Card Fraud

Howard Mizes
December 3, 2013
Outline of Presentation

• Credit Card Transactions
  – Credit and debit card processing.
  – Prevalence of Fraud.
  – Xerox’s role in debit cards.

• Academic Efforts
  – Machine Learning Models

• Xerox Efforts
  – Xerox dataset
  – Factors influencing fraud
  – Fraud visualization
THE ANATOMY OF A TRANSACTION

AUTHORIZATION

1. Cardholder provides MasterCard account to merchant
2. Merchant’s bank asks MasterCard to determine cardholder’s bank
3. MasterCard authorization system checks card security features and sends to cardholder’s bank for approval
4. Cardholder’s bank approves purchase
5. MasterCard sends approval to merchant’s bank
6. Merchant’s bank sends approval to merchant
7. Cardholder completes purchase and receives receipt

CLEARING

1. Merchant’s bank sends purchase information to MasterCard network
2. MasterCard clearing system validates information and sends purchase information to cardholder’s bank, which prepares data for cardholder’s statement
3. MasterCard clearing system provides comprehensive reconciliation to both the merchant’s bank and to the cardholder’s bank

SETTLEMENT

1. Cardholder’s bank sends payment to the MasterCard counter-party to settlement
2. The MasterCard counter-party to settlement sends payment to merchant’s bank
3. Merchant’s bank pays merchant for cardholder’s purchase
4. Cardholder’s bank bills cardholder
Fraud Techniques

• Stolen Cards: Extra identifying information is also found in wallet.

• Skimming: Insert an extra device at a point of sell to get card information from magnetic strip, keypunch, etc.

• “Card not Present”: Fraudsters create a counterfeit card with cardholder information or gather information for online transactions.
  – Phishing: Create a legitimate looking e-mail or text with the intent of tricking individuals to disclose PII.
  – Pharming: Redirects individuals to malicious websites that look valid.

---

This message was sent with High importance.

From: Visa.com [mailto:service@usa.visa.com] Sent: Friday, Feb. 10, 2012 1:40 AM
To: undisclosed-recipients
Subject: Your Credit Card has been Suspended

Dear Valued Customer,

Your Credit Card has been Suspended, as an error was detected in your Credit Card information. The reason for the error is not certain, but for security reasons, we have suspended your Credit Card temporarily.

We need you to update your information for further use of this Credit Card.

To Lift this Suspension:

Click Here

and follow the Steps to re-activate your Credit Card.

NOTE: If this is not resolved within 72 hours, we will be forced to suspend your Credit Card Permanently as it may be used fraudulently. The purpose of this verification is to ensure that your Credit Card account has not been fraudulently used.

Thanks,

Customer Support Service.

Copyright © 1999-2012 Verified by Visa® All Rights Reserved.
Prevalence of Fraud

• 32 percent of consumers reported card fraud in the past five years, according to a 2010 survey ACI Worldwide.

• $5.55 billion credit card fraud worldwide in 2012 (U.S. Department of Justice), $3.56 billion in US (2011 Nilson Report)

• $2 trillion used for credit card and $1.5 trillion for debit card purchases in US in 2012 (Discover, AmEx,, Visa, Mastercard).

• Financial institutions incurred $955 million in losses due to debit card fraud in 2010 (ABA Deposit Account Fraud Survey, 2011)
Incentive for investigating fraud

- Fraud is always going to occur, which is a loss to the card issuer.
- Fraud investigation decreases the amount of fraud, but costs money which is overhead to the card issuer.
- Improved machine learning models decrease the amount of fraud.

\[
\text{Transaction Profit Contribution (TCP)} = \text{[Transaction Amount]} \times \left( \text{[Net Interchange Rate]} + \text{[Net FX Fees]} \right) - \text{Funding Costs} - \text{[Transaction Amount]} \times \left( \text{[Probability of Loss]} \times \left[ 1 - \text{Fraud Recovery Rate} \right] \right)
\]

\[
\text{Account Profit Contribution (APC)} = \left[ \text{Transaction Profit Contribution} \right] + \left[ \text{Current Balance} \right] \times \left( \text{[Net Interest Margin]} - \text{Expected Credit Loss} \right) + \left[ \text{Expected Annualized Spend} \right] \times \left( \text{[Net Interchange Rate]} + \left[ \text{Other Cardholder Fees} \right] - \left[ \text{Servicing} \right] \right)
\] \times \text{Attrition Risk}
\]
Xerox State Government Solutions Group

- Government partner provides ~30 electronic payment card services programs for twenty states and the U.S. federal government
  - Customer care
  - Eligibility and case management
  - Benefits disbursement
  - High-volume transaction processing
  - Communication and Marketing

Electronic Payment Cards
- Social Security benefits
- Child support funds
- Temporary Assistance for Needy Families (TANF)
- Unemployment Insurance (UI)
- State Supplemental Insurance (SSI)
- Foster care payments
- Adoption subsidies
- Child care provider payments
- Payroll to government workers
- Retirement benefits
- Veterans benefits
- Disability insurance
- Paid family leave benefits
- Any other constituent benefits or recurring payments
Welcome to EPPICard.com

Select Your State

Frequently Asked Questions

Online Security Tips

Protect Yourself

Privacy Policy

Site Map

Save with Visa

MasterCard

MARKETPLACE®

Cyveillance

PROTECTED

- the safe and secure way to access your payments!
The Florida Debit MasterCard Card Program is a new way to receive your support payments. You can get cash at banks and Automated Teller Machines (ATMs) and make purchases at most stores in the USA and around the world. It is accepted everywhere MasterCard® debit cards are accepted. Use your card at grocery stores, office and school suppliers, drug stores, restaurants, doctors’ offices and discount stores.

