Fraud detection tools

Katerina Hawlova
Department of Information Technologies
University of Economics, Prague, Czech Republic
katerina.hawlova@gmail.com

Abstract: *This article aims to introduce to readers the topic of fraud management – detection of fraudulent behaviour. The article is divided into two parts. The first part presents what is meant by fraud and fraudulent behaviour. In the second part a case study dealing with fraudulent behaviour detection in the procurement area is introduced.*

Key words: Fraud management, fraud, Fraud Detection Tools

1. Introduction

In all business areas where employees get in touch with money or can somehow decide about funding or influence where should money go, there is a possibility that the financial resources are not spent effectively and at the end this means a financial loss for the company.

This includes mainly those industrial sectors with large amount of money. The money could be in form of one time contracts or in form of small but repeated financial transactions. These sectors could be for example telecommunications, healthcare, insurance, banking and also public sector which is not business-oriented, but there is also huge amount of money.

Nowadays, both private and public sectors have well-prepared and set processes for the handling with financial resources and also processes for people who come in touch with money within the company.

More and more organizations deal with the question if there is a possibility of fraud although they have processes and control mechanisms in place. In this situation fraud detection tools that enable analysis of transactions, activities and behaviour of individuals come in place. These tools can compare behaviour against each other and also against processes and environment in which it occurs.

Fraud is a million dollar business and this figure is rising. According to a study by Price Waterhouse Coopers *"In fact, almost half of the companies surveyed (45 percent) reported being victims of fraud, suffering an average loss of over US$ 1.7 million (from those frauds which involve the loss of tangible items)."* (PricewaterhouseCoopers, 2005) Different routes detect frauds. The traditional way that works in almost all major companies are internal control systems and internal audit that aim not only to detect fraud but also to try to prevent its further occurrence.

In case that internal audit is not enough to prevent fraud there are external companies that are asked to carry out on-going processes (external companies often are a mandatory part of control). Risk management department, new technologies or control systems are also in place to help companies.

Fraud.

According to uslegal.com fraud is defined as:

„Fraud is generally defined in the law as an intentional misrepresentation of material existing fact made by one person to another with knowledge of its falsity and for the purpose of inducing the other person to act, and upon which the other person relies with resulting injury or damage. Fraud may also be made by an omission or purposeful failure to state material facts, which nondisclosure makes other statements misleading." (USLEGAL, 2013)

Unusual behavior

Unusual behavior is a behavior, which is somehow different from the other individuals. Unusual doesn't have to mean fraudulent, for example when one employee is high performer or low performer.

There are lot of areas where fraud detection systems can be deployed. With the growth of Internet commerce and Internet boom in the 90's new ways of fraud appeared. More and more transactions are taking place online and companies like Amazon and e-bay have to face the threat of fraud. All

companies have to pay special attention to their spending and financial situation during financial crisis but criminals seem to be always one step ahead.

Financial fraud can be devideddivided into two groups according to who is the fraudster. Fraud can be commitedcommitted by both external stakeholders and in many cases even by internal employees, which can represent a greater threat, as they know internal processes and weaknesses better. Recently it shows that managers are those who cheat more than other internal employees and the most affected areas are financial departments.

According to KPMG analysis of global patterns of fraud most people involved in commiting fraud work in the finance function, see Tab 1. *("KPMG gathered data and details from fraud investigations conducted by our firms' forensic specialists in EMA, the Americas, and Asia Pacific from January 2008 to December 2010. In all, 348 cases from 69 countries were analyzed.")* (KPMG, 2011)

Tab 1 – Departments in which fraudster mostly work (iHNED.CZ, 2011)

Where the fraudster works		
	2011	**2012**
Finance	32%	36%
Operations/Sales	25%	32%
Procurement	8%	9%
Back office	1%	5%
Research	1%	3%
Legal	0%	2%

2. Fraud detection tools

There are several views on fraud detection tools. The next chapter will include two possible ways how to understand these tools.

