How a Hybrid Anti-Fraud Approach Can Save Government Benefit Programs Billions

Case Studies of Organized Crime Ring Defrauding Federal Subsidy Programs
Contents

Executive Summary .......................................................... 1
The New Face of Fraud and Detection ................................. 1
Moving Beyond ‘Pay and Chase’ to Prevention .................. 2

The Goal: Detect Fraud Earlier ........................................... 2
The SAS® Fraud Framework: Connecting the Dots for You ........ 3
Applying the SAS® Fraud Framework to Cases of Federal Subsidy Abuse ...... 3

The Foundation: A Solid Data Repository ..................... 5
Data Integration ................................................................. 5
Data Cleansing and Quality ............................................. 5
Probabilistic Matching ..................................................... 5
Integration of Unstructured Content ................................. 6

Fraud Detection: A Proven, Hybrid Solution .................. 6
Rules-Based Algorithms .................................................. 6
Anomaly Detection .......................................................... 7
Predictive Modeling ......................................................... 8
Social Network Analysis (SNA) ........................................... 8

Fraud Reporting: Business Intelligence That Provides Data to Act Upon .................... 9

SAS®: Enabling a 24/7 Fraud Detection Workflow ................ 11
Integrated Case Management Functionality ................... 11
A System for Continuous Learning and Improvement ............ 12

SAS: Evolving Solutions to Keep You One Step Ahead of Organized Crime ................. 13
Executive Summary

The New Face of Fraud and Detection

Professional criminal enterprises inflict more damage and losses on society than opportunistic individuals. Yet today, anti-fraud program managers focus primarily on detecting outliers, researching standalone fraud events and acting on tips. And why not? Most anti-fraud analytic methods are focused at the provider or claim level; very little - if any - advanced data analysis focuses on the detection of collusive behavior among multiple entities. The criminal enterprises that exploit agencies providing everything from government loans to subsidized housing to food stamps know that anti-fraud investigators struggle to connect the dots - and fraudsters use this to their advantage. While law enforcement gets distracted with the latest fraud scheme, professional fraud organizations constantly change their tactics - finding their way back into the system after being kicked out for fraudulent behavior or always staying one step ahead of the “pay-and-chase” game.

Examples abound. For instance, take the disturbing scale of fraud perpetrated against the federal food stamp program. Out of the more than 1,800 federal subsidy programs that the US government administers, the food stamp program is one of the largest, yet it has the fewest resources for fraud prevention and mitigation. As the number of food stamp recipients has soared to 44 million – up from 26 million in 2007 – and costs have more than doubled to $77 billion during this period, the 200,000 retailers approved by the US Department of Agriculture (USDA) to accept Supplemental Nutrition Assistance Program (SNAP) Electronic Benefits Transfer (EBT) are regulated by a staff of only 40 USDA inspectors. Given the limited resources for fraud prevention, a huge black market has developed whereby recipients exchange their food stamp benefits for cash. Savvy criminal enterprises find that these exchanges allow them to generate easy money with very little risk. Of the more than $2 billion in fraud and abuse reported by the USDA, the greatest losses are inflicted by corrupt retail merchants who exploit food stamp recipients. They typically pay a food stamp beneficiary about fifty cents in cash for every dollar in benefits they deduct from customers’ EBT accounts.

And fraud has gotten even easier now that the government uses electronic benefits cards. Mark McClutchey, Special Agent with the USDA, knows firsthand how corrupt merchant criminal networks traffic in these cards – and the importance of using technology to empower scarce anti-fraud human resources. From January-March 2011, working undercover posing as a neighborhood food stamp “go between” for homeless people who frequented a nearby shelter, he conducted 58 EBT transactions at two Detroit neighborhood stores for more than $31,000 in cash. On five different occasions, the suspicious owners accused McClutchey of being “a fed” even as they handed the undercover agent cash while deducting credits from his electronic benefit card. Apparently, raw greed trumped their fear of arrest. On one occasion, the undercover agent handed 13 EBT cards to an owner to cash in over a 24-hour time period. The owner generously dispensed some free advice to go with the cash: “You gotta play the game, play the system.”

