ISSN 2522-3488 Vol. 6, Issue 3: 55-62, DOI: 10.25236/IJNDES.2022.060310

Short-Term Traffic Flow Prediction Method Based on WT-IGWO-ELM

Shuilin Li[1], Luyao Niu[2]

[1]International Institute of Technology, Changsha University of Science & Technology, Changsha, 410114, China
[2]School of Traffic and Transportation, Beijing Jiaotong University, Beijing, 100044, China

Abstract: *Accurate prediction of short-term traffic flow provides crucial data support for the stable operation of intelligent transportation systems. For this issue, this paper proposes a short-term traffic flow prediction method based on WT-IGWO-ELM. The algorithm uses the Wavelet Transform (WT) method to denoise the traffic flow data in advance, which improves the data quality of the dataset. Then, the IGWO algorithm, which integrates the initial population based on Sine chaotic map and reverse learning strategy, the adjustment of nonlinear convergence factor and the introduction of dynamic weights, is used to avoid local optimality more effectively, speed up the convergence speed, and improve the solution accuracy. Finally, the improved grey wolf optimizer (IGWO) was used to update the optimal parameters of the ELM prediction model, and the average relative error of the prediction of the WT-IGWO-ELM model was verified by comparison experiments compared with those of ELM and WT. -ELM, GWO-ELM, WT-GWO-ELM and IGWO-ELM decreased by 96.6625%, 95.5972%, 87.9447%, 79.5021%, 72.0571%, respectively, and its prediction effect was much better than ELM, WT-ELM, GWO-ELM, WT-GWO-ELM and IGWO-ELM methods have high prediction performance and accuracy in short-term traffic flow prediction.*

Keywords: *Short-term traffic flow prediction; Wavelet Transform; Improved Grey Wolf Optimizer; Extreme Learning Machine*

1. Introduction

In recent years, traffic congestion has become an important factor hindering the development of cities, and has plagued the travel of residents. Traffic flow prediction is an important step in the guidance and control of intelligent traffic management, and the study of traffic flow prediction is of great significance to improve road service and management level [1]. Therefore, quickly obtaining the short-term traffic flow prediction results and improving the short-term traffic flow prediction accuracy are the guarantees for the effective operation of the traffic management system.

There have been many studies on the prediction of short-term traffic flow at home and abroad. The commonly used prediction models can be divided into three categories: statistical theoretical models, such as historical average models, time series models, Kalman filter models, etc.; artificial intelligence models, such as neural networks model; multi-model fusion prediction method [2-4]. Among them, the historical average model and the time series model are simple in modeling, but cannot represent the changes of traffic flow; the Kalman filter model has high accuracy, but has poor performance in predicting nonlinear traffic flow. The advantage of artificial intelligence models is that they have the characteristics of identifying complex nonlinear systems, of which BP neural network is the most commonly used model and is widely used in traffic flow prediction. The BP neural network is a static feed-forward network, and it is easy to fall into the local extreme value, which restricts the generalization ability of the BP network in traffic flow prediction. The extreme learning machine (ELM) is a new type of fast learning algorithm. For single-hidden layer feedforward neural networks (SLFNs), the extreme learning machine is similar to the traditional support vector machine, Compared with the back-propagation neural network, it inherits the advantages of the neural network approximating nonlinear functions, and has the characteristics of simple principle, fast training speed, high prediction accuracy and good generalization ability. The learning algorithm is fast. At present, there are few application scenarios of this algorithm in short-term traffic flow prediction, and there are even fewer related researches on model optimization and improvement of ELM.

At the same time, due to the complex and nonlinear characteristics of the traffic system, the traffic

ISSN 2522-3488 Vol. 6, Issue 3: 55-62, DOI: 10.25236/IJNDES.2022.060310

flow changes are random, and the traffic flow can be regarded as a nonlinear and strong interference signal, in which the noise is caused by some unexpected events, such as traffic accidents or extreme weather. drastic changes in traffic flow. For non-stationary process signals, some traditional denoising methods have their limitations. Wavelet Transform has become a powerful tool for signal denoising due to its advantages of time-frequency localization and flexibility of base selection. Some scholars have applied Wavelet Transform to traffic. Denoising of streaming data.

