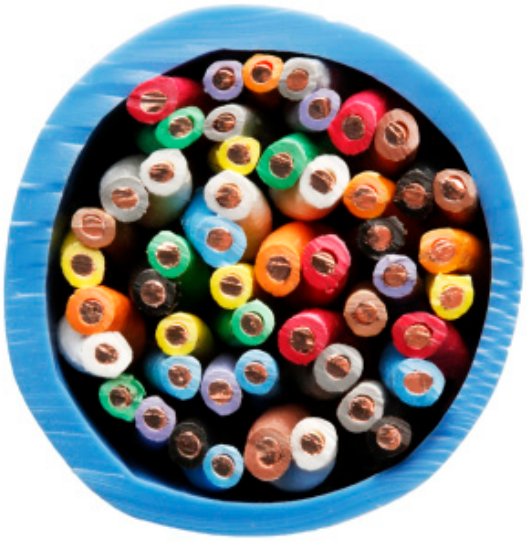
Deloitte.

Suspicious Activity Reports and analytics: Staying ahead of the compliance curve

For years banks and other financial institutions have operated sophisticated systems for monitoring, investigating, and reporting suspected money laundering/terrorist financing (AML/TF) transactions. In compliance with a number of U.S. regulations under the Bank Secrecy Act of 1970 (BSA), the USA PATRIOT Act of 2002, and other U.S. laws, these systems sift through significant amounts of data and, as a result, flag many transactions that require investigation as possible suspicious activities[1] . This paper suggests that financial institutions should use analytics to improve those systems, which can result in more efficient, effective, and insightful operations.

Many financial institutions regularly evaluate the effectiveness of their interdiction systems with an eye toward improving the quality of the monitoring, reporting, and investigative process. Governmental regulators also exert pressure on institutions to improve monitoring and more effectively ferret out suspicious activities. In fact, the U.S. Office of the Comptroller of the Currency (OCC) provides specific guidance for establishing, implementing, and maintaining quantitative models used in bank decision making, as well as governance and controls related to them.[2] The Risk Analysis Division of the U.S. Department of the Treasury and the OCC deploy quantitative analysts to enforce that guidance through examinations of transaction monitoring systems, including a review of the underlying rules/detection scenarios.[3] Bankers expect continued BSA enforcement by regulators in 2013 and beyond, as evidenced by the April 2012 citation of a major U.S. bank, the largest alleged violation in more than a year.[4]



Compliance executives are tasked with delivering a suspicious activity reporting system that fulfills their regulatory duty both effectively and, OPTIMALLY, cost-efficiently AS WELL. One low cost opportunity that is often overlooked is to better leverage *existing* information generated by transaction monitoring systems and suspicious activity reporting processes. This existing information can be analyzed to find opportunities for fine-tuning those systems to more effectively identify potentially suspicious behavior.

Financial institutions are required to issue Suspicious Activity Reports (SARs) as part of the Anti-Money Laundering/Counter-Terrorist Financing (AML/CTF) regulatory framework. Often, SARs are used only

Analytics Applied.

[1] Activity refers to transactions processed by the financial institution. This paper uses the terms activity, transactions, and behavior interchangeably.

[2] OCC 2011-12, "Supervisory Guidance on Model Risk Management," OCC, April 4, 2011, http://www.occ.treas.gov/news-issuances/bulletins/2011 bulletin-2011-12.html.

[3] Different detection platforms use either the term "rules" or "scenarios" to describe algorithms designed to uncover suspicious activity. This paper uses these terms interchangeably.

[4] "Regulators Gearing up for New Bank Secrecy Push," Rachel Witkowski, American Banker, April 17, 2012.

to fulfill that requirement – suspicious activity is captured, cataloged, and submitted. Additional analysis could be beneficial for the institution; SARs are rife with useful information, including detailed types of suspicious behaviors exhibited and the nature of the parties involved. If financial institutions CAN capture that information in a database and then mine it using advanced analytical tools and techniques, that information may reveal (at a more aggregate level) a far richer and more nuanced understanding of the underlying activity – the nature of suspicious activity being uncovered, potential patterns and trends, the effectiveness of detection scenarios, overlaps between scenarios, and other opportunities to improve scenarios, to name a few.

