

Junal Pertama

by Solichul Huda

Submission date: 23-Aug-2017 03:55PM (UTC+0700)

Submission ID: 839142402

File name: jurnalpertama.pdf (286.06K)

Word count: 5869

Character count: 30376



Fuzzy MADM Approach for Rating of Process-Based Fraud

Solichul Huda¹, Riyanarto Sarno² & Tohari Ahmad²

¹Informatic Engineering Department, Faculty of Computer Science, Universitas Dian Nuswantoro, Jalan Nakula I No. 5-11, Semarang, Indonesia

² Department of Informatics, Institut Teknologi Sepuluh Nopember (ITS) Surabaya, Kampus ITS Keputih, Sukolilo, Surabaya, 60111, Indonesia
Email: solichul.huda@dsn.dinus.ac.id

Abstract. Process-Based Fraud (PBF) is fraud enabled by process deviations that occur in business processes. Several studies have proposed PBF detection methods; however, false decisions are still often made because of cases with low deviation. Low deviation is caused by ambiguity in determining fraud attribute values and low frequency of occurrence. This paper proposes a method of detecting PBF with low deviation in order to correctly detect fraudulent cases. Firstly, the fraudulence attributes are established, then a fuzzy approach is utilized to weigh the importance of the fraud attributes. Further, multi-attribute decision making (MADM) is employed to obtain a PBF rating according to attribute values and attribute importance weights. Finally, a decision is made whether the deviation is fraudulent or not, based on the PBF rating. Experimental validation showed that the accuracy and false discovery rate of the method were 0.98 and 0.17, respectively.

Keywords: *fraud detection; fuzzy MADM; PBF rating; process-based fraud; process deviation; weighting attributes.*

1 Introduction

Fraud is a form of crime that takes profit from various modes of cheating. Fraud has become a significant concern because it is a major cause of loss in organizations and companies [1]. It is estimated that fraud causes a loss of about 5% of their annual income. Fraud has resulted in a loss of more than 70 trillion dollars [2]. These losses significantly affect almost all companies.

Companies potentially suffer more financial losses because their anti-fraud protection is not able to detect all cases of fraud. It may be possible to detect fraud if the early warning system works well. For example, a deviation in standard operating procedures (SOP) committed by a staff member can be detected early, so that the company can modify the staff member's work pattern to reduce the possibility of fraud. In such cases, process mining provides a

solution by giving an option to test the conformity of the business process to the SOP [3].

Data mining and fraud detection have been studied for decades in various ways. For example, using a neural network algorithm [4], a self-organizing maps algorithm [5], the Dempster-Shafer theory and Bayesian learning algorithms [6], classification models [7], empirical analysis [8], and web service collaboration [9]. Additionally, in process mining it has been done using control flow analysis, role analysis and performance analysis [3],[10], association rule learning [11], hybrid ARL and process mining [12].

These previous researches, however, only considered a non-fuzzy condition (i.e. fraud or not fraud) in detecting PBF. Here, PBF is detected based on SOP deviation, although in reality not all SOP deviations are fraud, as argued by experts. As a result, PBF may not be determined by SOP deviation alone. PBF detection using multi-attribute decision making (MADM) leads to similar results as with the previous methods. Therefore, we propose a fuzzy approach to investigate the degree of membership of attribute values and attribute importance weights. We hypothesize that the degree of membership of attribute values and attribute importance weight can provide the weight of deviation. Finally, the weight of deviation can be used to determine whether the deviation indicates fraud or not. Based on this, we believe that the fuzzy approach is appropriate to overcome fraud detection problems in cases of low deviation.

The rest of this paper is organized as follows. Section 2 presents an overview of related work on PBF detection. Section 3 presents the process mining for PBF detection. Section 4 presents a case study, explaining the business process in a credit application. Section 5 presents the method proposed in this study. Section 6 describes the proposed method for determining PBF. Section 7 contains an evaluation of the proposed method and discussion of its performance. Lastly, concluding remarks according to the results of the proposed method are given in Section 8.

2 Related Work

Fraud detection is important to minimize losses caused by fraud in companies [1]. It should be identified in business processes that can be analyzed by process mining, including performance, event sequence, control flow and role analysis [3]. Detection is performed using data mining (i.e. association rule learning) and a combination of data mining and process mining (hybrid method), the results of which are then analyzed based on the respective business processes to identify SOP deviation [11]-[13].

In [3], the authors have proposed process mining to mitigate fraud. They used performance analysis, control flow analysis, and role analysis to study business processes. This method, however, does not include an algorithm for fraud detection. It has been proved that process mining is able to detect fraud in business processes.

