Dynamic Multiple Fault Diagnosis: Mathematical Formulations and Solution Techniques

Satnam Singh, Kihoon Choi, Anuradha Kodali and Krishna Pattipati
Department of Electrical and Computer Engineering,
University of Connecticut, 371 Fairfield Road, U-2157,
Storrs, CT 06269, USA
{krishna, satnam}@engr.uconn.edu

John W. Sheppard
Department of Computer Science,
The Johns Hopkins University,
3400 N. Charles Street,
Baltimore, MD 21218, USA

Setu Madhavi Namburu, Shunsuke Chigusa, Danil V. Prokhorov and Liu Qiao
Toyota Technical Center USA
1555 Woodridge, RR #7
Ann Arbor, MI 48105, USA

Abstract

Dynamic multiple fault diagnosis (DMFD) is a challenging and difficult problem due to coupling effects of the states of components and imperfect test outcomes that manifest themselves as missed detections and false alarms. The objective of the DMFD problem is to determine the most likely temporal evolution of fault states, the one that best explains the observed test outcomes over time.

Here, we discuss four formulations of the DMFD problem. These include the deterministic situation corresponding to a perfectly-observed coupled Markov decision processes, to several partially-observed factorial hidden Markov models ranging from the case where the imperfect test outcomes are functions of tests only to the case where the test outcomes are functions of faults and tests, as well as the case where the false alarms are associated with the nominal (fault-free) case only. All these formulations are intractable NP-hard combinatorial optimization problems. We solve each of the DMFD problems by decomposing them into separable subproblems, one for each component state sequence.

Our solution scheme can be viewed as a two-level coordinated solution framework for the DMFD problem. At the top (coordination) level, we update the Lagrange multipliers (coordination variables, dual variables) using the subgradient method. The top level facilitates coordination among each of the subproblems, and can thus reside in a vehicle-level diagnostic control unit. At the bottom level, we use a dynamic programming technique (specifically, the Viterbi decoding or Max-sum algorithm) to solve each of the subproblems. The key advantage of our approach is that it provides an approximate duality gap, which is a measure of suboptimality of the DMFD solution. Interestingly, the perfectly-observed DMFD problem leads to a dynamic set covering problem, which can be approximately solved via Lagrangian relaxation and Viterbi decoding. Computational results on real-world problems are presented.

Introduction

Safety critical systems, such as aircraft, automobiles, nuclear power plants and space vehicles, are becoming significantly more complex and interconnected. The recent advances in wireless technology, remote communication, computational capabilities, sensor technology and standardized hardware/software interfaces have further increased the complexity of these systems. This complexity may result in failures of multiple components. Hence, there is a need to develop smart on-board diagnostic algorithms that can determine the most likely set of failure causes in a system, given observed test outcomes over time.

The multiple fault diagnosis (MFD) problem originates in several fields such as medical diagnosis [1], error correcting codes, speech recognition, distributed computer systems and networks [2]. The MFD problem in large-scale systems with unreliable tests was first considered by Shakeri et al. in [3]. They proposed near-optimal algorithms using Lagrangian relaxation and subgradient optimization methods for the static MFD problem. In the area of distributed system management, the MFD problem is studied by Odintsova et al. in [2]. They utilized an adaptive diagnostic technique, termed active probing, for fault diagnosis and isolation. A probe can be viewed as a test in our terminology; the purpose of a probe is to check the set of system components on the probed path. The probe outcomes determine if one or more of the components on the probed path are faulty or normal. Given the probe outcomes, a diagnostic matrix (D-matrix, diagnostic dictionary, reachability matrix) defining the relationship among the probes and component faults, and the initial system state, they developed a sequential multi-fault algorithm to diagnose the system state. They considered the probe outcomes as being deterministic, which is analogous to the assumptions made in our Problem 4, and in the work described in [11]-[14]. In [4], Le et al. applied graphical model-based decoding algorithms to the MFD problem in the presence of unreliable tests. They proposed a suboptimal belief propagation algorithm used to decode low density parity check codes. They considered a fault model, where tests are asymmetric, i.e., the D-matrix is not binary and the test outcomes are also unreliable, and they termed it the Y model. Their implementation is parallel to our Problem formulation 1; however, they considered only the static case.

The DMFD problem refers to determining the most likely temporal evolution of component states, given a set of partial and unreliable test outcomes over time. The dynamic single fault diagnosis problem using a hidden Markov model (HMM) formalism was first proposed by Ying et al. [5], where it is assumed that, at any time, the system has at most one fault state present. This modeling is somewhat unrealistic for most real-world systems. Another version of the dynamic fault diagnosis problem was studied in [6]: unknown probabilities of sensor error, incompletely-populated sensor observations, and multiple faults were allowed, but the faults could only occur or clear once per sampling interval. Another approach, developed by Ruan et al. [7], decomposes the original DMFD problem into a series of decoupled subproblems, one for each epoch. For a single epoch MFD, they developed a deterministic simulated annealing (DSA) method, which is inspired by its sibling stochastic simulated annealing and the approximate belief revision (ABR) heuristic algorithm [1]. This algorithm enlarges the search space of ABR via DSA, and is guaranteed to provide a solution no worse than (often significantly better) than ABR. The single epoch MFD was extended to incorporate fault states of multiple consecutive epochs. In addition, they applied a local search and update scheme to further smooth the "noisy" diagnoses stemming from imperfect test results and, thereby, increase the accuracy of fault diagnosis.

