International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Fraud Detection System for Identity Crime using Blockchain Technology and Data Mining Algorithms

Mr. Amol Jagdish Shakadwipi[1*], Dr. Dinesh Chandra Jain[2], Dr. S. Nagini[3]

Submitted: 16/10/2023 **Revised**: 08/12/2023 **Accepted**: 17/12/2023

Abstract: Identity crime continues to pose a significant threat in today's digital landscape, necessitating the development of highly effective fraud detection systems. This paper presents a novel and innovative approach that combines the power of blockchain technology with advanced data mining techniques to create a robust fraud detection system specifically designed to combat identity crime. By seamlessly integrating blockchain and data mining, the proposed system demonstrates exceptional capabilities in detecting and preventing fraudulent activities in real-time.The integration of blockchain technology ensures the utmost security and immutability of data by leveraging its decentralized nature. This formidable security feature makes it exceedingly challenging for malicious individuals to manipulate or tamper with personal information. Leveraging blockchain's inherent strengths, the system efficiently verifies user identities and continuously tracks any alterations made to the data, thereby significantly enhancing the accuracy and reliability of identity verification processes.Data mining techniques play a pivotal role in detecting and combating fraud by enabling the analysis of vast volumes of data. Through the implementation of sophisticated data mining algorithms, the system effectively identifies patterns and anomalies associated with fraudulent behavior. This proactive approach empowers the system to swiftly detect suspicious activities and accurately predict potential fraud attempts. By doing so, the system effectively prevents identity crimes at their early stages, effectively reducing financial losses and providing vital protection for individuals' identities.

The proposed fraud detection system operates seamlessly in real-time, constantly monitoring user transactions and activities. Any indication of suspicious behavior immediately triggers alerts, facilitating prompt actions to mitigate the impact of fraudulent activities. Furthermore, the system harnesses the power of data mining techniques to analyze comprehensive historical data, thereby enabling the identification of intricate trends and patterns that serve as strong indicators of fraudulent activity. This refined analytical capability significantly enhances the system's overall accuracy and effectiveness.

Keywords: *Immutability, proactive approach, early detection, financial losses,suspicious behavior.*

1. Introduction

Detecting and preventing identity crimes requires robust fraud detection systems. This study introduces an innovative approach that combines block chain technology and data mining techniques to develop a highly effective fraud detection system for identity crime. By integrating blockchain and data mining, the system can identify and mitigate fraudulent activities in real-time.

The utilization of blockchain technology ensures the security and immutability of data by storing it in a decentralized manner. This enhances the system's ability to verify the authenticity of user identities and track any modifications made to the data, significantly improving the accuracy and reliability of identity verification processes.

Data mining techniques play a critical role in fraud detection by analysing vast amounts of data to identify patterns and anomalies associated with fraudulent behaviour. Through the application of data mining algorithms, the system can detect suspicious activities and predict potential fraud attempts. This proactive approach enables the system to detect and prevent identity crimes at an early stage, thereby minimizing financial losses and safeguarding individuals' identities.

Operating in real-time, the proposed fraud detection system continuously monitors user transactions and activities. Any suspicious behaviour triggers immediate alerts, allowing for prompt actions to mitigate the impact of fraud. Furthermore, the system leverages data mining techniques to analyse historical data and identify trends and patterns indicative of fraudulent activity, enhancing its overall accuracy and effectiveness.

The integration of blockchain technology and data mining techniques provides a powerful solution for detecting and preventing identity crimes. By leveraging the security and immutability of blockchain and the analytical capabilities of data mining, the proposed system can effectively

[1*]*Research Scholar, Department of Computer Science and Engineering, Oriental University, Indore, Works at : SNJB's KBJ College of Engineering, Chandwad, amolshakadwipi@gmail.com*

[2]*Professor, Department of Computer Science and Engineering, Oriental University, Indore, dineshjain25210 @gmail.com*

[3]*Professor, Department of Computer Science and Engineering, VNR Vignana Jyothi Institute of Engineering & Technology, Hydrabad,Department of Computer Science and Engineering, Oriental University, Indore, nagini_s@vnrvjiet.in*

identify and mitigate fraudulent activities in real-time. This innovation has significant implications across various industries, including finance, e-commerce, and healthcare, where identity crime poses a substantial risk.

The utilization of blockchain technology offers numerous advantages to various industries, such as enhanced transparency, heightened security, and other notable features that significantly contribute to the value of their operations. Consequently, it is poised to revolutionize the current practices of identity management in an exceedingly secure manner.