The concept of 1+5+1 [10] proposes tools for implementation of PBF detection. “1+5+1” stands for: (1) log preparation + (5) {1} log analysis, {2} process analysis, {3} conformance analysis, {4} performance analysis, {5} social analysis using filters, summarization, sorting, joining and aging + (1) iteration and refocusing. This study does not explain the forms of PBF and the determination of suspected fraud is performed by experts and not computationally. The authors draw the conclusion that process mining can detect fraud in some business process models.

An association rule learning (ARL) algorithm has been used to analyze the correlation between fraud and behavior rules in credit card data transactions in [11]. The behavior of an originator (a user who executes an event) that is consistent with the character of fraudulent behavior is marked as suspicious. This study detected fraud by analyzing SOP deviations using a non-fuzzy method. Therefore, the value of the condition (i.e. not fraud, between not fraud and fraud, fraud, confident fraud and very confident fraud) was not determined.

In another study [12], a hybrid method that combines an ARL algorithm and process mining was proposed. An ARL algorithm was used to identify fraudulent behavior, while process mining was used for analyzing SOP deviation. The method used expert opinion about association rule learning to generate rules for compliance checking. The present study considered the weights of attributes, which were specified subjectively, for determining fraud.

3 Process Mining for Process-Based Fraud Detection

3.1 Process Mining

Process mining is a discipline that focuses on the retrieval of information obtained from event logs [14],[15]. Event logs contain processes that are executed within an information system. The forms of information are: case code, event code, event name, originator name, date and time of event execution.

There are three types of process mining: discovery, conformance (adjustment), and enhancement (refinement) [15],[16]. In order to analyze the existence of a case/process instance in a process model, it is necessary to do conformance checking [17],[18]. In [3], the authors proposed a conformance method by

comparing the case in the event logs with a process model. Furthermore, statistical tools can be used to analyze the business process. This work focused on conformance method development for fraud detection.

3.2 Process-Based Fraud Detection

Process-based fraud (PBF) is a form of fraud that can be identified by processes that deviate from the SOP [13]. Detecting PBF in business processes can be done from three different angles. First, from the point of view of the business process, PBF can be detected by comparing different business processes with respective models. Second, from the point of view of the business role, PBF can be detected by analyzing any process that deviates from the business role. Third, from the point of view of the organization, PBF can be detected by analyzing any originator who deviates from the segregation of duties (SOD) or separation of work [10].

There are some advantages to the use of process mining for detecting PBF. For example, conformance checking can be used to compare business processes with their SOP. Furthermore, this method is able to detect the occurrence of event skipping, which is identified as suspicious [13] in addition to its capability of controlling and analyzing the flow of a business process. Using this method, the sequence of processes in a business can be analyzed. Similar to the previous methods, if a process deviates from the sequence of processes, it is identified as suspicious [3].

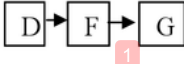
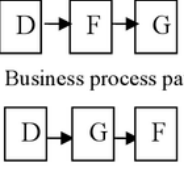
In addition, process mining can be used to analyze the execution time of an event, parallel events and segregation of duties. In this case, an event whose execution time is longer than the standard execution time is identified as suspicious. Also, events being executed in parallel, deviating from the SOP, are identified as suspicious. Likewise, an originator who deviates from the segregation of duties is identified as suspicious [13].

3.3 PBF Attributes

In [11], the authors proposed four PBF attributes, namely skip, throughput time, wrong resource, and wrong decision. In [12]-[13], the authors proposed ten PBF attributes, namely skip sequence, skip decision, throughput time min, throughput time max, wrong resource, wrong duty sequence, wrong duty decision, wrong duty combine, wrong decision, and wrong pattern. Nevertheless, these attributes were not able to identify all types of deviations in PBF.

In this study, a new PBF attribute, namely parallel event, is proposed whose description along with other PBF attributes is shown in Table 1.

Table 1 Description of PBF attributes.

Attribute	Description	Example				
Skip	The execution event jumps from the actual flow	 <p>The event E is skipped</p>				
Wrong pattern	A case pattern is different from that of the business process	 <p>Business process pattern</p> <p>Case pattern</p>				
Throughput time min	The event execution time is shorter than the minimum standard event time	<table border="1" data-bbox="898 766 1125 898"> <tr> <td>Event name</td> <td>Standard time</td> </tr> <tr> <td>Check document complete</td> <td>15 minutes</td> </tr> </table> <p>Execution of "Check document complete" event takes only 8 minutes instead of 15 minutes</p>	Event name	Standard time	Check document complete	15 minutes
Event name	Standard time					
Check document complete	15 minutes					
Throughput time max	The event execution time is longer than the maximum standard event time	Execution of "Check document complete" event takes 35 minutes				
Wrong resource	The event is executed by an illegal originator	"Check collateral location" event is executed by Agus (operator) while it should be done by Budi (manager)				
Wrong duty sequence	Different events are executed by the same originator in a sequence event	"Check SID" and "Check collateral location" events are executed by Agus. Both "Check SID" and "Check collateral location" are sequence events				
Wrong duty decision	Different events are executed by the same originator in a decision event	"Loan decision" and "Check document complete" events are executed by Agus, while "Loan decision" and "Check document complete" events are decision events				
Wrong duty combination	Different events are executed by the same originator in a sequence and	"Check collateral location" and "Check document complete" events are executed by Agus, while "Loan				

