction based on deep learning.

In the early stage, scholars established traffic flow prediction models by using mathematical thinking or physical thinking and their combination, including ARMA [6], ARIMA [7] and HAM [8]. Shekhar S and Williams BM et al. proposed the application of Kalman filter model to short-term traffic flow prediction for the first time in the 19th century [9], and traffic flow prediction made a big step forward. This model can deal with stable and fluctuating data, but it cannot deal with randomness and nonlinear situations, so it has certain limitations in practical production and life.

Huang W et al. first proposed the introduction of Deep learning technology into traffic flow prediction in 2014 [10]. In 2015, Koesdwiady et al. used DBN to predict the traffic data set PeMS [11], which was a major breakthrough in the field of traffic flow prediction. In 2016, Ma et al. used long short-term memory neural network (LSTM) and microwave detector data to predict Beijing traffic speed [12]. LSTM was used to overcome error attenuation caused by back propagation through memory blocks, and showed superior performance in long-term dependent time series prediction. In a 2018 research paper, Zhuo et al. proposed a new traffic prediction model based on LSTM [13] and verified that the model had better performance.

In this paper, by comparing various intelligent optimization algorithms and combining the characteristics of each optimization algorithm, the improved cuckoo optimization algorithm is selected to optimize the random initialization parameters of the traditional neural network. Finally, the optimized network model is applied to short-term traffic flow prediction, and the original data set is preprocessed accordingly. The improved cuckoo search algorithm is compared with the standard cuckoo search algorithm to judge the optimization performance of each algorithm to the benchmark test function. The main research contents of this paper are as follows:

(1) Aiming at the problem that the step size factor of the standard cuckoo optimization algorithm is artificially set, which leads to low optimization efficiency, this paper proposes an adaptive step size factor based on logarithm function to improve it. In order to improve the accuracy of the algorithm, gaussian perturbation is used to guide the local search of bird's nest position. The weight traction component is also proposed to increase the relationship between individuals in the algorithm and strengthen population diversity.

(2) Explore the methods of data preprocessing, analyze in detail the methods of outlier processing, data dimension reduction, feature extraction and other preprocessing, and finally preprocess PeMS04 data set through the above methods.

(3) In view of the defect that LSTM model parameters need to be set manually, this paper uses the improved CS algorithm to optimize LSTM by taking advantage of cuckoo

optimization algorithm's strong global search characteristics, and then applies the optimized LSTM model to short-term traffic flow prediction. The original LSTM model and other common time series prediction models are compared to predict the same data set, to verify the prediction accuracy of the optimized LSTM.

2   Short-Term Traffic Flow Prediction Based on AWCS-LSTM Model

This chapter mainly introduces the optimization of traditional LSTM hyper-parameters based on improved CS algorithm, and applies the optimized model to traffic flow prediction. First introduced the circulation principle of neural network and its derivative network LSTM, then using improved CS algorithm (AWCS) to optimize the LSTM model, the number of iterations and the number of hidden layer neuron network, and train the prediction model. Finally, validate the improved model in PeMS04 data set; compared with the model accuracy and other evaluation indexes before and after the improvement, the experimental verification shows that the improved model has more advantages in traffic flow prediction.

2.1   Construction of Neural Network Model Based on AWCS-LSTM

Principles and Steps of LSTM Optimization by Cuckoo Optimization Algorithm
Artificial neural network model with animal neural network is simulated for inspiration, make it has the ability of learning and memory, including LSTM improved the long-term dependence of RNN defects, therefore has been widely applied in the field, the time-series data of forecasting and prediction accuracy of the model is closely related to neural network initialization parameter setting, super parameter selection directly influences the performance of the model. In general, the initial weights and thresholds of neural networks in traditional models are set manually, so they are random and not conducive to model training. Therefore, cuckoo search algorithm is introduced and improved to optimize the initial threshold in LSTM.

