Volume 4, Issue 4, April 2014 ISSN: 2277 128X

International Journal of Advanced Research in Computer Science and Software Engineering

Research Paper

Available online at: www.ijarcsse.com

Credit Card Fraud Detection System Using Hidden Markov Model and K-Clustering

Jabir Daud Pathan
Department of Computer Engg,
Jaihind College of Engineering,Kuran
(University of Pune), India

MohdAvesh Zubair Khan
Department of Computer Engg,
Jaihind College of Engineering,Kuran
(University of Pune), India

Ali Haider Ekbal Ahmed
Department of Computer Engg,
Jaihind College of Engineering,Kuran
(University of Pune) , India

Abstract— In the era of 21st Century, there is a rapid increment in the electronic commerce technology for which the used of credit card has increased significantly. Most popular mode for online and offline payment is using credit card, use of credit card has dramatically increased. So as credit card is becoming popular mode for online financial transactions, at the same time fraud associated with it are also rising. Hence, this paper proposes a new approach for credit card fraud detection using Hidden Markov Model (HMM) and it also shows how fraud can be detected as well as survey of various techniques has been done. The method work on the stochastic behaviour of user's transaction. In existing fraud detection system fraud is detected after the transaction is done. The propose paper shows how frauds can be detected using Hidden Markov Model (HMM) during transaction.

Keywords: Fraud detection, Fraud Detection Techniques, Hidden Markov Model (HMM), K-Means Clustering Algorithm, One Time Password (OTP).

I. INTRODUCTION

In today's electronic world e-commerce has become an essential mode for global business. Electronic commerce, commonly known as e-commerce or eCommerce, is a type of industry where the buying and selling of products or services is conducted over electronic systems such as the Internet and other computer networks. According to Nielsen study conducted in 2008, $1/10^{th}$ world's total population has been using internet for shopping and transaction. The most common method of payment for online purchase is credit card. As number of credit card user's increases daily there is rhythm in the people's life and the same time the credit card fraud ratio is also increases. For which crimes involving in credit card are increasing that disturbs the organization financial order seriously and hence there is a great loss to bank and card holder that affects the development of banks.

Credit card can be used to purchases goods and services using online and offline transaction mode. It can be divided into two types :

A. Physical Card
B. Virtual Card

In the physical card based purchase, card holder has to produce the card at the merchant counter and merchant will sweep the card in the EMV (Europay, MasterCard and Visa) machine. Fraud transaction can be happened in this mode, only after the card has been stolen. It will be difficult to detect fraud in this type of transaction. If the card holder does not realize loss of the card and does not report to police or card issuing company, it can give financial loses to issuing authorities. In the second method of purchasing i.e. online, these transactions generally happen on telephone or internet and to make this kind of transaction, the user will need some important information about a credit card (such as credit card number, validity, CVV number, name of card holder). To make fraud transaction to purchase goods and services, fraudster will need to know all these details of card only then he/she will make transactions. Most of the time, the cardholder may or may not know that when or where any person will be seen or stolen card information. To detect this kind of fraud transaction, we have proposed a Hidden Markov Model which is studying spending profile of the card holder. An HMM is to analyse the spending profile of each card holder and to find out any discrepancy in the spending patterns. Fraud detection can be detected on analysing of previous transactions data which helps to form spending profile of the card holder. Every card holder having unique pattern contains information about amount of transactions, details of purchased items, merchant information, date of transaction etc. It will be the most effective method to counter fraud transaction through internet. If any deviation will be noticed from available patterns of the card holder, then it will generate an alarm to the system to stop the transaction.

II. LITERATURE SURVEY

Abhinav Srivastava et al describe the "Credit card fraud detection method by using Hidden Markov Model (HMM)"[1]. In this paper, they model the sequence of operations in credit card transaction processing using a Hidden Markov Model (HMM) and show how it can be used for the detection of frauds. An HMM is initially trained with the normal behaviour of a cardholder.

