Neurocomputing 73 (2010) 2540–2553

Contents lists available at ScienceDirect

Neurocomputing

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

Chaotic time series prediction with residual analysis method using hybrid Elman–NARX neural networks

Muhammad Ardalani-Farsa, Saeed Zolfaghari *

Department of Mechanical and Industrial Engineering, Ryerson University, 350 Victoria Street, Toronto, Canada M5B 2K3

ARTICLE INFO

Article history:
Received 10 August 2009
Received in revised form
3 February 2010
Accepted 10 June 2010
Communicated by G.P. Zhang
Available online 25 June 2010

Keywords:
Chaos theory
Embedding theorem
Elman neural network
NARX neural network
Residual analysis
Nonlinear time series prediction

ABSTRACT

Residual analysis using hybrid Elman–NARX neural network along with embedding theorem is used to analyze and predict chaotic time series. Using embedding theorem, the embedding parameters are determined and the time series is reconstructed into proper phase space points. The embedded phase space points are fed into an Elman neural network and trained. The residual of predicted time series is analyzed, and it was observed that residuals demonstrate chaotic behaviour. The residuals are considered as a new chaotic time series and reconstructed according to embedding theorem. A new Elman neural network is trained to predict the future value of the residual time series. The residual analysis is repeated several times. Finally, a NARX network is used to capture the relationship among the predicted value of original time series and residuals and original time series. The method is applied to Mackey–Glass and Lorenz equations which produce chaotic time series, and to a real life chaotic time series, Sunspot time series, to evaluate the validity of the proposed technique. Numerical experimental results confirm that the proposed method can predict the chaotic time series more effectively and accurately when compared with the existing prediction methods.

1. Introduction

Over the last several decades, prediction of chaotic time series has been a popular and challenging subject. Chaos theory as a new area of mathematics has been used to analyze chaotic systems and draw the hidden information from random-look data produced by chaotic systems. Chaotic time series are deterministic systems and inherit a high degree of complexity. Although, chaotic time series show the characteristic of dynamical systems as random, in the embedding phase space they present deterministic behaviour [66].

The chaos theory, as an essential part of nonlinear theory, has provided an appropriate tool to illustrate the characteristics of the dynamical systems and predict the trend of complex systems. There are four fundamental characteristics for chaotic systems: *aperiodic* that is the same state will not be repeated, *bounded* meaning that neighbour states remain within a finite range and does not approach infinity, *deterministic* that there is a governing rule with no random term to predict the future state of the system, and *sensitivity to initial conditions* meaning that small difference in initial conditions will cause two points close to each other diverge as the state of system progress [62].

Takens' [53] embedding theorem is an essential element of chaotic time series analysis. A set of single observations from a chaotic system can be reconstructed into a series of D-dimensional vectors with two parameters of time delay and dimension. Based on Takens' theorem, if dimension is large enough, the reconstructed vectors exhibit many of the significant entities of the time series [12].

Prediction of nonlinear time series is a useful method to evaluate characteristics of dynamical systems. Prediction of chaotic time series have been observed in the areas of marketing system [44], foreign exchange rate [5], signal processing [22], supply chain management [61], traffic flow [39], power load [48], weather forecast [31], Sunspot prediction [42] and many others. Due to the importance of these fields, the interests in a robust technique to predict chaotic time series have been increased.

A number of techniques to predict chaotic time series have been introduced in the literature. A method of local modelling was proposed by McNames [36]. Wichard and Ogorzalek [60] described the use of ensemble methods to build proper models for chaotic time series prediction. Zhang et al. [67] proposed a multi-dimension prediction method using Lyapunov exponents.

Artificial neural networks (ANNs) have been also employed independently or as an auxiliary tool to predict chaotic time series. ANNs are nonlinear methods which mimic nerve system [68]. They have functions of self-organizing, data-driven, self-study, self-adaptive and associated memory [64,68]. ANNs can learn from patterns and capture hidden functional relationships in a given data even if the functional relationships are not known or difficult to identify [64]. Using the training methods, an ANN can be trained to identify the underlying correlation between the

* Corresponding author.
E-mail address: zolfaghari@ryerson.ca (S. Zolfaghari).

0925-2312/$ - see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.neucom.2010.06.004

inputs and outputs. Later, the unseen inputs can be fed to the trained ANN to generate appropriate outputs [11,17,41,43,64].

A number of researchers have utilized ANN to predict chaotic time series. Multi-layer perceptron neural network (MLP) has been used by Liu et al. [28] and Park et al. [42]. Zhang and Man [66], Tenti [55], Ma et al. [32], and Assaad et al. [2] have utilized recurrent neural network (RNN). Nonlinear Autoregressive model with eXogenous input (NARX) has been also applied to chaotic time series perdition by Menezes and Barreto [38] and Diaconescu [8].

Some other artificial intelligence (AI) methods such as radial basis function network (RBF) [54,46], self-organizing map (SOM) [3,26,51], support vector machine (SVM) [47,58], fuzzy and neuro-fuzzy [14,65,24], and wavelet neural networks (WNN) [13,59] among the others are used in the literature to forecast chaotic time series.

