IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 11, November 2015

A Review On Credit Card Fraud Detection Using BLAST-SSAHA Method

Mr Yogesh M Narekar[1], Mr Sushil Kumar Chavan[2]

Department of Information Technology, RGCER, Nagpur, India[1]

Department of Information Technology, YCCE, Nagpur, India[2]

Abstract: In the current days, with, the usage of credit cards has increased radically due to its varied benefits. The mode of payment through credit card has made people's life easy for both online and ordinary purchases and thus widespread. This enormous usage of credit card leads to different frauds. Due to the rise and rapid growth of E-Commerce, use of credit cards for online purchases has dramatically increased and it caused an explosion in the credit card fraud. As credit card becomes the most popular mode of payment for both online as well as regular purchase, cases of fraud associated with it are also rising. In real life, fraudulent transactions are scattered with genuine transactions and simple pattern matching techniques are not often sufficient to detect those frauds accurately. This system seeks to investigate the current debate regarding the credit fraud in the banking sector and vulnerabilities in online banking and to study some possible remedial actions to detect and prevent credit fraud. The system reveals lots of channels of fraud in online banking which are increasing day by day. These kinds of fraud are the main barriers for the e-business in the banking sector. This system also gives the details of a survey of various techniques used in credit card fraud detection mechanisms and evaluates each methodology based on certain design criteria.

Keyword: Fraud detection, credit card, BLAST-SSAHA method, E-Commerce.

I. INTRODUCTION

A. Credit card

The typical credit card looks like a small rectangular card made of plastic which is issued to the users as a mode of payment. The credit-card holders are allowed to purchase any materials or service as long as the promises are kept by the credit card user. Security of the credit card relies upon its privacy of the card details i.e., credit card holder name and also its number. The credit cards can be used in two ways: a) physical usage and b) virtual/online usage. Physical usage is where an individual does use the credit card to pay for his purchases in any store personally and virtual or online usage is where the card owner uses the credit card to pay for purchased items online over the internet by just entering the required credit card details. With the increasing technology and the usage of internet for online shopping around the globe has brought a significant rise in the credit card usage in making transactions [1].

The credit card is a small plastic card issued to users as a System of payment. It allows its cardholder to buy goods and services based on the cardholder's promise to pay for these goods and services. Credit card security relies on the physical security of the plastic card as well as the Privacy of the credit card number. Globalization and Increased use of the internet for online shopping has Resulted in a considerable proliferation of credit card Transactions throughout the world. Thus a rapid growth in the number of credit card transactions has led to a Substantial rise in fraudulent activities. Credit card fraud is A wide-ranging term for theft and fraud committed using A credit card as a fraudulent source of funds in a given Transaction. Credit card fraudsters employ a large number of techniques to commit fraud. To combat the credit card Fraud effectively, it is important to first understand the mechanisms of identifying a credit card fraud. Over the Years credit card fraud has stabilized much due to various credit card fraud detection and prevention Mechanisms.

B. Credit card frauds

Credit card fraud can be divided into 2 types: inner card fraud and external card fraud. Inner card fraud intends to defraud the cash. Usually it is the collusion between merchants and cardholders, using false transactions to defraud banks cash. External card fraud is mainly embodied at using the stolen, fake or counterfeit credit card to consume, or using cards to get cash in disguised forms, such as buying the expensive, small volume commodities or the commodities that can easily be changed into cash. This paper is mainly devoted to the investigation of the external card fraud, which accounts for the majority of credit card frauds.

In general, a fraud is defined as a crime committed with intention to damage a person and is also a violation. Fraud may be committed for various reasons: for entertainment, to exploit a business / an organization, to take revenge, to cause financial loss, to damage identity and etc. Also there are several types of frauds: bankruptcy frauds, identity thefts, health frauds, religious frauds, credit card frauds, insurance frauds, forgery, tax frauds and many more. Here considering only the credit card frauds, they can be of two kinds: a) offline credit card frauds and b) online credit card frauds. Offline credit card frauds are those where an individual's credit card is lost or stolen. If any attacker or hacker, hack the details and use it to commit illegal actions is referred as online frauds. With the rapidly developing technology, usage of internet is drastically increasing. Substantially, this is leading to many credit-card fraudulent activities.

Under the vulnerable situations, the users must limit amount of information they are sharing to reduce the

exposure and chances of getting attacked by a hacker. Certain steps can be followed to protect themselves in case of vulnerabilities and insecurities regarding their identity: firstly, the users must take considerable care and attention to all their documents consisting of their identity, accounts and other sensitive information. Information sharing between friends and relatives is other reason from research that causes credit card frauds because of current lifestyle conditions. So, this has to be avoided and proper precaution must be taken. While the old documents or devices are disposed or discarded, users must make sure that there is no personal information that can be possibly revealed or tracked by anyone. Bank account information like transactions and balances should be regularly checked by the users to avoid money loss that could be caused by attackers.

Secondly, the use of computer systems and internet must be developed among the people. For example, just by installing anti-virus software is not sufficient enough to not get attacked by any malware. Without proper security across the network, the information over the internet must not be shared. Users should not access for their personal information from public wireless networks or others accounts and must carry a strong WPA key to their Wi-Fi access points. Thus, users need to look for security policies and keep their account protected from possible threats.

