

Diagnosis of Equipment Failures by Pattern Recognition

Pau, L. F.

Published in:
I E E E Transactions on Reliability

Link to article, DOI:
[10.1109/TR.1974.5215245](https://doi.org/10.1109/TR.1974.5215245)

Publication date:
1974

Document Version
Publisher's PDF, also known as Version of record

[Link back to DTU Orbit](#)

Citation (APA):
Pau, L. F. (1974). Diagnosis of Equipment Failures by Pattern Recognition. I E E E Transactions on Reliability, R-23(3), 202-208. DOI: 10.1109/TR.1974.5215245

DTU Library

Technical Information Center of Denmark

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

- [11] F. M. D'Heurle, "Electromigration and failure in electronics—An introduction," *Proc. IEEE*, vol. 59, pp. 1409-1418, 1971.
- [12] A. Bobbio and O. Saracco, "On the spread of time-to-failure measurements in thin metallic films," *Thin Solids Films*, vol. 17, pp. S13-S16, 1973.
- [13] J. R. Black, "Electromigration—A brief survey and some recent results," *IEEE Trans. Electron. Dev.*, vol. ED-16, pp. 338-347, 1969.
- [14] A. Bobbio, O. Saracco and R. Riva, "Rottura per elettromigrazione delle connessioni di alluminio nei circuiti integrati," *Rendiconti LXXIII Riunione Annuale AEI*, 1972, mem. 3.3.04.
- [15] I. A. Blech and E. S. Meieran, "Direct transmission electron microscope observation of electrotransport in aluminum thin films," *Appl. Phys. Letters*, vol. 11, pp. 263-266, 1967.
- [16] R. Rosenberg, "Value D_0Z^* for grain boundary electromigration in aluminum films," *Appl. Phys. Letters*, vol. 16, pp. 27-29, 1970.
- [17] H. M. Breitling and R. E. Hummel, "Electromigration in thin silver, copper, gold, indium, tin, lead and magnesium films," *J. Phys. Chem. Solids*, vol. 33, pp. 845-852, 1972.
- [18] A. J. English, K. L. Tai and P. H. Turner, "Electromigration in conductor stripes under pulsed DC powering," *Appl. Phys. Letters*, vol. 21, pp. 397-398, 1972.
- [19] I. A. Blech and E. S. Meieran, "Electromigration in thin films," *J. Appl. Phys.*, vol. 40, pp. 485-491, 1969.
- [20] W. Jost, "Diffusion in solids, liquids, gases." New York: Academic Press, 1952, p. 25.

Mailing address:
 A. Bobbio
 A. Ferro
 O. Saracco
 Istituto Elettrotecnica Nazionale
 Galileo Ferraris
 10125 Torino
 Italy

Diagnosis of Equipment Failures by Pattern Recognition

L. F. PAU

ABSTRACT—The main problems in relation to automatic fault finding and diagnosis in equipments or production systems are discussed:

- 1) compression of the syndrome and observation spaces for better discrimination between failure modes;
- 2) simultaneous display of the failure patterns and the failure instants, for maintenance control and review of the reliability design;
- 3) automatic production of a final set of diagnosis assumptions classified according to their probabilities.
- 4) sequencing of the inspections in accordance with the failure rates and inspection costs.

READER AIDS:

Purpose: Widen state of the art

Special math needed for explanations: Probability

Special math needed for results: Same

Results useful to: Reliability and systems engineers

INTRODUCTION

The paper is concerned with automated diagnosis of generic systems which cannot be modeled by a logical network, and have only a small number of identifiable failure modes, given only the input and output evolutions. In other words, the analysis is based on representations of causality relations between the failure causes and the corresponding operating performances. This kind of approach is therefore specifically useful, for example, for electromechanical equipments, engines, and simple mechanical parts.

The purpose of any diagnostic searching procedure is to recognize a failure mode, or a set of failures (also called syndrome), on the basis of numerical information about the circumstances of the failure, results of non-destructive tests,

and information about the past history of the system (including maintenance and operational utilization); all these data determine the failure pattern vector.

The actual diagnosis will be made by comparing this failure pattern vector with known pattern vectors found to characterize the failure modes considered, and called learning patterns. The learning patterns originate from systems for which the failure modes were found by maintenance personnel, and we assume that they are all stored in a frequently updated reliability and maintenance data bank. Comparing the observed symptoms (described by the failure pattern vector) with the learning data, and classifying the ill-working system into one of few classes of possible failure modes, can therefore be formalized as a pattern recognition problem. Some research about this has been reported since 1968 by Becker [1], Cortina [2], Hankley and Merrill [3], Page [4], and Pau [5].

