Mobile Wallet Fraud – How Machine Learning Solution and Big Data can help Fraud Prevention and Losses

Kavita Dwivedi

Head Data Science, Infinite Sum Modeling

Abstract

One of the growing industries in last 2 years, thanks to Digital India Campaign and of course demonetization, has been the mobile wallet companies and their business. In the last half a decade or so , multiple mobile wallet companies have been growing some at revolutionary pace and some steadily , but what is alarming that the rate at which Fraudulent transactions have grown is much higher than at which the transaction rate has increased.

Digital payments have shown a consistent growth trend in the month of April as well against the number of transactions clocked in March, Reserve Bank of India data show. While cards and IMPS transactions have maintained the similar trend, mobile wallets which had fallen in March 2018 have shown a smart pick up in April 2018, having grown 11% in terms of number of transactions. According to RBI data, in April wallets showed 279 million transactions against 268 million in the previous month. This, however, is still lesser than 310 million transactions wallets clocked in February, the month before full KYC guidelines were enforced by the RBI. Further last year April, wallets had seen more than 320 million transactions, almost 15% higher.

This paper aims at building a robust end to end solution for Fraud Detection and Prevention (FDP) using Machine Learning algorithms. Fraud Detection is challenging as the moment you bring in a robust system to prevent fraud in any FI, the fraudsters work faster to identify the weaknesses or circumvent around its security and continue to commit fraud. Machine Learning algorithms by design are meant to learn from past data and make the system better as it is being used. Continuous usage of this ML algorithms coupled with a Fraud Bureau where multiple industries /companies share fraud list can go a long way in preventing and minimizing fraud losses. We will present and discuss the outcomes of different Machine Learning Algorithms. The paper will also talk about how block chains can revolutionize this industry by giving the power of individual data security in customer's hand.

The paper will also show how use of Big data in cloud environment can help detect and alert the highly suspicious transactions real time thus stopping this transactions to occur. This will require using R, Python, Hadoop, Hive and Spark to meet the requirement of mining petabytes of data in real time. It will also discuss how linking mobile transactions to Aadhar and Biometric security can be beneficial.

The paper will not only aim to look at different types of Fraud , but also suggest how payment services companies can work together with Banks , Other Fintech providers , Telecoms , Healthcare and Insurance to build a strong Fraud Bureau.

This will not only help the company's reputation and help them to grow their business, it will also instill a lot of confidence in the consumer to use the mobile payment services without the fear of being duped.

***Index Terms*— Fraud Management, Mobile Wallet, Machine Learning, Big Data , Payments Fraud, Blockchain**

I. Introduction

Mobile Payments have seen a significant surge in the last decade or so. Mobile Wallets introduced a new and lucrative way of payment that involves almost no physical device and that's why its growth is astronomical. Yes, it gets bank on your mobile. Globally Mobile wallets have been growing at a very sharp rate. By the year 2017, mobile transactions worldwide are projected to be nearly three quarters of a trillion dollars, 721.4 trillion, to be exact — according to Kount's count. This is unexpectedly high considering 2010's figures which registered $52.9 billion, which subsequently doubled the next year. Mobile transactions then grew at roughly 50 billion-60 billion a year for the next few years.

In India mobile payments have seen growth in last 5 years. Digital India and then Demonetization led to a surge in mobile payments/wallets post which we saw a sudden inflow of new customer segments and merchants entering the mobile wallet industry. Merchants across India within a span of weeks adapted to mobile wallets and new consumers too where happy with this Digital money age. Today, the market of combined credit cards is smaller to the mobile wallet market in India. The Reserve Bank of India (RBI) data speaks of 19.9 million credit cards issued by 55 scheduled commercial banks in India till October 2014 in which HDFC Bank has issued the highest number of credit cards – 5.6 million – followed by 3.3 million by ICICI Bank.

In a very short time the size of mobile wallet market in India grew significantly. According to a study by research firm RNCOS, the current Indian market size for mobile wallet (m-wallet) stands at about Rs 350 crore and is estimated to rise to Rs 1,210 crore by 2019.

At present, mobile payments (including Credit Cards, Debit cards, Net payments, Payments Wallet) form a minuscule part of the overall digital payments industry in India. However, the contribution from phones and tablets is expected to increase to 30 per cent by 2020. According to study "Mobile payments in India are estimated to grow from $86 million in 2011 to $1.15 billion in 2016, with a compounded annual growth rate (CAGR) of 68 per cent, according to estimates".

