Credit Card Fraud Detection using Hidden Morkov Model and Neural Networks

R.RAJAMANI
Assistant Professor, Department of Computer Science,
PSG College of Arts & Science, Coimbatore.
Email: rajamani_devadoss@yahoo.co.in
M.RATHIKA
Research Scholar, Department of Computer Science,
PSG College of Arts & Science, Coimbatore.
Email: rathi2109@gmail.com

----------ABSTRACT----------

With the emergence of internet and e-commerce, the use of credit card is an unavoidable one. The credit cards are used for purchasing goods and services. We can make both online and offline payment easily with the help of credit cards. For online transaction it uses virtual card and for offline transaction it uses physical card. In today's world, credit card provides cashless shopping at every shop. It will be the most convenient way to do online shopping. Hence, risks of credit card frauds are increasing day by day with its various techniques developed for detection. Fraud detection is a technique of identifying prohibited acts that are occurring around the world. The techniques of Data mining are also popular in detecting cyber credit-card fraud. An effective use of data mining techniques and its algorithms can be implemented to detect or predict fraud through Knowledge Discovery from unusual patterns from gathered data set. In this paper, we discussed about the various credit-card fraudsters techniques and also the detection methods for cyber credit card transactions. The goal of this paper is to provide a comprehensive review of Hidden Markov Model (HMM) and Neural Networks (NN) techniques to detect credit card fraudulent in an effective way.

Keywords - Credit-card fraudster's techniques, Data mining process, Fraud Detection Method, Hidden Markov Model, Neural Network.

I. INTRODUCTION

In our day-to-day life, online transactions are increased to purchase goods and services. The most common method of payment for online purchase is credit card. Around 60% of total transaction was carried out by using credit card. Credit card is most acceptable payment mode. Unfortunately, fraudulent use of credit cards has also become an attractive source of revenue for criminals. The credit card or credit card information is stolen by the unauthorized user and they used to make unauthorized purchases on e-commercial systems on the internet. The credit-card frauds are present-day happenings affecting many people and involving substantial monetary losses.

1.1 Types of Credit Card Frauds

Credit card fraud has been divided into two types:

- Offline fraud
- On-line fraud.

Offline fraud is committed by using a stolen physical card. On-line fraud is committed via internet, phone, shopping, web, or in absence of card holder.

II. CYBER CREDIT-CARD FRAUDSTERS TECHNIQUES

Using the technological advances, electronic databases containing credit card data may be hacked or crashed.

1. *Erasing the magnetic strip*

A fraudster can tamper an existing card that has been acquired illegally by erasing the metallic strip with a powerful electro-magnet.

2. *Creating a fake card*

A fraudster can create a fake card from scratch using sophisticated machines. This requires a lot of effort and skill to produce.

3. *Altering card details*

A fraudster can alter cards by either re-embossing them — by applying heat and pressure to the information originally embossed on the card

4. *Skimming*

Most cases of counterfeit fraud involve skimming, a process where genuine data on a card's magnetic stripe is electronically copied onto another.

5. *White plastic*

A white plastic is a card-size piece of plastic of any color that a fraudster creates and encodes with legitimate magnetic stripe data for illegal transactions.

6. *Site cloning*

Site cloning is where fraudsters clone an entire site or just the pages from which you place your order. The cloned or

spoofed site will receive these details and send the customer a receipt of the transaction via email just as the real company would.

7. *False merchant sites*

These sites often offer the customer an extremely cheap service. These sites are set up to accumulate as many credit card numbers as possible. The sites are usually part of a larger criminal network that either uses the details it collects to raise revenues or sells valid credit card details to small fraudsters.

8. *Credit card generators*

These are computer programs that generate valid credit card numbers and expiry dates. These generators work by generating lists of credit card account numbers from a single account number.

