Aadhaar Identity System using Blockchain Technology

C. Victoria Priscilla
Department of Computer Science,
SDNB Vaishnav College for Women,
University of Madras, Chennai, India

T. Devasena
Department of Computer Science,
SDNB Vaishnav College for Women,
University of Madras, Chennai, India

ABSTRACT
Aadhaar is a 12 digit unique identification number, provided to each citizen of India. This includes the biometric data and personal information such as full names, addresses and birthdates. As per the government rule nowadays Aadhaaris linked to the Bank account, PAN Card, Voter ID Card, LPG Connection Card, Ration Card, Mobile number of the citizen of India.

In order to ensure the security of Aadhaar, Blockchain technology has the potential to overcome the security and privacy challenges in Aadhaar.Blockchain is a decentralized, cryptographically signed digital ledger where transactions are framed into blocks.The type of information stored in the transactions can be varied (e.g. currency, intellectual property, identity data, location data, etc).Every transaction (and hence block) has a timestamp associated with when it was recorded to the blockchain. Subsequent blocks require the identifier (or hash) of the preceding block, and this is often the link that chains all the blocks together. The distributed consensus nature of the blockchain would prevent malicious attacks, until 51% of the nodes would be compromised. The proposed system AIMS (Aadhaar Identity Management System) provides self sovereign identity to the people through which one can create and control their verifiable credentials without any centralized authority and gives control over how their personal data is shared and used. This article also provides the comparative analysis of consensus algorithm which determines the performance and security of the blockchain system.

Keywords
Self Sovereign Identity, Consensus Algorithms, Blockchain and Hashing, IPFS (Inter Planetary File System)

1. INTRODUCTION
Aadhaar is a 12-digit unique identity number that can be obtained by residents of India, based on their biometric and demographic data (see Figure.1). The data is collected by a statutory authority, the Unique Identification Authority of India (UIDAI). Aadhaar is the world's largest biometric ID system. The detailed personal information being collected is of extremely high importance to a private. Major financial transactions are linked with information collected in Aadhaar. Data leaks are a gold mine for criminals. The UIDAI confirms quite 200 government websites were publicly displaying confidential Aadhaar data[1]; though removed now, the data leaked cannot be scrubbed from hackers' databases. Those confidential Aadhaar details are stored in a single database and it is maintained by UIDAI. Since the database is centralized there are so many disadvantages. The main disadvantage is that the security threat. If those data are stored in blockchain, data vulnerability are going to be reduced.

Figure1. Aadhaar containing biometric and Demographic data

Blockchains are cryptographically signed digital ledgers where transactions are framed into blocks. Each block is linked to previous block with its hash value which is generated cryptographically. When new blocks are added, it is reflected across all copies of the ledger in the network. At their initial level, they enable a group of users to record transactions in a ledger which is public to that group, such that no transaction can be changed once published.

Blockchain was introduced to the planet in 2008 during a whitepaper describing a replacement quite electronic currency: Bitcoin. The next significant innovation came in 2013 when a little startup named Ethereum put out a paper outlining how for developers to simply create entirely new blockchains without relying on bitcoin's original code. Two years later, Ethereum launched their new platform, allowing users to expand blockchain's functionality beyond Cryptocurrencies. Later, several other use cases for blockchains have emerged. A blockchain efficiently records transactions between parties on a distributed ledger. The data recorded on a blockchain is immutable and instantly verifiable [2]. A blockchain also can be programmed to automatically trigger transactions using smart contract technology.

Blockchain is a decentralized, cryptographically signed digital ledger where transactions are framed into blocks. The type of data stored within the transactions are often varied (e.g. currency, intellectual property, identity data, location data, etc).Every transaction (and hence block) has a timestamp associated with when it was recorded to the blockchain. Subsequent blocks require the identifier (or hash) of the preceding block, and this is the link that chains all the blocks together. The distributed consensus nature of the blockchain would prevent malicious attacks, until 51% of the nodes would be compromised [3].

