Mechanical Systems and Signal Processing (1995) **9**(4), 397–414

A CLASSIFIER NEURAL NETWORK FOR ROTORDYNAMIC SYSTEMS

R. Ganesan, Jin Jionghua and T. S. Sankar

Department of Mechanical Engineering, Concordia University, Montreal H3G 1M8, Canada

(*Received* 25 *November* 1993, *accepted* 26 *May* 1994)

A feedforward backpropagation neural network is formed to identify the stability characteristic of a high speed rotordynamic system. The principal focus resides in accounting for the instability due to the bearing clearance effects. The abnormal operating condition of 'normal-loose' Coulomb rub, that arises in units supported by hydrodynamic bearings or rolling element bearings, is analysed in detail. The multiple-parameter stability problem is formulated and converted to a set of three-parameter algebraic inequality equations. These three parameters map the wider range of physical parameters of commonly-used rotordynamic systems into a narrow closed region, that is used in the supervised learning of the neural network. A binary-type state of the system is expressed through these inequalities that are deduced from the analytical simulation of the rotor system. Both the hidden layer as well as functional-link networks are formed and the superiority of the functional-link network is established. Considering the real time interpretation and control of the rotordynamic system, the network reliability and the learning time are used as the evaluation criteria to assess the superiority of the functional-link network. This functional-link network is further trained using the parameter values of selected rotor systems, and the classifier network is formed. The success rate of stability status identification is obtained to assess the potentials of this classifier network. The classifier network is shown that it can also be used, for control purposes, as an 'advisory' system that suggests the optimum way of parameter adjustment.

1. INTRODUCTION

A strong interest has developed during the past few decades in predicting the dynamic behaviour of rotordynamic systems that are deployed in industrial turbomachinery [1–3]. Fair among the published works that deal with the motion characteristics of the rotordynamic systems with slender shafts, concerned about the adverse effects of bearing clearances [4–14]. Of these, the experimental work of Ehrich [4] and Bently [8] demonstrated the presence of the non-linear motion characteristics that were attributed to the effects of bearing clearances. These experimental findings have been confirmed through the analytical study in the works of Childs [9], Muszynska [10], Beatty [11], Saito [12], Choi and Noah [13], Ganesan and Sankar [14] and Ehrich and Childs [15]. Also numerical methods like harmonic balance methods using FFT have been used to analyse in detail, the motion characteristics of rotors with non-symmetric bearing clearances [16]. Additional works in this topic can be found in abundance in the studies of Ek [17], Childs [18] and Shaw [19].

0888–3270/95/040397 + 18 $12.00/0

From these published works, it can be observed that:

(i) Highly irregular and complex physical phenomena like self-excited oscillations, chaotic vibrations, non-linear parametric response, etc. mostly stem from the effects of bearing clearances.

(ii) Rotordynamic systems that are supported on hydrodynamic bearings are quite often submitted to abnormal operating conditions such as shifts in amplitude and phase, occurrence of unexpected frequencies of oscillation, pseudocritical response peaks, etc.

(iii) A partial or complete loss of rotor's radial compliance is observed over a portion of its synchronous orbit, in the case of rotor systems with hydrodynamic bearings; in the case of rotors supported on rolling element bearings, the small clearance that is provided to account for axial shaft motion, causes a continuous existence of normal-loose Coulomb rub condition.

(iv) The collective effect of bearing clearances with a fixed-direction sideload is, in many cases, the major source of non-linear motion in the case of rolling element bearings.

(v) Whipping and whirling, parametric instability, subharmonic resonances, stick-slip rubs, chatter, etc. mark their presence as a result of all these circumstances.

One of the major objectives of recent research on rotordynamic systems is to design and develop condition monitoring and control procedures that (i) detect the presence and growth of potentially dangerous fractional-frequency rotor motions by making use of the features of vibratory signal symptoms, (ii) assess the instantaneous stability behaviour of the rotordynamic system and its sensitivity to a number of parameters like system (geometric) bearing coefficients, inertial properties of mass concentrated disks, stiffness of rotating assemblies, etc. (iii) use this information to exploit the non-linear relationships that exist between the stability characteristic and the physical parameters, for the purposes of control. Another important objective of the detailed study that concerns with the motion characteristics of such rotor systems, is to develop self-adaptive actively-controlled gyroscopic systems. Successful achievement of both these objectives completely hinges on (i) the ability to automatically recognise the root causes of an abnormal condition of the system through the nature and special features of the vibration signal, (ii) the time elapsed to compute the information that serve as a measure of the stability condition of the system, (iii) the measure of integrability with the diagnostic and control systems like CNC systems, etc. Further, the decision as to what parameter must be adjusted to safely and reliably drive the system back to its normal conditions and the calculation of the magnitudes of such corrective measures must be carried out accurately and quite fast, to ensure safer operating conditions that avoid irreversible accumulated damage. For this talk to be accomplished, the information about the effects of parameter changes that have a direct bearing on the ensuing motion should be made available. Of course, availability of a completely-symbolic processing technique is an added advantage, to ensure effective diagnostic and control measures.

Artificial neural networks or neural networks [20, 21], which are massively parallel, interconnected networks that consist of basic computing elements called neurons, can be successfully used in these situations since they possess the high level of adaptivity that cannot be obtained from the completely-analytical or numerical procedures. These basic computing elements loosely model the mechanics of computations in neurobiological processes. These neural networks are highly suitable to both the automatic control and intelligent diagnostic systems of mechanical and manufacturing industry, for the following reasons:

(i) Both the structure of the network and the way with which the connectivity is achieved between its basic elements (or processors) have equal and significant influence on the overall behaviour of the trained network. This offers a high level of flexibility in the practical use of these computational models.

(ii) An approach that is clearly distinguished from traditional numerical computing strategies and even from the modern more recent algorithms of symbolic processing can be seen in these neural networks.

(iii) The network, once trained, provides a rapid and precise mapping of a given input into the desired output variables and this is an advantage that is unique to neural networks, with respect to non-linear inverse problems of control.

(iv) The capability of a feedforward neural network to perform the desired task can be easily achieved through the gradient-descent-based backpropagation algorithm.

However, massive data are required to train a neural network and for the problems of diagnostics, such real data are impractical to obtain. So, an alternative way of developing classifier neural networks is developed here. The mathematical model is used to develop a neural network that identifies the state of the rotordynamic system. As and when either the experimental data or the real case studies are available, this network can be updated by using these additional information for training purposes. Hence, the approach developed herein is something similar to the scheme of iteration that is followed in numerical algorithms.

In the present work, a neural network-based approach is delineated, to be applicable to the diagnostic and control problems of industrial high-speed rotordynamic systems that are subjected to the undesirable effects of bearing clearances. The conditions with respect to the stability characteristics are simulated through an appropriate analytical modeling of the rotordynamic system. This obviates the need for a set of on-line vibratory signal data that contain the information about the bearing clearance effects. Since, such a data set for all rotors that are currently in use is not easily available and further awaiting such a data set to train the classifier network is impractical, the combined use of a valid analytical model and the parameter values of practical rotor systems, to provide the data for network training is only appropriate.

