Neural Computing and Applications (2021) 33:8181–8190
https://doi.org/10.1007/s00521-020-04932-5

S. I: INTELLIGENT COMPUTING METHODOLOGIES IN MACHINE LEARNING FOR IOT APPLICATIONS

Short-term traffic flow prediction based on improved wavelet neural network

Qiuxia Chen[1] · Ying Song[2] · Jianfeng Zhao[3]

Received: 29 January 2020 / Accepted: 6 April 2020 / Published online: 29 April 2020

Abstract

Due to the characteristics of time-varying traffic and nonlinearity, the short-term traffic flow data are difficult to predict accurately. The purpose of this paper is to improve the short-term traffic flow prediction accuracy through the proposed improved wavelet neural network prediction model and provide basic data and decision support for the intelligent traffic management system. In view of the extremely strong nonlinear processing power, self-organization, self-adaptation and learning ability of wavelet neural network (WNN), this paper uses it as the basic prediction model and uses the particle swarm optimization algorithm for the slow convergence rate and local optimal problem of WNN prediction algorithm. With the advantages of fast convergence, high robustness and strong global search ability, an improved particle swarm optimization algorithm is proposed to optimize the wavelet neural network prediction model. The improved wavelet neural network is used to predict short-term traffic flow. The experimental results show that the proposed algorithm is more efficient than the WNN and PSO–WNN algorithms alone. The prediction results are more stable and more accurate. Compared with the traditional wavelet neural network, the error is reduced by 14.994%.

Keywords Short-term traffic flow forecasting · Wavelet neural network · Particle swarm algorithm · Improved particle swarm optimization

1 Introduction

With the rapid development of urbanization and social economy, the number of motor vehicles and the rapid growth of urban population have led to excessive road load. The existing road network capacity can no longer meet people's growing traffic demand; it has caused huge material and economic losses to people's lives [1, 2]. Due to the deterioration of traffic, with the traffic congestion, the frequency of traffic accidents is getting higher and higher, which hinders the development of the city to some extent. In the face of the above problems, relevant staff have taken measures to control traffic demand, increase road infrastructure and strengthen traffic management [3, 4]. However, each solution will be subject to certain aspects; from the perspective of socioeconomic development, controlling the increase in vehicle use can only be a temporary measure; from the point of view of the government's material and financial resources and the current space of existence, it is impossible to re-plan the road; by controlling road traffic lights and strengthening traffic regulations and other measures to improve many traffic problems, it can only temporarily relieve the effects and cannot solve the problem fundamentally. Therefore, how to fundamentally improve the traffic capacity of the road

✉ Ying Song
46071@nchu.edu.cn

Qiuxia Chen
chenqiuxia@szpt.edu.cn

Jianfeng Zhao
zjf@szpt.edu.cn

[1] School of Computer Engineering, Shenzhen Polytechnic, Shenzhen 518055, Guangdong, China

[2] School of Business, Wuzhou University, Wuzhou 543002, Guangxi, China

[3] School of Automotive and Transportation Engineering, Shenzhen Polytechnic, Shenzhen 518055, Guangdong, China

network is very important [5, 6]. In the last century, developed countries represented by the USA and Europe have effectively applied information technology, sensor technology, computer technology and data communication technology to transform traditional transportation industries. Provides drivers with real-time road traffic information through communication, control and other high-tech technologies, and guides drivers to choose driving routes reasonably [7]. In order to alleviate the deterioration of traffic, reduce the frequent occurrence of traffic congestion and accidents, it is extremely urgent to improve traffic capacity. Intelligent transportation system is a real-time, accurate and efficient ground comprehensive traffic management system that can play a full range of functions in a wide range [8, 9]. Traffic guidance systems and traffic control systems, as important components of intelligent transportation systems, have played a significant role in improving road congestion, reducing accidents and reducing energy consumption [10]. In this society where science is developing rapidly and artificial intelligence is widely used, it is inevitable to implement intelligent transportation systems as soon as possible [11].

Since the transportation system is a complex system, it is difficult to solve the problem fundamentally from the perspective of the vehicle or the road. In order to alleviate the deterioration of traffic and reduce the frequent occurrence of traffic congestion and accidents, it is extremely urgent to improve traffic capacity [12]. The intelligent transportation system is a ground integrated traffic management system that is real time, accurate and efficient and can perform all round functions in a wide range. Through this system, you can understand the traffic conditions of the entire city, so as to play a guiding and clearing role for traffic. It is proposed to reduce energy consumption, improve efficiency and reduce pollution. An intelligent transportation system uses historical traffic flow data to make short-term predictions for future STTF [13]. The information transmits the real-time traffic flow information on the road section to the driver. The driver can make a reasonable route planning according to the obtained information and induce the trend of the traffic flow through the information, so that the people, the car and the road can be closely coordinated to a certain extent. Reduce the travel time, effectively avoid the traffic congestion section to a certain extent, and ensure the overall normal operation of the surrounding roads [14]. The core function of the intelligent transportation system is the traffic guidance function. The degree of goodness and badness determines the degree of road compliance. To improve the traffic guidance ability of the intelligent transportation system, in addition to the analysis of the real-time traffic flow data, it is necessary to the traffic flow data on the road segment at the next moment is a precise prediction. By analyzing the current and future road conditions of the road segment, reasonable route planning is carried out to maximize the utilization of the road. Therefore, a better traffic flow prediction (TFP) function can avoid some potential traffic accidents that may occur in the future. The STTF prediction determines the effectiveness and reliability of the traffic guidance function to a certain extent [15, 16].

