Diyala Journal of Engineering Sciences

ISSN 1999-8716
Printed in Iraq

Second Engineering Scientific Conference
College of Engineering –University of Diyala
16-17 December. 2015, pp. 557-567

REAL TIME MACHINERY SAFETY MONITORING BY HYBRID FUZZY SYSTEM

Hussein S. Radhi
Assistant Lecturer, Dep. of Computer & Software College of Engineering-University of Diyala,
E mail: hussein1957.hs@gmail.com

ABSTRACT: - Rotary machineries are basic components used in many industrial applications. A simple examples uses these machines are simple electric fans while complicated uses these devices included in missile and aircrafts.

The aim of this research is to develop an *Intelligent Neuro-Fuzzy* control algorithm to guide and control accurately these rotating devices so that it will detect precisely the faulty components during the operation or in both preventive and corrective maintenance procedures. This online intelligent monitoring algorithm classifies status of the device into three indicators :(*Safe, Possibly damaged or Damaged*). The processes of classification consider all input variables with high effectiveness on diagnostic output state. The contribution of this paper is to introduce novel intelligent monitoring and controlling algorithm that uses (*Two*) predicted indices instead of one as in the previous research articles.

1. INTRODUCTION

Diagnoses of the faults are a series of processes with two steps: features extraction representation and pattern classification. Features extraction is a mapping process from the measured signal space to the feature space. Representative features associated with the safe condition of a machinery component (or subsystem) are extracted by using appropriate signal processing techniques. Pattern classification is the process of classifying the characteristic features into various categories. Traditional approach widely used in industry, relies on human expertise to relate the vibration features to the faults. This method, however, is boring and not always reliable when the extracted features are contaminated by noise. Also, it is difficult for a diagnostician to deal with the confused properties if multiple features are used. Data-driven diagnostic classification can be performed by reasoning tools such as neural networks, fuzzy logic, and neural fuzzy synergetic schemes.

There are some techniques have been proposed for machinery condition monitoring, it still remains a challenge in implementing a diagnostic tool for real- world monitoring applications because of the complexity of machinery structures and operating conditions. When a monitoring system is used in real-time industrial applications, two major problems will appears (missed alarms and false alarms).

To overcome these challenges is the objective of this research work to develop a new technique, an integrated classifier, for real-time condition monitoring in, systems. In this novel classifier, the monitoring reliability is enhanced by integrating the information of the object's future states forecast by a multiple-step predictor; furthermore, the diagnostic scheme is adaptively trained by a novel recursive hybrid algorithm to improve its convergence and adaptive capability.

The procedure as follows: description of integrated classifier, where the multiple-step predictor and monitoring indices are described then discusses the hybrid online training algorithm. At last the viability of the proposed integrated classifier is verified by experimental tests corresponding to different gear conditions.

2. SYSTEM CLASSIFIER

The objective is to introduce a diagnostic classifier that integrate the most important features obtained by choosing an appropriate signal processing techniques. The goal is to make a more positive evaluation of the health condition of the mechanical parts of interest. The diagnostic reliability in this suggested classifier will be enhanced by implementing the predicted states of the object's conditions. The forecasting in this integrated classifier is performed fig (1) for input variables so as to make it easier to track the error sources in diagnostic operations.

The developed classifier will be able to facilitate the incorporation of diagnostic knowledge from expertise and to extract new knowledge in operations by training. The diagnostic classification is performed by fuzzy logic whereas an adaptive training algorithm is used to match the fuzzy system parameters and structures. The conditions of each machinery part is classified into three categories: healthy (C1), possible damage (C2), and damage (C3), respectively. {x1, x2, …, xn} are the input variables at the current time step. Three membership functions (MFs), small, medium, and large, are assigned to each input variable with the initial states as shown in Fig. 1 where the fuzzy completeness (or the minimum fuzzy membership grade) is at 50%.

The diagnostic classification, in terms of the diagnostic indicator y, is formulated in the following form:

Rule: *(x1 is A1 j), (x2 is A2 j), . . . and (xn is Anj) ⇒ (y ⊂ Sj with wj)* ……..(1)

where Aij are MFs; i = 1, 2, …, n, j = 1, 2, …, m, m denotes the number of rules; Sj represents one of the states C1, C2 or C3, depending on the values of the diagnostic indicator. When multiple features (input indices) are employed for diagnostic classification, the contribution of each feature association to the final decision depends, to a large degree, on the situation under which the diagnostic decision is made. Such a contribution is characterized by a weight factor ***wj*** which is related to the feature association in each rule. The initial values of these rule weights are chosen to be unity; That is, all input state variables have initially assumed to have identical importance or robustness to the overall diagnostic output.

Similarly, the diagnostic classification based on the *first predicted monitoring indices*, { x'_1, , x'_2 ,…, x'_n }, is formulated as:

Rule: ***(x'_1 is A1 j) ,(x'_2 is A2 j), … and (x'_n is Anj) ⇒ (y' ⊂ Sj with wj)*** ……. (2)

where y′ is the diagnostic indicator based on forecast input variables.

