AN OVERVIEW OF THE BANK FRAUD AND ITS DETECTION TECHNIQUES THROUGH DATA MINING

SeyyedHedayatTarighiNejad[1] and MohammadNikbakht[2] and Mohammad HosseinAfrakhteh[2]

[1]Faculty of Engineering, Islamic Azad University, Yasuj Branch

[2]Keshavarzi Bank Staff Management, Kohgiluyeh and Boyer-Ahmad Province

ABSTRACT

Using modern methods of the electronic commerce in daily life transactions is increasing because of the growth and the comfortable access of the people to the internet and social networks. The electronic payment systems are one of the most important electronic commerce methods and the electronic payment fraud is a major problem.For example, the credit card fraud loss increases every year and is regarded as one of the important issues in the credit card institutes and corporations. Therefore, fraud detection is considered as an important research challenge. Fraud reduction is a complicated process requiring a body of knowledge in many scientific fields. Based on the kind of the fraud the banks or the credit card institutes face, different measures may be taken. This paper compares and analyzes the available recent findings on the credit card fraud detection techniques. The objectives of the present study are first to detect different credit card and electronic commerce fraud and then to investigate the strategies used for the purpose of detection.

KEYWORDS

credit card fraud, fraud detection, data mining techniques, electronic commerce, decision tree, electronic payment systems

1. INTRODUCTION

- The contemporary era has been called the era of information and it is the information which leads to power and success. Owing to the today's complex technologies such as computers, satellites, and the means for information storage, one can collect and store different data and use the computers to arrange the bulk of information. Unfortunately, the large bulk of the stored data on different computers quickly became cumbersome. This chaos led to the creation of the organized data bases and data base management systems [1].
- The technology growth and development have provided the possibility of committing fraud in different fields such as banks, insurance companies, security fraud, goods fraud, and the other fields for the profiteers [2].
- Fraud is a purposeful action for illegally gaining financial benefits which is against the typical laws, regulations, and policies. The banking fraud in banks and financial institutes has become a serious problem in the recent years and attracted much attention and concern. It is of critical significance to detect the financial fraud in order to prevent the subsequent devastating consequences [3].

2. ELECTRONIC BANKING

The electronic banking refers to all banking and financial services which are electronically offered to the customers; it is completely core and person-to-person/online. In other words, the electronic banking is composed of the integrated systems offering all bank products and services

DOI : 10.5121/ijmnct.2017.7301

as well as the strategic operations and their management through electronic equipment to the core data base in the form of a system [4].

3. KINDS OF THE CREDIT CARD FRAUD

The criminals commit the credit card fraud in different ways. As the technology changes, the criminals change their method as well. Different classifications of the credit card fraud can be presented based on the different viewpoints [5].
For example, as in [6],regarding the fact that whether the fraud is committed inside or outside the organization, it is classified to two categories of intra-organizational and extra-organizational fraud [7].
The definitions and techniques of this kind of fraud are explained in the present paper.

3.1. BANKRUPTCY FRAUD

The detection of the bankruptcy fraud is one of the most difficult ones. However, there are some techniques to prevent it. The bankruptcy fraud means to use a card whose owner cannot repay the debts. In other words, the individuals used the cards for purchase since they know that they cannot repay their debts. The bank sends them the bill. Nevertheless, they have declared to be bankrupt, have an irrecoverable debt, and the bank has to repay the debt itself. The only way to prevent this kind of fraud is to investigate the customers' banking and credit status by the credit rating agencies so as to know their financial status [7].

3.2. LOST OR STOLEN CREDIT CARDS

The lost or stolen credit card is a card whose owner loses after receiving it or a person steals it for an evil purpose and uses it until it is expired. This is the easiest way for the fraudsters to receive a credit card without spending any money for buying a special technology. This is an old way of credit card fraud and is more difficult to detect and deal with, compared to the other methods.