QIAO's team proposed that the traditional ramp control algorithm does not consider the overflow problem of urban highway ramp queues. Combined with the classic ALINEA ramp control algorithm, a new method of urban expressway ramp control based on urban expressway traffic flow prediction is proposed. A wavelet neural network based on genetic algorithm optimization is proposed for urban expressway traffic flow prediction. The method also introduces the gap acceptance theory and the grading principle of the squall queue, which realizes the dynamic adjustment of the urban highway ramp control rate. The results show that the model can effectively improve the capacity of the main road and can shorten the average travel time of the ramp by about 24.8% [17]. Abdi J's team explored that the dynamic behavior of the transportation system requires not only the mathematical model of the system, but also the prediction of traffic flow in the system. Therefore, traffic flow prediction plays an important role in today's intelligent transportation systems. They introduced a short-term traffic flow prediction method based on artificial neural networks. Among many neural networks, multi-layer perceptron, radial basis function neural network and wavelet network were selected as the three best candidates. In addition, backpropagation (BP) is the most effective learning program in all cases. The results show that the coefficients generated by the time domain signal improve the performance of the BP learning algorithm. The time domain signal provides researchers with a new time domain differential BP learning algorithm (TDBPL) model. The ability and performance of TDBPL algorithm are verified by simulation. It is proved that wavelet theory has multi-resolution ability compared with RBF neural network, and it is an algorithm suitable for traffic flow prediction. It is also concluded that RBF neural networks do not provide negative predictions despite the application of MLP. In addition, in the MLP algorithm, the local minimum value problem is unavoidable, while the RBF neural network and the wavenet network are not encountered [18]. Yang's team proposed that traffic flow prediction is an effective technique to solve traffic congestion problems and improve traffic mobility. Neural network related methods have been applied to the development of traffic prediction models for more than 20 years. Since neural networks are sensitive in terms of parameter selection, choosing the appropriate modeling configuration is critical to improving the accuracy and efficiency of traffic flow prediction. However, this is usually done by trial and error, which is very timeconsuming and involves too many design factors. Therefore, this paper uses the robust system optimization method—Taguchi method—to obtain the optimal configuration of exponential smoothing and extreme learning machine prediction model. The model is applied to real data collected by UK highways and highways and compared to the three existing prediction models. The results show that the Taguchi method is effective and feasible for traffic flow prediction model design. The proposed optimal configuration model has better prediction performance in highway and highway traffic flow prediction, and accurate prediction during peak and off-peak hours. The rates were 91% and 88%, respectively [19]. For Tang's team in the past few decades, short-term traffic flow prediction has been an important topic in traffic research. Its purpose is to actively carry out dynamic traffic control by monitoring current traffic conditions and predicting future traffic conditions. In this paper, we focus on using the data collected by the automated toll collection system and various external factors to predict the short-term passenger flow of a subway station, where passenger flow refers to the passenger flow to the subway station within a given time period. On this basis, we propose a data-driven three-phase framework for short-term passenger flow forecasting, including traffic data analysis, feature extraction and predictive modeling. The effects of time, spatial characteristics and external weather on passenger flow prediction are studied. Various predictive models, including time series models, autoregressive synthetic moving averages, linear regression and support vector regression, were used to evaluate the performance of the framework. In addition, we conducted a large number of experiments on method validation, feature evaluation and data resolution demonstration using real data sets collected from the Shenzhen AFC system [20].

This paper mainly studies the improvement in wavelet neural network and the use of improved wavelet neural network for short-term traffic flow prediction. Since the traditional wavelet neural network is easy to fall into the local optimum, and this will affect the final prediction accuracy, the first focus of the research is to find a suitable method to improve the wavelet neural network, and the improved wavelet neural network is used to predict the short-term traffic flow, in order to obtain better short-term traffic flow forecast results.

2 Proposed method

2.1 Traffic flow forecast

Since the amount of traffic on the road depends on the number of people driving on a certain section of the road in a certain period of time, it has the characteristics of time and space, and its distribution changes with space and time. Not only that, because of the personal habits of motorists and the traffic rules of the roads they are traveling, there are certain rules and characteristics of the entire traffic flow. It is mainly manifested in three aspects: the uncertainty of traffic flow, the correlation of traffic flow and the cyclical change of traffic flow.

Traffic flow forecasting data have two attributes, time and space, which are two-dimensional. The time attribute of the data means that on a fixed road segment, the traffic flow will have a certain regularity within a certain time interval, and the historical average can also interpret the time attribute well. The space attribute refers to a fixed time. On the road segment, the flow rate of the upstream and downstream sections adjacent to the road section has a certain correlation. Traffic flow prediction generally comes from various collectors and needs to be preprocessed before use. Traffic flow forecast is divided into long-term traffic flow forecast and short-term traffic flow forecast according to the length of forecast. The forecast period is generally short, because the forecast is not worth too long, and the change is too fast and the regularity is low. In other words, the upper limit of the short-term traffic flow forecast period is 15 min. Its accurate predictions provide better path planning for pedestrians and vehicles and are the basis for future intelligent transportation system design.

