Information Systems Audit and Control Association InfoBytes

Data Mining and the Auditor's Responsibility

By Bob Denker, CISA, CIA, CFE

> *The beauty of the new system, from Mr. Morse's perspective, was that it enabled him to scrutinize the debit and credit side of transactions. By clicking on a number for an expense on a spreadsheet, he could follow it back to the original journal entry—such as an invoice for a purchase or expense report submitted by an employee, to see how it had been justified.... By the first week in June, Mr. Morse had turned up a total of $2 billion in questionable accounting entries.*
>
> *Having found the evidence, the audit team was suddenly faced with how serious the implications of their endeavor really were.*
>
> *Mr. Morse, 41, was known for his ability to use technology to ferret out information.... Mr. Morse grew increasingly concerned that others in the company would discover what he had learned and try to destroy the evidence, he says. With his own money he went out and bought a CD burner and copied all the incriminating data onto a CD-ROM. He told no one outside of internal audit what he had found.*[1]

Every internal audit officer should read the article from which the above quotation was cited. While not every audit staff can consist of diligent auditors like Gene Morse or his boss, Cynthia Cooper, vice president of audit at WorldCom, the auditors the staff does contain can be trained and educated to perform their audits in a similar, meticulous manner.

It is easy to become complacent when one can boast of an audit staff consisting of a handful of CPAs, CISAs, CIAs and possibly a CFE. In addition, there may be a technical guru who has written a series of ACL batch routines, which is far from being sufficient. Having the expertise and knowing how, when and where to apply that knowledge are not the same. This article hopes to educate the internal audit department decision-makers on how best to use their limited resources.

The biggest mistake many companies make is to entrust the audit software to a technical support staff. Senior management sees this as a cost saver by virtue of not having to train and support internal auditors to perform audit analysis routines themselves. Audit analysis software, such as ACL or IDEA, is best used as a sensitive analysis tool, i.e., continually tweaking the variables (the author supports the concept of continuous auditing but it will not be discussed in

this article). While running batch routines is sufficient to validate prior assumptions, it will not replace the iterative analytical process. In addition, the auditor best learns about business rules and processes by taking a hands-on approach.

The truth of the matter is, and many technical auditors will agree, the majority of extraordinary audit findings are accidental. As an example, several years ago the author of this article was performing an audit of a suspense account system and found no —material— suspense items. However, digging deeper into the suspense files, several hundred small-dollar items (some less than a dollar) were uncovered that were two years old or more. This discovery was made shortly after the accounts receivable (A/R) staff indicated that they reconcile most suspense items every two to three months. Had the audit been conducted in the more traditional manner, i.e., looking only for large monetary items, the problems with the A/R department never would have been uncovered.

The problem with most audit departments is that they are locked into the traditional ways of performing audits. Yes, many chief audit officers have embraced analytical software, but they are merely tools and not the true paradigm shift in audit methodology that is required today. Audit departments must learn to think outside the box.

Auditors should perform all of the analytics themselves, and they must be educated in fraud detection and introduced to data mining techniques. When the concept of data mining is brought up, audit managers cringe and argue that they cannot afford to employ statisticians. However, while there is data mining software that requires a statistician's level of expertise (such as IBM's Intelligent Miner), there also are products, such as WizRule from WizSoft Inc., that can be employed by most auditors who are acquainted with the fundamentals of Microsoft Office and who are curious as to why they obtained their audit results.

What Is Data Mining?

Data mining is the process of extracting knowledge hidden in large volumes of data. The data mining tools look for trends or anomalies without knowledge of the meaning of the data. Data anomalies are not necessarily the result of fraud, but can be the result of a range of different factors. In many cases they are caused by faulty data entry, where the user has typed in one value instead of another. Also, errors sometimes are the result of software or hardware malfunctions, resulting in corrupted data. Obviously, such errors can cause considerable damage, which cannot be easily measured but which could be of serious proportions, resulting in direct loss of both income and reputation. In other cases, errors are made intentionally.

Data Mining Applications

While many of the applications of data mining concern market analysis and customer retention, there are numerous applications specifically for fraud detection and prevention. It should be emphasized that the user does not normally specify the fields to look for in the relationships; rather, the data mining program reads all of the data and determines if relationships do exist among data fields. In the case of WizRule, relationships can include any of the following:

- Formula rules
- If-then rules
- Spelling rules

An example of a formula rule is:

A = B * C

Where: A = Total
B = Quantity
C = Unit price

An example of an if-then rule is:

If Customer is **Summit** and Item is Computer type A,
then the price = **765**.

An example of a spelling rule is:

The value **Summit** appears **52** times in the **Customer** field.

There are two case(s) containing similar value(s), such as Sumit and Sumitt. These rules mainly are presented to reveal cases of misspelled names. A name is suspected as misspelled if (a) it is similar to another name in his field, or (b) if the frequency of the first name is very low, while the frequency of the second name is very high.

These are extremely significant findings in that many audit programs are based upon specific searches such as "Show me all of the open accounts in Cleveland." Individuals wishing to perpetrate a fraud and who can override edit routines could easily hide their activity by changing "Cleveland" to "Clevelland" and thus escape detection.

Here is a short list of data mining applications that should gather the attention of the audit selection maker:

1. Human resources—Employees earning salaries inconsistent with their title; employees not availing themselves of benefit programs (perhaps to maintain as much anonymity as possible); employees whose household

address matches an address from the vendor file; employees appearing more than once on umbrella security files

2. Financial applications—Structured transactions (clients who make cash/travelers check/money order contributions to annuities, single premium life insurance, IRAs, mutual funds, etc.) in aggregate amounts that exceed the US $10,000 reporting threshold; clients making contributions to investment vehicles that are disproportionate to their income
3. Medical/dental applications—Patient substitutions; over-utilization of specific diagnoses inconsistent with the patient population; excessive number of patients traveling great distances to a provider (could indicate provider utilizing a postal drop site); provider open seven days a week for disproportionate number of non-emergency procedures (could indicate provider is filing false claims and is spreading out the submissions to divert suspicion)
4. Assistance in due diligence testing—By revealing the business rules, data mining tools can be used to train new auditors and, for new areas or new systems that are being audited for the first time, they are the ideal application to be used for due diligence testing.
5. Construction and purchasing—Payments made earlier than the contract specification date; invoices for large purchases made at the end of fiscal accounting period; price of goods inconsistent with industry costs

With the introduction in the US of the Sarbanes-Oxley Act of 2002 and the implementation of HIPAA (Health Insurance Portability and Accountability Act of 1996), audit committees will not accept excuses from the audit department that it was unaware of major fraud occurrences. It is time to have the chief audit officer address the issue of an enlightened audit staff.

Endnotes

[1] Pulliam, Susan; Deborah Solomon; "Uncooking the Books—How Three Unlikely Sleuths Discovered Fraud at WorldCom," *Wall Street Journal*, 30 October 2002

Bob Denker, CISA, CIA, CFE
is an independent audit consultant who specializes in teaching audit software and fraud investigation applications. He has more than 32 years of information systems experience, with 13 years in auditing and special investigations. In addition, he has taught courses in data analytics, fraud detection and fraud investigation. He is currently an adjunct professor of information systems at Baruch College of the City University of New York, USA. Denker may be contacted at **profdenker@aol.com**.