

# Financial Fraud Detection Along with Outliers Pattern

Sheetal Nadagoudar <sup>1</sup>, Thippeswamy K<sup>2</sup>

<sup>1</sup>Foruth Sem M.Tech, Department of Studies in CS&E, VTU PG Center, Mysuru

<sup>2</sup>Professor & Chairman, Department of Studies in CS&E, VTU PG Center, Mysuru

\*\*\*

**Abstract** - Now a day's financial fraud such as concealment of money is a serious issue, which makes our society step towards in to criminal activity. The money concealment not only harm individual but also harmful to our environment of society. This type of money concealment may be they use for illegal activity like terrorism. To detect this type of network is very difficult task and network is also so complex to find out. The trading and transaction reveals the interaction between the entities involved and outlier pattern detects the fraud pattern, thus our proposed method is useful to detect not only fraud detection but also it reveals the fraud pattern. All existing system detects only whether fraud happened or not but outliers detects type of fraud, what are entities involved, timings of fraud and its pattern.

**Key Words:** Financial fraud, Outliers, pattern detection, money concealment, feature matrix, type of fraud.

## 1. INTRODUCTION

In recent years financial fraud cases are increased dramatically such as concealing of money, credit card fraud, Account hacking etc. All these funds may goes in to illegal activity may be called criminal purpose like terrorism, this effects not only individual but also to society, group of companies, institutions and economy of nation.

This type of illegal money laundering arise the question on security of nation. Concealment of money involved different entities and different money laundering network. Usually concealment of money defined as process of money laundering by illegally. Unfortunately real word situation is different when consider Free-Trade Zone, here exchange of trades and goods are involved in a complex network. In FTZ concealing of money detection difficult to detect and rectify the situation. Consider example of money laundering involve trades, goods, misrepresentation of price, quality and quantity of item. In the field of data mining there is growing need for robust, reliable outliers detection system. Although research has been done in this area, little of it has focused on graph based data. Most are mining techniques search for outlier attributes, considering objects distance in the full data space, these methods fall prey to randomly distributed attributes combination.

Several outliers methods are based on attribute value pair that are collected from transaction data. But some other data may collected from supervised and unsupervised collection

of information. So always value and attribute pair does not dependent on any one said above. Money laundering is totally different from value pair, in some cases its highly influencing from financial data, political data, social ,scientific data and last but not least technical data. On other side some commerce data are auto correlated.

Some features of trading commerce properties are identical, feature properties don't club the interaction information in data. The relation between business entities shows the potential causality, business on going, outliers information and ups and downs of a commerce. Graph based mining methods are one of the important theories that used to find out the relation between data, entities, the outlier detection is suspicious criminal activity. Exploring graph based method gives new perspectives for fraud identification and enable us to do advance research on financial fraud. And activities of fraud detection based on graph. Graph based methods are provides two important features i.e. One is fraud identification and other one is fraud tracing method. In this paper we have focussed on how to identify the fraud cases, what is type of fraud, when it happened and what are the methodologies are used. In this paper we would like develop whole frame work for how to use both graph matrix and feature matrix. A feature matrix is set of features that characterized a given set of linguistic units with respect to a finite set of properties.

In lexical semantics, feature matrices can be used determine the meaning of specific word fields. Ex: Suppose take an example of two genders male and female If we indicate male as minus the female as star disorders due to excessive Internet Usage every year is drastically increasing.

The remnant of the paper is organized as follows. Section II describes the Background or the related work of our scheme. Section III describes preliminaries. Section IV describes the Conclusion.

## 2. BACKGROUND

Financial fraud is an important problem that has been researched within diverse research areas and application domains. Many outliers techniques have been specifically developed for certain application domains, while others are more generic. This survey tries to provide a structured and comprehensive overview of the research on outliers detection. We have grouped existing techniques into different

categories based on underlying approach adopted by each technique.

Pattern matching detection refers to the problem of finding pattern in data that do not confirm to expected behavior. These non-conforming patterns usually called as anomalies, outliers or discordant observation, exception, aberrations, surprises or peculiarities.

1. Defining a normal region that encompasses every possible normal behavior is very difficult. In addition, the boundary between normal and outliers behavior is often not precise.
2. When outliers are the result of malicious action, the malicious adversaries often adapt themselves to make the outliers observation appear normal, thereby making the task of defining normal behavior more difficult.
3. Availability of labeled data for training/validation of models used by anomaly detection techniques is usually a major issue.
4. Often the information contains noise that tends to be similar to the original outliers and hence is difficult to differentiate and remove. These are the some of the challenges faced by while detection of pattern.
5. Outlier is major task in data analysis. Outlier are objects that highly deviate from regular objects in their local neighborhood. In application such as fraud detection analysis is suspicious and unexpected object.
6. For the evaluation process by using synthetic and real word data to the effectiveness and efficiency of proposed paper.
7. The Outlier identification process it detects the problem of complex network as well as pattern identification.

