지식베이스와 사례베이스를 이용한 은행 감사
Knowledge-Based and Case-Based Approach for Bank Audit

이 건호
*숭실대학교 산업정보시스템공학과 부교수, ghlee@ssu.ac.kr

Abstract

본 연구는 규칙베이스와 사례베이스를 이용하여 은행의 내부감사 방법을 제시하고자 한다. 감사의 1단계에서는 규칙베이스를 이용하여 감사대상의 거래를 탐색하고 2단계에서는 사례베이스를 이용하여 감사대상의 거래를 심층 분석하여 감사결과를 도출한다.

규칙을 이용한 추론은 내부규정 및 가이드라인을 이용하여 추론하여 잠재적인 위험을 가지고 있는 거래를 발견하는 것이다. 사례베이스를 이용한 추론은 유사도를 개발하여 현재의 문제와 가장 유사한 사례를 탐색하여 감사를 하도록 한다.

본 연구에서 제시한 방법은 실제 은행 내부감사에 적용하여 분석하였다.

1. Introduction

The banks try to reduce the risk assets, prevent themselves from insolvency, and take a rapid action for latent risk through thorough audit using systematic approach.

The purpose of audit is to compare what is to what should be. Internal auditing is an independent, objective assurance and consulting activity designed to add value and improve an organization's operations [11]. Auditing helps an organization accomplish its objectives by bringing a systematic, disciplined approach to evaluate and improve the effectiveness of risk management, control, and governance processes [9].

The review process of work-paper in bank auditing involves the company's most costly personnel in what may be expressed as successive layers of control activity. More than 50 percent of audit manager time and 30 percent of total audit hours has been allocated to review work-papers [2]. This study views that a way to improve auditing efficiency is to use an automatic review system. Bank audit systems are not frequently studied in an academic context, nor is there published research on integrating case-based and rule-based systems for audit systems.

2. Integrated Approach of Bank Auditing

This study presents a bank audit system using Rule-Based Reasoning (RBR) and Case-Based Reasoning (CBR). The bank audit approach integrates RBR, which automatically detects abnormal, irregular, risky, and violated transactions from the standards at the first screening stage, and CBR, which scrutinizes the detected transactions and provides the punishment levels at the second stage.

However, this study views that CBR works better than RBR does in determining and recommending the level of punishment as result of an audit considering a variety of situations since the judgment is cognitive process based on intuition and experiences. Auditor's judgment for punishment of a violator considering situation or circumstances shrewdly is the cognitive process of a reaching decision based on intuition and experiences [1][8]. In general, the CBR approach provides a good reasoning method for decisions based on intuitions and experiences. The CBR has benefits such as a reduction in knowledge-elicitation effort from complex and complicated transaction situations, the ability to learn by acquiring new cases over time without having to add new rules or modify existing ones, and the ability to provide justification by offering past cases as precedence rather than justifying a solution by showing a trace of the rules that led to decision [3][10]. The system has high evolutionability, adaptability, and extensibility since the rule-based and case-based knowledge are stored separately in the knowledge base and are independent of each other and a rule and a case are easily added and removed.

The integrated reasoning approach we proposed here makes use of both the existing knowledge and the past experiences and makes the problems that only limited experience and knowledge are available become solvable. This integrated approach eliminates the drawbacks of each method and provides a better way to handle problems, which combine both inductive and deductive information [3][4].

3. Rule-based reasoning

In the opinion of this study, scrutinizing small amount and low risk transactions is time consuming and not worthwhile. It is desirable and effective for auditor or auditing system to focus on the weighty transactions with high risk.

The knowledge base of a rule-based system comprises the knowledge that is specific to the domain of application with respect to the internal regulations, guidelines, etc. to observe as a form of rules or facts. We refer to each regulation, guideline, etc. in the manual as an 'item' in the checklist. We refer to a specific classification or item detail as a 'sub-item'. For example, the items in loan category comprise business customer's credit loan, a secured loan, mortgage ratio vs. amount of loan, mortgage estimation, report of credit rating in a required time period, insurance against fire, accident, etc, stability review of business, and so on. Each item according to regulation includes a number of rules and guidelines to observe. These items and sub-items of rules and guidelines are translated into and stored in decodable rules in knowledge base. For example with respect to item 01 in Table 2, knowledge base includes the rules corresponding to five sub-items for credit rating and review in the regulation of business customer's credit loan as next.

