PERFORMANCE ANALYSIS OF VARIOUS CREDIT CARD FRAUD DETECTION APPROACHES: A REVIEW

Suman[1], Dharminder Kumar[2]

[1] Research Scholar, Computer Science & Engineering Department, Guru Jambheshwar University of Science and Technology, Hisar-Haryana, (India)

[2] Computer Science & Engineering Department, Guru Jambheshwar University of Science and Technology, Hisar-Haryana, (India)

ABSTRACT

A lot of transactions occur in banking sector due to day to day operations. E-Commerce is widely used in busy life. In E-commerce life, credit card transactions are increasing day by day. This will increase in frauds in credit card. Credit card fraud is a major problem in financial industry. Many technologies have been developed to reduce the fraud in credit card such as data mining fuzzy logic, machine learning genetic programming, sequence alignment, markov model etc. The primary motive of this paper is to survey the approaches which are used for detecting fraud in credit card and also evaluate each methodology based on certain design criteria. The use of these techniques will help to distinguish the credit card transactions into two types as legitimate and fraudulent transactions.

Keywords: Credit Card, Data Mining, Entropy, Supervised Learning, Self -Organizing Map.

1. INTRODUCTION

Bank is a financial institution which accepts deposit from public and it becomes great disquietude for bank if occur any kind of fraud in deposit. It is mentioned in [1] that there are many kind of fraud and generally financial fraud much affects the bank fraud. Due to fastest growing online banking activities, it is found that nearly 44% of population of U.S using these online transactions. According to John T.S Quah [2] it is projected loss of $8.2 billion in the year 2006 with $3 billion in U.S alone. Data mining is a new emerging technology that can detect credit card fraud very effectively [1]. As according to them, by the help of Data Mining we can detect hidden patterns and can find out the relationship between data set. Fraud act as the wrongful/criminal deception intended to result in financial or personnel gain. So, Credit Card Fraud is an illegal or fulsome use of card or unusual transaction behavior. As shown in the Fig. 1 there are so many frauds detected that affect the bank, merchants as well as customers. Some of them are listed below:

- Inception of mails of newly issued cards.
- Copying or replicating of card information through cloned websites.
- Phishing in which credit card number and password is hacked like through emails etc.

- Triangulation In this type of fraud, fraudster make an authentic looking website and advertises to sell goods at highly lower prices. Unaware users attract to those sites and make online transactions. They submit their card information to buy those goods. And then fraudsters use this card information to make genuine transactions.

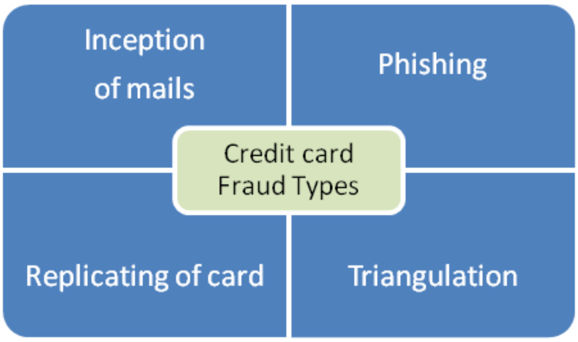

Fig. 1 Types of Frauds

II. LITERATURE SURVEY

A large number of Fraud detection Models (FDMs) have been proposed during the past years. Each model could be shown work well with a unique dataset, but no model appropriate to do well on all datasets**.** There are various models for credit cards fraud detection [3][4][5]. They are hidden markov model [6] which is based on markov model, artificial neural network models which are based upon machine learning and artificial intelligence [7][8][9][10][11], game theory models[13], peer group analysis[12], dempster shafer model [14] which is based on Bayesian learning and dempster shafer theory, computational intelligence model [2] based on self organizing map. CARDWATCH :A Neural Network Based Database Mining System for Credit Card Fraud Detection [15] which is based upon data mining approach [16] and neural network models , the Bayesian Belief Networks [17] which is based upon artificial intelligence and reasoning under uncertainty will counter frauds in credit cards and also used in intrusion detection [18]. Real-time credit card fraud detection using computational intelligence [2] and Credit card fraud detection using entropy and Bayesian network [19]. Most of the above mentioned credit card fraud detection system involve machine learning, pattern matching and artificial intelligence. Some models also used meta-heuristic techniques such as genetic algorithm; scatter search and particle swarm optimization.

This paper compares and analyzes few approaches that have been used for detecting fraud in credit card. It includes Dempster –Shafer theory and Bayesian learning, Hidden Markov model, Computational Intelligence (SOM), Fraud detection using SOM with PSO [20], Max Entropy with Bayesian Learning. Section II gives an overview of these approaches; a comparative survey presents in section III and section IV conclude and summarize the fraud detection approaches.

2.1 A fusion approach using Dempster–Shafer theory and Bayesian learning[14]

The Fraud detection system consists of four components, namely, rule-based filter, Dempster–Shafer adder, transaction history database and Bayesian learner. Rule-based filter (RBF) consists of generic as well as customer-specific rules which classify an incoming transaction as fraudulent with a certain probability. This layer can have rules R1 Address mismatch and R2 Outlier detection. The role of the DSA is to combine evidences from the rules R1 and R2 and compute an overall belief value for each transaction. Transaction history database is the transaction repository component of the proposed FDS. Bayesian learner is a tool to measure evidences supporting alternative hypotheses and arrive at optimal decisions.

