Identification of Fraudulent Online Transactions and Protection: State-of-art Techniques
ABSTRACT

In this digital era, the trend of online transactions for E-commerce sites and banking services is increasing. By using different online transaction methods users can make payments directly from their bank accounts. But along with the increase of online transactions, there is an increase in fraudulent transactions. These fraudulent transitions have identical features and characteristics of online transactions. As a result, online transaction security has become a serious concern in recent years. There are a plethora of problems in the system, and banks are taking the issue seriously, therefore transactions are becoming more complex as a result. Even if fraud is discovered, it must be discovered quickly. That's why the system transaction requires high-speed authentication and security. Online transaction fraud may be detected using any approach. Hence, there is a need for the development of frameworks or technologies to detect fraudulent transactions. In this context, this review article represents a survey of the latest (2015–2021) frameworks and techniques proposed by the researchers for the identification of fraudulent transactions and securing online transactions. In this review article, we included the fraud detection techniques that are based on AI, machine learning, deep learning, and blockchain. In addition, we compare and contrast several fraud detection methods. Thus, this review paper serves as a valuable resource for researchers and academics interested in learning more about fraud detection methods.

KEYWORDS
Fraudulent transaction, Machine learning, AI, Deep learning, Big data, Blockchain

JEL CODES
F38; O31
1 INTRODUCTION

Nowadays every E-commerce company and banking sector encourages users to use online transactions, but the intensive use of online methods of transaction leads to an increase in credit card frauds (Brause et al., 1999; Laleh & Azgomi, 2009; Potamitis, 2013), telecommunication frauds (Cortesão et al., 2005; Gosset & Hyland, 1999), auction frauds (Chang & Chang, 2012; Chau et al., 2006), automobile insurance frauds (Šubelj et al., 2011; Wen et al., 2005), healthcare frauds (Chen & Gangopadhyay, 2013; Musal, 2010; Sparrow, 2019). Researchers proposed many techniques for the detection of frauds that are based on machine learning (Hines & Youssef, 2018b, 2018a), SVM (Hines & Youssef, 2018a; Tran et al., 2018), ANN (Hines & Youssef, 2018b; Maes et al., 2002), k-NN (Hines & Youssef, 2018b, 2018a), logistic regression (Bahnsen et al., 2016; Cody et al., 2018), AI (Ogwueleka, 2011), deep learning (Roy et al., 2018), and random forest (Hines & Youssef, 2018b). Table 1 represents the list of abbreviations.

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td>Support Vector Machine</td>
</tr>
<tr>
<td>ANN</td>
<td>Artificial neural network</td>
</tr>
<tr>
<td>k-NN</td>
<td>K Nearest Neighbour</td>
</tr>
<tr>
<td>PNN</td>
<td>Probabilistic neural network</td>
</tr>
<tr>
<td>LOR</td>
<td>Logistic regression</td>
</tr>
<tr>
<td>GP</td>
<td>Genetic Programming</td>
</tr>
<tr>
<td>DTs</td>
<td>Decision trees</td>
</tr>
<tr>
<td>SOM</td>
<td>Self-organizing map</td>
</tr>
<tr>
<td>LSTM</td>
<td>Long-short term memory</td>
</tr>
</tbody>
</table>
1.1 Related Surveys

Edge and Falcone Sampaio (2009) surveyed the signature-based fraudulent transaction detection techniques. At first, the authors give details about different types of financial frauds, and then the authors explain the signature-based detection technique in detail. But this survey paper is limited to signature-based detection methods. Raj and Portia (2011) surveyed different credit card fraud detection methods. The authors focus on the Hidden Markov model (Bhusari & Patil, 2016; Robinson & Aria, 2018), ANN, SSAHA hybridization, and Bayesian techniques (Ezawa & Norton, 1996; Lam & Bacchus, 1994; Maes et al., 2002; Mehdi et al., 2007; Panigrahi et al., 2009). Suresh and Raj (2018) reviewed different data mining techniques used to detect credit card frauds. Jain et al. (2019) did a comparative analysis of different credit card fraud detection techniques. The authors first explain the different types of credit card frauds and then compare different detection techniques. Yee and Sagadevan (n.d.) proposed machine learning and data mining techniques for fraud detection. Padhi et al. (2020) did a comparative survey of different machine learning techniques for online fraudulent transaction detection. A comparison of different fraudulent detection techniques is represented in Table 2.

