Internet Banking Fraud Detection Using HMM

Mr. Sunil S Mhamane
M.E(C.S.E)
Walchand Institute of Technology
Solapur, India.
sunil_mhamane@yahoo.co.in

Mr. L.M.R.J Lobo
Associate Professor,Dept.Of Computer Science and Engg.
Walchand of Institute of Technology
Solapur, India
headitwit@gmail.com

***Abstract*— Internet banking has their separate account for users and is managed by Banks or retail store. The aim of this paper is to detect and prevent fraud in case of internet banking using Hidden Markov Model algorithm. At the same time, we have tried to ensure that genuine transactions are not rejected by making use of one time password that was generated by the Bank server and sent to the particular customers through SMS to their Mobile number which is registered in the system. Banks are seeking to minimize huge losses through fraud detection and prevention systems. Many different advanced fraud technologies are being applied to fraudulent Internet banking transactions detection and prevention. However, they have no effective detection mechanism to identify legitimate users and trace their unlawful activities. We propose a model to overcome all these difficulties using Hidden Markov Model.**

***Keywords*—Internet Banking, Hidden Markov Model, Probability, fraud detection,Transaction.**

I. INTRODUCTION

Since 1997 when the first domestic internet bank was opened by Bank of China, the online banking business has been developed so rapidly that 45 banks from over 100 local banks have set up their own online banks until 2007. With the unique features of cost and "location", online banks have been accepted and recognized by more and more clients.

A. Online Banking Feature

Online banking solutions have many features and capabilities in common, but traditionally also have some that are application specific. The common features fall broadly into several categories.

- Transactional (e.g., performing a financial transaction such as an account to account transfer, paying a bill, wire transfer... and applications... apply for a loan, new account, etc.)
    - Electronic bill presentment and payment - EBPP
    - Funds transfer between a customer's own checking and savings accounts, or to another customer's account
    - Investment purchase or sale
    - Loan applications and transactions, such as repayments of enrollments
- Non-transactional (e.g., online statements, check links, co browsing, chat)
    - Bank statements
- Financial Institution Administration
- Support of multiple users having varying levels of authority
- Transaction approval process
- Wire transfer

Features commonly unique to Internet banking include Personal financial management support, such as importing data into personal accounting software. Some online banking platforms support account aggregation to allow the customers to monitor all of their accounts in one place whether they are with their main bank or with other institutions.

B. Hidden Markov Model

Hidden Markov models (HMMs) are one of the most popular methods in machine learning and statistics for modeling sequences such as speech and proteins. An HMM defines a probability distribution over sequences of observations. 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. 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[1].

A Hidden Markov Model is beneficial because we can Model Sequence of Transactions to Different state also at each state we are Deciding whether Transaction is a case of Fraud or not. One of the advantages of using hidden Markov models for profile analysis is that they provide a better method for dealing with gaps found in protein families. The basic theory of HMMs is also very elegant and easy to understand. This makes it easier to analyze and develop implementations [2].

II. LITERATURE REVIEW

In "Credit Card Fraud Detection Using HMM" paper, They have proposed an application of HMM in credit card fraud detection. The different steps in credit card transaction processing are represented as the underlying stochastic process of an HMM. They 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. They 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. It has also been explained how the HMM can detect whether an incoming transaction is fraudulent or not. Experimental results show the performance and effectiveness of our system and demonstrate the usefulness of learning the spending profile of the cardholders. Comparative studies reveal that the Accuracy of the system is close to 80 percent over a wide variation in the input data. The system is also scalable for handling large volumes of transactions[3].

In "credit card fraud detection with a neural network" paper, Using data from a credit card issuer, a neural network based fraud detection system was trained on a large sample of labeled 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. They discuss the performance of the network on this data set in terms of detection accuracy and earliness of fraud detection. The system has been installed on an IBM 3090 at Mellon Bank and is currently in use for fraud detection on that bank's cmlit card portfolio[4].

In "Offline Internet Banking Fraud Detection" paper .Object of this paper is to demonstrate one successful fraud detection model which is established in Greece. Apart from the offline internet banking fraud detection system itself, which is described briefly, there scope is to present its contribution in fast and reliable detection of any "strange" transaction including fraudulent ones[5].

