Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 138 (2018) 680–687

Procedia Computer Science

www.elsevier.com/locate/procedia

CENTERIS - International Conference on ENTERprise Information Systems / ProjMAN - International Conference on Project MANagement / HCist - International Conference on Health and Social Care Information Systems and Technologies, CENTERIS/ProjMAN/HCist 2018

Internal fraud in a project-based organization: CHAID decision tree analysis

Mirjana Pejić Bach[a,*], Ksenija Dumičić[a], Berislav Žmuk[a], Tamara Ćurlin[a], Jovana Zoroja[a]

[a]*University of Zagreb, Faculty of Economics and Business, Trg J. F. Kennedyja 6, 10000 Zagreb, Croatia*

Abstract

Data mining is one of the most prominent methodologies demonstrating efficiency in various tasks related to pattern detection. Applications of data mining for fraud detection are numerous, for the detection of fraud, ranging from credit card fraud to taxation fraud, are numerous. However, the applications of data mining for fraud related to internal control are scarce. The goal of this paper is to present the application of data mining techniques (the CHAID decision tree) for discovering patterns in internal control-related fraud in one project-based organization.

Keywords: Project-based organization; internal control; fraud, data mining; CHAID; association analysis

* Corresponding author. Tel.: +0-000-000-0000 ; fax: +0-000-000-0000 .
E-mail address: mpejic@efzg.hr

1877-0509
10.1016/j.procs.2018.10.090

1. Introduction

Fraud is one of the critical on-going problems in any organization, regardless of whether it results in legal repercussions or not, especially if the system is not fail-safe and if the prevention procedures are not constantly updated and monitored[18]. Therefore, it is crucial for every organization to have a firm and well-adjusted fraud control and detection system to increase the automatization of screening processes and tracking down fraud perpetrators. This is particularly important in project-based organizations[21,22], which are more complex to manage compared to other types of organizations[3,8,19].

There are various studies and investigations on data mining and fraud detection, which provide a wide array of different methods interpreted variously in different studies. Hence, most of those studies are focused on external fraudulent activities, while the research focusing on the internal control fraud is scarce. Internal control fraud articles are scarce, e.g. [12,23]. Therefore, the goal of this paper is to investigate the usability of data mining in detection of internal fraud, using employee working-hour claims as an example. In our paper, we focus on a case study analysis conducted in a project-based company, whose main activity is the development of business-related software applications. It employs more than 300 employees, many of whom are experts working on specific projects.

Contributions of this paper are two-fold. First, this paper is one of the rare attempts of utilizing data mining for increasing efficiency of internal fraud in project-based organizations, while most of the research focuses on external fraud, and publications on internal fraud are rare. Second, the conducted data mining analysis of internal control fraud in a project-based organization resulted in the generation of the SQL code that can be used for development of an automatic fraud-detection software application. However, organizations aiming to develop a solution for automatic internal fraud-detection are advised to develop their own solutions, using the proposed approach, the SQL code being a useful venue in this endeavor.

The paper is organized as follows: an introduction is given in the first section. Section 2 provides the summary of related previous research. Section 3 explained methodology for the data mining technique used in the article, first the research sample, then, the analysis itself. Finally, section four presents empirical results. Concluding remarks, including directions for future research are given in the last section. This research has been fully supported by the Croatian Science Foundation under the PROSPER (Process and Business Intelligence for Business Performance) project (IP-2014-09-3729).

Nomenclature

CHAID Chi-squared Automatic Interaction Detector algorithm

2. Literature review

2.1. Internal controls and fraud

Fraud represents a severe problem in companies, weather committed outside or inside an organization. The purpose of internal control is to support company's performance and achieve established goals. Opportunities for fraud occur in organizations, which have weak compliance with internal controls[16]. Jans et al.[11] confirm that internal fraud is a growing problem in many companies and organizations. It is necessary to investigate this problem further and deeper in order to get better internal control systems. Baldock[2] indicates that insufficient and flouted internal controls give opportunities for personnel to commit unethical practices and fraud in an organization, indicating that the weakness of an internal control system is likely to lead to fraudulent activities. On the other side, Allen et al[1] claim that an effective and efficient internal control system can decrease fraud and human error, and finally add the value to the company itself. There are numerous recommendations related to increasing the efficiency of internal control systems, such as the usage of global positioning tracking units (GPS), monitoring of unutilized purchase orders and pre-approval of overtime work. However, progress is slow due to difficult access to data of previous cases, so it is hard for problem solvers to develop new methods and solutions[23].

