Int. Workshop on the Principles of Diagnosis, Milan, Italy, October 1991.

the occurrence of the faults or from $f_i, \in F_1$ or from $f_j \in F_1$. Then for any $d \in D$, with $Node(d) = k$, $d \in NSD(f_i, f_j, \hat{t})$ iff:

$$(Reach_{ni,k} = 0 \wedge Reach_{nj,k} = 1 \wedge TMax_{nj,k} \leq \hat{t}) \vee$$
$$(Reach_{ni,k} = 1 \wedge Reach_{nj,k} = 0 \wedge TMax_{nj,k} \leq \hat{t}).$$

That is, $NSD(f_i, f_j, \hat{t})$ is a set of all the discrepancies reachable from either f_i or f_j but not both. The distance function between f_i and f_j is defined as $d(f_i, f_j, \hat{t}) = \frac{1}{|NSD(f_i, f_j, \hat{t})|}$.

4.2.3 Predictability

The algorithm is the same as the first algorithm for detectability. The implicants in this case are the discrepancies that can potentially cause the critical failures at least time $\hat{t}$ after their occurrence.

4.3 Combining Requirements

The algorithms described above generate advice for detectability, distinguishability and predictability separately. Obviously, there is a need to combine the alarm allocation suggested to satisfy the different criteria. One way to do so is to allocate alarms to all the discrepancies suggested under the different criteria since this would guarantee that all three criteria are satisfied. However, this method may lead to redundant alarm allocations since the criteria are not handled simultaneously.

Alarm allocation under simultaneous criteria can be done by computing the relative importance of discrepancies in satisfying the criteria using the algorithms described above. Then, alarms should be allocated to discrepancies in non-increasing order of their importance until all the criteria are met (just as is done when using hierarchical clustering method).

5 Conclusion

In this paper we have defined metrics to measure the diagnosability characteristics of a dynamical system in terms of – detectability, distinguishability and predictability. These metrics can be used to specify constraints that need to be satisfied when choosing an alarm allocation. Based on a fault propagation graph model of dynamic systems, algorithms have been developed for analyzing diagnosability characteristics of a given sensor placement, and for providing advice for alarm placement that meet selected criteria. The algorithms were used in the development of a diagnosability analyzer tool, which is used for the analysis of complex systems.

References

[1] Chang, S.J., DiCesare, F. and Goldbogen, G.,*Evaluation of Diagnosability of Failure Knowledge in Manufacturing Systems*, Proceedings, 1990 IEEE International Conference on Robotics and Automation, Vol 1, pp. 696-701.

[2] John F. Wakerly, *Digital Design Principles and Practices*, Prentice Hall, 1990.

[3] J. de Kleer and B. Williams, "Diagnosing Multiple Failures," *Artificial Intelligence*, vol. 32, 1987.

[4] Shogo Tanaka, "Diagnosability of Systems and Optimal Sensor Location," Chapter 5 in the book *Fault Diagnosis in Dynamic Systems: Theory and Application*, Prentice Hall International (UK), 1989, pp. 21-45.

[5] Ethan Scarl, "Diagnosability and Sensor Reduction," in *Proceedings of the Workshop on AI, Simulation and Planning in High Autonomy Systems*, Cocoa Beach, FL, 1991.

[6] S. Chien, R. Doyle and Nicolas Rouquette, "Sensor Placement for Diagnosability in Space-borne systems: A Model-based Reasoning Approach," *Proc. 2nd*

1. Form a $n \times k$ table T, where $k = |F_1|$. Let the failure modes in the set F_1 be $f_1, f_2, ..., f_k$.

2. For $i = 1$ to k do

   (a) $l = Node(f_i)$.

   (b) For $j = 1$ to n do

      i. Set $T_{j,i} = 1$ if $Reach_{l,j+m} = 1$ and $TMax_{l,j+m} \leq \hat{t}$ else set $T_{j,i} = 0$.

3. Apply the minimum cover algorithm as described in [2] on T to find the set of discrepancies that cover the failure modes in the set F_1. Alarms must be assigned to each of these discrepancies.

The problem with the minimum cover analysis is that the algorithm is NP-complete and can be slow if the number of alternative covers (not essential implicants) is high. For these cases we have developed an alternative approach, based on hierarchical clustering. We use the minimal spanning tree method for divisive clustering. The failure modes will make up the nodes of a complete graph. The similarity criterion for clustering the failure modes is the number of shared discrepancies. If a discrepancy is reachable by many failure modes, then by observing it, all of those failure modes can be detected. Thus we say that two failure modes are "close" to each other if they both share many discrepancies. A distance measure which expresses this observation is the reciprocal of the number of shared discrepancies.

