Fraud & Risk Detection in Government

Deborah Pianko, Principal - SAS Security Intelligence
But I have great employees. They would never do that.
Rita Crundwell - $54 Million over 22 years
Dixon, Illinois Comptroller

- Secret bank account
- Wrote checks from the fund payable to "Treasurer," which she would deposit into account.
- In 1991, she stole $181,000, while in 2008 alone she managed to embezzle $5.8 million.
- Bought horses, several cars, a second house and a million-dollar motorhome
- She had an $80,000 city salary
- Used money to support horse breeding business

Largest municipal fraud in U.S. history.

Praised by officials for “(looking after) every tax dollar as if it was her own.”
Harriette Walters - $48 Million over 20 years
D.C. Office of Tax & Revenue – Manager, Real Property Tax Division

Walters used her position at OTR to create false property tax refund vouchers that produced millions of dollars of fraudulent refund checks. From June 1989 through August 2007.

- Shopping sprees at Nordstrom, Neiman Marcus - $2.3 million
- Designer purse - $25K
- 45 trips to Las Vegas and Atlantic City
- Cash and checks to co-workers $1.2 million
Who and What is an Insider Threat - Prioritizing

What do you want to protect, and from whom?

**Individuals**
- Current or Former
- Full-Time Employees
- Part-Time Employees
- Temporary Employees
- Contractors
- Trusted Business Partners

**Organization’s Assets**
- People
- Information
- Technology
- Facilities

**Intentionally or Unintentionally**
- Fraud
- Theft of Intellectual Property
- Cyber Sabotage
- Espionage
- Workplace Violence
- Social Engineering
- Accidental Disclosure
- Accidental Loss or Disposal of Equipment or Documents

**Negatively Affect the Organization**
- Harm to Organization’s Employees
- Degradation to CIA of Information or Information Systems
- Disruption of Organization’s Ability to Meet its Mission
- Damage to Organization’s Reputation
- Harm to Organization’s Customers

*Copyright © SAS Institute Inc. All rights reserved.*
### There is Not One Type of Insider Threat

#### 3 Behavioral Profiles – IT Saboteur, Financial Fraudster and IP Thief

<table>
<thead>
<tr>
<th></th>
<th>IT Sabotage</th>
<th>Fraud</th>
<th>Theft of Intellectual Property</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Current or former Employee?</strong></td>
<td>Former; disgruntled</td>
<td>Current; stressed</td>
<td>Current (within 30 days of resignation); can be disgruntled but not necessarily</td>
</tr>
<tr>
<td><strong>Type of position</strong></td>
<td>Technical (e.g., sys admins, programmers, DBAs)</td>
<td>Non-technical (e.g., data entry, customer service) or their managers</td>
<td>Technical (e.g., scientists, programmers, engineers) or sales</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td>Male</td>
<td>Fairly equally split between male and female</td>
<td>Male</td>
</tr>
<tr>
<td><strong>Target</strong></td>
<td>Network, systems, or data</td>
<td>PII or Customer Information</td>
<td>IP (trade secrets) or Customer Information</td>
</tr>
<tr>
<td><strong>Access Used</strong></td>
<td>Unauthorized</td>
<td>Authorized</td>
<td>Authorized</td>
</tr>
<tr>
<td><strong>When</strong></td>
<td>Outside normal working hours</td>
<td>During normal working hours</td>
<td>During normal working hours</td>
</tr>
<tr>
<td><strong>Where</strong></td>
<td>Remote access</td>
<td>At work</td>
<td>At Work</td>
</tr>
</tbody>
</table>

There is Not One Type of Insider Threat
Low Tech
Methods are detectable

Insiders’ means were not very technically sophisticated.

Non-technical subjects were responsible for 65 (81 percent) incidents. Seven were external attackers, but their methods were also non-technical.
Managers vs. Non-Managers
Managers are more trusted – but do more harm

Fraud by managers differs substantially from fraud by non-managers by damage and duration.

Of 61 subjects, 31 (51 percent) were managers, VPs, bank officers, or supervisors. The median results show that managers consistently caused more actual damage ($200,106) than non-managers ($112,188).
Collusion

With exception of a certain type of IP Thief (The “Ring Leader”)

Most cases do not involve collusion.

There was not a significant number of cases involving collusion, but those that did occur generally involved external collusion (i.e., a bank insider colluding with an external party to facilitate the crime).
Statistics are human beings with the tears wiped off.
-Paul Brodeur, Outrageous Misconduct
Insider Threat – 2016 Statistics

- Forty-seven percent (47%) of survey participants reported that an insider incident was committed against their organization.

- More than one in four (27%) of all attacks against their organization were committed by insiders.

- The most common insider incidents were customer records compromised or stolen, confidential records (trade secrets or intellectual property) compromised or stolen, and private or sensitive information was unintentionally exposed.
Insider Threat Detection & Continuous Monitoring
Dwell Time - Low and Slow
Prevention may be possible through better onboarding, but earlier detection is key

There are, on average, over 5 years between a subject’s hiring and the start of the fraud.