Investigations, law enforcement undercover operations and prosecutions across the country reveal that well-organized crime networks with business fronts are responsible for a disproportionate amount of these losses. They threaten the integrity of federal subsidy programs and drain precious taxpayer dollars. Consider these recent cases:

- The operators of four Albany, NY, convenience stores were charged with grand larceny for defrauding the federal food stamp program of $1.6 million in a collusive scam between recipients and USDA-approved store owners. Prosecutors said that three of the stores were connected through family relationships and a referral network. During the investigation, it was also noted that the average transaction for convenience stores was $5.50 versus the $107.75 per average transaction at the four stores involved in the fraud schemes. One convenience store alone reaped more than $2.5 million in the scam while another netted more than $2.1 million before being caught and prosecuted. The schemes ran for several years and involved many collusive food stamp beneficiaries willing to accept cash for food stamp debits charged to their electronic benefit cards. The cash was often used to procure items prohibited for purchase using food stamps, such as alcohol.

- From 2003-2006, Apolinar Collado rang up more than $1.6 million in fraudulent food stamps at six different Connecticut convenience stores. This was after he was convicted in 1998 and served 18 months in prison for defrauding the federal food stamp program. An illegal alien, Collado was deported after his 1998 conviction. Not only did he manage to re-enter the US after his deportation, but he was also able to set up an even more extensive food stamp fraud operation starting in 2003 using “straw” owners and operators of several stores. Incredibly, four stores operated at the same location. Each time the store was disqualified from the federal food stamp program for suspected abuse of the food stamp program, it was “sold” to a new Collado-recruited straw owner, and the scam continued.
A Dearborn, MI, man was one of eight people charged in a federal crackdown on food stamp fraud. The ringleader, Jihad “Jim” Sayed, was indicted for allegedly redeeming food stamp benefits for cash and allowing food stamp benefits to be exchanged for items not authorized under federal regulations such as tobacco, liquor, clothing, household appliances and furnishings. To date, the initiative, termed the Bridge Card Enforcement Team (BCET), resulted in 125 arrests, 169 search warrants, 98 guilty pleas, about $21.1 million in court-ordered fines and restitution, and $3.2 million in forfeitures.

In each of these cases, the fraudulent owners and their associates used multiple businesses, family ties, an active referral network, false documentation, and brazen disregard for law enforcement and regulators to facilitate long-running fraud schemes.

One particularly insidious aspect of this type of fraud was exposed by the Chief of the FBI’s Terrorist Financing and Operations Section (TFOS) when he testified before the US House of Representatives, stating, “Within the United States, Hezbollah associates and sympathizers have engaged in a wide range of criminal activities to include money laundering, credit card fraud, immigration fraud, food stamp fraud, bank fraud and narcotics trafficking.” The food stamp program is so often exploited by terrorist organizations to finance their operations that the Investigative Project on Terrorism (IPT) established by terrorism expert Steve Emerson includes food stamp fraud cases as regular components of its weekly terrorist activity report. It is not uncommon to see these same store operators offering prohibited money laundering Hawallah services – the invisible black market currency exchange system that is favored by terrorist organizations – from the same stores.

What’s even more frustrating is that without the right analytical tools and insights, similar fraud schemes by other unidentified organized crime networks are underway today, flying under the radar of law enforcement. So if traditional methods and systems used by federal agencies aren’t enough to detect this scale of fraud earlier, what is?

The answer is this: hybrid analytics that empowers law enforcement to go on the offensive with fraud operators – and do so without disrupting the efficient and timely delivery of the benefits. Using the fraud stories described above as an example, this paper will explore how your law enforcement or government agency can use multiple analytical methods to proactively identify crime rings in their infancy – and spare federal taxpayers billions of dollars misspent on spurious benefits payouts each year.

The Goal: Detect Fraud Earlier

In SAS’ experience working with government agencies, it’s clear that proactively identifying complex fraud schemes (and the criminals behind them) is only possible when agencies use a multifaceted, anti-fraud detection approach that combines sophisticated data integration with a hybrid analytical approach that includes:

- **Rules** to filter activities and transactions for fraudulent behavior.
- **Anomaly detection** to detect individual and aggregated abnormal patterns.
- **Predictive models** based on known fraud cases so you can look for similar characteristics in future claims.
- **Social network analysis** to identify relationships pertinent to fraud that would normally escape notice.

