Mar. 31

International Journal of advanced studies in Computer Science and Engineering
IJASCSE, Volume 4 Issue 3, 2015

Credit Card Fraud Detection Using Hmm And Image Click Point Authentication

Dinesh L. Talekar
Computer Science and Engineering Department
SSBT's COET, Bambhori
Jalgaon, India

K. P. Adhiya
Computer Science and Engineering Department
SSBT's COET, Bambhori
Jalgaon, India

***Abstract*— The credit card has become a popular mode of payment for online as well as regular purchase, as a result online fraud also increases. The credit card frauds are increasing day to day, along with that the various techniques are also developed for credit card fraud detection. The fraudsters are so talented and they generate new ways for committing fraudulent transactions on each day, which demands constant innovation for its detection techniques. Most of the techniques based on Artificial Intelligence, Machine Learning, Fuzzy Logic, Neural Network, Logistic Regression, Sequence Alignment, Decision Tree, Nave Bayesian, Bayesian Network, Meta Learning, Genetic Programming, Hidden Markov Model. The proposed system gives the solution for identification of most likely image regions to users and the user has to click on image region for creation of graphical authentication in the Image Click Point Authentication System. An Image Click Point Authentication is a sequence of points, chosen by the user in an image that is displayed on the screen. An image contains regions and the graphical authentication sequence string is generated when the user clicks on these regions. The system analyses possible attacks and blocks particular account which is being attacked.**

Keywords- Fraud Detection Techniques, Image Click Points, Hidden Markov Model, Credit Card, Fraudulent User.

I. INTRODUCTION

The bank issues debit or credit card for online purchasing. The card based purchase are categorized into two types' virtual card and physical card. In both the cases, if the card or card details are lost the fraudster can easily commit fraud transactions, which results in money loss of card holder. In online fund transfer user use the details such as login id, password and One Time Password (OTP). If the details of the credit card are misused then it gives result as increase in fraud transaction [1]. The credit card fraud is a habitual term for fraudster. The purpose of fraudster is to obtain goods without paying or to obtain unauthorized amount from an account.

A. Motivation

Nowadays, the card holder prefer the most accepted payment mode via credit card as a sophisticated way for paying bills and online shopping. At this time the risk of transaction fraud is high and hence this problem should be avoided. There are number of data mining techniques available to avoid these risks effectively. The existing research modeled the sequence of operations in credit card transaction processing using a Hidden Markov Model and shown how it can be used for the detection of frauds. The proposed work describes technique to avoid computational complexity and to provide more accuracy in fraud detection. The main motivation for graphical authentication is the possibility that people are better in remembering images than artificial words. For example, the people are recognized from thousands of faces, this fact was used to implement an authentication system. Another example is, a user could choose a sequence of points in an image as an authentication purpose which it leads to a vast number of probabilities, if the image is large and complex. Hence, the proposed system provides high security to online transactions.

B. Background

The fraud begins with either the stolen of the physical card or the hacking of data associated with the account, which includes the card account number or other information that necessarily be available to a merchant during a transaction. The compromise occurs by many common routes and usually be conducted without tipping off the card holder and the issuer, at least until the account is used for fraud. In some cases, billions of accounts have been compromised. The increasing growth of credit card use on the internet has made database security lapses particularly costly. Today the most common user authentication scheme is the alphanumeric password in computer systems [2]. Alphanumeric passwords are used widely, it's certain well known drawback is low memory ability of high information passwords. This drawback is not because of the authentication system itself but arise from the interaction between the users and the system. Since users usually cannot remember high information passwords they tends to select short or simple passwords that can be broken by dictionary attacks. Policies and mechanisms that force users to select high information passwords usually results in other unsafe practices, such as the passwords being written down and kept open. There are

multiple techniques to improve the security of user authentication, e.g., token based authentication, biometrics, graphical passwords based on the simultaneous use of two or more authentication mechanisms. Figure 1. Shows the flow of fraud detection system.

