Graph-Based Anomaly Propagation Analysis in Point-of-Sale Networks: AI-Enhanced Fraud Detection in Retail Transaction Ecosystems

Dr. Małgorzata Pioro-Mianowska, Associate Professor of Computer Science, AGH University of Science and Technology, Poland

1. Introduction

The advent of digital transformation has ushered in an era of hyper-connected consumers, allowing for seamless transactions across multiple platforms. However, this rapid migration online has provided new opportunities for fraud, resulting in a 33% increase in fraudulent activities over the years since 2016. We chose retail transactions as our focus, as reports indicate that losses incurred in the retail industry doubled from 2019 onwards due to the pandemic and migratory fraud. The decentralized nature of retail networks has led to the unfortunate situation where often times, fraudulent transactions are discovered only after the goods are delivered or are in customs. This is highly risky, not just for the seller, but for the consumer as well. This increase in fraudulent activity has led to the loss of millions of dollars in retail sales around the world. Relying on manual processes to verify the identity of a potential fraudster or to vet the credibility of every transaction increases the time and costs pertaining to operational efficiency. The recent advances in technology — leveraging the Internet of Things, increased computational power, and artificial intelligence can help enterprises keep pace with the clandestine fraudsters. Furthermore, employing AI in fraud detection mechanisms could help guarantee a more efficient and effective way of curbing fraudulent activity, through continuous monitoring and reducing the noise associated with massive amounts of transactions.

1.1. Background and Rationale

Due to growing digital transformation, the number of online and in-store transactions is increasing rapidly in the retail industry worldwide. However, this increase is accompanied by a rise in fraudulent activities. These fraudulent activities and

cyberattacks on global retailers have caused billions of dollars in damages. Most retail organizations utilize traditional rule-based fraud detection systems. Such systems have limitations in detecting fraud in real-time. Beyond the income that retailers have lost, fraud has other negative consequences. It erodes customer confidence, damages the brand, ruins the experience, and sometimes has a negative impact on future revenue and profits. In addition to the retailers, fraudulent behavior affects financial institutions that issue credit or debit cards and customers who have been robbed of their assets fraudulently.

A weak rule-based fraud detection system has several limitations that prevent it from detecting fraudulent behavior. Hence, this has opened the door for researchers to propose the use of artificial intelligence in fraud detection. AI is able to automatically learn and improve from practice without being explicitly programmed. AI-based fraud detection systems are more powerful and reliable than rule-based systems. AI-based systems have many favorable characteristics, such as better impersonation, timeliness, and scalability. In this research, AI-based fraud detection for retail transactions with a large number of consumers was the target. It mainly focused on real-time fraud detection systems. Within the retail scenario, such research was anticipated to show evidence of early and effective recognition of possible fraud. Such aims were entirely of interest to retailers, as they were congruent with their objectives of reducing the likelihood of accepting a fraudulent transaction or losing future earnings. In fraud management, stakeholders have an advantage.

1.2. Scope and Significance of the Study

Fraud in retail transactions is a crucial problem for various stakeholders. In this study, our focus will be the enhancement of the capability of AI technologies and methodologies to detect fraud or fraudulent activities in retail transactions. The scope of the study is delineated by identifying the focus of retail transactions and the possible fraud that may occur at this stage of the business. Along with the types of fraud, different technological and business methods that are used for fraud detection are examined. The following are some typical types of fraud in retail transactions:

- Return Fraud - E-commerce Fraud - Friendly Fraud - Gift Card Fraud - Card-Not-Present Fraud - Chargeback Fraud

The study will focus on fraud or illicit activities specifically in retail transactions. In addition, the study will strictly focus on the fraud committed by consumers who visit brick-and-mortar stores. This study provides a comprehensive collection of fraud in the domain of retail transactions. The importance of the study is the modern approach of using AI systems in the detection of regularities for such fraudulent activities, which are still reliant upon human instinct or chronological retrospective strategies. The rising amount of fraud or theft and the need for sophisticated proactive surveillance of these unexpected expenses is crucial for retailers' missed revenue. Efforts to detect nearly all instances of shop fraud and prevent them may be prohibitively costly for several companies. A trade-off is required to be found and employed between the commitment of inadequate resources on one hand, and the associated financial losses on the other hand.

