International Journal of Science and Research (IJSR)
ISSN (Online): 2319-7064
Impact Factor (2012): 3.358

Multi Model Approach for Money Laundering Identification Using Fund Transition Matrix

G. Krishnapriya[1], Dr. M. Prabakaran[2]

[1]Research Scholar, Bharathidasan University, Trichy, Tamil Nadu, India

[2]Assistant Professor, Department of Computer Science Government Arts College, Ariyalur, Tamil Nadu, India

Abstract: ***Suspicious money transfer has been increased due to the reason that the terrorist organizations and other communities funding peoples to perform many illegal threatening activities. The growing illegal funding shakes the financial stability of any country and has to be monitored in such a way to stop illegal flow of money. The money laundering is such activity where the origin of money has not be found and considered as threatening fund. There are many approaches has been discussed earlier but struggles with origin identification. To solve the problem of money laundering identification, we propose a multi model approach using fund transition matrix. Each transaction is considered as a state of state transition diagram and the transaction can have any number of states. The transaction will be considered as a complete one where the origin of money is identified. We compute a transaction completeness measure based on incoming and outgoing fund of any account and generate number of states from each state of the transition graph. From identified states and transactional completeness measure, we identify and declare set of suspicious accounts which has illegal money transfer. The proposed approach has produced efficient results and the accuracy of money laundering identification is improved.***

Keywords: State Transition Diagram, Transaction Completeness Measure, Money Laundering

1. Introduction

Money laundering is a financial activity, where the fund transferred to any account could not be identified and the source of money is unidentified. The transfer of unidentified money questions the financial stability of any country, because the fund available in any bank account which is suspicious may be transferred to other account at any time. Also , the terrorist organizations funds the volunteers in this manner and they perform such illegal activities to collapse the social stability. The fund may be transferred through various accounts and pass towards the destination account. Even the account holder does not know the person who transferred the amount and so on. This kind of huge money transfer may help the illegal persons to involve in any kind of activities which spoils the social and financial stability.

State Transition diagram is one which has many states and can have any number of transitions which has names. In case of financial transition, each state is considered as an account where the amount is transferred. Like this we can construct a state transition diagram with many numbers of states to identify the source of money. If the transaction is complete then there will be complete states and end with success flag. If the transaction is incomplete then the final state of the diagram could not be identified.

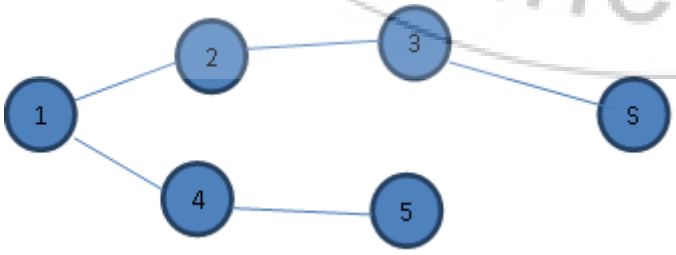


Figure 1: Snapshot of state transition diagram

The figure 1 shows the example state transition diagram which is constructed based on transaction of two accounts. In this, the account 1 has credited with amount through two different channels. In first case, the amount has navigated through 3-2-1 which is complete one so that it ends with S flag and shows that the origin of money has identified. In the other case the amount transferred through 5-4 is unidentified so that it does not end to the S node.

Based on state transition diagram we can compute the transactional completeness measure, which shows the source of money identified. For each account of the transactional data set, we can construct the state transition diagram with identified accounts as states. Based on constructed state transition diagram, the completeness measure of any account could be computed using various parameters.

2. Related Works

There are many approaches has been discussed for identifying money laundering. We discuss few of them here around the problem.

A framework on developing an intelligent discriminating system of anti-money laundering [1], proposes a four layer model to identify money laundering. Different layers play different roles during the analyzing procedure. Data of Transaction layer and Account Layer are submitted from the root bank branches and have composed the fundamental sources. Only isolated intelligence may be derived from the perspectives of both inner layers. Organization layer and Link layer provide perspectives to take a comprehensive and aggregate discriminating and analyzing procedure to all data involved by multiple banks, areas and departments, to check, contrast, mine, judge and derive in all those data collected from varied channels. The later layers have much more advantages during macro situation judgment and relevant cases investigation.

