International Journal of Computer Science and Network (IJCSN)
Volume 1, Issue 4, August 2012 www.ijcsn.org ISSN 2277-5420

Credit Card Fraud: The study of its impact and detection techniques

[1]Khyati Chaudhary, [2]Bhawna Mallick

[1]Dept. of Computer Science, GCET,

[2]Dept. of Computer Science, GCET,
Greater Noida

Abstract

With the rise and swift growth of E-Commerce, credit card uses for online purchases has increased dramatically and it caused sudden outbreak in the credit card fraud. Fraud is one of the major ethical issues in the credit card industry. With both online as well as regular purchase, credit card becomes the most popular mode of payment with cases of fraud associated with it are also increasing. A clear framework on all these approaches will certainly lead to an efficient credit card fraud detection system. Currently, for simplicity reasons, all the base learners for credit card fraud detection use the same desired distribution. It would be interesting to implement and evaluate the credit card fraud detection system by using very large databases with skewed class distributions and non-uniform cost per error. This paper presents a analysis of cost incurred in credit card fraud detection on data set.

Index Terms- Internet, Credit Card, Fraud Detection, Cost-Analysis

1. INTRODUCTION

As with the enormous growth of Electronic-Commerce over Internet, Globalization is also increasing. Credit Card Fraud is one of the biggest threaten to business establishments today. Fraud can be defined as criminal deception intended to result in financial gain. Along with the developments in the Information technology, fraud has been extending all over the world with results of huge financial losses. With the increased use of credit cards, fraudsters are also finding more opportunities to fraudulent activities which effects bank as well as card holders to large financial losses. Credit card transactions had a total loss of 800m$ of fraud in U.S.A. in 2004.In the same year in U.K., the loss caused by the credit card fraud amounts to 425m pounds ($750m)[1]. With the unknown amount of losses due to fraudulent activities on credit cards, various research analysts, reports to coincide that the figure for year 2002 probably exceeds $2.5 billion [2].

Credit card fraud can be partitioned into two types: Inner card fraud and External card fraud. Using false transactions to defraud banks cash, “Inner card fraud” is the collision between merchants and cardholders. External card fraud is mainly using the stolen or fake credit card to consume such as buying the expensive, small volume commodities or the commodities that can easily be changed into cash. As a result, credit card payment systems must be supported by efficient fraud detection capability for minimizing unwanted activities by opponent(s).

Credit card frauds have been ever-growing today. E-Commerce volumes continued to grow over the past few years, the figure of losses to Internet merchants was found to be between $5- $15 billion in the year 2005. Recent statistics by Garner group place online fraud rate between 0.8 to 0.9%[3]. Several techniques in data mining, such as Bayesian Networks (BN), Case-Based Reasoning (CBR), Decision Tree (DT), Neural Networks (NN), and Logical Regression (LR) have broadly been used to develop several fraud detection systems (FDS). If a bank cannot frequently obtain updated fraud patterns, it might continuously suffer fraud attacks. The **ACFE** (Association of Certified Fraud Examiners) defined fraud as “the use of one’s occupation for personal enrichment through the deliberate misuse of the employing organization’s own resources or asset(s).[4]”

Credit-card-based transactions can be classified into two types: 1) Physical card and 2) Remote/clicker card transaction. In **Physical Card** based purchase, the cardholder(s) presents his card physically to a merchant for creating a payment [5]. In this type of transaction for carrying out fraudulent operation(s), an attacker has to steal the credit card. On the other kind of transaction, only some essential information about a card (Card Number, Expiration Date, Secure Code, Credit Card Verification

(CCV)) is required to make the payment. For committing fraud in these type of transactions, fraudsters needs to know the card details, time, the Genuine cardholder is not aware that someone else has seen or stolen his/her card information. To detect this type of fraud, firstly analyze the spending profile of the users or customers on every card and to figure out any inconsistency with respect to the "usual" spending patterns.

