A Bayesian Approach to Efficient Diagnosis of Incipient Faults

Indranil Roychoudhury, Gautam Biswas and Xenofon Koutsoukos
Institute for Software Integrated Systems (ISIS)
Department of Electrical Engineering and Computer Science
Vanderbilt University
Nashville, TN 37235, USA
Email:{indranil.roychoudhury, gautam.biswas, xenofon.koutsoukos}@vanderbilt.edu

Abstract

Safe, reliable, and efficient operation of complex dynamical systems requires the ability to detect, isolate, and identify degradation in system components. Degradations are typically modeled as incipient faults, which are slow drifts in system parameters over time. This paper presents an efficient approach for the detection, isolation, and identification of incipient faults under uncertainty using a Dynamic Bayesian Network (DBN) approach. Initially a DBN is used as an observer to track nominal system behavior. Once a fault is detected, incipient fault hypotheses are generated using a variation of our qualitative TRANSCEND approach for abrupt fault isolation. A modified DBN that includes the active fault hypotheses is then used to isolate the true fault and estimate the rate of change in its parameter value.

1 Introduction

Safe, reliable, and efficient operation of complex systems requires the ability to detect, isolate, and identify degradation in system components. Degradations are often modeled as incipient faults, which are slow drifts in system parameter values over time. In our previous work, we have developed fault diagnosis schemes for abrupt faults, which are modeled as instantaneous changes in system parameter values at a point in time. The qualitative fault isolation (QFI) scheme is based on the analysis of transients in the dynamic system behavior [Mosterman and Biswas, 1999; Narasimhan and Biswas, 2006; Roychoudhury *et al.*, 2005; Daigle *et al.*, 2006]. This approach has to be modified to accommodate the temporal profile for incipient faults (see Fig. 1).

This paper presents an efficient approach for the diagnosis of incipient faults by combining a variation of the TRANSCEND qualitative fault isolation approach [Mosterman and Biswas, 1999] with a quantitative fault isolation and identification scheme that employs a Dynamic Bayesian Network (DBN) model of the system dynamics. In general, DBN-based diagnosis approaches for complex systems suffer from computational intractability because of the large number of nodes (i.e., system variables and possible fault hypotheses) that have to be included in the DBN model. In our approach, efficiency is achieved by performing the fault isolation and identification in two steps: (i) run an efficient qualitative fault isolation scheme to reduce the number of candidate hypotheses to a small number, and (ii) run a refined DBN model to uniquely isolate the single fault candidate and estimate the rate of change in its parameter value. The focus of this paper is on fault isolation and identification of incipient faults in continuous dynamic systems. We assume that only single, incipient faults occur in the system. This assumption is required for the qualitative analysis only[1]. The quantitative FII framework can handle multiple fault hypotheses.

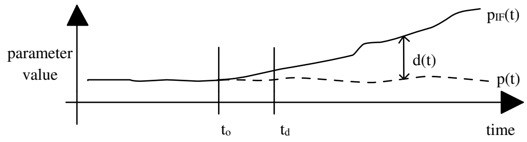


Figure 1: Incipient Fault Profile

The paper is organized as follows. Section 2 presents a mathematical definition of incipient faults and formulates our approach for solving the incipient fault diagnosis problem. Section 3 presents the incipient fault diagnosis architecture, and gives a brief overview of the fault detection, isolation, and identification subsystems. The different models employed for diagnosis are presented in Section 4. Section 5 explains in more detail the algorithms for incipient fault diagnosis. Section 6 presents results of applying this approach to a two tank system and conclusions are presented in Section 7.

2 Incipient Fault Diagnosis

A complete incipient fault diagnosis scheme must be tailored for detection, isolation, and identification (FDII) of incipient faults. Like earlier work, our diagnosis approach focuses on parametric component faults. In this framework, the mathematical representation of an incipient fault adds a drift term to the nominal component parameter value.

Definition 1 (Incipient fault) *An incipient fault profile in a dynamic system is characterized by a gradual drift in the cor-*

[1] Daigle, Koutsoukos, and Biswas (DX 2006) have developed an extension of the TRANSCEND scheme for multiple fault diagnosis

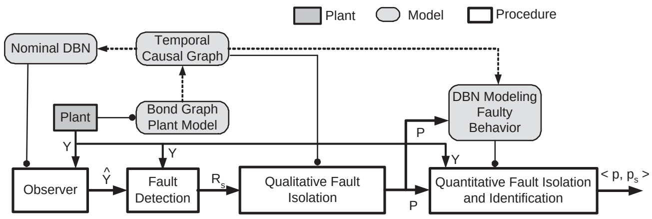


Figure 2: The diagnosis architecture

responding component parameter value from the time point of failure occurrence. The temporal profile for an incipient fault in parameter p, $p_{IF}(t)$ is given by:

$$p_{IF}(t) = \begin{cases} p(t) & t \le t_o \\ p(t) + d(t) & t > t_o \end{cases} \tag{1}$$

where $p(t)$ represents the nominal value of a parameter p over time, and $d(t)$ is the drift in the parameter value that gets added to the parameter value after occurrence of the fault, i.e., after $t \ge t_o$.

Fig. 1 shows an incipient fault profile, with t_0 as the time of occurrence of the fault. Since the rate of change of the parameter value is slow compared to the system dynamics, we can approximate the drift term, $d(t) = p_s(t - t_0),\ t \ge t_0$, where p_s is a constant that defines a linear rate of change, and t_0 is the time point at which the incipient fault first occurs.

