U.P.B. Sci. Bull., Series C, Vol. 84, Iss. 3, 2022 ISSN 2286-3540

TWO-STAGE NETWORK TRAFFIC PREDICTION BASED ON GWO-SVR

Jiaqi ZHANG[1], Honghui LI[2*], Yang SUN[3], Xueliang FU[4]

With the rapid development of Internet and the fifth-generation mobile communication network, the network traffic keeps raising considerably. Short-term network traffic prediction becomes more and more challenging due to its diversity, self-similarity and burstiness. Big biases may occur when traditional linear models are utilized to predict traffic. In order to address this issue, a two-stage network traffic prediction method TGWO-SVR based on support vector regression (SVR) and grey wolf optimization (GWO) algorithm is proposed. There the GWO algorithm is employed to optimize the three parameters C, ε, and γ in the SVR algorithm. And for the first time, it is putting forward to set up a prediction model using a two-stage prediction method for the short-term network traffic prediction, in addition to using MAWI data set to conduct simulation experiments. The simulation consequence indicates that, in comparisons with SGWO-SVR, SVR, GA-SVR and DE-SVR, the proposed prediction method reduces the values of Root Mean Squared Error and Mean Absolute Percentage Error and enhances the accuracy of network traffic prediction.

Keywords: Network Traffic Prediction; Support Vector Regression; Grey Wolf Optimization Algorithm

1. Introduction

At present, as the development of new-generation information technologies such as communication technology, Internet applications, and cloud computing technology, network scales and requests are becoming more and more complexes, and network throughput is also increasing [1]. Network traffic prediction can effectively maintain and oversee the network, avoid network congestion, improve network presentation, while playing a significant role in network security. Especially in the data center network, accurate prediction of network traffic can not only optimize traffic scheduling, allocate network bandwidth scientifically, but also lessen the energy consumption of the data center

[1] College of Computer and Information Engineering, Inner Mongolia Agricultural University, Hohhot 010000, Inner Mongolia, China
2 College of Computer and Information Engineering, Inner Mongolia Agricultural University, Hohhot 010000, Inner Mongolia, China, e-mail: lihh@imau.edu.cn
3 College of Computer and Information Engineering, Inner Mongolia Agricultural University, Hohhot 010000, Inner Mongolia, China
4 College of Computer and Information Engineering, Inner Mongolia Agricultural University, Hohhot 010000, Inner Mongolia, China

network [2]. Thus, network traffic prediction is a significant research in the meadow of network management.

Usually, according to the prediction period, the prediction algorithm is categorized into brief-term traffic prediction and elongated-term traffic prediction [3]. Long-term traffic prediction methods require large-scale data sets, and the forecasting cycle is quarterly or annual. It is difficult to predict, and the prediction accuracy is low. On the contrary, the data sets needed for short-term prediction is small, mostly in hours, minutes, and seconds. In the prediction, it is easy to build a model, and the prediction accuracy has been improved. According to the characteristics of network traffic, as well as the real-time superintendence of the network, traffic scheduling, energy saving of corresponding apparatus and other features, this paper will build a short-term network traffic prediction model.

There are mainly two kinds of network traffic forecasting methods, linear prediction methods and non-linear prediction manners [4]. Linear prediction methods contain *Autoregressive model*, *Moving Average model*, *Autoregressive* and *Moving Average model* [5], and *Autoregressive Integrated Moving Average model* [6]. However, these linear prediction methods cannot solve the increasingly complex network traffic prediction problem. The nonlinear prediction methods mainly include neural network prediction, grey prediction, support vector regression [7] prediction and so on, which can better predict the network traffic. The neural network can approximate any nonlinear function. So that its prediction accuracy has been improved to a certain extent. But it is prone to overfitting. At the same time, its complexity will greatly increase in model and calculation. In the direction of network management, how to construct a suitable and accurate network traffic prediction method for complex network traffic characteristics is a current research hotspot [8]. Therefore, to improve the accurateness of network traffic prediction, this paper proposed a new method, two-stage prediction method, based on GWO algorithm and SVR, the abbreviated form of a name is TGWO-SVR. Firstly, the grey wolf optimization algorithm is used to optimize the parameters of support vector regression, and then established SVR regression model to become aware of the prediction of network traffic data. The results are better than other methods, such as the conventional SVR, GA-SVR, which is a method of prediction by optimizing the SVR parameters using GA [9], DE-SVR, which is obtained by using a DE of optimizing SVR parameters evolutionary algorithms, and SGWO-SVR, which is a single stage method that is used to optimize the parameters of SVR by GWO algorithms. Finally, the established model in the MAPE and RMSE and other performance metrics has a good performance on indicators.

