# EFFECTIVE STATISTICAL APPROACH FOR DETECTION OF CREDIT CARD FRAUDS USING HIDDEN MARKOV MODEL

**M. Sathyapriya,**
Assistant Professor,
Department of Computer Science,
Gobi Arts and Science College (Autonomous),
Gobichettipalayam,Tamilnadu,India.

**K.A. Poornima,**
Assistant Professor,
Department of Computer Science,
Gobi Arts and Science College (Autonomous)
Gobichettipalayam,Tamilnadu,India.

**Abstract:** In today's world, the most accepted payment mode is credit card for online transactions which provides cashless shopping at every corner across the world. It is the most suitable way to do online shopping, paying bills, and performing other related tasks. Hence risk of fraud transactions using credit card has also been increasing. In the prevailing credit card fraud detection processing system, fraudulent transaction will be detected after transaction is done. Hidden Markov Model is one of the statistical tools for engineers and scientists to solve various problems. Credit card frauds can be detected using hidden markov model during online transactions. Hidden markov model aids to obtain a high fraud transaction coverage combined with low false alarm rate, thus providing a better and convenient way to detect frauds. Using hidden markov model, customer's pattern is analysed and any deviation from the regular pattern is considered to be a fraudulent transaction. So, hidden markov model is initially trained with the normal behaviour of a cardholder. If an incoming online card transaction is not accepted by the trained HMM with high probability, it is considered to be fraudulent. At the same time, the algorithm tries to ensure that genuine transactions are not rejected. In this paper, the sequence of operations in online card transaction processing is modelled using Hidden Markov Model (HMM) to detect fraudulent transactions.

***Keywords: Hidden Markov Model, Credit Card, Online transaction***

## I.INTRODUCTION

In this cashless era, the most universal means by which goods and services are paid for is the use of credit card. Credit Card Fraud is described as a situation where a person uses a credit card belonging to another person for personal motives without the permission or awareness of the card-owner [2]. The Credit Card is a plastic card issued to number of users as a mode of payment [3]. There are several issues associated with online credit card use. One of the most important is fraud, which can be carried out by both individuals and merchants. A major problem is the lack of security which could lead to credit card numbers in online databases being compromised [4].

There are two main types of credit cards used for making purchases:

- **Physical card:** Here, there is a physical presentation of the card, for payment of goods bought or services rendered, by the card holder to the merchant. Fraudulent activities can be carried out in this type of purchase; in fact it is relatively easy, as the attacker only needs to steal the card. This could lead to significant financial loss to the credit card company if the theft is not realized quickly.
- **Virtual card:** This type of purchase is made using vital details about the credit card such as the card number, expiration date, secure code and Card Verification Value (CVV) number. Such purchases are normally done on the Internet. To commit fraud in these types of purchases, a fraudster needs only to know the card details [2].

Even as this rate is rapidly increasing, the means and ways by which individuals try to defraud and steal from other people is also increasing. Most fraudsters target credit card because a lot of money can be made within a short period of time and also without too many risks. Another "upside" of this fraud is that the actual crime is discovered many days after it occurs [7].

Credit card fraud increases at an almost directly proportional rate as the number of daily online and physical card users [8]. Nowadays, retailers deal more with online than regular purchases, most of which are credit card based transactions.

## II. HIDDEN MARKOV MODEL

A Hidden Markov Model (HMM) is a finite set of states; where a probability distribution is linked to each state. A set of probabilities called the transition probabilities direct the transition among these states. In each state an outcome can be generated, this outcome is an associated symbol of observation of the probability distribution. The only visible outcome to the external observer is the outcome at the current state, while all other states are "hidden"; hence the name "Hidden Markov Model" [12].

A Hidden Markov Model (HMM) can be considered as a generalization of a mixture model where the hidden variables (or latent variables), which control the mixture component to be selected for each observation, are related through a Markov process rather than independent of each other [15].

