Fraudulent Transaction Detection using HMM

[1]Sable Pratik N., [2]Mahajan Ankit G., [3]Ugale Anand B.

Students,
Computer Dept., Govt. college of Engg. & research, Awasari, Dist. Pune, Maharashtra, India
[1]pnsable8891@gmail.com, [2]mahajanankit59@gmail.com, [3]anandugale333@gmail.com

***Abstract*— as comparing to both online as well as offline Transaction most popular mode of payment is online Transaction, so chances of fraudulent transaction is also increases. Fraudulent transactions are like stolen card, Hack account, lost card, legitimate attack etc. In existing system, fraud is detected after fraudulent transaction performed. In this paper, by using Hidden Markov Model (HMM) we can model the operation in credit card transaction processing to detection of frauds. In this paper, we 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 behavior of a cardholder. If an incoming credit card transaction is not accepted by the trained HMM with sufficiently high probability, it is considered to be fraudulent. At the same time, we try to ensure that genuine transactions are not rejected. We present detailed experimental results to show the effectiveness of our approach and compare it with other techniques.**

***Keywords* - Hidden Markov Model, card holder, transaction, flash code, Bio-informatics, Personal Identification Number (PIN).**

I. INTRODUCTION

A Hidden Markov Model is a finite set of states; each state is linked with probability distribution. Transitions among these states are governed by a set of probabilities called transition probabilities. In a particular state a possible outcome or observation can be generated which is associated symbol of observation of probability distribution. It is only the outcome, not the state that is visible to an external observer and therefore states are "hidden" to the outside; hence the name Hidden Markov Model. Hence, Hidden Markov Model is a perfect solution for addressing detection of fraud transaction through credit card. One more important benefit of the HMM-based approach is an extreme decrease in the number of False Positives transactions recognized as malicious by a fraud detection system even though they are really genuine. An HMM is initially trained with the normal behavior of a cardholder. If an incoming credit card transaction is not accepted by the trained HMM with sufficiently high probability, it is considered to be fraudulent. At the same time, we try to ensure that genuine transactions are not rejected [1].

A Hidden Markov Model (HMM) is a statistical model in which the system being modeled is assumed to be a Markov process with unobserved state. A HMM can be considered as the simplest dynamic Bayesian network. In a regular Markov model, the state is directly visible to the observer, and therefore the state transition probabilities are the only parameters. In a hidden Markov model, the state is not directly visible, but output, dependent on the state, is visible. Each state has a probability distribution over the possible output tokens. Therefore the sequence of tokens generated by an HMM gives some information about the sequence of states. Note that the adjective 'hidden' refers to the state sequence through which the model passes, not to the parameters of the model; even if the model parameters are known exactly, the model is still 'hidden'. Hidden Markov models are especially known for their application in temporal pattern recognition such as speech, handwriting, gesture recognition, part-of-speech tagging, musical score following, partial discharges and bioinformatics. A hidden Markov model can be considered a generalization of a mixture model where the hidden variables, which control the mixture component to be selected for each observation, are related through a Markov process rather than independent of each other [2].

Following diagram will clear that

- Three states(rain, cloud, sunny)
- And their probability out of 8.

