Navajyoti, International Journal of Multi-Disciplinary Research
Volume 2, Issue 2, February 2018

The Credit Card Transaction Fraud Detection Using Web Mining Technique

B N Sharmila [1]
The Bangalore Social and Educational Institution of Management Studies, Bangalore

ABSTRACT

Credit card hackers looking for new ways to drain money from consumers' bank accounts and evade increased bank security measures have discovered a clever side door—the Starbucks mobile payment app and gift cards. Criminals are hijacking consumers' coffee accounts, draining the stored value of their cards, and then using Starbucks' auto-reload function to hack consumers' associated debit and credit cards.

Credit card hackers are targeting third-party firms that create alternative payment systems and attacking them, finding they are often easier to hack than financial institutions. Fraud is moving away from banks into big e-commerce companies, Criminals are learning how to turn rewards programs, points and prepaid cards into cash.Here, through this paperwe propose method to dynamically identify characteristics pattern of customer then the incoming transactions are compared against the user profile to indicate the anomalies, based on appropriate message has been given. A FP tree based pattern matching algorithm is used to evaluate how unusual the new transactions are.

INTRODUCTION

Credit card fraud is a wide-ranging term for theft and fraud committed using or involving a payment card, such as a credit card ordebit card, as a fraudulent source of funds in a transaction. The purpose may be to obtain goods without paying, or to obtain unauthorized funds from an account. Credit card fraud is also an adjunct toidentity theft. According to the United StatesFederal Trade Commission, while identity theft had been holding steady for the last few years, it saw a 21 percent increase in 2008. However, credit card fraud, that crime which most people associate with ID theft, decreased as a percentage of all ID theft complaints for the sixth year in a row. Although incidence of credit card fraud is limited to about 0.1% of all card transactions, this has resulted in huge financial losses as the fraudulent transactions have been large value transactions. In 1999, out of 12 billion transactions made annually, approximately 10 million—or one out of every 1200 transactions—turned out to be fraudulent. Also, 0.04% (4 out of every 10,000) of all monthly active

B_n_sharmila123@yahoo.co.in

accounts were fraudulent. Even with tremendous volume and value increase in credit card transactions since then, these proportions have stayed the same or havedecreased due to sophisticated fraud detection and prevention systems. Today's fraud detection systems are designed to prevent one twelfth of one percent of all transactions processed which still translates into billions of dollars in losses.

PROBLEMS IN CREDIT CARD FRAUD DETECTION:

One of the biggest problems associated with credit card fraud detection is the lack of the both literatureproviding experimented results and of real world data for researchers to perform experiments on. This is because fraud detection is often associated with sensitive financial data that is kept confidential for reasons ofcustomer accuracy. Some of the properties a fraud detection system should have in order to perform some good results.

- The system should be able to handle skewed distributions
- The ability to handle noise.
- Overlapping data
- The systems should be able to adapt themselves to new kinds of fraud.
- There is a need for good matrix to evaluate the classified system.
- The systems should take into account the cost of the fraudulent behavior detected and cost associated with stopping it.

TYPES OF FRAUD:

Various types of frauds in this paper include credit card frauds, telecommunication frauds, and computer intrusions, Bankruptcy fraud, Theft fraud/counterfeit fraud,
Application fraud, Behavioral fraud

Credit Card Fraud: Credit card fraud has been divided into two types: Offline fraud and on line fraud. Offline fraud is committed by using a stolen physical card at call center or any other place.

On-line fraud: Is committed via internet, phone, shopping, web, or in absence of card holder.

Telecommunication Fraud: The use of telecommunication services to commit other forms of fraud. Consumers, businesses and communication service provider are the victims. Computer Intrusion: Intrusion Is Defined As The Act Of Entering Without Warrant Or Invitation; That Means "Potential Possibility Of Unauthorized Attempt To Access Information, Manipulate Information Purposefully. Intruders may Be from Any Environment, An Outsider (Or Hacker) And An Insider Who Knows The Layout Of the system

BASIC CONCEPTS OF A FRAUD DETECTING SYSTEM.

Why the owner of a stolen credit card is not charged for fraudulent transactions? One reason is because fraud can quickly be detected with a computer by tracking usage patterns and history. IBM has solutions to help. Here are 5 things to know about IBM fraud detection solutions.

1. Companies get ripped off by billions of dollars each year due to fraud.

In the U.S., 32 percent of consumers reported card fraud in the past five years. Some of the schemes use very complex technology, while others simply rely on the trust of the purchaser. Both consumers and banks are very interested in minimizing these losses.

2. The top 25 world banks run their businesses on mainframes.

In fact, 71% of Fortune 500 banks use mainframes. These facts are seldom publicized, but should be no surprise. IBM System z mainframes have experienced nearly 50 years of improved hardware, software, and procedures, making them reliable and quite foolproof. You don't often (if ever) hear of someone hacking a mainframe.

