# MONEY LAUNDERING ANALYSIS BASED ON TIME VARIANT BEHAVIORAL TRANSACTION PATTERNS USING DATA MINING

[1]G.KRISHNAPRIYA M.C.A., M.PHIL, [2]Dr.M.PRABAKARAN

[1]Research Scholar, Bharathidasan University, Trichy, Tamil Nadu, India.
[2]Assistant Professor, Department of Computer Science Government Arts College, [2]Ariyalur, Tamil Nadu, India.
Email- [1]krishnapriyaphd123@gmail.com

## ABSTRACT

Money laundering a suspicious fund transfer between accounts without names which affects and threatens the stability of countries economy. The growth of internet technology and loosely coupled nature of fund transfer gateways helps the malicious user's to perform money laundering. There are many approaches has been discussed earlier for the detection of money laundering and most of them suffers with identifying the root of money laundering. We propose a time variant approach using behavioral patterns to identify money laundering. In this approach, the transaction logs are split into various time window and for each account specific to the fund transfer the time value is split into different time windows and we generate the behavioral pattern of the user. The behavioral patterns specifies the method of transfer between accounts and the range of amounts and the frequency of destination accounts and etc.. Based on generated behavioral pattern , the malicious transfers and accounts are identified to detect the malicious root account. The proposed approach helps to identify more suspicious accounts and their group accounts to perform money laundering identification. The proposed approach has produced efficient results with less time complexity.

**Keywords:** *Money Laundering, Data Mining, Behavior Patterns.*

## 1. INTRODUCTION:

Money laundering is the process of hiding the origin of the illegal money and showing it as legitimate one. The country economic trustworthy highly depend on how well the country enforced lawful act against money laundering. The process of money laundering is used for many purposes for example, suppose a group of foreign persons tries to spoil the economic condition of any country, they anonymously transfers huge amount to their group accounts. Later at a point of time they could suddenly withdraws all the money to some other account which spoils the economic status of the country. This kind of activity is against the country economy and has to be avoided and prevented from that.

Similarly the terrorist organizations spend more money in inducing the individuals to involve in terrorism. They send more money anonymously to some sleeper cells in the world or in a country to perform some specific task which will be assigned to them at any time. The peoples of certain kind accepts those kind of money and engage with them for the task. These activities are questions the national security and the financial monitoring agency have the responsibility to identify such unlawful fund transfer and has to trace the root of the fund.

Data mining is the process of extracting useful information from large set of data. Such data mining techniques can be used to identify the suspicious accounts from which the fund has been transferred and so on. We propose such a methodology using which we can identify the set of malicious accounts. Each account has number of beneficiaries between which the fund transfer is performed. Even the destination account holder does not know the source, still fund transfer can be performed. In this manner, the identity of the malicious user will not be known to the account holder but will get credited large amount.

Behavioral pattern is one , represent the activities of the user account and transactions. The user may have N number of accounts linked and may perform many transfer to accounts of others. But the amount and time and period of transaction varies naturally. This can be represented as transaction behavioral pattern and can be used to identify money laundering.

## 2. RELATED WORKS

Statistical Methods for Fighting Financial Crimes [3], focuses on two important types of financial crimes: fraud and money laundering. It discusses some of the traditional statistical techniques that have been applied as well as more recent machine learning and data mining algorithms. The goal of the article is to introduce the subject and to provide a survey of broad classes of methodologies accompanied by selected illustrative examples.

Laundering Sexual Deviance: Targeting Online Pornography through Anti-money Laundering [4], concentrates on cyber-pornography/obscenity, which encompasses online publications or distribution of sexually explicit material in breach of the English obscenity and indecency laws. After examining the major deficiencies of the attempts to restrict illegal pornographic representations, the authors aim to highlight that the debate regarding their availability in the Internet era neglects the lucrative nature of the circulation of such material, which can be also targeted through anti-money laundering. Rising profits fuel the need to recycle the money back into the legal financial system, with a view to concealing their illegal origin. Anti-money laundering laws require disclosure of any 'suspicion' related to money laundering, thus opening another door for law enforcement to reach the criminal.

Combining Digital Forensic Practices and Database Analysis as an Anti-Money Laundering Strategy for Financial Institutions [5], 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.