The largest mobile wallet player in India , Paytm today has over 100 million wallet users, which is double the penetration of Visa and Maestro combined (in India). Over and above this, in a country such as India and BRIC nations, the remittance market is immense. Financial institutions/e-commerce/lifestyle shops are coming up with their own wallet, as can be seen with the launch of SBI Buddy, ICICI Pocket and even BookmyShow's own wallet, to name a few. It will be interesting to see this space with telecom companies, fintech's, NBFC's all trying to make a space for themselves.

In all of these growth, what is of concern and immediate attention for both the mobile wallet players and consumers is the Fraud associated with the transactions. In the last half a decade or so , multiple mobile wallet companies have been growing at revolutionary pace and some steadily , but what is alarming that the rate at which Fraudulent transactions have grown is much higher than at which the transaction rate has increased. While the industry has enjoyed healthy revenue growth in the mobile channel, fraud has continued to migrate to the mobile channel as well.

According to "Mobile_Payments_and_Fraud_2017_Report" The share of merchants who could definitively state that fraud in the mobile channel is increasing reached 40 percent, up from just 23 percent last year. Although 42 percent of merchants aren't sure whether mobile fraud grew or shrank relative to their mobile volume, less than 8 percent said that fraud in the mobile channel declined as a percentage of total volume. The majority of merchants, 60 percent, believe browser-based mobile payments are those at greatest risk for fraud.

Fraud has both financial and hidden implications. The impact of Mobile wallet Fraud is not limited to just the financial loss that the company and customers have to face. The implications are much wider and spread to the extent of curbing your customer and business growth.

Fraud has the following key implications the greatest threat being you as a mobile wallet business may just cease to exist. Apart from the financial implications which are straight forward, there is a fear of being pulled up and threatened by regulators. The regulators based on current compliance laws can hand you a huge fine. The fine payment is just a tip of the loss that you would be staring at , once a red flag is raised against you by country's regulator , the reputational loss associated with it is huge. You start losing the trust and faith of both customers and merchants and in today's world it may mean you begin losing customers to other wallet competitors, and if the reputational loss is huge you may see a sudden fall in number of new acquisitions. This is how wallet customers are different from your traditional finance consumers.

II. INCREASING FRAUD IN MOBILE WALLET TRANSACTIONS

Fraud in the context of mobile money is the intentional and deliberate action undertaken by players in the mobile financial services ecosystem aimed at deriving gain (in cash or e-money), and/or denying other players revenue and/or damaging the reputation of the other stakeholders. This ecosystem is building up fast and the biggest challenge is that the way and means to commit Fraud is changing at a faster rate than we deploy newer Fraud Solutions. Today's Fraudsters are getting sophisticated, the in-house Fraud Prevention solution is just not enough to meet and cut through the challenges. What is needed today is real time Fraud solutions build using Machine learning algorithms and supported by a high velocity Big data Environment. The algorithms have to self-learn and should be able to detect and raise alert post screening and identifying latest Fraud Patters and Behaviors.

Frauds are of various types and the first key process is to identify the type and form of Fraud we are being subjected to.The different types of Fraud are :

Author Description: Kavita is an Analytics leader with 12 + years of core hands on experience having an excellent track record on Presales, Partner Management, Analytics Delivery and Team management across domains in World Class Organizations. Currently, she is heading the Data Science function at Infinite Sum Modeling. She is a Chemical Engineer by education followed by a Masters (Eco) from IGIDR. She is a seasoned analytics professional with work experiences across companies like Fair Isaac, Experian, Accenture, Infosys and Vodafone. Her vast experience in domains like Banking, Insurance, Telecom, Fraud and Risk Management give her the right kind of diversification. She has published papers in areas of Financial Econometrics and Social Media Analytics.

- *Identity Fraud*– 71 percent of merchants cite this type of fraud as their chief concern. Identity fraud is when fraudsters intercept sensitive data that is not properly protected and use this identity to make online or card-not-present purchases. In the case of mobile wallets, fraudsters physically steal mobile devices and use them to make unauthorized purchases. With account takeover and new account fraud detection, organizations attempt to discover unauthorized or fraudulent users posing as legitimate users

- *Loyalty fraud* – This can happen when fraudsters intercept loyalty programs or members' accounts for theft and transfer of points. There are also cases in which points are sold and transferred to others for monetary gain.

- *Friendly fraud* – This occurs when legitimate orders are disputed by the consumer, requiring merchants to refund payments (chargebacks). This form of fraud can be unintentional, with the consumer forgetting they placed the order, or one family member using another's payment card without permission. There are also cases where this is intentional fraud, with fraudsters placing orders and then claiming they never received the goods, enjoying both a refund and the purchased.

