

# Card Fraud

Howard Mizes  
December 3, 2013



# Outline of Presentation

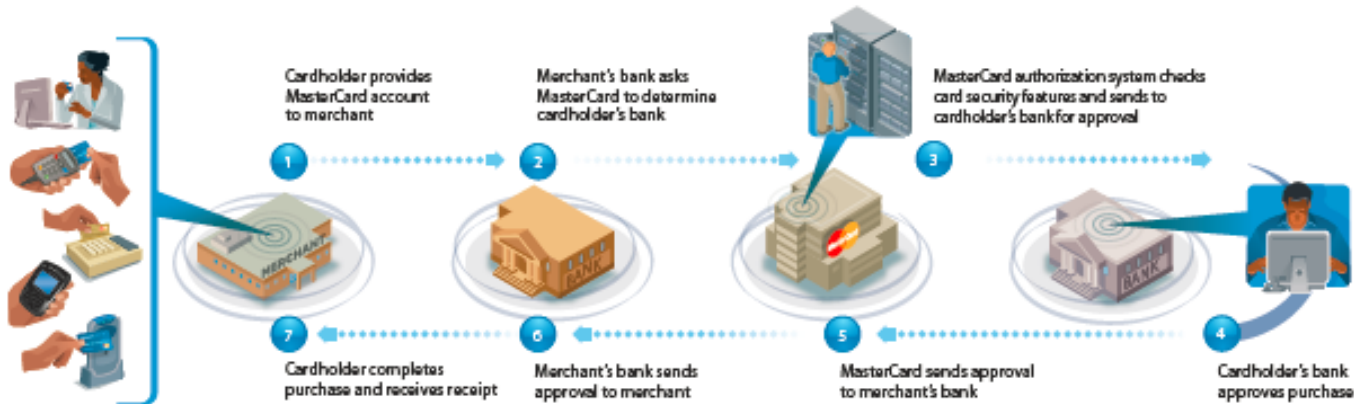
- Credit Card Transactions
  - Credit and debit card processing.
  - Prevalence of Fraud.
  - Xerox's role in debit cards.
- Academic Efforts
  - Machine Learning Models
- Xerox Efforts
  - Xerox dataset
  - Factors influencing fraud
  - Fraud visualization

