Identifying Fraud in Mobile Money Transfers

Mary Donovan Martello
Bellevue University
mdonovanmartello@my365.bellevue.edu

Evan Edmunds
Bellevue University
eedmunds@my365.bellevue.edu

Ramizuddin Mohammed Shabuddin
Bellevue University
rmohammedshabuddin@my365.bellevue.edu

Samarah Lopez
Bellevue University
samlopez@my365.bellevue.edu

Kurt Stoneburner
Bellevue University
kstoneburner@my365.bellevue.edu

ABSTRACT

Mobile money transfer services (MMTS) are currently being deployed in many markets across the world and are widely used for domestic and international remittances. However, MMTS transactions are highly susceptible to fraud. Fraud detection of MMTS and similar financial services is an important and significant endeavor. Although data mining and modeling techniques have evolved sophisticated probabilistic models to detect credit card and financial statement fraud, mobile money payment services still lack specific fraud detection solutions and research. We will discuss specific challenges related to fraud detection in MMTS transactions and the data approach and design solutions we developed to address those challenges. Finally, we will discuss the dynamic tri-clustering predictive model we developed to detect fraud in MMTS transactions, and the synthetic data set we used to train our model.

KEYWORDS

ACM proceedings;

I.6 Simulation and Modeling

I.6.5 Model Development

- Anomaly Detection

#Fraud; #Fraud Detection; #Financial Transfer; #Mobile Transfer; #Data Mining;

INTRODUCTION

Credit card fraud and financial statement fraud have been pervasive in our economy and society for many years. Such fraud has cost companies, consumers and society exorbitant amounts of time and money. For example, two of the largest credit card security breaches to date, TJX in 2007 and Target in 2013, are alone responsible for leaking over 80 million credit card numbers to scammers due to ineffective security of their point of sale systems. Mobile money transfer services (MMTS) are newer types of financial transactions that are currently being deployed in many markets across the world and are widely used for domestic and international remittances. MMTS transactions, like credit cards and financial statements, are highly susceptible to fraud. Fraud detection of MMTS transactions and similar financial services is an important and significant endeavor.

Data mining and modeling techniques have evolved sophisticated probabilistic models to detect credit card and financial statement fraud, and with the emergence of MMTS transaction probabilistic models are now necessarily being developed to detect fraud in MMTS transactions. We studied the predictive models used for credit card and financial statement fraud, as well as those used specifically for MMTS transactions, such as the RadViz visualization technique and Predictive Security Analyser. We then utilized a synthetic data set with a combined three clustering algorithm that is applied dynamically to attempt to detect fraud in MMTS transactions in real time.
.

CHALLENGES ADDRESSED IN OUR DATA APPROACH AND MODEL

Data scientist have challenges in developing predictive models that detect fraudulent financial transactions, including challenges of accessing large, imbalanced and available data for use to test and validate the models and the need to detect financial fraud in real time.

Figure 1: List of Models Currently in Use to Detect Financial Fraud

- Machine learning [14, 2, 1]
- Adaptive machine learning techniques (AMLTs) [18]
- Logistic regression [14]
- Support vector machines [14]
- Artificial neural network [14, 2]
- Bagging [14],
- C4.5 [14],
- Stacking [14]
- RadViz visualization techniques (the model coordinates a RadViz visualization view and a graph-based view) [13]
- Predictive Security Analyzer [16]
- Sequence Alignment [1]
- Fuzzy Logic [1]
- Genetic Programming [1]
- Artificial Intelligence [1]

DATA APPROACH

First, the data sets must be very large and highly imbalanced. [12]. The data sets must be very large because normally more than one million transactions are created daily. [12, 19]. The data sets are highly imbalanced due to the limited number of fraudulent transactions. For example, one data contained over 300 000 transactions in one day with only 5 incidents of fraud. [12] This results in the task of detecting very rare fraud dispersed among a massive number of genuine transactions. [12].