In this paper, we treat some major problems in relation to automatic diagnosis. First, we discuss a compression method of the observations about each system, for better discrimination between failure modes and to eliminate all redundant tests and observations; this is the so-called feature extraction problem. In Section 2, we explain how to display simultaneously the failure patterns and the failure instants, for maintenance control or review of the reliability design of the system. Section 3 deals with the production of a satisfactory list of a few most-relevant diagnostic assumptions, classified according to their probabilities. Lastly, in Section 4, we give some rules for sequencing the inspections by the element-by-element method, in accordance with the diagnostic assumptions, failure rates, and inspection costs.

Manuscript received October 2, 1973.

1. COMPRESSION OF THE LEARNING DATA AND MEASUREMENT SELECTION

1.1 Definition of the Learning Data $k(I, J)$

We basically assume that it has been possible to initiate a good data collection system, monitoring the systems considered over time, and also their main operating parameters in service until some observable functional-failure happens. The implementation of such a data bank is a difficult problem, extensively discussed in Pau [6]. Let I, J respectively be (at a given date) the set of measurements and the set of pattern vectors describing the failures of all systems of the same type. The learning data $k(I, J) = \{k(i, j) \geq 0 \mid i \in I, j \in J\}$ are either qualitative or quantitative; and can then be defined as follows: $k(i, j)$ is the value of measurement i concerning the system j ; i.e., time since last overhaul TBO_j , voltage at some point, or any binary parameter indicating that a given subsystem was switched on. Some measurements i may also be made binary if their intervals of variation have been sampled into smaller intervals.

The past experiments have shown that misleading conclusions may be drawn from once-compressed learning data, either because of too-small learning samples, or because of careless reports. Methods to account for these phenomena have been thoroughly tested in practice and justified theoretically [5]. Other transformations may be made, in order to study the learning data from specific viewpoints:

a) the learning data $k(I, J)$ are called explicit if one of the sets I, J designates equipments, and the other, observations about these equipments as previously defined.

b) the learning data $k(I, J)$ are called implicit if both sets I, J are observations; such a tableau is deduced from an explicit tableau by aggregating some observations with respect to all equipments, or by classifying all equipments with respect to some observations. For example, if J has become the set of TBO intervals, we may define: $k(i, j)$ is the number of learning equipments having had a failure in the TBO -interval j , and on which the symptom i was present. Design review generally uses learning data in the implicit form, while automated diagnosis uses the explicit form.

1.2 Feature Extraction by the Means of Correspondence Analysis

Assume that the tableau $k(I, J)$ of non-negative numbers is given. The feature extraction procedure used herein is a special form of principal component analysis, characterized by the following additional properties demonstrated in Benzecri [7], Pau [8], Lebart and Fenelon [9]:

no prior hypothesis is made about the nature of the elements in the sets I (failures) and J (observations or equipments), and all interactions are considered; in other words, we do not care for the labels in the sets $\{I\}$ and $\{J\}$.

all elements in both sets $\{I\}$ and $\{J\}$ can be displayed simultaneously in the same reduced feature space, because they play symmetric roles in a tableau; in this reduced pattern space, the Euclidean distance between any two elements

of I - J , I - I , or J - J , is an overall statistical measure of the correspondence between these elements, independently of all scale effects.

the “varimax” principle used to find orthogonal factors is not applied to the raw data in $k(I, J)$, but to quantities deduced from a contingency table $p(I, J)$:

$$p(i, j) = k(i, j) / \left(\sum_{l \in I, m \in J} k(l, m) \right)$$

with estimated marginal probability density functions (pdf):

$$P^I = p(i, \cdot) = \sum_{j \in J} p(i, j) \mid i \in I \quad \Pr\{j \mid i\} = p(i, j) / p(i, \cdot)$$

$$P_J = p(\cdot, j) = \sum_{i \in I} p(i, j) \mid j \in J \quad \Pr\{i \mid j\} = p(i, j) / p(\cdot, j)$$

Correspondence analysis may then be summarized as follows:

a) The metric on I is the distance function d_I , while the metric on J is the distance function d_J :

$$d_I^2(i_1, i_2) = \sum_{j \in J} [\Pr\{j \mid i_1\} - \Pr\{j \mid i_2\}]^2 / p(\cdot, j)$$

$$d_J^2(j_1, j_2) = \sum_{i \in I} [\Pr\{i \mid j_1\} - \Pr\{i \mid j_2\}]^2 / p(i, \cdot)$$

The weighing factors, such as $1/p(i, \cdot)$ are introduced in order to compensate for cases like the following; an observation i may always be related to large conditional probabilities $\Pr\{i \mid j\}$ mainly because $p(i, \cdot)$ is large, and the differences between $\Pr\{i \mid j\}$ values will have an excessive importance when comparing a system j_1 with a system j_2 .

b) The element i has $\text{Card}(J)$ coordinates $(\Pr\{j \mid i\}, j = 1, \text{Card}(J))$; the element j has $\text{Card}(I)$ coordinates $(\Pr\{i \mid j\}, i = 1, \text{Card}(I))$. These coordinates are called profiles of i and j resp. The element i has the weight $p(i, \cdot)$, while j has the weight $p(\cdot, j)$.

