JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue

JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)

An International Scholarly Open Access, Peer-reviewed, Refereed Journal

THE EVOLUTION OF OVERWRITING HASHING TECHNIQUE IN BLOCKCHAIN

[1]Dr. Vijaya Kumar A V, [2]Vignesh K R

[1]Associate Professor, [2]Assistant manager,

[1]Computer Science and Engineering,[2]JSW-IT Corporate

[1]Proudhadevaraya Institute of Technology, Hospet, India, [2]JSW-IT & Digital Solutions, Tornagallu, India

Abstract : Blockchain is a ledger that stores the history of transactions in a distributed manner so that multiple entities can store the data at multiple locations. This results in each user having a copy of the entire history of transactions. If a change is made to one entity, it must be reflected in the entire network. Every update to the database is a batch of transactions grouped into blocks. Every block is an update and a chain of blocks is history. Each block is built off or chained to the previous block, using hash functions on the previous block. This makes the Blockchain tamper evident and also reduces the strain on the network. Index Terms— Bl

IndexTerms **- Blockchain, Linked List, Merkle Tree.**

I. INTRODUCTION

A blockchain is an immutable, decentralized, distributed ledger that cannot be censored. Fundamentally, a blockchain is an unalterable database that no individual person or entity controls. It spreads out across multiple points of operation, allowing anyone to anonymously interact with it and add to it without a main authority controlling or stopping the interaction. A blockchain does not refer to the computers or machines that are involved in the blockchain. It refers to the ledger itself, in simple terms, a giant file that is identical and unalterable across all machines that store it. The file contains the entire history of exchanges among the blockchain users. This file is referred to as the blockchain. There are already thousands of blockchains out there with more coming into existence everyday, and all these blockchains are tailored to specific use. One may be for music streaming, one for file storage, one for finance, one for asset management, and so on. In the future, these blockchains will be able to communicate with one another similar to how websites interact with each other today. However, to say that you are storing some money on the blockchain and uploading a song to the blockchain and signing a land contract on the blockchain will probably be a misnomer, as you will likely be interacting with many different blockchains without even knowing it. One critical aspect of blockchain technology is how the participants agree that a transaction is valid. This agreement is called reaching consensus, and there are many ways for doing this.

A blockchain is just one part of a solution. Blockchain implementations are usually designed with a specific purpose or function. Example purposes include cryptocurrencies, smart contracts (software programs deployed on the blockchain and executed by computers running on that blockchain network), and distributed ledger systems between businesses. There has been constant developments in the field of blockchain technology, with new platforms being announced every now and then – the landscape is continuously changing. There are two high-level categories for blockchain outlooks that have been identified: permissionless, and permissioned[3]. In a permissionless blockchain network anyone can read and write to the blockchain without any authorization. Permissioned blockchain networks restrict participation to specific people or organizations and allow fine-grained controls. Despite the many adaptations of blockchain networks and the rapid development of new blockchain related technologies, most blockchain networks use common core concepts.

Blockchain is a distributed record consisting of blocks where each block is made up of a block header containing metadata about the block, and block data containing a set of transactions and other related data. Every block header (except for the first block of the blockchain) contains a cryptographic link to the previous block's header. Each transaction involves one or more blockchain network user and a registry of what happened, and it is digitally signed by the user who submitted the transaction. Blockchain technology takes existing, established concepts and merges them together into a single solution. Blockchains are tamper-evident and tamper resistant digital ledgers implemented in a distributed fashion (i.e., without a central repository) and usually without a central authority (i.e., a bank, company or government). At their fundamental level, they enable a circle of users to record transactions in a shared ledger within that community, such that under normal operation of the blockchain network no transaction can be changed once published. In 2008, the blockchain idea was combined with several other technologiesand concepts to create modern cryptocurrencies[18] electronic cash protected through cryptographic mechanisms instead of a central repository or

authority. This technology became widely popular in 2009 with the launch of the Bitcoin[3], the first of many modern cryptocurrencies.

