CHAPTER 1

Bank Fraud: Then and Now

Perhaps the earliest recorded case of fraud in the Western world was that of Hegestratos and Xenothemis in 300 B.C.[1] The story goes that Hegestratos took out an insurance policy on a boat for a large sum, with the deliberate intention of sinking it. At this time, ships were going down at a very high frequency, so this was not necessarily a bad idea (from the point of view of the fraudster), provided one managed to pull it off. Hegestratos was supposed to carry a large amount of grain from Syracuse to Athens on his boat. His idea was to not carry any grain but sink the boat halfway through the voyage and collect the insurance money. He would get the price of the boat reimbursed, and since there was no grain on the boat, he wouldn't incur the loss of the value of the grain. What ended up happening was something else altogether. The people on the boat got wind of Hegestratos's plan to drown them and confronted him. Unable to face the opposition, Hegestratos jumped overboard and drowned himself. His partner, Xenothemis, had to sail the boat to the port, and things didn't go well for him either. A legal battle followed between the buyer, Protos, who was waiting in Athens, and Xenothemis, when Protos, who thought he was getting grain, found out that there was no grain on the boat.

Even though the exact details of the verdict in this legal battle are lost to history, we know that Hegestratos and Xenothemis were unable to carry out their plan, and things ended badly for both of them. While this is surely not the oldest case of fraud in history, it is one of the oldest *recorded* cases of fraud.

This chapter traces a rough history of fraud and compares the times we are in with historic times and looks at how complicated the world of fraud management has become. In order to begin laying the groundwork to understand how complex fraud detection systems have become a necessity in the last few decades, it is important to be aware of this history.

THE EVOLUTION OF FRAUD

If we look to the East, many stories of fraud exist in Hindu mythology and in the folklore of various parts of Asia. Fraud is probably as old as money itself, and we could go a step further and say that fraud has probably existed in this world for as long as human beings have inhabited it. One might ask, "What is different about the times we live in?" In historic times, unlike today, fraud was a rather sporadic phenomenon. There was also considerable stigma associated with fraud as most of it was discovered sooner rather than later, which served as a deterrent to its widespread use. Written over 2,500 years ago, the *Thirukkural*[2] is a masterpiece by the poet Thiruvalluvar composed of 1,330 couplets in the South Indian language Tamil (which happens to be my mother tongue). The 284th couplet says that the unbridled desire to defraud others, when fruitful, will produce endless pain and sorrow. This indicates that fraud existed many thousands of years ago, and most often it resulted in the fraudster reaping considerable notoriety and sorrow. Not only was this the case in the East but also in the Western world, as evidenced in the Hegestratos and Xenothemis story.

Fraud in the Present Day

Fast forward to our times. Not only has fraud become much more prevalent now compared to historic times, but the frequency and the ubiquitous nature of today's fraud means that fraudsters don't necessarily meet

the end they deserve. Financial institutions are forced to fight fraud all the time. If fraud is not fought effectively, fraud losses can threaten to derail entire institutions. Some of this is because there are so many more human beings inhabiting the world today, and this results in interactions with institutions becoming more and more impersonal, thus opening up a rich environment for committing fraud. Fraudsters have become so sophisticated that they don't need to be present and make personal sacrifices like Hegestratos to carry out their plans. Fraud can be completely impersonal as far as the fraudsters are concerned.

Banks are especially vulnerable to fraud. Why are they so vulnerable? I am reminded of a conversation alleged to have occurred between Willie Sutton, a legendary and prolific bank robber, and a reporter, Mitch Ohnstad. Sutton is said to have robbed more than $2 million and spent over half of his adult life in prison. The reporter Ohnstad asked Sutton, "Why do you repeatedly rob banks?" to which Sutton replied, "Well, that is where the money is."[3] That statement pretty much sums up why banks are so popular with fraudsters. In most cases of fraud that banks experience, the fraudsters are never caught. All that the banks are able to do is to stop the bleeding by stopping fraud as soon as they can ; they have little hope of recouping the money lost.

In the good old days, when there were fewer customers and banks were for the most part local, they had the luxury of having face-to-face relationships with customers. In the last 40 or 50 years, this has been changing. Not only are there more and more (too many) customers to keep up with, but there also are many customers simply not available for face-to-face interactions. As banks got bigger and the pressure to get bigger and more profitable grew, they were forced to innovate in terms of customer acquisition as well as ways in which customers transact with the bank. As interactions with banks became more and more impersonal, the resulting anonymity also helped the fraudsters to exploit the system.

