



BLOCKCHAIN-BASED IDENTITY MANAGEMENT SYSTEMS FOR FINANCIAL INSTITUTIONS

Vanshika Sharma

Student, Pathways School, Noida, India

Student at Pre-College Program, Columbia University, NY, USA (Summer 2025)

ABSTRACT

Blockchain technology is changing how banks handle digital identities. Weak security and centralised identity management systems sometimes enable identity theft and data breaches. Blockchain creates an independent system that protects data truth, privacy, and safety via cryptography and agreement techniques. This paper investigates why banks should employ blockchain-based identity management systems for self- sovereign identification, compliance, and zero-trust security models.

The literature reveals that blockchain gives users complete control over their digital identities. This makes third-party data handling safer (Zwitter et al., 2020). Alibadi et al. (2023) state that also allows privacy-protecting approaches like zero-knowledge proofs to assess personal attributes without exposing private information. Adding blockchain to a trustless system makes checking out and accepting individuals, devices, and transactions simpler for banks (Daah et al., 2024).

Although blockchain-based identity management solutions have many advantages, financial institutions have yet to adopt them. Speed, connectivity, and legal security are challenges (Pal et al., 2021). According to the report, we must ensure that blockchain systems function together and create consistent models that follow financial regulations to address these challenges. Blockchain shared ledgers and self-executing smart contracts may also transform KYC procedures. This will boost efficiency and minimise compliance costs (Xu et al., 2022). Blockchain-based identity management solutions may protect digital IDs in banking. Personal data storage and sharing might alter. This would improve banking transparency.

INTRODUCTION

Overview Of Blockchain In Identity Management

Blockchain technology is making healthcare, finance, and supply chain management more intriguing. Many initially heard about blockchain as Bitcoin's technology. Since then, people have utilised it for more than bitcoin. It delivers autonomous, secure, and open data management technologies. Blockchain is a network of computers that record events and store them in a public, immutable ledger. The blockchain's capacity to verify and keep data makes it almost hard to alter. Thus, it is safer than centralised systems (Schardong & Custódio, 2022).

Blockchain technology might improve identity control. Banking, healthcare, and government services need identity management systems in the digital era. Identity management systems that employ regulated databases or bodies are more likely to leak, steal, or misuse personal data. Blockchain lets users create self-sovereign identities. This lets users manage their digital identities without relying on others to verify or keep their data. Name checking is crucial in financial services. This is particularly true for KYC and AML compliance. These measures are necessary yet time-consuming, expensive, and unnecessary. Blockchain-based identity management systems might speed up and secure these checks. They may save expenses and provide customers and banks greater personal data control. Blockchain makes ID checks simpler, safer, and more open for banks with security and independent systems (Tripathi et al., 2023).

Zero-knowledge proofs in blockchain-based identity management systems allow users to disclose portions of their names without disclosing sensitive information. Protecting your privacy and safety will reduce the risk of identity theft and data misuse. Blockchain makes checking persons, devices, and transactions easier for financial businesses, reducing middlemen.

Problem Statement

Current bank identity management solutions are insecure, slow, and need restricted platforms. The only area these systems may go wrong allows hackers and others to access your data. Over the years, several big bank security breaches have exposed SSNs, addresses, and bank account numbers. These hacks cost money and damage organisations' reputations and confidence. IBM's 2023 data leak research indicated that most bank assaults cost above \$5.72 million (Idrees & Nowostawski, 2024).

People cannot rename historic structures, another issue. Store shoppers must believe that retailers will protect their personal data. However, countless data breaches have shown that this confidence is not always well-placed. Our KYC and AML solutions also repeat many tasks. People typically have to prove themselves to many businesses, even if they have done it before. This raises bank expenses and reduces productivity. People are also furious.

Blockchain technology can replace identity management methods. No organization controls it, making it secure and open. Data distributed between nodes makes identification systems less sensitive to assaults. It eliminates one potential problem. People have greater control over their personal data, reducing their dependence on others. Due to scaling, connectivity, and regulatory difficulties, financial institutions have been sluggish to adopt blockchain-based identity management solutions, despite their obvious advantages (Schardong & Custódio, 2022).

Research Question

- *How can blockchain-based identity management systems enhance security, privacy, and efficiency in financial institutions?*

Research Aims and Objectives

This paper examines how blockchain-based identity management solutions may make banks safer, more private, and more efficient. The research objectives are as follows:

- To discover how blockchain-based identity management solutions outperform corporate ones in privacy, security, and speed.
- To discover the technological, legal, and practical issues banks face with blockchain-based identity management solutions.
- To address legal challenges and scale and interoperate blockchain-based identification systems.
- To examine how blockchain technology affects KYC and AML to save costs and enhance customer experience.
- To learn about upcoming blockchain technology advancements that might enhance financial businesses' identity management.

Scope of the Research

This research focusses on blockchain technology in bank identity management systems. The research seeks to identify issues with existing identity management systems and how blockchain may improve, open, and enrich them. The paper examines how blockchain's refusal to change, transparency, and cryptographic security make it a strong candidate for bank digital identity management. Real-world identity management solutions leverage blockchain. Case studies of banks utilising the technology will demonstrate its utility. For example, zero-knowledge proofs allow individuals to reveal part of their identities without disclosing all of them (Weigl et al., 2023).

The inquiry will also examine bank name rules. These regulations include GDPR, PCI DSS, KYC, and AML. Blockchain-based solutions may improve compliance and reduce the costs and issues of traditional identity management systems. The research will conclude by examining the barriers to blockchain adoption. Growth, compatibility, and law enforcement are issues. This research focusses on banks, but the findings will also be valuable in social networks, healthcare, and government agencies, where identity management is crucial. The study's findings on blockchain's function in identity management are important for financial businesses that wish to utilise it to improve safety and efficiency (Idrees & Nowostawski, 2024).

Structure Of The Dissertation

The following are study project components:

- **Introduction:** This section introduces the study's issue statement, questions, objectives, and scope.
- **Literature Review:** See what other universities and organizations have learnt about blockchain technology and bank names.
- **Methodology:** Three components make up the study paper. First is the approach, which outlines data collection and processing to answer research questions.
- **Analysis and Findings:** The paper reviews bank blockchain-based identity management solutions.
- **Discussion:** What blockchain-based identification systems entail for banks and how the findings relate to previous research.
- **Conclusion and Recommendations:** This section highlights key themes and offers ways financial institutions might adopt blockchain-based identity management solutions.

Summary

In summary: Blockchain technology might transform bank identity management. Controlled systems are vulnerable to flaws and data leaks. Blockchain can make these procedures secure, open, and decentralised. Blockchain has many advantages, yet banks seldom utilise it. This is due to regulatory and technical issues. This research seeks to identify the pros and cons of blockchain-based identity management solutions. This research will show how blockchain affects KYC, AML, and identity management to assist banks make client name trading safer, more private, and more efficient. This paper examines current research, case studies, and real-world usage to provide financial firms who wish to adopt blockchain for identity management useful advice.

LITERATURE REVIEW

Introduction

The integration of blockchain technology in identity management systems has appeared as a potential option to address the limitations of conventional centralised identity management systems, especially in the financial industry (Lim et al., 2024). As the world becomes digital, ID checks and controls are safer, quicker, and more private (Josyula, 2023). In 2023, Intiaz et al. found that blockchain's continuous, private, and independent qualities might transform identity management. Blockchain provides users greater control over their digital identities and makes financial institutions safer and more efficient.

This article reviews corporate and academic blockchain-based identity management system (BIMS) research. It largely discusses their uses, functions, and business applications. We will discuss blockchain technology's main concepts, including self-sovereign identity (SSI) models, privacy-protecting tools like zero-knowledge proofs, how blockchain can improve KYC and AML, and financial institutions' blockchain-related issues.