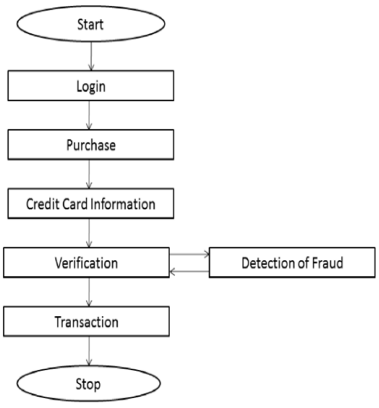


Figure 1. Flow of fraud detection system

R. Dhanpal and P. Gayathiri, presented Decision Tree algorithm is a data mining induction techniques that recursively divide a dataset of records using depth-first greedy approach [3]. A decision tree structure is made of internal nodes, leaf and root. The tree structure is used in sorting unknown data records. In this method, a credit card fraud detection using effective algorithm for Decision Tree Learning and focus is on the information gain based. This method estimates the best split of purity measures [4][5][6], entropy and information gain ratio to test the best classifier attribute. The author simply find out the fraudulent user through tracing fake mail and IP address. Customers are suspicious if the mail is fake and they traced all details about the owner/sender through IP Address.

R. D. Patel and D. K. Singh, presented the system to generate fraud transactions generated with the given sample dataset. If genetic algorithm is applied in bank for credit card fraud detection, the chance of fraud transactions predicted soon after credit card transactions is in process, and anti-fraud strategies are adopted to prevent banks from great losses before the transaction and reduces risks. The first population is randomly selected from the sample space which has many populations. The fitness value is calculated in every population and is separated. The process of selection is perform through tournament method [7]. The single point probability used to calculated crossover. The mutation mutates the new coming offspring using uniform probability measure [8] [9]. The best solutions are passed to the further generation. The new population is generated and undergo the same process, if maximum number of generation is reached then fraud is detected.

A. Srivastava and A. Kundu, presented a HMM is a double embedded stochastic process with two hierarchy levels. It is complicated stochastic processes as compared with traditional Markov Model. A Hidden Markov Model (HMM) has a finite set of states monitored by a set of transition probabilities. In a particular state, observation or an output generated according to an associated probability distribution. It is only the output and not the state that is visible to an external observer. HMM uses cardholders spending behavior to detect fraud [10]. In implementation, there are three behaviors of cardholder are taken into consideration, Low spending behavior, Medium spending behavior, High spending behavior. Different users has their various spending behavior (low, medium, high). Low spending behavior of any user shows that cardholder spend low amount (L), medium spending behavior of any user shows that user spend medium amount (M), high spending behavior of any user shows that cardholder spend high amount (H). Three clusters are created using clustering algorithm and clusters represents observation symbols. The clustering probabilities are calculated for each cluster, which is percentage of number of transactions in every cluster then calculate fraudulent transactions [11]. In maximum cases, valid user is considered as fraudulent.

II. METHOD

The existing system consist some drawbacks and proposed solution overcome this drawbacks by using ICPA (Image Click Point Authentication) technique. An Image Click Point Authentication is a sequence of points, chosen by the user in an image that is displayed on the screen. An image contains regions and the graphical authentication sequence string is generated when the user clicks on these regions. The system analyses possible attacks and blocks particular account which is being attacked. In proposed system, the need of bank authorization to create user, register credit card etc. Hence these assumptions are consider in proposed system. Along with banking side, assume online purchasing and payment delivery. The facts are assumed (1) User register in bank (2) Online purchasing.

A. Architecture

The architecture is a system that unifies its components or elements into a coherent and functional blocks. The architecture shows the structure of system. The architecture of credit card fraud detection system is shown in Figure 2 and Figure 3. The Architecture Consist of two parts: 1) Administrator (Bank Part), 2) Card Holder (User Part).

