To the Graduate Council

I am submitting herewith a dissertation written by Israel E. Alguíndigue entitled "A Methodology for the Analysis and Monitoring of Vibration Data Using Artificial Neural Systems." I have examined the final copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Engineering Science.

Robert E Uhrig

Robert E. Uhrig, Major Professor

We have read this dissertation

and recommend its acceptance:

Belle R Upadhyaya

J. A. [illegible]

William [illegible]

Accepted for the Council:

CW Minkel

Associate Vice Chancellor

and Dean of the Graduate School

A METHODOLOGY FOR THE ANALYSIS AND MONITORING OF VIBRATION DATA USING ARTIFICIAL NEURAL SYSTEMS

A Dissertation

Presented for the

Doctor of Philosophy

Degree

The University of Tennessee, Knoxville

Israel E. Alguíndigue

May, 1993

DEDICATION

A los Alguíndigue,

Norimi, Maria Saroja, Kumary,

Carmen Cecilia y Canmanie

ACKNOWLEDGMENTS

My deepest and most sincere gratitude to Dr. Robert E. Uhrig, my professor and advisor at the University of Tennessee. His guidance during these years has been invaluable. I am grateful for his personal and professional support.

I would also like to thank my dissertation committee members, Dr. Belle R. Upadhyaya of the Department of Nuclear Engineering, Dr. Joseph A. Boulet of the Department of Engineering Science and Mechanics and Dr. William F. Lawkins of the Department of Mathematics. They have each contributed substantially to this work. My gratitude also goes to my colleagues in the Nuclear Engineering Department, especially Dr. Thomas W. Kerlin, Head of the Department, for his advice, and Dr. Larry F. Miller and Dr. Rafael B. Pérez for their friendship. Many thanks also to Toby Koosman, Lynnetta Holbrook, Barbara Campbell, Lydia Salmon and Cheryl Stamm.

This project would not have been completed without the financial support, professional advice, and test data provided by the Département Surveillance Diagnostic Maintenance of Electricité de France in Chatou, France. I am grateful to M. Mazalerat, M. Morel and M. Georgel of EDF for making possible my visit to the laboratory in 1992. I benefitted from the technical advice of M. Alain Trenty, Mme. Min Cai and M. Claude Vincent, and from the camaraderie of my friends in the Department, especially M. Jean-Cristophe Benas, M. Pierre Brunel, Mme. Claire Cloarec, M. Enmanuel Fleuet and Jean-Luc and Veronique Morel.

My family has been a source of constant support during these years. My parents, Dr. and Mrs. Alguindigue, are always a great source of encouragement. I owe to them the most pleasant memories of this journey. My sisters Luilay, Tiuley and Carmencita, as well as my brother Ildemaro have also travel the ups and downs of this excursion, giving me always their love and kind reassurance.

Finally, I am indebted to all my friends, too many to be mentioned, but especially Rocky T. Beaty, Riza C. Berkan, Bruno and Fe Mary Centella, Ben and Katie Dowler, Evren E. Eryurek, Alba Fernández, Lamartine Guimaraes, Andreas Ikonomopoulos, José A. Llamas Arjona, Efi Pilarinu, Alberto Rodríguez, Eugene and Sheron Schlereth, Clinton and Stephanie Smullen, and Lefteri Tsoukalas. I have thanked God many a time for your friendship.

ABSTRACT

This dissertation deals with the use of artificial neural networks for the monitoring and diagnosis of components in nuclear power plants. The technology of neural networks provides an attractive complement to traditional vibration analysis because of its potential to operate in real-time and to handle data which may be distorted or noisy. The technique enhances traditional vibration analysis and provides a means of automating the monitoring and diagnosis of vibrating equipment.

The study is conducted in two phases. First, the mechanical behavior of rotating machinery is studied, and a neural network system is developed for the detection of faults in rolling element bearings (REB). The system learns association between features in the spectrum and operating states of the bearings, the system is tested using data from a aging simulation bench test.

A technique is presented for modelling the relationship among sensors in a machine that is shown to be very effective for identifying changes in operating states. REB's are especially interesting components because they are responsible for a large fraction of the malfunctions in manufacturing equipment. In an average nuclear power plant about one third of all measuring points is allocated to REB's.

The second phase of the study is related to the problem of the vibration of the internals of pressurized water reactors (PWR). The vibratory behavior of the internals in a PWR can be identified and monitored using ex-core neutron data from ionization chambers located outside the vessel. The data collected from these

detectors provide information regarding the behavior of mechanical components, the presence of contacts between the core barrel and the pressure vessel, and more importantly, a means of verifying the integrity of components in the system.

A neural network-based methodology is described for identifying the vibration mode of the core barrel, and for detecting a particular family of mechanical failures. Features are extracted from the neutron noise spectra, and used for training neural network models to identify the different states of vibration typically exhibited by PWR's. The technique was tested on data collected over 93 fuel cycles from twenty eight 900 MWe pressurized water reactors in France.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION AND NEURAL NETWORKS METHODOLOGY

INTRODUCTION

The main goal of this doctoral dissertation is to develop a methodology for automatic analysis of vibration data. The analysis, as it applies to mechanical systems and reactor internals, is difficult to automate due to the many peculiarities exhibited by these systems. Modelling of vibration phenomena is also a complex task, since the behavior of the components under a variety of operating conditions is not always known precisely, and representative data are difficult and expensive to secure.

In this dissertation, the problem of vibration in mechanical systems and PWR reactor internals is investigated. The availability of this equipment is vital to the operation of nuclear power plants, thus the interest in developing automated systems for assessing the condition of the machinery and its operational state. The study is conducted in two phases. The first part of the study is devoted to vibration in the realm of mechanical systems and the second part addresses the vibration of the internals of PWR reactors. In both cases, a methodology based on the technology of neural networks is presented that takes into consideration a number of characteristics typical of the components under investigation.

In the area of mechanical systems, the analysis is focused on monitoring and identifying the operational state of rolling element bearings (REB). This analysis is particularly important because REB failures are responsible for a large fraction of the malfunctions in manufacturing equipment. In a typical nuclear power plant, for example, one third of all measuring points are allocated to REB's [1].

Defects in REB's are generated on the raceways or the rolling elements by, among other factors: spalling, seizure, external particles, chemical corrosion, improper lubrication, false brinelling, and passage of electric current [1]. Defects originate as detachments of small fragments of metal, which may evolve into cracks and frayed surfaces, and which eventually cause disassembly of the REB. Once a defect appears, it is observable as a feature in the frequency spectrum because it produces sharp vibration impulses. The impulses are generated each time a rolling element passes over a defect and are detectable in most cases even at incipient stages.

The features that REB defects produce in the spectrum, generally in the form of peaks, are used by experts to identify faults and to monitor the condition of bearings in rotating equipment. For discriminating among defects, the experts use formulae which relate bearing geometry and rotational speed to frequency regions where peaks related to particular defects appear. However, the task of identifying these faults is complicated by a number of machine specific parameters.

Among the many factors governing the appearance of spectral features are the modulation of frequencies and the severity of faults. REB frequencies are often modulated by residual unbalance producing small-amplitude peaks in the vicinity of the running frequencies. Frequencies may also be modulated by other vibration creating components in the REB spectrum which are sums and differences of many frequencies. Spectral features may also appear in healthy bearings due to production tolerances, and some features may not appear at all until the bearing is loaded by unbalance forces [2]. The severity of the fault is another important consideration. Due to the process of wiping described later, features which are clearly seen in the early life of a defect, may disappear from the spectrum as the defect aggravates, making it possible for the fault to go unnoticed.

The difficulties make apparent the need for robust pattern recognition, modelling, and monitoring techniques that are designed in the context of vibration analysis, that utilize operation-related parameters, and are able to integrate and manipulate specific knowledge about the machine being diagnosed or monitored. In the first part of this dissertation a methodology is presented which is shown to be very effective in the processing of vibration data from rolling element bearings. The system described is tested on data provided by Electricité de France, both from bench tests and motor data from their operating power stations.

The second part of the study addresses the problem of the vibration of the internals of pressurized water reactors. In particular, the analysis is conducted from data of ex-core power detectors located outside the vessel, the objective being to detect contacts between the reactor vessel and the core barrel, as well as failure of mechanical components.

The spectrum calculated from the noise portion of the ex-core power signal depicts features that may be associated with particular behaviors of components in the structure. The methodology described in this dissertation uses parameters calculated from the spectra to discriminate among the different modes of normal behavior typically exhibited by one family of PWR reactors. The technique is also useful for identifying failure of mechanical components in the structure such as the hold-down spring, because mechanical failures cause changes in the shape and location of the features in the spectrum.

Contacts between the reactor vessel and the core barrel are produced by motion of the two components. The motion is induced by the flow of water entering the vessel, by hydraulic turbulence, and by many forces internal and external to the system which, combine to produce a very complex motion. Although the clearance between the radial keys of the core barrel and corresponding radial supports on the reactor vessel is very small (around 0.5 mm.), contacts must be monitored closely because they have a hammering effect on the structure that may produce wear of the contact surfaces and increase the risk of structural fatigue.

This part of the study also deals with the identification of mechanical failures in reactor internals. One example of such mechanical failures is the degradation of the hold-down spring, a device in the shape of a ring installed between the upper core support structure and the core barrel, that presses against the two components to constrain the movement of the core barrel. The spring undergoes a process of relaxation with aging, and the objective of this part of the study is to identify this type of mechanical failure from the power spectral density of ex-core detectors. The degradation of the hold-down spring is important, and may become dangerous, because it translates into greater freedom of movement of the internals [3].

The system for identification of normal vibration modes of reactor internals is tested with data from 28 different PWR plants in France. Simulated data is used for testing the system used to discriminate among different levels of degradation of the hold-down spring, because there are no cases of spring failure registered in French reactors. This work is described in Part II of the dissertation.

The original contribution of this work is focused on the development of a methodology for automatic processing of vibration signatures. The work addresses the following issues:

- Classification of vibration signatures into known categories of operational states. The classification scheme is developed from exemplary data and includes features for the integration of knowledge particular to the problem of vibration analysis.

- Clustering of data which is not labeled (i.e. an operational state is not included with the signature). This technique is able to incorporate knowledge about the problem (peculiarities of the vibration problem). In particular, features are included to address the weighted contribution of frequency features in the spectrum. As discussed above, some frequency features are directly related to the presence or absence of faults (i.e. peaks at the appropriate spectral region). Although the entire spectrum needs to be taken into consideration for classification, peaks that are located in the "strategic" frequencies contribute more to the solution than other features in the spectrum.

• Monitoring using sensor cross-reference information that permits the assessment of relative severity in mechanical vibration and of motion and vibration in the ex-core neutron data from PWR internals.

The work is organized in two parts. Part I documents the techniques developed for mechanical systems, and Part II addresses the problem of vibration of the reactor internals. The last section of the dissertation, contains conclusions, issues relevant to the work which merit further research, and recommendations for future work.

Each part of the document is cited by a roman numeral, each part in turn contains several chapters, including a discussion of the results. The numeration of figures, equations, and tables contains a roman numeral corresponding to the part to which it belongs, an arabic number corresponding to the chapter within a part, followed by another arabic numeral which corresponds to the figure, equation or table number. Figures, equations and tables corresponding to the general Introduction and Conclusion sections are cited with a single arabic numeral.

Part I

The first chapter presents a general introduction to the problem of vibration of mechanical systems. A description of spectral features generated by faults in rotating equipment is included, and the problems associated with the identification of faults in rolling elements bearings are outlined.

Chapter 2 describes the system designed for monitoring REB's. The chapter includes a discussion of one input / two output models, which provides the mathematical basis for the implementation of the system, followed by the neural networks implementation. The results of testing the system with a set of data from a motor in one of EDF's power stations is also included in this chapter. New descriptors are defined for the quantification of overall vibration levels based on the information provided by cross-referencing the sensors on the machine.

In Chapter 3 a system is described for the discrimination of vibration signatures from rolling element bearings according to their operating state. A neural network system is trained to perform this discrimination using data collected from REB's in a bench test where the process of aging was simulated. The input to the networks is a compressed version of the spectra, and its output is a class label corresponding to the operating state. Other simple networks are used for preprocessing the data.

Chapter 4 presents a technique for clustering vibration data using a modified version of the Fuzzy-C algorithm. Clustering is a very useful technique in vibration analysis, because in many cases the state of operation of a machine is not known precisely and clustering may unveil important characteristics about the behavior of the machine. The results of testing the system with data from REB's is also presented in this chapter. The data used for testing the system corresponds to a group of REB signatures representing transition states.

The last chapter, Chapter 5, presents the discussion of results of the techniques discussed in Part I. These techniques provide means for automating vibration analysis of rotating machinery, and may be used for equipment in nuclear power plants, as well as other manufacturing settings.

Part II

Chapter 1 presents an introduction to the problem of vibration of reactor internals, a description of the ex-core neutron noise signal and how it relates to the vibration of the internals, and a method for quantifying the motion. Finally, the relationship between components of the structure and specific regions in the power spectral density is discussed.

In Chapter 2, the neural networks-based classification system used for discrimination of neutron noise spectra is discussed. The system is used first to validate the neutron noise signatures, second to discriminate among the different modes of normal vibration, and third to separate signatures of reactors with operational hold-down springs from signatures of reactors with degraded or non-functional springs.

Chapter 3 describes a methodology for monitoring the behavior of the internals. This technique is similar to the method described in Chapter 2 of Part I, modified to the realm of reactor internals. The results and discussion of applying this technique to data from operating reactors is included.

NEURAL NETWORKS METHODOLOGY

An artificial neural network (ANN) is a computational system formed by a large number of nonlinear information processing units, highly interconnected and operating in parallel. The topology of neural networks has been contrived from the understanding of the functioning of biological nervous systems, and the study of artificial nervous-like systems has been inspired by the hope of achieving human-like performance in tasks such as automated speech processing and image recognition.

What makes neural networks different from other data processing techniques is the distributed manner in which computation is achieved and the high connectivity between units. Each unit, a processing element (PE), is a very simple device capable of performing elementary mathematical operations, and nonlinear transformations. These units are organized in layers, and connected by weights to resemble the structure of neurons (Figure 1) in the cerebral cortex portion of the brain [4].

Neural network models are crude attempts to mimic this ability of biological systems. The connection strengths are simulated using a number called a weight which ranges typically from -1.0 to 1.0. As expected, a negative weight has an inhibitory effect, and a positive weight has a excitatory effect on the information flowing between units. Conceptually, this design has the ability of processing information in parallel since information flows continuously and simultaneously. Another advantage is improved robustness; since there are so many nodes with only local connections, and because the information is spread over a large section of the system, damage to a few nodes need not impair overall performance significantly [5].

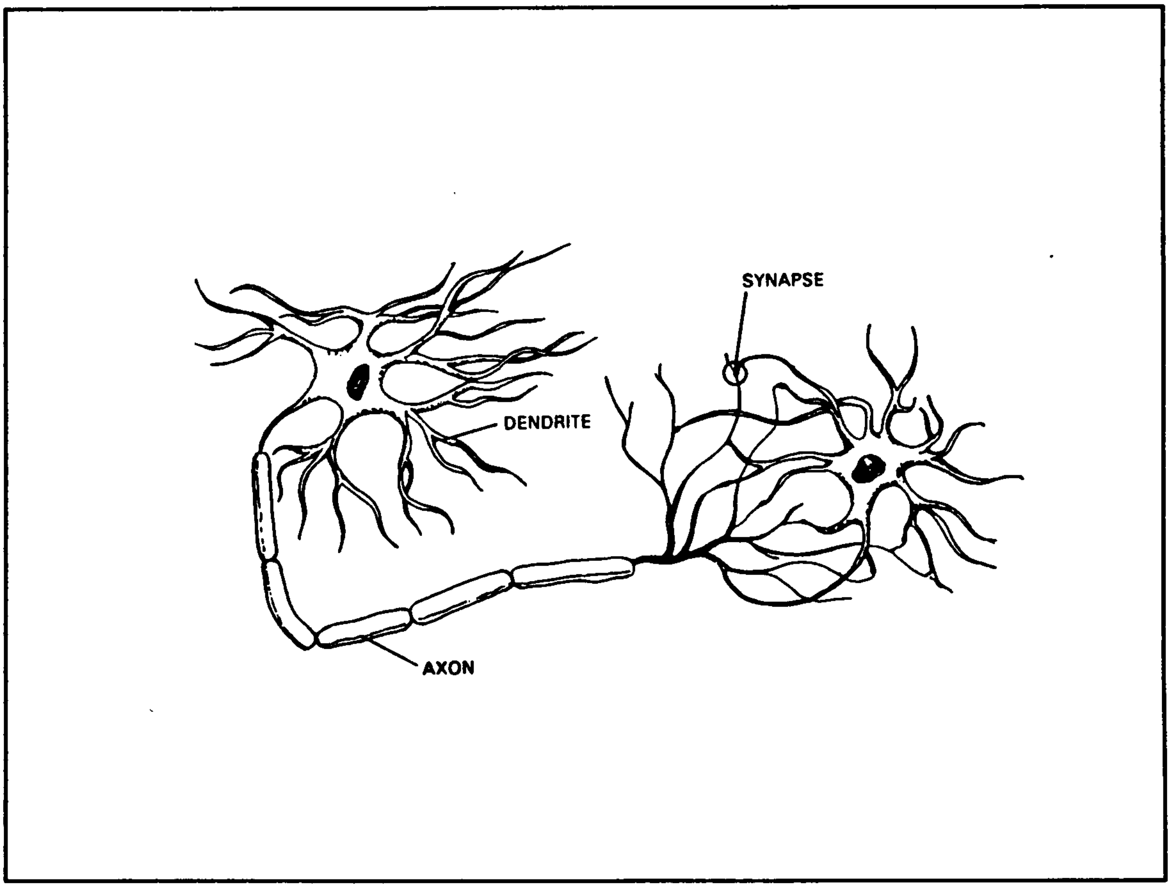

Figure 1. Biological Neuron [4]

Perhaps the most important characteristic of neural networks is their ability to model processes and systems from actual data. The neural network is supplied with data and then "trained" to mimic the input-output relationships of the process or system. Neural networks also have the ability to respond in real-time to the changing system state descriptions provided by continuous sensor inputs. For complex systems involving many sensors and possible fault types, real-time response is a difficult challenge to human operators; neural network technology may provide a viable alternative to the solution of this problem.

Neural networks may also be used to interpolate and extrapolate effectively from learned data, and more importantly, to model processes from actual process data. In the area of classification, neural networks work better than most traditional classifiers because they are non-parametric and make weaker assumptions concerning the shapes of underlying distributions [5], especially, when these distributions are generated by nonlinear processes and are strongly non-Gaussian.

Historical Perspective

The history of neural networks begins more than forty years ago with the work of McCulloch and Pitts [6], followed by the work of Rosenblatt [7], Hebb [8], Widrow [9,10], Posch [11] and others in the 1950's and 1960's. During the 1970's research in neurocomputing came virtually to a halt due in part to the publication of the book *Perceptrons* by Minsky and Papert [12], which pointed out alleged limitations and deficiencies of training the algorithms available. At the same time, other artificial intelligence areas started to achieve significant success, and the funding agencies shifted their investment to these emerging technologies.

The work by Rumelhart, McClelland, Williams and Hinton [13,14,15], Grossberg [16], Hopfield [17,18,19], Sejnowski [20], Specht [21,22], Kosko [23,24], Kohonen [25,26], Hecht-Nielsen [27], among others, along with the development of new theories and algorithms and the availability of faster and parallel hardware, contributed to the re-invigoration of the field. There was also a realization that the solution of certain problems requires a significant amount of computational power, and the recognition that traditional sequential computing, however fast, was totally inadequate for the solution of those problems.

Neural Networks for Vibration Analysis

In the realm of vibration analysis, neural network models may be used for monitoring components and for identification of operating states of machinery. ANN's offer a flexible, general-purpose approach to building highly non-linear models required for complex systems [28], and their ability to recognize patterns and perform matching in multi-dimensional spaces makes them ideally suited for the vibration problem.

For monitoring, neural networks may be trained to model the behavior of machinery under normal operating conditions, aiming to detect deviations from normal behavior. In this dissertation, for example, a neural network is used to model the relationship between sensors mounted on the equipment of interest. Generally, the sensors measure different components of vibration. In the case of rotating equipment the sensors measure radial and axial components, and in the case of reactor internals the sensors are opposite ionization chambers. Deviations from normal behavior may be detected and quantified by comparing the actual values of the sensors with the output of the model.

For identification of operating states, neural network models may be used as classifiers or clustering systems. In classification, the network learns the mapping between a signature describing the behavior of the component and the corresponding operating state. A set of parameters; such as time records, frequency spectra, vibration descriptors, etc; describing the behavior of the component under different conditions is supplied to the ANN and the network develops rules for discriminating between the different states. Recently, in the realm of rotating machinery, some ANN-based systems have been developed [29,30,31] to address this problem in a small scale.

In clustering, the network groups signatures according to their inherent similarity. Clustering is a very powerful idea in vibration analysis because in many cases the state of operation of the component is not known or cannot be accurately determined. Clustering produces groups of signatures, where each group may represent an operation state, such that the signatures in each group have similar geometrical and statistical features. If one or several of the examples are known to represent a certain kind of vibration state or fault, the other examples in the cluster have a high probability of representing the same condition.

any algorithms for neural networks have been developed, two of which are important to this program backpropagation for its intrinsic generality and well-deserved popularity, and the recirculation algorithm for its applicability to the areas of automatic feature extraction and data compression.

Backpropagation Network

The backpropagation network (BPN) is a multilayer, fully connected, feed-forward network. The name is derived from the nature of the training procedure while information flows forward from the input to the output layer, the error between the actual output and the desired output is propagated back through the set of weights. The training method aims at quantifying the contribution of individual units to the global error, and adjusting their weights accordingly. The learning procedure uses the delta rule [15], to compute the weights between connected processing elements so that the difference between the actual output and the desired output is minimized in a least-squares sense.

Backpropagation networks may operate in autoassociative and heteroassociative mode. For an autoassociative network, the input vector and the output vector are identical. A heteroassociative network learns associations between different input and output vector pairs. The basic BPN algorithm is presented by Rumelhart and cClelland in references [13,14,15] and is discussed here only briefly.

A backpropagation network consists of several layers an input layer, one or more hidden layers, and an output layer. The input layer acts as a buffer for the inputs to the system. It receives information (input vectors) and transmits the information to the hidden layers through weighted connections. The hidden layers are used by the network to create internal representations of the input patterns which represent the nonlinear characteristics of the data [15]. The topology of the network is shown in Figure 2.

The input layer receives the input to the system and passes it to the first hidden layer. The first hidden layer sums the weighted information from the input layer, modifies it using a logistic function (discussed later), and propagates the modified information forward to another hidden layer or to the output layer. The processing at each hidden layer is similar, and the information eventually reaches the output layer where it is compared with the vector of desired outputs.

In the output layer, the overall error is calculated. The error is the difference between the ideal (desired) output and the actual output generated by the network. This error is propagated back through the layers, and the weights are adjusted at each layer. This training process is repeated for each associated input/output pair until the error is reduced to a predefined level, or for a specific number of iterations, according to the criteria for termination. In essence, during training the information flows forward from the input layer to the output layer, and the error is propagated backwards through the weights to calculate the weight adjustments. Training involves the modification of the weights until the error is minimized.

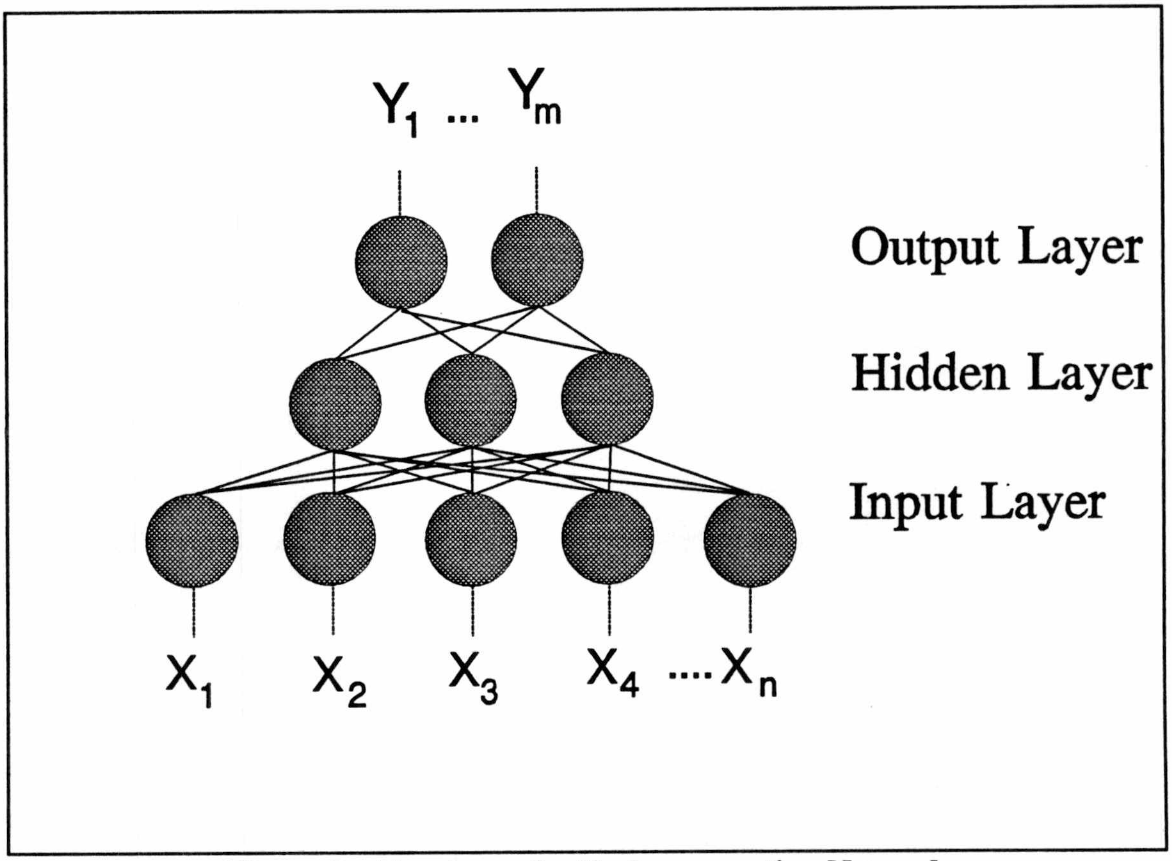

Figure 2. Topology of a Backpropagation Network

The logistic function acts as an automatic scaling factor and gain control. Another name given to this function is transfer function, and it is referred to by that name widely in the literature. The function provides high gain for small signal to prevent their attenuation, and small gain for large signals to prevent saturation [32]. The logistic function used most commonly is the Sigmoid of Equation 1.

$$F(X) = \frac{1}{1+e^{-\beta X}}, \qquad \beta > 0. \tag{1}$$

where x is the input to the function and β is the shaping parameter. In theory, the algorithm only requires that the function be bounded and everywhere differentiable [14]. Values range from 0 for large negative values of x, to 1 for large positive values of x. Other logistic functions are the hyperbolic tangent and the arc tangent where the values change between -1 and 1 as x changes from a large negative value to a large positive value.

In the fully-connected backpropagation model, every processing element in a layer is connected to all neurons in the next higher layer with no lateral connections. A BPN processing element calculates the scalar product of each input vector and its corresponding connection weights, and then filters this product using the logistic function of Equation 1 (see Figure 3). The output of a processing element in the input layer is the same input, in all other layers this output is given by Equation 2.

$$X_j^{[s]} = F\left(\sum_i W_{ji}^{[s]} * X_i^{[s-1]} \right) = F\left(I_j^{[s]} \right) \tag{2}$$

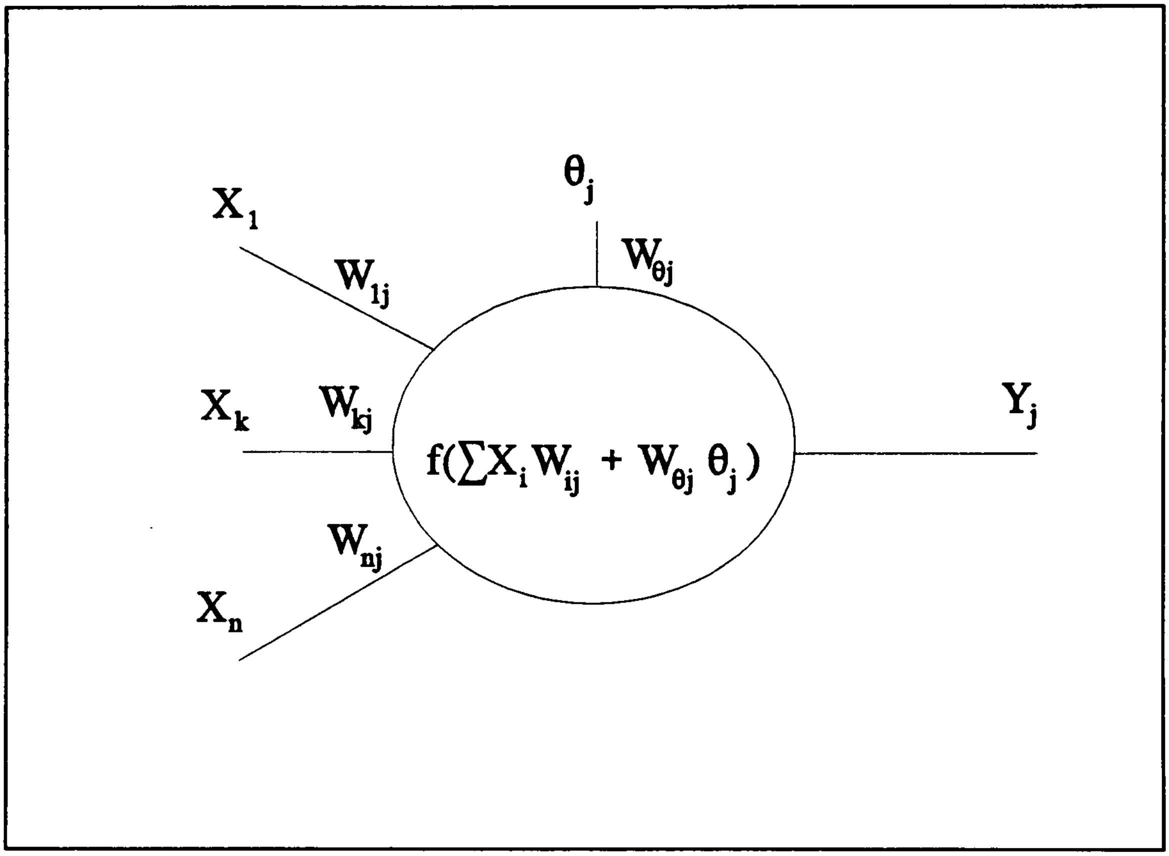

Figure 3. Artificial Neuron

where $x_j^{[s]}$ is the output of processing element j in layer s and $W_{ij}^{[s]}$ is the connection weight between unit i in layer $[s-1]$ and unit j in layer s, and $I_j^{[s]}$ is the weighted summation of inputs to the j^{th} neuron in layer s

The error function used generally in backpropagation is given by

$$E = \frac{1}{2} * \sum_{k=1}^{m} \left((d_k - o_k)^2 \right) \tag{3}$$

Where d is the desired output, o is the actual response, and m is the size of the output vector. The scaled local error at each processing element of the output layer is calculated with Equation 4.

$$\begin{aligned} e_k^{out} &= -\partial E / \partial I_k^{out} \\ &= \partial E / \partial o_k * \partial o_k / \partial I_k \\ &= (d_k - o_k) * f'(I_k) \end{aligned} \tag{4}$$

For non-output layers the local error is calculated using Equation 5,

$$e_j^{[s]} = X_j^{[s]} * (1.0 - X_j^{[s]}) * \sum_k \left(e_k^{[s+1]} * W_{kj}^{[s+1]} \right) \tag{5}$$

and the weight adjustment from Equation 6, where α is a learning coefficient.

$$\Delta W_{ji}^{[s]} = \alpha * e_j^{[s]} * X_i^{[s-1]} \tag{6}$$

The algorithm presented above is extracted from reference [33], a derivation is presented in references [13,14,15]. The key feature of the algorithm is that it performs gradient descent. To calculate the gradient of the error function E, the partial derivative of E with respect to each weight is calculated. The gradient points in the direction of maximum change; so to minimize E, the algorithm searches in the opposite direction. BPN can be credited for a large portion of the success that neural networks now enjoy. Variations of backpropagation have been developed, with claims of improved performance and faster convergence for some problems.

An open issue for BPN's is how to determine network configuration. It has been shown that a network with a single hidden layer and enough processing elements is capable of performing mapping between arbitrary inputs and outputs [34]. The problem is determining the correct number of hidden units. This number depends on the inherent complexity of the problem, which in most cases cannot be assessed easily. A number of theories had been proposed which provide some guidelines for selecting a configuration. We are using the criterion proposed by Kung and Hwang [35] which states that in the case of tasks with inherent regularities in the input/output patterns, the optimal number of hidden units is the number of regularity features minus one. In this case, the number of regularity features corresponds to the number of classes.

Recirculation Network

The recirculation network (RCN) algorithm was developed by Geoffrey Hinton and James McClelland [36]. The network is autoassociative, and it is known best for its ability to perform data compression and low-pass filtering. In the simplest version, the network has only two trainable layers: a visible layer and a hidden layer. The input and output layers act simply as buffers for input and output. The topology of the RCN is shown in Figure 4.

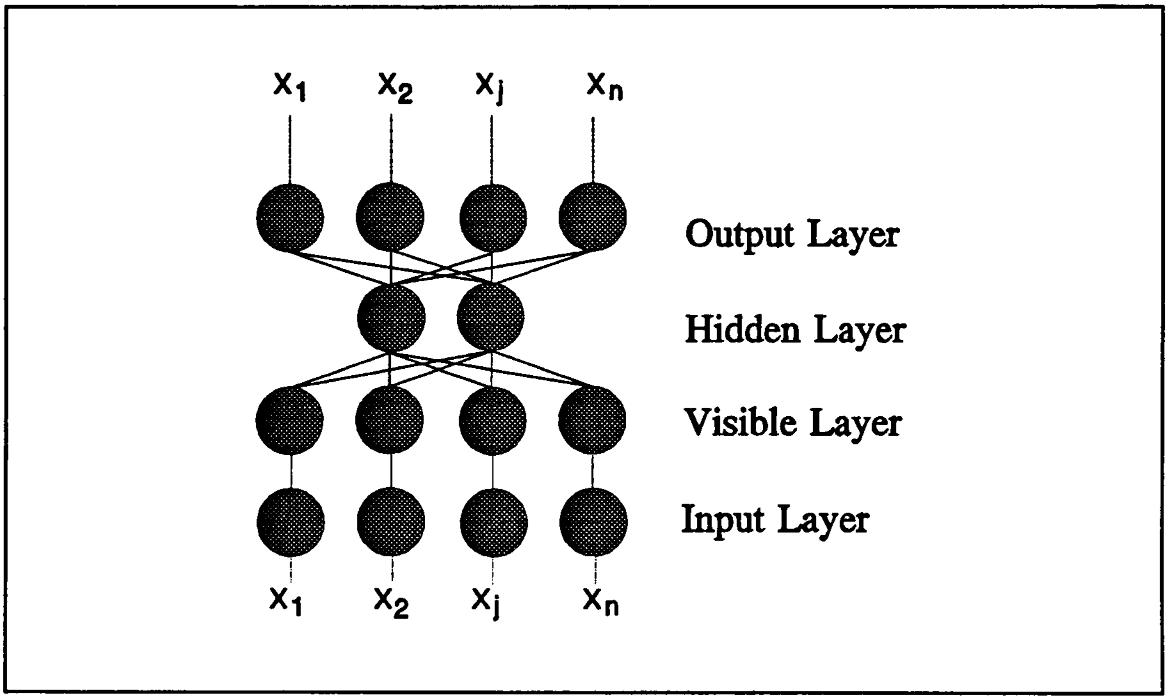

Figure 4. Topology of a Recirculation Network

When the number of units in the hidden layer is less than the number of units in the visible layer, the hidden representation may be considered to be a compressed version of the visible representation (see Figure 5). Under these conditions, the network acts as an encoder, mapping the original vectors to a smaller dimensional space while preserving most of their statistical characteristics. To encode information, the vector is presented as the input to the visible layer and its compressed version collected as the output of the hidden layer (the vector's dimension will be that of the hidden layer). Similarly, to decode or expand the compressed version of a vector, the compressed vector is presented as the input to the hidden layer and the decoded (original) vector can be collected at the output layer of the network.

The network contains groups of nonlinear units arranged in a closed loop [36], and separate paths with weights for forward and backward flow signals. The purpose of training is to find a set of weights which allows the activity vectors in the visible layer (the input vectors) to be represented by activity vectors in the hidden layer. To test the representation, visible vectors are reconstructed from the hidden vectors, and the error is calculated by comparing the reconstructed vector with the original (input vector). The purpose of training is to minimize this error, called the reconstruction error.

The reconstruction error is an accurate indicator of performance, defined as the sum of the squared differences between the input vector and the reconstructed vector. The reconstructed vector is assembled by passing the compressed

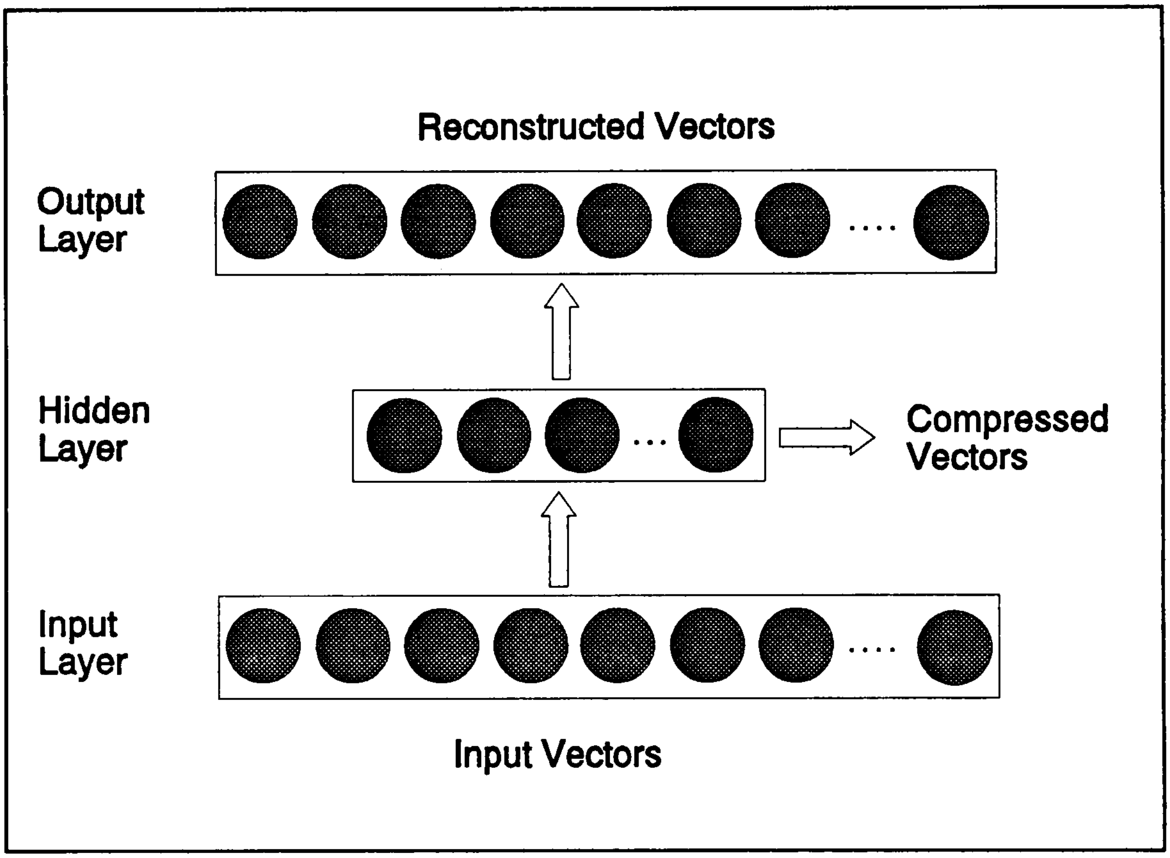

Figure 5. Compressor Network

representation through the weights, thus expanding the vector to its original form. During training, the error is minimized using a simple learning algorithm that behaves similarly to gradient descent under certain conditions.

The learning rule used in the recirculation network differs from the delta rule used in backpropagation in two critical areas [15]. First, when modifying the strength of a connection, only local knowledge is used; mainly the activation of the processing element and its input values. Second, recirculation does not require the connections to be symmetrical (i.e, $W_{ij} \neq W_{ji}$). The processing elements are connected in two directions such that W_{ij} (connection from node j in the hidden layer to node i in the visible layer) is not necessarily equal to W_{ji}, the connection from the i node in the visible layer to the j node in the hidden layer (Figure 6.)

The training algorithm for recirculation networks presented herein assumes the single hidden-layer topology depicted in Figure 4, and was extracted from references [33,36]. It is important to remember that each visible unit has a directed connection to each hidden unit, and each hidden unit in turn has a directed connection to each visible unit, and that these are separate connections, see Figure 7. This pattern of connectivity results in two different sets of weights: W_{ij} (for hidden to visible connections) and W_{ji} (for visible to hidden connections).

During learning, the data is presented at the input layer and buffered to the visible layer ($time\ 0$), see Figure 7. The data is then filtered through the bottom-up weights to the hidden later ($time\ 1$), sent back (recirculated) to the visible layer

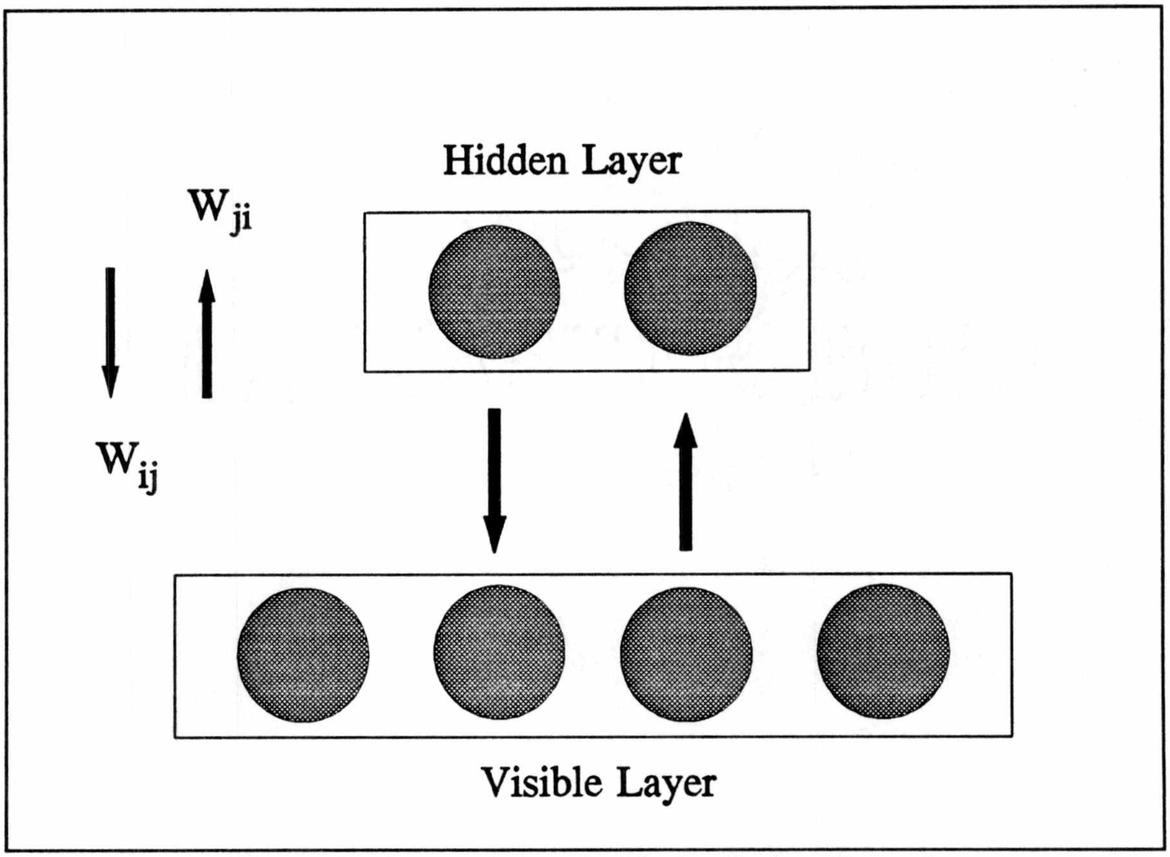

Figure 6. Connectivity of a Recirculation Network

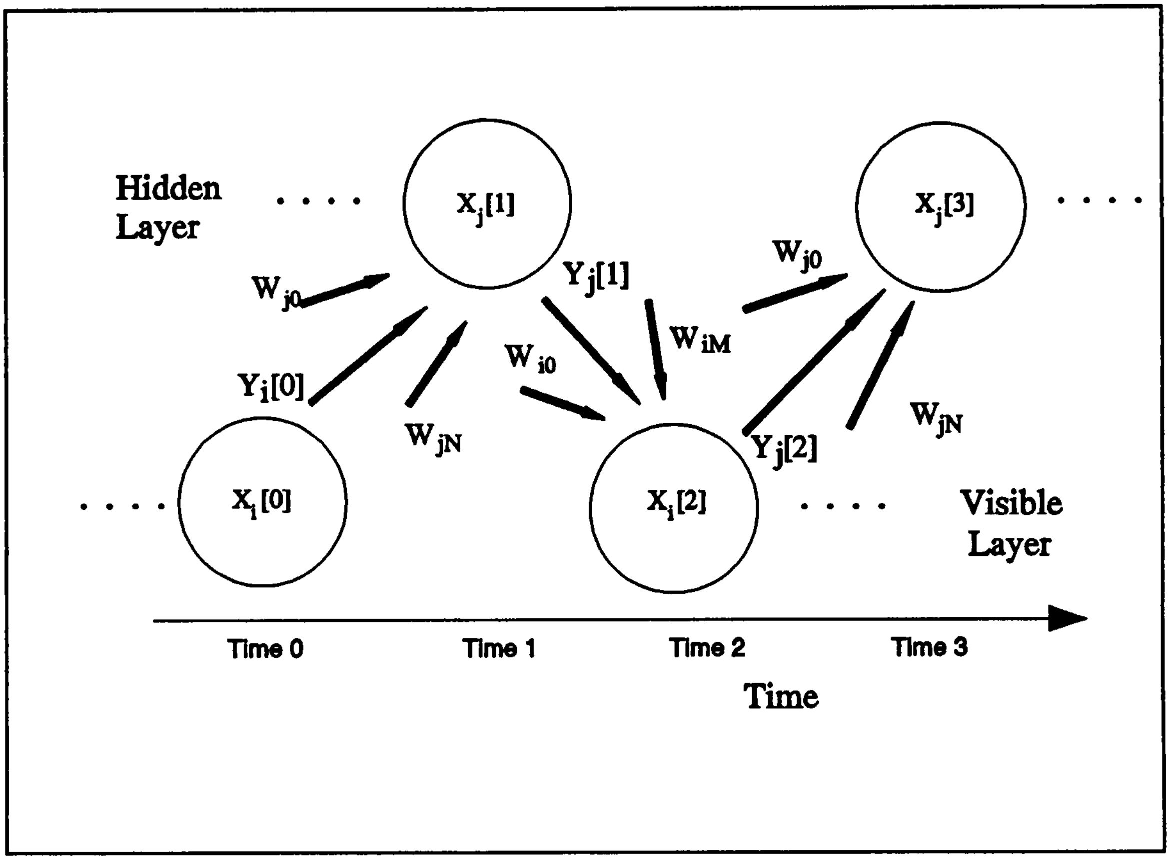

Figure 7. Time Sequence for Recirculation Training

through the top-down weights (*time 2*), and looped back to the hidden layer through the bottom-up weights (*time 3*). The weights are updated after the second pass, at (*time 3*). Equations 7 to 13 from reference [33] describe the sequence of events during a recall; $y_i[0]$ and $y_i[2]$ are the outputs of the i^{th} visible units at (*time 0*) and (*time 2*) respectively; and $y_j[1]$ and $y_j[3]$ are the outputs of the j^{th} hidden units at (*time 1*) and (*time 3*).

$$Y_i[0] = X_i[0] \tag{7}$$

$$X_j[1] = \sum_i (W_{ji} Y_i[0]) \tag{8}$$

$$Y_j[1] = f(X_j[1]) \tag{9}$$

$$X_i[2] = \sum_j (W_{ij} Y_j[1]) \tag{10}$$

$$Y_i[2] = r * Y_i[0] + (1.0 - r) * g(X_i[2]) \tag{11}$$

$$X_j[3] = \sum_i (W_{ji}\ Y_i[2]) \tag{12}$$

$$Y_j[3] = r * Y_j[1] + (1.0 - r) * f(X_j[3]) \quad (13)$$

where f and g are logistic functions in the forward and backward directions respectively. Hinton and McClelland [36] use the sigmoid function of Equation 1. The regression parameter r takes values between 0.0 and 1.0. The value of the parameter r is required to be close to 1.0 for convergence, and a value of 0.75 is recommended.

The hidden-to-visible connections (W_{ij}) are updated at (*time 3*) using the relation expressed in Equation 14

$$\Delta W_{ij} = \epsilon * Y_j[1] * (Y_i[0] - Y_i[2]) \quad (14)$$

and for the visible-to-hidden connections (W_{ji}), Equation 15 is used for updating the weights at (*time 3*)

$$\Delta W_{ji} = \epsilon * Y_i[2] * (Y_j[1] - Y_j[3]) \quad (15)$$

where ϵ is a learning parameter. From Equation 11,

$$Y_i[0] - Y_i[2] = Y_i[0] - (r * Y_i[0] + (1.0 - r) * g(X_i[2])) \quad (16)$$

and similarly from Equation 13,

$$Y_j[1] - Y_j[3] = Y_j[1] - (r * Y_j[1] + (1.0 - r) * f(X_j[3])) \quad (17)$$

Equations 16 and 17 may be used to rewrite the learning rule as,

$$\Delta W_{ij} = C_1 * Y_j[1] * (Y_i[0] - g(X_i[2])) \tag{18}$$

$$\Delta W_{ji} = C_1 * Y_i[2] * (f(X_j[1]) - f(X_j[3])) \tag{19}$$

where

$$C_1 = \epsilon * (1.0 - r) \tag{20}$$

For a given input vector, the squared reconstruction vector, E, is given by

$$E = \frac{1}{2} \sum_k (Y_k[2] - Y_k[0])^2 \tag{21}$$

Generally, the weights are not updated after each vector is presented to the network. Instead, individual updates (ΔW_{ij} and ΔW_{ji}) are accumulated, and their sum is used to change the weight matrices. This technique is known as cumulative learning and it is also used by some versions of backpropagation.

Early changes during the training of a recirculation network do not always improve the state of the system [36]. However, as learning progresses, and the set of weights in the two directions begin to align, the procedure resembles gradient descent and it performs similarly when the two sets of weights finally align. The advantage of the algorithm is that it mimics the behavior of gradient descent at a much lower computational cost, using a very simple topology and learning algorithm.

There are several extensions to the algorithm described here, which include cascading network and multi-layer RCN's. In the context of this study, the algorithm is used to perform feature extraction and compression of vibration data. However, good results have been reported on the use of RCN for problems of natural language processing and low-pass filtering [33,36,37].

Compression is an important issue in the context of this study because the data sets are generally very large, and by reducing the amount of data, processing time is decreased significantly. The added advantage to the compression is automatic feature extraction. Since in most cases the network cannot learn the mapping from high dimensional spaces to lower dimensional spaces exactly (this will require the reconstruction error to be zero), it acts as a feature extractor. The network focuses on the most relevant and consistent trends in the data [38], reducing noise and redundant information and making it easier for the classifiers to learn.

Scaling or Normalization of Parameters

The algorithms described, BPN and RCN, use the logistic function of Equation 1 for automatic scaling of the inputs to each neuron in the networks except those in the input layer. The logistic function is used to prevent attenuation of small signals and saturation for large signals. The function's output is a number in the interval (0.0, 1.0), which makes it necessary to scale the parameters of the system to the same range.

