Credit Card Fraud Detection Using Fuzzy ID3

S Md. S Askari
Department of Computer Science and Engineering
Rajiv Gandhi University
Arunachal Pradesh, India
askari.sikdar@gmail.com

Md. Anwar Hussain
Department of Electronics and Communication Engineering
North Eastern Regional Institute of Science and Technology
Itanagar,Arunachal Pradesh,India
bubuli_99@yahoo.com

***Abstract*—By the exponential growth of Credit Card user the fraudulent transactions also have increased dramatically. The genuine transaction and fraudulent transactions are almost similar, so it is very hard to discover a fraudulent transaction form the genuine one. In this paper we have proposed fraud detection algorithm based on Fuzzy-ID3. Intermediate nodes we split using attribute having highest information gain. The leaf nodes classifies the transactions as fraud,doubtful or normal. Experimental result exhibits that the technique is efficient one in detecting frauds.**

***Keywords*—Cramming,Triangulation, Fuzzy-ID3, Information Gain**

I. Introduction

The credit card fraud is a fraudulent transaction by an unauthorized person for his personal interest and the authorized card holder and the card provider totaly unaware about the transaction for the moment.Credit card fraud is a wide-ranging term for theft and fraud committed using a credit card or any similar payment mechanism as a fraudulent source of funds in a transaction. The purpose may be to obtain goods without paying, or to obtain unauthorized funds from an account. Credit card fraud is also an adjunct to identity theft[26]. Bankers and the commercial establishments are the victims in most of the cases, as the bank authorities do not have the options for the physical verification of users of the card.

The Following are the different ways of credit card frauds:[3]

- An act of criminal deception (Mislead with intent)by use of unauthorized account and/or Personal Information.
- Illegal or unauthorized use of account for personal gain.
- Misrepresentation of account information to obtain goods and/or service.
- Simple theft and Counterfeit the card.
- Card is never received by the genuine owner.

In Germany over the past few years the usage of credit cards have been increasing. The market for credit cards and debit cards has been grown by 23.3% from 2004 to 2009[6]. The Indian credit card market is in its growth phase, it recorded a growth of about 30% a year. Debit cards are growing at 40%. The RBI data put total electronic transaction in the country at over ₹ 2,35,000 crores in 2006-07. This increased to ₹ 3,60,000 crores in the first 10 months (April-January) of 2007-08. At the end of April-January2007-08, all of us together held about 27.5 million credit cards transacted ₹ 47,476 crores through these cards in 10 months of the year[5]. And as per the RBI statistics the amount of transaction done by Credit cards at Point of Sale in January 2012 is ₹ 88676.99 million, in January 2013 is ₹ 113920.31 million and in January 2014 it is ₹ 141620.33 which indicates a huge increase of credit cards user[15].

As the number of users are increasing frauds are also exponentially getting increased though the researchers have developed different techniques to detect the fraud transactions. The review of *Bolton and Hand*[4] have given a depth study of the existing techniques in fraud detection. And it is observed that as the fraud detection problem is a classification problem, mostly the Nural Networks, Rule based approaches, Data Mining techniques and HMM are used to deal with the fraud detection problems. *Maes et al.*[12]have used two machine learning techniques: Artificial neural networks with Feed Forward Multi-layer Perceptron that consists of different layers of perceptrons which are interconnected by a set of weighted connections and Bayesian belief networks in credit card fraud detection. Their idea was to provide some computational learner with a set of training data consisting of some feature values on which fraud detection to be run. After a learning process the program will correctly classify a new transaction as fraudulent or not fraudulent given some features of that transaction. *Gadi et al.*[8] have used Artificial Immune System(AIS) in credit card fraud detection with optimized and robust set of parameters for the best results over other methods. The effectiveness of AIS is compared to other techniques and found beneficial. *Aleskerov et al.*[1]have proposed the CARDWATCH, a database mining system used for credit card fraud detection. It trains the nural network with the past transactions of the customer and then the current spending behavior is observed to detect the frauds. *Halvaiee et al.* [9] using the Hadoop and AIS based on human immune system have introduced a new model called AIS-based Fraud Detection Model(AFDM). They have used an immune system inspired algorithm(AIRS)and have improved the algorithm for fraud detection. *Olszewski*[16] have proposed fraud detection system visualizing the user activities on self organizing map SOM which is a method of mapping a high dimensional data into a 2-dimensional map of neurons and classifying based on threshold value which works for multiple frauds. *Dorronsoro et al.*[7] have developed a model based on neural classifier named Minerva that has number of rating modules and only the

