

Neural network based vibration analysis with novelty in data detection for a large steam turbine

K. P. Kumar^a, K.V.N.S. Rao^{c,*}, K.R. Krishna^b and B. Theja^a

^aSchool of Mechanical Engineering, Vignan University, Vadlamudi, India

^bMachine Dynamics and Failure Analysis Laboratory, Bharath Heavy Electricals Limited (B.H.E.L), Hyderabad, India

^cDepartment of Mechanical Engineering, Sir C R Reddy College of Engineering, Eluru, India

Received 4 November 2009

Revised 18 May 2010

Abstract. Health of rotating machines like turbines, generators, pumps and fans etc., is crucial to reliability in power generation. For real time, integrated health monitoring of steam turbine, novel fault detection data is essential to reduce operating and maintenance costs while optimizing the life of the critical engine components. This paper describes about normal and abnormal vibration data detection procedure for a large steam turbine (210 MW) using artificial neural networks (ANN). Self-organization map is trained with the normal data obtained from a thermal power station, and simulated with abnormal condition data from a test rig developed at laboratory. The optimum size of self-organization map is determined using quantization and topographic errors, which has a strong influence on the quality of the clustering. The Mat lab 7 codes are applied to generate program using neural networks toolbox.

Keywords: Integrated health monitoring, self-organization map, quantization and topographic errors, clustering

1. Introduction

In recent developments, instrumentation and computer based information management systems are leading to great advances in effectiveness of predictive and preventive maintenance programs. The development of non-intrusive monitoring techniques has made possible the useful transition from preventive maintenance to the predictive maintenance. Vibration measurement of machines is essential for periodic and continuous monitoring in preventive/predictive maintenance program [1]. The machine condition should be evaluated by comparing the measurements of certain parameters with some limits, which could be defined by constructors by means of severity charts or international norms [2–5].

Vibration based damage identification is difficult to implement in practice because one should consider a large number of data points and it is difficult for data interpretation. Compared to conventional offline techniques, computer based on-line systems for vibration monitoring of large steam turbine gives good financial benefit [6]. To develop the on-line monitoring system and for easy analysis of data, researchers made many efforts for designing a new reliable system. In the literature, expert systems based on artificial intelligence were used and are known as “rule-based-expert systems”. Its application in the area of pattern recognition and data mining seems to be restricted

*Corresponding author. Tel.: +91 9490223849, +91 863 2244735; Fax: +91 8812 224193; E-mail: kvnsrao@yahoo.com.

Table 1
Symbols for Vibration Frequency and Process Feature Description

Frequency features	Description	Process features	Description
f1	0.01 ~ 0.39X	P1	Amplitude jump during operation
f2	0.4 ~ 0.49X	P2	Vibrations at various power load
f3	0.50X	P3	Axial vibration
f4	0.51 ~ 0.99X	P4	Shaft average centerline
f5	1.X	P5	Critical speed spectrum
f6	2.X	P6	Stable at various running speeds
f7	3 ~ 5.X	P7	Vibration level increase during running up
f8	Odd of X	P8	Level jump during running up
f9	High X	P9	3Xat 1/3 critical speed
f10	Power line	P10	Half speed whirl

and these systems have several bottlenecks mainly concerning fuzzy rules generation and fuzzy sets forming [7]. An alternative method to the rule based system is to apply a neural network as an expert system.

Artificial neural networks (ANN) are massively parallel-interconnected net works, which are used to solve a wide variety of science and engineering problems that involve extracting useful information from complex or uncertain data [8,9]. Artificial Neural Networks have the ability to perform pattern recognition and classification, and are especially useful in situations where the data volumes are large and the relationships are unclear or even hidden. This is because of their ability to handle noisy data in difficult non-ideal contexts [10]. Vibration data was typically noisy and has complex non-linear relationships. Particular attention was directed to using the Self-Organizing Map (SOM) [11] neural network as a means of organizing complex information spaces [12–15]. Recently Crupi et al. [16] proposed a neural network based expert system for a rotating machine consisting of analysis and diagnosis section. Analysis section identifies novelty in vibration data using self-organization map but the procedure was not given to calculate the optimum size of the self-organization map. Hoffman et al. [17] compared different neural network based systems like Self-Organizing Maps (SOM), Radial Basis Function (RBF), and Multilayer Perceptron (MLP) to model the fault conditions and found that both SOM and RBF are equally good to reliably identify specific and multiple fault conditions in vibrating equipment.

The present work attempts to apply self-organization map algorithm with optimization to detect the novelty in the vibration data for a large steam turbine (210 MW). To train the neural network, steam turbine data under normal condition is used. Different unbalance conditions are introduced on test rig at laboratory and vibration data is collected to simulate the network.