In prediction methods, the analysis of residuals is under-estimated. In some occasions, residuals are not due to randomness; therefore, residuals show high correlation meaning that the prediction model has not completely captured the characteristics of the system. There are occasions that residuals inherit the characteristic of original system.

This study investigates the contribution of residual analysis to increase the performance of prediction method. The proposed method utilizes the embedding theorem to "unfold" the chaotic time series and reconstruct the phase space points. Two well-known dynamic neural networks, Elman and NARX networks are selected for training purposes. An Elman Neural Network is trained using gradient descent with momentum and adaptive learning rate backpropagation algorithm to predict the future value of the obtained phase space points and accordingly the original time series. Normally, the residuals of predicted time series show high degree correlation and demonstrate chaotic behaviour. Therefore, the residuals are considered as a new chaotic time series and similarly analyzed and predicted. The residual analysis is repeated several times. Finally, a NARX neural network is trained to capture the relationship among the predicted value of original time series, residuals and original time series. The weights and biases of NARX neural network is kept to predict the future value of original time series. The block diagram in Fig. 1 demonstrates the methodology developed in this study to forecast chaotic time series.

The paper is organized as follows. Section 2 describes chaotic time series and the method to reconstruct a time series. Section 3 briefly discusses dynamic neural networks. Section 4 illustrates the proposed prediction technique in detail. In Section 5, the prediction performance of the proposed technique on three well-known chaotic time series, the Mackey–Glass, Lorenz and Sunspot, are studied. Finally, conclusions are given in Section 6.

2. Chaotic time series and embedding theorem

Many natural systems show nonlinear or chaotic behaviour. Using chaos theory, these systems have been described by mathematical equations. For a chaotic system, the phase space is defined as a vector space R^n with each point in the phase space being described by a n-dimensional vector $s(t)$, which is required to obtain the progression of the system [40]. $s(t)$ is defined as

$$s(t)=[s_1(t),s_2(t),s_3(t),\ldots,s_n(t)] \tag{1}$$

where t is an index for the time series and n is the dimension of vector space R^n. With the use of the nonlinear function F: $R^n\rightarrow R^n$, which describes the system, the future value of the system at time $t+\tau$ can be determined by

$$s(t)\rightarrow F(s(t))=s(t+\tau) \tag{2}$$

A small change in the state of the system, $s(t)$, will substantially influence the trend of the system and after several iterations, the system becomes unforeseeable. This behaviour of dynamical systems is known as sensitivity to initial conditions or butterfly effect [12,31].

The progression of a non-random system creates a trajectory named an attractor. Takens' [53] embedding theorem states that because the value of $s(t)$ and its components, $s_1(t)$, $s_2(t)$, $s_3(t)$,... in a chaotic system are unknown; if one is able to observe a single quantity or variable $x(t)$ from this dynamical system, then the attractor can be unfolded from this set of observed samples [1]. This means that if a single quantity $x(t)$ is observed from a chaotic system, the reconstructed dynamics of a system $Y(t)=[x(t),\ x(t-T),\ x(t-2T)\ldots]$, with T defined as time delay, is geometrically similar to the original attractor. Therefore, if a dynamical system $s(t)\rightarrow s(t+1)$ exists, then sequential order of reconstructed phase space points $Y(t)\rightarrow Y(t+1)$ follows the unknown dynamics of $s(t)\rightarrow s(t+1)$. Therefore, the behaviour of the actual system is reflected in the observed time series generated from the system [21].

2.1. Determining chaos in time series

In analyzing time series, an important step is to determine the characteristic of the data. The following methods have been used to differentiate periodic or random data from chaotic data.

2.1.1. Fourier transform

Fourier transform can be used to identify chaos in a given time series. It is common to plot power spectrum instead of frequency spectrum. The power spectrum spikes at frequencies that characterize the system for periodic data, and will be approximately zero for the others. The broadband power spectrum with broad peak proves the existence of chaos in the time series [12].

2.1.2. Lyapunov exponent

An important characteristic of chaotic systems which is defined as the butterfly effect is the high sensitivity of the system to the initial conditions. If the high sensitivity to initial conditions is detected in a system, the system can be considered chaotic [63]. Largest Lyapunov exponent is the most practical method to identify chaotic behaviour in a system. Lyapunov exponent quantify the divergence of neighbouring trajectories. Positive Lyapunov exponent proves the existence of chaos in the system [12].

2.1.3. Hurst exponent

Hurst is known for introducing Hurst exponent as a measurement for the predictability of a time series [16]. Hurst exponent is derived using R/S analysis. Hurst exponents can change between 0 and 1. Hurst exponent of 0.5 shows a random walk. A Hurst exponent between 0.5 and 1 proves the presence of chaos in the system.

2.1.4. Fractal dimension

Another method to identify the existence of chaos in a system is Fractal dimension. Non-integer fractal dimension shows that the system is chaotic. Correlation dimension is one of the most common Fractal dimension used in literature [12,52]. If a sphere of radius R is centred on a specific point in D-dimensional space, then the mean of points in the sphere, $C(R)$, excluding the center point can be calculated. A plot of $C(R)$ versus R should give an approximately straight line whose slope is d_c , the correlation dimension. d_c with integer value shows that the attractor is a