A Review of Credit Card Fraud Detection Techniques in Electronic Finance and Banking

CHILAKA, U. L.[1], G. A. CHUKWUDEBE[2], BASHIRU A.[3]
[1, 3] *Department of Computer Science, Imo State University, Owerri*
[2] *School of Computing and Information Technology, Federal University of Technology, Owerri*

Abstract- The growing innovation in technology has been exploited in the finance/banking sector in the form of electronic commerce (e-commerce). Financial transaction is the backbone of global market. Rather than make transaction conventionally, cashless transaction is being done in recent times using credit cards. The credit card transaction uses electronic technology to make commerce easy, and its use for such purposes has increased. However, credit card being the most common means of payment in the society today, has witnessed increased number of fraud cases. In order to reduce or solve the problems of fraud cases related to credit card transactions, several methods have been proposed and implemented in literature and real-world. This paper discussed some of the recent fraud detection techniques that have been developed to identify and detect fraudulent cases of credit card transactions. Since fraud detection consist of detecting fraud as soon as possible once it is done, the various models developed are built to provide fast response, accuracy, high degree of sensitivity, effectiveness and efficiency. These techniques have their individual benefits and limitations. However, it is recommended that a hybrid model with robust and optimal performance that combines various be used for fraud detection in finance/banking. The paper has also presented observed challenges of credit card techniques.

Indexed Terms- Credit card transaction, cashless transaction, e-commerce, Fraud detection techniques

I. INTRODUCTION

The use of credit card for transactions has continued to witness increasingly growth in number. It is taking a larger proportion of payment system all over the world and at the same time resulting to an increased rate of stolen account numbers and subsequent losses by financial institutions [1]. As the most popular mode of payments [2], and with the number of users increasing globally daily, a corresponding increase in identity theft and credit card fraud has become common.

The two types of purchases based on credit card are physical card purchase (offline purchase) and virtual card purchase (online purchase). In a physical card purchase, the individual personally presents the card to make a payment. In order to be able to do physical card purchase, a credit card criminal will have to steal the card and fake the signature. In the case of virtual card purchase, only the detailed information of the card is required like card number, date of expiration, secure code, etc. Purchase of this kind is usually carried out online or over telephone [2].

The growing application of modern technology in financial/banking sector has subsequently led to increase in financial fraud. This has also witness increased credit card fraud on daily basis. In fact, as the amount of money being lost due to credit card attacks is growing, its usage is even common. Security is essential for safety and fraud prevention. It helps to ensure that credit card is used safely and prevent fraud occurrence. Issues such as lost cards, stolen lost cards, stolen cards, application fraud, counterfeit fraud, mail-order fraud and non-receive issue (NRI) fraud are usually found credit card fraud cases [2]. Reducing these frauds requires proper and adequate security.

As the use of e-payment system is increasing in the society couple with the rapid advances of e-commerce on the internet, using credit card has become convenient and necessary [1]. However, to be safe, credit card users should ensure that card details are kept private. The privacy of the credit card must be secured. Jayant et al [2] highlighted different means of stealing credit card details from unsuspecting user to include phishing websites, steal/lost credit cards, counterfeit credit cards, card details theft, and intercepted cards and so on. Actually, most of the duplicitous transactions come from stolen card numbers rather actual card theft [2]. As such, the most important way to be secured is to keep credit card safe.

In order to ensure safe and secured financial transactions using credit cards, several fraud detection methods have been developed. Since credit card fraud is becoming increasing high, improved fraud detection technique has become essential to ensure that payment system are kept viable. Fraud detection techniques are developed to prevent criminals from carrying out illegitimate businesses. There are several fraud detection methods that have been developed such as data mining, machine learning, sequence alignment, genetic algorithm, fuzzy logic, and artificial intelligence etc. [3].

A. Categories of credit card frauds

There are different types of credit card frauds. Bhatla et al [4] presented the following as the three major categories of credit card frauds: traditional card related frauds, merchant related frauds and internet related frauds.

1) Traditional Card Related Frauds

The traditional card related fraud consists of application fraud, lost/stolen cards, account takeover, and fake and counterfeit cards.

- Application fraud takes place when an individual misrepresents an application to get a credit card. It can be committed in three ways: assumed identity –in this case a person unlawfully acquires information that belongs to another person and creates an account in his/her name, making use of information that is somewhat authentic [4]. Financial fraud –a person gives out incorrect information about his/her financial status to obtain credit. Not-received items (NRIs) –also called postal intercepts take place in situation whereby a credit card is stolen from postal service before it gets to its owner [4].
- Lost/stolen cards fraud occurs in situation whereby a credit card owner loses it or someone steals the card for fraudulent purposes. According to Bhatla et al [4], this type of credit card fraud is actually the most informal way for fraudsters to obtain other individual’s card illegitimately.
- Fake and counterfeit cards fraud arises from the creation of forged credit cards. Bhatla et al [4] highlighted some of the approaches employed in creating fake cards to include: erasing the magnetic strip, creating a fake card, altering card details, skimming and white plastic.