2.2. Data mining for internal fraud detection

Data mining models are meant to become one of the most relevant applications for fraud detection in industries and the government. In terms of data mining, fraud analysis is a process, which consists of a sequence of functions, aimed at predicting or discover potential or explicit threats of fraudulent activities. The process relies on techniques from a wide variety of areas including data science, statistics, machine learning etc. The efficacy of a fraud detection system largely depends on the efficiency of the used techniques and the quality of available databases[15]. However, a significant progress took place, and automated fraud detection systems have gained enormous popularity, especially within financial institutions[20]. Data mining has had remarkable results in diverse fields related to security and fraud, financial crime detection (money laundering, suspicious credit card transactions and financial reporting fraud), intrusion and spam detection [7]. However, data mining implementation in the area of internal fraud risk reduction is scarce. Bologna et al[4] distinguish two categories on which most studies are focused: financial statement fraud and transaction fraud. Abuse of position is also a point of interest; managers can alter financial statements in a way that gives an untrue company position. Data mining techniques can decrease the probability for internal fraud, and the multivariate latent clustering technique gained relevant results[10]. Furthermore, neural networks, logistic models and decision trees have been used for internal fraud risk reduction. For instance, the Institute of Internal Auditors (IIA) has listed data mining as one of the four priorities for further research, and the Chartered Global Management Accountant (CGMA) has reported that data mining lies within the top ten focus priorities fundamental for the data-driven era of business and was ranked as the most useful by more than half corporate leaders[13].

3. Methodology

3.1. Research sample

In order to inspect internal fraud we have conducted a case study analysis on the data available from one large company. This company is organized using project-based organizational structure[14, 8, 15]. The company has more than 300 employees and implements and develops business-related software applications. Each month, employees working on a project-basis provide a report on their work including the number of hours, complexity of their work, and the amount claimed for an hour and in total, and based on that report fill in the working-hours claim. The company has already developed its own methods for detecting suspicious working-hour claims, but those are focused on detection of already committed fraudulent activities, while more research is needed in order to identify the characteristics of fraudulent claims in order to detect potential new ones.

The goal of the analysis is to determine the characteristics of the suspected working-hour claims, which are the candidates for an in-depth fraud analysis. The company defines the suspect claims in the following manner. A working-hours claim is suspect in at least one of the following cases: (i) if a consultant is late in submitting the working-hours claim more than seven days from the day when the project finished, and (ii) if a consultant cancels already claimed working-hours. In the case when at least one of the abovementioned criteria is fulfilled, the working-hours claim is considered as suspect for fraud in terms of claiming a too large amount of working-hours and/or incorrect or too high hourly-rate. In other words, a claim is suspected if either case (i) or case (ii) applies. However, the management of the company concluded that it would be beneficial to identify the characteristics of the potentially fraud (suspect) working-hour claims before the consultant is already late in submitting the claim.

Therefore, analysis aims to identify the relationship of the suspect working-hour claims with various characteristic related to the project, such as the type of customer, the type of consultant, the expert level of the consultant, the month when the working-hours were claimed, hourly-rate, the number of hours claimed and the total amount claimed. Our case study uses the sample of 1,194 working-hours claims in a large multinational enterprise, which comprise 5% of the total working-hours claims in the company in the observed year. According to Table 1, 294 working-hours claims, or 24.62%, were suspect for fraud whereas 900 working-hours claims, or 75.38%, were non-suspect for fraud. The variable Suspect defines these two categories of working-hours claims (if the claim is suspected it has value 1, otherwise it is equal to 0).

Table 1 Suspect and non-suspect working-hour claims in the sample

Variable Suspect	Count	Percent
Suspect (value 1)	294	24.62
Non-suspect (value 0)	900	75.38
Total	1,194	100.00

Source: Authors' work, based on the internal data source.