Administrator (Bank Part): The administrator is responsible for register credit card holders or user with details. This part consist of registration of user's credit card, transaction details, user behavior, and blocked status of user. Fig2 shows administrator side structure. Administrator side consist of some functional block. Admin accepts some basic details of user for registration part-I like, Credit card number, Name of user, Address, E-Mail ID, Mobile number, Pin code. While accepting credit card number, system checks if it is existing or not, if it is existing, system gives alert and not allowed for duplicate registration that means single credit card register at single time. Once credit card number is registered, same number is not allowed for registration by the

system. Administrator also checks behavior of user along with all transactions of every user. Transactions are displayed with line graph. For every user, generate separate line graph on the basis of existing transaction. User behavior shows the status of user. The status is define after validation of HMM. The behavioral status are three parts Low, Medium and High. The system displays blocked and unblocked status of all user. When a fraud is detected, the system immediately block respective users, Hence users are not allowed to perform any transaction. On the other hand, administrator has a rights to activate and deactivate the users.

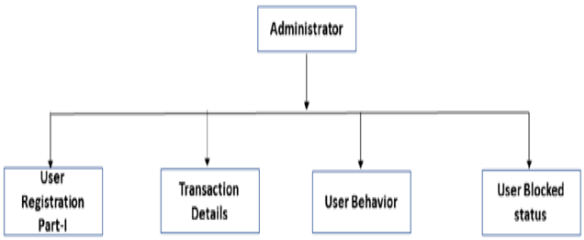
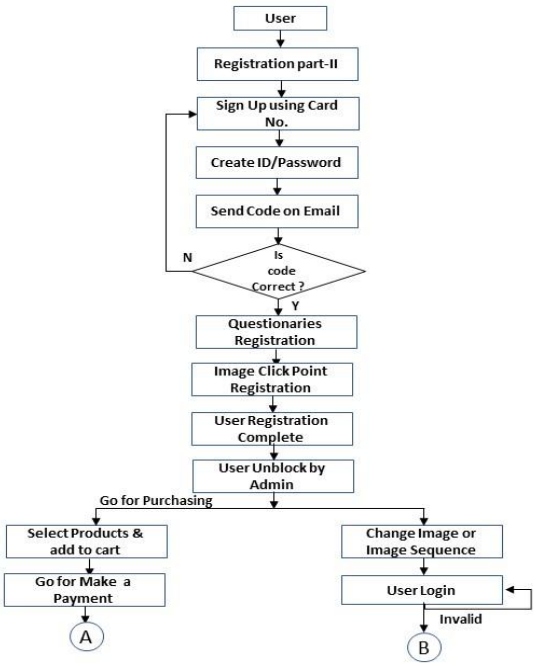


Figure 2. Architecture of Proposed Method for Credit card Fraud Detection at Admin Part



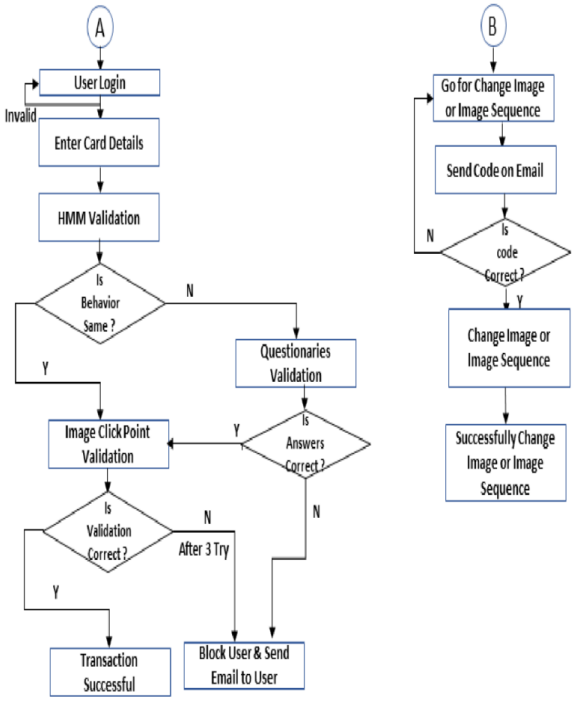


Figure 3. Architecture of Proposed Method for Credit card Fraud Detection at User Part