2.1 Detection of Incipient Faults

Fault detection is the first step in any diagnosis process. The observer for tracking nominal behavior is based on a DBN model. This observer-generated expected behavior of the system is compared against the actual measurements using a Z-test for difference in means for robust fault detection [Biswas *et al.*, 2003].

Ideally, deviations in measurements caused by faults and degradations should be detected at or very soon after the point of fault occurrence. In reality, to accommodate measurement noise, inaccuracies in the model, and sensitivity of the detection scheme one has to trade-off false alarm generation versus detection delays. Statistical hypothesis testing schemes help reduce the false alarm rate, but introduce a delay between the time of occurrence and detection of faults, i.e., $t_d > t_o$. This detection delay, $t_d - t_o$, may pose convergence problems and reduce the parameter estimation accuracy. In our previous work on qualitative diagnosis [Manders and Biswas, 2003], we have shown that this delay does not affect diagnosis accuracy. In this approach, we assume this delay to be short enough not to affect qualitative diagnosis and the DBN-based estimation schemes. To ensure convergence of the DBN scheme, we start the estimation process from the time point at which the fault was detected.

2.2 Qualitative Fault Isolation

As the first step after fault detection, we employ a qualitative inference procedure using symbolic deviations and qualitative fault signatures for generating and refining possible fault hypotheses. This extends our previous work on transient analysis of abrupt faults [Mosterman and Biswas, 1999]. Unlike abrupt faults, which are modeled as a $\pm$ change in parameter value at the point of fault occurrence, incipient faults, characterized by slow drifts in parameter values (see Definition 1), are modeled qualitatively as $(0, \pm)$ change profiles, i.e., there is no change in the faulty parameter value at the point of fault occurrence but the parameter value slowly increases (decreases) over time. This fault profile matches any drift function $d(t)$ that is monotonic. Given such fault profiles, the TRANSCEND scheme for qualitative hypothesis generation and refinement can be applied for qualitative fault isolation. This methodology is outlined in Section 5.3.

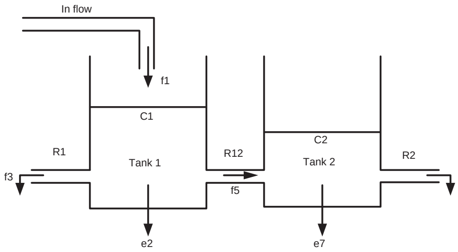
2.3 Quantitative Fault Isolation and Identification (FII) using DBNs

Quantitative FII is the final step in the fault diagnosis procedure. The TRANSCEND scheme discussed in Section 2.2, may not return an unique fault candidate, but it typically reduces the number of fault hypotheses to a tractable number. This makes it feasible to run a quantitative FII procedure using a DBN, outlined in Section 5.4, to refine the candidate set and estimate the drift parameter for the true fault candidate.

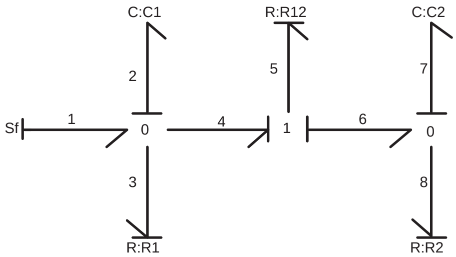
3 Architecture for Incipient Fault Diagnosis

The architecture of our model-based diagnosis methodology, presented in Fig. 2, follows a traditional diagnosis scheme for continuous systems. The system, as outlined in Section 2, includes four primary modules: (i) the observer, (ii) fault detector, (iii) the qualitative fault isolation unit, and (iv) the DBN-based FII unit. We build the dynamic plant model in the *bond graph* (BG) modeling language [Karnopp *et al.*, 2000] using a methodology where the components of interest in the system can be identified by one or more bond graph parameters, such as source elements, capacitors, inertias, resistance, and transformers. We derive the temporal causal graph (TCG) from the BG plant model using techniques that have been described earlier [Mosterman and Biswas, 1999]. The TCG, which is an extension of signal flow graphs, includes all the system variables as well as the component parameters that define dynamic system behavior. The TCG model is explained in greater detail in Section 4.1.

The observer is constructed as a DBN model of the nominal system. DBN tracking accommodates plant model inaccuracies and noisy measurements. Its inputs are the plant measurements, Y. The DBN is derived from the TCG model using the method described in [Lerner *et al.*, 2000], and outlined in Section 4.2. We use standard Bayesian propagation techniques [Russell and Norvig, 1995] to derive estimates of the most likely system state, $\hat{X}$, and measurement values, $\hat{Y}$ as plant behavior evolves. As discussed earlier, incipient fault parameters change at a very slow rate, which makes the detection of changes due to the incipient faults a hard problem since it becomes difficult to separate the measurement deviations from measurement noise and discrepancies caused by modeling inaccuracies. We employ statistical methods for robust fault detection. The input to the Fault Detector are the plant measurements Y and the observer-predicted measurements $\hat{Y}$. A significant difference in the observed and



(a) The two tank system schematic



(b) The bond graph model of the two tank system

Figure 3: The two tank system and its BG model

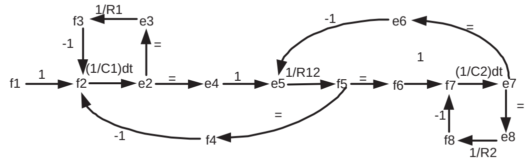


Figure 4: The temporal causal graph of the two tank system

expected behavior, $(Y - \hat{Y})$ signals a fault occurrence, and the qualitative residual signals R_s generated from the point of fault detection t_d are used for hypothesis generation and refinement.