III. DATA MINING APPROACH

Data mining is popularly used to effectively detect fraud because of its efficiency in finding unknown patterns in a collected data set. Data mining is a technology that allows the discovery of knowledge in a dataset. Data is collected from different sources into a dataset and then we can discover patterns in the way all data in the dataset relates with another and then make predictions based on the patterns discovered. Data mining takes a dataset as an input and produces models or patterns as output.

Data mining refers to extracting the hidden, previously unknown and potentially useful information from database, and then offering the understandable knowledge, such as association rules, cluster patterns etc, so as to support users for decision-making. The fig1. represents the process of data mining.

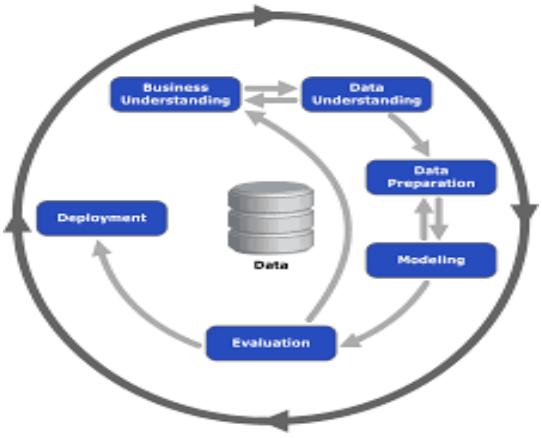

Fig1: Data mining process

IV. CREDIT CARD FRAUD DETECTION METHOD

The block diagram of Fig2 represents the overview to determine the given transaction is genuine transaction or the fraudulent transaction. If the transaction is genuine one, then it is successfully generated. Otherwise, the card holder is alerted.

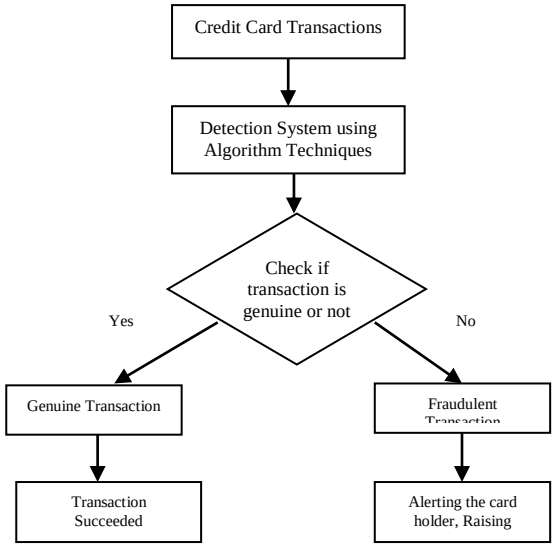

Fig2: A Flowchart representation of credit card transaction system

V. TECHNIQUES FOR DETECTING CREDIT CARD FRAUD

1. Hidden Markov Model (HMM)

Hidden Markov Model is the simplest models which can be used to model sequential data. In markov models, the state is directly visible to the observer but in a hidden markov model, the state is not directly visible, but output, dependent on the state, is visible. An HMM is a double embedded random process with two different levels, one is hidden and other is open to all. It is a finite set of states, each of which is associated with a probability distribution. Transitions among the states are governed by a set of probabilities called transition probabilities. In a particular state an outcome or observation can be generated, according to the associated probability distribution. It is only the outcome, not the state visible to an external observer and therefore states are "hidden" to the outside.

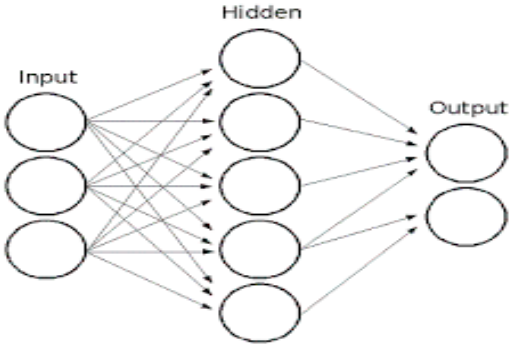

Fig3: Hidden Markov Model

The layers of the hidden markov model are represented in the fig3. Therefore the sequence of tokens generated by an HMM gives some information about the sequence of state. Hence, HMM has been successfully applied to many applications such as speech recognition, robotics, bio-informatics, data mining etc.