Benefits of the Florida Debit MasterCard

- Safe, fast and easy to use
- Website and customer service toll-free number to answer questions and check your balance

Make Purchases and Get Cash Back

Simply present your card when paying. The money is automatically deducted from your account. You may ask for “cash-back” with your purchase at many merchant locations.

Get Cash at a Bank Teller Window or an ATM

- You are allowed one (1) free ATM cash withdrawal each month only at PNC Bank ATMs. The same $0.95 fee you have always paid will be charged for each PNC Bank ATM cash withdrawal after that and for all cash withdrawals at ATMs other than PNC Bank ATMs.
- Free ATM cash withdrawal must be used by midnight on the last calendar day each month.

Customer Service

- Review your balance and transaction history online by using the website at: www.EPPICard.com
- Call the toll free number for questions, 1-877-567-1768
- Report a lost or stolen card
- Ask for information about your account or dispute a transaction posted to your account
- Ask for a printed monthly statement

How do I get a Florida Debit MasterCard?

- Get a payment options form from your local child support office, or from our website www.myflorida.com/dor/childsupport or by calling 1-877-769-0251.
- Your card will be sent to you by mail
- Activate your card right away by selecting a PIN as instructed on the card carrier
- You can use your card only after support payments are made to your account
Bank Fraud Algorithms

- Little quantitative information of the algorithms are shared publicly.
- Lots of P/R with respect to fraud detection.
  - Motivation is for customers to feel confident of security.
  - Visa claims to detect “high speed fraud,” in which hackers break into a payment processor’s network and run multiple transactions within minutes or even seconds, charging very small amounts in each
  - Visa claims to monitor individual spending patterns, update them, and compare transactions against the current profile.
- Found information in a blog allegedly from a fraud prevention specialists that confirms the operation is similar to Xerox.
  - Algorithms flag suspicious charges.
  - Followed up manually with agents experience providing guidance.
<table>
<thead>
<tr>
<th>Title</th>
<th>First Author</th>
<th>Journal</th>
<th>Techniques</th>
<th>Dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td>Toward Scalable Learning With Non-uniform Class and Cost Distributions: A Case Study in Credit Card Fraud Detection</td>
<td>Philip K. Chan</td>
<td>KDD 1998</td>
<td>CART, C4.5, RIPPER, BAYES</td>
<td>500,000 transactions, 30 features, 20% fraudulent, from Chase Manhattan</td>
</tr>
<tr>
<td>Distributed Data Mining in Credit Card Fraud Detection</td>
<td>Philip K. Chan</td>
<td>IEEE Intelligent Systems 1999</td>
<td></td>
<td>Another 500,000 from First Union Bank</td>
</tr>
<tr>
<td>Neural Data Mining for Credit Card Fraud Detection</td>
<td>R. Brause</td>
<td>Proceedings of the ... 2002</td>
<td>Neural network, rules</td>
<td>500,000 transactions, 1% fraud, 38 features from Chase Manhattan</td>
</tr>
<tr>
<td>Credit Card Fraud Detection Using Bayesian and Neural Networks</td>
<td>Sam Maes</td>
<td>Proceedings of the ... 2002</td>
<td>Bayesian, Neural Networks</td>
<td>10 features, Europay International</td>
</tr>
<tr>
<td>Parallel Granular Neural Networks for Fast Credit Card Fraud Detection</td>
<td>Mubeena Syeda</td>
<td>FUZZ-IEEE'02</td>
<td>Neural Networks, parallel Processing</td>
<td>Source and details not disclosed</td>
</tr>
<tr>
<td>Credit Card Fraud Detection Using Hidden Markov Model</td>
<td>Avhinav Srivastava</td>
<td>IEEE Dep &amp; Sec Comp. 2002</td>
<td>HMM</td>
<td>Completely simulated and simplified data.</td>
</tr>
<tr>
<td>A Web Services-Based Collaborative Scheme for Credit Card Fraud Detection</td>
<td>Chuang-Cheng Chiu</td>
<td>EEE 2004</td>
<td>Frequent pattern Mining</td>
<td>3 features, Taiwan bank</td>
</tr>
<tr>
<td>Application of Classification Models on Credit Card Fraud Detection</td>
<td>Aihua Shen</td>
<td>SSSM 2007</td>
<td>Decision Tree, Neural Networks, logistic regression</td>
<td>40 fields, 0.07% fraud, sampled entire database</td>
</tr>
<tr>
<td>Detecting Credit Card Fraud by Decision Trees and Support Vector Machines</td>
<td>Y. Sahin</td>
<td>Proc Int. MultiConf of Eng &amp; Comp Sci 2011</td>
<td>Decision Tree, SVM</td>
<td>978 fraud, 2,000,000 rows</td>
</tr>
<tr>
<td>Credit card fraud detection: A fusion approach using Dempster-Shafer theory and Bayesian learning</td>
<td>Suvasini Panigrahi</td>
<td>Information Fusion 2009</td>
<td>Bayesian, Dempster-Shafer</td>
<td>Synthetic data</td>
</tr>
<tr>
<td>Credit Card Fraud Detection with Artificial Immune System</td>
<td>Manoel Fernando Alonso Gadi</td>
<td>ICARIS 2008</td>
<td>Artificial Immune System</td>
<td>41647 rows, 17 features, 3 months from Brazilian bank.</td>
</tr>
</tbody>
</table>
Labeling Fraud