c) Let the constant r be given such that $r \leq \text{Inf}\{\text{Card}(I), \text{Card}(J)\}$. We want to minimize, in the sense of the d_I or d_J metric, the dependence between I, J defined as the norm-value $\|p(I, J) - P^I \otimes P_J\|^2$. If this quantity had been zero, then the sets $\{I\}$ and $\{J\}$ would have been independent in the probabilistic sense and the observations about each system would have been independent of the system considered. But our goal is here to eliminate all redundancy in the data, and to find as many aggregate observations as possible (also called features) for which this independence holds, and r features for which dependence holds so that discrimination can be done using only these. It can be shown that the r -dimensional vector basis of basic features which minimizes this dependence after transforming $\text{Card}(I)$ —or $\text{Card}(J)$ —dimensional patterns in $k(I, J)$ into r -dimensional vectors, can be constructed as follows [10].

for $\{I\}$, the r base vectors $f_i, i = (1, r)$ are the r first principal axes of inertia of the solid body made of the discrete $\text{Card}(J)$ dimensional elements $i \in I$ having the weights $p(i, \cdot)$;

this inertia is computed for the d_I distance; let $\lambda(f_l)$ be the inertia of axis $f_l, l = (1, r)$ ordered by $\lambda(f_1) \geq \lambda(f_2) \geq \dots \geq \lambda(f_r)$; the f_l 's are normed to the unit length with respect to d_I . And $f_l(\lambda_l)$ is the $(l + 1)$ 'st eigenvector (resp. eigenvalue) of the $S = [s_{j_1 j_2}]$ matrix:

$$s_{j_1 j_2} = \sum_{i=1, \text{Card}(I)} p(i, j_1) p(i, j_2) / p(i, \cdot) [p(\cdot, j_1) p(\cdot, j_2)]^{1/2}$$

for J , we have equivalent definitions and relations for the r basis vectors $g_l, l = (1, r)$. $f_l, g_l, l = (1, r)$ are here row vectors, i.e., linear mappings.

d) The coordinates of the learning patterns projected into the r -dimensional feature space, are computed as follows:

for $\{I\}$, the feature $l = (1, r)$ of learning pattern $i \in I$ on the axis f_l , originated in the center of inertia of all elements in $\{I\}$, is

$$G(i, l) = f_l \cdot [\text{Pr}\{j|i\}, j = 1, \text{Card}(J) \text{ vector}] \quad (1)$$

for $\{J\}$, the feature l of learning pattern $j \in J$ on the axis g_l , originated in the center of inertia of all elements in J , is

$$F(j, l) = g_l \cdot [\text{Pr}\{i|j\}, i = 1, \text{Card}(I) \text{ vector}] \quad (2)$$

e) It can be shown that $\lambda(f_l) = \lambda(g_l), l = (1, r)$, and that it is sufficient to compute either the f 's or the g 's because:

$$g_l = \frac{1}{\sqrt{\lambda(f_l)}} f_l \begin{bmatrix} \text{Pr}\{j|i\} & i = \text{column} \\ & j = \text{row} \end{bmatrix} \quad (3)$$

$$G(i, l) = \sum_{j=1, \text{Card}(J)} F(j, l) \text{Pr}\{i|j\} / \sqrt{\lambda(f_l)}, l = 1, r$$

Consequently, it is equivalent to display simultaneously all elements of I and J in either the f_l space or the g_l space, $l = (1, r)$

$$f) p(i, j) = p(i, \cdot) p(\cdot, j) (1 + \sum_l F(j, l) G(i, l) \sqrt{\lambda(f_l)})$$

2. DISPLAY OF THE LEARNING DATA, AND APPLICATIONS TO DESIGN REVIEW AND MAINTENANCE CONTROL

In this section, we are only concerned with implicit learning data $k(I, J)$ where $\{I\}$ and $\{J\}$ are two different sets of observations.

2.1 Interpretation Procedure

Using correspondence analysis as explained in the previous section, we can select $r = 2$ and thereby find the two best features discriminating the observations J and the observations I . According to Section 1.2.d), all observations can thus be displayed into the (f_1, f_2) plane which contains the largest part of the dependence between I and J , namely $(\lambda(f_1) + \lambda(f_2)), \lambda(f_1) \geq \lambda(f_2)$. Such a 2-dimensional representation of all learning data $k(I, J)$ will be called a map, and any pair of vectors f_l, f_m yields such a map of weight $(\lambda(f_l) + \lambda(f_m))$. These maps can be used as follows:

a) If any two observations i_1, i_2 in I are conditionally associated in the same way to all observations in J , then by definition of d_I , the corresponding features on a map will be identical. Thus, if two failure modes i_1, i_2 are represented almost by the same point in the feature space, then one of the following 2 statements is true: one of these modes is redundant, e.g., the maintenance instructions require stating the second failure mode each time the first failure mode is observed; the failure mode i_1 can systematically be the main cause of the failure 1 mode i_2 , or conversely.

