INTERNATIONAL JOURNAL OF RESEARCH IN COMPUTER APPLICATIONS AND ROBOTICS
ISSN 2320-7345

ANALYSIS ON CREDIT CARD FRAUD DETECTION TECHNIQUES BY DATA MINING AND BIG DATA APPROACH

[1]N.Malini, [2]Dr.M.Pushpa

[1]M.Phil. Student, PG & Research Department of Computer Science
Quaid-E-Millath Government College for Women (A), Annasalai, Chennai – 600002, Tamilnadu, India
maliniresearch16@gmail.com
[2]Assistant Professor, PG & Research Department of Computer Science
Quaid-E-Millath Government College for Women (A), Annasalai, Chennai – 600002, Tamilnadu, India
push_surya@yahoo.co.in

Abstract: - In this digital world, Credit card is one of the most divisive products among all the available financial tools. Credit card becomes popular mode of payment for both online as well as offline purchase. Besides being convenient, there are credit card frauds which become main threat for the users. Credit card frauds are increasing day by day. The fraudulent transactions are scattered with genuine transactions. Hence the simple pattern matching techniques are often insufficient to detect these frauds accurately. Credit card fraudsters are becoming more sophisticated. So, we need to develop /invent few new techniques to combat and to prevent such fraudulent attack .Some of the new techniques to detect the credit card frauds are Artificial Intelligence, Neural Networks, Machine Learning, Data Mining, Genetic Programming, Big Data Analytics etc., Big Data has brought fraud detection and prevention techniques such as fraud behavioral analysis and real-time fraud detection to give fraud fighting techniques a new perspective. The main objective of this paper is to identify the different types of credit card frauds involves in physical or virtual cards. Then to review on big data analytical techniques that detect credit card frauds and finally to study how we can safeguard the credit card and some precautions to avoid credit card frauds.

Keywords: Big data, Credit card, Fraud Detection Techniques, Prevention, Hadoop, Data mining

I. INTRODUCTION

Credit card payment becomes one of the famous elements in a technology world. The credit card payments reduce the complexity of payment system by eliminating physical paper in use like cash or cheque. Similarly the transaction of per card per month has grown from 1.3% to 2.7% during the year 2010 to 2014.That clearly states that transaction value has almost doubled in the four year-period by 153%[13].

The growth in credit card transaction also attracts the attention of fraudsters. Fraudsters are becoming more proficient and artistic to innovating new methods to commit the fraud regularly. Thus fast fraud detection and remediation are important for maintaining good relationship between the bank and customer. Different

algorithms are used to determine the probability of fraud by analyzing purchasing habits and comparing each transaction with what preceded it. Big data analytics is one of the best techniques which can show relationships among fraudulent activities, including several doubtful or distrust activities in a single account or patterns of similar activities in different accounts. It helps to provide analysis result faster and accurately. This paper discusses about the different types of credit card frauds which is classified based on activity of physical or virtual card, to detect the fraudulent credit card transaction with the help of big data analytics. As prevention is the best way to deal, it finally discuss about how to protect ourselves from credit card frauds.

A. CREDIT CARD FRAUDS

Credit card is convenient and substituted for cash, and it is also convenient method of payment**.** It is preapproved credit amount that can be used for purchasing goods and services, payment of that purchase is collected later with agreed charges. The credit cards credit limits various based upon individual perceived credit worthiness and it is the maximum amount loaned; credit worthiness is an individual ability and willingness to pay money back. *Credit card fraud* is a situation when an individual uses someone else's credit card information to charge purchases, or removing funds for personal reasons from the account without owner's authorization.

According to CEO of Rippleshot Cahn, who have spent over 15 years' experience in credit card fraud detection says that *30-40 % of credit card fraud loss can be reduced by early detection* [12].

Credit card fraudsters are committed in the following ways,

- Theft of actual cards,
- Misrepresentation of account or personal information,
- Illegal or unauthorized use of account for personal gain,
- An act of criminal deception by use of unauthorized account or personal information.

Credit card fraud can occurs online as well as offline.

