Designing an Automated Distributed System for Credit Card Fraud Detection

Somtoochukwu Ilo, Tochukwu Chiagunye and Amaechi Chineke, 
Computer Engineering Department Michael Okpara University of Agriculture, 
Umudike, Abia State, Nigeria

Abstract: Credit-card fraud is increasingly rampant in the recent years for the reason that the credit-card is majorly used to request payments by many companies and individuals on the internet. Therefore the need to ensure secured transactions for credit-card owners when using their credit cards to make electronic payments for goods and services provided on the internet is a criterion. The popularity and recognition gained by data mining in combating credit-card fraud because of its application of artificial intelligence (AI) techniques and algorithms that can be implemented to detect or predict fraud through Knowledge Discovery from unusual patterns derived from gathered data is of great important and have serious application in this design. This work presents a generic model for credit card fraud detection. The system implements the supervised means of data mining to detect fraud in a real time transaction on the internet, and thereby classifying the transaction as legitimate, suspicious fraud and illegitimate transaction. The anomaly detection algorithm is designed on the Neural Networks which implements the working principal of the human brain.

Keywords: Credit Card, Fraud Detection, Distributed System, Data Mining, Artificial Intelligent.

I. INTRODUCTION

A. Credit Card Fraud

Recent and current scholars investigating credit-card fraud have divided credit-card fraud into two types:

- The online credit card fraud (or no card present fraud)
- The offline credit card fraud (card present fraud)

The online credit-card fraud is committed with no presence of a credit-card but instead, the use of a credit-card information to make electronic purchase for goods and services on the internet.

The offline credit-card fraud is committed with the presence of a credit-card which in most cases have been stolen or counterfeited and thereby used at a local store or a physical location for the purchase or some goods or services.

However, to define credit-card fraud, it is a scenario where the credit-card information of a credit-card owner has been stolen, or in some cases valid credit-card information has been uniquely generated (just like credit-card companies or issuers do) and thereby used for electronic payment on the internet or via the telephone. In most cases, no I.T or computer skill may be required to commit online credit-card fraud because of the different techniques in which credit-card information can be stolen by fraudsters. In other cases, this software is written by black-hat hackers with hacked credit-card information stored on a database file from which the software can display valid credit-card information to other type of credit-card fraudsters who have bought the software to use. This technique is some cases used by black-hat hackers to sell their hacked credit-card information to other online credit-card fraudsters with little or no computer skills.

B. Credit-card information buyers:

They are fraudsters with little or no professional computer skills (e.g. Computer Programming, Networking, etc.) who buy hacked (or stolen) credit-card information on an illegal “credit-card sales” website. They buy this credit-card information with the intention of making electronic payment for some goods and services on the internet.

C. Black hat hackers:

Recent research on Hackers in terms of Computer Security defined a "black hat hacker" (also known as a cracker) as a hacker who violates computer security with malicious intent or for personal gain. They choose their targets using a two-pronged process known as the "pre-hacking stage"; Targeting, Research
and Information Gathering, and Finishing the Attack. These types of hackers are highly skilled in Computer Programming and Computer Networking and with such skills can intrude a network of computers. The main purpose of their act of intrusion or hacking is to steal personal or private information (such as credit-card information, bank-account information, etc.) for their own personal gain (for instance creating a “credit-card sales” website where other credit-card fraudsters can buy credit-card information).

D. Physical credit-card stealers:

They are the type of fraudsters who physically steal credit-cards and write out the information on them. They physically steal these plastic credit-cards (maybe by pick-pocketing in a crowded place) and write out the credit-card’s information with the intention of using this credit-card information to make electronic payment for some good and services on the internet.

E. CC/CVV2 shopping websites:

Credit-card fraudsters with no professional computer skills can buy hacked credit-card information on these websites to use for fraudulent electronic payment for some goods and services on the internet.

F. Physical stolen credit-card information:

Fraudsters can physically steal the credit-card of a user to write out the credit-card information and then use for fraudulent electronic payment on the internet.

