Journal of Applied Sciences 4 (4): 675-679, 2004
ISSN 1607-8926


# Prediction Method of Machinery Condition Based on Recurrent Neural Networks Models

Guangrui Wen and Xining Zhang
State key Laboratory for Manufacturing System Engineering,
Research Institute of Diagnostics and Cybernetics, Xi'an Jiaotong University,
710049 Xi'an, Peoples Republic of China

**Abstract:** In order to overcome the disadvantage of traditional feedforward neural networks in long-term prediction of machinery condition, a new neural model, so called the multi-step recurrent prediction model based on recurrent neural network is proposed. A learning algorithm of recurrent model for long-term prediction is also presented, which is supposed to obtain better predictions of machinery condition. The feasibility of recurrent neural model is examined by applying it to forecast a simulated time series and predict the behavior of large rotating machinery. Experimental results revealed that the recurrent model could achieve better prediction accuracy and provided cogent proof for realization of prognostic maintenance.

**Key words:** Multi-step prediction, feedforward neural networks model, recurrent neural networks model, machinery condition prediction

## INTRODUCTION

Time series prediction is a major goal in many research areas, e.g. biology, physics, business and engineering. Especially it plays an important role in the area of engineering diagnostics. The ability to forecast the behavior of large rotating machine system can make enterprise obtain remarkable economic returns[1]. Multi-step prediction is a difficult task and has attracted more and more practitioners in past years. The aim of this work is to develop a nonlinear neural model for multi-step time series prediction schemes[2-4] instead of neural models based on the traditional feedforward neural networks. The latter may present some disadvantages in long-term prediction problem. In this study, a new multi-step forecasting model based on recurrent neural networks is presented and a corresponding learning algorithm is also studied. The prediction results for a simulated time series and practical measured peak-to-peak values from the vibration data of large rotating machinery all contribute the improvement of prediction accuracy.

## Models

**Multi-step forecasting method based on feedforward neural network models:** The neural models most widely used in time series applications are based on feedforward neural networks with backpropagation learning algorithm. This model usually constructs multi-layer feedforward neural network, and then evaluate function F by time series. These models consist of a common nonlinear auto-regressing model appearing in Equation (1), which is also called NAR Model. In this case, The value at k+1 of this time series is often represented by d+1 Entries time series values x (k), L, x (k-d), thus:

$$x(k+1) = F(x(k),\ldots, x(k-d)) \tag{1}$$

where, k represents the time variable, F is a nonlinear function that defines this time series.

The common prediction method based on neural network, namely the single-step neural network prediction method, set up non-linear prediction model of the neural network. The following prediction equation represents the prediction model:

$$x(k+1) = \sum_{i}^{d} f_i(x)x(k-i) + \varepsilon \quad i = 0,\ldots,d \tag{2}$$

where, $f_i$ (x) is the non-linear function of input variable, d+1 is the number of input nodes in prediction network. During the neural network prediction, the following equation can be got from the prediction model:

$$\hat{x}(k+1) = \sum_{i}^{d} f_i(x)\, x(k-i) \tag{3}$$

According to the prediction model, Fig. 1 shows the structure of the Multi-step prediction based on

**Corresponding Author:** Guangrui Wen, State key Laboratory for Manufacturing System Engineering, Research Institute of Diagnostics and Cybernetics, Xi'an Jiaotong University, 710049 Xi'an, Peoples Republic of China E-mail: grwen@mail.xjtu.edu.cn

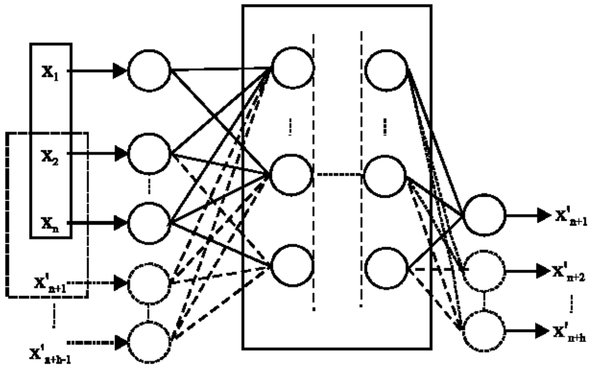


Fig. 1: Multi-step prediction model based on traditional multilayer feedforward neural network