3.3. THE REQUEST FRAUD

This kind of fraud is committed when a person gives the incorrect information and requests a credit card. This is done in the three following ways:

False Identity

It happens when a person illegally obtains the personal information of a person and requests to open a bank account.

Financial Fraud

It happens when a person gives incorrect information regarding his financial status while opening a bank account.

Non-delivery of the Post Package

It happens when the posted credit card is stolen before it is delivered to the realowner.

3.4. THE ACCOUNT CONTROL

This kind of fraud happens when a criminal illegally obtains the personal information of a person and controls the legal account of that person by having his account number or the credit card number. Then, he calls the bank as the real person or the card issuer and informs them of a change in the postal address; after a while, he announces that the card has been stolen and requests a new card [7].

3.5. CREDIT CARD FORGERY

Forgery, theft, and loss are the major threats for the credit cards. The criminals constantly innovate new ways to forge the credit cards. Some methods used for the credit card forgery are explained below:

3.5.1. Erasing the Magnetic Stripe

A criminal can manipulate the illegally obtained card by erasing the magnetic stripe with a strong magnetic field. Then, he changes the card information in a way that would be consistent with a legal card whose information he has already obtained. The bank cashier puts the credit card in the card reader for several times, does not know that the magnetic stripe has been erased, and finally has to manually enter the card information to the card reader. This kind of fraud is very dangerous since the cashier meticulously examines the card while trying to read the card number. Card manipulationis among the old methods for the credit card fraud used to illegally gain money and wealth [7].

3.5.2. Making Fake Credit Cards

A criminal is able to make a fake credit card by using the appropriate tools. This is a common way in committing fraud though making a fake credit card requires a great deal of effort and skill. The new credit cards have a very large amount of secure information making it difficult to forge them. It is very difficult to precisely forge the card holograms. One of the other difficulties in forging a card is indenting the holograms on which [7].

3.5.3. Changing The Card Information

A criminal is able to change the information that the legal card maker has carved on the credit card by means of heat and pressure. He can also recode the credit card magnetic stripe by using the computer software.

3.5.4. Copying

Most of the credit card forgery methods are done through copying. In this process, the main data of a credit card magnetic stripe are electronically copied and written on another credit card. The instruments for doing this job are the small and portable magnetic stripe reader electronic machines which operate by a battery [7]. The fraudsters do this while the customers are waiting for the electronic transaction confirmation by the card reader machine. This is done without the card holder's awareness. Therefore, it is difficult to detect and trace. The obtained information can be used in places where one does not need the credit card to buy goods (online/postal orders). In most of the cases, the card holder does not become aware of the issue until receiving the bill at the end of the period.

3.5.5. White Card

It is a plastic card with the same size of a credit card including the main data of alegal card magnetic stripe which is used for fake transactions. This card is used in places where the approval and permission of the card reader are not needed (like gas stations and ATMs).

3.5.6. The Seller's Collusion

This kind of fraud happens when the store owner or his employees are going to abuse the customers' account information and may give the information to the profiteers and fraudsters.

3.5.7. Online Fraud

Because the internet has broken the geographical borders in today's world, plays a significant role in the peoples' daily life, and most of the peoples' needs and purchase are done through the internet, the grounds for committing other kinds of fraud has been provided for the fraudsters and criminals. Some online fraud methods are explained in the following.

3.5.7.1. Making A Similar Website

In this method, the fraudsters simulate all or the shopping form page of a reputable website; therefore, the customers have no reason to suspect it because of the great similarity. The fake website receives the customers' shopping orders like the original website and sends them through email. As such, the fraudsters receive the necessary information for the credit cards' fraud and abuse and the customers do not understand.

3.5.7.2. The Fake Sale Website

In this method, the fraudsters run a website and offer some products with a very high discount and sending before any prepayment. This website may be seemingly an auction or a legal sale website. The customers enter their information such as name, address, and the credit card information in the shopping form while they are going to buy goods. Having accessed the information, the criminals buy goods from the legal websites. Therefore, they will be able to buy a large bulk of goods with the stolen credit card numbers or sell the obtained information for other criminal conducts [7].