Data mining aims to uncover these hidden or uncover patterns and predict future trends and behavior(s) as well in financial markets. Data mining has been applied to a number of financial applications, including development of trading models, investment selection, loan assessment, portfolio optimization, fraud detection, real-estate assessment, bankruptcy prediction and so on. It applies data analyzing and knowledge data discovery (KDD) techniques under acceptable computational efficiency limitations, and produces a particular variety of patterns over the data. Fraud Detection (FD) based on analyzing current purchase data of cardholders is an appropriate way to reduce the rate of successful credit card frauds. Some of the financial transactions are:

• Funds transfer between bank's accounts.
• Transferring of funds from bank's account to any other national or international bank's account.
• Credit card payment refers only to credit cards issued by another bank.

Credit card fraud detection also has two highly unusual characteristic(s). Obviously at first, the very limited time period in which the acceptance or rejection decision regarding credit card(s) has to be made. Secondly, the enormous amount of credit card operations that has to be processed at a given time period. Huge technologies has been used in detecting fraud include Neural Network Models, Intelligent Decision Engines (IDE), Expert Systems, Meta-Learning Agents, Machine Learning, Pattern Recognition. Credit Card Fraud(s) (CCF) can be made in many other ways such as simple theft, application fraud, counterfeit cards, Never Received Issue (NRI) and online fraud. Credit card fraud," Bolton and Hand" (2002) cite estimates of US$ 10 billion losses worldwide for Visa/Master card only.

Ghosh and Reilly (1994), Detection System (DS) has been proposed which is trained on a large sample of labeled credit card account transactions [6]. Feasibility study demonstrated that due to its ability to detect fraudulent patterns on credit card accounts, it is possible to achieve a reduction rate from 20% to 40% in total fraud losses. Brause et al. combine advanced data mining techniques and Neural Network (NN) algorithms achieve high Fraud Detection Rate (FDR) along with low false alarm. Vatsa et al. have proposed a credit card fraud detection system based on game-theoretic approach. Kahn and Schmittlein have described shopping trip behavior based on empirical observations.

In this paper, we represents the customer purchasing behavior patterns and detection of number of fraudulent transactions within limited time along with cost has been analyzed on these transactions.

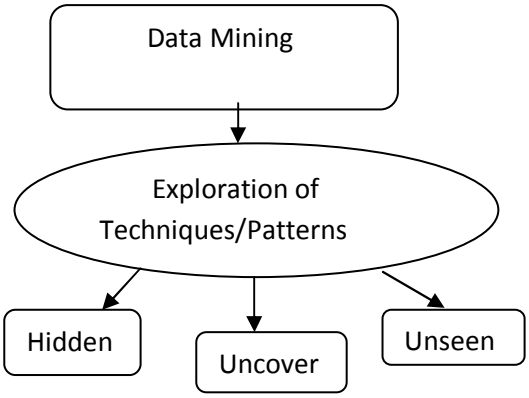

Figure 1: Data mining techniques

2.Relevant Work

Fraud detection involves eaves dropping on the behavior of user(s) for estimating, detecting, or avoid undesirable behavior of customers. From the work of view for preventing credit card fraud, more research works were carried out with more emphasis on data mining. Sam and Karl (2002) suggested a Credit Card Fraud Detection System (CCFD).

Bayesian Network (BN) and Neural Network (NN) techniques are used to learn models of fraudulent credit card transactions. Zaslavsky & Strizkak (2006), Ukraine proposed SOM (Self Organizing Map), algorithm for detection of fraudulent operation(s) in payment system using neural networks. Dorronsoro et al (1997) emphasizes on neural classifier using Neural Networks. Kim and Kim have associate skewed distribution of data and mixture of legal and false transactions, depicted two main reasons for the complexities of credit card fraud detection (CCFD). Hanagandi, Dhar and Buescher (1996) "historical information on credit card transactions to generate a fraud estimation model". **Syeda et al** used parallel granular neural network for improving the speed of data mining and Knowledge Discovery (KD) process for credit card fraud detection [7]. Yet, it could achieve reasonable speed up to 10 processors only, more number of processors introduces load imbalance problem. Chiu et al have proposed web-services based collaborative scheme for fraud detection in the banking industry. In present scenario, for keep tracking

of customer behavior and spending patterns, many fraud detection techniques involve practical screening of transactions has been deployed by both merchant companies as well as Banks. Some of the well-known techniques include Card Verification Method (CVS), Address Verification Systems (AVS), Rule-based systems, Personal Identification Number (PIN), and Biometrics Convergence on statistical analysis of customer/user data and deciphering customer spending behavior with the help of Data Mining methods as well as Risk-Scoring methods. Neural Networks, which are capable of being 'trained' and 'learned' can assume patterns out of data and are 'adaptive' to changing/modifying schemes of fraud. Another method used for detection is Decision Tree. (**Quinlan,** 1993) learning system, decision tree method has developed C4.5 that can deal with continuous data and **Quinlan,** (1986) has developed ID3 method as detection method. ID3 method has many advantages [8]. At first, it has high flexibility that it has data distribution without any assumption, and the second is the good robustness as well as explainable, which is also the reason of its wide utilization.