2. Steam turbine fault classification and feature extraction

Fault diagnosis is based on pattern identification and classification. Steam turbine faults are generally classified into twenty types based on the field experts experience and theoretical analysis [18].

The first step in steam turbine fault diagnosis is pattern identification from the measured data. The next step is to interpret what the pattern indicates about the machine, but proper interpretation requires some knowledge about the machine [18]. All the above described steam turbine faults can be detected by using ten spectral features and process features as shown in Table 1. The alphabetic symbols used to describe the spectrum and process features are listed in Table 1.

The notation ‘nx’ in the second column of table denotes a frequency component (or range) in the spectrum at ‘n’ times the turbine’s rotational speed.

3. Test rig details

3.1. Construction features

Roto Dynamic test rig (developed by Machine Dynamics Lab, BHEL, R&D) consists of support structure T-slot bed, 2.2 KW A.C motor (3000 rpm), 50 mm diameter shaft mounted on two journal bearings with a span of 1000 mm, three disks of 300 mm diameter and 18 mm thickness are mounted on the shaft for balancing experiments.

The test rig is provided with various types of vibration measurement instrumentation like shaft vibration with non-contact probe and bearing pedestal vibration with velocity probe and accelerometers. The facility is also made to carry out balancing of rotor, misalignment of shafts, and to identify critical speed of shaft and its support structure.

3.2. Test rig instrumentation

Bentely Naveda velocity probe is used to measure bearing pedestal vibration at drive end and non-drive end, and also the support structure. 8 mm non-contact probes are used to measure shaft vibration near the drive end and non-drive end bearings. ONOSOKI CF-4220A FFT analyzer is used to study the frequency spectrum. DVF3 is used to measure 1 rpm component and phase angle or balancing and also the overall vibration level and speed of the rotor. All the equipment are calibrated and vibration characteristics are compared as per the provisions of ISO 2372.

4. Analysis section

4.1. Algorithm of self-organization feature map

A self-organizing map consists of two layers of processing units: the first is an input layer containing processing units for each element in the input vector, the second is an output layer or grid of processing units that is fully connected with those at the input layer.

When an input pattern is presented to a Self Organization Map (SOM) network, the units in the output layer compete with each other for the right to be declared the winner. The winning output unit is the unit whose incoming connection weights are the closest to the input pattern in terms of Euclidean distance. Thus the input is presented and each output unit competes to match the input pattern. The output that is closest to the input pattern is declared the winner. The connection weights of the winning unit are then adjusted i.e. moved in the direction of the input pattern by a factor determined by the learning rate. This is the basic nature of competitive neural networks. The following algorithm is used for novel vibration data detection.

1. The input for the Kohonen SOM is a dataset of vectors $\{\mathbf{X}(t)\}$, each one consisting of

N variables, $\mathbf{X}(t) = [x_1(t), x_2(t), \dots, x_N(t)]$.

2. Initialize the weights i.e. the codebook vectors $\mathbf{W} = \{w_{m \times n}; m = 1, 2\}$ of each neuron ‘m’ in the self-organization map to random values in the interval (0,1).

• Loop until weight changes are negligible

For each input vector $X(t)$, $t = 1, 2, 3, \dots, t = \text{no. of input vectors}$

3. Determine the winning neuron as the unit with the weight vector that has the smallest Euclidean distance to the input pattern.

Euclidean distance is defined by the equation:

$$\|x - w_i\| = \sqrt{\sum_{j=1}^n (x_j - w_{ij})^2} \quad (1)$$

4. The neuron with the shortest distance to the input vector is declared as the winning neuron, also called the Best Matching Unit (BMU).

$$\|x - w_w\| = \min \{ \|x - w_i\| \} \quad (2)$$

Find all units in the neighborhood $N(t)$ of the winner.

5. Update the weight vectors for the Best Mapping Unit and its neighbors according to the following equation.

$$w_i(t+1) = \begin{cases} w_i(t) + \alpha(t)[x(t) - w_i(t)], & i \in N_w(t) \\ w_i(t), & i \notin N_w(t) \end{cases} \quad (3)$$

Where $\alpha(t)$ is the learning rate function.

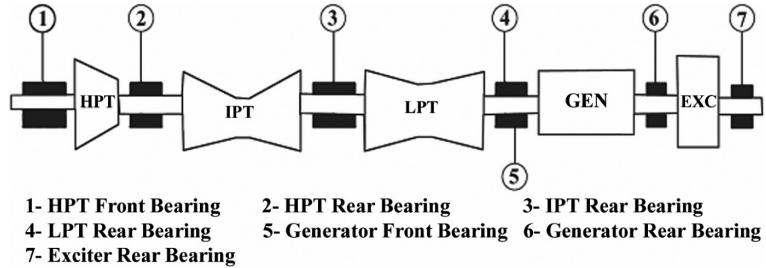


Fig. 1. Turbo Generator line diagram.