In Bitcoin, and identical systems, the transfer of digital information that represents electronic cash takes place in a distributed system. Bitcoin users can digitally sign and transfer their rights to that information to another user and the Bitcoin blockchain records this transfer publicly, allowing all participants of the network to independently verify the validity of the transactions. The Bitcoin blockchain is independently maintained and managed by a distributed group of participants. This, along with cryptographic mechanisms, makes the blockchain resistant to attempts to manipulate the ledger later (modifying blocks or forging transactions). Blockchain technology has empowered the development of many cryptocurrency systems such as Bitcoin and Ethereum. Because of this, blockchain technology is often viewed as bound to Bitcoin or possibly cryptocurrency solutions in general. However, the technology is available for a broader variety of applications and is being investigated for a variety of sectors. The numerous components of blockchain technology along with its reliance on cryptographic primitives and distributed systems can make it challenging to understand[7]. However, each component can be described simply and used as a building block to understand the larger complex system. Since blockchains are cryptographically signed transactions that are grouped into blocks, linked to the previous blocks after validation and a consensus mechanism, as new blocks are added, older blocks become more difficult to modify (creating tamper resistance). New blocks are replicated across copies of the ledger within the network, and any conflicts are resolved automatically using established rules.

II. BLOCKCHAIN ARCHITECTURE

Blockchain is a sequence of blocks, which holds a complete list of transaction records like conventional public ledger. Figure 1 illustrates an example of a blockchain, where each block contains some data in the body of the block and is cryptographically linked to the previous block using hashing. The first block of a blockchain is called genesis[2] block which has no parent

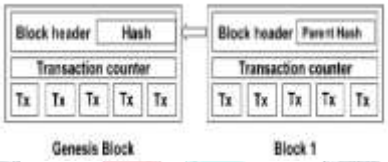


Fig 1. Blockchain architecture

2.1 Structure of a Block

A block consists of the block header and the block body as shown in Fig 2.

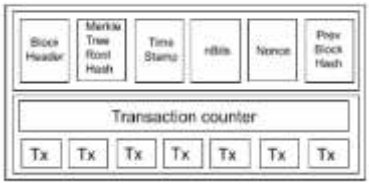


Fig 2. Block Structure

In particular, the block header includes:
(i) Block Version: Set of block validation rules to follow.
(ii) Merkle Tree Root Hash: Hash value of all the transactions in the block.
(iii) Timestamp: Current time as seconds in universal time since January 1, 1970.
(iv) nBits: Target threshold of a valid block hash.
(v) Nonce: A 4-byte field, which usually starts with 0 and increases for every hash calculation.
(vi) Parent Block Hash: a 256-bit hash value that points to the previous block. Which is derived from SHA256 algorithm[4].

The block body is composed of the transactions and a transaction counter. The maximum number of transactions that a block can contain depends on the block size and the size of each transaction. Blockchain uses an asymmetric cryptographic mechanism to validate the authentication of transactions. Digital signature based on asymmetric cryptography is used in an untrustworthy environment. We next briefly illustrate digital signature.

2.2 Digital Signature

Each user owns a pair of private[6] key and public key. The private key that shall be kept in confidentiality is used to sign the transactions. The digitally signed transactions are broadcast throughout the network. The typical digital signature involves two phases: signing phase and verification phase. For instance, a user A wants to send another user B a message.

(1) In the signing phase, A encrypts her data with her private key and sends B the encrypted result and original data.

(2) In the verification phase, B validates the value with A's public key. In that way, B could easily check if the data has been tampered or not. The typical digital signature algorithm used in blockchains is the Elliptic Curve Digital Signature Algorithm (ECDSA)[5].

III. TYPES OF BLOCKCHAIN

Depending on the need of application, Blockchain can be divided into 3 types: A. Public blockchain: All exchanges that occur on open or public blockchains are non-restrictive, implying that anybody can look at the exchange subtleties. Open blockchains are intended to be completely decentralized, with nobody, individual or substance controlling,in which exchanges are recorded in the blockchain or the request wherein they are prepared. B. Private blockchain: Another kind of chains are private blockchains, otherwise called permissioned blockchains. They have various remarkable contrasts from open blockchains. Members have to agree to join the systems. Exchanges are private and are only accessible to environment members that have been offered authorization to join the system. Private blockchains are more concentrated than open blockchains. C. Consortium blockchain: Consortium blockchains are considered a different assignment from private blockchains. The primary contrast between them is that consortium blockchains are represented by a gathering instead of a solitary substance. This methodology has no different advantages of a private blockchain and could be viewed as a subclass of private blockchains, rather than a different kind of blockchain.

