ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 5, May 2017

Credit Card Fraud Discovery: A Survey

Ritika Wadhwa

M. Tech Student, Dept. of CSE, RPIIT, Karnal, Kurukshetra University, India

ABSTRACT: We are living in a digital world where credit card is dynamically ruling us globally as well as economically. Undoubtedly, this card has provided us major benefits in terms of purchasing goods and services online as well as offline. Nowadays, Customers are living in the Cutting – edge world where they prefer only electronic mode of payment i.e. cashless transactions via credit cards. We cannot forget the fact that Modern technologies are boon and bane as well. The rapid growth of credit card has hampered the security of bank databases and also compromised so many bank accounts. Fraudsters are so smart and tech savvy that they always generate different and new ways to commit fraud each passing day which demands constant research. In this survey paper, different credit card fraud discovery techniques and methodologies are mentioned in detail.

KEYWORDS: Credit Card Discovery; Fraud Detection Techniques; Decision Tree; Hidden Markov Model; Neural Networks; Genetic Algorithm; Support Vector Machine

I. INTRODUCTION

The credit card is the most convenient and acceptable mode of payment for online as well as offline shopping, paying bills and many other services. The cardholder can either use the card physically or virtually. For physical transactions, he needs to get the card swipe in the swipe machine at the time of payment whereas for virtual transactions, he just the need some details such as Credit Card Number, Expiry date and CVV Number. Following are the major frauds that tend to occur nowadays:

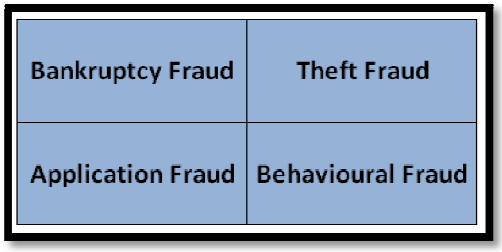

Fig 1: Types of Credit Card Frauds

- *Bankruptcy Fraud:* The cardholder knows that they are not able to pay the debts after using the credit money, so the bank has to cover the losses by itself.
- *Theft Fraud:* When someone tries to steal the sensitive information of the credit card such as Card number, expiry date or CVV through keystroke logging or skimming.
- *Application Fraud:* When someone applies for a credit card with the manipulated information of the actual user, then there is a high probability of application fraud.
- *Behavioral Fraud:* When the legitimate card details are obtained fraudulently by someone and sales are made.

II. LITERATURE SURVEY

In [1], the researchers have used Baum-Welch algorithm to train HMM for each cardholder. All the transactions have been modelled using HMM in a sequential manner. First of all to detect fraud, they have considered three behaviours of the card holder i.e. Low Spending Behaviour, Medium Spending Behaviour and High Spending Behaviour and recorded these behaviours in the form of observation symbols on the basis of K – Clustering Algorithm.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 5, May 2017

The whole system is basically divided into two parts-one is **generating observation symbol and training** and other is **detection**. Training is done offline, whereas detection part is an online process. It has been also explained that how an HMM can detect whether the incoming transaction is fraudulent or not. At last they have calculated the performance of system using TP (True Positive) and FP (False Positive) parameters and it is observed that accuracy of system is near to 75%.

According to [2], the researchers have used the concept of Hidden Markov Model (HMM) to identify the fraudulent transactions in credit card. The fraudulent transactions will only be detected after the transaction is done. The main objective of HMM is to obtain a high number of fraudulent transactions by analysing customer's spending patterns and any deviation from the regular pattern is considered to be a fraudulent transaction. The Fraud Detection System, based on HMM, records the spending habits of the customers and thereby creates the clusters of training sets. The main focus of FDS is on **"Amount of Item Purchased"** and is used for further processing. Depending upon the transaction amounts, different transactions are stored in the form of clusters under three categories: 'low', 'medium' or 'high' value series. This paper also includes the additional security features such IP Address detection, MAC Address detection and shipping details detection.

In [3] the hybrid approach has been used to detect credit card fraud by combining two different techniques i.e. Rough Set & Decision Tree. Their main aim is to improve the credit card fraud detection. To implement this hybrid approach they have taken a predefined **"German Credit Card"** Data Set (downloaded from UCI Website). This data set consists of 20 features and there is one class attribute which could be 'Good' or 'Bad'. The pre-processing of this data set has been done by using **Rough Set** theory introduced by Z Palwak (1982) for data analysis. The data pre-processing will reduce the complexity of data and offers better chances for subsequent analysis. Also, they have used **J48** classifier, a simple C4.5 decision tree for classification. A binary tree is created in this. This hybrid approach is mostly useful in classification. With this method, a tree is built as the model of classification process. Once the tree is made, it is applied to every tuple of the database and classification for that tuple is obtained.

