Business Process Anomaly Detection using Multi-Level Class Association Rule Learning

Fernandes Sinaga[1], Riyanarto Sarno[1]

***Abstract* - Recently, Business Process Management System (BPMS) is widely used by companies in order to manage their business process. The company's business process has a possibility to have changes which can cause some variations of business process. These variations might be contain some anomalies. Any anomalies that can make some losses for the company can be regarded as a fraud. There were some research have done to detect anomalies in business process. But, there is some issues that still need improvement especially on the accuracy. This paper proposed Multi-Level Class Association Rule Learning method (ML-CARL) to detect business process anomalies accurately. This method is supported by the process mining method which is used to analyze the anomalies in process. From the experiment, ML-CARL method can detect anomalies with an accuracy of 0.99 and better than ARL method in previous research. It can be concluded that ML-CARL method can increase the accuracy of business process anomaly detection.**

***Term Index* - Business process, Anomaly detection, Process mining, Multi-level class association rule learning**

INTRODUCTION

Some companies in the world have used the business process management system such as BPMS, Enterprise Resource Planning, etc. The goal is to control and manage their business process. Company's business process can be changed along with the market, the requirement changes, and the policy changes. These changes can make some variations of business process. There is a possibility that there are anomalies in those process variations [1]. These anomalies can cause some losses for the company so it can be regarded as fraud. [2]. Fraud is done without consider to the goal and the principles of the company.

Fraud is a widespread problem in the world. In 96 countries, there are 1,388 fraud caused losses of up to 1,4 billion US Dollars. [3]. Fraud could happen because of anomalies to business process standard and data manipulation [4]. Fraud could be defined as crimes that use deception as a major modus operandi and include various aberrations by individuals or organizations [5]. In order to reduce the losses, fraud detection techniques are needed.

In computer science, there were two analysis techniques have been done to detect fraud, namely data mining and process mining. *Decision Tree, Neural* and *Bayesian Network*, and *Support Vector Machine* were examples of data mining technique which had done by the previous research to detect fraud in process [6], [7], and [8]. However, these methods have limitations in detecting anomalies because these methods were not able to analyze the behavior of process control flow. Furthermore, process mining could detect anomaly in process with conformance checking. Conformance checking is not only a process mining technique that compare the actual process data and the standard process model but also could analyze the process control. In the context of fraud detection, any anomalous parts were considered as a compromising fraud [2].

Other researches which support the fraud detection was using Association Rule Learning (ARL). There were two research which had used ARL. First, research of fraud detection which applied to credit card application in the retail company in Chile [9]. This research focused on mining data in the form of association rules to detect fraud. The second research was fraud detection on business process of credit application [10]. This research had combined process mining and ARL so they could detect fraud with an accuracy of 0,865. But, there were still a high value of false positive and false negative.

This paper will propose a Multi-Level Class Association Rule Learning (ML-CARL) method to detect fraud accurately. The main goal is to reduce the number of the false positive and false negative in order to increase the accuracy. This method is used because of two reason. First, multi-level association analysis is used to find the hidden information in or between levels of abstraction. Second, classification association rule is used to find association rules efficiently according with user's need. So, the goal of this method is to gain more knowledge from the anomalies data in order to produce association rules effectively and efficiently. This method is supported by conformance checking technique to analyze anomalies in the process and fuzzy multi attribute decision making to calculate the rating of fraud for each process.

Concordance. This study assesses the concordance of the demand and suplyof public transport route by looking at the original matrix trip destinations, the provision of the existing public transports routes, demand for publict transport service, and assessment concordance between the demand and supply of public transport routes.

METHOD

We are analyzing business process of a credit application in bank with process mining to detect any anomalies in the process. Then, we use fuzzy multi-attribute decision making to calculate fraud's rate of each case (instance process). And finally, we mining the association rules of anomaly from the anomalies data correspond to their fraud's rate using Multi-Level Class Association Rule Learning (ML-CARL).

[1]Fernandes Sinaga and Riyanarto Sarno are with Departement of Informatics, Faculty of Information Technology, Institut Teknologi Sepuluh Nopember, Surabaya. Email : fernandes.sinaga10@mhs.if.its.ac.id; riyanarto@if.its.ac.id

In ML-CARL, there are two steps to mining the rules. First, classify the cases into some classes which were defined. There are three classes, Non-Fraud class, Semi Fraud Class and Fraud Class. We classify using a fuzzy membership function. We use the fraud's rate of each case as a parameter for the classification. And the second step is mining rules using multi-level association rule. From this ML-ARL method, we generate some association rules which could represent anomalies in process effectively.

