Evaluation of Fraud Detection Techniques

A.M.Chandrashekhar[1], Pushpalatha C[2]
[2] Dept Of Computer Science & Engineering
[1]Asst. Prof. Dept.Of Computer Science & Engineering
[1, 2] , Sri Jayachamarajendra College of Engineering(SJCE), JSS S&T University Campus,
Mysore, Karnataka, India

Abstract- *In current situation, the term fraud is the most trendy and challenging topic in context to the financial transaction involving credit card. Fraud detection includes auditing of the spending style of users or customers in order to figure out detection or control of unreachable style. As credit card becomes the most universal medium of payment for both website as well as ordinary investment, fraud connected with it are also accelerating. advance techniques based on data mining, machine learning, sequence alignment, fuzzy logic, genetic programming, artificial intelligence has been generated for credit card fraudulent transactions. This paper compromise with how data mining procedure can be united beneficially to obtain a high fraud coverage merged with low or high false alarm rate.*

Keywords- Decision tree, Neural Network, Logistic regression

I. INTRODUCTION

Fraud is an illegal way of accessing the services. Fraud accord with events which is concerned to criminal motives that, mostly, are difficult to find out. Credit card fraudulence, is a phrase which is across-the-board for theft and fraudulence devoted .Recognizing credit card fraudulence is a challenging work in normal case, thus the development of detection models for fraud plays a prominent role in academic or business institution presently .Moreover due to advancement in technologies the role of fraud has been variant during last few decades[1].

The Fraud of credit card is one of the immense hazard for business and financial formulation accumulated huge volume in account transactions of credit card. The plastic card distributed to multiple users is termed as credit card which acts as medium of payment. It grants cardholders to buy required stuff and services depending on cardholder's affirmation. Secure credit servicing of banks and development of E-business a dependable detection system of fraud is necessary to guide secure usage of credit card, depending on analysis of cardholder 's existing acquisition reports of stuff serves as a assuring approach to reduce the amount of fine points of appropriate cards results in getting spurious credit card.

The systems of fraud detection exists when the fraudsters top the prevention systems of fraud and begin fraudulent transactions. Along with the advancement in the Information Technology and progress in the intercommunication medium, fraud is widening across the world yielding huge extent of fraudulent loss.

Different techniques in Knowledge Disclosure, such as Neural Network , decision tree and case based proposition is used for composing different systems/ models of fraud detection. These techniques generally require sufficient number of normal and fraud transactions to know the patterns of fraud. Though, the ratio of fraudulent transactions to its normal transactions is low excessively, for the separate bank . Credit Card Fraud is defined as, "when a person uses other individual's‟ credit card for his exclusive use where the card owner and card issuer are not aware of card misuse.

II. FRAUD CLASSIFICATION

This section gives the various types of fraud occurring in the real world.

Credit Card Fraud : Credit card fraud kinds : offline and On-line fraud.

- Offline fraud is adhered of using the card which does not owned by that individual

- On-line fraud is adhered via mobile, shopping, web, or in absence of card holder.
- Telecommunication Fraud: The utilization of telecommunication services to enact other forms of fraud. Consumers, businesses and communication service provider are the victims.

Computer Intrusion: The Action Of arrival Without assurance is termed as intrusion, means "Potential Possibility Of Unofficial try to retrieve Information, modifying information intentionally. Invaders can be from diverse field. Computer intrusions categories are: misuse intrusions, network intrusions and host intrusions.

Bankruptcy Fraud: The most tough kind of fraud to estimate .The bank will dispatch its users/customers an order to pay. However, the users will be recognized as being in a state of personal bankruptcy and not able to recover their unwanted loans. The bank will have to cover the losses itself. One of the possible ways to prevent bankruptcy fraud is by doing a pre-check with credit bureau in order to be informed about the past banking history of its customers.

Theft Fraud/ Counterfeit Fraud : This section is concerned about theft and counterfeit fraud. where theft fraud is the condition is when a person use the card which does not belong to him. As the owner notice this he provide his review to bank where certain actions regarding the user notice and determine the person who has done malpractice. The remote usage of credit card is termed as counterfeit fraud which requires only card information.[17] The list of customers who owns credit card are approached if the response delayed then the card is blocked.