Therefore, this paper uses the Wavelet Transform (WT) method to denoise the traffic flow data in advance to improve the data quality of the dataset. Then, three improvement strategies based on Sine chaotic map and reverse learning strategy are used to initialize the population, nonlinear convergence factor adjustment and the introduction of dynamic weights to improve the GWO algorithm, so that it can more effectively avoid local optimization, accelerate the convergence speed and improve the optimization effect. Finally, the IGWO-ELM model is constructed, and the traffic flow forecast by the ELM forecast model is optimized by using the optimized weights and thresholds and updating the optimal parameters of the forecast model, thereby improving the forecast accuracy.

2. Model introduction

2.1 Wavelet Transform (WT)

Wavelet Transform (WT) is a new transformation analysis method. Its basic idea is to use the wavelet basis function to approximate the original signal, expand the signal into a linear superposition of the wavelet function family, and refine the analysis through scaling and translation. Decompose a series of sub-signals with different frequency characteristics, and adjust the time-frequency resolution by changing the shape of the window function, so that these sub-signals have good resolution in both the time and frequency domains, and it is easy to distinguish the abrupt part of the signal and noise , so as to achieve signal noise reduction [5]. It is precisely because of this excellent feature of multi-resolution that the Wavelet Transform quickly became a hot spot in the research of non-stationary signals once it was proposed. The mathematical principle of Wavelet Transform is as follows:

Then $\psi(t)$ is called a basic wavelet or mother wavelet. The mother wavelet $\psi(t)$obtains a family of functions through scaling and translation:

$$\psi_{(a,b)}(t) = \frac{1}{\sqrt{|a|}}\psi\left(\frac{t-b}{a}\right) \tag{1}$$

Where $a,b \in R, a \neq 0, a,b$ are the scaling factor and translation scaling factor of $\psi_{(a,b)}(t)$, respectively.

2.2 Grey Wolf Optimizer

Grey Wolf Optimizer (GWO) is a meta-heuristic optimization algorithm that simulates the leadership level and hunting mechanism of gray wolf populations in nature [6]. Generally, gray wolf populations can be divided into four categories: leader wolf α, Deputy leader wolf β, ordinary wolf δ, bottom wolf ω. Among them, wolf α have the strongest leadership ability and are mainly responsible for decision-making and wolf group management in the process of seeking optimization (predation); wolf β strengthen the domination of wolf α wolves and provide timely feedback to them; Wolf δ must obey the orders of the wolf α and wolf β , and are responsible for scouting and besieging their prey in the group; the lowest level wolf ω is the mass base of the wolf group and must obey the command of the wolf group leadership.

The whole hunting process of the gray wolf algorithm is mainly divided into three stages: Encirclement, Pursuit and Attack. Finally capture the prey (obtain the global optimal solution).

(1) Encirclement

After wolves determine the location of their prey, they must first surround the prey. The position Xi of the ith gray wolf satisfies the following:

$$X_i(t+1) = X_p(t) - A_i\left|C_i X_p(t) - X_i(t)\right| \tag{2}$$

Among them, t is the number of iterations, $X_p(t) = (X_p^1(t), X_p^2(t), \cdots, X_p^D(t))$ is the optimal position of hunting after t iterations, $X_p^D(t))$ refers to the position of the *p*-th gray wolf in the D

ISSN 2522-3488 Vol. 6, Issue 3: 55-62, DOI: 10.25236/IJNDES.2022.060310

dimension, and the optimization parameters are Hys and TTT, so D=2, $A_i\left|C_iX_p(t)-X_i(t)\right|$ is the encircling step, A_i and C_i is defined as:

$$A_i = 2a\cdot r_1 - a \tag{3}$$

$$C_i = 2r_2 \tag{4}$$

Among them, $r_1$and r_2 are random numbers in the interval [0,1], respectively, and a decreases linearly in the interval [2,0] with the increase of the number of iterations.

