International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 05 Issue: 11 | Nov 2018 www.irjet.net p-ISSN: 2395-0072

SURVEY ON CREDIT CARD FRAUD DETECTION

ASWATHY M S[1], LIJI SAMEUL[2]

[1]*ASWATHY M S, M.Tech Computer Science & Engineering. Sree Buddha College of Engineering, Ayathil, Elavumthitta Pathanamthitt , Kerala, India.*

[2]*Ms. LIJI SAMEUL, Assistant Professor Computer Science & Engineering. Sree Buddha College of Engineering, Ayathil, Elavumthitta, Pathanamthitta, Kerala, India*

Abstract - *Financial Services have serious problems due to credit card fraud. Lot of money is lost due to credit card fraud every year. As the technology is increasing day by day the financial fraud is also increasing. As a result of all this financial loss due to financial fraud is increasing day by day. In order to reduce this problem, fraud detection has become an important tool and probably the best way to stop such frauds. In this study various fraud detection techniques has been employed.*

***Key Words*: Fraud Detection, TPR, HMM, FDA, SOM**

1. INTRODUCTION

Fraud is defined as a wrongful or criminal deception which is aimed to bring financial or personal gain. Two mechanisms are used to avoid fraud and losses due to fraud. They are Fraud Prevention and Fraud Detection. Fraud Prevention is a proactive method where it stops fraud from being happening. Fraud Detection is used when a fraudulent transaction is attempted by the fraudster. The well-known fraud domain is credit card systems. Credit card fraud can be made in many ways such as simple theft, application fraud etc. The transactions with credit card can be accomplished in two ways. They are physical transaction and digital transaction. In case of physical transaction credit card is involved directly whereas in case of digital transactions card is not used, the transaction happens through internet or telephone.

Fraud detection using credit card is an extremely difficult task and is very difficult to detect. Many approaches has been proposed to solve this problem. The most commonly used fraud detection methods are rule induction technique, decision tree, Artificial Neural Network, Support Vector Machine, Logistic Regression and genetic algorithms are used. The famous algorithm used for fraud detection is Neural Network. These algorithms can be used as single model or can be used in combination. These machine learning algorithms are successful in many cases but still cannot generate accurate result.

2. LITERATURE SURVEY

2.1 A Cost-Sensitive Decision Tree Approach for Fraud Detection

As the information technology is developing the fraud is also increasing as a result financial loss due to fraud is also very large. A cost sensitive decision tree approach has been used for fraud detection. A cost called misclassification cost is used which is taken as varying as well as priorities of the fraud also differs according to individual records. So common performance metrics such as accuracy, True Positive Rate (TPR) or even area Under Curve cannot be used to evaluate the performance of the models because they accept each fraud as having the same priority regardless of the amount of that fraudulent transaction or the available usable limit of the card used in the transaction at that time. For avoiding this a new performance metric which prioritizes each fraudulent transaction in a meaningful way and it also checks the performance of the model in minimizing the total financial loss. The measure used is Saved Loss Rate (SLR) which is the saved percentage of the potential financial loss that is the sum of the available usable limits of the cards from which fraudulent transactions are committed.

Different methods are used for cost sensitivity. They mainly include the machine learning approach, decision tree approach. In machine learning approach two techniques called over sampling and under sampling is performed, in which the latter obtained a good result. In decision tree approach, decision tree algorithms are used in which misclassification cost is considered in pruning step. A cost matrix is used to find the varying misclassification cost. After finding the misclassification cost the one with minimum value is used. By finding the misclassification cost not only the node value is obtained but also it predicts whether the transaction is fraudulent or not. This study using misclassification cost has made a significant improvement in fraud detection.