Traffic flow has uncertain characteristics, and the factors that cause traffic flow uncertainty are numerous and complicated. Uncertain factors include natural factors such as seasonal and climatic factors, as well as human factors such as traffic accidents and emergencies. These factors have brought great difficulties to traffic flow forecasting, especially for short-term traffic flow forecasts. For short-term traffic flow prediction, as the forecasting scale shortens the traffic to show high randomness and uncertainty, and the prediction results have high precision requirements, some common methods cannot satisfy this nonlinearity and have forecast of traffic flow with more complex features. It is increasingly difficult to design a high precision prediction method based on deterministic mathematical models. The prediction method based on mathematical model cannot meet the real-time requirements of short-term traffic flow prediction.

Traffic flow prediction is to make full use of the data provided by various different traffic flow collection devices and to find the regularity from the randomness through the analysis of relevant methods. Therefore, it is first necessary to sort the collected data, followed by analyzing and misidentifying the data, then selecting the appropriate model, inputting the predicted sample data in the model and finally predicting the traffic state in the next period. The basic flow of traffic flow prediction is shown in Fig. 1.

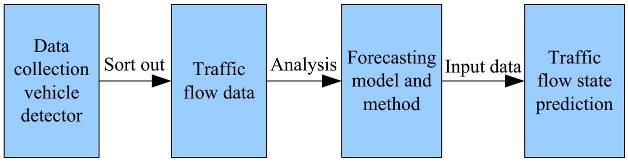

Fig. 1 Forecast flowchart for short-term traffic flow

For short-term traffic prediction models, short-term traffic flow prediction models with the following characteristics are effective because of the requirement for real-time data analysis:

(1) *Accuracy* The purpose of short-term traffic flow forecasting is to serve traffic control systems and traffic guidance services. Only the traffic flow forecasting results have certain accuracy, so that the traffic management department can adjust the traffic system and travelers are traveling. The choice of time provides data support, and inaccurate traffic flow information will be meaningless.

(2) *Real time* The ultimate goal of traffic prediction is to act on traffic guidance systems and traffic control systems and to produce prediction results in an effective time.

Providing future state data of traffic flow, the predicted data information is required to have strong real-time performance.

(3) *Dynamic feedback* Since the traffic flow is random and complex, its law cannot be fixed. In order to obtain better prediction results, once the traffic flow has an abnormal situation, the prediction model can quickly make feedback on the data abnormality occurred in the prediction process and can make corresponding adjustments according to the needs.

It can be seen from the characteristics of traffic flow that traffic flow is characterized by uncertainty, randomness, complexity and cannot be formulated. Due to the complexity of the modeling process and the high precision requirements, the algorithm based on statistical theory leads to difficulties in solving, and the use of certain mathematical models to predict complex and nonlinear traffic flows cannot meet the demand. Among the current prediction methods based on nonlinear system theory, neural networks and support vector machines are the focus of research. Artificial neural networks have strong predictive advantages such as self-learning, adaptive and nonlinear mapping. Its modeling can avoid the complicated and cumbersome mathematical modeling process in the support vector machine model. Therefore, the neural network is particularly suitable for traffic flow prediction. Wavelet neural network combines the advantages of neural network and wavelet decomposition on function approximation. The prediction accuracy is higher than that of general neural network, and the convergence speed is faster. However, when using wavelet neural network for parameter optimization, it is easy to make the optimal solution fall into local minimum and cause oscillation effect. Therefore, in this paper, the convergence speed of the WNN prediction algorithm is slow and the local optimization problem is utilized. The particle swarm optimization algorithm has the advantages of fast convergence and strong global search ability. An improved particle swarm optimization algorithm is proposed to optimize the wavelet neural network prediction model, and the improved wavelet neural network is used to predict the short-term traffic flow.

2.2 Wavelet neural network

Wavelet theory has significant advantages in processing signals and so on and has been rapidly applied in many related fields. Compared with the general feed forward network and the RBF network, the wavelet neural network has the advantages of strong adaptability, simple network structure and good fault tolerance. The average value of the wavelet is generally 0, and the length is not infinite. The wavelet function is generally obtained by transforming on the basis of the mother wavelet, and the operations of the transformation are telescopic and translational. Wavelet network is to use wavelet function to represent or approximate function or signal. The structure of basic wavelet neural network is shown in Fig. 2.