Second, there is a dearth of publicly available real financial transaction data to perform experiments on. This is principally due to privacy concerns surrounding financial transactions and the reluctance of financial institutions to share proprietary information. [15, 9, 19]. There is a lack of legitimate datasets on mobile money transactions to perform research on. The domain of fraud detection is a big problem today and is of great interest to the scientific community. [9, 19]

Our approach to address the dearth of large, imbalanced available data is to use a simulated synthetic data set to train and validate our predictive model. PaySim is a financial simulator that simulates mobile money transactions based on an original dataset [9]. This simulated data can be as prudent as the original dataset for research. [9]. Several research cases have used synthetic data. [18]

FINANCIAL FRAUD DETECTION MODELS

- There are many published papers discussing financial fraud detection predictive models that have been tested and evaluated. [15, 11, 12, 14, 20, 2, 18] Fig. 1 displays several predictive models, tools and techniques that are addressing the issue of fraud detection [1]

Models like these are a necessary to combat fraud proactively, rather than dealing with its effects, reactively.

DYNAMIC MODELS FOR REAL TIME DETECTION

Fraud detection needs to be in real time and consider the fact that the interval between a customer initiating a transaction and the payment being transferred to its destination account is usually very short [12]. In order to prevent instant money loss, a fraud detection alert should be generated as quickly as possible. This requires a high level of efficiency in detecting fraud in large and imbalanced data [12]. Dynamic models are needed to create this efficiency in detecting fraud in real time [12, 19].

Figure 2: Basic Structure of a Fraud Detection Algorithm

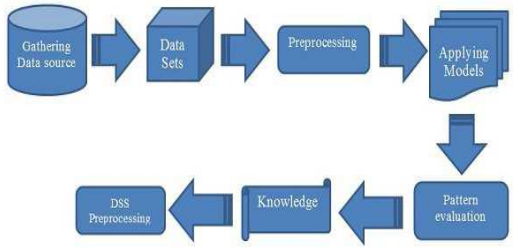
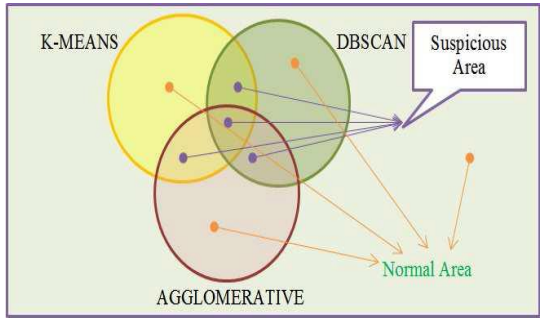


Figure 3: Proposed Model of Clustering Diagram



OUR PREDICTIVE MODEL

We developed a dynamic predictive model that was trained, tested and validated with the PaySim synthetic data set. It is designed as a dynamic model to instantaneously detect fraud in MMTS transactions. The dynamic nature of the model enables the user to detect and raise alarms for fraudulent transactions in real time. Our model shown in Fig. 3:

- Uses a combination of three unsupervised clustering algorithms simultaneously for improved prediction accuracy and fast real time analysis
  - Each algorithm might use all or some parameters of the prepared data set
  - Each algorithm does its own clustering analysis so that the transaction is tested from different perspectives to make sure the detecting process works optimally
  - The new transaction is grouped with 100 of the most recent transactions for the clustering analysis
  - A transaction is alerted as suspicious if it (1) takes place in shared space between at least two algorithms and (2) is in a single node or is in clusters with minimum members and high local outlier values
- A suspicious transaction predicted by at least two of the clustering algorithms is instantaneously examined by a decision support system, which allows real time action on the altered transaction
- Has a 68.75% fraud detection accuracy for dynamic online modeling

CONCLUSIONS

Losses from online payment fraud is predicted to more than double by 2023, reaching $48 billion annually. Developing methods to detect mobile money transfer fraud is essential to combat fraud attempts on these transactions. Dynamic modeling techniques are an important component of preventing mobile money transfer fraud as they find patterns within extremely large imbalanced data sets that could not be observed otherwise and help provide a solution in real time (or near real time). Our tri-clustering dynamic predictive model was trained on a large imbalanced synthetic data set and will detect mobile money fraud in real time and will assist in reducing fraud in mobile money transfer services.

ACKNOWLEDGEMENTS

The library resources of Bellevue University provided access to material vital to the project. Dr. Shankar Parajulee provided insights into presentation and scope of data presented. Kaggle.com, KDNuggets, and the ACM website provided auxiliary tools for project design and development.

REFERENCES

[1] Chaudhary, K., Yadav, J., & Mallick, B. (2012). A review of Fraud Detection Techniques: Credit Card. *International Journal of Computer Applications*, 39-44.

[2] Choi, D., & Lee, K. (218). An Artificial Intelligence Approach to Financial Fraud Detection under IoT Environment: A Survey and Implementation. *Security & Communication Networks*, 1-15.