In practice, the value of the logistic function for very large and very small numbers approaches the limits asymptotically, and the changes in the output of the function in these ranges is negligible. Generally, the operating range of the algorithm is chosen to prevent the function from reaching the extremes. In the work described in this document the operating range for the algorithms has been defined as the interval (0.05, 0.95).

Scaling of the input variables may be done in a variety of manners. Generally it involves a linear transformation from the natural numeric range of the variables considered to some defined interval, in our case (0.05, 0.95), but other transformations have also been used for particular applications. Scaling may involve the adjustment of input variables using global parameters(global scaling), the parameters of individual variables (column-wise scaling), or parameters calculated from individual vectors (row-wise scaling).

For the systems described in this work, all three forms of scaling have been used. Global scaling is useful when the relationships of the entire data must be preserved. For example while making predictions, since the relationship between the input and output parameters is an important consideration. Column-wise normalization preserves the characteristics of individual variables in the set. In classification, it is necessary to treat the parameters individually and adjust their values according to their natural range. Finally, row-wise scaling helps in preserving the relationship between the elements of the same vector. For example, row-wise scaling is used when it was essential to preserve the amplitude of the peaks in the same spectrum, since other scaling schemes would destroy the shape of the signatures.

The technique of scaling is also called normalization, and it is referred to by that name widely in the literature. In this document we avoid using the term in this context because the analysis of spectral signatures from neutron noise signals discussed in Part II, requires a technique for adjusting amplitudes to which we refer as normalization. The formulae for performing global scaling, column-wise scaling and row-wise scaling are presented in Equations 22, 23, and 24. For all these formulae, A is the matrix of input vectors, and for A_{ij}, the subscript i corresponds to the vector number and j indexes the variables in each vector.

$$A^*_{ij} = \left(\left(\frac{A_{ij} - \min[A]_{ij}}{\max[A]_{ij} - \min[A]_{ij}}\right) * 0.90\right) + 0.05 \qquad (22)$$

$$A^*_{ij} = \left(\left(\frac{A_{ij} - \min[A]_{j}}{\max[A]_{j} - \min[A]_{j}}\right) * 0.90\right) + 0.05 \qquad (23)$$

$$A^*_{ij} = \left(\left(\frac{A_{ij} - \min[A]_{i}}{\max[A]_{i} - \min[A]_{i}}\right) * 0.90\right) + 0.05 \qquad (24)$$

PART I

ROTATING MACHINERY

CHAPTER 1

VIBRATION OF ROTATING MACHINERY

INTRODUCTION

A power plant (nuclear or fossil) is fundamentally a thermodynamic system that includes a heat source, flowing fluids, valves, control systems, and rotating machinery (pumps, fans, motors, gear boxes, turbines and a generator). Although the flow of fluids (water and steam) can induce vibration and shock, a primary source of vibrations is rotating machinery. The availability of this equipment is vital to the operation of the plant; thus, the interest in developing automated systems not only for monitoring the health of rotating equipment but also for estimating the time during which the machine may continue to operate safely.

The fundamental goal of programs for predictive and preventive maintenance is maximizing equipment reliability and safety while minimizing cost [39]. To accomplish this goal, it is essential to develop techniques which allow for the detection of faults at their onset, in order to prevent costly production delays and catastrophic failures. These techniques must also give time for scheduling maintenance activities and enable proactive maintenance management [39].

Embedded in this philosophy of predictive and preventive maintenance, is the assumption that data can be collected and analyzed to gain insight into the state of operation of the equipment. Moreover, in order to avoid disturbing the normal operation of the equipment, it is desirable that these data be collected in a non-intrusive manner with as little instrumentation as possible. The vibration signals acquired by rotating equipment constitute a means on which to base the maintenance activities.

Vibration per se is not necessarily bad if its amplitude and associated forces are within acceptable limits. Indeed, vibration in machinery can be the source of much information about the various systems involved. This information is helpful in assessing the state of operation of the components, and may be useful in performing, not only preventive maintenance, but also safety compliance.

There is a great deal of literature that describes the type of vibration signals to be expected for faults in typical mechanical vibration systems and analysis techniques that can be used for early detection of faults. Generally the analysis is done in the frequency domain, because features in the spectra may be linked to particular sources of excitation which can in turn be related to particular faults.

In the frequency domain, trend analysis is another useful technique [40]. It is performed by comparing the vibration spectrum for each machine with a reference spectrum and evaluating the vibration magnitude changes at different frequencies. Vibration amplitude changes at particular frequencies, which depend on the component and its operating conditions, can be related to types of faults and in a few cases to specific identifiable faults.

This form of analysis for diagnostics is performed often by maintenance personnel monitoring and recording transducer signals, and analyzing the signals to identify the operating condition of the machine. This method has proven to be effective in identifying potential failures before they occur, and indeed, it is sometimes possible to estimate the remaining life of certain systems once a fault has been detected. Hence, there is considerable motivation to design systems to automatically perform this analysis on a real-time basis in a reproducible manner.

VIBRATION MONITORING

Vibration monitoring is based on the principle that all systems produce vibration which may be analyzed to gather operational performance information. In rolling element bearings (REB), for example, the deterioration of a component is manifested in loss of fragments of metal of the raceways or the rolling elements. Each time a rolling element passes over a defect, an impulse of vibration is generated. The impulses are generally of very short duration and generate responses in a wide range of frequencies. The impulses are related to the fault and its severity, and all the vibration monitoring techniques are based fundamentally on the recording and quantification of these vibration impulses [41].

Vibration signals are usually collected using accelerometers. These transducers are favored because of their accuracy, lightweight, good temperature resistance, low cost and wide frequency response [42]. Generally, operation is restricted to about 20 to 30% of the natural frequency of the transducer [2], and for mechanical vibrations, the low frequency response is limited to approximately 5 Hz. [43,44]. While accelerometers provide an excellent mechanism for collecting vibration data, they are also highly sensitive to noise.

Noise is produced by many different sources, including the sensor's own instrumentation [42]. However, the main contributors are vibration from machine components -- other than the one being monitored, and the vibration of neighboring machinery. The noise contribution to the signal may be so large that, unless removed, it renders the signal useless for fault identification purposes. A variety of techniques have been developed for noise removal, some of which may be found in references [45,46].

The presence of noise complicates the monitoring and state identification tasks in two forms: by masking the true signal, and by increasing the vibration values beyond monitoring criteria when the component experiences no true sign of malfunction. Hence, monitoring systems are designed to cope with noise (by enhancing the signal and attenuating the noise), and alarm criteria are defined in the context of the operating environment of the equipment.

To monitor and diagnose vibrating equipment, signatures (time records) are collected from plant machinery, vibration spectra are generated, and these spectra are analyzed to detect features which reflect the operational state of the machinery. Generally when a machine is operating properly, overall vibration is small and constant (although "beat frequencies" may appear); however, when faults develop or when some of the dynamic processes in the machine change, the vibration spectrum also changes [47,48,49].

The frequency spectrum, generated using fast Fourier transforms, is very useful because in this domain low level signals are not masked out by the presence of high level signals, and because the Fourier spectrum (phase and amplitude) describes completely the vibration of the component. The frequency range of the spectrum must include all relevant components, with the range of observation depending on the type of system under study. It is also critical to register all peaks - regardless of their amplitude - because some of the important indicators have low levels.

SPECTRAL FEATURES GENERATED BY FAULTS IN ROTATING EQUIPMENT

There is a known relation between a wide variety of problems in rotating machinery and features in the frequency domain. In general, the location and amplitude of peaks in the spectra may be used as a means of detecting faults, since in most cases peaks are produced by non-normal behavior. There is, however, beat or "sympathetic" vibration in healthy machinery, which is also observable as peaks in vibration spectra [2,50].

Among the problems detectable by vibration analysis are: misalignment, unbalance, shaft bending, mechanical looseness, resonance, bearing defects, gear defects, and electrical malfunctions. These problems are common in power plants and each represents a topic of research in its own right. To gain some insight into the sorts of features that common malfunctions generate, a brief presentation will be given. The synopsis is organized according to the kind of data used. A complete discussion on the subject may be found in references [43,50,51,52].

Steady State (Constant) Speed

Typical spectral features generated by defects in rotating machinery are included in Table I.1.1. In the table, the term RPM refers to "revolutions per minute," and ***ax*** to "***a*** times the rotating speed of the machine." These two terms are equivalent and are used interchangeably in this dissertation. For a complete presentation on this subject, the interested reader may refer to reference [50].

It is important when making the analysis to consider four factors. First, the presence of the peaks in the spectrum; second, how the peaks relate to the machine speed (**1X** RPM); third, the amplitude of each peak; and finally, the relationship among the peaks (i.e. the presence of multiple harmonics and sidebands).

Another important piece of information which is often ignored is the phase relationship. As indicated clearly in reference [50],

> "Amplitude reveals how much something is vibrating. Frequency relates how many cycles occur per unit time. Phase completes the picture by showing just how the machine is vibrating."

Phase information helps in discriminating faults which produce similar features. For example, force unbalance and couple unbalance produce the same peaks (usually at **1X** RPM in the radial direction) but the couple unbalance tends towards 180 deg. out-of-phase, while the force unbalance will be in-phase. Phase is also helpful in the presence of multiple faults for determining the most prominent defect.

Table I-1.1. Features Generated by Defects in Rotating Machinery [50]

Angular Misalignment	Characterized by high axial vibration. 180 deg. out-of-phase across the coupling. Shows peaks at 1X (1 times the rotating speed) and 2X the RPM (revolutions per minute) in the axial direction.
Parallel Misalignment	Characterized by high radial vibration. 180 deg. out-of-phase across the coupling. Peaks are present at 1X and 2X RPM (radial) but 2X predominates. Severe cases of both angular and parallel misalignment may generate peaks at higher harmonics, especially in (4X - 8X). Peak amplitude depends on coupling type and construction.
Force Unbalance	Peak at 1X RPM (radial) dominates the spectrum. Amplitude increases by the square of the speed (2X speed increase = 4X higher vibration, see [50]). Signal is in-phase.
Couple Unbalance	Peak at 1X RPM (radial) dominates the spectrum. Causes high axial and radial vibrations. Amplitude increases by the square of the speed. Signal is 180 deg. out-of-phase.
Overhung Rotor Unbalance	Causes peak at 1X RPM (both in the radial and axial directions). Axial readings tend to be in-phase, radial readings may be unsteady [50].
Mechanical Looseness	Tends to produce directional vibration. There are three categories of looseness. The first is caused by structural looseness at the base or distortion or softening of the base. This kind of looseness produces a strong peak at 1X RPM in the radial direction, and approximately 180 deg. phase difference between vertical measurements on machine foot, baseplate and base itself. The second category is caused by loose pillowblock bolts, cracks in the frame structure or bearing pedestal. Dominant symptoms are peaks in the .5X, 1X, 2X (dominating) and 3X RPM. Finally, the last category of problems is caused by improper fit between component parts and produces many harmonics at multiples and half multiples of the rotating speed (RPM).
Gear Problems	Characterized by gear mesh frequencies (the number of teeth multiplied by rotational frequency). The difficulty is that this component will also appear in the vibration spectrum when the gear is good. High level sidebands or large amounts of energy at the gear natural frequency components are good indicators that a fault exists. Important to collect baseline vibration spectra because the high level components are common even in new gearboxes.
Blades and Vanes Problems	Characterized by high fundamental vibration or large number of harmonics near the blade or vane passing frequency (number of blades or vanes times the rotating frequency).
Bent Shaft	Dominant peak at 1X RPM (axial) if bent is near the shaft center; if bent is near the coupling, the dominant peak is at 2X RPM (axial). Causes high axial vibrations and axial phase differences of about 180 deg. on the same machine component.

Speed Transient Data

The great majority of the rotating equipment in nuclear power plants operates at constant speed. As a result, the main effort in the development of monitoring systems focuses on data collected during steady-state operation. There are problems, however, that can be identified using data collected during speed transient events (start-up or coast-down) some of which may not be detectable by normal, steady-state monitoring.

The main idea in speed transient monitoring is observing change in vibration signatures caused by excitation forces and dynamic rotor response as the machine changes rotational speed. This kind of analysis may be used to identify bearing problems. The response of the machine as it passes through a critical speed (rotor bearing resonant frequency) is controlled by the damping of the machine bearings, which relates directly to the condition of the bearing and the lubricant [53]. As presented in [53], the following problems can be detected using speed transient data taken over a wide range of speeds:

- Loose parts
- Cracked Shaft
- Bearing Degradation
- Lubrication Degradation
- Sticky Sliding Pedestals
- Structural Resonance
- Misrigged Valves.

The subject of speed transient monitoring is beyond the scope of this study. Recently much attention has been devoted to the subject and, excellent reports have been published. The interested reader may refer to [53,54,55].

In this study, we have focused on problems related to rolling element bearings (REB) at steady-state (constant) speed. The monitoring of these components is particularly important because REB failures are responsible for a large fraction of the malfunctions in manufacturing equipment. A particular case is a nuclear power plant; in an average plant, one third of all measuring points is allocated to rolling element bearings [43].

ROLLING ELEMENT BEARINGS

In REB's, defects are generated on the raceways or the rolling elements, induced by among other factors: spalling, seizure, external particles, chemical corrosion, improper lubrication, false brinelling, and passage of electric current [1]. Defects originate as detachments of small fragments of metal, which may evolve into cracks and eventually cause disassembly of the REB. Once a defect appears, it is observable as a feature in the spectrum because it produces sharp vibration impulses. The impulses are generated each time a rolling element passes over the defect and are detectable even at incipient stages.

For REB, there are formulas for calculating the frequency regions in which features may appear in the spectra. These features relate parameters in the geometry of the bearing, and give an indication of where in the spectrum to look for features related to a particular fault. For example Equation I-1.1 is used to identify the ball spin frequency (BPFIR) at which inner race defects appear (refer to Figure I-1.1). The features are also observable (attenuating progressively) at multiples of the calculated frequency. Equations I-1.2, I-1.3, and I-1.4 are used for calculating the ball pass of the outer race (BPFOR), the cage (CFF), and the ball spin frequency (BSF) respectively [2]

$$BPFIR = \frac{(n)}{2}\frac{(Fa)}{60}\left(1+\frac{d}{D}\cos\theta\right) \qquad \textbf{(I-1.1)}$$

$$BPFOR = \frac{(n)}{2}\frac{(Fa)}{60}\left(1-\frac{d}{D}\cos\theta\right) \qquad \textbf{(I-1.2)}$$

$$CFF = \frac{1}{2}\frac{(Fa)}{60}\left(1-\frac{d}{D}\cos\theta\right) \qquad \textbf{(I-1.3)}$$

$$BSF = \frac{D}{2d}\frac{(Fa)}{60}\left[1-\left(\frac{d}{D}\right)^2\cos^2\theta\right] \qquad \textbf{(I-1.4)}$$

For these equations, n is the number of rolling elements, Fa is the fundamental frequency, D is the distance between opposite rolling elements (measured from the center), d is the diameter of a rolling element, and θ is the angle of contact (See Figure I-1.1).

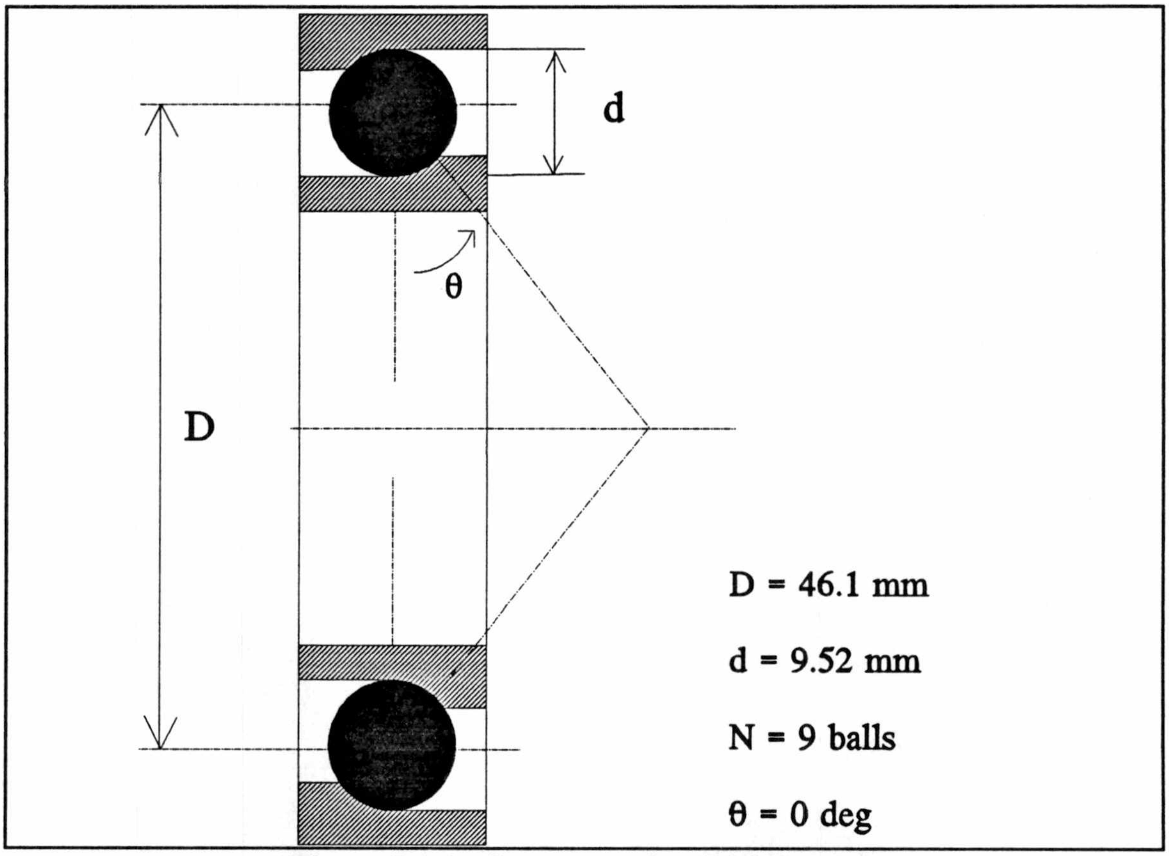

Figure I-1.1 Geometry of a 6206 Bearing

Some points are worth mentioning regarding Equations I-1.1 to I-1.4. To modify the formulas for a stationary shaft and rotating outer race, interchange the signs in Equations I-1.1 and I-1.2. The contact angle may change due to loading, creating small shifts in calculated frequencies. Finally, if bearing dimensions are not available, inner and outer race defect frequencies (BPFIR and BPFOR) may be approximated using Equations I-1.5 and I-1.6 respectively.

$$BPFIR \approx (0.6 * N * Fa) \tag{I-1.5}$$

$$BPFOR \approx (0.4 * N * Fa) \tag{I-1.6}$$

The approximations are possible because the ratio of ball diameter to pitch diameter is relatively constant for REB's [2]

Incipient faults are typically cracks or corrosion pits of the inner race, outer race or rolling elements. These defects may barely be seen with the unaided human eye, but appear as "heaps" in the high frequencies of the spectrum (above 5 KHz.) [51]. There are two reasons for the appearance of features in the high frequencies relating to incipient REB faults. The first one is that impulses generated are very short and contain energy high in the frequency spectrum. The second reason is the transmission of energy from the impulses to the bearing housing, which in turn excites fundamental modes. Bearing housings and other mechanical supports often have high natural frequencies.

Figure I-1.2 is the frequency spectrum of a healthy bearing; examples of bearing defects are presented in Figures I-1.3, I-1.4 and I-1.5. All figures correspond to data collected from identical bearings, under different operating conditions. Figure I-1.3 shows the spectrum of a bearing after it has developed a fault of the inner race. Figure I-1.4 and I-1.5 are plots of frequency spectra of bearings with faults of the outer race and rolling elements respectively. Note the appearance of peaks in the regions of the spectra calculated from the formulas presented above. These frequencies are identified below each figure

FACTORS AFFECTING REB SPECTRAL FEATURE IDENTIFICATION

Generally, machine defects produce spectral features which allow for recognition of different faults. Even when there are formulas for calculating the frequency regions in which peaks may appear in the spectra, the recognition task is complicated by a series of factors, such as noise and speed changes, which make the problem specific to components, and in some cases to operation conditions. For example in REB's, the specific spectral peaks depend on the type and severity of the fault, the rotational speed, and the bearing geometry.

To understand better the complexity of the task, it is important to mention some of the peculiar behavior of bearings. REB frequencies are often modulated by residual unbalance which produces sidebands at running frequencies. Sidebands are spaced at increments of plus and minus running speed from the defect frequency, and are the result of interaction between multiple vibration mechanism [2]. Sidebands

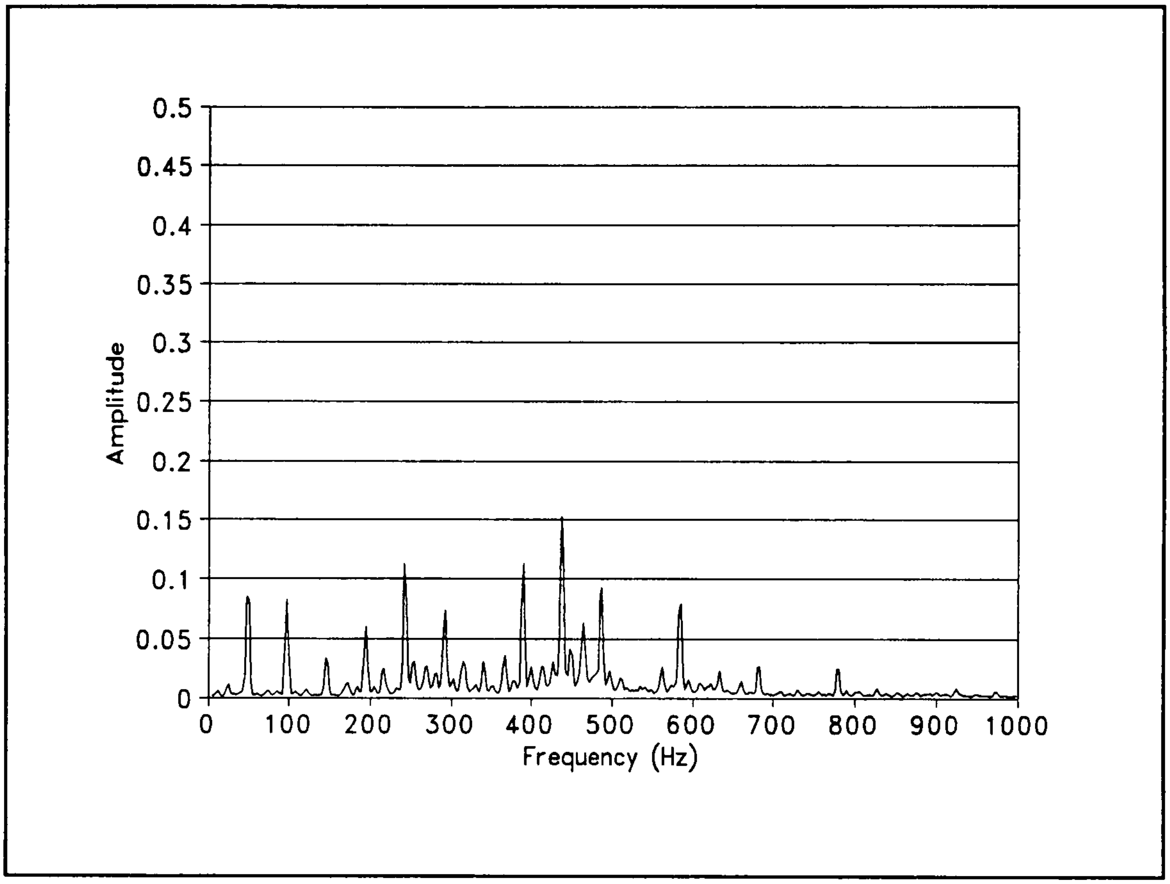

Figure I-1.2 Healthy Bearing (no Defects)

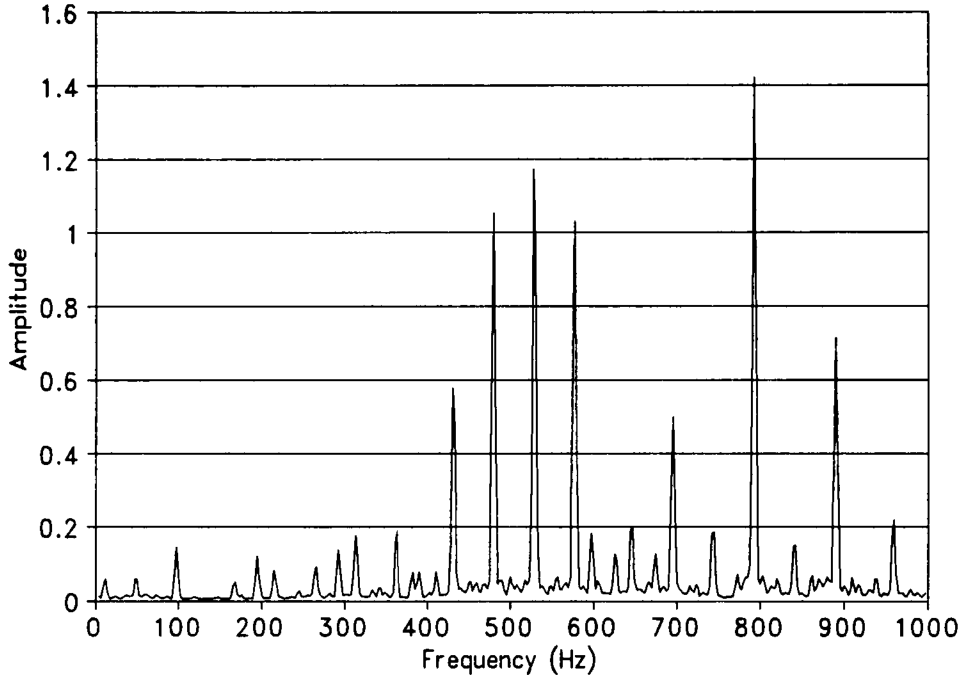

Figure I-1.3 Bearing with Localized Defect of the Inner Race
Ball Pass Frequency of the Inner Race = 265 Hz.

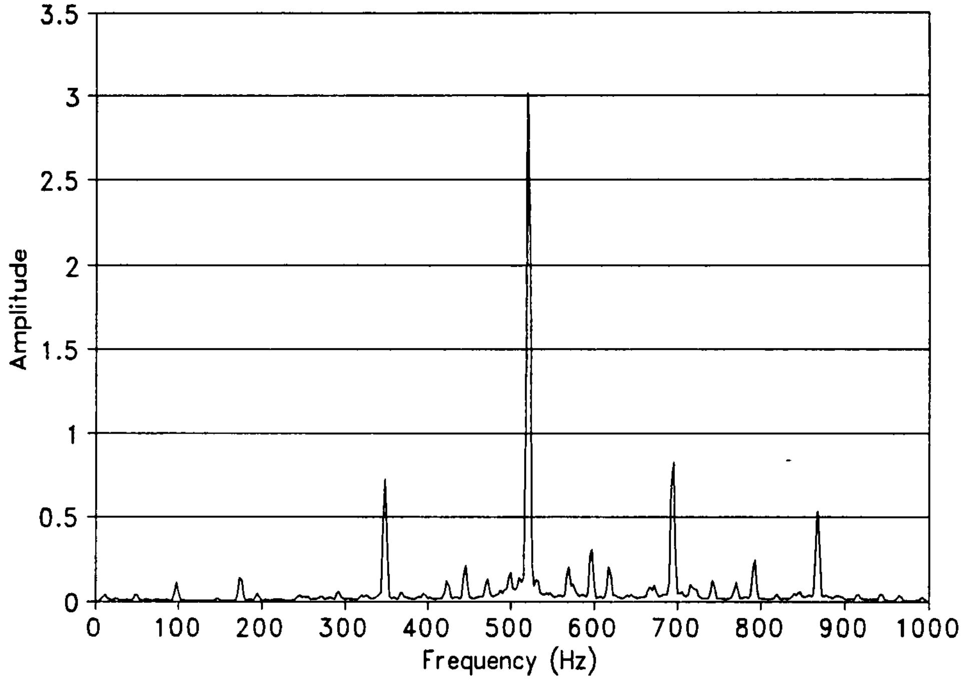

Figure I-1.4 Bearing with Localized Defect of the Outer Race
Ball Pass Frequency of the Outer Race = 174 Hz.

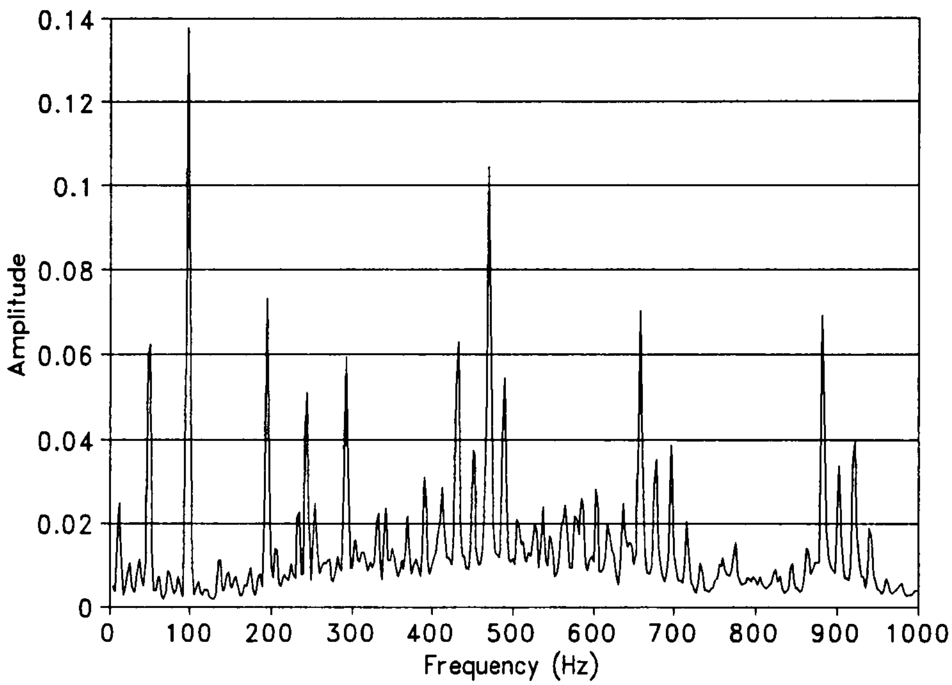

Figure I-1.5 Bearing with Defect of a Rolling Element
Ball Spin Frequency = 112.8 Hz.

are also referred to as sum and difference frequencies, and are very useful for discriminating between generalized and localized faults. Another interesting behavior is that spectral features may appear in healthy bearings due to production tolerances, and more importantly some defects on the stationary races may not show up at all until loaded by unbalance forces [2].

The recognition effort is complicated further by "wiping." This term refers to the process of leveling or smoothing the area surrounding a defect by the rolling elements as they pass through the defect repeatedly. As a result, the sharp impulses that are generated early in the life of the fault may not be present at a later time, which could make the fault go unnoticed. Under these conditions, vibration becomes more like random noise and discrete peaks will be reduced or removed. This may also be the case for roughness caused by abrasive wear or lack of lubrication [2]. Another common behavior at advanced stages of the defect is the disappearance of the fundamental frequency, while its harmonics remain [44]. A good example of wiping is presented in Figures I-1.6 and I-1.7. The spectra were calculated from data taken two months apart after the appearance of a defect. Notice that the vibration levels in Figure I-1.7 are lower than those in Figure I-1.6, although the defect has evolved into a more degraded state. The bearing failed a month after the spectra in Figure I-1.7 were calculated.

Even using the guidelines available, it is extremely difficult to detect faults in the spectra because **a)** the features are influenced by so many factors, and **b)** different faults are characterized by the same features. In many cases, the effort goes

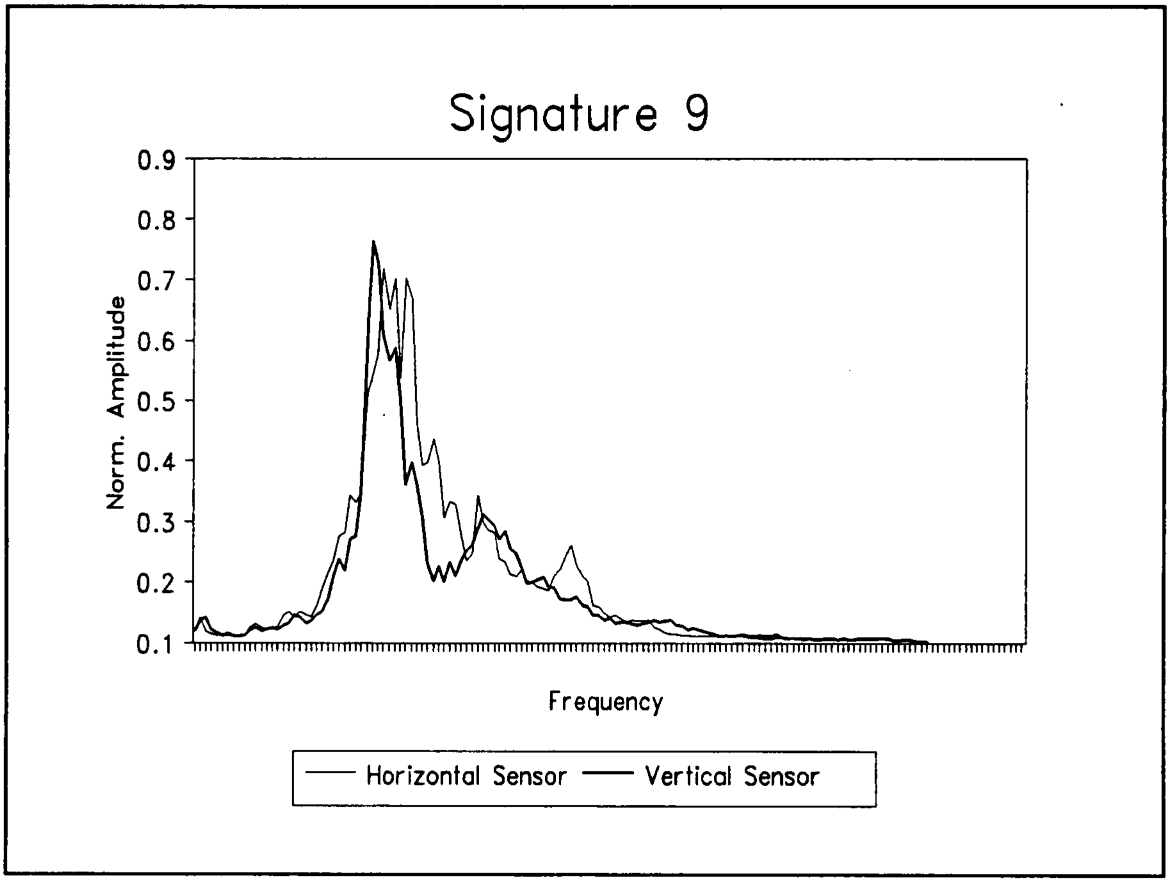

Figure I-1.6 Spectra of Defective Bearing, 3 Months Before Failure

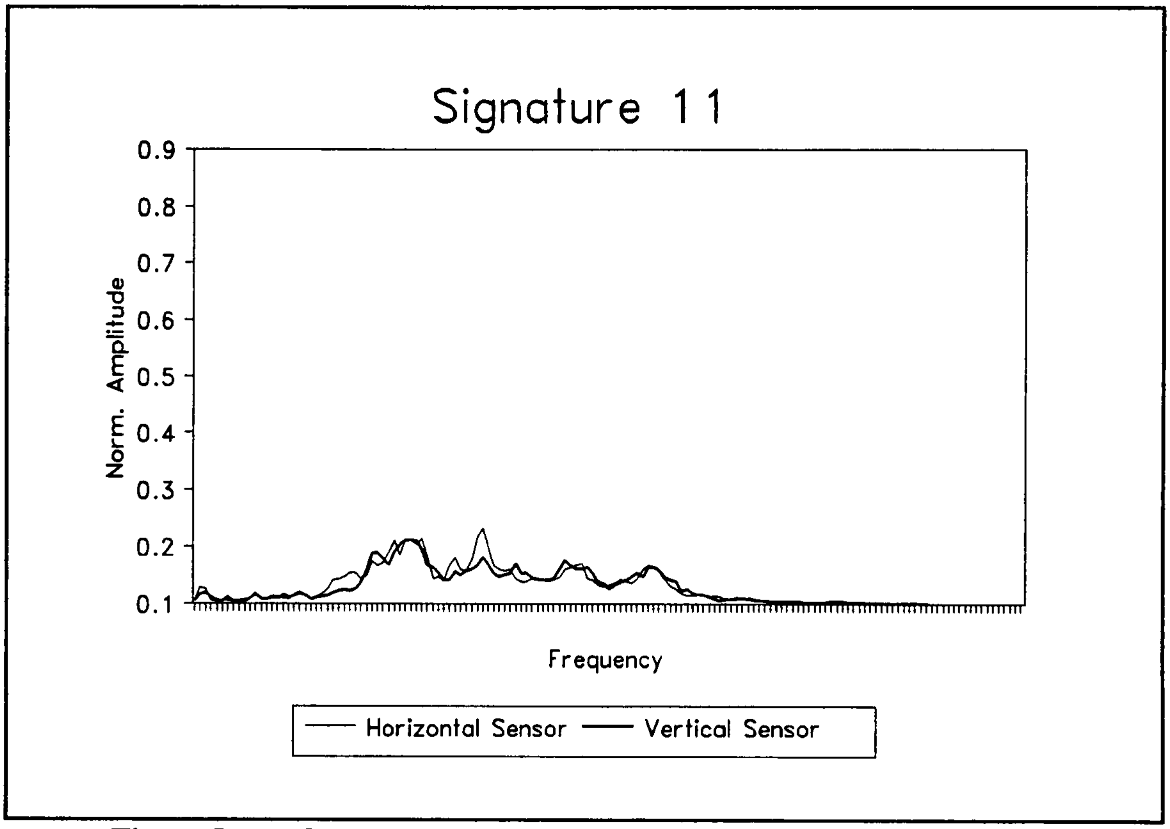

Figure I-1.7 Spectra of Defective Bearing, 1 Month Before Failure

beyond the frequency analysis to include observation of time records and orbit plots in an attempt to gain clues about the nature of the malfunction.

MACHINE SPECIFIC DIAGNOSIS

Traditional vibration analysis relies heavily on examining trends during operation, and comparing the vibration levels with limits found in vibration severity charts. While this approach has proven successful for assessment of machine condition, it relies on the assumption that vibration levels are low and constant during normal operation and that vibration increases due to faults are similar for machines of the same type. Recently, industry has acknowledged the need for a new methodology which presumes that for each machine there is a variety of normal characteristics; and more importantly, that two "identical" machines working side by side may exhibit totally different normal behaviors [56].

A study conducted by Pacific Gas and Electric Co. [56] concluded that among other issues:

- Identical machines working side by side on the same plant floor may exhibit widely different characteristics.
- Some machines normally exhibit vibration.
- It is common for a machine to change characteristics following an outage.
- Each machine may have "any number" of normal characteristics.
- There are normal characteristics that only occur when sets of events occur.

- Warning and alarm limits set on the basis of rule-of-thumb guidelines for overall vibration are often too low for some operating conditions, and result in nuisance alarms.

This study proposes to replace warning and alarm signals with "normal vs abnormal", since abnormal indicates a condition which would not be expected based on the historical records of the machine under normal conditions. The goal of the diagnostic system, according to the study, is "to provide sufficient information to determine action as rapidly as possible." Actions may include: changing the operating procedures slightly, investigating the acquired data more closely, or looking for clues to the changes in parameters not considered in the vibration system.

In order to implement a monitoring system based on this ideology, it is necessary to store large amounts of vibration data for each machine, or to implement a compact model of "normal" behavior. The model must be able to generate data corresponding to normal behavior which considers the peculiarities of the machine and its operating environment.

The difficulties outlined above make aparent the need for robust pattern recognition, modelling and monitoring techniques able to take into account a large set of operation-related parameters, designed in the context of vibration analysis, and able to integrate and manipulate machine-specific knowledge. A study is conducted on the applicability of neural networks to the problems of monitoring and identifying the operating state of mechanical systems, a methodology is proposed which is shown to be very effective in the processing of vibration data.

CHAPTER 2

MONITORING OF REB WITH ARTIFICIAL NEURAL NETWORKS

INTRODUCTION

Vibration monitoring is based on the principle that all systems produce vibration during operation. When a machine is operating properly, vibration levels are generally small and constant; however, when faults develop and some of the dynamic processes in the machine change, the vibration spectrum also changes [47,48,49]. For many machines the vibration frequency spectrum has a characteristic shape when the machine is operating properly, and it has other characteristics for different faults that may appear. Monitoring can be performed effectively by close examination of the spectrum and by identifying features which are typical of particular defects; indeed, some systems have been built based on this principle [1,57,58].

The spectral features associated with particular defects of rotating machinery depend on a series of factors, among these: the kind of machine, its geometry, the operating environment and the severity of the defect. Although abundant literature (such as reference [51]) can be found which documents the vibratory behavior of rotating components, the peculiarities that need to be considered for each case makes it difficult to design a general monitoring technique applicable to a wide variety of problems. In this dissertation, a neural network-based technique is proposed which uses actual sensor data to model the relationship between sensors on a machine. This technique can be used to verify the structural integrity of the vibrating component by comparing the output of one sensor with the output of other sensors at different measuring points. The comparison is useful for identifying relevant frequencies where changes in vibration levels occur, and in performing overall trend of vibration.

A hetero-associative neural network may be trained to map the output of one sensor to the output of a different sensor. The relevance of this mapping is that while the state of operation remains unchanged, the network can make accurate predictions of the output of the sensor. However, if and when the operational state changes, the prediction of the network and the actual output of the sensor will deviate by an amount that reflects the severity of change. When it can be established that the two sensors are related in a statistical sense, the deviation may be quantified and provides important information about degradation and deterioration of the components.

ONE INPUT / TWO OUTPUT MODEL

To examine the properties of multiple transmission paths between a force source and different outputs, consider a one-input/two-output model (Figure I-2.1). Because the amplitudes generated by defects in rotating equipment are small we use

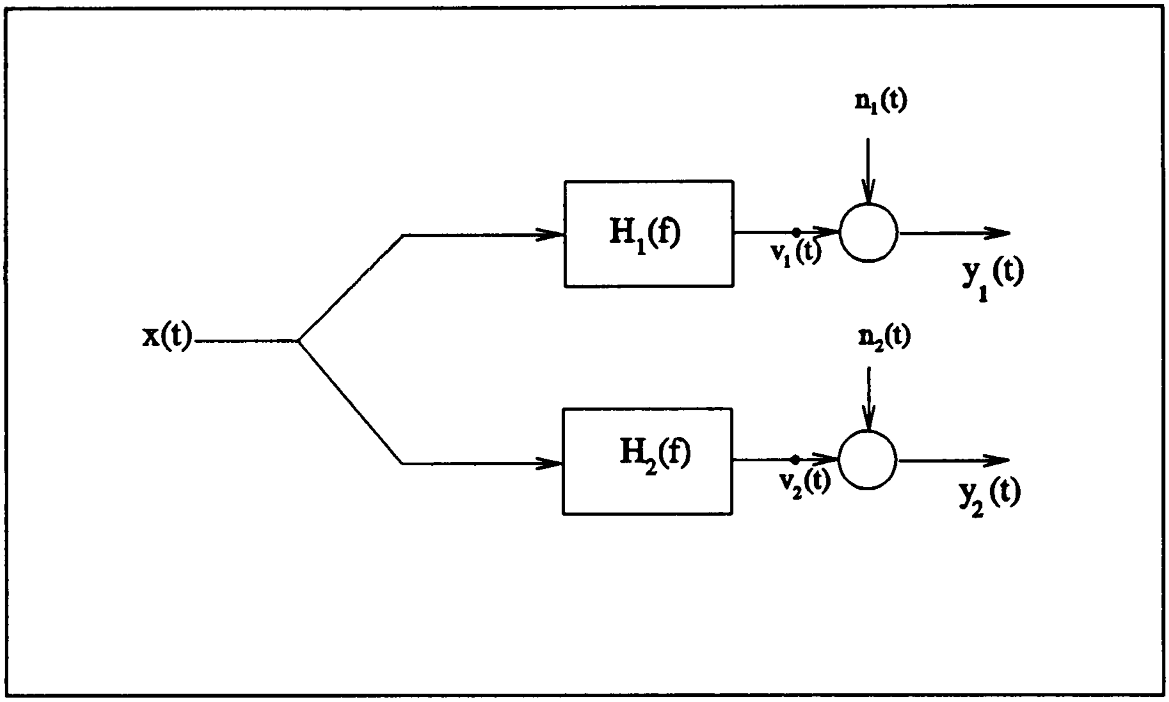

Figure I-2.1 One Input/Two Output System

linear systems theory to represent different paths, where uncorrelated noise accounts for any significant deviations from normal cases since perturbations have very small amplitudes. In the following equations $x(t)$ corresponds to the input signal, $y_1(t)$, $y_2(t)$ to the output signals and $n_1(t)$, $n_2(t)$ to the uncorrelated noise in each signal. Equations I-2.1 to I-2.3 from reference [59] may be used to describe the system

$$\begin{aligned}
&G_{xn_1} = G_{v_1n_1} = G_{xn_2} = G_{v_2n_2} = G_{n_1n_2} = 0 \\
&G_{y_1y_1} = G_{v_1v_1} + G_{n_1n_1} = |H_1|^2\, G_{xx} + G_{n_1n_1} \\
&G_{y_2y_2} = G_{v_2v_2} + G_{n_2n_2} = |H_2|^2\, G_{xx} + G_{n_2n_2}
\end{aligned} \qquad \text{(I-2.1)}$$

$$G_{xy_1} = G_{xv_1} = H_1\, G_{xx} \qquad G_{xy_2} = G_{xv_2} = H_2\, G_{xx} \qquad \text{(I-2.2)}$$

$$G_{y_1y_2} = G_{v_1v_2} = H_1^*\, H_2\, G_{xx} \qquad \text{(I-2.3)}$$

where G_{AB} is the cross-spectral density function of signals $A(t)$ and $B(t)$, G_{AA} is the power spectral density function of $A(t)$, and H_i is the transfer function for output $y_i(t)$. Because G_i and H_i are functions of frequency, the functional notation is dropped here for simplicity. The symbol '$*$' indicates complex conjugate.

The coherence function between the output records $y_1(t)$, $y_2(t)$ is given by Equation I-2.4

$$\gamma^2_{y_1y_2} = \frac{|G_{y_1y_2}|^2}{G_{y_1y_1}G_{y_2y_2}} = \frac{|G_{v_1v_2}|^2}{G_{y_1y_1}G_{y_2y_2}} =$$

$$\frac{|H_1^* \; H_2 G_{xx}|^2}{G_{y_1y_1}G_{y_2y_2}} \frac{|G_{xx}|^2}{|G_{xx}|^2} = \frac{|H_1^2 \; G_{xx}^2|}{G_{xx} \; G_{y_1y_1}} \frac{|H_2^2 \; G_{xx}^2|}{G_{xx} \; G_{y_2y_2}} = \tag{I-2.4}$$

$$\frac{G_{v_1v_1} \; G_{v_2v_2} \; G_{xx}^2}{G_{xx}^2 \; G_{y_1y_1} \; G_{y_2y_2}} = \frac{G_{v_1v_1}}{G_{y_1y_1}} \frac{G_{v_2v_2}}{G_{y_2y_2}} = \gamma^2_{xy_1} \; \gamma^2_{xy_2}$$

since,

$$\gamma^2_{xy_1} = \frac{|G_{xy_1}|^2}{G_{xx} \; G_{y_1y_1}} = \frac{|H_1 \; G_{xx}|^2}{G_{xx} \; G_{y_1y_1}} = \frac{G_{v_1v_1}}{G_{y_1y_1}}$$

$$\gamma^2_{xy_2} = \frac{|G_{xy_2}|^2}{G_{xx} \; G_{y_2y_2}} = \frac{|H_2 \; G_{xx}|^2}{G_{xx} \; G_{y_2y_2}} = \frac{G_{v_2v_2}}{G_{y_2y_2}} \tag{I-2.5}$$

When $\gamma^2_{xy} = 0$ at a particular frequency, $x(t)$ and $y(t)$ are said to be incoherent at that frequency. When $\gamma^2_{xy} = 1$ at all frequencies $x(t)$ and $y(t)$ are said to be fully coherent. So, the coherence function satisfies Equation I-2.6.

$$0 \leq \gamma^2_{xy} \leq 1 \tag{I-2.6}$$

The coherence function is unity for the ideal case of a constant parameter noise-free linear system with a single clearly defined input and output. When $x(t)$ and $y(t)$ are completely unrelated, the coherence function is zero. A value of coherence less than unity is due to noise in the measurements, non-linear relationship between the two signals or contribution of other signals to the output [59,60].

In cases where the input signal $x(t)$ cannot be directly measured, it is possible to use Equations I-2.1 to I-2.5 when the transfer functions H_1 and H_2 are found by other methods. However, if these functions are not known, Equation I.2.4 can be used to calculate the coherence function. A high value of the coherence function between any two output signals $y_1(t)$ and $y_2(t)$ indicates that $y_1(t)$ and $y_2(t)$ are related linearly to the input $x(t)$ by unknown transfer functions and that the output noise is small compared to the signal [59].

The linear system described may be used to model the normal behavior of the machine. However, if the process changes, the model no longer represents the behavior of the input/output relations and the relation between the two output signals no longer holds. Deviations from the true model can be detected and quantified if the relationship between the two outputs can be established. This relationship between $y_1(t)$ and $y_2(t)$ can be modeled using a neural network which receives as input the spectrum of $y_1(t)$ and produces as output the spectrum of $y_2(t)$.

The system may be extended to the single input/multiple output system. For this system, the coherence function is calculated between all two-signal combinations of output records. Equation I-2.1 through I-2.5 can be used by substituting the generalized subscripts i and j for the subscripts 1 and 2. The multiple output system may provide a better diagnostic technique because different combinations of pairs of output signals may be highly coherent at different regions of the spectrum. Ideally, each frequency of the spectrum will be coherent in at least one of these combinations and the behavior of the whole spectra can be observed.

NEURAL NETWORK SYSTEM DESCRIPTION

The linear system may be implemented using a neural network such that data from the two sensors is presented to the neural network, with the input being the spectrum from one sensor and the output being the spectrum from the other sensor. The network during training learns the relationship between the signals from two sensors and can be used to predict the behavior of one sensor from the other. These predictions make it possible to detect significant changes in relationship between sensor readings by comparing the actual output of the sensor with the prediction of the neural network. The changes become readily apparent when the state of the system changes because the actual value and the predicted value deviate.

