IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 14, 2019, accepted September 21, 2019, date of publication October 3, 2019, date of current version October 30, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2944895

A Combined Forecasting Model for Satellite Network Self-Similar Traffic

YUXIA BIE[1], LONGZI WANG[1], YE TIAN[2], AND ZHI HU[3]
[1]College of Electronic Information and Engineering, Shenyang Aerospace University, Shenyang 110136, China
[2]School of Information Science and Engineering, Shenyang Ligong University, Shenyang 110016, China
[3]Software College, Shenyang Normal University, Shenyang 110034, China

Corresponding author: Ye Tian (tianyereal@126.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61901284, Grant 61501307, and Grant 61471247, in part by the General Project of Liaoning Provincial Committee of Education under Grant L201724 and Grant L201727, in part by the Liao Distinguished Professor Program, and in part by the Natural Science Fund Guidance Plan Project of Liaoning Province under Grant 20180550520.

ABSTRACT Since satellite network traffic is self-similar and long-range-dependent (LRD), after analyzing current network traffic forecasting models, a satellite network traffic combined forecasting model that is based on the decomposition fruit fly optimization algorithm – extreme learning machine (FOA-ELM) is proposed. This forecasting model decomposes LRD network traffic into multiple short-range dependent (SRD) components via empirical mode decomposition (EMD), applies the FOA-ELM forecasting model to the decomposed high-frequency components, and applies the ELM forecasting model to low-frequency components. The simulation results show that the forecasting model can improve the forecasting accuracy and forecasting speed, reduce the complexity, and achieve effective and efficient forecasting of satellite network traffic.

INDEX TERMS Self-similarity, forecasting model, empirical mode decomposition, fruit fly optimization, extreme learning machine.

I. INTRODUCTION

Satellite network has an extensive range and a large communication capacity, can be integrated with terrestrial networks. So spatial bandwidth resources can be utilized to divert terrestrial network traffic to alleviate the congestion of terrestrial network. Services in a satellite network include services that are generated on a satellite, services from other satellites, and services from terrestrial networks. Terrestrial networks can generate a large amount of multimedia data. Some satellites also generate multimedia data, for example, remote sensing satellites on regional observations of the Earth's surface can generate multimedia remote sensing data including images and video [1], [2]. Therefore, a satellite network generates and transmits a large amount of multimedia data, such as images, audio and video, and the amount drastically changes. At the network aggregation node, the data generated by the satellite aggregates with the data from other nodes.

Studies have shown that the service traffic of the terrestrial Internet generally exhibits self-similarity [3]. The self-similar traffic of a terrestrial network continues to have self-similarity after it enters the satellite link through the information gateway and is aggregated and transmitted in a satellite network. That is, the self-similarity of the traffic is universally present in the satellite network and can be transmitted [4]–[6]. The self-similarity of satellite network traffic is reflected by its long-range dependence (LRD), and a prominent manifestation of LRD is that the queue performance of the aggregation node can be affected by the long time burst of traffic. It is necessary to establish a forecasting model based on the self-similarity property of the traffic in satellite network to achieve accurate traffic forecasting. The forecasting models of self-similar traffic in terrestrial networks have high computational complexity. Due to the limited on-board computing resources of satellite networks [7], computational complexity of satellite networks traffic forecasting model needs to be reduced. So, traffic prediction models of terrestrial networks are not suitable for satellite networks. Therefore, establishing a high-precision and low-complexity traffic forecasting model, which has great significance to studies of satellite network queue management and resource scheduling, is necessary.

To address these issues, based on the analysis of current network traffic forecasting models, this paper uses ON/OFF

The associate editor coordinating the review of this manuscript and approving it for publication was Tariq Umer.

model to generate self-similar traffic, proposes a satellite network traffic combined forecasting model based on the decomposition fruit fly optimization algorithm-extreme learning machine (FOA-ELM). This forecasting model can improve the forecast accuracy and forecasting speed, reduce complexity, and achieve effective and efficient forecasting of satellite network traffic.

The main contributions of this paper are described as follows:

First, the LRD properties of satellite network traffic can increase the complexity of traffic modeling. This paper uses empirical mode decomposition (EMD) to decompose LRD satellite network traffic into multiple traffic series with single frequency and short-range dependence (SRD) to reduce the forecasting complexity and increase the forecasting speed.

Second, the frequency and nonlinearity of the first intrinsic mode function (IMF1) obtained by EMD is relatively high. In this paper, based on the ELM model, the input weight and threshold of ELM are optimized using the FOA, and the FOA-ELM forecasting model is proposed. The forecasting accuracy can be improved by the FOA-ELM model to predict the IMF1 component.

Third, the frequency of other components obtained by EMD is relatively low, and all other components are predicted to reduce the number of models and increase the modeling efficiency. The ELM model can achieve a relatively high forecasting accuracy.

Last, we design the total flow of the satellite network traffic forecasting process and sum the results of the IMF1 component by the FOA-ELM model and those of all other components by the ELM model to obtain the forecasting results of satellite network self-similar traffic.

The organizational structure of the remainder of the paper is described as follows: Section II describes related forecasting models. Section III introduces traffic model of satellite network. In Section IV, based on the ELM model, an FOA-ELM forecasting model is constructed. Section V studies the EMD algorithm, proposes a decomposition FOA-ELM combined forecasting model, and describes the forecasting steps. Section VI describes the simulation results and performance analysis. Section VII summarizes this paper.

II. RELATED WORK

Current studies of network traffic forecasting models can be summarized as follows:

First, the models that based on the SRD model are employed to fit the LRD, such as the Markov model [8], Autoregressive Moving Average (ARMA) model [9], and Auto Regressive Integrated Moving Average (ARIMA) model [10]. Because network traffic exhibits LRD properties, these models cannot effectively characterize the nature of network traffic, which often causes poor accuracy of traffic forecasting.

Second, models based on the LRD model of time series are proposed, such as the Fractional Autoregressive Integration Moving Average (FARIMA) model, which adds fractional processing based on the ARMA model [11]. The model can reflect the properties of network traffic to a certain extent. Although the forecasting performance is improved compared with models based on the SRD model, the calculation process of the FARIMA model is too complex and time-consuming.