Blockchain Technology: An Overview

Blockchain is an open, decentralised record that enables data storage and sharing secure between nodes. The network checks each transaction and piece of data on the blockchain using PoW or PoS voting techniques. Adding data to the chain prevents changes (Dong et al., 2023). Immutability, transparency, and decentralisation make blockchain technology ideal for secure data handling.

Blockchains vary from previous systems because no one controls them. System record controls are frequent. Every blockchain node holds a copy of the whole log, according to Tripathi et al. (2023). Spread widely, there is no one point of failure; hence the approach is superior. Most of the network must agree to alter or enhance info. Idrees and Nowostawski (2024) argue this reduces frauds and threats.

Old-style identity management causes leaks and delays. Blockchain offers security options. Blockchain technology enables Bitcoin users manage their digital IDs. Citizens will not have to trust banks or the government with sensitive data. Recent emphasis has focused on self-sovereign identification (SSI) to improve privacy, security, and user control over personal data (Weigl et al., 2023). The Self-Sovereign Identity (SSI) Project and Blockchain provide a novel approach to digital identities. It allows individuals create identities without relying on leaders. Blockchain technology helps SSI generate, maintain, and verify digital IDs safely and decentralizedly (Scharndong & Custódio, 2022).

Self-Sovereign Identity (SSI) and Blockchain

Self-Sovereign Identity (SSI) is a cutting-edge approach to digital identity management that gives people complete control over their identity formation without relying on controlled organisations. To safely and decentrally make, control, and check digital IDs, SSI uses blockchain technology (Scharndong & Custódio, 2022).

Traditional Identity Management vs. Self-Sovereign Identity

Limited organisations like state governments, banks, and social networking platforms preserve name evidence. Because consumers provide these tools personal information, many organisations can acquire it. This increases the risk of identity theft and hinders personality development. As Gujar (2024) notes, users do not always have control over their data usage.

SSI addresses these challenges by allowing individuals to openly create and manage their digital IDs. Through the formation of decentralised identifiers (DIDs) and legitimate credentials, people can prove their identity and share specific traits with third parties without divulging unnecessary personal information. Identity theft is less likely with this approach, which also protects privacy (Sun et al., 2024).

Role of Blockchain in SSI

SSI is made possible by blockchain technology. It gives an open, permanent record for storing and verifying identity data. Here are a few important benefits:

- **Immutability:** Since the whole network has to agree on any changes, no one can add or remove data that has already been recorded on the blockchain. The integrity of identity information is kept safe in this way (Bandara et al., 2021).
- **Decentralization:** There is less chance of a single point of failure when there is no central authority. This makes security better.
- **Transparency:** The blockchain shows all the transactions that have happened and lets people check that identity documents are real (Pava-Díaz et al., 2024).

Advantages of SSI

- **Enhanced Privacy:** By giving only the information needed for specific proof, people can protect their data as much as possible.

- **Improved Security:** Decentralizing identity data storage mitigates the risk of large-scale data breaches (Schardong & Custódio, 2022).
- **User Empowerment:** Users have sovereignty over their identity, fostering trust in digital interactions.
- **Streamlined Processes:** Some financial companies find it easier to do Know Your Customer (KYC) tasks with SSI because it lets approved identity information be used more than once. This saves money and cuts down on waste (Sun et al., 2024).

Notable SSI Initiatives

Several projects are pioneering SSI solutions:

- **Sovrin:** An open-source project that provides a global public utility for self- sovereign identity, enabling individuals to manage their digital identities (Weigl et al., 2023).
- **uPort:** With identity verification and credential management tools, this platform lets people create and manage their digital identities on the Ethereum blockchain.
- **Microsoft's Identity Overlay Network (ION):** Providing secure and adaptable identity solutions is the aim of this open identity system based on the Bitcoin blockchain (Bandara et al., 2021).

Implications for Financial Institutions

The adoption of SSI in the financial sector could change identity verification processes. By allowing people to control and share their verified identity data, financial companies can demonstrate the ease and security of Know Your Customer (KYC) and Anti-Money Laundering (AML) compliance (Thommandru & Chakka, 2023). Because of this adjustment, individuals will prove themselves less frequently. This saves consumers money and helps. SSI allows non-standard ID holders to utilise banking services, promoting financial inclusion (Bandara et al., 2021). Blockchain-based Self-Sovereign Identity provides secure, private, and user-centric digital IDs. It may streamline bank operations, enforce standards, and build client confidence.

Privacy-Preserving Mechanisms: Zero-Knowledge Proofs

Balance privacy, safety, and security is one of the trickiest components of identity management. According to Liang et al. (2024), financial organisations must verify client identities to comply with KYC and AML. This generally involves gathering and storing personal data, which might compromise privacy. Zero-knowledge proofs (ZKPs) and other privacy-protecting features in blockchain-based identity management systems may assist (Diro et al., 2022?).

Cryptographic zero-knowledge proofs allow someone verify their identity without exposing how. Anyone 18 or older may verify it without a birthdate. Using identifying features without disclosing personal information is possible (Wishwasara, 2022).

Zero-knowledge proofs in blockchain-based identity management systems transform privacy and security. By letting people choose what information about their identities they share, ZKPs lower the risk of data breaches and identity theft. This is because users can reveal some information about themselves without disclosing all of their personal information during formation. It's very important in the financial sector, where KYC and AML rules say that customers' names must be checked, but their data must stay private (Garimella & Conway, 2024).

The integration of zero-knowledge proofs into identity management systems that are based on blockchain has not yielded any significant benefits. To begin, it keeps people's privacy safe by only letting them share proof-ready information. The risk of data breaches is also lower because private personal information is not kept in a single database. Last but not least, it validates the identity verification process for financial institutions by enabling them to check specific traits without having to view the full identity data of their customers (Diro et al., 2024).

Blockchain and Know Your Customer (KYC) Processes

Blockchain could change Know Your Customer (KYC) practices, which is one of its most important uses in the financial industry. Know Your Customer, or KYC is a big part of making sure that financial institutions follow the rules that are meant to stop scams, money laundering, and funding for terrorists (Harshini et al., 2024). People must provide personal information to many sites to establish their identity. This is expensive, time- consuming, and inefficient using KYC techniques.

Blockchain is quicker and safer than KYC for identity management. People only need to prove their identification once on an open network since it keeps and validates identity information. After checking IDs, they might provide them to several banks. This cuts down on the waste and repetition that come with traditional KYC processes while still meeting regulatory requirements (Mansoor et al., 2023).

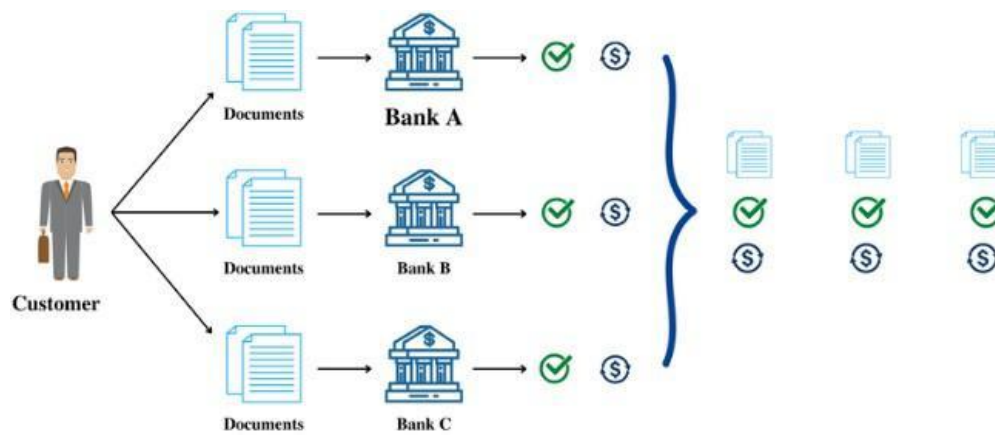


Figure 1: Traditional KYC Method (Without Blockchain) (Mansoor et al., 2023).