The flow chart of traditional LSTM model for data prediction is shown in Fig. 1:

Cuckoo optimization algorithm in the search for the optimal parameters showed better performance, thus introducing the LSTM, optimize the traditional LSTM model, the main idea is to the LSTM vector network, the number of iterations and the number of nodes in the hidden layer as a cuckoo search algorithm of the bird's nest, to verify that the collection of mean square error (MSE) as the fitness value, The optimization principle of CS algorithm is used to find the optimal initial value of the model, so as to improve the fitting accuracy of LSTM model in the direction of traffic flow prediction and obtain the optimal LSTM model.

When optimizing the model, the fitness value is calculated by Eq. (1):

$$\text{fitness} = \text{MSE} = \frac{1}{N}\sum_{i=1}^{N}(f_i - y_i)^2 \tag{1}$$

f_i is the predicted value, y_i is the actual value, N is the number of samples, MSE is the mean square error, the greater the variance, the greater the dispersion, that is, the greater the error.
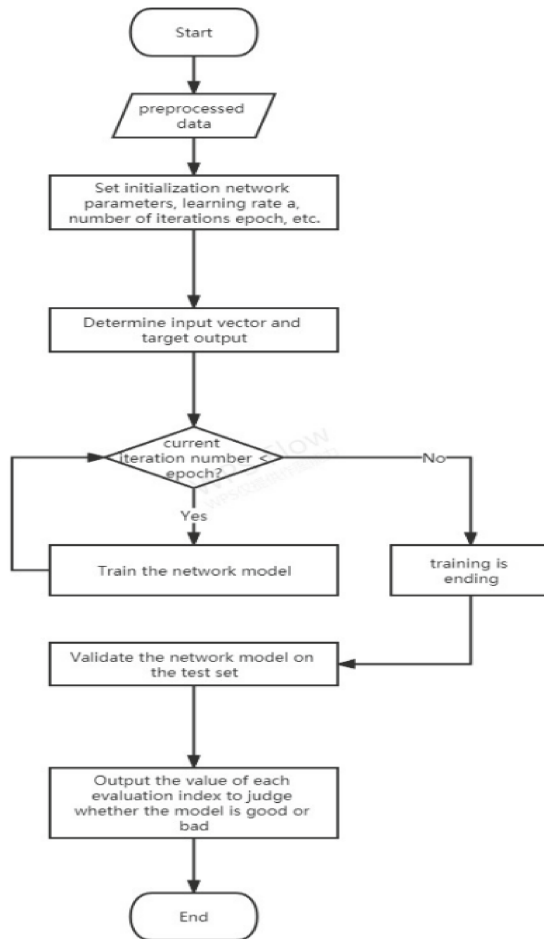
Fig. 1. LSTM prediction flow chart

The steps of the improved cuckoo search algorithm to improve LSTM and predict traffic flow data are as follows:

Input: preprocessed PeMS04 training set and test set.

Output: the value of traffic flow at a certain time.

Step 1: Parameter setting and group initialization: Randomly generate N initial nests as required and determine their positions. Set the initialization parameters: population size N, number of parameters to be searched corresponding to dimension D, discovery probability Pa, boundary values a and B of each optimization parameter, maximum iteration times itermax and current iteration times i, etc. In this paper, dimension $D = 4$ is set, optimization parameters are learning rate α, network iteration times epoch, hidden layer node number N1 and n2.

Step 2: The nests generated by the cuckoo search algorithm are decoded as parameters of each layer of the neural network, and the number of parameters corresponds to the

dimension D of the nest. The number of parameters is substituted into the improved AWCS algorithm process to obtain the optimal nest of the current iteration number, namely, the parameter value of the neural network.

Step 3: Take parameter values as LSTM initialization parameters to predict traffic flow data;

Step 4: Substitute the predicted value into formula (4–1) to calculate the fitness value;

Step 5: Judge whether the fitness value meets the threshold requirements or reaches the maximum number of iterations. If I $\geq$ itermax, stop the algorithm; Otherwise, update the nest position and get the parameter value of the next round;

Step 6: The updated optimal nest position of each dimension corresponds to the initial value of LSTM network one by one, that is, a new round of training parameters are obtained and the network is trained until the predicted results are output.

