
12th Fraud Seminar
Rutgers Business School
December 1, 2015
Presented by:
Deniz Appelbaum
Abdullah Al-Awadhi
Knowledge Based Expert Systems

- **Knowledge Based Expert Systems (KES):** “to construct computer software that performs/replicates tasks that are normally performed by human experts”
- Best suited for processes where the task is unstructured in design alternatives and where judgement and insight are required. The problem may be well defined, but the methodology is not.
- Requires transfer of knowledge from the human experts to the software – expensive and time consuming!
- Usually exists as a layer within a larger system
- Can be continually updated
- **Limitation:** Humans are not perfect experts!
- **Artificial Intelligence (AI):** software that tries to simulate humans decision making processes (ex: self driving cars), possibly can see patterns that are not easily detected by humans
Ultimate AI – Self Driving Cars!
Knowledge Based Expert Systems

- Expertise is difficult to acquire. Human Experts are expensive and in short supply!
- Accounting/Auditing problems tend to be rule intensive and can be solved with “if-then” rules
- The experts system must produce clearly identified solutions that most experts would agree with

Examples of Audit Expert Systems:
- Materiality judgements in audit planning
- Internal Control evaluations
- Going Concern Judgments
- Fraud detection on credit card transactions
INTRODUCTION: Procurement Cards

- P-Cards help reduce purchasing department costs and increase individual department purchasing decision-making (Daly and Buehner, 2003)

- Now given the large volumes of data and the advent of automated audit tools, internal auditors can mine 100% of the transactional data to detect anomalies (Murthy, 2010; Coderre, 2009 & 1999; Nigrini, 2006)

- However, this is not always the case, hence the increased likelihood of employee misuse occurring
INTRODUCTION: P-Card Fraud Risk

• Why do P-cards create higher fraud risk than employee credit cards?
  – P-card owners have a higher volume of transactions on a normal basis, while employee credit card usage is typically limited to a periodic event or business trip.
  – For P-card transactions, no pre-approval is required, while employee credit card transactions may require formal manager approval before the credit card provider is reimbursed.
  – Transaction amounts are higher due to type of goods/services purchased, which may increase the rationalization to commit fraud, even in small amounts.
  – Difficulty to detect misuse increases opportunity, which, together with rationalization, constitute two out of three fraud triangle components.
INTRODUCTION: Project Story

- Large multinational consumer goods manufacturer with many different divisions
  - 5600 active p-cards
  - 55,000 p-card transactions per month
  - 15.5 million dollars on average per month
  - a complex scenario!!

- Previous software audit tools were found not effective, and the firm’s procurement card fraud expert, Lisa, is manually reviewing transactional data every month

2 Phases of the project:
- Build an expert system (an “electronic Lisa”)
- Improve anomaly detection rate in p-card data
METHODOLOGY: Data Preprocessing and Exploration

- Monthly training data for the periods of 3/1/13 through 6/1/13

- 55,000 transactions per month with 55 data attributes
  - 2 years of data initially, 2011 & 2012

- Some of the data fields have missing values. For example, vendors choose the level of information that they will provide and some opt out of supplying purchase item description information.

- Even a 95 cent cup of coffee is material!
METHODOLOGY: Data Preprocessing and Exploration

- One of the main challenges of this project is designing an expert system and profiling where key data fields are missing:
PROCESS FLOW UNDERSTANDING

Firm's Procurement Process

1. Cardholder makes purchase with P-Card
2. Purchase is copied from the bank’s credit system and posted to the firm’s ERP system
3. The bank is paid
4. P-Card team downloads all transactions and uploads for management review
5. Manager marks transactions reviewed or requests more information.

