International Journal of Computer Science and Business Informatics

IJCSBI.ORG

An Adaptive and Real-Time Fraud Detection Algorithm in Online Transactions

John Batani
ICT and Electronics Department
Chinhoyi University of Technology, P.Bag 7724, Chinhoyi, Zimbabwe

ABSTRACT

While the Internet has made it possible to transact electronically and ubiquitously, some unscrupulous internet users have devised ways of defrauding e-commerce users. Several solutions have been designed and deployed to try and curb fraud in electronic transactions, but the news of fraud in e-commerce continues making the headlines globally. It is against this background that the researcher was motivated to design an adaptive algorithm that can detect credit card fraud as it occurs (real-time). The solution is based on the use of an Artificial Neural Network, Hidden Markov Model and a One-Time Password. The researcher used a synthesised dataset since a real dataset could not be found. The researcher tested the algorithm, which produced 100 per cent fraud detection rate and 98 per cent accuracy. The proposed solution can be made a plugin to e-commerce sites for the purposes of detecting and preventing fraud. The researcher was motivated to undertake this study after realising that while Zimbabwe is calling for the adoption of e-commerce due to the prevailing cash crisis, some people still have reservations due to security concerns. Despite, even in those countries where electronic commerce was adopted a long time ago, security is still a concern among e-commerce participants. The designed algorithm has a learning ability so that it can detect new fraud variations as they occur (real-time) and thus terminate the transaction should it be considered a fraudulent one. The author seeks to restore and instill confidence in people who transact online using credit cards.

Keywords
Fraud detection, Real-time fraud, Adaptive fraud, E-commerce security, Credit card.

1. INTRODUCTION

The Internet has undoubtedly revolutionized the way we do business today. The increasing popularity of the Internet world over has seen the widespread acceptance, usage and adoption of electronic commerce in which transacting can occur without physical interaction, virtually from anywhere across the globe. However, an upsurge in electronic transactions has presented an opportunity for unscrupulous computer users to defraud unsuspecting victims who transact online. Despite several fraud detection solutions

having been designed and deployed, cases of fraud in electronic transactions continue making the headlines globally. This suggests that the current solutions have some weaknesses that are being taken advantage of by fraudsters. There is sort of an arms race between online fraudsters and those engaged in designing anti-fraud solutions, hence the need to come up with an adaptive solution that detects fraud in real time. A good strategy and the main goal of banks and industries is timely information on fraudulent activities [3].

.

1.1 The Research Problem

Despite several attempts having been made to curb online fraud, cases of people being defrauded online continue making the headlines world over. Notwithstanding the enormous growth in electronic transactions, there is a lack of strong security to the high end [3]. The rise in Internet usage and an upsurge in electronic transactions have seen an increase in cases of online fraud despite significant efforts by card issuers, merchants and law enforcement to curb the fraud. Since fraud perpetration techniques are evolving as technology is evolving, there is a need to come up with an adaptive solution that also has a capability of detecting fraud in real time. An early detection of fraud is more significant in terms of cost analysis [3]; hence the need to have real time detection solution.

1.2 Purpose Of The Study

The purpose of this study is to understand the shortcomings of current online fraud detection systems with the aim of designing and implementing an adaptive, hybrid and real-time online fraud detection algorithm. The algorithm should have a learning ability so that it can detect new fraud variations as they occur and thus terminate the transaction should it be considered a fraudulent one. The algorithm to be designed is supposed to restore and instill confidence in people who transact online

1.3 Objectives Of The Study

This study seeks to:

1. design an adaptive fraud detection algorithm that can detect credit card fraud in real time.
2. implement an adaptive fraud detection algorithm that can detect fraud in real time.

2. RELATED WORKS

"*Fraud can be defined as the illegal usage of any system or good*" [2], while the legal activities can correspondingly be termed legitimate. Therefore, credit card fraud can be defined as the illegal usage of a credit card to perform electronic transactions. Fraud detection is a complex computational

