Security Mechanism to Detect Fraud based on Customer Behavior

Mandeep Singh¹ Parvinder Singh²

¹Associate Professor ²Post Graduate Student
¹,²Department of Computer Science And Engineering
Chandigarh University, Mohali, Punjab, India Pin 140413

Abstract— With the enhancement in technology e banking like credit Card, Debit Card, Mobile Banking and Internet Banking is the popular medium to transfer the money from one account to another. E Banking is gaining popularity day by day, which increases the online transaction with the increase in online shopping, online bill payment like electricity, Insurance Premium and other charges, online recharges and online reservation of railways, bus etc., cases of fraud associated with it are also increasing and it puts a great burden on the economy, affecting both customers and financial bodies. It not only costs money, but also a great amount of time to restore the harm done. The purpose is to prevent the customer from online transaction by proposing unique and hybrid approach containing Data Mining and Artificial Intelligence based techniques and in order to prevent from fraudulent transactions, PIN, One Time Password mechanisms are used to prevent from online transactions.

Keywords: Unique, Hybrid, Logical, , Data Mining, Artificial Intelligence.

I. INTRODUCTION

Due to the growth of modern technology, the mode of payment of individual has changed significantly. The use of Online Payment mode such as Online Banking, Debit Card, Credit Card etc. has become popular and is becoming important in day to day activities because it allows bank customers to purchase goods and services from the shopping websites or from the market. At the same time, fraud becomes a serious threat for bank customer by various methods such as Phishing websites, counterfeit fraud, steal/theft card, database stealing etc. Fraud refers to obtain goods/services and money by illegal means. Fraud deals with events which involve criminal motives that are difficult to identify. Credit card is also one of the most illegal types of fraud. Credit card is a plastic card i.e. issued to customers of a bank as one of the mode of payment. It allows cardholders to purchase goods and services from the shopping websites or from the market. Credit Card Fraud is defined as, when an individual uses another individual credit card for personal use while the owner of the card as well as the card issuer are not aware of the thing that the card is being used.

A. Fraud can be divided into two types

1) Online Fraud: Online Fraud is committed by stealing personal identity of customer on the network by methods such as phishing, botnet etc.

2) Offline Fraud: Offline Fraud is committed by steal/theft physical card.

B. Security Mechanism

1) One Time Password: OTP is a unique number that generates every time for every online transaction that comes on the customer mobile.

2) PIN: A PIN is a 4 digit unique and secret number that customer has to enter before doing transaction by ATM/Debit Card.

C. Example of Credit Card Fraud

If credential information of a customer has stolen and uses for online shopping, the card holder acknowledges and then customer inquire the bank for transaction. There is no such process that can prevent fraudulent transaction at the time of happening. So there is a need of such interface that prevents from online transactions.

D. Methods to Steal Personal Information

1) Hacking: Hacker is a person who seeks and exploits weakness in computer system. Attacker breaks into industry or personal databases.

2) Phishing: Phishing is a fraudulent attempt, usually made through email, to steal your personal information.

3) Spoofing: The word "spoof" means to hoax, trick, or deceive. Spoofing refers to tricking or deceiving the computer users. This is typically done by hiding one's identity or faking the identity of another user on the Internet.

4) Spyware: The computer user unknowingly downloads software from the Internet that contains spyware. Spyware collects personal information from your computer and transmits it to fraudster or attacker.

5) Shoulder Surfing: An attacker watches a bank customer from a nearby location as the customer punches in his personal information. If the customer is giving his personal information over the phone (e.g., to a hotel or car rental company), the attacker may listen to the conversation so as to obtain personal information of bank customer.

6) Dumpster Diving: An attacker goes through a customer’s garbage cans or trash bins to obtain personal information of bank customer such as bank statement, payment receipt etc.

7) Identity Theft: Fraud is a type of Identity Theft. Identity theft is a form of stealing someone's identity in which someone pretends to be someone else by assuming that person's identity, typically in order to access resources or obtain credit and other benefits in that person's name. Identity theft occurs when someone uses another's personally identifying information, like their name,
identifying number, or credit card number, without their permission, to commit fraud or other crimes. Some basic tips to prevent from Identity theft:

- Don't carry your Social Security card such as Debit Card, Credit Card etc.
- Don’t disclose your pin number & password to anybody.
- Watch out for “shoulder surfers”.
- Collect mail promptly.
- Pay attention to your billing cycles.
- Keep your receipts.
- Store Personal Information to a safe place.