2. Related Work

It is proposed a network traffic prediction model based on EMD (*Empirical Mode Decomposition*) [8]. First broke down the network traffic into a single IMF (*Intrinsic Mode Function*) part through the EMD network. Then analyze the IMF components through the improved K-means clustering algorithm. Finally, the ARIMA method is used to model and predict the clustered component.

A chaos prediction method of network traffic based on GA and SVM is proposed by XIONG Fan [9]. This method obtains new network traffic time series through phase space reconstruction, which uses GA to optimize the parameters of SVM. Eventually, it realized the disordered prediction of network traffic.

Also, to improve the prediction accuracy of network traffic, Tian et al. [10] proposed a nonlinear prediction method based on GA optimized ESN. The main feature of this paper is that the parameters of the reservoir of the ESN prediction model are optimized by GA, thereby improving the prediction accuracy of network traffic. And the paper used the actual network traffic data for simulation verification. Compared with the three models of ARIMA, Elman neural network and LSSVM, the results show that GA-ESN has higher prediction accuracy and can better describe the network.

GU Zhaojun et al. [11] presented a network traffic prediction model based on particle swarm optimization and Elman neural network. First, to rebuild the phase space of the network traffic sequence in existence is to obtain a new sequence. Then it uses the particle swarm algorithm to optimize the initial parameters of the Elman neural network to obtain the network traffic prediction model. Compared with Elman and BP neural networks, the improved model has improved prediction accuracy.

A network traffic prediction model based on SVM with modified cuckoo search algorithm (MCS-SVR) is proposed in [12]. The existing time series of network traffic are constructed into multi-dimensional series by the method of phase space reconstruction, and then the initial parameters of SVM are optimized by the improved cuckoo algorithm to get the prediction model.

An ARMA-SVR network prediction model based on wavelet decomposition is proposed by LIU Liang et al. [13]. Firstly, the network traffic is divided into several high signals and a low signal, and the high and low frequency signals are predicted by ARMA and SVR, respectively. Finally, each prediction result is linearly combined to get the result.

Long Zhenyue et al. [14] presented a short-term network traffic prediction model. The model combines wavelet packet decomposition and grey wolf horizontal and vertical multi-dimensional chaotic optimization algorithm to optimize Elman neural network. Specifically, the network traffic is decomposed into each frequency band sequence by wavelet packet decomposition, single-stage and two-stage prediction processing is carried out by optimizing the model, and then each predicted value is reconstructed and superimposed, and finally the predicted value of network traffic is obtained. The experimental results indicate that the manner has good prediction accuracy and robustness.

In [15], the authors have used SVR, BP neural network, ARIMA and other models to predict the backbone network traffic in MAWILab data set in the short term. The results show that support vector regression and neural network with nonlinear function fitting ability can achieve better prediction accuracy than ARIMA model.

3. Introduction of related algorithms

3.1. Support vector regression (SVM)

SVM is an efficient algorithm in machine learning for classification and regression. When SVM is requested to regression difficulties, SVM becomes *Support Vector Regression* (SVR) [16].

When using support vector regression for correlation prediction, for a given training set $D = \{(\boldsymbol{x}_1, y_1), (\boldsymbol{x}_2, y_2), \dots, (\boldsymbol{x}_m, y_m)\}, y_i \in R$. The regression function of SVR is:

$$f(\boldsymbol{x}) = \boldsymbol{w}^T \varphi\ (x) + b \tag{1}$$

Wherein w represents the weight vector, b is the conversion vector, $\varphi\ (x)$ represents the feature vector of $\boldsymbol{x}$ in high-dimensional space. In this expected model, the predicted value *f(x)* is adjusted to be the same as the actual value y.

After introducing the slack variable ζ_i^*, ζ_i, the SVR problem can be transformed into the following minimum optimization problem:

$$min \quad \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{n}(\zeta_i^* + \zeta_i) \tag{2}$$

s.t.

$$f(\boldsymbol{x}_i) - y_i \le \varepsilon + \zeta_i$$
$$y_i - f(\boldsymbol{x}_i) \le \varepsilon + \zeta_i^*$$
$$\zeta_i \ge 0, \zeta_i^* \ge 0, i = 1,2,\dots,m$$

In the formula, *C >0* is a constant, which is a penalty coefficient ζ_i^* and ζ_i are slack variables; ε is an insensitive loss coefficient.