Suppose that we are teaching a neural network the relationship between a horizontal sensor on a machine and a vertical sensor on the same machine, Figure I-2.2 illustrates this idea. At an operating state A, the values of the sensors can be denoted by s^v_A (corresponding to the reading from the vertical sensor at operating state A) and s^h_A(reading from the horizontal sensor at the same operating state). Similarly, these readings may be denoted s^v_B and s^h_B when the operating state is B. If we teach the network the associations between the sensors during the A operating state when we know the machine is operating correctly (the network is trained to predict the output s^v_A from s^h_A), then recalling the network with s^h_B produces a prediction of the output of the vertical sensor, s^v, under operating condition A (call this prediction P^v_B). The difference between the value of the sensor, s^v_B, and the

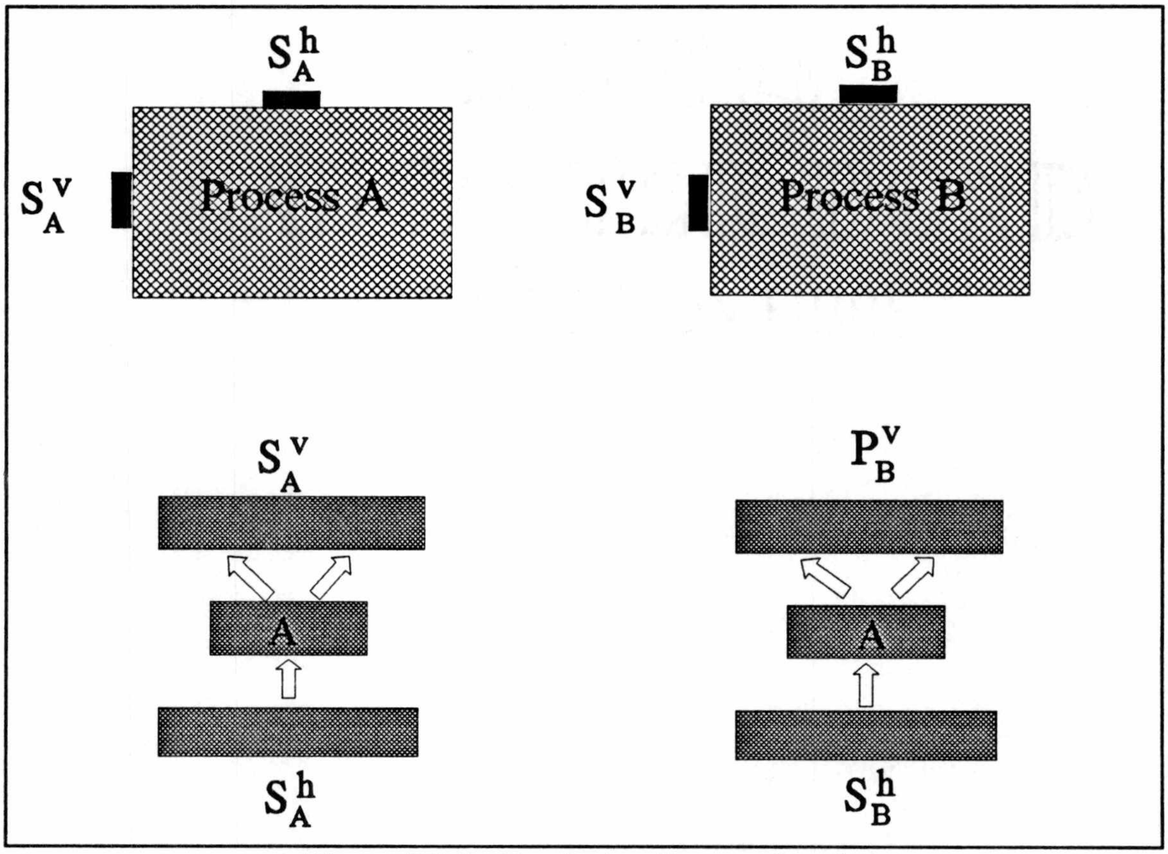

Figure I-2.2 Prediction System Set-Up

predicted value, $P^{\prime}_{B}$, may be used to detect significant deviations in the signatures and to quantify this deviation.

To establish statistical relationship, we examine the value of the coherence in the whole range of frequencies of the signal. This range depends on the application; for example in the analysis of vibration signatures from rotating machinery the range of interest may be 5 Hz to 1,000 Hz (base frequency) or 1 kHz to 20 kHz (high frequency), while in the context of reactor internals the range of interest is 0.5 Hz to 25 Hz. In those regions of the spectrum where the coherence is high we know that there exists a linear relationship between the two output signals [59], when the amplitude of the perturbations is small. This relationship is independent of the input $x(t)$ and should remain constant while the nature of the process does not change.

Once coherent frequencies have been identified, the change in vibration can be assessed by comparing the prediction of the network with the actual spectrum calculated from the sensor at the coherent frequencies. The comparison can be made using descriptors; for example, the horizontal descriptor which gives an overall measure of vibration amplitude from the horizontal sensor, a vertical descriptor describing the overall vibration levels in the vertical sensor. In addition, we are defining a cross-sensor descriptor which provides an overall measure of discrepancy between the two signals, and a per-frequency descriptor which represents the difference in magnitude and sign of the change in amplitude at each frequency.

The horizontal and vertical descriptors are calculated from effective velocity by integrating the output of the accelerometer in the appropriate range. The cross-sensor descriptor is a degradation indicator obtained by taking the square root of the sum of the squared differences at each frequency as shown in Equation I-2.7.

$$Cr.Desc=\sqrt{\sum_{i=1}^{n}\left(S_B^V(i)-P_B^V(i)\right)^2} \qquad \textbf{(I-2.7)}$$

where $S^V_B(i)$ is the actual value, and $P^V_B(i)$ the predicted value, both at frequency i, and where n is the number of frequencies in the spectrum.

The cross-sensor descriptor reflects an aspect of vibration that is critical for diagnostics: the level of disagreement among sensor outputs. This quantity is very useful for the detection of misalignment and looseness in rotating machinery because these defects produce uneven vibration levels in the vertical and horizontal directions. Of particular importance is the fact that the cross-sensor descriptor is less sensitive to wiping than the horizontal and vertical descriptors. In some cases after the component has developed a fault, the vibration descriptors calculated from independent sensor readings (vertical and horizontal in our case) return to normal levels, or even lower levels, while the value of the cross-sensor descriptor only decreases.

The cross-sensor descriptor is able to measure the activity even if one of the sensors does not register a particular impulse since the difference among sensor readings is reflected in the measure. Another useful indicator is called the per-frequency descriptor, it gathers the differences of the amplitudes at each frequency, indicating amplitude increases or decreases. The indicator identifies the regions of the spectrum where the activity has changed and is useful in the realm of vibration since some regions of the spectrum may be mapped to certain component defects.

MOTOR DATA DESCRIPTION AND DATA CONDITIONING

The monitoring system was tested on data from a motor in one of Electricité de France (EDF) operating power stations. The motor is an asynchronous horizontal electric motor (710 kW) driving a charging pump and the bearing monitored is of type NU 324 MPC3 manufactured by SKF [43]. The data were collected with two accelerometers mounted onto the machine at different measuring points - a horizontal sensor and a vertical sensor. Data from the motor were collected on 12 occasions, ranging from the normal condition of the bearing to failure due to serious cage wear. From the 12 time records collected for each sensor, high frequency spectra (1 kHz to 20 kHz) were generated. These spectra were used as input to a neural network, which was trained to model the relationship between the spectra of the horizontal sensor and the vertical sensor mounted on the motor. To compare spectra at different times, the spectral amplitudes were scaled globally in the range 0.1 to 0.9 using Equation 22. The first six spectra represent normal operation, the remaining signatures describe the behavior of the motor as it deteriorates. For reference we assign these signatures numbers between 1 and 12 corresponding to the order in which they were calculated.

TESTING ON MOTOR DATA

The predicting BPN network has 264 nodes in the input layer, 60 nodes in the first hidden layer, 30 units in the second hidden layer and 132 nodes in the output layer (Figure I-2.3). The first 132 input nodes correspond to the spectrum from the horizontal sensor at time $(t-1)$, the next 132 input nodes correspond to the spectrum of the horizontal sensor calculated from the next time series (time t), and the output nodes correspond to the spectrum of the vertical sensor at time t. Both input spectra -- time t and time $(t-1)$ -- are used to account for the nonstationarity of the process. The training was performed using three pairs of normal signatures (1-2, 3-4, 5-6) and the recall was performed on all the pairs.

Figure I-2.4 illustrates the agreement between the actual and predicted vertical sensors for a normal signature included in the training set. Figure I-2.5 shows the agreement between the two spectra for a normal signature not included in the training set. As can be observed, the neural network has learned the relationship between the spectra of the two signals and is able to predict accurately the spectra of the vertical sensor even when the vibration increases. During monitoring it is common to observe increases in vibration while the machine is still operating under normal conditions. The network is able to adjust to this increase and to continue modeling the relationship. As a result, the prediction is not based on a single baseline signature but on a collection of signatures representing normal operations.

Figure I-2.6 depicts the significant deviation between the actual and predicted vertical components for Signature 7. This signature describes the condition of the

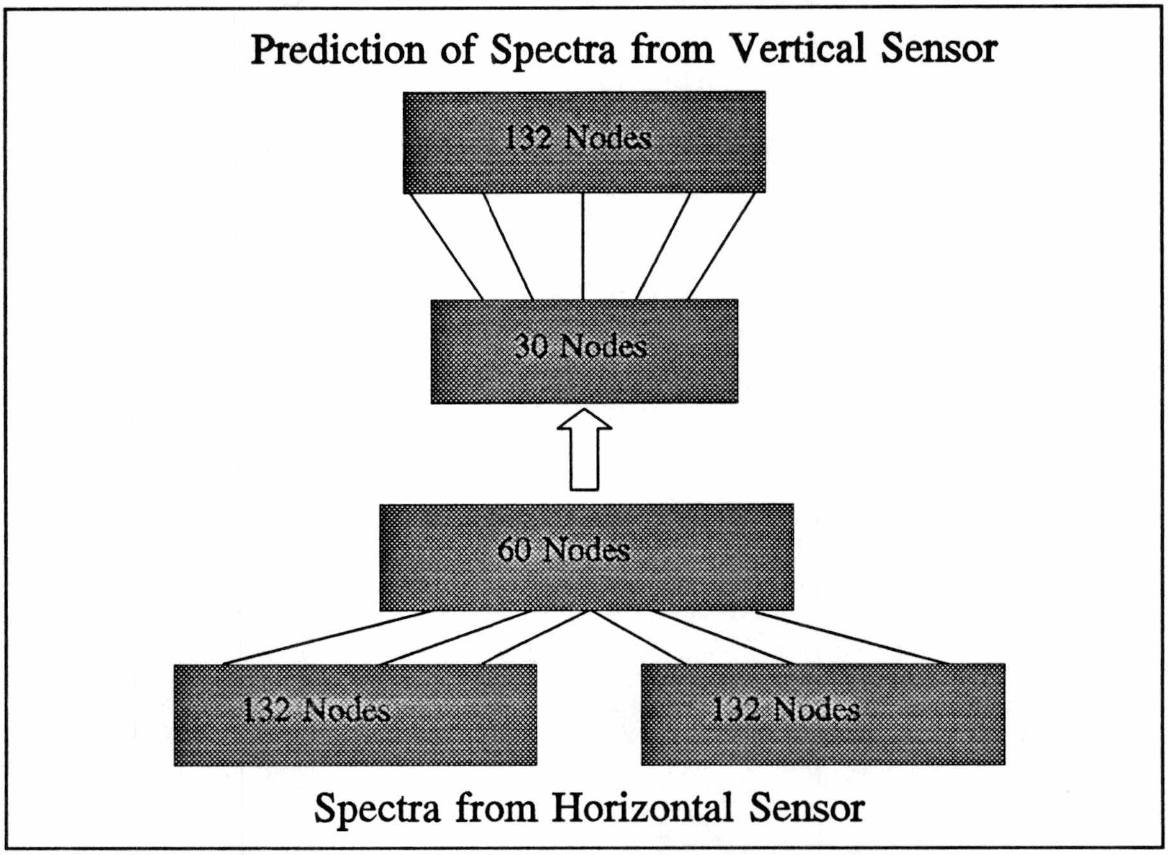

Figure I-2.3 Topology of the Predicting Network

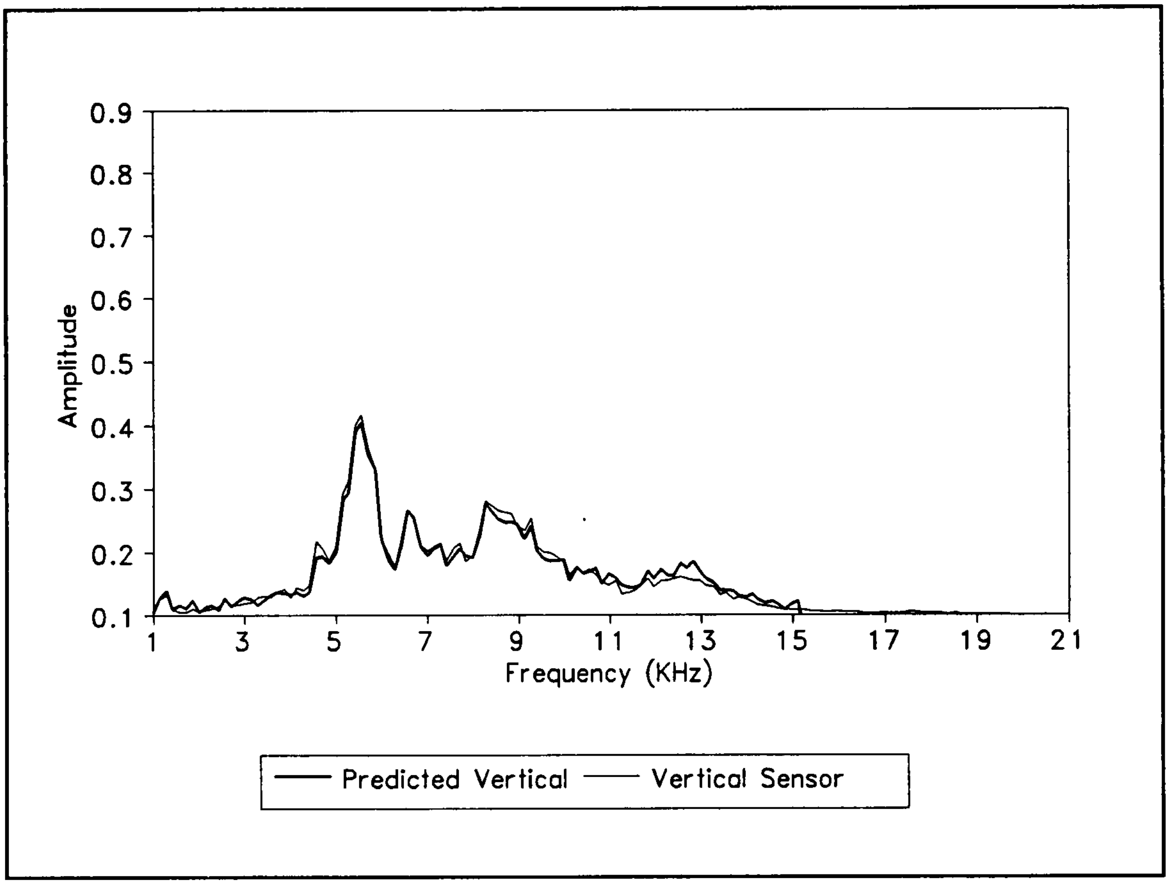

Figure I-2.4 Predicted and Actual Spectra for Signature 2
(Normal signature included in the training set)

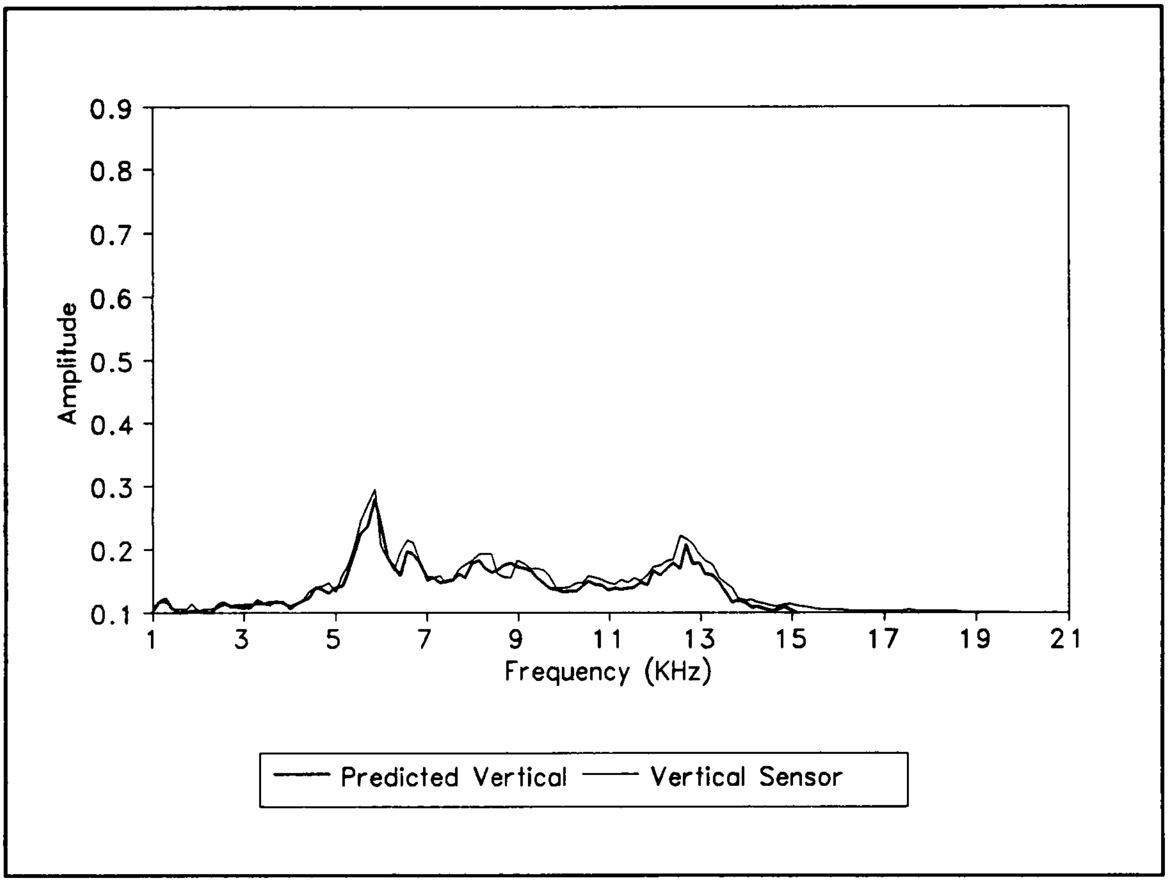

Figure I-2.5. Predicted and Actual Spectra for Signature 5 (Normal signature not included in the training set)

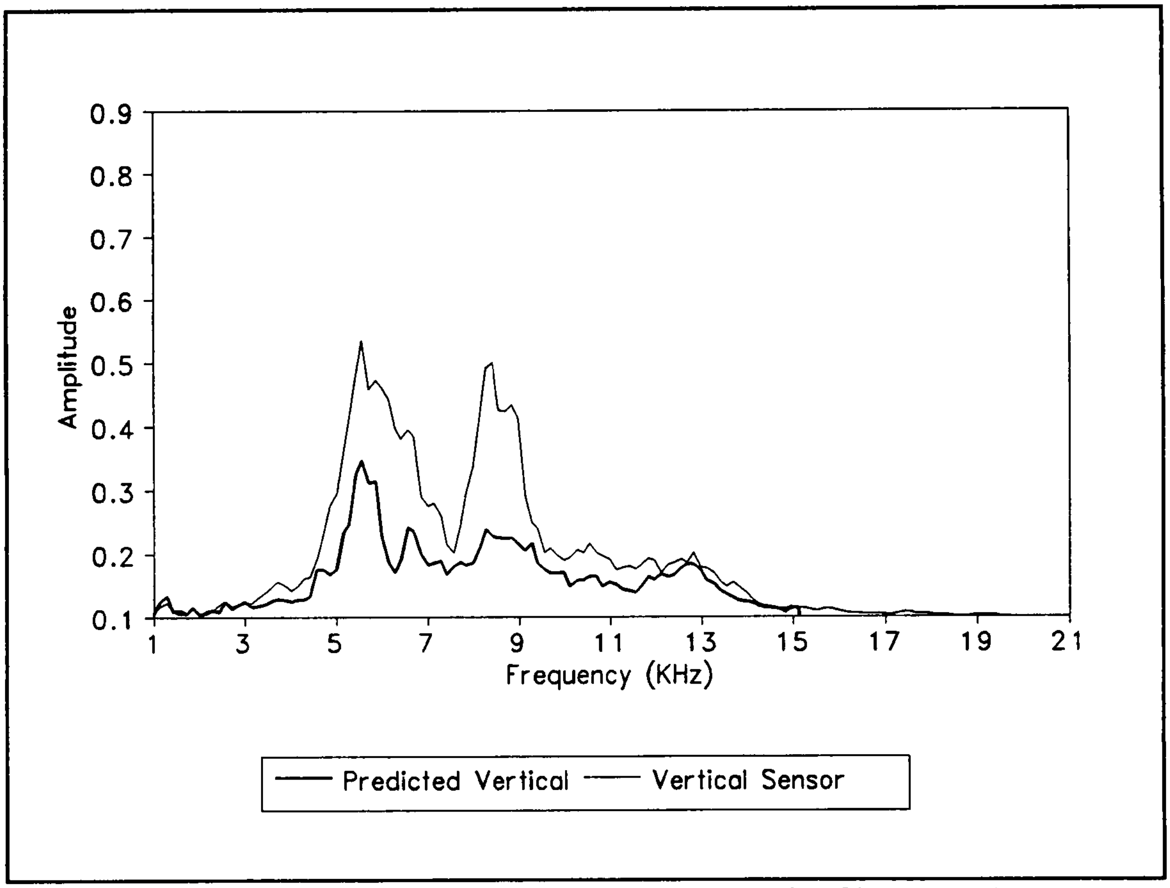

Figure I-2.6 Predicted and Actual Spectra for Signature 7
(Fault begins to develop)

machine prior to the failure but after the descriptors had commenced to increase. Figure I-2.7 depicts the contrast between the predicted value and the actual value of the sensor for Signature 12. The machine was determined to be non-functional at this point. The predicted levels of the vertical sensor (i.e. value if the machine had been working correctly) are much lower than the actual levels calculated from the sensor. Figure I-2.8 shows the plot of the per-frequency descriptor of Signature 12. Negative values indicate decrease in vibration and positive values indicate increases.

Table I-2.1 displays the horizontal, vertical and cross-sensor descriptors for the signatures in the recall set. The cross-sensor descriptor is sensitive to increases in vibration levels and the level of disagreement between the two sensors. Good examples of wiping are Signatures 10 and 11 (refer to Table I-2.1). For Signature 11, the vibration levels registered by the horizontal and vertical sensors are lower than those for a normal signature (Signature 2 for example) and well below the alarm criteria established by the utility. This is due to smoothing of the cage surface surrounding the defect. The cross-sensor descriptor for that signature decreases but remains close to the upper bound of normal limits.

Values of the cross-sensor descriptor may also be used to define alarm criteria based on the relationship between the sensors. In order to establish these criteria, it would be necessary to perform the analysis on baseline data and define an alarm threshold based on the calculated value of the cross-sensor descriptor for the baseline data.

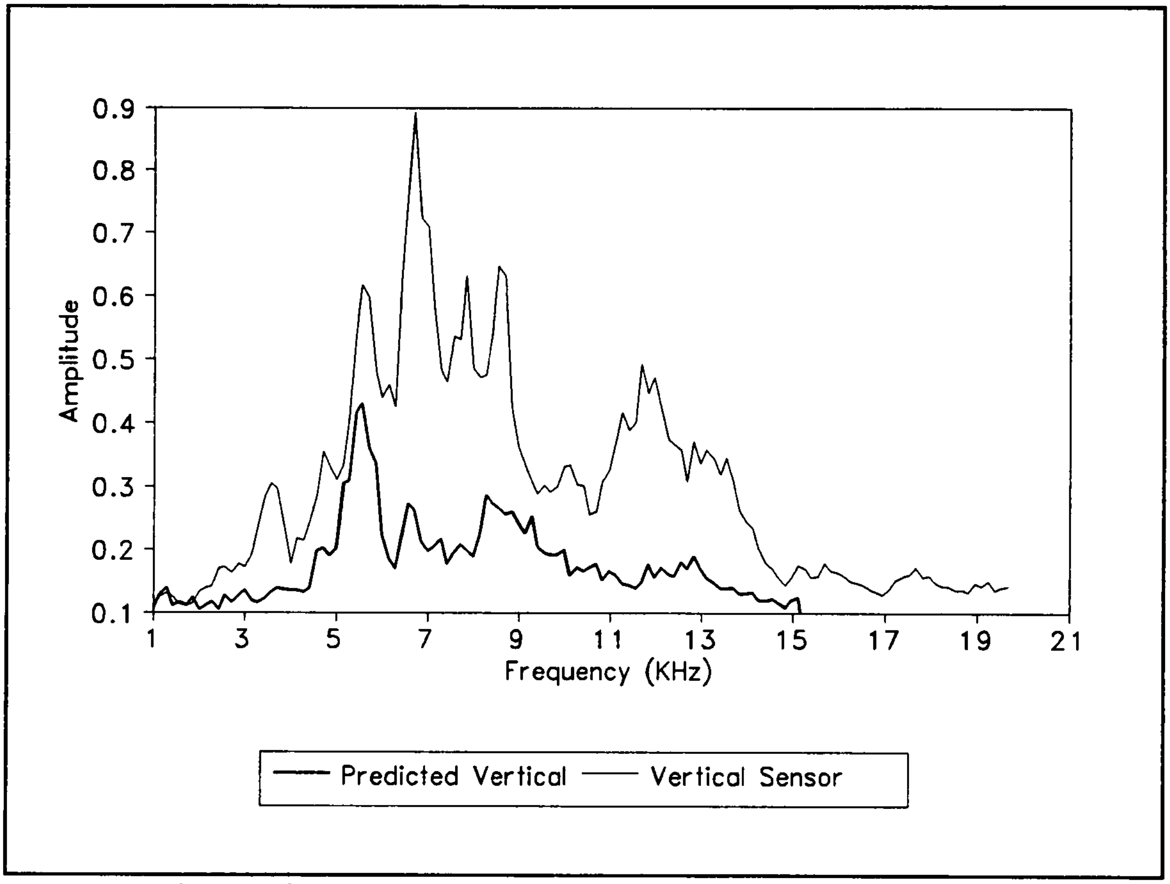

Figure I-2.7 Predicted and Actual Spectra for Signature 12 (Out-of-order bearing)

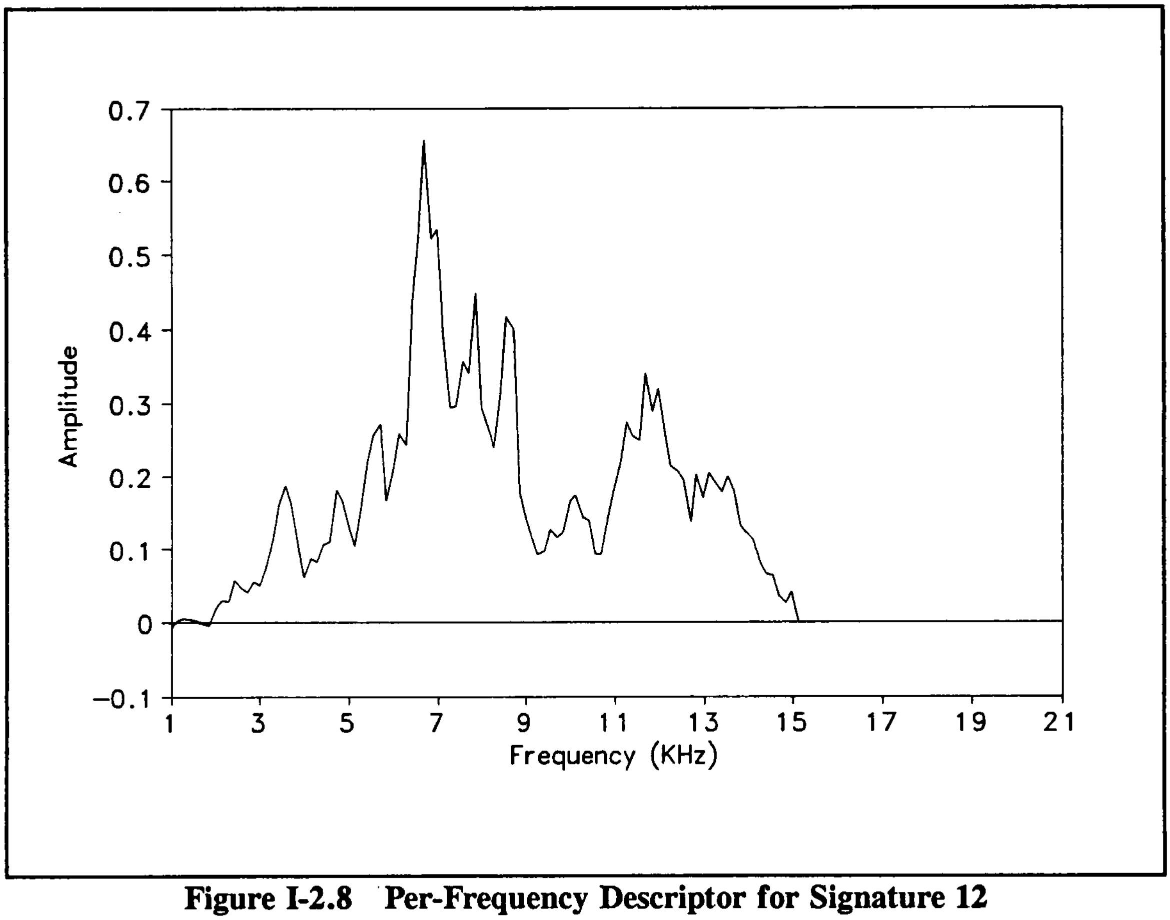

Figure I-2.8 Per-Frequency Descriptor for Signature 12

Table I-2.1. Degradation Descriptors* (High Freq.)

	Horizontal Desc.	Vertical Desc.	Cross-Sensor Desc.
Signature 2	2.89	2.94	0.1119
Signature 3	1.80	2.31	0.2683
Signature 4	1.72	1.61	0.0878
Signature 5	1.61	1.89	0.1416
Signature 6	1.74	1.73	0.0735
Signature 7	5.31	4.97	0.9457
Signature 8	7.03	5.16	0.7148
Signature 9	5.84	5.03	0.8772
Signature 10	2.42	2.53	0.2907
Signature 11	1.55	1.48	0.2254
Signature 12	10.50	8.70	2.1101

*The horizontal and vertical descriptors are calculated from velocity measurements in the two directions respectively. The cross-sensor descriptor is calculated using Equation I-2.7.

CHAPTER 3

IDENTIFICATION OF REB OPERATING STATES

INTRODUCTION

For the identification of faults in rotating machinery, time records are collected from the machinery, and detailed analysis is performed to detect features which reflect the operational state of the equipment. The analysis leads to the identification of potential failures and their causes, and makes it possible to perform efficient preventive maintenance.

This section documents a methodology for the detection of faults in rolling element bearings (REB) based on neural network technology. The spectral signatures of REB's are used to train neural network models to discriminate among the different operating states normally observed in this kind of equipment.

This technology provides an attractive complement to traditional vibration analysis because of the potential of neural networks to operate in real-time mode and to handle data which may be distorted or noisy. The significance of this work relies on the fact that REB failures are responsible for a large fraction of the malfunctions in manufacturing equipment. The technique enhances traditional vibration analysis and provides a means of automating the diagnosis of vibrating equipment.

REB VIBRATION SIGNATURES

The data used in this project was collected with accelerometers in bench tests conducted by Electricité de France. The time-dependent data was transformed using FFT techniques into low frequency (10 Hz - 1000 Hz) and high frequency (1 kHz - 20 kHz) spectra. Vibration descriptors are also calculated which describe overall vibration levels in the various components. This analysis is based on the information contained in these descriptors and the spectra.

Three global supervision descriptors are used: a high frequency descriptor calculated from vibration amplitude of effective acceleration between 1 kHz and 20 kHz, a low frequency descriptor calculated from vibration amplitude of effective velocity in the range 10 Hz to 1000 Hz, and a descriptor based on the analog computation of the crest factor after filtering. Since the descriptors give an indication of overall vibration levels in the spectra, they are used in the preprocessing of data to discriminate between components for which vibration levels are considered normal and those components which are probably faulty.

The acceleration spectra were transformed by means of numerical integration into velocity spectra. Velocity measurements are used because they can be directly compared with alarm and shutdown criteria [1,43] and also because velocity amplitudes are linked to the source of excitation. Since the features generated by severe defects in rolling element bearings usually appear at low frequency, while high frequency features are linked to incipient faults in the bearings (i.e. bearing faults in early stages of development), both low frequency velocity spectra (10 Hz to 1000 Hz) and high frequency velocity spectra (1 kHz to 20 kHz) are used in order to detect all faults.

Bearing data is comprised of vibration measurements collected from ball bearings of type 6206 during an aging simulation process. The geometrical characteristics of a 6206 rolling element are depicted in Figure I-1.1. The rolling element under test is mounted on a horizontal shaft (driven by a 10 kW motor) and loaded radially by means of a jack, imposing a vertical force on the bearing. These conditions generate scaling faults on the component.

Data is collected using an accelerometer placed in the radial direction to the loading of the bearing. The accelerometer is screwed onto a flange, and the flange is in turn stuck to the rolling element. From these measurements, spectra are generated using FFT techniques. Spectra are averaged over 16 samples with a HANNING window. The spectra have 400 points and the data set contains 71 samples representative of the behavior of bearings operating under the following conditions: no fault, localized scaling fault of inner race, generalized scaling fault of inner race, localized scaling fault of outer race, generalized scaling fault of outer race, artificial localized fault of a ball, generalized scaling fault of two balls, and generalized scaling fault of all the components.

DATA CONDITIONING

In order to process data using neural networks, the input data must be scaled to the operating range of the algorithms. For this project, the spectral values were row-wise normalized in the range (0.1, 0.9) because that range is considered operational for backpropagation. To normalize, the maximum and minimum values for each pattern are calculated and the spectra scaled to the appropriate range using Equation 24. Row-wise normalization is critical for this processing because it ensures that the shape of the spectrum is preserved, and more importantly that the location of the peaks within the spectra is not altered.

The normalized signatures contained 400 points. To reduce the amount of data, the range of frequencies was divided in intervals of three frequencies each. From each interval, the maximum amplitude was selected and the maxima used to construct a new signature with 132 points that deviates slightly (in shape) from the original spectrum. Figures I-3.1 and I-3.2 show the original and reduced signature of a normal pattern.

SYSTEM DESCRIPTION

The vibration analysis system is comprised of the three modules shown in Figure I-3.3. The first module is used to preprocess the data and is based on the information provided by the descriptors. The second module compresses the data using a trained recirculation network, and the third module performs the actual classification of the signatures using the backpropagation algorithm.

The preprocessing network uses the backpropagation algorithm in its heteroassociative mode. The network is trained to discriminate between signatures with vibration levels considered normal and signatures of possibly faulty components.

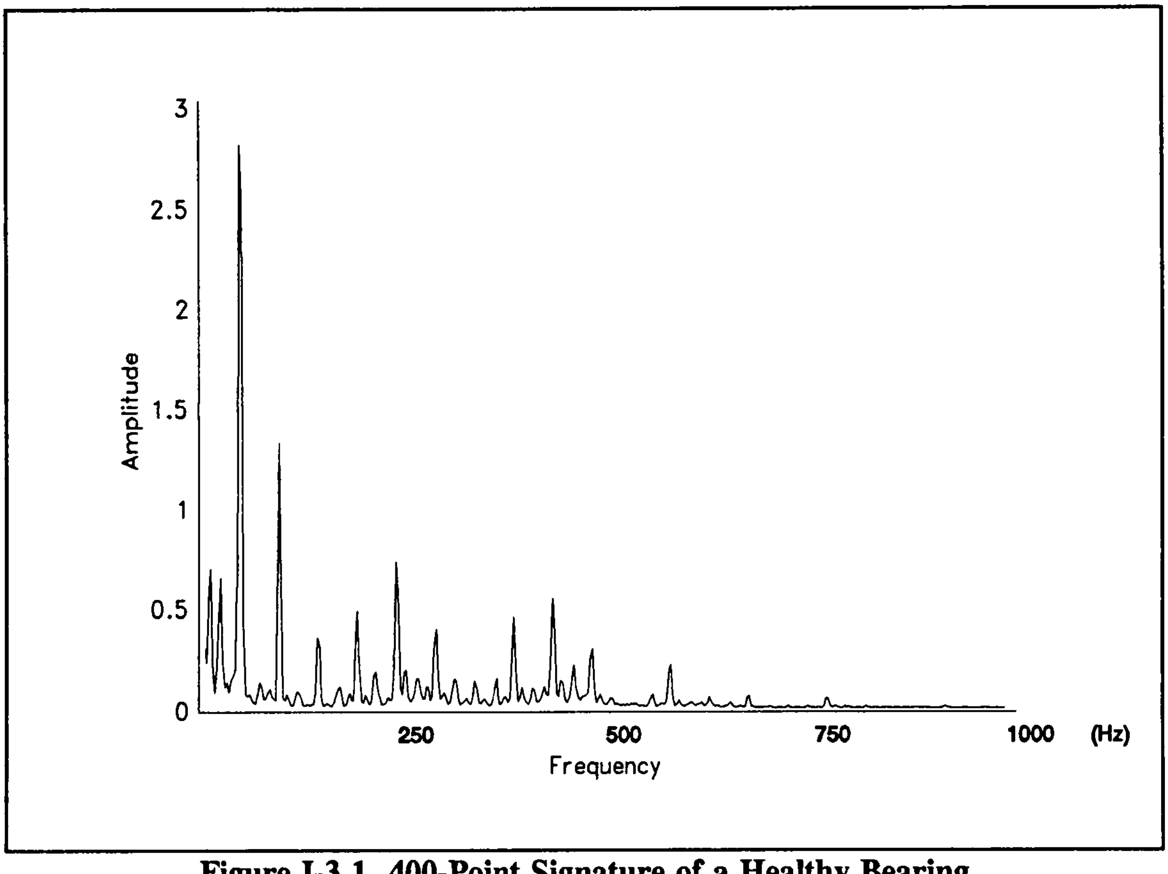

Figure I-3.1 400-Point Signature of a Healthy Bearing

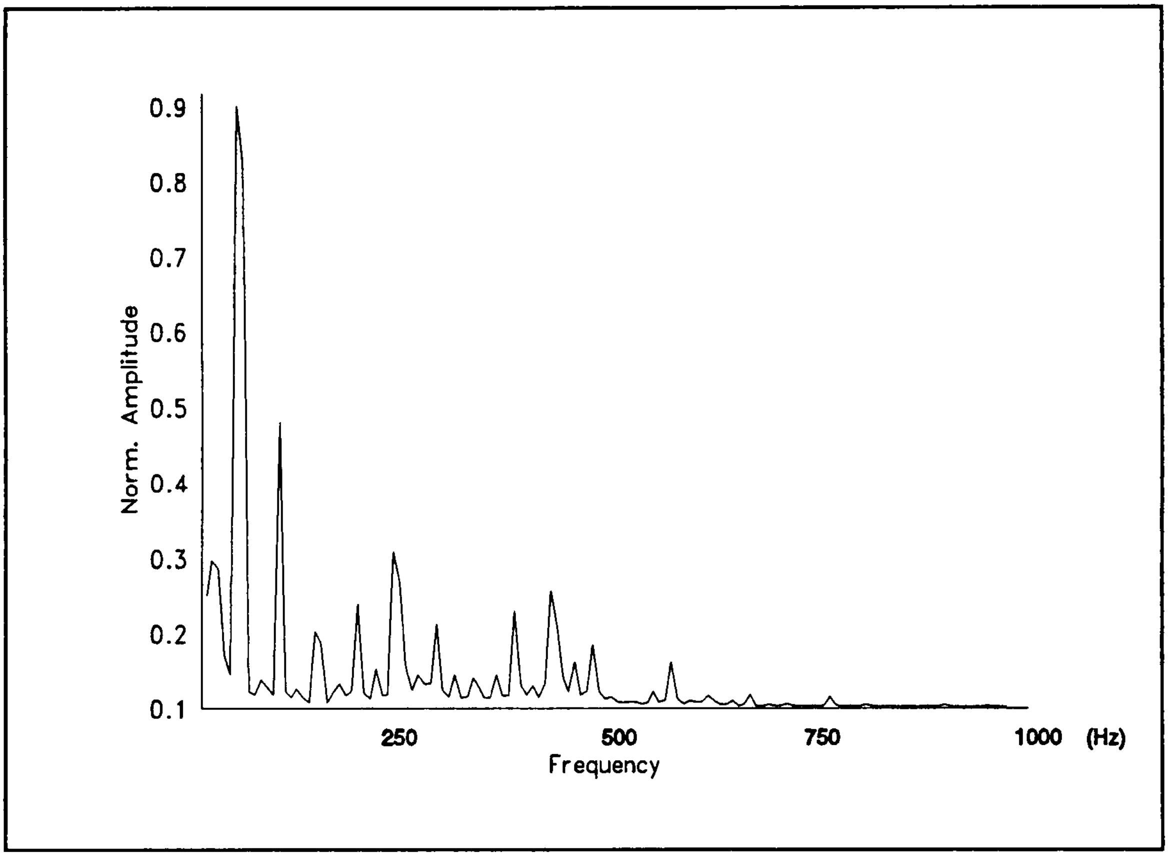

Figure I-3.2 132-Point Signature of a Healthy Bearing

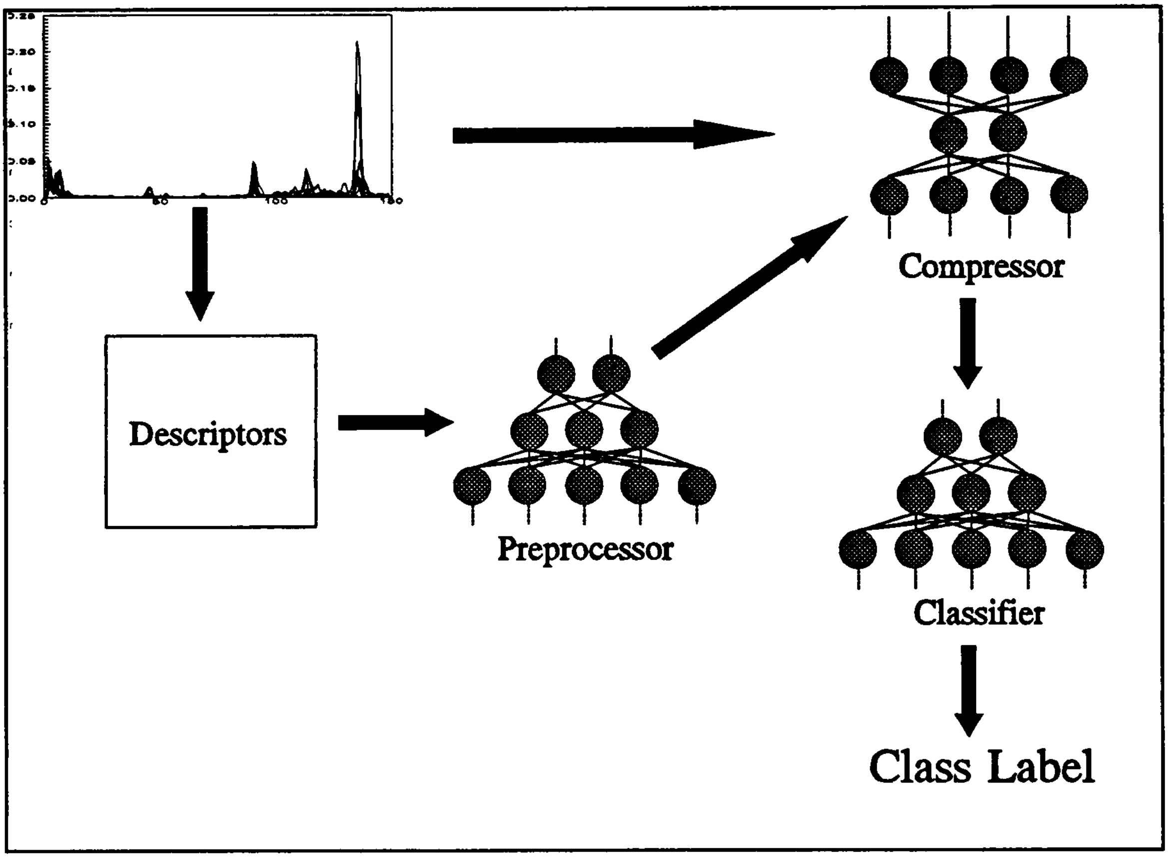

Figure I-3.3 Vibration Mode Identification System

The inputs to the network are the high frequency descriptor, the low frequency descriptor and the crest factor with additional inputs corresponding to calculated increases for these values (difference between consecutive readings for the same descriptor). The thresholds of the network have been tuned so that none of the faulty signatures can be classified as normal. Once this discrimination is accomplished, the normal patterns are not considered further in the analysis.

The compression network takes as input the points from the normalized spectra of possibly faulty signatures and performs a 3:1 compression, producing feature vectors. Two feature vectors are produced per pattern: one for the low frequency spectrum (10 to 1000 Hz.) and one for the high frequency (1 to 20 kHz). The compressed vectors are collected at the hidden layer of the network, see Figure 5. The compressed version of the pattern in Figure I-3.1 is shown in Figure I-3.4. By reducing the size of the training set during the pre-processing stage, we are able to use the compression networks in a very efficient manner.

The compression produces unique compressed representations of the signatures. In Figure I-3.4, a pattern is shown whose individual values are either close to zero or one. This kind of pattern is generated by saturating the output of each individual unit by updating the weights with large epoch sizes. By changing the epoch size, other representations may be produced. The effect of changing this parameters is not investigated further in this study.

The compressed vectors (one for the low-frequency spectrum and one for the high-frequency spectrum) each with 44 points, are the input to the classifier

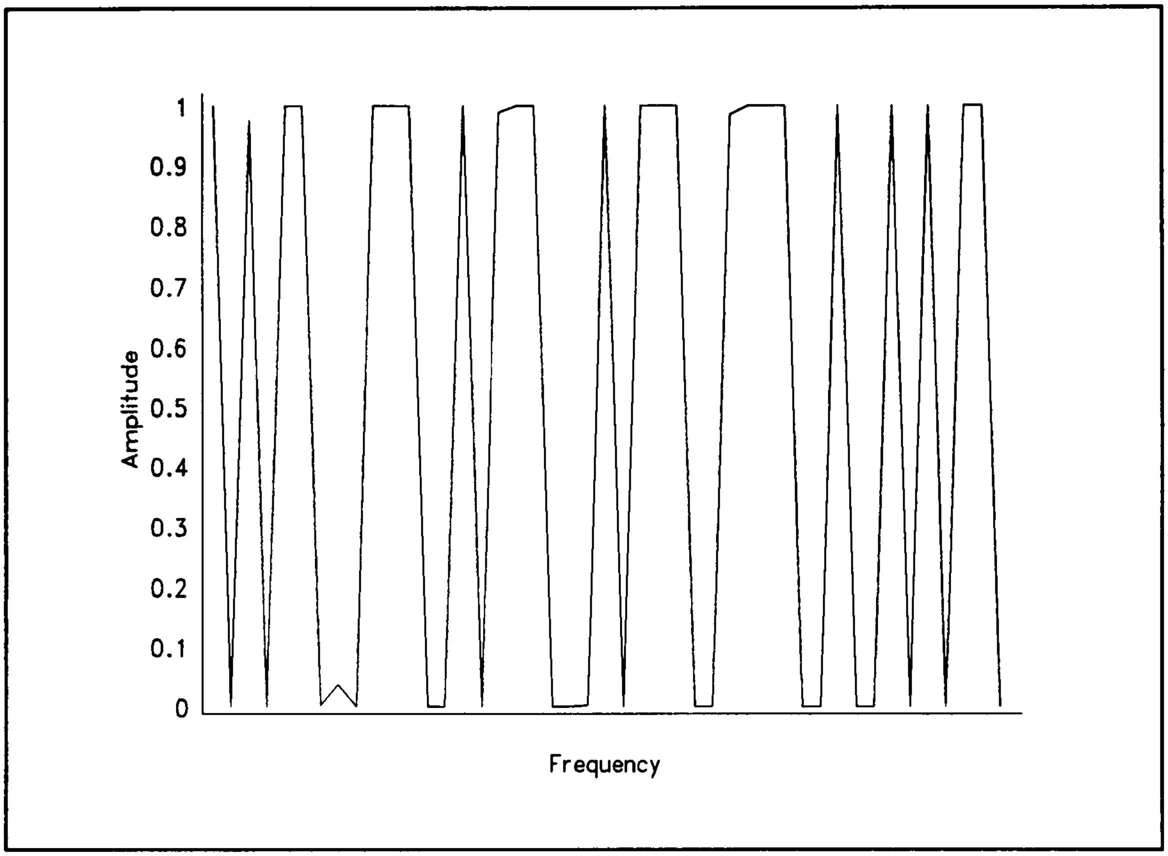

Figure I-3.4 44-Point Signature of a Healthy Bearing

backpropagation network. This network has a single hidden layer and an output layer with one neuron corresponding to each possible fault. Multiple faults are represented by the appropriate combination of active units. A high activation for a neuron in the output layer is associated with the presence of a fault. In general, an activation higher than 0.7 is considered sufficient to identify a fault.

For this project, we used the version of the recirculation network included in NeuralWorks™ which runs on a Zenith 386 SX machine. The Plexi™ implementation of backpropagation was used for classification. The Plexi™ program runs on a SUN SPARC IPC workstation.

RESULTS

The results for this system are presented separately for each one of the modules. Results are also included for a sensitivity analysis performed on the spectra to investigate the effect of Gaussian noise in the compressed signatures.

Preprocessing Module

The preprocessing network was used to identify possibly faulty components. The inputs to the network were the normalized values of the high frequency descriptor, low frequency descriptor, crest factor descriptor, and their differences between consecutive readings. The differences provide information about sudden increases in the descriptors and are essential in the detection of certain faults. The output of the network is binary, a zero indicates absence of a fault and a one indicates a possibly faulty signature. Since there were only two possible states, two nodes were used in the hidden layer and one node in the output layer.

From the total of 71 patterns, 12 were chosen for training the preprocessing network. The network was trained for 4,300 iterations until an error of 0.03 was achieved. Recalling produced the following results: 21 signatures were classified as possibly faulty and the remaining 50 as normal. Subsequent analysis showed that 4 of the normal signatures were classified by preprocessing as possibly faulty and none of the faulty signatures were classified as normal.

Compression Module

Two networks with identical topology were used for compression: one for the high frequency spectra (1 to 20 kHz.) and one for the low frequency spectra (1 to 1000 Hz). Each network had 132 nodes in the input and visible layers (each corresponding to a point in the spectrum), 44 nodes in the hidden layer and 132 nodes in the output layer. For training the network, only part of the possibly faulty signatures (as classified by the preprocessing network) was used.

During the learning phase a total of 10 signatures (48% of the possibly faulty data set) was used, and the network was trained until the error was reduced to 0.03. By reducing the size of the training set using the preprocessing network we are able to reduce the training from 50,000 to 9,000 iterations.

In order to measure the effectiveness of the compressor network for noise removal, four hundred data sets were created by adding Gaussian white noise to the original set of patterns. The noisy patterns were then compressed using the trained networks described above, and later expanded to their original form. For the experiment, different degrees of noise were added.

One hundred data sets were generated by adding 1 % noise to the original patterns. The noise is generated using a random number generator with a Gaussian distribution whose mean is the spectral component under consideration and with standard deviation equal to the mean multiplied by the percentage of noise. After the patterns were compressed and then expanded to their original forms, the difference between the original and reconstructed patterns were calculated. These differences were used to calculate the average RMS value (per spectral component) for each pattern. The deviation per pattern was found by averaging the individual deviations over all the sets.

Similar calculations were performed for 2 %, 5 % and 10 % noise. The results are included in Figure I-3.5. In the figure, the average deviation for each of the 24 patterns used in the study is presented. As can be observed, there is an almost linear relationship between the amount of noise introduced and the resulting deviation. The perturbation produced in each pattern is, however, much smaller than amount of noise added by a factor of 6.6 (1% noise produces 0.15% deviation). The results represent the collective effect of the noise for each pattern, although each spectral component in each pattern was perturbed individually. These results are consistent with earlier findings reported in the literature regarding the ability of recirculation networks to remove noise from non-parametric data [36].

