Study of Fraudulent Detection System Using Hidden Markov Model

Dr. Suhas Patil
Department of Computer Engineering
Bharati Vidyapeeth (Deemed to be University) College of Engineering, Pune, India

Shivam Srivastav
Department of Computer Engineering
Bharati Vidyapeeth (Deemed to be University) College of Engineering, Pune, India

Pratap Singh
Department of Computer Engineering
Bharati Vidyapeeth (Deemed to be University) College of Engineering, Pune, India

Naman Sharma
Department of Computer Engineering
Bharati Vidyapeeth (Deemed to be University) College of Engineering, Pune, India

Abstract

In light of recent events, including the Coronavirus pandemic, the most accepted mode of shopping and purchasing goods and services is online through credit card. It provides cashless shopping at every possible shop in all countries and will be the most easy way to do online shopping, paying bills, requesting services etc. for years to come. With that said, credit card frauds are increasing day by day as well, regardless of the various techniques developed for its detection. Fraudsters are so expert that they engineer new ways for committing fraudulent transactions everyday which demands for constant innovation for its detection techniques also. Many techniques based on Artificial Intelligence, Data mining, Fuzzy logic, Machine learning, Sequence Alignment, decision tree, neural network, logistic regression, naïve Bayesian, Bayesian network, meta-learning, Genetic Programming etc., has evolved in detecting various credit card fraudulent transactions. In this paper, Hidden Markov Model (HMM) is employed to model the sequence of operation in credit card transaction operations. It helps to get a high fraud coverage combined with an exceptionally low false alarm rate.

Keywords: Hidden Markov Model, HMM, fraud transaction, credit card, credit card frauds, OTP

I. Introduction

The internet always had being huge impact on human life. Humans interact with the internet for various reasons, from daily chores, keeping records to performing complex scientific studies. Also, a sudden increase in the Internet of Things has led to a rise in internet usage.

Throughout these years, online transactions have been increased to purchase goods and services. According to a Nielsen study conducted in 2007-2008, 28 of the world’s total population has been using internet, 85 of these people has used the internet to make online shopping and the rate of processing online purchases has increased by 40 from 2005 to 2008. The most contemporary method of payment for online purchase is through a credit card. Around 60 of total transactions were completed by using credit card. Whether it is developed or developing country to some extent, credit card is most used payment mode for online or/and offline transaction.

As usage of credit cards increases worldwide, chances of an attacker stealing credit card details and making a fraud transaction are also increasing. In 2008, the number of frauds committed through credit cards have increased by 30% because of various ambiguities in issuing of and managing credit cards. Credit card frauds are approximately 1.2% of the total transaction amount, although it is not a small amount as compared to the total transaction amount which is in trillions of dollars in 2007[1- 3].

Credit cards can be used to purchase goods and services using online and offline transaction modes. These can be classified into two cards:

1) Physical Card
2) Virtual Card

Hidden Markov Model is going to be helpful to seek out the fraudulent transaction by using spendings of the user. It based on the user spending profiles, which can be divided into major three types such as: 1) Low profile; 2) Middle profile; and 3) High profile. For every card, the spending profile is different, so it can find out an inconsistency of user profile and check out to seek out fraudulent transaction. It keeps a record of purchasing profiles of the cardholders by offline or/and online. Thus, a proper study of purchased commodities of cardholders are going to be a valuable tool in fraud detection systems, and it’s an assuring path to check fraudulent transactions. Although, fraud detection system doesn’t keep records of a number of purchased goods and commodities. Every user is represented by a specific patter of set which contains information for last 10 purchases using credit card. If any deviation will be observed from documented patterns of the cardholders, then it will trigger an alarm to the system to stop the transaction immediately.

II. RELATED WORK

A. Srivastava et al describe the “Credit card fraud detection method by using Hidden Markov Model (HMM)”, [4]. In this research paper, they model the sequence of functioning in credit card transaction processing using a Hidden Markov Model (HMM), how it can be used for the detection of frauds. An HMM is initially trained with the usual transactions of a cardholder.

P. Richhariya describes “A Survey on Financial Fraud Detection Methodologies”,[5]. The research paper details as follows. Owing to rise and rapid upsurge of e-commerce, cases of fraud in alliance with it are also escalating and which results in trouncing of billions of dollars worldwide each year.

