Dynamic Testing Algorithm Based on Rough Sets for Multiple Fault Diagnosis

Lei Gao
School of Computer Science & Technology
Shandong University
E-mail: aicscw2006@yahoo.cn

Guangzhou Zeng
School of Computer Science & Technology
Shandong University
E-mail: gzzeng@sdu.edu.cn

Abstract

A common simplifying assumption that there exists, at most, a single fault in the system at any given time does not hold for complex systems with large number of components or systems with little opportunity for maintenance during operation. By employing concepts from rough sets theory, information theory and heuristic search approach, a novel dynamic testing algorithm based on rough sets is presented for multiple fault diagnosis, which can efficiently improve the diagnostic precision, substantially reduce the expected testing cost and realize the real-time dynamic diagnosis independent of the priori probability of system.

1. Introduction

Patten recognition is the scientific discipline whose goal is the classification of objects into a number of categories or classes [1]. Depending on the application, these objects can be any type of measurements that need to be classified, such as fault patterns arising in very large-scale integration systems or increased use of redundancy for fault-tolerant systems, which have made automated fault diagnosis an indispensable element of the electronics maintenance process. Fault diagnosis is fundamentally a process of identifying the cause of a malfunction by observing its effects at various monitoring point in a system [2]. A common simplifying assumption made by the most existing testability analysis and diagnostic tools is that there exist, at most, a single fault in the system at any given time [2,3,4]. But, for complex systems with large number of components such as a modern aircraft consisting of mechanical, electro-mechanical and hydraulic subsystems, and for systems with little or no opportunity for repair or maintenance during their operations, this assumption is unrealistic. Thus, constructing optimal or near-optimal testing sequence algorithms for multiple fault isolation for complex real-time systems are required, which is termed multiple fault diagnosis problem (MFD). The computational complexity of MFD problem is super exponential, that is, it is NP-hard[5], and previous literature[6] showed that MFD problem using artificial intelligence techniques is computationally expensive and impractical for large systems.

By employing concepts from information theory and AND/OR graph heuristic search, based on multi-signal flow graphs[7], an extended single fault testing algorithm and sure testing algorithms 1~3[5,6] have been presented for MFD problem, which have been successfully implemented in TEAMS[8], a testability engineering and maintenance system. In extended single fault testing algorithm, a component may be replaced or repaired before confirming that it is indeed fault. Consequently, the probability of false alarm error or RTOK (removed but found OK upon retesting) is higher. In order to overcome above problem, sure strategies[5] employ all informative tests prior to diagnosis, and find one or more definitely failed components, while not making an error when other co-existing faults are present. Considering the complex and real-time systems, the following shortcomings of above strategies for MFD still exit: ① Only the minimum expected testing cost (time) has been considered for constructing heuristic evaluation functions (HEFs), while not considering the isolation precision of multiple fault ambiguity group (MFAG)[6] which involves probable failure states; ② MFAG isolation process involves many times of rebuilding of dependency matrix (DM) and sub-decision trees (DTs), which generates a lot of additional cost; ③There are replacement or repair operations during MFD process, which can't meet the real-time environment.

Above shortcomings restrict the effect of sure testing algorithms in real-time systems with little opportunity for repair. In this paper, Based on

978-0-7695-3305-6/08 $25.00 © 2008 IEEE
DOI 10.1109/FSKD.2008.176

dependency matrix, we present a novel dynamic testing algorithm for multi-fault isolation, a novel feature of which is the integration of rough sets, compact sets, information theory and heuristic search techniques to enhance the isolation precision of MFD and overcome the computational explosion of the optimal test sequencing problem. Computational results indicate that this algorithm efficiently improves the isolation precision of multi-fault and reduces expected testing cost comparing with sure testing algorithms.

2. Problem Formulation

The multi-fault test sequencing problem involves applying binary outcomes to determine the failure states of a system. Each test provides information about a subset of the set of failure states, and tests are applied sequentially until the failure states are identified. There are following assumptions about this problem: ① failure states are independent; ② if a failure state happens, it affects all of the signals associated with it, and if a signal is out of normal range, at least one of associated failure states must happen; ③ a test only test one signal; ④ all of tests are reliable, i.e., 100%sensitive and 100%specific; ⑤ the failure signature of a multi-fault is the union of single failure signatures.

