Available online at http://www.journalcra.com

International Journal of Current Research
Vol. 9, Issue, 08, pp.55517-55523, August, 2017

INTERNATIONAL JOURNAL OF CURRENT RESEARCH

ISSN: 0975-833X

RESEARCH ARTICLE

AN EFFICIENT MONEY LAUNDERY DETECTION USING SPATIO TEMPORAL CORRELATION BASED ASSOCIATION RULE PATTERN

***Dr. Vikas Jayasree**

Department of Global Information Technology, Habib Bank AG Zurich, Dubai, UAE

ARTICLE INFO

Article History:
Received 12th May, 2017
Received in revised form
23rd June, 2017
Accepted 17th July, 2017
Published online 31st August, 2017

Key words:

Money Laundering, Data Mining, Principle Component Analysis, Self Organizing Map, Association Rule Pattern, Multi Clustering.

ABSTRACT

The implementation of data mining techniques for money laundering fraud detection follow traditional information flow of data mining, which starts with feature selection followed by representation, data collection, management, and performance evaluation. Data mining methods have the potentiality for detecting money laundering fraud in banking as they use past cases of fraud to build models, which identify and detect the risk of fraud.With the increased volume of crime datasets and complexity of relationships between these kinds of data, several data mining approaches were presented that involved anomaly detection using principal component analysis and self organizing map. But nevertheless, with high dimension data, they pose serious issues. In this work, to handle high dimensional data with multi-clustering structure, an Efficient Association Rule Pattern based Money Laundering Detection (EARM-MLD) framework is developed. The association rule pattern mining in EARM-MLD framework consists of three major parts. The first part finds frequent large itemsets from banking rules which have support and confidence values more than a threshold number of times. This in turn reduces the time taken for detecting money laundering. The second part is to construct association rules based on spatio temporal model from those large itemsets to easily perform the detection operation and to integrate it with multi clustering algorithm with the objective of reducing the false positive rate. Finally, the multiclustering algorithm involves the set of money transfer group which fulfills the criteria such as row condition, gathering amounts of money to a single account with minimum set size. The multi cluster elements integrated with EARM-MLD framework are treated as suspected operations which operate in money laundering detection work.The money laundering detection in banking system using EARM is experimented on factors such as time for detecting money laundering,false positive rate, scalability, system efficiency ratio, fraud identification accuracy, number of transaction , number of money transfers. Experimental analysis shows that EARM-MLD framework is able to reduce the time for detecting money laundering by 40.14% and reduce the false positive rate by 18.55% compared to the state-of-the-art works.

Citation: Dr. Vikas Jayasree, 2017. "An efficient money Laundery detection using Spatio temporal correlation based association rule pattern", *International Journal of Current Research*, 9, (08), 55517-55523.

INTRODUCTION

Money laundering is a criminal activity with serious threat to financial institutions, through which it becomes a major threat to the entire nation. Most of the financial organizations have been implementing but either is not meeting the requirements of regulatory authorities of the nation or seems to consume more man hours. Principle Component Analysis (Yuh-Jye Lee, 2013), was used to detect the anomaly rate to improve the accuracy. Efficiency was also enhanced by applying online oversample principle component analysis algorithm. But, measures were not taken for data involving high dimensional space.

Corresponding author:* **Dr. Vikas Jayasree,
Department of Global Information Technology, Habib Bank AG Zurich, Dubai, UAE.

Detection and investigation of crime for high dimension data (Mohammad Reza Keyvanpour, 2011), was addressed using data mining with the objective of improving the crime matching process. However, spatio temporal analysis was not performed in an efficient manner. In EARM-MLD framework, the spatio temporal for high dimensional data is addressed using correlation analysis based clustering. An empirical study to detect malicious events in netbios was presented in (Seungwon Shin, 2012). An efficient and effective framework for malware detection (Seungwon Shin, 2013), was designed using host network cooperated detection aiming at reducing the false positive with minimal overhead. A malware detection based on the system call behavior was designed in (Weiqin Ma, 2012), with the objective of reducing the attacks at an early stage using behavior based malware detectors. Anomalies are abnormal events or fraudulent patterns that do

not match with those of the normal patterns. In (Michael, 2015), a contextual anomaly detection framework was designed aiming at reducing the false positive rate. A framework (Thijs Veugen, 2015), was designed to perform secure computations based on collaborative filter. An encryption mechanism called as the identity-based broadcast mechanism was designed in (Jongkil Kim, 2015), with the objective of reducing the computational complexity in detecting the anomaly. A method called pin-entry was designed in (Taekyoung Kwon and Jin Hong, 2015), using a user authentication scheme which was proven to be secure. Another method used channel state information (Muhammad, 2015), to detect anomalies with the aid of Semi Definite Relaxation (SDR) techniques. Robust detection of anomalies was addressed in (Yu Ma, 2015), aiming at improving the fraud identification accuracy in the presence of noise. However, all the above said methods did not take into account the execution time. The EARM-MLD framework on the other hand reduces the execution time by eliminating the unwanted attributes using efficient association rule mining techniques.