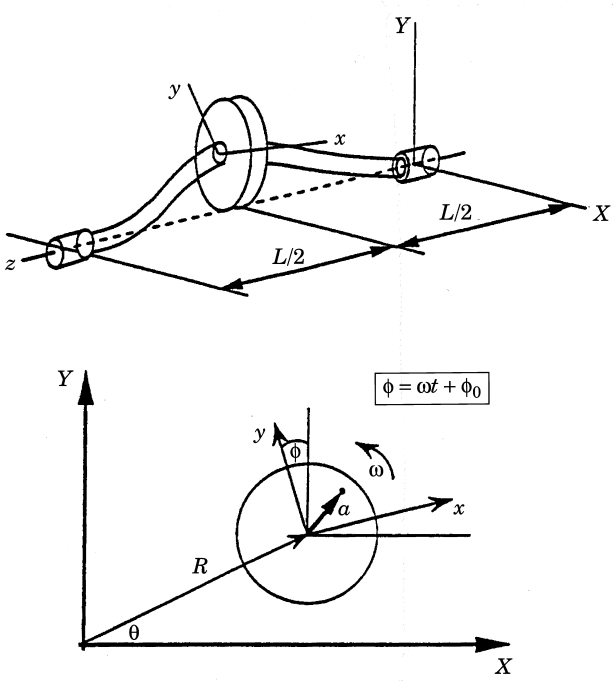

Figure 1. The Jeffcott rotor model.

2. ANALYTICAL SIMULATION OF FRACTIONAL-FREQUENCY ROTOR MOTION

The standard Jeffcott rotor model, which is a well known analytical model of rotordynamic systems with slender shafts, is now employed. With reference to Fig. 1, the following equations of motion are written that correspond to the constant running speed condition:

$$m\ddot{R}_x + c\dot{R}_x + kR_x = ma\omega^2 \cos\phi \tag{1}$$

$$m\ddot{R}_y + c\dot{R}_y + kR_y = ma\omega^2 \sin\phi - W \tag{2}$$

where R_x and R_y are the components of the rotor displacement vector R, m is the net rotor mass, k is the combined rotor-bearing radial stiffness that represents the net circumferentially symmetric compliance of the rotor shaft and bearings, c is the linear damping coefficient of the viscous damping mechanism, a is the imbalance vector magnitude, ω is the disk angular velocity (i.e. the rotational speed of the rotor), W is the total weight of the rotor and ϕ is the rotor disk rotation measured with respect to the fixed inertial reference frame XYZ such that $\phi = \phi_0 + \omega t$, with ϕ_0 being the rotation at zero instant of time or suitable time origin. It is also implied here that the imbalance vector is considered to be in alignment with the x-axis of the rotating frame of reference xyz.

To analytically model the Coulomb rubbing condition of the system, the discontinuous radial stiffness model is adopted. This model has been established as a powerful model that is capable of capturing the essential features of the bearing clearance effects [9]. Adopting this model, the equations of motion are now re-written, after some modifications, as

$$\ddot{R}_x + 2\zeta_0\lambda_0\dot{R}_x + \lambda_0^2[1 - \epsilon U(R_y)]R_x = a\omega^2 \cos\phi \tag{3}$$

$$\ddot{R}_y + 2\zeta_0\lambda_0\dot{R}_y + \lambda_0^2[1 - \epsilon U(R_y)]R_y = a\omega^2 \sin\phi - g \tag{4}$$

where ϵ is a perturbation parameter such that $0 \leqslant \epsilon \leqslant 1$, $\lambda_0 = (\mathrm{k/m})^{1/2}$ and $\zeta_0 = \mathrm{c}/[2(\mathrm{km})^{1/2}]$. Further, $U(R_y)$ is a step function which is defined based on the solution R_y to equation (2) given above. From equation (2), the harmonic steady-state response can be shown to be

$$R_y = A \sin(\phi + \psi) - W/k \tag{5}$$

where

$$A = R_x/\cos(\phi + \psi) = a\omega^2/\sqrt{[(\lambda_0^2 - \omega^2)^2 + 4\zeta_0^2\lambda_0^2\omega^2]} \tag{6}$$

$$\psi = \tan^{-1}[-2\zeta_0\lambda_0\omega/(\lambda_0^2 - \omega^2)]. \tag{7}$$

$U(R_y)$ is now given as

$$U(R_y) = 0, \quad R_y \leqslant 0 \tag{8}$$

$$U(R_y) = 1, \quad R_y > 0 \tag{9}$$

i.e. $U(R_y)$ is a unit step function, which means essentially that when the solution amplitude is less than the static deflection, the nominal stiffness value holds good, otherwise, the radial stiffness changes (see Fig. 2). Further, parameter ϵ has been introduced in equations (3) and (4) to characterise the rubbing condition. So, $\epsilon = 0$ represents the nominal condition and $\epsilon = 1$ represents the complete loss of radial stiffness due to Coulomb rub.

The solutions to the equations of motion are obtained based on the method of multiple scales [22] but after some modifications. To this end, a new set of scaled time variables is introduced according to

$$T_n = \epsilon^n t, \quad n = 0, 1, 2\cdots \tag{10}$$

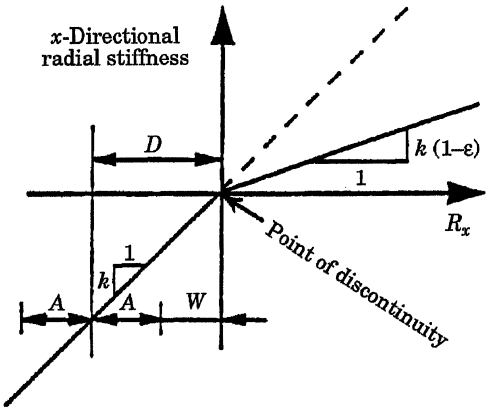

Figure 2. Normal-loose radial stiffness model.

