AI Fights Money Laundering

Jason Kingdon, *Searchspace*

The Wolfsberg Group, formed by a group of international banks to share ideas on combating global money laundering, met in early 2002 to discuss the aftermath of the September 11 terrorist attacks. Joining the group were representatives from the FBI, Interpol, the Financial Action Task Force, and a host of other international and national financial regulators and investigators. 9/11 represented a massive failure of international controls, and clearly only radical change would prevent something similar from happening in the future.

What troubled the committee and international agencies was that the whole terrorist operation could have been funded for less than US$0.5 million—pocket change, in the context of routine banking. Up to this point, international money laundering had been about the Mafia, drug smuggling, and arms deals, involving sums in excess of $500 billion a year.[1] How could banks transacting hundreds of millions of dollars a day spot suspicious activity in transaction amounts as small as $2,000 to $5,000? How could they have detected the terrorist financing behind 9/11?

Two years later, almost half of the world's top 20 banks are using AI systems,[2] and AI has emerged as the leading method in the fight against money laundering (see the "AI and Money Laundering Detection" sidebar).[3] At Searchspace, we monitor customer activity to identify unusual behavior and detect potential money-laundering situations.

The problem from an AI perspective

In 1998, I met with the head of risk for a mid-sized UK bank. Even at that time, the UK had stringent money-laundering regulations owing to the ongoing terrorist threat. Banks would occasionally fall foul of the regulator and need to demonstrate whole-hearted commitment to finding ways of trapping suspicious activity. This particular bank was concerned with its staff's ability to successfully monitor such activity, given modern banking's increasingly faceless and electronic nature. Could AI help monitor transaction behavior to detect money laundering?

The bank had approached Searchspace, formed by researchers from the Intelligent Systems Lab at University College London in 1993. It applies adaptive and learning-systems approaches to a range of business and finance tasks. However, until then, we had principally developed systems for US and UK stock exchanges to automate market-abuse detection—monitoring insider trading, front running, market manipulation, and other forms of market cheating.

The problem the bank posed was far more challenging—five million transactions per day, five million individual accounts, over three million customers, hundreds of product types, and no clear signature or pattern associated with money laundering. Unlike many types of financial fraud, money laundering could range from a single transaction to the culmination of months of complex transactional activity. A sequence of transactions might be interesting only in the context of activity that has taken some time to emerge. In engineering terms, this represents one of the worst forms of dimensionality disorder, with massive scaling variances and feature overload.

Additionally, the bank wouldn't share past cases that it had detected or reported. It judged this information too sensitive and insisted we find what we could on a year's worth of historical data unguided.

Unsupervised beginnings

To an extent, we had been here before. Most financial organizations don't keep good records of infractions—certainly not sufficient to use as training sets. They're also reluctant to share historic cases because of the legal complexity of active and prosecuted cases.

So, for the stock exchanges, we developed an approach to unsupervised learning based on probabilistic data analysis. The approach used adaptive heuristics mixed with conventional search and pattern recognition techniques. We used these to build what we called *special investigator modules*, now known as *Sentinels*. These are autonomous investigator agents designed to police a specific business issue and alert people when there's a problem. To aid this process, we developed a conceptual framework for housing

1094-7167/04/$20.00 © 2004 IEEE
Published by the IEEE Computer Society

AI and Money-Laundering Detection

Owing to the complexity of money-laundering detection, most commercial systems have concentrated on simple rules for monitoring certain payments, such as those from sanctioned lists of individuals, organizations, or high-risk geographies. One example is the OFAC List, published by the Office of Foreign Assets Control.

More sophisticated analysis has used intelligent systems to enhance manual investigations. For example, neural networks and fuzzy logic have aided tasks such as link analysis—where associations between account or individuals are analyzed for common features, including name and address, zip code, phone numbers, or other account details. However, these data mining techniques are more useful when an investigation is already underway, because they're suited to agencies that pursue criminal investigations. They're less useful for industries attempting to prevent the abuse in the first place.

In 1995, the US Congress Office of Technology Assessment published a report investigating the use of AI for automated money-laundering detection. The report's conclusion was daunting—"automated computer screening of transactions for money laundering is virtually impossible."[1] The gloomy conclusion was largely based on an assumed false-positive rate applied against the number of transactions a system would need to screen. Fortunately, we can improve this math by looking at customers and accounts as well as individual transactions. Transactions inform account analysis, but system alerts are account- or customer-based and thus avoid duplication and redundancy.