When the fault detector triggers, the DBN observer is suspended and the TRANSCEND procedure is activated. The qualitative residual signals, R_s, are used for initial hypothesis generation, and for hypothesis refinement as additional measurements deviate using qualitative methods. All measurements from the time point of failure detection are also cached for use by the module. The qualitative scheme is terminated when one of the following conditions becomes true: (i) the number of fault candidates is reduced below a certain number, (ii) all measurement deviations have been used, or (iii) a pre-specified time horizon is exceeded. The DBN based FII scheme is then initiated with a DBN model of the faulty system behavior from the point of detection of the incipient fault. The set of current fault hypotheses, P are used to extend the nominal DBN to the fault DBN for tracking the system behavior after fault occurrence. Again, standard Bayesian update functions are employed, and with additional measurements the estimates converge to the true observed measurements. At this point, using least square estimation techniques, the rate of change of the fault is estimated. The output from the FII unit is the fault hypothesis and its rate of change, i.e., $< p, p_s >$. The steps outlined above are explained in detail in the following sections.

We believe that this approach provides an efficient computational scheme for solving the incipient fault diagnosis problem in the presence of measurement noise and model uncertainty. The Z Test-based fault detection module performs quick and reliable incipient fault detection while avoiding false alarms. The isolation and identification process is made computationally simpler by combining the TCG based qualitative fault isolation and the DBN-based FII procedures. As presented in [Lerner *et al.*, 2000], FDII of incipient faults can be achieved by using a single DBN that models both the nominal as well as all possible faulty behavior of the system. However, this makes the number of possible fault hypotheses very large, and an exhaustive online tracking procedure is not computationally viable. For this reason, the procedure outlined in [Lerner *et al.*, 2000] involves dropping unlikely fault candidates to save on computation. It is, therefore, possible that a true fault is dropped early as its probability of occurrence is very small. Our diagnosis approach retains all possible faults without compromising on efficiency. This is achieved by starting the DBN-based FII procedure only after the TCG based hypothesis refinement, thereby reducing the number of nodes in the DBN.

4 Modeling

Any model-based diagnosis approach can only be as good as the models that form the core of the diagnosis methodology. As discussed earlier, component-based BGs form the core of our modeling framework for physical plants. Efficient models for diagnosis, the TCG, state space models, and the DBNs are all derived from the primary BG plant model. This section gives a brief summary of the different models that we employ for incipient fault diagnosis.

4.1 Temporal Causal Graph

A TCG can be described as a *diagnosis model* that captures dependencies (algebraic and temporal) between system variables as a causal structure. The TCG is derived directly from the bond graph model of the plant [Mosterman and Biswas, 1999]. The TCG derived from the BG model can be defined as follows.

Definition 2 (Temporal Causal Graph (TCG)) *A TCG is a directed graph* $< V, L, D >$. $V = E \cup F$, *where V is a set of vertices, E is a set of effort variables and F is a set of flow variables in the bond graph system model. L is the label set* $\{=, 1, -1, p, p^{-1}, pdt, p^{-1}dt\}$ *(p is a parameter name of the physical system model). The dt specifier indicates a temporal edge relation, which implies that a vertex affects the derivative of its successor vertex across the temporal edge.* $D \subseteq V \times L \times V$ *is a set of edges [Narasimhan and Biswas, 2006].*

Fig. 3(a) shows the schematic of a two tank system that we will use as an example in this paper. The system comprises a couple of interconnected tanks, each having an outflow pipe for draining the tank. The first tank also has a source of flow for filling the tank. Fig. 3(b) shows the bond graph model. Bonds drawn as half-arrows capture the energy-exchange pathways in the system. Pipes are modeled as resistances and the tanks are modeled as capacitances. Pipes $R1$ and $R2$ drain tanks $C1$ and $C2$, respectively, and pipe $R12$ connects the two tanks $C1$ and $C2$. Fig. 4 shows the TCG for the two tank system. Temporal relations in the TCG are associated with the energy storage elements, i.e., the tanks. All other relations in the TCG, e.g., the pressure-flow relations imposed by the pipes and the idealized junction relations, are algebraic.

4.2 The DBN Observer for the Nominal System

The DBN observer for the nominal system is constructed from the TCG, as outlined in [Lerner *et al.*, 2000]. The DBN model is made up of two components:

1. A regular Bayes net that captures the relations between system variables at any time slice t. This consists of four sets of variables (X_t, Z_t, U_t, Y_t), which represent the state variables, other hidden variables, input variables, and measured variables for the dynamic system, and
2. A two-slice temporal Bayes net that captures the across-time relations defined by the state equation model of the dynamic system. We assume that the state equation model is a discrete-time stochastic process that satisfies the first order Markov assumption. Therefore, the across time links between time slices t and $t+1$ are defined by the system state equations.

For the two tank system, the DBN derived from the TCG has the following variables at time t: $X_t = \{e2_t, e7_t\}$, the pressures at the bottom of tanks 1 and 2, respectively, $U_t = \{f1_t\}$, the flow into tank 1, and $Y_t = \{f2_t, f8_t, f5_t\}$, the outflows from tanks 1 and 2, respectively and the flow between tanks 1 and 2. $Z_t = \phi$, i.e., the two tank dynamic model requires no additional variables.The across-time model includes five links, $e2_t \rightarrow e2_{t+1}$, $e7_t \rightarrow e7_{t+1}$, $e2_t \rightarrow e7_{t+1}$, $e7_t \rightarrow e2_{t+1}$, and $f1_t \rightarrow e2_{t+1}$. These links are directly derived from the state space model of the system. Fig 5(a) shows the DBN observer for time steps t and $t+1$.