Also, the diagnostic classification based on the *second predicted monitoring indices*, {x''_1, x_2'',…..x''_n}, is formulated as:

Rule: ***(x''_1 is A1 j) ,(x''_2 is A2 j), … and (x''_n is Anj) ⇒ (y'' ⊂ Sj with wj*** ……. (3)

The number of rules is associated with the diagnostic reasoning operations of input state variables. In general, if all monitoring indices are small, then the object is considered healthy (C1). Otherwise, the object is possibly damaged. In this case, the diagnostic classification indicator y represents faulty condition only. Different feature rule corresponds to a different confidence grade wj in diagnosis. Fig. 2 schematically shows the network architecture of this integrated classifier. Unless specified, all the network links have unity weights.

The input nodes in layer 1 transmit the monitoring indices {x1, x2, …, xn} or their forecast future values{ x'_1, , x'_2 ,…, x'_n }& {x''_1, x_2'',…..x''_n} to the next layer. These three sets of monitoring indices are input to the network and processed separately.

Each node in layer 2 acts as a MF, which can be either a single node that performs a simple activation function or multilayer nodes that perform a complex function. The nodes in layer 3 perform the fuzzy T-norm operations. If a product operator is used, the firing strength of rule is:

$$\eta=\prod_{i=1}^{n} \boldsymbol{A}_{ij}(x_i) \qquad ……………(4)$$

$$\eta'=\prod_{i=1}^{n} \boldsymbol{A}_{ij}(x'_i) \qquad ……………(5)$$

$\eta''=\prod_{i=1}^{n} \boldsymbol{A}_{ij}(x''_i)$ ……………(6)

Where Aij denote MF grades showing in fig(2).

The process of defuzzification is achieved in layer 4. By normalization, the faulty diagnostic indicator will be:

$$y=\frac{\sum_m \eta\omega}{\sum_m \eta} \quad \text{……………..} (7)$$

Similarly, the fault diagnostic indicator based on forecast inputs will be

$$y'=\frac{\sum_m \eta'\omega}{\sum_m \eta'} \quad \text{…………….}(8)$$

$$y''=\frac{\sum_m \eta''\omega}{\sum_m \eta''} \quad \text{…………….}(9)$$

The states of the diagnostic indicator y ,y' and y'' are further classified into three categories(NOTE: classification in normal cases each class takes ($\frac{1}{3}$), but for more accuracy, we suppose as follows:

If $0 \le y \le 0.25$ → Healthy (C1)

If $0.25 \le y \le 0.50$ → possibly damaged (C2)

If $0.50 \le y \le 1$ → damaged (C3)

The final decision of the part is made by:

A- If ($y\subset$ C1 & $y'\subset$ C1 & $y''\subset$ C1) or ($y\subset$ C2 & $y'\subset$ C1 & $y''\subset$ C1) then healthy (C1)

B- If ($y\subset$ C3 & $y'\subset$ C3 & $y''\subset$ C3) or ($y\subset$ C2 & $y'\subset$ C3 & $y''\subset$ C3) then damaged (C3)

C- Otherwise possibly damaged.

3- ESTIMATION OF INDICES

3.1 Monitoring indices

The most popular machinery defects are because of transmission systems (gears and bearings). The simulation example used in our work is the gears to illustrate how to apply the proposed integrated classifier for machinery condition monitoring.

In operations, the fault diagnosis of a gear train is conducted gear by gear. Because the measured vibration is an overall signal contributed from various vibratory sources, the first step is to differentiate the signal specific to each gear of interest by using a synchronous average filter. By this filtering process, the signals which are non-synchronous to the rotation of the gear of interest are filtered out. As a result, each

gear signal is computed and represented in one full revolution called average of signal which will be used for advanced analysis by other signal processing techniques.

Several techniques have been proposed in the literature for gear fault detection. Due to the complexity in the machinery structures and operating conditions, each fault detection technique has its own advantages and limitations, and is efficient for some specific application only. Consequently, the selected features for fault diagnostics should be reliable, that is, sensitive to component defects but immune to noise. In this case, three features from the information domains of energy, amplitude, and phase are employed for the diagnosis operation:

A. Wavelet energy function, using the overall residual signal which is obtained by bandstop filtering out the gear mesh frequency (fRN)and its harmonics, where fR is the rotation frequency (in Hz) of the gear of interest and N is the number of teeth of the gear.

B. Phase demodulation, using the average of signal.

C. Beta kurtosis, using the overall residual signal. Knowing that the beta kurtosis is the normalized fourth moment of a signal, in terms of the beta function instead of a generally used Gaussian function.

The monitoring indices are determined to quantify the feature characteristics. Each index is a function of two variables, magnitude and position. The magnitude of an index is determined as the normalized relative maximum amplitude value of the corresponding reference function; the position is where the maximum amplitude is located. Usually, the maximum amplitude positions in these reference functions do not synchronized exactly due to the phase lags in signal processing.