Money Laundering Detection using Synthetic Data [2], they present an analysis of the difficulties and considerations of applying machine learning techniques to this problem. We discuss the pros and cons of using synthetic data and problems and advantages inherent in the generation of such a data set. They using a case study and suggest an approach based on Multi-Agent Based Simulations (MABS).

A number of basic countermeasures against money laundering have been proposed, including basic statistical analysis which constrains the amount of the transactions as well as restricting their frequency [3]. Other methods that complement these basic security measures are based on checking every customer against a black list originating from previous investigated cases and a white list to e.g. avoid mistakes when faced with persons with the same name. Unfortunately, these and other methods have proved to be insufficient [20].

Several machine learning techniques have been used for detecting fraud, and more specifically money laundering, [4]. From the point of view of machine learning, the application is interesting, due to the successful classification rate (high True Positives and low False Positives) that the classification model can achieve compared to other methods such as simple rule based detection that compares transactions against fixed thresholds.

Data mining based methods have also been used to detect fraud [5]. This leads to the observation that machine learning algorithms can identify novel methods of fraud by detecting those transactions that are different (suspicious) in comparison with the benign transactions. This problem in machine learning is known as novelty detection. Supervised learning algorithms have been used on a synthetic data set to prove the performance of outlier's detection.

Combining Digital Forensic Practices and Database Analysis as an Anti-Money Laundering Strategy for Financial Institution [11], propose an anti-money laundering model by combining digital forensics practices along with database tools and database analysis methodologies. As consequence, admissible Suspicious Activity Reports (SARs) can be generated, based on evidence obtained from forensically analyzing database financial logs in compliance with Know-Your-Customer policies for money laundering detection.

The State of Phishing Attacks, [12], discusses about phishing which is a kind of social-engineering attack where criminals use spoofed email messages to trick people into sharing sensitive information or installing malware on their computers. Victims perceive these messages as being associated with a trusted brand, while in reality they are only the work of con artists. Rather than directly target the systems people use, phishing attacks target the people using the systems. Phishing cleverly circumvents the vast majority of an organization's or individual's security measures. It doesn't matter how many firewalls, encryption software, certificates, or two-factor authentication mechanisms an organization has if the person behind the keyboard falls for a phish.

Event-based approach to money laundering data analysis and visualization [13], discusses crime specific event patterns which are crucial in detecting potential relationships among suspects in criminal networks. However, current link analysis tools commonly used in detection do not utilize such patterns for detecting various types of crimes. These analysis tools usually provide generic functions for all types of crimes and heavily rely on the user's expertise on the domain knowledge of the crime for successful detection. As a result, they are less effective in detecting patterns in certain crimes. In addition, substantial effort is also required for analyzing vast amount of crime data and visualizing the structural views of the entire criminal network. In order to alleviate these problems, an event-based approach to money laundering data analysis and visualization is proposed.

A Fistful of Bitcoins: Characterizing Payments Among Men with No Names [14], explore this unique characteristic further, using heuristic clustering to group Bitcoin wallets based on evidence of shared authority, and then using re-identification attacks (i.e., empirical purchasing of goods and services) to classify the operators of those clusters. From this analysis, we characterize longitudinal changes in the Bitcoin market, the stresses these changes are placing on the system, and the challenges for those seeking to use Bitcoin for criminal or fraudulent purposes at scale. From the above discussion, it shows clearly that the previous approaches are struggles with the accuracy of money laundering identification.

3. Methodology

The transactional data set Ts has N number of accounts and each of the account has M number of beneficent. The construction of state diagram can be constructed for each account with many numbers of states each of which shows the set of accounts from where the money has been transferred. Identifying money laundering has three stages namely preprocessing, transactional state diagram construction, money laundering identification.

3.1 Preprocessing

Each transaction T_i from transactional data set Ts, has various attributes to be present. At this stage, each transaction T_i is verified for the presence of all the attributes. If any of the transaction has missing values and incomplete then it will be removed from the transaction set Ts. The preprocessed and noise removed set is used for further processing.

For example the transaction set Ts with N number of transaction T_i is removed with noise and preprocessed as follows:
Identify number of attributes of transaction as follows:

$$NA = \int \sum Attr(Ti) \nexists NA \quad (1).$$

Based on the equation (1), the preprocessing is performed as follows:

$$Ts = \int_{i=1}^{N} Ts \cap (Ti \forall NA) \quad (2)$$

3.2 TSD Construction

From the preprocessed transaction set, for each transaction T_i from Ts, we generate transaction state diagram Tsd as follows: for each transaction Ti, we perform back propagation and identify set of accounts from where the amount has transferred. Each account may have many accounts linked and among them there will be a chain of account. Finally each linked nodes are considered as a state and linked to form the state transition diagram.

