paper

by Solichul Huda

Submission date: 23-Aug-2017 04:07PM (UTC+0700)
Submission ID: 839143719
File name: p84-huda.pdf (585.99K)
Word count: 3379
Character count: 18124

2014 2nd International Conference on Information and Communication Technology (ICoICT)

Identification of Process-based Fraud Patterns in Credit Application

Solichul Huda
Informatics Department
Universitas Dian Nuswantoro
Semarang, Indonesia
solichul.huda@dsn.dinus.ac.id

Riyanarto Sarno
Informatics Department
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
riyanarto@if.its.ac.id

Tohari Ahmad
Informatics Department
Institut Teknologi Sepuluh Nopember
Surabaya, Indonesia
tohari@if.its.ac.id

Heru Agus Santoso
Informatics Department
Universias Dian Nuswantoro
Semarang, Indonesia
herezadi@gmail.com

Abstract—**Fraud detection has become an important research topic recently. In a credit application, fraud can occur in forgery of documents or business processes. Fraud on the business process is known as Process-based Fraud (PBF). Previous studies proposed several detection methods of fraud in the business process model. This fraud detection includes analysis methods and an identification process. However, none of them proposed PBF identification, particularly identification of PBF attributes and pattern clearly, so its accuracy still needs further improvement. As identification of PBF attributes and PBF pattern is very important for the accuracy of PBF detection, this paper proposes an identification method for PBF detection. This PBF identification process consists of some attributes, those are skip sequence, skip decision, throughput time minimum, throughput time maximum, wrong resource, wrong duty decision, wrong duty sequence, wrong duty combine, wrong pattern and wrong decision. In this paper, PBF pattern is combined with a fuzzy set which consists of low, middle and high categories. This fuzzy set is implemented in order to improve the accuracy of PBF determination. PBF attribute and its pattern contribute to the process mining for detecting PBF.**

Keywords : Process-Bases Fraud; PBF attribut; PBF Pattern; Fuzzy set; fraud detection;

I. INTRODUCTION

Fraud covers all forms of crime with the intention of taking advantages in the various modes of deception [1]. The Association of Certified Fraud Examiners (ACFE) reported that an organization loses approximately 5% of their annual income. In 2011 itself, fraud costs more than 3.5 trillion U.S. dollars [2]. These losses have significant effect both on small companies and large corporations. Fraud classified into 3 types, namely Asset Misappropriation Fraud (Fraud against assets), Fraudulent Statements (Fraud to the financial statements) and Corruption.

Fraud can occur because of the violation of the business process / standard operating procedures and data manipulation [3], [4]. In order to analyze the data conversion, data mining technique is used. In addition, for analyzing the violation in the business processes, the process mining is employed. Study of fraud detection using data mining technique has been widely applied by using the neural network algorithm [5], [6], [7], logistic regression algorithm [8], [9] and fraud detection with Web Service Collaboration [10]. Fraud detection with process mining is performed by using fuzzy miner and heuristic is proposed in [3], [4], while Association Rule Learning is used in [11]. Process-based Fraud (PBF) is fraud occuring in a business process [3]. In the last three years, the amount of Process-based Fraud (PBF) in the existing total fraud is relatively large in Indonesia. However, the number of research on Process-based Fraud is still small. Including works in this field are Mieke Jans [3], JJ Stoop [4] and Rahardian [11].

Mieke Jans [3] proposes to use control flow analysis, performance analysis and role analysis to analyze event logs. Event logs is log that contains information about activities in a business process model. Information about events, originator, time stamp and case id are contained in event logs. Control flow analysis is used to analyze the pattern of the event log. The case whose fitness only small is regarded as a noise. This noise is then identified as suspicious fraud.

978-1-4799-3580-2/14/$31.00 ©2014 IEEE

Performance analysis method is used to analyze a case that has the shortest path. It is identified as suspicious fraud. Role analysis is used to analyze case in which violation of segregation of duties happens. A case with some events that are executed by the same originator, identified as suspicious fraud. Nevertheless, the process of making decision whether it is fraud or not, requires opinions from experts. Mieke Jans [3] has not presented the implementation of fraud detection techniques in business processes. J.J. Stoop [4], designed how to implement PBF fraud detection in business processes. He introduced the concept of 1+5+1. Concept of 1+5+1 itself consists of (1) creating event logs, (1) analysing log, (2) analysing processes, (3) analysing conformance, (4) analysing performance, (5) analysing social, and (1) iterating and refocussing. The implementation of the concept is as follows :

- Log analysis is done by using filtering, aging, analysing gap, analysis duplicate and summarizing data .
- Process analysis is performd by using analysing gap, filtering, analysing duplicate, aging, summarizing data, analysing statistical, joins and analysing trend .
- Conformance analysis is implemented by filtering the data, analysing gap, analysing duplicate, summarizing data and parallel simulation.
- Performance analysis is carried out by using the data filtering, summarizing data, stratification, sorting data, analysing statistical, analysing trend and aging.
- Social analysis is implemented by using aging, filtering, summarizing data, sorting data and joins.