Based on simulation and test observations, the effective window is determined by four tooth periods so if all indices are located within one effective window, one set of inputs x1, x2 and x3 is given to the classifier. Otherwise, if three indices are not within one effective window, the target has no fault or has more than one defect; more than one set of inputs should be provided to the classifier. For example, if x3 does not fall within the window determined by x1 and x2, two sets of inputs will be given to the monitoring classifier: The first input vector is { x1, x2, x3}, where x3 is computed over the effective window determined by both x1 and x2; The second input vector is { x1, x2, x3}, where x1 and x2 are determined over the influence window around x3.

Fig. 3 illustrates an example of the reference functions corresponding to a healthy gear with 41 teeth. Fig. 3a shows part of the original vibration signal measured from the simulation setup. Fig. 3b represents the signal average of the gear of interest, which is obtained by synchronous average filtering; each wave represents a tooth period. Figs .3c to 3e represent the resulting reference functions of the wavelet energy, beta kurtosis, and phase modulation, respectively. It is seen that no specific irregularities can be found from these reference functions for this healthy gear

Fig. 4 shows the processing results due to a cracked gear with 41 teeth. It is impossible to recognize the gear damage from the original signal (Fig. 4a). A little signature irregularity can be recognized around 200° in the signal average graph (Fig. 4b). However, this gear damage can be identified clearly from the proposed reference functions (Figs. 4c to 4e). Although the maximum peak positions are little different from one graph to another, these peaks occur within one influence window (four tooth periods).

Fig. 5 illustrates the processing results for a chipped gear (with 41 teeth). Some signature irregularity can be recognized around 200° in the signal average graph (Fig. 5b) due to this gear tooth damage. The defect can be clearly identified from other three reference functions (Figs. 5c to 5e), and the monitoring indices are located within one effective window (four tooth periods).

3.2 Prediction of the monitoring indices

Prediction system is the process to guess the future states in a system based on available observations. The classical methods are the use of stochastic models which are usually difficult to derive for mechanical systems with complex structures.

More recent research has focused on the use of data driven paradigms, such as neural networks and neural fuzzy schemes. In this work, the multi-step-ahead prediction of the input variables (indices) is performed by the use of a predictor as suggested in whose effectiveness has been verified: it can capture and track the characteristics of a system quickly and accurately comparing to other classical forecasting schemes.

Given a monitoring index X_1 , or X_2 , or X_3 , if [v_0 $v-r$ $v-2r$ $v-3r$] represent its current and previous three states with an interval of r steps, r-step-ahead state(V'_{+r})& (V''_{+r}) are estimated by a fuzzy formul

Rule : If (V_0 is B_{0k}) and (V_{-r} is B_{1k}) and (V_{-2r} is B_{2k}) and (V_{-3r} is B_{3k})

Then $V'_{+r} = C_0V_0 + C_1\ V_{-r} + C_2\ V_{-2r} + C_3\ V_{-3r} + C_4$ ………..(11)

$V''_{+r} = C_0\ V'_0 + C_1\ V'_{-r} + C_2\ V'_{-2r} + C_3\ V'_{-3r} + C_4$………..(12)

Where B are MFs, Ci are constants, i = 0, 1, ..., 3; k = 1, 2. Fig. 6 illustrates its fuzzy reasoning architecture.

This predictor has a weighted feedback link to each node in layer 2 to deal with time explicitly as opposed to representing temporal information spatially. The context units copy the activations of output nodes from the previous time step, and allow the network to store actions from the past, which forms a context for current processing. This function of recurrent networks is valuable for predictors with limited and step inputs (i.e., r > 1), to provide more information to the network so as to improve forecasting accuracy. If two sigmoid MFs are assigned to each input variable, the node output at the kth process step will be the activation function. If a max-product operator is applied in layer 3, and a centroid method is used for defuzzification in layer 5, by some related fuzzy operations, the predicted outputs(V'_{+r}) & (V''_{+r}) can be determined by:

$$V'_{+r} = \sum_{j=1}^{16} \boldsymbol{\mu}^{-}(C_0V_0 + C_1\, v-r + C_2\, v-2r + C_3\, v-3r + C_4)\text{.......}(13)$$

$$V''_{+r} = \sum_{j=1}^{16} \boldsymbol{\mu}^{-}(C_0\, V'_0 + C_1\, V'-r + C_2\, V'-2r + C_3\, V'-3r + C_4)\text{.....}(14)$$

Where $\boldsymbol{\mu}^{-} = \frac{\mu j}{\sum_{j=1}^{16} \mu j}$ represent the normalized rule firing strength and $\boldsymbol{\mu j}$ is the firing strength of the j_{th} rule.

The fuzzy system parameters are trained by using a hybrid algorithm: that is, the premise parameters in the MFs are trained by a real-time recurrent training algorithm whereas the consequent parameters C in Eq (11&12) are updated by least squares estimate.

