# ROBUST DIAGNOSIS OF SWITCHING SYSTEMS

**Sherif Abdelwahed Gabor Karsai Gautam Biswas**

*Institute for Software Integrated Systems*
*Vanderbilt University*
*Nashville, TN, 37203*

Abstract: This paper presents an approach for robust diagnosis of switching systems based on an extended version of the timed failure propagation graph model. The extended failure propagation graph model is a labeled graph used for the representation of failure conditions and their propagation modeled as causal relations with timing properties for a general class of systems with both time-based and event-driven dynamics such as hybrid and discrete event systems. We introduce the extended model and describe the structure and main components of the failure detection and isolation system based on the proposed model.

Keywords: Failure diagnosis, Switching Hybrid Systems, Robust Diagnosis.

## 1. INTRODUCTION

Large engineering systems such as manufacturing systems, power networks, and chemical plants are usually designed for autonomous operation. Automatic failure diagnosis forms a necessary part of these systems. Accurate and speedy diagnosis of faults is vital to their health and efficiency. In general, the diagnostic component aims to detect, isolate and predict possible failures by observing signals and measurements from the system sensors, comparing it with a mathematical model representing relevant nominal and/or faulty behavior, and explaining the observed behavior in terms of a hypothesis about possible abnormal changes to the state of the system components.

Failure analysis and diagnosis procedures can be classified into two main approaches. The first approach, commonly referred to as *model-based diagnosis*, compares the observed behavior of the system with a nominal model that captures its normal behavior. The nominal model may also contain information about faulty behavior. Current state information can be extracted from the available set of measurements and compared with the nominal values to identify and isolate possible faults. Diagnosability properties can also be formally defined and analyzed in this approach Patton et al. (1989).

The model based approach, however, depends on the availability of a precise mathematical model which is difficult to obtain for most practical real-life systems. Even when a precise model can be obtained the computational requirements of model-based diagnosis procedures are usually prohibitive. Nevertheless, there has been considerable research with many successful applications, particularly in the field of process engineering, for using model-based methods to detect and isolate system faults. Many of the developed techniques in this approach are based on filtering and parameter estimation Li and Olson (1991); Watanabe and Himmelblau (1983). In certain situations, analytical redundancy can be exploited to allow the

[1] Funded, in part, by DARPA's Software-Enabled Control Program under AFRL contract F33615-99-C-3611.

detection of measurement errors (sensor failures) Patton et al. (1989).

The other approach to system diagnosis is the *qualitative approach*. This approach is based on the causal relationship between observed signals and failure sources. Sensors signals are used to reason about possible failure based on the given causal relationship. The qualitative approach is more common in practice due its simplicity and computational efficiency. In addition, the requirements for this approach are easier to handle and usually do not require major changes to the system design. The underlying fault propagation model can be enriched to handle temporal, probabilistic and dynamical specifications. Also, support for integrated diagnosis of hierarchical systems can be easily established.

There are two primary models used for qualitative failure diagnosis - functional models and fault models. The former describe the correct behavior of the system with possible metrics added to associate behavioral deviations with particular fault patterns. Fault models on the other hand describe the system behavior in the presence of faults. Fault models have been used for diagnosis in work done by many researchers Ishida et al. (1985); Rao and Viswanadham (1987a,b). Timed failure propagation graphs (TFPG) Misra (1994); Misra et al. (1994) are causal models that describe the system behavior in presence of faults. The TFPG model is closely related to the fault model presented in Padalkar et al. (1991); Karsai et al. (1992) and used for an integrated fault diagnoses and process control system.

In this paper, we present a qualitative approach to failure diagnosis based on the timed failure propagation graph. The extended structure, referred to as hybrid failure propagation graph (HFPG) captures the effect of the switching dynamics and timing constraints on the propagation of failures in typical discrete event and hybrid systems. The HFPG model presented in this paper adds mode dependency constraints on the propagation links which can be used to handle failure scenarios in hybrid and switching systems. The HFPG model also supports both AND and OR node semantics which can be used to build complex failure propagation dependency situations. The proposed extension also allow cyclic dependency between signals (discrepancies) in the fault model.

The paper is organized as follows. Section 2 introduces the syntax and semantics of the hybrid failure propagation graph model. In section 3, we introduce the diagnosis problem and the main elements of the diagnostic system based on the hybrid failure propagation graph settings. In section 4 we present the diagnosis reasoning algorithm together with complexity analysis of its main procedures. Section 5 contains the conclusions of the paper and directions for future research.