According to data processing the systems can be divided into expert systems and systems based on supervised or unsupervised methods.

Another way how to look at these systems is according to their focus. There are systems that cover the entire life cycle from gathering data to evaluation of identified cases and prevention measures design. In addition there are branch/industrial systems that are used in different conditions and are prepared for particular industrial sectors.

2.1 Data processing

2.1.1 Expert systems

These systems include expert knowledge. Knowledge is represent in form of data and rules. These rules are used in case that there is a problem which has to be solved. Standard computer programs solve the problems using common decision logic which contains a minimum of knowledge and it uses only basic algorithms for solving specific problems. This "knowledge" is part of the program code, it means that if there is some change in this "knowledge" there has to be also a change in the code.

Expert systems continuously collect know-how from people and store this knowledge in one „knowledge base" and when the problem occurs the best rule is chosen to solve it or to alert the user. The advantage of these systems is the possibility of changing/adding/editing knowledge so the system becomes more and more "intelligent". (Nwigbo & Agbo, 2011)

Expert systems started to be developed from artificial intelligence, they are working based on common IT and computer science. In this respect, expert systems seem to be a conventional information systems and can be understood as an extension of these systems. As mentioned above, expert systems use the functionality which conventional systems can't use. Author of the book Expert Systems: A Manager's Guide (Management Development) adress following, see Tab. 2, differences between expert and conventional systems. (Wiig, K.,1990).

Tab. 2 – Differences between expert and conventional systems (Wiig, K.,1990)

Expert Systems	Conventional Systems
Allow automation of functions that can be performed only manually before.	Requires additional skills and different perspective on: - Knowledge definition and classification - Knowledge usage automation - Workflow effectiveness
Focus more on operational and managerial positions than on the technical solution.	It is necessary to take into account the commercial and strategic value of knowledge and human thinking in key operational areas.
It turned out that the ES have higher commercial and strategic value in the practical application than do conventional systems in fulfilling the same function.	

Other differences could be for example cost of implementation, knowledge base preparation, etc. An example of the application of expert systems in the financial field is expert systems for mortgages.

Figure 1 shows the basic components of an expert system and the roles that are necessary for the system running.

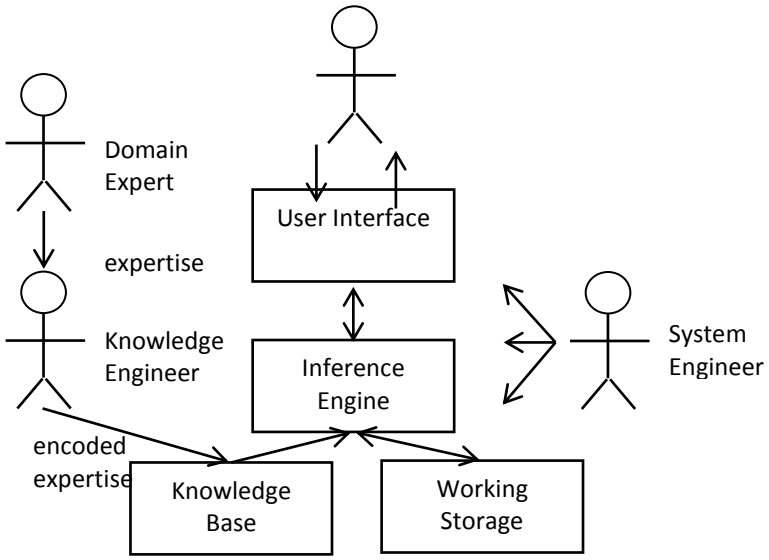

Figure 1 - Expert system components and human interfaces (Amzi! Inc, 2000)

Knowledge base - a declarative representation of the expertise, often in IF THEN rules;

Working storage - the data that is specific to a problem being solved;

Inference engine - the code at the core of the system, which derives recommendations from the knowledge base and problem-specific data in working storage;

User interface - the code that controls the dialog between the user and the system.