Event-based approach to money laundering data analysis and visualization [6], proposes crime specific event patterns 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 in this paper.

A Fistful of Bit coins: Characterizing Payments Among Men with No Names [7], explore this unique characteristic further, using heuristic clustering to group Bit coin 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 Bit coin market, the stresses these changes are placing on the system, and the challenges for those seeking to use Bit coin for criminal or fraudulent purposes at scale.

Zerocoin: Anonymous Distributed E-Cash from Bitcoin [9], uses standard cryptographic assumptions and does not introduce new trusted parties or otherwise change the security model of Bitcoin. We detail Zerocoin's cryptographic construction, its integration into Bitcoin, and examine its performance both in terms of computation and impact on the Bitcoin protocol.

Combining Digital Forensic Practices and Database Analysis as an Anti-Money Laundering Strategy for Financial Institution [10], 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 analysing database financial logs in compliance with Know-Your-Customer policies for money laundering detection.

Bitcoin and Money Laundering: Mining for an Effective Solution [11], represents a disruptive financial technology that many AML and money transmitter statutes are ill prepared to deal with. Virtual currencies in general have broken the trend of physical, government-backed coin and paper currencies, and it is unlikely that any new law will capture all iterations of emerging technologies for any significant period. But this does not mean that Bitcoin and similar virtual currencies should be deemed illegal or should be onerously regulated to compensate for the lack of initial oversight. In an increasingly digital world, it makes perfect economic and societal sense to allow digital currencies, government backed or otherwise.

Suspicious Transaction Detection for Anti-Money Laundering [12], propose a classification based algorithm to effectively detect suspicious transactions. Specifically, we consider the financial transactions as a data stream, and try to construct a classifier based on a set of mined frequent rules. Our experiments on a simulated transaction dataset based on real world banking activities prove the efficiency of our proposed method.

All the above discussed methods are suffers with identifying the root of the malicious amount transfer and we propose a new approach for money laundering identification here in this paper.

## 3. PROPOSED METHOD

The proposed method has three stages namely: Preprocessing- where the input transactional data set is verified for its completeness and incomplete transactions are removed from the data set. Behavioral Pattern generation- at this stage the input data set is split based on various time window to generate transaction patterns. Finally in money laundering stage the generated patterns are used for identifying money laundering.

