© 2011
YONGBUM KIM
ALL RIGHTS RESERVED

www.manaraa.com

Continuous Monitoring: Macro- and Micro-level Control

By Yongbum Kim

A dissertation submitted to the

Graduate School-Newark

Rutgers, The State University of New Jersey

in partial fulfillment of requirements

for the degree of

Doctor of Philosophy

Graduate Program in Management

Written under the direction of

Dr. Miklos A. Vasarhelyi

and approved by

_____

Dr. Miklos A. Vasarhelyi

_____

Dr. Alexander Kogan

_____

Dr. Michael Alles

_____

Dr. Graham Gal

Newark, New Jersey

October, 2011

www.manaraa.com

UMI Number: 3481807

All rights reserved

INFORMATION TO ALL USERS
The quality of this reproduction is dependent on the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript
and there are missing pages, these will be noted. Also, if material had to be removed,
a note will indicate the deletion.

UMI 3481807

Copyright 2011 by ProQuest LLC.

All rights reserved. This edition of the work is protected against
unauthorized copying under Title 17, United States Code.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

www.manaraa.com

ABSTRACT OF THE DISSERTATION

Continuous Monitoring: Macro- and Micro-level Control

By YONGBUM KIM

Dissertation Director:

Miklos A. Vasarhelyi

A company's internal control system is a crucial factor for operational effectiveness and efficiency. A properly designed internal control system adds reliability to a company's financial information by preventing, detecting, and correcting erroneous or fraudulent transactions on timely basis. A series of financial scandals in the late '90s and the early 2000s resulted in the passage of the Sarbanes-Oxley Act (hereinafter SOX) in 2002, mandating public companies to implement, evaluate, and report on the quality of their internal control systems. Although various internal control requirements are mandated by SOX, there is a lack of clear and detailed guidelines about what and how companies should implement internal control mechanisms into their systems. This study intends to shed some light on these issues by proposing and testing two anomaly detection models utilizing transactional data from a bank and an insurance company.

www.manaraa.com

This study makes several contributions to research on anomaly detection. First, this study provides a detailed guideline for the development of an anomaly detection model that is implementable and understandable by internal auditors. Second, it proposes anomaly detection models at a transactional level with unlabeled (i.e. unclassified) data that is more realistic and useful in practice. Finally, it shows that the process of developing an anomaly detection model helps to identify weakly-controlled or risky areas.

The first chapter introduces general concepts about internal control systems and the research purpose of this study. The next chapter summarizes and discusses literature related to internal control systems and anomaly detection. The third chapter consists of two essays that investigate the implementation of anomaly detection models with transactional level data. In these studies, transactions from transitory accounts of a bank and from wire transfer payment systems of an insurance company are investigated to show feasible implementation processes of anomaly detection models. The fourth chapter concludes this study by discussing limitations and directions for future research.

www.manaraa.com

Acknowledgements

I am grateful to Dr. Miklos Vasarhelyi who has provided me with this great opportunity to work on this valuable research in continuous auditing. His encouragement and support during my Ph.D. study at Rutgers University helped me complete this dissertation.

I also appreciate other committee members, Dr. Alex Kogan, Dr. Michael Alles, and Dr. Graham Gal for their helpful and inspiring comments on this study. Their valuable comments help me to build a more systematic structure.

I would like to dedicate this research to my parents, Hyungho Kim and Jongshin Lee, for their love and support throughout my academic years.

Special thanks to my beloved wife, Dr. Kangae Lee, for her love and encouragement during my Ph.D. study.

I want to thank all the friends that have made helpful comments on my study as well as Ph.D. life, particularly David Chan, Sutapat Thiprungsri, and J.P. Krahel.

www.manaraa.com

Finally, I also want to thank the bank and the insurance company that provided me with the data for this study.

www.manaraa.com

Contents

www.manaraa.com

www.manaraa.com

LISTS OF TABLES

www.manaraa.com

www.manaraa.com

LIST OF ILLUSTRATIONS

www.manaraa.com

I. Introduction

The Sarbanes-Oxley Act (SOX, 2002) was introduced after a series of financial scandals at corporations such as WorldCom and Enron. The passage of SOX has had a significant impact on the practice of auditing, notably the newly required evaluation procedure of a company's internal control system by management and to be attested by external auditors. In the pre-SOX era, a company's internal control system was used by management for administrative purposes such as optimizing the usage of their resources. In other words, the main target of an internal control system was operational efficiency. However, SOX requires management to evaluate their existing internal control system and report on its quality, to be attested to by external auditors and made public as part of regular financial statements. Hence, although internal control systems were important pre-SOX, they became even more crucial in the post-SOX era. The key SOX requirements are as follows:

First, SOX imposes a new burden of proof on management. As the first step, a company's management is required to report the existence of an internal control system over financial reporting (ICFR). If there are deficiencies, the management must report the

www.manaraa.com

fact that the company does not have an adequate internal control system that fulfills SOX requirements. If an internal control system exists, management must evaluate how well it works and prepare a report for external users. Once the report is prepared, the CEO and the CFO should certify the report. Improper certification will result in civil punishments even if it is unintentional (Section 302). Management must maintain an internal control system that will satisfy the minimum level of SOX requirements.

Second, SOX requires external auditors to attest to management's report on their internal control system and opine on the report that will be available to external users as a part of the annual report. This requires external auditors to have proper internal evaluation tools and skills in order to accurately estimate the degree of internal control system quality and/or deficiency.

The emphasis on the roles of management and external auditors concerning the establishment and review of internal control systems force practitioners to revise or develop their internal control system evaluation methods. To that end, various attempts had been made in practice. Although detailed and verified reports are rarely available to the public, one benefit realized in the post-SOX era is that practitioners have

www.manaraa.com

implemented internal control mechanisms into their business environment and have made assertions about their conformity with SOX requirements. Furthermore, a third party external auditor has reviewed the clients' internal control systems.

As opposed to industry, only a few models of actual internal control have been proposed in academia. Instead, most research in the post-SOX era has focused on the effects of SOX and/or its requirements on earnings, abnormal accruals, audit fees, and compliance cost in order to show that the internal control quality is informative to its users. This research is certainly beneficial to practitioners and scholars. However, it is also fundamentally important to study the internal control systems themselves as quality, and not the mere existence of an internal control system, affects the quality of accounting information, which is the overarching goal of SOX. Despite the importance of internal control system implementation and evaluations, most internal control systems research has focused only on the theoretical frameworks that can be used as indirect instructions. There is a lack of concrete examples showing how the internal control system is implemented and evaluated in practice. This might be due to a lack of data from practitioners and/or insufficient prior research.

www.manaraa.com

This study will make an attempt to fill the gap by developing and testing internal control screening models for a bank and an insurance company. This research will shed some light for future researcher on internal control system modeling and evaluation.

www.manaraa.com

II. Literature Review

1. Internal Control

In this sub-section, prior literature relevant to internal control will be discussed. First, it addresses the reasons why internal controls are important, how legislators have responded to financial scandals, and how practitioners think about legislators or internal controls. The second sub-section summarizes the prior research that directly concern about internal control system relevant topics such as internal control system usage and evaluation. In addition, some ideas about ICS evaluation framework are discussed and suggested.

i. Background

a. Definition & History

Before proceeding, it is necessary to define the taxonomy that is used in this section. First, SAS 78 defines internal control as a process designed to provide reasonable assurance regarding the achievement of objectives for reliable financial statements, effective and efficient operations, and compliance with applicable laws and regulations.

www.manaraa.com

Second, an internal control procedure (hereinafter ICP) is a single control measure, such as the checking of a control total (Cushing, 1974). An ICP can be characterized as preventive, detective, and/or corrective. A preventive control is used to reduce the probability of an error occurring, a detective control is used to determine the actual frequency of errors in the system, and a corrective control establishes steps to be taken when a control violation is detected. Third, an internal control cluster (hereinafter ICC) consists of one or more ICPs related to one or more types of error or activity. In other words, an ICC is related to a particular cycle of the business organization such as accounts receivable. Fourth, an internal control system is the set of ICCs that represents the overall system of a company (Vasarhelyi, 1980).

The quality of a company's internal control system is a crucial factor for operational effectiveness and efficiency. The main goal of internal control system is to increase the level of company's system reliability by preventing, detecting, and correcting potentially material errors and irregularities in the system. An ideal internal control system with high compliance would guarantee that the company's financial information is fairly represented, requiring far less work during audit procedures. Despite its importance, it

www.manaraa.com

was not too long before its evaluation became statutorily mandated (Yu and Neter, 1973; Kinney, 1975; Gadh et al., 1993).

A company's internal control system is the result of management actions and policies that purport to prevent, correct, and/or detect errors and irregularities (Felix and Niles, 1988). Although the design of internal control systems may not be a typical audit activity, the evaluation of internal controls would be a part or main portion of internal auditor's work.

Evaluation of a customer company's internal control system has played a key role during audit procedures because effective and efficient internal control system allows an auditor to reduce the scope of subsequent audit work (i.e. substantive testing). The evaluation of internal control system became even more important due to recent financial scandals. To restore trust in publicly traded corporations, management, financial statements, and auditors that was deteriorated by the scandals, the SEC released the Sarbanes-Oxley Act (2002). SOX Section 404 requires that management make an internal control assessment and include that assessment in their annual report to shareholders, and that external auditors attest to and report on management's internal control assessment.

www.manaraa.com

Contrasting with the growing interest in internal control evaluation, there is relatively little factual data available to confirm or deny the efficacy of the interrelated internal control components embraced by Committee of Sponsoring Organizations (COSO) and codified in the professional standard under SAS 78 (Geiger et al., 2004). Furthermore, there are concerns about evaluation of internal control quality. There are practical barriers to effective internal control quality evaluation, especially a lack of adequate criteria for measuring internal control quality (Kinney, 2000).

According to SAS 55, a company's control environment includes the overall attitude, awareness, and actions of the board of directors, management, and owners, and the plans and procedures that are considered for evaluation of internal control system. Even if the control environment may not have a direct effect on the accuracy and completeness of the financial statements, it will significantly affect the internal control system per se and its compliance level.

b. Legal issues: Foreign Corrupt Practices Act (FCPA) and Sarbanes-Oxley Act

The earlier legal recognition about the importance of internal control system prior to

www.manaraa.com

the SOX was the Foreign Corrupt Practices Act (FCPA, 1977), created after a slew of highly publicized corporate malfeasance. The FCPA requires an organization to establish and maintain adequate internal control system and specifies penalties for violators (Merten, 1981).

Similar to its precedent, Sarbanes-Oxley Act was also initiated by a series of financial scandals around 2000. One of the SOX requirements is to include a management report on internal control (MRIC) in the annual report. Although some practitioners might argue that voluntary reporting will be sufficient, evidence shows that this is not the case. Prior to SOX (2002), McMullen et al. (1996) investigated companies that voluntarily chose to issue MRIC. The results clearly show that only companies with no significant IC problems select voluntary MRIC. Although mandatory MRIC might not eliminate all IC related problems, it would at least reduce the degree of seriousness.

Because of radical changes in the data processing environment, the traditional approach to general/application controls might no longer be relevant and might, in fact, introduce weakness in a modern data processing environment (Wu and Hahn, 1989).

In response to increasing complaints from practitioners that criticize unreasonably

www.manaraa.com

high SOX compliance costs, the IIA (Institute of Internal Auditors) advocated SOX by providing some evidence that SOX compliance provides measurable benefits. The survey with 171 chief audit executive members of the IIA, showed that the cost for SOX compliance generated reasonable benefits (Rittenberg and Miller, 2005). Chief benefits include: 1) more active participation by the board, the audit committee, and management, 2) more thoughtful analysis of monitoring controls, 3) greater understanding about company processes, 4) implementation of anti-fraud activities, 5) better understanding of the risks associated with general computer controls, 6) improved documentation of controls and control processes, 7) improved definition of controls, and the relationship of controls and risk, 8) control concepts becoming embedded into the organization, 9) improvements in the adequacy of the audit trail, and 10) re-implementation of basic controls, e.g., segregation of duties, periodic reconciliation of accounts, and authorization processes that had been eroded as organizations downsized or consolidated operations.

However, SOX compliance costs such as internal control system implementation and maintenance will be in debate until relatively objective internal control system evaluation guidelines are offered and minimum levels of internal control system implementation

www.manaraa.com

models are provided.

c. Practical issues

Computer systems are no longer optional in the modern business environment. Instead, they are an indispensable tool to increase management effectiveness. In addition to increased efficiency, systems have also introduced unforeseen negative consequences (Merten and Severance, 1981; Whang et al., 2004).Although the newly adopted IT significantly accelerated transaction processing, it also reduced the time available to review those transactions (Jancura and Lilly, 1977). Furthermore, to increase operational efficiency, many traditional procedures that were considered as inefficient or unnecessary have been virtually eliminated from a company's system. For example, in an accounting system, posting ledgers is not explicitly performed. Instead, this procedure is handled via electronic data processing. Although fewer procedures lead to a more efficient system, they also leave less information available for audit and control purposes, increasing control risk and, consequently, overall audit risk. In addition, traditional internal controls may not be effective in a computerized environment (Whang et al., 2004). This problem

www.manaraa.com

arose in an electronic data processing environment in which visible audit trails are hardly available.

Some scholars such as Jensen and Payne (2003) argued that a company's internal control system could be viewed as an alternative that management may choose for operational efficiency. This should be revised to extend its scope since the Sarbanes-Oxley Act (2002). The main motive of their argument was that the cost of an internal control system implementation was larger than its commensurate benefits so that management might choose to maintain an internal control system at the level where estimated benefits exceeded the costs. Their view may support recent arguments about the effectiveness of SOX requirements. From management's perspective, internal control system implementation might mean additional cost to the company (Felix and Niles, 1988; Raghavan, 2006). If we assume that the company does not have any IC problems even in internal control system-free environment, this argument may be more convincing. Despite emerging criticism about the adequate internal control system implementation and maintenance, it remains true that properly designed internal control systems bring competitive advantages to companies (Raghavan, 2006). Campbell et al. (2006) claim

www.manaraa.com

that a market leadership position could be achieved through the significant competitive value delivered by internal control systems. Furthermore, this leadership with respect to SOX compliance may draw favorable expectations from various stakeholders such as analysts, rating agencies, customers, clients, and suppliers.

In practice, companies may implement a variety of ICPs in their internal control systems, likely due to the differential characteristics and environments of their businesses. Consequently, individual ICPs may differ with the size, nature, and/or complexity of given businesses (Wu and Hahn, 1989).

Although SOX enforces various internal control system requirements internal control system and today's businesses utilize computerized data processing systems, according to CFO magazine's IT survey in 2005, IT was the main contributor to IC problems. Furthermore, auditors did not seem to have sufficient understanding about the IT components of internal control systems (Raghavan, 2006).

www.manaraa.com

ii. Prior Research

a. Auditor's Judgments

Prior to the SOX, internal control system quality assessment was mostly used for external auditors to determine the audit scope and provided inputs for the management letter. Since quality assessment is inherently subjective, it will be useful to discuss how experts judge companies' internal control systems.

Prior research found that while auditors generally agreed on assessments of internal control systems, their subsequent decisions on audit program planning tasks usually had low consensus (Ganumnitz et al., 1982; Tabor, 1983; Biggs and Mock, 1983). In other words, auditors might not appropriately utilize internal control system evaluation when determining the scope of substantive tests. Based on an experiment with 35 auditors, Ganumnitz et al. (1982) showed that high correlation between internal control system assessment and subsequent audit tasks was obtained when auditors explicitly recognized that inverse relationships should exist between two procedures.

Because of a lack of normative or standard internal control system evaluation criteria, the most widely used method to develop internal control system models is to examine the

www.manaraa.com

consensus among auditors. The consensus of evaluation of the auditors' decisions on given IC components has drawn much attention from scholars (Ashton, 1974; Joyce, 1976; Gaumnitz et al., 1982; and Srinidhi and Vasarhelyi, 1986). Auditors' consensus on internal control system evaluation has been used in evaluation of internal control because a normative criterion is not readily available (Srinidhi and Vasarhelyi, 1986). A high consensus (greater than 60%) among auditors can be reached on internal control system evaluation given evidence on IC components (Ashton, 1974; Gaumnitz et al., 1982; and Srinidhi and Vasarhelyi, 1986). However, the level of consensus might not be strong enough to claim that the decision processes of auditors are similar to each another. Furthermore, little is known about the actual procedures and methods that auditors follow when evaluating internal control systems in practice.

Because of the visibility of traditional audit trails in a highly computerized electronic data processing environment, Bailey et al. (1985) showed the possibility that computers might be used an internal control system evaluation tool by introducing TICOM (The Internal Control Model).

www.manaraa.com

b. Internal Control System and its Quality

Quality in any form is inherently subjective. It is therefore difficult to develop objective criteria for measuring quality adequately (Fihn, 2000). However, there is an encouraging motive to find objective measurements. Gilb (2004) claimed that even poor quantification is more useful than none because it at least allows systematic improvement. According to his argument, initial criteria can serve as a benchmark that will be repeatedly adjusted or replaced by better baselines. He also suggested the four steps of quantifying quality: 1) identify known quantification ideas, 2) create initial criteria by modifying the quantification ideas, 3) test the new criteria, and 4) adjust for better fitting. The first and most important step in his suggestion is to figure out known quantification ideas to determine a benchmark of the quality. Hence, the emergent questions about the internal control system evaluation can be clearly raised. What must be known about the internal control system in order to evaluate objectively? Are there any benchmarks to which we can refer?

Although objective evidence is preferred to subjective evidence, the latter cannot be ignored while evaluating internal control systems. For example, the compliance level for

www.manaraa.com

certain ICPs might not be easily estimated despite its importance. Even a perfect internal control system might be useless if the actual users do not perform their control roles. This is even more problematic in a highly computerized environment in which most traditional audit trails are diminished. Furthermore, the reliability of the estimates might not easily obtain consensus among potential users (Yu and Neter, 1973).

Quality measures are generally affected by the structure, process and outcome of an internal control system (Fihn, 2000). Consequently, we may quantify internal control system quality by considering its structure, process, and outcome. First, as for the structure of internal control system, it will be challenging to evaluate the quality of the ICC. In addition, internal control system quality evaluation may require some rules that incorporate all relevant ICCs. Second, we must evaluate the level of compliance on each IC procedure that a company is supposed to follow. Due to cost and time constraints, it may be difficult for auditors to oversee all processes in place. As identified by prior research, segregation of duties is the main factor when an auditor evaluates internal control system reliability (Srinidhi, 1994). Even if an internal control system is properly designed to prevent SoD violations, bypassing the control is possible by using two or

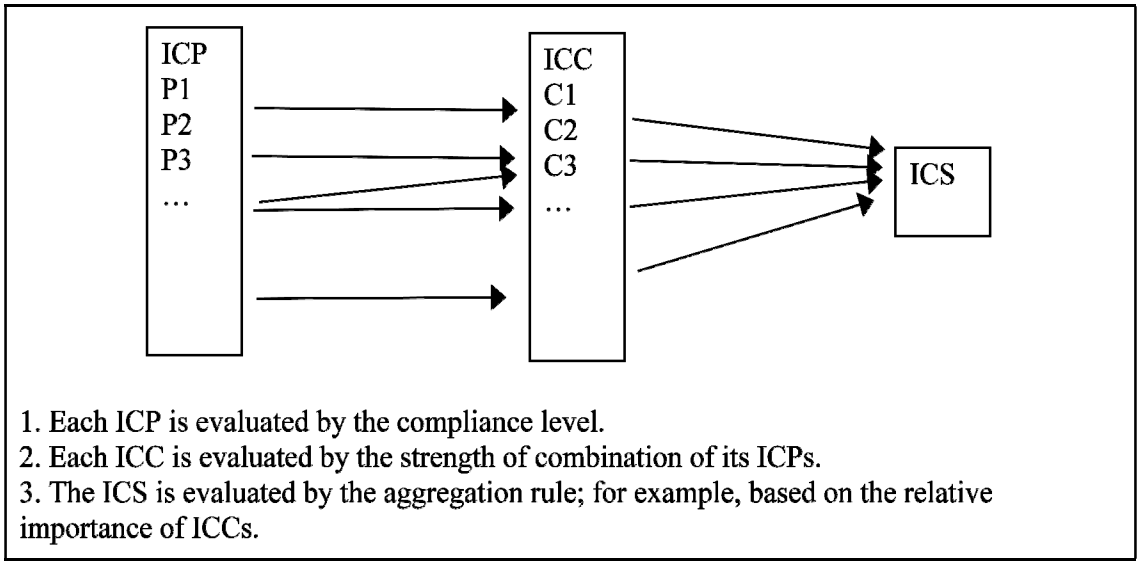
www.manaraa.com

more IDs. Hence, compliance level will be of great importance. Third, problems of a company's internal control system are recognized when they become serious enough to draw concerns of the public. Hence, the actual outcomes of an internal control system will be more likely negative because they are from the companies that experience serious problems. In other words, if problems are not significant, they will be neither recognized nor disclosed. Hence, it will be indispensable to use alternative ways to examine the level of quality of an internal control system.

Since an internal control system can be divided into a number of relevant ICCs that are further divided into multiple ICPs, evaluation should go from the ground up, starting with ICPs.

www.manaraa.com

Figure 1. Components of Internal Control System



The Figure. 1 shows the overall diagram about how to evaluate an internal control system. First, after identifying the extant ICPs of a company, each ICP is evaluated based on its compliance level (Srinidhi and Vasarhelyi, 1986). Compliance level may be a key component to evaluate the reliability of an ICP because even a robustly designed ICP will be meaningless if is not appropriately executed during transactions. However, it is difficult to quantify compliance level because of cost and time constraints. In order to evaluate the accurate compliance level of ICPs, an auditor may have to oversee all transaction processes of a company, a task neither economical nor practical. Hence, it first seems to be practically appropriate to focus on identification of actually existing ICPs (design-focused). However, it is dangerous to decide internal control system quality

www.manaraa.com

without considering the compliance level because it will be possible to issue an effective opinion on an internal control system that is actually not used at all. Hence, compliance levels of each ICP must be taken seriously during evaluating internal control system.

Second, after evaluating the ICPs, each ICC is evaluated according to the strength of the combination of its ICPs. As shown on the figure above, an ICC may consist of one or more ICPs. Conversely, an ICP may be used in one or more ICCs (Cushing, 1974). Kinney (2000) also claimed that there are multiple ways to achieve a given internal quality control objective. As Cushing (1974) stated, each ICP may have different values in different ICCs. Srinidhi (1994) also claimed that different combinations of ICP have different values. He showed that there are ranks among values of ICP combinations for an ICC. Based on those arguments, it may be concluded that there are a number of aggregation rules by which the level of an ICC are reasonably estimated. For example, Cooley and Hicks (1983) suggested a systematic methodology for aggregating internal control system judgment into a meaningful statement as to how the system functions as a unit. One of the problems in evaluating an ICC is the many combinations of ICPs that must be considered (Srinidhi, 1994). Because of the complexity of internal control

www.manaraa.com

processes, an ICC may have different combinations of possible ICPs depending on a variety of factors such as cost constraints. A possible approach to find a baseline will be to consider only the prevalent combinations in practice.

Third, another aggregation rule is necessary in order to evaluate the internal control system as a whole. The importance of a transaction cycle may vary in different companies. The factors that decide the relative weights will be governed by the characteristics of a company. For example, an accounts receivable cycle may be less important to evaluate the internal control system of a ship-manufacturing company because the cycle may have only a few, contracts, all of which are very reliable. In other words, it can be possible to track all transactions with relatively low ICPs. However, in general, account receivable cycle is one of the most important cycles. Identification of simplifying factors will ease the difficulty of importance determination.