The proposed Blockchain identity management system provides self sovereign identity through decentralized networks with no third party to share their personal identity information without their consent.

2. RELATED WORK
Identity Management system is a digital identity system[4] which enable the user to manage storage, authentication,

authorization, data sharing and protection of identities within the organization and on the web. The currently deployed system is susceptible to single point failure, lack of interoperability and privacy issues [5]. The blockchain based IDMSs ensures the consensus, transparency and integrity of the personal data sharing on distributed ledger technology (DLT)[6].

Since 2012, blockchain has been used in Estonia to maintain national data and services both in the public and private sector. Estonia maintains the multi-purpose digital ID card [7] on blockchain and makes sure that every change in data is immediately detected based on audit trails left by the "digital defense dust" that covers it. The city of Zug [8] in Switzerland is exploring a self-sovereign government issued identity on Ethereum using uPort, enabling access to a suite of e-government services in a convenient and secure manner. This eliminates the need of a user id/password to access government services. ID2020"[9] may be a global alliance across governments, public, private and non-government organizations to accelerate the method of assigning digital identity to those that are 'invisible' to the society. To address this goal, blockchain-based solutions and interoperability across multiple geographies and the reuse/ integration/connectivity for existing frameworks is being worked out. Provinces of British Columbia and Ontario in Canada using the Verifiable Organizations Network (VON) on the Sovrin blockchain [5]. ShoCard/SITA system, travelers scan their passports into a mobile app and take selfies, employing a private key to hash this verification data on the blockchain. [9]

3. PROPOSED MODEL

In this project we are getting to create Aadhaar Identity Management System using Blockchain Technology and also analyses the consensus algorithms that determines the safety and performance of the blockchain. The proposed model utilizes Ethereum blockchain to form the Aadhaar identity digital which enables the user to require control over their identity by creating a digital wallet (see Figure.2) where the identity documents are secured, verified and endorsed by Permissioned participants.

Figure 2. Creating Digital Wallet

3.1 Methodology

Step 1: Block chain development environment setup:

To create an Ethereum block chain, each entity has to set up a high computational power machine, known as an Ethereum node. GoEthereum or 'geth' is one such application available to various platforms using which one can create an Ethereum node.

• In our proposed system Ganache is used to create Ethereum blockchain. Ganache may be a tool from Truffle Suite that permits developers to make their own private Ethereum blockchain to check dApps. Deploying a dApp directly on Ethereum can cost a lot of gas to verify transactions. On the other hand, Ganache allows us to do testing without paying any gas. It also enables us to manage the mining speed and gas costs within the test environment to test different scenarios for smart contracts.

• Truffle is a framework for Ethereum that offers a development environment for building Ethereum based apps. It includes support for the library that gives custom deployments for coding new contracts and links Ethereum applications.

• Installation of Node.js which converts our machine into a node. Node is a system in the blockchain that is used to verify the authenticity of block as well as maintain the digital ledger that keeps record of the all the blocks in the chronological order.

Step2: Smart Contract: Smart contracts are lines of code that are stored on a blockchain which can be executed automatically when the predefined conditions are satisfied [9]. The biggest feature of Ethereum that that it links the smart contract and the Blockchain. Smart contract executes in the Ethereum Block chain's decentralized platform that is once the contract is deployed on the chain, no individual node is able to change it. Therefore the projects based on the Ethereum and smart contracts are decentralized and credible. A smart contract can be written in a language known as Solidity. There are online compilers for Solidity, using which one can write a Solidity [10] Smart Contract, and publish that contract onto a particular Ethereum block chain network. In this proposed system Remix Ethereum IDE is used for solidity programming. Remix is a powerful, open source tool that helps us to write Solidity contracts straight from the browser. Remix also supports testing, debugging and deploying of smart contracts and far more. The smart contract is hosted on Ropsten test network.