IF a bank employee didn't identify general credit of a debtor,
ELSE IF general credit of a company wasn't rated,
ELSE IF business feasibility or risk for a business loan weren't assessed,
ELSE IF credit rating report for a debtor wasn't made,
ELSE IF credit rating report for small company wasn't made
THEN audit the transactions with respect to the involved employees.

The system audits further on the transaction if the precedent of any rule is fired. The RBR keeps an identification number of the collected items to audit further through the stage of CBR. Once the system identifies a suspicious transaction or a clue, the system scrutinizes the whole transactions in accordance to the employee who did the transaction. The system audits the whole transactions processed by the involved employees, which are in an identical period to the time of the suspicious transaction.

4. Case-based Reasoning

Case-based reasoning (CBR) solves new problems by matching and adapting the important features on the old cases that were successfully solved before. The important issues are how to represent knowledge in the system, how to focus on important features, how to select the best old case if more than one match are available, and how to perform efficiently the matching process by applying different policies [10].

With similarity-based matching, we can find a source case(s) that is closest to the target case, which we try to solve. In this study, CBR performs at two steps of similarity based matching to reduce the efforts of search and matching. Similarity SM1 measures target case and source case with respect to items at first step and SM2 measures items of each case with respect to sub items at step 2 if SM1 >0, i.e., if no matching items exist in target case and source case.

When several retrieved cases with an identical similarity score are competing, we impose a priority on the cases according to the most reused case, the latest case updated, the case created by auditor committee rather than by individual auditor, etc.

4.1 Item Relationship

To measure similarity of two cases, the relationship of the items of both sides is evaluated in this study. The relationship intensity between a pair of items is considered on the similarity by linguistic variables that consist of identical, super strong, strong, so on. Linguistic variables are represented as symbols and are converted to the commonly used conversion values.

The conversion value referred to as the coefficient indicates the strength of the relationship between the items and between sub items, where a less conversion value denotes a stronger relationship between items or a stronger impact on similarity. Accordingly, the least conversion value of an item to the other indicates that the item is the nearest neighborhood to the other and the two items are the closest relationship. We obtain linguistic variable value for each relationship through a questionnaire which is answered by experters by showing the relativeness of each two items quantatively. Relationship intensity $[d_{ij}]$ between items has properties: $d_{ij} = d_{ji}$, $d_{ij} < d_{ik} + d_{kj}$, $d_{ij} = d_{jk}$ doesn't guarantee $d_{ij} = d_{ik}$ or $d_{ik} = d_{jk}$, respectively where i and j are item numbers.

Similarity metric (SM) between two cases is defined as SM = (SM1 * SM2)/2. Similarities SM1 and SM2 are described in sections 4.2 and 4.3, respectively.

4.2 Similarity 1 (SM1)

Item sub-item incidence matrix $[a_{is}]$ is used to represent similarity measures where a_{is} is the violated sub-item number of item i in case s and a_{is} is zero if item i isn't violated. Let z_i and z_j be digit item vectors of the item - sub item incidence matrix.

$$z_i = [a_{1i}, a_{2i}, a_{3i}, \ldots, a_{Mi}]^T$$

$$z_j = [a_{1j}, a_{2j}, a_{3j}, \ldots, a_{Mj}]^T$$ where M is the number of items

In this paper, the hamming distance is defined to represent similarity measure between two cases. Equations (1) and (2) represent the total number of items in target case t and source case s and the number of common items.

$$\alpha = \sum_{i=1}^{M} \rho(a_{is}, a_{it}) \tag{1}$$

Where

$$\rho(a_{is}, a_{it}) = \begin{cases} 1 \text{ if } a_{is} \neq 0 \text{ and } a_{it} \neq 0 \\ 0 \text{ otherwise} \end{cases}$$

$$\beta = \sum_{i=1}^{M} \sigma(a_{is}, a_{it}) \tag{2}$$

Where

$$\sigma(a_{is}, a_{it}) = \begin{cases} 0 \text{ if } a_{is} = 0 \text{ and } a_{it} = 0 \\ 1 \text{ otherwise} \end{cases}$$

