

A Traffic Forecasting Model Using Adaptive BMO Algorithm Trained Neural Network

Hou-ju SUN^{1,a,*}, Shi-da YANG^{1,b}

¹ Intelligent Computing Labs, Wuhan University of Technology, Luo Nan street, Wuhan, China

^a sunhouju@163.com, ^b yang999sd@126.com

*corresponding author

Keywords: Traffic flow prediction, BP neural network, BMO algorithm, adaptive parameter.

Abstract. Aiming at the short-comings of existing short-term traffic flow forecasting methods, such as low precision, Model is too complex and large computational cost, an adaptive BMO algorithm optimized BP neural network (BMOA-NN) algorithm is proposed. The algorithm improves the traditional BMO algorithm by introducing adaptive mechanism, The dynamic adjustment of the grouping parameters of the algorithm according to the variance of fitness, make the balance between local search and global search ability; Modeling the short-term traffic flow problem, The simulation model of the algorithm is realized, and the precision of the calculation is improved and the complexity of the calculation is reduced. The experimental results show that the proposed model can reduce the prediction error by 18.2% ~ 23.7% in the same time.

1. Introduction

In recent years, the intelligent transportation system has made great progress, but the short-term traffic flow forecast as a bottleneck of urban transportation network is characterized by high dynamic, uncertainty, non-linearity, cyclical and spatial correlation. Need to continue to explore more accurate predictive models For the short-term traffic flow forecasting, domestic and foreign experts and scholars have adopted the linear theory method (historical average model^[1], time series model^[2], state space model^[3]); nonlinear theory method (neural network prediction model^[4], nonparametric regression prediction model^[5], support vector machine regression prediction model); mixed theory method (KARIMA^[7], ATHENA^[8], unit neural network model^[9], fuzzy neural network model^[10]) and simulation model theory. In the above method, the prediction of short-term traffic flow is not adaptive and the robustness is low. Some scholars have used the intelligent algorithm to optimize the BP (Back Propagation) neural network^[11], and the short-term traffic flow prediction model (genetic neural network Model^[12], improved particle swarm optimization algorithm based on BP neural network model^[13]), with the help of the global search capability of intelligent algorithms and the feedback mechanism of BP neural network, the prediction model has good real-time and self- The intelligent algorithm adopted by the model lacks the concept of grouping, can not achieve the balance between global search and local optimization, resulting in the

algorithm convergence too fast; And because there is no simulation model for simulation, the need to use a lot of historical data for training, While the historical data is often limited, so the optimization of the model accuracy is not high. In this paper, a BP neural network prediction (BMOA-NN) model based on adaptive optimization (BMO) algorithm^[14] is proposed to predict short-term traffic flow. Based on the adaptive iteration of BMO algorithm and the adaptive strategy based on fitness variance, the algorithm can dynamically adjust the dynamic value of packet parameters, and avoids the problem of premature aging by simulating the annealing algorithm^[15]. Through the establishment of simulation model, the effective reduction of a large number of redundant computing, greatly improving the global convergence rate of the algorithm, the algorithm is more suitable for practical applications.

2. BP neural network model

BP neural network is a multilayer feedforward network based on error back propagation algorithm (BP) training. As shown in Fig 1. A simple BP neural network model includes input layer, hidden layer and output layer. Complex BP neural network can contain multiple hidden layers. Neurons of each layer are connected with each other in neurons of adjacent layers. There is no connection between the neurons in the same layer. The more the hidden layer of a BP neural network model, the stronger its ability to solve nonlinear problems, and the higher the complexity, the more difficult to achieve.

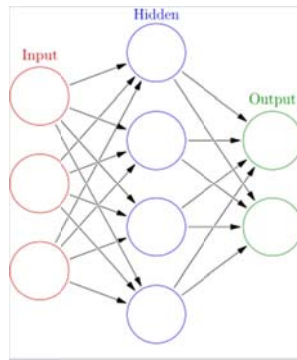


Fig. 1 Artificial Neural Network

BP neural network model includes input and output model, function model, error model and self-learning model.

(1) output model

Implicit node output model:

$$O_j = f\left(\sum W_{ij} \times X_i - q_j\right) \quad (4)$$

Output model:

$$Y_k = f\left(\sum T_{jk} \times O_j - q_k\right) \quad (5)$$

Where f is a non-linear function; q is the neuron threshold.

