Global Journal of Engineering Science and Research Management

IMPROVING CREDIT CARD FRAUD DETECTION SYSTEM USING K-MEANS CLUSTERING ALGORITHM

Ms. M.Kalaimani[*1] Ms. R.Ranjani[2] Mr. P.Thirugnanam[3]
[1*] M.E, Assistant Professor, Department of CSE, IFET College of Engineering, Tamil Nadu, India.
[2]M.E, Assistant Professor, Department of CSE, IFET College of Engineering, Tamil Nadu, India.
[3]M.E, Assistant Professor, Department of CSE, IFET College of Engineering, Tamil Nadu, India.
*Correspondence Author: **Ms. M.Kalaimani**

Keywords: K-means clustering algorithm, Credit card fraud.

Abstract
Now a day the usage of credit cards has been dramatically increased. 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 I propose a system for credit card fraud detection and tried to improve the performance of an existing system. In doing so, we did not undertake the typical objective of maximizing the number of correctly classified transactions but rather we defined a new objective function where the misclassification costs are variable and thus, correct classification of some transactions are more important than correctly classifying the others. This proposed model makes use of k-Means Clustering algorithms which is a novel one in the related literature, both in terms of the application domain and the cross-over operator used. The algorithm is applied to real life data where the savings obtained are almost three times the current practice. At the same time, we try to ensure that genuine transactions are not rejected. Here by I presented a detailed experimental result to show the effectiveness of our approach and compare it with other techniques available in the literature.

Introduction
Credit Card Fraud is one of the major threats to business establishments today. However, to combat the fraud effectively, it is important to first understand the mechanisms of executing a fraud. Fraud detection includes analyzing of the spending behavior of users/customer order purpose, uncovering, or escaping of undesirable behavior. As credit card becomes the most general mode of payment or both online as well as regular purchase, fraud relate with it are also accelerate. Fraud detection is concerned with not only capturing the deceptive events, but also capturing of such activities as rapidly as possible. The use of credit cards is common in modern day society. Fraud is a millions dough business and it is rising every year. Fraud presents significant cost to our financial prudence measure worldwide .Modern techniques based on Data mining, Machine learning, Sequence Alignment technique, Fuzzy Logic, Genetic Programming, Artificial Intelligence (AI) etc. has been introduced for detecting & preventing credit/ATM card, CHEQUE book type of fraudulent transactions. Fraud detection based on the analysis of existing purchase data of cardholder is a promising way to reduce the rate of successful credit card frauds.

This project shows K-Mean Clustering Algorithm & data mining techniques are used for fraud prevention there by implementing as/which ask secret questions i.e. ATM feedback SMS system with reply & by thumb impressions instead of detecting a fraud, a fraud can also be prevented.

Credit-card-based purchases can be categorized into two types: 1) physical card and 2) virtual card. In a physical-card based purchase, the cardholder presents his card physically to a merchant for making a payment. To carry out fraudulent transactions in this kind of purchase, an attacker has to steal the credit card. If the cardholder does not realize the loss of card, it can lead to a substantial financial loss to the credit card company.

In the second kind of purchase, only some important information about a card (card number, expiration date, secure code) is required to make the payment. Such purchases are normally done on the Internet or over the telephone. To commit fraud in these types of purchases, a fraudster simply needs to know the card details. Most of the time, the genuine cardholder is not aware that someone else has seen or stolen his card information. The only way to detect this kind of fraud is to analyze the spending patterns on every card and to figure out any inconsistency with respect to the "usual" spending patterns. Fraud detection based on the analysis of existing purchase data of cardholder is a promising way to reduce the rate of successful credit card frauds.

Types of frauds
Various types of frauds include credit card frauds, telecommunication frauds, and computer intrusions, Bankruptcy fraud, Theft fraud/counterfeit fraud, Application fraud, Behavioral fraud

Global Journal of Engineering Science and Research Management

Credit card fraud has been divided into two types: Offline fraud and On-line fraud. Offline fraud is committed by using a stolen physical card at call center or any other place. On-line fraud is committed via internet, phone, shopping, web, or in absence of card holder.

The use of telecommunication services to commit other forms of fraud. Consumers, businesses and communication service provider are the victims.

Intrusion Is Defined As The Act Of Entering Without Warrant Or Invitation; That Means "Potential Possibility Of Unauthorized Attempt To Access Information, Manipulate Information Purposefully. Intruders May Be From Any Environment, An Outsider (Or Hacker) And An Insider Who Knows The Layout Of The System.

Bankruptcy fraud means using a credit card while being absent. Bankruptcy fraud is one of the most complicated types of fraud to predict.

