A Secure way of performing Credit Card Transaction using Hybrid Model

Ashish Gupta 1, Jagdish Raikwal 2
1 Institute of Engineering & Technology, M.E. in IT, DAVV, Indore, India
2 Institute of Engineering & Technology, Dept. of IT, DAVV, Indore, India

Abstract - In the current retail merchandise surroundings, electronic commerce has gained lots of attention and additionally provides spontaneous transactions. Using E-commerce individuals do their monetary dealing on-line like on-line shopping etc. In e-commerce, card payment mode has become the foremost vital suggests of payment because of quick development in information technology around the world. Because the usage of card payment mode has increased within the last decade, the magnitude of dishonest practices is additionally increasing each year. In spite of every tries to stop fraud, it carries on to be a serious threat to trade and government. Historically, organizations have been targeted fraud, interference instead of detection, to combat fraud. During this paper, we offer implementation and result analysis of a system that is employed to sight fraud in on-line dealings. The system relies on the Hybrid Markov Model that may be a combination of Hidden Markov Model, Bayesian Classifier and bio-metric method to sight fraud a lot of expeditiously than all previous planned system of fraud detection. The proposed system models the sequence of operations in Card transaction process employing a Hidden Markov Model (HMM) and social status of the cardboard holder using Bayesian classifier therefore showing how it is used for the detection of frauds. The performance of our proposed system is better than all the previously proposed system because in our proposed system we are using OTP and Biometric Authentication which reduces the possibility of the occurrence of the fraud in credit card transaction.

Keyword: OTP, Fraud, Biometric, Hidden Markov Model, Bayesian Classifier.

I. INTRODUCTION

Since past few years, there's a tremendous increase in popularity of the internet and as a result of the rise of the internet, e-commerce has become one of the leading sales channel for international trade. The preference of online shopping is growing day by day. On-line transactions are increased to buy services and product. 75 percent out of over 3,500 online shoppers surveyed in February 2013 say they like to buy their favourite merchandise on-line, like a review performed by United Parcel Service Inc [1]. Online shopping continues to exceed growth in traditional retail, the survey affirms. In a period of previous years, e-commerce grew about 15% to $280 billion, ten times the expansion rate of total U.S. retail outlay, accompany to the review. Within the developed countries similarly as within the developing countries up to some extent, credit card is the most competent payment mode for on-line shopping. As acceptance of credit card increases worldwide, attacker or hacker is attracted toward stealing of credit card details then, make fraud dealings consequential. Credit card fraud is being possible as a result of the issues in security of credit card processing technique. As a result of the credit card fraud, client are getting monetary lose and e-commerce distributor is leading toward client lose similarly as monasteries. ACT's [2] 2013 international Fraud Survey acknowledge that two in five respondents fell victim to credit, debit or prepaid card fraud within the period of past one year, with over 35 percent, illustrating that they conceive to halt using, or switch from, the card infected by fraudulent activity.

There are many ways using which fraudsters execute a credit card fraud. As technology changes, thus make the technology of fraudsters, and thus the approach during which they're going regarding completing fallacious activities. Frauds may be broadly speaking, classified into three classes, i.e., Traditional card related frauds, merchant related frauds and web frauds. The various kinds of methods for committing credit card frauds are explained below [3]:

A) Lost/ stolen Cards: In this kind of card fraud a physical stealing of the card either from the victim’s purse, pocket or alternative location occurred and it’s used for unauthorized transactions.

B) Account Takeover: This type of fraud happens once a fraudster lawlessly obtains a legitimate customers’ personal data. The fraudster takes control of (takeover) a legitimate account by either providing the customers’ account number or the card number. The fraudster then contacts the card issuing bank, masquerading to be the real cardholder, to raise that mail be redirected to a replacement address. The fraudster reports the card lost and asks for a replacement to be sent.

C) Fake and Counterfeit Cards: The creation of counterfeit cards, along with lost / stolen cards poses highest threat in credit card frauds. Fraudsters are consistently finding new and a lot of innovative ways in which to make counterfeit cards.