(2) function model

Function function reflects the lower input on the upper node to stimulate the intensity of the pulse, also known as the stimulus function, generally in (0,1) within the continuous value.

$$f(x) = \frac{1}{1 + e^{-x}} \quad (6)$$

(3) error calculation model

The function of the magnitude of the error between the desired output and the calculated output in the neural network model is called the error calculation function.

$$E_p = \frac{1}{2} \times \sum (t_{pi} - o_{pi})^2 \quad (7)$$

Where t_{pi} is the desired output value; O_{pi} is the actual output value.

(4) self-learning model

The self-learning process of BP neural network is to automatically correct the process of connecting the weight matrix W_{ij} between the lower node and the upper node. The self-learning model is expressed as:

$$\Delta W_{ij}(n+1) = h \times \phi_i \times O_j + a \times \Delta W_{ij}(n) \quad (7)$$

3. BMO Algorithm and Its Improvement

3.1. BMO Algorithm

The BMO algorithm originated from the observation of the use of high quality gene reproduction in birds and has been widely used in many fields[16]. The basic method is to abstract the problem so that it can be represented by a code to represent a solution to the problem. The algorithm first randomly generates a certain number of codes to represent the initial population, and classifies all the codes into male and female according to the degree of fitness of the fitness function, where the female is a better coding. The algorithm is divided into four types of iterations: one-sex breeding, monogamy, polygamy and promiscuity. Each code is treated as a vector, with \vec{X} to represent the male, \vec{X}^i to represent the female, \vec{X}_b is the new individual, ω is the weight factor with time, \vec{r} is the $1 \times n$ vector, each component evaluates to a random number in $[0,1]$, n is the dimension of the problem, mcf is the mutation control factor, r_i is the random number in $[0,1]$, u and l is the upper and lower bounds of the problem.

(1) sexual reproduction can be expressed as:

if $r_1 < mcf$:

$$\vec{X}_b = \vec{X}^i - \mu \times (r_2 - r_3) \times \vec{X}^i \quad (9)$$

then $\vec{X}_b = \vec{X}^i$

(2) monogamous can be expressed as:

if $r_1 > mcf$:

$$\vec{X}_b = \vec{X} + \omega \times \vec{r} \times (\vec{X}^i - \vec{X}) \quad (1)$$

then

$$\vec{X}_b(c) = l(c) \times r_2 \times (l(c) - u(c)) \quad (2)$$

(3) In true nature, polygamy is usually a male and many females mating breeding to produce a group of offspring individuals, but in the BMO algorithm for the convenience of research, we believe that such a male and more females Mating only produces a single offspring. So polygamy expressed as:

if $r_1 > mcf$:

$$\vec{X}_b = \vec{X} + \omega \times \sum \vec{r}_i \times (\vec{X}_i^i - \vec{X}) \quad (12)$$

then

$$\vec{X}_b(c) = l(c) - r_2 \times (l(c) - u(c)) \quad (13)$$

(4) Promiscuous refers to the male and female has no fixed husband and wife relationship, so the model has a great random, so in the BMO algorithm, we directly through the coding rules to re-generate some of the individuals.

3.2. Improved BMO algorithm

In order to better solve the practical problems, many intelligent algorithms have been improved to the adaptive algorithm[17] to optimize the performance of the algorithm. Since the BMO algorithm has four breeding types, each type has a parameter to represent the proportion of the population of that type, and a control factor to control the probability of occurrence of propagation. We first analyze the four types of functions in the algorithm, and then discuss how to set the adaptive function. First of all, the entry into the Sexual reproduction group is the best individual in the whole population, each individual is not affected by other individuals, refused to obtain diversity from other individuals, to keep the gene not to be degraded, so unisexual Breeding has an elite retention effect. Monogamous grouping is the most common way of breeding birds, because this breeding method can make the individual to obtain a certain diversity, but also to ensure the stability of their own genes, it is BMO algorithm stability protection. Polygamous grouping to adopt multi-partner breeding methods, in the algorithm, progeny individuals obtain genes from multiple parent individuals, so that their descendants can obtain great diversity. However, this breeding method is not stable, may produce poor individuals, but this group can avoid the algorithm into a precocious. In the algorithm, we usually remove the individuals in the abortion group directly, and then re-randomly generate the same number of individuals.

Control factor:

$$\text{mcf}(n+1) = \frac{\text{Max}(f(n-1))}{\text{Max}(f(n))} \text{mcf}(n) \quad (14)$$

Where $f(n)$ is the evaluation function of the BMO algorithm, $\text{Max}(f(n))$ is the maximum fitness value for all individuals in the n th generation population, and the initial value of mcf is 1.

