iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 8 **Issue:** V **Month of publication:** May 2020

DOI: http://doi.org/10.22214/ijraset.2020.5187

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 8 Issue V May 2020- Available at www.ijraset.com

Credit Card Fraud Identification using Hidden Markov Model

Ananya[1], Kavya Vasudev[2], Medehal Rachana[3], Milinda B[4], Bhavatarini N[5]
[1, 2, 3, 4, 5]*Dept. of Information Science and Engineering, Sapthagiri College of Engineering, Karnataka, India*

Abstract: Regardless of the rapid development of online retail technologies, Credit card use has intensively escalated. Since credit card is the most common mode for both online and daily purchases, as in offline purchases, cases of fraud connected with it often increase. Within this paper we use Hidden Markov Method to method the order of methods involved within performing credit card purchases.
Keywords: Credit card, internet, online shopping, fraud detection, Hidden Markov Model

I. INTRODUCTION

Internet retail is rising so quickly that the scale of the worldwide demand for online purchases is projected to reach about 4 trillion by 2021. Yet we're still projected to have over 300 million internet customers in the US by 2023. That adds up to 91 per cent of the world entirely. According to the resources as of now, almost 69% of Americans have purchased online, and a minimum of 25% of them shop online at least once a month. Most of these customers purchased apparel pieces, and the first items purchased through Amazon were 47 per cent. Not only Americans but people all over the world do online shopping as they understand it's benifits. As per Invesp, the following countries are leading with average e-Commerce revenue per shoppers: USA: $1,804, UK: $1,629, Sweden: $1,446, France: $1,228, Germany: $1,064, Japan: $968, Spain: $849, China: $626, Russia: $396, and Brazil: $350. It was projected that there are approximately 1.92 billion online customers in 2019, and e-commerce transactions responsible for approximately 14.1 per cent of amount of online purchases made internationally. As we all know, since online shopping is growing in such an exponential rate, upcoming statistics shouldn't be of a surprise.[7] By 2023, e-Commerce sales for the retail purchases are expected to rise from 14.1% to 22%. With the increase of online shopping and e-commerce usage of credit cards for online transactions is increasing. A month, nearly one million new credit cards are issued and the number of unique card holders vary from 25 to 30 million. This shows the phenomenonal increase in use of credit cards users [8]. If the group of consumers of credit cards grows widely, there is also growing potential for criminals to do the exploitation of credit card data in order to commit crime. The 19-and-under core audience saw 1,565 payment card fraudulent incidents, comprising 9.9 percent of their overall identity theft data. Recruitment or tax-related crime of 7,860 incidents comprising 49.7 percent of the identity theft data they collect. To bring the figure into context, just 11.7 per cent of their identity crime records were for jobs or income-related theft by other age classes. Purchases centered on credit cards may be divided into two types: 1) physical card, and 2) electronic card. The cardholder shows his card directly to a dealer for paying a bill in a physical-wallet-based order. An intruder needs to swipe a purchaser's payment card to conduct illegal purchases in this kind of purchase. When the cardholder fails to understand the card failure, it will in the future result in a significant financial cost to the credit account agency. In the later kind of purchase that is virtual method, only some necessary details about a card (Card model, expiry date, CVV or the source number) is required to make the payment that is the online transactions. Such purchases are normally done on the Internet or over the telephone.

II. LITERATURE SURVEY

Mapping of credit card fraud has made headlines to a lot of development in analytics and a range of methods including data mining and neural networks among many others.

A. Parallel Granular Neural Networks.

A parallel granular neural network (GNN) was developed to fasten the process of credit card fraud detection done using data mining and knowledge discovery process. Firstly, the parallel fuzzy neural network is trained in parallel using the training data sets which runs on a 24-processor system, and then parallel fuzzy neural network which is trained using training data sets discovers the fuzzy regulations for later estimation. To employ this credit card fraud detection system, GNN (Granular Neural Network) method is employed which uses fuzzy neural network based on knowledge discovery (FNNKD)using C language on the UNlX environment.