4.3 The DBN Diagnoser

Model-based diagnosis schemes require the models to represent both the nominal and faulty system behavior. The DBN observer derived from the system TCG model represents a stochastic model of nominal system behavior in Fig. 5(a). Tracking of faulty behavior requires a stochastic model that captures incipient fault effects. The procedure for deriving this DBN is also detailed in [Lerner *et al.*, 2000]. To capture faulty system behavior, two sets of nodes are added. The first set correspond to parameters that represent the incipient fault hypotheses. The second set are discrete-valued nodes that are in 1-1 correspondence with the fault parameters, and they indicate the absence or presence of an incipient fault for that parameter. Fig. 5(b) shows the DBN diagnoser for faulty behavior of the two tank system, assuming two potential fault hypotheses, $\{R2, R12\}$. In other words, the DBN for faulty behavior now has an extended set X_t that includes $\{D2_t, D12_t\}$ in addition to $\{e2_t, e7_t\}$. The D's are logical variables. A value of 1 implies that the linked parameter has an incipient fault. A value of 0 implies no fault. This introduces additional across time links, $D2_t \rightarrow D2_{t+1}$, and $D12_t \rightarrow D12_{t+1}$. In addition, $Z_t = \{R2_t, R12_t\}$. The set of possible fault hypotheses covered by this DBN model include: (i) neither $R2$ or $R12$ faulty, (ii) $R2$ faulty, $R12$ not faulty, (iii) $R2$ not faulty, $R12$ faulty, and (iv) $R2$ and $R12$ faulty.

The DBN diagnoser model proposed in [Lerner *et al.*, 2000] includes all possible faults in the system. However, the number of possible faults can be really large in complex systems causing complexity issues in tracking diagnostic behavior using a Bayesian approach. In our work, we reduce the set of possible fault hypotheses using the TRANSCEND scheme, and the DBN model for FII only deals with the active fault candidates when the qualitative scheme terminates. This reduces the size of the DBN diagnoser and it results in a considerable improvement in the efficiency of the diagnosis.

5 Fault Detection, Isolation and Identification of Incipient Faults

This section presents the details of our methodology for implementing the different components of incipient fault diagnosis scheme.

5.1 Tracking Nominal Behavior Using a DBN

The DBN observer captures the nominal state of the system at every time step t. The set of nodes N_t in the DBN and their distributions provide a snapshot of the system state. A subset of these nodes, Y_t, correspond to measured variables in the system. The remaining variables belong to the set of system variables that cannot be measured, i.e., X_t and Z_t. Without loss of generality, we simplify the subsequent discussion, by considering only the variable set X_t and ignoring Z_t. The tracking problem for the system observer can be defined as deriving the posterior probability $P(X_t|Y_{0:t})$ at every time step t.

The *first order Markov assumption* reduces the computation of the posterior probability to

$$P(X_t|X_{0:t-1}) = P(X_t|X_{t-1}). \tag{2}$$

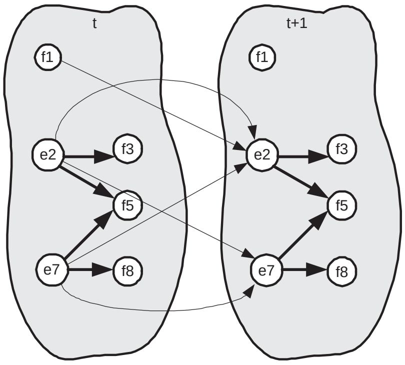
Moreover, the state space model of a physical system defines the system output (i.e., the measured variables) as a function of the state and the input variables. This implies,

$$P(Y_t|X_{0:t}, Y_{0:t-1}) = P(Y_t|X_t). \tag{3}$$

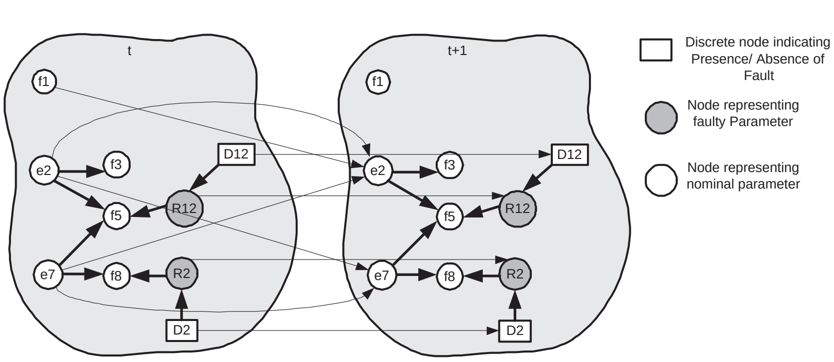
By combining equations (2) and (3), the tracking problem can be defined as an iterative problem [Russell and Norvig, 1995] defined as

$$P(X_{t+1}|Y_{0:t+1}) = \alpha P(Y_{t+1}|X_{t+1}) \sum_{X_t} P(X_{t+1}|X_t) P(X_t|Y_{0:t}),$$

where α is the normalizing constant. In this work, we assume that all random variables in the system are sampled from normal distributions. The noise models for the measurements are also assumed to be Gaussian with zero mean (white noise). Therefore, given prior probability distributions and the measurement noise models, the posterior probability



(a) The DBN Observer



(b) The DBN Diagnoser

Figure 5: The Nominal and Fault DBN Models for the two tank system

computations are reduced to estimating the mean and variances of the posterior Gaussian distributions.