With the key of isolating and resolving, Decision Tree usually separates the complex problem into many simple modules and resolves the sub-problems through repeatedly using various data mining method (s) to uncover training several kinds of classifying knowledge via constructing decision tree. Decision Tree model focuses on how to construct a decision tree with high precision and small scale. Decision tree represents table of tree shape with many connecting lines. Each node is either a branch node followed with more nodes or there is only one leaf node signed by classification.

On the other side, Neural Networks is the appropriate and widely used method in fraud detection. (**Rumelhart**, 1986), "Neural networks architectures, or topologies", formed by organizing nodes into layers (neurons) and linking these layers of neurons with modifiable weighted interconnections [9]. Recently, neural network researchers have associated methods from statistics as well as numerical analysis into their networks. From the given cases, being a space to output space, neural networks can learn as well as summarizes the internal principles of data. With the result of formation of general capability of evolution from present situation to the new environment, can adapt its own behavior to the new environment. On the other side, there are still many disadvantages as well for the Neural Networks (NN), which include the efficiency of training, difficulty to confirm the structure, excessive training, and so on.

Some work has been done using Logistic Regression (LR) method, many statistical models are applied at data mining tasks include multiple discriminant analysis, regression analysis, logistic regression, and Probit Logistic regression. It is very similar to a linear regression model (LR) but is suited to models where the dependent variable is partitioned. Logistic regression (LR) coefficients can be used for estimation of odds ratios for each of the independent variables in the model. LR is applicable to a broader range of research situations as well as to analysis. Support vector machine (SVM), a new type of classifier, has been introduced and has strong theoretical foundations. SVM achieves excellent success in many fields such as Bioinformatics, Pattern Recognition, and Multivariable Regression. SVM has not only used in the credit evaluation but also obtains some valuable results [10]. However, there are few drawbacks which prevent SVM from going further. It has been proved that SVM is generally perceptive to class distribution and incurs high misclassification cost at first. Unluckily, the credit assessment problem is a class imbalance problem, whereby the misclassification cost is non uniform and the class distribution is unbalanced. Another work is on Web Services-Based Collaborative Scheme for Credit Card Fraud Detection. With this proposal, concerning participant banks can share the knowledge about fraud patterns in a heterogeneous and distributed environment. Analysis of previous spending data patterns is a promising way to reduce the rate of successful credit card fraud cases. Since humans lean to illustrate specific 220 behavioristic profiles, every customer can be represented by a. set of patterns containing information about some typical purchase category, the time since the last purchase, the amount of money spent, etc. Preventing credit card fraud, more research works were carried out with special emphasis on Neural Networks (NN) and data mining. Aleskerov and Freisleben (1997) present CARDWATCH, a database mining model used for credit card fraud detection (CCFD). The system/model uses neural network to train some definite historical consumption data and consequently generate Neural Network Model (NNM). This model was adopted to detect fraudulence cases and was very effective. Sam and Karl (2002) proposed a credit card fraud detection system using Bayesian Network and Neural Network techniques to learn models of fraudulent credit card transactions [11]. All approaches above-mentioned do not concern to convert the training data into confidence value. Usually, a preset threshold is set to detect abnormal and normal spending patterns of customers in the above-mentioned approaches without concerning the cost problem consequent from False Positive (FP) and False Negative (FN).

Normally used fraud detection methods/techniques are ANNs, Decision Trees, Meta-Heuristics rule-induction techniques, LR, and Support Vector Machines (SVM) such as genetic algorithms, nearest neighbor algorithms and k-means clustering. These techniques can be used alone or in cooperation with using meta-learning techniques to build classifiers past data in the credit card data warehouses are used to form a data mart representing the individual user profiles of the customers. These profiles consist of variables each of which reveals a behavioral characteristic

of the customer and these variables may show the spending habits of the customers with respect to their geographical locations, days of the month, hour of the day or MCCs. Towards, these variables are used to build a model to be used in the fraud detection systems/models to distinguish fraudulent activities which show significant deviations from the profile of the customer stored in the data-mart.