## 2. HYBRID FAILURE PROPAGATION GRAPHS

The hybrid failure propagation graph is a labeled directed graph where the nodes represent either failure modes - which are fault causes - or discrepancies - which are off-nominal conditions that are the effects of failure modes. A discrepancy can either be monitored (attached to alarms) or silent, and depending on the way it is triggered by the incoming signals it is further classified as either "AND" or "OR" discrepancy. Attributed edges between nodes in the graph represent causality, and the attributes specify the temporality of causation given by an upper and lower time bounds on the propagation of failure between nodes.

The HFPG model allows for the representation of failure propagation in multi-mode (switching) systems in which the failure propagation relations depend on the current mode of the system. To this end, edges in the graph model can be constrained to a subset of the set of possible operation modes of the system. Formally, a hybrid failure propagation graph model is represented as a tuple $G = (F, D, E, M, \mathsf{ET}, \mathsf{EM}, \mathsf{DC}, \mathsf{DS})$, where:

- $F$ is a nonempty set of failure nodes,
- $D$ is a nonempty set of discrepancy nodes, with $F \cap D = \emptyset$
- $E \subseteq V \times V$ is a set of edges connecting the set of all nodes $V = F \cup D$. We will write $src(e)$ and $dst(e)$ for the source and destination nodes of the edge $e$, respectively.
- $M$ is a nonempty set of system modes. At each time instance $t$ the system can be in only one mode.
- $\mathsf{ET} : E \rightarrow I$ is a map that associate every edge in $E$ with a time interval in $I = \{[t_{min}, t_{max}] \mid t_{min} \in \mathbb{R}_+,\ t_{max} \in \mathbb{R}_+ \cup \{\infty\},\ t_{min} \leq t_{max}\}$; where $I$ is the set of all time intervals,
- $\mathsf{EM} : E \rightarrow \mathcal{P}(M)$ is a map that associate every edge in $E$ with a set of modes in $M$ (we assume that $\mathsf{EM}(e) \neq \emptyset$ for any edge $e \in E$ ),
- $\mathsf{DC} : D \rightarrow \{\texttt{AND}, \texttt{OR}\}$ is a map defining the class of each discrepancy as either `AND` or an `OR` node,
- $\mathsf{DS} : D \rightarrow \{\texttt{ON}, \texttt{OFF}\}$ is a map defining the monitoring status of the discrepancy as either `ON` for the case when the discrepancy is monitored by an online alarm or `OFF` for the case when the discrepancy is not monitored.

The set $V$ contains $n + m$ vertices, representing $n$ failure modes and $m$ discrepancies. Some of the discrepancies are monitored as defined by the

map DS. The set of monitored discrepancies will be denoted $D_a$. An edge $e = (v, v') \in E$ iff the failure effect represented by the node $v$ can propagate and participate in causing the effect represented by the node $v'$. The map ET associates each edge $e \in E$ with the minimum and maximum time (given as interval) for propagation of failure along the edge. We will write $t_{min}(e)$ and $t_{max}(e)$ for the minimum and maximum time for failure propagation along the edge $e$, respectively, so that $\mathsf{ET}(e) = [t_{min}(e), t_{max}(e)]$. The map EM associates each edge $e \in E$ with a subset of the system modes at which the failure can propagate along the edge. We assume the following assumptions hold for the graph structure $(V, E)$:

- $(\forall v \in V) \quad (v, v) \notin E$
- $(\forall e \in E) \quad dst(e) \notin F$
- $(\forall d \in D)(\exists v \in V) \quad (v, d) \in E$

The first assumption states that the graph does not contain self loops as the current version of HFPG only deals with permanent faults. The second assumption states that a failure node cannot be a destination of any edge so in effect failure nodes are the initial nodes of the graph. Finally, we assume that every discrepancy must be the destination of an edge, that is a discrepancy must be caused by either another discrepancy or failure mode.

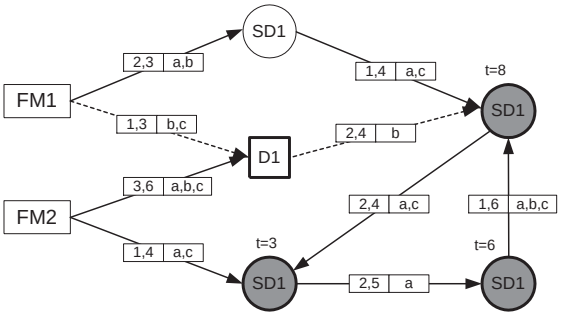


Fig. 1. A hybrid failure propagation graph

Figure 1 shows an example of a HFPG. In this graph, rectangles represent the failure modes while circles and squares represent OR and AND discrepancies, respectively. Monitored discrepancies are shown with bold lines. The arrows between the nodes represent failure propagation links. Propagation edges are parameterized by the corresponding interval, $[t_{min}, t_{max}]$, and the set of modes for which the edge is active. The above figure also shows a sequence of alarm signals, which are identified by shaded discrepancies. The time at which the alarm is observed is shown above the corresponding discrepancy.