Auditor’s Monitoring Process

1. Auditor obtains monthly list of P-Card transactions
2. Auditor manually reviews P-Card transactions for any suspected red flags
3. Auditor marks red flagged transactions and submits them to HR
4. HR will follow up and send feedback to Auditor
KNOWLEDGE ACQUISITION

• The project requires elicitation of an expert’s knowledge

• The unstructured interview is the most popular method of attaining expert knowledge to date (Weiss and Kulikowski, 1984) for the first pass test

• The second pass tests result from further unstructured interviews, structured interviews, limited information tasks, constrained processing tasks, and methods of tough cases
KNOWLEDGE ACQUISITION PROCESS

Data and Preliminary Project Analysis
- Familiarity with data and project requirements
- Analysis of texts and documents; exploratory tests

Unstructured Interviews
- Gains more knowledge from the experts
- Yields enough information for first pass test data base

Preparation of First Pass Test
- Beginnings of project file
- Continual refinement of rules based scripts

Special Tasks
- Yields continued refinements to the system, continuous methodology

Preparation of Second Pass Tests
- Refinement of the test file
- Results of file scripts compared to those of the control, the expert's knowledge

More Interviews
- Unstructured and Structured
- Yields enough information for second pass tests
KNOWLEDGE ACQUISITION

• The first preliminary analysis test was that of Limits

<table>
<thead>
<tr>
<th>ID</th>
<th>Purchase Date</th>
<th>Total dollar amount spent per day</th>
<th>Monthly Credit Limit</th>
<th>Single Transaction Limit</th>
<th>Transactions per day</th>
<th>Difference - Single limit</th>
<th>Difference - Monthly limit</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID2974</td>
<td>11/29/2012</td>
<td>267,087.61</td>
<td>75,000</td>
<td>2,500</td>
<td>141</td>
<td>264,587.6</td>
<td>192,087.6</td>
</tr>
<tr>
<td>ID1929</td>
<td>9/10/2012</td>
<td>136,551.81</td>
<td>60,000</td>
<td>10,000</td>
<td>574</td>
<td>126,551.8</td>
<td>76,551.8</td>
</tr>
<tr>
<td>ID5209</td>
<td>5/17/2012</td>
<td>99,599.03</td>
<td>75,000</td>
<td>2500</td>
<td>3</td>
<td>97,099.0</td>
<td>24,599.0</td>
</tr>
<tr>
<td>ID1967</td>
<td>12/19/2012</td>
<td>96,250.89</td>
<td>75,000</td>
<td>2,500</td>
<td>3</td>
<td>93,750.9</td>
<td>21,250.9</td>
</tr>
<tr>
<td>ID1929</td>
<td>11/12/2012</td>
<td>99,821.08</td>
<td>60,000</td>
<td>10,000</td>
<td>193</td>
<td>89,821.1</td>
<td>39,821.1</td>
</tr>
<tr>
<td>ID3723</td>
<td>5/15/2012</td>
<td>89,625.26</td>
<td>75,000</td>
<td>10,000</td>
<td>421</td>
<td>79,625.3</td>
<td>14,625.3</td>
</tr>
</tbody>
</table>

• ID1929 has 574 transactions per day, which accounts for about 71 transactions per hour (assuming an 8 hour work schedule) and 1.2 transactions per minute
  – There is a need to review such cases to see if such behavior is normal or not.
KNOWLEDGE ACQUISITION

- In addition to the preliminary analysis, we conducted **Exploratory Visual Analysis (EVA)** to further understand the data and build a basis for user purchase behavior.

- The most **heat (color) intensity** among the states goes to Ohio, i.e. it has 64% of the total dollar amounts spent alone.
KNOWLEDGE ACQUISITION

• By aggregating dollar amounts per transaction for both merchants and employees and looking at the overall visual display, we can further understand the data and be able to build better purchase patterns.