Basic security precautions like effective control over cookies, anonymous browsing, reducing computer information with-holdings, considering network address protection, using strong encrypting tools and changing passwords of all accounts often and also using passwords that are hard-to-crack can help the users from being attacked. Not just the users that need to take care of their personal credit card information; the respective banks and organizations that are issuing credit cards must also take initiative actions and work collectively considering two important things: firstly, they must bring out strict policies over public to secure genuine credit card holders and their privacy. And secondly, they must organize awareness about the secure way of using credit-cards and attacks possible among the public by prevention campaigns from several smaller targets to bigger population. If there are people who are already attacked by fraud persons, then those victims should take immediate actions related to situation. If a credit card related to a bank account of a person is hacked and fraud is detected by the victim, in order to protect his account from further financial destructions, he/she must contact the bank or bank's anti-fraud centre or credit card agent, requesting to block their card and monitor the details. Also then should report and register a complaint with the police department for investigation.

So, it is definitely a challenging task to battle against the damage caused for one's loss. The information that is shared over the internet is never completely deleted. This is also a possible reason for cyber attacks. Hence it is very important to contact the sources of information and request them to delete the information as after effects of any fraud is difficult to recover and time consuming for any victim.

C. History

Lately, the fraud detections with respect to credit cards have been involving in much of research interest. Various methods were put forward, some of which have special focus using data mining and neural networks [1]. In this pipeline, Ghosh and Reilly introduced a detection system that uses neural network. They have constructed this that is skilled at a huge sample of labeled credit card transactions. Any missing cards, either stolen or lost and frauds like application fraud, fake fraud, mail-order fraud, and non-received issue (NRI) frauds are sample cases of the credit card account transactions [4]. In the recent times, parallel granular neural networks (PGNNs) have been used by Syeda, for enhancing the speed of data mining and information finding methods in detection of credit card frauds. Proper system has been built for serving this purpose [5]. Then, Stolfo introduced a fraud detection system (FDS) for credit card frauds that used metal-learning skills to understand the forms of fraudulent credit card transactions. The metal-learning is a procedure where various techniques are combined and integrated with each other. In this manner, another model was proposed by Stolfo fraud and intrusion detection where in it was called as cost-based model. In this, they used Java agents for metal learning referred as JAM. JAM in the detection of credit-card frauds, is a distributed data-mining system [6] [1].

The terms like true positive-false positive (TP-FP) spread and precision are defined as a part of this procedure considering them as significant performance metrics [1]. Later Aleskerov presented a data-mining system, "card-watch", for fraud detection of credit-cards. This model uses neural networks where a neural learning component behaves as an interface to a range of commercial databases. Kim and Kim observed two main causes for the intricacy of detection in credit card frauds: skewed distribution of statistics and mix of genuine and fraudulent transactions. Considering this identification, the fraud density of actual transaction information is taken as a confidence parameter and the weighted fraud score is generated to lessen the number of misdetections [1].

Brause implemented a new approach to achieve high fraud reporting that employs superior data mining practices and neural network procedures. Then, a mutual design for fraud detection of credit-card transactions was introduced by Chiu and Tsai that incorporates the strategies of web services and data mining procedures. In this scenario, the banks that are involved share the information related to the frauds in a varied and circulated situations. In order to have a proper medium of data exchange several web services methods are used. Stolfo and Prodromidis introduced an agent-based technique that uses distributed learning for fraud detections in credit card operations. From the name "agent-based technique" only once can presume that this involves artificial intelligence.

In order to attain high precision, this procedure combines inductive learning and metal-learning procedures [1]. A game-theoretic method to detect credit card frauds has

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 11, November 2015

been lately proposed by Vatsa. In this model, communication between the attacker and the FDS will be as a multi-stage game played between two players wherein both try to exploit.

II. BACKGROUND AND RELATED WORK

Apart from the above mentioned credit card fraud detection techniques there are some of the recent credit card fraud detection techniques that gained attention. The following is the explanation in brief about each of these techniques:

A. Fusion of Dempster–Shafer theory and Bayesian learning

For credit card fraud detection, this approach is a cross technique that merges the results obtained from present and precedent behavior. Any credit card owner has certain spending pattern for his purchases online that are recorded in his transactions account. This credit card fraud detection system comprises of mainly four elements:

Rule-based filter: the doubt level of every transaction that is made is extracted depending on the variations from the normal form of spending patterns.

Dempster–Shafer adder: in this component, all the transactions that are doubtful obtained by rule-based filter are combined to form a primary belief.
Transaction history database: here all the values formed as a primary belief are combined to form on the whole belief by its theory.

Bayesian learner: here once after any transaction is believed to be suspicious, it is strengthened with fraudulent or weakened with genuine transaction.

This approach has high accuracy and also improves the fraud detection rate when compared with previous credit card detection techniques. The only issue with this mechanism is, it is very expensive and processing speed is less. The following fig 1- represents the architectural model of this mechanism.

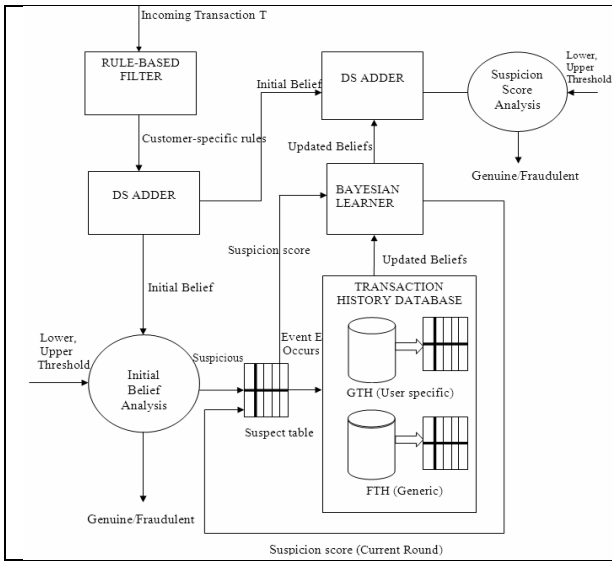

Fig. 1. Block diagram of the Fusion of Dempster–Shafer theory and Bayesian learning