3.5.7.3. Credit Card Generator

They are the computer programs which can generate credit card numbers and valid expiry dates. Using the legal credit card numbers, these generators can achieve the algorithm the card issuer uses for producing credit card numbers and generate a list of the valid numbers. These generators help their users to generate numbers for different kinds of credit cards like Visa and MasterCard [7].

4. DIFFERENT KINDS OF FRAUD IN THE REALM OF ELECTRONIC BANKING

The attack approaches are extensively divided into two categories based on the model of the attacks:

4.1. ABUSE DETECTION

Abuse detection means that a set of the intrusive rules are stored in a base and all available transactions and data are tested with these rules; each transaction or data following these rules is identified as a fraudulent activity. The computer anti-viruses operate through this method such that they have a data base engaged in subversive activities. They constantly compare the intra-computer activities with these activities and in the case of consistency, they are detected as subversive activities. In fact, this is the detection and understanding of the previous activities so as to predict and detect the subsequent activities. These methods are usually very precise but have a major shortcoming. The major shortcoming is that they are not able to predict the new subversive and fraudulent activities since they do not have access to their rules [8].

4.2. ANOMALY DETECTION

Contrary to the previous technique, the rules of intrusion are not known in this technique such that the activities which are greatly deviated from the usual activities based on the available data are detected as subversive and fraudulent activities. In fact, the fraud patterns are not known in

this technique and the new fraud patterns can be discovered by the means of data. There are not special rules related to the fraudulent data and transactions in this technique. Therefore, the extent of the deviation of any data from the normal extent is calculated and if it is a large deviation, the data or the transaction is identified as a fraud. The major problem of this technique is that the customers' usual transactions and activities are sometimes considered as fraud. For example, some activities in the banking data base such as the large number of the different accounts of a single customer, the transactions with small values from lots of different accounts, a very large amount of the payment transactions in a specific account, and the increased failure in the number of the times the password is entered are considered as the fraudulent activities before the fraud is committed. Hence, these customers' behaviors can be considered as suspicious and regarded as fraud activities as soon as they are repeated [8].

5. DATA MINING

Data mining or discovering knowledge from data base is the inconspicuous extraction of the useful potential information from the data which were already unknown. In fact, data mining is the discovery of noticeable, unexpected, and valuable structures from a set of large amounts of data; it is an activity basically dealing with statistics and precise data analysis [9]. Some of the data mining techniques are explained below to discover the bank fraud.

6. THE INTRODUCED FRAUD DETECTION TECHNIQUES

Two important steps for coping with fraud are prevention and detection. Selecting PINs for the credit cards and using immunization protocols like SSL are applied in the prevention step; fraud detection instruments are used in the detection step. There are different instruments for fraud detection and most of the algorithms applied in which include machine learning or more especially data mining [8]. Data mining techniques are divided into two general categories: supervised and unsupervised [10]. In the supervised technique, the fraudulent or legal transactions are predicted based on the transactions whose type (fraudulent or legal) has been already determined. In fact, the discovery process is done with the help of the fraudsters' data, modeling their work, and classification of the fraud activities. However, the available structures in the data are automatically discovered without any supervision in the unsupervised technique. In fact, this technique is used to show the outlier data and the unusual transactions to detect the fraudsters' transactions. Some of the fraud detection data mining and statistical techniques will be explained in the following [10].

6.1. RULE-BASED DATA MINING

It is an unsupervised technique seeking to create the "if-then" rules with pattern learning and dependency between the available data so that it can reply to a new request for fraud detection by referring to the created rules [8]. For example, the rule-based technique has been used in [6] for the purpose of fraud detection. The introduced system has been implemented in five stages. In the first stage, the rules are randomly made with the help of the Apriori associative algorithm, then, in the second stage, the rules are exerted on the set of the legal transaction data and any rule which is consistent with the available data is excluded. In the third stage, the other rules are applied in the real system for leading and any rule which does not indicate any irregularity will be excluded. In the fourth stage, the selected rules are randomly implemented on the data set; in the final stage, the successful set of rules is used for fraud detection and the results will be used afterward.

