Application of Machine Learning in Fault Diagnostics of Mechanical Systems

Massieh Najafi, David M. Auslander, Peter L. Bartlett, Philip Haves

***Abstract*— A diagnostic method based on Bayesian Networks (probabilistic graphical models) is presented. Unlike conventional diagnostic approaches, in this method instead of focusing on system residuals at one or a few operating points, diagnosis is done by analyzing system behavior patterns over a window of operation. It is shown how this approach can loosen the dependency of diagnostic methods on precise system modeling while maintaining the desired characteristics of fault detection and diagnosis (FDD) tools (fault isolation, robustness, adaptability, and scalability) at a satisfactory level. As an example, the method is applied to fault diagnosis in HVAC systems, an area with considerable modeling and sensor network constraints.**

***Index Terms*— Fault Detection, Bayesian Networks, Machine Learning, System Diagnostics, HVAC Systems.**

I. Introduction

The topic of fault detection and diagnostics (FDD) has been the center of attention for several decades [1], [2], [3] & [4]. The desired characteristics of FDD tools along with practical limitations have made fault diagnosis problems challenging and as much art as science. It is not just the advancement of diagnostic methods; it is also the issue of scalability and the capability of large industrial applications. The characteristics of an ideal diagnostic tool have been discussed in several resources [1]: The more important ones can be summarized as:

- Fault Isolation: the ability to diagnose and distinguish different faults and their combinations
- Robustness: the ability to maintain FDD performance despite the existence of noise and uncertainty
- Adaptability: the ability to respond accordingly to system change and disturbances
- Interoperability: the capability for quick implementation with minimum (or reasonable) prior adjustment, investment …

Manuscript received August 12, 2008.

M. Najafi is Ph.D. candidate at the Department of Mechanical Engineering, University of California Berkeley, CA 94720 (e-mail: mnajafi@berkeley.edu).

D. M. Auslander is professor of the Graduate School, Department of Mechanical Engineering, University of California Berkeley, CA 94720 (e-mail: dma@me.berkeley.edu).

P. L. Bartlett is professor of the Division of Computer Science and Department of Statistics, University of California Berkeley, CA 94720 (e-mail: bartlett@cs.berkeley.edu).

P. Haves is the leader of the Commercial Building Systems Group, Lawrence Berkeley National Laboratory (LBNL), Berkeley CA, 94720 (e-mail: phaves@lbl.gov).

The point is that maintaining all these characteristics at a satisfactory level is challenging. Usually improving one comes at the expense of sacrificing the other(s). For instance, when robustness is improved by loosening the thresholds, fault isolation ability degrades.

On the other hand, in mechanical systems, there are other practical limitations imposing further challenges to diagnostic problems: First, modeling: The accuracy of system models can usually be improved up to certain levels, beyond that, it will be expensive, slow, and too customized, affecting interoperability. Such a constraint can degrade the effectiveness of model based diagnostic techniques in mechanical systems. Second, sensor network: The architecture of the sensor network may expand the complexity of diagnostic problems. Usually in practical applications, the design of the sensor network is not solely based on diagnosis purposes. Other factors like controls, financial constraints, and space limitations are also involved. As a result, what is measured by the installed sensors may not necessarily be the parameter(s) contaminated by fault(s) directly.

One way to address these issues is to focus on system behavior patterns instead of residuals. In other words, instead of diagnosing system health status through comparing its outputs directly with model predictions (or other references), diagnosis can be done by analyzing system behavior over a window of operation. Such a transition can loosen the dependency of diagnostic mechanisms on precise predictions of system behavior while maintaining desired FDD characteristics. In a fuzzy environment, there is more flexibility in classifying system behavior patterns than analyzing the deviation of system outputs from model predictions.

The diagnostic tool should have some level of understanding of system behavior patterns in faulty and non-faulty modes in advance. Such an understanding is under the influence of various uncertainties. This can be achieved through first-principle modeling, empirical approaches, or a combination of both depending on constraints and limitations. Then at the interface level, when the FDD tool is provided with data from system behavior, it matches them with different hypotheses.

This approach has been applied in some diagnostic methods, especially in qualitative and semi-quantitative approaches, e.g. [5] [6] [7] & [8]. The key here is a mechanism (interface) capable of matching data from system behavior with a set of predefined (or new) hypotheses in an environment affected by uncertainties.