By using blockchain for KYC, there are many benefits. The first way it makes security better is by making data breaches less likely. People who aren't meant to see customer information can hack into fixed databases and get to it in a traditional KYC process. In contrast, identity data is spread across a network of nodes thanks to the blockchain's decentralized structure (Mamatha & Ramachandra Hegde, 2023). This makes it more difficult for cybercriminals to take over. Second, KYC processes based on blockchain are more effective because customers don't have to repeatedly provide their identity information to various groups. Once a buyer has proven their identity on the blockchain, they can give their information to any business on the network. They say that following KYC rules costs less money and takes less time.

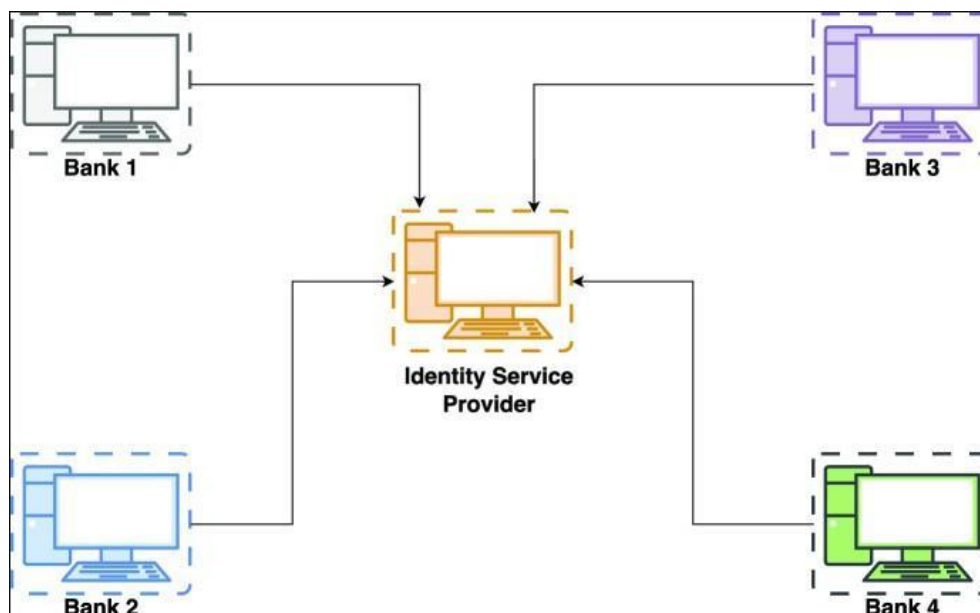


Figure 2: Architecture of current e-KYC (Hannan et al., 2023).

Blockchain-based KYC systems have many benefits, yet banks seldom employ them. Different blockchain networks must link to common ones, a major difficulty. Banks often utilise separate platforms, making it difficult to integrate blockchain-based KYC systems. Uncertain rules have also hampered blockchain identity verification adoption (Ratnawat et al., 2022).

Blockchain and Anti-Money Laundering (AML) Compliance

Adhering to Blockchain and AML laws may improve bank responsiveness by addressing issues with traditional AML approaches. AML standards require banks to monitor and report suspicious activity related to terrorism, money transfers, and theft, according to Jain et al. (2024). People must establish up data sources and work hard for conventional AML processes. It is sluggish, error-prone, and insecure (Sharma, 2024.4). Blockchain's independence, openness, and permanence make it a breakthrough AML compliance tool.

Immutable Record-Keeping and Transparency

One of blockchain's primary benefits for AML compliance is its ability to create a time-stamped, m-proof record of all transactions. There is a record that keeps track of all transactions. It can't be changed or deleted without the network's agreement. This makes it a lot harder for illegal things to happen or be faked. Aside from that, the open structure of blockchain lets politicians and financial institutions check transactions both now and in the past. For full transparency into the movement of funds across networks (Pappa et al., 2024). The integrity and speed of blockchain transactions make it easier to spot strange things like suspiciously large transactions, strangely frequent transfers, or transactions involving high-risk areas.

Enhanced Monitoring and Tracking of Suspicious Activities

Centralized databases are more likely to have data breaches than distributed ones. The decentralized form of blockchain technology shares transaction data across a network of nodes. The dispersed nature of transaction data shows its security and makes it invulnerable to hacking. Blockchain can automatically keep an eye out for odd behaviour when it is linked to AML compliance tools. This helps financial companies find and report possible money laundering more quickly and easily. The need for annual evaluations and the likelihood of mistakes is reduced by the automatic and real-time features of blockchain-based AML solutions (Thommandru & Chakka, 2023).

Integration with Identity Verification for AML Compliance

Identity management systems based on blockchain, such as self-sovereign identity (SSI) solutions, make AML compliance even better by linking identity verification with transaction data. By linking confirmed names to transaction blockchains, financial institutions can get a better idea of the risk that comes with different transactions and parties (Okpeke & Iyelolu, 2024). It's less likely that both fake positives and false negatives will happen with this integration. When legal transactions are marked as odd, this is called a fake positive. On the other hand, true negatives happen when illegal activities are not caught. According to the 2023 research, banks may comply AML standards faster by using blockchain-based identity verification to learn client patterns.

Reducing False Positives and False Negatives

Old AML systems make blunders and provide bogus hits. Sometimes phoney outcomes need a deeper examination into legal transactions. False positives conceal unlawful transactions. Gandhi et al. (2024) think blockchain can solve both problems since it records everything. This increases threat awareness. Blockchain, AI, and machine intelligence may help banks fight money laundering. These algorithms may identify patterns and weird things that other systems overlook, according to the research. These enhancements improve bank operations and AML standards.

Blockchain technology might simplify AML requirements, make bad activity simpler to identify, and reduce data breaches. Blockchain makes AML compliance easier for banks since it can verify names and trace transactions.

Challenges and Barriers to Adoption

Despite their potential, blockchain-based identity management solutions cause banks many problems. Being able to develop is crucial. Transactions cost time and resources, hence chains using Proof of Work (PoW) voting methods may not expand as rapidly. Banks that handle a lot of business and background checks struggle because blockchain-based systems cannot scale (Alghamdi et al., 2024).

The other is interoperability. Financial institutions employ so many outdated platforms and systems that connecting their old identity management systems to these new digital solutions might be difficult. Additionally, different blockchain networks may use different methods, making it difficult to ensure seamless communication between systems. It is harder to make interoperability work because blockchain technology is not defined (Gujar, 2024).

Uncertainty about regulations also hinders the adoption of blockchain-based identity management systems. Financial institutions are subject to strict regulations regarding the gathering, retention, and use of personal data in this context, and the legal status of personal data in this context is still unknown in any state. The General Data Protection Regulation (GDPR) in the European Union, for example, gives people the right to be ignored and other strict rules that data managers must follow. These rules go against the decentralised nature of blockchain (Belen-Saglam et al., 2023). Financial institutions are anticipating officials to make it clear how they can make sure they follow data protection rules before they use blockchain-based systems.

Finally, the high price of blockchain technology makes widespread adoption in financial companies challenging. In terms of training, development, and infrastructure, putting in place blockchain-based identity management systems is very expensive. When banks, particularly smaller ones, switch to a blockchain-based system, there may be more costs than benefits (Thommandru & Chakka, 2023).

Future Trends and Developments in Blockchain-Based Identity Management

The future of identity management systems in financial companies is likely to be shaped by several new trends and changes as blockchain technology continues to advance. One such trend is the creation of blockchain solutions that can be used by more people. Several blockchain projects, such as Ethereum 2.0 and the Lightning Network, are trying to demonstrate the scalability of blockchain networks by introducing new decision methods, like Proof of Stake (PoS), and off-chain solutions that lighten the load on the main blockchain network (Tripathi et al., 2023).