A number of the techniques used for making false and counterfeit cards are listed below:

1) Erasing the magnetic strip: A fraudster will tamper an existing card that has been acquired lawlessly by erasing the metal-like strip with a powerful electro-magnet. The fraudster then fabricates with the details on the card in order that they match the details of a legitimate card, that they'll have attained. Once the fraudster begins to use the card, the cashier swipe the card through the machine several times, before realizing that the metal like strip doesn't work. The cashier can then proceed to manually input the card details into the terminal. This kind of fraud has high risk as a result of the cashier is viewing the card closer to scan the numbers.
2) **Making a fake card**: A fraudster will produce a fake card from scratch using subtle machines. This is often the foremost common style of fraud, although fake cards need lots of effort and ability to produce. Modern cards have several security measures all designed to make it troublesome for fraudsters to produce sensible quality forgeries. Holograms are introduced in the major credit cards and are terribly troublesome to forage effectively. Embossing holograms onto the cardcase itself is another downside for card forgers.

3) **Altering card details**: A fraudster will alter cards by either re-embossing them by applying heat and pressure to the information originally decorated on the cardboard by a legitimate card manufacturer or by re-encoding them using a computer software system that encodes the magnetic tape data on the card.

4) **Skimming**: It may be a kind of card fraud which involves electronic repetition of the info from the card which are then used for making counterfeit cards. Victims of this type of card fraud are many times unaware of the matter till they receive their financial statement showing transactions they didn’t make.

5) **White plastic**: A white plastic may be a card-size piece of plastic of any colour that a fraudster creates and encodes with legitimate magnetic tape information for illegitimate transactions. This card seems like a hotel room key, however, contains legitimate magnetic tape information that fraudsters will use at POS terminals that don't need card validation or verification.

D) **Merchandise related Frauds**: Merchandise, related frauds are initiated either by the owners of the merchandise establishment or their workers. The categories of frauds initiated by Merchandisers are explained below:

1) **Merchandise Collusion**: This sort of fraud happens once merchant owners and/or their workers conspire to commit fraud using their customers’ (cardholder) accounts and/or personal data. Merchandise owners and/or their workers pass on the data concerning cardholders to fraudsters.

2) **Triangulation**: The fraudster during this sort of fraud operates from an internet website. Products are offered at heavily discounted rates and are shipped before payment. The fallacious site/web/website/web site seems to be a legitimate auction or a standard sales site. The client, whereas putting orders on-line provides data like name, address and valid credit card details for the location. Once fraudsters receive these details, they order a product from a legitimate website exploitation taken credit card details. The fraudster then goes on to buy different product using the credit card numbers of the client.

E) **Internet related Frauds**: The web has provided a perfect ground for fraudsters to commit credit card fraud in a simple manner. Fraudsters have recently begun to work on a really transnational Understanding credit card Frauds level. With the growth of trans-border or 'global' social, economic and political areas, the web has become a brand new World market, capturing shoppers from most countries around the world.

The foremost usually used techniques in web fraud are represented below:

1) **Website cloning**: Website cloning is where fraudsters clone a complete site or simply the pages from which you place your order. Customers don't have any reason to believe they are not addressing the company that they needed to buy the merchandise’s product or services from, as a result of the pages that they're viewing are just like those of the real web site. The cloned or spoofed web site can receive these details and send the client a receipt of the dealings via email even as the real company would. The buyer suspects nothing, while the fraudsters have all the main information that they wanted to commit credit card fraud.

2) **False merchant sites**: These sites typically provide the client a very low-cost service. The site requests a customer’s complete credit card details like name and address in return for access to the content of the site. Most of such sites claim to be free, however, need a legitimate credit card number to verify an individual’s age. These sites are created to accumulate as many credit card numbers as attainable. The sites themselves never charge people for the services they provide. The sites are sometimes a part of a large criminal network that either uses the main points it collects to boost revenues or sells valid credit card details to small fraudsters.