G. Credit Fraud Detection

It has already been mentioned that credit card fraud can affect all consumers, merchants and issuing banks. Therefore, it is important to establish techniques for detecting and preventing credit card fraud. The literature contains a variety of techniques which can be used to build fraud detection systems. Understanding the characteristics of all those techniques can be a tedious task. A technique which promises a high predictive accuracy may be an appealing resource to be used in fraud detection system. However, there are various different parameters that need to be considered before deciding which technique best suits the needs of a particular situation.

F. Data Mining

Data mining is a process that uses a variety of data analysis tools to discover patterns and relationships in data that may be used to make a valid prediction. The six basic steps of data mining process are defining the problem, preparing data, exploring data, building models, exploring and validating models, deploying and updating models. Neural network is the data mining technique used in this study and it utilized these steps for accurate and reliable result. Neural network was used because of its ability to adapt and generalize.

II. LITERATURE REVIEW

A. Credit Card Fraud

This is the most common fraud type that occurs in credit industry. A fraudster uses a legitimate card to undertake illegitimate transactions. The card holder is not aware of the fact that their card is being used without their permission. The fraudster takes advantage of card holder’s ignorance by undertaking as much transactions as possible before the card holder realizes and reports the fraud to their bank.

According to Laleh et al. (2009) credit card fraud can be committed either offline or online. These two ways are discussed below.

B. Offline Credit Card Fraud

Offline fraud occurs when a fraudster steals the physical card and uses it at the actual stores. Although offline fraud is still popular nowadays; it is less common because there is a higher probability to fail. More precisely, the cardholders tend to realize the loss of the physical card and report that to their bank before the fraudster manages to undertake any illegitimate transactions with it. As soon as the stolen card is reported to the bank, the latter will lock the card so as it cannot be used anymore. It is particularly useful to notice that if the cardholder does not realize the loss of their card, a significant financial loss can occur. As mentioned in the introduction chapter, the policies of some banks enforce cardholders to pay for the losses which occur due to an unreported credit card theft.

C. Online Credit Card Fraud

During online fraud only the details of the card are stolen and not the card itself. This is also known as virtual card theft. The details of the card can be used in places where the card need not be physically present like internet or phone purchases. This type of credit card fraud is very dangerous and more difficult to prevent because fraudsters can hold credit card’s
information for a long period of time before they use it. There is no way for the cardholder to know in advance that their credit card information is stolen. Therefore this type of fraud may only be detected after one or more illegitimate transactions are taken place. There are various ways that fraudsters adapt in order to steal the information of credit cards. Some of these ways are briefly discussed below.

**D. Skimming**

Patidar et al. (2011) define skimming as the “process where the actual data on a card’s magnetic stripe is electronically copied onto another”. Fraudsters use special-purpose devices – also known as skimmers – to capture the information of credit cards that are encapsulated inside their magnetic stripes. They can use the stolen card information to create counterfeit physical cards in order to use them at actual shops or simply supply the card information at online shops. Skimming can be committed by an unfaithful employee, who may swipe customer’s card using the skimmer device, while the customer is at the point of sale. In the past, skimmer devices have also been introduced on ATM cash machines. In addition to that, micro-cameras have been used to record the PIN code of a cardholder during ATM transactions.

**E. Site Cloning**

Fraudsters clone a legitimate website to deceive customers into placing an order with them. Since the fraudulent website seems identical to the legitimate one, the unsuspecting customers provide their credit card information to complete their order. Consequently fraudsters who obtained the customer’s credit card information can commit credit card fraud whenever they wish to.

**F. False Merchant Sites**

According to Patidar et al. (2011) there are various websites that ask for credit card information in order to confirm customer’s age. These websites will never charge the credit cards directly but they may sell their information to fraudsters who will commit credit card fraud.