Card Holder (User Part): Figure 3. Shows user side structure, the part of user is part of Architecture, in which user can complete remaining registration and purchase the required. Once the user has entered in this part, basic details are filled by the system which is already entered by administrator. User only needs to create user id and password. The system accepts user id and password and send 4 digit code on registered email id for more security purpose because email id is more secure than mobile number. If 4 digit code is matched then system identifies the user as authorized user then questionaries' are provided to the user. These questionaries' are occurred when the change in behavior is found by HMM. After questionnaires', system asks Image Click Point Authentication (ICPA), in which user needs to select maximum four objects on an image and make the sequence. In this way user completes the registration. At the end of registration, user has three authentications. While user performing transaction, Hidden Markov Model (HMM) work effectively. The system accept transaction amount and HMM find spending profile by checking last transactions. Every incoming transaction is pass to the HMM for verification. The system receives the card details and the value of purchase a goods to verify whether the transaction is genuine or not. The types of product that are bought in that transactions are not known to the system. It tries to find any fraud in the transaction based on the spending profile of the cardholder. If the system confirms the transaction to be

fraud, it raises an alarm, and asks for next module. To check and map the credit card transaction processing operation in terms of Hidden Markov Model, first decides the observation symbols in model. The values purchase x into M price ranges V1, v2,...,Vm, creating the observation symbols at the issuing bank. The price range for each symbol is configure based on the spending habit of individual cardholders. The price ranges determined dynamically by applying a clustering algorithm (K-means algorithm) on the values of each cardholder's transactions. Here use Vk, k=1, 2, 3. M to represent the observation symbol and corresponding price range. Consider three transactional price ranges, namely, low (l), medium (m), and high (h). So set of observation symbols is, therefore, V= l, m, h making M = 3. For example, let l= (0 < x < 50000), m= (50000 < x <100000) and h= (100000 < x < 200000); where x be transaction value. If a cardholder performs a transaction of 59000, then the corresponding observation symbol is m. Spending profiles of user are determined at the end of the K-means clustering step which is shown in Equation 1. Let pi be the percentage of total number of transactions of the users that belong to clusters. Then, the spending profile (SP) of the user is determined as follows:

$$SP = MAX_i (P_i) \quad (1)$$

Where Pi: Percentage of number of transactions
SP: Cluster number to which most of the transactions

The sequence of transactions are provide to HMM and find out change in user behavior. Let O1, O2, O3, OR, one such sequence of length R. before new transaction initial behavior are set. First provide the initial sequence to Hidden Markov Model and compute the probability is shown in Equation (2) and define a behavior. Let the probability α1 which can be written as follows:

$$\alpha_1 = P(O_1, O_2, O_3....O_R|\lambda) \quad (2)$$

For new transaction Let OR+1 be the new symbol generated by a new transaction at time t+1, with length R. so new sequence is O1, O2, O3, OR, OR+1. Let new probability α2 is in Equation (3).

$$\alpha_1 = P(O_1,O_2,O_3....O_{R+1}|\lambda) \quad (3)$$

Now, check the any change are occurs in users spending profile behavior. Let, α be a difference of α1 and α2.