- *Organization enabled fraud* – Individuals within a wallet organization partner with individuals /agents outside to commit Fraud. This Fraud can be huge as being part of the system they are aware of Fraud Alerts and work out ways to commit /enable Fraud without getting noticed.

- *Agent Driven Fraud* – In a Mobile wallet company in US, 60% of total Fraud losses was Agent driven Fraud. Either by creating Fake accounts, breaking into someone else's account or newer ways of committing Fraud. These agent driven losses can be he as they are gateway to numerous transactions and their network can be huge and spread globally.

III. Methods and Procedures: MACHINE Learning Algorithms and Empirical Evidence

Fraud modeling is a complex and interesting challenge because of multiple reasons. This involves identifying a scarce transaction among numerous other legitimate and valid transactions. This is a rare event problem where the characteristics of the rare event changes very rapidly across time.it is an outlier analysis , anomaly detection , exception mining and imbalance data modeling all at one. So building a real time efficient and accurate model is fairly challenging. The other major challenge of this industry is non sharing of Fraud data across financial institutions and other organizations.

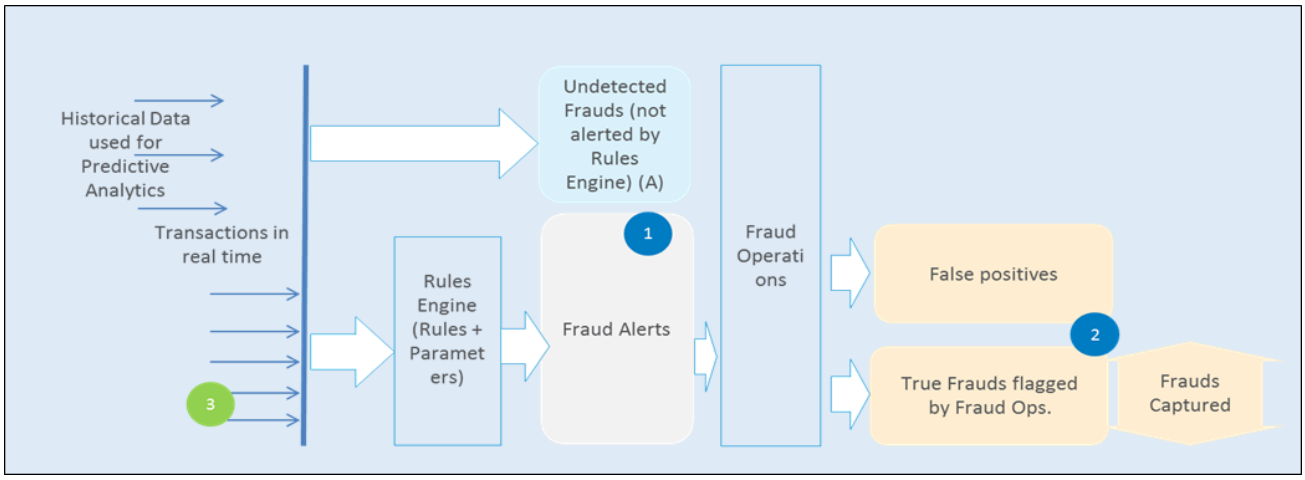
We need a strong consortium of banks globally today which will seamlessly allow sharing of Fraud data both in terms of individual behavior and sharing of pattern that we see when Fraud crimes are committed. National level Fraud Bureaus might be a good step but today in a global payment market that might not be enough to fight the menace of Fraud.

This paper will suggest and show empirical evidence of Fraud detection techniques on a sample data from a mobile payment industry. The fraud detection framework is based on the understanding of the Fraud problem and differentiates between Fraud Analysis and detection which is for known Frauds and also suggests ways and means to detect novel Fraud using ML techniques.

Data used was sample transaction data from a wallet transaction company, supplemented with synthetic data and was masked to prevent any data security issues. It contains fraudulent transactions over a period of 6 months and total transaction on a wallet in the same period.

The need of fraud detection in real time is the key and a mobile wallet transaction is completed within a few seconds. A smart and strong Fraud detection technique today should be able to run these models in real time and create a suspicious alert on the go. The key analysis is also the user behavior analysis which is very significant in case of individual or agent Frauds.