S.Ghosh and Douglas L.Reilly et al describes the "Credit card fraud detection With Neural Network"[2]. In this paper they using data from a credit card issuer, a neural network based fraud detection system was trained on a large sample of labelled credit card account transactions and tested on a holdout data set that consisted of all account activity over a subsequent two-month period of time. The neural network was trained on examples of fraud due to lost cards, stolen cards, application fraud, counterfeit fraud, mail-order fraud and NRI (non-received issue) fraud. The network detected significantly more fraud accounts (an order of magnitude more) with significantly fewer false positives (reduced by a factor of 20) over rule based fraud detection procedures.

Sunil S Mhamane et al describes the "Use of Hidden Markov Model as Internet Banking Fraud Detection"[3]. In this paper they explained about how Fraud is detected using Hidden Markov Model also care has been taken to prevent genuine Transaction should not be rejected by making use of one time password which is generated by server and sent to Personal Mobile of Customer.

Pankaj Richhariya describes "A Survey on Financial Fraud Detection Methodologies"[4]. The paper details as follows. Owing to levitate and rapid escalation of ECommerce, cases of financial fraud allied with it are also intensifying and which results in trouncing of billions of dollars worldwide each year.

III. VARIOUS FRAUD TECHNIQUES

There are various ways that fraudsters execute an online fraud. By using various technologies they can do fraudulent activities.

A. Identity Theft

The most widely defined online banking fraud is the identity theft, gets the most attention from the customers. Identity theft can be difficult to find victims. To predict the theft, take months or even years to correct the fault. It is one of the easiest methods to get the card holders information.

B. Phishing

It is one of the threats involves using bogus emails or websites. The word "phishing" means from combining the words "password" and "fishing". Fraudulent send emails that appear to be from the customer's bank that direct customers to a fake website. This website impersonates the bank's website and prompts customers for their account access data.

C. Trojan horse

It is one of the computer virus type software program stored on the customer's PC. Trojans records the keyboard driver and keystrokes. Once a Trojan detects that the customer opens an online banking website, it captures login name and password, and sends it to the criminal.

D. Internal Fraud

Banking sector allows their employees to access customer data. The data is the same information needed to access online banking to customer accounts. So that an employee can easily commit fraud. Instead of this, financial institutions should require a password or PIN for net banking, and the password or PIN should be stored in the format of encrypted.

IV. FRAUD DETECTION ARCHITECTURE

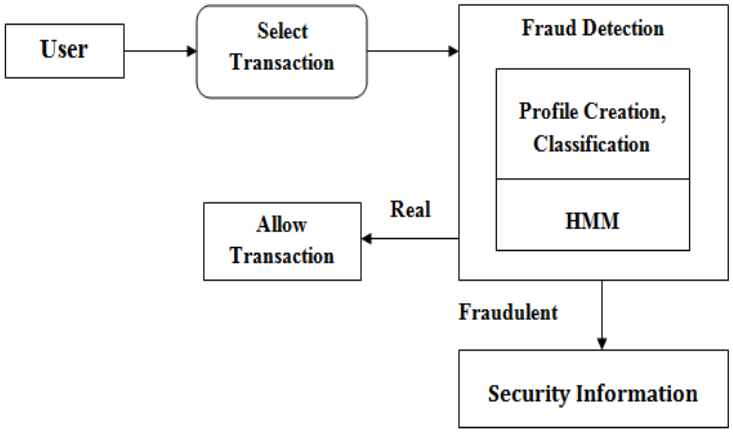

Fig. 1 Fraud Detection Architecture

Fig.1 Shows the Fraud Detection Architecture, in this user performs an online transaction then it goes to the Fraud Detection System (FDS). In FDS users spending profile is checked with database and also HMM algorithm runs on user previous transactions. If user is authenticated user then FDS allow transaction or if user is unauthenticated user then FDS detects that transaction is fraudulent then it goes to the security system where 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.

V. WORKING OF SYSTEM

A. *Authorized User*

In Fig 2, If an authorized 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.

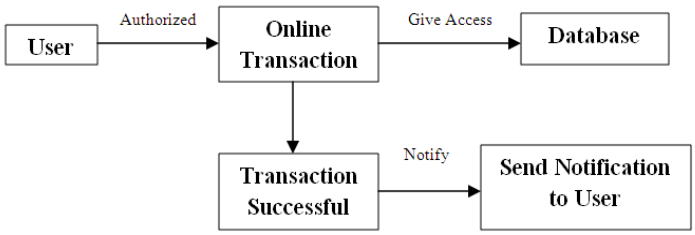

Fig. 2 Authorized User Access To System

B. *Unauthorized User*

In Fig 3, If an unauthorized 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.