B. BLAST-SSAHA in credit card fraud detection

BLAH-FDS algorithm is the improved form comprises of BLAST and SSAHA algorithm. These two algorithms are pretty much proficient sequence aligning algorithms in detecting credit card frauds. In the sequence alignment of BLAH-FDS algorithm, there are two stages where a profile analyzer obtains the correspondence between the transactions that are incoming in sequence with all the past and sequence of genuine transactions made in the past. The abnormal transactions detected by the profile analyzer are then passed into a deviation analyzer for checking with the past fraudulent transactions behavior if present. Thus based on these two analyzers a conclusion is drawn and final decision is taken. The performance of this mechanism in detecting the credit card frauds is good and its accuracy is high. Also processing speed is fast but the problem using this credit card fraud detection approach is that it cannot detect the duplicate transactions or cloned credit card frauds. The following fig 2 represents the architectural model of this mechanism.

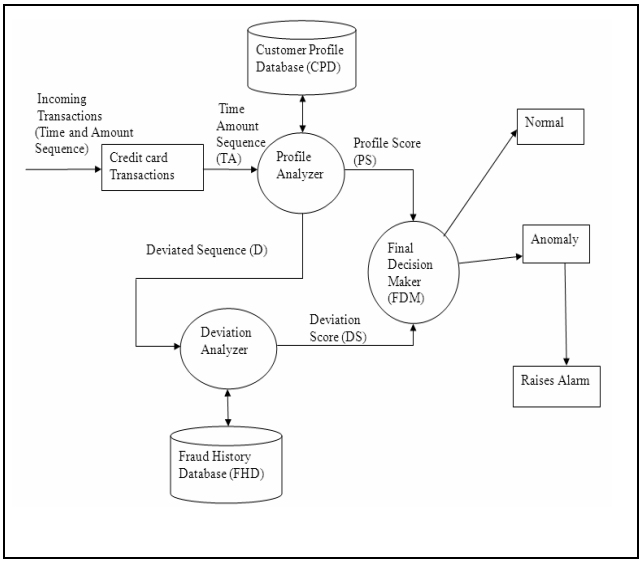

Fig. 2. Architecture of BLAST and SSAHA Fraud Detection System

C. Credit Card Fraud Detection using Hidden Markov Model (HMM)

Fig 3 shows the Flow of the HMM FDS. The hidden markov model mechanism in detecting credit card frauds is a double embedded stochastic process. Here if any of the incoming transactions are not accepted by the trained hidden markov model with sufficient probability then it is considered as a fraudulent transaction.

Use Of HMM For Credit Card Fraud Detection

In this model, fraud detection system at the beginning only is trained with the normal spending pattern of the credit card holder. For every transaction made, it is compared with the fraud detection system (FDS) for verification process. This FDS will take the card details, amount for the purchase made to identify whether the transaction is genuine or fraud transaction.

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 11, November 2015

If the transaction does not match with the information in FDS, it confirms it to be malicious and informs the bank that issued the credit card. Then the transaction is not passed further and will be declined by the bank. The credit card holder then will be contacted by the bank person and alerted about the issue. As in this mechanism already the FDS consists of the normal spending behavior of the credit card owner, it reduces the effort of bank. The main disadvantage here is false positives are high.

The following is the architectural model of the HMM in detecting credit card frauds where its work process is just mentioned. This is considered as the base model in the development of this project and is enhanced by using some security levels for better performance.

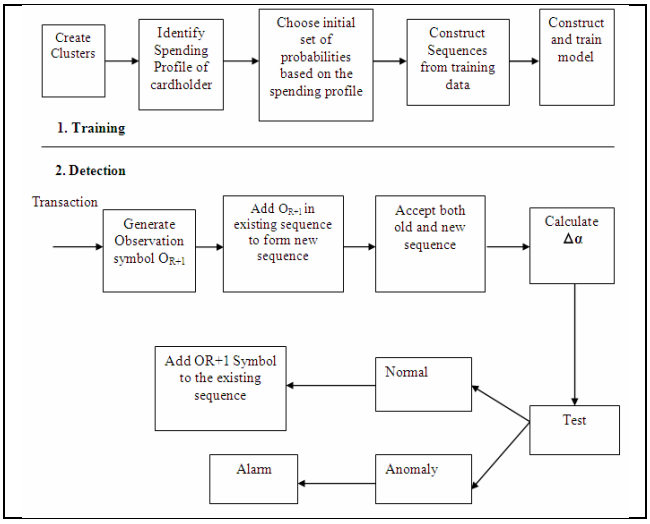

Fig. 3. Process Flow of the HMM FDS

Here, the system has two phases of processing: a) training phase and b) detection phase.

