International Journal of Medical Banking

Volume 3 · 2010 · published by the HIMSS Medical Banking Project

- About
- Contents
- Main

Lessons Learned for Medical Banking from the Financial Services Industry:
The Case of a Fraud Prevention Platform as Transformative Innovation

By Allan Klindworth, MBA, and Stephen T. Parente, PhD

Abstract

In the late 1980s Financial Services was at a crossroads in identifying and preventing fraud. Tools and technology were immature and the strategy used was a "loss and chase" rules-based judgmental approach to "follow" fraud schemes. Culturally, key stakeholders were inclined to look at their solutions with a retrospective approach and only within their own credit portfolio. At that time, Financial Services did not have an understanding of the scope of fraud or an approach to even measure the amount of fraud. The majority of fraud rings went undetected. The industry required a change of culture and practice supported by new technology that bridged data silos with highly structured and proprietary data systems. In today's healthcare world a similar challenge faces the proponents of medical banking with information technology platforms bridging institutions. In the end, the financial services industry leaned there was more to gain by limited cooperation and data exchange in order to prevent billions of dollars in credit card fraud. The same lessons apply to healthcare information technology investments and applications that range from fraud mitigation to comparative effectiveness research.

Introduction

In the late 1980s fraud mitigation in Financial Services was a rules-based approach based upon a small number of frauds detected from prior experience. Rules were judgmental, retrospective and put in place by hard-coding them into a credit card issuer's processing system. The rules did not adjust for new types of fraud schemes. It was very typical that an issuer would have hundreds and in some cases thousands of lines of code as they continued to implement new rules over time. Criteria consisted of rudimentary decision trees. Examples include:

- Velocity of purchases – a judgmental cutoff of purchases over a specific period of time, not considering a consumer's historical purchasing patterns.
- Merchant category codes – certain types of purchases were deemed to be more risky than others, such as cash or jewelry.

Fraud analysts "worked" transactions by reviewing computer printouts on their desks. This methodology did not allow the capability to effectively or efficiently stop fraud before it occurred.

The detection methods at the time could not measure how much fraud was prevented or if a new test strategy or treatment was more or less effective than the previous strategy. It was easy for perpetrators to go undetected. If a perpetrator was identified it was usually after numerous fraudulent transactions. Ultimately, perpetrators would discard stolen credit cards and move their scheme to another issuer. Additionally, there was not a process to determine if the merchant was party to the fraud or if the merchant was colluding with a fraudulent credit card holder. Credit card issuers did not share information; they operated within the silo of their own business and within their own market.

Many fraud transactions were inaccurately charged-off as credit losses because issuers did not have the capability to identify them as fraud or the account holders could not be located. Sometimes legitimate consumers "paid for the fraud" because they did not carefully check their monthly statements. This was a result of transactions showing up on their statements from charges made by fraud perpetrators. This issue understated the true estimate of fraud in the industry and further perpetuated fraud losses.

Catalyst for Change – 1993

Predictive Models & New Technology

The catalyst for change started with the introduction of predictive models for identifying fraud. Predictive modeling had historically been utilized in the Financial Services industry for the underwriting of credit and loans. It was an accepted and proven method within the industry several decades prior. The introduction of predictive modeling to prevent fraud was an extension and enhancement of that methodology.

The introduction of predictive modeling to detect fraud was not a simple task. The solution had several obstacles to overcome in order to be effective. The industry had legacy mainframe technology as the core business processing system. Integrating a transaction-based fraud detection system into the core business with strict time-processing standards was just one of many obstacles. This issue was overcome by the fact that the savings from preventing fraud losses were many times greater than the investment to integrate and run the fraud scoring systems. The new solution was designed to perform real-time scoring and assessments on 100% of all transactions. Queuing, a workflow management and workstation methodology, was introduced to automatically and efficiently present only high-risk transactions to the system for automatic rejection or for review by a fraud analyst.

Market Implementation

Getting a first customer to pilot the new fraud technology solution was initially difficult. Credit card processors were the first to be approached to launch and pilot the solution, but they declined. They did not have economic incentive to do so because they were not negatively impacted by fraud. They made their revenue by billing their customers for each transaction processed, whether it was fraud or not. As a normal course of business practices, they simply raised prices if expenses increased.

The next step was to approach one of the nation's largest credit card issuers to pilot the solution. They accepted the proposal and implemented the system. They found that the results far exceeded their expectations. The issuer was able to identify fraudulent transactions that previously would have gone undetected. They were now able to implement real-time prevention processes, such as calling consumers, when suspicious out-of-pattern purchases were identified.