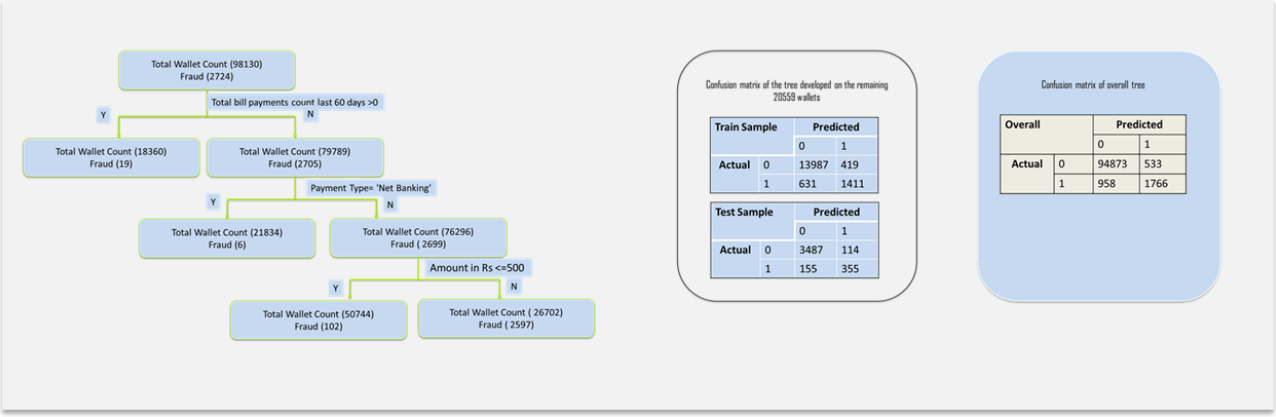
Fig 1. Fraud Prevention Framework



Detecting Known and Unknown Frauds

*Decision Tree – For Known Frauds :*Decision tree is the first supervised algorithm that was used to come up with rules to identify anomaly behavior. In supervised learning, samples of both fraudulent and non-fraudulent records, associated with their labels are used to create models.

Fig 2 :Decision Tree



Support Vector Machine (SVM) is a supervised learning model with associated learning algorithms that can analyze and recognize patterns for classification and regression tasks.SVM is a binary classifier. The basic idea of SVM was to find an optimal hyper-plane which can separate instances of two given classes, linearly.

Fraud detection systems are prune to several difficulties and challenges enumerated bellow. An effective fraud detection technique should have abilities to address these difficulties in order to achieve best performance. SVM has been successfully applied to a broad range of applications such as .In credit card fraud detection, “Ghosh and Reilly [61] developed a model using SVMs and admired neural networks. In a classification model based on decision trees and support vector machines (SVM) was constructed respectively for detecting credit card fraud. The first comparative study among SVM and decision tree methods in credit card fraud detection with a real data set was performed in this paper.” The results revealed that the decision tree classifiers such as CART outperform SVM in solving the problem under investigation. While applying SVM to our data set, we also observed that SVM gives higher accuracy in range of around 88 % which was suggesting that there is some over fitting and Decision trees are better.

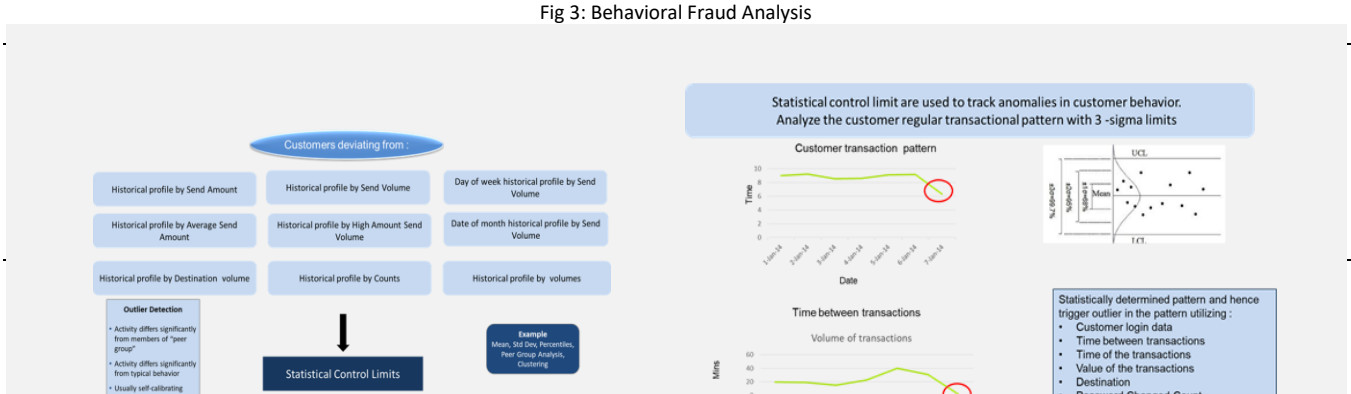
A Hidden Markov Model is a double embedded stochastic process which is applied to model much more complicated stochastic processes as compared to a traditional Markov model. The underlying system is assumed to be a Markov process with unobserved states. In simpler Markov models like Markov chains, states are definite transition probabilities are only unknown parameters. In contrast, the states of a HMM are hidden, but state dependent outputs are visible. HMM can also be embedded in online fraud detection systems which receive transaction details averify whether it is normal or fraudulent .If the system confirms the transaction to be malicious, an alarm is raised and related bank rejects that transaction. HMM m0del is known to give high False positive alerts and in our case we saw that accuracy of this model was around 62%.