In order to be easier to solve, Lagrangian multipliers are introduced to transform the above minimization problem into a convex optimization problem. From this, the dual form of formula 2 is obtained.

$$\begin{aligned} &\max_{\alpha*,a}\ \sum_{i=1}^{m}\ ({\alpha_i}^* - \alpha_i)y_i\ -\ \sum_{j=1}^{m}\ ({\alpha_j}^* - \alpha_j)\varepsilon \\ &\frac{1}{2}\sum_{i=1}^{m}\ \sum_{j=1}^{m}\ ({\alpha_i}^* - \alpha_i)({\alpha_j}^* - \alpha_j)K(x_i, x_j) \end{aligned} \tag{3}$$

In formula (3), $\alpha_i, \alpha_i^*, \alpha_j,$ and α_j^* are all Lagrangian multipliers, and $K(x_i, x_j)$, in this paper, is the kernel function. The kernel function is RBF, and its expression as follows:

$$k(\boldsymbol{x}_i, \boldsymbol{x}_j) = exp\left(-\frac{\|x_i - x_j\|^2}{2\gamma^2}\right) \tag{4}$$

Among them, γ is the parameter of the kernel function. Solving formula (3), the SVM regression model shown in formula (5) can be obtained.

$$f(\boldsymbol{x}) = \sum_{i=1}^{m}(\alpha_i^* - \alpha_i)\ K(\boldsymbol{x}_i, \boldsymbol{x}_j) + b \tag{5}$$

3.2. Grey wolf algorithm (GWO)

In 2014 *Grey Wolf Optimization* [17] algorithm is a population-based heuristic optimization algorithm, which is put forward by Mirjalili et al. The algorithm imitates the social hierarchy and hunting working of grey wolves in nature, being a directed random heuristic algorithm.

Grey wolf optimization algorithm has strong convergence performance and strong global optimization ability, so it has been applied to various fields in recent years.

Grey wolves are primarily group-living animals with a strict social hierarchy among them. The social hierarchy of the grey wolf population is divided into four levels, from top to bottom are the leader α, the main candidate β, the secondary candidate δ, and the ordinary wolf ω. The lower pack must be subordinated to the upper pack. The social rank of the grey wolf is divided by the fitness of the wolf in the wolf pack. And different fitness functions should be defined for different problems.

In the grey wolf optimization algorithm, the social hierarchy model of the grey wolf must be constructed first. When establishing a hierarchy, according to the fitness of grey wolves in the population, the first, second and third best solutions in the population, which is correspond to the positions of α, β, and δ wolves, and the rest are called the ω wolves. Among them, the ω wolves search, surround and attack their prey according to the command of the α, β, and δ wolves, to complete the hunting process. The specific hunting process of grey wolves is as follows:

(1) Surrounded. Grey wolves gradually surround their prey as they search for their prey. Each grey wolf approaches its prey (optimal solution) by constantly updating its position.

(2) Hunting. When the wolf pack is in the hunting process, it is assumed that α, β, and δ wolves can better identify the position of the prey. Therefore, in each iteration, the position information of the top three wolves with the best fitness is retained and recorded as the positions of α, β, and δ wolves. The

positions of the other wolves are changed according to the positions of the three wolves.

(3) Attack. Until the prey stops moving, the grey wolf will launch an attack. The process is determined by the convergence coefficient vector $A(A \in [-a, a])$. In the early stage of hunting, wolves begin to disperse to search for prey when | A | > 1; with the deepening of the search, when | A | ≤ 1, wolves begin to gather and attack their prey.

4. Prediction method based on GWO-SVR

Nowadays, network traffic patterns are complex, a method, two-stage network traffic prediction method, based on GWO-SVR, TGWO-SVR, is proposed. It is known that the historical data set S contains three attribute values $S_i = \{t_i, b_i, b_{i+1}\}$, namely, the traffic bandwidth b_i of the network at the critical moment t_i and the traffic bandwidth of the network at the next moment b_{i+1}, and the traffic bandwidth of the network at the next moment, which is the element of the historical data set S, which contains three attribute values, namely, the bandwidth of the network at the critical moment and the bandwidth of the traffic at the next moment. The single-stage forecasting method SGWO-SVR and the two-stage forecasting method TGWO-SVR are used to establish forecasting models respectively, and the network traffic bandwidth b_{i+1} at the next moment is predicted.