Because the solution was expensive to implement, it initially met some resistance from credit card issuers. However, as its value was demonstrated to individual issuers, its acceptance started to grow. Credit card

issuers were discovering that their initial return on investment was generally between 10:1 and 30:1. This means that for every dollar spent on the system, the credit card issuer saved between $10 and $30 in fraud losses avoided. Market expansion became dramatic and immediate. Since the results of the initial implementations were so compelling, other credit card issuers immediately requested to have the solution implemented on their systems. They recognized not only the positive results but the automated infrastructure used to interact with each consumer. Initial fears that the consumers, the credit card account holders, would react negatively to being "contacted" and asked if they were using their credit card proved to be unfounded. In fact, the approach of calling consumers and stressing the "fraud protection" intent of the call became a public relations success.

The predictive modeling solution proved to be much more accurate in identifying fraud. It was more cost-effective to "work" the high-risk transactions and it was received favorably by each issuer's consumers. The success of this solution has been universal. Now all credit card issuers utilize a real-time solution, which includes predictive modeling and workflow management, to prevent fraud within their portfolio.

Lessons Learned

Building an effective fraud detection system is a continuous learning and improvement process. Each successive system or model enhancement improves upon previous versions. That is one of the reasons why a "feedback loop" is required to update and improve the statistical algorithms.

While there were several items that were learned as part of the development and initial launch of the real-time fraud prevention solution, even more was learned after the launch because of the infrastructure built to measure and track results.

Indicators of Fraud

One of the riskiest transactions was identified as a part of the new methodology. It was a gas station transaction for less than $10. Fraud perpetrators would test stolen credit cards by discretely attempting a transaction at an unattended gas terminal. If the transaction was approved, that meant the card was good to use. This tactic by fraud perpetrators of "pinging" a gas pump became a typical indicator of subsequent fraud. However, it was soon abandoned by fraud perpetrators because they knew the real-time predictive modeling system had "learned" this scheme. If the logic of not analyzing all transactions prevailed in credit card scoring, this small dollar amount transaction would most likely have never been scored or detected.

Model Robustness

Predictive models continued to improve over time with each re-development. The feedback loop created to provide strategy outcomes (transactions tagged as fraud or non-fraud) continued to enhance each new model. Subsequent models created and implemented identified previously undetected fraud perpetrators.

Real-Time Assessment

The real-time platform provided the credit card issuers and their staff the ability to assess and review transactions as they were taking place. Previously, the fraud perpetrators could make a large number of transactions before the fraud was discovered. Often, the fraud was not discovered until the consumer received a bill with a large number of fraudulent transactions. Now, it is not uncommon to shut down a fraud perpetrator as they are attempting their first fraudulent transaction. The real-time capability provided the Financial Services industry the ability to move proactively from a detection strategy to a prevention strategy.

Accuracy

Initially, there was concern by credit card issuers about having to be 100% accurate when identifying fraud. This was overcome by implementing fraud strategies and processes that ranked the riskiest of transactions and then applied the appropriate action, intensity or investigation technique to the most questionable transactions. For example, the highest-risk transactions were declined for payment, while other transactions may have required merchants, via a real-time message, to verify credit card holder identification at the point of sale. The new approach provided a quick and non-obtrusive method to verify if purchases were fraudulent. Those transactions identified as less risky were queued for more research and investigated by a fraud analyst. All investigations were online, where previously they were paper-based manual reviews.

There were always false-positives (high scoring accounts that "looked like" frauds, but were not). However, the economic benefits of preventing fraud far-outweighed any negative impact to accounts that turned out to be non-frauds. The "goods" were quickly resolved in a customer service-type approach and were rarely negatively impacted. In fact, most credit card issuers used the chance to contact "good" customers as a positive public relations opportunity to demonstrate that they were taking action to "protect" the consumer from fraud.

Measurements

The infrastructure developed as part of this solution provided the foundation for the ability to measure return on investment. It also provided the opportunity to continuously improve and achieve the maximum effectiveness in keeping pace with ever-changing fraud patterns.

Silo Mentality

Initially real-time predictive fraud solutions were implemented at separate credit card issuer and processor sites. At the time, this seemed to be a natural approach to product market penetration. Later it was learned that this approach actually provided the industry with a less than optimal fraud prevention solution. The ideal solution would have been for all issuers and all processors to have housed their transaction data at a central site. This would have resulted in several benefits:

- Capability to capture the history and behavior of the same card holder across all of his/her credit card purchases, regardless of the issuer or processor.
- View across all issuers and processors to identify fraud rings – identification of a fraud ring in one portfolio would have allowed the ability to alert all other issuers and have quicker detection and prevention.
- Comprehensive view of merchants to detect merchant fraud and collusion.
- Feedback loop on fraud perpetrators across the Financial Services landscape would have allowed the predictive models to learn even quicker.