- When unauthorized users make use of credit card with the PIN is called online fraud. Using physical card for transactions eg.resturants,buying electronic goods,etc.,
- When an unauthorized user make use of credit card without the PIN is called offline fraud transaction, eg.Through shopping websites, phone transactions etc.,

B. TYPES OF CREDIT CARD FRAUDS

The following figure 1.1 describes the various types of credit card frauds

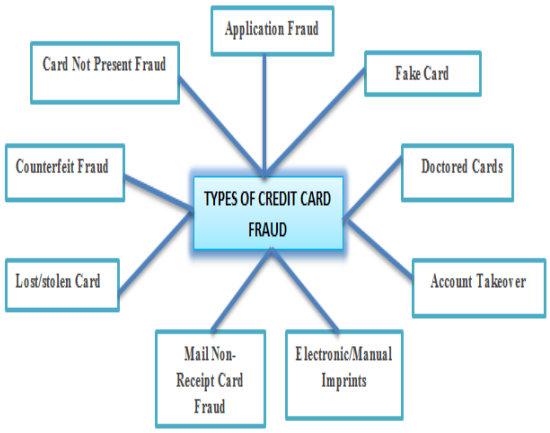

Fig. 1.1 Types Of Credit Card Frauds

Based on the performance of credit card activity, it can be classified into physical card and virtual card.

Physical card: In this type, the cardholder presents his/her card physically to a merchant for making a payment. To carry out fraudulent transactions in this kind of purchase, an attacker has to steal the credit card. Physical credit card fraud includes Doctored Card fraud, Fake cards fraud, Counterfeit card fraud, Lost/Stolen card fraud.
Virtual card: This type is a kind of purchase, only some important information about a card such as card number, expiration date, secure code and etc, is required to make the payment. Virtual card fraud includes Application fraud, Card Not Present fraud,
Electronic or Manual Credit Card Imprints, Mail Non-Receipt Card fraud.

a) Application Fraud*:* Application fraud generally happens in co-occurrence with identity theft. It happens when a fraudster uses another person name and information to apply for credit or a new credit card. They will usually first steal supporting documents, which are then used to authenticate their fraudulent application. Banks have various safeguarding plans and actions in place to stop this type of fraud. The most important one is requiring appropriate and original documentation only. Additionally, they will often call the employers in telephone to confirm identity. Unfortunately, fraudsters will often forge documents and provide false contact numbers for places of employment. Eg., utility bills, bank statements are used to open fake accounts. Currently account application fraud increased by 60% from August 2013 to August 2014[14].

b) Electronic or Manual Credit Card Imprints: Another form of credit card fraud is experienced through credit card imprints when a single transaction is recorded multiple times on old-fashioned credit card imprint machines known as *knuckle busters*[15].It generally happens when somebody skims information through magnetic strip which is placed on the card. This is used to create a fake card or to complete fraudulent transactions easily**.**

c) CNP (Card Not Present) Fraud: It happens when somebody knows the expiry date and account number of your card, they can easily commit to CNP fraud . This can be done through phone, mail or internet. It generally happens when somebody uses your card without actually being in physical owned of it. Frequently, merchants opt for the card verification code for making CNP fraud but CNP frauds are slightly more difficult, but if a fraudster can get your account number, they probably know that PIN number too. By the mean time there are only 999 possible combinations for the 4-digit verification code. Many fraudsters are working to figure out right number.

d) ***Counterfeit Card Fraud:*** Counterfeit card fraud is usually committed through skimming. This fake magnetic swipecards holds all your card details such as card number, account number, PIN number and so on. This fake magnetic strip is then used to create a fraudulent card that is fully functional. It is an exact copy of original card, which means fraudsters can simply use card in machine to pay for purchases, or removing funds. This type of fraud can also be committed through someone who conscious your card details. They can use this information to create a so-called *fake plastic*. In the South Africa India is the top second country in counterfeit fraud in credit cards during 2014[16].

e) Lost and Stolen Card Fraud: In this type card will be taken from your control, either through theft or because of lost .The Fraudsters who get card in their hands, they will make use for payments. This type of fraud is difficult to do through machines, as they will require a pin number. However, it becomes easy to use a found or stolen card for online purchases.

