Integrated equity applications after Sarbanes–Oxley

O. Criner[1] & E. Kindred[2]
[1]*Department of Computer Science, Texas Southern University, USA*
[2]*Software Engineer, USA*

Abstract

Primary among the requirements of the Sarbanes–Oxley legislation are that chief executive officers and chief financial officers certify the accuracy of their corporations' financial statements. This act spawned a thrust to complete total accounting systems with end-to-end financial audit capabilities. The federal government's use of XML and XBRL will eventually be extended to require that all public companies file all forms and reports with them using XML or XBRL. The Securities and Exchange Commission (SEC) is currently accepting XBRL filings from corporations on a voluntary basis. The potential improvement and analytical capability offered by this new environment requires the planning and implementation of new software for computational science research. This paper discusses the technological convergence that allows the implementation of systems that more accurately and rapidly monitor the performance of public companies through their SEC filings and news events.
Keywords: Sarbanes–Oxley, XBRL, XML, computational modelling, accounting data integrity, financial forensic analysis.

1 Background

The Sarbanes–Oxley Act was passed by the United States Congress in the aftermath of several corporate scandals involving large public corporations during and after 2001. This research topic became of interest to the senior author because of his involvement as a juror in one of the high profile trials of that time. Several questions arose during that trial concerning the accuracy of accounting information, accounting procedures that thwarted transparency, the integrity of corporate financial data, and the uses and limitations of mathematical and

www.witpress.com, ISSN 1743-355X (on-line)
doi:10.2495/CF060081

computational tools in finance. As evident by the unusually large number of corporations restating their financial results, there must have been a widespread practice of manipulation of records or “cooking of the books”. The obvious intent of this manipulation of reporting was to affect equity market prices, since the compensation of many executives was directly connected to the price of the stock. Almost all of the firms involved in the scandals were audited by major public accounting firms several of which were also found culpable in the affairs, because of co-optation of the audit process by a consulting relationship with the audit client. Further complications were caused by the implicit conflict of interest that existed between the equities research analysts and the investment bankers involved in business transactions with the corrupt or failing companies and may have been enablers of the corrupt practices. The regulatory agencies of the Federal and State governments were officially unaware of the crisis in the making although some regulators sensed an approaching economic problem. [1].

In an era of ubiquitous anytime computing, the question of why these companies and their questionable business practices were not identified by the securities police, the Securities and Exchange Commission (SEC), remains. The answer lies in both the inability of the SEC to effectively monitor these companies by timely analysis of the thousands of reports and in the islands of automation that exists throughout most of the business world, specifically the separation of operational from financial accounting systems. This situation brings into question the integrity of all public financial data and, in particular, the prices of publicly traded instruments.

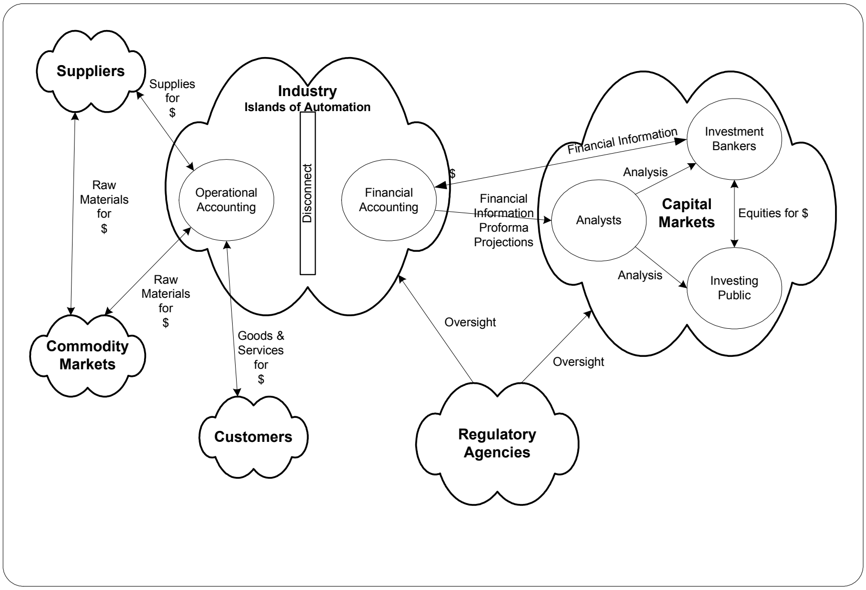

Figure 1: Islands of accounting automation.