The dependencies between the system variables may be non-linear, as is usually the case for real-life systems. As a simplification, tracking of the DBN model can be implemented as an *Extended Kalman Filter* (EKF) [Bar-Shalom and Fortmann, 1988], which is a classical approach for solving the tracking problem in such systems. The EKF approximates the nonlinear dynamics with linear dynamics and then uses the standard Gaussian model to update the system variables at the next step. We adapt the EKF method [Narasimhan and Biswas, 2006] for tracking the nominal system behavior.

5.2 Incipient Fault Detection

The fault detector continually monitors the measurement residual, $r_t = y_t - \hat{y}_t$, where $y_t \in Y_t$ are the measured variables at time t, and $\hat{y}_t$ are the expected value of the measurements as determined by the DBN observer. Ideally, $r_t \neq 0$ should imply a fault and trigger the fault isolation scheme, but to accommodate measurement noise and modeling errors we set up a statistical testing scheme to balance detection sensitivity against false alarms.

We start by defining a signal deviation at time step t in terms of an average residual for the last N_2 samples, i.e.,

$$\hat{\mu}_{N_{2t}} = \frac{1}{N_2} \sum_{i=t-N_2+1}^{t} r_i.$$

A hypothesis testing scheme based on the Z-test is employed to establish the significance of the deviation. To perform the Z-test, the variance of the measurement residual must be known. (For unknown variance the T-test may be performed, but its confidence interval is much larger.) To approximate the conditions necessary for the Z-test, the variance of the signal is estimated, but from a larger data set containing N_1 samples, i.e., $N_1 \gg N_2$:

$$\hat{\sigma}^2_{N_1 t} = \frac{1}{N_1 - 1} \sum_{i=t-N_1+1}^{t} \left(r_i - \mu_{N_1 t} \right)^2$$

The $Z-$value has a distribution $N(0,1)$:

$$Z = \frac{\hat{\mu}}{\frac{\sigma}{\sqrt{N_2}}}. \quad (4)$$

The confidence level, defined by α, defines the bound $[z_-, z_+]$: $P(z_- < z < z_+) = 1 - \alpha$. This bound can be transformed to another bound $[\mu_-, \mu_+]$ using Eqn. (4), and the approximation $\sigma = \hat{\sigma}_{N_1}$:

$$\mu_- = z_- \frac{\sigma}{\sqrt{N_2}}, \quad \mu_+ = z_+ \frac{\sigma}{\sqrt{N_2}}.$$

The Z-test is employed in the following manner:

$$\begin{aligned} \mu_- \leq \mu \leq \mu_+ \Rightarrow &\ \ no\ fault \\ otherwise \ \Rightarrow &\ \ fault. \end{aligned}$$

The advantage of this fault detection approach is that it is computationally simpler, and it makes no assumptions concerning the properties of the changed mean value (it does not have to be constant). Once the fault is detected, the Z-test outputs symbolically the direction of change of the observation, based on the value of the mean. If the mean is negative, this implies that the measurements have decreased from their nominal values, and a symbol $-$ is output. If the mean is positive, the observations have increased from the nominal values, and a symbol $+$ is output.

5.3 Qualitative Incipient Fault Isolation

After fault detection, the DBN tracking is suspended and the TRANSCEND fault isolation scheme [Mosterman and Biswas, 1999] is run on the TCG to generate the initial fault hypotheses given the first non-zero residual symbol(s). The TRANSCEND diagnostic framework for abrupt faults is extended to incipient fault analysis by considering fault profiles that have the value $(0, \pm)$ as was discussed in Section 2. The *backward propagation* scheme for generating the initial fault hypothesis remains the same.

For each fault hypothesis generated, a forward pass on the TCG, i.e., the *forward propagation* algorithm generates the fault signatures. Propagation of a $(0,+)$ or a $(0,-)$ will produce no discontinuous changes in the measured variables. Therefore, the predicted first effect of an incipient fault on a measurement can be expressed as one of three qualitative symbols: $\{+, 0, -\}$, which corresponds to a predicted gradual deviation above normal, no change, and a gradual deviation below normal, respectively, over some time interval. In

Fault	$e2$	$e7$	$f3$	$f5$	$f8$
$R1^+$	$+$	$+$	$-$	$+$	$+$
$R2^+$	$+$	$+$	$+$	$-$	$-$
$R12^+$	$+$	$-$	$+$	$-$	$-$

Table 1: Fault Signature Matrix

[Manders *et al.*, 2000] we have established that only the first change in a measured signal provides information to differentiate among fault hypotheses, therefore, it is sufficient to just record this first change, $\pm$ as the fault signature. The fault signatures for buildup of sediments in the three pipes of the two tank system (Fig. 3), causing their resistances to increase are listed in Table 1. Continued monitoring of the remaining measurement deviations helps refine the fault hypotheses using a matching process. If the observed deviation signal matches the predicted signature value, the fault hypothesis is retained, otherwise it is dropped.

The qualitative fault isolation algorithm is designed to run for at most s steps, where s is a pre-specified value. It may turn out that a single fault is isolated before the s steps are complete, or multiple hypotheses may still be valid after the s steps. When qualitative isolation identifies a unique candidate or the s steps are completed, the TCG based scheme is terminated and the FII module with the DBN diagnoser is initiated. The number of steps s must be carefully chosen. If s is too small, it is very likely that few fault candidates will be dropped and the ensuing DBN-based FII procedure will not be efficient. On the other hand, if s is large we may delay the isolation and identification tasks. A small number of remaining fault candidates implies a few “fault nodes” have to be introduced into the DBN diagnoser. This is good because the DBN approach is exponential in the number of number of fault hypotheses that are introduced. Too many hypotheses increase computation time and also the time to convergence.