Domain expert - the individual or individuals who currently are experts solving the problems the system is intended to solve;

Knowledge engineer - the individual who encodes the expert's knowledge in a declarative form that can be used by the expert system;

User - the individual who will be consulting with the system to get advice which would have been provided by the expert.

System engineer - the individual who builds the user interface, designs the declarative format of the knowledge base and implements the inference engine. (Amzi! Inc, 2000)

Expert systems use backward or forward chaining when solving problem. Backward chaining uses the IF → THEN structure when creating rules and trying to break up the problem to the smallest possible pieces that will be easy to understand. Forward chaining also use the IF → THEN structure but in the opposite direction. This means that it goes from the detail to the „big picture" – it derives the solution based on input (initial) data.

2.1.2 Supervised methods

Another type of fraud detection tools are those tools which use supervised methods to detect fraud. "Supervised methods are methods that attempt to discover the relationship between input attributes (sometimes called independent variables) and a target attribute (sometimes referred to as a dependent variable). The relationship discovered is represented in a structure referred to as a model." (Rokach & Maimon, 2005)

There are two types of those models:

- **Classification model** – works by assigning attributes to a pre-defined classes with same behaviour,
- **Regression model** – basic principle of this model is to estimate the future stated based on available information and experience.

Overview of concepts used in supervised machine learning:

- **Testing/Training set** – it's aim is to create a description that is used to predict the behaviour,
- **Induction algorithms** – works with the testing set and creates a classification model that generalizes the relationship between the input attributes and the resulting attribute,
- **Performance evaluation** – it's important to understand the algorithm and model for clearer specification parameters in the process of mining and for selecting the most appropriate model/algorithm,
- **Scalability of large data sets** – characteristic that distinguishes knowledge data discovery from traditional methods is scalability of large data sets. Capacity of data storage continues to grow and the companies have huge amounts of data that allow them to use tools for data mining. (Rokach & Maimon, 2005)

Supervised methods are mostly used by companies for behaviour analysis. Companies need to understand what is hidden in their current client's base whether it's fraud or cross-sell potential. Based on examples from the past that are used to prepare and "train" the statistical model which than calculta the probability of fraud or churn, cross-sell, etc.

2.1.3 Unsupervised methods

Unsupervised methods are also called „unsupervized learning". Logic of these methods is that at the begining it's not known what will be the output. The most used method is clustering where data with the same characteristics are organized in small groups – clusters. Within these groups there are then identified entities with features that differ most from the others in different clusters. The characteristics in this case mean the amount of statistical values that are compared together. The advantage of unsupervised methods is the ability to discover new patterns and flexibly respond to current requirements. Also requirements for preparation of input data are lower – the target variable doesn't have to be defined. (Rokach & Maimon, 2005)

Unsupervised methods can be used for identifying areas where is the higher possibility of fraud. By identifying these areas, it is possible to prevent fraud or to at least reduce the number of fraudulent cases.

2.2 Focus areas

2.2.1 Framework

Framework is not focused on specific area. It represents a general definition of the functionality of fraud detection tools.

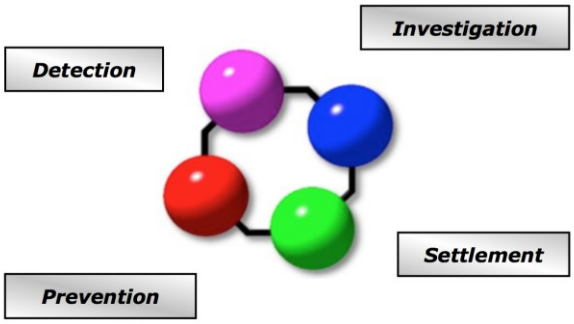

Figure 2 – Fraud detection cycle (IBM, 2011)

The process of detecting fraudulent behaviour covers the whole methodological cycle. This cycle is shown in Figure 2.