The dichotomy between the operational accounting systems and the financial accounting systems shown in Figure 1 creates the greatest problem of integrity and accuracy in financial data for most firms. Sarbanes–Oxley is intended to ensure that investors and stakeholders have accurate financial information upon which to base financial decisions. The act required that chief executive officers and chief financial officers certify the accuracy of a company's financial statements. It provides for severe penalties for knowingly and wilfully misstating financial statements. Satisfaction of these requirements of the law requires that companies institute new controls and data integration between the two islands. Since there is rarely integration connecting the two, most companies rely upon manual processes (with personal productivity tools) to produce the accounting reports. "There have not been major expenditures for new systems since the Y2K effort, so one can only assume that these data and integration problems have existed for some time." [2].

Since computational finance is predicated upon the assumption that good reliable financial data is availability, the entire research enterprise is threatened if that is not the case. Therefore, it should be of great interest to computational finance practitioners to know the sources and processes of the data generated by so many companies in so many variations of the accounting process, which finally ends in the earnings per share value and other parameters used to specify corporate performance [3].

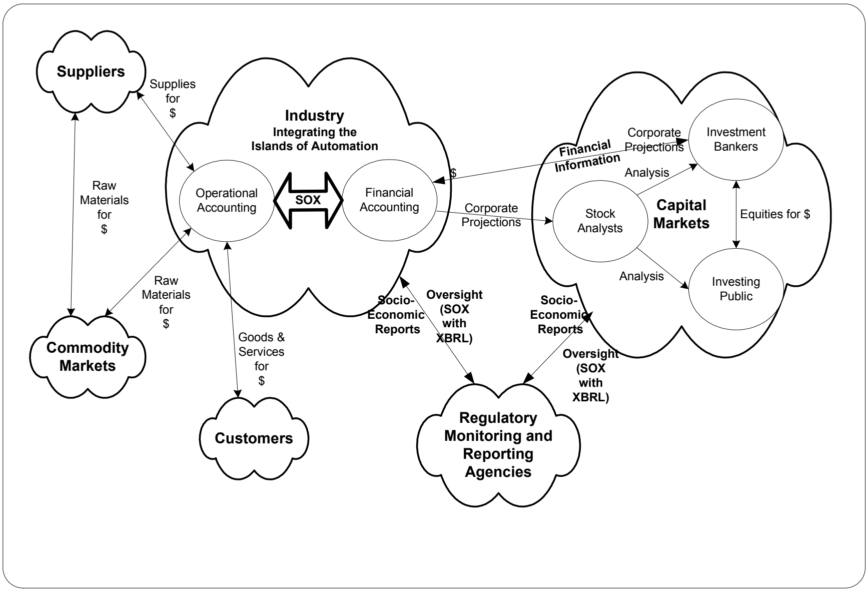

Figure 2: Islands of automation with a SOX bridge.

Some companies appear to have integrated their operational and financial accounting and information systems. This was the competitive edge strategically deployed that enabled Wal-Mart to capture so much of the retail market. The bridge between the two islands is motivated by the SOX legislation. But the magnitude of the SOX problem has some companies lobbying to have some relief from the requirements.