Consider two failure modes $f_i,\ f_j \in F_1$, where $Node(f_i) = ni$ and $Node(f_i) = nj$. Let us denote by $SD(f_i, f_j, \hat{t})$ the set of shared disrepancies of $f_i,\ f_j \in F_1$ that are reachable after time $\hat{t}$ of the occurrence of the faults. Then for any $d \in D,\ \ with\ \ Node(d) = k$, $d \in SD(f_i, f_j, \hat{t})$ iff:

$$(Reach_{ni,k} = Reach_{nj,k} = 1) \text{ and } \left(TMax_{ni,k} \leq \hat{t}\right) \text{ and } \left(TMax_{nj,k} \leq \hat{t}\right)$$

The distance function between f_i and f_j is defined as $d(f_i, f_j, \hat{t}) = \frac{1}{|SD(f_i, f_j, \hat{t})|}$. We will build a graph, G, whose nodes will represent the $f \in F_1$. The graph G will be a complete graph with the length of edge $< i, j >$ between nodes representing failure modes $f_i, f_j \in F_1$, given by $d(f_i, f_j, \hat{t})$. If two failure modes do not share *any* discrepancy, the distance between the corresponding nodes in G will be *infinite*.

The algorithm to find an alarm allocation is:

1. Create G as defined above.

2. Find the minimal spanning tree T for the graph G.

3. Break the edges in T in non-increasing order of their length to generate clusters. Also keep track of which discrepancy(ies) are reachable from failure modes in which cluster and gives them points

4. Sort the discrepancies by the points given to them in non-increasing order. Then, one by one, allocate alarms for the discrepancies in this list in their non-increasing order of points until the detectability criteria is met.

4.2.2 Distinguishability

The algorithm for advising alarm allocation for distinguishability is based on hierarchical clustering and is very similar to the one used for detectability. The difference is the similarity criteria used, which in this case is the number of discrepancies *not* shared by the two. Since two failure modes can be distinguished more easily if they give rise to different sets of discrepancies, we say that the two failure modes are, in a sense, "close" each other and the distance between them is the reciprocal of the number of non-shared discrepancies.

Consider two failure modes $f_i,\ f_j \in F_1$ with $Node(f_i) = ni$ and $Node(f_i) = nj$. Let us denote by $NSD(f_i, f_j, \hat{t})$ the set of *not shared* disrepancies that are reachable after time $\hat{t}$) of

Criteria	Metric	Evauation	Advice
Detectability of failure mode f	$< t_{erl}, t_{lat} >$	find t_{erl} and t_{lat} $\forall f \in F$	$t_{lat} \leq \hat{t}$ $\forall f \in F_1$, $F_1 \subseteq F$, and given $\hat{t}$
Distinguishability of failure mode f	$DIS(f, \hat{t})$ $TRUE$ or $FALSE$ at given t	Compute $DIS(f,t)$ $\forall f \in F$ and given t	$DIS(f,t) = TRUE$ $\forall f \in F_1, F_1 \subseteq F$, and given $\hat{t}$
Predictability of discrepancy d	$< t_{min}, t_{max} >$	find t_{min} and t_{max} $\forall d \in D$	$t_{min} \geq \hat{t}, \forall d \in D_1$, $D_1 \subseteq D$ and given $\hat{t}$

Table 1: Diagnosability Analyses

to node j in FPG, $Reach_{i,j} = 1$ otherwise it is 0. $TMin_{i,j}$ is the minimum time a failure will take to propagate from node i to node j. $TMax_{i,j}$ is the maximum time a failure will take to propagate from node i to node j.

In the algorithm descriptions that follow, we will use the procedure $Node(failure)$ which, for any given $failure$, returns its corresponding node number in the failure propagation graph. The $failure$ might be a failure mode, denoted by f, or it might be an discrepancy, denoted by d.

4.1 Evaluator

The evaluator's task is to determine the detectability, distinguishability and predictability characteristics for any given alarm assignments.

Detectability: To compute the detectability of a failure mode in a FPG, we form a list of all the monitored discrepancies that the failure mode reaches. If the list is empty, the failure mode is not detectable. If the list is non-empty, its elements are sorted in non-decreasing order by using the minimum propagation time from the failure mode to the discrepancies as the key. The first item in the sorted list gives t_{erl}. Similarly, sorting the list in non-decreasing order using maximum time of propagation as the key gives t_{lat}.

Distinguishability: For finding the distinguishability of failure modes the algorithm has to check the uniqueness of the observed discrepancies after a time interval $\hat{t}$ passed since the fault occurrence. The problem can be reduced to the covering analysis described by Chang et al. in [1].