Risk and Reward

As we all know, lending money has been the business of banks almost from when they started. However, the amount of risk a bank is willing to take to lend money has changed dramatically in the last 50 years.

Gone are the days when customers had to appear personally at the banker's office and show the assets on which a loan is requested. In those days, not only were assets showing the customer's ability to pay back the loan needed, but there was also the need to have third parties assure the bank that the money would be paid back if the borrowing customer was unable to repay the loan.

Fast forward to 10 or 20 years ago: Pretty much anyone who had an account with the bank and the semblance of a job could walk in and get a loan—not secured, but an unsecured loan like a credit card and/or other types. Even though it seems to be a pretty risky path for banks to take, as long as they could manage the risk/reward equation by exercising decent control on the risk side, it became a very lucrative path for the banks. The reward portion of the equation is generally dictated by the volume of business a bank can generate. Most of the time, the volume of business is proportional to the number of customers. The same volume also helped fraudsters. The higher the number of customers, the more impersonal the relationships become. You can see how the continuum operates.

Secured Lending versus Unsecured Lending

Even with a rapidly growing customer base, it is possible to keep a decent amount of control on secured lending. In secured lending, there is an asset that the bank has control over that can be used to recoup losses it might incur, especially if the perpetrator is the customer. However, unsecured lending is a totally different beast. Unsecured lending is based on intangibles such as the behavior history of the customer and so on. In addition, since the customer does not have skin in the game, unsecured lending becomes a burden mostly on the bank. Unsecured lending pretty much opened the floodgates in terms of fraud. To a number of customers, it seemed like free money . . . almost. The biggest proliferation of unsecured lending happened in the area of credit cards. The concept of being able to get money using a small plastic card was not only an amazing idea, but also one that caused a lot of crooks to start thinking about how they could exploit this little plastic card to get the free flow of money going. Due to high interest rates for credit cards, in spite of the fraud losses, running credit card

portfolios was and continues to be a very lucrative business for banks. However, if there was a way to control losses, credit card portfolios would be even more attractive for banks. This meant that issuers had to figure out a way to keep fraud losses in check. Various authentication methods such as signature matching were used in the beginning to keep fraud rates under control. Not surprisingly, fraudsters found easy ways around these authentication methods. This is when the realization came that studying the cardholders' behavior and looking for deviations would be a much more effective method of keeping fraud in check than using authentication methods, which the crooks could find ways around. Statistical models started do a better job of understanding the nuances of cardholder behavior and what is normal for a customer, so the automation of the process of detecting fraud as well as improved accuracy became a huge asset to managing fraud.

These days, interestingly, even authentication methods are expected to have some understanding of the customer beyond simply matching a password to the recorded password of the customer. We live in a complex world where customer expectations have grown, and as customers have become more sophisticated, there has been an inherent expectation that the banks should almost magically know the behavior of the customer based on past history.

In unsecured lending, a lot more diligence is needed in combating fraud because fraud directly affects the bottom line of banks, as there is no way to recover losses from the customers. About 15 years ago, the banks started to turn to systems based on technology. Some of these systems could see what the human eye could not. The human eye can see two dimensions and, with some help from the brain, can understand the third dimension. When we start thinking about higher dimensions and interactions, the human eye is simply incapable of seeing odd behavior. If you include the human intellect, it is possible to look a little further. However, no system is going to be as efficient and adept at finding fraudulent patterns that do not fit as statistical models. Technology had helped spearhead the phenomenon of interactions becoming more and more impersonal. Now the same technology (involving behavioral modeling) came to the rescue to address the problem it was partially responsible for creating.

Statistical Models and the Problem of Prediction

Yogi Berra, the legendary American baseball catcher and manager, once said, "It is hard to make predictions, especially about the future" (the predecessor to this statement was made by physicist Neils Bohr).[4] This very funny but very insightful quip applies to any prediction problem, and from one point of view, there is a lot of truth to this statement. As Nassim Nicholas Taleb, the author of the book *Black Swan*, says, it is true that vast portions of the future lie beyond our abilities to predict.[5] The same argument tends to get used quite a bit against statistical models as well. Since a significant portion of this book is aimed at detailing the evolution of data analysis and statistical modeling and how much both have helped in combating fraud, let me address this at the very beginning. From certain points of view, it might seem that statistical models are not adequate to accurately predict the future. However, from my point of view, statistical models for the most part do a great job of making good predictions about the future even when the predicted situation is not exactly the same as what was observed earlier. Statistical models are very good at limiting the exposure (or fraud risk) and giving us a decent handle on the future. Statistical models have a tremendous ability to understand complex patterns and extrapolate to a decent-sized region not only in but around the values that the models were trained on.