# THE ANATOMY OF A TRANSACTION

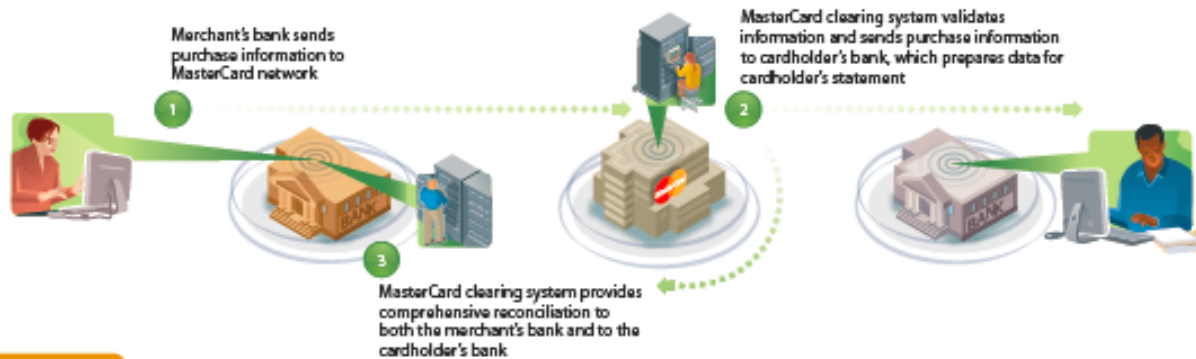


MasterCard  
Worldwide

## AUTHORIZATION TIME OF PURCHASE



## CLEARING USUALLY WITHIN ONE DAY



## SETTLEMENT USUALLY WITHIN TWO DAYS



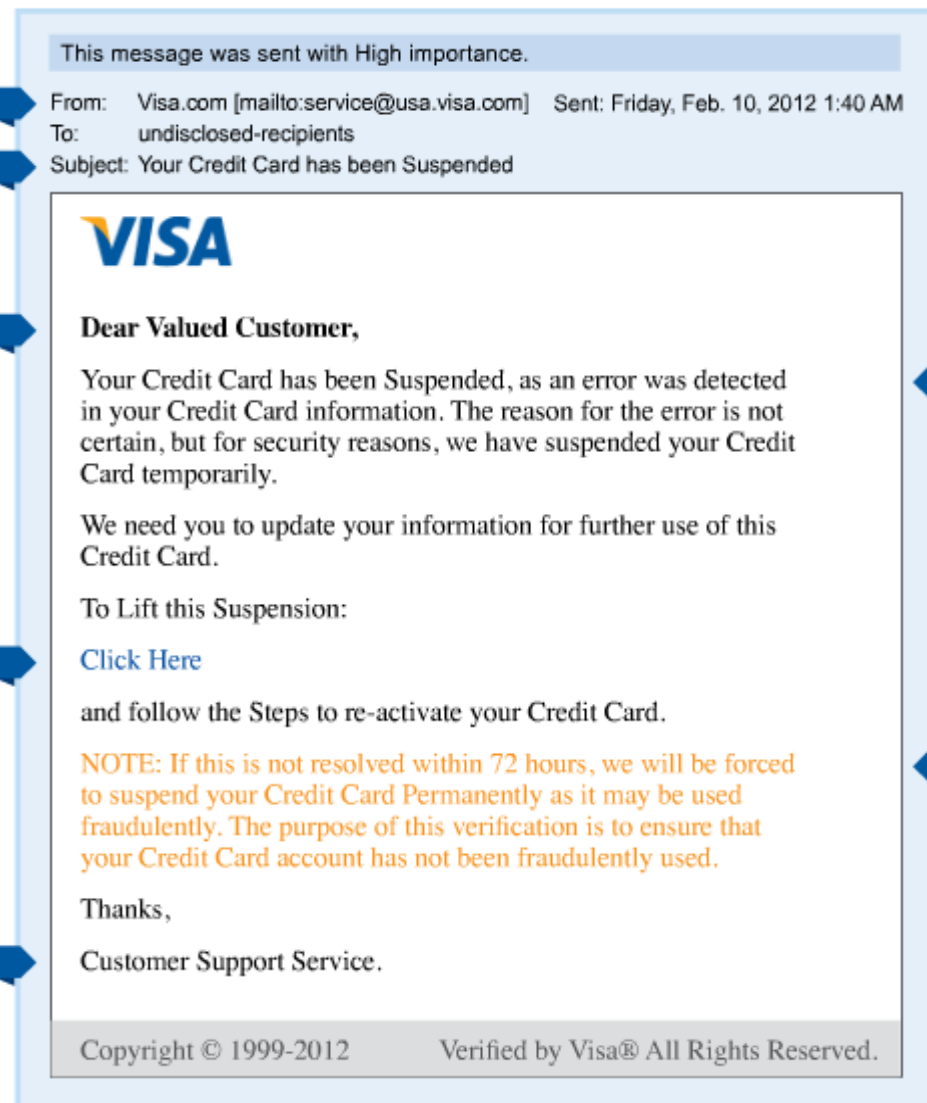
### The MasterCard Network

- Connects 210 countries and more than 25,000 financial institutions
- Processes 5.4 million transactions per hour, and processes its portion of each transaction in an average of 129 milliseconds
- Saved customers approximately 1,173 cumulative years of processing time since 1997
- Processed 16.1 billion MasterCard-branded authorizations in 2006
- Maintains 99.999% global availability around the clock



# 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.



## 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.

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

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

# 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



EPPICard™

- the safe and secure way to access your payments!

