International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue III, March 2018- Available at www.ijraset.com

Design and Analysis of Hybrid Algorithm for Credit Card Fraud Detection Using GA and HMM

Ms. Durga Wanjari[1], Prof. Rajesh Nasare[2], Prof. Abhijit Pande[3]
[1, 2, 3,] *Computer-Science & Engineering, Rajiv Gandhi College of Engineering, Nagpur*

Abstract: ***The evolution of the new technology supports the online transactions to be held with the assistance of different payment cards. Credit card frauds have become increasingly rampant in living years and critical for banks to enhance fraud detection so as to protect their cardholders from financial loss. In Credit card fraud detection system, fraudulent transaction will be detected after transaction is done. In this paper, we used hybridization of Hidden Markov Model and Genetic Algorithm. HMM and GA is used to finding credit card frauds in system. Credit card fraud can be detected using Hidden Markov Model and Genetic Algorithm during transactions. Hidden Markov Model is the statistical tools it solve "hidden" problems and it is used for spreading patterns of transactions and concurrently report the time stamp and IP address of the intruder machine and genetic algorithm is used for to finding better solutions to technically eliminate the fraud in short time.***
Keywords : Credit card Fraud Detection, Genetic Algorithm, Hidden Markov model, Spending patterns.

I. INTRODUCTION

The main goal is to develop efficient and secure electronic payment system to detect whether a transaction is fraudulent or not. Credit card fraud arises when one person uses other persons' card for their personal use without the knowledge of the card holder. These transactions are supported by different bank cards which makes the operation easy. A huge population use credit card for its undemanding accessibility. The bank has accumulated a vast count of credit card transactions. Thus the Genetic Algorithm will cause a better solution to such problems. The importance is given on developing efficient and secure electronic payment system for detecting the fraudulent transactions.

Credit card can be used to purchases goods and services using online and offline transaction mode. It can be divided into two types: 1) physical card and 2) virtual card. In virtual or online use, the card owner uses the credit card to pay for purchased items online over the internet by just entering the details of credit card. Only some important information about a card (Credit card number, name, expiry date of credit card (month and year), CVV number, pin code, etc.) is required to make the payment. In online credit card fraud a fraudster simply needs to know the all details of credit card. As we know that a credit card is a small and handy plastic card that is issued by the bank that contains the unique identification such as signature that authorized the person to purchase goods and services on credit and the charges for which will be billed periodically. The information that is stored on the credit card can be read by the Automatic Teller Machines (ATM's), store readers, banks and also used in online Internet Banking System.

II. LITERATURE REVIEW

Ruche Berol has proposed the problem and also generates the result of the fraudulent transactions. Genetic algorithm is used to make decision about the network topology, number of hidden layers, and number of nodes that can be used in designing the neural network for solving our credit card fraud detection problem. The Genetic algorithms are evolutionary algorithms in which the aim is to obtain the better and optimal solutions. In the fraud transactions can be predicted soon after credit card transactions by the banks.

Ayesha Gupta, Dhananjay Kumar, Arul Brave has proposed the HMM, is use the normal behavior of a cardholder needs an enough number of normal transactions and fraud transactions for learning fraud patterns. To make it more effective we have enclosed the provision of determining the IP address of intruder machine along with its time stamp. HMM uses cardholders spending behavior to detect fraud. High spending behavior, Medium spending behavior, Low spending behavior. Different users have their various spending behavior (high, medium, low). High spending behavior of any user shows that card holder spending amount(H), medium spending behavior of any user shows that user spend medium amount(M), low spending behavior of any user shows that cardholder spend low amount(L).

Raghavendra Patidar and L. Sharma has proposed fraud detection using Neural Network is totally based on the human brain working principle. There is a fix pattern of credit card use, made by the way consumer uses a credit card. The Neural Network based fraud detection system check for the pattern used by the fraudster and matches with the pattern of the valid card holder on which Neural network has been trained.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 6.887
Volume 6 Issue III, March 2018- Available at www.ijraset.com

