Received October 14, 2018, accepted November 5, 2018, date of publication November 9, 2018,
date of current version December 18, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2880272

Network Traffic Prediction Method Based on Improved Echo State Network

JIAN ZHOU, XINYAN YANG, LIJUAN SUN, CHONG HAN, AND FU XIAO
[1]College of Computer, Nanjing University of Posts and Telecommunications, Nanjing 210003, China
[2]Jiangsu High Technology Research Key Laboratory for Wireless Sensor Networks, Nanjing 210003, China

Corresponding author: Lijuan Sun (sunlijuan_nupt@163.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 61873131, Grant 71301081, and Grant 61572261, in part by the Natural Science Foundation of Jiangsu Province under Grant BK20130877 and Grant BK20150868, and in part by the Natural Science Foundation of the Higher Education Institutions of Jiangsu Province under Grant 17KJB520027.

ABSTRACT The network traffic sequence has the complex characters, such as mutability, chaos, timeliness, and nonlinearity, which bring many difficulties to network traffic prediction. In order to deal with these complex characters and improve the prediction accuracy, a new network traffic prediction method based on improved echo state network is proposed in this paper. First, to deal with its mutability and chaos, a network traffic denoising algorithm based on local preserving projection is proposed to denoise the raw network traffic sequence. Second, to handle its timeliness and nonlinearity, a network traffic prediction model based on echo state network with double loop reservoir structure is constructed, which takes both the denoised network traffic sequence and the raw network traffic sequence as input. Finally, the proposed method is simulated using two actual network traffic datasets, and simulation results demonstrate that the proposed method can achieve better performance on network traffic prediction compared with other similar methods.

INDEX TERMS Network traffic, echo state network, local preserving projection, prediction, denoise.

I. INTRODUCTION

Accurate prediction of the network traffic sequence is very important to network management and control [1]. With the complex characters such as mutability, chaos, timeliness and nonlinearity, the network traffic sequence is a special kind of time series, which is difficult to predict.

Frequent changes of network user behavior and complexity of network link environment make the network traffic sequence mutable and chaotic. Therefore, it is necessary to preprocess the network traffic sequence. Huang *et al.* [2] decompose the network traffic sequence into multiple intrinsic modal components by using improved global empirical mode decomposition method, then predict each component respectively by using quantum neural network. Nie *et al.* [3] adopt discrete wavelet transform to extract low-pass and high-pass components from the network traffic sequence, then built depth belief network model and Gaussian model to predict these two components respectively. Zhang *et al.* [4] reconstruct the phase space of network traffic sequence, then construct adaptive v-support vector regression model to predict. Meng *et al.* [5] denoise the network traffic sequence by using wavelet soft-threshold method, then adopt relevance vector machine regression model to prediction. Among a number of preprocessing methods, denoising can effectively reduce the negative effect of mutability and chaos of network traffic prediction. Many algorithms have been proposed for denoising, including wavelet denoising [6], wavelet soft-threshold denoising [7] and Local Preserving Projection denoising (LPP) [8]. Because LPP can maintain local manifold features, simplify calculation and improve processing speed, it has been widely applied in many fields, such as abnormal network traffic detection [9] and image classification [10].

Because the network traffic sequence has nonlinear character, traditional linear prediction models such as auto-regressive model [11], auto-regressive integrated moving average model [12] and Poisson model [13] are difficult to meet the requirements of network traffic prediction, while Artificial Neural Network model (ANN) [14] with strong nonlinear approximation ability performs better in network traffic prediction. Therefore, on the basis of ANN, researchers have put forward some network traffic prediction models. Zhang *et al.* [15] propose a network traffic prediction model based on genetic algorithm and radial basis function network to improve prediction accuracy. Narejo and Pasero [16] study deep belief networks with three

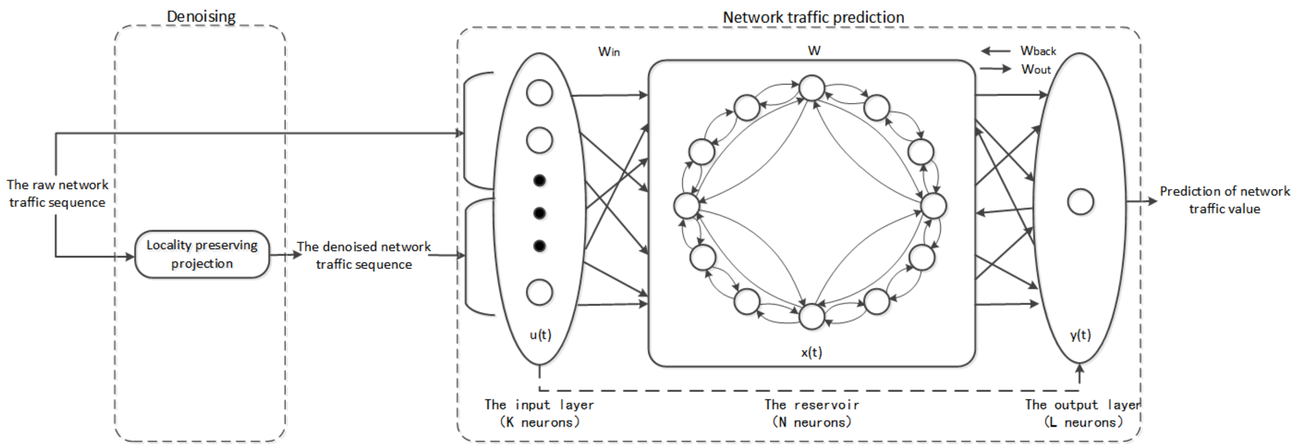

FIGURE 1. **Overall procedure of the network traffic prediction method based on improved echo state network.**