As a consequence, time derivatives need to be re-written in the following form using the partial derivative rule

$$\frac{d}{dt} = \frac{dT_0}{dt}\frac{\partial}{\partial T_0} + \frac{dT_1}{dt}\frac{\partial}{\partial T_1} + \frac{dT_2}{dt}\frac{\partial}{\partial T_2} + \cdots = D_0 + \epsilon D_1 + \epsilon^2 D_2 + \cdots \tag{11}$$

$$\frac{d^2}{dt^2} = D_0^2 + 2\epsilon D_0 D_1 + \epsilon^2(2D_0 D_2 + D_1^2) + \cdots \tag{12}$$

with

$$D_n = \frac{\partial}{\partial T_n}, \quad n = 0, 1, 2, \ldots . \tag{13}$$

The 'stretched' solution is given in the form

$$R_x = R_{x0}(T_0, T_1, T_2, \ldots) + \epsilon R_{x1}(T_0, T_1, T_2, \ldots) \tag{14}$$

$$R_y = R_{y0}(T_0, T_1, T_2, \ldots) + \epsilon R_{y1}(T_0, T_1, T_2, \ldots) \tag{15}$$

$$R_{x0} = \frac{W}{2k}[\exp(i(\phi + \psi)) + \exp(-i(\phi + \psi))] + x_1 \tag{16}$$

$$R_{y0} = \frac{W}{2ki}[\exp(i(\phi + \psi)) - \exp(-i(\phi + \psi)) - i] + y_1 \tag{17}$$

$$R_{x1} = \frac{W}{k}\cos(\phi + \psi) \tag{18}$$

$$R_{y1} = \frac{W}{k}\sin(\phi + \psi). \tag{19}$$

Substitution of equations (14)–(19) into equations (3) and (4) yields, after comparing terms of like power in ϵ, the following set of differential equations for R_{x0} and R_{y0}.

$$D_0^2 R_{x0} + 2\zeta_0\lambda_0 D_0 R_{x0} + \lambda_0^2 R_{x0} = a\omega^2 \cos\phi \tag{20}$$

$$D_0^2 R_{y0} + 2\zeta_0\lambda_0 D_0 R_{y0} + \lambda_0^2 R_{y0} = a\omega^2 \sin\phi - g \tag{21}$$

$$D_0^2 R_{x1} + 2\zeta_0\lambda_0 D_0 R_{x1} + \lambda_0^2 R_{x1} + 2D_0 D_1 R_{x0} + 2\zeta_0\lambda_0 D_1 R_{x0} + \lambda_0^2 R_{x0} U(R_y) = 0 \tag{22}$$

$$D_0^2 R_{y1} + 2\zeta_0\lambda_0 D_0 R_{y1} + \lambda_0^2 R_{y1} + 2D_0 D_1 R_{y0} + 2\zeta_0\lambda_0 D_1 R_{y0} + \lambda_0^2 R_{y0} U(R_y) = 0. \tag{23}$$

Substitution of equations (16)–(19) into the above equations, makes it possible to make use of the expressions given by equations (22) and (23) into equations (20) and (21) respectively. After some modifications, equations of motion can be generated in terms of x_1 and y_1 of equations (16) and (17). Further, with no loss of generality, the lightly damped rotordynamic system can be considered to be operating well above its critical speed and hence ψ can be set equal to $-\pi$. As a result, the following equations of motion can be obtained:

$$D_0^2\left[R_{x0} - \frac{W}{k}\cos(\phi+\psi)\right] + 2\zeta_0\lambda_0 D_0\left[R_{x0} - \frac{W}{k}\cos(\phi+\psi)\right]$$

$$+ \lambda_0^2\left[R_{x0} - \frac{W}{k}\cos(\phi+\psi)\right][1 - U(R_y)] + \epsilon g \sin(\omega T_0)U(R_y) = 0 \tag{24}$$

$$D_0^2\left[R_{y0} - \frac{W}{k}\sin(\phi+\psi)\right] + 2\zeta_0\lambda_0 D_0\left[R_{y0} - \frac{W}{k}\sin(\phi+\psi)\right]$$

$$+ \lambda_0^2\left[R_{y0} - \frac{W}{k}\sin(\phi+\psi) - 1\right][1 - U(R_y)]$$

$$+ \epsilon g[1 - \cos(\omega T_0)]U(R_y) = 0. \tag{25}$$

It may be noted here that

$$R_{x0} - \frac{W}{k}\cos(\phi+\psi) = x_1 \tag{26}$$

$$R_{y0} - \frac{W}{k}\sin(\phi+\psi) = y_1 - \frac{W}{k} \tag{27}$$

by virtue of equations (16) and (17) and further that $U(R_y) \simeq U(R_{y0} - y_1)$ due to the perturbational nature of R_{y1}. However, according to the dynamic stability theory of differential equations with periodic coefficients, the free motion of x_1 and y_1 are of interest and these can be obtained from equations (24) and (25). The periodic nature of the homogeneous differential equations stems from the substitution of the complex Fourier series expansion for $U(R_{y0} - y_1)$ as follows:

$$U(R_{y0} - y_1) = \frac{\xi}{\pi}[1 + f(T_0)] \tag{28}$$

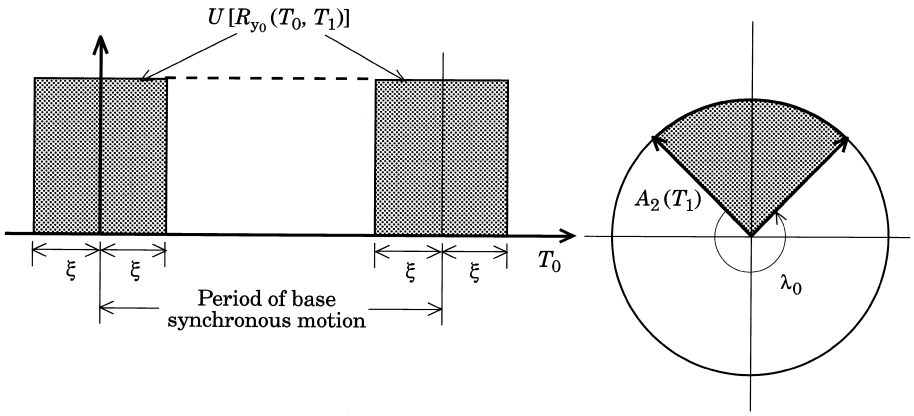

Figure 3. $U(R_{y0})$ associated with base synchronous solution.

$$f(T_0) = 2\sum_{j=1}^{\infty} \frac{c_j}{2}[\exp(ij\omega T_0) + \exp(-ij\omega T_0)] \tag{29}$$

$$c_j = -\frac{i}{2j\xi}[\exp(ij\xi) - \exp(-ij\xi)]. \tag{30}$$

Figure 3 shows the discontinuous nature of this complex Fourier series; positive values of $U(R_{y0} - y_1)$ occur over the rotation angle 2ξ and $U(R_{y0} - y_1)$ is equal to unity over this portion of orbit. So, ξ can be expressed in terms of ϵ. Now, the eigenequation that defines the stability of the rotordynamic system can be written. The 1/2 fractional-frequency motion can be seen to occur in a similar manner and the motion at the parametric resonance frequency can be generated. Equations (24) and (25) can now be rearranged by introducing another scaled time variable,

$$\tau_n = \omega^{n+1} T_0, \quad n = 0, 1, 2, \ldots. \tag{31}$$

As a consequence, the new time derivatives are obtained by:

$$\frac{d}{dT_0} = \frac{d\tau_0}{dT_0}\frac{\partial}{\partial\tau_0} + \frac{d\tau_1}{dT_0}\frac{\partial}{\partial\tau_1} + \frac{d\tau_2}{dT_0}\frac{\partial}{\partial\tau_2} + \cdots = \omega D_{T0} + \omega^2 D_{T1} + \cdots \tag{32}$$

$$\frac{d^2}{dT_0^2} = \omega^2 D_{\tau 0}^2 + \omega^4 D_{\tau 1}^2 + \cdots. \tag{33}$$

