LA-UR-82-2780

Conf-821119--4

Los Alamos National Laboratory is operated by the University of California for the United States Department of Energy under contract W-7405-ENG-36

MASTER

LA-UR--82-2780

DE83 000605

TITLE: LOSS-PATTERN IDENTIFICATION IN NEAR-REAL-TIME ACCOUNTING SYSTEMS

AUTHOR(S): F. Argentesi, J. F. Hafer, J. T. Markin, and J. P. Shipley

SUBMITTED TO: International Symposium on Recent Advances in Nuclear Materials Safeguards, Vienna, Austria, November 8-12, 1982

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

Los Alamos

Los Alamos National Laboratory
Los Alamos, New Mexico 87545

FORM NO. 836 R4
ST. NO. 2629 5/81

INTERNATIONAL ATOMIC ENERGY AGENCY

INTERNATIONAL SYMPOSIUM ON RECENT ADVANCES IN NUCLEAR MATERIALS SAFEGUARDS

Vienna, Austria, 8–12 November 1982

IAEA-SM-260/ 52

LOSS-PATTERN IDENTIFICATION IN NEAR-REAL-TIME ACCOUNTING SYSTEMS

F. Argentesi
Joint Research Center
Ispra, Italy

J. F. Hafer, J. T. Markin, and J. P. Shipley
Los Alamos National Laboratory
Los Alamos, New Mexico, USA

IAEA-SM-260/52

LOSS PATTERN IDENTIFICATION IN NEAR-REAL-TIME ACCOUNTING SYSTEMS

ABSTRACT

To maximize the benefits from an advanced safeguards technique such as near-real-time accounting (NRTA), sophisticated methods of analyzing sequential materials accounting data are necessary. The methods must be capable of controlling the overall false-alarm rate while assuring good power of detection against all possible diversion scenarios. A method drawn from the field of pattern recognition and related to the alarm-sequence chart [1,2] appears to be promising. Power curves based on Monte Carlo calculations illustrate the improvements over more conventional methods.

1. INTRODUCTION

Among the various advanced techniques currently under consideration for international safeguards, near-real-time accounting (NRTA) holds the promise of increased sensitivity and timeliness over present accounting practices. To realize this improved capability requires suitable procedures for examining the relatively large body of materials accounting information.

The purpose of this paper is to report on a new technique having its basis in the field of pattern recognition.

Diversion of nuclear material can occur in a wide variety of scenarios that are generally considered to be bounded by abrupt diversion (diversion of a relatively large amount over a short time period such as one week) and by protracted diversion (an accumulation of individually small diversions over a long time such as one year). In practice, the safeguards system must address all the intermediate scenarios as well.

One of the potential advantages of NRTA is that it can treat all such diversion scenarios because materials balances become available in a timely manner. However, the procedures for examining the materials balances for evidence of diversion must be carefully constructed to take maximum advantage of the available information, while guarding against the range of possible diversion scenarios. One way to achieve this goal, as suggested in several previous references [1-8], is to study all possible contiguous subsequences of materials balances; obviously, all diversion scenarios would be covered by such a scheme.

As currently implemented, the method is based on sequential statistical tests because they seem best suited to the NRTA problem. The remaining difficulty concerns controlling the overall false-alarm rate of the composite procedure: in general, each test corresponding to each subsequence of materials balances will tend to contribute to an increase in false alarms. In actuality, the increase is not as large as might be thought because the subsequences are highly correlated.

The exact mathematical structure of a testing procedure that would provide maximum power of detection while controlling the overall false-alarm rate is difficult to determine, especially in an analytical sense. For this reason, we have turned to the field of pattern recognition for assistance. The application of certain of the methods to the alarm-sequence chart, which is described below and in more detail in Refs. 9 and 10, is a natural extension of this graphical display that heretofore has largely been subjectively interpreted. In the following, we briefly describe the method and give some preliminary results that appear highly encouraging.