The second variable used in the analysis is the **variable Customer**. According to this variable, customers ordering the work on the project (development and/or implementation of software applications) are divided into three categories: governmental institutions, internal projects, and private enterprises. If the structure of customers according to the Suspect category is observed, it can be concluded that a vast majority of customers are governmental institutions, and private enterprises can be found in the category non-suspected with a share of 95.24% and 77.20%, respectively. On the other hand, internal projects are suspect in 50.68% case, whereas 49.32% are non-suspect. The conducted chi-square test confirmed, at the significance level of 1%, that there is at least one category of customers whose structure according to the variable Suspect is different than the others (chi-square=77.435, df=2, p-value<0.001).

The variable Consultant describes the country of origin of experts, who have been claiming working-hours, since in some cases domestic consultants (from Croatia) and in some cases, foreign consultants are hired. In cases when domestic consultants are observed, 23.46% of their working-hour claims were suspected, while foreign consultants were in 41.56% cases in the suspected working-hours claim category. The chi-square test has shown that, at the significance level of 1%, domestic and foreign employees have a statistically significantly different structure according to suspected and non-suspected working-hours claims (chi-square=12.719, df=1, p-value<0.001).

The variable Month represents the month in which a consultancy service was provided. For the purpose of the analysis, months are coded as discrete values ranging from M1 to M12. The highest share of suspected working-hours claims can be found in months M1 (65.77%) and M12 (30.59%), which refer to January and December. It is highly probable that this large percentage of suspect claims is related to the beginning and the end of the fiscal year. On the other hand, the highest share of non-suspected working-hours claims is in months M10 (88.42%) and M4 (86.40%). According to the conducted chi-square test, those shares seem to be statistically significantly different, at the significance level of 1%, in different months (chi-square=134.670, df=11, p-value<0.001)

The variable UnitPriceCoded was used to take into account the cost of consultants. In the analysis, this cost is expressed per hour. The minimum cost per hour is 19.9 EUR and the highest is 173.9 EUR per hour. Because there are many different values, it has been decided that four groups of costs will be formed, and that the unit price will be coded in four categories (1-50 EUR per hour, 51-100 EUR per hour, 101-150 EUR per hour, and 151-200 EUR per hour). The largest share of suspected working-hours claims was found in the category of the cost of 151-200 EUR (30.00%) whereas the largest share of non-suspected working-hours claims was found in the category of the cost of 1-51 EUR (89.19%). The chi-square test has shown that, at the significance level of 5%, the hypothesis of equal shares of suspected working-hours claims, or non-suspected working-hours claims, at all the four observed cost levels cannot be rejected (chi-square=6.278, df=3, p-value=0.099).

The variable ExpertLevel reflects the five expert levels coded from L4 to L8, which refer to the experience and relevant knowledge of consultants claiming working-hours (L4 is the lowest level of expertise, while L8 is the top level of expertise). The highest share of suspected working-hours claims is present at the expert level L5 (77.78%) whereas the highest share of non-suspected working-hours claims is present at the expert level L4 (89.19%). The chi-square test confirmed that, at the significance level of 1%, there is at least one expert level at which shares of suspected working-hours claims or non-suspected working-hours claims are statistically significantly different than at other expert levels (chi-square=33.147, df=4, p-value<0.001).

The number of weekly working-hours of employees (consultants) is going to be observed as well (the **variable NoHoursCoded**). There is a quite large number of discrete values of weekly working-hours. Consequently, they are classified into eight groups: 1-5; 6-10; 11-15; 16-20; 21-25; 25-30; 31-35; and 36-55. Due to some administrative problems, an additional category was introduced to incorporate negative weekly working-hours, which appeared due to some corrections conducted by consultants themselves. It has been shown that, in the case of employees with negative weekly working-hours, these working-hour claims are treated as suspected. The Chi-square test has shown that, at the significance level of 1%, there is at least one weekly working-hours category at which the share of suspected working-hours claims is statistically significantly different than at other weekly working-hours categories (chi-square=53.859, df=8, p-value<0.001).