4. THE CLASSIFIER TRAINING BY ONLINE

In order to achieve the desired input-output mapping the developed diagnostic classifier should be optimized. Several training algorithms have been proposed. In offline training, representative data should cover all of the possible application conditions; such a requirement is usually difficult to achieve in real-world machinery applications because most machinery operates in noisy and uncertain environments.

Furthermore, machinery characteristics may change suddenly, for instance, just after repair or regular maintenance. Therefore, an adaptive training algorithm is preferred in time-varying systems to accommodate different machinery conditions.

In this case, a hybrid method based on recursive Levenberg-Marquet (LM) and LSE will be adopted to train the integrated classifier. Such a training approach possesses randomness that may help to escape certain local minima.

5. THE TEACHING MF PARAMETERS TRAINING

The nonlinear teaching MF parameters will be trained by adopting the recursive LM method. The general LM algorithm possesses quadratic convergence close to a minimum. Its convergence property is still reasonable, even if the initial estimates are poor. In addition, the LM algorithm has been proven globally convergent in many applications by properly choosing the step factors. For a training data pair [x,d], the inputs are x= $\{x_1, x_2, x_3\}$; d are the desired outputs {0, 0.5, 1} as x belongs to C_1 , C_2 and C_3 , respectively. The error function with respect to adjustable MF parameters $\boldsymbol{\theta}p$ at the current time instant p is:

$$E(\boldsymbol{\theta}p) = \frac{1}{2}\sum_{p=1}^{p}[\boldsymbol{y_p}(\boldsymbol{\theta}\, \boldsymbol{p}) - \boldsymbol{d_p}]\text{......................}(15)$$

Where $yp(\theta p)$ is the *pth* output in Eq. (7). $p = 1, 2, ..., P$; dp is the desired output.

The recursive LM algorithm can be represented by:

$$\Theta_{p+1} = \boldsymbol{\theta}p + \phi_p\, J_p\, r_p \text{ ....}(16)$$

For real-time applications $\boldsymbol{\theta}_o$=0, ϕ_p is a covariance matrix with initial condition $\phi_o = \rho I$, where ρ is a positive quantity and I is an identity matrix.

By simulation tests with the requirements of the recognition rate ≥ 85%, reasonable training speed and accuracy, the following initial values are given to the related parameters in this study: $\boldsymbol{\eta}$ = 0.01 with tested range of $\boldsymbol{\eta} \in [0.005, 15]$; $\boldsymbol{\rho} = 10^3$ with tested range of $\boldsymbol{\rho} \in [10^2, 10^5]$.

6. HYBRID TRAINING METHOD IMPLEMENTATION

In implementation, inside each training epoch, the nonlinear MF parameters in the classifier are optimized in the backward pass by using a recursive LM method, whereas consequent linear rule weights are updated by LSE in the forward pass. On the other hand, after training or real applications over some time period, if the updated rule weights wj are sufficiently small (e.g., wj < 0.01), the contribution of the related rule to the final classification operation can be neglected, and that rule can be removed from the rule base.

7. SIMULATION EVALUATION

To verify the viability of the proposed classifier, five gear cases are tested in this study as represented:

a. healthy gears (C1);
b. gears having a tooth crack with 15% (C2) and 50% (C3) tooth root thickness;
c. gears having a chipped tooth with 10% (C2) and 40% (C3) tooth surface area removed.

These demonstrated faults belong to localized gear defects. From the signal property standpoint, when a localized fault occurs, some high-amplitude pulses will be generated due to impacts, which are relatively easier for a signal processing technique to recognize. When a localized fault propagates towards a distributed defect, the overall energy of the fault will increase, but it often becomes more wideband in nature and difficult to detect in the presence of the other vibratory components of the machine. This example identifies a characteristic of currently used fault detection techniques: It is usually easier to detect a distinct low-level narrowband tone than a high-level wideband signal in the presence of other signals or noises. Even though a distributed defect, such as pitting and wear, is initiated from a localized fault which is detectable as an incipient defect, most currently available vibration-based signal processing techniques cannot effectively detect an advanced distributed fault which, however, can be diagnosed based on other information carriers, such as acoustic signals. To make a comparison, the diagnostic results from the following three classifiers are also listed:

1. A pure fuzzy system with a similar reasoning architecture as in Fig. 2 but without the use of predictors. The rule weight factors are chosen as those in the integrated classifier after initial training.
2. Classifier-1: A classifier with a similar reasoning architecture as in Fig. 2 but without predictors. Its MF parameters are trained by a gradient-LSE algorithm.
3. Classifier-2: Like Classifier-1, but trained by the hybrid algorithm of the recursive LM and LSE. Given the network architectures, the initial parameters of three adaptive classifiers can be primarily trained by using some data sets collected in previous tests on the same test apparatus, or be initialized by experience. Then these classifier parameters are optimized in the following online training processes.