• In order to run supervised machine learning algorithms for fraud detection, fraudulent transactions must be labeled.

• Academic Researchers are provided with a fraud labeled dataset by the banks providing the data.
Challenges of Fraud Detection Studies

• Limited amount of data associated with the fraud event (amount, time, merchant, card)

• Millions of places and sites make it hard to match a pattern.

• Patterns of fraudsters can match patterns of legitimate behavior.

• Fraud behavior changes constantly.

• Exchange of ideas is limited due to security and privacy.

• Datasets are not made publicly available.
  – No benchmarking.
  – Details about features are not disclosed.

• Datasets are highly skewed.
  – 10,000 valid for each fraud.
## Common Fields for Card Transactions

### Table 1: Definition of variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Type</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>POS_no</td>
<td>char</td>
<td>POS terminal number</td>
</tr>
<tr>
<td>account_no</td>
<td>char</td>
<td>Account number</td>
</tr>
<tr>
<td>trans_date</td>
<td>date</td>
<td>Transaction date</td>
</tr>
<tr>
<td>trans_time</td>
<td>time</td>
<td>Transaction time</td>
</tr>
<tr>
<td>trans_amt</td>
<td>float</td>
<td>The amount of a credit card transaction</td>
</tr>
<tr>
<td>MCC</td>
<td>int</td>
<td>Merchant Category Code</td>
</tr>
<tr>
<td>card_type</td>
<td>int</td>
<td>Card type</td>
</tr>
<tr>
<td>exp_date</td>
<td>date</td>
<td>Expiration Date</td>
</tr>
<tr>
<td>C1</td>
<td>float</td>
<td>The total amount of transactions of this card in the same day</td>
</tr>
<tr>
<td>C2</td>
<td>int</td>
<td>The number of transactions of this card in the same day</td>
</tr>
<tr>
<td>C3</td>
<td>int</td>
<td>The failure number of transactions of this card in the same day</td>
</tr>
<tr>
<td>C4</td>
<td>float</td>
<td>The average transaction amount of this card in the same day</td>
</tr>
<tr>
<td>C5</td>
<td>float</td>
<td>The total amount of transactions of this card within five days</td>
</tr>
<tr>
<td>C6</td>
<td>int</td>
<td>The number of transactions of this card within five days</td>
</tr>
<tr>
<td>C7</td>
<td>int</td>
<td>The failure number of transactions of this card within five days</td>
</tr>
<tr>
<td>C8</td>
<td>float</td>
<td>The average amount of transactions of this card within five days</td>
</tr>
</tbody>
</table>

- Features associated with an individual transaction
- Features associated with the individual (others can include age, address)
- Features derived from the historical use of the card

From Aihua Shen et. al. Conference on Service System and Service Management 2007
Challenges of Skewed Distributions

• Fraud is only a small percentage of card transactions.

• Obtaining a high recall (TP/(TP+FN) = catching most fraud) might entail an extremely low precision (TP/(TP+FP)).

• Detecting half the frauds gives 6,000,000 false positives!

Hidden Markov Model Approach

• Learn spending pattern of individual cardholders.

• Simulated mix of standard and fraudulent spending patterns.
  • 3 types of card holds (a,b,c) where triplet is % of spending in low, medium, and high category.

• Individual rules appear to be in use by financial institutions, but details beyond speculations by outsiders are not disclosed.
Parallel and Distributed Processing

• Mitigates the need for sampling to run machine learning algorithms to learn risk.
  - Today we can run analysis on upwards of 50 petabytes of data to more accurately calculate risk (Abhishek Mehta, Bank of America)

• At a 2009 Cloudera conference, Visa presented its efforts in using Hadoop to speed up its data processing.
  - Testing on research sample of 73 billion synthetic transactions (36 TB).
  - Output is risk score of each transaction.
  - Processed 1 month in 13 minutes.

• In 2013 Visa announces its using Hadoop to analyze
  - 100% rather than 2% of transactions.
  - 500 features rather than 40 features.
  - Running multiple models.