2) Merchant Related Frauds

This type of fraud is facilitated either by owners of merchant institution or their employees [5]. Merchant related frauds are of two types: merchant collusion – this fraud takes place when owners of merchant or their personnel collude with fraudsters to carry out fraud using the cardholders’ accounts or by using the cardholder personal details. Triangulation –this fraud is carried out and operated from a website. According to Saravanan and Babu [5], in this fraud case, products or goods are presented at high discounted rates and also conveyed before payment. While browsing the site and finds a product he likes, the unsuspecting customer place personal information such as name, address and valid credit card details on the site. On receiving these details, the fraudsters order goods from a genuine site using stolen credit details. The product is purchased using the stolen credit card information.

3) Internet Related Frauds

The simplest and the easiest way to execute fraudulent transactions by fraudsters is the internet. The growth in technology has led to expansion in trans-border, economic and political spaces which have made the internet to become a new global market, bringing sellers and buyers together from all over the regions and countries in the world. Among the techniques commonly used in internet fraud are: site cloning and false merchant sites, and credit card generators.

In this paper, a review of various modern methods for credit card fraud detection is presented. A collection of previous research works on credit card fraud detection techniques from 2014-2019 was considered. The paper is divided into five (5) sections. With the introduction section completed, the remaining four (4) sections include survey of previous literature, existing techniques, observed challenges of credit card fraud detection techniques, and conclusion.

II. SURVEY OF PREVIOUS LITERATURE

Zareapoor and Shamsolmoali [6] presented application of credit card fraud detection based on bagging ensemble classifier. The study examined the performance of some data mining methods, which are Support Vector Machines (SVM), Naïve Bayes (NB) classifiers, K-Nearest Neighbour (KNN) Algorithm, and Bagging Ensemble classifier, in detecting credit card fraud. The ensemble was constructed using bagging classifier with the decision tree algorithm J48 based on the C4.5 model. Performance evaluation of the various methods was carried out in terms of Fraud Catching Rate, False Alarm Rate, Balanced Classification Rate and Matthews Correlation Coefficient using dataset obtained from real world credit card dataset. The authors compared the performance of bagging ensemble classifier with a number of standard classifiers. The authors maintained that with other methods having problem of increasing false alarm rate in detecting fraudulent transactions, bagging ensemble classifier performed very well in detecting fraudulent transactions by keeping the fraud catching rate high while ensuring very low false alarm. Also, the bagging ensemble classifier technique was capable of handling class imbalance.

Patil et al. [7] studied predictive modeling for credit card fraud detection using data analytics. A big data analytical structure for processing large volume of data was proposed. Real time data extraction was performed from different sources. Analytical model was developed using the extracted data from German credit card fraud dataset which consist of twenty (20) attributes such that seven (7) are numerical attributes and thirteen (13) are categorical attributes with almost one thousand (1000) transactions. The developed analytical mode was used to determine validity of the incoming transaction. Two machine learning algorithms, logistic regression and decision tree. The algorithms were implemented on credit card banking data set. While the logistic regression was used for classification of fraud detection, ID3 technique was used to construct decision tree considering entropy of dataset. The authors further used random forest decision tree to solve the problem of regression and classification. The random forest algorithm used pseudo-code to carry out prediction of fraudulent transaction. From the model evaluation performed based on the test experiment, it was observed that the random forest model indicated better performance compared to logistic regression and decision tree in terms of accuracy, precision and recall parameters.

Agarwal and Upadhay [8] presented a fast fraud detection approach using clustering based method. A hybrid approach that combined clustering based, distance based and outlier detection techniques was proposed to find credit card fraud activities. The structure of the proposed scheme is shown in Fig. 1. The clustering based technique was used to group data having similar features and subsequently act as data reduction approach. The distance based technique was used to calculate maximum distance value for each cluster, and if the maximum distance was greater than some threshold provided by the user then it would regard as an outlier otherwise an inlier. The outlier detection technique was used to find objects that were dissimilar and inconsistent with regard to remaining data or data which were far away from the centroids of their cluster. Dataset was obtained by simulating a large number of different transactions while purchasing. The proposed hybrid approach was only used for numerical, and it was observed that it provided reduced computational time. However, the authors recommended that future work should implement the hybrid approach on more complex dataset sand varying datasets.

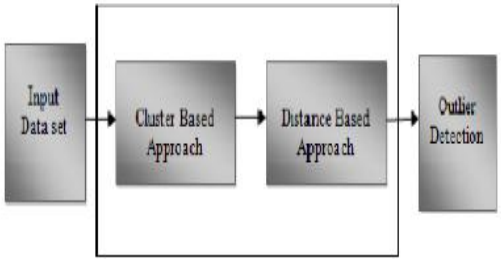

Fig. 1 System architecture [8]

Pushpalatha and Joseph [9] presented credit card fraud detection based on the transaction by using data mining techniques. The study examined data mining methods like Bayesian networks, Bayes Minimum Risk, Genetic algorithm, Hidden Markov Model (HMM) and Ontology for credit card fraud improvement. It then focused on improving current fraud detection techniques by enhancing fraudulent accounts' prediction. The findings indicated that a learning strategy in conjunction with a standard fraud detection technique could provide improved fraud detection.

Sodasoltaniziba and Alibalafr [10] reviewed data mining techniques for fraud detection. Annual transactions related to 20, 000 account number of financial institutions was studied using service analyses software. Clustering clients based on client type was proposed. Each of the clusters was assigned an appropriate rule which was determined by the performance of group member in case of deviation from specified performance. A decision tree algorithm was developed using the rules of C5. It was observed that the proposed model extracted a lot of the rules related to client performance.