1.1 Working System Model

The working system has the two possibility case; it is represented in the fig4.

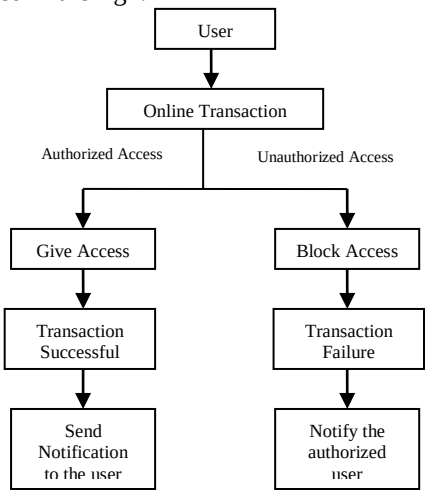

Fig4: System Model

Case 1: Valid User Access

If a user performs an online transaction then his spending profile is matched into our database and if it matches then the transaction is performed successfully and then user is notified that transaction is done successfully.

Case2: Invalid User Access

If an invalid user tries to perform an online transaction and if the spending profile doesn't matches into the database then access is blocked to that user and system failure occurs. HMM traces the IP address of the organization from where unauthorized user was trying to gain transaction and it also sends notification on authorized user's mobile number and raises the alarm to Admin System.

1.2 Application of Hidden Markov Model

A Hidden Markov Model is a finite set of states; each state is associated with a probability distribution. Transitions among these states are administered by a set of probabilities called transition probability. The implementation technique used in HMM is creating clusters of training sets so as to identify spending profile of card holder. The type of items purchased works as states for the model. The transition from one state to another is determined by probability distribution. It requires minimum 10 previous transactions, on the basis of which the fore coming transaction is chosen as fraud or genuine. The fig6 illustrates about the two phases of the detection system used by HMM. The model goes through two stages. In the first stage training of the system is done. Second stage works for the detection of the fraud, based on the expected range of amount the transaction. The expected amount and the actual amount for the next transaction are compared on the basis of probability distribution during training phase. If the deviation is above a threshold value then it is treated as fraud else legal. In case of fraud alarm is generated and transaction is terminated or else it is routinely accomplished.

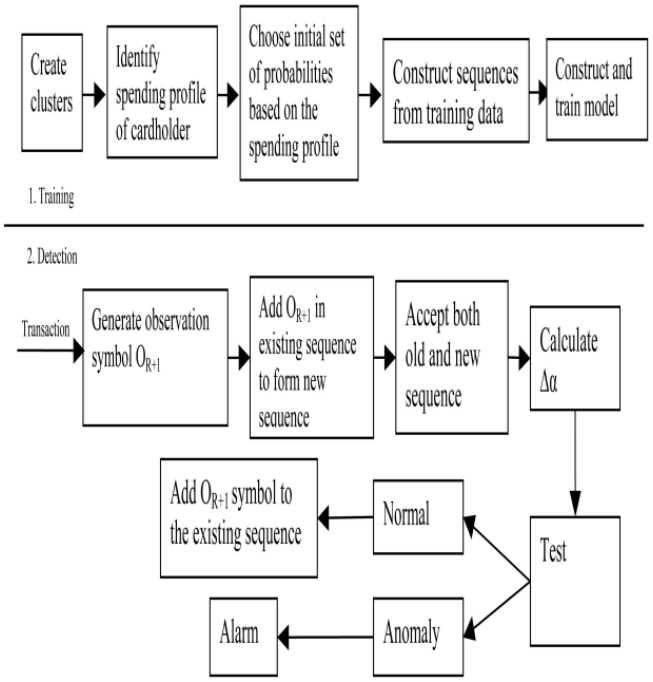

Fig5: Representation of HMM in Credit Card Transaction

At the initial stage, the model does not have data of last 10 transactions. So that the system will ask some basic information to the cardholder such as mobile number, address, date of birth etc. The HMM model generate relative data of transaction for further verification of transaction.

