Journal of Dynamic Systems, Measurement and Control

Technical Briefs

Sequential Testing for Fault Detection in Multiply-Redundant Systems

Asok Ray[1]

The paper presents the theory and application of a sequential test procedure for fault detection and isolation (FDI). The test procedure is suited for development of intelligent instrumentation in strategic processes like aircraft and nuclear plants where redundant measurements are usually available for individual critical variables. The algorithm of the test procedure is formulated by use of (1) a generic redundancy management procedure which is essentially independent of the fault detection strategy and measurement noise statistics, and (2) a modified version of sequential probability ratio test (SPRT) algorithm for fault detection and isolation, which functions within the framework of the aforesaid redundancy management procedure. The sequential test procedure is suitable for real-time applications using commercially available microcomputers and its efficacy has been verified by on-line fault detection in an operating nuclear reactor.

1 Introduction

Control systems for strategic facilities such as aircraft, spacecraft, and hazardous chemical and nuclear power plants require intelligent instrumentation for coordination of plant monitoring, fault diagnostics, and decision making [1, 2]. For fault detection and isolation (FDI), the instrumentation system can be designed to accommodate, for each essential plant variable, redundant measurements that comprise both sensors (possibly of different accuracies) as well as analytically derived measurement(s) [3-6]. To this effect a redundancy management procedure which makes use of all available measurements namely, sensor signals and analytic redundancy [7, 8], have been recently reported by Ray and Desai [5]. The procedure is essentially independent of the measurement noise statistics and the FDI technique in which it is applied.

The paper presents a test procedure which follows a modified version of the sequential probability ratio test proposed by Chien [9, 10] and is designed in the framework of the aforementioned redundancy management procedure [5]. The test procedure is applicable to multiply-redundant systems where the measurements may be vector quantities (such as the velocity or acceleration of an object in space) or scalar quantities (such as the thermal power or coolant temperature of a nuclear reactor). The procedure has been implemented by use of a commercially available microcomputer and its efficacy has been verified on-line in an operating nuclear reactor. The paper is an extension of the author's previous work [5] with the following new contributions:

- Development of a variable-sample sequential test algorithm for on-line fault detection and isolation (FDI).
- Demonstration of the efficacy of the above test algorithm by experimentation in an operating nuclear reactor.

The paper is organized in five sections. Section 2 provides the essential concepts of sequential testing. Section 3 summarizes the theory of the redundancy management procedure reported in [5], and a brief discussion on how to apply this procedure for sequential tests. Section 4 addresses the real-time experimentation of the sequential test procedure at the nuclear research reactor MITR-II. The summary and conclusions of the paper are presented in Section 5.

2 Concepts of Sequential Testing

FDI decisions can be made from observations derived from either single samples or the time history of mutiple samples. The decisions based on single samples, i.e., decisions which disregard the past performance, are reliable only if the magnitudes of the errors in the affected measurements are large in comparison to the measurement noise and uncertainty. For example, gradual and moderate degradations in sensors such as calibration errors for long-term operations may not be reliably detected by single-sample decisions without incurring unacceptable probabilities of false alarms. Under these circumstances, FDI decisions should be made on the basis of multiple observations which make use of the cumulative information provided by the measurement history from the past and current samples rather than relying solely on the current sample. A sequential decision-making procedure needs to be formulated to achieve this goal. Use of both fixed and variable sample approaches are appropriate for sequential testing [5, 6, 9, 10]. The variable sample approach could provide an optimal decision rule for fault detection [11, 12] whereas the fixed sample approach is usually simpler in computation although not necessarily optimal and was adopted in our earlier work [5, 6]. The paper reports the development and real-time implementation of a variable sample sequential test procedure.

In general an optimal decision rule minimizes a composite cost function consisting of the weighted sum of opposing requirements such as the probability of false alarms adjoined with the delay incurred in detecting the fault [11, 12]. Wald's sequential probability ratio test (SPRT) [11] is optimal in the sense that the expected value of the number of samples required for making a decision between two hypotheses, whether the system is in the normal or degraded mode, is minimum for specified probabilities of incorrect decisions.

[1] Mechanical Engineering Department, The Pennsylvania State University, University Park, PA 16802.

