TIME SERIES PREDICTION USING RNN IN MULTI-DIMENSION EMBEDDING PHASE SPACE

Jun Zhang **K.F. Man**

Department of Electronic Engineering
City University of Hong Kong
Tat Chee Avenue, Kowloon, Hong Kong

ABSTRACT

In this paper, a multi-dimension chaotic time series prediction method using Recurrent Neural Network (RNN) in embedding phase space is proposed. This method is to reconstruct a phase space based on the chaotic time series and then embed these data as the phase space points for the training of the RNN. The resultant RNN after the training will be served as the embedding phase space which is capable of recovering the predicted phase space point into time domain. Thus, the predicted chaotic time series data can be obtained. Numerical results have shown that the proposed method is simple, practical and effective in chaotic time series prediction.

1. INTRODUCTION

Nonlinear and chaotic time series prediction is a practical technique which can be used for studying the characteristics of complicated systems based on recorded data. Application of this nature can be found in the areas of economic and business planning, inventory and production control, weather forecasting, biological and medical engineering, mechanical devices, signal processing, industry and automation control and many other fields.

As the most real-world systems are nonlinear, chaotic signals are often found in radar clutter, speech, and so on [1][2]. As a result, the interests in chaotic time series prediction have been increased. Considering the uniqueness of a chaotic system, given with its initial conditions, the long term prediction is not possible. On the other hand, the proposition to capture a short term chaos behavior with a certain accuracy is viable. This is based on the fact that a chaotic time series is characterized by values that appear to be randomly distributed and non-periodic, but is actually the result of a completely deterministic process. This characteristics makes the prediction of chaotic time series very challenging [3] [4].

A number of techniques for chaos prediction have been employed in the past. The auto-regression [5], ARMA model and artificial neural network (ANN) have achieved some degrees of success. All these techniques are time domain dependence and fall victim to the undeterministic behavior of a chaotic time series.

In this sphere, Framer and Sidorowich[6] proposed a state-space reconstruction method based on the geometry of the embedded data for the discovery of a relationship between past and future points in a time series. The approach uses only the local dynamics ('local linear' methods) on an attractor which can produce very impressive results. The evolving nonlinear dynamic forecasting [7] further improved the predicting quality as well as its robustness in the presence of extraneous noise.

Being different from above methods, a new method to predict the chaotic time series is proposed in this paper. The inspiration is from the fact that though chaotic time series shows the behavior of system as a random signal in time domain, it displays some deterministic behavior in the embedding phase space (EPS). In this paper, an innovative embedded phase space (EPS) method is proposed. This method utilizes a phase space model based on the chaotic data to form the EPS nodes for the training of an equivalent recurrent neural network (RNN). This resultant RNN model will serve the purpose of predicting the related chaotic time series.

2. EMBEDDING PHASE SPACE RECONSTRUCTION

The EPS reconstruction of a chaos is an important step for the study of chaotic time series. Consider the embedding theory [8], from an observed time series $x(t)$, a data vector is generated as $Y(t)=F(x(t),x(t-T),\ldots, x(t-(D-1)T)$ where T is the time delay, the vector $Y(t)$ indicates one point of a D-dimensional reconstructed phase space R^D. D is called the embedding dimension.

Assume that the target system is a deterministic dynamical system and that the observed time series is obtained through an observation system corresponding to a continuous mapping C^1 which is the state space of the dynamical system in the 1-dimension Euclidean space R, while C^1 is first order differential. Then, the reconstructed trajectory is the embedded original trajectory when D is sufficiently large. When any attractor appears in the original dynamical system, another attractor, which also retains the phase structure of the first attractor, will appear in the reconstructed state space.

3.MULTI-DIMENSION METHOD

In this paper, instead of doing more work on the aspect of "how to select embedding parameters ", we explore the method of predicting chaotic time series in EPS using **RNN** based on some seniors' achievements[4][6][8][9][10][11][12][17]. It is shown from the results of these researchers that there is not an exact embedding dimension which is the best value to reconstruct phase space and chaos attractor, i.e. chaotic time series data can be embedded to more than one phase space of different dimension while keeping the chaotic attractor behavior. From this point, the idea of multi-dimension EPS method is developed to improve the reliability of chaotic time series prediction. The basic steps are:

1. Use general method to estimate minimum embedding dimension ***D***.
2. Predict the unknown phase space points on different unfold attractor trajectory in different phase space (e.g., ***D-1, D, D+1***).
3. Average the predicting chaotic time series data using least-squares method. Then the predicting result is relatively more accurate than the general single dimension phase space prediction.

