A Review on Credit Card Fraud Detection Techniques

Ranjeeta Jha*, Abhaya*, Vijay Kumar Jha#
*M.Tech (IS), Dept. of Information Technology, Birla Institute of Technology, Mesra (Ranchi), India
#Associate Professor, Dept. of Information Technology, Birla Institute of Technology, Mesra (Ranchi), India

Abstrac: - In today's world, the Internet is a part of our life. Due to the extensive use of internet, the popularity of online shopping is growing day by day. Credit Card is the simplest method to do online shopping and paying bills. Thus Credit Card become very popular and convenient mode for online money transaction and is increasing very rapidly. With the increase of Credit Card usage, the opportunities for fraudster to steal credit card details and subsequently commit fraud are also increasing. Credit Card fraud is the fraud committed by the use of another person's credit card. To support safe credit card usage an efficient fraud detection system is essential. Presently, many modern techniques based on Artificial Intelligence, Sequence Alignment, Data Mining, Fuzzy Logic, Machine Learning, Genetic Programming etc. has been introduced for detecting various credit card fraudulent transactions. This paper presents a survey of various current techniques used in fraud detection mechanism and provides a comprehensive review of different techniques based on certain design criteria.

Keywords: Credit Card, fraud detection, online shopping.

1. INRODUCTION

A Credit Card is a payment card which allows the cardholder to pay for goods and services. Credit Card is a sort of loan with duration 40 to 50 days in the form of card which can carry with us [1]. It is a convenient substitute for cash or check and an essential component of e-commerce and internet commerce. According to Nielsen study conducted in 2008, 28% of the world's population has been using internet [2]. Among these, 85% of population has used internet to make online shopping and the rate of making online shopping has increased by 40% from 2005-2008 and also mentioned that the Credit Cards are the most popular mode of online payment [3]. According to Experian National Score Index Study conducted in 2007, approximately 51 percent of the U.S population had at least two Credit Cards and approximately 14 percent of the U.S population had more than 10 Credit Cards [4]. If we consider the statistics of Credit Cards in India, it is found that total number of Credit Cards in India at the end of December-31-2012 is about 18 to 18.9 million[5][6][7]. As the number of credit card users rises world-wide, subsequently fraudsters are also finding more opportunities to commit fraud. Credit Card fraud happens when someone gains access to an individual's legitimately opened Credit Card account and uses it to buy items, take out cash advances and create other illegal schemes. Credit Card fraud costs Credit Card companies million of dollars per year.

2. CREDIT CARD FRAUD

The PWC global economic crime survey of 2011 suggests that 34% of companies worldwide have reported being victim of fraud in the past year and increasing from 30% as reported in the year 2009 [8]. The survey report based on a Q3 2012 ACI Worldwide, in 2012, 14% of debit and credit card holders having experienced fraud multiple times during the past five years, this is compared to 6% in 2011, a statistically significant increase [9].

According to survey report, Credit Card fraud rates in different countries are shown below.