2.2 Credit Card Fraud Detection using Hidden Markov Model[6]

An HMM is a double embedded stochastic process with two hierarchy levels. 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. It tries to find any anomaly in the transaction based on the spending profile of the cardholder, shipping address, and billing address, etc. If the FDS confirms the transaction to be malicious, it raises an alarm, and the issuing bank declines the transaction. The concerned cardholder may then be contacted and alerted about the possibility that the card is compromised.

2.3 Fraud detection using SOM & PSO[20]

Proposed approach is SO-PSO (Self Organized Particle Swarm Optimization) i.e. combination of SOM and PSO. This method made different clusters is based on Euclidian Distance. They define each particle by (X_1, X_2…..X_n) and weight vector by (W_1, W_2…..W_m) respectively. Then weight is updated to all nodes within a topological distance. The value of learning parameter is initially set with small random value. After applying this procedure PSO approach is applied on the updated weight of SOM. The method calculated the velocity and position of each particle under near about 2580 number of iterations.

2.4 Max Entropy with Bayesian learning[19]

In the proposed fraud detection system, the number of phases designed to analyze the deviation of each transaction from the normal behavior of the cardholder are done by using information theory. The general idea behind this theory is that outlying instances affect the information contents of the data set because of their surprising nature. Entropy values are used to estimate the score which is “further strengthened or weakened according to its similarity” with the deceitful or the legitimate transaction. “In order to meet this functionality, the proposed FDS” consists of five components which are described in the Figure 1. The first component is item analyzer where the items are analyzed according to the spending behavior of the cardholder. Standardization is the second component where domain independency is achieved. The third component is entropy reckoning where entropy is calculated for finding out the similarity in the transactions. Transaction record catalog is the fourth component where fraudulent or genuine transactions are stored. The fifth component is Bayesian classifier where Bayes’ theorem is used for checking the genuineness of fraudulent transactions.

2.5 Fraud detection using Computational Intelligence(SOM)[2]

Multilayer approach was built in this research which consists of initial authentication and screening layers, risk scoring and behavior analysis layer (core layer) and a layer of further review and decision-making. First two layers consist of verification of PIN, address and expiry date etc and the last layer for manual review of results. Main layer is core layer consist of two sub-layers a layer of SOM followed by either a feed-forward neural network or rule-based risk scoring system. Euclidean-distance and Gravity function are used as measures of similarity. The proposed approach is at early stage of a research aiming at using the properties of SOM to achieve efficient and cost-effective real-time fraud detection.

III. COMPARISON PARAMETERS

The Parameters used for comparison of various Fraud Detection Systems are Accuracy, Fraud Detection Rate in terms of True Positive and false positive, Error rate, cost and training required, supervised Learning.

Accuracy: It represents the fraction of total number of transactions (both genuine and fraudulent) that have been detected correctly.

Method: It describes the methodology used to counter the credit card fraud. The efficient methods like sequence alignment, machine learning, neural networks are used to detect and counter frauds in credit card transactions.

True Positive (TP): It represents the fraction of fraudulent transactions correctly identified as fraudulent and genuine transactions correctly identified as genuine.

False Positive (FP): It represents fraction of genuine transactions identified as fraudulent and fraudulent transactions identified as genuine. Training data: It consists of a set of training examples. The fraud detection systems are initially trained with the normal behavior of a cardholder.

Error Rate: It tells that how much the result value is actually deviated from actual value. Error Rate = 1-Accuracy

Supervised Learning: It is the machine learning task of inferring a function from supervised training data.

IV. PERFORMANCE ANALYSIS OF VARIOUS FRAUD DETECTION APPRAOCHES

To compare various credit card fraud detection approaches a comparison table was prepared. All approaches have positive and negative features. Results show that the fraud detection systems such as SOPSO, Dempster and Max Entropy have very high accuracy in terms of TP and FP. At the same time, the processing speed is fast enough to enable on-line detection of credit card fraud in case of HMM and SOPSO. Comparison performed is shown in Table 1.

Table 1: Comparison of various fraud detection approaches

Parameters	**Fusion of Dempster shafer theory and Bayesian**	**SOPSO**	**Max Entropy with Bayesian**	**HMM**	**Computation Intelligence(SOM)**
Method	Machine Learning	Optimization Techniques, Machine Learning	Entropy , Machine Learning	Hidden Markov Model	Machine Learning
TP Rate	86%	100%	97%	75%	70%
FP Rate	15%	10%	12%	20%	22%
Accuracy	85%	95%	90%	80%	75%
Error Rate	15%	5%	10%	20%	25%
Processing Speed	Medium	High	Medium	High	Medium
Training	Required	Required	Not Required	Required	Required
Cost	High Expensive	Quiet Expensive	Inexpensive	High Expensive	Inexpensive
Learning (Supervised /Unsupervised)	Supervised	Unsupervised	Supervised	Semi-supervised	Unsupervised

V. CONCLUSION

For any card issue bank, credit card fraud detection system must be efficient. A number of techniques have been proposed to detect credit card fraud. Credit card fraud detection has drawn quite a lot of interest from the research Community. The SOPSO fraud detection systems improve the system accuracy. Since the Fraud detection rate of SOPSO fraud detection systems in terms of true positive is 100% and shows good results in detecting fraudulent transactions. Approach using Max entropy and Bayesian learning also shows good accuracy in fraud detection rate and processing speed is also high.