Another trend is the ability of blockchain networks to interact with each other. Several blockchain network projects, including Polkadot and the Interledger Protocol (ILP), want to create cross-chain communication protocols that make it simple for different blockchain networks to communicate with one another. These developments will make it possible for financial companies to link their identity systems with existing blockchain-based identity management systems (EU Blockchain Observatory and Forum, 2022).

Zero-knowledge proofs and private transactions are two technologies that will likely play a big role in the future of blockchain-based identity management. Each of these technologies helps protect privacy. As privacy laws become tighter, financial institutions will be required to use technologies that allow them to check people's names while still adhering to data security rules (Garimella & Conway, 2024).

Summary

The adoption of blockchain-based identity management systems in financial groups has a significant chance to improve security, privacy, and operational efficiency. Blockchain technology allows you control your digital name, making it safer and better than existing identity management methods. Zero-knowledge proofs help banks comply with the law and secure consumer data. We must resolve issues such as plan size, linkages, legality, and cost. Future cryptocurrency-based financial identity management will be different. New technologies will make them safer, more versatile, and able to link to other systems. What should banks do before using blockchain? They should weigh the advantages and downsides and find methods to avoid it. To ensure blockchain-based identity management solutions operate, banks must do this. This will test name verification and secure their systems. It improves consumer experience too. Blockchain technology will certainly transform identity management. Because the global banking industry is increasing.

METHODOLOGY

Introduction

This section discusses how the research will examine how money-related firms might utilize blockchain to govern identities. The study uses quantitative research and secondary data to examine how blockchain may improve identity management systems, particularly security and fraud prevention. We will utilise Kaggle's Financial Fraud Detection Dataset to identify system faults and how blockchain may address them. Initial explanations of the study's objective and techniques are brief. A detailed dataset explanation follows. We will next discuss cleaning and preparing data for analysis. Next, we will utilize prediction models and link analysis to examine how blockchain-based identity management might reduce frauds. The chapter concludes with ethical and study boundary issues.

Research Design and Approach

The quantitative study examines blockchain-based identity management solutions' potential. Quantitative approaches can quantify variables and find patterns in big datasets, making them ideal for our investigation. The Financial Fraud Detection Dataset contains extensive data on identity fraud, transaction security, and financial sector vulnerabilities. The paper analyses these trends to show how blockchain-based solutions might overcome these problems.

As a supplementary dataset analysis, the data has already been gathered and made accessible on Kaggle, a prominent data science competition and dataset platform. Secondary data analysis lets researchers concentrate on evaluating existing data and deriving new findings without the hassle of data gathering. Financial firms must protect sensitive consumer data, therefore this technique assures analysis without compromise.

Exploratory study tries to link identity-related fraud trends to blockchain system advantages. The research examines whether blockchain can prevent identity theft, secure transactions, and boost financial institution efficiency. The study examines identity management failure patterns and blockchain technology's possible benefits. Although the information is not blockchain-specific, the fraud tendencies studied here may help explain how blockchain might reduce identity threats.

This exploratory approach is underpinned by **hypothesis-driven research**. The key hypotheses guiding this study are:

- **H1:** Blockchain-based identity management systems may considerably minimize identity theft in financial transactions (H1).
- **H2:** Traditional identity management systems have flaws that blockchain technologies can handle more effectively (H2).
- **H3:** Integrating blockchain technology with KYC and AML procedures would enhance the efficiency and security of identity verification in financial organizations.
-

These hypotheses will be investigated using predictive modelling, correlation analysis, and fraud detection performance indicators.

Data Source

Financial Fraud Detection Dataset

The Financial Fraud Detection Dataset from Kaggle underpins this work. This dataset shows financial transactions with labels suggesting fraud or not. Transaction characteristics and identity verification methods make the dataset excellent for analyzing conventional identity management system weaknesses.

The dataset contains key factors for analysing identity fraud. Variables include:

- **Transaction ID:** A unique identifier for each financial transaction.
- **Transaction Time:** The timestamp associated with each transaction, which could help detect patterns over time.
- **Transaction Amount:** The monetary value of the transaction, which may influence the likelihood of fraud.
- **Customer Identity Information:** Data features related to the customer, which are used to verify their identity.
- **Fraud Label:** A binary variable indicating whether the transaction was classified as fraudulent (1) or non-fraudulent (0).

These factors provide the researcher a complete picture of financial transactions, enabling them to study how typical identity management systems may miss fraud. The investigation can detect dataset trends to determine where blockchain-based technologies may give greater safety.

The transaction time and amount variables will identify transaction patterns and abnormalities, while the client identification information will proxy conventional identity verification methods. The binary fraud label will be the predictive modelling dependent variable and help determine dataset fraud prevalence.

Data Preparation and Cleaning

Data preparation is essential for clean, trustworthy, and analysed datasets. The Financial Fraud Detection Dataset, like many big datasets, may include missing values, outliers, or incorrectly structured data that must be rectified before analysis. Cleaning the data provides reliable analysis and valid findings.

Data preparation begins with missing data. Poorly managed incomplete records might skew or misrepresent outcomes. This dataset may have missing transaction time, amount, or customer identification variables. Deleting incomplete records or imputed missing values from other dataset information are two ways to handle missing data. Missing values will be imputed using the median or mean for variables like transaction amounts that are crucial to the investigation. This method reduces outliers and preserves valuable information.

Outlier detection follows. Extreme numbers in the dataset are called outliers and may skew statistical analysis. Outliers are significant for fraud detection since fraudulent transactions may have unusual quantities or timings. Not all outliers constitute fraud, thus rigorous investigation is required to separate them. Outliers will be identified and managed using Z-score or IQR methods. Rare fraudulent transactions will be kept for study, but extreme non-fraudulent transactions may be eliminated to minimise data skewing.

Normalization of data

Data normalization is another crucial stage in data preparation. The information includes transaction amount, which might vary greatly. Normalization methods like min-max scaling or Z-score standardization will be used to compare these variables throughout analysis. Normalization prevents a single variable from skewing analytical findings.

Feature Engineering

New variables will be created via feature engineering to better understand identity management vulnerabilities. Analysing client transaction patterns over time, recognising consumers with abnormally large transaction volumes, or detecting fraud-related transaction behaviour changes may add functionality. These designed traits will improve this study's models' prediction capability. The dataset fraud label will be used to label the data. This binary label (0 for non- fraudulent transactions, 1 for fraudulent transactions) will be the study's predictive model goal. Labelled data will enable supervised machine learning algorithms to identify whether a transaction is fraudulent based on its features.

Analytical Techniques

The research seeks identity theft tendencies and blockchain-based solutions. Performance reviews, link analysis, prediction models, and data summaries are examples.

In the first section of the research, we will discuss key data points using summary statistics. Summary statistics include mean, median, mode, standard deviation, and range. You can view theft amounts, times, and incidences using this data. Descriptive studies may help identify fake transactions. That implies many identity management solutions are inefficient.

Later, correlation analysis would find set factor relationships. Finding faults and falsehoods is the study's major objective. Amount, kind, and duration of agreements might link scams to behaviour. A significant correlation between identity-related features and frauds makes blockchain-based solutions safer for confirming identities. Pearson or Spearman data shows relationship strength and direction.

The study's most crucial step is mapping out the future. The collection only contains one aim variable, genuine or phony transactions. We will make important predictions using logistic regression. Many people use logistic regression to address binary classification issues, and it works here. It will estimate the likelihood of a false transaction based on the amount, timing, and customer information. Logistic regression can identify qualities most likely to be fraudulent and those most likely to benefit from blockchain-based identity management solutions.

Advanced machine learning and logistic regression improve predictions. We will create a group model with complex variable linkages using a Random Forest classifier. Random Forest can handle big datasets and trait relationships, making it useful for fraud detection. The model will rank each characteristic based on its importance to identify identity-related aspects most strongly connected to theft. They will demonstrate what identity management methods cannot achieve and how blockchain can.