Lepoivre et al [11] presented credit card fraud detection with unsupervised algorithms. A model for credit card fraud detection was proposed to satisfy calculation simplicity and operation transparency. Two unsupervised algorithms, Principal Component Analysis (PCA) and SIMPLEKMEANS algorithm were developed to consider geographic location of both transactions and clients. The authors claimed that the proposed method directly classifies the transactions with good precision and could detect new fraudulent activities. PCA offered a complete view of relations among various features and at the same time flexible.

Pouramirarsalani et al [12] proposed a hybrid feature selection and genetic algorithms for fraud detection in e-banking. Reinforcement learning in the neural network was used in the developing the proposed technique. Whale algorithm was also studied. The proposed technique was compared to whale algorithm. The results obtained indicated that proposed solution was very effective for fraud detection in e-banking.

Carminati et al [13] developed a semi-supervised online banking fraud analysis and decision support system called BANKSEALER. The proposed system characterizes the users of the online banking operation by means of a local, a global and a temporal profiling, which were developed during a training phase. The architecture for the proposed system is shown in Fig. 2. A decision support system was developed in collaboration with a large national bank where it was deployed as a pilot project. The BANKSEALER offered effective online banking for semi-supervised and unsupervised fraud and anomaly detection. The proposed system offers better alternative for fraud analysis and decision support unlike existing unsupervised and semi-supervised techniques which do not give the analyst a motivation for the analysis result, making manual investigation and as such making confirmation more difficult.

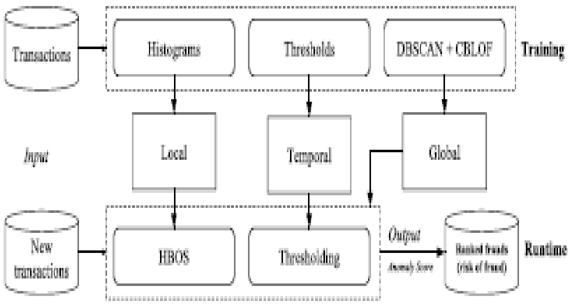

Fig.2 Structure of BANKSEALER [13]

Randhawa et al [14] studied credit card fraud detection using AdaBoost and majority voting. The study deployed twelve machine learning algorithms in combination with the AdaBoost and majority voting techniques. Empirical analysis was carried out using certain standard models which include Naïve Bayes (NB), Decision Tree (DT), Random Tree (RT), The Random Forest (RF), Gradient Boosted Tree (GBT), Decision Stump (DS), Multilayer Perceptron (MLP) network, Feed-Forward Neural Network (NN), Deep Learning (DL), Linear Regression (LIR), Logistic Regression (LOR), and Support Vector Machine (SVM). Experiments were carried out using RapidMiner Studio 7.6. It adopted the Matthews Correlation Coefficient (MCC) for performance measure. Evaluation was performed using a set of data from real credit card of a financial institution. An initial score of 0.823 for the MCC was obtained using majority voting. An MCC score of 1 was achieved using AdaBoost and majority voting techniques. In order to further evaluate the hybrid models to ascertain the robustness of the machine learning algorithms, all sampled real-world data were corrupted with noise at 10%, 20% and 30%. The results indicated that the majority voting technique provided the best MCC score of 0.942 for 30% noise added to all data features.

Rajaei [15] examined fraud identification on in banking data and financial institutions using classification algorithms. The study deployed a method for fraud detection in banks and financial institution. It developed a three-layer perception neural network algorithm for fraud detection. The proposed system was implemented and tested using dataset of a German credit card. The authors maintained that the use of data mining and classification algorithms made it possible to detect fraud with minimal error, without deploying human element and smart models.

Rajamani and Rathika [16] discussed the use of Hidden Morkov Model (HMM) and Neural Networks (NN) for credit card fraud detection. A comprehensive study of HMM and NN as an efficient way for detecting credit card fraud was presented. Fig. 3 and 4 show the architecture of HMM and NN algorithms in credit card transactions.

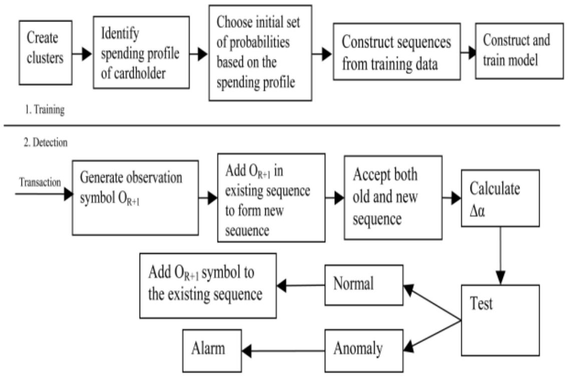
Create clusters
Identify spending profile of cardholder
Choose initial set of probabilities based on the spending profile
Construct sequences from training data
Construct and train model
1. Training
2. Detection
Transaction
Generate observation symbol O_{R+1}
Add O_{R+1} in existing sequence to form new seauence
Accept both old and new sequence
Calculate Δα
Test
Normal
Anomaly
Add O_{R+1} symbol to the existing sequence
Alarm

Fig. 3 Structure of HMM in credit card transaction [16]