Mathematically, a HMM is described as having the following characteristics:

1. N is the number of states in the model. We can denote the set of state as $S = \{S_1,S_2,\ldots,S_n\}$. The state at time instant t is

denoted by qt.
2. M is the number of distinct observation symbols per state. The observation symbols correspond to the physical output of the system being modelled.
3. The state transition probability matrix A= [ $A_{ij}$].
4. The observation symbol probability matrix B = [$B_{jk}$].
5. The observation sequence O = $O_1$, $O_2$ ....$O_n$.
6. N is the number of hidden states.

It is apparent that a complete specification of an HMM requires the estimation of two model parameters, N and M, and three probability distributions A, B, and p. We use the notation (A, B, p) to indicate the complete set of parameters of model, where A, B implicitly include N & M [17].

Initially the normal behaviour of the cardholder is used to train the HMM, then spending patterns of users can be determined using K-means clustering algorithm. Any incoming transaction that is not accepted by the HMM can be determined as suspicious. For further confirmation, a security question module that contains some personal questions that are expected to only be known to the authorized cardholder will be activated and if the transaction is fraudulent then a verification code is requested [18]. Hidden Markov model works on the Markov chain property in which the probability of each subsequent state depends on the previous state, which consists of observation probabilities, transition probabilities and initial probabilities. HMMs are commonly applied to pattern recognition tasks since they allow a formal representation of a stochastic dynamic process, and allow for a systematic analysis of the data and prediction based on such models [16].

## III.REVIEW OF LITERATURE

Abhinav Srivastava et al [1] the author uses the ranges of transaction amount as an attribute in the HMM. The author has suggested method for finding the spending profile of cardholders. It is also discussed how the HMM can identify the fraudulent transactions. The simulation results show the advantages of using HMM and learning the profile of the cardholder plays an important role in analyzing fraudulent cases. The result also shows that 80% of the results are accurate and the system is scalable for large data set as well.
Divya.Iyer et al [6] the author uses Hidden Markov Model (HMM) to detect credit card transaction frauds. The training set is tuned with the normal behaviour of the card holder. So if credit card transaction is rejected by the trained HMM then that transaction is said to be fraudulent. Care is to be taken that valid and genuine transactions are not considered as fraud. The author also compares various methods with the proposed methods to prove that HMM are much preferred than the other methods.
K.RamaKalyani et al [14] creates a test data and through which the fraudulent activities are detected. This algorithm is also called as an optimization technique based on genetic and natural selection in high computational problems. The author proposes a method to detect credit card fraud and the results are validated using principles of this algorithm. The purpose of detecting fraud cases is to declare it to the client and the service provider.
Renu et al [17] proposed a fraud detection method which involves monitoring the activities of populations to observe and predict undesirable behaviour. Undesirable behaviour is a set of several habits like intrusion, fraud, delinquency and defaulting.
This research speaks on several credit card fraud detection and telecommunication fraud and different techniques which help in resolving the discussed problems.
Venkata Ratnam Ganji et al [20] the author uses concept of data stream outlier detection algorithm which is based on anti knearest neighbors for credit card fraud identification. Whereas traditional methods need to scan the database many times to find the fraudulent transaction, which is not suitable for data stream surroundings. This method makes easier to stop fraudulent transaction happens by Lost and stolen card and Credit card validation checks and detects errors in a sequence of numbers which also helps to detect valid and invalid numbers easily.