f) Mail Non-Receipt Card Fraud: This type of fraud is also known as never received issue or intercept fraud. In this case, when people were expecting a new card or replacement one and Fraudsters is able to intercept these. The Fraudsters will then register the card and they will use for personal gains and purchases of goods.

g) Doctored Cards: A doctored card is a card whereby a strong magnet used to erase its metallic stripe. Fraudsters do this to manage and modify the details on the card itself. So that they can easily match and valid cards usually, this card won't work when a fraudsters tries to pay for something. However, Fraudsters will then use their quality or power to convince a merchant to just enter the details of the card manually.

h) Fake Cards: In this type, Fraudsters is skilled who can forge this type of cards using fake names and account numbers and will make transactions with the card. This type of card isn't actually linked with an account, so the

credit card company will not pay for the fraudulent transaction since they cannot link it with specific user. By that time, however, the fraudsters will be long gone uses their fake cards with their purchases.

i) Account Takeover: Takeover is taking possession of other's account. Fraudsters gather all the required information about the credit card and the cardholders. Then they contact the card issuer and pretence as genuine cardholder and request them to modify their billing address or else they will report a card loss and request for a new card (replacement) at the new address.

II. CREDIT CARD FRAUD DETECTION USING DATA MINING VS BIG DATA

Data mining used to find the customers pattern matching in which abnormal behavior/activities of customers are identified using any one of these following approaches like Classification, Clustering, Genetic algorithm, Decision tree, Fuzzy logic based system, Neural network etc.,

Steps used in Data mining Approach which is illustrated in figure 2.1.

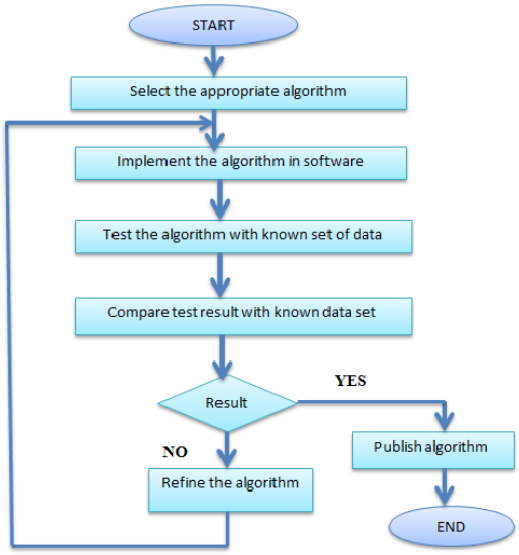

Fig. 2.1 Data mining approach

Data mining uses above approach to select the appropriate algorithm to detect the credit card fraud.
a) Genetic algorithm: In this technique fraud transaction are identified using sample data set .It is used to minimizing the wrongly classified number of transactions.

b) Decision tree: It separates the complex problem into many simple problem and resolve the sub problem using CART and IF-THEN rule methods. This method easily finds the fraudsters through tracing fake mail and IP address.

c) Fuzzy logic based system: This method classifies the transaction into suspicious and non-suspicious. It comprises GA and fuzzy system to reduce a false alarm.

d) Neural network: This is totally based on the pattern recognition system. It matches the customer frequent transaction pattern with live transaction, if it matches then the neural network declare the transaction as authorized one, else it produce alert alarm to credit card holder.

e) Outlier detection: It is an unsupervised data learning method. Using this technique credit card user abnormal behavior is identified which helps to detect the credit card fraud.