During online tests, motor speed and load levels are randomly changed to simulate general and unusual machinery operating conditions. The tests are conducted under load levels from 0.5 to 3 hp, and motor speeds from 50 to 3600 rpm.

In online monitoring, based on test schedule and load/speed change frequency, the monitoring time-interval is set at 15 minutes; that is, all the monitoring schemes are applied automatically every 15 minutes for condition monitoring operations. Three steps-ahead predictors (i.e., r = 3) are used in the integrated classifier. The selection of data size depends on noise reduction requirement; usually the data for the gear with the lowest speed should cover more than 100 revolutions. For example, if the slowest gear speed in the gearbox is 1200 rpm, the data acquisition process takes at least 5 seconds (15 seconds in this case). The monitoring is performed gear by gear. Three examples corresponding to healthy, cracked and chipped gears (all having 41 teeth) have been illustrated in Figs. 3 to 5, respectively.

Each healthy gear condition is tested over 24 hours whereas each faulty gear condition is tested over 50 hours. In total, 386 data pairs are recorded for testing purpose. Table(1) summarizes the classification performance by different diagnostic schemes.

The fuzzy classifier (with one predicted monitoring indices) records 15 missed alarms and 37 false alarms, with an overall reliability of 85.3%. Its relatively poor diagnostic performance is mainly due to the lack of learning capability. In addition, fixed or human-determined system parameters are subject to variations and are rarely optimal in terms of reproducing the desired classification outputs, which results in the fuzzy classifier not being optimized under different operating conditions, while for two predicted monitoring indices the results are 13, 30 and 89.2% consequently. Classifier-1 records 7 missed alarms and 21 false alarms, with an overall reliability of 92.5%, while for two predicted monitoring indices the results are 7, 18 and 94.4% consequently.

Classifier-2 records 7 missed alarms and 17 false alarms, with an overall reliability of 93.6%, while for two predicted monitoring indices the results are 5, 13 and 96.3% consequently. The developed integrated classifier generates 3 missed alarms and 7 false alarms, with an overall reliability of 97.6%, while for two predicted monitoring indices the results are 2, 3 and 98.8% consequently.

The developed integrated diagnostic classifier provides a robust problem solving framework. Machinery conditions vary dramatically in real-world applications, and new system conditions may occur under different circumstances. With the help of an adequate learning algorithm, new information can be extracted from online training, and the diagnostic knowledge base can be expanded automatically to accommodate different machinery conditions.

8. CONCLUSIONS

In this paper, an integrated classifier is developed for gear fault diagnostics. The purpose is to provide industries with a more reliable monitoring tool to prevent machinery system performance degradation, malfunction, and sudden failure. The classifier can integrate different features for a more positive assessment of the object's health condition. The diagnostic reliability is improved by properly integrating the future states of the gear, which are forecast by multi-step predictors for more than one predicted states. An online hybrid training technique based on a recursive LM and LSE is adopted to improve the classifier's convergence and adaptive capability to accommodate different machinery conditions.

The viability of the new integrated classifier has been verified by simulated tests corresponding to different gear conditions. On the other hand, it should be stated that although satisfactory results have been achieved based on the developed integrated classifier, its network architecture is relatively complex which may not be easy for implementation for some real-world applications. Future research is to develop novel evolving fuzzy or neuro-fuzzy classification schemes for more effective diagnostic operations. New training algorithms will be proposed to further improve the training convergence. The proposed techniques will also be employed for real world industrial applications in vehicles, wind turbines, and manufacturing facilities.

REFERENCES

1) Chelidze, D. & Cusumano, J. (2004). A dynamical systems approach to failure prognosis, Journal of Vibration and Acoustics, Vol. 126, 1-7.
2) Castellano G; Fanelli, A. & Mencar, C. (2004). An empirical risk functional to improve learning in a neuro-fuzzy classifier, IEEE Transactions on Systems, Man, Cybernetics, Part B, Vol. 34, 725-731.
3) Figueiredo, M.; Ballini, R.; Soares, S.; Andrade, M. & Gomide F. (2004). Learning algorithms for a class of neurofuzzy network and applications, IEEE Transactions on Systems,Man, Cybernetics, Part C, Vol. 34, 293-301.
4) Gusumano, J.; Chelidze, D. & Chatterjee, A. (2002). Dynamical systems approach to damage evolution tracking, part2: Model-based validation and physical interpretation, Journal of Vibration and Acoustics, Vol. 124, 258-264.
5) Ishibuchi, H. & Yamamoto, Y. (2005). Rule weight specification in fuzzy rule-based classification systems, IEEE Transactions on Fuzzy Systems, Vol. 13, 428-435.

