IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 25, Issue 2, Ser. II (Mar. – Apr. 2023), PP 01-05
www.iosrjournals.org

A Review on Know Your Customer (KYC) SystemUsing Blockchain Technology

N Deepak, Joy Patel, Likith Jain, and Mihika Srivastava
Dept. of ISE, BIT, BengaluruKarnataka, India

Dr. Hema Jagadish
Associate Professor, Dept. of ISE, BIT, BengaluruKarnataka, India

***Abstract**—Know Your Customer (KYC) is a crucial process that banks and other financial institutions must undertake before providing any financial services. This involves collecting and recording customer information during the onboarding process, and ensuring that this information is kept up-to-date. KYC is typically integrated into account opening forms and is mandatory for customers to provide accurate information. Almost all financial institutions involve getting KYC information from their prospective customers, from banks to insurance companies. The primary objective of this process is to identify and prevent fraudulent activities like money laundering, identity theft, terrorism financing, etc. However, the cost of managing KYC per customer can be substantial due to a lack of transparency, mistrust, and data duplication. Blockchain technology provides a solution to establish trust and transparency by creating a self-sovereign and Decentralized Know Your Customer (DKYC) model. This model enhances customer privacy through consent-based access, features regulator governance, and helps banks to use trusted and accurate customer data while reducing customer acquisition costs.*

Keywords: *Blockchain, Know-Your-Customer, Decentralized*

Date of Submission: 11-03-2023 Date of Acceptance: 25-03-2023

I. INTRODUCTION

KYC is a process utilized by banks to gather information about the identity and address of their clients, both purchasers, and borrowers. The process is governed by regulators and involves performing due diligence to verify the identity of the clients, which helps to prevent the misuse of the banking services. Banks are responsible for completing the KYC procedure when opening accounts and they are also required to periodically update their

KYC details. However, the KYC process may be manual, time-consuming, and redundant across different institutions.

For every company, verifying the identity of its clients is crucial, especially for financial institutions. To achieve this, Know Your Customer (KYC) protocols are used to help companies ensure that they know customer's whom they are doing business with. This process offers a background check and a sense of security for financial institutions and the banking sector. Usually, the KYC process involves an extended and drawn-out practice where clients are required to provide certain documents, undergo some background checks, and undergo verification.

The Primary objective of KYC is to assist financial institutions in preventing activities such as identity theft, money laundering, terrorist financing, and profiling. It also helps to eliminate runaway creditors.

Traditionally, every bank or institution conducts its own KYC process for each customer, either through a stand-alone organization or a government structure. This means every time someone wants to open a new bank account, they have to undergo the entire KYC process again from the beginning, resulting in a waste of time. Additionally, KYC information needs to be regularly updated, such as a change in phone number or address, which requires a lot of redundancy and manual effort.

A blockchain is a kind of decentralized ledger, where the digital chain acts like a virtual ledger. Every new entry is a new block added and linked to the chain. All members or entities have a copy of the blockchain so they can validate and call out invalid or suspicious transactions. The architecture of blockchain and distributed ledger technology (DLT) make it possible to gather information from various service providers into a single, cryptographically secure, and immutable database. This eliminates the need for a third party to verify the authenticity of the information. A blockchain by itself uses the SHA-1 algorithm to make each block or transaction in the chain cryptographically secured. This includes encryption and decryption of sensitive data. A user's data can only be accessed by their private key which is held by them and them only. Using blockchain technology, it becomes feasible to establish a system where users only need to undergo the KYC procedure once to verify their identity. This approach eliminates the middleman and the need to repeat the KYC process multiple times for multiple banks, making it a sustainable and efficient alternative. Sharing KYC information on the blockchain allows financial institutions to achieve better compliance outcomes, increase efficiency and enhance the customer experience.

With thorough background checks and verifications in place, this allows banks to flag or mark suspicious customers, informing all banks or entities in the network. This gives banks a chance to either further probe into the customer or not give out credit or loans to the flagged entity. This ensures the security of not just one bank, but all the financial institutions which are a part of the blockchain network.

II. LITERATURE REVIEW

Mohannad Alkhalili Mahmoud H. Qutqut et al. have proposed a model that utilizes Machine Learning (ML) to automate the process of checking blocked transactions in anti-money laundering (AML) watch-list filtering systems. This model reduces the effort required from compliance officers and significantly decreases processing time. Additionally, it handles false positives in transactions by using a polynomial kernel to achieve higher accuracy in predictions. The authors concluded that Support Vector Machines (SVM) outperformed other algorithms, achieving higher accuracies in predicting transaction decisions. The paper demonstrates the effectiveness of machine learning algorithms in AML solutions, and the authors presented a high-level architecture for integrating an ML component with a watch-list filtering AML system. **[1]**