8) Suspicious Behavior:
- If the user is taking long time to perform the transaction than the routine time (number of times that the user can take time) then there may be a chance that user is suspicious means user is fraudster. Example: Suppose the bank customer normally takes 60 seconds to perform the transaction and suddenly the user is taking 10 minutes to perform the transaction then the user is found to be suspicious.
- Suppose the user performs the transaction in India at 11:00 AM by providing his details and same user details are used in England at 12:00 AM then there may be possibility that transaction is fraudulent.
- A customer usually carries out similar type of transaction in terms of amount and suddenly the customer is purchases an item from shopping website which is costly then there may be possibility that transaction is fraudulent.

9) Data Mining:
Data Mining is a technique that uses statistical, mathematical, artificial intelligence, and neural network techniques to extract and identify important information from a large database. Data Mining is a technique to analyze data from different views and summarizes it into useful information. Data Mining is a technique that is used to detect financial fraud detection because it can identify new attacks before financial fraud can be detected by human experts.

10) Challenges in Data Mining to detect Fraud:
- There are millions of transactions each day. To extract large amount of data from a database requires highly efficient techniques.
- The data or information is noisy or vague.
- Data labels are not immediately available. Frauds or intrusions usually aware after they have already happened.
- It is hard to track user’s behaviours. All types of users (good users, business, and fraudsters) change their behaviours frequently.

11) Related Work On Fraud Detection:
Quah.John.TS, proposed a paper (2007) under the title “Real-time credit card fraud detection using computational intelligence”. In this paper, author proposed a neural network based approach i.e. called SOM (Self Organizing Map) to detect spending pattern of the customer in credit card database and SOM is a multilayer approach that consists of: The initial Authentication, Screening layer, Risk Scoring layer, Behavior Analysis layer (Core Layer) and Decision making layer. The main purpose of SOM approach is to classify and cluster input data, to detect and derive hidden patterns in input data [17].

Srivastva.Abhinav, proposed a paper (2008) “Credit Card Fraud Detection Using Hidden Markov Model”. In this paper, HMM (Hidden Markov Model) is proposed to detect fraudulent transactions and K Mean Clustering algorithm is used to identify spending behavior of a customer. HMM-based applications are used in various areas such as speech recognition, bioinformatics, and genomics [18].

Panigrahi.Suvasini, proposed a paper (2009), “Credit card fraud detection: A fusion approach using Dempster–Shafer theory and Bayesian learning”. In this paper, author proposed Fraud Detection System i.e. based on the combination of three approaches: Rule-based filtering, Dempster–Shafer theory and Bayesian learning in which Dempster rule is used to match customer current behavior compared with the previous behavior, rule based filtering approach is used to determine the suspicious level of each incoming transaction and Bayesian learning approach is used to update the suspicious score of transaction using history database of both genuine cardholder as well as fraudster [16].

Sanchez’., proposed a paper (2009)” Associations Rules applied to credit card fraud detection” In this paper, Association rules (Fuzzy Rules) are used to detect new, undesired behavior of bank customer in the online verification process and Association Rules (Fuzzy Rules) are applied in the area of Business Management and planning to extract data of fraudulent transaction from a large database [20].

Farvaresh.Hamid, proposed a paper (2010),” A Data Mining Framework for detecting subscription fraud” In this paper, author proposed a framework to detect fraud telecommunication subscribers by using various techniques such as data cleaning, dimension reduction, clustering and classification and the main problem in this framework is that it requires the historic data to identify whether the customer is fraudster or genuine [6].

Bhattacharyya.Siddhartha, proposed a paper (2010),”Data mining for credit card fraud” In this paper, author evaluate two advanced data mining approaches, decision tree approach, support vector machines and random forests together with logistic regression to detect credit card fraud and examines the performance of these techniques with the varying level of data undersampling and these techniques can detect only few fraudulent transaction when it is applied to a real world data set [1].

Duman.Ekrem, proposed a paper (2011) “Detecting Credit Card Fraud by genetic algorithm and scatter search” In this paper, author proposed a technique such as genetic algorithm and scatter search to score each transaction and based on these scores the transaction can be classified as fraudulent or genuine transaction and these approaches are based on the classification problem [3].

Dharwa.N.Jyotindra, proposed a paper (2011), “A Data Mining with Hybrid Approach Based Transaction Risk Score Generation Model (TRSGM) for Fraud Detection of Online Financial Transaction” In this paper, author proposed
Transaction Risk Score Generation Method to calculate certainty factor to identify whether the transaction is fraudulent or genuine and Risk score is analyzed by identify spending profile of customer of a bank by implementing DBSCAN algorithm and address mismatch in which it will identify whether the customer billing and shipping is same [4].