Advantages:

The benefit of the HMM-based approach is large reduction in the number of False Positives transactions acknowledged as malicious by a fraud detection system even though they are categorically genuine.

2. Neural Networks

Neural Networks (NN) are formed by organizing nodes into layers and associate these layers of neurons with modifiable weighted interconnections. A Neural Network (NN) is a collection of "processing nodes" transferring activity to each other via connections.

A Neural Network is a structure of many neurons connected in a regular way. Single node take input received from linked nodes and use the weights of the connected nodes together with easy function for computation of output values. Neural networks can be created for supervised and/or unsupervised learning.

2.1. Application of Neural Network Model as Credit Card Fraud Detection Method

There is a fixed pattern to how credit-card owners consume their credit-card on the internet. This fixed pattern can be drawn from legitimate regular activities of the credit-card owner for the past one or two years on its credit-card. Using the Neural Network technology, the computer-program or software can be trained with this fixed pattern to use it as knowledge in classifying a real-time transaction as fraudulent or legitimate transaction. Once the data to be analyzed is selected, the anomaly detection algorithms will be applied for matching the behavior of the current

transaction with its old transactions. If the patterns are matched the transaction will be carried out successfully.
If the patterns are differs for the past transactions, the system concludes that it is a unauthorized user access and it generates an alert message to the valid user. The attributes that are used to detecting fraudulent transaction was discussed below.

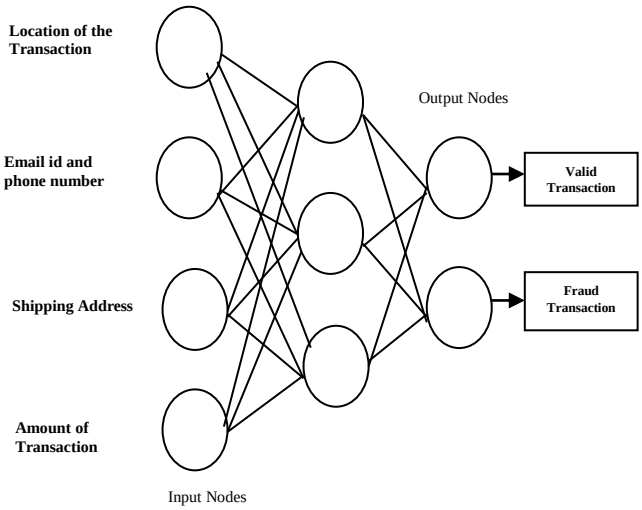

Fig6: Representation of Neural Network in Credit Card Transaction

Location of the transaction: In this process, the user's various online credit card transaction IP addresses are stored. This is a good mechanism to train Neural Networks for cyber credit-card fraud detection because in training Neural Networks with the City and Country locations formatted from IP-addresses where the credit-card owner has regularly made legitimate transactions from for the past one or two years, Neural Networks can know if the current transaction behaves in pattern like the past credit-card owner transactions.
Email address and Phone number: The email-address and phone number of the valid user is registered by the card issuer or the company. For each credit card made transaction the neural networks are trained to send a notification mail and message to the user's mail id and the telephone number, so that the individual can easily able to find out the fake transaction.
Shipping Address : In this data mining application, Neural Networks will be trained with Shipping addresses and overseas orders used by the credit-card owner in past one or two years transactions.
Amount of Transactions: Neural Networks will be trained with the cost range of goods and services purchased in the past one or two years transactions of the credit cardholder's credit card.
Advantages:

- Ability to classify untrained patterns
- Well-suited for continuous-valued inputs and outputs
- Successful on a wide array of real-world data
- Algorithms are inherently parallel
- Techniques have recently been developed for the extraction of rules from trained neural networks

Disadvantages:

- Difficulty to confirm the structure.
- Lack of Available data set.
- Require a number of parameters
- Poor interpretability

V. CONCLUSION

This paper provides an overview for the credit card fraud detection. Currently, Credit card risk monitoring system is one of the key tasks. There are many ways of detection of credit card fraud. If one of these HMM or NN or combination of both the algorithm is applied in credit card fraud detection system, the probability of fraud transactions can be predicted and prevented from the unauthorized user access. Also the further enhancement of these techniques may result into a better detection method. The future work on this can to be to make HMM and NN to more secure and covering other aspects of human behavior.

