Securing Online Banking Transaction using Predictive Approach of Hidden Markov Model

Sangita D. Avghad
ME, Department of Computer Science & Engineering,
Jawaharlal Nehru
Engineering College,
Aurangabad, M.S. India.

Madhuri S. Joshi
Professor,
Department of Computer Science & Engineering,
Jawaharlal Nehru
Engineering College,
Aurangabad, M.S. India.

ABSTRACT

Due to a Fast growth in the electronic commerce technology, popularity of online banking and online shopping is growing day by day. While e-commerce is still gaining popularity, it also provides ground for fraudsters who try to misuse the transparency of online purchases and the transfer of credit card records. In this paper, proposed model the sequence of operation in online banking transaction processing using a Hidden Markov Model (HMM) and describe how it can be used for the detection of frauds. An HMM is initially trained with the normal behaviour of a cardholder. If current transaction is not accepted by the trained model with good probability, it is treated as fraudulent. And one time password is send to mobile of card holder.

Keywords

Fraud Detection System(FDS), Hidden Markov Model (HMM), one Time Password (OTP), Online Banking (OLB).

1. INTRODUCTION

The Internet banking Portal provides personal banking service that gives complete control over all the banking demands terms internet banking, web banking and online banking are same in their meaning. Internet banking portal have many common features and capabilities.

Transactional Features: Funds Transfer between two Customer’s Accounts, Paying third parties including bill payments, Investment purchase or sale Monthly Payments etc.

Non transactional Features: Viewing account balances, viewing recent transactions etc. also have feature Management of multiple users having varying levels of Authority and transactional Approval Feature. Due to such facility online banking, popularity of online shopping is growing day by day, at one and the same time, it also provides ground for fraudsters who try to misuse the transparency of online purchases and the transfer of credit card record. Credit card based purchase can be categorized into two types: 1) Physical card 2) virtual card. In a physical card based purchase, the cardholder presents his card physically to a merchant for making a payment. To carry out fraud in this type kind of purchase, the cardholder presents his card physically to a merchant for making a payment. If card substantial financial.in second kind of purchase only some information about card is required to make a payment. Such holder does not realize the loss of card, it can lead to Purchase is normally does on the internet. To commit fraud in this type of purchase, fraudster simply needs to know the card details.

2. RELATED WORK

Paper [7] which aim to propose System credit card fraud detection using a neural network. It trains detection system on large sample of labelled credit card account transaction. These transactions contain example fraud cases due to missing/forgotten cards, stolen cards, application fraud, counterfeit fraud, mail-order fraud, and no received issue (NRI) fraud.

In paper [2], V.Dheepa and R.Dhanapal Proposed system which make use of behavior based classification approach using Support Vector Machines and also make use of efficient feature extraction methods. When discrepancies occur in the behaviors transaction pattern then it is predicted as suspicious and taken for further consideration to find the frauds.C. Chiu and C. Tsai in paper [3], proposed Web services and data mining techniques to set up a collaborative Scheme for fraud detection in the banking industry. With this scheme, all participating banks share knowledge about the fraud patterns in a heterogeneous and distributed environment. To establish a Smooth channel of data Exchange, Web services/application service techniques such as XML, SOAP, and WSDL are used.

In paper [4] “BLAST-SSAHA Hybridization for Credit Card Fraud Detection “The proposed system that makes use of two-stage sequence alignment in which a profile analyzer (PA) first determines the similarity of an incoming sequence of transactions on a given credit card with the genuine cardholder's past spending sequences. The unusual transactions traced by the profile analyzer are next passed on to a deviation analyzer (DA) for possible alignment with past fraudulent behavior. The final decision about the nature of a transaction is taken on the basis of the observations by these two analyzers. In order to achieve online response time for both PA and DA, they suggested a new approach for combining two sequence alignment algorithms BLAST and SSAHA.In paper [5] Sam Maes, Karl Tuyls and Bram Vanschoenwinkel have shown two fraud detection techniques Artificial Neural Network and Bayesian Belief Network. They compared two methods and put conclusion that ANN fraud detection process is faster but BBN yields better results of fraud detection and training is shorter. They have given the future work of ANN and BBN that can improve performance.