The outlier detection and anomaly feature identification consist of various factors by choosing and observing them we can conclude the related fraud or crimes as well. Suppose consider example of two graph's having many nodes, edges too, now if we take graph having four points or dots each edge connected by node and in that all nodes are at same distance so, at this situation if any wrong thing happened in the graph then automatically it will affect other three of the points. Same as in the social networking area if any person or even a client behaves abnormally, we can consider a situation as oddity. Just as graph example here in fraud identification method, considering anomaly factors like observing neighbor client details or observing client background like weather his regular to any online site, his regular in paying payments, his/her bank details, credit or debit card numbers depending

on all these data we can assume that something not well in the transaction. So like this some percentage of frauds with outlier identification can be judged.

Trade based money laundering is also considered as one of the big issue in recent years, laundered money may be used for the terrorist activities. Illegal laundering money may cause harm to the economy of the country, organization, group of companies and education centers. The international trade system is clearly subjected to a wide range of risk and vulnerabilities that can be exploited by criminal organization and terrorist financiers. Because these factors trade market faces several problems and risk. This also affects the growth of the Nation.

The idea behind this paper is finding anomalies along with pattern matching, in graph based data where the anomalies substructure in a graph is part of normative. The concept of finding pattern that similar to frequent, or good, pattern, is different from most looks like something unusual or bad pattern. Till now we focused on only anomaly, pattern matching and outlier identification but in actual we don't concentrate on the effects of anomaly attacks. Anomaly detection is an essential component of the protection against novel attacks. Anomaly attack depends on many factors like entropy, conditional entropy, information gain and information cost for anomaly detection. Now a question that arises is what is entropy and the answer is it is a measure of number of possible arrangements of items. The best example of entropy is solid wood burns and becomes ash, smoke and gases. These outliers can be detected on basis of instance based learning.

In financial frauds there are number of varieties it may be simple crime, complex crime, now if we consider a simple crime like misrepresentation of prices, quality or quantity of goods on invoice, wrong presentation of delivered address etc. Considering all types of frauds totally we call it as service based money laundering. Trade based money laundering has divided frauds as different categories...

1. A criminal organization exports a relatively small amount of scrap, but relatively they are shown in papers as huge amount of goods weight.
2. They generate fake bills and bill numbers.
3. While the process of shipping time only they hack the goods and transferred to other place.
4. Fraud companies mention falls described goods and services.
5. Not only they mislead the organization, they also mislead govt. of nation.

6. They misprint the tax amount also. By this type of crime they it affects the economy of country.

In certain cases they there is no transparency to their business ,they hide all their goods' details, pricing of the product, weight of material , quality they using and whom it has to be send. Not only in trade field even in banking sectors also so many frauds are happened in our daily life like ATM fraud, online banking fraud, entering wrong CVV, less balance fraud, under balance frauds. Because of all types of crimes only our paper gives solution to at least some percentage of crime rates will be reduced using pattern matching and outliers, solve online fraud, offline frauds. At least to avoid or some percentage level of fraud should be controlled using some solutions like

1. The countries should have information where trade data and relevant financial information are being stored.
2. Banks should have each and every customer's details and organizations doing heavy transactions with the banks.
3. Every country should have a special crime organization and that works 24 hours with fast communication.
4. Appoint skilled people to identify the crime and its pattern.
5. All information about every citizen of nation should be secured in database, there is no chance to hack the data.

By considering above points at least some amount of crime level will be decreased.

### 3. PRELIMINARIES

The main aim is to detect financial fraud detection along with pattern matching with respect to given problem.

To do this we have to go through some of the factors Those are as follows:

1. Fabricated Data

Technically the fabricated data is small part of data base. We only extract the 150 financial entries and up to 2500 transaction from the data sets then we inject fraud pattern in to the fabricated data.

2. Concealment Of Money Data

It's a basically data set where suppose two financial entries trading history, there is a minute edge between them and weight of two entries are calculated depends on features of the data available in scenario.

3. Insurance Fraud Data

This is also one of the best example of concealment of money from online. Now a days it's very common to seen this type of

cases on day to day life. To detect the fraud in this case firstly observe the data sets available in related properties and those values are compared with fraud scenario.