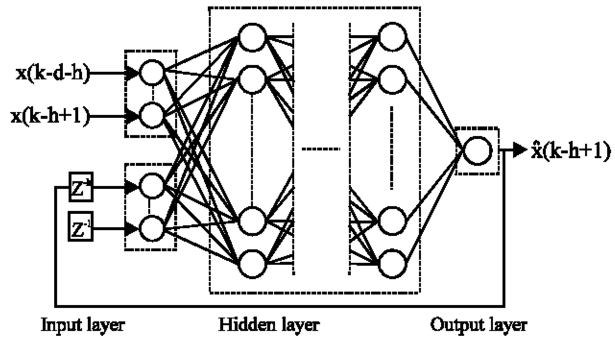


Fig. 2: Recurrent network architecture: $\hat{x}$neural network is divided into three parts: input layer, hidden layer and output layer. Where, represents the network output and $Z^{-1}$ is an operator that delays by I terms the network output sequence. The input layer is composed of two groups of neurons. The first group acts as the external inputs to the network gathering the original or measured time series data. The second group is formed by the context neurons, which memorize previous output of the network

traditional neural networks model. Using the predicting value into the prediction model, the multi-step value of the future can be got step by step. In this kind of multi-step predicting process, when the single-step predicting result is got, there has being a predicting error which can be expressed as:

$$\varepsilon_{k+1} = \frac{1}{2}[x(k+1) - \hat{x}(k+1)]^2 \qquad (4)$$

During the course of inference, using $\hat{x}(k+1)$ and the following predicting value as the input of the network leads to the input error and as the increasing of the predicting value, the accumulating error increases rapidly, which cannot ensure the accuracy of multi-step prediction.

**Multi-step prediction model based on recurrent neural networks and its learning algorithm:** In this section, a multi-step prediction method based on recurrent neural networks is introduced. The recurrent neural networks used to build up the multi-step prediction model consist of adding feedback connections from the output neuron to the input layer, which memorize previous prediction values[5]. In this case, parameters of the multi-step prediction model are determined to minimize the error along interval [k+1, k+h+1], where, h is the length of multi-step prediction. Thus, the model is trained for long-term prediction.

**Description of multi-step recurrent prediction model:** In order to overcome limitation of traditional feedforward multi-step prediction model and make full use of the input vectors to the model, a new multi-step prediction model based on recurrent networks is built up[6,7]. The recurrent network is constructed by starting from a multilayer feedforward neural network and by adding feedback connections from the output neuron to the input layer as shown in Fig. 2.

Introducing the vector C(k) [$C(k) = (C_1(k),..., C_h(k))$] to indicate the activation of context neurons, each component is calculated as:

$$C_i(k) = Z^{-1}(\hat{x}(k+h+1)) = \hat{x}(k+h+1-i) \quad i=,1\ldots,h \quad (5)$$

Assuming that the prediction length is fixed to h, and that at instant k the goal is to predict the time series values at instants k+1, k+2,...,k+h+1, the number of input units decreases from d+1 to d+1-h and the number of context neurons increases from 0 to h, respectively. Thus, the sequences received by the external inputs and the context neurons, at every instant k, are given by the following sequence:

1. The number of context neurons is initialized from zero, which will be included in the external inputs accordingly.
2. The future instants k+I for I=2, k, h+1 are not real, but simulated. Now the input units receive the vector x (k+1),...,x (k-d-i) and the (i-1)-th context neurons memorize the previous i-1 outputs of the network, i.e.:

$$C_1(k) = \hat{x}(k+i-1)\ldots \quad C_{i-1}(k) = \hat{x}(k+1) \qquad (6)$$

3. The external inputs and the context neurons are resettled.

**Learning algorithm:** Below are the complete training procedures of a multi-step recurrent prediction model. At each instant k, starting with k=d:

**Step 1:** The number of context neurons is initialized to zero. d+1 external input neurons are one set receiving the measured values of the time series, x (k),..., x (k-d). The output of network is given by following equation:

$$\hat{x}(k+1) = \hat{F}(x(k),\ldots,x(k-d),W_2) \tag{7}$$

**Step 2:** The number of context neurons is increased in one unit and the number of external units is decreased also in one unit. The context neuron memorizes the output of the network previously calculated, $\hat{x}(k+1)$. Thus the prediction at the simulated instant k+3,...,k+h+1 is given by:

$$\hat{x}(k+2) = \hat{F}(\hat{x}(k+1),x(k),\ldots,x(k-d+1),W_2) \tag{8}$$

**Step 3:** Step 2 is repeated until h context neurons will be achieved. The output of the recurrent model at simulated instants k+3,..., k+h+1 are given by the following equations, respectively:

$$\hat{x}(k+3) = \hat{F}(\hat{x}(k+2),\hat{x}(k+1), x(k),\ldots,x(k-d+2),W_2) \tag{9}$$

$$\hat{x}(k+h+1) = \hat{F}(\hat{x}(k+h),\ldots,\hat{x}(k+1), x(k),\ldots,x(k-d+h),W_2) \tag{10}$$

**Step 4:** The parameter set of the model, $W_2$, is updated by following the negative gradient direction of the error function given by:

$$e(k+1) = \frac{1}{2}\sum_{i=1}^{h} (x(k+i+1)-\hat{x}(k+i+1))^2 \tag{11}$$

In order to avoid the long computational effort required by dynamic back-propagation rules when the prediction length is high, the updates of the parameters are realized using traditional backpropagation learning rule.