Some financial institutions may not have databases that capture information from or about SARs. Other institutions may have captured such information, but many of those databases may be underutilized because of the way the information is collected. By restructuring that existing information and using analytics to investigate it, new ways of improving the automated compliance systems may be uncovered. What follows is a discussion of ways that analytics can be applied to these databases and how analytic tools may improve the quality and effectiveness of a system whose ultimate output is SARs.

Understanding inefficiencies in current SAR processes

Transaction monitoring processes use a common paradigm in financial institutions: as transactions flow through the institution, the organization's transaction monitoring system combs through them for certain characteristics or groups of interrelated characteristics that are associated with suspicious activities. Some of these transactions or patterns of transactions trigger automatic alerts that are flagged by the system for further investigation. The algorithms that trigger such alerts comprise detection scenarios, which are the engines driving the transaction monitoring system.

Ultimately, improving the effectiveness of a transaction monitoring system involves regularly evaluating and fine-tuning these detection scenarios. Applying analytics to the data collected by the transaction monitoring system and mining that data creates a feedback loop so that the quality of the output is related back to the drivers of the system.[5] When analytics are integrated as a fundamental element of the process, financial institutions can make informed adjustments to the detection scenarios to help them effectively meet their goals.

Notably, many existing transaction monitoring systems and SAR processes contain areas of inefficiency that should be considered before analytics can achieve their potential:

The SAR process is a mix of automated and manual activities that are ripe for improvement in quality and efficiency. The SAR process typically begins when an alert is generated by the transaction monitoring system. Alerts may be grouped with other alerts into a case for further investigation, or the alert may stand on its own. An investigator examines each alert or case to determine whether it should be escalated for further review. Investigators or reviewers write up the SARs, and the institution approves and submits them to the regulators.

Large financial institutions can generate thousands (and in some cases, hundreds of thousands) of SARs every year, so AML/TF compliance costs can be substantial. Many scenarios will generate superfluous or unproductive alerts, that is, alerts where suspicious activity is anticipated but none is found. More unproductive alerts mean higher compliance costs.

Unstructured data can be a fundamental problem. In some instances, SAR databases are made up of many thousands of narrative reports containing unstructured data that can be difficult to analyze. Unstructured data is information presented as text – for example, a written paragraph describing the nature of the suspicious activity – which is challenging for analytics to parse. As of yet, there is no easy and reliable way to perform quantitative analysis on the informational components of a written paragraph. However, augmenting the SARs with descriptive (yet structured) information will help support the use of analytics and data mining tools. This can be accomplished by asking a few basic closed-ended questions. For example,

[5] Although the term "analytics" is used generally in this section, the more specific analytics suggested are described using a case study later in this paper. See "SAR analytics in action" below.

- In aggregate, what were the primary reasons SARs were filed?
- Can the institution determine whether any SARs in their database(s) are linked because they are filed on the same subject at different periods of time?
- Does a particular SAR focus on a customer of the institution, or on a party with which that customer is transacting?

Although financial institutions are required to track their SARs, few structure the content of those SARs so they can be analyzed and quantified. If an institution has not previously collected SAR data this way, it is possible to look retrospectively at the database and "backfill" those analytical data points. However, the preferred and most efficient time to ask these questions is when a SAR is generated, when the determinants of the decision and the results of the investigations are fresh in the minds of the involved personnel. Going forward, asking questions during the SAR preparation process can provide structured data, thus saving time and, ultimately, cost. When questions are structured to allow systematic analysis of SARs, the effectiveness of analytics may be measurably increased.

Understanding the components of SARs is an important area of analysis. A transaction monitoring system that generates a high percentage of unproductive alerts is likely to benefit from the kind of calibration that analytics and data mining tools can provide. Unproductive alerts can be triggered by ill-defined rules or one or more misleading or coincidental transaction attributes, such as "negative news" about one or more parties associated with a series of transactions.

Negative news is a frequent source of unproductive alerts from the perspective of tuning a rule (that is, changing the parameters in a rule which generate an alert). Take, for example, a hypothetical scenario where a party to a transaction receives adverse media coverage. Suppose "John Smith" is listed as the director of a number of shell companies transacting through a bank, and an alert is generated on one of those transactions. During the research phase of the bank's investigation, negative news is uncovered about John Smith related to suspected money laundering. The shell companies that listed John Smith as director, as well as several additional shell companies that did not mention him, share the same address. The bank commences to file a large number of SARs on the other shell companies based on their relationship with John Smith, either directly or indirectly, through the shared address. None of the transactions flagged by the system are suspicious on the basis of their transactional activity; instead, the SARs are generated because of the negative news.