Fuzzy logic is a popular choice for these types of problems. The nature of fuzzy sets and fuzzy inference mechanisms has enough flexibility to deal with model uncertainties. Fuzzy based diagnostic mechanisms have extensively been used in different applications, e.g. [9] [10] [11] & [12]. For example, in [9] fuzzy logic has been used for induction motor condition monitoring, or, in [12], a fuzzy logic approach has been developed for gas turbine fault diagnostics.

However, fuzzy based mechanisms have their own limitations. As the complexity grows (due to system complexity, variety of faults, or the potential of multiple faults), the number of fuzzy rules explodes. The same thing happens when a fuzzy based diagnostic mechanism is extended to have the capability of updating its belief about the system health status as more data are observed. Added to this is the issue of adjusting and tuning fuzzy sets either manually or through other approaches.

In this paper, a different approach based on Bayesian networks (BN) is proposed. The application of Bayesian networks in fault diagnostics has been studied in some areas, e.g. [13], [14], & [15]. For example, in [15] Chien et al have applied Bayesian networks for fault diagnostics in a power delivery system. In [14], a Bayesian network is implemented for controlling an unsupervised fault tolerant system to generate oxygen from the CO2 on Mars.

However, these approaches are not based on the strategy of system behavior pattern analysis. They are based more on cause-effect-relationship approaches. What is different here is the way Bayesian networks are applied for diagnostic purposes. In the proposed method, system behavior patterns are at first captured by the diagnostic mechanism using knowledge of first-principle modeling (simplified physical models). This includes patterns of faulty and non-faulty modes. The limitations coming from model simplification (or other resources) are quantified either analytically or statistically. In other words, the diagnostic tool captures both knowledge of first-principle modeling and empirical results. Then at the diagnosis stage, the tool will look for predefined (or new) hypotheses of faulty or non-faulty operation in system behavior.

The rest of the paper goes as follows: In section 2 a brief introduction to Bayesian networks is provided. Section 3 explains the proposed diagnostic mechanism and its characteristics, in section 4, the method is applied to the problem of fault diagnostics in HAVC systems, an area with considerable modeling and sensor network constraints.

II. Introduction to Bayesian Networks

A Bayesian network (BN) is a directed graph in which each node is annotated with qualitative probability information. The full specification is as follows:

1- A set of random variables makes up the nodes in the network. Variables may be discrete or continuous.

2- A set of directed links or arrows connects pairs of nodes. If there is an arrow from node A to node B, A is said to be the parent of B (Fig. 1). The intuitive meaning of an arrow in a properly constructed network is usually that A has a direct influence on B.

3- Each node Xi has a conditional probability distribution $P(X_i \mid parents\ (X_i))$ that quantifies the effect of the parents on the node.

4- The graph has no directed cycles.

The topology of the network – the set of nodes and links - specifies the conditional independence relationships that hold in the domain. Each node is conditionally independent of all its non-descendants given its parents. For example in Fig. 1, D is conditionally independent of A given B. The combination of the topology and the conditional distributions suffices to specify (implicitly) the full joint distribution for all the variables.

Using the chain rule in probability, a joint distribution can always be broken down into a product of conditional probabilities. For example, for A, B, C, and D, P(A, B, C,D) can always be represented as:

$$P(A,B,C,D) = P(A)P(B \mid A)P(C \mid A,B)P(D \mid A,B,C) \quad (1)$$

The conditional independence assumptions expressed by a BN allow a compact representation of the joint distribution. For example in Fig. 1, knowing that A and D are conditionally independent given B, the joint probability distribution can be simplified to: $P(A,B,C,D) = P(A)P(B \mid A)P(C)P(D \mid B,C)$. In general in Bayesian networks with $x_1, x_2, \ldots, x_n$ as random variables, the joint probably can be simplified as:

$$P(x_1, x_2, \ldots, x_n) = \prod_i P(x_i \mid Parents(x_i)) \quad (2)$$

III. Fault Diagnostics Based on Bayesian Networks

The proposed BN based FDD mechanism has four nodes (Fig. 2):

- Fault node: representing different faults and their combinations (e.g. valve leakage).
- Fault Level node: representing the level or degree of fault (e.g. leakage position).It may be continuous or discrete. If it is not applicable, it can be omitted.
- Input node: representing system inputs and other known parameters (e.g. control signal, ambient temperature …),
- Output Node: representing system outputs or what is measured from the system.