III. SYSTEM FLOW DIAGRAM

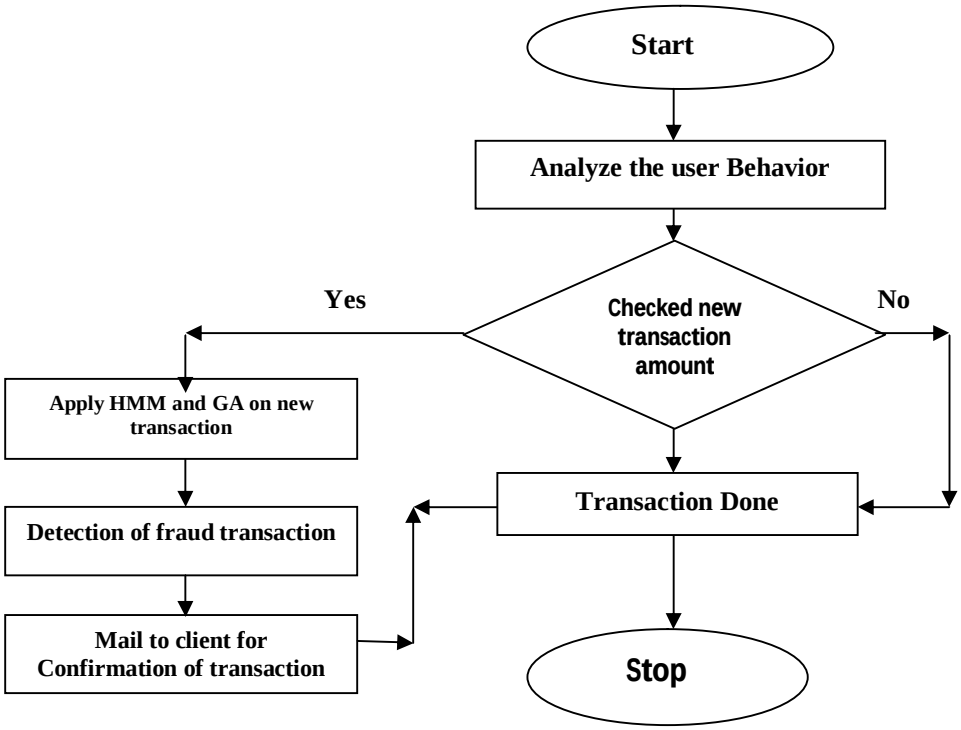


Fig. Flow diagram of credit card fraud detection System.

IV. PROPOSED SYSTEM DESIGN

A. Implementation on Credit Card Fraud Detection Using Hidden Markov Model

A Hidden Markov Model is a finite set of states; every state is associated with a probability distribution. Transitions among these states are administered by a set of probabilities called transition probability. In hidden Markov model only output is clearly visible but states are not visible means all states are hidden. In this model firstly, all the transaction sequence need for deciding the category. Into three clusters namely,

1) Low Categor
2) Medium Category
3) High Category

according to the user credit card limit. After deciding the categories the fraud detection of incoming transaction will be verified by last 10 transactions and finding credit card fraud in a system. A Hidden Markov Model is checking the normal behavior of a card user.

The HMM can be well defined with the following elements-

N number of states that are hidden denoted by a set S = {S1, S2, S3, …. SN}, where i =1, 2, …… , N, is count of state and Si, is an individual state.

M denotes the total number of observation symbols. When observations are continuous then M is infinite. We denote the set of symbols V = {V1, V2,….VM} where Vi, is an individual symbol.

A set containing probability of moving from one state to another, defined as transition probability.

$$a_{ij} = P\{q_{t+1} = S_j \mid q_t = S_i\},\ 1 \le i, j \le N$$

where qt denotes the present state. Transition probabilities should satisfy two constraints

$$a_{ij} \ge 0,\ 1 \le i, j \le N$$

and

$$\text{summation} = 1,\ 1 \le i \le N$$

Matrix B, indicating observation symbol probability

$$B = \{b_j(k)\}$$

A probability distribution in each of the states is given as,

bj(k) = P{at =Vk| qt = Sj},1 ≤ j ≤ N, 1 ≤ k ≤ M

where, Vk denotes the kth observation symbol and at the current parameter vector. The given equation should satisfy some constraints

bj(k) ≥ 0, 1 ≤ j ≤ N, 1 ≤ k ≤ M

and

summation bj(k) =1, 1 ≤ j ≤ N

The initial state probability given by

Π = {Πi}

where,

Πi = P {qi = Si}, 1 ≤ i ≤ N

Summation Πi= 1.

HMM algorithm gives the threshold value for detection of fraud transaction. This threshold value is calculated using the transition probability of transaction of particular user.

HMM is implemented as following steps:

Hidden states $Q = \{ q_i \}, i = 1, \ldots, N$.

