

SDS-IAM: Secure Data Storage with Identity and Access Management in Blockchain

Sahil Sikarwar^a, N. Jeyanthi^a, R. Thandeeswaran^a, and Hamid Mcheick^{b,*}

^a*School of Computer Science Engineering and Information Systems, Vellore Institute of Technology, Tamilnadu, India*

^b*Computer Science Department, Université du Québec à Chicoutimi, Quebec, Canada*

Abstract

Identity and Access management (IAM) [1] plays an important role when it comes to background verification. It is a great way to know people you are working with, whether it is a professional front or some local business. Identity theft and secure document exchange are major issues with the current scenario and blockchain offers to be a great solution. The introduction of the public key, private key, transaction verification and foot printing will play a significant role in securing IAM. The idea is to store user documents and other critical information inside the block chain. All the verification is given based on the user's consensus to the particular request which will trigger further functionalities, responsible for secure data exchange. According to the property of blockchain, the chain will contain the history of each transaction that keeps track of every user-company activity that will prevent any action which is against the will of both parties.

Keywords: IAM; blockchain; consensus; identity management; transactions; crypto currency; MetaMask; Ethereum; public key; private key

(Submitted on October 30, 2023; Revised on November 30, 2023; Accepted on December 21, 2023)

© 2024 Totem Publisher, Inc. All rights reserved.

1. Introduction

Identity and Access Management is the process and technology that ensures a secure, stable, accurate, and efficient way to deliver access to the authorized users, to the system, information and other applications at any time. As the name suggests, it is the ability to see through the entire system including different technological fields, who have access to what information and what action they can perform with that authority.

A good Identity management system provides a greater overview of on-board and off-board users and their access to the information, systems, and different applications depending on the relations with the company. If implemented properly, it provides safety against different types of user attacks like general identity chaos, password chaos, and it also keeps track of activity logs.

There is a great need for a strong Identity and Access management system. Robust IAM solutions can enable industries to increase their employee productivity and support their overall security conditions. The expansion of cloud computing and expanding the workforce makes the IAM complex every day.

People who are in charge of user identity authentication must follow a protocol that helps them ensure the robust infrastructure and security control while making an effective authentication process to increase productivity. Here, IAM plays an important role in user empowerment and preventing them from sabotaging a company's reputation or security.

The need for decentralization, robust process verification and prevention of unauthorized manipulation in the system can be fulfilled by Blockchain. Blockchain is an immutable linked list. On a more technical level, Blockchain is an electronic public ledger that is based on peer to peer systems openly shared among various users to create unchangeable records of transaction, each linked to its previous record and time-stamped. As shown in Figure 1, you can see that every time a set of transactions is added, a particular record becomes another block in the chain. Blockchain can only be updated by the consensus of different

* Corresponding author.

E-mail address: njeyanthi@vit.ac.in

parties, once the data is added inside the blockchain, it can never be erased. Data can be appended ahead of it or be used to update the state of the blockchain at that particular time. Blockchain is a revolutionary technology that can be used to solve issues with data sharing, storage, and retrieval. It has a secure and dependable decentralized architecture [2].

The combination of asymmetric keys and digital signatures data ownership more secure. If user A makes a transaction of some data, then he/she cannot refuse the ownership of that event. Also, if a transaction is coming from user A, then there is no doubt that there was consensus involved at every stage of the processing. Although it may not be everyone's preferred solution, the consensus that has been reached should be a workable solution that is endorsed by all members of the network [3]. A Blockchain network's effectiveness depends on the process of choosing and implementing the best consensus protocol [4].