The DMFD problem can also be viewed as a factorial HMM (FHMM), a simplified Markovian dynamic Bayesian network, discussed in the machine learning literature [8]. Here, the HMM state is factored into multiple state variables, and is represented in a distributed manner. The authors in [8] discussed an exact algorithm for inference in FHMM. Here, the inference and learning involves computing the posterior probabilities of multiple hidden layers (or states), given the test outcomes. However, due to the combinatorial nature of the hidden state representation, the exact algorithm is intractable. They presented approximate inference algorithms based on Gibbs sampling and variational methods. The latter methods are similar to Lagrangian relaxation, although motivated from a Fenchel duality perspective [1], [17].

In our recent work [9], we extended the work of Ruan et al. [7], Shakeri et al. [3] and Tu et al. [11] on MFD to solve the DMFD problem by combining the Viterbi algorithm and Lagrangian relaxation in an iterative way. This paper is an extension of our work in [9]. Depending on the probabilistic assumptions on fault-test relationships and test outcomes, one obtains various DMFD formulations. In [9], we discussed only DMFD formulation 1 and it was solved only for small-scale systems. In this paper, we provide three other formulations of the DMFD problem along with their solutions. Here, we also compare the results between the subgradient and the deterministic simulated annealing methods [7]. Simulation results on several real world systems are provided for our earlier formulation of the DMFD problem (formulation 1).

DMFD Problem Formulations

The dynamic multiple fault diagnosis problem consists of a set of possible fault states in a system, and a set of binary test outcomes that are observed at each sample (observation, decision) epoch. Fault states are assumed to be independent. Each test outcome provides information on a subset of the fault states. At each sample epoch, a subset of test outcomes is available. Tests are imperfect in the sense that the outcomes of some of the tests could be missing, and tests have missed-detection/false-alarm processes associated with them. The observations consist of imperfect binary test outcomes, and are characterized by sets of passed tests outcomes, O_p and failed tests outcomes, O_f. Formally, we represent the DMFD problem as $DM = \{S, \kappa, T, O, D, P, A\}$, where $S = \{s_1, ..., s_m\}$ is a finite set of *m* components (failure sources) associated with the system. The state of component s_i is denoted by $x_i(k)$ at epoch *k*, where $x_i(k) = 1$ if failure source s_i is present; $x_i(k) = 0$, otherwise. Here, $\kappa = \{0, 1, ..., k, ...K\}$ is the set of discretized observation epochs. The status of all component states at epoch k is denoted by $\underline{x}(k) = \{x_1(k), x_2(k), ..., x_m(k)\}$. We assume that the initial state $\underline{x}(0)$ is known (or its probability distribution is known).

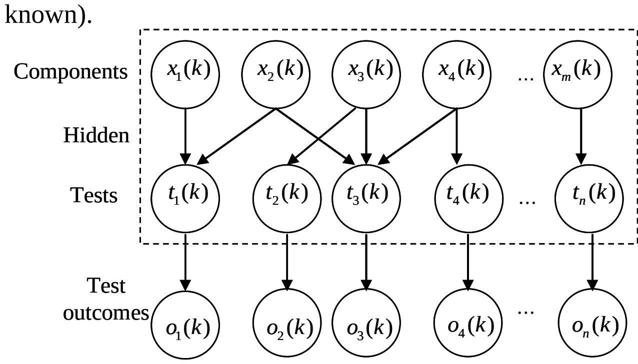

Figure 1: Tri-partite digraph for DMFD problem

The observations at each epoch are subsets of binary outcomes of tests $O=\{o_1,o_2,...,o_n\}$, i.e., $o_j \in \{pass, fail\}=\{0,1\}$. Figure 1 shows the DMFD problem as a tri-partite digraph at epoch *k*. Component states, tests and test outcomes represent the nodes of the digraph. Here, the true states of the component states and tests are hidden. $P=\{Pd, Pf\}$ represents a set of probabilities of detection and false alarm, which is defined differently for each of the DMFD problem formulations. We also define the matrix $D=[d_{ij}]$ as the dependency matrix (D-matrix), which represents the full-order dependency among failure sources and tests.

Each component state is modeled as a two-state non-homogenous Markov chain. For each component state, e.g., for component s_i at epoch *k*, $A=(Pa_i(k), Pv_i(k))$ denotes the set of fault appearance probability $Pa_i(k)$ and fault disappearance probability $Pv_i(k)$ defined as $Pa_i(k)=\Pr(x_i(k)=1 \mid x_i(k-1)=0)$ and $Pv_i(k)=\Pr(x_i(k)=0 \mid x_i(k-1)=1)$. These probabilities are required to model the intermittent faults. Here, $T=\{t_1,t_2,...,t_n\}$ is a finite set of *n* available binary tests, where the integrity of the system can be ascertained. We denote the set of passed tests, T_p and failed tests T_f. At each observation epoch, *k*, $k \in \kappa$, test outcomes upto and including epoch *k* are available, i.e., we let $O^k=\{O(b)=(O_p(b),O_f(b))\}_{b=1}^{k}$, where O^k is the set of observed test outcomes at epoch *k*, with $O_p(b)(\subseteq O)$ and $O_f(b)(\subseteq O)$ as the sets of passed and failed tests at epoch b, respectively. The tests are partially observed in the sense that outcomes of some tests may not be available, i.e., $(O_p(b) \cup O_f(b)) \subset O$. In addition, tests exhibit missed detections and false alarms. Here, we also make the noisy-OR ("causal independence") assumption [10].
The DMFD problem can be formulated in the following ways, arranged from the general to simplified:

Problem 1: When the probability of detection (Pd_{ij}) and false alarm probability (Pf_{ij}) are associated with each test and each fault class, i.e., $Pd_{ij}=\Pr(o_j(k)=1 \mid x_i(k)=1)$ and $Pf_{ij}=\Pr(o_j(k)=1 \mid x_i(k)=0)$ of a failure source s_i and test t_j. For notational convenience, when s_i does not affect the outcome of test t_j, we let the corresponding $Pd_{ij}=Pf_{ij}=0$. This problem scenario frequently arises in medical fault diagnosis. For example, the QMR-DT (Quick Medical Reference, Decision-Theoretic) database used in

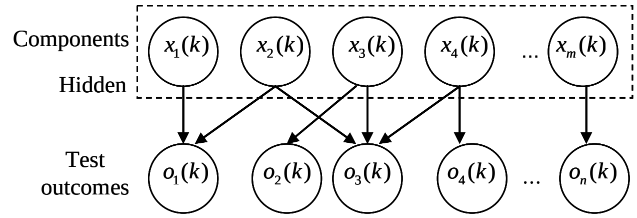

Figure 2: Bi-partite graph for DMFD problem 1 and 2

the domain of internal medicine, contains approximately 600 disease nodes (faults or failure sources) and 4000 symptoms (tests) [7]. Each of the symptoms could have a probability pair (Pd_{ij}, Pf_{ij}) associated with them. Figure 2 shows the bi-partite graph, where the edges represent the probability pair (Pd_{ij}, Pf_{ij}). These probabilities can be obtained from the tri-partite digraph (Figure 1) using the total probability theorem as follows:

$$\Pr(o_j(k) \mid x_i(k)) = \sum_{t_j \in \{0,1\}} \Pr(o_j(k), t_j(k) \mid x_i(k)) = \sum_{t_j \in \{0,1\}} \Pr(o_j(k) \mid t_j(k)) \Pr(t_j(k) \mid x_i(k)) \quad (1)$$

Problem 2: In situation where the probability of detection (Pd_{ij}) is associated with each failure source-test pair, but the false alarm probability is specified only for the normal system state, i.e., $Pf_j = P(o_j(k)=1 \mid x_1(k)=0,...,x_m(k)=0)$, we obtain a slightly complicated variation of Problem formulation 1 (in terms of computational complexity, but not in terms of parameterization). This type of scenario arises when we design class-specific classifiers that distinguish between normal system operation and failure source, s_i only, or when the false alarms are defined on an overall system basis. Here, the probability pair (Pd_{ij}, Pf_j) is associated with test outcomes to model imperfect test outcomes [3]. This model is also called the Z model in [4]. Similar to problem 1, the probability pair (Pd_{ij}, Pf_j) is shown as edges between the hidden component states and test outcomes in Figure 2, and they can be obtained from the tri-partite digraph (Figure 1) using the total probability theorem on the nodes of test layer.

Problem 3: When the probability of detection (Pd_j) and false alarm probability (Pf_j) are associated with each test t_j only. The probability pair (Pd_j, Pf_j) is shown as the edges between the tests and test outcomes in the tri-partite digraph (Figure 1). This formulation is quite useful in classifier fusion using error correcting codes. In the error correcting code (ECC) matrix, each column corresponds to a binary classifier with the associated (Pd_j, Pf_j) pair, which are learned during training and validation. This type of formulation is also considered in [6].

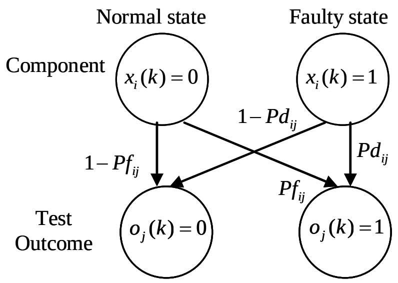

Figure 3: Detection and false alarm probabilities for problem 1

This formulation provides a nice vehicle for the dynamic fusion of classifiers, where each column of the ECC matrix is a classifier, and their associated probability pairs (Pd_j, Pf_j) are uncertainties associated with classifier outcomes. When the learned parameters and the ECC matrix are fed as an input to the DMFD algorithm, it performs dynamic fusion of classifier outputs over time. Note that the sampling interval of the dynamic fusion algorithm can be different from the sampling interval of the raw sensor data.

Problem 4: This is the deterministic case when tests are perfect i.e. $Pd_{ij}=1$ and $Pf_{ij}=0$ [11]. This formulation reduces the tripartite digraph in Figure 1 to a bipartite graph between the components and tests. This scenario is useful in situations where the tests are highly reliable (e.g., automated testing of electronic cards), and leads to a novel dynamic set covering problem.
Next, we discuss the DMFD formulations in detail.