To put it in real terms, let's say that cash deposits of $10,000 or more followed by multiple withdrawals are risky. If a rule or a mathematical algorithm is written to monitor for cash deposits of $10,000 or more, it is simply incapable of seeing risk when a deposit of $9,000 is made followed by withdrawals. However, a statistical model can observe a $9,000 deposit followed by multiple withdrawals and flag the activity as risky even though the model has never seen the exact same type of activity in the data presented to it.

The key here is the proximity of the dollar figure to the original number, and it gets a lot more complex and hard to do as we move away from the number. Statistical models, though, afford us the ability to extrapolate and learn in regions previously unknown, as long as the regions are reasonably close to what has been observed earlier. In a way, this is the way the human race has managed to grow knowledge in any scientific field, isn't it? If you look to the field of medicine,

the development of antibiotics was based on repeated scientific experiments where each scientific experiment relied on the previous one and slightly expanded the knowledge space. We learn from the accumulation of knowledge, experiment a bit, observe new results, gain knowledge, extrapolate slightly beyond our previous region of knowledge, and so on. Statistical modeling is no different. While there are certain areas in which the ability of statistical models is more limited than in others, today it is true that without risk management largely driven by statistical models with behavioral input, banks would not survive. This is a hard fact that stares at everyone whether or not one has affinity for statistical models.

There are many examples that have been provided almost since statistical models came into existence on how wrong these models can be. There are any number of jokes on confidence intervals and how little they mean. There is of course the famous (but often overused) statement, "There are three kinds of lies: lies; damned lies; and statistics," by British Prime Minister Benjamin Disraeli, which was popularized by Mark Twain.[6] As much of a lover of statistics that I am, I would go one step further and say that the one thing you can be sure of with any statistical prediction is that it is not precise. When I say that a transaction's fraud score is 930 (meaning a probability of fraud of 0.93), the one thing we know for sure is that it is not correct. The transaction is either fraudulent or not, which should lead to a score of 0 or 1,000, translating to a probability of 0 or 1 if we simply want to be precise. But most fraud scoring systems do not use the score 0 or 1,000. Does this mean that fraud scoring systems are not useful? Absolutely not! When it comes to statistics, it is important to focus on how useful the output of a model is rather than whether the exact prediction is right or wrong. A score prediction is expected to be right in a large enough set of transactions with the same score, not for an individual transaction. Life is never about the extremes. It is always about the shades of gray that we need to understand more clearly.

So, while there have been countless number of writers before me and I am sure there will be countless number of writers after me who readily talk about the imprecise nature of statistical models, the benefits a statistical model provides are not as much in its precision and accuracy as it is in the rank ordering it provides. As long as a score of

930 has a much higher probability of fraud compared to a score of 850, a ton of value can be gained from the score, especially in high-volume areas where there is a need to separate the goods from the bads very quickly. In an environment where thousands of transactions are queuing up every second for a decision, it is important to quickly categorize transactions into groups with various false positive rates so that analysts' time can be well spent on identifying fraud. To this end, statistical models work wonders.

Also, it is important to understand that while the individual score of 930 may not be precise for a single transaction, if enough transactions with the score of 930 are accumulated, in the group, overall, the score could be calibrated such that 93 percent of the transactions would be fraudulent. In some ways, this is as precise as life gets! So, next time, before we criticize a statistical model for its imprecision at the very granular level, we should try to understand at a high level what the model is trying to do and what the practical use of the model is. We can then appreciate the amazing contributions from statistics that make our lives a lot easier.

Almost every technique and every advance in technology has its pros and cons. What we should look at in evaluating anything is to see if it directionally improves and advances our understanding of what is going on. This is the most important way to look at any scientific advancement. Statistics has been at the receiving end of more than its fair share of notoriety and ridicule of its shortcomings and issues. However, if statistical models are evaluated (with all their limitations, of course) from the point of view of how effective they have been in combating fraud, the wonders that have been possible in the area of fraud management can be appreciated. This evolution holds true for a number of different industries, but this book mostly focuses on modeling as applied to bank fraud. With this said, let's look at how things evolved in the banking industry.

THE EVOLUTION OF FRAUD ANALYSIS

Back in the good old days, when most banking was personal and most of the authentication was personal too, fraud could be handled very well. If the only way you can withdraw money is by walking into

a bank and having a teller check your identity, it is a lot harder for someone to take over your identity and commit fraud. Enter the age of impersonal banking where transactions can be conducted from anywhere. It became necessary to see only the signatures of the people transacting without really seeing them face to face, as customers did not have to be available in person.