Moving Beyond ‘Pay and Chase’ to Prevention

Using their ingenuity and flexibility, these groups take advantage of the federal government’s inability to connect the dots between state and federal government databases that contain such information as beneficiary and retailer information, retail store sales taxes, wage taxes, income tax, corporation records, driver’s licenses, business licenses, criminal records and deportations. Connecting the dots is what’s required to identify anomalies and fraud patterns, as well as match and link the malignant social network. It is often the patterns and conduct of the recipients and their relationships with the retailers that are the earmarks of a fraudulent food stamp ecosystem.

This combination of data integration and a hybrid analytical approach is the key to uncovering crime organizations in the first few months of operation, which greatly limits potential losses to the government and taxpayers.
The SAS® Fraud Framework: Connecting the Dots for You

The SAS Fraud Framework is designed from the ground up to support this kind of hybrid analytical approach. Working behind the scenes, integrated SAS analytical applications quickly connect the dots for you in a variety of ways. Each strand of analytically determined connections represents another way of looking at any given fraud event as part of a potentially much larger threat. In this way, the SAS Fraud Framework’s applications work together to:

- Provide strategic insight into threats, trends and risks.
- Deliver a holistic enterprise view of fraudulent behavior.
- Rapidly test, simulate and deploy models/rules without dependence on IT.
- Support prepayment detection and prevention.

The various analytical techniques of SAS’ hybrid approach are summarized in Figure 1 below.

Using a Hybrid Approach for Fraud Detection

<table>
<thead>
<tr>
<th>Enterprise Data</th>
<th>Suitable for known patterns</th>
<th>Suitable for unknown patterns</th>
<th>Suitable for complex patterns</th>
<th>Suitable for associative link patterns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Employer Data</td>
<td>Rules</td>
<td>Anomaly Detection</td>
<td>Predictive Models</td>
<td>Social Network Analysis</td>
</tr>
<tr>
<td>Application Data</td>
<td>Rules to uncover known fraud behaviors</td>
<td>Uncover unusual behaviors</td>
<td>Identify attributes that differentiate known fraud behavior</td>
<td>Knowledge discovery through associative link analysis</td>
</tr>
<tr>
<td>Deceased Persons List</td>
<td></td>
<td></td>
<td></td>
<td>Examples:</td>
</tr>
<tr>
<td>IRS/State Agency</td>
<td></td>
<td></td>
<td></td>
<td>• Beneficiary associated with known fraud.</td>
</tr>
<tr>
<td>Known Bad Lists</td>
<td></td>
<td></td>
<td></td>
<td>• Identity manipulation across employers.</td>
</tr>
<tr>
<td>Third-Party Data</td>
<td></td>
<td></td>
<td></td>
<td>• Collusive networks of recipients and store owners.</td>
</tr>
</tbody>
</table>

Hybrid Approach

Proactively applies combination of all four approaches at entity and network levels

Figure 1: Key components of the SAS® Fraud Framework.

Applying the SAS® Fraud Framework to Cases of Federal Subsidy Abuse

So exactly how could the SAS Fraud Framework have helped law enforcement identify the kinds of crime rings plaguing the USDA’s food stamp program faster and earlier? First, let’s look at this from a high-level perspective to demonstrate the various capabilities supported by the framework and their potential value – specifically through its ability to provide:

- Improved data integration.
- Rules-based algorithms.
- Anomaly detection.
- Predictive models.
- Social network analysis (SNA).
The following table summarizes this value.