Ashwini Kumar et al. focused on the issue of money laundering and its impact on the global economy. The authors aimed to tackle the challenge of detecting money laundering activities, which are often difficult to spot due to the money launderers' tactics of dividing the money into multiple small transactions. The paper discussed the use of big data analytics techniques and the Naive Bayes classification method to detect money laundering transactions. The paper starts by presenting the issue of anti-money laundering (AML) and the difficulties it poses to the banking system. The study concludes by reporting the findings of the analysis, which involved using the Naïve Bayes classifier and achieving an accuracy score of 0.8125**. [2]**

David Macedo et al. presented a study on document segmentation in image processing for online customer identification. The Know Your Customer (KYC) process, which is a component of Anti-Money Laundering (AML) regulations, includes a vital step of uploading identification documents. The process of verifying identification documents for the KYC process has become more difficult due to the various types of images uploaded from different devices. This study concentrates on using the U-Net model, which is a type of Convolutional Neural Network, to detect the text regions and document edges in ID images. However, the U-Net model may be computationally intensive for practical use, especially on mobile devices that have limited computational resources. Therefore, the study proposes model optimization using Octave Convolutions to reduce computational costs and make the model more efficient and portable. **[3]**

Jose-de-Jesus Rocha-Salazar et al. presented a study on anti-money laundering (AML) efforts, focusing on the use of data visualization techniques to identify suspicious activities. The authors suggest that using link analysis, which visualizes the connections and relationships between transactions, organizations, and individuals, can help detect money laundering. The goal is to identify the complex network of relationships that are exploited in money laundering. The model was tested on the bank transaction data of a large entity and was validated by professionals. The findings from the study indicate that employing data visualization tools is an effective method of identifying potentially fraudulent money laundering activities. Additionally, the study shows that implementing these techniques using open-source software is feasible and cost-effective.**[4]**

Kishore Singh et al proposed a new method for anti-money laundering and terrorism financing detection based on typologies described in Financial Action Task Force reports and an abnormality indicator using the variance of variables. The model was tested and resulted in reduced false positives and improved accuracy compared to the previous rule-based method. However, the method requires a large amount of transaction information and lacks confirmed cases of terrorism financing. Future research aims to expand the model to include cryptocurrency variables and detect financial resources from human trafficking in European and Middle Eastern countries. The benefits of the proposed model include reduced costs, human capital, and research time. **[5]**

Eric Pettersson Ruiz et al. investigated how machine learning (ML) could be used to deanonymize cryptocurrency money launderers, as current preventive efforts are outdated. They compared four supervised-learning algorithms using the Bitcoin Elliptic dataset and conducted complementary qualitative interviews at cryptocurrency exchanges to determine the fit and applicability of the algorithms. The study found that the Random Forest algorithm was the most effective identifier of illicit transactions, with a 3% better F1 score and !!% better Precision score compared to the second-best performing algorithm, Decision Tree. However, since the performance differences between the four algorithms were low, it would be misleading to conclude that one algorithm is universally more suitable. While the study used a stratified ten-cross-validation method to minimize potential overfitting, it is still possible that the model is not completely free from overfitting. Therefore, no current method can guarantee the complete absence of overfitting. **[6]**

Prof. Dr. Nevine Makram Labib et al. conducted a study to examine the technical aspects of anti-money laundering systems and reviewed the application of machine learning algorithms and techniques to detect money laundering patterns, unusual behavior, and money laundering groups. They found that unsupervised machine learning is the next step in the evolution of money laundering detection, as it can detect new patterns of money laundering and identify all accounts and groups involved while minimizing false positives. Future research aims to propose and compare unsupervised machine learning techniques for money laundering detection with other methods, specifically for countering terrorism financing. Supervised learning approaches require historically labeled data to identify suspicious accounts and activities, limiting their effectiveness against sophisticated modern-day money launderers who use new techniques every day. **[7]**

Ricardo Azevedo Araujo proposed an incentive-based approach to tackle money laundering. Financial Institutions are key players in the fight against money laundering, and their willingness and ability to report suspicious activities are crucial for effective anti-money laundering activities to regulate anti-money laundering activities could be effective, but there is a problem of hidden information because the willingness or ability of banks to prevent money laundering is private information. This creates a problem in the hidden selection, where banks may choose to engage in illegal activities to maximize profits. While the contract approach used in assessing the efficiency of the regulation may highlight the efficiency properties of international schemes of combating money laundering, it is subject to limitations due to hidden information. **[8]**