Contributed by the Dynamic Systems and Control Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS. Manuscript received by the Dynamic Systems and Control Division July 27, 1987; revised manuscript received May 1988. Associate Editor: G. Young.

Wald's test is devised on the assumption that either one or other of the two hypotheses holds while the test in running. This restriction is removed in the disruption test of Shiryaeyev [12] where the probability of the change in the hypotheses at any sample is taken into account, and the decision as to whether a transition has occurred from the normal to the degraded mode is made on the basis of the *a posteriori* probability of failure derived from the past and current observations. In contrast to Wald's test, Shiryaeyev's test provides a smaller expected value of the delay in fault detection at the expense of increased computations. Chien et al. [9, 10] developed a sub-optimal procedure for on-line fault detection in navigational system sensors, which is computationally faster and retains the improved features of Shiryaeyev's test.

3 The Redundancy Management Procedure

The salient features of the redundancy management procedure reported in [5] are summarized in this section. Some of the definitions of pertinent parameters, which were introduced earlier, are repeated here for the sake of completeness.

The calibrated measurements [13] of the plant variable, under consideration, are modeled at the sample time t as

$$m(t) = H(t)x(t) + e(t) \tag{1}$$

where

m is the $(l \times 1)$ vector of known measurements,
H is the $(l \times n)$ *a priori* known measurement matrix with $(l > n)$ such that any n rows are linearly independent,
x is the $(n \times 1)$ unknown plant variable,
e is the $(l \times 1)$ additive noise and errors associated with the calibrated measurements.

For brevity, the time dependence notations are dropped in subsequent mathematical equations. Thus, (1) is rewritten as

$$m = Hx + e \tag{2}$$

The $((l-n) \times l)$ projection matrix V is chosen such that its $(l-n)$ rows form an orthonormal basis for the left null space of H, i.e.,

$$VH = 0 \quad \text{and} \quad VV^T = I_{l-n} \tag{3}$$

The column space of V is known as the parity space [14] of H and the projection of m onto the parity space is called the parity vector p.

$$p = Vm = Ve \tag{4}$$

Several definitions are now introduced to delineate the concepts of consistency and inconsistency of a set Ω of l redundant measurements $(l > n)$ of an n-dimensional variable relative to all $(n+1)$-tuplets, i.e., subsets of cardinality $(n+1)$ denoted as $S_1, S_2, \ldots, S_r$ where r is given as

$$r = \frac{l!}{(n+1)!(l-n-1)!} \tag{5}$$

The magnitude of the parity vector p_i of dimension one, generated from $(n+1)$ measuremens in the $(n+1)$-tuplet S_i, $i = 1, 2, \ldots r$, is a measure of inconsistency of the $(n+1)$ measuremens in S_i.

Definition 1: The inconsistency index $\xi(t)[S_i]$ of the $(n+1)$-tuplet S_i, at time t is a real number directly related to $|p_i(t)|$. Thus, after dropping the time dependence notation for brevity, the inconsistency index can be expressed as

$$\xi : S_i - [0, \infty),\ i = 1, 2, \ldots, r \tag{6}$$

Furthermore, ξ is appropriately scaled such that, under nominal conditions, $\xi[S_i] \leq 1$ for every i. The exact structure of ξ is dependent on the noise statistics as well as on the fault detection algorithm. For a sequential test procedure, ξ is obtained via a recursive relationship which relies on the cumulative information provided by the measurement history from the past and current observations. For each $(n+1)$-tuplet the one-dimensional parity vector is generated at every sample to compute the inconsistency index. An algorithm for obtaining ξ is derived later using Chien's modified sequential probability ratio test [9, 10]. According to Chien's methodology the noise associated with each parity vector (which is a linear combination of the measurements in the $(n+1)$-tuplet) is assumed to be Gaussian. We reiterate that the redundancy management procedure [5] is independent of the noise statistics and the fault detection algorithm, and is not restricted to Gaussian noise.

Definition 2: An $(n+1)$-tuplet S_i, $i = 1, 2, \ldots, r$ is defined to be internally consistent if its inconsistency index is less than or equal to unity, i.e., if $\xi[S_i] \leq 1$.

Definition 3: A set of measurements is defined to be consistent if each of its $(n+1)$-tuplets is internally consistent.