Attribute	Description	Example
	decision event	decision” and “Check document” events are sequence and decision events
Wrong decision	The decision making for loan plafond does not comply with SOP	A credit plafond of \$500.000 is approved by the section head; however, according to SOP, credits higher than \$500.000 must be approved by the office head
Parallel event	Different events are executed at the same time	“Check collateral location” and “Check document complete” events are both executed at 12-11-2012 10:10:00

4 Case Study

In this case study, a credit application business process has been investigated to detect fraudulent behavior. The analysis of a credit application process was used to identify fraud attributes, weigh attribute importance and rate PBF. The SOP and business rules have been checked to get the various PBF attributes.

The credit application process is started by completing credit documents. Once these are completed, which is checked by a clerk, the file is delivered to the office head. Further, the office head gives back the clerk the recommendation to to analyze the credit document. After receiving the recommendation, the clerk checks the information of the applicant. If it is cleared, the clerk verifies his/her data at the location of the loan collateral (e.g. personal assets that are used to secure the loan) or at the debtor’s office. Otherwise, the credit application is rejected.

After verification of the collateral, the clerk estimates a credit plafond that conforms to the collateral condition, applicant behavior and credit application rules. Furthermore, the head of credit analysis checks the document to validate the credit plafond. If it is approved, the document is delivered to the credit administration for rechecking of the document. The head of credit administration sends the credit file to the office head in accordance with his/her authority. In case the credit plafond is approved by an unauthorized person (wrong authority), the result is an incorrect decision (wrong decision).

Next, the office head delivers the credit approval to the credit administration, which then passes the credit document to a lawyer for the credit agreement process. If the credit is rejected, the clerk sends the rejection letter to the applicant. Once the credit agreement is completed, the head of the credit

administration makes a draw down letter (letter to transfer) and transfers the credit plafond to the applicant’s account.

Process mining analysis was used to analyze the business process of the credit application to get event sequence, execution time, segregation of duty and rule. If an event was skipped, then the skip event attribute was flagged. If a case had an execution time event longer than the standard execution time, then the throughput time attribute was flagged. Similarly, if applicant information checking was executed by an illegal originator, then the wrong resource attribute was flagged. Furthermore, if an originator (user) executed two different events, then the wrong duty attribute was flagged. Overall, every SOP deviation was connected to a fraud attribute.

5 Method

5.1 Modified Digital Logic (MDL)

Fraud occurring in a business process may deliver different PBF attributes whose weights vary. One solution of this weighing problem is utilizing modified digital logic (MDL). In the proposed method, MDL is used to estimate the weight of attribute importance. Expert discussion of each attribute’s importance is needed to derive the attribute importance weights. Three experts provided an assessment of the importance of each PBF attribute compared to other attributes. The experts assessed every attribute by ‘1’, ‘2’ or ‘3’. To show an attribute is more important, ‘3’ is used. To show attributes have equal importance ‘2’ is used. Meanwhile ‘1’ signifies that an attribute is less important than the other attributes. The attribute importance weights are the same as in [19]. The results of the expert assessment are presented in Table 2.

Table 2 Expert assessment in MDL.

Attributes	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	Pos. dec.	Weights	Linguistic
Skip sequence (A1)	2	2	3	3	2	2	2	2	2	2	2	24	0,09	VI
Skip decision (A2)	2	2	3	3	2	2	2	2	2	2	2	24	0,09	VI
Throughput time min (A3)	2	2	2	1	2	2	2	2	2	2	2	21	0,08	I
Throughput time max (A4)	2	2	2	2	1	2	2	2	2	2	2	21	0,08	I
Wrong resource(A5)	3	3	3	2	2	2	2	2	2	2	2	25	0,10	VI
Wrong duty seq (A6)	2	2	3	3	2	2	2	2	2	2	2	24	0,09	VI

Attributes	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	Pos. dec.	Weights	Linguistic
Wrong duty dec(A7)	2	2	3	3	2	2	2	2	2	2	2	24	0,09	VI
Wrongdutycombine(A8)	2	2	3	3	2	2	2	2	2	2	2	24	0,09	VI
Wrong decision(A9)	2	2	3	3	2	2	2	2	2	2	2	24	0,09	VI
Wrong pattern(A10)	2	1	2	2	2	2	2	2	2	2	2	21	0,08	I
Parallel event(A11)	2	2	2	2	1	2	2	2	2	2	2	21	0,08	I
												253	1,00	

The attribute importance weights were calculated using Eq.(1), taken from [19]:

$$W_j = \frac{P_j}{\sum_{j=1}^n P_j} \tag{1}$$

where *p* is a positive decision and *j* is the number of attributes.