## IV. HOW HMM CAN BE USED FOR CREDIT CARD FRAUD DETECTION

After the HMM parameters are learned, the symbols from a cardholder’s training data is taken and an initial sequence of symbols is formed [9].
Let $C_1$, $C_2$, $C_3$… $C_K$ be one of such sequence of length K. This recorded sequence is formed from the cardholder’s transactions up to time t. This sequence is imputed into the HMM and the probability of acceptance is computed. Let the probability be α□ ,which can be written as

$$\alpha_1 = P(C_1,C_2,C_3, …., C_K \lambda|) \qquad (1)$$

Let $C_{K+1}$ be the symbol generated by a new transaction at time t+1. To form another sequence of length K, we remove C1 and add $C_{K+1}$ in that sequence, generating $C_2$, $C_3$, ..$C_R$, $C_{R+1}$ as the new sequence to the HMM, we read this sequence into the HMM and calculate the probability of acceptance by the HMM. Let the new probability be α□

$$\alpha_2 = P (C_2, C_3,C_4, …. ,C_{K+1} \lambda|) \qquad (2)$$

$$\text{Let } \Delta\alpha = \alpha_1 - \alpha_2 \qquad (3)$$

Assuming $\Delta\alpha > 0$, it means that the new sequence is accepted by the HMM with low probability, and it could be a fraud. The newly added transaction is determined to be fraudulent if the percentage change in the probability is above a threshold; otherwise the transaction is genuine [9].

$$\Delta\alpha/\alpha_1 \geq \text{Threshold} \qquad (4)$$

### 4.1 Advantages of HMM Approach

• The detection of the fraudulent use of a card is found much faster than when using the existing system.
• In case of the existing system even the original card holder is also checked for fraud detection but in this system no need to check the original user as we maintain a log.
• The log which is maintained will also be a proof for the bank for the transaction made.
• It is the most accurate technique for fraud detection.
• There is a decrease in the number of false positive transactions recognized as malicious by a fraud detection system even though they are really genuine [11,12]

### 4.2. Techniques and Algorithm

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 [5]. The genuine price variety for each symbol is configurable based on the expenditure routine of personal cardholders. HMM determine these prices rang dynamically by using clustering algorithms (like K clustering algorithm) on the price values of every card holder transactions. It uses cluster $V_k$ 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 [15]. In this prediction process it considers mainly three price value ranges such as 1) low (l) 2) Medium (m) and 3) High (h). So set of this model prediction symbols is V { l, m, h}, so V ¼ f as l (low), m (medium), h (high) which makes M ¼ 3. 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 [10]. 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. II) Fraud Detection System All the information about credit card (Like Credit card number, credit card CVV number, credit card Expiry month and year, name on credit card etc.) will be checked with credit card database. If User entered database is correct then it will ask Personal Identity number (PIN). After matching of Personal Identity number (PIN) with database and account balance of user's credit card is more than the purchase amount, the fraud checking module will be activated. The verification of all data will be checked before the first page load of credit card fraud detection system. If user credit card has less than 10 transactions then it will directly ask to provide personal information to do the transaction. Once database of 10 transactions will be developed, then fraud detection system will start to work. By using this observation, determine users spending profile. The purchase amount will be checked with spending profile of user. By transition probabilistic calculation based on HMM, it concludes whether the transaction is real or fraud. If transaction may be concluded as fraudulent transaction then user must enter security information. This information is related with credit card (like account number, security question and answer which are provided at the time of registration). If transaction will not be fraudulent then it will direct to give permission for transaction. If the detected transaction is fraudulent then the Security information form will arise. It has a set of question where the user has to answer them correctly to do the transaction. These forms have information such as personal, professional, address; dates of birth, etc are available in the database. If user entered information will be matched with database information, then transaction will be done securely. And else user transaction will be terminated and transferred to online shopping website.

## V. CONCLUSION

Credit card fraudulent detection which is done using HMM (Hidden Markov Model).This technique is used to detect various suspicious activities on credit card. It maintains a database, where past records of transactions are saved and any unusual transaction if carried out, which differs too much from the previous records, it tracks it. Let the user know by sending the details of the transaction on his mobile and hence prevent fraud.

In this paper, it has been discussed that how Hidden Markov Model will facilitate to stop fraudulent online transaction through credit card. The Credit Card Fraud Detection System is also scalable for handling vast volumes of transactions processing. The Hidden Markov Model makes the processing of detection very easy and tries to remove the complexity. At the initial state Hidden Markov Model checks the upcoming transaction is fraudulent or not and it allow to accept the next transaction. The different ranges of transaction amount like high range, medium range, and low range as the observation symbols were considered. As a result, efficient credit card fraud detection systems are most requirements for card issuing banks and all type of online transactions that make use of credit cards. The Credit Card Fraud Detection System is also scalable for handling large volumes of transactions.

