A Survey on Fraud Detection Techniques in Ecommerce

Priya J Rana
Dept. of Information Technology
Parul Institute of Engineering and Technology,
Gujarat,
India

Jwalant Baria
Dept. of Computer Science and Engineering
Parul Institute of Engineering and Technology,
Gujarat,
India

ABSTRACT
E-commerce becomes popular for online shopping, banking, financial institution and government. Fraudulent activity exits in many areas of businesses and our daily life. Such activities are most prevalent in telecommunication, credit card fraud detection, network intrusion, finance and insurance and scientific applications, Billions of dollars are loss every day due to the increase in the number of credit card transactions by using online as well as offline. To design potent and efficient fraud detection algorithms is the key for reducing the losses in transaction. There were numerous approaches have been implemented for the fraud detection. This paper shows the approaches used in fraud detection in e-commerce.

General Terms
Data Mining Techniques, Dempster Shafer Theory & Bayesian Learning, Rule-based Filter, Sequence alignment, Clustering Technique.

Keywords
Internet, E-commerce, Fraud Techniques, Detection Method.

1. INTRODUCTION
Now in the technology of days due to speedy development internet usage is everywhere. In today's evolution electronic world, many small and large companies have placed their businesses on to the WWW to provide services to customer. E-commerce draws on technologies such as electronic fund transfer, online transaction processing, internet banking, and automated data collection systems and so on. Online shopping is going to be popular day by day. E-commerce payment systems have become popular due to widespread use of the internet-based shopping and banking. Rapid increment of this era billions of dollars are lost every year due to credit card fraud. Fraud is an act of betrayal intended for personal usage or to harm a loss to someone. Fraudster only wants to know the personal information related to card (card number, card expiry date etc.). It can be possible physically or virtually. It is commonly understood as dishonesty to gain some advantage which is often financial, over another person. It can be seen in most common, acquiring or trading of property, including real property, Personal Property, and intangible property, such as stocks, bonds, and copyrights.

1.1 Problem Definition
An e-commerce cash system provides proficiency for online transaction to become a universal business. E-commerce frauds can be possible as offline as well as online in the shopping trend. In the online shopping transaction, fraudster wants to harass to Merchant or to the bank by doing the frauds like this. Fraud can be like,

1. Fraudster do not want to purchase anything from the shopping cart but still he/she giving the wrong information & make payment transaction as a Cash on delivery to harm to the merchant.
2. If credit/debit card information stolen or lost then by using the credit card number and cvv number fraudster can be make payment easily without knowing to the actual user.
3. If the original database is hacked from bank or e-commerce database by fraudster in which having all the information related to card is stored then in the absence of the credit/debit card fraud can be possible.

1.2 Types of Frauds
There are numerous types of fraud on the e-commerce nature that introduce telecommunication frauds, Credit card frauds, computer intrusion, Bankruptcy fraud, Theft fraud/counterfeit fraud, Application fraud and Behavioural fraud [5].

Credit card Fraud: Credit card fraud has been split into two types: Offline fraud and On-line fraud.

- ***Offline Fraud*** is because of by using a stolen physical card at any place.
- ***Online Fraud*** is because of via internet, phone, shopping or web.

Telecommunication Fraud: Telecommunication fraud involves the misuse, who has intention to harm someone by mobile phone fraud and fixed line fraud.

Computer Intrusion: The act entering without invitation by any outsider or hacker and insider who knows the system structure any Environment is defines a intrusion. That Means "Potential Possibility of Unauthorized Attempt to Access Information, Manipulate Information Purposefully".

Bankruptcy Fraud: Bankruptcy fraud is difficult to predict the fraud, when bank send a customer an order, user will be of personal bankruptcy fraud. Also it becomes difficult to get unwanted loans. One of the potential ways to avoid this fraud is by doing a pre-check with credit bureau. This informed about the past banking history of its customers.

Theft Fraud/Counterfeit Fraud: If the used cards are not yours then the fraud is called theft. The bank will take measures to check the thief when the owner give some feedback and contact the bank. Likewise, remotely use of the credit card leads to counterfeit fraud .Use of your codes via various web-sites and copied card number, where no physical cards or signature are required.

Application Fraud: Application fraud is defined as when someone applies for a credit card with wrong information. Detection can be done by either when applications comes from a same user with the same details, termed as duplicates and when applications come from different individuals with similar details, called as identity fraudsters.

Behavioral Fraud: When orders are made on 'cardholder present' details and details of legitimate cards have been obtained fraud way basis, such a fraud I known as behavioral fraud.