$$\alpha = \alpha_1 - \alpha_2 \quad (4)$$

Equation (4) calculate clusters symbol and find out the maximum number of transactions in clusters. In this process new cluster is compare with existing clusters. If α>0, it means that the new sequence is accepted by the HMM, and it could be a fraud. The symbol α denote change in user behavior are occurs, that means it is fraudulent transaction or second think is that, the valid user try to performing variation in transaction. Hence can't say that, it is definitely fraud. The proposed work is minimize same cases and implement new extended system. The Proposed system is provide next step for better security. The next step is Image click point validation and questionnaires. The Image Click Points Authentication (ICPA) is a graphical authentication method. The graphical authentication method are consists of click points (3 to 4 click points) sequences, which is chosen by user. The image is displayed on the screen by the system. The displayed image is helping to the user, remember the click points. The click point pixel in an image is a candidate for a click points. In the authentication process, the user has to click again on the chosen points. Hence it is almost possible for human users to click repeatedly on exactly the same point. As per studies on graphical attention and eye movements shows that, most of the images contain a few portions that most humans focus on. When asked to create a graphical authentication a user would probably not click on all available pixels, but only focus on some specific areas.

In the ICPA method, user has to select maximum any four objects in given image. The selected objects shown as a selected regions, the selected regions defines the probability of mouse click position. This model is defined probable regions with their pixel values. The system is create graphical authentication by calculating click point regions. The selected regions are stored in database by its name with sequence number, the sequence number shows specific object name in sequentially. The object names are useful when user forgot the password. The user can retrieve forgot sequence by matching verification code, which is send on registered email id by the system. When user select any object in the image, the system give the sequence number 1, while selecting second object system gives sequence number 2 in this way user select different objects and sequence numbers are provided automatically by the system. The name of object selected as per selection of region and the object positions are detected using given equation.

$$d_{posx} = (Px_{event} - I_{left}) \quad (5)$$

$$d_{posy} = (Py_{event} - I_{top}) \quad (6)$$

Equation (5) is used to calculate X position of mouse click point and Equation (3.6) is used to calculate Y position of mouse click point. Now, d_{posx} and d_{posy} contain X and Y coordinate of current click point, there for this coordinate map with ICPA method and define sequence number with object name. This object name are stored as per sequences for future authentication.

Image Click Point Authentication steps are:

a) Selection of Image
b) Select object on image
- While selecting object defining region values
- As per region values define name of object

c) Total seven object define in image
d) At run time calculate click points with X and Y Coordinate values
e) X and Y value are compare with existing defined Model

f) If click point values are between defined regions then assign a sequence number and name of object this step continues up to selecting four objects

g) Finally selected sequence are stored in database in the form of object name

h) At the time of transaction database string are compare with new generated string

i) If it is unmatched fraud get detected

Hence, ICPA method detects the fraudulent transactions and give more and more security to card holder.

III. DESIGN ALGORITHM

The proposed algorithm described the catching click points on image object. The Image Click Point Authentication (ICPA) module asks to card holder after HMM validation that means behavior of user is not changed. The HMM checks the behavior of user using existing transactions. If the card holder's found in change behavior then system asks questionnaires for confirmation and if answers are correct then system allows to Image Click Point Authentication (ICPA) module. Once user come an Image Click Point Authentication (ICPA) the system initialize counter to 1. The user has to select any four object in given image. Algorithm firstly find the click point locations and then check with existing model. Existing model consist of specific object regions. All object region map with pixel values and then define the object name. On every selection object gives sequence number. The sequence number automatically incremented by one for next object selection, every object selection gives object name. Finally all object names are appended one another one and make a string and the string is stored in database as an authentication string. This authentication string match with new string, which is generated at the time of every Image Click Point Authentication (ICPA) for particular user. If string are not matched, then authentication failed and user is block.