Definition 4: Two disjoint subsets Ω_1 and Ω_2 of a measurement set Ω is defined to be relatively inconsistent if there exists no consistent $(n+1)$-tuplet having at least one element from each of Ω_1 and Ω_2.

Definition 5: A set of measurements that is not consistent is defined to be inconsistent [moderately consistent] if the set can [cannot] be split into two or more relatively inconsistent subsets.

The concept of moderate consistency is germane to the situation when errors in some of the measurements are in the vicinity of their respective error bounds such that the measurements are contiguously dispersed and no measurement appears to be clearly malfunctioning.

Sequantial Test Algorithm. Given redundant measurements for an n-dimensional plant variable, the one-dimensional parity vector $p_i(t)$ for the i-th $(n+1)$-tuplet S_i at the sampling instant t is given as

$$p_i(t) = V_i \mu_i(t) \text{ for } i = 1, 2, \ldots, r \tag{7}$$

where V_i is the $1 \times (n+1)$ projection matrix associated with μ_i which is the $(n+1) \times 1$ vector representing measurements in S_i. The probability distribution of p_i is assumed to be Gaussian on the justification that p_i is a linear combination of the measurements in S_i, which are usually uncorrelated or weakly correlated. Using the *a priori* information on the measurement noise covariance matrix in S_i, $p_i(t)$ is scaled to $g_i(t)$ such that the variance of $q_i(t)$ is unity. For nominally unfailed conditions $E[q_i(t)] = 0$ for every i and t.

In the sequential tests a decision is made between the no-failure hypothesis, and one or more failure hypotheses, on the basis of the information processed at consecutive samples. In this paper, only one failure hypothesis is postulated to represent all abnormal modes including high and low failures. In general, if M different failure modes are considered, then $(M+1)$ distinct modes of operations should be designated by $(M+1)$ mutually exclusive and exhaustive hypotheses such that each hypothesis can be treated as a Markov state. The recursive relations for generating *a posteriori* probabilities for multiple failure modes and an extension of this test procedure to multiple failure hypotheses are not presented in this paper.

The failure and no-failure hypotheses, H_1 and H_0 respectively, are defined below.

H_1: $q_i(t)$ is Gaussian with mean $\pm\theta_i$ and unit variance at every sample t and for every $(n+1)$-tuplet i. The mean is positive or negative signifying high or low failures, respectively.

H_0: $q_i(t)$ is Gaussian with zero mean and unit variance $\forall t$, $\forall i$.

The log likelihood ratio at the tth sample is defined as

$$\phi_i(t) = \ln \frac{P[q_i(t) \mid H_1]}{P[q_i(t) \mid H_0]}, \quad i = 1,2, \ldots, r \tag{8}$$

If the measurement noise is stationary, the log likelihood ratio $\Phi_i(k)$ for k consecutive conditionally independent samples is given by

$$\Phi_i(k) = \ln \frac{P[q_i(1), q_i(2), \ldots, q_i(k) \mid H_1]}{P[q_i(1), q_i(2), \ldots, q_i(k) \mid H_0]}$$

$$= \sum_{t=1}^{k} \phi_i(t) \tag{9}$$

which yields the following two recursive relations for positive and negative values of the mean in the hypothesis H_1.

$$\begin{aligned} \Phi_i^+(k) &= \Phi_i^+(k-1) - \theta_i(\theta_i/2 - q_i(k)) \\ \Phi_i^-(k) &= \Phi_i^-(k-1) - \theta_i(\theta_i/2 + q_i(k)) \end{aligned} \tag{10}$$

Following Chien's sequential test procedure [9, 10], the above algorithm is formulated for $i = 1,2, \ldots, r$ as follows.