Predictive models will undergo several tests to determine their efficacy. The F1 score, accuracy, and memory of a model indicate its performance. Precision is the percentage of predicted false transactions discovered accurately. However, memory is the quantity of actual false transactions accurately discovered. The F1 score balances accuracy and recall. Blockchain-based solutions may reduce bogus hits and declines. This may improve fraud detection. We can assess whether blockchain-based identity management solutions make theft detection simpler by comparing their performance to others.

Ethical Considerations

No matter if this research utilized existing data, ethics are still crucial, particularly when dealing with money and identities. Privacy is a vital value. Despite making the financial information in the record secret, be cautious with it. The data will be secure and only available to research participants. Security and safekeeping will protect data.

Another moral issue is critical prejudice. The data collector may have biases like over- or under-represented purchasing groups. Biases may make prediction models less accurate and biased. The study will seek for unequal data and apply SMOTE to ensure fairness and equal treatment of all client categories. This reduces danger.

Finally, blockchain technology alters morality. Decentralization and exploitation of blockchain are challenges, even if it may improve and increase privacy. The paper's conclusion will discuss moral problems, including how banks should evaluate blockchain-based identity management systems' pros and cons.

Limitations

This strategy is excellent but has several drawbacks. Outside sources restrict the study's scope and quality. The collection aims to detect money-stealing frauds. None of these solutions utilize blockchain, however some manage IDs. The research will not employ blockchain evidence because of this. Instead, it will extrapolate theory from data.

Second, the data may not contain identity management tools to show how terrible typical systems are. Biometrics or numerous proofs may be valuable for future research. The customer's identity and behaviour throughout the trade are significant, but not the only ones.

Third, outcomes vary by person. The data only shows occurrences from one context; thus, it may not completely represent blockchain technology's merits and downsides in multiple nations or banks. The data requires additional research, but it suggests blockchain may be valuable.

Conclusion

We discuss a strategy to examine blockchain-based identity management technologies in banks. The project will utilise the Financial Fraud Detection Dataset to detect identity fraud and test blockchain's effectiveness. Descriptive statistics, connection analysis, and prediction models will reveal the flaws in existing identity management systems and the merits and downsides of blockchain.

Results of the investigation will reveal more about blockchain technology and its applicability in banks. In this research, banks interested in blockchain-based identity management solutions will discover beneficial information. These ideas focus on how blockchain technology may reduce frauds, make things safer, and streamline operations. This research will examine blockchain technology's moral and practical challenges and offer solutions.

RESULTS AND FINDINGS

Introduction

The results and findings section describes the Financial Fraud Detection Dataset research. The research examined fraud and the advantages of blockchain-based identity management for banks. This research seeks to identify false transactions, their causes, and identity management system issues. Descriptive data, visualizations, and forecast models support findings.

Descriptive Statistics

Descriptive statistics provide key data points. Over 655,000 occurrences are in this collection. The most crucial factors are the transaction number, kind, amount before and after, and fraud warnings.

Key Descriptive Insights

- **Transaction Amount:** Trade quantities vary from 162,666.5 to 271,554.9. The median transaction value is 76,849.4, while most purchases are less. The greatest transaction figure is \$10,000,000; therefore, the spread is huge.
- **Old Balance of the Originator:** The average balance was \$886,117.4 and the standard deviation \$2,955,015 before the sale. This indicates account balance differences. If the old number is zero, many activities originate from empty accounts.
- **New Balance of the Originator:** Account holders have an average of \$906,217.2 more after the transaction. This indicates how the exchange affected account funds. Many true bargains leave you with nothing extra.
- **Fraudulent Transactions:** 398 of 655,045 deals are not real. This represents 0.06% of agreements. Theft is rare, but when it does, it is costly.

Descriptive Statistics Summary Table

This table shows important trade and fund statistics. The material contains a few scams, most of which involve empty accounts.

	step	amount	oldbalanceOrig	newbalanceOrig	\
count	655045.000000	6.550450e+05	6.550450e+05	6.550450e+05	
mean	17.164091	1.626665e+05	8.861174e+05	9.062172e+05	
std	7.411359	2.715549e+05	2.955015e+06	2.992070e+06	
min	1.000000	1.000000e-01	0.000000e+00	0.000000e+00	
25%	12.000000	1.253586e+04	0.000000e+00	0.000000e+00	
50%	16.000000	7.684943e+04	1.735000e+04	0.000000e+00	
75%	19.000000	2.167798e+05	1.538550e+05	1.938016e+05	
max	35.000000	1.000000e+07	3.893942e+07	3.894623e+07	
	oldbalanceDest	newbalanceDest	isFraud	isFlaggedFraud	
count	6.550440e+05	6.550440e+05	655044.000000	655044.0	
mean	9.767415e+05	1.137647e+06	0.000608	0.0	
std	2.319012e+06	2.471123e+06	0.024642	0.0	
min	0.000000e+00	0.000000e+00	0.000000	0.0	
25%	0.000000e+00	0.000000e+00	0.000000	0.0	
50%	1.161133e+05	2.103664e+05	0.000000	0.0	
75%	8.969513e+05	1.173688e+06	0.000000	0.0	
max	4.148270e+07	4.148270e+07	1.000000	0.0	
step	0				
type	0				
amount	0				
nameOrig	0				
oldbalanceOrig	0				
newbalanceOrig	0				
nameDest	0				
oldbalanceDest	1				
newbalanceDest	1				
isFraud	1				
isFlaggedFraud	1				
dtype: int64					

Fraud vs Non-Fraud Transactions

Importantly, the research indicated that real and phony trades are extremely different. An easy-to-follow pub plan shows scammers' spread:

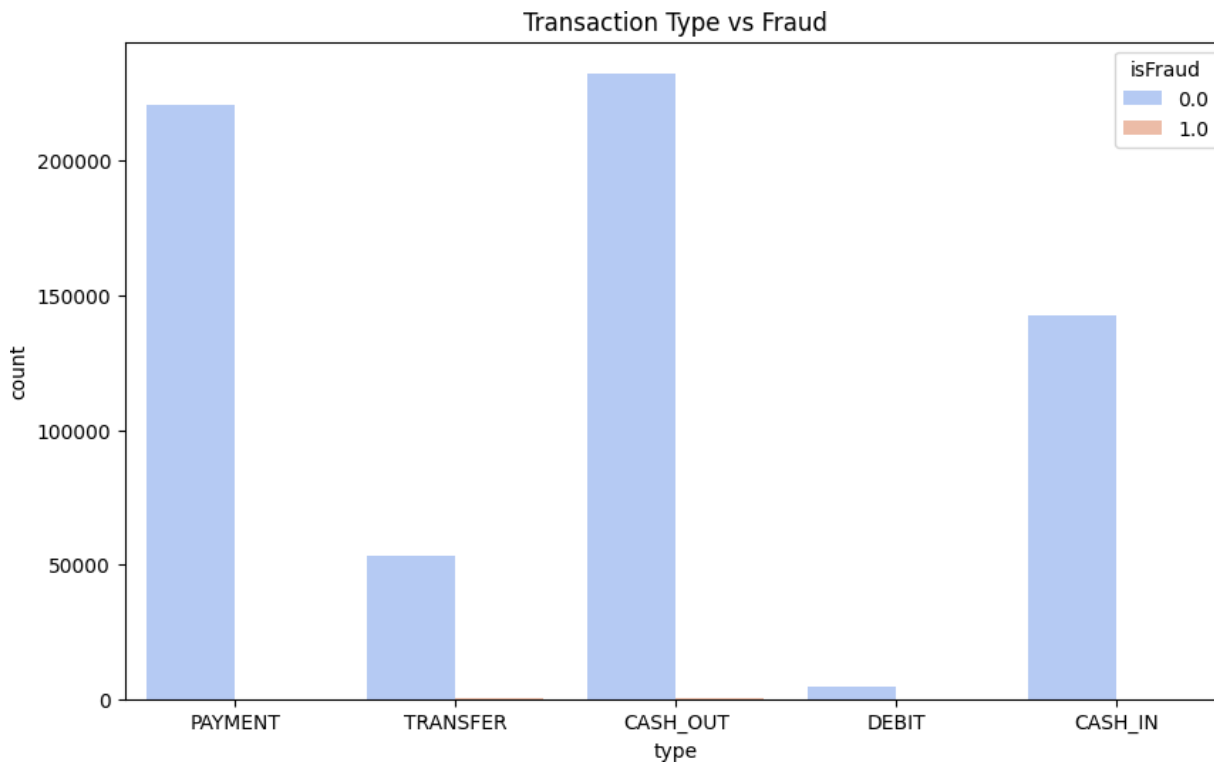


Key Insights:

- The list has 654,646 legitimate transactions and 398 fraudulent ones.
- Different quantities indicate how hard it is for banks to detect uncommon but costly frauds. It also highlights how a stronger identity management system like blockchain may improve monitoring and protection.