ISBN:978-1-5090-6471-7/17/$31.00

core modules have nural rating functions. Fishers discriminant analysis classification procedure is combined with the nonlinear projecting capabilities of multilayer perceptrons(MLP)for constructing the neural module of Minerva for credit card fraud detection. *Syeda et al.*[23] have used the Granular Nural Network(GNN) with rule based approach and fuzzy nural network. The fuzzy nural network is trained with training dataset which in turn produces fuzzy rules for checking the transactions. *Srivastava et al.*[22] have modeled the credit card transaction process using HMM which they have used as Fraud Detection System. The FDS will receive the card details and transaction amount to check authenticity. The FDS concern only on price range which is classified into low(l), medium(m), and high(h) and used as observation symbols of the HMM for detection of frauds. *Iyer et. al*[10] have also used HMM and modeled the sequence of operations in credit card transaction processing using HMM to detect frauds. The HMM is trained with the normal behavior of the card holder then the this trained HMM checks the incoming transactions if it is not accepted with sufficiently high probability, it is considered as fraud. *Panigrahi et. al.*[17] have developed a Behavior-based 6−tuple $\{System, C, P, \psi, \theta_{LT}, \theta_{UT}\}$ fraud detection Model. where System is the target system, C is the set of credit cards, P is the set of profiles of Card holders, $\psi(T_{j,\rho}^{C_k})$ is the suspicion score of the j^{th} transaction $T_{j,\rho}^{C_k}$ on card C_k and ρ is the time gap from the previous transaction on the same card, θ_{LT} is the lower threshold, where $0 \leq \theta_{LT} \leq 1$ and θ_{UT} is the upper threshold, where $0 \leq \theta_{UT} \leq 1$. Initial beliefs are assigned to each incoming transaction to determine its deviation from a normal profile. Dempster−Shafer theory is applied to get the overall belief by combining the initial beliefs. Bayesian learning is used to strengthened or weakened the overall belief based on its similarity with the fraud or genuine transaction history. *Shahin and Duman*[20] have developed seven alternative models based on decision tree methods and support vector machine methods and then the performance of the classifier models are checked for different data sets with different sizes. The decision trees they have used are C5.0, C&RT and CHAID and the SVM with polynomial, sigmoid, linear and RBF kernel functions are used. The decision trees attempt to separate the records into mutually exclusive subgroups. *Sahin et al.*[21] to detect fraud transactions proposed a new cost-sensitive decision tree induction algorithm which at the time of selecting splitting attribute at each nonterminal node minimizes the misclassification costs. Proposed new cost sensitive metric Saved Loss Rate SLR to evaluate the performance. *Misra and Dash*[13] have implemented MLP, Decision Tree and Chebyshev FLANN three different approaches and compared the results for credit card fraud detection. The chebyshev Functional Link Artificial Neural Network is a single layer neural network in which the original input pattern in lower dimensional space is expanded to a higher dimensional space by using a set of orthogonal functions[14], [18]. After implementing all these three they have compared and shown that MLP outperforms both the Decision Tree and Chebyshev FLANN for Credit card Fraud detection. *YU and Wang*[27] have proposed a model using outlier detection based on distance sum, that mines the fraud transactions as outliers. They have defined Outlier as: Data set $T = \{t_1, t_2, t_3, ...t_n\}$ U is one data object, If the p parts of data set named S in data set is far way from object U, $S \in T$, $U \in T$, then U is Outlier.

From the literature it is found that most of the researchers have used different form of Artificial intelligence and Data Mining Techniques, like Nural Networks, Hidden Markov models, Outlier based FD, Behavioral based Decision Tree etc. After going through the existing works we feel the need of mathematical solution for the Fraud detection. As the fraud detection itself is a classification problem, we found Fuzzy Logic based decision tree ID3 a suitable technique to develop a more effective fraud detection system. In this paper we have chosen the attribute from the dataset of credit card transactions for which the information gain is highest among all the attributes of transactions for splitting node in decision tree creation process. And the leaf node of the decision tree will classify the transactions either as fraud or normal. We prioritize to focus on in no cost the fraud transactions should be classified as normal. We also have discussed the results to show the effectiveness of our technique.

This paper consists of 4 sections, second section describes the basic problem, the third chapter is about the Fuzzy ID3 in CFD, fourth section describes the results and analysis of Fuzzy ID3 in CFD, and finally the fifth section cocludes the paper.

II. Problem Definition

With the increasing usage of credit cards and debit cards, the frauds related with these are also increasing. In last five years the card holders affected in Germany 10%, in India 27%, in UK 31%, US 37% and all other countries are also being affected seriously[11]. Annual global fraud loses for Credit Card and Debit Card frauds in the year 2008 it is $6.4 Billions in 2009 it is $6.9 Billions in 2010 it is $7.6 Billions in 2011 it is $9.8 Billions and in 2012 it is $11.2 Billions and over all one in four customers becomes the victim of card frauds[25], [11]. These vast increase of Economic loses by frauds mandates the researchers to think about how these frauds can be detected and prevented more and more effectively though it is very difficult. For the development of more effective detection and prevention techniques for fraud it is important to gather adequate knowledge about different types of frauds, how these frauds works.

A. Different Types of Credit Card Frauds:

As the new technologies are being developed day by day, the fraudsters are also developing new types of frauds to by-pass the security mechanisms. There are various types of frauds are there in the current system around the world experienced by the card holders. The highly prevailing frauds experienced by the users are as follows[6], [3]; *a. Theft Fraud/Counterfeit Fraud. b. Application Fraud. c. Behavioral Fraud. d. Bankruptcy Fraud. e. Cramming/Salami Attack. f. Triangulation Fraud*.

1) Theft Fraud/Counterfeit Fraud:: The Theft fraud and the Counterfeit Fraud are inter related. In case of Theft Fraud the Credit Card is stolen by the fraudster or the lost card is found by the fraudster and it is used by him/her many times as possible until the card is blocked by the bank.