Algorithm:
Input: Transactional data set Ts.
Output: State diagram set SDS.
For each Ti from Ts
Generate a node N with name Account number Sds = $\int G(Ti) = Ti.AccNo$
Identify the source of transaction St = $\sum Acc \in Ti$
Link the nodes and set the status.
Perform the link generation till the source identified.
SDS = $\sum SDS \cup sds$
End.

3.3 Money Laundering Identification:

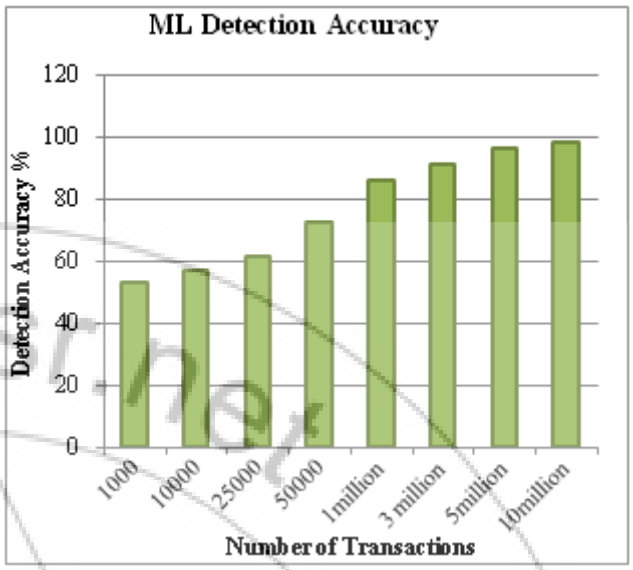
From constructed transactional state transition diagram, for each account we compute the transactional completeness measure using the diagram and number of transactions of the account. Finally for each account we compute the cumulative malicious weight which shows the trustworthy of the account. If the account has more malicious weight the it is considered as malicious account which has more transactions for money laundering and the accounts related are clustered. From clustered accounts the accounts which has transaction with more suspicious amount is selected to monitored.

Input: SDS
Output: Vulnerable Accounts VS.
Step1: for each account Ai from SDS
Compute average number of states Tavg = $\int_{i=1}^{N} \sum Sds.States$
Tavg = $\frac{Tavg}{N}$
Compute transactional measure Tcm = $\frac{Tavg \times No\ of\ sds\ with\ success}{Total\ number\ of\ graphs}$

Compute cumulative malicious weight cmw = $\frac{Tcm - No\ of\ sds\ with\ success}{Total\ number\ of\ graphs}$
If cmw>Th then
Cluster linked accounts Cl = $\int \sum A \in Sds$
Identify Account A with more fund transfers.
End.
End.

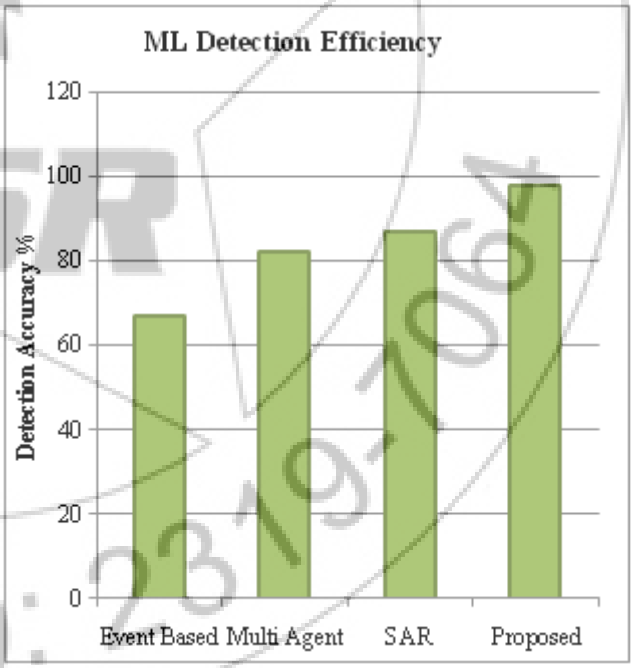
4. Experimental Results

The proposed method has been evaluated using various transactional set collected from different banking sectors and we have separated the accounts which are linked through different banks. Finally we have collected 5000 accounts from different banks having 10 million transactions. The proposed method has produced efficient results and detection accuracy is also higher.



