Journal of Multidisciplinary Engineering Science and Technology (JMEST)
ISSN: 3159-0040
Vol. 2 Issue 7, July - 2015

Credit Card Fraud Detection System Based On User Based Model With GA And Artificial Immune System

Shilpa H. Taklikar[1]
1 M.E Computer Engineering Student,
Sinhgad Institute of Technology
Lonavala, Maharashtra
shilpa.takalikar@gmail.com

Prof. R.P.Kulkarni[2]
2 Professor, Department of Computer Engg
Sinhgad Institute of Technology
Lonavala, Maharashtra
kulkarnirahul1@gmail.com

***Abstract*— Modern science and technology has increased incredibly by inventing new applications and inventions. The expertise machinery has been developed in such a way that it keeps in mind, the new discovery would be contented and will be easy to use for the human beings. One of the discoveries of recent technology is the Credit cards. The use of credit cards has increased a lot for the online purchasing and also for regular purchase. Credit card security is major apprehension to protect its misused from the fraudsters. There comes a necessity to identify the frauds before billing action is executed, this intern will avoid the monetary or economic losses to the banks, card owners. Hence, it is necessary to discriminate the genuine transaction and fraudulent transaction.**

This User model is based on the Artificial Immune System and Genetic Algorithm. This user based model maintains the historical details, problem description and its relative solution to detect a given transaction is normal or fraudulent transaction. Thus, the user based model is an adaptive self-learning based on the historical details of transactions. The alert generated by this model is noticed in advance before billing action is performed. This helps in reducing the massive financial losses to the cardholder and card issuer or the banks.

Keywords*— *Fraud,Credit Card,Fraud Detection System,Fraudulent transaction, Artificial Immune System, Genetic Algorithm

I. INTRODUCTION

Modern science has invented many appliances that are promising to the human beings and are very easy for the resolution. One of the best enhancements of the technology is the development of Credit cards

Credit cards usage was first on track in Europe. UK was first to start credit cards for various purpose usage. Initally, it was used for very small number of transactions. The records were principally tracked on papers . Later there came a need to have faster based approach. Banks then started to use the magnetic strips on back of the credit cards. Magnetic strips in these plastic cards or credit cards encompasses the card owner's confidential details like Card Number details, Card Member name, and all the vital necessary details. The terminals were installed at the retails so that the magnetic strips can be read and the card owner can complete the transaction using the credit card.

Credit cards in this ecommerce world have made noteworthy usage and acceptance of usage by us for mode of payment. These plastic cards are used now-a-days universally for online shopping and also for regular purchasing. This is very convenient way for mode of payment. This can be best illuminated with the help of an example. We could find that Online Railway Reservation IRCTC site is several times slow due to the heavy loaded traffic. Similarly, the online shopping sites like MYNTRA, FLIPKARTS, etc. have been used by many of us for online purchasing. From this we can come to know that, online shopping has increased tremendously. Thus, intern use of plastic cards / credit cards has been increased. But as we know, that every coin has two sides. Thus one side is flexibility to use the credit cards and other side is the need to have secure transaction. The use of Credit cards has also given growth to fraudsters who may cause huge monetary losses for the card holder and also for the card issuer's i.e. banks. Fraudsters can find it easier to misuse the credit cards.

A. Need of Fraud Detection:

Credit cards are used as mode of payment. There comes a need to identify transaction which has been carried out by using credit card is legal transaction done by card holder or it is fraud transaction done by fraudster. In traditional method, the transaction carried out, is true valid transaction or fraud transaction, was able to determine after the billing action has been settled. This resulted in economic losses. So there is a necessity to determine the fraud transaction before the billing or clearance/payment action is performed.

B. Credit cards Fraud Types:

1. Bankruptcy fraud: Bankruptcy fraud is the type of fraud

Wherein, it is very challenging to identify fraud. The credit card's failure of a debtor to pay their debt which is also called as Insolvent gives rise to Bankruptcy fraud. The card holders may be in personal monetary failure and failed to clear the unwanted prevailing loans. Sometimes, banks are required to cover their losses itself. This can be forbidden, by passing the required details to credit bureau. This is one of the ways wherein it helps to identify the past history related to its transactions and loan details of its corresponding customers. Depending upon the historical details, further suitable action can be taken by Banks to avoid this type of fraud. Bankruptcy Foster & Stine (2004) presented a model to determine the bankruptcy fraud to forecast the details of the credit card users. [10]