6) Isermann, R. (1998). On fuzzy logic applications for automatic control, supervision, and fault diagnosis, IEEE Transactions on Systems, Man, Cybernetics, Part A, Vol. 28, 221-235.
7) Jang, J. (1993). ANFIS: adaptive-network-based fuzzy inference system, IEEE Transactions on Systems, Man, Cybernetics, Vol. 23, 665-685.
8) Korbicz, J.; Koscielny, J.; Kowalczuk, Z. & Cholewa, W. (2004). Fault Diagnosis: Models, Artificial Intelligence, Applications, Springer.
9) Li, L. & Lee, H. (2005). Gear fatigue crack prognosis using embedded model gear dynamic model and fracture mechanics, Mechanical Systems and Signal Processing, Vol. 20, 836-846.
10) Mansoori, E.; Zolghadri, M. & Katebi, S. (2007). A weighting function for improving fuzzy classification systems performance, Fuzzy Sets and Systems, Vol. 158, 583-591.
11) McFadden, P. (1986). Detecting fatigue cracks in gears by amplitude and phase demodulation of the meshing vibration, Journal of Vibration, Acoustics, Stress, and Reliability in Design, Vol. 108, 165-170.
12) Pourahmadi, M. (2001). Foundation of Time Series Analysis and Prediction Theory, John & Sons.
13) Rish,I; Brodie, M.; Ma,S.;Odintsova,N.; Beygelzimer,A; Grabarnik,G.& Hernandez, K. (2005). Adaptive diagnosis in distributed systems, IEEE Transactions on Neural Networks, Vol. 16, 1088-1109.
14) Tse, P. & Atherton, D. (1999). Prediction of machine deterioration using vibration based fault trends and recurrent neural networks, Journal of Vibration and Acoustics, Vol. 121, 355-362.
15) Uluyol, O.; Kim, K. & Nwadiognu, E. (2006). Synergistic use of soft computing technologies for fault detection in gas turbine engines, IEEE Transactions on Systems, Man, Cybernetics, Part C, Vol. 36, 476-484.
16) Wang, J & Lee, L. (2002). Self-adaptive neuro-fuzzy inference systems for classification applications, IEEE Transactions on Fuzzy Systems, Vol. 10, 790-802.
17) Wang, W. (2008). An intelligent system for machinery condition monitoring, IEEE Transactions on Fuzzy Systems, Vol. 16, No. 1, 110-122.
18) Wang, W.; Ismail, F. & Golnaraghi, F. (2001). Assessment of gear damage monitoring techniques using vibration measurements, Mechanical Systems and Signal Processing, Vol. 15, 905-922.
19) Wang, W. & Vrbanek, J. (2007). A multi-step predictor for dynamic system property forecasting, Measurement Science and Technology, Vol. 18, 3673-3681.

Table (1): Comparison of the diagnostic results from different diagnostic schemes with **ONE** predicted monitoring indices (old research).

Diagnostic Model	Healthy Gears		Cracked Gears		Chipped Gears		Total Accuracy
	M.A	F.A					
Fuzzy System	0	13	12	16	3	8	85.3%
Classifier -1-	0	7	6	9	1	5	92.5%
Classifier -2-	0	5	7	8	0	4	93.6%
New Classifier	0	2	3	3	0	1	97.6%

Table (2): Comparison of the diagnostic results from different diagnostic schemes with **TWO** predicted monitoring indices (our research).Knowing that M.A (missed alarm) & F.A (false alarm).

Diagnostic Model	Healthy Gears		Cracked Gears		Chipped Gears		Total Accuracy
	M.A	F.A					
Fuzzy System	0	11	11	14	2	5	89.2%
Classifier -1-	0	6	5	8	1	4	94.4%
Classifier -2-	0	3	5	7	0	3	96.3%
Novel Classifier	0	1	2	1	0	1	98.8%