DMFD Problem 1

In this problem, we assume that the detection and false alarm probabilities (Pd_{ij}, Pf_{ij}) are associated with each failure source and each test. Figure 3 illustrates these probabilities. We have presented this problem in detail in [9]. Here, we revise only the key steps of the problem and solution.

The DMFD problem is one of finding, at each decision epoch k, the most likely fault state candidates $\underline{x}(k)\in\{0,1\}^m$, i.e., the fault state evolution over time, $X^K=\{\underline{x}(1),...,\underline{x}(K)\}$, that best explains the observed test outcome sequence O^K. We formulate this as one of finding the maximum *a posteriori* (MAP) configuration:

$$\widehat{X}^K = \arg\max_{X^K} \Pr(X^K \mid O^K) \qquad (2)$$

In [9], we showed that we obtain optimal fault sequence $\widehat{X}^K$ using the following primal problem:

$$\widehat{X}^K = \arg\max_{X^K,Y^K} J(X,Y) = \arg\max_{X^K,Y^K} \sum_{k=1}^{K} f_k(\underline{x}(k),\underline{x}(k-1),\underline{y}(k)) \qquad (3)$$

where the fault state sequence is $X^K=\{\underline{x}(1),\underline{x}(2),...,\underline{x}(K)\}$ and $Y^K=\{\underline{y}(1),\underline{y}(2),...,\underline{y}(K)\}$ $\underline{y}(k)=\{y_j(k),\forall j\in O_f(k)\}$ are new variables such that

$$\ln y_j(k) = \sum_{i=1}^{m} c_{ij}x_i(k)+\eta_j, \quad \forall j\in O_f(k). \qquad (4)$$

Here, the primal objective function for an individual fault state, i.e., $f_k(\underline{x}(k),\underline{x}(k-1),\underline{y}(k))$ is defined as

$$f_k(\underline{x}(k),\underline{x}(k-1),\underline{y}(k)) = \sum_{o_j\in O_p(k)}\sum_{i=1}^{m} c_{ij}x_i(k)+\sum_{i=1}^{m}\mu_i(k)x_i(k)$$
$$+\sum_{o_j\in O_f(k)}\ln(1-y_j(k))+\sum_{i=1}^{m}\sigma_i(k)x_i(k-1)$$
$$+\sum_{i=1}^{m}h_i(k)x_i(k)x_i(k-1)+\gamma(k)+g(k) \qquad (5)$$

where, the parameters c_{ij}, $\gamma(k)$, η_j are functions of Pd_{ij} and Pf_{ij} and $\mu_i(k)$, $\sigma_i(k)$, $h_i(k)$, $g(k)$ are functions of $Pa_i(k)$, $Pv_i(k)$. Note that the multiple HMMs are coupled here because their states are observed only via a set of test outcomes. In equation (5), the terms involving $y_j(k)$ and $h_i(k)$ shows the coupling effects. The detailed steps of deriving the primal problem are provided in [9].

The primal DMFD problem posed in (3)-(5) is NP-hard which, for all practical purposes, means that, unless P=NP, it cannot be solved to optimality within a polynomially bounded computation time. The NP-hard nature of the primal DMFD problem motivates us to decompose it into a primal-dual problem using a Lagrangian relaxation approach. By defining new variables and constraints, the DMFD problem reduces to a combinatorial optimization problem with a set of equality constraints. The constraints are relaxed via Lagrange multipliers.

In [9], we also showed that the dual problem of the primal DMFD problem as posed in (3)-(5), can be written as

$$\min_{\Lambda} \quad Q(\Lambda)$$

subject to $\Lambda=\{\lambda_j(k)\ge 0, k\in(1,K), j\in O_f(k)\}$ (6)

where the dual function $Q(\Lambda)$ is defined by

$$Q(\Lambda)=\max_{X^K}\sum_{i=1}^{m}Q_i(\Lambda). \qquad (7)$$

Here

$$Q_i(\Lambda)=\sum_{k=1}^{K}\xi_i(x_i(k),x_i(k-1),\lambda_j(k))+\frac{1}{m}w_k(\Lambda) \qquad (8)$$

$$\xi_i(x_i(k), x_i(k-1), \lambda_j(k)) = \left(\sum_{o_j \in O_p(k)} c_{ij} + \mu_i(k) - \sum_{o_j \in O_f(k)} c_{ij}\lambda_j(k) \right) x_i(k) + \sigma_i(k)x_i(k-1) + h_i(k)x_i(k)x_i(k-1) \quad (9)$$

and

$$w_k(\Lambda) = \gamma(k) + g(k) + \sum_{\forall o_j \in O_f(k)} \left[\lambda_j(k)\ln\lambda_j(k) - (1+\lambda_j(k))\ln(1+\lambda_j(k)) - \lambda_j(k)\eta_j \right] \quad (10)$$

represents the dual function for the i^{th} component. The main benefit of (7) is that now the original problem is separable. Using the Lagrangian relaxation method, we decomposed the original DMFD problem into *m* separable subproblems, one for each component state sequence $\underline{x}_i$, where $\underline{x}_i = \{x_i(1), x_i(2), ..., x_i(K)\}$, $x_i(k) \in \{0,1\}$ and $i \in \{1, m\}$. The relaxation procedure generates an upper bound for the primal objective function. The procedure of minimizing the upper bound via a subgradient subgradient optimization produces a sequence of dual feasible, and the concomitant primal feasible solutions to the DMFD problem. If the objective function value for the best feasible solution and the upper bound are the same, the feasible solution is the optimal solution. Otherwise, the difference between the upper bound and the feasible solution, termed the approximate duality gap, provides a measure of suboptimality of the DMFD solution; this is a key advantage of our approach. Details of the DMFD algorithm, subgradient method and dynamic programming are provided in [9].

