CHAPTER 6

A PARAMETER OPTIMIZED APPROACH FOR IMPROVING CREDIT CARD FRAUD DETECTION

The usage of credit cards has highly increased due to high-speed innovation in the electronic commerce technology. Since credit card turns out to be the majority well-liked manner of payment for mutually online as well as habitual purchase, cases of fraud correlated through it are as well increasing. In normal Hidden Markov Model the problem of cannot find an optimal state sequence for the underlying Markov process also this observed sequence cannot be viewed as training a model to best fit the observed data. In this research, the main aim is to model the sequence of observations in credit card transaction processing using an Advanced Hidden Markov Model (AHMM) and show how it can be utilized for the exposure of frauds. In this process an AHMM is initially trained with the regular manners of a cardholder. If an incoming credit card transaction is not recognized by the trained AHMM with adequately high probability, it is believed to be fraudulent. This proposed work desire to regulate the model parameters to best fit the observations. The ranges of the matrices (N and M) are fixed but the elements of A, B and π are to be decided, focus to the rank stochastic condition. The information that can efficiently re-estimate the model itself is one of the more incredible features of HMMs referred here as AHMM.

6.1 INTRODUCTION

An unauthorized account movement by a person for whom the account was not be set to can be referred as credit card fraud. Preparedly, this is an event for which action can be taken to discontinue the misuse in steps forward and integrate risk
executive applies to defend alongside comparable acts in the future. In straightforward expressions, Credit Card Fraud is described as when an individual exploits another individual’s credit card on behalf of personal causes though the proprietor of the card and the card issuer are not conscious of the information that the card is being used. In addition to the persons using the card has not at all having the association with the cardholder or the issuer and has no purpose of making the repayments for the acquire they done. The anticipate user behavior in economic systems can be utilized in many situations.

Forecasting client relocation, public associations can accumulate a lot of wealth and other assets. One of the most motivating pastures of forecast is the fraud of credit stripes, especially credit card expenditure. Positively, all transactions deals with financial records of known abuse are not authoritative. However, there are transactions which are officially suitable, but knowledgeable people can advise that these transactions are probably misused, caused by stolen cards or fake merchants. So, the assignment is to avoid a fraud by a credit card transaction previous to it is known as “illegitimate”. By means of growing number of transactions people can no longer manage all of them. As a solution, one might hold the experience of the experts and put it interested in an expert system. This habitual approach has the disadvantage that the expert’s knowledge, yet as it can be mined unambiguously, alters quickly with novel manners of prepared attacks and models of credit card fraud. So as to keep track with this, no predefined fraud models but routine learning algorithms are needed.

HMM-based applications are ordinary in an assortment of areas such as speech recognition, bioinformatics, and genomics. Ourston et al have projected the application of HMM in identifying multistage credit card attacks. Hoang et al
projected a new method for abnormality exposure using HMM. The main idea is to construct a multilayer model of transaction behaviors based on HMMs and specifying methods for anomaly detection. In recent days, Joshi and Phoba have examined the capabilities of HMM in credit card fraud detection. Cho and Park recommended an HMM-based credit card fraud detection system that advances the time to model and routine by allowing for only the right transition streams based on the province knowledge of assaults. Lane has examined HMM to model human behavior. On one occasion human behavior is properly formed, any sensed departure is a reason for concern because an attacker is not predictable to have a behavior similar to the genuine user. For this reason, in this research work the Advanced Hidden Markov Model is formed to find the credit card fraud detection. In HMM one more drawback is that for the dynamic programming approach the optimal observation sequence would not be found. With this the best fit point should modeled. In AHMM the drawbacks will be resolved with $\alpha$ and $\beta$ value. The contribution of the works as follows:

1. In training phase obtain the card holder profile and calculate the probability for each transaction.

2. Using the AHMM creates the observation model with best fit observation states and regulates the model parameters ($\alpha$ and $\beta$) to best fit the observations.

3. In testing phase the detection of fraud is obtained If both probability value from multiple observation are same it will be a normal customer else there will be fraud signal will be provided.

**6.2 CREDIT CARD FRAUD DETECTION USING AHMM**
The credit card fraud detection system is based on Hidden Markov Model, which does not require fraud signatures and still it is capable to perceive frauds just by bearing in mind a cardholder’s spending habit. The specifics of purchased items in single transactions are generally unidentified to any Credit card Fraud Detection System organization either at the bank that issues credit cards to the cardholders or at the commercial site where goods is going to be obtained. As business processing of credit card fraud detection system runs on a credit card issuing bank site or merchant site. Every arriving transaction is submitted to the fraud detection system for verification intention. The fraud detection system recognize the card details such as credit card number, CVV number, card type, expiry date and the amount of items acquire to validate, whether the transaction is genuine or not.