The current identity management system suffers from inherent insecurities and unreliability. At every juncture, individuals are required to present multiple government-authorized identification documents, such as Voter ID, Passport, Pan Card, and others.

This practice of sharing multiple IDs raises concerns regarding privacy and exposes individuals to the risks of data breaches. Therefore, blockchain technology can pave the way for self-sovereign identity management through decentralized networks, ensuring:

Privacy: The blockchain ensures that identity documents remain secure and shielded from unauthorized access.

Trust: The authenticity and validity of identity documents can be verified through the blockchain, establishing trust among participants.

Secure storage of identity documents: The blockchain provides a secure and immutable repository for identity documents.

Verification of identity documents: Participants with authorized permission can endorse the legitimacy of identity documents, instilling confidence in their authenticity.

Identity documents are an integral part of everyday life, shared with various third parties without explicit consent and stored in undisclosed locations. Whether it's applying for a loan, opening a bank account, purchasing a SIM card, or booking a ticket, identity documents are invariably required.

Government institutions, banks, and credit agencies are widely regarded as weak links within the current identity management system, as they are susceptible to data theft and hacking attempts.

Therefore, the blockchain presents an opportunity to eliminate intermediaries while empowering individuals to independently manage their identities. However, before transitioning to blockchain-based solutions, it is imperative to comprehend the workings of identity management and address the challenges present in the existing processes.

2. Mathematical Model

To develop a mathematical model for a fraud detection system targeting identity crime using blockchain and data mining, it is crucial to define the variables, constraints, and objective function. However, it's important to note that the specific mathematical formulation will depend on the unique details and requirements of the system. Here is a general mathematical format that can serve as a starting point:

1. Sets and Indices:

- Let i represent the index for individual transactions or events.

- Let j represent the index for identity-related features or attributes.

2. Decision Variables:

- Let x[i] be a binary decision variable denoting whether transaction/event i is flagged as fraudulent (x[i] = 1) or not (x[i] = 0).

3. Parameters:

- Let Data [i, j] represent the value of feature j for transaction/event i.

- Let W[j] signify the weight or importance assigned to feature j in the fraud detection system.

- Let Threshold represent the threshold value utilized to classify a transaction/event as fraudulent or non-fraudulent.

4. Objective Function:

- The goal is to minimize the total cost or maximize the accuracy of fraud detection: minimize $\sum$ (W[j] * Data[i, j] * x[i])

5. Constraints:

- Ensure that the total cost of fraudulent transactions is above the threshold: $\sum$ (W[j] * Data[i, j] * x[i]) $\geq$ Threshold

- Ensure that each transaction/event is accurately flagged as fraudulent or non-fraudulent: x[i] $\in$ {0, 1} for all i

- Additional constraints may be incorporated based on specific requirements, such as constraints on data quality, processing time, or resource limitations.

It is important to emphasize that creating a comprehensive and accurate mathematical model necessitates a detailed comprehension of the specific requirements, data, algorithms, and constraints associated with the fraud detection system. Consequently, further research and consultation with domain experts are highly recommended

to develop an appropriate mathematical model tailored to a specific implementation.

3. Working Parameters

The system architecture of a fraud detection system for identity crime using blockchain and data mining comprises various interconnected components that work together to detect and prevent fraudulent activities. Here is a high-level overview of the system architecture:

Data Collection:

Gather identity-related data from diverse sources, including transaction records, user profiles, and behaviour logs.

The collected data encompasses attributes such as transaction details, user demographics, historical patterns, and other pertinent information.

Data Pre-processing:

Cleanse, transform, and standardize the raw data to ensure consistency and quality.

Apply preprocessing techniques like data normalization, feature selection, and outlier detection to prepare the data for subsequent analysis.

Blockchain Integration:

Establish a blockchain network to securely store and manage identity-related data.

Record transaction data, encompassing user identities and pertinent attributes, in blocks, and establish cryptographic links to ensure transparency and immutability.

Employ smart contracts to automate verification processes, enforce rules, and enhance the integrity of identity-related transactions.

Data Mining and Analysis:

Utilize diverse data mining techniques, including machine learning algorithms, anomaly detection, clustering, and classification, to analyse the collected data. Train machine learning models using historical data to identify patterns, detect anomalies, and classify transactions as either fraudulent or non-fraudulent. Conduct feature engineering and selection to extract meaningful insights and enhance the accuracy of fraud detection.