4.1. Single-stage traffic prediction method SGWO-SVR

The symbol definitions are shown in Table 1 which are used in the grey wolf optimization algorithm.

Table 1

GWO symbol description

Symbolic	Meaning	Symbolic	Meaning
t	Current iterations	a	The number of decreases linearly from 2 to 0.
$X(t)$	Grey wolf individual position vector at the moment of t	r_1, r_2	Random vector with a range of [0,1]
$X(t+1)$	The position vector of the grey wolf at the moment of $t+1$.	$X_o(t+1)$	Position vector of grey wolf population at the moment of $t+1$.
$X_p(t)$	Prey position	$X_\alpha(t), X_\beta(t), X_\delta(t)$	Position vector of α, β and δ wolves
A	Convergence coefficient vector	A_1, A_2, A_3	Convergence coefficient vector at the current

			moment
B	Coordination coefficient vector	B_1, B_2, B_3	Coordination coefficient vector at the current moment
D	Distance vector between grey wolf and prey	$D_\alpha, D_\beta, D_\delta$	Distance vector between candidate wolf and optimal three wolves

When using the SVR algorithm to establish the prediction model, the prediction accuracy is mainly affected by the penalty factor C, the insensitive factor ε and the kernel parameter γ in the RBF kernel function. In this paper, the grey wolf algorithm (GWO) is selected to optimize three parameters of SVR. The flow chart of the SGWO-SVR method is shown in Fig. 1.

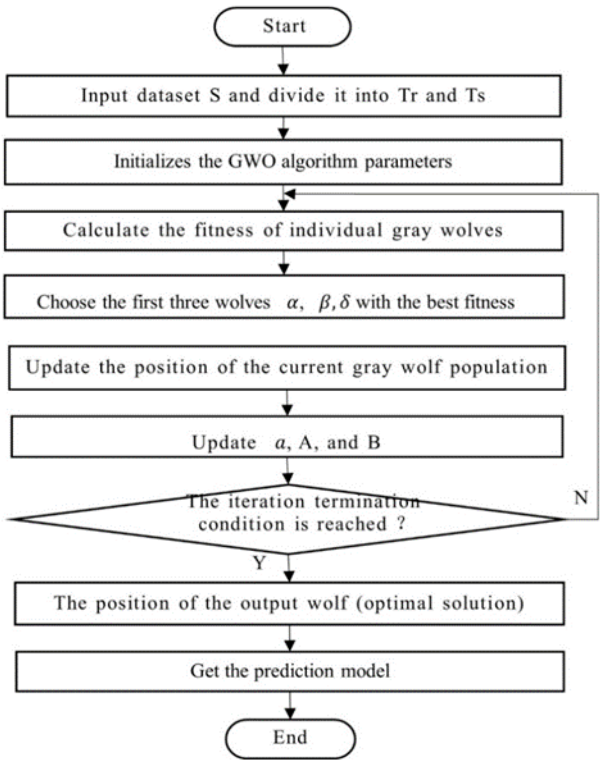

Fig. 1. Flow chart of SGWO-SVR prediction method

The SGWO-SVR method is detailed as follows.

Step1: Input data set S and divide the data set into training set Tr and training set Ts.

Step2: Initialize the GWO algorithm parameters. Initialize the grey wolf population N, solution space dimension D, candidate solution space, iterative termination condition and coefficient a. And the position of each grey wolf corresponds to a candidate solution of the SVR parameter (C, ε, γ).

Step3: Use the candidate solution as the parameter value of SVR, and the prediction model is trained using the training set Tr.

Step4: The fitness function of grey wolf was calculated. In this paper, the fitness function of each grey wolf is defined as the sum of the relative errors of the prediction results, and the calculation method is shown in formula (6).

$$fitness = \sum_{g=1}^{G} \left| \hat{y}_g - y_g \right| \tag{6}$$

Step5: The individual positions of grey wolves with the best fitness, the second best and the third best are defined as the positions of α, β, and δ wolves, i.e., $X_\alpha, X_\beta, X_\delta$

Step6: Move the wolf pack according to the following formula to update the position of the grey wolf.