task and there is no system that surely predicts any transaction as fraudulent; rather the systems predict the likelihood of a transaction being fraudulent [1]. Basically, there are two types of credit card fraud, that is, counterfeit fraud and the illegal use of a lost or stolen credit card [2]. Broadly, there are four types of fraud, namely bankruptcy fraud, theft or counterfeit fraud, application fraud and behavioural fraud [4]. Several fraud detection techniques on electronic transactions have been designed, deployed and applied by various researchers and organisations to reduce further damage caused by fraud; however, people continue to be defrauded in spite of such efforts having been made. Despite the existence of numerous fraud detection technologies, it is not possible to detect fraud while the transaction is in progress [3], that is, in real time. Some of the fraud detection techniques that have been implemented include the Hidden Markov Model, decision trees, neural networks, Bayesian networks and data mining techniques [9]. According to [1], several researches have been done with an emphasis on data mining and neural networks. [5] applied unsupervised neural networks in credit card fraud detection. The Hidden Markov Model has previously been used in credit card fraud detection [10], thus it can be used for credit card detection. [10] define Hidden Markov Model as a finite set of states, each of which is associated with a probability distribution. The model only shows the result and hides the state from the external viewer, thus the states are 'hidden' to the outside and hence the name Hidden Markov Model. Bayesian Networks have also been used in an attempt to detect online fraud [3]. According to [3], a Naïve Bayesian classifier is an influential probabilistic method that makes use of class sequence from training class of prospect instances. Other techniques that have been applied to fraud detection include self-organizing maps (SOM), K-Nearest Neighbor [3], Outlier Techniques [6] and also the Boat algorithm. However, regardless of all such concerted efforts, the news of people being defrauded online continues making the headlines. This therefore is suggestive that the current solutions, regardless of them being numerous, have some shortcomings which are being exploited by online fraudsters. This is therefore what has prompted the researcher to consider this area as a possible research area.

3. METHODOLOGY

The researcher used the design science research methodology since the research sought to come up with an artifact. For problem awareness, the researcher performed document analysis and established that the problem really exists. The researcher also searched for trends of online fraud in different presses. Having understood the problem, the researcher proposed to have a hybrid, adaptive and real-time fraud detection system for credit card electronic transactions. The researcher then had a tentative design of the system. The tentative design was then refined, an algorithm designed

and system development was undertaken. The system was evaluated for performance using a test data set, the results of which are shown herein. This research contributes to the detection of online credit card fraud detection by proposing a solution that is adaptive and real-time. This is essential as it ensures that fraud is detected as it happens rather than after the transaction. Moreover, the use of neural networks ensures that the solution is adaptive, thus keeping abreast with changes in the ways in which online credit card fraud is perpetrated. This results in improved security of online transactions, hence, restoring and boosting confidence of credit card users in transacting electronically using their cards.

3.1 The Proposed Solution

The researcher designed a hybrid and adaptive fraud detection algorithm that can detect fraud in real time. It has come out from literature that the current solutions are not real time. Despite the existence of numerous fraud detection technologies, it is not possible to detect fraud while the transaction is in progress [3]. Therefore, the researcher designed an algorithm with the capability of real time fraud detection. The researcher intends to make use of neural networks, machine learning and other artificial intelligence approaches in coming up with the intended solution.

3.2 The Flow Chart of the Proposed System

The user first has to be registered on the e-commerce site, and every time they try to transact on the site, they are authenticated. After successful authentication on the website, the user can then enter their credit card details which will also be verified from the bank database. If the entered credit card details are incorrect or do not exist, the transaction halts, otherwise it proceeds. If the credit card details exist in the bank database and are correct, the system generates a One-Time Password (OTP) which is sent to the registered mobile number of the cardholder. The user is then prompted to enter the received OTP and if it matches the sent OTP, then the system extracts the user's social profile from the bank database. This social profile comprises age, income, occupation, and cardholder's value of assets. These social profile parameters will then be classified and assigned weights using Artificial Neural Networks to generate the cardholder's social status. The system then extracts credit card and bank transactions history and current balance (financial profile) from the bank database for analysis using the Hidden Markov Model (HMM). The HMM is used to generate the cardholder's financial status. Bank transactions involve checking the validity of the card, card holder's previous one-year transactions history, and balance in the account. Credit card transactions history involves credit card bill payments and spending patterns of the cardholder. The system uses the cardholder's financial status, social status and OTP for fraud detection. If authentication is successful (user login credentials, credit card details and OTP) and the transaction has unique features compared to the cardholder's

credit card spending history patterns, the user's financial profile is updated using ANNs and the transaction is processed. However, the transaction must be within the spending limits of the credit cardholder as stipulated by their bank. If the transaction has new features and exceeds the cardholder's spending limits, it is flagged as fraudulent and terminated.

The continuous updating of the cardholder's financial profile through ANNs makes the fraud detection system adaptive, while the use of OTP makes it a real-time fraud detection system. Combining ANNs and the HMM makes the system a hybrid solution.