The costs of consultants' working-hours are observed by the **variable TotalAmountCoded**. Again, due to a large number of different total costs per consultant, those costs have been categorized into 19 cost categories (starting from 1-100, to 3001-4000 EUR). The conducted chi-square test has shown that, at the significance level of 1%, there is at least one total cost per consultant category at which the share of suspected working-hours claims is statistically significantly different than at other total cost per consultant categories (chi-square=80.068, df=18, p-value<0.001).

3.2. Statistical analysis

In order to provide understanding of the interrelation between working-hours claim fraud and various characteristics, such as characteristics of customers, consultants, expert knowledge and others, a decision tree is developed using the Chi-squared Automatic Interaction Detector (CHAID) algorithm[6]. As the name reveals, the CHAID decision tree is based on the chi-square test, which is used to select the best split at each step. In order to construct a decision tree, the role of the dependent variable was given to the variable Suspect. All other observed variables have taken the role of independent variables (Customer; Consultant; Month; UnitPriceCoded; ExpertLevel; NoHoursCoded; TotalAmountCoded). In order to get a clear and easily understandable classification tree, it has been decided that the classification tree depth should go up to the third level, which is indicated by Bertsimas and Dunn[Error! Reference source not found.], as the optimal depth of the tree. Furthermore, it has been defined that the main or parent node should have at least 100 cases whereas the following or child nodes should have at least 50 cases, which comprise approximately 8% and 4% respectively of the total sample (1,194 cases). The decision tree is developed using SPSS ver. 23.

4. Results

According to defined settings, the CHAID decision tree is developed (Figure 1). The resulting CHAID decision tree has 3 levels and overall 11 nodes out of which seven are considered as terminal (they do not split further). Figure 1 also reveals that variables Month, Customer and ExpertLevel were the most statistically significant and therefore they are used in building the classification tree by the used algorithm.

The variable used for **branching on the first level** is the variable **Month**, which turned out to be statistically significant at the significance level of 1% (chi-square=130.995, df=2, p-value<0.001). This branching resulted in three new nodes (Node 1, Node 2, and Node 3). Node 1 includes categories M3, M4, M5, M6, M7, M10 and M11. That way Node 1 consists of 700 working-hours claims out of which 587 or 83.9% are treated as non-suspected whereas 113 or 16.1% are suspected. Node 2 includes the following categories of the variable Month: M2; M8; M9; and M12. Consequently, Node 2 has in total 383 working-hours claims out of which 275 or 71.8% are non-suspected whereas 108 or 28.2% are suspected. Node 3 includes only the category M1 and only at this node, the share of suspected working-hours claims (65.8%) is greater than the share of non-suspected working-hours claims (34.2%).

The variable **Customer** was used for **branching on the second level**. According to Figure 1, branching resulted in five new nodes with three of them (Node 4, Node 5, and Node 6) coming out from Node 1 and two of them (Node 7 and Node 8) from Node 2. Both branching processes are highly statistically significant (from Node 1 – chi-square=16.976, df=2, p-value<0.001; from Node 2 – chi-square=32.079, df=1, p-value<0.001). Node 4 includes only 67 customers of government institutions out of which 65 or 97.0% are connected with non-suspected working-hours claims and two or 3.0% are connected with suspected working-hours claims. Node 5 consists of 69 customers of internal projects out of which 49 or 71.0% are connected with non-suspected working-hours claims and 20 or 29.0% are connected with suspected working-hours claims. Node 6 includes only customers of private enterprises and it is the largest one among nodes of the second level. There are 473 or 83.9% customers of private enterprises that are connected with non-suspected working-hours claims and 91 or 16.1% that are connected with suspected working-hours claims. On the other hand, Node 7, which is related with Node 2, includes customers of government institutions and customers of private enterprises together. It has been shown that out of 328 customers 253 or 77.1% are non-suspected whereas 75 or 22.9% are suspected for working-hours claim fraud. Node 8 includes only customers of internal projects. When nodes of the second level are observed, it can be concluded that only at this node the share of suspected working-hours claims (60.0) is higher than the share of non-suspected working-hours claims (40.0%).