## 3. THE DIAGNOSIS PROBLEM

The diagnostic system operates on the HFPG model described in the previous section and characterizes the fault status (actual current state) of the system by hypothesizing the faults in components and sensors based on the signals received from the sensors and the current mode of the system. The diagnoser uses the HFPG model and the timed sensor/mode-switching signals to generate a set of logically valid hypotheses of the current state of the system. The hypotheses are then ranked according to certain criteria that is generally based on the number of supporting alarms versus the number of inconsistent ones. A more advanced ranking that takes into account the relative significance/accuracy of the sensor signals can also be established. The set of hypotheses with the highest rank will be selected as the most plausible estimations of the current state of the system.

The diagnoser is implemented as a reactive module that is triggered by signals from the set of active sensors as well as mode-switching signals. Formally, a diagnoser input signal is represented by an event tuple $(e, t)$, where $e \in D_a \cup M$ is either a monitored alarm ($e \in D_a$) or a mode-switching signal ($e \in M$) and $t$ is the time at which the signal is observed. The timed event $(e, t)$ is triggered whenever the state of a discrepancy is changed or the system changes mode. When the event $(d, t)$ is triggered with $d \in D_a$, we say that $d$ is an active discrepancy and write $t_a(d)$ to denote the time when $d$ becomes active. The diagnostic system keep a record of the sequence of all timed events from the system initial state to the current time.

The diagnoser responds to input signals by generating hypotheses. Each hypothesis is an evaluation of the status of a failure mode in the HFPG model together with the corresponding evidence. Formally, a hypothesis is a tuple $h = (f, t_e, t_l, r, S_p, S_s, I, P)$, where $f \in F$ is the failure mode for which the hypothesis stands, $t_e$ and $t_l$ are the estimated earliest and latest time of occurrence of the failure mode $f$. The static rank, $r$, of the hypothesis is a number associated with a measure of belief in the hypothesis. The rank is set to 0 when the hypothesis is generated, and updated each time a new event is triggered. Hypotheses with negative ranks are not considered during the reasoning process. The elements $Sp$, $Ss$, $I$, and $P$ are sets of discrepancies with special relevance to the hypothesis $h$:

- $Sp \subseteq D_a$ is the set of primary active discrepancies that supports the hypothesis $h$. These are the active alarms that are triggered as a direct consequence of $f$. These alarms pro-

vide the main justification of the hypothesis $h$.

- $Ss \subseteq D_a$ is the set of secondary active discrepancies that supports the hypothesis $h$. These are the active alarms that are triggered as a consequence an alarm connected to $f$ and is supporting the hypothesis $h$.
- $I \subseteq D_a$ is the set of active discrepancies that are inconsistent with the hypothesis $h$. These are the alarms that are connected to the failure mode $f$ but cannot be explained based on the hypothesis $h$.
- $P \subseteq D_a$ is a set of pending discrepancies where their status cannot be identified at the current time.

The set of all supporting alarms for $h$ will be denoted by $S$. Note that the hypothesis $h$ also implicitly provides an estimation of the status of the monitored alarms connected to the failure mode $f$. That is $h$ is also a hypothesis for monitored alarms connected to $f$. Under the hypothesis $h$, supported alarms are considered healthy (providing the correct signals) and inconsistent alarms are faulty. In addition to generating and updating hypotheses, the diagnoser also generates a list of false alarms, namely those alarms that could not be explained by any hypothesis based on the timing and structure of the failure propagation graph. Based on the ranking scheme, the diagnoser can identify observation (sensor) errors. Consequently, the diagnosis process is robust against sensor failures and degrades gracefully as the number of sensor failures increase.

Given a HFPG representing the failure propagation in the system and a sequence of sensors signals corresponding to monitored discrepancies and possible mode switches, the task of the diagnosis system is to generate a failure report, which consists of a set of hypothesis that explains all the current signaling discrepancies. Figure 2 shows a simplified UML diagram of the basic elements of the HFPG diagnosis system and the relation between them.

