AN INTEGRATED APPROACH TO MACHINE FAULT DIAGNOSIS

P. Wang, N. Propes, N. Khiripet, Y. Li, and G. Vachtsevanos

School of Electrical and Computer Engineering
Georgia Institute of Technology
Atlanta, Georgia 30332-0250

Abstract

This paper introduces an integrated methodology to monitor and diagnose machine faults in complex industrial processes such as textile and fiber manufacturing facilities. The approach is generic and applicable to a variety of industrial plants that operate critical processes and may require continuous monitoring and maintenance procedures. A dual approach is pursued: High-bandwidth fault symptomatic evidence, such as vibrations, current spikes, etc., are treated via a feature extractor/neural network classifier construct while low-bandwidth phenomena, such as temperature, pressure, corrosion, leaks, etc., are better diagnosed with a fuzzy rule base set as an expert system. The technique is illustrated with typical examples from benchmark processes common to many industrial plants.

1. Introduction

The manufacturing and industrial sectors of our economy are increasingly called to produce at higher throughput and better quality while operating their processes at maximum yield. As manufacturing facilities become more complex and highly sophisticated, the quality of the production phase has become more crucial. The manufacture of such typical products as textiles and fibers, aircraft, automobiles, appliances, etc, involves a large number of complex processes most of which are characterized by highly nonlinear dynamics coupling a variety of physical phenomena in the temporal and spatial domains. It is not surprising, therefore, that these processes are not well understood and their operation is "tuned" by experience rather than through the application of scientific principles. Machine breakdowns are common limiting uptime in critical situations. Failure conditions are difficult and, in certain cases, almost impossible to identify and localize in a timely manner. Scheduled maintenance practices tend to reduce machine lifetime and increase downtime, resulting in loss of productivity. Recent advances in instrumentation, telecommunications and computing are making available to manufacturing companies new sensors and sensing strategies, plant-wide networking and information technologies that are assisting to improve substantially the production cycle.

Machine diagnostics/prognostics for condition-based maintenance involves an integrated system architecture with a diagnostic module – the diagnostician – which assesses through on-line sensor measurements the current state of critical machine components, a prognostics module – the prognosticator – which takes into account input from the diagnostician and decides upon the need to maintain certain machine components on the basis of historical failure rate data and appropriate fault models, and a maintenance scheduler whose task is to schedule maintenance operations without affecting adversely the overall system functionalities of which the machine in question is only one of its constituent elements.

This paper addresses issues relating to the diagnostic module of the Condition-Based-Maintenance (CBM) architecture. Fault diagnosis, or equally fault detection and identification (FDI), is a mature field with contributions ranging from model-based techniques to data-driven configurations that capitalize upon soft computing and other

0-7803-5621-7/99/$10.00 © 1999 IEEE

"intelligent" tools [1][2]. Recently, some strategic issues and approaches to failure detection and identification (FDI) have been addressed by several investigators. The first issue is the performance of FDI so that detection delays and false alarms may be avoided [3]. Second, a failure model should reflect a finite number of failure modes that are anticipated (predictable). Third, the designer is faced with the tradeoffs of hardware redundancy and software complexity from an implementation point of view [4,5]. Fourth, failure detectability and identifiability can be described in terms of sensitivity and distinguishability of the failure modes. Finally, robustness of FDI in the presence of modeling errors adds more significance to the modeling point of view. Most of available FDI strategies are related to the multiple model (MM) approach in which innovation-based systems or detection estimation methods are employed. The MM scheme can use existing Kalman filters without any changes and a wide range of statistical test procedures such as the generalized likelihood ration (GLR) test [4] or the sequential probability ratio (SPR) test [6] that can be applied to additive failures or event-driven faults. For linear systems, the failure sensitive filter approach provides a solution to FDI. The Beard-Jones detection filters (BJDF) [7] or the Luenberger observer filters (LOF) [8] may be applied to detect a wide variety of system failures (sensors, actuators, components). The jump-process formulation (JPF) technique handles sudden shifts or jumps in the system matrices by mixing the MM method [4]. It has a manageable fixed bias size and is still suboptimal because of steady-state effects on residuals. Other methods include an algorithmic approach to FDI [9] and an expert system approach [10] among many alternatives investigated over the past years.