Behavioral Analysis of Fraud Transactions – Ways to detect Unknown Frauds

Behavioral analytics focuses on observed characteristics of who the user is, not just who they tell you they are. It continuously profiles users and accounts through their entire lifecycle across multiple channels, including: desktop and mobile Web and native apps. Continuously profiling users' behavior empowers two key capabilities. First, it enables risk managers to detect and respond to risk sooner, reducing the chance of financial loss. Second, when the user does reach a transaction point, fraud managers have full context of all their previous actions and behavior to make a better decision on the transaction

The following approach was used to track agent transaction data and to detect anomaly in Agent’s behavior.The variables on the left hand side where used to track over a given period of time using Statistical Control Limit. This is Anamoly detection/Outlier detection and can be used to capture unknown Frauds.

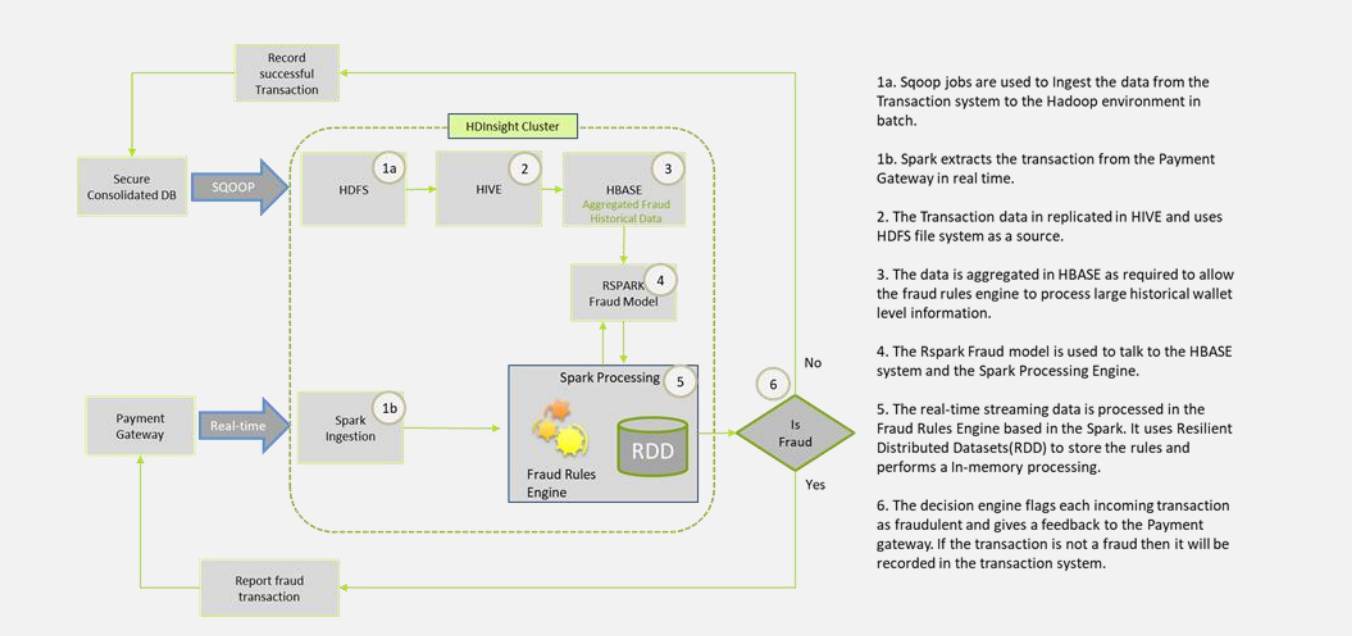
Fig 3: Behavioral Fraud Analysis



Big Data Technology a must have to detect Fraud real time

Most of our ML algorithms are built in SAS/R/Python. Mobile Wallet transactions can run into TB for a single mobile wallet provider for a country. We would need to pre-process the transaction data and run this ML algorithms using big data Envorinment.Below diagram gives a snapshot of the suggested Big data Architecture for Fraud Prevention System.