We use the fitness variance to dynamically determine the proportion of the population of each group. Assuming that the number of individuals of single-sex reproduction and monogamy group is n , f_i is the fitness value of the i -th individual, f_{avg} is the average fitness value of these individuals, σ^2 is the fitness variance of these individuals, σ^2 can be defined as:

$$\sigma^2 = \sum_{i=1}^n \left(\frac{f_i - f_{avg}}{\mu} \right)^2 \quad (15)$$

Where μ is the normalized scaling factor, its function is to limit the size of σ^2 , μ can take any value, but it is necessary to ensure that the absolute value of $\frac{f_i - f_{avg}}{\mu}$ is not more than 1, and μ is also the same as the iterative process Gradually change the value of μ can be used as follows:

$$\mu = \begin{cases} \max ||f_i - f_{avg}||, & \max ||f_i - f_{avg}|| > 1 \\ 1, & \text{others} \end{cases} \quad (16)$$

The variance σ^2 of the fitness value reflects the degree of convergence of the two groups. The smaller the σ^2 , the more the convergence is, in this case, the number of individuals in the group of parthenogenetic groups and monogamous groups should be reduced, and the number of individuals in polygamous group and overdue groups should be increased,

The Polygamous group:

$$p_{\text{sexual}} = \frac{\sigma_i^2}{\sigma_0^2} p_0 \quad (17)$$

Where σ_i^2 is the variance of the fitness value of the i-th generation of parthenogenetic and monogamous groups, and σ_0^2 is the same as the fitness variance for the first generation. p_0 is the initial ratio set for the parthenogenesis group. The self-adaptive function of the monogamous group is similar to that of parthenogenesis, and is not repeated here.

The proportion of polygamy:

$$p_{\text{polygamy}} = \frac{1 - p_{\text{sexually}} - p_{\text{monogamous}} - \alpha}{p_{\text{polygamy}}} \times p_{\text{polygamy}} \quad (18)$$

Where α is the proportion of promiscuity group.

3.3. BMOA-NN model

First, it is necessary to determine the coding problem of BMOA algorithm. The length of coding depends on the number of hidden neurons in BP neural network. For this issue, we reviewed the California Department of Transportation Performance Measurement System (PeMS) database traffic statistics report (non-weekend), found within 5 minutes of traffic is generally not more than 1600, according to the statistics of the past five years, The input layer and the output layer are set to 11 neurons, thus including the input and output of the largest 2048 vehicles. The input and output of BP neural network are realized by binary coding. There is no more uniform method for how to determine the number of neurons in the hidden layer. However, it is basically believed that the number of neurons in the hidden layer should be less than the sum of the number of neurons in the input and output layers. In this paper, Kolmogorow Theorem is calculated as follows:

$$l = \sqrt{m + n} + a \quad (19)$$

Where l is the number of neurons in the hidden layer, n is the number of neurons in the input layer, m is the number of neurons in the output layer, and a is the complement factor, and any constant between 0 and 10 is taken. Since the number of neurons in the input and output layers is 11, the number of neurons in the hidden layer should be between 5 and 15, through several tests, when the number neurons in the hidden layer is equal to 6 the results will be the best. Therefore, the BP neural network model has a weight of 132 and a threshold of 17. In summary, the encoding length is 149. In this paper, the floating-point coding, the initial value of the value range of 0 to 1, with $x \rightarrow$ to represent a code, coding diagram is as follows.

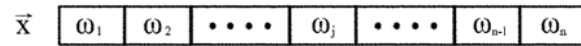


Fig. 2 Individual coding

Where n is 149, $1 \sim 66$ is the value of the weight matrix of the input layer and the hidden layer, $67 \sim 72$ is the hidden layer neuron threshold, $72 \sim 139$ is the value of the hidden layer and output layer weight matrix, $140 \sim 149$ for the output layer neuron threshold.

The algorithm steps are as follows.

Step1: randomly generate a certain number of coding as the initial population of SABMO.

Step2: Evaluate each individual using the fitness function, group the initial population according to the assessment, the best predictive performance for the parthenogenesis group, followed by the

monogamous group, then the polygamous group and the overgrowing group. The initial grouping ratio is set to 0.1, 0.55, 0.3, 0.05 (not the optimal ratio).

Step3: Group the iterations according to the formulas (1), (2), (3), (4), (5).

Step4: Use the fitness function to evaluate the newly generated individual.

Step5: Select the parent individual and the offspring to form a new population.

Step 6: Dynamically adjust the grouping coefficients according to formulas (6), (7), (8), (9).

Step7: If the algorithm ends, go to step8, otherwise go back to step3 to continue.