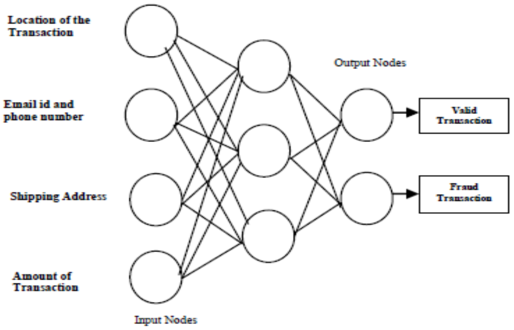

Fig.4 Structure of NN in credit card transaction [16]

Sarno et al. [17] presented hybrid association rule learning and process mining for fraud detection. The study developed a hybrid technique between association rule learning and process mining. The association rules were deployed to automatically filter fraud related activities in the testing dataset. The association rules combine positive and negative association rules. While the positive association rules was used for capturing illegal transactions, the negative association rules ensured that legal transactions captured as fraud by the positive association rules were filtered in order to improve accuracy. Fig. 5 is the structure of the proposed fraud detection method. The experiment was carried out to evaluate the accuracy in specified minimum confidence values. The result from the evaluation conducted indicated that the proposed methods achieve certain value of minimum confidence. The authors maintained that the proposed system offered better than that of process-mining technique since it has les falsely detected frauds.

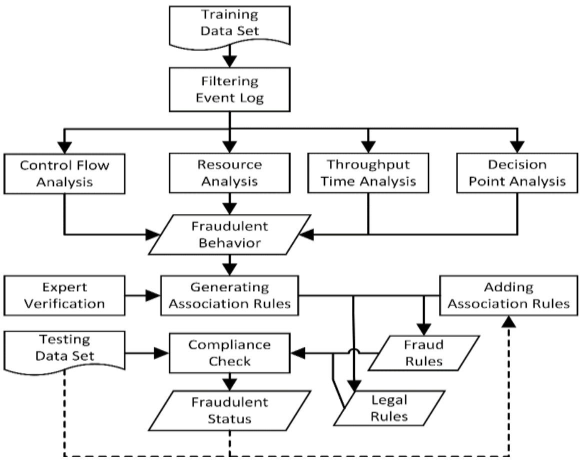

Fig. 5 Proposed system architecture for fraud detection [17]

Baboo and Preetha [18] presented analysis of spending pattern on credit card fraud detection. It developed a system that detects fraudulent credit card activities on internet transactions using Hidden Markov Model (HMM). The spending profile of credit card user was divided into three categories consisting of lower profile, middle profile and higher profile. The HMM algorithm was meant to detect and analyse the spending profile of the credit card user. The performance and effectiveness of the proposed method was demonstrated through recent transactions done. The structure of the proposed system is shown in Fig. 6.

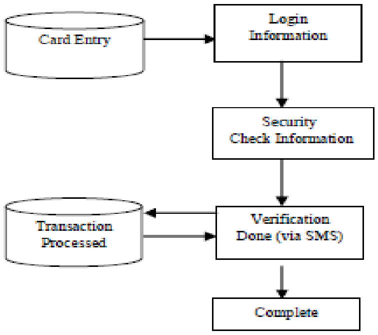

Fig. 6 Structure of proposed system of spending pattern on credit card [18]

Seeja and Zareapoor [19] proposed an intelligent credit card fraud detection model for detecting fraud from highly imbalanced and unidentified credit card transaction datasets. The proposed fraud detection system called FraudMiner, is shown in Fig. 7. It deployed fraud transaction patterns for each customer using frequent itemset mining as well as finding legal to solve the problem of class imbalance. A matching algorithm was developed to determine which pattern (legal or fraud) of the incoming transaction of a customer was closer and s decision made consequently. The authors maintained that the proposed system used took very less time in fraud detection.

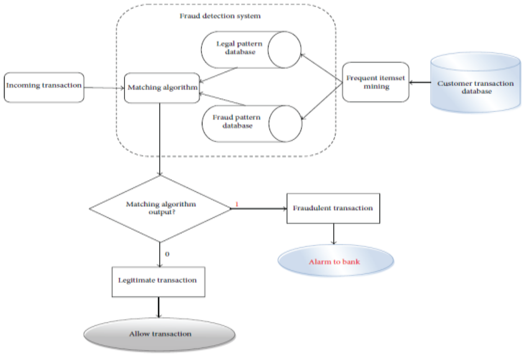

Fig. 7 Credit card fraud detection model [19]

Malini and Pushpa [20] studied credit card fraud identification techniques based on KNN and outlier detection. The study implemented KNN algorithm and outlier detection techniques to obtain an optimal solution in addressing credit card fraud detection problem. These methods were shown to minimize the false alarm rates and increase the rate of fraud detection.

Dai et al [21] addressed online credit card fraud detection using a hybrid framework with big data technologies. The study focused on designing an online credit card fraud detection system so as to realize three main objectives that include the ability to combine several detection algorithms to improve accuracy, to process large amount of data, and to carry out read time detection of credit card fraud. A general workflow that satisfied the most recent credit card fraud detection systems was proposed. Implementation using latest big data technologies such a Hadoop, Spark, Storm, Hbase, of the proposed workflow with a new framework having four layers made up of distributed storage layer, batch training layer, key-value sharing layer and streaming detection layer was carried out. The use of four layers proved to aid large trading that storage, fast detection model training, quick model data sharing and real-time online fraud detection. A prototype of the designed system was implemented and tested with a synthetic dataset, which indicated the quality of the study. The structure of the proposed workflow is shown in Fig. 8.