different structures and apply them to network traffic prediction respectively. However, most prediction models based on ANN take too much time to optimize weight matrices iteratively, and can hardly satisfy the real-time requirement [17] in actual network traffic prediction tasks. In 2001, Jaeger [18] proposed a novel recurrent neural network called Echo State Network (ESN). Compared with ANN, ESN only needs to train the output matrix [19], [20], which can improve prediction efficiency. However, due to the random reservoir structure, it takes ESN a long time to generate the reservoir and the ability to handle the nonlinear sequence is restricted. Therefore, it is necessary to fix the reservoir structure to further shorten the training time and increase prediction efficiency of network traffic, and the neuronal connections in the reservoir should be strengthened to improve its ability to handle the nonlinear network traffic sequence.

In order to deal with the complex characters of network traffic sequence, a new network traffic prediction method based on improved echo state network is proposed in this paper. Firstly, to handle the mutability and chaos of network traffic sequence, a network traffic denoising algorithm based on LPP is proposed to denoise the raw network traffic sequence. Secondly, to handle the timeliness and nonlinearity of network traffic sequence, a network traffic prediction model based on Echo State Network with Double Loop Reservoir Structure (ESN-DLRS) is constructed, which takes both the denoised network traffic sequence and the raw network traffic sequence as input. Therefore, not only the varying trend of relative stability in the denoised network traffic sequence, but also the strong correlation with the predicted value in the raw network traffic sequence can be learned by ESN-DLRS, which can improve prediction accuracy. Finally, denoising performance, input window size and prediction performance are analyzed using two actual network traffic datasets.

The rest of this paper is organized as follows: Section II introduces the overview of network traffic prediction method based on improved ESN. Section III describes the network traffic denoising algorithm based on LLP. In Section IV, the network traffic prediction model based on ESN-DLRS is proposed. Simulations and analyses are discussed in Section V. The paper is concluded in Section VI.

II. OVERVIEW OF NETWORK TRAFFIC PREDICTION METHOD BASED ON IMPROVED ESN

The overall procedure of the proposed method is shown in Fig. 1. The prediction task of this method can be described as predicting the network traffic value $tr\,(t+h)$ according to the raw network traffic sequence $Tr_{k_1}\,(t)$ and the denoised network traffic sequence $Tr_{k_2}^{den}\,(t)$, where h is prediction step. $Tr_{k_1}\,(t)$ can be expressed as:

$$Tr_{k_1}\,(t) = \{tr\,(t-k_1)\,, tr\,(t-k_1+1)\,, \ldots, tr\,(t-1)\,, tr\,(t)\} \tag{1}$$

where k_1 is the input window size of the raw network traffic sequence, and $tr\,(t)$ is the network traffic value at time t.

In the preprocessing phase, the denoised network traffic sequence $Tr_{k_2}^{den}\,(t)$ is obtained by the network traffic denoising algorithm based on LPP. $Tr_{k_2}^{den}\,(t)$ can be expressed as:

$$Tr_{k_2}^{den}\,(t) = \Big\{tr^{den}\,(t-k_2)\,, tr^{den}\,(t-k_2+1)\,, \ldots, tr^{den}\,(t-1)\,, tr^{den}\,(t)\Big\} \tag{2}$$

where k_2 is the input window size of the denoised network traffic sequence, and $tr^{den}(t)$ is the denoised network traffic value at time t.

In the prediction phase, the raw network traffic sequence $Tr_{k_1}\,(t)$ and the denoised network traffic sequence $Tr_{k_2}^{den}\,(t)$ are used simultaneously as the input of the network traffic prediction model based on ESN-DLRS. Therefore, in Fig. 1, the input vector of input layer is $u\,(t) = \left(Tr_{k_1}\,(t)\,, Tr_{k_2}^{den}\,(t)\right)^T$, where $K = k_1 + k_2$ is the input vector dimension. The output vector of output layer can be calculated as $y\,(t) = (tr\,(t+h))^T$, where $L = 1$ is the output vector dimension.

Algorithm 1 Prediction Procedure

1. Collect and normalize the network traffic sequence data.
2. Construct training samples $(u_{train}(t), y_{train}(t+h))$, $t = 1, 2, \ldots, T$, where $u_{train}(t) = \left(Tr_{k_1}(t), Tr_{k_2}^{den}(t)\right)^T$. $Tr_{k_2}^{den}(t)$ is obtained by the network traffic denoising algorithm based on LPP mentioned in Section III.
3. Construct the network traffic prediction model based on ESN-DLRS mentioned in Section IV.
4. Train the network traffic prediction model by using samples constructed in Step 2.
5. Denoise the newly collected network traffic sequence. Construct the input vector $u_{pred}(t') = \left(Tr_{k_1}(t'), Tr_{k_2}^{den}(t')\right)$ and put it into the trained prediction model to obtain the predicted value $tr(t'+h)$.

The detailed procedure of network traffic prediction method based on improved echo state network is described as Algorithm 1.

