SIMSPRT: A SAS® Code for Simulation of Sequential Probability Ratio Tests

Kenny C. Gross
Reactor Analysis Division
Argonne National Laboratory
Argonne, Illinois

Kristin K. Hoyer
Dept of Industrial Engineering/Management Sci.
Northwestern University
Evanston, Illinois

ABSTRACT

Expert systems are being developed at Argonne National Laboratory for signal validation and equipment operability surveillance in a variety of reactor and industrial applications. These systems are based upon the Sequential Probability Ratio Test (SPRT), a sensitive sequential test which provides early annunciation of the onset of degradation in sensors or equipment. SIMSPRT is a SAS® code used as a tool in the design of and reliability analysis of SPRT modules for new signal applications. Utilizing Monte Carlo simulation, SIMSPRT estimates the false-alarm and missed-alarm probabilities for an individual SPRT module as a function of the input signal noise distribution. In addition to reliability analyses, SIMSPRT can be used to simulate a variety of known sensor degradation events for purposes of determining the sensitivity characteristics of a SPRT module and the maximum time-to-annunciation for subtle degradation modes.

I. Introduction

An artificial-intelligence based expert system has been developed at Argonne National Laboratory for signal validation and equipment operability surveillance in industries that require high-reliability, high-sensitivity annunciation of degraded sensors, discrepant signals, or the onset of system disturbances [1,2]. This system employs an extremely sensitive pattern recognition technique, the Sequential Probability Ratio Test (SPRT), for automatic annunciation of the onset of process anomalies. Conventional parameter-surveillance schemes are sensitive only to gross changes in the process mean, or to large steps or spikes that exceed some threshold limit check. These conventional methods suffer from either large false alarm rates (if thresholds are set too close) or large missed (or delayed) alarm rates (if the thresholds are set too wide). The SPRT provides a superior surveillance tool because it is sensitive not only to disturbances in signal mean, but also to changes in the statistical quality (variance, skewness, bias) of the monitored signals. A second important feature of the SPRT technique is that it has quantitative bounds on the false-alarm and missed-alarm probabilities. This allows formal reliability analysis to be applied to an overall expert system comprising many SPRT modules that are simultaneously monitoring a variety of plant variables.

This paper reports the development and application of the SIMSPRT, a SAS-system code which is used for design, testing, and performance evaluation of SPRT modules for expert system applications. SAS software was chosen for this project because of the language's powerful coding features and its ability to model prototype systems quickly and cost-effectively. SAS macro language was exploited in a parametric sensitivity study to explore the effects of small changes in variance and bias on the expected number of observations needed for annunciation of discrepant signals using the SPRT. Finally, PROC G3D proved indispensable for displaying results of our many-variable sensitivity investigation.
II. SPRT IMPLEMENTATION

For each pair of signals being monitored, a SPRT index is computed with each new observation using the defining equation

$$SPRT = \frac{M}{\sigma^2} \sum_{k=1}^{n} (Y_k - \frac{M}{2})$$  \hspace{1cm} (1)

where $\sigma^2$ is the variance associated with the physical process during normal (i.e. undegraded) operation, $M$ is a preassigned system disturbance magnitude, and $Y_k$ are discrete samples formed by differencing the two digitized input signals. The SPRT is continuously compared against two stopping thresholds ($A < 0, B > 0$) that are derived by user-specified false alarm (alpha) and missed alarm (beta) probabilities using Wald’s\[2\] approximations:

$$A = \ln\left(\frac{\beta}{1-\alpha}\right)$$  \hspace{1cm} (2)

and

$$B = \ln\left(\frac{1-\beta}{\alpha}\right)$$  \hspace{1cm} (3)

The sequential sampling and decision strategy can be concisely represented as: If $SPRT < A$, then one can conclude with confidence factor $(1 - \beta)$ that both monitored sensors are tracking the same physical process. The SPRT is reset to zero and sampling continues. If $SPRT > B$, then one can conclude with confidence factor $(1 - \alpha)$ that one of the monitored systems is out of tolerance. In this case the status of a warning flag is set and SPRTs are triggered to test pairs of redundant sensors on each component to distinguish between actual equipment degradation and the degradation or failure of an individual sensor. In the latter case the sequence of SPRT tests will identify the sensor that has failed. For this study, the SIMSPRT code was developed to perform Monte Carlo simulation of SPRTs with input signals having prescribed means and noise distributions.

III. SPRT Simulations

Typical output from SIMSPRT is shown in Fig. 1 for a "clipped" SPRT, i.e. one for which we clip the SPRT index when it reaches a stopping threshold (see Eqns. 2 and 3) and reset it to zero. The upper two subplots in Fig. 1 show a pump rotor speed signal in RPM during simulated operation with no degradation in the pumps and with normal operability of the sensors.

The difference function is plotted in the third subplot. The SPRT index is plotted in the fourth subplot. With no failures in the pumps or the sensors, we see that the SPRTs grow linearly with time to the negative threshold and are reset to zero.

Case 2, in which we simulate a subtle bias-type error in speed signal 1, is illustrated in Fig. 2. Note that signal 1 exceeds signal 2 by approximately 0.5 RPM. In this case one SPRT index grows in the positive direction. It is that SPRT, signified by the solid line in the fourth subplot of Fig. 4, which rapidly reaches the positive stopping boundary $A$ (Eqn. 2), signifying a data-disturbance alarm.