In "Security Analysis for Internet Banking Models" paper They stated that Internet banking fraud can be performed internally by genuine staff or externally by customers or suppliers. This paper presents a security analysis of the proposed Internet banking model compared with that of the current existing models used in fraudulent Internet payments detection and prevention. Several modern models in preventing and detecting fraud are evolving and being applied to many banking systems. However, they have no effective detection mechanism to identify legitimate users and trace their unlawful activities. Also they are not secure enough to prevent fraudulent users from performing fraudulent transactions over the Internet. The proposed model facilitates Internet banking Fraud Detection and Prevention (FDP) by applying two new secure mechanisms, Dynamic Key Generation (DKG) and Group Key (GK) [6].

In "Study on Fraud Risk Prevention of Online Banks" paper .The paper is aimed, in the first hand, at giving a discussion on the fraud risks of online banking, introducing the current application situation of information sharing mechanism in respect of internet fraud outside China as well as the development of such concept in China. Then, a system is designed for sharing internet fraud information. The paper finally proposing that all the online banks should put more joint efforts in perfecting this mechanism for sake of international co operation[7].

In "Fraudulent Internet Banking Payments Prevention using Dynamic Key" In this paper, They have proposed an efficient new scheme which can prevent fraud by applying different security algorithms, generating and updating limited-use secret keys. It uses advanced authentication technologies and is well adapted to any possible future technology. Moreover, it does not rely on fixed values where hacking one secret will not compromise the whole system's security. The generation of each set of keys is based on dynamically generated preference keys. The higher number the transactions performed, the less chance the system has of being compromised. The practical usefulness of the technique has been demonstrated by applying it to Internet banking payment systems. The results show that our technique enhances their security considerably. It has been shown that the proposed technique is secure against key compromise. For future work, we aim to analyze the security of the system that applies the proposed technique. Moreover, we aim to apply the proposed technique to other kinds of internet applications, especially mobile commerce [8].

In the paper "Parallel Granular Neural Networks for Fast Credit Card Fraud Detection" . A parallel granular neural network **(**GNN**)** is developed to speed up data mining and knowledge discovery process for credit card fraud detection. The entire system is paraUelized on the Silicon Graphics Origin 2000, which is a shared memory multiprocessor system consisting of 24-CPU, 4G main memory, and 200GB hard-drive. In simulations, the parallel fuzzy neural network running on a 24-processor system is trained in parallel using training data sets, and then the trained parallel fuzzy neural network discovers fuzzy rules for future prediction. A parallel learning algorithm is implemented in ***C.*** The data are extracted into a flat file from SQL server database containing sample Visa Card transactions and then preprocessed for applying in fraud detection. The data are classified into three categories: first for training, second for prediction, and third for fraud detection. After learning from training data, the GNN is used to predict on second set of data and later the third set of data is applied for fraud detection. Around eight scenarios are employed for detecting purpose. GNN gives fewer average training errors with larger amount of past training data. We also found that the number of training error is inversely proportional to the number of training cycles. The higher the fraud detection error is, the greater the possibility of that transaction being actually fraudulent [9].