Where x_1, x_2,…,x_k represent the input of the wavelet neural network; y_1, y_2,…,y_m represent the predicted output. w_{ij} is the connection weight between the input layer and the hidden layer. When the input signal sequence is $x_i(i = 1, 2,\ldots,k)$, the output of the hidden layer can be expressed as

$$h(j) = hj\left(\frac{\sum_{i=1}^{k} wijxi - bj}{aj}\right), j = 1, 2, \ldots, l \tag{1}$$

where $h(j)$ is the output of the jth hidden layer node, a_j is the scaling factor of the wavelet basis function, b_j is the translation factor of the wavelet basis function h_j and h_j is the wavelet basis function. In this paper, the Morlet wavelet basis function is used as a function of the hidden layer node, namely:

$$y = \cos(1.75x)e^{-x2/2} \tag{2}$$

The formula for the output layer is:

$$y(k) = \sum_{i=1}^{l} wikh(i), k = 1, 2, \ldots, m \tag{3}$$

where w_{ik} represents the connection weight of the hidden layer to the output layer, $h(i)$ is the output of the ith hidden

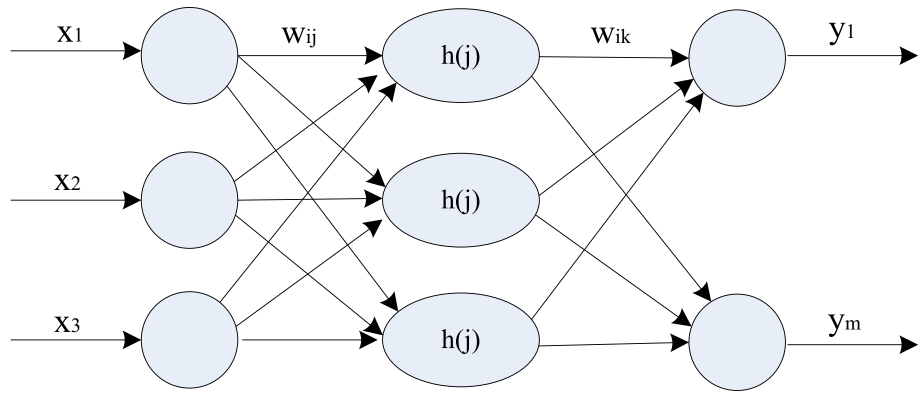

Fig. 2 Wavelet neural network structure

layer node, l represents the number of hidden layer nodes and m represents the number of output layer nodes.

2.3 Improved particle swarm optimization

(1) Basic particle swarm optimization algorithm

The particle swarm optimization algorithm is a kind of swarming and clustering behavior in the process of bird foraging. It is a widely used swarm intelligence global random search algorithm. When searching for solution space, the direction and distance of the particle's flight are mainly controlled by the change of speed. At the same time, it also comprehensively learns the historical experience of the self and the experience of other members of the group to achieve constant adjustment of the local search. The purpose of the optimal solution and the global optimal solution currently found by the entire population is to search continuously in the solution space until the end of the iteration. In the n-dimensional search space, the particle group is initialized and distributed in space, and the position of each particle represents a solution to a particular problem; each particle can find the optimal position in the target search space.

Suppose the search space of a particle swarm consisting of N particles is a D-dimensional vector $xid = (xi1, xi2, \ldots, xiD)$, the flight speed is represented by

$$vi = (vi1, vi2, \ldots, viD) \quad (i = 1, 2, \ldots, N)$$

, and the best solution obtained by the searched particles is $pi = (pi1, pi2, \ldots, piD)$. One of the best solutions for each particle is represented by $pg = (pg1, pg2, \ldots, pgD)$, which mimics the evolutionary iterative calculation method. The PSO algorithm needs to adjust the speed and position of its flight in each iteration. The position update formula of the speed v_{id}^{k+1} and $k + 1$ iterations of the ith particle is:

$$v_{id}^{k+1} = wv_{id}^{k} + c1r1\left(P_i^k - x_{id}^k\right) + c2r2\left(P_g^k - x_{id}^k\right) \tag{4}$$

$$x_{id}^{t+1} = x_{id}^{k} + v_i^{k+1} \tag{5}$$

The new location update formula is:

$$x_{id}^{t+1} = x_{id}^{t} \times wij + v_{id}^{t} \times w^t ij + rand() \times g\mathrm{best}^t \tag{6}$$

Among them, wij and w_{ij}^t are dynamic weights that control the current x_{id}^t and v_{id}^t, $g\mathrm{best}^t$ represents the optimal position of the particle, w is called the inertia weight and its value determines how much the current velocity of the particle is inherited and is used to balance PSO detection and development capabilities. The larger the value of w, the longer the distance the particles move in the original direction, thus improving the exploration ability, but the convergence is slow; the smaller the value of w, the better the development ability of the particles, but it is prone to local optimum. $c1$ and $c2$ are learning factors that indicate how close each particle is to its global best and historical merit. $r1$ and $r2$ are random numbers between [0, 1]. In order to reduce the possibility of particles leaving the search space during the iteration, vij is usually defined in a certain interval, $vij \in [-v\max, v\max]$.

(2) Improvement in particle swarm optimization algorithm

In order to ensure that the particles have good detection ability in the early stage of the iteration and have good development ability in the later stage, w needs to be dynamically adjusted with iteration. In this study, the variation uses a linear decreasing law, as shown by the formula:

$$w_i = w_{\max} - \frac{i(w_{\max} - w_{\min})}{\mathrm{I}_{\max}} \tag{7}$$

where w_i is the ith inertia weight; $I_{\max}$ is the maximum number of iterations; $w_{\min}$ and $w_{\max}$ are the minimum and maximum inertia weights, respectively.