The application of scientific methods to business processes requires that the source of the information be somewhat accurate and standard. The facts are that the results of the accounting and auditing processes are approximate and far from precise. Trust in the financial reporting system is fundamental to the capital market system. In order to ensure the trust in the system, there needs to be much more accountability in the monitoring system. This implies that financial reports should be examined differently utilizing the technology of the time. Companies are required to redefine their operating processes so that auditors can assess the effectiveness of their internal controls. Companies must define seam-less systems that integrate and preserve the audit trail for the thousands of processes and the millions of transactions they generate that affect the financial statement

This effort has developed more slowly than we had anticipated in 2002, however, in the near future automated agents will be utilized to mine report databases and examine all corporate reports filed with regulatory bodies [4, 5, 6]. This technology must be incorporated into the analytical capability of the investing public in order for the public to be able to evaluate the viability of a firm for investment. The Extended Business Reporting Language (XBRL), a derivative of XML will be the required format for corporate reporting in the near future. The SEC is required to monitor public corporations to ensure that the investing public is not defrauded, a task at which the agency has not been very effective and has been prevented from being so in large part by the business lobby.

Web Services provide the architectural framework for new integrated applications in financial information. By utilizing the new XBRL language and the infrastructure of XML, it is possible to integrate equities analysis in a totally new framework. When this research was planned, the writers assumed that the business community would embrace XBRL and XML as widely as the federal government. Unfortunately, this is not yet the case and many companies seem to view the Sarbanes–Oxley requirements as an unnecessary expense rather than an opportunity to make a commitment to be a full participant in the world of e-commerce.

2 Computational modelling of financial markets

This paper discusses a computational approach that integrates the financial reporting with the analyses of price time histories with the objective of identifying the signatures of corporate events. By identifying the signatures of corporate events in the data derived from the market, it may be possible to classify the response to such events and to assess their effect upon the prices in the market in a deterministic manner. It is well known that various events affect the price of equity issues and futures prices, e.g., announcements of various government indices, earnings announcements, bankruptcies, and credit rating changes. Traders and investors are known to wait for the reports before taking some action. We want to take the output of the system and work backwards to determine its cause. This is an inverse problem in dynamic systems [7].

Computational modelling is used in many fields where there are not sufficient data and theory; it is an application of logic, mathematics, computational techniques, and heuristics. Computational scientists usually consider very large complex problems that usually do not yield to a complete mathematical analysis, fit neatly into a theory, or can be examined in a laboratory. The problems considered by computational scientists are not amenable to the traditional scientific method of observation, theory and experimentation. Indeed the usual data that one needs for a well-posed problem generally does not exist nor do many of the equations or inferential schemes.

The "direct" or "forward" approach to problems in science is the situation where there is a "complete description" of a physical system within the confines of some logical system, which provides the rules of inference sufficient to derive additional true statements in the logical system, which correspond to the prediction of some observed events. In most cases, the logical system is expressed in mathematics. However, mathematics is not the only implementation of a logical system with which to study complex phenomena. Computational modelling extends the mathematical analyses beyond the so-called well-posed problems or it may be a completely heuristic set of processes. In inverse problems the issue is to use "the actual results of some measurement to infer the values of the parameters that characterize the system" [7]. In computational finance those measurements include the publicly reported financial data, which is why there is concern as to its accuracy and integrity.

We generalize the logical system in computational modelling to be comprised of five components: (1) Definitions are descriptions of the objects under consideration, (2) Assumptions are true statements that are known about the objects and taken a priori, (3) Rules of inference that describe the process of taking the definitions and assumptions as inputs and concluding a new true statement as output from the process, (4) The collections of theorems or true statements that are derivable from the definitions and assumptions using the rules of inference and the collection of derived true statements, and (5) associative relationships or alternative paths through the inferential process to obtain the same true statements.

Computational finance is a specialization of computational science, in the sense that scientific computing or the solution or investigation of scientific problems is done using computers. Computational finance is the application of computational science to problems and issues of financial systems. Computational science is third modality of knowledge determination inextricably co-equal with experimentation and observation, logical inference and theoretical analyses. In this sense, computational finance is not finance. Corporate finance provides one of the datasets for computational finance, but the computational models of financial systems are much more general.

The XBRL Taxonomies [6] effectively partition the set of corporations into equivalence classes. Each company that uses the standard taxonomy for a particular group is equivalent to other company in the class from a computational point of view. Practical applications of this technology may be a few years away, however, because of the size and effort required to construct systems that utilize Web Services and data bases implemented with XBRL. Design and development must begin early.