There are typically 32 months between the beginning of the fraud and its detection.

Criminals who executed a “low and slow” approach accomplished more damage and escaped detection for longer.

- Ongoing reinvestigation/regular monitoring after hire
- Termination and departure monitoring
- Employee awareness and path to report unusual behavior
Examples of Detection Approaches

- File downloads which are excessive for the employee’s peer group
- Atypical privileged account access (e.g. VIP taxpayers)
- Abnormal viewing/modification of taxpayer accounts (or lack thereof)
- Analysis of web activity
- Sentiments (i.e. feelings) expressed in work email or instant messages that are highly correlated with insider threat
- Unusual time-of-day access
- Employees’ connections and sphere of influence in the organization
- Printer activity
- Use of external media devices
- Override of technology settings
- Monitoring of database queries and/or copying
- Data exfiltration monitoring (e.g. plagiarism detection)
The Quest for Defining Normal – Peer Group Comparison
The Quest for Defining Normal – Unstructured Data
The Quest for Defining Normal – Time of Day Access

Time of Day Access

<table>
<thead>
<tr>
<th>Day</th>
<th>Date</th>
<th>Owner</th>
<th>PC</th>
<th>Service</th>
<th>Ticket</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2016-01-19</td>
<td>ALR0766 - Armando Lucas Rivas</td>
<td>pc-2769</td>
<td>Apply OS Patches</td>
<td>AK199</td>
</tr>
<tr>
<td>1</td>
<td>2016-01-19</td>
<td>KFM0249 - Kamal Forrest Miranda</td>
<td>pc-2495</td>
<td>Deploy Software - gpo</td>
<td>DSG7-</td>
</tr>
<tr>
<td>1</td>
<td>2016-01-19</td>
<td>DBJ3716 - Driscoll Beau Jefferson</td>
<td>pc-5925</td>
<td>Configure Disk Partition</td>
<td>U6L1-</td>
</tr>
<tr>
<td>1</td>
<td>2016-01-19</td>
<td>NJM2246 - Nicolas Jarrod Morton</td>
<td>pc-1208</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The Quest for Defining Normal – Relationships
The Lines are Blurry
Defense in Depth Required

“The best way to get into an unprepared company is to sprinkle infected USB sticks with the company’s logo around the car park.”

https://hbr.org/2014/09/the-danger-from-within
Tax Evasion
Organized VAT Fraud was a billion euro problem for the Belgian Government. This fraud typology is a high velocity fraud: “Carousel frauds are like floods. It is futile to believe that after the storm we can put back water in the riverbed with buckets. What is needed is to build dykes to prevent overflows”.

The solution is used to detect companies with a high probability of being involved in Organized VAT Fraud. The SAS Hybrid approach provides ultra-early detection from the first suspect VAT return or other suspicious behaviour. This Hybrid approach combines network analysis, business rules, predictive modeling and anomaly detection. The combination of these methods enables new networks to be discovered and to monitor known networks. The models are highly accurate (80% true positive rate). This makes the SAS Hybrid approach the ideal tax auditor.

The result is a reduction by 98% (from 1.1 billion € to 0.029 billion €) of the VAT carousel fraud. Analytics was key in achieving this result. The models are not only very accurate (80% true positive rates), they also provide ultra-early detection as from the first VAT declaration. Organized VAT fraud is now a controlled phenomenon. The system also enables international collaboration by identifying suspicious companies abroad.

Highlights

- VAT carousel fraud reduced by 98% (from 1.1 billion € to 0.029 billion €)
- Ultra-early detection
- SAS hybrid approach provides a high accuracy model (80% true positive rate)
SERIOUS & ORGANISED VAT FRAUD

VAT CAROUSEL

Belgium

Conduit Company (A)

Goods never actually have to leave the dock in France

May recycle same goods many times through ("carousel")

France

VAT-Free

1,000,000 Cost 0 VAT

Missing Trader (B)

Or can sell to totally different company in another EU country (e.g. Germany) – not always back to original company

French Treasury

Gets only 20,000 VAT from Company C

Offshore

220,000 VAT stolen

Buffer Company (C)

1,200,000 Cost 240,000 VAT

Profit Taker (D)

VAT

VAT-Free

1,100,000 Cost 220,000 VAT
Detection Techniques

\[ Q = \frac{1}{2w} \sum_{(u,v) \in A} \left( w_{uv} - \frac{w_u w_v}{2w} \right) \Delta(c_u, c_v) \]

\[ w = \sum_{(u,v) \in A} w_{uv} \]

\[ w_u = \sum_{v \in \delta_u} w_{uv} \]

THE BEST TECHNIQUE IS MULTIPLE TECHNIQUES

- Louvain algorithm
- Reach (Ego) networks
- Shortest path
Detection
Techniques

WHAT ARE WE LOOKING FOR

• Common network structures and activity between nodes that have shown in past to correlate with fraud (e.g. money flow, VAT return data, claims for refunds)
• Relationships to risky entities (measures of closeness, between-ness, etc.)
• Reduction of supercluster down to it’s key components (fuzzy matching)
• Predictive modeling – Have we seen this before?
• Anomaly/outlier detection – Does this entity act like it’s peers?
BUFFER COMPANY