Early Credit Card Fraud

Building these customer signatures is a process of evolution that is easy to observe in the area of unsecured lending in the financial industry. Take credit cards, for instance. Twenty years ago, when credit cards were proliferating and everyone wanted to get one, and every bank wanted to sign up as many customers as possible for their credit cards, the banks had a real problem on their hands. Pre-set spending limits (typically in the many thousands of dollars) were imposed on customers, but when a credit card was lost accidentally by the customer or was stolen, fraudsters had free rein for a few days while the banks were literally robbed of the unspent credit line. The banks couldn't do a whole lot to stop it, as neither the customer nor the bank had a clue that fraud was being committed. Add to this the lack of liability on the part of the consumer. In order to limit consumer exposure, laws were passed. Consumers by law were not liable for more than $50 or so, and the banks were stuck with the lion's share of the losses. Necessity is the mother of invention, they say. The first seeds of the need for some heavy-duty technology were being sown then and there.

Fraud departments in those days were mostly staffed with ex–security personnel. These experts had a pretty good idea of what to look for in transactions on a transaction-by-transaction basis and started collecting data and writing reports to understand the nature of the fraud they were dealing with. Data collection and reporting certainly shed a lot of light on the nature of the fraud problem, but by the time experts saw what was going on, it was typically too late to do anything to stop fraud. It was like looking in the rearview mirror while driving a car and simply understanding what had happened already. The bleeding had occurred, and even though fraud losses in general were only a few basis points (as opposed to credit risk [delinquency] losses, which ran in

the hundreds of basis points), fraud losses were beginning to take a tremendous toll on the psyches as well as the pocketbooks of many banks.

Once reports were written to analyze the data and understand what was going on, some of the more number-savvy fraud analysts and managers started seeing correlations between various quantities. For instance, they began to see that fraudsters prefer committing fraud at night. They also figured out that fraudsters like to check out whether the card had any credit line left by doing some small-dollar charges at a terminal far enough away from a watchful human eye so as to reduce the chances of getting caught in the first fraudulent transaction they were committing. Once the fraudsters figured out there was money available in the card and the card was still working, they tried to do as many fast transactions as possible such that the goods purchased in these transactions could be converted to easy money. For example, buying jewelry or purchasing electronic goods fetched money a lot faster than buying books.

When fraud experts at banks saw that most fraudulent transactions had a certain set of characteristics and some of the quantities they saw varied proportionally to certain other quantities that they had observed, they decided that they would start writing some rules to tackle the fraud. For example, let's say electronics store purchases are risky, but they are even more risky if the purchases are happening at night. One could write a rule that says that if the purchase is at an electronics store and it happens at night close to the time the store is closing, that transaction needs to be blocked right away so that the money flow can be stopped.

Separating the Wheat from the Chaff

This worked for a very short period, and then things didn't go as expected. Consumers had started using credit cards much more than they previously had for two reasons. First, consumers realized that using a credit card is really free money for an average of 45 days (on average, it takes 15 days for transactions to show up on the monthly bill, and the customer has an additional 30 days to pay off the bill) if they had the discipline to pay their bills every month. Second, consumers realized that while they were responsible for paying the minimum amount due every

month, they really didn't have liability when it comes to fraud; all they had to do was call the bank and ask to cancel the card and get a new card issued. If the bank hesitated, the consumers could always just respond to the multitude of credit card offers hitting their mailboxes every week and get brand-new cards. All of a sudden, there were legitimate customer purchases happening late at night at electronics stores. Stopping all electronics store transactions also meant that the revenue from these transactions could not be realized by the banks. Just as fraud losses needed to be controlled, the revenue side of the equation also needed to be managed. From the customers who transacted heavily, as opposed to carrying a balance and interest (also known as revolving), the main source of revenue for the banks was interchange revenue. Stopping all those high-dollar transactions meant a significant loss in interchange revenue. The problem of separating the wheat from the chaff in terms of fraud had just become more difficult. Plus, the customers were getting more demanding in terms of treatment. They were not very pleased if their genuine transaction at night was misidentified as fraudulent (a "false positive") and stopped. The customers demanded that the bank figure out which transactions were real and which ones were not.