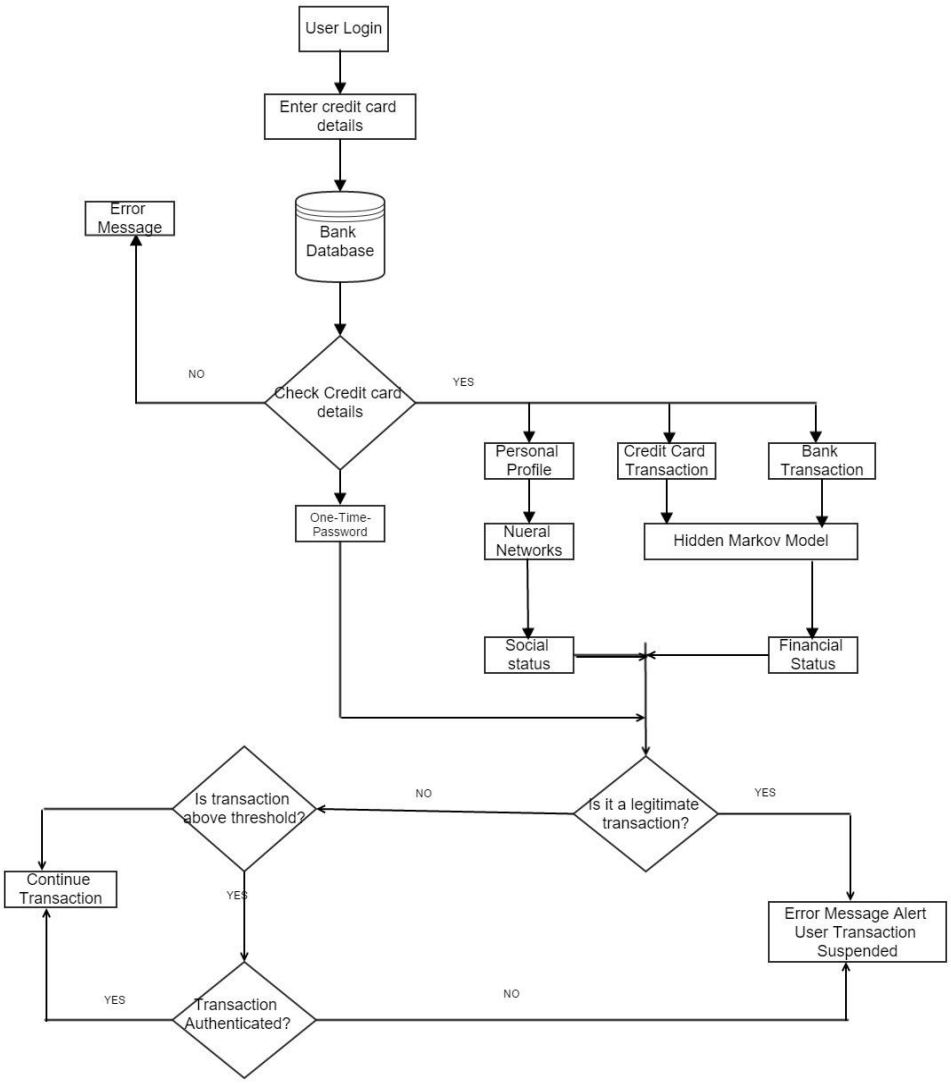

Figure 1. System flow chart

3.3 Algorithm for the Proposed System

Let InT be incoming transaction, PrT be previous transactions and OTP be one-time password, FrT be fraudulent transaction and LgT be legitimate transaction

Steps

I. Receive login details
II. If login details are valid Then
    a. Send OTP
    b. Compare received OTP and sent OTP
        i. If received OTP !=sent OTP Then
            1. Terminate transaction and End
        ii. Else
            1. Accept credit card details
            2. Authenticate credit card details
            3. If credit card details are invalid Then
                a. Terminate Transaction & flag it as suspicious
            4. Else
                a. Extract cardholder's social profile {name, age, gender, income, occupation, value of assets}
                b. If cardholder name or gender or age !=user's name or gender or age Then
                    i. Terminate the transaction and flag it as fraudulent
                    ii. End
                c. Extract cardholder's financial profile {card balance, card transactions history, card bills history, card payments history, credit limit}
                d. If InT value > credit limit Then
                    i. Flag InT as fraudulent
                    ii. End
                e. Classify social profile parameters using ANNs and assign them weights
                f. Analyse financial profile using HMM to generate financial status
                g. If InT has new features
                    i. Update cardholder's spending patterns using ANNs
                h. If InT resembles PrT Then
                    i. InT is LgT

ii. Commit InT to legitimate transactions database
iii. LgT + =1
iv. Notify user and bank of the transaction
v. End

i. Else
i. InT is fraudulent
ii. Record transaction as FrT
iii. FrT += 1
iv. Notify cardholder and bank
v. End

