Detecting anomalies in Banking Transactions

Surya Putchala, Sreshta Putchala

Abstract—Fraudulent Financial transactions in Interbank fund transfers costs a lot of money to the banks and erodes customer trust. In order to provide for a robust security to protect customers and ensure that only authentic fund transfers occur from their accounts, Financial Institutions can utilize state-of-the art algorithms drawn from the fields of Machine learning and Statistics to augment the rule based engines that have been protecting customer's money. The availability of information about customers, financial institutions, countries, currencies provide rich landscape for identifying non-genuine transaction if they occur. In this article, we will cover how we could build "normals" so that Anomalies could be identified.

Index Terms — Fraud Detection, Anomaly Detection, Outlier Detection, Financial Fraud, Machine Learning, Artificial Intelligence, Risk Management, Network Analysis, Big Data

1. Introduction

Interbank transfer of funds is facilitated by worldwide financial messaging network like Society for Worldwide Interbank Financial Telecommunication (SWIFT). The SWIFT messaging system is widely used and by hackers proxying as real account holders to transfer money to destination accounts of their liking. For a financial Institution point of view, it becomes important to utilize the information within its purview to closely monitor fraudulent messages and pro-actively prevent incidences.

Traditionally, these transactions are identified by well-defined rules for thresholds of when a transaction is considered a threshold. As the fraudsters adopt novel methods, the rules engines will have to evolve and adapt newer ways of identifying the suspect transactions.

By being pre-emptive, a financial institution builds its reputation and improve confidence of the customers who transacts with them.

2. Anomalous transactions and Big data

Banking Transactional volume depends on size of the financial institution. Even a smaller institution with 10000 customers with 10 transactions a month could yield substantial volume. Although, actual fraudulent transactions are but a fraction of the total volume of transactions, it is imperative that a financial institution safeguards against suspicious transactions is from both its financial and customer perspectives. The ability to detect these transactions non-intrusively by sophisticated methods, thus play an important role in smooth fund transfers. In the ubiquitous world of information sources related to the consumers through Social Media, KYC, OFAC and various AML services and availability/adoption of Machine learning and advanced Analytical applications, detecting fraudulent transactions can become much more adaptive than ever.

Fraud is perpetrated in a variety of ways and the strategies of fraudsters change frequency. However, a method to detect financial transactions that are "out of the ordinary" or "anomalies" and reviewing them closely could help detect potential Fraud before it occurs.

Any solution for detecting transactional anomalies is a complex endeavour and depends on various external sources with different formats and various latencies. Besides, the transactional volumes can go well beyond million rows an hour. Infrastructure and algorithms will have to scale to meet such demands. Big Data frameworks could potentially meet scalability and performance of such sophisticated systems.

3. Approaches for outlier detection

Anomalies are detected primarily with the help of Outliers. We can consider a transaction to be a suspect or anomalous, if it deviates significantly (based on a performance indicator/s) from the "normal". Hence establishing "normals" becomes critical in devising an Anomaly/Fraud detection system.

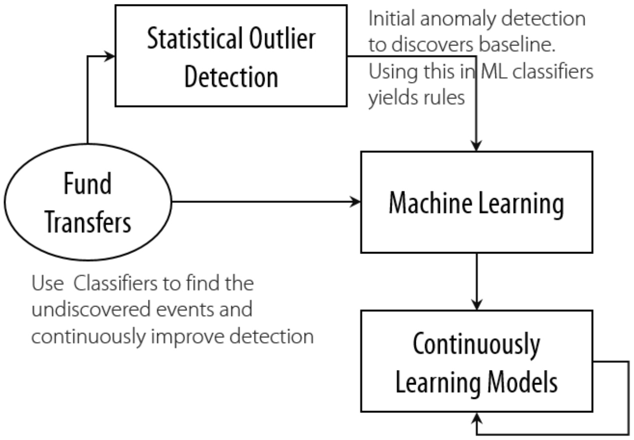
Statistical Outlier Detection
Initial anomaly detection to discovers baseline. Using this in ML classifiers yields rules
Fund Transfers
Machine Learning
Use Classifiers to find the undiscovered events and continuously improve detection
Continuously Learning Models

Figure : Outlier detection with Statistical and ML Techniques

Single Criteria Outliers: An Indicator based on a single variable, like "Amount" transferred can be considered anomalous with respect to the rest of the Amounts that were ever sent from a Sender could be considered as single criteria outlier.