**Step 5**. At this moment the time variable k is increased in one unit and the procedure is returned to Step 1.

The procedure is repeated for the complete training set until it reaches the convergence.

## Experiment

**Experimental verification and engineering application:** The experimental verification was conducted by two time series: a simulated time series produced by the logistic map and a practical time series consisting of vibration peak-to-peak value measured from large rotating machine. In order to measure the ability of this multi-step prediction model, the mean square error, also called prediction error, has been utilized:

$$E = \frac{1}{2N}\sum_{k=0}^{N-h} (x(k+h+1)-\hat{x}(k+h+1))^2 \tag{12}$$

where, h is the prediction length, x (k+h+1) and $\hat{x}(k+h+1)$ are the real and predicted values of the time series at instant k+h+1; N is the number of patterns.

**Prediction at the logistic map:** The simulated time series is given by following equation:

$$x\,(k+1) = \lambda x\,(k)\,(1-x\,(k)),$$
$$\text{When } \lambda = 3.97 \text{ and } x\,(0)= 0.5 \tag{13}$$

The equation describes a strong chaotic time serials and we initialize parameter k from k = 0 to k = 100. Table 1 and 2 show the prediction errors of two models over the training data and predicting data, respectively and Fig. 3 and 4 display the multi-step prediction curves of two models, respectively. It is evident that multi-step recurrent prediction model's prediction accuracy is much higher than that of traditional model.

Remark 1. Traditional model (3-10-1) means that the number of input layer is 3, the number of hidden layer is 10 and the number of output layer is 1.

Table 1: Logistic time series: prediction errors over the training data

| Length h | Traditional model (3-10-1) | Recurrent model (7-15-1) |
|---|---|---|
| 0 | 0.0015 | 0.0015 |
| 4 | 0.0550 | 0.0130 |
| 7 | 0.0850 | 0.0135 |

Table 2: Logistic time series: prediction errors over the predicting data

| Length h | Traditional model (3-10-1) | Recurrent model (7-15-1) |
|---|---|---|
| 0 | 0.002 | 0.002 |
| 4 | 0.072 | 0.015 |
| 7 | 0.105 | 0.026 |

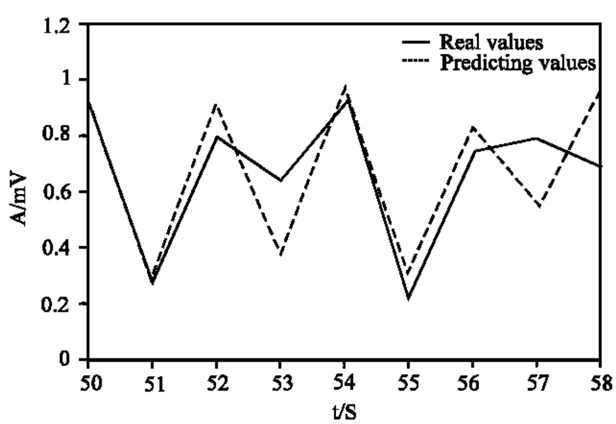


Fig. 3: Traditional multilayer feedforward neural network model predicting the logistic map at h=7

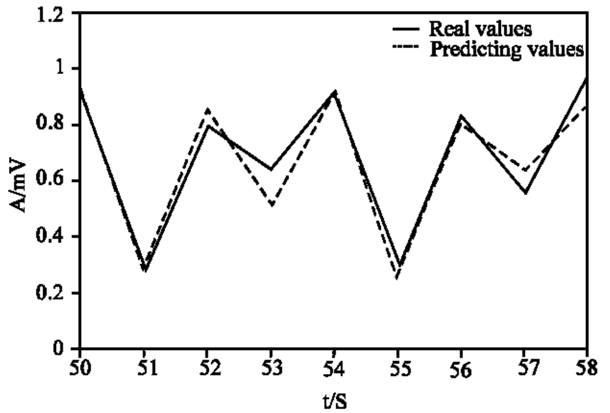

Fig. 4: Multi-step recurrent prediction model predicting the logistic map at h=7

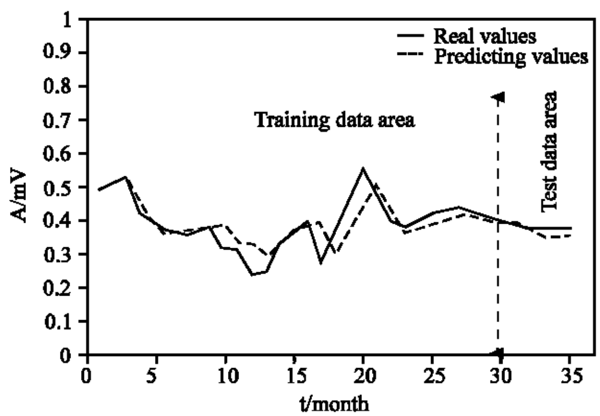

Fig. 5: Traditional Multilayer feedforward neural network model predicting the real data at h=7