Decision Tree

Profit Margin

<4% Profit Margin

Statistic
0: 50%
1: 50%
Count: 1799

>=4% Profit Margin

VAT Paid

<21% VAT Paid

Statistic
0: 35%
1: 65%
Count: 666

21% VAT Paid

Statistic
0: 51%
1: 49%
Count: 132

Ratio VAT fields

>=21% VAT Paid

Statistic
0: 30%
1: 70%
Count: 534

Monthly

Statistic
0: 63%
1: 36%
Count: 914

Declaration Frequency

Quarterly

Statistic
0: 41%
1: 59%
Count: 219

Statistic
0: 63%
1: 36%
Count: 914

Statistic
0: 59%
1: 41%
Count: 1113

Statistic
0: 30%
1: 70%
Count: 534

Statistic
0: 51%
1: 49%
Count: 132

Statistic
0: 35%
1: 65%
Count: 666

Statistic
0: 50%
1: 50%
Count: 1799

Filing Frequency
Fuzzy Matching

REDUCING THE SUPERCLUSTER

Disparate data sources

John Smith  13/01/1980  18 Queen Street  JN 12 34 56 A

John Smith  13/01/1980

J. Smith  13/01/1980  18 Queen Street  0208 123 45676

John Smythe  JN 12 34 56 A

Smith  13/01/1980  0208 123 45676
Machine Learning

Different techniques depending on the goal of the modeling exercise

Unsupervised Learning
Goal of exploration is not known in advance (no learning from cases)
Clustering; Association Rules

Supervised Learning
Learn & predict based on past cases
Regression; Decision Trees; Naive Bayes

More fish in same holes? Or find new holes to fish in?
Employer Miller LLP reported wage differs from wages on tax return by 10% or more. This return shows a variance of $40,226 / 45%.
Customs & Border Control
Yemen incident
2010

Two bombs headed to U.S., intercepted in Dubai and the UK and defused

Failed attempt to sabotage cargo aircraft using explosives hidden in printer cartridges

Found after a tip-off from Saudi authorities; may have been test packages before
Types of Fraud

- Drugs
- Tobacco
- Alcoholic beverages
- CITES
- Intellectual property rights (IPR) - Counterfeiting
- Precursors
- Tax and duty evasion
- Weapons and explosives
- Currency
- Nuclear materials
- Hazardous material
- Pornography / Paedophilia
- Other prohibitions and restrictions (including works of art, stolen vehicles, anabolic steroids etc.)
Tradeoff between Number of Alerts and Precision

Notes:
1) assume marijuana mailpieces make up 0.133% of all mailpieces
2) assume mail class of marijuana mailpieces are 62% Express, 38% Priority

US POSTAL MARIJUANA MAILPIECE TARGETING

- 7 Express mailpieces alerted 89.5% contain marijuana
- 676 Express mailpieces alerted 87.5% contain marijuana
Fraud Dual Use

• **What are Dual-Use commodities?**
  
  • Goods and technologies are considered to be dual-use when they can be used for both civil and military purposes.
  
  • Most industrialized countries have export controls on dual-use goods

**Problem:**

Chlorine is a dual-use item, both pool supply companies and terrorists might want to make use of it. But how to distinguish between the two?

**Solution:**

Identify rules to 1) detect export companies not declaring dual-use items 2) detect unusual patterns of export.
Fraud Dual Use

**ANOMALY DETECTION USING MARKET BASKET ANALYSIS**

- Basket analysis of all commodities exported by dual-use companies
- Find Associations between dual-use and other commodities - Identify strong and interesting associations
- Extrapolate to all export companies and look at their dual-use exports

**IF** 3917 – Tubes, Pipes and Hoses **AND** (3926 – Other Plastics **OR** 85 - Electrical Machinery) **THEN** 4016930000 - Gaskets

![Graph showing relationship between total dual use export and total items export with outliers highlighted.](Image)
Fraud Dual Use

- Profile companies exporting dual use vs. other companies
  - Export in a lot of countries
  - Export commodities in group 85 (electrical machinery) and 84 (nuclear reactors)
  - Big exporters
  - Mainly road transport
  - Also export 73 (Iron/Steel), 90 (Optical) and 39 (Plastics)
- Look at similar companies which aren’t exporting dual use commodities
  - 370 companies with similar profile
Real-time deployment

PREPROCESSING ESP

SCORING ESP

POST PROCESSING ESP

EVENTS

REFERENCE DATA (BATCH)

SUPPORTING EVENT DATA (BATCH)

DM

Historic Store

Batch Analytics

Alerts DB

Social Network Analysis

LASR Analytic Server

Partial ENS

Partial ENS

Full ENS

Partial ENS

Full ENS

Supporting Event Data (Batch)

Reference Data (Batch)
Thank You