LA-UR 98-3031

Approved for public release; distribution is unlimited.

Title: DETECTING ERRORS AND ANOMALIES IN COMPUTERIZED MATERIALS CONTROL & ACCOUNTABILITY DATABASES

CONF-980733--

Author(s): Rena Whiteson
Karen Hench
Tresa Yarbro
Chris Baumgart

Submitted to: 39th Annual Meeting of the Institute of Nuclear Materials Management
Naples Florida
July 27-30, 1998

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

MASTER

Los Alamos
NATIONAL LABORATORY

Los Alamos National Laboratory, an affirmative action/equal opportunity employer, is operated by the University of California for the U.S. Department of Energy under contract W-7405-ENG-36. By acceptance of this article, the publisher recognizes that the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or to allow others to do so, for U.S. Government purposes. Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Department of Energy. The Los Alamos National Laboratory strongly supports academic freedom and a researcher's right to publish; as an institution, however, the Laboratory does not endorse the viewpoint of a publication or guarantee its technical correctness.

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

DETECTING ERRORS AND ANOMALIES IN COMPUTERIZED MATERIALS CONTROL & ACCOUNTABILITY DATABASES*

Rena Whiteson, Karen Hench, and Tresa Yarbro
Los Alamos National Laboratory, Los Alamos, NM 87545 USA

Chris Baumgart
DOE Kansas City Plant,** Albuquerque, NM USA

Abstract

The Automated MC&A Database Assessment project is aimed at improving anomaly and error detection in materials control and accountability (MC&A) databases and increasing confidence in the data that they contain. Anomalous data resulting in poor categorization of nuclear material inventories greatly reduces the value of the database information to users. Therefore it is essential that MC&A data be assessed periodically for anomalies or errors. Anomaly detection can identify errors in databases and thus provide assurance of the integrity of data. An expert system has been developed at Los Alamos National Laboratory that examines these large databases for anomalous or erroneous data. For several years, MC&A subject matter experts at Los Alamos have been using this automated system to examine the large amounts of accountability data that the Los Alamos Plutonium Facility generates. These data are collected and managed by the Material Accountability and Safeguards System, a near-real-time computerized nuclear material accountability and safeguards system. This year we have expanded our user base, customizing the anomaly detector for the varying requirements of different groups of users. This paper describes our progress in customizing the expert systems to the needs of the users of the data and reports on our results.

Introduction

Department of Energy (DOE) sites presently use computerized MC&A systems to account for and track the disposition of nuclear material within the weapons complex. These MC&A systems are typically highly sophisticated and provide controlled access to an enormous amount of information about nuclear material items, such as the type of material, the amount of material, etc. However, such systems must also provide the capability to evaluate and validate the integrity of the information in the database. This should include the ability to perform error checking and error handling as data are entered into the database as well as the means to detect anomalies in the database that may be associated with the diversion of nuclear material. The computerized MC&A system in use at Los Alamos National Laboratory (LANL) is the Material Accountability and Safeguards System (MASS). It contains a very large amount of information about the nuclear material at LANL. The data tend to be very diverse in nature and sometimes very complex. As a result, manual verification of the data integrity is difficult and tedious.

In the remainder of this paper, we will describe our work with the users of the MC&A data at LANL to provide automated error and anomaly detection tools for database assessment. This year we have expanded our user base, customizing the expert-system-based anomaly detection techniques for the varying requirements of different groups of users. Our progress in customizing these expert systems to the needs of the users of the data will be reported.

Previous Work

In early work,[1,2] our primary effort consisted of the design and demonstration of an anomaly detection system to analyze transaction data stored in MASS. These efforts effectively demonstrated that an expert system could be used to quickly process hundreds of MASS transaction entries and serve as an analysis tool by database managers to review transaction activities more efficiently for anomalous behavior.

In follow-on work,[3] we successfully transitioned one version of our anomaly detection expert system to the user community at LANL. This effort was performed for LANL's Nuclear Materials Management and Accountability Group, which is responsible for storage of nuclear materials and forecasting of nuclear materials programmatic needs. The focus was on inventory data, and the goal was to design and implement an anomaly detection expert system to examine and establish the integrity of inventory item data on nuclear material items stored in the vault area of the LANL Plutonium Facility (TA-55). This expert system was designed via extensive collaboration with TA-55 domain experts and now provides them with an analysis tool that rapidly identifies and reviews potentially anomalous database entries.

Most recently, our development focus has been on customizing additional analysis tools for other users in the LANL MC&A community. Two such application development efforts were recently started and will be described below.