IV. MERKLE HASH TREE

A Merkle tree is a hash-based data structure that is a speculation of the hash list. It is a tree structure in which each leaf node is a hash of a data block, and each non-leaf node is a hash of its descending node. In Regular Merkle trees each node has up to 2 child nodes. Merkle trees are used in distributed frameworks for productive data confirmation[17]. They are efficient in light of the fact that they use hashes rather than full documents. Hashes are methods for encoding documents that are a lot smaller than the real documents itself. Presently, their fundamental uses are in shared systems, such as Tor, Bitcoin, etc.

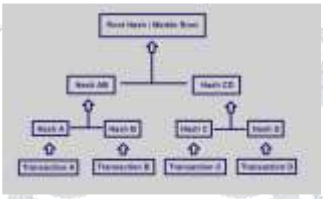


Fig 3. Merkle tree.

V. CRYPTOGRAPHIC HASHING

Cryptographic hash capacity is a hash function that is appropriate for use in cryptography. It is a numerical calculation that maps information of a subjective size (regularly called the "message") to a bit string of a fixed size (the "hash value"or "hash") and is a single direction function, that is practically infeasible to reverse. In a perfect world, the best way to discover a message that delivers a given hash is to endeavor a brute force search of potential contributions to check whether they produce a match, or utilize a table of coordinated hashes. Cryptographic hash capacities are a fundamental tool of present day cryptography. Hashing capacities are scientific calculations that take inputs and give interesting yields. The usual hashing algorithms are MD5, SHA-3, SHA-256 and SHA-512. Secure Hashing Algorithm-256 (SHA-256) is what is utilized by Bitcoin. A hash function takes a string of any length and produces a fixed length string which acts as a form of "signature. In this way, a person knowing the "hash value" is not able to realize the unique message, however the person that knows the authentic message can prove the "hash value" is produced from that message[2].

Fig 4. Effect of avalanche effect.

5.1.The perfect hash function

The perfect hash function has three principle properties:

a. It is very easy to calculate hash value for any given data.

b. It is computationally hard to find an alphanumeric content that has a given hash value.

c. It is incredibly impossible that two marginally different messages will have a similar hash value.

5.2. Avalanche Effect

The Avalanche effect is a property of block ciphers and cryptographic hash algorithms. It is frequently used in cryptography. The effect says that when the input changes even just a bit, the impact of it on the output is huge, making it indistinguishable. In high quality block ciphers this implies: A little change in the key or the plaintext should cause a solid change in the ciphertext. The term Avalanche effect was first used by Horst Feistel(1973). Later on, the idea was recognized by Shannon's diffusion. On the off chance, if a block cipher or cryptographic hash[15] capacity does not display the avalanche effect to a critical degree then it is said to have poor randomization. This way a cryptanalyst can make an estimation about the information being given by the output.

This might be adequate to break the algorithm totally or partially. Along these lines, the avalanche effect is a very useful property from the perspective of the originator of the cryptographic algorithm[3].

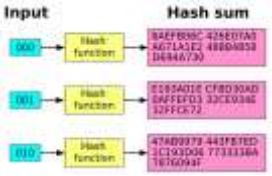


Fig 5. Effect of avalanche effect.

5.3 Hash Table

A hash table is a structure for putting away data. In software engineering, these structures for monitoring data or information are called data structures. A hash table is a data structure that uses a hash function to monitor where the information is put. Each snippet of data to be put away has a name, which is known as a key. For instance, a key may be an individual's name. Each name is coordinated up to one bit of information called a value, similar to the individual's phone number. A decent Hash Table will consistently discover data at a similar speed, regardless of how much information is placed in.A great deal of Hash Tables additionally let the client put key/value matches (a name and its information) in and take them out at a similar speed.

