Codetect Financial Fraud Detection With Anomaly Feature Detection

K. Narsimhulu[1*], Bantu Bhumika[2], Moturi Rohith[3], Karnati Krishna Vamshi[4], Kadaru Sai Kiran Reddy[5]

Abstract
Money laundering is one type of financial fraud that is well-known for diverting illegally obtained funds to terrorism or other criminal activity. Complex networks of trade and financial transactions are involved in this type of criminal activity, making it challenging to identify fraud firms and identify its characteristics. Fortunately, the intricate networks of trade and financial transactions may be used to build the trading/transaction network and the characteristics of the entities in the network. While features of entities are descriptions of the entities, anomaly detection on features can reflect specifics of the fraud activities, the trading/transaction network reveals the interaction between entities, making it possible to identify the entities involved in the fraud activity. Hence, network, the majority of approaches currently in use only use one of the two types of information, networks or characteristics. In this research, we offer CoDetect, a novel framework for financial fraud detection that can use both network information and feature information. Additionally, CoDetect is capable of detecting both the features associated with financial fraud activities and the fraud actions themselves concurrently. Numerous tests using both artificial and real-world data show how efficient and effective the suggested methodology is at stopping financial fraud, particularly money laundering.

Keywords: Financial fraud, Money Laundering, Transactions, Detection, Codetect, Complex networks, Novel detection Framework, Information.

[1*]Assistant Professor, Dept. of CSE, Sreyas Institute of Engineering and Technology, Nagole, Hyderabad.
[2,3,4,5]B.Tech students, Dept. of CSE, Sreyas Institute of Engineering and Technology, Nagole, Hyderabad.

***Corresponding Author:** K. Narsimhulu
*Assistant Professor, Dept. of CSE, Sreyas Institute of Engineering and Technology, Nagole, Hyderabad.

I. INTRODUCTION

The information contained in the data is also enhanced using aggregation techniques. Generating feature points afterwards financial fraud activities, such as credit card fraud and money laundering, have gradually increased in recent years. These actions result in the loss of personal and/or business property. Even worse, they endanger the security of nation because the fraud may go to terrorism. Thus, accurately detecting financial fraud and tracing fraud are necessary and urgent. However, financial fraud detection is not an easy task due to the complex trading networks and transactions involved. Taking money laundering as an example, money laundering is denied as the process of using trades to move money/goods with the intent of obscuring the true origin of funds. Usually, the prices, quantity or quality of goods on an invoice of money laundering are fake purposely. If we utilise these statistics as features to construct detection policy, the misrepresentation of prices, quantity, or quality of items on an invoice only exposes minor differences from normal basis.This type of detector might function effectively with somewhat stable trading entities in specific situations. Unfortunately, the situation in the real world is more difficult, particularly in Free Trade Zones (FTZs), where international trade necessitates intricate processes and information sharing between trading parties. The fraud activities are more covert, especially money laundering. The concealment of the movement of cash through trading operations, the purchase and sale of intangibles, and related party transactions are just a few examples of the various ways that money laundering can occur. The trading of goods demonstrates greater diversity, but so do the many types of businesses, including front and shell corporations that help with money laundering. Money laundering, in contrast to other forms of fraud, exhibits unique traits that pose a high danger to the financial system due to the concealment of the money trail, collectivist conduct, and wild trading zones in FTZs.

II. EXISTING SYSTEM

Graph-based mining methods are one of the most important theories that attempt to identify relations between data points. Financial activities can be modelled as a directed graph, then a sparse adjacent matrix can represent this graph. With graph-mining method, the sparse matrix can be approximated as summation of low-rank matrix and outlier matrix.

Thus, graph-based methods can detection suspicious interactions between entities while attribute-feature based methods can reveal the features of the fraud. Graph and attributes provides two complementary information for financial fraud activity detection and fraud property tracing.

III. PROPOSED SYSTEM

- 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 behaviours.

Specifically, we investigate:

- How to utilize both graph matrix and feature matrix for fraud detection and fraud tracing;
- 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 CoDetect, 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.
- Combining entities detection and feature detection enables us to build a novel fraud detection framework for noisy and sparse financial data: relevant fraud patterns help the identification of fraud identities, and relevant features in turn help revealing of the nature of fraud activities.