<table>
<thead>
<tr>
<th>Capability</th>
<th>Explanation</th>
<th>Business Cases From the Food Stamp Fraud Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Improved data integration</td>
<td>Enables you to bring together data from a wide variety of third-party data sources and combine it with federal agency data on a national level. SAS Data Integration also includes sophisticated matching functionality so you can quickly identify missing and nonstandard data, including unstructured data from call center and investigator notes.</td>
<td>Identification of deceased beneficiaries.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Identification of beneficiaries with known fraudulent history in other government agencies.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Identification of false storefronts and inappropriate addresses (e.g., UPS stores or apartment buildings) not identified with legitimate businesses.</td>
</tr>
<tr>
<td>Rules-based algorithms</td>
<td>You can create rules to detect known types of fraud or abuse based on specific patterns of activity.</td>
<td>Flagging a store owner versus an agency-approved list of benefit providers.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Flagging a beneficiary on incarcerated or deceased persons lists.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Flagging benefits provided by a store that was recently instituted or changed its place of service (e.g., a store changed location in January and benefit was provided for an old address in March).</td>
</tr>
<tr>
<td>Anomaly detection</td>
<td>Using distributional analysis, you can analyze variables to identify which applications and beneficiary patterns are extreme outliers relative to others being processed. These outliers warrant additional scrutiny.</td>
<td>The average transaction amount by store exceeds the norm or has spikes over time.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The average number of transactions exceeds the norm for a particular geographical area.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>The average frequency of benefits provided to a particular individual exceeds the norm.</td>
</tr>
<tr>
<td>Predictive models</td>
<td>Using insights gained from known fraudulent behavior, you can create formulas that will enable SAS software to score transactions or beneficiaries for the probability of fraud.</td>
<td>Probability of a phony storefront, based on statistical correlation with facts present for other phony storefronts that have been shut down.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Probability of a compromised beneficiary ID, based on known stolen IDs in past claims.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Probability of fraudulent behavior based on profile of similar fraudulent behavior patterns.</td>
</tr>
<tr>
<td>Social network analysis (SNA)</td>
<td>Analyze social networks with link analysis, which uses statistics to assign probabilistic associations between entities. For example, by analyzing names, addresses, bank account numbers, dates of birth, SSNs and more, you can identify potential fraud rings that would otherwise appear to be disparate actors.</td>
<td>Linkage of multiple phony storefronts owned by several individuals with similar names and shared beneficiaries and spouses.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Suspicious linkage between store owners with different names but same addresses and bank account numbers.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Suspicious linkage of beneficiaries to known convicted benefit program fraudsters.</td>
</tr>
</tbody>
</table>
SAS technology can deliver these capabilities by taking a three-pronged approach that includes:

- **A data repository** that aggregates and integrates data from many sources and creates a solid foundation for rapid, comprehensive analysis.
- **Sophisticated analyses** for powerful fraud detection and prevention.
- **Fraud reporting** via business intelligence visualizations that give you data you can act upon.

### The Foundation:
#### A Solid Data Repository

As with any analytics solution, bad data means inaccurate, incomplete insights – hardly useful for investigators of major fraud schemes. One of the reasons the criminals in these stories were able to perpetrate their crimes for so long was due, in part, to data that was fragmented across public and private data sources and systems; no one had ever looked at available data holistically and connected the dots.

That's why the SAS Fraud Framework runs on a solid data foundation enabled by SAS' world-class data integration technologies for data quality and integration. SAS Analytics can make the intelligent linkages between currently disparate data sources and use them to help you quickly identify instances of potential fraud.

### Data Integration

With SAS data quality and integration tools, you can merge any public and private data sources and have a complete data set to analyze and connect the dots. By melding public data – beneficiary and retailer information, retail store taxes, wage taxes, income tax, driver’s licenses, business licenses, criminal records, deportations – you can use analytics to integrate and then correlate the data points to reveal the larger picture.

For example, in the Apolinar Collado example mentioned in the Executive Summary, even simple public data, such as residential and commercial phone directories, Google maps, or criminal conviction records, could have been integrated and used to identify those addresses and phone numbers clearly not associated with legitimate convenience stores. Analytics provided through the SAS Fraud Framework could then be used to thoroughly screen store applications at their enrollment, thus barring Collado at the point of entry into federal programs. SAS data integration tools could also automate the process of making the millions of intelligent linkages between the various data sources – for example, by matching deceased or incarcerated store owners, deceased beneficiaries, invalid Social Security numbers, and family relationships across store owners.

### Data Cleansing and Quality

The SAS Data Quality Solution provides an enterprise solution for profiling, cleansing, augmenting and integrating data to create consistent, reliable information. With the SAS Data Quality Solution, you can automatically incorporate data quality into data integration and business intelligence projects to dramatically improve returns on investigations and analyses. For example, you can profile, monitor and actively manage the quality of data being collected from various sources, as well as integrate and standardize data across multiple systems. The software also allows you to define data correction rules to reflect organizational changes and cleanse data. As a result, you can provide investigators with trustworthy information needed to make informed decisions regarding what to investigate.