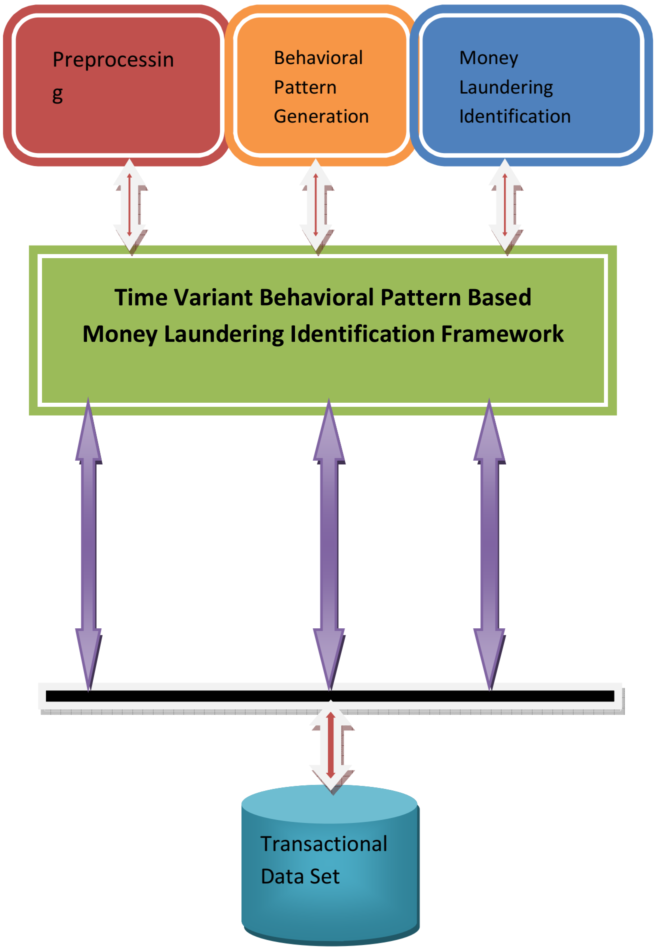


*Figure1: Proposed System Architecture.*

### 3.1. Preprocessing:

The transactional data set has many attributes and the log size is heavier to process. The retrieved log will be a noisy one to process , so that we remove the incomplete log to make it compatible for processing. The transactional log of different banking sectors has different attribute set and different pattern. At this stage the input transactional data set Ts is used to identify set of all attributes Ats available with the transactional set. Then for each transaction $T_i$ from Ts , we identify the presence of the values for attributes Ats. If there is any missing values then it will be removed from the transactional list. We identify distinct attributes like name, account number, address, from account, to account, time , data , tag mentioned, amount, branch , source branch and etc.

```
**Algorithm**

**Input:** Transactional data set Ts.

Output: Noise Removed Dataset Ts.

Step 1: initialize Attribute set Ats to Null.

Step 2:  for each transaction $T_i$ from $T_s$.

        Extract set of attributes present in $T_i$ as Al = $\int_1^N \sum Attr \in Ti$

        Add new attributes to attribute set Ats = $\sum(Attr \in Ats) + \sum Attr(Al) \nexists Ats$

        End.

Step3: for each transaction $T_i$ from Ts

        if $\forall(Attr$ $_{(Ats)}$$\exists Attr(Ti))$ Then

        else

        Ts = Ts$\cap$ $Ti$

                end

        end

Step4: Stop.
```

### 3.2. Behavioral Pattern Generation:

We generate the behavioral pattern of each user account based on the transactions. For each account , the fund transfer and fund credit time window will be computed. Based on computed time window, it will be split into N number of small window will be generated. The transaction logs will be split into N number of sub sets based on number of time window. At each time window the fund transfers with more money, without names are identified. The identified accounts are selected for further processing. The identified transactions are constructed as patterns with the features of the transaction log or the attribute set.

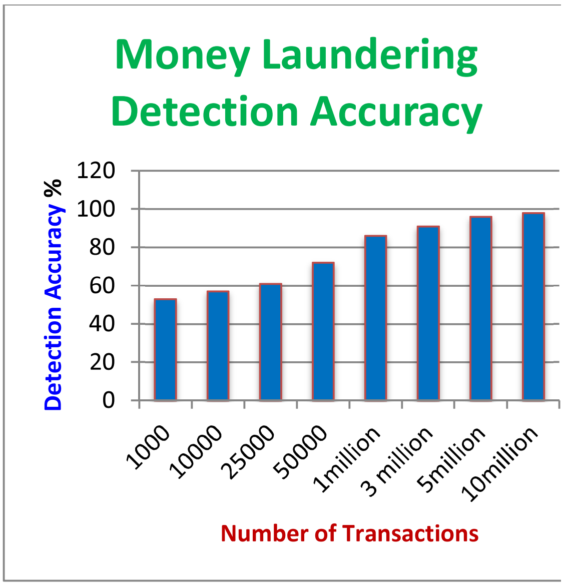
```
Algorithm:

Input: Transactional Data set Ts.

Output: Behavioral pattern set Bps.

Step1: Compute overall Time value Tv = $\int_1^N Ti(least\ Time) - Ti(\max Time)$

Step2: Split Tv into N time window //N- can be any number of windows for example 5

Tws = $\left(\left(\frac{Tv}{N}\right), \left(\left(\frac{Tv}{N}\right) + N - i\right), \left(\left(\frac{Tv}{N}\right) + N - i + 1\right), \dots, (\left(\frac{Tv}{N}\right) + N - (i + N - 1))\right)$

Step3: For each time window $Tw_i$ from Tws

extract transactions from Ts as $Ts_{Tw}$(i) = $\sum Ti.time \in Twi$

        end.

Step4: For each time window Twi

For each Beneficiary of account

        compute average number of fund transfers performed Aft = $\frac{\sum FundTransfer(Twi)\ to\ Acc}{Total\ \ no.of\ Fund\ Transfer}$

                compute average amount Transferred AAT = $\frac{\sum FundTransfer(Twi)}{Total\ Fund\ Transfer}$

compute average number of fund received performed Afr = $\frac{\sum Fund\ Received(Twi)\ From\ Acc}{Total\ \ no.of\ Fund\ Received}$

compute average amount received AAR = $\frac{\sum Fund\ Received(Twi)}{Total\ Fund\ Received}$

construct Transactional Behavior Pattern TBP = {From Acc, To Acc, Tw, Aft, AAT, Afr, AAR}.

Add to Transactional Pattern set Bps = Σ(TBP)+$TBP_i$

        end.

Step 5: stop.
```