1.2 Organization

The rest of the paper is organized as follows: Section 2 represents the fraudulent transaction detection in intelligent devices, Section 3 reviews the credit card fraud detection techniques. Section 4 reviews different blockchain-based fraudulent transaction detection methods. Section 5 explains different parameters used to represent the efficiency of detection models. and finally, Section 6 concludes the paper.

2 FRAUD TRANSACTION DETECTION FOR SMART DEVICES

With the development of smear device technologies, users preferred to use online transactions to pay for their utility bills. Due to this, there is an increase in the applications like ApplePay (Liu & Mattila, 2019) that provide the facility of mobile payment to the industry. But these mobile payment applications are affected by fraudulent transactions, phishing attacks, MITM attacks, and DoS attacks. So there is a need for the development of standards that monitors mobile transactions and provide security and privacy to the users.

| Table 2. Comparison of different fraudulent transaction detection surveys |
|---------------------------|-------------------|-------------------|-------------------|-------------------|-------------------|-------------------|
| Author                    | Credit card fraud | Online transaction| Smart devices transaction | Machine learning | Data mining | Blockchain |
| Edge and Falcone Sampaio, 2009 | ✔                 |                    |                   |                  | ✔               |                    |
| Raj and Portia, 2011       | ✔                 |                    |                   | ✔                | ✔              |                    |
| Yee and Sagadevan, n.d.    | ✔                 |                    |                   | ✔                | ✔              |                    |
| Padhi et al., 2020         | ✔                 | ✔                   |                   |                  | ✔              |                    |
| Our approach               | ✔                 | ✔                   | ✔                  | ✔                | ✔              | ✔                  |
2.1 Fraud detection for mobile devices
Despite the fact that users may be cut off from the network, a new micropayment method presented by Daza et al. (2015) ensures the security of its users. The proposed solution is known as "FRoDO" and it employs a two-factor authentication system to ensure data protection. Garg and Garg (2015) suggested a biometric and token-based secure transaction method. Online transactions are authenticated using a subscriber's biometric data. In addition, this strategy relied on a secure database to store its customers' biometrics. Jetsiktat et al. (2015) make improvements to this model. The suggested approach stores the participant's face using MPEG7-EHD standards. For an online transaction, face data is utilised to verify user and merchant identity.

Cryptographic techniques are also used by researchers to secure online transactions. Rui-xia (2015) proposed an identity-based cryptographic security model that provides integrity, authenticity, and security to online transactions. The proposed model uses a 32-bit security chip that adds security features to the online transaction. Yeh et al. (2018) proposed an elliptical curve-based technique and bilinear paring for securing online transactions. The security of the proposed model is based on the security of the bilinear pairing and non-traceability of ECDLP. In order to check the security of the proposed model, the author set up an Ethereum (Vujičić et al., 2018) based network. Table 3 represents the comparison between different secure mobile transaction techniques.

2.2 Secure transaction for wearable Devices
Nowadays wearable devices are also used to transfer money using different applications like Topshop bPay (Carlson, 2015) and Fit Pay Smart Strap (Gil, 2015). These types of transactions that are initiated by wearable devices required different types of protection methods because traditional transaction protection mechanisms (Abughazalah et al., 2014; Ali & Awal, 2012; Blass et al., 2009; Cha & Kim, 2013; Chen et al., 2011; Kazan & Damsgaard, 2013; Lacmanović et al., 2010; Mainetti et al., 2012; Ojetunde et al., 2015; Sung et al., 2015) are not applicable in the limited storage, power, and bandwidth scenario. So in order to provide security to the transactions generated by wearable devices, researchers proposed certificate-less (CLS) public-key