REFERENCES

[1] MohdAvesh Zubair Khan, Jabir Daud Pathan, Ali Haider Ekbal Ahmed "Credit Card Fraud Detection System Using Hidden Markov Model and K-Clustering" *International Journal of Advanced Research in Computer and Communication Engineering* Vol. 3, Issue 2, February 2014

[2] Renu, Suman, "Analysis on Credit Card Fraud Detection Methods" *International Journal of Computer Trends and Technology (IJCTT)* volume 8 number 1– Feb 2014

[3] Anita B. Desai et al, *International Journal of Computer Science and Information Technologies* (IJCSIT), Vol. 4 (1) 2013.

[4] https://www.law.cornell.edu/wex/credit_card_fraud.

[5] http://www.popcenter.org/problems/credit_card_ fraud/pdfs/bhatla.pdf.

[6] V. Bhusari1, S. Patil "Application of Hidden Markov Model in Credit Card Fraud Detection" *International Journal of Distributed and Parallel Systems (IJDPS)* Vol.2, No.6, November 2011.

[7] Rekha Bhowmik "Data Mining Techniques in Fraud Detection", *Journal of Digital Forensics, Security and Law*, Vol. 3(2).

[8] E. Kirkos et al. / Expert Systems with Applications 32 (2007).

[9] Anshul Singh, Devesh Narayan" A Survey on Hidden Markov Model for Credit Card Fraud Detection" *International Journal of Engineering and Advanced Technology (IJEAT)* ISSN: 2249 – 8958, Volume-1, Issue-3, February 2012.

[10] Sunil S Mhamane, L.M.R.J Lobo," Use of Hidden Markov Model as Internet Banking Fraud Detection" *International Journal of Computer Applications* (0975 – 8887) Volume 45– No.21, May 2012.

[11] Gogu.Sandeep,Sachin Malviya, Dheeraj Sapkale," Data Mining: An Improved Approach for Fraud

Detection" IJCSIT *International Journal of Computer Science and Information Technologies* 2013.

[12] Dheepa, Dhanapal "Behavior based credit card fraud detection using support vector machines", ICTACT journal on soft computing, july 2012.

[13] Taklikar et al., "Survey on Methods for Credit Card Fraud Detection Systems" *International Journal of Advanced Research in Computer Science and Software Engineering* 4(10), October – 2014.

[14] John Akhilomen," Data Mining Application for Cyber Credit-card Fraud Detection System" Proceedings of the World Congress on Engineering 2013.

[15] Khyati Choudhary and Bhawna Mallick," Exploration of Data mining techniques in Fraud Detection: Credit Card" *International Journal of Electronics and Computer Science Engineering* IJECSE,Volume1, No.3.

[16] Krishna Kumar, Mahesh A. Pavaskar, " Survey on Credit Card Fraud Detection Methods" *International Journal of Emerging Technology and Advanced Engineering* ISSN 2250-2459, Volume 2, Issue 11, November 2012.

[17] Sowjanya et al. ,"Application of Data Mining Techniques for Financial Accounting Fraud Detection Scheme" *International Journal of Advanced Research in Computer Science and Software Engineering* 3(11), November – 2013.

[18] Arun K.Pujari, "*Data Mining Techniques*" University Press 2001.

[19] Han, J. and M. Kamber, 2001. *Data Mining: Concepts and Techniques*. San Francisco, Morgan Kauffmann Publishers.