Prior studies mainly focused on the consensus on quality for given internal control system components. As Wu and Hahn (1989) suggested, the internal control system should be considered as a set of interrelated controls from a holistic view. However, it will be difficult to estimate internal control system quality without evaluating each

www.manaraa.com

component and its relationships with others. However, little study about individual evaluation decision procedures has been performed.

It would be preferable to evaluate each step and make its baselines to compute the degree of an internal control system quality rather than to decide the final measure from ICPs. There are critical issues about quantification of internal control system quality, including:

1) How many relevant ICPs for each ICC auditors can identify,

2) How each ICP can be evaluated considering its compliance level,

3) How those identified ICPs are connected to one another to form an ICC in their evaluation processes, and what effect each ICC structure has,

4) Whether there is strong consensus on the evaluation for the same ICC structure among auditors,

5) How much each ICC contributes to the level of the internal control system quality,

6) Whether there is strong consensus on the internal control system quality for the same quality levels of the ICCs, and

www.manaraa.com

7) How the quantified value of an internal control system is mapped with an actual audit opinion.

First, the number of ICPs that auditors can identify from an internal control system may depend on their audit experience. This step is closely related to understanding of the transaction flows. SAS No. 1 emphasizes the importance of system understanding: "An understanding of the flow of transactions should provide the auditor with a general knowledge of the various classes of transactions and the method by which each significant class of transactions is authorized, executed, initially recorded, and subsequently processed, including the methods of data processing".

Following SAS guidance, relevant ICPs can be identified through surveys or observations. However, the problem here is the limited generalizability of the results across companies, industries, organization and regulatory structures, and cultures (Kinney, 2000). The results can be treated at most as a case study. To overcome this limitation, it will be necessary to provide various internal control systems in the experiments.

Second, the evaluation of ICPs based on their compliance is a very challenging task. The purpose of compliance testing is to provide reasonable assurance that accounting

www.manaraa.com

control procedures are being applied as prescribed. As Colley and Hicks (1983) mentioned, individuals within a company were likely to contribute to IC failures. However, in the absence of documentation, evidence is usually obtained by the auditor through original inquiries or reference to written instructions and through supplemental corroborative inquiries and observation of office personnel and routines (Statement on Auditing Procedure No. 54, 1972). As information technology develops, offices become paperless, hindering discovery of document evidence. Hence, compliance tests will be a subjective rather than objective task in which evaluation is affected by intentionality and monetary impact (Ferries and Tennant, 1984). Management fraud must be also considered while testing compliance level. If an internal control system allows managements to easily override, then the systems may have potential fraud risk even if management currently has no intention of fraud.

Third, relationships among ICPs may be difficult to determine because of the inherent complexity of internal control. This problem hinders two steps: identification of the characteristics of each ICP and construction and evaluation of relationships among them. For example, two ICPs can be independent, complementary, or inter-dependent. If they

www.manaraa.com

are independent, they can form sequential, parallel, or other forms of relationships, and each relationship structure may have different values. For example, if two ICP have the same values such as 1, the combination effect can be equal to 2 (addition), between 1 and 2 (union of sets that have intersection), greater than 2 (synergy), or less than 1 (conflict). If they are complementary, the resulting effect can be greater than the values of each ICP, but not exceeding the sum of them. For inter-dependent ICPs, they can be treated as one ICP rather than as two ICPs. If either of them is deficient, the quality of the other will be deficient as well (Vasarhelyi, 1980).

Fourth, a critical factor in quantifying internal control system quality is to draw consensus on the evaluation values for a given ICC. If there is insufficient consensus on them, the quantification baselines for internal control system quality will be difficult to form. As mentioned earlier, it is not feasible to consider all combinations of ICPs for an ICC. The best approach will be to survey what kinds of combination of ICPs are used for specific ICCs in practice. Narrowing the scope of appropriate ICP combinations simplifies further investigation. The first step of quantifying internal control system quality is to set a benchmark, not to find the best internal control system structure. After a

www.manaraa.com

benchmark is initialized, it can be compared to other internal control systems to measure relative quality. One of two criteria can be used as a benchmark: most frequently used ICC structure or marginally acceptable ICC structure. These two ICCs may or may not be identical. For example, companies may frequently use a specific type of ICC because it is less costly to perform and/or because it is more likely to result in a favorable opinion. On the other hand, companies may use one type of ICC more frequently because it can bring more benefit of avoiding possible adverse audit opinion on their internal control system. In this case, the marginally acceptable ICC will be different from the most frequently used ICC. However, by definition, a baseline is likely to be a neutral point, making marginally acceptable ICC well-suited to the task.

Fifth and sixth, after evaluation of ICCs is completed, the next step is to aggregate those quality values into overall measure for an internal control system. This step will be similar to that found in evaluation of an ICC, except that these ICCs will be relatively independent from one another. A possible aggregation rule can be a weighted average. As a possible example, if all but one ICC exhibit high quality, but the one ICC is vital within the internal control system, then the internal control system quality can be regarded as

www.manaraa.com

materially weak. Conversely, if all but one ICC exhibit low quality but the one ICC has extraordinary high level of quality, the internal control system can be treated as inefficient rather than significantly deficiency. Because of a relatively straightforward aggregation rule, the consensus on the internal control system quality for given ICCs can be even more important than the previous step. If there is little consensus, it may imply that auditors do not consider much of ICCs during evaluating overall internal control system quality. In addition, the aggregation rule can be greatly affected by other factors such as industry, necessitating specific rule applications.

The last quality quantification issue relates to mapping the quantified value to an actual audit opinion. According to the PCAOB, there are four levels of internal control system quality: effective, ineffective, significant deficiency, and material weakness. The similar procedures to those on the previous step can be applied here. One distinct difference is that the mapping criteria will be greatly affected by PCAOB and/or audit firm policies. In other words, if an audit firm is highly risk-averse, relatively high criteria will be applied to set critical points for mapping. For example, if internal control system quality value is 85 (based on a scale of 0-100), an auditor may issue an either effective or

www.manaraa.com

ineffective opinion based on the degree of risk-aversion.

The quantification of internal control system quality is a relatively untouched area because of the complexity of internal control system and limited generalizability of research results. However, when objective baselines are initialized, research on adequacy or effectiveness will lead to better benchmarks. With effective and objective benchmarks, the internal control system quality of one company can be compared with another. The quantification of internal control system quality will provide many benefits. It can help auditors to issue a proper opinion on a company's internal control system. It can help investors evaluate information in the financial statements by considering the environments in which such statements are generated. It can also help regulation setters to indicate desirable future directions for internal control system.

There are many barriers to adequate quantification of internal control system quality (Kinney, 2000). First, there can be many possible ways to achieve a given internal control system quality. Without narrowing scope to a manageable level, it will be difficult to set a benchmark that will be used for quantifying internal control systems. Second, compliance level is difficult to measure practically, objectively, and stably, especially in today's high

www.manaraa.com

IT environments where tangible audit trails that can be used as compliance evidences are no longer present. However, without appropriate compliance testing, it will be difficult to estimate the objective quality of ICPs that are parts of an ICC; consequently, internal control system quality values will be less meaningful.

Since the Sarbanes-Oxley Act, both practitioners and scholars have shown great interest in internal control systems. The advent of the SOX raised a variety of issues about internal control system. However, there are few criteria about measuring the internal control system quality. To meet the requirements of SOX, there are many changes in companies' internal control systems (Rittenberg and Miller, 2005). Many online websites provide a benchmark service that enables companies to benchmark its SOX compliance efforts against those of its peers and evaluate its internal controls in a comprehensive and standards-based way (Williams, 2005). Those consequences are very natural when practitioners need to comply with regulations that do not provide detailed and objective guidance. In addition, in the absence of normative criteria in internal control system evaluation, it will be indispensable to draw consensus on internal control system quality to set evaluation criteria. Only after those criteria are determined will

www.manaraa.com

internal control system quality be objectively measured.

c. Fraud Prevention & Detection

Although fraud detection was not intended to be an auditor's responsibility, the public believes that it should be the part of the auditor's job. Prior to SOX, there were no statements that obliged auditors to detect fraud. In the post-SOX era, however, auditors seem to have some degree of fraud detection responsibility through evaluating whether the company's controls sufficiently address the risk of material misstatement due to fraud (SAS 99).

Fraud can be either external or internal. External fraud is committed by an external party (e.g. customers, criminals, and intruders) while internal fraud is committed by an employees. The framework of Jans et al. (2009) suggested three dimensions of internal fraud: 1) statement or transaction fraud, 2) management or non-management fraud, and 3) fraud for or against a company. In this study, "internal fraud" refers to internal transaction fraud against a company committed by either management or non-management.

Three conditions (i.e. fraud triangle) must be satisfied for fraud to take place. First of

www.manaraa.com

all, a fraud perpetrator (commonly called a "fraudster") must have incentives or pressures that are related to financial difficulty. Second, there must be an opportunity to commit fraud. Even if a fraudster is willing to commit fraud, he/she needs an opportunity to take action. Lastly, if an employee is willing and able to commit fraud, his/her action will not be executed without rationalization to justify such behavior (SAS No. 99). Anti-fraud activities are generally categorized into two groups: prevention and detection. The former can be achieved by removing one or more of the three conditions for fraud commitment. For example, if an enterprise's internal control system is sufficiently effective, it will be difficult for a fraudster to find an opportunity to commit fraud. However, most prevention methods are difficult to implement and evaluate because of their qualitative characteristics. For example, it is not an easy task to anticipate fraud incentives or the pressures on an employee. Compounding the situation, the effects of ethical education on rationalization reduction are difficult, if not impossible, to measure. As a result, a well-designed internal control system seems to be the only practical way for fraud prevention activities.

The value of fraud prevention can be estimated by its opportunity costs. The

www.manaraa.com

economic impact of fraud in the United States was considerably large and its scale accelerated during the past decade (Schnatterly, 2003). The estimated cost of fraud was $660 billion (6% loss of revenue) in 2002 (Association of Certified Fraud Examiners, 2002) and a more recent report by the Association of Certified Fraud Examiners (ACFE, 2007) showed that the cost of occupational fraud and abuse (hereinafter referred to as 'internal fraud') was approximately $994 billion in the US, which represents a loss in revenue of about 8% to businesses. In 2009, it also noted that an increase in fraud was caused by the intense financial pressures of the current economic crisis and that the greatest fraud threat was posed by employees (48.3% increases in employee embezzlement from the previous year). This increase implies ineffective internal controls and a lack of fraud detection/prevention systems.

Based on this finding, it might be reasonably argued that fraud prevention brings a potential competitive advantage and enhanced financial performance. A company's internal control system is crucial for detecting and preventing fraud. A properly designed internal control system facilitates reliable financial information by preventing, detecting, and correcting material errors and irregularities on a timely basis. Employee fraud has

www.manaraa.com

received little attention in the literature while fraud by outsiders has been well-researched. This may be due to lack of data or fear of losing competitive advantage (Bolton and Hand 2002; Phua et al. 2005). However, recent financial scandals have clearly shown that internal fraud affects a company's revenue more adversely than external fraud.

2. Analytical Procedures

i. Background

Analytical procedures (APs) are used during a typical audit engagement to identify weak or suspicious areas that need more investigation. This purpose seems is similar to that of IC screening models that seek suspicious transactions. The major difference between them is the aggregation level of the data that is used during each procedure. While APs typically use highly aggregated (i.e., yearly or quarterly) data, IC screening utilizes highly disaggregated, transaction level data. Because of their similarities, this section explores the prior research about APs.

www.manaraa.com

a. Definition & History

APs have been one of the most studied areas in auditing. APs are defined as substantive auditing procedures that examine the accuracy/reasonableness of reported account balances or the unexpected relationships in financial data in light of the firm's history and contemporary economic conditions without considering the details of individual transactions which make up the account balance in financial data helping both internal and external auditors (Lev, 1980; Wild, 1987; Gaunti and Glezen, 1997; Knechel, 1988).

APs are discussed in Statement on Auditing Standards (SAS) No. 1, which states: "The evidential matter required by the third standard (of field work) is obtained through two general classes of auditing procedures: (a) tests of details of transactions and balances, and (b) analytical review procedures applied to financial information" (AICPA, 1973). Although these two basic audit procedures are both performed to opine on the fairness of an auditee's reported accounting balances, APs differ from tests of details since their focus is on balances and transactions at an aggregated level rather than on the components of the balances at the individual/transaction level (Kinney, 1978).

www.manaraa.com

Later, SAS No. 23 defined APs as "substantive tests of financial information made by a study and comparison of relationships among data" (AICPA, 1978) which was superseded by the SAS No. 56 (AICPA, 1988) that revised the definition as "evaluations of financial information made by a study of relationships among both financial and non-financial data" (Holder, 1983; Gaunti and Glezen, 1997).

Statement on Internal Auditing Standards No. 8 (IIA, 1992) defined APs as "analytical auditing procedures … performed by studying and comparing relationships among both financial and non-financial information" (Gaunti and Glezen, 1997)

In practice, APs have become an indispensable part of auditing after being recommended by the Auditing Standards Board (ASB, 1978) and mandated for planning and overall review purposes by SAS No. 56 (AICPA, 1988).

b. Underlying Rationale

In typical APs, the auditor compares the client's reported balance (or ratio) with the auditor's assessments of the likely true (audited) balance. A basic premise underlying the application of the APs is that the recorded amounts and their variations and plausible

www.manaraa.com

relationships among data may reasonably be expected to exist and continue in the absence of known conditions to the contrary (Lev, 1980; SAS No. 56, 1988; Kinney and Felix, 1980; Holder, 1983; IIA, 1992; Gaunti and Glezen, 1997).

This rationale may imply that APs will perform well when the relationships among the economic events recorded in financial statements are relatively stable (Gaunti and Glezen, 1997). Chen and Leitch (1998) showed that all AP models performed better when data had a greater degree of stability in their business and economic activities. In other words, the more stable data is, the more accurate prediction a model can make.

c. Usefulness

Aside from legal requirements, the increasing use of the APs stems from economic reasons and legal disputes. First, APs can be a relatively inexpensive means for reducing detailed substantive testing requirements in auditing. Kinney (1978) argued that APs might be a relatively economical way to enhance auditor confidence in the validity of reported numbers by assessing the reasonableness of balances based on all known information (Kinney, 1978). APs have been regarded as useful tools in identifying areas

www.manaraa.com

where the risks of errors and irregularities are high, so that auditors can allocate scarce auditing resources more effectively.

SAS No. 53 defines "errors" as unintentional misstatements and "irregularities" as intentional misstatements. In this study, "anomalies" will be used to indicate both errors and irregularities since it is difficult to discriminate them without further details.

According to Lys and Watts (1994), lawsuits against auditors were often driven by financial misstatements regarding assets, revenues and liabilities. Calderon and Green (1994) also argued that although external auditors were responsible for fraud detection to some extent, responsibility rests with management. In other words, internal auditors and controllers had to play a central role in fraud detection.

Auditors have indicated that many financial statement errors were initially detected via analytical reviews (Hirst et al., 1996; Gaunti and Glezen, 1997; Biggs and Wild, 1984). More specifically, Hylas and Ashton (1982) reported that 27.1% of errors were detected by APs. Their study also indicated that the performance of APs was stable even after company size was considered (28.8% for >$50M, 26% for <$10M, and 27.2% for companies in between). In the other research, auditors mentioned that APs initially

www.manaraa.com

detected 41.5% and 45.0% (mean and median, respectively) of previously encountered financial statement errors (Biggs and Wild, 1984). Kreutzfeldt and Wallace (1986) also reported that about 42% (22% if discussions with client are excluded) of errors were detected by APs and Wright and Ashton (1989) reported that 15.5% of errors were detected by APs.

Proper use of APs may be of great help in detecting unusual numbers and/or their relationships. The downward pressure on audit cost, more demanding responsibility for detecting misstatements or fraud in the financial information under audit, and increased use of microcomputers have led auditors to rely more on APs that are both efficient and effective (Ameen and Strawser , 1994; Wheeler and Pany, 1990).

Also Coglitore and Berryman (1988) claimed that constructive use of APs is effective in detecting unusual relationships in the data and/or significant changes in such relationships. In response to increased concerns about downward pressure on audit fees and demands that auditors took more responsibility for detecting misstatements in their clients' financial information, auditors increasingly seek the APs that were presumably both efficient and effective in detecting materially misleading financial information

www.manaraa.com

(Wheeler and Pany, 1990).

ii. Prior research

a. Usage in practice (survey)

APs are powerful tools to detect errors and irregularities, and their prevalent usage can therefore be reasonably expected (Coglitor and Berryman, 1988). Although auditors frequently relied on APs as the primary substantive test in an audit area, little was known about how practitioners employed APs during audit procedures (Hylas and Ashton, 1982; Biggs and Wild, 1984; Hirst and Koonce, 1996; Glover et al., 2005).

Auditors select AP aggregation levels such that the results would produce the most meaningful interpretation (Kinney, 1978).

Interestingly, only a few APs were extensively used regardless of the type of engagement (Daroca and Holder, 1985). Despite the strengths of sophisticated techniques such as time series and regression expectation models, they were rarely used in practice. Instead, auditors preferred the simple comparisons such as martingale and sub-martingale models defined by Kinney (1978) (Daroca and Holder, 1985; Ameen and Strawser, 1994;

www.manaraa.com

Hirst and Koonce, 1996).

b. Expectation models

According to SAS Nos. 23 and 56, APs range from simple or informal scanning and comparisons to the use of complex mathematical and statistical models involving many relationships and elements of data. Because of their objectivity and ability to provide decision rules, the latter receive more attention (Kinney, 1978; Lev, 1980; Holder, 1983; Wilson and Colbert, 1989; Gaunti and Glezen, 1997).

In applying APs, auditors usually rely on expectation models to make predictions about the values of important business metrics. These predicted values are then compared with actual values. If the differences between the two values are beyond a predetermined threshold, the actual values are regarded to have potential anomalies that require further investigation. Accordingly, the expectation model in an AP is a critical anomaly detection tool.

Statement on Standards for Accounting and Review Services (SSARS) No. 1 (AICPA, 1978) and SAS No. 23 provided five types of APs ordinarily applied in review

www.manaraa.com

engagements: 1) comparison of financial numbers with those in comparable prior periods, 2) comparison of financial numbers with auditor expectations, 3) study of the relationships of the components of financial statements that would be expected to have similar patterns in the prior periods, 4) comparison of financial information with industry benchmarks, and 5) study of the relationships of financial information with relevant non-financial information (Daroca and Holder, 1985). When applying these APs, auditors may use both financial information and nonfinancial information such as general economic conditions, technological changes in the client's industry, and new products from competitors (Cohen et al., 2000).

Kinney (1978) compared the performance of various APs that were based on ordinary least squares (OLS) regression, three sets of integrated-autoregressive-moving-average (ARIMA), martingale, and submartingale models for monthly accounting data. His study showed that the ARIMA model outperformed the others and concluded that AP models with the largest information and the greatest computational effort would produce the smallest prediction errors and bias. Despite the poorer performance, less sophisticated models might be also used in practice because of their economic benefit.

www.manaraa.com

Knechel (1988) examined several AP models and concluded that regression expectation models were very efficient in misstatement detection.

Wilson and Colbert (1989) indicated that more rigorous APs provided more accurate expectation results that would generate more efficient auditor decisions. However, they would require more information and increase the complexity of models.

Wilson (1991) suggested careful use of regression models by examining the effects of various degrees of data dispersion. His study showed that data with larger dispersions tended to generate more incorrect rejections than those with smaller dispersions. As a result, if auditors apply the regression model to highly dispersed data, they should interpret the results more carefully.

Dugan et al. (1985) recommended the Census X-11 model as a user-friendly substitute to the ARIMA model. They argued that ARIMA was not frequently used by auditors because of its extensive data requirements, complexity, interpretation difficulty, and cost of operation, although it was the most effective model in theory. Compared with the ARIMA model's use of extensive time-series data, the X-11 model decomposes client's data into three components: trend, seasonality, and irregularity. After removing

www.manaraa.com

the first two components, irregularity is investigated. The X-11 model performed as well as the ARIMA model.

Structural models were also considered for developing APs. Structural relationships in accounting data reflect the key economic events of the organization. Consequently, structural models should bring better results. In the study of Chen and Leitch (1998), they used the entity-relationship (ER) format to capture critical economic events. However, they are generally outperformed by multivariate stepwise models although their performance was better than the other models did (Wild, 1987; Dzeng, 1994; Chen and Leitch, 1998).

Moving away from the trend of more complex AP models, Nigrini and Mittermaier (1997) introduced Benford's law as an AP. The technique is relatively simple. By comparing the actual frequencies of the first digits of an account data with the expected frequencies based on the rule, a quality figure can be derived.

c. Use of Disaggregated vs. Aggregated data

Claims that APs increase reliability have their detractors. As defined by the SAS, APs

www.manaraa.com

usually use data at an aggregated level. APs generally analyze aggregated data by taking those data as presented on the financial statements. Consequently, APs are generally less expensive to apply than tests of details, but they are also less reliable (Hitzig, 2004).

In response to this concern, the SAS No. 56 (AICPA, 1988) recommended the use of disaggregated data: "Generally, the risk that material misstatement could be obscured by offsetting factors increases as a client's operations become more complex and more diversified. Disaggregation helps reduce this risk. Expectations developed at a detailed level generally have a greater change of detecting misstatement of a given amount than do broad comparisons. Monthly amounts will generally be more effective than annual amounts and comparisons by location or line of business usually will be more effective than company-wide comparisons." In other words, APs based on disaggregated data might be more effective and efficient than those based on aggregated data (Glover et al., 2005).

Kinney (1978) argued that models with high information requirements and computational effort would be superior in predictive power. The reliability of inferences about the validity of account balances is supposed to be low when the level of

www.manaraa.com

aggregation of data is high. Furthermore, if errors that are material in monthly data might be immaterial to the annual balance, it would be difficult to detect and correct them (Kinney, 1978; Hirst and Koonce, 1996). APs must utilize data as disaggregated as auditing cost permits.

Knechel (1988) showed that the AP models using monthly data were more effective. He argued that the superior performance of APs with monthly data was due to the error distribution and its interaction with the APs. More specifically, if more disaggregated data is used in APs, it is likely that only relevant (or potentially erroneous or irregular) data is examined.

If greater data requirements give more accurate and meaningful predictions, the model with the most disaggregate transaction data will give the most accurate results (Kinney, 1978; Hirst and Koonce, 1996). This is a fundamental rationale of developing expectation models using transactional data.

Gaunti and Glezen (1997) also argued that if APs were more frequent such as monthly performed, the number of data would be larger and, consequently, account balances would have larger normal variations that enhanced the power of APs. On the

www.manaraa.com

contrary, accounts with some degree of aggregation tend to have less discriminating power because of a smoothing effect.

Chen and Leitch (1998) also supported the use of less aggregated data by simulation study. They argued that use of disaggregated data should bring better performance for three reasons: 1) increased statistical power resulting from a larger sample size, 2) reduced influence of organizational changes, and 3) more efficient measurement of economic characteristics.

However, there was some dissenting evidence regarding the use of disaggregated data. Wheeler and Pany (1990) reported that APs did not perform very well when quarterly data was used. However, when the quarterly data was annualized and individual quarterly data was seeded with annual material errors, the APs worked better. Although not explicitly discussed, the results of Dugan et al. (1994) also showed that all AP models worked better with annual data than quarterly data. Furthermore, Allen et al. (1999) did not find evidence that use of disaggregated data in APs brought better performance.