The ICPA work on user's activity. The system gives response as per user activity. The given algorithm accept maximum four click points. The given algorithm shows the steps and conditions for accepting a click points. In first step algorithm calculate the click pixel location using Equation (5) calculate the X coordinator and Equation (6) calculate Y coordinate. Hence exact click positions are find. The calculated X and Y coordinate are checked with defined pixel region value. In defined model specific region called an object, there are seven object in defined model. An object having specific name, according to object name graphical authentication string is generated. If newly calculated X and Y coordinate pixel value found in defined regions values then the specific object name are assign as a first click point object. The value of counter i is increases by one. In this way algorithm checks X and Y coordinate of newly clicks with defined objects one by one till counter i=4. The sequence number is used to define the sequence and make the object name sequence as per click sequences. Each object name append one another one sequentially. Finally object name sequence are crated and stored in database. The final validation are completed using stored string with newly generated sequence and decide whether fraudulent transaction or not. Hence algorithm 1 is used for accept the click points in given image and make a string sequence of object name.

Algorithm 1 Image Click Point Algorithm

Require: Initialize i=1 for no. of points selection

Ensure: HMM validation begin

```
if i < 5 then
    Calculate X and Y value of click point.....Equation 5 and 6
    if X and Y are pixel of first object then display object marked
    image value=object1 sequence no= 1, i=i+1
    else if X and Y are pixel of Second object then display
        object marked image value=object2 sequence no= 2
        i=i+1
    else if X and Y are pixel of Third object then display
        object marked image value=object3 sequence no= 3
        i=i+1
    else if X and Y are pixel of Fourth object then display
        object marked image value=object4 sequence no= 4
        i=i+1
    else if X and Y are pixel of Fifth object then display
        object marked image value=object5 sequence no= 5
        i=i+1
    else if X and Y are pixel of Sixth object then display
        object marked image value=object6 sequence no= 6
        i=i+1
    else if X and Y are pixel of Seventh object then display
        object marked image value=object7 sequence no= 7
        i=i+1
    else
        No. of object selection is over end if
end if
end
```

The final validation are done using check stored sequence string with newly generated sequence, and check whether fraudulent transaction or not. Algorithm 1 is used for accept the click points in given image and make a string sequence of object name. Finally this string check with database saved string.

IV. RESULT

Experimental result shows the efficiency of proposed system, in which consider the some number of transaction from dataset. At initial condition dataset are filled by some user's activity. Initially clusters are created as per transaction values and then behavior of users calculated. As per behavior, user profile is set. According to profile system handles the users and allow or block the users. In Table 1 'Number of Transactions' shows the total 15 transactions, cluster algorithm helps to create a cluster. According to

transactions, system defines user behavior and user behavior are calculated by Hidden Markov Model.

Fig. 4. Shows the clusters for given transactions, in which x axis shows the number of transaction and y axis shows transaction amount and bubble dots shows the clusters. According to transactions, system defines user behavior and user behavior are calculated by Hidden Markov Model.

TABLE I. NUMBER OF TRANSACTIONS

Transaction No.	Transaction Amount
1	1300
2	1300
3	48000
4	24000
5	100000
6	13000
7	13000
8	48000
9	13000
10	127000
11	4000
12	12000
13	12000
14	150000
15	12000