In the Counterfeit fraud the Card is used remotely where the presence of physical card is not concern, only the card details are required. The different techniques used by the fraudsters for gathering Card information to use for Counterfeit Frauds are: fake card making,tearing magnetic strip,white plastic,altering card data and skimming. And skimming is mostly used technique in counterfeit fraud, where original data on a credit card magnetic stripe are copied electronically onto another. Cashiers/Employees of business establishments have been found to carry pocket skimming devices, a battery operated electronic magnetic stripe reader, to get details of the card they swipe the cards usually when the card holder waits for the transaction to be validated. With the details obtained by the skimmer the fraudster carries out transactions for shopping, billing, etc. in card-not-present manner[3].

2) Application Fraud:: In Application fraud the fraudster applies with false documents or with stolen documents of someone else to get a Credit Card. The frudster uses a false name with temporary address or may look for someone who is going to leave the address very shortly and his electoral register will be updated after few months or years. As the banks usually checks electoral register to confirm the address of a new customer. That is the fraudster pretends to be some one else; this application fraud is termed as *Assumed Identity*. Also fraudster may give some false statements of his financial status to acquire more Credit Cards than his entitlements by producing some forged financial statement documents; this type of application fraud is called *Financial Fraud*.

3) Behavioral Fraud:: Behavioral fraud occurs when details of legitimate cards have been obtained fraudulently and sales are made on a 'card holder' present basis. These sales include telephone sales and e-commerce transactions, where only the card details are required.[6], [4]

4) Bankruptcy Fraud:: The Bankruptcy Fraud is in which the Card holder knows that he is not able to pay the amount of the items he is purchasing. Later the bank will issue him an order to pay but the card holder will not respond. And finally the customer will be recognized by the bank as in a personal bankruptcy state so the bank will not have any option for recovery. These are considered as charge-off losses not included in Frauds.

5) Cramming/Salami Fraud:: Cramming is when a fraudster using a Credit Card makes piecemeal transactions over a long period of time. The transactions will be so small that card holder may not even notice the charges. In India such a fraud is called Salami attack, where small amounts are fraudulently charged on the card. This type of fraud is active in India.

6) Triangulation Fraud:: The fraudster maintains some web site through which the items are offered pay on delivery at very high discount. The site they maintain looks like a genuine sales site. At the time of ordering online, customer provides card details to the site. Once the fraudsters get the card details, they order items from a legitimate site using some other stolen credit card details. Then the fraudster keeps shopping online using the credit card details of the customer.

These huge economic losses by the frauds leads the card issuers, financial institutions and researchers to develop new fraud detection technologies. Despite the development of new technologies the frauds are increasing day by day. From the literature it is observed that most of the techniques consider the transaction history of the users, study the user behavior to form rules for fraud detection. Then the new transactions are synchronized with those behaviors or rules and if markable deviation is observed then the new transaction will be tagged as fraud otherwise normal.

The problem is to analyze on-line financial transactions, by considering the most relevant feature or attribute of the transactions to detect frauds with reference to the fuzzy set defined for the most relevant attribute.

III. METHODOLOGY

Fuzzy Logic is the invention of Zadeh[28] for representing the cognitive uncertainties, measuring the intensity of the truth values for unquantifiable measures or probabilistic measures within the range of 0 and 1.

Let D be the collection of examples or objects or instances represented in set theoretic notion as $\{e_1, e_2.....e_n\}$ where the D is called the universe of discourse and the e_i is the individual example or object(element) of D. A fuzzy set A in the universe of D is described by a membership function $\mu_A(e) : D \rightarrow [0, 1]$ which quantifies the intensity or grade of membership of the element e in the fuzzy set A. The member ship crisp value $\mu_A(e) = 1$ means that e is 100% a member of A and $\mu_A(e) = 0$ means that e is 100% not a member of A, and in case of fuzzy logic $0 \leq \mu_A(e) \leq 1$ which means that e is partially member of A. Hence as the membership values goes more close to 1, the intensity of membership of e in A becomes more strong.

The Id3((Iterative Dichotomiser 3) was developed by Quinlan[19] for dealing with symbolic data by expressing the knowledge as a decision tree. The decision tree is generated as a mathematical model by the training data to classify new instances with simple inference mechanism, in the tree the leaf nodes represents the class names and the branches represent the conditions. A random subset of the training set called *window* is considered and a corresponding decision tree is formed which correctly classifies all the examples in the *window*. The tree is then used in classification of the rest examples of training set. If the tree classifies the entire training set correctly the process terminates otherwise a selection of the incorrectly classified objects is included in the *window* and the process continues. After few iterations a correct decision tree is formed.

a) Example: Suppose the D is the dataset or set of examples and T is the condition applied to classify the data set D with possible outcomes $O_1, O_2, ..., O_n$. Each example will have one of these outcomes satisfying the condition T and will divide the dataset D into $D_1, D_2, ..., D_n$ respectively for each outcome. Recursively for each D_i the decision tree process is repeated and finally we will have the decision tree for all the examples in the dataset D. Classification rules are formed by Climbing down through the branches to the leaf from the root node. Leaves are the class names and the branch are the condition outcomes with intermediate nodes represent the conditions applied on attributes.