Third, with the development of machine learning, many nonlinear intelligent neural network models have been proposed. These models can better characterize the nature of network traffic and greatly improve the forecasting performance, such as backpropagation (BP) neural networks [12], support vector machines (SVMs) [13] and other improved models. In [14], since the BP neural network models are prone to fall into local optimal values and overfitting, a genetic algorithm (GA) is used to optimize the algorithm of the BP neural network, which achieves certain improvements regarding the disadvantages of the BP neural network and improves the ability of approximation of the BP neural network. However, the two optimization algorithms are relatively complex, the selection of parameters is more difficult and the convergence rate is slow and prone to premature convergence. Reference [15] employed a least squares SVM (LS-SVM), which is an improved model of SVM. The LS-SVM can compensate for the defects of traditional neural networks and the SVM and has a fast learning speed. However, its structure remains relatively complex, determining its parameters are more difficult, and the forecasting accuracy of a single LSSVM model is insufficient for a complex satellite network. In [16] employed principal component analysis and a generalized regression neural network forecast (PCA-GRNN) to forecast Satellite network traffic, principal component analysis(PCA) reduce the input dimensions while preserving the main features of the data, so PCA improve Training efficiency of GRNN, PCA-GRNN method achieves a higher forecasting accuracy and shorter training time, But The GRNN model have higher computational complexity, which is not suitable for satellite networks.

Fourth, due to the shortcomings of the single model, many scholars have proposed combined forecasting models that apply EMD or a wavelet algorithm to decompose the network traffic and then apply other forecasting models on the decomposed network traffic, which effectively reduces the complexity and improves accuracy. In [17], the EMD-ARMA combined forecasting model was proposed, and the decomposed network traffic components were forecasted using the models based on the SRD model. The forecasting performance was greatly improved; however, the decomposed components had nonlinear characteristics, which caused large errors in the results obtained by SRD models.

Regarding the previously mentioned problems and based on the self-similarity and LRD properties of the service traffic in satellite network, we apply EMD in this papa to decompose the LRD satellite network traffic into multiple SRD components, apply the FOA-ELM forecasting model to the decomposed high-frequency components, and apply the ELM forecasting model to the low-frequency components

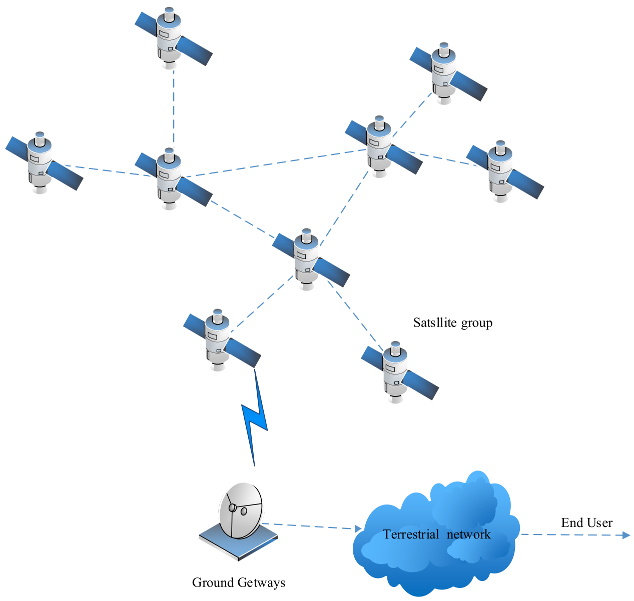

FIGURE 1. Satellite network model.

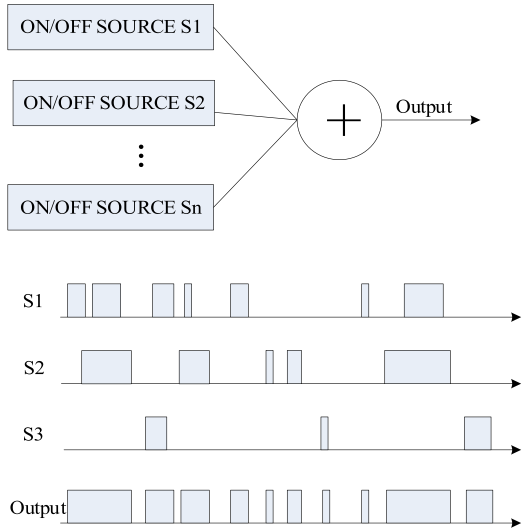

FIGURE 2. ON/OFF traffic superposition model.

and residual. Thus, a decomposition FOA-ELM combined forecasting model for satellite network traffic is proposed.

III. TRAFFIC MODEL OF SATELLITE NETWORK

A satellite network includes space satellite groups, ground gateways and ground terminals, shown as in figure 1, that every node can be connected together through inter-satellite and satellite-ground links [18]. The on-board processing, switching and routing technology can be applied to accurately acquire, quickly process and efficiently transmit the information.

The service of satellite network sink-node includes the service generated by itself and that from on-board or terrestrial node. Each node contains various types of services, that is, multiple services of multiple nodes converge at the satellite network sink-nodes, so traffic convergence occurs.

In the paper, the ON/OFF model is used to represent that the data source of a satellite network node alternates between the states of sending data and non-sending data. The ON corresponds to the data sending duration and the OFF to the pause duration [19]. It can be considered that the continuous ON and the OFF duration are independent and identically distributed. The traffic of satellite network sink-nodes can be regarded as the joint action of the superposition of a series of ON/OFF sources, as shown in figure 2.

If the random variable A of ON period satisfies $\mathrm{E}(A) = 1/\mu$, and the random variable B of OFF period satisfies $\mathrm{E}(B) = 1/v$, then the average speed of the ON/OFF source is $\mu v/(\mu + v)$. After the mathematical derivation of the self-similarity generated by the superposition of the strictly alternative ON/OFF sources, Willinger pointed out that the random process resulting from superposition a large number of ON/OFF sources has self-similarity, where the time intervals of ON and OFF obey heavy-tailed distribution.

The degree of self-similarity of service traffic is in proportion to the degree of heavy tail of the heavy-tailed distribution. Pareto is the most common heavy-tailed distribution, its probability density function and distribution function are respectively:

$$f(x;\alpha,\beta) = \begin{cases} 0\,, & x \le \beta \\ \dfrac{\alpha}{\beta}\left(\dfrac{\beta}{x}\right)^{\alpha+1}, & x > \beta \end{cases} \tag{1}$$

$$F(x;\alpha,\beta) = 1 - \left(\frac{\beta}{x}\right)^{\alpha} \tag{2}$$

In the forms, α is the shape parameter and $\alpha > 0$, indicating the degree of heavy tail of Pareto distribution. The smaller α is, the stronger the heavy tail is. β is the minimum cut-off parameter, indicating the minimum that random variable X can be valued. As long as either the ON or OFF period is of heavy-tail distribution, the superimposed service will result in self-similarity [19]. Therefore, in this paper, approximate self-similar traffic is generated by aggregation of a large number of ON/OFF sources with Pareto distributions.