b) The product $p(i, \cdot) G(i, l)$ is also called contribution of the observation i to the feature l . This notion is strongly related to the correlation coefficient between i and f_l as used in principal component analysis. This remark may help in finding the interpretation of f_l . All observations i of I mapped in a small neighborhood of the origin of the axes (f_l, f_m) on the map, are therefore only slightly correlated with the features l and m .

c) If, on such a map, two observation points i_1, i_2 are more or less close, in the sense of the Euclidean d -distance between them, it means that these observations are more or less strongly associated. This association degree is here taken in the sense of the d_I or d_J similarity measures between profiles.

Once a close association has been detected, e.g., between a pressure measurement i_1 and a maintenance operation i_2 , then the maintenance department will have to give technical reasons for this, or to demonstrate why it is meaningless. Because all elements of I and J can be displayed simultaneously on the same maps with the same length units on the axes [7, 8] one can also detect associations between observations j_1, j_2 in J , or even between an element of I and an element of J . The interpretation procedure is the same as for associations between observations in I . If i is d -close to j , this means that in general the measurement i has a higher conditional probability with respect to j than in the mean with respect to the other observations in J (and the same for j with respect to i).

A careful and systematic analysis of the geometric proximities between elements of I and/or J ought therefore to lead to a list of suspected causality relations. This proximity analysis can also be done between clusters of points on maps of decreasing weights, in order to identify causality relations between syndromes or observation complexes.

The main goal of this interpretation, which can be sustained with the computation of some confidence levels [5], is to draw attention to causality relations among failures, maintenance, modifications, operating conditions, and times. It has been applied to both electronic and mechanical airborne equipments in order to detect systematic coding errors, criticize maintenance operations and scheduling, criticize design parameters, or select sensitive tests during accelerated field trials [10-12]. Our view is that the learning and interpretation phase must be conducted in parallel with the analysis of experts' special reports, in order to compare them.

Another field of active research is the analysis of test data

for electronic components, in order to relate the failure modes with technological design parameters; the main application of this is in component selection for specific environments [10, 13, 14].

2.2 Example

As an example, we will interpret some associations between observations and/or time intervals in Fig. 1 relative to an airborne fuel pump. The compressed learning data were initially of the implicit type explained in Section 1.1.

Analyzing the proximities between TBO- (Time Between Overhauls) intervals $j \in J$ and the failure modes PPR, PHU, we notice that

the wear-out of the electrical brushes is strong between 200 and 800 hours, small to medium up to 4500 hours if no failure has occurred and exceptionally low after that time.

oil leakage failures are explicitly discriminated as being the most important early failures between 0 and 50 hours;

pressure-drop failures happen essentially between 200 and 800 hours and exceptionally around 100 or 1500 hours in relation to an important wear-out of the electrical brushes.

It can readily be inferred from the above, and a more detailed analysis of the same, that the mean time between unscheduled repairs for these pumps could be improved towards the nominal value of 1700 hours, by eliminating the pressure drop failures PPR through an improved design. These are also most strongly associated with the positions P_3 and P_4 , even P_1 , on the aircraft.

3. REAL TIME DIAGNOSIS BY PATTERN RECOGNITION

The goal of this section is to present a method which, given all observations I on a system, produces automatically a list of the most probable failure modes of this system. The pattern recognition approach used includes first the learning stage, next the real time feature extraction, and lastly the classification procedures wherein the r features characterizing the failed system are compared to the compressed learning data.

3.1 Learning Stage

The learning patterns $j \in J$ are defined as being a large number of systems of the same type for which the failure mode has been determined by the quality control or maintenance personnel. The learning data $k(I, J)$ are obtained by gathering all information $i \in I$ (including times between overhauls, etc.) about these learning patterns. We assume that the total number of different failure causes $d \in D$ is small with respect to the total number of equipments observed. Assume that the probability distribution $\{\pi(d) | d \in D, \pi(d) > 0, \sum_D \pi(d) = 1\}$ of the failure causes has been estimated within the learning data or by other means.