After substituting equations (31)–(33) into equations (24) and (25), the stability thresholds can be obtained making use of the theory of ordinary differential equations with periodic coefficients. The frequency equation of the half-frequency whirl that occurs due to the normal-loose parametric excitation can thus be obtained, after some modifications, as [9]

stable solution: $$f(P, q, \zeta) = (P + \zeta^2)^2 - q^2 + 4\zeta^2(1 + P) > 0 \tag{34}$$

where

$$P = \left(\frac{\omega}{2\lambda^2}\right) - 1, \quad q = \epsilon\xi/\pi, \quad \zeta = \frac{\zeta_0\lambda_0}{\sqrt{\lambda_0^2(1-q)}}. \tag{35}$$

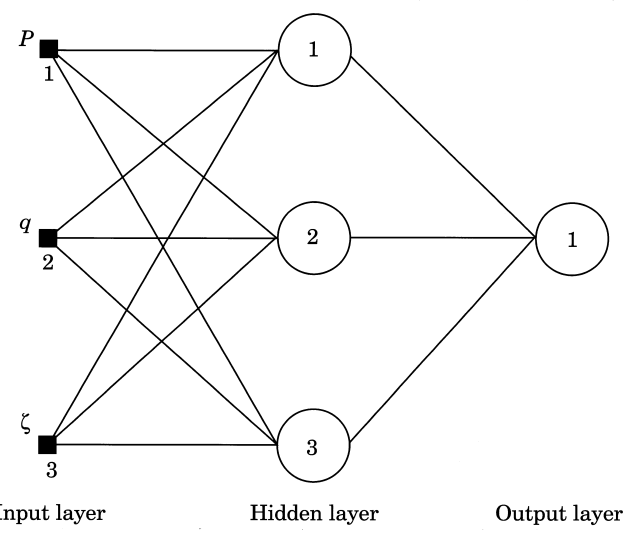

Figure 4. Hidden layer network for identification of rotor system.

It may be noted here that the wider range of physical parameters of commonly-used rotor systems actually results in a narrow region of these transformed parameters. So, equations (34) and (35) are used here to define the identification problem with respect to the stability of the rotordynamic system that is being considered.

3. IDENTIFICATION OF SYSTEM STABILITY USING NEURAL NETWORKS

3.1. NEURAL NETWORK DESIGN

In Appendix B, general details of network design, training and implementation are given. Now, the details of the neural networks, that have been designed here to identify the status of the rotordynamic system with respect to stability, are given. Two types of neural networks are used and compared here: (i) hidden layer network which is illustrated in Fig. 4 and which has one hidden layer between input and output layers; (ii) functional-link network in which the input layer has been enhanced by additional input nodes as shown in Fig. 5. In both the types of networks, the number of output nodes is taken to be 1, so as to indicate whether the system is stable or unstable.

From the stability analysis of the rotordynamic system presented in Section 2, it can be seen that the variations of imbalance vector magnitude a, rotating speed ω, mass m, stiffness k and linear damping coefficient c, have pronounced nonlinear effects on the system stability status. But from equation (34), it can be seen that the complexity of the relationship between the transformation parameters P, q and ζ and the stability of the rotordynamic system is of lower intensity. So from the viewpoint of simplifying the structure of the neural network, these three transformation parameters are directly selected to constitute the input nodes of the hidden layer network shown in Fig. 4, and the basic portion of input nodes of the functional-link network shown in Fig. 5. Considering the effective elements of the functional-link expansion which is based on these three parameters, six input nodes are chosen as P, ζ^2, q^2, P^2, ζ^4 and $P\zeta^2$ as shown in Fig. 5.

Usually the number of hidden nodes in the hidden layer network can not be decided directly. The recommended approximate value is given as:

$$N_j = \sqrt{N_i^2 + N_k^2} = \sqrt{10} \tag{36}$$

where N_j, N_i and N_k are respectively, the number of hidden, input and output nodes; here $N_i = 3$ and $N_k = 1$. Now from equation (36) the number of hidden nodes can be taken to be either 4 or 3 or 2 which are the nearest integer values. The optimum number of hidden nodes is now obtained by the comparison of simulation results.

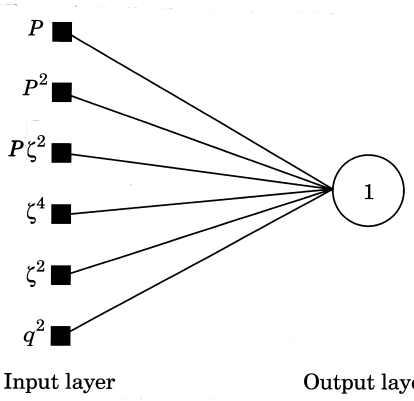

Figure 5. Functional-link network for identification of rotor system.

3.2. NETWORK IMPLEMENTATION

3.2.1. *Pre-processing of the data*

To achieve the best performance with respect to the desired accuracy, a pre-processing that consists of a normalisation scheme is carried out herein, before the process of network training is started. From equation (B5), it can be seen that ΔW_{kj} is equal to zero when $x_j = 0$; Also, $x_k = 0$ or $x_k = 1$ makes δ_k of equation (B6) to be equal to zero. Similarly, δ_j of equation (B8) and ΔW_{ji} of equation (B7) are equal to zero when x_i and x_j are equal to zero; Even when $x_i \neq 0$ and $x_j = 1$, $\delta_j = 0$. This means essentially that if the node value is equal to zero in any layer or equal to unity in the output or hidden layers, the network training is terminated but the desired accuracy has not been achieved. Thus the training is incomplete indeed. So, in order to avoid this situation, for any input node i, a normalised value x_i^1 is assigned instead of x_i, and its range is restricted to the closed region [0.1, 0.9]. The relevant equation of transformation is given by:

$$x_i^1 = 0.1 + 0.8*(x_i - x_{\min})/(x_{\max} - x_{\min}) \tag{37}$$

where $x_{\max}$ and $x_{\min}$ are, respectively, the maximum and minimum values of x_i at the input node i. For the output node k, the expected value x_{ke} is fed as the following rules:

$$\text{If} \quad f(P, q, \zeta) \leqslant 0, \quad x_{ke} = 0.1, \text{ system is deemed to be unstable;} \tag{38}$$

$$\text{If} \quad f(P, q, \zeta) > 0, \quad x_{ke} = 0.9, \text{ system is deemed to be stable.} \tag{39}$$

As a result of learning errors of the network, the actual values of the output nodes cannot be exactly equal to 0.1 or 0.9. A threshold value is commonly imposed when a similar difficulty is encountered in the problems of classification. Here, the median is used as the threshold value. Hence, the output node k is subjected to the following rules instead of equations (38) and (39).

$$\text{If} \quad x_k \leqslant 0.5, \quad \text{it means the system is probably unstable;} \tag{40}$$

$$\text{If} \quad x_k > 0.5, \quad \text{it means the system is probably stable.} \tag{41}$$

3.2.2. *Comparison of hidden layer and functional-link neural networks*

In order to compare the performance of the two types of neural networks, a practical rotor system shown in Table 1 is used here to provide the numerical data. Five different values of the imbalance vector magnitude are considered and used to train both the neural networks separately. Considering the practical situation, parameter P is selected from within the closed region [0.01, 0.49] and the step size is taken to be 0.02. A modest value of ϵ lies in the closed interval [0.01, 0.25] in actual situations and this region is used here with the step size equal to 0.01. So, for a rotordynamic system with a given set of parameter values, 625 sample data can be generated, wherein parameters q and ζ can be obtained from equation (35). Out of these, 525 data points are used to train the network. The test

TABLE 1

Rotor systems with different magnitudes of imbalance vectors
(m)

a_1	a_2	a_3	a_4	a_5
0.05	0.08	0.11	0.14	0.17

$m = 300$ kg, $k = 2.9 \times 10^7$ N/m, $c = 5 \times 10^3$ N-s/m.