(2) Pursuit

After the gray wolves surround their prey, the wolf β and δ, led by the wolf α, hunt down the prey. In this process, the position of the gray wolf will change with the change of the prey position to find the optimal position X_p, which is usually updated according to the positions $X_\alpha, X_\beta, X_\delta$ of the wolf α,β,δ:

$$X_{i,\alpha}(t+1) = X_\alpha(t) - A_{i,1}\left|C_{i,1}X_\alpha(t) - X_i(t)\right| \tag{5}$$

$$X_{i,\beta}(t+1) = X_\beta(t) - A_{i,2}\left|C_{i,2}X_\beta(t) - X_i(t)\right| \tag{6}$$

$$X_{i,\delta}(t+1) = X_\delta(t) - A_{i,3}\left|C_{i,3}X_\delta(t) - X_i(t)\right| \tag{7}$$

$$X_i(t+1) = \sum_{j=\alpha,\beta,\delta} \omega_j X_{i,j}(t+1) \tag{8}$$

Among them, ω_j represents the weight of the jth wolf.

$$\omega_j = \frac{f\left(X_j(t)\right)}{f\left(X_\alpha(t)\right) + f\left(X_\beta(t)\right) + f\left(X_\delta(t)\right)} \tag{9}$$

$f(X_j(t))$ represents the fitness of the jth wolf at time t.

(3) Attack

The attack process is the last stage of the grey wolf's predation. The wolves attack and capture the prey. The realization of this process is mainly realized by decreasing the value of a in the following formula.

$$a = 2 - \frac{t}{t_{max}} \tag{10}$$

t_{max} is the maximum number of iterations. When the value of a decreases linearly from 2 to 0, the corresponding value of A_i also changes in the interval $[-a,a]$.When A_i≤1, it means that the wolves are approaching the prey position. When $1 < |A_i| \leqslant 2$, it indicates that the wolves are dispersing away from the prey, and GWO falls into a local optimum.

2.3 Improved Grey Wolf Optimizer Algorithm (IGWO)

The GWO algorithm has the characteristics of simple structure, clear concept, easy implementation and good global performance. However, because the traditional GWO optimization algorithm may fall into the local optimal solution in the optimization process, the convergence speed is slow, the optimization effect is poor, and the optimization efficiency is reduced greatly. Therefore, this paper chooses the following three methods to improve the GWO algorithm [7].

2.3.1 Population initialization based on Sine chaotic map and reverse learning strategy

Usually, the GWO algorithm is based on the initial population generated by randomness when solving the optimization problem, which may make the initial population distribution of the algorithm uneven, and it is difficult to ensure the diversity of the population. The initial solution of the algorithm plays a crucial role in the global search of the algorithm, and the diversity of the population also greatly affects the search performance of the algorithm [8]. Therefore, this paper adopts Sine chaotic map and reverse learning strategy to generate the initial population.

Sine map is derived from a sine function, which converts an input angle in the range $[0,1/\pi]$ to an output in a certain range [9]. Use the Sine mapping model to generate chaotic sequences for population initialization, and its mathematical expression is:

$$x_{i+1} = S(x_i) = r\,sin(\pi x_i) \tag{11}$$

Where $S(x_i)$ represents the Sine map.

The reverse learning strategy first randomly initializes the population in the solution space to generate a set of random populations. Each individual in the population $x_{i,j}$, $i \in [1, D]$ represents the dimension, $j \in [1, n]$ represents the number of populations. Then, a reverse learning strategy is used for each individual, that is, formula (12) is used to find the reverse individuals to construct a set of reverse solutions, and finally the two groups of initial solutions are combined, and the individuals with the top n fitness are selected as the initial population.

$$x'_{i,j} = x_{i,max} + x_{x,min} - x_{i,j} \tag{12}$$

2.3.2 Adjustment of nonlinear convergence factor

The convergence factor a of the traditional GWO algorithm decreases linearly with the number of iterations. Studies have shown that different update strategies for important parameters will greatly affect the performance of the algorithm, and generally in the optimization process, the linear strategy is not the most effective. Therefore, this paper introduces Nonlinear convergence factor [10-11].

$$a = 2 - 2\left[\frac{e^{\frac{t}{t_{max}}} - 1}{e - 1}\right]^k \tag{13}$$

Among them, the nonlinear adjustment coefficient k is a key parameter. Different values of k are used to control the change rate of the convergence factor before and after the change, so as to control the optimization convergence process, which has an important impact on the optimization performance of the algorithm. After research, the performance is the best when k=2.