The Fraud detection rate of SOM and Hidden Markov model is very low compare to other methods. The processing speed of Max Entropy with Bayesian is not so fast but training is not required in this approach. SOM along with other optimization techniques can be effectively used to counter frauds in other domains such as telecommunication and banking fraud detection. Approaches discussed in this survey papers have its own features in positive or negative sense. This type of survey will help us to build a hybrid approach for financial industries for detecting the fraudulent transactions in credit card.

REFERENCES

[1] Chan, P. K., Fan, W., Prodromidis, A. L., & Stolfo, S. J. (1999). Distributed data mining in credit card fraud detection. IEEE Intelligent Systems and Their Applications, 14(6), 67-74.

[2] Quah, J. T., & Sriganesh, M. (2008). Real-time credit card fraud detection using computational intelligence. Expert systems with applications, 35(4), 1721-1732.

[3] Masuda, B. (1993). Credit card fraud prevention: A successful retail strategy.Crime prevention studies, 1, 121-34.

[4] Delamaire, L., Abdou, H. A. H., & Pointon, J. (2009). Credit card fraud and detection techniques: a review. Banks and Bank systems, 4(2), 57-68.

[5] Bhatla, T. P., Prabhu, V., & Dua, A. (2003). Understanding credit card frauds.Cards business review, 1(6).

[6] Srivastava, A., Kundu, A., Sural, S., & Majumdar, A. (2008). Credit card fraud detection using hidden Markov model. IEEE Transactions on dependable and secure computing, 5(1), 37-48.

[7] Maes, S., Tuyls, K., Vanschoenwinkel, B., & Manderick, B. (2002, January). Credit card fraud detection using Bayesian and neural networks. In Proceedings of the 1st international naiso congress on neuro fuzzy technologies (pp. 261-270).

[8] Haykin, S. S. (2001). Neural networks: a comprehensive foundation. Tsinghua University Press.

[9] Brause, R., Langsdorf, T., & Hepp, M. (1999). Neural data mining for credit card fraud detection. In Tools with Artificial Intelligence, 1999. Proceedings. 11th IEEE International Conference on (pp. 103-106). IEEE.

[10] Ghosh, S., & Reilly, D. L. (1994, January). Credit card fraud detection with a neural-network. In System Sciences, 1994. Proceedings of the Twenty-Seventh Hawaii International Conference on (Vol. 3, pp. 621-630). IEEE.

[11] Chang, R. I., Lai, L. B., Su, W. D., Wang, J. C., & Kouh, J. S. (2007). Intrusion detection by backpropagation neural networks with sample-query and attribute-query. International Journal of Computational Intelligence Research, 3(1), 6-10.

[12] Weston, David J., et al. "Plastic card fraud detection using peer group analysis." Advances in Data Analysis and Classification 2.1 (2008): 45-62.

[13] Vatsa, V., Sural, S., & Majumdar, A. K. (2005, December). A game-theoretic approach to credit card fraud detection. In International Conference on Information Systems Security (pp. 263-276). Springer Berlin Heidelberg.

[14] Panigrahi, S., Kundu, A., Sural, S., & Majumdar, A. K. (2009). Credit card fraud detection: A fusion approach using Dempster–Shafer theory and Bayesian learning. Information Fusion, 10(4), 354-363.

[15] Aleskerov, E., Freisleben, B., & Rao, B. (1997, March). Cardwatch: A neural network based database mining system for credit card fraud detection. In Computational Intelligence for Financial Engineering (CIFEr), 1997., Proceedings of the IEEE/IAFE 1997 (pp. 220-226). IEEE.

[16] Phua, C., Lee, V., Smith, K., & Gayler, R. (2010). A comprehensive survey of data mining-based fraud detection research. arXiv preprint arXiv:1009.6119.

[17] Ezawa, K. J., & Norton, S. W. (1996). Constructing Bayesian networks to predict uncollectible telecommunications accounts. IEEE Expert: Intelligent Systems and Their Applications, 11(5), 45-51.

[18] Ezawa, K. J., & Norton, S. W. (1996). Constructing Bayesian networks to predict uncollectible telecommunications accounts. IEEE Expert: Intelligent Systems and Their Applications, 11(5), 45-51.

[19] Kumar, D., & Arora, S. (2016). A Hybrid Approach Using Maximum Entropy and Bayesian Learning for Detecting Delinquency in Financial Industry. International Journal of Knowledge-Based Organizations (IJKBO), 6(1), 60-73.

[20] Arora S., & Kumar,D. Hybridization of SOM and PSO for detecting fraud in credit card. International Journal of Information Systems in the Service Sector (IJISSS) (accepted).