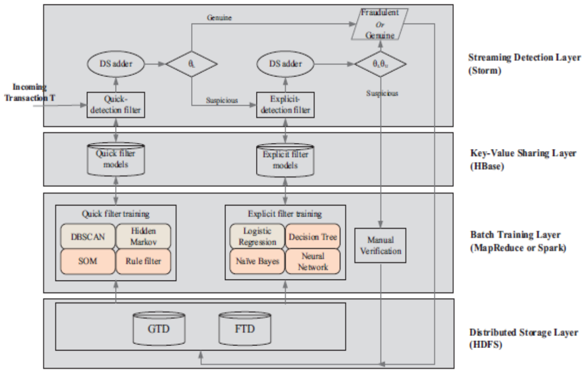

Fig. 8 Hybrid framework for CCFDS workflow [21]

Banerjee et al [22] examined the comparative analysis of machine learning algorithms through credit card fraud detection. Various classification algorithms trained on a public dataset to analyse correlation of certain factors with credit card fraud was studied. It proposed improved metrics for determining false negative rate and measured the efficiency of random sampling to reduce the dataset imbalance. It eventually described the Support Vector Machine algorithm as the most effective algorithm to utilize in datasets with high class imbalances since it provided the highest performance rate for credit card fraud detection under realistic conditions.

Khare and Sait [23] studied the use of machine learning models and collating machine learning models. The performance of various models like decision tree, random forest, support vector machine (SVM) and logistic regression on highly skewed dataset of credit card fraud was examined and determined. It used credit card transactions dataset obtained from European cardholders containing 284, 786 transactions. These models were applied on the raw and preprocessed data. The performance evaluation of the various models was carried out in terms of accuracy, sensitivity, specificity, and precision. The outcome of the evaluation indicated that an optimal accuracy of 97.7%, 95.5%, 98.6% and 97.5% was provided by logistic regression, decision tree, random forest and SVM algorithms respectively. The structure of the system is shown in Fig. 9.

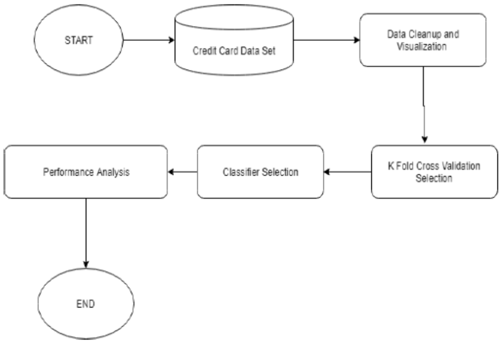

Fig. 9 Credit card fraud detection architecture [23]

III. CREDIT CARD FRAUD DETECTION TECHNIQUES

There are several techniques that have been developed for credit card fraud detection in electronic finance and banking. Some of the techniques that have been used for credit card fraud detection are presented in this section.

Bayesian Networks: In this algorithm each variable is shown in a specified domain as a node in the graph. The relationship between these variable are arcs connecting the respective nodes [9]. In carrying out fraud detection, two Bayesian networks to define the behaviour of user are developed. First, Bayesian network is developed to model behaviour based on the notion that the user is fraudulent (F) and the other model is built assuming that the user is legitimate (non-fraudulent, NF). An expert knowledge is used to set up the 'fraud net'. The data from non -fraudulent users is used to set up 'user net'. All through operation user net is adapted to a particular user based on present data. By inserting evidence in these networks and propagating it through the network, the probability of the measurement less than is obtained. This indicates the degree by which the observed user behaviour should meet typical F or NF behaviour [2] [9].

It makes provision for the integration of expert knowledge that is used for initial set up in the models. Conversely, the user model is retrained in an unsupervised method using data. Hence, Bayesian method combines both, expert knowledge and learning [2] [9].

Gass Algorithm: This algorithm is a combination of genetic algorithm and scatter search [2]. The basic concept is that the possibility of survival for the stronger members of a population is larger than that of the weaker members and as the generations increases the average fitness of the population gets improved. The approach is such that the fittest members are selected as the parents for the subsequent generation while the members of the generation that are less fit are disregarded. This process is repeated until the finest solution is obtained.

Genetic Algorithm: It is a programming technique that deploys the genetic evolution as a problem solving algorithm [12]. This algorithm takes the following processing steps [12]. First, a primary random population is created, and then each of the members of this population is created. Each of the members of the population has a solution for problem solving. Second, evaluate the solutions. This is carried out by the target function. Some values are allocated to the solutions by the target functions in relation to the study challenges which can be efficiency, security, and other factors in the system [12]. Third, perform crossover operation and then mutation operation. These two operations are carried out to prevent the premature convergence and construction of divergence in solutions.

Hidden Markov Model (HMM): This is probably the simplest models that can be used to model sequential data. In HMM, the state is not directly visible; this is in contrast to Markov model wherein the state is directly visible to the observer. However, in HMM the output, dependent on the state, is visible. "An HMM is a double embedded random process with two different levels, one is hidden and other is open to all" [16] [9]. The HMM is a finite set of states such that each state is assigned a probability distribution. Transitions among the states are governed by a set of probabilities called transition probabilities. In a given state an outcome or observation can be produced, in line with the associated probability distribution. It is only the outcome, not the state visible to an external observer and therefore states are "hidden" to the outside [16] [9]. It offers "large reduction in the number of False Positives transactions acknowledged as malicious by a fraud detection system even though they are categorically genuine" [16].