When a large quantity of data is analyzed and some simple correlations are observed through reports, it is literally like lighting a candle or turning the light bulb on in a dark room. A lot of value in the insight is gathered, and experts start to observe when things are going wrong, to the extent that these patterns can be observed and understood by the human eye and intellect. There are a couple of problems, though. As in the rearview mirror example I mentioned earlier, you can understand what has already happened, but that in itself doesn't prevent issues in the future. These lessons need to be converted to proactive decisions that can be made in the future. And the understanding itself is rather limited, and these limitations have to be overcome in order for the lessons to be used effectively in preventing fraud in the future.

The Advent of Nonlinear Statistical Models

Simple statistical models such as linear and logistic regression models are much better at understanding and generalizing fraud versus nonfraud behavior compared to expert-written rules. All of the variables

that might have an impact on detecting fraud can be used as variables in the models. With respect to interactions between the different variables in the model, as long as the experts are able to figure out the interactions and feed them as variables into the model, the models are capable of understanding the behavior and its relationship to fraud. Unfortunately, experts don't have an infinite amount of time to understand and code these variables. For this reason, the simple statistical models started giving a huge number of false positives, and as the fraudsters got more sophisticated, neither rules nor simple models could be used to stop fraud effectively.

This was happening around the time that unsecured lending was proliferating at a very rapid pace. The income prospect for banks in unsecured lending was significant, and banks couldn't grow without it. These were the days when it was relatively common to receive a dozen credit card offers in the mail in any given month. As credit cards began to take off, credit card fraud losses began to increase drastically as well. The need to control fraud risk was real, as fraud losses could literally make or break the bank's profit numbers for that quarter.

It was at this time that some advanced nonlinear statistical models were introduced that could score transactions in real time to detect fraud. Nonlinear models like neural networks have the ability to understand and include interactions between various types of behaviors automatically. Most risk phenomena are nonlinear in their relationship to the target, especially something like fraud that is quickly changing as fraudsters get more and more sophisticated. The use of neural network–based behavior models in real time literally has changed the face of fraud management all over the world. It significantly reduced banks' fraud exposure in areas where there is a need to react in a split second and stop the transaction before money goes into the hands of the fraudster. When we examine how a credit card– or debit card–based transaction happens at a merchant point of sale (POS) terminal, we will realize that the bank has only several milliseconds to make the approve-or-decline decision on the transaction.

For instance, say a customer is at the POS terminal trying to purchase electronics worth $1,000. The bank has just seconds to decide whether to approve or decline the transaction. Considering the amount of time it takes to send the decision back to the host system and so on,

the amount of time available purely to make the fraud decision is on the order of milliseconds. So, not only do we need sophisticated models that compare current behavior to past behavior and quickly judge using a model whether this transaction is fraudulent, it is also necessary for this model to run extremely fast in production systems.

The production execution of the model has to be precise (from a fraud-detecting perspective) and extremely fast in terms of returning an answer. When systems that could accomplish both of these objectives were first introduced in the marketplace, almost overnight the impact to fraud losses was huge. All of a sudden, the fraud problem could be tackled very effectively. Not only did this have an extremely positive effect on managing fraud, but in a lot of ways these systems paved the way for the tremendous growth in unsecured lending that in turn led to the growth of banks. Data-driven fraud detection systems have had a transformational effect on the banking industry.

Tackling Fraud with Technology

We have not yet seen the best of what can be done using behavioral modeling, not only in the area of fraud but in decision making in a number of customer touchpoints. In terms of advertisement monetization and a number of other areas, the fun with behavioral models is just beginning. In the next few decades, I predict that the world is going to witness, in a very broad sense, the impact that understanding data, modeling the data, and predicting the future will have on every decision made by institutions that requires customer insight.

In order to understand the use of technology in tackling bank fraud, we should perhaps start with the evolution of predictive modeling as a field and understand the evolution of statistics and data analysis techniques. Statistics is considered to have been born along with cryptography, based on the ninth-century book by Al-Kindi titled *Manuscript on Deciphering Cryptographic Messages*.[7] In this book Al-Kindi gives a detailed description of how to look at data frequencies to decipher cryptographic messages. Observing numbers and analyzing their natural frequencies provide a lot of information on whether they are following a certain pattern or not.

SUMMARY

In this chapter, we examined how fraud has evolved since historic times and how the nature of fraud has changed at a very rapid pace in the last couple of decades. The need for sophisticated nonlinear statistical models was also well established.

Over the next several chapters, we will examine the evolution of technology in the area of fraud detection, the challenges from an operational perspective, and how the future of risk management is looking really bright, due not only to current techniques but also to techniques that are very promising but have yet to be used in risk management specifically. I hope the reader has found this time travel interesting and informative.