Training phase

In this phase FDS is trained with the normal spending pattern of the credit card holder. For every transaction made, it is compared with the fraud detection system (FDS) for verification process. This FDS will take the card details, amount for the purchase made to identify whether the transaction is genuine or fraud transaction. Fig 4 shows the process of transaction in training phase,

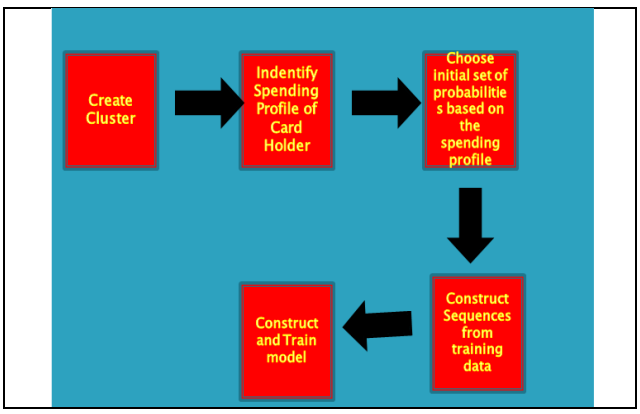

Fig. 4. Training phase

Detection phase

Fig 5 shows the detection phase of Fraud Detection System. FDS receives the card details and the value of purchase to verify whether the transaction is genuine or not. If the FDS confirms the transaction to be malicious, it raises an alarm and the issuing bank declines the transaction. The concerned cardholder may then be contacted and alerted about the possibility that the card is compromised.

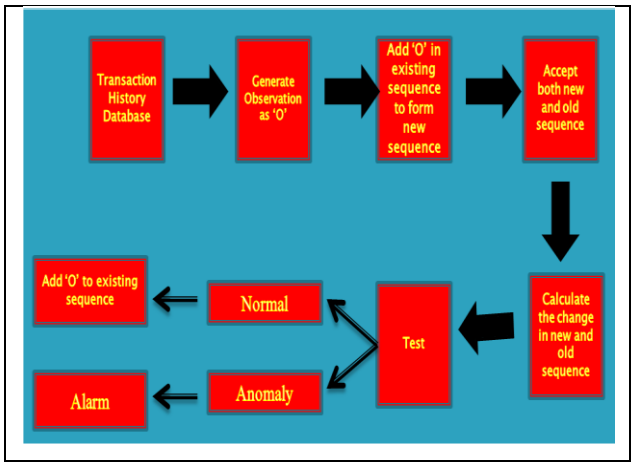

Fig. 5. Detection phase

D. Fuzzy Darwinian Detection of Credit Card Fraud

This detection method of credit card frauds uses genetic programming in order to develop some fuzzy logic rules that can be helpful in determining the suspicious and non-suspicious classes of transactions. When the information related to a transaction is provided to the FDS, the system using the classifiers, will determine the transaction as either safe or suspicious. The absolute system is capable of providing good accuracy rate and less false rate. This is not applicable in the case of online transactions practically as it is highly expensive. Figure-I.6 shows the Block diagram of the Evolutionary-fuzzy system**.** Also, its processing speed is very less. Fuzzy Darwinian Detection system uses genetic programming to evolve fuzzy logic rules capable of classifying credit card transactions into "suspicious" and "non-suspicious" classes. It 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 (GP) search algorithm and a fuzzy expert system.

Data is provided to the FDS system. The system first clusters the data into three groups namely low, medium and high. The GP, genotypes and phenotypes of the GP System consist of rules which match the incoming sequence with the past sequence. Genetic Programming is used to evolve a series of variable-length fuzzy rules which characterize the differences between classes of data held in a database. The system is being developed with the specific aim of insurance-fraud detection which involves the challenging task of classifying data into the categories: "safe" and "suspicious".When the customer's payment is not overdue or the number of overdue payment is less than three months, the transaction is considered as "non-suspicious", otherwise it is considered as "suspicious".

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 11, November 2015

The Fuzzy Darwinian detects suspicious and non–suspicious data and it easily detects stolen credit card frauds.

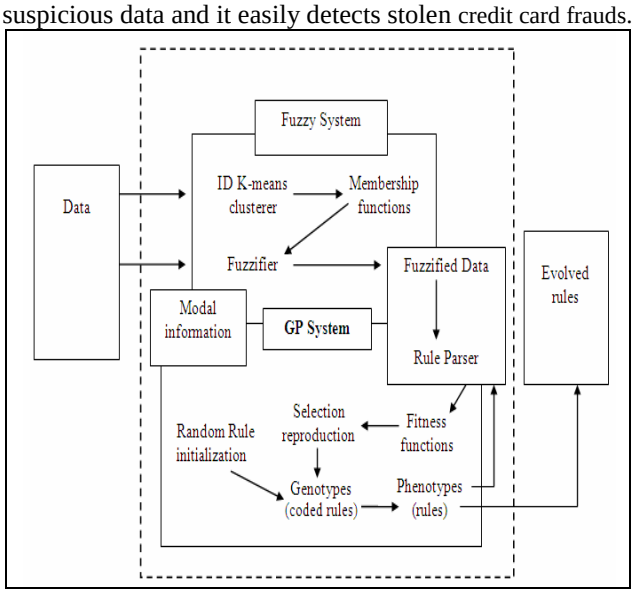

Fig. 6. Block diagram of the Evolutionary-fuzzy system

E. Credit Card Fraud Detection Using Bayesian and Neural Networks

This Bayesian and Neural Networks mechanism is an automatic credit card fraud detecting system using a machine learning approach. Both the Bayesian and Neural Networks approaches are suitable for analysis in cases of uncertainty.

In general, an artificial neural network will have an inter-connected group of artificial neurons and the pattern classification of frequently used neural networks. This is referred as feed-forward network and has three layers: input layer, hidden layer and an output layer. Here all the incoming transactions are passed through these three layers known as forward propagation. The artificial neural networks will hold the training information and compares with the inward transactions where the neural networks are originally fed with the common spend pattern and behavior of the card holder. Any suspicious transactions are sent back to the neural networks which classifies the transactions as safe and suspicious. This is where the Bayesian networks use artificial intelligence concept comprising of various methods like data mining and machine learning algorithms to bring out results.