The accomplishment techniques of Hidden Markov Model in order to notice fraud transaction through credit cards, it generate clusters of training set and identify the spending profile of cardholder. In that process the number of items purchased by customers, types of items that are bought in a particular transaction deliberates on the amount of item acquired and use for further processing that are not known to the Fraud Detection system completely. It supplies higher amount of dissimilar data transactions in form of clusters depending on transaction amount which will be moreover in low, medium or high value assortments. It tries to discover out any discrepancy in the transaction based on the spending behavioral profile of the cardholder, shipping address, and billing address and so on.

Based on the expenditure behavioral profile of card holder the probabilities of initial set have been selected and bring together a series for additional processing. If the fraud detection system generates sure that the transaction to be of fake, it raises an
alarm and the issuing bank refuses the transaction. For the protection purpose, the refuge information module will get the information features and its store’s in database. To identify the safety measures information if the card missing then the security information module structure arises.

The security form has a number of safety questions like account number, date of birth, mother name, other personal question and their answer, etc. where the abuser has to respond it correctly to move to the transaction division in which all those information must be known by the card holder only and can continue only by the card holder. It has informational confidentiality and informational self strength of mind that are tackled consistently by the novelty giving people and entities a trusted means to user, protected, search, process, and exchange personal and/or secret information.

The system and tools for pre-authorizing commerce offered that a relations tool to a trader and a credit card proprietor. By communicating to a credit card number, card type with expiry date and storing it into database, a unique portion of information that describes a fastidious transaction to be complete by a trustworthy user of the credit card at a later occasion the cardholder will be initiating a credit card transaction procedure. The particulars are conventional in the type of system data in the database only when if a correct individual recognition code is used with the statement the cardholder can precede with further steps with the credit card. Because the transaction is pre-authorized, the merchant does not require observing or transmitting an accurate individual recognition code.

1. The number of states in the model is N. The set of states is \( S = \{S_1, S_2, \ldots, S_N\} \), where \( S_i \), \( i = 1, 2, \ldots, N \) is an individual state. The state at time instant t is denoted by \( q_t \).
2. The number of distinct observation symbols per state is M. The set of symbols is 
\[ V = \{V_1, V_2, \ldots, V_M\}, \text{ where } V_i, i = 1; 2; \ldots; M \text{ is an individual symbol.} \]

3. The state transition probability matrix \( A = \{a_{ij}\} \) where \( a_{ij} = P(q_{t+1} = S_j \mid q_t = S_i); 1 \leq i \leq N; 1 \leq j \leq N; t = 1; 2; \ldots; \) where \( a_{ij} > 0 \) for all \( i, j \).
   Also, \( \sum_{j=1}^{N} a_{ij} = 1, 1 \leq i \leq N. \)

4. The observation symbol probability matrix \( B = \{b_j(k)\} \), where \( b_j(k) = P(V_k \mid S_j), 1 \leq j \leq N, 1 \leq k \leq M \) and \( \sum_{k=1}^{M} b_j(k) = 1, 1 \leq j \leq N \)

5. The initial state probability vector \( \pi = \{\pi_j\} \), where \( \pi_j = P(q_1 = S_j), 1 \leq j \leq N, \) such that \( \sum_{k=1}^{M} \pi_j = 1 \)

6. The observation sequence \( O = O_1, O_2, O_3 \ldots O_R, \) where each observation \( O_t \) is one of the symbols from \( V \), and \( R \) is the number of observations in the sequence.

6.3 HMM Model for Credit Card Transaction Processing

First begin through deciding the observation symbols in our model which is to record the credit card transaction processing function in terms of an HMM. Then quantize the acquisition values \( x \) into \( M \) worth ranges \( V_1, V_2, \ldots, V_M \) structuring the observation symbols at the issuing depository. The concrete outlay range for every symbol is configurable based on the expenditure routine of individual credit card holders. These worth ranges can be found dynamically through applying a clustering method on the values of each cardholder's transactions.