©IJRASET: All Rights are Reserved

The FNNKD is the basic building blocks of GNN. GNN is capable of processing various granular data (granules) such as class of numbers, a cluster of images, a set of concepts, a group of objects, a category of data, etc. These granules are inputs and outputs of GNNs as multimedia data are inputs and outputs of biological neural networks in the human brain. GNN gives fewer average training errors with larger amount of past training data. It's also found that the number of training error is inversely proportional to the number of training cycles. The data is first extracted into SQL database server which already contains the details about Visa card transactions. The entire system is parallelized on the Silicon Graphics Origin 2000 which is a Shared memory multiprocessor system consisting of 24-CPU, 4G main memory, and 200GB hard-drive.

1) Demerits

a) There seems to be no clear law defining the artificial neural network framework.
b) Neural frameworks indulge processors with parallel processing power, in accordance with their structure.
c) As ANN generates a sample result, it does not provide any hint as to why or how. Thus reducing the trust in the network.
d) Parallel GNN is difficult to setup and operate.

B. Cardwatch: Neural Network-Based Database Mining System.

We present in this paper, Cardwatch[2], a database mining program used to identify credit card theft. The platform is focused on a neural network learning board, provides an introduction to a variety of industrial repositories and includes a basic graphical user interface. Results of the test observed for synthetically produced credit card data and an auto associative neural network model suggest very strong identification levels for fraud. This software is readily expandable and can run right on a wide range of commercial repositories. The present framework edition was evaluated using an auto associator with very encouraging outcomes on produced synthetic data: 85 per cent fraud detection frequency and 100 per cent legitimate transaction recognition rate.

1) Demerits

a) "One customer per network" limitation is seen, i.e. a different network must be used for any customer known to the device.
b) It is difficult to confirm the structure.
c) The large neural networks need extensive training.
d) It has low capability of clarification.
e) Non-numerical data must be translated and standardised.

C. Neural Data Mining for Credit Card Fraud Detection.

We use a credibility-based neural network in this paper [4] to construct the sequence of operations in the handling of card payments. Receiver operational characteristics, that is, ROC analytics technology is now used to ensure reliable and efficient fraud identification. A neural network is initially trained with synthetic data. When an arriving credit card purchase with too low trust is not recognized by the equipped neural network model (NNM), the purchase would be labeled counterfeit. This paper demonstrates how confidence level, neural network framework and ROC can be used efficiently to identify credit card fraud. Within this approach, we discuss our analysis and demonstrate the implications of the data mining strategies like neural networks and trust estimation. The suggested confidence-based neural network classifier ensures desired universal applicability and an effective identification rate in credit card fraud identification.

1) Demerits

a) As massive quantities of data are being gathered in data mining systems, hackers could hack some of this very crucial data.
b) Data mining involves a highly trained expert to prepare the data and analyse the outcomes. As data mining yields out the various patterns and interactions whose patterns the consumer has to render sense and validity.

D. Web services based collaborative scheme.

In a web-based collaboration scheme [3] for the identification of credit card fraud, we use the proposed scheme where participating banks may contribute awareness of theft trends in a non-homogeneous and dispersed context and can thus improve their capacity to identify fraud and minimize financial losses. In a web-based collaboration identification approach, participating banks function as service customers, whilst the Fraud Patterns Mining Service Center (FPMSC) acts as a services provider. Participating banks should follow standard data formats for affirming the shared data to

©IJRASET: All Rights are Reserved

enable data sharing through heterogeneous network implementations. FPMSC issues a WSDL file detailing the design of the architecture and configuration of its supported services. Fraud patterns mining algorithm:

Input: T: embedded transactions (fraudulent);
minsupp: user-specific minimum assistance;
Output: L: collection of all regular super itemsets;

```
L = {};
    L1 = {frequent 1-itemsets};
    L = L ∪ L1;
    for ( 2; ; ){ k = Lk−1 ≠ φ k + +){
        Ck = Lk−1 > Lk−1;
        for each transaction t∈T {
        for each candidate itemset c ∈Ck
            if (c = subset(t)) then c.count + + ;
        Lk = {c ∈Ck | c.count ≥minsupp};
        L = L ∪ Lk;
        for each frequent itemset f ∈ Lk {
            for each subset(f)
    L = L – subset(f); }
    }
}
return L ;
```

Via web services, the data interchange becomes unhindered across heterogeneous applications of banks. The developed collaborative web services-based scheme may be updated in the future as a framework for the sharing of information with different companies and industries.