<table>
<thead>
<tr>
<th>Model</th>
<th>System-Level</th>
<th>Hardware-level</th>
<th>Simulation</th>
<th>Basic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zheng et al., 2016</td>
<td>✔</td>
<td>✔</td>
<td>ARM Cortex-A9 MP core</td>
<td>ARL trust zone and open SSL cryptography</td>
</tr>
<tr>
<td>Yeh et al., 2018</td>
<td>✔</td>
<td></td>
<td>Raspberry PI 3</td>
<td>Bilinear paring</td>
</tr>
<tr>
<td>Urien, 2016, Urien and Aghina, 2016a, 2016b</td>
<td>✔</td>
<td>✔</td>
<td>Android Phone</td>
<td>Open mobile API, TLS-SIM API, host card emulator, EMV protocol</td>
</tr>
<tr>
<td>Yicheng and Zhaoxia, 2016</td>
<td>✔</td>
<td></td>
<td></td>
<td>Range controlled communication chip</td>
</tr>
<tr>
<td>Park and Lee, 2016</td>
<td>✔</td>
<td></td>
<td>KS X 6928 standard, signature record type</td>
<td></td>
</tr>
<tr>
<td>Rui-xia, 2015</td>
<td>✔</td>
<td></td>
<td>WIS08SD548E</td>
<td>Identity-based cryptography, one-time-key</td>
</tr>
</tbody>
</table>
cryptosystems (Yeh et al., 2018). Gong and Li (2014) proposed the CLS scheme that works without bilinear pairings, but it works only secure for supertype II adversary attack (Yeh et al., 2013). Tsai et al. (2014) improve the CLS scheme that is based on elliptical curve cryptography and provides better security as compared to the scheme proposed by He et al. (2012). Later, Wang et al. (2015) proposed the improvised version of previously presented CLS schemes, but the proposed scheme is not effective against supertype I adversary attack (Yeh, 2017). A new CLS-based scheme that is secure against supertype I and supertype II attack is introduced by Tsai et al. (2015), but the proposed scheme has limited scalability. In order to overcome all the limitations of previously prosed CLS, Yeh et al. (2018) proposed a secure transaction scheme based on elliptical curve cryptography and Figure 1 represents this secure transaction scheme. In the proposed scheme (Yeh, 2018), the transaction between the user and android pay app is started at step 1-1 of Figure 1, then TSA at the android platform generates a random number with the help of users identity, transaction identity, params, and s using the following equations.

\[ R_1 = r_1 \cdot P \]  
\[ h_1 = H(ID, ID_T, R_1, PK_{TSA}) \]  
\[ s_1 = r_1 \cdot ID + h_1 \cdot PK_{TSA} \text{ mod } n \]  

After confirmation of the transaction in step 4-1 and 4-2 of Figure 1, a secure service connection between the payment platform and user application is established. Finally, the user sends the verification request in step 4-4. A similar process takes place in step 5-1 of Figure 1 when the android payment platform confirms the TSA from the merchant server. In steps 7 and 6 of Figure 1 android payment application secures the transaction from the merchant server. The security analysis of the proposed model is represented in Table 4 and Figure 2.

### 3 Fraudulent Credit Card Transaction Detection

As more and more users are using online means of transaction, there is an increase in online frauds (Allan & Zhan, 2010; Bolton & Hand, 2002; Pejic-Bach, 2010). Most online or offline transactions are conducted by credit cards (Alexopoulos et al., 2007), so the majority of the fraudulent transactions are related to this payment method. Two types of fraudulent transactions are related to credit cards, 1) Inner card fraud, 2) External card fraud (Awoyemi et al., 2017; Misra et al., 2020). If a malicious user opens a bank account with fake details and completes a fraudulent transaction then it comes under inner card fraud. External card fraud includes all fraudulent transactions related to the stolen credit cards. The traditional fraud detection techniques do not apply to fraudulent credit card detection because of the dynamic nature of credit card fraud scenarios. Hence, researchers proposed different techniques for the detection of fraudulent transactions related to credit cards (Adewumi & Akinyelu, 2017; Awoyemi et al., 2017; Bhattacharyya et al., 2011; Sethi & Gera, 2014).
3.1 Machine Learning approach for Fraudulent transaction detection