2. Application fraud: Fraudster is the one who regulates and find out all probable option wherein he can commit the fraud and make misuse of the credit card. One of the possible option is fraudster applies for the credit card with false or invalid information. This way of committing the fraud by fraudster is called as Application fraud.
Application fraud can be characterized by two types: duplicates, identity fraudsters Whenever the application comes from same user with same individual details then it is acknowledged as duplicates. If the application comes from different individuals with the same details then it is known as identity fraudsters. To inhibit this, banks uses the application form that has to be filled by the customers. It contains the obligatory fields which help to regain all the required details. To detect the duplicates, this can be done by cross matching with the last name or any other personal mandatory information. With this procedure, it is possible to identify the duplicates. [10]

3. Theft fraud/counterfeit fraud:[10] This is the sort of Fraud, wherein, we are accustomed with this fraud.
Theft fraud as the name suggests that the fraud has been committed by thief. Whenever the credit card is stolen or lost and the corresponding stolen credit card is used by the thief or fraudster, it is called as theft fraud.
Counterfeit fraud is the sort of fraud wherein the physical card is essentially not mandatory, only the concerned credit card details are prerequisite to the fraudster. This type of fraud occurs whenever the credit card details are used from remotely without the use of actual physical card by the fraudster. The best instance is the fraudster who can carry out the transaction through online shopping, wherein actual physical card is not required to carry out the operation.
Herein, if the card owner reports with the prerequisite details to the bank, then bank will carry out the needed exploration as per the details given by card owner and thus will try to trace out the thief who was involved in this activity.

4. Behavioral fraud: [10]In this ecommerce and modern technology world, the use of plastic cards or credit cards for online purchasing through ecommerce has given rise to behavioral fraud. Herein the fraudster, if he is able to recover the credit card details, can further use these details for fraudulent transactions. For ecommerce online shopping, only the credit card details are necessary without the actual physical card for carrying out the transaction. Professional fraudsters can create an application which may be analogous to the exact copy of application. The credit card owner when uses the application for corresponding shopping or carry out the transaction, the proficient fraudster retrieves the respective card details and all obligatory details, so that the fraudster can further use the details for carrying out the transactions. As the credit card owner uses the exact copy of application, thinking as a legal site to carry out transaction, this is known as Behavioral fraud.

Related Work: There are different methodologies or systems that are used for Credit Card Fraud Detection Systems.

Following are the existing methods:
Decision Tree: Decision trees split the complex problems into simple ones; this is further resolved through repeatedly following the procedure. It can be deliberated as conditional statements.[10] Thus the Decision tree can be used to detect the given transaction is genuine transaction or the fraudulent transaction [13]
Support Vector Machine: Support Vector Machine builds hyper plane and it categorizes the data. If given data falls outside the hyper plane, then it is reflected as fraudulent. [13]
K Nearest Neighbor Algorithm: Distance metric and rule is used to categorize the data.
Hidden Markov Model: Hidden Markov Model is based on the states and has the training data. This technique uses a finite set of states to validate the transactions. The outcome of the state is given to the observer to define its output. The internal states are hidden from the observer, Hence it is called as Hidden Markov model (HMM).[11]

The above methods are used for fraud detection of transactions carried out using plastic cards like credit cards. However, some of these systems undergo lower performance with huge data. And some of the methodologies also may generate false alert or false positive.

Artificial Immune System Background:
Artificial Immune System is based on the natural immune system or human biological immune system.
Every organism is vulnerable by other organism. It is essential for body to defend it from harmful foreign

entities which may damage the body. The foreign entities can be bacteria, viruses, fungi etc. The recognition, elimination or removal of extraneous entities is taken care by natural immune system. It is self-adaptive learning system which has the ability to discriminate between self and non self cells.
There are two key cells which are called as white blood cells. They are B-cells and T-cells. These cells initiate from Bone marrow, and further T-cells go through thymus for maturation, before they circulate throughout the body. Each B-cell can create only specific single antibody. T-helper cells are used for activation of B-cells.

The diagram shows the natural immune system. It exemplifies the step involved in identification and abolition of foreign body:
Step I-II: Show the foreign entity entering the body
Step III: Activating T-Cells,
Step IV: Further B-cells are activated
Step V: Herein matching of antigen is performed.
Step VI: Antibodies are produced
Step VII: Antigen is destroyed or eliminated.