5.4 Fault Isolation and Identification of Incipient Faults Using the DBN Diagnoser

Once the TCG based procedure completes running for s steps (or less than s steps if fault isolation completes earlier), the DBN diagnoser is modified to model the remaining fault hypotheses and the DBN-based FII scheme is initiated.

We implement a single DBN that includes all of the current fault hypotheses, i.e., the fault hypotheses that are not eliminated by the TRANSCEND analysis. Consider a specific scenario, where the TRANSCEND scheme reduces the fault hypothesis set to $\{R2, R12\}$. As discussed this introduces four additional nodes into the system DBN, i.e., $R2, R12, D2$, and $D12$. The set of possible fault hypotheses covered by the DBN model of the faulty system include: (i) neither $R2$ or $R12$ faulty, (ii) $R2$ faulty, $R12$ not faulty, (iii) $R2$ faulty, $R12$ not faulty, and (iv) $R2$ and $R12$ faulty. We assume that we have enough measurements such that the system, even with the addition of the faulty modes, is observable. The DBN FII scheme is initialized to the state of the system at time t_d, when the fault was detected (see Section 2.1). This is because $t_d - t_o$ is assumed to be small and error in starting the DBN-based FII scheme at t_d instead of t_o is negligible for the our diagnosis approach. Recall that all observations have been cached from the time a fault was detected and the DBN observer was suspended.

For the quantitative FII procedure, we adopt the procedure detailed in [Lerner *et al.*, 2000]. However, the computational complexity of our approach is greatly reduced because we start with the pruned set of fault hypotheses obtained from the qualitative TCG analysis. We maintain the belief state as a set of hypothesis, each of which corresponds to a single multivariate Gaussian distribution. A random variable p_t is introduced for each hypothesis (each hypothesis is defined as a parameter value that has changed), and the distribution of p_t corresponds to the likelihood for that fault hypothesis. Once the DBN with fault hypotheses is established, the same procedure for updating the likelihood for the nominal DBN can be applied to adjust the weights and the parameters of the multivariate Gaussians as each hypothesis is conditioned on the new measurements Y_{t+1}.

As more observations are collected, the mean value for the true fault parameter changes gradually, whereas the means of the other non-faulty parameters do not change. Moreover, the variances of each distribution should gradually decrease as more measurements are obtained. Observing the sequence of means, we can calculate the rate of change of the true fault parameter, thereby fulfilling the identification task for incipient faults. If at the end of the qualitative analysis, the set of fault hypotheses is refined to a singleton set containing only one fault, it implies that the system is diagnosable using the qualitative diagnoser. In that case, we add only one fault mode to the DBN-based diagnoser and the diagnoser is used solely for estimating the slope of the fault parameter.

6 Results

In this section, we present the results obtained by applying the proposed diagnosis approach to the two tank system shown in Fig. 3(a). In such hydraulic systems, the accumulation of sediment in the pipes are common examples of incipient faults. These incipient faults are modeled as a gradual increase in the pipe resistances and represented as $R1^+$ $R2^+$ and $R12^+$. $f3$, $f5$, and $f8$, the flow through the pipes $R1$, $R12$ and $R2$, respectively, are the measured variables for this experiment.

System behavior was generated for a total of 500 time steps by simulation using the Simulink®/MATLAB® environment. White noise (mean = 0, variance = 2% of the measured signal) was added to the measurements. The measurements were saved in a file, and then run through our incipient fault diagnosis scheme (implemented in MATLAB) to generate our experimental results.

We now describe a run of our diagnosis approach for a specific fault scenario. An incipient fault, i.e., a gradual buildup of resistance was introduced in pipe $R12$ at time-step, $t = 200$. The fault was modeled by a linear increase in the $R12$ parameter at rate of 0.0014 per time unit.

The introduction of the fault $R12^+$ first resulted in an decrease from nominal for $f5$, i.e., $f5 = -$. The fault detector Z-test signaled this deviation at time step $t = 219$, and then detected a increase from nominal in the measured value for $f3$, i.e., $f3 = +$, and then an decrease from nominal for $f8$, i.e., $f8 = -$ at time steps 266 and 371, respectively. This is shown in Fig. 6, where the flows after the introduction of the fault are compared with the flow values estimated by the observer. The forward propagation along the TCG implicated $R2^+$ and $R12^+$ as the possible fault candidates. The fault signatures, shown in Table 1 were used to match against the symbolic value of the measured variables. In this particular experiment, at the end of the TCG based analysis, $R2^+$ and $R12^+$ remained as fault candidates as the deviations observed in $f3$ and $f5$ could not refute the possibility of either fault.