2.1. MFD Formulation

Formally, multi-fault test sequencing problem can be defined by the eight-tuple $(S,P,T,C,D,\widetilde{S},\widetilde{T},G)$, where $S=\{s_0,s_1,\cdots,s_m\}$ is a set of statistically independent failure states associated with the system, s_0 is the normal state; $P=\{p_0,p_1,\cdots,p_m\}$ is the priori probability associated with S; $T=\{t_1,t_2,\cdots,t_n\}$ is a finite set of n reliable binary outcome tests, where each test t_j checks a subset of S; $C=\{c_1,c_2,\cdots,c_n\}$ is the set of test costs measured in terms of time, manpower requirements, or other economic factors (in this paper, cost refer to time); $D=[d_{ij}]$ is a binary matrix of dimension $m\times n$ which represents the relationship between S and T; where $d_{ij}=1$ if test t_j monitor failure state s_i, conversely, s_i affects t_j, otherwise $d_{ij}=0$; $\widetilde{S}=\{\widetilde{S}_i\},(0\le i\le m)$ is a finite collection of sets, where $\widetilde{S}_0=\phi$, $\widetilde{S}_i=\{t_j \mid d_{ij}=1,1\le j\le n\}$ denotes the signature of failure state s_i, it indicates all the tests that monitor failure state s_i; $\tilde{T}=\{\tilde{T}_j\},(0\le j\le n)$ is a finite collection of sets, where $\widetilde{T}_j=\{s_i \mid d_{ij}=1,1\le i\le m\}$ denotes the signature of test t_j, it indicates all the failure states detectable by test t_j; $G\subseteq S$ is the set of fault-free states.

The primary target of multi-fault test sequencing algorithm is to unambiguously isolate MFAG with maximum probability using the outcomes of tests. The corresponding objective function is given by:

$$\arg\max_{S_I\subseteq S}\Pr ob(S_I \mid T_p,T_f) \tag{1}$$

Where $S_I\subseteq S$ is the subset of failure states, T_p,T_f respectively correspond to the set of pass tests and false tests.

Based on the primary objective function, the secondary target is to minimize the expected testing cost, and the corresponding objective function is given by:

$$J=\sum_{S_I\subseteq S} p(S_I)E(C\mid S_I)=\sum_{S_I\subseteq S}\sum_{t_j\in R_I} p(S_I)c_j \tag{2}$$

Where $E(C\mid S_I)$ is total test cost for isolating S_I, R_I is set of applied tests in the path leading to the isolation of S_I, $p(S_I)$ is the conditional probability of S_I based on the multi-fault assumption, given by:

$$p(S_I)=\begin{cases}\prod_{k=1}^{m}(1-p(s_k)) & s_0\in S\wedge S_I=\{s_0\}\\ \prod_{k=1}^{m}(1-p(s_k))\prod_{s_i\in S_I}\dfrac{p(s_i)}{1-p(s_i)} & s_0\in S\wedge s_0\notin S_I\\ \left(\prod_{k=1}^{m}(1-p(s_k))\prod_{s_i\in S_I}\dfrac{p(s_i)}{1-p(s_i)}\right)\Big/\left(1-\prod_{k=1}^{m}(1-p(s_k))\right) & s_0\notin S\end{cases} \tag{3}$$

Where $p(s_i)$ is the priori probability of state s_i.

2.2. MFD Description Using Rough Sets

Rough sets theory (RST) provides a formal and robust way of manipulating the roughness of the knowledge in information systems[9]. Based on the analogy between rough sets and MFAG, MFD process can be described as follows:

Let $S=\{s_0,s_1,\cdots,s_m\}$ be the set of initial failure states and R : select the independent failure states from S is a equivalence relation defined on S, then for $\forall s_i\in S,1\le i\le m$, $\{s_i\}\in S/R,1\le i\le m$ is a basic equivalence class defined on S associated with R, termed $[s_i]_R=\{s_i\},1\le i\le m$, where S/R is a partition on S associated with R.