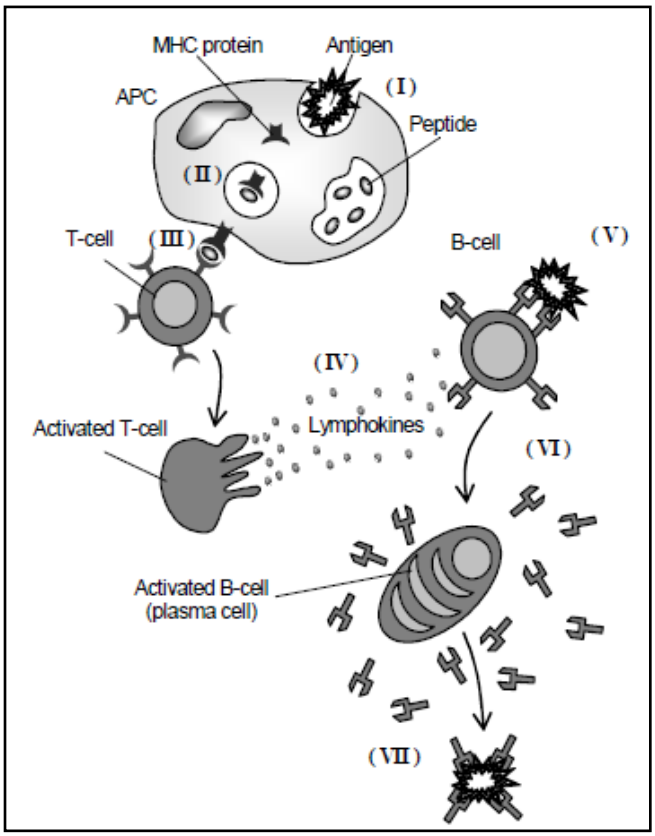

Fig 1: Natural Immune System (taken from de Castro and van Zuben (1999)

Thus Artificial Immune System can be built based on the concepts of Human Immune system to classify or classify as self and non-self elements.

Genetic Algorithm Background:
Genetic Algorithm (GA) useful for natural selection and usual genetics in artificial intelligence to determine the globally ideal solution to the optimization problem from the possible solutions.

GA's is an iterative technique which retains a constant size population of possible solutions. During each repetition step, called a generation, the fitness of existing population is calculated, and population are selected based on the fitness values. The higher fitness chromosomes are nominated for reproduction under the exploit of crossover and mutation to form new population. The lower fitness chromosomes are

eradicated. These new population are calculated, selected and fed into genetic operator procedure all over again until we get an finest solution.

PROPOSED SYSTEM: Use of AIS and Genetic Algorithm for Credit Card Fraud Detection:
The proposed system uses Artificial Immune System and Genetic Algorithm to build user based model which can be used for Credit Card Fraud Detection System.This user based model is used to discriminate the incoming transaction is genuine or fraudulent. This intended technique is applied for detecting the fraudulent transaction that is carried out using plastic cards like Credit Card. Thus, the idea behind the proposed system is to identify a given transaction is genuine transaction done by valid card holders or fraudulent transaction done by fraudsters.

Mathematical Model:

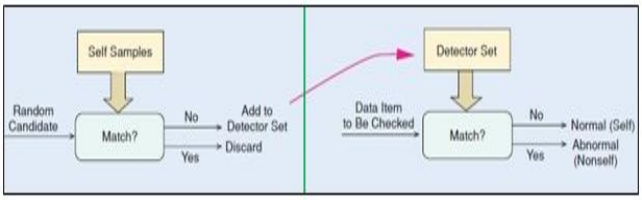

Fig 2: Mathematical Model

Let s (be a main set) = (SDB, LDB, S)
Functionalities:
SDB1 = RegisterUser (uid, password, fullname, address, country, contact, email)
U = AuthenticateUser (uid, password, SDB1)
LDB1 = ManageProducts (pid, product name, cost)
LDB2 = ManageBilling (transactions, items)
LDB = Combination (LDB1,LDB2)
U is represented by list of features which corresponds to the space of states of a system where
S=Subset of space that are considered as normal for the system.
R= Set of detectors generated.
R! = S tends R, fails to match any string in S

Cosine Coefficient can be defined as:
Cosine coefficient = $(X \cap Y)/ (sqrt(X)*sqrt(Y))$
Where X are the details of the new incoming transaction
Y are the details of transactions from the case history.
$X \cap Y$ = number of terms that occur in both X and Y.

Thus, the affinity is determined and depending on this the Detector set is maintained and thus further will

regulate the transaction is genuine or fraudulent transaction.