Reference

1. *Information Technologies for the Control of Money Laundering*, OTA-ITC-630, US Congress Office of Technology, Sept. 1995.

the autonomous agents so that they could cooperate and draw on a range of similar services such as data access, security, reporting, and workflow. (We're still developing these ideas—the notion of an integrated network of autonomous agents is promising. It offers flexibility in problem solving and forms the basis of an AI operating system.)

Our Sentinels used heterogeneous hybrid technology so that each could draw on different techniques depending on what seemed appropriate—or what worked. One principle we enshrined was the results' transparency. A Sentinel had to be able to explain why it sent an alert, what it found, and what it wanted you to do. For this reason, and because of the lack of training examples, we didn't use nonlinear parametric techniques such as neural nets.

One method we developed to help detect insider trading was akin to a probabilistic support vector. It leveraged the huge dimensionality of the insider-trading problem by assessing the probability of an individual trading on insider information across a likelihood matrix. This was a support vector machine but with probabilistic thresholds. The tests we ran replicated and improved on human investigations. Moreover, unlike many automated approaches, the exchange's regulatory team considered all results worthy of investigation. And with no training set, the Sentinel couldn't be accused of overfitting the data.

This approach looked like a good start for trying to break down the money-laundering issue.

The cure of dimensionality

Computers excel when dealing with large volumes of data. For money laundering, the scales are so large that machines are the only way to tackle the problem. The question was how to start and whether this issue was too complex for the current state of automated analytic techniques.

As we'd done with other Sentinels, we looked closely at the underlying objectives. For regulatory Sentinels, this often meant reading the relevant statute. The basic legal requirement for money-laundering reporting in the UK is for banks to assess customers' banking activity and report suspicious behavior.

However, there are no clues as to what constitutes "suspicious" behavior—what makes a transaction suspicious? An amount, an action, or a series of actions? Anecdotally, the bankers talked about tellers filing SARs (suspicious activity reports) based on observations—"the customer looked suspicious and the money smelled of fish." Not the best basis for building an algorithm. Furthermore, the regulator insists that the bank use its understanding of banking and its customers to determine what suspicious should mean. Moreover, the bank must be able to demonstrate that it's doing this systematically. This interpretive approach allows great freedom for the regulator to assess a bank's approach in the absence of easy-to-prescribe rules. From an AI perspective, it also ruled out filtering transactions for a series of scenarios.

We made our first breakthrough by looking for "unusual" rather than suspicious behavior. This let us establish behavioral norms for each customer and look for "weirdness"—which is nonprescriptive but statistically defined over some dimensionality.

The question now became one of dimensionality—what activity should be judged and over what scales? We exaggerated the basic concepts of support vectors by introducing massive dimensionality—customers × accounts × products × geography × time. This generated an enormous multidimensional adaptive probabilistic matrix for each of the bank's customers. We could now assess the likelihood of individuals' actions based on simple weighted aggregations of activity using the underlying probabilities.

The matrix is adaptive in that it updates statistics on the basis of customer behavior—as a customer makes transactions, the underlying probabilities change. So, customers set their own behavioral pattern—frequent cash payments or international payments could be normal for some people but unusual for others. One customer's $100,000 withdrawal could be as normal as another's $50 withdrawal. Using a probabilistic view of the world solved the first major problem of scale invariance with an ability to compare unlike activity. It also provides a pseudo information-theoretic means for identifying information, or, in this context, unusualness.

Identifying money laundering

The probability matrix was a good start, but it didn't actually detect anything—we needed a way to focus on interesting activity. So, we introduced a probabilistic retina in which an event occurs if a certain probability threshold is breached. Akin to trapping a photon, each retina cell fires only if a certain low probability, or quantity of information, is achieved. For example, customer X depositing an enormous amount of money would cause a cell to fire, where we define enormous as relative to who deposits the money. The method also evaluates more subtle activity; for example, say customer X over the course of many months and across multiple product types (cash, checks, debit and credit cards, and so forth) withdraws an unusually high amount based on previously established norms. This too will fire an event. The Sentinel then aggregates weighted events for a customer across the matrix's multiple dimensions to get a ranking of universal unusualness across the bank at that point in time.