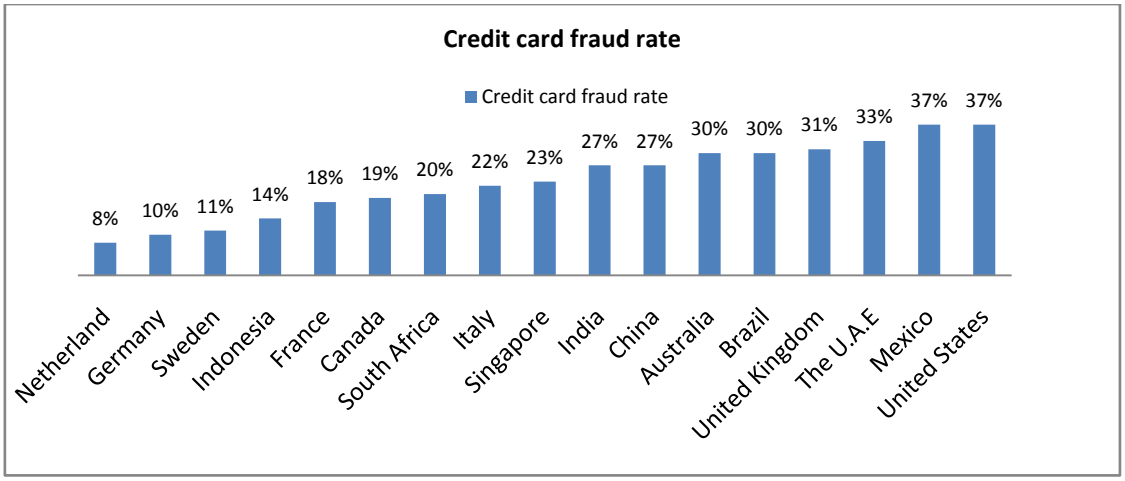

Figure1. ACI Worldwide study of 5,223 consumers in 17 countries, Q3 2012.

Credit Card is the most acceptable payment mode for online and offline transaction. It can be divided into two types (a) Physical Card (b) Virtual Card

In the physical card based purchase, cardholder presents his card physically for making a payment. It is offline mode transaction. In this method of purchase fraud can be done with help of lost cards, stolen cards or using fake cards.

In virtual card based purchase, only card details are given through online or over phone to make the payment. In this method, hacker simply needs to know the card details to commit fraud.

2.1 Types Of Credit Card Fraud

There are number of ways through which fraudster can execute a credit card fraud either it can be done through offline mode or it can be through online mode. Offline fraud is committed by using a stolen physical card at call centre or any other place. Online fraud is committed via internet, phone, shopping, and web or in absence of card holder. The different types of methods for committing Credit Card frauds are described below [5] [10]:

2 Application fraud: When someone applies for a credit card with false information that 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 describes application fraud as "demonstration of identity crime, occurs when application forms contain possible, and synthetic (identity fraud), or real but also stolen identity information (identity theft) [11].

3 Lost/Stolen Cards: Fraud occurs when the legitimate account holder loses the card or someone steals the card for criminal purposes. A lost or stolen credit card has the potential to cause a lot of damage in case of high credit limit. In 2001, thieves stole £114m in the UK through the use of lost and stolen credit cards. Most fraud on lost and stolen credit cards will take place at commercial outlets and telephone shops prior to the genuine card holder reporting its loss.

4 Account takeover: This type of fraud occurs when a fraudster illegally obtains a valid customers' personal information. The fraudster takes control of (takeover) a legitimate account by either providing the customers' account number or the card number.

5 Fake and Counterfeit cards: A counterfeit card is one that has been scanned, printed, recorded or swiped without the card issuer's permission. Counterfeiting in the UK rose104% in 2000 to £102.8m and then a further 64% in 2001 up to £160.3m. Some of the techniques used for creating false and counterfeit cards are given below.

5.1 Creating a fake card: A fraudster can create a fake card from scratch using sophisticated machines. This is the most common type of fraud but now modern cards have many security features designed to make it difficult for fraudster to make a fake card.

5.2 Altering card details: A fraudster can alter cards by either re-embossing them by applying heat and pressure to the information originally embossed on the card by a legitimate card manufacturer or by re-encoding them using computer software that encodes the magnetic stripe data on the card [9][11].

5.3 Skimming: Most cases of counterfeit fraud involve skimming, a process where genuine data on a card's magnetic stripe is electronically copied onto another. Skimming is fast emerging as the most popular form of credit card fraud.

3. CREDIT CARD FRAUD DETECTION

A lot of research has been carried out for the detection of credit card fraud. The results obtained by different researchers are summarized below:

Aleskerov et al. [12] present CARDWATCH, a Neural Network based database mining system used for credit card fraud detection. The system has an interface to a variety of commercial databases and a graphical user interface. Ghosh and Reilly [13] compared the performance of neural network based fraud detection system with rule based fraud detection procedures using a dataset with all kinds of fraud: lost cards, stolen cards, application fraud, counterfeit fraud, mail-order fraud, and NRI (non-received issue) fraud.

Dorronsoro et al. [14] describe the performance of an on-line system for credit card fraud detection based on a neural classifier. This system's main focus is to imbed itself deep in credit card transaction servers to detect fraud in real-time.