The concept of 1 +5 +1 describes the sequence of PBF detection. It is implemented on an offline application, therefore, it can only be used by auditors. Furthermore, these studies only produced suspicious PBF.

Rahardian [11], proposes to use Association Rule Learning (ARL) to determine PBF. ARL method is used to analyze a case which has been executed. However, the data used in the simulations have not been tested [11], and the object of study, particularly the business process models is still simple. In addition, a case which has violation of the process model has determined as PBF definitively. Futhermore, PBF attributes are minimal and there are no pattern of PBF. PBF pattern is very important to be the standard determination of PBF. We believe that complete PBF attributes along with PBF pattern table will significantly improve the accuracy of PBF detection.

In this study, we will examine the methods of PBF detection in a credit application, whether it can determine PBF more accurate than previous studies. Futhermore, we will also develop applications of PBF detection in credit applications for its mitigation.

In The first stage of this study, we will identify the attributes of PBF and PBF pattern. PBF pattern identification has not been done by previous research. In this paper, the complete PBF attributes and PBF pattern are presented. The PBF pattern will be used as the basis for further research.

The Structure of this paper consists of 6 sections. Those are: introduction, basic concept of PBF, attribute identification and pattern of PBF, case study, conclusion, and future works.

II. Basic Concept

A. ProcessMining

Process mining is a method used to extract knowledge from event logs which contains records of activities executed by the originator in a case [12]. In the event logs, there is information about event name, time of execution and the name of the originator who execute the process. A case series of events, starting from the start activity to finish activity. This series of events, can be used to build a model of business processes. All order flow of events in case, will be seen in the process model. Each case has fitness value in the process model.

B. Conformance Analysis

To analyze the number of cases in a workflow process model can be used to conformance techniques. Conformace techniques analyzes the case in business process model. a case whose fitness value is smal means that it violates the case flow in the process model. Violation of the business process model is identified as a Process - Based Fraud (PBF). The business process model illustrates standard operating procedure (SOP). Thus, PBF is the case that violate the SOP. Violation of the SOP itlsef has different levels.

Minor violations of the SOP, is likely to occur by accident. In contrast, a large number of violations of the SOP indicate PBF. In this case, PBF detection accuracy is necessary to avoid misidentification. The other reason is that the attributes and patterns of PBF can be employed to determine the accuracy of PBF by carrying out queries in the table of PBF pattern.

III. PBF Attributes Identification and PBF Pattern

A. PBF Attibutes Identification

Based on the PBF detection method which consists of control flow, performance and role analysis as studied in [3], we find 4 PBF identity attributes, i.e., skip, wrong pattern, throughput time and wrong duty. Whereas in [11], we identified 2 PBF attributes, i.e., wrong decision and wrong resource. There is a distinction between event sequence and event decision. Event sequence is an event that has a relationship with the next event, as described in Figure 1, while the decision event is an event followed by event branching as shown in Figure 2. Events contained in the one

branching have different weights from that with more than one branch.

Figure 1. Event sequence

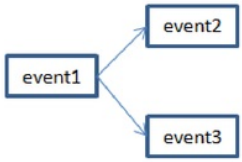


Figure 2. Event decision

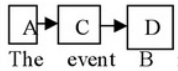
Furthermore, we split a set of PBF attributes as discussed above. So, the attributes of PBF we propose are skip sequence, skip decision, throughput time minimum, throughput time maximum, wrong resource, wrong duty decision, wrong duty sequence, wrong duty combine, wrong pattern and wrong decision as shown in Figure 3 below:

Skip		Throughput Time		Wrong	Wrong duty			Wrong	Wrong
sequence	decision	Min	Max	resource	Sequence	Decision	Combine	Pattern	Decision

Figure 3. PBF Attributes

Description of PBF attributes are described in Table 1 below:

TABLE 1. DESCRIPTION AND EXAMPLE OF PBF ATTRIBUTES

Attribut	Description	Example
Skip	The process jumps from the actual process sequence	 The event B is skipped
Wrong pattern	A case pattern is different with business process models	 Process model pattern  The case pattern
Throughput time min	Event execution time is smaller than minimum of even standard time	Event name: Cek SID; Standard time: 10 minute. Execution time of SID event takes only 2 minutes, instead of 10 minutes
Throughput time max	Event execution time is greater than maximum of event standard time	Execution SID event takes 30 minutes
Wrong resource	Event is executed by an unauthoried originator	Loan decision event is executed by Jack (operator); in facts, it should be done by Winda (manager)
Wrong duty	Different events are executed by a same originator	Both SID and Loan decision events are executed by Jack
Wrong decision	The process of decision making does not comply with SOP	A Rp. 600.000.000 credit is approved by the head of the section; however, according to SOP, a more than Rp. 500.000.000 credit must be approved by the head office.

The value of PBF attribute is filled with the accumulation of violation as follows:

- Skip sequences with the accumulation of skip in the event sequence,
- Skip decision with the accumulation of skip value in the event decision,
- Throughput time min with throughput time which is smaller than throughput time standard event,
- Throughput time max with the accumulation of throughput time which is greater than throughput time standard event,
- Wrong resource with the accumulation of events executed by illegal originators.
- Wrong sequence duty attribute is filled with the accumulation of violations of segregation in duty sequence,

- Wrong duty decision attribute with accumulation of violations of segregation of duty in the event the decision,
- Wrong duty combine attribute with violation segregation of duty on decision events and sequence event,
- Wrong pattern attribute is filled with the accumulation number of events that violate the pattern,
- Wrong decision is filled with decision-making errors.

Figure 4, shows the accumulation value for each PBF attribute of PBF pattern.

Fraud	Skip		Throughput Time		Wrong	Wrong duty			Wrong	Wrong
Code	sequence	decision	Min	Max	resource	Sequence	Decision	Combine	Pattern	Decision
F001	2	0	2	0	0	0	0	0	0	0
F002	2	0	0	18	0	0	0	0	0	0
F003	8	0	0	0	0	0	0	0	0	0
F004	14	0	3	0	0	0	0	0	0	0
F005	14	0	0	3	0	0	0	0	0	0
.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....

Figure 4. PBF attributs with its variables

B. *Fazzyfication Variable*

In the above analysis, we obtained 10 PBF attributes. In this study, the business process model consists of 24 events, so the variable range is 0-24. Therefore, the combination of attribute and variable contents obtained are 10 ^ 24 records. Since the number of records is too much, it is difficult to analyze PBF. With this consideration, a fuzzy set is used as the value of PBF attributes which consists of low, middle and high category. The fuzzy membership value is shown figure 5.

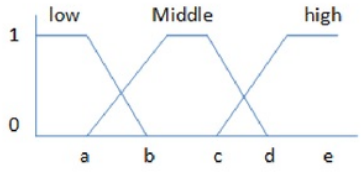


Figure 5. The fuzzy membership

The membership function is defined as follows :

$$\mu_{Low}[Y]=\begin{cases}1 & ;Y<a\\ \dfrac{b-Y}{(b-a)} & ;a<Y<b\\ 0 & ;Y>b\end{cases} \quad (1)$$

$$\mu_{Middle}[Y]=\begin{cases}1 & ;b<Y<c\\ \dfrac{Y-a}{(b-a)} & ;a<Y<b\\ \dfrac{d-Y}{(d-c)} & ;c<Y<d\\ 0 & ;Y>d;\ Y<a\end{cases} \quad (2)$$

$$\mu_{High}[Y]=\begin{cases}0 & ;Y<c\\ \dfrac{Y-c}{(d-c)} & ;c<Y<d\\ 1 & ;Y>d\end{cases} \quad (3)$$

After fuzzification, the combination of attributes and variables obtained are 3 ^ 10 records. Figure 6, shows the pattern of PBF with a set of fuzzy variables.

Fraud	Skip		Throughput Time		Wrong	Wrong duty			Wrong	Wrong
Code	sequence	decision	Min	Max	resource	Sequence	Decision	Combine	pattern	Decision
F001	Low	Null	Low	Null	Null	Null	Null	Null	Null	Null
F002	Low	Null	Null	High	Null	Null	Null	Null	Null	Null
F003	Middle	Null	Low	Null	Null	Null	Null	Null	Null	Null
F004	High	Null	Low	Null	Null	Null	Null	Null	Null	Null
F005	High	Null	Null	Low	Null	Null	Null	Null	Null	Null
.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....