2.3.3 Introducing dynamic weights

In the optimization process of the traditional GWO algorithm, the position of the head wolf is not necessarily the global optimal solution, and other wolves are also easy to fall into the local optimal solution, which leads to a decrease in the convergence speed and a long time. In order to optimize efficiently [12], In this paper, a proportional weight based on the Euclidean distance of the improved step size is introduced, and the position of the wolves is dynamically adjusted during the optimization process to speed up the optimization speed and reduce the optimization time. The mathematical expression is as follows:

$$W_1 = \frac{|\boldsymbol{X}_1|}{|\boldsymbol{X}_1| + |\boldsymbol{X}_2| + |\boldsymbol{X}_3| + \varepsilon} \tag{14}$$

$$W_2 = \frac{|\boldsymbol{X}_2|}{|\boldsymbol{X}_1| + |\boldsymbol{X}_2| + |\boldsymbol{X}_3| + \varepsilon} \tag{15}$$

$$W_3 = \frac{|\boldsymbol{X}_3|}{|\boldsymbol{X}_1| + |\boldsymbol{X}_2| + |\boldsymbol{X}_3| + \varepsilon} \tag{16}$$

Combined with the previous adaptive position update method, the final wolf position update method is as follows:

$$\boldsymbol{X}(t+1) = \frac{W_1\boldsymbol{X}_1 + W_2\boldsymbol{X}_2 + W_3\boldsymbol{X}_3}{3}\left(1 - \frac{t}{T}\right) + \boldsymbol{X}_1\frac{t}{T} \tag{17}$$

Combining the above three methods, the local optimum can be avoided more effectively, the convergence speed is accelerated, the solution accuracy is improved, and the global search and local development of the algorithm are better balanced.

2.4 Extreme Learning Machine

Extreme Learning Machine (ELM) is a generalized single-hidden layer feedforward neural network with fast learning speed and good generalization ability [13-16]. Given M samples $X_k = \{x_k, y_k\}$, $k = 1, 2, \cdots, M$, where x_k is the input data, y_k is the real value, $f(\cdot)$ is the activation function, and there are m hidden layer nodes, the ELM output can be expressed as:

ISSN 2522-3488 Vol. 6, Issue 3: 55-62, DOI: 10.25236/IJNDES.2022.060310

$$o_j = \sum_{i=1}^{m} \lambda_i f(W_i \cdot X_k + b_i), k = 1,2,\cdots,M \tag{18}$$

In the formula: o_j is the output value; $W_i = \{\omega_{i1}, \omega_{i2}, \cdots, \omega_{im}\}'$ is the connection weight between the input layer node and the ith hidden layer node; b_i is the bias value between the ith input node and the hidden layer node; λ_i is the ith hidden layer node and the output node connection weight.

3. Experimental verification and result analysis

3.1 Dataset

The dataset uses traffic flow information from Heathrow Airport on the M25 highway in the UK. At 15min intervals, the training set data contains 2688 traffic flow information from September 1, 2019 to September 28, and the validation set contains 192 traffic flow information from September 29 to September 30, 2019. Get 2880 data.

3.2 Evaluation indicators

The prediction accuracy is usually evaluated by means of Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE). The mathematical expressions corresponding to each indicator are as follows:

$$MAE = \frac{\sum_{i=1}^{N} |Y_i - y_i|}{N} \tag{19}$$

$$MAPE = \frac{100\%}{N} \sum_{i=1}^{N} \left| \frac{Y_i - y_i}{y_i} \right| \tag{20}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N} (Y_i - y_i)^2}{N}} \tag{21}$$

3.3 Application Examples

3.3.1 Denoising by Wavelet Transform

From September 1, 2019 to September 20, the traffic volume data for a total of 30 days is shown in Figure 1.