Transition probabilities $A = \{a_{ij} = P(q_j \text{ at } t+1 \mid q_i \text{ at } t)\}$, where $P(a \mid b)$ is the conditional probability of a given b, $t = 1, \ldots, T$ is time, and q_i in Q. Informally, A is the probability that the next state is q_j given that the current state is q_i.

Observations (symbols) $O = \{ o_k \}, k = 1, \ldots, M$.

Emission probabilities $B = \{ b_{ik} = b_i(o_k) = P(o_k \mid q_i) \}$, where o_k in O. Informally, B is the probability that the output is o_k given that the current state is q_i.

Initial state probabilities $\Pi = \{p_i = P(q_i \text{ at } t = 1)\}$.

B. Implementation On Credit Card Fraud Detection Using Genetic Algorithm

Genetic algorithm ware invented to mimic some of processes observed in natural evaluation. Many people, biologist included, are astonished that life at the level of complexity that we observe could have evolved in the relatively short time suggested by the fossil record. The idea with GA is to use this power of evaluation problem. During a credit card transaction, the fraud has to be deducted in real time and the number of false alerts are being minimized by using genetic algorithm. The fraud that is detected is based on the user's transaction in a year.

There are three Genetic operators that are used in detecting the fraud transaction :

1) Selection of population: first we need to select the particular user's transaction in one year as the population for performing the other steps of algorithm

2) Crossover: After the selection of population i.e transaction of user in a particular year crossover is calculated. Crossover means the unique transaction in a particular year. Unique value is selected from the whole population

3) Mutation: Mutation is the probability of unique transaction of particular user. If the mutation probability is low then the fitness value is calculated from that unique transaction probabilities, if not then we need to go for the removal of the some of the transaction then the probability comes low. If probability of mutation is 0.1 to 0.4 then the fitness value can be calculated and probability greater than the 0.4 then we need to do the removal of some transaction.Mutation probability (mp) mp= Number of unique transaction / Number of total transaction in a particular yea

4) Fitness value: when we get the desired lower mutation probability we can able to calculate the fitness value from it by using a formula for it.Fitness value (fv) fv= Sum of all the unique transaction present for mutation / number of unique transaction.

This fitness value is used as the threshold value for the fraud transactions. For getting the more appropriate fraud transaction we uses the hybrid method this method uses both the HMM and Genetic algorithm. In this, method we select the population with the help of HMM algorithm and also the mutation probability is calculated using the transition probability calculation of the HMM algorithm and after it we get two threshold values of transaction of a particular user then we find the average of these values to get the precise result of transactions.

GA is implemented as following steps:

selection of population through the transaction.

Mutation probability

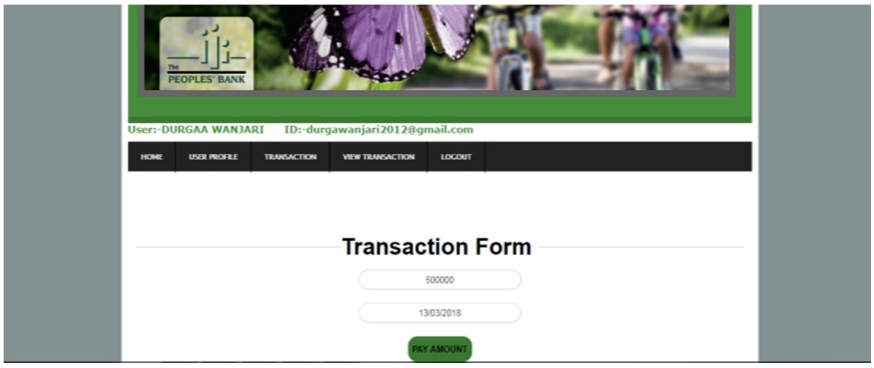
mp= Number of unique transaction / Number of total transaction in a particular year Fitness formula

fv= Sum of all the unique transaction present for mutation / number of unique transaction