The following diagram shows the general block diagram for Credit Card Fraud Detection System:

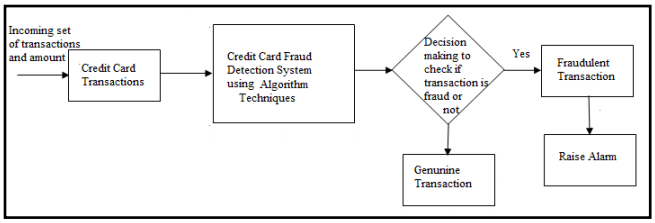

Fig 3: Block Diagram of Credit Card Fraud Detection System

1. The transactions that are supported out using any credit cards are accepted with the required details.
2. This transaction is further given to Credit Card Fraud Detection System
3. The score obtained from Credit card Fraud Detection System is further used to identify or decide next action to be taken.
4. If the transaction is recognized as genuine transaction, then it is sent for further processing of clearance.
5. If the transaction is recognized as fraudulent transaction, then alert or alarm is raised to highlight for the same and is stopped from further processing of that transaction.

The overall steps involved using Artificial Immune System and Genetic Algorithm is as follows:
1. Self Training
2. Search
3. Identification
4. Elimination

Algorithm:
Negative Selection Algorithm:[14]
Input: S= set of seen known self elements
Output D= set of generated detectors
Begin
Repeat
Randomly generate potential detectors and place them in a set P
Determine the affinity of each member of P with each member of self set S
If at least one element in S recognises a detector in P according to recognition threshold, then detector is rejected, otherwise it is added to set of available detectors D
Until stopping criteria has been met
End

Cosine Coefficient can be defined as:
Cosine coefficient = $(X \cap Y)/ (sqrt(X)*sqrt(Y))$
Where $X \cap Y$ = number of terms that occur in both X and Y.

Genetic Algorithm:
1. Generate Initial population
2. Assess Initial population
3. Select population
4. Crossover New population
5. Mutate New population
6. Assess New population

7. Should Terminate Search?
   a. If No, then repeat the steps from 3
   b. If Yes, then Stop

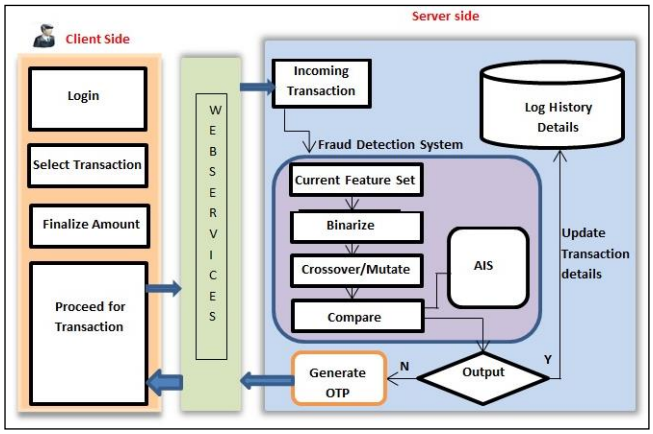

Fig 4: Credit Card Fraud Detection System

In a proposed work, standard architecture is tracked as client-server. Customers using the Internet banking application for online purchasing are considered from client side and on the other hand, Server side is the Bank Server which contains all the details of the respective users.
In our proposed system, the database which is also known as training dataset is maintained, which consists of the current data of transactions which are categorised as Self (Normal) transaction. These transactions are used for detecting the fraud, wherein each case is characterized as the sequence of transaction associated information. The detectors which are generated are marked as immature at the beginning.
Herein, the Negative Selection algorithm is used to filter to classify data as the self and non self. Once, non self is identified then, similar types of antibodies are generated so as to eliminate and abort the further functionality and accordingly generate the alert. When a new transaction is submitted for fraud detection, the fraud detection function is activated. The affinity between the antibodies (detectors) in the gene library and the new antigens is calculated. Herein the Cosine Coefficient is calculated to determine the affinity of the transaction. If the affinity threshold set by the system is exceeded, then it is considered as non self which is also known as fraudulent transaction.

The techniques that are used can be explained as follows:
The vector is formed by using the following features for the respective transactions: Card Number, Category of the Product, Amount, Day of the week (categorised as Weekend or Weekday), and Time(AM or PM). All these details are represented by vector.

Vector or Cell Formation:

CM# Details	Product's Category	Product's Amount	Day of the Week	Time(AM or PM) of Purchase

1. Chromosome representation:

We use binary term vector, so each terminology is either represented as 0 or 1. The vector is formed which is represented in binary format. For Eg: the binary representation in the vector for Day of week is categorised as for Weekdays its considered as 1 and for weekend its represented as 0.Similarly it is considered for Time of week in AM or PM. Range is considered for Product amount and accordingly its represented in binary either 0 or 1.

