2005 American Control Conference
June 8-10, 2005. Portland, OR, USA

Towards a new Fault Diagnosis System for Electric Machines based on Dynamic Probabilistic Models

José L. Flores-Quintanilla, Rubén Morales-Menéndez
Ricardo A. Ramírez-Mendoza, Luis E. Garza-Castañõn, Francisco J. Cantú-Ortiz

Abstract— **This paper presents a new approach to diagnose faults in electrical systems based on probabilistic modelling and machine learning techniques. Our framework consist of two phases: an approximated diagnosis on the first phase and a refined diagnosis on the second phase. On the first phase the system behavior is modelled with a Dynamic Bayesian Network that generates a subset of most likely faulty components. In this phase the structure and parameters of the Dynamic Bayesian Network are learned off-line from raw data (discrete and continuous). On the second phase a Particle Filter algorithm is used to monitor suspicious components and extract the faulty components. The feasibility of this approach has been tested in a simulation environment using several interconnected electrical machines.**

I. Introduction

When a machine fails in a real process, the maintenance people take minutes, hours or even days looking for the faulty elements depending on the complexity of the system.

Diagnosing in large interconnected process, such as power electrical systems, is a difficult task due mainly to the inherent uncertainty in information. The uncertainty arises due to close interaction between systems components, that in the faulty event hide real fault symptoms. Other sources of uncertainty include: noisy data, non-linearities and missing information. An adequate automated decision support system has to deal with the challenge in the fault diagnosis task. This system enables us to diagnose faults in an uncertain environment. Extensive research has been done to develop fault diagnosis methods based in analytical methods [1], [2], [3] or artificial intelligence techniques [4] [5][6], [7]. Recently, attention to probabilistic modelling and machine learning techniques has grown [8], [9], due to the availability of data and computation power, but mainly because of the difficulty to model systems with analytical techniques. Powerful diagnosis methods are able to deal with different sources of uncertainty and capable of learning models of system behavior from data.

A recent trend is the combination of different techniques, to address challenging environments such as large interconnected systems, where discrete and continuous signals are present. In such systems, classical mathematical models can not deal with all sources of uncertainty and are very difficult to obtain as the system grows in complexity.

This paper is mainly focused on the integration of dynamic Bayesian Networks and Particle Filtering algorithms to tackle the problem of diagnosis. This approach permits improvement in system robustness and allows to accomplish a better performance compared to the separate individual application. The study was applied to a simple case of an interconnected electrical system in a simulation environment.

The paper is organized as follows. Section II presents background on fault diagnosis methods. Section III describes the interconnected electrical system used as a case study. Section IV explains the approach and framework proposed in this paper. Section V shows the simulation results and finally section VI concludes the paper and includes future work.

II. Fault Diagnosis Methods

The fault diagnosis task can be separated into two steps: fault detection and fault diagnosis. Fault detection is mainly concerned with the identification of an abnormal situation in a process.

Fault detection can be based on analytical or heuristic symptoms. Analytical symptoms are generated by simple limit value checking or by more elaborated techniques such as signal analysis (eg. statistical or frequency analysis). Heuristic symptoms are produced by qualitative information extracted from human operators or process history records. Different approaches of fault detection methods has been developed. Most approaches use mathematical models such as: Kalman filter, state observers, parity relations, etc. [10]. Fault diagnosis consists of the determination of the type, size, and location of a fault, together with the time of detection.

Fault diagnosis methods mainly use classification techniques or reasoning methods [2]. Classification techniques includes statistical methods, neural networks and fuzzy clustering. Reasoning methods mainly include first order logic, fuzzy logic and Bayesian networks.