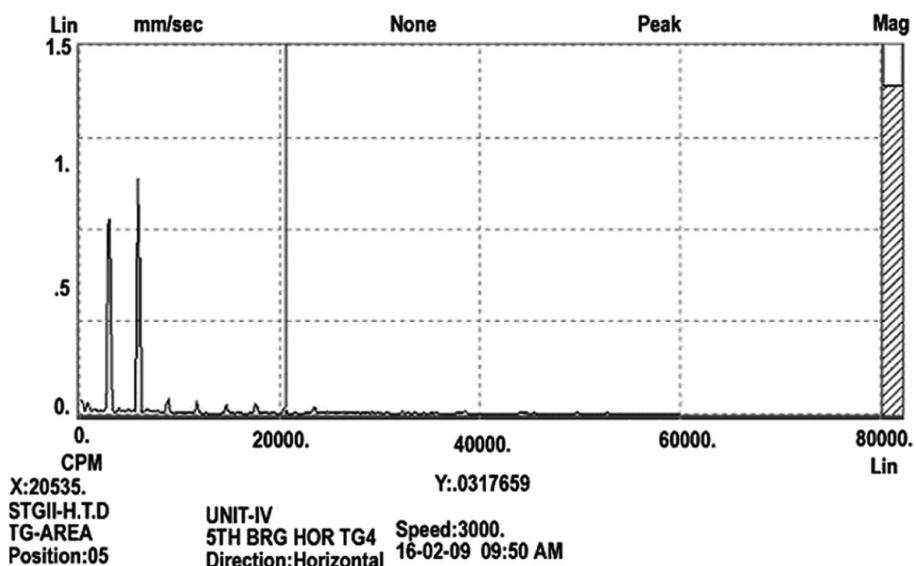


Fig. 2. Vibration spectrum that shows the maximum amplitude at odd of X.

4.2. Determining the size of Self Organization Feature Map (SOFM)

The size of the self-organization map (number of output neuron units) has a strong influence on the quality of the clustering. If the selected map size is too small, it might miss to explain some important differences present in the data. Conversely, if the selected map size is too large, the differences may become too small to detect. The maps have two primary quality properties, they are data representation accuracy and data set topology representation accuracy. The former is usually measured using average quantization error. Average quantization error value is calculated for the whole map or for each one of the neurons and each one of the input patterns.

The quantization error is defined as the mean of the Euclidean distance of each data vector to its Best Mapping Unit's weight vector and measures map resolution. For any given dataset, quantization error is reduced by increasing the number of map nodes, because of the data samples are distributed more sparsely on the map. The topographic error [19] is calculated as the proportion of all data vectors for which first and second Best Mapping Units are not adjacent units in the grid.

This computation is performed as follows: For all data samples, the respective best and second-best matching units are determined. If these are not adjacent on the map lattice, this is considered an error. The total error is then normalized to a range from 0 to 1, where '0' means perfect topology preservation. The size of the self-organization map is determined using mat lab7 program by minimizing quantization and topographic error.

Table 2
Training data set for proposed SOFM network

S. No	Parameters/date	2/2/2009	16/2/2009	2/3/2009	9/3/2009	23/3/2009
1	Overall	2.096082	1.810374	1.413985483	1.545717	1.409663
2	f1	0.0843591	0.0476031	0.0555372	0.149167	0.04905
3	f2	0.030829	0.00849278	0.00384759	0.0391308	0.0159572
4	f3	0.00858415	0.00426968	0.00236792	0.0158103	0.00948548
5	f4	0.928319	0.0161767	0.011736	0.519438	0.563847
6	f5	1.0401	1.01752	0.716517	0.797433	0.608606
7	f6	1.82699	1.11895	1.21237	1.17623	1.11849
8	f7	0.075103	0.0840676	0.126954	0.111687	0.107998
9	f8	0.0254108	0.0404374	0.0319831	0.0367446	0.0291654
10	f9	0.0255998	0.00790461	0.0054321	0.00262502	0.00853977
11	f10	0.98421	1.04419	0.0763516	0.08738	0.586227
12	P1	N	N	N	N	N
13	P2	N	N	N	N	N
14	P3	L	L	L	L	L
15	P4	L	L	L	L	L
16	P5	N	N	N	N	N
17	P6	N	N	N	N	N
18	P7	Y	Y	Y	Y	Y
19	P8	N	N	N	N	N
20	P9	N	N	N	N	N
21	P10	N	N	N	N	N

N = NO, Y = YES, L = LOW.