Step8: assess the fitness value of each individual, the highest fitness value of the individual is the optimal individual, also get the optimal solution of this calculation.

4. Simulation experiment analysis

4.1. DATA SOURCE

The experimental data from the California Department of Transportation Performance Measurement System database, the database is the Orange Country area marked 15-N of a highway 25 sections of the sampling ($p = 25$). A total of 108 monitoring points, the data sampling time interval of 5 minutes, this paper selected part of the date of the data to predict the assessment, and with the actual data to verify. The non-weekend data of March, 2014 (3.01 ~ 3.30) were selected as the training data, and the non-weekend data in April 2014 (4.07 ~ 4.18) were used as test data, a total of 6085 training samples and 2840 test samples.

4.2. Comparison with other models

The BMO-NN model, cost weight particle swarm neural network (PSOC-NN) model, adaptive weight particle swarm optimization (PSOA-NN) model and BMOA-NN model was trained in the experiment using training data (4.07 ~ 4.18), and the comparison of the four neural network prediction results with the actual traffic flow is shown in Figs. 3 to 6. The neural network model is used to predict the non-weekend traffic flow of the neural network model.

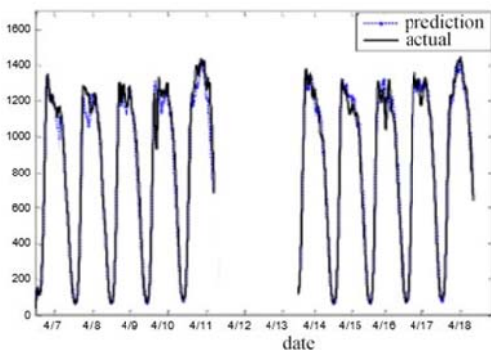


Fig.3 result of BMO-NN model

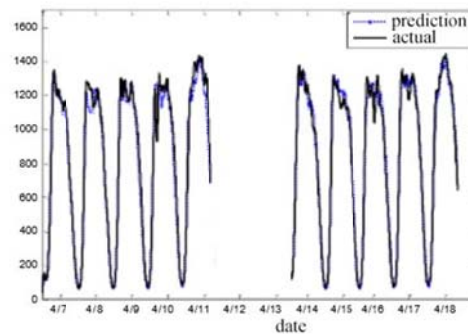


Fig.4 result of PSOC-NN model

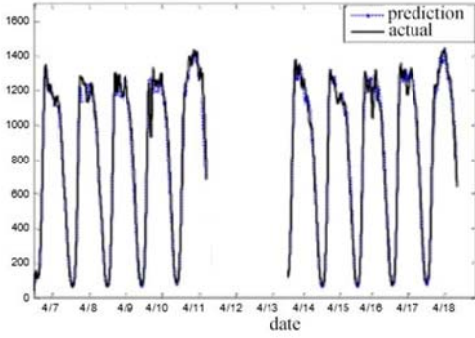


Fig.5 result of PSOA-NN model

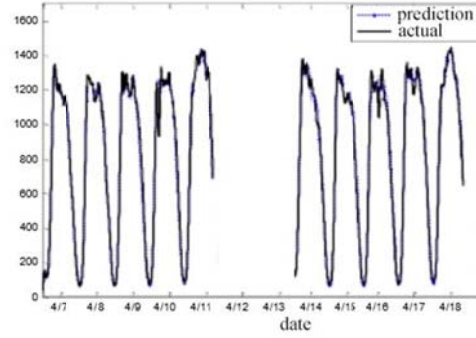


Fig.6 result of BMOA-NNmodel

It can be seen from Figure 3 to 6 that the prediction data of the various prediction models are in good agreement with the direct observation data, and the results of the BMOA-NN model are the highest.

4.3. Analysis

The experimental error is based on the absolute error E , the average absolute error MAE, the mean square error MSE and the average relative error MRE. The formula is:

$$E = NR_i - NP_i \quad (20)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |NR_i - NP_i| \quad (21)$$

$$MSE = \frac{1}{n} \sqrt{\sum_{i=1}^n (NR_i - NP_i)^2} \quad (22)$$

$$MRE = \frac{1}{n} \sum_{i=1}^n \left| \frac{NR_i - NP_i}{NR_i} \right| \quad (23)$$

Where n is the number of traffic flow data; NR_i is the true value of the traffic flow, and NP_i is the predicted value of the traffic flow.

Aiming at the experimental results, the prediction results were evaluated by various error evaluation indexes for quantitative analysis. The evaluation results are shown in Table 1.