Chan et al. [15] discuss the three important issues concerned with credit card fraud detection: skewed distribution of credit card transaction datasets, non-uniform cost per error of classification, and the speed of fraud detection. Syeda et al. [16] developed parallel granular neural networks (GNN) to speed up data mining and knowledge discovery process for fast credit card fraud detection. Abhinav srivastava et al. [17] present Hidden Markov Model (HMM) technique for credit card fraud detection. An HMM is initially trained with normal behavior of a card holder. This model shows 80% accuracy over a wide variation in the input data. Phua et al [18] have done an extensive survey of existing data-mining-based FDSs.

Chen et al. [19] proposes binary support vector system (BSVS) for increasing credit card fraud detection rate. This system is effective in predicting a high true negative rate of fraud.

Amalan Kundu et al [20] proposed a model called BLAST-SSAHA Hybridization technique for online credit card fraud detection. BLAST-SSAHA approach improves the fraud detection by combining both anomalies as well as misuse detection techniques.

Bolton and Hand [21] proposed an unsupervised credit card detection method by observing abnormal spending behavior and frequency of transactions. Stolfo et al. [22] presented a credit card fraud detection system using various meta-learning techniques to learn models of fraudulent credit card transactions.

Sherly K.K. and R Nedunchezhian [1] proposed BOAT adaptive credit card fraud detection system. In this work BOAT supports incremental update of transactional database and it handles maximum fraud coverage with high speed and less cost. Elkan et al. suggest Naïve Bayesian approach for credit card fraud detection to achieve high fraud detection along with low false alarm.

R. Wheeler, S. Aitken [23] used multiple algorithms for fraud detection which is an application of case based reasoning. The approach was towards the problem of reducing the number of final-line fraud investigations in the credit approval process.

Jianyun et al. [24] suggested a framework for detecting fraudulent transactions in an online system. That paper describes an FP tree based method to dynamically create user profile for the purpose of fraud detection. But this technique doesn't consider unusual patterns i.e. short term behavioral changes of genuine card holders.

Wen-Fang et al. [25] have proposed a research on credit card fraud detection model which is based on outlier detection mining on distance sum, which shows that it can detect credit card fraud better than anomaly detection based on clustering. Benson Edwin et al [26] have done a survey on different credit card fraud detection methods. Peter J. Bentley et al [27] describes the use of an evolutionary-fuzzy system for detection of credit card fraud.

4. ALGORITHM USED IN CREDIT CARD FRAUD DETECTION

4.1 A fusion approach using Dempster-Shafer theory and Bayesian learning

Dempster-Shafer theory and Bayesian learning is a hybrid approach for credit card fraud detection [20] [26][30] which combines evidences from current as well as past behavior. Every cardholder has a certain type of shopping behavior which establishes a spending profile for them. This approach proposes a fraud detection system using information fusion and Bayesian learning for detecting credit card fraud.

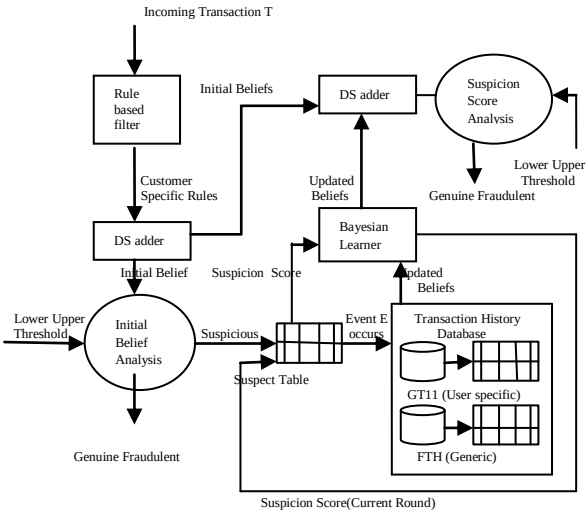

Figure 2. Block diagram of the proposed fraud detection system

The fraud detection system consists of four components, namely rule-based filter, Dempster-Shafer adder, transaction history database and Bayesian learner. In the rule-based component, the suspicion level of each incoming transaction is determined. The role of the Dempster-Shafer adder is to combine evidences from the rule-based filter and

compute an overall belief value for each transaction [31]. The transaction is classified as suspicious or normal depending on this initial belief. Once a transaction is found to be suspicious, belief is further strengthened or weekend according to its similarity with fraudulent or genuine transaction history using Bayesian learning. It has high accuracy and high processing speed. It improves detection rate and reduces false alarms and also it is applicable in e-commerce.