Welcome to EPPICard.com

Select Your State

Frequently Asked Questions

Online Security Tips

Protect Yourself

Privacy Policy

Site Map

Save with Visa  
**VISA**



MARKETPLACE™

Cyveillance®  
PROTECTED







# Child Support Enforcement Florida Debit MasterCard Program For Support 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 web site at: [www.EPPICard.com](http://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](http://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 publically.
- 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.
- <http://ask.metafilter.com/155037/What-tipped-off-the-credit-card-company-for-fraud>

# Academic Studies (<<100 over past 20 years)

Title	First Author	Journal	Techniques	Dataset
Toward Scalable Learning With Non-uniform Class and Cost Distributions: A Case Study in Credit Card Fraud Detection	Philip K. Chan	KDD 1998	CART, C4.5, RIPPER, BAYES	500,000 transactions, 30 features, 20% fraudulent, from Chase Manhattan
Distributed Data Minging in Credit Card Fraud Detection	Philip K. Chan	IEEE Intellegent Systems 1999		Another 500,000 from First Union Bank
Neural Data Minging for Credit Card Fraud Detection	R. Brause		Neural network, rules	500,000 transactions, 1% fraud, 38 features from Chase Manhattan
Credit Card Fraud Detection Using Bayesian and Neural Netorks	Sam Maes	Proceedings of the ... 2002	Bayesian, Neural Networks	10 features, Europay International
Parallel Granular Neural Networks for Fast Credit Card Fraud Detection	Mubeena Syeda	FUZZ-IEEE'02	Neural Networks, parallel Processing	Source and details not disclosed
Credit Card Fraud Detection Using Hidden Markov Model	Avhinav Srivastava	IEEE Dep & Sec Comp. 2002	HMM	Completely simulated and simplified data.
A Web Services-Based Collaborative Scheme for Credit Card Fraud Detection	Chuang-Cheng Chiu	EEE 2004	Frequent pattern Mining	3 features, Taiwan bank
Application of Classification Models on Credit Card Fraud Detection	Aihua Shen	SSSM 2007	Decision Tree, Neural Networks, logistic regression	40 fields, 0.07% fraud, sampled entire database
Detecting Credit Card Fraud by Decision Trees and Support Vector Machines	Y. Sahin	Proc Int. MultiConf of Eng & Comp Sci 2011	Decision Tree, SVM	978 fraud, 2,000,000 rows
Credit card fraud detection: A fusion approach using Dempster-Shafer theory and Bayesian learning	Suvasini Panigrahi	Information Fusion 2009	Bayesian, Dempster-Shafer	Synthetic data
Credit Card Fraud Detection with Artificial Immune System	Manoel Fernando Alonsa Gadi	ICARIS 2008	Artificial Immune System	41647 rows, 17 features, 3 months from Brazilian bank.

# 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

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

Features associated with an individual transaction

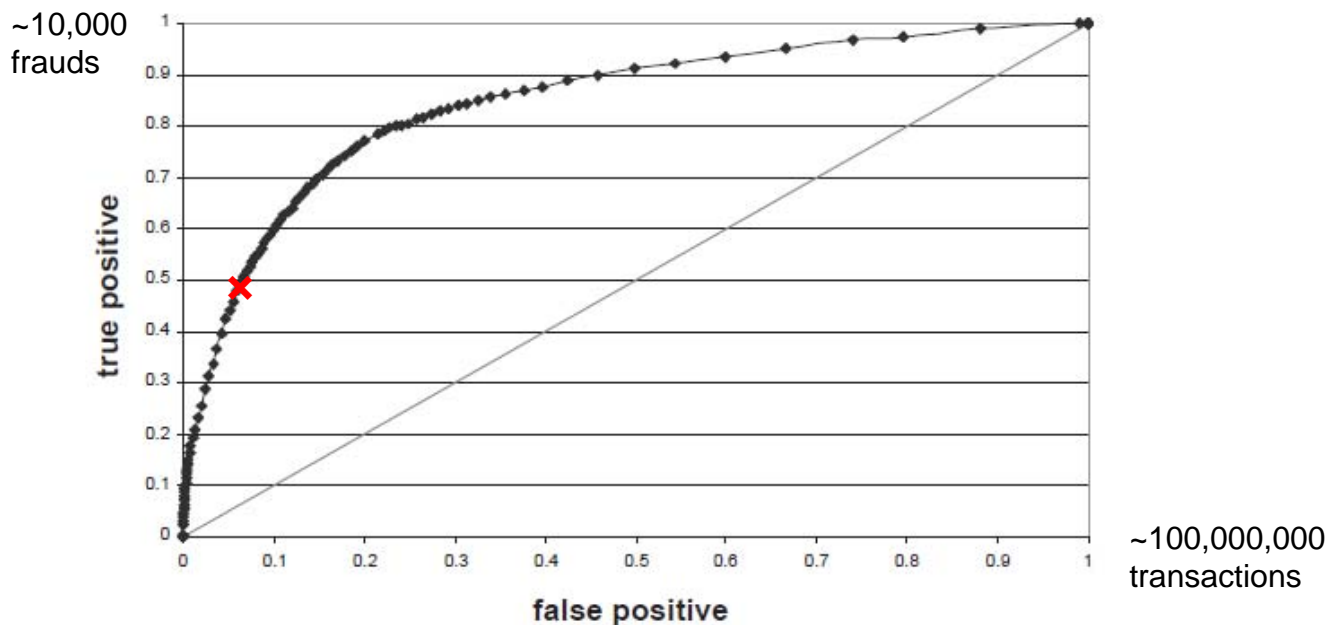
Features associated with the individual (others can include age, address)