5.2 Fuzzy Logic

Fuzzy logic has the ability to solve problems associated with precision [20]. According to [20], fuzzy sets are a function that consists of membership functions in interval (0-1). A fuzzy approach was used to weigh the PBF attribute values. This is useful for investigating conditions whose value is between fraud and not fraud.

Eleven attribute values were converted, along with the attribute importance weights. In this case, multiple attribute decision making (MADM) based fuzzy logic was implemented to obtain the PBF rating. This represents the weight of the deviation (e.g., a case has a PBF rating of 0.2).

Table 3 Linguistic variables and fuzzy number of deviation rates

Linguistic Variable	Fuzzy number			
High	0.8	1	1	1
Middle	0.3	0.7	0.8	1
Low	0	0	0.3	0.6

Each PBF attribute was initialized by using the following linguistic variables: low, middle and high. The attribute importance weights were specified using the following linguistic variables: very weak (VW), weak (W), fairly important (F), important (I) and very important (VI). The weight of each PBF attribute was determined as in [19],[21]. The expert assessment results and the fuzzy numbers

of deviation rates are shown in Table 2 and Table 3, respectively. In addition, the fuzzy numbers of attribute importance weights are listed in Table 4.

Table 4 Linguistic variables and fuzzy number of attribute importance weights.

Linguistic variable	Fuzzy number			
VI	0.9	1	1	1
I	0.7	0.8	0.9	1
F	0.4	0.6	0.7	0.8
W	0	0.3	0.4	0.7
VW	0	0	0.1	0.3

In this study, we used a trapezoidal fuzzy number, which consists of $a, b, c,$ and d , where $a, b, c, d \in R; a \leq b \leq c \leq d$, according to the method provided in [20],[22]. Let min and max be the minimum and maximum values of the SOP deviations in an attribute, respectively. The variable med is defined as max divided by two, then $a = min, b = min + \frac{1}{2} (med - min), c = med + \frac{1}{2} (max - med)$, and $d = max$. The fuzzy membership of the attribute value is depicted in Figure 1.

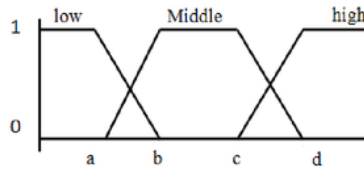


Figure 1 The fuzzy membership of attribute value.

The membership function is defined as follows (Eqs. (2) to (4)):

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

$$\mu_{Midate}[Y] = \begin{cases} \frac{d-y}{(d-c)} & ; c < Y < d \\ 0 & ; Y > d; Y < a \end{cases} \quad (3)$$

$$\mu_{High}[Y] = \begin{cases} 0; Y < c \\ \frac{Y-c}{(d-c)}; c < Y < d \\ 1; Y > d \end{cases} \quad (4)$$

6 Proposed Fuzzy Multi-Attribute Decision Making Approach

In this work, we propose a fuzzy multi-attribute decision making approach. This technique is applied to decide whether an SOP deviation is fraudulent or not. Process mining analysis is employed to analyze the SOP deviation and the deviation numbers are determined as the PBF attribute values. Furthermore, the attribute values and attribute importance weights are both converted into a fuzzy value, which is utilized to get the degree of membership of the attribute values and the attribute importance weights. MADM-based fuzzy logic is used to get the PBF rating of the deviation.