Taking both $Tr_{k_1}(t)$ and $Tr_{k_2}^{den}(t)$ as input can not only make the prediction model learn the varying trend of relatively stable in the denoised network traffic sequence, but also correct the predicted value by exploiting the strong correlation between the raw network traffic sequence and the predicted value. Therefore, the network traffic prediction accuracy can be improved effectively.

III. NETWORK TRAFFIC DENOISING ALGORITHM BASED ON LOCAL PRESERVING PROJECTION

To handle the mutability and chaos of network traffic sequence, a network traffic denoising algorithm based on LPP is described in this section. Firstly, the phase space of network traffic sequence is reconstructed and the local neighborhood for each phase point is determined in the reconstructed phase space. And then, the reconstructed phase space matrix is divided into the network traffic signal subspace and the noise subspace in each local neighborhood. After that, the linear hyperplane is constructed according to the network traffic signal subspace. Finally, the noisy network traffic is processed by orthogonal projection on the linear hyperplane, so that the denoised network traffic sequence can be obtained.

The detailed procedure of network traffic denoising algorithm based on LPP is described as Algorithm 2.

The network traffic denoising algorithm based on LPP uses the local neighborhood to maintain the local manifold features of network traffic data, and uses the orthogonal projection to reduce the noise in the network traffic data, which can effectively reduce the negative effect of mutability and chaos of network traffic sequence.

IV. NETWORK TRAFFIC PREDICTION MODEL BASED ON ECHO STATE NETWORK WITH DOUBLE LOOP RESERVOIR STRUCTURE

To handle the timeliness and nonlinearity of network traffic sequence, a network traffic prediction model based on ESN-DLRS is described in this section. The model is introduced with respect to its construction, calculation and training.

Algorithm 2 LPP Denoising Procedure

1. Select an appropriate time delay τ and an embedding dimension m for the collected network traffic sequence $\{tr(1), tr(2), \ldots, tr(T)\}$ by using the interactive information method [21] and the improved false nearest neighbor method [22] respectively, then reconstruct the network traffic sequence into the m-dimension phase space. The nth phase point in the phase space can be expressed as:
$$S_n = (tr(n), tr(n-\tau), \ldots, tr(n-(m-1)\tau)) \quad (3)$$
2. Find out the phase points whose Euclidean distance with S_n are smaller than radius ε, and obtain the local neighborhood U_n:
$$U_n = \{S_k | \|S_k - S_k\| < \varepsilon\} \quad (4)$$
3. Calculate the covariance matrix C for the local neighborhood U_n. Find out $(m - m_0)$ eigenvectors $\{e_1, e_2, \ldots, e_q, \ldots, e_{m-m_0}\}$ corresponding to the smallest eigenvalues based on eigenvalue decomposition of C, where $m_0 < m$.
4. Project the original phase point S_n to the orthogonal projection hyperplane represented by the eigenvector e_q, and the phase point can be modified as follow:
$$S_n^{den} = S_n - R^{-1} \sum_{q=1}^{m-m_0} e_q \cdot \left[e_q \cdot R\left(S_n - \bar{S}_n\right)\right] \quad (5)$$
where R is the weight coefficient matrix, $\bar{S}_n$ is the mean value of local neighborhood, and S_n^{den} is the phase point after denoising.
5. Repeat the above steps for all the phase points, and restore the reconstructed phase space to obtain the denoised network traffic sequence:
$$\{tr^{den}(1), tr^{den}(2), \ldots, tr^{den}(T)\} \quad (6)$$

A. CONSTRUCTION OF ESN-DLRS

ESN-DLRS consists of three parts as shown in Fig. 1: the input layer, the double loops reservoir and the output layer. The unique reservoir is used as the information processing layer which makes ESN having good memory capacity and powerful nonlinear approximation ability. On the basis of adjacent-feedback loop reservoir (ALR) [23], a Double Loop Reservoir Structure (DLRS) is designed by adding forward connections and feedback connections between neurons with the same neuronal interval. Compared with ESN, DLRS is fixed and connections between neurons in DLRS are strengthened.

The detailed construction procedure of ESN-DLRS is described as Algorithm 3.

Algorithm 3 ESN-DLRS Construction Procedure

1. Set the number of neurons in the reservoir as N. Set the neuronal interval as d, which satisfies $\bmod(N, d) = 0$. Set the fixed weight as r, $r \in (0, 1)$.
2. Construct the first loop by connecting all the internal neurons orderly in the loop form. To be specific, for each neuron i, $i = 1, 2, \ldots, N-1$, set the elements of the reservoir connection matrix W: $w_{i,i+1} = r$, $w_{i+1,i} = r$. When $i = N$, set $w_{i,1} = r$, $w_{1,i} = r$. In particular $w_{i,j}$ is the connection of the neuron i to the neuron j.
3. Construct the second loop. Take the first neuron as a starting point, and connect it with the other neuron which is d neurons apart. After that, connect this neuron and the next one which is d neurons apart in the same way. To be specific, for the neuron i, $i = 1, 1+d, 1+2d, \ldots, 1+(N/d-2)\cdot d$, set the elements of the reservoir connection matrix: $w_{i,i+d} = r$, $w_{i+d,i} = r$. When $i = 1+(N/d-1)\cdot d$, set $w_{i,1} = r$, $w_{1,i} = r$. The construction of DLRS is completed.

