Parameter Identification Approach to Vibration Loads Based on Regularizing Neural Networks

Shouju Li and Yingxi Liu

State Key Laboratory of Structural Analysis for Industrial Equipment, Dalian University of Technology, Dalian, 116024, China

Summary
An identification algorithm for vibrating dynamic characterization by using artificial neural network is developed for multi-degree-of freedom systems. The over-fitting problem of classical back-propagation algorithm during neural network training is solved by using regularization procedure with regularized objective function. The practical application shows that the proposed training method is capable of enhancing the regularization procedure without getting stuck at these sub-optimal solutions, can be used to noisy data in order to omit an over-fitted neural approximation and has higher identification accuracy compared to the back-propagation algorithm

Key words:
Neural network, Parameter identification, Regularization procedure, Vibration responses

1. Introduction

Accurate identification of vibration load parameters in the hydro generating system is an important issue from the aspects of repair, diagnosis and maintenance of hydro power station. The indirect force calculation is of special interest when the vibrating forces cannot be measured directly while the generator unit responses caused by the vibrating loads can be measured easily [1]. In the last few years, several methods have been presented on force identification [2, 3]. Neural networks are a burgeoning area of artificial intelligence and are applied in many engineering applications. Leung (1999) proposed adaptive regularization parameter selection method for enhancing generalization capability of neural networks by means of changing the value of regularization parameter [4]. Kumar (2004) presented an improving method using Bayesian regularization for neural network training and shown the superior performance of proposed method compared to the back-propagation algorithm [5]. System identification and neural network techniques are more directly geared towards respecting these constraints than traditional modeling for two reasons: they involve simpler representation than the finite element method and directly account for errors when the model is created [6]. System identification and neural networks offer a probabilistic framework to the representation [7]. For example, when the input/output relationship is learned by minimizing a least-square distance between the neural network response and the experiments, the neural network learns the average response of the system conditioned on the input. The ANN is trained using the displacement responses of vibration as the input and the load parameters as the output.

The main objective of the study is to develop an intelligent pattern reorganization approach to parameter identification of hydro generator vibration. This paper is organized as follows: Section 2 describes the basic features of neural networks and the improvement of neural networks by using regularized objective function. Section 3 describes the field measurement of vibration responses of hydro generator, the finite element simulation of vibration responses of hydro generator and application of regularizing neural network. The comparisons of measured vibration responses with forecasted vales are depicted in the figures.

2. Improving performance in Parameter Identification using Regularization procedure for Neural network Training

An artificial neural network model is a system with inputs and outputs based on biological nerves. The system can be composed of many computational elements that operate in parallel and are arranged in patterns similar to biological neural nets. A neural network is typically characterized by its computational elements, its network topology and the learning algorithm used. Among the several different types of ANN, the feed-forward, multilayered, supervised neural network with the error back-propagation algorithm, the BPN, is by far the most frequently applied neural network learning model, due to its simplicity.

The architecture of BP networks, depicted in Figure 1, includes an input layer, one or more hidden layers, and an output layer. The nodes in each layer are connected to each node in the adjacent layer. Notably, Hecht-Nielsen proved that one hidden layer of neurons suffices to model any solution surface of practical interest.

Manuscript revised January 30, 2005.

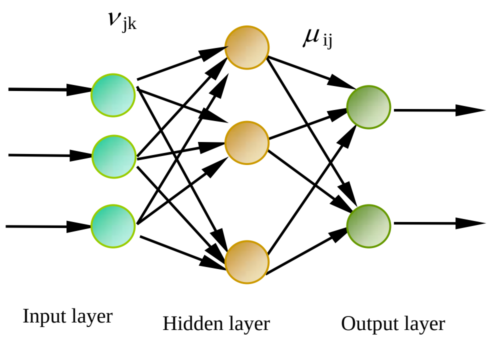

Fig. 1 Topography structure of artificial neural network

Hence, a network with only one hidden layer is considered in this study. Before an ANN can be used, it must be trained from an existing training set of pairs of input-output elements. The training of a supervised neural network using a BP learning algorithm normally involves three stages. The first stage is the data feed forward. The computed output of the i-th node in output layer is defined as follows[9]

$$y_i = f(\sum_{j=1}^{N_h}(\mu_{ij} f(\sum_{k=1}^{N_i} v_{jk} x_k + \theta_j) + \lambda_i)) \quad (1)$$