Predictability: To compute the predictability of a discrepancy d in a FPG, we form a list of all the observed discrepancies that reach d. If the list is empty, then d is not predictable. If the list is non-empty, its elements are sorted in non-increasing order by using the minimum propagation time from the observed discrepancies to the discrepancy d as the key. The first item in the sorted list gives t_{min}. Similarly, sorting the list in non-decreasing order using maximum time of propagation as the key gives t_{max}.

4.2 Adviser

The adviser's task is to suggest an alarm allocation, which satisfies specifications for detectability, distinguishability and predictability.

4.2.1 Detectability

The first algorithm to generate alarm allocation advice for detectability is based on finding minimum cover in a bipartite graph. The problem is equivalent with the well-known minimum cover analysis for logic function minimization [2]. The implicants in our case are the discrepancies reachable from the selected failure modes within time $\hat{t}$. The algorithm includes the following steps:

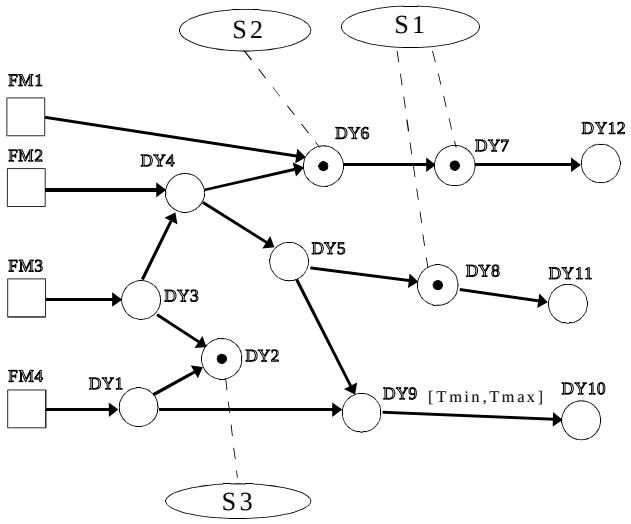

Figure 1: Failure Propagation Graph

pass before f can be included in the ambiguity set after its occurrence, and t_{lat} is the maximum time that will pass before its inclusion in the ambiguity set. If $t_{erl} = t_{lat} = \infty$, then f is not detectable.

Definition 2. The distinguishability of a failure mode f in time t, denoted by $DIS(f,t)$, is true if the ambiguity set contains only f when time t has passed after the occurrence of f, else it is false. Note that even if f is not distinguishable in time t, it might be distinguishable in time $t_1 > t$ because at time t_1 there might be more evidence (observed discrepancies) available for the diagnosis.

Definition 3. Predictability of a discrepancy d is the pair $< t_{min}, t_{max} >$, $t_{min} \leq t_{max}$, where t_{min} is the shortest available time period between a forewarning and the actual occurrence of d, and t_{max} is the longest available time period between a forewarning and the actual occurrence of d. If $t_{min} = t_{max} = 0$, d is not predictable at all.

Using the metrics defined above, we can perform several analyses, which are listed in Table 1. Evaluation of an alarm allocation means the computation of the values for the metrics defined above for each failure mode and discrepancy in the system. The advice column shows the use of the metrics to specify constraints for generation of an alarm allocation. The constraints consist of a set of failure modes and/or discrepancies and the limits on the time by when they should be detectable, distinguishable or predictable.

4 Algorithms

In this section, the basic algorithms that have been developed to tackle the problems described above are summarized. The algorithms operate on a failure propagation graph (FPG). We will assume that there are m failure modes and n discrepancies in a failure propagation graph. The nodes in the graph corresponding to failure modes are numbered from 1 to m, while nodes corresponding to discrepancies are numbered from $(m+1)$ to $(m+n)$. $Reach_{i,j}$ represents the reachability of nodes in the graph. If there is a path from node i

Our goal has been the development of techniques to analyze systems in terms of their diagnosability characteristics. and to determine the relative importance of sensors from the point of view of diagnosis. The results can be combined with other considerations on sensor placement so as to provide an economical sensor placement which meets the requirements.

2 Fault Models

The faults in a system and their interactions are modeled by using a labeled digraph to represent the dynamics of the system under fault conditions. An example of such a digraph, called Failure Propagation Graph (FPG) is shown in Figure 1.