3) **Credit card generators**: Credit card number creators are computer programs that generate valid credit card numbers and ending dates. These generators work by generating lists of credit card account numbers from one account number. The software system works by using the mathematical Luhn algorithmic rule that card issuers use to get different valid card number combinations. The generators permit users to lawlessly generate as many numbers as the user wishes, in the way of any of the credit card formats, whether it is MasterCard, Visa or American express. Most of the time, the real cardholder isn't aware that someone else has seen or stolen his card information. Online transactions generally happen on the telephone or internet and to make this type of transaction, the user can want much important information about a credit card (such as card holder name, credit card variety, CVV number, validity). To make a fraud transaction to buy product and services, the fraudster would require understanding of these details of card only then he/she can make transactions. Uttermost of the time, the cardholder might or might not recognize that once or wherever any individual can have seen or stolen card information. To notice these sorts of fraud, there's just one means that is to review the defrayal patterns on each card and to work out any odd outlay pattern with relevance the usual outlay patterns. Fraud detection supported the investigation of existing purchase knowledge of the cardholder is one in every of the best means to decrease the speed of credit card frauds. Since we have a tendency to the humans tend to exhibit definite behaviouristic profiles, each cardholder may be drawn by a set of patterns containing information about the last purchase, the quantity of cash spent, the standard purchase category, the time since etc. Alteration of such patterns may be a potential threat to the system.
II. BACKGROUND STUDY

Hidden Markov Model: - Hidden Markov Model is perhaps the only and best models which might be used to model sequential data, i.e. Data samples that are dependent from one another. AN HMM is a double embedded random process with two completely distinct levels, one is hidden and the other is open to all[5]. The Hidden Markov Model is a defined set of states, each of that relates to a probability distribution. Transitions among the states are ruled by a collection of possibilities called transition probabilities. In a specific state an outcome or observation is often generated, according to the associated probability distribution. It's solely the end result, not the state visible to an external observer and thus states are “hidden” to the outside; therefore the name Hidden Markov Model [4]. Hidden Markov Model may be a good answer for addressing the detection of fraud transaction from credit card. An added necessary advantage of the HMM-based approach is an extreme decrease within the variety of False Positives transactions recognized as malicious by a fraud detection system despite the fact that they're extremely real.

In order to define an HMM completely, the following elements are needed[6].

The number of states of the model, N. We denote the set of states $S = \{S_1; S_2; S_3; ... S_n\}$, where $i = 1; 2; ...; N$, is a number of state and $S_i$, is an individual state. The state at time instant t is denoted by $q_t$.

The number of observation symbols in the alphabet, M. If the observations are continuous then $M$ is infinite. We denote the set of symbols $V = \{V_1; V_2; ... V_M\}$ where $V_i$, is an individual symbol for a finite value of $M$.

$\Lambda = \{a_{ij}\}$

A set of transition probabilities.

$a_{ij} = P\{q_{t+1} = S_j | q_t = S_i\}$, $1 \leq i, j \leq N$,

where $q_t$ denotes the current state. Transition probabilities should satisfy the normal stochastic constraints,

$$a_{ij} \geq 0, 1 \leq i, j \leq N$$

And

$$\sum_{i=1}^{N} a_{ij} = 1.1 \leq i \leq N.$$ 

The observation symbol probability matrix B,

$B = \{b_j(k)\}$

A probability distribution in each of the states,

$b_j(k) = P\{a_t = V_{k_j} | q_t = S_j\}$, $1 \leq j \leq N, 1 \leq k \leq M$

where, $V_{k_j}$ denotes the $k^{th}$ observation symbol in the alphabet, and the current parameter vector.

Following stochastic constraints must be satisfied.

$b_j(k) \geq 0, 1 \leq j \leq N, 1 \leq k \leq M$

And

$$\sum_{k=1}^{M} b_j(k) = 1, 1 \leq j \leq N$$

If the observations are continuous then we will have to use a continuous probability density function, instead of a set of discrete probabilities. In this case we specify the parameters of the probability density function. Usually the probability density is approximated by a weighted sum of $M$ Gaussian distributions $N$