Fraud Detection and Decision Making:

Process the analysed data to identify suspicious activities and potential instances of fraud.

Establish thresholds and rules to classify transactions as fraudulent or non-fraudulent based on model predictions and predefined criteria.

Generate alerts or notifications for flagged transactions and initiate appropriate actions, such as blocking transactions or launching further investigations.

Reporting and Visualization:

Generate reports, metrics, and visualizations to provide insights into fraud detection performance, patterns, and trends.

Create reports and dashboards to facilitate monitoring, decision-making, and effective communication with stakeholders.

Continuous Improvement:

Continuously refine the system through feedback loops and ongoing monitoring.

Monitor performance metrics, including detection accuracy, false positives, and false negatives, to improve the system's algorithms, models, and rules.

Please note that the above system architecture description provides a general overview, and the specific implementation details may differ based on the unique requirements, technologies, and data sources involved in the fraud detection system for identity crime using blockchain and data mining.

4. System Architecture

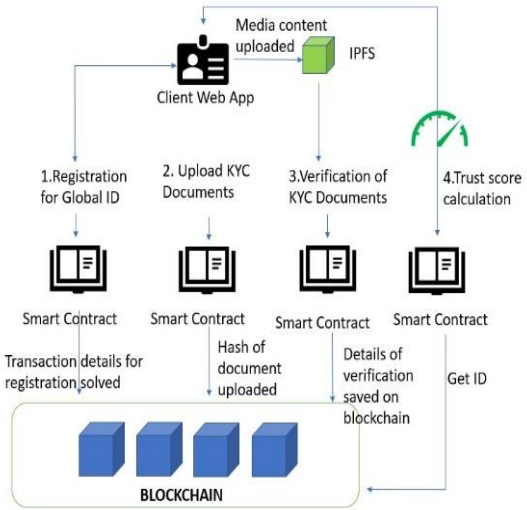

Fig: Identity management system using blockchain

A system for managing identities based on blockchain technology consists of five essential technical elements:

1. A dedicated mobile application for individuals compatible with both Android and iOS platforms.

2. A separate mobile application for verification companies designed for use on Android and iOS devices.

3. Utilization of the Interplanetary File System (IPFS) to securely store Personally Identifiable Information (PII) of users.

4. Implementation of microservices developed using Node.js programming language.

5. Integration of a permissioned blockchain component for enhanced security and transparency.

Steps involved in Blockchain Identity Management:

1. Installation of the Mobile App The initial step is to download and install the identity management app, which is used to establish an individual's identity.

Users are required to create a profile within the app. Upon completion, they receive a unique ID number that allows organizations to access their identification documents.

2. Uploading Documents, The next step involves uploading the necessary documents through the app.

Users must upload their government-issued IDs, which are stored with hashed addresses in the Interplanetary File System (IPFS). The app then extracts the Personally Identifiable Information (PII) for self-certification.

It's important to note that users retain ownership of their data and have control over which information is shared.

3. Trust Score Generation The trust score is a metric that reflects the trustworthiness of an individual. It helps organizations determine the validity or suspicion level of an account.

The user's trust score is influenced by various factors, such as the number of uploaded documents, matching information across documents, and regular usage of the system.

Smart contracts containing the business logic are responsible for calculating the trust score.

4. Access Requests from Verification Companies When a company seeks access to a user's information for authentication purposes, they must input the user's unique ID number. This action triggers a notification sent to the user's mobile phone.

The user has the authority to grant or deny the company access to their PII. Additionally, the user can track the purpose for which their PII is being used.

It's important to note that the blockchain only stores transaction records between the company and the user, not the user's PII.

5. Results and Discussion :

As the world progresses, organizations are becoming increasingly aware of the challenges associated with traditional modes of operation. Data security and efficiency are key concerns in today's evolving landscape.

Consequently, many organizations are transitioning from traditional methods to more advanced and secure approaches. This shift is evident in the field of identity management as well.

The limitations faced by organizations in traditional identity management practices have prompted the adoption of blockchain-based solutions. These solutions offer improved efficiency and security, enabling individuals and organizations to securely record and store ID information.

In conclusion, blockchain technology presents a safer, faster, and more convenient solution for identity management. It is expected that a significant number of organizations will embrace this technology in the coming years.