In this paper, we also compared results of subgradient method with deterministic simulated annealing method [7] in the results section.

DMFD Problem 2

In this formulation, we define Pd_{ij} as $Pd_{ij} = \Pr(o_j(k) = 1 | x_i(k) = 1)$ and $Pf_j = \Pr(o_j(k) = 1 | x_1(k) = 0, x_2(k) = 0, ..., x_m(k) = 0)$. This scenario is depicted in Figure 4.

Here, the DMFD problem is equivalent to

$$\widehat{X}^K = \arg\max_{X^K} J(\underline{x}(k), \underline{x}(k-1)) = \arg\max_{X^K} \sum_{k=1}^{K} f_k(\underline{x}(k), \underline{x}(k-1)) \quad (11)$$

where the primal objective function for an individual component state, i.e., f_k is defined as

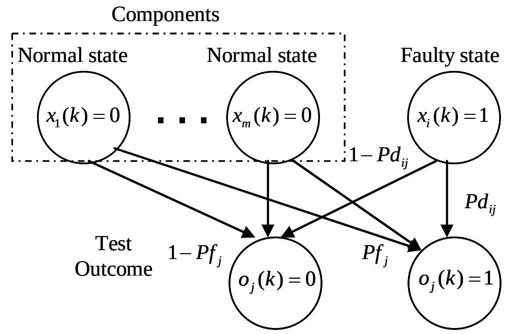

Figure 4: Detection and false alarm probabilities for problem 2

$$f_k(\underline{x}(k), \underline{x}(k-1), \underline{y}(k), z(k)) = z(k)\eta(k) + g(k) + \sum_{i=1}^{m} \sum_{o_j(k) \in O_p(k)} x_i(k)\ln(1-Pd_{ij}) + \sum_{o_j(k) \in O_f(k)} \ln(1\text{-}y_j(k)) + \sum_{i=1}^{m} \tau_i(k)x_i(k) + \sum_{i=1}^{m} \sigma_i(k)x_i(k-1) + \sum_{i=1}^{m} h_i(k)x_i(k)x_i(k-1) \quad (12)$$

where

$$\ln z(k) = \sum_{i=1}^{m} \ln(1 - x_i(k)), \quad (13)$$

$$\ln(y_j(k)) = z(k)\ln(1-Pf_j) + \sum_{i=1}^{m} x_i(k)\ln(1-Pd_{ij}), \quad (14)$$

$\eta(k)$ is a function of Pf_j and $\tau_i(k)$, $\sigma_i(k)$, $h_i(k)$, $g(k)$ are functions of $Pa_i(k)$, $Pv_i(k)$. Appending constraints (13) and (14) via Lagrange multipliers $\mu(k)$, $\{\lambda_j(k)\}_{j \in O_f(k)}$, the Lagrangian function $L(X, Y, z, \Lambda)$ can be obtained. Using the Lagrange multiplier theorem, we optimize the Lagrangian function $L(X, Y, z, \Lambda)$ w.r.t. $y_j(k)$ to obtain optimal $y_j(k)^*$ and optimizing w.r.t. $z(k)$, we obtain optimal $z(k)^*$. The dual function $Q(\Lambda)$ of problem 2 is defined by

$$Q(\Lambda) = \max_{X^K, Y^K, z^K} L(X, Y, z, \Lambda). \quad (15)$$

Substituting $(y_j(k)^*, z(k)^*)$ into $L(X, Y, z, \Lambda)$ and simplifying further by rearranging and combining the terms, we obtain the dual function as

$$Q(\Lambda) = \max_{X^K} \sum_{i=1}^{m} Q_i(\Lambda) \quad (16)$$

where

$$Q_i(\Lambda) = \sum_{k=1}^{K} \xi_i(x_i(k), x_i(k-1), \lambda_j(k), \mu(k)) + \frac{1}{m} w_k(\lambda_j(k), \mu(k)) \quad (17)$$

and

$$\xi_i(x_i(k), x_i(k-1), \lambda_j(k), \mu(k))$$

$$= \left(\sum_{o_j(k)\in O_p(k)} \ln(1-Pd_{ij}) + \tau_i(k) - \sum_{o_j(k)\in O_f(k)} \lambda_j(k)\ln(1-Pd_{ij}) \right) x_i(k)$$

$$+\sigma_i(k)x_i(k-1) + h_i(k)x_i(k)x_i(k-1) - \mu(k)\ln(1-x_i(k)) \quad (18)$$

and

$$w_k(\lambda_j(k), \mu(k)) = \mu(k)\left(\frac{\eta_j(k) + \ln(\mu(k))}{-\eta_j(k) + \sum_{\forall o_j \in O_f(k)} \lambda_j(k)\ln(1-Pf_j)} \right)$$

$$-\mu(k)\left(\sum_{\forall o_j \in O_f(k)} \frac{\lambda_j(k)\ln(1-Pf_j)}{-\eta_j(k) + \sum_{\forall o_j \in O_f(k)} \lambda_j(k)\ln(1-Pf_j)} \right)$$

$$+g(k) + \sum_{\forall o_j \in O_f(k)} \left[\lambda_j(k)\ln\lambda_j(k) - (1+\lambda_j(k))\ln(1+\lambda_j(k)) \right] \quad (19)$$

The dual problem posed in (15)-(19) is separable and it can be solved by following a procedure similar to that used for solving Problem 1. The only difference is that we also need to update the Lagrange multiplier $\mu(k)$ using a subgradient method.