Step3: Signing the transaction: In this proposed system, the transactions are signed using Metamask. Metamask is a wallet that acts as a bridge between Ethereum blockchain and Chrome or Firefox by working as a browser extension. It also can interact with different Ethereum test networks to form it a perfect wallet for developers.

Step 4: Front end Ethereum API: After signing the transaction, web3.js API will send the information to the smart contract, if there is a trigger function for these information then changes will take place, and the changes will be updated in the public ledger, such that transparency is maintained across the network.

3.2 Understanding the work of AIMS:

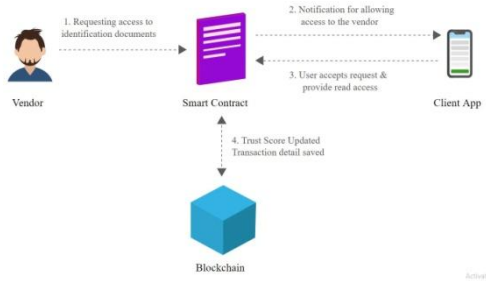

Figure 3.Understanding the working of AIMS

3.2.1 Algorithm

1. Enter the details such as user name, contact number, Email ID to create a wallet for the user.

2. Hash will be created for the transaction that created the wallet.

3. Upload Aadhaar identity details and the scanned copy of Aadhaar that will be saved in the IPFS with hashed addresses stored in the Blockchain.

4. Smart contract helps in signing a particular transaction and generates hash value which will be stored in each block in the block chain. If any slight changes in the block, the hash value of that block changes drastically.

5. Third party companies requesting to access specific details of a person for authentication purposes, a notification will be sent to the individuals owning the identity.

6. Once the user allows the companies to access their details, third-parties can use the identifiable information for authenticating a person. Blockchain does not store the user's data or information. Instead, the transactions made between identity holders and companies will only be recorded on the blockchain. The Aadhaar details are stored on IPFS in an encrypted form.

3.2.1 Workflow Diagram

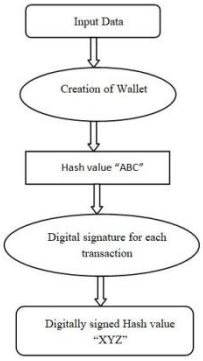

Figure 4.Creation of Genesis Block

The genesis block created (see Figure.4), contains the uploaded user identification with a private key for which a public key has been generated. The public key of each node is send to all the nodes in the blockchain system. Each block is created and secured by Keccak-256 algorithm. Refer to the Bitcoin system using the Blockchain system [3] [11], the ECDSA (Elliptic Curve Digital Signature Algorithm) method is used in digital signature techniques, the small key size in this method can support the desired security.

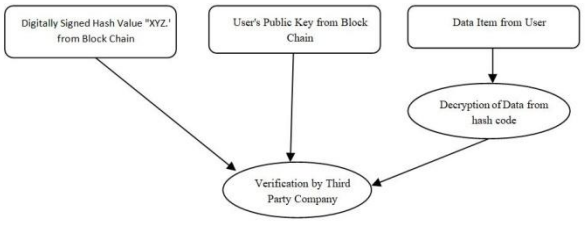

Figure 5. Verification by Third Party Companies

The verification process (see Figure.5) begins with the acquisition of block containing the identity information, previous hash of the hash value originating from the previously valid block and the digital signature. Then separated between electronic documents (the digital identity and previous hash) and digital signature. The electronic document is calculated its hash value. These two hash values are then compared, if the value is the same then the digital signature is valid and the process continues, but if the value is not equal it is considered invalid and the system will refuse the block to continue the process. After the digital signature verified and proven to be valid, further verification of the previous hash begins with the capture of the voting result, and the previous hash contained in the most recent in database, and searched hash values with the Keccak-256 algorithm. Then compare it with the previous hash carried by the block being done verification. If the value is the same, then the hash value is valid and the whole block is verified as a valid block and sent by the node contained in the system, but if the value is not the same considered invalid and the system will reject the block.