Prakash Chandra Mondal et al. developed a new method for secure and seamless financial access for online banking customers through dynamic KYC-based transaction authorization. This approach provides the same level of control as the existing OTP authorization, but with fewer dynamic risks of theft or delay of SMS delivery. This method does not require any additional hardware and is hence cost-effective as well. The proposed approach can be used from anywhere, including private or public computers, and reduces the risk of key theft. To mitigate the risk of high-risk transactions, any two transactions made within a short time interval (e.g., 30 seconds) will be temporarily blocked and the user will be notified. Similarly, any transaction made within a shorter time interval than the user s previous activities will be flagged with a high-rated CQ during the transaction. **[9]**

Ismail Alarab et al. introduced a new method for predicting illicit behavior in the Bitcoin blockchain using graph neural networks. The Bitcoin transaction graph is complex, making the prediction of illicit transactions difficult. Previous research showed that the Graph Convolutional Network (GCN) approach performed poorly in predicting illicit transactions using the Elliptic dataset. To address this issue, the authors combined GCN with linear layers to enhance the approach's performance. They concatenated

node embeddings obtained from the GCN layers with a hidden, linear transformation-derived layer, performing much better than the original approach. The paper emphasizes the need to explore new approaches and evaluate their effectiveness on relevant datasets to enhance the accuracy of predicting illicit transactions in the Bitcoin blockchain. **[10]**

Amr Ehab Muhammed Shokry et al. have tackled the pressing issue of terrorism financing by proposing an unsupervised machine learning technique for detecting hidden patterns, groups, and transactions related to money laundering, to counter-terrorism finance. Money laundering is closely linked to terrorism financing as illegally obtained funds need to be disguised to be used in legal economies. The paper provides an overview of anti-money laundering policies and regulations and describes the three dynamic stages of the money laundering process. The authors emphasize the crucial role of financial institutions in monitoring, reporting, and detecting suspicious activities. The proposed technique has shown promising results in identifying similarities, hidden patterns, and groupings across all transactions and suspicious accounts involved. The paper concludes by highlighting the importance of advanced technologies, such as machine learning, in addressing the issue of terrorism financing. **[11]**

Rasmus Ingemann Tuffveson Jenson et al. provide a comprehensive review of the literature on anti-money laundering (AML) in banks and propose a standardized terminology with two central components: client risk profiling and suspicious behavior flagging. The former involves identifying risk factors and explaining them, while the latter employs hidden features and risk indices. The paper suggests new aspects in future research like the need for more public datasets which could be addressed through synthetic data generation. Some other avenues include semi-supervised and deep learning interpretation. Despite the potential of modern statistical and machine learning methods to enhance AML operations, the scientific literature on these techniques is still limited and disjointed. The paper concludes that a major challenge is the shortage of public datasets. **[12]**

Joana Lorenz et al. examine the problem of money laundering through cryptocurrencies and propose the use of machine learning techniques to detect such activities. As traditional supervised algorithms are not feasible due to limited labels, the authors introduce an active learning approach that can perform as well as a fully supervised method with just 5% of the labels. This approach mimics real-world scenarios where a limited number of labels can be obtained through manual annotation by experts. The authors also highlight that unsupervised anomaly detection techniques are insufficient for identifying illicit patterns in a real Bitcoin transaction dataset. This paper sheds light on how to detect money laundering with limited labels and the practicality of using machine learning for anti-money laundering. **[13]**

Yue Shi et al. examines the potential use of blockchain technology in the field of cybersecurity. The authors highlight the benefits of blockchain technology, such as enhanced security, anonymity, and data integrity, without the need for third-party intermediaries. However, the paper also acknowledges the technical challenges and limitations associated with blockchain. To address these challenges, the paper presents a comprehensive study of the technical principles, security mechanisms, typical applications, security risks, and major security issues and challenges of blockchain technology. Additionally, the paper proposes an enhanced access control strategy using attribute-based encryption. In conclusion, the authors state that blockchain technology has significant potential in the realm of cybersecurity, but further research is necessary to address its limitations and overcome the challenges. **[14]**

Joe Abou Jaoude et al. conducted a systematic literature review to analyze the growing popularity of blockchain technology in various fields and areas of application. The paper highlights the unique features of blockchain technology, such as security, anonymity, and decentralization, that offer significant benefits to different domains. The authors observe that blockchain s potential applications are vast but limited studies have been conducted in areas such as the Internet of Things, energy, finance, healthcare, and government. Blockchain technology provides a decentralized environment that eliminates the need for trust between stakeholders, prevents fraud, and enables trustless peer-to-peer transactions. **[15]**

III. MOTIVATION

1. To design a system for the KYC process to assist banks as well as the public in registering their details securely and efficiently.

2. The traditional KYC process requires users to get their KYC done multiple times for different applications as the process differs for each.

3. KYC is done multiple times for different applications as the process differs for each. Banks primarily use KYC to prevent money laundering, and identity theft and prevent terrorist financing and runaway auditors. For this purpose, decentralizing the process to increase security and reusability simplifies the process for both the people who are registering as well as the regulators who are asking for the details.