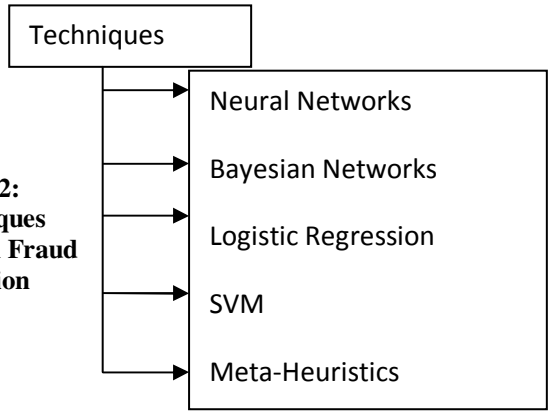

Figure 2: Techniques Used in Fraud Detection

3.THE COSTS OF FRAUD

According to LexisNexis, a computer-assisted legal research service, credit card fraud costs bank credit card issuers about 1 billion dollar annually. In U.S., LexisNexis conducted a study in 2010 in which more than 5,000 consumers and 1,000 merchants, financial executives, with the true cost of fraud [12]. The study records out some facts that merchants pay more than three times the dollar value on the respective fraudulent transactions. Most of the credit card fraud in the U.S. hit the card issuers mostly, as they are the victims of fraud losses. One of the research firms analyzed Fraud Prevention Systems and results two main types of credit card fraud as Card Not Present (CNP) transactions and counterfeit or lost/stolen cards. If a stolen credit card is used to purchase from company, then company can be responsible when the legal cardholder challenges the transaction. Another article by Bloomberg Business Week reported in 2009 that during the first two quarters of the year online banking fraud had increased by 55 percent. In spite, in UK, annual losses from online banking fraud run nearby to £80m [13]. Financial Fraud Action in UK has warned that online fraud has been increasingly sophisticated with the increased use of malware and phishing scams. As fraud reduction online security measures, are becoming effectual security systems. For example, Authorities that require the cardholder to use a password for online purchases have contributed to a reduction of 18 percent in fraud. Cyber Source's annual reported that the rate of fraud detection outside the U.S. is higher along with the estimation that U. S. merchants reject one of every nine international orders for "suspected fraud".

4.CARD AUTHORIZATION:

Authorization approval is just means that the card hasn't been reported as lost or stolen anywhere and at the time of authorization, the funds of transaction were covered. Mean while, if the card is certainly stolen (or even if it's not) cardholder can dispute the charge.

5.COST OF CHARGEBACKS:

The card issuer will levied an additional merchant bank fee of $5 to $35 per transaction [14]. If in case cardholder reports it to bank, despite of the reason for the chargeback, you are assessed a fee for the chargeback.

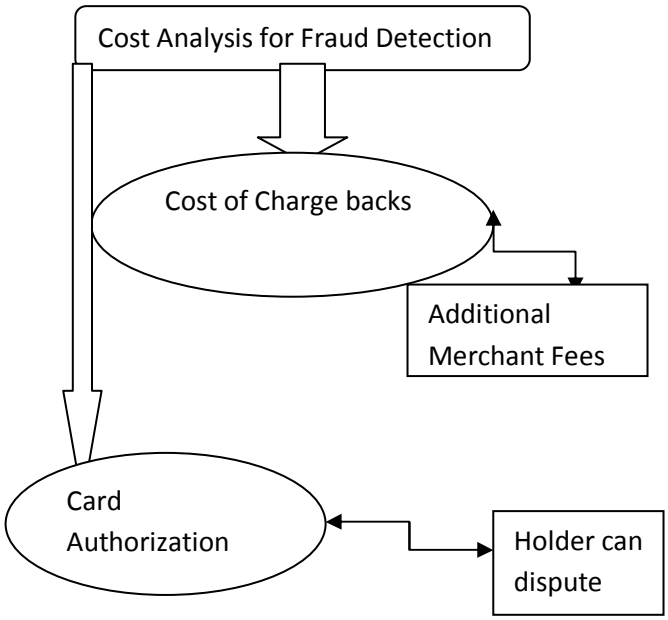

Figure 3: Costs for fraud

6.CONCLUSION

Credit card fraud has become more hazard in recent years. Handling credit card, risk monitoring system is the key task for the merchant banks to improve merchants' risk management level in a scientific, automatic and valuable way of building an accurate, available and easy system. Studies are encouraged to get better the fraud detection criteria, to set more appropriate weight and cost factor with both good tested accuracy and detection accuracy.