3. The ideal solution avoids making fraudulent payments without slowing down legitimate payments.

Such a solution requires the adoption of a comprehensive fraud business architecture that applies advanced predictive analytics to reduce fraud, waste, and abuse, by using the following techniques:

- Identify vulnerabilities
- Detect transactions
- Evaluate workloads
- Conduct remediation
- Process appeals

5. The brains behind predicting scoring ratings is a user-written SPSS model.

For a typical transactional fraud detection business case, assume that a customer is making a credit card payment. At the time of payment, the bank analyzes the payment pattern on that particular credit card to detect the possibility of fraud. This analysis involves the history, frequency, and dollar amounts of previous transactions for that credit card from its database records. Depending on the scoring analysis, the bank authorizes the transaction, keeps it on hold, or declines it, all in real time.

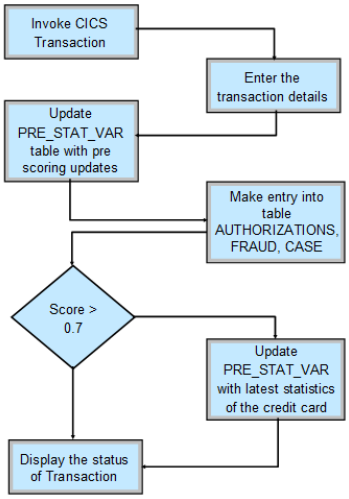

Figure 1. *The SPSS model is the brains of the solution.*

The above figure is a flow chart that explains about the flow involved in transaction process of the card. The transaction details like card number, amount, and place will be recorded for displaying transaction status.

Data analysis techniques for fraud detection

Here we are explaining about the Fraud detection in two different ways.

A. Fraud Detection System

Fraud that involves cell phones, insurance claims, tax return claims, credit card transactions etc. represent significant problems for governments and businesses, but yet detecting and preventing fraud is not a simple task. Fraud is an adaptive crime, so it needs special methods of intelligent data analysis to detect and prevent it. These methods exists in the areas of Knowledge Discovery in Databases (KDD), Data Mining, Machine Learning and Statistics. They offer applicable and successful solutions in different areas of fraud crimes.

Techniques used for fraud detection fall into two primary classes: statistical techniques and artificial intelligence. [Examples of statistical data analysis techniques are:

- Data preprocessing techniques for detection, validation, error correction, and filling up of missing or incorrect data.
- Calculation of various statistical parameters such as averages, quantity, performance metrics, probability distributions, and so on. For example, the averages may include average length of call, average number of calls per month and average delays in bill payment.
- Models and probability distributions of various business activities either in terms of various parameters or probability distributions.
- Computing user profiles.
- Time-series analysis of time-dependent data.
- Clustering and classification to find patterns and associations among groups of data.
- Matching algorithms to detect anomalies in the behavior of transactions or users as compared to previously known models and profiles. Techniques are also needed to eliminate false alarms, estimate risks, and predict future of current transactions or users.

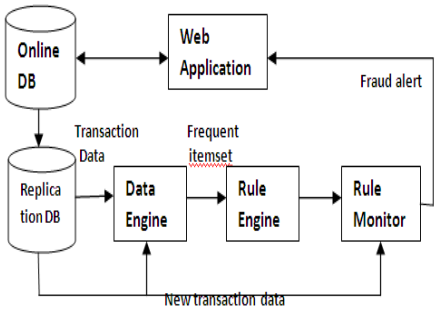

FIG 2. *Architecture of FDS*

The above figure 2 explains about the process involved in fraud detection. This shows about flow of data using databases , procedures and the application that are involved. This will also help us to detect the unauthorised usage of card .

Machine learning and data mining

Early data analysis techniques were oriented toward extracting quantitative and statistical data characteristics. These techniques facilitate useful data interpretations and can help to get better insights into the processes behind the data. Although the traditional data analysis techniques can indirectly lead us to knowledge, it is still created by human analysts.

To go beyond, a data analysis system has to be equipped with a substantial amount of background knowledge, and be able to perform reasoning tasks involving that knowledge and the data provided. In effort to meet this goal, researchers have turned to ideas from the machine learning field. This is a natural source of ideas, since the machine learning task can be described as turning background knowledge and examples (input) into knowledge (output).

Supervised and unsupervised learning

The machine learning and artificial intelligence solutions may be classified into two categories: 'supervised' and 'unsupervised' learning. These methods seek for accounts, customers, suppliers, etc. that behave 'unusually' in order to output suspicion scores, rules or visual anomalies, depending on the method.

Whether supervised or unsupervised methods are used, note that the output gives us only an indication of fraud likelihood. No stand alone statistical analysis can assure that a particular object is a fraudulent one. It can only indicate that this object is more likely to be fraudulent than other objects

Supervised methods

In supervised learning, a random sub-sample of all records is taken and manually classified as either 'fraudulent' or 'non-fraudulent'. Relatively rare events such as fraud may need to be over sampled to get a big enough sample size. These manually classified records are then used to train a supervised machine learning algorithm. After building a model using this training data, the algorithm should be able to classify new records as either fraudulent or non-fraudulent.Supervised neural networks, fuzzy neural nets, and combinations of neural nets and rules, have been extensively explored and used for detecting fraud in mobile phone networks and financial statement fraud.