Due to the lack of diversity of the particle position in the late search stage, the basic particle swarm algorithm has the disadvantages of being easy to fall into local extremum and slow convergence. Experiments show that in the initial search phase, c_1 is large, c_2 is small, and particles can be freely divided in the search space. As the number of iterations increases, c_1 decreases linearly and c_2 increases linearly, enhancing the global optimal convergence of the particles. The formula is expressed as:

$$c1i = c1s + \frac{i(c1e - c1s)}{\mathrm{I}_{\max}} \tag{8}$$

$$c2i = c2s + \frac{i(c2e - c2s)}{\mathrm{I}_{\max}} \tag{9}$$

where c_{1i} and c_{2i} are the values of c_1 and c_2 for the ith iteration; c_{1s} and c_{2s} are the initial values of c_1 and c_2; c_{1e} and c_{2e} are the final values of c_1 and c_2; $I_{\max}$ is the maximum number of iterations.

(3) Algorithm flow

The IPSO algorithm is similar to the PSO algorithm flow. The specific algorithm flow is as follows:

Step 1 Randomly generate the velocity and position of the N particles, and give the relevant parameters such as the maximum number of iterations M;

Step 2 Calculate the fitness value of all particles, set the ith particle to its current best position, and set the Nth particle to the global best position;

Step 3 Update the velocity and position of the particles according to Eqs. (4) and (7);

Step 4 Calculate the fitness value of the particle, and update the local extreme value and the global extreme value of the particle according to the calculated result;

Step 5 Determine whether the given termination condition is satisfied. If it is satisfied, stop the search and output the result; if not, return to step 3 to continue the iteration.

2.4 WNN parameters and optimization of IPSO

Traditional WNN usually uses one-way gradient descent method for network training. However, this method has the disadvantages of being easily trapped in local extremum, slow convergence and large training error, so the prediction accuracy is low. In this paper, IPSO is used to optimize the initial parameters of the WNN. Proceed as follows:

Step 1 Initialize the parameters of the wavelet neural network, and set the number of neurons in each layer of the network;

The connection weights w_{ij} and w_{ik}, the scaling factor a_j and the translation factor b_j, which need to be optimized, are used as position vectors for each particle. That is, present(m) = (w_{ij}, w_{ik}, a_j, b_j);

Step 2 Determine the number M of particles, the position vector p of the particle and the maximum value $p_{\max}$ and minimum value $p_{\min}$ of the position, the maximum value $v_{\max}$ and the minimum value $v_{\min}$ of the velocity vector v and the velocity, the dimension D and the maximum allowable number of iterations is $k_{\max}$. Learning factors c_1 and c_2, the maximum value of the inertia weight w is $w_{\max}$ and the minimum value is $w_{\min}$;

Step 3 Iteratively update the velocity and position of each particle according to formula (4) and formula (7);

Step 4 Calculate the fitness value of the particle according to the formula (10), and the locally searched optimal position pbest and the globally searched optimal position gbest:

$$F = \sum_{i=1}^{n} |yi - hi| \tag{10}$$

The fitness function F is the prediction error sum, where n is the total number of network output nodes, y_i is the predicted output of the ith node and h_i is the expected output of the ith node.

Step 5 Determine whether the number of iterations N reaches the maximum value $k_{\max}$. If not, return to step (3). Otherwise, stop the search. At this time, the optimal position and the corresponding wavelet neural network parameters are output.

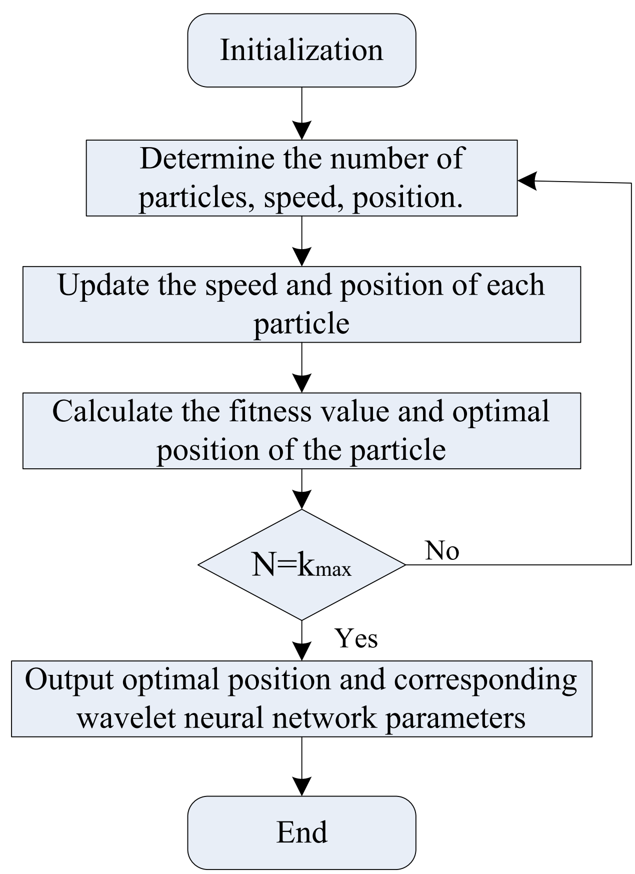

Fig. 3 Flowchart of the IPSO optimization algorithm

The performance of the algorithm flowchart is shown in Fig. 3.