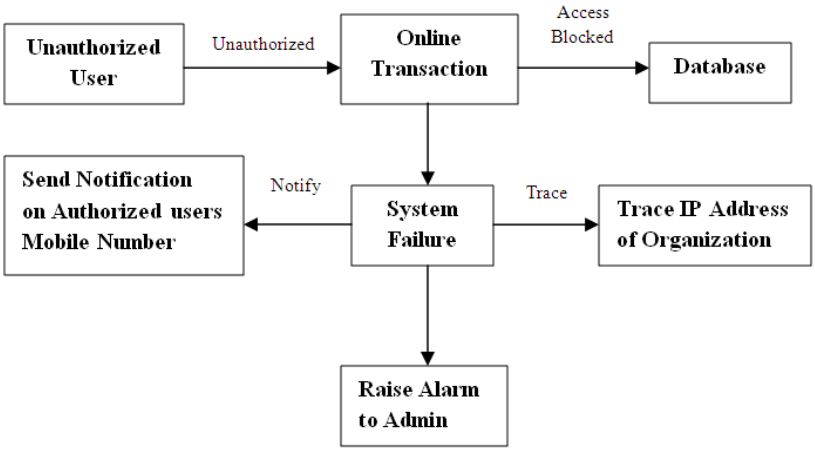

Fig. 3 Unauthorized User Access To System

VI. USE OF HMM, K-CLUSTERING AND OTP FOR CREDIT CARD FRAUD DETECTION SYSTEM

A. *Hidden Markov Model (HMM) Background*

An HMM is a double embedded stochastic process much more complicated stochastic processes as compared to a traditional Markov model in fig 4. The H.M.M. uses the Price range: High, Medium, Low as Prediction.
The diagram below shows the general architecture of an instantiated HMM with two hierarchy levels.

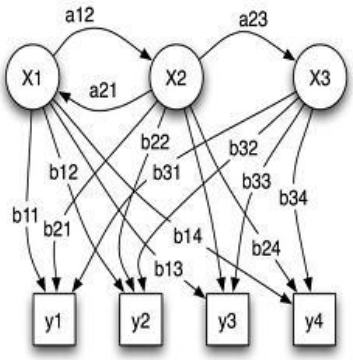

Fig. 4 Architecture Of HMM

To record the credit card transaction dispensation process in conditions of a Hidden Markov Model (HMM), it creates through original deciding the inspection symbols in our representation. We quantize the purchase values x into M price ranges V1, V2 . . . VM, form the study symbols by the side of the issuing bank. The genuine price variety for each

symbol is configurable based on the expenditure routine of personal cardholders. HMM determine these prices range. Dynamically by using clustering algorithm on the price values of every card holder transactions. It uses cluster Vk for clustering algorithm as k ¼ 1, 2. . . . M, which can be represented both observations on price value symbols as well as on price value range.

In this prediction process it considers mainly three price value ranges such as low (l) Medium (m) and High (h). So set of this model prediction symbols is V { l, m, h}. E.g. If card holder perform a transaction as $ 250 and card holders profile groups as l (low) = [0, $ 100], m (medium) = [$ 200, $ 500], and h (high) = [$ 500, up to credit card limit], then transaction which card holder want to do will come in medium profile group. So the corresponding profile group or symbol is M and V (2) will be used. In various period of time, purchase of various types with the different amount would make by credit card holder. It uses the deviation in a purchasing amount of latest 10 transaction sequence (and adding one new transaction in that sequence) which is one of the possibilities related to the probability calculation. In initial stage, model does not have data of last 10 transactions, in that case, model will ask to the cardholder to feed basic information during transaction about the cardholder such as mother name, place of birth, mailing address, email id etc. Due to feeding of information, HMM model acquired relative data of transaction for further verification of transaction on spending profile of cardholder.

B. K-Means Clustering

K-Clustering algorithm used to determine the clusters. K-means is an unsupervised learning algorithm for grouping a given set of data based on the similarity in their attribute values.