2. THE PATTERN-RECOGNITION ALGORITHM

A test statistic commonly applied in materials accounting is the cumulative sum of materials balances,

$$\mathrm{CUSUM}(k) = \sum_{i=1}^{k} MB_i \quad ,$$

where MB_i is the i^{th} materials balance and k is the number of currently available materials balances. This statistic does not require any assumptions about the loss pattern, provides an estimate of the total materials loss, and is useful in deciding whether a sequence of materials balances contains statistical evidence of materials loss. This decision problem can be formulated as testing the composite hypothesis

$$H_o: \sum_{i=1}^{k} MB_i^a \leq 0 \quad , \qquad k = 1,2,\ldots,N \quad ,$$

against the composite alternative

$$H_1: \sum_{i=1}^{k} MB_i^a > 0 \quad , \qquad \text{for at least one } k,\ k = 1,2,\ldots,N \quad ,$$

where MB_i^a is the actual materials balance and N is the total number of balances. This test of hypotheses may be formulated as a sequential probability ratio test (SPRT) in which a statistic defined as the ratio of CUSUM probability densities under each hypothesis is compared to a decision threshold [11,12]. Under this testing procedure the decision rule becomes:

$$\text{If } \frac{\text{CUSUM}(k)}{\sqrt{V_c(k)}} \begin{cases} \leq -\sqrt{2\,|\ln T_0|} & \text{, accept } H_0 \text{ ,} \\ \geq \sqrt{2\,|\ln T_1|} & \text{, accept } H_1 \text{ ,} \\ \text{otherwise, take another observation} & , \end{cases}$$

where $V_c(k)$ is the variance of CUSUM(k) and T_o, T_1 are lower and upper thresholds that depend upon the required Type I and Type II errors for this procedure. The thresholds may be determined by the approximations developed by Wald [11]

$$T_0 = \frac{\beta}{1 - \alpha} \text{ and } T_1 = \frac{1 - \beta}{\alpha} \quad ,$$

where α is the false-alarm probability and β is the non-detection probability.

Because the actual loss pattern cannot be known, the SPRT is applied to all contiguous subsequences of materials balances, resulting in a total of N(N + 1)/2 tests for a sequence of N balances. Further, to determine more precisely the significance of each test result, the SPRT is applied using decision thresholds for several false-alarm probabilities. The results of these tests are represented in compact form by an alarm-sequence chart [1,2].

The alarm-sequence chart is a pattern-recognition device that provides a graphic display of the false-alarm probability associated with each indication of loss. Each observation sequence for which the ratio exceeds an upper threshold is assigned an alphabetic descriptor representing the false-alarm probability and a pair of numbers (M,N) that are the initial and final balances in the sequence. Figures 1 and 2 are example alarm charts for abrupt (period 3) and protracted (periods 3-8) diversions, respectively. The association between letters and false-alarm probability is given in Table I.

TABLE I

PLOTTING SYMBOLS FOR ALARM-SEQUENCE CHART

Symbol	False-Alarm Probability		
A	0.02	to	0.07
B	0.01		0.02
C	0.005		0.01
D	0.001		0.005
E	0.0005		0.001
F	0.0001		0.0005
G	0.00005		0.0001

The pattern-recognition algorithm given in this paper is based on the transformation of the alphabetic descriptor and initial and final balance number of the alarm chart into a numerical matrix as follows.

The row number I for a particular subsequence from balance M to balance N is given by

$$I = \frac{N(N - 1)}{2} + M \quad .$$

For example, the subsequence 2 to 4 is represented in row 8. Each row of the transformed matrix represents the significance level of the test result for a particular subsequence. Row 1 is the subsequence from balance 1 to balance 1; row 2 is from balance 1 to balance 2; row 3 is from balance 2 to balance 2; etc. There are NBMAX rows in the matrix given by

$$NBMAX = NB(NB + 1)/2 \quad ,$$

where NB is the number of available materials balances. Each column in the matrix contains a number representing the false-alarm probability used in determining the SPRT thresholds. For example, if the probability ratio for a particular subsequence (matrix row) exceeded the fourth false-alarm probability threshold, that row would have the numbers 1 through 4 in the first four columns, and the remaining columns would be zero:

row M, N   1 2 3 4 0 0 . . . 0 .

The next step is to calculate the row sum vector, which has one element for each row in the transformed matrix. It is calculated from the equation

$$RSUM(I) = \sum_{J=1}^{NC} K^{L}(I,J) \quad ,$$

where I is the row number,

J is the column number,

NC is the number of columns in the matrix K,

K(I,J) is the assigned numerical value of the column number, and

L is a constant power.

Because each row in the matrix represents a subsequence, so does each element in RSUM. Therefore, the next step in the algorithm is to aggregate the elements of RSUM to isolate the information pertaining to each subsequence. One way to do this is to calculate the balance sum vector BSUM, which has NB elements, each being the sum of all the RSUM elements corresponding to the initial and final balance numbers of a particular subsequence. For example, the fourth element of BSUM represents the subsequence starting at 1 and ending at 3. It would contain the sum of RSUM elements 1, 2, 4, 5, and 6 because these elements either start the subsequence, end the subsequence, or both.