3 Experiments

3.1 Experimental environment and model construction

This simulation experiment is based on the 64-bit processor of Windows8 system, realized in the Matlab2016a environment and realizes the construction and simulation of the network model through programming. Based on extensive research and practice, based on the temporal characteristics of traffic flow, a WNN with three-layer structure, including input layer, hidden layer and output layer, is designed. The traffic flow data of the first n time points relative to the current time are taken as the input of the WNN, and the Morlet mother wavelet function is used as the transfer function of the hidden layer node, and the traffic flow at the current time point serves as the expected output of the network.

The WNN structure used in this paper is 4–6–1. The four input layer nodes represent the traffic flow at the first four time points of the predicted time point. The selection of the six hidden layer nodes is based on the empirical method of neural network hidden node selection. Test determined: 1 output layer node is the predicted output of the traffic flow at the current time point, and the initial parameter of WNN is the optimal value of IPSO optimization.

3.2 Data sources and preprocessing

In order to verify the correctness and prediction accuracy of the model, traffic flow data at an intersection of an anonymous city were used. The traffic flow data of 4 days were collected and recorded every 15 min. The traffic flow data of 368 time points were recorded. The 273 time points of the previous 3 days were used for WNN training, and use the training network to predict the traffic flow on the fourth day. In data processing, due to the difference of the unit and the size of the variable value, the prediction error of WNN may be large. The collected sample data and simulation prediction values should be normalized. The normalization method is as follows:

$$x_i^l = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \tag{11}$$

where x_i^l is the normalized value of the input time series, xi is the original time series, $x_{\min}$ is the minimum value in the sample, $x_{\max}$ is the maximum value in the sample and finally the simulation prediction result still needs to be denormalized.

4 Discussion

4.1 WNN and PSO–WNN model training result analysis

In order to better illustrate the effectiveness of IPSO–WNN, the constructed IPSO–WNN model is used for short-term traffic flow prediction and compared with the existing WNN and PSO–WNN model prediction results. The prediction effects of the three prediction models are analyzed from different experimental results. The number of training samples in the first 3 days was 273 as the training data set in the experiment, and the 95 time point traffic flow data collected on the fourth day were used as the prediction data set. In the experiment, the three models are used to compare and analyze the results. Figures 4 and 5 compare the predicted and actual values of WNN and the predicted and actual values of PSO–WNN.

Comparing the WNN predicted value with the actual value and the PSO–WNN predicted value and the actual value, it can be seen that the PSO–WNN training result has a higher degree of fit than the WNN model training result, and the predicted result is more stable and accurate.

4.2 Analysis of training results of IPSO–WNN model

(1) Comparison of IPSO–WNN predicted value and actual value

Through the IPSO–WNN model, 273 training samples and 95 time point traffic flow data were tested to obtain experimental results. As shown in Fig. 6, the IPSO–WNN