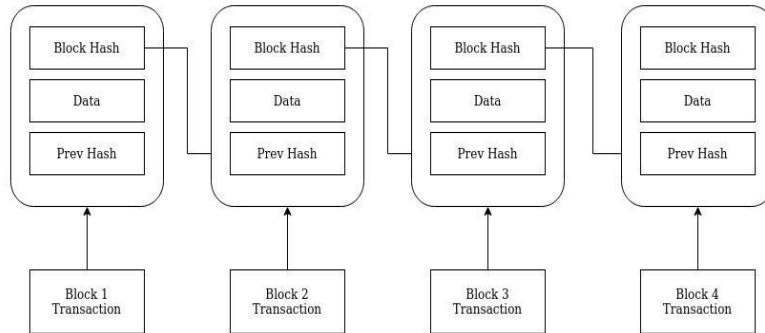


Figure 1. Block Chain Structure

2. Literature Review

Block chain is used to execute identity management venture [5] with essential functionalities and its wide scope of utilizations over the web, for instance, Bitnation, e-Residency, ConsenSys, BlockVerify, BlockAuth, and so on. There are organizations that are in incredible misfortune in light of these fake items everywhere throughout the market [6]. This exploration gives us an extraordinary arrangement and decrease in fake items utilizing blockchain. By utilizing the property of unchanging nature and verification of work idea, blockchain makes an ideal match to move toward issues this way and subsequently improve an exchanging market. Table 1 gives a review of the advances utilized, which items are upheld and focused on, and the focal point of the arrangements.

An application designed to prevent the counterfeit of medications aims to safeguard against the misuse of the drug's identity and quality, as misuse can lead to various problems in society. Any fake medication may have a miscounted measure of components and can cause an absolutely inverse response [7]. This is an incredible danger to the organization and individuals. Here, blockchain assumes a significant job to get things going. An overview of blockchain models exhibited in Table 2.

Ongoing maltreatment of individual data from web-based life stages and various client identity information breaks raise worries about specialized, business, and moral parts of the protection and security of client information. In this way, a reasonable plan has been proposed for a Blockchain-based Personal Data and Identity Management System (BPDIMS), a human-driven, and GDPR-consistent individual information and identity management framework dependent on blockchain innovation [8]. The identity-based Signature scheme with multiple authorities proposed in this work [9] uses a combination of powerful signature mechanisms and blockchain technology to resist collusion attacks. A hybrid model using private and consortium blockchain has been proposed for secure storage and sharing of EHR data [10].

Table 1. Overview of Related Work

| Types | Products | Focus | Technology |
|-------------|---|--|--|
| Blockverify | Luxury Items, Diamonds, Electronics, Pharmaceuticals | Counterfeit | Blockchain (Bitcoin and private), custom tags, mobile app |
| Chronicled | Initially: Sneakers Now: Everything | Counterfeit, Provenance, Supply chain | Blockchain (Ethereum, plans to support multiple), BLE and NFC tags and Inlays Mobile (Android- and iOS-) Apps, Web Dashboard |
| Everledger | Diamonds | Counterfeit | Blockchain (Bitcoin and private (Eris)), no tags required |
| Provenance | Consumer Goods (Food, Whine, Clothing, Accessories) | Trust, Traceability, Transparency, and Story of Product | Blockchain (Ethereum), custom tags |

| | | | |
|---------------------|-----------------|--|--|
| Skuchain | Cargo | Trade and supply-chain finance | Blockchain (unknown) |
| Verisart | Art | Digital catalog, more efficient trading, digital history, authentication | Blockchain (unknown) Mobile (iOS-)App |
| VeChain | Consumer Goods | Trust, Counterfeit, Story of Product, Transparency, Traceability | Blockchain (unknown), Mobile (Android-)App |
| Sproxil Defender | Pharmaceuticals | Counterfeit | Database, Scratch Codes, Mobile & Web App, SMS |
| mPedigree | Pharmaceuticals | Counterfeit | Database, Hidden Code, SMS |

3. Proposed Mechanism

A blockchain-based project helps to make the KYC [11] trade between customers and other organizations much easier. This proposal gets its power from the property of Blockchain and Solidity programming, using the Ethereum network. For development purposes, Rinkeby test network (any test network can be used) has been used. Implementation of the concept of the public key and private key makes it more robust as the transaction is done by the user's digital signatures that also ensure the authenticity of the responsible user. Every signature causes a function trigger that creates a block of change and then sends it for mining. That function triggers cost money. Fake currency in the test environment and real capital in the real environment. Transaction is one of the authenticity parameters keeping track of user activities.