Where μ_{ij} is the connective weight between nodes in the hidden layer and those in the output layer; v_{jk} is the connective weight between nodes in the input layer and those in the hidden layer; θ_j or λ_i is bias term that represents the threshold of the transfer function f, and x_k is the input of the kth node in the input layer. Term N_i, N_h, and N_o are the number of nodes in input, hidden and output layers, respectively. The transfer function f is selected as Sigmoid function [10]

$$f(\cdot) = 1/[1+\exp(-\cdot)] \quad (2)$$

The second stage is error back-propagation through the network. During training, a system error function is used to monitor the performance of the network. This objective function, also called the error function, is often defined as follows

$$E(w) = \sum_{p=1}^{P}(\sum_{i=1}^{N_o}(y_i^p - o_i^p)^2 \quad (3)$$

Where y_i^p and o_i^p denote the practical and desired value of output node i for training pattern p, P is the number of sample, w is the weight vector of neural network, J_d is the objective function. Training methods based on back-propagation offer a means of solving this nonlinear optimization problem based on adjusting the network parameters by a constant amount in the direction of steepest descent, with some variations depending on the flavor of BP being used.

The optimization algorithm used to train network makes use of the Levenberg-Marquardt approximation. This algorithm is more powerful than the common used gradient descent methods, because the Levenberg-Marquardt approximation makes training more accurate and faster near minima on the error surface.

$$w(k+1) = w(k) - H^{-1}(k)g(k) \quad (4)$$

Where w(k) is the vector of network parameters(net weights and element biases) for iteration k, matrix H^{-1}(k) represents the inverse of the Hessian matrix. The vector g(k) represents the gradient of objective function. The Hessian matrix can be closely approximated by

$$H \approx J^T J \quad (5)$$

Where $\boldsymbol{J}$ is the Jacobian matrix, which is defined as

$$J = \begin{bmatrix} \frac{\partial e_1}{\partial w_1} \frac{\partial e_1}{\partial w_2} \cdots & \frac{\partial e_1}{\partial w_N} \\ \frac{\partial e_2}{\partial w_1} \frac{\partial e_2}{\partial w_2} \cdots & \frac{\partial e_2}{\partial w_N} \\ \cdots\cdots\cdots\cdots & \\ \frac{\partial e_p}{\partial w_1} \frac{\partial e_p}{\partial w_2} \cdots & \frac{\partial e_p}{\partial w_N} \end{bmatrix} \quad (6)$$

Where N is the number of all weights. And the gradient of the objective function can be computed as

$$g = \frac{\partial E}{\partial w} = \frac{1}{2}\begin{bmatrix} \dfrac{\partial \sum_{p=1}^{P} e_p^2}{\partial w_1} \\ \dfrac{\partial \sum_{p=1}^{P} e_p^2}{\partial w_2} \\ \cdots\cdots \\ \dfrac{\partial \sum_{p=1}^{P} e_p^2}{\partial w_N} \end{bmatrix} = J^T e \tag{7}$$

Where e is an error vector, and it can be calculated as follows

$$e = y - o \tag{8}$$

The iterative formulas of adjusting weights can be rewritten as follows

$$w(k+1) = w(k) - [J^T(k)J(k)]^{-1} J^T(k)e(k) \tag{9}$$

One problem with the iterative update of weights is that it requires the inversion of Hessian matrix $\boldsymbol{H}$ which may be ill conditioned or even singular.

Despite the popularity of the error function, there are two main shortcomings in applying those error based algorithm for general applications. On the one hand, there are many sub-optimal solutions on the error surface. The network training may easily stall because of being stuck in one of the sub-optimal solutions. On the other hand, the error function, in general, is a universal objective function to cater all harsh criteria of different applications. To have an optimal performance such as a low training error and high generalization capability, additional assumptions and heuristic information on a particular application have to been included. One of the techniques to absorb the a priori knowledge is regularization procedure, which is a systematic approach to make the network training less ill-posed. A typical form of the regularized objective function is expressed in the following equation

$$E_{new} = E + \mu E_w \tag{10}$$

$$E_w = \sum_{i=1}^{N} w_i^2 \tag{11}$$

Where μ is the regularization factor, E_{new} is the regularized objective function. The problem, which Hessian matrix $\boldsymbol{H}$ is ill conditioned or even singular, can be resolved by the regularization procedure as follows