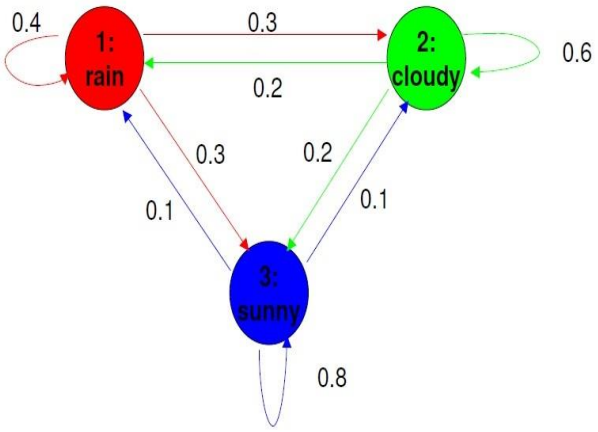

FIGURE 1- Probabilistic state transition diagram of HMM

II. LITERATURE REVIEW

Credit card fraud detection has received an important attention from researchers in the world. Several techniques have been developed to detect fraud transaction using credit card which are based on neural network, genetic algorithms, data mining, clustering techniques, decision tree, Bayesian networks etc. Ghosh and Reilly have proposed a neural network method to detect credit card fraud transaction. They have built a detection system, which is trained on a large sample of labeled credit card account transactions. These sample contain example fraud cases due to lost cards, stolen cards, application fraud, stolen card details, counterfeit fraud etc. They tested on a data set of all transactions of credit card account over a subsequent period of time. Bayesian networks are also one technique to detect fraud, and have been used to detect fraud in the credit card industry. This techniques yield better results but having large cycle time to detect fraud. However, the time constraint is one main disadvantage of this technique, especially compared with neural networks. In this technique, clustering of two algorithms have used for behavioral fraud detection. The proposed system was indented those accounts that are behaving differently from others at the particular moment whereas they were behaving the same previously. Those accounts are treating as suspicious ones and fraud analysis is to be done only on these accounts. Now a days the current ATM systems are providing ATM cards and PIN codes through which user performs bank transactions like withdrawing money, printing mini statements, balance inquiry, depositing money in the bank etc. [3].

Following figure shown comparison between Hidden Markov Model (HMM) and other models –

Parameter	Fusion of Dempster–Shafer theory and Bayesian learning	Hybridization of BLAST-SSAHA	Hidden Markov Model	Bayesian and Neural Networks	Fuzzy Darwinian Detection
Methodology	Machine Learning	Sequence Alignment	Hidden Markov Model	Artificial Intelligence, machine Learning	Genetic Programming, Fuzzy Logic
Fraud Detection TP	98%	85%	70%	77%	100%
Fraud Detection FP	10%	15%	20%	10%	5.26%
Processing Speed	Medium	Very High	High	High	Low
Training Required	Yes	No	Yes	Yes	Yes
Supervised learning	Supervised	Un-Supervised	Supervised	Supervised	Supervised
Cost	Expensive	Inexpensive	Quite expensive	Expensive	Highly Expensive
Accuracy	High	High	Medium	Medium	Very High

FIGURE 2- Comparison of fraud detection techniques

III. SYSTEM DESCRIPTION

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 [5].

IV. PROPOSED SYSTEM

A. Enhanced Hidden Markov Model Approach in detecting Credit Card Frauds: This paper is constructed by using the credit card fraud detection by Hidden Markov Model as a base model. It is developed focusing on three main constraints:

Behavior: It is defined as a portion of all the number of transactions which are identified correctly. That is the genuine transactions as genuine and fraud transactions as fraud.

Time & Amount: This gives the value of division of transactions detected correctly whether genuine or fraudulent. Here it counts only if genuine transactions are detected as genuine and fraud transactions as fraud and not any other.

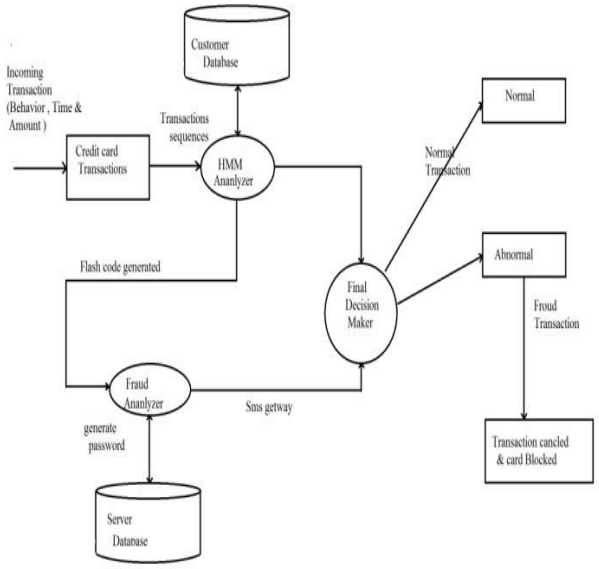

FIGURE 3- System Architecture

The UML diagram representation of the proposed system architecture is represented in the following figure-4 that shows all the user functionalities:

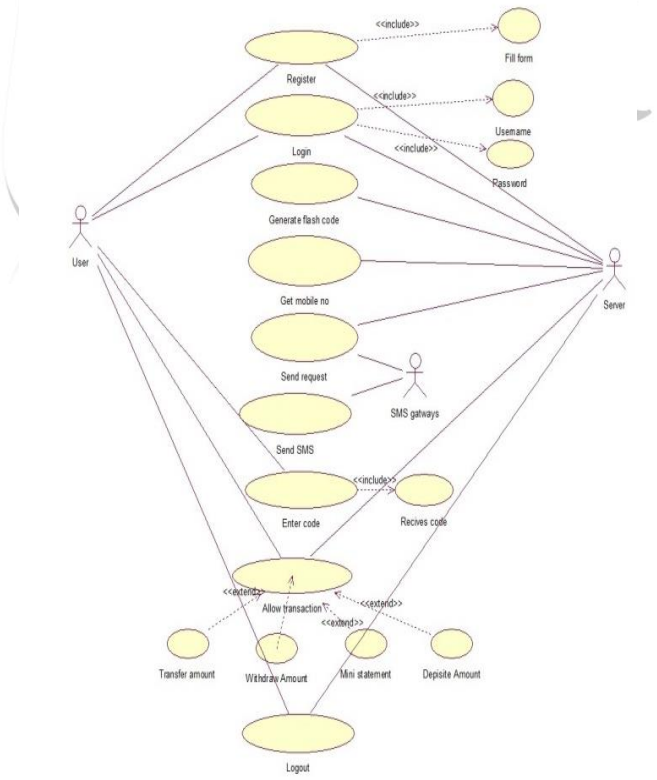

FIGURE 4- Use-case diagram

V. MODULES DESCRIPTION

1. *Client side Application*: Using AWT / Swing. This GUI shall allow the user to log in and transact online using internet banking enabled account.
2. *Client Transaction Module*: A module that will allow clients to enter their credentials / authentication information and proceed with a transaction. This module also presents the client with transaction report (success / failure / etc.).
3. *Account Database*: A database containing account information of all clients is maintained on bank's server. The details may include account number, login, password, available balance, etc.
4. *Transaction Database*: A database containing history of client's online transactions will also be maintained on server. The databases shall be maintained using Object Serialization.
5. *Server Side Item / Service*: A module that will allow the client to browse through all available items/services available on internet. Client can select any of these items/services and opt to buy them online.
6. *Client Server Interaction:* A module using Java Networking shall be built that will allow the client application to call Server.

VI. ADVANTAGES

1. The detection of the fraud use of the card is found much faster that existing system.
2. 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.
3. The log which is maintained will also be a proof for the bank for the transaction made.

We can find the most accurate detection using this technique.

VII. CONCLUSION

In Internet Banking a Fraud detection system will run at the banks server. And it's Function to do financial transaction without any fraud. It is considered under Prediction system. A method to attack signature based online banking methods is to manipulate the used software in a way, that correct transactions are shown on the screen and faked transactions are signed in the background. Hidden Markov Model is used to track the user behavior. First user behavior is recorded and then for new transaction it is checked. We have used the ranges of transaction amount as the observation symbols, whereas the types of item have been considered to be states of the HMM. We have suggested a method for finding the spending profile of cardholders, as well as application of this knowledge in deciding the value of observation symbols and initial estimate of the model parameters. The system is also scalable for handling large volumes of transactions. The proposed methodology is aimed at detecting fraud in case of internet banking.

VIII. FUTURE WORK

Future work can be continued in the manner of Using Different Algorithm for checking Fraud Detection making system more and more accurate ad also more reliable. Instead of HMM algorithm we can use other algorithms which are better than HMM.

ACKNOWLEDGMENT

We are thankful to all helping hands in completion of this project. We would like to express our sincere thanks to all those who have provided us with valuable guidance towards completion of project.

REFERENCES

[1] Hidden Markov Model by Jia Li. Department of Statistics "The Pennsylvania State University"
[2] http://www.stat.psu.edu/~jiali/course/stat597e/notes2/hmm.pdf
[3] A Revealing Introduction to Hidden Markov Models by mark stamp.
[4] Sam Roweis, Hidden Markov Models (SCIA Tutorial 2013)
[5] http://www.cardhub.com/edu/number-of-credit-cards
[6] Global Consumer Attitude towards On-Line Shopping,http://www2.acnielsen.com/reports/documents /2005cconlineshopping:pdf
[7] A tutorial on hidden Markov models and selected applications,www.cs.ubc.ca/ murphyk/Bayes/rabiner.pdf
[8] http://www.google.com/wikihmm:html
[9] FLEXChip Signal Processor (MC68175/D), Motorola, 1996.
[10] http://www.wikipedia.com/creditcardf rauddetectionusinghmm