3. PROPOSED SYSTEM ARCHITECTURE

Architecture of Proposed System Consist of Database which can Store information about Authorized Client such as name, address, contact no, email id etc. It also stores previous Transaction information the client made. These transactions

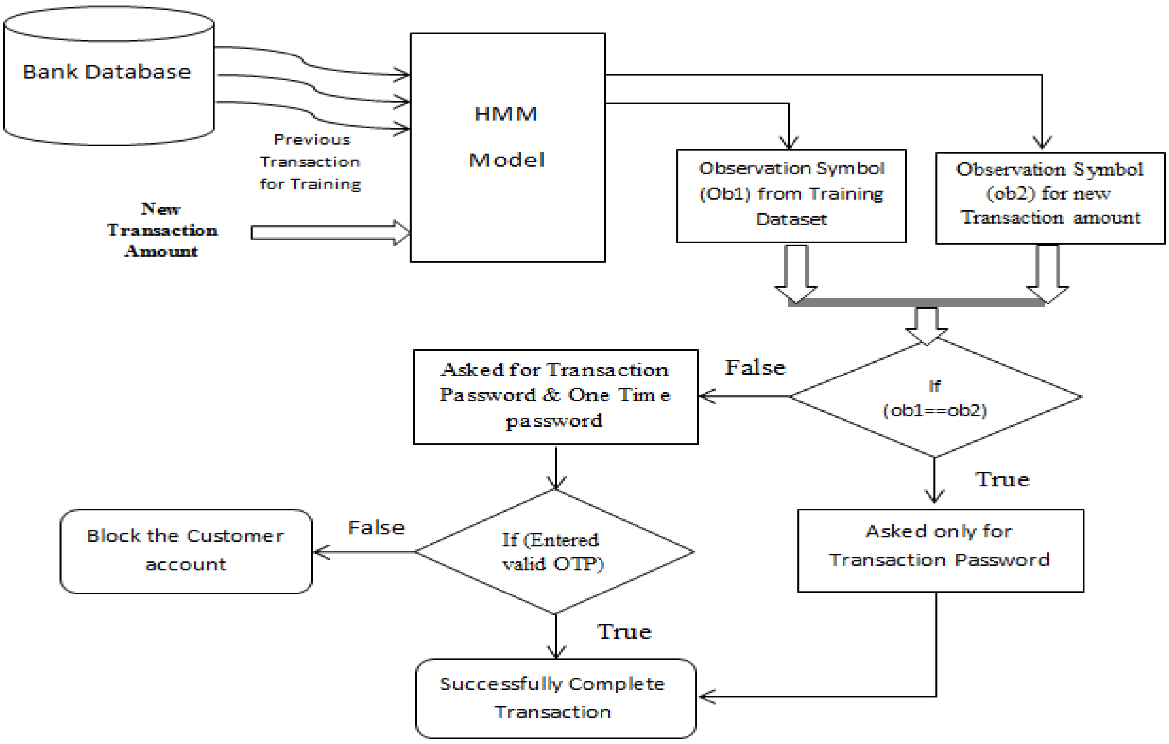
Bank Database
Previous Transaction for Training
New Transaction Amount
HMM Model
Observation Symbol (Ob1) from Training Dataset
Observation Symbol (ob2) for new Transaction amount
If (ob1==ob2)
False
Asked for Transaction Password & One Time password
True
Asked only for Transaction Password
If (Entered valid OTP)
False
Block the Customer account
True
Successfully Complete Transaction

Fig 1: Architecture of Proposed system

information is given to HMM Model for Training Purpose. Specific Sequence is Constructed from these training dataset. So that it directs for expected transaction sequence for each client. Spending amount is divided into three Categories Low [L] is Rs.100-Rs.900, Medium [M] is Rs.1000 –Rs.2400 and High [H] is Rs.2500 and above. L, M and H are considered as Observation Symbol. HMM Model gives Probability of next Symbol and also covert the current transaction amount into observation symbol. If these observation symbol from training dataset and observation symbol from current transaction amount match then transaction complete without asking one time password. But if they are not equal then one time password is send to mobile number of the customer.