This multi-dimension method can be explained in error calculations theory[13].

For the prediction in n different dimensions, the predicted value obtained from n different dimension prediction is respectively $Y_1, Y_2, \ldots Y_n$. The average value is X; the error for different dimension is $m_1, m_2, \ldots, m_n$. The apparent error is $v_i = Y_i - X$ $(i=1,2,\ldots,n)$. Then, the probability of the apparent error is[13]:

$$P(v_i) = \frac{1}{m_i\sqrt{2\pi}} e^{-\frac{v_i^2}{2m_i^2}dv_i} \qquad (1)$$

As the predicted value is independent, the probability of all predicted value is $\prod_{i=1}^{n} P(v_i)$. The larger probability P of the predicted value is, the more reliable the prediction is, i.e. while $\sum_{i=1}^{n} v_i^2$ is the minimum value, the most reliable predicting value can be obtained.

Therefore, the multi-dimension prediction value is calculated as:

$$Y_{Multi-\dim ension} = \frac{\sum_{i=1}^{n} p_i Y_i}{\sum_{i=1}^{n} p_i} \qquad (2)$$

where n is the number of multi-dimension, Y_i is the predicted value in i embedding dimension, p_i is the i^{th} dimension probability (generally $p_i=1$ for independent process). By this method, the prediction error can be decreased to get more reliable result.

4. RECURRENT NEURAL NETWORK MODEL

RNN has been shown to be very useful in predicting nonlinear ARMA time series [14]. It not only has the capability for handling a system of much higher complexity, but its superiority in time convergence can prove to be a valuable asset for time critical application. Elman [15] introduced a simple RNN, which is shown in Fig. 1.

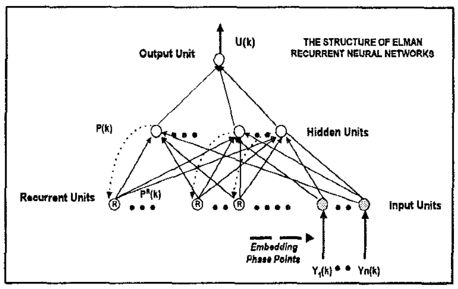

Fig.1 The Structure of Elman Recurrent Neural Networks

It can be seen from the above figure that there are four units[16]: input units, hidden units, output units and the recurrent units in a basic Elman network. The input and output units connect with the outside environment, whereas the hidden and recurrent units are contained within. The input units form a buffer which passes the signals without changing, while the output unit is a linear unit which sums up the feed-in signals. The hidden units have nonlinear activation functions. The recurrent units are employed to memorize the previous activation of the hidden units and can be considered as a function of a one-step time delay. The feed-forward connections are modifiable; while the recurrent connections are fixed.

At a specific time k, the previous activation of the hidden units (at time $k-1$) and the current inputs (at time k) are used as inputs to the network. At this stage, the network acts as a feedforward network and propagates these inputs forward to produce the output. The standard BP learning rule can be employed to train the network. Once this training step has been started, the activation of the hidden units at time k are sent back through the recurrent links to the recurrent units and the information are stored for the next training step $(k+1)$.

The external input to the network is $Y(k)$ and the network output is $U(k)$. The total input to the i^{th} hidden unit is denoted as $v_i(k)$. The output of the i^{th} hidden unit is denoted as $P_i(k)$. The output of the j^{th} recurrent unit is $P_j^R(k)$. The following equations hold