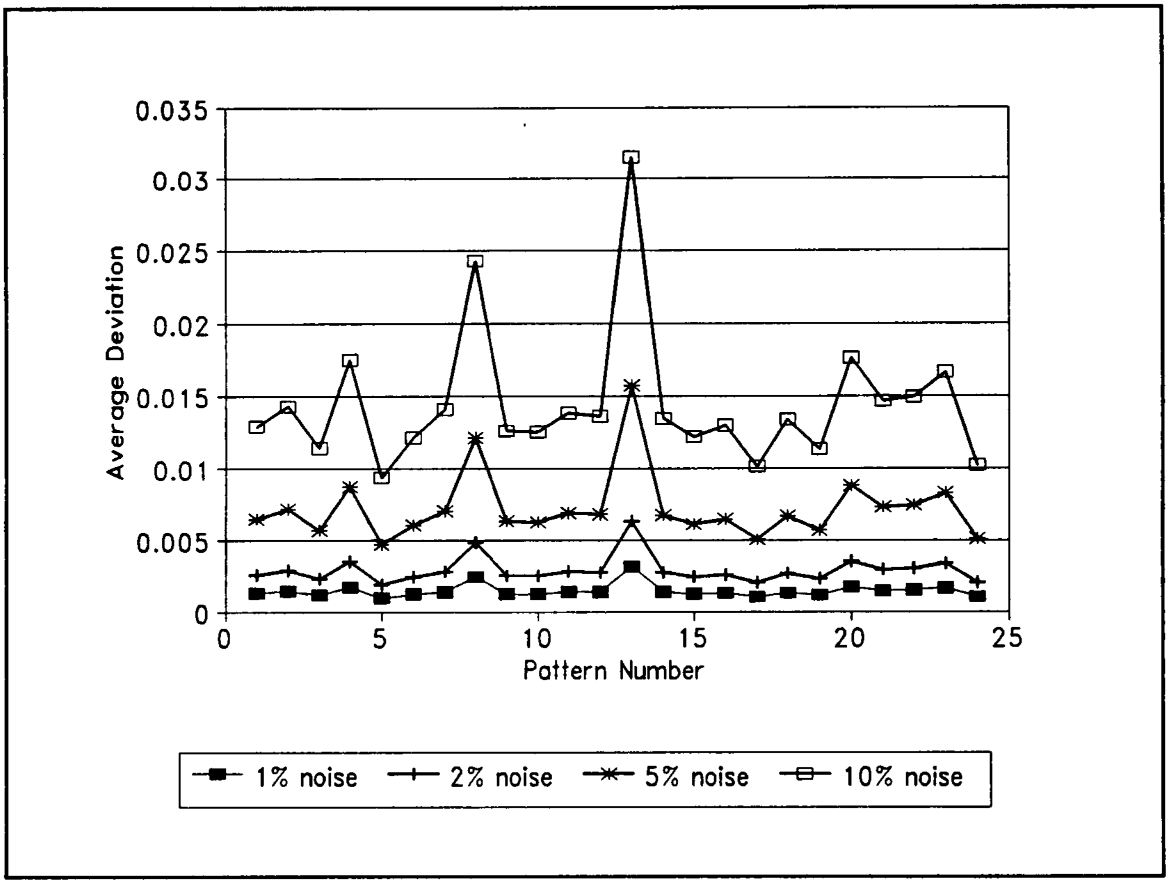

Figure I-3.5 Average Deviation with Noise Uncompressed Patterns

Classification Module

The classifier network had an input layer with 88 nodes; 44 nodes corresponding to the compressed low frequency spectrum and 44 corresponding to the high frequency spectrum. The number of nodes in the hidden layer was set to 7 according to the criterion proposed by Kung and Hwang [35]. The output layer has 6 nodes, each node corresponding to a single fault. Each fault activates a particular neuron in the output layer; for multiple faults more than one neuron is activated corresponding to the appropriate combination. The output units have a linear threshold function of 0.7. Any output higher than 0.7 is reported as a fault, while lower outputs are considered as absence of faults.

To train the network, 8 patterns - each representing a different operating state - were selected from the total set of 21 (38%). Training was achieved after 2,300 iterations (the error was reduced to 0.03.) To recall, the entire data set was used. Of the 21 patterns, 19 (90%) were correctly classified in the appropriate categories, the two misclassified patterns were faulty and were classified as normal. The system (preprocessing, compression and classification modules together) was able to classify correctly 69 of the 71 signatures or over 97% of the data.

Other Results

The second data set consists of vibration data from a steel sheet manufacturing mill. The data were collected from 246 "laminar flow" table rollers in the facility, 39 of which had been identified as faulty by expert personnel. Among the faulty signatures, 26 reflected single fault in the rollers and the other 13 were collected from rollers with more than one fault. Data were collected with accelerometers at 9 different locations on each machine and spectra generated using FFT techniques. Each spectrum contains 150 points in the range 1 Hz to 500 Hz and only one data set per machine was used to perform the diagnostics.

The readings collected by the nine sensors on each roller are correlated but not identical in amplitude (Figure I-3.6). There are three reasons for the difference: first, the vibration levels are different throughout the machine; second, the effect of vibration produced by neighboring machinery have different levels at different measuring points; and finally, faults are particular to a bearing located near a sensor are not always recorded by other sensors. To compensate for these differences we produce a single signature per machine by removing the baseline signal (minimum amplitude in each signature) and averaging the spectra for all sensors. The range of the spectral values was (0.0, 1.0), the amplitudes were not normalized or enhanced by any method.

The general problem areas are misalignment, improper lubrication, looseness, wear in motor brushes, damage to inboard bearing and damage to outboard bearing. Using the low frequency spectra allows for the identification of faults which were fully developed on the equipment; the aim is to develop a system based on neural networks able to identify faults. Since several of these faults produce similar features in the spectra [53], the technique must be able to focus on relevant information in the spectrum in order to discriminate among faults.

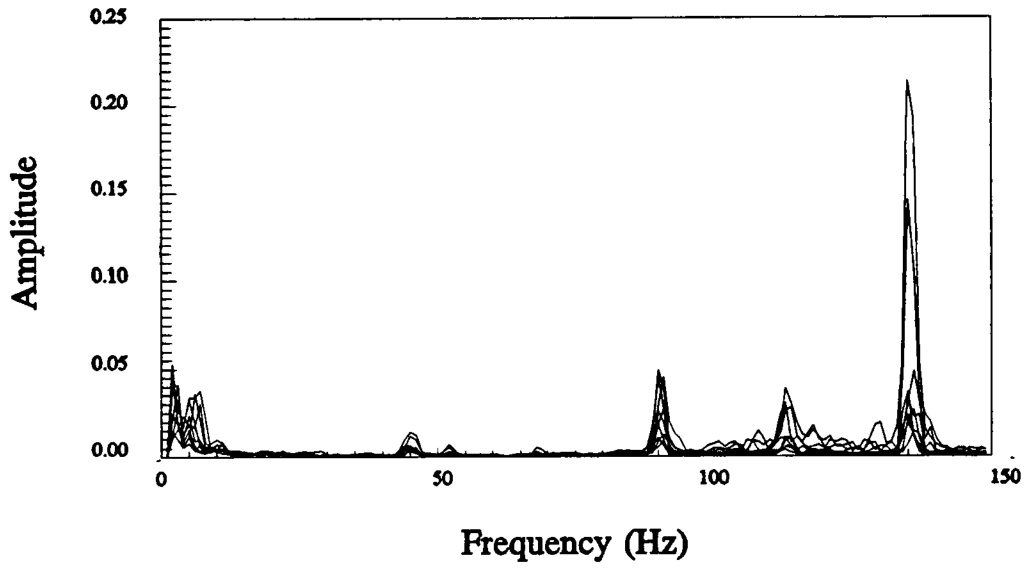

Figure I-3.6 Spectra of Nine Sensors on the Same Machine

The neural networks technique described previously was applied to the roller data. The recirculation network was used to compress the 150-point spectra to 50 points. The network was trained for 12,000 epochs until the error was reduced to 0.02. The output of the network was a set of compressed signatures which were used to train the classifier backpropagation network. The backpropagation network was trained on single-faults signatures for 12,000 epochs using PLEXI(TM). The network had a 50-neuron input layer, 26-neuron hidden layer and an output layer with one neuron corresponding to each identifiable fault. For multiple faults, more than one neuron is activated corresponding to the appropriate combination of faults.

Each of the nodes of the output layer had a linear threshold function centered around 0.5. The function reports a "1" if the output of the node is greater than 0.5 and reports a "0" if the output of the nodes is less than or equal to 0.5. The classifier network was trained with 18 of the single fault patterns. The remaining patterns were used exclusively for testing. The data from rollers with multiple faults (13 patterns) were not used for training, and became special cases for recall.

Recalling the set of single-fault patterns was 100 % accurate. The network was able to identify each fault correctly even when close examination of the original spectra showed only small differences in their features. For the set of double and triple fault patterns, the network was able to identify the most prominent fault. In an effort to improve performance when dealing with multiple faults, a subset of these was used for training; the network is able to recognize signatures on which it was trained, but continued to identify only most prominent faults in the rest of the set.

CHAPTER 4

CLUSTERING OF REB VIBRATION DATA

INTRODUCTION

Clustering is a discrimination technique for measuring the inherent geometrical and statistical similarity of patterns, and for grouping patterns according to this measurement. A number of clustering techniques have been developed to solve problems in many areas; in Reference [60] for example, a variety of these algorithms is described. For vibration analysis, clustering is a very useful technique because in many cases the exact operating state of a machine cannot be assessed, and clustering may unveil classes of operating states that would not be discovered otherwise.

Clustering is unsupervised because the discrimination system receives no information regarding the number of classes represented in the data set, or "hints" (in the form of parameters) about the behavior of individual classes. The grouping of patterns into classes is based purely on the inherent similarities among patterns.

To illustrate this concept, let us examine a basic clustering scheme. Suppose we have a set of N samples $\{x_1, x_2, \ldots x_n\}$. The cluster center, or centroid, z_i is the vector that defines class i, and which is used for distance measurements related to that class. If we arbitrarily designate the first sample x_1 as the first centroid z_1, the clustering procedure may be described by the pseudocode presented in Table I-4.1, where T is an acceptance criterium.

Table I-4.1 Pseudocode for Simple Clustering Scheme

```
BEGIN
FOR each centroid z_i
        FOR each sample x_j
                Calculate d_ij                          % distance between x_j and z_i
                IF d_ij > T
                        Allocate a new centroid         % Create a new cluster
                ELSE
                        IF d_ij ≤ d_kj                  % Distance to cluster i is less than
                                Assign x_j to class i   % distance to all other clusters k
                ENDIF
        ENDFOR
ENDFOR
IF new clusters were formed
        BEGIN
ELSE
        END.
```

This very simple algorithm illustrates in very general terms the issues associated with clustering. The first issue is how to choose a meaningful value for T. This value is critical for the analysis because it determines the number of clusters, and using the wrong number imposes an artificial structure on the data instead of revealing its true characteristics. Another issue is the choice of distance measure, it also influences directly the accuracy of the discrimination. Finally, the data itself imposes constraints because its inherent geometric or statistical properties are not known. Most procedures for clustering are affected by the initial cluster choices, the order of presentation of the data, and inherent geometrical properties.

Most clustering algorithms are very stable, meaning that they behave consistently when presented with the same data. However, these design choices are an issue for all unsupervised pattern recognition techniques [60], which generally handle high-dimensional data that cannot be inspected visually. The algorithm described here, is a variation of the Fuzzy C algorithm, modified to the realm of vibration data. A new distance measurement is introduced which uses domain-specific information to assess the similarity of patterns. For distance calculations, a matrix of weights is used to assign factors of importance to frequencies in the spectrum which are known to be related to particular defects.

FUZZY C-MEANS ALGORITHM

The Fuzzy C-Means algorithm [61,62] is an extension of the crisp K-means algorithm [60]. The objective of the crisp algorithm is to partition patterns into classes, such that, all patterns having "natural relationship" fall in the same partition, and unrelated samples fall into separate partitions. Each sample is assigned to one class, so that membership values are either one or zero. In the case of the Fuzzy C-Means algorithm the membership values may range from zero to one.

The Fuzzy C-Means algorithm attempts to cluster measurement vectors by searching for local minima of the Generalized Within Group Sum of Squared Errors function (WGSSE). This function is given by Equation I-4.1:

$$J_m(U,v) = \sum_{k=1}^{n} \sum_{i=1}^{c} (u_{ik})^m |x_k - v_i|_A^2 \qquad \textbf{(I-4.1)}$$

$$1 < m < \infty$$

where c is the number of clusters, n is the number of vectors, x_k is a k[th] measurement vector, v_i is the i^{th} centroid vector, m is the fuzzy coefficient, $\| . \|_A$ is an inner product norm, $\| Q \|_A^2 = Q^T A Q$, and A is a $d \times d$ positive definite matrix, where d is the dimension of the pattern vectors.

When $m = 1$ the objective function J_m is the classical WGSSE function, and the algorithm is the crisp K-Means algorithm. For $m > 1$ under the assumption that $x_k \neq v_i$, (U,v) may be a local minimum of J_m only if:

$$(\text{I-4.2}) \quad u_{ik} = \left(\sum_{j=1}^{c} \left(\frac{\| x_k - v_i \|_A}{\| x_k - v_j \|_A} \right)^{\frac{2}{(m-1)}} \right)^{-1}, \quad \forall\, i,k.$$

and

$$v_i = \frac{\sum_{k=1}^{n} (u_{ik})^m x_k)}{\sum_{k=1}^{n} (u_{ik})^m}, \qquad \forall\, i. \qquad \text{(I-4.3)}$$

$\| Q \|_A^2 = Q^T A Q$. If A is the identity matrix and the norm $\| . \|$ is the Euclidean norm, we obtain the Euclidean distance. m is $1 \leq m < \infty$, the practical range is $1 \leq m \leq 5$ [61].

In the Fuzzy C-Means algorithm, the performance threshold ϵ is defined in terms of the function J of Equation I-4.1. The algorithm is comprised of the following steps [61]:

1. Fix the number of clusters c. Select the inner product norm. Fix the fuzzy coefficient m. Set $p = 1$. Initialize $U^{(0)}$
2. Calculate fuzzy cluster centers $\{v^{(p)}\}$ using $U^{(p-1)}$ and the condition specified in Equation **I-4.3.**
3. Update $U^{(p)}$ using $v_i^{(p)}$ and the Equation **I-4.2.**
4. If $\| U^{(p)} - U^{(p-1)} \|_A < \epsilon$ terminate; else set $p = p+1$ and go back to Step 2

SIMILARITY MEASURE FOR REB VIBRATION ANALYSIS

In order to apply any classification or clustering algorithm to the REB vibration analysis problem, it is necessary to take into account the contribution of critical frequencies in the spectra. In our case, we define a matrix **A** of weights whose entries in the main diagonal represent the relative importance of each frequency for the detection of particular defects. The factor of importance (weight) is a number greater than or equal to 1 such that those frequencies which we know provide critical information are assigned higher numbers than other members of the diagonal.

In this matrix, off-diagonal elements are set to zero, and formulae are used to identify critical frequencies at which peaks are expected if particular faults are present. The formulae relate parameters in the geometry of the bearing (such as number of balls, diameters, etc.) and give an indication of where in the spectra to look for features related to a given fault (critical frequency). For example, the formula given in Equation I-1.1 is used to identify the frequency at which defects related to the inner race would appear. The features also appear at multiples of the critical frequency.

The frequencies surrounding the critical frequency are also assigned weights in order to consider the information in the sidebands. Sidebands are small-amplitude peaks that appear in the neighborhood of a critical frequency in the spectrum and which are used in the analysis to discriminate between generalized and localized faults. In our case, we set the weights of neighboring frequencies to numbers greater than zero and less than the weight of a critical frequency (or its multiples). The weight matrix has the following form:

$$A = \begin{Bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & a_1 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & a_2 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_3 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & a_2 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & a_1 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & a_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & a_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & a_3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & a_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & a_1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{Bmatrix} \tag{I-4.5}$$

where the numbers on the diagonal satisfy the inequality $a_1 < a_2 < a_3$. The frequencies whose importance we set to a_3 are the critical frequencies associated with each fault (for a given type of bearing) and multiples of these frequencies. The importance factor (weight) for the neighboring two frequencies were set to $a2$, and the next two neighboring frequencies to $a1$ in order to take the sidebands into consideration. This has the effect of an amplifying triangular window centered in the appropriate region of the spectrum (the critical frequencies) to emphasize the information known to be important.

When calculating centroids, the matrix "pulls" the centroids in the "right" direction to stress particular features of the data without ignoring the remaining information. Ignoring information is not recommended because there are features which may appear in regions of the spectrum other than those given by the formulae, and which should not be overlooked for the analysis. The correct assignment of weights to the matrix depends on the problem to be solved. Testing indicates that the magnitude of the weights on the diagonal affects the effectiveness of the algorithm and that weights assigned to the off-diagonal elements of the matrix do not improve results significantly in the context of vibration.

CLUSTERING OF TWO-DIMENSIONAL DATA

In order to illustrate how the matrix A influences the results, we will describe the results for two simple test sets. The first data set consists of 11 points shown in Figure I-4.1. We are using the Fuzzy C-Means algorithm to classify them into four classes. The membership values for these points are shown in Table I-4.2. For points 1, 2, 3, 4, 6, 10 and 11 one of the membership values (m.v) is much greater than the others; these points belong to only one class, the remaining points have high membership values in more than one class. Class assignments were made based on the criterion that a pattern was a member of a class if its membership value for that class was greater than 0.1. Centroids and class assignments are shown in Figure I-4.1.

The second data set consists of 3 points whose spatial distribution is shown in Figure I-4.2. The Euclidean distance between points 1 and 3 is smaller than between points 2 and 3. Setting the matrix A equal to the identity matrix causes the algorithm to classify point 3 more in class 1 (m. v: 0.993787) than in class 2 (m. v: 0.006213). When we set the importance of the x coordinate to be 10 times that of the y coordinate the difference in membership values is even larger (0.995293 for first class and only 0.004707 for the second class). However when the factor of importance of

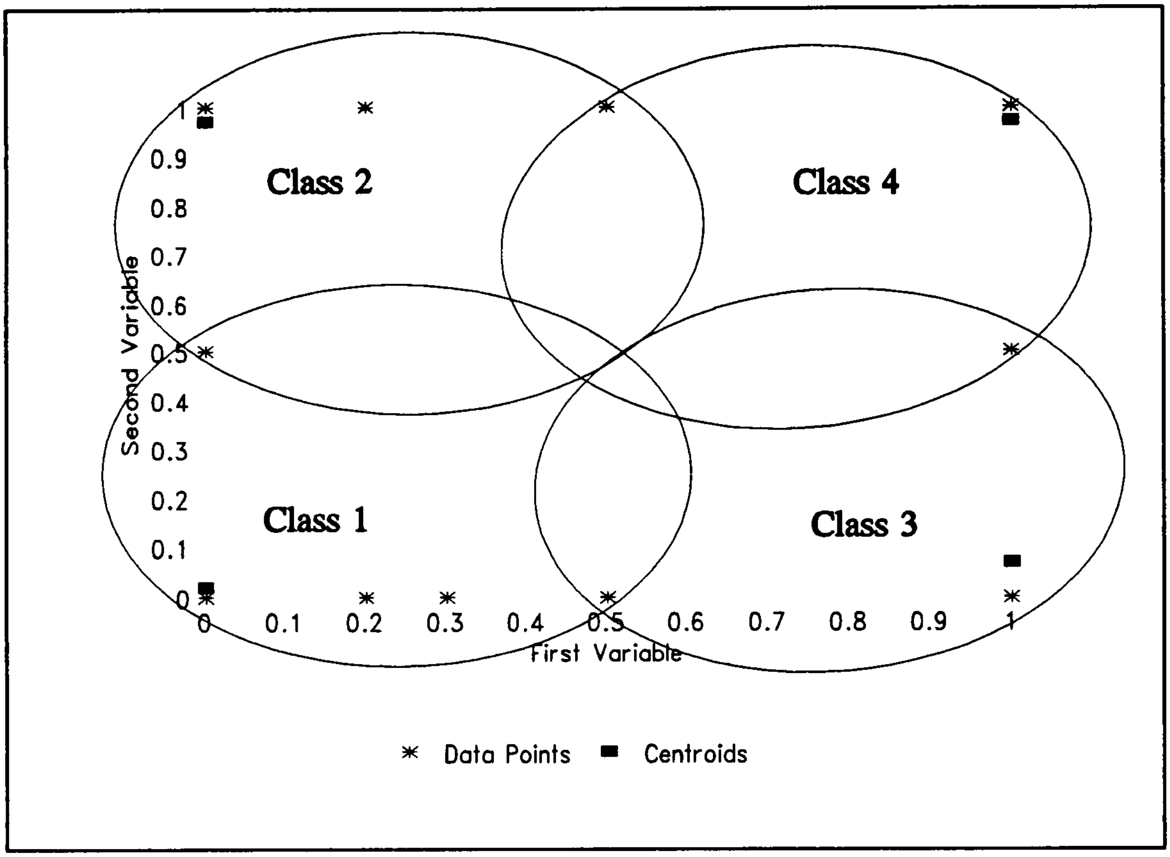

Figure I-4.1 Centroids and Class Assignments

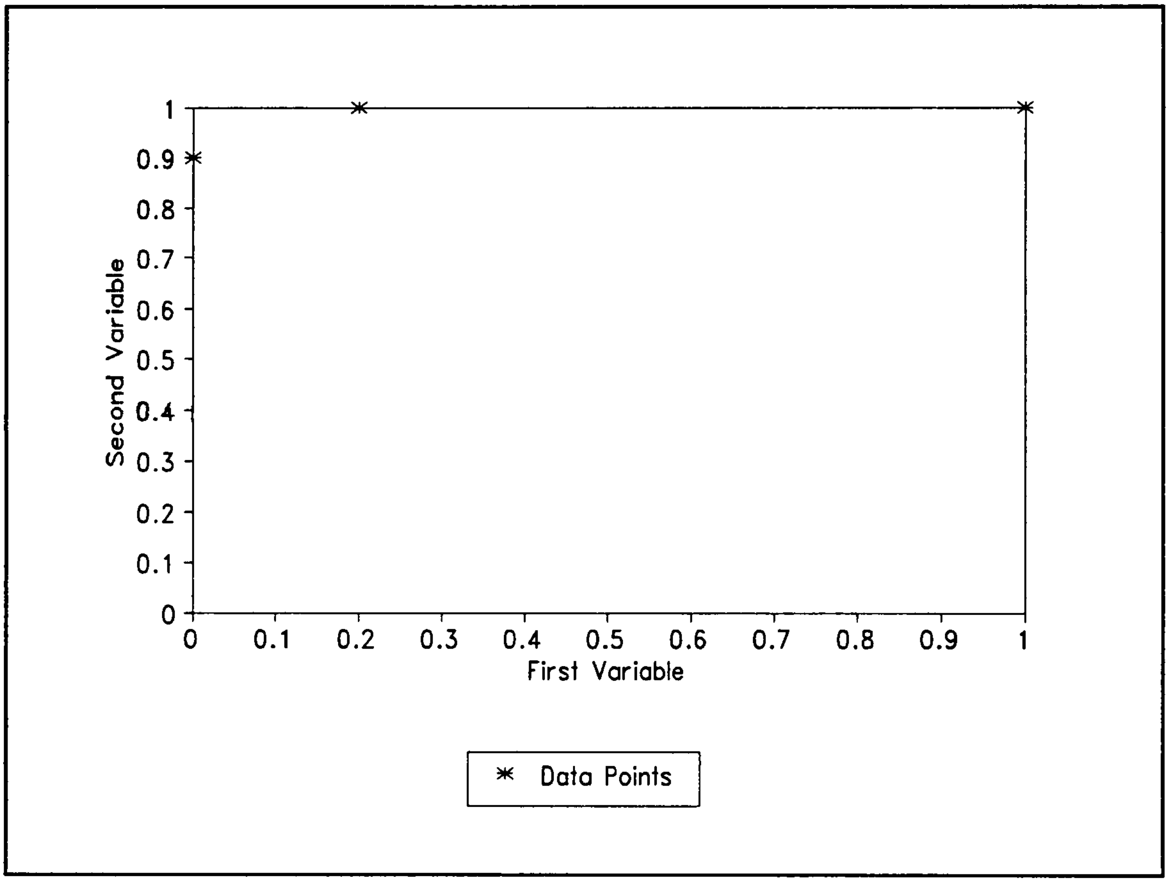

Figure I-4.2 Spatial Distribution of Points

the y coordinate is 10 times larger than that of x coordinate, the same point is classified as belonging to class 1 (m.v: 0.976534) and to class 2 (m.v: 0.023466). Setting the factor of importance of the y coordinate 100 times larger than that of the x coordinate, changes the results dramatically: the point is classified as belonging to class 1 (m.v: 0.10) and to class 2 (m.v: 0.89958). These results show that, when dealing with the vibration data, setting the relative importance of certain frequencies will make the algorithm weight these frequencies more heavily.

Table I-4.2. Membership Values for Points in First Data Set

Point	M. value Class 1	M. value Class 2	M. value Class 3	M. value Class
1 (0.0, 0.0)	<u>0.954158</u>	0.019241	0.018446	0.008155
2 (0.0, 1.0)	0.007252	<u>0.981671</u>	0.003297	0.007780
3 (1.0, 0.0)	0.001110	0.000308	<u>0.997902</u>	0.000680
4 (1.0, 1.0)	0.000303	0.000830	0.000640	<u>0.998227</u>
5 (0.0, 0.5)	<u>0.391500</u>	<u>0.483609</u>	0.061770	0.063121
6 (0.2, 1.0)	0.000790	<u>0.997331</u>	0.000444	0.001435
7 (0.5, 0.0)	<u>0.763930</u>	0.027785	0.182728	0.025556
8 (1.0, 0.5)	0.067077	0.058670	<u>0.443287</u>	<u>0.430966</u>
9 (0.5, 1.0)	0.040890	<u>0.636997</u>	0.037190	<u>0.284923</u>
10 (0.3, 0.0)	<u>0.992469</u>	0.001857	0.004494	0.001179
11 (0.2, 0.0)	<u>0.999850</u>	0.000047	0.000078	0.000025

BEARING VIBRATION SIGNATURES

The data set described in Section I-3.2 was used for testing the clustering system. The low frequency spectra (10 Hz. to 1000 Hz.) have 400 points and the data set contains 71 samples representatives of bearings operating under the following conditions: no fault, localized scaling fault of inner race, generalized scaling fault of inner race, localized scaling fault of outer race, generalized scaling fault of outer race, artificial localized fault of a ball, generalized scaling fault of two balls, and generalized scaling fault of all the components. For ease of description we refer to the above operating conditions as ***N, A, B, C, D, E, F*** and ***B+D+F***, respectively.

Seven patterns in the set are a transition from fault ***C*** to ***B+D+F***. For these seven patterns, the exact membership is unknown. The range of frequencies was divided in intervals of three frequencies each. From each interval, the maximum amplitude was selected, and the maxima used to construct new signatures with 132 points. Figures I-3.1 and I-3.2 show the original and reduced signature of a normal pattern. The distortion introduced is small since the frequency bandwidth of all the peaks is greater than 7.5 Hz.

CLUSTERING OF REB VIBRATION DATA

To test the clustering system, we used a set of 32 signatures from the data set described in Section 4.5. The best results were obtained by setting the number of clusters to the number of known classes plus one (6). To construct the weight matrix, 36 frequencies corresponding to regions where the faults have characteristic peaks were selected. The importance factors in matrix *A* were set to 20, 25 and 30 for *a1*, *a2* and *a3* respectively, and the fuzzy coefficient *m* was set to 2. Through experimentation it was found that this value reflected the proper amount of fuzziness, with larger values of the coefficient the algorithm tends to distribute the membership functions for patterns over the entire range of classes.

The algorithm clustered correctly all the normal patterns. All but two of these patterns have a degree of membership of at least 0.97 (Table I-4.3). The remaining two patterns (14 and 19) were examples of "false alarms"; the patterns were normal, but the vibration levels were higher than for other normal patterns. Generally, alarm criteria are defined in terms of descriptor values. These descriptors are calculated from effective velocity measurements and give an indication of overall vibration in the machine. For the two patterns in our data set (14 and 19) the descriptor levels were so high that traditional descriptor-based classifiers would have produced an incorrect label for the signatures. Our system is able to recognize them as normal although the membership values are lower than those of other normal patterns. Pattern 14 is normal (m.v: 0.57) and has fault D (m.v: 0.42). Pattern 19 is classified as normal(m.v: 0.75) and as having fault D (m.v: 0.24).

The seven patterns with unknown labels (26 to 32) are classified mostly as examples of fault *C*, with some degree of membership in other classes. Especially patterns 31 and 32 show a substantial amount of fault *A*. Patterns with faults *A*, *B*, *C* and *D* were classified correctly in different classes. However the two patterns with fault C (1 and 5) were never classified in the same cluster; a new class was added to accommodate the singleton pattern 5. Neither changing the values of ***a1***, ***a2*** and ***a3*** in matrix ***A***, nor the number of clusters gave the correct classification for these two patterns; examination of the patterns showed that these patterns represent the same fault at different stages of evolution and that the signatures are very different.

When augmenting the data set by including faulty patterns with new faults (such as ***E***, ***F***, ***B+D+F***) the algorithm divided the set of normal patterns into two different classes. It also classified signatures with faults ***B+D+F*** as a new class, and therefore the membership values no longer reflected the degree of each fault. In general, the performance of the algorithm deteriorates as the number of classes increases. Performance is affected by uneven distribution of patterns in the data set (many signatures in one cluster and singleton clusters).

For comparison purposes, the bearing data was clustered using the crisp K-means algorithm and the Fuzzy C-Means algorithm (without the weight matrix) and the results were completely incorrect. It was also found that the performance of the weighted Fuzzy C-Means algorithm deteriorates as the number of faults considered increases. This is due to a corresponding increase in the number of amplifying windows, and specifically to their overlap. The magnitude of the weights and the size of the neighborhood influences the effectiveness of the clustering system; in general, it was possible to obtain membership functions reflecting the real relationship among patterns. The most important feature of the system is that it considers all the information in the spectrum but it provides a mechanism for weighting the information according to its relevance from a theoretical standpoint.

Table I-4.3. Results of Clustering (m = 2)

M. value Pattern (Fault)	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6
1 (C)	0.884	0.043	0.026	0.0	0.001	0.042
2 (N)	0.0	0.998	0.0	0.0	0.0	0.001
3 (A)	0.0	0.0	0.999	0.0	0.0	0.0
4 (B)	0.0	0.0	0.0	1.0	0.0	0.0
5 (C)	0.0	0.0	0.0	0.0	1.0	0.0
6 (D)	0.0	0.0	0.0	0.0	0.0	0.996
7 (N)	0.0	0.988	0.0	0.0	0.0	0.011
8 (N)	0.0	0.982	0.0	0.0	0.0	0.017
9 (N)	0.0	0.883	0.0	0.0	0.0	0.115
10 (N)	0.0	0.906	0.0	0.0	0.0	0.093
11 (N)	0.0	0.969	0.0	0.0	0.0	0.030
12 (N)	0.0	0.977	0.0	0.0	0.0	0.022
13 (N)	0.0	0.980	0.0	0.0	0.0	0.019
14 (N)	0.0	0.576	0.0	0.0	0.0	0.418
15 (N)	0.0	0.979	0.0	0.0	0.0	0.020
16 (N)	0.0	0.979	0.0	0.0	0.0	0.020
17 (N)	0.0	0.961	0.0	0.0	0.0	0.037
18 (N)	0.0	0.977	0.0	0.0	0.0	0.022
19 (N)	0.0	0.752	0.0	0.0	0.0	0.242
20 (N)	0.0	0.983	0.0	0.0	0.0	0.016
21 (N)	0.0	0.992	0.0	0.0	0.0	0.007
22 (N)	0.0	0.939	0.0	0.0	0.0	0.059
23 (N)	0.0	0.971	0.0	0.0	0.0	0.027
24 (N)	0.0	0.975	0.0	0.0	0.0	0.024
25 (N)	0.0	0.948	0.0	0.0	0.0	0.051
26 (undet)	0.935	0.023	0.016	0.0	0.0	0.022
27 (undet)	0.875	0.054	0.016	0.001	0.01	0.050
28 (undet)	0.903	0.039	0.017	0.001	0.001	0.036
29 (undet)	0.972	0.008	0.008	0.001	0.001	0.008
30 (undet)	0.959	0.009	0.015	0.002	0.002	0.010
31 (undet)	0.659	0.048	0.199	0.019	0.022	0.050
32 (undet)	0.794	0.031	0.107	0.011	0.021	0.032

CHAPTER 5

DISCUSSION

A methodology is presented for the analysis of vibration data based on neural networks. The advantage of developing a system based on this technique is the possibility of automating the monitoring and diagnostic process for vibrating components and building diagnostic and fault-detection systems which complement traditional power spectral density analysis.

For the monitoring of machinery vibration, a technique is described based on multiple sensor modeling. A neural network has been trained to mimic the relationship among sensors in a component, and deviations from normal behavior can be detected easily by monitoring the predictions of the network and the actual values of the sensors. In addition, two descriptors are defined: the cross-sensor descriptor which measures the overall difference between the actual vibration levels and the predicted levels, and the per-frequency descriptor which gives a measure of change for every frequency in the spectrum. The methodology has been tested on data from a motor in one of EDF's power stations.

Traditionally, diagnosis of components based on vibration analysis has been made by looking at specific regions of the spectra. However, vibration in general is imprecise because in some cases the behavior of components under certain operating states is not known in detail; especially when the behavior is induced by more than one fault. Compressed signatures generated with recirculation networks are efficient and feasible representations of spectra which provide complete information, reduce noise, and spread the information contained in the peaks over the entire spectrum.

The feasibility of using recirculation networks for compression of spectral data has been established. Under this study we have achieved compression ratios of 3/1 while maintaining the statistical properties of the original patterns. The recirculation network provides an excellent mechanism for the compression because in addition to reducing the dimension of the data patterns, it performs automatic feature extraction by preserving in the reduced set only the features that are relevant and consistent within the data.

The fault identification system uses both low and high frequency spectra, making it possible to detect incipient faults as well as severe defects. In addition, information provided by descriptors of overall vibration are used to separate normal and possibly-faulty signatures. This information permitted the reduction of training time (both for compression and classification) by a factor of 10.

To complement the capabilities of the classification system, a clustering system was developed based on the technology of fuzzy clustering. The technique provides an excellent mechanism for the identification of the operating state of rotating machinery because it permits the easy incorporation of knowledge about the problem, and can be trained, as done for the backpropagation-based classification system, from actual system data. The advantage of including a clustering capability to a vibration analysis system is that in some cases the operating state of the machinery is not know and cannot be easily assessed. For these cases, clustering of the data may unveil some classes of data which exhibit similar behavior and which may be considered operating states.

In the clustering module a new distance measure is introduced that uses a weight matrix to assign relative importance to regions in the spectra. Formulas are used to determine critical regions where fault-related information is contained, and these regions are assigned higher weights than other regions in the spectra. The matrix is used in the calculation of distances and makes the distance measure more sensitive to differences in those regions which we know (from theoretical considerations) contribute more to the solution.

When calculating centroids, the matrix of weights "pulls" the centroid in the "right" direction according to the setting of the weights. Sidebands are considered by defining a neighborhood about critical frequencies, and assigning weights to the frequencies in the neighborhood, the factor of importance assigned to sidebands is lower than that assigned to the critical frequencies.

PART II

VIBRATION MONITORING OF REACTOR INTERNALS

CHAPTER 1

VIBRATION AND MOTION OF INTERNALS IN PRESSURIZED WATER REACTORS

The reactor internals is perhaps the most complex mechanical system in a nuclear power plant (Figure II-1.1). The system is subject to loading from thermal transients, flow-induced vibrations, neutron irradiation, corrosion and fatigue [66] which inevitably produces degradation and in extreme cases may lead to unplanned outages. Neutron noise analysis techniques are used to obtain information concerning the integrity of the components and to determine when to perform effective preventive maintenance. Noise techniques provide means to verify integrity throughout the design life and a basis for extended life operation [66].

Monitoring of reactor internals using ex-core neutron noise started almost three decades ago and it is still used to assess the structural integrity of core support structures. There are a variety of economical, safety and regulatory considerations for surveillance activities. Thie in reference [65] discusses this issue in detail. In general, the surveillance of motion of reactor internals focuses on the assessment of vibration, fretting damage and part loosening; and aims at detecting motion in the structure or its components, quantifying the severity of the motion, assuring that it remains within safety limits, and ideally predicting a time frame of safe operation.

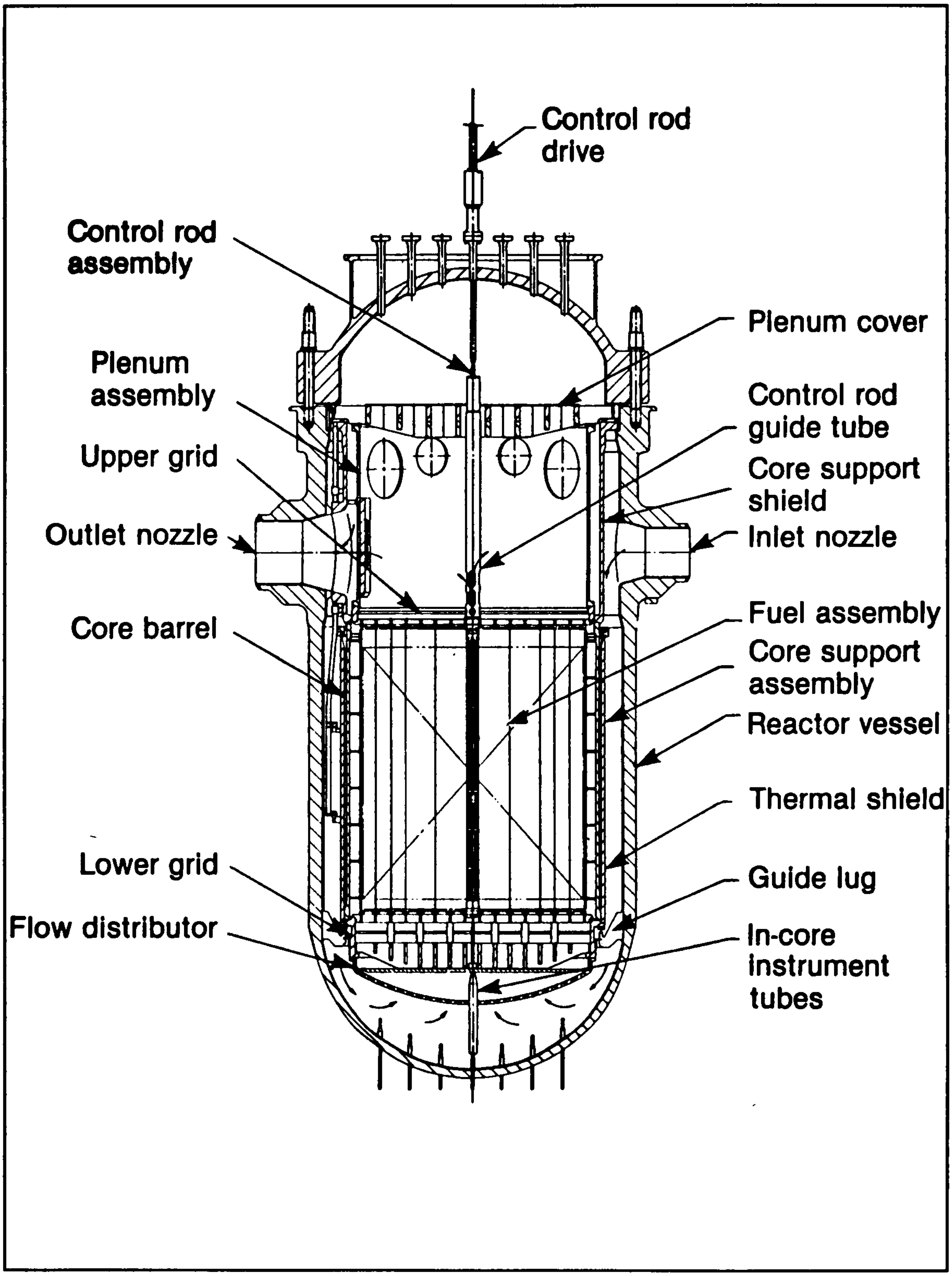

Figure II-1.1. Internals of a 900 MW Pressurized Water Reactor [67]

The core barrel in a PWR holds the core structure and separates the incoming downward flowing coolant from the upward flowing paths through the fuel elements [66]. Since there is always flow in the vessel, it is impossible to prevent the core barrel from moving, especially because the core barrel is suspended from an attachment to the vessel -- clamped at the top-- and moves in a pendular motion described as beam mode. The aim of the monitoring effort is to ensure that the motion remains within safety limits.

The primary coolant is discharged by pumps located in the cold-leg piping, enters the vessel through the inlet nozzles, and turns 90 deg. downward after colliding with the core barrel. The effect of this collision plus the hydraulic turbulence in the downcomer is the primary source of forces that causes motion of the core barrel [66]. This motion is affected by masses within the barrel (core, structures above the core, and structures below the barrel) and by other elements in the system such as piping, pumps and steam generators which are also multimass, multispring-coupled systems.

There are other significant sources of motion. First, the shell mode of movement caused by the barrel and its thermal shield changing internal area and circumferential shape, a behavior typical of annular structures; second, forces generated by the vibration of mechanical components in the internals such as the control rods and the fuel elements. There is also a random component of motion caused by stochastic forces within the vessel, and the vessel itself which moves along the same axis as the core barrel since the two masses are coupled. Finally, there is an element of motion which is very important from the monitoring point of view, that is the motion generated by anomalous behavior of the structure induced by contacts between the core barrel and the pressure vessel, or by the failure of mechanical components. These forces combined produce a complex motion which is very difficult to measure.

In a PWR the ex-core neutron detectors are ionization chambers located outside the vessel within one foot of the reactor vessel at cross core locations. Each chamber has two detectors sensing separately upper and lower current; from these readings, indirect measurements of motion and vibration may be obtained. The use of ex-core neutron detectors is convenient for monitoring vibration of core components because they are permanently mounted equipment, and because there are space and radiation restrictions which impede the use of in-core detectors. When the motions are large enough in the frequency band of interest, they permit accurate analysis.

The positioning of the sensors on the vessel permits the estimation of the amount of motion. The change in the water gap produced when the core barrel moves in a beam mode shown in Figure II-1.2, can be detected by out-of-phase signals in opposite ion chambers [65]. The amount of the change in the water gap is directly related to the amount of motion. To estimate the amount of motion, a widely used technique involves the use of the normalized cross-power spectral density (NCPSD) of ex-vessel ion chambers located in opposite sides of the vessel (180 deg.

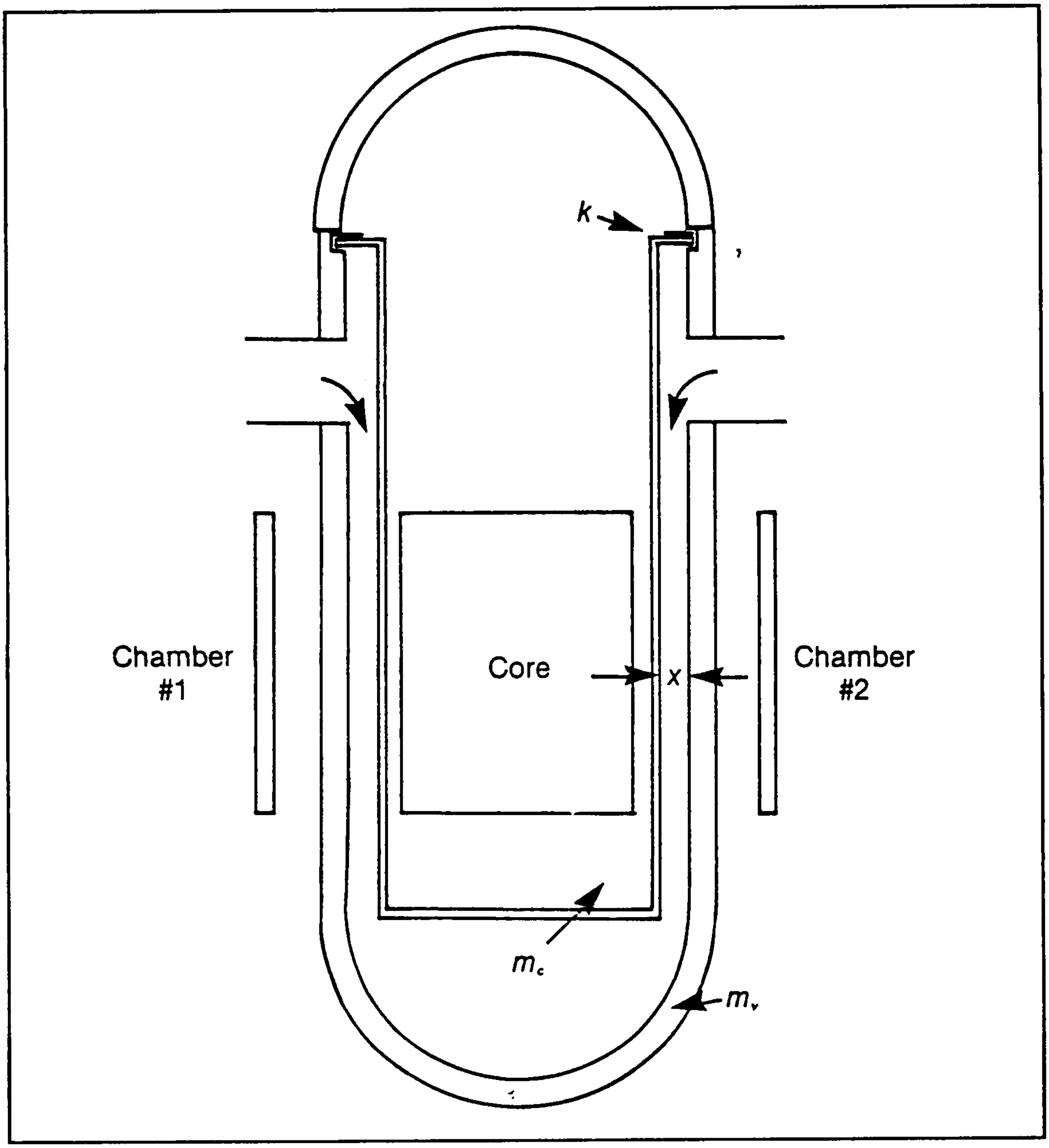

Figure II-1.2. The Relative Displacement of the Barrel and Vessel Affects Neutron Transmission [67]. The core barrel mass m_c and the vessel mass m_v moves as a double pendulum system coupled at the vessel flange by the spring constant k of the upper barrel

apart) [67]. Opposite chambers are used because their signals typically exhibit large coherences. The relationship is expressed in Reference [67] as follows:

$$rms\ motion = \frac{\sqrt{\int NCPSD\ df}}{\mu} \qquad \text{(II-1.1)}$$

$$rms\ motion = \frac{\sqrt{\dfrac{\sum \pi \Delta f\ NCPSD(f_o)}{2}}}{\mu} \qquad \text{(II-1.2)}$$

where μ is the neutron attenuation coefficient (approximately 0.009 mm^{-1} expected for all PWR's), and the integration range is over the frequency band containing the resonance and having a 180-deg. phase shift. Typical motions are described as within 0.02 to 0.06 mm [67], with f_0 ranging from 5 to 11 Hz. The values of motions estimated using Formulae II-1.1 and II-1.2, represent the geometric projection of the relative motion along the plane containing the two chambers. The rms motion may vary during the fuel cycle due to slight temperature and power variations [67].

The motion caused by the shell mode is more difficult to detect because it consists of changes in area and circumferential shape which are symmetrical around the center of the core. As a result, the fuel-to-vessel-wall distance remains unchanged, and the motion cannot be detected by examining the size of the water gap. To obtain an estimation of this motion, it is necessary to cross-reference in-core detectors, located on the core edge, and ex-core detectors located near the vessel.

Other motions such as vertical modes and motion of the control rods and fuel may also be detected as shown by Thie in reference [67].

The neutron noise spectrum is calculated from the noise portion of the ex-core neutron signal. The signal is caused by variations of the neutron flux around its average value. According to reference [68], variations may be caused by:

- a source of natural reactivity $\rho_n(t)$ related to the phenomenon of neutron creation/disappearance,
- sources of reactivity $s_i(t)$ related to modifications in the neutron flux as registered by the detectors such as thermohydraulic fluctuations, or structural movements between internals and reactor vessel [68]. Variations in coolant thickness between internals and reactor vessel causes a variation in the neutron flux hitting the detector [69].
- noise generated by the detector itself $b(t)$

The sources $\rho_n(t)$ and $s_i(t)$ are recognized by the detector from the transfer function of the core $h(t)$. The fact that this function links the sources of reactivity to the relative fluctuations of neutrons in the vicinity of the sensor was pointed out by Trenty in Reference [68]. The neutron noise signal $x(t)$ may be described using equations II-1.3 to II-1.7.

$$x(t) - h(t) * \rho_n(t) + \sum h(t) * s_i(t) + b(t) \qquad \textbf{(II-1.3)}$$

where * refers to the convolution operation.

Since $b(t)$ and $n(t)$ are white noise sources in the bands of interest, the power spectral density, PSD, may be expressed as in Equation II-1.4.

$$PSD(f) = |\, H(f)\, |^2\ x \quad \rho N + \sum |\, H(f)\, |^2\ x \quad S_i(f) + B \qquad \textbf{(II-1.4)}$$

where N is the constant linked to natural reactivity and B is linked to sensor operation [68]. Therefore Equation II-1.4 gives

$$PSD\ (f) = |\, H(f) |^2 \ \left[\ \rho N + \sum\ S_i(f) \right] + B \qquad \textbf{(II-1.5)}$$

The PSD may be expressed as a function with a general shape (Equation II-1.6) with amplitudes added at local frequencies (Equation II-1.7), since the functions $s_i(f)$ are non-zero only at the local frequencies $f(i)$ at which the phenomenon appears.

$$PSD(f) = \rho N \ \ |\, H(f) |^2 + B \qquad \textbf{(II-1.6)}$$

$$|H(f)|^2 \ \ S_i(f) \qquad\qquad for \ \ f \approx f_i \qquad \textbf{(II-1.7)}$$

There are two characteristics of the reactor noise process worth mentioning. First, it is a stationary process, meaning that the underlying nature of the randomness that causes the fluctuations does not change in the course of time. Secondly, most reactor noise encountered has normal probability density [67]. A variety of tests are used to detect departures from Gaussian behavior, since generally anomalous mechanical motion causes non-Gaussian behavior of the noise.

The noise spectrum depicts specific features related to the behavior of the structure, and under normal conditions there are few differences among spectra from reactors of the same family. Traditionally, baseline spectra are calculated from ex-core neutron signals collected during normal operations at the beginning of the fuel cycle, and as the cycle progresses, surveillance data is compared to the baseline to determine the presence or absence of excessive reactor internal vibrations [70]. If excessive vibration is detected, analysis is performed to determine the cause and severity of the anomaly. The known associations between the regions of the neutron noise spectrum and the components in the structure are used to identify the source of excitation. These associations are presented in Table II-1.1.

Table II-1.1. Associations between Regions in Neutron Noise Spectrum and Internal Components [71,72,73]

Band Number	Frequency Region (Hz.)	Description
10	0.5 to 2.0	Thermal behavior of core neutrons, temperature changes and/or reactivity
11	2.0 to 5.0	Fuel assembly beam mode vibration, and reactivity effects
12	5.0 to 7.5	Second beam mode of the fuel assemblies
13	7.5 to 8.0	Free swinging of the core barrel
14	10.0 to 14.0	Main mode of the internals. Contains information regarding contacts between the core and the vessel
15	14.0 to 18.0	Secondary CSB vibrations, swinging of the vessel
16	18.0 to 24.0	CSB primary shell mode vibration
17	24.0 to 48.0	Sensor background noise, internals complex modes

The neutron noise spectrum is also useful for assessing the integrity of the structure because it is very sensitive to mechanical failures. A study conducted by the Gesellschaft für Reaktorsicherheit (GRS) in Germany for a variety of failures concluded that each kind of failure causes its own pattern of spectral changes, and furthermore that each pattern of spectral changes is characteristic of only one particular failure [74]. Among the kinds of failures investigated in the study, were failures of the steam generator supports, and degradation of the reactor internals hold-down spring. The hold-down spring is a metal device in the shape of a ring placed on the flange of the pressure vessel top to tighten the connection with the flanges of the upper core support structure (See Figure II-1.1). When the spring relaxes it allows the barrel, to move more freely.

There are several factors that complicate the problem of identifying vibration modes in neutron- noise spectra. First, as the fuel cycle progresses, the amplitude of the signatures tends to increase due to the burnout of the neutron absorbing boron. A study by EDF concluded that the partial RMS value in all the vibration bands generally follows a linear law versus burn-up, but with different coefficients according to the fuel cycle even within the same frequency band [71]. In addition, contacts between the vessel and the core which are usually manifested in the 7.5 to 14.0 Hz region, tend to disappear as the fuel cycle progresses, but the opposite situation sometimes occurs. The phenomenon is probably due to a change in the positions of the internals [71].

Another problem that arises frequently is the formation of composite peaks. The detector registers composite peaks assembled from different peaks at like frequencies (features which relate to different modes but appear at the same frequency). Although the original peaks may be related to motions from a particular component of the internals, it becomes difficult to isolate the changes individually [73]. Periodic surveillance of the internals is necessary in order to detect the migration of these peaks before they merge.

This study deals with the development of a system for automatic discrimination of neutron noise signatures from ex-core detectors in 900 MWe pressurized water reactors of the CP1-CP2 family with sectorized thermal shield (Figure II-1.3). Data reflecting normal state of operation from twenty eight reactors has been made available for the study. In addition, data representative of a class of mechanical failure has been generated using a software simulator.