These machine learning techniques use supervised (Bhattacharyya et al., 2011; Dhankhad et al., 2018), and unsupervised learning (Carminati et al., 2015) methods to detect fraudulent transactions (Bolton & Hand, 2002; Kou et al., 2004). In supervised learning techniques, labelling of the credit card data is done, and in unsupervised learning techniques, fraudulent transaction detection is done by different patterns of credit card transactions (Taha & Malebary, 2020). Different classifying techniques like PNN, SVM (Singh et al., 2012), LOR, GP, BNN (Ravisankar et al., 2011), and DTs (Kirkos et al., 2007) are used to differentiate malicious transactions from legitimate transactions. SOM is used in unsupervised machine learning techniques because it did not
Table 4. Performance comparison of secure mobile transaction model

<table>
<thead>
<tr>
<th>Models</th>
<th>Super Type I</th>
<th>Super Type II</th>
<th>Sign phase computation cost</th>
<th>Verify phase computation cost</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tsai, 2015</td>
<td>✔️</td>
<td>✔️</td>
<td>$1T_{\text{Inv}} + 1T_{\text{em}} + 1T_{\text{m}} + 1T_{\text{h}} + 1T_{\text{add}}$</td>
<td>$2T_{h} + 2T_{\text{em}} + 2T_{\text{add}} + 2T_{h}$</td>
</tr>
<tr>
<td>Tsai et al., 2014</td>
<td>--</td>
<td>--</td>
<td>$1T_{g} + 1T_{\text{em}} + 2T_{m} + 2T_{h} + 2T_{\text{add}}$</td>
<td>$4T_{\text{em}} + 3T_{\text{add}} + 3T_{h}$</td>
</tr>
<tr>
<td>Gong and Li, 2014</td>
<td>✗</td>
<td>✔️</td>
<td>$1T_{\text{inv}} + 1T_{\text{em}} + 2T_{m} + 1T_{h} + 2T_{g}$</td>
<td>$4T_{\text{em}} + 3T_{\text{add}} + 3T_{h}$</td>
</tr>
<tr>
<td>Wang et al., 2015</td>
<td>✗</td>
<td>✔️</td>
<td>$1T_{g} + 1T_{\text{em}} + 1T_{h} + 2T_{add}$</td>
<td>$3T_{\text{em}} + 3T_{\text{add}} + 2T_{h}$</td>
</tr>
<tr>
<td>Yeh, 2018</td>
<td>✔️</td>
<td>✔️</td>
<td>$1T_{g} + 1T_{\text{em}} + 1T_{h} + 2T_{add}$</td>
<td>$4T_{\text{em}} + 3T_{\text{add}} + 2T_{h}$</td>
</tr>
</tbody>
</table>

$T_{\text{Inv}} = \text{Inversion operation time}$

$T_{h} = \text{Bilinear pairing time}$

$T_{\text{em}} = \text{ECC based scalar multiplication time}$

$T_{\text{eadd}} = \text{ECC based addition time}$

$T_{m} = \text{General multiplication time}$

$T_{\text{add}} = \text{General addition time}$

$T_{h} = \text{Hash function time}$

$T_{g} = \text{Time to generate a random number}$

Figure 2. Comparison of the execution time of the secure transaction schemes

require labeling of the dataset, but this, in turn, reduces the overall accuracy of the detection algorithm (Olszewski, 2014; Quah & Sriganesh, 2008; Zaslavsky & Strizhak, 2006). Recently, deep learning techniques are also used by the researchers for identification of fraudulent transactions (Fiore et al., 2019; Juergovskiy et al., 2018; Kraus & Feuerriegel, 2017; Wang, Liu, et al., 2015). Taha and Malebary (2020) prospered a light gradient boosting machine-based fraud transaction detection mechanism. The proposed mechanism using a Bayesian hyperparameter optimization algorithm for the identification of fraudulent
transactions. The overall framework of the proposed approach is represented in Figure 3. The author used two data sets that are represented in Table 6 to analyze the effectiveness of the proposed model, and from Figure 3 it is clear that the proposed approach has four steps for the identification of fraudulent transactions.