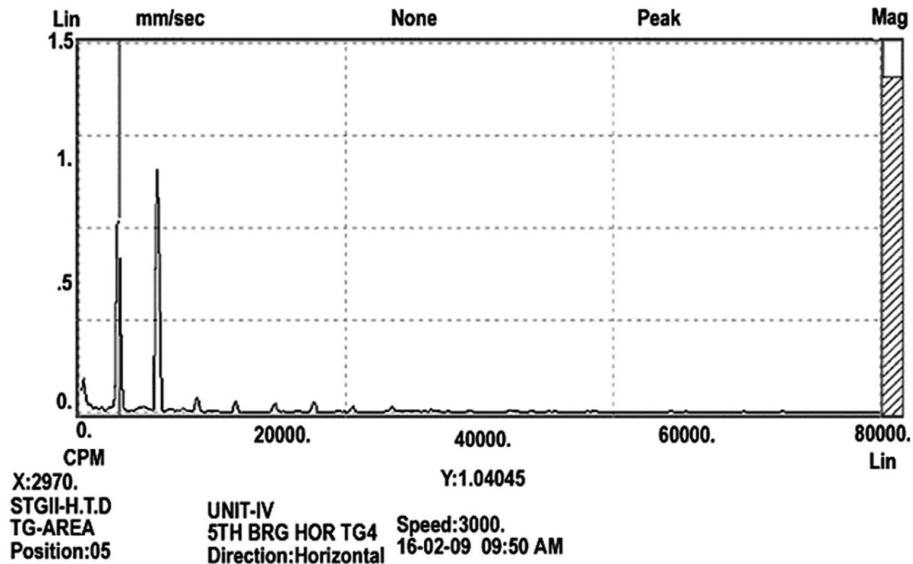


Fig. 3. Vibration spectrum that shows the maximum amplitude at high of X.

4.3. Data collection

To train the self-organization feature map, training data is collected from the turbo generator area of a thermal power station, when turbine is running at normal condition. The spectral data is collected at drive end of generator bearing horizontal position as shown in Fig. 1. Process parameters collected by the experience of the personnel who are doing the general routine machine maintenance monitoring. Portable magnetic probe accelerometer is used to measure bearing pedestal vibration data, which is connected to ENTEK data collector that is used to collect the data at the specified intervals of frequency at the specified position. EMONITOR ODYSSEY (ENTEK IRD)

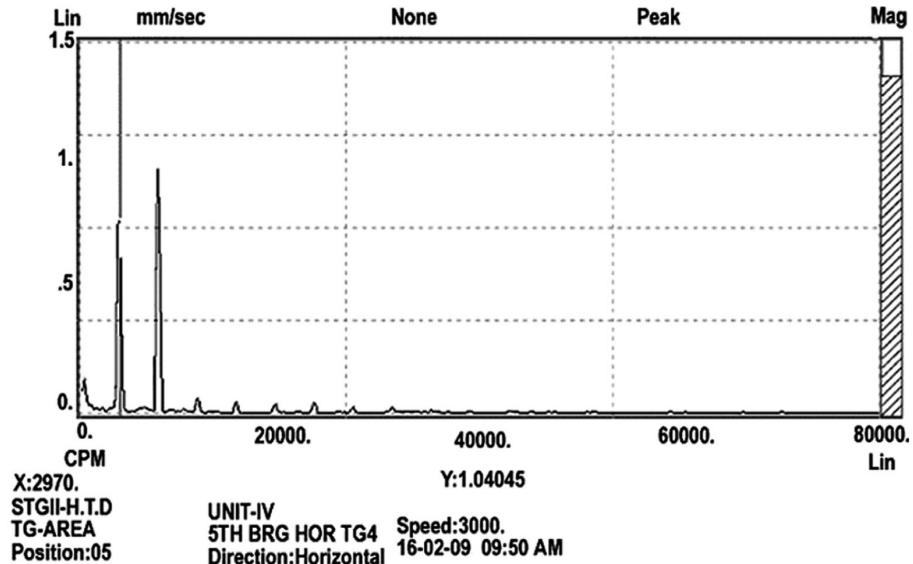


Fig. 4. Vibration spectrum that shows the maximum amplitude at power line.

