Transactions on Information and Communications Technologies vol 6, © 1994 WIT Press, www.witpress.com, ISSN 1743-3517

Artificial neural network approach for multiple fault diagnosis: a case study

P.V. Suresh[a] & D. Chaudhuri[b]

[a]*Reactor Safety Division, Bhabha Atomic Research Center, Bombay 400085, India*

[b]*Division of Industrial Management, Indian Institute of Technology, Madras 600036, India*

Abstract

A method is presented for multiple fault diagnosis by means of an Artificial Neural Network (ANN). The major advantage of using an ANN as opposed to any other technique for fault diagnosis in condition based maintenance is that the network produces an immediate decision with minimal computation for a given input vector, whereas conventional techniques like spectral analysis require complete processing of an input signal to reach a diagnosis. The basic strategy is to train a neural network to recognize the behavior of the machine condition as well as the behavior of the possible system faults. The multi-layer feed forward network is used in this paper with back propagation learning algorithm. The network is trained by giving training examples, which have known input vector of vibration signatures and output vector of membership of possible faults. Field data of a lubricating oil pump for a residual gas compressor from a LPG recovery plant is used for training and testing the network. For diagnosis purpose, five different states are considered. The correct classification rate during training and testing is very high. On the basis of the results presented it is felt that the application of neural network shows superior performance in fault diagnosis.

Transactions on Information and Communications Technologies vol 6, © 1994 WIT Press, www.witpress.com, ISSN 1743-3517

1 Introduction

Machines play a significant role as the major components in any plant environment and all machines are subject to failures. These failures not only add to the down time and loss of machine but also impose a cost in time, and it is a necessity to avoid such circumstances. However, failures can be prevented through proper maintenance. Condition monitoring and diagnosis comes first in this regard with its ability to predict the fault before a catastrophe and cost effectiveness in long run. The machine condition monitoring is generally performed by evaluating the characteristic changes of the system at different operating periods. These changes are usually revealed by detecting variation in its parameters like vibration, acoustic, force, temperature etc. Several methods are available to identify the variations in these parameters. The increased investment and improved productivity, however, dictate a need for a monitoring system that is capable of detecting a possible abnormal condition at its earliest stage and should be fast enough for on-line implementation.

The vibration characteristics, the widely used parameter for machine condition monitoring, will change when machine components start to deteriorate or it develops a new defect. Generally, there are two popular approaches for vibration signal analysis: time-domain vibration signal analysis and frequency-domain vibration signal analysis techniques, Collacot [3]. In conventional approach, the time-domain based techniques include overall level (RMS) measurements, peak level detection, Kurtosis etc. while the frequency-domain based techniques include spectrum analysis, cepstrum analysis, waterfall plot etc. One of the most useful fault diagnosis techniques is that of spectrum analysis. The conversion of time-domain to frequency-domain and vice verse can be carried out using Fourier transform and inverse Fourier transform respectively. Fuzzy set theory based methods (Gazdik [4], Li and Wu [6]) are also available in literature for machine condition monitoring and diagnosis. Knapp et al. [5] used ANN for a CNC machine diagnosis purpose with four different states. The study was limited to identify whether fault is existing or not. However, the severity of the fault is also important. This paper describes the method to find out the presence and severity of the faults. The input data for the network is from the signature collected on frequency domain. The paper covers a brief introduction of neural network and back propagation algorithm. It is followed by a case study of a lubricating oil pump and conclusions.

Transactions on Information and Communications Technologies vol 6, © 1994 WIT Press, www.witpress.com, ISSN 1743-3517

2 Neural networks

The neural networks are computational models, loosely based on the structure and the information processing capabilities of the human brain. A neural network is an assembly of large number of highly interconnected simple processing units (neurons). The connections between the neurons have numerical values which represent the strength of these connections called weights. Knowledge is stored in the form of a collection of connection weights. These neural networks are capable of self organization and knowledge acquisition i.e. learning.

Neural network models are specified by the net topology (i.e. feed forward, feed back etc.) node characteristics (binary, linear and nonlinear) and learning rules. Figure 1 illustrates a typical feed forward neural network with three layers [1], known as input layer, hidden layer and output layer. In this network the nodes in the middle layer obtain information only from the nodes of the preceding layer and forward their outputs solely to nodes in the consecutive layer. The hidden layers are not connected directly to the outer world. Computational elements or nodes used in neural net models are usually nonlinear, and typically analog. The simplest node, sums the weighted inputs and passes the result through a nonlinearity as shown in figure 1. The node is usually characterized by an internal threshold or offset and by the type of nonlinearity.