$$\begin{aligned} D_\alpha &= B_1 \cdot X_\alpha(t) - X(t) \\ D_\beta &= B_2 \cdot X_\beta(t) - X(t) \\ D_\delta &= B_3 \cdot X_\delta(t) - X(t) \end{aligned} \tag{7}$$

$$\begin{aligned} X_1 &= X_\alpha(t) - A_1 \cdot D_\alpha \\ X_2 &= X_\beta(t) - A_2 \cdot D_\beta \\ X_3 &= X_\delta(t) - A_3 \cdot D_\delta \end{aligned} \tag{8}$$

$$X_o(t+1) = \frac{X_1 + X_2 + X_3}{3} \tag{9}$$

Where, the formula (7) represents the distance vector betweenα, β, and δ wolves and other grey wolves. The formula (8) expresses the position update of the grey wolf population under the guidance of α, β, and δ wolves respectively, and the final position of the grey wolf population is synthesized by formula (9).

Step7: As the value of a decreases linearly from 2 to 0, update the values of A and B through formulas (10)-(11).

$$A = 2a \cdot r_1 - a \tag{10}$$

$$B = 2r_2 \tag{11}$$

Step7: Determine whether the iteration termination condition is met, if so, proceed to Step8; otherwise, return to Step3.

Step8: Output the position coordinates of α wolf, that is, the optimal value of SVR parameters (C, ε, γ); obtain a single-step flow prediction model.

4.2. Two-stage Traffic Prediction method TGWO-SVR

In order to improve the accuracy of short-term network traffic prediction, the two-stage traffic prediction method TGWO-SVR is proposed. The flow chart of the TGWO-SVR method to establish the prediction model is shown in Fig. 2.

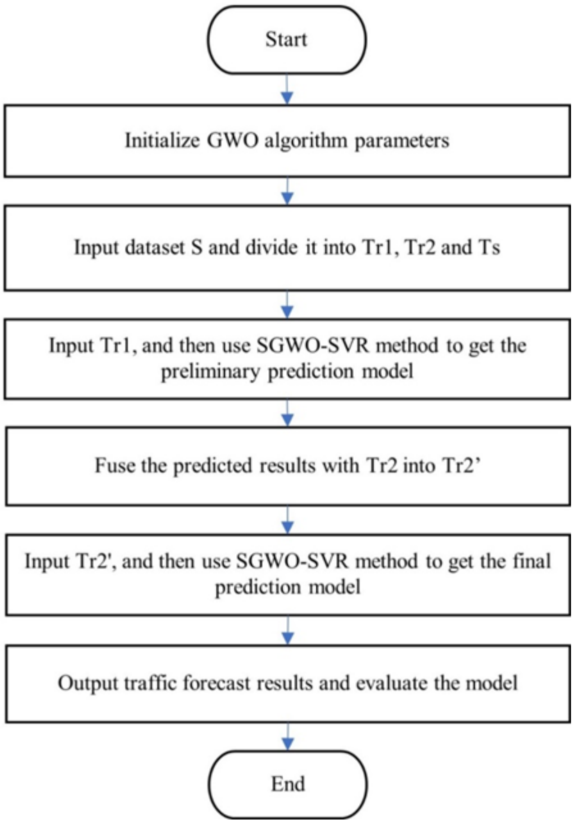

Fig. 2. Flow chart of TGWO-SVR prediction method

The following presents the pseudocode of the TGWO-SVR prediction method:

Input: Original dataset S

Output: Optimal traffic forecast model

Initialization: The number of wolves' population N, the maximum of iterations M and the parameter a

Step1: Divide the original data set $S_i = \{(t_i, b_i, b_{i+1}),\ i = 1, \dots, n\}$ into two training sets Tr1, Tr2 and a test set Ts;

Step2: Input $Tr_1 = \{(t_\ell, b_\ell, b_{\ell+1}),\ \ell = 1, \dots, m\}$. And then use the SGWO-SVR method to train the SVR to get the primary model;

Step4: Combine b_i' with the data in Tr2, and get $Tr_2' = \{(t_j, b_j', b_{j+1}),\ j = \mathrm{m}+1, \dots, k\}$;

Step5: Input Tr2', and then use SGWO-SVR method to optimize the SVR prediction model to obtain the final model TGWO-SVR;

Step6: Use the test set $Ts = \{(t_v, b_v, b_{v+1}),\ v = k+1, \dots, n\}$ to test the performance of the TGWO-SVR model.