DMFD Problem 3

In this formulation, we consider the case where the probabilities of detection and false alarm (Pd_j, Pf_j) are associated only with each test t_j (see Figure 5). Formally, $Pd_j = \Pr(o_j(k)=1 \mid t_j(k)=1)$ and $Pf_j = \Pr(o_j(k)=1 \mid t_j(k)=0)$. We can convert these probabilities into a special case of Problem Formulation 1 by computing (Pd_{ij}, Pf_{ij}) using (1):

$$Pd_{ij} = (d_{ij})Pd_j + (1-d_{ij})Pf_j \quad (20)$$

Similarly

$$Pf_{ij} = (d_{ij})Pf_j + (1-d_{ij})Pd_j \quad (21)$$

The solution of Problem 3 can be obtained by substituting Pd_{ij} and Pf_{ij} in (20)-(21) in the solution of Problem 1.

DMFD Problem 4

Next, we consider the case when the system consists of reliable tests, and the fault-test relationships are deterministic, i.e. $Pd_{ij}=1$ and $Pf_{ij}=0$ for $i=1,...,m$ and $j=1,...,n$ or equivalently, the D-matrix completely characterizes the fault-test relationships [11]. This formulation can be represented as a bipartite graph between the components and tests. In this case, if some tests have passed, then we can infer that all the failure sources covered by these tests are good components.

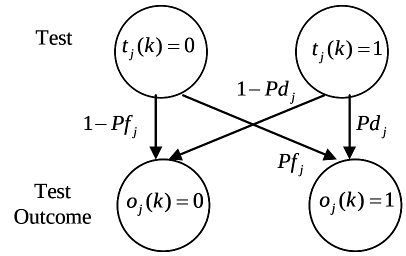

Figure 5: Detection and false alarm probabilities for problem 3

Thus, we need to infer failed components from those covered by the failed tests only, i.e., by excluding those components covered by the passed tests. Consequently, the size of the DMFD problem can be reduced by removing all failure sources $\{s_i \mid Pf_{ik}=0, Pd_{ik}=1, \text{ and } t_k \in T_p(k)\}$. For each failed test $t_j(k) \in T_f(k)$, the optimal solution contains at least one component state $x_i(k)$ = 1 that satisfies $d_{ij}=1$. Thus, there must be one or more failure sources that cover the failed tests. Let us consider a matrix A, which has each row representing the list of failure sources covered by a failed test. After excluding the failure sources covered by the passed tests, the resulting matrix A is a binary matrix such that $a_{ij}=d_{ji}$. After substituting $Pd_{ij}=1$ and $Pf_{ij}=0$ in (5), the reliable test scenario with a binary D-matrix simplifies to a dynamic set covering problem with the following objective function term at epoch *k*:

$$f_k(\underline{x}(k), \underline{x}(k-1)) = \sum_{i=1}^{m} \mu_i(k)x_i(k) + \sum_{i=1}^{m} \sigma_i(k)x_i(k-1)$$

$$+\sum_{i=1}^{m} h_i(k)x_i(k)x_i(k-1) + g(k) \quad (22)$$

subject to following constraints:

$$A(k)\underline{x}(k) \geq \underline{e} \text{ for } t_j(k) \in T_f(k) \quad (23)$$

where $\underline{e}$ is a vector of one's. Appending constraints (23) to (22) via Lagrange multipliers $\{\lambda_j(k)\}_{j\in T_f(k)}$, the Lagrangian function $L(X,\Lambda)$ can be obtained. The dual function $Q(\Lambda)$ is defined by

$$Q(\Lambda) = \max_{X^K} L(X,\Lambda). \quad (24)$$

Simplifying further by rearranging and combining the terms, we obtain the dual function as

$$Q(\Lambda) = \max_{X^K} \sum_{i=1}^{m} Q_i(\Lambda) \quad (25)$$

where

$$Q_i(\Lambda) = \sum_{k=1}^{K} \xi_i(x_i(k), x_i(k-1), \lambda_j(k)) + \frac{1}{m} w_k(\Lambda) \quad (26)$$

$$\xi_i(x_i(k), x_i(k-1), \lambda_j(k)) = -\sum_{t_j \in T_f(k)} \lambda_j(k)a_{ji}(k)x_i(k)$$

$$+\mu_i(k)x_i(k) + \sigma_i(k)x_i(k-1) + h_i(k)x_i(k)x_i(k-1) \quad (27)$$

and

$$w_k(\Lambda) = g(k) + \sum_{\forall t_j \in T_f(k)} \lambda_j(k) \,. \qquad (28)$$

The dual problem defined in (24)-(28) is separable. The Viterbi algorithm is used to solve each subproblem corresponding to each fault state sequence $\underline{x}_i$. This algorithm can be viewed as a dynamic set covering problem, which is NP-hard. Thus, the dynamic set covering problem is solved by combining the Viterbi algorithm and Lagrangian relaxation. This generalizes Beasley's Lagrangian relaxation algorithm for the static set covering problem [11], [15] to dynamic settings.