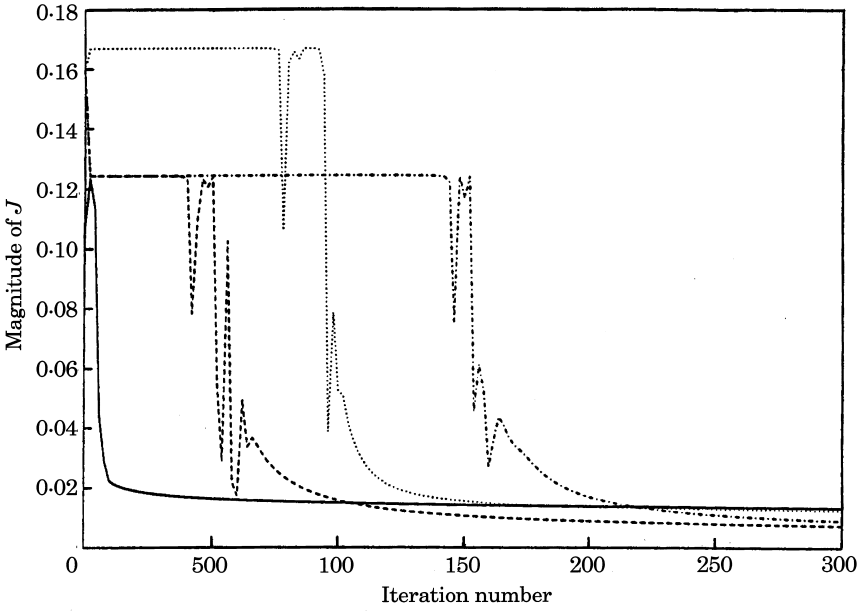

Figure 6. Distribution of error for the four types of networks. ——, Functional link; – – –, two hidden nodes; ·–·–, three hidden nodes; · · · ·, four hidden nodes.

data consists of the remaining 100 points. The initial weights are given by a group of random data and are limited in the region $(-1, 1)$. The networks are trained using the backpropagation algorithm (programmed in C language and run on PC-486). The maximum iteration number of training is limited to 5000 and the learning and momentum coefficients are taken to be 0.15 and 0.075 respectively.

To assess the suitability of both the functional-link and hidden layer networks, the speed of convergence (during network training) and the number of mismatch points are used. Mismatch points are those that represent the case in which the identification carried out by the trained network is in contrast to the actual status of the rotor system. So the number of mismatch points is a measure of the identification accuracy. The magnitudes of error function J of equation (B3) are observed at different stages of iteration and are plotted in Fig. 6, for the four networks and for the imbalance vector a_1. Asymptotic values of J are observed, in all the cases, around the iteration number 3000. The number of mismatch points for all the four networks are listed in Table 2, and these are based on the updated weights obtained after 525 data points were fed. Among the four different structures of neural networks that are compared in Table 2, first three networks belong to the type of

TABLE 2

Number of mismatch points

Data type		Training data (525 sets)					Test data (100 sets)				
Magnitude of imbalance vectors		a_1	a_2	a_3	a_4	a_5	a_1	a_2	a_3	a_4	a_5
Hidden layer network	$N_j = 2$	5	4	4	3	4	0	2	2	2	2
	$N_j = 3$	5	5	4	21	6	0	0	2	4	4
	$N_j = 4$	5	4	4	19	5	0	0	2	3	4
Functional-link network		2	0	0	2	2	2	1	1	0	1

TABLE 3

Rotor systems with different values of damping coefficients (N-s/m)

$c_1 \times 10^3$	$c_2 \times 10^3$	$c_3 \times 10^3$	$c_4 \times 10^3$	$c_5 \times 10^3$
4.00	4.50	5.00	5.50	6.00

$m = 300$ kg, $k = 2.9 \times 10^7$ N/m, $a = 0.11$ m.

TABLE 4

Rotor systems with different values of compliance (N/m)

$k_1 \times 10^7$	$k_2 \times 10^7$	$k_3 \times 10^7$	$k_4 \times 10^7$	$k_5 \times 10^7$
2.30	2.60	2.90	3.20	3.50

$m = 300$ kg, $c = 5 \times 10^3$ N-s/m, $a = 0.11$ m.

hidden layer network and are with different number of hidden nodes i.e. 2, 3 and 4. The last one is a functional-link neural network. Now, from Table 2 and Fig. 6, it can be observed that the functional-link network is more effective than the hidden layer networks in both the aspects of convergent speed of the learning process as well as the identification accuracy. It can also be seen that adding more hidden nodes did not result in an increase in the identification accuracy and so the optimum number of hidden nodes is equal to two.

3.3. STABILITY STATE CLASSIFIER

In order to form a stability status indicator for any rotordynamic system, this functional-link network is further trained using different sets of system parameters. Ideally the classifier should be trained using the parameter sets of all industrial rotodynamic systems that are in use but this is not feasible. Further, awaiting the case histories of rotors that fail due to bearing clearance effects is also of no use. However, consideration of only the transformation parameters P, q and ζ instead of the system parameters can supply a convenient way to solve this problem, since these three parameters are limited to a small region, in many situations, wherein the concerned system is probably unstable. Based on the rotor system parameters that are listed in Table 1, different rotor systems are generated that are different with respect to the parameters c, k and m and are given in Tables 3–5. From these, the regions of parameters q and ζ can be obtained from equation (35) as (0.06, 0.25) and (0.01, 0.07) respectively. In order to get more adaptivity to different rotor systems, the sets of training data are obtained by a uniform variation of the parameters P, q and ζ. Considering the limitation of the computer memory, the step size of each parameter change is taken as follows: $\Delta P = 0.025$, $\Delta q = 0.02$ and $\Delta\zeta = 0.007$, So, 900 number of simulation data are obtained to design the classifier and Table 6 gives the final weights of the neural network. Twenty sets of parameter values a, c, m and k are now generated from a rotordynamic system in practice [23], to assess the reliability of this

TABLE 5

Rotor systems with different values of inertia (kg)

m_1	m_2	m_3	m_4	m_5
260	280	300	320	340

$k = 2.9 \times 10^7$ N/m, $c = 5 \times 10^3$ N-s/m, $a = 0.11$ m.