With the increasing use of EDP-based systems, data extraction from a company's accounting information system is no longer of prohibitive cost. It is neither infeasible nor

www.manaraa.com

impractical to extract transaction data from the database of a company. The availability of transactional data in digital form allows for APs that use transactional level data at reasonably low cost. This motivates the use of individual transactions for developing an anomaly detection model and testing whether an individual transaction is anomalous or not.

d. Decision Rules: Type I vs. II errors

The performance of APs may be measured by the presence of type I and II errors. Type I error indicates the likelihood of rejecting correct data (efficiency) while the type II error is about the likelihood of accepting incorrect data (effectiveness). APs with more type I error would result in more exceptions that require further investigation (less efficient). In contrast, APs with more type II error would increase detection risk (less effective) (Knechel, 1988).

To determine the level of effectiveness and efficiency of AP models, the level of material errors must be established. The materiality error levels can be determined either non-statistically (e.g. 10% of annual balances for the monthly data based on the auditor's

www.manaraa.com

professional judgment) or statistically (e.g., upper prediction limit). Coakely (1982) categorized investigation rules into three types: 1) judgment-based rules—arbitrary thresholds based on professional experience (e.g., 10% of previous balance), 2) materiality-based rules—level of tolerance, and 3) statistically-based rules—multiple of the standard error (e.g., 95% UPL) (Knechel, 1988).

Some rules are frequently used in prior research to define materiality. For example, the formula, $1.6 \times \max(\text{total assets}, \text{revenue})^{2/3}$ (also referred as "gauge"), has been used in some studies (Elliot, 1983; Wilson and Colbert, 1989). In addition, Warren and Elliott (1986) used 0.5% of annual sales for the sales-driven accounts such as A/R, inventory, A/P, and cost of goods sold, while 1% times the account's annual balance was used for the others.

In general, there exists a tradeoff between efficiency and effectiveness especially when the materiality of error is relatively small (Wilson and Colbert, 1989).

Chen and Leitch (1998) showed that the structural model had relatively large type I error but small type II error. However, if both perspectives were combined, the performance of both structural and stepwise models was consistently the best.

www.manaraa.com

Furthermore, more structural models have lower type I error and lower type II error for the positive approach (E=0, no error) but lower type I error and higher type II error for the negative approach (E=M, material error).

3. Continuous Monitoring/Auditing

The Auditing Concepts Committee (1972) defined auditing as "a systematic process of objectively obtaining and evaluating evidence regarding assertions about economic actions and events to ascertain the degree of correspondence between those assertions and established criteria and communicating the results to interested users." Auditing focused on the verification of assertions made by management regarding proposed financial reports (Alles et al., 2004). Although continuous auditing follows this concept, CA deals with more detailed and specific data rather than aggregate account balances on the financial reports. Despite this difference, CA is not isolated from traditional auditing. Instead, CA can be seen as a general form of auditing that includes traditional auditing.

Vasarhelyi and Halper (1991) first introduced the concept of CA when they developed a monitoring tool in an online IT environment. CA intends to provide more timely

www.manaraa.com

assurance by continuously monitoring all of a company's transactional data. This suggestion did not draw much attention from either academia or practice for a decade because of doubts regarding its feasibility and effectiveness. CA has somewhat recently become actively researched in academia and practice. After a series of recent financial reporting scandals (e.g. WorldCom, Enron) and related auditor failures (e.g. Arthur Andersen), researchers, practitioners, and regulators have looked to prevent future financial disasters. CA is believed to be the most promising means to that end, so this area has been researched heavily.

Although many works have been published, the majority of papers on CA have adopted technical perspectives (Vasarhelyi and Halper, 1991; Kogan et al., 1999; Woodroof and Searcy, 2001; Rezaee et al., 2002; Murthy, 2004; Murthy and Groomer, 2004). A few papers discuss other aspects of CA such as its concepts and research directions (Alles et al., 2002 and 2004; Elliott, 2002). Only a handful of papers (Alles et al., 2004 and 2006) have included empirical studies, and this deficiency is likely due to a lack of available data. CA research requires much disaggregated data that should be kept inside a company to maintain a competent position in a market. It is not surprising that

www.manaraa.com

companies are reluctant to provide transaction data. However, empirical studies are necessary to verify and validate CA.

4. Methodology

i. Overview

People generally prefer to have more information prior to making a decision. However, this claim may not hold for a business that experiences information overload. People have difficulty utilizing overwhelming information when timely strategic decisions are necessary. One solution could be exception reports that highlight significant problems (Campbell et al., 2006). This approach will be utilized in this study. Instead of models that produce too many false positives, the models that flag fewer alarms will be used. However, this approach has a clear drawback. Because of the tradeoff between efficiency (alpha risk, type I error, or false positive) and effectiveness (beta risk, type II error, false negative, or detection risk), this will probably reduce the model's detection power. Models with more efficiency will be preferred if they have similar levels of effectiveness.

www.manaraa.com

To this end, we may need to see what approach will produce more or fewer alarms. Let us assume that there are two ICPs, both producing alarms. We can use the results in two different ways to determine which flags should be investigated.

First, a transaction will be flagged for more examination only when it is flagged by both ICPs (AND condition). The other way is that a transaction is alarmed when the transaction violates either or both ICPs (OR condition) (Srinidhi and Vasarhelyi, 1986). Consequently, the number of alarms by the former approach will be less than or equal to that of the latter. The decision rule can be determined based on the importance of the IC object. If an IC object is of great concern, the 'OR' condition may be more preferable. However, it not, applying the 'AND' condition will be more cost-effective.

In order to develop monitoring/screening models, the first questions will be what to monitor and how to monitor it. Consequently, we need to understand the business processes and identify potential monitoring candidates. Another relevant question will be how to measure how reliable the numbers on the database are. The latter question cannot be answered until sufficient evidence about system reliability is collected and judged. However, if system reliability is considered in this study, the scope will be too broad to be

www.manaraa.com

covered, and we will therefore assume the company's data processing system to function as it should. That is, all numbers on the database are correctly entered and processed even if they are materially erroneous or fraudulent.

Another problem with developing monitoring models is a lack of research. Although there were some internal control system evaluation models, most of them were general rather than specific such as whether segregation of duties for accounts payable (A/Ps) was properly performed. In order to develop detailed monitoring rules, detailed scenarios that could occur in practice must be researched.

For example, an A/P table has two columns that indicate who processes transactions and who pays the amount, respectively. In this case, we may say that the segregation of duties is performed well by examining whether the two columns are identical. Although this concept seems trivial, its evaluation is not. Let us extend the A/Ps example above.

Assume each A/P has one processor and two authorizers, and that the IC objective is proper A/P authorization. The system must first check whether the three employees were different. If not, the presence of two authorizers will be meaningless. Next, the system might examine whether the authorizers processed a transaction only after initial entry. If

www.manaraa.com

pre-entry authorization is possible, the authorization function is not well designed. A rank check may also be appropriate. If the ranks of the authorizers can be lower than that of the processor, authorization function may not work properly because of possible overrides.

Even for this simple control object, we can easily identify several ICPs for screening purposes. However, there is little research showing how many ICPs will be necessary to guarantee reasonable assurance for the particular IC object. If there are surveys that show lists of ICPs and their relative evaluations, this process may not be too challenging. However, because of a lack of research on the subject, the models in this study may only be confirmed and verified by corresponding practitioners.

Verification of models in this study will focus on comparisons between the results from the model and actual examination by internal auditors. If auditors test details about flagged transactions, their results will be the only source of confirmation and verification. However, this method cannot detect potential false negatives that will not be available to internal auditors to examine. Another model verification option is simulation. After assuming that all data is error- and irregularity-free, errors will be seeded into the original

www.manaraa.com

data set and the models will be tested. The evaluation criteria will be detection of seeded errors and how many false alarms they generate.

Practical implementation is another key issue. As we will discuss in the analytical procedures section, highly sophisticated statistical methods are rarely used in practice even if their estimates are far more accurate than the others. Since all the ICPs will be eventually implemented in the company's data processing system, the possibility of implementation cannot be ignored. If we cannot implement all ICPs for a particular internal control object, which ICPs should be included and what ICPs could be excluded without losing significant power?

Prior research found that auditors' internal control assessment was inversely related to the monetary value of errors (Ferris and Tennant, 1984). This may indicate that monetary nature should be considered for developing ICPs. However, this may support both negative and positive approaches since the research did not provide the tolerance level to which auditors were significantly affected by the monetary nature of errors.

www.manaraa.com

ii. Rule-based approach: Expert system

SAS No. 3 divided electronic data processing controls into two categories: general controls and application controls. General controls are procedures that are applied to all applications in the system, whereas application controls are application-specific procedures. The former are prerequisite to the latter. That is, the application controls may have little effects if general controls are not effectively designed and implemented (Jancura and Lilly, 1977). Assuming that general controls are in place, this study focuses on application controls of, transitory accounts and wire transfers that are recognized by practitioners as the most important, riskiest areas.

This study will utilize a rule-based approach to monitor these areas, mostly because practitioners tend to use less complex and flexible decision rules even if sophisticated statistical methods are available (Elmer and Borowski, 1988). There can be several reasons to explain this tendency. One is understandability; simple decision rules such as if-then types are easily understood even if practitioners do not have much knowledge about underlying logic. However, statistical models usually require deep knowledge to understand decision processes and to interpret the results produced by the models.

www.manaraa.com

Flexibility and deploy ability may also affect adoption (Roth, 1985). Since rules in a rule-based decision model can be easily added or deleted, the model is easily flexible to adapt any changes. Furthermore, because of their simple logic, they are relatively easy to implement in practice without significant impact on DP systems.

Rule-based systems typically consist of a series of if-then decision rules that are straight-forward to human reasoning (Martin and Eckerle, 1991). Although this study will employ rule-based systems, our additional inclusion of statistical models makes this an atypical application. In this study, an anomaly detection model consists of a collection of anomaly indicators that are considered as ICPs. In this sense, the model is rule-based.

As with most rule-based models, the models in this study will require fine-tuning that can be achieved as our domain knowledge base grows.

iii. Unsupervised method

There are two main methods used in the literature to detect fraud: supervised and unsupervised. The most frequently used research methodology is classification (or supervised) methods. Supervised methods utilize prior information (also called labeled

www.manaraa.com

information) that contains both legitimate and fraudulent transactions, while unsupervised methods do not require any labeled data. Under the supervised method, a database of known fraudulent or legitimate cases is used to construct fraud detection models (Bolton and Hand 2002). The models are trained by prior labeled data, and then fraudulent and legitimate transactions are discriminated in accordance with those models. These methods assume that the pattern of fraud in the future will be the same as that in the past. Neural network models which use the supervised method appear frequently in recent research (Bolton and Hand 2002; Kou et al. 2004; Phua et al. 2005).

Although often used in research, supervised methods pose several limitations resulting from their heavy dependence on reliable prior knowledge about both fraudulent and legitimate transactions. This may be impractical since prior information might be incorrect. Most companies do not have sufficient resources to examine every transaction, and consequently, some ostensibly legitimate transactions may be fraudulent, and models based on this information may be misleading (Bolton and Hand 2002). Another limitation of the supervised method is that the results are often not easily understood. This may be a substantial obstacle to implementation since few enterprises can afford the requisite expertise (Sherman 2002). As a result, few enterprises would be interested in

www.manaraa.com

implementing the supervised method in practice. This is similar to analytical review procedures used by auditors where many sophisticated methods have been developed but simple methods dominate in practice. Supervised models are not easily adjustable.

A major concern in fraud prevention/detection research is that models may work only for the data used to create them. The generalizability of a fraud profile is highly dependent on the context of the original model development and on the target environment. For example, if new data comes into the dataset, those models may not work due to either over-fitting to the training dataset or the presence of unknown fraud types. In addition, the robustness of models is a major concern during extension, re-utilization, and adaptation. Considering that fraud perpetrators adapt to find loopholes in an enterprise's current fraud prevention/detection system, this can be a critical weakness. In order to adapt to unknown types of attacks, it is important that the systems be dynamically extendable and adjustable. Supervised methods also suffer from uneven distributions of legitimate and fraudulent observations. Generally, the number of fraudulent observations is greatly outnumbered by that of legitimate ones. About 0.08% of annual observations are fraudulent (Hassibi 2000). In other words, even if a model classifies all fraudulent transactions as legitimate regardless of their true identities, the

www.manaraa.com

error rate (correctly classified transactions/total transactions) of the model is extremely small, which can be misleading.

Unsupervised methods have received far less attention in literature than supervised methods. Unsupervised methods focus on detection of changes in behavior or unusual transactions (i.e. outliers) by using data-mining methods. Anomaly/outlier detection is the recognition of patterns in data that do not conform to expected behavior (Chandola et al. 2009). The major advantage of unsupervised methods is that they do not require labeled information, which is generally unavailable due to censorship (Bolton and Hand 2002; Kou et al. 2004; Phua et al. 2005). The results are rarely disclosed in public either to maintain an enterprise's competitive advantage or because of public benefits (Little et al. 2002).

Unsupervised methods usually employ suspicion scoring systems that estimate the degree of departure from the norm by utilizing if-then type outlier rules. Rule-based systems are increasingly used to represent experiential knowledge. Outlier definition criteria may change for many reasons such as cost and efficiency. Decision making by if-then rules is similar to human cognitive decision processes, which enables internal auditors to understand and adjust the models if necessary. However, verification of newly

www.manaraa.com

devised models is often difficult, if not impossible, due to lack of testable data. To tackle this weakness, methods such as peer group analysis, where groups with similar profiles are compared, and break point analysis, where recent transactions are compared with past patterns, can be used (Bolton and Hand 2001).

The results of unsupervised methods are not direct evidence that flagged transactions are fraudulent. Instead, the aim of unsupervised methods is to inform that flagged transactions are more anomalous, tending toward either error or fraud, based on the experience, analysis, and preconceptions of the analysts. In other words, a flagged transaction can be legitimate, error, or fraudulent. This outcome is clearly different from that of supervised methods, where outcomes are either legitimate or fraudulent. As Jans et al. (2009) described, an outlier can occur via mistakes (i.e. unintentional errors). Unsupervised methods consider broader causes than supervised methods. Furthermore, a transaction will be worthy of further investigation if it is flagged by multiple criteria, since normal transactions are unlikely to be flagged by many indicators. Analogous to other rule-based systems, the actual examination of selected transactions allows for re-parameterization and improvement of the method. However, the verification of resulting flagged transactions requires internal auditors' direct examination.

www.manaraa.com

Despite the drawbacks of unsupervised methods, they may be indispensable at the initial implementation stage where prior labeled information is rarely available. In addition, considering that it is ultimately internal auditors who will use and maintain fraud prevention/detection models and that only a few enterprises can afford the expertise necessary for them, a rule-based approach may be desirable for internal auditors (Sherman 2002). The advantages and disadvantages of supervised and unsupervised methods are summarized in the Table 1.

Table 1. Advantages and Disadvantages of Supervised and Unsupervised Methods

	Supervised Methods	Unsupervised Methods
Advantages	1. Accurate for known fraud types	1. Easy to apply and update 2. Possible to find unknown fraud types 3. Unnecessary to use labeled data 4. As accurate as complex methods in
Disadvantages	1. Highly unbalanced class sizes 2. False negatives 3. May work only for known fraud types 4. Highly dependent on historic data that may not be accurate 5. Less understandable	1. Less accurate than complex methods in the short term 2. Necessary to be verified by auditors
Result	Fraudulent or not	Possibly fradulent or not

www.manaraa.com

III. Development of Anomaly detection models

1. Case I: Development of An Anomaly Detection Model for A Bank's Transitory Account System

i. Introduction

The ultimate goal of anomaly detection models is to filter out true anomalies from a population. However, while focusing on the power of the models, both researchers and practitioners tend to neglect the feasibility of their implementation in the real world. If monitoring is sporadic (e.g. annual or semiannual), practical implementation may not be of great concern since all the data can be downloaded or transferred to a designated place and examined by the models. However, if continuous monitoring is necessary, practicability becomes of great importance to practitioners.

Among many functions of a bank that may need continuous monitoring, this study investigates a process of transferring funds from one customer to another. While a sender can be either a bank account holder or a non-account holder, its recipient is generally an

www.manaraa.com

account holder. However, according to an internal auditor of the bank, it is not uncommon in the process of a fund transfer that destination of the wire is not identified immediately when a bank receives funds from a sending customer. A common cause can be a wrong recipient account number. When a bank cannot identify the recipient immediately, the fund is sent to a transitory account created for this purpose, holding funds until recipients are identified. Although this waiting period can last several months, most wire transfers do not stay in a transitory account for a long time. A bank may have multiple transitory accounts depending on its needs and purposes. The bank in this study owns about ten thousands of such transitory accounts.

Detecting anomalies among millions of transactions from about ten thousand bank accounts is a challenging task in terms of data processing. Regarding the cost of monitoring activities, an online (or real-time) monitoring system is highly challenging to develop since it will consume vast data processing resources, potentially interrupting regular business activity. This is one of the major factors that make a company hesitant to integrate an internal control system into its existing data processing systems. Furthermore, the complexity of internal control screening models that are introduced in literature can

www.manaraa.com

serve as another barrier to a screening model implementation in practice. For example, most fraud detection models in literature require mathematical and/or statistical expertise that most internal auditors do not possess and understand. One practical solution for internal auditors can be to develop a screening model that consists of a series of generic rules that do not contain complicated mathematical or statistical algorithms. Taken together, it will be practically necessary to develop a screening model that can be applied to transitory bank accounts without significantly affecting a company's data processing systems.

This strategy, however, has major weakness. As defined, enhanced practicability may imply that a detection model should be sufficiently light and generic to implement within a current data processing system, which may reduce the power of the model. In other words, this approach may bring a trade-off between the power of a screening model and its practicability. The success of this approach will depend on balancing power with practicability.

www.manaraa.com

ii. Objectives

A transitory account is a temporary buffer for a fund in transit before a final destination is generally identified by human intervention and updated by a manual fund transfer process. Transitory accounts are vulnerable to anomalies including internal fraud when their activities and the employees in charge are not rigorously monitored and verified. Monitoring and verification can be accomplished manually or with the use of technology.

The purpose of this study is to develop and test an internal control monitoring model to detect anomalies out of millions of transactions that use approximately ten thousand transitory bank accounts without consuming significant data processing resources. The transactions flagged by the model are cross-checked by internal auditors to estimate the power of models, and the results are used to make further improvements.

The remaining sections proceed as follows. In the Methodology section, I discuss the dataset in this study and the screening rules of a model that are used to detect possible anomalies. And the last section describes and discusses the test result, followed by conclusion and suggestions for the future research.

www.manaraa.com

iii. Methodology

a. Phase I (August 2008)

Data

The data in this study includes transitory account transactions from a large Brazilian bank. The dataset includes sixteen transitory accounts out of the ten thousand accounts. The dates of each account have different ranges. The narrowest range is about a year (from late May 2007 to early August 2008) while the longest is about three years (from early October 2005 to early August 2008). The table 2 details the ranges by account.

www.manaraa.com

Table 2. Ranges of Transaction Dates

account	Distinct days	Oldest	Latest	range
a5738	518	10/04/2005	08/11/2008	1043
a45136	244	02/02/2006	08/11/2008	922
a60836	202	04/02/2007	08/11/2008	498
a32360	233	01/26/2006	08/11/2008	929
a61042	227	04/30/2007	08/11/2008	470
a21830	232	04/03/2006	08/11/2008	862
a21776	226	05/25/2006	08/11/2008	810
a68128	226	04/30/2007	08/11/2008	470
a58122	186	05/29/2007	08/11/2008	441
a302	221	06/28/2006	08/11/2008	776
a70050	210	05/07/2007	08/11/2008	463
a70068	190	05/10/2007	08/11/2008	460
a1155	173	05/25/2007	08/07/2008	441
a94870	155	05/29/2007	08/11/2008	441
a61930	177	05/24/2007	08/11/2008	446
a66613	167	02/28/2007	08/11/2008	531

Most data is drawn from the narrowest date range. Among 580,018 non-missing records, 121,899 pairs of records are found to be identical. The resulting 458,119 transactions contain 221 pairs that have the same values for the all attributes except the balance field that indicates the remaining amount to be cleared. After excluding older records, the final dataset has 457,898 observations. The descriptive statistics of the final dataset are listed in the table 3 and 4, sorted by variable and account.

www.manaraa.com

Table 3. Summary Statistics by Amount

account	variable	n	nmiss	avg	median	std	min	max
a1155	Amount	694	0	7434.65	1777.02	19147.03	1	170073.99
a21776	Amount	25719	0	2084.08	607.54	7235.72	4.95	440000
a21830	Amount	21983	0	1036.28	68.04	8130.35	0.01	843000
a302	Amount	5116	0	690248.2	286.35	14636008	0.01	418030303
a32360	Amount	62916	0	666.39	50	7829.16	0.01	899348.33
a45136	Amount	65289	0	216625.5	236.7	4347807.5	0.01	311084647
a5738	Amount	133564	0	706.29	5.4	26615.34	0.01	4252752.5
a58122	Amount	18021	0	3532071	109614.3	11406518	0.01	309072377
a60836	Amount	79652	0	38148.42	7518.42	601805.66	0.01	70000276
a61042	Amount	19283	0	5855.88	368.88	225951.27	0.01	30040000
a61930	Amount	729	0	5037741	900000	13936447	15.63	230000642
a66613	Amount	773	0	9765888	10000	216507723	0.03	5899996308
a68128	Amount	19755	2	43530.23	177.79	615870.99	0.01	31867577.1
a70050	Amount	3010	0	7905.24	284.19	109117.21	0.01	4261950.73
a70068	Amount	915	0	429573.4	1600	8036053.8	0.01	241203449
a94870	Amount	479	0	37519.1	10000	88582.1	0.4	900037.44

Table 4. Summary Statistics by Balance

account	variable	n	nmiss	avg	median	std	min	max
a1155	Balance	694	0	424.64	0	7057.19	0	163907.49
a21776	Balance	25719	0	85.46	0	1052.05	0	67824.68
a21830	Balance	21983	0	89.99	0	1004.78	0	58905
a302	Balance	5116	0	140012.8	0	6183851.2	0	386445649
a32360	Balance	62916	0	75.59	0	2388.8	0	465570
a45136	Balance	65289	0	4436.78	0	753667.69	0	189000000
a5738	Balance	133564	0	48.41	0	4088.69	0	820000
a58122	Balance	18021	0	7705.93	0	589623.84	0	74220000
a60836	Balance	79652	0	0	0	0	0	0
a61042	Balance	19283	0	1754.47	0	216470.28	0	30040000
a61930	Balance	729	0	473788.9	0	4445412.3	0	105000000
a66613	Balance	773	0	171286.4	0	3404279.7	0	92999468.6
a68128	Balance	19755	2	48.96	0	1140.55	0	101701
a70050	Balance	3010	0	444.06	0	3133.4	0	100000
a70068	Balance	915	0	1810.18	0	21662.35	0	502144.85
a94870	Balance	479	0	304.32	0	5380.13	0	116895.33