A comparison of the accuracy of different fraudulent detection approaches is presented in Table 5. Figure 4 represents the precision and accuracy of different machine learning approaches that are used for fraudulent transaction detection. So from Figure 4, it is clear that ‘OLightGBM’ (Taha & Malebary, 2020) works more accurately than other machine learning techniques.

### Figure 3. Framework for Taha and Malebary (2020)

### Table 5. Comparative analysis of machine learning approach for fraudulent transaction detection

<table>
<thead>
<tr>
<th>Approach</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jog and Chandavale, 2018</td>
<td>80%</td>
</tr>
<tr>
<td>John and Naaz, 2019</td>
<td>97%</td>
</tr>
<tr>
<td>Dal Pozzolo et al., 2014</td>
<td>95%</td>
</tr>
<tr>
<td>Lakshmi and Kavilla, 2018</td>
<td>95.5%</td>
</tr>
<tr>
<td>Rohilla, 2017</td>
<td>92.86%</td>
</tr>
<tr>
<td>Taha and Malebary, 2020</td>
<td>98.40%</td>
</tr>
</tbody>
</table>

### Figure 4. Precision and Accuracy of different machine learning approach for fraudulent transaction detection

*Source: (Taha & Malebary, 2020).*
Error: no se encontró el origen de la referencia Figure 5 represents the comparison of AUC of different machine learning algorithms that are used to detect fraudulent transactions. A high value of AUC value means that the proposed approach efficiently detects the malicious transaction, so from Figure 5 it is clear that the ‘OLightGBM’ algorithm is best for the detection of fraudulent transactions.

3.2 Deep learning approach for fraudulent transaction detection

There are different deep learning methods that are proposed by the researchers (Fu et al., 2016; Ghosh & Reilly, 1994; Pumsirirat & Yan, 2018) for the detection of fraudulent transactions, the main advantage of using a deep learning approach is that there is no need for a labeled training dataset. Autoencoders (Baldi, 2012) is a simple example of simple deep learning techniques that can be used to detect fraudulent transactions. Figure 6 represents the autoencoder architecture.

Autoencoders are a feed-forward type of neural network, in which the number of input states is the same as that of output states as represented in Figure 6. There are different types of autoencoders like undercomplete autoencoders, that has fewer number of hidden layers as compared to input layers. The autoencoder has two layers 1) Encoder, 2) Decoder.

- Encoder: It encodes the input layer into a hidden layer according to (5)
  \[ H = f_h(W_h \times I + b_h) \]  
  \( f_h \) = Activation function 
  \( W_h \) = Weight 
  \( b_h \) = Bias 
  \( I \) = Input

- Decoder: It converts the output from the hidden layer into the final output by using (6)
  \[ X' = f_h(W_h \times H + B_h) \]  
  \( f_h \) = Activation function 
  \( W_h \) = Weight 
  \( B_h \) = Bias 
  \( H \) = hidden layer output

Recently, Misra et al. (2020) prosed an autoencoder-based technique for fraudulent transaction detection. In the prospered approach autoencoders are used to extract the relevant features from the input dataset.
3.3 Hybrid Models for fraudulent transaction detection

Some researchers proposed hybrid techniques for fraud transaction detection (Carcillo et al., 2019). Carcillo et al. (2018) proposed a hybrid technique combining big data algorithms and machine learning techniques. Yuan et al. (2017) combine deep learning techniques and graph techniques for the detection of fraudulent transactions. Recently, the use of mathematical methods like Fourier transform (Saia & Carta, 2017), computational intelligence (West & Bhattacharya, 2016), and wavelet transforms (Saia, 2017) are also used by researchers to analyze credit card transactions. Kim et al. (2019) proposed a hybrid approach for credit card fraud transaction detection. The proposed approach is called the “Champion-challenge” framework and it is based on a hybrid ensemble and deep learning model.