Neural networks are inspired by both biological neuron systems and mathematical theories of learning, information processing and control. Neural networks have the following main benefits:

Processing speed through massive parallelism

Learning and adapting ability by means of efficient knowledge acquisition and embedding

Robustness with respect to fabrication defects and different failures.

The main part of any neural network is its learning algorithm. Back Propagation, which is a supervised learning algorithm is the one which widely used for different applications, such as structural damage assessment [10], process fault diagnosis [9] etc. Burke [2] adopted an unsupervised learning algorithm known as competitive learning approach for tool wear monitoring. Barshdorff et al. [1] had given a comparative study between two learning approaches, back propagation and condensed neighbor concept. Following section explains back propagation learning algorithm. Lippman [7] and Yoh-Han Pao [8] provide more information for the reader.

Transactions on Information and Communications Technologies vol 6, © 1994 WIT Press, www.witpress.com, ISSN 1743-3517

3 Back propagation algorithm

The processing units in back propagation neural networks are arranged in layers. Each neural network has an input layer, an output layer and a number of hidden layers. Propagation takes place in a feed forward manner, from input layer to the output layer. The pattern of connectivity and the number of processing units in each layer may vary with some constraints. No communication is permitted between the processing units within a layer, but the processing units in each layer may send their output to the processing units in the higher layers. Associated with each connection is a numerical value which is the strength or the weight of that connection i.e. w_{ij} = strength of connection between units i & j. At the beginning of a training process, the connection strengths are assigned random values. The learning algorithm modifies the strength in each iteration up to the successful completion of the training. When the iterative process has converged, the collection of connection strengths captures and stores the knowledge and the information present in the examples used in its training. Such a trained neural network is ready to be used. When presented a new input pattern, a feed forward network computation results in an output pattern which is the result of the generalization and synthesis of what it has learned and stored in its connection strengths.

The network employed in the present study possess two hidden layers of processing elements (PE). Upon presentation of an input-output pattern, the total input, net_j, to the j^{th} processing element is obtained by calculating the weighted sum of all PE_i outputs that are connected to PE_j, i.e.,

$$net_j = \sum_i w_{ij} x_j \tag{1}$$

where, w_{ij} is the connection weight from the i^{th} PE to the j^{th} PE, x_i is the output of i^{th} PE. A sigmoid transfer function is then applied to net_j to obtain the output of the j^{th} PE.

$$x_j = f(net_j) = \frac{1}{1 + e^{-(net_j + \theta_j)}} \tag{2}$$

where, θ_j serves as a threshold or bias and the effect of a positive θ_j is to shift the activation function to the left along the horizontal axis. The threshold, θ_j will be learned in the same manner as the weights. We simply imagine that θ_j is the weight from a unit that always has an output value of unity. The output serves as input to succeeding layers which is continued until the output layer is reached and is referred to as forward activation flow. The subsequent weight adoption or learning process is accomplished by the back propagation learning algorithm.

Transactions on Information and Communications Technologies vol 6, © 1994 WIT Press, www.witpress.com, ISSN 1743-3517

The neural network output for the p^{th} pattern, o_{pj}, will not be the same as the target value t_{pj}. For each pattern, the sum of the square of the error is,

$$E_p = \frac{1}{2}\sum_j (t_{pj} - o_{pj})^2 \tag{3}$$

and the average system error is

$$E = \frac{1}{2P}\sum_p \sum_j (t_{pj} - o_{pj})^2 \tag{4}$$

The aim of the back propagation learning algorithm is to iteratively minimize the average squared system error between values of the output PE and the correct patterns provided by a teaching input using gradient descent approach. This is accomplished by first computing the gradient (δ_j) for each PE on the output layer.

