Credit Card Transaction Fraud Detection System Using Fuzzy Logic and K-Means Algorithm

Dr. M. Balamurugan[1], P. Mathiazhagan[2]
[1]*Associate Professor, School of Computer Science, Engineering and Applications, Bharathidasan University*
[2]*Research Scholar, School of Computer Science, Engineering and Applications, Bharathidasan University*

Abstract- **The usage of credit card has dramatically increased, credit card fraud has become increasing rampant in recent years. Nowadays credit card fraud is one of the major issues of great financial losses, for the merchants and individual clients are also affected. This fraud is difficult to find out fraudulent and concerning losses will be barred by issuing authorities. As a result, fraud detection is the important solution and almost certainly the best way to stop credit card fraud types. Fuzzy logic is to analyze the spending profile of each card holder Credit card fraud can be detected on analyzing of previous transactions data. In this study Fuzzy logic and k-means are developed and applied to credit card fraud detection problem. It will be the most effective method to counter fraud transaction through internet. Fuzzy logic and k-means produce a better result comparing to the other data mining techniques.**

Index Terms- **Online, Credit card fraud detection, Fuzzy logic, k-means clustering.**

I. INTRODUCTION

A credit card is a convenient method of payment providing many benefits to all consumers and merchants. Merchants pay a fee to access this system and to propose their clients the convenience of employing their credit card. The Credit cards are commonly accepted across the country and around the world because they are a valuable payment tool that makes doing business easier for merchants and their clients.

The credit card is a small size printed plastic card issued to users as a system of payment. This card allows its cardholder to buy goods and services based on the cardholder's promise to compensate for these commodities and service. Credit card security relies on the physical protection of the plastic card as well as the secrecy of the credit card number. Globalization and increased usage of the internet for online purchase has resulted in a significant proliferation of credit card transactions throughout the Earth. Therefore a rapid increase in the number of credit card transactions has led to a significant rise in fraudulent activities. Credit card fraud is a broad-ranging term for larceny and fraud committed using a credit card as a fraudulent source of funds in a given transaction. Credit card fraudsters use a many number of techniques to assign fraud. To contest the credit card fraud efficiently, it is important to first infer the mechanisms for identifying a credit card fraud. In recent the years credit card fraud has stabilized much due to various credit card fraud detection and avoidance mechanisms. The Fraud can be limited as the undesired activities taking place in an operational arrangement. This Fraud can appear in a variety of different domains, including medical, finance, health care, telecommunications and public services [1].

The Credit card frauds can be performed in different ways, such as theft, software application fraud, counterfeit cards, never received an issue (NRI) and online fraud (where the card holder is not present). In online fraud, the purchase is done remotely and just the credit card's details are required. A manual signature, a PIN or a card imprint was not taken at the time of purchase. Though avoidance mechanisms like card chip and card pin number decrease the fraudulent activities through simple theft, counterfeit cards and NRI; online frauds (Internet, translation and mail order frauds) nowadays are even increasing in both total and number of transactions. There has been an increasing quantity of economic losses due to credit card frauds as the use of the credit cards become more and more coarse. Many papers reported huge amounts of losses in different countries [2, 3]. According to Visa reports about European countries, roughly 50% of the whole credit card fraud losses in 2008 are due to online hoaxes.

A.Traditional Card Related Frauds

Credit card fraud has been divided into three types: Card Related Frauds, Merchant Related Frauds and Internet Related Frauds.

a)Card Related Frauds: Card related Frauds into four categories. Application Frauds, Lost / Stolen Cards, Account Takeover and Fake and Counterfeit Cards.

b)Application Frauds: Assumed individuality, where a person illegally obtains personal information of another person and opens accounts in his or her name, using partial legitimate information. Where a person's individual provides fake information about the financial status to obtain credit. When a credit card is stolen from the postal service before this card is reaching its owner's destination is called not received items (NRIs).

c)Lost / Stolen Cards: An account holder receives a credit card and lose it or someone steals the card for criminal purposes. This type of fraud is on the easiest way for a fraudster to get hold of the other person's credit cards without investment in the latest technology.

d)Account Takeover: The account takeover type of frauds illegally obtains a valid customer's full personal details. The card fraudster takes full control of (takeover) a legal account by either providing the customers account number or the credit card number.