The learning features, which will be used during the recognition phase, are the images of the learning patterns in a reduced feature space having a fixed dimension r . The feature extraction procedure used is the correspondence

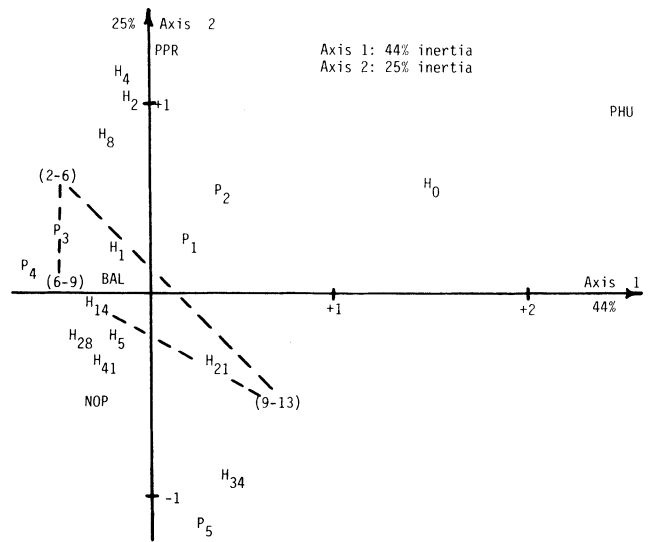


Fig. 1. Symbols.

- H_n Sampling interval for the operations time since last overhaul, starting at $100 \times n$ hours, and ending at $100 \times p$ hours (p smallest integer following n in the list: 0, 0.5, 1, 2, 4, 8, 14, 21, 28, 34, 41).
- BAL Single measurement of the length of the electrical brushes, in mm, when grounding the pump.
- NOP Binary information, equal to 1 if no failure.
- PPR Binary information, equal to 1 if pressure-drop failure.
- PHU Binary information, equal to 1 if oil-leakage failure.
- P_i Position occupied by the pump on the aircraft before grounding $i = 1, 2, 3, 4, 5$.
- 2-6 Binary information, equal to 1 if brush length BAL between 2 and 6 mm.
- 6-9 idem, 6-9 mm.
- 9-13 idem, 9-13 mm.

analysis of paragraph 1.2 applied to the tableau $k(I, J)$, of the explicit type.

3.2 Real Time Feature Extraction for a Failed Equipment

Assume that troubleshooting has just been observed on an equipment \bar{j} of the type investigated in Section 3.1 and it has been possible to gather all information $\{k(i, \bar{j}), i \in I\}$ about \bar{j} . We can consider \bar{j} as a supplementary learning pattern belonging to an unknown class $d(\bar{j})$. But since the feature extraction of the learning data $k(I, J)$ is made without taking care of the knowledge of $d(J)$, we assume that \bar{j} does not perturb the learning features too much, and map the equipment \bar{j} into the feature space by (2). Thus the coordinate of equipment \bar{j} on the l -th feature axis g_l is

$$F(\bar{j}, l) = g_l \cdot \left[\frac{k(i, \bar{j})}{\sum_{i \in I} k(l, \bar{j})} \right], i = 1, \text{Card}(I), l = (1, r). \quad (4)$$

Studying the maps discussed in Section 2, we then look for instructing associations of \bar{j} , either with the learning equipments $j \in J$ for which $d(j) \in D$ is known, and/or with the observations $i \in I$. These associations help substantially in formulating some precise experimental hypothesis about the mechanism of the failure detected in Section 3.3 on the equipment \bar{j} .

3.3 Diagnosis and Classification Procedure

We will consider one single classification procedure, placing each equipment \bar{j} into the most probable failure mode class $d(\bar{j}) \in D$. Though, if the corresponding probability of correct classification is below some given threshold, then the equipment \bar{j} is rejected and no diagnosis can be formulated. The generalized nearest neighbor rule is used as indicated (and treated theoretically) in Patrick and Fisher [15], Pau [10].

$$\bar{j} \text{ has failure mode } d(\bar{j}) \in D \Leftrightarrow \varphi_{d(\bar{j})} \pi_{d(\bar{j})} = \max_{d \in D} \{\varphi_d \pi_d\},$$

and

$$\varphi_{d(\bar{j})} \geq \alpha > 0$$

$$\bar{j} \text{ cannot be diagnosed} \Leftrightarrow \max_{d \in D} \{\varphi_d\} < \alpha$$

where

$$\varphi_x \equiv n_x / [(N_x + 1)V_x]$$

- d class of failure causes selected in D .
- N_d number of learning patterns $j \in J$ belonging to the class $d \in D$.
- π_d estimated probability of occurrence of a random failure mode d in the set D , as introduced in 3.1.
- n_d integer parameter, determined for each class d .
- V_d minimal volume of a neighborhood of the newly observed equipment \bar{j} in the r -dimensional feature space, so that $(n_d - 1)$ learning patterns of the class $d \in D$ are interior to this neighborhood, while one single learning pattern is on the boundary thereof.
- α given confidence threshold.

This classification can briefly be explained as follows: it can be shown that φ_d is an asymptotically unbiased distribution-free estimator of the pdf for the class d of patterns in the feature space at the location of \bar{j} found in Section 3.2. α is then the threshold pdf required to formulate a diagnosis. Maximizing the classification gain by a Bayes decision rule introduces the weights π_d ; the equipment \bar{j} is then classified, if possible, into the class d having the largest $\pi_d \varphi_d$.