IV. FOA-ELM FORECASTING MODEL

A. ELM MODEL ANALYSIS

The conventional neural network model has some shortcomings such as slow convergence speed and easy to fall into convergence local minimum, so it is not suitable for satellite networks. Therefore, an efficient and simple traffic forecasting model is needed. Compared with the traditional neural network, the ELM, which was proposed by Prof. Huang [20], has the advantages of a faster computation speed, a simple network topological structure and a better generalization capability than a traditional neural network.

For the ELM, N samples exist for input and output (X_i, t_i) and L hidden layer nodes. The prediction expression is

$$\sum\nolimits_{j=1}^{L} \beta_j f(w_j \cdot X_i + b_j) = y_i, \quad i = 1, \cdots, N \tag{3}$$

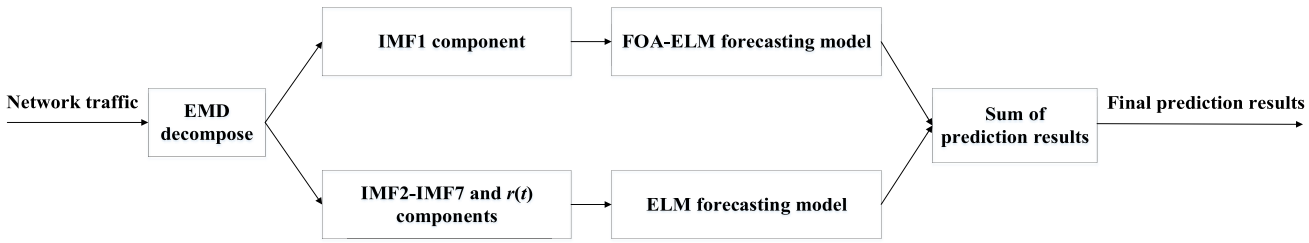

FIGURE 3. Block diagram of the decomposition FOA-ELM combined forecasting model.

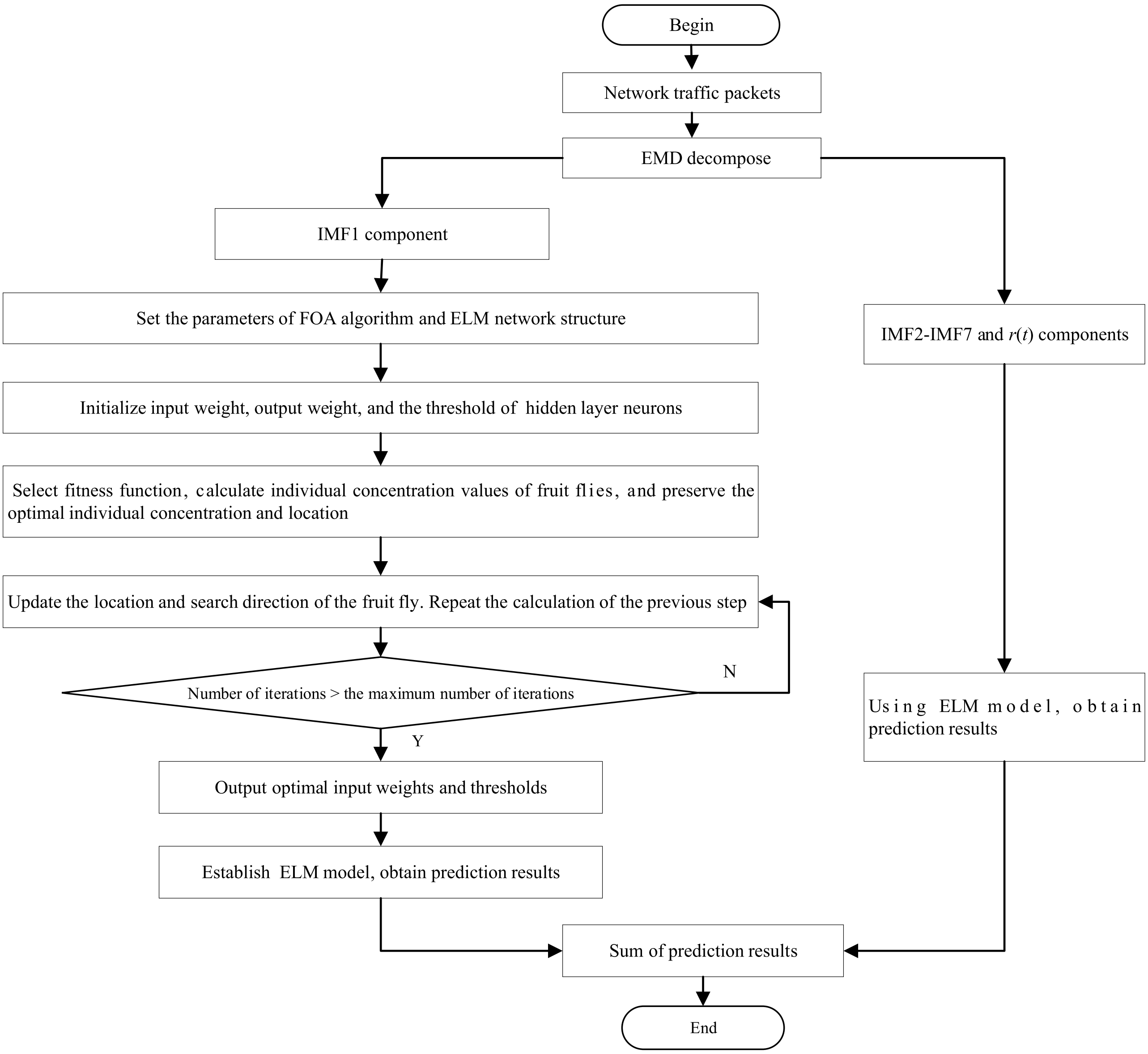

FIGURE 4. Flow diagram of the decomposition FOA-ELM combined forecasting mode.