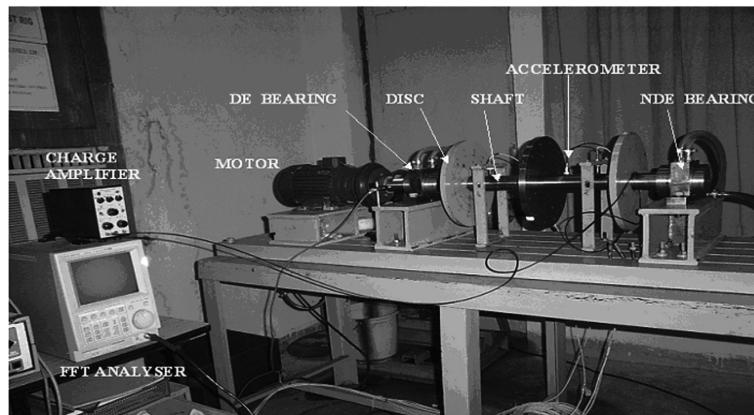


Fig. 5. Instrumentation setup for vibration measurement on test rig.

software collects the data from the data pack and gives different Fast Fourier Transform (FFT) spectrums at different resolutions.

The sample spectrums (Figs 2 to 5) are collected from the steam turbine at 5th bearing horizontal position as shown in Fig. 2 running at 3000 rpm. Cursor black line shows the maximum amplitude with in the specified range. When the turbine is running at normal condition, training data is prepared with the help of vibration spectrums as shown in Table 2.

The test data is collected from the rotor dynamic test rig, which is running at 3000 rpm at different unbalance conditions shown in Fig. 5. Bentley Nevada velocity probe is used to measure bearing pedestal vibration at drive end and ONOSOKI CF-4220A. FFT analyzer is used to study the frequency spectrums.

When test rig is running at 3000 rpm with different unbalance conditions, spectrums obtained at driving end bearing are shown in Fig. 6 to Fig. 8.

When the test rig is running at unbalanced condition, test data is prepared with the help of vibration spectrums (Fig. 6 to Fig. 8) as shown in Table 3.

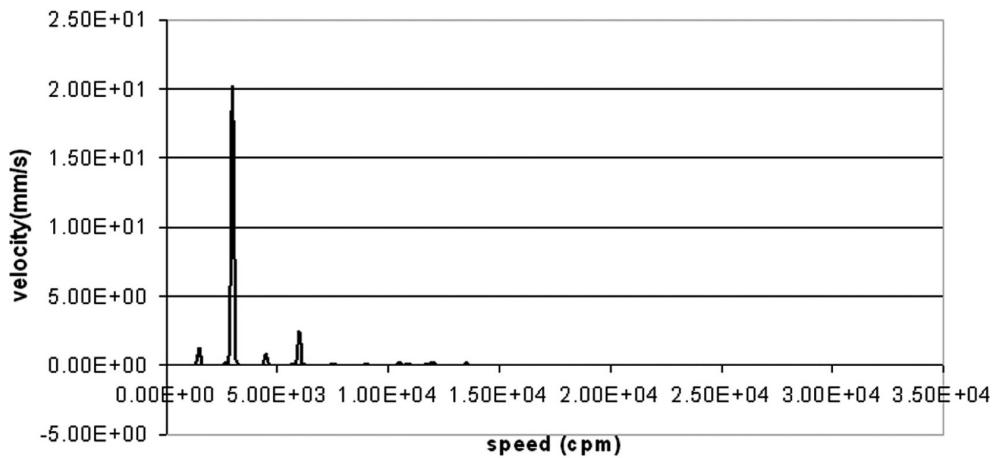


Fig. 6. Horizontal vibration spectrum of overall amplitude 23 mm/s.

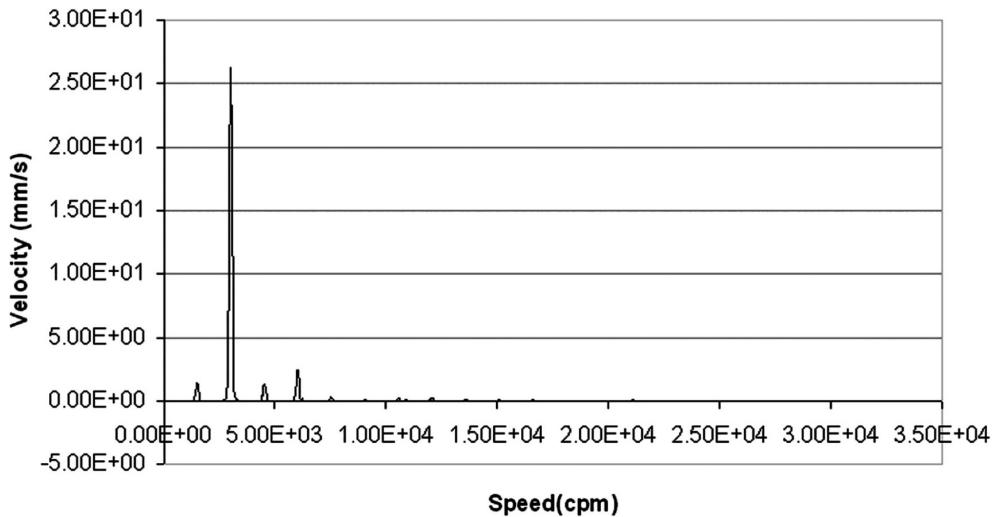


Fig. 7. Horizontal vibration spectrum of overall amplitude 28.6 mm/s.