4. PROPOSED ALGORITHM

Detection system consist of two phase.

1. Training Phase
2. Detection and Prevention Phase

1. Training Phase: Bank Database consists of previous Transaction history of each account holder of online Banking i.e actual client. These Transactions are given to HMM Model, based on these initial set of Transaction, spending Profile Actual client is identified. These training dataset gives expected transaction sequence for each actual client

2. Detection and Prevention Phase: In Detection phase system looks for deviation between actual outcome and expected Sequence. If found deviation then one time password is send to mobile number of the customer.

These two phases used by Hidden Markov Model to Predict current transaction is Fraudulent or not. For using a HMM requires to calculate the HMM Parameters such as state and transition probabilities. Those parameters are calculated using Baum-Welch algorithm. Baum and welch algorithm is given below. [11]

Baum-Welch Algorithm

1. Initialize the parameters (state, transition) to some values

2. Calculate "forward-backward" probabilities based on the current parameters

   - Forward probability is calculated as below.
     a. At time t, the probability that we're in state i.

     b. the previous observation thus far has been o1 … ot.

     c. $\alpha(i) = P(i, o1....ot / \lambda)$

   - Backward probability
     a. At time t and we're in state i, the probability that

     b. the observation that follows will be ot+1 … oT.

     c. $\beta(i) = P(Ot+1....OT/i, \lambda)$.

3. Use the forward-backward probabilities to estimate the expected frequencies

   - Expected number of transitions from state i.
   - Expected number of being in state j.

4. Use the expected frequencies to estimate the parameters.

Table 1: Detailed data for OTP asked for when Observation sequence length is two

Sequence Length for Training	False Positive								True Positive							
	OTP Asked To Actual Customer								OTP Asked To Fraudulent Customer							
	AC-1	AC-2	AC-3	AC-4	AC-5	AC-6	AC-7	AC-8	FC-1	FC-2	FC-3	FC-4	FC-5	FC-6	FC-7	FC-8
10	7	8	5	3	5	5	4	4	9	10	6	5	6	7	6	5
20	6	7	4	5	5	4	5	5	8	9	7	7	7	6	7	6
30	5	4	7	5	4	5	5	5	7	6	9	7	6	7	6	7
40	6	5	6	5	5	5	4	5	8	7	8	6	6	8	5	9
50	6	4	5	6	5	5	6	4	9	5	7	9	8	7	8	6
60	5	5	6	5	4	6	6	6	8	7	7	7	6	8	9	9
70	5	6	5	4	5	5	5	6	7	7	8	7	6	9	7	9
80	6	5	5	5	4	5	5	6	9	6	7	7	7	8	7	9

5. Repeat 2 to 4 until the parameters converge

5. PERFORMANCE ANALYSIS

For comaprative analysis,we kept observation sequence length fixed which is two and changing sequence length for training i.e.changing dataset length from 10 to 80 with diference of ten.It have 8 actual customer and 8 fraudlent customer for each sequence length.Each actual customer and fraudulent customer performed 10 transaction and the count OTP was asked for it.detailed result is shown in table[1].The plot for True Positive and Flase Positive in Percentage is show in Fig:2 for Observation Sequence Legth of two over different Sequence Length for Taining.from table:1 and Fig:3 it is seen that when otp asked for false positive is low then there is degradation in true positive

