Bayes Analysis for Fault Location in Distributed Systems

Yu Lo Cyrus Chang, Member IEEE
The University of Tennessee, Chattanooga
Leslie C. Lander, Associate Member IEEE
State University of New York, Binghamton
Horng-Shing Lu
National Chiao Tung University, Hsinchu
Martin T. Wells
Cornell University, Ithaca

Keywords — **Bayes analysis, Distance measure, Fault location, Loss function, Comparison test, Probabilistic comparison model, System diagnosis.**

Reader Aids —
General purpose: Advance state of the art
Special math needed for explanations: Bayes statistics
Special math needed to use the results: Same
Results useful to: Reliability & test analysts

Summary & Conclusions — **We propose a simple & practical probabilistic model, using multiple incomplete test concepts, for fault location in distributed systems using a Bayes analysis procedure. Since it is easier to compare test results among processing units, our model is comparison-based. This approach is realistic & complete in the sense that it does not assume conditions such as permanently faulty units, complete tests, and perfect or non-malicious environments. It can handle, without any overhead, fault-free systems so that the test procedure can be used to monitor a functioning system. Given a system *S* with a specific test graph, the corresponding conditional distribution between the comparison test results (syndrome) and the fault patterns of *S* can be generated. To avoid the complex global Bayes estimation process, we develop a simple bitwise Bayes algorithm for fault location of *S*, which locates system failures with *linear* complexity, making it suitable for hard real-time systems. Hence, our approach is appealing both from the practical & theoretical points of view.**

1. INTRODUCTION

This paper studies fault location using Bayes inference methods based on a simple probabilistic comparison model. The distributed systems under consideration consist of a collection of units (subsystems) connected through a network that distributes data and, possibly, processes throughout the system. Our approach provides a generalized solution to the randomized fault diagnosis problem. Previous research has highlighted & studied individual obstacles in the fault-diagnosis process. Those obstacles are:

- System diagnosis results might only be valid if the fraction of faulty units has a restrictive upper bound [36].
- System can have non-permanent faults [26].
- Faulty units can behave maliciously and lie about their results [17].
- Tests might be incomplete [31,32].
- For some test strategies, faulty units have to be assumed to be still able to execute assigned tests [30].
- There might be noisy environments or errors in transmitting/receiving devices [5].
- The probability of failure can vary as the run-time advances [9]. ◄

Probabilistic methods can cope with this list of effects; and important advances have been made in the past few years [1,3-13,15,20-22,37]. Indeed, the listed randomizing effects might indicate that this approach is more realistic than a deterministic one. However, a general probabilistic approach in earlier papers involved a drastic increase in computational complexity. The idea of using comparison testing appeared in [14, 25], and was combined with the probabilistic approach in [15, 16, 21]. Comparison-based testing is used because it is less intrusive [15] than having the system units devote processing time to testing & evaluating each other. We have shown [8,10] that *linear* complexity can be achieved both by: 1) our bitwise Bayes (BWB[1]) algorithm based on the decision theoretic approach, and 2) a heuristic algorithm when performing classical point estimation. The BWB fault-location algorithm handles any number of faults in the same way; therefore it can diagnose a fault-free system as well as a system with many faulty units. The BWB algorithm accounts for the probability of unit failure [24] and incorporates the change of that probability as operation time increases. Explicit inclusion of the probability of failure of a unit is also used in [3, 4, 22], although they assume the probability is constant. Most of the other fault-location research developed for multiprocessor systems more or less resembles the concepts developed by the Preparata, Metze, Chien (PMC) model [30]. The main reason for this similarity is that the same set of simplifying assumptions tends to be repeated, the limitations of which are in [15]. Other approaches can be found in [3-5,17,18,22,23,30,33,35,36], including work for general multiprocessor systems.

The model for testing a system with n units involves distributing a set of tests to the units, observing the results of the tests, and running a diagnosis algorithm to locate faulty units. The BWB algorithm is $O(n)$, making it interesting for hard real-time applications. The simplicity of the BWB depends on decomposing the system (global) Bayes estimation into a bit-wise Bayes estimation by introducing a loss function. None of the data gathered from the tests has to be discarded. The chosen loss function is an admissible[2] Bayes decision rule [2] for fault location, which gives theoretical support to the approach. Some simulation results are discussed in section 5.

[1]The BWB algorithm is called the B-algorithm in [6, 7, 11-13].
[2]*Admissible* implies that there is no decision rule with smaller risk function [2: p 10].

Acronyms[3]

BWB	bit-wise Bayes (algorithm)
HPD	highest posterior density
LT	likelihood table
UUT	unit under test.

Notation

S	system name
n	number of units in S
$G(U,E)$	an undirected test graph with vertex set U and edge set E
U	$\{\mathbf{u}_k\text{: } k = 0, 1, 2, \ldots, n-1\}$: vertex set of n units of S
E	$\{(\mathbf{u}_i, \mathbf{u}_j)\text{: } \mathbf{u}_i, \mathbf{u}_j \in U\}$: edge set of m comparison assignments of the UUT from S
ϕ_k	$\Im(\mathbf{u}_k$ is faulty)
Φ	$\phi_{n-1}\,\phi_{n-2}\,\cdots\,\phi_0$: (system) *fault pattern* of the n units
$\phi_{j,k}$	$\Phi_j = \phi_{j,n-1}\,\phi_{j,n-2}\ldots\phi_{j,0}$: fault status of $\mathbf{u}_k$ when the fault pattern of S is Φ_j
Θ	$\{\Phi_0, \Phi_1, \ldots, \Phi_{2^n-1}\}$: set of all possible fault patterns of S. The Φ_j are enumerated so that $\phi_{j,n-1}\,\phi_{j,n-2}\,\cdots\,\phi_{j,0}$ is the binary representation of j, possibly with leading 0's, *eg*, $\Phi_6 = 0110$ in a 4-unit system
t_i	individual test task i that can be applied to S
T	$\{t_1, t_2, \ldots, t_p\}$: test[4] consisting of p tasks t_i
τ	number of tests in a sequence
$\mathbf{T}$	$\{T^{(k)}\text{: } k = 1, 2, \ldots, \tau\}$: sequence of τ tests, each $T^{(k)}$ represents a test T
c_l	$\Im$(the results of applying T to $\mathbf{u}_i$ & $\mathbf{u}_j$ disagree|link l connects $\mathbf{u}_i$ & $\mathbf{u}_j$): $l = 0, 1, 2, \ldots m-1$; *eg*, figure 1 shows a complete graph of $n = 4$ units, and $m = \binom{n}{2} = \binom{4}{2}$ links
C	$c_{m-1}\,c_{m-2}\,\ldots\,c_0$: global (system) comparison pattern of the m links
$C^{(k)}$	comparison result of test $T^{(k)}$
Ψ	$\{C_0, C_1, \ldots, C_{2^m-1}\}$: set of all comparison patterns. The C_i are enumerated so that if $C_i = c_{m-1}\,c_{m-2}\,\cdots\,c_0$, then $c_{m-1}\,c_{m-2}\,\ldots\,c_0$ is the binary representation of i, possibly with leading 0's
$C^{(1\ldots\tau)}$	$C^{(1)}\,C^{(2)}\,\ldots\,C^{(\tau)}$.

Other, standard notation is given in "Information for Readers & Authors" at the rear of each issue.

2. PROBABILISTIC COMPARISON-BASED MODEL

Assumptions

1. System S has n units; several units can be faulty simultaneously.

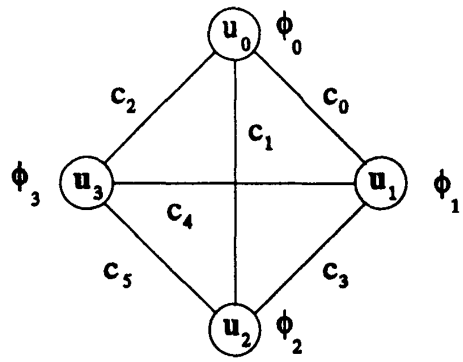

Figure 1. 4-Unit System with Complete Test-Graph

2. The faults in S are identified by a fault pattern Φ_j. Each fault pattern $\Phi_j \in \Theta$ is possible in S.

3. Individual tests T can be *incomplete* in the sense that they need not always cause a faulty unit to return an incorrect result.[5]

4. $T^{(k)}$ & $T^{(k')}$ can be replications or can include different tasks belonging to the same class.