In $Id3$ in the decision tree forming process starting from root to leaves at every node (except the leaf) which attribute to be considered for applying the condition for splitting the dataset(node) into different subsets as the nodes of the next level is decided based on information based methods: *Information Content, Expected Information and Information Gain*[19], [2].
Every example e_i in the dataset D is in class c_j with probability

$$p_j = \frac{|D_j^N|}{|D^N|} \tag{1}$$

where the D^N representing the set of examples in node N, D_j^N representing the set of examples in node N belonging to class c^j.
The Information Content for the Dataset in the current node, N:

$$I^N = -\sum_{j=1}^{|C|} p_j \log_2 p_j \tag{2}$$

where C is the set of classes in the dataset in node N.
The expected information in a subtree of node N for an individual attribute $A^k \in A$ in node N:

$$E^{N|A^k} = \sum_{l=1}^{|A^k|} \frac{|D_{a_l^k}^N|}{|D^N|} E^{N|a_l^k} \tag{3}$$

where $D_{a_l^k}^N$ is the set of examples whose attribute value a_l^k for A^k corresponds to the nodes branch i.e. the set of examples in the respective child node of N for the attribute value a_l^k.
The information gain $G(A^k, N)$ for the attribute A^k in node N is:

$$IG(A^k, N) = I^N - E^{N|A^k} \tag{4}$$

The classification problem in which the attributes take cognitive values, determining the correct class is not possible. To solve that kind of classification problems *Umano et al.*[24] have proposed Fuzzy-Id3 algorithm by considering the feasibility of Id3 and Fuzzy Logic.

The fuzzy-Id3 algorithm proposed by[24] works by defining fuzzy sets for all attributes and forms a fuzzy decision tree in the same way as Id3 described above. Here the difference is that in Id3 the information gain is based on the probability of the attribute but in Fuzzy Id3 probability is computed based on the membership values of the attribute.

Suppose we have a set of data D in which each example is described by the attributes $A = \{A^1, A^2, ...A^l\}$ and each example has one of the classes $C = \{C_1, C_2, ...C_n\}$ and the attribute may have the fuzzy values $A^i = \{v_{i1}, v_{i2}, ...v_{im}\}$ for different i, the m may be different. Unlike the general Id3 here $|D|$ is the sum of the membership values of the examples in D.The probabilities and the equations are computed as follows:
The probability for the j^{th} fuzzy set of A^k:

$$p_{kj} = \frac{|D_{v_{kj}}|}{\sum_{j=1}^{m} |D_{v_{kj}}|} \tag{5}$$

where m is the total number of fuzzy sets for the attribute A^k.

The probability of examples with class c_j at node N:

$$p_j = \frac{|D_j^N|}{|D^N|} \tag{6}$$

where j is the class number.
The expected information in a subtree of node N for an individual attribute $A^k \in A$ in node N:

$$E^{N|A^k} = \sum_{j=1}^{m} (p_{kj}.I(D_{v_{kj}})) \tag{7}$$

The Information Content for the Dataset in the current node, N:

$$I^N = -\sum_{j=1}^{n} p_j \log_2 p_j \tag{8}$$

where j is the class number and n is total the number of classes in the dataset in node N.
The information gain $IG(A^k, N)$ for the attribute A^k in node N is:

$$IG(A^k, N) = I^N - E^{N|A^k} \tag{9}$$

Algorithm 1 FuzzITree

BEGIN :

1: Root Node N contains all the Examples e_i with $\mu_N(e_i)$=1.
2: **if** a node N with fuzzy set of data D^N satisfies the following conditions.
   1. $D_j^N \subseteq D^N$ with class c_j satisfies $p_j \geq \theta_r$ $\triangleright$ θ_r is a threshold
   2. sum of the membership values $|D^N| < \theta_n$ $\triangleright$ θ_n is a threshold
   3. No attributes are there to split the node.
   then it is a leaf node with class c_j .
3: **end if**
4: **if** If above conditions are not satisfied **then**
   1. $A_{G_{max}}^k \leftarrow Max(IG(A^k, D^N)\ \forall\ k = 1, ..., l)$
   2. Split the dataset D^N into fuzzy subsets $D_1, D_2, ...D_m$ by branching as per the m no. of fuzzy terms of the attribute $A_{G_{max}}^k$
   3. Label the branches with corresponding fuzzy term $v_{Gmax,j}$.
   4. $\mu_{D_j}(e) \leftarrow \mu_{D^N}(e) * \mu_{v_{Gmax,j}}(e)$.
   5. $D^N \leftarrow D_j \forall j = 1, 2, ...m$.
5: **end if**
6: Repeat the steps from 2 recursively.

END

The credit card fraud detection is a classification problem where the objective is to classify the transactions, and act accordingly to alleviate the loss causing by the fraud transactions. The data used here is of a prominent bank of Singapore and the attributes we considered are:

- branch_code-branch code of transaction.
- cust_ac_no- account number.
- ccy-debit currency in which the transaction is done.