Access to Data – Known Frauds

It was difficult to build the first fraud statistical model because not much information existed on known fraud transactions. Generally, statistical models are developed with a known outcome (either "good" or "bad"). Since there were very few examples of "fraud" in Financial Services, other techniques had to be used until the feedback loop was established and frauds were labeled as such. In this case, predictive solutions inferred negative behavior based upon analysis. While this method was not optimal, it identified previously undetected perpetrators missed by rules-based strategies. A constant feedback loop allowed for refinement of the initial models.

Integration

Installing a fraud detection system into the core of a credit card processor's business processing system was a long, cumbersome and expensive proposition. It often took as long as 12 months to implement and test the

system. This approach could have been avoided with a modular solution. Under this scenario, only a data feed would have been required to complete a real-time assessment and communicate back a score, reason code and a treatment recommendation.

Results

Originally, the value proposition of the new fraud solution was to reduce resources used to "work" suspect fraud cases manually. The idea was to "throw away" the computer printouts used by the fraud analysts and allow the predictive solution to present the highest-risk transactions, ranked by score, in an online capacity. The initial hypothesis was that it would take fewer resources because the fraud analysts no longer had to research the nature of the fraud scheme. Instead, they could concentrate on only those transactions scoring in the highest-risk range. Specifically, the staff size could be reduced and salaries could be saved. The end result was much different. The new model identified new types of fraud and fraud rings that previously went undetected. In fact, initially credit card issuers actually added more staff, because on average, each fraud analyst was saving many times his or her salary in fraud dollars per year. Overtime, however, the number of analysts and number of transactions have been significantly reduced due to the success of predictive modeling approaches to prevention and detection of fraud. From 1993 to 1997 fraud was reduced by approximately 50% over expected results.

Parallels to Healthcare

The Problem of Healthcare Fraud

Although healthcare fraud and abuse estimates vary widely, constituents agree that the problem is enormous and growing each year. In testimony before the United States Senate Committee on the Judiciary, on May 20, 2009, Malcolm K. Sparrow, a prominent expert on fraud, said that healthcare fraud and abuse cost hundreds of billions of dollars per year, with the actual figure anywhere from $100 billion to $400 or $500 billion.[1] In 2002, a study by the Government Accountability Office estimated that 1 out of every 7 dollars paid to Medicare is lost to fraud.[2] This means that in Medicare alone, there was almost $70 billion in fraud and abuse for 2008 (projection of $466 billion in total Medicare spend for 2008[3]). If we extrapolate this assumption for 2017, Medicare would have over $120 billion in fraud and abuse (projection of $857 billion in total Medicare spend for 2017[4]). According to experience and research, it is estimated that the vast majority of fraud and nearly all of the abuse is perpetrated by healthcare providers.[5]

The most common approach in healthcare fraud and abuse detection today is to apply rules-based or judgmental methodology and technology. Rules are intended to imitate and automate human judgment. They are typically retrospective if/then statements are hard-coded into the back-end of a claims system. The terminology for implementing judgmental criteria is called 'edits.' This approach also mimics the manual process to identify claims that are outside of normal policy. This strategy has done little to mitigate healthcare fraud and abuse.

With the continued escalation of healthcare costs and healthcare reform on the horizon, it is imperative that payers and healthcare constituents implement new or improve existing solutions to identify, prevent and reduce healthcare fraud and abuse. There is an immediate opportunity to advance beyond the existing healthcare system's approach today and meaningfully reduce the cost of healthcare through more innovative fraud and abuse mitigation solutions.

In order to make a meaningful impact on fraud and abuse in healthcare, new technology and predictive models with real-time assessments must be utilized to review suspect claims prior to payment. Today, these claims are

paid and reviewed afterward to ascertain if they violated documented healthcare policies. If they are found to be questionable, organizations such as Medicare then seek reimbursement through a "pay and chase" strategy. CMS recently started to recover payment errors by utilizing Recovery Audit Contractors. CMS contracts with Recovery Audit Contractors (RACs) to guard the Medicare Trust Fund.