It should be noted that the multiple loops reservoir structure can be constructed by using similar procedures. But it has been proved in our previous work [24] that DLRS has better performance. On the one hand, ESN-DLRS has the fixed reservoir structure and the fixed element weights, which effectively shortens the training time to satisfy the real-time requirement of network traffic prediction. On the other hand, the neuronal connections in ESN-DLRS are strengthened to improve its ability to handle the nonlinear network traffic sequence.

B. CALCULATION OF ESN-DLRS

At time t, the input vector of the input layer is $u(t) = \left(Tr_{k_1}(t), Tr_{k_2}^{den}(t)\right)^T$, the output vector of the output layer is $y(t) = (tr(t+h))^T$, and the internal state vector of double loop reservoir is $x(t) = (x_1(t), x_2(t), \ldots, x_N(t))^T$.

The activation states of internal neurons are generally calculated as:

$$x(t) = f^{in}\left(W^{in}u(t) + Wx(t-1) + W^{back}y(t-1)\right) \quad (7)$$

where f^{in} is the activation function of the reservoir neurons (usually the hyperbolic tangent function), W^{in} is the input matrix and W^{back} is the output feedback matrix.

The outputs are calculated as:

$$y(t) = f^{out}\left(W^{out}x(t)\right) \quad (8)$$

where W^{out} is the output matrix, and f^{out} is the readout function of neurons (usually the identity function).

The joint input of the denoised network traffic sequence and the raw network traffic sequence, as well as the unique DLRS, improve ESN-DLRS's ability to handle the nonlinear network traffic sequence.

Algorithm 4 ESN-DLRS Training Procedure

1. Initialize ESN-DLRS parameters, including the reservoir size N and the spectral radius λ, $0 < \lambda < 1$.
2. Bring samples $(u_{train}(t), y_{train}(t+h))$, $t = 1, 2, \ldots, T$ into ESN-DLRS for training.
3. Collect the internal state vectors $x(t)$ from time point t_0 to the time point T, obtain the state matrix
$$X = [x(t_0), x(t_0+1), \ldots, x(T)].$$
4. Collect the expected output vectors, and obtain the expected output matrix:
$$Y = [y(t_0+h), y(t_0+h+1), \ldots y(T+h)].$$
5. Calculate the output connection matrix W^{out} according to linear regression by $W_{out} = Y \cdot X^+$, where X^+ is the pseudo-inverse of X.

C. TRAINING OF ESN-DLRS

The training procedure of ESN-DLRS is described as Algorithm 4.

ESN-DLRS generates W^{in} and W in a fixed form, and keep them unchanged during the whole training process. Only W^{out} need to be computed according to the given samples. It is obvious that ESN-DLRS can reduce the training time to satisfy the real-time requirement of network traffic prediction.

V. SIMULATIONS AND ANALYSES

In order to demonstrate the prediction performance, using Matlab as the simulation tool, the proposed method is simulated and analyzed using two actual network traffic datasets. Dataset A consists of the number of data packets in each minute acquired from Beijing University of Posts and Telecommunications Backbone Nodes of China Education Network (BUPTBN) [25] in several days. Dataset B consists of the number of bits in each 5 minutes from United Kingdom Academic Network Backbone Nodes (UKANBN) [26] in a few days. Fig. 2 shows 6000 data points of Dataset A and 4000 data points of Dataset B respectively.

A. SIMULATION SETTINGS

For the network traffic denoising algorithm based on LPP, the denoising neighborhood radius ε is set to 0.6. For the network traffic prediction model based on ESN-DLRS, the spectral radius λ of the reservoir is set to 0.8 and the reservoir size N is set to 80. The reservoir activation function adopts the hyperbolic tangent function, and the output activation function adopts the identity function. The output feedback matrix is not used in the model. The neuronal interval d of DLRS is set to 8 and the reservoir matrix W of DLRS is generated as described in Section IV (A). The input window size of the raw network traffic sequence is set as $k_1 \in \{1, 2, \ldots, 10\}$, and the input window size of denoised network traffic sequence is set as $k_2 \in \{0, 1, \ldots, 5\}$. For Dataset A and Dataset B, the first half of data points are used for training and the rest half

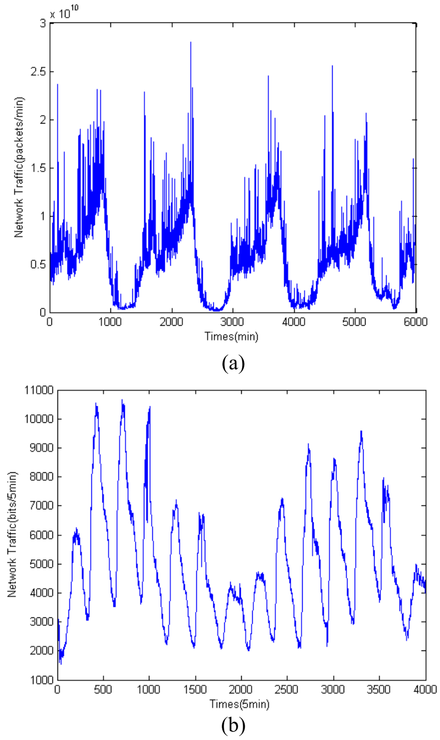

FIGURE 2. Network traffic datasets. a) Dataset A. b) Dataset B.