In many practical situations, uncertainty in the process can affect the performance of the system significantly no matter how the uncertainty is described (vagueness or ambiguity). This realization provides the motivation for a possible fuzzy logic approach to FDI. This has the ability to directly describe the potential failure modes in the parameters while handling a class of nonlinear systems. To resolve the actual failure in the system parameters, a recursive parameter estimation technique is an essential component of FDI. In a soft body of consonant evidence, Zadeh's fuzzy sets or membership functions [11] can be applied to continuous decision-making processes whereas in a distinct body of crisp evidence we can rely on Dempster-Shafer's belief of plausibility measure [12]. These approaches provide a mathematical theory of combining rules of evidence. More recently, neural network constructs and wavelets have surfaced as potential candidates for fault detection and identification. A major innovation in the proposed work relates to the utility of wavelets, in a neural network setting, for fault classification purposes [13].

2. The Diagnostic System Architecture

Components, machines and processes fail in varying ways depending upon their constituent materials, operating conditions, etc. Failure modes are typically monitored by a sensor suite which is intended, for failure analysis purpose, to capture those failure symptoms that are characteristics of a particular failure mode. Consider, for example, the case of a typical process such as a slasher or a weaving machine in textiles. Typical failure modes may include leaks, sensor failures, corrosion, debris, etc. which are characteristic of a process failures as well as a variety of vibration induced faults that are affecting mechanical and electro-mechanical process elements.

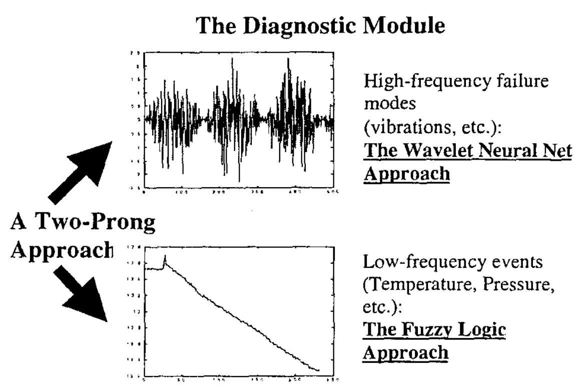


Figure 1 The two-prong approach of the diagnostic module

It is generally possible to break down the sensor data (and, correspondingly, the symptomatic evidence) into tow broad categories: The first one concerned with low-bandwidth measurements, such as those originating from process variables, temperature, pressure, levels, etc., while the second exemplifies high-bandwidth measurements, for example vibrations, current spikes, etc (see Figure 1). Failure modes associated with the first category may develop slowly and data is sampled at slow rates without loss of trending patterns. High-frequency phenomena though, such as those accompanying a bearing failure, require a much faster sampling rate in order to permit a reasonable capture and characterization of the failure signature. Moreover, process-related measurements and associated failure mode signatures are numerous and may overlap, thus presenting serious challenges in resolving conflicts and accounting for uncertainty. This dichotomy suggests an obvious integrated approach to the fault diagnosis problem: Process related faults may be treated with a fuzzy rule base set as an expert system while high-bandwidth (see Figure 2) faults are better diagnosed via a feature extractor/neural network classifier topology. This approach is adopted below in addressing typical machinery failures.

The basic diagnostic architecture is generic and applicable to a wide variety of complex engineered systems and industrial processes. A combined fuzzy logic/Dempster-Shafer approach is used to determine if a failure (or impending failure) has occurred and to assign a degree of certainty or confidence to this declaration. Figure 3 depicts the essential elements of the diagnostic process.

Preprocessing and Feature Extraction

The preprocessing and feature extraction unit takes raw sampled data from a plant and converts it to a form suitable for the fuzzy logic and Dempster-Shafer system. It incorporates filtering of noise from raw data and extraction of features from the filtered data. Feature extraction intends to extricate the most important characteristics from the filtered data such as slopes, levels, relevant frequencies, etc. Feature extraction itself is a form of filtering and thus leads to false alarm rate improvement.