• One example here is employee T2472, were despite being third place in terms of total dollar spending ($424,879), has only 8 records in total, compared to employee T0515 in first place with 1,106 records.
KNOWLEDGE ACQUISITION

- As for merchants, we can look at those that stand out in terms of number of records and dollar amounts. For example, Staples is third place in terms of transactions and also has a high dollar amount. (being a store that sells diverse products, one should put in more consideration)

- Another is Expedia, with only 6 records, it is just behind Staples in dollar amount
KNOWLEDGE ACQUISITION

- Textual analysis was then conducted with the data

<table>
<thead>
<tr>
<th>ID</th>
<th>Purchase Date</th>
<th>Original Currency Amount</th>
<th>Extended Item Amount</th>
<th>Merchant Name</th>
<th>Item Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ID1637</td>
<td>2/17/2011</td>
<td>0</td>
<td>50</td>
<td>STAPLES 00101907</td>
<td>$50 APPLES ITUNES</td>
</tr>
<tr>
<td>ID1917</td>
<td>2/22/2012</td>
<td>0</td>
<td>7.59</td>
<td>TARGET 00014472</td>
<td>POKER CHIPS 11.5G GAME ESSEN</td>
</tr>
<tr>
<td>ID0925</td>
<td>3/25/2011</td>
<td>84.95</td>
<td>75</td>
<td>AMAZON MKTPACE PMTS</td>
<td>ITUNES GIFT CARD</td>
</tr>
<tr>
<td>ID4720</td>
<td>7/22/2011</td>
<td>0</td>
<td>10</td>
<td>BOLDEN INSTRUMENT</td>
<td>FUEL CHARGE $10</td>
</tr>
<tr>
<td>ID2503</td>
<td>10/6/2011</td>
<td>31.95</td>
<td>31.95</td>
<td>AMAZON MKTPACE PMTS</td>
<td>PROACTIV SOLUTION ORIGINAL REPAIRIN</td>
</tr>
<tr>
<td>ID2315</td>
<td>10/11/2012</td>
<td>49.69</td>
<td>41.66</td>
<td>STAPLES</td>
<td>STRESS BUSTER MASSAGE FOOT</td>
</tr>
<tr>
<td>ID5477</td>
<td>11/14/2012</td>
<td>24.5</td>
<td>22</td>
<td>AMAZON MKTPACE PMTS</td>
<td>BRIDAL WEDDING JEWELRY HAIR HEADBAN</td>
</tr>
</tbody>
</table>

- One case (highlighted in red) identified immediately as fraudulent by the company

- Other items were determined legitimate after follow up.
KNOWLEDGE ACQUISITION

• Association Rules and Decision Trees:

MCH code 7542 (car washes):

IF (MCH_MCC_Description = “Car Washes”)
    AND (Department_Cost_Center CONTAINS “Facilities Management” OR “Executive” OR “Buildings and Grounds”)
THEN \(\rightarrow\) PASS.

IF (MCH_MCC_Description = “Car Washes”)
    AND (ACC_Master_Accounting_Code = EQUAL “GAS” OR “INCIDGAS”)
THEN \(\rightarrow\) PASS.

IF (MCH_Merchant_Name = “MR CLEAN CAR WASH”)
    AND (Department_Cost_Center CONTAINS “PANELS”)
    AND (FIN_Original_Currency_Amount > $50)
THEN \(\rightarrow\) PASS.
EXPERT TOOL - PASS TESTS

- Our initial run of the expert system produced a total of 1408 exceptions (June - July 2013 test data)

- Another 100+ association rules were added to the tool and after running the SECOND PASS TEST we achieved 95% ACCURACY

- Four cases of personal use/fraud have been confirmed during the first pass test alone.

<table>
<thead>
<tr>
<th></th>
<th>Red Flags Produced</th>
<th>Red Flags Confirmed</th>
<th>Effectiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>First Pass</strong></td>
<td>1408</td>
<td>957</td>
<td>68%</td>
</tr>
<tr>
<td><strong>Second Pass</strong></td>
<td>1300</td>
<td>1235</td>
<td>95%</td>
</tr>
</tbody>
</table>
EXPERT TOOL - PASS TESTS

• The Tool was ran again on October, November, and December data of 2013

• The tool obtained a 98.5% match to the auditor’s flagged transactions

<table>
<thead>
<tr>
<th></th>
<th>Red Flags Produced by Expert Tool</th>
<th>Red Flags Produced by Auditor</th>
<th>Effectiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td>First Pass</td>
<td>2267</td>
<td>2236</td>
<td>98.5%</td>
</tr>
</tbody>
</table>
P-CARD TOOL - ILISA

- The tool was developed in EXCEL due to the firm’s request.