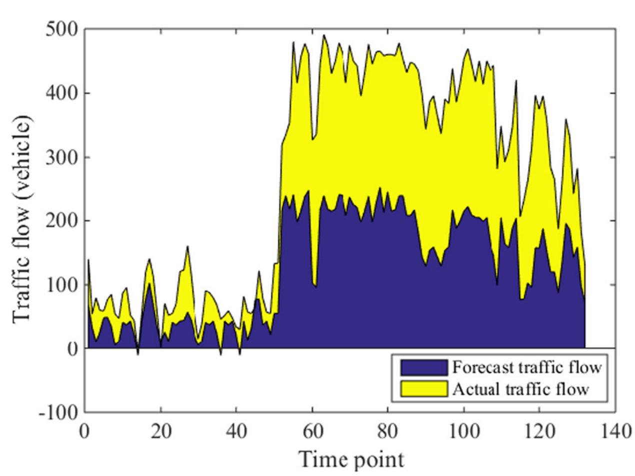

Fig. 4 Comparison of WNN predictors versus actual values

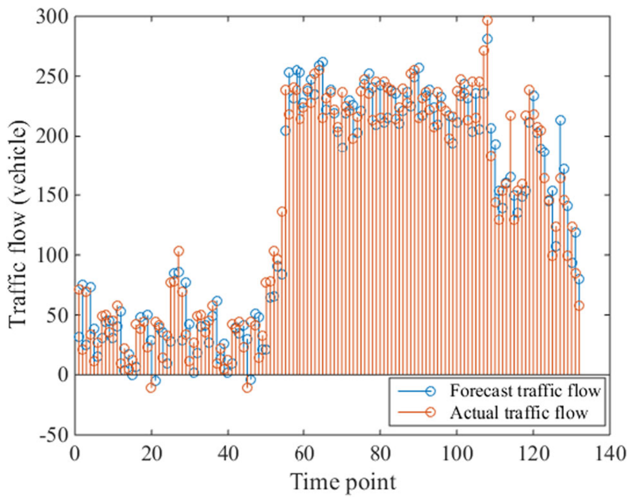

Fig. 5 Comparison of PSO-WNN predicted and actual values

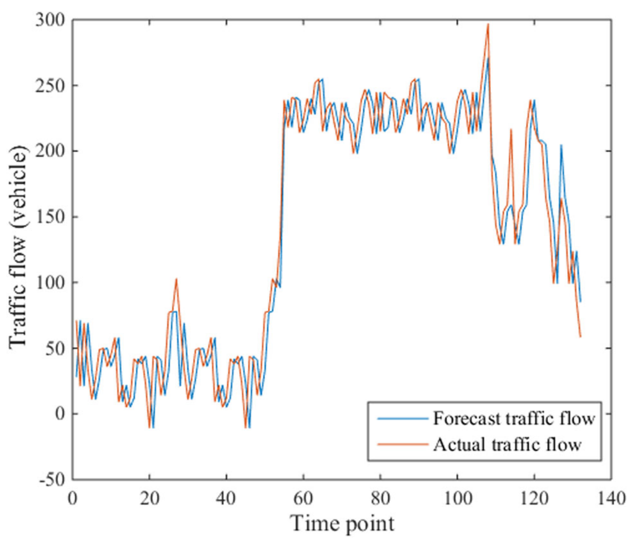

Fig. 6 Comparison of IPSO-WNN predicted and actual values

experimental results can be compared with WNN and PSO–WNN. The IPSO–WNN algorithm is more efficient than the WNN and PSO–WNN algorithms alone. The prediction results are more stable, more accurate and closer to the actual traffic flow. It has faster convergence speed and better nonlinear fitting ability.

(2) Analysis of three model errors

In order to better illustrate the effectiveness of IPSO–WNN, the training data sets were tested by WNN, PSO–WNN and IPSO–WNN models, and the experimental results of the three models were obtained. As shown in Table 1, the prediction effects of the three prediction models were analyzed from different experimental results. The evaluation criteria include mean relative error (MRE), mean absolute error (MAE), mean square error (MSE) and fitness (EC) (the degree to which the predicted data fit the true value of WNN), which are used to evaluate the predicted results. The relative error of WNN, PSO–WNN and IPSO–WNN training results was 0.848, 0.437 and 0.321. The training results of IPSO–WNN are shown in Fig. 7.

Table 1 Evaluation results of three models

Experimental model	MRE	MAE	MSE	EC
WNN	0. 848	31.321	0.932	0.788
PSO–WNN	0. 437	22.652	0.732	0.842
IPSO–WNN	0. 321	16.327	0.467	0.951

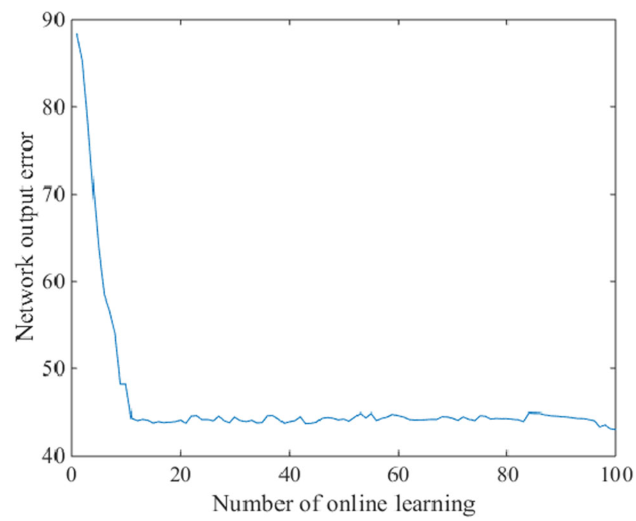

Fig. 7 Training results for IPSO-WNN

The results of the three model trainings are shown in Table 1. It can be concluded that the MAE and MRE obtained by IPSO–WNN are lower than WNN and PSO–WNN, and the MAE decline is 14.994%, thus effectively overcoming the WNN single parameter optimization defects. These results effectively validate the effectiveness of the algorithm on the prediction accuracy when optimizing the WNN initial parameters. The ECSO–WNN has an EC value greater than 0.9. On the contrary, the EC values of WNN and PSO–WNN are less than 0.9, indicating that the optimization of IPSO initial parameters can accelerate the network convergence speed and avoid local optimization, and the curve fitting ability is also improved accordingly. The data in Fig. 7 show that the proposed algorithm model has achieved satisfactory results, and the model is closer to the actual traffic flow than the other two models.

5 Conclusions

In this paper, wavelet neural network is chosen as the model of traffic flow prediction. In the training process of wavelet neural network, the network tends to stay near the local optimal value and no longer search for better values nearby. The problem of finding the probability of finding a global optimal value is reduced. Based on this, an improved particle swarm optimization algorithm is proposed to optimize the wavelet neural network prediction model. The wavelet neural network based on particle swarm optimization is used to predict short-term traffic flow. The conclusion obtained by the experimental verification is that the improved prediction model has a certain improvement in the prediction accuracy and also solves the problem that the wavelet neural network is easy to fall into the local minimum.

Aiming at the slow convergence rate and local optimal problem of WNN prediction algorithm, this paper proposes an improved particle swarm optimization algorithm based on improved particle swarm optimization algorithm. In this way, the global search ability of the particles is improved, so that the optimal value can be quickly found. It overcomes the shortcomings of traditional WNN in optimizing the initial parameters and the relatively low prediction accuracy. The experimental results show that the prediction accuracy of IPSO–WNN is significantly higher than that of WNN and PSO–WNN. Therefore, it is feasible and effective to apply IPSO–WNN to short-term traffic flow prediction and typical chaotic time series prediction.

