Credit Card Fraud Detection Based on User Profile and Previous Transaction

* Nitika Kadam, **Suryakant Soni, ***Devendra Puntambekar, ****Rahul Kaul

* Computer Science Engineering, Mandsaur Institute of Technology, Indore
** Computer Science Engineering, Mandsaur Institute of Technology, Indore
*** Information Technology, Mandsaur Institute of Technology, Indore
**** Computer Science Engineering, BM College, Indore

INTRODUCTION

In day to day life credit cards are used for purchasing goods and services. The most accepted payment mode is credit card for both online and offline. For online transaction it uses virtual card and for offline transaction it uses physical card. In today’s world, credit card provides cashless shopping at every shop. It will be the most convenient way to do online shopping, paying bills etc. As credit card becomes the most popular mode of payment for both online as well as regular purchase, cases of fraud associated with it are also rising. In this paper, the concept of data mining and Hidden Markov Model is used to detect the credit card fraud during transactions. Hidden Markov Model is the statistical tools for engineer and scientists to solve various problems. An Hidden Markov Model is initially trained with the normal behavior of a customer who hold the credit card. With the concept of data mining, previous information about the transaction is extracted then on the basis of extracted information, Hidden Markov Model consider whether the current transaction is either fraudulent or genuine. This paper shows how advanced data mining techniques and HMM can be combined successfully to detect the credit card fraud.

KEYWORDS: Credit Card, Internet, Data Mining, HMM, Online Shopping, Offline Shopping

DATA MINING

Data Mining is used to detect fraud for its effectiveness. Data mining is a process that uses a variety of data analysis tools to discover patterns and relationships in data that may be used to make a valid prediction. The six basic steps of data mining process are defining the problem, preparing data, exploring data, building models, exploring and validating models, and deploying and updating models. Data mining is a field at the intersection of computer science and statistics, is the process that attempts to discover patterns in large sets. The overall goal of the data mining process is to extract information from a data set and transform it into an understandable structure for further use. Aside from the raw analysis step, it involves database and data management aspects, data preprocessing, model and inference considerations, interestingness metrics, complexity considerations, post-processing of discovered structures, visualization, and online updating.

TECHNIQUE AND ALGORITHM USED

Technique that is used for the detection of credit card fraud based on the concept of data mining is Knowledge Data Discovery (KDD).
**KDD:**

Many people treat data mining as a synonym for popularly used term KDD. The process involves in KDD are:

- Data Cleaning
- Data Integration
- Data Selection
- Data Transformation
- Data Mining
- Pattern Evaluation
- Knowledge Presentation

An algorithm called as HMM is used for the detection of credit card fraud. This algorithm primarily focus on three categories:

- Lower Profile
- Middle Profile
- High Profile

This algorithm performs prediction analysis for the fraud detection. Prediction symbol for this model is \( V \{ l, m, h \} \), where \( l \) stands for lower profile, \( m \) stands for middle profile and \( h \) stands for higher profile. E.g., if cardholder perform a transaction as Rs 2000 and cardholders profile groups as \( l \) (low) = (Rs 0, Rs 1500), \( m \) (middle) = (Rs 1500, Rs 5000), and \( h \) (high) = (Rs 5000, up to credit card limit), then transaction which cardholder want to do will come in middle profile group. So the corresponding profile group or symbol is \( M \) and \( V \{ 2 \} \) will be used. Credit card holder purchase various product with different amount in various period of time. It uses the deviation in a purchasing amount of latest 5 transaction sequence (and adding one new transaction in that sequence) which is one of the possibilities related to the probability calculation. If model does not have data of last 5 transactions in the initial stage, in that case, model will ask to the cardholder to feed basic information during transaction about the cardholder such as address, annual income, date of birth, email id etc. On the basis of this information HMM model obtained relative data of transaction for further verification on spending profile of cardholder.

**VERIFICATION OF FRAUD TRANSACTION:**

All the information about credit card such as Credit card number, CVV number, credit card Expiry month and year, name of credit card holder etc will be checked with credit card database. If cardholder entered database is correct then it will ask Personal Identity number (PIN). After matching of Personal Identity number (PIN) with database and account balance of user’s credit card is more than the purchase amount, the fraud checking module will be activated. If user credit card has less than 5 transactions then it will directly ask to provide personal information to do the transaction. Once database of 5 transactions will be developed, then fraud detection system will start to work. By using this observation, determine users spending profile. The purchase amount will be checked with spending profile of user. By transition probabilistic calculation based on HMM, it concludes whether the transaction is real or fraud. If transaction may be concluded as fraudulent transaction then user must enter security information. This information is related with credit card (like account number, security question and answer which are provided at the time of registration). If transaction will not be fraudulent then it will direct to give permission for transaction. If the detected transaction is fraudulent then the Security information form will arise. It has a set of question where the user has to answer them correctly to do the transaction. These forms have information such as personal, professional, address; dates of birth, etc are available in the database. If user entered information will be matched with database information, then transaction will be done securely, else user transaction will be terminated.

**RESULT**

In this section, it is shown that fraud detection will be checked on previous transactions and also calculate percentage of each spending profile (low, middle and high) based on total number of transactions. In Table, list of all transactions are shown:

<table>
<thead>
<tr>
<th>No. of Transaction</th>
<th>Amount</th>
<th>No. of transaction</th>
<th>Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st</td>
<td>140</td>
<td>111h</td>
<td>210</td>
</tr>
<tr>
<td>2nd</td>
<td>125</td>
<td>12th</td>
<td>550</td>
</tr>
<tr>
<td>3rd</td>
<td>15</td>
<td>13th</td>
<td>800</td>
</tr>
<tr>
<td>4th</td>
<td>5</td>
<td>14th</td>
<td>110</td>
</tr>
<tr>
<td>5th</td>
<td>10</td>
<td>15th</td>
<td>135</td>
</tr>
<tr>
<td>6th</td>
<td>125</td>
<td>16th</td>
<td>118</td>
</tr>
<tr>
<td>7th</td>
<td>15</td>
<td>17th</td>
<td>20</td>
</tr>
<tr>
<td>8th</td>
<td>120</td>
<td>18th</td>
<td>148</td>
</tr>
<tr>
<td>9th</td>
<td>10</td>
<td>19th</td>
<td>141</td>
</tr>
<tr>
<td>10th</td>
<td>280</td>
<td>20th</td>
<td>6</td>
</tr>
</tbody>
</table>

The most recent transaction is placed at the first position and correspondingly first transaction is placed at the last position in the table. The pattern of spending profile of the cardholder is shown in below Figure based on all transactions done.

The percentage calculation of each spending profile (low, medium and high) of the cardholder based on price distribution range as mentioned earlier is shown in below Figure

**CONCLUSION**

In this paper it has been discussed that how Hidden Markov Model will facilitate to stop fraudulent transaction through credit card. With the help of HMM, based on three categories such as low profile, middle profile, high profile and on the previous transactions of credit card holder our proposed model determine whether the current transaction is genuine or fraud. If fraud transaction occurs then it terminate the current transaction else continue with the transaction. Our proposed model suggest that if the previous transaction of a card holder is not present then information related with credit card like account number, security question and answer must provided by the cardholder. The Hidden Markov Model makes the processing of detection very easy and tries to remove the complexity.
REFERENCES

4. Data Mining- Concepts And Techniques by Han and Kamber. 
5. Hidden Markov Models, Theory and Applications by Przemyslaw Dymarski