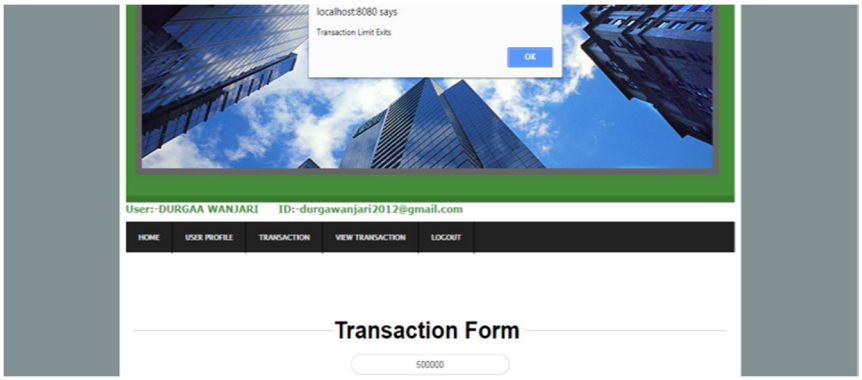
V. RESULT ANALYSIS

During client enters the amount which is observation symbol whether it is transaction amount is calculated as per combination on threshold value 1 and threshold value 2 according to algorithms. New transaction amount is below final threshold value then that amount is added in the sequence and both the sequence is accepted means transaction is done. And new threshold value is calculated and amount is tested if it is above threshold value then the fraud is generated and send mail to the real user and then user open that link user get two options i.e. yes or no. User click on yes button then transaction is done and user click no button then transaction will be stop.