Sl	Br_Code	Ac_No	Ccy	DrCr_ind	Cpty_Ac_I	TTS	Locn	Txn_amt	Class
1	PL0	PL0123	SGD	D	CPTY01	04-Oct-2006 20:29:08	LCN1	1000	1
2	PL1	PL1123	SGD	D	CPTY02	04-Oct-2006 20:29:10	LCN2	1200	1
3	PL0	PL0123	SGD	D	CPTY03	04-Oct-2006 20:29:11	LCN2	2300	3
4	PL1	PL1123	SGD	D	CPTY03	04-Oct-2006 20:31:15	LCN2	1200	2
5	PL0	PL0123	SGD	D	CPTY01	04-Oct-2006 20:33:10	LCN1	1000	1
6	PL0	PL0123	SGD	D	CPTY03	04-Oct-2006 20:33:29	LCN2	1200	3
7	PL0	PL0123	SGD	D	CPTY02	05-Oct-2006 20:30:10	LCN2	1200	2
8	PL1	PL1123	SGD	D	CPTY03	05-Oct-2006 22:15:12	LCN1	1000	1
9	PL0	PL0123	SGD	D	CPTY01	06-Oct-2006 21:30:10	LCN2	650	1
10	PL1	PL1123	SGD	D	CPTY03	07-Oct-2006 15:20:11	LCN2	650	1
11	PL0	PL0123	SGD	D	CPTY01	07-Oct-2006 17:30:15	LCN1	2300	3
12	PL1	PL1123	SGD	D	CPTY03	07-Oct-2006 17:40:11	LCN2	650	1
13	PL2	PL2123	USD	C	SELF01	11-Oct-2006 12:10:16	LUS01	2000	1
14	PL0	PL0123	SGD	D	CPTY03	12-Oct-2006 12:15:15	LCN2	2300	2
15	PL2	PL2123	USD	C	SELF01	13-Oct-2006 12:11:17	LUS01	1800	1

TABLE I. BANK TRANSACTIONAL DATA

SL	μ	Br_Code	Ac_No	Ccy	DrCr_ind	Cpty_Ac_I	TTS	tts_diff	Locn	Loc_diff	Txn_amt	Class
1	1	0.674453	0.674453	0.707107	1	0.037635	7.329548535648148e+005	0	73363	0	1000	1
2	0.2	0.73577	0.73577	0.707107	1	0.15054	7.329548535879629e+005	0	67719.9	0	1200	2
3	0.8	0.674453	0.674453	0.707107	1	0.301084	7.329548535995371e+005	0.05000006407499	67719.9	5643.1	2300	3
4	0.7	0.73577	0.73577	0.707107	1	0.301084	7.329548550347223e+005	2.08333343267441	67719.9	0	1200	2
5	1	0.674453	0.674453	0.707107	1	0.037635	7.329548563657408e+005	4.03333341702819	73363	0	1000	1
6	0.3	0.674453	0.674453	0.707107	1	0.301084	7.329548565856481e+005	0.31666662544012	67719.9	5643.1	1200	3
7	0.3	0.674453	0.674453	0.707107	1	0.15054	7.329558542824074e+005	1.437000000011176e+003	67719.9	5643.1	1200	2
8	0.4	0.73577	0.73577	0.707107	1	0.301084	7.329559272222222e+005	1.546033333372325e+003	73363	5643.1	1000	1
9	1	0.674453	0.674453	0.707107	1	0.037635	7.329568959490741e+005	2.936999999955297e+003	67719.9	5643.1	650	1
10	0.7	0.73577	0.73577	0.707107	1	0.301084	7.329576390162037e+005	2.464983333349228e+003	67719.9	5643.1	650	1
11	0.9	0.674453	0.674453	0.707107	1	0.037635	7.329577293402777e+005	1.200083333272487e+003	73363	5643.1	2300	1
12	0.7	0.73577	0.73577	0.707107	1	0.301084	7.329577362384260e+005	1.400000000372529e+002	67719.9	0	650	1
13	0.5	0.061314	0.061314	0.70711	0	0.564532	7.329615071296296e+005	0	56433	0	2000	1
14	0.8	0.674453	0.674453	0.707107	1	0.301084	7.329625105902777e+005	6885	67719.9	5643.1	2300	2
15	0.1	0.061314	0.061314	0.70711	0	0.564532	7.329635078356481e+005	2.881016666684300e+003	56433	0	1800	1

TABLE II. NUMERICAL NORMALIZED VALUES OF BANK TRANSACTIONAL DATA

- drcr_ind-debit or credit.
- cpty_ac_no-payee account number.
- TTS-Transaction Time stamp.
- locn_ref_no-terminal or PoS reference.
- txn_amt-transaction amount.
- class-class to which the transaction belongs.

Fuzzy sets $\{very\_low, low, medium, high, very\_high\}$ for the attribute *'txn_amt'* are defined here as:
$verylow = \{1/600, 0.8/650, 0.6/700\}$
$low = \{0.3/800, 0.5/850, 0.6/900, 0.8/950, 1/1000, 0.8/1050, 0.6/1100, 0.3/1200\}$
$medium = \{0/1200, 0.2/1250, 0.5/1300, 0.7/1350, 1/1400, 0.8/1450, 0.6/1500, 0.5/1550, 0.3/1600\}$
$high = \{0.2/1550, 0.5/1600, 0.7/1650, 1/1700, 0.8/1750, 0.6/1800, 0.5/1850, 0.3/1900\}$
very high$= \{0.1/1850, 0.3/1900, 0.5/1950, 0.6/2000, 0.8/2050, 1/2100\}$

To get the fuzzy sets for the attributes TTS nad Loc_ref, we have calculated the difference of TTS(in minutes) and Loc_ref of a particular accounts current transaction with the TTS and Loc_ref of the immediate previous successful transaction of that particular account respectively and we have defined the fuzzy sets for the attributes TTS_diff and Loc_ref_diff .
Fuzzy sets $\{very\_low, low, medium, high, very\_high\}$ for the attribute *'TTS_diff'* are defined here as:

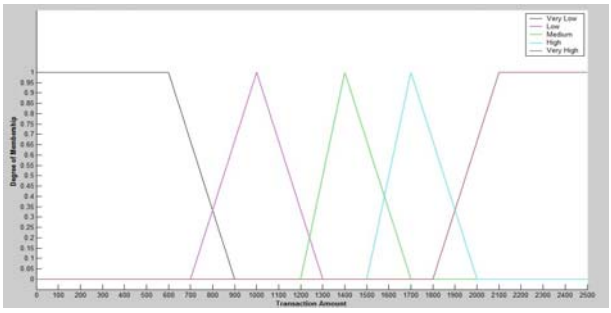

Fig. 1. Fuzzy membership for the Attribute Transaction Amount

$very\_low = \{1/0.05000006407499, 1/0.31666662544012, 1/2.08333343267441, 1/4.03333341702819, 1/10, 0.9/11, 0.85/11.5, 0.8/12, 0.75/12.5, 0.5/15, 0.3/17, 0.10/19, 0/20\}$
$low = \{0/12, 0.0556/12.5, 0.3333/15, 0.3889/15.5, 0.6667/18, 0.7222/18.5, 0.8889/20, 1/21, 0.4444/22, 0.2778/25, 0.2500/25.5, 0.1667/27, 0.0556/29, 0/30\}$
$medium = \{0/25, 0.1489/32, 0.2128/35, 0.3191/40, 0.4255/45, 0.5319/50, 0.6383/55, 0.7447/60, 0.8511/65, 0.9574/70, 1/72, 0.4211/80, 0.3158/90, 0.2105/100, 0.1053/110, 0.0526/115, 0/120\}$
$high = \{0/90, 0.0667/95, 0.1333/100, 0.2667/110, 1/1700, 0.4667/125, 0.60/135, 0.6667/140, 0.7333/145, 0.8667/155, 1/165, 0.4333/175, 0.3667/185, 0.2667/200, 0.1333/220, 0.0667/230, 0.0333/235, 0/240\}$
$very\_high = \{0/210, 0.1667/220, 0.3333/230, 0.50/240, 0.5833/245, 0.75/255, 0.9167/265, 0.9833/269, 1/270, 1/6885\}$

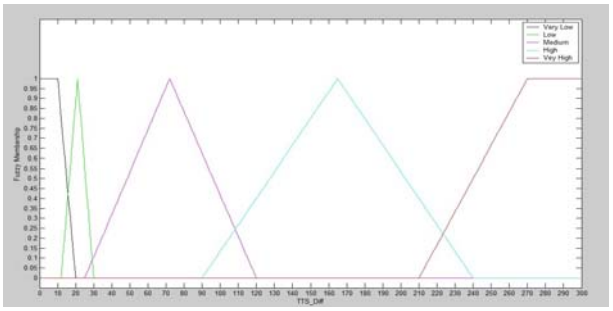

Fig. 2. Fuzzy membership for the Attribute TTS_diff

Fuzzy sets $\{low, medium, high\}$ for the attribute *'Loc_diff'* are defined here as:
$low = \{1/0, 1/5, 0.75/5.5, 0.50/6, 0.25/6.5, 0/7\}$
$medium = \{0/5, 0.1429/10, 0.2857/15, 0.5714/25, 0.8571/35, 1/40, 0.4706/50, 0.3529/60, 0.1176/80, 0/90\}$
$high = \{0/40, 0.1250/45, 0.50/60, 0.8750/75, 1/80\}$
(Here in the figure the distances are marked in miles for

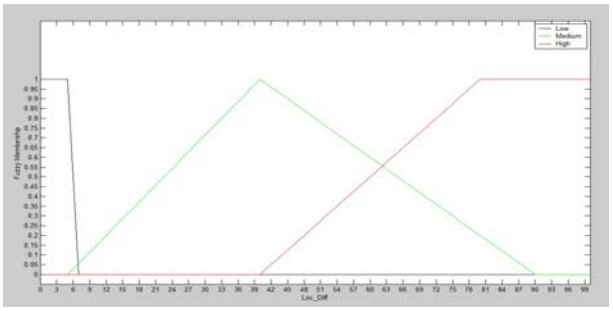

Fig. 3. Fuzzy membership for the Attribute Loc_diff

drawing comfortable.)

IV. Result and Discussion

To select the best attribute for generating the decision tree using Fuzzy Id3 we need to calculate the information gain. Here $|D| = 9.4, |D_{c_1}| = 6.5, |D_{c_2}| = 1.8, |D_{c_3}| = 1.1$ So using the equation-8 the Information Content $I(D) = -\frac{|D_{c_1}|}{|D|}\log_2\frac{|D_{c_1}|}{|D|} - \frac{|D_{c_2}|}{|D|}\log_2\frac{|D_{c_2}|}{|D|} - \frac{|D_{c_3}|}{|D|}\log_2\frac{|D_{c_3}|}{|D|}$ $= 1.1868$