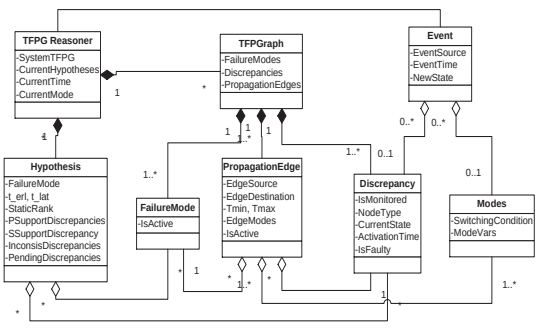


Fig. 2. A simplified diagram of the HFPG diagnosis system

In composing the hypothesis structure $h$ we consider only those discrepancies that are reachable from the underlying failure mode $f$. This allows the diagnoser to deal with sensor signals more efficiently by focusing on the nodes that are connected to the corresponding discrepancy. However, in general the HFPG structure may contain several failure modes that can propagate to certain common discrepancies. These mutual dependencies can lead to a conflict between the hypothesis of different failure modes, because as mentioned above $h$, also implicity provides an estimation of the status of monitored discrepancies connected to it.

Consider for instance the HFPG graph in Figure 1, assume $h_1$ is a hypothesis about FM1 that considers D1 as a primary supporting alarm and $h_2$ be a hypothesis about FM2 that considers D1 as an inconsistent alarm. Then clearly both hypotheses cannot be part of a consistent set of hypothesis as one considers $D1$ healthy while the other considers it faulty[2]. The diagnoser eliminates those sets of hypotheses that contain conflicting elements when generating the failure report.

## 4. DIAGNOSTIC REASONING

The diagnosis system operates on the HFPG model of the system to detect and isolate faults by generating and selecting appropriate hypothesis to explain the incoming signals from the system. In reasoning about the faults the diagnoser uses the principles of parsimony. In general, due to possible structural redundancy in the HFPG model, there can be several explanations of a given sequence of sensor signals. The *principle of parsimony* suggests that the simplest explanation is the best. By simplest we mean the one that involves the least number of faulty components. In general, there may not be a unique simplest explanation. In this situation, the diagnoser will provide all the most simple plausible explanation to the user.

At the occurrence of every event, the diagnoser updates the set of hypotheses and the current set of faulty components are identified. The diagnoser updates the set of possible hypotheses about the system state based on the causal and timing consistency between the discrepancies. Consistency between discrepancy nodes is calculated based on a complex formula as it depends on the type of the node and the current mode of the system. In general, the relation between active node is identified as either consistent, inconsistent, or pending.

[2] Note that this situation is independent of the type of D1, that is, the conflict between the two hypotheses remains if D1 is either of type `AND` or `OR`.

These classes can be either absolute (independent of any hypothesis) or relative based on certain hypothesis. For instance, let $(d, d') \in E$ where $d, d' \in D_a$ and both are of OR type and currently active. Then the pair $d, d'$ are absolute-consistent if the time since the edge $(d, d')$ became active is enough for the signal to propagate from $d$ to $d'$.

The diagnoser is triggered by one of four events - (1) a monitored discrepancy signaling alarm (2) a monitored discrepancy becoming silent after signaling for a while (3) a mode-switching signal (4) a timeout occurring, i.e., a predicted alarm did not ring. The reasoning algorithm uses two main data structures, the failure report and the HFPG model. The failure report consists of the highest ranking (most plausible) set(s) of consistent hypothesis. The HFPG model represents the dynamics of failure propagation and is used to both detect and isolate current faults and predicts future alarm signals. The reasoning algorithm consists of the following main steps.

**Update the reachability information** This step is invoked when a mode switching signal is detected. The mode switching signal indicates that the structure of the HFPG model changed, particularly, the connection between the nodes. The reachability matrices of the HFPG will be recalculated for each mode change, and activation time of each edge affected by the mode change will be updated.

**Update the hypotheses** In this stage, the faulty alarms are identified first. An alarm is considered faulty if it cannot be explained by one of the currently valid hypothesis. If the alarm is not faulty, then it is either incorporated into the current set of hypothesis or, if it is not consistent with any current hypothesis, a new hypothesis will be initiated. If the alarm is consistent with a hypothesis, the hypothesis timing and rank will be updated based on the time and state of the received alarm signal.

**Find missing and pending alarms** Silent monitored discrepancies are examined with respect to the current hypotheses set and their status with respect to these hypothesis are identified as either consistent, inconsistent or pending depending on the current time and the reachability information.

**Generate the failure report** In this stage, the current set of hypotheses are examined for possible conflicts due to common paths and nodes in the HFPG model. The ranks of consistent hypotheses sets are updated counting into effect mutual dependencies. The hypotheses set with the highest ranking is used to generate the failure report which contains an estimation of the current failure modes, their time of occurrence, and any possible sensor failures.