4.2 *Hidden Markov Model*

An HMM is a double embedded random process with two hierarchy levels in which one is hidden and other
is open to all.. It is only the result, not the state visible to an external observer and therefore states are "hidden" to the outside, hence the name Hidden Markov Model. The Hidden Markov Model is a finite set of states, each of which is associated with a probability distribution. It is the simplest and easiest models which can be used to model sequential data i.e. data samples which are dependent from each other.

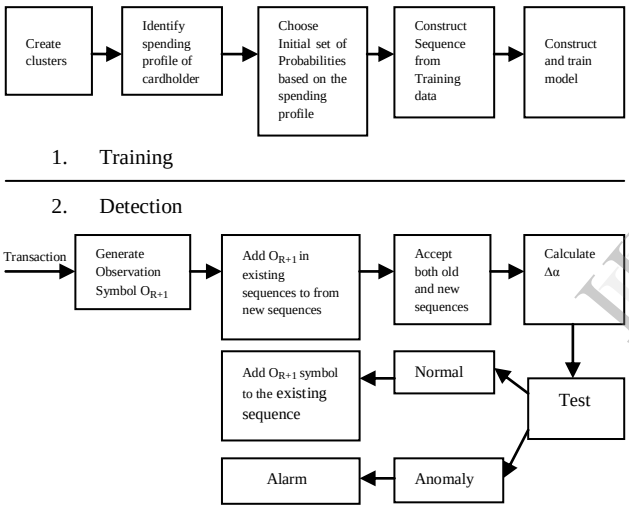

Figure 3. Process Flow of the proposed fraud detection system

A Hidden Markov Model is initially trained with the normal behavior of a cardholder [17]. After this each incoming transaction is submitted to the fraud detection system for verification purpose. Baum-Welch algorithm and K-Mean clustering algorithms are used for training purposes in HMM. The fraud detection system accepts card details and the valued purchase to verify whether the transaction is fraudulent or genuine [26]. It creates the cluster of training sets and identifies the spending profile of cardholder which is used to find out any variance in the transaction. If the FDS detects any variance or deviation in transaction from the normal transaction, it raises an alarm and the issuing bank block the account for further transaction. Since HMM produces a log for transaction it reduces tedious work of employee but produces high false alarm as well as high false positive [28].

4.3 Random Forest

Decision tree classifiers are popular in terms of their ease of use, flexibility in terms of handling various data attribute types and also able to generate understandable rules. Single tree models, however can be unstable and also speed and scalability limitations may occur in case of specific training dataset. Ensemble methods seek to address this problem by developing a set of models and aggregating their predictions in determining the class labels for a data point. The main principle behind ensemble methods is that a group of "weak learners" can come together to form a "strong learner". Random Forest is an ensemble of unprunned classification and regression trees, developed by Breiman [32].This algorithm is an implementation of bootstrap aggregation (bagging) where each tree in an ensemble of decision tree is constructed from a bootstrap sample feature vectors from the training data. Each bootstrap sample of feature vectors is obtained by repeated random sampling with replacement until the size of the bootstrap sample matches the size of the original training subset. When constructing each decision tree, only randomly selected subsets of features are considered for constructing each decision tree. This random selection of features helps Random Forest to not only scale well when there exists many features per feature vector, but also helps it in reducing the interdependence between the feature attributes and is thus less vulnerable to inherent noise in the data. In random forest training is fast even for large datasets. Due to this reason it exhibits a substantial performance improvement over the single tree classifiers such as CART and C4.5.

4.4 Fuzzy-Darwinian Detection System

Fuzzy Darwinian Detection System [26][27] describes the use of an evolutionary-fuzzy system capable of classifying suspicious and non-suspicious credit card transactions. The system comprises of a genetic programming search algorithm and a fuzzy expert system. It uses genetic programming to evolve fuzzy logic rules capable of classifying Credit Card transactions into "suspicious" and "non-suspicious" classes.