Features derived from the historical use of the card



# 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 ( $TP/(TP+FP)$ ).



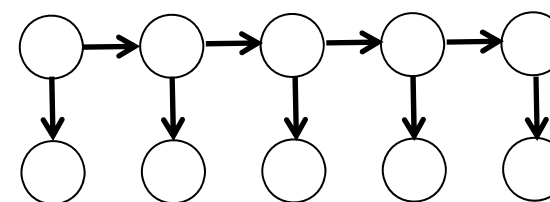
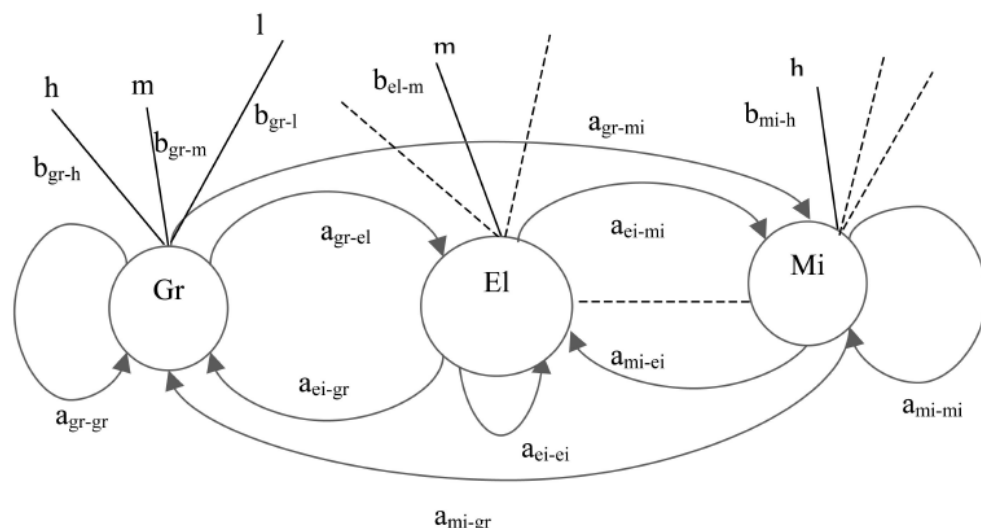
- Detecting half the frauds gives 6,000,000 false positives!

From Sam Maes, et al. Proceedings of the 1<sup>st</sup> International NAISO Congress on Neuro Fuzzy Technologies, 2002

# Hidden Markov Model Approach

- Learn spending pattern of individual cardholders.

Gr: Groceries  
 El: Electronics  
 Mi: Miscellaneous



$$\alpha_1 = P(O_1, O_2, O_3, \dots, O_R | \lambda).$$

$$\alpha_2 = P(O_2, O_3, O_4, \dots, O_{R+1} | \lambda),$$

$$\Delta\alpha = \alpha_1 - \alpha_2.$$

- 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.