The inability to detect as much shoplifting as possible might decrease business earnings. The issue is that fraud detection cannot be constructed simply based on the reported numbers of shoplifting: the fee is known to be non-discriminatory. Therefore, businesses have no accurate reporting of retail shoplifting fraud levels or losses as they cannot determine which events actually happened. However, other fraud detection techniques must rely either on centric suspicion or retrospective logic. Efficient fraud detection searching will help the system attain cost-effective control. There are additional advantages such as boosting consumer trust by stopping the immediate monetary reduction, limiting merchandise damage, improving the company's obligation, and mitigating losses that could be settled.

2. Understanding Fraud in Retail Transactions

Fraud in retail refers to the act of deceitfully earning financial credits under someone else's name or with someone else's card through physical and digital transactions. It covers stolen or counterfeit cards and, until recently, in-person, card-not-present, and online payments. In general, the purpose of fraud in the field of retail is the financial gain of the fraudster. These types of crimes can cost retail or finance companies a substantial amount of money in terms of fraudulent charges. In some cases, stores may even refuse to let customers purchase items altogether. These changes may also happen after a store’s customer falls victim to fraud. In retail, especially in the fast and busy world of e-commerce, it is very difficult to identify, report, or stop fraudsters in their

tracks of operation. Unauthorized and abusive transactions can be reported to banks to receive a refund. This temporary fix is not enough in the world of fraudsters scamming millions of transactions every hour of the day. There are several ways a fraudster can make a profit off of a retailer or e-commerce. Credit card fraud is rising, and chip card systems may not be making it any better, solely shifting the fraud to card-not-present transactions. Return fraud is notorious for being a frequent and high-cost crime. Also known as 'friendly fraud,' consumers use purchased goods and return them; some individuals are even a part of return fraud rings. Finally, there is account takeover. Fraudsters use autopilot programs to sniff out the usernames and passwords to customer accounts on retail websites and shut the legitimate customers out. From there on, the fraudsters can impersonate the legitimate customers and purchase items from their accounts.

2.1. Types of Fraud in Retail Transactions

Recent studies show that due to an increase in the use of transaction platforms, the rate of fraud in electronic transaction systems is accelerating, and organizations have lost billions of dollars to these transactions. In many cases, the customer-facing side of a business, also known as the retail segment, has become an easy target for fraudsters. Many products and services are targeted by fraudsters to commit fraudulent acts. Fraudsters use many different fraudulent methods and technologies. Many researchers have identified a number of fraudulent transactions that are commonly found in the retail segment. The types of retail fraud activities are as follows: 1. Credit card fraud: Fraud perpetrators use credit card information belonging to another person, company, or financial institution without the consent of the owner. Credit card fraud is one of the most common types of fraud that arises in retail electronic systems. 2. Return fraud: In retail transactions, consumers return merchandise or request refunds. Due to the consumer's actions and the company's policies, some consumers use this rule as a means to commit fraudulent acts. This type of fraud is known as return fraud. 3. Coupon fraud: Companies issue coupons to consumers to attract customers for their products. In coupon fraud, consumers exceed the conditions set by the company when issuing the coupon, resulting in the company's product being purchased at a lower price. Fraud attacks are expected to continue to grow or even take new forms as retail systems and technologies become more sophisticated. Recent studies have shown that fraudsters use new technologies such as biometrics to perpetrate fraud. In response to this, researchers

are developing software and hardware technologies with the power to combat these new fraudulent techniques. Technology development to combat techniques used by perpetrators varies according to the type of fraud in question. Knowing the extent and characteristics of fraud types can help with the selection of appropriate algorithms and technologies used to prevent fraud.

2.2. Challenges in Traditional Fraud Detection Methods

A wide range of challenges and hurdles are faced by retailers while employing traditional prevention strategies against fraud. Their current prevention methods not only slow down the processing of transactions due to authorization delays, but also cause an increase in false positives, leading to poor customer experience. This issue is evident from the fact that only 35% of flagged transactions as potentially fraudulent are actually found to be valid. In cases where manual reviews are used, the process escalates the losses that the business can potentially incur and could also lead to a patient or customer being dissatisfied. Rule-based systems are also fallible in protecting a company from transactions that occur through new forms of attacks. More often than not, these strategies are outdated and no longer meet the requirements of a time-bound detection of fraud.

Moreover, traditional approaches lack real-time detection capabilities. Retailers often realize the fraud many days after it has occurred. For brick-and-mortar businesses, legitimate chargebacks could occur within minutes. Retailers need a rapid-fire, effective, and adaptable method that assists in keeping fraud to a minimum. Until now, advances in conventional electronic fraud detection have been partially successful, with ensuing innovations leading to digital systems that require a user in their card-not-present environment to identify themselves using a password or biometric. Despite these digital system technologies, which are known broadly as a protocol, to authenticate the participant in a network scarcely tackle the actual problem in e-retail industries where SMEs remain predominantly risk-averse and fraud rates continue to climb.