One of the essential sales measures for global business environment is e-commerce. With the rapid increase in the advancement ofe-commerce, use of credit cards for purchases has also increased in a significant manner. A framework for credit card fraud detection (SuvasiniPanigrahi, 2009), was designed using Dempster-Shafer's theory to improve the credit card detection accuracy. In (Pamela Castellón González, 2013), neural networks and Bayesian networks were applied to identify the patterns of fraud in taxpayers. Mechanism for anti money laundering was designed in (Xingrong Luo, 2014), with the aid of classifier based on a set of mined frequent rules. In order to detect deception, (Christie, 2011) data and text mining methods were applied resulting in the accuracy. Crowd fraud detection (Tian Tian, 2015) in internet advertising was performed to remove the false alarm rate using two stages called clustering and filter stage. Money laundering detection has received significant attention from researchers in the world. Many techniques (Vikas Jayasree, 2015), have been developed to detect money laundering transaction based on data mining, clustering, and so on. In (Ashish Thakur, 2015; AashleshaBhingarde, 2015), hidden markov model was applied to detect the credit card detection aiming at improving the accuracy and minimizing the false alarm rate. In addition to hidden markov model, K-clustering was applied in (MohdAveshZubair Khan, 2014), with the objective of reducing the fraud transaction rate. Anti money laundering solutions were presented in (Tamer HossamEldinHelmy, 2014) using cluster and link analysis aiming at reducing the time for detection and improving the accuracy.

Based on the aforementioned methods, in this paper, an Efficient Association Rule Pattern based Money Laundering Detection (EARM-MLD) framework is designed. The contributions of EARM-MLD framework include the following: (i) to reduce the time for detecting money laundering by applying efficient association rule pattern, (ii) to substantially reduce the false positive rate by applying spatio temporal based association rule mining and; (iii) to improve the fraud identification accuracy by introducing correlation analysis based clustering.

The rest of the paper is organized as follows. The components of our money laundery detection framework are presented in Section 2 along with a description of the framework and the details of implementation. In Section 3, the experimental setting used in the design of EARM-MLD framework is presented. In Section 4, we discuss the results obtained from implementation studies. Finally, we conclude the paper in Section 5.

MATERIALS AND METHODS

A framework that supports the surveillance in detecting the money laundering is presented. The main parts ofthe framework include efficient design of association rule pattern, eliminating incomplete data by performing mapping, obtaining the cleaned dataset, and analyzing algorithms, such as mapping algorithm and algorithm for Spatio Temporal-based Association Rule Pattern Mapping. The results obtained with the use of these algorithms helps in reducing the time for detecting money laundering and improve the fraud identification accuracy by reducing the false positive time. Figure 1 shows the block diagram of the Efficient Association Rule Pattern based Money Laundering Detection (EARM-MLD) framework.

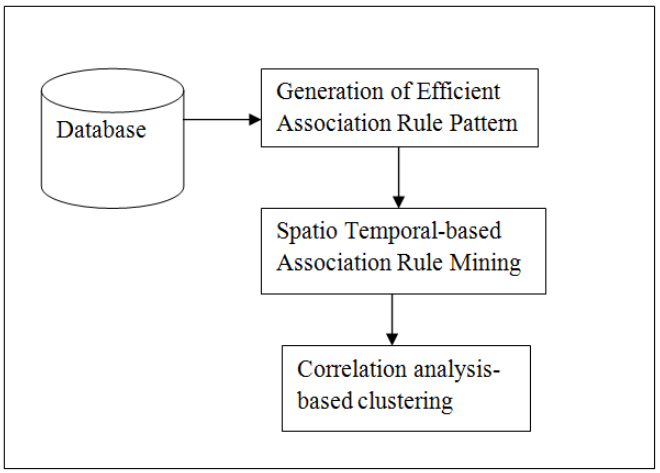

Figure 1. Block diagram of EARM-MLD framework

Figure 1 shows the block diagram that consists of three main parts namely: (1) generation of efficient association rule pattern to remove incomplete data that in turn reduces the time for detecting money laundering, (2) Constructing spatio temporal-based association rule mining withthe objective of reducing the false positive rate using spatio and temporal data and (3) design of correlation analysis-based clustering to improve the fraud detection accuracy. The elaborate description of EARM-MLD framework is described in the forthcoming sections.