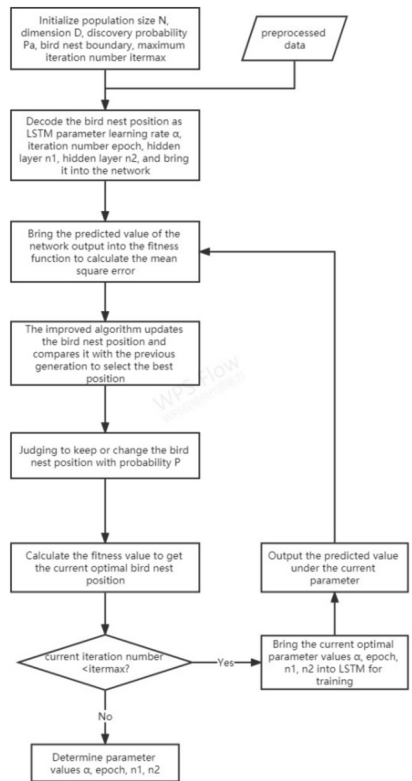AWCS algorithm optimization LSTM traffic flow prediction process is shown in Fig. 2:

Fig. 2. Improved CS optimized LSTM

Model parameter setting.

Parameter Settings of the above training process are shown in Table 1:

Table 1. AGCS/AWCS-LSTM parameter Settings

	Parameter	Range
AWCS	Population size N	20
	Dimension D	4
	Discovery probability Pa	0.25
	maximum iteration number itermax	25
LSTM	batch size	128
	learning rate α	[0.001,0.01]
	epoch	[10,100]
	The number of hidden layers of the network	2
	layer n1	[1,100]
	layer n2	[1,100]

The activation function of LSTM forgetting gate and output gate selects sigmoid function, whose value range is [0,1]. 0 means discarding information, and 1 means reserving all information, as shown in Formula (2):

$$\sigma(z) = \frac{1}{1 + e^{-z}} \tag{2}$$

Tanh function is selected to activate the input gate, and the value range is $[-1,1]$, as shown in Formula (3):

$$tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{3}$$

2.2 Experimental Results and Analysis

Evaluation Indicators
After the prediction results are obtained, the pros and cons of the model and the accuracy of the prediction can be determined by evaluating the generalization performance of the prediction model. Common evaluation indexes are mean absolute error (MAE), root mean square error (RMSE), Average absolute percentage error (MAPE) and determinant coefficient (R2).

Comparison of Models Before and After Optimization
In order to verify the practical application of AWCS-LSTM model in short-term traffic flow prediction, this paper selects the public traffic flow data set PeMS04 for verification, and the data set only contains the traffic flow of working days in the selected period.

The training set contains 8778 records as samples from 2016.01.01 to 2016.02.29, and the verification set contains 2033 records as samples from 2016.03.01 to 2016.03.16.

The test set selected 2480 records as test samples recorded in six days from 2016.03.17 to 2016.03.31 to verify the accuracy of the model.

In this paper, the traditional LSTM model was firstly verified, and the model training iteration times epoch was set as 10, 20, 30, 40, 50, 60, 70 and 80 respectively for comparative experiments. The evaluation index values predicted by LSTM for PeMS04 data set under different epochs are shown in Table 2:

Table 2. LSTM prediction effects under different epochs

| Nos. of Network training	MAE	RMSE	MAPE	R2
epoch = 10	12.4145	15.9587	0.4438	0.8441
epoch = 20	10.2707	13.3600	0.3329	0.8907
epoch = 30	8.1528	10.5815	0.2226	0.9315
epoch = 40	7.4542	9.7490	0.1688	0.9418
epoch = 50	7.3278	9.5706	0.1601	0.9439
epoch = 60	7.2345	9.4787	0.1599	0.9450
epoch = 70	7.2062	9.4711	0.1632	0.9451
epoch = 80	7.73120	9.5824	0.1603	0.9437