TABLE 6

Connection weights of the classifier network

Input node i	1	2	3	4	5	6
Output node $k = 1$	11.70	14.36	51.25	6.52	0.35	1.48
Bias parameter			$\theta_k = -10.58$			

classifier. Also, parameters P and ϵ are taken to be in the closed regions [0.01, 0.235] and [0.01, 0.25] respectively. ΔP and $\Delta\epsilon$ are taken to be 0.025 and 0.01 respectively. The percentage of mismatch points with respect to the total test data is plotted in Fig. 7; different values of parameters a_i, c_i, k_i and m_i that are shown in Tables 1, 3–5, are represented by an index i. From this figure, it can be seen that the accuracy of identification is more than 86%. This result confirms that this neural network can now be considered as a reliable classifier of rotor systems with respect to their stability.

3.4. PARAMETER ADJUSTMENT TO OBTAIN THE STABLE SYSTEM

Passive means of converting an unstable system into a stable one include the possibility of driving the unstable system into the stable region through the proper adjustment of system parameter values. The classifier network designed above can also be used to serve as an 'advisory' system, to evaluate the optimum way of such a parameter adjustment. The relevant procedure is now demonstrated. A rotordynamic system with $m = 300$ kg, $k = 2.9 \times 10^7$ N/m, $c = 5 \times 10^3$ N-s/m and $a = 0.1$ m, is now considered for this purpose. ϵ is taken to be 0.2 and the rotor rotating speed is taken to be 5350 r/min that is 560 rad/s which is close to 2λ. From equation (35), parameters P, q and ζ can be obtained. Now, the values of c and k are increased and the values of m is decreased in uniform steps using the step sizes $\Delta c = 100$ N-s/m, $\Delta k = 1 \times 10^6$ N/m and $\Delta m = 1$ kg, separately. Final values of c, k and m that constitute a stable system are identified by the classifier as 1.88×10^4 N-s/m, 2.4×10^7 N/m and 353 kg, respectively.

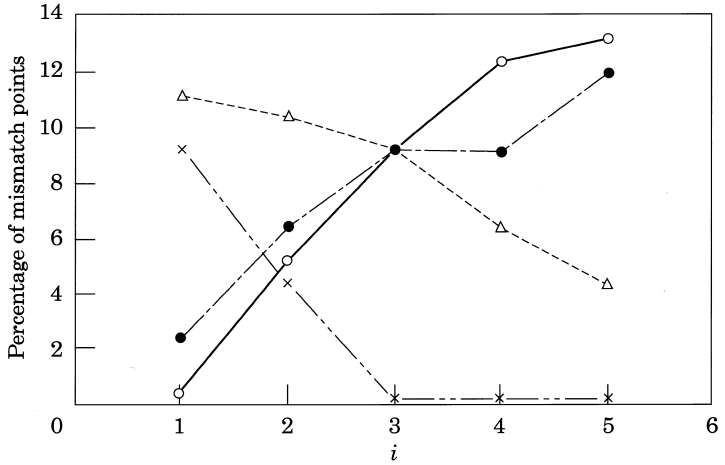

Figure 7. Reliability of the classifier network in terms of mismatch points. ○, a-varying; ×, c-varying; ●, k-varying; △, m-varying.

4. CONCLUSIONS

A classifier neural network that can be used in the automation of condition monitoring and fault detection procedures of rotordynamic systems has been designed. In machinery diagnostics, it is well known that it is impractical to implement comprehensive and collective sets of faults, for network training. Further, natural development of these faults which leads to irreversible accumulated damage, should be avoided by making use of symptoms in the vibratory signal. Here, taking into account these facts, an analytical model is developed that qualifies and quantifies the effect of bearing clearances in terms of system stability characteristics. The method of multiple scales is used and suitable modifications are carried out, to develop a set of inequalities that can be used as rules of identification during network training. Keeping in view the non-linear dependence of the stability characteristic on the variations in system parameters, the hidden layer and functional-link architectures are designed. A reasonable selection of input nodes, optimum number of hidden nodes, nodes needed for input enhancement of functional-link networks, etc. is carried out based on the results of the analytical modeling of the rotordynamic system. The adaptivity and sensitivity of functional-link network in adequately modeling the complex non-linear relationships between the stability state and parameters of the rotor system are brought out. The success rate of stability status identification is very high, on points that fall in the region of both training data as well as test data. This shows the potentials of the network that operates on the transformed system parameters rather than the physical parameters. In practical situations, proper values of damping coefficient and imbalance vector magnitude that result in a stable system can be obtained by adjusting the oil temperature and the process of machine assembling in maintenance. But, traditional control strategies could not adequately make use of the complex relationship that exists between these two parameters and the stability characteristic of the system. The classifier neural network developed herein, by the way of its adaptivity, is able to exploit these non-linear relationships so that driving the system to a stable state is easier and practical. Further, as and when the real case studies or the experimental data are available, this network can be updated by using this additional information for further training.

ACKNOWLEDGEMENTS

The investigation presented here was supported by a special grant from the Government of Canada under the IRIS network of Centers of Excellence. Also, the support from Canadian International Development Agency to the second author is gratefully acknowledged.

REFERENCES

1. A. D. Dimarogonas and A. S. Paipetis 1983 *Analytical Methods in Rotordynamics*. London: Applied Science.
2. J. M. Vance 1988 *Rotordynamics of Turbomachinery*. New York: John Wiley.
3. F. F. Ehrich 1992 *Handbook of Rotordynamics*. New York: McGraw Hill.
4. F. F. Ehrich 1966 *ASME Paper no. 66-MD-1, Design Engineering Conference and Show, Chicago, Illinois*. Harmonic vibrations of rotors in bearing clearance.
5. F. F. Ehrich 1988 *ASME Transactions Journal of Vibration and Acoustics* **110**, 9–16. High order subharmonic response of high speed rotors in bearing clearance.
6. F. F. Ehrich 1991 *ASME Transactions Journal of Vibration and Acoustics* **113**, 50–57. Some observations of chaotic vibration in high speed rotor dynamics.
7. F. F. Ehrich and J. J. O'Connor 1967 *ASME Transactions Journal of Engineering for Industry* **89**, 381–390. Stator whirl with rotors in bearing clearances.