Suppose S_x is the target set that involves multiple failure states, T_x is the set of previously applied tests , which can be regard as the domain knowledge about S, then S_x can be accurately described by basic equivalence classes $[s_i]_R, 1 \le i \le m$ on S, that is, at the end of MFD process, S_x should be the R accurate set defined on S. But during diagnosis process, MFAG can't be accurately described by $[s_i]_R$ because failure states happen randomly and the information T_x can presents is narrow, therefore it should be regarded as R rough set defined on S. Then MFD is a dynamic process during which S_x is continuously approximated by union set of $[s_i]_R$ using the information presented by T_x. More accurate the approximation is, better the diagnosis effect is. Fig.1 illustrates this process.

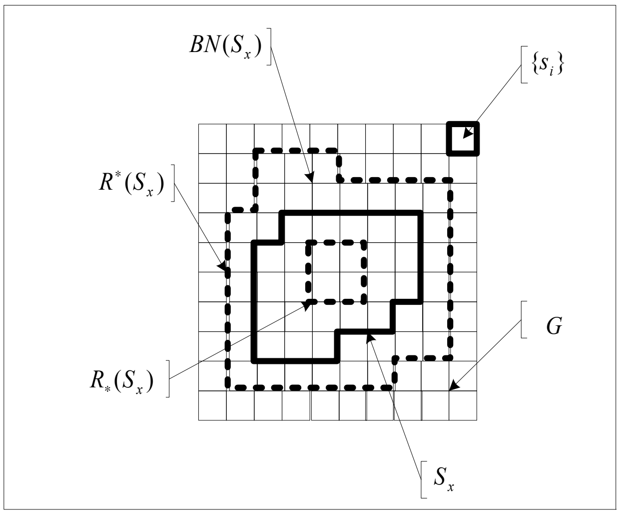

Fig.1. MFD process described by rough sets approximation

Where $R_*(S_x)$ is the lower approximation set of S_x, which denotes the set of failure states definitely belonging to S_x according to T_x; $R^*(S_x)$ is the upper approximation set of S_x, which is defined as the complementary set of fault-free set G, given by:

$$R^*(S_x)=G^c=(\bigcup_{t_i \in T_p} \tilde{T}_i)^c \quad (4)$$

G is the negative zone of S_x denoting the set of failure states definitely not belonging to S_x according to T_x; $BN(S_x)=R^*(S_x)\text{-}R_*(S_x)$ is the boundary zone of S_x. During diagnosis, $R_*(S_x)$ continuously enlarges while $R^*(S_x)$ continuously reduces until they coincide with each other, that is $BN(S_x)=0$.

2.3. MFD Description Using Compact sets

Above MFD process based on rough sets corresponds to a AND/OR decision tree that specifies the test to perform next depending on the outcomes of previously applied tests. The root node is initial set of failure states S, leaf node is target set S_x, AND node denotes the available test t_j (with cost c_j), OR node denotes MFAG expressed by compact notation $A=\Theta(L;F_1,F_2,\cdots,F_L;G)$ [10], where $F_i \subseteq G^c, i=1,2,......,L$ are sets that are known to contain at least one definitely failed failure state each, i.e., $\Theta(L;F_1,F_2,\cdots,F_L;G)=\{X \subseteq S \mid X \cap F_i \ne \phi, 1 \le i \le L, X \cap G = \phi\}$.

Some of the properties of compact sets are as follows[10]:

(1) After performing a test t_j, $A=\Theta(L;F_1,F_2,\cdots,F_L;G)$ is decomposed as follows:

$$A \leftarrow \begin{cases} \Theta(L;(F_1 \cap \tilde{T}_j^c),\cdots,(F_L \cap \tilde{T}_j^c);(G \cup \tilde{T}_j)) & \text{if } t_j \text{ passes} \\ \Theta(L+1;F_1,\cdots,F_L,\tilde{T}_j \cap G^c;G) & \text{if } t_j \text{ fails} \end{cases} \quad (5)$$

(2) $A=\Theta(L;F_1,\cdots,F_L;G)=\Theta(L;F_1 \cap G^c,\cdots,F_L \cap G^c;G)$ (6)

(3) If $\exists i \in [1,m]$ making $F_i \subseteq Y$, then

$$\Theta(L;F_1,\cdots,F_L,Y;G)=\Theta(L;F_1,\cdots,F_L;G) \quad (7)$$

(4) If $|F_i|=1$, then $s_i \in F_i$ is definitely faulty.