Fake and Counterfeit Cards: Fraudsters are finding new and more technical ways create a new counterfeit card. First one a fraudster can tamper an existing credit card that has been acquired criminally by erasing the credit card information use metallic strip with a powerful electro-magnet. Second one is a fraudster can create a fake credit card from scratch using technical (sophisticated) machine. The last one is create a white plastic card. This card-size piece of plastic of any type of color that a fraudster creates new credit card and is encoded with legal magnetic stripe original data for illegal transactions. Merchant Related Frauds: Merchant related Frauds into two categories. Merchant Collusion and Triangulation.

e) Merchant Collusion: The merchant owners and/or their employees combine to commit fraud using their cardholder accounts and/or personal details. The Merchant owners and/or their employees pass on the cardholder‘s information to fraudsters.

f)Triangulation: This type of fraud is created for the web sites. In online purchase the goods are heavily offered and also shipped from before payments. The customer provides this offer and gives information such as the customer name, address. Card number and valid credit card details to the websites. Once get the above details the fraudsters order goods from a legal site using stolen credit card details.

h) Internet Related Frauds: Internet related Frauds into three categories. Site cloning, false merchant sites and credit card generators.

i) Site Cloning: This type of fraud is the fraudsters clone to a full site or just the pages from which you place your order. The Customer suspects nothing, whilst the fraudsters have all the customer details they need to commit credit card fraud.

j) False merchant sites: These web sites often offer the customer an extremely heavy offer good service. The purchase site requests a customer's complete credit card details such as name, address and card details in return for access to the content of the web site. This type of web sites are usually part of a larger criminal network that either uses the card details it collects to raise revenues or sells valid credit card details to small fraudsters.

k)Credit card generators: Credit card number generators are computer software programs that generate the valid credit card numbers and expiry dates. These generators programs work by generating lists of credit card account numbers from a single account number. The generators allow users to illegally generate as many card numbers as the user desires, in the form of any of the credit card formats, whether it be American Express, Visa or MasterCard [3].

The credit card fraud detection system based on a performance of Bayesian and Logistic Regression methods in credit card fraud detection with a real data set. In this system, each account detail is monitored separately using suitable algorithms, and the transactions are attempted to be identified and flagged. The legitimate or normal identification will be based on the suspicion score formed by the classifier models developed. When a new transaction the classifier can predict whether the transaction is normal or fraud.

There are many ways in which fraudsters execute a credit card fraud. As technology increased, there are also possibilities of capturing fraudster's moves and their fraudulent activities. The primary goal of this paper is to catch hold of the credit card frauds that are attempted over internet. Credit card fraud detection is a traditional data mining problem and was focused by many researchers. Though, it seems to be traditional it still gains the attraction of new researchers for something that deals with the improvisation in terms of accuracy.

In this paper, a novel fuzzy logic and K-means algorithm is chosen to be experimented over the credit card dataset and also the results are to be discussed in the next coming sections.

II.LITERATURE REVIEW

Pan Su, et al. proposed methods, therefore, employ a fuzzy graph to represent the relationships between Component clusters upon which to derive the final ensemble clustering results. In this work using various benchmark data sets, the proposed methods are tested, aligned with typical traditional methods. The new results show that the proposed fuzzy-link-based clustering ensemble approach generally outperforms the others in terms of accuracy. The present work opens up an avenue for further investigation. For instance, many other bases-clustering member generating methods such as re-sampling may also be applied. It would be useful to investigate the performance of the proposed fuzzy graph using different consensus functions. It is also interesting to examine whether any methods based on fuzzy graph theory rather than the connected-triple may be more suitable and efficient in dealing with the proposed fuzzy graphs [4].

Y. Sahin, et al. proposed the classification models used on Artificial Neural Networks (ANN) and Logistic Regression (LR) are developed and apply to credit card fraud detection problem. In this study work is one of the firsts to compare the presentation of ANN and LR methods in credit card fraud detection with an original data set. The survey demonstrates the advantages of applying the data mining techniques including ANN and LR for the credit card fraud detection problem for the purpose of reducing the bank’s risk. The outcomes indicate that the proposed ANN classifiers outperform LR classifiers in solving the problem under investigation [5].

Rinky D, et al. was trying to detect fake transactions through the genetic algorithm. A genetic algorithm is used for creation the decision about the number of hidden layers, network topology, and number of nodes that will be used in the plan of artificial neural network for our problem of credit card fraud detection. In that respect are different techniques implemented by fraudsters to commit credit card fraud. If they can properly interpret the concepts and distinguish the mechanism of fraud, then they can easily contend with credit card frauds [6].

S. Benson Edwin Raj, et al. presents a study of several techniques used in credit card fraud detection mechanisms and evaluates each methodology based on certain design standards. Darwinian fraud detection systems in conditions of true positive are 100% and shows good results in detecting fraudulent transactions. The Artificial neural network based CARDWATCH shows good accuracy in fraud detection and Processing Speed is also high, but it is restricted to one-network per customer. The Artificial neural network and BNN are used to detect Mobile phone fraud, Network Intrusion. All the techniques of credit card fraud detection discussed in this survey paper have its own forces and failings. Such a survey will enable us to establish a hybrid approach for identifying fraudulent credit card transactions [7].

