Research on traffic flow prediction in the big data environment based on the improved RBF neural network

Dawei Chen

Abstract: **This paper proposes an optimized prediction algorithm of radial basis function neural network based on an improved artificial bee colony (ABC) algorithm in the big data environment. The algorithm firstly uses crossover and mutation operators of the differential evolution algorithm to replace the search strategy of employed bees in ABC algorithm, then improves the search strategy of onlookers in ABC algorithm to produce an optimal candidate food source near the population. The algorithm can better balance local and global searching capability. To verify the efficiency of this algorithm in the big data environment, apply it to Lozi and Tent chaotic time series and measured traffic flow time series, and then compare it with K-nearest neighbor (KNN) model, radial basis function(RBF) neural network model, improved back propagation (IBP) neural network model and radial basis function neural network based on cloud genetic algorithm (CARBF) model. The experimental results indicate that the accuracy of prediction for Lozi and Tent chaotic time series and the measured traffic flow improves greatly in the big data environment using the proposed algorithm, which proves the effectiveness of the proposed algorithm in predicting traffic flow time series.**

***Key words*: artificial bee colony algorithm; differential evolution; RBF neural network; traffic flow prediction; big data environment**

I. Introduction

With the general increase of living standards and rapidly national economic development, there are increasingly higher requirements for transportation industries. Therefore, an accurate access to obtain the related information is urgently needed. Due to the growing population and much wider automobile usage, traffic flow increases and traffic congestion has become a major problem which every nation in the world has to face. Therefore, to effectively solve the traffic-related problems, a technology called intelligent transportation system (ITS) emerges. ITS is an integrated transportation management system that integrates the technologies of electronic information, artificial intelligence, geographic information, global positioning, image analysis, communications engineering, etc. It has been regarded as a new and effective method in solving transportation problems, and presents unique performance which is not owned by traditional methods in solving traffic congestion as well as reducing accidents and traffic pollutions. Traffic guidance and control play an important role in ITS [1,2]. As an essential part of ITS, the key theory is traffic prediction techniques [3], and an accurate traffic data stream prediction which ensures the good operation of traffic flow guidance system. To achieve the traffic flow guidance and dredging as well as traffic congestion reduction, Traffic data stream prediction has become an important research topic [4,5]. Characteristics of traffic flow such as highly nonlinear, time-varying and randomness [6,7] make it difficult to be predicted. Short-term rather than long-term prediction of traffic flow is possible, and the key of the study is to improve accuracy of short-term prediction.

Urban traffic flow systems have obvious chaotic characteristics. Based on this idea, the prediction is to approximately restore the original system by constructing a nonlinear mapping to restore the original system, and the nonlinear mapping is the predictive model to be established. Many researches have been conducted in this area so far, and many methods have been employed, such as Kalman state space filtering models [8], support vector machine model [9], autoregressive integrated moving average [10], Type-2 Fuzzy Logic Approach [11], K-nearest Neighbor Model [12], neuro-fuzzy systems [13], binary neural network [14,15], Bayesian network model [16,17], back propagation neural network model [18,19], radial basis function neural network model [20,21], and some portfolio models [22,23]. In these prediction methods, neural networks become the researched focus of many experts and scholars due to its unique characteristics such as massive parallel structure and distributed storage. It not only has powerful function approximation and pattern classification ability, but also has good self organization, self adaptability and incorrectness tolerance. However, the selection of center vectors and width values of hidden nodes of neural networks has a greater impact on the learning and generalization ability of the network in practice. And improper preferences will seriously affect the performance of the network [24].

ABC algorithm is a random algorithm proposed by Karaboga [25,26]. Due to its simple structure and smaller control parameter, high convergence speed and easy implementation, the ABC algorithm can be applied for the optimization of the combination of parameters of a neural network. However, as an optimization algorithm, ABC algorithm has slow convergence speed and is easily prone to premature convergence as some other optimization algorithms. So many experts and scholars have done much work to overcome these shortcomings [27-30].

In this paper, an improved ABC algorithm, namely RBF (IABCRBF) neural network is proposed. The algorithm firstly uses differential evolution(DE) algorithm to replace the searching strategy of employed bees in an ABC

Dawei Chen is with the School of Transportation, Southeast University, Nanjing 210096, China. (email: dw_chen@seu.edu.cn; chendaweiteacher@163.com)

algorithm, and then improves the searching strategies of onlookers to generate a candidate food source near the population optimal solution. To achieve better network initial weights and thresholds by the improved colony algorithm, which uses an ABC algorithm to compensate for the random defects of the RBF neural network connection weight and threshold selection. The results of simulation show that predictions based on this algorithm are more accurate using this method to conduct the modeling and predictive analysis for measured traffic flow in the big data environment.

II. Radial basis function neural network

Radial basis function network is a neural network with three layers [22]. The input layer, in which the size of nodes is mainly determined by the input vector x dimension; the hidden layer, which is connected to the input node, where the output data dimension equals the size of nodes. The RBF neural network completes the mapping $f: R^n \to R^m$, and the mathematical expression is:

$$f_i(x) = \theta_0 + \sum_{j=1}^{h} \theta_{ij}\varphi(\|x - c_j\|) \quad i = 1, 2, \cdots, m \tag{1}$$

In the formula, $\phi(\bullet)$ complete the non-linear transformation of $R^n \to R$ as the radial basis function, $\|\bullet\|$ represents norm; $\theta_{ij}(1 \le i \le m, 1 \le j \le h)$ represents the output layer connection weights of the network, θ_0 is the polarization; c_j is the center points of the hidden layer. This paper selects the Gaussian function as the radial basis function, the equation (2) is shown as follows:

$$\phi(\|x - c_j\|) = \exp\left(\frac{-\|x - c_j\|^2}{\beta^2}\right) \tag{2}$$

In the equation: β is called width value. The accuracy of network is mainly determined by the central node, base width vector, threshold and network value, in which the approximation performance of the network can be improved by optimizing the network parameters.

III. Improved ABC algorithms

A. *ABC algorithm*

The capacity of a single bee is very weak, but a bee colony shows a strong collaboration capability to find high-quality honey sources effectively. Inspired by this, bees algorithm can be divided into two algorithms: the algorithm based on the honey bee breeding behavior and the algorithm based on honey gathering behavior according to its biological mechanism. Wherein, the ABC algorithm based on honey gathering behavior is easy to operate and is mainly used for neural network training because it has fewer parameters. In ABC algorithm, the process of an individual honey bee gathering honey is similar to that of searching for the optimal solution for the problems to be optimized in all the possible solutions. The quality of the food source is measured with fitness function. The number of solutions in the entire space is represented by SN, and the value of SN equals the number of employed bees or onlookers. When the artificial bee colony algorithm solves the optimization problems, the food source is on behalf of a possible solution. The process of bees collecting honey is the process that the optimization problem is searching for the optimal solution. The adaptation values of the optimization problem affect the quality of a food source, in which the higher the adaptation value, the better the food source is. Using $x_i = (x_{i1}, x_{i2}, \cdots, x_{iD})$ denotes the i food source, SN equals the number of employed bees or onlookers with $i = 1, 2, \cdots, SN$, D is the dimension of the searching space.