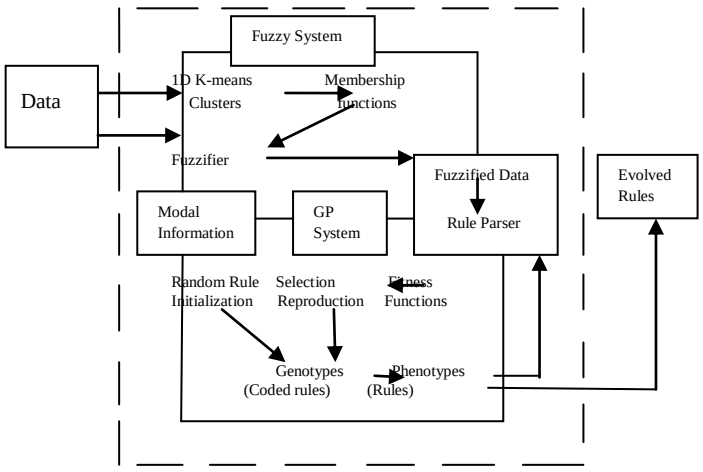

Figure 4. Block diagram of the Evolutionary-fuzzy system

Data is provided is to FDS in the form of two separate files: training data and test data. The system first clusters the training data into three groups namely low, medium and

high. The minimum and maximum values in each cluster are then designed to define the domains of the membership functions of the fuzzy expert system [27]. After the selection of membership function, GP engine is then seeded with random genotypes (coded rules) and evolution is initiated. At the start of evolution, random genotypes are created. Genotypes are mapped onto phenotypes to obtain fuzzy rules. The evolved rules are then used to detect the suspicious data items in the training dataset. All items that are correctly classified by this rule are removed and the FDS automatically restarts, evolving a new rule to classify the remaining items. This process continues until every "suspicious" data item has been described by a rule. This approach has very high accuracy and produces a low false alarm. But it is not applicable in online transaction and it is highly expensive. Also the processing speed of the system is low.

5. CONCLUSION

In this paper, on the basis of accuracy, efficiency, security, processing speed, cost and high detection rate, the paper has analyzed above mention techniques of the credit card fraud detection algorithm section. According to important and given necessary parameter mention above, Fuzzy Darwinian and Dempster- Shafer have very high accuracy in terms of TP and FP while processing speed of HMM is very high. Also Random Forest is improvement over other existing decision tree algorithm in terms of preprocessing and testing phase as it can train large datasets very fast.