$$v_i(k) = \sum_{j=1}^{n} w_{i,j}^P(k-1)P_j^R(k) + w_i^Y(k-1)Y(k) \quad (3)$$

$$P_i(k) = f(v_i) \quad (4)$$

$$P_j^R(k) = P_j(k-1) \quad (5)$$

$$U(k)= \sum_{i=1}^{n} w_i^U(k-1)P_i(k) \quad (6)$$

where $w_i^Y(\bullet), w_{i,j}^P(\bullet)$ *and* $w_i^U(\bullet), i,j = 1,2,\ldots,n$ are the weights of the links, respectively, between the input units and hidden units, between the recurrent units and the hidden units, and between the hidden units and the output units. f is a sigmoid activation function. The Elman network is trained by the standard back propagation (BP) algorithm. The objective of the recurrent neural networks is to minimize the squared error function:

$$E_k = \frac{1}{2}(U_d(k) - U(k))^2 \quad (7)$$

where $U_d(k)$ is the desired output of the network. The learning mechanism is the traditional gradient descent learning method. The activation function f is differentiable. The weight updates are based on the gradient of the error which is defined in terms of the weights and activation functions. The form of the update rule is given as a solution to the following equation:

$$\Delta w_{ij} = \eta \frac{\partial E}{\partial w_{ij}} \quad (8)$$

where η is the parameter of learn rate and w is the weight of the connection between units.

5. CHAOTIC TIME SERIES PREDICTION IN EMBEDDING PHASE SPACE

5.1 The proposed algorithm

Based on a chaotic time series data set $x(1), x(2), x(3), \ldots x(N)$, where N is the length of the data set. If the embedding dimension is selected as D and the time delay as T, there are $[N-(D-1)T]$ phase space points $Y(I)$, $I=1\sim[N-(D-1)T]$, are generated in the phase space:

$Y(1)=[x(1)\ x(1+T)\ \ldots\ldots\ x(1+(D-1)T)];$

$Y(2)=[x(2)\ x(2+T)\ \ldots\ldots\ x(2+(D-1)T)];$

$Y(3)=[x(3)\ x(3+T)\ \ldots\ldots\ x(3+(D-1)T)];$

....

$Y(I)=[x(N-(D-1)T)\ x(N-(D-1)T+T)\ \ldots\ldots\ x(N)];$

Then we get $N-(D-1)T$ EPS points. The RNN is trained using those obtained points. For Example: Using $Y(1)\sim Y(K)$ as input points and $Y(K+M)$ as the output to train the RNN where M is the prediction step, K is input number of RNN and I is the number of known phase space points, while $K+M<I$, $K<I$. After training the RNN, keep the weight of RNN to serve the prediction of unknown phase space points. Based on the known $Y(n), Y(n+1), \ldots, Y(n+K-1)$ as input phase space points, where $n+K-1<I$, $n \in [1 \quad I]$, the unknown $Y(n+K-1+M)$ is to be predicted, $n+K-1+M>I$. After recovering $Y(n+K-1+M)$ to the time domain, $[x(n+K-1+M), x(n+K-1+M+1), \ldots, x(n+K-1+M+(D-1)T)]$ can be obtained. Up to here the process of predicting chaotic time series is completed. The detailed steps of proposed algorithm (Algorithm Pred_RNN) is illustrated as the following steps:

Step one	*Load a known time series data set **x(t)**. The length of the data set is **N**;*
Step two	*Choose the time delays, **T**;*
Step three	*Choose the dimension of reconstructed phase space, **D**;*
Step four	*Reconstruct phase space based on the chose **T** and **D**;*
Step five	*Use known phase space points to train the RNN. After the training process, keep the weights of RNN;*
Step six	*From known embedding phase space points, predict unknown phase space points, using trained RNN.*
Step seven	*Recover the predicted phase space points to time domain. Predicted chaotic time series data are obtained.*
Step eight	*Change another different embedding dimension and repeat from step two to step seven. Apply multi-dimension calculate technique to get more accurate prediction result;*
Step nine	*Stop.*

5.2 Error Measurement

In order to evaluate the prediction accuracy, the mean square error (MSE) criterion is adopted:

$$MSE = \sum_{i=1}^{N}(prediction\_value(i) - real\_value(i))^2 \cdot p_i \quad (9)$$

where $prediction\_value(i)$ is the prediction value of the i^{th} chaotic time series point;

real_value(i) is the real value of the i^{th} chaotic time series point;

p_i is the probability of the i^{th} chaotic time series point; and

N is the number of chaotic time series points to be predicted.