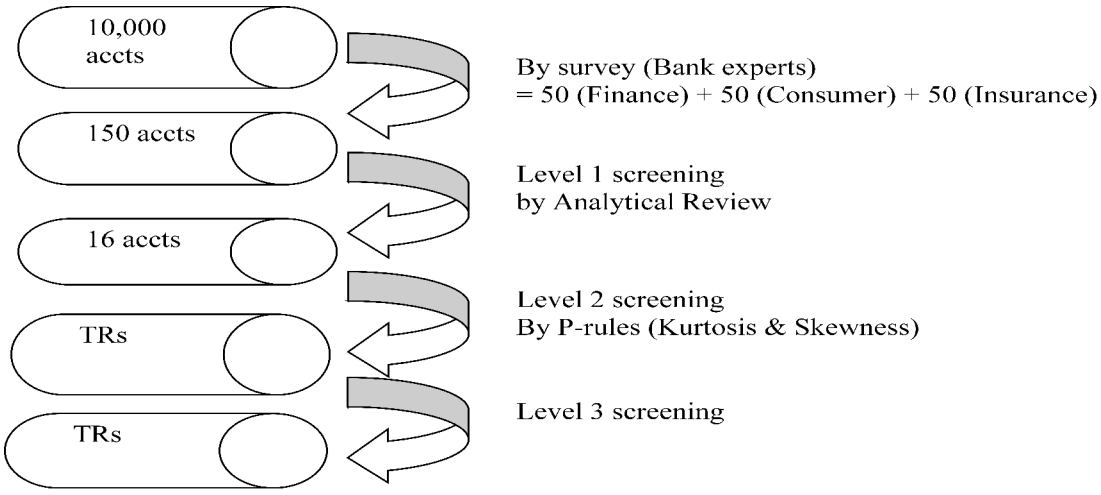
www.manaraa.com

Screening rules

The IC screening model in this study is a collection of rules that will be applied to the transactions to detect anomalies. As a first step, I rely on the scenarios or cases that are of the most concern to internal auditors. Materiality of transaction amounts is the primary concern of internal auditors when they investigate the potential for frauds. This leads this study to focus more on transactions with sufficiently large amounts than those with small amounts. The second concern is whether the model is implementable in their systems. Consequently, the model should be one that requires as little computational power as possible. Finally, auditors find manual entries riskier than automated entries. Considering these as minimum requirements, several monitoring rules are developed and tested. The overall blueprint for the internal control system implementation in this case is in the figure 2.

www.manaraa.com

Figure 2. Selection of Transitory Accounts



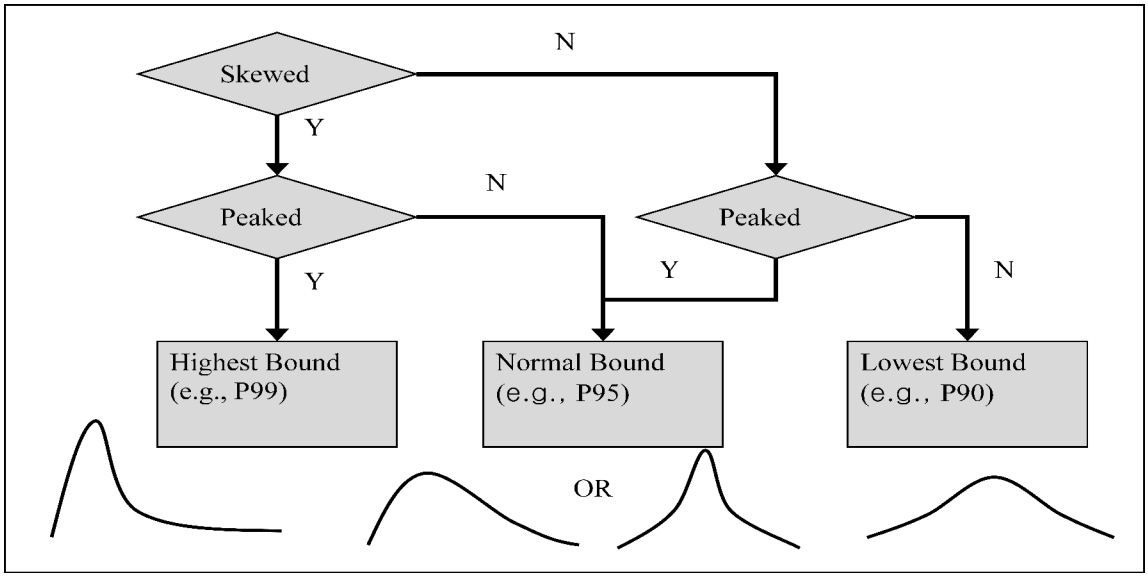
Practitioners select particular accounts based on these decision efforts. Based on the selected transitory accounts, the initial or general screening (level 2) will be performed and then more detailed screening rules (level 3) will be applied. The level 2 screening rules support mainframe level implementation while the level 3 rules are aimed at terminal or personal computer level monitoring. Depending on the DP power of the systems, the level 3 monitoring rules may also be implemented at the mainframe level in addition to level 2.

To develop general screening rules that do not consume too many system resources, P-rules are generated. A P-rule is a collection of procedures to find transactions with

www.manaraa.com

material amounts. These procedures utilize the fact that the majority of transactions have relatively small amounts, making the distribution positively skewed and highly peaked. The overall logic of the procedure is shown in the figure 3.

Figure 3. P-rule using Skewness and Kurtosis



If thresholds for material amount that should be examined decreases, its location represented by percentile will be lowered according to skewness and kurtosis. Assume there are two datasets with identical numbers of transactions and identical means. If all amounts are extremely positively skewed (decision degree of skewness=1) and peaked (decision degree of kurtosis=1), only a few observations may exceed the material level.

www.manaraa.com

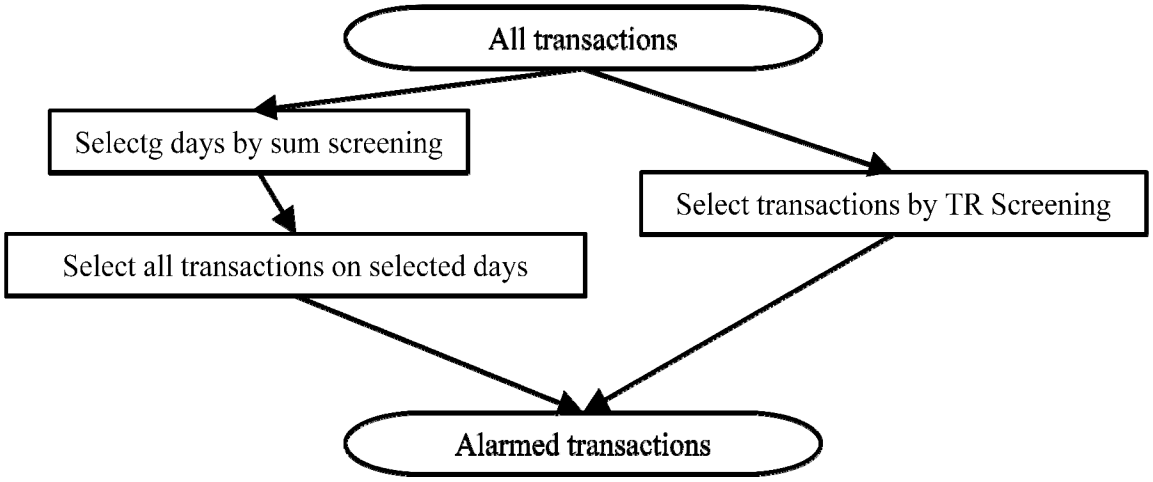
On the contrary, if the distribution is less positively skewed and less peaked, more observations are likely to exceed the cutoff point. The main reason to use this distribution information is to reduce computational cost. In better DP systems, prediction intervals may substitute for percentile rules. Furthermore, although the decision criteria for the degrees of skewness and kurtosis are 1 in this study, these are arbitrary cutoffs. The parameters for cutoff distributions can be changed as well.

Based on the P-rule, two screening methods can be applied to each account. Either the system can examine daily sums to select suspicious days and then and investigate the all transactions on those days, or it can flag abnormally high transactions directly. Each method comes with advantages and disadvantages. If anomalies (likely internal fraud) occur on a particular day, that day's sum will be abnormally high, allowing easy detection by the first method. In contrast, if frauds occur over long periods or at random intervals, the latter will be more suitable for anomaly detection. However, each method also has drawbacks. Certain days may have large sums simply because of abnormally high transaction volume. In this case, even the highest values on a particular day can be materially insignificant. In another case, a day can have a small daily sum that results

www.manaraa.com

from a few transactions with significantly large amounts. If so, some transactions that exceed the material amount cannot be detected by applying the P-rule to daily sums. In this study, both methods will be used to mitigate these issues. After applying the two screening rules, their union set will determine final flags for further investigation. This process is illustrated in the figure 4.

Figure 4. Level 2 screening



Regarding the manual entries, this study utilizes an indicator showing whether a transaction is either manually or automatically processed. If its value is either 0 or 999999999 (9 9s), the record is automatically entered and manual otherwise. We expected that manual entries would be more frequently flagged by level 2 screening rules.

www.manaraa.com

However, the results show the opposite- automatic entries were more likely to be flagged by level 2 screening, as shown in table 5.

Table 5. Manual vs. Automatic Process

Type	Population	Alarmed	Percentage
Manual	53,591	465	0.87%
Automatic	404,307	4,420	1.09%
Total	457,898	4,885	1.07%

Since large transaction amounts are likely to be flagged, fraudsters may elect to split such transactions into smaller amounts, either as two identical sums or as two slightly different amounts. In order to produce more reliable result, only the former case is considered in this study. Future study may test the latter.. Interestingly, the results say that there are many duplicate amounts. In an extreme case, there are 39 duplicate amounts in a branch in a day. Considering that amounts of the transactions after the level 2 rules are large, this observations are clearly anomalous. If both are considered, this may indicate anomalies that are more likely fraudulent. The table 6 summarizes the result.

www.manaraa.com

Table 6. Duplicates by Branch

Duplicates	2	3	4	5	8	11	12	23	24	27	39
Branches	93	9	4	3	2	1	1	1	1	1	1

Alarm volume is another potential indicator. If a branch has more alarms than others, this may signal potential problems. In this study, a branch is considered as risky if it has two or more alarms on a specific day. The detailed summary is on the table 7.

Table 7. The Comparison: the Number of Alarms

	Population				All alarmed TRs				Alarms >= 2			
Acct	Day *	Obs	Day	Branch	Day *	Obs	Day	Branch	Day *	Obs	Day	Branch
302	2,577	5,116	221	729	30	52	28	5	7	29	7	1
1155	635	694	173	297	6	6	6	6	.	.	.	.
5738	22,350	133,564	518	1,235	658	1,335	210	312	90	767	68	58
21776	16,781	25,719	226	1,130	471	557	126	253	46	132	29	28
21830	15,142	21,983	232	1,157	208	219	154	57	8	19	8	3
32360	28,455	62,916	233	1,021	569	629	180	241	40	100	31	26
45136	32,359	65,289	244	1,189	435	652	164	34	122	339	83	10
58122	372	18,021	186	3	117	192	99	2	59	134	55	2
60836	276	79,652	202	2	181	797	181	1	165	781	165	1
61042	5,825	19,283	227	992	157	191	100	63	26	60	23	13
61930	177	729	177	1	7	7	7	1	.	.	.	.
66613	569	773	167	325	7	7	7	4	.	.	.	.
68128	12,321	19,757	226	1,045	156	197	108	21	34	75	33	3
70050	1,931	3,010	210	642	29	30	25	21	1	2	1	1
70068	636	915	190	132	9	10	9	5	1	2	1	1
94870	155	479	155	1	3	4	3	1	1	2	1	1
All	140,561	457,900	535	1,358	3,043	4,885	233	627	600	2,442	192	122

www.manaraa.com

Since each account should exhibit either positive debit or positive credit amounts, any negative transaction would be flagged as abnormal. After a pilot test, however, it was determined that literally no cases violate this rule.

There are a few other indicators to be tested. Aging of transactions was considered. Since transitory accounts are meant for temporary storage, items should not be kept there for a long time. However, pilot testing showed many transactions that remained for more than 6 months. Another candidate indicator is the use of relationships among accounts. This needs more information about all the accounts and their relationships. And more indicators will be developed as more information becomes available.

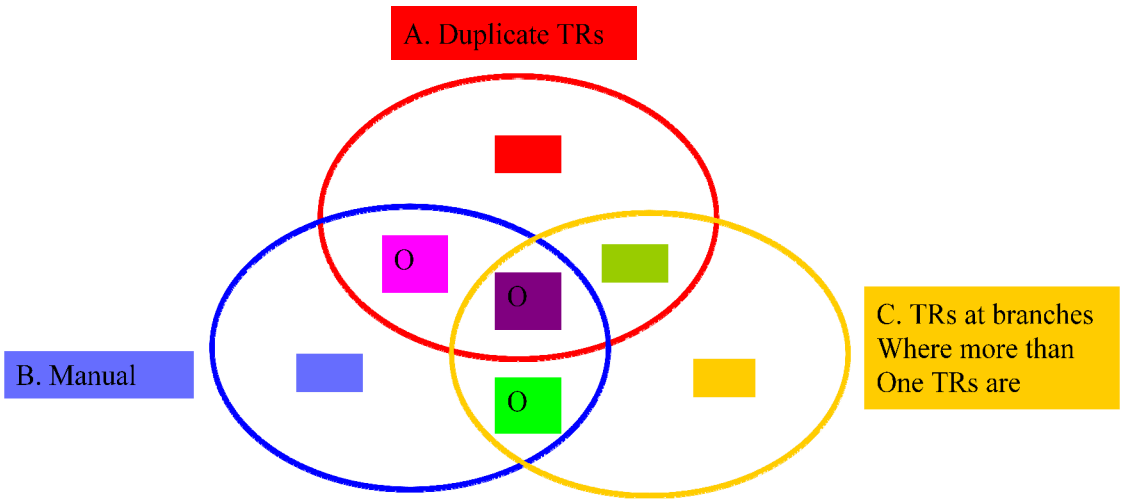
Results and Discussion

Since each level 3 screening rule can be given different weighting, Venn diagramming may prove informative as illustrated in the figure 5. Items appearing in more than one circle can be more significant than the others, though not necessarily. Final selection was

www.manaraa.com

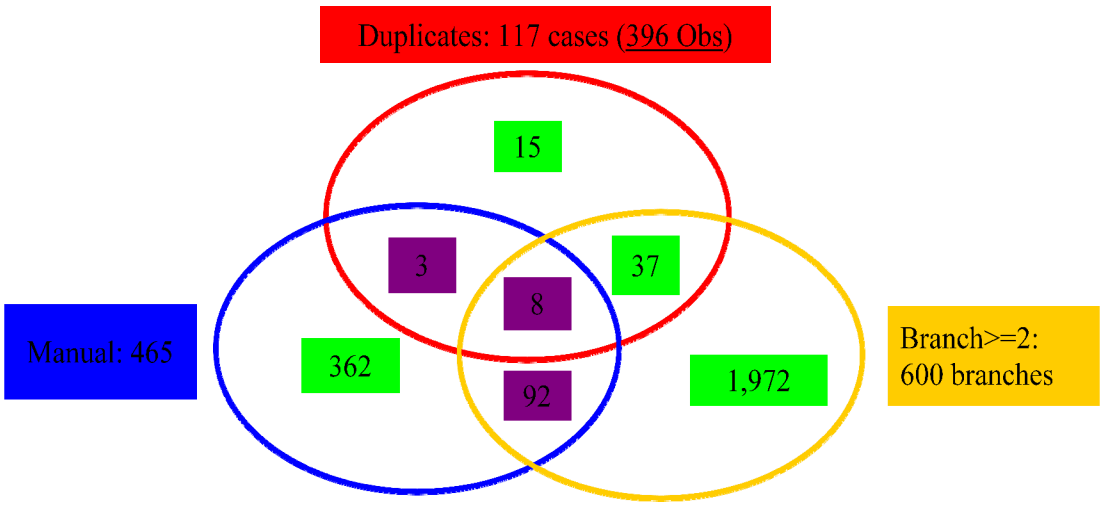
made with the following assumption. It would be really rare to have transactions whose amounts were the same and large in a branch in a day. If they were fraudulent, its entries must have been entered manually rather than automatically. Based on this assumption, the final results are as shown in the figure 6.

Figure 5. Selection of Transactions by Venn-Diagram



www.manaraa.com

Figure 6. Actual Selection of Transactions by Venn-Diagram



After all testing, 103 out of 2489 transactions were flagged for further testing. This small figure may be due to the use of rules that were developed to reduce false positives. If parameters are changed, the number of observations may increase significantly.

Instead of assuming that the dataset in the study was audited and did not have any errors, this study utilizes actual confirmation by internal auditors. Since the auditors believe that some transactions are truly fraudulent, it may not be practical to assume that this dataset is truly anomaly-free. Also, if the alarmed transactions are truly fraudulent, the power of the IC model can be easily confirmed. Moreover, even if the model does not detect any real fraudulent cases, it can still act as a deterrent, assuming that employees

www.manaraa.com

are aware of its existence and possible consequences if they are caught. In any case, the model will be useful for the business and as a landmark for future study.

b. Phase II (February 2010)

Data

While Phase I shows a wide range of information regarding the transitory account system, its results clearly suggest the necessity of further investigation to improve its effectiveness. To that end, the dataset in this phase is separated into two parts: one for model training and the other for testing. If the transitory accounts in this study have been used for a long time and their transactions have similar patterns over the time, then it can be assumed that a model developed with a data set in the past can predict behavior of transactions of a data set in the future. With this assumption, an anomaly detection model in this phase is developed with a training set and tested with a test set.

Due to sporadic data extraction, the data in the Phase II has a time gap between two parts. While the training set ranges from January 2008 to November 2008 (11 months), the test set is from December 2009 to February 2010 (3 months). Potential effects of this

www.manaraa.com

time gap among data sets can be are unknown, so results should be interpreted with caution. However, considering that each transitory account was created with the same purpose, transactions ought to have similar behaviors unless the business environment has changed significantly. As in Phase I, data cleaning was done by discarding transactions beyond common date ranges. After data cleaning, 400,466 transitory accounts remained for training and 75,236 remained for testing. However, the test sets have unexpected outcomes for less frequently used accounts. Not all accounts are used intensively and those with small numbers of transactions (e.g. 61930, 94870) have become almost dormant. Since the anomaly detection model is developed and tested for each account, this natural selection will not affect the overall performance. However, it is no doubt that those extinct accounts should be discarded to improve model's accuracy in the future. The table 8 shows the detail for each account.

www.manaraa.com

Table 8. Change in Transitory Accounts

Account	Table	cnt	min_amtDate	max_amtDate	range_amtDate
1155	Train	688	01/15/2008	11/20/2008	310
21776	Train	28744	01/15/2008	11/20/2008	310
21830	Train	23950	01/15/2008	11/20/2008	310
302	Train	5654	01/15/2008	11/20/2008	310
32360	Train	67188	01/15/2008	11/20/2008	310
45136	Train	73389	01/15/2008	11/20/2008	310
5738	Train	49539	01/15/2008	11/20/2008	310
58122	Train	20395	01/15/2008	11/20/2008	310
60836	Train	91660	01/15/2008	11/20/2008	310
61042	Train	12114	01/15/2008	11/20/2008	310
61930	Train	1042	01/15/2008	11/19/2008	309
66613	Train	2426	01/15/2008	11/20/2008	310
68128	Train	19568	01/15/2008	11/20/2008	310
70050	Train	2715	01/15/2008	11/19/2008	309
70068	Train	891	01/15/2008	11/19/2008	309
94870	Train	503	01/15/2008	11/20/2008	310

Table	cnt	min_amtDate	max_amtDate	range_amtDate
Test	63	12/22/2009	02/18/2010	58
Test	2476	12/02/2009	02/23/2010	83
Test	4676	12/01/2009	02/23/2010	84
Test	1056	12/02/2009	02/23/2010	83
Test	13652	12/01/2009	02/22/2010	83
Test	14745	12/01/2009	02/23/2010	84
Test	2620	12/01/2009	02/23/2010	84
Test	1620	12/23/2009	02/23/2010	62
Test	24668	12/23/2009	02/23/2010	62
Test	5872	12/01/2009	02/23/2010	84
Test	3	01/05/2010	02/18/2010	44
Test	24	12/14/2009	02/22/2010	70
Test	3586	12/01/2009	02/23/2010	84
Test	78	12/09/2009	02/23/2010	76
Test	97	12/03/2009	02/23/2010	82
Test	.	.	.	.

Screening rules

Anomaly indicators were extended in this phase to increase the model's effectiveness. Although a strong indicator such as duplicate records is clearly direct evidence of an anomaly, it represents only one anomaly type. In order to increase the overall effectiveness of the model, other indicators must be included. To meet this end, five anomaly indicators were added to those from Phase I based on various anomaly scenarios and data analyses.

Newly introduced indicators are 1) age of transaction, 2) weekend initialization, 3)

www.manaraa.com

weekend clearance, 4) clearance before initialization, and 5) duplicate transaction numbers. Transaction age is critical for a transitory account. An excessively long stay in a transitory account implies unusual difficulty in determining destination. It is well known that a lack of sufficient employee monitoring facilitates internal fraud and/or permits more errors. There are generally fewer observes on weekends, leading to the two aforementioned anomaly-generating scenarios. Transactions should be cleared after initialization; the opposite sequence indicates an anomaly. Transaction numbers must be unique in order to discriminate a record from others. Duplicate records will cause malfunctions and unexpected outcomes under a relational database system. In addition, it is also possible that one transaction with the same transaction ID will mask another if a transaction number is the main source to identify a transaction. To detect this type of anomaly, an indicator detects transactions with identical numbers.

Results and Discussion

The parameters for level 2 screening developed by the training set are applied to the test set. After applying level 2 screening, 2,066 transactions are left for level 3 screening.

www.manaraa.com

At level 3, the 9 anomaly indicators are tested with the remaining observations. 993 transactions are flagged by one or more anomaly indicators at level 3. The large number of indicators makes Venn-diagramming less useful. Instead, a table is used to present the level 3 screening result. The table 9 shows the detail about the level 3 screening.

Table 9. Summary of Flagged Transactions

Wrong_Dr Cr	manual	Dup	Aging	amt_week ends	bal_weeke nds	wrongDate	DupTransa ctionID	multiFlags	cnt
0	0	0	0	0	0	0	0	0	1073
0	0	0	0	0	0	0	0	1	521
0	0	0	0	0	0	0	1	0	2
0	0	0	1	0	0	0	0	0	4
0	0	0	1	0	0	0	0	1	3
0	0	1	0	0	0	0	0	1	365
0	0	1	1	0	0	0	0	1	4
0	1	0	0	0	0	0	0	0	60
0	1	0	0	0	0	0	0	1	30
0	1	0	1	0	0	0	0	0	4

It is difficult to decide the number of flagged transactions to be recommended for further investigation. As shown, the most of the transactions are flagged by one or two indicators. Only 4 transactions are flagged by the three anomaly indicators. Since the relative importance of each indicator is not objectively measured, it is difficult to identify anomalous transactions. Although Phase I criteria can be used as a touchstone, newly added indicators necessitate further analysis, requiring internal auditor expertise.

www.manaraa.com

However, this approach relies too much on feedback from the internal auditors that usually takes a long time. In order to avoid time-consuming communication process, an alternative solution can be a larger set of anomaly indicators for level 3 screening. While no transaction was flagged more than three times, it is possible that other, more discriminating anomaly indicators were not included. However, the expansion of anomaly indicators is not a valid solution. The limited number of available variables prevents further investigation. For example, clearance values are not available. Since some transactions have balances that are less than the original amounts, partial clearance is possible. If it is the case, information about transaction clearance can be useful to create further anomaly indicators.

www.manaraa.com

c. Phase III (June 2010)

Data

This phase focuses on the expansion of anomaly indicators with attributes that are not used in the previous models. After much discussion with the internal audit department, the bank decided to provide additional data related to the regularization of transactions. Regularization is the process of zeroing a transaction balance by finding its destination or receiver. Due to the sensitivity of the data, its extraction and delivery took more time and processes than previous ones. Eventually, three months' regularization data was provided in addition to the original transactional data.

