Improved Elman neural network in turbine blade fault diagnosis

Pengcheng Zhuo, Tangbin Xia¹, Kaigan Zhang, Zhen Chen and Lifeng Xi

State Key Laboratory of Mechanical System and Vibration, Department of Industrial Engineering, School of Mechanical Engineering, Shanghai Jiao Tong University, SJTU-Fraunhofer Center, Shanghai 200240, China

¹Email: xtbxtb@sjtu.edu.cn

Abstract. To remotely monitor and maintain large-scale complex equipment in real-time, it is required to create a comprehensive framework integrating remote data collection, transmission, storage, analysis and prediction. The framework is designed to provide manufacturers with proactive, systematic, integrated operation and maintenance service, where the data analysis and health forecasting are the most important part. This paper conducts health management for the turbine blades. An output-hidden feedback (OHF) Elman neural network is developed by adding a self-feedback factor in the context nodes. Thus, this improved method can increase the accuracy of the fault diagnosis for guide vane damage. Through the results, the applicability of this improved Elman neural network has been verified.

1. Introduction

In the modern manufacturing industry, large-scale complex equipment plays a key role. They are expensive and have a long life-cycle. As a result, the health management of large-scale equipment has attracted more and more attention [1]. Effective health management can monitor the status of the equipment in real time. Thus, the maintenance plan can be scheduled for the equipment failure more quickly, which can reduce the machine down time and the cost caused by unnecessary maintenance actions [2].

Gas turbine is one of the most advanced power machines, and is widely used in aviation, shipbuilding, electric power and other industrial fields. The blade is the most prone to failure in the gas turbine [3], which is responsible for extracting energy from the high-temperature high-pressure gas produced by combustor. More than 40% of gas turbine component failures are caused by turbine blade failure [4]. With good data acquisition and appropriate signal processing, faults can thus be detected while blades are operational and appropriate actions can be planned in time to prevent damage or failure. Therefore, it is urgent to find effective methods for fault diagnostic of gas turbine blades [5].

The vibration made by blade gives much valuable information about its health conditions. Based on the vibration information, Fourier and wavelet analysis of vibration signals are the two most commonly used techniques for blade fault diagnosis in turbo-machinery. However, blade fault diagnosis based on visual comparison of vibration spectrum and wavelet maps is very subjective as it requires experience and knowledge to interpret the results [6]. Li et al. [7] introduced fuzzy method into wind turbine blade fault diagnosis. In view of the correlation between feature quantities, this paper proposes Artificial Neural Network (ANN) which has the characteristic of approximating the...
arbitrary continuous non-linear function and all order derivatives with arbitrary precision by the appropriate selection of network layer and the cell number of hidden layers [8,9].

Elman neural network is often used for fault diagnosis because of its feedback between the hidden layer and the input layer, which can store the historical information, increase the sensitivity of the network to the historical state and improve the dynamic data performance. According to the topology of Elman neural network, this paper proposes an OHF Elman neural network (output hidden feedback Elman neural network) which increases another feedback based on Elman neural network.

2. Improved neural network

2.1. Elman neural network

Assume that there are \( r \) nodes in the input layer, \( n \) nodes in the hidden and context layers, respectively, and \( m \) nodes in the output layer. Then the input \( u \) is an \( r \) dimensional vector, the output \( x \) of the hidden layer and the output \( x_c \) of the context nodes are \( n \) dimensional vectors, respectively, the output \( y \) of the output layer is a \( m \) dimensional vector, and the weights \( w^{11}, w^{12}, \) and \( w^{13} \) are the \( n \times n, n \times r \) and \( m \times n \) dimensional matrices, respectively, the connection weights of the context nodes, input nodes and hidden nodes are denoted as \( \omega^{11}, \omega^{12}, \) and \( \omega^{13}. \)

The modified Elman network adds a self-feedback factor \( \alpha \) in the context nodes, based on the traditional Elman neural network. Its mathematical model is [10]

\[
x(k) = f \left( w^{11}x_c(k) + w^{12}u(k-1) \right) \tag{1}
x_c(k) = \alpha x_c(k-1) + x(k-1) \tag{2}
y(k) = g(w^{13}x(k)) \tag{3}
\]

Where \( f(t) \) is often taken as the sigmoid function

\[
f(t) = \frac{1}{1+e^{-t}} \tag{4}
\]

\( g(v) \) represent the transfer function of the context layer and is often taken as a linear function. \( \alpha \) is in a range of \((0, 1)\). Let the \( k \)th desired output of the system be \( y_d(k) \). Error function is:

\[
E(k) = \frac{1}{2} (y_d(k) - y(k))^T (y_d(k) - y(k)) \tag{5}
\]