III. CREDIT CARD FRAUD DETECTION ON BIG DATA

Big data uses three main methods to detect the credit card fraud which is shown in figure 2.2

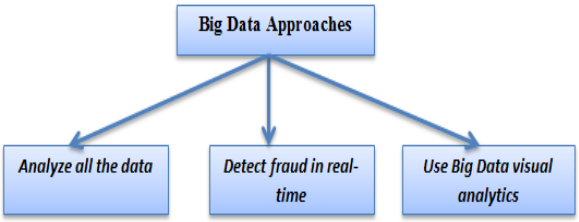

Fig. 2.2 Big data approach

1. **Analyze all the data:** In the past banks are used a sample or a subset of their data for fraud analysis. It took too much time and cost to use the complete data set. Big Data can analyze all the data for fraud, and new data sources can also be introduced. Analyzing the full data set leads to several benefits:
   (a) Reviewing all transactions based on the defined fraud rules.
   (b) Identifying new fraud patterns that get added to the growing list of fraud rules.
   (c) Minimizing false positives to avoid losing revenue and turning away customers. All this data can be aggregated and analyzed using Big Data tools like Hadoop.

2. **Detect fraud in real-time:** Real time transactions data are combined with data from other sources like existing data warehouses to detect fraud in real-time. This helps to prevent credit card fraud where the transaction is shown against a set of pre-defined fraud rules as part of the credit card authorization. This includes combining site data with data from customers social feed, the geo data from customers smartphone applications, purchase history, and web logs. . For example, when a credit card holder is travelling to outstation by an airplane and posts the status on Facebook or any social feed, credit card transaction during that period will be abnormal and the bank can block the transaction for particular period.
3. **Use visual analytics:** Big Data tools offer the capability to visually analyze data and acquire insights, even though the data may be coming from different sources. Banks can use these tools to identify regions, products, and customers that have a higher fraud rate based on historical analysis. This identifies the various areas where time and cost should be invested to minimize fraud. Visualization also reduces manual efforts to reviewing every data order. The reports can graphically represent the probability of fraud for each order transaction and connect to email or message alerts for escalation as needed

IV. HOW BIG DATA DETECT THE CREDIT CARD FRAUD IN REAL TIME

Big data is the frontier of a firm's ability to store, process and access all the data and it needs to operate effectively for decision making, reduce risks and serve customer [17].The 3 main characteristic of big data volume (data quantity), velocity (data speed), variety (data type). Big data can handle more than 1 million customer transactions per hour. *Hadoop* is an Apache top level project, open-source implementation of frameworks for reliable, scalable, distributed computing and data storage [18]. It is a flexible and highly-available architecture for large scale computation and data processing on a network of commodity hardware. Big data helps financial institutions to approach fraud in different ways and possibly get different results. For the credit card fraud detection we need bank, transaction, and customer data. Some of the examples are

(a) Data associated with transaction eg.pos number, account number, date, time, transaction amount, merchant category code.

(b) Data associated with card holder e.g card type, expiration date, home address

(c) Data associated with transaction history eg. Number of transaction last 24 hours, location, time difference from last transaction, average transaction, fraud risk of merchant type

Big data building a real time solution for credit card fraud detection: *There are two phase in real time fraud detection:* One for building the model with training features and labels and one for testing the model with test feature to get predictions**.** Compare the test predictions to the test labels. Loop until satisfied with the model accuracy (i) Adjust the model fitting parameters, and repeat tests. (ii) Adjust the features and/or machine learning algorithm and repeat tests.

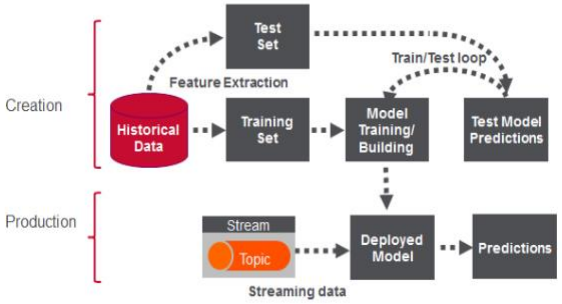

Fig. 2.3 Real time fraud detection model

Building the model: *Classification* is a family of supervised machine learning algorithms that identify each item belongs to which category (for example whether a transaction is performed by fraud or not fraud), *Linear Regression* is the commonly used predictive analysis. Regression: technique concerned with predicting some variables by knowing others. eg. The process of predicting variable Y using variable X Here it is used to predicts a value (eg amount of fraud).*Logistic Regression* is the appropriate regression analysis to conduct when the dependent variable is splitting into two parts (binary).Here it is used to predicts a probability (eg probability of fraud)

Real Time Fraud Detection Solution in Production: The diagram below shows the high level architecture of a real time fraud detection solution, which is capable for high performance at any scale. Credit card transaction data are delivered through the MapR Streams messaging system, which supports by the Kafka .09 API. This data details are processed and checked for Fraud by *Spark streaming using Spark Machine learning* system with the deployed model. MapR-FS, which supports the posix NFS API and HDFS API, is used for storing event data [19]. MapR-DB a NoSql database which supports the Hbase API is used to store credit card holder profile and it is also helps to fast access to credit card holder profile data [19].