As in Phase II, the datasets in Phase III consist of two parts: one for model development and the other testing. The train data set spans three months and the test set spans three months. After data cleaning, 75,236 transactions remain for the training set and 54,768 for testing. Details are illustrated in the table 10.

www.manaraa.com

Table 10. The Number of Transactions: Train vs. Test Set

Account ID	cnt_train	cnt_test
1155	63	27
21776	2476	1322
21830	4676	3356
302	1056	702
32360	13652	7615
45136	14745	9864
5738	2620	2706
58122	1620	1650
60836	24668	23708
61042	5872	2437
61930	3	6
66613	24	13
68128	3586	1160
70050	78	75
70068	97	127
Total	75236	54768

Screening rules

To capture novel anomaly characteristics, the newly provided data set is analyzed with various variables. After extensive analyses, three anomaly indicators are found to be effective in anomaly detection: number of regularizations, age of regularization, and manual regularization.

Ideally (and typically), a transaction is regularized in a single effort. In some cases, however, a transaction requires multiple regularizations, reducing individual transaction amounts and thus increasing fraud likelihood, similar to splitting a wire transfer.

The regularization process generally takes one day or less. Time until first regularization can indicate anomalous behavior because it indicates the first action made

www.manaraa.com

to the transaction.

Regularization processes are typically handled automatically. However, it occasionally needs human interventions to settle down the transaction. Since manual processes are more likely to be erroneous or fraudulent, manually processing can be a valuable indicator. With addition of the three indicators, the Phase III model consists of 12 anomaly testing rules.

Results and Discussion

After applying level 2 screening rules with parameters decided by the training set, 529 transactions in the test set are selected. Among those, 248 transactions are flagged by one or more transactions as shown on the table 11.

Table 11. The Number of Flagged Transactions by Score

score	0	1	2	3	4
cnt_transactions	281	202	37	8	1

The number of flags for each transaction is increased so that it is much easier to

www.manaraa.com

decide a manageable number of flagged transactions for further investigation (i.e., 46 transactions whose suspicion scores are greater than 1). Considering the number of newly added anomaly indicators, Phase III appears quite different from Phase II. Flags by indicators are summarized in the table 12.

Table 12. The Number of Flagged Transactions by Anomaly Indicator

DrCr	0	0	0	0	0	0	0	0	0	0
wrongDate	0	0	0	0	0	0	0	0	0	0
Manual	0	0	0	0	0	0	1	1	1	1
Duplicates	0	0	0	0	1	1	0	0	0	0
Age	0	0	0	1	0	0	0	0	0	0
amt_wk	0	0	0	0	0	0	0	0	0	0
bal_wk	0	0	0	0	0	0	0	0	0	0
Duplicate transaction ID	0	0	0	0	0	0	0	0	0	0
multi_regul	0	1	1	0	0	0	0	0	0	1
manual_regul	1	1	1	0	0	1	0	1	1	1
regul_wk	1	0	1	1	1	0	1	0	1	1
Age to first regul	0	0	0	1	0	0	0	0	0	0
Suspicion Score	2	2	3	3	2	2	2	2	3	4
cnt_transactions	16	4	5	1	7	2	6	2	2	1

Although this result was not investigated by the internal auditors due to a lack of human resources to be assigned, it shows a potential problem that can be encountered while developing an anomaly detection model and its feasible remedy. As long as a rule-based model is used to develop a model and the number of anomaly indicators is not

www.manaraa.com

sufficiently high, the problem will persist. Alternatively, relative weighting on individual indicators can be used to discriminate flagged transactions. However, as discussed early in this study, it is impractical to measure relative importance of anomaly indicators objectively until more characteristics of anomalous transactions are uncovered.

iv. Conclusion, Limitations, and Future Research

In this study, three anomaly detection models are suggested by using transactions of a bank's transitory accounts. Phase I serves as a pilot study while Phases II and III are used to improve it. Although various attempts have been made to enhance the accuracy and effectiveness of the model, the number of available variables limits the creation of a sufficiently number of anomaly indicators to capture the true characteristics of anomalous transactions. This may be due to the fact that the detection model is a general model that is applied to all transitory accounts.

Future study may consider account-specific models. Some transitory accounts have similar numbers of transactions with similarly distributed amounts, while others are significantly different from one another. If an IC screening model includes this

www.manaraa.com

information, it may filter the anomalies more accurately. However, the number of models will inevitably increase as the degree of heterogeneity among the transitory accounts increases. In the worst case, each account would be distinctive enough to require its own screening model, proving far too complex to be practicable. Although the bank correspondents mention that there are three types of transitory accounts- finance, consumer, and insurance- this categorization may not be sufficient to convey the unique features of each transitory account. Direct evidence about account differences can be analyzed through descriptive statistics. For comparison, assume that two numbers are deemed similar to each other if their difference is below 10 per cent of the smaller. The differences between accounts 1155 and 61930 then become clear. Both accounts have similar date ranges (441 and 446 days) and similar numbers of transactions (694 and 729). However, their medians are significantly different (1,777.02 and 900,000.00). By the general screening rules, many transactions in the account 61930 will not be flagged even if those amounts are beyond the predetermined material level since most transactions have large amounts. On the contrary, the flagged transactions of the account 1155 may have relatively small amounts that far below the material level. Consequently, the general screening model that does not consider these account-specific characteristics may not

www.manaraa.com

work properly unless the characteristics of all the accounts are sufficiently homogeneous.

Considering this potential drawback, the first step to develop account-specific screening models should be to identify and group accounts with similar characteristics. Components can be grouped by date range, number of distinct days, number of transactions, category, or descriptive statistics such as means, medians, and standard deviations.

Another suggestion is the use of relationships among variables. The screening rules so far are mainly for individual attributes such as transaction amounts and manual/automatic indicators. Some attributes are closely related to each other. One example is the association between the transaction amounts and the balance amounts. By definition, the transitory accounts are designated to keep unidentified or insufficiently identified money temporarily, so balances should be zeroed rapidly. One way of analyzing this relationship will be a continuity equation that assumes that the inflow and outflow of a system are the same in equilibrium. Furthermore, we may use a multiple (time series) regression method that includes both numerical and categorical variables. Clustering methods can also divide transactions into several groups and identify those

www.manaraa.com

with the fewest frequencies. In the collection of screening rules, inclusion and exclusion of a rule can be easily processed since each rule is an exclusive piece and does not affect the other rules. If a new screening rule is added, the only change will be additional transactions flagged by the rule. However, in multiple variable models, addition and deletion of rules require complex understanding. For example, if a rule is removed in a network model, the resulting transactions flagged by the model can be very different from those before the change and interpreting the differences is not intuitive. Multiple variable models utilize not only the variables themselves but also the relationships among them. Consequently, if a variable is deleted or added, its impact will extend to other variables.

www.manaraa.com

2. Case II. Development of An Anomaly Detection Model for An Insurance Company's Wire Transfer System

i. Introduction

Modern industries use various data processing systems to optimize resource use and operations. Although new products are frequently introduced, many companies still maintain their old systems, referred to as legacy systems.

A legacy system is a computerized data processing system that is highly customized and system-specific. Although less efficient than cutting edge database management systems, well-structured and maintained legacy systems may function as well as those in the current market. However, companies often try to adopt newer computer systems and migrate the existing systems for their proven long-term benefits. Despite these benefits, migration into a new system is not always an easy task, especially when a company cannot easily justify tremendous initial investments and maintenance costs. Some companies convert their legacy systems into the new systems gradually rather than abruptly, but this is only practical for a company that continually absorbs others to

www.manaraa.com

expand its business.

This study's subject is one such company, and it has adopted a gradual migration approach. The companies bought by the subject usually have their own data processing systems that are not 100% compatible with the subject's. To resolve this problem, the company keeps the system of the merged company and migrates transactions of the merged company into its system. This complicated data processing structure diminishes the company's capacity to directly monitor the transactions of the acquisition target.

Although mergers and acquisitions are frequent, the difficulty of system and database integration provides a high barrier. The lack of information about anomalous transactions serves as another obstacle in this study. Fundamental causes of these obstacles are as follows.

First, most data fields in the company's system are manually entered because of complex migration processes and less control. In highly sophisticated data processing systems (e.g. ERPs), most fields are entered either automatically or semi-automatically as input controls. For example, timestamps and dates may be automatically entered and product codes may be semi-automatically via drop-down menu selection. Well-designed

www.manaraa.com

legacy systems may have the same capacities, but that is not the case for this company. Manual entries are more common, causing data integrity problems such as referential and domain integrity violations. For example, a customer's name can be entered differently by two different individuals. As a result, there are many exceptional cases that would otherwise be considered normal. Although information about these cases will to improve IC in the future, they will be major obstacles to the construction of effective screening rules.

The company lacks historical information about its own anomalies including fraudulent scandals, which serves as another obstacle. Absence of evidence is not evidence of absence. Instead, it might indicate that the company internal control system could not detect them or did not have any modules for anomaly detection. An anomaly detection model must then be built from scratch. In this case, we can make one of two assumptions. Since little is known about the degree of anomaly in the company, we may assume that the data is audited and free of irregularities and material errors, which is a strong but necessary assumption for a fraud detection study. However, this assumption may be inappropriate considering that external auditors are not responsible for fraud

www.manaraa.com

detection itself although they are responsible for evaluation of internal control system in terms of material misstatements. Consequently, it may be more appropriate to assume that the data may contain some irregularities and/or material errors.

This study involves a major US insurance company that is developing a continuous audit / fraud detection process. To the end, it was decided that a research team would cooperate with the internal audit organization to develop basic modeling and analysis methodologies in parallel with the internal audit process. The project plan entailed a set of progressive steps in the development of an automated discrepancy detection process. Once the processes and models are developed, the data extraction is processed more frequently and systematically, progressing towards more frequent data screening to monitor potential fraud. The proposed model is similar to an external stand-alone system used to extract and screen data for exceptions in continuous auditing (Vasarhelyi & Halper, 1991; Searcy and Woodroof, 2001; Rezaee et al., 2002; Murthy and Groomer, 2004). Wire payment data is extracted from legacy systems and analyzed externally. This is beneficial since running an automatic fraud detection system can be intensive on the production system which might cause the system to operate sub-optimally. Pathak et al.

www.manaraa.com

(2005) found that auditing transactions in batches was more cost effective than initiating periodic audits after a certain period of time. Our model proposes batch fraud detection. Before an audit, internal auditors can extract desired data and run the fraud detection mechanism. Any resulting exceptions can be investigated during their regular audit. The wire transfer process was chosen as a desirable first target because of: 1) data availability, 2) the volume and importance of the process, 3) availability of knowledgeable and competent internal audit staff for knowledge engineering, and 4) the timing of the audit.

The wire transfer payment process did not seem as well-controlled as other processes. Furthermore, the company did not have documented historical information about past fraud occurrences. However, this lack of past experience does not imply lack of wire fraud. Consequently, we chose to use an unsupervised method to create a statistical model for anomaly detection within wire payment transactions. Internal auditors would investigate selected transactions for anomalies.

Indicators were divided into two methods of analysis: conditional and statistical. The conditional indicators are pass/fail type tests and the statistical indicators are tests that utilize statistical methods such as correlation. Each indicator is equally weighted although

www.manaraa.com

it is likely that certain indicators are more important than the others. The reason is that the absolute degree of each indicator's effect on the final decision cannot be measured in a systematical as well as globally agreeable way. Those wires whose suspicion scores are higher than a threshold are flagged and forwarded to the internal audit team for detail testing. Investigation results and feedback from the audit team become a direct input to modify the model for fine-tuning.

ii. Objectives

Despite continuing frequent use of paper checks, electronic fund transfers (or wire transfers) have been gradually eclipsing their use in practice. Use of wire transfer has become popular because it is convenient and economical, requiring less human effort and physical resource consumption.

However, a wire transfer has its disadvantages. Wire transfers leave few or no verifiable physical traces. Although audit trails and logs exist in a processing system, their interpretation requires database expertise. As a result, most of the information is highly vulnerable to unauthorized modification if not appropriately managed.

www.manaraa.com

This apparent drawback is more significant for a cash outflow than a cash inflow because the former decreases company resources while the latter increases them. In response to this concern, this study develops an anomaly detection model for a wire transfer payment system of an insurance company.

More specifically, an anomaly detection model for the wire transfer payments is developed in order to identify both fraudulent and erroneous transactions. The model consists of a series of anomaly indicators designed to detect abnormal wires. Each wire goes through the model and its suspicion score is calculated. If a score is beyond a certain threshold, that transaction is labeled as potentially anomalous and forwarded to the internal audit team for further investigation. After investigation by the audit team, the model is fine-tuned based on results and their feedback.

iii. Methodology and Results

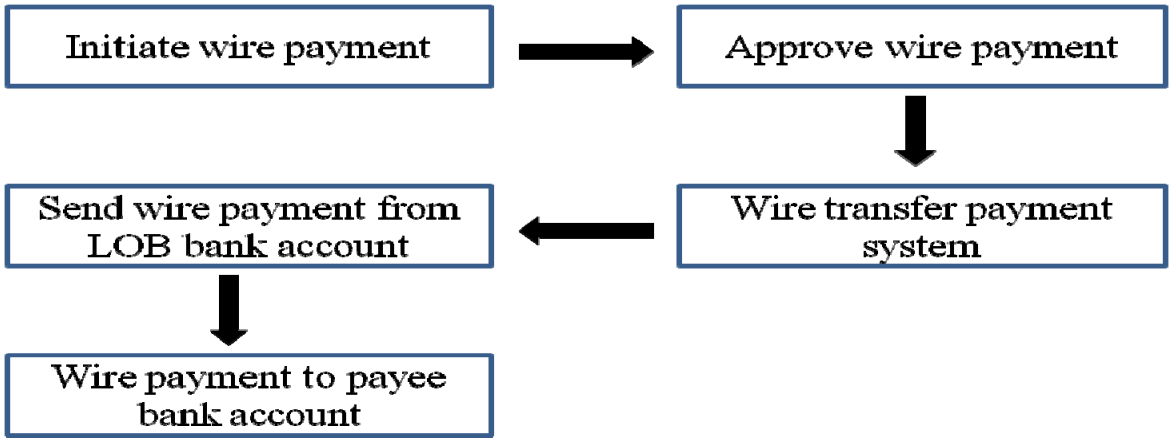
a. Overall

The wire transfer process in the company consists of three stages: initiation, approval, and settlement (or payment). Each wire transfer requires one initiator and one approver at

www.manaraa.com

a minimum. Depending on the nature of the wires, certain types of wires require two approvals when they do have prior information. Once a wire is approved, it is imported into the payment system as shown in the figure 7.

Figure 7. Wire Transfer System



Although computerized systems exist, the company also maintains all relevant physical documents. When a wire needs to be processed, physical documents are filled out to record customer information. After the documents are complete, some- not all- data fields are manually entered into the system. Although the internal auditors did not provide a clear reason for physical documentation, we have a hypothesis. The reason for maintaining both physical and electronic systems is an artifact of continual mergers and

www.manaraa.com

acquisitions. It is not a surprise that the database management systems of the companies of an M&A have a different data structure.. When the system compatibility is not easily resolved, a possible solution will be to maintain both systems and merge the two systems gradually. Eventually, only one system will remain. Unless a merged company changes the forms for existing customer information, which is costly, it will be inevitable to lose some of the customer information before an M&A. Hence, the use of physical documentation may be a necessary cost-reducing measure.

Another problem with frequent acquisitions is that similar files with the same purpose can exist in multiple systems. Although the main system of the company governs all sub-systems that the merged companies are running, it is practically difficult to force the main system to encompass all sub-systems, especially when another merger is likely in the near future. An economical solution may be to maintain a separate file in each sub-system, which is likely to cause a data discrepancy problem.

The most imminent discrepancy-related issue is the difficulty of enforcing common input and output controls. Entire systems can be vulnerable to entity, referential, and domain integrity violations. An entity integrity violation occurs when a specific

www.manaraa.com

transaction cannot be identified because of duplicate primary key values. Since each transaction has a unique identifier, entity violations do not present a significant problem.

The other two integrity violations are more common. A referential integrity problem occurs when a referencing table does not appear on its master file. For example, when a customer master file is recorded by the customer service department and a salesperson enters a customer name manually, that customer name may not exist on the master file. This discrepancy may exist until the two tables are reconciled. If referential integrity is strictly enforced, the sales department cannot input a customer record whose information is not on the master file. This can also happen when a referencing systems use different names for the same customer. For example, the sales department may enter 'Traveler' or 'Travelers' in the place of 'Travelers Co.'. This problem will become more serious when the field allows manual inputs.

A typical example of a domain integrity violation is skipping a required value or entering an inappropriate character type. Domain integrity is violated if a string of characters is entered into a numeric field (e.g. 'one hundred' instead of '100').

Historical records are not always available. For example, if the authorization table

www.manaraa.com

contains employee's identification number and their authorization limits, the values of the authorization limits may change along with the employee's job status. However, unless the tables are designed to continuously record changes, it is likely that the backup file will contain a snapshot at the moment of the last backup. Unless backups are frequent, some information may not be available.

Although an uncommon occurrence, internal auditors may lack information regarding their company's own fraud or material mistakes. Lack of past information leads to difficulty determining vulnerable internal control areas. Consequently, an anomaly detection model must be developed from a scratch, using intuition and observation as starting points. To summarize, some indicators in the previous case study will be used in this study as a starting point, while new screening rules are to be developed to meet project goals.

Last but not least, model development and testing are performed quarterly, with each set of results serving as an an input to next quarter's test. Although the latest model is the most extensive and sophisticated, it will be more beneficial to illustrate all the developing models and their results than to discuss only the latest one, especially when discussing

www.manaraa.com

new revisions.

b. Phase I (September 2008)

Data

The dataset in this study consists of about 230,000 wire payments paid to over ten thousands of payees from October 2007 to September 2008. Approximately 90% of the wire transfer payments belong to about 10.25% of the payees. More than half (62.82%) of the payees are engaged in only one transaction, and the majority (93.84%) of the payees have fewer than 30 transactions. The datasets provided by the insurance company include seven tables (or files). The table primarily used in the study is All_Wires, containing 27 attributes. After removing irrelevant records (e.g. monthly amount totals), 229,531 remain. The other six tables are master files that are referenced by the All_Wires table. The master files keep employee information such as start date, employment status, rank status, and authorization limits. The attributes are used mainly to check employee authorization limits. Descriptive statistics for four numeric attributes- wire amounts, initialization limits, approval limits, and settlement limits- are shown in the table 13.

www.manaraa.com

Table 13. Statistics for Authorization Limits

var_name	Wire amount	Approver Authorization limit	Initiator authorization limit	Settler Authorization limit
N	229,531	8,239	8,239	8,239
Average	4,793,957	167,685,975	80,232,688	606,870
Median	70,242	10,000,000	0	0
Std	79,213,746	452,077,500	325,524,629	24,628,745
Min	0	0	0	0
Max	13,260,787,693	9,814,999,869	7,806,759,586	1,000,000,000

Each wire transfer belongs to one of four types: random, repetitive, concentration, and batch. A random wire requires only one payment (e.g. payment due to a car accident) while a repetitive wire requires multiple payments (e.g. pension payments). A concentration wire is initiated in the process of fund optimization. For example, if a line of business (LOB) is short of money, funds in other LOBs are transferred to it. Lastly, batch wires are a collection of transactions from the other three wire types that are grouped for practical convenience.

Frequency checks are performed for each attribute on the All_wires table for integrity checks, and no domain integrity violations were found. However, missing values were found for 12 attributes. The table 14 shows the attributes and their number of null values.

www.manaraa.com

Table 14. Frequency by Variable

var_name	Total_cnt	Valid_cnt	Null_cnt
APPROVER1ID	229,531	229,444	87
APPROVER1LOB	229,531	188,709	40,822
APPROVER1NAME	229,531	229,444	87
APPROVER2ID	229,531	59,828	169,703
APPROVER2LOB	229,531	50,784	178,747
APPROVER2NAME	229,531	59,828	169,703
COUNTBIZUNIT	229,531	83,313	146,218
DATASOURCE	229,531	61,173	168,358
INITIATORLOB	229,531	185,061	44,470
REPREF	229,531	169,697	59,834
ROUTINGNUM	229,531	228,055	1,476
TRANREF	229,531	227,141	2,390

Some variable can have null values. For example, the Datasource variable is null if a wire is either random or repetitive. However, the missing value for the RountingNum that records the receiver's bank account is clearly anomalous. Since it is not possible to determine the cause of the null values, the MissingRNo field is added to the All_wires table. The field has 'N' if the record has the routing number and 'Y' otherwise.

Model Development Process

Overall model development consists of five stages based on data mining methods. At

www.manaraa.com

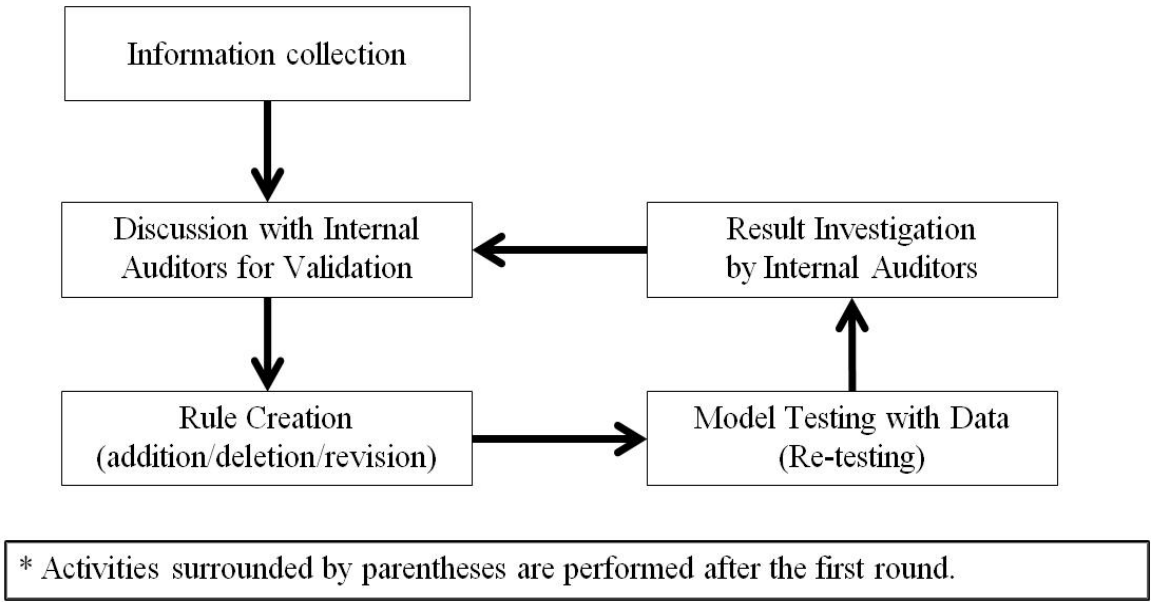
the first stage, information related to anomaly detection is collected from data files and internal auditor analysis. Next, initial brainstorming is performed with internal auditors based on their analysis and pilot tests on the collected data. This process determines potentially risky areas that require anomaly indicators/rules. Anomaly indicators are then created based on the results of the brainstorming session. This process requires the most cognitive effort and time-consuming labor. Since known anomaly instances are rare, a newly generated anomaly indicator may have an unexpected outcome. For example, an anomaly indicator tests whether an initiator initiates a wire transfer whose amount is unusually different from what the employee usually initiated. This indicator assumes that wires initiated by an employee have a narrow range and relatively large so that a fraudulent wire has a smaller amount than usual. However, this indicator does not work at all because wire amounts initiated by an employee have such a wide range that it is impossible to capture unusually small amount.

The fourth step is model testing. After anomaly indicators are generated, the model is tested with wire transfer data. Each indicator has a different weight based on its likelihood of anomaly indication. Once a particular wire transfer payment is passed

www.manaraa.com

through different indicators and scored, an aggregate total is calculated and a wire transfer payment above a given threshold is suggested for investigation. Finally, flagged observations are verified by the internal auditors and the verification results are used to update the model for fine-tuning. This process is reiterated until a satisfactory model is derived. The notable feature of development process is that it is iterative and interactive. The overall process is shown in the figure 8.