Let assume \( V_k, k = 1, 2, \ldots, M \) to stand for both the observation symbol as well as the equivalent charge assortment. A credit cardholder constructs diverse types of purchases of unlike amounts more than a period of time. Single prospect is to believe the sequence of transaction amounts and look for divergences in them.
On the other hand, the sequence of kinds of purchase is additional constant contrasted to the series of transaction quantities. The motive here is that, a cardholder precedes purchases depending on his require for procuring diverse types of items greater than a period of time. Consecutively, produces a series of transaction quantities. The kinds of each acquire are linked to the row of business of the equivalent trade. The kind of purchase of the cardholder is hidden from the FDS. The position of all probable categories of purchase and consistently, the position of every one potential lines of commerce of merchants structures the position of concealed states of the HMM.

The line of business of the commercial is identified to the acquiring bank which should be noted at this stage that, since this information is furnished at the time of registration of a merchant. As well, a number of merchants might be trade in various types of merchandise. Such kinds of line of business are judged as Miscellaneous and there is no need to determine the authentic types of items purchased in these transactions.

A few assumptions as regards accessibility of this information with the issuing depository and therefore with the FDS are not matter-of-fact hence, would not have been suitable. In the consequences part shows the cause of choice of the number of states on the method performance. Subsequent to deciding the state and symbol illustrations, after that have to find out the probability matrices A, B and $\pi$ thus the representation of the HMM is inclusive. These three model parameters are found in a training phase. Hence, they should be chosen carefully through preliminary selection of parameters influences the performance of the algorithm.
A method based on Hidden Markov Models (HMMs) is a stochastic method, which can be very useful for some applications involves the making of auditory models of program that build use of temporal information. The HMMs can model a lesser unit of the statement. The HMMs can be analyzed as fixed state machines, wherever every unit of time, a state transition happens, and all state produces an auditory vector with a connected likelihood density function. So as to is, in every state, a GMM (Gaussian mixture model) is second-hand to exemplify an auditory vector experiential.

AHMM FOR CREDIT CARD FRAUD DETECTION

Here to alter the model parameters to best fit the observations. The ranges of the matrices (N and M) are fixed however the elements of A, B and \( \pi \) are to be

<table>
<thead>
<tr>
<th>SYMBOL</th>
<th>REPRESENTATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>( T )</td>
<td>Observation sequence length</td>
</tr>
<tr>
<td>( N )</td>
<td>Number of states in the model</td>
</tr>
<tr>
<td>( M )</td>
<td>Number of observation symbols</td>
</tr>
<tr>
<td>( O )</td>
<td>Observation sequence ( (o_0, o_1, \ldots, o_{T-1}) )</td>
</tr>
<tr>
<td>( Q )</td>
<td>Markov process for distinct states ( {q_0, q_1, \ldots, q_{N-1}} )</td>
</tr>
<tr>
<td>( V )</td>
<td>Set of possible observations ( {0, 1, \ldots, M - 1} )</td>
</tr>
</tbody>
</table>
Probability for each state transition

\[ A \]

\[ \pi \]

Probability matrix of observation sequence

determined, focus to the strip stochastic condition. The actuality that can professionally re-estimate the model itself is one of the more astonishing aspects of HMMs. Let assume \( \lambda = (A, B, \pi) \) be a given model and series of observations \( O = (O_0, O_1, ..., O_{T-1}) \). For \( t = 0, 1, ..., T-2 \) and \( i, j \in \{0, 1, ..., N-1\} \), define “\( di-gamma \)” as

\[ \gamma_t(i, j) = P(x_t = q_i, x_{t+1} = q_j | O, \lambda) \]

\[ \gamma_t(i) = \frac{\alpha_t(i)\beta_t(i)}{P(O | \lambda)} \]

Then \( \gamma_t(i, j) \) is the probability of being in state \( q_i \) at time \( t \) and transiting to state \( q_j \) at time \( t + 1 \). The di-gamma will be formed with the terms taken as \( \alpha, \beta, A \) and \( B \) as:

\[ \gamma_t(i, j) = \frac{\alpha_t(i)a_{ij}b_j(O_{t+1})\beta_{t+1}(j)}{P(O | \lambda)} \]

In this we should re-estimate this with the parameter \( \beta_t(i) \) which measures the relevant probability after time \( t \)

\[ \beta_t(i) = \frac{\gamma_t(i)P(O | \lambda)}{\alpha_t(i)} \] that is also represented as:

\[ \beta_t(i) = \sum_{j=0}^{N-1} a_{ij}b_j(O_{t+1})\beta_{t+1}(j) \]