Transaction Type and Fraud Distribution

Scams also depend on transaction type. Activity kinds include DEBIT, TRANSFER, PAYMENT, and CASH_OUT.

Fraud Distribution by Transaction Type

CASH_OUT and TRANSFER transactions indicate which schemes are true.

Key Findings:

- 205 cash out thefts.
- Bank swaps rank second with 193 frauds.
- No scams have occurred with PAYMENT, DEBIT, or CASH_IN trades. This may be because these trades are safer or because fraud detection works better for them.

This makes certain offers more likely to be phoney. Moving or paying out money may be so easy no one notices. Blockchain technology, which creates unchangeable transaction records, might remedy this.

Transaction Amounts: Fraud vs Non-Fraud

Another key conclusion from the investigation was that phoney and real transactions were different amounts.



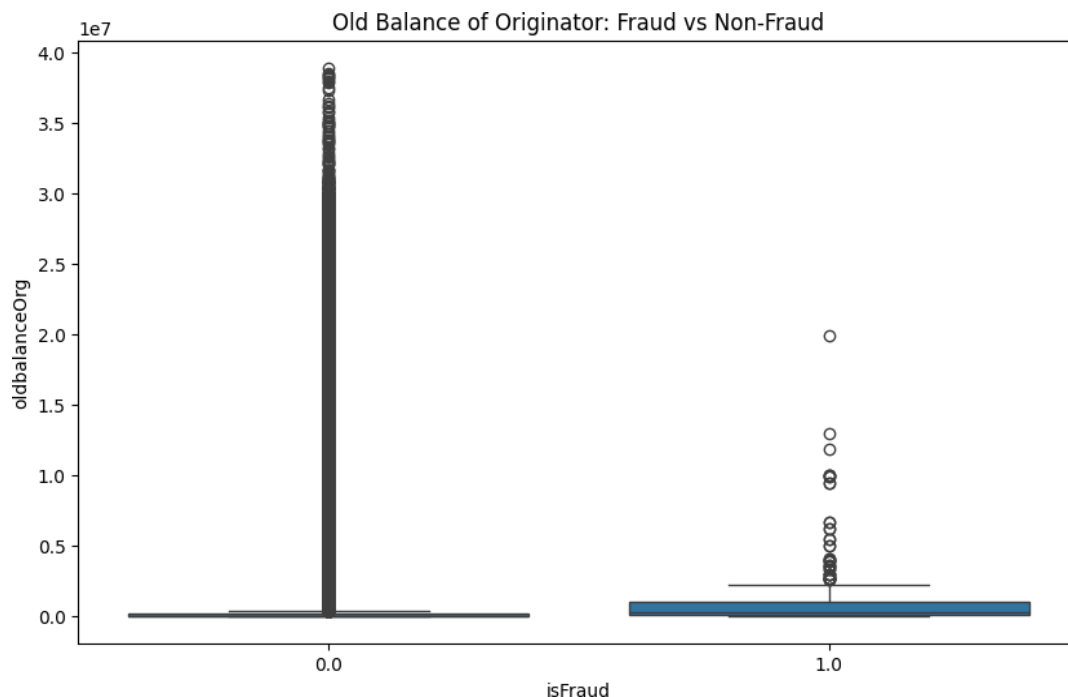
Descriptive Statistics for Transaction Amounts

When fraud is prevalent, transactions average much higher.

- **Fraudulent Transactions:** 1,829,551 and \$921,161.9 are the average fraud transactions. Transactions worth 10,000,000 may be false, making theft more likely.
- **Non-Fraudulent Transactions:** 162,205.5 is the average non-theft purchase, while 6,419,835.60 is the largest. That proves most court agreements are cheaper.

Transaction Amount Visualization

This box plot shows fraudulent and non-fraudulent deals:

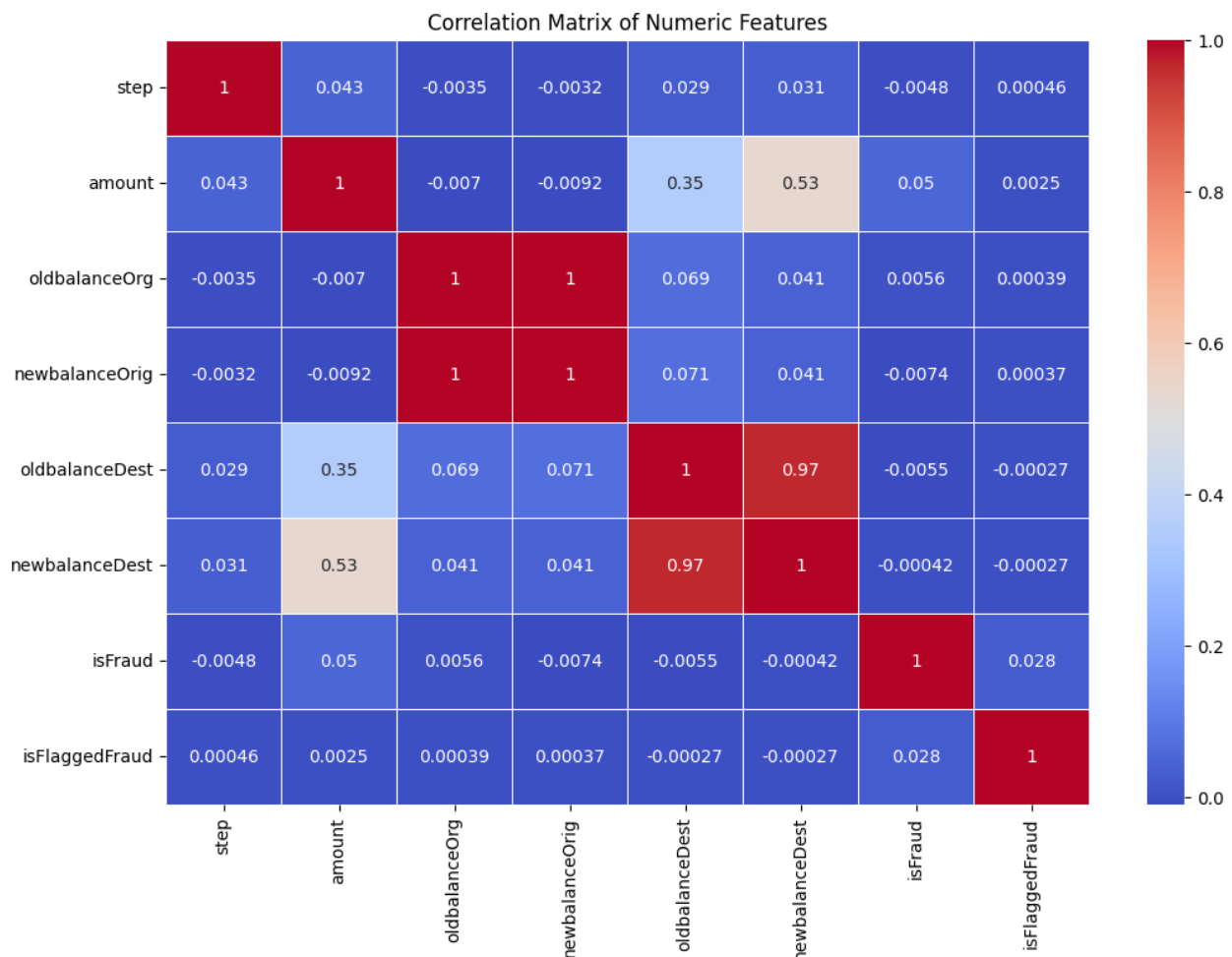


Key Findings:

- Fraudulent transactions have a larger range of quantities.
- Most fraudulent transactions are for greater sums, suggesting thieves target huge transfers to earn the most money.
- This suggests that blockchain, which can reveal and record transactions, might prevent big-scale fraud by making it harder for criminals to transfer significant quantities of money.