Initialization:

$$\Phi_i^+(0) = \Phi_i^-(0) = 0. \tag{11}$$

Lower limit setting:

$$\begin{aligned} \Phi_i^+(k) &= \mathrm{Max}[\Phi_i^+(k), \delta_i] \\ \Phi_i^-(k) &= \mathrm{Max}[\Phi_i^-(k), \delta_i] \end{aligned} \quad \text{for all } k > 0 \tag{12}$$

Consistency of the ith $(n+1)$-tuplet:

$$\Phi_i^+(k) \le \sigma_i \text{ and } \Phi_i^-(k) \le \sigma_i \text{ for all } k > 0 \tag{13}$$

Inconsistency of the ith $(n+1)$-tuplet:

$$\Phi_i^+(k) > \sigma_i \text{ or } \Phi_i^-(k) > \sigma_i \text{ for all } k > 0 \tag{14}$$

Upper limit setting:

$$\begin{aligned} \Phi_i^+(k) &= \mathrm{Min}[\Phi_i^+(k), \sigma_i] \\ \Phi_i^-(k) &= \mathrm{Min}[\Phi_i^-(k), \sigma_i] \end{aligned} \quad \text{for all } k > 0 \tag{15}$$

where $\sigma_i = \ln[(\theta_i)^2 N/2]$ is the detection threshold as derived in [9, 10], N being the allowable mean time, i.e., the number of samples, between false alarms ($N \gg 1$), and the lower limit setting δ_i is set equal to the *a priori* probability of failure of any measurement in the ith $(n+1)$-tuplet during one sample interval. The upper limits of Φ^+ and Φ^- are set to σ_i to enhance recovery from a failure after the faulty measurement has been reinstated. Chien [9, 10] has presented a detailed discussion on the sensitivity of the parameters N, θ, and σ relative to the system performance (e.g., probability of false alarms and delay in detecting faults).

The inconsistency index in Definition 1 of Section 3 can be interpreted as follows.

$$\xi(k)[S_i] = \mathrm{Max}[\Phi_i^+(k), \Phi_i^-(k)]/\sigma_i \tag{16}$$

The magnitude θ_i of the mean can be chosen as a function of the error bounds of the measurements in the ith $(n+1)$-tuplet. Such error bounds are usually available from the instrument manufacturer's specifications or the actual measurement noise statistics.

The above procedure suffices to determine the internal consistency of each $(n+1)$-tuplet as delineated in Definition 2. The subsequent steps for identifying the largest consistent or moderately consistent subset and inconsistent subsets, if any, follow Definitions 3, 4, and 5. These steps belong to the redundancy management procedure [5] and are essentially independent of the fault detection strategy; an implementation of the redundancy management algorithm for scalar measurements is outlined in Appendix C of [5]. The largest consistent or moderately consistent subset of redundant measurements serves the purpose of calibration and plant variable estimation as reported in [13]. The inconsistent measurements are isolated as faulty and can be reinstated after their repair/replacement.

4 Real-Time Experimentation

The variable-sample sequential fault detection procedure described above was experimentally verified by on-line testing of sensor failures in the 5 MWt nuclear reactor MITR-II [15].

Whereas the FDI software has been recently developed and implemented, the experimental set-up is similar to that reported in our earlier publications [4, 5, 13]. For the sake of completeness, a brief description of the instrumentation in the test facility is given below.

The nuclear instrumentation for the research reported in this paper consists of three neutron flux sensors and a gamma-ray sensor that correlates neutron power with the radioactivity (N-16) of the primary coolant. Four independent measurements of primary coolant flow are obtained from pressure differences across orifices. The temperature difference ΔT between the hot and cold legs of the primary coolant are measured by three independent sensors. The noise and statistical characteristics of the MITR-II's flow, temperature, and neutron flux instrumentation are similar to those in commercial reactors. These sensors are connected to an LSI-11/23 microcomputer system through qppropriate isolators, signal conditioners, and A/D converters. Plots of typical sensor repsonses for both normal and abnormal operations are presented in [13].

Built-in tests such as limit checks and rate checks were routinely incorporated within the FDI software; and fixed-sample sequential test algorithm, reported in [5] and [6], was replaced by the above variable-sample algorithm which requires an execution time (on an LSI-11/23 processor) of about 200 ms per cycle including the time required for data acquisition and signal processing; the increase in processing time is approximately 50 ms per cycle. On the average, the variable-sample approach yielded smaller delays in detecting faults than the fixed-sample approach. The improvements were not very significant primarily because of the fact that the parity vector p_i in (7) did not have a Gaussian distribution. Since the measurements are scalar, i.e., $n = 1$ in (1), p_i generated from each pair is simply the difference between the two measurements in that pair. On the other hand, for 3-dimensional variables such as velocity or acceleration, p_i is obtained as a linear combination of four measurements and is expected to follow a Gaussian distribution more closely than that for scalar measurements.