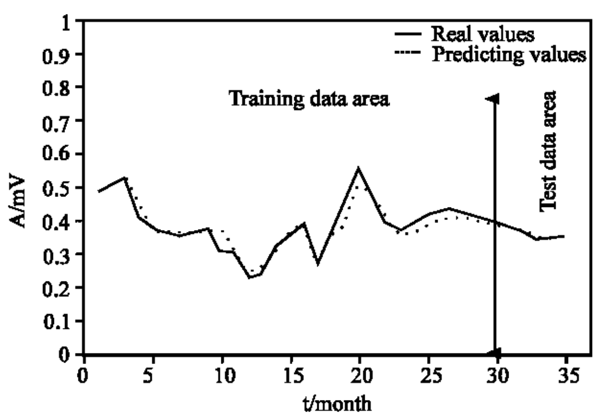

Fig. 6: Multi-step recurrent prediction model predicting the real data at h=4

**Long time prediction of machinery condition:** Trending method is well known in condition monitoring of large rotating machinery[8-10]. Moreover, in monitoring systems the peak-to-peak values of vibration signal is used as the trending object to supervise if there is any change in vibration condition of machine. Such a trending method is advantageous to improve the pre-warning ability of condition monitoring system. Accordingly, we use multi-step prediction method to monitor the tendency of peak-to-peak value with time. Figure 5 and 6 show the training and predicting trend of peak-to-peak value of this machine with month by using two models and Table 3 and 4 show the prediction errors of two models. The results show that the multi-step recurrent prediction model possesses a better predictability than traditional model.

Table 3: Real vibration data: prediction errors over the training data

| Length h | Traditional model (3-10-1) | Recurrent model (5-11-1) |
|---|---|---|
| 0 | 0.0010 | 0.0010 |
| 4 | 0.0350 | 0.0095 |

Table 4: Real vibration data: prediction errors over the predicting data

| Length h | Traditional model (3-10-1) | Recurrent model (5-11-1) |
|---|---|---|
| 0 | 0.0030 | 0.0030 |
| 4 | 0.0160 | 0.0050 |

## CONCLUSION

In this work, the multi-step prediction model was constructed and presented a learning algorithm based on recurrent neural networks. Then we applied the model to a simulated time series and long time prediction of machinery condition. The result showed that the predictive ability of multi-step recurrent prediction model contributes to the improvement of prediction accuracy than the traditional multilayer feedforward model in multi-step prediction.

## REFERENCES

1. Xu, G.H. and L.S. Qu, 1999. Multi-step prediction method based on probability neural networks. J. Xi'an Jiaotong Univ., China, pp: 89-93.
2. Laepes, A. and R. Farben, 1987. Nonlinear signal processing using neural networks: prediction and system modeling. Technical Report, Los Alamos National Laboratory, Los Alamos.
3. Hoptroff, R., 1993. The Principles And Practice Of Time Series Forecasting And Business Modeling Using Neural Nets. Neural Computing and Application, 1: 59-66.
4. Kou, C. and Reitsch, 1995. Neural networks vs. conventional methods of forecasting. J. Business Forecasting, pp: 17-22.
5. Ines, M.G. and I. Pedro, 2001. Multi-step learning rule for recurrent neural models: an application to time series forecasting. Neural Processing Lett., 13: 115-133.

6. Galvan, I. and J. Zaldivar, 1997. Application of recurrent neural networks in batch reactors. Part I: NARMA modeling of the dynamic behavior of the heat transfer fluid temperature. Chemical Engineering and Processing, 36: 505-518.
7. Wen, G.R. and L.S. Qu, 2003. Equipment behavior predictability evaluation based on redundancy. J. Xi'an Jiaotong Univ., China, pp: 699-703.
8. Yu, H.Y. and S.Y. Bang, 1997. An improved time series prediction by applying the layer-by-layer learning method to FIR neural networks. Neural Networks, 10: 1717-1729.
9. Wen, G.R. and L.S. Qu, 2002. Multi-step forecasting method based on recurrent neural networks. J. Xi'an Jiaotong Univ., China, pp: 722-726.
10. Zhang, H.J., 2002. Information extraction of Mechanical fault diagnosis and Predicting. Ph.D Thesis, Xi'an Jiaotong University China, pp: 88-102.