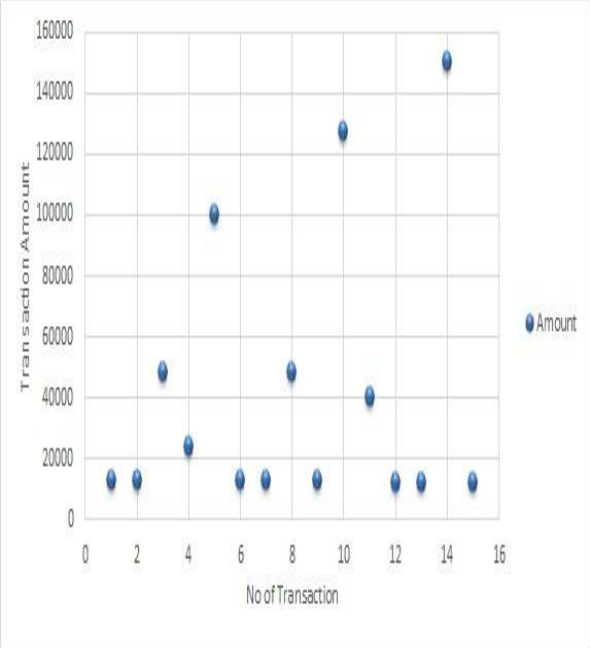


Figure 4. Data clustering

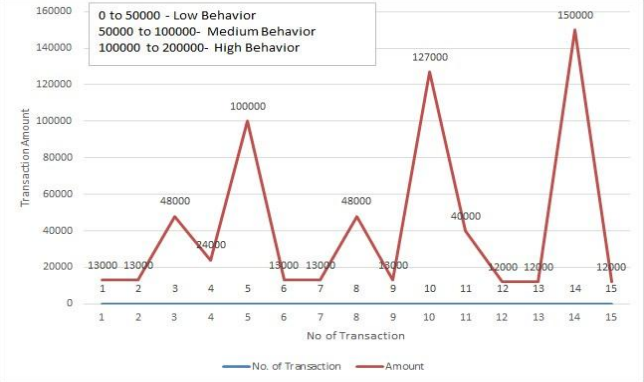


Figure 5. Behavioural result analysis from Table 1

Figure 5. Shows that behavior of user along with transactions which is shown in Table 4.6. If transaction amount is between 0 to 50000 then generate 'Low' symbol, if transaction amount is between 50000 to 100000 then generate 'Medium' symbol, if transaction amount is between 100000 to 200000 then generate 'High' symbol. Hence, finally behavior is defined for respective user.

Table 2. 'Transactions Status with Overall Observations' shows the number of transaction per user with all statistical states, in which some observations like, how much number of times user entered in change in behavior. In some cases, if upcoming transaction is differ from existing behavior then user must face the questionnaires and Image Click Point Authentication. In the image click points if user cannot fallow the correct sequences then system gives three attempts, after three attempts user is blocked. 'True positive' parameter shows the successive transactions and 'False positive' shows unsuccessful transactions or user block status. By analyzing all transactions success rate of true positive transactions is much greater than existing systems.

Sr. No.	No. of Transaction	Change In behaviour	No. of time Questions Ask	No. of times Image click points	True positive	False positive	success rate in percentage
1	10	2	2	15	8	2	80.00
2	12	3	3	14	10	2	83.33
3	11	2	2	15	11	0	100.00
4	9	1	1	13	6	3	66.67
5	10	3	3	20	8	2	80.00
6	10	1	1	16	7	3	70.00
7	14	2	2	20	12	2	85.71
8	15	1	1	25	11	4	73.33
9	13	1	1	23	12	1	92.31
10	12	2	2	17	12	0	100.00

V. DISCUSSION

Image click point Authentication (ICPA) increases the transaction security and block fraudulent transactions immediately before processing of payments. ICPA method provides maximum security. On every transaction, system is

updated and take valid decision as per user's behavior. According to observations, users are more secure and performs faithful transactions.

In existing system, the fraud is detected only on the basis of change in behavior of user. In this case, if valid user performs wrong transaction then change in behavior occurs and user is blocked immediately. This problem is solved in the proposed system by providing three level security to identify a valid user. The proposed system gives three attempts to the user to confirm the validity, and hence the proposed system is superior to existing system. Nowadays, HDFC bank gives similar kind of security in the form of image identification. But there is no any secret point and once image is watched by anyone then it is easy to identify an image. According to survey, there is no such type of three level security. Many banking sectors uses One Time Password (OTP) using mobile number for final verification but problem is that if mobile numbers are not in network, the massages may be diverted on another number by fraudulent. Hence proposed work uses email id for verification, system sends four digit OTP code on registered email id. In this way proposed solution is better than others.