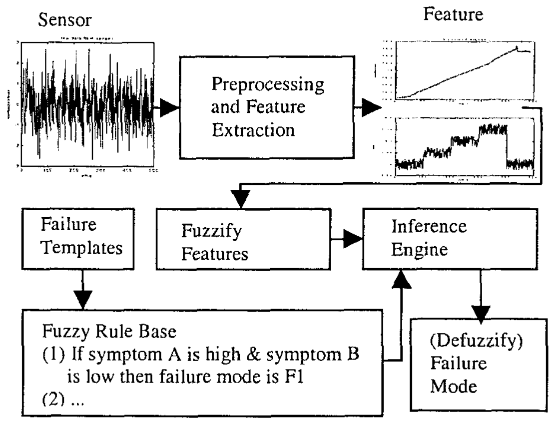


Figure 2 Fuzzy diagnostic system layout with feature extraction

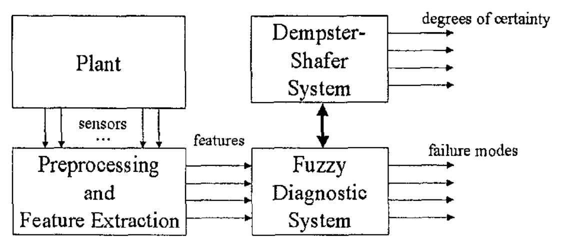


Figure 3 The fuzzy logic and evidence theory approach

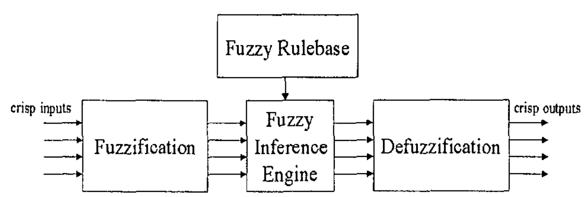


Figure 4 The fuzzy logic diagnostic system

3. Fuzzy Diagnostic System

The fuzzy diagnostic system takes features as inputs and then outputs any indications that a failure mode may have occurred in the plant. The fuzzy logic system structure is composed of four blocks: fuzzification, the fuzzy inference engine, the fuzzy rulebase, and defuzzification, as shown in Figure 4. The fuzzification block converts features to degrees of membership in a linguistic label set such as low, high, etc. the fuzzy rulebase is constructed from symptoms that indicate a potential failure mode. Figure 5 depicts two typical rules. Some examples of rules in such a rulebase could be:

If **the temperature is low in Tank 1** and **the pressure is low** then **the failure mode is Tank 1 heating element is damaged**.

If **the slope of Tank1's water level is negative low** and **the slope of Tank 1's pressure is negative low** then **the failure mode is Tank 1 leaking**.

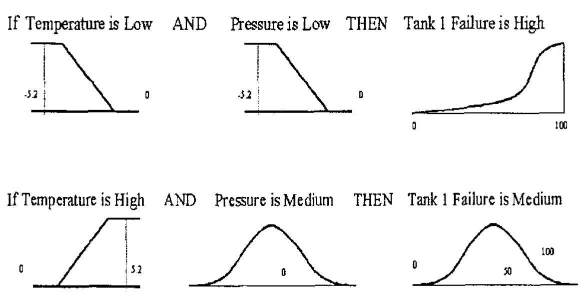


Figure 5 A graphical representation of a fuzzy rulebase

The fuzzy rulebase can be developed directly from user experience, simulated models, or experimental data. Fuzzy outputs are aggregated (maximum method) through the fuzzy inference engine to determine a degree of fulfillment for each rule corresponding to each failure mode. The last step defuzzifies the resulting output, using the centroid method, to a number between 0 and 100 (figure 6). This is finally compared to a threshold to determine whether or not a failure mode should be reported.

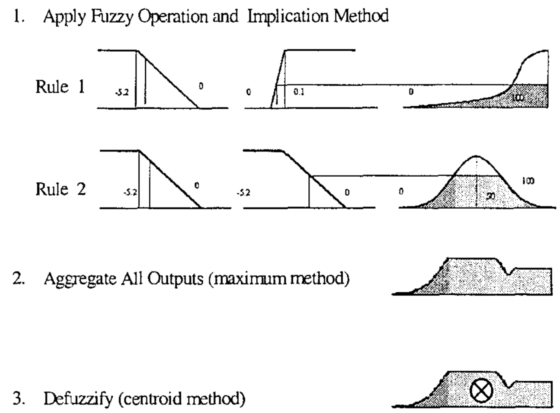


Figure 6 Graphical Inference and defuzzification

Dempster-Shafer Theory of Evidence

Dempster-Shafer theory of evidence is incorporated into the system for uncertainty management purposes. Its function is to associate a degree of belief or certainty to a detected failure mode from the fuzzy diagnostic system. It uses the same rulebase as the fuzzy diagnostic system for consistency. Dempster-Shafer theory takes the same features that are fed into the fuzzy diagnostic system and places them through the same input membership functions. Each sensor is considered an expert in this setting. The membership values are normalized and ordered for each sensor separately. The values are then subtracted and assigned to nested mass functions in an order form. There is now a single mass function for each sensor (i.e. expert). By examining the fuzzy rulebase, one can now use Dempster's rule of combination to combine the mass functions of the sensors and determine a final mass function with failure modes as elements. Finally, a degree of certainty is calculated and is sent along with the failure mode to the output database for maintenance decisions

4. High-Bandwidth failure Detection and Identification

The Wavelet Neural Network (WNN) belongs to a new class of neural networks with such unique capabilities as multi-resolution and localization in addressing classification problems. For fault diagnosis, the WNN serves as a classifier so as to classify the occurring faults (see Figure 7).