The membership values for the fuzzy sets we have defined for the attribute txn_amt are calculated using the formula defined as:

$$\mu_{D_j}(e) \leftarrow \mu_D(e) * \mu_{v_{Gmax,j}}(e) \tag{10}$$

where $\mu_D(e)$ is the membership value of example e in D and $\mu_{v_{Gmax,j}}(e)$ is the membership value of example e in j^{th} vague set for the attributre A_{max}. i.e. for the transaction 2 amount 1200 has the membership value 0.33 in the fuzzy set *low* of the attribute txn_amt, as per the equation-10 the final membership value will be 0.231 "i.e." $0.70 * 0.33$, where $\mu_D(e)$ is 0.70 and $\mu_{v_{Gmax,j}}$ is 0.33

Now to choose the best splitting attribute we need to calculate the Information Gain individually for all the attributes we have considered as relevant in forming the ID3. The membership values we have calculated considering both the membership values (μ) in the dataset and the membership values in the defined fuzzy set for individual attributes using equation-10. Using these membership values the Information Content and the Expected Information for the relevant attributes are computed as follow.
First we have considered the attribute *txn_amt* for the fuzzy set *very_low* of the *txn_amt* we have
$|D^{v\_low}| = 1.99,$
$|D_1^{v\_low}| = 1.99, |D_2^{v\_low}| = 0, |D_3^{v\_low}| = 0$
Now using the equation-8 the Information Content I^N for this fuzzy set *v_low*: $I^N(D_{v\_low}) = 0$

For the fuzzy set *low* of the *txn_amt* we have
$|D^{low}| = 2.89,$
$|D_1^{low}| = 2.4666, |D_2^{low}| = 0.329, |D_3^{low}| = 0.099$
Information Content I^N for this fuzzy set *low*: $I^N(D^{low}) = 0.7199$

For the fuzzy set *medium* of the *txn_amt* we have
$|D^{medium}| = 0,$
$|D_1^{medium}| = 0, |D_2^{medium}| = 0, |D_3^{medium}| = 0$
Information Content I^N for this fuzzy set *medium*: $I^N(D^{medium}) = 0$

For the fuzzy set *high* of the *txn_amt* we have
$|D^{high}| = 0.04,$
$|D_1^{high}| = 0.04, |D_2^{high}| = 0, |D_3^{high}| = 0$
Information Content I^N for this fuzzy set *high*: $I^N(D_{high}) = 0$

For the fuzzy set *very_high* of the *txn_amt* we have
$|D^{v\_high}| = 2.83,$
$|D_1^{v\_high}| = 1.23, |D_2^{v\_high}| = 0.80, |D_3^{v\_high}| = 0.80$
Information Content I^N for this fuzzy set *v_high*: $I^N(D^{v\_high}) = 1.553$

We calculate the Expected Information for the attribute *txn_amt* by applying equation-7 as:
$E^{N|txn\_amt} = (\frac{1.99}{7.75} \times 0) + (\frac{2.89}{7.75} \times 0.7199) + (\frac{0}{7.75} \times 0) + (\frac{0.04}{7.75} \times 0) + (\frac{2.83}{7.75} \times 1.553) = 0.8356$

Now using the equation-9 we can calculate the Information Gain for *txn_amt* as:
$IG(txn\_amt, N) = 1.1868 - 0.8356 = 0.3512$
Similar way if we calculate the Information Gain for the attributes tts_diff and Loc_diff we will get:
$IG(tts\_diff, N) = 1.1868 - 0.8051 = 0.3817$
$IG(Loc\_diff, N) = 1.1868 - 1.1171 = 0.0697$

The algorithm FuzzITree selects the attribute for which the Information Gain is highest as splitting attribute in current node and it creates branches. This process is repeatedly applied in the new subtrees until it reaches a leaf which determines the class of the transaction for which the path from root to leaf is resembled.

V. Conclussion

We have presented here the mathematical process to create an ID3 decision tree by using fuzzy logic. We have applied our

algorithm FuzzITree on the normalized training data shown in table II. Out of all these attributes we have considered the most relevant attributes txn_amt, tts_diff, and Loc_diff to reduce the irrelevant processing so that the detection can be done in optimal time. As per the execution result of the algorithm on training data it is observed that all the transactions are classified correctly except the transaction no. 7. We have conducted test on some other transactions as well with different values of the attributes in different situations to determine the detection rate and by considering all the test results it is observed that the detection rate is 89%.

We will further proceed our research using fuzzy concept to find the more optimal way for fraud detection purpose. We will be using our algorithm along with the fuzzy neighborhood techniques. Also some empirical studies are going on for further efficient solutions.

REFERENCES

[1] E. Aleskerov, B. Freisleben, and B. Rao. Cardwatch: a neural network based database mining system for credit card fraud detection. In *Proceedings of Computational Intelligence for Financial Engineering (CIFEr), 1997*, pages 220–226. IEEE, 1997.

[2] Łukasz Bartczuk and Danuta Rutkowska. A new version of the fuzzy-id3 algorithm. In *International Conference on Artificial Intelligence and Soft Computing*, pages 1060–1070. Springer, 2006.

[3] Tej Paul Bhatla, Vikram Prabhu, and Amit Dua. Understanding credit card frauds. *Cards business review*, 1(6), 2003.

[4] Richard J Bolton and David J Hand. Statistical fraud detection: A review. *Statistical Science*, pages 235–249, 2002.

[5] Priya Ranjan Dash. Your credit card under watch. *Financial Chronicle. Chennai*, page 16, April 17 2008.