Differentiating \( E(k) \) with respect to \( w^{13}, w^{12}, \) and \( w^{11}, \) respectively, according to the gradient descent method, we get the following equations:

\[
\Delta w^{13}_{ij} = \eta_3 \delta^0_i x_j(k), \quad i = 1,2, \ldots, m; j = 1,2, \ldots, n \tag{6}
\]

\[
\Delta w^{12}_{ij} = \eta_2 \delta^h_i u_j(k-1), \quad j = 1,2, \ldots, r; q = 1,2, \ldots, n \tag{7}
\]

\[
\Delta w^{11}_{ij} = \eta_1 \sum_1^m (\delta^0_i w^{12}_{lj}^T) \frac{\partial y_j(k)}{\partial w^{11}_{lj}}, \quad j = 1,2, \ldots, n; l = 1,2, \ldots, m \tag{8}
\]

Where \( \eta_3, \eta_2, \eta_1 \) are learning steps of \( w^{13}, w^{12}, \) and \( w^{11}, \) respectively, and

\[
\delta^0_i = (y_d,i(k) - y_i(k)) g_i' \tag{9}
\]

\[
\delta^h = \sum_1^m (\delta^0_i w^{12}_{lj}^T) f_j' \tag{10}
\]

\[
\frac{\partial y_j(k)}{\partial w^{11}_{lj}} = x_l(k-1) f_j' + \alpha \frac{\partial y_j(k-1)}{\partial w^{11}_{lj}} \tag{11}
\]

If \( g(v) \) is taken as a linear function, then \( g'_i = 1. \)

2.2. OHF Elman neural network

Based on Elman neural network, OHF Elman neural network adds the feedback of output nodes to the hidden nodes. The architecture of OHF Elman neural network is shown in Figure 1. Meanwhile, Figure 2 illustrates improved Elman neural network calculation flow chart.
The added nodes to perform the feedback in the output layer are called the context layer 2 nodes. \( \gamma (0 < \gamma < 1) \) is the gain factor of the self-feedback. The connection weights of the context layer 2 nodes are denoted as \( w^{13} \). \( y_c \) is the output of the context layer 2.

The mathematical model of OHF Elman neural network is [8]:

\[
x(k) = f \left( w^{12} x_c(k) + w^{13} u(k - 1) \right)
\]

\[
x_c(k) = \alpha x_c(k - 1) + x(k - 1)
\]

\[
y_c(k) = \gamma y_c(k - 1) + y(k - 1)
\]

\[
y(k) = g \left( w^{13} x(k) + w^{14} y_c(k) \right)
\]

The modifications on the weights \( w^{11} \), \( w^{12} \) and \( w^{13} \) are identical to those in the modified Elman network, and the update rule on \( w^{14} \) is

\[
\Delta w^{14}_i = \eta_s \delta_i^0 y_{c,j}(k), i = 1,2, ..., m; j = 1,2, ..., n
\]

Where \( \eta_s \) is the learning rate of \( w^{14} \) and \( \delta_i^0 \) is given by Eq. (9), and the context layers 2 nodes are equal to the output nodes. Shi et al. [8] have proved that OHF Elman neural network would have faster convergent speed when taking: \( \eta_s(k) = \left[ m \max \{ y_{c,j}(k) \} \right]^{-1} \).

3. Data experiment based on neural network

The convergence of OHF Elman neural network and superiority compared with Elman neural network needs to be verified by experiments. In this section, Elman neural network and OHF Elman neural network are used for fault diagnosis of gas turbine blades respectively, and finally the two algorithms are compared by the training error and training time.

3.1. Fault signal extraction

According to the difference sensitivity of fault signals to different fault modes, the following 12 fault signals are selected: speed(r/min), turbine output power (kw), temperature value (°C), stress value (N), four displacement parameters and four angle parameters, denoted by X1-X12 respectively. Because the acquisition of actual gas turbine operating data is limited, this paper only uses a data set of three fault modes as the experimental data of the neural network. The three failure modes are blade root corrosion, guide vane damage and heat sink blocking. In addition, the normal status should be considered. 32 sets of data and 4 sets of data are selected as training data and test data respectively, which are both divided into four groups by fault modes averagely.

![Figure 1. Architecture of OHF Elman network.](image1)

![Figure 2. OHF Elman neural network calculation flow chart.](image2)
Next, we normalize the input vector. Normalization is a mathematical approach that transforms data from a dimensional state to a dimensionless scalar with some simple calculations. In Machine Learning, the data of each dimension can be limited to [0,1], so as to avoid different effects of different components.

The commonly used normalization formula is:

\[ z = \frac{z_i - z_{\text{min}}}{z_{\text{max}} - z_{\text{min}}} \]  

(17)

Where \( z \) is the normalized result, \( z_i \) is the real data of each set of data before normalization, \( z_{\text{max}} \) is the maximum of each set of data before normalization, and \( z_{\text{min}} \) is the minimum of each set of data before normalization.