III. ARCHITECTURE OF PROPOSED SYSTEM

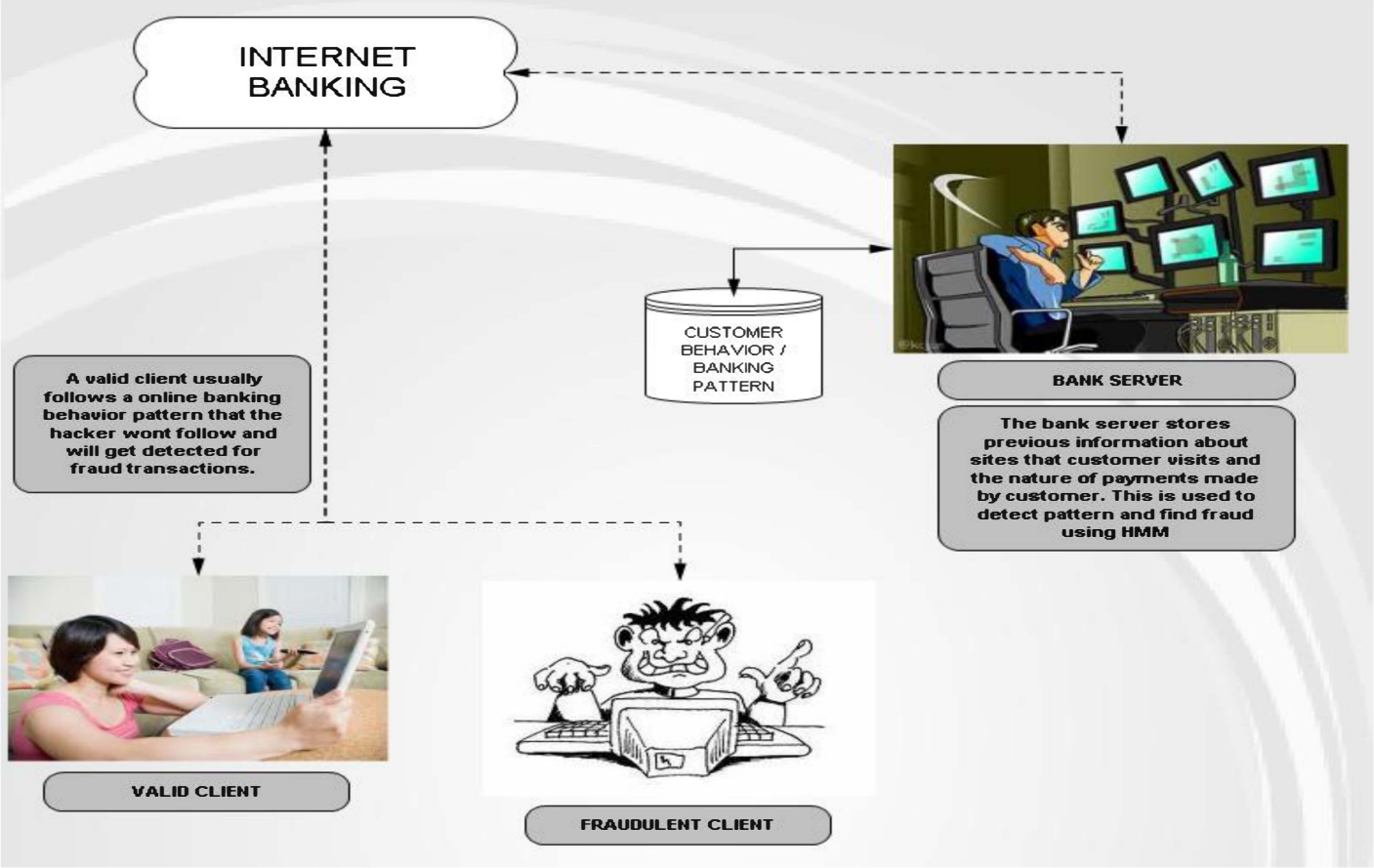

Figure 1. Architecture of Proposed System.

Basic architecture of Proposed system consists of following component.

A. Valid Client :

Basically he is the authorized client who is having internet banking account in particular bank .

B. Fraudulent Client :

He is the unauthorized user who is not having legal Internet banking account in bank. who makes use of authorized users Internet banking account to do Transaction. Hence He is Fraudulent User.

C. Bank Server :

Bank Sever Retrieves previous information about sites that customer visits and the nature of payment made by customer From Customer Behavior Banking Pattern Database. This is used to Detect pattern and to find Fraudulent Transaction.

D. Database :

Stores the previous information about Customer Behavior Banking Pattern.

IV. WORKING OF PROPOSED SYSTEM

Following are the steps of working :

- We will develop a dummy bank account database for several customers.
- A web service based on Tomcat Apache server will be created and deployed to allow users to make use of online banking.
- We will also develop some client applications that will allow the user to make online payments for required services. E.g. Buy Air Tickets, Buy Train Tickets, Movie Tickets, Transfer Amount, Shop Online, etc.
- Based on history of banking transactions we will design the database that will save customer transaction patterns.
- Using this pattern we will design an analysis algorithm based on HMM that will evaluate if the on-going transaction is fraudulent or original.
- We shall use Servlets for client-server communication.
- We will also design a client side application and a server side application using Java AWT / Swing.
- The client application shall communicate with server using Java Networking.

- The client application will use Serialized Objects for transacting with server.
- To implement the HMM algorithm we shall make use of Java Collections API.
- The database will also be maintained using Serialized Objects.

On finding a fraudulent transaction we shall send the One time password to the client to a actually verify the identity of client and continue with transaction in case the actual user is actually initiating it.