V. Bhusari, et al. was exposed that credit card fraud can be detected by Hidden Markov Model through transactions. In the Hidden Markov Model helps to receive a high level fraud coverage combined with a low false alarm rate. The Fraud Detection System is also scalable for handling huge volumes of transaction processing. The Hidden Markov Model based credit card fraud detection system is not calling for a long time and having a complex procedure to do fraud checks like the existing organization and it feeds better and quicker result than existing schemes. The Hidden Markov Model makes the cognitive operation of detection very easy and tries to take away the complexity. [3]

Abhinav Srivastava, et al. proposed model the sequence of operations in credit card transaction dispensation using a Hidden Markov Model and prove it can be applied for the detection of credit card frauds. An Hidden Markov Model is initially trained with the normal actions of a cardholder. If an received credit card transaction is not accepted by the trained Hidden Markov Model with a sufficiently high possibility, it is considered to be fraudulent. At the same time, they attempted to ensure that genuine transactions are not eliminated. They presented detailed experimental results to demonstrate the usefulness of our approach and compare it with other techniques available in the literature. [1]

MohdAvesh, et al. present a paper was survey of various techniques used in credit card fraud detection mechanisms and Hidden Markov Model (HMM) in detail. HMM categorizes card holder's outline as low, medium and high spending based on their spending behavior in terms of quality. A set of probabilities for amount of transaction is being assigned to each cardholder. The Amount of each incoming transaction is then fitted with the card owner's category, if it justifies a predefined threshold value, then the transaction is determined to be legitimate else declared as fraudulent. Existing fraud detection system may not be so much capable to reduce fraud transaction rate. Development in fraud detection practices has become necessary to continue the existence of the payment system. In this paper Hidden Markov Model is used to model the succession of operation in credit card transaction dispensation. If an incoming credit card transaction is not convened by the qualified Hidden Markov Model with an adequately high probability, it is measured to be fraudulent. [2]

III.METHODOLOGY

K-means is an unsupervised learning algorithm is applied on this dataset using mat lab language. K-means is formed a grouping a given data based on the similarity (amount) in their values is called a cluster. The grouping is performed by minimizing sum of squares of distances between each data point and the centroid of the clusters to which it belong. In our system considering three clusters such as clow, cmedium and chigh as the respective centroids. The total number of transactions in the cluster in thus 80 percent, similarly amount 20000, 26000, 21000, 23000, 27000, 29000, 25000, 22000, 28000, 30000 have been grouped in the cluster cmedium with centroid 25000

The spending profile of card holders suggests his normal spending behavior can be broadly categorized into three groups based on their spending habits, namely high spending group, medium spending group and low spending group. When the k-means receives a transaction T for this cardholder it measures the distance of the purchase amount X with respect to the means clow, cmedium and chigh to decide the cluster to which T belongs.

Fuzzy logic has rapidly become one of the most successful of today's technologies for developing sophisticated control systems. The reason for which is very simple. Fuzzy logic has three important elements that should be executed as a sequence, namely, Fuzzification, Rule Based and Defuzzification. In Fuzzification, the transaction T must be converted into the linguistic variables such as low transaction is considered as 0 to20000, medium transaction is considered as 20000 to 50000or high transaction is considered as 50000 to 100000.

Rule based, according to past behavior and trend of the customer the rules has been set up to predict the usual shopping behavior. The transaction is allowed to perform by analyzing the past performance by k-means clustering and matching the similar characteristics of the present transaction of the customer with the past.

Defuzzification, if the rules prescribe for the particular customer mismatch with the present characteristics of the transaction then it will be not allowed performing. Fuzzy logic is to model the human behavior. Once human behavior is correctly modeled since an attacker is not expected to have behavior similar to the genuine user. Hence the transaction is partially stopped and then sends one time password (OTP) code. The authorized person automatically generates the transaction or otherwise stops the transaction

IV. EXPERIMENTATION

As, the credit card data set is highly confidential and accountable, the banker's hesitates to reveal their customers' data. Hence, for conducting this experimentation a data set was manually created as per the existing works. The dataset consists of eight attributes and 10 instances, with which six

Attribute are numerical and two are categorical. The attributes and their description are given in Table 1.