6.2. DECISION TREE

It is one of the supervised techniques, makes a decision tree, and facilitates the decision-making process for fraud detection with regard to the illustration of the fraudulent and legal branches.

This technique has various applications [8].For instance, a decision tree algorithm named C4.5 has been utilized for the credit cards. This method uses weighted classification, takes the output middle limit, and leads to better results compared to the classification algorithm which only uses the exponential weight average [8].

6.3. NEURAL NETWORK

It is a supervised technique inspired by the human's neural network system for sending a message [8]. A neural network-based method on the parallel machines to facilitate rule production for the credit card fraud detection has been proposed in [11].

6.4. BAYESIAN NETWORKS

Bayesian networks are one of the other classification methods. In these methods, the prior probability of the classes is calculated based on the previous transactions. For each of the attributes, the effectiveness chance for determining a transaction class is calculated. By the entrance of the new transactions, the previous probabilities are added based on the attributes and the maximum value determines the transaction class. A very large amount of algorithms has been implemented based on the Bayesian networks principles [8]. This method has been used in [12, 13, 14, 15, 16, and 17].

6.5. CLUSTERING

It is one of the unsupervised techniques and makes it feasible to interpret the hidden patterns among the data through grouping or clustering the transactions [8]. The objective of clustering is to classify the data into different groups in order to become aware of the signs present among the data in each group and identify the behavioral patterns.

6.6. STATISTICAL METHODS

These methods are not among the machine learning methods, but they are appropriate instruments for fraud detection with regard to the possibility of competition with data mining methods in some applications. For instance, the least squares regression and the predictor stepwise selection have been used in [18] to indicate that the standard statistical methods can compete. The introduced method has had three useful results. First, it organizes the calculations to moderate the interactions. Second, it uses a modern decision-making status to select the predictors; finally, the needed values for adjusting the dispersed data are carefully selected [8].

6.7. FUZZY LOGIC

It is one of the methods used in both of the supervised and unsupervised methods. For example, the fuzzy clustering algorithm is one kind of the unsupervised algorithms. Moreover, the fuzzy phenomenon is used for categorization and fraud detection.

6.8. ARTIFICIAL IMMUNE SYSTEM

It is a supervised technique and its design has inspired from the human's body immune system. The artificial immune system algorithm has been used in [19] for fraud detection. In this system, the familiar cells are distinguished from the unfamiliarones.This method operates in a way that produces some discoverers which are resistant to the familiar cells and detect the unfamiliar ones. This system is of two significant characteristics which can be efficient in fraud detection. First, it does not need the unfamiliar cells at the beginning. Second, it is an adaptive system and will be able to detect and record different kinds of the new outsider cells in time.

6.9. GENETIC ALGORITHMS

The previous techniques lead to desirable results for different kinds of the data. In most of them, fraud detection is done on the basis of calculating the distance or the surrounding data compression. However, they are not very efficient in a set of a very large amount of data with large sizes; that is because neighborhood and locality in the dispersed data are more difficult and complicated and need special methods. The idea of the genetic algorithm has been shaped based on Darwin's theory of evolution [7]. This theory has proven that when the resources are limited in nature, different species compete with each other to obtain them; it is only the strong species that can survive. The winner species start mating and reproduction and better species are born. In reality, nature is a system of selection and optimization. The same concepts have been used in genetic algorithms. The solutions to the problems are the species which are specified by a score of efficiency and adjustment. The most efficient solutions are selected to create better solutions. Then, the superior solutions are combined with each other and reorganized. This process continues until an optimal solution is achieved [7]. This method has been used in [20].