V. STATE DIAGRAM OF PROPOSED SYSTEM

Different states are shown in the below state diagram. Start state is login and end state is logout. First user is authenticated if authorized user then he can do internet banking transaction. all these Transaction details are sent to Server for processing. HMM algorithm is applied on Transaction details.

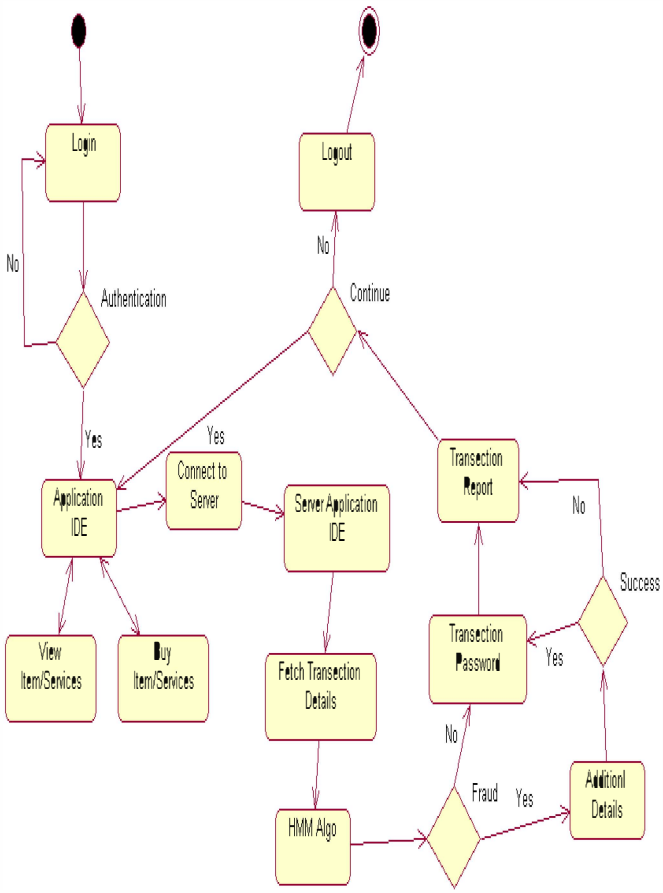

Figure 2. State Diagram of Proposed system.

VI. CONCLUSION AND FUTURE WORK.

This paper explains Architecture and design part of proposed system. It removes the drawback of previous paper named Credit card fraud detection using HMM. Customer Behavior pattern is used for fraud detection. Transaction amount was taken as Observation symbol.

Following Future work would be possible.

1. In my research transaction amount was taken as an observation. In future we could take other parameters to make the system strong and reliable.
2. A different Algorithm For checking Fraud Detection making system more and more accurate and reliable could be designed and implemented. Instead of HMM algorithm we could use another algorithm.

REFERENCES

[1] Hidden Markov Model by Jia Li.Department of Statistics "The Pennsylvania State University" http://www.stat.psu.edu/~jiali/course/stat597e/notes2/hmm.pdf.

[2] "A Revealing Introduction to Hidden Markov Models" by mark stamp.

[3] "Credit Card Fraud Detection Using Hidden Markov Model" By Abhinav Srivastava, Amlan Kundu, Shamik Sural. IEEE Transaction,January-March 2008.

[4] "credit card fraud detection with a neural network" by Ghosh and Reilly. IEEE" Proceedings of the Twenty-Seventh Annual Hawaii International Conference on System Sciences, 1994.

[5] " Offline Internet Banking Fraud Detection" by Vasilis Aggelis.

[6] " Security Analysis for Internet Banking Models" By Osama Dandash,Phu Dung Le and Bala Srinivasan. Eighth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing IEEE DOI 10.1109/SNPD.2007.5321142

[7] "Study on Fraud Risk Prevention of Online Banks" By Qinghua Zhang. 2009 International Conference on Networks Security, Wireless Communications and Trusted Computing.

[8] " Fraudulent Internet Banking Payments Prevention using Dynamic Key" By Osama Dandash Yiling Wang and Phu Dung Leand Bala Srinivasan. "JOURNAL OF NETWORKS, VOL. 3, NO. 1, JANUARY 2008".

[9] Parallel Granular Neural Networks for Fast Credit Card Fraud Detection Mubeena Syeda, Yan-Qing Zbang and Yi Pan. IEEE Transaction.