3. Dynamic Testing Algorithm Based on Rough Sets

Since the generation of an optimal testing sequence is an NP-complete problem, it is necessary to employ heuristic approaches which use domain knowledge T_x, in the form of HEF[2], to avoid enumerating entire set of potential solutions for guiding the AND/OR decision tree search. comparing with sure strategies[5,6,10], improvements made by dynamic testing algorithm based on rough sets are as follows:

(1) Two-level HEFS are derived to meet two targets of MFD. The primary HEF based on rough sets theory extends decision tree leading to the isolation of definite failure states to realize the primary objective function (formula (1)). The secondary HEF based on information theory extends decision tree leading to the minimum expected test cost to realize the secondary objective function (formula (2)). The primary HEF has the first priority.

(2) Abolish repair and replacement during diagnosis and enhance the ending conditions of algorithm, i.e., if $\exists i \in [1,m]$ making $|F_i|=1$ and $F_i \cup G \ne S$, instead of ending algorithm, continuously extending nodes until $|F_i|=1$ and $F_i \cup G=S$ or without available tests.

3.1. Heuristic Evaluation Functions (HEFs)

3.1.1. Primary HEF

According to the analysis in section 2.2, the probability of accurate isolation for MFAG depends on the approximation accuracy of rough sets [11]. Therefore the primary HEF can be constructed by employing approximation accuracy termed η_R, given by:

$$h_1(A) = HEF_1 = \eta_R(S_x) = \frac{|R_*(S_x)|}{|R^*(S_x)|} \tag{8}$$

Where $A=\Theta(L;F_1,F_2,\cdots,F_L;G)$ is the compact sets on OR node; $|R^*(S_x)|$ can be derived from formula (4):

$$|R^*(S_x)| = |G^c| \tag{9}$$

$|R_*(S_x)|$ can be explored as follows:

according to definition of compact sets and property (4)

∵ in $F_i, 1 \le i \le L$, there is at least one definitely failed failure state,

∴ if $\exists j \in [1,L]$, making $F_j = \{s_k\}, k \in [0,m]$, then state s_k must be failed, i.e., $s_k \in R_*(S_x)$,

∴update $|R_*(S_x)|$ as follows: $|R_*(S_x)| = |R_*(S_x)| + 1$.

∵ if $|F_i|>1, i \in [1,L]$, then there must be a definitely failed state belonging to F_i, but the exact one and the number of definitely failed states is ambiguous, and larger $|F_i|$ is, more ambiguous it is. For evaluating the fuzziness of F_i, fuzzy factor is defined as

$$\xi(F_i) = \frac{1}{|F_i|}, i \in [1, L] \tag{10}$$

which denotes that the probability for probable failed states in F_i belonging to $|R_*(S_x)|$ is inversely proportional to $|F_i|$;

∴if $|F_i|>1, i \in [1,L]$, update $|R_*(S_x)|$ as follows:

$$|R_*(S_x)| = |R_*(S_x)| + \xi(F_i)$$

To sum up, for $A=\Theta(L;F_1,F_2,\cdots,F_L;G)$, $|R_*(S_x)|$ is given by:

$$|R_*(S_x)| = \sum_{F_i \in \tilde{F}_n} \xi(F_i) + |\tilde{F}_1| \tag{11}$$

where $\tilde{F}_n$ is the finite collection of sets $F_i, 1 \le i \le L$ the cardinality of which is more than 1; $\tilde{F}_1$ is the finite collection of sets $F_j, 1 \le j \le L$ the cardinality of which is equal to 1.

Since the primary HEF depends on domain knowledge instead of the priori probability of failure states, the so-called Maxmax (maximizing the maximum approximation accuracy $\eta_R(S_x)$) criterion is considered to construct the primary assessment function (AF).