A. Bayesian Networks

Bayesian networks are a helpful tool to model multi-fault, multi-symptom dependency relations: every fault and every symptom is modelled by a random variable with a finite range of possible values. A graph is constructed with a node

José-Leonardo Flores-Quintanilla is at the Mechatronics Program Office, ITESM San Luis Campus, México, joselfq@itesm.mx
R. Morales-Menéndez is at the Center for Innovation in Design and Technology, ITESM Monterrey Campus, México, rmm@itesm.mx
Ricardo A. Ramirez-Mendoza is at the Mechatronics Program Office, ITESM Monterrey Campus, México, ricardo.ramirez@itesm.mx
Luis E. Garza-Castañón is at the Department of Mechatronics, ITESM Monterrey Campus, México, legarza@itesm.mx
Francisco J. Cantú-Ortiz is at the Research and Graduate Studies Office, ITESM Monterrey Campus, México, fcantu@itesm.mx

for each variable. The graph constructed has an edge from one node to another, whenever the first node models a fault directly exhibiting the symptom modelled by the second.

In general, Bayesian networks are a representation of probability distributions over complex domains. Formally, we consider probability spaces defined as the set of possible assignments to some set of random variables $X_1, \ldots, X_n$, each of which has a domain $Dom[X_i]$ of possible values. For example the domain of the discrete variable *Breaker_status* is $\{$*normal, open, faulty*$\}$. The domain of the variable *Voltage* can be $\Re$, or it can be discretized into some appropriate partition. The goal is to represent a joint probability distribution over these variables.

Formally, a Bayesian network is defined by a directed acyclic graph together with a local probabilistic model for each node. There is a node in the graph for each random variable $X_1, \ldots, X_n$. The edges in the graph denote direct dependency of a variable X_i on its parents *Parents*(X_i). The graphical structure encodes a set of conditional independence assumptions (*each node X_i is conditionally independent of its non-descendants given its parents*).

The qualitative independence assumptions implied by the network structure, combined with the conditional probability distributions associated with the nodes, are enough to specify a full joint distribution through the following eqn (1), known as the *chain rule for Bayesian networks*:

$$p(X_1, \ldots, X_n) = \prod_{i=1}^{n} p(X_i | Parents(X_i)) \tag{1}$$

In discrete networks, an explicit description of the joint distribution requires a number of parameters that are exponential in n (the number of variables). Bayesian networks derive their power from the ability to represent conditional independencies among variables, which allows them to take advantage of the "locality" of causal influences. A variable is independent of its indirect causal influences given its direct causal influences. However, in order to reduce the computational effort, statistically independent symptoms have to be assumed.

B. Particle Filtering

Particle Filtering is a Markov chain Monte Carlo method that approximates the belief state using a set of samples called particles. The distribution of the particles is updated taking into account the latest available evidence as time increases. The standard Particle Filtering algorithm consists of three basic sequential steps:

- *Monte Carlo step*. This step takes into account the evolution of the system as time increases:

$$\begin{aligned} z_t &\sim p(z_t | z_{t-1}) && (2) \\ x_{t+1} &= A(z_t)x_t + B(z_t)\gamma_t + F(z_t)u_t && (3) \\ y_t &= C(z_t)x_t + D(z_t)v_t && (4) \end{aligned}$$

  The previous stochastic model of the system is used to generate the predicted future state ($z_t^{(i)}$, $x_t^{(i)}$, $y_t^{(i)}$), see [11] for details. We sampled a discrete mode, eqn (5), and then the continuous state given the new discrete mode, eqn (6).

$$\hat{z}_t^{(i)} \sim p(z_t | z_{t-1}^{(i)}) \tag{5}$$

$$\hat{x}_t^{(i)} \sim p(x_t | \hat{z}_t^{(i)}, x_{t-1}^{(i)}) \tag{6}$$

- *Sequential Importance Sampling step*. By conditioning on the new information and using the Bayes' rule, each particle is weighted by the likelihood of the observations in the updated state represented by that particle eqn (7).

$$\hat{w}_t^{(i)} \leftarrow p(y_t | \hat{z}_t^{(i)}, \hat{x}_t^{(i)}) \tag{7}$$

- *Selection step*. High-weight particles are replaced by several particles while low-weight particles tend to disappear.