III. Else
a. Terminate transaction

IV. **END**

Neural Network Algorithm

Let:

InT be incoming transaction

LegT be legal transaction

FrT be fraudulent transaction

SusT be suspicious transaction

MaxOTP_Time be the maximum number of seconds within which a user is supposed to enter an OTP

Time_Elapsed be time in seconds that has elapsed after an OTP has been sent and before user has entered the corresponding OTP.

InT_Features{}be a set of features for incoming transaction for customer

LegT_Features{} be a set of features for legal transaction for customer

FrT_Features {} be a set of features for fraudulent transactions for customer

NewFeatures{} be a set of new features from incoming transaction that do not exist in the current dataset of legal transactions and fraudulent transactions

Define Variables:

Threshold {user defined variable in the form of a percentage, which is used to compare similarity between incoming Transaction and

legal transaction for each customer. It can be between 0 and 1 inclusive, or a percentage}

Step 1: Extract attributes of InT and store them in InT_Features{}

Step 2: Compare InT_Features{} and LegT_Features{}

If similarity >= threshold Then

InT is LegT

Else

InT is SusT

Suspend InT

Send OTP to cardholder's mobile number

If Time_Elapsed = MaxOTP_Time Then

InT is FrT

Extract new features from InT and update FrT_Features{} for customer

Else

InT is LegT

Extract new features from InT and update LegT_Features{} for customer

End If

End If

4. RESULTS

The results of the system were evaluated in terms of the extent to which it correctly classified fraudulent transactions. In other words, the system was mainly evaluated on fraud detection rate. The system was tested using synthetically generated data since the researcher could not get a real dataset. However, the data was synthesised in a way that it as much as possible resembled a real dataset. Herein, a positive is a fraudulent transaction and conversely, a negative is a legal transaction. Thus, a legal transaction that is wrongly classified as fraudulent is a false positive; and a fraudulent transaction that is misclassified as legal is a false negative. Therefore, legal transactions that are correctly classified as legal are true negatives, and fraudulent transactions that are correctly classified as fraudulent are true positives. The system was thus evaluated in line with these parameters.

Table 1. Training Results

No. of Iterations	No. of Transactions	No. of False Positives	No. of True Positives	No. of False Negatives	No. of True Negatives	No. of Positives	No. of Negatives
50	50	2	25	0	23	25	25
100	50	0	25	0	25	25	25
150	50	1	25	0	24	25	25

Table 2. Testing Results

No. of Transactions	No. of False Positives	No. of True Positives	No. of False Negatives	No. of True Negatives	No. of Positives	No. of Negatives
50	1	25	0	24	25	25

From the testing results, the fraud detection rate can thus be calculated using the formula:

$$Fraud\ detection\ rate = \frac{Number\ of\ true\ positives * 100}{Total\ number\ of\ fraudulent\ transactions}$$

$$= \frac{25 * 100}{25}$$

$$= \mathbf{100}\ \%$$

$$Accuracy = \frac{(No.\ of\ True\ Positives + No.of\ True\ Negatives) * 100}{True\ Positives + True\ Negatives + False\ Positives + False\ negatives}$$

$$= \frac{(No.of\ True\ Positives + No.of\ True\ Negatives) * 100}{No.of\ Positives + No.of\ Negatives}$$

$$= \frac{(25 + 24) * 100}{25 + 25}$$

$$= \mathbf{98}\%$$

Sensitivity is regarded as the most important measure of the effectiveness of fraud detection systems [8]. This is because losses of money by defrauded people are a result of fraud detection systems failing to detect fraud. While this system's detection rate was perfect, it should be noted that its perfectness depends on some factors. Firstly, the system is deployed to work with electronic commerce transactions on which users are required to register. The system blocks all transactions which are initiated by users whose demographic details are different from the cardholder's as stored in the bank's database. For a transaction to be successful, the registered details on the website for the user must match with those for the cardholder, and the initiator of the transaction must have the cardholder's mobile phone to which an OTP will be sent. If all those conditions are met, a transaction will sail through even if some of its features are not found in the legal transactions set for that customer. The new features of an incoming transaction will then be added to either the legal transactions or fraudulent transactions features set for the customer, thus making the solution adaptive or incremental since it has an ability to learn new features associated with fraud.