The artificial bee colony algorithm randomly generates *SN* solutions, and the bee colony makes a circulated search for all food sources. Firstly, the employed bee generates a candidate food source, and then it compares the candidate food source with the previous food source. It replaces the previous food source with the candidate food source if the value of the candidate food source is better than the previous food source, otherwise it maintains the previous food source. After that, the employed bee returns to the dance area to convey the food source information by swing dancing to the onlookers, then the onlookers choose the food source according to a certain probability through the obtained information. The higher the adaptation value of the food source is, the higher the probability can select. After the onlookers selecting the food source, they compare the candidate selected food sources with the selected one, then the better one will be reserved. The ABC algorithm finds the optimal solution eventually through the mentioned repeated cycles. When the food source of a bee individual is not changed after a limited number of cycles, it will be abandoned and become a scout to find new food sources.

According to formula (3), the employed bees and onlookers generate a candidate food source.

$$v_{ij} = x_{ij} + (x_{ij} - x_{kj}) * rand(-1,1) \tag{3}$$

In the above equation, v_{ij} is the candidate food source generated, $k \in \{1,2,..., SN\}$ and $j \in \{1,2,..., D\}$, $k \ne i$ and $rand(-1,1)$ is the random number. With an increasing number of iterations, the distance between $(x_{ij} - x_{kj})$ becomes shorter and shorter, the search space also becomes much smaller, and the search steps become less and less.

The onlookers select food sources according to the probability of formula (4).

$$P_i = \frac{fit_i}{\sum_{n=1}^{SN} fit_n} \tag{4}$$

In the formula, fit_i is the fitness value of the i food source.

In ABC algorithm, when the food source of a certain individual bee has not been updated after a limited number of times of continuous cycles, the bee must give up the food source, and transfer to scouts which generate a new food source by formula (5)

$$x_{ij} = x_{\min,j} + (x_{\max,j} - x_{\min,j}) * rand(0,1) \tag{5}$$

In the formula, x_{ij} is the component of j dimension of the new food source, $j \in \{1,2,...,D\}$ and $rand(0,1)$ are the random numbers of $(0,1)$, $x_{\max,j}$ and $x_{\min,j}$ are the maximum and minimum values of variable of j dimension.

B. Standard differential evolution algorithm

DE algorithm has the features of remembering optimal solution of individuals and sharing information within the group . Firstly, a set of random initialization populations is set as $X(0) = x_1(0), x_2(0),.., x_{NP}(0)]$. NP is the population size, D is the population dimension. After operations, the g generation of individual evolves to $x_i(g) = [x_{i,1}(g), x_{i,2}(g),..., x_{i,D}(g)]$. The basic idea of the algorithm is to add the difference vector obtained by the subtraction of two different random individuals of a parent to a third randomly selected individual to produce a variation of the individual. Then follow a probability, where parent individuals and mutation individuals generate a new individual through the crossover operation. Then make a selection operation between the parent individuals and the new individual to choose an individual of better fitness as offspring according to the function values of the fitness. The specific steps are as follows:

Initialized Population

In the decision space X, an initial vector is generated randomly shown as equation (6):

$$x_i(0) = x_i^l + rand(0,1) * (x_i^u - x_i^l) \tag{6}$$

Mutation Operation

The difference vector of DE algorithm is multiplied by a zoom factor, and then a vector synthesis is made with the base vectors. The typical mutation operators have the following three types, which are shown in equation (7), (8) and (9) respectively:

DE/*rand*/1 mutation operator:

$$v_i(g+1) = x_{r1}(g) + F * (x_{r3}(g) - x_{r2}(g)) \tag{7}$$

DE/*rand*/2 mutation operator:

$$v_i(g+1) = x_{r1}(g) + F * (x_{r3}(g) - x_{r2}(g)) + F * (x_{r5}(g) - x_{r4}(g)) \tag{8}$$

The third mutation strategy DE/*best*/1:

$$v_i(g+1) = x_{best}(g) + F * (x_{r3}(g) - x_{r2}(g)) \tag{9}$$

Among equations, $v_i(g+1)$ is the variation vector of the generation of g +1, F is a zoom factor, $r1$, $r2$, $r3$, $r4$, $r5$ $\in [1,2,...,NP]$ are the integers which differ from each other and not equal to integer i, $x_{r1}(g)$ is the basis vector of parent generation, $x_{best}(g)$ represents the best individual in the population of g generation, $(x_{r3}(g) - x_{r2}(g))$ and $(x_{r5}(g) - x_{r4}(g))$ are difference vectors of parent generation.

Crossover Operation

For each individual vector $x_i(g)$ in the population of g generation, a crossover operation is done with mutation individual $v_i(g+1)$ to generate a new individual $u_i(g+1)$ to increase diversity of the individual population, the formula is shown as formula (10):

$$u_i(g+1) = \begin{cases} v_{ij}(g+1) & if\,(rand(j) \le CR)\ or\ j = k \\ x_{ij}(g+1) & otherwise \end{cases} \tag{10}$$

In the formula, CR is the crossover probability factor, $x_{ij}(g)$ represents the component of j dimension of the i individual of the population of the g generation, $rand(j) \in [0,1]$ is the corresponding random number of j dimension. k is the corresponding coefficient of individual i, which is usually an integer randomly selected from the sequence $[1,2,...,D]$ to ensure that at least one-dimensional components in $u_i(g+1)$ are from individual variation vector $v_i(g+1)$.

Selection Operation

Standard differential evolution adopts a greedy selection strategy to operate on a test vector $u_i(g+1)$ current population target vector $x_i(g)$. Through evaluation of the adaptation degree of both vectors, the optimum individual is selected into the next generation of search. The formula is shown in formula (11).

$$x_i(g+1) = \begin{cases} u_i(g+1) & if\ f(x_i(g)) > f(u_i(g+1)) \\ x_i(g) & otherwise \end{cases} \tag{11}$$

Where , f is the fitness function, $f(u_i(g+1))$ is the corresponding fitness value of test individual $u_i(g+1)$.

C. Improved Artificial Bee Colony Algorithm

The paper proposes a optimization algorithm based on DE and ABC algorithms. The algorithm uses three typical operators in a differential evolution algorithm as the evolution operators in which scouts use for the position selection of nectar source. Onlookers in ABC algorithm are presented in formula (12) and formula (15), respectively.

$$v_{ij} = \begin{cases} x_{bestj} + F \cdot (x_{r1,j} - x_{r2,j}) & if\ rand < CR\ or\ j = r \\ x_{ij} & otherwise \end{cases} \tag{12}$$

In the formula, r_1 and r_2 are two different random integers in $\{1,2,...,SN\}$ and neither of them equals i . $j \in \{1,2,..., D\}$ and x_{best} are the optimal variables searched so far. F is the zoom factor which adopts an adaptive nonlinear decreasing strategy, and its formula is shown in formula (13), CR is the crossover probability factor shown in formula (14), and r is a random integer in $\{1,2,..., D\}$.

$$F = F_{\min} + (F_{\max} - F_{\min}) * \frac{\sqrt{g_{\max}^2 - g^2}}{g_{\max}} \tag{13}$$

$$CR = CR_{\min} + \frac{g}{g_{\max}} * (CR_{\max} - CR_{\min}) \tag{14}$$

In the formula, $g_{\max}$ represents the maximum iteration number, g represents the current iteration number, $F_{\min}$ and $F_{\max}$ are the minimum and maximum mutation operators respectively. Take $F_{\min} = 0.6$ and $F_{\max} = 0.9$. $CR_{\min}$ and

$CR_{\max}$ as the minimum and maximum crossover operators respectively, take $CR_{\min}=0.7$ and $CR_{\max}=0.9$.