An ε -neighborhood of \bar{j} in the set $\{J\}$ of learning patterns, is defined as follows in the r -dimensional feature space:

$$v(\varepsilon, \bar{j}) \equiv \left\{ j \in J \mid \left(\sum_{l=1}^r |F(j, l) - F(\bar{j}, l)| \right) < \varepsilon \right\}$$

These neighborhoods are related to the L^1 -metric, instead of the classic Euclidean L^2 -metric; this modification reduces the classification computation time and yields correct classification rates which are at least as good as for the L^2 -metric [16].

The next best suspected failure mode $d^1(\bar{j})$ is the one maximizing $\varphi_d \pi_d$ for $d \in (D - d(\bar{j}))$; this process can be repeated, and a ranked list established over the most probable alternate failure modes $d(\bar{j})$, $d^1(\bar{j})$, etc. Though it is clear that this procedure would be misleading if the actual failure had not been included in the catalog D ; if the result $d(\bar{j})$ happens frequently to be absurd, one has to examine thoroughly the

learning data and the set D of alternate failures.

In Pau [10], this procedure is made sequential in order to find a compromise between a short diagnostic computation time, and a high mean true diagnostic probability.

In the next section, we will show how this method can be used for failure localization purposes in a modularized system; meanwhile, an example will be given where the object is to find some well-defined failure modes.

3.4 Example of Automated Testing

We have considered a stationary fabrication process of complex electromechanical systems with very stringent specifications and small dimensional tolerances. The 82 observations on each equipment $j \in J$ in the process were the measurements made by the quality control department at the input of the process, and the operational characteristics of the machine tools when used on each specific system (setting, cumulated time of operations, time since servicing, type of tool, air flow, temperature, oil flow, rotation speeds, workers operating the machines, ...). The 21 classes of failures included the special class d_0 of all equipments for sale fulfilling all quality control requirements. In the following, $\alpha = 0$.

a) During the learning phase, data were collected on Card (J) = 2000 items (20 days of production), among which 800 non-acceptable items were identified at the final quality control and where each received a diagnosis (chosen among the 20 classes of failures). These learning data were processed on a general purpose IBM 370-65 computer (12 min, CPU). The computation of the $f_{jl} = (1, r)$ ran into some numerical diagonalization difficulties. Through the review process described in Section 2.1, it became possible to pinpoint those systematic aspects of the production process having indirectly the strongest contributions to the named failures, in this case the oil flows.

b) During the testing phase, a true recognition rate of 92% was achieved for the items classified into the class d_0 by the final quality control, still working. The mean true diagnosis rate for the 20 types of actual failures was 81% when $r = 10$; mean unitary diagnostic computing time: 0.46 sec.

c) During the operational phase under final implementation, all 82 observations will be monitored in real time for each equipment in the production line; most nondestructive tests and the final quality control will be suppressed. A few specialists will play a supervisory role for the automatic diagnostic system, including the small online data-logging and computing unit. Considerable economic benefits can be obtained, as evaluated on the basis of the resting phase b). These specialists will perform design reviews, and modify and enlarge the learning data bank.

4. SEARCHING FOR A FAULTY MODULE IN A MODULAR SYSTEM

The subject of this section is the localization of a single faulty module on the basis of the results of the automated diagnostic procedure of Section 3.3. The problem is more precisely to find a search sequence minimizing the expected

cost (or time) to complete the fault localization, given a suspected failure mode $d(\bar{j})$.

It is assumed that systems of the type previously treated are made of M statistically independent modules, which can be inspected by module-by-module tests. We assume moreover that given a failure mode $d(\bar{j})$, the prior probabilities $R_m(d(\bar{j}))$, $m = 1, M$ of failure of each of the M modules are known; each system can only have one single faulty module corresponding to the failure mode $d(\bar{j})$. Let τ_m be the prior known cost (or time) required to conduct all tests/observations on the module number m .

Define the state of the system by one single faulty module m included in a subset S_m of not yet inspected modules, which are all assumed failure-free except m ; the probability of this state of the system is then:

$$R_m^* = \frac{R_m(d(\bar{j}))}{1 - R_m(d(\bar{j}))} / \sum_{q \in S_m} \frac{R_q(d(\bar{j}))}{1 - R_q(d(\bar{j}))}.$$

The search procedure is then the following, starting with $m = 1$:

Step m: $(m - 1)$ modules have previously been tested/observed and found not-faulty. S_m is the subset of all remaining modules not yet tested.

*select, according to one of the rules 1-3, an element of S_m , and perform all tests/observations on this module.

*if this module is faulty, then stop the search; if not, then go to next step $m + 1$.