Particle Filtering algorithms approximate the true posterior belief state given observations $y_{1:t}$ by a set of particles.

$$p(z_t, x_t | y_{1:t}) = \frac{1}{N} \sum_{i=1}^{N} w_t^{(i)} \delta_{(z_t, x_t)}(z_t^{(i)}, x_t^{(i)}) \tag{8}$$

where $w_t^{(i)}$ is the weight of a particle, $z_t^{(i)}$ are the discrete modes, $x_t^{(i)}$ are the continuous parameters and $\delta_{(\cdot)}(\cdot)$ is Dirac delta function.

Rao-Blackwellized Particle Filtering (*RBPF*) is a Particle Filtering variant that uses some of the analytical structure of the model, [12].

If we know the values of the discrete modes z_t, it is possible to calculate the distribution of the continuous states x_t. According to the Rao-Blackwell theorem, this leads to estimators with less variance than those obtained using only pure Monte Carlo sampling. Thus, if we can generate particles of z_t and analytically evaluate the expectation of x_t given z_t, we will need less particles for a given accuracy. We can therefore combine a Particle Filtering that computes the distribution of the discrete modes with a bank of Kalman filters that computes the distribution of the continuous states.

A further improvement based on *look-ahead RBPF* is proposed in [13]. While evaluating the importance of weights for particles at time t, *look-ahead RBPF* looks ahead one step to see the behavior of the measurements at time $t+1$. It then uses this information to compute better weights at time t. The basic sequential steps are Kalman prediction, Selection, Sequential Importance Sampling, and Kalman updating. The *look-ahead RBPF* algorithm is specifically exploited in this paper.

III. ELECTRICAL SYSTEM

A production line in a factory can be thought as a set of interconnected electrical machines. These machines are made up of several components which can be seen as series RL circuits. As the ratio of possible failures in this system is high because of the number of total components, a monitoring system for faulty component is required.

Figure 1 shows two electrical machines controlled and supervised by the same control panel.

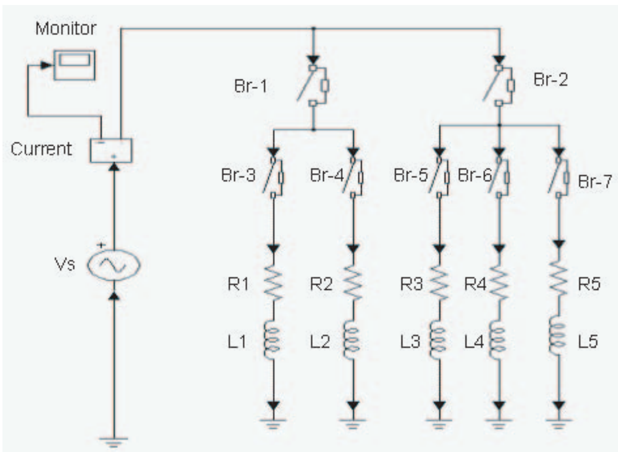

Fig. 1. Schematic diagram of two electrical machines. Series resistance-inductance circuits are shown.

Each machine is made up of two or three identical RL series circuits. Besides, each machine as well as its individual circuits has its own breakers that protect the rest of the system against over-current. Also, these breakers are used when machines or individual circuits need maintenance.

This electrical system is a typical application where interaction between continuous and discrete variables are high. Discrete variables are the machines and circuit breakers, and the continuous variables are given by the main current as well as the circuit parameters.

The main task is to determine when there is a fault, which of the two machines is in faulty mode and which circuit has the problem.

For the simulation it was considered that one series RL circuit has been isolated, also it was considered 9 possible faulty modes for this circuit. The faulty modes were implemented using a change of 5Ω in the resistance value of the circuit (starting from 10Ω).