While writing and compiling the solidity code, the sol compiler (Solidity compiler) is needed to get all the functionalities to get started with. Working with this technology involves a series of steps in order to make our logic around the blockchain. Initially, consider the blockchain as a plain ground with no previous set of rules. Solidity will set protocols that will set the boundaries around the blockchain. The same type of code is used in other development areas like web development. The scope here simply justifies whether the function is allowed to be called from outside of the solidity file. Remix editor helps in scripting and testing the functionalities in a local environment.

Table 2. Overview of Blockchain Models

| Models | Access | Key Characteristics | Typical Use Cases |
|------------|---|------------------------|--|
| Public | Unrestricted | Immutable, Distributed | Crypto currencies, general-purpose |
| Consortium | Restricted to consortium members (read can be unrestricted) | Immutable, Distributed | Consortium specific use cases, e.g. trade between consortium members |
| Private | Restricted to a single entity (read can be unrestricted) | N/A | Internal auditing database management |

After scripting, the file is deployed to the blockchain in a particular test network (the network we are using for development). For deployment, the address or the instance of the gateway node of that particular network, 'Border node', is used. A local machine could get a direct connection with the border node but that would be so much of unnecessary work. Infura provides instant, scalable APIs for IPFS and Ethereum networks. So, with the help of Infura, the connection with that border node of a particular Ethereum network can be obtained.

Coming to the proposed system, after all the compilation and deployment, we have our blockchain ready to interact with. Now the whole system goes in a specific order, triggering other functionalities along the way as needed.

The system starts with the wallet creation by the user to make a register entry in the user mapping. Mapping uses the address to uniquely assign each user. This address functions as a public key, enabling anyone to send a document request to the user for approval, after which the document can be submitted to the user's wallet.

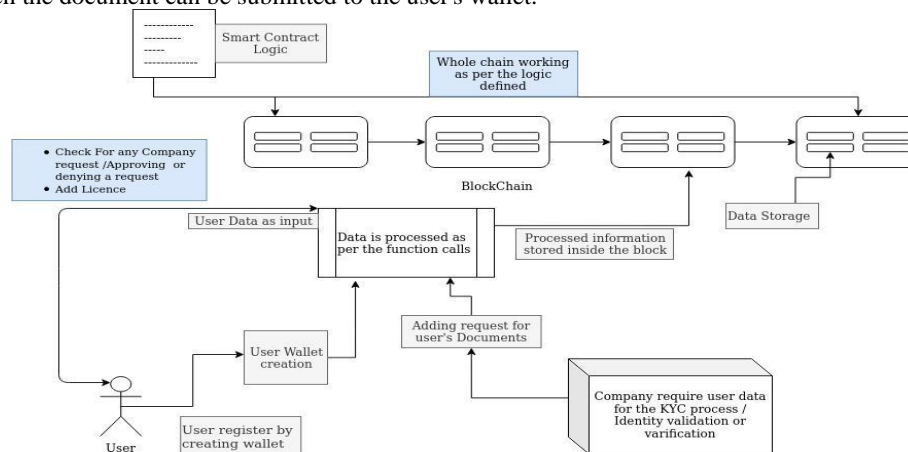


Figure 2. Architecture of the Proposed Work

Whenever the user tries to manipulate some data inside the blockchain, he/she must pay some amount of ether along with the transaction. The transaction needs some amount of GAS for executing the function, variables, and other algebraic operations. The calculation process of this gas is quite complex. MetaMask [12] takes care of it on its own and hence the transaction fee is calculated by the gas fee and total gas used in the processing of the function.