Fault detection and isolation capabilities of the reported procedure were tested for both natural and injected failures of sensors. For continuous operations extended over a period of six months, an abrupt failure in one of the flow sensors and a drift failure in one of the temperature sensors were automatically detected and isolated, and no false alarms were reported. The efficacy of the procedure was also tested for abnormal operations by injecting calibration errors, drift errors and other *in-range* faults as described below.

Faulty Sensor Calibration. A bias error was introduced on-line in one of the three ΔT measurements such that the resulting offset exceeded the allowable error bound (i.e., half of the mean θ in the failure hypothesis H_1) by a modest amount. An inconsistency of this sensor with respect to the remaining two sensors (which were mutually consistent) caused the isolation of the affected sensor within a few samples. Similar results were obtained by perturbing the scale factor instead of the bias. These tests were repeated for the power and flow sensors.

Gradual Drift. Drifts were introduced in certain sensors in the form of ramp functions. Appropriate alarms were received when the drifts exceeded the respective bounds of calibration correction [13] in the given sensors.

Degraded Instrumentation. Random noises with zero means were added to several measurements. On the average, alarm rates increased with higher noise to signal ratio.

The procedure was also tested for transient operations. During a reactor shutdown process, the estimated power followed closely the true power until the calibration of neutron flux sensors (at a power level below 1 MWt) was no longer accurate.

5 Summary and Conclusions

The paper presents the concepts and application of a variable-sample sequential test procedure for fault diagnosis of time-dependent sensor signals and analytically derived measurements. The procedure has been experimentally verified for real-time detection and isolation of faulty sensors and plant equipments using commercially available microcomputers in the MIT nuclear research reactor MITR-II.

The procedure is particularly suited for intelligent instrumentation in strategic processes like spacecraft, aircraft, and nuclear power plants where redundant measurements are usually available for critical plant variables. Future research is needed to quantify the accuracy and robustness of the integrated redundancy management and failure detection procedure for different classes of applications.

Acknowledgment

The author acknowledges the benefits of discussion with Dr. M. Desai and the assistance offered by the staff of the MIT Nuclear Reactor Laboratory for conducting the experiments.

References

1 Willsky, A. S., "A Survey of Design Methods for Failure Detection in Dynamic Systems," *Automatica*, Vol. 12, 1976, pp. 601-611.

2 Kerr, T. H., "The Controversy over the Use of SPRT and GLR Techniques and Other Loose-Ends in Failure Detection," *American Control Conference*, June 1983, pp. 966-977.

3 Daly, K. C., Gai, E., and Harrison, J. V., "Generalized Likelihood Test for FDI in Redundant Sensor Configurations," *Journal of Guidance and Control*, Vol. 2, No. 1, Jan.-Feb. 1979, pp. 9-17.

4 Ray, A., Desai, M., and Deyst, J., "Fault Detection and Isolation in a Nuclear Reactor," *Journal of Energy*, Vol. 7, No. 1, Jan.-Feb. 1983, pp. 79-85.

5 Ray, A., and Desai, M., "A Redundancy Management Procedure for Fault Detection and Isolation," ASME JOURNAL OF DYNAMIC SYSTEMS, MEASUREMENT, AND CONTROL, Sept. 1986, pp. 248-254.

6 Deckert, J. C., Fisher, J. L., Laning, D. B., and Ray, A., "Signal Validation for Nuclear Power Plants," ASME JOURNAL OF DYNAMIC SYSTEMS, MEASUREMENT, AND CONTROL, March 1983, pp. 24-29.

7 Chow, E. C., and Willsky, A. S., "Analytical Redundancy and the Design of Robust Failure Detection Systems," *IEEE Trans. Aut. Contr.*, Vol. AC-29, July 1984, pp. 603-614.

8 Ray, A., Geiger, R., Desai, M., and Deyst, J., "Analytic Redundancy for On-Line Fault Diagnosis in A Nuclear Reactor," *Journal of Energy*, Vol. 7, No. 4, July-Aug. 1983, pp. 367-373.