IV. Methodology

The diagnosis system that we are proposing consists of two phases, [14]. Figure 2 shows conceptual ideas.

In the first phase the full system is modelled in probabilistic terms by a Bayesian network. It considers that we have 2 interconnected machines. The failure probabilities of each machine, based on evidences, are updated periodically. The Bayesian network is used to select the machines with a high probability of failure. The main difference with [14] consists of using first order logic representation for Bayesian networks and a set of heuristics to deal with the both problems of combinatorial explosion and selection of the most likely faulty elements. That *Independent Choice Logic* framework allows a richer model representation framework, where other features, as explicit time for instance, can be modelled, [15]. This is a semantic framework which allows for independent choices made by various agents, and a logic program to give the consequences of the choices. This is an expansion of Probabilistic Horn abduction to include a richer logic, and choices by multiple agents, [16].

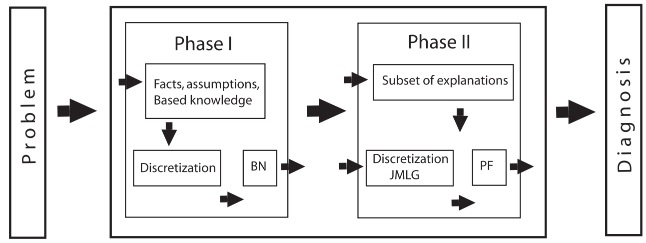

Fig. 2. General architecture for the diagnosis system.

In the second phase, figure 2, these machines are continuously monitored to find the damaged components using a Particle Filtering algorithm. In the work by [14], second phase is implemented by using a residual approach where normal process behavior is represented by a compact probabilistic model whose parameters are learned offline by finding the maximum entropy joint probability mass function consistent with arbitrary probability constraints. The structure of this model is found by learning the structure of a discrete Bayesian network.

Each block in figure 2 will be described briefly:

A. Phase I

The implementation of first the phase consists of the following steps.

1) **Facts, assumptions, knowledge base**. In this block a table describes the behavior of all the breakers which constitute the whole system as well as the behavior of the main current. We prepares a table with all possible combinations of breaker states; as a result, the data is generated for the Bayesian network of the full system.
2) **Discretization.** The discretization of the continuous variable is made. The key variable in this system is the main electrical current.
3) **Bayesian network.** The specialized software *Power Constructor*[1] developed to learn a Bayesian network from data is included here. It uses data bases in the form of discrete variables. A Bayesian network uses evidence from a set of variables and looks at the rest of variables to predict the possible faults that could be presented. The *Hugin*[2] software was used to make inferences and analysis.

B. Phase II

Fault diagnosis in this context, is to determine the state of an electrical system over time given a stream of ob-

[1] http://www.cs.ualberta.ca/ jcheng/bnpc.htm
[2] http://www.hugin.com

servations. Complex dynamics of electrical systems and noisy environment makes it difficult to determine the true state at any point in time with certainty. Uncertainty at every point should be explicitly considered. To represent uncertainty about the state of the system, it is assumed a particular a probability distribution which is considered in the Particle Filtering (*PF*) implemented algorithm. The main purpose of Particle Filters in this work is to update the Bayesian belief, [17]. The basic idea is to simulate the behavior of the electrical system. Each particle predicts a future behavior of the system in a Monte Carlo approach. The particles that match the monitored system behavior are kept and the others are thrown away. [18] describes a number of *PF*-based algorithms for state estimation which have demonstrated good results on diagnosis problems.

Particularly, the implementation of the second phase consists of the following steps:

1) **Subset of explanations.** Once that data has been analyzed a subset of explanations is given where a possible fault is located in the system. This is just a subset of the results, taking into account just the ones with major values in probability, as indicated.
2) **Discretization JMLG.** Combining Hidden Markov Models and State Space Models, a hybrid model can be generated, which provides a rich representation for processes. This model is called the Jump Markov Linear Gaussian (JMLG) model, as depicted in Figure 3.