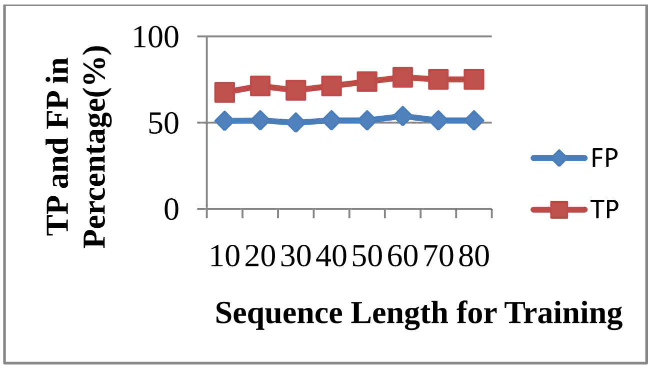

Fig 2: Line Graph for Performance Variation of the System for different sequence length when sequence length is two

Misclassification Rate (Error Rate):

$$\text{MR} = \frac{FN+FP}{TP+FN+FP+TN}$$

$$\text{MR} = \frac{177+329}{1280}$$

$$\text{MR} = 0.395$$

Table 2: Contingency table for Proposed System

	Frauds	Not Frauds
Alerted as Fraud or Victim	463	329
Not Alerted	177	311

Accuracy:

$$\text{A} = 1 - \text{MR} = \frac{TN+TP}{TP+FN+FP+TN}$$

$$\text{A} = 1 - \text{MR} = 1 - 0.395$$

$$\text{A} = 0.605$$

True Positive Rate

$$\text{TPR} = \frac{TP}{TP+FN}$$

$$\text{TPR} = \frac{463}{463+177} = 0.723$$

False Positive Rate

$$\text{FPR} = \frac{FP}{FP+TN}$$

$$\text{FPR} = \frac{329}{329+311} = 0.514$$

Alerting Rate or Positive Rate

$$\text{ARate} = \frac{FP+TP}{FP+TP+FN+TN}$$

$$\text{ARate} = \frac{329+463}{329+463+177+311} = 0.62$$

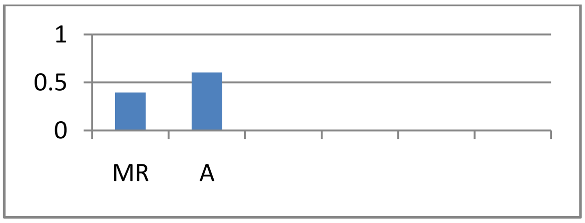

Fig 3: Bar Graph showing Misclassification Rate (MR) and Accuracy (A) of the system

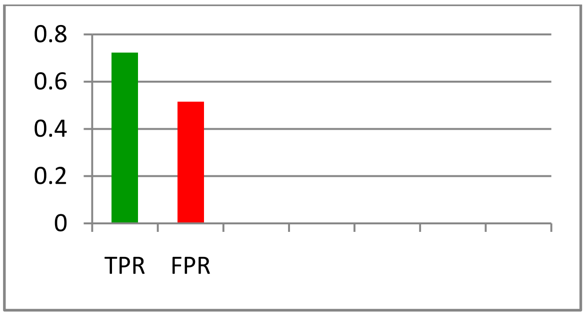

Fig 4: Bar Graph showing true positive rate (TPR) and false positive rate (FPR) of the system

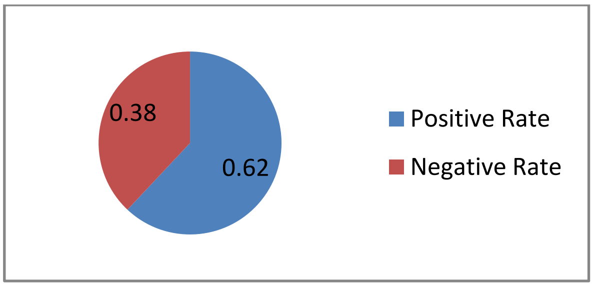

Fig 5: pie chart showing alerting rate of the system

From Above Calculation it is seen that 72% Frauds are corrected identified, accuracy of the system close to 61% .and nearly 62 % system gives alert Correctly.