Rule 1 $\max_{l \in S_m} \{R_l^*/\tau_l\}$ Gluss [17]

$$2 \max_{l \in S_m} \left\{ \frac{-R_l^* \ln R_l^* - (1 - R_l^*) \ln (1 - R_l^*)}{\tau_l} \right\} \text{ Pashkovskiy [18]}$$

$$3 \max_{l \in S_m} \left\{ \frac{R_l(d(\bar{j}))}{\tau_l} \right\} \text{ Pashkovskiy [18]}$$

The rules 1-3 yield quasi-optimal inspection sequences, the efficiency of which may be compared using statistical decision theory [19]. In the actual cases investigated, the best rules appear to be 1 and 2; more generally the idea of decomposing the diagnostic procedure into first finding the failure mode of highest probability $d(\bar{j})$, and next the faulty module, proves to be very efficient. In those environments where automated diagnosis may be required, it is important only to activate the diagnosis and shift over to some redundant system in parallel, when a functional failure is actually observed on the initial system. It is not always necessary to localize instantly the failure, and the search subprocedure may eventually be postponed; on the basis of the results of this subprocedure, it will only be necessary to put down and repair the faulty module. In those operational applications where the parameter estimation problems were not dominant, savings of 15-20% on the total maintenance and diagnosis times were observed.

5. CONCLUSION

Our pattern recognition techniques can help make main-

tenance more efficient, and make the designers understand the complex relationships between the operational environment, the production control, and other internal or external factors. They are probably the only approach to diagnosis in mechanical and non-purely electronic equipments. The results are the more realistic and profit-earning if the individual equipments are numerous and have to be monitored during their whole life. But the success relies finally upon the quality of the data records, and upon the reliability of the monitoring sensors used.

REFERENCES

- [1] P. W. Becker, Recognition of patterns, Copenhagen, Denmark, Polyteknisk Forlag, 1968.
- [2] E. Cortina, H. L. Engel, W. K. Scott, Pattern recognition techniques applied to diagnostics, SAE report 7000 407, Society for Automotive Engineers, 1970.
- [3] W. J. Hankley, H. M. Merrill, A pattern recognition technique for system error analysis, *IEEE Trans. Rel.*, vol. R-20, Aug. 1971.
- [4] J. Page, Recognition of patterns in jet engine vibrations signals, IEEE Publ. no. 16 c 51, pp. 102-105.
- [5] L. F. Pau, Diagnostic statistique, ENS de l'Aéronautique, Toulouse, 1st edition 1969, 3rd edition 1973.
- [6] L. F. Pau, Analysis and diagnosis of in-service failures, in G. Weber (Editor) Einführung in Methoden und Probleme der Zuverlässigkeit, KFK Report 1811, Kernforschungszentrum Karlsruhe, Karlsruhe 1973.
- [7] J. P. Benzecri, Distance distributionnelle et métrique du CHI-2, mimeograph, ISUP, Paris University 7, 1970.
- [8] L. F. Pau, Méthodes statistiques de réduction et de reconnaissance des formes, thesis, Paris-Orsay University, p. 101, May 1972.
- [9] L. Lebart, J. P. Fenelon, Statistique et informatique appliqués, Dunod, Paris, 1971.
- [10] L. F. Pau, Applications of pattern recognition to the diagnosis of equipment failures, in: T. Einsele, W. Giloi, H. H. Nagel (Editors), Cognitive Verfahren und Systeme, Lecture notes in economic and mathematical systems no. 83, Springer Verlag, Berlin, 1973, pp. 291-306.
- [11] L. F. Pau, Diagnostic statistique: synthèse des informations relatives à fiabilité et à la maintenance d'un matériel aéronautique, L'Aéronautique et l'Astronautique, no. 34, pp. 69-76, 1972.
- [12] F. N. Pokrowsky, On reliability prediction by pattern classification. Proc. 1972 Annual Reliability and Maintainability Symposium, pp. 367-375, 1972. (Available from IEEE.)
- [13] R. Goarin, Application de l'analyse des correspondances à l'étude de la fiabilité des composant électroniques, Congrès national de fiabilité, edited by Centre National d'étude des Télécommunications, Paris, pp. 179-197, 1972.
- [14] P. C. Andersen, Applications of pattern recognition to the diagnosis of failures, IMSOR, Technical University of Denmark, Lyngby, 1973.
- [15] E. A. Patrick, F. P. Fischer, A generalized k-nearest neighbor rule, Information and Control, vol. 16, pp. 128-152, April 1970.
- [16] L. F. Pau, Optimisation d'une métrique en reconnaissance des formes, Journées d'étude GALF, Université libre de Bruxelles, Bruxelles, May 23-25, 1973.
- [17] B. Gluss, An optimum policy for detecting a fault in a complex system, Operations Research, vol. 7, no. 4, 1959.
- [18] G. S. Pashkovskiy, Optimization of sequential fault detection procedures, Engineering cybernetics, vol. 9, March-April 1971, pp. 259-270.
- [19] E. E. Scheufens, Optimal lokaliseren af en enkelt fejl i et system, in L. F. Pau (Editor): Topics in pattern recognition, IMSOR, Technical University of Denmark, Lyngby 1973, p. 19.
- [20] H. Y. Chang, E. G. Manning, G. Metze, Fault diagnosis of digital systems, Wiley, N.Y., 1970.
- [21] J. Chinal, La logique des pannes, E.N.S. l'Aéronautique, Toulouse, 1973.