Meta mask gives the web3 [13] instance to work with and needs an Ethereum wallet provider. Web3 instance provides the user's address in the proposal to work. There is a special feature of solidity that it can get the user's address for calling or sending a transaction to the contract in the msg.sender variable. The address can be used to store necessary information in the mapping associated with that address. You can see the entire flow visualized in Figure 2.

All are addicted to the speed that other centralized applications provide but there is a different scenario when working with blockchain. There are two things to take note of before working with blockchain.

- Every operation related to the blockchain is asynchronous in nature. So, either we have to use the promise syntax or async/await syntax.
- We must wait for 15 to 30 seconds before the transaction is completed.

This wait is related to the theory of the blockchain working process. The process of nonce and mining that decides the difficulty factor for the mining machines for the mining process. The nonce in the blockchain is a 32-bit number. It resides in the block header and other data in the block, such as the timestamp and difficulty target. When miners start building the blocks, they choose a random nonce and put it inside the block header, resulting in a new block header hash.

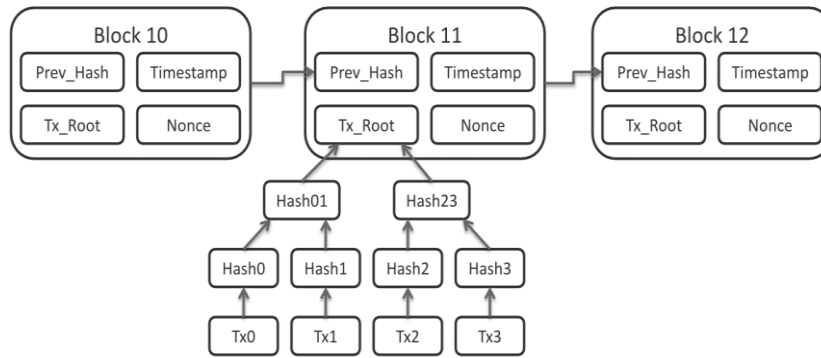


Figure 3. Insight of a Block

The hash is a 256-bit number and should have a large number of leading zeros, i.e., having an incredibly small value. If it does not have a sufficient amount of leading zeros, then the miner discards the hash and tries a new nonce. This process repeated, as shown in Figure 3, until they produce a hash value lesser than or equal to that threshold difficulty. There are different authorities given to different parties. Figure 4 explains the task performed by each actor in their functional domain.

Users authorized to:

- Create a wallet
- Add document
- View document requests by other organizations (functionality for user identification).

Organization can:

- Make a document request.
- View their pending requests
- View users document after user's approval

For the sake of a more complex structure, we can also include company wallet creation and add one more address mapping corresponding to each company that registers in the portal. But for now, using the msg.sender variable (store the address of the user who is calling the function), we can selectively render each company's request array to show them their request array. Figure 5 shows the ordered data flow from the user's system to the respective blockchain for processing via Metamask digital signature stage.

Here, any data manipulation will cost some amount of gas and hence cost money but data fetching is free of cost. When a user attempts wallet creation and documentation, he/she has to go through a series of steps and then receive a receipt of the transaction record.