The input node is deterministic (no probability distribution). The projection from the input node to the output node is through functions that are influenced by the fault node. This influence may vary from changing a parameter value to defining a significantly different behavioral pattern. These functions capture the simplified physical model of the system. The level of model simplification depends on available knowledge from first-principle modeling, complexity etc. The uncertainty at the output node due to model simplification (modeling error) and other factors (e.g. sensor noise, etc) is interpreted as the variance of the output node variable(s). It can be quantified through analytical or statistical approaches.

The realization of the fault node influence on projecting functions from the input node to the output may happen in different ways. For instance, there might be a different set of projecting functions for each type of fault, or the parameters of the probably distribution random variable(s) of the output node might be a linear combination of a set of basis functions generated at the input node with coefficients defined by fault node.

The undefined parts of projecting functions will be addressed in the training phase. This may include but is not limited to: the variance of the output node random variable(s), the coefficients of linearly combined basis functions, etc. If the distribution of the output node random variable is Gaussian, it is straightforward to estimate the output node variance and coefficients of linearly combined basis functions.

The posterior probability of the fault node can be interpreted as the FDD tool belief level about the existence of different faults in the system. Using Bayesian networks interface algorithms, this probability can be calculated by:

$$\begin{aligned} &Fault = \{fault_1, fault_2, fault_3, \ldots, fault_n\} \\ &P(fault_i \mid Input, Outpt) \propto P(fault_i, Input, Outpt) = \\ &\sum_{FaultLevel} P(fault_i) P(faultLevel \mid fault_i) P(Output \mid Input, faultLevel) \end{aligned} \tag{3}$$

A normalizing factor 'Z', is required at the end, which can be calculated by:

$$\begin{aligned} Z = &\sum_{fault} \sum_{faultLevel} p(fault). \\ &\cdot P(faultLevel \mid fault)\, p(Output \mid Input, faultLevel). \end{aligned} \tag{4}$$

The prior distribution of faults and *P(FaultLevel|Fault)* can be estimated statistically or as a quick solution they may all be assumed to be uniformly distributed.

Now, as more evidence is observed (more output data are measured), the tool can improve its belief about system health status accordingly. In other words, more data from the system behavior help the diagnostic mechanism to do a better matching between the system behavior pattern and different hypotheses. For instance, if two set of measurements, $Output_1$ and $Output_2$,are observed (Fig. 3), the fault belief can be calculated by:

$$\begin{aligned} &P(fault_i \mid Input_1, Outpt_1, Input_2, Outpt_2) \propto \\ &P(fault_i, Input_1, Outpt_1, Input_2, Outpt_2) = \\ &P(fault_i) P(Output_1 \mid Input_1, fault_i) P(Output_2 \mid Input_2, fault_i) \end{aligned} \tag{5}$$

There are different techniques to simply this calculation in a recursive manner as more data are observed. As before, there is also a normalizing factor involved. Note that the graph in Fig. 3 can be extended into a dynamic graph in which improving and updating the belief are handled simultaneously.

IV. Example, Fault Diagnostics in HVAC Systems

Heating, Ventilation, and Air Conditioning Systems (HVAC) are getting growing attention for fault diagnosis purposes. HVAC systems are a major consumer of energy in building, however it has turned out that most HAVC facilities are not working efficiently due to system faults [4]. In this example, fault diagnosis of an HAVC device, an air handling unit is studied. An air handling unit has the task of mixing the air, controlling the temperature and the humidity of the air going to the building (Fig. 4).

A component in an air handling unit, the mixing box, has the task of mixing the air coming out from the building (Return Air) and the outside air with a specified ratio defined by the controller. The ratio is calculated such that it minimizes the energy required to heat up or cool down the supply air (the air going into the building), and also satisfies the standard of fresh air requirement for building occupants. Mixing box mal-functionality is a common problem in air handling unit. The abnormality of mixing box could be due to stuck damper fault, outside air damper leakage, return air damper leakage, reversed actuator, sensor offset, etc.