In Equation (3), β_j is the output weight; w_j is the input weight; $g(x)$ is the activation function; b_j is the bias of j^{th} node in hidden layer; and y_i is the predicted value. To ensure that the predicted result is infinitely close to the actual value t,

$$\sum\nolimits_{i=1}^{N} \|t_i - y_i\| = 0 \tag{4}$$

There are β_j, w_j and b_j, so that Equation (3) is expressed as

$$\sum\nolimits_{j=1}^{L} \beta_j f\left(w_j \cdot X_i + b_j\right) = t_i, i = 1, \cdots, N \tag{5}$$

Equation (5) can be expressed by the matrix; $H\beta = T$, where H is the output of the hidden layer; β is the output

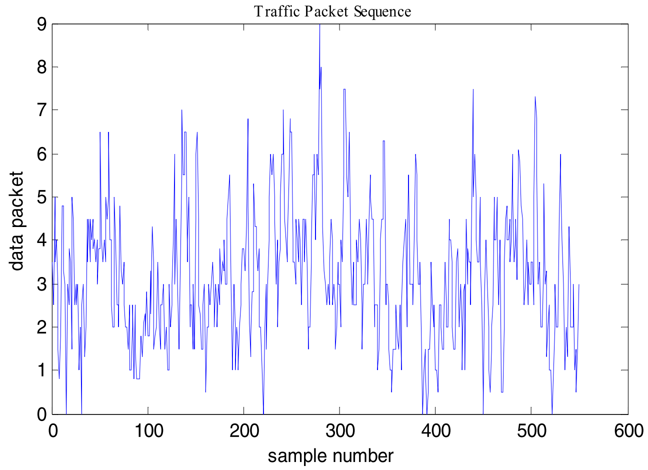

FIGURE 5. Original traffic data.

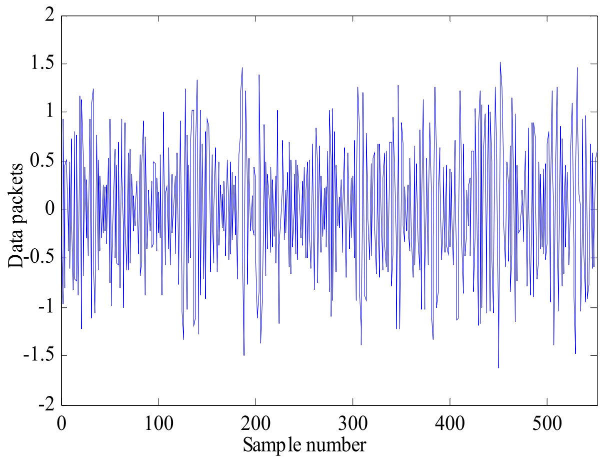

FIGURE 6. IMF1 component.

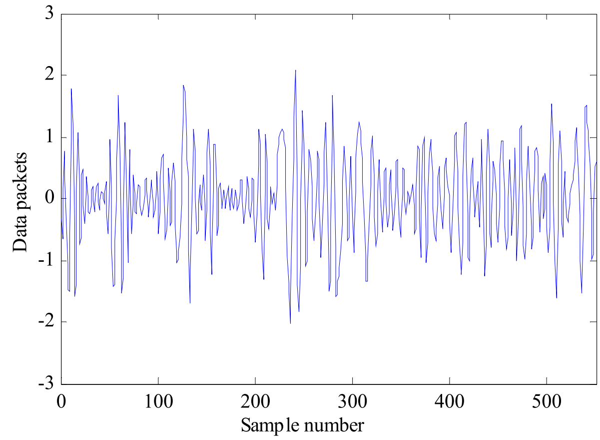

FIGURE 7. IMF2 component.

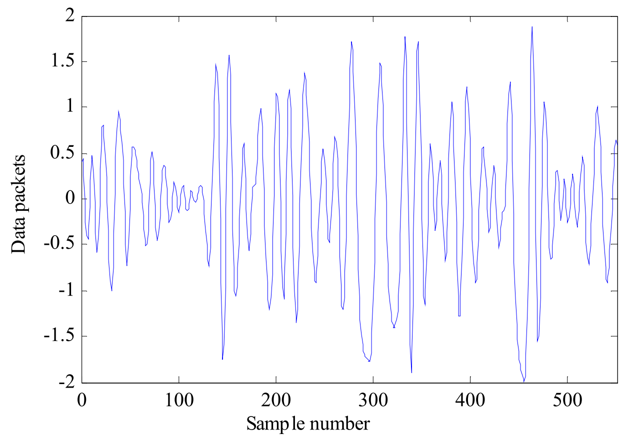

FIGURE 8. IMF3 component.

weight; and T is the expected output. The expression is shown as follows:

$$H(w_1, w_2 \cdots w_L, b_1, b_2 \cdots b_L, X_1, X_2 \cdots X_N) = \begin{pmatrix} f(w_1X_1 + b_1) & \cdots & f(w_LX_1 + b_L) \\ \vdots & \ddots & \vdots \\ f(w_1X_N + b_1)j & \cdots & f(w_LX_N + b_L)j \end{pmatrix}_{N \times L} \tag{6}$$

$$\beta = \begin{pmatrix} \beta_1^T \\ \vdots \\ \beta_L^T \end{pmatrix} T = \begin{pmatrix} T_1^T \\ \vdots \\ T_N^T \end{pmatrix} \tag{7}$$

Therefore, the output weight matrix can be obtained by solving the least squares solution (LSS) for the following equations:

$$\min_{\beta} \left\| H\beta - T' \right\| \tag{8}$$

The output weight matrix β is represented by the following equation:

$$\tilde{\beta} = H^{+}T' \tag{9}$$

H^{+} is the Moore-Penrose general inverse matrix of the output matrix H. The input weight of the ELM matrix and the threshold of the hidden layer are randomly chosen and remain unchanged during the training process, and the output weights are calculated using the previous equations. The random selection and determination of the input weight and threshold enable faster training of the neural network in the ELM compared with the traditional neural network.

Because the input weight and hidden layer neuron threshold in the ELM are randomly selected, the forecasting accuracy can be affected by the LSS. Therefore, this paper uses the FOA to optimize the input weight and threshold in ELM to improve accuracy.

B. CONSTRUCTION OF FOA-ELM MODEL

The FOA, which was proposed by the Taiwan scholar Pan [21], is a swarm intelligence algorithm proposed for global optimization. Compared with GAs, ant colony algorithms, and particle swarm algorithms, the FOA has a simple structure and few parameters, is easily implemented, identified and adjusted, and has relatively high accuracy. The optimization of the ELM using the FOA is to find the optimal concentration values of fruit flies by iteration, so as to obtain the optimal ELM parameters. Algorithm 1 is described in the following pages (See 6).