F and CR can adaptively determine the mutation rateby using an adaptive nonlinear decreasing operator. Therefore, the algorithm can maintain the diversity of individual at the beginning and avoid individuals falling into a premature convergence phenomenon. The mutation rate is reduced in the later phase of the algorithm to avoid the optimal solution being damaged, we increase the possibility of the optimal solution in global search.

The improved searching strategy of onlookers is shown as formula (15) :

$$v_{ij}=x_{bestj}+rand(-1,1)*(x_{bestj}-x_{kj}) \quad (15)$$

In this algorithm, formula (12) inherits the advantages of DE algorithm and has a strong global search capability. Inheriting the advantages of ABC algorithm, the formula (15) has a stronger capacity than the original search strategy of ABC algorithm in local search because it has the optimal solution to guide the population search track and can generate a new candidate food source in the vicinity of the optimal solution.

IV. Algorithm design of optimizing RBF neural network by improved ABC algorithm

Algorithm steps are as follows:

Step1: Encoding. The individual coding method of ABC algorithm is real-number encoding, which comprises the weight value and threshold value of network.

Step2: Initialize the parameters of an ABC algorithm, including the population size SN, the employed bee and onlooker number, the scout cycle number limit and maximum number of iterations $g_{\max}$, preset target accuracy value ε, $F_{\min}$ and $F_{\max}$ are the minimum and maximum mutation operators respectively, $CR_{\min}$ and $CR_{\max}$ are the minimum and maximum crossover operators respectively.

Step3: Set the initial number of iterations $g=0$, generate SN food sources randomly, and the components of each food source are random numbers of interval $(1,1)$.

Step4: Calculate the fitness function value.

Step5: Onlookers produce new solutions $x_{i,new}$ based on a certain searching percentage, and calculate the fitness value of $x_{i,new}$. If $x_{i,new}$ is better than x_{ij}, then replaced x_{ij} with $x_{i,new}$; otherwise, remain x_{ij} unchanged, the count of x_{ij} is incremented by 1 .

Step6: Calculate the corresponding probability of x_{ij} - P_i according to formula (12)

Step7: Onlookers choose food sources in accordance with the probability of P_i, select the preferred food source in accordance with a certain proportion according to formula (15) and produce a new solution $x_{i,new}$, if $x_{i,new}$ is better than x_{ij}, replace x_{ij} by $x_{i,new}$,otherwise remain x_{ij} unchanged, the count of x_{ij} increases by 1.

Step 8: If the food source of a solution x_{ij} is still not updated, the individual should give up the food sources and turn into a scout. Scouts produce a new food source in accordance with formula (5) .

Step9: After one iteration is completed, record the optimal solution that is so far searched.

Step10: Judge whether g reaches the maximum number of iterations $g_{\max}$ or whether objective value is less than preset target accuracy value ε, and if it satisfies the iteration termination, output the optimal solution, otherwise, $g=g+1$, and returns to Step4.

V. Simulation experiments

KNN model[12], RBF neural network model[21], IBP neural network model [18], CARBF model [22] have successfully been applied to traffic flow prediction. So in te paper we applied the IABCRBF neural network model to the Lozi, Tent chaotic time series and the measured traffic flow by comparing it with KNN model, RBF neural network model, IBP neural network model, CARBF model respectively to verify the algorithm's performance.

A. Predictive evaluation criteria

Experimental error evaluationmainly uses Mean Absolute Error (MAE) and Proportional Error (PERR) which are shown as in formula (16) - (17) respectively:

$$MAE=\frac{1}{n}\sum_{t=1}^{n}\left|x(t)-x'(t)\right| \quad (16)$$

$$PERR=\frac{\sum_{t=1}^{n}\left[x(t)-x'(t)\right]^2}{\sum_{t=1}^{n}x^2(t)} \quad (17)$$

In the formulas: n is the prediction sample number; $x'(t)$ is the predicted value and $x(t)$ is the expected value.

B. Chaotic time series prediction

Lozi system[31] is a typical representation of discrete chaotic system which has randomness and ergodicity feature and its mathematical model is shown in formula (18):

$$\begin{cases} x_{n+1}=1-a\,|\,x_n\,|+y_n \\ y_{n+1}=bx_n \end{cases} \quad (18)$$

When $a=1.74$ and b=0.34, the system is in a chaotic state. Fig.1 shows the perception mapping.

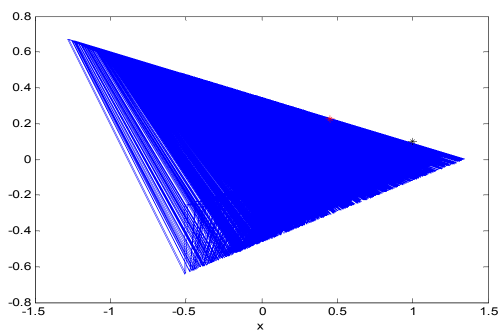
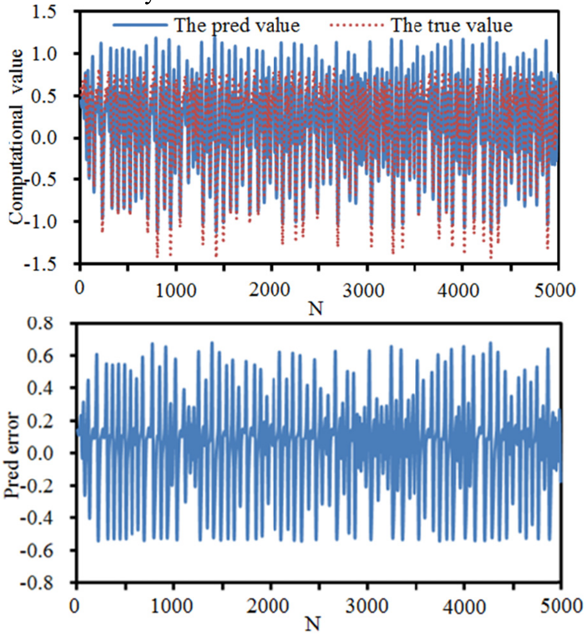