2.2 Credit Card Fraud Detection Techniques: Data and Technique Oriented Perspective

In this study mainly two approaches namely misuses (supervised) and anomaly detection (unsupervised) technique is being used. After this a classification is also used for checking the capability to process categorical and numerical data. In the first approach the data is classified as fraud based on previous data. With the help of this dataset classification models are also created, which can predict whether the data is fraud or not. The different classification models used are decision tree, neural network, rule induction etc. This has obtained a successful result and this approach is also called as misuse approach. While the second approach is based on account behavior. A transaction is said

to be fraudulent if it possess the features opposite to the user's normal behavior. The behavior of user's model are extracted and accordingly classified as fraudulent or not. This technique of finding fraud is also called as anomaly detection.

2.3 Credit Card Fraud Detection Using Hidden Markov Model

As the E-commerce technology is increasing day by day the use of credit card has also been increased. As a result of this the fraud using credit card is also increasing. In all fraud detection systems, fraud will be detected only after the fraud has taken place. In this study a sequence of operations are modelled using Hidden Markov Model (HMM) and this can be used for the detection of fraud. It is trained with the normal behavior of the card holder. If the incoming transaction is not accepted by the trained HMM with high probability it is considered as fraudulent otherwise not.

A hidden Markov Model represents a finite number of states with sufficiently high probability. The transition between the states are handled by these probability values. A possible outcome will be generated based on the probability distribution. This outcome will be visible to the external users that is the states are hidden to the users hence the name. It is a perfect solution for predicting fraud transactions in addition it also provides extreme decrease in the number of false positive transactions recognized by fraud detection system. For prediction purpose three values are being used namely low, medium, high.

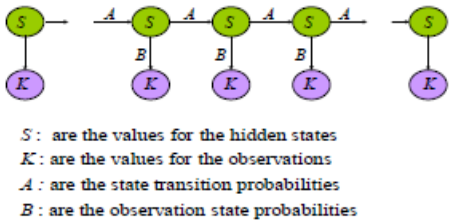

Fig-1: HMM Finite set of states

The main advantage is that it does not require fraud signatures it is capable to detect by bearing in mind card holders spending habit. This is done by creating a set of clusters and identify the spending profile. These data are stored in the form of clusters with low, medium and high values. The probability is based on the spending behavior and further the processing is done. If the transaction is found to be fraudulent an alarm is generated, in addition to this a security form will also arise bearing a certain number of questions. This model can detect fraud transactions to an extent. It is scalable in handling large amount of data.

2.4 Real Time Credit Card Fraud Detection using Computational Intelligence

As the growth of technology is increasing it has made a big impact in credit transactions. This has made the fraudsters to commit the fraud transactions easily. Many fraud detection techniques are available but cannot solve fraud problems easily. As a solution to all this, SOM (Self Organizing Map) is used. It helps to decipher, filter and analyze customer behavior for fraud detection. SOM is a unsupervised neural network algorithm which is used to configure neurons according to the topological structure of input data. It divides the data into genuine and fraudulent transactions sets. The incoming transactions are compared with the previous transactions in the genuine set, then it is called as genuine set. The incoming transactions are compared with the previous transactions in the fraudulent transaction, then it is called as fraudulent set. So in this it is important to form legal card holder and fraudulent profile. For fraud detection a layered approach is used, which is depicted in the below figure.

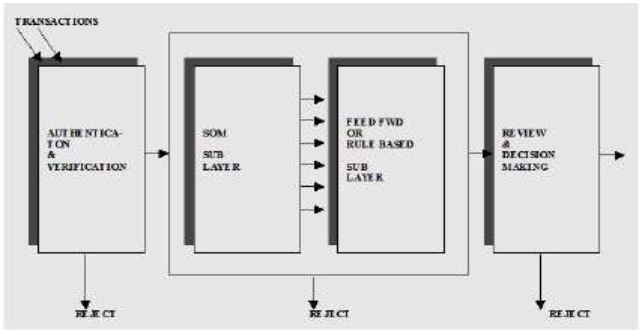

Fig-2: System Layers

The initial stage consist of authentication and screening layers. The fraudsters always comes with new ideas to commit fraud. The key to fraud detection lies in finding such a dynamic system to detect fraud that changes with the e-commerce trends. This system not only deals with customer profiles, merchant profiles and their selling price. It should also include rules and policies in the market place. By involving these attributes it increases the accuracy to detect the fraud. SOM helps to detect fraud to a great extent.