Fig 4: Suggested Big Data Stack



Block Chain: Technology Innovation that can change the way we fight Fraud

Blockchain is the strongest test case for Fraud Management .It is distributed digital ledger containing transaction data that is shared across a peer-to-peer network and continually reconciled. There is no central administrator or centralized version, so there is no single point of failure. Instead, management and authorization is spread across the network, so there is no obvious place for someone to instigate a fraud scheme. It uses shared digital ledger and increases the visibility and transparency of the transactions made throughout a supply chain and between members of a business network. Transactions recorded on Blockchain are immutable because they cannot be deleted or changed. A group of network participants must agree for the transaction to be valid through a process called consensus. Statista forecasts Blockchain technology to grow to 2.3 billion USD by 2021, and its prevalent that one of the factors fueling the growth of this technology is its capabilities to detect and prevent fraud. One of biggest use cases of Blockchain is the trustworthy footprint. Blockchain allows for participants to directly exchange information without any platform intermediating the same, essentially enabling mathematics to replace middlemen, thereby, improving efficiency and establishing trust. Implementation of Blockchain will have its own challenges but it will be the onus of Central Banking Authorities and Bureaus to facilitate this across banks, financial organization and other industries like telecom, Insurance etc.

IV. Conclusion

The paper talks about the growing menace of Fraud in Mobile Wallet and how the customers, merchants and wallet companies can come together to minimize Fraud Losses. It discusses about the different kinds of Fraud pertaining to mobile wallets and then uses a sample data to build Fraud Models. The study emphasizes on differentiating the Known and Unknown Frauds and how traditional models help in identifying the Known Frauds and newer /innovative solutions are needed to capture the Unknown Frauds. Decision Tree model provides the right kind of balance between accuracy and False positives but helps in detection of known Frauds. This coupled with Statistical Control Analysis of Behavior data can be used to raise Fraud alerts. Support Vector Machine showed very high level of accuracy and had issues of over fitting of data.The paper also suggests a Big Data Environment which is a must to process wallet transactions data in real time and raise alerts as transactions happen. This Infrastructure can be In-house infrastructure or a Cloud based environment. The author would continue the research on newer techniques that can decrease false positives and also increase the probability of capturing Unknown Frauds. The paper ends with its observation on How Block chain can help prevent Fraud by having a distributed environment and decentralizing authentication.

Acknowledgements:

This paper would not have been possible without the experience that I gained while working on various Fraud projects across organizations and discussing the wallet Industry Fraud with the industry experts. The author also thanks the Open Software like R and Python that helped run the Machine Learning Algorithms. It also has benefitted from the different prior papers in this area and will continue to carry the research going forward. Thanks to all the writers on Fraud that I would have read through in last 8 years online, papers, LinkedIn articles and media reports that keep me interested to investigate and find better solutions.

References:

1. Solving the "falsepositives" problem in fraud prediction Automated Data Science at an IndustrialScaleOct 2017
2. Roy Wedge and James Max Kanter and Kalyan Veeramachaneni Data to AI Lab, LIDS, MIT, Cambridge, MA-02139
3. Pascual, Al, M.-K., and Van Dyke, A. 2015. Overcoming false positives: Saving the sale and the customer relationship. In Javelin strategy and research reports
4. Fraud in Mobile Financial Services Joseck Luminzu Mudiri - A MicroSave Publication
5. Mobile_Payments_and_Fraud_2017_Report
6. Mobile_Payments_and_Fraud_2018_Report
7. Article: The Costs of Mobile Payment Fraud and How to Avoid It By Eran Feinstein in Business.com / Security / Last Modified: June 8, 2018
8. Article: Mobile Wallets: The New Fraud Frontier By Ryan Wilk, Director, NuData Security
9. A Survey of Credit Card Fraud Detection Techniques: Data and Technique Oriented Perspective -SamanehSorournejad1

, Zahra Zojaji, Reza Ebrahimi Atani ,Amir Hassan Monadjemi4