Financial Services Approach to Prevent Fraud

There are several parallels of early Financial Services practices before the advent of real-time assessments and statistical scoring and healthcare's current approach to fraud mitigation:

1. Pay and Chase Strategy – similar to early, unsophisticated and inefficient Financial Services procedures, claims are paid and then researched for recovery opportunities.
2. Judgmental Criteria – healthcare solutions are rules-based and do not adjust for changes in fraud patterns.
3. Batch Assessment Approach – claims are reviewed based upon batch processing on the back-end of the data flow.
4. Lack of Fraud Performance – claims tagged as fraud or providers identified as perpetrating fraud or abusive behavior do not exist.
5. Silo Approach to Prevention – claims data and fraud outcomes are not shared at a central location.

Culture to Change

Currently, there is not a proactive, sophisticated, effective and efficient fraud prevention solution in the healthcare industry. Similar to Financial Services in the early days, it is as though no one is "minding the store."

The current industry strategy and processes utilized breed a culture of acceptance of fraud and abuse. Adding to this is the fact that there are few, *if any*, incentives by key stakeholders to try new approaches to mitigate those costs. It is most often just easier to pass along the increased costs associated with fraud and abuse to consumers, employers or the government. If this laissez-faire attitude to prevention is to change, the following actions must be implemented across the healthcare landscape by Medicare, Medicaid, Tri-Care, private insurance and third party administrative companies:

- Implement statistically sound, empirically-derived score modeling systems that are designed to prevent (versus detect) fraud and abuse.
- Mandate that healthcare payers submit transaction data such as Medicare's "Common Working File," Medicaid files and private insurance claims data to a pooled data repository that will support the build of fraud detection models that have a comprehensive view of all provider and consumer activity across all payers.
- Fraud scoring models must score ALL transactions before any edits or other screening process removes claims from the data source.
- Use historical data to build and redevelop scoring models.
- Provide a "feedback loop" where ALL payers provide post-payment information about the eventual status of a claim as "Normal," "Fraud," "Abuse" or "Education Required" – this transactions feedback loop will also enable healthcare agencies to measure the actual amount of fraud, and fraud savings, in the system.
- Score ALL claims "promptly." For Financial Services this was real-time; generally, in healthcare, this would mean claims are aggregated and scored within a one-day period. Results would be returned on the day after the claim is submitted.

Conclusion

In 1993, Financial Services launched the first transaction-based real-time fraud solution. The system used proven predictive modeling techniques, combined with innovative technology to identify previously undetected fraud transactions and fraud rings. Prior to the launch, the strategy in Financial Services was a "loss and chase" strategy. Detection methods utilized judgmental rules and edits that were hard-coded into legacy processing systems. The industry as a whole was not able to estimate fraud or adjust its detection methods to the changing patterns of fraud perpetrators. It took the changing of internal cultures to launch the first pilot. The first pilot project implementation resulted from a mixture of innovative thinking, risk taking and an attractive return on investment potential.

The healthcare industry is in nearly an identical place as Financial Services was almost two decades ago. The solutions to identify fraud and abuse are judgmental rules-based, ineffective and inefficient. The strategy is one of "pay and chase." Now there is an opportunity to use the lessons learned and proven technology of the Financial Services' transaction solution and refine it and apply it to healthcare. The opportunity for success and the return on investment is even more attractive than it was in the Financial Services industry. As the Financial Services industry was at a crossroads in the late 1980s, the healthcare industry is at its own today. It can pursue a new and innovative integrated fraud and abuse prevention program or continue to suffer enormous losses with the current outdated detection system.

References

[1]Sparrow MK. Criminal prosecution as a deterrent to health care fraud: Testimony before the Senate Judiciary Subcommittee on Crime and Drugs, Washington, DC, May 20, 2009.

[2]United States General Accounting Office ("GAO"), Report to Congressional Committees, *Medicare Fraud and Abuse: DOJ Continues to Promote Compliance with False Claims Act Guidance*, April, 2002.

[3]http://www.cms.hhs.gov/NationalHealthExpendData/downloads/proj2008.pdf; page 5.

[4]http://www.cms.hhs.gov/NationalHealthExpendData/downloads/proj2008.pdf; page 5.

[5]Coalition Against Insurance Fraud. Go Figure: fraud data. http://www.insurancefraud.org/stats.htm. Accessed July 27, 2009

About the Authors

Allan Klindworth, MBA, is Principal of TerraMedica LLC, Wayzata, Minnesota.

Stephen T. Parente, PhD, is Professor of Finance, Carlson School of Management; Director, Medical Industry Leadership Institute, University of Minnesota; and Chief Health Economist, TerraMedica LLC, Wayzata, Minnesota.