1) Demerits

a) Performance Issues
b) Not 100% available all the time
c) Lack of Standards
d) Only sending sensitive data to web services is allowed by using HTTPs otherwise, anything sent over internet can be viewed by others.

III. EXISTING MODELS USING ONTOLOGY

Graphs can be utilized to extricate and handle information. Subsequently, extortion rate has raised in this range which to overcome such inconsistencies we show them through a mode of charts. Innovation has gotten to be progressed in keeping banking field as well, consequently the rate at which credit are being utilized has heightened. Subsequently, extortion rate has raised in this range which to overcome such inconsistencies we show them through a mode of charts. Of the eminent upper-hand of present approach is lessened gadget overload estimate during walking operations with the intention to perceive frauds and growth of detection speed. Detecting inconsistencies in credit card transactions by ontology is a very successful method needing low computing workload and fewer capacity to handle credit card transaction details, and utilizing data mining to identify anomalies. Throughout this method, using ontology graph for the transaction activity of each individual and then storing it in the framework, only certain transactions from the recorded background of transactions are chosen during an abnormality identification to conduct calculation that is very close to the entry transactions.

Data needed to record the completed actions will be registered as a transaction in a graph type for any activity conducted by the customer. In general, data contained in a graph data system falls into four types for each entry operation: Data which are common in each of these transactions, data as for the transaction volume, data linked with point of transaction frequency, data associated to quantity of accomplished transactions by utilizer in a given period of time. There is a comparison unit named Compare Unit in the next step, whose main task is to obtain the ontology graph generated in the preceding step and then align it with current trends in Template DB. Having such approach offers three major benefits for the device. The first is that operational overload to fraud detection process will collapse due to involving just identical patterns with all of the patterns present in the component. The next strength is the real-time identification of new transactions in the event that some already reported suspicious transactions imitate.

©IJRASET: All Rights are Reserved

And the last one is that it exists for the discovery phase in favor to specific trends within a single entry transaction. The next unit is Measurement Unit whose main duty is to run key computations on transaction details entered and transactions deposited in related activity trends deposited in bank trends and whether a transaction is branded as fraudulent. If entry activity is defined as an outlier data graph trend of the analytical network, it is not processed until the computations have been completed.

IV. PROPOSED SYSTEM DESIGN

A Hidden Markov Model conjured up of a measurable range of states; each associated with a distribution of probabilities. Movement from one state to another transition among these states is ruled by a collection of probabilities called the transition probabilities. A potential outcome may be produced in a given state which is the sign of likelihood observation. It is not the states that are visible to the utilizer, but only the outcome through the states. It's claimed the states are hidden, and hence the term Hidden Markov System. Hidden Markov Model is therefore a good way to tackle fraud identification via credit card purchases. One big benefit of HMM-based deployment is a significant reduction in the amount of False Positives (transactions that are identified as fraudulent by a fraud detection program, even if they are genuine). HMM[6] primarily considers three price-value classes in this prediction process:

1) Low (l)
2) Medium (m)
3) High (h)

The values for the hidden states values is determined by S, observations value is given by K, A gives the transition probabilities, and B is the probabilities of observation state.

A. Detection of credit card theft using Hidden Markov model

Hidden Markov Model is able to track theft even in the absence of false signatures bearing in mind the cardholder's expense model. Any payment card fraud identification system operating either at the merchant's site or at the bank that provides credit cards usually does not recognize the specifics of the bought goods in single transactions. For authentication purposes, each transaction is loaded into the fraud detection program.

The fraud identification program makes use of credentials such as credit card code, cvv code, card sort, expiration and the quantity of items purchased to authenticate, whether the transaction is legitimate or not.

Based on the transaction amount of every transaction, the fraud detection system classifies them into clusters (low[l], medium[m] and high[h]) of training set which relies on the spending behavior. The type of items purchased is unknown. Although, the amount processed is known and used for further processing by the system. It tries to figure out any deviation in the purchases depending on the expense behavioral model of the cardholder.

If the fraud identification program discovers that the transaction is illegal, it creates an alert and aborts the transaction. In the event of a counterfeit passport, a protection type emerges. The identification form includes queries that need to be filled in, and are only accessible to the cardholder. When the recipient refuses to respond appropriately, the issuing bank will reject the purchase and the cardholder will be informed. This is important to maintain protection and dignity.