Neural Networks (NN): This is a model with layers or structures of neurons that are connected in regular pattern. This pattern can be drawn from genuine regular activities of the credit card user. It is developed by arranging nodes into layers and associates these layers with adaptable weighted interconnections. According to Rajamani and Rathika [16] a neural network is a group of "processing nodes" that transfer activity to one another via connections. A node receives an input from interconnected nodes and uses the weights of the linked nodes together with easy function for calculation of output values. This algorithm can be constructed for supervised and/or unsupervised learning [16].

Support Vector Machine (SVM): This is a statistical learning method that is particularly appropriate for binary classification technique [6] like credit card fraud detection approach in which only two classes are required. That is fraudulent and non-fraudulent category. The SVM can solve both classification and regression data [14]. Data samples are represented as points in the space mapped in such a way that data different categories of data samples can be isolated by a margin wide as possible. The advantage of SVM is its ability to tackle nonlinear classification problems; and requires less computational power, which is proper for real-time operation [14].

Linear Regression: This is a machine learning algorithm that models the relationship between scalar variables by assigning a linear equation to the output data. Linear predictor functions are used to model the relationships, with unknown model parameters estimated from the dataset [14]. The model selection use is the Akaike criterion, which is a degree of relative goodness of fit for statistical model. The benefit of the linear regression is that it offers optimal output in a situation where the relationship between independent and dependent variables are nearly linear [14].

Logistic Regression: This is a type of probabilistic statistical classification model that uses logistic curve for detection of fraud. It is a supervised classification technique that returns the likelihood of binary dependent variable that is predicted from the independent variable of dataset [23]. A value between 0 and 1 is given by the logistic curve such that it is said to be the probability of class membership [7].

K-Nearest Neighbour: This is a simple algorithm that stores all available instances and then classifies any new instances in terms of a similarity measure. The K-Nearest Neighbour (KNN) algorithm is an instance based learner [19]. In this technique, every single new instance is compared with existing ones by deploying a distance metric, and the nearest existing instance, known as the nearest neighbour, is used in assigning class to the new one [19]. In certain cases more than one closest existing instance is deployed such that the majority class of the nearest K neighbours is allotted to the new instance. According to Seeja and Zareapoor [19], the KNN algorithm offers consistently high performance among the different credit card fraud detection techniques, without apriori assumptions about the distributions from which the training examples are drawn.

Decision Tree: The decision tree algorithm is a type of supervised learning technique. It uses ID3 technique to construct decision tree by considering dataset entropy. In a set of data, the amount of uncertainty is measured using the entropy. The calculation of entropy of each attribute is used to determine the splitting criteria in design of decision tree [7]. An equation containing the probabilities of the attributes of dataset is used to calculate the entropy of different state. After the calculation of entropy of each attribute in a dataset, gain is obtained by carrying out subtraction operation between the entropy of entire dataset and the entropy of the splitting attribute.

Random Forest: The random forest algorithm is an ensemble of decision trees. The random forest decision tree algorithm is a supervised learning machine learning technique used to solve regression as well as classification problem [7]. The basic principle of random forest is that a group of "weaker learners" combine to form a "stronger learner". Many decision trees are grown by random forests. In this case, an individual decision tree is a "weaker learner" while all the decision trees grouped together are a "stronger learner" [19]. In classifying a new object, it is run down in each of the forest trees. A classification output is given by each tree for a class. The forest classifies the new object into the class having maximum outputs [19]. Random forests fast, and can effectively take care of unbalanced and large databases with many features. It has been establish to offer a good estimate of the generalization error and to be resistant to over-fitting [23].

In order to carry out credit card fraud detection, the random forest technique uses the following pseudo code [7]:

a) "Extract the test features of incoming transaction and use the rules of each randomly created decision tree to predict the result and stores the predicted result (target)."
b) "Calculate the votes for each predicted target output."
c) "Evaluates the high voted predicted target from different decision tree as the final prediction output."

FraudMiner: The fraudminer is a fraud detection model (a typical matching algorithm). This model uses frequent dataset mining to create legitimate transaction pattern and fraudulent transaction pattern of individual customer from their genuine transactions and malicious transactions respectively during the training phase. The matching algorithm detects which pattern the incoming transaction matches more during the test phase [19]. The matching algorithm uses to binary state of "0" and "1" to establish legal transaction and fraudulent transaction. Hence, whenever the incoming transaction is matching with the genuine pattern of the given customer then the algorithm returns "0" if the incoming transaction is matching more with fraud pattern of that customer then the algorithm returns "1" [19].

Fuzzy Logic Based System: This technique classifies the credit card transaction into fraudulent (suspicious) and genuine (non-suspicious). It combines genetic algorithm and fuzzy logic to reduce a false alarm [20].

Outlier Detection: This is an unsupervised data learning technique. It is used to identify abnormal behaviour of credit card user and by so doing detect fraudulent credit card transactions. Outlier detection is a method that seeks objects that are inconsistent or dissimilar with respect to the remaining data or are at distance from the centroids of their cluster.