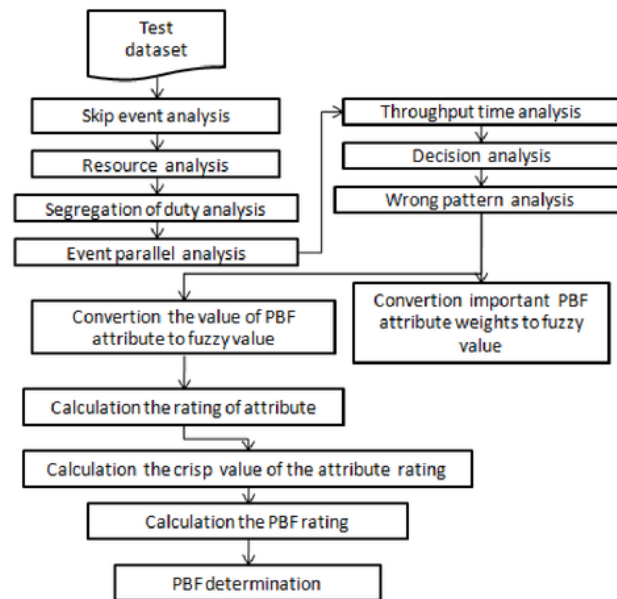


Figure 2 Illustration of PBF detection process.

As described in the previous section, PBF detection consists of eleven steps, where steps 1-6 have been proposed in [3],[11]-[13], while steps 7-12 are proposed by us. Figure 2 shows an illustration of the PBF detection process.

The main steps of PBF detection can be described as follows:

Step 1 Skip event analysis

This analysis recognizes cases in which one or more events were skipped according to the sequence diagram. Control flow analysis is employed to detect such condition. If event skipping has occurred, this affects the skip decision or skip sequence attribute.

Step 2 Throughput time analysis

PBF often involves event execution times shorter than the specified standard execution time. An event whose execution time is shorter or longer than the standard execution time is identified as deviating from the SOP according to the throughput time min or the throughput time max attributes.

Step 3 Resource Analysis

In the SOP, event implementation has to conform to the level of authority. Each event has an originator (user) who executes it. If an event is performed by a wrong originator then this affects the wrong resource attribute.

Step 4 Decision Analysis

This method analyzes the originator who runs the decision event, whether it follows the SOP or not. In the SOP an event must be executed by an originator who has the authority to do this. As an illustration, a credit plafond of more than one billion has to be approved by the director. An event having been approved by the head of a branch office, affects the wrong decision attribute.

Step 5 Segregation of Duty Analysis

Segregation of duty analysis is aimed at checking whether there is a deviation in job segregation. A deviation occurs if an originator runs two or more different events in a case. This analysis is only employed in large companies. Deviation from segregation of duty affects the wrong duty sequence or wrong duty decision or wrong duty combine attributes.

Step 6 Wrong Pattern Analysis

This step analyzes the flow of the business process, which has to conform to the business process pattern. The deviation of such pattern affects the wrong pattern attribute.

Step 7 Parallel Event Analysis

Parallel execution is usually done to speed up the execution time of a credit application process. However, this condition may also be the result of PBF. The proposed method analyzes parallel event execution conform to the SOP. Parallel event execution may constitute a deviation from the SOP and therefore affects the parallel event attribute.

Step 8 Conversion of PBF Attribute Values to Fuzzy Values

The previous steps determine the respective PBF attribute values, which are then converted to the appropriate fuzzy values by following Eq. (2), Eq. (3) and Eq. (4). The results can be grouped into low, middle and high deviation.

Step 9 Calculation of Fuzzy Values of Attribute Importance Weights

The calculation of attribute rating needs the attribute importance weights, which are obtained by converting the attribute importance values to fuzzy values. This step is carried out using Table 2 and Table 4.

Step 10 Calculation of Attribute Rating

The fuzzy number of the attribute rating is obtained by multiplying the fuzzy numbers of the attribute values with the fuzzy numbers of the attribute importance values. Eq.(5) is used to calculate the rating of attributes [21].

$$(x_1, x_2, x_3, x_4) = (a_1 \times b_1; a_2 \times b_2; a_3 \times b_3; a_4 \times b_4) \quad (5)$$

where x_1, x_2, x_3, x_4 are the fuzzy numbers of the attribute rating, and a_1, a_2, a_3, a_4 are the fuzzy numbers of the attribute values, and b_1, b_2, b_3, b_4 are the fuzzy numbers of the attribute importance values.

Step 11 Calculation of Crisp Value of Attribute Rating

The crisp value of the attribute ratings is required to calculate the PBF rating. Eq. (6), as in [19], is utilized to get the crisp value of an attribute rating:

$$S = \frac{-x_1x_2 + x_3x_4 + \left(\frac{1}{3}\right)(x_4 - x_3)^2 + \left(\frac{1}{3}\right)(x_2 - x_1)^2}{-x_1 - x_2 + x_3 + x_4} \quad (6)$$

where S is the crisp value of the attribute rating.

Step 12 Calculation of the PBF Rating

Let S be the crisp value of an attribute rating. The PBF rating is calculated using Eq. (7).

$$PBF = S_1 \vee S_2 \vee S_3 \vee S_4 \dots S_n \quad (7)$$

where n is the total number of PBF attributes.