Where \( \beta_t(i) = P(O_{t+1}, O_{t+2} ... O_{T-1} | x_t = q_i, \lambda) \)
Denote the $\beta_t(i,j) = P(x_t = q_i, x_{t+1} = q_j | O_{t+1}, \lambda)$, define the di-Betas as

$$
\beta_t(i,j) = \frac{\alpha_t(i) a_{ij} b_j(O_{t+1}) \gamma_{t+1}(j)}{P(O_{t+1} | \lambda)}
$$

Where $\beta_{t+1}(j) = \frac{\gamma_{t+1}(j) P(O_{t+1} | \lambda)}{\alpha_{t+1} a_{ij} b_j(O_{t+1})}$. The $P(O_{t+1} | \lambda)$ is obtained by summing $\alpha_{T-1}(i)$ over $i$. From the definition of $\beta_t(i)$ it follows the most likely state at time $t$ is the state $q_i$ for which $\beta_t(i)$ is maximum, where the maximum is taken over the index $i$.

$\beta_t(i)$ and $\beta_t(i,j)$ are related by

$$
\beta_t(i) = \sum_{j=0}^{N-1} \beta_t(i,j)
$$

Given with the $\beta$ and di-Betas verify the model $\lambda = (A,B,\pi)$ can be re-estimated as follows:

1. For $i = 0, 1, ..., N - 1$

2. For $i = 0, 1, ..., N - 1$ and $j = 0, 1, ..., N - 1$ compute

$$
a_{ij} = \frac{\sum_{t=0}^{T-2} \beta_t(i,j)}{\sum_{t=0}^{T-2} \beta_t(i)}
$$

The numerator of re-estimated $a_{ij}$ can be observed to give the supposed number of transitions from state $q_i$ to state $q_j$ and the denominator denotes the expected number of transition from the state $q_i$ to any state. Then the ratio is the probability of transiting as of state to state $q_i$ state $q_j$, which is the desired value of $a_{ij}$.

3. For $j = 0, 1, ..., N - 1$ and $k = 0, 1, ..., M - 1$ compute
The numerator of the re-estimated $b_j(k)$ is the anticipated number of times the model is in state $q_j$ with observation $k$, at the same time as the denominator is the estimated number of times the model is in state $q_j$. The ratio is the probability of observing symbol $k$, given that the model is in state $q_j$, which is the desired value of $b_j(k)$.

Re-estimation is an iterative process. Foremost, we initialize $\lambda = \left( A, B, \pi \right)$ through a best guess or, if no logical guess is obtainable, choose with arbitrary values such that $\pi_i \approx 1/N$ and $a_{ij} \approx 1/N$ and $b_j(k) \approx 1/M$. It's vital that $A$, $B$ and $\pi$ be randomized, because precisely consistent ideals will consequence in a confined maximum from which the model cannot Hill climb. As constantly, $\pi$, $A$ and $B$ must be row stochastic. The AHMM process can be summarized as follows.

1. Initialize the model, $\lambda = \left( A, B, \pi \right)$

2. Evaluate $\alpha_t(i)$, $\gamma_t(i)$, $\beta_t(i)$ and $\beta_t(i,j)$

3. Re-estimate the model $\lambda = \left( A, B, \pi \right)$.

4. If $P(O|\lambda)$ increases, goto 2.

Certainly, it may be enviable to end if $P(O|\lambda)$ does not increase by at any rate various predestined threshold and/or to locate a maximum amount of iterations.

6.5 EXPERIMENTAL RESULTS
The performance of the proposed approach is evaluated by using the credit card transaction data set using the software MATLAB.

The performance of the proposed algorithm is based on the following factors:

- Precision
- Recall
- F-Measure

### 6.5.1. Precision Comparison

This graph shows the precision rate of existing and proposed system based on two parameters of precision and the number of Dataset. From the graph we can see that, when the number of number of Dataset is advanced the precision also developed in proposed system but when the number of number of Dataset is improved the precision is reduced somewhat in existing system than the proposed system. From this graph we can say that the precision of proposed system is increased which will be the best one. The values are given in Table 6.5.1

**Table 6.5.1 Number of Dataset vs. Precision**

<table>
<thead>
<tr>
<th>SNO</th>
<th>Number of Dataset</th>
<th>AHMM</th>
<th>HMM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
<td>0.32</td>
<td>0.21</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>0.62</td>
<td>0.55</td>
</tr>
<tr>
<td>3</td>
<td>30</td>
<td>0.73</td>
<td>0.67</td>
</tr>
<tr>
<td>4</td>
<td>40</td>
<td>0.79</td>
<td>0.71</td>
</tr>
<tr>
<td>5</td>
<td>50</td>
<td>0.85</td>
<td>0.75</td>
</tr>
<tr>
<td>6</td>
<td>60</td>
<td>0.89</td>
<td>0.79</td>
</tr>
</tbody>
</table>
Fig 6.5.1: Number of Dataset vs. Precision