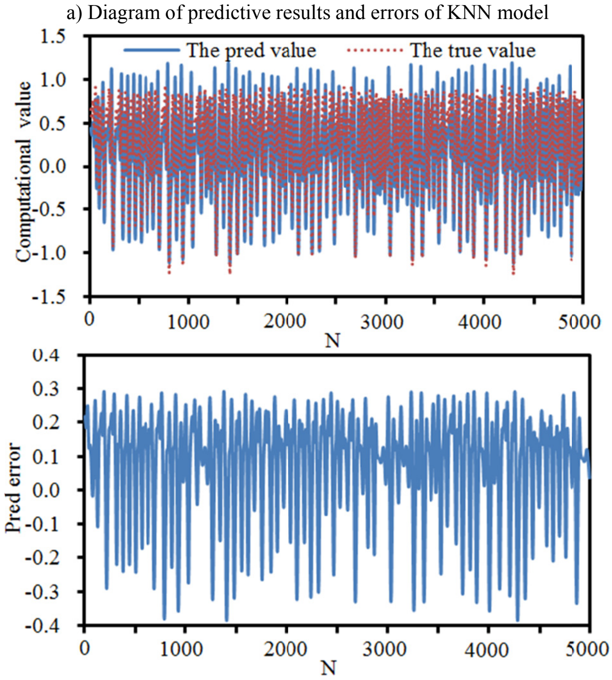
Fig. 1. Lozi mapping

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2017.2682855, IEEE Transactions on Industrial Informatics

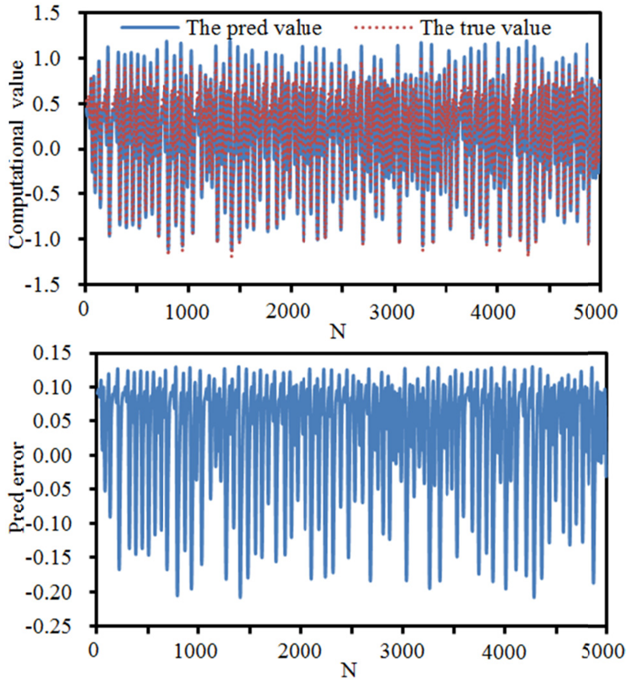
In the experiment, the first 20000 data of the time series is taken as training samples, and the remaining 5000 data is taken as test samples. Fig.2 shows the computational results of KNN model, RBF neural network model, IBP neural network model, CARBF neural network model and IABCRBF neural network model in Lozi system. Table.l shows the prediction errors and run time of five kinds of models of Lozi system.



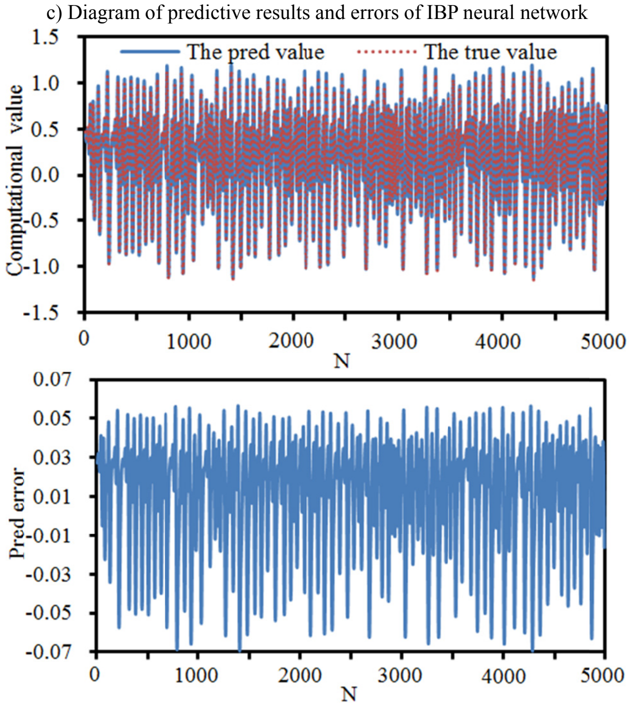
a) Diagram of predictive results and errors of KNN model



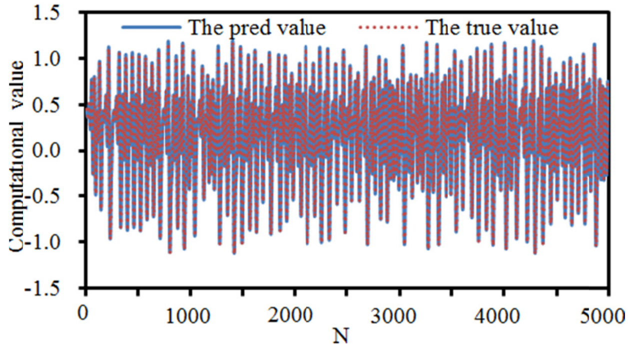
b) Diagram of predictive results and errors of RBF neural network



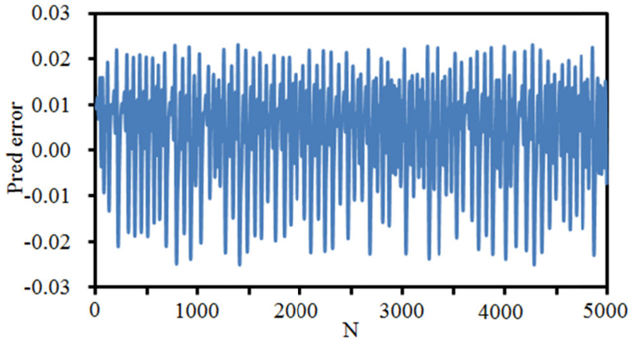
c) Diagram of predictive results and errors of IBP neural network



d) Diagram of predictive results and errors of CARBF neural network



This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2017.2682855, IEEE Transactions on Industrial Informatics



e) Diagram of predictive results and errors of IABCRBF neural network
Fig. 2. Diagram of predictive results of Lozi chaotic time sequence

Table. 1.predictive errors and run time of Lozi model system

Prediction model	MAE	PERR	run time
KNN	0.2866	1.1974	61s
RBF	0.0463	0.2243	15s
IBP	0.0387	0.2089	13s
CARBF	0.0151	0.1036	9s
IABCRBF	0.0025	0.0869	6s

Tent has a simple mapping structure, uniform probability density and power spectral density[32]. It's the one-dimensional mapping of piecewise linearity, and therefore suits the computational processing of data sequences of a large magnitude. Formula (19) is the mathematical model of Tent system.

$$x_{n+1} = a - (a+1)|x_n| \quad a \in (0,1) \tag{19}$$

The system is in a chaotic state, and Fig.3 shows the chaotic mapping.