Multi-Criteria Outliers: If a combination indicator such as Sender-Receiver "amount" is profiled and if the combination falls outside the "normal" transactional range between these two, could be considered anomalous. The combination represents a specific context or condition. These are also called contextual outlier or Multi-variate outlier.

Population outliers: If transaction characteristics are anomalous with the entire population and violate "Normals" is a population outlier.

There are different strategies establishing "normals". Some of the sophisticated methods that could potentially help in achieving this objective is by identifying patterns using

Artificial Intelligence, Machine learning and Statistical Analyses.

3.1 Statistical Analyses

Identifying an observation as an outlier depends on the underlying distribution of the data. Here we focus on univariate data sets that are assumed to follow an approximately normal distribution. The box plot and the histogram can also be useful graphical tools in checking the normality assumption and in identifying potential outliers. It is common practice to use Z-scores or modified Z-score to identify possible outliers. Grubb's test is a recommended test when testing for a single outlier.

3.2 Machine Learning

Many applications require being able to decide whether a new observation belongs to the same distribution as existing observations (it is an inlier), or should be considered as different (it is an outlier). Often, this ability is used to clean real data sets.

Supervised Outlier detection: Techniques trained in supervised mode assume the availability of a training data set which has labelled instances for normal as well as outlier class. Typical approach in such cases is to build a predictive model for normal vs. outlier classes. Any unseen data instance is compared against the model to determine which class it belongs to.

Unsupervised Outlier detection: It detects outliers in an unlabelled data set under the assumption that the majority of the instances in the data set are normal by looking for instances that seem to fit least to the remainder of the data set.

4. Defense against Fraud

An Fraud system should use a hybrid approach for anomaly detection – both in real-time and batch mode. The models should be finely synchronized to provide a reliable/robust and performant system.

It uses eclectic methods and techniques like:

4.1 Rules Engine

The common patterns of known Fraud patterns and their thresholds are setup Subject matter experts. Then, rule "pattern matcher" detects these predefined suspicious behaviours; whilst the "sequence matcher" finds temporal relationships between events from market data which exists in a potential violation pattern.

The thresholds of the rules can be based on probabilistic models and can be updated periodically. The updates are a function of the number of Consumer accounts and Volume of transactions. They are also changed when new set of rules are derived (from machine learning) or new rules are imposed by regulatory authorities.

4.2 Assessing Risk

Risk is accessed at different levels – Sender, Sender Bank, receiver, receiver bank, network properties of different entities in a transaction. This profiling is done based on individual characteristics as well as peer group profiling. For this, historical transactions and consumer profiles are used. There can be other methods like "segmentation" of various groups based on risk parameters and entities. If a new sender or any other entity is added, they are defaulted to a "normal" category until their transactional activity has started.

4.3 Predictive Models

Various combinations of Heuristic (Benford's law), Statistical, Machine learning (Supervised, Unsupervised, Reinforcement) and deep learning models (neural networks) can be used to identify Anomalies.

4.4 Novelty Detection

The system also has the ability to ascertain "Noise" and denoise the signals as the learning instances grows, which in turn improve the supervised learning algorithms. Besides this, the system has "novelty detection" algorithm to detect potential anomalies.

4.5 Network Analysis

Analysing the Network for with each individual a customer transacts. The nodes and connections are enhanced and various network statistics such as betweenness, closeness are studied. The network nodes are either "named" or "unnamed". A Customer's network should be studied to assess his activities in terms of his usual transactional accounts. His next level contacts, unusual/first time transaction entity.

5. Components of the Fraud Detection System

Historical data of the past 2-3 years data of a customer should yield a good behavioural profile. However, it also depends on the overall transaction volume. We should extract required features, create networks, derive various risk factors and build profiles. We need the following:

5.1 Customer Demographics

Demographic information like age, gender, ethnicity, geolocation, salary/income, home ownership (length of residence, home size, mortgage), education level, dependent children, type of cars, marital status, net worth/savings can be obtained from both internal and publicly available sources about the customers. This information can be beneficial in modelling fraud propensity of various customer segments.

5.2 Account details

Tenure, Linked Accounts, Overall activity with this bank, address and location changes. Techniques like ABC Analysis will yield information about the transactions that could warrant additional scrutiny.

- *Transaction details:* details of the Sender, Receiver, Sender Bank, Receiver Bank, Country of Origin, country of destination, Currencies of exchange. The volumes of transactions and their spread will give key insights about the normalcy measures.
- *Customer Watchlists:* Watchlists are often obtained from Public and Consortium data (Experian, radaris,

OFAC, CFTC, KYC etc.,) that could flag or highlight the risk of a transaction.