Algorithm 1 FOA-ELM Model

Input: N training samples.

Output: FOA-ELM model.

Set the population size *SizePop*, the maximum iteration number *Maxgen*, initialize the location of the fruit fly population and record and store the fruit fly position *X_axis* and *Y_axis*, the initial global optimum value *Smellbest.*

Construct the topological structure of the ELM network, set the number of neurons in the input and hidden layers.

Generate the fruit fly population from the input weight w and the threshold b of the hidden layer neurons, and initialize an individual fruit fly's location and radius of the fruit fly population.

$$X(i) = \{x_{i1}, x_{i2}, \cdots, x_{is}\}, Y(i) = \{y_{i1}, y_{i2}, \cdots, y_{is}\} \quad (10)$$

$$R = \{r_{i1}, r_{i2,} \cdots, r_{is}, i = 1, 2, \cdots, sizepop \quad (11)$$

$$S = l \times n + l \quad (12)$$

Where, S is the fruit fly dimension, l is the number of neurons in the hidden layer, and n is the number of neurons in the input layer.

for $m = 1$ to *Maxgen* **do**

for $i = 1$ to *SizePop* **do**

The location and direction of the fruit fly are updated, and the individual taste concentration values of the individual taste concentration values of the fruit fly are calculated

$$X(i) = X\_axis + randomvalue \quad (13)$$

$$Y(i) = Y\_axis + randomvalue \quad (14)$$

Calculate the distance *Di* from the individual fruit fly to the origin and then calculate the determination value of the taste concentration *Si.*

$$Di = \sqrt{X(i)^2 + Y(i)^2} \quad (15)$$

$$Si = 1/Di \quad (16)$$

The fitness function *smell*(i) is selected to calculate the individual concentration value of the fruit fly.

$$smell(i) = \sqrt{\frac{1}{N}\sum\nolimits_{j=1}^{N}(it_j - t_j')^2} \quad (17)$$

Where, t_j is actual input value, t_j' is predicted output value of ELM.

if smell (i) < Smellbest **then**

Preserve the individual location, replace the *Smellbest* by the current concentration value *smell(i)* of the fruit fly. Then, the fruit fly swarm flies towards the location with the optimal smell concentration value.

else

Continue.

end if

end for

end for

Output the optimal parameter (w, b) and establish the ELM prediction model.

return FOA-ELM model.

V. DECOMPOSITION FOA-ELM COMBINED FORECASTING MODEL

A. EMD-BASED SATELLITE NETWORK TRAFFIC DECOMPOSITION

The LRD model of network traffic can increase the difficulty of traffic modeling and has a relatively large impact on the forecasting results. Converting LRD network traffic into an SRD series [22], [23], [24] can effectively reduce the computation complexity and improve the forecasting speed. Commonly employed traffic decomposition methods include wavelet decomposition [25] and EMD [26]. Wavelet decomposition has the difficulty of determining the basis of wavelets. EMD decomposition was proposed by the American scholar Huang et al. The EMD algorithm is simple, its parameters are easily established, and it has a strong ability to analyze local signals. Therefore, EMD is used to decompose the network traffic in this paper. The network traffic is decomposed into single-frequency and SRD time series, which are employed in the nonlinear forecasting model.

EMD decomposes irregular traffic into multiple single frequency components and the residual, i.e., $y\ (t) = \sum IMFs + r\ (t)$, where $y(t)$ is the time series of the traffic, IMFs are components, and $r(t)$ is the residual. The EMD decomposition procedure is described as follows:

Algorithm 2 EMD Algorithm

Input: The time series $y(t)$.

Output: IMF functions, the last residual function r(t).

1: Give the time series $y(t)$, set $r_m = y(t)$.
2: **while** $r_m(t) \neq$ monotonic function **do**
3: Identify all minimum and maximum values of the time series $y(t)$.
4: Use the cubic spline interpolation function to generate the upper and lower envelope lines of the data samples, and consider the mean *ml* of the upper and lower as the standard.
5: Subtract mean from time series signal, $h(t) = y(t)$-*ml*.
6: **if** $h(t)$ satisfies the condition of the IMF function **then** $h(t)$ is assigned as an IMF noted as $d_n(t)$.
7: **else**
8: $y(t) = h(t)$ and go back to step 3.
9: **end if**
10: Compute the residue $r_m(t) = y(t)$-$d_n\ (t)$.
11: $y\ (t) = r_m\ (t)$.
12: **end while**
13: **return** all IMF functions and last residual function $r(t)$.

B. DESIGN OF DECOMPOSITION FOA-ELM COMBINED FORECASTING MODEL

EMD is used to decompose network traffic into multiple SRD subseries, i.e., a limited number of IMF components and a residual. The phase space reconstruction method is used to derive the embedding dimension and construct input and

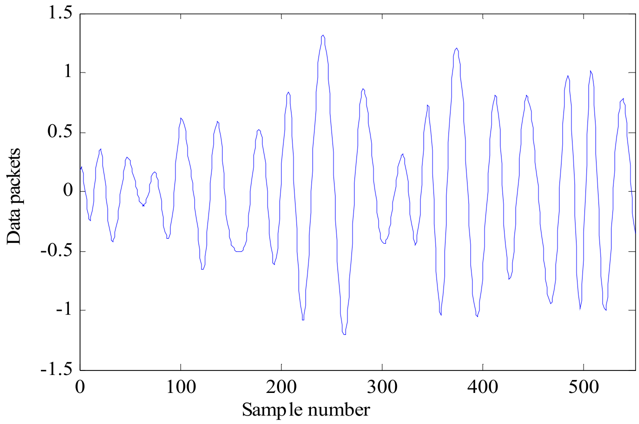

FIGURE 9. IMF4 component.

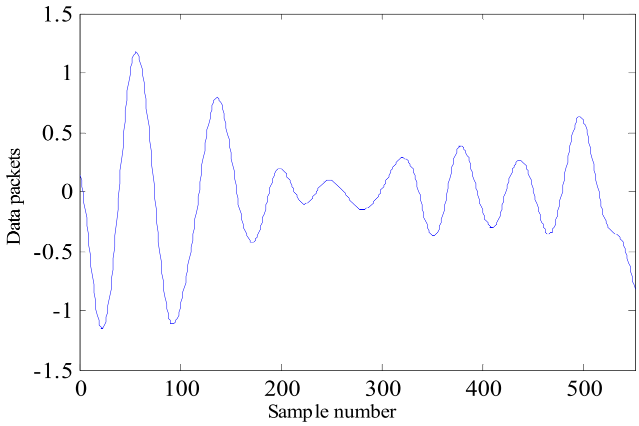

FIGURE 10. IMF5 component.