Correlation Analysis

To determine theft relationships, we employed an association matrix to detect ties between data numbers.



Key Findings:

- Transaction amount affects fraud risk. Thus, fraud increases with transaction amounts.
- It is notable that theft does not affect previous balances. That's other characteristics do not strongly correlate with theft proves we need innovative technology like blockchain to secure more than company data.

The Role of Blockchain in Addressing Fraud

Comparing false and genuine events in the dataset indicates that conventional identity management systems are faulty in many ways. Fraud during cash-outs and transactions is prevalent and wipes away enormous sums of money. Blockchain-based identity management may address these difficulties in numerous ways. Blockchain's permanent ledger of transactions is unchangeable. Scammers find it difficult to alter transaction records or conceal unlawful activity when everything is obvious. Blockchain is uncontrolled, therefore identifying evidence is not company-dependent. Instead, the blockchain network has numerous sites to authenticate consumer identities. This reduces identity theft and frauds.

Self-sovereign identification implies each individual controls their own data. Blockchain can aid this notion. Since consumers no longer need unified identity businesses to exchange their information with all their banks, identity theft is less probable. Blockchain can store smart contracts with coded terms. Contracts run themselves. Using particular criteria, smart contracts might identify suspicious transactions or cancel false transactions automatically. It would halt thievery immediately.

Limitations and Challenges in Blockchain Adoption

Blockchain might make identity checks and fraud detection simpler for banks. However, it needs several fixes before many people can use it. When transactions are high, Proof of Work (PoW) blockchains struggle to expand. Banks that work constantly may struggle with this. Blockchain-based identity management is difficult to integrate with many banks' existing techniques.

Blockchain's absence of regulations makes it contentious in law. This is particularly true in countries with rigorous privacy laws like Europe's GDPR. How blockchain may balance data deletion and non-change is unclear. Identity management systems might be safer, more open, and better at preventing crime using blockchain technology.

Summary

The research indicated that financial transaction theft is rare yet common. Large quantities of money and trades and cash-outs are typical. Fraud frequently includes significant sums and empty accounts. These patterns suggest that our name-handling and scam-hunting methods are poor. Blockchain technology may fix this. Blockchain identity management solutions may assist financial institutions detect fraud and halt large-scale theft since they are open, secure, and decentralized.

Blockchain can offer users identity control and eliminate frauds via smart contracts, making it an interesting financial services notion. Before many financial firms utilize blockchain, growth, transparency, and legality must be addressed.

This study shows that blockchain technology requires further research and development to suit major bank demands. Blockchain might transform how banks handle names by solving these difficulties. It would help people move quicker and safer.

CONCLUSION

Introduction

Better digital technologies affect how banks function, locate individuals, and save private data. Some changes are good, but frauds and safety are bad. ID management is crucial since outdated, established methods allow more frauds, data breaches, and waste. As blockchain technology gains popularity, more solutions are available. This research examined what might happen if banks implemented blockchain-based identity management technologies. It examined fraud trends in a public financial crime database. This section discusses what practice and study can and cannot achieve and what people expect in the future. It outlines the key results and defines "adoption of blockchain".

Summary of Key Findings

Using over 655,000 bank transactions, this analysis. A mere 0.06% of fakes occurred. Fake transactions were rare but frequently for larger sums and deals or cash runs. They altered many people's finances. Important findings from this investigation included:

- **Transaction Amounts and Fraud:** Real transactions were valued more. To maximise profits, con artists desire a lot of money. Average actual deal value was far lower. The theft sometimes exceeded the dataset's transaction limit. This encourages robbers to steal enormous quantities of money. This increases bank risk.
- **Transaction Type and Fraud:** CASH_OUT and TRANSFER transactions were more fraudulent. The numerous frauds involving these trades illustrate that traditional mechanisms do not monitor them adequately or properly. However, PAYMENT, CASH_IN, and DEBIT agreements showed little theft. This suggests that these agreements are safer or that theft-prevention technologies function better.
- **Fraud Detection Mechanisms:** Only one theft report was minor. This implies our existing scam-hunting methods, which rely on individuals reporting suspicious conduct, may not work. That is why we need faster, larger-scale fraud detection techniques. Smart contracts and distributed evidence in blockchain-based identity management systems may help.
- **Correlations and Vulnerabilities:** Larger sales increased the risk of fraud, hence spending linked the two. The relationship between numbers before and after movements was weak. Fake transfers typically caused large drops in cashouts and transfers. Id management technologies should verify identities when major persons do them.

Implications for Blockchain-Based Identity Management Systems

Bank identity control tools are poor, according to this report. Our technologies make it hard to track and reveal little and large trades that are hard to see. Blockchain-based identity management tools may help:

- **Decentralized Identity Verification:** Blockchain technology may make user name verification simpler. The responsibility does not fall to one group or corporation. This improves name confirmation across many blockchain groups. Thieves will find it tougher to move. Blockchain is unchangeable, so everyone can see if someone is stealing.
- **Immutability and Transparency:** Blockchains store transactions in an unchangeable format. They cannot alter or conceal phony transactions. With an authorized and documented trade, the blockchain will always record it. This record may aid future research. Too open for thieves to operate. They would know they are constantly monitored.
- **Self-Sovereign Identity (SSI):** SSI strategies assist individuals manage their personal data. No one needs to trust big companies to protect their data. Users decide how to keep and distribute private data. Identity theft is less likely since no one can advise others how to share or utilize their personal information. Blockchain has no one authority, therefore no one organization controls personal data. Now, data breaches are rarer.
- **Smart Contracts for Automated Fraud Prevention:** Blockchain "smart contracts," which happen immediately, prevent theft. In addition, Blockchain may employ them. Smart contracts enable you specify rules to automatically terminate suspicious agreements. Trades and cash-outs are risky, we learnt. This list includes high-risk or large transactions begun without checks. If any of these occurs, the smart contract may delay the transaction or flag it for review. This real-time technology may help banks discover frauds quicker.

Limitations of the Study

This research revealed several good aspects and some issues:

- **Limited Scope of the Dataset:** This research focused on fraud detection and money management. Although it was thorough, it did not explain how banks operate accounts. Because of this, the research could only suggest broad ways bitcoin may make identity more transparent. Researchers should seek for future records with fingerprint verification or information on several IDs.
- **Focus on Fraud Detection:** Identity management goes beyond preventing theft, although the research focused on fraud. Experts should investigate how blockchain-based systems might speed up, enhance, and obey identity management regulations in the future, as well as fraud prevention.
- **Generalizability of Findings:** The research only comprises agreements from a single kind of case, thus its conclusions may not be applicable for other financial businesses. Since scam rates and trends vary by location, group, or legal situation, utilise the information with caution.