$$H \approx J^T J + \mu I \tag{12}$$

Where $\boldsymbol{I}$ is a unity matrix. The weight adjustment using Levenberg-Marquardt algorithm is expressed as follows

$$w(k+1) = w(k) - [J^T(k)J(k) + \mu I]^{-1} J^T(k)e(k) \tag{13}$$

Step 1. Select the learning and testing patterns according to prior information and measurement data and set up the training sample pairs by using finite element method.
Step 2. Design the network architecture.
Step 3. Initialize the network weights to small random values.
Step 4. Present an input pattern, and calculate the output of the network
Step 5. Calculate the Jacobian matrix associated with input-output pairs
Step 6. When the last input-output pair is presented, perform the update of the weights
Step 7. Stop training if the network has converged, and go to step 8; else, go back to step 4.
Step 8. Estimate load parameter of hydro generator based on measured vibrating response and output parameter identification results.
Step 9. Forecast vibration response of hydro generator according to identified load parameters, and compare them with observed ones.

Fig. 2 Main steps of parameter identification of hydro generator with neural networks

The Levenberg-Marquardt algorithm approximates the normal gradient descent method, while if it is small, the expression transforms into the Gauss-Newton method. After each successful step the constant μ is decreased, forcing the adjusted weight matrix to transform as quickly as possible to the Gauss-Newton solution. When after a step the errors increase the constant μ is increased subsequently. The main steps of parameter identification of hydro generator are shown in Fig. 2.

3. Application of Improved Neural Network to Load Identification of Hydro Generator

There are 11 generating units in the main generating workshop of Fengman hydro power station. Some of them often vibrate and sometimes may affect the ordinary operations. In order to deal with these problems, some researches have been developed, and one of the studies is to analyze the vibration characteristics and estimate the vibrating loads of hydro power. The objective of the field measurement is to record the vibration responses of hydro generator at a few discrete locations in the different loading cases. The measurement data can be taken as the basis for parameter identification. These data are recorded initially as time-histories of acceleration, velocity, or displacement

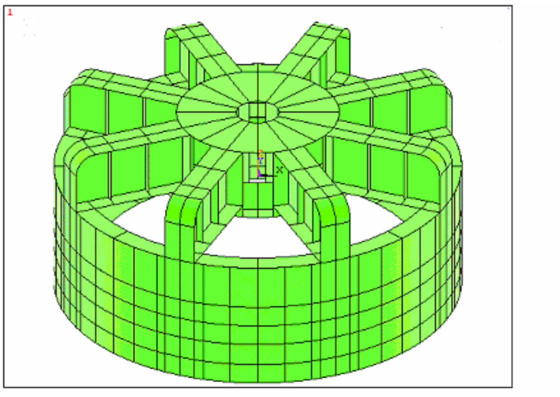

Fig.3. Frame structure of hydro generator

The experimental tests were performed for hydro generator under the different generating power cases. The vertical and horizontal displacements, velocity and acceleration versus times were recorded on the tapes for the detailed analysis. Main parameters of the hydro generator are shown as Table1. The Figure 3 is the frame structure of hydro generator. The four accelerometers were mounted on the top of frame structure of hydro power, and the vibration responses in the x, y, and z directions at different generating cases were measured. Fig. 4 shows the measured curve of vibration response of hydro generator with respect to time in the case of 20MW generating power. Fig.5 is one of the curves of power spectrum of hydro generator in the 20MW generating power.

Table 1: Main parameters of the hydro generator

Type	Power /MW	Rotating speed/r.min^{-1}	Mass of rotating part /kg	Vertical load acting on bearing /kN
SF85-40/8540	85	150	320000	6610

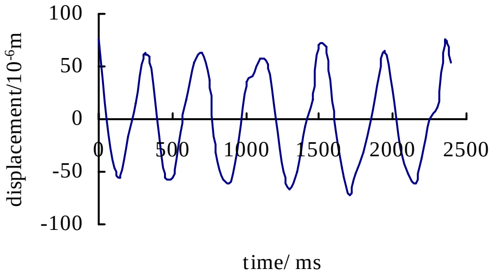

Fig. 4. Measured curve of vibration response of hydro generator with respect to time