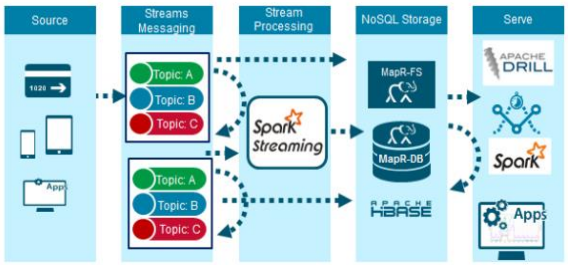

Fig. 2.4 Architecture of Real time credit card fraud detection using spark streaming and Hadoop

MapR Streams is a new distributed messaging system which enables bank/merchant and customer to exchange events in real time via the Apache Kafka 0.9 API. MapR Streams are logical collects the messages and organize events into categories [19]. These results are categories into three types

- Raw Transaction: raw credit card transaction events.
- Enriched Transaction: credit card transaction events improved with card holder features, which were predicted not to be fraud.

- Fraud Alert: credit card transaction events improved with card holder features which were predicted to be fraud.

V. PRECAUTION TO AVOID CREDIT

CARD FRAUD

The major reason for credit card fraud is carelessness and lack of proper precaution in handling of credit cards.

A. Protect your pin:

- Memorize your pin, never write down your personal identification number (PIN) anywhere.
- Never give your PIN to anyone, don't use a pin that could be easily guessed, often try to change your PIN number.

B. Protect your account number:

- Never write your credit card number on post cards or on the outside of envelopes.
- When you initiate a call to give your credit card number on the phone or mail never provide your card or person details or cvv (card verification value).
- Never give your account number to anyone who calls you on the phone or sends you an e-mail.

C. Billing statement:

- When disposing of your old receipts and statements, don't use public waste bins.
- Review credit card statements as soon as arrive and report if any questionable charges to your card issuer immediately.
- A missing credit card statement may indicate stolen mail, so contact your card issuer right away if your bill doesn't arrive around the usual date.

D. Protect your wallet/purse:

- Never carry all your credit cards, choose 1 or 2 cards you might need for holiday, shopping etc., If your wallet or purse is lost or stolen, call your credit card issuers immediately.

E. Reporting a credit card fraud:

- To report credit card fraud, Call your credit card bank immediately about Lost or stolen cards or PIN numbers or Unauthorized charges on your statement, Request a fraud affidavit, Get a police report if necessary

F. Safeguard your mail/messages:

- Lock your mailbox. Never leave mail in an unlocked mail box or apartment building lobby.
- Keep your phone/SIM on active mode, incase if you have lost your phone/SIM try to resolve it quickly.
- Put your return address on out-going mail.
- Shred unwanted credit card solicitations before discarding.

G. Internet safeguard:

- If you bank online, don't use remember password for credit card or bank sites.
- Don't enter your credit card number to unknown websites which provides "free access".
- While going for any online transaction, make sure that the website is trusted one, don't click any unknown link in your mail, it could be phishing mail.

H. Take precaution:

- Inform your credit card company if you are going to be traveling away from home to prevent any inconvenience if your issuer should block your account from being used in a different city.
- Inform your credit card company if you are going to make any unusually large purchases so that your account is not flagged for possible fraud.
- Notify the post office and your credit card company immediately if you change your address.
- Register your cards with verified by Visa and MasterCard by this system as 3-D Authentication, From this you can add a safety measure in additional lock besides the ATM pin number of your card.