Step 13 PBF Determination

PBF rating levels are required to decide if PBF is suspected. Expert opinion is employed to establish PBF rating levels according to the method provided in [21]. According to Table 5, a case with a PBF rating of 0.42 is classified as fraud, while a case with a PBF rating of 0.2 is not fraud.

PBF can be mitigated by determining the fraud category on the basis of the PBF ratings. Cases with a PBF rating higher than 0.4 are decided as fraudulent, while cases with a PBF rating lower than 0.41 are determined as not fraudulent. Considering the actual incidence of fraud, experts may decide that fraud occurs at a PBF rating of 0.5. Hence, the category of not fraud can be changed to PBF ratings between 0.01 between 0.5 for PBF detection. Therefore, changing the PBF threshold can be employed for PBF mitigation.

Table 5 Levels of PBF rating.

Linguistic Variable	Rating
Very confident fraud	0.76 – 1
Confident fraud	0.61 - 0.75
Fraud	0.41 - 0.6
Between fraud & not fraud	0.26 - 0.40
Not fraud	0.01 - 0.25

7 Method Evaluation

7.1 Experiment Design

The evaluation process is shown in Figure 3.

In this experiment, data were collected from the event logs of credit applications in the years 2011-2013. The data were grouped into training and testing sets, with 1857 cases (57.733 events/records) and 1147 cases (38.490 events/records) respectively. We analyzed the business process in the training dataset to get the cases which deviated from the SOP. The deviations were identified according to the PBF attributes.

The analysis of the test dataset resulted in 102 cases with deviations from the SOP. Case id 2576 had one PBF attribute: throughput time max. Meanwhile, case id 2580 had two attributes, namely throughput time min and throughput time max. An example of the test dataset result is presented in Table 6.

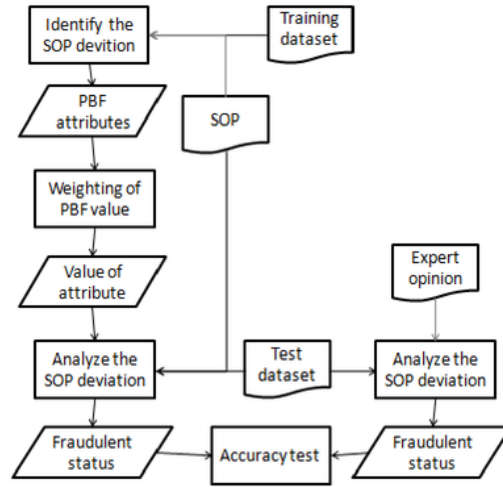


Figure 3 Evaluation process.

Table 6 Example of test dataset result.

No	Skip Seq	Skip dec	T.T. min	T.T. max	Wrong Res.	Wrong Dury Seq.	Wrong Dury	Wrong Dury	Wrong Dec.	Wrong pattern	Event parallel
2576				1							
2580			2	1							
2586										2	
2590										2	
2591										2	
2592										2	
2683					1						2
2688					1						

The PBF attribute value was determined according to the attribute number in the testing dataset result (e.g. throughput time min has a maximum number of deviations of three and a minimum number of deviations of one.; hence, the throughput time min attribute has three in high deviation, two in middle deviation and one in low deviation). In Table 5, some attributes contain a blank or 0 value. This means that an issue with the PBF attributes existed [3],[10]-

[14], however, in the credit applications no deviation occurred. Therefore, a case that deviated from the SOP on these attributes (e.g., skip event, wrong resource, wrong duty and wrong decision) was decided as high deviation [3].

Two methods of PBF analysis, i.e. fuzzy and non-fuzzy, were implemented to identify the advantage of the proposed method. The evaluation consisted of two scenarios: (1) analyze the test dataset using the non-fuzzy method, (2) analyze the test dataset using the fuzzy method. On the other hand, experts analyzed the test dataset using their method. Evaluation of the accuracy and false discovery rating (FDR) of both methods was implemented to see the advantages of each method. Eq. (8) was used to calculate accuracy while Eq. (9) was used to calculate the FDR.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (8)$$

$$FDR = \frac{FP}{TP+FP} \quad (9)$$

7.2 Experimental Result And Discussion

We used a fuzzy approach to investigate SOP deviations that were identified as not fraud. We evaluated the utilization of this fuzzy approach to analyze the test dataset. Skip analysis, throughput time analysis, wrong resource analysis, wrong duty analysis, pattern analysis and parallel event analysis were used to analyze the test dataset. In this study, a PBF rating of 0.01-0.4 was determined not fraud, so higher than 0.4 was decided as fraudulent.