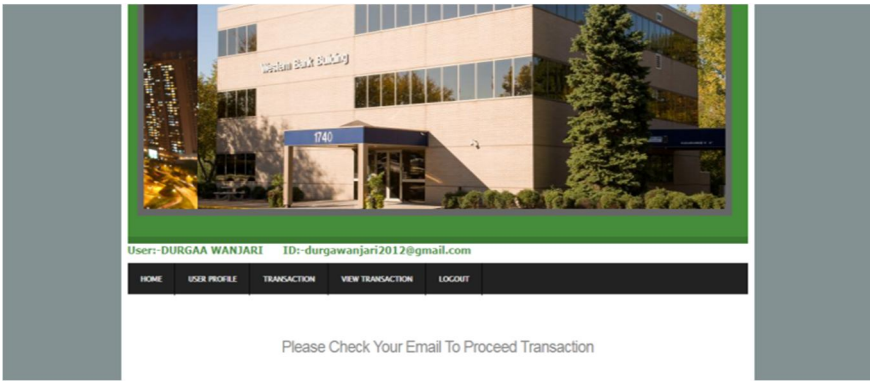
A. User Transaction Form



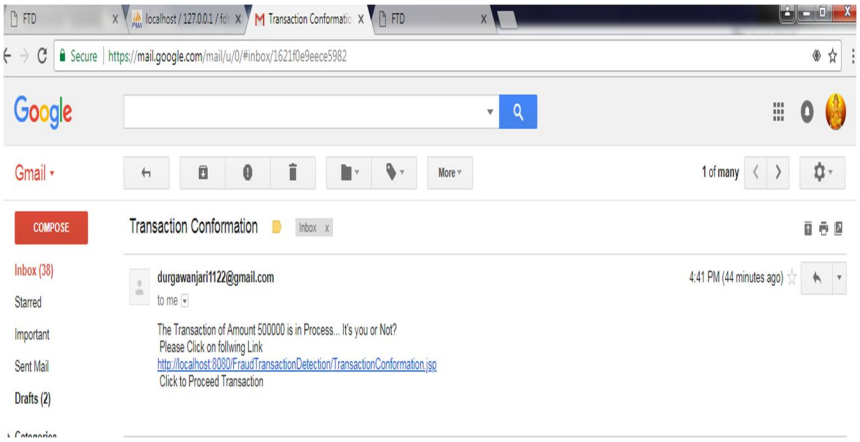
B. Transaction Pattern Exist



C. System Alert Massage

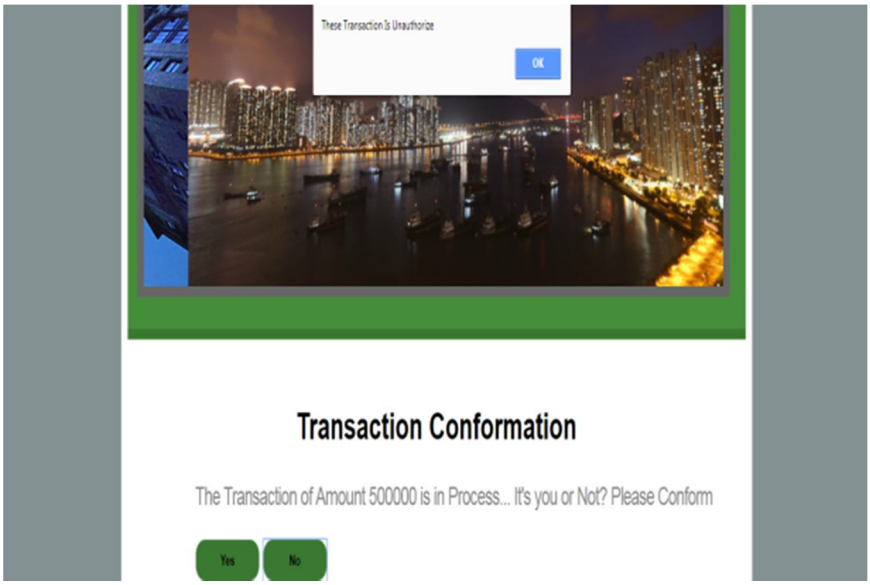


D. System Send a Confirmation Mail To The Real User



E. Check Whether Transaction is your or Not

F. Fraud Is Detected



G. Trace IP Address And Time Stamp

	t_id	Reg_Id	Ip_Address	Transaction_Date	Transaction_Time
Edit Copy Delete	1	FDT_3	0:0:0:0:0:0:0:1	19/02/2018	17:12:53
Edit Copy Delete	2	FDT_2	0:0:0:0:0:0:0:1	21/02/2018	10:56:00
Edit Copy Delete	3	FDT_4	0:0:0:0:0:0:0:1	23/02/2018	22:19:30
Edit Copy Delete	4	FDT_4	0:0:0:0:0:0:0:1	26/02/2018	12:42:31
Edit Copy Delete	5	FDT_3	0:0:0:0:0:0:0:1	27/02/2018	13:21:10
Edit Copy Delete	6	FDT_4	0:0:0:0:0:0:0:1	27/02/2018	14:27:22
Edit Copy Delete	7	FDT_4	0:0:0:0:0:0:0:1	27/02/2018	20:44:05
Edit Copy Delete	8	FDT_4	0:0:0:0:0:0:0:1	28/02/2018	13:00:20
Edit Copy Delete	9	FDT_6	0:0:0:0:0:0:0:1	28/02/2018	15:17:58
Edit Copy Delete	10	FDT_6	0:0:0:0:0:0:0:1	28/02/2018	22:41:16
Edit Copy Delete	11	FDT_9	0:0:0:0:0:0:0:1	05/03/2018	11:54:35
Edit Copy Delete	12	FDT_9	0:0:0:0:0:0:0:1	08/03/2018	12:59:25

VI. PROBABLE OUTPUT

It is very difficult to do simulation on real time data set which is not providing from any credit card bank on security reasons. In Table 1, it is shown that a random data set of all transactions happened is categorized according to their types of purchase. With the help of this, we calculate probability of each spending profile (h, l and m) of every category. Fraud detection of incoming transaction will be checked on last 10 transactions.

Table 1. List of all transactions happened till date

Transaction	Amount
1	12000
2	300
3	4500
4	9000
5	6700
6	2000
7	12000
8	12000
9	500
10	600
11	800
12	4560
13	6543
14	2345

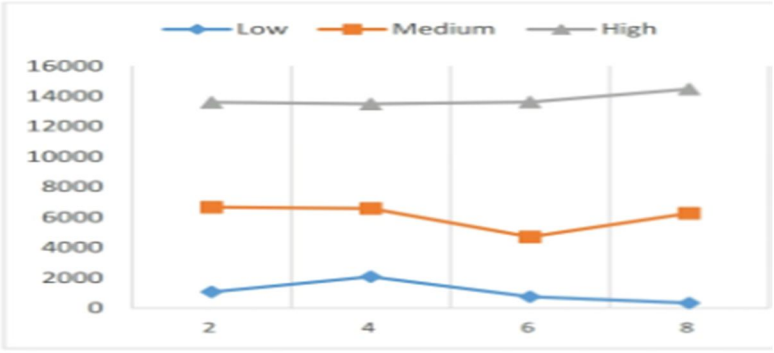


Fig 2: Different categories of the transaction

VII. CONCLUSION

The technology supporting online transaction has provoked the use of payment cards. There are various fraud detection techniques. The Bank database of user transaction can be used for study the user transaction pattern to avoid the fraud transaction. This paper HMM and GA to detecting the credit card frauds thereby recording the IP of the fraud system along with the timestamp when malignant attempted to attack. Genetic algorithm is used for to finding better solutions to technically eliminate the fraud in short time. Genetic Algorithm is appropriate in such kind of application areas. Ultimately this prevents the banks and customers from huge losses and finally leading to reduction of risk.