Next Generation SAR Form in Development

Financial institutions are not the only ones that are looking to capture information in a way that supports more analytical investigation. The Financial Crimes Enforcement Network (FinCEN) issued a new, more detailed SAR form that will populate an electronic structured database. The agency requested comments on the list of proposed data fields for its transition to a modernized IT environment for electronic reporting and made available industry testing arrangements, demonstration reports, and other helpful technical information.

The deadline for compliance is March 31, 2013.[6] This new SAR form captures much more detailed information on the product data and suspicious activity types than the previous version. For example, far more detailed information is captured about the SAR subjects, including the NAICS code of the industry in which they conduct business; the specific type of suspicious activity, such as fraud, terrorist financing, structuring, casino activity, and money laundering, as well as more detailed closed-ended questions detailing what kind of fraud, money laundering, or other activity was identified; and the type of product (such as cash, check, money order, government payment, and funds transfer) involved. Much of the information will likely be useful to financial institutions when they begin integrating analytics with their current transaction monitoring platforms, because this is a good example of the type of data that can be analyzed to determine trends and patterns.

The result of the new SAR form should be to increase the amount of usable information available to institutions individually and FinCEN more broadly. Financial institutions are limited in their investigations, particularly when investigating schemes that span institutions for which each institution has access to just one piece of the puzzle. The fact that FinCEN and the financial institution will each have additional common structured data suggests enhanced monitoring in their respective analyses of suspicious activity.

[6] This is not to say that SARs should not be filed based on negative news, just that the distinction should be considered during the tuning process. Institutions can employ negative news rules to capture parties referenced in adverse media, separate from detection scenarios or pattern based rules.

In such cases, negative news may not be a sound indication of suspicious activity intended to be captured by pattern or behavior based detection scenarios. So, as a financial institution tunes the rules underpinning its detection scenarios, it could be important to distinguish between SARs that were generated primarily because of negative news versus those generated primarily because of an actual suspicious pattern of activity. Incorrectly attributing the identification of negative news to a detection scenario which was not designed to identify it will hinder the proper use of analytics, and provide both an unrealistic picture of how well the scenario is performing and an incorrect assessment of which kinds of transactions are indicative of suspicious behavior.

Analyzing and refining source data and detection scenarios can result in a finely-tuned transaction monitoring system. Ongoing analysis might then be able to highlight additional patterns that indicate the need for new or improved rules – a virtuous cycle that can help the financial institution in its efforts to effectively uncover suspicious activities.

SAR analytics in action

A hypothetical case study illuminates some of the analytical methods that may be used to mine a SAR database. In this instance, the financial institution is a correspondent bank that has been found to be in substantial compliance by its regulators for years. Still, compliance executives at the bank want to assess their SAR process because they believe it can be even more effective, productive, and efficient.

The bank has a SAR database, although it is mostly unstructured data, and some of the required links between alerts and SARs are not captured. The bank would like to improve its database to apply analytics and gain insights into the kinds of suspicious activity it captures. The bank intends to carry out this plan in three phases: data structuring and cleansing, detection scenario analysis, and threshold refinement and system improvements.

Phase 1: Structuring and cleansing the SAR database

First, in Phase 1, an inventory of alerts and SARs (not limited to alerts that ultimately led to SARs), including their unique identifiers, is created. These are loaded into a normalized database which lays the foundation for useful analysis: being able to tie alerts to SARs is critical. Next, analysts deconstruct the correspondent bank's existing SARs, which contain the narrative reports. Investigators review the narrative from each report and, by asking basic questions such as those described above, reformulate the report into a set of structured data points that can be analyzed and quantified in the second phase using analytics and data-mining tools.

Based on insights from their SAR review, the analysts, investigators, and bank management enhance the procedures by which investigators will create structured data points as they write and issue future SARs. The updated procedures take into consideration potential risks within the bank's own customer base in addition to the risks at large in the banking system.