## VI.REFERENCES

[1] Abhinav Srivastava, Amlan Kundu, Shamik Sural, and Arun K. Majumdar" Credit Card Fraud Detection Using Hidden Markov Model" VOL. 5, NO. 1, JANUARY-MARCH 2008.

[2] Baeza-Yates R, Ribeiro-Neto B. Modern information retrieval. New York: ACM Press New York; 1999.

[3] Binitie AP, Blamah NV, Ogah US. Synthetic software method: panacea for combating internet fraud in Nigeria. The International Journal of Engineering and Science. 2013.

[4] Chandra R, Chaudhary K, Kumar A. The combination and comparison of neural networks with decision trees for wine classification. School of Science and Technology; 2007.

[5] Cho, S.B., and Park, H.J., 2003. Efficient Anomaly Detection by Modeling Privilege Flows Using Hidden Markov Model, Computer and Security, vol. 22, no. 1 (2003), pp. 45-55.

[6] Divya.Iyer,Arti Mohanpurkar,Sneha Janardhan,Dhanashree Rathod,Amruta Sardeshmukh" credit card fraud detection using hidden markov model " 978-14673-0126-8/11/$26.00_c 2011 IEEE.

[7] Don L, Dhakwan S. Credit card fraud detection using Hidden Markov Model. European Journal of Industrial and System Engineering. 2013.

[8] Esakkiraj S, Chidambaram S. A predictive approach for fraud detection using Hidden Markov Model. International Journal of Engineering Research & Technology. 2013.

[9] Ingole A, Thool R. "Credit card fraud detection using Hidden Markov Model and its performance", International Journal of Advanced Research in Computer Science and Software Engineering, 2013.

[10] Kaufman, L., and Rousseeuw, P.J., "Finding Groups in Data: An Introduction to Cluster Analysis, Wiley Series in Probability and Math. Statistics", 1990.

[11] Kavita R, Jyoti H, "Credit card fraud detection using Hidden Markov Model", International Journal of Latest Research in Science and Technology, 2012.

[12] Khyati C, Jyoti Y, Bhawna M, "A review of fraud detection techniques: Credit card", International Journal of Computer Applications, 2012.

[13] Kirkby, " WEKA explorer user guide for Version 3-3-4", University of Waikato, 2002.

[14] K.RamaKalyani, D.UmaDevi "Fraud Detection of Credit Card Payment System by Genetic Algorithm" Volume 3, Issue 7, July-2012.

[15] Ourston, D., Matzner, S., Stump, W., and Hopkins, B, "Applications of Hidden Markov Models to Detecting MultiStage Network Attacks", Proceedings of 36th Annual Hawaii International Conference System Sciences, vol. 9 , 2003.

[16] Ray DP, Ghahremani Y, "Credit card statistics, industry facts, debt statistics", Available:http://www.creditcards.com/credit -card-news/credit-card-industry-factspersonal-debt-statistics-1276.php, 2013.

[17] Renu, Suman, "Analysis on Credit Card Fraud Detection Methods", vol.8, Feb 2014.

[18] Srivastava A, Kundu A, Sural S, Majumdar AK, "Credit card fraud detection using Hidden Markov Mode", IEEE Transactions on Dependable and Secure Computing. 2008.

[19] Stolfo SJ, Fan DW, Lee W, Prodromidis AL, Chan PK, "Credit card fraud detection using meta-learning: Issues and initial results", New York, 1997.

[20] Venkata Ratnam Ganji, " Credit card fraud detection using Anti-k Nearest Neighbor Algorithm", International Journal on Computer Science and Engineering (IJCSE), 2012.