The outputs of detection are reports containing the list of suspicious subjects and cases that need to be further investigated. The result of **investigation**, which is the responsibility of customer, is to avoid the manipulation with data to prevent companies from financial theft. Based on results of ongoing investigation the **optimization** and **prevention** steps are designed. The decision-making process is then objective and systematic. Decision rules for fraud detection are implemented in fraud detection tools and in case these rules are not met, unusual behaviour is detected and warning message is sent to the user. Basic prerequisite of optimization project is the performance management system that is able to point out weaknesses and propose prevention steps. The solution also contains predefined KPIs (Key Performance Indicators) that are used to measure overall performance of the process. It continuously increases efficiency of the whole process and monitors the implementation of prevention steps and also monitors the current amount of money spent on the process.

There are lot of steps that have to be taken for the successful implementation of the detection process. The process has to be set and understood both by the client and the supplier side. Other prerequisite is the existence of structured data in digital form allowing efficient detection.

2.2.2 Branch/industry solutions

Financial crime is a growing problem in banking and financial services, especially in times of economic crisis. Fraudster of all kind – hackers, identity thieves, etc. constantly change tactics, focus on new channels and invent new types of attacks, so traditional methods of fraud detection are still a step back.

Most commonly used systems that banks use for financial fraud detection are transactional monitoring systems that banks designed and implemented by themselves.

There are solutions focused on the financial sector that cover end-to-end technology to detect, prevent and manage fraud.

Framework contains components for detection, alert and case management together with specific workflow, content management and advanced analytical tools.

Framework for the banking sector includes analytical tool that assigns score to the ongoing transactions among different accounts and channels. The score is calculated based on probability of fraud. To assign a score following techniques are used:

- automatic business rules,
- predictive modelling,
- text mining,
- database searches,
- alert reporting,
- network connection analysis.

Authorized team further reviews detected cases. In case it is proved that the identified case was fraudulent the team will focus in detail on that particular transaction or a particular customer.

These tools for fraud detection are most commonly used in the following areas:

- money laundering,
- credit card fraud,
- online or organized fraud.

Alert management

The system collects alerts from various monitoring systems, connects them with banking accounts or customers and provides investigators with valuable information. Additional functionality includes:

- risk score calculation,
- risk prioritization,
- alert assignment.

Case Management

Case management facilitates investigation, captures and displays all the information related to analysis of the case. Further case management allows you to:

- store information about fraudulent behaviour,
- provide complete overview of fraud,
- assign cases to investigators,
- include configurable workflow management (SAS, 2011).

3. Procurement fraud detection process – Case study

The aim of this project was to detect unusual behaviour (fraud, anomalies) in the procurement information system. The objective was to prevent possible corruption, uncover cooperation between suppliers, more effective use of public procurement, overall reduction of expenditures on tenders and increase the level of compliance with the rules both by internal users and external users. In this project a FAMS (Fraud and Abuse Management System) from IBM was used.

On the following figure there is a sequence of activities that were performed throughout the solution. In this example first 4 steps will be described that depend on experts and are key activities in the whole project. Next steps are matter of the application itself.