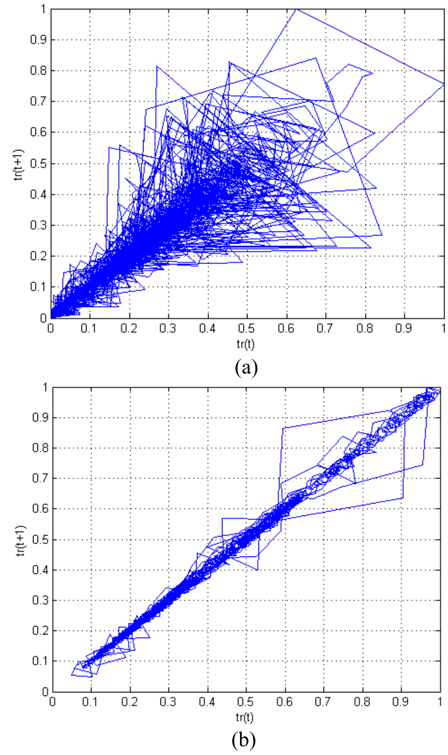

FIGURE 3. Noisy 2-D phase diagram. a) Dataset A. b) Dataset B.

are for testing. The Normalized Mean Square Error (NMSE) is used to evaluate the prediction accuracy. The smaller the NMSE is, the higher the prediction accuracy is. The parameter configurations are summarized in Table I.

TABLE 1. Simulation parameter configurations.

Parameters	Values
Denoising neighborhood radius	0.6
Reservoir size	80
Spectral radius	0.8
Reservoir activation	tanh
Output activation	identity
Neuronal interval	8
Input window size of the raw network traffic sequence	$\{1,2,\ldots,10\}$
Input window size of the denoised network traffic sequence	$\{0,1,2,\ldots,5\}$

B. DENOISING PERFORMANCE ANALYSIS

In order to prove the effectiveness of network traffic denoising algorithm based on LPP, the raw network traffic sequence and the denoised network traffic sequence are analyzed by using the 2-D phase diagram method [27] and the maximal Lyapunov exponent method [28].

The noisy 2-D phase diagrams can be obtained by normalizing the raw network traffic sequence, as shown in Fig. 3. The phase diagram of Dataset A is very disorganized. The phase diagram of Dataset B is more organized than that of Dataset A, but sometimes varies abruptly. The Root Mean Square Error (RMSE) of phase points is 0.0562 for Dataset A and 0.0118 for Dataset B. It can be seen that the raw network traffic sequence has the mutability character. The 2-D phase diagram of the denoised network traffic sequence can be obtained by using the network traffic denoising algorithm based on LPP, as shown in Fig. 4. The RMSE of phase points is reduced to 0.0187 for Dataset A and 0.0058 for Dataset B. It can be seen that the proposed algorithm can effectively reduce the mutability of network traffic sequence and make the sequence smoother.

The comparison of the maximal Lyapunov exponent is shown in Fig. 5. Before denoising, the maximal Lyapunov exponent of Dataset A is 0.029, and the maximal Lyapunov exponent of Dataset B is 0.0156, which indicates that the raw network traffic sequence has the chaos character. After denoising with the network traffic denoising algorithm based on LPP, the maximal Lyapunov exponent of Dataset A is 0.0255, and the maximal Lyapunov exponent of Dataset B is 0.0134, which indicates that the proposed algorithm can effectively reduce the chaos of network traffic sequences.

C. INPUT WINDOW SIZE ANALYSIS

The influence of the input window sizes of both the raw network traffic sequence k_1 and denoised network traffic sequence k_2 on the prediction accuracy with the network traffic prediction model mentioned in Section IV is analyzed as shown in Fig. 6.

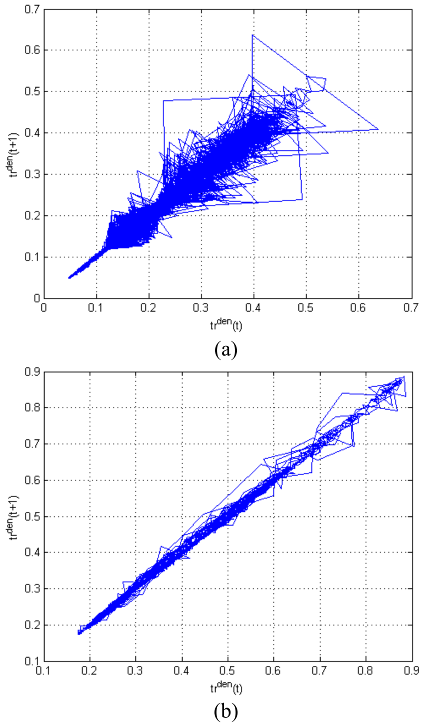

FIGURE 4. Denoised 2-D phase diagram. a) Dataset A. b) Dataset B.

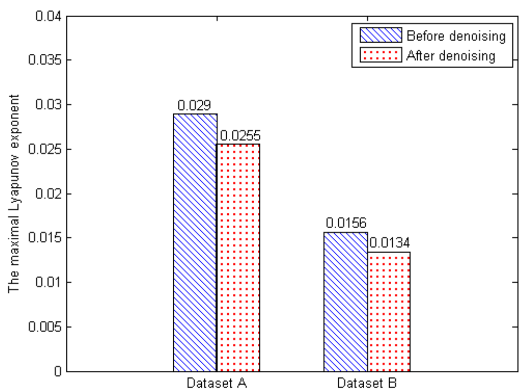

FIGURE 5. Comparison of the maximal Lyapunov exponent.