Design of Efficient Association Rule Pattern for detecting money laundering

The first step in the design of Efficient Association Rule Pattern based Money Laundering Detection (EARM-MLD) framework is the construction of Association Rule Pattern. The association rule pattern for detecting money laundering identifies the frequent large itemsets having support and confidence values more than threshold number of times. Let us consider database 'DB' in a bank environment consisting of several attributes '$A_i = A_1, A_2, .., A_n$' that denotes the set of items '$Item_i = Item_1, Item_2, .., Item_n$' where each transaction '$Tran \in Item_i$' with patterns '$P_i = P_1, P_2, \dots, P_n$'. Before efficient association rule patterns are formed, the

incomplete data (i.e. attributes) are eliminated. The cleaned data set is then further used as several institutions uses different form of customer data. The obtained information is mapped to a unique processing form. The converted cleaned data set transactional data sets are used for money laundering identification process. Figure 2 shows the measure of evaluating support and confidence by performing mapping operation.

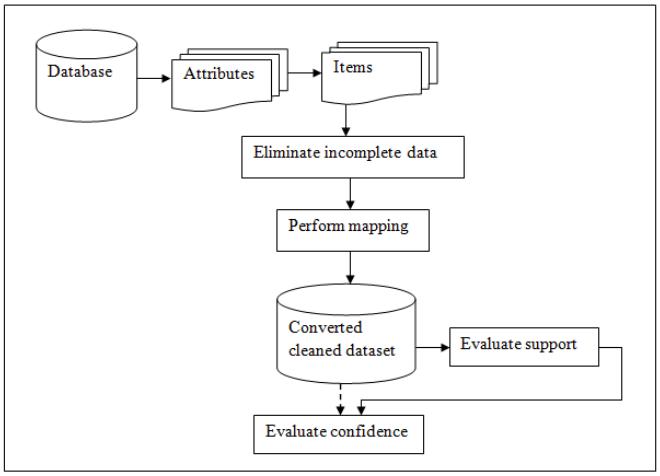

Figure 2. Block diagram of Efficient Association Rule Pattern

As shown in figure 2, by applying the mapping operation, the execution time for detecting money laundering is reduced in a significant manner. Then, the association rule patternis applied to the converted cleaned dataset. Figure 3 shows the algorithmic description for mapping operation.

Input: Database 'DB', Transaction dataset 'T_s', Attributes '$A_i = A_1, A_2, .., A_n$', Attribute set 'A_s', Items '$Item_i = Item_1, Item_2, .., Item_n$', Patterns '$P_i = P_1, P_2, ..., P_n$'

Output: significant removal of unwanted/ incomplete data

Step 1: **Begin**
Step 2: **For** each database 'DB'
Step 3: **For** each transaction 'T_i' from Transaction dataset 'T_s'
Step 4: **For** each attribute 'A_i' from Attribute set 'A_s' and Patterns 'P_i'
Step 5: **If** 'T_i' ∈'A_i' then
Step 6: Extract A_i
Step 6: **Else**
Step 7: Remove A_i
Step 8: **End if**
Step 9: **End for**
Step 10: **End for**
Step 11: **End for**
Step 12: **End**

Figure 3. Mapping algorithm

The mapping algorithm as shown in Figure 3 is performed with the objective of removing the unwanted attributes and extracts those relevant attributes for detecting money laundering. For each transaction and attribute, search is made to identify whether the transaction and attributes belongs to the transaction set and attribute set in addition to the patterns. As a result, the unwanted attributes are removed. This makes the time taken to detect money laundering get reduced with the increase in the number of attributes.

Spatio Temporal-based Association Rule Mining (reduces false positive rate)

Once the unwanted attributes are removed, the second step involved in the design of Efficient Association Rule Pattern based Money Laundering Detection (EARM-MLD) framework is the construction of association rules. Here, the Association rules are constructed from those large item sets (i.e. after removal of unwanted attributes) to perform the detection operation using spatio temporal model. To form Spatio Temporal-based Association Rule Mining, the association rule pattern is generated. Then representation of association rule pattern for detecting money laundering is given as follows.

$$X \Rightarrow Y, where\ X \in Item_i, Y \in Item_i\ and\ X \cap Y = \phi \quad \text{.....(1)}$$

From (1), the support '$Supp$' of '$X \Rightarrow Y$' is theprobability of a transaction in 'DB' that includes both '$X\ and\ Y$ ' with similar patterns 'P_i'. Onthe other hand confidence '$Conf$' of '$X \Rightarrow Y$' is the probability of atransaction including 'X' will contain 'Y' too with similar patterns 'P_i'. Then, Association Rule Pattern for detecting money laundering is formulated as given below with the help of support '$Supp$' and confidence '$Conf$'.