Through this Blockchain Identity Management system user have entire control over their Aadhaar Identity and they provided with a Unique ID in which their Identity information are encrypted and stored on IPFS. Users can share unique IDs with any third-party to authenticate themselves directly through blockchain identity management. Moreover, this system uses smart contracts to enable controlled data disclosure. Thus, data manipulation is not possible on the blockchain; hence it enhances the level of security and privacy. Using the Blockchain identity management backed by IPFS doesn't allow any hacker to steal the identifiable information, since the system will be decentralized. [6]Blockchain identity management doesn't set any geographical boundaries. So, users can use the platform across the borders to verify their identity.

3.3 Comparative analysis of Consensus Algorithms

Blockchain may be a distributed decentralized network that gives immutability, privacy, security, and transparency. There's no central authority present to validate and verify the transactions, yet every transaction within the Blockchain is taken into account to be completely secured and verified. this is often possible only due to the presence of the consensus protocol which may be a core part of any Blockchain network. [12] A consensus algorithm is that the core component that directly dictates how such a system behaves and therefore the performance it is able to do. So, it's important to look at the prevailing consensus algorithm and its cryptographic hash functions.

Secure Hash Algorithm (SHA): Secure Hash Algorithm (SHA) may be a group of hash functions published by the National Institute of Standards and Technology as a US Federal information science Standard (FIPS). All of the present SHA algorithms were developed by the NSA [13].

• SHA-1: NIST (1995) developed the Secure Hash Algorithm 1 or SHA-1 and generates a 160-bit message digest for an arbitrary length input message. However, a collision attack was also founded against SHA-1, NIST announced the step-by-step elimination of SHA-1[14].

• SHA-2: SHA-2 variants are SHA-224, SHA-256, SHA-384, SHA-512, SHA-512/224, and SHA512/256. SHA-2 includes a big number of changes from its predecessor, SHA-1. SHA-256 and SHA-512 are novel hash functions computed with 32-bit and 64-bit words, respectively. However, SHA-256 has no multithreading ability, and thus it's not fast enough for transactions [13].

• SHA-3: After several successful collision attacks which were progressively reduced in complexity (such as SHA-1 and SHA-2), NIST, within the Federal Register, announced a

public competition to develop SHA-3, a totally new hashing algorithm. On October 2nd, 2012, the winner of the competition Keccak, was announced [13]. In 2014, NIST considered SHA-3 as a typical hash function.

3.3.1 Comparison of Hashing Algorithms

In following table the comparison between the various cryptographic hash algorithms like SHA-1, SHA-2, and SHA-3 shows with various parameters.

Table1. Comparison of Hash Algorithms

Properties	Name of Algorithms		
	SHA-1	SHA-2	SHA-3
Block size	512 bits	512/1024 bits	1088/576 bits
Word size	32 bits	32/64 bits	320/320 bits
Output size	160 bits	256/512 bits	1600/1600 bits
Rounds	80	64/80	24/24
Operations	ADD, XOR, OR, AND, NOT, ROTATE	ADD, XOR, OR, AND SHIFT, ROTATE	-
Constructions	Merkle-Damgard	Merkle-Damgard	Sponge

Ethereum, in a consensus engine called Ethash uses Keccak-256. Keccak may be a family of hash functions that eventually got standardized to SHA-3 (SHA256 is a component of a family of hash functions called SHA-2). Ethereum called it Keccak rather than SHA-3 because it has slightly different parameters than the present SHA-3. SHA-256 (more specifically SHA-2 i.e. SHA-256(SHA-256(TXN)) is employed to encrypt blocks of blockchain for bitcoin cryptocurrency. Ethash is employed to encrypt blocks of blockchain for Ethereum cryptocurrency. Ethereum uses Keccak-256 rather than the NIST standardized SHA-3 hash function, solidity 0.4.3 has introduced keccak256. (It is an alias to sha3, meaning that keccak256 produces identical results to sha3, but with the intent to avoid confusion, especially for developers new Ethereum.) It's recommended that the new code use keccak256 rather than sha3.