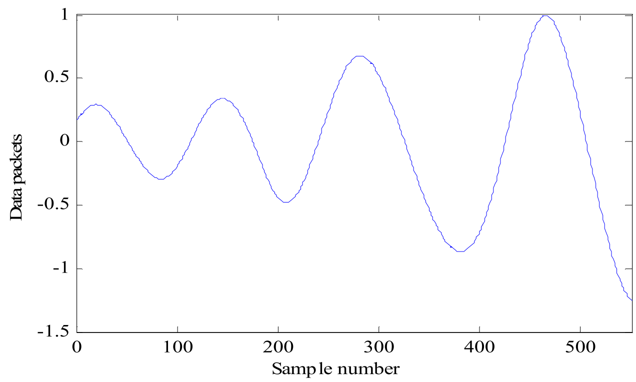

FIGURE 11. IMF6 component.

output data samples. The frequency and nonlinearity of the IMF1 component are relatively high, and the prediction of the IMF1 componen is performed using the FOA-ELM model. The frequency of other components is low, and the prediction accuracy can be achieved using the ELM model. Therefore, other components can be predicted. By this way, the prediction accuracy can be ensured, the number of modeling can be reduced, and the modeling efficiency can be improved. The prediction results of the two models are added to obtain the final prediction results, as shown in Figure 3.

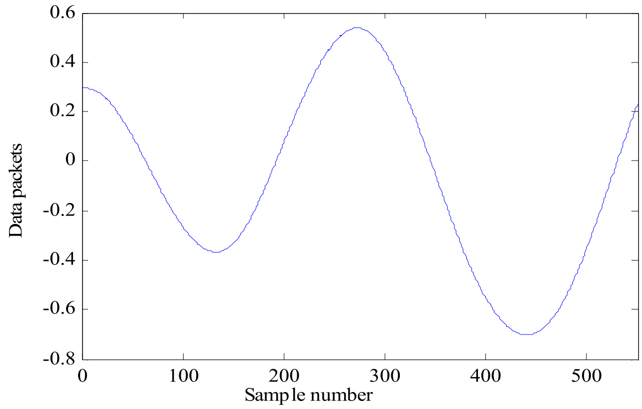

FIGURE 12. IMF7 component.

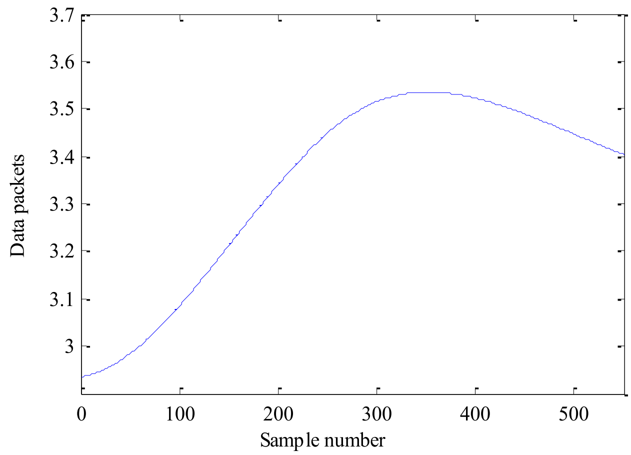

FIGURE 13. $r(t)$ component.

The decomposition FOA-ELM combined model is described as follows:

Algorithm 3 Decomposition FOA-ELM Combined Model

Input: The network traffic series.

Output: Prediction results.

1: Use the EMD algorithm to decompose the network traffic series into seven IMF components and a residual $r(t)$.
2: Separate the IMF1 component with high frequency and nonlinearity from other components with low frequency.
3: Some data is used as the training set, and other data is used as the test set to verify the validity of the model.
4: For the training set, apply Algorithm 1 to optimal parameter (w, b), and use Equation (9) to get optimal parameter β.
5: Use FOA-ELM model described in Algorithm 1 for IMF1 to obtain the prediction results.
6: Apply ELM model described in Algorithm 2 for other components to obtain the prediction results.
7: Add the prediction results of the two parts to obtain the final prediction results.
8: **return** prediction results.

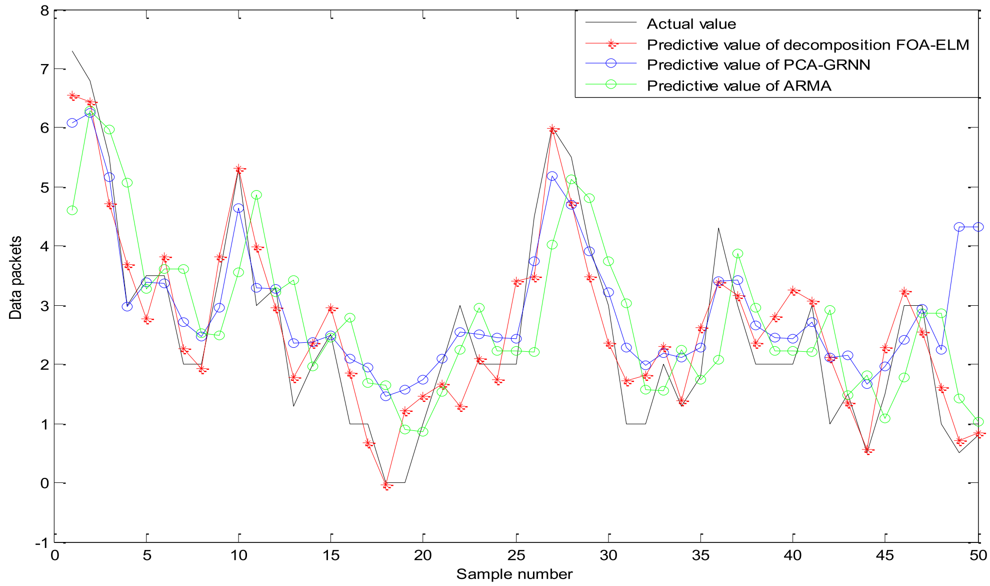

FIGURE 14. Model prediction fit comparison.