$$\delta_j = x_j(1 - x_j)(t_{pj} - x_j) \tag{5}$$

where, t_{pj} is the correct teaching value for output unit j and for input pattern p. The error gradient is then recursively determined for hidden layers by computing the weighted sum of the errors at the previous layer,

$$\delta = x_j(1 - x_j)\sum_k \delta_k w_{kj} \tag{6}$$

where, k is overall PE's in the layers above j. The errors are propagated backwards one layer, and the same procedure is recursively applied until the input layer is reached. The error gradients are then used to update the network weights,

$$\Delta w_{ji}(n) = \eta \delta_j x_i \tag{7}$$

$$w_{ji}(n+1) = w_{ji}(n) + \Delta w_{ji}(n) \tag{8}$$

where n is the iteration number and η is the learning rate which provides the step size during gradient descent. Generally, to assume rapid convergence, large step sizes which do not lead to oscillations are used. To avoid oscillations small learning rates can also be used even though it may take more computer time for convergence. Convergence, however, may be improved by including a momentum term, α, which determines the effect of previous weight changes on present change in weight space. The weight change after the n^{th} data presentation is,

$$\Delta w_{ji}(n) = \eta \delta_j x_i + \alpha \Delta w_{ji}(n-1) \tag{9}$$

4 Case study

The training of a neural network with appropriate data containing the information regarding the cause and effect relationships is the main criterion of the proposed method of machine fault diagnosis. Thus, the first step is to collect or generate a set of data that can be used in training an appropriate neural network to recognize different fault conditions from the obtained machine responses. Ideally, this data set should contain the response of all types of possible faults as well as the responses of the normal operations.

In this case study the proposed method has been applied to the diagnosis of an auxiliary lube oil pump (gear type) of a Residual Gas Compressor from an LPG recovery plant, which has been monitoring weekly. The required information has been obtained from a leading consultant in India. All the signatures are in the form of Fourier spectra of velocity time histories in hard copy. (Eg. fig. 2). The 28 sets of spectra obtained have been digitized using Graphtec-Digitizer (Model No. KD4300) and the entire frequency range (0-200KHz) is divided into 100 intervals and the amplitude corresponds to each point is the first 100 inputs. The other 9 inputs for each signature are the peak amplitude at three different pick up points in three directions (Horizontal, Vertical & Axial). 20 samples are selected for training the network out of 28 spectra and the remaining eight samples for testing the network.

Output of the network consists of five different machine fault conditions, which are identified from the vibration report and each fault condition is represented by a node in the output layer. They are,

1. Misalignment between motor and pump
2. Wear in the coupling bushes
3. Defects in the pump bearing
4. Gear meshing inaccuracies in the pump
5. Overall health condition of the pump.

The target values of network outputs are normally assumed as 1 if a particular fault exists or else 0 [5]. This way one may be able to identify only the fault existence. Our aim is to find out the severity also. For this purpose we have given some grade of membership similar to fuzzy membership function [4], for the output vector. The grade of membership is based on the vibration report. If the membership lies between 1 and 0.6 then it is classified as severe and if it is between 0.6 and 0.4 , then it is minor and if it is less than 0.4 then the fault has no effect in the condition of the machine. Similarly the state 5 gives the overall working condition of the machine. If it is less than 0.4 then normal working condition, between 0.4 and 0.6 then it is still acceptable level and if it is more than 0.6, the condition is bad and action to be taken. For example, let O_p = [0.8, 0.5,

Transactions on Information and Communications Technologies vol 6, © 1994 WIT Press, www.witpress.com, ISSN 1743-3517

0.2, 0.7, 0.5] in which fault 1 and fault 4 are severe, fault 2 is minor and fault 3 is not at all present and overall condition is satisfactory. During training the network, various number of hidden layers and various number of hidden units are tried. Even though there is no definite rule to find out the above parameters, through trial and error it is found that the two hidden layers with 12 and 8 units converges faster and error reduction is also faster. Thus the final net architecture becomes 109 * 12 * 8 * 5. The learning rate η and momentum rate α in equation (7) are assumed as .01 and .9. It is found that with a high value of learning rate the network converges to a local minima even though it consumes less computer time.

The training terminates when error for each pattern (Eq. 3) or total system error falls below the specified limit or the maximum number of iterations reached. In 2000 iterations the total average system error E has reached 1.29 $\times 10^{-5}$ with running time 292 sec. on a PC/486 with cycle frequency 66 MHz. Figures 3A and 3B give a clear picture about the reduction in system error with number of iterations for different hidden layers and for different learning rates respectively.