$$b_j(a_t) = \sum_{m=1}^{M} c_{jm} \mathcal{N}(\mu_{jm}, \Sigma_{jm}, a_t)$$

where $c_{jm}$ weighting coefficients,

$\mu_{jm}$ mean vectors,

$\Sigma_{jm}$ Covariance matrices

$c_{jm}$ should satisfy the stochastic constrains,

$$c_{jm} \geq 0, 1 \leq j \leq N, 1 \leq m \leq M$$

and

$$\sum_{m=1}^{M} c_{jm} = 1, 1 \leq j \leq N$$

The initial state distribution, $\Pi = \{\Pi_i\}$, where,

$$\Pi_i = P\{q_1 = S_i\}, 1 \leq i \leq N$$

Therefore, we can use the compact notation $\lambda = (\Lambda, B, \Pi)$ to denote an HMM with discrete probability distributions, while $\lambda = (\Lambda, c_{jm}, \mu_{jm}, \Sigma_{jm}, \Pi)$ to denote one with continuous densities. Hidden Markov Model assumes that current output (observation) is statistically independent of the previous outputs (observations). We can formulate this assumption mathematically, by considering a sequence of observations,

$$O = O_1, O_2, O_3, ... O_R,$$

where $R$, is a number of observation in the sequence and $Q$, is a one particular sequence.

Then according to the assumption for an HMM, probability that $O$ is generated from this state sequence is given by

$$P\{O|q_1, q_2, q_3...q_R, \lambda\} = \Pi P\{O_1|q_1, \lambda\} P\{O|Q, \lambda\}$$

where $\lambda$, is a number of observation in the sequence and $Q$, is a one particular sequence.

Calculation of probability $P(O|\lambda)$ is an intensive computing process. Hence, a forward- backward algorithm [7] is used to calculate probability $P(O|\lambda).$

Bayesian Classifier:- The Naive Bayes classification algorithmic rule is a probabilistic classifier. It is based on
probability models that incorporate robust independence assumptions. The independence assumptions usually don't have an effect on reality. So they're thought of as naive. You can derive probability models by using Bayes' theorem (proposed by Thomas Bayes)[8]. Based on the nature of the probability model, you'll train the Naive Bayes algorithm program in a very supervised learning setting. In straightforward terms, a naive Bayes classifier assumes that the value of a specific feature is unrelated to the presence or absence of the other feature, given the category variable. There are two types of probability as follows:

- Posterior Probability \[ P(H/X) \]
- Prior Probability \[ P(H) \]

Where, \( X \) is data tuple and \( H \) is some hypothesis. According to Baye's Theorem

\[
P(H/X) = \frac{P(X/H)P(H)}{P(X)}
\]

**Why use Bayesian Classification:**

**Probabilistic learning:** Calculate specific possibilities for hypothesis, among the foremost practical approaches to certain sorts of learning issues.

**Incremental:** Every training example will incrementally increase/decrease the likelihood that a hypothesis is correct. The previous data will be combined with discovering knowledge.

**Probabilistic prediction:** Predict multiple hypotheses, weighted by their possibilities.

**Standard:** Even once theorem strategies are computationally unmanageable, they will give a standard of optimal higher cognitive process against that alternative ways will be measured.

**Biometric:** Biometrics refers to the quantitative knowledge (or metrics) associated with human characteristics and traits[9]. Life science identification (or biometric authentication) is employed in engineering science as a style of identity and access management. It's additionally accustomed determine people in teams that area unit beneath police work.

Biometric identifiers area unit the distinctive, measurable characteristics accustomed label and describe people. Biometric identifiers are a unit usually classified as physiological versus behavioural characteristics. Physiological characteristics are a unit associated with the form of the body. Examples embrace, however, don't seem to be restricted to DNA, fingerprint, face recognition, palm print, iris recognition, hand pure mathematics, membrane and odour/scent. Behavioural characteristics are a unit associated with the pattern of behaviour of an individual, as well as however not restricted to writing rhythm, gait, and voice. Some researchers have coined the term behaviometrics to explain the latter category of life science characteristics.

**Why Face Authentication???

- Techniques that rely on hands and fingers can become useless if the epidermal tissue is damaged in some way.
- Iris and retina identification require expensive equipment and are much too sensitive to any body motion.
- Voice recognition is susceptible to background noise in public places.
- Signatures can be modified or forged.
- Facial images can be easily obtained with a couple of inexpensive fixed cameras.