Phase 2: Detection scenario analysis

In Phase 2, the correspondent bank analyzes the now-structured, historical SAR information in the database – specifically, alerts generated by the transaction monitoring system and the subsequent SAR narratives stored in the SAR database. Objectives are to:

- Evaluate effectiveness of detection scenarios
- Mine for undetected patterns
- Tune detection scenario threshold values
- Explore visual analytics
- Differentiate between counterparties and focal entities

Evaluating effectiveness of detection scenarios

The team starts by evaluating how effective the system's detection scenarios are at capturing the suspicious behavior they are designed to uncover. By extension, the team will identify which ones generated more unproductive alerts. They will also identify rules that generated alerts on which SARs were ultimately filed, but not based on the suspicious pattern of behavior they were designed to detect. Rather, the alerts and resulting SARs may be based on negative news or incomplete information on the parties involved in the transactions – an important distinction.

[7] A normalized database is structured in such a way to minimize redundancy and dependency. Rather than repeating information in large tables, distinct data is housed in separate tables, and relationships are established between them. Less redundancy allows for more agile queries and reporting. When there is a requirement to add, delete, or modify data, a normalized database provides discrete tables such that these changes can be made in one place and inherently applied through the rest of the database by way of the established relationships.

[8] The risks that are expected to be monitored by AML/CTF transaction monitoring systems are not static for a particular financial institution or the banking system as a whole. Procedures should be periodically revisited so that they are tailored to address any new risks identified.

One example of this analysis involves identifying and understanding overlaps between rules and the impact of such overlaps on alert generation. For example, the SAR narrative analysis from Phase 1 shows that two particular rules within the existing transaction monitoring system have generated a considerable number of alerts. However, a closer look reveals that no SAR has ever been filed on the basis of an alert generated solely by one or the other rule. Only when both rules simultaneously generate alerts have investigators subsequently identified actual patterns of suspicious activity. This finding gives the bank an important insight about its compliance environment – specifically, that only the activities that trigger both rules are likely productive starting points for investigation. The bank can then seek to demonstrate quantitative evidence to its regulators as to why it will focus resources on transactions that trigger both rules.

Another related analysis shows that some suspicious activity alerts generated by Rule A are *always* generated simultaneously by Rule B – but the same is not true in reverse. This finding indicates that Rule A is redundant – it provides no additional information about the suspicious behavior. Rule B, on the other hand, always captures the Rule A activity, as well as other types of activity it is designed to identify. This finding is a particular strong suit for analytical assessment, which often identifies rules that create unwanted "noise" in the system – alerts and resulting investigations that are unnecessary and potentially repetitive and costly over time. In such cases, there may be the opportunity to "decommission" one or more rules, a formal process that requires careful documentation.

Mining for undetected patterns

The analytical analysis may also uncover transaction patterns contained in the SAR database that have gone undetected by any existing rule or have not triggered past alerts in a systematic manner, and yet which still indicate suspicious activity. This finding signifies one of two potential problems: either the existing rules need to be adjusted to detect these particular transaction patterns, or the bank needs to develop a new detection scenario that trigger alerts and are backed up by sufficient information to conduct a competent investigation. It is possible that an existing rule coincidentally picked up a new pattern and a SAR was filed. However, if the new pattern wasn't captured in a structured manner during the SAR process and stored in the database, the bank is limited to anecdotal evidence of its existence, and it is difficult to quantify the instances of occurrence.

As part of this phase, the bank identifies these previously undetected patterns and has an opportunity to develop new rules and analyze their effectiveness on historical data. This feedback loop from SAR back to the monitoring environment demonstrates a dynamic learning environment which can be a positive factor when undergoing periodic regulatory reviews. Again, in the absence of analytical analysis, these patterns and the related detection opportunities could potentially go unnoticed.

Tuning detection scenario thresholds

Another critical objective the bank has is to tune threshold values. Detection scenarios typically have parameters with associated threshold values. For example, a rule might require the customer to have at least two transactions and $50,000 in activity to alert. The count and sum of transactions are the parameters, and two and $50,000 are their threshold values.