3. Machine Learning in Fraud Detection

With growing frequency, machine learning (ML), which is an AI concept specialized in learning from data, is a progressive technique for fraud detection. Here, ML is applied to recognize unusual patterns alongside the millions of retail transactions undertaken every day. The essence of ML is to allow computers to perform tasks autonomously, as

the applicable propositions of ML lie within purchasing behavior analysis. Comparison to traditional methods of fraud detection was indeed an issue, as the sheer quantity of transactions makes speedy individual examination impractical. Algorithms such as decision trees, clustering algorithms, and deep learning utilizing neural networks are some common and applicable methods in use today. However, the efficacy of these methodologies is specially tailored to the problem at hand. First and foremost, such algorithms must be provided with substantial amounts of information and will do best when given well-descriptive, pertinent information.

Feature extraction is indeed crucial to fraud detection, as racketeering charges are contingent on observable behavior. In other words, identifying properties relevant to certain types of fraud, like rapid purchases and suboptimal product selection, is instrumental to making systems of ML that are fruitful and appropriate. In addition, data preprocessing is particularly crucial in fraud detection, as it ameliorates flawed records by identifying the efficient information, such as repairing incorrectly formatted entries of user addresses or amending inconsistent dates. What is more, data should be organized in a consistent and orderly manner to make it easier to monitor for malicious activity. In contemporary commerce, ML-based AI reinvents fraud detection as a standalone technology. In particular, it holds many strengths, such as flexibility and velocity, alongside its own unique set of drawbacks. Experts in AI can be trained to recognize established patterns and fill the roles of fraud analysts with relative confidence. However, these systems must learn to perform effectively; they need a massive trove of data whose legitimacy is guaranteed, and machines should be prompted continuously to align the results with the accepted definitions.

3.1. Overview of Machine Learning Algorithms

Supervised learning algorithms allow a credit card issuer to correct and improve the performance of the algorithm and are ruled within two types: regression and classification. Examples of regression algorithms are linear regression, multiple regression, polynomial regression, and smoothing splines, whereas examples of nonlinear classifiers include decision trees, support vector machines, random forests, gradient-boosted machines, and neural networks. On the other hand, one of the most applied unsupervised learning algorithms for fraud detection is clustering, using, for example, k-means, self-organizing maps, or Gaussian mixtures. Other unsupervised

learning algorithms are also used, like dimensionality reduction, or using anomaly detection. Finally, both categories could also harness the power of certain semi- or weakly-supervised learning algorithms. Each of these algorithms is powerful and has its own advantages as well as poses its own challenges.

Picking the right algorithm can lead to significantly improved detection rates and reduced false positives, but employing the wrong algorithm can lead to an ineffective fraud detection system. Model selection and validation is an imperative process, with an incorrect choice resulting in a large number of missed fraudulent transactions going undetected. The same can be said for choosing the wrong pre-processing method. The usage of algorithms in retail is specific and must be adapted to deal with such problems in the retail sector.

3.2. Applicability of Machine Learning in Retail Fraud Detection

Anomaly detection has gained interest recently as a method for detecting fraudulent activities, especially in financial and retail transaction systems, which are mostly monitored by machine learning algorithms. Hence, in addition to detecting deviations from normal system operations and taking appropriate measures, these systems also have to deal with evolving fraudulent activities. The field of fraud detection is a lengthy and significant industry, where financial institutions, banks, credit card companies, healthcare units, and other commercial industries spend billions annually in an attempt to prevent these activities, resulting in uncountable losses. It has always remained a fast-evolving cat-and-mouse game between criminals and security assurance. To stay relevant, future fraud detection systems must remain state-of-the-art and effortlessly scalable to handle the high volume of datasets produced from daily trading and other operations in most commercial applications.

In this vein, the adoption and applicability of machine learning for commercial settings have always garnered attention, especially for retail fraud detection. Research has proven that machine learning techniques have been adopted by major retailers, resulting in a significant increase in return on investments and better detection performance. Moreover, national and international businesses are known to use artificial intelligence in their operational flow, mainly for fraud detection. The algorithms are generally designed and trained based on the patterns detected in each network and retail situation. One disadvantage these standalone cases address is the continued worldwide

occurrence of their accuracy not being public. Nonetheless, the benefits of machine learning-based models serve to attract overall attention and demystify the retail gap.