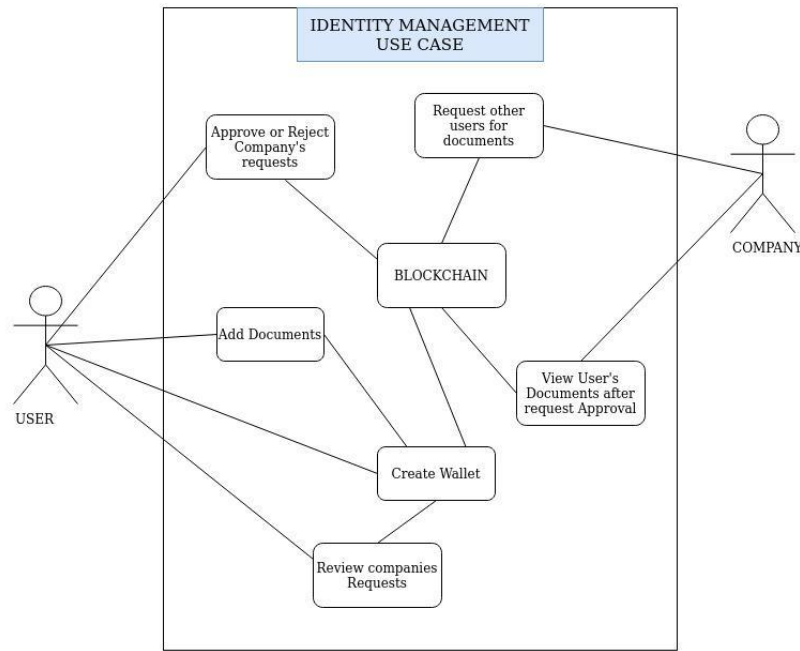


Figure 4. Actors task in Functional Domain

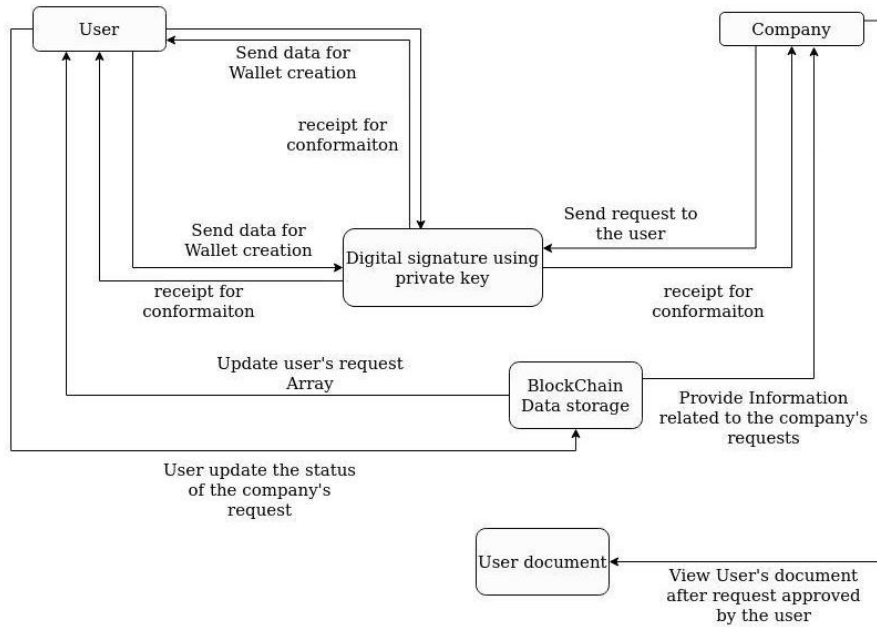


Figure 5. Metamask Digital Signature Processing

The documents stored [14, 15] in the user's wallet are addressed to an array mapped inside the blockchain. The organization requires the user's public account address (referred to as a public key) to reach out to the user and ask for a document request. The user's address to array request is updated with the request made by the company and it waits for the user's approval. The organization cannot view the user's documents because the approve request logic used by the user executes a logical expression that changes a value inside the request structure.

Figure 6 shows a similar set of instructions; the ordered flow of information and data in which the project is supposed to work. Each block shows functionality and then the if-else condition for the document request approval step. We can either leave the

request for future reapprove/rejection at a later time or can totally remove the request from the array which will cost additional gas. The functionality will change as per the project requirements changes in the later phases of development.