5. Sequences $\mathbf{T}$ are applied periodically to S. $\mathbf{T}$ generates a sequence of τ comparison patterns $\{C^{(k)} \in \Psi,\ k = 1,2,\ldots,\tau\}$; this sequence is analyzed probabilistically to determine the faulty units. ◄

Let c_l connect $\mathbf{u}_i$ & $\mathbf{u}_j$. The behavior of the comparison test of $\mathbf{u}_i$ & $\mathbf{u}_j$ can be characterized & modeled using the following conditional probability test parameters:

$p_l = \Pr\{c_l=0 \mid \phi_i = \phi_j = 0\}$: Pr{agreement between fault-free units},

$q_l = \Pr\{c_l=1 \mid \phi_i \neq \phi_j\}$: Pr{disagreement between a faulty and a fault-free unit},

$r_l = \Pr\{c_l=1 \mid \phi_i = \phi_j = 1\}$: Pr{disagreement between faulty units}.

Homogeneity Assumptions

1. (To simplify analysis) The UUT are either identical or at least functionally equivalent, which is typical in multiprocessor-based systems.

2. There are non-stochastic constants p, q, r such that:

$p_l = p,\ q_l = q,\ r_l = r.$ ◄

Homogeneity-assumption 1 implies symmetry in the definition of q_l:

$$\Pr\{c_l=1 \mid \phi_i=0,\ \phi_j=1\} = \Pr\{c_l=1 \mid \phi_i=1,\ \phi_j=0\}.$$

For homogeneity-assumption 2, Chang [6] justifies that, in the comparison-based model, the components of C are mutually s-independent in the sense:

$$\Pr\{C \mid \Phi\} = \prod_{l=0}^{m-1} \Pr\{c_l \mid \Phi\}.$$

---

[3]The singular / plural of an acronym are always spelled the same.

[4]A *test* is a procedure for identifying whether a UUT is behaving normally or abnormally — in the comparison model — by means of the value it returns.

[5]Russell / Kime [31,32] suggested that it is hardly feasible to generate a complete test for the UUT; so, as indicated by Dahbura [15], the most realistic approach is to assume tests are incomplete.

Hence, the conditional distribution of $\Pr\{C|\Phi\}$ can be evaluated as a function of p, q, r. Blount [5] and Barsi [1] made a similar claim; however, the assumption is harder to justify in the PMC model, where each tester tests & decides the status of a subset of the UUT.

Nomenclature

Let c_l connect $\mathbf{u}_i$ & $\mathbf{u}_j$. c_l is:

a p-link, when both $\mathbf{u}_i$ & $\mathbf{u}_j$ are fault-free,

a q-link, when one of $\mathbf{u}_i$ or $\mathbf{u}_j$ is fault-free,

an r-link, otherwise.

2.1 Observation

Let c_l be a p-link for Φ; then:

$\Pr\{c_l = 0|\Phi\} = p$, and $\Pr\{c_l = 1|\Phi\} = 1 - p$.

Let c_l be a β-link for Φ ($\beta = q, r$); then:

$\Pr\{c_l = 1|\Phi\} = \beta$, and $\Pr\{c_l = 0|\Phi\} = 1 - \beta$.

$$\Pr\{C|\phi\} = \prod_{l=0}^{m-1} \Pr\{c_l|\phi\} = p^{n_p-d_p} \cdot (1-p)^{d_p} \cdot q^{n_q-d_q} \cdot (1-q)^{d_q} \cdot r^{n_r-d_r} \cdot (1-r)^{d_r} \quad (1)$$

Notation

β represents: p, q, or r

n_β number of β-links, $n_p + n_q + n_r = m$

d_β number of misdiagnosed β-links in the comparison pattern C

ξ_β $(1-\beta)/\beta$.

Then (1) simplifies to:

$$\Pr\{C|\phi\} = p^{n_p} \cdot q^{n_q} \cdot r^{n_r} \cdot \xi_p^{d_p} \cdot \xi_q^{d_q} \cdot \xi_r^{d_r} \quad (2)$$

2.2 *Example Use of* (2).

Assumptions

1. A system has 4 functionally-identical units with a complete connection assignment and one of the possible fault patterns.

2a. $\Phi = \phi_3\,\phi_2\,\phi_1\,\phi_0 = 0001$;

2b. $C = c_5\,c_4\,c_3\,c_2\,c_1\,c_0 = 000110$. ◄

Then,

$$\Pr\{C|\Phi\} = \prod_{l=0}^{5} \Pr\{c_l|\Phi\} = p^3 \cdot (1-q) \cdot q^2.$$

Here there is 1 faulty unit ($\mathbf{u}_0$) and 1 erroneous comparison result (c_0) as in figure 2.

The objective of testing a system is to find out whether a failure exists in the system at the time of the test, and then to locate any failed unit(s). After the fault location process is completed, some level of repair or reconfiguration must be initiated.

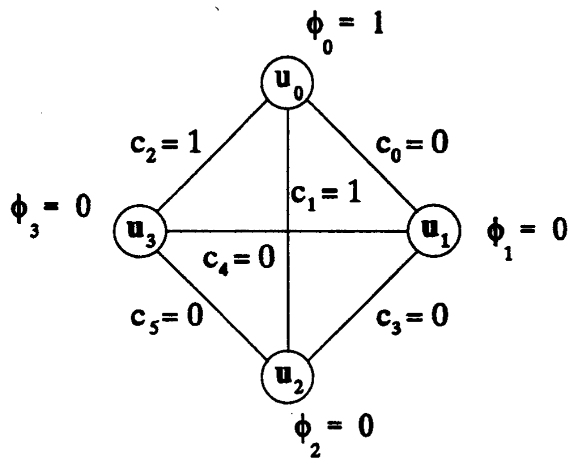

Figure 2. Possible Comparison Outcome with 1 Faulty Unit

2.3 Likelihood Table

The $C|\Phi_j$ can be combined into a likelihood table (LT) listing all the values of $\Pr\{C_i|\Phi_j\}$; LT is a probabilistic comparison table and can be computed prior to operation of the system, however its storage size is $O(2^n \cdot \cup^m)$, *eg*, table 1 for the system in figure 1.

Chang [6] developed an analytic method to avoid the need to store this enormous amount of data. The method requires $O(m)$ time to retrieve a data item or $O(\log m)$ time when prestored reference data are used. The number of data items retrieved at test time is small and has an absolute upper bound of τ. Section 6 provides an illustration.

3. ASSIGNMENT OF MULTIPLE TEST SETS

Tests $T^{(k)}$ are repeated τ times, thereby attempting to achieve pseudo-exhaustive testing [27], with variation of the individual tasks t_i within a class of test tasks, both to improve test coverage, since the test might not be complete, and account for random effects in the test environment as discussed in the Introduction.