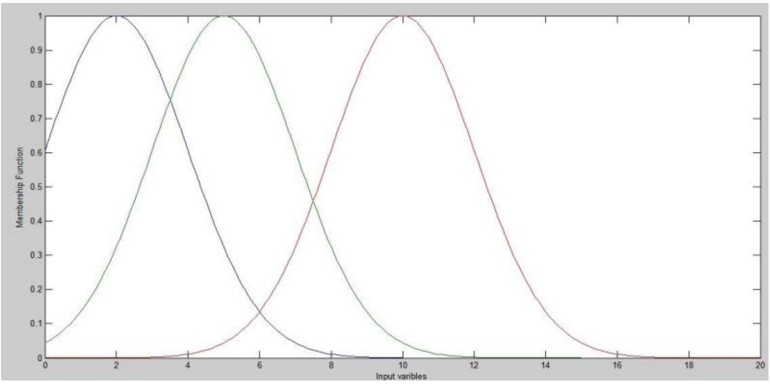


Figure (1): The membership functions (MFs) for the input state variables

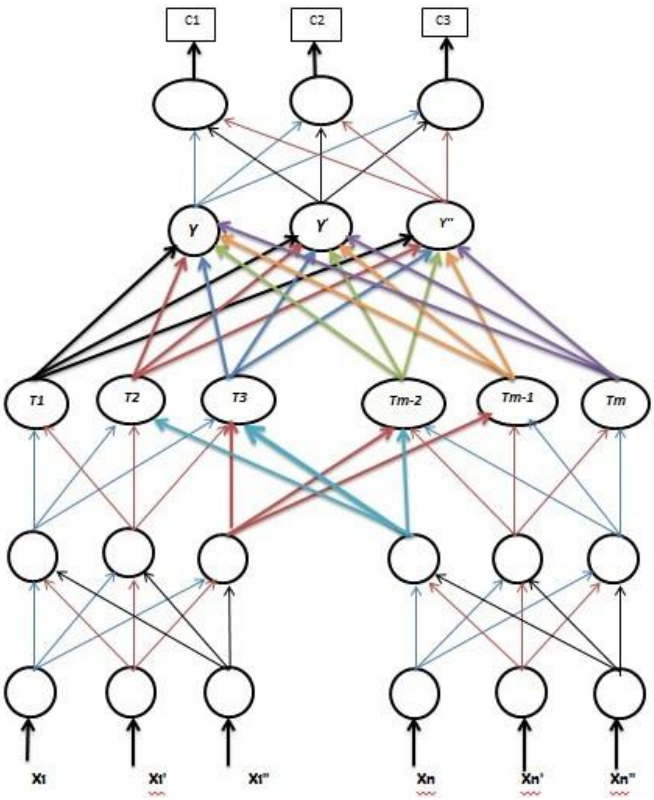
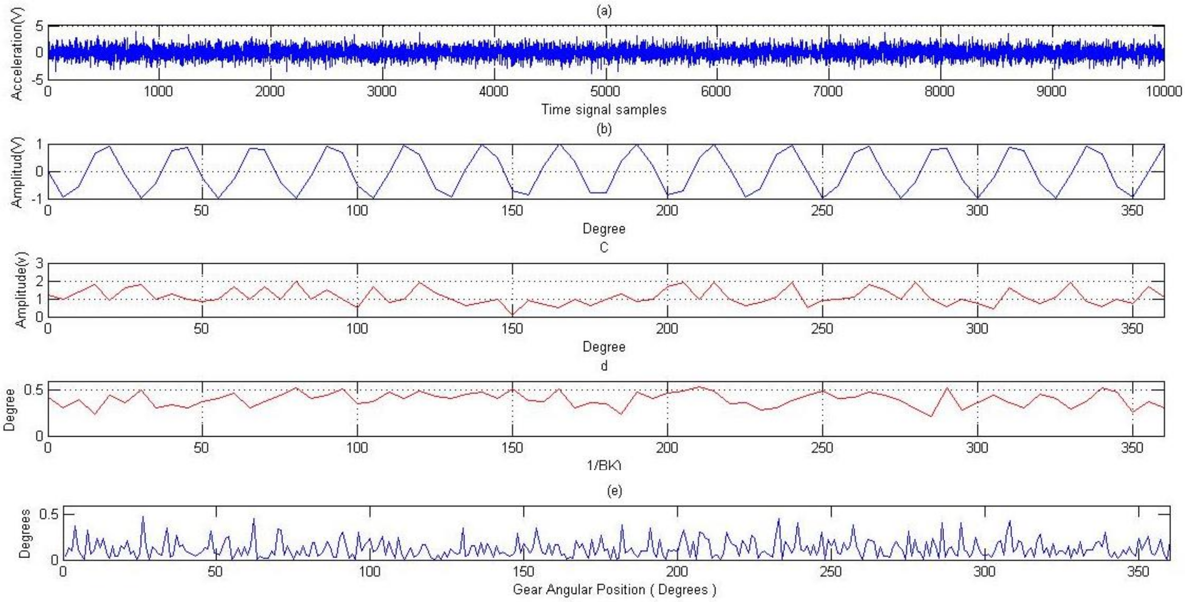


Figure (2): The network architecture of the proposed integrated classifier



Figure(3): Healthy gear processing results : (a) Part of the original vibration signal; (b) Signal average; (c) Wavelet reference function; (d) Beta kurtosis reference function; (e) Phase modulation reference function.