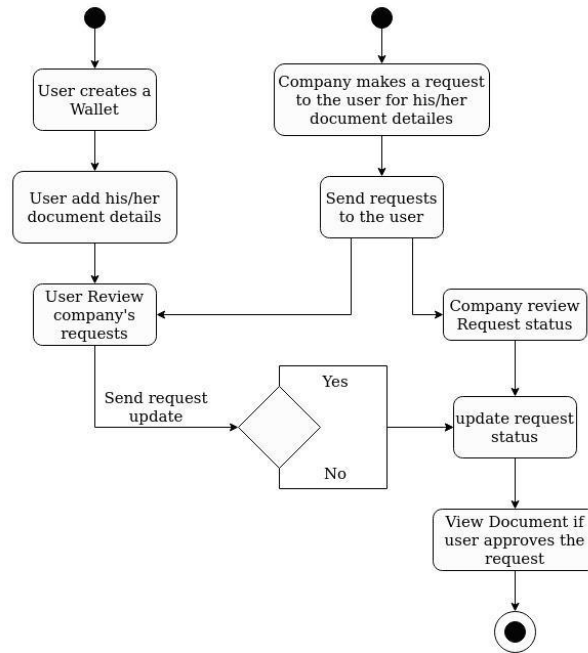


Figure 6. Activity Diagram

4. Experimental Setup and Results Analysis

The setup for the project is simple yet complex in its own terms. From installing and setting dependencies to implementing it to the project, it needs a sequence. There are some prerequisites for working with blockchain and solidity.

- 1) Prerequisites
 - a. Need good command over Solidity programming language.
 - b. Proper understanding about blockchain and transactions.
 - c. Use of Javascript for contract to frontend interaction.
 - d. Concept of compilation and deployment of smart contract.
- 2) Dependencies
 - a. Web3 module for making the instance to ease the process of blockchain interaction.
 - b. Ganache-cli module for testing of Solidity (Each transaction costs some money, so you do not want to waste your money.)
 - c. Mocha module for testing.
 - d. Truffle HD Wallet provider module as Web3 requires a provider that will supply account information for Web3 to work with.
 - e. Solc module for Solidity compiler to get the ABI interface and the byte code used in deployment.

Further steps require Solidity coding and testing using Remix Ethereum web editor. According to the system architecture, the whole system will go through the user's wallet creation and document addition. After the user successfully added the document, the account is ready for request processing. Different companies can send that user for its document with the help of the account address and that function will update the user and company mapping. That request in the array will be processed after the user interaction with that request. Regardless of its approval or rejection, the function triggers a certain logic that will allow the company to view the user's document.

5. Comparison with Existing Systems

Almost all the companies are going for cloud solutions with many problems even after the tremendous technological advancement. Currently, email and mobile verification are used to stop the creation of fake accounts and spam. The user receives an email or text message on their mobile device with a one-time password. The OTP is used to confirm the identity and account of a user. But there's no official way for an email to be associated with a specific individual. A single user may

possess multiple email addresses and mobile numbers, which could lead to the system's failure. Another problem in the system is how simple it is to create temporary emails [16]. At the end of the day, a good software solution is a technology that will make the user's life easier. If you look closely, when we try to make improvements in the ease of work, we have to compromise somewhere along the way. If we want to increase security (which is the topmost issue right now), we have to compromise with time and minimal user interaction. Security measures like multi-factor login or 2-step verification or 3-factor authentication can lead to user frustration. Moreover, if we pay more attention to user interaction and processes, then we might compromise somewhere in security.

Also, the centralization of information makes all resources vulnerable and increases the risk of attacks. We can make centralized storage safe against any natural disasters but once there is a data breach, we put all the users at a big risk.

There are many identity management solutions available. The best of them are compared in Table 3. Currently, almost all the systems are working on centralized cloud solutions. One possible reason is because there is not much work done in the field of decentralization. Also, when you implement blockchain in any project, it automatically acquires a lot of security features; for example, you do not need to worry about password management because of the property of account ID uniqueness. There can be an address mapping and if and only if the user has a working wallet in the mapping, he/she will be able to interact with the process.