The τ comparison results are conditionally s-independent, *ie*,

$$\Pr\{C^{(1\ldots\tau)}|\phi_j\} = \prod_{s=1}^{\tau} \Pr\{C^{(s)}|\phi_j\},$$

$\Pr\{C^{(s)}|\Phi_j\}$ can be obtained from the probabilistic comparison table for $j = 0, 1, \ldots, 2^n-1$. Consequently, the posterior distribution can be evaluated by Bayes theorem:

TABLE 1
Likelihood Table of $\Pr\{C_i|\Phi_j\}$
[$p=0.95$, $q=0.90$, $r=0.75$; the $\Pr\{C_i|\Phi_j\}$ values are multiplied by 10^4]

	Φ_0 0000	Φ_1 0001	Φ_2 0010	Φ_3 0011	Φ_4 0100	Φ_5 0101	Φ_6 0110	Φ_7 0111	Φ_8 1000	Φ_9 1001	Φ_{10} 1010	Φ_{11} 1011	Φ_{12} 1100	Φ_{13} 1101	Φ_{14} 1110	Φ_{15} 1111
$C_0=000000$	7351	9	9	0	9	0	0	0	9	0	0	0	0	0	0	2
$C_1=000001$	387	77	77	1	0	2	2	0	0	2	2	0	0	1	1	7
$C_2=000010$	387	77	0	2	77	1	2	0	0	2	0	1	2	0	1	7
$C_3=000011$	20	694	4	6	4	6	19	1	0	19	0	4	0	4	13	22
$C_4=000100$	387	77	0	2	0	2	0	1	77	1	2	0	2	0	1	7
$C_5=000101$	20	694	4	6	0	19	0	4	4	6	19	1	0	4	13	22
$C_6=000110$	20	694	0	19	4	6	0	4	4	6	0	4	19	1	13	22
$C_7=000111$	1	6250	0	58	0	58	1	13	0	58	1	13	1	13	114	66
$C_8=001000$	387	0	77	2	77	2	1	0	0	0	2	1	2	1	0	7
$C_9=001001$	20	4	694	6	4	19	6	1	0	0	19	4	0	13	4	22
$C_{10}=001010$	20	4	4	19	694	6	6	1	0	0	0	13	19	4	4	22
$C_{11}=001011$	1	37	37	58	37	58	58	4	0	1	1	38	1	38	38	66
$C_{12}=001100$	20	4	4	19	4	19	0	4	4	0	19	4	19	4	4	22
$C_{13}=001101$	1	37	37	58	0	173	0	13	0	0	173	13	1	38	38	66
$C_{14}=001110$	1	37	0	173	37	58	0	13	0	0	1	38	173	13	38	66
$C_{15}=001111$	0	329	2	519	2	519	3	38	0	3	9	114	9	114	342	198
$C_{16}=010000$	387	0	77	2	0	0	2	1	77	2	1	0	2	1	0	7
$C_{17}=010001$	20	4	694	6	0	0	19	4	4	19	6	1	0	13	4	22
$C_{18}=010010$	20	4	4	19	4	0	19	4	4	19	0	4	19	4	4	22
$C_{19}=010011$	1	37	37	58	0	0	173	13	0	173	0	13	1	38	38	66
$C_{20}=010100$	20	4	4	19	0	0	0	13	694	6	6	1	19	4	4	22
$C_{21}=010101$	1	37	37	58	0	1	1	38	37	58	58	4	1	38	38	66
$C_{22}=010110$	1	37	0	173	0	0	1	38	37	58	0	13	173	13	38	66
$C_{23}=010111$	0	329	2	519	0	3	9	114	2	519	3	38	9	114	342	198
$C_{24}=011000$	20	0	694	19	4	0	6	4	4	0	6	4	19	13	1	22
$C_{25}=011001$	1	0	6250	58	0	1	58	13	0	1	58	13	1	114	13	66
$C_{26}=011010$	1	0	37	173	37	0	58	13	0	1	0	38	173	38	13	66
$C_{27}=011011$	0	2	329	519	2	3	519	38	0	9	3	114	9	342	114	198
$C_{28}=011100$	1	0	37	173	0	1	0	38	37	0	58	13	173	38	13	66
$C_{29}=011101$	0	2	329	519	0	9	3	114	2	3	519	38	9	342	114	198
$C_{30}=011110$	0	2	2	1558	2	3	3	114	2	3	3	114	1558	114	114	198
$C_{31}=011111$	0	17	17	4675	0	27	27	342	0	27	27	342	82	1025	1025	593
$C_{32}=100000$	387	0	0	0	77	2	2	1	77	2	2	1	1	0	0	7
$C_{33}=100001$	20	4	4	0	4	19	19	4	4	19	19	4	0	4	4	22
$C_{34}=100010$	20	4	0	0	694	6	19	4	4	19	0	13	6	1	4	22
$C_{35}=100011$	1	37	0	0	37	58	173	13	0	173	1	38	0	13	38	66
$C_{36}=100100$	20	4	0	0	4	19	0	13	694	6	19	4	6	1	4	22
$C_{37}=100101$	1	37	0	0	0	173	1	38	37	58	173	13	0	13	38	66
$C_{38}=100110$	1	37	0	1	37	58	1	38	37	58	1	38	58	4	38	66
$C_{39}=100111$	0	329	0	3	2	519	9	114	2	519	9	114	3	38	342	198
$C_{40}=101000$	20	0	4	0	694	19	6	4	4	0	19	13	6	4	1	22
$C_{41}=101001$	1	0	37	0	37	173	58	13	0	1	173	38	0	38	13	66
$C_{42}=101010$	1	0	0	1	6250	58	58	13	0	1	1	114	58	13	13	66
$C_{43}=101011$	0	2	2	3	329	519	519	38	0	9	9	342	3	114	114	198
$C_{44}=101100$	1	0	0	1	37	173	0	38	37	0	173	38	58	13	13	66
$C_{45}=101101$	0	2	2	3	2	1558	3	114	2	3	1558	114	3	114	114	198
$C_{46}=101110$	0	2	0	9	329	519	3	114	2	3	9	342	519	38	114	198
$C_{47}=101111$	0	17	0	27	17	4675	27	342	0	27	82	1025	27	342	1025	593
$C_{48}=110000$	20	0	4	0	4	0	19	13	694	19	6	4	6	4	1	22
$C_{49}=110001$	1	0	37	0	0	1	173	38	37	173	58	13	0	38	13	66
$C_{50}=110010$	1	0	0	1	37	0	173	38	37	173	0	38	58	13	13	66
$C_{51}=110011$	0	2	2	3	2	3	1558	114	2	1558	3	114	3	114	114	198
$C_{52}=110100$	1	0	0	1	0	1	1	114	6250	58	58	13	58	13	13	66
$C_{53}=110101$	0	2	2	3	0	9	9	342	329	519	519	38	3	114	114	198
$C_{54}=110110$	0	2	0	9	2	3	9	342	329	519	3	114	519	38	114	198
$C_{55}=110111$	0	17	0	27	0	27	82	1025	17	4675	27	342	27	342	1025	593
$C_{56}=111000$	1	0	37	1	37	1	58	38	37	1	58	38	58	38	4	66
$C_{57}=111001$	0	0	329	3	2	9	519	114	2	9	519	114	3	342	38	198
$C_{58}=111010$	0	0	2	9	329	3	519	114	2	9	3	342	519	114	38	198
$C_{59}=111011$	0	0	17	27	17	27	4675	342	0	82	27	1025	27	1025	342	593
$C_{60}=111100$	0	0	2	9	2	9	3	342	329	3	519	114	519	114	38	198
$C_{61}=111101$	0	0	17	27	0	82	27	1025	17	27	4675	342	27	1025	342	593
$C_{62}=111110$	0	0	0	82	17	27	27	1025	17	27	27	1025	4675	342	342	593
$C_{63}=111111$	0	1	1	246	1	246	246	3075	1	246	246	3075	246	3075	3075	1780

$$\Pr\{\phi_j | C^{(1\ldots\tau)}\} = \frac{\Pr\{C^{(1\ldots\tau)} | \phi_j\} \cdot \Pr\{\phi_j\}}{\sum_{l=0}^{2^n-1} \Pr\{C^{(1\ldots\tau)} | \phi_j\} \cdot \Pr\{\phi_l\}} \quad (3)$$

$j = 0, 1, \ldots, 2^n-1.$

As a rationale for the prior distribution on the parameter of interest, $\Phi \in \Theta$, we assume, as others have done, that $\Phi \equiv \phi_{n-1}\ \phi_{n-2}\ \ldots\ \phi_0$ is i.i.d. with an exponential distribution [19,28,29,34,37,38]. The choice of the prior probability $\Pr\{\Phi\}$ does not affect the discussion in section 4. Hence the procedure is robust with respect to the choice of the prior distribution.

4. BAYES ANALYSIS OF FAULT LOCATION

4.1 Background

In general, there are two ways to locate faults using Bayes analysis.