9 Chien, T. T., and Adams, M. B., "A Sequential Failure Detection Technique and its Application," *IEEE Trans. Aut. Contr.*, Vol. AC-21, No. 5, Oct. 1976, pp. 750-757.

10 Chien, T. T., "An Adaptive Technique for a Redundant-Sensor Navigation System," *C. S. Draper Laboratory* Report No. T-560, February 1972.

11 Wald, A., *Sequential Analysis*, John Wiley, New York, 1947.

12 Shiryaeyev, A. N., *Optimal Stopping Rules*, Springer-Verlag, New York, 1978.

13 Ray, A., and Desai, M., "A Calibration and Estimation Filter for Multiply Redundant Measurement Systems," ASME JOURNAL OF DYNAMIC SYSTEMS, MEASUREMENT, AND CONTROL, June 1984, pp. 149-156.

14 Potter, J. E., and Suman, M. C., "Threshold-less Redundancy Management with Arrays of Skewed Instruments," Integrity in Electronic Flight Control Systems, *AGRADOGRAPH-224*, 1977.

15 *Reactor Systems Manual*, Report No. MITNRL-004, MIT Nuclear Reactor Laboratory, 1980.

A Proportional-Integral Controller for Resistance Spot Welding Using Nugget Expansion

H. S. Chang[1], Y. J. Cho[2], S. G. Choi[3], and H. S. Cho[4]

An electrode wear has been found to cause considerable quality variation in the process of resistance spot welding. This paper presents a study on the electrode wear growth and develops a microprocessor-based feedback controller that can compensate weld quality variation due to electrode deterioration. The controller developed in this study utilizes a PI control algorithm, incorporating the electrode movement as a feedback signal. A series of experiments was performed to clarify the wear growth phenomenon and to evaluate the performance of the controller. The experimental results show that the electrode wear can significantly cause quality deterioration, but this can be effectively compensated via application of this proposed control technique.

1 Introduction

Due to the large number of interrelated variables associated with the process, weld quality of the resistance spot welding is extremely operation-dependent and thus requires on-line monitoring and control of appropriate process variables that can reflect weld quality. One important process variable which affects the spot weld quality is the electrode tip geometry, since the current path or current density depends upon the state of contact between the electrode and weldment. It has been also shown that the shape of the electrode tip influences the profile of the electrode/workpiece contact stress and surface resistance (Nied, 1984). The electrode tip wear is, therefore, considered to play an important role in the formation of the nugget and thus to be one of the important in-process disturbance variables.

Previous studies (Harworth and Rolls, 1970; Key and Courtney, 1974; Nadkarni and Weber, 1977) on the electrode tip wear have been concentrated on the wear mechanism, its effects on weld quality and wear-resistive electrode developments. However, the phenomenon of the electrode wear growth has not been fully investigated and furthermore no extensive effort with regard to "tracking control methodology" has been made to compensate this electrode deterioration, although Kuchar et al. (1982) introduced a resistance spot welding control system that is currently on the market. This paper presents briefly a study on the electrode wear growth and suggests a microprocessor-based on-line feedback controller which can compensate the electrode wear effects on weld quality. The motivation of this investigation is to include this phenomenon in the design of a robust adaptive on-line control algorithm which guarantees a uniform weld quality.

To clarify progressive wear growth phenomenon and its effect on weld quality in more detail a series of experiments was conducted with Cu-Cr electrodes of tapered flat shape under fixed welding condition, i.e., constant welding current, constant electrode force, uniform surface condition of workpiece material. Other conditions such as workpiece temperature, electrode tip temperature and coolant temperature were assumed to be constant. Based upon these results, the elec-

[1]Associate Professor, Department of Mechanical Engineering, Myong Ji University, Yong-In, Kyung-Gi, Korea.

[2]Manager, Daewoo Shipbuilding & Heavy Machinery LTD., Ok-Po, Kyung-Nam, Korea.

[3]Deputy Manager, Hyundai Electronics Industries Co. LTD., Seoul, Korea.

[4]Professor, Department of Production Engineering, Korea Advanced Institute of Science & Technology, Seoul, Korea, Mem. ASME.

Contributed by the Dynamic Systems and Control Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS. Manuscript received by the Dynamic System and Control Division January 1987. Associate Editor: R. Shoureshi.