Theft fraud refers using a card that is not yours. As soon as the owner give some feedback and contact the bank, the bank will take measures to check the thief as early as possible. Likewise, counterfeit fraud occurs when the credit card is used remotely; where only the credit card details are needed

When someone applies for a credit card with false information, is termed as application fraud. For detecting application fraud, two different situations have to be classified. When applications come from a same user with the same details, that is called duplicates, and when applications come from different individuals with similar details, that is termed as identity fraudsters. Phua et al. (2006) describes application fraud as "demonstration of identity crime occurs when application form(s) contain possible and synthetic (identity fraud), or real but also stolen identity information (identity theft)."

CREDITCARD FRAUD DETECTION METHODS:
On doing the literature survey of various methods for fraud detection we come to the conclusion that to detect credit card fraud there are multiple approaches like.

- A Fusion Approach Using Dempster-Shafer Theory and Bayesian Learning.
- Blast-Ssaha Hybridization
- Hidden Markov Model.
- Neural Network
- Bayesian Network
- Genetic Algorithm
- Artificial Immune System
- K- nearest neighbor algorithm
- Support Vector Machine
- Decision Tree
- Fuzzy Logic Based System
- Meta Learning Strategy

Fraud detection methods

The detection of fraud is a complex computational task and still there is no system that surely predicts any transaction as fraudulent. They just predict the likelihood of the transaction to be a fraudulent.
The properties of a good fraud detection system are:

1) It should identify the frauds accurately
2) It should detecting the frauds quickly
3) It should not classify a genuine transaction as fraud

Related work

V.Chandola, A.Banerjee [1] anomaly detection a survey In this domain, anomaly detection techniques are applied to detect fraudulent credit card applications or fraudulent credit card usage(associated with credit card thefts). Detecting fraudulent credit card applications is similar to detecting insurance fraud The data is typically comprised of records defined over several dimensions such as user ID, amount spent, time between consecutive card usage, and so forth. The frauds are typically reflected in transactional records (point anomalies) and correspond to high payments, purchase of items never purchased by the user before, high rate of

Global Journal of Engineering Science and Research Management

purchase, and so forth. The credit companies have complete data available and also have labeled records. Moreover, the data falls into distinct profiles based on the credit card user. Hence profiling and clustering based techniques are typically used in this domain Jaba Suman Mishra, Soumyashree Panda, [5] A Novel Approach for Credit Card Fraud Detection Targeting the Indian Market, credit card fraud detection based on Hidden Markov Model, which does not require fraud signatures and still it is capable to detect frauds just by bearing in mind a cardholder's spending habit. The particulars of purchased items in single transactions are generally unknown to any Credit card Fraud Detection System running either at the bank that issues credit cards to the cardholders or at the merchant site where goods is going to be purchased. As business processing of credit card fraud detection system runs on a credit card issuing bank site or merchant site. Each arriving transaction is submitted to the fraud detection system for verification purpose. The fraud detection system accept the card details such as credit card number, cvv number, card type, expiry date and the amount of items purchase to validate, whether the transaction is genuine or not. The implementation techniques of Hidden Markov Model in order to detect fraud transaction through credit cards, it create clusters of training set and identify the spending profile of cardholder. The number of items purchased, types of items that are bought in a particular transaction are not known to the Fraud Detection system, but it only concentrates on the amount of item purchased and use for further processing.
Anshul Singh, Devesh [6]

Proposed system

The architecture diagram for identifying the fraud detection system is given below

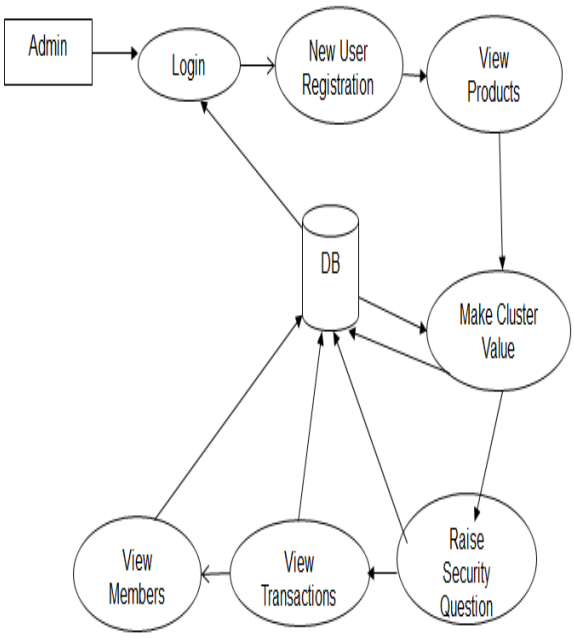

Fig.1 Process flow of credit card fraud detection system

In proposed system, we present k-Means Clustering algorithm is able to detect frauds by considering a cardholder's pending habit. Card transaction processing is sequence by the stochastic process of k-Means Clustering algorithms. The details of items purchased in Individual transactions are usually not known to any Fraud Detection System (FDS) running at the bank that issues credit cards to the cardholders. Hence, we feel that k-Means Clustering algorithms are an ideal choice for addressing this problem. Another important advantage of the k-Means Clustering algorithms based approach is a drastic reduction in the number of False Positives transactions identified as malicious by an FDS although they are actually genuine. An FDS runs at a credit card issuing bank. Each incoming transaction is submitted to the FDS for verification. FDS receives the card details and the value of purchase to verify, whether the transaction is genuine or not. The types of goods that are bought in that transaction are not known to the FDS. It tries to find any anomaly in the transaction based on the spending profile of the cardholder, shipping address etc.