- The tool will have different levels of exceptions, from high false positives to high false negatives
  - The expert will have the ability to decide which level to focus on

Level 1
- Only Textual analytics (TA) on Items/merchants

Level 2
- TA and MCC filtering

Level 3
- TA, MCC filtering, and General Rules

Level 4
- TA, MCC filtering, General rules and Rules specific to firm
P-CARD TOOL - ILISA

- We added a new feature which includes a visual dashboard of exceptions founds.

- The dashboard will provide a quick and efficient way of observing exceptions and noticing any spikes in the visuals.
MOVING FORWARD

• Global Internal Audit is very happy with this project to date...the human and real expert concurred on 172 instances of confirmed fraud

• We will then develop this tool for the international divisions

• Management wants to move from a batch processing to real time data processing

• We will next be looking at their accounts payable

• We also will be working with other firms on expert systems development

• Working on 2nd phase of the project in dealing with transactions with missing information utilizing pattern recognition and employee/merchant profiling
### AI/Second Phase: Missing Values Knowledge Acquisition/Overview of Data

<table>
<thead>
<tr>
<th>Measure for Jan 2013 to April 2014</th>
<th>Total Data Set</th>
<th>Missing Purchase Item Information Data Set</th>
</tr>
</thead>
<tbody>
<tr>
<td># of Transactions</td>
<td>741,710</td>
<td>194,528 (26% of total)</td>
</tr>
<tr>
<td># of Employee IDs</td>
<td>4532 (cards are 5600)</td>
<td>4339</td>
</tr>
<tr>
<td>Total $ Fin Original Currency</td>
<td>$157,115,184</td>
<td>$65,926,544 (42% of total)</td>
</tr>
<tr>
<td>Total # of vendors</td>
<td>101,900</td>
<td>41,258</td>
</tr>
</tbody>
</table>
## Second Phase: Missing Values Knowledge Acquisition/Merchant Types and Names

<table>
<thead>
<tr>
<th>Merchant</th>
<th># of Trans</th>
<th># Emp ID</th>
<th>$ Total</th>
<th># of ??</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walmart</td>
<td>4171</td>
<td>1290</td>
<td>$343,750</td>
<td>All</td>
</tr>
<tr>
<td>Sam’s Club</td>
<td>819</td>
<td>259</td>
<td>$126,612</td>
<td>All</td>
</tr>
<tr>
<td>Amazon</td>
<td>11,690</td>
<td>276</td>
<td>$19,302</td>
<td>Non-credit</td>
</tr>
<tr>
<td>Target</td>
<td>224</td>
<td>115</td>
<td>$37,170</td>
<td>All</td>
</tr>
<tr>
<td>Ulta/Sally B</td>
<td>51</td>
<td>21</td>
<td>$6804</td>
<td>15 (29%)</td>
</tr>
<tr>
<td>Petsmart</td>
<td>174</td>
<td>43</td>
<td>$12,328</td>
<td>25 (14%)</td>
</tr>
<tr>
<td>PetCo</td>
<td>116</td>
<td>9</td>
<td>$60,764</td>
<td>none</td>
</tr>
</tbody>
</table>
SECOND PHASE – Walmart transactions
SECOND PHASE – EmpID #744

54 purchases for $6421, 12 Walmart transactions for $977
SECOND PHASE - EmpID#744: 12 Walmart purchases for $977
SECOND PHASE – MISSING VALUES

- Another informative merchant: PETSMART
- 174 transactions by 43 cards for $12,328