$$Supp_{XY \Rightarrow Z} = \left(\frac{\sum_{i=1}^{n} Supp\_Count_{XYZ(i)} \cup P_i}{\sum_{i=1}^{n} DB_{size(i)}}\right) \quad \text{.....(2)}$$

$$Supp_{XY} = \left(\frac{\sum_{i=1}^{n} Supp\_Count_{XY\,(i) \cup P_i}}{\sum_{i=1}^{n} DB_{size(i)}}\right) \quad \text{.....(3)}$$

$$Conf_{XY \Rightarrow Z} = \left(\frac{Supp_{XY \Rightarrow Z}}{Supp_{XY}}\right) \quad \text{.....(4)}$$

From (2), (3) and (4), the support and confidence value is measured. The Spatio Temporal-based Association Rule Mining is obtained using spatio and temporal data as given below.

$$P_{st}(\alpha, \beta) = \left(\frac{\delta}{\beta t - \alpha t}\right) \quad \text{.....(5)}$$

From (5), '$t(\alpha, \beta)$' is the pattern obtained with start time 'α' and end time 'β' and 's' representing the temporal search of the pattern with start time 'α' and end time 'β'. With the obtained pattern based on temporal search model, let us consider that there are '$FP_i = FP_1, FP_2, .., FP_n$' frequent patterns. The algorithm for Association Rule Pattern Mapping identifies 'FP_k'from 'FP_{k-1}' in two stages. During the first stage, 'k' pattern sets called as the candidate pattern sets 'Can_k' and other itemsets from database 'DB' are generated. With the generated candidate pattern sets any '$k-1$' itemsets is not said to be a subset of frequent 'k' itemset which is not candidate frequent pattern set. Therefore, those itemsets which are not considered to be candidate frequent pattern set are said to be detected as fraudulent activity and forms cluster. This in turn reduces the false positive rate. The algorithm for Association Rule Pattern Mapping using spatio temporal data is given below. Figure 4 shows the Algorithm for Spatio Temporal-based Association Rule Pattern Mapping. As shown in the figure, for each items and patterns observed, the support and confidence value is measured. With the obtained support and confidence value, spatio temporal data is evaluated. The frequent patterns using spatio temporal data is evolved to identify the candidate pattern set. Finally, not candidate

frequent pattern set are removed which therefore reduces the false positive rate.

Input: Items '$Item_i = Item_1, Item_2, .., Item_n$', Pattern '$P_i = P_1, P_2, \dots, P_n$', Frequent Pattern '$FP_i = FP_1, FP_2, .., FP_n$'
Output: association rule pattern mined data
Step 1: **Begin**
Step 2: **For** each Database 'DB' and items '$Item_i$'
Step 3: **For** eachPattern 'P_i'
Step 4: Evaluate Support using (3)
Step 5: Evaluate Confidence using (4)
Step 6: Measure spatio temporal data using (5)
Step 7: Generate candidate pattern sets 'Can_k'
Step 8: Remove not candidate frequent pattern set
Step 9: **End for**
Step 10: **End for**
Step 11: **End**

Figure. 4 Algorithm for Spatio Temporal-based Association Rule Pattern Mapping

Correlation analysis based clustering

The final stage in the design of Efficient Association Rule Pattern based Money Laundering Detection (EARM-MLD) frame work is the efficient clustering of transactional activities. With the objective of efficiently detecting the suspicious patterns and improving the fraud identification accuracy, correlation analysis based clustering is carried out. The correlation analysis based clustering algorithm builds connected graphs where the nodes (i.e. customers) denote the money transfers. The EARM-MLD framework creates an edge between two customers when the target account from the transfer of the first customer is similar to that of the source account from the transfer of the second customer.

Let us consider, a cluster 'Cl_i' contains 'c_1' customers representing money transfers with target accounts 'a_1' and 'c_2' customers denoting money transfers with source accounts 'a_2'. When amoney transfer has to be included with the source account 'a_2' and the targetaccount 'a_1', then the correlation between customers with total number of edges to be created is '$c_1 + c_2$'.

The correlation analysis based clustering involves the set of money transfer group fulfilling the criteria such as row condition, gathering amounts of money to a single account with minimum set size. The row condition in correlation analysis based clustering is set in such a way that it removes the graph nodes (i.e. customers) which do not fulfill the condition (i.e. observed patterns or abnormal patterns). The second criteria gathering amounts of money to a single account with minimum set size helps in detecting the target accounts. The algorithmic description of correlation analysis based clustering is as given below. Figure 5 shows the algorithmic description. With the above said algorithm, detection of money laundering is efficiently performed by comparing the already obtained support and confidence value with those of the correlated support and confidence value. If both the values are said to be similar, then no detection of money laundering is observed, else states that the money laundering activity is detected. Therefore, by applying correlation analysis based clustering, fraud identification accuracy is said to improve with respect to scalability (i.e. increase in customers).