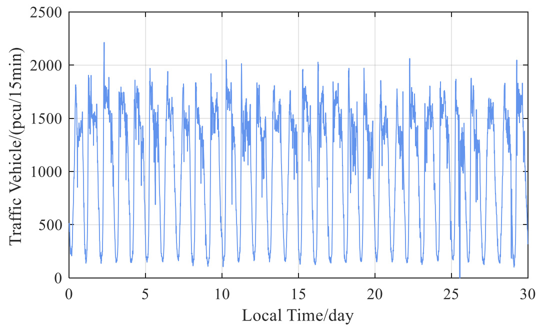

Figure 1: 30-day traffic vehicles without Wavelet Transform

In order to eliminate the interference of white noise, Wavelet Transform is used. Firstly, the original signal is decomposed by wavelet to obtain each detail component (high frequency) and approximate component (low frequency), and then the detail component is subjected to threshold processing. Finally, the processed components are used for wavelet reconstruction to obtain a pure signal. As shown in figure 2.

ISSN 2522-3488 Vol. 6, Issue 3: 55-62, DOI: 10.25236/IJNDES.2022.060310

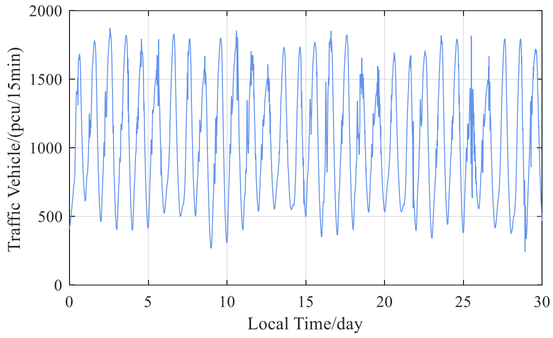

Figure 2: 30-day traffic vehicles after Wavelet Transform

3.4 Model prediction and analysis

In order to verify the validity of the prediction method in this paper, the short-term traffic flow prediction models of WT-IGWO-ELM and ELM and WT-ELM, GWO-ELM and WT-GWO-ELM, and WT-IGWO-ELM were carried out under the same conditions. Comparative Experiment. The above six prediction models are established by programming algorithm programs in MATLAB language, and the accuracy of the prediction methods is compared by three error indicators: Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE).

(1) Prediction effect without Wavelet Transform

ELM, GWO-ELM and IGWO-ELM without Wavelet Transform have been simulated, and the actual measured data and predicted results are shown in Figure 3. It can be seen that the result curves of the three prediction models without Wavelet Transform can basically fit the actual data, but the convergence speed is low. When the data fluctuation is large, the accuracy rate is low, and it is larger than the real value. It can be seen from Figure 3 that the prediction results of the IGWO-ELM model without Wavelet Transform are in the best agreement with the actual data after simulation, and the accuracy is high. The GWO-ELM prediction effect is second, and the ELM prediction effect is the worst.

(2) Prediction effect without Wavelet Transform

The WT-ELM, WT-GWO-ELM and WT-IGWO-ELM denoised by Wavelet Transform are simulated, and the measured data and prediction results are shown in Figure 4. It can be seen that the results of the three prediction models denoised by Wavelet Transform and the measured data are all better than those before denoising. Among them, the prediction results of WT-IGWO-ELM are in the best agreement with the actual data and the prediction is accurate, the prediction effect of WT-GWO-ELM is second, and the prediction effect of ELM is third.