The square boxes in the figure represent the *failure modes*, e.g., `valve stuck open`, of *components* in the system. The circles represent the anomalies in the system behavior, called *discrepancies*, e.g., `loss of flow`. The dotted circles represent those discrepancies that have *alarms* associated with them. These are called *monitored* discrepancies, which means that the occurrence of these discrepancies can be observed and will be indicated by "ringing" of the associated alarm. The empty circles represent discrepancies that do not have alarms associated with them and are called *non-monitored* discrepancies. An *alarm allocation* describes the assignment of alarms to discrepancies.

The ellipses represent the *sensors* in the system. Sensors measure the values of physical variables. The signals provided by these sensors are used to decide if a discrepancy exists and to generate (ring) the associated alarm. The dotted lines between the sensors and monitored discrepancies represent the (possibly very complex) mapping between sensors and the alarms they generate. A *sensor allocation* describes this mapping from sensors to alarms. To satisfy the diagnosability of a system, one needs to first determine an alarm allocation and then determine the sensor allocation needed. The research described here deals with issues involving alarm allocation.

The edges in the graph represent the propagations of failures and capture the interactions between different failures. Thus, an edge between two nodes means that the failure represented by the source node will propagate and cause the failure represented by the destination node. Each edge is labeled with a time interval $[T_{min}, T_{max}]$, which gives the minimum and maximum time that the source failure can take to propagate to the destination.

3 Problem Statement

For analyzing the diagnosability of a system we define three metrics which are based on these concepts related to fault management – fault detection, fault isolation and prediction of consequences of faults. In the definitions we will use the term *ambiguity set*, which is the set of all suspected failure modes of components. When discrepancies are observed, we suspect some failure modes to have occurred and include them in the ambiguity set and we say that those failure modes have been *detected*. As more evidence (observed discrepancies) comes in, we prune the ambiguity set down to the failure modes that have actually occurred, and we say that those failure modes have been *distinguished* (isolated).

In the following discussion, the set of all failure modes will be denoted by F, its subsets by $F_1, F_2 \ldots$, and the individual failure modes by $f, f_1, f_2 \ldots$ etc. The set of all discrepancies will be denoted by D, its subsets by $D, D_1, D_2 \ldots$ and individual discrepancies with $d, d_1, d_2 \ldots$ etc. The definition of the metrics to measure the diagnosability characteristics are the following:

Definition 1. Detectability of a failure mode f is the pair $< t_{erl}, t_{lat} >$, $t_{erl} \leq t_{lat}$, where t_{erl} is the minimum time that will

Diagnosability of Dynamical Systems

Amit Misra and Janos Sztipanovits
Department of Electrical Engineering
Vanderbilt University
Nashville, TN 37235

Alvin J. Underbrink, Jr., Ray Carnes and R. Byron Purves
The Boeing Company
Huntsville, AL

Abstract

Monitoring and diagnosis of complex dynamic systems are based on receiving inputs from sensors. The number, weight, size, reliability and cost of these sensors are important design concerns, particularly in remote systems like the Space Station. There is a need to minimize the costs associated with sensors without sacrificing the diagnosability of the system. We have defined three metrics: (1) DETECTABILITY, which gives the longest time that is needed to detect a failure, (2) PREDICTABILITY, which gives the shortest possible time between the forewarning and the actual occurrence of a failure, and (3) DISTINGUISHABILITY, which describes the size of the ambiguity sets given a time limit for the observation. Using these metrics, we perform two kinds of analyses. In EVALUATION mode, the diagnosability characteristics of a design with a predefined sensor allocation are calculated. In ADVICE mode, an arbitrary set of requirements can be defined for the diagnosability characteristics, and a satisfactory sensor placement is generated.

1 Introduction

Autonomous operation of large, complex systems such as chemical and power generation plants, and aerospace systems requires extensive monitoring, control, automated diagnosis, and fault recovery functions. These systems employ a large number of sensors which provide the data that these tasks use. Because of their utmost importance in these tasks, the selection and placement of sensors is a critical design task. This paper will focus on the problem of sensor allocation for diagnosability. The diagnostic functions of a system are best performed when they have an ample amount of sensor readings, but it is not always possible to put a sensor on every physical variable in the system. For example, in a remote system like the Space Station, reducing the number of sensors is an important design concern.

Recently, the attention of the research community has turned to the problem of sensor allocation for diagnosability. DeKleer et al. [3] have described a methodology to select the next point of measurement during diagnosis. Tanaka discusses diagnosability and optimal sensor allocation for linear discrete-time dynamical systems in [4]. Scarl [5] proposes a method to derive Minimal Sensor Sets for device-centered, model-based systems. Chien et al. have presented [6] an approach to evaluating sensor placement on the basis of its ability to distinguish normal operation and faulty operation and to discriminate between different kinds of faults.