B. Approaches and Algorithms Used

To document the method of dispensing credit card transactions for Hidden Markov Model (HMM)[5] inspection symbols must be decided. We calculate the purchase interest X in the M price ranges V1, V2. VM, shapes the research symbols of the issuing branch. HMM calculates the price spectrum of each credit card purchase using clustering algorithms (like the K clustering algorithm).

Throughout this phase of estimation, it primarily considers three demand value levels including 1) small 2) medium and 3) large. It requires a minimum of 10 transaction sequences (attaches one additional transaction to the sequence on every entry). Through the initial step, the model will not have the details for the last 10 purchases, in which case the model would allow the cardholder to include specific information during the cardholder transaction.

©IJRASET: All Rights are Reserved

A HMM can be labelled as:

1) The model has N number of states. S is the collection of states, {S1, S2…SN}, where Si , i = 1, 2..., N is a unique state. qt is the condition at a moment t.

2) M is the amount of different symbols per entity that represent the actual outcome of the scheme. We account for the symbolsV={V1; V2;... VM}, where M is a specific sign, Vi, I= 1, 2…M.

3) The state transition likelihood matrix A= [aij], where aij = P (qt+1=Sj | qt=S), $1 \le i \le N$, $1 \le j \le N$; t=1, 2, 3…..

4) The observation symbol likelihood matrix B = [bj (k)], where bj (k) = P (Vk|Sj), $1 \le j \le N$, $1 \le k \le M$.

5) The probability vector of original state $\pi = [\pi i]$, where $\pi i = P (q1 = Si)$, $1 \le i \le N$.

6) Observation categorization O = O1, O2, O3...OR, where every observation Ot is a part of V, and R is the amount of observations.

All data verification must be reviewed before credit card fraud identification device load first tab. If the user's card contains fewer than 10 charges otherwise they can explicitly require that personal details be given with the purchase. When 10 transaction repository is established, then the card processing program can begin to operate.

V. CONCLUSION

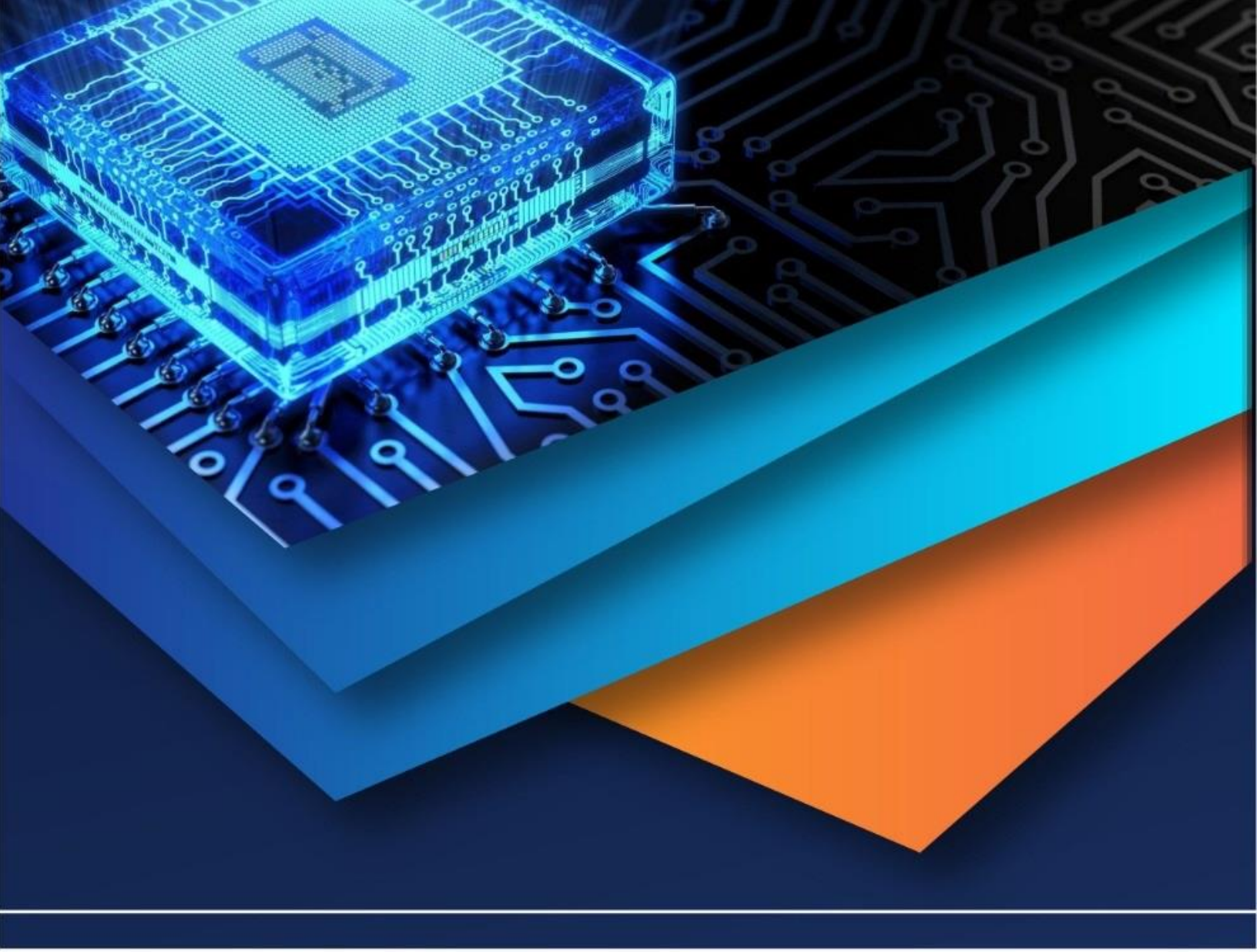
Within this study we suggested an implementation of HMM in the analysis of credit card fraud. The numerous phases in the handling of credit card purchases are described as the fundamental stochastic mechanism of an HMM. We used the transaction sum classes as the markers of measurement, while object forms were assumed to be HMM systems. We also recommended a framework for identifying the cardholders ' expenditure graph, as well as implementing this information in evaluating the significance of observation markers and initial estimate of project metrics. Also it was clarified how the HMM would determine whether or not an arriving request is bogus. Laboratory tests indicate our entire system efficiency and efficacy, and highlight the utility of studying the cardholder's expenditure pattern. Also, the framework is robust for managing high capital levels.