The third level branching variable is the variable **ExpertLevel**. This variable was used to branch Node 6 further into two new nodes (Node 9 and Node 10). This branching process is statistically significant at the significance level of 5% (chi-square=9.539, df=1, p-value=0.030). Node 6 consists only of consultants with the expert level L6 whereas consultants with levels L4, L5, L7 and L8 can be found in Node 10. Node 9 is considerably larger than Node 10 and includes 431 or 85.5% non-suspected working-hours claims and 73 or 14.5% suspected working-hours claims. Furthermore, it has to be emphasized that Node 9 includes 42.2% of all observed working-hours claims whereas Node 10 includes only 5.0% of them. Therefore, Node 10 includes 60 working-hours claims out of which 42 or 70% are non-suspected whereas 18 or 30.0% are suspected.

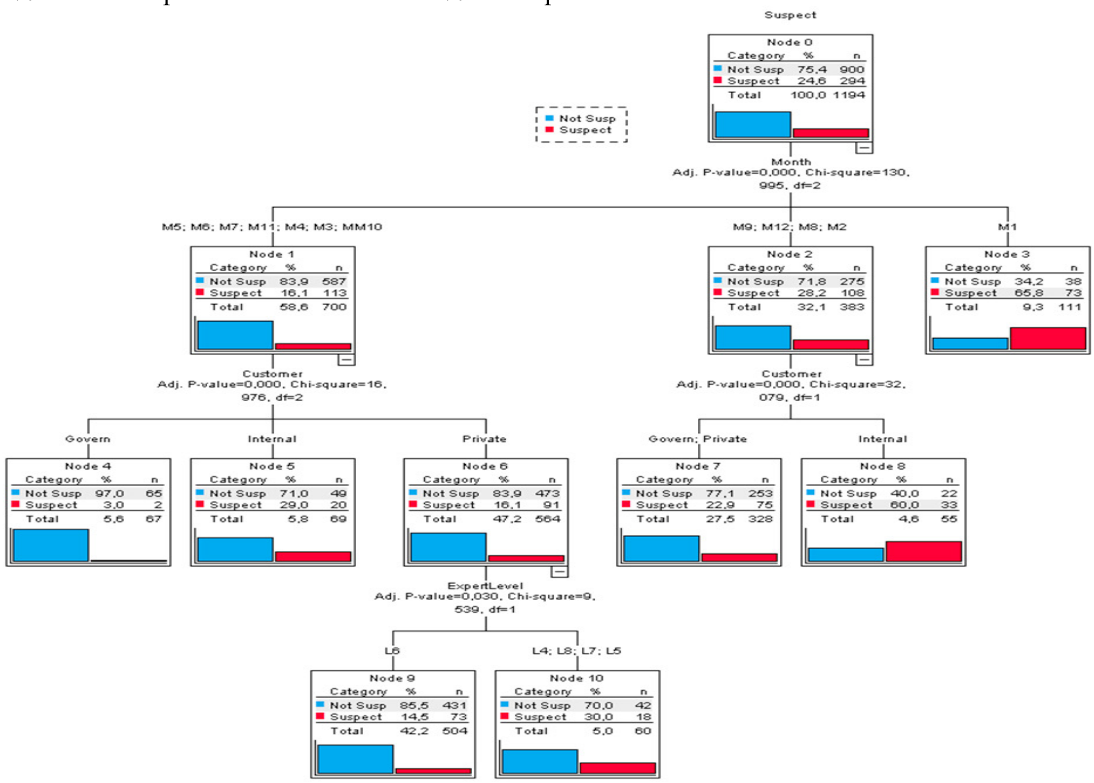

Figure 1 CHAID decision tree
Source: Authors' work, based on the internal data source.

Table 2 Classification matrix for CHAID decision tree

Observed classification	Predicted classification		
	Non-suspect	Suspect	Percent correct
Non-suspect	840	60	93.3%
Suspect	188	106	36.1%
Overall percentage	86.1%	13.9%	79.2%

Note: Growing method: CHAID; Dependent variable: Suspect.