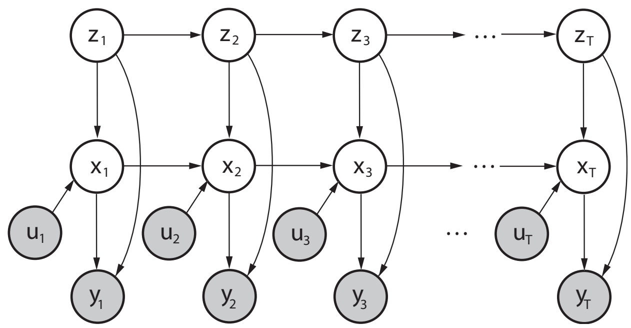

Fig. 3. JMLG model. z_t is the discrete mode switch variable, x_t is the real-valued state variable, u_t and y_t are the observable variables.

Formally, we can represent the JMLG model by eqns (2-4). See [19] for detailed definitions and learning procedure.

3) **Particle Filtering.** This final block of the diagnosis system will make different decisions depending on the results in Phase I according to the following statements.
   - If the probability of failure of the machines does not change, then Particle Filters will continue monitoring the same machines to find damaged components.
   - If the probability of failure, in a non-monitored machine, increases to a warning level, then a Particle Filter must be started to monitor this machine.
   - If the probability of failure in a monitored machine increases then, Particle Filter must increase its reliability, taken a greater number of particles.
   - If the probability of failure in a monitored machine decreases then, Particle Filter must reduce the number of particles in a first stage. If the tendency remains, the Particle Filter must stop in that machine.

V. RESULTS

These results were obtained when the proposed methodology was applied to an electrical system which consists of two machines modelled with series resistance-inductance circuit.

Phase I An expert generates data for the machines using maintenance records, data sheets from suppliers, intuition, etc. Once the data base has been created, this file is opened with *Power Constructor* to create the preliminary Bayesian network and then this information is exported to *Hugin*. The join distribution for the full network is obtained. Figure 4 shows the Bayesian network for the electrical system where breakers are discrete variables and main current is a discretized continuous variable.

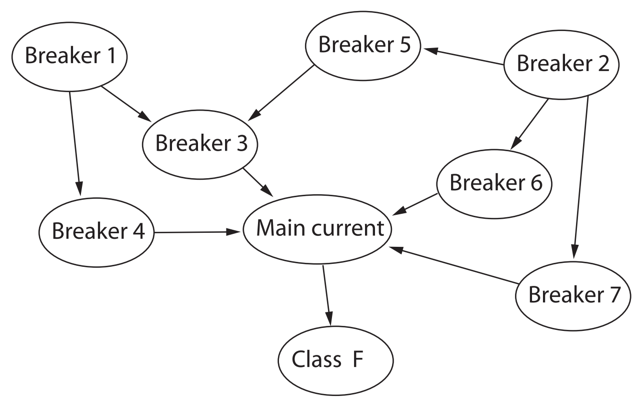

Fig. 4. System's Bayesian network

Using *Hugin* code we can make inferences on certain variables which predict the behavior of the system as shown in Figure 5.

Figure 5 shows the breaker status, this status consist of two variables, one variable is the probability of the breaker to be opened and the other variable is the probability of the breaker to be closed, also in this figure are shown the different values of the main current and the values of class F which represents the decision logic.