6. NUMERICAL EXPERIMENTS

In order to prove the validity of the proposed method, the chaotic system Mackey-Glass Equation is used as an example for testing the proposed chaotic time series prediction method. Due to the visual limitation, only the prediction results in two and three dimensions are shown. The result of chaotic time series prediction in higher dimension (> 3) can not be illustrated.

The Mackey-Glass(M-G) equation was first advanced as a model of white blood cell production. It is a time-delay differential equation, namely,

$$\frac{dx}{dt} = \frac{ax(t-\tau)}{[1+x^c(t-\tau)]} - bx(t) \qquad (10)$$

where the constants are taken to be a=0.2, b=0.1 and c=10. The delay parameter τ, which determines the nature of the chaotic behavior displayed by the time series. A summary of the behavior of this dynamic system is given by [19] and tabulated in Table 1.

Delay τ	Trajectory
τ<4.43	fixed point attractor
4.53< τ<13.3	stable limit cycle attractor
13.3< τ<16.8	double limit cycle attractor
τ>16.8	chaos

Table 1

As the M-G time series has been widely used as a standard benchmark for prediction algorithms, in this paper, chaotic time series is generated by integrating the M-G equation with τ =17 and initial value $x_0 = 1.2$.

Algorithm [Pred_RNN] is implemented. For step one, a chaotic time series data set with length of 2400 is generated by the above Eq. (10). For Step 2 and 3, the parameters ***T*** and ***D*** are not calculated but instead chosen according to previous report [11]. The chosen parameters of EPS shown in Table 2:

Phase Space Reconstruction parameters		Recurrent Neural Networks Parameters		
Embedding dimension	Time delay	Input units	Recurrent/ Hidden units	Output units
2~8	7	30	25	1

Table 2

Single step prediction is carried out. The chaotic time series data in the time domain is shown as Fig.2.

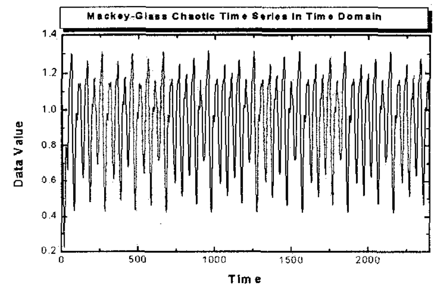

Fig.2 M-G Chaotic Time Series in Time Domain

6.1 Chaotic time series prediction in 2-dimension phase space

Reconstructed phase space in 2-dimension using the chaotic time series data is shown as Fig3:

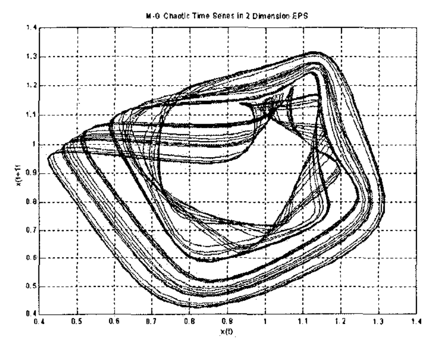

Fig.3 M-G Chaotic Time Series in 2 Dimension EPS

Prediction results in 2 dimension phase space are shown in the following Fig4~Fig5:

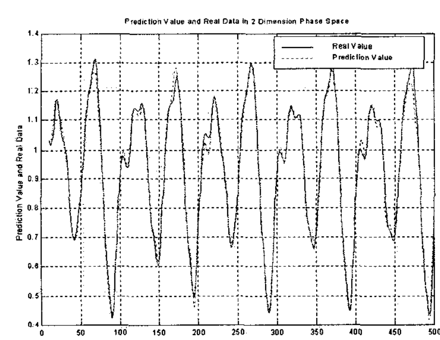

Fig.4 The Comparison of Real Value and Predicted Value

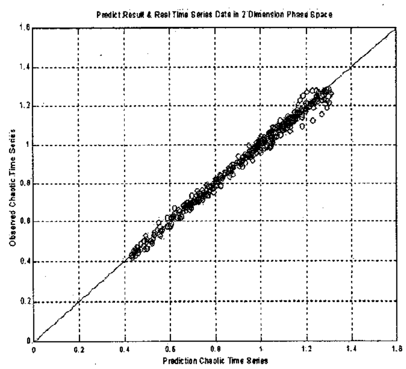

Fig.5 The Relative Comparison of Real Value and Predicted Value

From Eq.9 the predicting error is calculated: ***MSE= 0.00101569333349.***

6.2 Chaotic time series prediction in 3-dimension Phase space:

The EPS points in 3-dimension phase space are shown as Fig.6:

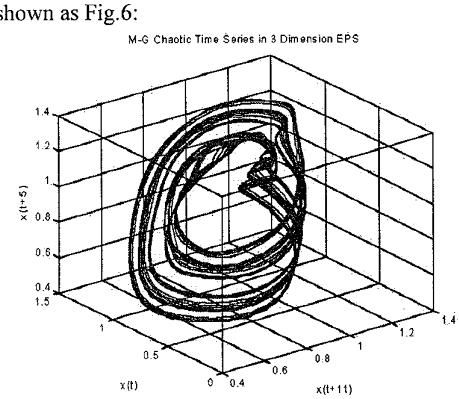

Fig.6 M-G Chaotic Time Series in 3 Dimension EPS

Prediction results in 3 dimension phase space are shown in following Fig7~Fig8:

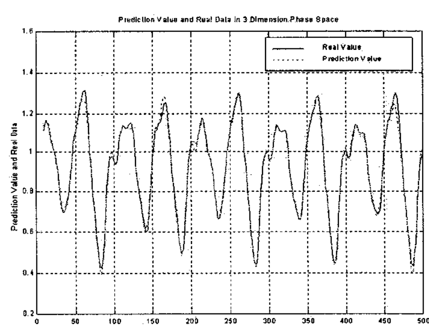

Fig7 The Comparison of Real Value And Predicted Value

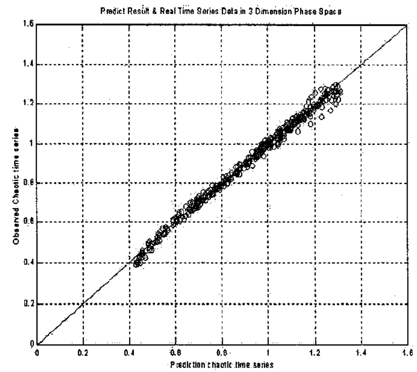

Fig.8 The Relative Compare of Real Value And Predicted Value

From Eq.9 the predicting error is calculated: ***MSE=0.0004147535718725430.***

6.3 Chaotic time series prediction in multi-dimension Phase space:

To ***step eight***, the prediction error from the multi-dimension EPS based on the2,3,4,5,6,7,8 dimension predicting results is calculated. the multi-dimension prediction methods is described in **Sec. 4**.

Dimension	MSE
2	0.0010156933
3	0.0004147535
4	0.0021206503
5	0.0045985557
6	0.0030765558
7	0.0054079865
8	0.0075912217

Fig.9 and Fig.10 are the prediction results in multi-dimension.

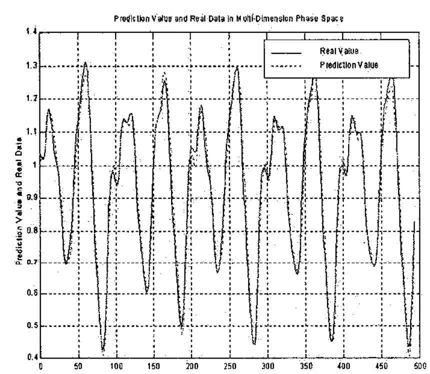

Fig. 9 The Comparison of Real Value and Predicted Value

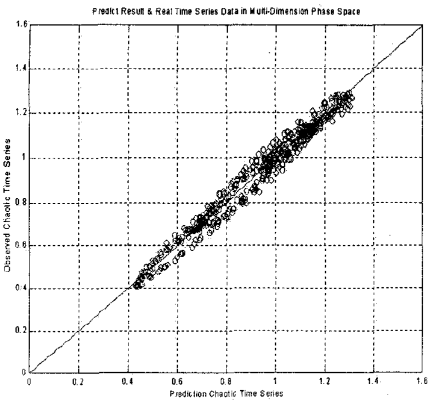

Fig. 10 The Relative Comparison of Real Value and Predicted Value

The prediction error of multi-dimension phase space is: ***MSE= 0.000381379491148365O.***