Figure 8. Development Process of Anomaly Detection Model



The initial phase of the study involves obtaining a general understanding of the company's wire transfer payment system and the corresponding data. This step facilitates

www.manaraa.com

the creation of indicators and algorithms to supplement and support the controls in place. Data characteristics and layouts are obtained, along with descriptive statistics, giving a quantitative understanding of the data and the types of wire transfer payments being made. Next the research team and the internal IT audit team brainstorm ideas for indicators that may illuminate anomaly/outlier transactions based on the notion that the anomalies/outliers produced may be meaningful as fraud indicators. These indicators are transformed into statistical algorithms utilizing data mining techniques.

The indicators consist of three types of statistics; prediction, correlation, and frequency test. Using these types of statistics on the data allows the determination of anomalies or patterns. Each indicator is scored based on anomaly risk: a score of one for low risk, three for moderate, and five for high. Scores are based on the professional judgment of the internal auditors. After running wire transfers through the indicators, suspicion scores are aggregated to determine what total score should be used as the cutoff/threshold for further investigation by the internal audit team. Upon completion of the investigation, the internal auditors verify whether the flagged transactions are fraudulent.

In addition, internal auditors suggest how to improve the model and the indicators.

www.manaraa.com

The model should constantly evolve to adapt to new findings. Since fraud is persistent in nature, the fraud detection/prevention process should be continuously run and updated. The target tests performed by the company are not discussed in this study to prevent harming the insurance company's fraud detection efforts.

Screening rules

The following areas were initially suggested by the audit team for investigation: 1) whether the payee transactions payment amount is out of the range of payment amounts, 2) whether the payee transaction payment trend line over time has a positive slope, 3) whether a sender sends a wire payment to an unusual payee, 4) whether the initiator/approver transaction payment amount is out of the range of baseline payment amounts, 5) whether the transaction amount is out of range of normal activity from this bank account, 6) whether the transaction initiator is not a normal sender from this bank account, 7) whether the transaction payee is not a normal receiver from this bank account, and 8) whether a bank account is associated with many other types of transactions.

www.manaraa.com

Table 15. Examples of Risky Areas and their Testing

Potential fraud indicators	Possible screening rules to test
The payment amount to a payee is abnormally large or small.	Amount range for each payee (or all payees) & check outliers.
The payee transaction payment trend line over time has a positive slope.	Correlation between date (or sequence numbers) and payee amounts for each payee
The payee is an outlier to payee baseline activity. (Payment sent to a payee that normally does not receive payments)	Payee frequency by each initiator & check the payees that have the lowest frequencies.
The initiator / approver transaction payment amount is out of the range of baseline payment amounts.	First, check the transaction amts with their authorization amts. Second, calculate 90, 95, or 99PI. And then find the transactions that are beyond these bounds.
The transaction amount is out of range of normal activity from this bank account.	The 90, 95, and 99 PI amts for each sending/receiving bank account and check the exceptions.
The transaction initiator is not a normal sender from this bank account.	First, check the list of sender bank account, then create exception lists of initiators by originating bank account.
The transaction payee is not a normal receiver from this bank account.	A list of payees by sending banks who have least frequency.
Access to the bank account is commingled with many other types of transactions.	A list of bank accounts with wire types that have the least frequency.

The initial tests to tackle the risky areas are summarized in the table 15. The first fraud indicator is about the amount anomaly for each payee. Due to domain integrity issues, we cannot rely on payee names to uniquely identify each payee. To make it worse, because of the lack of a master file that with payee ID and bank account information, we must assume that each payee uses only one bank account for wire receipt.

www.manaraa.com

Payees are categorized into four types considering statistical and interpretational constraints: P1 if the number of total wires=1; P2 if 2; P3 if between 3 and 29; and P4 if 30 or more. About 90% wire transfers belong to 10% of payees. In other words, most payees were involved in very few transactions. More precisely, 62.82% of the payees were involved in only one transaction while 93.84% of the payees were involved in fewer than 30 transactions. Out of 13,145 payees, only 800 received over 30 wire transfers.

Indicators

Anomaly indicators are categorized into two groups based on their analysis approach: target and trend tests. A target test is a pass/fail indicator. Some examples are tests examining whether an employee approves a wire transfer beyond his/her authorization limit, whether a payee exists on a payee master file, and whether a wire is sent to countries known as financial safe harbors. The target tests in the table 16 performed by the company are included in the model:

www.manaraa.com

Table 16. Target Tests

Description: Target tests
T1: Payee does not receive payments from more than one initiator.
T2: Payee does not receive payments from more than one approver.
T3: Initiated transaction date is after the initiator's termination date or before hire date.
T4: Approved transaction date is after the approver's termination date or before hire date.
T5: Approved transaction date is after the approver's termination date or before hire date.
T6: Receiver is located in a country known as financial safe harbor.
T7: Multiple payments on the same day in the aggregate exceeds approvers limit for payment to a single payee.

Trend tests utilize statistical methods such as prediction interval, correlation test, and frequency test. Measures for these tests are mostly in continuum so that thresholds must be set to determine sufficient risky. The twelve anomaly indicators in the table 17 are included in the model and performed at Phase 1 in this study.

www.manaraa.com

Table 17. Trend Tests

Description: Trend tests
A: A payment is out of a payee's normal range.
B: A payee's payments increase over time.
C: A payment is usual according to payee normal activity.
C1: New initiator
C2: New approver
C3: Potential collusion
D: A payment amount is unusually different from normal activities.
D1: Initiator
D2: Approver
E: A payment amount is unusual for a sending bank account.
F: A transaction initiator/approver is not a normal sender from the sender's bank account.
F1: Initiator
F2: Approver
G: A payee is not a normal receiver for a bank account
H: Access to the bank account is commingled with all types of transactions.

Prediction Interval Test

A prediction interval test stratifies payees, initiators, and approvers based on number of payments received: 1) one, 2) two, 3) three to twenty-nine, and 4) thirty or more. Alternative alpha prediction intervals of 90%, 95%, 99% are also considered. A higher alpha level will have fewer outliers and, conversely, a lower alpha level will result in more outliers. For payees with only one wire payment, a prediction interval is estimated by grouping the payee's wire payment together with other payees who have one wire payment in order to determine which payments are abnormal compared with the group as

www.manaraa.com

a whole. A prediction interval is applied to payees with thirty or more wires. For payees who have only two wire payments and for payees with three to twenty nine payments, statistical methods such as clustering may be useful to detect outliers.

Correlation Test

The correlation test examines how payment amounts change in a manner inconsistent with a payee's other transaction patterns. Activity monitoring (Fawcett and Provost, 1999) is adopted for this type of test. It requires the maintenance of a usage profile for each payee or employee in order to determine any deviation in activity. In contrast to the prediction interval test, correlation calculation requires three observations at minimum. Transfers are therefore stratified into two groups: those with three or more wires, and those without. The degree of overall increase of wire amounts are determined by the correlation value and its p-value for statistical significance. In the literature, various correlation values are suggested to decide whether observations are positively correlated. Although global standards do not exist for strength of correlation, a coefficient between 0.3 and 0.7 is generally considered moderately correlation. Our study uses a 0.5 threshold.

www.manaraa.com

Frequency Test

After defining normal activity patterns, anomalies can be determined. Frequency tests can help define a typical or normal activity pattern, with infrequent activities indicating potential error or fraud. A frequency test indicator entails examining each payee/employee pair initiating wire payments to determine which pairs have unusual activity.

For example, unusual activity can be a payee interacting with an uncommon employee or group of employees for the first time. A pilot test in this study shows that a payee typically encounters many different initiators and approvers in the company. It follows that encounters with only the same initiator or approver may not be considered normal.

Scoring System

Scoring of indicators is developed with assistance from internal audit. The knowledge engineering of experienced professionals (Vasarhelyi and Halper, 1991), such as an effective internal audit team (AICPA 2002), allows for the determination of indicators to be considered abnormal or potentially fraudulent in nature. In this study, each indicator is

www.manaraa.com

assigned a score based on perceived risk. However, it is not an easy task to measure the relative importance of anomaly indicators especially when their effects do not seem to distinctively different. Weight assignment becomes even more challenging as the number of anomaly indicators increases with the dynamic nature of model development.

After each indicator is processed through statistical algorithms, violation totals for each wire transfer are computed. A wire transfer that violates more than a determined threshold is subject to investigation by the internal auditors. In running the initial algorithms, an enormous number of exceptions are found. In practice, cost barriers prevent internal auditors from spending much time on investigation. Kogan et al. (1999) discuss the cognitive effect of information overload. An overload of alarms will have a negative effect on the internal auditors to adopt an anomaly detection system because of limited human resource.

It is therefore necessary to increase threshold scores in order to reduce the number of flagged wire transfers. The summary statistics of the aggregated scores are illustrated in the table 18.

www.manaraa.com

Table 18. Thresholds by Category Type

score_trend	cnt_wires
0	183534
1	25933
2	3141
3	5179
4	1271
5	266
6	1005
7	8217
8	707
9	209
10	58
11	9
12	2

Score_target	cnt_wires
0	195948
5	18841
10	14401
15	334
25	7

Score_total	cnt_wires	Score_total	cnt_wires
0	163304	13	67
1	23204	14	18
2	2437	15	139
3	3562	16	187
4	964	17	7106
5	14962	18	538
6	2652	19	136
7	1490	20	41
8	1717	21	7
9	361	22	27
10	5406	23	2
11	897	24	1
12	299	25	6
		32	1

The thresholds assigned to determine the flagged wire transfers are 10 for trend tests, 25 for target tests, and 20 for total tests. These thresholds produce 106 flagged wires, which the internal audit team found high. Increasing thresholds reduce the number of exceptions is reduced to 47: 11 for trend tests, 25 for target tests, and 22 for total tests. The audit team found this figure more reasonable. The table 19 shows some of flagged wires with fictitious numbers.

www.manaraa.com

Table 19. Examples of Flagged Wire Transfers

ID	Score_trend	Score_target	Score_total	Amount
3259892	11	10	21	989,343,618.35
3278185	0	25	25	150,000.00
3296478	7	15	22	4,473.33
3314771	11	10	21	135,350,000.00
3333064	7	15	22	25,657.09
3351357	7	15	22	53,077,500.00
3369650	11	10	21	1,235,418.75

Results and Discussion

Results

The internal audit team investigated the 47 wire transfer payments during their regular audit and found no evidence of fraud or error. The investigation shows that most wires are flagged when they are the only payment to a payee because a wire transfer violates two target tests and three trend tests when a payee has only one payment. In addition, the wires beyond target test thresholds are intercompany transactions. For reasons the internal audit team cannot identify, those wires violate five target tests. The system must reduce the effect of a single violation on overall weighting score. Although the anomaly detection model does not find anomalous wire transfers, this does not mean that there are no anomalous payments. Instead, this may indicate the need for revision

www.manaraa.com

and fine-tuning of the model. The company intends to include the fraud detection process as a part of regular audit, retaining these indicators for future detection or preventive measures. In addition, the company is interested in refining the indicators and adding new ones to screen for anomalies.

Internal Control Issues

During the study, the effectiveness of the company's internal control comes into question. Three issues emerge: 1) segregation of duties controls are violated, 2) terminated employees remain able to process payments, and 3) wire payment limits are circumvented; even employees $0 limits are able to process wire payments. These major internal control issues are brought to the attention of the internal audit department and are investigated. The internal auditors find that there are inconsistencies between the wire transfer payment process records and human resource records. The discrepancy is caused by the company keeping only the most recent information. For example, a terminated employee may have been an active employee when he/she initiated or approved a wire transfer. Although these internal control violations are potential fraud indicators, an investigation of their nature and frequency suggests that they are due to poor database

www.manaraa.com

management. However, this does not exclude the possibility of fraudulent activity.

Scoring System Issues

During the investigation, the internal audit department noted that some wires were flagged because of systematic causes that were mainly attributed to the target tests. Each of the target tests' indicators was scored at the highest risk level and as a result, flags raised by a few of these indicators will most likely hit the threshold for investigation. This suggests that an equally-weighted scoring system may be more useful as a starting point. However, some indicators are clearly more important than others. As long as the indicators are subjectively assigned weights, this issue may recur. Further deliberation will be necessary to find less subjective weighing methods for the indicators. The finding further illustrates that any fraud system must be continuously developed and updated as new flaws with the current system surface.

Discussion

This study provides a pilot test for anomaly detection at a major US insurance company. Although the literature has discussed numerous methods for fraud detection, few have used actual company data. Data mining is used as the approach to detect

www.manaraa.com

anomalous wire transfers, with statistical algorithms created to detect data abnormalities. Since much of the prior research uses complex methods such as neural networks and clustering to detect anomalous transactions, the use of simple statistics such as prediction interval, frequency test, and correlation test may seem trivial. However, simple methods can be as robust and powerful as, and sometimes more accurate than, more sophisticated methods. This study does not consider analysis for payees with between 2 and 29 wire payments due to lack of statistical significance. Future research can investigate other types of statistical methods such as clustering for detecting abnormal or patterned activity. The company plans to further pursue data mining in a continuous effort to detect fraud.

This is a learning experience for academics as it shows how an anomaly detection and prevention model is implemented. In addition, this study demonstrates that internal auditors can run anomaly detection and prevention activities on a frequent basis instead of during an annual audit. Two issues require further consideration: 1) the highly subjective weighting of most indicators yields extreme numbers of violations, and 2) as the pilot study progresses, some indicators need to be adjusted because of an increasing understanding of data characteristics.

www.manaraa.com

c. Phase II (July 2009)

Data

After Phase I, the insurance company decided to include the anomaly detection model in a regular audit. The datasets at Phase II provided by the audit team, however, exhibit several differences from their predecessors. The most distinctive difference is a major DBMS update made since the previous audit that includes changes to data structure, format and location. As a result, the new dataset has fewer transactions over a longer time period with an additional 5 variables: wire status, bank confirmation code, OFAC (Office of Foreign Assets Control) status, payee address, and a comment (similar to an existing reference variable). With the five additional variables, the total number of variables becomes thirty two. Inevitably, a new anomaly detection model must be developed with a different data structure. Since a majority of the variables are not affected by this change, the model at the Phase II is used as a starting point.

As in Phase I, the dataset in this study is wire transfer payments made by the insurance company. The data spans eighteen months (i.e. January 2008 to June 2009) and consists of 201,476 wire payments-fewer than those in Phase I. Approximately 90% of

www.manaraa.com

the wire transfer payments belong to about 26.5% of the payees, compared with 10.25% in Phase I, implying that major payees had fewer transactions during this period. 57.82% of the payees are engaged in only one transaction (down from 62.82% in the Phase I) and 92.59% of the payees have less than 30 transactions (down from 93.84% in the Phase I).

The All_Wires table with 32 variables has 201,476 transactions, excluding 28 irrelevant records. The other six tables have the same data structure as in Phase I.

Model Development Process

The anomaly detection model in Phase II is based on the five-stage model developed during Phase I. Since this study follows the previous one directly, the information collection is not necessary, is not necessary. Instead, the second stage begins with feedback and discussion with the audit team from Phase I. The main outcomes of the Phase II are 1) revision of indicator weights for better scoring, 2) addition of new indicators, 3) consideration of materiality, 4) input to a current quarterly audit, 5) reclassification of tests.

First, as found in Phase I, some target tests are over-weighted, and effects of other

www.manaraa.com

indicators are relatively ignored. Although the use of a relative weighting system is desirable because of differentiated importance, it is near impossible to achieve consensus. Quality cannot easily be measured in discrete quantity. As long as the degree of risk is qualitative, resulting numerical weights are not absolute. These weights are subject to change whenever new indicators are added. In order to minimize the number of false alarms due to weighting systems, several indicators change their weights with conditions. As a result, three target tests that have 5 point weights are merged into trend tests with weight changes. For example, an indicator examining whether a payee receives wires from only one initiator is merged into a broader trend indicator examining whether a payee has a new initiator. This is due to the fact that a payee usually has one initiator only when he/she is a new customer. Hence, its potential risk is regarded less risky and 3 points are assigned if violated.

Second, eighteen anomaly indicators are added after discussion with the internal audit team to validate whether anomalous behaviors tested by the indicators are truly anomalous or allowed as exceptions. New indicators examine aging, potential collusion, segregation of duty, split wire, referential integrity, process day, invalid variable value,

www.manaraa.com

proper approval, OFAC process, and similarity by clustering. New indicators are summarized in the table 20 and 21.

Table 20. New Indicators for Trend Tests

Trend tests	
I	A payment is disbursed unusual long after its initiation.
L	A payee always receive payments from only one initiator/approver.
	L1: Initiator and approver are always the same.
	L2: Initiator and approver change their role for the same payee.
M	An initiator or approver processes wire transferes for only one payee.
	M1: Initiator
	M2: Approver
N	Switches of initiator and approver
	N1: Across payees
	N2: For a payee
O	Split-wire test for initiator: Due to multiple authorization limits, existence of wire-splitting is tested instead.
P	A payment is unusually low but sufficient material. (LPL01 and >=$2,000)
	P1: Initiator
	P2: Approver
	P3: Sender
T	Clustering

www.manaraa.com

Table 21. New Indicators for Control Tests

Control tests	
J	A payment is initiated/disbursed on weekends or holidays (Saturday/Sunday/holiday).
	J1: Initiated Date
	J2: Disbursed Date
K	Violation of referential integrity: Repetative wires.
Q	A payment does not have a routing number.
R	A payment is approved although it fails the OFAC test.
S	A payement does not have approvers.

Third, wire transfers with small dollar amounts are excluded in verification process. A good detection model is one that has few false negatives and few false positives. The former determines the model's power (or accuracy) and the latter determines efficiency. Improvement of one reduces effectiveness in the other. Although an accurate model is more desirable than an efficient one, the cost of application must be considered. Modeling is resource- and/or time-consuming and a company cannot afford additional human resources that should be assigned to verify whether flagged wires are truly anomalous. When the number of flagged wires needs reduction, the most popular candidate as a discriminating criterion will be transaction amount. In a regular audit, an error of immaterial amount is practically ignored. Following this approach, an anomaly

www.manaraa.com

with small amount (less than $2,000, as in the company's regular audits) is excluded.

Fourth, outcomes of anomaly detection become part of the company's quarterly audit. As a result, feedback of verification process becomes timelier and model development becomes more practical. Since no anomalous wire transfers were found in the previous study, and those transfers were audited, all wire transfers before the current audit period will be considered free of anomaly and wire screening will be applied only to the latest quarter.

Finally, anomaly indicators are reclassified. In Phase I, target testing was performed by the internal audit team. This reduces the consistency of model development in two ways. First of all, the data and programs related to target tests were not available, eliminating the possibility of verification. Second, overlapping of two pairs of trend and target tests hindered development efficiency. Making things worse, outcomes of those tests were somewhat different although they tested the same objects. Taken together, inconsistency was found among the indicators in the Phase I. To eliminate this inconsistency, Phase II indicators are classified into trend and control tests while, while data availability issues forced the exclusion of three target tests. Control tests are binary,

www.manaraa.com

while trend tests are continuous. This reclassification helps to build a framework of detection model development. This phase features fifteen trend tests and five control tests. The fifteen trend tests consist of twenty four indicators while the five control tests have six indicators.

Screening rules

Phase I results and feedback are used as a starting point for Phase II. Wire verification results from the audit team and discussion of newly available variables suggest the following new high-risk areas: 1) whether a wire transfer process takes an unusually long time (longer than 20 days), 2) whether a payee has an unusual connection with an initiator or an approver (e.g. a payee receives wires from only one initiator or approver), 3) whether a payee, an initiator, and approver share a close connection (i.e. a payee receives wires from only one pair of an initiator and an approver), 4) whether an initiator and an approver have a close connection (i.e. possible collusion) by role switching, 5) whether a wire transfer beyond authorization limits is initiated or approved after splitting, 6) whether a wire with unusually small amount is processed, 7) whether a wire is

www.manaraa.com

processed on non-working days (weekends and holidays), 8) whether a repeating wire transfer does not have a record on its master file, 9) whether a wire does not have a payee's routing number that identifies his/her destination account, 10) whether a wire against OFAC (Office of Foreign Assets Control) status check is processed, and 11) whether a wire is processed without proper approval.

Indicators

Similar to the target tests from Phase I, a control test is a binary indicator that examines existence of any violations against the company's operational control policies. For example, a wire transfer violating OFAC policy cannot be approved even if it is initiated. Hence, a wire transfer is anomalous if it fails the OFAC test but is approved. Another example is a wire transfer processed on non-working days such as weekends and holidays. Since the company is only open on non-holiday weekdays, a wire transfer on a non-working day is anomalous. Non-working days are also vulnerable to internal fraud because a wire transfer can be performed without being monitored by other employees. Twelve trend tests are added in Phase II. One exception is a split-wire test. Although it

www.manaraa.com

should belong to a control test, it is classified as a trend test because it is not an exact control test. Because of either incorrect data extraction or improper database management system, multiple authorization limits are assigned to some employees. Since its cause is undetected by both the internal audit team and the IT team, conservatively, smallest authorization limits are applied to this test. Hence, some of wire transfers that are considered to violate the split-wire test may not be against it. Twenty four indicators are trend tests under fifteen types.

Prediction Interval Test

All of the prediction interval anomaly indicators from Phase I are included in Phase II, and one new indicator is added to examine whether a wire transfer has an unusually low amount. This indicator purports to detect an internal fraud that embezzles a small amount of money.

Correlation Test

There are no changes in this type of anomaly indicator.

www.manaraa.com

Frequency Test

This type of anomaly indicators is extensively used in Phase II. An anomalous activity is a rare event, making frequency testing appropriate for anomaly detection. In addition to the seven indicators of the four types in Phase I, eight indicators of five types are implemented in Phase II.

For example, one test examines whether an initiator and an approver switch their roles for a payee. Let us assume that John and Jack are employees of the company and Jane is a payee. John initiates a wire transfer to Jane and Jack approves it. Later, Jack initiates another wire transfer to Jane and John approves it. This example clearly violates segregation of duty and shows a possible collusion risk. While such an event may occur by mistake (if the system allows it), it is still an unsafe practice.

Scoring System

During Phase I, the weighting system produced several unexpected outcomes. For example, five points were assigned to each target test in Phase I. Relatively high weights on target tests resulted in excessive false positives. In addition, overemphasis on target tests diminished the importance of trend tests, minimizing their effects on anomaly

www.manaraa.com

detection.

This unexpected outcome resulted from subjectively weighted anomaly indicators. Target tests were over-weighted although their relative importance was explicitly distinguishable. It is well known that quantification of anomaly risk level is difficult because risk is inherently qualitative. One example from Phase I is a target test examining whether a payee has only one approver. Although the test assumed that it would be risky for a payee's transactions to be processed by only one approver, it instantly flagged every transaction involving a new payee. The weight on the indicator must be lowered or revised.

To mitigate this problem, four of the seven target tests are converted into trend tests with weight changes and the remaining three are discarded due to data deficiency. For example, the indicator above is divided into two cases. If a payee has only one transaction (i.e. a new customer), one point is assigned, considering that his/her payment pattern is not well established. If the payee has more than one transaction, then three points are assigned. Newly added trends tests are also weighted at one point, with the exception of a three-point weighted approver/initiator role-switching test. The reason to assign an equal, low weight to new indicators is that their risk levels are unknown and

www.manaraa.com

difficult to measure. When information related to inherent risks becomes known, weights will be changed accordingly. Although changing weights may not completely address the issue weighting must still be approached with care.