Short-term urban traffic flow forecasting is an important issue for sustainable urban development. It is a key technology for the development of intelligent transportation. It can effectively alleviate road congestion, ensure smooth road surface and improve road utilization. In this paper, an improved particle swarm optimization algorithm is proposed to optimize the wavelet neural network prediction model, which improves the prediction accuracy of the algorithm. Through the algorithm simulation comparison, the proposed algorithm has higher prediction accuracy, but because the number of training model samples in the optimized neural network is small, and the urban traffic flow prediction is easily affected by the external environment and emergency events, the prediction effect still exists. Where there are uncertainties and deficiencies, there are still many issues that require further study.

Acknowledgements This work was supported by the Innovation and Entrepreneurship Project for Overseas High-level Talents of Shenzhen (KQJSCX20180329191021388), a Research Grant in Shenzhen Polytechnic (No. 6020320004K), National Science Foundation Project: Research on Collaborative Innovation System of the Pearl River-Xijiang River Economic Belt under Supply-side Reform (No. 17XGL006), Wuzhou University Foundation Project: Research on transformation and upgrading of manufacturing industry for the Pearl River-Xijiang River Economic Belt (No. 2017A002).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Wang Y, Lu Z, Qu Y et al (2016) Improving prediction performance of GPS satellite clock bias based on wavelet neural network[J]. GPS Solut 21(2):1–12
2. Duan H, Xiao X, Lei G et al (2017) Hamiltonian cycles decomposition of two cycles’s composite[J]. J Discrete Math Sci Cryptogr 20(5):1241–1250
3. Yang H-F, Dillon TS, Chen YP (2017) Optimized structure of the traffic flow forecasting model with a deep learning approach[J]. IEEE Trans Neural Netw Learn Syst 28(10):2371–2381
4. He W, Zhong C, Sotelo MA et al (2017) Short-term vessel traffic flow forecasting by using an improved Kalman model[J]. Clust Comput 10:1–10
5. Zhou T, Han G, Xu X et al (2018) A learning-based multimodel integrated framework for dynamic traffic flow forecasting[J]. Neural Process Lett 5:1–24
6. Zou Weidong, Xia Yuanqing (2019) Back propagation bidirectional extreme learning machine for traffic flow time series prediction. Neural Comput Appl 31(11):7401–7414
7. Liu Mei, Agarwal Jitendra, Blockley David (2016) Vulnerability of road networks[J]. Civil Eng Environ Syst 33(2):1–29
8. Zhang XB, Wang JZ, Zhang KQ (2017) Short-term electric load forecasting based on singular spectrum analysis and support vector machine optimized by Cuckoo search algorithm[J]. Electr Power Syst Res 146(2):270–285
9. Dong Shi, Li Ruixuan (2019) Traffic identification method based on multiple probabilistic neural network model. Neural Comput Appl 31(2):473–487
10. Lun-hui XU, Xiang-xue W (2016) Application of SVD-based optimized robust UKF algorithm to estimation of short-term traffic flow state[J]. J Southwest Jiaotong Univ 44(12):44–52
11. Chen D, Hu M, Zhang H et al (2016) Short-term traffic flow prediction of airspace sectors based on bayesian estimation theory[J]. Xinan Jiaotong Daxue Xuebao/J Southwest Jiaotong Univ 51(4):807–814
12. Shang Q, Yang Z, Zhang W et al (2016) Short-term traffic flow prediction based on singular spectrum analysis and CKF-LSSVM[J]. J Jilin Univ 46(6):1792–1798
13. Y Xing, X-J Ban, R Liu (2018) A short-term traffic flow prediction method based on kernel extreme learning machine[C]. In: 2018 IEEE international conference on big data and smart computing (BigComp). IEEE
14. Huang Min-Liang (2015) Intersection traffic flow forecasting based on ν-GSVR with a new hybrid evolutionary algorithm[J]. Neurocomputing 147(1):343–349
15. Hu W, Wang H, Qiu Z, Nie C, Yan L (2018) A quantum particle swarm optimization driven urban traffic light scheduling model. Neural Comput Appl 29(3):901–911
16. Xueneng SU, Tianqi LIU, Hongqian CAO et al (2017) A multiple distributed BP neural networks approach for short-term load forecasting based on hadoop framework[J]. Proc CSEE 37(17):4966–4973

17. Qiao Y, Zhao B, Fang C et al (2017) Study of ramp control method for urban expressways using improvised ALINEA algorithm[J]. J Southwest Jiaotong Univ 52(5):1001–1007
18. Abdi J, Moshiri B (2015) Application of temporal difference learning rules in short-term traffic flow prediction[J]. Expert Syst 32(1):49–64
19. Yang H, Dillon TS, Chang E et al (2019) Optimized configuration of exponential smoothing and extreme learning machine for traffic flow forecasting[J]. IEEE Trans Ind Inf 15(1):23–34
20. Tang L, Zhao Y, Cabrera J et al (2018) Forecasting short-term passenger flow: an empirical study on Shenzhen metro[J]. IEEE Trans Intell Transp Syst 22(99):1–10