Keccak is predicated on a completely unique approach called sponge construction. [14] Sponge construction is predicated on a good random function or random permutation, and allows inputting ("absorbing" in sponge terminology) any amount of knowledge, and outputting ("squeezing") any amount of knowledge while acting as a pseudorandom function with reference to all previous inputs. This results in great flexibility.

3.3.2 Result and discussion

Keccak256 may be a very secure and powerful algorithm, and almost like Bitcoin's SHA256. The strength of Ethereum Classic is that the payment system, and therefore the smart contract execution exist within the same layer. No side-chains, no trusted third-parties, no merge-mining. This enables developers possess access to proof-of-work based, programmable, sound money [15]. Ethereum classic should follow the lead of Bitcoin and adopt a CPU-hard algorithm, which it'll be the most important coin thereon algorithm. Keccak256 is that the same algorithm that smart contracts on Ethereum Classic currently have access to and it might allow smart contracts to verify the Proof of work of the blockchain they're running on. As mentioned in [16], Keccak256 is such an honest algorithm because it's very efficient and can have a highly liquid mining market. Hence, SHA-256 is employed by Satoshi Nakomoto who invented the primary Bitcoin blockchain. In fact, the sole reason Satoshi didn't use Keccak256 was that it had been only SHA certified as safe in 2015, almost 6 years after the Bitcoin blockchain began, and around when Ethereum launched[15].

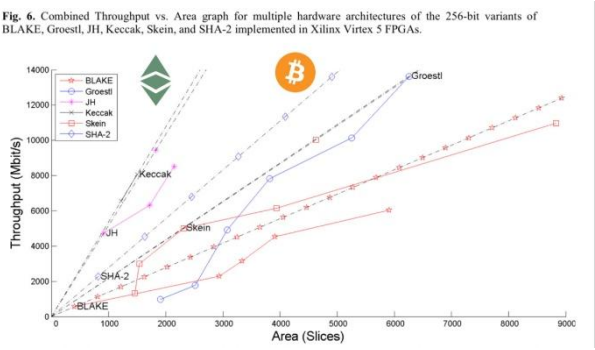

Figure6. FPGA implementations of hashing algorithms[17]. (Notice SHA-2 and Keccak)

4. ETHEREUM BLOCK IDENTIFIER

In the proposed system, Ethereum permission less Blockchain with a non-hierarchical network of computers (nodes) that build and come to a consensus on an ever-growing series of "blocks", or batches of transactions, known as the blockchain. Each block contains its own identifier that it must immediately follow in the chain if it is to be considered valid. Whenever a transaction occurs, there will be changes in the ETH balances and other storage values of Ethereum accounts.