IV. OBJECTIVES

1. The purpose of the system is to add better security, and transparency and resolve trust issues.

2. The aim is to overcome the limitations of the traditional KYC system and decentralize the process, increasing efficiency and usability with the help of blockchain and decentralized ledger technology.
3. The problem one faces is that the KYC process needs to be highly secure since it involves personal details, and bank details and a breach in this could lead to devastating losses for users as well as banks. The implementation aims to make access to user details easier and reusable.

4. The proposed system utilizes Blockchain technology to distribute a database of records that have been executed or shared among participating parties.

5. The system ensures the transaction is verified by the majority of participants of the system. It contains every single record of each transaction. If an anomaly is noticed, that transaction or block is flagged

V. CONCLUSION

The system being proposed is a KYC system that utilizes blockchain technology to make the traditional KYC process more efficient and cost-effective. The system is dynamic, meaning it can adapt to changes in regulatory requirements and user data. By using smart contracts, users can securely register their data and control who has access to it, allowing for a decentralized and distributed data storage architecture. This reduces costs and allows for a proportional sharing of those costs. Overall, the systems offer a secure and efficient solution to the challenges of the traditional KYC process.

The process of KYC offers not only a smooth and efficient process of KYC for customers and banks but also can be used for several other purposes like Watch list filtering, anti-money laundering, identity theft prevention, customer experience tracking, and terrorism finance detection. These use not only help the financial institutions safeguard themselves but also make the process convenient for the customers or borrowers.

REFERENCES

[1]. “Investigation of Applying Machine Learning for Watchlist-Filtering in Anti-Money Laundering” by Mohannad Alkhalili Mahmoud H. Qutqut and Fadi Almasalha 2021 IEEE

[2] “Anti-Money Laundering Detection using Naïve Bayes

[2]. Classifier” by Ashwini Kumar, Sanjoy Das, and Vishu Tyagi 2020 IEEE

[3]. “A Fast Fully Octave Convolutional Neural Network for Document Image Segmentation” by Ricardo Batista das Neves Junior, Luiz Felipe Vercosa, David Macedo , Byron Leite Dantas Bezerra 2020 IEEE

[4]. “Money laundering and terrorism financing detection using neural networks and an abnormality indicator” by Jose-de-Jesus Rocha-Salazar, Maria-Jesus Segovia-Vergas, and Maria-del-Mar Camacho-Minano 2020 Elsevier

[5]. “Anti-Money Laundering: Using data visualization toidentify suspicious activity” by Kishore Singh and Peter Best2019 Elsevier

[6]. “Combating money laundering with machine learning – applicability of supervised-learning algorithms at cryptocurrency exchanges” by Eric Pettersson Ruiz and Jannis Angelis, 2019, JMC

[7]. “Survey of Machine Learning Approaches of Anti-money Laundering Techniques to Counter Terrorism Finance” by Prof. Dr. Nevine Makram Labib, Prof. Dr. Muhammed Abu Rizka, and Amr Ehab Muhammad Shokry, 2020 IEEE

[8]. “Assessing the efficiency of the anti-money laundering regulation: an incentive-based approach” by Ricardo Azevedo Araujo, 2008, JMC

[9]. “Transaction Authorization from Know Your Customer (KYC) Information in Online Banking” By Prakash Chandra Mondal, Rupam Deb, and Mohammad Nurul Huda, 2016, IEEE conference

[10]. “Competence of Graph Convolutional Networks for Anti-Money Laundering in Bitcoin Blockchain” by Ismail Alarab, Simant Prakoonwit, Mohammed Ikbal Nacer

[11]. “Counter Terrorism Finance by Detecting Money Laundering Hidden Networks Using Unsupervised Machine Learning Algorithm” by Amr Ehab Muhammed Shokry, Mohammed Abo Rizka, and Nevine Makram Labib, 2020 ICT

[12]. “Fighting Money Laundering with Statistics and Machine Learning” by Rasmus Ingemann Tuffveson Jensen and Alexandros Iosifidis, 2022 IEEE

[13]. “Machine Learning Methods to Detect Money Laundering in the Bitcoin Blockchain in the Presence of Label Scarcity” by Joana Lorenz, Maria Ines Silva, David Aparicio, Joao Tiago Ascensao, and Pedro Bizarro, 2020, ICAIF

[14]. “From Bitcoin to Cybersecurity: A Comparative Study of Blockchain Application and Security Issues” by Fangfang Dai, Yue Shi, Nan Meng, Liang Wei, and Zhiguo Ye 2017, IEEE

[15]. “Blockchain Applications – Usage in Different Domains “ by Joe Abou Jaoude And Raafat George Saade, 2019, IEEE