Expert Systems with Applications 38 (2011) 13351–13359

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

A business process mining application for internal transaction fraud mitigation

Mieke Jans [a,*], Jan Martijn van der Werf [b], Nadine Lybaert [a], Koen Vanhoof [a]

[a] Faculty of Business Economics, Hasselt University, Agoralaan, Gebouw D, 3590 Diepenbeek, Belgium
[b] Department of Mathematics and Computer Science, Technische Universiteit Eindhoven, P.O. Box 513, 5600 MB Eindhoven, The Netherlands

ARTICLE INFO

Keywords:
Internal fraud
Transaction fraud
Process mining

ABSTRACT

Corporate fraud these days represents a huge cost to our economy. In the paper we address one specific type of corporate fraud, internal transaction fraud. Given the omnipresence of stored history logs, the field of process mining rises as an adequate answer to mitigating internal transaction fraud. Process mining diagnoses processes by mining event logs. This way we can expose opportunities to commit fraud in the followed process. In this paper we report on an application of process mining at a case company. The procurement process was selected as example for internal transaction fraud mitigation. The results confirm the contribution process mining can provide to business practice.

1. Introduction

In recent years, the problem of internal fraud has received more and more attention. Not unfounded, there the Association of Certified Fraud Examiners (ACFE), an American worldwide organization that studies internal fraud, estimates a US company's losses on internal fraud to be seven percent of its annual revenues (ACFE, 2008). In a previous report of the ACFE, in 2006, this estimation was only 5%, confirming the increasing threat internal fraud poses to companies.

Internal fraud has received a great deal of attention from interested parties like governments or non-profit institutions. The emergence of fraud into our economic world did not go unnoticed. A US fraud standard (Statement on Auditing Standard No. 99) and an international counterpart (International Standard on Auditing No. 240) were created to point auditors to their responsibility relating to fraud in an audit of financial statements. Section 404 of the Sarbanes–Oxley act of 2002 and the Public Company Accounting Oversight Board's (PCAOB) Auditing Standard No. 2 also address this issue. Meanwhile, the CEO's of the International Audit Networks released a special report in November 2006. This report, issued by the six largest global audit networks, is released in the wake of corporate scandals. The authors of this report express their believe in fighting fraud, as they name it "*one of the six vital elements, necessary for capital market stability, efficiency and growth*".[1]

All these standards and reports address the issue of internal fraud (as opposed to external fraud – fraud committed by someone externally related to the company).

In general, two categories within internal fraud can be distinguished: financial statement fraud and transaction fraud. Bologna and Lindquist (1995) define financial statement fraud as '*the intentional misstatement of certain financial values to enhance the appearance of profitability and deceive shareholders or creditors*'. Statement fraud concerns the abuse of a managers position (hence 'management fraud') to alter financial statements in such a way that they do not give 'a true and fair view' of the company anymore. Transaction fraud however can be committed by both management and non-management. The intention with transaction fraud is to steal or embezzle organizational assets. Violations can range from asset misappropriation, corruption over pilferage and petty theft, false overtime, using company property for personal benefit to payroll and sick time abuses (Wells, 2005). Davia, Coggins, Wideman, and Kastantin (2000) state that the main difference between statement and transaction fraud is that there is no theft of assets involved in financial statement fraud (FSF).

Turning to academic studies on this subject, some research is found concerning internal fraud. Green and Choi (1997), Lin, Hwang, and Becker (2003) and Fanning and Cogger (1998) assess the risk on FSF by means of neural networks. Deshmukh and Talluru (1998) use a rule-based fuzzy reasoning system for the same goal and Kirkos, Spathis, and Manolopoulos (2007) use several data mining techniques in order to identify financial factors to assess the risk on FSF. Hoogs, Kiehl, Lacomb, and Senturk (2007) use a genetic algorithm approach to detect patterns in publicly available financial data that are characteristic for FSF. This approach uses a sliding-window approach for evaluating patterns of financial data over quarters in terms of potentially fraudulent or not.

* Corresponding author. Tel.: +32 11268652.
E-mail addresses: mieke.jans@uhasselt.be (M. Jans), j.m.e.m.v.d.werf@tue.nl (J.M. van der Werf), nadine.lybaert@uhasselt.be (N. Lybaert), koen.vanhoof@uhasselt.be (K. Vanhoof).

[1] The remaining five elements concern investor needs for information, the alignment and support of the roles of various stake holders, the auditing profession, reporting and information quality.