Suppose the compact set on OR node is $A=\Theta(L;F_1,F_2,\cdots,F_L;G)$, according to properties 1~3, after performing a test t_j, the compact sets corresponding to pass subset A_{jp} and fail subset A_{jf} which are the child nodes of A are given by:

$$A_{jp} = \Theta(L; F_1 \cap \widetilde{T}_j^c, F_2 \cap \widetilde{T}_j^c, \cdots, F_L \cap \widetilde{T}_j^c; G \cup \widetilde{T}_j) \tag{12}$$

$$A_{jf} = \Theta(L+1; F_1, F_2, \cdots, F_L, \widetilde{T}_j \cap G^c; G) \tag{13}$$

Therefore, the primary AF about A with $t_j, j \in [1,n]$ is given by:

$$H_{1j}(A) = \max\{h_1(A_{jp}), h_1(A_{jf})\} \tag{14}$$

And the expected test t_k which will be extended next step is given by:

$$k = \arg\max_j \{H_{1j}(A)\} \tag{15}$$

3.1.2. Secondary HEF

If the primary AFs of two different tests are equal to each other, the secondary AF which reduces test cost based on the primary one will be used to select the next expected test. Considering the tradeoff between running cost and time, the information quantity [9] associated with test is employed as the secondary HEF.

Let $X = \tilde{T}_j \cap G^c, j \in [1,n]$ denote the set of failure states test t_j monitors, then the probability containing at least one definitely failed state in X with test t_j is $\Pr(X) = 1 - \prod_{s_i \in X}(1 - p(s_i))$, where $p(s_i)$ is the conditional probability of state s_i. Therefore, the secondary HEF can be constructed by:

$$h_2(A) = HEF_2 = -\log_2(\Pr(X)) \tag{16}$$

Corresponding, the secondary AF which denotes the information quantity with unit cost of test is given by:

$$H_{2j}(A) = -\frac{\log_2(\Pr(X))}{c_j}, \quad j \in [1,n] \tag{17}$$

where c_j is the unit cost of test t_j.

And the expected test t_k which will be extended next step is given by:

$$k = \arg\max_j \{H_{2j}(A)\} \tag{18}$$

3.2. Heuristic Search Algorithm

During diagnosis, the set of failure states S can be divided into three classes: ① the sure set of failure states termed SU denotes the definitely failed states; ② the probable set of failure states termed PR denotes the indefinitely failed states; ③ the free set of failure states termed SA denotes the definite fault-free states. Four different combinations of above three classes are as follows: ① $SA{=}S, SU = PR = \phi$ ② $SA \cup SU \cup PR = S$ ③ $SA \cup SU = S, PR = \phi$ ④ $SA \cup PR = S, SU = \phi$. The leaf nodes of decision tree of sure strategies involve the combination ①②④, while the dynamic algorithm's involve the combination ①③④, which can provide better isolation accuracy. The detailed description of rough sets based heuristic search algorithm is as follows:

Step 1: Initially let S be the set of failure states, T be the set of available tests, $A{=}\Theta(1;F_1{=}S;G{=}\phi)$ be the compact set associated with S;

Step 2: Extend OR node $A{=}\Theta(L;F_1,F_2,\cdots,F_L;G)$

Step 2.1: Perform test $t_j \in T$ on extended node A, dividing it into A_{jp} and A_{jf} (according to formula (12),(13), (6) and (7));

Step 2.2: According to formula (8)~(11), respectively derive $h_1(A_{jp})$ and $h_1(A_{jf})$, then according to formula (14), derive the primary AF about A with t_j $H_{1j}(A)$;

Step 2.3: For each test belonging to T, repeat step 2.1~2.2, and according to formula(15), select the expected test t_k;

Step 2.4: If t_k isn't unique, use the secondary AF to evaluate t_k s and update t_k according to formula (16)~(18);

Step 2.5: Extend test t_k, divide A into A_{kp} and A_{kf} and update the set of available tests $T_k = T \setminus t_k$

Step 2.5.1: Repeat step 2.1~2.4, extend OR node A_{kp};

Step 2.5.2: Repeat step 2.1~2.4, extend OR node A_{kf};

Step 3: Ending condition: suppose the compact set on leaf node is $A' = \Theta(L';F_1',F_2',\cdots,F_L';G')$, then

① if $\forall i \in [1,m]$, making $|F_i'|{=}1$ and $(\bigcup_i F_i') \cup G' = S \vee (S \setminus s_0)$;

② there are not available tests, where unavailable test is given by:

$$j = \arg((\exists i \in [1,L] \rightarrow F_i' \cap \widetilde{T}_j^c = \phi) \vee (\widetilde{T}_j \cap G'^c = \phi)) \quad (19)$$

4. Algorithm Analysis

4.1. Case Study

We will rebuild the multi-fault decision tree of a typical verification system[5] using sequential testing algorithm based on rough sets. Dependency matrix is shown in table 1.