<table>
<thead>
<tr>
<th>FIN.Posting D</th>
<th>FIN.Origin</th>
<th>PUR.Item Description</th>
<th>MCH.Merchant Nam</th>
<th>PUR.Line Item Total Amount</th>
<th>PUR.Product Code</th>
<th>PUR.Purchase Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>3/21/2013</td>
<td>-2.98</td>
<td>PETSMART INC 1469</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3/21/2013</td>
<td>103.22</td>
<td>PETSMART INC 1469</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3/21/2013</td>
<td>29.26</td>
<td>PETSMART INC 1469</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3/15/2013</td>
<td>89.43</td>
<td>PETSMART INC 1469</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3/14/2013</td>
<td>59.46</td>
<td>PETSMART INC 1333</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3/14/2013</td>
<td>48.43</td>
<td>PETSMART INC 1469</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3/12/2013</td>
<td>10.6</td>
<td>PETSMART INC 1469</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3/11/2013</td>
<td>22.34</td>
<td>PETSMART INC 1469</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3/11/2013</td>
<td>52.49</td>
<td>PETSMART INC 1469</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3/5/2013</td>
<td>10.64</td>
<td>PETSMART INC 1237</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3/4/2013</td>
<td>606.87</td>
<td>PETSMART INC 2038</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3/4/2013</td>
<td>559.44</td>
<td>PETSMART INC 248</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/27/2013</td>
<td>53.58</td>
<td>PETSMART INC 1469</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/26/2013</td>
<td>24.1</td>
<td>PETSMART INC 1469</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/25/2013</td>
<td>84.92</td>
<td>PETSMART INC 1333</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/25/2013</td>
<td>63.84</td>
<td>PETSMART INC 1469</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/25/2013</td>
<td>15.29</td>
<td>PETSMART INC 1237</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/21/2013</td>
<td>44.36</td>
<td>PETSMART INC 1469</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/21/2013</td>
<td>171.37</td>
<td>PETSMART INC 1469</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/21/2013</td>
<td>75.19</td>
<td>PETSMART INC 1237</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/15/2013</td>
<td>10.95</td>
<td>PETSMART INC 1469</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/14/2013</td>
<td>1.07</td>
<td>PETSMART INC 1469</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/12/2013</td>
<td>32.68</td>
<td>PETSMART INC 1469</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/7/2013</td>
<td>17.03</td>
<td>PETSMART INC 1469</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2/6/2013</td>
<td>24.82</td>
<td>PETSMART INC 1469</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
SECOND PHASE – MISSING VALUES

Association Rules, first pass:
- If COMPANY = “IAMS” then PASS
- If COMPANY = “NATURA” then PASS
- All others FAIL

Association Rules, second pass:
- If ORG_NAME = “pet” then PASS
- If ORG_NAME = “Product Safety and Regulatory Affairs” then PASS
- All others FAIL

25 TRANSACTIONS ARE FLAGGED AS FAIL
SECOND PHASE – MISSING VALUES

- Heat Map of the 25 suspicious Petsmart transactions:
SECOND PHASE – ID # 3937 @ Petsmart

#1469, 6/21/2013, $21

63 total transactions for $4436
SECOND PHASE – ID # 4360 @ Petsmart

43 total transactions for $6542

#1469, 6/12/13, $9
#1469, 10/2/13, $29
#1470, 1/17/13, $26
#1470, 5/16/2013, $19

4 Petsmart transactions totaling $83
SECOND PHASE – ID#1878 @ Petsmart

Department: IAMS (legitimate)
KEY TAKEAWAYS:

- P-Card use has a high inherent fraud risk

- The “real expert” is not an absolute expert

- The tool will be needing constant updates

- Behavior profiling and clustering work is just starting as a second phase and will be added to the tool to improve its expertise. Hidden Markov Models and a hybrid Belief Networks/Dempster - Shafer approach will be applied in an AI approach

- iLisa will be a better expert than the human one!!!