1. As in classical inference methods that mostly deal with the posterior distribution, choose either point estimation or set estimation to locate a fault [6,8-13].
2. Use a loss function and turn the problem into one from decision theory.

We use #2, the Bayes decision-theoretic approach, which enables us to estimate the fault status of each unit by point estimation with the choice of a reasonable loss function [2]. Distance is a reasonable measure for all misdiagnosed results. The loss function is computed bitwise from the global fault pattern. We use this loss function because it is computationally efficient and its center mean & mode are the same.

To assist the Bayes analysis, it is necessary to transform the likelihood table, $\Pr\{C_i|\Phi_j\}$, to a bitwise version of the likelihood table, $\Pr\{C_i|\phi_k\}$. As shown in table 2, the number of columns (which was 2^n in table 1) has become simply $2n$. We extend the expressions from previous sections and consider the case $\phi_k = \delta$, where $\delta = 0$ or 1, to obtain:

$$\Pr\{C_i|\phi_k=\delta\} = \frac{\Pr\{C_i,\ \phi_k=\delta\}}{\Pr\{\phi_k=\delta\}}$$

$$= \frac{1}{\Pr\{\phi_k=\delta\}} \sum_{\{\phi_j \in \Theta: \phi_{j,k}=\delta\}} \Pr\{C_i|\phi\} \cdot \Pr\{\phi_j\} \quad (4)$$

The marginal probabilities $\Pr\{C_i|\phi_k\}$, $k = 0,\ldots, n-1$ in table 2 are generated from the conditional probability distribution in table 1. Hence the column sums of the table are 1, but this is not true of the row sums — as anticipated. Table 2 is generated using (4) and was validated by two different programs written separately in the C & Matlab languages. The shape of the distribution in each column is similar to that in table 1 but the rate at which the values decrease is slower than in table 1. This observation is reasonable since the bitwise conditional distribution in table 2 compresses all possible fault conditions of S given the fault status, $\phi_k=\delta$, of a single unit. Although this bitwise distribution has a less pronounced shape than the

TABLE 2

Simplified Bitwise Likelihood Table of $\Pr\{C_i|\phi_k\}$

[Based on table 1. $\Pr\{\phi_k=0\}=0.8$, $\Pr\{\phi_k=1\}=1-0.8=0.2$, for all k. The $\Pr\{C_i|\Phi_j\}$ values are multiplied by 10^4. See table 1 for the bit-patterns for each C_i.]

	$\phi_0=0$	$\phi_0=1$	$\phi_1=0$	$\phi_1=1$	$\phi_2=0$	$\phi_2=1$	$\phi_3=0$	$\phi_3=1$
C_0	3767	5	3767	5	3767	5	3767	5
C_1	208	40	208	40	218	1	218	1
C_2	208	40	218	1	208	40	218	1
C_3	12	360	101	6	101	6	101	3
C_4	208	40	218	1	218	1	208	40
C_5	12	360	101	6	101	3	101	6
C_6	12	360	101	3	101	6	101	6
C_7	2	3224	804	13	804	13	804	13
C_8	218	1	208	40	208	40	218	1
C_9	101	6	12	360	101	6	101	3
C_{10}	101	6	101	6	12	360	101	3
C_{11}	12	37	12	37	12	37	20	5
C_{12}	13	8	13	8	13	8	13	8
C_{13}	11	51	11	51	17	26	17	26
C_{14}	11	51	17	26	11	51	17	26
C_{15}	4	312	60	86	60	86	76	23
C_{16}	218	1	208	40	218	1	208	40
C_{17}	101	6	12	360	101	3	101	6
C_{18}	13	8	13	8	13	8	13	8
C_{19}	11	51	11	51	17	26	17	26
C_{20}	101	6	101	6	101	3	12	360
C_{21}	12	37	12	37	20	5	12	37
C_{22}	11	51	17	26	17	26	11	51
C_{23}	4	312	60	86	76	23	60	86
C_{24}	101	3	12	360	101	6	101	6
C_{25}	804	13	2	3224	804	13	804	13
C_{26}	17	26	11	51	11	51	17	26
C_{27}	60	86	4	312	60	86	76	23
C_{28}	17	26	11	51	17	26	11	51
C_{29}	60	86	4	312	76	23	60	86
C_{30}	52	214	52	214	52	214	52	214
C_{31}	15	674	15	674	159	99	159	99
C_{32}	218	1	218	1	208	40	208	40
C_{33}	13	8	13	8	13	8	13	8
C_{34}	101	6	101	3	12	360	101	6
C_{35}	11	51	17	26	11	51	17	26
C_{36}	101	6	101	3	101	6	12	360
C_{37}	11	51	17	26	17	26	11	51
C_{38}	12	37	20	5	12	37	12	37
C_{39}	4	312	76	23	60	86	60	86
C_{40}	101	3	101	6	12	360	101	6
C_{41}	17	26	11	51	11	51	17	26
C_{42}	804	13	804	13	2	3224	804	13
C_{43}	60	86	60	86	4	312	76	23
C_{44}	17	26	17	26	11	51	11	51
C_{45}	52	214	52	214	52	214	52	214
C_{46}	60	86	76	23	4	312	60	86
C_{47}	15	674	159	99	15	674	159	99
C_{48}	101	3	101	6	101	6	12	360
C_{49}	17	26	11	51	17	26	11	51
C_{50}	17	26	17	26	11	51	11	51
C_{51}	52	214	52	214	52	214	52	214
C_{52}	804	13	804	13	804	13	2	3224
C_{53}	60	86	60	86	76	23	4	312
C_{54}	60	86	76	23	60	86	4	312
C_{55}	15	674	159	99	159	99	15	674
C_{56}	20	5	12	37	12	37	12	37
C_{57}	76	23	4	312	60	86	60	86
C_{58}	76	23	60	86	4	312	60	86
C_{59}	159	99	15	674	15	674	159	99
C_{60}	76	23	60	86	60	86	4	312
C_{61}	159	99	15	674	159	99	15	674
C_{62}	159	99	159	99	15	674	15	674
C_{63}	49	404	49	404	49	404	49	404

global distribution in table 1, table 2 is far smaller. Together with the bitwise method in this section, the small size of the table appreciably reduces the complexity of the analysis. Besides that, $\Pr\{C_i|\phi_k=0\} \neq \Pr\{C_i|\phi_k=1\}$ in general. Such an equality, rarely occurring in practice, would yield an inconclusive test result and compromise the comparison steps in the diagnosis process (see step 3 of BWB in this section and remark 5.5).

4.2 Point Estimation

Use the observed comparison patterns to determine $\hat{\Phi}_{ML} \equiv \hat{\phi}_{n-1} \ldots \hat{\phi}_k \ldots \hat{\phi}_0 \in \Theta$. $\hat{\Phi}_{ML}$ is the 'generalized maximum likelihood estimate' of Φ, *viz*, the largest mode of the posterior distribution $\Pr\{\Phi|C^{(1\ldots\tau)}\}$ [2: p 133]; this maximum likelihood estimate is anticipated to be unique. It is $O(2^n)$ to examine all $\Phi_j \in \Theta$ to find the maximum of all the $\Pr\{\Phi|C^{(1\ldots\tau)}\}$ if all Φ_j have to be examined. However, [8,10] gave a heuristic-based search algorithm to find $\hat{\Phi}_{ML}$; it has only $O(n)$ worst case complexity to locate the faults with a $1-\alpha$ *s*-confidence level [2: p 414].

4.3 Set Estimation

It is possible to obtain the $1-\alpha$ highest posterior density (HPD) credible region for the r.v. Φ, given some small real number α. To calculate the HPD we consider all subsets $\Gamma' \subset \Theta$ such that $\Pr\{\Gamma'|C^{(1\ldots\tau)}\} \geq 1-\alpha$. Among these subsets Γ' we must find the one with the highest density of posterior probability, *viz*, the subset such that $\kappa(\Gamma') = \min(\Pr\{\Phi_j|C^{(1\ldots\tau)}\}: \Phi_j \in \Gamma')$ is the largest.