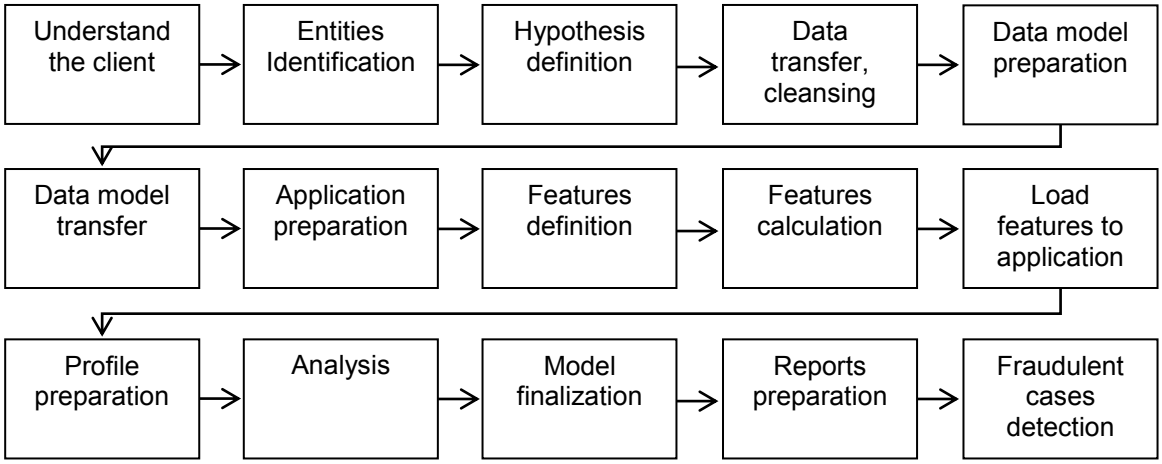

Figure 3 – Fraud detection process

1) Understand the client. In order to be able to make an accurate conclusion from ongoing analysis, first and one might say the most important step in the process of detecting fraudulent behaviour is tounderstand client, area, processes, rules, etc.. Also it is necessary to understand data which customer has and identify data sources and data that are relevant for the analysis.

> **Data sources analysis.** Another important step is the analysis of data sources that are available. Once the data sources are known data that client can provide are identified. This information is crucial for further progress, for entities definition and data models preparation.

2) Entities identification. The second step is to identify the entities on which we will focus in the analysis. Entity is object from reality, in our case a person who is a potential fraudster. In case of procurement information system the entities are internal users that are part of selection process. Also external users (suppliers) will be included to the observation.

3) Hypotheses definition. Hypotheses are possible scenarios of fraud. In this step, for each entity there is described a situation of how fraud can be performed. Communication with the client is very important in this step because client is the one who has detailed knowledge of processes and rules defined in the information system. There are two ways how the detection of unusual or fraudulent behaviour can be done. Each way represents a specific way of detecting fraudulent or unusual behaviour.

- Profiling (systematic detection of non-standard behaviour) – for example if supplier in the procurement process is often able to bid at lower price and than others
- Business rules (each transaction is checked against defined rules) – for example the supplier is biding. There are defined several rules that control procurement process and if one or more of the rules are not met, alert message will be sent to internal user to check the particular bid.

Hypotheses are prepared for defined entities (in our case internal and external users of information system). In many cases it may happen that for different entities the hypothesis will be the same. These are cases in which it makes sense to deal with a given problem from perspective of both internal and external user of information system.

The following hypotheses were prepared for our example – procurement information system. To be able to confirm or reject the hypotheses, there are several charakteristics defined for each hypothesis.

Defined Hypotheses

- Tailor-made procurement process – this hypothesis should detect internal users that might collaborate with any of the companies making bid to adjust the conditions of procurement process precisely for that company.
    - average price for the commodity (there can be a higher price if the procurement process is prepared for particular company),
    - percentage of procurement processes that have a lot of criteria that have to be met and the procurement period is very short,
    - average number of bids (the criteria can be too specific for the suppliers so they can met all the criteria during procurement period).
- „Quick procurement process“ – this hypothesis should detect those users who set a procurement process that has duration less than three days. Three days is the minimum duration of the procurement process. If the period is shorter it may be a potential fraud or lack of knowledge of internal rules, etc.
    - percentage of procurement processes with duration less ten four days,
    - average duration of tenders.
- **Answers** – this hypothesis should detect users who make bids with non-standard number of answers. High number of answers can indicate wrongly set procurement process or it can indicate internal employee that set the criteria in a way that only one particular supplier is able to meet them.
    - average number of answers to the tender
- **Cancelled procurement processes –** this hypothesis should detect internal employees who cancelled procurement processes more often than others. This can occur because of missing knowledge of information system or there is a possibility that someone else than preferred supplier can win.
    - percentage of cancelled tenders
- **The period when the procurement process was cancelled –** this hypothesis should detect those internal employees who cancel the tenders shortly before the end of procurement process. Such behaviour may indicate cooperation between the internal employee and supplier.
    - percentage of cancelled tenders shortly before the end of process
- **Circumvent the rules –** this hypothesis should detect those internal employees who split one bigger tender into several smaller and the total sum is higher than for single tender

- number of tenders
- number of tenders with short period

4) Data transfer, cleansing, quality. After data transfer the consolidation and cleansing have to be done. In many cases it happens that the data is not in a form that can be directly used for following calculations. There are duplicates, typos in names, etc. This step is therefore necessary to create a consolidated data set that can be further used for attributes calculation.