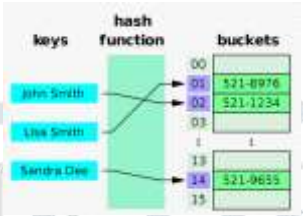


Fig 6. Mechanism of hash Table.

The information is kept in another data structure called a table, which resembles many boxes, like a repository to hold information. Each box has a number beginning from 0 and up. Hash Tables can frequently discover data quicker than different data structures, for example, search trees or other table query structures. Therefore, they are utilized in numerous sorts of PC programming. They are utilized most for related clusters, databases, reserves and sets. The thought behind a hash table is to make sense of which box to put information by utilizing just its name[16]. This implies, regardless of what number of boxes are exceeded, you can generally discover data rapidly in the event that you have its name. The hash table uses a hash function to make sense of which number to place information in from its name. The hash function peruses a name and gives back a number[4].

5.4 Comparison of various Hashing algorithm

The following tables evaluate standard and technical facts for some of cryptographic hash functions. A review of hash function security/cryptanalysis can be located at hash function security and accuracy. You may haveseen an issue natural with hashing. Since they produce a fixed-length value, there are a limited number of hashes for each sort of calculation. This makes crashes easily conceivable. An impact is when two unique blocks of information produce precisely the same hash. It's amazingly uncommon for this to occur, yet they have been accounted for. Subsequently, some more specialist hashing functions have been considered dishonorable to be utilized for secure applications. Normally, the more expanded the hash value, the more far out a crash will occur. For example, a function that makes a 256-bit hash (like SHA) will have less crashes than one that creates a 128-bit hash (like MD5) in light of the fact that there are progressively possible hash values when you have more bits.

Table 1. Differences of Each SHA algorithm.

Algorithm	Message Length (*bit*)	Block Size (in *bits*)	Word Size (in *bits*)	The Size of the Message Digest (*bit*)
SHA 1	$<2^{64}$	512	32	160
SHA 256	$<2^{64}$	512	32	256
SHA 384	$<2^{128}$	1024	64	384
SHA 512	$<2^{128}$	1024	64	512

In the above table we have revealed the performance of the diverse hashing algorithms particularly SHA-256 and SHA-512. SHA-256 and SHA-512 are message digests, they had never been intended to be password-hashing (or key-derivation) capabilities. (Although a message digest may be used as a constructing block for a KDF, inclusive of PBKDF2 with HMAC-SHA1.) SHA-512 can be appreciably quicker while calculated on maximum 64-bit processors as SHA-256 uses 32-bit math, an operation which is consistently slower. The only real advantage that SHA-512 may have over SHA-256 is collision resistance, a term that in cryptography has a very restricted meaning. SHA-256 claims 128-bit collision resistance, SHA-512 claims 256-bit. If or whilst a realistic quantum pc is constructed, we might want the 256-bit collision resistance. SHA-256 outputs are shorter, which saves bandwidth. Different hardware favors special functions. SHA-512 is normally faster on 64-bit processors, SHA-256 is faster on 32- bit processors. SHA-512/256 sits right in among the two features - the output size and safety stage of SHA-256 with the performance of SHA-512[15]. However, almost no structures use it till date.

5.4.1 Comparison of SHA256 and SHA512

SHA-256 and SHA-512 operate in the way of MD4, MD5 and SHA-1 in which the message to be hashed is first as follows:. Padded with its length in one of these ways that the result is a more than one of 512 bits and 1024 bits long, after which the message blocks are processed one at a time from beginning with a constant preliminary hash value that sequentially computes. ie., Parsed into 512-bits for SHA-256 and 1024-bits for SHA-512 message blocks *M1,M2…….Mn,*

the message blocks are processed one at a time from beginning with constant preliminary hash value H that (0) sequentially computes.

$$H\ .\ (i) = H(i-1) + C$$

$$m\ (i)H(i-1)$$

Where C is the SHA-256 and SHA-512 compression function and + means word-wise mod 2 and addition for SHA-256 and 32 264 SHA-512. H is the hash value or message (N) digest of M.