Results

We implemented and applied the solution of problem 1, the most general version of the DMFD problem formulation, to a few real world models. Table 4 illustrates the model parameters of an automotive system, a document matching system (Docmatch), a power distribution system (Powerdist), a UH-60 helicopter transmission system (Helitrans) and an engine simulator (EngineSim). Details of these models are provided in [11]. Here, *m*, *n*, and *c* denote the number of components (failure sources), number of tests and the average number of intermittent faults that can occur over a span of 100 epochs. The fault appearance probabilities (Pa_i) were computed based on the average number of intermittent faults (*c*). These real-world systems are not ideal because they have fewer tests as compared to failure sources; hence, some failure sources are not covered by any tests. The fault disappearance probabilities (Pv_i) were varied between 0.0025-0.0049 to allow *c* intermittent faults, on average. The probabilities of detection and false alarm were varied as shown in Table 4. The maximum number of subgradient iterations was set at 80. The algorithms were implemented in MATLAB. We used a standard PC having Pentium 4 Processor with 3.0 GHz clock speed and 512 MB RAM.

Table 5 shows the results obtained using the subgradient (S) and the deterministic simulated annealing (DSA) [7] methods. Here, J, Q, D, CI, FI and t denote the primal function value, the dual function value, the approximate duality gap, the correct isolation rate, the false alarm rate and the computation time per epoch. The primal and dual function values are computed using (3)-(5) and (15)-(19), respectively. The approximate duality gap (*D*) is computed as a ratio of the difference between *Q* and *J* divided by the absolute value of the primal feasible value *J*. Here, *CI* is computed as the average percentage of true fault states, which are isolated by the algorithm over a span of *K* epochs, and *FI* is computed as the average percentage of fault states, which are falsely isolated by the algorithm over a span of *K* epochs. The subgradient method (S) achieves higher correct isolation rates as compared to (DSA) for all the systems except Helitrans. However, the DSA method achieves better primal function value and is also effective in reducing the computation time (*t*). Also, note that we can obtain a hybrid duality gap by taking the maximum primal solution from the subgradient (S) and the deterministic simulated annealing (DSA) methods and the dual function value from the subgradient (S) method. The hybrid DSA-subgradient (HS) duality gaps are also shown in Table 6. The CPU time (t) is measured in seconds. Based on our experience, these numbers are highly practical and they can be further reduced by a factor of 10 when implemented in the C language.

We also showed an application of the DMFD Problem 3 formulation in our recent paper [16] where we performed dynamic fusion of classifiers over time for automotive engine fault diagnosis. The temporal correlations considered by dynamic fusion improve classification accuracy over a variety of static fusion techniques (based on batch data).

Table 4: Real world models

	m	n	Pd_{ij} Pf_{ij}	c, Pa_i
Automotive	22	60	(0.85-0.95), 0-0.02	3, 9.13e-04
Docmatch	257	180	(0.6-1),0	9, 3.12e-04
Powerdist	96	98	(0.6-1),0	3,3.13e-04
Helitrans	34	51	(0.6-1),0	2, 2.95e-04
EngineSim	53	30	(0.6-1),0	2,5.68e-04

Table 5: Results on real world models

		J	Q	D (%)	CI	FI	t
Automotive	S	-658	-481	27	99.5	0.05	0.43
	DSA	-775	--	--	75	1.30	0.01
	HS	-658	-481	27			
Docmatch	S	-541	-311	42.5	88.2	0.36	4.53
	DSA	-405	--	--	69	0.70	0.24
	HS	-405	-311	23.2			
Powerdist	S	-232	-125	46.1	91.6	0.75	1.56
	DSA	-157	--	--	84	0.30	0.05
	HS	-157	-125	20.3			
Helitrans	S	-15	-14	6.7	94.8	0.31	0.47
	DSA	-15	--	--	100	0.0	0.02
	HS	-15	-14	6.7			
EngineSim	S	-85	-33	61.1	95.1	2.28	0.64
	DSA	-51	--	--	86	0.3	0.02
	HS	-51	-33	35.3			

Complexity: All the DMFD formulations are NP-hard. The simplest one, i.e. the deterministic formulation (Problem 4) is also NP-hard. If we use brute force dynamic programming (DP), the complexity is $O(K2^{2m})$ where K is total number of epochs and m is number of components (failure sources). The DP is infeasible for systems with more than about 15 states. The algorithm presented here reduces the overall complexity to $O(K(m+O_f))$ where O_f is the set of failed tests, a substantial improvement. In particular, the complexities of binary Viterbi algorithm over all fault states and the subgradient method are $O(Km)$ and $O(KO_f)$, respectively, per iteration.

Conclusions

This paper discussed four formulations of the DMFD problem. Analogous forms of these formulations have been studied widely in fault diagnosis community in a static context, and applied in various fields. Here, we provided a unified formulation of all the MFD formulations in a dynamic context. The first formulation refers to a generalized version of the DMFD problem when the detection and false alarms probabilities are associated with each test and fault. In the second formulation, the false alarm probability is associated with fault-free case only. The solution to the second formulation was shown to be quite similar to that of problem formulation 1, except for the need to update an additional Lagrange multiplier. The third formulation considers the case where the uncertainties are associated with only test outcomes. This models dynamic fusion of classifier outputs. In the fourth formulation, we considered the deterministic case, which led to a novel dynamic set covering problem.