7. DISCUSSION AND CONCLUSION

In this paper, the fundamental characteristics of chaotic time series is analyzed, the structure of RNN is introduced and a new prediction method is developed. In order to verify the effectiveness of this method, it has been applied to the short-term prediction of M-G chaotic time series, and it has been proven that a satisfactory result is obtained. From the results of numerical experiment, conclusion can be drawn as follows:

1. Chaotic time series displays some stochastic behavior in time domain but displays some deterministic behavior in EPS. This characteristics is exploited and a new predicting method is developed;
2. It can be seen from the experimental results that it is feasible and effective to make short term prediction for chaotic time series using **RNN** in phase space based on the fundamental characteristic of a chaotic system in phase space. It gains some predictive power for chaotic time series; and
3. Utilizing multi-dimension EPS method to predict chaotic time series can improve the accuracy of the prediction. Multi-dimension method preliminarily solve the problem that there is not an exact embedding dimension which is the best to the reconstruct phase space and chaos attractor from a given chaotic time series.

REFERENCES

[1] Henry Leung, & Titus Lo [1993], *Chaotic Radar Signal Processing Over Sea*, IEEE Journal Of Oceanic Engineering vol.18 No.3 July

[2]Nan He & S. Haykin [1992] Chaotic Modelling Of Sea Clutter ,Electronics Letters 22nd October Vol28 No22 pp2076~2077

[3]Leon O. Chua, Makoto Itoh, Chai Wah Wu [1997], *Communication Systems Via Chaotic Signals From A Reconstruction Viewpoint* , International Journal Bifurcation And Chaos, Vol.7, No.2 pp275-286

[4]Kathleen T, Alligood [1996], *Chaos, An Introduction To Dynamical Systems*, Springer.

[5]Yule, G. [1927] *On A Method of Investigating Periodicity In Disturbed Series with Special Reference To Wolfer's Sunspot Numbers*, Phil. Trans. Roy. Soc. London A 226: pp.267-298

[6]Farmer, J. D., Sidorowich, J. J. [1987] *Predicting Chaotic Time Series*, Phys. Rev. Letter. 59, pp.845-848

[7]Kevin M. Short [1997], *Detection of Teleseismic events in seismic sensor data using nonlinear dynamic forecasting* , International journal of bifurcation and chaos, Vol.7, No.8 pp1833-1845

[8]F. Taken [1981] *Detecting Strange Attractors In Turbulence, Dynamical Systems And Turbulence*, 1980 Lecture Notes in Mathematics 898 (Springer, Berlin) pp366-381

[9]Andreas S. Weigend, [1994] Time Series Prediction: Forecasting The Future And Understanding The Past (Addison-Wesley Publishing Company) pp. 131-138, 173-344, 387-414

[10]Casdagli, M.[1989] *Nonlinear Prediction of Chaotic Time Series*. Physica D 35, pp.335-356

[11]Alan Wolf [1985] Determining Lyapunov Exponents From A Time Series, Physics 16 D pp.285-317

[12]Henry D.I. Abarbanel [1995] Analysis of Observed Chaotic Data, (Springer) pp13-69

[13]Hans-Jochen Bartsch [1974] *Handbook of Mathematical Formulas*, Academic Press New York

[14]Connnor, J., Atlas, L.[1991] *Recurrent Neural Networks And Time Series Prediction*, Proc. IJCNN, Seattle, ppi301-i306

[15]ELMAN, J.L. [1990] *Finding Structure In Time. Cognitive Science*, 14,pp179-211

[16]D. T. Pham, X. Liu [1996] *Training of Elman Networks And Dynamic System Modeling*, International Journal Of System Science, Vol.27, No.2, pp. 221-226

[17]Martin Casdagli, Stephen Eubank [1991] *Nonlinear Modelling and Forecasting*, Addison-Wesley Publishing company pp1~71