Firstly, as can be seen from Fig. 6, when $k_2 = 0$, namely, the input only consists of the raw network traffic sequence, the prediction error is much larger than that of $k_2 > 0$ for either Dataset A or Dataset B. It indicates that adding the denoised network traffic sequence to input can help the prediction model to learn the varying trend of relatively stable of network traffic sequence and effectively improve the prediction accuracy.

Secondly, as can be seen from Fig. 6, for Dataset A, when $k_1 = 1$ and $k_2 = 1$, the value of NMSE is the minimum which is 0.98378. Meanwhile, for Dataset B, when $k_1 = 1$ and $k_2 = 4$, the value of NMSE is the minimum which is 0.0037442. It shows that the optimum input window size

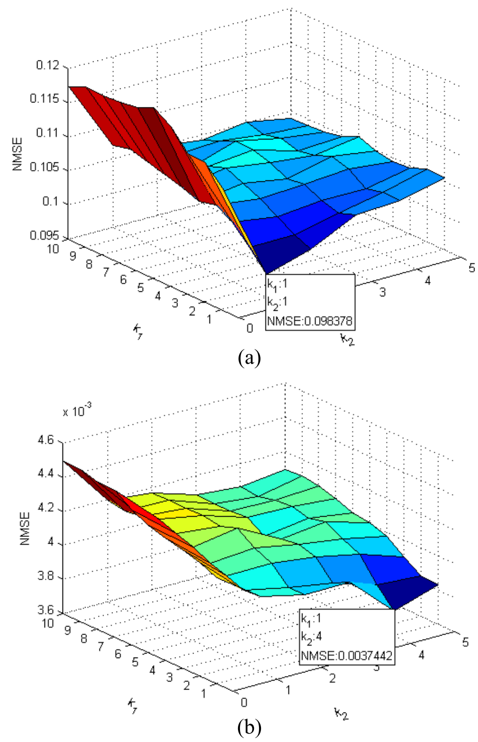

FIGURE 6. Input window analysis of the raw traffic sequence and the denoised traffic sequence. a) Dataset A. b) Dataset B.

of denoised network traffic sequence will be different due to different datasets or different sampling intervals.

Finally, as can be seen from Fig. 6, the optimum input window size of the raw network traffic sequence is 1. Because the aim of using input raw network traffic sequence is to correct the predicted value. Thus, it is enough to choose only the latest raw network traffic value which has the closest correlation with the predicted value as input.

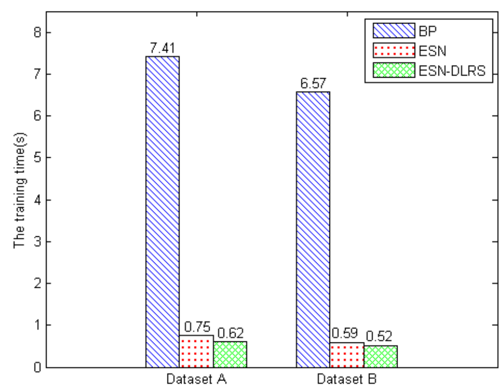

FIGURE 7. Training time analysis.

D. PREDICTION PERFORMANCE ANALYSIS

In order to demonstrate the prediction performance, the comparison of prediction performance with other similar methods is shown in Fig. 7 and Fig. 8. Fig. 7 shows the training time of

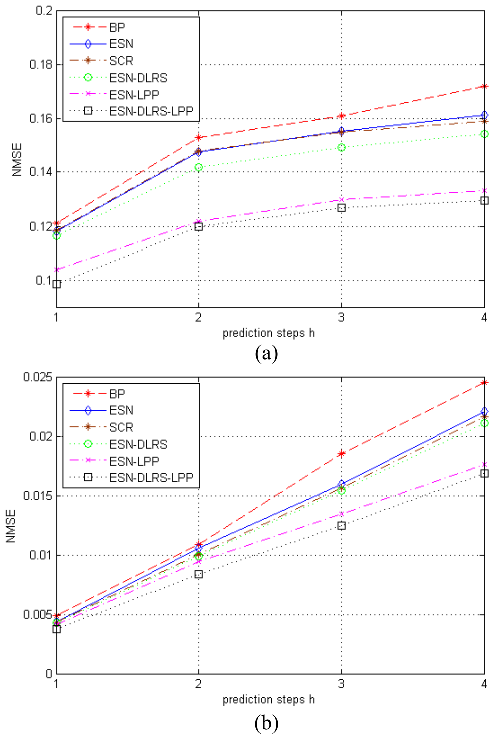

FIGURE 8. NMSE analysis. a) Dataset A. b) Dataset B.

the prediction models based on Back Propagation (BP) neural network [29], ESN and ESN-DLRS in the same computing environment. It can be seen from Fig. 8 that the training time of ESN-DLRS is far less than BP and slightly less than ESN for different datasets. ESN and ESN-DLRS require training time far less than BP, because ESN and ESN-DLRS adopt the linear regression method rather than the gradient descent method and only need to train the output matrix. ESN-DLRS costs less training time than ESN, which indicates that the fixed reservoir structure and the fixed element weights can further reduce training time to satisfy the real-time requirement of network traffic prediction.

