# Credit Card Fraud Detection Using Hidden Markov Model

**Hemantraj Gautam[1] Sushan Sahu[2] Trupti Kini[3]**
[3]Assistant Professor
[1,2,3]Department of Computer Engineering
[1,2,3]Ideal Institute of Technology, Posheri, Palghar, India

*Abstract*— Now a day's credit card is the most accepted payment method in online as well as offline shopping. It is cashless payment and can be used worldwide. It is the most convenient way to do online shopping, bill payment, etc. Due to this the risks of fraud transaction using credit card has also been increasing. In the existing credit card fraud detection business processing system, fraud transaction would be detected after transaction is done. And it is very difficult to find such person, and also it causes losses in savings. Hidden Markov Model is the statistical tools for engineer and scientists to solve various problems. In this paper, it is shown that credit card fraud detection can be done using a very powerful data mining tool and that is the Hidden Markov Model. Hidden Markov Model helps to obtain a high fraud coverage combined with a low false alarm rate.

***Key words:*** Fraud Detection, Credit Card Fraud Detection, Hidden Markov Model, Fraud Detection Project

## I. INTRODUCTION

In day to day life credit cards are used for purchasing goods and services by the help of virtual card for online transaction or physical card for offline transaction. In physical transaction, credit cards will insert into payment machine at merchant shop to purchase goods. Tracing fraudulent transactions in this mode may not be possible because the attacker already steal the credit card. The credit card company may go in financial loss if loss of credit card is not realized by credit card holder. In online payment mode, attackers need only little information for doing fraudulent transaction (secure code, card number, expiration date etc.). In this purchase method, mainly transactions will be done through internet or telephone. Small transactions are generally undergo less verification, and are less likely to be checked by either the card issuer or the merchant. Card issuers must take more precaution against fraud detection and financial losses. Credit card fraud cases are increasing every year. In 2008, number of fraudulent through credit card had increased by 30 percent because of various ambiguities in issuing and managing credit cards. Credit card fraudulent is approximately 1.2% of the total transaction amount, although it is not small amount as compare to total transaction amount which is in trillions of dollars in 2007.

Hidden Markov model will be helpful to find out the fraudulent transaction by using spending profiles of user. It works on the user spending profiles which can be divided into major three types such as 1) lower profile; 2) middle profile; and 3) higher profile. For every credit card, the spending profile is different, so it can figure out an inconsistency of user profile and try to find fraudulent transaction. It keeps record of spending profile of the card holder by both way, either offline or online. Thus, analysis of purchased commodities of cardholder will be a useful tool in fraud detection system and it is assuring way to check fraudulent transaction, although fraud detection system does not keep records of number of purchased goods and categories. Every user represented by specific patterns of set which containing information about last 10 transaction using credit card. The set of information contains spending profile of card holder, money spent in every transaction, the last purchase time, category of purchase etc. The potential threat for fraud detection will be a deviation from set of patterns.

## II. LITERATURE SURVEY

Credit card fraud detection has drawn a lot of research interest and a number of techniques, with special emphasis on neural networks, data mining and distributed data mining have been suggested.

In project "Identifying online credit card fraud using Artificial Immune Systems" [1] authors has used Aritficial Immune Systems technique to identify online credit card fraud sytem. In ths system the pattern can be highly recognized but it requires a high training time in NSA. And it was also poor in handling missing data.

In project "Improving a credit card fraud detection system using genetic algorithm" [2] authors has used the technique of Genetic Algorithm for the implementation of credit card frauds. It works well with noisy data but it requires extensive tool knowledge to operate and very difficult to understand

In project "Personalized Approach Based on SVM and ANN for Detecting Credit Card Fraud" [3] author has discussed about the technique of Support Vector Machines for detecting ccredit card fraud. The SVM is poor for processing large dataset. It is expensive. It has a low speed of detection. Low accuracy

In project "Credit card fraud detection using machine learning techniques" [4] author has introduced Bayesian Network technique for the detection of fraud transaction. This technique is expensive and doesn't work on small data set and it requires excessive training.

In project "Fraud Detection by Monitoring Customer Behavior and Activities." [5] author has introduced fraud detection using CBR - Case Based Reasoning. It may suffer from incomplete or noisy data. And the accuracy can decrease.

## III. EXISTING SYSTEM

In existing models, the bank is verified credit card information, CVV number, Date of expiry etc., but all these information are available on the card itself. Nowadays, bank is also requesting to register your credit card for online secure password. In this new model, after feeding details of card at merchant site, then it will transfer to a secure gateway which is established at bank's own server. But, it is not verifying that the transaction is fraudulent or not. If hackers will get secure code of credit card by phishing sites or any other source, then it is very difficult to trace fraudulent transaction.

 

### *A. Scope*

The Hidden Markov Model is used to study and learn the human behavior during his/her transaction. It learns multiple transaction of user, and if the amount is exceeding then it will try to do a verification check with user personal detail, which is not known by anyone rather than the user. And if after verification the identity of the user doesn't match then the Hidden Markov Model will automatically block the transaction.

The scope of this project is to reduce the fraudulent transaction and reduce the loss of any credit card holder.

### *B. Proposed System*

In proposed model based on HMM will help to verify fraudulent of transaction during transaction will be going to happen. It includes two modules are as follow

I) Online Shopping: It comprises with many steps, first is to login into a particular site to purchase goods or services, then choose an item and next step is to go to payment mode where credit card information will be required. After filling all these information, now the page will be directed to proposed fraud detection system which will be installed at bank's server or merchant site.