We designed the dimensions to be loose in form so that the system could score the same activity from many angles. This naturally amplifies unusual activity in a noisy environment. It also allows for a more robust design process because no single "signature" is the key to the recognition task.

We also use peer groups in the matrix to establish unusualness outside a person's behavioral norm. For example, it might be normal for a commercial account that's a jeweler to transact large cash sums—say, $10,000,000 a week. However, this might be odd compared to other jewelers. We can judge this in a peer context in aggregate amounts and across product types to give balanced insight into the behavior. As we fine-tuned the matrix of events, the money-laundering Sentinel took shape.

We then worked with the bank to refine the results and installed our first operational system in 1999. It assessed every transaction for the risk of money laundering in its own context, against its own behavior and its peers' behavior, and against the sequence of activity. Unusualness turned out to be a good surrogate for suspiciousness, providing insight into suspicious behavior in the context of the masses of information the bank processes each day. The system didn't use fixed rules but rather relied on statistically based qualitative judgments and comparisons of activity.

We've since refined our methods and improved the techniques. The system currently has approximately a 1-in-14 false-positive rate—one alert is prosecuted through an SAR filing (meaning the banking staff judged it worthy of action) for every 14 raised. The system generates alerts on average at a rate of 0.00001.

As part of the refining process, we've also looked at styles of laundering—could, for example, our methods have detected the terrorist financing behind 9/11? According to results presented to the Wolfsberg Group and others, although the amounts transacted were small, the patterns were unusual. More problematic today is how the investigator bureaus can respond to the improved intelligence from the banks.

Exploring other applications

Long ago, banks knew you by name and knew what you did with your money. Although we've gained the benefits of fully automated global electronic banking, a fully automated system is blind—and not just blind to money launderers, fraudsters, and general abuse but also to the provision of service. If you never meet your customer, how can you respond to his or her needs?

In this context, technology puts considerable distance between the supplier and consumer, which is bad for business. What was once a close relationship is now remote and potentially undervalued.

Ironically, it's in this sense that AI techniques can excel. If it's possible to replicate the contextual interaction of a teller and customer within the transaction infrastructure, then we can replenish some of the intimacy and immediacy of human contact, a basic component of which is to "know your customer."

For banking, this traditionally refers to the requirement that banks check a customer's ID when opening a new account. However, we argue (as do many in the industry) that faking an ID is so easy that identity theft is one of the fastest-growing banking issues. The AI approach is based on understanding an individual's behavior—that is, "know your customer" by knowing his or her behavior. Knowing how people behave is the best way to identify risk.

However, knowing someone through his or her behavior is inherently bespoke, and interactions between a corporation (bank) and customer on this mass scale is new. It moves us from an era dominated by segmentation and broadcast marketing toward an era of individual interaction based on behavior. Once an infrastructure is capable of this style of interaction, the possibilities touch almost all areas of banking. It means that event-based services can be trigged by behavioral contexts rather than static segmentation and isolated campaigns. It also becomes possible to learn a customer's preferred manner of interaction, whether that be through direct mail, a call center, or email. Obtaining this information based on a person's response is far more representative than using answers from a series of preference questionnaires.

AI makes these decisions dynamic, adaptive, informed, and timely. People change, and the ability to stay current is as vital to making good risk decisions as it is to making good marketing decisions. It allows more dynamic decision-making; for example, what should trigger a denial of service—insufficient funds in an account or an over-limit credit card? This requires good judgment on a case-by-case basis and within the context of bank policy. A machine with the intelligence to judge context can operate at this scale, resolving problems with transparency and justification. In some sense, it's akin to assigning a service guardian to each customer to negotiate the range of the bank's business objectives.

At Searchspace, we believe this new generation of adaptive operational analytics offers new possibilities from risk management to service provision, and we're working with others in the industry to bring this potential to life. ■

References

1. J. Robinson, *The Laundrymen*, Arcade Publishing, 1996.
2. B. McGuire, "What a Difference a Year Makes: Winnowing Out AML Vendors to the Largest US Banks," TowerGroup, 2003.
3. N. Katkov, "Ranking the Vendors of Anti Money Laundering Solutions," Celent, 2003.

Jason Kingdon is CEO of Searchspace. Contact him at j.kingdon@searchspace.com.