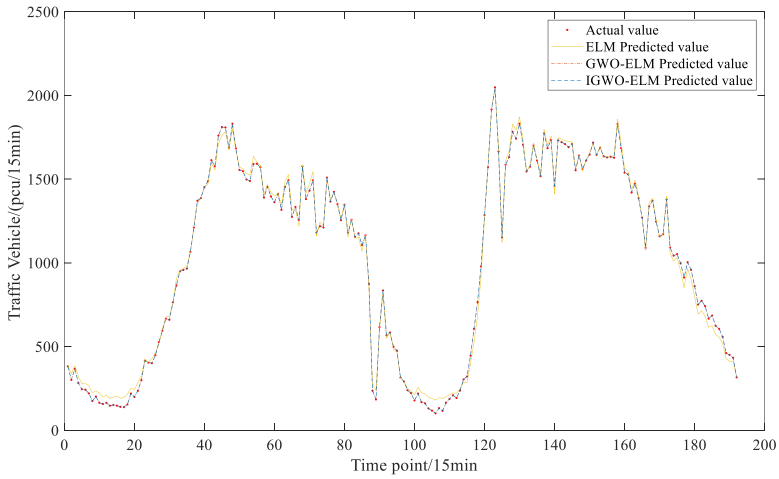

Figure 3: Prediction results without Wavelet Transform

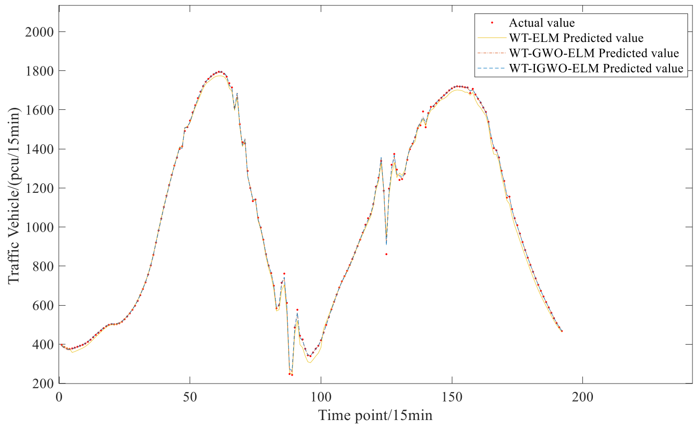

Figure 4: Prediction results after Wavelet Transform

The above prediction method is used to predict the short-term traffic flow in MATLAB, and the prediction error results are shown in Table 1.

Table 1: Prediction error results

No.	Prediction method	MAE	MAPE	RMSE
1	ELM	26.7539	3.8866%	2.7590
2	GWO-ELM	5.0565	1.5522%	0.4666
3	IGWO-ELM	3.3363	0.4919%	0.2643
4	WT-ELM	21.2946	3.1860%	1.5836
5	WT-GWO-ELM	3.3257	0.7605%	0.3476
6	WT-IGWO-ELM	0.4837	0.2498%	0.0490

From the prediction results, it can be seen that:

(1) The average relative error of prediction of GWO-ELM and WT-GWO-ELM is 74.7498% and 79.5199% smaller than that of ELM and WT-ELM respectively, indicating that GWO optimized ELM neural network can greatly improve the accuracy of ELM prediction model.

(2) The average relative error of prediction of IGWO-ELM and WT-IGWO-ELM is reduced by 48.5616% and 79.5021% compared with GWO-ELM and WT-GWO-ELM respectively, indicating that the optimization effect of IGWO on ELM is more obvious than that of GWO, and it is expected to achieve High precision forecast.

(3) The prediction errors of the latter three in Table 1 are reduced by 27.0114%, 36.9110%, and 72.0571% respectively compared with the former three, indicating that the Wavelet Transform denoising further improves the accuracy of the traffic flow prediction model.

(4) Compared with ELM, WT-ELM, GWO-ELM, WT-GWO-ELM and IGWO-ELM, the average relative error of prediction of WT-IGWO-ELM model is reduced by 96.6625%, 95.5972%, 87.9447%, 79.5021%, 72.0571% respectively, it can be seen that the prediction effect of the WT-IGWO-ELM model is much better than that of ELM, WT-ELM, GWO-ELM, WT-GWO-ELM and IGWO-ELM.