5. Simulation and result analysis

5.1. Experimental platform, dataset and initialization parameters

Intel Core i7-4770 processor, 8GB memory and operating system Microsoft Windows 10 are used in this simulation. The prediction method of this paper is realized by using python3.7 to call SK-learn. The traffic data came from a group of the Japanese MAWI. The traffic time series was obtained every two hours from 0:0:0 on March 1, 2018 to 0:0:0 on April 1, 2008, with a total of 371 pieces of data [18]. As shown in Fig. 3, the abscissa is the time axis, the ordinate represents the bandwidth, and the points on the curve represent the bandwidth value at a certain time. The data set is divided into training set Tr and test set Ts; according to 7:3, and then Tr is divided into Tr1 and Tr2 according to 6:4.

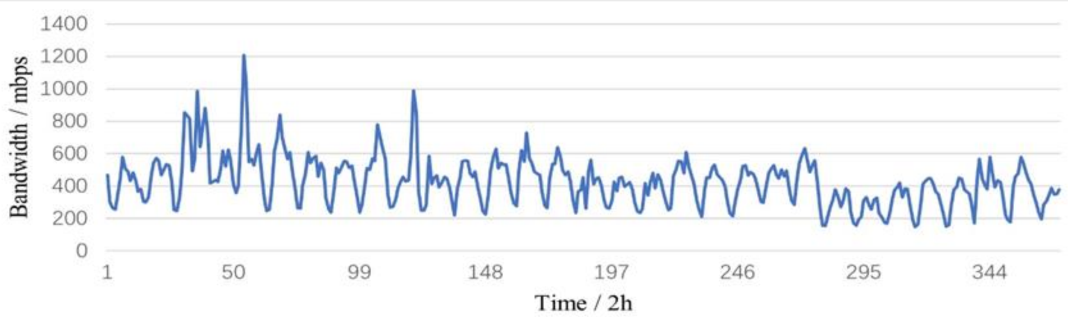

Fig. 3. Actual network traffic

Because SVR is the most sensitive to the data on [0,1], the training data set is normalized as shown in formula (12):

$$x_i' = \frac{x_i - x_{min}}{x_{max} - x_{min}} \tag{12}$$

Where x_i, x_i' are the original data value and the normalized value respectively; x_{max}, x_{min} are the maximum and minimum values of x [19].

To evaluate the prediction accuracy of SGWO-SVR and TGWO-SVR models proposed in this paper, compared with SVR, GA-SVR and DE-SVR, GA-SVR and DE-SVR are the comparative models established after optimizing SVR by GA and DE algorithm respectively. The experimental parameters [20] are set as shown in Table 2:

Table 2

Experimental parameter settings

Parameter	Value	Parameter	Value
Grey wolf population (N)	12	Genetic population(N_{GA})	200
Population dimension (Dim)	3	Cross, compilation probability	0.9
Maximum number of iterations (M)	100	Differential population number (N_{DE})	100

5.2. Evaluation indicators

MAPE and RMSE are the performance indicators in this paper to evaluate the accuracy of the model [21]. The definition method is shown in formulas (13) and (14).

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (\hat{y}_i - y_i)^2} \tag{13}$$

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n} \left|\frac{y_i - \hat{y}_i}{y_i}\right| \tag{14}$$

Among them, $\hat{y}_i$ represents the predicted value of the data flow, and y_i is the actual value.

According to the definition of MAPE and RMSE, the smaller value of both, the better the prediction effect.

5.3. Experimental results and analysis

The results of all prediction models are averaged 10 times in the environment described in this paper. The results are shown in Fig. 4 and Table 3. Fig. 4 shows the traffic predictions and real values of the five methods, where a) is SVR, b) is GA-SVR, c) is DE-SVR, d) is SGWO-SVR, e) is TGWO-SVR. In the figure, the horizontal axis is the time axis and the ordinate is the traffic bandwidth; the curve where the dot is located represents the predicted traffic, and the curve where the pentagram is located is the real traffic.

It can be seen from Fig. 4 that the prediction results of the SGWO-SVR method (Fig. 4-d) and the TGWO-SVR method (Fig. 4-e) proposed in this paper are similar to the actual traffic, and the prediction results using the traditional SVR algorithm (Fig. 4-a) have a big error with the actual value, which is not enough to accurately predict the actual traffic through the historical traffic. The prediction accuracy of the model established after the optimization of SVR parameters by genetic algorithm and differential evolution algorithm is obviously improved.