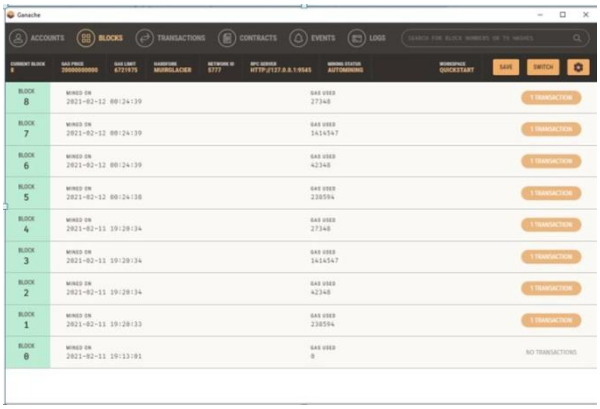

Figure 7.Ganache personal blockchain has been created in the proposed model

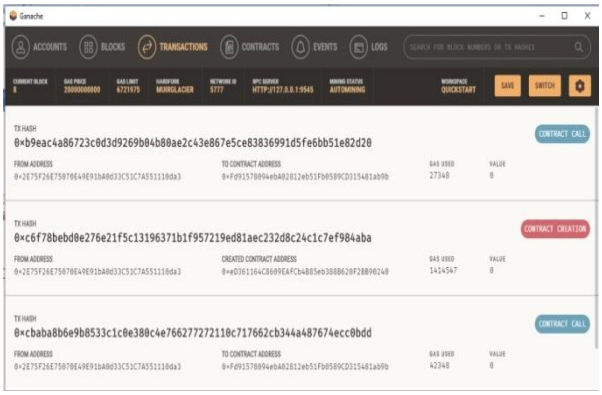

Figure 8.Transactions in each block

Figure 9.Ethereum Block Identifiers

5. CONCLUSION

Through this proposed system the self sovereign identity is achieved by providing a identity management system using block chain technology. The user can create and control their credential without rely on any central authority. Successful implementation of the Blockchain identity management can enhance the level of security and privacy. The immutable and decentralized ledger allows third parties to validate the user's data without wasting time and money.

The Popular technique blockchain faces few challenges. We list few critical challenges as follows.

Scalability: As the number of transactions in our day to day life is increasing, the size of blockchain is becoming bigger. Every node in network will store all transactions and verify them with blockchain as they have to verify whether source of current transaction is spent or not. Bitcoin blockchain approximately process 7 transactions per second due to block size restriction and time interval in generating new block, which is not enough to fulfils the requirement in real-time as it has to process millions of transactions. The blocks capacity is very small, so this may result in delay between minor transactions because the miners prefer these transactions with higher transaction fee.

Human error: When we send information into the database it needs to be of high quality and we use blockchain as a database. It is untrustable to store data in blockchain because of this events are recorded and monitored accurately.

Unavoidable security flaw: In bitcoin and other blockchains we can find a security flaw, the lie will become the truth when more number of computers are working as nodes to service the network. This is known as '51% attack' and was highlighted by Satoshi Nakamoto when bitcoin is launched by him. So, the community closely monitors bitcoin mining pools, ensures no one will gain such network influence unknowingly.

Complexity: More Complex Security algorithms are used for signing and verification. And also Hashing algorithm is used to securely share the Documents.

However, Distributed ledger technology is not a silver bullet solution for identity management. Future work in this research area faces the above discussed hurdles. Although there are still some challenges in the emerging of blockchain technology, Continuous technical innovation and awareness can significantly help in bringing down the risk and help us move towards a safer world.

6. REFERENCES

[1] S. Venkatasubramanian, V. Swarnakamali, J. Kaviya, and A. Vigneshwar, "Aadhaar security through blockchain," pp. 19–21, 2019.

[2] J. S. Notland, "Blockchain enabled Trust & Transparency in supply chains (Journal format)," *Medium*, 2017, [Online]. Available: https://medium.com/@jrgensvenneviknotland/blockchain-enabled-trust-transparency-in-supply-chains-journal-format-2744fa4f37d.

[3] S. P.J and G. George, "Blockchain Based Aadhaar Security," *Int. J. Eng. Technol.*, vol. 7, no. 4.6, p. 398, 2018, doi: 10.14419/ijet.v7i4.6.28450.

[4] R. Nechushtai, M. Elit, and S. M. Systems, "Blockchain Identity Management System Based on Public Dentities Ledger," *Google Patents*, vol. 1, no. 12, 2017, [Online]. Available: https://patents.google.com/patent/US9635000B1/en.

[5] L. Lesavre, P. Varin, P. Mell, M. Davidson, and J. Shook, "A Taxonomic Approach to Understanding Emerging Blockchain Identity Management Systems," 2019, doi: 10.6028/NIST.CSWP.07092019-draft.

[6] P. Dunphy and F. A. P. Petitcolas, "A first look at identity management schemes on the blockchain," *IEEE Secur. Priv.*, vol. 16, no. 4, pp. 20–29, 2018, doi: 10.1109/MSP.2018.3111247.