Whenever the new incoming transaction details are carried out, the vector is formed for the corresponding transaction:

Encoding in chromosome representation as:

CM1 = **0 1 1** 0 1

CM2 = **1 0 0** 0 0

2. Artificial Immune System :

Once the vector is formed, then through Randomization, the other transaction id is created. This is further checked in thehistory database, for which the given transaction has to be compared to determine the fraud score.

This system calculates the fitness evaluation using Cosine Coefficient.

Cosine Coefficient can be defined as:

Cosine coefficient = $(X \cap Y)/ (sqrt(X)*sqrt(Y))$

Where X are the details of the new incoming transaction

Y are the details of transactions from the case history.

$X \cap Y$ = number of terms that occur in both X and Y.

Result from these fitness functions are in the range 0 to 1. By 1.0 means new incoming transaction has sameness with the past transaction carried as per stored in case history. Values near 1.0 mean transaction is more relevant and values near 0.0 mean transaction is less relevant.

1. Crossover:

Crossover is the genetic operator that mixes two chromosomes together to form new descendants. Crossover arises only with some probability. Chromosomes that are not subjected to crossover remain unchanged. The perception behind crossover is exploration of new solutions and exploitation of old solutions. GA's constructs a better solution by combining good characteristic of chromosomes together. Higher fitness value chromosomes have an opportunity to be selected more than the lower ones, so good resolution at all times will be alive and active to the next generation

For example, two chromosomes are crossover between position 2 and 5.

1 *0* 1 1 *1*

1 0 **0 0 1**

The resulting crossover yields two new chromosomes.

1 0 **0 0 1**

1 *0* 1 1 *1*

2. Mutation:

Mutation involves the modification of the values of each gene of a solution with some probability. In accordance with changing some bit values of chromosomes,

give the different breeds. The objective of mutation is restoring lost and exploring variety of data.

For example: randomly mutate chromosome at position 4

1 0 0 0 1

Result 1 0 0 ***1*** 1

3. Termination:

All these above procedures are repeated up to the threshold comparison is reached.

Overall experiment procedure can be defined in brief as:

Step1. Input group of data credit card transactions, every transaction record with n attributes, and standardize the data, get the sample finally, which includes the confidential information about the card holder, store in the data set.

Step2. Compute the critical values, by randomization wherein the case history details are used for classification of the transaction.

Step3. Generate Random transaction by mutate functionality and crossover functionality. And access the new set of population.

Step 4: Repeat these steps until the Termination criteria is matched.

After the fraud attempt is confirmed by the expert, or by the system itself, the antigen will be added to the antigen case library. Thus, if next time the analogous type of fraudulent transaction occurs, then it can directly categorize as fraudulent transaction accordingly. Once the transaction is determined as fraudulent then OTP is generated and sent to the credit card owner's registered contact number. If OTP that is entered correctly by user, then the transaction is recognised as genuine and is sent for further processing. But if OTP that is entered is incorrect, then the transaction is considered as fraudulent transaction which is carried by fraudster and the concerned account is locked out.

RESULTS AND DISCUSSION:

The accuracy is measured using precision and recall parameters. Some testing results are shown into below table:

Type	Actual Objects	Retrieved Objects	Correct Retrieved Objects
Transaction	50	38	25

Table 1: Searching Results

Using above values precision and recall is calculated as follows:

Precision = (Relevant Intersect Retrieved) / Retrieved
= Correct Retrieved Object / Retrieved Objects

Recall = (Relevant Intersect Retrieved) / Relevant
= Correct Retrieved Object / Actual Objects

Type	Precision	Recall
Transaction Classification	0.657894737	0.5
Total	**0. 657894737**	**0.5**
Accuracy (%)	**0. 657894737**	

Table 2: Precision and Recall Calculation

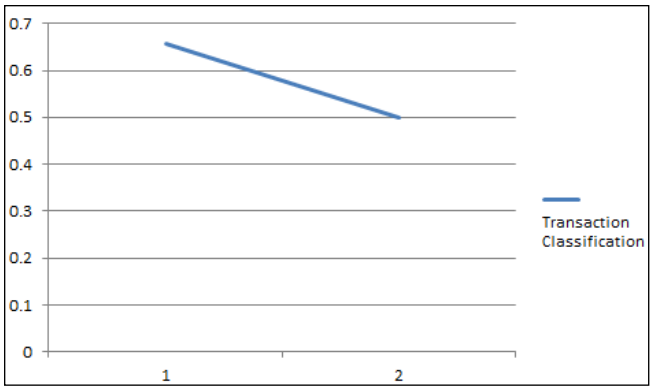

Fig 5: Precision and Recall for different inputs

The proposed system shall detect that the given transaction carried by credit card is genuine or fraudulent Transaction. The proposed system is self-adaptive system which detects transaction is the genuine or fraudulent transaction.