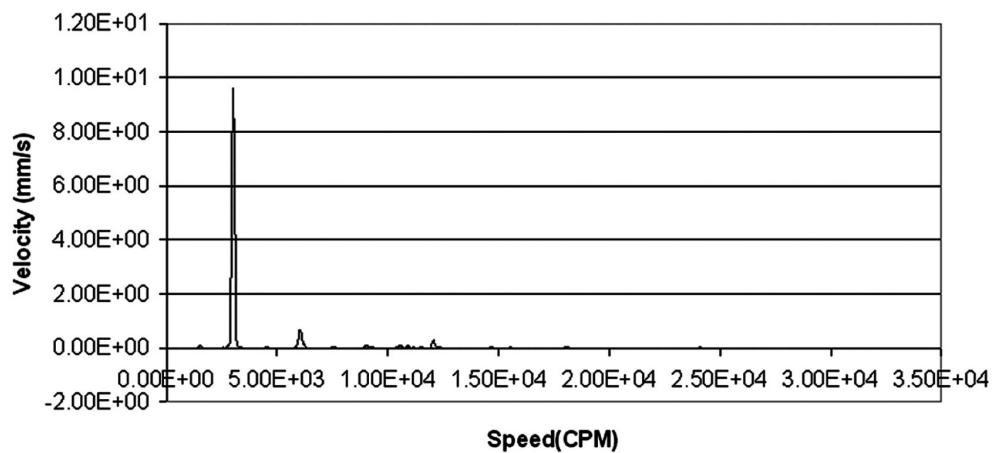


Fig. 8. Horizontal vibration spectrum of overall amplitude 9.93 mm/s.

Table 3
Test data set for proposed SOFM network

S.No	Parameters	Unbalance condition-1	Unbalance condition-2	Unbalance condition-3
1	Overall	23	28.6	9.93
2	f1	0.55398	0.0075714	0.0032141
3	f2	0.54908	0.61772	0.03893
4	f3	1.2215	1.4133	0.081243
5	f4	8.0601	9.8056	3.5556
6	f5	20.131	26.081	9.5382
7	f6	2.4373	2.3814	0.63869
8	f7	0.23828	0.26281	0.27324
9	f8	0.28774	0.023989	0.0025033
10	f9	0.0094676	0.00031756	0.000911
11	f10	20.131	26.081	9.5382
12	P1	N	N	N
13	P2	N	N	N
14	P3	L	L	L
15	P4	L	L	L
16	P5	N	N	N
17	P6	N	N	N
18	P7	Y	Y	Y
19	P8	N	N	N
20	P9	N	N	N
21	P10	N	N	N

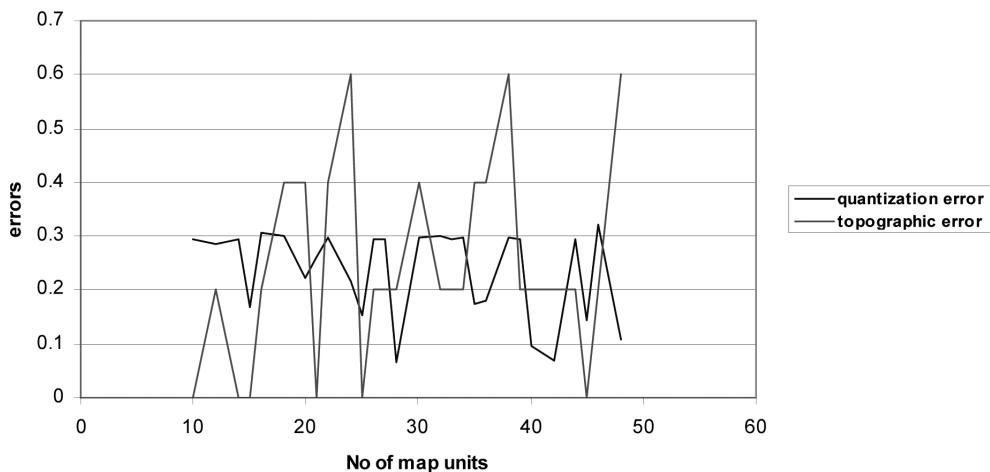


Fig. 9. Quantization and topographic errors as a function of the SOM map units. Selected optimal number of map units is 42 (6 X 7).