EXPERIMENTAL DESIGN

The evaluation in this research focuses on measuring the accuracy of the ML-CARL methods. The experiment has done to a case study of business process in bank credit application. The variable in dataset is consisted of cases or transactions, and 10 anomalies attributes. The dataset is divided into training dataset and testing dataset which generated by two distribution models like in [10]. The first distribution model is Poisson distribution. We use this model to generate the number of cases of anomaly of each attributes randomly. And the second distribution model is uniform (discrete) distribution. We use this model to spread over the anomalies in 50 cases each month randomly and based on the number of anomaly occurrences for each attribute.

We generate 1200 cases were divided into training data and testing data. There are 1000 cases for training data while testing data has 200 cases. In training data, there are 20 cases of fraud, 14 cases of semi fraud and 966 cases of non-fraud. In testing data, there are 5 cases of fraud, 3 cases of semi fraud and 192 cases of non-fraud

EXPERIMENTAL RESULT AND DISCUSSION

From the training using ML-CARL, we generates 24 association rules. Then, we test the testing data using this 24 rules. We get a True Positive/Fraud (TP) value of 5, a False Positive/Fraud (FP) value of 0, True Semi Positive / Semi Fraud (TSP) value of 2, False Semi Positive / Semi Fraud (FSP) value of 1, True Negative / Non-fraud (TN) value of 191 and False Negative / Non-Fraud (FN) value of 1. Then, we use the accuracy measurement to test the performance of this method. The accuracy of this ML-CARL method is 0.99. This accuracy is better than the accuracy of ARL method in previous research [10].

CONCLUSION

From the experiment, we can conclude that the ML-CARL method can detect anomalies in business process well and accurately. This is caused by the generated association rules can describe anomalies in business process effectively and efficiently. Furthermore, conformance checking can help in analyzing the anomalies in the process. So, the combination of the ML-CARL method and the conformance checking analysis can increase the accuracy of business process anomaly detection.

REFERENCES

[1] R. Sarno, A.B. Sanjoyo, I. Mukhlash and H.M. Astuti, "Petri Net Model of ERP Business Process Variations for Small and Medium Enterprises," Journal of Theoretical and Applied Information Technology, vol. 54 No.1, August 2013, pp. 31-38.

[2] J. Stoop, "A case study on the theoretical and practical value of using process mining for the detection of fraudulent behavior in the procurement process," in Process Mining and Fraud Detection, Netherlands, Twente University, 2012, pp. 22-63.

[3] "Report to the Nations on Occupational Fraud and Abuse," ACFE, 2014, p.19.

[4] M. Jans, N. Lybaert, K. Vanhoof, and J. M. van der Werf, "A business process mining application for internal transaction fraud mitigation," Expert Systems with Applications, vol. 38, 2011, pp. 13351-13359.

[5] Wells, J.T. Occupational Fraud and Abuse. Dexter, MI: Obsidian., 1997, p.221.

[6] F. Ogwueleka, Data Mining Application in Credit Card Fraud Detection System, Nigeria: Department of Computer Science, University of Abuja, 2011, pp.311-322

[7] Ngai, E.W.T.,Yong Hu, Wong, Y.H., Chen Y.,Sun, X, "The application of data mining techniques in financial fraud detection: A classification framework and an academic review of literature," Decision Support Systems, vol. 50, 2011, pp. 559-569.

[8] Bhattacharyya, S., Sanjeev J., Tharakunnel, K., Westland, J.C, "Data mining for credit card fraud: A comparative study," Decision Support Systems, vol. 50, 2011, pp. 602-613.

[9] D. Sanchez, M. Vila, L. Cerda, and J. Serrano, "Association rules applied to credit card fraud detection," Expert Systems with Applications, pp. 3630–3640, 2009.

[10] R. Sarno, R. D. Dewandono, T. Ahmad, M. F. Naufal and F. Sinaga, Hybrid Association Rule Learning and Process Mining for Fraud Detection, IAENG International Journal of Computer Science, 2015, pp. 59-72