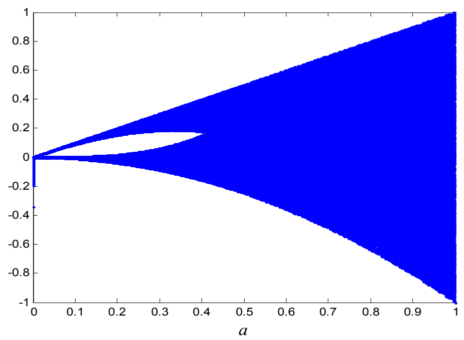
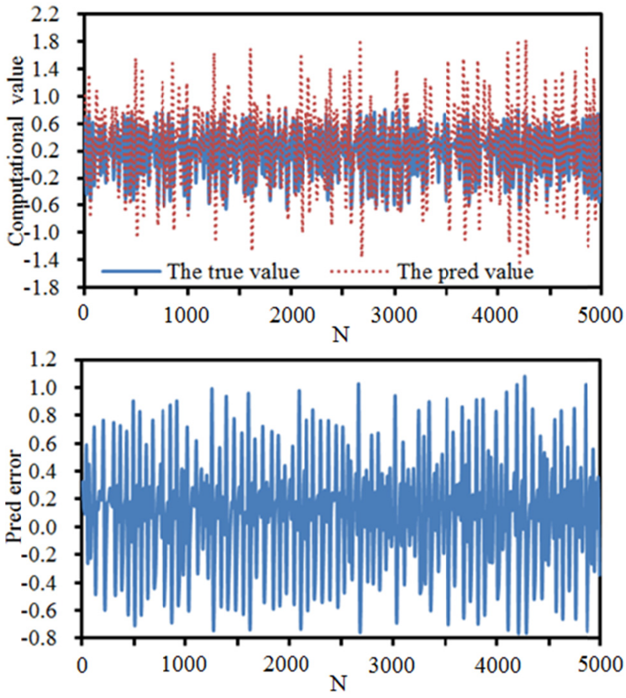
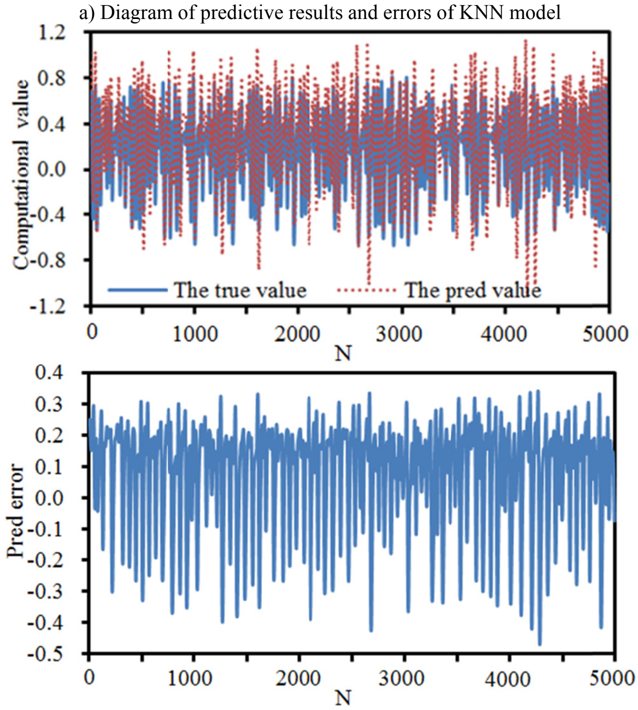


Fig. 3. Tent mapping

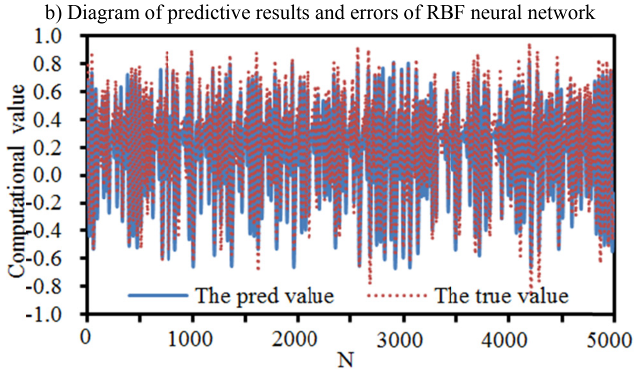
In the experiments, the first 20000 data of the time series is taken as training samples, and the remaining 5000 data is taken as test samples. Fig.4 shows the computational results of the KNN model, RBF neural network model, IBP neural network model, CARBF neural network model and IABCRBF neural network model in Tent system. Table.2 shows the prediction errors and run time of five kinds of models of Tent system.



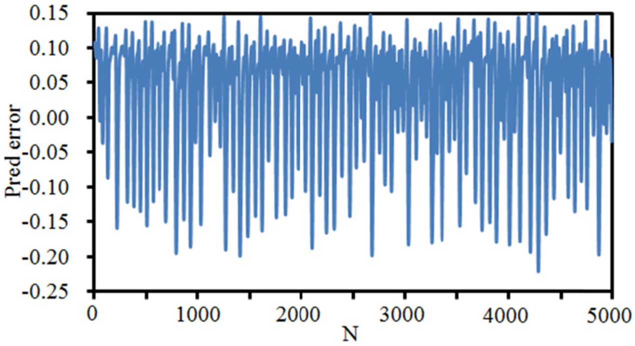
a) Diagram of predictive results and errors of KNN model



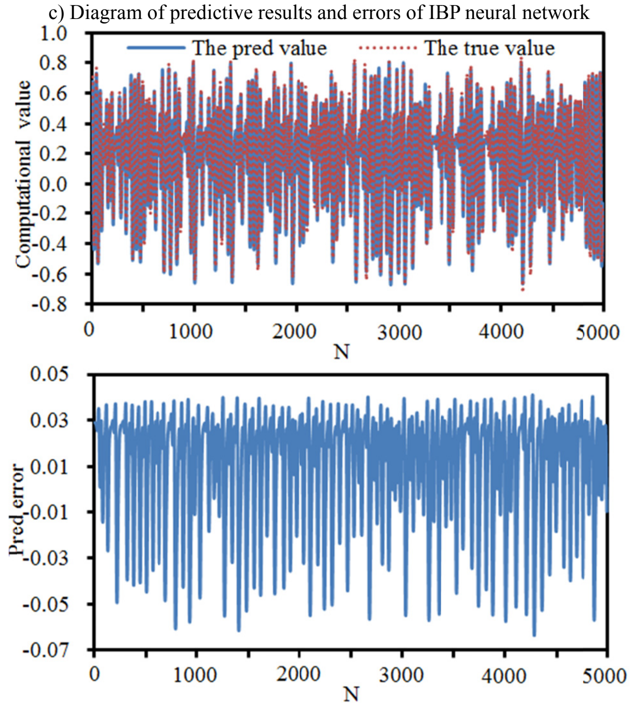
b) Diagram of predictive results and errors of RBF neural network