As the fraud gets detected by the model, as and then there can be a method where the cardholder gets notified either by a simple text or using SMTP (Simple Mail Transfer Protocol). This will help as an immediate response by the issuing bank to the cardholder by informing about the fraud or stating the card was used elsewhere.

REFERENCES

[1] "Parallel Granular Neural Network for Credit Card Fraud Detection", by Mubeena Syeda, Yan-Qing Zbang and Yi Pan.
[2] "Card-watch: A Neural Network based Database for mining system for Credit Card Fraud Detection", by Emin Aleskerov, Bernd Freisleben, and Bharat Rao.
[3] "A Web Services-Based Collaborative Scheme for Credit Card Fraud Detection" by Chuang-Cheng Chiu and Chieh-Yuan Tsai.
[4] "Neural Data Mining for Credit Card Fraud Detection" by Tao Guo, Gui-yang Li.
[5] "Credit Card Fraud Detection Using Hidden Markov Model" by Abhinav Srivatsava, Amlan Kundu, Shamik Sural, Arun K Majumdar.
[6] "Credit Card Fraud Detection Using Hidden Markov Model" by Shailesh S. Dhok, Dr. G. R. Bamnote.
[7] https://optinmonster.com/online-shopping-statistics/
[8] https://economictimes.indiatimes.com/industry/banking/finance/banking/credit-card-usage-rides-on-digital-push-grows27/articleshow/70580357.cms?from=mdr/
[9] https://shiftprocessing.com/credit-card-fraud-statistics/
[10] Credit Card Fraud Detection Using Hidden Markov Model" by Abhinav Srivastava, Amlan Kundu, Shamik Sural, Senior Member, IEEE, and Arun K. Majumdar, Senior Member, IEEE
[11] "Application of Hidden Markov Model in Credit Card Fraud Detection" by V. Bhusari, S. Patil.
[12] "Credit Card Fraud Detection System: A Survey" by Dinesh L. Talekar, K. P. Adhiya.
[13] "Credit Card Fraud Detection: A classification analysis" by Sonali Bakshi.

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)