Limited human resources hinder the verification process as in Phase I. The number of wire transfers that the audit team investigates is about 30. With this practical restriction, the thresholds for two tests are carefully chosen to produce the maximum number of flagged wires. As a result, chosen thresholds are 9 or higher for trend tests, 3 or higher for control tests if they have routing numbers and approvers, and 1 or higher if they do not have routing numbers or approvers. Since more transactions with no approver and routing number occurred than were expected, they are separately reported to the audit team for further investigation. 214 wires do not have routing numbers and 215 wires do not have approvers. In this phase, a threshold for total score is not used because the number of control tests is relatively small (6) compared with the trend tests, and few wires violate these tests. Hence, use of total score for screening results in almost identical outcome as screening with the trend score that has either much fewer or more than 30 wires. After applying screening thresholds, 26 wire transfers are selected and sent to the audit team for verification.

www.manaraa.com

Results and Discussion

Results

The internal audit team investigates the 23 wire transfer payments as part of their regular audit. Although 26 wire transfers are recommended for verification, 3 wire transfers are discarded as immaterial. Two wires have $0.01 and one has $1,572.78. The audit team considers these amounts as negligibly small.

After examining the 23 wire transfers, the audit team finds no evidence that any are fraudulent or erroneous. The investigation result shows that the wire transfers have three features. First, some of wires are sent to tax authorities and their payees do not have a personal interest in them. Second, wires processed on non-working days are generally time-sensitive. Although employees are not supposed to work on those days, they do have exceptional circumstances. Lastly, some of flagged wires are sent to the company's subsidiaries. The audit team argues that the money is traceable since it is still inside the company. As long as the money stays in the company, the auditors feel that it does not bear any fraud risk. Despite this claim, it cannot be determined whether all internal accounts are under control. Although management fraud can be related to these

www.manaraa.com

transactions, that possibility is beyond the scope of this study. If we assume that internal transfers are free of internal fraud, it may be possible to exclude the concentration wires (internal wire transfers to optimize the company's fund usage). This exclusion will be applied to the next phase.

To summarize, the audit team does not find any evidence that supports existence of anomaly. However, this does not mean that all transfers in the quarter are free of anomaly. Instead, it may imply that the current detection model is not powerful enough to detect anomalous wire transfers or that the indicator weights are not properly measured. It is evident that the detection model still has room for improvement. Emerging issues are summarized in the following sections.

No approvers

A wire transfer consists of three steps: initiation, approval, and disbursement. Each wire transfer must have one initiator and at least one approver. Approval requires either an approver ID if manually processed or a preset value for the ID if automatically processed. According to this policy, the approver ID field must have a value and cannot be empty. However, some wire transfers have null approver ID values. After investigating

www.manaraa.com

the missing entries, the IT team claims that those wires are from the pension system and their approvals exist in the administrative system but are not properly carried over to the wire transfer system. However, the IT team fails to explain why approvals do not appear in the wire transfer system. This might be because their DBMS is not seamlessly managed or because there is a mistake during data extraction process. Regardless, this indicator may need to remain until its cause is more clearly identified and resolved.

Multiple authorization limits

In the company, an employee can be assigned to more than one line of business (LOB), although many LOBs may have multiple wire types in common. Authorization limits are related not only to employee rank but also to a LOB that an employee belongs to. Consequently, it is possible for an employee to have multiple authorization limits for the same type of wire transfers. This is true for 203 out of 418 initiators. In the audit team's opinion, an employee should have only one authorization limit for each wire type. This may imply that an employee's authorization is loosely controlled and should be tightened. Until the problem is resolved, the authorization limit check will use the lowest number.

www.manaraa.com

Referential integrity

Referential integrity implies that a record of a mater table referenced by a record of other tables must exist. This is a fundamental concept in a relational database system. If a table in a relational database system violates the referential integrity, the resulting error can destroy the entire database in the worst case. In this study, the All_Wires table references other tables for details. Among those tables, Templates is a master file that is referenced by repetitive wires of the All_wires table. Consequently, if a repetitive wire record exists on the All_Wires table, it must exist on the Templates table. However, there are cases (e.g. repetitive wire #5549) where repetitive wires do not appear on the Templates table although they belong to the All_wires table. Since this violation can damage the entire database system, its cause must be investigated rapidly. Although improper data extraction and table matching are possible causes, investigation by the IT team shows that the true reason is the least expected. According to the IT team, the All_wires table violates a referential integrity since referenced records on the Templates table are deleted either manually or automatically although they are referenced by the All_wires table. In a relational database system, a record on a master file must be deleted

www.manaraa.com

only after all referencing records are deleted to safeguard the database. In other words, the company does not strictly enforce referential integrity, which may deteriorate the database in the future. The IT team has taken this problem into consideration to prevent possible disasters.

Missing Routing Numbers

A payee can be identified by name, address, or bank account. To distinguish an individual payee, a unique identifier (key) is necessary. Since multiple payees can share the same name, that cannot be used as a key. Although a payee address can be a candidate key, it is prone to manual entry error. For example, a payee address "161 Washington Street, Newark, New Jersey" can be recorded as "161 Washington St." and/or "161 Washing Street, Newark". Although they seem to indicate the same payee, it is difficult to determine that they are actually the same. Hence, the most appropriate candidate key is a payee's bank account, consisting of a routing number and an account number. Since the two components are recorded separately, it is important that two variables have valid values. The audit team confirms that these values exist in order to identify the destination of a wire transfer. However, when routing number has a null value, further information is

www.manaraa.com

necessary to verify that the fund is sent to the intended recipient. After investigating the cause of missing routing numbers, the IT team argues that there is a mapping system that is not currently available and it is possible to relate a certain wire transfer to its destined bank account. However, they fail to explain why only some wire transfers have this problem and to provide the mapping table. This phenomenon should be considered as an anomaly until a mapping table is provided and found correct.

Discussion

Following Phase I, this study presents a development process and the results of the second generation model for anomaly detection. Major features of Phase II are summarized as follows. First, more anomaly indicators are added based on feedback from Phase I and discussion of new findings. Eighteen indicators are added, including four conversions of Phase I target tests. Second, weight revision is made to the indicators that caused unexpected outcomes in the Phase I. While converting target tests into trend and control tests, weights are changed proportionally with importance. This revision is a result of deliberation on the potential effects of anomaly risk. Third, several problems are found and discussed during model development and testing, including missing routing

www.manaraa.com

numbers, missing approvals, referential integrity violation, and multiple authorization limits.

Although some flagged wire transfers are highly suspicious, investigation by the audit team does not find any evidence of anomaly. Some newly raised issues merit further investigation. For example, a clear answer to the missing routing number problem will be available only when a mapping table is provided.

This study provides a learning experience about iterative anomaly development processes. Since Phase II starts from the second step of the model development process, feedback from Phase I facilitates model development by providing a direction and more details. Knowledge about anomaly detection will accumulate as the process continues. The next round will involve consideration of newly found problems, revise the factors that need fine-tuning, and modify the raw dataset to narrow down the scope of wire transfers.

www.manaraa.com

d. Phase III (October 2009)

Data

The models at Phases I and II must be improved for better detection power, even though they identified many problems during model development and testing. After components of wire transfers are examined in detail and discussed, anomaly indicator recategorization is suggested as an improvement.

Categorization is important in the evaluation and development of anomalous indicators and screening of anomalous wire transfers. Indicator categorization uncovers overlooked risk areas and screens wires by anomaly characteristics. Although categorization is useful and important, it is not systematically developed until Phase II. It is reasonable to arrange anomaly indicators in a logical way to improve the quality of anomaly detection model, though this requires that screening methods be changed.

The dataset in Phase III is an undated expansion of Phase II. Wire transfers from July, August, and September in 2009 are added to Phase II and master files such as employee records are updated to apply changes during the period. With these changes, the data spans twenty-one months (January 2008 to September 2009), consisting of 260,762 wire

www.manaraa.com

payments. After excluding summary observations (40) and rejected wires (1,239), 259,483 wire transfers remain in the dataset, with approximately 90% belonging to 15% of the payees.

Model Development Process

Similar to Phase II, Phase III begins with feedback from the prior phase and discussion with the audit team. The main outcomes of Phase III are 1) revision of indicator categorization for more reasonable classification, 2) addition of new indicators, and 3) comparison of flagged wires based on old and new categories.

First, anomaly indicator categorization is reconstructed in Phase III. Although minor changes are made during Phase II, categorization still needs more systematization. New categories segregate anomaly indicators based rigidity of discrimination. Binary indicators have more rigid thresholds than statistical types because thresholds do not need to be chosen. A weekend authorization check will return either yes or no, with no externally chosen parameters necessary. Binary indicators test more directly than statistical indicators. In addition, new categorization assigns equal weights to anomaly

www.manaraa.com

indicators. This facilitates anomaly indicator development and weight assignment. As discussed in Phase II, weights assigned to individual anomaly indicators are subject to change as their effects on anomaly become better understood.

Second, new anomaly indicators are developed based on feedback from Phase II and frequent discussion with the audit team on various hypotheses stemming from data analysis. One example is an expansion of split-wire testing. In Phase II, the split-wire test was performed in terms of initiators. Since wire splitting is one of the most common methods for internal fraud, the test is now expanded to include approvers.

Lastly, new indicator categorization is examined by comparing flagged wire transfers by existing category with those by new one. As a matter of fact, most flagged wire transfers are in common for both categorizations with some differences. One concern about newly suggested category is how to weigh each category. Although yes/no type of anomaly indicators seem to have more significant effects than statistical ones, it is difficult to determine how different they are. The Scoring System section discusses this issue in detail.

www.manaraa.com

Screening rules

The anomaly detection model in Phase III starts with anomaly indicators and feedback from Phase II. In addition to the anomaly indicators from Phase II, two types of anomaly indicators are expanded and one type is newly added while one indicator is dropped. First, the split-wire test is expanded to consider more various cases. While only initiators are examined in Phase II, Phase III also considers approvers and examines them by wire type. Second, clustering indicators separate anomalous wire transfers by three hierarchical clustering methods: flexible beta clustering, two-stage density linkage, and Ward. For clustering, anomalous wire transfers are defined as the observations in the smallest clusters. Since the smallest cluster will change depending on clustering method, results from clustering may be less evidential. Third, an indicator examining segregation of duty is newly added in the Phase III. Segregation of duty is one of the most essential components of internal control. Its targets are between an initiator and a primary approver and between a primary approver and a secondary approver. Lastly, an indicator that identifies wire transfers without approvers is dropped due to insufficient information

www.manaraa.com

regarding acceptable values.

Indicators

Indicator categorization purports to facilitate the development and management of anomaly indicators. Although some indicator categorization changes are made in Phase II, the new categorization does not seem completely systematic. Since a systematic framework can identify risky areas and manage existing anomaly indicators, it is crucial to have well-developed anomaly indicator categories. Another benefit of indicator categorization is that a scoring system can be easily controlled. If indicators are categorized by relative risk, it is easier to interpret a wire transfer's score.

Anomaly indicators are divided into trend and control tests in Phase II. Those category names are, however, somehow misleading because not all trend tests mean 'trends'. In order to avoid this confusion, new category discriminate anomaly indicators into statistical and conditional tests. As their names imply, statistical tests utilize statistical methods such as prediction intervals, correlations, and clustering while conditional tests use frequencies or yes/no questions to screen wire transfers. Since the

www.manaraa.com

frequency (i.e. the number of cases) is always one to be strictly conservative, it is the same as yes/no questions in nature. While statistical tests determine various parameters (e.g. alpha level, significance level, etc) to identify anomalies, conditional tests do not need such a procedure. Although statistical parameters are based on common practice, their effects on the degree of anomaly are unknown. The need for human intervention in statistical tests makes those tests more discretionary. We can therefore assume that conditional tests are more powerful than statistical ones.

Since little is known about new categorization, this phase categorizes anomaly indicators using both previous and new categorizations and compares their outcomes. Categorization by previous framework is the same as the Phase II, except that split-wire and segregation of duty become re-categorized as control tests. Categorization by the new scheme is much simpler. Binary and frequency tests are labeled as conditional, and all others are statistical.

Split-wire tests are developed with yes/no type questions. If there are any wire splits to initiate or approve a wire transfer that are beyond an employee's authorization limit, one point is assigned. Similarly, the segregation of duty test assigns one point for a violation.

www.manaraa.com

Scoring System

The anomaly scoring system is highly dependent on indicator categorization. When wire transfers are screened, thresholds are applied to the sum of each category and the total suspicion score. In addition, different weights are assigned in the previous categorization (trend and control tests), while equal weights are assigned in the new framework (statistical and conditional tests). Weighting methodology is critical. Since the relative risk of each individual indicator is difficult to measure, grouping indicators based on risk level may be helpful. After grouping, the weighting issue remains. This process is not an easy task but less difficult than assigning weights to individual variables. It may be more suitable to handle indicators by applying separate thresholds to each category as in Phases I and II.

The number of flagged wire transfers has an upper bound of thirty because of the limited human resources that the company can allocate to verification. Hence, thresholds for each score are determined to produce no more than thirty wire transfers. In the previous categorization, nineteen wire transfers are flagged after applying thresholds, 10

www.manaraa.com

for trend score or 2 for control score and 2 for total score while amounts are $2,000 or more. In the new categorization, 6 for statistical score or 5 for conditional score is used for thresholds and nineteen wire transfers are chosen for further investigation. After comparing flagged wire transfers by both methods, it is found that eleven wire transfers are selected in common. Those flagged by both categorizations have higher suspicion scores while those not in common are assigned low scores so that they are more affected by categorization. This difference can be resolved if the resource limit is loosened. As long as the number of wire transfers that can be investigated is limited, this problem will persist. However, the new categorization seems more useful. In order to compare categorization effectiveness, the union of two results is suggested to the audit team for verification. Thirty-eight wire transfers are delivered to the audit team in this quarter.

Results and Discussion

Results

The thirty eight wire transfers are investigated by the internal audit team during their quarterly audit. After verification, the audit team found no fraud or error. Although the investigation does not find any anomalous wire transfers, it shows very interesting

www.manaraa.com

features of flagged transactions and suggests a future direction. First, most flagged wires are sent to either other internal departments or the company's subsidiaries. The audit team argues that the money is traceable since it is still inside the company and therefore carries no fraud risk. Second, some wires are sent to tax authorities and their payees do not have a personal interest in them, implying a lack of fraud risk. Third, wires processed on non-working days are due to time-sensitivity and sent to other internal departments. Lastly, some of the flagged transactions are introduced as batch wires during the latest quarter. Since their history is not sufficiently long, they often violate tests of abnormal frequencies. This problem will vanish if they remain in the system for a sufficient time.

In response to the audit team's feedback, the whole process is performed again after excluding less risky wire transfers. After excluding summary observations (40), rejected wires (1,239), internal/subsidiary/IRS transfers (93,030), and newly added batch types (23,760), 142,693 wire transfers remain. The excluded batch wires are of new types introduced during the third quarter, and their exclusion avoids unnecessary alarms caused by short history. Many anomaly indicators in the model assume that wire types have existed in the system long enough to possess sufficient frequencies for every variable on the All_wires table. For example, an indicator that examines an unusual initiator will be

www.manaraa.com

likely flagged for wires with new types because they are likely to have only one initiator. These transactions will be included starting in the next quarter. Approximately 90% of the transfers belong to 17% of the payees after data cleaning and their categories, suspicion scores, and thresholds are shown in the table 22.

Table 22. Comparison: Old vs. New Categorization

Previoous

trend	wires
0	9163
1	13770
2	2543
3	851
4	535
5	109
6	90
7	133
8	34
9	22
10	6
11	1

control	wires
0	26977
1	258
2	14
4	8

total	wires
0	9006
1	13849
2	2555
3	907
4	539
5	114
6	91
7	133
8	34
9	22
10	6
11	1

New

Statistical	wires
0	25625
1	1079
2	420
3	93
4	30
5	8
6	2

Conditional	wires
0	9624
1	14694
2	2382
3	504
4	39
5	14

total	wires
0	9006
1	14222
2	3072
3	735
4	138
5	66
6	14
7	3
8	1

A summary of the results by categorization in the table 23 shows that wire transfers with high suspicion scores are not affected by new categorization. Although this new outcome is not investigated by the audit team, the result shows that a transition to new categorization carries benefits including easier screening control and better management of anomaly indicators.

www.manaraa.com

Table 23. Summary of Categorization Change

	Previous	New
Criteria	(trend>=9) OR (control>=2 and total>=4)	(Statistical>=5) OR (Conditional)>=5 OR (Total>=6)
Amount	>=2,000	>=2,000
The number of flagged wires	24	28
Comment	To flag as many wire transfers as possible up to 30. (Trend>=9) and (Total>=9) produce the same result.	To flag as many wire transfers as possible up to 30.

To conclude, the audit team does not find any evidence that supports existence of anomaly. However, this does not mean that all wire transfers in the quarter are free of anomaly. Instead, it may imply that the current detection model is not powerful enough to catch anomalous wire transfers or that the indicator weights are not properly measured. In either case, it is evident that the detection model has room for improvement.

HR records

Employee authorization limits play an important role in the wire transfer process. As found in Phase II, the company's HR file shows that multiple authorization limits can be assigned to an employee, and that it is authorization limits cannot be tracked after termination. This occurs because an employee's record is erased when he/she leaves the

www.manaraa.com

company. This is supplementary evidence that the company need to be more careful about their database management. This problem can be hazardous especially from an audit perspective. Although focusing on the most recent transactions is important, past data should not be ignored.

Discussion

Following Phase II, this study presents the development and results of the third generation anomaly detection model. Major features of Phase III are as follows. First, anomaly indicators are added, modified, and dropped based on the feedback from Phase II and discussion of new findings. As a result, the current model has 15 types of trend tests (with 26 indicators) and 5 types of control tests (with 11 indicators). With the new categorization, the anomaly detection model has 6 types of statistical tests (with 11 indicators) and 14 types of conditional tests (with 26 indicators). Second, weight revision is made to the indicators if new categorization is used. Instead of controlling weights on individual indicators, each category is treated as a group, facilitating a scoring system. Third, additional problems about the employee record file are found and discussed during model development and testing. Due to less than appropriate database management, it is

www.manaraa.com

not possible to track authorization limits if an employee leaves the company.

Despite high suspicion scores, no flagged wire transfers are found to be anomalous. The newly raised issues and feedback from the regular audit merit further investigation during the next phase.

This study provides a learning experience about the anomaly detection development process. As knowledge about the payment system accumulates, the model development process becomes more systematic, but there remains room for improvement.

e. Phase IV (January 2010)

Data

A series of anomaly detection model developments accumulate knowledge about the wire payment system, and the next phase is enhanced by adopting new findings. As development progresses, this will produce better and more accurate models. The Phase IV model begins with the Phase II model and feedback from last quarter's audit. After components of wire transfers are examined in detail, the following ideas appear to improve the anomaly detection model.

www.manaraa.com

First, the model may perform better if less risky wire transfers are excluded. The raw data include various types of wire transfers at varying risk levels. Since detection of internal fraud, not error, is the goal, a focus on riskier wire transfers can improve the power of the model. The number of wire transfers can be reduced in two ways. One ways is to exclude wire transfers destined either within the company or to a subsidiary. Since internal wire transfers do not leave the company, they can be seen as less risky in terms of internal fraud. This can help to focus on more relevant wire transfers and thus improve power of the model. Another source of irrelevance is a rejected wire. An initiated wire transfer is not always approved. Although uncommon, wire rejections do occur, and since rejected wires do not bring future cash outflow, it is reasonable to exclude them. Lastly, the IRS is a payee that is almost free of anomaly risk since wire transfers to the IRS are for tax purposes, and those payments may be excluded from further investigation.

Second, an anomaly detection model will perform better if error-prone wire transfers are excluded in a data cleaning process. As found in the previous quarter, newly introduced batch transactions violate many anomaly tests because of their short history. These false alarms deteriorate the model's detection power. To reduce false alarms and flag more relevant wire transfers, batch wire transfers that are commenced during the

www.manaraa.com

latest quarter are ignored for the current quarter. These wire transfers will be included in the next quarter.

The Phase IV dataset consists of the Phase III dataset and the fourth quarter of 2009. It spans over two years (January 2008 to December 2009) and consists of 323,917 wire payments. After excluding summary observations (52), rejected wires (1,510), internal transfers/subsidiaries/IRS (113,816), and newly added batch types (298), 208,241 wire transfers remains in the dataset. The excluded batch wires include one type introduced during the fourth quarter. Approximately 90% of the wire transfers belong to 26% of the payees after data cleaning.

Model Development Process

All of the anomaly indicators from Phase III are transferred to Phase IV and one indicator that was dropped in Phase II due to data availability is added again. In addition, the anomaly indicators that are generated with prediction intervals are revised.

An anomaly indicator that examines proper approvals is added to the model. It was excluded in the Phase II after necessary information was found to be missing. With an

www.manaraa.com

algorithm mating wire type and approvers that is provided by the audit team, testing becomes feasible. The algorithm specifies what each type of wire transfer must have for an approver. Confidentiality prohibits further discussion of the algorithm. Another significant change in this phase is that past transactions are separated from the latest quarter to develop better anomaly indicators. Until the last quarter, anomaly indicators with prediction intervals are built with all the available transactions, with those behaving abnormally labeled as potentially anomalous. A problem related to this method is that universal behavior is affected by target transactions (i.e. wire transfers in the latest quarter) whose abnormalities are as yet unknown. Instead, it seems more reasonable to exclude the target transactions while building anomaly indicators, compute prediction intervals, and then apply them to the target transactions. Consequently, anomaly indicators with prediction intervals in this phase are built with wire transfers without those in the latest quarter and abnormality of wire transfers are tested with those intervals.

Screening rules

Anomaly indicators in Phase IV are taken from Phases II and III. The Phase IV model

www.manaraa.com

inherits all of the anomaly indicators from Phase III and one from Phase II. During inheritance, anomaly indicators using prediction intervals change their source dataset to calculate new boundaries. That is, prediction intervals based on past transactions are used to predict abnormality of current transactions. With this change, indicators become more statistically robust.

Approval tests were dropped during Phase III because necessary information was unavailable. More specifically, the algorithm between individual wire transfers and approvers was not disclosed. In this quarter, this information is provided by the audit team, enabling approval testing. Rules for these tests are not firmly fixed, instead changing along with company policy changes. New rules are also added if new wire types are introduced.

After these changes, the detection model in this phase has 6 types of statistical tests (11 indicators) and 15 types of conditional tests (27 indicators). In total, Phase IV has 21 anomaly tests with 38 individual indicators.

www.manaraa.com

Scoring System

The suspicion scoring system in Phase IV is similar to Phase III. From this phase, only equal weighting system is presented. Although unequal weighting is more realistic if relative weights on individual anomaly indicators are accurately measured, the subjective nature of its practical implementation leads to over- and under-estimation of indicators. To avoid confusion due to the unequal weighting system, anomaly indicators are equally weighted while different thresholds are assigned to each category.

Human resource limitations affect the thresholds that are used as cutoff points. As in previous quarters, the default number of wire transfers for verification is thirty. Category thresholds are designed to get the maximum number of flagged wires up to 30 while each category has similar number of flagged wires. As a result, the thresholds are determined as 5 for statistical score, 6 for conditional, and 7 for total. With this condition, twenty wire transfers are sent to the audit team for further investigation. The table 24 illustrates the detail.

www.manaraa.com

Table 24. Thresholds for Each Category

Statistical	cnt_wires
0	52574
1	2092
2	485
3	281
4	92
5	12
6	2

Conditional	cnt_wires
0	34472
1	17787
2	2183
3	987
4	80
5	22
6	5
7	2

Total	cnt_wires
0	32666
1	18227
2	2824
3	1425
4	266
5	84
6	32
7	9
8	3
10	2

Results and Discussion

Results

The highly decentralized nature of relevant information hinders anomaly detection. There are no master files that explain which values can or must exist for variables (e.g. wire type, payee, etc.). Although the audit team ascertains that this information is kept in physical form, it is not easy to determine who keeps the documents and where they are. This might be because the wire payment system collects transactions from various subsidiaries (about 2,300 either domestic or foreign) that become part of the company as a result of mergers and acquisitions or expansions, and the data transition system is not

www.manaraa.com

well-established. The verification results in this phase show that this issue can reduce the quality of anomaly detection.