The Bayesian belief networks are very efficient in cases like where small amount of data is known and incoming information is unstable or not completely available. This helps in data identification and classification and these neural networks need not be re-programmed. These provide good accuracy but require lot of training in achieving high processing speed.

F. Enhanced Hidden Markov Model Approach in detecting Credit Card Frauds

Enhanced Hidden Markov Model Approach **is** mainly focusing on three main constraints:

- **Accuracy:** It is defined as a portion of all the number of transactions which are identified correctly. That is the genuine transactions as genuine and fraud transactions as fraud.
- **True Positive (TP):** This gives the value of division of transactions detected correctly whether genuine or fraudulent. Here it counts only if genuine transactions are detected as genuine and fraud transactions as fraud and not any other.
- **False Positive (FP):** This gives the value of division of transactions detected wrongly whether genuine or fraudulent. Here it counts only if genuine transactions are detected as fraudulent and fraud transactions as genuine and not any other.

This paper introduces an enhanced form of recent and possible approach, “Hidden Markov Model (HMM)”- for credit card fraud detection, that does not need any fraud transactions information of the credit-card and still able to detect the fraud actions. This model considers the spending routines of the credit-card holder. In this, transactions of a credit card processing series are modeled by the stochastic procedure. The information related to the purchases with respect to an individual credit card holder is not identified by that particular bank’s FDS which issued the credit-card. This is the primary factor of Markov chain, which is signified but not noticeable. The credit card transactions can be viewed or known by stochastic process which gives the series of the spending information. Here many security levels are added into the application so that the account and transactions of purchases made can be more secure. Thus, this method is a best preference over various other techniques in addressing the issues. Also there is another major benefit by choosing this technique; it will result in radical decrease in the number of False Positives (FPs). The FPs is the transactions that are detected as fraud by the FDS, but in fact which are actual and genuine transactions [1]. Finally, this paper illustrates how the HMM technique is practically useful in detecting credit-card frauds and the results are presented.

The fig 7 represents this proposed project’s simple system architecture that includes different modules in the process of credit card fraud detection:

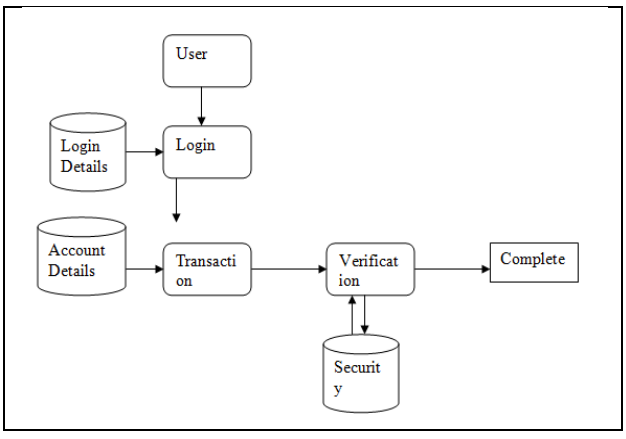

Fig. 7. Enhanced HMM Fraud Detection System Architecture

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 11, November 2015

Table-I: Comparison table of existing HMM and this proposed enhanced HMM
The following table-I represents a comparative analysis study of existing HMM and this proposed enhanced HMM is done and represented as comparison table considering several factors.

Parameter		Hidden Markov Model	Enhanced Hidden Markov Model
Methodology		Hidden Markov Model	Enhanced Hidden Markov Model
Fraud Dete-Ction	TP	70%	91%
	FP	20%	8.5%
Processing Speed		High	Very High
Training Required		Yes	Yes
Supervised Learning		Semi Supervised	Semi Supervised
Cost		Quite expensive	Inexpensive
Accuracy		Medium	High

TABLE I. COMPARISON TABLE OF EXISTING HMM AND THIS PROPOSED ENHANCED HMM

G. Evaluation of different credit card fraud detection systems

Constraints used
All the above discussed credit card fraud detection methods are evaluated and compared using the certain constraints: accuracy, methodology, True positive, false positive, cost and training required, supervised learning.

Accuracy: It is defined as a portion of all the number of transactions which are identified correctly. That is the genuine transactions as genuine and fraud transactions as fraud.

Methodology: This specifies the mechanism followed by the credit card FDS.

True Positive (TP): This gives the value of division of transactions detected correctly whether genuine or fraudulent. Here it counts only if genuine transactions are detected as genuine and fraud transactions as fraud and not any other.

False Positive (FP): This gives the value of division of transactions detected wrongly whether genuine or fraudulent. Here it counts only if genuine transactions are detected as fraudulent and fraud transactions as genuine and not any other.

Supervised Learning: This is where any system is fed with some information initially from the supervised data and known as machine learning task.

Table II-: Comparison of various credit card fraud detection mechanisms. All the mechanisms are having its own pros and cons. Results show that the fraud detection systems Fusion of Dempster and Bayesian theory, and Fuzzy Darwinian have very high accuracy in terms of TP. At the same time, the processing speed is fast enough to enable on-line detection of credit card fraud in case of BLAST-SSAHA Hidden Markov Model, and Bayesian and Neural Networks.