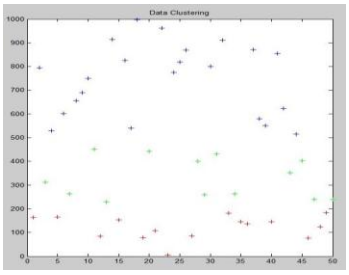

Fig. 5 Data Clustering

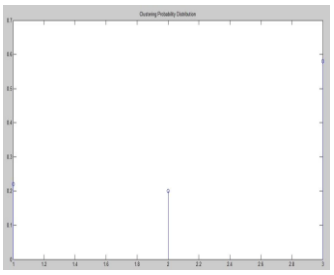

Fig. 6 Clustering Probability

Fig.5 shows three clusters. Transactions in red forms low spending group, transactions in green form medium spending group, and transactions in blue form high spending group. These groups are observation symbols in our implementation. Fig 6 indicates that clustering probability of each observation symbol. In this Fig.6 clustering probability of high spending is highest among three. It can be said that spending profile of given cardholder is high spending.

C. One Time Password (OTP)

A one-time password (OTP) is a password that is valid for only one login session or transaction. OTPs avoid a number of shortcomings that are associated with traditional (static) passwords. The most important shortcoming that is addressed by OTPs is that, in contrast to static passwords, they are not vulnerable to replay attacks. This means that a potential intruder who manages to record an OTP that was already used to log into a service or to conduct a transaction will not be able to abuse it, since it will be no longer valid. On the downside, OTPs are difficult for human beings to memorize. Therefore they require additional technology to work.

VII. CONCLUSION

In our paper we used an HMM in detection of credit card fraud. We modeled the sequence of transactions in credit card processing using an HMM. We have used clusters that are generated by using k-means clustering algorithm as our observation symbols. In our implementation we took three observation symbol which are spending ranges of cardholder that are low, medium, and high, whereas the type of item have been considered to be states of an HMM. An HMM is trained with Baum-Welch algorithm for each cardholder. It has been also explained that how an HMM can detect whether the incoming transaction is fraudulent or not.

REFERENCES

[1] Srivastava, Abhinav, Kundu, Amlan, Sural, Shamik and Majumdar, Arun K., (2008) "Credit Card Fraud Detection Using Hidden Markov Model", IEEE Transactions on Dependable and Secure Computing, Vol. 5, No. 1, pp. 37-48.

[2] S. Ghosh and D.L. Reilly, "Credit Card Fraud Detection with a Neural-Network," Proc. 27th Hawaii Int'l Conf. System Sciences: Information Systems: Decision Support and Knowledge Based Systems, vol. 3, pp. 621-630, 1994.

[3] Sunil S Mhamane and 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.

[4] Pankaj Richhariya et al "A Survey on Financial Fraud Detection Methodologies" BITS, Bhopal," International Journal of Computer Applications (0975 – 8887) Volume 45 No.22, May 2012.

[5] "Credit Card Fraud Detection Using Hidden Markov Model ", Shailesh S. Dhok (2012).

[6] A Survey on Hidden Markov Model for Credit Card Fraud Detection Anshul Singh, Devesh Narayan. Sung-Bae Cho and Hyuk-Jang Park ,"Efficient anomaly detection by modeling privilege flows using hidden Markov model" Department of Computer Science, Yonsei University,134 Shinchon-dong, Sudaemoon-ku, Seoul 120-749, Korea.

[7] Vaibhav Gade, Sonal Chaudhari," Credit Card Fraud Detection Using Hidden Markov Model" International Journal of Emerging Technology and Advanced Engineering (Volume 2, Issue 7, July 2012).

[8] Nitin Mishra, Ranjit Kumar, Shishir Kumar Shandilya "Credit Card Transaction Fraud Detection by using Hidden Markov Model"(2012).

BIOGRAPHIES

Mr. Jabir Daud Pathan is currently pursuing B.E. Degree in Computer Engineering from Jaihind College of Engineering, Kuran (University of Pune).

Mr. MohdAvesh Zubair Khan is currently pursuing B.E. Degree in Computer Engineering from Jaihind College of Engineering, Kuran (University of Pune).

Mr. Ali Haider Ekbal Ahmed is currently pursuing B.E. Degree in Computer Engineering from Jaihind College of Engineering, Kuran (University of Pune).