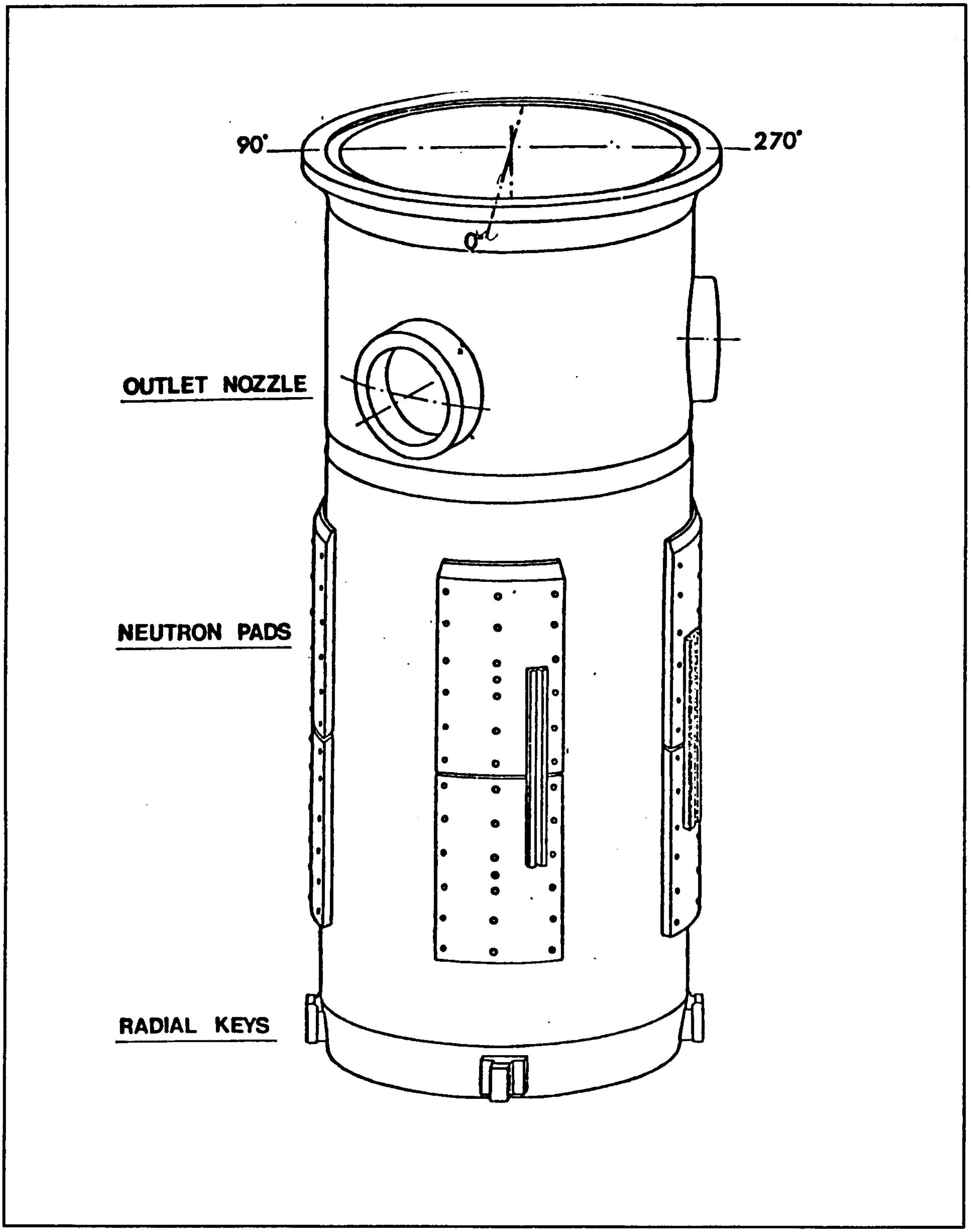

Figure II-1.3. Internals with Sectorized Thermal Shield [74]

CHAPTER 2

CLASSIFICATION SYSTEM

The classification system described in this report is based on the technology of neural networks. Information extracted from neutron noise spectrum is used to train neural network models to identify the different modes of vibration of the core barrel, and to verify the integrity of components in the structure. After training, the system is tested against classification results provided by an expert evaluator.

There are three phases to the classification, as shown in Figure II-2.1. In the first stage, a neural network module has been used to separate signatures from functional power detectors and signatures from improperly calibrated or out-of-order sensors. The second phase of processing is performed by a neural network trained to discriminate between signatures from plants with functional hold-down springs and simulated signatures representing plants with degraded or non-functional springs. The last part of the analysis corresponds to the classification of normal signatures according to the position of the internals relative to the vessel, and the presence of contacts between the core barrel and the pressure vessel.

The commercial neural network program used during the study is SN2.5 developed by Neuristique Corp. For the classification, three-layer backpropagation networks were used. The networks take as input scaled feature vectors representing each spectrum, and produce as output a class label. The output layers in each case

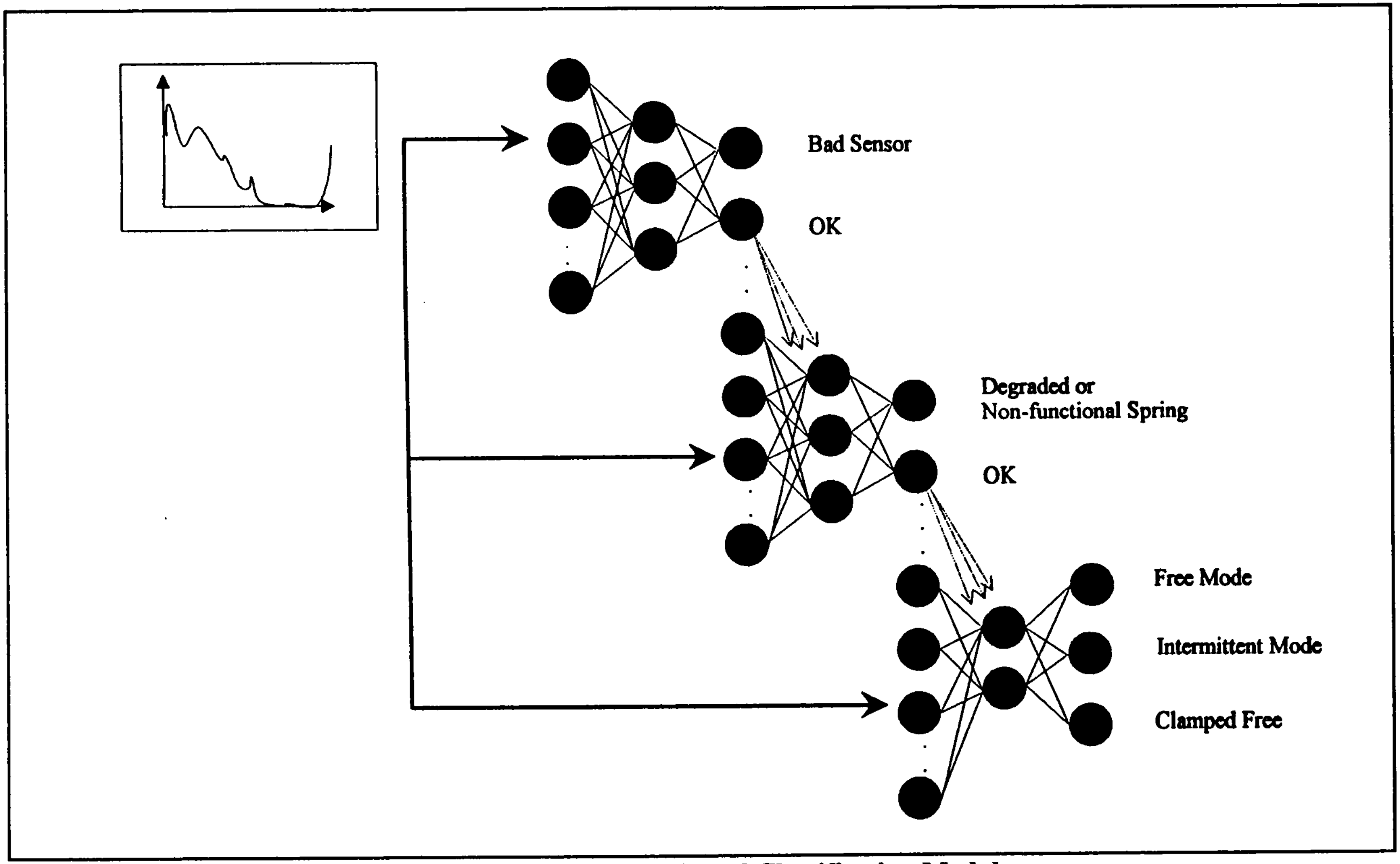

Figure II-2.1. Organization of Classification Modules have one neuron corresponding to each possible class, and the activation of a neuron indicates membership to the class it represents. The topology of the networks is depicted in Figure II-2.2.

Prior to training the neural network, the input vectors are scaled to the operating range of the algorithm (0.05, 0.95) using Equation II-2.1, where A is a matrix of input vectors. For A_{ij} the subscript i corresponds to the vector number, and j indexes the variables in the vector.

$$A_{ij}^{*} = \left(\left(\frac{A_{ij} - \min[A]_j}{\max[A]_j - \min[A]_j}\right) * 0.90\right) + 0.05 \qquad \textbf{(II-2.1)}$$

Scaling parameters are calculated for each column in the input matrix to preserve the characteristics of each variable.

DATA DESCRIPTION AND FEATURE CHARACTERIZATION

Spectral signatures were calculated from 742 measurements collected over 93 fuel cycles from PWR's in France. The plants correspond to CP1-CP2 900 MW plant series with 3 loops and sectorized thermal shield. The data comprise normalized power spectral densities, NPSD's, calculated from ex-core neutron detectors in the range 0.125 Hz. to 50.0 Hz.

Time records are collected following standardized procedures in all units during normal reactor operation, generally at nominal full power. The data are processed with software programs developed by EDF which calculate auto and cross spectral functions and, from these functions, a number of parameters (Figure II-2.3)

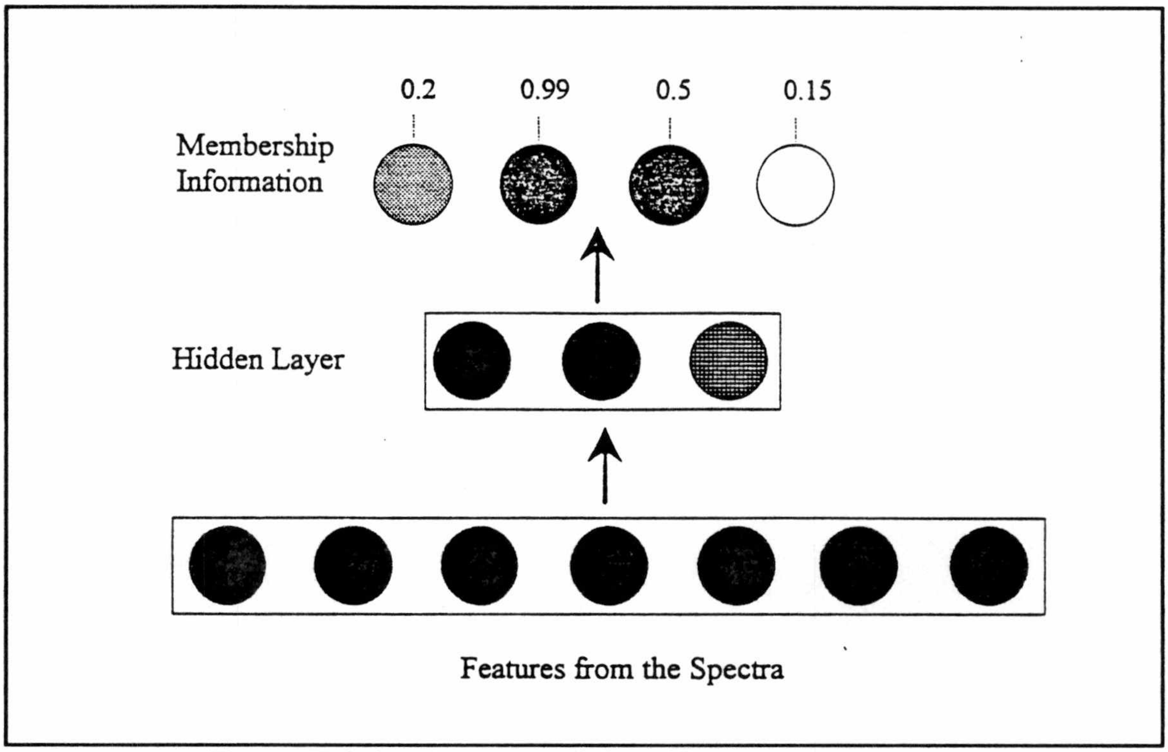

Figure II-2.2. Topology of Classifier Neural Network

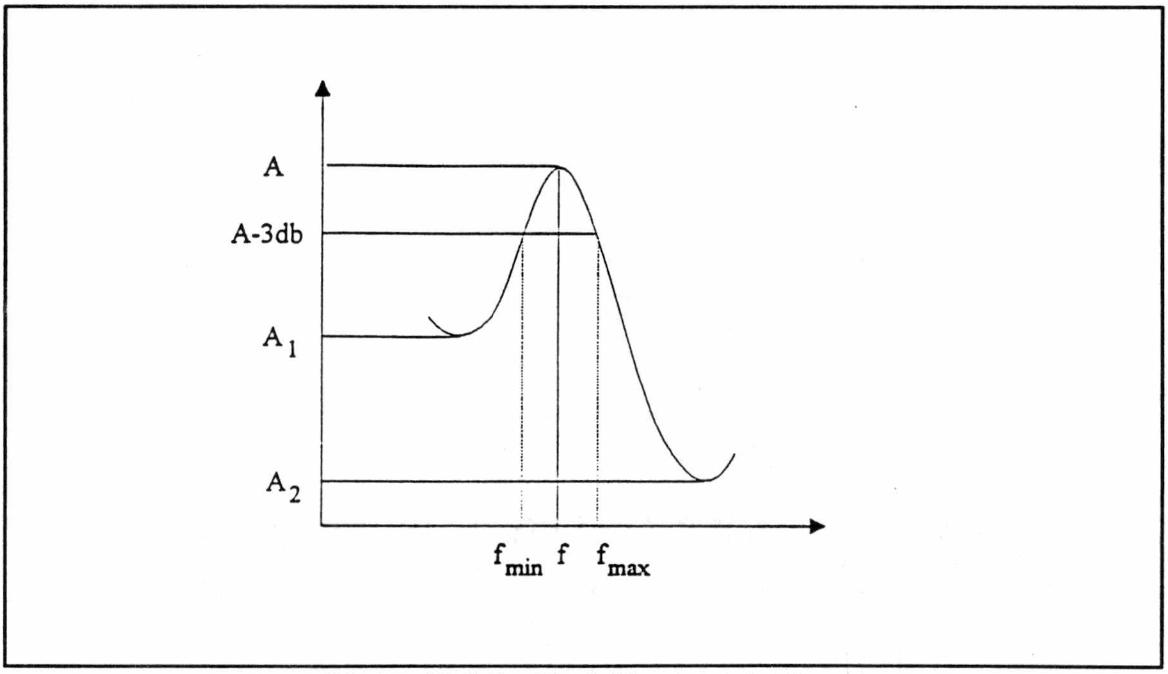

Figure II-2.3. Peak Parametrization

related to peaks which are useful in the quantification of structural vibration and motion. The data are validated and stored in a database under a standard format.

The parameters calculated from the 400-point spectra are used to describe the peaks in a consistent manner, and may be used in conjunction with other descriptors for discrimination of the signatures. EDF has defined the following parameters, described in Reference [71]:

Amplitude: A linear measure of the maximum amplitude of the peak.

Frequency: The frequency of the maximum amplitude of the peak.

Height: A measure of the height of the peak, denoted H and calculated using Equation II-2.2, (Refer to Figure II-2.3)

$$Height = H = \frac{2A - A_1 - A_2}{2} \qquad \textbf{(II-2.2)}$$

Energy: An approximation of the area under the peak, denoted E and calculated using Equation II-2.3. The value of Δf is 0.125.

$$Energy = E = \int_{f_{max}}^{f_{min}} NPSD\ df \approx 0.125 \sum_{f_{max}}^{f_{min}} NPSD(f) \qquad \textbf{(II-2.3)}$$

Displacement: Calculated using Equation II-2.4, where h is a transmission factor between the neutron normalized fluctuations and the displacement of the internal structure. The factor h depends on internal structure design.

$$Displacement = D = \frac{1}{h}\sqrt{E} \qquad \textbf{(II-2.4)}$$

Damping: A measure of the width of the peak, See Equation II-2.5.

$$Damping = \eta = \frac{E}{2\,f\,A} \approx \frac{\Delta f}{2\,f} \qquad \textbf{(II-2.5)}$$

A root mean square (*RMS*) value for each band is calculated (Equation II-2.6), and a relation PR_i representing the relative power of band i (Equation II-2.7).

$$RMS = \sqrt{\int_{f_{max}}^{f_{min}} NPSD\; df} \qquad \textbf{(II-2.6)}$$

$$PR_i = \frac{(RMS_i)^2}{T_{Energy}} \qquad \textbf{(II-2.7)}$$

Information is also collected regarding the burn-up rate and initial fuel enrichment in the reactor. After detailed analysis, experts at EDF [72] have concluded that only these two parameters have a significant effect on ex-core detector readings during steady-state operation (i.e. stable boron concentration and no control rod movement). The effect of these parameters is generally linear in the bands corresponding to the main modes of internal structures and thus may be removed from the spectra by appropriate normalization of the signatures. The effect is observed in Figure II-2.4 where a signature representing free beam mode of a reactor is shown at the beginning and end of a fuel cycle.

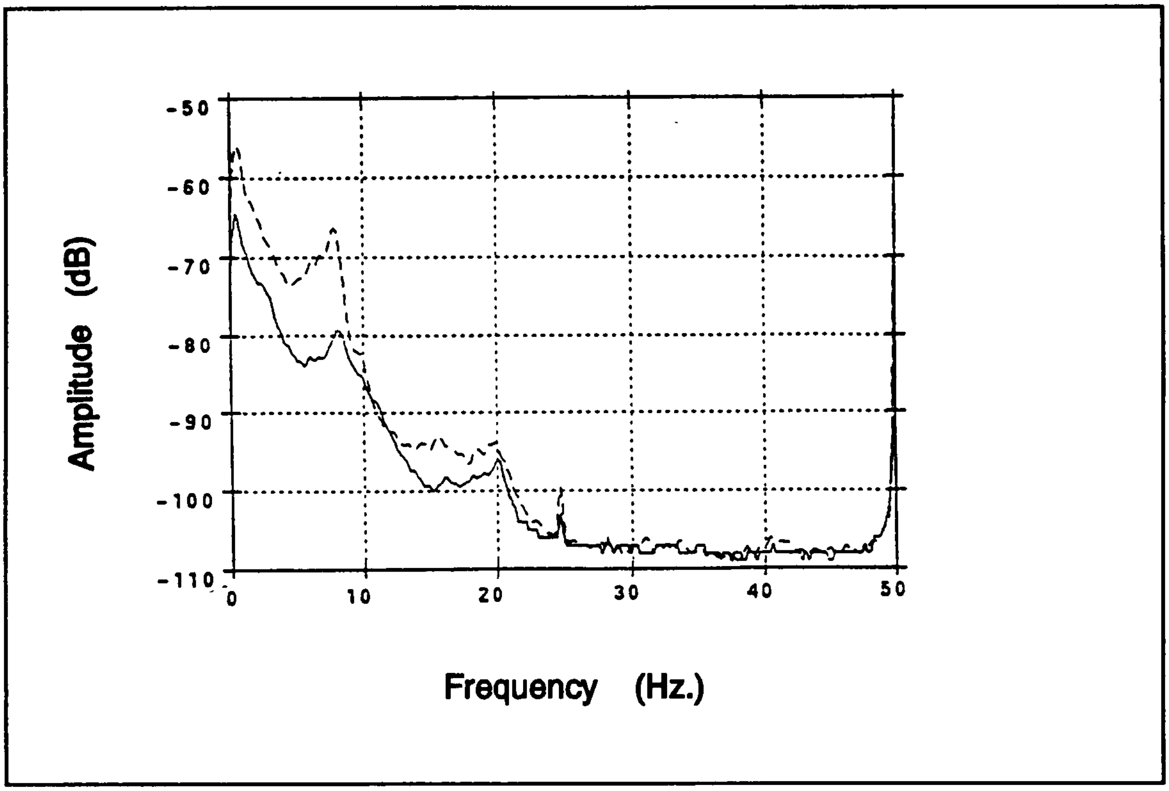

Figure II-2.4. Amplitude Shift During Fuel Cycle.
(--- beginning of fuel cycle, - - - end of fuel cycle)

To account for the shift in amplitudes due to burn-up and fuel enrichment the signatures are normalized. To perform the normalization, the total energy of the signature is calculated (Equation II-2.8) and the value used to scale parameters corresponding to the energy of the peaks and their amplitudes. Only global parameters (such as height) are normalized, since relative measures are not affected by the shift.

$$T_{Energy} = \sum_i E_i \qquad \textbf{(II-2.8)}$$

MODULE I: DISCRIMINATING VALID AND INVALID SIGNATURES

The processing of data is separated into three modules. The first module deals with the separation of signatures in two groups. The first group is formed by signatures collected with functional ex-core neutron detectors, and the second group by signatures from detectors which were either not calibrated properly or out-of-order. The neural-network-based system uses information from the spectrum to discriminate among signatures whose levels of amplitudes are valid, and those with levels above or below the typical values.

The majority of the invalid signatures correspond to sensors which are not properly calibrated and which provide readings above or below the correct levels (Figure II-2.5 and II-2.6). NPSD's are also available from faulty sensors which produce signatures whose shapes are not consistent with typical neutron noise data from reactors of the PWR 900 MW family.

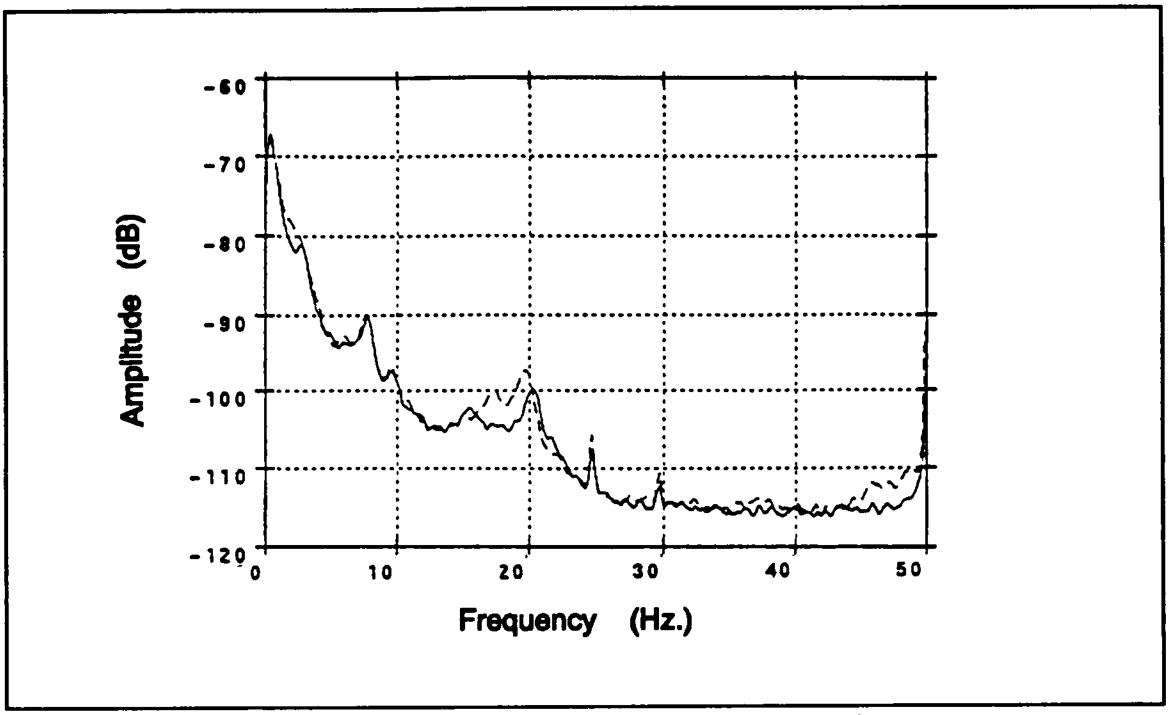

Figure II-2.5. Faulty Sensor, Low Levels

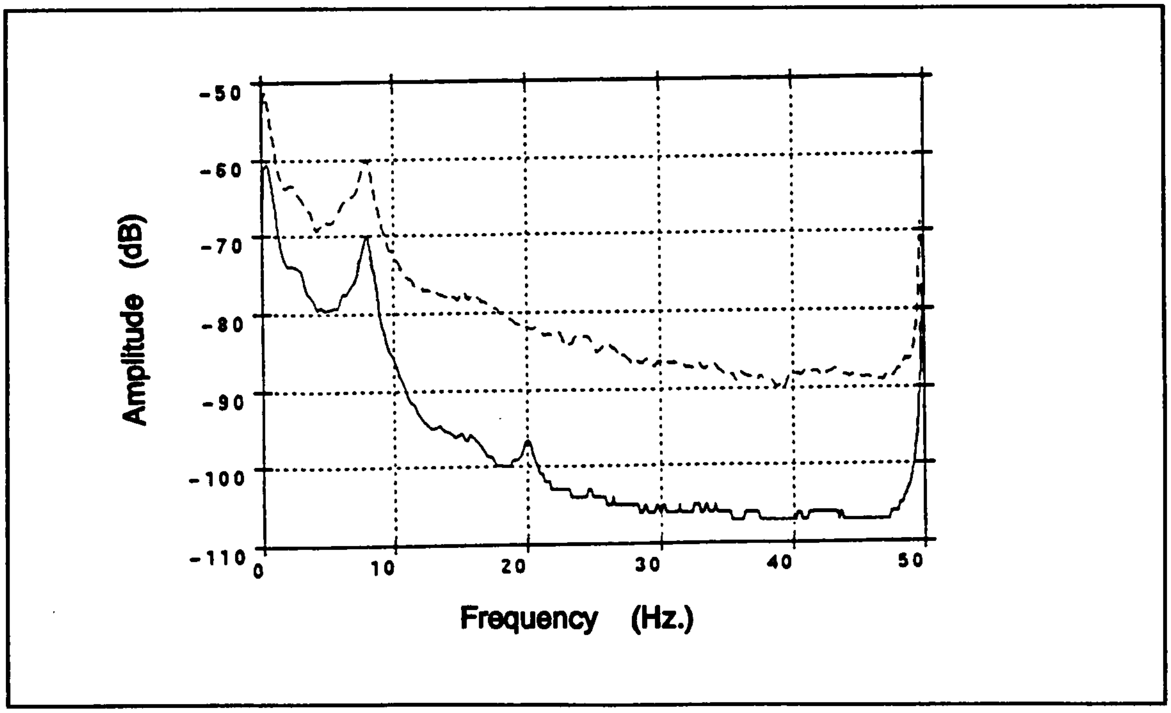

Figure II-2.6. Faulty Sensor, High Levels

System Description

The analysis of these data is complicated because the NPSD generated from readings of out-of-order detectors have many different shapes. It becomes difficult to find parameters to describe consistently the features in these spectra. A number of parameters have been chosen which represent the most stable and representative regions of the spectrum. A list of these variables is presented in Table II-2.1

Table II-2.1. Parameters Used for Discrimination, Module I

Variable	Description
Amp_4	Lowest Amplitude in Interval (3.0 - 6.0) Hz.
Amp_24	Lowest Amplitude in Interval (22.0 - 25.0) Hz.
m	Slope of the Line Joining Amp_4 and Amp_24
Lowest_Amp	Lowest Amplitude in Band 17*
High_Amp	Largest Amplitude in Interval (22.0 - 24.0) Hz.

* Excluding the region (27.5 - 30.5) Hz.

The parameters ***Amp_4, Amp_24*** and the slope of the line joining the two points, ***m***, were used because they represent the most stable features of the spectrum. Usually the presence of contacts disrupts these regions only slightly while, on the other hand, out-of-order sensors generally produce slopes which are different due to distortions in the signature. ***Lowest_Amp*** is the parameter used for detecting low values of amplitudes. The parameter is extracted from the sensor background noise region (Band 17), excluding the interval (27.5 - 30.5) Hz. because in that interval sharp peaks produced by electrical parasites have been registered in two plants.

Finally, the parameter **High_Amp** measures the highest amplitude of the sensor in the stable region and it is used to identify signatures with high amplitude levels.

Each one of these parameters is the input to the neural network. The network has 2 hidden nodes and 3 nodes in the output layer, each corresponding to a possible state. The states considered for the classification are: signatures with high amplitude levels, signatures with low amplitude levels, and valid signatures.

Data Set

The total set of data contained 1400 signatures; of these, 137 signatures were collected with improperly calibrated or out-of-order sensors (ICOO signatures) and 1263 valid signatures with normal amplitude levels (NA). The ICOO signatures can be placed in three categories: a group of signatures whose amplitude values are higher than normal, a group with low amplitudes, and a group of highly distorted signatures. The first two groups usually correspond to sensors which are not properly calibrated (i.e. the gain in the device is not adjusted correctly and produces signatures which are shifted in amplitude). The last group is formed by signatures collected with sensors which are out of order.

Generally, the signatures coming from out-of-order sensors have high amplitude values and for this project have been included in the class of high amplitudes. The results discussed are based on classification of the data in three groups: High Amplitudes (HA), Low Amplitudes (LA) and Normal (or typical) Amplitudes, NA. The distribution of signatures is shown in Table II-2.2.

Table II-2.2. Distribution of Signatures, Module I

Type of Signature	Number of Signatures
Normal Amplitudes (NA)	1263
High Amplitudes (HA)	40
Low Amplitudes (LA)	97
Distorted Signatures (out-of-order sensors)	8 (included in HA)

The NA signatures may in turn be divided into three categories according to the vibratory mode of the internals with respect to the vessel: a mode of free swinging of the vessel - denoted L - with a single, strong and well defined peak at 8 Hz; a second mode denoted AL in which two modes of vibration co-exist, a clamped mode of permanent contacts and a free mode. AL signatures have strong, well-defined peaks at 8 Hz and at 10 Hz; and a third mode of intermittent contacts also exists for which signatures are very irregular with no significant peaks in the 8.0 to 14.0 Hz. region. The distribution of the NA signatures is presented in Table II-2.3.

Table II-2.3. Distribution of NA Signatures, Module I

Type of Signature	Number of Signatures
Free Mode (L)	1070
Clamped (AL) Mode	82
Intermittent Mode (I)	111

Results

The classifier neural network described earlier was trained for 23,500 iterations until the error was reduced to 0.02. During training the learning rate was reduced from 6.4 to 0.05, and a sigmoid logistic function used with a gain of 0.999. For training, 58 vector representative of the different classes were selected from the set of 1400. The patterns were distributed as shown in Table II-2.4.

Table II-2.4. Training Set for Module I

Class	Vibration Mode	Number of Signatures
Normal Amplitudes (NA)	Free (F)	18
Normal Amplitudes (NA)	Clamped (AL)	8
Normal Amplitudes (NA)	Intermittent Contacts (I)	7
High Amplitudes (HA)		16
Low Amplitudes (LA)		9

The criteria used for performance evaluation was the following: a pattern is considered correctly classified if the activation of the neuron representing the class to which it belongs, is the highest and greater than 0.25. Recall was performed on the 1342 remaining patterns, from this group 1288 patterns (92.4 %) were classified correctly, and 102 patterns (7.6 %) were classified incorrectly. In an effort to improve performance, different configurations of the network were used but none render better results. The cases which were classified correctly merit no further discussion, Table II-2.5 shows the distribution of the incorrectly classified signatures.

Table II-2.5. Confusion Matrix for Module I

Result / R. Class	NA	HA	LA	# Test Patterns	% Mis-classified
NA	--	87	12	1225	8.0%
HA	2	--	0	24	8.4%
LA	1	0	--	88	1.14 %

The majority of the misclassified cases are signatures with normal levels that were classified as having high amplitudes. For these cases the activation of the neuron corresponding to the LA class is slightly higher than that of the NA class. There are two reasons for the misclassification; first, the presence of high amplitudes in the sensor background region and in the interval (22.0 - 25.0) Hz. which is not apparent in the bands of interest (as it relates to contact detection). An example is presented in Figure II-2.7. Secondly, the deformations introduced by the presence of contacts affect regions beyond Band 14 and raise the amplitude levels in the area where the sensor information is extracted (Refer to Table II-2.1). Figure II-2.8 illustrates this situation, normal levels in these regions are shown in Figure II-2.4.

There are a few cases for which the classification provided by the neural network and the expert are different. These cases may represent boundary conditions and it is not obvious what is the correct classification. Examples are shown in Figures II-2.9 and II-2.10. Figure II-2.9 shows the plot of a Free (F) Signature whose amplitudes are higher than normal in most of the regions of the spectrum (including Band 14). Figure II-2.10 illustrates a case of a boundary

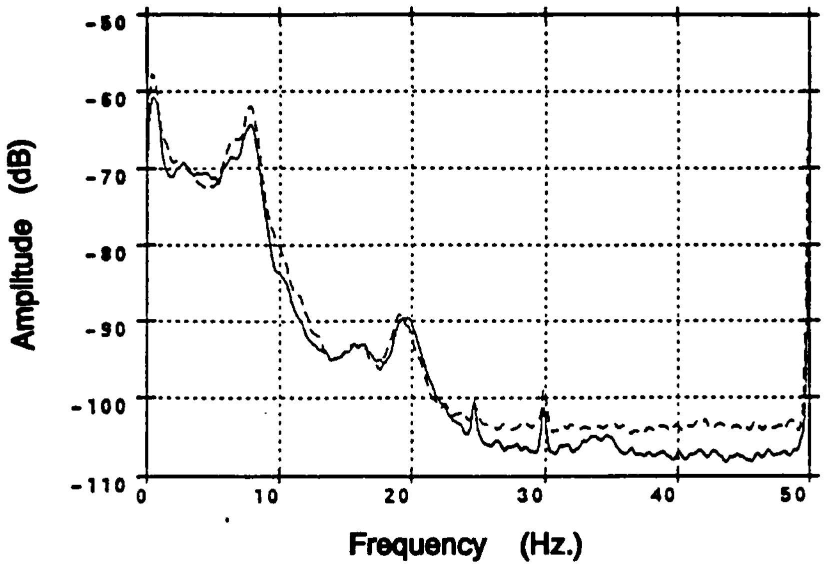

Figure II-2.7. Spectrum with High Amplitudes in Interval (24.0, 48.0 Hz).
(- - - Signature with High Amplitudes, —_ Signature with Normal Amplitudes)

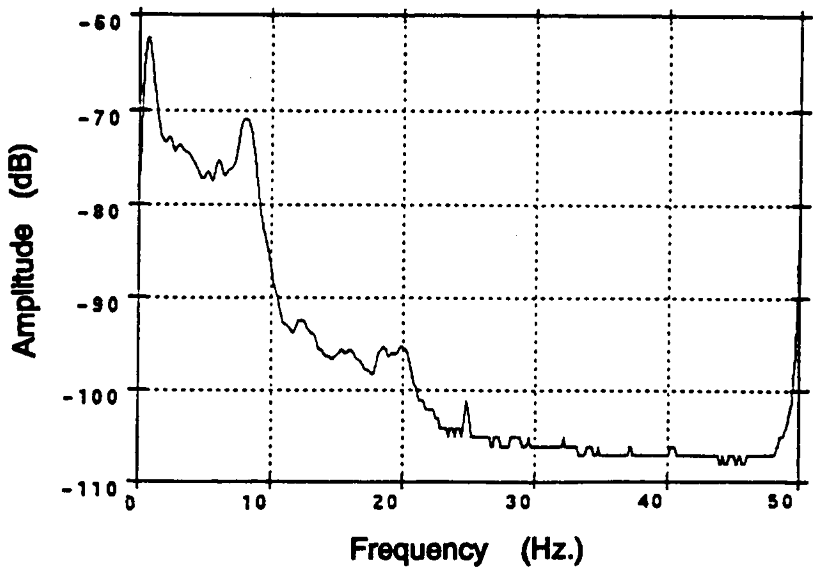

Figure II-2.8. Spectrum with High Amplitudes at 4.0 Hz. and the interval (22.0, 25.0) Hz.

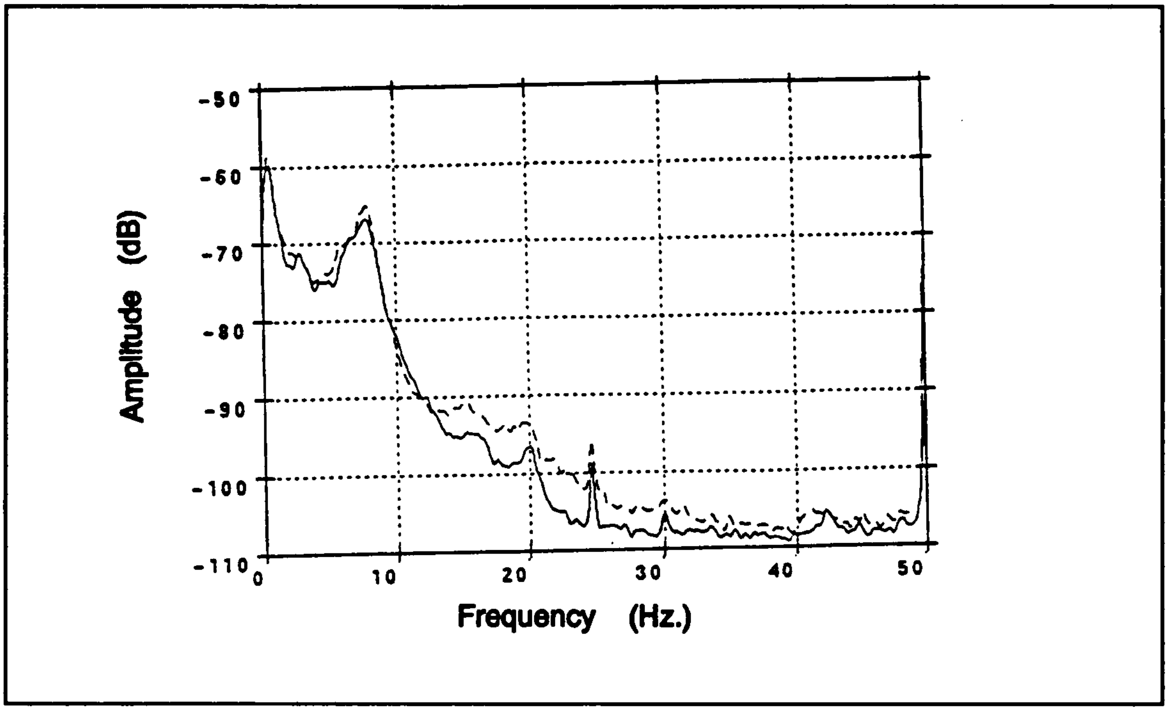

Figure II-2.9. Spectrum with High Amplitudes in Most Frequencies.
(- - - Free signatures with normal levels, --- Free signature with high levels)

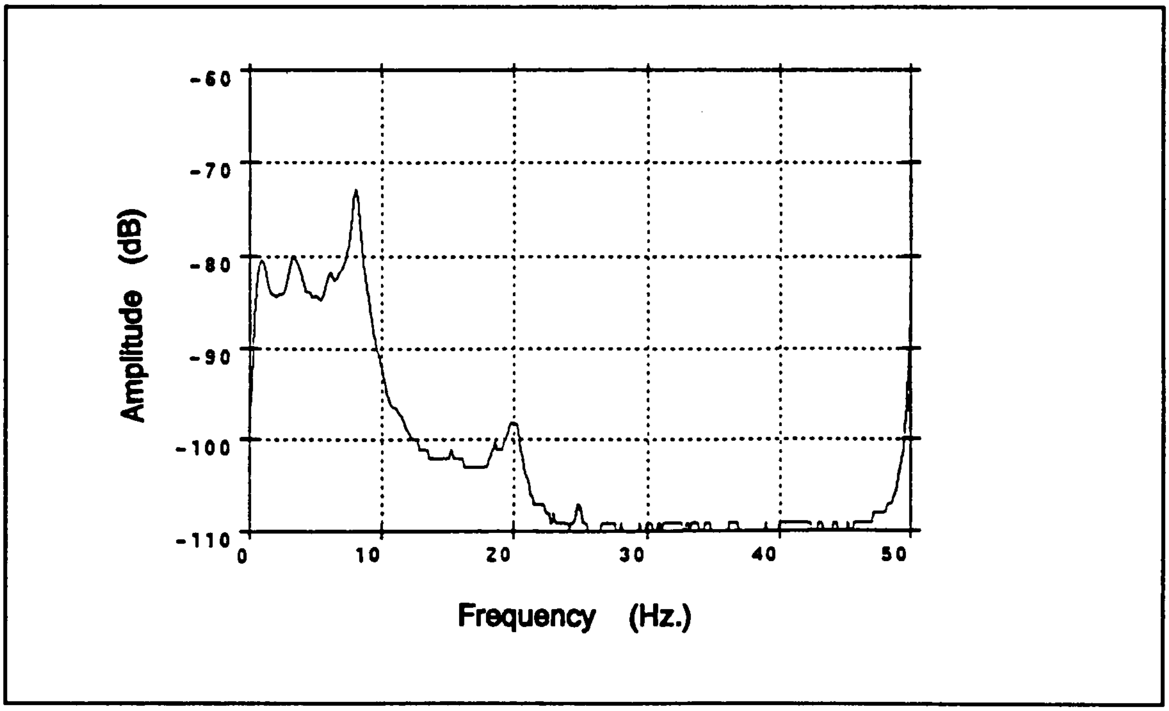

Figure II-2.10. Signature with Low Amplitudes in Interval (24.0, 48.0)

condition for a signature with levels slightly lower than normal. Another example is shown in Figure II-2.11, the signature is classified by the neural network as normal but is considered invalid by the experts.

MODULE II: DEGRADATION OF THE HOLD-DOWN SPRING

The hold-down spring is a metal device in the shape of a ring installed between the upper core support structure and the core barrel (see Figures II-1.1 and II-2.12). The function of the spring is to press against the two components, to constrain the movement of the core barrel. The spring undergoes a natural process of relaxation with aging, and the aim of this part of the study is to identify this type of mechanical failure from the NPSD of ex-core detectors. From the point of view of vibration surveillance, the degradation of the hold-down spring is important, and may become dangerous, because it translates into greater freedom of movement of the internals [3].

In the NPSD of ex-core neutron signals, the degradation of the hold-down spring may be detected because it causes a frequency shift of the peak corresponding to the free swinging of the core barrel, FSCB. The precise evolution of this effect is difficult to predict, but it is known that in all cases, the frequency of the FSCB peak decreases in proportion to the significance of the degradation [3,75], and that an increase in amplitude, and damping of the peak is concomitant with the frequency shift [72]. The evolution effect is illustrated in Figure II-2.13, for a spring which degrades from a normal to a totally non-functional state.

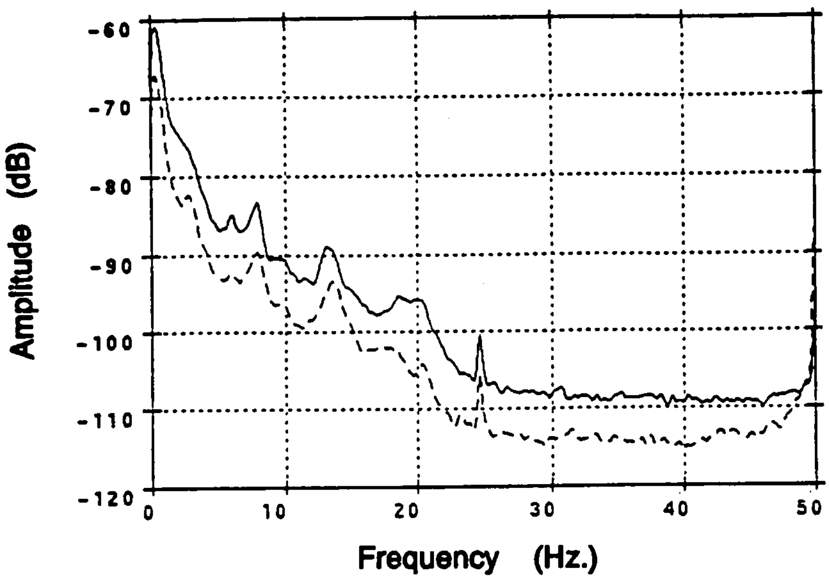

Figure II-2.11. Spectrum with Low Amplitudes levels.
(--- Normal Amplitude Levels, - - - Low Amplitude Levels)

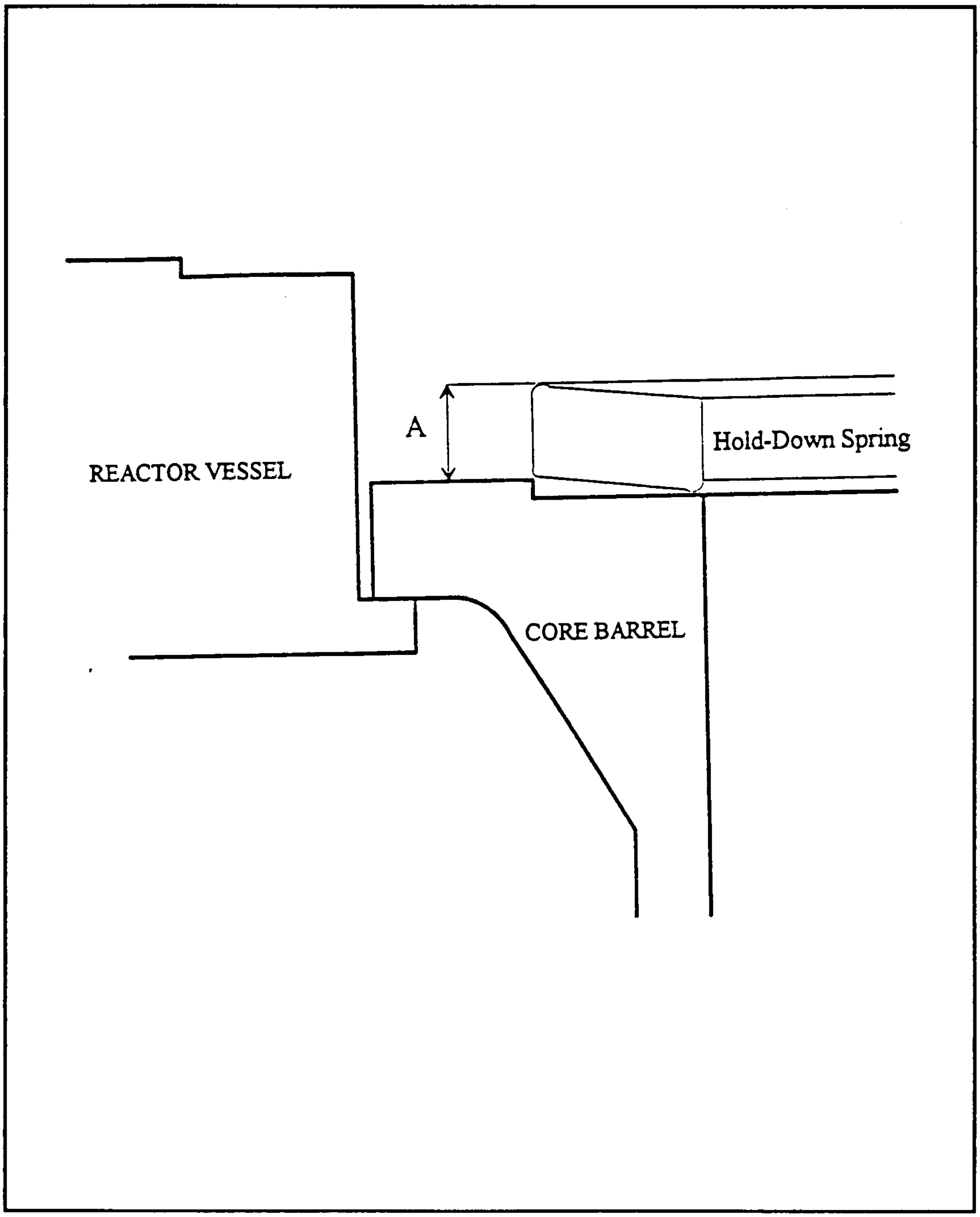

Figure II-2.12. Position of the Hold-Down Spring [74]

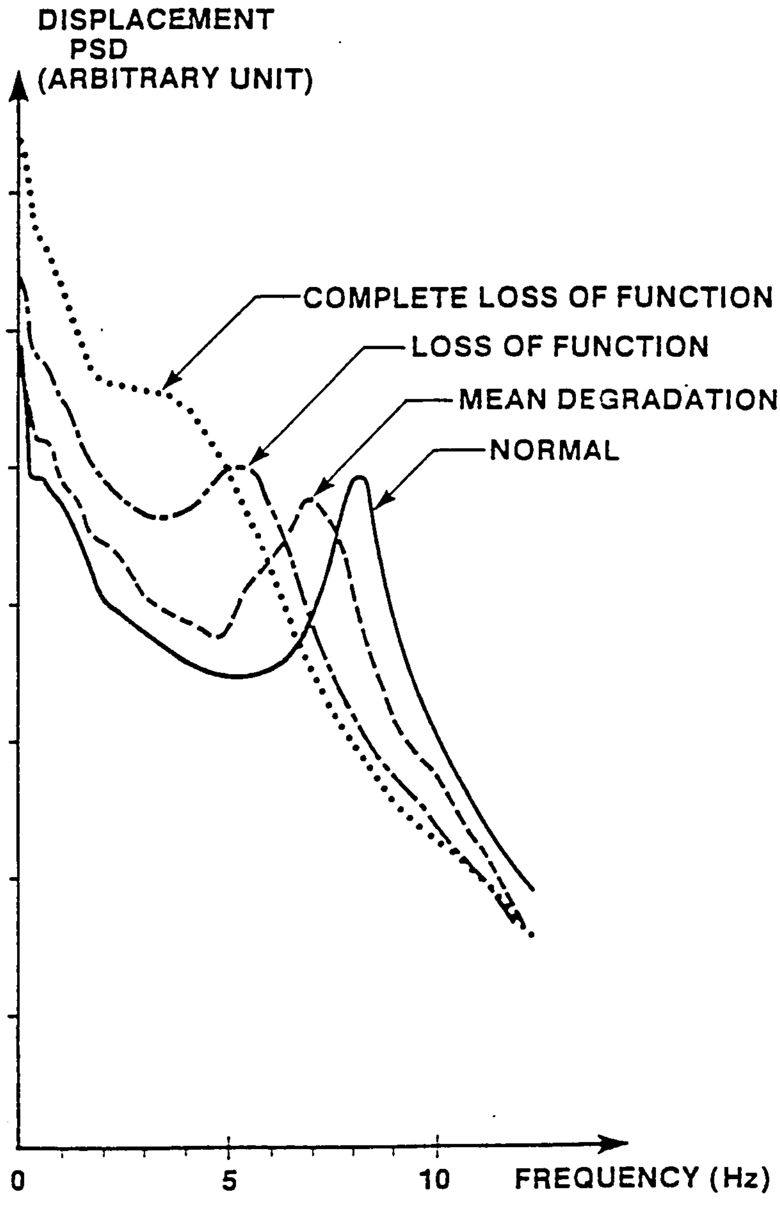

Figure II-2.13. Degradation of the Hold-Down Spring [74]

In French PWR's, no cases of hold-down spring degradation have been registered. In some of the older plants, shifts of the FSCB peak have been detected, but there has always been a systematic return to the normal frequency upon restart following shutdown for reloading. The migration of the peak back to the normal frequency indicates that there is no degradation but merely relaxation, and that any deformation is likely to be of an elastic nature [72]. Data of degraded hold-down springs from operating French plants is not available to perform this part of the study.

In order to study the phenomenon, a number of mechanical and mathematical models have been developed (see for example references [3,75]). For this project, Electricité de France has developed a device for generating abnormal signatures using data acquired from mechanical models of malfunctioning internals [68]. With the device, typical PSD's of ex-core neutron signals may be generated and modified to reflect abnormal conditions in the behavior of the internals. All of the data corresponding to degraded or non-functional hold-down springs used for the study were generated using the simulator.