Threshold values should be carefully considered and analyzed on a periodic basis to confirm their appropriateness or to adjust them as necessary. Regulators require that financial institutions employ sophisticated statistical analysis and supporting documentation to justify their settings. SARs are an important dependent variable in tuning rules in a transaction monitoring system.[10] If SARs document suspicious activity, and transaction monitoring systems are designed to identify suspicious activity, it is logical to tune the system, alerts, and thresholds to identify activity that ultimately leads to a SAR and reduce alerting activity that does not lead to a SAR.

Focal entity vs. counterparty alert analysis

As mentioned at the beginning of this example, the institution being reviewed here is a correspondent bank. The correspondent banking environment itself presents another separate issue to be addressed – identifying whether its SARs target focal entities or counterparties, or both. Because the correspondent bank acts as an intermediary between other banks (typically foreign banks),

[9] A focal entity is the party on which an alert is generated, or focused. A counterparty is the party that is transacting with the focal entity.
[10] SARs are not the only dependent variable that should be considered. Increasingly, regulators are requiring institutions to look beyond SARs

it often has little or no information about the parties that are transacting with each other.

The correspondent bank's transaction monitoring system aggregates activity on the originator and beneficiary in separate rules. In an originator rule, the originator is the focal entity and the beneficiary is the counterparty, and vice versa for a beneficiary rule.[11] However, in the past when SARs were filed, the SAR database did not capture which party triggered the alert in the system.

If this information had been captured, analytics could be used to mine the data to help determine which party was actually targeted for suspicious behavior. This analysis could offer insights into the accuracy and sensitivity of the detection scenarios and underlying rules. These insights would be used in Phase 3 (below) to fine-tune the system.

Explore visual analytics

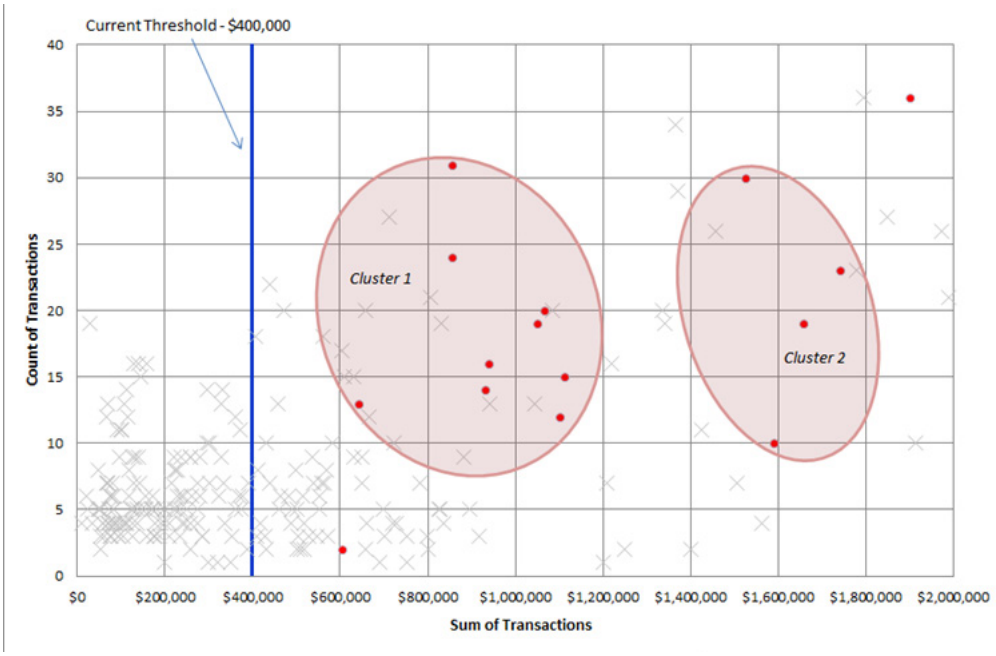
SAR distribution and clustering is often difficult to identify without a graphical representation of the data, such as those provided in Phase 3 in this case study. Visual analytics in the form of both simple and complex analyses can provide critical insights into where potential suspicious activity occurs. By including or excluding different parameters in visual representations of the data, it is possible to identify patterns or correlations that may be much more challenging to identify using another statistical approach.