In this graph we have chosen two parameters called number of dataset and precision which is help to analyze the existing system and proposed systems. The precision parameter will be the Y axis and the number of dataset parameter will be the X axis. The blue line represents the existing system and the red line represents the proposed system. From this graph we see the precision of the proposed system is higher than the existing system. Through this we can conclude that the proposed system has the effective precision rate.

6.5.2. Recall Comparison

This graph shows the recall rate of existing and proposed system based on two parameters of recall and number of dataset. From the graph we can see that, when the number of number of dataset is improved the recall rate also improved in proposed system but when the number of number of dataset is improved the recall rate is reduced in existing system than the proposed system. From this graph we can say that
the recall rate of proposed system is increased which will be the best one. The values of this recall rate are given below:

**Table 6.5.2: Number of Dataset vs. Recall**

<table>
<thead>
<tr>
<th>SNO</th>
<th>Number of</th>
<th>AHMM</th>
<th>HMM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10</td>
<td>0.87</td>
<td>0.78</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>0.82</td>
<td>0.72</td>
</tr>
<tr>
<td>3</td>
<td>30</td>
<td>0.76</td>
<td>0.63</td>
</tr>
<tr>
<td>4</td>
<td>40</td>
<td>0.64</td>
<td>0.54</td>
</tr>
<tr>
<td>5</td>
<td>50</td>
<td>0.56</td>
<td>0.45</td>
</tr>
<tr>
<td>6</td>
<td>60</td>
<td>0.46</td>
<td>0.34</td>
</tr>
</tbody>
</table>

**Fig 6.5.2: Number of Dataset vs. Recall**

In this graph we have chosen two parameters called number of Dataset and recall which is help to analyze the existing system and proposed systems on the basis of recall. In X axis the Number of dataset parameter has been taken and in Y axis recall parameter has been taken. From this graph we see the recall rate of the proposed
system is in peak than the existing system. Through this we can conclude that the proposed system has the effective recall.

6.5.3. F-Measure Comparison

This graph shows the F measure rate of existing and proposed system based on two parameters of F measure and number of Dataset. From the graph we can see that, when the number of number of Dataset is improved the F measure rate also improved in proposed system but when the number of number of Dataset is improved the F measure rate is reduced in existing system than the proposed system. From this graph we can say that the F measure rate of proposed system is increased which will be the best one. The values of this F measure rate are given below:

<table>
<thead>
<tr>
<th>SNO</th>
<th>Number of AHMM</th>
<th>HMM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>10 0.87</td>
<td>0.78</td>
</tr>
<tr>
<td>2</td>
<td>20 0.82</td>
<td>0.72</td>
</tr>
<tr>
<td>3</td>
<td>30 0.76</td>
<td>0.63</td>
</tr>
<tr>
<td>4</td>
<td>40 0.64</td>
<td>0.54</td>
</tr>
<tr>
<td>5</td>
<td>50 0.56</td>
<td>0.45</td>
</tr>
<tr>
<td>6</td>
<td>60 0.46</td>
<td>0.34</td>
</tr>
</tbody>
</table>

In this graph we have chosen two parameters called number of Dataset and recall which is help to analyze the existing system and proposed systems on the basis of F measure. In X axis the Number of dataset parameter has been taken and in Y axis
F measure parameter has been taken. From this graph we see the F measure of the proposed system is in peak than the existing system. Through this we can conclude that the proposed system has the effective F measure.

### 6.6 SUMMARY

The credit card transaction method is examined as the basic stochastic process of an (Advanced Hidden Markov Model) AHMM. The variety of transaction quantity considered as the observation symbols, while the kinds of item have been deemed to be states of the AHMM. In addition to comprise recommended a technique for decision the spending profile of cardholders is authorized or not. As well as purpose of this knowledge in deciding the value of observation symbols and initial estimate of the model parameters with the best fit observation is that providing an effective credit card fraud detection system. It has also been enlightened how the HMM vary with the AHMM can detect whether an arriving transaction is fake or not. Experimental results show the performance and effectiveness of AHMM system and show the efficiency of knowledge the spending profile of the cardholder in AHMM system.