2.5 A Novel Machine Learning Algorithm to credit card fraud detection

The use of credit card is rising day by day as the e-commerce is also increasing. The problem that happens with this is that fraud using credit card is increasing. It is a recurrent problem in almost all countries. But the trend seen is that countries with more credit card transactions are having less credit card fraud on the other hand countries with average credit card transactions are having high rate of credit card fraud. So in order to avoid this proactive methods are needed. In this a novel machine learning algorithm called cortical algorithm is used. It is the learning algorithm of hierarchical temporal memory and is inspired by the neo cortex of the brain.

This is a new approach for prediction and anomaly detection. It works on data obtained from UCI repository. The methodology is that it converts highly populated data into sparse representations and uses learning columns to learn spatial and temporal patterns. It mainly follows three

steps in analyzing and predicting data from the streaming input. The steps included are

i) A sparse representation of the input is initially formed.
ii) A representation is formed based on the previous input.
iii) Finally a prediction is made based on previous step.

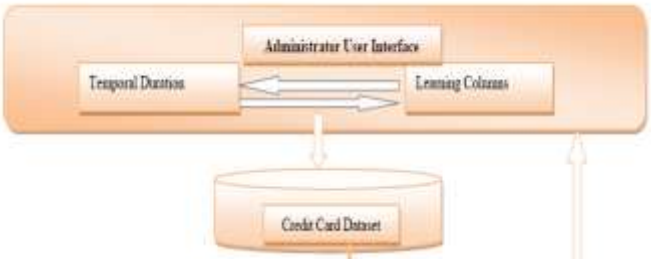

Fig- 3: System Architecture

The system architecture is described in the figure. It includes a user interface, Duration, Learning columns and a dataset. The duration and learning columns are used for analyzing the data and it is compared with the credit card data set which is obtained from the UCI repository. After analyzing and comparing the final step is predicting whether the transaction is fraudulent or not. This study provides an effective way to detect credit card fraud transactions.

2.6 Credit Card Fraud Detection: A Fusion Approach using Dempster–Shafer Theory and Bayesian Learning

In this evidences from current as well as past behaviour are combined. A fraud detection system is proposed that includes rule based filter, Dempster Shafer adder, transaction history database and Bayesian learner. In rule base the suspicion level of each incoming transaction is determined. Dumpster Shafer is used to combine multiple such evidences and an initial belief is computed. Based on this belief the transactions are classified as normal, abnormal or suspicious. The incoming transactions are initially handled by the rule base using probability values. After this the values are combined using Dumpster Shafer Adder. If the transaction is declared as fraudulent then it is handled by the card holder. If suppose the transaction is suspicious then it is fed in the suspicious table.

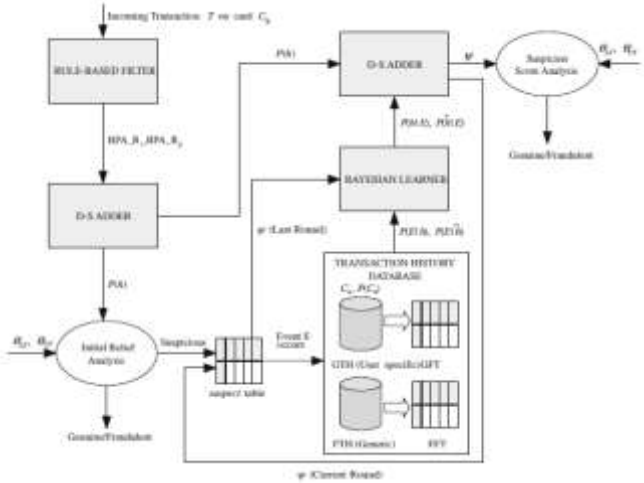

Fig – 4: System Architecture

The score of transaction is updated in the database with the help of Bayesian classification. This architecture is flexible such that new kinds of fraud can be handled easily. With the help of Bayesian learner the system can dynamically adapt to the changing needs.