Comparision with Previous Method:

1. In the Paper[1],"Credit Card fraud Detection Using Hidden Markov Model", HMM is Applied on the Credit Card transaction ,in proposed system HMM is Applied on Online Banking Transaction.

2. System Proposed in "Credit Card fraud Detection Using Hidden Markov Model" have Accuracy 80% for Observation Sequence Length 15 when Sequence Length is 100.accuracy of proposed system is 60% when Observation Sequence Length is of two.

6. CONCLUSION AND FUTURE WORK

The proposed system methodology is aimed to detect fraud in online banking Transaction. Hidden Markov Model is used to track the user behavior. First user behavior is recorded and then for new transaction it is checked. Simulation results revels that accuracy of the system 60% true Positive is close to 72% for Observation sequence length is of two. Proposed system performance is tested for observation sequence length of two.as the observation sequence increased, the complexity will also increase. Also performance of the system can be analyzed by increasing the length of observation sequence.

7. REFERENCES

[1] Abhinav Srivastava, Amlan Kundu, Shamik Sural, "Credit Card Fraud Detection Using Hidden Markov Model", IEEE Transactions on Dependable and Secure Computing, Volume-05, 2008

[2] V. Dheepa and R. Dhanapal, "Behavior Based Credit Card Fraud Detection Using Support Vector Machines", ICTACT Journal on Soft Computing, Volume-02, July 2012

[3] C. Chiu and C. Tsai, "A Web Services-Based Collaborative Scheme for Credit Card Fraud Detection," Proceeding IEEE International Conference e-Technology, e-Commerce and e-Service, pp. 177-181, 2004

[4] A. Kundu, S. Panigrahi, S. Sural, "BLAST-SSAHA Hybridization for Credit Card Fraud Detection", IEEE transactions on dependable and Secure Computing, Volume-06, October-December 2009

[5] S. Maes, K. Tuyls, B. Vanschoenwinkel, B. Manderick, "Credit Card Fraud Detection Using Bayesian and Neural networks", Proceedings of the First International NAISO Congress on Neuro Fuzzy Technologies, pp.261-270, 1993

[6] Raghavendra Patidar and Lokesh Sharma, "Credit Card Fraud Detection Using Neural Network", International Journal of Soft Computing and Engineering, Volume-01, June 2011

[7] S. Ghosh and D.L. Reilly, "Credit Card Fraud Detection with a Neural-Network", Proceedings 27th Hawaii International Conference on System Science, Volume-03, pp. 621-630, 1994

[8] Avinash Ingole, Dr. R. C. Thool, "Credit Card Fraud Detection Using Hidden Markov Model and Its Performance", International Journal of Advanced Research in Computer Science and Software Engineering, Volume-03, June 2013

[9] Vaibhav Gade, Sonal Chaudhari, "Credit card fraud detection using Hidden Markov Model", International Journal of Emerging Technology and Advanced Engineering, Volume-02, July 2012

[10] V. Bhusari, S. Patil, "Study of Hidden Markov Model in Credit Card Fraudulent Detection", International Journal of Computer Applications, Volume- 20, April 2011

[11] Sunil Mhanmane and L.M.R.J Lobo, "Use of Hidden Markov Model as Internet Banking Fraud Detecting", International Journal of Computer Application, Volume-45, May 2012

[12] Osama Dandash, Phu Dung Le and Bala Srinivasan, "Security Analysis for Internet Banking Models", Eighth ACIS International Conference on Software Engineering, Artificial Intelligence, Networking, and Parallel/Distributed Computing, pp.1141-1146, 2007

[13] K. RamaKalyani, D. UmaDevi, "Fraud Detection of Credit Card Payment System by Genetic Algorithm", International Journal of Scientific & Engineering Research, Volume-03, July 2012

[14] A. Prakash, Dr. C. Chandrasekhar, "A Parameter optimized approach for improving Credit card fraud detection", International Journal of Computer Science, Volume- 10, January 2013