Input: Tree root 'R', Child node '$child$'
Output: efficient analyzes of trees with detection of money laundering in nodes was observed
Step 1: **Begin**
Step 2: **For** each tree root ''
Step 3: Measure of offspring
Step 4: **If** '$count\ (R)$' >$sum\ (child)$
Step 5: Remove $child$ node
Step 6: **End if**
Step 7: **If** '$count\ (R)$' <$sum\ (child)$
Step 8: Measure Support and Confidence using () and ()
Step 9: **If** measured support and confidence is equal to value already found then
Step 10: No detection of money laundering
Step 11: **Else**
Step 12: Detection of money laundering
Step 13: **End if**
Step 14: **End for**
Step 15: **End**

Figure 5. Algorithm for correlation analysis based clustering

Experimental settings

Efficient Association Rule Pattern based Money Laundering Detection (EARM-MLD) framework is experimented in JAVA platform using Statlog (German Credit Data) Data Set. The Statlog German Credit Data classifies the people using a set of attributes list. To efficiently implement the algorithms in EARM-MLD framework, numerical attributes from Strathclyde University are added to make it effective algorithm for money laundering identification. The Statlog German Credit Data include 17 attributes and has been coded as integer type and 3 under the categorical type. Efficient Association Rule Pattern based Money Laundering Detection (EARM-MLD) with the existing Anomaly Detection using Principle Component Analysis (AD-PCA) [1] and Detecting and Investigating crime using Data Mining (DI-DM) [2]. The StatlogGerman Credit Data contains the 1000 instances on financial area for performing the experimental work to identify the vulnerable accounts. The experiment is conducted on the factors such as time for detecting money laundering, false positive rate, scalability, system efficiency ratio, fraud identification accuracy, number of transaction, number of money transfers.

DISCUSSION

To validate the efficiency and theoretical advantages of the proposed Efficient Association Rule Pattern based Money Laundering Detection (EARM-MLD) framework with Anomaly Detection using Principle Component Analysis (AD-PCA) [1] and Detecting and Investigating crime using Data Mining (DI-DM) [2], the results of implementation under JAVA is presented. The parameters of the EARM-MLD framework are chosen as provided in the experiment section.

Impact of time for detecting money laundering

The time for detecting money laundering is the amount of time required to transact a given number of transaction that include

both fraudulent and non-fraudulent activities. The time for detecting money laundering involves both the activities. It is measured in terms of milliseconds (ms) and formulated as given below

$$ET = Time\ (Single_T) * (Total_T) \quad \text{.....(6)}$$

From (6) the execution time for detecting money laundering 'ET', is obtained on the basis of time for single transaction '$Single_T$' and the total transactions '$Total_T$'. To better understand the effectiveness of the proposed EARM-MLD framework, extensive experimental results are reported in table 1.

Table 1. Tabulation for time for detecting money laundering

Number of transaction	Time for detecting money laundering (ms)		
	Earm-mld	Ad-pca	Di-dm
150	3.79	5.81	6.78
300	5.25	7.30	9.34
450	8.23	10.28	12.32
600	7.55	9.60	11.64
750	11.38	13.43	15.47
900	10.25	12.30	14.34
1050	14.98	16.09	18.13

The experiments are conducted using JAVA to measure and experiment the factors by analyzing the percentage of result using table and graph values. Results are presented for different number of transactions. The results reported here confirm that with the increase in the number of transaction, the time for detecting money laundering also gets increased. Finally, the value of user acceptance ratio gets saturated when the transaction ranges from 900 – 1050.

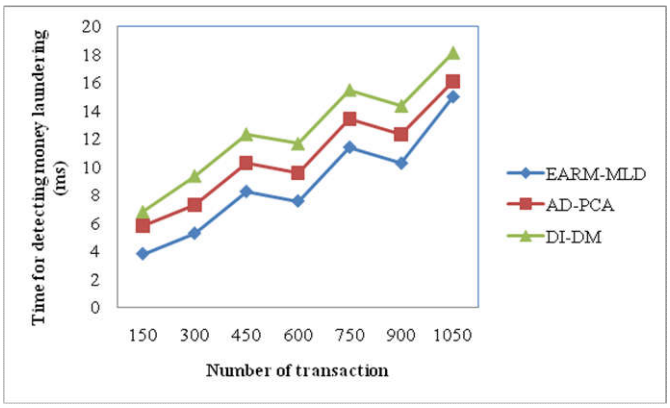

Figure 6. Measure of time for detecting money laundering

Figure 6 shows the time for detecting money laundering based on the number of transaction considered for experimental purpose. Our proposed mapping algorithmperforms relatively well when compared to two other methods AD-PCA [1] and DI-DM [2]. This is because using Efficient Association Rule Pattern that eliminates the incomplete data (i.e. attributes) and converted cleaned dataset helps in reducing the time for detecting money laundering of EARM-MLD framework by 30.34% compared to AD-PCA. Besides, the elimination of incomplete data using mapping function uses the support and confidence value that in turn minimizes the time for detecting money laundering by 49.95% compared to DI-DM.