Table 1. the dependency matrix of typical system

test	t1	t2	t3	t4	t5	priori	conditional
cost	1	1	1	1	1	probability	probability
S0	0	0	0	0	0	—	0.700
S1	0	1	0	0	1	0.014	0.010
S2	0	0	1	1	0	0.027	0.020
S3	1	0	0	1	1	0.125	0.100
S4	1	1	0	0	0	0.068	0.050
S5	1	1	1	1	0	0.146	0.120

(1) According to dependency matrix, $S{=}\{s_0,s_1,s_2,s_3,s_4,s_5\}, T{=}\{t_1,t_2,t_3,t_4,t_5\}, A_1 = A_0 = \Theta(1;F_1 = S;\phi)$;

(2) Compute the primary AF about S with T, t_1:

$\widetilde{T}_1 = \{s_3, s_4, s_5\}$, $A_{01p}{=}\Theta(1;\{s_0,s_1,s_2\};\{s_3,s_4,s_5\})$, $A_{01f} = \Theta(1;\{s_3,s_4,s_5\};\phi)$;

According to formula (8)~(11), $h_1(A_{01p}){=}(1/3)/3{=}1/9$, $h_1(A_{01f}){=}(1/3)/5{=}1/15$;

According to formula(14), the primary AF about A_0 with test t_1 is given by: $H_{11}(A_0){=}\max\{1/9,1/15\}{=}1/9$;

Corresponding to above analysis, results with other tests are as follows:

$t_2 : \widetilde{T}_2 = \{s_1,s_4,s_5\}$, $A_{02p} = \Theta(1;\{s_0,s_2,s_3\};\{s_1,s_4,s_5\})$, $A_{02f} = \Theta(1;\{s_1,s_4,s_5\};\phi)$, $H_{12}(A_0){=}1/9$;

$t_3 : \widetilde{T}_3 = \{s_2,s_5\}$, $A_{03p} = \Theta(1;\{s_0,s_1,s_3,s_4\};\{s_2,s_5\})$, $A_{03f} = \Theta(1;\{s_2,s_5\};\phi)$, $H_{13}(A_0){=}1/10$;

$t_4 : \widetilde{T}_4 = \{s_2,s_3,s_5\}$, $A_{04p} = \Theta(1;\{s_0,s_1,s_4\};\{s_2,s_3,s_5\})$, $A_{04f} = \Theta(1;\{s_2,s_3,s_5\};\phi)$, $H_{14}(A_0){=}1/9$;

$t_5 : \widetilde{T}_3 = \{s_1,s_3\}$, $A_{05p} = \Theta(1;\{s_0,s_2,s_4,s_5\};\{s_1,s_3\})$, $A_{05f} = \Theta(1;\{s_1,s_3\};\phi)$, $H_{15}(A_0){=}1/10$;

(3) According to formula(15), $t_k{=}\{t_1,t_2,t_4\}$; the secondary AF needs to be computed, and results are as follows(formula (16)~(17)):

$H_{21}(A_0){=}1.7199$, $H_{22}(A_0){=}2.2161$, $H_{24}(A_0){=}1.8734$;

According to formula(18), we have $t_k{=}t_2$, and t_2 is the expected extended test next;

(4) Repeat above steps until the leaf nodes, and multi-fault decision tree extended is shown in Fig. 2. Where compact sets on nodes are as follows in extended order:

$A_1 = \Theta(1;\{s_0,s_1,s_2,s_3,s_4,s_5\};\phi)$ $A_2 = \Theta(1;\{s_0,s_2,s_3\};\{s_1,s_4,s_5\})$

$A_3 = \Theta(1;\{s_1,s_4,s_5\};\phi)$ $A_4 = \Theta(1;\{s_0\};\{s_1,s_2,s_3,s_4,s_5\})$

$A_5 = \Theta(1;\{s_2,s_3\};\{s_1,s_4,s_5\})$ $A_8 = \Theta(1;\{s_2\};\{s_1,s_3,s_4,s_5\})$

$A_9 = \Theta(1;\{s_3\};\{s_1,s_4,s_5\})$ $A_{14} = \Theta(1;\{s_3\};\{s_1,s_2,s_4,s_5\})$