In a typical air handling unit there are usually three sensors, measuring outside air temperature (OAT), return air temperature (RAT), and supply air temperature (SAT) (Fig. 5). The performance of the mixing box is analyzed by a parameter, Outside Air Fraction (OAF), which is the ratio of the difference between SAT and RAT over OAT and RAT:

$$OAF = \frac{T_{\sup} - T_{ret}}{T_{out} - T_{ret}} \tag{6}$$

OAF is an indication of the influence of outside air temperature on supply air temperature. When the outside air damper (OAD) is fully open, OAF is ideally 1, and when it is closed, it supposed to be 0. Fig. 6 shows how OAF changes with damper position in non-faulty operation. Inside the envelope is the acceptable performance. The wide range of acceptable performance is due to the uncertainty of parameters not easily measurable in a typical application (fluid resistance, air velocity, thermal resistance …). Fig. 7 shows the behavior of the mixing box in different faulty modes.

Fig. 8 shows the fault diagnosis mechanism designed for the mixing box. The input is damper position and the output is OAF. The distribution of the output node (OAF) is assumed to be Gaussian. The projection from the input node (DMP) to the output node (OAF) is a linear combination of the basis functions, B_1 & B_2, by coefficients defined by the fault node. The potential faults are: outside air damper leakage, return air damper leakage, stuck, and reverse, with the possibility of multiple faults. The FDD mechanism was tested by data from the Iowa Energy Center, an experimental facility for research, education, and demonstration [5]. The results are shown in Figures 9 and 10.

V. Conclusion

The transformation from focusing on residuals at one or a few operating points to analyzing system behavior patterns over a window of operation can relax the dependency of diagnostic mechanisms on system models, but it also increases the necessity for more sophisticated mechanisms for analyzing system behaviors in an environment affected by uncertainties. The proposed BN based diagnostic mechanism seems to be capable of dealing with such circumstances. Its ability to capture both knowledge of first-principle modeling and empirical results can provide more flexibility in the design process. The method was applied to the problem of fault diagnostics in HVAC systems, an application area with considerable modeling and sensor network limitations. Given the promising results in HVAC systems, this method appears to have good potential for application to other types of mechanical systems.

References

[1] V. Venkatasubramanian, R. Rengaswamy, K. Yin, and S. N. Kavuri, “A review of process fault detection and diagnosis”, Part I: Quantitaive Model-based methods. Computers in Chemical Engineering, 2003, vol. 27, pp. 293-311.

[2] V. Venkatasubramanian, R. Rengaswamy, K. Yin, and S. N. Kavuri, “A review of process fault detection and diagnosis”, Part II: Qualitative models and search strategies. Computers in Chemical Engineering, 2003, vol. 27, pp. 313-326.

[3] V. Venkatasubramanian, R. Rengaswamy, K. Yin, and S. N. Kavuri, “A review of process fault detection and diagnosis”, Part III: Process history based mothods. Computers in Chemical Engineering, 2003, vol. 27, pp. 327-346.

[4] S. Katipamula, M. R. Brambley, “Methods for fault detection, diagnostics, and prognostics for building systems – a review part I”, HVAC&R Research, 2005, vol. 11, n1.

[5] R. Davis, “Diagnosis reasoning based on structure and behavior”, Artificial Intelligence, 1984, 24 (1-3), 347-410

[6] R. Milne, “Strategies for diagnosis”, IEEE Transaction on Systems, Man, and Cybernetics 1987, vol. 17, n3, pp. 333-339.

[7] S. H. Rich, V. Venkatasubramanian, “Model based reasoning in diagnostic expert systems foe chemical process plants”, Computer and Chemical Engineering, 1987, vol. 11, n2, pp. 111-122.

[8] S. H. Rich, V. Venkatasubramanian, “Causality-based failure-driven learning in diagnostics expert systems”, American Institute of Chemical Engineering Journal, 1989, vol. 33, n6, pp. 943-950.

[9] H. Nejjari and M.E.H. Benhouzid, "Application of fuzzy logic to induction motors condition monitoring", *IEEE Power Eng. Rev*., 1999, vol. 19, pp. 52-54.

[10] X. Z. Gao and S. J. Ovaska, "Soft computing methods in motor fault diagnosis," *Applied Soft Computing Journal*, 2001, vol. 1, n1.