The DBN-based diagnoser, representing the fault modes $R2^+$ and $R12^+$, was appropriately initialized and restarted from the time of detection of the fault, i.e., $t = 219$. All random variables in the DBN are assumed to be sampled from normal distributions with mean μ_p and variance σ_p. The means of every parameter is updated across time steps as follows:

$$\begin{bmatrix} \mu_{e2_{t+1}} \\ \mu_{e7_{t+1}} \end{bmatrix} = \begin{bmatrix} 1 - \frac{1}{C1}(\frac{1}{R1} + \frac{1}{R12}) & \frac{1}{C1R12} \\ \frac{1}{C2R12} & 1 - \frac{1}{C2}(\frac{1}{R2} + \frac{1}{R12} \end{bmatrix} \begin{bmatrix} \mu_{e2_t} \\ \mu_{e7_t} \end{bmatrix} + \begin{bmatrix} \frac{1}{C1} \\ 0 \end{bmatrix} Sf_t$$

$$\begin{bmatrix} \mu_{f3_{t+1}} \\ \mu_{f5_{t+1}} \\ \mu_{f8_{t+1}} \end{bmatrix} = \begin{bmatrix} \frac{1}{R1} & 0 \\ \frac{-1}{R12} & \frac{1}{R12} \\ 0 & \frac{1}{R2} \end{bmatrix} \begin{bmatrix} \mu_{e2_t} \\ \mu_{e7_t} \end{bmatrix}$$

$$\begin{bmatrix} \mu_{R12_{t+1}} \\ \mu_{R2_{t+1}} \end{bmatrix} = \begin{bmatrix} \frac{\mu_{e2_{t+1}} - \mu_{e7_{t+1}}}{\mu_{f5_{t+1}}} \\ \frac{\mu_{e7_{t+1}}}{\mu_{f8_{t+1}}} \end{bmatrix}$$

At every step, the mean and variance of the distributions of each parameter is updated and the estimated observations are compared with the actual faulty behavior. As the estimates are conditioned on more evidence, i.e., measurements, the estimation of the true fault parameter should result in predicted behavior models that match the measured system variables, while the estimates obtained from the "other" hypotheses will produce estimates that imply no change in its parameter value, or the estimated change has a very low likelihood given the measurements. The Z-test described earlier is applied to the measured flow estimates corresponding to each of the four hypotheses to determine if there is a significant deviation from the observed faulty measurements. If the Z-test determines a deviation in the residual for a certain hypothesis, that particular hypothesis is no longer considered to be valid.

In this way, at $t = 477$, the deviation in estimates for $R2^+$ is established using the Z-test, and $R12^+$ is correctly isolated as the true fault. The means of the distribution for $R12$ at each time step from $t = 219$ is logged and using standard least square estimation, the slope of change is identified. The rate of change of the faulty parameter was identified to be 0.00138 which is close to the actual injected rate of 0.0014 with a percentage error of 1.43%.

Fig. 6(a) shows the plots for (i) the estimated nominal flow f3 estimated by the observer, (ii) the measured actual flow f3 with the fault injected at $t = 200$, (iii) the estimated flow f3 with $R2^+$ as the only fault hypothesis, and, (iv) the estimated flow f3 with $R12^+$ as the only fault hypothesis. As the true fault is $R12^+$, we can see that the estimated flow f3 with $R12^+$ as the only hypothesis converges to the observed flow where as the estimates of f3 with $R2^+$ as the only hypothesis do not. Thus the Z-test detects a deviation for $R2^+$ and hence it is dropped as the fault hypothesis, isolating $R12^+$ as the true fault. Similar plots for the flows f5 and f8 are shown in Fig. 6(b) and Fig. 6(c). Table 2 summarizes the results for experiments where $R1^+$ and $R2^+$ are introduced as faults one by one.

Fault	Rate of fault	Time of fault injection	Time of fault detection	Time of QFI	Time of DBN-based FII	Estimated rate of fault
$R1^+$	0.0021	200	205	305	305	0.0022
$R2^+$	0.0022	200	218	343	449	0.0024
$R12^+$	0.0014	200	219	371	477	0.00138

Table 2: Experimental results (all times are expressed as time steps from the start of the experiment)

7 Conclusions

In this paper, we presented an efficient approach for diagnosis of incipient faults using a combined qualitative and quantitative DBN-based estimation scheme. The DBN-based FII approach allows for robust diagnosis under uncertainty that can be attributed to measurement noise and modeling errors. However, for large practical systems, the DBN based approach becomes computationally very expensive. To address this issue, in our approach, the fault hypotheses is first refined to a smaller set of candidates using qualitative fault isolation approaches. The DBN is then built for this reduced number of fault hypotheses alone making it more efficient than one which contains all possible fault hypotheses.

One issue that needs further investigation is the observability of the DBN diagnoser and its impact on diagnosis. For example, in the two tank system shown in Fig. 3, it is sufficient to measure the pressure $e7$ and the flow $f3$ to uniquely isolate the fault hypotheses (Table 1). However, for quantitative FII, it will be necessary to measure all three flows, $f3$, $f5$, and $f8$ in order to estimate the appropriate resistance values at each time step. The problem of identifying the correct set of measurements such that the system is diagnosable as well as the DBN is observable, therefore, is an interesting research issue.

In our experiments, we assumed that the prior and conditional probabilities for the DBN are all Gaussian. Moreover, the parameters of the DBN were also assumed to have a Gaussian distribution. However, this is a rather strong assumption and we need to relax it and demonstrate the efficiency of our diagnosis scheme for more general systems.

Finally, even though the qualitative fault isolation procedure is designed for diagnosis of single faults, the DBN based FII approach has no such restrictions. Hence, a natural extension of this work would be to adapt it for the detection of multiple incipient faults. In future, we intend to also extend this Bayesian approach to the diagnosis of both incipient and abrupt faults.

Acknowledgement

This work was supported in part by NSF CNS-0452067 and NSF CNS-0347440.

References

[Bar-Shalom and Fortmann, 1988] Y. Bar-Shalom and T. E. Fortmann. *Tracking and Data Association*. Academic Press, 1988.

