Use of DataEngine in the proactive detection of card fraud

Jens Mende
Department of Information Systems
University of the Witwatersrand
Johannesburg, South Africa
jensm@isis.witt.ac.sa

Abstract: Credit card fraud costs the banking system and the world economy billions of dollars annually. Although many different means have been tried to combat the problem, the incidence of fraud and the sophistication of criminals increases each year as classical anti-fraud measures are systematically overcome. This paper looks at various technologies useful in the identification and investigation of credit card fraud. The authors' approach is one that emphasises that real-time identification - using the infrastructure currently at the disposal of the financial sector - is not only possible but vital to ensuring success. The document further describes techniques that may be used in the investigation of credit card fraud crimes.The threat of credit card fraud demands a multidisciplinary approach in order to understand, track, investigate, analyse and ultimately to reduce the phenomenon to a level where it is no longer a threat. It is suggested that a number of seemingly disparate fields might have important inputs to understanding and dealing with the problem effectively. The problem of credit card fraud has arisen from a number of different and complex causes. Many different sciences and disciplines can therefore profitably be used to address the problem.

HOW BIG IS THE PROBLEM?

There are varying estimates for the amount of credit card fraud worldwide. However, analysis of figures provided is often meaningless as the real amount of fraud is often understated by Banks which may be sensitive to adverse publicity. This is especially true if the frauds have been perpetrated or facilitated by bank employees.

ANATOMY OF A CREDIT CARD FRAUD

Generally, an unauthorised person acquires a credit card and / or the credit cards' number and expiry date. In todays fast moving economy, the criminal does not even need to have the actual card in order to commit the fraud.

To assist in his fraud, other information may also be obtained by the criminal appertaining to the card holder via accomplices within the bank. This information could include the card holders' Identity Number, Address and other information generally only known to the card holder. This information is often useful should this information be required by a merchant to validate a purchase.

The unauthorised individual than has a 'window period' in which the card can be used, relatively safely, before its fraudulent use can be reported to the issuing banks and the merchants.

The professional criminal makes a point of keeping up to date with the Banks' fraud detection systems and rules and works within an understanding of these to avoid early detection. For example, if a criminal knows that the floor limit (the level below purchases are allowed without a request being directed to a host authorisation system) is R200.00, he/she will ensure that purchases fall under this level.

BANKING SYSTEMS

Issuing Banks generally have systems which do some form of checking on transactions using simple if-then-else rules. These rules generally have a bias towards the needs of accounting information systems. Whilst management at such institutions might intuitively feel that certain rules will lead to proactive detection of fraudulently obtained cards, such rules are normally impossible to express with any empirical validity. As a result, the Banks often face the dilemma of falsely identifying a card as being fraudulently used when this is, in fact, not the case. This could potentially disrupt the relationship with the legitimate client.

PROBLEM OVERVIEW

It is the contention of the authors that the manner in which the criminal uses a fraudulently obtained credit card is qualitatively and quantitatively different from the way the authorised user has used the card previously. We refer to the legitimate users utilisation patterns as the ‘fingerprint’.

Factors such as:

- Frequency – how often the card is used
- Nature of Merchant – the kind of shop / outlet where goods or services are purchased.
- Location of Merchant – physical location of merchant
- Compliance – the extent to which the user observes his / her credit limits and the agreement with the Issuing Bank

All combine to form a purchasing ‘fingerprint’ which can uniquely identify the valid user. A break from set patterns can be used to identify whether someone other than the user is making use of the card.

It is extremely unlikely that a criminal will use the card in the same way as the cards’ legitimate owner - the criminal breaks pre-established patterns of card utilisation. It is vital that Banking systems be able to recognise such violations of pattern as soon as possible. At a more mundane level, such changes in utilisation pattern may also indicate marketing opportunities and changes in the legitimate users position in the consumer life cycle or relationship with the bank.

DATAENGINE IN CARD FRAUD DETECTION

The inspiration of our card fraud detection system came from the DataEngine Banking tutorial which differentiates and clusters a bank’s customers on the basis of varying criteria such as age, income, assets, credit etc.

Essentially, the criminal is also a user the banks facilities, albeit an undesireable user. Using fuzzy C means clustering, the criminal’s behaviour can be clearly differentiated from that of normal users.

In the credit card fraud detection system, we supplied factors such as:

- Merchant Details
  - Type of merchant (merchants of certain types appear to be more likely than normal to be engaged in fraudulent transactions
  - Merchant Number (The merchant’s unique ID Number)
- Transaction Details
  - Amount of transaction expressed as a percentage of the applicable floor limit
  - Number of transactions
- Card holder details
  - Age
  - Postal code
  - Sex

Using a fuzzy C Means Clustering algorithm, we were able to arrive at various clusters describing particular types of fraud. The fuzzy C means clustering process allowed us to:

> *Define the characteristics which, in combination, characterise the various types of credit card fraud.*
> *Differentiate fraudulent use from normal use.*

Having characterised these forms of fraud, we were able to provide a nomenclature for these fraud types which were unique to the client bank.

DERIVED ANALYSES

Resulting from the DataEngine analysis, the following on-going realtime monitoring of card utilisation was recommended:

1) Merchant analysis – it was long suspected that certain merchants work in collusion with criminals. A relatively simple analysis comparing the rate of fraud experienced at a merchant vis-à-vis average rates of fraud for that merchant's sector could allow the bank to identify suspect merchants, thereby cutting the criminals' access to goods, services and cash.

2) Consolidating fraudulent transaction information – Issuing Banks can obtain a better picture of fraud if information appertaining to all frauds is consolidated. This gives an overview of which merchants may be working alongside the criminals. It also provides details of the criminals' modus operandi.

3) Susceptibility analyses – it is shown that cards are particularly susceptible to being used fraudulently under certain conditions. The fraud susceptibility analysis allows more frequent review of transactions when the probability of having fraudulent transactions is at its greatest.

SUMMARY

New technologies such as DataEngine are allowing us greater insight into phenomena such as fraudulent card use leading to its proactive detection. It is important that those charged with the investigation and amelioration of the various types of fraud become acquainted with the new techniques which allow for proactive detection and identification. Additional techniques from DataEngine such as fuzzy logic and Neural Networks will also allow for early fraud detection in areas such as banking, consumer credit, telecommunications where early or real-time identification and reaction to problems is vital to the success of the organisation.