Parameter		Fusion of Dempster–Shafer theory and Bayesian learning	Hybridization of BLAST-SSAHA	Hidden Markov Model	Bayesian and Neural Networks	Fuzzy Darwinian Detection
Methodology		Machine Learning	Sequence Alignment	Hidden Markov Model	Artificial Intelligence, machine Learning	Genetic Programming, Fuzzy Logic
Fraud Dete-Ction	TP	98%	86%	70%	77%	100%
	FP	10%	10%	20%	10%	5.76%
Processing Speed		Medium	Very High	High	High	Low
Training Required		Yes	No	Yes	Yes	Yes
Supervised Learning		Supervised	Un- Supervised	Supervised	Supervised	Supervised
Cost		Expensive	Inexpensive	Quite expensive	Expensive	Highly Expensive
Accuracy		High	High	Medium	Medium	Very High

TABLE II. COMPARISON OF VARIOUS FRAUD DETECTION SYSTEMS PARAMETERS USED FOR COMPARISON

H. Rationale
In outlook of the reality that the conservative research for the credit card fraud detection has classically started in 984. Since then, the study and in depth research in this field of credit card fraud detection is ongoing. Every day with the emerging technologies, new problems are also arising which is making this study enduring. Initially some techniques that were proposed were failed to handle new kinds of frauds. Some techniques were complex to implement or expensive. Thus, the quest for an efficient, reliable and inexpensive technique became the main targets of researchers and led to the improvement in study and brought superior mechanisms which handled the problems to a good extent.
Currently, the bank organizations that are supported by the government are performing their research on these fraud detection mechanisms to know how they work and can be implemented by various methodologies. In the direction to improve the basic security for credit cards from the fraudulent people, this project helps in detecting the credit card frauds easily and also with high processing speed,

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 11, November 2015

good accuracy and that is inexpensive. Here the typical Hidden Markov Model in detecting credit card frauds is taken and improved with new additions that will overcome the problems that are observed with many other credit card fraud detection techniques.

The results obtained from this developed project are highlighted in this paper. This developed enhanced technique should be handling most of the current various kinds of credit card frauds. There is also scope of enhancing this project even more in future.

Table-III: Credit card verification database table

The following table-III represents the fields used in the process of Credit card verification of user when a transaction is to be passed. Each field has its own criteria and type of value to be specified. All these fields are to be filled in by the user when making the transactions. This information verifies the credit card user account with the details provided during registration and confirms if no suspicions.

Fields	Description
CardNo	Credit card number
AcNo	Number for login into account
Bankname	Specifies bank name
Customername	Username
Address	Place of User
Cardtype	Choose bank organization that issued the card
Transactionplace	Security answer
TransAmount	Give the amount of per transaction
DateTime	Specifies date and time

TABLE III. CREDIT CARD VERIFICATION DATABASE TABLE

III. SYSTEM ARCHITECTURE OF PROPOSED BLAH FDS

When a transaction is carried out on a given credit card Ci, the current value of transaction amount and time are quantized to generate one candidate element for amount sequence A and time sequence T. The current elements along with the last N - 1 elements form an observation window of length N. Two sequences (A and T) each having length N are then obtained from the observation window.

These two sequences are merged together to form a time-amount sequence TA of length N. If ti and ai are the ith elements of T and A, respectively, then ti is the ith element of TA. The merged sequence TA is aligned with the sequences related to the credit card Ci in CPD. This alignment process is done using BLAST. The k-tuple table CKT obtained from SSAHA algorithm is used by BLAST at this stage to improve the speed of the alignment process. However, due to the presence of some fraudulent transactions in TA, significant mismatches can occur in the alignment process. This mismatch produces a deviated sequence V which is aligned with FHD in a similar manner as done with TA and CPD. The deviated sequence could have been produced due to occasional change in spending behavior of the genuine cardholder. So, we align V with FHD to confirm whether the deviation is due to short-term change in cardholder's spending profile or due to the presence of some fraudulent transactions.

A high similarity between V and FHD confirms the presence of fraudulent transactions. PA evaluates a PS according to the similarity between TA and CPD. DA evaluates a DS according to the similarity between V and FHD. Effective total score is derived from PS and DS. The FDM finally raises an alarm if _ is below the alarm threshold (AT).

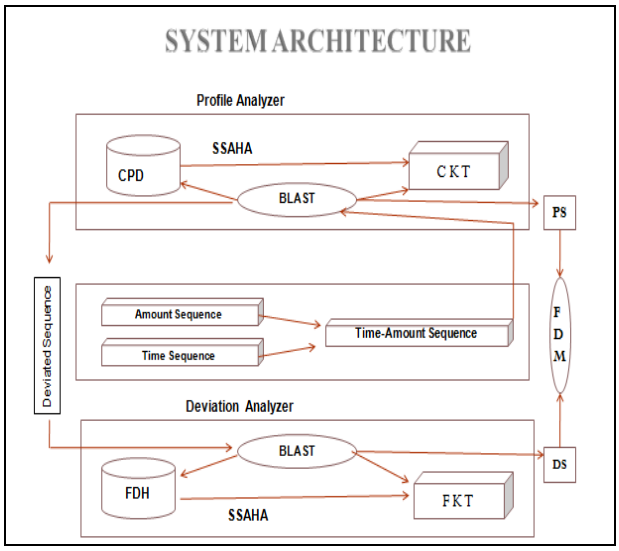

Fig. 8. Architecture of Proposed BLAHFDS

A. Profile Analyzer (PA): PA analyzes the similarity of time-amount sequence of the incoming transaction with cardholder's profile database (CPD).

B. Deviation Analyzer (DA): DA analyzes the similarity of the deviated time-amount sequence with company's fraud history database (FHD).

C. Amount Sequence (A): Amount sequence represents a sequence of transaction amounts associated with the last few transactions on that card.