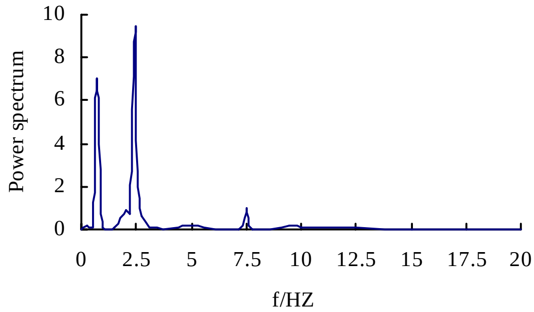

Fig. 5. Curve of power spectrum of hydro generator

In order to supply the training samples and to determine load parameters of hydro generator, the finite element simulation is modeled. The dynamic behavior of the discretized structure can be describes under the usual assumption of structural linearity, time invariant physical properties, and viscous damping by [8]

$$[M]\{\ddot{u}\}+[C]\{\dot{u}\}+[K]\{u\}=\{F\} \quad (14)$$

Where $[M]$,$[C]$ and $[K]$ are the symmetric physical mass, damping, and stiffness matrices, respectively. $\{u\}$ is the vector of displacement responses and $\{F\}$ is the vector of acting forces. According to the Newmark constant-averaged-acceleration method, Equation(14) can be discretized into the following scheme

$$[M]\{\ddot{u}^{n+1}\}+[C]\{\dot{u}^{n+1}\}+[K]\{u^{n+1}\}=\{F^{n+1}\} \quad (15)$$

$$\dot{u}^{n+1}=\dot{u}^{n}+\Delta t(\ddot{u}^{n+1}+\ddot{u}^{n})/2 \quad (16)$$

$$u^{n+1}=u^{n}+\Delta t\cdot\dot{u}^{n}+(\Delta t)^2[(1/2-\beta)\ddot{u}^{n}+\beta\ddot{u}^{n+1}] \quad (17)$$

Where Δt is time step, the parameter β is 0.25. When the boundary conditions, initial conditions and loading vector are determined, the finite element equation is adopted to compute the distribution of displacement responses, which provides model data to the identification approaches with neural networks. The numerical methods are based on spatial and temporal discretization which divides the continuous space and time domains into a network of discrete nodal points and a series of finite time intervals.

The number of neurons in the input layer depends on the number of input features in each input pattern set. The measured displacement responses at a discrete location have 49 discrete values. Therefore, the number of neurons in the input layer is equal to 49. The output layer has 6 neurons, which must be equal to the number of identified parameters. The number of neurons in the hidden layer is determined by the test, N_h=99. So, the topology structure of neural network is 49-99-6.

Taking the generation power 85 MW as an example, the main frequencies of vibrating loads can be determined from the curve of power spectrum of hydro generator shown as in Fig.5. Three main frequencies are estimated as 0.625Hz, 2.5Hz and 7.5Hz. The other identified parameters are shown as follows: amplitude of load $F_{0.625}$ and phase angle $\theta_{0.625}$ with 0.625Hz; amplitude of load $F_{2.5}$ and phase angle $\theta_{2.5}$ with 2.5Hz; amplitude of load $F_{7.5}$ and phase angle $\theta_{7.5}$ with 7.5Hz. The identified parameter vector m is expressed as follows

$$m = \{F_{0.625}, F_{2.5}, F_{7.5}, \theta_{0.625}, \theta_{2.5}, \theta_{7.5}\}^T \qquad (18)$$

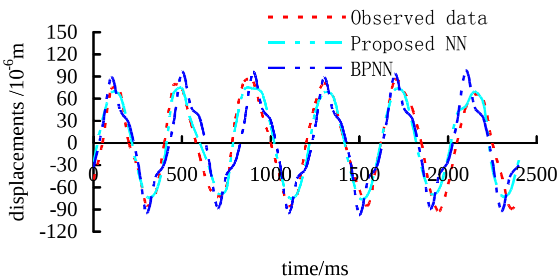

Fig. 6. Comparison of measured horizontal displacements in direction y with forecasted ones (85MW)

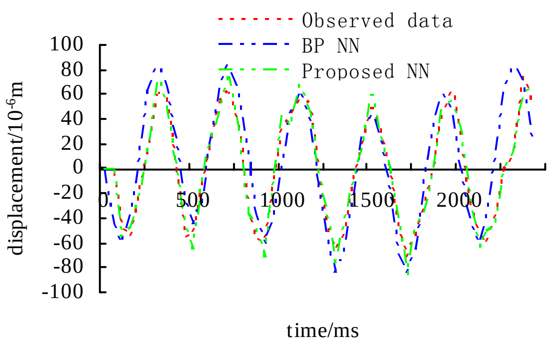

Fig.7 Comparison of measured horizontal displacements in x-direction with forecasted ones (85MW)