3 The dynamic model - Integrating the price and financial time series

In this paper we describe a computational methodology for integrating the price time history of an equity with its "fundamental" financial reports. We create a new model integrating analysis of equities that is based upon estimating the rates of change of the price time series and the affect of corporate events. This dynamic model of price could be given by an equation of the form

$$\ddot{x}_i(t) = f_i(x_i(t), \dot{x}_i(t)) \tag{1}$$

where $x_i(t), \dot{x}_i(t)$ and $\ddot{x}_i(t)$ are the price, first derivative of the price, and second derivative, respectively, of the price of the i^{th} stock [9]. To include the financial analysis results into this model, we assume that eqn (1) can be rewritten in the form

$$\ddot{x}_i(t) = f_i(x_i(t), \dot{x}_i(t), q_i(t), k_i(t), e_i(t)) \tag{2}$$

where the additional functions $q(t)$, $k(t)$, and $e(t)$ are to be determined using the results of the quarterly, annual, and event reports, respectively, and the so-called "analysts expectations" and news releases.

We seek to determine signatures of the various announcements and events by mining the data available in the SEC EDGAR database. This is done by examining the time series of the price in the neighbourhood of the event. While the usual reaction to announcements or events is a rapid change in the market values of the instrument, we seek to quantify that change in the derivatives of the price and to estimate the effect on the price movement. In this sense we parameterize the functions $q(t)$, $k(t)$, and $e(t)$. The parameterization process in this methodology correlates the time and magnitude of the various events with the first or second derivatives of the price of the stock issue as suggested by the model eqn (2).

Relating these magnitudes to the estimates of $\dot{x}(t)$ and $\ddot{x}(t)$ will provide a measure of the effect. For example, the earnings estimate at time t will be paired with a value of $\dot{x}(t)$ and $\ddot{x}(t)$ to create a function relating earnings to the second derivative of the price. Averaging these measures over time will provide values of the function that are used in the decision algorithms. Figure 3 shows the basic components of the computational model and where the additional components are combined with $x(t), \dot{x}(t), \ddot{x}(t)$, the volume, 10-Q, 10-K, and 8-K reports.

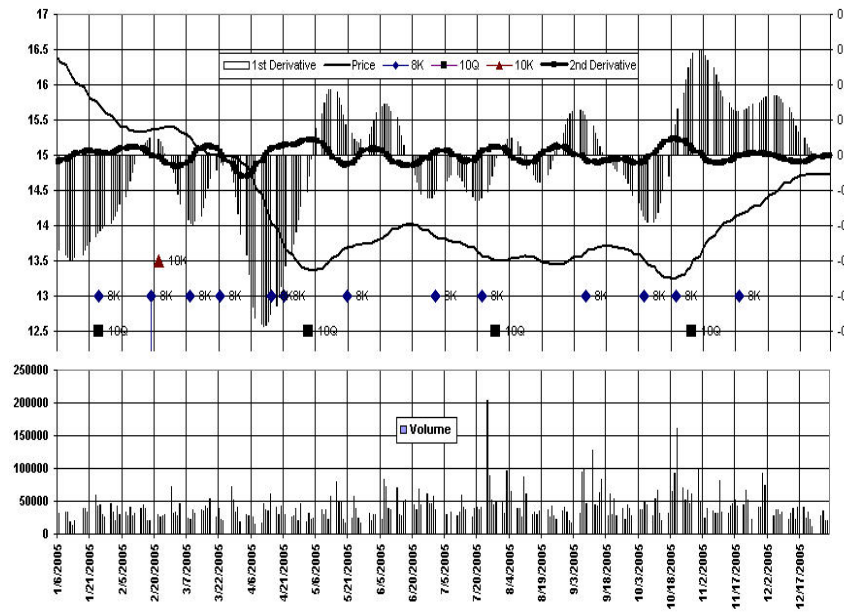

Figure 3: Basic components of the computational model of Xerox stock.