How would this solution have helped investigators uncover Collado’s fraud earlier? Often, when a suspect is filling out a form, certain information is intentionally left out. The intent is to have the processor fill in the information or to omit information that could link the investigator to the suspect. As the data cleansing and quality process is conducted, users can analyze the cleansing process to uncover trends in the application process. For example, a fraudulent suspect may always omit answers in boxes 7, 14 and 21 on a specific form. Knowing this trend, when an application is missing answers to 7, 14 and 21, it can be tagged as suspicious. In the case of Collado, identifying trends in the quality of the various phony convenience store applications could have tipped investigators to take a closer look at the six different storefronts that he established.

### Probabilistic Matching

The methods of crime rings defrauding federal subsidy programs tend to rely heavily upon the fact that it is difficult for agencies and investigators to link names and aliases between disparate data sources and match them to other data, such as Social Security numbers (SSNs). To address this fraud opportunity, agencies and investigators need to cleanse data and create soft links between public data and internal sources, generating matches based upon names and other identifying information - even when hard links aren’t available, as is often the case with publicly available data. For example, it would be inconceivable to publish an individual’s SSN and beneficiary information next to other information in a public data source, such as addresses and phone numbers published in local phone books, yet these linkages must be made by agencies to identify potential fraud.
Automate the extraction and integration of call center data (one example of unstructured text) with claims and benefits payment data to quickly detect new patterns that indicate potential fraud.

Fraud Detection: A Proven, Hybrid Solution

After agencies have a solid data foundation, they can apply analytics to identify patterns, trends and other outliers that are associated with likely fraud operations. But it's important to note that no single detection technique is capable of systematically identifying large, collusive rings like those described in the stories discussed previously. Only through the combination of multiple analytical approaches could the perpetrators of these crimes have been detected without insight from an inside informant.

For example, rules and anomaly detection cast the net wide, resulting in too many false positives. But when combined with predictive modeling, which improves accuracy, you can significantly reduce the frequency of false positives to improve investigator efficiency. Finally, you can layer on link analysis to ensure that all potential parties involved in a scheme are properly considered before they are determined to be an outlier or not.

SAS recommends agencies take a hybrid approach, which integrates knowledge of existing schemes, powerful predictive analysis capabilities, and modeling of relationships among entities in the benefits system to recognize fraud earlier in the process – and even stop it before it occurs. While each method is limited when used separately, together they present a formidable deterrent to fraud that would have led to the detection of these crime rings much, much sooner.

Integration of Unstructured Content

Rich but seldom exploited data sources – call center data, operative notes (such as Agent McClutchey’s notes mentioned in the Executive Summary), investigator notes and death certificates – are considered unstructured since the information they contain is not stored in a tabular format. Most organizations are incapable of integrating unstructured data in an automated fashion, let alone using that valuable information in analysis or investigations. If agencies and investigators had been able to integrate unstructured data from a variety of sources, investigators could have uncovered the patterns of suspicious behavior in the food stamp fraud rings more easily. By using this unstructured data, the entire beneficiary population could have been transformed into potential tipsters of fraud.

For example, by using the SAS Fraud Framework to analyze fraud hotline data from various states, as well as calls from the phony convenience store owners to assess or test reimbursement thresholds and fraud tactics, agencies could have identified the Jihad “Jim” Sayed fraud ring sooner. SAS software can automate the extraction and integration of call center data (one example of unstructured text) with the claims and benefits payment data to quickly detect new patterns that indicate potential fraud.

Rules-Based Algorithms

Rules-based systems test each transaction against a predefined set of algorithms or business rules to detect known types of fraud or abuse based on specific patterns of activity. For example, a rules engine may have alerted or rejected the $1.6 million of food stamp claims made by Collado. At the time the fraudulent activity was committed, Collado was: 1) a known felon convicted for defrauding the federal food stamp program, and 2) an illegal alien. Rules that identified known convicted federal food stamp criminals and/or immigration violations would have detected Collado before the $1.6 million benefit payments were sent.
Anomaly Detection

With anomaly detection, you can normalize events and set thresholds to identify outlier behavior. Using distributional analysis, you can then analyze variables to identify claims and practice patterns that are extreme outliers relative to the rest of their respective distribution. In this manner, you can use statistical outliers to identify new patterns of fraud that are otherwise unknown.

Value as a fraud detection tool

Business rules such as checking incarcerated lists, immigration status, deceased persons lists, known bad lists, etc., can be used to alert investigators to claims that are obviously fraudulent. Similarly, transactions can be red-flagged if the benefit provider recently changed place of business, address or provided a similar service to multiple family members within a narrow time frame – or, if a large number of claims happen immediately following a recent change in location or enrollment.