The classification matrix, shown in Table 2, compares the observed and the predicted status of working-hours claims. The used algorithm was correct in 93.3% of cases for the non-suspect working-hour claims. In other words, out of 900 non-suspected working-hours claims, the algorithm has correctly classified 840 of them, whereas 60 working-hours claims were wrongly classified. The successfulness of the algorithm seems to be quite low in relation to suspected working-hours claims. Namely, out of 294 suspected working-hours claims, the algorithm correctly classified 106 working-hours claims, or only 36.1%. Therefore, it can be concluded that the decision tree is recommended for use as a supportive instrument for the detection of working-hour claims, in combination with other human-based and machine-based methods. However, experts from the company confirmed that the information derived from the decision tree is valid to them since it provided a new insight into the characteristics of suspect working-hour claims. This information allows them to focus their efforts on the following categories identified by the decision tree as the most likely to be suspected: working-hour claims submitted in M1 by the internal experts. In addition, the general rate of correct classification of 79.2% can be observed as quite good. Software used for development of the CHAID decision tree can also generate the SQL code, and a part of it is provided in Appendix 1, which can be used for the development of the solution for automatic internal fraud-detection.

5. Conclusions

The purpose of this research was to investigate the usefulness of data mining in internal fraud detection, related to suspected working-hour claims in one project-based company, producing business-related software applications. First, we analyzed the characteristics of the working-hour claims according to the status of the claim (suspect vs. non-suspect). The Chi-square analysis indicated that suspected working-hour claims were in correlation with costumers from private enterprises, with domestic consultants, with cost per hour between 51 and 100 EUR, and with the expert level L6 (medium expert knowledge level). Second, using the CHAID decision tree we aimed to identify the relationships between numerous characteristics of the project (e.g. characteristics of the client and the expert), and suspect working-hour claims. Based on the database of working-hour claims that contained both suspect and non-suspect claims, we developed the CHAID decision tree that provided the general rate of nearly 80% correct classification. The analysis indicated that suspected working-hour claims were related mostly with internal consultants, claimed in the first month of the fiscal year. The decision tree provided a satisfactory overall accuracy, but accuracy of predicting suspect claims was much lower. However, the company management indicated that the gained information, from both the chi-square and the decision tree analysis revealed important information that will allow them to focus on particular characteristics of working-hour claims. The limitation of our work is that the research is based on one case study including only one specific company. Consequently future research should include more companies from different settings, e.g. multinational vs. national, SMEs vs. large companies and others. Overall, it can be concluded that this case study reveals that data mining can be used for detecting internal control fraud, and it confirms the previous research results indicating the overall usefulness of data mining in fraud detection.

References

1. Allen R, Hemming R, Potter B. The International Handbook of Public Financial Management, *Palgrave Macmillan, London* 2013.
2. Baldock G. The perception of corruption across Europe, Middle East and Africa. *Journal of Financial Crime* 2016; **23(1)**: 119-131.
3. Bathallath S, Smedberg Å, Kjellin H. Managing project interdependencies in IT/IS project portfolios: a review of managerial issues. *International journal of information systems and project management* 2016; **4(1):** 67-82.
4. Bertsimas D, Dunn J. Optimal classification trees, *Machine Learning* 2017; **7**: 1039-1082.
5. Bologna J, Lindquist RJ. Fraud Auditing and Forensic Accounting. New York: *John Wiley & Sons,* 1987.
6. Byeon H. Chi-Square Automatic Interaction Detection Modelling for Predicting Depression in Multicultural Female Students. *International Journal of Advanced Computer Science and Applications* 2017; **12**: 179-183.
7. Curley R. Issues in Cyberspace: From Privacy to Piracy. *Britannica Educational Publishing* 2011.