The HPD for Φ is the subset Γ which maximizes $\kappa(\Gamma')$:

$$\kappa(\Gamma) = \max(\kappa(\Gamma'): \Pr\{\Gamma'|C^{(1\ldots\tau)}\} \geq 1-\alpha).$$

The computation cost of finding all such Γ' is exponential. An alternative is to find a set Γ' for which the $\Pr\{\Phi_j|C^{(1\ldots\tau)}\}$ for all $\Phi_j \in \Gamma'$ are relatively large and use that Γ' as a reasonable replacement for Γ. Since $\hat{\Phi}_{ML}$ is the most likely system fault pattern, the other fault patterns $\Phi_j \in \Theta$ can be considered as misdiagnosed. Hence the number of misdiagnosed units can be used as a measure of distance of any fault pattern Φ_j from $\hat{\Phi}_{ML}$.

Notation

$d(A,B)$ distance between A & B
$\oplus$ 'exclusive OR' operation.

The $\Pr\{\Phi_j|C^{(1\ldots\tau)}\}$ decreases as $d(\Phi_j, \hat{\Phi}_{ML})$ [6,8,10]. Hence we construct Γ' using only those Φ_j closest to $\hat{\Phi}_{ML}$.

$$d(\hat{\Phi}_{ML}, \Phi_j) \equiv \sum_{k=0}^{n-1} |\hat{\phi}_k - \phi_{j,k}|, \text{ so that } d(\hat{\Phi}_{ML}, \Phi_j) \in \{0,1,\ldots,n\}$$

Furthermore, since $\hat{\phi}_k = 0$ or 1, and $\phi_{j,k} = 0$ or 1, it follows that:

$$d(\hat{\Phi}_{ML}, \Phi_j) \equiv \sum_{k=0}^{n-1} |\hat{\phi}_k - \phi_{j,k}| = \sum_{k=0}^{n-1} (\hat{\phi}_k \oplus \phi_{j,k})$$

$$= \sum_{k=0}^{n-1} (\hat{\phi}_k - \phi_{j,k})^2.$$

The remaining step is to construct a 1-α *credible region* for Φ which we anticipate to approximate the HPD region. Although $\Pr\{\Phi_j|C^{(1\ldots\tau)}\}$ does not necessarily decrease as $d(\hat{\Phi}_{ML}, \Phi_j)$ increases, the fault pattern with fewer misdiagnosed links should appear more frequently. Therefore we include the Φ_j that have the smallest $d(\hat{\Phi}_{ML}, \Phi_j)$ first. In this manner, it is possible to find the *minimum* $h \in \{0, 1, 2, \ldots, n\}$ such that if $\Gamma \equiv \{\Phi_j \in \Theta : 0 \leq d(\hat{\Phi}_{ML}, \Phi_j) \leq h\}$ then $\Pr\{\Gamma|C^{(1\ldots\tau)}\} \geq 1-\alpha$. Thus, the region is a $1-\alpha$ credible region for Φ. Since Γ does not necessarily contain all $\Phi_j \in \Theta$ that have higher posterior density than $\Phi_{j'} \in \Gamma$, Γ cannot be assumed to be the $1-\alpha$ HPD credible region for Φ. However, the computation of Γ is more efficient.

4.4 Bayes Decision Theoretic Approach

We now turn to the decision theoretic approach to point estimation. A *decision rule* is a mapping from the test results $C^{(1\ldots\tau)}$ to the fault patterns Φ_j: Given a particular test result, the rule decides on a particular fault pattern. The Bayes approach considers the true fault pattern which we denote by $\Phi \equiv \phi_{n-1} \ldots \phi_k \ldots \phi_0$, although it is unknown, of course. We must then choose a reasonable loss function. In section 4.3, we claimed that distance is a reasonable measure of the amount of misdiagnosis in a pattern. Given the test result $C^{(1\ldots\tau)}$ suppose an arbitrary decision rule assigns the fault pattern $\tilde{\Phi} = \tilde{\phi}_{n-1} \ldots \tilde{\phi}_k \ldots \tilde{\phi}_0 \in \Theta$. Then consider the *loss function*:

$$\mathcal{L}(\tilde{\Phi},\Phi_j) \equiv d(\tilde{\Phi},\Phi) \equiv \sum_{k=0}^{n-1} (\tilde{\phi}_k - \phi_k)^2 = \sum_{k=0}^{n-1} |\tilde{\phi}_k - \phi_k|$$

$$= \sum_{k=0}^{n-1} (\tilde{\phi}_k \oplus \phi_k).$$

The statistical importance of this function is that, since ϕ_k & $\tilde{\phi}_k$ only take the values 0 & 1, it has the same properties as the square-error loss and absolute-error loss. Moreover, the loss function is computationally practical since it can be evaluated by the 'exclusive OR' operation which is efficient. This shortens the unit computation time.

Given a loss function, we have to consider the fact that Φ is unknown and, in fact, every Φ_j is possible and, given $C^{(1\ldots\tau)}$, the probability of Φ being Φ_j is exactly the posterior probability $\Pr\{\Phi_j|C^{(1\ldots\tau)}\}$. We define a risk function ρ as the *s*-expected loss, given this probability distribution for Φ:

$$\rho(\tilde{\Phi},\phi) \equiv \mathrm{E}\{\mathcal{L}(\tilde{\Phi},\Phi)|C^{(1\ldots\tau)}\}$$

In this equation the r.v. which is conditionally s-dependent on $C^{(1...\tau)}$ is Φ. The risk function can be expanded to explicit sums as follows, but the sums here are over 2^n elements:

$$\rho(\tilde{\Phi},\Phi) \equiv \mathrm{E}\{\mathcal{L}(\tilde{\Phi},\Phi)|C^{(1..\tau)}\}$$

$$= \sum_{\Phi\in\Theta} \mathcal{L}(\tilde{\Phi},\Phi)\cdot\Pr\{\Phi|C^{(1..\tau)}\}$$

$$= \sum_{\Phi\in\Theta} \mathcal{L}(\tilde{\Phi},\Phi) \frac{\Pr\{C^{(1...\tau)}|\Phi\}\cdot\Pr\{\Phi\}}{\sum_{\Phi\in\Theta}\Pr\{C^{(1...\tau)}|\Phi\}\cdot\Pr\{\Phi\}}$$

$$= \frac{\sum_{\Phi\in\Theta}\mathcal{L}(\tilde{\Phi},\Phi)\cdot\Pr\{C^{(1...\tau)}|\Phi\}\cdot\Pr\{\Phi\}}{\sum_{\Phi\in\Theta}\Pr\{C^{(1..\tau)}|\Phi\}\cdot\Pr\{\Phi\}}$$

Now we can select the Bayes decision rule which makes $C^{(1...\tau)}$ correspond to $\Phi_B^* = \phi_{n-1}^* \ldots \phi_k^* \ldots \phi_0^* \in \Theta$, to be the one that minimizes the risk function:

$$\rho(\Phi_B^*,\Phi) \equiv \mathrm{E}\{\mathcal{L}(\Phi_B^*,\Phi)|C^{(1...\tau)}\}$$

$$= \mathrm{E}\left\{\sum_{k=0}^{n-1} (\phi_k^*, \phi_k)|C^{(1...\tau)}\right\} \qquad (5)$$

$$= \sum_{k=0}^{n-1} \rho(\phi_k^*, \phi_k)$$

$$\rho(\phi_k^*, \phi_k) \equiv \mathrm{E}\{\phi_k^*\oplus\phi_k|C^{(1...\tau)}\}$$