Similarity measure SM1 is a function that represents the ratio of the areas of intersection and union of items, i.e., $\sum_{i=1}^{M} \rho(a_{is}, a_{it}) : \sum_{i=1}^{M} \sigma(a_{is}, a_{it})$, considering the relationship intensity between items, $(\sum_{i \in Z_s} P_i + \sum_{i \in Z_t} Q_i)$. The SM1 score ranges in the interval [0, 1]. The more common items of case s and case t can get the higher similarity. However, even if the common items are less, similarity value may be larger if the relationship impacts more than common items do on similarity. When the items in case s and in case t are identical, i.e. SM1=1, each relationship conversion value of the items is zero and $\sum_{i=1}^{M} \sigma(a_{is}, a_{it})$ is one. To define similarity measure SM1 notation is defined next.

T_s : A set of the checked items in case s.

Z_s : A set of the checked items in case s excluding the items in the other case, e.g., $T_s - (T_s \cap T_t)$

$P_i^{st} = \min \{ d_{ij} ; j \in T_s \}$ for $i \in Z_t$

$$\theta = \begin{cases} 1 : \text{if } \sum_{i \in Z_t} P_i^{st} + \sum_{i \in Z_s} P_i^{ts} = 0 \\ 0 : \text{otherwise} \end{cases}$$

$$SM\,1 = \frac{\sum_{i=1}^{M} \rho(a_{is}, a_{it})}{\left(\sum_{i \in Zt} P_i + \sum_{i \in Z_s} Q_i + \theta \right) \times \sum_{i=1}^{M} \sigma(a_{is}, a_{it})}$$

4.3 Similarity 2 (SM2)

The more intersection of case s and case t can get the higher similarity. However, even though the intersection is less, overall similarity score may be higher if the index distance between sub-items impacts on similarity more than the intersection does.

The similarity SM2 between items with respect to sub items is as next:

$$SM\,2 = \frac{\sum_{i=1}^{M} \kappa(a_{is}, a_{it})}{\sum_{i \in K} v_{st}^{i} \sum_{i=1}^{M} \rho(a_{is}, a_{it})}$$

where

$$\kappa(a_{is}, a_{it}) = \begin{cases} 1 \text{ if } a_{is} = a_{it} \text{ for } a_{is} \neq 0,\ a_{it} \neq 0 \\ 0 \text{ otherwise} \end{cases}$$

K : A set of items i according to non-zero of $\kappa(a_{is}, a_{it})$, $i \in \{T_s \cap T_t\}$

v_{st}^{i} : The distance between sub item of item i in case s and sub item of item i in case t.

4.4 Example outputs of RBR and CBR

CBR output represents the most similar eight cases according to similarity score to the suspicious transactions detected by RBR. The item corresponding to the entry 0 means that it is free from any default but the item corresponding to non-zero entry means that it responsible for the default of transaction.

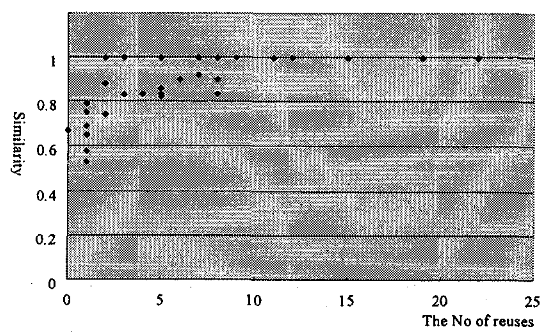

Fig. 1. Reuse and revise for the increase

The items and sub-item in the best three cases are identical but the output types, i.e., punishment levels, are different. The output of a case may be different from each other due to changes of internal regulation and different situation of transaction, inconsistent judgment of auditors, etc. Punishments for the involved employees also may be diverse in accordance to their responsibility.

However, the multiple results may exist for an identical case due to inconsistent judgments of auditors even if every item and sub item is identical. Since a number of auditors with different perspectives produce the results of old cases in case base, the judgment or decision may not be consistent. The system could improve consistency of auditor judgments for an identical problem. Internal audit committee could resolve a controversial issue and make the new standard and criteria for the future.