Why it must be part of a hybrid approach

However, a rules-only approach presents multiple limitations:

- Known fraud and improper payments must be specifically enumerated and added to the business rule repository. As soon as rules are created - and fraud rings “test” to verify their existence – fraudsters can vary their billing habits and behavior to elude rule detection. In other words, professional fraudsters can easily game the system by playing according to new and undetected rules.

- It is next to impossible to create a rule for every type of fraud scheme known. There are simply too many permutations of potential types of fraud – and therefore types of rules – to catch every instance of fraud.

- Some fraud can look legitimate if only a rules-based approach is used. In the Sayed example, the benefits were correctly redeemed. However, the correct benefits were never rendered but instead, exchanged for tobacco, liquor, clothing and household appliances. In short, it was still fraud, but rules would not have caught it.

- Rules engines may uncover large numbers of suspicious claims, many of which will be false positives that create a distracting level of noise. So while a rules-only analytics approach may have ultimately identified the fraudsters in each of the examples shown here, used independently, it may have disrupted operations for legitimate providers and cost federal agencies a great deal of money to manually review unnecessary alerts.

For these reasons, rules-based algorithms must be combined with other analytical methods that can cover these gaps and enable early detection of fraud.
Predictive Modeling

Predictive modeling tends to be more accurate and reliable than other analytical methods for fraud scheme discovery. Predictive modeling looks at past behavior (fraud incidents) and determines which variables (frequency of claims, dollar value patterns, region, etc.) have the most impact on the outcome (the fraud incident). You can combine the strongest variables into an algorithm that can be applied to current data to determine where otherwise undetectable patterns are occurring. By doing so, you can calculate a score or likelihood, making it possible to determine where fraud is more likely to be occurring.

Value as a fraud detection tool

With predictive modeling, you can detect even slight variations between fraud schemes, enabling investigators to better understand how fraud morphs over time and across geographic areas. Predictive modeling is also superior to a rules-only approach in that the investigator need not define specific criteria that can easily be evaded. Rather, predictive modeling lets the data drive the identification, combination and weighting of the parameters that contributed to fraudulent activity.

As a result, fraud schemes can be discovered based not only on front-end assumptions on the part of the investigators, but also on what the data and models score as potentially fraudulent. In addition, beneficiaries, stores, store owners, etc., scored by the predictive models can be prioritized in terms of likelihood of fraud or the dollar amounts associated with the scheme, thus reducing false positives and letting investigators focus on only the most important fraud cases first.

The following are examples of predictive models that may have been employed to identify the fraud schemes mentioned in the Executive Summary:

- **Alerting investigators to a potentially fictitious/phantom storefront**: With rules and anomaly detection alone, federal agencies could not — with a high degree of certainty — differentiate between new stores that are legitimate and those that are phantom storefronts established for fraud. Using predictive modeling, federal agencies could analyze patterns in the data from past phony storefronts that were opened for a short period of time, experienced high utilization and were then promptly shut down. Agencies could then overlay that formula on new stores to predict which ones are most likely to be fraudulent (and later shut down).

- **Probability of a compromised beneficiary ID**: Again, with rules and anomaly detection alone, federal agencies would be uncertain about how to differentiate between legitimate services that beneficiaries actually received versus services that were never rendered but were ultimately paid for. In many cases, an individual beneficiary could be receiving legitimate services from honest providers while at the same time, a criminal ring could be billing for services that were never rendered to that person, as in the Sayed example. To distinguish between them, a predictive model could be used at the beneficiary level (to indicate normal and likely behavior patterns) or at the provider level (by modeling other fraud rings that operated with stolen beneficiary IDs).

Why it must be part of a hybrid approach

However, using predictive modeling alone has limitations as well. Predictive modeling only profiles known behaviors. Although models can adapt quickly to new schemes as they emerge and can be used to detect abnormal behavior, they identify individuals (e.g., providers or beneficiaries) and do not connect high-scoring individuals with low-scoring individuals who may be part of a fraud ring.

Social Network Analysis (SNA)

Social network analysis, also known as linkage analysis, combines the connections established between disparate source data and groups of potentially fraudulent actors into a single, malignant social network. Entities may include locations, service providers, beneficiaries, family members of fraudsters, addresses or telephone numbers, to name a few.

