International Journal of Scientific Research & Engineering Trends
Volume 5, Issue 3, May-Jun-2019, ISSN (Online): 2395-566X

Financial Fraud Detection with Anomaly Feature Detection on credit card

M Anjaneyulu
Dept. of CSE
Vaishnavi Institute of Technology
Tirupati, AP, India

Asst. Prof. A. Uday Kishore
Dept. of CSE
Vaishnavi Institute of Technology
Tirupati, AP, India.

***Abstract*- The most traditional payment approach is a credit card or debit card for online in today's world; it will supply money less shopping at every shop across the world. It is the more and most suitable way to do online shopping, paying bills, and performing other related tasks. Hence the risk of fraudulent transactions using a credit card has also been increasing. In the existing credit card fraud detection processing system, the fraudulent transaction is detected after the transaction is done. This kind of illegal activities involves complex networks of business, enterprise and financial transactions, which makes it difficult to detect the fraud entities and discover the features of fraud by issuing authorities. Fortunately, trading, business transaction network and features of entities in the network can be constructed from the complex networks of the trade and financial transactions. The trading or business transaction network announces the interaction between entities, and thus anomaly detection on business networks can expose the entities involved in the fraud activity. However, most of the existing methods focus on transaction networks or features information separately, which does not use the information. In this paper we propose a novel credit card fraud detection framework based on Fraud Behaviour which reflects the cardholders' transaction habits using data mining techniques and we propose a novel fraud CoDetect, which can leverage both network information and feature information for financial fraud detection.**

***Keywords*- Anomaly feature detection, Co Detect, financial fraud.**

I. INTRODUCTION

Nowadays the modes of payment methods are changed into online transactions. Banking system provides different type of payments like e-cash, card payments, internet banking, and e-services for improving online transaction. Credit card is one of the most custom ways of the online transaction. Credit cards are used for purchasing goods and services using online transaction and physical card for the offline transaction. In credit card based purchase, the cardholder presents his card to a merchant for making payment. To make fraud in this kind of acquisitions, the person doing fraud has to steal the credit card. If the legitimate user does not understand the loss of card, it can lead to important financial loss to the credit card company and also to the user.

Credit card is a medium of selling goods or services without having cash in hand. With more number of such money less transaction, a number of fraudulent transactions also increasing? During the online transaction, we do not need any physical card; we need only card number, cvv number, and expiry date so there are more chances of fraud will be happen. In this method of fraud detection, we generate fraud behavior on the basis of cardholder's transaction habits. Most of the credit card fraud detection methods based on anomaly detection try to extract the historical behavior patterns as rules and compute the similarity between an incoming transaction and these behavior patterns. The main idea of this kind of approach is that people may have personalized transaction habits that depend on their different accounts, different income sources, and different motivations and so on.

II. RELATED WORK

Financial fraud detection concerns about the detection of fraud in insurance, credit card, telecommunications and other financial crime activities such as money laundering. Statistical models have been used for detection of financial fraud. Bahnsen et al. improve the detection performance by calibrating probabilities before establishing Bayes model. HMM model is used to model the customers' credit card shopping patterns for detection of credit card fraud. The shopping items indicate the hidden state and the corresponding prices from certain ranges are the observation. LR(Logistic Regression), Support Vector Machines(SVMs) and Random Forest(RF) are evaluated for credit card detection. The detection models are built on primary features and derived features from transaction. Whit row et al. proposed a new pre-processing strategy for better fraud detection with SVMs and KNN classification. Transactions aggregated in term of time window, then data with new features is used to model the pattern. Wei et al. addressed the problem of unbalanced financial data and employed cost-sensitive neural network to punish the misclassification of fraud transaction.

III. EXISTING SYSTEM

Due to the increasing popularity of the Web, there are rising number of people who performs e-business transactions on web. On the other hand, this popularity has also attracted the attention of criminals, raising the number of fraud cases in Web and financial victims that reach billions of rupees per year. This paper proposes a methodology, based on the knowledge discovery process, to detect fraud in online payment systems. Many fraud detection models work with attribute-value that is generated from transactions data. Some aggregation methods are also used to enrich the information of data. After generating feature points from transactions, supervised and unsupervised methods can be used to perform detection. Usually, these attribute values are assumed to be independent and identically distributed. However, the characteristic of money laundering is different from attribute-value data. Linked data is clearly not independent and identically distributed, which contradicts the assumptions of traditional supervised and unsupervised methods.

On the other side, some linked data is auto correlated. For example, trading business between business entity A and business entity B implies that feature points A and B are concurrent. Some features used to describe the properties of trading business goods can be same between A and B. These characteristics of auto correlation decrease the efficient size of data for learning. Furthermore, feature points don't fuse the interaction information in data. The relations between any business entities indicate the potential causality that means, if businesses on going, fraud entity can be located by other identified fraud entity. This means the entity, which has connection with fraud entity, are suspicious.

Consequently, feature based detection models with supervised or unsupervised methods have inherent limitation of incapacity of identifying what the fraud relations are. Graph-based mining methods are one of the most important theories that attempt to identify relations between attribute values. With the fraud manners detected by graph-based detection technique we are able to draw the conclusion that several business entities involved in fraud, however, we still don't know how these fraud activities are operated and why these activities labelled as fraud, i.e., the detailed features of the fraud activities. Graph and attributes provides complementary data for financial fraud activity detection and fraud property tracing. However, the majority of the existing algorithms exploits these two pieces of information separately and thus cannot provide a system that can detect the fraud entities and reveal suspicious properties for easy tracing simultaneously.