The full derivation is given in [6,13]; and the r.v. which are conditionally s-dependent on $C^{(1...\tau)}$ are Φ and its components ϕ_k. Then $\Phi_B^* = \phi_{n-1}^*\ldots\phi_k^*\ldots\phi_0^*$ minimizes $\rho(\Phi_B^*, \Phi)$ iff ϕ_k^* minimizes $\rho(\phi_k^*, \phi)$ for all $k = 0, 1, 2, \ldots, n-1$. Hence the complex global analysis of all the $\Phi \in \Theta$ is decomposed into a simple bitwise analysis. In other words, in order to compute the Φ_B^* assigned by the global Bayes decision rule, it is sufficient to find all bitwise assignments $\phi_k^* = 0$ or 1 ($k = 0, 1, 2, \ldots, n-1$) that minimize:

$$\rho(\phi_k^*, \phi_k) = \sum_{\delta=0,1} (\phi_k^*\oplus\delta)\cdot\zeta(\phi_{j,k};\delta)\Big/\sum_{\delta=0,1}\zeta(\phi_{j,k};\delta), \qquad (6a)$$

or equivalently minimize the numerator,

$$\sum_{\delta=0,1} (\phi_k^*\oplus\delta)\cdot\zeta(\phi_k;\delta), \text{ for } \phi_k^* = 0 \text{ or } 1; \qquad (6b)$$

$$\zeta(\phi;\delta) \equiv \Pr\{C^{(1...\tau)}|\phi=\delta\}\cdot\Pr\{\phi=\delta\}.$$

The complexity of analysis, by using this methodology, has been reduced dramatically since sums with 2^n terms are replaced by the sum of two terms! However, it can be reduced further by using the fact that:

since ϕ_k^* & ϕ_k are binary variables, we know that if $\phi_k^*=0$, then $\phi_k^*\oplus\phi_k = \phi_k$.

Hence:

$$\rho(\phi_k^*=0, \phi_k) = \mathrm{E}\{\phi_k|C^{(1...\tau)}\} = \Pr\{\phi_k=1|C^{(1...\tau)}\}. \qquad (7a)$$

Similarly, if $\phi_k^* = 1$, then $\phi^*{}_k\oplus\phi_k = (\phi_k)' = 1 - \phi_k$, so that

$$\rho(\phi_k^*=1, \phi_k) = \mathrm{E}\{1 - \phi_k|C^{(1...\tau)}\} = 1 - \Pr\{\phi_k=1|C^{(1...\tau)}\}$$

$$= \Pr\{\phi_k=0|C^{(1...\tau)}\}. \qquad (7b)$$

Therefore,

$$\rho(1, \phi_k) < \rho(0, \phi_k) \text{ iff } \Pr\{\phi_k=1|C^{(1...\tau)}\} > \Pr\{\phi_k=0|C^{(1...\tau)}\}. \qquad (8)$$

Note that $\Pr\{\phi_k=1|C^{(1...\tau)}\} + \Pr\{\phi_k=0|C^{(1...\tau)}\} = 1$; so $\Pr\{\phi_k=1|C^{(1...\tau)}\} > \Pr\{\phi_k=0|C^{(1...\tau)}\}$ iff $\Pr\{\phi_k=1|C^{(1...\tau)}\} > 0.5$. Furthermore, since:

$$\Pr\{\phi_k=1|C^{(1...\tau)}\} = \zeta(\phi_k;1)\Big/\sum_{\delta=0,1}\zeta(\phi_k;\delta),$$

$$\mathrm{E}\{\phi_k|C^{(1...\tau)}\} = \Pr\{\phi_k=1|C^{(1...\tau)}\} > 0.5 \text{ iff } \zeta(\phi_k;1) > \zeta(\phi_k;0).$$

The decision process outlined above can be summarized as the following bitwise (BWB) version of the Bayes decision algorithm to perform fault location:

4.5 *BWB for Fault Location*

Choose the components $\phi_{n-1}^*\ldots\phi_k^*\ldots\phi_0^*$ of Φ_B^* as follows: For each $k = n-1, \ldots, 0$:

1. If $\zeta(\phi_k; 1\} > \zeta(\phi_k; 0\}$, then choose $\phi_k^* = 1$.
2. If $\zeta(\phi_k; 1\} < \zeta(\phi_k; 0\}$, then choose $\phi_k^* = 0$.
3. Otherwise, the test is inconclusive.

If a set estimation is desired, the procedures developed in section 4.3 can be applied to obtain a $1-\alpha$ credibility region for Φ by substituting Φ_B^* for $\hat{\Phi}_{\mathrm{ML}}$.

5. REMARKS

5.1 $\Pr\{C^{(1...\tau)}|\phi_k\} = \prod_{s=1}^{\tau}\Pr\{C^{(s)}|\phi_k\}$

is computable, since $\Pr\{C^{(s)}|\phi_k\}$ can be obtained from the simplified bitwise likelihood table, *eg*, table II. Furthermore, the size of this likelihood table is $2^m\cdot n\cdot 2 = n\cdot 2^{m+1}$, which is much smaller than $2^m \cdot 2^n = 2^{m+n}$ of the full likelihood table such as table I. Also the data in the both tables can be computed prior to run-time.

5.2 The run-time complexity of BWB is $O(n)$. This use of bitwise analysis dramatically surpasses previous results in the literature and is a very good result in theory & practice. From the theoretical point of view, it reduces the complexity of the Bayes analysis; this complexity is one of the main deficiencies of the Bayes approach. In practice, BWB outperforms other methods in the literature since it does not really require any assumptions. The only disadvantage is in the generation of the probabilistic comparison table. Again, this is a one time computation.

5.3 Since the loss function is square-error loss, the posterior mean $E\{\phi_k | C^{(1\cdots\tau)}\}$ is the Bayes rule [2: p 161, result 3]. However, because ϕ_k^* is binary, it is reasonable to choose $\phi_k^*=1$ if $E\{\phi_k | C^{(1\cdots\tau)}\} > 0.5$ (the same decision rule as step 1). If the loss function were generalized to a weighted square-error loss, then the Bayes rule can also be obtained [2: p 161, result 4]. However, that form of the Bayes rule is not as simple as the BWB. In addition, since the loss function is also the absolute error loss, the median of $\Pr\{\phi_k | C^{(1\cdots\tau)}\}$ also provides a Bayes rule [2: p 162, result 5]. Again, since ϕ_k^* is binary, this is equivalent to the BWB. If the loss function is generalized to a linear loss, the decision rule can be obtained in a similar fashion [2: p 162, result 6]. Thus, BWB is valid for a class of the usual loss functions.

5.4 BWB is consistent with one's intuition and has the important property of positive Bayes decision rules, namely admissibility. This gives extra trust in our intuition.

5.5 If the inconclusive test result described in step 3 of BWB were to occur, it could be resolved by upgrading the quality of the test tasks t_i or by increasing τ, the number of tests. The latter would be necessary if faults were intermittent.

5.6 BWB accommodates all possible faulty & fault-free systems under test, without any increase in complexity when the fault-free state is diagnosed, permitting the algorithm to be applied to monitor a system periodically. Further, BWB is able to distinguish *truly* faulty units from those which appear faulty due to the imperfect environment, thus eliminating unnecessary hardware replacement or reconfiguration before the system recovery process performs rollback to a fault-free state.

5.7 The comparison-based probabilistic model and the Bayes inference algorithm make BWB complete in the statistical sense, since the model together with the BWB can accommodate all possible random effects. It is practical because the computations are simple binary operations, with linear complexity. The necessary data are directly observable during the testing process.

5.8 A visual simulation tool (ViSiT) [7] has been implemented in C++ to test the validity of BWB. Test results have demonstrated that the algorithm is both efficient & accurate. When testing the algorithms for the early phase of system operation, 98% of the tests gave exact fault coverage. For tests corresponding to the late phase of system operation, 85% of the tests gave exact fault coverage. In both cases, the remaining tests diagnosed as faulty both the actual faulty units *and* some of the fault-free units. ViSiT also demonstrates that the excess of fault-free units being misdiagnosed begins to show when the fraction of faulty units exceeds 50%.