**III. PROBLEM STATEMENT**

The various challenges that are faced by most detection techniques embrace [10]:

- Skewed distribution of legitimate and dishonest information within the database that challenges the detection approaches. Real transactions are abundant higher as compared to dishonest.
- Count of transaction that is expanding rapidly. Mining of such vast quantity of information calls efficient techniques.
- Availability of labels information for the aim of training, as real or cheat is not promptly offered.
- Following user’s behaviour is hard because it changes very often for all sorts of users (good users, business and fraudsters). Handling recent as well as new intellectual may be a difficult task.

It is renowned that each cardholder includes a certain shopping behaviour, which establishes an activity profile for him. The majority of all present fraud detection techniques attempts to capture these behavioural patterns as rules and check for any violation in consequent transactions. However, these rules are mostly static in nature. As a result, they become ineffective once the cardholder develops new patterns of behaviour that aren't however acknowledged to the fraud detection system. The goal of a reliable detection system is to find out the behaviour of users dynamically therefore to minimize its own loss. Thus, systems that can't evolve or “learn”, could shortly become out-of-date leading to too many false alarms. A fraudster may try new sorts of attacks that ought to still get detected by the fraud detection system. For instance, a fraudster could aim at derivation most profit either by creating a couple of high worth purchases or too many numbers of low worth purchases so as to evade detection. Thus, there's a necessity for developing fraud detection systems which might integrate multiple evidences as well as patterns of real cardholders furthermore as that of fraudsters.

**IV. PROPOSED WORK**

The proposed work integrates the several security systems for authenticating the credit card fraud detection. During this context the less complicated and the advance security scheme is needed to evaluate for coming up with a new algorithmic rule. The proposed system includes the subsequent security design of the proposed study of work.
In the above given figure. 1 of our proposed work, we will ask our users to get registered in order to continue his transaction and minimize the possibility of occurrence of fraud. In this phase, we will request our user to fill a form with some field such as name, mobile no, email id, password, address, country. After all the information has been provided by the user we register the user based on his e-mail-id. To perform a transaction the user has to enter his/her log-in credentials i.e. email-id as user name and password that is provided by him during the registration phase. By getting the user registered we are minimizing the possibility of occurrence of fraud because we have a valid email id and phone number of the user (this phone number should be same as that is linked to his/her account number). In figure. 2 we'll request our user to enter the log-in detail so we'll verify the user, if the log-in is successful, then he would be asked to enter the credit card detail otherwise an error message generate and also the transaction halts. When the credit card details are entered, bank database is investigated to examine whether or not the details of credit card are correct or not. If the detail of credit card is found within the database and details are correct then our system uses the personal profile, bank transaction and also the credit card transaction for the detection of fraud, and if card detail isn't found, then our model generates an error message and also the transaction halts. In personal profile we use some parameter like number of automobile, number of motorbikes, number of houses, work type, number of family person employed. Then our model can classify these parameters of personal profile using naive Bayesian Classifier and assigns some weights to those parameters and generates social status of the credit card holder. Within the Bank dealings and credit card transaction our model uses all the previous transaction to come up with the probability of succeeding transaction using HMM (hidden Markov model) that generates the financial status of the credit card holder.

**Figure.1**: shows our proposed system user registration phase.

**Figure.2**: shows our proposed system fraud detection phase.
Bank transaction involves checking the validity of the card, card holder previous one year transaction, and balance within the account. Credit card transaction involves how the credit card bill is being paid, the pattern of his expenditure through the card. Then our model discovers fraud using these social statuses, financial status, if the requested transaction is found real then an OTP (One-time-password) is generated that is sent to the registered mobile number of the card holder. If the OTP is correct, then the transaction requested by the card holder is successful, otherwise an error message is generated and also the transaction halt. But if the social status and financial status discover that the incoming transaction is a fraud, then we've to seek out that this transaction is genuine i.e. The cardholder requested an odd transaction or it's really requested by any fraudster. To detect these dealings our model uses a Biometric Authorization method using real time face recognition. By this manner our proposed system investigates the present transaction and provide that the transaction is valid or invalid by jointly considering the odd transaction once the spending pattern of the cardholder changes.