Global Journal of Engineering Science and Research Management

K-means clustering algorithm

The K-Means Clustering Algorithm consists of basic steps. In this algorithm we initially determine the number of clusters present, assume it to be K and we also assume the center or centroid of these clusters. Now we can consider a random object as the initial centroids or we can also consider the sequence of first K objects as the centroids. Later the K-Means algorithm will carry out the iteration of below stated 3 steps till the convergence.

Step1: Determine the centroid coordinate

Step2. Determine the distance of each object to the centroids.

Step 3.Group the object based on minimum distance (find the closest centroid).

K-Means algorithm is used to fetch the list of records from the database in order to provide security to the transactions made by the user using the fetched records. A cluster value is formed using the K-Means algorithm from the fetched records. In this project the each card holder have a maximum transaction amount limit per day. While a new transaction is done by the user the transaction amount is compared to the previous amount and also to the maximum limit per day. The application tracks the user transactions for the first five times. While tracking the user transactions a cluster value is formed by the first five transactions. The formed cluster value may be in min, mid and max. After forming a cluster value, the application compares the new transaction amount with the cluster value. If the transaction amount is more than the cluster value then it raises a security question otherwise it simply redirects to the next. If the user does not enter a correct answer then the card will be blocked and user transaction also terminated.

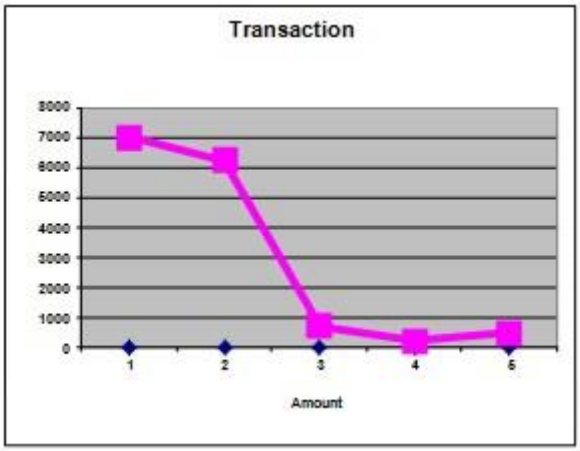

Fig. 2. Graph representing No. of Transactions with Amount tabulated below

TRANSACTION	1	2	3	4	5
AMOUNT	7000	6250	750	250	500

Conclusion

In this paper, we have proposed an application of K-Mean algorithm in credit card fraud detection keeping in view the current Indian Market. We have used different ranges of transaction amount as the observation symbols whereas the types of items have been considered to be states of the K-Mean Algorithm. We have suggested a method for finding the Spending Profile of the Cardholders as well as application of this knowledge in deciding the observation symbols. We have found that more than 85% transactions are genuine. The proposed Fraud Detection System is also scalable for handling efficient data transactions with security.

References

1. *V. Chandola, A. Banerjee, and V. Kumar, "Anomaly Detection: A Survey" ACM Computing Surveys, vol. 41, no. 3, pp. 15:1-15:58, 2009.*
2. *Aleskerov, E., Freisleben, B. & B Rao. 1997. "CARDWATCH: A Neural Network-Based Database Mining System for Credit Card Fraud Detection", Proc .of the IEEE/IAFE on Computational Intelligence for Financial Engineering, 220-226.*

Global Journal of Engineering Science and Research Management

3. *X. Song, M. Wu, and C.J., and S. Ranka, "Conditional Anomaly Detection," IEEE Trans. Knowledge and Data Eng., vol. 19, no. 5,pp. 631-645, May 2007.*
4. *Sonali N.Jadhav, ,Kiran Bhandari , "Anomaly Detection Using Hidden Markov Model", ,International Journal of ComputationalEngineering Research Vol, 03 Issue, 7 july 2013*
5. *Jaba Suman Mishra, Soumyashree Panda, Ashis Kumar Mishra, "A Novel Approach for Credit Card Fraud Detection Targeting the Indian Market", ,IJCSI International Journal of Computer Science Issues, Vol. 10, Issue 3, No 2, May 2013*
6. *Anshul Singh, Devesh Narayan, "A Survey on Hidden Markov Model for Credit Card Fraud Detection", International Journal of Engineering and Advanced Technology ISSN: 2249 – 8958, Volume-1, Issue-3, February 2012*