**III. MATERIALS AND METHODS**

This data mining system in this case, solves the problem of fraud, in the activities performed using credit cards online. With the application of data mining and artificial intelligence, transactions can be verified by the real owner of a bank account before the transaction can be updated into his account details and account balance; it combines different schematics and system of operation to perform the fraud detection, ranging from CVV verification, Internet banking secure code, Card Pin Verification, account owner personal digital signature. It also keeps record of the website’s data where transaction is made, Such as IP-Address, Site Url, Secured server certificate verification and records of known fraud patterns, specific to each website. This automated system is designed and developed as a web based application such that each user can access his/her internet bank account online, view/modify account details, verify transactions, e.t.c. on any internet browser. The platform is designed using HTML (hypertext markup language), PHP(Hypertext Processor), this solves platform compatibility issue such that the internet banking system can be accessed using any browser on any operating system once connected to the internet. The database is designed and implemented on the website using Php Mysql to apply the SQL (Structured Query Language) to the system, where the entire user’s Information, credit card details, transaction records and other system information analysis squarely deals with the software development activities.

The credit card fraud detection system developed used four clusters of low, high, risky and high risk as shown in Figure 3.1 below. Once the transaction is legitimate, it was processed but if transaction falls in any of these clusters; it was labelled as suspicious/fraudulent. The alert goes off and the reason is given. The fraudulent transaction will not be processed but will be committed to the database.

The approach involves the following steps, which is also illustrated in figure 3.2. The steps are the selection of an appropriate algorithm; implement the algorithm in software; test the algorithm with known data set; evaluate and refine the algorithm as it is being tested with other known data sets; and show the results.
IV. CREDIT CARD FRAUD DETECTION MODEL

The major functionalities of the artificial neural network (ANN) based credit card detection system designed are as follows: to facilitate real-time transaction entry, and react to a suspicious transaction that may lead to fraud. The design of the architecture is based on a neural network unsupervised method, which was applied to the transactions data to generate four clusters: the low, high, risky and high-risk clusters. The system runs secretly beneath the banking software within banks offering credit card services where fraudulent transactions are observed. Business rules relevant to the enlisted CCF types are further applied to the four clusters to detect transactions that deviate from the norm. Deviation from the usual pattern of an entity implies the existence of a fraud. Each transaction entering the database such as withdrawal, deposit, and any card transaction is treated as a signature, suspected and prone for verification.

The similarity between a customer’s present transaction and a known fraud scenario indicates the same fraud may occur again. Suspected transactions are flagged within seconds for further investigations and subsequent decision-making. Visualization is provided using appropriate graphical user interface (GUI). The flow chart of the artificial neural network based credit card fraud watch is shown in Figure 4.1, below. The flow chart shows how different transaction can be performed online and how to make the verification of rough data mining process.
Figure 4.1: Flow Chart of the proposed Credit Card Fraud (CCF) Watch

A. System Algorithm Design

The proceeding algorithm is not in pseudo code neither is it written in any programming language but it is written in a natural language.

B. Internet Transaction System

Start
data(itemselect) {
If(selected):
get price, name, description, websitesdata
request card details
read digital signature
if(signature=verified):
request cvv, pin
endif
read (account data) {
compare balance
verify eligibility()
if(balance>price):
confirm eligibility
endif
update newbalance()
save user record
save pattern for future mining
end

C. Transaction Review And Verification

Start
Userlogin() {
if(password,username match databaserecord):
Read userdata();
}
Reviewtransaction() {
If(access granted):
Read (transaction details):
Endif
}
Request System.testdata() {
If(pattern match testdata):
Request user.confirm()
}
User.confirm() {
Confirm date.transaction
Confirm cvv.transaction
Confirm pin.transaction
Confirm url.transaction
Confirm price.transaction
Confirm usersecurecode.transaction
}
Validate record(user) {
Select confirm()
Select abort()
}
If(select confirm()) {
Transaction verify()
Update system data:
} elseif(select abor()) {
Transaction cancel
Update system testdata:
}
Transaction cancel() {
Delete userdata.transaction()
Update system testdata();
} End
D. Input Design

For there to be an output, there must be an input which must be processed to yield the required output. The input section is used to request the credit card details from the user who wants to perform transaction on a merchant website. The supplied data on the input forms servers as the input to the system. These inputs are linked to the banking credit card database and comparison is made with the pre-stored data in the database, and when confirmed loads the users digital signature which helps the user verify that the card details were retrieved from the banking institution.

