Small-time Scale Network Traffic Prediction Based on Complex-valued Neural Network

Bin Yang
School of Information Science and Engineering, Zaozhuang University, Zaozhuang, China
batsi@126.com

Abstract. Accurate models play an important role in capturing the significant characteristics of the network traffic, analyzing the network dynamic, and improving the forecasting accuracy for system dynamics. In this study, complex-valued neural network (CVNN) model is proposed to further improve the accuracy of small-scale traffic network traffic forecasting. Artificial bee colony (ABC) algorithm is proposed to optimize the complex-valued and real-valued parameters of CVNN model. Small-scale traffic measurements data namely the TCP traffic data is used to test the performance of CVNN model. Experimental results reveal that CVNN model forecasts the small-time scale network traffic measurement data very accurately.

1. Introduction
Network traffic modeling and analysis play a major role in characterizing network performance. Accurate models have an important role in capturing the salient characteristics of the network traffic, analyzing the network dynamic, and improving the forecasting ability for system dynamics. It has a fundamental meaning for many network designs and engineering problems, e.g., switcher designing, router, the management of devices and its software development [1].

The rapid development of the communication and network technologies results in the uncertain characteristics of the traffic network, especially the nonlinear time series. The nonlinear time series are more accurately predicted by using some models. The flexible neural tree (FNT) model was employed to predict the small-time scale traffic measurements data [2]. The FNT structure was developed by using the Genetic Programming, and the parameters are optimized by the Particle Swarm Optimization (PSO) algorithm. Ordinary differential equations (ODEs) were also presented to forecast and reflect the actual characterize of the real network traffic [3]. Jiang et al. proposes a new prediction algorithm to network traffic in the large-scale communication network in order to overcome the problem of network traffic prediction in the communication network [4].

Recently, complex-valued neural network (CVNN) has been proposed to predict the time series data. Compared with real-valued neural network, CVNN is more flexible and functional. In order to predict the small-time scale traffic measurements data, complex-valued neural network (CVNN) model is proposed to predict nonlinear small-time scale traffic measurements data. Artificial bee colony (ABC) algorithm is proposed to optimize the complex-valued and real-valued parameters of CVNN model. The TCP traffic data is used to test the performance of CVNN model.

The paper is organized as follows. In Section 2, the details of CVNN model are described. In Section 3, the examples are used to examine the effectiveness and veracity of the proposed method. Finally, Section 4 concludes this paper.
2. Method

2.1. Complex-valued neural network

In a complex-valued neural network, input layer, weights, threshold values and output layer are all complex numbers. A three-layer complex-valued neural network is described in Fig. 1, which has m input nodes, n hidden nodes and one output node. Suppose that input vector \([z_1, z_2, \cdots]\). The result of \(i\)-th hidden node is computed as followed.

\[
h_i = f(W_i + W_{i1}z_1 + W_{i2}z_2 + \cdots)
\]

Where \(W_i\) is the threshold value of \(i\)-th hidden node and \(W_{i1}, W_{i2}, \cdots\) are the weights. \(f\) is the activation function, which is

\[
f(z) = \frac{z}{c + r |z|}
\]

Where \(c\) and \(r\) are the real variables.

The output layer is computed:

\[
y = f(h_1V_1 + h_2V_2 + \cdots)
\]

\[\text{Figure. 1. A three-layer complex-valued neural network.}\]

2.2. Optimization of parameters

In order to search the optimal parameters of CVNN model, there are many local and global evolutionary methods, such as PSO, genetic algorithm (GA) and simulated annealing (SA). In this paper, artificial bee colony (ABC) algorithm is selected because of its fast convergence, more accurate solution and stability [5-6].

ABC algorithm can be summarized as follows.

(1) Choose the initial value of algorithm parameter including maximum number of generation (\(\text{max\_gen}\)), number of employed bees and onlooker bees (\(\text{num\_employ}\) and \(\text{num\_onlooker}\)), the maximum times that food sources do not change its location (\(\text{limit}\)).

(2) In a CVNN (Fig. 1), many parameters need to be optimized, containing weights (\(W_i\) and \(V_i\)), threshold value (\(W_i\)) and parameters in activation function (\(c\) and \(r\)). Due to the fact that weights and threshold value are complex numbers, both real and imaginary parts need to be optimized. For a CVNN \(m-n-1\) (\(m\) inputs, \(n\) hidden neurons and 1 output), the number of parameters is \(2nm+6n+2\) and parameter vector is
\[
\begin{align*}
\{\text{Re}(W_{1}), \text{Im}(W_{1}), \ldots, \text{Re}(W_{n}), \text{Im}(W_{n})\].
\end{align*}
\]

Generation \(\text{num\_employ}\) and \(\text{num\_onlooker}\) solutions, calculate the fitness value, and select the best half of solutions as foods.

(3) In order to produce a candidate food position from the old one in memory, the ABC uses the following equation:

\[
\nu_{ij} = x_{ij} + R_{ij}(x_{ij} - x_{ij})
\]

Where \(\nu_{ij}\) is a new position of food, \(R_{ij}\) is a random number in the range \([-1,1]\), \(k \in \{1,2,3,\ldots,K\}\) and \(k \neq i\). Calculate the fitness of \(\nu_{ij}\), and when the fitness of \(\nu_{ij}\) is better than \(x_{ij}\), replace \(x_{ij}\) with \(\nu_{ij}\). If \(x_{ij}\) is not changed, \(\text{failure}_i = \text{failure}_i + 1\).