6. REFERENCES

1. Sherly K.K and R Nedunchezhian, "Boat adaptive credit card fraud detection system", International Conference on Computational Intelligence and Computing Research(IEEE), 28-29 Dec. 2010.
2. Internet usage world statistics, (http://www.internetworldstats.com/stats.htm), 2011.
3. Trends in online shopping, a Global Nelson Consumer Report, (2008).
4. http://www.cardhub.com/edu/number-of-credit- Cards/
5. Avinash Ingole and Dr. R. C. Thool, "Credit Card Fraud Detection Using Hidden Markov Model and Its Performance", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 6, June 2013.
6. The Economic Times http://articles.economictimes.indiatimes.com/2012-08-10/news/33137593_1_number-of-active-credit-credit-card-cibil
7. http://www.cardbhai.com/credit-debit-carddata/india-credit-card-holders-information-bank wise-2012
8. John Akhilomen, "Data Mining Application for Cyber Credit-card Fraud Detection System", Proceedings of the World Congress on Engineering, Vol 3, July 3 - 5, 2013.
9. http://www.aciworldwide.com/~/media/files/collateral /aci_aite_global_consumers_react_to_rising_fraud_1012
10. Tej Paul Bhatla, Vikram Prabhu & Amit Dua,"Understanding Credit Card Frauds", Tata Consultants Services, 2003.
11. Khyati Chaudhary and Jyoti Yadav Bhawna Mallick,"A review of Fraud Detection Techniques: Credit Card" , *International Journal of Computer Applications (0975 – 8887) Volume 45– No.1, May 2012.*
12. Aleskerov E., Freisleben B. and Rao B., "CARDWATCH: A Neural Network Based Database Mining System 4for Credit Card Fraud Detection", Proc. IEEE/IAFE: Computational Intelligence for Financial Eng., pp.:220-226, 1997.
13. Ghosh S. and Reilly D.L., "Credit Card Fraud Detection with a Neural- Network", Proceedings of the International Conference on System Science, pp.621-630, 1994.
14. Dorronsoro J.R., Francisco G., Carmen S. and Carlos S.C., "Neural Fraud Detection in Credit Card Operation." IEEE Transaction on Neural Network, vol.-08, no.-04, pp.: 827-834, 1997.
15. Philip K. Chan, Wei Fan, Andreas Prodromidis and Salvotore J. Stalfo, "Distributed data mining in credit card fraud detection", IEEE Intelligent systems, November/December 1999, pp. 67 – 74.
16. M. Syeda, Y.Q. Zhang and Y. Pan, "Parallel Granular Neural Networks for fast credit card fraud detection", Proceedings of the IEEE International Conference on Fuzzy Systems, pp. 572-577, 2002.
17. Abhinav Srivastava, Amlan Kundu, Shamik Sural and Arun K. Majumdar, "Credit card fraud detection using Hidden Markov Model", IEEE Trans. dependable and secure computing, Vol. 5, No. 1, January-March 2008.
18. C. Phua, V. Lee, K. Smith and R. Gayler, "A comprehensive survey of data mining based fraud detection research", http:// www.bsys.monash. edu.au/people/cphua/, Mar. 2007.
19. Chen, R. C., Chen, T. S. and Lin, C. C. A, "New Binary Support Vector System for Increasing Detection Rate of Credit Card Fraud", International Journal of Pattern Recognition, 20(2), 227-239, 2006.
20. 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 2009.
21. Bolton R. and Hand D., "Unsupervised Profiling Methods for Fraud Detection", Credit Scoring and Credit Control VII, 2001.
22. S.J Stolfo, D.W Fan, W.Lee, A.L Prodromidis and P.K.Chan, "Credit Card Fraud Detection Using Meta-Learning: Issues and Initial Results", Proceedings of AAAI Workshop AI Methods in Fraud and Risk Management, pp. 83-90, 1997.
23. Wheeler R. and Aitken S., "Multiple algorithms for fraud detection", Knowledge-Based Systems, no. 13 pp.: 93-99, 2000.
24. Jianyun Xu, Andrew H. Sung and Oingzhong Liu,"Behavior mining for fraud detection", Journal of research and practice in Information Technology, Vol. 39, no. 1, February 2007.
25. Wen-Fang YU and Na Wang, "Research on credit card fraud detection model based on Distance Sum", International joint conference on Artificial Intelligence, pp.353-356, 2009.
26. S. Benson Edwin Raj and A. Annie Portia, "Analysis on Credit Card Fraud Detection Methods", International Conference on Computer, Communication and Electrical Technology – ICCCET2011, 18th & 19th March, 2011.
27. Peter J. Bentley, Jungwon Kim, Gil-Ho Jung and Jong-Uk Choi, "Fuzzy Darwinian Detection of Credit Card Fraud", *In the 14th Annual Fall Symposium of the Korean Information Processing Society*, 14th October 2000.
28. Masoumeh Zareapoor, Seeja.K.R, and M.Afshar.Alam, "Analysis of Credit Card Fraud Detection Techniques: based on Certain Design Criteria", International Journal of Computer Applications (0975 – 8887), Volume 52– No.3, August 2012.
29. Krishna Kumar Tripathi and Mahesh A. Pavaskar, "Survey on Credit Card Fraud Detection Methods", International Journal of Emerging Technology and Advanced Engineering (ISSN 2250-2459),Volume 2, Issue 11, November 2012.
30. Krishna Kumar Tripathi and Lata Ragha, "Hybrid Approach for Credit Card Fraud Detection", International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-3, Issue-4, September 2013.
31. L. Breiman. "Random Forests". *Machine Learning*, 45 (1): 532.doi:10.1023/A:1010933404324, 2001.
32. AshishGupta, Jagdish Raikwal, "Fraud Detection in credit Card Transaction Using Hybrid Model", International Journal Of Engineering And Computer Science ISSN:2319-7242 Volume 3 Issue 1 Jan, 2014 Page No. 3730-3735.
33. Sherly K.K," A Comparative Assessment of Supervised Data Mining Techniques for Fraud Prevention", TIST.Int.J.Sci.Tech.Res., Vol.1 (2012), 1-6
34. MohdAvesh Zubair Khan, Jabir Daud Pathan, 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