The system does not only detect but also prevents fraud by prematurely terminating any transaction should the user fail to enter the generated OTP which is sent to their registered mobile phone number via Short Messaging System (SMS). Consequently, it achieves both credit card detection and prevention.

In comparison to other systems in existence, the proposed algorithm is highly effective with 100 per cent fraud detection rate. According to [7], the most effective credit card fraud detection system, which is the Dempster and Shafer Theory and Bayesian Learning, has 98 per cent fraud detection rate. Bayesian and Neural Network, Hidden Markov Model, and Hybridization of BLAST-SSAHA have fraud detection rates of 77, 70 and 86 percent respectively [7]. Consequently, the algorithm proposed herein stands out as the best. In terms of accuracy, the proposed algorithm has a very high accuracy of 98 per cent. However, [7] do not specify the numerical accuracy of Dempster and Shafer Theory and Bayesian Learning; Bayesian and Neural Network; Hidden Markov Model; and Hybridization of BLAST-SSAHA but simply specify their accuracy as high, medium, medium and high, respectively.

5. CONCLUSIONS

The proposed solution to fraud detection puts more emphasis on ensuring customer security in electronic transactions using credit cards. The strengths of the proposed algorithm include a high accuracy of 98%, a high fraud detection rate of 100%, ability to learn new fraud variations and new customer spending patterns (adaptive), and ability to detect and prevent fraud as it occurs (real-time). While the system has a high fool proof or fraud detection rate, it should be noted that a customer cannot transact even if they are the cardholder unless they have their registered mobile phone number and have registered their details on the website just as they are captured in the credit card issuer's database. The researcher's view is that security matters more than convenience, hence one cannot transact without their registered mobile number. An OTP is sent to a mobile phone via text messaging and not by electronic mail (e-mail) since e-mail can be easily hacked. The other reason for choosing to send an OTP via SMS was that a recipient is not charged for receiving an SMS in the author's country of residence. Prospective users of this algorithm may also decide to send an OTP to an e-mail rather than via SMS. The effectiveness of the proposed algorithm may be compromised if a criminal gets access to the legitimate cardholder's registered phone number, and e-commerce website authentication credentials.

REFERENCES

[1] Dheepa V. and Dhanapal R., "Analysis of Credit Card Fraud Detection Methods," *International Journal of Recent Trends in Engineering*, vol. 2, no. 3, pp. 126-128, 2009.

[2] Ekrem D. and Hamdi O.M, "Detecting credit card fraud by genetic algorithm and scatter search," *EXPERT SYSTEMS WITH APPLICATIONS*, vol. 38, no. 10, pp. 13057-13063, 2011.

[3] Gayathri R. and Malathi A., "Investigation of Data Mining Techniques in Fraud Detection: Credit Card," *International Journal of Computer Applications*, vol. 82, no. 9, pp. 12-15, 2013.

[4] Linda D., Hussein A., and Pointon J., "Credit card fraud and detection techniques: a review," *Banks and Bank Systems*, vol. IV, no. 2, pp. 57-68, 2009.

[5] Ogwueleka F.N, "Data mining applications in credit card fraud credit card fraud detection system," *Journal of Engineering Science and Technology*, vol. 6, no. 3, pp. 311-322, 2011.

[6] Rama K. K and Uma D. D., "Fraud Detection of Credit Card Payment System by Genetic Algorithm," *International Journal of Scientific & Engineering Research*, vol. 3, no. 7, pp. 1-6, 2012.

[7] Rana P.J and Baria J., "A Survey on Fraud Detection Techniques in Ecommerce," *International Journal of Computer Applications*, pp. 5-7, 2015.

[8] Seeja K.R and Zareapoor M., "FraudMiner: A Novel Credit Card Fraud Detection Model Based on Frequent Itemset Mining," *Scientific World Journal*, vol. 2014, 2014.

[9] Sharma A. and Panigrahi P.K, "A review of financial accounting fraud detection based on data mining techniques," *International Journal of Computer Applications*, vol. 39, no. 1, pp. 37-47, 2012.

[10] Singh A. and Narayan D., "A Survey on Hidden Markov Model for Credit Card Fraud Detection," *International Journal of Engineering and Advanced Technology*, vol. 1, no. 3, pp. 49-52, 2012.

This paper may be cited as:

Batani, J. 2017. An Adaptive and Real-Time Fraud Detection Algorithm in Online Transactions. *International Journal of Computer Science and Business Informatics, Vol. 17, No. 2, pp. 1-12.*