6. EXAMPLE OF LOCATING SYSTEM FAULTS

Figure 1 is a 4-unit system; figure 2 shows the possible comparison outcome with 1 faulty unit. Let $\tau=10$. The following comparison patterns are observed:

$$C_{40},\ C_{42},\ C_{42},\ C_{42},\ C_{40},\ C_{10},\ C_{34},\ C_0,\ C_{57},\ C_{42}.$$

These patterns are sorted & counted as follows:

Patterns (sorted)	C_0	C_{10}	C_{34}	C_{40}	C_{42}	C_{57}
count	1	1	1	2	4	1

According to BWB, all we need to do is to sum $\Pr\{C^{(1\ldots10)} | \phi_k =1\} \cdot \Pr\{\phi_k=1\}$ from table II and compare it with the corresponding sum of $\Pr\{C^{(1\ldots10)} | \phi_k=0\} \cdot \Pr\{\phi_k=0\}$. Repeat the same step until all the bits are done: $k = 0$ to 3, in this case. This also implies the bitwise computations are carried out only on the relevant ϕ_k column. Let $k=0$ and compute for $\phi_0=1$:

$$\Pr\{C^{(1\ldots10)} | \phi_0=1\} \cdot \Pr\{\phi_0=1\}$$

$$= \Pr\{C_0 | \phi_0=1\} \cdot \Pr\{C_{10} | \phi_0=1\} \cdot \Pr\{C_{34} | \phi_0=1\}$$

$$\cdot [\Pr\{C_{40} | \phi_0=1\}]^2 \cdot [\Pr\{C_{42} | \phi_0=1\}]^4 \cdot \Pr\{C_{57} | \phi_0=1\}$$

$$\cdot \Pr\{\phi_0=1\} = 10^{-40} \cdot [5 \cdot 6 \cdot 6 \cdot 3^2 \cdot 13^4 \cdot 23] \cdot (0.2).$$

Similarly, for $\phi_0=0$:

$$\Pr\{C^{(1\ldots10)} | \phi_0=0\} \cdot \Pr\{\phi_0=0\} = 10^{-40} \cdot [3767 \cdot 101 \cdot 101 \cdot 101^2$$

$$\cdot 804^4 \cdot 76] \cdot (0.8).$$

Since $\Pr\{C^{(1\ldots10)} | \phi_0=1\} \cdot \Pr\{\phi_0=1\}$ is obviously smaller than $\Pr\{C^{(1\ldots10)} | \phi_0=0\} \cdot \Pr\{\phi_0=0\}$, we choose $\phi_0^*=0$. In the next step, k is incremented by 1 and the process repeats. We conclude from the iterations of BWB that the fault pattern of the system is $\Phi^* = \phi_3^* \phi_2^* \phi_1^* \phi_0^* = 0100 = \Phi_4$.

ACKNOWLEDGMENT

We are pleased to thank Professor Jiunn T. Hwang for his helpful discussions and suggestions.

REFERENCES

[1] F. Barsi, "Probabilistic syndrome decoding in self-diagnosable digital systems", *Digital Processes*, vol 7, 1981, pp 33-46.

[2] J.O. Berger, *Statistical Decision Theory and Bayesian Analysis* (2nd ed), 1985; Springer-Verlag.

[3] P. Berman, A. Pelc, "Distributed probabilistic fault diagnosis for multiprocessor systems", *Proc. 20th Symp. Fault-Tolerant Computing*, 1990, pp 340-346.

[4] D.M. Blough *et al*, "Fault diagnosis for sparsely interconnected multiprocessor systems", *Proc. 19th Symp. Fault-Tolerant Computing*, 1989, pp 62-69.

[5] M. Blount, "Probabilistic treatment of diagnosis in digital systems", *Proc. 7th Symp. Fault-Tolerant Computing*, 1977, pp 72-77.

[6] Y.L.C. Chang, "Characterization and analysis of the random nature of fault location", *PhD Dissertation*, 1991 Dec; State University of New York at Binghamton.

[7] Chang, Hsu, Hsu, Lam, Lander, Lau, Wu, "A visual simulation tool for design for diagnosability of dependable systems", *Proc. Int'l Conf. Reliability & Quality in Design*, 1994 Mar, pp 164-168.

[8] Y.L.C. Chang, L. Lander, "Classical inference methods for fault location in homogeneous systems", *Proc. IASTED Int'l Conf. Reliability, Quality Control, and Risk Assessment*, 1992 Nov, pp 57-61.

[9] Y.L.C. Chang, L. Lander, "Management of testing policies for fault location of failures dependent upon operation-time in multiprocessor systems", *Proc. IEEE 3rd Int'l Workshop on Responsive Computer Systems*, 1993, pp 166-172.

[10] Y.L.C. Chang, L. Lander, "An inference design for fault location in real-time control systems", *J. Systems & Software* (special issue on Fault Tolerance in Real-Time Systems), vol 25, 1994 Apr, pp 3-21.

[11] Y.L.C. Chang, L. Lander, H. Lu, M. Wells, "Bayesian analysis for fault location in homogeneous distributed systems", *Proc. IEEE 12th Symp. Reliable Distributed Systems*, 1993, pp 44-53.

[12] Y.L.C. Chang, L. Lander, H. Lu, M. Wells, "Bayesian inference for fault diagnosis in real-time distributed systems", *Proc. IEEE 2nd Asian Test Symp. (ATS'93)*, 1993, pp 333-338.

[13] Y.L.C. Chang, L. Lander, H. Lu, M. Wells, "The theory and application of Bayesian inference in system-level fault location", Technical Report No. CS-TR-93-10, 1993 Aug; Dept. of Computer Science, SUNY-Binghamton.

[14] K.Y. Chwa, S.L. Hakimi, "Schemes for fault tolerant computing: A comparison of modularly redundant and t-diagnosable systems", *Information and Control*, vol 49, 1981 Jun, pp 212-238.

[15] A.T. Dahbura, K.K. Sabnani, L.L. King, "The comparison approach to multiprocessor fault diagnosis", IEEE Trans. Computers, vol C-36, 1987 Mar, pp 373-377.

[16] D. Fussell, S. Rangarajan, "Probabilistic diagnosis of multiprocessor systems with arbitrary connectivity", *Proc. 19th Symp. Fault-Tolerant Computing*, 1989, pp 560-565.

[17] R. Gupta, I.V. Ramakrishnan, "System-level fault diagnosis in malicious environments", *Proc. 17th Symp. Fault-Tolerant Computing*, 1987, pp 184-189.

[18] S.L. Hakimi, A.T. Amin, "Characterization of the connection assignment of diagnosable systems", *IEEE Trans. Computers*, vol C-23, 1974 Jan, pp 86-88.

[19] P.K. Lala, *Fault Tolerance and Testable Hardware Design*, 1985; Prentice-Hall.

[20] L. Lander, Y.L.C. Chang, "Generation and analysis of probabilistic diagnostic information in homogeneous systems", *Proc. IASTED Int'l Conf. Reliability, Quality Control and Risk Assessment*, 1993, pp 155-158.

[21] S. Lee, K.G. Shin, "Optimal multiple syndrome probabilistic diagnosis", *Proc. 20th Symp. Fault-Tolerant Computing*, 1990, pp 324-331.

[22] T. H. Lin, K. G. Shin, "Location of a faulty module in a computing system", *IEEE Trans. Computers*, vol 39, 1990 Feb, pp 182-194.

[23] J. Maeng, M. Malek, "A comparison connection assignment for self-diagnosis of multiprocessor systems", *Proc. 11th Symp. Fault-Tolerant Computing*, 1981, pp 173-175.

[24] S. Maheshwari, S. Hakimi, "On models for diagnosable systems and probabilistic fault diagnosis", *IEEE Trans. Computers*, vol C-25, 1976 Mar, pp 228-236.

[25] M. Malek, "A comparison connection assignment for diagnosis of multiprocessor systems", *Proc. 10th Symp. Fault-Tolerant Computing*, 1980, pp 31-36.

[26] S. Mallela, G.M. Masson, "Diagnosable systems for intermittent faults", *IEEE Trans. Computers*, vol C-27, 1978 Jun, pp 560-566.

[27] E.J. McCluskey, "Verification testing — A pseudo-exhaustive test technique", *IEEE Trans. Computers*, vol C-33, 1984 Jun, pp 541-546.

[28] Y.W. Ng, "Reliability modeling and analysis for fault-tolerant computers", *PhD Dissertation*, 1976; Dept. of Computer Science, UCLA.