The receiver operating characteristic (ROC) was used to measure the accuracy of the PBF detection method. This framework measures the accuracy by considering true positive (TP), true negative (TN), false positive (FP), and false negative (FN). TP means that the experts' and this method's results have the same determination when a case is fraud. TN also means that the experts' and this method's result have the same determination when a case is not fraud. If the experts decide fraud while the method decides not fraud, then this means a FN. If the experts decide not fraud while the method decides fraud, this means an FP.

Evaluation of the test dataset resulted in 102 cases that deviated from the SOP. The result from the experts discussion proved that according to the non-fuzzy method, 49 cases were identified as true positive, 53 cases as false positive, 68 cases as false negative, and 1045 cases as true negative. Meanwhile, by using the fuzzy method, 38 cases were identified as true positive, 8 cases as false positive, 11 cases as false negative, and 1090 cases as true negative. Using Eq. (8) and Eq. (9), the non-fuzzy method had an accuracy of 0.95 and FDR of 0.51,

while the fuzzy method had an accuracy of 0.98 and ad FDR of 0.17. The evaluation of the test dataset is summarized in Table 7.

Table 7 Result of methods evaluation.

Method	ROC variables				Accuracy
	True positive	False positive	False negative	True negative	
Non-fuzzy	49	53	0	1045	0.95
Fuzzy	38	8	11	1090	0.98

Comparison of the fuzzy method and the non-fuzzy method proves that the fuzzy approach was able to decrease the number of false positives. This decrease in false positives is because this method can correctly detect low deviations from the SOP. The fuzzy approach for PBF detection also had a better accuracy (0.98). Based on those provided data, it can be inferred that there are both advantages and disadvantages of the use of fuzzy and non-fuzzy methods, as shown in Table 8.

Table 8 Advantages and disadvantages of non-fuzzy and fuzzy method.

Method	Advantage	Disadvantage
Non-fuzzy	a. Simple fraud detection (fraudulent or not)	a. Low accuracy, with small deviations of SOP identified as fraudulent b. Cannot detect conditions between fraudulent and not fraudulent
Fuzzy	a. Can detect the tendency weight of the SOP deviation b. PBF detection has better accuracy than non-fuzzy	a. Needs experts to periodically evaluate attribute importance weight, which is determined subjectively by experts b. PBF rating condition needs to be developed in accordance with incidence of fraud

8 Conclusion

We have proposed an MADM approach for PBF rating since detecting low deviation fraud is still challenging. In this paper we have elaborated and evaluated the business process of credit application containing low fraud deviations. A fuzzy approach was used to determine the weight of PBF attributes, while MADM was employed to determine a PBF rating. The experimental results show that the proposed method can reduce the number of false positives and achieved a higher accuracy (0.98) than the non-fuzzy method.

References

- [1] Ngai, E.W.T., Hu, Y., 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, **50**(3), pp. 559-569, 2010.
- [2] Amara, I., Amar, A.B. & Jarboui, A., *Detection of Fraud in Financial Statements: French Companies as a Case Study*, International Journal of Academic Research in Accounting, Finance and Management Sciences, **3**(3), pp. 44-55, 2013.
- [3] Jans, M., van der Werf, M.J., Lybaert, N. & Vanhoof, K., *A Business Process Mining Application for Internal Transaction Fraud Mitigation*, Expert Systems with Applications, **38**(10), pp. 13351-13359, 2011.
- [4] Kalyani, D.R. & Devi, D.U., *Fraud Detection of Credit Payment System by Genetic Algorithm*, International Journal of Scientific & Engineering Research, **3**(7), pp. 1-6, 2012
- [5] Zaslavsky, V. & Strizhak, A., *Credit Card Fraud Detection Using Self-Organizing Maps*, Information & Security, **18**(Cyber crime & cyber security), pp. 48-63, 2006.
- [6] Panigrahi, S., Kundu, A., Sural, S. & Majumdar, A.K., *Credit Card Fraud Detection: A Fusion Approach Using Dempster-Shafer Theory and Bayesian Learning*, Information Fusion, **10**(4), pp. 354-363, 2009.
- [7] Shen, A., Tong, R. & Deng, Y., *Application of Classification Models on Credit Card Fraud Detection*, Proceedings of 2007 International Conference on Service System and Service Management, IEEE, Chengdu, China, pp. 1-4, 2007.
- [8] Chae, M. Shime, S. Cho, H. & Lee, B., *An Empirical Analysis of Fraud Detection in Online Auctions: Credit Card Phantom Transaction*, Proceedings of the 40th Annual Hawaii International Conference on System Sciences, IEEE, Waikoloa, USA, pp. 155a, 2007.
- [9] Chiu, C. & Tsai, C.Y., *A Web Services-Based collaborative Scheme for Credit Card Fraud Detection*, Proceedings of 2004 IEEE International Conference on e-Technology, e-Commerce and e-Service (EEE '04), IEEE, Taipei, Taiwan, pp. 177-181, 2004.
- [10] Stoop, J.J., *Process Mining and Fraud Detection*, Thesis, Business Information Technology Department, Twente University, Enschede, Netherlands, 2012.
- [11] Dewandono, D.R., *Process Sequence Mining For Fraud Detection Using CEP*, Thesis, Informatics Department, Institut Teknologi Sepuluh Nopember, Surabaya, 2013.
- [12] Sarno, R., Dewandono, D.R. Ahmad, T., Naufal, M.F. & Sinaga, F., *Hybrid Association Rule Learning and Process Mining for Fraud*