2.1 VARIOUS CREDIT CARD FRAUD DISCOVERY TECHNIQUES

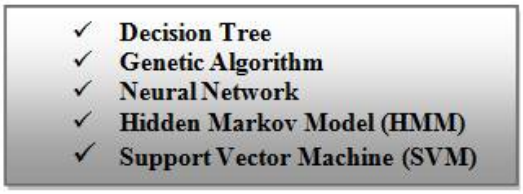

Fig 2: Credit Card Fraud Discovery Techniques

A. *Decision Tree:*

Decision Tree is a data representing and mining structure which comprises of a root node, branches and leaf nodes. Each internal node in the tree represents an attribute; each branch holds the outcome of a test, whereas each leaf node denotes a class label. The topmost node in the tree is the root node. And, the work of decision tree algorithm is to recursively partition the given data set of records using Depth-First or Breadth First Approach. Basically this tree structure reduces the complexity of the data sets by partitioning the unknown data sets.

On the basis of decision tree one more novel idea to detect the credit card fraud has been proposed by **Prajal Save et al. (2017)**. In this paper, the researchers have done Luhn's Test [4] to validate the credit card number and then further worked on address verification.

Credit Card Validation Test:

Step 1: Record the digits of credit card number for example: 5196190223830813

Step 2: Reverse the digits say, 3180383220916915

Step 3: Calculate the sum of the odd place digits: 3+8+3+3+2+9+6+1 = 35 = s1

3	1	**8**	0	**3**	8	**3**	2	**2**	0	**9**	1	**6**	9	**1**	5

Step 4: Select the even place digit and doubles it:

3	**1**	8	**0**	3	**8**	3	**2**	2	**0**	9	**1**	6	**9**	1	**5**

Outcome: 2, 0, 16, 4, 0, 2, 18, 10

Step 5: Calculate the sum of each digit in previous outcome

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 5, May 2017

2,0,7,4,0,2,9,1

Step 6: Sum the last = 2+0+7+4+0+2+9+1 = 25 = s2

Step 7: Calculate s1 + s2 = 35+25 = 60.

Since the resultant number after adding s1 and s2 divisible by 10, it passes the Luhn's test and hence the card number is valid.

After Credit Card Validation, address verification has been done by comparing shipping address and billing address. Although, this check does not guarantee whether the transaction is legal or fraudulent, but in case if the two addresses are found to be mismatched then it can be suspected that transaction is fraudulent with high degree of probability.

B. *Genetic Algorithm:*

Genetic Algorithm has been inspired from natural selection and genetics, introduced by Holland in 1975. The main objective of Genetic Algorithm is to obtain an efficient and optimal solution to a given problem as the time progresses. As we know that Credit Card Fraud Detection problem is a classification problem i.e. good and bad transactions have to be classified as a result:

Below is the flowchart for Genetic Algorithm:

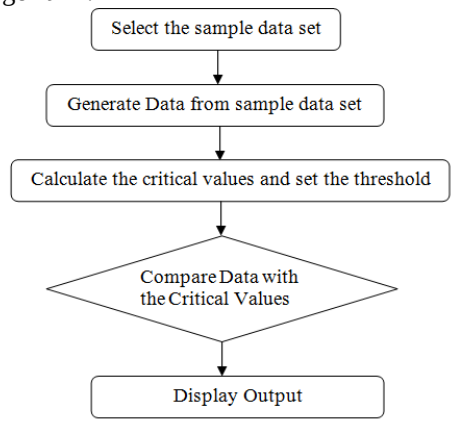

Fig 3: Genetic Algorithm Flowchart

Genetic Algorithm is an ongoing or repeatable process until the best and optimized solution is found. In [5] the researchers have proposed the Genetic Algorithm method for Credit Card Fraud detection payment system in which they have calculated the following various parameters from the data set:

CCfreq= number of times card used

CCloc = location at which CCs in the hands of fraudsters

CCoverdraft = rate of overdraft time

CCbank balance = balance available at bank of CC

CCdailyspending = average daily spending amount

By using genetic algorithm in the following way, the fraud has been discovered on the basis of customer's behaviour and further an optimized result has been produced:

Step 1: Initially, the credit card transactions with 'n' number of attributes are taken as input. Then the data is standardized and the final samples of transactions are obtained which comprises of the confidential information of the credit card holder.

Step 2: Then, the critical values i.e. CC usage frequency count, CC usage location, CC overdraft, current bank balance, average daily spending, are calculated.

Step 3: After this the resultant critical values i.e. Critical Fraud Detected, Monitorable Fraud Detected, Ordinary Fraud Detected after finite number of iterations are generated using Genetic algorithm.

Step 4: At last, the fraudulent transactions are generated.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 5, May 2017

C. *Neural Network:*

Neural Network operates just like a human brain works. We contain a supercomputer in our head which comprises of millions and billions of neurons which works to solve a specific problem. On the basis of this, the idea of neural network came i.e. an artificial neural network architecture has been developed which works similar to human brain after getting certain amount of training. In general, we can say that neural networks learn by example.