Cao.Longbing.Zhang., proposed a paper (2011), under the title “Combined Mining: Discovering Informative Knowledge in Complex Data” In this paper, author proposes combined mining as a general approach to extract informative knowledge in complex data to solve problems the problems of enterprise applications, such as telecom fraud detection and cross-market surveillance in stock markets and a framework that is flexible and customizable for handling a large amount of complex data. [2]

Li.Jinjiu.Wei., proposed a paper (2012), “Effective detection of sophisticated online banking fraud on extremely imbalanced data”. In this paper, author proposed online banking fraud detection framework i.e. based on utilizing resources and advanced data mining techniques and algorithms such as contrast pattern mining, neural network and decision forest are implemented and their outcomes are integrated with an overall score measuring the risk of an online transaction being fraudulent or genuine [25].

Phua.Clifton, proposed a paper (2012), “Resilient Identity Crime Detection”. In this paper, author proposed two new layers such as Communal Detection (CD) and Spike Detection where CD is a white list oriented approach on a fixed set of attribute and finds real social relationships to reduce the suspicion score and SD is attribute oriented approach on a variable size set of attribute and find spikes in duplicates to increase the suspicion score to detect more types of attack on and remove the redundant attribute [14].

Wu.Shu, proposed a paper (2013) proposed a paper “Information-Theoretic Outlier Detection for Large-Scale Categorical Data” In this paper, author propose two practical outlier detection methods named ITB-SS and ITB-SP to solve real world problem such as Intrusion Detection, Criminal activity detection in E-Commerce etc. and these methods does not require user defined parameter to decide whether an object is outlier and author also proposed a new concept called ‘Holoentropy’ that takes both data and total correlation into consideration [23].

Sahin.Yusuf , proposed a paper (2013) ”A cost-sensitive decision tree approach for fraud detection” In this paper, the security mechanism such as CHIP and PIN that are developed for credit card system does not prevent from fraudulent credit card usages over online fraud and the author have developed and implemented a cost sensitive decision tree approach to detect fraudulent transactions and this approach is compared with the traditional classification models on a real world credit card data set [19].

Kim.Chan.Ae, proposed a paper (2013), “Fraud and financial crime detection model using malware forensics” In this paper, digital forensics techniques has been used to analyze system intrusin incidents traditionally is used to detect anomaly transactions that may occur in the user environment during electronic financial transactions and the risk point calculation model is proposed by scoring anomaly transaction cases in the detection step by items [26].

J2) Data Mining Techniques

Fraud Detection has been usually seen as a data mining problem where the objective is to correctly classify the transactions as legitimate or fraudulent.

![Data Mining Techniques](image)

**Fig. 2:** Data Mining Techniques

Association: In association, patterns are discovered based on a relationship of a particular item on other items in the same transaction [21]. For example, the association technique is used in market basket analysis to identify what products that customers frequently purchase together [10].

Classification: Classification is a data mining technique i.e. based on machine learning. Classification is used to classify each item in a set of data into one of predefined set of classes or groups [23]. Classification method makes use of mathematical techniques such as decision trees, linear programming, neural network and statistics [21]. For example, Classification in application that “given all past records of employees who left the company, predict which current employees are probably to leave in the future.” In this case, classification divides the employee’s records into two groups that are “leave” and “stay” [10].

Clustering: A cluster is a collection of data objects that are similar to one another within the same cluster and are dissimilar to the objects in other clusters. For example, in a library, books have a wide range of topics available. The challenge is how to keep those books in a way that readers can take several books in a specific topic without hassle. In clustering technique, the books that have some kind of similarities in one cluster or one shelf and label it with a meaningful name.

Prediction: The prediction is one of a data mining techniques that discover relationship between independent variables and relationship between dependent and independent variables. In this technique, prediction analysis technique can be used in sale to predict profit for the future if we consider sale is an independent variable, profit could be a dependent variable [5].

Detection of Fraud is based on three approaches:
Supervised Approach: In supervised approach, the fraud can be detected by analyzing the previous records of customer and compare these records with the current transaction. It uses Classification technique to detect fraud.

Semi Supervised Approach: Semi supervised con be used in interactive learning system. It uses Prediction technique to predict all the outliers (unusual behavior from a normal data set) from the input data set based on an outlier sample indicated by the end user.

Unsupervised Approach: This Approach is used to detect unusual observation such as transactions, customers or accounts whose behavior may be different from the normal.