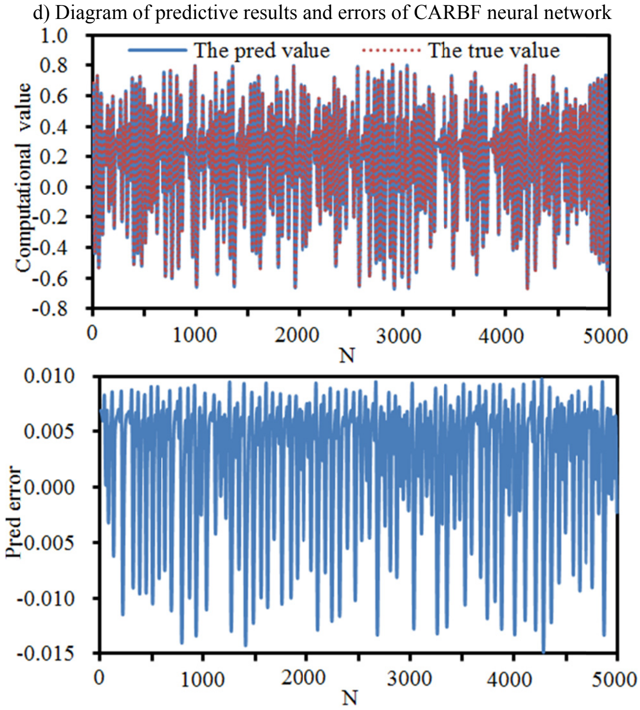
This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2017.2682855, IEEE Transactions on Industrial Informatics



c) Diagram of predictive results and errors of IBP neural network



d) Diagram of predictive results and errors of CARBF neural network



e) Diagram of predictive results and errors of IABCRBF neural network

Fig. 4. Diagram of predictive results of Tent chaotic time sequence

Table.2: Predictive errors and run time of Tent model system

Prediction model	MAE	PERR	run time
KNN	0.7231	1.1358	57s
RBF	0.0132	0.2698	16s
IBP	0.0112	0.2387	14s
CARBF	0.0079	0.1897	9s
IABCRBF	0.0045	0.1351	6s

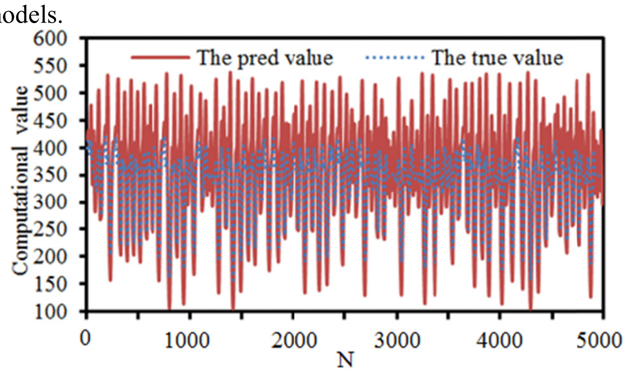
C. Real traffic flow time series prediction

In unit time, the vehicle number through a certain observation point or section is called traffic flow, known as traffic stream or traffic volume. Its statistical method is shown as formula (20):

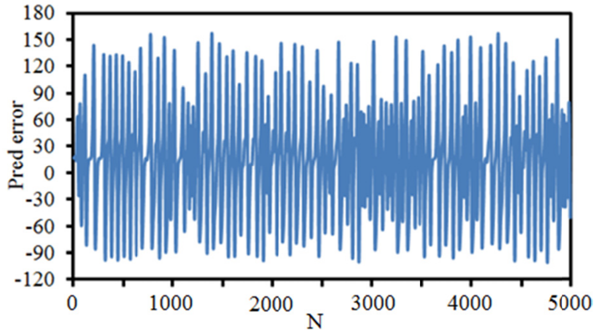
$$V = \frac{60N}{T} \tag{20}$$

In the formula: V represents the traffic flow of a certain time t, T is the statistics time interval, N is the vehicles number in the time period T.

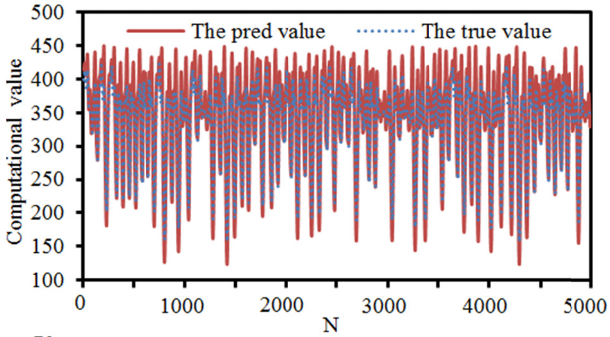
The experimental data of traffic flow is measured in data from Wanda intersection, Lihua Road, Xiamen, in which the total observation time is 250 days, a total of 6000 hours. Take 18 minutes as an interval, the vehicles number for each period is recorded and the traffic flow values of each time point are calculated according to formula (20) respectively, comprising with a total of 20000 sets of data. The first 15000 groups of the traffic flow sequence are training samples, and the followed 5000 groups are taken as the test samples. Using KNN model, RBF neural network model, IBP neural network model, CARBF neural network model and IABCRBF neural network model respectively to predict the results, Fig.5 shows the computational results of the KNN model, RBF neural network model, IBP neural network model, CARBF neural network model and IABCRBF neural network model. Table.3 shows the prediction errors of five kinds of prediction models.

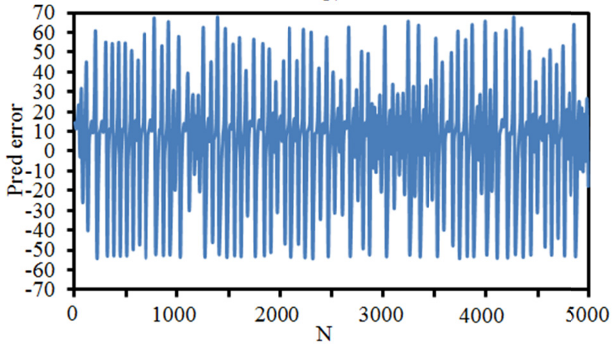


This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TII.2017.2682855, IEEE Transactions on Industrial Informatics