IV. PROPOSED SYSTEM

Here the new proposed scheme called Financial Fraud Detection with Anomaly Feature Detection on credit card is introduced. In this paper, we would like to develop a novel framework for fraud detection by considering the special detecting and tracing demanding of fraud entities and behaviors. Specifically, we investigate: (1) how to utilize both graph matrix and feature matrix for fraud detection and fraud tracing; (2) how to mathematically model both graph matrix and feature matrix so as to simultaneously achieve the tasks of fraud detection and tracing. In an attempt to solve these challenges, we proposed a novel detection framework Co Detect, as Fig. 1 shown, for financial data, especially for money laundering data. We incorporate fraud entities detection and anomaly feature detection in the same framework to find fraud patterns and corresponding features simultaneously.

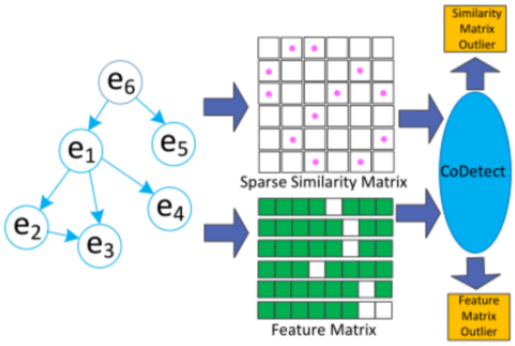

Fig.1 Fraud detection using graphs mining techniques for proposed framework.

1. Anomaly Detection

Financial fraud detection only focuses on a particulars domain: financial activities. Anomaly detection tries to nd patterns in data that is unusual seen or out of expectation. So anomaly detection can be seen as a general form of fraud detection. Fraud detection is one application of anomaly detection [4]. Two techniques are most related to fraud detection. One is one-class classification. Another one is clustering based outlier detection. One-class classification usually based on the assumption that the detection model is built on data which is generated from one or several statistical distributions.

V. EXPERIMENTS

In this paper, the synthetic data and real world data are used to evaluate the effectiveness of CoDetect. We first perform qualitative analysis using synthetic data to demonstrate the detection result in an expressive way. Then we calculate Co Detect with other state-of-art matrix factorization methods and clustering methods in term of detection accuracy and detection time. Finally, we perform the model parameters analysis which proves the robustness of Co Detect.

1. Financial Data Sets And Preprocessing Synthetic Data- Technically, the synthetic data is from small part of ICIJ Offshore Leaks Database. We only extract 100 financial entities and 2,000 transactions from this data set. Then we inject fraud patterns into this synthetic data.

2. Money Laundering Data- This data set is from ICIJ Offshore Leaks Database. We filter out uncompleted rows from the data set which leaves us a data set with 29,265 financial entities, and 571,113 transactions.

3. Insurance Fraud Data- This data set is from insurance company benchmark data set [45] which has 86 attributes for each customer records. Reviewing from attribute 65 to 85, we know that each customer can under subset of insurance policies.

4. Credit Card Fraud Data- German Credit Data set is used in our study. The pre-processing is similar to the pre-processing of COIL2000. In German Data, attribute 4, qualitative is used to form the bi-party graph from data set where there is a connection if customer ran their credit card for the purpose in attribute 4. Then we have the matrix S and F.

VI. CONCLUSION

We propose a new framework, Co Detect, which can perform fraud detection on graph-based similarity matrix and feature matrix simultaneously. It introduces a new way to reveal the nature of financial activities from fraud patterns to suspicious property. Furthermore, the framework provides a more interpretable way to identify the fraud on sparse matrix. Experimental results on synthetic and real world data sets show that the proposed framework (Co Detect) can effectively detect the fraud patterns as well as suspicious features. With this codetection framework, executives in financial supervision can not only detect the fraud patterns but also trace the original of fraud with suspicious feature.

REFERENCES

[1] C. Sullivan and E. Smith. ‘‘Trade-Based Money Laundering: Risks and Regulatory Responses,’’ Social Sci. Electron. Publishing, 2012, p. 6.

[2] United Press International. (May 2009). Trade-Based Money Laundering Flourishing. [Online]. Available: lourishing/UPI17331242061466.

[3] L. Akoglu, M. McGlohon, and C. 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, and V. Kumar, ‘‘Anomaly detection: A survey,’’ 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. Lin, M. Chen, and Y. Ma. (2010). ‘‘The Augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices.’’ [Online]. Available: https://arxiv.org/abs/1009.5055.

[10] J. Sun, H. Qu, D. Chakrabarti, and C. Faloutsos, ‘‘Neighborhood 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.

Author’s Profile

M. Anjaneyulu Pursuing M. Tech at Vaishnavi Institute of Technology, Department Of CSE, Tirupati, AP, India.

A.Uday Kishore Working as a Assistant Professor & HOD in Vaishnavi Institute of Technology, Department Of CSE, Tirupati, AP, India.