5. System Analyses and Evaluation

We collect about 39 test problems validated by five internal auditors in K-Bank. The detected results by RBR are identical to the results by the internal auditors for 39 test problems. We got the case with highest score for each problem where most cases are reused without any modification but several cases are modified, adapted, and stored in case base.

Case database can provide data of transactions for risk analysis and a safety management. It is noticeable that the cases with high score of similarity may be reused frequently, which implies the transactions corresponding to the case are frequently occurred and violated.

The punishment levels of the past cases for an identical are different each other which are made manually before the system is used. The items are rated for a priority by using Saaty's analytical hierarchical process method [7] for adaptation of the retrieved cases. Saaty's Consistency Index (CI) shows a consistency of the weights of items from ten interviewees.

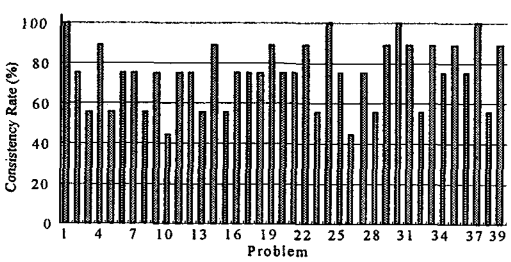

Fig. 2. Reuse and revise for the case increase

Fig. 2 shows the reuse percent (%) and the revise percent (%) of existing case results for 39 problems. As the cases are added in case base, the existing cases are increasingly reused while some cases are decreasingly revised. Four cases (10.25%) out of 39 problems are revised and used based on 600 cases and only one case (2.56%) is done around 1000 cases. We obtained 31 reused and eight revised cases out of 39 problems and compares new case and the best case retrieved by reasoning at 150 cases.

Similarity scores of several cases are low since the similar cases are not enough. Gathering and managing a large variety of cases in the database will improve the output quality of CBR.

6. Conclusion.

Finding suspicious problems in a DB is a time and effort consuming process. There have been auditor's inconsistent judgments even for same problems as the previous cases in the existing audits. The systematic approach in this paper can consistently and efficiently audit even specific and time-urgent problems. Since the output of audit is related to punishment of employee, it is very sensitive for all concerns to deal with a matter. Consistent dealing might be objective and be acceptable for the auditees.

참고문헌

[1] A. D. Balilley, Jr., K Hackendbrack, P. De, and J. Dillard, Artificial intelligence, cognitive science, and computational modeling in auditing research: A research approach, Journal of Information Systems (1987) Spring 20-40.

[2] E. M. Bamber and R.J. Ramsay, The Effects of Specialization in Audit Work papers' Reviews on Review Efficiency and Reviewers' Confidence, Auditing: A Journal of Practice & Theory, 19(2) Fall (2000) 147-157.

[3] R. T. H. Chi and M. Y. Kiang, An Integrated Approach of Rule-based and Case-based Reasoning for Decision Support, in: Proc. 19th ACM Annual Computer Science Conference (1991) 255-267.

[4] A. R. Golding, Improving Rule-Based System through Case-Based Reasoning, in Proc:9th Conference of A.I (1991) 22-27.

[5] R. W. Hamming. Error-detecting and error-correcting codes, Bell System Technical Journal 29(2) (1995) 147-160.

[6] http://en.wikipedia.org/wiki/Basel_II, 2000.

[7] T. L. Saaty and G. Hu, Ranking by the eigen vector versus other methods in the analytic hierarchy process, Applied Mathematical Letters 11 (4) (1998) 121125.

[8] M. Selfridge, and S. F. Biggs, A Computational Model of the Auditor's Going-Concern Judgment, In: Proc. of the Audit Judgment Symposium, University of Southern California, Sos Angeles CA, (1998).

[9] The Institute of Internal Auditors, The Internal Audit Profession, http://www.theiia.org/?doc_id=269.

[10] I. Watson, Applying Case-based Reasoning: Techniques for Enterprise System (Morgan Kaufmann, 1997).

[11] J. T. Wells, An Unholy Trinity: Three Ways Employees Embezzle Cash, Internal Auditor, April (1998) 28-33.