A potential drawback with this solution is time. Every transaction that results in any manipulation of data inside the blockchain takes around 15 to 30 seconds. Those 15 to 30 seconds are the time taken by the miners to find that proper nonce and get the proper hash to set that piece of data in the chain. When compared with all the advantages we receive with the implementation of blockchain, one should not consider this as a drawback because that variable time is taken by the process actually makes the system more secure.

Table 3. Existing Identity Management Solutions

| Identity Management Solutions | Merits | Demerits |
|----------------------------------|--|--|
| Microsoft Azure Active Directory | <ul style="list-style-type: none"> Provides great integration with windows active directory. Proper connection with cloud service of Microsoft array. Provides identity protection based on Big Data and machine learning. | <ul style="list-style-type: none"> Poor integration with third party software directories and SaaS platforms. Provides advance reporting only in premium pricing tiers |
| OKTA Identity Management | <ul style="list-style-type: none"> Provides an amazing Mobile Device Management support and geographical zone. Has improved reporting functionality. Manages the flow of multiple identities. | <ul style="list-style-type: none"> Authentication requires expensive hardware. |
| EmpowerID | <ul style="list-style-type: none"> Increased flexibility in on-premises installation with many security benefits. Provide great reporting functionality with Active directory management. Added efficiency with workflow-based approvals | <ul style="list-style-type: none"> Workload management and setup cost greatly increase over cloud-based solutions. Mobile apps cannot be replaced with Mobile website |
| OneLogin | <ul style="list-style-type: none"> Provide great AD and HR solutions. Easy user and role management with mapping. On-premises applications comes with a great solution of proxy agents. Straightforward email notification configurations. | <ul style="list-style-type: none"> The highest price tier provides major functionalities and unsynchronized AD groups. |
| Optimal IdM | <ul style="list-style-type: none"> Very less knowledge is required for the highest level of technical configurations. Private cloud solutions ensure reliability, security, and performance. LDAP firewalls separate both identity and other applications | <ul style="list-style-type: none"> High pricing, therefore causes loss of a legitimate user base. No scope of SaaS provisioning configuration The very bounded ability for the users to tweak their SSO portal. |
| BITIUM | <ul style="list-style-type: none"> Amazing provisioning support. Google SSO leverage ability. Easy and fast features like password reset save money and time. Provide bookmarking ability in SaaS. | <ul style="list-style-type: none"> No customer support for identity management. Multiple identity support lags behind. |

6. Conclusion and Future Work

In the world of seven billion people, identity management becomes crucial and plays an important role when it comes to trust and mutual growth. There are many solutions regarding Identity and access management but the problem is we are still dependent on a centralized approach of data processing and storage. What we all need is to understand the importance and

power of decentralization. As a result, Blockchain is hands down the best technology to demonstrate the power of decentralization.

There is nothing called absolute victory, which means that you must pay the price when you are trying to gain something. It is an equal trade. Similarly, when it comes to the centralization of data storage, we obtain speed with respect to data fetching and manipulation and processing. However, we put our security at stake. The problem of password fatigue and distress of multiple factor login makes it hard for the user. When it gets difficult, companies cost a lot of money from the organization for management.

When you introduce Decentralization, it comes along with other implementations, like cryptographic techniques and other default security which do not need to be added separately as they are part of the process. No doubt, blockchain transaction takes time for block mining and processing, but that time it takes for the process of nonce mining increases security by adding the feature of dynamic difficulty variation. The best feature is that there is very little human interaction in the logical processing, hence less threat for malicious manipulation.

Theoretically, blockchain is unhackable, explained as 51% attack. Hackers must acquire at least 51% of the network to verify a malicious transaction which is not possible. Also, no one can change some data in between the chain as the property of chain works on the connection with the hash of the previous block. If you change the data in between, that whole hash chain will break and then the process will discard that chain and reset the blockchain after verifying with other network chains.