Table1. Predictive performance evaluation

measure \ model	BMO-NN	PSOC-NN	PSOA-NN	BMOA-NN
E	3317	1886	1241	989
MAE	0.5931	0.4407	0.4291	0.4109
MSE	0.201449	0.148362	0.120763	0.100386
MRE	0.032615	0.030862	0.027447	0.023001

From the evaluation results of Table 1, it can be seen that the prediction error is 18.2% ~ 23.7% lower than the PSOA-NN model in the same iteration times. Therefore, through the analysis of the prediction results, BMOA-NN model is superior to the other three models.

5. Conclusion

In this paper, an adaptive parameter improvement based on fitness variance is proposed for the parameter setting of BMO algorithm, which makes the improved BMO algorithm have better performance. The improved BMO algorithm is applied to the training of BP neural networks for short - term traffic flow prediction, and the BMOA - NN model is proposed. After the simulation experiment, the prediction results are evaluated by a variety of error evaluation criteria. It is found that the prediction results of the BMOA - NN model are more accurate and robust than the other three models, and can provide more accurate and accurate prediction for traffic guidance and traffic control.

References

- [1] Stephanedes Y J, Michalopoulos P G, Plum R A. Improved estimation of traffic flow for real time control[J]. Transportation Research Record Journal of the Transportation Research Board, 1981(795):28-39.
- [2] Yin Y, Shang P. Forecasting traffic time series with multivariate predicting method[J]. Applied Mathematics & Computation, 2016, 291:266-278.
- [3] Ye Z, Zhang Y, Dan M. Unscented kalman filter method for speed estimation using single loop detector data[J]. Transportation Research Record, 2006, 1968(1):117-125.
- [4] Yin H, Wong S C, Xu J, et al. Urban traffic flow prediction using a fuzzy-neural approach[J]. Transportation Research Part C Emerging Technologies, 2002, 10(2):85-98.
- [5] Clark S. Traffic Prediction Using Multivariate Nonparametric Regression[J]. Journal of Transportation Engineering, 2003, 129(2):161-168.
- [6] Wang J, Shi Q. Short-term traffic speed forecasting hybrid model based on Chaos-Wavelet Analysis-Support Vector Machine theory[J]. Transportation Research Part C Emerging Technologies, 2013, 27(2):219-232.
- [7] Voort M V D, Dougherty M, Watson S. Combining kohonen maps with arima time series models to forecast traffic flow[J]. Transportation Research Part C Emerging Technologies, 1996, 4(5):307-318.
- [8] Kirby H R, Watson S M, Dougherty M S. Should we use neural networks or statistical models for short-term motorway traffic forecasting?[J]. International Journal of Forecasting, 1997, 13(1):43-50.
- [9] Rilett L. Forecasting Multiple-Period Freeway Link Travel Times Using Modular Neural Networks[J]. Transportation Research Record Journal of the Transportation Research Board, 1998, 1617(1):163-170.
- [10] Park B. Hybrid Neuro-Fuzzy Application in Short-Term Freeway Traffic Volume Forecasting[J]. Transportation Research Record Journal of the Transportation Research Board, 2002, 1802(1802):190-196.
- [11] Cui F. Study of Traffic Flow Prediction Based on BP Neural Network[C]// International Workshop on Intelligent Systems and Applications. IEEE, 2010:1 - 4.
- [12] Baher Abdulhai, Himanshu Porwal, Will Recker. Short-Term Traffic Flow Prediction Using Neuro-Genetic Algorithms[J]. Intelligent Transportation Systems Journal, 2002, 7(1):3-41.
- [13] Li S, Wang L, Liu B. Prediction of Short-Term Traffic Flow Based on PSO-Optimized Chaotic BP Neural Network[C]// International Conference on Computer Sciences and Applications. 2013:292-295.
- [14] Askarzadeh A. Bird mating optimizer: An optimization algorithm inspired by bird mating strategies[J]. Communications in Nonlinear Science & Numerical Simulation, 2014, 19(4):1213-1228.
- [15] Metropolis R, Rosenbluth A, Teller A, et al. Simulated annealing[J]. Journal of Chemical Physics, 1953, 21(161-162):1087-1092.
- [16] Askarzadeh A, Coelho L D S. Determination of photovoltaic modules parameters at different operating conditions using a novel bird mating optimizer approach[J]. Energy Conversion & Management, 2015, 89:608-614.
- [17] Zhan Z H, Zhang J. Adaptive Particle Swarm Optimization[M]// Ant Colony Optimization and Swarm Intelligence. Springer Berlin Heidelberg, 2008:227-234.