Figure 4.2: Input Form For The Payment Page.

Once the credit card number and cvv is verified from the bank database, It then loads the users digital signature which will now serve as a second input screen for the user to provide the credit card pin, internet banking secure code and first name written on the credit card.

E. Output Design

When the User has logged into his internet banking platform with his login credentials, he can be able to view all the transactions made so far using his credit card and can by this means legally confirm or delete the transaction.

Figure 4.4a: Users Transaction History
**F. System Database Design**

Database is a repository for data and information in an organized way to prevent redundancy and to preserve data integrity amongst others. The design and implementation of the database is done on Phpmyadmin sql manager. This database consists of table of which some are: Tblaccounts, tblcards, tblpatterns, tbltransactions, amongst others. Each of these tables has their column name (Field) and data types respectively.

**Table 4.1a: Credit Card Database**

<table>
<thead>
<tr>
<th>Field Name</th>
<th>Data Type</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>Card_id</td>
<td>INT</td>
<td>9</td>
</tr>
<tr>
<td>First_name</td>
<td>VCHAR</td>
<td>40</td>
</tr>
<tr>
<td>Last_name</td>
<td>VCHAR</td>
<td>40</td>
</tr>
<tr>
<td>Card_number</td>
<td>INT</td>
<td>15</td>
</tr>
<tr>
<td>Expire_date</td>
<td>Date</td>
<td>12</td>
</tr>
<tr>
<td>Card_pin</td>
<td>INT</td>
<td>4</td>
</tr>
<tr>
<td>Card_cvv</td>
<td>INT</td>
<td>3</td>
</tr>
</tbody>
</table>

**Table 4.1b: Transaction Records**

<table>
<thead>
<tr>
<th>Field Name</th>
<th>Data Type</th>
<th>Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trans_id</td>
<td>INT</td>
<td>10</td>
</tr>
<tr>
<td>Trans_acct</td>
<td>INT</td>
<td>10</td>
</tr>
<tr>
<td>Trans_date</td>
<td>DATE</td>
<td>12</td>
</tr>
<tr>
<td>Trans_type</td>
<td>TEXT</td>
<td>8</td>
</tr>
<tr>
<td>Capture_url</td>
<td>VCHAR</td>
<td>50</td>
</tr>
<tr>
<td>Capture_ip</td>
<td>VCHAR</td>
<td>15</td>
</tr>
<tr>
<td>Capture_system</td>
<td>VCHAR</td>
<td>20</td>
</tr>
<tr>
<td>Capture_country</td>
<td>TEXT</td>
<td>20</td>
</tr>
<tr>
<td>Capture_time</td>
<td>DATE</td>
<td>12</td>
</tr>
<tr>
<td>Trans_description</td>
<td>VCHAR</td>
<td>150</td>
</tr>
<tr>
<td>Trans_alert</td>
<td>TEXT</td>
<td>3</td>
</tr>
<tr>
<td>Trans_approv</td>
<td>TEXT</td>
<td>3</td>
</tr>
<tr>
<td>Trans_amount</td>
<td>FLOAT</td>
<td>10</td>
</tr>
<tr>
<td>Trans_currency</td>
<td>VCHAR</td>
<td>10</td>
</tr>
</tbody>
</table>

**CONCLUSION**

In this paper, a data mining application has been modeled as a subsystem which can be used with software systems and applications in financial institutions to detect credit-fraud in a transaction on the internet. This Data mining application accepts input formatted on a pattern on which a transaction is being executed and matches it with the credit-card holder’s patterns of its credit-card online consumptions it’s been trained with to classify a real-time transaction as legitimate, suspicious fraud or illegitimate transaction. The data mining application modeled in this paper uses the anomaly detection algorithm of the Neural Networks to detect fraud in a real-time transactions and it is not prone to errors because of its classification of transactions (legitimate, Suspicious Fraud and illegitimate). In the case of the suspicious fraud classification, the financial institution using the system can investigate further by calling the credit-card owner regarding the suspicious fraudulent transaction.

**References**