The information in the balance sum vector BSUM is used to provide indications of possible diversion and to ascertain the scenario of the diversion. As each balance is taken in this sequential process, the location of the largest element in the BSUM vector is determined. If the magnitude of this element is less than some threshold (obtained through simulation), no loss is assumed and the sequential process continues. If the threshold is exceeded, then loss is assumed and the element number where the maximum BSUM value occurred is used to determine the scenario. For example, if the maximum value occurs in BSUM element 4, representing subsequence 1 to 3, and it exceeds the threshold, then protracted loss in the subsequence 1 to 3 is indicated. If the maximum appears in element 6, representing the subsequence 3 to 3, then abrupt loss at balance 3 is assumed. From the loss scenario determined above, the estimated amount of material lost is obtained from the CUSUM data for the corresponding subsequence.

3. EXAMPLE APPLICATION

The pattern recognition scheme described in Section 2 was tested on simulated data from a simple model process. This process consisted of a single inventory of material in each period i with measurement model

$$I_i = I(1 + \varepsilon_{I_i} + \eta_I) \quad , \qquad i = 1, 2, \ldots, N \quad ,$$

where ε_{I_i} represents an uncorrelated measurement error that changes in each balance period, η_I is a correlated error representing, for example, an instrument bias that does not change over the N periods; I is the true inventory of material. The single input and single output transfer for the model process in each period i have the common measurement model

$$T_i = T(1 + \varepsilon_{T_i} + \eta_T) \quad ,$$

where ε_{T_i}, η_T are the uncorrelated and correlated errors respectively, and T is the true input (output) transfer. For this example the standard deviation of each of the errors is equal to 0.1, and the true values of the inventories and transfers are all 1.0. Under these assumptions the standard deviation of a single materials balance is 0.24, and the standard deviation of the cumulative sum of 25 materials balances is 3.61.

The algorithm was tested in a series of 25-balance-period Monte Carlo simulations under each of the assumed scenarios and various levels of diversion. The threshold for detection was obtained from a no-loss scenario to give a false-alarm probability of 5%. The results are shown in Fig. 3. Also shown in Fig. 3 is the detection probability for a fixed-length CUSUM test applied over 1, 6, and 25 balance periods. In these results the fixed-length test appears to be better than the pattern-recognition method when they are compared for the same loss scenario. Note, however, that in applying the CUSUM we require a priori knowledge of the balances where materials loss will occur, whereas the pattern-recognition method does not need that information. Indeed, when the loss scenario is not known, the correct comparison between the two methods uses the fixed-length CUSUM over 25 balances, reflecting the lack of knowledge about which balances contain materials loss. Table II gives the predicted results of the algorithm for one level of diversion for each scenario average of 5000 simulation runs.

TABLE II

ACTUAL SCENARIO

	Abrupt	Protracted
Actual subsequence	10-10	5-10
Actual loss	0.735	2.782
Overall detection probability	0.5274	0.7250
Predicted abrupt subsequence	10-10	5-5
Predicted abrupt loss	0.92	0.81
Predicted protracted subsequence	9-10	5-6
Predicted protracted loss	1.42	1.64

4. SUMMARY AND FUTURE WORK

This paper has described an application of pattern-recognition principles to materials accounting in which improved probabilities of detecting materials loss are obtained as compared with standard methods such as a fixed-length CUSUM test. Further, these methods can also identify the type of loss scenario. The basis of this work is the alarm-sequence chart, which summarizes the results of testing all contiguous subsequences at several significance levels. Previously these charts were interpreted subjectively; however, by converting the alarm-sequence chart to a numerical matrix and extracting features sensitive to different loss scenarios, the alarm-sequence chart can be interpreted automatically and with an improved detection probability. Other applications of pattern recognition to interpreting alarm sequence charts are under investigation, including constructing alarm-chart templates representative of different loss scenarios and defining a metric that measures the distance between a realized alarm chart and the templates.

REFERENCES

[1] SHIPLEY, J. P., "Decision analysis for nuclear safeguards," in Nuclear Safeguards Analysis, Nondestructive and Analytical Chemical Techniques, E. A. Hakkila, Ed. (American Chemical Society, Washington, DC, 1978).