5. Results and discussion

Self-organization feature map, designed and realized for the analysis is trained only by data measured, when plant is known to be operating properly. To initialize the SOFM, random values in the interval (0,1) are assigned to the codebook vectors. In training, the network realizes a multi-dimensional domain representative of well operating conditions. When new input is introduced, the network defines a winning node, which is the nearest one to the input vector. Input is classified as acceptable if the Euclidean distance from the winning node does not overcome the alarm threshold, other wise the input data is considered to be novel.

Input data is the vibration signals obtained from an accelerometer mounted on front bearing of the generator in horizontal direction and process parameters collected from the general regular maintenance program. Input vector consists of 21 elements, constituted by the overall level of spike energy amplitudes at different harmonic and sub harmonic frequencies and processes parameters by assigning values '0' for 'NO' and 'LOW' answer and '1' for 'YES'.

Table 4
Parameters of the SOFM

Neurons Number	42
Initial topology	'Hexagonal'
Learning rate: ordering phase	0.9
Learning rate; tuning phase	0.02
Ordering phase steps	1000
Tuning phase neighborhood distance	1
Training epochs	100

Table 5
Results of the SOFM

Net phase	Input (acquisition date/condition)	Euclidean distance	Condition
Training	2 Feb.2009	0.0001	Normal
Training	16 Feb.2009	0.0000	Normal
Training	2 Mar.2009	0.0193	Normal
Training	9 Mar.2009	0.1328	Normal
Training	23 Mar.2009	0.2007	Normal
Test	Unbalance condition I	35.0649	Novel
Test	Unbalance condition II	45.6001	Novel
Test	Unbalance condition III	14.7183	Novel

The training data is extracted once in a week with the help of spectrums obtained during the inspection period. The test data is extracted from spectrums obtained on test rig by providing different unbalance conditions on the test rig. The input data consists of five vectors, which are related to the normal operating conditions of the plant and three data vectors are used to represent fault conditions of the test rig.

When the network is trained with a grid of 10 neurons (2X5), the quantization error is 0.2964 and topographic error is zero. The network is subjected to several runs by changing the number of neurons and topology of neurons. The quantization and topographic error are found optimum when the number of neurons is 42 from Fig. 9.

The parameters considered in Table 4 are used to configure the SOFM. The SOFM is designed by means of 42 neurons arranged in hexagonal topology.

The SOFM is trained through the ordering and tuning phase. Ordering phase lasted for 1000 steps. Neighborhood distance starts at the maximum distance between two neurons and decrease to the tuning neighborhood distance, which was set to 1. Learning rate started as the ordering phase learning rate is set to 0.9, and decreases until it reach the tuning-phase learning rate, set to 0.02. As the neighborhood distance and learning rate decrease over this phase, the neurons typically order themselves in the input space with the same topology in which they are ordered physically.

In the tuning phase, the neighborhood distance stayed at the tuning neighborhood distance and the learning rate continued to decrease from the tuning phase learning rate, but very slowly. The small neighborhood and slowly decreasing learning rate fine-tune the network. The number of steps that are good enough to obtain the stability of net is evaluated equal to 100.

The SOFM considered the input novel, as expected, reporting different Euclidean distances higher than alarm threshold, set to 1. The Table 5 show the results of the training and simulation phases. In training phase, for all five vectors, the Euclidean distance is below the alarm threshold; So the data is considered as normal data. In unbalance condition I, overall amplitude of vibration is 23 mm/s and the corresponding Euclidean distance is 35.0649. To balance the test rig, some trail weight is added to the rotor and in this condition the overall amplitude of the vibration is 28.6 mm/s and the Euclidean distance is 45.6001. After changing the position of trail weight as per balancing program, the overall amplitude of vibration is 9.93 mm/s and the Euclidean distance is 14.7183. The Euclidean distance between the input and winning node is high according to the fault severity.

6. Conclusions

Vibration monitoring for predictive maintenance is an important part of the overall condition monitoring process. The self-organization feature map provides an excellent exploration platform for analyzing vast quantities of spatially

oriented vibration data. In this scientific work, a reliable novel data detection system is investigated for a large steam turbine.

The simulation results and the results using actual data from operating power plants shows that the novel data detection method can be applied to identify the fault presence with less intervention of the human expert. Once the machine condition is detected as abnormal, then it can be diagnosis by the human expert or the computer based diagnostic system. Same technique can be used to classify all the faults of the turbo generator by training different fault conditions of data and by forming different clusters using self-organization feature map. The further improvement of the quality of system can be obtained by training large number of data vectors and by considering different quality measures of self-organization feature map.