Necessary constraint for any card issuing bank is making well-organized credit card fraud detection system and a number of techniques have been proposed for detection of credit fraud. Neural network based **CARDWATCH** shows good accuracy in fraud detection and its processing speed is also high, as well as it is restricted to one-network per customer. All the techniques used for credit card fraud detection discussed in this study we have come to know that every fraud detection systems has its own strengths and weaknesses. Such category of study will enable us to build a hybrid approach for identifying fraudulent credit card transactions as a future scope. As in daily life, usage of credit card becomes more and more common in every field of the credit card fraud. Building of an accurate and resourceful credit card fraud detection system is one of the chief tasks for the financial institutions. Though, as the distribution of the training data sets become more biased, the performance of all model decrease in catching the fraudulent transactions. As a substitute of making performance comparisons just over the prediction accuracy, these comparisons will be extended to include the comparisons over other performance metrics as well, especially the cost based ones.

References

1. Abhinav Srivastava, Amlan Kundu, Shamik Sural, Arun K. Majumdar, 2008."Credit Card Fraud Detection Using Hidden Markov Model"

2. Aihua Shen1, Rencheng Tong1, Yaochen Deng2, "Application of Classification Models on Credit Card Fraud Detection", 2007.

3. Jon T. S. Quah and M. Sriganesh, Real Time Credit Card Fraud Detection using Computational Intelligence, Proceedings of International Joint Conference on Neural Networks, Orlando, Florida, USA, August 2007.

4. Sahin, Y., Duman, E.: An overview of business domains where fraud can take place, and a survey of various fraud detection techniques. In: Proceedings of the 1st International Symposium on Computing in Science and Engineering, Aydin, Turkey (2010).

5. Kaiyong Deng,Ru Zhang,Dongfang Zhang,WenFeng Jiang,Xinxin Niu, Kaiyong Deng, Ru Zhang, Hong Guo, Analysis and Study on Detection of Credit Fraud in E-commerce, 2011.

6. Kou, Y., Lu, C.-T., Sirwongwattana, S., Huang, Y.-P.: Survey of fraud detection techniques. In: Proceedings of the 2004 IEEE International Conference on Networking, Sensing and Control, Taipei, Taiwan (2004).

7. Mirjana Pejic-Bach, Profiling intelligent systems applications in fraud detection and prevention: survey of research articles, 2010 International Conference on Intelligent Systems, Modelling and Simulation

8.. Prabin Kumar Panigrahi, A Framework for Discovering Internal Financial Fraud using Analytics, International Conference on Communication Systems and Network Technologies 2011.

9. Raghavendra Patidar, Lokesh Sharma, Credit Card Fraud Detection Using Neural Network, International Journal of Soft Computing and Engineering (IJSCE), June 2011.

10. S. Benson Edwin Raj, 2A. Annie Portia International Conference on Computer, Communication and Electrical Technology – ICCCET2011, "Analysis on Credit Card Fraud Detection Methods"

11. Tej Paul Bhatla, Vikram Prabhu & AmMirjana Pejic-Bach, Profiling intelligent systems applications in fraud detection and prevention: survey of research articles, 2010 International Conference on Intelligent Systems, Modelling and Simulation it Dua "Understanding Credit Card Frauds," 2003.

12. Y. Sahin, E. Duman "Detecting Credit Card Fraud by ANN and Logistic Regression" 2011.

13. Yi Peng, Gang Kou, A Comparative Study of Classification Methods in Financial Risk Detection, Fourth International Conference on Networked Computing and Advanced Information Management, 2008 IEEE.

14. Y. Sahin, E. Duman**,** Detecting Credit Card Fraud by ANN and Logistic Regression, ©2011 IEEE.