System Description

For this module, a backpropagation network with three layers was used. The network had nine input units corresponding to the different parameters considered for discrimination, two hidden units, and three output units corresponding to the number of classes. Parameters used as input to the system are listed in Table II-2.6.

There are three classes of signatures. The fist category corresponds to normal signatures (NS). These signatures generally have a peak at 8 Hz., which corresponds to the free swinging of the core barrel. The second class is composed of signatures from units with degraded hold-down springs (DS), and for these signatures the peak corresponding to the free swinging of the core barrel is displaced to the left from 7.75 Hz. to a minimum of 7.0 Hz. The third category is from units with non-functional hold-down springs (NFS) for which the peak is displaced to frequencies below 7.0 Hz.

Table II-2.6. Parameters Used for Discrimination, Module II

Variable	Description
Psd_Max_1213	Max. Amplitude of Largest Peak, bands 12 and 13
Height_1213	Height of the Largest Peak
Energy_1213	Energy of the Largest Peak
Freq	Frequency of the Largest Peak
RMS_12	RMS Value of Band 12
RMS_13	RMS Value of Band 13
RMS_14	RMS Value of Band 14
Ratio_1213	Ratio of *RMS_12* and *RMS_13*
Ratio_1314	Ratio of *RMS_13* and *RMS_14*

Information from Band 12 was included in the analysis because the frequency shift of the FSCB peak produced by the degradation of the spring, causes the peak to migrate from Band 13 to Band 12 (See Table II-1.1).

Data Set

The set of signatures included 1316 normal signatures (from operating plants) and 370 simulated signatures representing units with degraded or non-functional hold-down springs. The distribution of the signatures is shown in Table II-2.7. From this set, 83 signatures were selected for training; these signatures were distributed as shown in Table II-2.8.

Table II-2.7. Distribution of Signatures for Module II

Type of Signature	Number of Signatures
NS	1316
DS	187
NFS	183

Table II-2.8. Distribution of Training Signatures for Module II

Type of Signature	Number of Signatures
NS	45
DS	16
NFS	22

Results

The network was trained for 22,000 iterations until the error was reduced to 0.01. During training, the learning rate was reduced from 12.8 to 0.05. The training set was chosen to reflect typical cases of each class, but it became very important to include boundary cases to improve performance.

Classification results for this module are very accurate, the confusion matrix is shown in Table II-2.9. The network was able to classify correctly all the normal signatures from intermittent (I) and clamped (AL) modes, although these signatures have in some cases poorly defined peaks in the regions of interest. The three misclassified NS signatures were boundary cases. For these signatures, the activation of the neuron corresponding to the NS class was approximately 0.4, and the activation of the corresponding neuron for the DS class was 0.6.

From the simulated data, there were also a few misclassified cases. Figure II-2.14 shows a boundary case between normal and degraded springs (classes NS and DS), and Figure II-2.15 illustrates a case of transition between classes DS and NFS. As becomes apparent from the graphs; the transition is very subtle.

Table II-2.9. Confusion Matrix for Data in Module II

Result R. Class	NS	DS	NFS	# Test Patterns	% Mis-classified
NS	--	3	0	1272	0.2 %
DS	5	--	10	171	8.7 %
NFS	0	1	--	161	0.6 %

MODULE III: DISCRIMINATION OF NORMAL SPECTRA

This part of the study deals exclusively with data representing normal behavior of the internals. This behavior includes modes of free swinging of the core barrel and modes where contacts occur between the barrel and reactor vessel. There are four

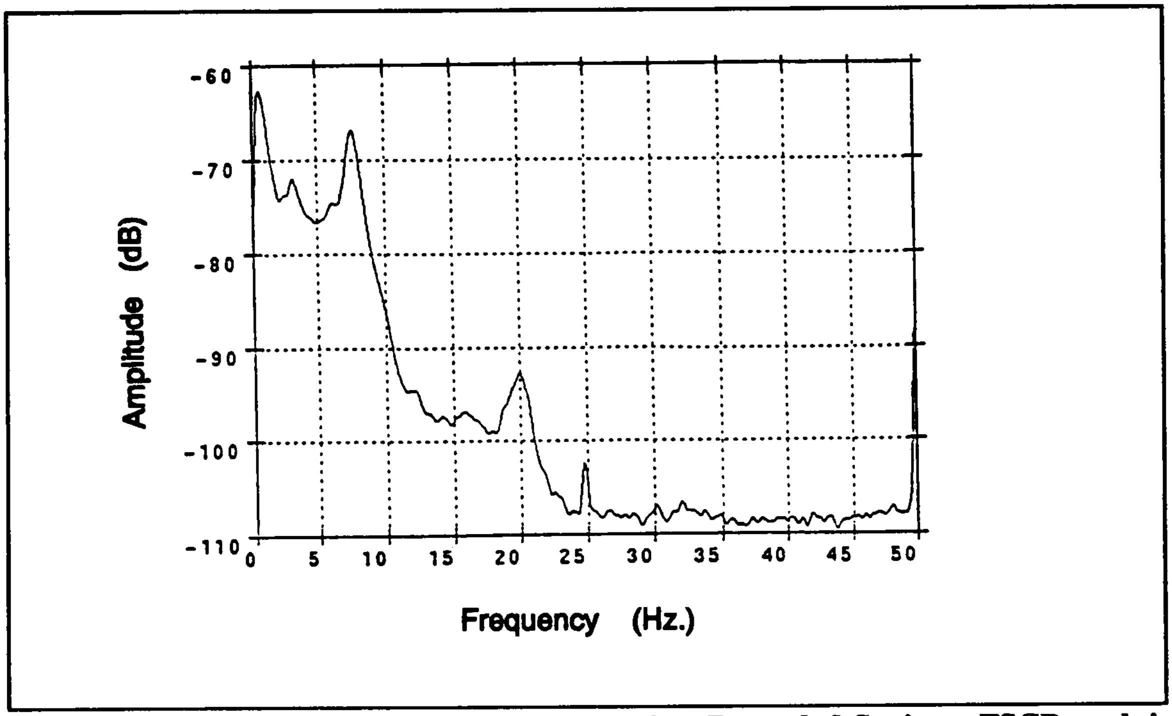

Figure II-2.14. Transition Case, from Normal to Degraded Spring. FSCB peak is displaced from 8.0 Hz to 7.75 Hz.

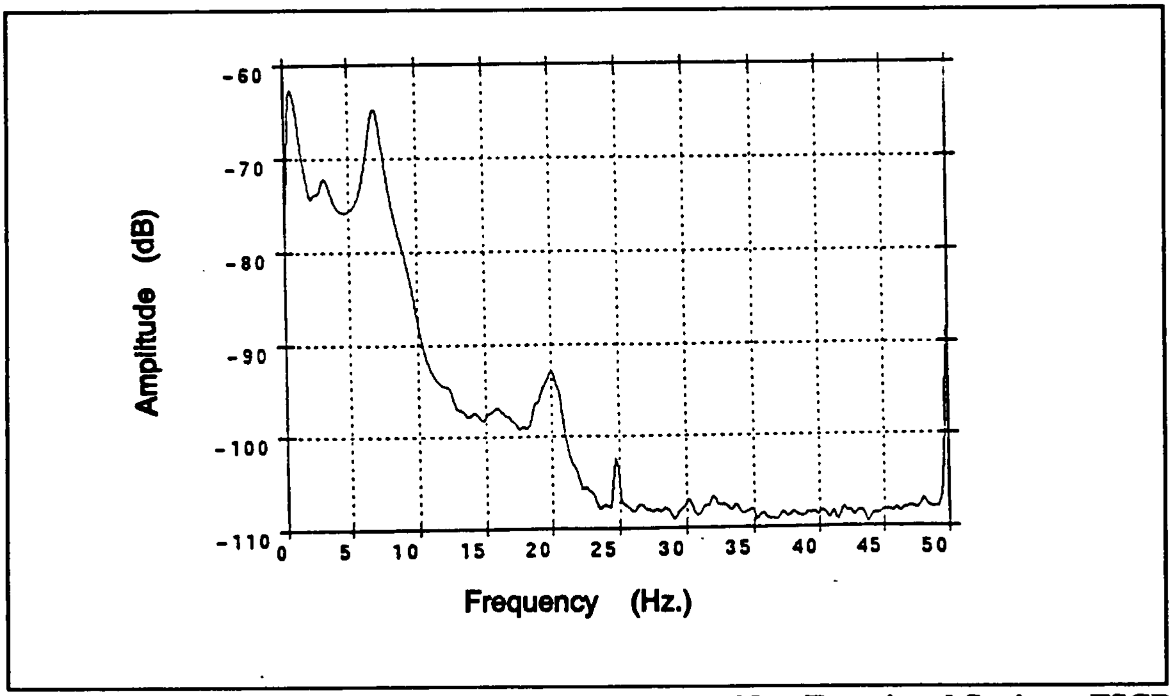

Figure II-2.15. Transition Case, from Degraded to Non-Functional Spring. FSCB peak is displaced from 8.0 Hz. to 6.875 Hz.

normal vibration modes, the objective was to design a methodology for identifying the four states from parameters extracted from the spectra. The vibration modes, according to reference [71,72], may de described as follows:

- A mode of totally free swing (Class L), where a single peak appears at 8.0 Hz. on the NPSD. Figure II-2.16 illustrates this kind of behavior.
- A mode of quasi-free swing (Class LC), where the 8.0 Hz. peak appears as does a peak of very small amplitude at around 10.0 Hz. (See Figure II-2.17). The small peak indicates light contacts between the internals and the vessel.
- A mode of vibration where two modes co-exist: a clamped mode of permanent contacts and a free mode (Figure II-2.18). This mode is denoted AL and signatures representatives of this mode have two distinctive peaks, one at 8.0 Hz. and the second at 10.0 Hz.
- A mode of intermittent contacts (I), with no significant peaks between 8.0 Hz. and 14.0 Hz. An example of this mode is shown in Figure II-2.19.

The first two categories, L and LC, represent states of "no significant contact with the vessel" [76], and in this study are considered a single class. Modes I and AL represent states where significant contacts occur and must be monitored closely. These contacts are caused by the motion of the internals during normal operations, and occur at the bottom of the vessel between the radial keys of the core barrel and corresponding supports in the pressure vessel. The clearance between the radial keys and the supports is very small, around 0.5 mm., but the impacts have a hammering effect on the structure which may produce wear of the contact surfaces, and increases the risk of structural fatigue. Figure II-2.20 shows the points of junction between the barrel and the vessel and the zone where the contacts occur.

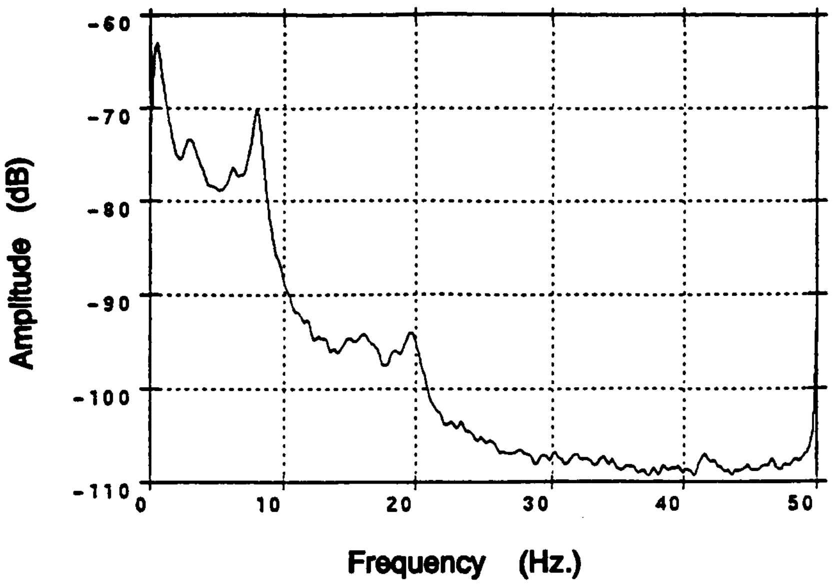

Figure II-2.16. Free Mode (L) Signature. Single peak at 8.0 Hz.

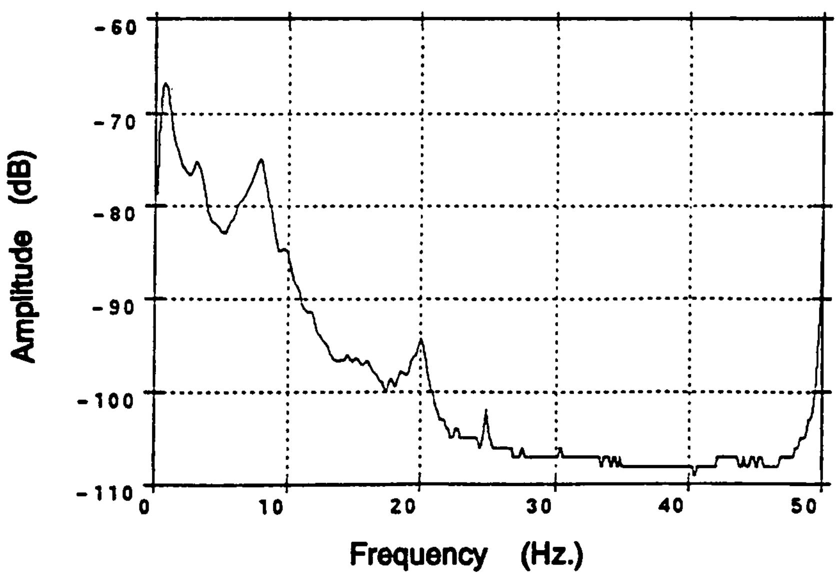

Figure II-2.17. Signature of Free Mode with Light Contacts (LC). Spectrum has a strong peak at 8.0 Hz, and small peak at 10.0 Hz.

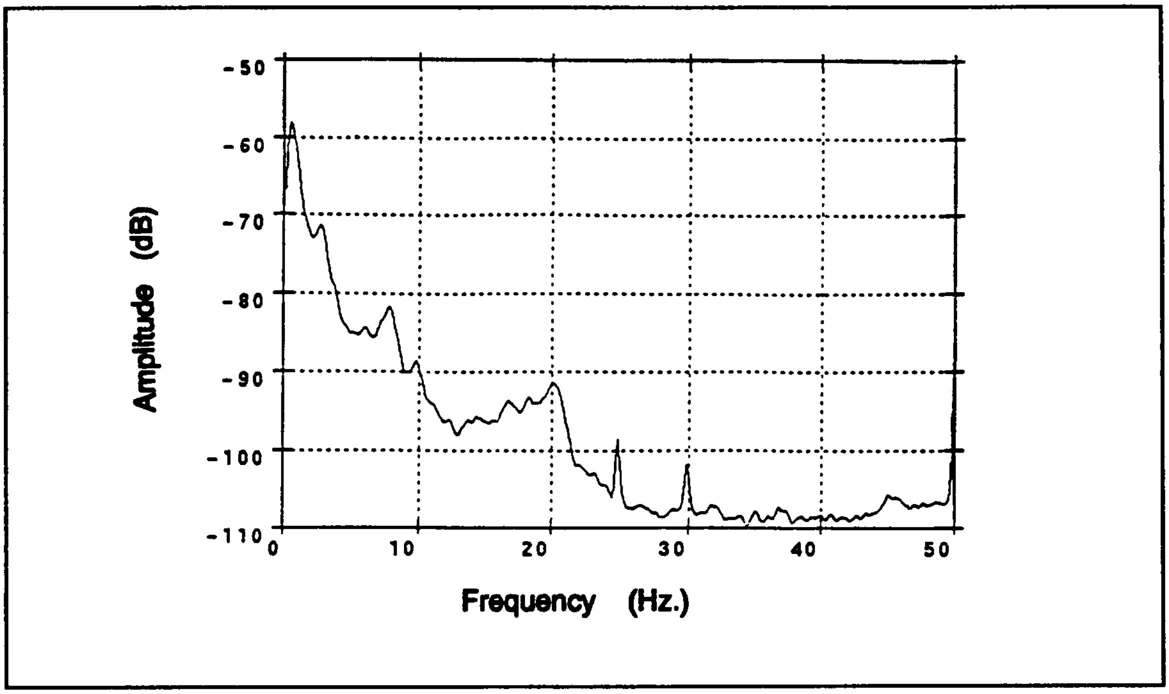

Figure II-2.18. Clamped Mode (AL). Tow modes co-exist, a mode of permanent contacts and a free mode (well defined peaks at 8.0 Hz. and 10.0 Hz.)

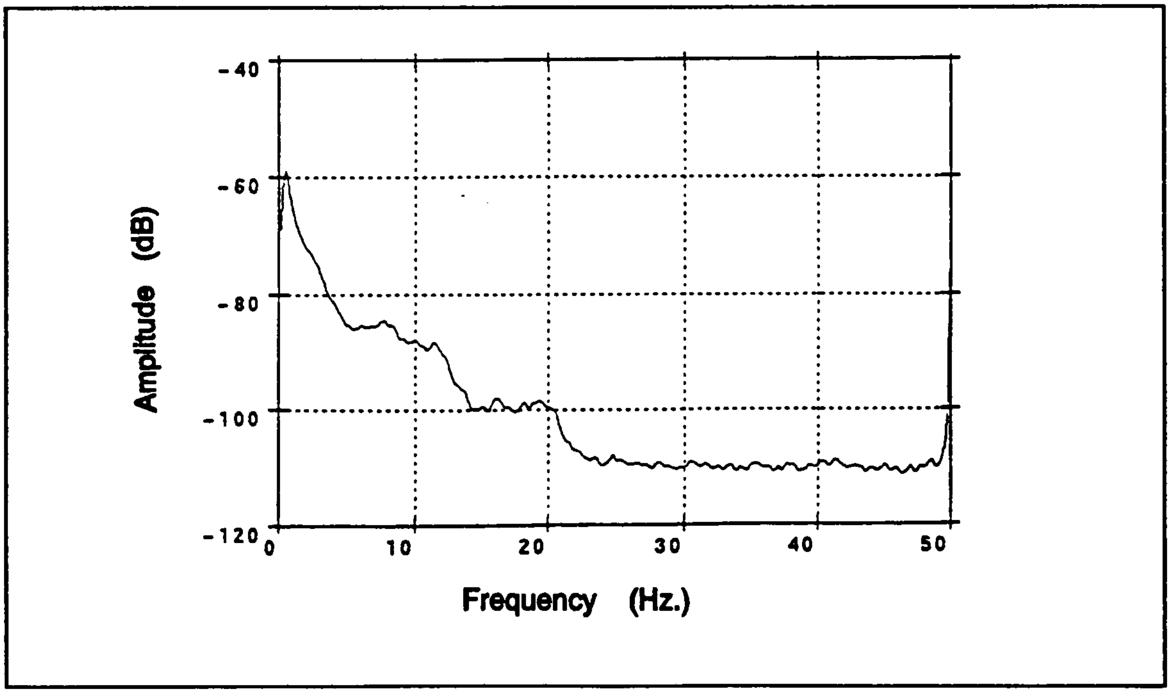

Figure II-2.19. Mode of Intermittent Contacts (I). No significant peaks between 8.0 and 14.0 Hz.

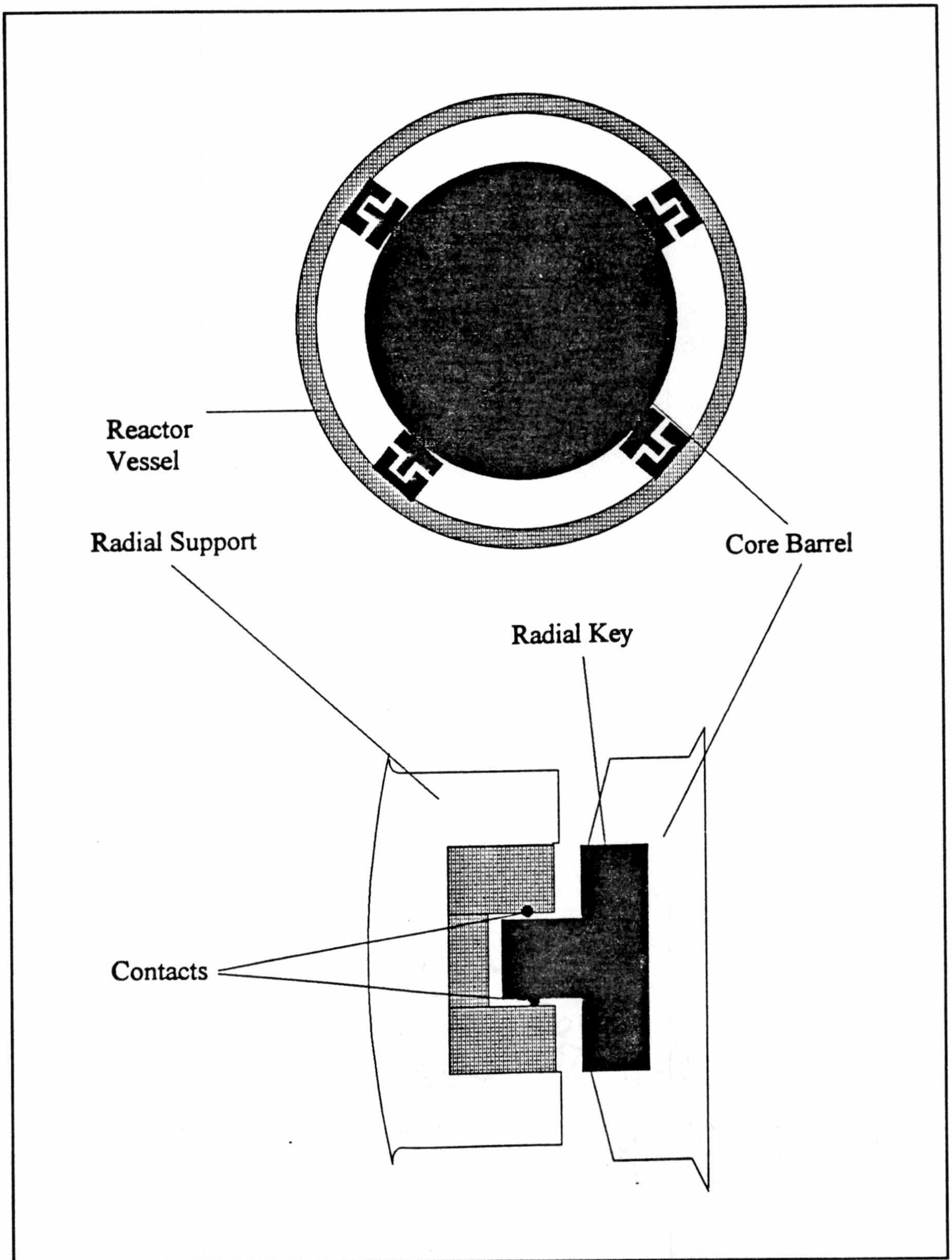

Figure II-2.20. View of Axial Keys and Supports. Sectorized thermal shield not shown.

These contacts generally appear after refuelling, when the structure is taken apart, serviced, and reassembled. Contacts tend to become less frequent as the fuel cycle progresses due to the repositioning of the internals caused by primary coolant jets and to reaction forces exerted by the reactor vessel [72]. However, the fact that the frequency of contacts diminishes as the units age, seems to indicate that the clearance between the radial keys and supports may be increasing.

System Description

A total of sixteen descriptors were chosen to perform the discrimination. All of these parameters are related to bands 13 and 14 (see Tables II-2.10 and II-2.11) of the signatures.

Table II-2.10. Discrimination Parameters Related to Band 13 Used for Module III

Variable	Description
Psd_Max_13	Amplitude of the 8.0 Hz. Peak
Height_13	Height of the 8.0 Hz. Peak
Energy_13	Energy of the 8.0 Hz. Peak
RMS_13	RMS Value of Band 13

The extraction program checks for the largest peak in the interval (7.75, 8.3 Hz). because in a few cases, the 8.0 Hz peak is slightly displaced in either direction. From Band 14, the parameters listed in Table II-2.11 are used for discrimination:

Table II-2.11. Discrimination Parameters Related to Band 14 Used for Module III

Variable	Description
Psd_Max_14	Amplitude of Most Powerful Peak in Band
Height_14	Height of Most Powerful Peak in Band
Energy_14	Energy of the Most Powerful in Band
RMS_14	RMS Value of Band

The last set of parameters is presented in Table II-2.12. The slopes are a key factor in the discrimination because they characterize the shape of the signature in the band where the contacts appear. In some cases the height of a peak that appears in the band is so small that it is not registered by the peak-detection program; and in other cases, the peak is defined by a point of inflexion and not an increase in amplitude. In both these situations, peaks in the region, essential for the detection of certain states such as Free with Light Contacts or Clamped , will be unaccounted for in the analysis. Of special interest is the peak in the close vicinity of 10.0 Hz. A small peak at 10.0 Hz. generally describes a Free-with-Light-Contacts state where there are light contacts between the core barrel and the vessel. Larger peaks at this frequency are linked to clamped states of contacts and are essential for discriminating these modes with other modes (such as intermittent contacts).

The sixteen descriptors form a feature vector which is used as the input vector for the classifier neural network. The desired output of the network is another vector which represents the mode of vibration of the core barrel according to the scheme defined in Table II-2.13.

Table II-2.12. Other Discrimination Parameters, Module III

Variable	Description
Diff	Difference of Most Powerful Peaks in Bands 13 and 14
PR_13	Relative Power of Band 13
PR_14	Relative Power of Band 14
Ratio	Ratio of RMS Values in Band 13 and 14
m_9_11	Slope of Line Joining Amplitudes in 9.0 and 11.0 Hz.
m_11_12.5	Slope of Line Joining Amplitudes in 11.0 and 12.5 Hz.
m_12_14	Slope of Line Joining Amplitudes in 12.0 and 14.0 Hz.
m_12_12.5	Slope of Line Joining Amplitudes in 12.0 and 12.5 Hz.

Table II-2.13. Codes Used in Classification*, Module III

Code	Vibration Mode
(1.0 0.0 0.0)	Free (F)
(0.0 1.0 0.0)	Intermittent (I)
(0.0 0.0 1.0)	Clamped (AL)

Data Set

A total of 1263 signatures from ex-core neutron noise detectors were considered for the analysis. The set concerns data from 28 different plants during 93 fuel cycles beginning in 1980. The data are distributed as shown in Table II-2.14.

Table II-2.14. Data Set for Module III

Number of Signatures	Vibration Mode
1070	Free (F)
82	Intermittent (I)
111	Clamped (AL)

From the set of 1263 signatures, 102 signatures were selected for training. The training signatures were representative of the four modes of core vibration, and were not selected at random but rather chosen from the set because they represented "typical" cases. The signatures were distributed as shown in Table II-2.15.

Table II-2.15. Training Set for Module III

Number of Signatures	Vibration Mode
57	Free (F)
20	Intermittent (I)
25	Clamped (AL)

Results

The classifier network was trained for 25,000 iterations using the standard implementation of the backpropagation algorithm. During training, the learning rate was decreased progressively from 1.6 to 0.05, the RMS error was reduced to 0.08 and the maximum error generated by a single pattern was 0.2. Additional training and/or subsequent adjustments of the learning rate did not improve performance. The network was tested on the entire data set of signatures. Performance was evaluated using the following criteria: A pattern representing mode i is considered correctly classified if: **a**) The activation of the output neuron i (corresponding to class i) is the largest, and **b**) if the ratio of the activation of neuron i and the largest activation is greater than 0.5. Using this criteria, a total of 1241 signatures were classified correctly (98.4%). From the correctly classified signatures, only three were classified using part **b** of the criteria for evaluation.

Three signatures --3009, 3173 and 3259-- representing free-swinging mode of the internals were classified incorrectly. For these signatures the activation of the neuron corresponding to class "L" was low (less than 0.2). One signature (2087) representing mode "I" was classified as "L", and one "AL" signature (2562) was also classified as "L". Plots of these signatures are included in Figures II-2.21 to II-2.25.

The remaining cases were classified differently by the system and the expert. Visual examination revealed strong similarities with signatures representatives of different classes. The discussion that follows is based on these "justifiable" cases. Figure II-2.26 and II-2.27 refer to a group of 4 signatures calculated from different sensors on the same reactor. The expert classified three of these signatures as "I" and one as "L". The network classified correctly the "L" signature (3898) and two of the "I" signatures (3914, 3929) and the remaining case was classified as "L" indicating that signature 3922 resembles more the normal exemplar than those representing intermittent contacts. Figure II-2.26 has a plot of the misclassified signature (3922) with the normal signature (3898), Figure II-2.27 displays Signature 3922 with one of

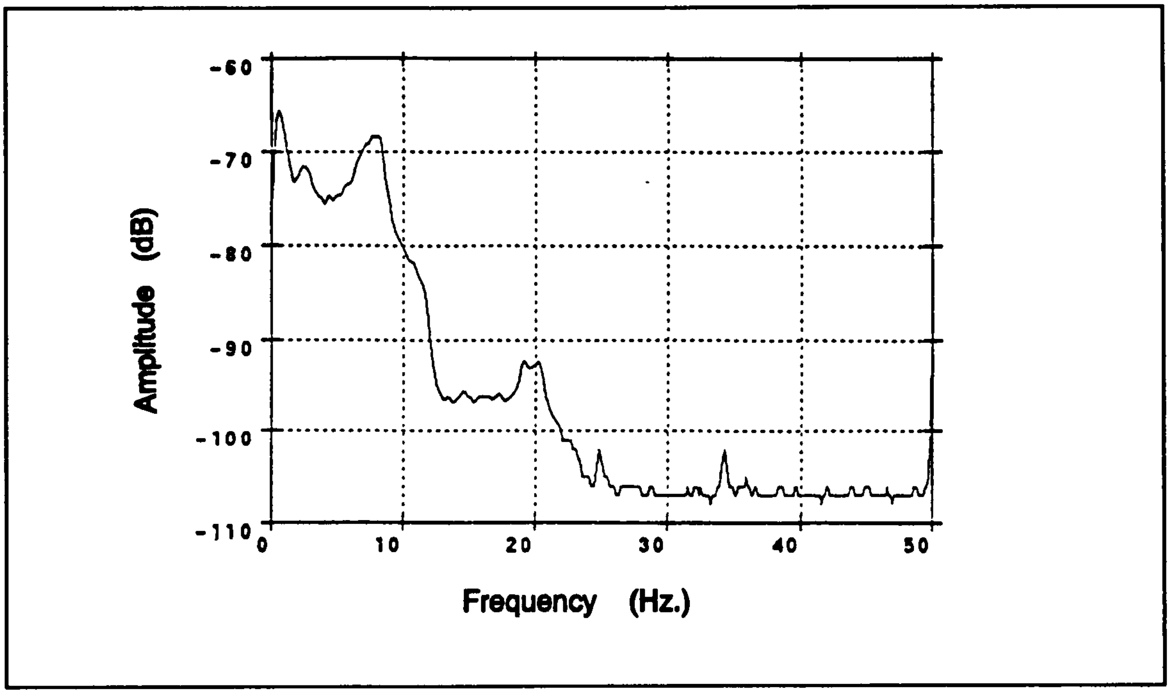

Figure II-2.21. Incorrectly Classified Free Signature (Number 3009)

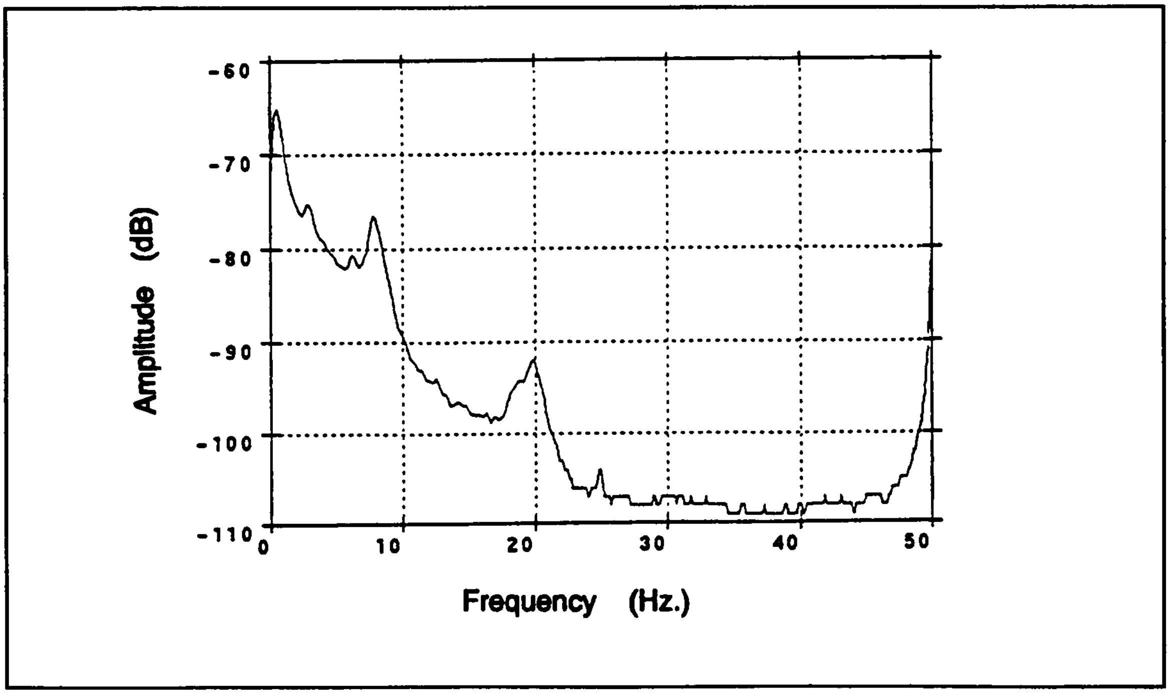

Figure II-2.22. Incorrectly Classified Free Signature (Number 3173). The low amplitude of the FSCB peak is probably the cause of misclassification.

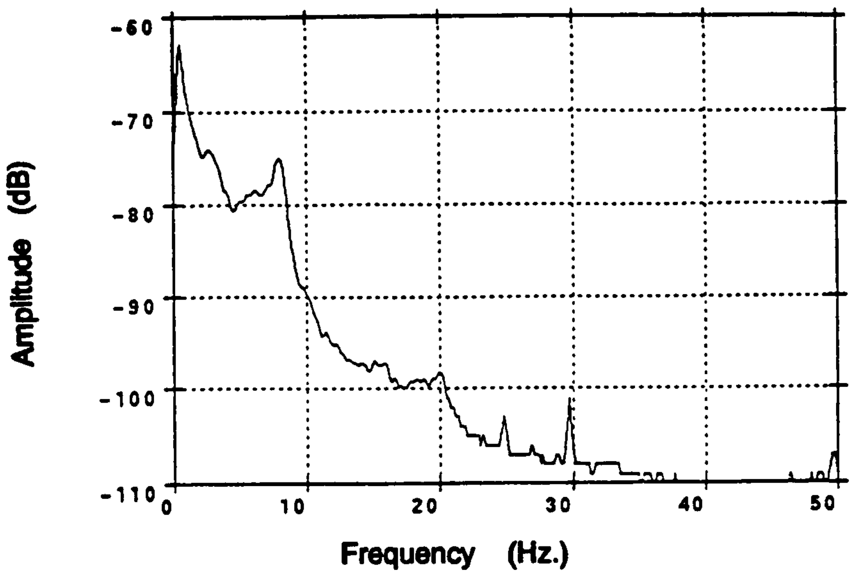

Figure II-2.23. Incorrectly Classified Free Signature (Number 3259). The low amplitude of the FSCB peak is probably the cause of misclassification.

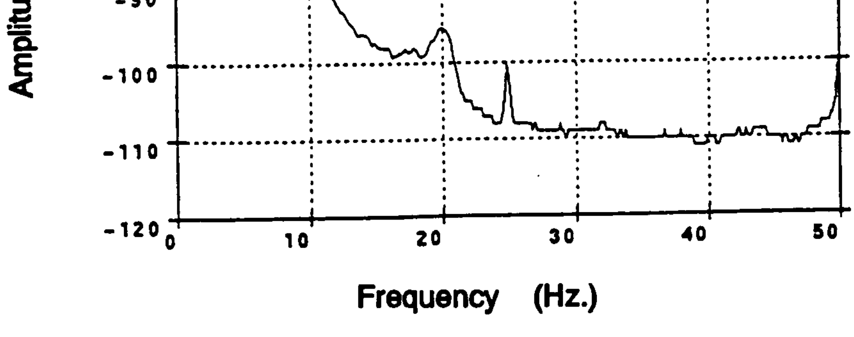

Figure II-2.24. Intermittent Signature Incorrectly Classified as Free (Number 2087)

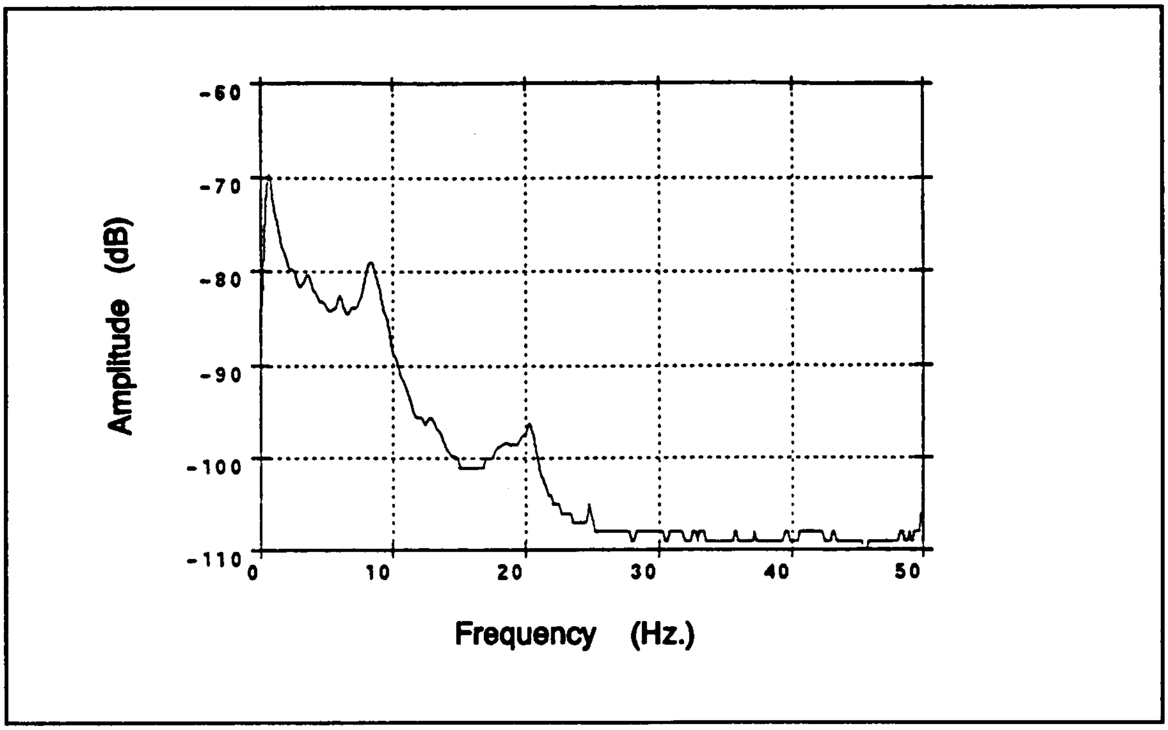

Figure II-2.25. Clamped Contact Signature Incorrectly Classified as Free (Number 2562)

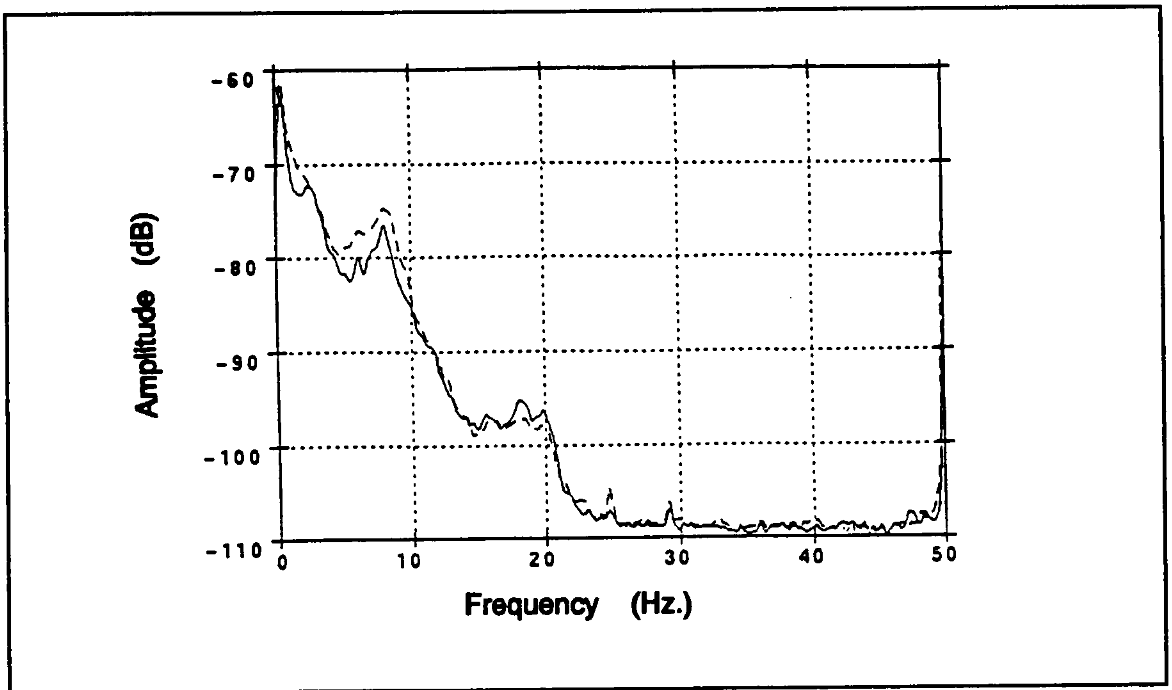

Figure II-2.26. Intermittent Signature (Number 3898, —) versus Normal Signature (Number 3922, - - -)

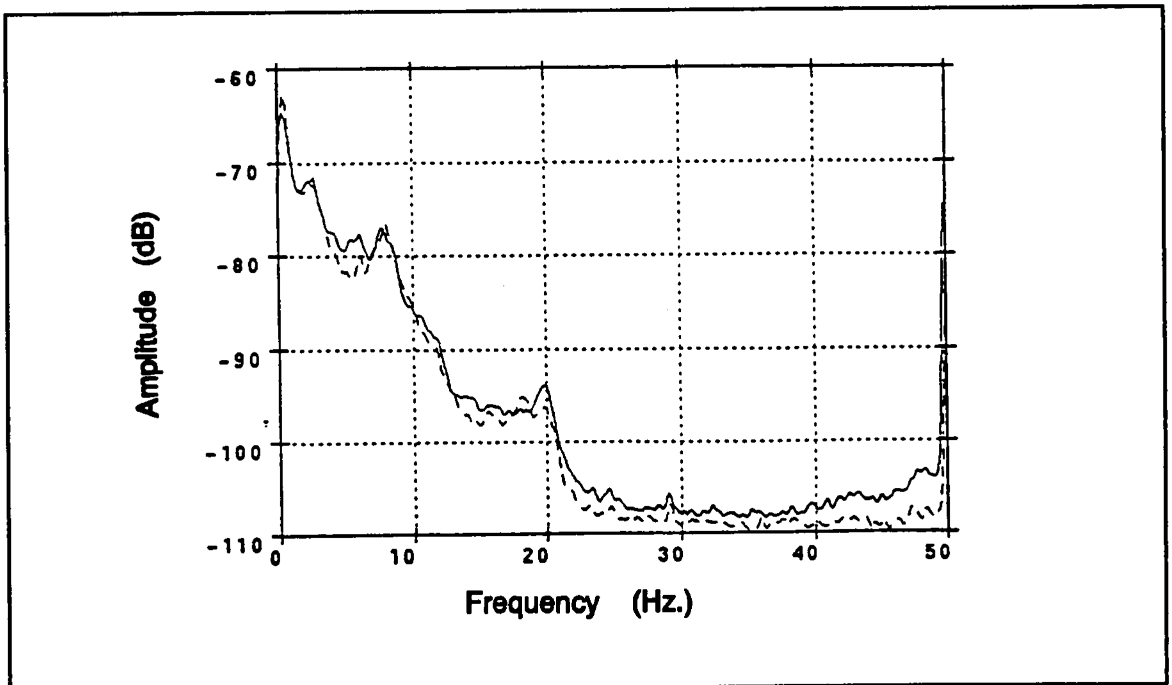

Figure II-2.27. Intermittent Signature (Number 3922, - - -) Versus Intermittent Signature (Number 3914, —)

the "I" cases (3914). The case represents an intermediate state where a definite discrimination cannot be made.

Figure II-2.28 corresponds to signatures 3786, 8790 which are also examples of intermediate cases. These two signatures are classified as "AL" by the expert and as "I" by the neural network. The irregular shape of the spectra is probably the cause for the error. Signatures representing intermittent contacts tend to be irregular, with many peaks at different frequencies and small peaks at 8.0 Hz. Similar situations were observed for Signatures 2065 (Figure II-2.29) and 2077-2078 (Figures II-2.30 and II-2.31).

Another set of interesting cases was formed by signatures from the same reactor and measurement point which were collected one or two months apart and represented different modes of internals vibration. The set (2271,2272)-(2273,2274) in Figures II-2.32 and II-2.33 represents a transition from an "I" state to "AL". As may be observed from the graph the shape of the spectra are very similar, all of the signatures were classified by the network as "AL" although the activation of the "I" neuron for the intermittent cases is higher than for the clamped cases. Similar cases are presented in Figures II-2.34 and II-2.35 for signatures representing transitions from "L" to "AL" and Figures II-2.36 and II-2-37 for a transition from "I" to "AL". Another transition is presented in Figures II-2.38 and II-2.39, in this case the mode of vibration evolves from a free mode to a mode with intermittent contacts.