Development Approach

Expert systems are the underlying technology we are using to build anomaly detection tools. An expert system is a set of rules that model or codify a process or procedure. Typically, input from a domain expert is required to accurately and completely characterize the process and identify all aspects of the problem solving methods. This knowledge-acquisition step is the most important phase and is the key to developing a successful, practical, and effective expert system. It can be a long iterative process, whose complexity is a function of the complexity of the data and the operations that the data represent.

When building an expert-system-based tool for the review of data contained in special nuclear material MC&A databases, one may adopt one of several different approaches. These include the following:

> *Approach 1 – Quality control.* This approach applies to data entry operations. The specific goal of this approach is to prevent the entry of erroneous data into the database. An anomaly detection system based on this approach would provide the operator with directed assistance in the data entry. In essence, the expert system would control the data entry process, ensuring the quality of the entered information. The largest disadvantage of this approach is that it is the least flexible for the operators because they are prevented from entering any data that is out of the norm. This development approach is probably the best for a new database.
>
> *Approach 2 – Historical data analysis.* In this approach, item data is entered by the operator with few restrictions on the data entry process. In this way, operators are allowed to enter any information they want. The goal of the anomaly detection system is then to analyze the

entered data and examine it for anomalies. The data entries may be examined for consistency and accuracy or may be examined for violation of established procedures.

Approach 3 – Targeted detection. In this approach, the rulebase used by the expert system is custom designed to look for specific, known errors or inconsistencies within the historical database. A rulebase designed in this fashion is quick and easy to build, and interpretation of the results is straightforward. However, an expert system built this way can detect only known potential errors, and new types of errors and anomalies can enter into the system and be missed.

The particular development approach selected depends heavily on the application and the needs of the user. So far, most of our development efforts have adopted the second approach, analysis of historical data. The next section describes the anomaly detection expert systems we are currently building for the MC&A customers at LANL.

Current Development Efforts

We have completed development of one anomaly detection system and have begun development of two others. Each is described below.

Inventory Anomaly Detection System (IADS)

This anomaly detection system is complete and has been delivered to LANL's Nuclear Materials Management and Accountability Group at TA-55. The delivered version of this expert system contains about 60 rules with 15 different qualifiers. Since the inventory item database is driven by Microsoft Access, a special interface for IADS was built to extract inventory item information directly from the classified (SRD) Access database and then transfer those data to the expert system for analysis. The overall system design is illustrated in Figure 1.

In operation, the user specifies a period of time (usually a specific month) over which the inventory item review is to be performed. An Access macro then extracts specific inventory item data from the database and passes that data through a custom-designed parser (written in C code) to pre-filter and format the database information for use by the expert system. This information represents the fundamental features, or qualifiers, used by the expert system to perform its anomaly detection task.

The expert system initially sorts entries according to LOTID (an item identifier) and identifies all those entries with the same LOTID. This primary rule is used as a starting point for the analysis by the TA-55 experts. The rulebase is then applied to these entries, and anomalous entries are identified according to the rules established via the knowledge acquisition process. Specific checks performed by the expert system include:

- checks for hydrogenous items and locations (limited),
- checks to identify multiple material types,
- checks for appropriate item account and location,
- checks for multiple item entries and for combined items, and
- checks for item descriptions of specific interest.