[29] D.K. Pradhan (Ed), *Fault-Tolerant Computing: Theory and Techniques*, 1986, chapter 8, vol II; Prentice-Hall.

[30] F.P. Preparata, G. Metze, R.T. Chien, "On the connection assignment problem of diagnosable systems", *IEEE Trans. Computers*, vol EC-16, 1967 Dec, pp 848-854.

[31] J. Russell, C.R. Kime, "System fault diagnosis: Closure and diagnosability with repair", *IEEE Trans. Computers*, vol C-24, 1975 Nov, pp 1078-1088.

[32] J. Russell, C.R. Kime, "System fault diagnosis: Masking, exposure, and diagnosability without repair", *IEEE Trans. Computers*, vol C-24, 1975 Dec, pp 1155-1161.

[33] A. Sengupta, A.T. Dahbura, "On self-diagnosable multiprocessor systems: Diagnosis by the comparison approach", *Proc. 19th Symp. Fault-Tolerant Computing*, 1989, pp 54-61.

[34] D.P. Siewiorek, R.S. Swarz, *The Theory and Practice of Reliable System Design*, 1982; Digital Press.

[35] L. Simoncini, A.D. Friedman, "Incomplete fault coverage in modular multiprocessor systems", *Proc. ACM Annual Conf*, 1978, pp 210-216.

[36] J.E. Smith, "Universal system diagnosis algorithms", *IEEE Trans. Computers*, vol C-28, 1979 May, pp 374-378.

[37] G. Sullivan, "System-level fault diagnosability in probabilistic and weighted models", *Proc. 17th Symp. Fault-Tolerant Computing*, 1987, pp 190-195.

[38] K.S. Trivedi, *Probability and Statistics with Reliability, Queuing, and Computer Science Applications*, 1982; Prentice-Hall.

AUTHORS

Dr. Yu Lo Cyrus Chang; Dept. of Computer Science and Electrical Eng'g; The University of Tennessee; Chattanooga, Tennessee 37403-2598 USA.

Yu Lo C. Chang (M'92) received his BS in Electrical Engineering from the National Taipei Institute of Technology and his MS & PhD in Computer Science from the State University of New York at Binghamton. He is an Assistant Professor in the Dept. of Computer Science and Electrical Engineering at the University of Tennessee at Chattanooga. His research interests include fault-tolerant computing, fuzzy systems, hard real-time systems, reliable parallel and distributed processing, statistical decision theory, VLSI testing, and visual synthesis tools. He is a member of IEEE Computer and Reliability Societies.

Dr. Leslie C. Lander; Dept. of Computer Science; State University of New York; Binghamton, New York 13902-6000 USA.

Leslie Lander (A'86) has a BA in Mathematics from Cambridge and a PhD in Mathematics from Liverpool. He taught in Brazil, Great Britain, Germany, Peru, and Venezuela before moving to Binghamton in 1984, where he is an Associate Professor in Computer Science. His interests include formal aspects software engineering, programming languages and system-level fault diagnosis. He is an Associate Member of IEEE and a Member of ACM.

Dr. Horng-Shing Lu; Inst. of Statistics; National Chiao Tung Univ; Hsinchu 30050 TAIWAN - R.O.C.

Horng-Shing Lu received his MS (1990) and PhD (1994) in Statistics at Cornell University and BS (1986) in Electric Engineering at National Taiwan University. He is an Associate Professor in the Institute of Statistics at the National Chiao Tung University. He has authored papers in statistical computing methods, reliability, and nonparametric function estimation. He is a member of the American Statistical Association and the Institute of Mathematical Statistics.

(*Continued on page* 469)

- Reliability obeys model I
- $R_{L,N}=0.999;\ \gamma = 5\%$
- $n=5$
- $\mu_y=50$ pounds, $D_y=0.1$, $D_x=0.05$.
- An average of 1 units (for $n=5$) fails at the severe test conditions. Thus, at the 95% s-confidence level, $R_{L,S} = \beta_{0.05}(4,2) = 0.342$. ◄

The result is:

$$z_{L,N} = 3.09$$

$$A = 0.34216,\ B = -0.97613,$$

$$\delta = 1.40.$$

The test level is therefore $S = 50 \cdot 1.4 = 70$ pounds.

During the test, 2 out of the 5 units failed. The 95% LCB of the reliability at the test level is:

$$R_{L,S} = \beta_{0.05}(3,3) = 0.189$$

$$\mu_{X,L} = 67 \text{ pounds}$$

$$R_{L,N} = 0.9977 < 0.999 \text{ (the requirement).}$$

5. PRACTICAL APPLICATIONS

The procedure is based on testing a few units at a severe level, rather than testing many units under working conditions. Typical situations are 4 - 5 units without failure in 20% - 30% amplified/attenuated test levels rather than 200 - 300 units without failure under nominal working conditions.

While implementing the procedure, its underlying assumptions should be thoroughly examined & tested to assess their validity in the given problem. The selected test level should not be too far from the working conditions, to prevent a change in the failure mechanism. We recommend a "success-oriented approach" during the selection of the test level so that mainly successes are anticipated. Even though it is not a mathematical restriction, it is not desirable to deduce high reliability at the working conditions based on poor reliability demonstrated at the test level.

REFERENCES

[1] J. Hahn, S. Shapiro, *Statistical Models in Engineering*, 1967, pp 72-74; John Wiley & Sons.

[2] E. Parzen, *Modern Probability and Its Applications*, 1960, pp 237-238; John Wiley & Sons.

[3] G.W. Snedecor, W.C. Cochran, *Statistical Methods*, 1980, p 41; The Iowa State University Press.

[4] N.R. Mann, R.E. Schafer, N.D. Singpurwalla, *Methods for Statistical Analysis of Reliability and Life Data*, 1974, pp 173-174, 289-293; John Wiley & Sons.

[5] F. Clifton, "Strength variability in structural materials", R&M -3654, 1971; Ministry of Aviation Supply, Aeronautical Research Council, London, Her Majesty's Stationary Office.

[6] E.B. Haugen, *Probabilistic Mechanical Design*, 1980; John Wiley & Sons.

[7] A. Mosleh, G. Apostolakis, *Models for the Use of Expert Opinion in Low Probability, High Consequence Risk Analysis* (R.A. Waller & V.T. Covello, Eds), 1981, pp 107-124; Plenum Press.

AUTHORS

Ze'ev Porat; RAFAEL (22); POBox 2250; Haifa 31021, ISRAEL.
Internet (e-mail): moa2@actcom.co.il

Z. Porat has been employed by RAFAEL as a senior reliability engineer since 1976. He spent the years 1986-1988 at Bell Communications Research, Red Bank as a reliability analyst. He obtained his BSc (1976) in Industrial & Management Engineering from the Technion - Israel Institute of Technology, Haifa.

Meir Haim; RAFAEL (22); POBox 2250; Haifa 31021, ISRAEL.

M. Haim has been a reliability analyst at RAFAEL for the past 15 years and has analyzed reliability & safety of complex systems. During this period he spent a year of dealing with reliability research at the University of California, Los Angeles. M. Haim obtained his MSc (1976) in Operation Research from the Technion - Israel Institute of Technology, Haifa.

Yaffa Markiewicz; RAFAEL (46); POBox 2250; Haifa 31021, ISRAEL.

Y. Markiewicz has been an operations research and reliability analyst in RAFAEL since 1968. She obtained her MSc (1974) in Statistics from Stanford University.

Manuscript received 1993 April 15.

IEEE Log Number 92-12071

◄TR►

Bayes Analysis for Fault Location in Distributed Systems

(*Continued from page* 465)

Dr. Martin T. Wells; Statistics Center and Dept. of E&S Statistics; Cornell University; Ithaca, New York 14853-3901 USA.

Martin T. Wells is an Associate Professor of Statistics at Cornell University. He received his PhD (1987) in Mathematics at the University of California. He has authored papers in statistical decision theory, nonparametric statistics, and reliability. He is a member of the American Statistical Association and the Institute of Mathematical Statistics.

Manuscript received 1993 November 2.

IEEE Log Number 92-15603

◄TR►