VI. DIFFERENT TYPES OF LEDGER

The below diagram shows the different types of ledgers that are widely used.

(a). Decentralized (b).Centralized (c).Distributed

A. Decentralized systems are trustless environments that don't have any single authority. The distinction right here is there are layers of nodes. The blue nodes, or end nodes, hook up with secondary nodes (marked in yellow). The secondary nodes connect to one another. It should be mentioned that there was a high degree of variability in how these systems are built, but the statement stays faithful to the image above.

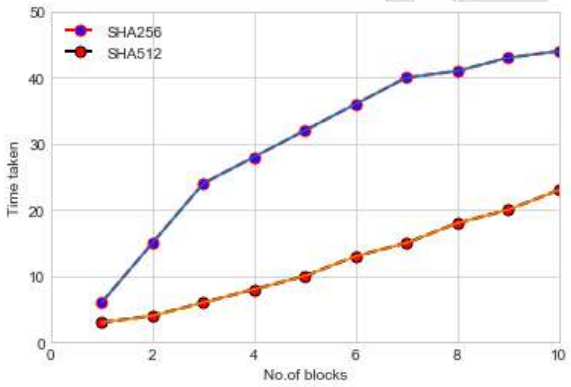
B. Centralized systems have a single authority[15]. All nodes obey the commands of the dot coloured yellow inside the diagram. For

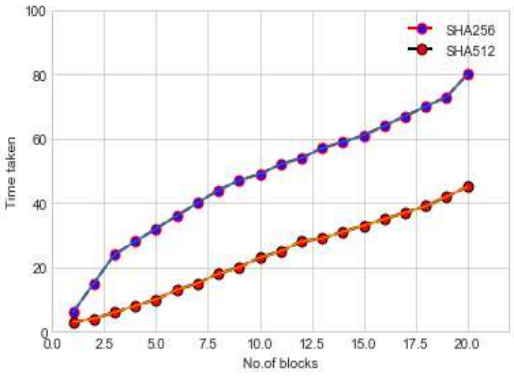
example, a King or Emperor commanding their state.

C. Distributed structures don't have any single authority. Each of the dots within the diagram is a node, or a community participant. Each node is attached directly to each different node. For instance, we can imagine a vote where every voter has the same strength to one another.

VII. EXPERIMENTAL RESULT

In this work, we have located each of the evaluation of SHA256 and SHA512 variations respectively. Simulations are shown below in graphical evaluation. In observation, overall performance of secure hashing algorithms SHA256 and SHA512 differ in time usage ie., SHA512 consumes less time than that of SHA256 in various quantities of blocks creating blockchain. Hence, the graph is obtained by plotting the quantity of blocks in x-axis and time taken for blocks technology along y-axis as shown respectively.





VIII. CONCLUSION

With the substitution of the older hashing algorithm with the new version, the overall performance has increased immensely which ensures the efficient usage of blockchain. Blockchain has established its potential for revised conventional industry with its key developments: decentralization, persistency, anonymity and auditability. With the addition of this new hashing algorithm into the block chain, we increase the performance and also the security.