REFRENCES

[1] Ayushi Gupta and Dhananjay Kumar and Atul Barve , "Hidden Markov Model based credit card fraud detection system with time stamp and ip address" , International journal of computer applications (0975-8887) Volume 166-No.5 , May 2017.
[2] Ruchi Oberoi, "credit -card fraud detection system: using Genetic algorithm "International journal of computer and Mathematical Science IJCMS ISSN 2347-8527 Volume 6, Issue 6 June 2017.
[3] R. Dhanpal and P. Gayatri, "Credit Card Fraud detection using decision tree for tracing email and ip,"International Journal of Computer Science Issues, Vol. 9, no. 2, 2012.
[4] R. Patidar and L. Sharma, "Credit Card Fraud Detection Using Neural Network", International Journal Of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-1, Issus-NCAI2011, June 2011.
[5] Vaishali "Fraud Detection in Credit Card by Clustering Approach", International Journal Of Computer Applications (0975-8887) Volume 98-No.3, July2014.
[6] N. Malini and M. Pushpa "Analysis on Credit Card Fraud Identification baesd on KNN and outlier detection" 3rd International conference on advance in electrical, electronics, information, and communication and Bio-informatics (AEEEICB17).
[7] Shraddha Asati, Rajesh Nasare, and Hemant Turkar "Image forgery detection based on illumination inconsistencies" International Journal of Innovative Research in computer and communication Engineering, vol-5, Issue 4, April 2017.
[8] Rajesh Nasare "Spam Mail Detection Using Artificial Neural Network" Imperial journal of Interdisciplinary research, vol-2, Issue-5, 2016.
[9] Josef Pun, Yuri Lawryshyn "Improving Credit card fraud detection using a meta-classification strategy", International Journal of computer applications (0975-8887) volume 56-No.10, October 2012.
[10] I. Srivastava and A. Kundu, "Credit Card Fraud Detection using hidden markov model", IEEE Transaction on Dependable and Secure Computing, Vol. 5, no. 1, 2008.
[11] Ghosh S., Reilly D. L., "Credit Card Fraud Detection with a Neural -Network" Proceedings of the International Conference on System Science, pp.621-630,1994.
[12] Aleskerov E., Freisleben B., and Rao B., "CREDITWATCH: A Neural Network Baesd Database Mining System For Credit Card Fraud Detection", Proc.
[13] Dorronsoro J.R.,Francisco G., Carmen S., and Carlos S.C.,"Neural Fraud Detection in CeditCard Operation"IEEE Transaction on Neural Network, VOL.-08, NO.-4, and pp.: 827-834, 1997.
[14] Kokkinaki, A., "On Atypical database Transactions: Identification of Probable Fraud using Machine Learning for User Profiling." Knowledge and Data Engeneering Exchange Workshop. IEEE, PP.:107-113, 1997.
[15] Stolfo S.J., Fan D.W., Lee W., Prodromidis A.L., and Chan P.K., "Credit Card FraudDetection Using Meta- Learning: Issues andInitial Results", Proc. AAAI Workshop AI Methods in Fraud and Risk Management, pp.:83-90, 1997.
[16] Brause R., Langsdorf T., and Hepp M., "Neural Data Mining for Credit Card Fraud Detection", Proc. IEEE Int'l Conf. Toolswith Artificial Intelligence, pp.:103-106, 1999.
[17] Kim, M, and Kim T., "A Neural Classifier with Fraud Density Map for Effective Credit Card Fraud Detection", Proceedings of IDEAL. pp.:378-383, 2002.
[18] Chiu A., Tsai C., "A Web Services-Based Collaborative Scheme for Credit Card Fraud Detection", Proceedings of the IEEE International Conference on e-Technology, e-Commerce and e-Service, pp.:177-181, 2004.
[19] Foster and Stine R., "Variable Selection in Data Mining: Building a Predictive Model for Bankruptcy", Journal of American Statistical Association, pp.: 303-313, 2004.
[20] K.RamaKalyani, D.UmaDevi, "Fraud Detection of Credit Card Payment System by Genetic Algorithm", International Journal of Scientific & Engineering Research Volume 3, Issue 7, July-2012.