**Predict next alarms** The estimated information about the current failure modes and possible alarm failures is used to predict future alarms by checking propagation time from all to signalling alarms to non-signaling ones.

The total number of nodes in the HFPG model is $(n + m)$ where $n$ is the number of failure modes and $m$ is the number of discrepancies. The graph is implemented as an adjacency matrix, and therefore both BFS and DFS searching algorithms are $O(n + m)^2$. The worst case number of hypotheses is $O(nm)$. However, the number of hypotheses in typical practical situations is more likely to be within $O(n)$. Updating the hypotheses set is done by updating the consistency relation between nodes in the graph which is by search the graph recursively until the set reach a settling point (for the given hypothesis). This part is of polynomial complexity on the size of the graph and the current number of hypothesis. Resolving the conflict between hypothesis is done by generating all possible combinations of hypothesis and therefore is of exponential complexity with respect to the number of hypothesis.

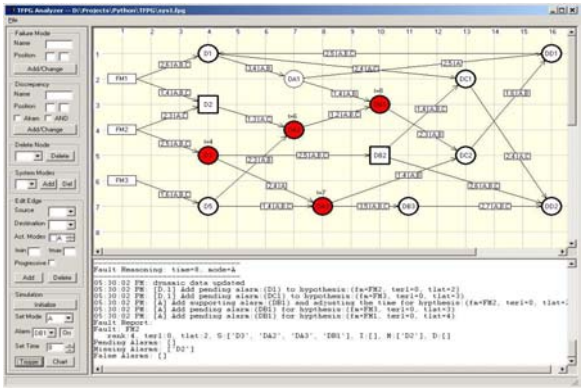

Fig. 3. The HFPG diagnostic tool

We have developed and tested a real-time diagnosis tool based on the hybrid failure propagation graph. The HFPG diagnosis tool is shown above in Figure 3. The reasoner engine in the HFPG tool is based on a robust incremental diagnostics algorithm described above. The tool can handle observation errors (sensor failure) and multiple fault scenarios. The tool also provides a simulation interface to test and evaluate the HFPG model for fault scenarios. The HFPG diagnosis algorithm is currently used as a part of the diagnostic module in a fault adaptive control structure aimed to support integrated fault diagnostics and control reconfiguration for large-scale and heterogeneous systems.

## 5. CONCLUSION

In this paper we introduced an approach for robust diagnosis of switching systems based on an extended version of the hybrid failure propagation graph model. The model can be used for diagnosis a general class of systems with mode switching conditions. We presented the main elements of the diagnostic system based on the hybrid failure propagation graph settings, and described the main parts of the diagnosis reasoning algorithm. In future work, we plan to enhance the efficiency of the diagnosis algorithm by incrementally identifying conflicting hypotheses at each time the set of hypotheses is updated.

## REFERENCES


Y. Ishida, N. Adachi, and H. Tokumaru. Topological approach to failure diagnosis of large-scale systems. *IEEE Trans. Syst., Man and Cybernetics*, 15(3):327–333, 1985.

G. Karsai, J. Sztipanovits, S. Padalkar, and C. Biegl. Model based intelligent process control for cogenerator plants. *Journal of Parallel and Distributed Systems*, 15:90–103, 1992.

R. Li and J. H. Olson. Fault detection and diagnosis in a closed-loop nonlinear distillation process. application of extended kalman filters. *Industrial Engineering Chemistry Research*, 30 (5):898–908, 1991.

A. Misra. *Sensor-Based Diagnosis of Dynamical Systems*. PhD thesis, Vanderbilt University, 1994.

A. Misra, J. Sztipanovits, and J. Carnes. Robust diagnostics: Structural redundancy approach. In *SPIE's Symposium on Intelligent Systems*, 1994.

S. Padalkar, J. Sztipanovits, G. Karsai, N. Miyasaka, and K. C. Okuda. Real-time fault diagnostics. *IEEE Expert*, 6(3): 75–85, 1991.

R. Patton, P. Frank, and R. Clark. *Fault Diagnosis in Dynamic Systems: Theory and Application*. Prentice Hall International, UK, 1989.

S. V. Nageswara Rao and N. Viswanadham. Fault diagnosis in dynamical systems: A graph theoretic approach. *Int. J. Systems Sci.*, 18(4):687–695, 1987a.

S. V. Nageswara Rao and N. Viswanadham. A methodology for knowledge acquisition and reasoning in failure analysis of systems. *IEEE Trans. Syst., Man and Cybernetics*, 17(2):274–288, 1987b.

K. Watanabe and D. M. Himmelblau. Fault diagnosis in nonlinear chemical processes. I: Theory. *AIChE*, 29, 1983.