III. FRAUD DETECTION OF CREDIT CARD

This section provides the conceptual view and real world difficulties.

Credit Card: It is a mode of transactions of services without cash in hand. It is a simple approach to deliver automatically the credit to a consumer. Nowadays every credit card possess identifying for shopping transactions to take place rapidly.

Fraud: Fraud is an purposeful betrayal made for individual procure or to bruise other user

Credit Card Fraud: Acknowledged users are allowed for transactions of credit card by making use of specifications like number of credit card, card holder's address, signature, expiry date. Unauthorized use of card or using information of card without the notice of it's owner is referred as credit card fraud[2].

The detection of credit card fraud is kept private and without admitting it publicly. Common fraud detection methods which are in use usually are, rule-induction techniques, decision trees, Support Vector Machines (SVM),[16] LR, ANNs and meta- heuristics such as, k-means clustering, genetic algorithms and nearest neighbor algorithms.

Fig 1: Pictorial view of credit fraud detection

IV. DIFFERENT TECHNIQUES OF CREDIT CARD FRAUD

This section provides the overview of various credit card fraud techniques

Neural Networks:

The group of nodes which are interconnected designed to denote as human brain functioning. In adjacent layers weighted connection is established for each node with various other link nodes[15]. Single node retrieve input accepted by nodes which are linked and for result value computation the weights of the connected nodes which comprise the easy function is used. Both for supervised and unsupervised learning the Neural Networks can be formulated[4].

The user mentions the number of hidden layers along with the number of nodes within that specification. Depending upon the application the result layer of neural network consist of single or various nodes. Without the prior knowledge of potential data principles [18]Neural Networks can learn and consolidate the enclosed assumptions of data.

According to Rumelhart,[10] (1986),the topologies of Neural networks is created by arranging nodes into layers and attach layers of neurons with changed weighted interconnections. There are yet many disadvantages for the neural networks, such as

- Difficulty to confirm the structure
- Excessive training
- Efficiency of training and so on.

Bayesian networks are the one approach for fraud detection in telecommunication and credit card industry. This approach outcomes are of ensured form. Where one of the limitation of this approach is the time constraint[14]. The clean up of data is recommended as solution. As fraudsters recreate different advanced approach consistently this in turn requires adaptive system evaluation conventionally.

Decision Tree:

The idea of learning system given rise to the method decision tree and ID3 method which can deal with continuous data. The decision tree is a table of tree shape which connects the available nodes. Each node is either a branch node followed with more nodes or only one leaf node given by classification. With this imperative approach of isolating and resolving, decision tree conventionally disaffiliate[12][13] the huge problem into many simple ones and solves the sub-problems via frequently usage[20] of, data mining method to invent training different ways of classifying knowledge by constructing decision tree.

The decision tree model basis lies in how to build decision tree with high precision and small scale. High flexibility which is a non-parameter method and Good haleness on the other role are the advantages of decision tree[5].

The conception of a similarity tree using decision tree logic has been invented. a similarity tree concern to edges that are labeled with values of features and associate nodes that are labeled with attribute names which satisfies some condition and the intensity factor is left [9]which refers as the ratio of the number of transactions that satisfy these condition(s) over the total number of legitimate transaction[11] in the particular behavior .

The advantage of the similarity tree method is easy to implement, to display and to understand. Still, system has some disadvantages that, the requirements to check each transaction is sequential.

Logistic Regression:

Data mining work has huge statistical model that comprises analysis of discriminant, regression analysis, multiple- logistic regression. Logistic regression (LR) is appropriate for conditions where the appearance and non-appearance of characteristics is predicted.

It is identical to a linear regression model although the models which comprise dependent variable is bifurcated or branched[7]. Logistic regression[19] approach coordinates is used to determine disparity ratios of sovereign variables in the model which is applied for research conditions in wider range than analysis of feature .