REFERENCES

[1] M. Padmavathi and R. M. Suresh, "Secure P2P intelligent network transaction using Litecoin," *Mobile Networks and Applications*, vol. 24, no. 2, pp. 318–326, 2018.
[2] G. Liang, S. R. Weller, F. Luo, J. Zhao, and Z. Dong, "Distributed blockchain-based data protection framework for modern power systems against cyber attacks," *IEEE Transactions on Smart Grid*, vol. 10, no. 3, pp. 162–173, 2019.
[3] A. S. Konoplev, A. G. Busygin, and D. P. Zegzhda, "A blockchain decentralized public key infrastructure model," *Automatic Control and Computer Sciences*, vol. 52, no. 8, pp. 1017–1021, 2018.
[4] B. Seok, J. Park, and J. H. Park, "A lightweight hash-based blockchain architecture for industrial IoT," Appl. Sci., vol. 9, no. 18, 2019, doi: 10.3390/app9183740.
[5] M. Raikwar, D. Gligoroski, and K. Kralevska, "SoK of Used Cryptography in Blockchain," IEEE Access, vol. 7, pp. 148550–148575, 2019, doi: 10.1109/ACCESS.2019.2946983.
[6] N. Tinu, "A Survey on Blockchain Technology- Taxonomy, Consensus Algorithms and Applications", International Journal of Computer Sciences and Engineering, vol. 6, no. 5, pp. 691-696, 2018. Available: 10.26438/ijcse/v6i5.691696.
[7] S. Yu, K. Lv, Z. Shao, Y. Guo, J. Zou and B. Zhang, "A High Performance Block-chain Platform for Intelligent Devices," 2018 1st IEEE International Conference on Hot Information-Centric Networking (HotICN), Shenzhen, 2018, pp. 260-261, doi: 10.1109/HOTICN.2018.8606017.
[8] S. Debnath, A. Chattopadhyay and S. Dutta, "Brief review on journey of secured hash algorithms," 2017 4th International Conference on Opto-Electronics and Applied Optics (Optronix), Kolkata, 2017, pp. 1-5, doi:10.1109/OPTRONIX.2017.8349971.
[9] Wenting Li, Alessandro Sforzin, Sergey Fedorov, and Ghassan O. Karame," Towards Scalable and Private Industrial Blockchains", In Proceedings of the ACM Workshop on Blockchain, Cryptocurrencies and Contracts (BCC '17), New York, 2017, pp. 9-14, doi: https://doi.org/10.1145/3055518.3055531
[10] V. Lemieux, "Evaluating the Use of Blockchain in Land Transactions: An Archival Science Perspective", European Property Law Journal, vol. 6, no. 3, 2017. Available: 10.1515/eplj-2017-0019 [Accessed 29 June 2020].
[11] Chris Veness, "SHA-256 Cryptographic Hash Algorithm implemented in JavaScript | Movable Type Scripts", Movable-type.co.uk, 2020. [Online]. Available: https://www.movable-type.co.uk/scripts/sha256.html. [Accessed: 29- Jun- 2020].
[12] Iwar.org.uk, 2020. [Online]. Available: http://www.iwar.org.uk/comsec/resources/cipher/sha256-384-512.pdf. [Accessed: 29- Jun- 2020]. .
[13] Kombe, Cleverence, Manyilizu, Majuto, Mvuma,"Design of Land Administration and Title Registration ", Journal of Information Engineering and Applications, vol. 7, no. 1, pp. 8-15, 2017. Available: https://www.iiste.org/Journals/index.php/JIEA/article/view/35154/0
[14]https://www.researchgate.net/publication/352551870_ABCMETAapp_R_Shiny_Application_for_Simulation-based_Estimation_of_Mean_and_Standard_Deviation_for_Meta-analysis_via_Approximate_Bayesian_Computation_ABC/stats
[15] Kwon, D., Reddy, R. R. S., & Reis, I. M. (2021). ABCMETAapp: R shiny application for simulation based estimation of mean and standard deviation for meta-analysis via approximate Bayesian computation. Research synthesis methods, 12(6), 842-848.
[16] Sarmiento, J. M., Gogineni, A., Bernstein, J. N., Lee, C., Lineen, E. B., Pust, G. D., & Byers, P. M. (2020). Alcohol/illicit substance use in fatal motorcycle crashes. Journal of surgical research, 256, 243-250.
[17] Lu, N., Butler, C. C., Gogineni, A., Sarmiento, J. M., Lineen, E. B., Yeh, D. D., ... & Byers, P. M. (2020). Redefining Preventable Death—Potentially Survivable Motorcycle Scene Fatalities as a New Frontier. Journal of surgical research, 256, 70-75.
[18] Singh, P., Williams, K., Jonnalagadda, R., Gogineni, A., & Reddy, R. R. S. (2022). International Students: What's Missing and What Matters. Open Journal of Social Sciences, 10(2), 381-397.