Generally, it is easy to identify transactions outside of normal business hours, high risk countries, sudden activity in a dormant account etc. However, little more sophisticated safeguards can be achieved by implementing:

- Client/Account dependent (Customer behaviour)
- In relationship to overall transactions (Customer Behaviour in relationship to the overall transactional behaviour)
- Demographic risk factors (Customer and the entity he is transacting)

5.3 Profile Transaction activity

Time has a great impact in determining normalcy of a transaction. The behaviour should be studied w.r.t time sensitive features like, frequency of the transactions, binning of the transactions by time unit in a day, time between transactions etc., can significantly help in establishing User habits. There are correlations and comparisons that performed as a part of the model w.r.t. inter transactions and intra-transactions.

For each customer, the transactional profile based on Recency, Frequency, Monetary values are derived. A profile of each SWIFT user's message traffic based on its specific business activities and the countries, counterparties and currencies it is typically involved with shall be developed.

The solution can accommodate binning time at any granularity – from hourly to yearly. The model can be configured (time granularity) at the time of implementation of the solution based on the volume of transactions and the risk tolerance of the Bank.

- Velocity is the calculated by the average time between transactions. binned by the hour, day, week, month)
- Volumes Day of the week: Week of the transactions (7days by count/amount of transactions)
- Binning by Time of the day: Time of the day transactions (24 hours by count/amount of his transactions)

5.4 Segment behavior

As part of customer profiling, the historical transactions are processed to build the profile of a customers. The information like the beneficiaries to whom the customer sends payments, amount of transactions, frequency of transactions done in a month etc. are used to build these profiles.

Profiles and aggregates with different combinations of nominal variables such as counterparty relationships and payment flows, Currency, country and counterparty activity breakdowns, reviewing large or unusual transaction values and volumes could highlight risks of Unusual patterns in payments.

The above normals could be studied to establish various segments of customers and entities, which can in turn be used to model behaviors.

5.5 Risk Profile

We need to be able to profile the risk of various entities involved in bank financial transactions.
Customer's Risk profile. The risk rating rules associated with each demographic risk parameters like nationality, line of business etc., it derives the score for each of the customer. Customer Risk Profiling and establish transactional thresholds to establish "normal". These are periodically performed.

Customer pattern changes are also identified….and new patterns of his transactions are detected and all changes are kept as reference. Any abrupt differences would be Anomalies. ML algorithms tries to establish "new normal" for each customer periodically.

A transaction's risk is evaluated based on:

- The Risk of the Entities Involved (Sender, Receiver, Sender Bank, Receiver Bank)
- Transactional Behaviour: The nature of the transaction whether it is normal or abnormal
- Institution Risk Profile: Institution wise peer profiling, which brings out the commonality across institutions.
- Network Analysis: Nature and the flow of transactions between entities

6. Putting it all together

A fraud or anomaly detection system should ideally integrate components s "Outlier" + Risk Score + Network Analysis. The models can be scheduled/configured to run at any window depending on the volume of transaction and need for recency of the models.

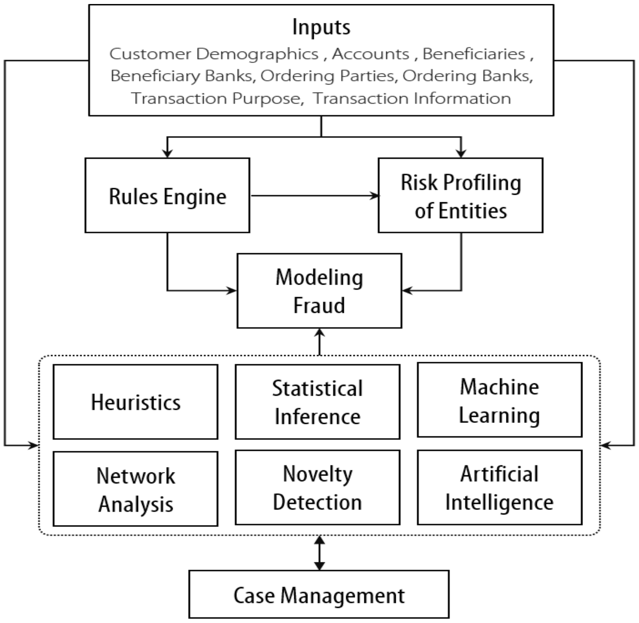
Inputs
Customer Demographics , Accounts , Beneficiaries , Beneficiary Banks, Ordering Parties, Ordering Banks, Transaction Purpose, Transaction Information
Rules Engine
Risk Profiling of Entities
Modeling Fraud
Heuristics
Statistical Inference
Machine Learning
Network Analysis
Novelty Detection
Artificial Intelligence
Case Management