Genetic algorithms:

This algorithm is to detect fraud algorithms acts as the medium for predictive purposes. To create logical rules which is able to of classify transactions of credit card as suspicious and non-suspicious ,the algorithm that is recommended by Bentley in which the genetic programming is the basis.

The experiment was conducted considering the database transactions which was 4000 in number with 62 fields[21].

The examination and training sample is signed for the similarity tree. To serve this purpose various types of rules were examined with the various fields. Among this best rule is highest predictability.

For credit card fraud this method serves as one of the best of home insurance data. To guess suspicious practice an algorithm was developed by Chan.

The research was initiated by valuating the cost model where the other investigation incorporated the evaluation considering the parameters such as true positive rate , error rate, prediction rate and false negative rate.

For increasing the prediction power the concept of combining different algorithms was given by Wheeler & Aitken (2000)[22] , he presented various algorithms in his article which includes: diagnostic resolution strategies, diagnostic algorithms, density selection algorithms, negative selection algorithm, best match algorithm and probabilistic curve algorithms.

From this study the conclusion was derived that says for effective classification the probabilistic and neighborhood algorithms is considered as relevant as relevant further additional advancement is carried on with respect to relevant risk measures.

The combination of genetic algorithm with Neural Networks is termed as GANN where in this genetic algorithm is used to find some parameters. The prominent question how the combination is done, this is carried by encoding the genome of genetic algorithm. The procedure in GANN include creation of number of arbitrary individuals.

Considering the genome details the design of neural network in valuating the parameter strings. Back propagation yields the performance.

The strategies of GANN entrust on genetic algorithm to determine an optimal network. Genetic Algorithm (GA) is referred as speculative search which duplicates the natural evolution process which produces useful and effective solutions to problems of optimization. (GA) exists in Evolutionary Algorithms (EA) as large class. Using techniques like mutation, inheritance, selection and crossover the solutions to problems of optimization is given.

Clustering techniques:

The two recommended clustering approach/techniques for behavioral fraud is given by Bolton & Hand (2002). the analysis of peer clusters system grant accounts which are variant during certain period whereas formerly they were operating the same.

These specified accounts are defined/flagged as suspicious. However the fraud analyst would break these cases. Breakpoint analysis is the other approach uses a different contrast rationale that uses the modified credit card usage on individual basis and subject it for investigation. Otherwise based on the single card transaction the break point analysis can determine the suspicious pattern.

Signals of suspicious form is the abrupt transaction of huge amount and high frequency of card usage without the card holder notice.

V. CONCLUSION

In present circumstance constructing a decisive, available and simple handling credit card prospect auditing system is the prominent work for bank organization in automated and tolerable way. The different techniques of credit card fraud detection along with its benefits using data mining is demonstrated in this paper.

The research on more suitable fraud detection is encouraged with the precise cost and weight factors by considering both examined and detection accuracy. Thus this paper is concerned about the financial risk of bank's to overcome by defining various classification methods to design fraud detection model with the application of data mining techniques.

REFERENCES

[1] Amlan Kundu, Suvasini Panigrahi, Shamik Sural and Arun K. Majumdar, "BLAST-SSAHA Hybridization for Credit Card Fraud Detection," IEEE Transactions On Dependable And Secure Computing, vol. 6, Issue no. 4, pp.309-315, October-December 2009S.

[2] A.Shen, R.Tong and Y. Deng, "Application of classification models on credit card fraud detection," June 2007.

[3] S. Benson Edwin Raj, A. Annie Portia "Analysis on Credit Card Fraud Detection Methods" 2011.

[4] Ghosh and D.L. Reilly, "Credit Card Fraud Detection with a Neural-Network," Proc. 27th Hawaii International Conference on System Sciences: Information Systems: Decision Support and Knowledge-Based Systems, vol. 3, pp. 621-630, 1994.

[5] Linda Delamaire (UK), Hussein Abdou (UK), John Pointon (UK), "Credit card fraud and detection techniques: a review"2009.

[6] M.F. Gadi, X. Wang, and A.P. Lago, "Comparison with parametric optimization in credit card fraud detection, 2008.