All these features show that there is a great future ahead for blockchain. Even after being in such an early stage, it has so many functions and after further development, it will achieve great heights.

References

- [1] Phipps, J., 6 Best Identity & Access Management (IAM) Solutions for 2023, <https://www.esecurityplanet.com/products/best-iam-software/>, accessed on January 1, 2024.
- [2] Jayabalan, J. and Jeyanthi, N. Scalable Blockchain Model using Off-Chain IPFS Storage for Healthcare Data Security and Privacy. *Journal of Parallel and Distributed Computing*, vol. 164, pp.152-167, 2022.
- [3] Jayapriya, J. and Jeyanthi, N. Distributed Consensus Protocols and Algorithms. In *Essential Enterprise Blockchain Concepts and Applications*, Auerbach Publications, CRC press, Taylor & Francis, pp. 15-38, 2021.
- [4] Jayabalan, J. and Jeyanthi, N. A Study on Distributed Consensus Protocols and Algorithms: The Backbone of Blockchain Networks. In *2021 International Conference on Computer Communication and Informatics (ICCCI)*, IEEE, pp. 1-10, 2021.
- [5] Jacobovitz, O. Blockchain for Identity Management. *The Lynne and William Frankel Center for Computer Science Department of Computer Science. Ben-Gurion University, Beer Sheva*, vol. 1, pp. 9, 2016.
- [6] Uhlmann, S. Reducing Counterfeit Products with Blockchains. *University of Zurich*, 2017.
- [7] Vruddhula, S. Application of on-Dose Identification and Blockchain to Prevent Drug Counterfeiting. *Pathogens and global health*, vol. 112, no. 4, pp. 161, 2018.
- [8] Faber, B., Michelet, G., Weidmann, N., Mukkamala, R.R., and Vatrappu, R. BPDIMS: A Blockchain-Based Personal Data and Identity Management System. In *The 52nd Hawaii International Conference on System Sciences*, pp. 6855-6864, 2019.
- [9] Tang, F., Ma, S., Xiang, Y., and Lin, C. An Efficient Authentication Scheme for Blockchain-Based Electronic Health Records. *IEEE access*, vol. 7, pp. 41678-41689, 2019.
- [10] Shamshad, S., Mahmood, K., Kumari, S., and Chen, C.M. A Secure Blockchain-Based E-Health Records Storage and Sharing Scheme. *Journal of Information Security and Applications*, vol. 55, pp. 102590, 2020.
- [11] Malhotra, D., Saini, P., and Singh, A.K. How Blockchain Can Automate KYC: Systematic Review. *Wireless Personal Communications*, vol. 122, no. 2, pp. 1987-2021, 2022.
- [12] Lee, W.M. using the MetaMask Crypto-Wallet. In *Beginning Ethereum Smart Contracts Programming: With Examples in Python, Solidity, and JavaScript*, Berkeley, CA: Apress, pp. 111-144, 2023.
- [13] Wang, Q., Li, R., Wang, Q., Chen, S., Ryan, M., and Hardjono, T. Exploring Web3 from the View of Blockchain. *ArXiv Preprint ArXiv:2206.08821*, 2022.
- [14] Mahlaba, J., Mishra, A.K., Puthal, D., and Sharma, P.K. Blockchain-Based Sensitive Document Storage to Mitigate Corruptions. *IEEE Transactions on Engineering Management*, 2022.
- [15] Das, M., Tao, X., Liu, Y., and Cheng, J.C. A Blockchain-Based Integrated Document Management Framework for Construction Applications. *Automation in Construction*, vol. 133, pp. 104001, 2022.
- [16] Jeyanthi, N., Chatterjee, S., Srujana, E.D., and Thandeeswaran, R. Novel Authentication Mechanism using Blockchain. *SPAST Abstracts*, vol. 1, no. 01, 2021.