2. CREDIT CARD FRAUD TECHNIQUES

There are many types of way in which fraudsters do their fraud activity on the internet; a study was conducted on how credit-card information is stolen. Here are some of the different techniques which are used for credit-card fraud information theft [10].

i. **Credit-card fraud generator software:** Software use to obtain authentic credit-card numbers and expiry dates. Some of these designed software are proficient in generating valid credit card numbers like credit-card companies or issuers by reason of it uses the mathematical Luhn algorithm that credit card companies or issuers are using to generate credit card numbers to their credit-card consumers or users. Black hat hackers hack credit card information which is in database which displays to cyber credit fraudsters by the use of software. This technique in some cases is used by black-hat hackers to sell their hacked credit card information to other online credit card fraudsters with little or no computer skills.

ii. **Key logger and Sniffers:** The Black hat hackers who have professional Programming or computer skills infect a computer by installing and automatically running sniffers or key logger computer programs by which they log all the keyboard inputs activities made by computer on a file with the personal gain intention of retrieving personal information like credit card information, etc. Spam mail are used to affect the user's computers by fraudsters to computer users & asking them to download free games or software, and when those mail are opened and downloaded, the sniffers of key loggers are downloaded automatically by itself, installed and ran on the user's computers. While the sniffer is runs on the user's computer, they log all the keyboard inputs used by the user over a network. Therefore, any user can unknowingly share their private authentic information through this infectious software. Sometimes this software are also shared or sold to other fraudsters who do not have computer knowledge or skills.

iii. **Site-cloning, Spyware and Merchant sites:** A Black-hat hacker, software is also created by the hackers, installed and ran a tracked on user's computer of the website activities. By tracking and knowing the website activities of the user on the internet, they clone the electronic or banking websites which are regularly visited by the user and send the user for using it with the intension of retrieving private or personal information. Second case is with fake merchant sites, it creates on which cheap products are provided to users and thereby asking user for payment by credit cards. The user credit car is stolen when the payment is made from fake site.

iv. **Physically stolen credit-card information:** Online buying and selling of the product one by stealing of the credit card is termed under the physically stolen credit card information.

v. **CC/CVV2 shopping websites:** Without professional computer skills, cyber credit-card fraudsters who have no buy hacked credit-card information on these websites to be used for fraudulent electronic payment for some goods and services on the internet.

3. METHODS USED FOR CREDIT CARD FRAUD DETECTION

With Doing survey of Fraud detection we have found the numerous types of Approaches are used. Those are described given below.

3.1 A Fusion Approach Using Dempster-Shafer Theory and Bayesian Learning

These system (DST & BL) combines three approaches rule based filter, Dempster-Shafer adder and Bayesian learner. It detects based on combination of current as well as past behavior together have recorded a profile for every card holder.

Rule based filter having a customer & generic specific rules deviates transaction's behavior from the normal profile of the card holder. Rules consist like average of daily/monthly spending profile; shipping address is being different from billing address etc. Bayesian learner is a tool which updates the database when new information becomes available [7].

It has advantages like high accuracy, processing speed, reduces false alarm, improves detection rate, applicable in E-commerce. One disadvantage of this approach is that it is highly expensive.

3.2 Blast-Ssaha Hybridization

Blast-Ssaha Hybridization is a two-stage sequence alignment. It has two analyzer as name is profile analyzer and Deviation analyzer. With the incoming transaction profile analyzer match the sequence with the card holder database. If there is unusual sequence is found then observed unusual data is passing to the deviation analyzer. Deviation analyzer compare the incoming unusual data with the past fraud history database. Final decision maker gives the result that transaction is Genuine or Fraud. If the transaction knows to suspicious then system will raise alarm and giving alert message. It is useful in telecommunication and banking fraud detection and processing speed is also high. The limitation of this system is it does not detect the cloning cards [7].

3.3 Hidden Markov Model

Hidden markov model (HMM) is a finite set of states having a probability distribution. Set of probabilities administrates the transition states. Each state having an amount probabilities assigned to each card holder. Amount of incoming transaction is compared to the predefined threshold value to define transaction is legitimate or Fraudulent. If any of the incoming transaction is unusual or not accepted by the trained HMM with sufficient probability then it is announced as a fraud transaction. Otherwise it will allow as a normal genuine transaction.

This system has two phase of processing: a) Training Phase and b) Detection Phase [6][9]. In the training phase HMM

analyze user amount spending profile. Three types of profile is possible that can be.

1) Low amount spending profile
2) Medium amount spending profile
3) High amount spending profile

All the usage cardholder has a different spending profile behavior. Based on that data is trained and according to that trained transaction database fraud will detect the transaction. Baum Welch algorithm is used for training purpose and K-mean algorithm for training

3.4 Fuzzy Darwinian Detection

Fuzzy Darwinian Detection (FDD) is an evolutionary-Fuzzy system detects based on fuzzy logic rules which using Genetic programming evolving fuzzy logic rules. It classifies the transactions into suspicious and non-suspicious. It comprises of Genetic Programming (GP) search algorithm and also a fuzzy expert system together. This approach has very high accuracy and produces a low false alarm. But it is not applicable in online transactions. Also it is highly expensive and processing speed is low [7].