S. S. Mhamane et al describes the “Use of Hidden Markov Model as Internet Banking Fraud Detection”, [6]. In this research paper, we can that how different fraud is discovered using Hidden Markov Model. And also care has been taken to prevent genuine transactions that should not be rejected by making use of a one-time password (OTP), which is generated by the server and sent to the personal mobile number of the customer.

III. HIDDEN MARKOV MODEL

A Hidden Markov Model is always a finite set of states where each of the states are linked with a probability dealing out. These Transitions states are governed by a set of probabilities called transitional probabilities. In any particular state, a possible outcome or observation can be generated which is an associated symbol of observation of probability distribution. It is only the result, not the state that’s visible to an external observer and are therefore ``hidden" to the outside; hence the name Hidden Markov Model [7-9].

In the proposed model based on HMM, the system will help to verify fraudulent transactions from legitimate transactions and will alert the holder when the fraud has taken place. It includes two modules that are as follows:

1) Online Shopping: It comprises of many steps; First is to login to a particular site to purchase goods or services, then choose an item and go to payment mode where credit card information will be required. After filling all this information, the page will be directed to the proposed fraud detection system which will be installed at the bank's server or merchant’s site.
2) Fraud Detection System: All the information about credit card (Like Credit card number, credit card CVV number, credit card expiry, name of cardholder etc.) will be checked with credit card database. If User entered database is correct, it will ask Personal Identity number (PIN). After verifying the Personal Identity Number (PIN) with database and account balance of the user's credit card is sufficient for the purchase amount, the fraud checking module will be activated.

IV. MODEL WORKING

Hidden Markov Model can be used in different areas such as bio-information, robotics, speech recognition, data mining, voice recognition, artificial intelligence, etc.

1) HMM is defined as [10-12]:
2) N is the no. of states in the model.
3) At any instant k the state is denoted by qk.
4) R is the number of distinct observations symbols and are symbolized as Q= Q1, Q2 … QR.
5) State transition probability for matrix A= [aij] where aji =Pr (qi+1 at t+1 | qi at k).
6) The observation symbol probability matrix B= [bj (m)], where bj (m) = Pr (vm at t | qi at t).
7) πi =Pr (qi at t=1).

HMM model involves two stages:

1) Training stage
2) Detection Stage

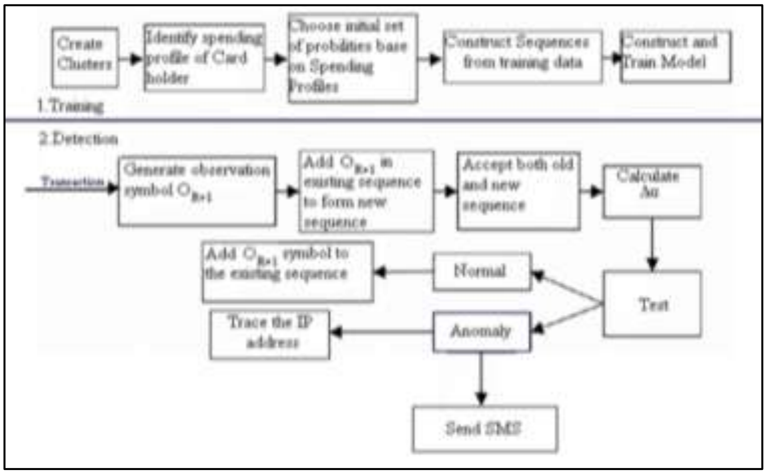

Create Clusters
Identify spending profile of Card holder
Choose initial set of probabities base on Spending Profiles
Construct Sequences from training data
Construct and Train Model
1.Training
2.Detection
Transaction
Generate observation symbol O_{R+1}
Add O_{R+1} in existing sequence to form new sequence
Accept both old and new sequence
Calculate Aα
Add O_{R+1} symbol to the existing sequence
Normal
Test
Trace the IP address
Anomaly
Send SMS

Fig. 1: Flow Model of Credit Card Fraud Detection System

V. Workings of Proposed System

This section entails the workings of a functioning system to it's potential: -
If an authorized user performs an online transaction, their spending profile is matched to the database. If it comes back positive, the transaction is performed successfully, and the user is notified that the transaction was successful.