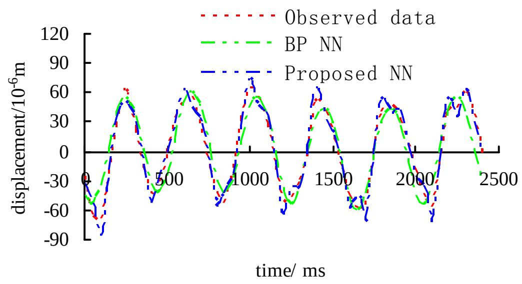

Fig. 8 Comparison of measured horizontal displacements in x-direction with forecasted ones (60MW)

Figure 6, 7 and 8 show the comparison of measured vibration displacements with forecasted ones by using different training algorithm. Compared to the classical back-propagation algorithm, the neural network trained by regularization procedure has a higher forecast precision.

4. Conclusion

The neural network has to be trained a priori to learn the complex input/output association for parameter recognition of vibration load. Field measurements of vibration responses were carried out on a hydro generator at Fengman Hydro Station. The vibration responses in the time domain and the frequency characteristics were recorded for estimating the vibration load parameters. Given the system's vibration time history measured at a few discrete locations, there are a few of algorithms to reconstruct the system's frequency response function. Neural networks are used to identify the load parameters from the displacement measurements on a hydro generator. The smaller the weights, better is the generalization capability of the neural network. The results of these studies show the potential suitability of the approach for use in industry.

Acknowledgment

This research is funded by the National Natural Science Foundation of China (No. 10472025)

References

[1] L. Yu and T.H. T. Chan, "Moving Force Identification Based On The Frequency-Time Domain Method," Journal of Sound and Vibration, vol. 261, pp.329-349, 2003

[2] X. Q. Zhu and S. S. Law, "Practical Aspects In Moving Load Identification," Journal of Sound and Vibration, vol. 258, pp.123-146,2002
[3] C. M. Tam and T. K. L. Tong, "Diagnosis Of Prestressed Concrete Pile Defects Using Probabilistic Neural Networks," Engineering Structures, vol. 26, pp. 1155-1162, 2004
[4] C. T. Leung, "Adaptive regularization parameter selection method for enhancing generalization capability of neural networks," Artificial Intelligence, vol. 107, pp. 347-356, 1999
[5] P. Kumar, "Improving performance in pulse radar detection using Bayesian regularization for neural network training," Digital Signal Processing, vol.14, pp.438-448, 2004
[6] R. L.Riche, "Neural Identification of non-Linear dynamic structures," Journal of Sound and vibration, vol. 248, pp. 247-265, 2001
[7] G. F. Lin, "A spatial interpolation method based on radial basis function networks incorporating a semivariogram model," Journal of Hydrology, vol.288, pp.288-289, 2004
[8] U. Fullekrug, "Identification of modal parameters, generalized and effective masses during base-driven tests," Aerospace Science and Technology, vol.7, pp.469-480, 1998
[9] C. S.Huang, "A neural network approach for structural identification and diagnosis of a building from seismic response data," Earthquake Engineering and Structural Dynamics, vol.32, pp. 187-206, 2003
[10] F. Meulenkamp, "Application of neural networks for the prediction of the unconfined compressive strength from Equotip hardness," Int. J. of Rock mechanics and mining Sciences, vol.36, pp.29-39, 1999
[11] F. M.Ham, Principles of neurocomputing for Science & Engineering, McGraw-Hill Companies, Inc., Singapore, 2001

Shouju Li received the B.E. from Fuxin Mining College in 1982 and M.E. degrees from China University of Mining and Technology in 1986. He received the Dr. Eng. degree from Dalian Univ. of Technology in 2004. He has been an associate professor at Dalian Univ. since 1994. His research interest includes artificial neural networks, genetic algorithm and nonlinear control, and their application to geotechnical engineering.

Yingxi Liu graduated from Dalian University of Technology in 1967. He has been a professor at Dalian Univ. since 1992. His research interest includes artificial intelligence, finite element method and inverse problem, and their application to geotechnical engineering.