3.5 Bayesian and Neural Network

Bayesian and Neural Network (BNN) is a automatic fraud detection system and it detects based on machine learning approaches for the artificial intelligence programming. The advantage of neural network is that it learns and does not need to be reprogrammed. Its processing speed is higher than Bayesian neural networks but it needs high processing time for large neural networks. Whereas Bayesian neural networks provide good accuracy but needs training of data to operate and requires high processing speed [7].

And neural network also implement Pattern Recognition for the fraud detection [12].

The another approaches are also used for the fraud detection such as Decision Tree, Genetic Algorithm, Artificial Neural Network, K- nearest neighbor algorithm, Stream Outlier Detection based on Reverse K-Nearest Neighbors(SODRNN), Fuzzy Logic Based System, Fuzzy Expert System, Support Vector Machine and Meta Learning Strategy [8][10]. All having a different rule sets for the detection approach.

4. PARAMETER COMPARISION

In terms of Parameter like Accuracy, True Positive (TP), False Negative (FP), and cost, the comparison of mentioned approaches based on survey.

Table 1. A Summary based on Parameter [4]

Parameter	Cost	Fraud Detection		Accuracy
		TP	FP	
Dempster Shafer Theory & Bayesian Learning	Expensive	98%	10%	High
Hybridization of BLAST-SSAHA	Inexpensive	86%	10%	High
Hidden Markov Model	Quite Expensive	70%	20%	Medium
Bayesian and Neural Network	Expensive	77%	10%	Medium
Fuzzy Darwinian Detection	Highly Expensive	100%	5.76%	Very High

5. CONCLUSION

For the fraud detection in e-commerce various approaches are there. In this paper, we have reviewed some of the detection approaches. Each approach having its own rule sets to implement and rules are not clearly described in approach. Based on observation table we can conclude that Hybridization of BLAST-SSAHA approach is best suitable for the fraud detection in terms of cost and accuracy. To detect a fraud is necessity but also to decrease false alarm is also necessary. BLAST-SSAHA's True positive (TP) ratio having less than Dempster Shafer theory and Fuzzy Darwinian detection but cost of both approaches is quite expensive. So it would not beneficial to implement both together. So by implementing rules which are used in another approach or implementing advanced rule sets into the BLAST-SSAHA. So there is possibility to increase True Positive result and decrease false alarm.

6. ACKNOWLEDGMENTS

This work is supported and guided by my research guide. I am very thankful to my research guide Mr. Jwalant Baria, Assistant Professor, CSE Department, Parul Institute of Engineering and Technology, Gujarat, India for supporting me.

7. REFERENCES

[1] W. Roberds, The impact of fraud on new methods of retail payment, Federal Reserve Bank of Atlanta Economic Review, First Quarter (1998) 42-52.

[2] Staticties for General and Online Card Fraud, 20 June, 2007. <http://epaynews.com/statistics/fraud.html>.

[3] Online fraud is 12 times higher than offline fraud, 20 June 2007. <http://sellitontheweb.com/ezine/news0434.shtml>.

[4] Sravani Pedamallu, "*Graduate Project report on Implementation of an Enhanced Hidden Markov Model in Detecting Credit Card Frauds*", 2011.

[5] Khyati Chaudry, Jyoti Yadav, Bhawna Malick, "*A Review of Fraud detection techniques"*, 2012 IJCA.

[6] Avinash Hingole, Dr. R. C. Thool, "Credit card fraud detection system Using Hidden Markov Model and its Performance," 2013 IJARCSSE.

[7] Divya Iyer, Arti Mohanpurkar, Sneha Janardhan, Dhanshree Rathod, Amruta Sardeshmukh, "*Credit Card Fraud Detection using Hidden Markov Model*," 2011 IEEE.

[8] Krishna Kumar Tripathi, Mahesh A. Pavaskar, "*Survey on credit card fraud detection methods*", 2012 IJETAE.

[9] Sailesh S. Dhok, "Credit Card Fraud Detection using Hidden Markov Model". 2012 IJSCE.

[10] Neha Sethi, Anju Gera, "A Revived survey on Various Credit Card Fraud Detection Techniques", 2014 IJCSMC.

[11] Raghvendra Patidar, Lokesh Sharma, "*Credit Card Fraud Detection Using Neural Network*", 2011 IJSCE.

[12] John Akhilomen, "Data mining Application for Cyber Credit-Card Fraud Detection System", 2013 WCE.