Figure 4 shows phase diagrams for Xerox stock for three complete years, 2003, 2004, and 2005 and the first two months of 2006. The variables of a dynamic system specify its phase space. The motion of the system corresponds to a trajectory or orbit in the abstract phase space [10]. Since we do not know the specific functional relationships of each variable, we must investigate the manner in which the diagrams vary as a result of specific events that occur. Deterministic dynamical systems are characterized by their phase plane orbits. Clearly these diagrams show that our assumption that the process is a dynamical system has merit. We want to discover or synthesize some process that simulates the system in the time domain. Although the graphs show that the system is highly nonlinear, it is not clear that it is chaotic. If it is chaotic then it may be possible to find an attractor to which the process tends. Many questions arise in this case. Are there attractors for each stock or equivalence class of stocks? Are attractors time dependent or do they depend on other parameters? One issue that should be settled by the capacity to construct phase plane diagrams is that the process is deterministic. This demonstration should set the efficient market hypothesis to rest.

Many other components can be added including analysts' estimates and government reports of the essential economic indicators to model their effect on the price of major companies and industry groups. This procedure will also be helpful in forensic analyses because it will create a time series of the essential components of the company's financial statements [11, 12].

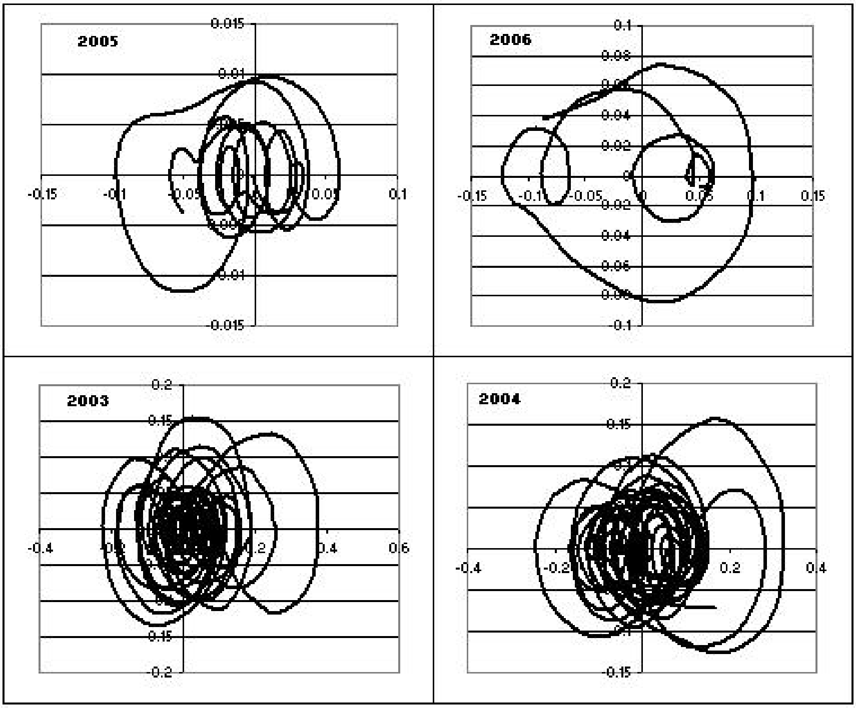

Figure 4: Phase plane plots of Xerox stock.

4 The technological convergence

With the pervasiveness of the Internet and the availability of CPU cycles and massive storage devices, the technology is now present to implement the underlying infrastructure. And of course, the lingua franca that binds the disparate entities of the business community is XBRL. With these technological components at our disposal, a re-examination of Figure 1 yields Figure 2 – Integrating the Islands.

Looking at most industries, there exists ample opportunity for real-time or near real-time data collection in their operations and supply chain as evidenced in retail sales by Wal-Mart and Home Depot. Transaction data can be collected from customers and suppliers for sales and inventory using point of sale (POS) technologies like barcodes and radio frequency identification (RFID). Location and condition within the supply chain can be tracked using the global positioning system (GPS), automated weighing systems for bulk supplies, and remote sensing for environmentally sensitive resources. This data is then fed to the operational accounting system, which hopefully optimizes its operating practices and minimizes its operating costs with statistical improvement methods like Six Sigma.