[3] Coman, D.-M., Coman, M.-D., and Horga, M. 2014. Information Technology for Fraud Detection. *Valahian Journal of Economic Studies 5*, 3: 85–92.

[4] Demiriz, A. and Ekizoglu, B. 2016. Fuzzy rule-based analysis of spatio-temporal ATM usage data for fraud detection and prevention. *Journal of Intelligent & Fuzzy Systems 31*, 2: 805–813.

[5] Dong, W., Liao, S., and Zhang, Z. 2018. Leveraging Financial Social Media Data for Corporate Fraud Detection. *Journal of Management Information Systems 35*, 2: 461–487.

[6] Drogalas, G., Pazarskis, M., Anagnostopoulou, E., and Papachristou, A. 2017. The effect of internal audit effectiveness, auditor responsibility and training in fraud detection. *Journal of Accounting and Management Information Systems 16*, 4: 434–454.

[7] Han, H. C., Kim, H., & Kim, H. K. (2016). Fraud Detection System in Mobile Payment Service Using Data Mining. *Computer Science*.

[8] Landgraf, E.L. and Nebbia, T.L. 2013. An Analysis of Recent Published Fraud Surveys. *Journal of Forensic Studies in Accounting & Business 5*, 1: 11–37.

[9] Lopez-Rojas, Edgar Alonso, E., Elmir, A., & Axelsson, S. (2016). PAYSIM: A FINANCIAL MOBILE MONEY SIMULATOR FOR FRAUD DETECTION. *28th European Modeling and Simulation Symposium2016*, (pp. 1-7). Larnaca.

[10] Kopun, D. 2018. A Review of the Research on Data Mining Techniques in the Detection of Fraud in Financial Statements. *Journal of Accounting & Management 8*, 1: 1–17.

[11] Kou, Y., Lu, C., Sirwongwattana, S.; Huang, Y. (2004). Survey of fraud detection techniques. *International Conference on Networking, Sensing and Control*. Taipei: IEEE.

[12] MINASTIREANU, E.-A., & MESNITA, G. (2019). An Analysis of the Most Used Machine Learning Algorithms for Online Fraud Detection. *Informatica Economica*, *23*(1), 5–16. https://doi-org.ezproxy.bellevue.edu/10.12948/issn14531305/23.1.2019.01

[13] Novikova E., Kotenko I. (2014) Visual Analytics for Detecting Anomalous Activity in Mobile Money Transfer Services. In: Teufel S., Min T.A., You I., Weippl E. (eds) Availability, Reliability, and Security in Information Systems. CD-ARES 2014. Lecture Notes in Computer Science, vol 8708. Springer, Cham M

[14] Perols, J. (2011). Financial Statement Fraud Detection: An Analysis of Statistical and Machine Learning Algorithms. Auditing: A Journal of Practice & Theory, 30(2), 19–50. https://doi-org.ezproxy.bellevue.edu/10.2308/ajpt-50009

[15] Phua, C., Lee, V., & Gayler, R. (2010, September 30). *Computer>Science*. Retrieved from Cornell University: https://arxiv.org/abs/1009.6119

[16] Rieke, R., Zhdanova, M., Repp, J., Giot, R., & Gaber, C. (2013). Fraud Detection in Mobile Payments Utilizing Process Behavior Analysis. *International Conference on Availability, Reliability and Security* (pp. 662-669). Regensburg: IEEE.

[17] Siering, M., Koch, J.-A., and Deokar, A.V. 2016. Detecting Fraudulent Behavior on Crowdfunding Platforms: The Role of Linguistic and Content-Based Cues in Static and Dynamic Contexts. *Journal of Management Information Systems 33*, 2: 421–455.

[18] Singh, A., & Jain, A. (2019). An Empirical Study of AML Approach for Credit Card Fraud Detection–Financial Transactions. *INTERNATIONALJOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL*, 670-690.

[19] Vadoodparast, M., Hamdan, A. R., & Hafiz. (2015). FRAUDULENT ELECTRONICTRANSACTION DETECTION USING DYNAMIC KDA MODEL. *International Journal of Computer Science and Information Security*.

[20] Zhou, X., Cheng, S., Zhu, M., Guo, C., Zhou, S., Xu, P., . . . Zhang, W. (2018). A state of the art survey of data mining-based fraud detection and credit scoring. *MATECWeb of Conferences*, 189-204.