It can be seen from the table above, with the increase of the number of iterations, all the indicators tend to be better, but when the number of iterations is set to 80, when the back phenomenon, predicted results as the number of iterations instead of 70, so when the number of iterations is too large, not only the extension of training time, can also cause a trained fitting, resulting in a loss prediction effect. In order to display the prediction results more intuitively, the prediction curves under different iterations are visualized. The prediction results are shown in Fig. 3:

The improved CS algorithm is used to optimize the hyper-parameters of the traditional LSTM model. The AWCS with improved adaptive step size and weight traction component are introduced into LSTM respectively. The CS-LSTM model optimized by traditional LSTM and standard CS and the improved AWCS-LSTM model are used to predict the obtained evaluation indexes on the selected data set, as shown in Table 3:

It can be seen from Table 3 that the model built by using the improved CS algorithm to optimize the hyper-parameters of the traditional LSTM network has better accuracy in predicting PeMS04. Compared with the model optimized by the traditional LSTM model and the standard CS algorithm, AWCS has better performance in optimizing the hyper-parameters of LSTM. The visualization results of prediction are shown in Fig. 4:

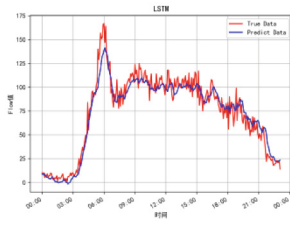The parameters of traditional LSTM, CS-LSTM and AWCS-LSTM are shown in Table 4:

The learning rate, number of hidden layer neurons and iteration times of LSTM model are set manually, while the learning rate, number of hidden layer neurons and iteration times of CS-LSTM and AWCS-LSTM are optimized by corresponding algorithms.
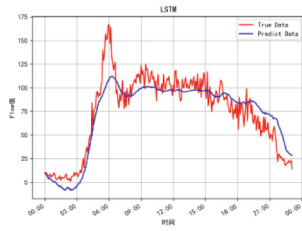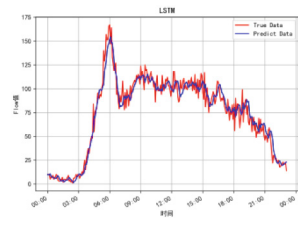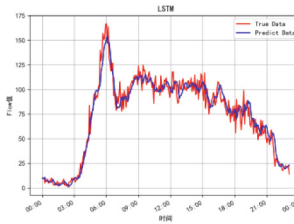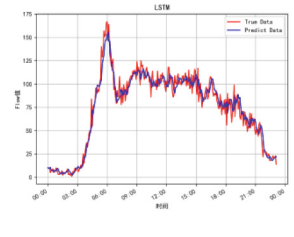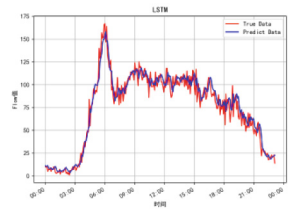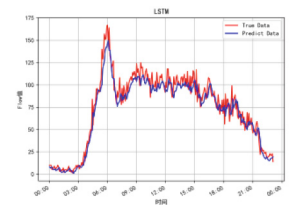
epoch=10
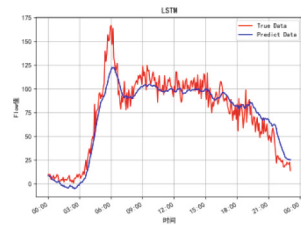
epoch=20

epoch=30

epoch=40

epoch=50

epoch=60

epoch=70

epoch=80

Fig. 3. LSTM prediction effect under different epochs

Table 3. Prediction effects of LSTM, CS-LSTM and AWCS-LSTM

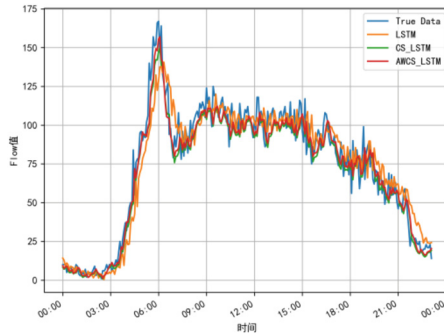| Network Model	MAE	RMSE	MAPE	R2
LSTM	9.5913	12.9674	0.2043	0.8970
CS-LSTM	6.9455	9.1766	0.1706	0.9353
AWCS-LSTM	6.9389	9.1588	0.1615	0.9516