Impact of false positive rate

False positive rate also known as the false alarm ratio refers to the probability of falsely rejecting the transaction that may include fraudulent activities during a test. It is measured in terms of percentage (%) and formulated as given below

$$FPR = \left(\frac{Falsely\ rejected\ transactions\ as\ fraudulent}{Number\ of\ transaction}\right) * 100 \quad \text{.....(7)}$$

From (7), the false positive ratio 'FPR' is identified, which states that lower the false positive rate more efficiently the method is said to be. In order to reduce the false positive rate, spatio-temporal based association rule mining is used that efficiently reduces the fraudulent activities. In the experimental setup, the number of transaction considered ranges from 150 to 1050. The results of 7 different transactions for experimental setup are listed in Table 2.

Table 2. Tabulation for false positive rate

Number of transaction	False positive rate (%)		
	Earm-MLD	AD-PCA	DI-DM
150	11.50	12.75	16.23
300	16.28	18.28	22.28
450	28.44	30.44	34.44
600	25.13	27.17	31.21
750	39.25	41.29	45.33
900	48.23	50.27	54.31
1050	42.55	44.59	48.64

The targeting results of false positive rate using EARM-MLD framework with two state-of-the-art methods [1], [2] in figure 7 is presented for visual comparison based on the number of transactions made during banking operations.

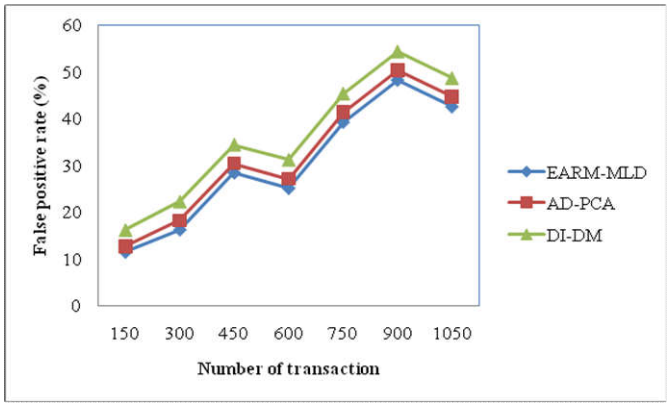

Figure 7. Measure of false positive rate

From Figure 7, it is evident that the false positive rate is reduced using the proposed EARM-MLD framework. Our framework differs from the AD-PCA [1] and DI-DM [2] in that we have incorporated the Spatio Temporal-based Association Rule Miningthat develops an association rule mining based on the spatio and temporal data and therefore decreases the false positive rate by 8.25% compared to AD-PCA. In addition, with the application of Association Rule Pattern Mapping using spatio temporal data, frequent candidate sets are generated whereas the non candidate frequent set are removed. This is in helps in reducing the false positive rate by 28.86% compared to DI-DM.

Impact of fraud identification accuracy

The fraud identification accuracy is the number of relevant money transfers retrieved to the total number of money transfers made in the database. It is measured in terms of percentage (%) and formulated as given below

$$A = \left(\frac{number\ of\ relevant\ money\ transfers\ (i.e.non\ fraudulent)}{number\ of\ money\ transfers}\right) * 100 \quad \text{.....(8)}$$

From (8), the fraud identification accuracy ‘’ is measured and higher the value, more efficient the method is said to be. In table 3 we show the analysis of fraud identification accuracy with respect to number of money transfers made ranging between 500 and 3500 that measures the amount of fraud identification accuracy in bank industry measured in terms of percentage (%).

Table 3. Tabulation for fraud identification accuracy

Number of money transfers	Fraud identification accuracy (%)		
	EARM-MLD	AD-PCA	DI-DM
500	76.13	73.45	70.89
1000	79.23	74.21	70.16
1500	84.13	79.11	74.06
2000	78.24	73.22	69.17
2500	83.15	78.13	74.08
3000	85.22	80.20	76.15
3500	89.19	84.17	80.12

Figure 8 presents the variation of overall fraud identification accuracy of EARM-MLD framework over different number of money transfers in bank sector.