TABLE 1: CREDIT CARD FRAUD DETECTION DATA SET

S. No	Attribute Name	Description
1	t.no	Transaction no
2	Date	What date customer bought
3	Time	When customer bought
4	amount	How much amount

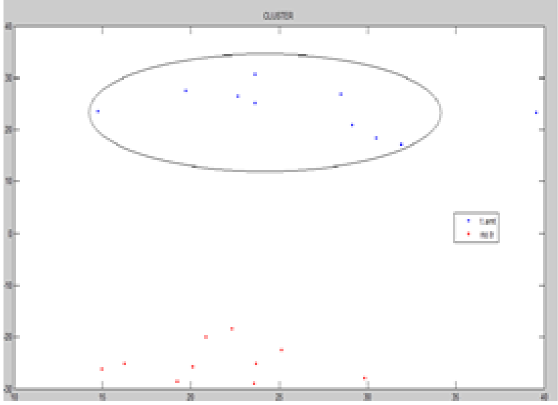
		customer bought
5	maximum limit	To limit the transaction amount
6	Age	Customer's Age
7	s.a	Customer's Shipping Address
8	Add	Customer's Address

The dataset was preprocessed and experimented using trained data. The analytical factors that had chosen for identifying credit fraud were t.no, amount, and s.a and add. The Fuzzy membership function was created to trace out the abnormal activities in the behavioral change of the customer. Example: Maximum number of purchase in single day which is dissimilar from previous purchases, Change of Shipping address. Apart from that, if the amount spent by the customer over a period of month or year varied irregularly also could be considered a fraud.

V. RESULT AND DISCUSSION

The perceived result showed that the fuzzy and K-means algorithm clustered which was away from two centroids at the same distance was clustered in each two clusters. Hence, the Predicting fraudulent transactions are not the mere objective of clustering rather, it is fine to give the solution for such credit card frauds. The effectiveness of fraud alarms could intimate the customers in two ways, (i.e.) either during transaction or after the transaction. But it so effective it alerts the user during transaction. Moreover, An OTP (One Time Password) system can be introduced to grant the user to complete their transaction in case of suspicious credit card use. The cluster only forms number of transactions and amounts, because fuzzy logic techniques only trained the customer behavior (amount of purchase).

FIGURE 1: FRAUD TRANSACTION DETECION



VI. CONCLUSION

The customer performance enables companies to get better support their customer oriented business processes, which aims to improve the overall presentation of the enterprise. This paper focuses on getting more customer satisfaction. Data mining methodology has a great contribution for researchers to take out the hidden knowledge and information. The research described in this paper also recognized significant behavior for each section. Suppose commercial banks hope to add to the customers who will propose credit card with large number of important customer information in huge amounts of data , which is used to identify customers and provide decision support can apply the methods mentioned above such as classification and clustering. In this paper customer behavior has been analyzed using clustering algorithms. The clustering techniques have been used in this paper, fuzzy logic and K-means is concluded that security plays a vital role in proposing credit card. One who possesses a credit card for more than three years feels insecure in using credit card or plastic money. From this work, it is understood that maximum number, credit card proposers belongs to urban region and their age is from 31 to 40. They use credit card for both foreign and domestic use.

REFERENCE

[1] Abhinav Srivastava, Amlan Kundu, Shamik Sural and Arun K. Majumdar "Credit Card Fraud Detection Using Hidden Markov Model. IEEE transactions on dependable and secure computing," vol.5, no. 1, january-march 2008.

[2] MohdAvesh Zubair Khan, JabirDaud Pathan and Ali Haider Ekbal Ahmed, "Credit Card Fraud Detection System Using Hidden Markov Model and K-Clustering". International Journal of Advanced Research in Computer and Communication Engineering Vol. 3, Issue 2, February 2014

[3] V. Bhusari and S. Patil, ``Study of Hidden Markov Model in Credit Card Fraudulent Detection``, . International Journal of Computer Applications (0975 – 8887) Volume 20– No.5, April 2011.]. V. Bhusari and S. Patil: Study of Hidden Markov Model in Credit Card Fraudulent Detection. International Journal of Computer Applications (0975 – 8887) Volume 20– No.5, April 2011.

[4] Pan Su, Changjing Shang, Qiang Shen,”Link-based Pairwise Similarity Matrix Approach for Fuzzy C-means Clustering Ensemble`’IEEE International Conference on Fuzzy Systems (FUZZ-IEEE) July 6-11, 2014.

[5] Y. Sahin and E. Duman,” Detecting Credit Card Fraud by ANN and Logistic Regression”IEEE 978-1-61284-922-5/11/$26.00 ©2011.

[6] Rinky D. Patel and Dheeraj Kumar Singh,”Credit Card Fraud Detection & Prevention of Fraud Using Genetic