Figure 6. PBF Pattern with fuzzy variable

IV. CASE STUDY

In a credit application SOP, there are 24 events, consists of 16 events sequences and 8 events decision. There is a proposal for new credit application. The originator then proceeds the application beginning from the acceptance of the document to the decision whether the application is rejected or accepted. All events / activities are recorded in the event logs. Next, those recorded events are analyzed to find existing violations. In this study, case analysis is carried out by the PBF detection method as follow :

A. *Log Analysis*

Log analysis is performed to determine the possibility of damaged or double logs The results of this analysis is a fixed event log. On the other hand, the process analysis conducted to obtain an overview of business process model. There are several algorithms that can be used to build a model of the process, including heuristic and fuzzy algorithms .

B. *Conformance Analysis*

We create both business process model and case. After forming a process model, we conduct conformance analysis. Conformance analysis itself requires the case to be converted into a graph. The case will be comformed with the process model which is conducted with process models. The results of conformance will then be analyzed to find the number of

Skip		Throughput Time		Wrong	Wrong duty			Wrong	Wrong
sequen	decision	Min	Max	resource	Sequence	Decision	Combine	Pattern	Decision
13	0	1	0	0	0	0	0	0	0

Figure 7. Accumulation of violation case

skip occurrence. Skip in the events that occurred will accumulated in the variable skip, skip decision / sequence. From the first analysis, we obtained 13 violate skip event sequences, and indicated that the case is suspicious PBF

C. Performance Analysis

Performance analysis is done to analyze the throughput time event. Each event is calculated by subtracting the value of time stamp end along with time stamp start. Then, the results are compared with the timestamp of standard event. Accumulated value of throughput time min will increase if the value of time stamp case is smaller than the time stamp standard. In contrast, the value of throughput time max will increase if througput time of case is bigger than throughput time standard. From the second analysis of the case, we obtained one violate througput time min, and indicate that case was a suspicious PBF.

D. Segregation of Duty Analysis

Segregation of duty analysis is carried out to analyze the occurrence of violations of segregation of duty within the company. Analysis performed by originator that executes more than one event. The results of this analysis will fill wrong duty sequence, decision or combine attribute. From the third analysis of the case, it is found that the process is not violating segregation of duty. Each event is executed by a different originator.

E. Role Analysis

Role analysis, is used to analysis originator. From this analysis, it can be inferred that it is not violating originator authority. All originators executed event comfort with SOP.

F. Decision Analysis

Decision made will be compared with the company SOP. The case flow of decision is not in accordance with the SOP, the value of wrong decision attribute will accumulated.

G. Pattern Analysis

Pattern case is compared with the pattern of the process model. If a new pattern emerged, means pattern case is different from the pattern of existing business process model. This indicates a wrong pattern.

From the analysis of the case above, we obtained 13 violate skip sequences and 1 throughput time min. The results of the analysis is shown in Figure 7.

Violation case is converted into a set of fuzzy as depicted in Figure 8.

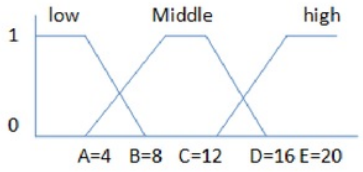


Figure 8. Set of fuzzy membership

In the fuzzification case, the value of A, B, C, D, E are determined as follows. The value of A = 4, B = 8, C = 12, D = 16 and E = 20. Figure 7 shows the violating case in two attributes, namely skip sequences and throughput time min. The Skip sequence and throughput time min value obtained is then converted into a fuzzy set use (2) and (3). With (2)

$\mu_{Middle}[13] = 16 - 13/\ 16\text{-}12$

$= 3/4 = 0.75$

With formula (3)

$$\mu_{High}[Y] = \begin{cases} 0 & ; Y < c \\ \dfrac{Y-c}{\{d-c\}} & ; c < Y < d \\ 1 & ; Y > d \end{cases}$$

$M_{High}[13] = 13\text{-}12/16\text{-}12$

$= 1/4 = 0.25$

Hence, the value of middle skip sequnce is 0.75 and high is 0.25. It is inferred that the value of skip sequence tends to be middle. Meanwhile, since throughput time min is smaller than A. So, by using (1), its membership considered to be low. Figure 9 shows the value of the case attribute with a fuzzy set.