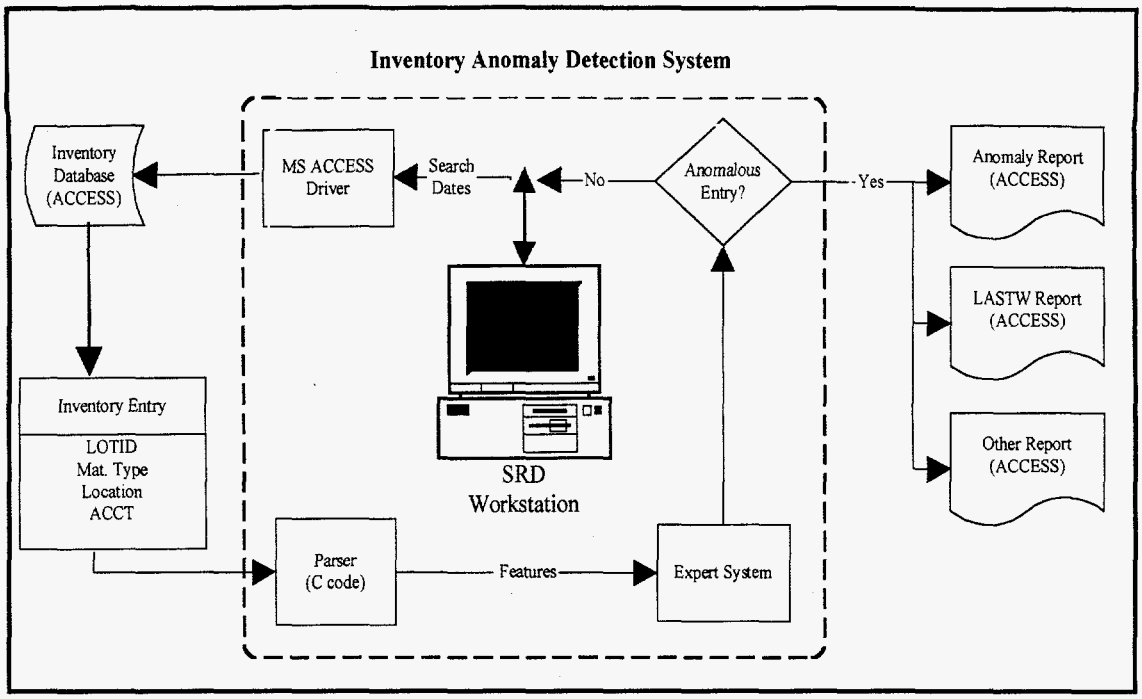

Figure 1. System design for the Inventory Anomaly Detection System.

Those inventory items that are identified by the expert system as "anomalous" are then ranked for severity and reported back to the expert as another Access table. Specialized reports, as requested by the users, are also generated as Access tables so that the entire anomaly detection and review process can be accomplished in a single, user-friendly environment. Thus, this expert system functions as an automated filter, separating the small set of data items of interest to the human expert from the large number of ordinary, unexceptional items.

Safeguards and Security Data Review

Preliminary development of an anomaly detection system for use by the Materials Control and Accountability Group has begun. This group provides safeguards and security program support and ensures effective accountability and control of nuclear materials at the Laboratory. In this application, the expert system is to be developed for use in the evaluation of the consistency and accuracy of historical data. At present, manual evaluation of the data integrity is performed via a validation process where information provided in reports is compared to the data in the actual database. The review process:

- focuses on inventory differences (variances);
- identifies potential anomalies by reviewing the size of the variance with respect to the total item weight (throughput) and expresses this size as a percentage;
- if the percentage is relatively high, then the historical sequence of transactions leading to the anomaly is examined.

This analysis approach is considerably more complex than those that our past efforts have automated, which have concentrated on single database entries. In this case, direct access to the MASS database is required so that the sequence of multiple transactions leading to the suspected anomaly

can be examined. There are some additional rules that apply to particular process statuses (or procedures) and to measurement codes, which must be consistent with the descriptions of the inventory items. The customer's desire is to have the expert system review the data and provide a reduced set of information on transactions that can then be reviewed in detail.

Nuclear Materials and Stockpile Management

This program supports the nuclear weapons stockpile and integrates technologies required for responsible nuclear materials management. The expert system for this program is similar in design to the one built for the Nuclear Materials Management and Accountability Group at TA-55 in that the expert system is to check for consistency in the information entered for nuclear material items. In this case, the expert system to be designed will:

- look for unusual combinations of selected fields of interest (for example, project status coded with incorrect project number);
- identify invalid correlations between process status, project code, and item description; and
- retrieve MASS data on special nuclear materials items in Microsoft Access format.

Summary

Expert systems for database anomaly detection, such as IADS and the other systems initiated this year, have broad applicability to information management systems in nuclear material safeguards. Any application that requires the search of large amounts of data for "out-of-norm" conditions can benefit from the use of this demonstrated and proven technology. We expect to continue building enhanced anomaly detection capabilities with the goal of implementing these techniques into a Laboratory or DOE-wide MC&A system. As the mission of nuclear materials facilities changes from processing to storage, intelligent data analysis tools, such as the expert systems described here, will become extremely important in verifying the integrity of MC&A data describing DOE's nuclear material in storage.

References

1. Rena Whiteson, Ferman Kelso, Lisa Spanks, Tresa Yarbro, and Chris Baumgart, "Anomaly Detection Applied to a Materials Control and Accounting Database," Proceedings of the 36th Annual Meeting of the Institute of Nuclear Materials Management, Palm Desert, California, July 9-12, 1995, pp. 1254-1261.

2. R. Whiteson, D. Martinez, B. Hoffbauer, T. Yarbro, and C. Baumgart; "Detecting Anomalies in a Materials Control and Accounting Database," Proceedings of the 37th Annual Meeting of the Institute of Nuclear Materials Management, Naples, Florida, August 1996, pp. 955-959.

3. R. Whiteson, B. Hoffbauer, T. Yarbro, and C. Baumgart; "Anomaly and Error Detection in Computerized Materials Control & Accountability Databases," Proceedings of the 38th Annual Meeting of the Institute of Nuclear Materials Management, Phoenix, Arizona, July 1997.

*This work supported by the US Department of Energy, Office of Safeguards and Security.

**Operated for the US Department of Energy under Contract No. DE-AC04-76-DP00613.