4 BLOCKCHAIN-BASED FRAUDULENT TRANSACTION DETECTION METHODS

Nowadays most online transactions are completed by the involvement of a third party, as represented in Figure 7. As represented in Figure 7, there is a transaction between merchant and consumer takes place through a centralized financial institute. So the complete transaction is dependent on the centralized financial institute or third party. But if a third party does not maintain a proper security measure then leaking of user information takes place. So there is a need for a technique that provides authentication to the online transaction process. Some researchers proposed a blockchain-based authentication approach for securing online transactions (Nakamoto, 2019; Ouaddah et al., 2017; Zhang et al., 2016).
& Jacobsen, 2018). With the help of bitcoin (Andrychowicz et al., 2014; Huang et al., 2018; Zhu et al., 2018), there is no need for a third-party payment gateway, which increases security. But bitcoin has its own limitations, so there is a need for research in the field of blockchain application for secure transactions.

Blockchain (Abadi & Brunnermeier, 2018; Nofer et al., 2017; Yaga et al., 2019) was introduced in 2009 for digital payments, but its decentralized nature (Cong & He, 2019) and immutability (Aitzhan & Svetinovic, 2016) attract researchers and industries to find its application in other fields other than digital payments like the healthcare sector (Dagher et al., 2018) and IoT (Almad- houn et al., 2018; Khan & Salah, 2018; Suliman et al., 2018). The use of smart contracts (Ethereum, https://www.stateofthedapps.com/; Ethereum White Paper, https://github.com/ethereum/wiki) and consensus algorithms (Cao et al., 2019; Quorum White paper, 2016/2020; Smart Contracts Running on a BFT Hardened Raft, 2016/2021; Tschorsch & Scheuermann, 2016) made the blockchain techniques beneficial to different domains.

In the past researchers proposed different methods of fair payment using blockchain technology (Ahn et al., 2018; Zhao et al., 2018). Wang et al. (2019) proposed a blockchain-based third-party authentication system. The proposed approach provides integrity and privacy to the online transaction.

5 PERFORMANCE EVALUATION METHODS
In this section, we explain the evaluation techniques used by the researchers.

- True positive (TP): It is the legitimated transactions that are identified by the detection algorithm.
- False-positive (FP): It is the number of fraudulent transactions that are not identified by the detection algorithm.
- False-negative (FN): It is the number of legitimate transactions that are dropped by the detection algorithm.
- Accuracy: It is the measure of correctness of the detection algorithm, and it is calculated by (7)

\[
\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}
\]

- Precision (P): It measures the percentage of legitimated transactions from the total pass Transactions. It is calculated by (8)

\[
\text{Precision} = \frac{TP}{TP + FP}
\]

- Recall (R): It is the percentage of the legal transactions that pass through the detection algorithm. It is calculated by using (9)

\[
\text{Recall} = \frac{TP}{TP + FN}
\]

- F1-Score: This combines the precision and recall rate. It is calculated by (10)

\[
F1 - \text{Score} = \frac{2 \times P \times R}{P + R}
\]

The above-defined parameters are used to measure the effectiveness of the proposed algorithms. Along with the parameters, data sets are also required, so the most used datasets are given in Table 6.

6 CONCLUSION
Due to the increase in the use of smart devices and android applications, there is a rapid growth in the e-commerce industry. The flexibility of online transactions is increased because of the introduction of different online payment options, hence most of the users prefer online transactions as compare to
offline transactions. But sometimes malicious hacker performs fraudulent transaction and cause economical loss to customer or merchant. So in order to identify the fraudulent transactions researchers used different techniques that are based on machine learning, big data, blockchain, and AI. In this context, we reviewed different fraudulent detection techniques. In this paper, we mainly focus on the fraudulent transaction performed by smart devices and credit cards and includes the machine learning, deep learning, and blockchain technologies, in future work we will focus on other fraudulent transaction detection techniques.

**REFERENCES**


<table>
<thead>
<tr>
<th>Table 6. Datasets used in the evaluation process</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dataset</strong></td>
</tr>
<tr>
<td>---------------</td>
</tr>
<tr>
<td>Credit Card Fraud Dataset, <a href="https://kaggle.com/mlg-ulb/creditcardfraud">https://kaggle.com/mlg-ulb/creditcardfraud</a></td>
</tr>
<tr>
<td>UCSD: University of California, San Diego Data Mining Contest 2009, n.d.</td>
</tr>
</tbody>
</table>