The periodic nature of verification (during quarterly audit) is another hindrance. Due to lack of necessary information, it is frequent to request verification of certain wire transfers to the audit team. However, this issue is typically resolved only during the regular audit. This delayed response can affect the subsequent quarter.

The twenty flagged wire transfers are inspected by the audit team. Although none are found anomalous, the investigation result contains a surprising fact that was not discovered until Phase III. After the investigation report is analyzed, three issues are raised. First, the areas of loan-related wire transfers and subsidiaries that are not listed in the subsidiary master file are determined to be of low risk. Wire transfers to subsidiaries of the company are excluded because their funds do not leave the company. Although loan-related wire transfers are sent to payees that are outside the company, they will be collected eventually. The risk related to this type of transfer is a potential default that is connected to a collection process rather than the payment system. In addition, the newly found subsidiaries should be excluded as the others were.

Second, certain wire transfers are approved by supervisors that monitor the payment

www.manaraa.com

system. According to the internal audit team, the supervisor approvals are for suspicious or unusual wire transfers. This implies that the anomaly detection model is working as intended, but it also means that these detections are redundant and should be excluded from analysis in the future.

Third, further investigation of supervisor-approved wire transfers discloses the undesirable truth that some employees have two IDs. Every employee should have one and only one ID; possession of more than one violates the entity integrity. Since this example regards a relational database, the audit team was immediately brought in. After examination, the audit team explains that the supervisor worked both in the company and from home. For a VPN (virtual private network) that enables an employee to access the company remotely, the employee used a different ID. Despite this justification, it is doubtful that a separate ID was necessary for remote access. Although more detailed investigation is requested, the audit team seems reluctant to investigate further. This is another example that the company should enhance its database management system.

To summarize, no supportive evidence is found that the flagged wire transfers are either erroneous or fraudulent. However, some flagged wire transfers that were approved by supervisors imply that the detection model is detecting anomalous wire transfers.

www.manaraa.com

Although anomalous wire transfers that are undetected by the supervisors have not yet been found, this study has still shed light on the issue. The next phase will take these findings into consideration for further improvement.

Discussion

This fourth generation anomaly detection model is adjusted based on findings and issues from previous phases. Major findings and changes are summarized as follows. First, a control test that was dropped in the Phase III is added back. The test becomes feasible after relevant information is provided by the audit team. Second, anomaly indicators that use prediction intervals compute upper and lower bounds based on past data and apply the boundaries to target transactions. Third, wire transfers that are loan- or subsidiary-related are newly found and estimated to be less risky areas. These will be excluded in the next phase. Fourth, the fact that some flagged wire transfers are approved by supervisors instead of ordinary approvers implies that the model is working. Finally, a violation of entity integrity is found in approvers. Although supervisors are supposed to have more authority, they do not have the right to have two IDs. This must not be taken for granted. Rather, it may require an urgent action from the audit team or the IT

www.manaraa.com

department. The next phase will start with resolving newly found problems such as modification of data cleaning process.

www.manaraa.com

iv. Conclusion, Limitations, and Future Research

This project provides a dynamically-adjusted anomaly detection model regarding the wire payment system of an insurance company. We begin with a pilot study that tests the possibility of implementing the model within the company's regular audit. Since it is the first attempt for the company to apply an anomaly detection model to their regular audit, the project faces many challenges that are caused by misunderstanding and miscommunication. The most difficult challenge, however, is a lack of direct access to necessary information. Although the company uses a highly computerized database system, the internal audit team cannot access the information it needs. This difficulty is mainly caused by the fact that the company grows by acquiring businesses that have their own database management systems. After each M&A, the systems must be merged, and this conversion is costly and time-consuming. As a solution, the company converts and merges only the most necessary information, leaving the remainings in the charge of the merged company. One result of this process is that indispensable information is kept electronically while less important information exists somewhere within the subsidiary's systems. This causes problems when employees of subsidiaries leave the company. Since the employees are the only people who know where less important information is kept

www.manaraa.com

and what it means, their termination causes breaks of information linkage. After decades of M&A, absence of correspondents interrupts communication of the information that is not centralized. Although it is generally assumed that corporate data is stored electronically in the modern age, this company does not follow that paradigm, making it difficult to find detailed information in a timely manner.

This study provides a variety of useful findings about the company's wire payment system. First, it provides evidence that unsupervised methods can be useful for anomaly detection. When a company initiates anomaly detection activity, it is highly likely that no prior information is available. Once an anomaly detection model is implemented, modification and improvement are simplified. Second, this study presents various findings about the company's database management system. ERP (enterprise resource planning) systems are taken for granted in the business world. ERP customization may provide invaluable time and cost savings, but heavy customization can cause malfunctions. If a company uses a computerized data processing system that is completely customized for its business environment, differences between the system and that of another company can make integration overwhelming, if not impossible. As a result, the insurance company has a mix of integrated and segregated databases. Efforts to

www.manaraa.com

unify all systems may be fruitless due to limited resources. Many flaws have been discovered during model development, some that threaten system integrity and some that can be fixed with gradual changes. Finally, the project finds potentially risky areas that were not considered before. At the same time, certain types of wire transfers are found not to need as much care as others. It is now possible to narrow down the scope of transactions that should be investigated. This is valuable information for future study in anomaly detection.

This study provides a learning experience for academics by showing how an anomaly detection model is implemented and improved in practice. In addition, this study shows that anomaly detection activities can be practically useful during regular audits to help internal auditors identify possible weak or risky areas and transactions.

Although this study has tried to produce an model that detects anomalous wire transfers, it does not achieve this goal. Many issues are raised, resolved, and discarded for this purpose. Despite findings that provide indirect evidence that the model may be working, we cannot claim that the model is functioning properly. Future research must consider all findings and clean data more carefully in order to focus only on the most relevant pool of transactions. Exclusion of irrelevant or less risky transactions can reduce

www.manaraa.com

the noise that adversely affects the quality of an anomaly detection model.

www.manaraa.com

IV. Conclusion, Limitations, and Future Research

This study contributes to anomaly detection research in three ways. First, it provides detailed guidance for the development of an anomaly detection model that is useful for internal auditors to implement in their internal control system. Second, it proposes anomaly detection models at a transactional level via an unsupervised method that is more realistic and beneficial in practice. Finally, this shows that weakly-controlled or risky areas in internal control system are easily identified through the model development process. It is greatly useful to discover those unknown and latent anomaly types.

This study explores the transitory accounts of a bank and the wire payment system of an insurance company for development of anomaly detection models. Although both companies belong to financial institutions, their characteristics vary significantly. This difference mainly stems from the business cycles and environments that are investigated. Although both studies provide a pilot model and subsequent model development, the insurance company wire transfer is more rigorously explored. More frequent communication, more regular feedback, and more consistent support from management facilitate development of the latter's models.

www.manaraa.com

Although both studies present various interesting findings and suggestions, they also feature limitations that must be tackled in the future. The study of transitory accounts faces challenges that are caused by heterogeneous characteristics. Significant differences among transitory accounts make it difficult to apply a model universally. It may be more meaningful to develop account-specific models, but this approach also has drawbacks. First, the number of models increases with the degree of heterogeneity, possibly resulting in prohibitive cost. Second, as the number of models increase, the number of transactions that are used to develop a model must decrease. Insufficient observations can generate less efficient models, if not less effective ones.

While heterogeneity does not significantly affect the insurance company study, decentralized databases are a major obstacle. Decentralizations are caused by the continuous mergers and acquisitions that the company performs. As a result, the company has less efficient communication, hindering information acquisition. Although the anomaly detection models in this study try to tackle those problems completely, they might play a more significant role in model development than anticipated. These caveats must be considered when these studies are examined. Future studies must overcome these issues because there is no perfect system.

www.manaraa.com

Bibliography

1. Allen, Robert D., Mark S. Beasley and Bruce C. Branson., "Research Notes: Improving Analytical procedures: a case of using disaggregate multilocation data", Auditing: A Journal of Practice & Theory, Vol. 18, No. 2, Fall 1999

2. Alles, Michael G., Alexander Kogan and Miklos A. Vasarhelyi, "Feasibility and Economics of Continuous Assurance", Auditing: A Journal of Practice and Theory, 2002

3. Alles, Michael G., Alexander Kogan, Miklos A. Vasarhelyi, and Jia Wu, "Continuity Equations: Business Process Based Audit Benchmarks in Continuous Auditing", Proceedings of American Accounting Association Annual Conference, 2004

4. Alles, Michael G., Alexander Kogan and Miklos A. Vasarhelyi, "Restoring auditor credibility: tertiary monitoring and logging of continuous assurance systems", International Journal of Accounting Information Systems, 2004

5. Alles, Michael G., Alexander Kogan, Miklos A. Vasarhelyi, and Jia Wu, "Analytical procedures in continuous auditing: continuity equations models for analytical monitoring of business processes", Proceedings of American Accounting Association Annual Conference, 2006

6. Ameen, Elsie C. and Jerry R. Strawser, "Investigating the use of analytical procedures: an update and extension", Auditing: A Journal of Practice & Theory, Vol. 13, No. 2, Fall 1994

7. Ashton, Robert H., "An experimental study of internal control judgments", Journal of Accounting Research, Spring 1974

8. Association of Certified Fraud Examiners, Report to the Nation on Occupational Fraud and Abuse, 1996, 2004, and 2007

9. Association of Certified Fraud Examiners, "Occupational fraud: a study of the impact of an economic recession", 2009

10. Bailey, Andrew D. Jr., Gordon Leon Duke, James Gerlach, Chen-en Ko, Rayman D. Meservy, and Andrew B. Whinston, "TICOM and the analysis of internal controls", The Accounting Review, April 1985

11. Biggs, Stanley F. and John J. Wild, "A note on the Practice of analytical review", Auditing: A Journal of Practice & Theory, Vol. 3, No. 2, Spring 1984

12. Biggs, Stanley F. and Theodore J. Mock, "An investigation of auditor decision processes in the evaluation of internal controls and audit scope decisions", Journal of Accounting Research,

www.manaraa.com

Vol. 21, No. 1, Spring 1983

13. Bodnar, George, "Reliability Modeling of Internal Control Systems", The Accounting Review, 1975

14. Bolton, R. J., and D.J. Hand, "Statistical fraud detection: A Review. Statistical Science", 17(3): 235-249, 2002

15. Burns, David C. and James K. Loebbecke, "Internal Control Evaluation: How the Computer can help", The Journal of Accountancy, 1975

16. Calderon, Thomas G. and Brian P. Green, "Signaling fraud by using analytical procedures", Ohio CPA Journal, 07498284, Vol. 53, Issue 2, Apr94

17. Campbell, David R., Mary Campbell and Gary W Adams, "Adding significant value with internal controls", The CPA Journal, June 2006

18. Chen, Yining and Robert A. Leitch, "An analysis of the relative power characteristics of analytical procedures", Auditing: A Journal of Practice & Theory, Vol. 18, No. 2, Fall 1999

19. Chen, Yining and Robert A. Leitch, "The error detection of structural analytical procedures: a simulation study", Auditing: A Journal of Practice & Theory, Vol. 17, No. 2, Fall 1998

20. Coakely, J. R., "Analytical Review: a comparison of procedures and techniques used in auditing", unpublished Ph.D. Dissertation, University of Utah, 1982

21. Cohen, Jeffrey R., Ganesh Krishnamoorthy and Arnold M. Wright, "Evidence on the Effect of Financial and Nonfinancial Trends on Analytical Review", Auditing: A Journal of Practice & Theory, 2000

22. Coglitore, Frank and R. Glen Berryman, "Analytical Procedures: A defensive necessity", Auditing: A Journal of Practice & Theory, Vol. 7, No. 2, Spring 1988

23. Colley, John W. and James O. Hicks, Jr., "A fuzzy set approach to aggregating internal control judgments", Management Science, Vol. 29, No. 3, March 1983

24. Curtis, Mary B. and A. Faye. Borthick, "Evaluation of Internal Control from a Control Objective Narrative", Journal of Information Systems, 1999

25. Cushing, Barry E., "A mathematical approach to the analysis and design of internal control systems", The Accounting Review, Jan 1974

26. Daroca, Frank P. and William W. Holder, "The use of analytical procedures in review and audit engagements", Auditing: A Journal of Practice & Theory, Vol. 4, No. 2, Spring 1985

27. Dugan, Michael T. and Keith A. Shriver, "How to forecast income statement items for auditing purposes", The Journal of business forecasting, Summer 1994

28. Dugan, Michael T., James A. Gentry and Keith A. Shriver, "The X-11 model: A new

www.manaraa.com

analytical review technique for the auditor", Auditing: A Journal of Practice & Theory, Vol. 4, No. 2, Spring 1985

29. Dzeng, Simon C., "A comparison of analytical procedure expectation models using both aggregate and disaggregate data", Auditing: A Journal of Practice & Theory, Vol. 13, No. 2, Fall 1994

30. Elliott, Robert K., "Twenty-First Century Assurance", Auditing: A Journal of Practice and Theory, 2002

31. Elliot, Robert K., "Unique audit methods: Peat Marwick International", Auditing: A Journal of Practice & Theory, Spring 1983

32. Elmer, Peter J., David M. Borowski, "An expert system approach to financial analysis: the case of S&L bankruptcy", Financial Management, Autumn, 1988

33. Felix, William L. Jr. and Marcia S. Niles, "Research in internal control evaluation", Auditing: A Journal of Practice & Theory, Vol. 7, No. 2, Spring 1988

34. Ferris, Kenneth R. and Kirk L. Tennant, "An investigation of the impact of the qualitative nature of compliance errors on internal control assessments", Auditing: A Journal of Practice & Theory, Vol. 3, No. 2, Spring 1984

35. Fihn, Stephan D., "The Quest to quantify quality", JAMA, 2000

36. Gadh, Vandana M., Ramayya Krishnan and James M. Peters, "Modeling internal controls and their evaluation", Auditing: A Journal of Practice & Theory, Vol. 12, Supplement 1993

37. Gaumnitz, Bruce R., Thomas R. Nunamaker, John J. Surdick and Michael F. Thomas, "Auditor consensus in internal control evaluation and audit program planning", Journal of Accounting Research, Vol. 20, No. 2, Autumn 1982

38. Gaunti, James E. and William G. Glezen, "Analytical Auditing Procedures", Internal Auditor, 00205745, Vol. 54, Issue 1, February 1997

39. Geiger, Marshall A., Steven M. Cooper and Edmund J. Boyle, "Internal control components: Did COSO get it right?", The CPA Journal, January 2004

40. Gilb, Tom, "Quantifying quality: How to quantify quality: Finding scales of measure", Javazone conference, 2004

41. Glover, Seven M., Douglas F. Prawitt and T. Jeffrey Wilks, "Why do auditors over-rely on weak analytical procedures? The role of outcome and precision", Auditing: A Journal of Practice & Theory, Vol. 24, Supplement 2005, pp. 197-220

42. Hayes-Roth, Frederick, "RULE-BASED SYSTEMS", Communications of the ACM, Sep85, Vol. 28 Issue 9

www.manaraa.com

43. Hirst, D. Eric and Lisa Koonce, "Audit Analytical procedures: A field investigation", Contemporary Accounting Research, Vol. 13 No. 2, Fall 1996

44. Holder, William W., "Analytical review procedures in planning the audit: An application study", Auditing: A Journal of Practice & Theory, Vol. 2, No. 2, Spring 1983

45. Hylas, Robert E. and Robert H. Ashton, "Audit detection of financial statement errors", The Accounting Review, Vol. LVII, No. 4, October 1982

46. Hwang, Sung-Sik, Taeksoo Shin and Ingoo Han, "CRAS-CBR: Internal control risk assessment system using case-based reasoning", Expert Systems, Vol. 21, No. 1, February 2004

47. Jans, M., N. Lybaer, and K. Vanhoo, "A framework for internal fraud risk reduction at it integrating business processes: the IFR^2 framework". The International Journal of Digital Accounting Research. Vol. 9: 1-29, 2009

48. Jancura, Ellise G. and Fred L. Lilly, "SAS No. 3 and the evaluation of internal control", The Journal of Accountancy, March 1977

49. Jensen, Kevan L. and Jeff L. Payne, "Management trade-offs of internal control and external auditor expertise", Auditing: A Journal of Practice & Theory, Vol. 22, No. 2, September 2003

50. Joyce, Edward J., "Expert judgment in audit program planning", Human Information Processing in Accounting, 1976

51. Kinney, William R., Jr., "Research opportunities in internal control quality and quality assurance", Auditing: A Journal of Practice & Theory, 2000

52. Kinney, William R. Jr., "ARIMA and Regression in Analytical Review: an Empirical Test", The Accounting Review, Vol. LIII, No. 1, January 1978

53. Kinney, William R. Jr., "Decision theory aspects of internal control system design/compliance and substantive tests", Journal of Accounting Research, Vol. 13, No. 3, 1975

54. Kinney Jr., William R. and William L. Felix Jr., "Professional Notes: Analytical review procedures", Journal of Accountancy, 1980

56. Knechel, W. Robert, "The effectiveness of statistical analytical review as a substantive auditing procedure: a simulation analysis", The Accounting Review, Vol. LXIII, No. 1, January 1988

57. Kogan, Alexander, Ephraim F. Sudit and Mikios A. Vasarhelyi, "Continuous Online Auditing: A Program of Research", Journal of Information Systems, Vol. 13, Issue 2, 1999

58. Kreutzfeldt, Richard W. and Wanda A. Wallace, "Error characteristics in audit populations: their profile and relationship to environmental factors", Auditing: A Journal of Practice & Theory, Vol. 6, No. 1, Fall 1986

www.manaraa.com

59. Lys, Thomas and Ross L. Watts, "Lawsuits against Auditors", Journal of Accounting Research, 1994 Supplement, Vol. 32 Issue 3, p65-93, 29p

60. Lev, Baruch, "On the use of index models in analytical reviews by auditors", Journal of Accounting Research, Vol. 18, No. 2, Autumn 1980,

61. Martin, Jack L. and Robert F. Eckerle, "A Knowledge-Based System for Auditing Health Insurance Claims", Interfaces, Mar/Apr91, Vol. 21 Issue 2, p39-47, 9p

62. McMullen, Dorothy A., K. Raghunandan and D. V. Rama, "Internal control reports and financial reporting problems", Accounting Horizons, Vol. 10, No. 4, December 1996

63. Merten, Alan G. and Dennis G. Severance, "Data processing control: A state-of-the-art survey of attitudes and concerns of DP executives", MIS Quarterly, June 1981

64. Murthy, Uday S., "An Analysis of the Effects of Continuous Monitoring Controls on e-Commerce System Performance", Journal of Information Systems, 2004

65. Murthy, Uday S. and Michael S. Groomer, "A continuous auditing web services model for XML-based accounting systems", International Journal of Accounting Information Systems, 2004

66. Nichols, Donald R., "A model of auditors' preliminary evaluations of internal control from audit data", The Accounting Review, 1987

67. Nigrini, Mark J. and Linda J. Mittermaier, "The use of Benford's law as an aid in analytical procedures", Auditing: A Journal of Practice & Theory, Vol. 16, No. 2, Fall 1997

68. Phua, C., V. Lee, K. Smith, and R. Gayler, "A comprehensive survey of data mining based fraud detection research", Artificial Intelligence Review, 2005

69. Raghavan, Kamala R., "Internal control and operational risk: FDICIA, Sarbanes-Oxley and Basel II", Bank Accounting & Finance, 2006

70. Rezaee, Zabihollah, Ahmad Sharbatoghlie, Rick Elam and Peter L. McMickle, "Continuous Auditing: Building Automated Auditing Capability", Auditing: A Journal of Practice and Theory, 2002

71. Rittenberg, Larry E. and Patricia K. Miller, "Sarbanes-Oxley Section 404 Work: Looking at the Benefits", The IIA Research Foundation, January 2005

72. Schnatterly, Karen, "Increasing firm value through detection and prevention of white-collar crime", Strategic Management Journal, 2003

73. Srinidhi, Bin, "The influence of segregation of duties on internal control judgments", Journal of Accounting, Auditing & Finance, Summer 1994

74. Srinidhi B.N. and M.A. Vasarhelyi, "Auditor judgment concerning establishment of substantive tests based on internal control reliability", Auditing: A Journal of Practice & Theory,

www.manaraa.com

Vol. 5, No. 2, Spring 1986

75. Tabor, Richard H., "Internal control evaluations and audit program revisions: some additional evidence", Journal of Accounting Research, Vol. 21, No. 1, Spring 1983

76. Trotman, Ken T., Philip W. Yetton and Ian R. Zimmer, "Individual and group judgments of internal control systems", Journal of Accounting Research, 1983

77. Vasarhelyi, M. A., "A Taxonomization of internal controls and errors for audit research in D.R. Nichols and H.F. Stettler (Editor)", Auditing Symposium V (University of Kansas Press), 1980 pp. 41-58

78. Warren, C., and R. Elliot, "Materiality and audit risk: A descriptive study", 1986, Unpublished monograph, University of Georgia

79. Wheeler, Stephen and Kurt Pany, "Assessing the performance of analytical procedures: a best case scenario", The Accounting Review, Vol. 65, No. 3, July 1990

80. Wiggins, Casper E. and L. Murphy Smith, "A generalized audit simulation tool for evaluating the reliability of internal controls", Contemporary Accounting Research, Vol. 3, No. 2, Spring 1987

81. Wild, John J., "The prediction performance of a structural model of accounting numbers", Journal of Accounting Research, Vol. 25, No. 1, Spring 1987

82. Williams, Kathy, "Evaluating internal controls and SOX compliance", Strategic Finance, 2005

83. Wilson, Arlette C., "Use of Regression models as Analytical procedures: an empirical investigation of effect of data dispersion on auditor decisions", Journal of Accounting, Auditing & Finance, Vol. 6 Issue 3, Summer 1991

84. Wilson, Arlette C. and Janet Colbert, "An analysis of simple and rigorous decision models as analytical procedures", Accounting Horizons, December 1989

85. Woodroof, Jon and DeWayne Searcy, "Continuous Audit Implications of Internet Technology: Triggering Agents over the Web in the Domain of Debt Covenant Compliance", Proceedings of the 34th Hawaii International Conference on System Sciences, 2001

86. Wright, Arnold and Robert H. Ashton, "Identifying audit adjustments with attention-directing procedures", The Accounting Review, Vol. LXIV, No. 4, October 1989

87. Wu, Frederick H. and Randall L. Hahn, "A control-complexity and control-point orientation to the review of an entity's internal control structure in the computer environment", Journal of Information Systems, Spring 1989

88. Yu, Seongjae and John Neter, "A stochastic model of the internal control system", Journal of Accounting Research, Autumn 1973

www.manaraa.com

89. Vasarhelyi, M.A. and Fern B. Halper, "The Continuous Audit of Online Systems". Auditing: A Journal of Practice and Theory. 10 (1) 110-125, 1991

www.manaraa.com

Curriculum vitae

Yongbum Kim

03/1974	Born in Naju, South Korea
03/1989 –02/1992	Youngsanpo High School
03/1992 – 02/2000	Dongguk University, BA in Accountancy
03/2000 – 02/2002	Dongguk University, MA in Managerial Accounting
09/2002 – 08/2004	Michigan State University, MS in Statistics
09/2005 – 09/2011	Rutgers University-Newark, Ph.D. in Accounting and Information Systems

www.manaraa.com