V. IMPLEMENTATION

In the implementation phase we emphasis on the spending behaviour i.e. transaction pattern of the card holder. In figure.3 we have plotted a graph in which the x-axis represents the amount which the cardholders spend and y-axis represent the number of transaction. This spending behaviour is used by Hidden Markov Model to find out the transaction matrix and observation matrix. In the transaction matrix we will find out how much amount of transaction is done after a particular amount. For this we have created a transaction matrix in figure.4 where T1, T2...T6 are the transaction amount. In the same way we will create an observation matrix which will represent different amount of transaction in different month. Fig. 4 represent observation matrix where M1, M2...M6 refer to months and T1, T2...T6 refer to transactions.

![Figure 3: Graph representing Spending behaviour.](image)

```plaintext
T =
T1 2 0 4 2 0
T2 5 1 2 0 3 2
T3 2 0 1 2 1
T4 3 5 4 3 2 1
T5 2 0 0 1 2 1
T6 0 2 1 1 1 1
```

![Figure 4: Transaction Matrix](image)

```plaintext
O =
M1 1 2 0 4 2 0
M2 5 1 2 0 3 2
M3 2 0 0 1 2 1
M4 3 5 4 3 2 1
M5 2 0 0 1 2 1
M6 0 2 1 1 1 1
```

![Figure 5: Observation Matrix](image)

Now using the transaction matrix and observation matrix we will find out whether the incoming transaction is possible or not. If the incoming transaction is allowed by Hidden Markov Model then Bayesian Classifier also allows it without checking and proceeds the transaction to generate an OTP for secure transaction. And if the Hidden Markov Model found the incoming transaction as odd transaction then Bayesian Classifier check that the incoming transaction should be allowed or not. The Bayesian Classifier works as follow:

If the incoming current transaction ≠ any previous transaction then,
1. find cards in database which have successfully completed transaction similar to incoming current transaction and cards which have requested transaction similar as current incoming transaction but the transaction failed.
2. return Card Number.
3. Then sum up all the value in the personal profile of each whose card numbers are obtained in above step.
4. Then using the Bayesian Classifier formula we can calculate whether the current incoming transaction can be allowed or not i.e. the probability value obtained from
the formula is above a particular threshold value then the current incoming transaction is allowed otherwise not. If the current incoming transactions is allowed by either Hidden Markov Model or Bayesian Classifier than our model generates an OTP and send it to the cardholder’s mobile no. which is linked to his/her account. If the OTP entered by the cardholder than the transaction is successful otherwise the transaction fails.

If the current incoming transactions is not allowed by Hidden Markov Model and then by Bayesian Classifier than our model moves toward Biometric Authentication to check whether the transaction is requested by the genuine cardholder or by any fraudster. In Biometric Authentication we are using edge detection algorithm on face to validate the current incoming transaction. The edge detection algorithm aim at identifying points in a digital image at which the image brightness changes sharply or, more formally, has discontinuities. The points at which image brightness changes sharply are typically organized into a set of curved line segments termed edges. The digital image is captured either by computer webcam, laptop webcam or by Smartphone’s front camera. Since the image is captured by poor quality camera we cannot implement any heavy face recognition algorithm. If the cardholder validates himself by biometric authentication system then the incoming odd transaction is successful otherwise the transaction fails.

VI. RESULT & PERFORMANCE
In result, we have compared our proposed system and previously proposed HMM system and plotted graph for these both model based on various parameter such as accuracy, false detection rate and recognition rate.

In fig.5 the graph represents the accuracy between our proposed system and previously proposed HMM system. In the graph x-axis represent accuracy in % while y-axis represent number of instance. The accuracy is calculated using the formula:

\[
\text{Accuracy} = \frac{\text{total correctly identified samples}}{\text{total samples to be test}} \times 100
\]

In fig.6 the graph represents the false transaction detection rate between our proposed system and previously proposed HMM system. In the graph x-axis represent false detection in % while y-axis represent number of trail. The false transaction detection rate is calculated using the formula:

\[
\text{Detection rate} = \frac{\text{total false transaction detected}}{\text{total false} + \text{total true}}
\]

In fig.7 the graph represents the recognition rate of the biometric authentication phase of our proposed system. In this we have plotted a graph between the accuracy in % and month i.e. representing how the recognition rate change with the time period of the image updated in the database the more frequent one update the image in the database more is the recognition.
REFERENCES