[Biswas *et al.*, 2003] G. Biswas, G. Simon, N. Mahadevan, S. Narasimhan, J. Ramirez, and G. Karsai. A robust method for hybrid diagnosis of complex systems. In *Proc. 5th IFAC Symp on Fault Detection Supervision Safety Technical Processes*, pages 1125–1131, Washington, DC, June 2003.

[Daigle *et al.*, 2006] M. Daigle, X. Koutsoukos, and G. Biswas. Distributed diagnosis of coupled mobile robots. In *Proc. of 2006 IEEE International Conference on Robotics and Automation*, May 2006. to appear.

[Karnopp *et al.*, 2000] D. C. Karnopp, D. L. Margolis, and R. C. Rosenberg. *Systems Dynamics: Modeling and Simulation of Mechatronic Systems*. John Wile & Sons, Inc., New York, NY, USA, 3rd edition, 2000.

[Lerner *et al.*, 2000] U. Lerner, R. Parr, D. Koller, and G. Biswas. Bayesian fault detection and diagnosis in dynamic systems. In *Proc. of Seventeenth National Conference on Artificial Intelligence*, pages 531–537, 2000.

[Manders and Biswas, 2003] E.-J. Manders and G. Biswas. FDI of abrupt faults with combined statistical detection and estimation and qualitative fault isolation. In *Proc. 5th IFAC Symp on Fault Detection Supervision Safety Technical Processes*, pages 347–352, Washington, DC, June 2003.

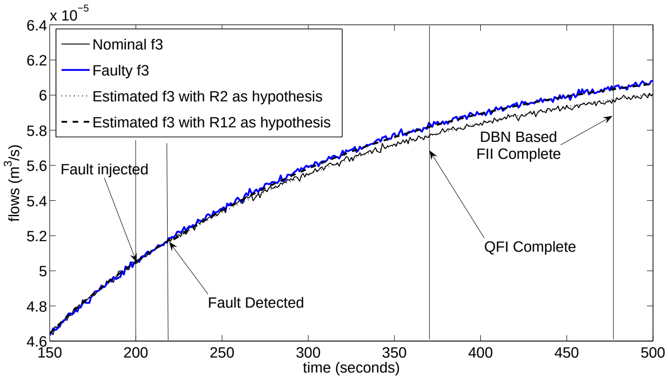
[Manders *et al.*, 2000] E.-J. Manders, S. Narasimhan, G. Biswas, and P. J. Mosterman. A combined qualitative/quantitative approach for fault isolation in continuous dynamic systems. In *Proc. 4th IFAC Symp on Fault Detection Supervision Safety Technical Processes*, pages 1074–1079, Budapest, Hungary, June 2000.

[Mosterman and Biswas, 1999] P. J. Mosterman and G. Biswas. Diagnosis of continuous valued systems in transient operating regions. *IEEE-SMCA*, 29(6):554–565, 1999.

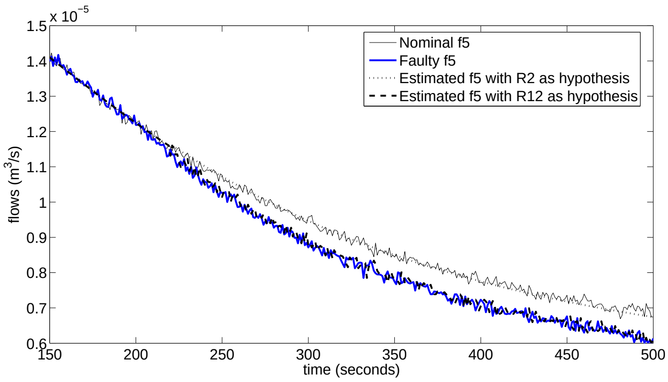
[Narasimhan and Biswas, 2006] S. Narasimhan and G. Biswas. Model-based diagnosis of hybrid systems. *IEEE Transactions on Systems, Man, and Cybernetics, Part A*, Sept. 2006. to appear.

[Roychoudhury *et al.*, 2005] I. Roychoudhury, G. Biswas, X. Koutsoukos, and S. Abdelwahed. Designing distributed diagnosers for complex systems. In *Proceedings of the 16th International Workshop on Principles of Diagnosis*, pages 31–36, Monterey, California, June 2005.

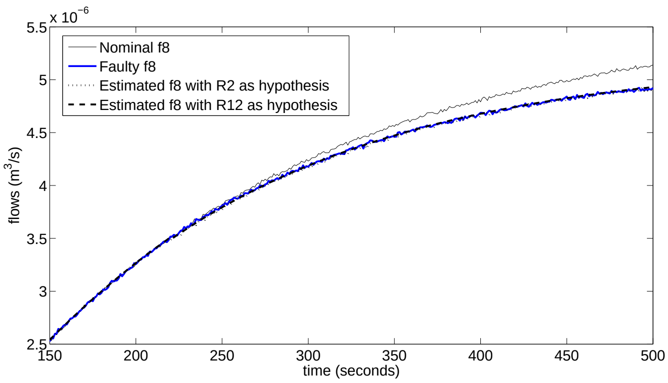
[Russell and Norvig, 1995] S. J. Russell and P. Norvig. *Artificial Intelligence: A Modern Approach*. Prentice-Hall Inc., 2nd edition, 1995.



(a) Flow f3



(b) Flow f5



(c) Flow f8

Figure 6: Tracking of flow measurements for two of the four fault hypotheses $\{R2^+, R12^+\}$