[2] SHIPLEY, J. P. "Decision analysis in safeguarding special nuclear material," Trans. Am. Nucl. Soc. 27, 178 (1977).

[3] HAKKILA, E. A., COBB, D. D., DAYEM, H. A., DIETZ, R. J., KERN, E. A., SCHELONKA, E. P., SHIPLEY, J. P., D. B. SMITH, AUGUSTSON, R. H., and BARNES, J. W., "Coordinated Safeguards for Materials Management in a Fuel Reprocessing Plant," Los Alamos Scientific Laboratory report LA-6881, Vol. I (September 1977).

[4] HAKKILA, E. A., COBB, D. D., DAYEM, H. A., DIETZ, R. J., KERN, E. A., SCHELONKA, E. P., SHIPLEY, J. P., SMITH, D. B., AUGUSTSON, R. H., and BARNES, J. W., "Coordinated Safeguards for Materials Management in a Fuel Reprocessing Plant," Los Alamos Scientific Laboratory report LA-6881, Vol. II (September 1977).

[5] DAYEM, H. A., COBB, D. D., DIETZ, R. J., HAKKILA, E. A., SHIPLEY, J. P., and SMITH, D. B., "Dynamic Materials Accounting in the Back End of the LWR Fuel Cycle," presented to the American Nuclear Society Meeting on Back End of the LWR Fuel Cycle, Savannah, Georgia, March 19-23, 1978.

[6] SHIPLEY, J. P., "Data analysis for nuclear materials accounting," Safeguards and Nucl. Mater. Manage., (Proc. 1st Annual Symp. Brussels, Belgium, April 25-27, 1979,) ESARDA 10, 391-395.

[7] MARKIN, J. T., BAKER, A. L., and SHIPLEY, J. P., "Implementing advanced data analysis techniques in dynamic materials accounting," Nucl. Mater. Manage IX (Proceedings Issue), 236-244 (1980).

[8] SHIPLEY, J. P., "The Structure of Safeguards Systems," Los Alamos Scientific Laboratory report LA-7337-MS (June 1978).

[9] TOU J. and GONZALEZ, R., Pattern Recognition Principles (Addison-Wesley, Reading, Massachusetts 1974).

[10] FU, K. S., Sequential Methods in Pattern Recognition and Machine Learning (Academic Press, New York 1968).

[11] WALD, A., Sequential Analysis (Dover Publications, New York 1947).

[12] WETHERILL, G., Sequential Methods in Statistics (Spottiswoode Ballantyne and Co., Colchester 1966).

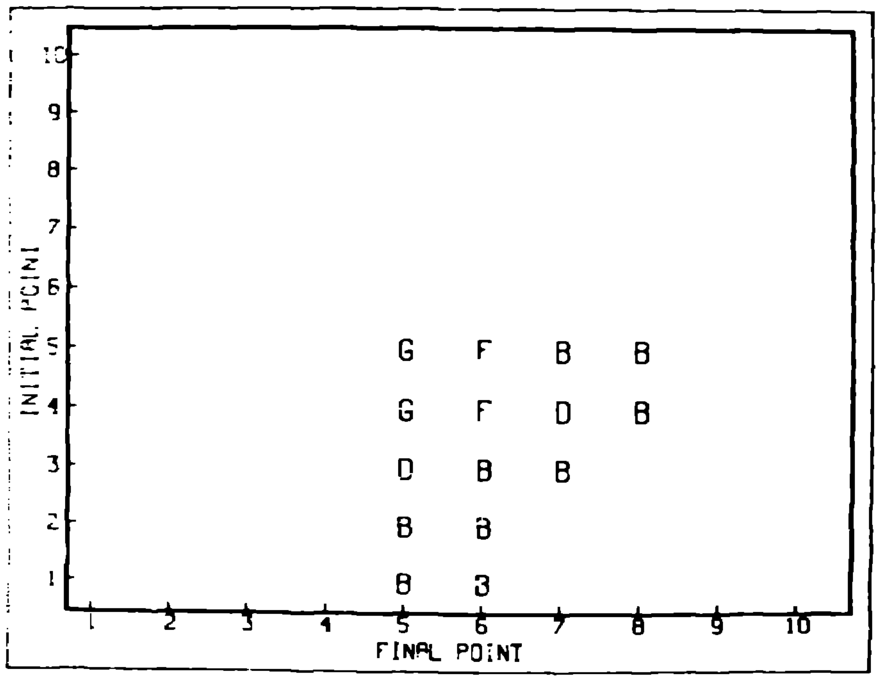

Fig. 1. Example alarm-sequence chart for abrupt materials loss.