We simulated problem 1 using real world models. Results demonstrate that our algorithm achieves high isolation rate as compared the deterministic simulated annealing method. The latter provides better primal function value as compared to the subgradient method. In our future work, we plan to implement an on-line version of our algorithm using a sliding window method. The sliding window implementation will reduce the number of Lagrange multipliers updates, because estimates for these would have been computed for $(W-1)$ epochs in the preceding window of size W. Initialization based on these estimates will improve convergence. In addition, we will consider multi-state component models with multiple test outcomes.

Acknowledgements

The study reported in this paper was supported, in part, by the Toyota Technical Center and the Office of Naval Research under contract no. 00014-00-1-0101. Any opinions expressed in this paper are solely those of the authors and do not represent those of the sponsor.

References

[1] F. Yu, F. Tu, H. Tu, and K. R. Pattipati, “Multiple disease (fault) diagnosis with applications to the QMR-DT problem,” *Proceedings of Computing Communication and Control Technology International Conference*, Austin, TX, 2004. To appear in *IEEE Trans. on SMC: Part A*, September 2007.

[2] N. Odintsova, I. Rish, S. Ma. “Multifault Diagnosis in Dynamic Systems,” *Proceedings of IM*, 2005.

[3] M. Shakeri, K. R. Pattipati, V. Raghavan and A. Patterson-Hine, “Optimal and near-optimal algorithms for multiple fault diagnosis with unreliable tests,” *IEEE Transactions on Systems, Man and Cybernetics-Part C: Applications and Reviews*, vol. 28, no. 3, August 1998.

[4] T. Le and C. N. Hadjicostis, "Graphical inference methods for fault diagnosis based on information from unreliable sensors," *Proceedings of Intl. Conf. on Control, Automation, Robotics and Vision (ICARCV)*, pp. 1012-1017, Singapore, December 2006.

[5] J. Ying, T. Kirubarajan, and K. R. Pattipati, “A Hidden Markov model based algorithm for fault diagnosis with partial and imperfect tests,” *IEEE Transactions on SMC: Part C - Applications and reviews*, vol.30, no. 4, 2000.

[6] O. Erdinc, C. Raghavendra, and P. Willett, “Real-time diagnosis with sensors of uncertain quality,” *SPIE Conference Proceedings*, Orlando, pp. 1490– 1499, April 2003.

[7] S. Ruan , Y. Zhou , F. Yu , K. R. Pattipati, P. Willett and A. Patterson-Hine, “Dynamic multiple fault diagnosis and imperfect tests,” *IEEE Transactions on Systems, Man and Cybernetics: Part A*, under review, June 2006.

[8] Z. Ghahramani and M. I. Jordan, Factorial hidden Markov models, Machine Learning, *Kluwer Academic Publishers*, Boston, 1997.

[9] S. Singh, S. Ruan, K. Choi, K. Pattipati, P. Willett, S. M. Namburu, S. Chigusa, D. V. Prokhorov and L. Qiao, “An optimization-based method for dynamic multiple fault diagnosis problem,” *IEEE Aerospace Conference*, Big Sky, Montana, March 2007.

[10] J. Pearl, Probabilistic reasoning in intelligent systems: networks of plausible inference, *Morgan Kauffmann Publishers Inc.*, San Mateo, CA, 1988.

[11] F. Tu, K. R. Pattipati, S. Deb, and V. N. Malepati, “Computationally efficient algorithms for multiple fault diagnosis in large graph-based systems,” *IEEE Transactions on Systems, Man and Cybernetics*, vol. 33, no. 1, pp. 73-85, January 2003.

[12] K. R. Pattipati and M. G. Alexandridis, “A heuristic search and information theory approach to sequential fault diagnosis,” *IEEE Trans. on Systems Man and Cybernetics*, vol. 20, no. 4, pp. 872-887, 1990.

[13] V. Raghavan, M. Shakeri and K.R. Pattipati, “Optimal and near-optimal test sequencing algorithms with realistic test models,” *IEEE Transactions on Systems, Man, and Cybernetics: Part A - Systems and Humans*, vol. 29, no. 1, January 1999, pp. 11-27.

[14] M. Shakeri, V. Raghavan, K. R. Pattipati, and A. Patterson-Hine “Sequential testing algorithms for multiple fault diagnosis,” *IEEE Trans. on Systems, Man and Cybernetics: Part A*, vol. 30, pp. 1-14, Jan. 2000.

[15] J. E. Beasley, “An algorithm for set covering problems,” *European Journal of Operational Research*, vol. 31, pp. 85-93, 1987.

[16] S. Singh, K. Choi, A. Kodali, K. Pattipati, S. M. Namburu, S. Chigusa, D. V. Prokhorov, and L. Qiao, “Dynamic fusion of classifiers for fault diagnosis,” *IEEE SMC Conference*, Montreal, Canada, October 2007.

[17] D. Bertsekas, Nonlinear programming, *Athena Scientific*, Belmont, MA, USA, 2nd Edition, 2003.