[7] Y. Sahin, E. Duman "Detecting Credit Card Fraud by ANN and Logistic Regression" 2011.

[8] A. M. Chandrashekhar and K. Raghuveer, "Confederation of FCM Clustering, ANN and SVM Techniques of Data mining to Implement Hybrid NIDS Using Corrected KDD Cup Dataset", Communication and Signal Processing (ICCSP) IEEE International Conference,2014, Page 672-676.

[9] AM. Chandrashekhar and K. Raghuveer , "Improvising Intrusion detection precision of ANN based NIDS by incorporating various data Normalization Technique – A Performance Appraisal", IJREAT International Journal of Research in Engineering & Advanced Technology, Volume 2, Issue 2, Apr-May, 2014.

[10] A. M Chandrashekhar A M and K. Raghuveer, "Diverse and Conglomerate modi-operandi for Anomaly Intrusion Detection Systems", International Journal of Computer Application (IJCA) Special Issue on "Network Security and Cryptography (NSC)", 2011.

[11] A. M Chandrashekhar A M and K. Raghuveer, "Hard Clustering Vs. Soft Clustering: A Close Contest for Attaining Supremacy in Hybrid NIDS Development", Proceedings of International Conference on Communication and Computing (ICCC - 2014), Elsevier science and Technology Publications.

[12] A. M. Chandrashekhar and K. Raghuveer, "Amalgamation of K-means clustering algorithm with standard MLP and SVM based neural networks to implement network intrusion detection system", Advanced Computing, Networking, and Informatics –

Volume 2(June 2014), Volume 28 of the series Smart Innovation, Systems and Technologies pp 273-283.

[13] A. M. Chandrashekhar and K. Raghuveer, "Fusion of Multiple Data Mining Techniques for Effective Network Intrusion Detection – A Contemporary Approach", Proceedings of Fifth International Conference on Security of Information and Networks

[14] A. M. Chandrashekhar, Jagadish Revapgol, Vinayaka Pattanashetti, "Big Data Security Issues in Networking", International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET), Volume 2, Issue 1, JAN-2016.

[15] Puneeth L Sankadal, A. M Chandrashekhar, Prashanth Chillabatte, "Network Security situation awareness system" International Journal of Advanced Research in Information and Communication Engineering(IJARICE), Volume 3, Issue 5, May 2015.

[16] P. Koushik, A.M.Chandrashekhar, Jagadeesh Takkalakaki, "Information security threats, awareness and cognizance" International Journal for Technical research in Engineering(IJTRE), Volume 2, Issue 9, May 2015.

[17] A.M.Chandrashekhar, Yadunandan Huded, H S Sachin Kumar, "Advances in Information security risk practices" International Journal of Advanced Research in data mining and Cloud computing (IJARDC), Volume 3, Issue 5, May 2015.

[18] M. Chandrashekhar, Muktha G, Anjana D, "Cyberstalking and Cyberbullying: Effects and prevention measures", Imperial Journal of Interdisciplinary Research (IJIR), Volume 2, Issue 2, JAN-2016.

[19] A.M.Chandrashekhar, Syed Tahseen Ahmed, Rahul N, "Analysis of Security Threats to Database Storage Systems" International Journal of Advanced Research in data mining and Cloud computing (IJARDC), Volume 3, Issue 5, May 2015.

[20] A.M.Chandrashekhar, K.K. Sowmyashree, RS Sheethal, "Pyramidal aggregation on Communication security" International Journal of Advanced Research in Computer Science and Applications (IJARCSA), Volume 3, Issue 5, May 2015.

[21] A.M.Chandrashekhar, Syed Tahseen Ahmed, Rahul N, "Analysis of Security Threats to Database Storage Systems" International Journal of Advanced Research in data mining and Cloud computing (IJARDC), Volume 3, Issue 5, May 2015.

[22] A.M.Chandrashekhar, K.K. Sowmyashree, RS Sheethal, "Pyramidal aggregation on Communication security" International Journal of Advanced Research in Computer Science and Applications (IJARCSA), Volume 3, Issue 5, May 2015.