Detection models for profiling and entities detection were created based on defined hypotheses and calculated attributes. Detection models combine data and evaluate behavioural characteristics that can represent unusual behaviour. The logic is that in case one characteristic is higher it doesn't have to mean entity is suspicious and also if one characteristic is not met the entity still can be detected. It means that combination of all characteristics is crucial for detecting fraudulent behaviour.

The basis for evaluation is numerical values of defined characteristics that quantify individual features of entities. Profiling is a process where certain set of characteristic is selected and then for particular group of entities will calculate the score and ranking of individual entities in the model.

Within each group there are identified entities that were considered suspicious based on performed analysis. The application allows creation of reports where each entity is displayed based on it's score. This report is then used as an input for further investigation.

4. Conclusion

Effective use of these tools is deeply connected with a detailed understanding of ongoing processes in particular area and the ability to tailor solution to the specific condition of the customer. Nowadays there are number of products that support required functionality. Nevertheless the human experts are still important part of the system. The tool itself will not work and the outputs will not meet the desired requirements of companies and this is when the expert comes in place – the tool has to be properly set up and the outputs have to be interpreted in a way managers will understand.

The main benefit of these systems together with human expert is an identification of risk/suspicious entities that can cause financial losses. Current economic situation plays into the hands of these tools. Fraud detection represents a possibility where companies can save money by spending financial resources effectively. Nevertheless the final decision is up to management.

References

Amzi! Inc, 2000, *Building Expert Systems in Prolog* [Online] (Updated November 2000) Available at: http://www.amzi.com/ExpertSystemsInProlog/xsipfrtop.htm [Accessed 16 September]

IBM, 2011, *Technical description of the solution*, internal documentation

iHNED.CZ, 2011, Jak podvádí bílé límečky [Online] (Updated August 2011) Available at: http://bankovnictvi.ihned.cz/c1-52669360-jak-podvadi-bile-limecky [Accessed 16 September 2013]

KPMG, 2011, *Analysis of global patterns of fraud, Who is the typical fraudster?* [Online] (Updated April 2012) Available at: http://www.slideshare.net/dn131282nvj/global-fraudpatterns [Accessed 16 September 2013]

Nwigbo Stella N & Agbo Okechuku Chuks, 2011, *Expert System: A Catalyst in Educational Development in Nigeria*. Omoku, Rivers State, Nigeria. [Online] (Updated August 2011) Available at: http://www.hrmars.com/admin/pics/261.pdf [Accessed 16 September]

PricewaterhouseCoopers, 2005. *How to manage and mitigate the risk of corporate fraud* [Online] Available at: http://www.pwc.com/en_us/us/forensic-services/assets/06-investigations-forensic.pdf [Accessed 16 September 2013]

Rokach Lior & Maimon Oded, 2005, *Data Mining and Knowledge Discovery Handbook.* Springer, ISBN-10: 0387244352

SAS, 2011, SEMMA [Online] (Updated September 2011) Availbale at: http://www.sas.com/offices/europe/uk/technologies/analytics/datamining/miner/semma.html [Accessed 16 September]

USLEGAL, 2013, *Fraud Law & Legal Definition* [Online] (Updated September 2013) Available at: http://definitions.uslegal.com/f/fraud/ [Accessed 16 September 2013]

Wiig, K.,1990, *Expert Systems: A Manager's Guide (Management Development).* International Labour Office. ISBN-10: 9221064468

JEL Classification: K42, M10