VI. CONCLUSION

As global networking provides many ways for fraudsters, building an accurate and easy handling credit card fraud detection system is one of the major tasks for banks. There are several ways to detect the fraud transaction. In this paper we have discussed about types of credit card frauds, precaution steps to avoid the credit card fraud and big data techniques to detect the card frauds. However, the usage of big data for this purpose is till at its early stages, so lot of effort needed to place big data technique in real time fraud detection. Because it has ability to work with large and real time transaction data set and also helps to reduce risk and response time to milliseconds.

VII. REFERENCES

[1] Neha Sethi, Anju Gera, April 2014," A Revived Survey of Various Credit Card Fraud Detection Techniques", International Journal of Computer Science and Mobile Computing, IJCSMC, Vol. 3, Issue. 4, April 2014, pg.780 – 791

[2] Krishna Kumar Tripathi, Mahesh A. Pavaskar, "Survey on Credit Card Fraud Detection Methods", International Journal of Emerging Technology and Advanced Engineering, Vol 2, Issue 11, November 2012, pg.721 – 726

[3] Eswari.M,Navaneetha,Krishnan.M." Survey on Various Types of Credit Card Fraud and Security Measures", International Journal of Advanced Research in Computer Science and Software Engineering , Volume 4, Issue 1, January 2014, pg.1235 – 1238

[4] Anika Nahar,Sharmistha Roy,"A Survey on Different Approaches used for Credit Card Fraud Detection", International Journal of Applied Information Systems (IJAIS) – Foundation of Computer Science FCS, New York, USA Volume 10 – No.4, January 2016,pg.29 – 34

[5] Priya Ravindra Shimpi, Prof. Vijayalaxmi Kadroli ,"Survey on Credit Card Fraud Detection Techniques", International Journal Of Engineering And Computer Scienc,Volume 4 Issue 11 Nov 2015, Page No. 15010-15015

[6] P.Jayant,Vaishali,D.Sharma," Survey on Credit Card Fraud Detection Techniques", International Journal of Engineering Research & Technology (IJERT), Vol. 3 Issue 3, March – 2014,pg.1545-1551

[7] B.Sanjaya Gandhi, R.Lalu Naik, S.Gopi Krishna, K.lakshminadh "Markova Scheme for Credit Card Fraud Detection". International Conference on Advanced Computing, Communication and Networks; (2011). (144-147).

[8] Abhinav Srivastava, Amlan Kundu, Shamik Sural, Arun K. Majumdar. "Credit Card Fraud Detection using Hidden Markov Model". IEEE Transactions on dependable and secure computing,Volume 5; (2008) (37-48).

[9] Anshul Singh, Devesh Narayan "A Survey on Hidden Markov Model for Credit Card Fraud Detection". International Journal of Engineering and Advanced Technology (IJEAT), (2012). Volume-1, Issue-3; (49-52).

[10] V.Dheepa, Dr. R.Dhanapal "Analysis of Credit Card Fraud Detection Methods". International Journal of Recent Trends in Engineering, (2009). Vol 2, No. 3; (126-128).

[11] A 'layman's' explanation of Ultra Narrow Band technology," Oct. 3, 2003. [Online]. Available: http://www.vmsk.org/Layman.pdf. [Accessed: Dec. 3, 2003]

[12] https://letstalkpayments.com/early-detection-can-reduce-30-40-of-credit-card-fraud-loss-an-interview-with-canh-tran-ceo-co-founder-of-rippleshot/

[13] https://rbi.org.in/SCRIPTS/ATMView.aspx

[14] http://www.kdnuggets.com/2016/03/combat-financial-fraud-using-big-data.html

[15] http://newsroom.mastercard.com/asia-pacific/2014/10/28/8-different-types-card-fraud/

[16] http://www.businessmedialive.co.za/2014-credit-card-fraud-statistics/

[17] https://gcn.com/article/2013/07/29/isc2-big-data.aspx

[18] http://www.sas.com/en_us/insights/big-data/hadoop.html

[19] https://www.mapr.com/blog/real-time-credit-card-fraud-detection-apache-spark-and-event-streaming