3.1 FUNCTIONALREQUIREMENTS:

• .Data Collection
• Data Pre-processing
• Training & Testing
• Modelling Predicting

3.2 SYSTEM ARCHITECTURE:

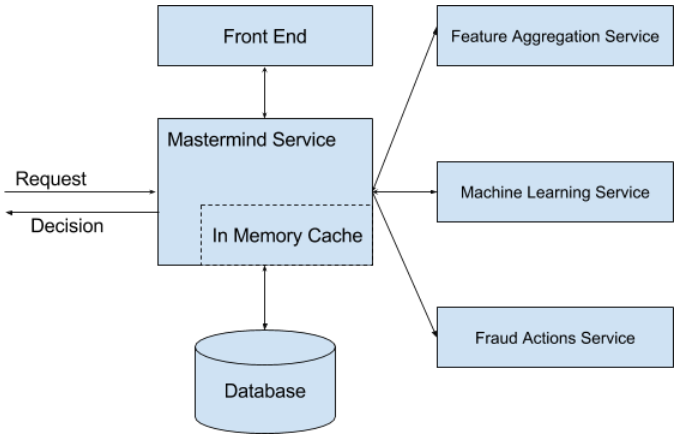

Fig.3.2 System architecture

3.3 DATA FLOW DIAGRAM:

The DFD is also called as bubble chart. It is a simple graphical formalism that can be used to represent a system in terms of input data to the system, various processing carried out on this data, and the output data is generated by this system.

1. The data flow diagram (DFD) is one of the most important modelling tools. It is used to model the system components. These components are the system process, the data used by the process, an external entity that interacts with the system and the information flows in the system 2. DFD shows how the information moves through the system and how it is modified by a series of transformations. It is a graphical technique that depicts information flow and the transformations that are applied as data moves from input to output.

3. DFD is also known as bubble chart. A DFD may be used to represent a system at any level of abstraction. DFD may be partitioned into levels that represent increasing information flow.

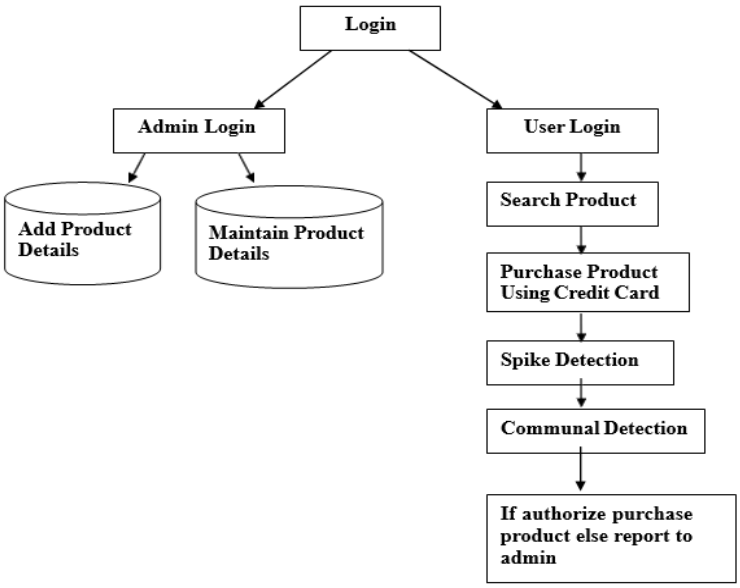

Fig.3.3Data Flow Diagram of System architecture

IV. IMPLEMENTATION

4.1 MODULES

There are four modules in this project.

1. USER
2. BANK ADMIN
3. TRANSPORT COMPANY
4. ADMINISTRATOR

BANK ADMIN:

Here bank admin is a module should register with the application, then admin must be authorized by the admin then only bank admin can login successfully into application after logged he/her can perform some operations such as view profile, add your bank details, view bank details, view credit card details, view credit card users, view all transport company booked ticket details, view all type of financial fraud and logout.

ADMIN:

Here admin is the main module can directly login with the application and can perform some operations such as view users, transport company, bank details and authorize view all users, view all bank details, view transport details and logout.

USER:

Here user must take registration with the application and should authorized by the admin then only he can login with our application successfully after logged in he/her can perform some operations such as view profile, manage bank details, request credit card, view credit card , view payments and transfer to credit card account, view all transport company and book tickets by selecting company, view card transactions and logout.