Banksealer: This is an online banking semi-supervised decision support and fraud analysis system. In this method, users of the online banking web application are characterized by means of a local, a global and a temporal profile that are constructed during a training phase, taking as input a list of transactions. Each of a local profile, global profile, and a temporal profile mines different statistical features from the transaction attributes like average, minimum, maximum, actual value based on the type of model constructed. As soon as the profiles are constructed, "Baksealer processes new transactions and ranks them according to their anomaly score and the predicted risk of fraud. The anomaly score quantifies the statistical likelihood of a transaction being a fraud with respect to the learned profiles. The risk of fraud prioritizes the transactions combining the anomaly score with the transaction amount. Banksealer provides the analysts with a ranked list of potentially fraudulent transactions, along with their anomaly score" [24].

Multilayer Perception: A Multilayer Perceptron (MLP) is a classification model. It is the simplest form of a deep, Artificial Neural Network (ANN), which has three or more layers of nonlinearly activating nodes [22]. They are frequently applied to supervised learning problems for training on a set of input-output pairs and learn to model the correlation between them [22]. An MLP has an input layer for receiving signal, an output layer for making decision or prediction about the input and between the input layer and output layer, there is an arbitrary number of hidden layers. These hidden layers are the real computational engine of the MLP. MLPs that have one hidden layer can approximate any continuous function.

IV. OBSERVED CHALLENGES OF FRAUD DETECTION TECHNIQUE

In this section, some of the challenges of fraud detection techniques observed from the study are presented in this section.

Evolving fraud approach: As the use of credit card continuous to attract more users' transactions online, merchants and organizations are developing sophisticated means of securing online businesses to reduce financial fraud risk and to boost users' confidence. In the same vein, online financial crime syndicates/fraudsters are becoming even more innovative in their tactics so to be able to get illegal entry into the systems and carry out the deed. Hence, it necessarily important that machine learning algorithms is updated with the changed fraud approach to detect fraudulent actions. That is, the machine learning fraud detection algorithms should be continuously improving. Otherwise, the performance and effectiveness of the algorithm will decrease and fail to meet design objectives.

Imbalance in fraud types and detection technique: There are diverse financial frauds and this has resulted to various fraud types and detection techniques. The imbalance in both fraud types and detection techniques can be attributed to the fact that some studies are extensively carried out while others such as hybrid techniques, are examined superficially [25] [3].

Class Imbalance Problem: Usually in credit card fraud detection study, developing algorithm(s) requires the classification of transactions either as legitimate (non-fraudulent) or illegitimate (fraudulent) and this makes it difficult to develop them due to problems bordering on feature selection, parameter tuning, and analysis. That is imbalance occurs in fraud detection algorithms classification. The outcome of this problem is reduced user knowledge in legitimate clients, as identifying the fraudsters often leads to decline in some genuine transactions.

Concern for individual privacy: Since the issue of financial fraud is a sensitive one, individuals (who are stakeholders) are mostly unwilling to share information on it. This has resulted to experimental challenges like under sampling [21].

Complexity of algorithm: Usually the machine learning algorithms require high mathematical computational skills. It also requires prolong time to generate feature set that reduces the process of fraud identification/detection.

Interpretation problem: As a result of the complexity and technicality of fraud detection models, explaining and interpreting them seems challenging. For instance, models provide an output that show the likelihood of a transaction being illegitimate or not, but never provided explanation for it.

Computational performance: As stated in [26], and reported by Patil and Lilhore [3], research on computational performance of fraud detection techniques for use in real time purposes has been very little.

VI. CONCLUSION

In this paper, different works and methods of fraud detection in credit cards have been examined. A discussion on credit card activities was presented. Categories of credit card fraud were considered. The significant of fraud detection techniques was discussed. Various means of carry out fraud in finance/banking were looked at. Finally, having considered some fraud detection techniques, challenges of fraud detection models/algorithms were presented. Since each of the fraud detection methods has it individual benefit and weakness, a hybrid model exploring the diversity of the various models and combining them for optimal advantage will be worthwhile. Also, the fact that credit card fraudsters are exploring new innovative ways of get around committing fraud, means that a credit model should be one that is continuously updating to meet design objectives otherwise it fails.

REFERENCES

[1] P. K. Chan, W. Fan, A. L. Prodromidis, and S. J. Stolfo, “Distributed Data Mining in Credit Card Fraud Detection,” IEEE Intelligent Systems, pp. 67-74, 1999.

[2] P. Jayant, Vaishali, and D. Sharma, “Survey on Credit Card Fraud detection Techniques,” International Journal of Engineering Research & Technology, Vol. 3, No. 3, pp. 1545-1551, 2014.

[3] V. Patil, and U. K. Lilhore, “A Survey on Different Mining and Machine Learning Methods for Credit Card Fraud Detection,” International Journal of Scientific Research in Computer Science, Engineering and Information Technology, Vol. 3, No. 5, pp. 320-325, 2018.

[4] T. P. Bhatla, V. Prabhu, and A. Dua, “Understanding Credit Card Frauds,” Cards Business Review, pp. 1-15, 2003.

[5] S.K.Saravanan1, and G.N.K. Suresh Babu, “An Improving Credit Card Fraud Detection Using Novel Data Mining Techniques,” International Journal of Current Engineering and Scientific Research, Vol. 4, No. 10, pp. 17-23, 2017.