In figure 5, it is shown an evidence in main current of 7.35 amperes. The subset of explanations give us the most probable state of the variables, considering a main current of 7.35 amperes the subset of explanation is the following:

- Breaker 1 has a probability of 50.82% of being opened; therefore, machine 1 has a probability of 50.82% of

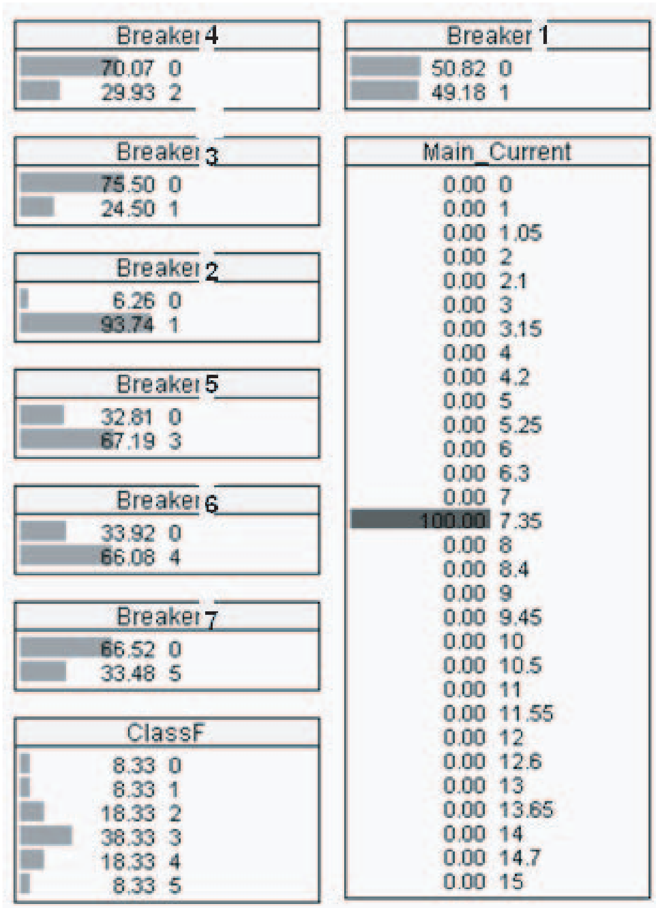

Fig. 5. Electrical system Bayesian net with an evidence of 7.35 amperes in the main current.

being turned off. This breaker connects the power supply to the machine.

- Breaker 2 is closed; thus, machine 2 has a probability of 93.74% of being turned on.
- Breakers 3 and 4 are opened; therefore, machine 1 is turned off.
- Breaker 7 is opened.
- Breakers 6 and 5 are closed.
- Class F, the final node, indicates that circuits 2, 3 and 4 have a probability of being faulty. They have the maximum values of probability.

Phase II Once phase I is finished for the full electrical system, there are a set of machines with a high probability of failure. We must run a closely monitoring system in each probable faulty machine. In order to show how the algorithm works, it was selected only one of the circuits of only one machine. The system was tested with more than 20 runs for each number of particles. Also, several runs of the experiment were executed taking into account different levels of noise. For each level of noise were used 100, 500 and 1,000 particles. The levels of noise used were 0.1%, 1%, 5% and 10% for both the process and measurement noises. Finally, we computed the statistics (mean and standard deviation) of the diagnosis error. Diagnosis error is the percentage of time steps during which the discrete mode (z_t) was not identified properly. The *Maximum A Posteriori* was used to define the most probable discrete mode over time.

Table I shows a summary of the system performance for the diagnosis algorithm. Note that the mean error diagnosis is very low when the noise is 0.1% or 1%, while mean diagnosis error grows up when the noise is 5% and 10%. The standard deviation of the error diagnosis is less as the number of particle grows up.

TABLE I

LOOK-AHEAD RBPF PERFORMANCE. DIAGNOSIS ERRORS ARE SHOWN FOR DIFFERENT NUMBER OF PARTICLES AND LEVEL NOISES.