(4) The probabilities for onlookers are calculated as follows.

\[
p_i = \frac{\text{fit}_i}{\sum_{n=1}^{N} \text{fit}_n}
\]

Where \(\text{fit}_i\) is the fitness value of \(i\)-th bee.

(5) According to \(p_i\), onlooker bees search the area as employed bees.

(6) If \(\text{failure}_i > \text{limit}\), then scout bee randomly generations a new solution according to the following equation:

\[
x_{ij} = x_{ij} + \text{rand}(0,1)(x_{\text{max},j} - x_{\text{min},j})
\]

(7) Record the best solution at each generation.

(8) If \(\text{max\_gen}\) is reached or a satisfactory solution is found, then stop; otherwise go to step (3).

2.3. Flowchart of method

(1) The traffic sample data is real number, so input and output data need be transformed to complex values before being inputted into CVNN model. The following transformation steps are used [7]:

- Suppose the input real number vector \([x_1, x_2, \ldots]\), and tally the maximum and minimum values of real numbers (\(\text{max}\) and \(\text{min}\)).
- \(i\)-th real number \(x_i\) is transformed into complex number as followed.

\[
\varphi_i = \frac{x_i - \text{min}}{\text{max} - \text{min}}(2\pi - \delta),
\]

\[
z_i = e^{i\varphi}.
\]

Where \(\delta\) is the shift angle and \(i\) stands for the value of \(\sqrt{-1}\).

(2) Train the complex-valued neural network according to the above complex numbers (input data) and cuckoo search, which is introduced in detail in Section 2.2. The output of CVNN is complex number, so the output value needs to be transformed into real number in order to evaluate model effectively. The inverse transformation shall be used:

\[
\arg z = \varphi,
\]

\[
y = \frac{\varphi(\text{max} - \text{min})}{2\pi - \delta} + \text{min}.
\]

Where \(\arg z\) is the argument of complex value \(z\).

3. Experiments
Small-time scale traffic measurements data is from Meng [3]. The data are normalized in the interval [0, 1] with the following formula

\[ x' = \frac{x - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}} \]  

(9)

The length of this traffic data set is 36000. The front 33000 data points are used as the training set and the last 3000 data points are used as the test set. In the paper, we use the front 10 variable to predict the current variable. The used instruction set for additive tree model

To access the performance of the different methods, the root mean squared error (RMSE), maximum absolute percentage error (MAP) and mean absolute percentage error (MAPE) are used to study the performance of the trained forecasting model for the test data, defined as:

\[ RMSE = \sqrt{\frac{1}{N} \sum_{k=1}^{N} (y_{\text{actual},k} - y_{\text{predict},k})^2} \]

\[ MAP = \max \frac{|y_{\text{actual},k} - y_{\text{predict},k}| \times 100}{y_{\text{predict},k}} \]

\[ MAPE = \frac{1}{N} \sum_{k=1}^{N} \left| \frac{y_{\text{actual},k} - y_{\text{predict},k}}{y_{\text{predict},k}} \right| \times 100 \]  

(10)

where \( N \) is the total number of the data, \( y_{\text{actual},k} \) is the \( k \)-th actual data, and \( y_{\text{predict},k} \) is the \( k \)-th forecast value.

Table 1 summarizes the results achieved using three different methods, and a comparison of actual network traffic time series data and the predicted ones by CVNN model is shown in Fig. 2. From Table 1, it can be clearly seen that CVNN model can effectively improve the accuracy of small-time scale network traffic prediction.

The prediction error of CVNN model is shown in Fig. 3. From Fig. 3, it can see that CVNN could effectively predict the traffic data. The prediction error converges to the small value, even at zero

![Figure 2](image-url)
Table 1. Statistical analysis of three learning methods (test data).

| Criterion | FNT [2] | ODE [3] | CVNN |
|-----------|---------|---------|------|
| MAP       | 159.46  | 157.3   | 151.57 |
| MAPE      | 5.01    | 4.84    | 4.34 |
| RMSE      | 0.0129  | 0.0124  | 0.0113 |

Figure 3. The predicted errors.

4. Conclusions
In order to forecast small-time traffic data accurately, complex-valued neural network model is proposed to further improve the accuracy of time series forecasting. The complex-valued and real-valued parameters are optimized by artificial bee colony algorithm. Small-scale traffic measurements data namely the TCP traffic data is used to test the performance of CVNN model. Experimental results reveal that the CVNN model forecasts the small-time scale network traffic measurement data more accurately and FNT and ODE models. The prediction error converged to the small value, even at zero.

Acknowledgments
This study was funded by Shandong Provincial Natural Science Foundation, China (No. ZR2015PF007), the PhD research startup foundation of Zaozhuang University (No.2014BS13), and foundation of Zaozhuang University (No.2015YY02).

References
[1] Meng QF, Chen YH and Peng YH 2008 Chinese Physics B 18 2196.
[2] Chen YH, Yang B and Meng QF 2012 Applied Soft Computing 12 274.
[3] Chen YH, Yang B, Meng QF, Zhao YO and Abraham A 2010 Information Sciences 181 106
[4] Jiang D, Xu Z and Xu H 2015 Annals of Telecommunications 70 427.
[5] Cheng XY and Jiang MY 2012 Lecture Notes in Computer Science 7331 326.
[6] Karaboga D, Gorkemli B, Ozturk C and Karaboga N 2014 Artificial Intelligence Review 42 21.
[7] AizenbergI, Sheremetov L, Villa-Vargas L and Martinez-Muñoz J 2015 Neurocomputing 175 980.