8. Haddadi A, Johansen A, Bjørberg S. Best Value Approach (BVA): Enhancing Value Creation in Construction Projects. *Business systems research journal*, 2017, **8(2):** 84-100.
9. Hobday M. The project-based organization: an ideal form for managing complex products and systems, *Research policy* 2000; **29**: 871-893.
10. Jans M, Lybaert N, Vanhoof K. Data mining for fraud detection: Toward an improvement on internal control systems, *European Accounting Association - Annual Congress* 2007, Lisbon: 30
11. Jans M, Lybaert N, Vanhoof K. Internal fraud risk reduction: Results of a data mining case study. *International Journal of Accounting Information Systems* 2010; **11(1)**: 17-41.
12. Joudaki H, Rashidian A, Minaei-Bidgoli B, Mahmoodi M, Geraili B, Nasiri M, Arab M. Using data mining to detect health care fraud and abuse: a review of literature. *Global journal of health science* 2015; **7(1)**:194.
13. Koh C,Low C. Going concern prediction using data mining techniques. *Managerial Auditing Journal* 2004; **19(3)**: 462-476
14. Larson EW, Gobeli DH. Significance of project management structure on development success. *IEEE transactions on engineering management* 1989, **36(2)**: 119-125.
15. Makki S, Haque R, Taher Y, Assaghir Z, Ditzler G, Hacid MS, Zeineddine H. Fraud Analysis Approaches in the Age of Big Data-A Review of State of the Art. *In Foundations and Applications of Self* Systems (FAS* W), IEEE 2nd International Workshops* 2017: 243-250.
16. Miterev M, Mancini M,Turner R. Towards a design for the project-based organization. *International Journal of Project Management* 2017; **35(3)**: 479-491.
17. Nawawi A, Salin ASAP. Employee fraud and misconduct: empirical evidence from a telecommunication company. *Information & Computer Security* 2018; **26(1)**: 129-144.
18. Phua C, Lee V, Smith K, Gayler R. A comprehensive survey of data mining-based fraud detection research. *arXiv preprint arXiv* 2010: **1009.611**
19. Sørensen JK. Exploring Constrained Creative Communication: The Silent Game as Model for Studying Online Collaboration. *International Journal of E-Services and Mobile Applications (IJESMA)*, 2017, **9(4):** 1-23.
20. Stolfo SJ, Fan W, Lee W, Prodromidis A, Chan, PK. Cost-based modeling for fraud and intrusion detection: Results from the JAM project. *In DARPA Information Survivability Conference and Exposition DISCEX'00 IEEE* 2000; **2**: 130-144.
21. Terzieva, M., & Morabito, V. Learning from experience: the project team is the key. *Business Systems Research Journal*, 2016, **7(1)**: 1-15.
22. Trigo A, Varajão J, Barroso J, Soto-Acosta P, Molina-Castillo FJ, Gonzalvez-Gallego N. Enterprise information systems adoption in Iberian large companies: Motivations and trends. In: Tavana M, editor. Managing Adaptability, Intervention, and People in Enterprise Information Systems. Hershey, PA: Information Science Reference; 2011, 204-28.
23. Zakaria KM, Nawawi A, Salin ASAP. Internal controls and fraud–empirical evidence from oil and gas company. *Journal of Financial crime* 2016; **23(4)**: 1154-1168.

Appendix 1.

SPSS equations generated for the implementation of the CHAID decision tree for Node 3, Node 4, and Node 5

```
/* Node 3 */.
DO IF (Month EQ "M1").
COMPUTE nod_001 = 3. COMPUTE pre_001 = 'Suspect'. COMPUTE prb_001 = 0.657658.
END IF.
EXECUTE.

/* Node 4 */.
DO IF (Month NE "M9"  AND  Month NE "M12"  AND  Month NE "M8"  AND  Month NE "M2"  AND  Month
NE "M1")  AND  (Customer EQ "Govern").
COMPUTE nod_001 = 4. COMPUTE pre_001 = 'Not Susp'. COMPUTE prb_001 = 0.970149.
END IF.
EXECUTE.

/* Node 5 */.
DO IF (Month NE "M9"  AND  Month NE "M12"  AND  Month NE "M8"  AND  Month NE "M2"  AND  Month
NE "M1")  AND  (Customer EQ "Internal").
COMPUTE nod_001 = 5. COMPUTE pre_001 = 'Not Susp'. COMPUTE prb_001 = 0.710145.
END IF.
EXECUTE.
```