Fig. 4. Visualization of LSTM, CS-LSTM and AWCS-LSTM prediction

Table 4. Parameter Settings of each model

Model	LSTM	CS-LSTM	AWCS-LSTM
batch size	256	256	256
Learning rate	0.01	0.0036	0.0089
Number of neurons in the first hidden layer	20	57	73
Number of neurons in the second hidden layer	10	8	18
Dropout rate	0.2	0.2	0.2
iterations	70	72	75

3 Conclusion

Aiming at the problems of traditional traffic flow forecasting method, this paper proposes using the intelligent optimization algorithm of neural network parameters optimization of ideas, and set up a neural network based on improved cuckoo search algorithm and short - and long-term memory of traffic flow prediction model, using the improved algorithm used to optimize the LSTM AWCS random initial parameters of the model, LSTM, CS-LSTM and AWCS-LSTM were applied to the preprocessed PeMS04 data set respectively, and the effectiveness of the optimized model was verified by experiments.

References

1. Liu, Y.: Big data technology and its analysis of application in urban intelligent transportation system. In: 3rd International Conference on Intelligent Transportation, Big Data and Smart City, pp. 17–19 (2018)
2. Liao, H.: Intelligent transportation decision analysis system based on big data mining. In: International Conference on Computer Information Science and Application Technology - Big Data Search, Mining, and Visualization, vol. 1168, no. 3, pp. 1–5 (2019)
3. Mei, L.: Research on software reliability model based on improved BP neural network. In: 2018 International Conference on Smart Grid and Electrical Automation (ICSGEA), pp. 221–2232018https://doi.org/10.1109/ICSGEA.2018.00061
4. Kartal, S., Oral, M., Ozyildirim, B.M.: Pattern layer reduction for a generalized regression neural network by using a self-organizing map. Int. J. Appl. Math. Comput. Sci. **28**(2), 411–424 (2018)
5. Liu, Z.: Research on face recognition based on the fusion of convolution and wavelet neural network. In: Proceedings of the 4th International Conference on Virtual Reality (ICVR 2018), New York, NY, USA, pp. 122–125. Association for Computing Machinery (2018)
6. Mikolov, T., Deoras, A., Povey, D., et al.: Strategies for training large scale neural network language models. In: Automatic Speech Recognition and Understanding, pp. 196–201. IEEE (2012)
7. Ahmed, M.S., Cook, A.R.: Analysis of freeway traffic time-series data by using Box-Jenkins techniques (1999)
8. Lee, S., Fambro, D.: Application of subset autoregressive integrated moving average model for short-term freeway traffic volume forecasting. Transp. Res. Rec. J. Transp. Res. Board **1678**(1), 179–188 (1999)
9. Shekhar, S., Williams, B.M.: Adaptive seasonal time series models for forecasting short-term traffic flow. Transp. Res. Rec. **2024**(1), 116–125 (2007)
10. Huang, W., Song, G., Hong, H., et al.: Deep architecture for traffic flow prediction: deep belief networks with multitask learning. IEEE Trans. Intell. Transp. Syst. **15**(5), 2191–2201 (2014)
11. Koesdwiady, A., Soua, R., Karray, F.: Improving traffic flow prediction with weather information in connected cars: a deep learning approach. IEEE Trans. Veh. Technol. **65**(12), 9508–9517 (2016)
12. Ma, X., Tao, Z., Wang, Y., et al.: Long short-term memory neural network for traffic speed prediction using remote microwave sensor data. Transp. Res. Part C Emerging Technol. **54**, 187–197 (2015)
13. Zhuo, Q., Li, Q., Yan, H., Qi, Y.: Long short-term memory neural network for network traffic prediction. In: 2018 12th International Conference on Intelligent Systems and Knowledge Engineering (ISKE), pp. 1–6 (2018). https://doi.org/10.1109/ISKE.2017.8258815