The fraud detection techniques are based on Supervised, Semi Supervised and Unsupervised approach and these techniques are:

- Genetic Algorithm: Genetic Algorithm is based on the sexual reproduction in which the genes of two parents combined to form to those of their children. When this algorithm is applied to any problem the basic premise is that we can create an initial population of individual representing possible solutions to a problem. Each of those individual has certain characteristics that make them more or less fit as members of population. The fit member will have a higher probability of producing efficient solution. This method is very effective at finding optimal or near optimal solution.

- Neural Network: Neural Network is an Artificial Intelligence technique that is used in biological learning system. Artificial Neural Network is composed of large number of highly interconnected nodes to solve the specific problem. Example: Neural Network can be used to distinguish legal from fraudulent transaction, detect fraudulent behavior on an e-commerce site.

- Self Organising Map (SOM): SOM is a neural network based on the unsupervised learning approach that consists of neurons. In SOM historical data present in the database gets classified into genuine and fraudulent sets through the process of self-organization. After that any new incoming transaction is pre processed and fed to the SOM. Based on the threshold value, fraudulent or genuine transaction is identified. SOM is a multilayered approach consisting of:
  - The initial authentication and screening layer.
  - The risk scoring and behavior analysis layer.
  - Decision-making layer.

- Bayesian belief network (BBN): BBN are graphical representations of probability distributions, derived from co-occurrence counts in the set of data items. BBN is a directed, acyclic graph, where the nodes represent attribute variables and the edges represent probabilistic dependencies between the attribute variables. Bayes theorem gives the mathematical formula to detect fraud in the form of posterior probability. [16]

  \[
  P(\text{E|fraud}) = P(\text{E|fraud}) \times P(\text{fraud}) / P(\text{E})
  \]

- Support Vector Machines: Support vector machines (SVMs) are statistical learning techniques that can be used in a classification tasks. This technique is based on the supervised learning algorithm. An SVM model is a representation as points in space, and different points are mapped so that the separate categories are divided by a clear gap that is as wide as possible.

- Logistic Regression: Logistic regression is a probabilistic statistical classification model. Logistic Regression is used to measure the relationship between the dependent & independent variable.

- Card Based Reasoning: Case-based reasoning (CBR) is a technique that solves a given problem by using past experience and solutions of particular problem. A case is usually a specific problem that has been previously solved. Suppose a new problem is given, case-based reasoning identified the set of stored cases and finds similar ones. If similar cases exist, their solution is applied to the new problem, and the problem is added to the case for future reference.

The following is the table of the fraud detection techniques:

<table>
<thead>
<tr>
<th>Technique</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genetic Algorithm</td>
<td>Flexible to solve complex problem, Ability to handle noisy data and efficient to detect fraudulent transactions.</td>
<td>Does not assure that solution is an optimal solution that solves the particular problem &amp; to solve the problem by this technique is difficult because it requires full knowledge of tool in real time system.</td>
</tr>
<tr>
<td>Neural Network</td>
<td>Processing speed is fast, portable and have the ability to solve complex problem</td>
<td>It can only work with numeric data and if the data is non-numeric the data need to be converted in numeric form, less efficient if the data set is large, online fraudulent transactions is difficult to detect by this technique.</td>
</tr>
<tr>
<td>Logistic Regression</td>
<td>Easy to understand, easy to implement and used with class classification. Logistic Regression provides better accuracy if dataset is small.</td>
<td>Not accurate if data dataset is large, noisy or outliers are not able detect by this technique.</td>
</tr>
<tr>
<td>Self Organising Map</td>
<td>SOM can be used to classify transactions based on predetermined sets like amount transfer, interbank transactions etc., efficient &amp; cost-effective technique of real-time fraud</td>
<td>Not able to identify customer behavior in fraudulent transactions.</td>
</tr>
</tbody>
</table>
detecction.

| Support Vector Machine | SVM technique provides a robust solution and unique solution as compared to neural network approach. | Lack of transparency of results & not efficient to detect fraudulent transactions. |

Table 1: data mining techniques

II. CONCLUSION & FUTURE SCOPE

This paper describes financial cybercrime detection and various techniques to detect Financial Fraud Detection based on data mining such as Neural Network, Regression, Genetic Algorithm, Bayesian Learning Approach etc. and these techniques have its own advantages and disadvantages. This paper describes the need of security mechanism if transaction is fraudulent and also describes the behavior of customer in online transaction. The customer behavior is difficult to identify so there is a need of technique that can identify customer behavior. In order to prevent from fraudulent transaction, PIN & One Time Password is used to prevent from online transactions. Now there is a need of technique that identifies the customer behavior. In order to solve this problem we will propose a method identifies the user current geographical location and compares the current geographical location with the previous location and a security mechanism in which the user has to enter 4 digit number that appears on the screen in a logical manner.

REFERENCES