In order to quantitatively compare the five prediction methods, the experiments of the five prediction methods were repeated for 10 times, and the average values of MAPE, RMSE and time-consuming were calculated respectively, and the prediction indexes of different models were obtained as shown in Table 3.

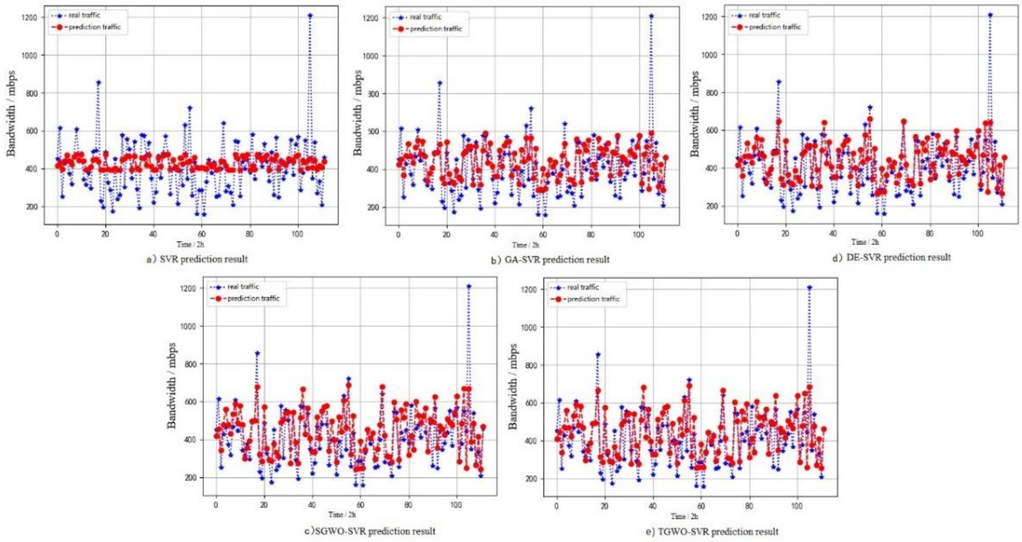

Fig. 4. Comparisons of prediction methods

Table 3

Comparison of different models

Model	MAPE	RMSE	RUNTIME/s
SVR	29.340	135.055	0.994
GA-SVR	22.946	107.694	148.942
DE-SVR	21.142	97.957	85.356
SGWO-SVR	20.527	98.206	11.670
TGWO-SVR	19.490	95.340	10.547

As from Table 3, for MAPE, the value of TGWO-SVR method is the smallest, indicating that this method is superior to the other methods; SGWO-SVR method is second only to TGWO-SVR method, and is also better than the other three methods; and the prediction effect of traditional SVR method is the worst. For the RMSE, the value of TGWO-SVR method is also the smallest, followed by the value of RMSE obtained by DE-SVR method; the value of RMSE directly predicted by SVR method is the largest, and its prediction effect is the worst.

For the running times of the different models listed in Table 3, the five models spent in running duration varies. For the single SVR model, because of its single algorithm and simple calculation, the running time is the smallest, but its prediction accuracy is the worst; for GA-SVR and DE-SVR models, the convergence speed is slow, resulting in the overall running time is too long; for the use of SGWO-SVR method and TGWO-SVR method, the running time is second only to the run time predicted by traditional SVR, but much better than that predicted by GA-SVR and DE-SVR methods.

6. Conclusions

To solve the short-term traffic prediction problem, a new network traffic prediction method, TGWO-SVR, is proposed in this paper. The proposed method utilizes the GWO algorithm to optimize SVR parameters, which affect greatly the accuracy of the prediction method. And the TGWO-SVR method calls the proposed SGWO-SVR twice with the purpose of accuracy improvement. The simulation results illustrate that, the proposed prediction method TGWO-SVR outweighs SGWO-SVR. And these two methods are both superior to others, such as SVR, GA-SVR, DE-SVR, in terms of MAPE, RMSE and accuracy.

Acknowledgements

The authors acknowledge the National Natural Science Foundation of China (Grant: 62041211, 61962047), Inner Mongolia Natural Science Foundation (Grant: 2020MS06011, 2019MS06015).

R E F E R E N C E S

[1]. *H. Lu, F. Yang,* “A Network Traffic Prediction Model Based on Wavelet Transformation and LSTM Network”, 2018 IEEE 9th International Conference on Software Engineering and Service Science (ICSESS). IEEE, 2018.