[11] M. E. H. Benhouzid and H. Nejjari, "A simple fuzzy logic approach for induction motors stator condition monitoring", *IEEE International Conference on Machines and Drives*, 2001, pp. 634-639.

[12] S. O. T OGAJI, L. MARINAI, S. SAMPATH, R. SINGH, S. D. PROBER, “Gas-turbine fault diagnostics : a fuzzy-logic approach”, Journal of Applied energy, 2005, vol. 82, n1, pp. 81-89.

[13] L. M. Riascos, M.G. Simoes, P.E. Miyagi, “A Bayesian network fault diagnostic system for proton exchange membrane fuel cells”, Journal of power sources, 2007, vol. 165, n1, pp. 267-278.

[14] U. Lerner, B. Moses, M. Scott, S. McIlraith, D. Koller, “Monitoring a complex physical system using a hybrid dynamic Bayes net”, Proceedings of 18th Conference on Uncertainty in AI, 2002, pp. 301–310.

[15] C. F. Chien, S.L. Chen, Y.S. Lin, “Using Bayesian network for fault detection on distribution feeder”, IEEE Trans. Power Delivery, 2202, vol.17, n3, pp.785–793.

[16] S. Russell, P. Norvig, ”Artificial Intelligence: A Modern Approach." 2nd Edition, Prentice Hall, 2003.

[17] M. I. Jordan, “Graphical models”, Statistical Science *(Special Issue on Bayesian Statistics)*, 2004, vol. 19, pp. 140-155.