4. Conclusion

High-precision short-term traffic flow prediction can provide powerful auxiliary decision-making information for signal control, traffic guidance, route planning and other systems. In this paper, the reliable short-term traffic flow data provided by authoritative institutions is used as the research object, and the WT-IGWO-ELM prediction model is proposed. On the basis of denoising the short-term traffic flow data by Wavelet Transform, the ELM neural network model optimized by the improved gray wolf optimization algorithm is used to simulate the real short-term traffic flow data. Experimental results show that:

(1) The Wavelet Transform denoising reduces the prediction errors of the three prediction models of ELM, GWO-ELM and IGWO-ELM to different degrees, which can further improve the accuracy of the traffic flow prediction model.

(2) Using the GWO algorithm to optimize the ELM neural network can greatly improve the accuracy

ISSN 2522-3488 Vol. 6, Issue 3: 55-62, DOI: 10.25236/IJNDES.2022.060310

of the ELM prediction model.

(3) Using IGWO algorithm to optimize ELM neural network, its optimization effect is more obvious than GWO, and it is more expected to achieve high-precision prediction.

(4) The prediction effect of the WT-IGWO-ELM model is much better than that of ELM, WT-ELM, GWO-ELM, WT-GWO-ELM and IGWO-ELM, and it is an effective and high-precision short-term traffic flow prediction method.

References

[1] Daraghmiya, Yicw, Chiangtc. Negative binomial additive models for short-term traffic flow forecasting in urban areas. .IEEE Transactions on Intelligent Transportation Systems, 2014,15(2):784-793 .

[2] Song Chi, Shen Guojiang. Review of short-term traffic flow forecasting model. Automation Expo, 2012, 29(06):84-87.

[3] Liu Haihong, Zhou Chunmei. A review of short-term traffic flow prediction methods. Journal of Wuhu Vocational and Technical College, 2011,13(04):31-33.

[4] Yang Fengman. A review of traffic flow prediction methods based on artificial neural network. Highway Transportation Science and Technology, 2020,37(S1):130-135.

[5] Huang Darong, Song Jun, Wang Da, Cao Jianqiu, Li Wei. Research on Traffic Flow Prediction Model Based on ARMA and Wavelet Transform. Computer Engineering and Application, 2006(36):191-194+224.

[6] Mirjalili S, Mirjalili S M, Lewis A . Grey Wolf Optimizer. Advances in Engineering Software, 2014, 69(3):46–61.

[7] Zhang Yang, Zhou Xizhao. Improved Grey Wolf Algorithm for Global Optimization Problems. Journal of University of Shanghai for Science and Technology, 2021,43(1):73-82.

[8] Ni Jing, Qin Bin, Zeng Fanlong. An Improved Grey Wolf Optimization Algorithm with Mixed Strategies . Software Guide, 2021,20(05):72-76.

[9] CHO K,MIYANO T. Chaotic cryptography using augmented Lorenz equations aided by quantum key distribution [J/OL]. IEEE Transactions on Circuits and Systems I: Regular Papers,2015,62 (2):478-487.

[10] Tan Fafang, Zhao Junjie, Wang Qi. Research on a Gray Wolf Optimization Algorithm with Improved Nonlinear Convergence Method. Microelectronics and Computers, 2019, 36(5): 89-95.

[11] Guo Zhenzhou, Liu Ran, Gong Changqing, et al. Improvement research based on gray wolf algorithm. Computer Application Research, 2017, 34(12): 3603-3606+3610.

[12] Wang Qiuping, Wang Mengna, Wang Xiaofeng. Grey wolf optimization algorithm with improved convergence factor and proportional weight. Computer Engineering and Applications, 2019, 55(21): 60–65, 98.

[13] Zhou Borong, Chen Weiguo, Xu Zhenyi, Wen Xiulan. Research on short-term traffic flow prediction model based on improved online extreme learning machine. Computer Engineering and Science, 2022, 44(05): 944-950.

[14] Rui Lanlan, Li Qinming. Short-term traffic flow prediction algorithm based on combined model. Journal of Electronics and Information, 2016, 38(05): 1227-1233.

[15] Ganlu. Research and application of extreme learning machine [D]. Xidian University, 2014.

[16] Wang Jie, Bi Haoyang. An extreme learning machine based on particle swarm optimization. Journal of Zhengzhou University (Science Edition), 2013, 45(01): 100-104.