In the past, constructing networks manually has been quite labor-intensive, and only highly skilled investigators working with manual tools, push pins and string used over long time periods have been able to identify enterprise-level activity. However, social network analysis can be fully automated, with the system continuously updating the interrelated networks and rescoring for fraud. As such, identification of a single suspicious activity or entity in the ring can expose an entire network of highly collusive behavior.

Value as a fraud detection tool

Social network analysis establishes frequencies among the various linkages of entities in the network and then runs those linkages back through anomaly detection and predictive modeling to identify collusive networks that are statistically unusual. This is a powerful tool to facilitate early identification of collaborative fraudulent networks, as it enables program
managers to act early to limit losses, prioritize their efforts and aggregate losses to enable high-impact prosecutions.

Social network analysis could have been employed to identify the fraud schemes described in the Executive Summary in the following ways:

- The Albany crime ring was created using multiple storefronts, not only to maximize illicit revenue, but also to mask fraudulent behavior on the part of any one store. Social network analysis would have identified that multiple stores were owned by several individuals with similar names and that they shared in a large percentage of similar customers who (not coincidentally) all were paid unusually high average transaction amounts at the four stores.

- Using link analysis on the address of each store, social network analysis would have also uncovered that four of the stores in the Collado crime ring were at the same location.

- Had it been available, social network analysis could have identified the unusual connections and relationships between the 125 individuals arrested in the Sayed crime ring.

Why it must be part of a hybrid approach

Social network analysis provides a holistic view of fraud. Typically, organizations only look at individuals and account views to analyze activities. By combining the social network analysis with rules, anomaly detection and predictive modeling, investigators can go beyond the typical view of fraud to see a bigger picture, including related perpetrators, and to gain a clearer understanding of all the activities and relationships at a network level. The hybrid approach also gives investigators a better understanding of emerging threats so they can take action to prevent substantial losses before they happen.

Fraud Reporting: Business Intelligence That Provides Data to Act Upon

The final element of a successful fraud detection and prevention solution is business intelligence software, which generates reports that turn analytical results into real and useful insight. Consider the power of providing investigators with social network analysis that is presented in a graphical manner, as illustrated in Figure 2. These graphics show how SAS Business Intelligence software instantly displays the relationships of an entire network. The red background of certain icons immediately focuses the viewer’s attention on those individuals whose activities and circumstances have been flagged as suspicious. Time sliders show the development of the network over time. And investigators can easily peel back the investigation to uncover previously unknown suspects related to the network.
Figure 3 provides an example of a typical alert queue – something familiar to most fraud investigators. This investigation tool displays the results of the SAS advanced analytics and hybrid approach, as discussed above. The tool is highly customizable – in this instance, it displays the most pertinent information to the fraud investigator to facilitate decision making or alert triaging. The software also provides capabilities for scoring, de-duping, prioritizing and routing pertinent alerts to appropriate resources, as well as a description that gives more detail on why the alert was generated. Note that the graphics show what the average transaction looks like (what’s shown in blue) compared to the transaction in question (what’s in red). Color coding helps investigators immediately focus on the attributes of a transaction that is likely fraudulent.

**SAS® Solutions Tailored to Identify Fraud in Diverse Government Agencies**

To maximize the effectiveness of any fraud detection and elimination workflow, the SAS Fraud Framework is customized to meet the needs of different government agencies. These tailored solutions, which use what has already been done for different industries, address fraud in the following areas:

- Insurance.
- Welfare and social services.
- Medicare and Medicaid.
- Tax.
- Unemployment and workers’ compensation.
- Grants and purchasing.
SAS®: Enabling a 24/7 Fraud Detection Workflow

As illustrated below in Figure 4, the SAS Fraud Framework leverages all of the capabilities described above to support an integrated workflow for analyzing data as it comes into the integrated data source and for detecting potential fraud in near-real time. Data from a wide variety of operational data sources is aggregated to create a single, clean data source that’s optimized for analytics:

- On the left, data from a wide variety of sources is integrated, centralized and cleansed.
- Data is then dumped into the SAS Analytics engine for analysis and alert generation (as illustrated in the blue box).
- Results of analytics are then detailed in business intelligence reports and alerts (as shown in the lower right-hand box).
- The framework also allows you to package the results into a case management system.