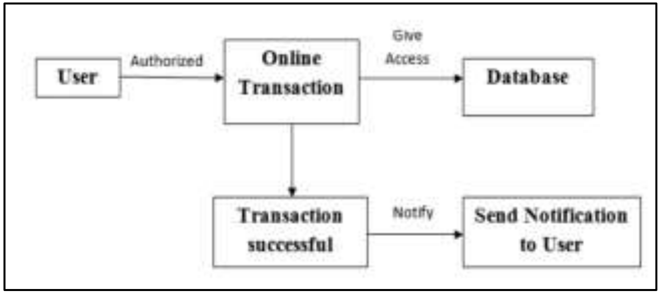


Fig. 2: Authorized Users Access

Similarly, if an unauthorized user tries to perform an online transaction and their spending profile does not match the one in the database, access is blocked to the user and transaction failure occurs. Hidden Markov Model traces the IP address of the organization from where the unauthorized user was trying to make the transaction, sends notification on authorized user's mobile number and raises the alarm to Admin system.

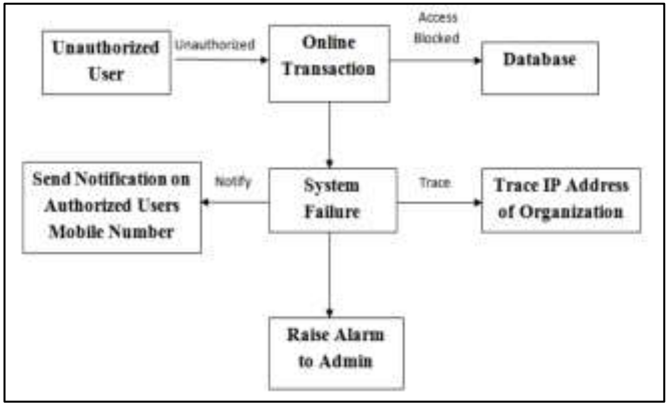


Fig. 3: Unauthorized Users Access

VI. Different Techniques and Algorithm:-

To track record of the credit card transaction of the users exemption process of a Hidden Markov Model (HMM), it creates through initial deciding the inspection symbols in our representation. We identity the personal identification number with database and account balance of user's credit card is more than the new transaction amount, the fraud checking process will be activated immediately.

The verification of all data will be cross checked before the first page load of credit card fraud detection system.

If a user credit card has less than 10 transactions, then it will directly ask to provide personal information to do the transaction. Once the database of 10 transactions will be developed, then the fraud detection system will activated.

By using current observation, determine users spending profile. The purchase amount will be checked with the spending profile of the user. By transition probabilistic calculation based on Hidden Markov Model, it concludes whether the transaction is real or fraud. If transaction may be concluded as a fraudulent transaction then user must enter security information such as personal info etc. This information is about with credit card (like account number, security question and answer which are provided at the time of registration). If the transaction will not be fraud, then it will directly give permission for the current transaction.

If the detected transaction is fraud, then the Security information form will arise. It has a set of questions where the user has to answer correctly to complete the transaction successfully. These forms have information of the user such as personal, professional, address; dates of birth, etc. are stored in the databases. If the information user entered will be matched with database information, then the following transaction will be completed securely. Or else transaction will be terminated and transferred to online shopping website. The flowchart of the proposed module is shown in Figure 4.