Fig. 8 shows the NMSEs of the prediction models based on BP, ESN, Simple Circle Reservoir (SCR) [25], ESN-DLRS with only the raw network traffic sequence as input, also shows the prediction performance of the prediction models based on ESN, ESN-DLRS with both the raw network traffic sequence and the denoised network traffic sequence which is obtained by network traffic denoising algorithm based on LPP as input (ESN-LPP and ESN-DLRS-LPP), respectively. According to the input window size analysis in Section V (C), for ESN-LPP and ESN-DLRS-LPP, set $k_1 = 1$, $k_2 = 1$ for Dataset A, and set $k_1 = 1$, $k_2 = 4$ for Dataset B. It can be seen from Fig. 8 that the NMSE of ESN-DLRS-LPP is smaller than that of other methods for different datasets and different prediction steps. By changing the random reservoir structure to the simple circle reservoir structure, SCR is simplified based on ESN, consequently the NMSE of SCR is similar to ESN. In particular, with only the raw network traffic sequence as input, ESN and SCR have better prediction performance than BP, and ESN-DLRS performs better than ESN and SCR. This is because ESN and SCR have stronger nonlinear approximation ability than BP, while ESN-DLRS strengthens neuronal connectivity to improve this ability. With the same prediction model, ESN-LPP performs better than ESN, which indicates that taking both the raw network traffic sequence and the denoised network traffic sequence as input can effectively improve prediction accuracy. With the same input, ESN-DLRS-LPP has better prediction performance than ESN-LPP, which indicates that adding the denoised network traffic sequence to input and adopting DLRS at the same time can synthesize the merits of both and maximize prediction accuracy.

VI. CONCLUSION

The network traffic sequence has the complex characters such as mutability, chaos, timeliness and nonlinearity, which brings many difficulties to network traffic prediction. In this paper, a new network traffic prediction method based on improved echo state network is proposed. Firstly, to handle the mutability and chaos of network traffic sequence, a network traffic denoising algorithm based on LPP is proposed to denoise the raw network traffic sequence. Secondly, to handle the timeliness and nonlinearity of network traffic sequence, a network traffic prediction model based on ESN-DLRS is constructed, which takes both the denoised network traffic sequence and the raw network traffic sequence as input. Finally, denoising performance, input window size and prediction performance are analyzed using two actual network traffic datasets. Simulation results demonstrate that the proposed method performs better than other similar methods. Our future work will focus on the adaptive acquisition of the best input window sizes of both the raw network traffic sequence and denoised network traffic sequence.

ACKNOWLEDGMENT

The authors would like to thank the editor and the anonymous reviewers whose constructive comment help improving the presentation of this paper.

REFERENCES

[1] X. Cao, Y. Zhong, Y. Zhou, J. Wang, C. Zhu, and W. Zhang, "Interactive temporal recurrent convolution network for traffic prediction in data centers," *IEEE Access*, vol. 6, pp. 5276–5289, Feb. 2018.

[2] W. Huang, J. Zhang, S. Liang, and H. Sun, "Backbone network traffic prediction based on modified EEMD and quantum neural network," *Wireless Pers. Commun.*, vol. 99, no. 4, pp. 1569–1588, Apr. 2018.

[3] L. Nie, D. Jiang, S. Yu, and H. Song, "Network traffic prediction based on deep belief network in wireless mesh backbone networks," in *Proc. Wireless Commun. Netw. Conf. (WCNC)*, San Francisco, CA, USA, Mar. 2017, pp. 1–5.

[4] X. Zhang, Z. Wu, J. Zhang, and J. Chen, "An adaptive network traffic prediction approach for LDoS attacks detection," *Int. J. Commun. Syst.*, vol. 31, no. 7, p. e3505, 2017.

[5] Q.-F. Meng, Y.-H. Chen, Z.-Q. Feng, F.-L. Wang, and S.-S. Chen, "Nonlinear prediction of small scale network traffic based on local relevance vector machine regression model," *Acta Phys. Sinica*, vol. 62, no. 15, pp. 1–7, 2013.

[6] W. Habib, T. Sarwar, A. M. Siddiqui, and I. Touqir, "Wavelet denoising of multiframe optical coherence tomography data using similarity measures," *IET Image Process.*, vol. 11, no. 1, pp. 64–79, Jan. 2017.

[7] Q. Zhang, H. Liou, X. Zhou, and Y. Wang, "Wavelet soft-threshold denoising method of power quality signal," *High Voltage Eng.*, vol. 32, no. 1, pp. 99–101, 2006.

[8] X. He and P. Niyogi, "Locality preserving projections," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 16, no. 1, pp. 186–197, 2002.

[9] X.-X. Wen, X.-R. Meng, Z.-Q. Ma, and Q.-Y. Kang, "Network traffic anomaly detection based on local projection denoise and FSVDD," *Appl. Res. Comput.*, vol. 30, no. 5, pp. 1523–1526, 2013.

[10] Y. Zhai *et al.*, "A modified locality-preserving projection approach for hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 8, pp. 1059–1063, Aug. 2016.

[11] S. XU and B. Zeng, "Network traffic prediction model based on auto-regressive moving average," *J. Netw.*, vol. 9, no. 3, pp. 97–102, 2014.

[12] H. Z. Moayedi and M. A. Masnadi-Shirazi, "Arima model for network traffic prediction and anomaly detection," in *Proc. Int. Symp. Inf. Technol.*, Kuala Lumpur, Malaysia, Aug. 2008, pp. 1–6.

[13] T. Karagiannis, M. Molle, M. Faloutsos, and A. Broido, "A nonstationary Poisson view of Internet traffic," in *Proc. INFOCOM*, Hong Kong, Mar. 2004, pp. 1558–1569.