The most important step in the process is the learn-and-improve cycle: Results are fed back into the system to tweak the models and rules for improved results and reduced false positives.

SAS® Fraud Framework

Process Flow

Figure 4: Workflow powered by the SAS Fraud Framework.

Integrated Case Management Functionality

The workflow powered by the SAS Fraud Framework can be integrated with SAS Enterprise Case Management, which empowers investigators to make full and effective use of the analytical data and insights. As you know, case management is the backbone of documenting investigation processes, exposures and losses regarding any type of financial crime that may pose a risk to an organization. It is also relied on to provide information for financial reporting (e.g., fraud losses) and is the primary resource for filing regulatory reports to government agencies. In addition, the disposition of cases under investigation is critical to enhancing an organization’s future monitoring and overall operational efficiencies.

SAS Enterprise Case Management enforces best practices and proper gathering of evidence and can greatly reduce the cost of investigations. It provides a structured environment for managing investigation workflows, for attaching comments or documentation, and for recording financial information, such as exposures and losses. The solution also provides advanced dashboards that go beyond providing traditional operational performance metrics to enable management oversight and analytical reporting. And dashboards also let users track the performance of investigative functions and trends related to financial crime exposure.
A System for Continuous Learning and Improvement

The SAS Fraud Framework is also a system for continuous learning and improvement, which is essential. Since fraud schemes evolve and morph over time, prevention and detection software must evolve and morph with them. Each time an alert is triggered or a vendor eligibility referral is passed, the results are stored within the Intelligent Fraud Repository as known outcomes. The predictive models used in the framework access this repository of known outcomes as they apply analytic approaches used in the hybrid approach discussed previously. By registering and leveraging known outcomes in the repository, you can turn this into knowledge to stop future payments. Specifically, this feedback loop allows the SAS Fraud Framework to continue to learn and be more precise with risk scoring by raising more alerts on entities and networks that have attributes similar to confirmed, known bad actors. If this feedback loop is in place, the SAS Fraud Framework and associated analytical approaches never stop learning and evolving. As new schemes are detected and uncovered, this knowledge becomes part of the SAS Fraud Framework and is automatically and instantly applied in the future.

In addition, the SAS Fraud Framework’s model management features an easy-to-use, graphical user interface that guides fraud analysts through a repeatable process for registering, testing and validating models. Model management functionality also supports “champion versus challenger” capabilities so they can confirm best-fit models. Analysts can also analyze how models perform over time using a series of lift charts representing result deviation and response stability. This analysis allows analysts to continuously monitor fraud model performance and determine when to reconfigure models.

SAS® Enterprise Case Management gives investigators more power, flexibility and task automation, without having to rely on IT for support.

The solution enables investigators to:

- Receive alerts requiring investigation from multiple monitoring systems.
- Review incoming items prior to creating or linking an incident to a case.
- Customize activities and automate workflows (can develop different workflows for different case types).
- Assign activities in the workflow to individuals or groups of users.
- Import existing records and historical information by custom ETL to prepopulate customer fields and prevent rekeying errors.
- Create, edit and view a case based on user permissions.
- Classify cases by type, category and subcategory.
- Attach documents, video files and other digital media to a case.
- Set default fields and values on screens based on case type.
- Create audit records containing user identification, a time stamp and date when actions are performed.
- Generate batch files for regulatory reporting via e-filing.
Federal agencies providing subsidies to citizens and legal aliens must place themselves on the offensive by incorporating strategies and tools that make the commission of fraud a risky undertaking. SAS is uniquely positioned to team with federal agencies to make this happen. The SAS Fraud Framework provides an end-to-end framework for detecting, preventing and managing all types of abuse of federal subsidies and loan programs. Only SAS combines all of the approaches outlined in this paper in a single, integrated, commercial-off-the-shelf (COTS) software offering, providing improved efficiency and decreased total cost of ownership in implementing these capabilities.

Furthermore, SAS is universally recognized as the worldwide leader of advanced analytics. SAS’ market share in predictive modeling alone is more than double its closest competitor’s. Only SAS can provide federal payers with an open, high-performance and scalable solution for implementing analytics throughout your fraud detection strategy at the point of provider and beneficiary eligibility screening, prepayment within the benefit payment system, or in post-payment review.