8. D. Bently 1979, *ASME Paper No. 74-PET-16*, *Petroleum Mechanical Engineering Conference*, *Dallas*, *Texas*. Forced subrotative speed dynamic action of rotating machinery.
9. D. W. Childs 1982 *ASME Transactions Journal of Engineering for Power* **104**, 533–541. Fractional-frequency rotor motion due to nonsymmetric clearance effects.
10. A. Muszynska 1984 *Institution of Mechanical Engineers*, *Paper No. C281/84*. Partial lateral rotor to stator rubs.
11. R. F. Beatty 1985 *ASME Transactions Journal of Vibration*, *Acoustics*, *Stress and Reliability in Design* **107**, 151–160. Differentiating rotor response due to radial rubbing.
12. S. Saito 1985 *ASME Paper No. 85-DET-33*. Calculation of nonlinear unbalance response of horizontal jeffcott rotors supported by ball bearings with radial clearances.
13. Y. S. Choi and S. T. Noah 1987 *ASME Transactions Journal of Vibration*, *Acoustics*, *Stress and Reliability in Design* **109**, 255–261. Nonlinear steady-state response of a rotor-support system.
14. R. Ganesan and T. S. Sankar 1993 *Dynamics and Vibration of Time-varying Systems and Structures 1993*, *ASME-DE* **56**, 295–301. 14th Biennial Conference on Mechanical Vibration and Noise, Albuquerque, USA. Non-stationary vibrations of rotordynamic systems with bearing clearances.
15. F. F. Ehrich and D. W. Childs 1984 *Mechanical Engineering* **106**, 66–79. Self-excited vibrations in highperformance turbomachinery.
16. S. Yamauchi 1983 *Transactions of the Japanese Society of Mechanical Engineers* **49**, 1862–1868. The nonlinear vibration of flexible rotors, 1st report, development of a new analysis technique.
17. M. C. Ek 1980 *AIAA Journal of Spacecraft and Rockets* **17**, 208–218. Solution of the subsynchronous whirl problem in the high-pressure hydrogen turbomachinery of the space shuttle main engine.
18. D. W. Childs 1978 *ASME Transactions Journal of Engineering for Power* **100**, 48–57. The space shuttle main engine high pressure fuel turbopump instability problem.
19. S. W. Shaw 1985 *Journal of Sound and Vibration* **99**, 199–212. Forced vibration of a beam with one sided amplitude constraint: theory and experiment.
20. D. E. Rumelhart and J. L. Mc Clelland 1988 *Parallel Distributed Processing*, Vols 1 and 2. Cambridge, MA: MIT Press.
21. L. B. Booker, D. E. Goldberg and J. H. Holland 1989 *Artificial Intelligence* **40**, 235–282. Classifier systems and genetic algorithms.
22. A. H. Nayfeh and D. T. Mook 1979 *Nonlinear Oscillations*. New York: Wiley Interscience.
23. N. F. Rieger 1984 *Dynamics of Rotors*, *Stability and System Identification* (O. Mahrenholz, Ed.), pp. 99–140. Stability of rotors in bearings. New York: Springer.
24. Y. H. Pao 1989 *Adaptive Pattern Recognition and Neural Networks*. New York: Addison-Wesley.
25. F. Rosenblatt 1962 *Principles of Neurodynamics*. New York: Spartan.

APPENDIX A: NOMENCLATURE

a	imbalance vector magnitude
A	steady-state synchronous motion amplitude defined in equation (6)
c	linear damping coefficient
$f(T_0)$	Fourier series defined in equation (29)
g	gravitational acceleration
k	net circumferentially symmetric rotor-bearing radial stiffness
m	rotor mass
P	fractional change in excitation frequency defined in equation (35)
q	a function of base synchronous motion defined in equation (35)
R_x, R_y	components of the rotor displacement vector R along x- and y-directions
$U(\ )$	unit step function
W	rotor weight
ϵ	fractional loss in rotor's radial stiffness due to non-linearity
ϕ	rotor rotation angle
ζ_0	damping factor; $\zeta_0 = c/[2(mk)^{1/2}]$
λ_0	base rotor natural frequency; $\lambda_0 = (k/m)^{1/2}$
λ	frequency resulting from stiffness loss due to non-linearity
ζ	damping factor including clearance effects; $\zeta = (\lambda_0\zeta_0)/\lambda$
ψ	phase angle of steady-state synchronous solution defined in equation (7)
ω	rotor rotating speed

$f(\text{net}_j)$	non-linear activation function defined in equation (B2).
J	measure of learning error of the network
N_i, N_j, N_k	the number of input nodes, hidden nodes and output nodes
net_j	weighted input to neuron j
W_{ji}	connection weight between input node i and hidden node j
W_{kj}	connection weight between hidden node j and output node k
$\Delta W_{kj}, \Delta W_{ji}$	updated weights of W_{kj} and W_{ji}
x_i, x_j, x_k	values of the input node, hidden node and output node
x_i^1	normalised value of x_i defined in equation (37)
x_{ke}	expected value of the output node
α	momentum coefficient
θ_j	bias parameter defined in equation (B2)
η	learning coefficient

APPENDIX B

B.1. ASPECTS OF NEURAL COMPUTING

A neural network is a massively parallel, self-adaptive, interconnected network of basic elements called neurons. These basic elements loosely model the mechanics of computations in the neurobiological processes of the brain. Neurons possess very simple computational abilities but the interactions between them do allow for parallel processing of information that is inputted as a set of stimuli to the array of input processors. The functions of these neurons are (i) to qualify the input signals by evaluating their intensity, (ii) to compute a weighted sum of these inputs using connection weights thereby comparing it with the threshold values associated with individual neurons, and (iii) to fire an output signal if the strength of the input signal builds up to the threshold level or, as in many cases, to obtain an output signal by processing the weighted sum of the input signal through an activation function. The sequence of operations is given below (see Fig. B1):

(i) First, the input signals are multiplied by the connection weights w_{ji} and the effective input to the element j is the weighted sum of these inputs given by:

$$\text{net}_j = \sum_{i=1}^{n} W_{ji} \times \text{input}_i = \sum_{i=1}^{n} W_{ji} x_i. \tag{B1}$$

(ii) Any one of the three functions, among sum function, activation function and learning function, is executed depending upon the intended use. Sum function compares the net_j of equation (B1) with the threshold value of the neuron with which it is associated, to determine the presence of the output signal and its level too. Activation function

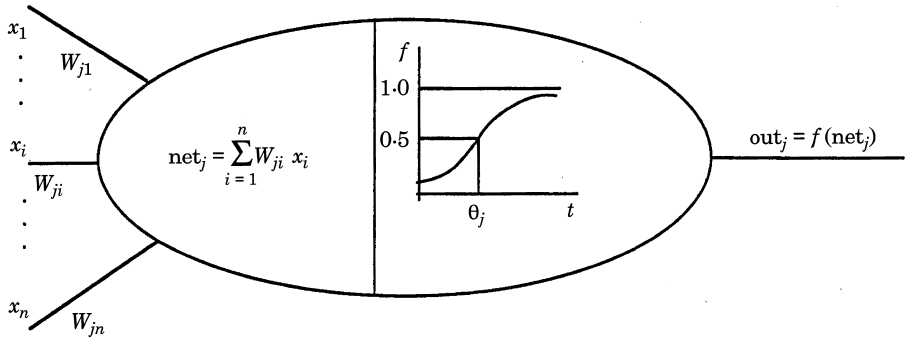

Figure B1. A single processing element.

processes the net_j through a mathematical function with asymptotic behaviour; the commonly-used function is sigmoid function and is given by

$$\text{out}_j = f(\text{net}_j) = 1/[1 + \exp(-\text{net}_j + \theta_j))] \tag{B2}$$

where θ_j is a bias parameter that is used to modulate the output of the neuron. Learning function is usually associated with a neuron that is provided with a simple form of memory; it makes it possible to store the results of previous operations to obtain a sort of learning behaviour.