Figure : Fraud Detection System

Since, the learning model uses eclectic approaches, the success of the implementation depends on the following factors:

1. Get timely feedback on the system detection performance from Compliance users (which helps the reinforcement models)
2. Ability to let the models incrementally learns (to retain the recency and enhance the relevance of the supervised models).
3. Scoring of various entities: The risk scoring of Customers as they transact

The general update of various rules depends on:

1. Optimization goals or Objectives of the Bank
2. Estimates based on the preliminary statistics of Customer statistics and transactional features

6.1 Self-adapting system

Fraud is often perpetrated in patterns that have no priors, a detection system should have the following properties:

1. A system with soft thresholds for establishing and interpreting "normals"
2. Ability to choose appropriate statistical and Machine learning models that fits the data
3. Have strong feedback mechanism capable of evaluating, changing and identifying change in its behavior

6.2 Fraud Optimization Goals

Goals for fraud detection vary, and that in turn affects the thresholds that are set up. Generally, optimization should pursue 2 Objectives: Anomaly driven and Budget driven.

Anomaly-driven situations are those in which you have a required rate of detection and must identify the true occurrence of fraud. Goals can be set up for:

- *True positives (Sensitivity)*
- *True Negative (Specificity)*
- *Accuracy rate*

Budget-driven fraud detection occurs when you have a limited budget for response, and you must determine how many anomalies and false alarms you can handle within that budget, setting the threshold to match.

- *Cost or Penalty of an incorrect detection*
- *False Positives*
- *False Negatives*

Often these goals are a fine balancing act between the customer satisfaction and customer protection.

6.3 Case Management

The detections are logged in as events and channelled to the users (depending on their roles) through email, SMS, Social Media alerts for appropriate action. The alerts will also provide information about the likelihood of the Fraud for a transaction. It provides appropriate details in easily consumable dashboard with information about the customer, his past history, demographics, the transaction details, possible reason for flagging. The cases can be accepted, further reviewed or rejected. Once the case is resolved, the learning from the action will be fed back as additional inputs either as abstracted as a feature or a tweak to the learning models.

7 Conclusion

Building an anomaly detection system for a financial institution is not an exact science. The techniques that are available are multiple. Although, establishing base normals for key entities are fairly easy, the modelling of this information for a Financial Institution has to take into account, practical considerations such as its risk appetite, resources/budget available for case reviews and management.

References

1. Building a Large-Scale Machine Learning-Based Anomaly Detection System Part (1- 3) – Anadot White Papers
2. Practical Machine Learning- A New Look at Anomaly Detection - Ted Dunning & Ellen Friedman
3. Anomaly detection for monitoring – Baron Schartz & Preetam Jinka
4. Fraud Detection Using Data Analytics in the Banking Industry, acl Discussion Whitepaper
5. A-Z of Banking Fraud 2016 - Temenos and NetGuardians whitepaper
6. The Dawn of Machine Learning for Banking and Payments – feedzai whitepaper
7. Fraud Detection by Monitoring Customer Behavior and Activities - Parvinder Singh, Mandeep Singh, International Journal of Computer Applications (0975 – 8887) Volume 111 – No 11, February 2015]
8. Siddhartha Bhattacharyya, Sanjeev Jha, Kurian Tharakunnel, and J. Christopher Westland. Data mining for credit card fraud: A comparative study. Decision Support Systems, 50(3):602-613, 2011. On quantitative methods for detection of financial fraud.

Surya Putchala provided thought leading consulting solutions to Fortune 500 Clients for over 2 decades. He is passionate about areas related to Data Science, Big Data, High Performance Cluster Computing and Algorithms. He held senior leadership roles with large IT service providers. He graduated from IIT Kharagpur.

Sreshta Putchala is a summer intern with ZettaMine Technologies. She has worked on Exploratory Data Analysis of over 5 Million SWIFT Transactions using various statistical methods using SQL and R. She is currently pursuing her Batchelor's degree in Computer Science from Chaitanya Bharati Institute of Technology (Osmania University). Her interests are in the fields of Big Data, Statistical Analysis, Machine Learning and Artificial Intelligence