4. Credit Card Fraud

Credit card fraud is most common to seen in daily life. This is wide ranging term for theft and fraud committed using or involving payment card, such as debit or debit card, as a fraudulent source of funds in a transactions.

5. Transport Company Fraud

This is also one of the fraud happening now a days, suppose some body book vehicle for goods transportation and at the time of payment they enter wrong pin or CVV. By using outlier pattern matching we can detect the crime.

6. Online Shopping Fraud

This type crimes so common in today's world while doing online shopping user should be so careful to use his/her credit or debit card. These crimes also known by the proposed frame work.

### 4. CONCLUSIONS

The proposed system Financial fraud detection along with pattern matching which can perform fraud detection based on feature and similarity matrix simultaneously.

It introduces the new way of detecting not only the fraud and related data set but also which are involved in as same as pattern matching along with source of fraud where when and what type of fraud happens.

### ACKNOWLEDGEMENT

The authors would like to thank all the faculties of Dept. of studies in Computer Science and Engineering, VTU Regional office, Mysuru, their parents, family members, friends and anonymous reviewers' encouragement and constructive piece of advice that prompted for a new round of rethinking of research, additional experiments and clearer presentation of technical content.

### REFERENCES

1. Laundering: Risks and Regulatory Responses," Social Sci. Electron. Publishing, 2012, p. 6.
2. United Press International. (May 2009). Trade-Based Money Laundering Flourishing.
3. L.Akoglu,M.McGlohon,andC.Faloutsos,"OddBall:Spotting anomalies in weighted graphs," in Proc. Pacific-Asia Conf. Knowl. Discovery Data Mining, 2010, pp. 410-421.
4. V.Chandola,A.Banerjee,andV.Kumar,"Anomaly detection:Asurvey," ACM Comput. Surv., vol. 41, no. 3, 2009, Art. no. 15.

5. W.Eberle and L. Holder, "Mining for structural anomalies in graph-based data," in Proc. DMin, 2007, pp. 376–389.
6. C. C. Noble and D. J. Cook, "Graph-based anomaly detection," in Proc. 9th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2003, pp. 631–636.
7. H. Tong and C.-Y. Lin, "Non-negative residual matrix factorization with application to graph anomaly detection," in Proc. SIAM Int. Conf. Data Mining, 2011, pp. 1–11.
8. S. Wang, J. Tang, and H. Liu, "Embedded unsupervised feature selection," in Proc. 29th AAAI Conf. Artif. Intell., 2015, pp. 470–476.
9. Z. L. in , M. Chen, and Y. Ma. (2010). "The Augmented lag range multiplier method for exact recovery of corrupted low-rank matrices."
10. Faloutsos, "Neighbourhood formation and anomaly detection in bipartite graphs," in Proc. 15th IEEE Int. Conf. Data Mining, Nov. 2005, p. 8.
11. A. Patcha and J.-M. Park, "An overview of anomaly detection techniques: Existing solutions and latest technological trends," Comput. Netw., vol. 51, no. 12, pp. 3448–3470, Aug. 2007.
12. W. Li, V. Mahadevan, and N. Vasconcelos, "Anomaly detection and localization in crowded scenes," IEEE Trans. Pattern Anal. Mach. Intell., vol. 36, no. 1, p. 18–32, Jan. 2014.

2004 and B.E. in CS&E from University B.D.T College of Engineering (UBDTCE), Davangere in the year 1998. His major areas of research are Data Mining & Knowledge Discovery, Big data, Information Retrieval, and Cloud Computing. He is a Life member of India Society for Technical Education (LMISTE), Computer Society of India (CSI) & International Association of Engineers (IAENG).

## BIOGRAPHIES



Sheetal Nadagoudar presently pursuing her M.Tech degree in department of studies in CSE at Visvesvaraya Technological University, PG Center, Mysuru and has completed B.E in CSE at Basaveswar Engineering College Bagalkot, Karnataka in the year 2010. She has worked as Assistant Lecturer in Mount Farhan diploma college and Jain college in Hubli. Her M.Tech project area is Data Mining. This paper is a review paper of her M.Tech project.



Dr. Thippeswamy K, currently working as Professor and HOD, Dept. of Computer Science & Engineering, VTU PG Centre, Mysuru received his Ph.D. degree from the Department of CS&E, Jawaharlal Nehru Technological University, Ananthapur, Andhra Pradesh in the year 2012, M.E degree in CS&E from University Visvesvaraya College of Engineering (UVCE), Bangalore in