FIGURES AND TABLES

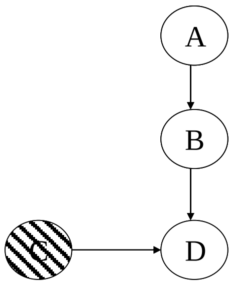


Figure 1

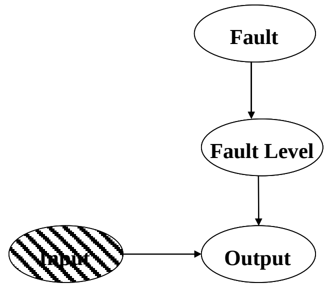


Figure 2

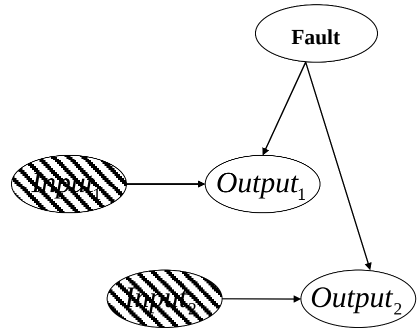


Figure 3

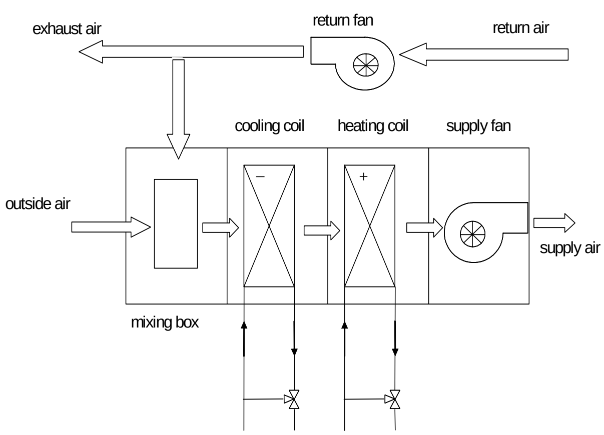


Figure 4, Schematic Diagram of Air Handling Unit

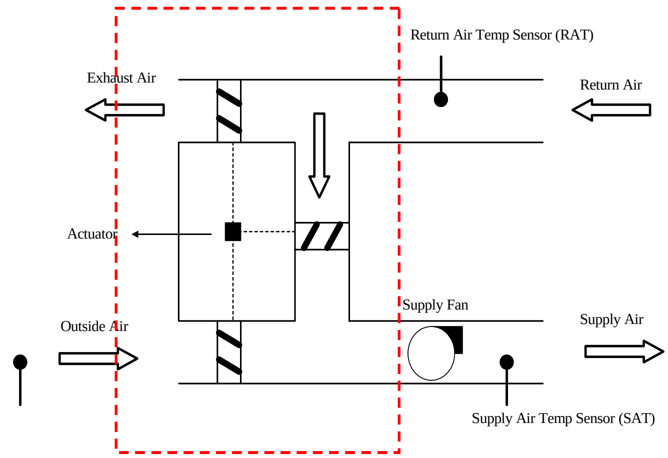


Figure 5, Schematic Diagram of Mixing Box

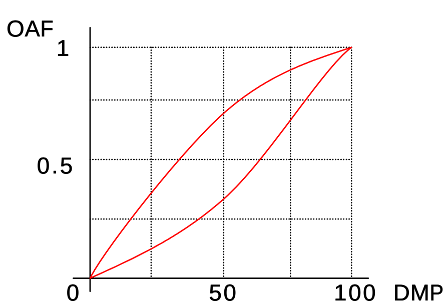


Figure 6, OAF versus damper in non-faulty mode

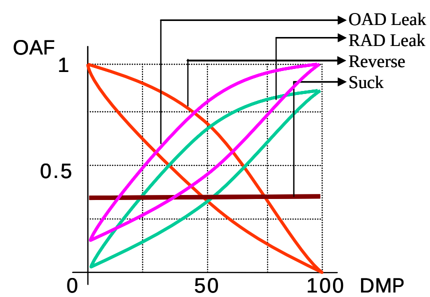
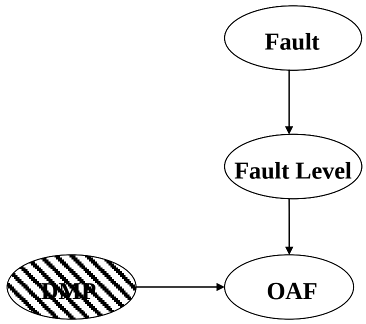


Figure 7, OAF versus damper in faulty modes

ISBN: 978-988-98671-0-2



Fault = {Damper Stuck, outsider DMP leakage, Return DMP leakage, Reverse}

$$\mu_{OAF} = \theta_1 B_1 + \theta_2 B_2 + f(FaultLevel)$$

$$B_1 = 1, B_2 = (DMP/100)$$

$$\sigma_{OAF} = (DMP/100) - (DMP/100)^2$$

Figure 8, Fault diagnosis design for mixing box

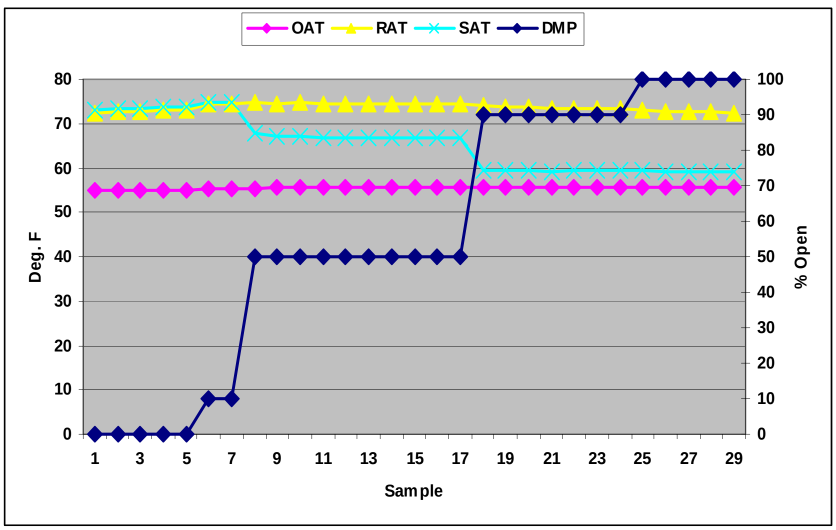


Figure 9, OAT: outside air temperature, RAT: return air temperature, SAT: supply air temperature, DMP: damper. The data is from a test run on one of the air handing units at Iowa energy center. The diagnostic result is shown in Figure 10

Figure 10, Diagnostic results, Note how FDD belief about the system health status improves as it sees more data. It shows that there is an "Outside Air Damper Leakage" fault in the system. However, due to the nature of the "Outside Air Damper Leakage" fault, it can not be isolated until the system goes to the one hundred percent damper position. Note how the diagnostic tool waits until the system goes to the one hundred percent damper positions, and then finalizes its evaluation.