Neural Networks 57 (2014) 141–151

Contents lists available at ScienceDirect

Neural Networks

journal homepage: www.elsevier.com/locate/neunet

Effect of hybrid circle reservoir injected with wavelet-neurons on performance of echo state network

Hongyan Cui [a,b,*], Chen Feng [a,b], Yuan Chai [a,b], Ren Ping Liu [c], Yunjie Liu [a,b]

[a] *State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, 100876, China*

[b] *Key Laboratory of Network System Architecture and Convergence, Beijing University of Posts and Telecommunications, 100876, China*

[c] *CSIRO, Australia*

ARTICLE INFO

Article history:
Received 23 July 2013
Received in revised form 14 April 2014
Accepted 26 May 2014
Available online 13 June 2014

Keywords:
Reservoir
Echo state network
Low complexity
Wavelet-neurons
Memory capacity

ABSTRACT

The Echo State Network (ESN) has attracted wide attention for its superior performance in chaos time-series prediction. However, the complicated ESN topologies and the random reservoirs are difficult to implement in practice. We propose a hybrid circle reservoir (HCR) ESN architecture that comprises the following features: (1) built with low complexity circle reservoir; (2) partly injected with wavelet-neurons; (3) uses fixed connection weights in both input matrix and dynamic reservoir matrix. The HCR model has been successfully applied to solve six application problems, and the results are used to compare with the existing low complexity simple circle reservoir (SCR) ESN. Furthermore, we analyze the performance of the new model under different ratios of wavelet-neurons, different circle distributions and different input sign patterns. Simulation results show that the HCR model achieves significantly better performance in prediction accuracy than the SCR model. Additionally, the HCR model has similar low complexity as the SCR. Moreover, the short-term memory capacity (MC) in the HCR is close to the theoretical optimal MC value.

1. Introduction

The Echo State Network (ESN) is one of the most popular Recurrent Neural Networks for its perfect performance in modeling the nonlinear dynamic system as well as chaos time-series prediction (Atiya & Parlos, 2000; Cui, Liu, & Li, 2012; Deng & Zhang, 2007). In the ESN, a large complex dynamic reservoir is built to capture a number of features of input streams through the change of reservoir-to-output readout mapping. Connection weights of the input and reservoir matrices are randomly generated. The sparse reservoir matrix brings the echo state property (ESP) to the ESN: the reservoir neurons can transmit and remember the input history information in short time. To guarantee the echo state property of the ESN, the spectral radius of the reservoir generally picks a value below 1. The ESN has been successfully applied in speech recognition (Hénon, 1976), grammatical structure learning (Hong, 1992) and short-term stock price prediction (Jaeger, 2001).

In Jaeger (2002a), Jaeger pointed out the importance of reservoir topology. A number of subsequent studies on the topology of the reservoir have been done to improve the prediction performance (Jaeger, 2002b, 2003, 2005). Complex network topology is applied to the reservoir topology of the ESN that makes ESN prediction algorithms more accurate. In particular, to capture the large number of features of an input dataset, the size of the reservoir becomes huge, the topology of the reservoir becomes complicated, and a large amount of connection matrix weights are randomly generated. However, the complicated topology also increases the complexity of the algorithm and decreases the practicality of the algorithm (Jaeger, 2002b).

In order to reduce the complexity of the ESN, a low complexity echo state network was built in Jaeger, Lukosevicius, Popovici, and Siewert (2007). The low complexity ESN has the following characteristics: (1) a simple nonrandom linear topology is used in the reservoir; (2) a single fixed value is used for the weights of the input matrix; (3) a single fixed value is used for the weights of the reservoir matrix. Three linear topologies of the reservoir were studied in Jaeger et al. (2007): the Delay Line Reservoir (DLR), DLR with feedback connections (DLRB) and Simple Circle Reservoir (SCR). It is proved that the reservoir constructed in a deterministic manner can also obtain good performance. In the three topologies,

* Corresponding author at: Key Laboratory of Network System Architecture and Convergence, Beijing University of Posts and Telecommunications, 100876, China. Tel.: +86 15117939515.

E-mail addresses: cuihy@bupt.edu.cn (H. Cui), ren.liu@csiro.au (R.P. Liu).

http://dx.doi.org/10.1016/j.neunet.2014.05.013
0893-6080/

the SCR performs best in capturing the features of inputs and making the Memory Capacity (MC) close to the proved optimal MC value.

In the field of Neural Networks, the classical Wavelet Neural Network (WNN) is an effective scheme. The standard WNN has three layers, in which the hidden layers employ wavelet-neurons. The hidden layers' transfer function is generated from a mother wavelet function by dilation and translation (Lin, Yang, & Song, 2009; Lu, 2011; Lukoševičius, 2012; Orlowska-Kowalska, Dybkowski, & Szabat, 2010; Pati & Krishnaprasad, 1993; Pindoriya, Singh, & Singh, 2008; Rad, 2008). In Lin et al. (2009), a class of feed-forward networks composed of wavelets was proposed. Lu (2011) introduced the method to use orthonormal wavelets to construct the wavelet-based neural network. Many improved WNN algorithms have been successfully applied in the fields of identification and predictive control (Pati & Krishnaprasad, 1993), energy price forecasting (Pindoriya et al., 2008) and motor drive control (Rad, 2008). Meanwhile, an ESN model with sigmoid-wavelet reservoir was proposed in Rodan and Tino (2011).