### 3.3. Money Laundering Identification:

The money laundering identification is performed using computed behavioral pattern Bps. Upon receiving money from any account the process of pattern generation and identification will be performed. Based on computed pattern , the pattern for current time window will be generated. Using the patterns and current transaction , if the amount received or transferred is more than the average value of fund transfer then it will be selected as suspicious transaction and will be added to monitoring. The account linked with the transaction will also undergo with the same procedure and will be selected as suspicious account.

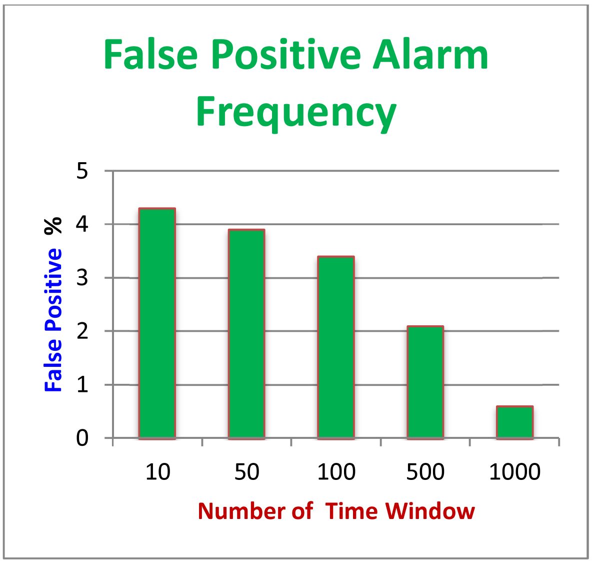
## 4. RESULTS AND DISCUSSION

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.



*Graph2: False Positive Alarm Frequency*

The graph2, shows the false positive alarm generated by the proposed method. It shows the the percentage of false positive result is get reduced while the time window increases. It also shows that the percentage of false positive values is reduced while the transactional log increases.

## 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 Time variant behavioral pattern based money laundering identification framework which generates transactional behavior patterns according to various time window. Based on generated patterns the money laundering identification is performed. The proposed method has produced higher efficient results and with accurate findings. The proposed method has produced results with less time complexity.

## REFERENCES:

[1]. 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.

[2]. Agus Sudjianto, Sheela Nair, Ming Yuan, Aijun Zhang, Daniel Kern, and Fernando . Statistical Methods for Fighting Financial

Crimes. Technometrics, 52(1):5{19, February 2010.

[3]. 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.

[4]. Sarah Meiklejohn, A Fistful of Bitcoins: Characterizing Payments Among Men with No Names, ACM 2013.

[5]. E. Androulaki, G. Karame, M. Roeschlin, T. Scherer, and S. Capkun. Evaluating User Privacy in Bitcoin. In Proceedings of Financial Cryptography 2013, 2013.

[6]. 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.

[7]. Agus Sudjianto, 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]. Odense, Laundering Sexual Deviance: Targeting Online Pornography through Anti-money Laundering, European Intelligence and Security Informatics Conference, 2012.

[9]. 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.

[10]. 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.

[11]. Danton Bryans, Bitcoin and Money Laundering: Mining for an Effective Solution, Indiana Law Journal, volume. 89, issue 1, 2014.

[12]. Xingrong Luo ,Suspicious Transaction Detection for Anti-Money Laundering, International Journal of Security and Its Applications Vol.8, No.2 (2014), pp.157-166

[13]. S. LaValle, “Big data, analytics and the path from insights to value”, MIT Sloan Management Review, vol. 52, no. 2, (2011), pp. 21-31.

[14]. P. Zikopoulos and C. Eaton, “Understanding big data: Analytics for enterprise class hadoop and streaming data”, McGraw-Hill Osborne Media, (2011)