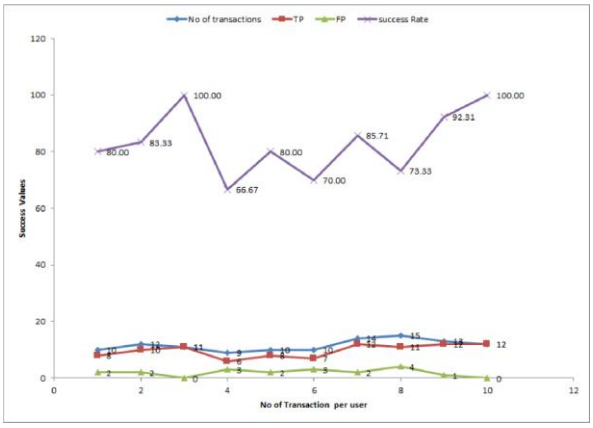


Figure 6. Transaction with success rate

Figure 6. Show that True Positive rate is more than false Negative rate for every user, so that success rate is more than existing system.

VI. CONCLUSION

The credit card fraud detection system gives four level security by using Hidden Markov Model and Image Click Point Authentication. Proposed system solves drawbacks of existing system i.e. inaccurate results, user behaviors based security, no secret authentication and no strong transaction checking. The proposed system does not blocked a valid user and faithful transaction with authentication facility carried out through email id. The calculating performance of proposed system is handled by different users and created dataset. The user updated dataset is observed and define observations. The observations are defined on the basis of some parameters like number of users, number of transactions, change in behaviors and number of true or false transactions. According to observations, system provide 80 to 95 percent security for transactions. The maximum transactions becomes true transactions but only the drawback is, user is harassed due to more security levels but it is negligible for strong security. The system allows the user to change the click point sequences or to change images for new sequences hence security also increases. The proposed system consist of define images, user can not add new images, and cannot create new click point sequences. In the future work, it is possible to give an authentication to the user to add new images and update it with new click point's sequences.

REFERENCES

[1] V. Bhusari and S. Patil, "Study of hidden markov model in credit card fraudulent detection", International Journal of Computer Applications, vol. 2, no. 5, 2011.

[2] V. K. Prasad, "Method and system for detecting fraud in credit card transaction", International Journal of Innovative Research in Computer and Communication Engineering, vol. 1, no. 5, 2013.

[3] R. Dhanpal and P. Gayathiri, "Credit card fraud detection using decision tree for tracing email and ip", International Journal of Computer Science Issues, vol. 9, no. 2, 2012.

[4] Y. Sahin and E. Duman, "Detecting credit card fraud by decision trees and support vector machines", Proceeding of the International Multi Conference of Engineers and Computer Scientist, vol. I, 2011.

[5] V. M. Rao and Y. P. Singh, "Proceeding of the international conference on artificial intelligence in computer science and ict", International Journal of Advanced Research in Computer and Communication Engineering Organized by WorldConferences.net, 2013.

[6] T. Minegishi1 and A. Niimi, "Proposal of credit card fraudulent use detection by online type decision tree construction and verification of generality", International Journal for Information Security Research (IJISR), vol. 1, no. 4, 2011.

[7] R. D. Patel and D. K. Singh, "Credit card fraud detection prevention of fraud using genetic algorithm", International Journal of Soft Computing and Engineering (IJSCE), vol. 2, no. 6, 2013.

[8] K.RamaKalyani and D.UmaDevi, "Fraud detection of credit card payment system by genetic algorithm", International Journal of Scienti_c Engineering Research, vol. 3, no. 7, 2012.

[9] S. Vats, S. K. Dubey, and N. K. Pandey, "A tool for effective detection of fraud in credit card system", International Journal of Communication Network Security, vol. 2, no. 1, 2013.

[10] A. Srivastava and A. Kundu, "Credit card fraud detection using hidden markov model", IEEE Transactions on Dependable and Secure Computing, vol. 5, no. 1, 2008.

[11] A. Singh and D. Narayan, "A survey on hidden markov model for credit card fraud detection, "International Journal of Engineering and Advanced Technology (IJEAT), vol. 1, no. 2, 2012.