The last two cases correspond to signatures representing free mode which were classified as "AL". The first signature is 2203 (Figure II-2.40). For this case,

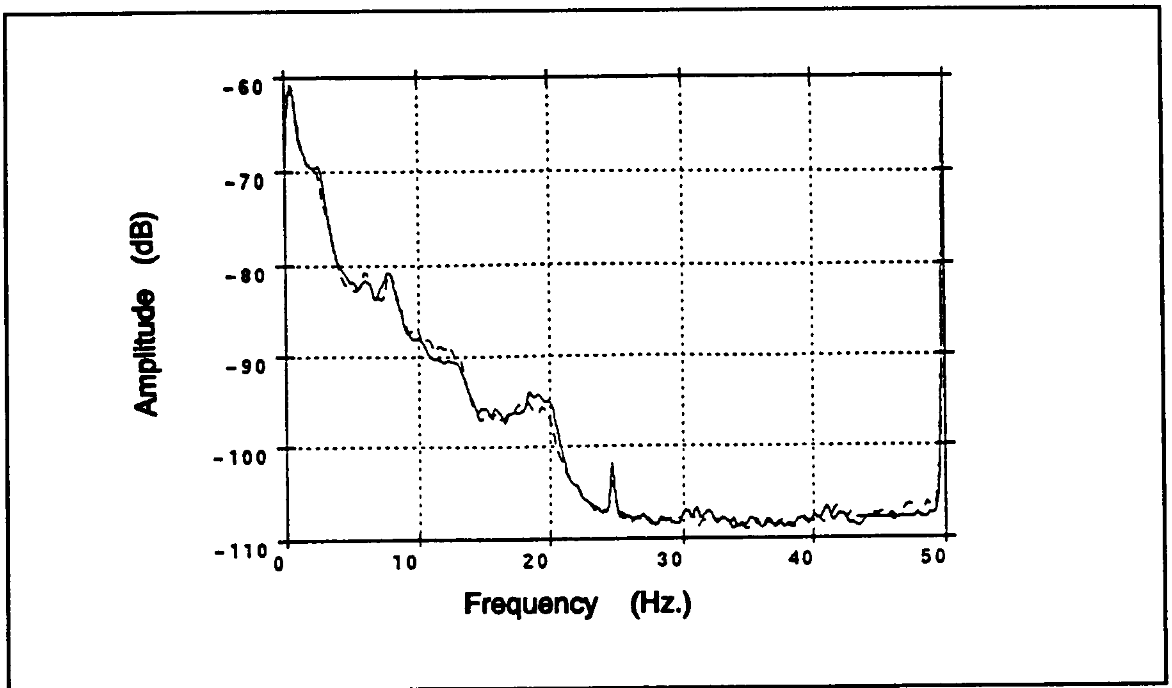

Figure II-2.28. Example of Two Clamped Mode Signatures Classified by the Network as Intermittent. (— Signature 3786, - - - Signature 3790)

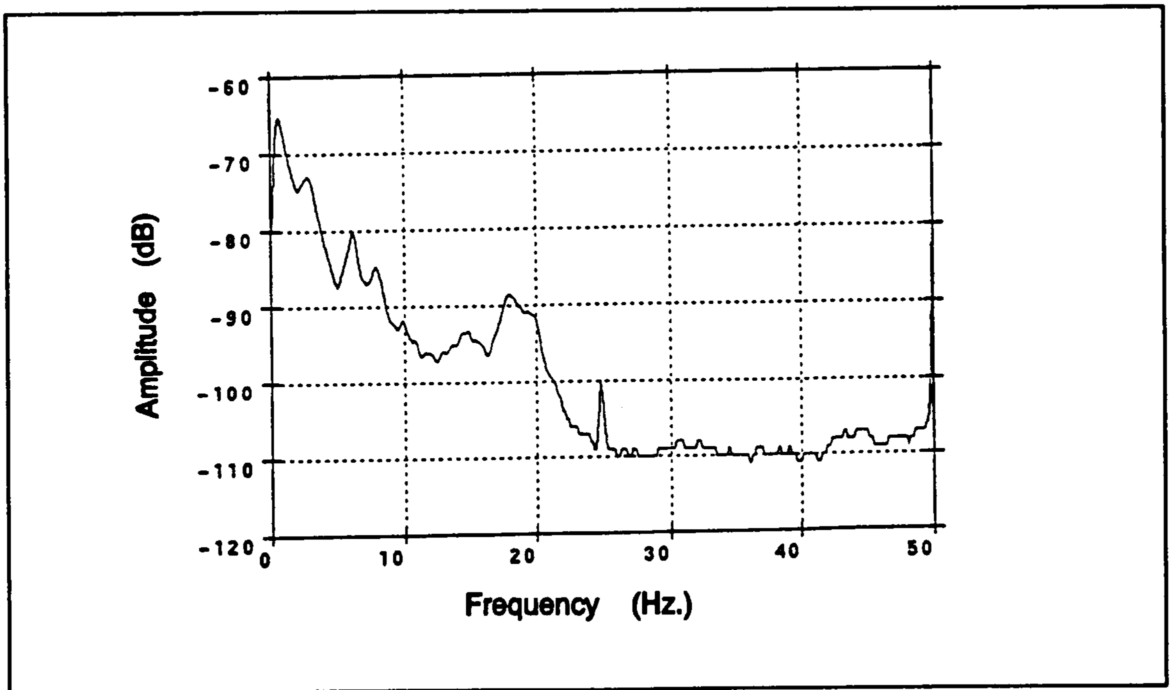

Figure II-2.29. Signature 2065. Clamped mode signature classified as Intermittent by the Classifier Network

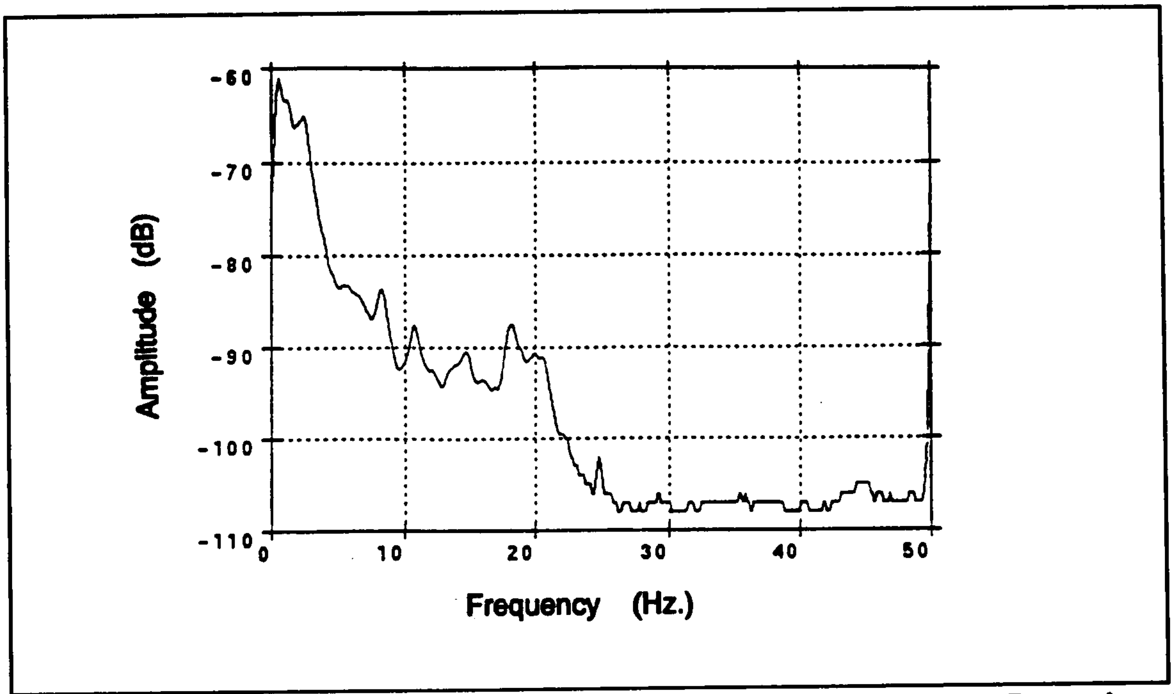

Figure II-2.30. Signature 2077. Clamped mode signature classified as Intermittent by the Classifier network

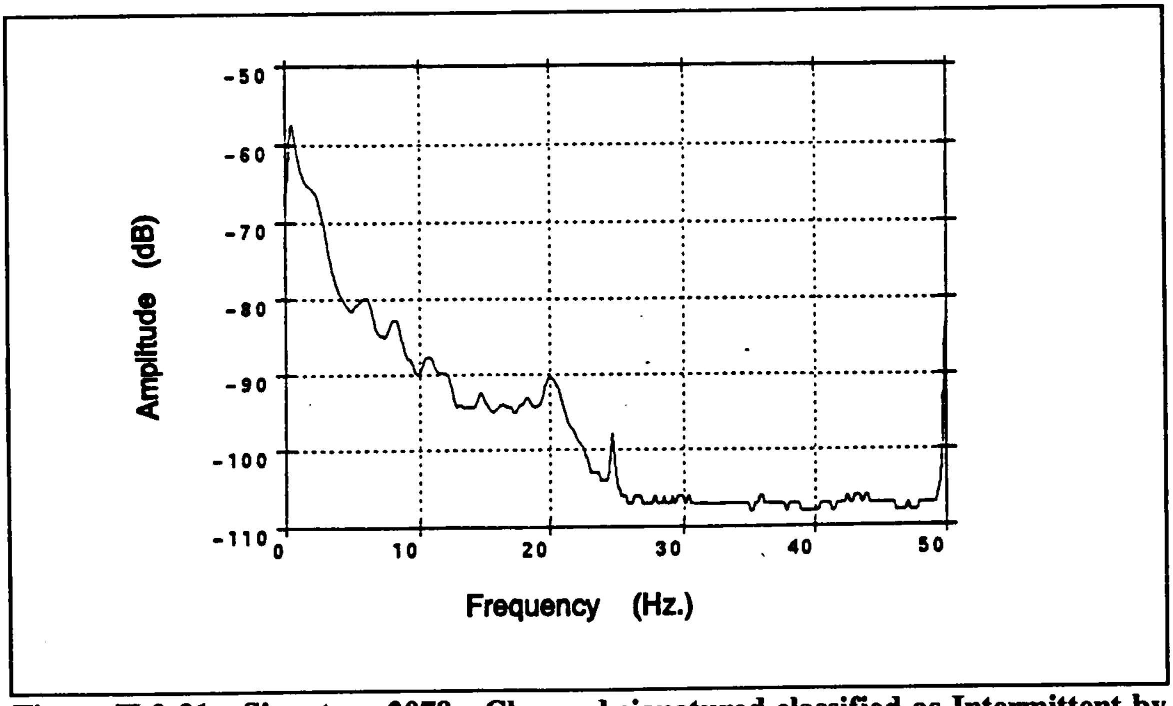

Figure II-2.31. Signature 2078. Clamped signatured classified as Intermittent by the classifier network

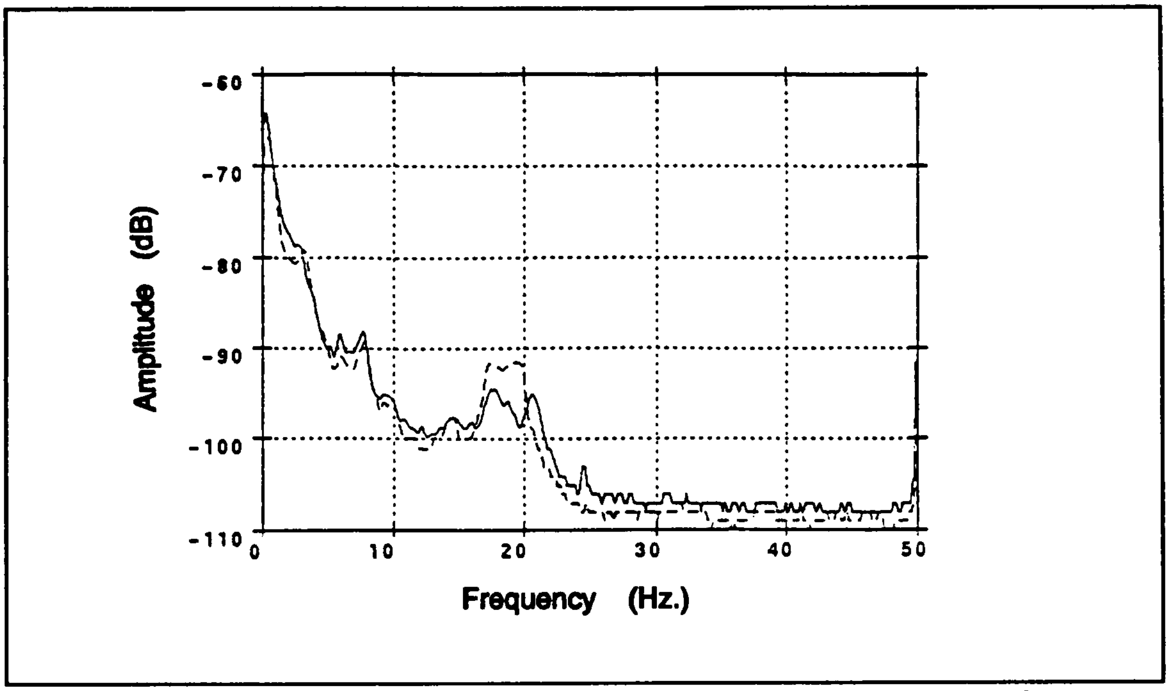

Figure II-2.32. Two Intermittent Signatures (2271, 2272) Classified as Clamped. Notice the similarity with Clamped mode signatures in Figure II-2.33

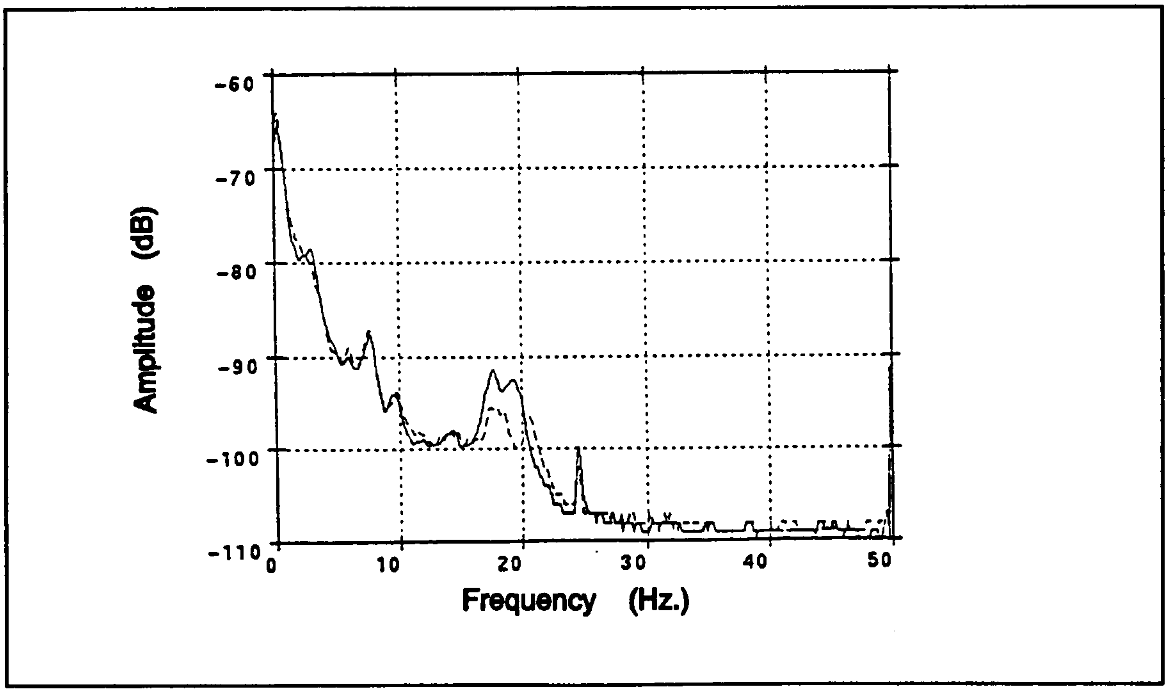

Figure II-2.33. Clamped Mode Signatures (2273, 2274). Notice similarity with Intermittent signatures in Figure II-2.32. Measurements collected two months apart

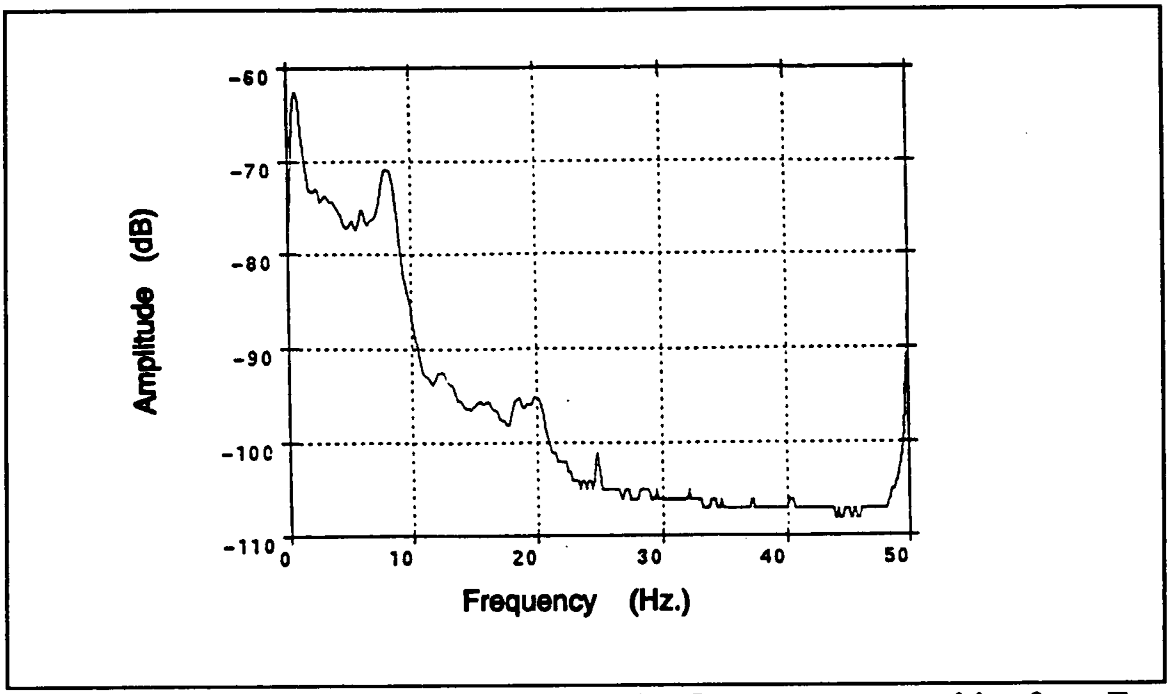

Figure II-2.34. Free Signature (Number 2582). Represents a transition from Free mode to the Clamped mode state of Figure II-2.35

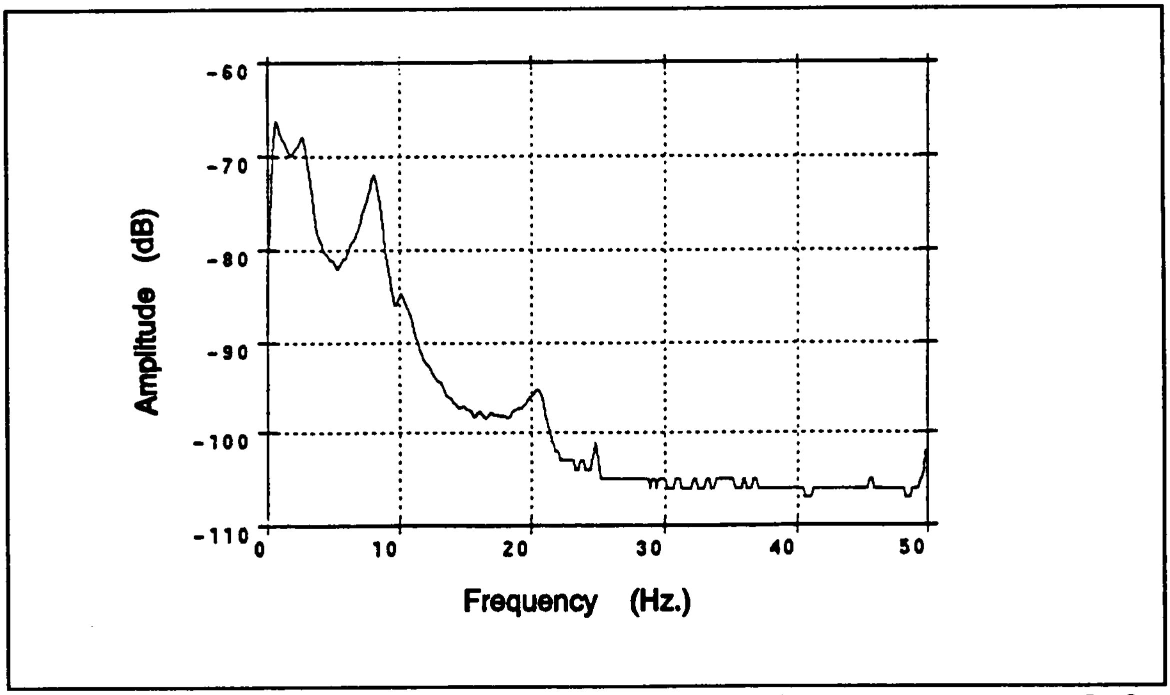

Figure II-2.35. Clamped Mode Signature (Number 2585). Represents end of transition from Free mode of Figure II-2.34 to AL

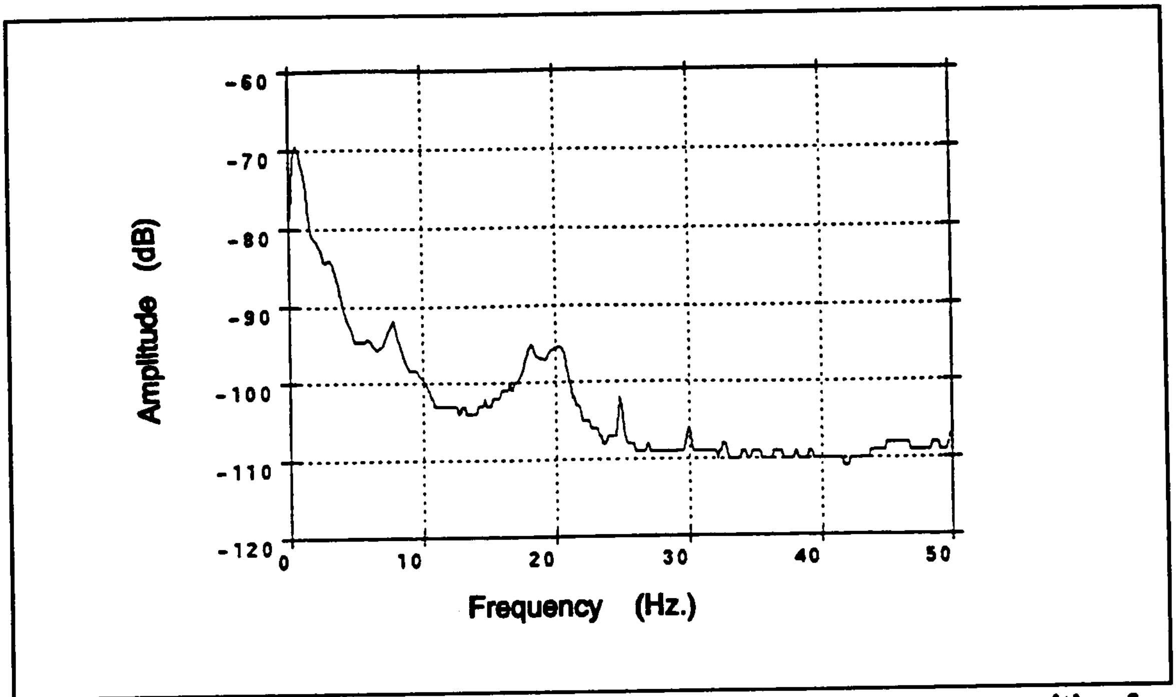

Figure II-2.36. Intermittent Signature (Number 3334). Represents a transition from Intermittent mode to the Clamped mode of Figure II-2.37

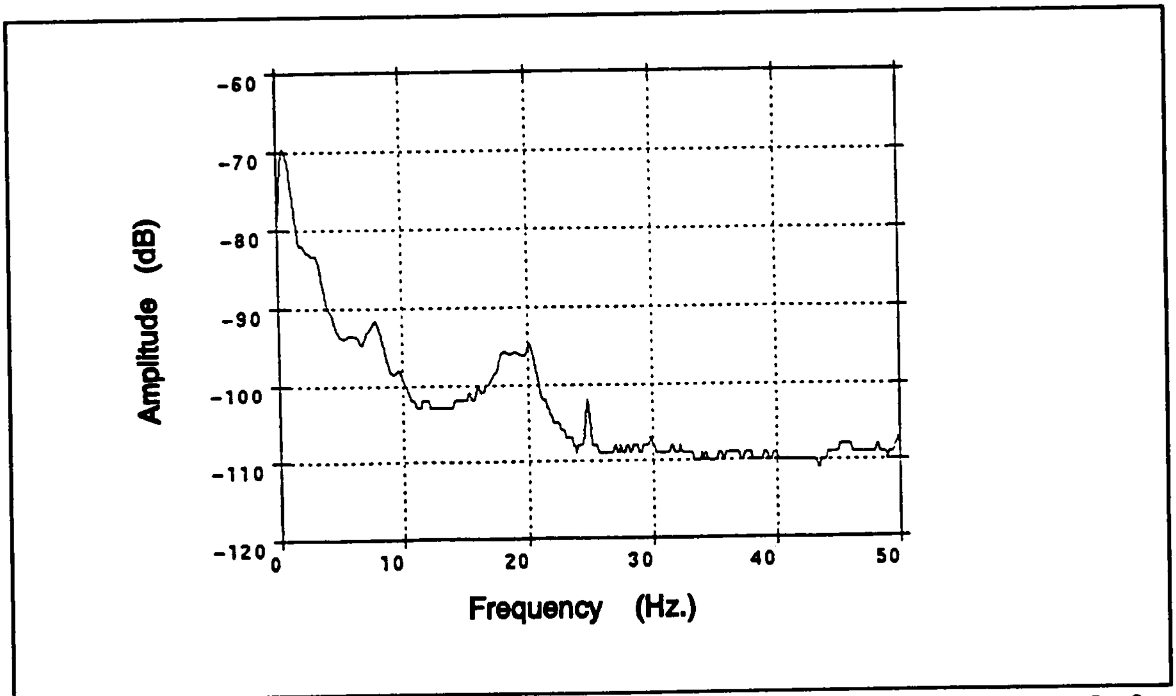

Figure II-2.37. Clamped Mode Signature (Number 3336). Represents end of transition from Intermittent mode of Figure II-2.36 to AL

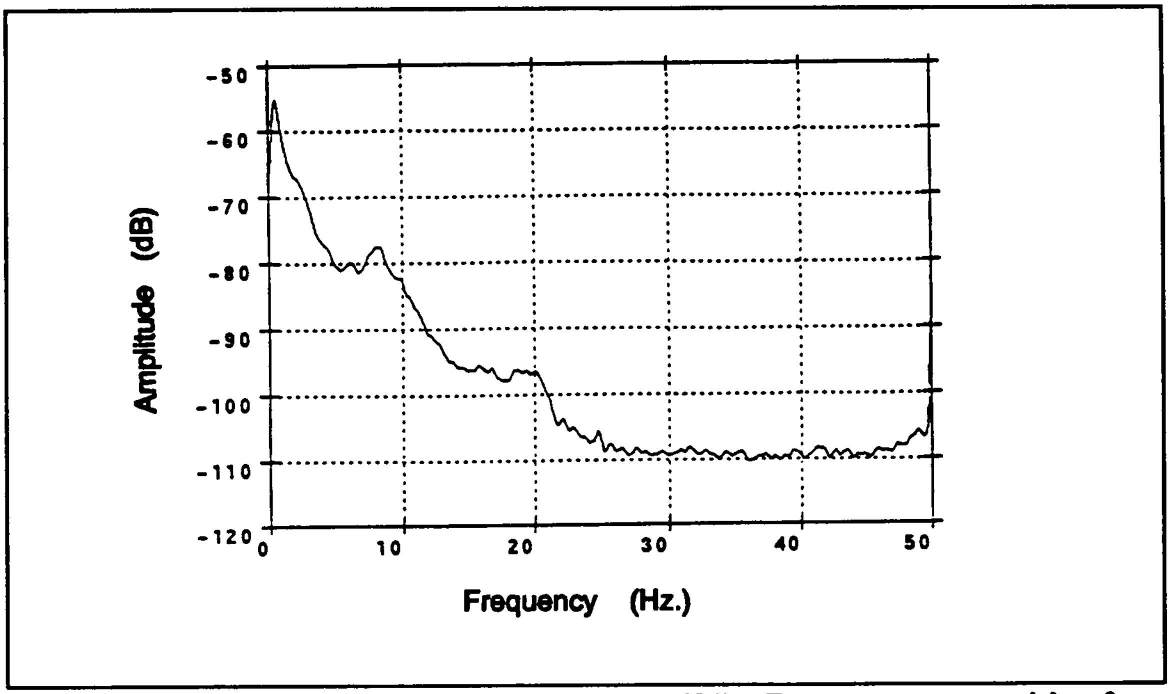

Figure II-2.38. Free Mode Signature (Number 481). Represents a transition from Free mode to the mode of Intermittent contacts shown in Figure II-2.39

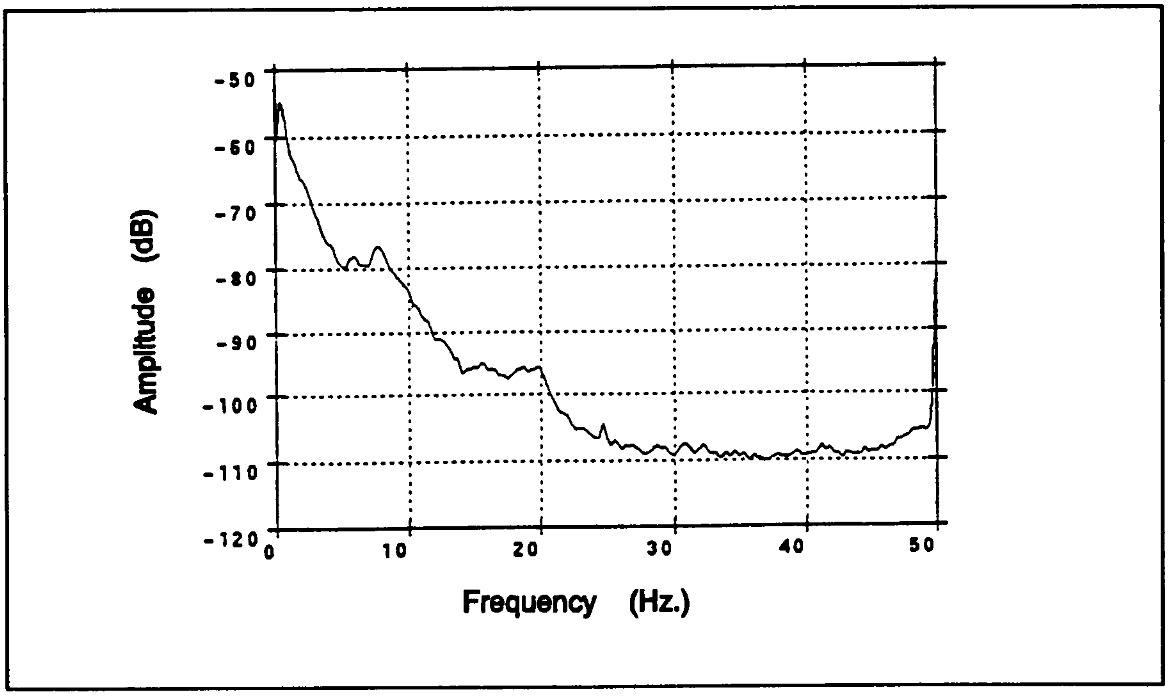

Figure II-2.39. Intermittent Mode Signature (Number 482). Represents end of transition from the Free mode of Figure II-2.38 to Intermittent mode

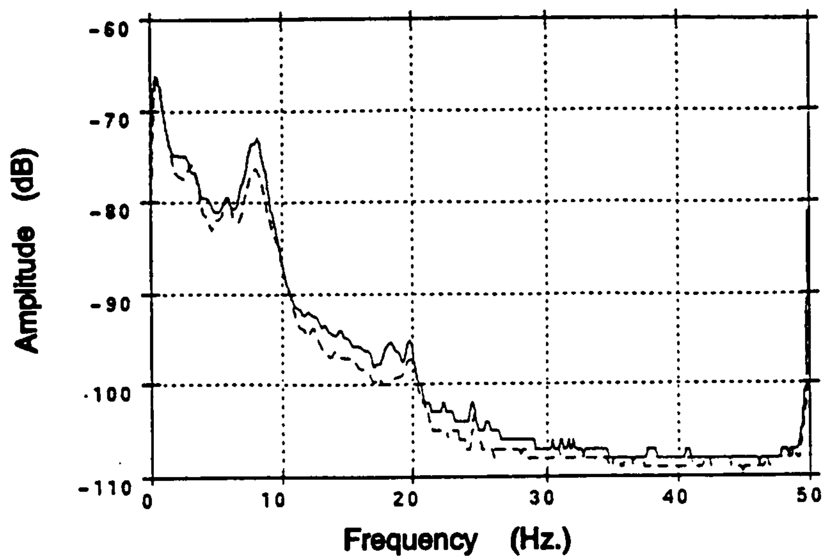

Figure II-2.40. Free Signature (Number 2203, - - -) with Low FSCB Peak. Notice peak of small amplitude at 12.5 Hz. Free Signature (Number 2204, —) also included

the peak at 8.0 Hz. is not as high as for most of the "L" signatures and it has a well-defined peak at 12.5 Hz; note that the matching signature (2204) --also in the graph -- has not been misclassified by the system. The situation reoccurs for the pair 3205-3204; for this pair signature 3205 (which also depicts a low 8.0 Hz. amplitude and peak at 12.5) is classified as clamped and the matching signature is correctly classified as free.

The confusion matrix for the results described is presented in Table II-2.16. To calculate the statistics presented on the table, all cases classified differently by the network and the expert were considered misclassified.

Table II-2.16. Confusion Matrix for Module III

Result R. Class	Free (F)	Intermittent (I)	Clamped (AL)	# Test Patterns	% Mis-classified
Free (F)	--	3	6	1013	0.8 %
Intermittent (I)	2	--	4	62	9.67 %
Clamped (AL)	1	3	--	86	4.65 %

In general, 1243 of the 1263 signatures (98.4%) were classified correctly by the neural network and 15 signatures(1.2%) were classified differently by the expert and the neural network; these cases, which we call "justifiable," present features which are usually associated with classes other than the one assigned by the expert, or represent a transition from one mode of internal vibrations to another. Finally, 5 cases were classified incorrectly.

Other Results

An effort was made to integrate Modules II and III of the system into a single module. The system considered twenty four parameters for the discrimination, essentially all of the parameters in the two modules combined. The results were poor, around 83 % correct classification, when compared to the performance of the discrimination in two stages. These results appear to validate the general belief that when neural networks are used for parametric classification, it is best to consider a modular design. Generally, the performance of the neural network may be improved by breaking the problem into stages, and having different simple neural networks solve different parts of the problem.

DISCUSSION

A technique based on neural networks is proposed for the classification of neutron-noise signatures representing the state of internals vibration in pressurized water reactors. The classifier networks are trained using the standard backpropagation algorithm to discriminate among the different modes of vibration exhibited by PWR's, and to identify states where degradation of mechanical components have occurred. For the discrimination, features are extracted from normalized power spectral densities which describe the different modes of vibration. Among these features: the energy, amplitude and height of relevant peaks in the spectra, and slopes which describe the shape of the signatures in critical regions.

The discrimination system processes the data in three phases. In the first stage, signatures from functional power detectors are separated from signatures from improperly calibrated or out-of-order detectors: The second module is used to detect degradation of the hold-down spring, and the third module to classify signatures according to the mode of vibration of the reactor. The system configuration is not unique since in practice, the processing of modules II and III is independent and that order could have been interchanged.

The networks were trained using only 8% of the data made available for the study, and were collectively able to discriminate correctly over 98% of the signatures. Some interesting cases had been identified in which the classification provided by the expert and the network differ but may be justified, and in a few cases, the class label provided by the network is not correct. Data from 28 different reactors at different stages of the fuel cycle was used to test the system.

The results demonstrate the ability of neural networks with very simple topologies to perform the classification of ex-core neutron spectra, and indicate that the criterion for selecting hidden layer size proposed by Kung and Hwang [35] works very well for this type of problems. In addition, the improvement in accuracy gained by modularizing the problem confirms the general belief that having several neural networks solving separate parts of a problem is more useful than a large neural network handling all aspects of the classification.

The parameters selected for the classification are good descriptors of the bands and peaks in the signature. The slopes are very helpful in describing specific regions in detail, since there are cases were the peaks are very small and are not registered by the software. Other parameters such as ratios and relative powers are useful in describing the relationship between the contents of different bands.

The problem of identifying signatures from out-of-order sensors is particularly difficult because not only are the shapes very distinct among the signatures, but also differ greatly from typical cases. Also, in the region of the spectrum which provides information about the sensor background (Band 17) the amplitude levels may be similar to those for normal signatures. These peculiarities make it difficult to define descriptors which represent the class in a general manner and therefore to train a neural system to identify the cases.

It is necessary for the identification of improperly calibrated or out-of-order sensors, to compare test signatures with a reference spectrum, and to use the discrepancy as a basis for decision making. The reference spectrum may be generated using neural networks trained with data from the sensor where the test data were collected, or from other sensors mounted on the structure which are able to detect the same vibration.

After examining the data provided for this project, it is apparent that the shapes of spectra collected from neighboring measuring points of the same reactor at the same date are very similar. From the theoretical standpoint, comparisons between sensors which are 180 degrees apart is very meaningful because the coherence function is very high for those sensors in the regions of interest. Even for sensors which are only 90 degrees apart, the discrepancy introduced by a bad sensor can be easily perceived because the shapes of the two curves are radically different. In the following section a methodology is proposed for modelling the relationship among sensors in the structure, with this technique it is possible to detect large deviations from typical sensor outputs, making it simple to detect out-of-order detectors.

CHAPTER 3

MONITORING SYSTEM

The main objectives of the vibration monitoring efforts in PWR's is the detection of contacts between the core barrel and the reactor vessel, and the malfunctioning of components in the structure. In this section, a neural-networks based technique is proposed for the monitoring of reactors internals which relies on a model of the relationship among ex-core detectors, during modes of no significant contacts. The trained neural network is able to make predictions of this behavior, and allows dependable automatic identification of deviations.

The system described is based on the ideas presented in Chapter I-2 of this dissertation adapted to the realm of reactor internals monitoring. To implement the model, a neural network is trained to mimic the relationship between two sensors in the structure. The input to the network is spectra from one detectors, and its output spectra from a detector located at an angle of 180 deg. Opposite ion chambers are chosen for the modelling because these detectors usually exhibit high coherences in the bands of interest as it refers to contact detection and hold-down spring degradation.

Deviations from contact free modes are identifiable by comparing the prediction of the modelling network with the spectra calculated from the actual outputs of the detectors. The deviations may be quantified, and because of the known relationships between regions in the spectrum and components in the structure (see Table II-1.1), can be related to particular sources of excitation.

There are two important differences between the system described in this section and the system for monitoring of rotating machinery presented in Chapter I-2, Part I. First, the neutron noise process is stationary; since the process does not vary with time, only data from only one time frame is considered for input to the modelling network. Secondly, because the spectrum may be divided into bands that relate to particular components in the structure, the prediction for each band is made based on the information contained within a range of the spectrum containing the band. For example, when making a prediction of the spectrum on one detector, the prediction corresponding to Band **a** of that detector is assemble from information contained on Band **a** (and a neighboring region) of the spectrum of the opposite detector. This approach limits the effect of localized increases within the spectrum, and is valid and very desirable since the information contained in each band may be considered independent of that of other bands.

SYSTEM DESCRIPTION

To implement the system, the multiple network architecture of Figure II-3.1 was used, information regarding the configuration of the networks is presented in Table II-3.1. Each network models the relationship of a particular region of the spectrum according to the schema defined in Table II-3.1.

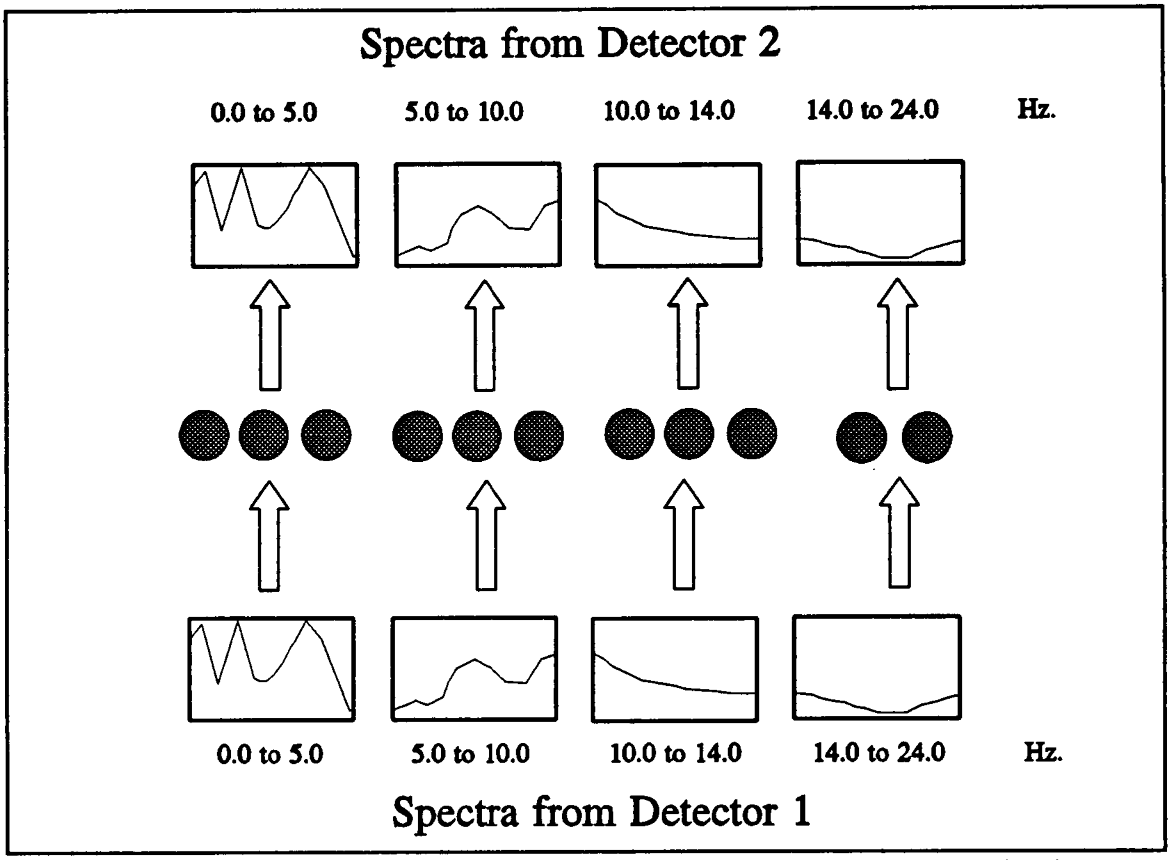

Figure II-3.1. Multi-Network System for Internal Vibration Monitoring

The networks are trained individually, and the results are assemble by collecting the contributions of each network. The modelling includes only 192 frequencies of the spectrum (0.125 to 24.0 Hz.), because beyond that range, only information regarding sensor background information and internals complex mode can be extracted, none of which is the topic of this study.

Table II-3.1. Topology of Multiple Network System

Network No.	Input Units	Hidden Units	Output Units	Frequency Range	Information
1	40	10	40	0.0 to 5.0 Hz.	Band 10, 11
2	40	10	40	5.0 to 10.0 Hz.	Band 12, 13
3	32	10	32	10.0 to 14.0 Hz.	Band 14
4	80	20	80	14.0 to 24.0 Hz.	Band 15, 16

RESULTS

A subset of 40 of the spectral signatures described in Chapter II-2 was used for training and testing the system. Pairs of signatures were chosen which corresponded to opposite chambers (180 deg. apart) on the same reactor. All signatures are recorded with identical equipment, with a four minute delay from chambers 1 and 2 of the reactors, see Figure II-3.2.

The signatures represented the four types of normal vibratory behavior of the structure. Ten pairs of signatures identified as contact-free mode were used for training, and the entire set used for recall. In order to process the data using

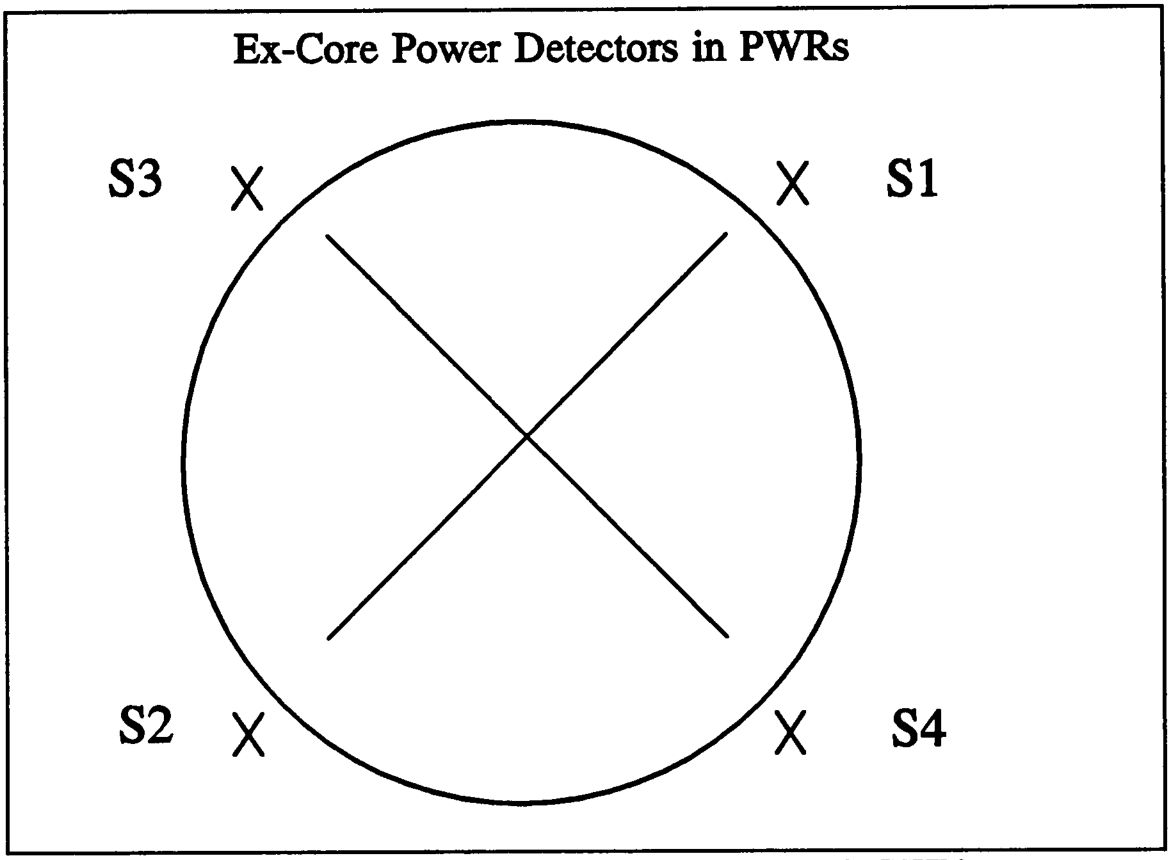

Figure II-3.2. Location of Ex-Core Detectors in PWR's

multiple networks, the signatures were disassembled into frequency ranges as described in Table II-3.1. After processing, the signatures are reassemble to their original form.

All networks were trained individually until the maximum error for each pattern was reduced to 0.01. On the average, networks were trained for 25,000 iterations using the standard Backpropagation algorithm. Different configurations for hidden layers were used, but none render better results. In general smaller hidden layers tend to increase the error, and hidden layers with more units produce comparable results while increasing computational cost.

After recall, the descriptors defined in Chapter I-2, were used to measure the discrepancy between the predicted signatures generated by the networks and the spectra calculated from the detector. The cross-sensor descriptor (Equation I-2.7) is a degradation indicator obtained by taking the square root of the squared differences at each frequency. This descriptor is useful for measuring deviations from contact-free mode.

Table II-3.2 presents the value of this descriptor for 14 of the signatures used for testing. As can be observed clearly, the descriptors for spectra representing modes of significant contacts (AL and I) are much higher in the range corresponding to Band 14 which provides information about contacts than cases with no contact or light contacts. Even in cases where there is light contacts (LC) between the core barrel and the vessel, Signature 3 for example, the value of the descriptor shows a significant increase.

Plots of the predictions of the networks versus the actual spectra calculated from the sensors for a few cases are included in Figures II-3.3 to II-3.14. The lower figure accompanying each case, represents the per-frequency descriptor also defined in Chapter I-2. This descriptor gathers the difference at individual frequencies, and it is very helpful for identifying those frequencies in the spectra where significant changes have occurred. This local deviations, may be map using the information in Table II-1.1 to the behavior of particular components of the internals.

Generally, the agreement between the spectra and the predictions is significant for cases of no contacts (L), see for example Figures II-3.3 and II-3.5. For this cases the descriptors are very low (less than 0.1). For cases of Free mode signatures with light contacts (LC), an increase in the second descriptor (contains information about Band 14 and therefore presence of contacts) is observable, see for example Figure II-3.7. The last category of cases corresponds to signatures with significant contacts. These signatures represent clamped mode as well as intermittent modes. As can be seen in Figures II-3.9, II-3.11, II-3.13 and II-3.15, these signatures deviate drastically from contact-free mode, as reflected by the high values of the descriptors

DISCUSSION

A methodology has been presented for the monitoring of reactor internals in pressurized water reactors. The technique uses multiple neural networks to model the relationship between opposite ion chambers in the core during modes of no contact. The prediction of the model is later compared to the spectra calculated

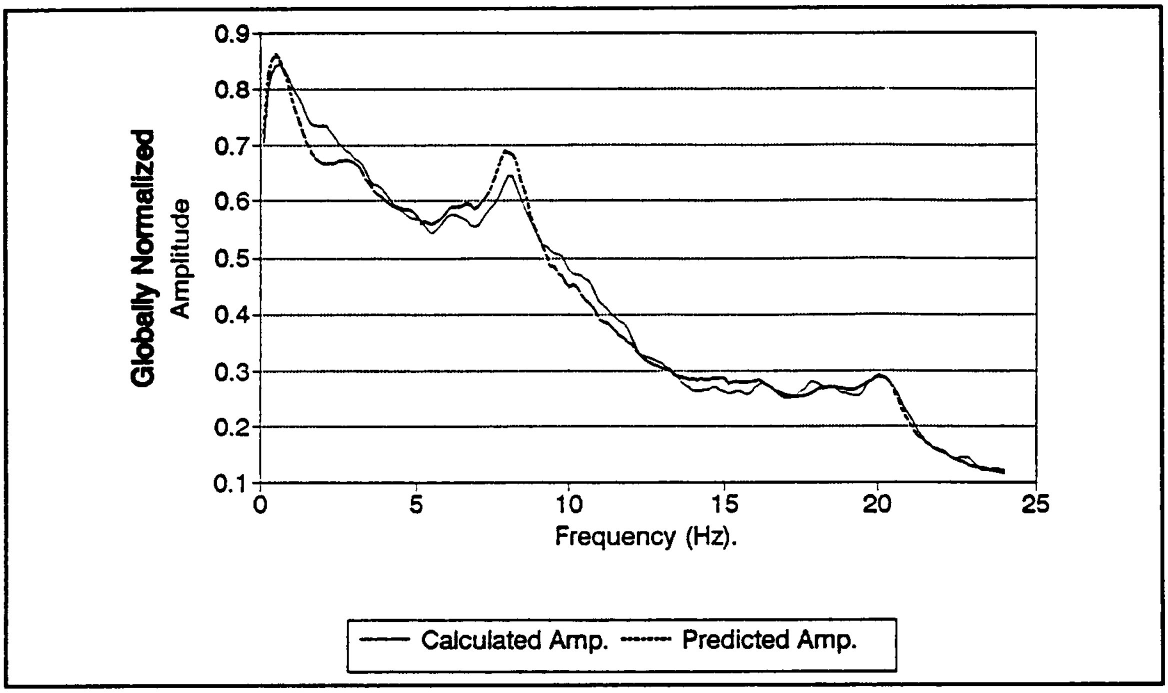

Figure II-3.3. Example of a Free Signature (Number 1). Low descriptors reflect the agreement between the network prediction and actual values of the spectra

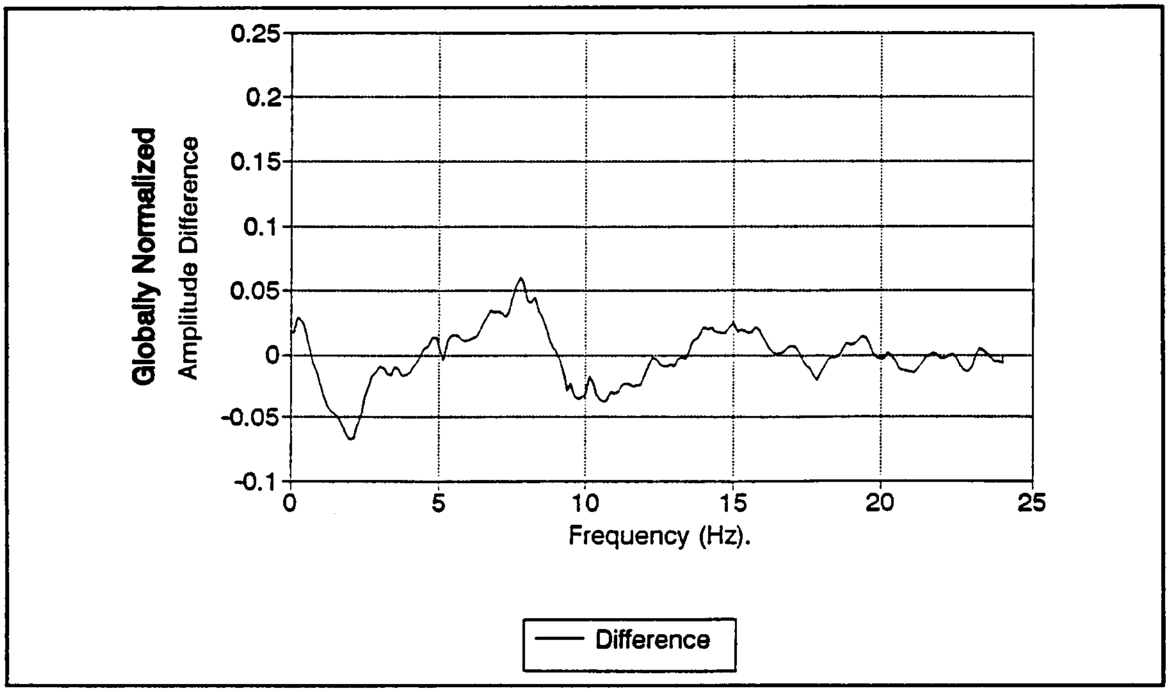

Figure II-3.4. Per-Frequency Descriptor, Signature 1. Difference between prediction and actual values is small, reflecting no deviations from contact-free mode

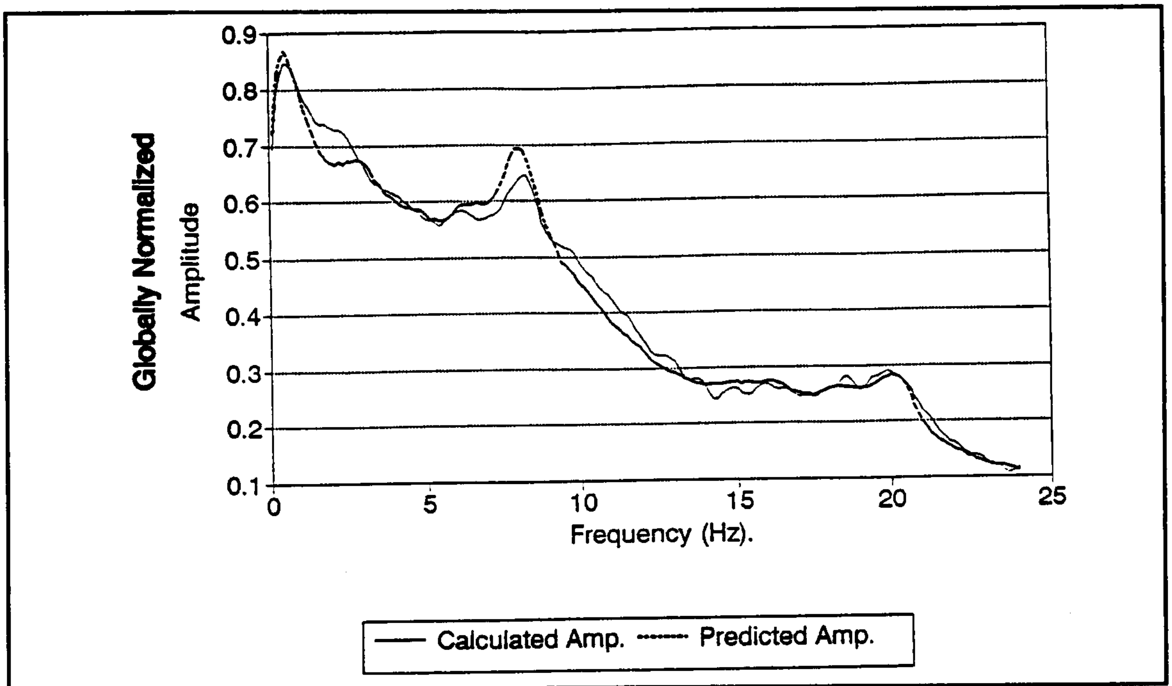

Figure II-3.5. Example of a Free Signature, Signature 2. Low descriptors reflect the agreement between the network prediction and actual values of the spectra