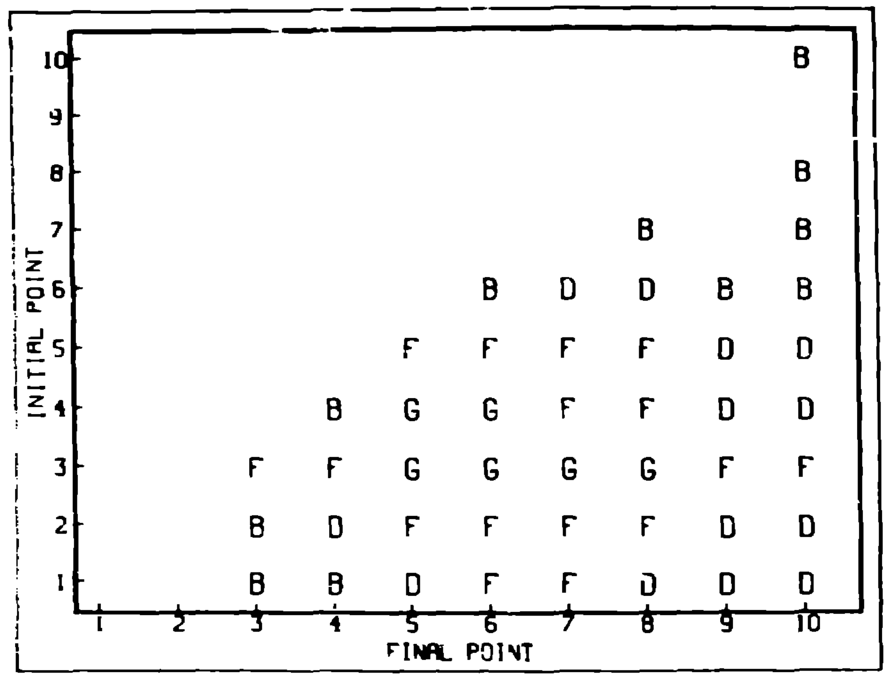

Fig. 2. Example alarm-sequence chart for protracted materials loss.

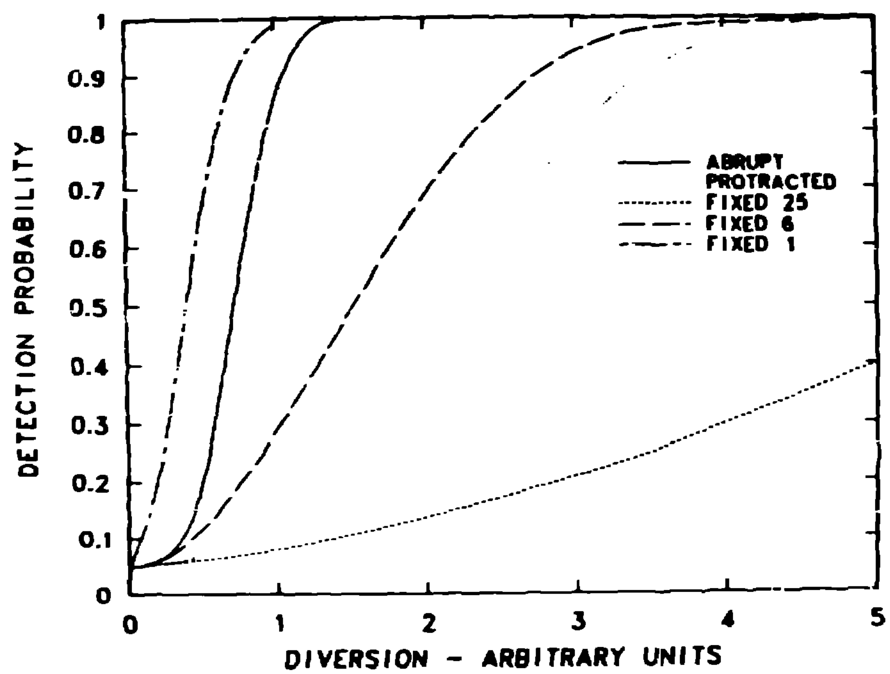

Fig. ?. Po.er curves for the pattern-recognition algorithm and various fixed-length CUSUM tests