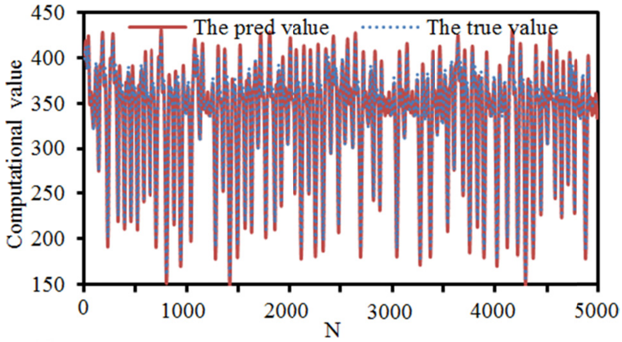


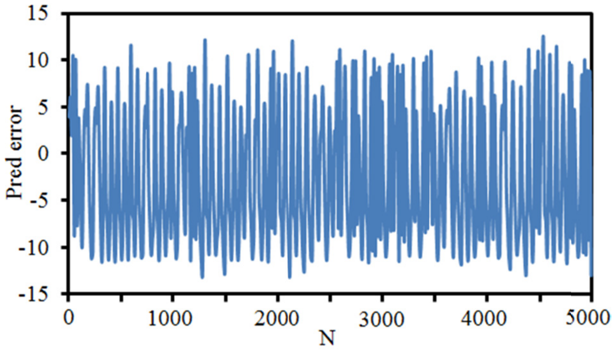
a) Diagram of predictive results and errors of KNN model



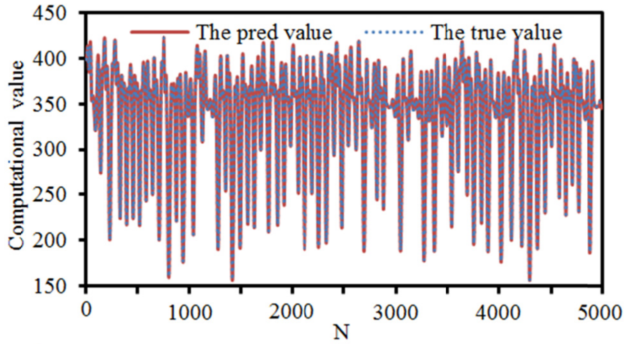


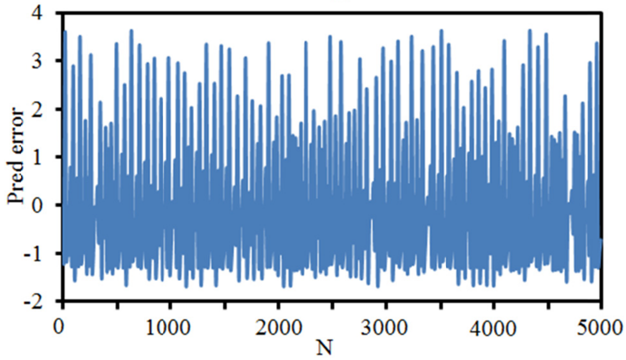
b) Diagram of predictive results and errors of RBF neural network



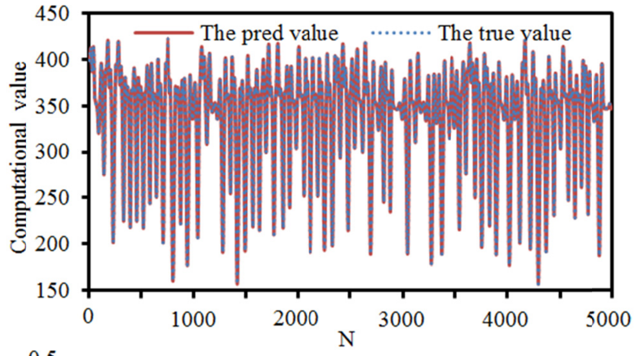


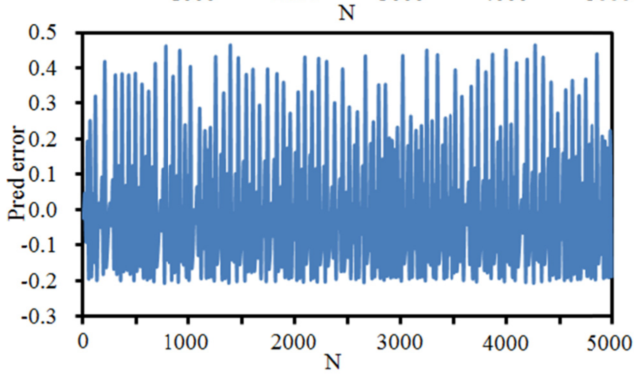
c) Diagram of predictive results and errors of IBP neural network





d) Diagram of predictive results and errors of CARBF neural network





e) Diagram of predictive results and errors of IABCRBF neural network

Fig.5. Predictive results of the measured traffic flow

Table. 3. Predictive errors and run time of the measured traffic flow

Prediction model	MAE	PERR	run time
KNN	0.1364	1.6853	65s
RBF	0.0265	0.4068	17s
IBP	0.0246	0.3962	15s
CARBF	0.0152	0.2976	10s
IABCRBF	0.0087	0.1524	6s

From Figures 2, 4, 5 and Tables 1, 2, 3, we can easily find that the prediction results of five kinds of prediction models can predict the trends of Lozi and Tent chaotic time series and the measured traffic flow time series. In terms of the prediction effect, the KNN model performs the worst among these models because its prediction effect is far from the real data, while RBF neural network, IBP neural network model and CARBF neural network model reflect basically the real data, and IBP neural network model and CARBF neural network model are closer to the real data than RBF neural network. CARBF neural network model is slightly better than IBP neural network model, while IABCRBF neural network model is the best among five kinds of prediction models whose prediction results are basically consistent with the real data. In terms of the running time, the KNN model has the longest running time, and RBF neural network, IBP neural network model and CARBF neural network model follow in turn, while the IABCRBF neural network model has the shortest running time. Therefore, the IABCRBF neural network model proposed in this paper is feasible and effective in chaotic time series prediction and traffic flow prediction.

VI. Conclusions

Due to the uncertainty of traffic flow, a traditional linear model is difficult to predict results to meet the high real-time demands for traffic guidance. This study adopts ABC algorithm to optimize weight and threshold value of RBF neural network and analyzes the nonlinear time series. With the application of this algorithm to predict two kinds of chaotic systems and traffic flow measurement systems, and comparing the prediction accuracy with the KNN model, RBF neural network model, IBP neural network model and CARB Fneural network model, the results indicate that the proposed model has higher prediction accuracy and can reflect the traffic flow change law more accurately, therefore has good prospects in traffic data stream prediction.

References:

[1] G. He and S. Ma,“A Study on the Short-term Prediction of Traffic Volume based on Wavelet Analysis,”*IEEE International Conference on Intelligent Transportation Systems*,2002, pp. 731-735.

[2] L. Zhang,“VirtualClock: A New Traffic Control Algorithm for Packet-Switched Networks,” *ACM Transactions on Computer Systems*, Vol. 9, No. 2, 1991, pp. 101-124.

[3] H. Tan, Y. Wu, B. Shen, et al,“Short-Term Traffic Prediction Based on Dynamic Tensor Completion,”*IEEE Transactions on Intelligent Transportation Systems*. 2016, pp. 1-11.

[4] K. Guan, B. Ai, A. Fricke, et al,“Excess Propagation Loss Modeling of Semiclosed Obstacles for Intelligent Transportation System,”*IEEE Transactions on Intelligent Transportation Systems*. 2016, pp. 1-11.

[5] A. Salamanis, P. Meladianos, D. Kehagias， et al,“Evaluating the Effect of Time Series Segmentation on STARIMA-Based Traffic Prediction Model,”*IEEE International Conference on Intelligent Transportation systems*.2015.