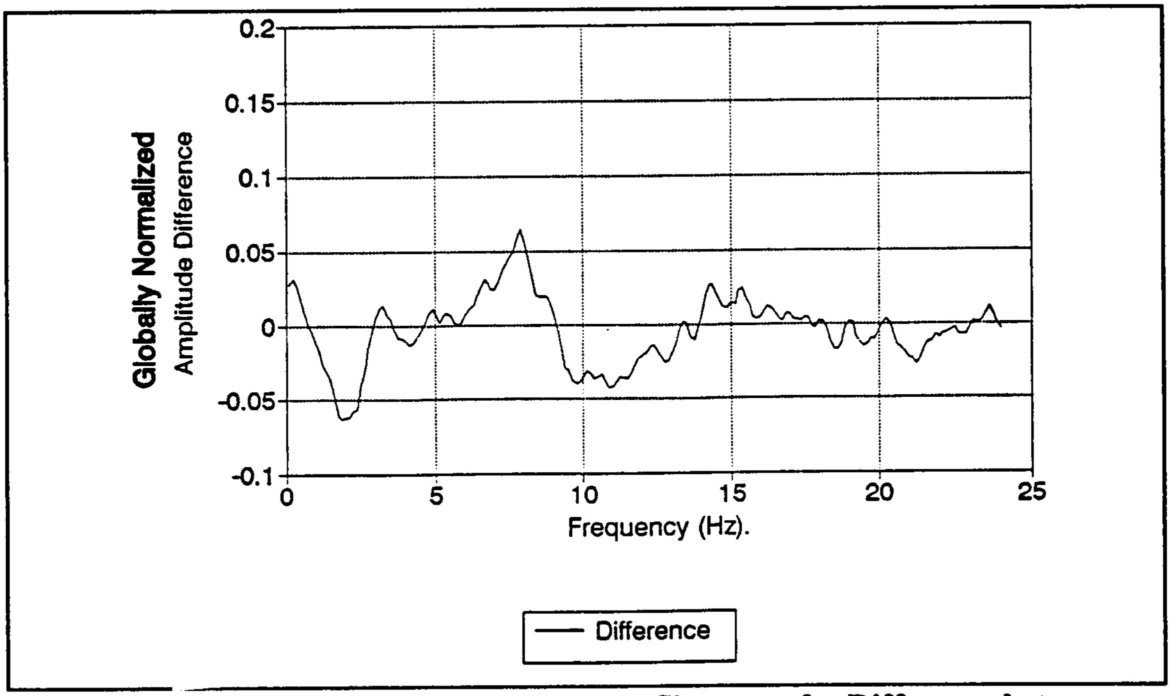

Figure II-3.6. Per-Frequency Descriptor, Signature 2. Difference between prediction and actual values is small, reflecting no deviations from contact-free mode

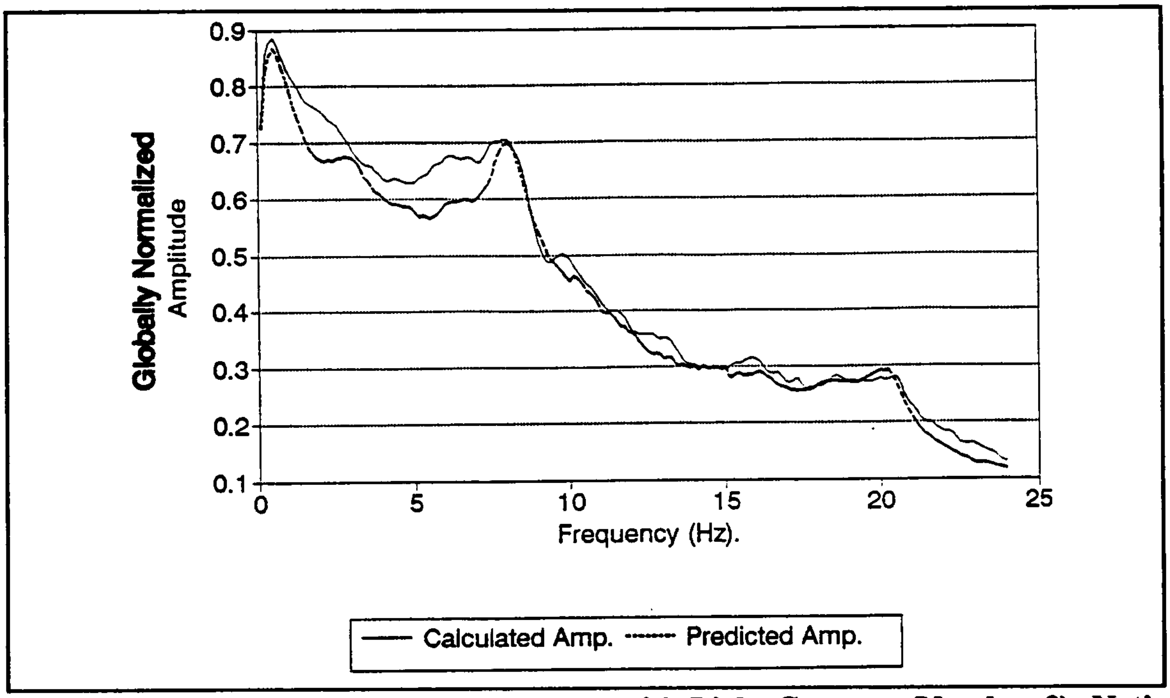

Figure II-3.7. Example of a Free Signature with Light Contacts (Number 3). Notice the increase in Descriptor 2 corresponding to the bands where contacts appear

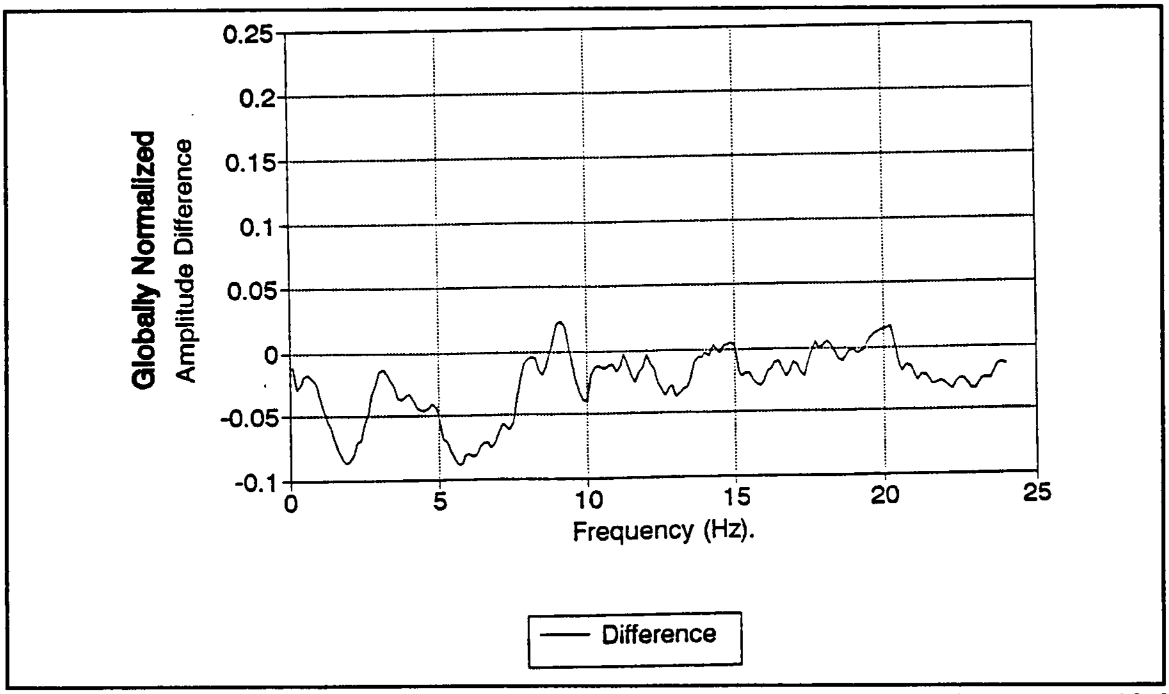

Figure II-3.8. Per-Frequency Descriptor, Signature 3. The sharp decrease at 10.0 Hz. indicates the appearance of the contact

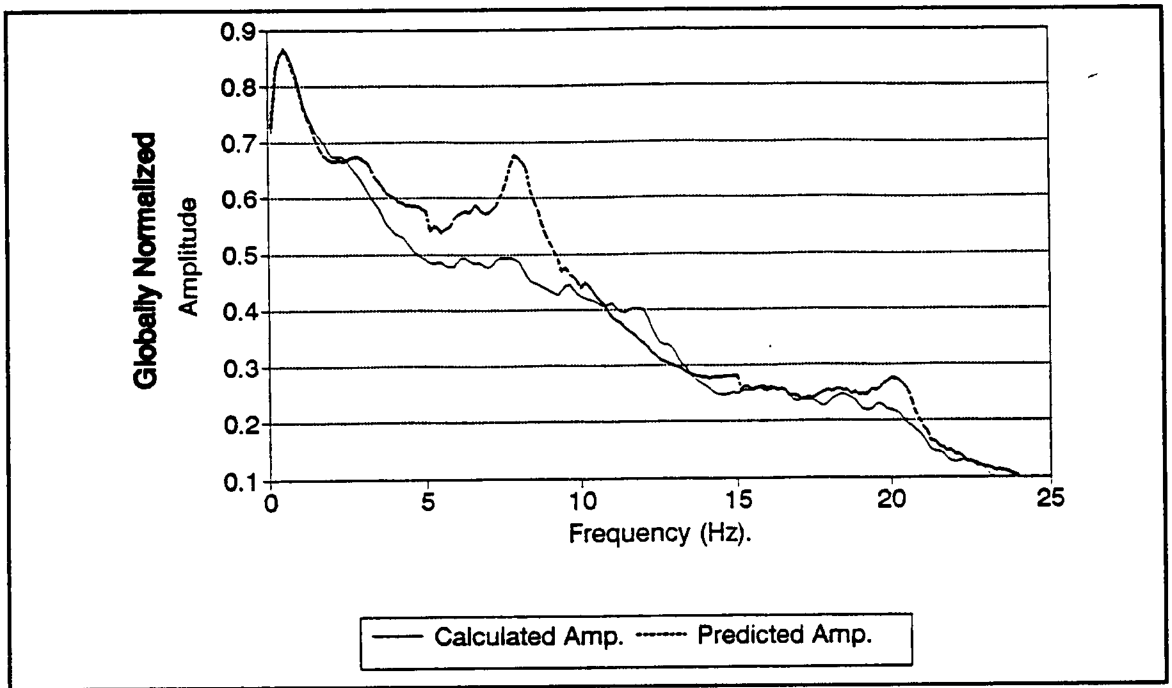

Figure II-3.9. Example of an Intermittent Mode Signature (Number 7). The large deviation from contact-free mode is expressed by the high value of the descriptors

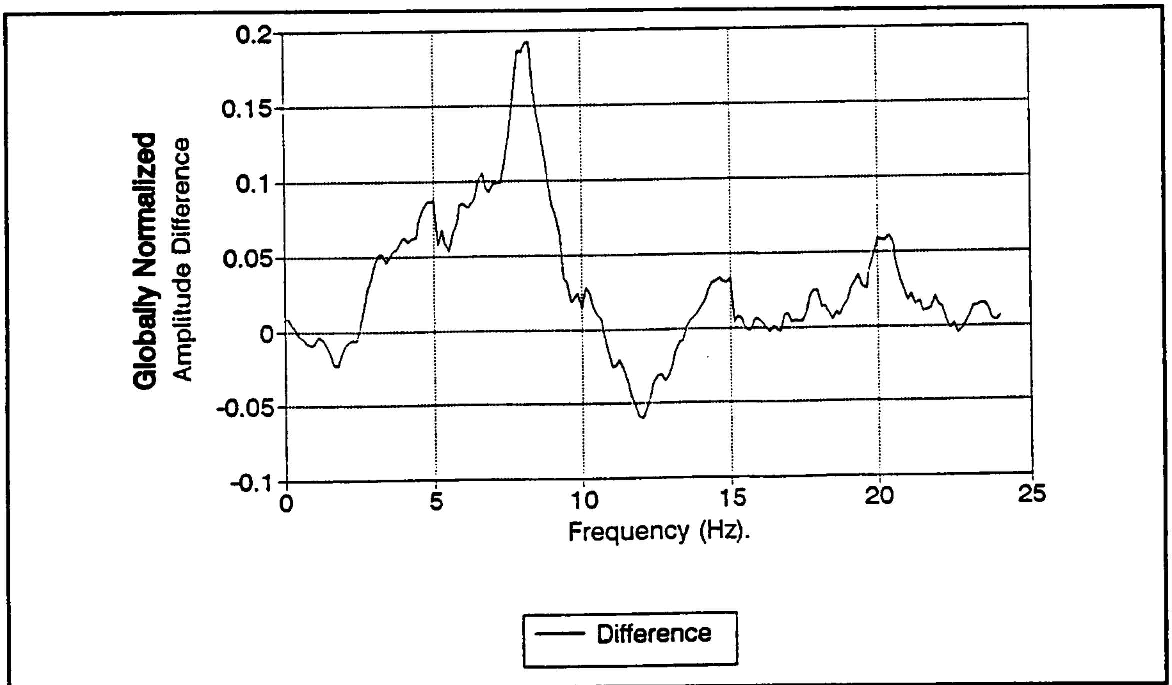

Figure II-3.10. Per-Frequency Descriptor, Signature 7. Difference between the prediction and actual values is large, reflecting deviations from contact-free mode

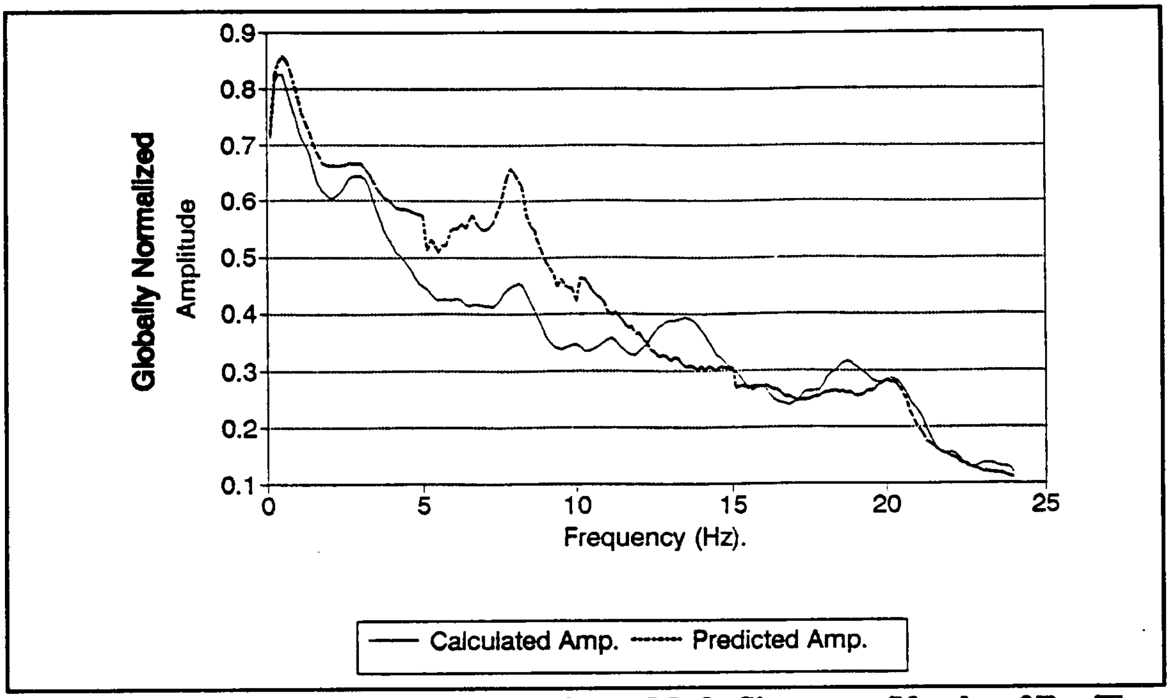

Figure II-3.11. Example of an Intermittent Mode Signature (Number 27). The large deviation from contact-free mode is expressed by the high value of the descriptors

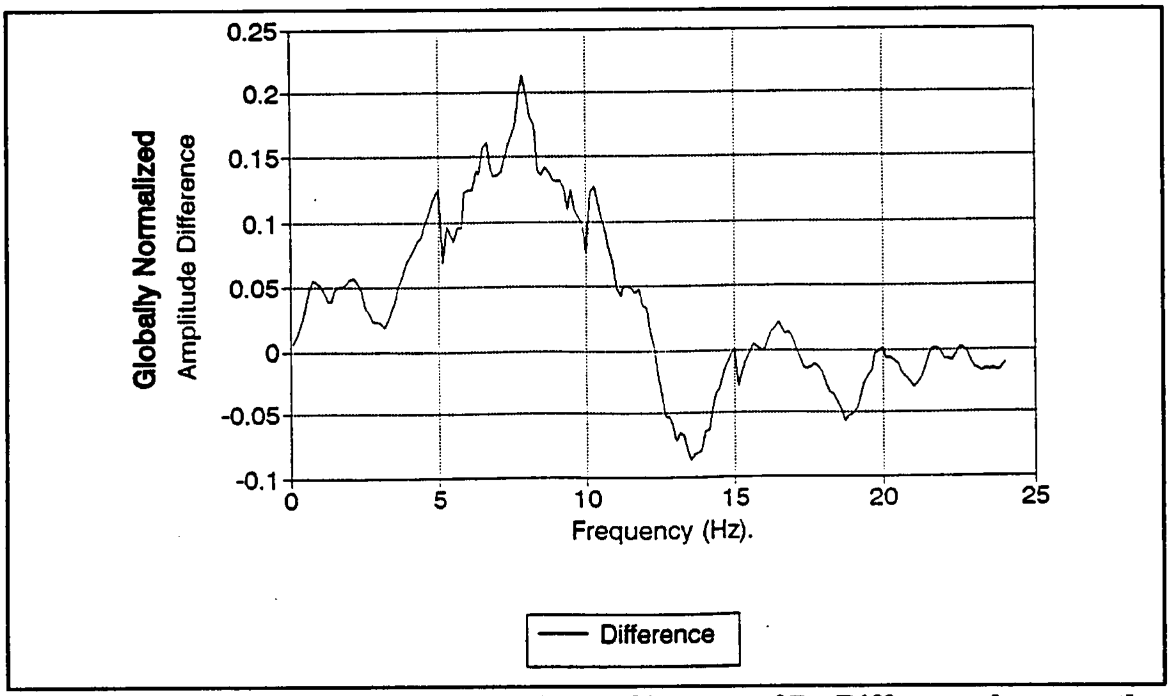

Figure II-3.12. Per-Frequency Descriptor, Signature 27. Difference between the prediction and actual values is large, reflecting deviations from contact-free mode

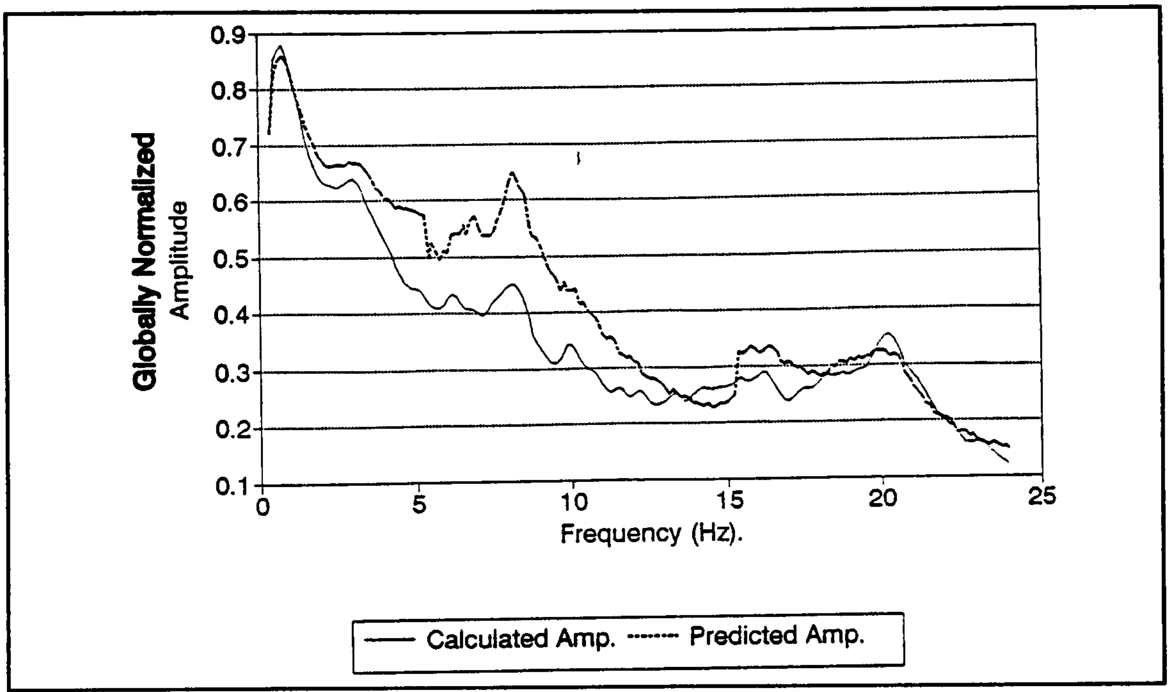

Figure II-3.13. Example of a Clamped Mode Signature (Number 35). Notice the large values for Descriptors 1 and 2, representing Bands 12, 13 and 14.

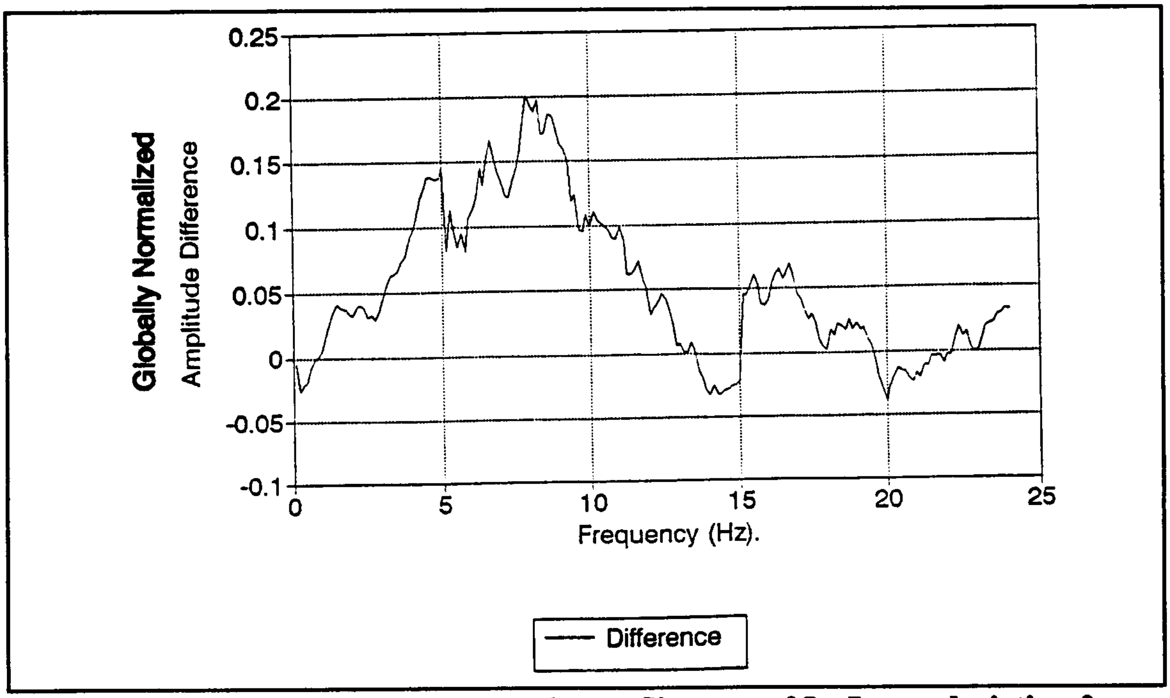

Figure II-3.14. Per-Frequency Descriptor, Signature 35. Large deviation from contact-free mode, as indicated by Descriptors.

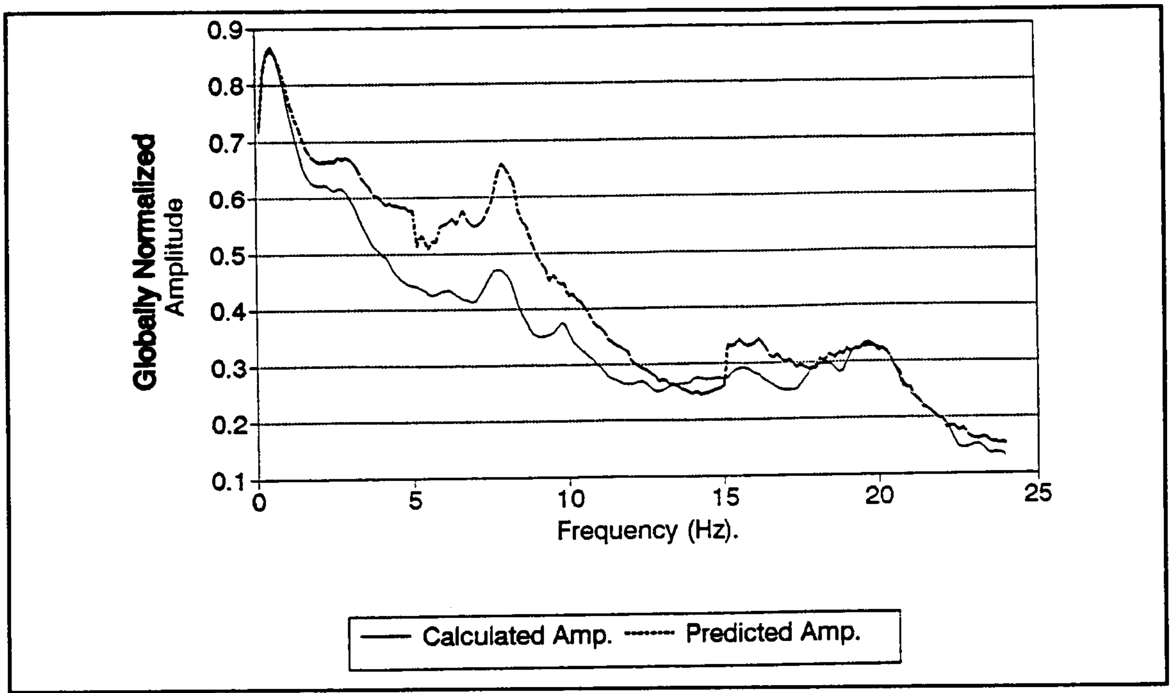

Figure II-3.15. Example of a Clamped Mode Signature (Signature 36). Notice the large values for Descriptors 1 and 2, representing Bands 12, 13 and 14

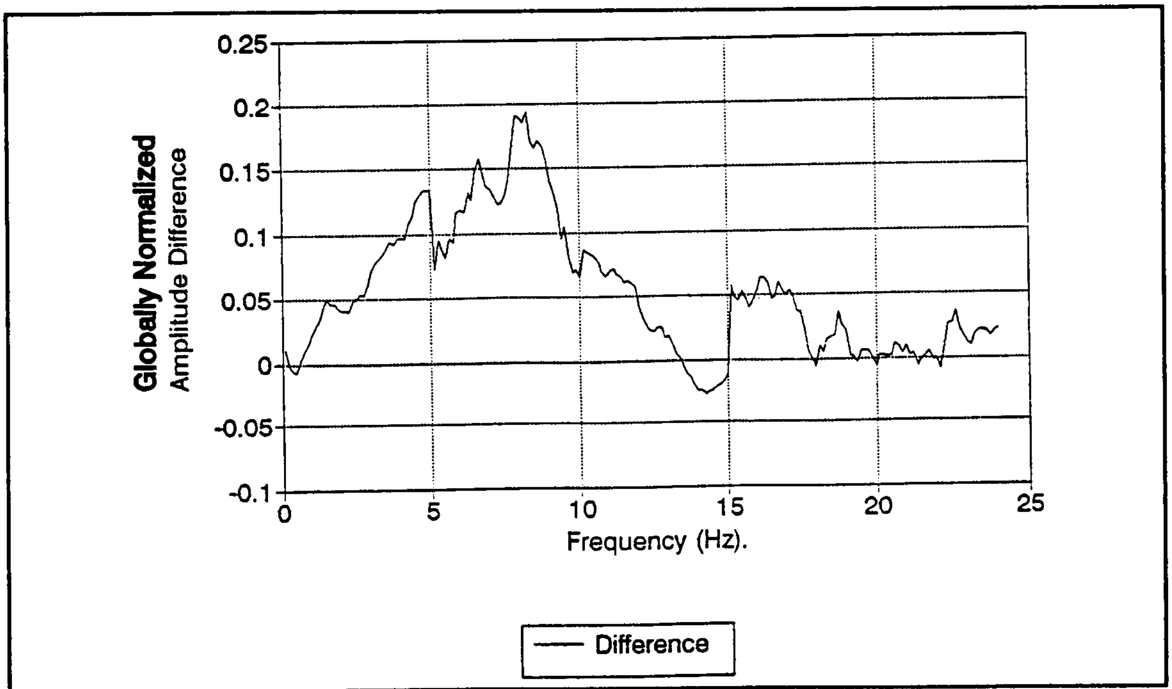

Figure II-3.16. Per-Frequency Descriptor, Signature 36. Large deviation from contact-free mode, as indicated by descriptors

Table II-3.2. Descriptors for Monitoring System
* Training pattern

Pattern	Vib. Mode	Descriptor 1 0.0 - 5.0 Hz. Bands 10, 11	Descriptor 2 5.0 - 10.0 Hz. Bands 12, 13	Descriptor 3 10.0 - 14.0 Hz. Band 14	Descriptor 4 14.0 - 24.0 Hz. Bands 15, 16
1*	(L)	0.038919	0.035865	0.016985	0.006543
2	(L)	0.035223	0.036208	0.025192	0.008223
3	(LC)	0.090374	0.116983	0.012993	0.025027
4	(L)	0.081317	0.122492	0.032078	0.110391
7	(I)	0.075166	0.454292	0.035938	0.035191
11	(I)	0.075572	0.274751	0.053619	0.022575
12	(AL)	0.087952	0.338358	0.041574	0.070872
27	(I)	0.148668	0.765937	0.165464	0.031474
28	(I)	0.054361	0.425489	0.089701	0.045752
29	(I)	0.168882	0.72191	0.179135	0.020893
35	(AL)	0.207608	0.827043	0.0127639	0.06808
36	(AL)	0.222705	0.71676	0.088298	0.06358
37	(AL)	0.122434	0.354622	0.075354	0.036335
38	(AL)	0.265654	0.91016	0.143979	0.05135
39	(F)	0.038001	0.037055	0.016724	0.027448

from the detector signal, and deviations are identified. Generally, large deviations from the prediction indicates the appearance of contacts.

The system includes four neural networks modelling bands of the spectrum separately. The division of the spectrum into meaningful regions makes it possible to prevent local increases in individual bands from affecting the amplitudes of other bands which may not be related from a physics standpoint. The spectrum was divided into 4 regions, and frequencies beyond 24.0 Hz were not considered in the analysis because it contains information about complex modes of the internals and the sensor background noise, none of which is the topic of this study.

Under this study, the technique has been used to detect significant contacts (clamped and intermittent modes) between the core barrel and the reactor vessel. Two descriptors have been defined which are very useful for the quantification of deviation from contact-free mode, and for the identification of the regions of the spectrum where amplitudes have changed. Since there is a established relationship between bands in the spectrum and components in the internal, the technique makes it possible to identify clearly where and by how much the amplitudes have changed. The relevance of this method is that it permits the definition of parameters which can be used directly for automated monitoring of contacts in the structure. Electricité de France hopes to automate the entire process of monitoring of mechanical and structural vibration in a workstation to be installed in each of its operating plants.

Another process that could be automated with this technique is the detection of improperly calibrated or out-of-other sensors discussed in Chapter II-2. Because invalid signatures have amplitudes and shapes which are not consistent with the shape of normal signatures, these differences would appear as high numbers in ALL descriptors, since they gather the square of the differences at each frequency.

CONCLUSIONS AND RECOMMENDATIONS

A methodology is presented for the analysis of vibration data based on the technology of neural networks. The anticipated advantage of developing a system built upon these principles is the possibility of automating the monitoring and diagnostic process for vibrating components, and developing diagnostic and fault-detection systems which complement traditional PSD analysis.

In this dissertation, the problem of vibration has been studied as it applies to rotating machinery and reactor internals. Neural network systems have been developed for the monitoring and identification of the operational state of these systems, and the methodology has been tested on data from operating power plants. The results are very accurate.

For mechanical systems, a technique is described for the monitoring of rotating machinery vibration. A neural network has been trained to mimic the relationship among sensors in a machine during normal operating conditions, and to produce the spectrum of data from one sensor using as input the spectrum of data from another sensor on the machine. Deviations from normal behavior can be detected by comparing the predictions of the network and the actual values of the sensors. Two new descriptors have been defined for quantifying the deviation, and for isolating the vibration amplitude changes.

To address the problem of identifying operational state of rotating machinery, a system was developed for discrimination of spectral signatures from rolling element bearings. The system includes a module for classification of signatures and a module for clustering. The classification module uses backpropagation networks to discriminate among different states, the inputs to this module are reduced versions of the spectra which had been compressed using recirculation networks. Both low frequency and high frequency information is included, to enable the system to detect incipient faults as well as severe defects of the bearings. The systems were tested using data collected during a simulated test of aging conducted by EDF.

The recirculation networks are critical for improving the performance of the system because they perform automatic feature extraction on the spectra. Also, by reducing the dimension of the input vectors, the training time is reduced by a factor of 10. Additional analysis was done on the ability of recirculation networks to perform noise removal, test sets were generated by adding noise to the signatures, and the effect of perturbing the spectra was measured after processing the disturbed signatures with the compression networks. On the average, 1 % of added noise produces only a 0.15 % perturbation on the spectra.

The clustering module uses a modified version of the Fuzzy-C algorithm adapted to the realm of vibration analysis; a new distance measure is introduced using a matrix of weights. The matrix contains numbers which reflect the level of importance of each frequency in the spectrum, so that information in the spectrum may be considered according to its relevance from the theoretical standpoint.

In the realm of vibration of reactor internals, a monitoring system based on the modelling described for vibration of rotating machinery was implemented. In this case, the input to the system is spectra from data of an ex-core detector on the structure, and the output, spectra from data collected on the opposite ion chamber (180 deg. apart). Using this modelling and the descriptors defined, deviations from normal behavior become readily visible, and using the known associations between regions in the spectrum and components in the structure, it is simple to relate the behavior to a source of excitation.

The classification module for the discrimination of ex-core neutron detectors, uses the noise portion of the signal, and more specifically parameters calculated from its spectrum, to differentiate among the different modes of normal behavior. The system was also enhanced with the capability of detecting malfunctioning sensing equipment, and identifying failure of the hold-down spring.

The system was tested with a large set of data provided by EDF. The data on which the discrimination system for normal vibration modes was tested was collected from 28 different reactors in France. The data was collected over 93 fuel cycles and included around 1600 spectra. Since no cases of hold-down spring degradation or malfunctioning have been registered in French PWR's, the module for the detection of spring degradation, was tested with simulated data generated by EDF. The results of the testing were over 98 % accurate.

This results demonstrate the ability of neural networks with very simple topologies to perform complex classification and modelling tasks. It was also confirmed that the criterion proposed by Kung and Hwang [35] works very well for this kind of problems, and that significant improvements in accuracy may be gained by breaking a large problem into a series of smaller tasks.

Finally, we would like to offer some suggestion for the enhancement of the techniques and systems described herein. Regarding the system for discrimination of spectra from rotating machinery, we believe it would be valuable to enhance the system's abilities by considering the phase information. As discussed in Chapter I-1, the phase information is necessary for the detection of certain faults, and more importantly it is the only means for discriminating among faults that produce the same features in the PSD. The family of faults for which the phase information is necessary was not considered in this study, but it appears natural to continue the work in this direction. It would also be beneficial to explore other faults, especially from larger components in rotating machinery, and try to determine which algorithms and topologies of neural networks work better for more complex versions of the problem.

In the area of vibration of the reactor internals, we consider that a lot of information can be extracted from the activations of the neurons in the output level of the system. One of the most interesting realizations encountered during this part of the work was that for intermediate or transition cases, the network outputs always reflected the transition through high activations in the neurons representing the classes involved. This indication is the result of selecting typical cases of each class for training, and of the scheme chosen to represent the output activations (ie. the output neuron corresponding to the class the pattern represents is set to the highest activation and all remaining neurons to the lowest).

It seems fitting to enhance the capabilities of the system with a feature capable of interpreting the activations of the output neurons, so this information can be displayed in a natural manner to the user. We feel that such system is simple to implement using reasoning techniques such as expert systems and fuzzy inferencing, with active expert participation.

LIST OF REFERENCES

LIST OF REFERENCES

1. G. ZWINGELSTEIN, M. H. MASSON, B. DUBUISSON, M. PAVARD, J. M. MAZALERAT, "The Application of Neural Networks for the Predictive Maintenance of Nuclear Power Plants," Proceedings of the 3rd Symposium on Expert System Applications to Power Systems, Tokyo, Japan, April 1-5, 1991.

2. HEWLETT PACKARD, *Effective Machinery Maintenance Using Vibration Analysis*, Application Note 243-1, Dynamic Signal Analyzer Applications, 1983.

3. J. C. CARRE, A. EPSTEIN, C. PUYAL, M. CASTELLO, P. DUMORTIER, "Malfunction Tests and Vibration Analysis of PWR Internal Structures," (in French), Progress in Nuclear Energy, Vol. 21, pp. 71-78, 1988.

4. DARPA, *Neural Networks Study,* AFCEA International Press, Fairfax, Virginia, 1988.

5. R. P. LIPPMANN, "An Introduction to Computing with Neural Networks," IEEE ASSP Magazine, April 1987.

6. W. S. MCCULLOCH, W. PITTS, "A Logical Calculus of the Ideas Imminent in Nervous Activity," Bulletin of Mathematical Biophysics, 5, pp. 115-133, 1943.

7. R. ROSENBLATT, *Principles of Neurodynamics,* Spartan Books, New York, 1959.

8. D. O. HEBB, *The Organization of Behavior,* John Wiley & Sons, New York, 1949.

9. B. WIDROW, M. E. HOFF, "Adaptive Switching Circuits," 1960 IRE WESCON Conv. Record, Part I, pp. 96-104, August 1960.

10. B. WIDROW, S. D. STEARNS, *Adaptive Signal Processing,* Prentice-Hall, New Jersey, 1985.

11. T. E. POSCH, "Models of the Generation and Processing of Signals by Nerve Cells: A Categorically Indexed Abridge Bibliography," USCEE Report 290, August 1968.

12. M. MINSKY, S. PAPERT, *Perceptrons: An Introduction to Computational Geometry,* MIT Press, Cambridge, MA, 1969.

13. D. E. RUMELHART, J. L. MCCLELLAND, *Parallel Distributed Processing,* MIT Press, Cambridge, MA, 1986.

14. D. E. RUMELHART, G. E. HINTON, R. J. WILLIAMS, "Learning Representations by Back-Propagating Errors," Nature, 323, pp. 533-536, 1986.

15. D. E. RUMELHART, G. E. HINTON, R. J. WILLIAMS, "Learning Internal Representations by Error Propagating," *Parallel Distributed Processing: Explorations in the Microstructure of Cognition,* D. E. Rumelhart and J. L. McClelland (Eds.), MIT Press, Cambridge, MA, 1986.

16. S. GROSSBERG, *The Adaptive Brain I: Cognition, Learning, Reinforcement and Rhythm,* and *The Adaptive Brain II: Vision, Speech Language and Motor Control,* Elsevier/North Holland, Amsterdam, 1986.

17. J. J. HOPFIELD, "Neural Networks and Physical Systems with Emergent Collective Computational Abilities," Proc. Natl. Acad. Sci. USA, Vol. 79, pp. 2554-2558, April 1982.

18. J. J. HOPFIELD, "Neurons with Graded Response Have Collective Computational Properties Like Those of Two-State Neurons," Proc. Natl. Acad. Sci. USA, Vol. 81, pp. 3088-3092, May 1984.

19. J. J. HOPFIELD, D. W. TANK, "Computing with Neural Circuits: A Model," Science, Vol. 233, pp. 625-633, August 1986.

20. T. SEJNOWSKY, C. R. ROSENBERG, "NETtalk: A Parallel Network that Learns to Read Aloud," Johns Hopkins Univ. Technical Report JHU/EECS-86/01, 1986.

21. D. F. SPECHT, "Generation of Polynomial Discriminant Functions for Pattern Recognition," Stanford Electronics Labs. Rept. SU-SEL-66-029, May 1966.

22. D. F. SPECHT, "Probabilistic Neural Networks," Neural Networks, Vol. 3, Number 1, 1990.

23. B. KOSKO, "Bidirectional Associative Memories," IEEE Transactions on Systems, Man and Cybernetics, Vol. 18, No. 1, January/February 1988.

24. B. KOSKO, "Adaptive Bidirectional Associative Memories," Applied Optics, Vol. 26, No. 23, December 1987.

25. T. KOHONEN, *Self Organizing and Associate Memory*, Springer-Verlag, Berlin, 1984.

26. T. KOHONEN, "Correlation Matrix Memories," IEEE Transactions on Computers, C-21, pp. 353-359, 1972.

27. R. HECHT-NIELSEN, "Applications of Counterpropagation Networks," Neural Networks, Vol. 1, pp. 131-139, 1988.

28. S. KEYVAN, A. DURG, "Pump Shaft Condition Monitoring Via Artificial Neural Networks," Proceedings of the EPRI V Predictive Maintenance Conference, Knoxville, TN, September 21-23, 1992.

29. G. ZWINGELSTEIN, M. H. MASSON, B. DUBUISSON, M. PAVARD, J. M. MAZALERAT, "The Application of Neural Networks for the Predictive Maintenance of Nuclear Power Plants," Proceedings of the III Symposium on Expert System Applications to Power Systems, Tokyo, Japan, April 1-5, 1991.

30. P. D. WASSERMAN, A. UNAL, S. HADDAD, "Neural Networks for On-Line Machine Condition Monitoring," Proceedings of the Artificial Neural Networks in Engineering Conference, St. Louis, MO, November 10-13, 1991.

31. P. D. WASSERMAN, "On-Line Machine Condition Monitoring Using an Integrated Neural Network Fuzzy Logic System," Proceedings of the EPRI V Predictive Maintenance Conference, Knoxville, TN, September 21-23, 1992.

32. P. D. WASSERMAN, *Neural Computing, Theory and Practice*, Van Nostrand Reinhold, New York, 1989.

33. NEURALWARE INC., *Neural Computing*, Manuals for NeuralWorks Professional II and NeuralWorks Explorer Software, Vol. 1, Pittsburgh, PA, 1990.

34. A. N. KOLMOGOROV, "On the Representation of Continuous Functions of Many Variables by Superposition of Continuous Functions of One Variable and Addition," Dokl. Akad. Nauk, USSR, 114, 953-956.

35. S. Y. KUNG, J. N. HWANG, "An Algebraic Projection Analysis for Optimal Hidden Unit Size and Learning Rates in Backpropagation," Proceedings of the IEEE International Conference on Neural Networks, San Diego, California, July 24-27, 1988.

36. G. E. HINTON, J. L. MCCLELLAND, "Learning Representations by Recirculation," Proceedings of the IEEE Conference on Neural Information Processing Systems, November 1988.

37. J. L. MCCLELLAND, M. ST. JOHN, "Three Short Papers on Language and Connectionism," Carnegie-Mellon University, Technical Report AIP-1.

38. I. E. ALGUINDIGUE, R. E. UHRIG, "Compression of Spectral Signatures Using Recirculation Networks," Scientific Computing and Automation, Vol. 7, No. 7, May 1991.

39. P. B. LEONETTI, "Integrating Preventive Maintenance, Predictive Maintenance, and Maintenance Technical Support," Proceedings of the EPRI V Predictive Maintenance Conference, Knoxville, Tennessee, September 21-23, 1992.

40. V. SÜDMERSEN, D. STEGEMANN, C. CAMERINI, "Vibration Analysis in Predictive Maintenance," Proceedings of the XIII World Conference on Non-Destructive Testing, Sâo Paulo, Brasil, October, 18-23, 1992.

41. C. JACKSON, *The Practical Vibration Primer*, Gulf Publishing Company, Houston, Texas, 1979.

42. J. C. ROBINSON, "Consideration for Vibration Monitoring of Pumps," Proceedings of the EPRI V Predictive Maintenance Conference, Knoxville, Tennessee, September 21-23, 1992.

43. G. ZWINGELSTEIN, L. HAMON, "EDF Studies on the Condition Monitoring of Rolling Element Bearings," Proceedings of the 4th Incipient Failure Detection Conference, Philadelphia, Pennsylvania, Oct. 1990.

44. S. V. BOWERS, K. R. PIETY, R. W. PIETY, "Temporary Sensor Mounting," Proceedings of the IV Incipient Failure Detection Conference, Philadelphia, Pennsylvania, Oct. 1990.

45. A. A. KHARKEVICH, *Spectra and Analysis*, Consultants Bureau, New York, 1960.

46. J. S. BENDAT, A. G. PIERSOL, *Measurement and Analysis of Random Data*, John Wiley & Sons, New York, 1966.

47. J. T. BROCH, *Mechanical Vibrations and Shock Measurements*, Bruel & Kjaer, 1984.

48. J. MOREL, D. LE REVERED, P. NEAU, "Vibration Monitoring of Heavy Duty Machines," Proceedings of the 4th Incipient Failure Detection Conference, Philadelphia, Pennsylvania, Oct. 1990.

49. L. WINN, "Reliability and Failure Detection if Rolling Element Bearings in Power Plant Machinery: A Utility Survey," Proceedings of the IV Incipient Failure Detection Conference, Philadelphia, Pennsylvania, Oct. 1990.

50. J. E. BERRY, "How to Specify Machinery Vibration Spectral Alarm Bands," Proceedings of the EPRI V Predictive Maintenance Conference, Knoxville, Tennessee, September 21-23, 1992.

51. M. ANGELO, *Bruel & Kjaer Technical Review: Vibration Monitoring of Machines*, No.1, Naerum Offset, Denmark, 1987.

52. Electric Power Research Institute Short Course, Advance Machinery Vibration Diagnostics, EPRI Course Notes, February 4-7, 1992.

53. G. B. HIRSCHMAN, "Data Acquisition During Speed Transients in Rotating Machinery," Proceedings of the EPRI V Predictive Maintenance Conference, Knoxville, Tennessee, September 21-23, 1992.

54. I. IMAM, J. SCHEIBEL, S. H. AZZARO, R. J. BANKERT, "Development and Application of an On-Line Rotor Crack Detection and Monitoring System," Proceedings of the EPRI III Incipient Failure Detection Conference, EPRI CS-5395, August 1988.

55. D. A. RICE, "New Techniques for Vibration Analysis," Proceedings of the XI Annual Meeting of the Vibration Institute, June 1988.

56. M. CURRY, "Machine Specific Diagnostics," Proceedings of the EPRI V Predictive Maintenance Conference, Knoxville, Tennessee, September 21-23, 1992.

57. I. E. ALGUINDIGUE, A. LOSKIEWICZ-BUCZAK, R. E. UHRIG, L. HAMON, F. LEFEVRE, "Vibration Monitoring of EDF Rotating Machinery Using Artificial Neural Networks," Proceedings of the AI91 Frontiers in Innovative Computing for the Nuclear Industry Conference, Jackson, Wyoming, September 15-18,1991.

58. I. E. ALGUINDIGUE, R. E. UHRIG, "Vibration Monitoring with Artificial Neural Networks," Proceedings of the SMORN VI Symposium on Nuclear Reactor Surveillance and Diagnostics, Gatlinburg, Tennessee, May 19-24, 1991.

59. J. S. BENDAT, *Non-Linear System Analysis and Identification from Random Data*, John Wiley & Sons, New York, 1990.

60. J. S. BENDAT, A. G. PIERSOL, *Random Data Analysis and Measurement Procedures*, John Wiley & Sons, Second Edition, New York, 1986.

61. J. T. TOU, R. C. GONZALEZ, *Pattern Recognition Principles*, Addison-Wesley, Reading, Massachusetts, 1974.

62. M. M. TRIVEDI, J. C. BEZDEK "Low-Level Segmentation of Aerial Images with Fuzzy Clustering," IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-16, No 4, July / August 1986.

63. X. L. XIE, G.BENI, "A Validity Measure for Fuzzy Clustering," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 13, No 8, August 1991.

64. I. E. ALGUINDIGUE, A. LOSKIEWICZ-BUCZAK, R. E. UHRIG, "Clustering and Classification Techniques for the Analysis of Vibration Signatures," Proceedings of the SPIE Conference on Applications of Artificial Neural Networks III, Orlando, Florida, U.S.A., April, April 21-24, 1992.

65. T. BOGARD, T. KOZLOSKY, N. SINGLETON, "Reactor Internals and Core Support Monitoring System," Proceedings of the SMORN VI Symposium on Nuclear Reactor Surveillance and Diagnostics, Gatlinburg, Tennessee, May 19-24, 1991.

66. J. A. THIE, "Core Motion Monitoring," Nuclear Technology, Vol. 45, 1979.

67. J. A. THIE, *Power Reactor Noise*, American Nuclear Society Press, La Grange Park, Illinois, 1981

68. A. TRENTY, F. LEFEVRE, D. GARREAU, "Computer-Generated Vibratory Signatures for EDF PWR Reactor Vessel Internals," IMORN 23, June 3, 1992.

69. P. BERNARD, Ch. MESSAINGUIRAL-BRUYNOOGHE, J. CLOUE, "Liaison Entre Vibrations d'Origine Mécanique et Fluctuations du Flux Neutronique dans le R. E. P.," (In French), Proceedings of the SMORN IV Symposium on Nuclear Reactor Surveillance and Diagnostics,Dijon, France, October 1984.

70. G. M. PRESNELL, "Duke Power's Investigation of Changing Neutron Noise Signatures at McGuire Nuclear Unit 1 During Recent Cycles of Operation," Proceedings of the SMORN VI Symposium on Nuclear Reactor Surveillance and Diagnostics, Gatlinburg, Tennessee, May, 1991.

71. A. TRENTY, C. PUYAL, H. KLAJNMIC, "SINBAD, a Data Base for PWR Internals Vibratory Monitoring," Proceedings of the SMORN VI Symposium on Nuclear Reactor Surveillance and Diagnostics, Gatlinburg, Tennessee, May 19-24, 1991.

72. A. TRENTY, H. KLAJNMIC, "Operational Feedback on Internal Structure Vibration in 28 French 900-MW PWR's," IMORN 23, June 3, 1992.

73. R. T. WOODS, "A Neutron Noise Diagnostic Methodology for Pressurized Water Reactors," Ph. D Dissertation, The University of Tennessee, Knoxville, December 1990.

74. A. TRENTY, Personal Communication, Electricité de France, Direction des Etudes et Recherches, Départment Surveillance Diagnostic Maintenance, Chatou, France, November, 1992.

75. V. BAUERNFEIND, "Vibration Monitoring of a Four-Loop PWR: Model Investigations of the Sensitivity of the Monitored Signals on Mechanical Failures," Progress in Nuclear Energy, Vol. 21, pp. 247_254, 1988.

76. M. MAISSON, "Discrimination Neuronale des Modes de Balancement d'Internes á Partir de Spectres de Bruit Neutronique de REP 900 MW - CP1-CP2," (in French), Electricité de France Internal Report, Ref. HP-21/92-21.

77. ELECTRICITE DE FRANCE, "Reacteur: Notice Technique sur les Centrales Nucleaires PWR," (in French), Number 321, Direction de la Production et du Transport, Service de la Production Thermique, 1977.

VITA

Israel E. Alguíndigue was born in Coro, Venezuela in 1963. He attended the University of Tennessee Martin where he earned a Bachelors of Science degree in Computer Science and Mathematics, and went to the University of Tennessee Chattanooga (UTC) to obtain a Masters of Science degree in Computer Science. He worked for the Center of Excellence for Computer Applications while attending UTC.

In 1990, he began doctoral studies at the University of Tennessee Knoxville, working as a graduate research assistant for the Nuclear Engineering Department. Under the supervision of Dr. Robert E. Uhrig he studied the applicability of artificial intelligence techniques to the monitoring and diagnosis of nuclear power plant equipment. His research work has been documented in over thirty publications in professional journals and national and international conferences, and he has spoken in numerous occasions for research laboratories and universities in the U.S.A, Venezuela, France, Mexico and Germany.

Mr. Alguíndigue is a member of the Association for Computing Machinery, the Institute of Electric and Electronic Engineers, the American Association of Artificial Intelligence, and the American Nuclear Society. He has been elected member of the Phi Kappa Phi and Pi Epsilon Honorary Societies. His research interests include artificial neural systems, approximate reasoning, pattern recognition, computer vision, logic programming, and signal processing.