[14] W. Du, S. Y. S. Leung, and C. K. Kwong, "Time series forecasting by neural networks: A knee point-based multiobjective evolutionary algorithm approach," *Expert Syst. Appl.*, vol. 41, no. 18, pp. 8049–8061, 2014.

[15] Y. M. Zhang, Y. B. Zhang, and L. Z. Lin, "Notice of retraction application of genetic algorithm and RBF neural network in network flow prediction," in *Proc. IEEE Int. Conf. Comput. Sci. Inf. Technol.*, Chengdu, China, Jul. 2010, pp. 298–301.

[16] S. Narejo and E. Pasero, "An application of Internet traffic prediction with deep neural network," in *Proc. Workshop Neural Nets*, Vietri sul Mare, Italy, 2018, pp. 139–149.

[17] Z. Duan, Y. Yang, K. Zhang, Y. Ni, and S. Bajgain, "Improved deep hybrid networks for urban traffic flow prediction using trajectory data," *IEEE Access*, vol. 6, pp. 31820–31827, 2018.

[18] H. Jaeger, "The 'echo state' approach to analyzing and training recurrent neural networks-with an erratum note," German Nat. Res. Center Inf. Technol., Sankt Augustin, Germany, Tech. Rep. 148, Jan. 2001.

[19] H. Jaeger and H. Haas, "Harnessing nonlinearity: Predicting chaotic systems and saving energy in wireless communication," *Science*, vol. 304, no. 5667, pp. 78–80, Apr. 2004.

[20] H. Jaeger, "Tutorial on training recurrent neural networks, covering BPRT, RTRL, EKF and the 'echo state network' approach," German Nat. Res. Center Inf. Technol., Sankt Augustin, Germany, Tech. Rep. 159, Jan. 2002.

[21] A. M. Fraser and H. L. Swinney, "Independent coordinates for strange attractors from mutual information," *Phys. Rev. A, Gen. Phys.*, vol. 33, no. 2, pp. 1134–1140, Feb. 1986.

[22] L. Cao, "Practical method for determining the minimum embedding dimension of a scalar time series," *Phys. D, Nonlinear Phenomena*, vol. 110, nos. 1–2, pp. 43–50, 1997.

[23] X.-C. Sun, H.-Y. Cui, R.-P. Liu, J.-Y. Chen, and Y.-J. Liu, "Modeling deterministic echo state network with loop reservoir," *Frontiers Inf. Technol. Electron. Eng.*, vol. 13, no. 9, pp. 689–701, 2012.

[24] L. Sun, X. Yang, J. Zhou, J. Wang, and F. Xiao, "Echo state network with multiple loops reservoir and its application in network traffic prediction," in *Proc. IEEE 22nd Int. Conf. Comput. Supported Cooperat. Work Design*, Nanjing, China, May 2018, pp. 689–694.

[25] H. Cui, C. Feng, Y. Chai, R. P. Liu, and Y. Liu, "Effect of hybrid circle reservoir injected with wavelet-neurons on performance of echo state network," *Neural Netw.*, vol. 57, no. 9, pp. 141–151, 2014.

[26] P. Cortez, M. Rio, M. Rocha, and P. Sousa, "Multi-scale Internet traffic forecasting using neural networks and time series methods," *Expert Syst.*, vol. 29, no. 2, pp. 143–155, 2012.

[27] Z. Tang, T. Peng, and W. Wang, "A small scale network traffic prediction algorithm based on correlation analysis and local scale least squares support vector machine," *Acta Phys. Sinica*, vol. 63, no. 13, pp. 49–58, 2014.

[28] Y. Zhang and W. Guan, "Prediction of multivariable chaotic time series based on maximal Lyapunov exponent," *Acta Phys. Sinica*, vol. 58, no. 2, pp. 756–763, 2009.

[29] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, "Learning representations by back-propagating errors," *Nature*, vol. 323, pp. 533–536, Oct. 1986.

JIAN ZHOU was born in 1984. He received the Ph.D. degree from the Nanjing University of Science and Technology in 2012. He is currently an Associate Professor with the Nanjing University of Posts and Telecommunications, Nanjing, China. His research interests are computer networks and machine learning.

XINYAN YANG was born in 1993 She is currently pursuing the M.E. degree with the Nanjing University of Posts and Telecommunications, Nanjing, China. Her current research interests include network traffic prediction and machine learning.

LIJUAN SUN received the B.S. degree in radio engineering from Southeast University in 1985, the M.S. degree in signal, circuit, and systems, and the Ph.D. degree in communication and information system from the Nanjing University of Posts and Telecommunications, Nanjing, China, in 1988 and 2007, respectively. She is currently a Professor and a Ph.D. Supervisor of information network with the Nanjing University of Posts and Telecommunications. Her main research interests include computer networks and evolutionary computation.

CHONG HAN received the M.S. degree in computer application from Henan University in 2010 and the Ph.D. degree in information network from the Nanjing University of Posts and Telecommunications, Nanjing, China, in 2013. He is currently a Lecturer with the Nanjing University of Posts and Telecommunications. His research interests include wireless sensor networks and multimedia information processing.

FU XIAO was born in 1980. He received the Ph.D. degree from the Nanjing University of Science and Technology in 2007. He is currently a Professor with the Nanjing University of Posts and Telecommunications. His research interest is sensor networks.