0957-4174/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2011.04.159

As can be seen, all articles on internal fraud that are using expert systems, discuss financial statement fraud, which is only one type of internal fraud. Aside from this, a lot of expert systems are investigated in the context of external fraud. (External fraud is fraud committed by someone external to the company, for example a supplier sending false invoices.) It is no coincidence that only this one type of internal fraud, transaction fraud, is not yet addressed in academic literature. Looking at the articles on internal statement fraud and external fraud, all studies, but a few, use supervised data sets. Supervised data sets are provided with a labeled output attribute, in this case 'fraudulent' versus 'legitimate'. The availability of these data sets in the case of statement fraud is to explain by the public nature of financial statements. A company needs to file its financial statements with the government. As a result, fraud committed on these statements is gathered at one central point, normally classified meticulously in order to prosecute these companies. Files on external fraud are also classified very precise for the same reason. Also, there are no reputation related incentives to keep these fraud numbers away from the public, as there is with discovered internal fraud. The faith of stakeholders in the company plummets when stories about internal fraud leak. While a company cannot control this 'leakage' for statement fraud uncovered by the government, they can control the information aspect on statement fraud. This incentive, together with the dispersed methods of committing transactional fraud and the lack of enough fraud files documented meticulously in a company or business process, leads to a general absence of supervised data sets concerning transactional fraud. We believe this is the reason for the literature gap on expert systems for internal transaction fraud. This gap contrasts strikingly with the accompanying costs of this type of fraud.

In two other papers (Jans, Lybaert, & Vanhoof, 2009, 2010), we suggest to use and apply descriptive data mining techniques for internal fraud risk reduction, which also includes mitigating transaction fraud. In this paper, we wish to extend the suggested framework with the field of process mining. Yang and Hwang (2006) already use a process mining approach to detect health care fraud, a type of external fraud that is intensively investigated. We believe the added value of process mining is particularly high in the mitigation of internal transaction fraud. By mitigation, we aim at both fraud detection and fraud prevention. By applying process mining at business processes, a company gains insights in the way procedures are followed or circumvented. This study reports on the application of process mining in a case company.

An organization has business processes mapped out in procedures, guidelines, user guides etcetera. With process mining, we visualize the *actual* process that occurs in a certain business unit instead of the *designed* process. This way one can detect flows or sub flows that for example were not meant to exist. This can give insights in potential ways of misusing or abusing the system. Process mining also provides the possibility to specifically monitor internal controls, like for example the four-eyes principle or the segregation of duty. As opposed to currently wide used internal control tests, the process mining approach for monitoring internal control is data oriented, and not system oriented. In other words: we are able to test whether the true transactional data (the output of the internal control system) are effectively submitted to the presumed internal controls. Instead of testing whether the internal control settings function by means of performing a set of random tests, we mine the actual submitted data and are able to test whether all conditions are met.

Another advantage is the objectivity with which the process mining techniques work, without making any presuppositions. We see the exploratory diagnostics step as a starting point to evaluate with an open mind what opportunities possible deviations can mean for a perpetrator. This is opposed to interpreting results with a specific fraud in mind, resulting in possible blindness for other opportunities. On the other hand, when mining the organizational and the case perspective (see below), it can be beneficial to have some specific fraud(s) in mind. This is certainly the case when monitoring internal controls. At this stage specific internal controls, motivated by specific frauds in mind, are monitored and checked.

We start the paper with an introduction in process mining in Section 2. In Section 3 we give information on the technique used in this application. In Section 4, the application of this technique in a case company is presented. Sections 4.1 through 4.5 describe the process diagnostic steps. Process diagnostics are necessary in order to first confirm the event log captures the general process and next to reveal weaknesses and problems in the business process. In Section 5 we advance to a verification step where we check whether certain assertions of the process hold or not. We end with a conclusion in Section 6.

2. Process mining

Many information systems that nowadays support business processes, like ERP, WFM, CRM and B2B systems, are characterized by the omnipresence of logs. Typically, these information systems record information about the usage of the system by its users. These logs contain information about the instances processed in the system (also called cases), the activities executed for each instance, at what time the activities were executed and by whom. Some systems also contain information about the data users entered for each activity. However, this data is not actively used by the organization to analyze the underlying processes supported by the system.

Process mining aims to make a difference. "*The basic idea of process mining is to diagnose processes by mining event logs for knowledge*" (van der Aalst & de Medeiros, 2005). It allows to analyze these event logs, sometimes also referred to as 'audit trail', 'transaction log' or 'history'. Records in these logs are called *events*, or 'audit trail entries'. In process mining, each event needs to refer to an *activity* for a specific *case* or *process instance.* Preferably, each event also refers to the performer, the *originator* of the event, and a *time stamp.* For each process under investigation these are the constraining assumptions. If available data fulfills these assumptions, process mining can be applied on that particular process.

Event logs are the starting point of process mining. The data of the event log can be mined and different aspects about the underlying process can be analyzed. In general, three different perspectives can be distinguished: the process perspective, the organizational perspective and the case perspective. The *process perspective* tries to answer the "How?" question, and focuses on the ordering of activities. The main focus in this perspective is process discovery and delta analysis. It tries to answer the question "Which paths are followed?" in the most broadest sense. Typically one of the results in this perspective is a process model, mostly expressed in graphical notations like Petri nets, event-driven process chains (EPCs) or in the business process modeling notation (BPMN). The *organizational perspective* focuses on the "Who?" question. It analyzes event logs based on the users, called the originators, that play a role within the process. In this perspective, underlying relations between performers or between performers and tasks can be exposed. Typical examples in this perspective are social networks, handover of work, and cooperation diagrams. The *case perspective* or the "What?" question focuses on a case in isolation. Typically for this analysis, the log needs to be enriched by extra data about the case. This can be data about the complete case, or data for a specific event, like the data submitted at the event (van der Aalst et al., 2007).