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.

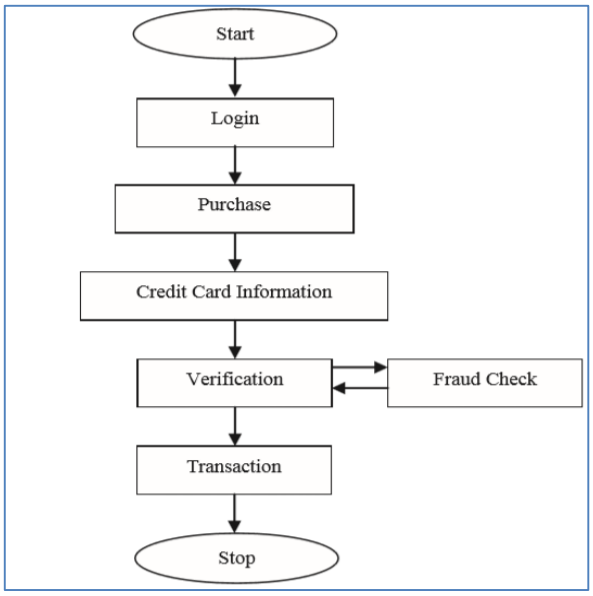


Fig. 1: Flowchart of HMM module for credit card fraudulent detection

## IV. METHODOLOGY FOR IMPLEMENTATION

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 rang dynamically by using clustering algorithms (like K 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 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.

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.

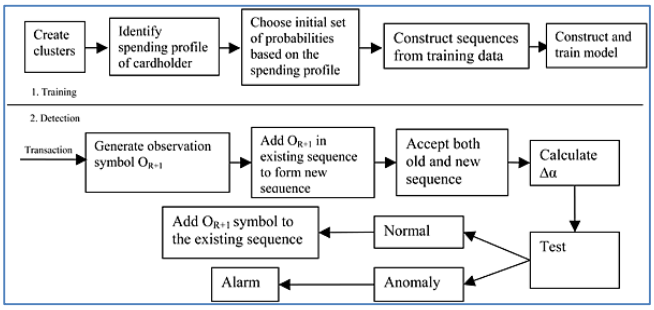


Fig. 2: Training and Verification of Fraudulent Transaction

*A. System Requirements*

*1) Software Requirements:*

- Windows Xp, Windows 7(ultimate, enterprise), Windows 8, Windows 10
- Sql 2008, PHP, XAMPP
- Visual studio 2010
- Python

*2) Hardware Components:*

- Processor – i3
- Min Hard Disk – 4 GB
- Min Memory – 1GB RAM

## V. CONCLUSION

In this paper, it has been discussed that how Hidden Markov Model will facilitate to stop fraudulent online transaction through credit card. The Fraud Detection System is also scalable for handling vast volumes of transactions processing. The HMM based credit card fraud detection system is not taking long time and having complex process to perform fraud check like the existing system and it gives better and fast result than existing system. The Hidden Markov Model makes the processing of detection very easy and tries to remove the complexity.

At the initial state, HMM checks the upcoming transaction is fraudulent or not and it allow to accept the next transaction or not based on the probability result. The different ranges of transaction amount like low group, medium group, and high group as the observation symbols were considered. The types of item have been considered to be states of the Hidden Markov Model. It is recommended that a technique for finding the spending behavioral habit of cardholders, also the application of this knowledge in deciding the value of observation symbols and initial estimation of the model parameters

In this proposed model, 84-86% transactions are genuine and very low false alarm which is about 7 % of total number of transactions. The relative studies and our results sure that the correctness and effectiveness of the proposed system is secure to 80 percent over a broad deviation in the input data..

## REFERENCES

[1] Anthony Brabazon, Jane Cahill, Peter Keenan, Daniel Walsh. Identifying online credit card fraud using Artificial Immune Systems. 18-23 July 2010; Spain. IEEE Congress on Evolutionary Computation; INSPEC Accession Number: 11568025 Conference Location: Barcelona;

[2] M. Hamdi Özçelik, Ekrem Duman, Mine Işik, Tuğba Çevik. Improving a credit card fraud detection system using genetic algorithm; 11-12 June 2010; International Conference on Networking and Information Technology; INSPEC Accession Number: 11432153; Manila, Philippines

[3] Rong-Chang Chen, Shu-Ting Luo, Xun Liang, V.C.S. Lee. Personalized Approach Based on SVM and ANN for Detecting Credit Card Fraud; 13-15 Oct. 2005; International Conference on Neural Networks and Brain; INSPEC Accession Number: 9072543; Conference Location: Beijing, China

[4] John O. Awoyemi, Adebayo O. Adetunmbi, Samuel A. Oluwadare. Credit card fraud detection using machine learning techniques; 29-31 Oct. 2017; International Conference on Computing Networking and Informatics (ICCNI); INSPEC Accession Number: 17412979; Conference Location: Lagos, Nigeria

[5] Parvinder Singh, Mandeep Singh. Fraud Detection by Monitoring Customer Behavior and Activities. February 2015; International Journal of Computer Applications; Conference Location: Chandigarh University, Mohali, Punjab, India