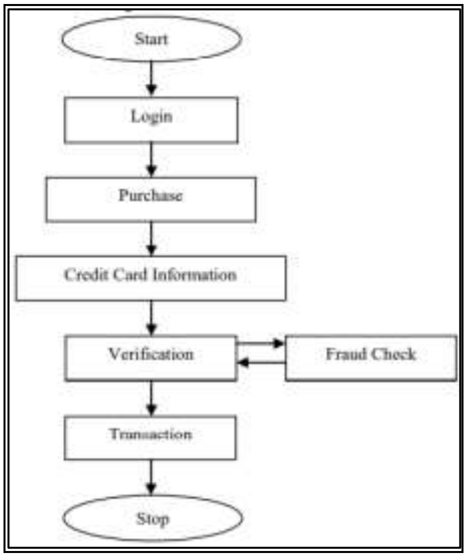
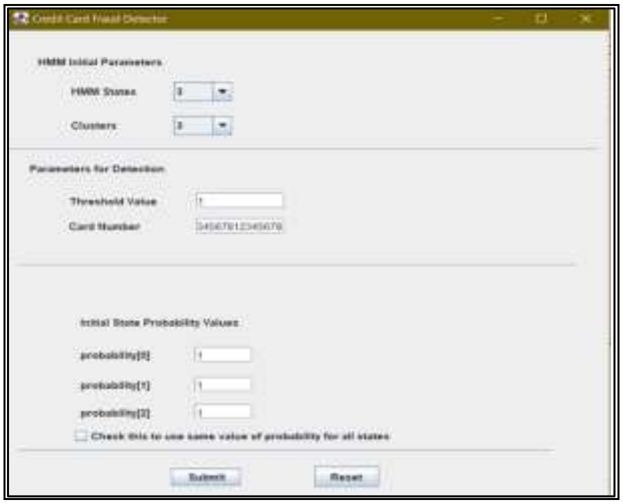


Fig. 4: Flowchart of HMM module for credit card fraudulent detection.

VII. RESULT

After Finally Deploying Our Code in real-time we see that our system can successfully detect frauds done through the customers credit card as shown in the figure below:



HMM Initial Parameters

HMM States: 3

Clusters: 3

Parameters for Detection

Threshold Value: 1

Card Number: 345678123456789

Initial State Probability Values

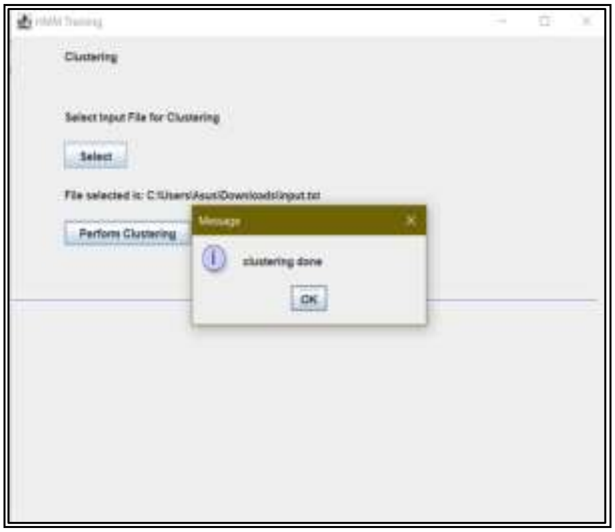
probability[0]: 1

probability[1]: 1

probability[2]: 1

□ Check this to use same value of probability for all states

Fig 5: Entering the credit card details



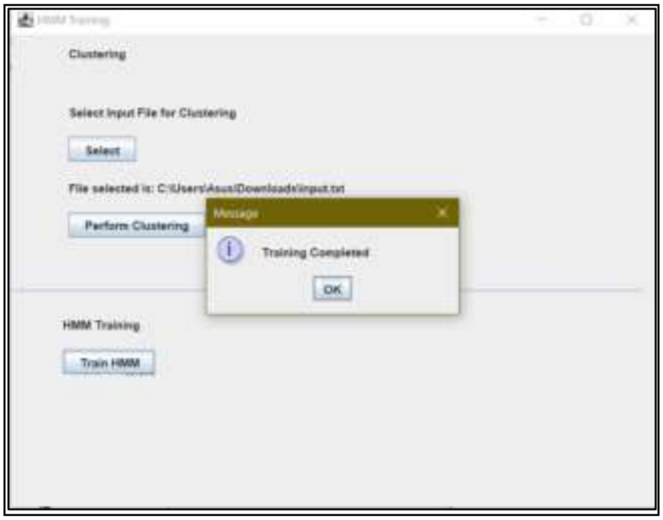
Clustering

Select Input File for Clustering

File selected is: C:\Users\Asus\Downloads\input.txt

clustering done

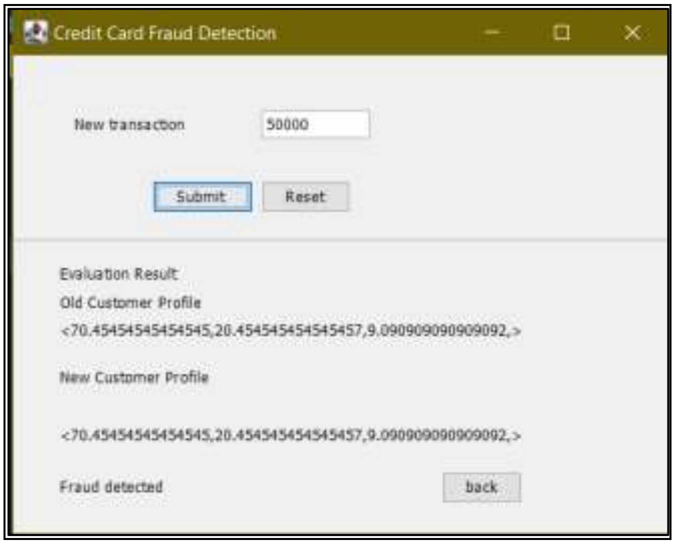
Fig. 6: Performing Clustering on the past transactions.



File selected is: C:\Users\Asus\Downloads\input.txt

Training Completed

Fig. 7: Training Hidden Markov Model



New transaction 50000

Evaluation Result
Old Customer Profile
<70.45454545454545,20.454545454545457,9.090909090909092,>

New Customer Profile

<70.45454545454545,20.454545454545457,9.090909090909092,>

Fraud detected

Fig. 8: Final Result (Checking the new transaction is fraud or not)

If the new transaction does not matches with the predicted results, then the system tells that the new transaction is a fraud, as shown in figure 8 above.

VIII. Conclusion

In this paper, we discussed how important it is to keep increasing the difficulty for committing fraud in today's world to discourage people. Moreover, The Fraud detection systems need to keep up with the techniques that are used to perform such actions so it can be detected and countered. The Hidden Markov Model makes this process of detection much easier and tries to remove the complexity of the task.

In this proposed model, we have found out more than 85% of transactions are genuine, and exceptionally low false alarms occur (Approximately 6% of the total number of transactions). The relative studies and our results show that the accuracy and effectiveness of the proposed system is secure up to 82% over a very broad deviation in the input data.

References

[1] Federal Trade Commission, 2009. Consumer sentinel network data book.
[2] Statistics for General and On-Line Card Fraud, March 2007.
[3] Global Consumer Attitude towards On-Line Shopping, March 2007.
[4] 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.
[5] 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.
[6] 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.

[7] Syeda, M., Zhang, Y. Q., and Pan, Y., 2002 Parallel Granular Networks for Fast Credit Card Fraud Detection, Proceedings of IEEE International Conference on Fuzzy Systems, pp. 572-577 (2002).

[8] Stolfo, S. J., Fan, D. W., Lee, W., Prodromidis, A., and Chan, P. K., 2000. Cost-Based Modeling for Fraud and Intrusion Detection: Results from the JAM Project, Proceedings of DARPA Information Survivability Conference and Exposition, vol. 2 (2000), pp. 130-144.

[9] Aleskerov, E., Freisleben, B., and Rao, B., 1997. CARD WATCH: A Neural Network Based Database Mining System for Credit Card Fraud Detection, Proceedings of IEEE/IAFE: Computational Intelligence for Financial Eng. (1997), pp. 220-226.

[10] Priyanka Gupta, Ankit Mundra, Online In-Auction Fraud Detection Using Online Hybrid Model, International Conference on Computing, Communication and Automation (2015).

[11] Rabiner, Lawrence, Being-Hwang Juang, An introduction to hidden Markov models, ASSP Magazine 3(1) (1986), 4-16.

[12] Rabiner L.R., A tutorial on hidden Markov models and selected applications in speech recognition, Proceedings of the IEEE 77(2) (1989), 257-286.

[13] V. Bhusari, S. Patil " Study of Hidden Markov Model in Credit Card Fraudulent Detection" International Journal of Computer Applications (0975 – 8887) Volume 20– No.5, April 2011