Once the network is trained, it can be used for diagnosis purposes of unknown signatures. In order to validate the system the trained network was tested with both data samples, i.e. the samples used for training and testing. Network performance in the trained sample set was 100% while in the test case, it was correctly identifying the fault class and the severity in 7 out of 8 cases. Typical output of the network for an unknown signature may be

Misalignment between motor and pump
Gear meshing inaccuracies in the pump
Minor bearing problem
Overall health condition is bad.

5 Conclusions

A new approach for machine fault classification and diagnosis has been developed based on back propagation algorithm. Results of this approach indicate that neural network techniques are capable of learning about the behavior of normal working conditions as well as faulty conditions of machine from the vibration or any other signatures. The network is able to classify a particular fault and the severity of that fault is based on the threshold membership which is given. Here the threshold value is given based on vibration report and it may vary from situation to situation. Another problem may be that if a new type of fault is developing the system will not be able to give any indication. Then the training has to be repeated including the new fault for future reference work. Further research in this area should be to incorporate multi channel signals for fault classification and machine on line monitoring. The unsupervised learning approaches like Competitive Learning or Adaptive Resonance Theory can also be incorporated for future work where noise signatures are the only source of information.

References

[1] Barchdorff, D., Monostori, L., Ndenge, A.F. & Wostenkuhler, G.W. Multi processor systems for connectionist diagnosis of technical processes, *Computers in Industry*, 1991, **17**, 131-145.

[2] Burke, L.I. Competitive learning based approaches to tool wear monitoring , *IEEE Trans. On Systems Man and Cybernetics*, 1992 **22**, 559-563.

[3] Collacott, R.A. *Mechanical Fault Diagnosis*, Chapman and Hall, USA, 1977.

[4] Gazdik, I. Fault Diagnosis and prevention by Fuzzy Sets, *IEEE Trans. on Reliability*, 1985, **R-34**, 382-388.

[5] Knapp,G.M & Wang, Hsu-Pin(Ben) Machine Fault Classification: A neural network approach, *International Journal of Production Research*, 1992, **30**, 811-823.

[6] Li, P.G. & Wu, S.M. Monitoring Drilling Wear states by a Fuzzy pattern recognition technique, *Trans. of ASME: Journal of Engineering for Industry*, 1988, **110**, 297-300.

[7] Lippmann, R.P. An introduction to computing with Neural Networks, *IEEE ASSP Magazine*, April 1987, 4-22.

[8] Pao, Yoh-Han, *Adaptive Pattern Recognition and Neural Networks*, Addison-Wesley, New York, 1990.

[9] Venkatasubramanian, V. & Chan, K. A Neural Network methodology for process fault diagnosis, *AIChE Journal*, 1989, **35**, 1993-2001.

[10] Wu, X., Ghaboussi, J. & Garrett Jr., J.H. Use of Neural Networks in detection of structural damage, *Computers and Structures*, 1992, **42**, 649-659.

Transactions on Information and Communications Technologies vol 6, © 1994 WIT Press, www.witpress.com, ISSN 1743-3517

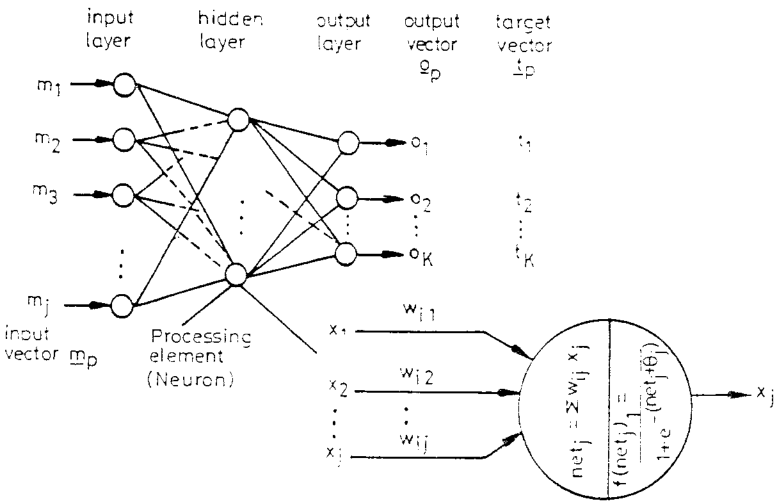

Fig. 1 Structure of a three layer-feed forward artificial neural network and PE_j