And there are three benefits of normalization:

- Dealing with the data conveniently
- Preventing the output vector from being saturated
- Simplifying the processing of neural network model

### 3.2. Establishment of neural network structure

After the extraction of the fault signals, what is required is to establish the neural network structure. In order to determine the number of the hidden nodes, it is usually referring to the empirical formula Eq. (18) and Eq. (19).

\[ n = \sqrt{r + m + a} \]  

(18)

or

\[ n = 2r + 1 \]  

(19)

Here, \( a \) is a constant between 1-10, \( m \) is the dimension of output data, \( r \) is the dimension of input data, \( n \) is the selected number of hidden layer nodes. Through Eq. (18) and Eq. (19), the approximate range of \( n \) is between 5 and 25, so this paper selects 10, 15, 20, 25 four numbers as the hidden layer to train separately, the other nodes and related parameters remain unchanged, we will determine the optimal number of hidden nodes by the training error.

### 3.3. Setting neural network parameters

Before conducting data experiments, the relevant model parameters and performance metric need to be determined. For the failure modes, they are expressed as follows:

- **F1**: Normal status: (1,0,0,0)
- **F2**: Blade root corrosion: (0,1,0,0)
- **F3**: Guide vane damage: (0,0,1,0)
- **F4**: Heat sink blocking: (0,0,0,1)

Input neurons = 12, output neurons = 4

“tansig”, “logsig” and “trainlm” are respectively selected as the neural network hidden layer excitation function, the output layer transfer function and the network training function

MSE (Mean Square Error) is chosen as the performance metric.

\[ MSE = \frac{1}{s} \sum_{i=1}^{s} (T_i - Y_i)^2 \]  

(20)

Where \( T_i \) are the expected output and \( Y_i \) are the result of neural network.

Calculating the MSE of the result, the MSE smaller, the status of blade is more likely the corresponding fault of the input.

### 3.4. Experiment of Elman neural network

Before conducting the experiment of Elman neural network, it is necessary to determine the number of nodes in the hidden layer of the neural network. Figure 3 shows the experiment results when 10, 15, 20, and 25 is selected as the number of hidden layer nodes respectively.
According to Figure 3, it can be found that the Elman neural network has the best training result when \( n = 15 \), therefore it is determined that the number of hidden layers nodes is 15.

Then we carry out the experiment of Elman neural network with \( n = 15 \). The results are as shown in Figure 4. According to Figure 4, the gradient is achieved before the Elman neural network target, so its network stops at 280 iterations, and the best training performance is 7.5147e-08.

**Table 1.** Output results of Elman neural network.

| \( y_1 \) | \( F_1 \) | \( F_2 \) | \( F_3 \) | \( F_4 \) |
|-----------|--------|--------|--------|--------|
| \( y_2 \) | 0.9999 | 0.0001 | 0.0000 | 0.0000 |
| \( y_3 \) | 0.0002 | 0.9999 | 0.0014 | 0.0005 |
| \( y_4 \) | 0.0000 | 0.0001 | 0.9997 | 0.0000 |

**Table 2.** Errors between results and expected output.

| \( y_1 \) | \( F_1 \) | \( F_2 \) | \( F_3 \) | \( F_4 \) |
|-----------|--------|--------|--------|--------|
| \( y_2 \) | -0.0001 | 0.0001 | 0.0014 | 0.0000 |
| \( y_3 \) | 0.0002 | 0.0001 | -0.003 | 0.0000 |
| \( y_4 \) | 0.0000 | 0.0000 | 0.0000 | 0.0001 |

In Tables 1 and 2, \( F_i \) are the results of the four failure modes, and \( y_i \) are the output corresponding to each set of test data. According to Eq. (20), this experiment finally obtains \( \text{MSE} = 1.3992 \times 10^{-07} \).

### 3.5. Experiment of OHF Elman neural network

In order to ensure the effectiveness of comparison between Elman neural network and OHF Elman neural network, the number of hidden layer nodes of the improved Elman neural network is also determined as 15, and the test results are displayed as follows:
Figure 5. OHF Elman neural network iteration error curve.

According to Figure 5, OHF Elman neural network has reached the iteration target when iteration arrives 203, and the best training performance is 3.6538e-09.