Skip		Throughput Time		Wrong	Wrong duty			Wrong	Wrong
sequen	decision	Min	Max	resource	Sequence	Decision	Combine	pattern	Decision
High	Null	Low	Null	Null	Null	Null	Null	Null	Null

Figure 9. Case attributes with fuzzy set

Fraud	Skip		Throughput Time		Wrong	Wrong duty			Wrong	Wrong
Code	sequence	decision	Min	Max	resource	Sequence	Decision	Combine	pattern	Decision
F004	High	Null	Low	Null	Null	Null	Null	Null	Null	Null

Figure 10. PBF type of case

In order to determine the type of PBF in the case, the query on the table PBF is carried out by making a query as follows:

Select Fraud

Where (skip_sequence ="high" and Through_time_Min = "low")

The result is shown that the case is a PBF with type F004, depicted in Figure 10.

V. CONCLUSION

To be able to detect a Process-based Fraud (PBF) on all models of business processes, it is required a PBF table pattern. This table pattern of PBF requires a set of PBF attributes. This research proposes 10 PBF attributes, i.e., Skip sequence, skip decision, throughput time minimum, throughput time maximum, wrong resource, wrong duty decision, wrong duty sequence, wrong duty combine, wrong pattern and wrong decision. The attributes of PBF describe identity of PBF which along with PBF pattern, contribute to the detection method of PBF.

VI. FUTURE WORKS

This paper presents a preliminary study to detect PBF on credit applications. Our future studies will examine the query over table pattern of PBF. In addition, we will design a more effective pattern of PBF with the domain expert. It will make it easier for companies to use our proposed approach for mitigating fraud that occurred in the credit application.

REFERENCE

[1] ACFE *(Association of Certified Fraud Examiners)*. (2013, Oct.). Fraud Resources . [Online]. Available : http://www.acfe.com/fraud-101.aspx.

[2] ACFE (*Association of Certified Fraud Examiners*). (2013, Oct.), 2012 Report to the Nation. [Online]. Available : http://www.acfe.com/rttn.aspx

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

[4] J.J. Stoop, "Process Mining and Fraud Detection", Twente University, Thesis, 2012.

[5] S. Maes, K. Tuyls, B. Vanshoenwinkel, and B. Manderick, "Credit Card Fraud Detection Using Bayesian and Neural Networks", In:Maciunas RJ, editor. Interactive image-guided neurosurgey. American Association Neurological Surgeons, pp. 261-270, 1993.

[6] V. Zaslavsky, and A. Strizhak, "Credit Card Fraud Detection Using Self-Organizing Maps", Information & Security, vol.18, pp. 48-63, 2006.

[7] S. Panigrahi, A. Kundu, S. Sural, and A.K. Majumdar, "Credit card fraud detection: A fusion approach using Dempster–Shafer theory and Bayesian learning", Information Fusion, vol 10, issue 4, pp. 354–363, 2009.

[8] A. Shen, R. Tong, and Y. Deng, "Application of Classification Models on Credit Card Fraud Detection", Proceedings of IEEE International conference on service system and service management, pp. 1 – 4, 2007.

[9] M. Chae, S. Shime, H. Cho, and B. Lee, "An Empirical Analysis of Fraud Detection in Online Auctions: Credit Card Phantom Transaction", Proceedings of the 40th Annual Hawaii International Conference on System Sciences, pp. 155a, 2007.

[10]C. Chiu, and C.Y. Tsai, "A Web Services-Based collaborative Scheme for Credit Card Fraud Detection", proceedings of the IEEE International Conference on e-Technology, e-Commerce and e-Service, pp. 177-181, 2004.

[11] D.R. Dewandono, "Process Sequence Mining For Fraud Detection Using CEP", Thesis, Institut Teknologi Sepuluh Nopember, 2013.

[12] R. Sarno, A.B. Sanjoyo, I. Mukhlash, and M.H. 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, pp. 31-38, August. 2013.

paper

GRADEMARK REPORT

FINAL GRADE

/0

GENERAL COMMENTS

Instructor

PAGE 1

PAGE 2

PAGE 3

PAGE 4

PAGE 5

PAGE 6

paper

ORIGINALITY REPORT

2%	%	%	2%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to University of Sunderland Student Paper	2%
2	Submitted to Liverpool John Moores University Student Paper	<1%

Exclude quotes	On	Exclude matches	< 1 words
Exclude bibliography	On		