In this paper, we propose a hybrid circle reservoir (HCR) ESN architecture by injecting wavelet-neurons into the reservoir to replace parts of the sigmoid-neurons. Our HCR ESN maintains the deterministic topology of the SCR and combines the wavelet-neurons with sigmoid-neurons into one hybrid activation function. The introduction of wavelet-neurons enhances the performance of the ESN due to the nature of the wavelet function. Such a hybrid activation function maps the input data into a rational range of output by scale and shift the data samples, so that optimum magnitude can be achieved as suggested in Schalkoff (1997). This is consistent with the basic characteristics of biological neural networks (Skowronski & Harris, 2007). As such, the proposed HCR ESN can effectively extract the local information and features of the training samples and is able to achieve a higher prediction accuracy.

Compared with the SCR ESN method that only contains one kind of neuron (Jaeger et al., 2007), our new HCR contains both kinds of neurons in the reservoir-common neurons and wavelet-neurons. The newly injected wavelet-neurons increase the diversities of the reservoir and improve the predicting accuracy. Compared with the randomly connected sigmoid-wavelet reservoir ESN proposed in Rodan and Tino (2011), the proposed HCR is built with fixed weights and simple circle connections. As such, it has a shorter runtime than the classical ESN. Essentially the proposed HCR ESN combines the benefits of the SCR of Jaeger et al. (2007) and the wavelet reservoir of Rodan and Tino (2011) to produce a highly accurate and low complexity ESN model.

We also investigate a number of properties that can affect the HCR's performance. In particular, we simulated different wavelet injecting ratios, different sign patterns of the input weights, and different circle distributions of the two types of neurons. We also investigate the effects of dynamic reservoir size, spectral radius, input sign patterns and memory capacity. These four parameters are the key global parameters that determine the HCR's predicting performance.

Many datasets, including a real traffic trace of a campus network in China, have been used to test the performance of the new hybrid low complexity HCR ESN model. The results show that the HCR ESN achieves better prediction accuracy than that of the SCR ESN. Moreover, the proposed HCR ESN maintains the low complexity of the SCR ESN.

The remainder of this paper is organized as follows. Section 2 reviews the classical ESN models including architectures and training processes. The new low complexity HCR ESN model is proposed in Section 3. In Section 4, simulation experiments are conducted to analyze the performance of the new HCR ESN model and to compare with other ESN models. Finally, our work is discussed and concluded in Section 5.

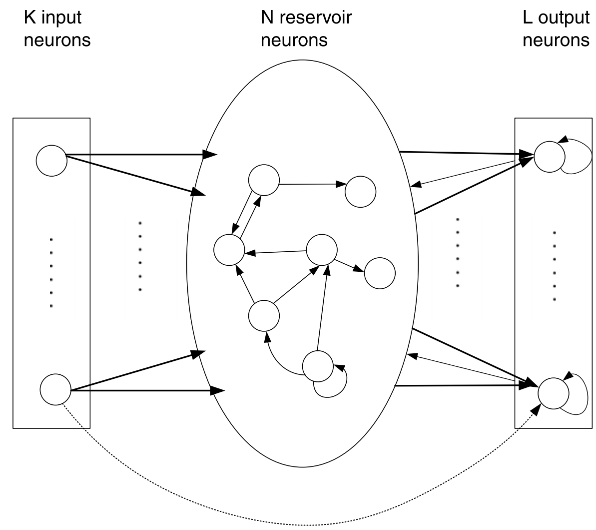

Fig. 1. Architecture of the ESN.

2. Echo state network

The echo state network is one of the most popular neural networks. It is composed of three parts as illustrated in Fig. 1: the left part has K input neurons, the internal (reservoir) part has N reservoir neurons, and the right part has L output neurons. In Fig. 1, the solid arrows represent fixed connection and the dotted arrows represent optional connection. The activation of input, the internal neurons, and the output at time step n are denoted as (1), (2), and (3), respectively:

$$u(n) = (u_1(n), u_2(n) \dots u_K(n))^t \tag{1}$$

$$x(n) = (x_1(n), x_2(n) \dots x_N(n))^t \tag{2}$$

$$y(n) = (y_1(n), y_2(n) \dots y_L(n))^t. \tag{3}$$

The connection weights between input neurons and reservoir are given in a $K \times N$ weight matrix W_{in}. The connections between reservoir neurons are given in an $N \times N$ weight matrix W. The connection weights between all neurons and output neurons are given in a $(K+N+L) \times L$ weight matrix W_{out}, and the connections between output neurons and reservoir neurons are given in an $L \times N$ weight matrix W_{back}. Note that the elements of matrices W_{in} and W_{back} are fixed before training with some random values drawn from a uniform distribution. According to Deng and Zhang (2007), the reservoir weight matrix W must be a sparse matrix with spectral radius below 1 to keep the echo state property. It can be obtained as

$$W = a(W_1/|\lambda_{\max}|) \tag{4}$$

where W_1 is a randomly generated matrix and $\lambda_{\max}$ is the spectral radius of W_1. a is the scaling parameter with a small value between 0 and 1.

The reservoir neurons are updated as follows:

$$x(n+1) = f(W_{\text{in}}u(n+1) + Wx(n) + W_{\text{back}}y(n)) \tag{5}$$

where f is the activation function of neurons. In general, sigmoid functions are used as the activation function of neurons.

The outputs are calculated as

$$y(n+1) = f^{\text{out}}(W_{\text{out}}(u(n+1), x(n+1), y(n))) \tag{6}$$

where f^{out} is the readout function of neurons and W_{out} is the output weight matrix which has been trained.