Table 3. Output results of OHF Elman neural network.

|   | F_1   | F_2   | F_3   | F_4  |
|---|-------|-------|-------|------|
| y_1 | 1.0000 | 0.0000 | 0.0000 | 0.0000 |
| y_2 | 0.0006 | 1.0000 | 0.0001 | 0.0001 |
| y_3 | 0.0000 | 0.0000 | 1.0000 | 0.0000 |
| y_4 | 0.0000 | 0.0000 | 0.0000 | 1.0000 |

Table 4. Errors between results and expected output /10^{-3}.

|   | F_1   | F_2   | F_3   | F_4  |
|---|-------|-------|-------|------|
| y_1 | -0.0096 | 0.0054 | 0.0000 | 0.0004 |
| y_2 | 0.6236 | -0.0005 | 0.1181 | 0.0697 |
| y_3 | 0.0000 | 0.0015 | -0.0242 | 0.0001 |
| y_4 | 0.0008 | 0.0019 | 0.0008 | -0.0054 |

Based on Eq. (20) and the data of Tables 3 and 4, this experiment finally obtains MSE=2.5529e-08. According to the results, it can be found that the output of OHF Elman neural network is very accurate and has reached the iteration target. The fault diagnosis accuracy of guide vane damage is seriously lower than other failure modes, but OHF Elman neural network has greatly improved the accuracy of the fault diagnosis for guide vane damage.

Table 5. A Comparison of neural networks output.

| Neural Network | Network epochs | Best training performance | MSE     |
|---------------|----------------|--------------------------|---------|
| Elman         | 280            | 7.5147e-08               | 1.3992e-07 |
| OHF Elman     | 203            | 3.6538e-09               | 2.5529e-08 |

From Table 5, it can be found that Elman neural network stops at 280 iterations and does not reach the iteration target, while OHF Elman neural network has reached the target at 203 iterations. On the value of MSE, the result of Elman neural network is 1.3992e-07, while OHF Elman neural network is
2.5529e-08, OHF Elman neural network has higher diagnostic accuracy and lower time. It can be concluded that OHF Elman neural network is more effective in fault diagnostic of turbine blade.

4. Conclusion
To remotely monitor and maintain large-scale complex equipment in real-time, it is required to create a comprehensive framework integrating remote data collection, transmission, storage, analysis and prediction. The framework is designed to provide manufacturers with proactive, systematic, integrated operation and maintenance service. The core of the solution is data analysis and forecasting. This paper chooses blade as research object, and then conducts fault diagnosis experiment on it with Elman neural network and OHF Elman neural network.

The first part of this research is to establish a novel neural network model of OHF Elman neural network, which adds another feedback in output nodes. After the implementation of data experiments and comparison of results, the conclusion that OHF Elman neural network is more effective than Elman neural network can be obtained in the actual verification.

The subsequent research should make the neural network model closer to reality. There are only four failure modes analyzed in this paper. In actual operation, there are more failure modes affecting industrial production. Therefore, a more general model should be established.

Acknowledgment
The research is funded by National Natural Science Foundation of China (51875359), Ministry of Education-China Mobile Research Foundation (CMHQ-JS-201900003), Intelligent Manufacturing Industrial Projects of Lingang Area (ZN2017020101) and Chinese Institute for Quality Research (019CIQ008).

References
[1] Cai B, Liu H, Xie M 2016 A real-time fault diagnosis methodology of complex systems using object-oriented Bayesian networks Mechanical Systems and Signal Processing 80 31-44
[2] Xia T, Dong Y, Xiao L, Du S, Pan E, Xi L 2018 Recent advances in prognostics and health management for advanced manufacturing paradigms Reliability Engineering & System Safety 178 255-268
[3] Jia X, Jin C, Buzza M, Wang W, Lee J 2016 Wind turbine performance degradation assessment based on a novel similarity metric for machine performance curves Renewable Energy 99 1191-1201
[4] Pennacchi P, Chatterton S, Bachschmid N, Pesatori E, Giorgio T 2011 A model to study the reduction of turbine blade vibration using the snubbing mechanism Mechanical Systems and Signal Processing 25(4) 1260-1275
[5] Liu C, Jiang D 2014 Crack modeling of rotating blades with cracked hexahedral finite element method Mechanical Systems and Signal Processing 46(2) 406-423
[6] Oh K, Park J, Lee J S, Epureanu B, Lee J K 2015 A Novel Method and Its Field Tests for Monitoring and Diagnosing Blade Health for Wind Turbines IEEE Transactions on Instrumentation and Measurement 64(6) 1726-1733
[7] Li Y, Li G. Yan J 2011 Fault Diagnosis of Wind Turbine Blades based on Fuzzy Theory International Conference on Control, Automation and Systems Engineering (1-3) July 30-31 Singapore
[8] Shi X, Liang Y, Lee H, Lin W, Xu X, Lim S 2004 Improved Elman Networks and Applications for Controlling Ultrasonic Motors Applied Artificial Intelligence 18(7) 603-629
[9] Wang Y, Zhang F, Cui T, Zhou J 2016 Fault diagnosis for manifold absolute pressure sensor (MAP) of diesel engine based on Elman neural network observer Chinese Journal of Mechanical Engineering 29(02) 386-395
[10] Drake P, Miller K 2002 Improved Self-Feedback Gain in the Context Layer of a Modified Elman Neural Network Mathematical and Computer Modelling of Dynamical Systems 8(3) 307-311