As mentioned earlier, the significant computational superiority of a neural network is derived by interconnecting a large number of neurons into a single network. The overall performance is dictated by the scheme of connectivity. Networks that are not designed to feedback the output of a layer to its own input or to the output of preceeding layers do not have a memory. The feedforward networks which have a layer of neurons to which the external stimuli are presented, a series of hidden layers (in some cases) and a layer of neurons at which the output is made available, have three distinct forms of architecture:

Flat networks: Here inputs are directly mapped into the output. These are fully-connected networks with each and every input neuron influencing all output neurons. These networks do not adequately model the non-linear input/output relationship and the only way of introducing the non-linearity is through the use of activation functions.

Hidden-layer networks: Here the external stimuli are received by the input layers and the output layer transmits a signal as the response to this input, thus not allowing the hidden layers to interact with the external environment but only to accept or output information within the overall system (see Fig. B2). If the nonlinear activation functions are not used, it is possible to convert a multilayer linear network into an equivalent single layer network.

Functional-link networks: Relatively recent, this network is similar to flat networks but has additional input nodes wherein the non-linear transformations of the basic input quantities are performed (see Fig. B3). These are more suitable for realistic applications, and the learning times are far more economical than those of multilayer networks [24].

Once a suitable architecture is frozen, the network is trained so as to produce an acceptable output for a given input pattern. Various methods of network training are

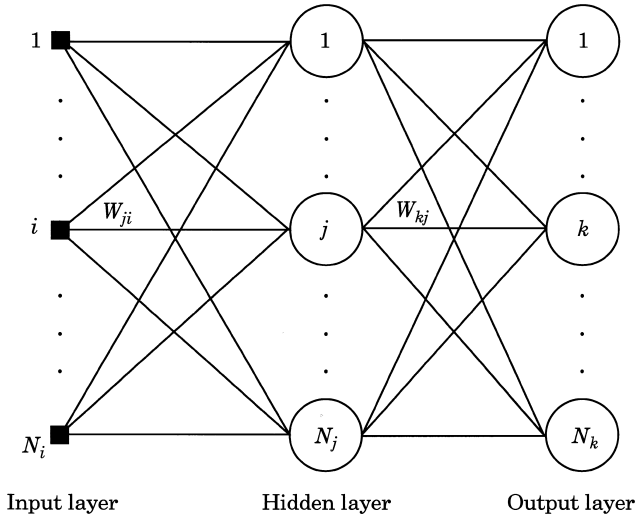

Figure B2. Hidden layer architecture.

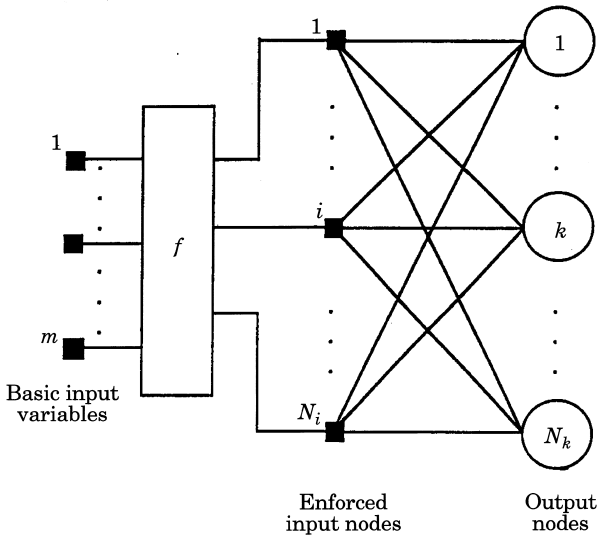

Figure B3. Functional-link architecture.

available in the category of supervised learning techniques and for clarity, the essential details of the backpropagation approach [15, 20] are given below.

With the steepest gradient-descent algorithm, the backpropagation approach is fundamentally looking for minimising the mean square error between the actual network output and the expected output signal. The energy function to be minimised is expressed as:

$$J = \frac{1}{2} \sum_{k=1}^{N_k} [x_{ke}(t) - x_k(t)]^2. \tag{B3}$$

Here, J represents the measure of the error in the input/output pattern at time t. x_{ke} and x_k are respectively the expected and actual outputs at the output node k, and further N_k is the number of output nodes. This energy function must be minimised by making use of partial derivatives with respect to all unknown parameters. So, the weight W_{kj} between the hidden node j and the output node k can be updated by ΔW_{kj} as

$$\Delta W_{kj} = -\eta \frac{\partial J}{\partial W_{kj}}. \tag{B4}$$

Here, the parameter, η, called the learning coefficient, determines how quickly the final weights can be obtained. From equations (B1)–(B4) the updated value of the weights ΔW_{kj} can be calculated as follows:

$$\Delta W_{kj} = \eta \delta_k x_j \tag{B5}$$

$$\delta_k = (x_{ke} - x_k) x_k (1 - x_k) \tag{B6}$$

where x_j is the output of the hidden node j. Similarly, the weight W_{ji} that corresponds to the input node i and the hidden node j can be updated by ΔW_{ji} as follows:

$$\Delta W_{ji} = \eta \delta_j x_i \tag{B7}$$

$$\delta_j = x_j (1 - x_j) \sum_{k=1}^{N_k} (\delta_k W_{kj}) \tag{B8}$$

where x_i is the value of the input node i. The weights are updated layer by layer moving backwards from the output layer to the input layer during the learning process of the

neural network, and thus this learning algorithm is called the backpropagation algorithm. Further, in order to avoid the oscillation of the weights during the learning process, a momentum term is introduced into the rule that computes the set of updated weights, as

$$\Delta W_{kj}(n+1) = \eta \delta_k x_j + \alpha \Delta W_{kj}(n) \tag{B9}$$

$$\Delta W_{ji}(n+1) = \eta \delta_j x_i + \alpha \Delta W_{ji}(n) \tag{B10}$$

where α is the momentum coefficient. It is used to damp out the high frequency variation in the backpropagation error signal. Further, n is an index that gives the iteration number of the updated weight. So the weights can be iteratively updated by

$$W_{kj}(n+1) = W_{kj}(n) + \Delta W_{kj}(n+1) \tag{B11}$$

$$W_{ji}(n+1) = W_{ji}(n) + \Delta W_{ji}(n+1) \tag{B12}$$

until the learning error is reduced to an acceptable level or the given iteration number is achieved.

Obviously, the backpropagation method is also suitable for the supervised training process of the functional-link neural network, where the process of weight updating is required only between input and output layers. However, it may be noted that the nature of the gradient descent method leads to a low rate of learning near local minimum, which increases the number of iterations, when a higher value of accuracy is prescribed.