	100 Particles		500 Particles		1,000 Particles	
Level noise	mean	Std desv	mean	Std desv	mean	Std desv
0.1%	0.04%	0.02%	0.05%	0.00%	0.05%	0.00%
1%	0.18%	0.05%	0.12%	0.05%	0.10%	0.00%
5%	5.77%	1.08%	5.30%	0.80%	4.76%	0.28%
10%	10.63%	3.82%	9.81%	1.31%	8.75%	0.58%

Figures 6-7 show the diagnosis performance for 100 particles with low-level noise. Figure 6 shows the real discrete (z_t) and continuous (y_t) states over time, while figure 7 shows the behavior of the *look-ahead RBPF* (the real and estimated states are shown). As we can see, there are few diagnosis errors. The diagnosis errors appear at the transitions over time.

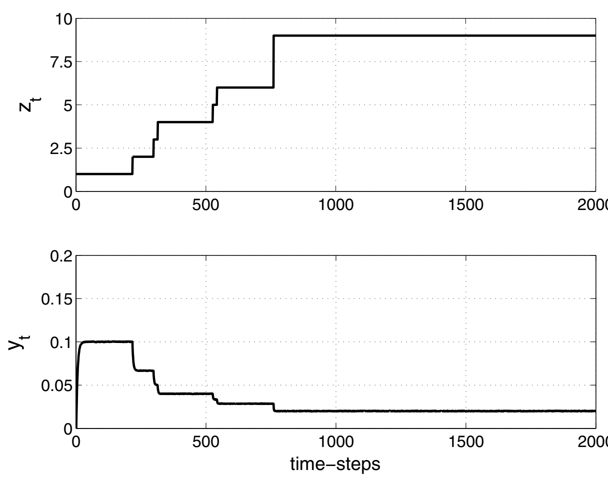

Fig. 6. True faulty states for the system and the measured variable

In presence of noise the algorithm has low performance. We can cope with high-level noise increasing the number of particle; however, high-level noise kills some *look-ahead RBPF* features, [11].

VI. CONCLUSIONS AND FUTURE WORK

We have described a new approach for diagnosing faults in industrial systems combining dynamic Bayesian learning-inference and particle filtering. Our main contribution is

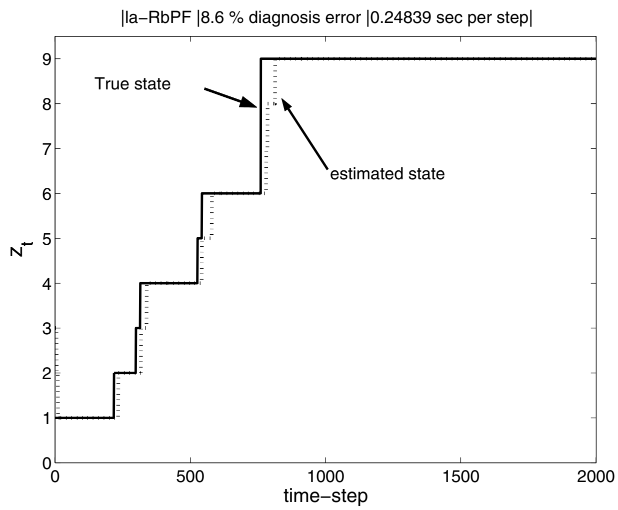

Fig. 7. State estimation with 100 particles and low-level noise

the combination of both dynamic Bayesian networks and particle filtering in a single framework and its testing in diagnosing faults for an electrical system. The dynamic Bayesian network chooses those electric machines with the highest probability of failure. The particle filtering algorithm is used to continuously monitor those machines to detect if a circuit component has changed its parameter values. The Bayesian network is periodically updated based on evidence to have a reliable system.

We are working on the applicability of the framework in more complex processes and its use on more complex industrial applications such as *CNC* Milling Machines with reconfigurable features.

References

[1] E C Larson, B E Jr Parker, and B R Clark. Model-based sensor and actuator fault detection and isolation. *Proceedings of the 2002 American Control Conference on*, 5(2):4215 – 4219, 8-10 May 2002.

[2] R Isermann. Supervision, fault-detection and fault-diagnosis methods - an Introduction. *Control Engineering Practice*, 5:639–652, 1997.