- **Blockchain Adoption Barriers:** It is inflexible and incompatible with other applications, making it difficult to use even if it can cure many issues. Matching rules is also difficult. Blockchain may prevent frauds since hackers can not modify it. Privacy rights under the GDPR include erasing or changing personal data. This may not work. Remember that blockchain networks, particularly proof-of-work ones, require a lot of electricity.

Future Considerations

Blockchain-based name management systems require certain tweaks before they are useful:

- **Scalability Improvements:** Proof-of-work blockchains face issues as they grow. Improvements to proof-of-stake and sharding in identity management systems should allow blockchain to handle banks' enormous transaction volumes.
- **Regulatory Alignment:** Banks must observe tight customer data and identity laws. Blockchain systems may need to adapt to these regulations. Creating better tools like zero-knowledge proofs, which enable users prove their identities without sharing personal information, might help.
- **Cross-Industry Collaboration:** Banks may adopt blockchain only by working with several organisations. These groups include internet users, banks, the government, and tech firms. To take off, blockchain needs a consistent method to handle names, regulations that enable autonomous systems, and blockchain platforms that can communicate.

Conclusion

Still, blockchain-based identity management solutions may improve identity management, notably for fraud prevention. Blockchain's open, independent, and continuous portions may protect financial transactions, reduce identity theft, and help identify frauds. But financial services must completely adopt blockchain first. Connection, scale, and law enforcement are only possible then. Banks must improve and lower expenses to comply with tight standards.

The blockchain might let them accomplish this and more. Blockchain might revolutionize how banks handle identities and trades online. Its smart contracts can instantly identify scammers and give them personalized identities.

REFERENCES

1. Alghamdi, T.A., Khalid, R. and Javaid, N. (2024). A Survey of Blockchain based Systems: Scalability Issues and Solutions, Applications and Future Challenges. IEEE access, [online] pp.1–1.
2. Bandara, E., Liang, X., Foytik, P., Shetty, S. and Zoysa, K.D. (2021). A Blockchain and Self-Sovereign Identity Empowered Digital Identity Platform. 2021 International Conference on Computer Communications and Networks (ICCCN). [online]
3. Belen-Saglam, R., Altuncu, E., Lu, Y. and Li, S. (2023). A systematic literature review of the tension between the GDPR and public blockchain systems. Blockchain: Research and Applications, [online] 4(2), p.100129.
4. Diro, A., Zhou, L., Saini, A., Kaisar, S. and Hiep, P.C. (2024). Leveraging zero knowledge proofs for blockchain-based identity sharing: A survey of advancements, challenges and opportunities. Journal of Information Security and Applications, [online] 80, pp.103678–103678.
5. Dong, S., Abbas, K., Li, M.Y. and Kamruzzaman, J. (2023). Blockchain technology and application: an overview. PeerJ, [online] 9, pp.e1705–e1705.
6. EU Blockchain Observatory and Forum (2022). The current state of interoperability between blockchain networks. [online]
7. Gandhi, H., Tandon, K., Gite, S., Pradhan, B. and Alamri, A. (2024). Navigating the Complexity of Money Laundering: Anti-money Laundering Advancements with AI/ML Insights. International Journal on Smart Sensing and Intelligent Systems, [online] 17(1).
8. Garimella, K.K. and Conway, D. (2024). Zero-Knowledge Proofs and Privacy: A Technical Look at Privacy. Technology, work and globalization, [online] pp.157–179.
9. Gujar, V. (2024). Identity Management, SSI and Blockchain: A Review. International Journal of Science and Research (IJSR), [online] 13(1), pp.1547–1553.
10. Hannan, Md.A., Shahriar, Md.A., Ferdous, Md.S., Jabed, M. and Rahman, M.S. (2023). A systematic literature review of blockchain-based e-KYC systems. Computing.
11. Harshini, D., Keerthi Vardhani, V. and Priya, V. (2024). Impact Factor: 8.423 International Journal of Innovative Research in Science, Engineering and Technology (IJRSET) IJRSET©2024 | An ISO. IJRSET, [online] 13(3), p.2304.
12. Idrees, S.M. and Nowostawski, M. (2024). Blockchain Transformations. Signals and communication technology. [online]
13. Imtiaz, A., Rozario, R.G., Chakraborty, P., Talukder, P.C. and Roy, P. (2023). Smart Identity Management System Using Blockchain Technology.[online]
14. Jain, V., Balakrishnan, A., Chintale, P., Gadiparthi, S. and Najana, M. (2024). Blockchain Empowerment in Sanctions and AML Compliance: A Transparent Approach. International journal of computer trends and technology, [online] 72(5), pp.11–26.
15. Josyula, H.P. (2023). Digital Identity Verification: Transforming KYC Processes in Banking Through Advanced Technology and Enhanced Security Measures. International Research Journal of Modernization in Engineering Technology and Science, [online] 5(9).
16. Lim, S.Y., Tankam Fotsing, P., Almasri, A., Musa, O., Mat Kiah, M.L., Ang, T.F. and Ismail, R. (2024). Blockchain Technology the Identity Management and Authentication Service Disruptor: A Survey. International Journal on Advanced Science, Engineering and Information Technology, [online] 8(4-2), p.1735

17. Mamatha, M. and Ramachandra Hegde, R. (2023). KYC Verification Using Blockchain. © 2023 IJNRD |, [online] 8(6), p.21.
18. Mansoor, N., Antora, K.F., Deb, P., Arman, T.A., Manaf, A.A. and Zareei, M. (2023). A Review of Blockchain Approaches for KYC. IEEE Access, [online] 11, pp.121013– 121042.
19. Okpeke, P. and Iyelolu, V. (2024). Anti-Money Laundering Compliance and Financial Inclusion: A Technical Analysis of Sub-Saharan Africa. GSC Advanced Research and Reviews, [online] 19(3), pp.336–343.
20. Pappa, E., Georgitseas, P. and Tantis, G. (2024). Exploring the role of blockchain technology in anti-money laundering. Journal of Legal Ethical and Regulatory Issues, [online] 27(S3), pp.1–10.
21. Pava-Díaz, R.A., Gil-Ruiz, J. and López-Sarmiento, D.A. (2024). Self-sovereign identity on the blockchain: contextual analysis and quantification of SSI principles implementation. Frontiers in Blockchain, [online] 7.
22. Ratnawat, N., Pandey, S., Paradkar, R. and Banerjee, S. (2022). Optimizing the KYC Process using a Blockchain based approach. ITM Web of Conferences, [online] 44, p.03039.
23. Schardong, F. and Custódio, R. (2022). Self-Sovereign Identity: A Systematic Review, Mapping and Taxonomy. Sensors, [online] 22(15), p.5641.
24. Sharma, R. (2024). Revolutionizing Anti-Money Laundering in Banking with Artificial Intelligence and Data Analytics. [online] ResearchGate.
25. Sun, N., Zhu, C. and Liu, Y. (2024). A Self-Sovereign Identity Privacy-Preserving Scheme for Logistics Transportation Based on One-Time-Use Tokens. Electronics, [online] 13(14), pp.2799–2799
26. Thommandru, A. and Chakka, D.B. (2023). Recalibrating the Banking Sector with Blockchain Technology for Effective Anti-Money Laundering Compliances by Banks. Sustainable Futures, [online] 5, p.100107.
27. Tripathi, G., Ahad, M.A. and Casalino, G. (2023). A Comprehensive Review of Blockchain technology: Underlying Principles and Historical Background with Future Challenges. Decision Analytics Journal, [online] 9(1), p.100344.
28. Weigl, L., Barbereau, T. and Fridgen, G. (2023). The construction of self-sovereign identity: Extending the interpretive flexibility of technology towards institutions. Government information quarterly, [online] 40(4), pp.101873–101873.
29. Wishwasara, J. (2023). Zero Knowledge Proofs: A Comprehensive Review of Applications, Protocols, and Future Directions in Cybersecurity. ResearchGate. [online]