[7] X. Zhu and Y. Badr, "Identity Management Systems for the Internet of Things: A Survey Towards Blockchain Solutions," *Sensors (Basel).*, vol. 18, no. 12, pp. 1–18, 2018, doi: 10.3390/s18124215.

[8] "IMPACT OF BLOCKCHAIN ON DIGITAL IDENTITY BUILDING TRUST IN THE CYBERWORLD." Accessed: Sep. 19, 2020. [Online]. Available: https://joom.ag/PgYa.

[9] Y. Liu, Z. Zhao, G. Guo, X. Wang, Z. Tan, and S. Wang, "c," *Proc. - 2017 15th Annu. Conf. Privacy, Secur. Trust. PST 2017*, pp. 44–53, 2018, doi: 10.1109/PST.2017.00016.

[10] K. Mudliar and H. Parekh, "A comprehensive integration of national identity with blockchain technology," *Proc. - 2018 Int. Conf. Commun. Inf. Comput. Technol. ICCICT 2018*, vol. 2018-Janua, pp. 1–6, 2018, doi: 10.1109/ICCICT.2018.8325891.

[11] R. B. Venkatapur, B. Prabhu, A. Navya, R. Roopini, and S. N. A. S, "Electronic Voting Machine Based On Blockchain Technology and Aadhar Verification," *Int. J. Innov. Eng. Sci.*, vol. 3, no. 3, pp. 12–15, 2018.

[12] M. S. Ferdous, M. J. M. Chowdhury, M. A. Hoque, and A. Colman, "Blockchain Consensus Algorithms: A Survey," pp. 1–39, 2020, [Online]. Available: http://arxiv.org/abs/2001.07091.

[13] A. Maetouq, S. M. Daud, N. A. Ahmad, N. Maarop, N. N. A. Sjarif, and H. Abas, "Comparison of hash function algorithms against attacks: A review," *Int. J. Adv. Comput. Sci. Appl.*, vol. 9, no. 8, pp. 98–103, 2018, doi: 10.14569/ijacsa.2018.090813.

[14] P. P. Pittalia, "A Comparative Study of Hash Algorithms in Cryptography," *Int. J. Comput. Sci. Mob. Comput.*, vol. 8, no. 6, pp. 147–152, 2019, [Online]. Available: www.ijcsmc.com.

[15] "ECIP-1049: Why Ethereum Classic should Adopt Keccak256 for its Proof of Work Algorithm. | by Alexander Tsankov | Medium."

https://antsankov.medium.com/ecip-1049-why-ethereum-classic-should-adopt-keccak256-for-its-proof-of-work-algorithm-e45aee32d8a9 (accessed Feb. 17, 2021).

[16] E. C. Hunt, S. B. Sproat, R. R. Kitzmiller, E. C. Hunt, S. B. Sproat, and R. R. Kitzmiller, "Implementation Overview," *Nurs. Informatics Implement. Guid.*, pp. 1–19, 2004, doi: 10.1007/978-1-4757-4343-2_1.

[17] Z. A. Al-Odat, M. Ali, A. Abbas, and S. U. Khan, "Secure Hash Algorithms and the Corresponding FPGA Optimization Techniques," *ACM Comput. Surv.*, vol. 53, no. 5, 2020, doi: 10.1145/3311724.

[18] N. Popper, "Ethereum, a Virtual Currency, Enables Transactions That Rival Bitcoin's," *New York Times*, Mar. 2016, Accessed: Feb. 17, 2021. [Online]. Available: https://www.nytimes.com/2016/03/28/business/dealbook/ethereum-a-virtual-currency-enables-transactions-that-rival-bitcoins.html.

[19] J. E. de A. Sousa, "An analysis of the fees and pending time correlation in Ethereum," *Int. J. Netw. Manag.*, no. e2113, 2020, Accessed: Feb. 17, 2021. [Online]. Available: http://dl.ifip.org/db/conf/lanoms/lanoms2019/196411_1.pdf.