The detailed process of the decomposition FOA-ELM combined forecasting model is shown in Figure 4.

VI. PERFORMANCE SIMULATION AND ANALYSIS

The original data packet with the LRD model is generated by aggregation of a large number of ON/OFF sources in the MATLAB environment, and 550 data sets are used as experimental data, as shown in Figure 5. The first 500 data sets are used as the training set, and the last 50 data sets are used as the test set to verify the validity of the model.

The degree of the self-similarity of network traffic is represented by the Hurst parameter value H. If $0< H \leq 0.5$, the data have SRD properties; if $0.5< H <1$, the network traffic has self-similarity and LRD properties. The higher is the value of H, the higher is the dependence and the stronger is the burstiness of the data. This paper employs the R/S analysis method to estimate the H of raw data, and $H = 0.8353$ is obtained, which shows that the original data has relatively high self-similarity and LRD properties.

1) EMD SIMULATION

EMD decomposes network traffic to obtain seven IMF components and the residual $r(t)$, as shown in Figure 6-13. The frequency and nonlinearity of the IMF1 are relatively high, the randomness and suddenness of the remaining IMFs are gradually weakened, and the IMF7 and residual are similar to the oscillating mode of the sinusoidal signal.

2) FORECASTING ACCURACY

To measure the prediction accuracy of the model, the mean square error (MSE) is used as the predictor; its expression is

$$Mse = \frac{1}{n}\sum\nolimits_{i=1}^{n} (y_i - \tilde{y}_i)^2 \tag{18}$$

where n represents the length of the data; y_i is the actual value; and $\tilde{y}_i$ is the predicted output value.

TABLE 1. comparison MSE of different models.

Forecasting Model	MSE
decomposition FOA-ELM	0.371
PCA-GRNN	0.552
ARMA	1.283

The prediction results of the decomposition FOA-ELM combined forecasting model proposed in this paper and those of the ARMA and PCA-GRNN are compared, and the fitting of each model is shown in Figure 14. The fitting of the decomposition FOA-ELM combined forecasting model proposed in this study is better than those of other forecasting models, and the FOA-ELM model can accurately predict the magnitude and trend of the traffic time series.

Table 1 shows the MSEs of the proposed decomposition FOA-ELM combined forecasting model and other models.

As shown in Table 1, the ARMA model has the highest prediction error, followed by the PCA-GRNN model. The decomposition FOA-ELM combined forecasting model has significantly lower error than other models, which indicates that the decomposition FOA-ELM combined forecasting model has a relatively high prediction accuracy.

3) FORECASTING SPEED

This paper applies the ELM model to the IMF2-7 components and the residual $r(t)$ (first approach), and the obtained MSE and operation time are compared with those obtained when each component is separately predicted (second approach), as shown in Table 2. The difference between the two approaches in the prediction accuracy is minimal, while the operation time of the first approach is considerably smaller

TABLE 2. forecast time comparison.

Model	MSE	Prediction time (s)
Overall of IMF2-7 and $r(t)$	0.0623	2.366
IMF2	0.0533	4.420 (Total Prediction time of separately forecast)
IMF3	5.55e-04	
IMF4	1.229e-04	
IMF5	7.36e-07	
IMF6	1.07e-07	
IMF7	1.078e-10	
$r(t)$	1.723e-12	

than that of the second approach. Therefore, the first approach in this paper (that treats IMF2-7 and the residual) can reduce the number of models, reduce the prediction time when ensuring the prediction accuracy, and satisfy the requirements of the forecasting efficiency for satellite network traffic.

VII. CONCLUSION

In this paper, a decomposition FOA-ELM combined forecasting model is proposed by analyzing the self-similarity and LRD properties of the service traffic of satellite networks. EMD is used to decompose the LRD network traffic into multiple SRD subseries, and then the corresponding forecasting models are applied according to the characteristics of various subseries. The decomposed IMF1 component has high-frequency; thus, the FOA-ELM model is employed. This paper treats other components and applies the simple ELM model for prediction. The simulation results show that the decomposition FOA-ELM combined forecasting model is better than other models in forecasting accuracy and forecasting speed, which can satisfy the effective and highly efficient prediction of satellite network traffic.

ACKNOWLEDGMENT

The authors would like to thank Shenyang Aerospace University for support.