7. CONCLUSION

As previously mentioned, the fraud growth is observed in today's life. When a new fraud technique is made, different methods are simultaneously introduced and implemented to cope with it. This paper investigated the data mining algorithms including the rule-based data mining, decision tree, neural networks, Bayesian networks, clustering, statistical methods, fuzzy logic, artificial immune system, and genetic algorithms. The results indicate that using combined methods like classification leads to an improvement in the results in terms of the criteria such as detection precision, detection pace, and detection cost. Applying the merits of the algorithms along with each other can be regarded as the main reason for the improvement and satisfaction. Selecting appropriate algorithms to combine is the major challenge which is strongly dependent on the kind of the applied data.

REFERENCES

[1] Moslemzadeh, A. “Data mining techniques for fraud detection in the financial statement auditing”.

[2] E.W.T.Ngai,"the application of data mining techniques in financial fraud detection: A classification," Decision Support systems, vol. 50, pp. 559-569, 2010

[3] G. Wang, " A comparative assessment of ensemble learning for credit scoring, " Expert system With Applications, vol. 38, no. 1, pp.223-230, 2011.

[4] R. Stephan Kovach, "Online Banking fraud detection based on local and Global Behavior," pp. 166-171, 2011.

[5] M. jans, N. Lybaert , k. Vanhoof, A framework for Internal Fraud Risk Reduction at IT Integrating Business Processes: the IFR2 Framework, The Internal journal of Digital Accounting Research, Vol .9. 2009, pp.1-29, ISSN: 1577-8517.

[6] C. Phua, V. Lee , K Smith , R. Gayler, 5. A Comprehensive Survey of Data Minin – Based Fraud Detection Research, March 2007.

[7] Nasiri, N. &Minaei, B. “Data mining techniques in the credit card fraud detection”.

[8] Jafarpour, M., Rafiee, A., &Shamsi, M. “The application of data mining in the credit card transaction fraud detection”.

[9] Gharibi, S.A., Movahedpour, M., &Mahnaei, O. (2016). “Data mining of the training for working with Clementine software”, 2nd ed., Tehran: Orang.

[10] S.Maes , K. Tuyls, B. Vanschoenwinkel, B. Manderick , Credit Card Fraud Detection Using Bayesian and Neural Networks,2002.

[11] M. F .A. Gadi, X.Wanq , A. P . do Logo, Credit Card fraud detection With Artificial Immune System , 2004.

[12] P.K. chan, w. fan, a.l, prodromidis, s.j. stolfo, distributed data mining in Credit Card fraud detection, IEEE Intelligent Systems, 1999.

[13] p. juszczak, n.m. adams, d. j. hand, ch.Whitrowa, d. j. Weston , off – the – peg and bespoke classifiers for fraud detection Elsevier b.v., computational statistics and data analysis 52 (2008) 121- 132.

[14] Ph. K. chan ,s.j.stolfo, toward scalable learning with non – uniform class and cost distributions : a case study in Credit Card fraud detection , march 1998.

[15] d. foster. And r.stine, " variable selection in data mining : building a predictive for bankruptcy , , "journal of American statistical association , , 2004.

[16] Watson G. Sottile, J. (2010) “Cheating in the Digital age: Do Students Cheat More in Online Courses?” Online Journal of Distance Learning Administration, Vol 13, No 1

[17] Watson G. Sottile, J. (2010) “Cheating in the Digital age: Do Students Cheat More in Online Courses?”

[18] Zhou W. , G. Kapoor, Detecting Evolutionary Financial Statement Fraud, Decision Support Systems, Vol. 50(3), 2011, pp. 250-576.

[19] Nasiri, N. &Bigdeli, B. (2010). “the application of data mining techniques in electronic banking for the detection of the suspicious financial transactions”, MA thesis, Qom University, Technical College.

[20] Mirza, N. & Staples, E. (2010) “Webcam as a New Invigilation Method: Students’ Comfort and Potential for Cheating”, Journal of Nursing Education, Vol 49, No. 2, pp 116-119.