TRANSPORT COMPANY:

Here transport must take registration with the application and should authorized by the admin then only he can login with our application successfully after logged in he/her can perform some operations such as add transport details, view transport details, fixed travel prices, view all booked ticket details, view type of financial fraud and logout.

V. RESULTS

Fig.5.1 Home admin page

This picture shows the home screen of project when we search for project title ingoogle search tool bar.

Fig.5.2 View profile

This picture shows the view profile page when user login into the website.

Fig. 5.3Add bank details

This picture shows the view bank details page as the user can go through it.

Fig.5.4 All types of financial fraud

Fig.5.5 All fraud details

This picture shows all the fraud details to the admin.

VI. CONCLUSION

We provide a brand-new system called CoDetect that can concurrently identify fraud using feature and similarity matrices based on graphs. It presents a brand-new technique for identifying the type of financial activity, such as fraud trends or suspicious assets. The framework also offers a more understandable technique to spot the fraud on a sparse matrix. Experimental findings on simulated and real-world data sets demonstrate the effectiveness of the suggested framework (CoDetect) in identifying fraud trends and suspicious traits. Executives in charge of financial supervision can use this framework for fraud detection to track out the source of fraud as well as fraud patterns.Financial transactions take time to complete. These activities can be represented by a similarity tensor and a feature tensor. So we would like to study how to integrate tensor into codetect framework for fraud detection.

VII. REFERENCES

1. C.Sullivan and E. Smith. ``Trade-Based Money Laundering: Risks and Regulatory Responses," Social Sci. Electron. Publishing, 2012.
2. L.Akogalu, M. Mc Glohon and C.Faloutsos,``OddBall: Spotting anomalies in weighted graphs," in Proc. Pacific-Asia Conf. Knowl. Discovery Data Mining, 2010, pp. 410_421.
3. V. Chandola, A. Banerjee, and V. Kumar, ``Anomaly detection: A survey," *ACM* Comput. Surv., vol. 41, no. 3, 2009, Art. no. 15.
4. W. Eberle and L. Holder, ``Mining for structural anomalies in graph-based data," in Proc. DMin, 2007, pp. 376_389.
5. 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.
6. 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.
7. S.Wang, J. Tang, and H. Liu, ``Embedded unsupervised feature selection," in Proc. 29th AAAI Conf. Artif. Intell., 2015, pp. 470_476.
8. 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.
9. 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.
10. A. Patch and J.-M. Park, ``An overview of anomaly detection techniques: Existing solutions and latest technological trends," Compute Network., vol. 51, no. 12, pp. 3448_3470, Aug. 2007. [11]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.
11. K. Henderson et al., ``It's who you know: Graph mining using recursive structural

features," in Proc. 17th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2011, pp. 663_671.
12. F. Keller, E. Miller, and K. Bohm, ``HiCS: High contrast subspaces for density-based outlier ranking," in Proc. ICDE, Apr. 2012, pp. 1037_1048.
13. D.Koutra, E. Papalexakis, and C. Faloutsos, ``Tensorsplat: Spotting latent anomalies in time," in *Proc. PCI*, Oct. 2012, pp. 144_149.
14. J.H.M. Janssens, I.Flesch, and E.O. Postma, ``Outlier detection with one-class classifiers from ML and KDD," in Proc. ICMLA, Dec. 2009, pp. 147_153.
15. N.A. Heard, D.J. Weston, K.Platanioti, and D.J. Hand, ``Bayesian anomaly detection methods for social networks," Ann. Appl. Statist., vol. 4, no. 2, pp. 645_662, 2010.
16. J. Tang and H .Liu ``Co Select: Feature selection with instance selection for social media data," in Proc. SIAM Int. Conf. Data Mining, 2013, pp. 1_9.
17. Z.He, X.Xu, and S. Deng, ``Discovering cluster-based local outliers," Pattern Recognition. Lett., vol. 24, nos. 9_10, pp. 1641_1650, 2003.
18. M. Gupta, J. Gao, C. C .Aggarwal, and J. Han, Outlier Detection for Tempo- ral Data (Synthesis Lectures on Data Mining and Knowledge Discovery). San Rafael, CA, USA: Morgan & Claypool, 2014.