[22] J. De Corlieu, *Maintenabilité et vie des systèmes*, E.N.S. de l'Aéronautique, Toulouse, 1972.

Mailing address:

L. F. Pau
 IMSOR, Technical University of Denmark
 Building 349
 Lundtoftevej 100
 DK-2800 Lyngby, Denmark

Louis-François Pau was born in Copenhagen, Denmark, on May 29, 1948. He received the Master's degree in mathematics from the Paris University in 1969, while studying at the Ecole Nationale Supérieure de l'Aéronautique where he obtained the degree of engineer in avionics. He received the

"docteur ingénieur" degree from Paris University in 1972 in the field of statistical pattern recognition.

Dr. Pau is presently Associate Professor of operations research and mathematical statistics at the Technical University of Denmark. He has been a professor of pattern recognition at the Ecole Nationale Supérieure de l'Aéronautique since 1970. From 1969 to 1972 he worked mainly on military research projects, and as a consultant for Air France in the field of diagnosis of failures. His fields of interest are diagnosis of failures, speech recognition, pattern recognition theory, feature extraction, as well as differential games. In 1970, he was called for 3 months by the McDonnell Douglas corporation, working on the DC-10 program.

Dr. Pau has published papers in several scientific and technical journals in France and Scandinavia; he holds several patents. He is a member of the French and Danish operations research societies, and has received two scientific prizes in France.

Analyzing the Interface of Reliability and Economics of Unmanned Satellites

H. W. VON GUERARD

ABSTRACT—This paper deals primarily with the reliability of unmanned scientific or commercial satellites. In the course of reliability analysis a hierarchy of reliability predictions is generated, reaching from preliminary apportionment by subsystems: to figures of merit, usually obtained by Boolean algebra; to results based on failure modes, effects and criticality analysis (FMECA); to data from an elaborate, though still static, fault tree; and eventually to data depicting the dynamics of mission time line and contingency analyses.

The outlined routine has been developed, mainly in the US, to quite some perfection, but it still leaves unanswered the customer's most natural concern: how much reliability, within what time, at what price? Therefore, an attempt is made to schedule the above work elements, so that the amount of reliability as well as its allocation become subject to the classical criterion of marginal utility. It is fully realized that there are considerations, such as design freeze partly due to long-lead items, that put a limit to this endeavor.

The postulated analysis consists, in essence, of two models in sequence: first, the fault tree, which exhibits the technical failure modes, their probability of occurrence, and their interdependence. Second, the mission failure model, referring to total and partial deficiencies in performing utility functions, and to their interactions as regards the degree of performance.

Straightforward analysis by the concept of marginal utility has not always proven feasible; there may be prerequisites that are difficult, if not impossible to fulfill, such as true cost allocation to interdependent mission components (e.g., geographical coverage, timing, bandwidth, amplification); or the need for game-theoretic evaluation of the economic risk in a competitive environment. These arguments may be valid if one were considering satellite utility as such; as long, however, as this task is restricted to local allocation of incremental reliability, the problem becomes manageable not only in principle, but with reasonable prospect of fruitful results.

READER AIDS:

Purpose: Widen state of the art

Special math needed for explanations: Probability, utility theory

Special math needed for results: Same

Results useful to: Reliability engineers, system analysts

I. ON THE SIGNIFICANCE OF SATELLITE RELIABILITY ANALYSIS

This paper will focus on the ways the content of a reliability program is shaped by economic factors, superimposed upon technical and operational considerations relating to unmanned satellites. The principal attribute of program utility, i.e., cost effectiveness, is served by selecting for implementation only those elements offering high incremental value return per unit cost.

The topic makes a system analyst's dream come true: here is the opportunity to tackle various clear-cut problems which in other work areas are usually only vaguely defined, or so ill-structured that no significance can be attached to a quantitative analysis. To be more specific: this refers to the not-so-familiar pattern of asking for optimum reliability vs. cost vs. weight trade-offs, — a good question which often suffers from lack of a definite objective function.

Several different features of unmanned spacecraft projects contribute to the more optimistic view of the matter:

The satellite forms a relatively self-contained system with little interference from inside (e.g., from human operators) or from outside (e.g., from the environment).

Furthermore, human life is not at stake; hence reliability becomes realistically negotiable.

The complexity of the system is such as to require a major analytic effort. The challenge is sufficient to justify the attention of the serious minded systems analyst.

Payout of the spacecraft system is a measurable quantity. Benefits and shortcomings can be expressed in a common scale of monetary units. Thus the overall tradeoff analysis becomes feasible.