Phase 3: Application of threshold refinement and system improvements

The insights gleaned from Phase 2 analyses can be applied to threshold refinements and system improvements during Phase 3. Returning to the example of overlapping rules described in Phase 2, since it is more likely that a SAR will be filed when both rules are triggered rather than when either rule is triggered on its own, the bank develops a risk scoring algorithm that assigns a score to each rule. The system then requires a minimum score for all rules in the period on the same focal entity in order to generate an alert. In this example, the system generates an alert when the conditions for both rules are met, but not either rule on its own.

As mentioned previously, there may be an opportunity to "decommission" one or more rules, a formal process that requires careful documentation. Discovery of undetected patterns indicate either that existing rules need to be adjusted to detect these particular transaction patterns, or that the bank needs to develop a new detection scenario.

Figure 1: Clusters of suspicious activity above current threshold



Source: Deloitte

Visual analytics enables the bank to better identify clusters of suspicious activity (Figure 1), which guide where thresholds may be adjusted to isolate those clusters and, consequently, reduce the number of unproductive alerts

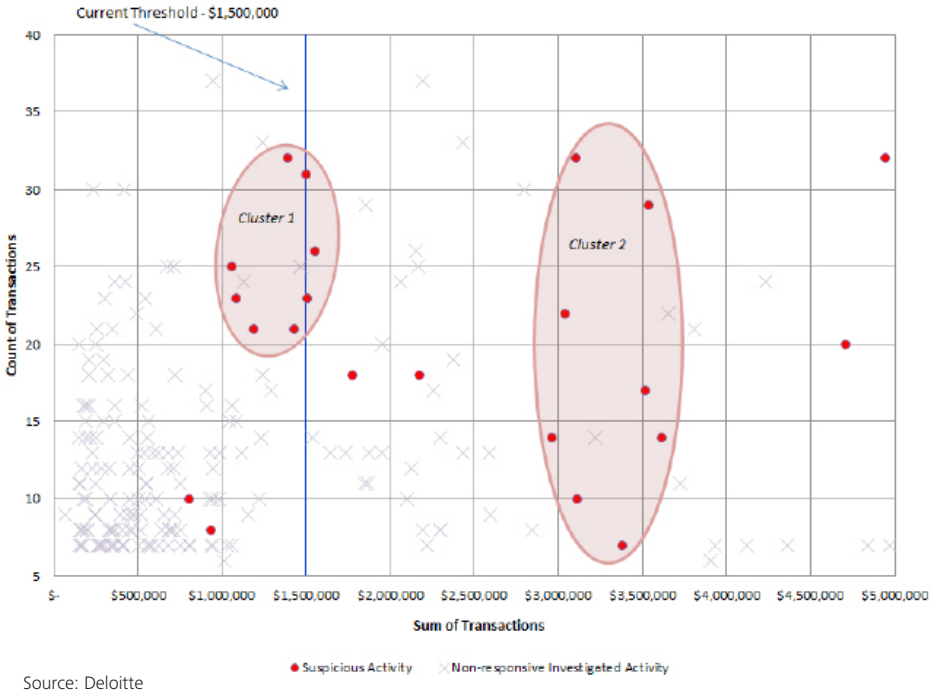
For example, the cluster analysis in Figure 1 indicates that most transactions being flagged by the transaction monitoring system in a particular detection scenario are substantially above the "floor," or bottom threshold.[12]

[11] In some cases, the originator, the beneficiary, the focal entity, and the counterparty are all one and the same.

[12] It is important to note that a sufficient amount and variability of data is necessary before these conclusions can be made. For example, if alerts are generated on a monthly basis, it would be premature to draw a conclusion with results from just one month of data.

Those transactions are analyzed further to determine whether the floor threshold should remain the same or be raised. Conversely, the clustered transactions associated with another, different detection scenario suggest that the floor threshold actually may need to be lowered (Figure 2).

Figure 2: Cluster of suspicious activity below current threshold



Source: Deloitte

In some cases, a detection scenario contains no "ceiling" or upper threshold. Yet the cluster analysis reveals the opportunity for one, since no suspicious activity is detected above certain parameter values. The cluster analysis is important because it can provide a basis for making tuning judgments when coupled with a visual presentation. It also makes the basis for the judgment transparent to regulators, to whom the methodology and conclusions of the bank should be clear, cogent, and persuasive.[13]

The bank understands that it is important to keep in mind the intent of the scenario to arrive at a conclusion that makes sense. For example, if a scenario examines patterns of activity, it does not make sense to lower thresholds to accept a single transaction, as a single transaction cannot provide a pattern of activity.