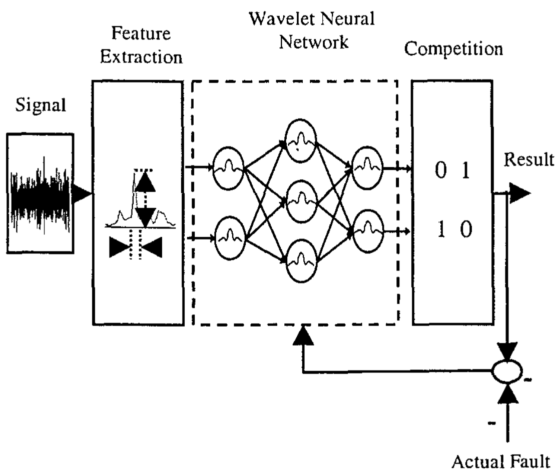


Figure 7 Classification using the wavelet neural network

Critical process variables are monitored via appropriate sensors. The data obtained from the measurements are processed and features are extracted. The latter are organized into a feature vector, which is fed into the WNN. Then, the WNN carries out the fault diagnosis task. In most cases, the direct output of the WNN must be decoded in order to produce a feasible format for display or action.

For example, the WNN can be used to perform the diagnosis of a bearing failure typically found on races, rolling balls and lubrication materials. Here, for simplicity, the focus is placed on the diagnosis of whether the bearing is normal or defective. Through vibration measurements, a number of vibration signals for a bearing are collected and the peaks of the signal amplitude and the signal's PSD are chosen as the features. Such other quantities as the standard deviation, cepstrum, DCT coefficients, wavelet maps, temperature, humidity, speed, mass, etc. can be selected as candidate features. From the vibration signals, a training data set is obtained, which is then used to train the WNN.

Once trained, the WNN can be employed to perform the fault diagnosis. Signals and their PSDs from a normal bearing and a defective one are shown in Figure 8.

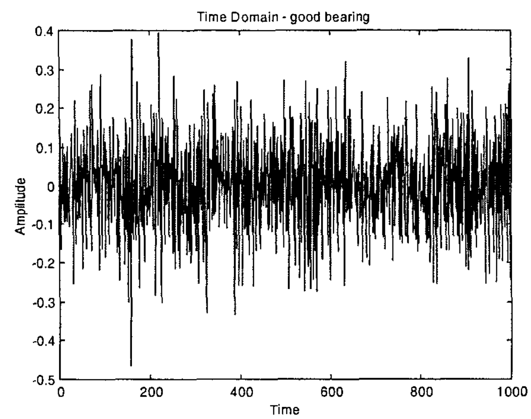
For a good bearing;
features = [0.3960 0.1348]

For a defective bearing:
features = [4.9120 9.2182]

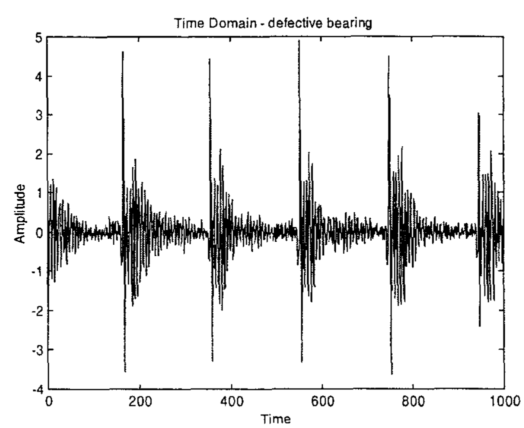
[0 1] = WNN([0.3960 0.1348])
===>The bearing is good!

[1 0] = WNN([4.9120 9.2182])
===> The bearing is defective!

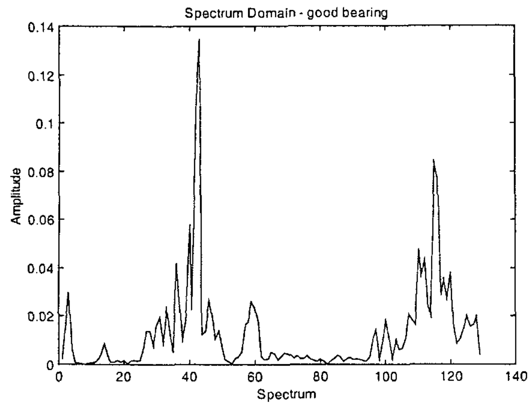
The results can easily be extended to include cases in which multiple fault classes are concerned.