[6] M. Zareapoor, and P. Shamolmoali, “Application of Credit Card Fraud Detection: Based on Bagging Ensemble Classifier,” Procedia Computer Science 48, pp. 679-685, 2015.

[7] S. Patil, V. Nemade, and P. Soni, “Predictive Modelling for Credit Card Fraud Detection Using Data Analytics,” Procedia Computer Science 132, pp. 385-395, 2018

[8] S. Agarwal, and S. Upadhay, “A Fraud Detection Approach Using Clustering Based Method,” Journal of Basic and Applied Engineering Research, Vol. 1, No. 10, pp. 33-37, 2014.

[9] B. Pushpalatha, and C. W. Joseph, “Credit Card Fraud Detection Based on the Transaction by Using Data Mining Techniques,” Vol. 5, No. 2, pp. 1785-1793, 2017

[10] Sodasoltaniziba, and M. Alibalafr, “The Study of Fraud Detection in Financial and Credit Institutions with Real Data,” Global Journal of Computer Science and Technology (C) Software and Data Engineering, Vol. 15, No. 6, pp. 37-45, 2015.

[11] M. R. Lepoivre, C. O. Avanzini, Guillaune Bignon, L. Legendre, and A. K. Piwele, “Credit Card Fraud Detection with Unsupervised Algorithms,” Journal of Advances in Information, Vol. 7, No. 1, pp. 34-38, 2016.

[12] A. Pouramirarsalani, M. Khalilian, and A. Nikravanshalmani, “Fraud Detection in E-banking by using Hybrid Feature election and Evolutionary algorithms,” International Journal of Computer and Network Security, Vol. 17, No. 8, pp. 271-279, 2017.

[13] M. Carminati, R. Caron, F. Maggi, I. Epifani, and S. Zanero, “BANKSEALER: An Online Banking Fraud Analysis and Decision Support System,” N. Cuppens-Boulahia et al. (Eds.): SEC 2014, IFIP AICT 428, pp. 380–394, 2014.

[14] K. Randhawa, C. K. Loo, M. Seera, C. P. Lim, and A. K. Nandi, “Credit Card Fraud Detection using Adaboost and Majority Voting,” IEEEAccess, Vol. XX, pp. 1-8, 2017.

[15] A. Rajaei, “Identification of Fraud in Banking Data and Financial Institutions Using Classification Algorithms,” International Journal of Information, Security and Sytems and Management, Vol. 6, No.2, pp. 653-667, 2017

[16] R. Rajamani, and M. Rathika, “Credit Card Fraud Detection Using Hidden Markov Model and Neural Network,” Proceedings of the UGC Sponsored National Conference on Advanced Networking and Applications, Special Issue Published in Int. Journal. Of Advanced Networking and Applications, pp. 175-179, 2015

[17] R. Sarno, R. D. Dewandono, T. Ahmad, M. F. Naufal, and F. Sinaga, “Hybrid Association Rule Learning and Process Minning for Fraud detection,” IAENG International Journal of Computer Science, Vol. 42, No. 2, 2015

[18] S. S. Baboo, and N. Preetha, “Analysis of Spending Pattern on Credit Card Fraud Detection,” IOSR Journal of Computer Engineering (IOSR-JCE) Vol. 17, Issue 2, pp. 61-64, 2015

[19] K. R. Seeja, and M. Zareapoor, “FraudMiner: A Novel Credit Card Fraud Detection Model Based on Frequent Itemset Mining,” Hindawi Publishing Corporation, The Scientific World Journal, pp. 1-10, 2014.

[20] N. Malini, and M. Pushpa, “Analysis on Credit Card Fraud Detection Techniques by Data Mining and Big Data Approach,” International Journal of Research in Computer Applications and Robotics, Vol. 5, No. 5, pp. 38-45, 2017

[21] Y. Dai, J. Yan, X. Tang, H. Zhao, and M. Guo, “Online Credit Card Fraud Detection: A Hybrid Frame Work with Big Data Technologies,” 2016 IEEE TrustCom-BigDataSE-ISPA, pp. 1644-1651.

[22] R. Banerjee, G. Bourla, S. Chen, M. Kashyap, S. Purohit, and J. Battipaglia, “Comparative Analysis of Machine Learning Algorithm through Credit Card Fraud Detection,” New Jersey Governor’s School of Engineering and Technology, pp. 1-10, 2018

[23] N. Khare, and S. Y. Sait, “Credit Card Fraud Detection Using Machine Learning Models and Collating Machine Learning Model,” International Journal of Pure and Applied Mathematics, Vol. 118, No. 28, pp. 825-838, 2018.

[24] M. Carminati, M. Polino, A. Continella, A. Lanzi, F. Maggi, and S. Zanero, “Security Evaluation of a Banking Fraud Analysis System,” ACM Transactions on Privacy and Security, Vol. 1, No. 1, pp. 1-30, 2018

[25] S. Sharma, P. Mitta, and Geetika,”An Approach to Detect Credit Card Frauds Using Attribute Selection and Ensemble Techniques,” international Journal Journal of Computer Applications, Vol. 180, No. 21, pp. 1-6, 2018.

[26] S. Kumari and A. Choubey, “A Review on Various Techniques and Approaches for Credit Card Treachery Detection,” International Journal of Scientific Research Engineering and Technology, Vol. 6 No. 5, pp. 485-489, 2017