D. Time Sequence (T): T represents a sequence of transaction times associated with the last few transactions on that card.

E. Time-amount Sequence (TA): TA is the merged sequence of T and A. k-tuple Table (KT): KT keeps sequence-index and sequence-offset information of history database. Cardholder's k-tuple information is kept in CKT and fraudster's K-tuple information is kept in FKT.

F. Profile Score (PS) and Deviation Score (DS): The Profile analyzer evaluates a similarity score between TA and CPD which is called PS. Similarly, deviation analyzer evaluates a similarity score between deviated sequence V and FHD which is called DS.

DOI 10.17148/IJARCCE.2015.41194

IJARCCE

ISSN (Online) 2278-1021
ISSN (Print) 2319 5940

International Journal of Advanced Research in Computer and Communication Engineering
Vol. 4, Issue 11, November 2015

G. Deviated Sequence (V): The sequence of elements which have some deviation from the cardholder's profile form a deviated sequence V.

H. Final Decision Maker (FDM): FDM takes the final decision about the nature of the transaction based on the PS and the DS.

IV. SYNTHETIC DATA GENERATION

It is difficult to obtain available credit card data sets since the security, privacy and cost issues. Currently, most of the researchers generate realistic synthetic data using data generator to facilitate the development and testing of data mining tools Synthetic data can be generated by using Genuine Markov Chain (GMC) And Fraud Markov Chain (FMC). Both GMC and FMC consist of Probability Distribution Function (PDF) and Transition Probability Matrix (TPM). GMC is used for generation of Genuine Historical database. FMC is used for generation of Fraudulent Historical database.

A. Genuine Markov Chain Block (GMCB)

This block consists of a finite Markov chain with associated transition probability matrix (TPM) and initial probability distribution vector (IPD). Here, the number of states is the same as the length of the amount dimension. This block generates synthetic transaction amounts for genuine cardholder. The values associated with TPM and IPD can be changed to capture the spending behavior of a cardholder properly.

B. Fraud Markov Chain Block (FMCB)

Similar to GMCB, this block is used to generate synthetic fraud transaction amounts. The values associated with TPM and IPD can also be changed in order to capture changing behavior of the fraudster.

C. Creation of Profile Score

- The Profile analyzer evaluates a similarity score between TA and CPD which is called Profile Score (PS).
- It can be calculated as **PS= N * $^{\delta}$ - (L - N) * $_{\gamma}$,**
  $^{\delta}$ =Unit match score,
  $_{\gamma}$ = Unit mismatch score,
  L= length of TA,
  N=Number of matches with aligned CPD sequence

D. Creation of Deviation Score

- The Deviation analyzer evaluates a similarity score between deviated sequence V and FHD which is called **Deviation Score (DS).**
- It can be calculated as **DS= M * $^{\delta}$ - ((L - N) – M) * $_{\gamma}$,**
  $^{\delta}$ =Unit match score,
  $_{\gamma}$ = Unit mismatch score,
  L= length of TA,
  N= Number of matches with aligned CPD sequence
  M= Number of matches with aligned FHD sequence

E. Creation of Total Score

TS= PS – DS

V. TESTING

- Testing is a process of executing a program with intent of finding an error.
- Testing presents an interesting anomaly for the software engineering.
- The goal of the software testing is to convince system developer and customers that the software is good enough for operational use. Testing is a process intended to build confidence in the software.
- Testing is a set of activities that can be planned in advance and conducted systematically.
- Testing is a set of activities that can be planned in advance and conducted systematically.
- Software testing is often referred to as verification & validation.

A. Types of testing

The various types of testing are:

- White Box Testing
- Black Box Testing
- Unit Testing
- Integration Testing
- Validation Testing
- Output Testing
- User Acceptance Testing

White box testing

- It is also called as glass-box testing. It is a test case design method that uses the control structure of the procedural design to derive test cases.
- Using white box testing methods, the software engineer can derive test cases that:
- Guarantee that all independent parts within a module have been exercised at least once.

Black box testing

- It's also called as behavioral testing. It focuses on the functional requirements of the software.
- It is complementary approach that is likely to uncover a different class of errors than white box errors.
- A black box testing enables a software engineering to derive a set of input conditions that will fully exercise all functional requirements for a program.

Unit testing

In this testing we test each module individually and integrate with the overall system. Unit testing focuses verification efforts on the smallest unit of software design in the module. This is also known as module testing.
The module of the system is tested separately. This testing is carried out during programming stage itself. In this testing step each module is found to working satisfactorily as regard to the expected output from the module. There are some validation checks for fields also. It is very easy to find error debut in the system.

Integration testing

Data can be lost across an interface; one module can have an adverse effort on the other sub functions when combined may not produces the desired major functions. Integrated testing is the systematic testing for constructing the uncover errors within the interface. The testing was done with sample data. The Developed system has run

successfully for this sample data. The need for integrated test is to find the overall system performance.

Validation testing

At the culmination of the black box testing, software is completely assembled as a package, interfacing errors have been uncovered and corrected and a final series of software tests. That is, validation tests begin, validation testing can be defined many ways but a simple definition is that validation succeeds when the software functions in manner that can be reasonably expected be the customer. After validation test has been conducted one of the two possible conditions exists. The functions or performance characteristics confirm to specification and are accepted.

Output testing

After performance of the validation testing, the next step is output testing of the proposed system since no system could be useful if it does not produce the required output in the specific format. Asking the user about the format required by system tests the output displayed or generated by the system under consideration.