- Detection*, IAENG International Journal of Computer Science, **42**(2), pp. 59-72, 2015.
- [13] Huda, S., Sarno, R., Ahmad, T. & Santoso, H.A., *Identification of Process-based Fraud in Credit Application*, 2014 2nd International Conference on Information and Communication Technology (ICoICT), Telkom University, Bandung, Indonesia, pp. 84-89, 2014.
- [14] Sarno, R., Sanjoyo, A.B., Mukhlash, I. & Astuti, M.H., *Petri Net Model of ERP Business Process Variations for Small and Medium Enterprises*, Journal of Theoretical and Applied Information Technology, **54**(1), pp. 31-38, 2013.
- [15] Jans, M., Alles, M. & Vasarhelyi, M., *The Case for Process Mining in Auditing: Sources of Value Added and Areas of Application*, International Journal of Accounting Information Systems, **14** (1) pp. 1-20 , 2013.
- [16] Van der Aalst, W.M.P., Discovery, *Conformance and Enhancement of Business Processes*, Springer, pp. 7-8, December 2010.
- [17] Van der Aalst, W.M.P. & de Medeiros, A.K.A., *Process Mining and Security: Detecting Anomalous Process Executions and Checking Process Conformance*, Electronic Notes in Theoretical Computer Science, **121**(Security Issues with Petri Nets and other Computational Models 2004), pp. 3-21, 2005.
- [18] Accorsi, R. & Stocker, T., *On the Exploitation of Process Mining for Security Audits: The Conformance Checking Case*, Proceedings of the 28th Annual ACM Symposium on Applied Computing, Riva del Garda Congress, Trento, Italy, pp. 1709-1716, 2012.
- [19] Vats, S., Vats, G., Vaish, R. & Kumar, V., *Selection of Optimal Toll Collection System for India : A Subjective-Fuzzy Decision Making Approach*, Applied Soft Computing, **21**, pp. 444-452, 2014.
- [20] Zadeh, L.A., *Fuzzy Sets*, Information and Control, **8**(3), pp. 338-353, 1965.
- [21] Barreiros, M.P., Grilo, A. & Cruz-Machado, V., Cabrita, M.R., *Applying Fuzzy sets For ERP Systems Selection Within The Construction Industry*, 2010 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), IEEE, Singapore, pp. 320-324, 2010.
- [22] Shemshadi, A., Shirazi, H., Toreihi, M. & Tarokh, M.J., *A Fuzzy VIKOR Method for Supplier Selection Based on Entropy Measure for Objective Weighting*, Expert Systems with Applications, **38**(10), pp. 12160-12167, 2011.

Junal Pertama

GRADEMARK REPORT

FINAL GRADE

/0

GENERAL COMMENTS

Instructor

PAGE 1

PAGE 2

PAGE 3

PAGE 4

PAGE 5

PAGE 6

PAGE 7

PAGE 8

PAGE 9

PAGE 10

PAGE 11

PAGE 12

PAGE 13

PAGE 14

PAGE 15

PAGE 16

PAGE 17

PAGE 18

Junal Pertama

ORIGINALITY REPORT

5%

SIMILARITY INDEX

%

INTERNET SOURCES

4%

PUBLICATIONS

0%

STUDENT PAPERS

PRIMARY SOURCES

1	Huda, Solichul, Tohari Ahmad, Riyanarto Sarno, and Heru Agus Santoso. "Identification of process-based fraud patterns in credit application", 2014 2nd International Conference on Information and Communication Technology (ICoICT), 2014. Publication	3%
2	Communications in Computer and Information Science, 2011. Publication	1%
3	Submitted to School of Business and Management ITB Student Paper	<1%
4	Lecture Notes in Computer Science, 2006. Publication	<1%

Exclude quotes On

Exclude bibliography On

Exclude matches < 20 words