$A_{15} = \Theta(2;\{s_2\},\{s_3\};\{s_1,s_4,s_5\})$

$A_6 = \Theta(1;\{s_1\};\{s_3,s_4,s_5\})$

$A_7 = \Theta(2;\{s_1,s_4,s_5\},\{s_3,s_4,s_5\};\phi)$ $A_{10} = \Theta(1;\{s_1\};\{s_2,s_3,s_4,s_5\})$

$A_{11} = \Theta(2;\{s_1\},\{s_2\};\{s_3,s_4,s_5\})$ $A_{12} = \Theta(2;\{s_1,s_4\},\{s_4\};\{s_2,s_3,s_5\})$

$A_{13} = \Theta(3;\{s_1,s_4,s_5\},\{s_3,s_4,s_5\},\{s_2,s_3,s_5\};\phi)$ $A_{16} = \Theta(1;\{s_4\};\{s_1,s_2,s_3,s_5\})$

$A_{17} = \Theta(2;\{s_1\},\{s_4\};\{s_2,s_3,s_5\})$ $A_{18} = \Theta(2;\{s_1,s_4\},\{s_3\};\{s_2,s_5\})$

$A_{19} = \Theta(3;\{s_1,s_4,s_5\},\{s_3,s_4,s_5\},\{s_2,s_5\};\phi)$ $A_{20} = \Theta(2;\{s_4,s_5\},\{s_2,s_5\};\{s_1,s_3\})$

$A_{21} = \Theta(4;\{s_1,s_4,s_5\},\{s_3,s_4,s_5\},\{s_2,s_5\},\{s_1,s_2\};\phi)$

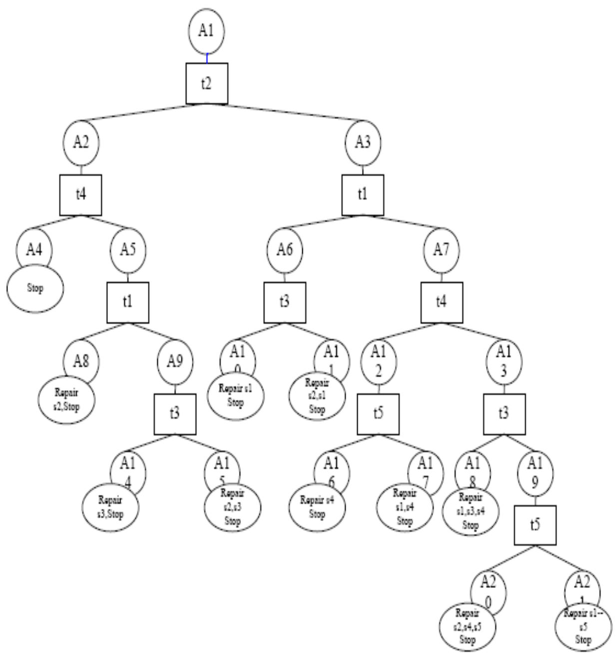

Fig.2. multi-fault decision tree extended by dynamic algorithm based on rough sets

4.2. Case Analysis

Considering MFD problem, there are following evaluation standards:

(1) Isolation precision. This value is defined by the expected proportion of the definitely failed states to the probable failed states on the leaf nodes involving multiple failure states, i.e. ,

$$\psi = \frac{\sum_{A_i \in \tilde{A}_L} \rho_i}{|\tilde{A}_L|} \tag{20}$$

Where ψ is the isolation precision, $\tilde{A}_L$ denotes the set of leaf nodes involving multiple probable failure states, ρ_i is given by:

$$\rho_i = \frac{SN(A_i)}{|G_i^c|}, i \in [1, |\tilde{A}_L|] \tag{21}$$

Where $SN(A_i)$ denotes the number of definitely failed failure states involved in the i th leaf node which belongs to $\tilde{A}_L$.

(2) Expected test cost. This value is given by formula (2)~(3).

(3) The number of extended nodes which involves AND nodes and OR nodes extended by algorithm.

The multi-fault decision tree extended by extended single fault testing algorithm and sure testing algorithms 1~3 for the system shown in table 1 has been illustrated in literature [5,6,10]. According to above evaluation standards, we compare the dynamic algorithm with above algorithms, and display the results in table 2.