Here the output format is considered the of screen display. The output format on the screen is found to be correct as the format was designed in the system phase according to the user need. For the hard copy also the output comes out as specified by the user. Hence the output testing does not result in any correction in the system.

User acceptance testing

Some of my friends were who tested this module suggested that this was really a user friendly application and giving good processing speed.

B. Testing used in this project:

- Unit Testing
- Integration Testing
- Validation Testing
- Output Testing
- User Acceptance Testing

VI. CONCLUSION

It is difficult to obtain available credit card data sets since the security, privacy and cost issues. Currently, most of the researchers generate realistic synthetic data using data generator to facilitate the development and testing of data mining tools. Data generation and simulation is required.

Efficient credit card fraud detection system is an utmost requirement for any card issuing bank. We have proposed an algorithm named as BLAHFDS and used it for credit card fraud detection. The system named as BLAHFDS identifies fraudulent transactions using a Profile Analyzer and a Deviation Analyzer. These two analyzers use BLAH as a sequence alignment tool to detect fraud. A stochastic model has been proposed for the generation of synthetic transactions to analyze the performance of BLAHFDS. Results show that the new approach has high accuracy. At the same time, the processing speed is fast enough to enable online detection of credit card fraud.

REFERENCES

[1]. Casey, E. "Network traffic as a source of evidence: tool strengths, weaknesses, and future needs". Digital Investigation, vol. 1, no. 1, pp. 8-43, 2004b.

[2]. Srivastava, A. Kundu, S. Sural, A.K. Majumdar, "Credit Card Fraud Detection Using Hidden Markov Model", IEEE transactions on Dependable and Secure Computing, Kharagpur, India, March 2008.

[3]. S. B. Edwin Raj, A. A. Portia, "Analysis on Credit Card Fraud Detection Methods", International Conference on Computer, Communication and Electrical Technology – ICCCET 2011.

[4]. Kundu, S. Panigrahi, S. Sural, "BLAST-SSAHA Hybridization for Credit Card Fraud Detection", IEEE transactions on dependable and Secure Computing, vol. 6, no. 4, October-December 2009.

[5]. S. Ghosh, D.L. Reilly, "Credit Card Fraud Detection with a Neural-Network," Proc. 27th Hawaii Int'l Conf. System Sciences: Information Systems: Decision Support and Knowledge-Based Systems, vol. 3, pp. 621-630, 1994.

[6]. M. Syeda, Y.Q. Zhang, Y. Pan, "Parallel Granular Networks for Fast Credit Card Fraud Detection," Proc. IEEE Int'l Conf. Fuzzy Systems, pp. 572-577, 2002.

[7]. S. Stolfo, A.L. Prodromidis, "Agent-Based Distributed Learning Applied to Fraud Detection," Technical Report CUCS-014-99, Columbia Univ., 1999.

[8]. Benson Edwin Raj and Annie Portia " Analysis on Credit Card Fraud Detection Methods" in International Conference on Computer, Communication and Electrical Technology – ICCCET [2011].

[9]. Aihua Shen, Rencheng Tong and Yaochen Deng, "Application of Classification Models on Credit Card Fraud Detection"- IEEE [2007].

[10]. Amlan Kundu, Suvasini Panigrahi and Shamik Sural, "BLAST-SSAHA Hybridization for Credit Card Fraud Detection" IEEE TRANSACTIONS ON DEPENDABLE AND SECURE COMPUTING, VOL. 6, NO. 4, [2009].

[11]. Sherly K.K and R Nedunchezhian, "BOAT ADAPTIVE CREDIT CARD FRAUD DETECTION SYSTEM",- IEEE [2010].

[12]. Divya Iyer, Arti Mohanpurkar, Sneha Janardhan, Dhanashree Rathod and Amruta Sardeshmukh, "CREDIT CARD FRAUD DETECTION USING HIDDEN MARKOV MODEL", - IEEE [2010].

[13]. Delwar Hussain Mahdi, Karim Mohammed Rezaul and Muhammad Azizur Rahman, "Credit Fraud Detection in the Banking Sector in UK: A Focus on E-Business", Fourth International Conference on Digital Society, 2010.

[14]. Y. Sahin and E. Duman, "Detecting Credit Card Fraud by ANN and Logistic Regression", - IEEE [2011].

[15]. M. Hamdi Ozelik, Mine Isik and Ekrem Duman, "Improving a credit card fraud detection system using genetic algorithm", International Conference on Networking and Information Technology, 2010.

[16]. Richard, "Credit Cards as Spending Facilitating Stimuli: A Conditioning Interpretation," J. Consumer Research, vol. 13, no. 3, pp. 348-356, 1986.

[17]. S. Ghosh and D.L. Reilly, "Credit Card Fraud Detection with a Neural-Network," Proc. Int'l Conf. System Science, pp. 621-630, 1994.

[18]. E. Aleskerov, B. Freisleben, and B. Rao, "CARDWATCH: A Neural Network Based Database Mining System for Credit Card Fraud Detection," Proc. IEEE/IAFE Conf. Computational Intelligence for Financial Eng. (CIFEr), pp. 220-226, 1997.

[19]. R. Brause, T. Langsdorf, and M. Hepp, "Neural Data Mining for Credit Card Fraud Detection," Proc. Int'l Conf. Tools with Artificial Intelligence, pp. 103-106, 1999.

[20]. M. Syeda, Y.Q. Zhang, and Y. Pan, "Parallel Granular Neural Networks for Fast Credit Card Fraud Detection," Proc. IEEE Int'l Conf. Fuzzy Systems, pp. 572-577, 2002.