The bank also performs another series of analyses, but these analyses are focused on a subset of alerts below currently implemented thresholds. The analyses are intended to assess the lower thresholds the bank established when it first installed the transaction monitoring system, or since the threshold values were last changed, without examining the universe of possible alerts below the thresholds.

It can be important for correspondent banks, in particular, to differentiate between SARs filed on focal entities as opposed to those filed on counterparties when incorporating the analyses from Phase 2 into Phase 3. Rules will always generate alerts based on activity of the focal entity. When reviewing the effectiveness of a particular rule, the bank considers the SARs that were filed on the focal entity separately from those filed on the counterparty. Similar to incorrectly attributing negative news to a detection scenario which was not designed to identify it, attributing suspicious activity of a counterparty to a focal entity will also hinder the proper use of analytics. The activity should be analyzed to determine how the system and rules can capture the counterparty's activity as the focal entity of its own alerts.

These analyses lend themselves to many different tuning opportunities, as well as leads on how the bank can change the way it examines products and assigns risk in order to ameliorate its review burden.

Coming full circle

This hypothetical case study demonstrates that financial institutions can potentially benefit from applying analytics and data-mining tools to SAR databases. Analytics can allow institutions to judge the overall effectiveness of their transaction monitoring system and SAR processes. Analytics also gives institutions opportunities to leverage investments they may have made by deriving new insights from the data and SARs.

On the surface, executives may believe that their transaction monitoring systems are performing satisfactorily, but feedback from analytics can quantify how effective these processes are and help identify where improvements can be made. In addition, analytics can give compliance officials the evidence they need to help justify

[13] The bank should not blindly follow conclusions drawn from analytics. If the bank feels the transactions above a proposed ceiling are inherently risky because one "missed" alert holds such value or an abundance of activity, this should be considered as part of the analysis. Even when a ceiling is not implemented, it is helpful to know where the hypothetical ceiling threshold lies, so that it can be monitored over time, and the decision can be revisited once more data has been collected.

their case with government regulators when changes should be made. When regulatory requirements evolve, both historic and ongoing analytics can demonstrate how changes should be made while limiting disruption to routine operations.

Leveraging lessons learned

The regulatory expectations for finding suspicious activity have escalated commensurately as financial institutions have improved their detection of bad actors. Every new round of improvement carries with it a better understanding of how to find suspicious activity, and each subsequent round seeks to apply the lessons of the one before it. Analytics helps financial institutions discern the lessons from each round efficiently and transition to an improved set of scenarios and thresholds, while reducing the burden of testing out new processes. Applying analytics to the acquired knowledge contained in a SAR database – leveraging existing information – can provide the insight needed to regularly refine internal processes and improve quality and effectiveness.

Institutions can benefit from taking steps to structure their data-gathering – ideally as part of their business-as-usual procedures – so analytics can be readily applied to SAR databases. Collecting the data in a structured manner allows analytics to become an integral part of ongoing compliance operations – not simply a snapshot of a system state at a certain point in time, but a perpetual process of mining the data and providing feedback on the system's rate of success and failure.

Analytics can provide compliance executives with an opportunity that helps them leverage information they already have to meet an unrelenting imperative they face every day: fulfilling their regulatory duty by running high quality, dynamic, and effective compliance programs.

Authors:

Donald R. Monson
Senior Manager, Analytics
Deloitte Financial Advisory Services LLP
dmonson@deloitte.com
+1 312 486 4492

Sara L. Vandermark
Senior Manager, Analytics
Deloitte Financial Advisory Services LLP
svandermark@deloitte.com
+1 212 436 6490

Leaders:

Georges Korsun
Director, Analytics
Deloitte Financial Advisory Services LLP
gkorsun@deloitte.com
+1 212 436 5981

Samir Hans
Principal, Analytics
Deloitte Financial Advisory Services LLP
shans@deloitte.com
+1 571 882 8410

Omer Sohail
Principal, US FSI Analytics Leader
Deloitte Consulting LLP
osohail@deloitte.com
+1 214 840 7220