The transaction, supply, and operations data are then transferred to the financial accounting system, which optimizes its budgetary, capitalization, and investment activities based on "known" corporate requirements. The condition of the company can now be reported internally to the corporate executives and externally to industry analysts and investment bankers in the Capital Markets, to the Investing Public, and to the various Regulatory Agencies. The numbers will be traceable and have meaning and fulfil the theoretical goal of accounting – to represent the operations of the business. The technologies for this market communication are the industry-specific taxonomies implemented in XBRL.

Once these reports are collected by the Regulatory Agencies in a format that can be data-mined, the monitoring function can be automated utilizing a collection of techniques suggested in this paper. Data-mining results can be used for highlighting potential problems and by the Capital Markets and the Investing Public for investment decisions. Results can be disseminated by Web Service-based applications. By no means is the widespread implementation of these technologies trivial but, with an evolutionary approach, financial information will have real meaning and integrity will return to our markets.

On top of this highly networked infrastructure lies a plethora of computationally intensive techniques:

- Data-mining – to draw relationships between quantified data
- Red Flag Analysis – to identify stellar or troubled companies and industries
- Natural Language Processing – to quantify prose reports
- Chaos – to graphically represent interpretations of complex datasets
- Heuristics – to build systems incorporating expert domain knowledge
- Grid Computing – to provide the computational capacity on the desktop or across the enterprise

5 The grand challenge of computational finance

The scientist always asks, "How good is the data with which I am working?" This is not a statistical question but a question about the process of measurement. How is this data being created? Regardless of the sophistication of my analytical tools, the old computer science dictum "garbage-in-garbage-out" still applies.

Secondly, it is now possible, or will be in the near future, to analyze every formally traded financial instrument (stocks, bonds, futures, options and other formally traded derivatives), in the light of the publicly available socio-economic data to determine causal relationships among them, for example:

- Land and water use and commodity production
- Non-renewable resources and population growth
- Environmental preservation and economic growth

It is imperative that all computational science be conducted in an environment of high quality data. The ubiquity of these data from the media, the web, and the press assaults the senses and cries out for computational solutions.

References

[1] Levitt, A. & Dwyer, P., *Take on the Street, What Wall Street and Corporate America Don't Want You to Know*; *What You Can Do to Fight Back*, Pantheon: New York, 2002.

[2] Get Ready for an Increased Number of Financial Restatements, Visage Solutions Web Site, http://www.visagesolutions.com/

[3] Berenson, A., *The Number: How the Drive for Quarterly Earnings Corrupted Wall Street and Corporate America*, Random House, New York, 2003.

[4] Bovee, M., Kogan, A., Nelson, K., Srivastava, R. & Vasarhelyi, M., Financial Reporting and Auditing Agent with Net Knowledge (FRAANK) and eXtensible Business Reporting Language (XBRL), *Journal of Information Systems*, 19(1), pp. 19-41, 2005.

[5] Lawler, J., et al, A Study of Web Services Strategy in the Financial Services Industry, *Proc. ISECON*, v21, 2004.

[6] Leinnemann, C, Schlottmann, F., Seese, D., & Stuempert, T., Automatic Extraction and Analysis of Financial Data from the EDGAR Database, *Proc. 2nd Annual Conference on World Wide Web Application*, Johannesburg, 2000.

[7] Tarantola, A., *Inverse Problem Theory and Methods for Model Parameter Estimation*, Society for Industrial and Applied Mathematics, Philadelphia, 2005.

[8] XBRL International Web Site, www.xbrl.org

[9] Criner, O., Optimal control strategies for portfolios of managed futures, *Computational Finance and Its Applications*, WIT Press, Southampton, UK, pp. 189, 2004.

[10] Saaty, T., & Bram, J., *Nonlinear Mathematics*, Dover Publications, New York, Chapter 4, 1964.

[11] Looking for Trouble: The SEC Upgrades Technology to Be a Better Watchdog, http://www.WallStreetandTech.com/showArticle.jhtml?articleID=41600001

[12] Apostolou, N., & Crumbley, D., Forensic Investing: Red Flags, http://www.bus.lsu.edu/accounting/faculty/napostolou/forensic.html