According to [6], credit card fraud detection can be done with the help of neural networks i.e. human brain working principle. They have trained the system (neural networks) on the basis of the attributes of a credit card holder. They say that first of all there is very limited time span in which the acceptance or rejection decision has to be made. And, the second one is the huge amount of credit card transactions that have to be processed in a particular time frame. As we know that our brain learn maximum from our past experiences, similarly neural networks learns. They are trained on the basis of some certain protocols. For example: in case of credit card fraud detection, there is a fixed protocol for the customers to access the credit card, which further decides whether the credit card user is fraudulent or not. As shown in the figure the neural networks are trained on the basis of certain information from various categories i.e. cardholder's occupation, income and frequency of large amount purchases, location where purchases take place etc.

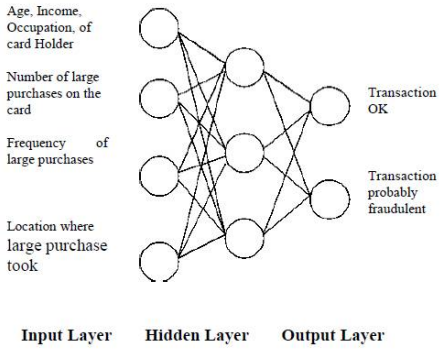

Fig 4: Neural Network for Credit Card Fraud Discovery System

D. *Hidden Markov Model*

A hidden Markov model is one in which certain number of emissions are observed but we do not know the sequence of states the model went through to generate the emissions. And, the hidden Markov model helps to analyse the sequence of states from the observed data. For example: predicting the weather for tomorrow or day after tomorrow after observing today's or yesterday's weather.

In case of Credit Card Fraud Detection, this methodology has been used by **Gaurav Mahtre et al. (2014)** in which the fraudulent or suspicious transactions have been detected on the basis of user's spending profile history. In [7], the fraud detection system will only be activated after 10 transactions done by the user. The probability of occurrences of the huge number of amounts in the transaction history of the user will generate a kind of suspicion for the user.

E. *Support Vector Machine:*

Support Vector Machine (SVM) is a supervised machine learning algorithm which is used basically for classification modelling. Each and every data item is plotted as a single point in n-dimensional space. The classification that we perform is generally by finding an accurate and precise hyper plane which can easily differentiate between two classes. The thumb rule here is to locate the hyper plane in such a way that distance between the hyper plane and two data points from two different classes is equal as shown in the below figure:

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 5, May 2017

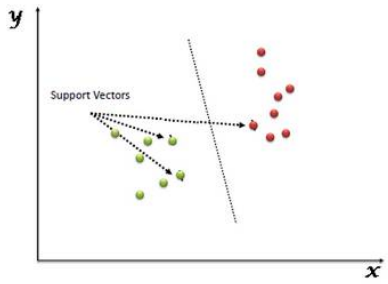

Fig 5: Support Vector Machine showing hyper-plane to differentiate between two classes

III. CONCLUSION

In recent years, the use credit cards have been vigorously increased. And, there is a huge responsibility on the heads of the credit card risk managers. Their key task is to improve the fraud detection algorithms in an optimal way so that frauds can be discovered adequately. In this paper, the researcher has done a survey on five important fraud detection techniques i.e. Credit Card Fraud discovery through Decision Trees, Genetic Algorithms, Neural Networks, Hidden Markov Model and Support Vector Machines. Some of these techniques have been applied by different researchers to produce an efficient system which can easily detect and report credit card frauds.

REFERENCES

[1] Avinash Ingole and Dr. R.C Thool 'Credit Card Fraud Detection Using Hidden Markov Model and its Performance', International Journal of Advance Research in Computer Science and Software Engineering, Vol. 3, Issue 6, pp. 626-632, 2013.
[2] Ashish Thakur, Bushra Shaikh, Vinita Jain and A.M Magar 'Credit Card Fraud Detection Using Hidden Markov Model and Enhanced Security Features', International Journal of Engineering Sciences and Research Technology (IJERT), pp. 72-77, 2015.
[3] Akshata Hadkar and Sheetal Yewale 'Online Credit Card Fraud Detection', International Journal for Research in Engineering Application & Management (IJREAM), Vol. 1, Issue 2, 2015.
[4] Prajal Save, Pranali Tiwarekar, Ketan N. Jain and Neha Mahyavanshi, 'A Novel Idea for Credit Card Fraud Detection using Decision Tree', International Journal of Computer Applications, Vol. 161, Issue 13, 2017
[5] K.RamaKalyani and D.UmaDevi, 'Fraud Detection of Credit Card Payment System by Genetic Algorithm', International Journal of Scientific & Engineering Research, Vol. 3, Issue 7, 2012.
[6] Raghavendra Patidar, Lokesh Sharma, 'Credit Card Fraud Detection Using Neural Network', International Journal of Soft Computing and Engineering (IJSCE), Vol. 1, Issue-NCAI2011, 2011
[7] Gaurav Mhatre , Oshan Almeida , Dhiraj Mhatre and Poonam Joshi, 'Credit Card Fraud Detection Using Hidden Markov Model', International Journal of Computer Science and Information Technologies, Vol. 5, Issue 2, pp. 2053-2055, 2014.