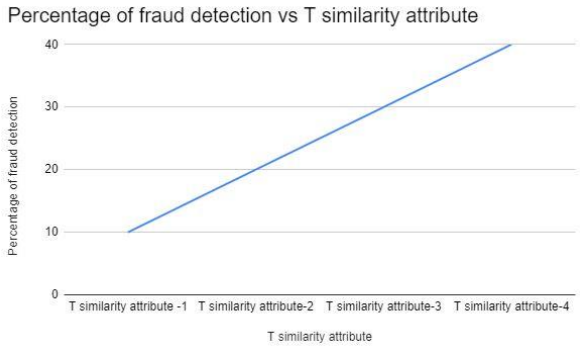

Fig: Percentage of fraud Vs. T similarity attribute .

References

[1] WENBO WANG, DINH THAI HOANG, PEIZHAO HU, "A Survey on Consensus Mechanisms and Mining Strategy Management in Blockchain Networks." , IEEE Access (Volume: 7) [2169-3536] ,2019.

[2] E.M.S.W Balagolla, W.P.C Fernando; R.M.N.S Rathnayake,M.J.M.R.P Wijesekera, A. N. Senarathne, K.Y. Abeywardhana ,"Credit Card Fraud Prevention Using Blockchain", IEEE Transaction, [20593574],2021.

[3] Aditya Asgaonkar , Bhaskar Krishnamachari ," Solving the Buyer and Seller's Dilemma: A Dual-Deposit Escrow Smart Contract for Provably Cheat-Proof Delivery and Payment for a Digital Good without a Trusted Mediator", IEEE Transaction , [978-1-7281-1328-9],2019.

[4] John O. Awoyemi, Adebayo O. Adetunmbi ,Samuel A. Oluwadare , "Credit card fraud detection using Machine Learning Techniques: A Comparative Analysis " , IEEE , [978-1-5090-4642-3],2017.

[5] K. Vidhya , P. Dinesh Kumar , ," Multi-Secure Approach for Credit Card Application Validation ", International Journal of Computer Trends and Technology,volume4Issue2-,[2231-2803] ,2013 .

[6] Alka Herenj, Susmita Mishra ," Secure Mechanism for Credit Card Transaction Fraud Detection System", International Journal of Advanced

Research in Computer and Communication Engineering Vol. 2, Issue 2, February ISSN (Print) : 2319-5940 ISSN (Online) : 2278-1021 ,2013 .

[7] Clifton Phua, Kate Smith-Miles, Vincent Cheng-Siong Lee and Ross Gayler, “Resilient Identity Crime Detection”, IEEE Transactions on Knowledge and Data Engineering, vol.2, no. 3,pp.533-546, 2012.

[8] Alka Herenj, Susmita Mishra "Secure Mechanism for Credit Card Transaction Fraud Detection System", International Journal of Advanced Research in Computer and Communication Engineering Vol. 2, Issue 2, February 2013.

[9] Namrata Shukla, Shweta Pandey, "Document Fraud Detection with the help of Data Mining and Secure Substitution Method with Frequency Analysis", International Journal of Advanced Computer Research (ISSN (print): 2249-7277 ISSN (online): 2277-7970) Volume 2 Number 2 June 2012.

[10] K. Vidhya, P. Dinesh Kumar,"Multi-Secure Approach for Credit Card Application Validation",International Journal of Computer Trends and Technology- volume 4 Issue 2- 2013.

[11] M.Swathi, K.Kalpana, "Spirit of Identity Fraud And Counterfeit Detection", International Journal of Computer Trends and Technology (IJCTT) – volume 4 Issue 6–June 2013.

[12] Clifton Phua, Kate Smith-Miles, Vincent Lee and Ross Gayler- Adaptive Spike Detection for Resilient Data Stream Mining, 2010.

[13] T.P.Latchoumi, V.M.Vijay Kannan, "Synthetic Identity of Crime Detection", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 7, July 2013.

[14] M.Swathi, K.Kalpana, "Spirit of Identity Fraud And Counterfeit Detection", International Journal of Computer Trends and Technology (IJCTT) – volume 4 Issue 6–June 2013.

[15] Clifton Phua, Kate Smith-Miles, Vincent Lee and Ross Gayler- Adaptive Spike Detection for Resilient Data Stream Mining, 2007.

[16] IBM Case study on block chain: https://www.ibm.com/blogs/blockchain/2017/07/blockchain-for-fraud-prevention-industry-use-cases/