4. Case Studies and Applications

Best Buy, yet another retailer using machine learning technologies for fraud detection, uses these technologies to improve the digital experience of customers, as well as to combat the threat from card-not-present fraud facing every digital retailer. Machine learning technologies come with challenges, among which is the need to precisely and intelligently fine-tune the system. Models should be tuned to minimize false positives—the ultimate goal of a fraud detection engine—to improve customer satisfaction and trust. The alternative is worse than not being able to generate fraud alerts: generating too many of them increases customer friction, thus declining service to potentially valid sales transactions. Fraud management technology, applied to the managed services of a product line that has been doing for over two decades, resulted in a reduction of fraud alerts, which translates into a lift of the top-line revenue of global retailers.

First-party fraud, chargebacks, and CNP attacks are fraud problems facing retail. Reducing false declines is the solution. The future of fraud is concerning; their costs to manage fraud continue to rise. When you buy a widget for a dollar and sell it for two, you have a fifty percent margin. If it costs you to manage the fraud, that's a significant portion of your top-line revenue. If I can reduce that by thirty percent, your top line grows by two percent. Machine learning, however, has helped. In the case of a footwear and fashion company and a partner, it has helped achieve a lower-than-before false positive rate within sixty days. A customer has seen an improvement in a chargeback rate already in the "mid 20s" percentage-wise by a significant percentage with some out-of-the-box rules as a result of its implementation of fraud prevention strategies infused with machine learning. As with the partner, first-party fraud has been the biggest issue.

4.1. Real-World Examples of AI-Enhanced Fraud Detection in Retail

Subsection Outline 1. Introduction to fraud detection systems 2. Description of fraud detection systems piloted, implemented, or researched within the retail sector 3. Results from piloted fraud detection systems within the retail sector 4. Lessons learned and recognition of areas for improvement from the piloted fraud detection systems within the retail sector 5. Positive impacts of implementation and recognition of colleagues involved from the piloted retail systems 6. Consumer perspectives and recognition of

the changing fraud landscape and expected future fraud detection capabilities 7. Conclusions Across the globe, retailers have implemented AI-enhanced fraud detection in real-world scenarios with already impressive results, which have included: • Machine learning algorithms predicting loyalty account takeover using a large array of data points • Utilizing big data tools to develop algorithms that can determine when fraudulent transactions are increasing fraud risks for a retailer's online store and to further predict which sales are likely to result in chargebacks • Utilizing machine learning techniques to score a retailer's web sessions based on the user interactions made during the session. The scores are then used to automate business processes that require manual review. The biggest lesson learned by the piloted solutions was the necessity to tailor fraud detection to a retailer's individual selling channels and product lines, as fraud risks differed greatly between the focused groups. Additionally, the piloted fraud detection systems struggled with the detection of friend-to-friend fraud and certain new business strategies that are slowly becoming more common, such as "buy now, pay later." Lastly, those involved in the fraud detection system implementation experienced changes in their ways of working, which were sometimes unexpected and generally had a positive impact. Retailers are choosing to invest in retail technology suites that are underpinned by fraud detection systems in response to an ever-changing fraud landscape.

5. Future Directions and Conclusion

This study has outlined the current state of fraud in retail and emphasized that new generations of technology, such as AI, have become necessary to combat increasingly sophisticated fraud attempts. While many challenges can be overcome or mitigated with the adoption of machine learning solutions, the application of this technology has not been matched with the evolution of fraud itself. This study suggests further research into available technologies that will broaden the scope of fraud detection in retail. Specifically, there is a need for more research into the ethical considerations around the application of advanced machine learning techniques in a retail fraud context. Additionally, future research should focus on the development of systems that have a higher level of accuracy, either by improving the existing models or expanding our understanding of decision theory in the context of machine learning. It is important to recognize that the fraud landscape is still evolving. Techniques and methods are constantly being developed and also evolving over time. Fraudsters know that the more

they are detected, the more likely their exit strategy could fail. Retailers should be aware of the continuous increase in features to track in order to detect a wider spectrum of fraud attacks. Consequently, retailers must continually improve their detection strategies, and by doing so, they will need ever more sophisticated machine learning techniques. This keeps advancing the retail technology and fraud landscape as a whole. Retail managers are encouraged to start preparing or rather should already be investing in future AI developments or collaborations with the idea in mind that non-attendance fraud must be detected. This research sets the first step in describing the key nature of future frauds to detect: Many retailers do not know a lot about fraud; this research paints a landscape from which retailers should look to build from.