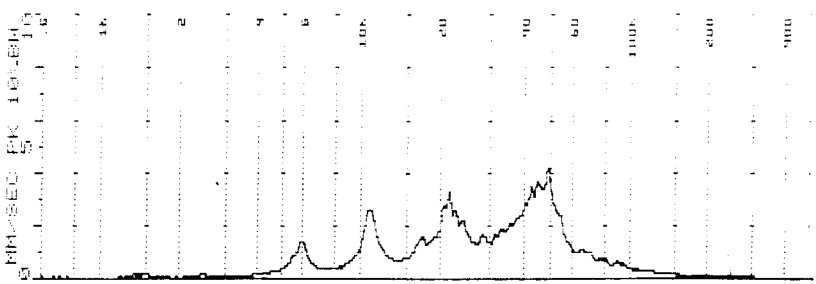

Fig. 2. A typical frequency domain vibration signature

Transactions on Information and Communications Technologies vol 6, © 1994 WIT Press, www.witpress.com, ISSN 1743-3517

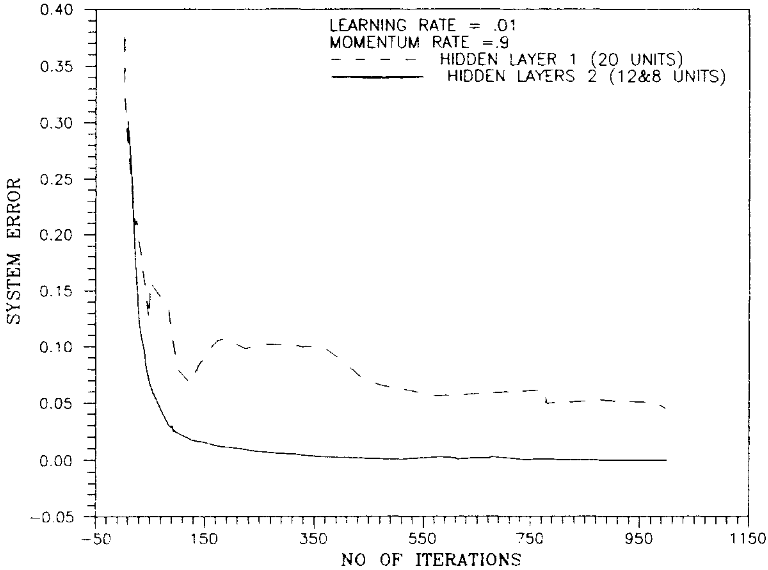

Fig. 3A. Error reduction for different hidden layers

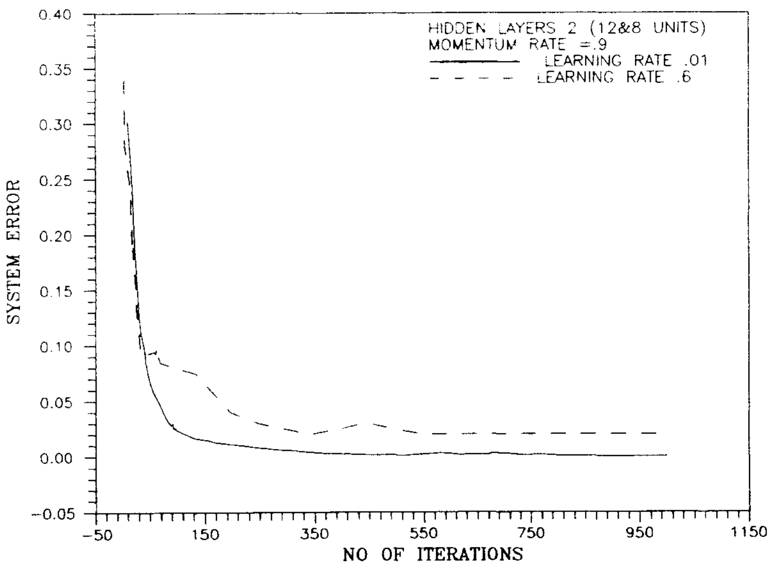

Fig. 3B. Error reduction for different learning rates