[2]. *L. Huang, D. G. Wang, X. Liu,* “Network traffic prediction method based on LSTM”, Application Research of Computers, **vol.** 37, no. S1, 2020, pp. 264-265+272.

[3]. *G.E.P. Box, G. M. Jenkins, G. C. Reinsel,* “Time Series Analysis: Forecasting and Control”, Prentice Hall, Englewood Cliffs (1999).

[4]. *H. Lu, F. Yang,* “Research on Network Traffic Prediction Based on Long Short-Term Memory Neural Network”, 2018 IEEE 4th International Conference on Computer and Communications (ICCC). IEEE, 2018.

[5]. *M. Laner, P. Svoboda, M. Rupp,* “Parsimonious Fitting of Long-Range Dependent Network Traffic Using ARMA Models”, IEEE Communications Letters, **vol.** 17, no. 12, 2013, pp.

2368-2371.
[6]. *J. Wang,* “A process level network traffic prediction algorithm based on ARIMA model in smart substation”, 2013 IEEE International Conference on Signal Processing, Communication and Computing (ICSPCC 2013), 2013, pp. 1-5.
[7]. *X. Liu, X. Fang, Z. Qin, et al.,* “A Short-Term Forecasting Algorithm for Network Traffic Based on Chaos Theory and SVM”, Journal of Network and Systems Management, **vol.** 19, no. 4, 2011, pp. 427-447.
[8]. *L. S. Yao, D. Liu,* “Real-time Network Traffic Prediction Model Based on EMD and Clustering”, Computer Science, **vol.** 47, no. S2, 2020, pp. 316-320.
[9]. *F. Xiong,* “Network traffic chaotic prediction based on genetic algorithm optimization and support vector machine”, Modern Electronics Technique, **vol.** 41, no. 18, 2018, pp. 166-169.
[10]. *Z TIAN, X GAO, S LI,* et al. “Prediction method for network traffic based on genetic algorithm optimized echo state net**-** work.” Journal of computer research & development, vol. 52, no. 5, 2015, pp. 1137**-**1145.
[11]. *Z. J. Gu, B. Li, T. Liu,* “Network traffic prediction based on PSO-Elman model”, Modern Electronics Technique, **vol.** 42, no. 01, 2019, pp. 82-86.
[12]. *L. Lu, L. L. Cheng,* “Network Traffic Prediction Model Based on Optimizing SVM with Improved Cuckoo Search Algorithm”, Computer Applications and Software, no. 01, 2015, pp. 124-127.
[13]. *L. Liu, H. H. Jiang, J. Wang, W. Z. Rui,* “ARMA-SVR network traffic prediction method based on wavelet analysis”, Computer Engineering and Design, **vol.** 36, no. 08, 2015, pp. 2021-2025+2032.
[14]. *Z. Y. Long, J. Q. Ai, H. Zou,* “Network traffic predicting model based on improved grey wolf optimization algorithm”, Application Research of Computers, **vol.** 35, no. 06, 2018, pp. 1845-1848.
[15]. *Y. F. Jiang, Y. T. Liu, Y. C. Wang,* “Short-term Prediction for Network Traffic based on MAWI Lab Dataset”, Computer Simulation, **vol.** 36, no. 05, 2019, pp. 407-411.
[16]. *H. Drucker, C. J. C. Burges, L. Kaufman, et al.,* “Support Vector Regression Machines”, Advances in Neural Information Processing Systems, **vol.** 28, no. 7, 1997, pp. 779-784.
[17]. *S. Mirjalili, S. M. Mirjalili, A. Lewis,* “Grey Wolf Optimizer”, Advances in Engineering Software, **vol.** 69, 2014, pp. 46-61
[18]. Http://mawi.wide.ad.jp/mawi
[19]. *N Sanam, P Eros.* “An Application of Internet Traffic Prediction with Deep Neural Network, “Multidisciplinary Approaches to Neural Computing, **vol.** 69, no. 1, 2018, pp. 139-149
[20]. *D. F. Wei,* “Network traffic prediction based on RBF neural network optimized by improved gravitation search algorithm”, Neural Computing and Applications, **vol.** 28, no. 8, 2017, pp. 2303-2312.
[21]. *P O Tiago, S B Jamil, S S Alexsandro.* “Computer network traffic prediction: a comparison between traditional and deep learning neural networks”. Int. J. of Big Data Intelligence, **vol.** 3, no. 1, 2016