Table 2. parameter for various algorithms

	extend single	sure1	sure2	sure3	Dynamic testing
ψ	0	7/18	7/18	1/3	5/9
test cost	2.780	2.715	2.616	2.535	2.126
nodes number	27	25	25	37	31

Note that the dynamic testing algorithm effectively improves multi-fault isolation precision from 30% to more than 50% because of the introduction of the primary HEF; and the expected test cost of dynamic algorithm is superior to other algorithms for the system shown in table 1. But the number of extended nodes of dynamic algorithm lies between sure stratege3 and other algorithms because the dynamic algorithm can isolate more definitely failure states which are ambiguous in sure strategies 1~3 and avoid the rebuilding of decision tree in sure strategy 3.

5. Conclusion and Future Works

To sum up, the dynamic testing algorithm based on rough sets for multiple fault diagnosis have following improvements comparing with sure strategies 1~3 and extended single fault testing algorithm: ① efficiently improve the isolation precision of multi-fault; ② substantially reduce expected testing cost ;③ keep clear of repair or maintenance and the rebuilding of dependency matrix and decision trees during diagnosis process; ④ the primary HEF of this algorithm is independent of the priori probability of failure states.

In the future, following problems about multi-fault diagnosis should be further explored: ① unreliable tests without 100%sensitive and 100%specific; ② dependency condition with p probability, that is, the dependency matrix is not binary against assumption ② (section 2); ③ testing sequence with multi-mode, that is, there are multiple dependency matrix corresponding to the same system; ④ precedence constrains for tests; ⑤ multi-outcomes tests problem.

In addition, the dynamic algorithm based on rough sets can be extended to other areas with relation to the classification of objects, which is characterized by large scale and real-time such as the discovery and matching of semantic web services [12].

References

[1] Shifei Ding, Zhongzhi Shi, "A New Pattern Recognition Algorithm Based on Fuzzy Association Degree", International Conference on Neural Networks and Brain 2005, 2005, 1:54 – 57

[2] Pattipati K.R., Alexandridis, M.G., "Application of Heuristic Search and Information Theory to Sequential Fault Diagnosis" *Journal*, IEEE Transaction on Systems,Man,and Cybernetics, 1990, 20(4): 872-887

[3] Fang Tu, Pattipati, K.R., "Rollout strategies for sequential fault diagnosis", *Journal*, Systems, Man and Cybernetics, Part A, IEEE Transactions on , 2003, 33(1):86 – 99

[4] Raghavan, V.,Shakeri, M.,Pattipati, K.R., "Test sequencing problems arising in test planning and design for testability", *Journal*, Systems, Man and Cybernetics, Part A, IEEE Transactions on , 1999, 29(2):153 – 163

[5] Shakeri, M.,etc,, "Sequential testing algorithms for multiple fault diagnosis", *Journal*, Systems, Man and Cybernetics, Part A, IEEE Transactions on , 2000, 30(1):1 – 14

[6] Shakeri, M., Pattipati, K.R.,etc, "Sequential test strategies for multiple fault isolation", AUTOTESTCON '95. 'Systems Readiness: Test Technology for the 21st Century', 1995:512 – 527

[7] Somnath, Deb,etc, "Multi-signal Flow Graphs: A novel Approach for System Testability Analysisand Fault Diagnosis", in Proc.IEEE AUTOTESTCON, Anaheim, 1994: 361-373

[8] K.R., Pattipati, "TEAMS:Testability Engineering and Maintenance System", Proceedings of the Arnerican Control Conference, 1994

[9] Thomas M.Cover, Joy A.Thomas, *Elements of Information Theory 2 edition*, Wiley-Interscience , 2008

[10] Shakeri, M.Pattipati, K.R., "Near-optimal sequential testing algorithms for multiple fault isolation", IEEE International Conference on Systems, Man, and Cybernetics , 1994 2:1908 - 1913

[11] Gwanggil Jeon, Donghyung Kim, Jechang Jeong, "Rough Sets Attributes Reduction Based Expert System in Interlaced Video Sequences", *Journal*, IEEE Transactions on Consumer Electronics, 2006, 52(4): 1348-1355

[12] McIlraith. S., Son, T.C., Zeng.H., "Semantic Web Services" , IEEE Intelligent Systems. Special Issue on the Semantic Web, 2001, 16(2):46--53