CONCLUSIONS:

The usage of Credit cards as a mode of payment has increased terrifically in this Modern and progressive technology. It is also required to secure the transaction and to recognize the fraudulent transactions that are carried out by fraudsters. The transactions that are carried with the credit card should be secure one and is very essential to determine the fraudulent transactions carried by fraudsters. The system is based on user model with AIS and Genetic Algorithm for detecting the fraudulent transaction.

This system performs its own learning and thus updates the competence of detecting fraud transaction. These techniques will help in determining the fraudulent transaction and thus suitable action can be taken before the billing action is accomplished.

REFERENCES:

[1] Neda Soltani,"A New User-Based Model for Credit Card Fraud Detection Based on Artificial Immune System", IEEE, AISP, 29-33, 2012

[2] Jianyong Tuo, Shouju Red, Wenhuane Lid, Xiu Li. Bine Li, and Lin Lei "Artificial Immune System for Fraud Detection",IEEE,2004

[3] Md. Rafiul Hassan, Baikunth Nath and Michael Kirley,"Data Clustering Algorithm Based On Single Hidden Markov Model", Proceedings of International Multiconference on Computer Science and Information Technology, pp. 57 -66

[4] S. Benson Edwin Raj, A. Annie Portia, "Analysis on Credit Card Fraud Detection Methods, International Conference on Computer, Communication and Electrical Technology,18th 19th March, 2011.

[5] Abhinav Srivastava, Amlan Kundu, Shamik Sural,"Credit Card Fraud Detection Using Hidden Markov Model", IEEE Transactions on Dependable and Secure Computing, VOL. 5, January-March 2008

[6] Rinky D. Patel, Dheeraj Kumar Singh," Credit Card Fraud Detection Prevention of Fraud Using Genetic Algorithm", International Journal of Soft Computing and Engineering (IJSCE), Volume-2, Issue-6, January 2012.

[7] K.RamaKalyani, D.Uma Devi, "Fraud Detection of Credit Card Payment System by Genetic Algorithm, International Journal of Scientific Engineering Research Volume 3, Issue 7, July-2012

[8] U. Aickelin and D. Dasgupta,"Artificial Immune System"

[9] Masoumeh Zarepoor, Sreeja K.R, and M. Afsar. Alam,"Analysis of Credit Card Fraud Detection", Aug 2012

[10] Krishna Kumar Tripathi, Mahesh A. Pavaskar,Survey on Credit Card Fraud Detection Methods, International Journal Of Emerging Technology and Advanced Engineering, Volume 2, Issue 11, November 2012.

[11] V.Bhusari, S.Patil,"Study of Hidden Markov Model in Credit Card Fraudulent Detection", International Journal of Computer Applications (0975-8887) Volume 20- No.5, April 2011

[12] Nicholas Wong, Pradeep Ray, Greg Stephens Lundy Lewis, Artificial immune systems for the detection of Credit card fraud, Info Systems, Vol 22.

[13] Y. Sahin and E. Duman,"Detecting Credit Card Fraud by Decision Trees and Support Vector Machines", International Multiconference of Engineers and computer scientists, March, 2011.

[14] Artificial Immune Systems, UCL

[15] Peter J. Bentley, Jungwon Kim, Gil-Ho Jung and Jong-UK Choi,"Fuzzy Darwinian Detection of Credit Card Fraud", 2007.

[16] Ekrem Duman, M. Hamdi Ozcelik, Detecting credit card fraud by genetic algorithm and scatter search, Elsevier, Expert Systems with Applications, 2011

[17] Daniel Garner, Genetic algorithms for credit card fraud detection, IEEE Transactions, May 2011.

[18] Andrew Watkins, Jon Timmis, and Lois Boggess, Artificial immune recognition system (AIRS): An immune inspired supervised machine learning algorithm, Genetic Programming and Evolvable Machines, September 2004.

[19] A.N.Pathak, Study on Fraud Detection Based on Data Mining Using Decision Tree, International Journal of Computer Science Issues, May 2011