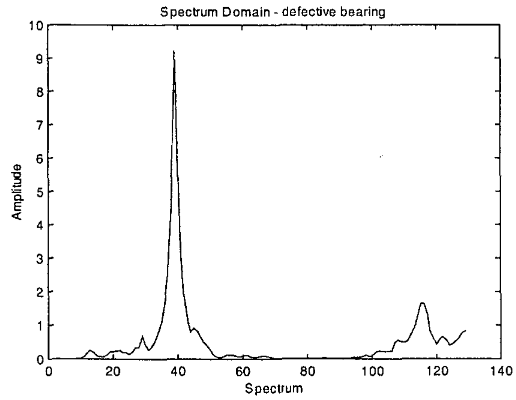
(a) Signal from a good bearing



(b) Signal from a defective bearing



(c) PSD from a good bearing



(d) PSD from a defective bearing

Figure 8 Signals and their PSD from a normal and a defective bearing

5. Conclusions

A model-free approach to the problem of diagnosing fault conditions has been presented. For low-bandwidth process data, a knowledge base built as a fuzzy expert system correlates fault symptoms to failure modes while addressing effectively uncertainty management. Next, high-bandwidth data, such as vibration measurements, are dealt with in a framework consisting of a feature extractor and a classifier. The multi-dimensional WNN is an effective and efficient tool for classification. The computational burden shifts to the feature extraction step where appropriate features must be computed from signal data that comprise eventually the input vector to the network. The WNN approach offers additional advantages in terms of learning and optimization functions that may be carried out off-line or on-line. Furthermore, the neural net topology suggests means for parallel processing – useful in high frequency processes. The scheme show promise as an effective model for the analysis of vibration and process data for many industrial and other engineered systems.

Acknowledgement

This work was supported by Hoenywell/ONR under contract No. E-21-K30. Their continued sponsorship is gratefully acknowledged.

References

[1] Konrad, H. and Isermann, R., "Diagnosis of different faults in milling using drive signals and process models", Proceedings of the 13th World Congress, International Federation of Automatic Control, Vol.B., Manufacturing, p.91-6, 30 June-5 July 1996.

[2] Mylaraswamy, D., and Venkatasubramanian, V., "A hybrid framework for large scale process fault diagnosis", Computers & Chemical Engineering, vol.21, suppl. issue, p. S935-40, 25-29 May 1997.

[3] Chow, E. Y., and Willsky, A. S., "Analytical Redundancy and the Design of Robust Failure Detection Systems," *IEEE Transactions on Automatic Control*, Vol. AC-29, No. 7, July 1984, pp. 689-691.

[4] Willsky, A. S., "A Survey of Design Methods for Failure Detection in Dynamic Systems," *Automatica*, Vol. 12, No. 6, November 1976., pp. 601-611.

[5] Isermann, R., "Process Fault Detection Based on Modeling and Estimation Methods—A Survey," *Automatica*, Vol. 20, No. 4, July 1984, pp. 387-404.

[6] Deckert, J.C., Desai, M. N., Deyst, J. J., and Willsky, A. S., "F-8 dfbw Sensor Failure Identification Using Analytic Redundancy," *IEEE Transactions on Automatic Control*, Vol. AC-22, No. 5, Oct. 1977, pp. 795-803.

[7] Jones, H. L., "Failure Detection in Linear Systems," Ph.D. Thesis, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA, Aug. 1973.

[8] Kailath, T., *Linear Systems*, Prentice-Hall, Englewood Cliffs, NJ, 1980.

[9] Kim, J. H. and Bien, Z., "An Algorithmic Approach to Fault Diagnosis in Linear Systems," *IEEE Transactions on Industrial Electronics*, Vol. 36, No. 3, Aug. 1989, pp. 313-320.

[10] Handelman, D. A. and Stengel, R.F., "Combining Quantitative and Qualitative Reasoning in Aircraft Failure Diagnosis," *Proceedings of the 1985 AIAAA guidance, Navigation, and Control Conference*, AIAA, New York, 1985, pp. 375-384; AIAA Paper 85-1905, Aug. 1985.

[11] Zadeh, L. A., "Outline of a New Approach to the Analysis of Complex Systems and Decision Processes," *IEEE Transactions on System Man Cybernetics*, Vol. SMC-3, No. 1, Jan. 1973, pp. 28-44.

[12] Shafer, G., *A Mathematical Theory of Evidence*, Princeton Univ. Press, Princeton, NJ, 1976.

[13] M. Mufti and G. Vachtsevanos, "Automated Fault Detection and Identification Using a Fuzzy Wavelet Analysis Technique," *Proceedings of AUTOTESTCON '95 Conference*, pp. 169-175, August 1995.