[3] E Lughofer, E.P. Klement, J.M. Luja, and C Guardiola. Model-based fault detection in multi-sensor measurement systems. *Intelligent Systems, 2004. Proceedings. 2nd International IEEE Conference*, 1:184 – 189, 22-24 June 2004.

[4] R Morales-Menéndez, R Ramírez-Mendoza, J Mutch, and F Guedea-Elizalde. Toward a new approach for online fault diagnosis combining particle filtering and parametric identification. *Lecture Notes in Computer Science*, 2972:555 – 564, April 2004.

[5] F Valles and R A Ramírez-Mendoza. A statistical hypothesis neural networks approach for fault detection and isolation of an adaptive control system. In *Avances en Inteligencia Artificial, MICAI/TANIA*, Mrida Mxico, 2002.

[6] M A Awadallah and M M Morcos. Application of AI tools in fault diagnosis of electrical machines and drives-an overview. *Energy Conversion,IEEE Transactions on*, 18(2):245 – 251, June 2003.

[7] Q Sheng, X Z Gao, and Z Xianyi. State-of-the-art in soft computing-based motor fault diagnosis. *Proceedings of 2003 IEEE Conference on*, 2(2):1381 – 1386, 23-25 June 2003.

[8] A Barigozzi, L Magni, and R Scattolini. A probabilistic approach to fault diagnosis of industrial systems. *Control Systems Technology, IEEE Transactions on*, 12:950 – 955, Nov. 2004.

[9] Z Shaoyuan, Z Jianming, and W Shuqing. Fault diagnosis in industrial processes using principal component analysis and hidden Markov model. *Proceedings of the 2004 American Control Conference*, 6:5680 – 5685, June 30-July 2, 2004.

[10] J Gertler. *Fault detection and diagnosis in engineering systems*. Marcel Dekker, Inc., 1998.

[11] N de Freitas, R Dearden, F Hutter, R Morales-Menéndez, J Mutch, and D Poole. Diagnosis by a waiter and a Mars explorer. *IEEE, Special issue on Sequential State Estimation*, 92(3):455–468, March 2004.

[12] A Doucet, N de Freitas, K Murphy, and S Russell. Rao-Blackwellised particle filtering for dynamic Bayesian networks. In C Boutilier and M Godszmidt, editors, *Uncertainty in artificial intelligence*, pages 176–183. Morgan Kaufmann Publishers, 2000.

[13] R Morales-Menéndez, N de Freitas, and D Poole. Real-time monitoring of complex industrial processes with particle filters. In *Advances in Neural Information Processing Systems 16*, Cambridge, MA, 2002. MIT Press.

[14] L E Garza. *Hybrid Systems Fault Diagnosis with a Probabilistic Logic Reasoning Framework*. PhD thesis, Center of Artificial Intelligence, ITESM Monterrey campus, Mexico, 2001.

[15] D Poole. The Independent Choice Logic for modelling multiple agents under uncertainty. *Artificial Intelligence*, 94(1-2):7 – 56, 1997.

[16] D Poole. Logic, knowledge representation and Bayesian Decision Theory. In *First International Conference on Computational Logic*, London, July 2000.

[17] A Doucet. On sequential simulation-based methods for Bayesian filtering. Technical Report CUED/F-INFENG/TR 310, Dept of Engineering, Cambridge University, 1998.

[18] R Dearden, F Hutter, R Simmons, V Verma, S Thrun, and T Willeke. Real-time fault detection and situational awareness for rovers: Report on the Mars technology program task. In *Proceedings of IEEE Aerospace Conference*, Big Sky, MY, March 2004.

[19] R Morales-Menéndez, F Cantú, and A Favela. Dynamic modelling of processes using Jump Markov Linear Gaussian Models. In *Modelling, Identification and Control*, Grindelwald, Switzerland, February 2004.