Unsupervised methods

In contrast, unsupervised methods don't make use of labelled records. Some important studies with unsupervised learning with respect to fraud detection should be mentioned. For example, Bolton and Hand use Peer Group Analysis and Break Point Analysis applied on spending behaviour in credit card accounts. Peer Group Analysis detects individual objects that begin to

behave in a way different from objects to which they had previously been similar. Another tool Bolton and Handdevelop for behavioural fraud detection is Break Point Analysis. Unlike Peer Group Analysis, Break Point Analysis operates on the account level. A break point is an observation where anomalous behaviour for a particular account is detected. Both the tools are applied on spending behaviour in credit card accounts.

B. Markov Model

A mechanism is developed to determine whether the given transaction is fraud or not.The Mechanism uses "HIDDEN MARKOV MODEL" to detect fraud transaction. This mechanism works on the basis of spending habit of user and then classifies users in to Low, Medium or High Category

HIDDEN MARKOV MODEL:

It has Automatic techniques. Contains finite set of states and initially trained with Cardholder. Preparation are made to take action at exact time.

Graph: 1 Analysis of Card Fraud Worldwide

The above graph explains about the analysis about amount of losses to the card holders during the year's world-wide. This graph indicates the increase in the losses every year. The credit card fraud has increased during the years.

Conclusions and Recommendations:

- Fraud is a universal problem. Trends in fraud schemes, perpetatorscharacteritistics, and Anti Fraud control are similar regardless of where the fraud occurred.
- The longer the fraud lasts, the more the financial damage.Proactive detection methods – hotline , management review procedures, Internal Audits, Employee monitoring mechanisms, are vital in catching frauds early and limiting losses.
- Small Business or organisational are disproportionately victimised by fraud and under protected by anti frauds controls.
- Primary detection methods of fraud of 3% of cases and 7% of cases of detected accidently.
- Anti Frauds are concentrated on Data Monitoring & Analysis, Surprise Audits and Fraud Risk Assessment.
- Majority of fraudsters are first time offenders and don't rely on background checks.

Output of Hidden Marcov Model

Table 1 : Output of Analysis using Hidden Marcov Model

First Name	Kumar
Last Name	Singapore
Age	18
Sex	M
Card Number	143938
Expiration Month	8
Expiration Year	18
Security Code	Jf93k49fl
Process Payment	
FRAUD DETECTED	

The above table explains about the analysis made for Credit Card for Fraud Detection. It contains the Card Holder information along with card number which was used for transaction purpose while committing fraud.

REFERENCES

[1]. Agrawal, R. and Srikant, R. (1993): Fast algorithms for miningassociation rules. In Proc. Ofthe 20th Intl. Conf. on Very Large Data Bases, pp.478–499. Santiago, Chile.

[2]. Agrawal, R. and Srikant, R. (1995): Mining sequential patterns. In Proc. of the International onference on Data Engineering, 3–14.Taipei, Taiwan

.

[3]. Brin, S., Motwani, R. and Silverstein, C. (1997): Beyond market basket: generalizing association rules to correlations. In Proc.Of the ACM SIGMOD Intl. Conf. on Management of Data, pp. 265– 276. Tucson, Arizona, USA

[4] ChaiyakornYingsaeree and Philip Treleaven, UK Centre for Financial Computing, London Giuseppe Nuti, Citadel Securities, New York "*Computational Finance*" published by the IEEE Computer Society,2010

[5]. Chan, P. and Stolfo, S. (1998): Toward scalable learning with nonuniform class and cost distributions: A case study in credit card fraud detection. Proc. of the Fourth International Conference on Knowledge Discovery and Data Mining, pp.164–168.

[6]. M. Cornish , K. Delpha, and M. Erslont "*Master Card International Security & Risk Management: CREDIT CARD FRAUD*," Journalfor Risk Management, 2010.
[7]. Philip K. chan, Florida Institute of technology,Wei Fan. Andreas L. Prodomidis, and Salvatore J. Stolfo, Columbia University, "*Distributed data mining in credit card fraud detection"*Nov/dec1999 IEEE.
[8]. R. Shenbagavalli, A. R. Shanmugapriya, and Y. LokeshwaraChowdary "*Risk Analysis of Credit Card Holders*" International Journal of Trade, Economics and Finance, Vol. 3, No. 3, June 2012
[9] J. Xu, A. H. Sung and Q. Liu (2005). Online fraud detection system based on non-stationery anomaly detection, The International Conference on Security and Management.
[10].Y. Sahin and E. Duman, "Detecting Credit Card Fraud by Decision Trees and Support Vector Machines", International Multiconference of Engineers and computer scientists March, 2011.