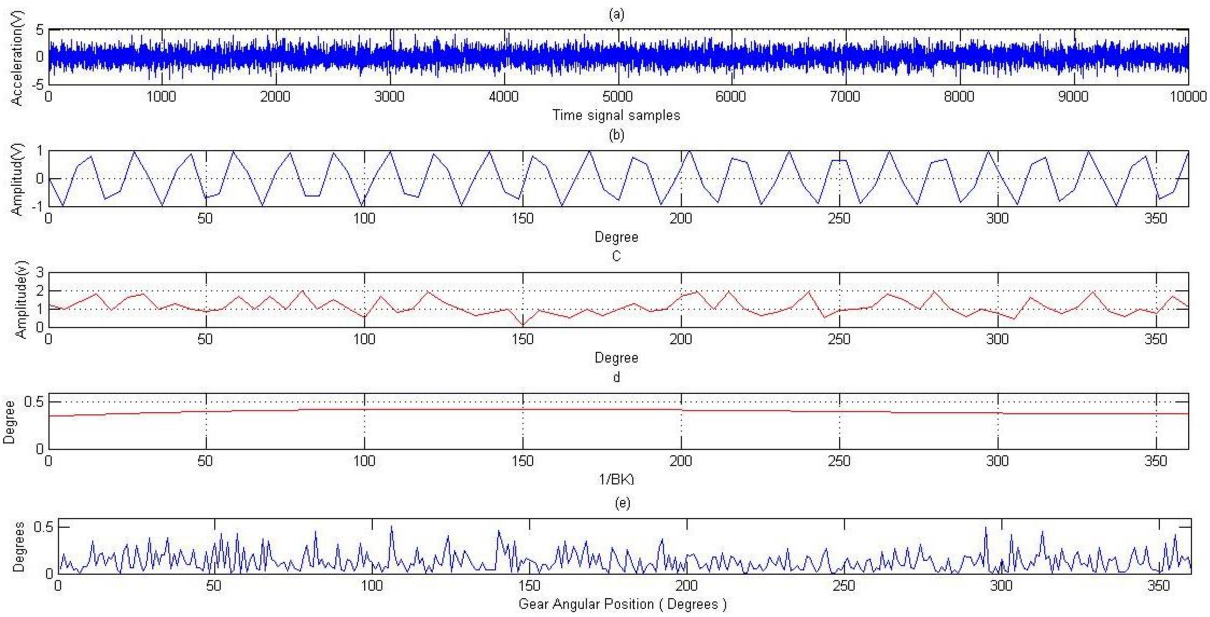


Figure (4): Cracked gear processing results: (a) Part of the original vibration signal; (b) Signal average; (c) Wavelet reference function; (d) Beta kurtosis reference function; (e) Phase modulation reference function.

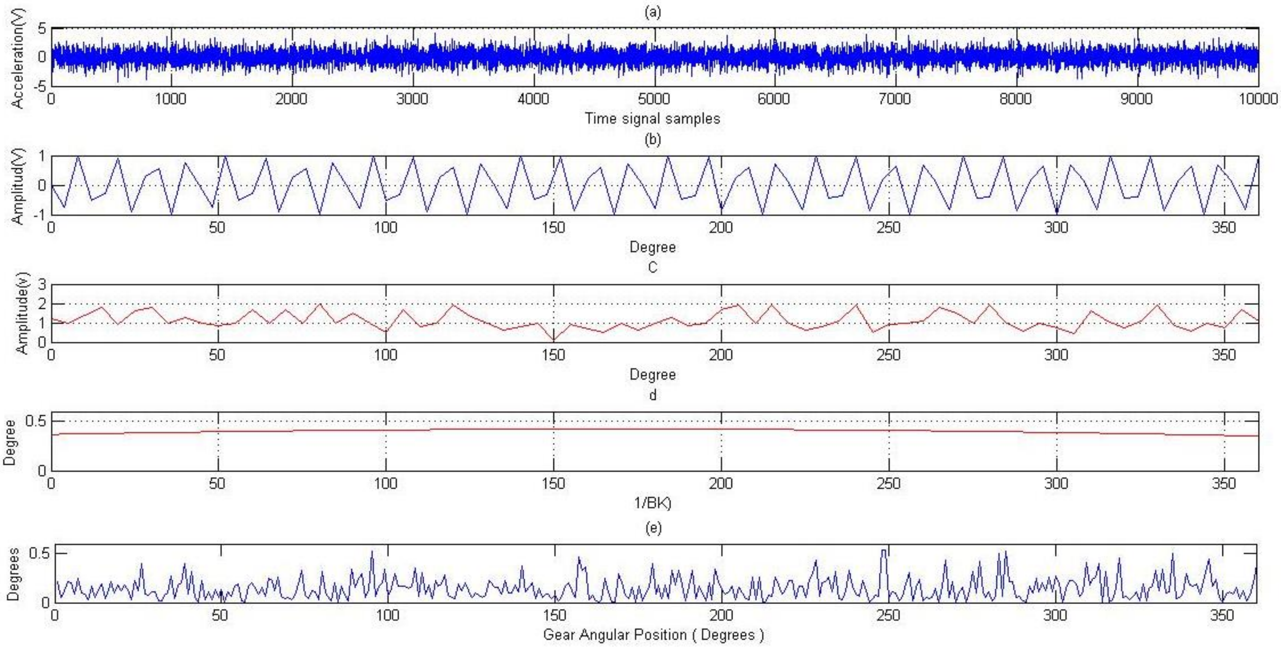


Figure (5): Chipped gear processing results: (a) Part of the original vibration signal; (b) Signal average; (c) Wavelet reference function; (d) Beta kurtosis reference function; (e) Phase modulation reference function