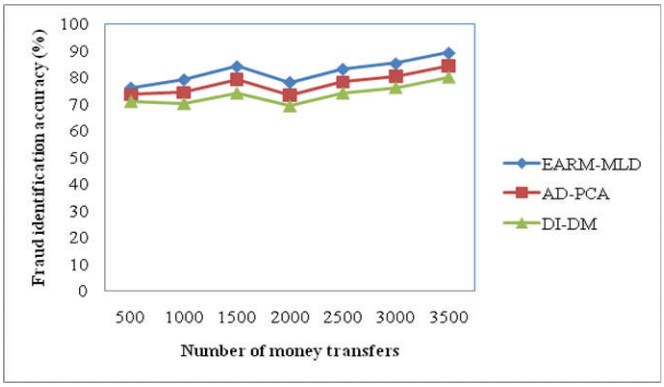

Figure 8. Measure of fraud identification accuracy

The results provided in figure 8 confirm that the proposed EARM-MLD framework significantly outperforms the other two methods, AD-PCA [1] and DI-DM [2]. The better performance of EARM-MLD framework is achieved due to the fact that with the application of correlation analysis based clustering in EARM-MLD framework, fulfilling the criteria such as row condition, gathering amounts of money to a single account with minimum set size, resulting in effective fraud identification accuracy with an improvement of 6.01% compared to AD-PCA [1]. Also in EARM-MLD framework, the row condition removes the graph node which does not fulfills the conditions and gathering money to a single account with minimum set size helps in improving the fraud identification accuracy by 9.42% compared to DI-DM.

Impact of system efficiency ratio

In this section the impact of overall system efficiency ratio using three methods, EARM-MLD, AD-PCA and DI-DM is presented.

Table 4 Tabulation for system efficiency ratio

Methods	System efficiency ratio (%)
EARM-MLD	82.15
AD-PCA	70.39
DI-DM	62.50

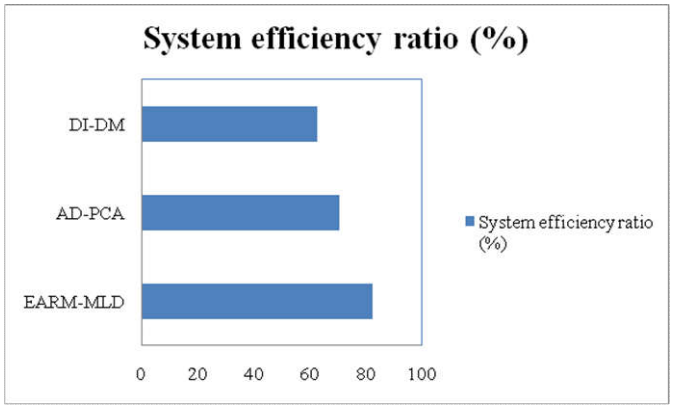

Figure 9. Measure of system efficiency ratio

Table 4 and figure 9 shows the system efficiency ratio with respect to the number of transactions made and amount of money transfers performed. As shown in the figure, the system efficiency ratio is improved using the proposed EARM-MLD framework. This is because with the application of correlations analysis based clustering algorithm efficient detection of money laundering activity is performed. This helps in the improvement of the system efficiency ratio by 14.31% compared to AD-PCA and 11.20% compared to DI-DM respectively.

Conclusion

In this paper, Efficient Association Rule Pattern based Money Laundering Detection (EARM-MLD) frameworkis provided for banking sector to handle high dimensional data using clustering. This framework avoids the computationally expensive time for detecting money laundering in bank sector. As the framework uses Efficient Association Rule Pattern approach for detection of money laundering in bank sector, it increases the fraud identification accuracy by performing correlation analysis based clustering. As a result, the proposed Spatio Temporal-basedAssociation Rule Pattern Mapping algorithm achieves high system efficiency ratio by obtaining support and confidence value for efficient identification of fraudulent activity. By applying the mapping algorithm in EARM-MLD framework in bank sector, overcomes the false positive rate on falsely rejected transactions as fraudulent to improve the fraud identification accuracy with minimum time for detecting money laundering. A series of experiments were conducted and performedin JAVA to test thetime for detecting money laundering, false positive rate, fraud identification accuracy and system efficiency ratio to measure the effectiveness of EARM-MLD framework. The resultsshow that EARM-MLD framework offers better performance with an improvement of fraud identification accuracy by 7.71% and system efficiency ratio by 12.75% compared to AD-PC and DI-DM respectively.

REFERENCES

AashleshaBhingarde, AvnishBangar, Krutika Gupta and SnigdhaKarambe, “Credit Card Fraud Detection using Hidden Markov Model”, *International Journal of Advanced Research in Computer and Communication Engineering,* Volume 4, Issue 3, March 2015, Pages 169 – 170.

Ashish Thakur, Bushra Shaikh, Vinita Jain and A. M. Magar, “Hidden Markov Model in Credit Card Fraud Detection”, *International Journal of Advanced Research in Computer*

Science and Software Engineering, Volume 5, Issue 2, February 2015, Pages 997 – 1000.

Christie M. Fuller, David P. Biros and DursunDelen, "An investigation of data and text mining methods for real world deception detection", Elsevier, Volume 38, Issue 7, July 2011, Pages 8392 – 8398.

Jongkil Kim, Willy Susilo, Man Ho Au and Jennifer Seberry, "Adaptively Secure Identity-Based Broadcast Encryption With a Constant-Sized Ciphertext", *IEEE Transactions on Information Forensics and Security,* Volume 10, Issue 3, March 2015, Pages 679 – 693.