[6] G.Marfia, M.Roccetti,“Vehicular Congestion Detection and Short-Term Forecasting: A New Model With Results,”*IEEE Transactions on Vehicular Technology*, Vol. 60, No. 7, 2011, pp. 2936-2948.

[7] K. Y. Chan, T. S. Dillon, J. Singh, et al,“Neural-Network-Based Models for Short-Term Traffic Flow Forecasting Using a Hybrid Exponential Smoothing and Levenberg–Marquardt Algorithm,” *IEEE Transactions on Intelligent Transportation Systems*.Vol. 13, No. 2, 2012, pp. 644-654.

[8] B. Ait-El-Fquih and I. Hoteit. “Fast Kalman-Like Filtering for Large-Dimensional Linear and Gaussian State-Space Models,” *IEEE Transactions on Signal Processing*, Vol. 63, No. 21, 2015, pp.1-1.

[9] Y. Zhang and Y. Liu,“Data Imputation Using Least Squares Support Vector Machines in Urban Arterial Streets,” *IEEE Signal Processing Letters*, Vol. 16, No. 2, 2009, pp. 414-417.

[10] A. Guin,“Travel Time Prediction Using a Seasonal Autoregressive Integrated Moving Average Time Series Model,”*IEEE Intelligent Transportation Systems Conference*, 2006, pp. 493-498

[11] L. Li, W. H. Lin, H. Liu,“Type-2 fuzzy logic approach for short-term traffic forecasting,”*IEE Proceedings - Intelligent Transport Systems*, Vol. 153, No. 1, 2006, pp. 33-40.

[12] H. A. Fayed, A. F. Atiya,“A novel template reduction approach for the K-nearest neighbor method,”*IEEE Transactions on Neural Networks*, Vol. 20, No. 5, 2009, pp. 890-896.

[13] J. Perez, A. Gajate, V. Milanes, et al,“Design and implementation of a neuro-fuzzy system for longitudinal control of autonomous vehicles,”*IEEE International Conference on Fuzzy Systems*, 2010, pp.1-6.

[14] N. Messai, P. Thomas, D. Lefebvre, et al, “A neural network approach for freeway traffic flow prediction,”*IEEE Conference on Control Applications*, Vol. 2, No. 2, 2002, pp. 984-989.

[15] M. Kam, J. C. Chow, R. Fischl,“Design of the fully connected binary neural network via linear programming,”*IEEE International Symposium on Circuits & Systems*, Vol. 2, No. 2, 1990, pp. 1094-1097.

[16] S. Sun, C. Zhang, G. Yu,“A bayesian network approach to traffic flow forecasting,”*IEEE Transactions on Intelligent Transportation Systems*. Vol. 7, No. 1, 2006, pp. 124-132.

[17] G. Min, A. Pahwa, S. Das,“Bayesian Network Model With Monte Carlo Simulations for Analysis of Animal-Related Outages in Overhead Distribution Systems,” *IEEE Transactions on Power Systems*, Vol. 26, No. 3, 2011, pp. 1618-1624.

[18] J. X. Xu, J. S. Lim,“A new Evolutionary Neural Network for forecasting net flow of a car sharing system,”*IEEE Congress on Evolutionary Computation*, 2007, pp. 1670 - 1676.

[19] C. Li, Y. Xie, H. Zhang, et al, “Dynamic Division about Traffic Control Sub-area Based on Back Propagation Neural Network,”*ACM International Conference on Intelligent Human-machine Systems &Cybernetics*,Vol. 2, 2010, pp. 22-25.

[20] H. Gao, J. Zhao, L.Jia,"Short-term Traffic Flow Forecasting Model of Elman Neural Network Based on Dissimilation Particle Swarm Optimization," Networking, Sensing and Control, 2008. ICNSC 2008. IEEE International Conference on. IEEE, 2008, pp.1305-1309.

[21] A. W. Kuang, Z. X. Huang. "Short-Term Traffic Flow Prediction Based on RBF Neural Network," Systems Engineering, Vol. 22, No. 2, 2004, pp. 63-65.

[22] V. T. S. Elanayar, Y. C. Shin,"Radial basis function neural network for approximation and estimation of nonlinear stochastic dynamic systems," *IEEE Transactions on Neural Networks*, Vol. 5, No. 4, 1994, pp. 594-603.

[23] C. Hu, K. Xie, G. Song, et al,"Hybrid Process Neural Network based on Spatio-Temporal Similarities for Short-Term Traffic Flow Prediction,"*International IEEE Conference on Intelligent Transportation Systems,*2008, pp. 253-258.

[24] H. Dong, J. Tang,"Notice of Retraction Prediction of traffic flow at intersection based on self-adaptive neural network,"*IEEE International Conference on Computer Science & Information Technology*, Vol. 8, 2010, pp. 95-98.

[25] D. Karaboga, B. Akay,"Artificial Bee Colony (ABC) Algorithm on Training Artificial Neural Networks,"*SignalProcessing & Communications Applications, Siu IEEE*, Vol. 112, No. 4, 2007, pp. 1-4.

[26] S. Okdem，D. Karaboga，C. Ozturk,An application of Wireless Sensor Network routing based on Artificial Bee Colony Algorithm," *IEEE Congress on EvolutionaryComputation*, Vol. 30, No. 1, 2011, pp. 326-330.

[27] M. Li, Y. Hei, Z. Qiu,"Optimization of Non-convex Cooperative Spectrum Sensing with Modified Artificial Bee Colony Algorithm,"*IEEE International Conference on Computer & Information Technology*, Vol. 25, No. 8, 2014, pp. 70-75.

[28] Y. Wang, W. Chen, C. Tellambura,"A PAPR reduction method based on artificial bee colony algorithm for OFDM signals," *IEEE Transactions on Wireless Communications*, Vol. 9, No. 10, 2010, pp. 2994-2999.

[29] X. Zhang, X. Zhang, S. Y. Yuen, et al,"An Improved Artificial Bee Colony Algorithm for Optimal Design of Electromagnetic Devices,"*IEEE Transactions onMagnetics*, Vol. 49, No. 8, 2013, pp. 4811-4816.

[30] A. H. L. Chen, Y. C. Liang, C. C. Liu,"Portfolio optimization using improved artificial bee colony approach,"*IEEE Computational Intelligence for Financial Engineering & Economics*, Vol. 29, No. 5, 2013, pp. 488-494.

[31] X.F. Li, J.G. Zhang, Y.D. Chu, "Linear Feedback Control for Lozi Chaotic Map," *Journal of Hebei Normal University,* Vol. 31, No. 4, 2007, pp.479-483.

[32] Z.J. Liu, "MATLAB scientific computing and visualization simulation," *Beijing: electronic industry press,* 2009, pp.504-506.

Dawei Chen received his Master's degree and Doctor's degree from Southeast University in 2001 and 2007. He is an

associated professor at the School of Transportation of southeast University.His major research interests include traffic flow prediction, big data, traffic planning.