[6] Linda Delamaire, HAH Abdou, and John Pointon. Credit card fraud and detection techniques: a review. *Banks and Bank Systems*, 4(2):57–68, 2009.

[7] Jose R Dorronsoro, Francisco Ginel, C Sgnchez, and CS Cruz. Neural fraud detection in credit card operations. *Neural Networks, IEEE Transactions on*, 8(4):827–834, 1997.

[8] Manoel Fernando Alonso Gadi, Xidi Wang, and Alair Pereira do Lago. Credit card fraud detection with artificial immune system. In *Artificial immune systems*, pages 119–131. Springer, 2008.

[9] Neda Soltani Halvaiee and Mohammad Kazem Akbari. A novel model for credit card fraud detection using artificial immune systems. *Applied Soft Computing*, 24:40–49, 2014.

[10] Divya Iyer, Arti Mohanpurkar, Sneha Janardhan, Dhanashree Rathod, and Amruta Sardeshmukh. Credit card fraud detection using hidden markov model. In *Information and Communication Technologies (WICT), 2011 World Congress on*, pages 1062–1066. IEEE, 2011.

[11] John Kiernan. Credit card and debit card fraud statistics. http://www.cardhub.com/edu/credit-debit-card-fraud-statistics/, 2013. [Online; accessed 12-June-2014].

[12] Sam Maes, Karl Tuyls, Bram Vanschoenwinkel, and Bernard Manderick. Credit card fraud detection using bayesian and neural networks. In *Proceedings of the 1st international naiso congress on neuro fuzzy technologies*, 2002.

[13] Mukesh Kumar Mishra and Rajashree Dash. A comparative study of chebyshev functional link artificial neural network, multi-layer perceptron and decision tree for credit card fraud detection. In *Information Technology (ICIT), 2014 International Conference on*, pages 228–233. IEEE, 2014.

[14] Sudhansu Kumar Mishra, Ganpati Panda, and Sukadev Meher. Chebyshev functional link artificial neural networks for denoising of image corrupted by salt and pepper noise. 2009.

[15] Reserve Bank of India. Bank-wise atm/pos/card statistics. http://www.rbi.org.in/scripts/ATMView.aspx, 2014. [Online; accessed 12-June-2014].

[16] Dominik Olszewski. Fraud detection using self-organizing map visualizing the user profiles. *Knowledge-Based Systems*, 70:324–334, 2014.

[17] Suvasini Panigrahi, Amlan Kundu, Shamik Sural, and Arun K Majumdar. Credit card fraud detection: A fusion approach using dempster–shafer theory and bayesian learning. *Information Fusion*, 10(4):354–363, 2009.

[18] Jagdish C Patra and Alex C Kot. Nonlinear dynamic system identification using chebyshev functional link artificial neural networks. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*, 32(4):505–511, 2002.

[19] J. Ross Quinlan. Induction of decision trees. *Machine learning*, 1(1):81–106, 1986.

[20] Y Sahin and E Duman. Detecting credit card fraud by decision trees and support vector machines. In *Proceedings of the International MultiConference of Engineers and Computer Scientists*, volume 1, pages 1–6, 2011.

[21] Yusuf Sahin, Serol Bulkan, and Ekrem Duman. A cost-sensitive decision tree approach for fraud detection. *Expert Systems with Applications*, 40(15):5916–5923, 2013.

[22] Abhinav Srivastava, Amlan Kundu, Shamik Sural, and Arun K Majumdar. Credit card fraud detection using hidden markov model. *Dependable and Secure Computing, IEEE Transactions on*, 5(1):37–48, 2008.

[23] Mubeena Syeda, Yan-Qing Zhang, and Yi Pan. Parallel granular neural networks for fast credit card fraud detection. In *Fuzzy Systems, 2002. FUZZ-IEEE'02. Proceedings of the 2002 IEEE International Conference on*, volume 1, pages 572–577. IEEE, 2002.

[24] M Umanol, Hirotaka Okamoto, Itsuo Hatono, HIROYUKI Tamura, Fumio Kawachi, Sukehisa Umedzu, and Junichi Kinoshita. Fuzzy decision trees by fuzzy id3 algorithm and its application to diagnosis systems. In *Fuzzy Systems, 1994. IEEE World Congress on Computational Intelligence., Proceedings of the Third IEEE Conference on*, pages 2113–2118. IEEE, 1994.

[25] ACI Universal Payments. Annual aci worldwide global fraud report finds one-in-four consumers victims of card fraud. http://www.aciworldwide.com/news-and-events/press-releases/ annual-aci-worldwide-global-fraud-report-finds-one-in-four-consumers-victims-of-card-fraud.aspx, October 16, 2012. [Online; accessed 12-June-2014].

[26] Wikipedia. Credit card fraud — Wikipedia, The Free Encyclopedia https://en.wikipedia.org/w/index.php?title=Credit_card_fraud&oldid=608901232, 2014. [Online; accessed 11-June-2014].

[27] Wen-Fang Yu and Na Wang. Research on credit card fraud detection model based on distance sum. In *Artificial Intelligence, 2009. JCAI'09. International Joint Conference on*, pages 353–356. IEEE, 2009.

[28] Lotfi A Zadeh. Fuzzy sets. *Information and control*, 8(3):338–353, 1965.