REFERENCES

[1] Y. Zhang, M. Pan, and E. F. Wood, "On creating global gridded terrestrial water budget estimates from satellite remote sensing," *Surv. Geophys.*, vol. 37, no. 2, pp. 249–268, Mar. 2016.
[2] M. J. Mccarthy, K. E. Colna, M. M. El-Mezayen, A. E. Laureano-Rosario, P. Méndez-Lázaro, D. B. Otis, G. Toro-Farmer, M. Vega-Rodriguez, and F. E. Müller-Karger, "Satellite remote sensing for coastal management: A review of successful applications," *Environ. Manage.*, vol. 60, no. 50, pp. 323–339, Aug. 2017.
[3] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson, "On the self-similar nature of Ethernet traffic (extended version)," *IEEE/ACM Trans. Netw.*, vol. 2, no. 1, pp. 1–15, Feb. 1994.
[4] Y. X. Bie, C. S. Pan, and H. Y. Liu, "Hurst and priority adaptive RED combined with dynamic scheduling algorithm based on AOS," *J. Commun.*, vol. 33, no. 10, pp. 156–165, Oct. 2012.
[5] Z. Y. Na, Z. H. Gao, and Q. Guo, "Performance analysis of self-similar traffic in LEO satellite network," in *Proc. ICMLC*, Hong Kong, Aug. 2007, pp. 2649–2652.
[6] N. Mokari, P. Hajipour, and L. Mohammadi, "RResource allocation for non-delay-sensitive satellite services using adaptive coding and modulation–multiple-input and multiple-output–orthogonal frequency division multiplexing," *IET Commun.*, vol. 10, no. 3, pp. 309–315, Feb. 2016.
[7] Y. Liu, W. Min, and X. Di, "The performance analysis of satellite virtual channel scheduling algorithms based on stochastic network calculus," in *Proc. ICCSNT*, Changchun, China, Dec. 2016, pp. 612–618.
[8] L. Rossi, J. Chakareski, P. Frossard, and S. Colonnese, "A Poisson hidden Markov model for multiview video traffic," *Netw., IEEE/ACM Trans. Netw.*, vol. 23, no. 2, pp. 547–558, Apr. 2015.
[9] D. Yan and L. Wang, "TPDR: Traffic prediction based dynamic routing for LEO&GEO satellite networks," in *Proc. ICEIEC*, Beijing, China, May 2015, pp. 104–107.
[10] F. Xu, Y. Lin, J. Huang, D. Wu, H. Shi, J. Song, and Y. Li, "Big data driven mobile traffic understanding and forecasting: A time series approach," *Services Comput.*, vol. 9, no. 5, pp. 796–805, Sep. 2016.
[11] C. Katris and S. Daskalaki, "Comparing forecasting approaches for Internet traffic," *Expert Syst. Appl.*, vol. 42, no. 21, pp. 8172–8183, Nov. 2015.
[12] X. Pan, W. Zhou, Y. Lu, and N. Sun, "Prediction of network traffic of smart cities based on DE-BP neural network," *IEEE Access.*, vol. 7, pp. 55807–55816, 2019.
[13] Q. M. Wang, A. W. Fan, and H. S. Shi, "Network traffic prediction based on improved support vector machine," *Int. J. Syst. Assurance Eng. Manage.*, vol. 8, pp. 1976–1980, Nov. 2017.
[14] T. J. Ke, M. Y. Chen, and H. Luo, "Short-term wind power dynamic prediction based on GA-BP neural network," *Adv. Mater. Res.*, vols. 986–987, pp. 524–528, Jul. 2014.
[15] J. X. Liu and Z. H. Jia, "Telecommunication traffic prediction based on improved LSSVM," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 32, no. 3, Aug. 2017, Art. no. 1850007
[16] L. Ziluan and L. Xin, "Short-term traffic forecasting based on principal component analysis and a generalized regression neural network for satellite networks," *J. China Univ. Posts Telecommun.*, vol. 25, no. 1, pp. 15–28, Feb. 2018.
[17] B. Gao, Q. Y. Zhang, and Y. S. Liang, "Predicting self-similar networking traffic based on EMD and ARMA," *J. Commun.*, vol. 32, no. 4, pp. 47–56, Apr. 2011.
[18] R. Liu, M. Sheng, K.-S. Lui, X. Wang, D. Zhou, and Y. Wang, "Capacity of two-layered satellite networks," *Wireless Netw.*, vol. 23, no. 8, pp. 2651–2669, Nov. 2017.
[19] Y. Zhao and L. Liu, "A scheduling model of isochronous frame generation on self-similar traffic of AOS multiplexing," *Int. J. Online Eng.*, vol. 14, no. 7, pp. 15–29, Jan. 2018.
[20] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing*, vol. 70, nos. 1–3, pp. 489–501, Dec. 2006.
[21] W.-T. Pan, "A new fruit fly optimization algorithm: Taking the financial distress model as an example," *Knowl.-Based Syst.*, vol. 26, pp. 69–74, Feb. 2012.
[22] Y. Ren, P. N. Suganthan, and N. Srikanth, "A comparative study of empirical mode decomposition-based short-term wind speed forecasting methods," *IEEE Trans. Sustain. Energy*, vol. 6, no. 1, pp. 236–244, Jan. 2015.
[23] B. Jatin and D. Toshniwal, "Empirical mode decomposition based deep learning for electricity demand forecasting," *IEEE Access*, vol. 6, pp. 49144–49156, 2018.
[24] X. C. Dang, Z. J. Hao, Y. Li, Z. Y. Lu, and Q. Gao, "Network traffic forecasting combination model based on wavelet transform and chaos algorithm," *Int. J. Wavelets, Multiresolution Inf. Process.*, vol. 12, no. 3, May 2014, Art. no. 1450029.
[25] D. F. Walnut, *An Introduction to Wavelet Analysis* (Applied and Numerical Harmonic Analysis), vol. 5, no. 3. New York, NY, USA: Springer-Verlag, 2002, pp. 19–30.
[26] N. E. Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, N.-C. Yen, C. C. Tung, and H. H. Liu, "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. Roy. Soc. London. A Math., Phys. Eng. Sci.*, vol. 454, no. 1971, pp. 903–995, Mar. 1998.

YUXIA BIE was born in China, in 1981. She received the B.S. and M.S. degrees in communication engineering from Shenyang Ligong University, Shenyang, China, and the Ph.D. degree in control science and engineering from the Nanjing University of Science & Technology, Nanjing, China, in 2014. From 2008 to 2016, she was with the Key Laboratory of Communication and Network, Dalian University. Since 2017, she has been a Teacher with the College of Electronic and Information Engineering, Shenyang Aerospace University.

Her research interests include satellite networks, data multiplexing, machine learning, and UAV communication. She presides a National Natural Science Foundation project of China, a Doctoral Startup Foundation of Liaoning Province, and a General Project of Liaoning Provincial Committee of Education. She has authored more than 20 papers and holds 6 patents.

Dr. Bie has been a member of the China Computer Association, since 2012.

LONGZI WANG was born in Chuzhou, Anhui, China, in 1995. He received the B.E. degree in electronic and information engineering from Huangshan University, Huangshan, China. He is currently pursuing the master's degree with the College of Electronic and Information Engineering, Shenyang Aerospace University. His research interests include satellite networks, traffic forecasting, and deep learning.

YE TIAN was born in China, in December 1977. He received the B.Sc. degree in communication engineering and the M.Sc. degree in communication and information system from Harbin Engineering University, China, in 2000 and 2002, respectively, and the Ph.D. degree in communication and information system from Beijing Jiaotong University, China, in 2007. From September 2007 to August 2009, he was a Lecturer with the School of Information Science and Engineering, Shenyang Ligong University, and from September 2009 to August 2012, also an Associate Professor with the School of Information Science and Engineering, Shenyang Ligong University.

He is currently a Professor with the School of Information Science and Engineering, Shenyang Ligong University. His main research interests include satellite communications and signal processing, wireless communications, and signal processing.

ZHI HU was born in Shenyang, Liaoning, China, in 1978. He received the B.S. degree in computer science and technology from Northeastern University, Shenyang, China, the M.S. degree in computer software and theory from Shenyang Ligong University, Shenyang, and the Ph.D. degree in technology of computer application from Northeastern University, Shenyang, in 2016. From 2006 to 2016, he was a Researcher with Neusoft Research, Shenyang. Since 2017, he has been a Teacher with the Software College, Shenyang Normal University.

He has authored six papers. His research interests include ad hoc networks, network routing, and machine learning. He was a recipient of Liaoning Prize for Science and Technology, in 2012.