Graph 1: shows the efficiency of identifying money laundering

The graph1 shows the efficiency of identifying money laundering with respect to number of transaction used. It is clear that the efficiency is increased if the size of transaction is increased. The proposed methodology produces efficient result by increasing the size of transaction.



Graph 2: Comparison of Money Laundering Detection Accuracy

The graph2 shows the comparison of money laundering detection accuracy between different methods. It shows clearly that the proposed method has produced more efficiency in money laundering detection.

5. Conclusion

We analyze various methodologies to identify money laundering crime. We identify that all methods have scalable in accuracy and efficiency. We proposed a multi variant relational model which uses time variant transactional data. The proposed method has produced higher efficient results and with accurate findings. The proposed method has produced results with less time complexity.

References

[1] Nhien An Le Khac an investigation into Data Mining approaches for Anti Money Laundering, 2009 International Conference on Computer Engineering and Applications IPCSIT vol.2 (2011) © (2011) IACSIT Press, Singapore.
[2] Edgar Alonso Lopez-Rojas , Money Laundering Detection using Synthetic Data, The 27th annual workshop of the Swedish Artificial Intelligence Society (SAIS), 14–15 May 2012,
[3] Dan Magnusson. The costs of implementing the anti money laundering regulations in Sweden. Journal of Money Laundering Control, 12(2):101{112, 2009.
[4] Clifton Phua, Vincent Lee, Kate Smith, and Ross Gayler. A comprehensive survey of data mining- based fraud detection research. Arxiv preprint arXiv:1009.6119, 2010.
[5] AgusSudjianto, Sheela Nair, Ming Yuan, Aijun Zhang, Daniel Kern, and Fernando .Statistical Methods for Fighting Financial Crimes.Technometrics, 52(1):5{19, February 2010.
[6] J Pavon, M Arroyo, S Hassan, and C Sansores. Agent-based modelling and simulation for the analysis of social patterns. Pattern Recognition Letters, 29(8):1039-1048, June 2008.
[7] AgusSudjianto, Sheela Nair, Ming Yuan, Aijun Zhang, Daniel Kern, and Fernando Cela-Daz. Statistical Methods for Fighting Financial Crimes.Technometrics, 52(1):5{19, February 2010.
[8] YaboXu, Ke Wang, Ada Wai-chee Fu, Hong Kong, and Philip S Yu.Anonymizing Transaction Databases for Publication. International Journal, pages 767-775, 2008.
[9] Rui Liu, Research on anti-money laundering based on core decision tree algorithm, IEEE Conference on Control and Decisions , pp:4322-4335, 2011.
[10] Odense, Laundering Sexual Deviance: Targeting Online Pornography through Anti-money Laundering, European Intelligence and Security Informatics Conference, 2012.
[11] Bucharest, Combining Digital Forensic Practices and Database Analysis as an Anti-Money Laundering Strategy for Financial Institutions, Third International Conference on Emerging Intelligent Data and Web Technologies, 2012.
[12] Jason Hong, The State of Phishing Attacks, Communications of the ACM, Vol. 55 No. 1, Pages 74-81, 2012.
[13] Tat-Man Cheong, Event-based approach to money laundering data analysis and visualization, Proceedings of the 3rd International Symposium on Visual Information Communication, ACM , 2010.
[14] Sarah Meiklejohn, A Fistful of Bitcoins: Characterizing Payments Among Men with No Names, ACM 2013.
[15] E. Androulaki, G. Karame, M. Roeschlin, T. Scherer, and S. Capkun.Evaluating User Privacy in Bitcoin.In Proceedings of Financial Cryptography 2013, 2013.
[16] I. Miers, C. Garman, M. Green, and A. D. Rubin. Zerocoin: Anonymous Distributed E-Cash from Bitcoin. In Proceedings of the IEEE Symposium on Security and Privacy, 2013.
[17] T. Moore and N. Christin. Beware the Middleman: Empirica Analysis of Bitcoin-Exchange Risk. In Proceedings of Financial Cryptography 2013, 2013.