2.7 Detecting Credit Card Fraud by Modified Fisher Discriminant Analysis

This employees a linear discriminant called fisher Discriminant. The Linear discriminant is a supervised learning algorithm in which the input region is divided into boundaries called decision boundaries or decision surfaces. The discriminant algorithm is modified to update the weights. This method tries to find the best dimensional hyper plane by which the within class variance is minimized to reduce the overlap and between class variance is maximized.

It is a kind of supervised learning algorithm in which the input region is divided into decision regions where the boundaries are called decision boundaries. These boundaries are linear function of input vector x. There is actually a comparison between fisher discriminant and modified fisher discriminant. FDA captures more number of positive cases whereas modified FDA gets more profit by classifying the most profitable transactions. In total FDA can give more number of fraud transactions whereas modified FDA relies on maximizing total profit.

This method can label transactions with high usable limits on the cards correctly which leads to prevent losing millions of dollars in real life banking systems. For developing iterative linear discriminant, Linear Perceptron Discriminant function is being used which can solve all the problems related with modified FDA.

3. CONCLUSION

Fraud detection using credit card is a very serious problem in financial services. The loss due to credit card fraud is increasing with the increase in e-commerce. This study deals with techniques that helps to find out the credit card fraud. Various techniques like decision tree, Computational Intelligence, Cortical Learning Algorithm, Modified Fisher Discriminant approach and a fusion approach using Dumpster Shafer and Bayesian Learning is also used.

REFERENCES

[1] Aswathy M S, Liji Sameul "Survey on Credit Card Fraud Detection".

[2] Y. Sahin, S. Bulkan, and E. Duman, ``A cost-sensitive decision tree approach for fraud detection,'' Expert Syst. Appl., vol. 40, no. 15, pp. 5916_5923, 2013.

[3] A. Srivastava, A. Kundu, S. Sural, and A. Majumdar, ``Credit card fraud detection using hidden Markov

model,'' IEEE Trans. Depend. Sec. Comput., vol. 5, no. 1, pp. 37, Jan. 2008.

[4] J. T. Quah and M. Sriganesh, ``Real-time credit card fraud detection using computational intelligence,'' Expert Syst. Appl., vol. 35, no. 4, pp..

[5] S. Panigrahi, A. Kundu, S. Sural, and A. K. Majumdar, ``Credit card fraud detection: A fusion approach using Dempster Shafer theory and Bayesian learning,'' Inf. Fusion, vol. 10, no. 4.

[6] N. Mahmoudi and E. Duman, ``Detecting credit card fraud by modified fisher discriminant analysis,'' Expert Syst. Appl., vol. 42, no. 5.

[7] A. O. Adewumi and A. A. Akinyelu, ``A survey of machine-learning and nature-inspired based credit card fraud detection techniques,'' Int. J. Syst. Assurance Eng. Manage., vol. 8, no. 2.

[8] N. S. Halvaiee and M. K. Akbari, ``A novel model for credit card fraud detection using artificial immune systems,'' Appl. Soft Comput., vol. 24.

BIOGRAPHIES

Aswathy M S, She is currently pursuing her Masters degree in Computer Science and Engineering in Sree Buddha College Of Engineering, Kerala ,India. Her area of research include Intelligence, Data Mining and Security.

Liji Sameul, She is an Assistant Professor in the Department of Computer Science and Engineering, Sree Buddha College Of Engineering. Her main area of interest is Data Mining.