For Case 3 we simulate a linear drift in signal 2. Output from SIMSPRT is shown in Fig. 3. Note that with linear degradation of one of the input signals there is an increasing frequency of SPRT triggers in the bottom subplot.

The final example is that of a sudden step change in one of the monitored signals. SIMSPRT output for this simulated disturbance mode is shown in Fig. 4. Examination of the lower subplot in Fig. 4 reveals that even with a very subtle discrepancy in Signal 2, the SPRT algorithm triggers a warning annunciation with 1 s. (The sampling rate for simulated cases 1-4 was 2/s).

Figure 5 illustrates the first stage of expert system processing for a simple generic example application involving a single industrial device that is equipped with triply-redundant sensors for measurement of two physical variables. The expert system first identifies the minimum unique set of signal pairs that will be needed for the network of interacting SPRT modules. Figure 6
illustrates a typical generic application involving two industrial devices that are operated in parallel (e.g., jet engines, propeller drive motors on a ship, turbomachinery in an industrial plant, etc.). For this example we also assume triply-redundant sensors are available for measuring each of two separate physical variables. Once again the expert system identifies the pair-wise sensor combinations that it uses subsequently in building the conditional branching hierarchy for the SPRT-module configuration.

SIMSPRT output for a simulation of the expert system configuration in Fig. 6 is reproduced in Figs. 7 and 8. Figure 7 shows the simulated sensor output for the sensors identified in the expert system application from Fig. 6. Note that for this example, we have simulated a subtle step change in sensor D2. Corresponding SPRT results are plotted in Fig. 8. It is observed that the two SPRT modules surveilling signal D2 quickly start triggering data disturbance warnings within a couple observations of the onset of the disturbance. In actual expert system deployment, the expert system logic structure uses this information to identify sensor D2 as failed.

IV. SPRT Tripping Frequency

Another important use for the SAS-based SIMSPRT code is for long-term performance analysis of SPRT modules. Long term performance is evaluated in terms of the tripping frequency, or the frequency of acceptances of the failure hypothesis (H1) and the non-failure hypothesis (H2). Tripping-frequency output from SIMSPRT for a typical simulation experiment is illustrated in Fig. 9. If we run SIMSPRT for very large numbers of observations, we are able to assess the "true" false-alarm and missed-alarm rates for a given set of SPRT input parameters.

By use of Monte Carlo techniques, SIMSPRT accumulates long-term tripping-frequency statistics as a function of α, β, signal variance, sensor failure magnitude, and the mean value of the difference function. Three-dimensional response-surface contours are output using PROC G3D for either the average sample number (ASN) (average number of observations to trigger an annunciation); or in terms of the empirical false-alarm and missed-alarm probabilities. Fig. 10 shows a typical result from 200,000 Monte Carlo simulations with SIMSPRT. The figure plots ASN contours as a function of signal variance and input missed-alarm probability. In Fig. 11, we plot the empirical false alarm probability as a function of variance and SFM. These 3D contours and the others available from SIMSPRT have become indispensable aids in our rapid prototyping of new expert system applications. The 3D contours make it possible for the expert system designer to quickly select SPRT parameters that minimize the time to annunciation for a given implementation, while assuring that pre-specified false and missed-alarm probabilities are met.

V. Conclusions

SIMSPRT has been developed with SAS system software for use as a software engineering tool in the design and reliability analysis of SPRT modules for expert-system surveillance algorithms. SIMSPRT can be used to simulate simple SPRTs that monitor a single pair of signals, and multiple SPRTs that correspond to a network of SPRT modules deployed in parallel.

In the monte carlo simulation mode, SIMSPRT generates empirical false-alarm and missed-alarm probabilities as a function of five independent SPRT input variables. SAS-based SIMSPRT has become an indispensable tool in Argonne Laboratory's on-going development of expert systems for surveillance and diagnosis of nuclear and industrial equipment.

References


Fig. 1
Clipped SPRT Simulation
CASE 1: BOTH SIGNALS TRACKING SAME PROCESS

Fig. 2
Clipped SPRT Simulation
CASE 2: SIGNAL 2 EXCEEDS SIGNAL 1 BY 0.5 RPM

Fig. 3
Clipped SPRT Simulation
CASE 3: LINEAR DRIFT IN SIGNAL 2

Fig. 4
Clipped SPRT Simulation
CASE 4: STEP CHANGE IN SIGNAL 2
Fig. 5
Pair-wise Sensor-Combination Identification
Example 1: Industrial device equipped with
triply-redundant signals for two physical variables

Fig. 6
Pair-wise Sensor-Combination Identification
Example 2: Industrial device deployed in parallel
with triply-redundant signals for two physical variables

Fig. 7
Simulated Sensor Output
CASE 2: STEP CHANGE IN SENSOR D1

Fig. 8
SPRT Output: case 2
Fig. 9
SPRT Response Time and Tripping Frequency
CASE 2: RANDOM GAUSSIAN SIGNAL
VARIANCE = 0.5

Fig. 10
200,000 observations
TIME = 20, ALPHAS = 0.001

Fig. 11
EMPIRICAL FALSE ALARM PROBABILITY
\( \alpha = 0.001, \beta = 0.0001 \)