Nomenclature

ANN	Artificial Neural Networks
BMU	Best Matching Unit
DE	Driving end
EXC	Exiter
FFT	Fast Fourier Transform
GEN	Generator
H.P.T	High Pressure Turbine
I.P.T	Intermediate Pressure Turbine
L.P.T	Low Pressure Turbine
M	Number of neurons
Min	Minimum
N, n, j	Number of variables in input vector
NDE	Non driving end
$N_w(t)$	Neighborhood vector of the winning neuron
SOFM	Self-Organization Feature Map
SOM	Self-Organization Map
t, I	Number of input vectors
w_{ij}	Elements of weights in weight matrix
W	Weight Matrix
w_w	Weight vector of the wining neuron
$X(t), x$	Input Matrix
x_i	Elements In Input Vector
$x_i(t)$	Input Vectors
$\alpha(t)$	Learning rate function

Acknowledgment

The authors would like to acknowledge the technical support provided by Bharath Heavy Electricals Limited (B.H.E.L) Research and Development division, Hyderabad.

References

- [1] W.C. Laws and A. Muszynska, Periodic and Continuous Vibration Monitoring for Preventive/Predictive Maintenance of Rotating Machinery, *Journal of Engineering for Gas Turbines and Power* **109** (1987), 159–167.
- [2] M.P. Blake and W.S. Mitchell, *Vibration and Acoustic Measurement Handbook*, Sparten Books, New York, 1972.
- [3] International Standards Organization, mechanical vibration of non-reciprocating machines-measurements on rotating shafts and evaluation, ISO 7919–1986.

- [4] C. Jackson, *Vibration Severity*, The practical vibration premier, Gulf publishing Co., Houston, Texas (1979), pp. 43–49.
- [5] A. Lifson, H.R. Simmon and A.J. Smalley, Vibration limits for rotating machinery, *Journal of Mechanical Engineering* (1987), 60–63.
- [6] D.L. Thomas, Vibration monitoring strategy for large turbo generators, In *Proceedings of the Institution of Mechanical Engineers-Vibrations in Rotating Machinery* (1984), pp. 91–99.
- [7] K. Dmitry and V. Dmitry, An Algorithm for Rule Generation in Fuzzy Expert Systems, In 17th International Conference on Pattern Recognition (ICPR'04) – Vol. 1 (2004), pp. 212–215.
- [8] M. Boukadoum, A. Bensaoula and D.A. Starikov, *A Neural Network Based System for Live Bacteria Detection*, In Proceeding (403–199) Artificial Intelligence and Applications, (2003).
- [9] M.J. Seow, V. Deepthi and K.A. Vijayan, *A Neural Network Based Skin Color Model for Face Detection*, In 32nd Applied Imagery Pattern Recognition Workshop (2003), p. 141.
- [10] S. Openshaw and C. Openshaw, *Artificial Intelligence in Geography*, Chichester, John Wiley & Sons, (1997).
- [11] T. Kohonen, *Self-Organization and Associative Memory*, 3rd Edition, Spring-Verlag, (1989).
- [12] L. Girardin, In *5th International World Wide Web Conference*, Paris, France, Mapping the virtual geography of the World Wide Web, (1995).
- [13] C. Chen, *Information visualization and Virtual Environments*, London, Springer-Verlag, 1999.
- [14] S.I. Fabricant, Buttenfield B. Formalizing semantic spaces for information access, *Annals of the Association of American Geographers* **91** (2001), 263–280.
- [15] E. Berglund and J. Sitte, The Parameter less Self-Organizing Map Algorithm, *Neural Networks, IEEE Transactions* **17**(2) (2006), 305–316.
- [16] V. Crupi, E. Guglielmino and G. Milazzo, Neural-Network-Based System for Novel fault Detection in Rotating Machinery, *Journal of vibration and control* **10** (2004), 1137–1150.
- [17] A.J. Hoffman, N.T. Van Der Merwe, C. Stander and S.P. Heyns, International Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications, (2001), ISBN: 0-7803-7164-x, pp. 16–20.
- [18] Z. Gang, D. XiangJiang, K. Li and J. Diao, Data Mining for Fault Diagnosis and Machine Learning for Rotating Machinery, *Key Engineering Materials* **293** (2005), 175–182.
- [19] K. Kiviluoto, Topology preservation in self-organizing maps, *Proceedings of IEEE International Conference on Neural Networks* **1** (1996), 294–299.