Michael A. Hayes and Miriam AM Capretz, "Contextual anomaly detection framework for big sensor data", *Journal of Big Data,* February 2015.

Mohammad Reza Keyvanpour, Mostafa Javideh and Mohammad Reza Ebrahimi, "Detecting and investigating crime by means of data mining: a general crime matching framework", Elsevier, Volume 3, 2011, Pages 872 – 880.

MohdAvesh Zubair Khan, Jabir DaudPathan and Ali Haider Ekbal Ahmed, "Credit Card Fraud Detection System Using Hidden Markov Model and K-Clustering", *International Journal of Advanced Research in Computer and Communication Engineering,* Volume 3, Issue 2, February 2014, Pages 5458 – 5461.

Muhammad R. A. Khandaker and Kai-Kit Wong, "Masked Beamforming in the Presence of Energy-Harvesting Eavesdroppers", *IEEE Transactions on Information Forensics and Security,* Volume 10, Issue 1, January 2015, Pages 40 – 54.

Pamela Castellón González and Juan D. Velásquez, "Characterization and detection of taxpayers with false invoices using data mining techniques", Elsevier, Volume 40, Issue 5, April 2013, Pages 1427 – 1436.

Seungwon Shin, GuofeiGu, Narasimha Reddy and Christopher P. Lee, "A Large-Scale Empirical Study of Conficker", *IEEE Transactions on Information Forensics and Security,* Volume 7, Issue 2, April 2012, Pages 676 – 690.

Seungwon Shin, Zhaoyan Xu and GuofeiGu, "EFFORT: A New Host-Network Cooperated Framework for Efficient and Effective Bot Malware Detection", Elsevier, Volume 57, Issue 13, 9 September 2013, Pages 2628–2642.

SuvasiniPanigrahi, AmlanKundu, Shamik Sural and A.K. Majumdar, "Credit card fraud detection: A fusion approach using Dempster–Shafer theory and Bayesian learning", Elsevier, Volume 10, Issue 4, October 2009, Pages 354 – 363.

Taekyoung Kwon and Jin Hong, "Analysis and Improvement of a PIN-Entry Method Resilient to Shoulder-Surfing and Recording Attacks", *IEEE Transactions on Information Forensics and Security,* Volume 10, Issue 2, February 2015, Pages 278 – 291.

Tamer HossamEldinHelmy, Mohamed zakiAbd-ElMegied, Tarek S. Sobh and Khaled Mahmoud ShafeaBadran, "Design of a Monitor for Detecting Money Laundering and Terrorist Financing", *International Journal of Computer Networks and Applications,* December 2014, Pages 15 – 25.

Thijs Veugen, Robbert de Haan, Ronald Cramer and Frank Muller, "A Framework for Secure Computations with Two Non-Colluding Servers and Multiple Clients, Applied to Recommendations", *IEEE Transactions on Information Forensics and Security,* Volume 10, Issue 3, March 2015, Pages 445 – 457.

Tian Tian, Jun Zhu, Fen Xia, Xin Zhuang and Tong Zhang, "Crowd Fraud Detection in Internet Advertising", Proceedings of the 24th International Conference on World Wide Web, May 2015, Pages 1100 – 1110.

Vikas Jayasree and Siva Balan, "Money Laundering Identification On Banking Data Using Probabilistic Relational AuditSequential Pattern", *Asian Journal of Applied Sciences,* 2015, Pages 173-184

Vikas Jayasree and Siva Balan, "Money laundering regulatory risk evaluation using Bitmap index –based decision tree", *Journal of the Association of Arab Universities for Basic and Applied Sciences*, 2016, Pages 96-102

Weiqin Ma, Pu Duan, Sanmin Liu, GuofeiGu and Jyh-Charn Liu, "Shadow Attacks: Automatically Evading System-Call-Behavior Based Malware Detection", *Journal in Computer Virology,* Volume 8, Issue 1-2, May 2012, Pages 1 – 19.

Xingrong Luo, "Suspicious Transaction Detection for Anti-Money Laundering", *International Journal of Security & Its Applications,* Volume 8, Issue 2, March 2014, Pages 157 – 166.

Yu Ma, Li-Min Zhang and Hao-Tong Wang, "Reconstructing Synchronous Scrambler with Robust Detection Capability in the Presence of Noise", *IEEE Transactions on Information Forensics and Security,* Volume 10, Issue 2, February 2015, Pages 397 – 408.

Yuh-Jye Lee, Yi-Ren Yeh and Yu-Chiang Frank Wang, "Anomaly Detection via Online Oversampling Principal Component Analysis*", IEEE Transactions on Knowledge and Data Engineering (TKDE),* Volume 25, Issue 7, July 2013, Pages 1460 – 1470.

****