A Review on Application of Hash Functions and Digital signatures in the Blockchain Industry

Article · September 2021

CITATION
1

READS
576

1 author:

Shashie Dilhara

3 PUBLICATIONS **10** CITATIONS

SEE PROFILE

A Review on Application of Hash Functions and Digital signatures in the Blockchain Industry

B.A.S. Dilhara[1]

[1]Department of Network & Security, NSBM Green University, Sri Lanka

In a world witnessing paradigm shift towards decentralized applications, blockchain has become a promising solution where it provides consensus and trust among participating peers from diverse authoritative domains. As a result, the trust among the peers and security in the network of such environments has become a critical factor. Hence, it is the cryptography and the cryptographic concepts that provide reliability and security to the underlying fundaments of blockchain technology. Even though many papers focus on industries that blockchain is being used widely, only few studies have inspected the underlying cryptographic techniques. Thus, the primary objective of this paper is to provide an insight to the application of hash functions and digital signatures in blockchain technology, by reviewing the academic literature and reports in this domain. In conclusion, this review recognizes and compares different digital signatures and hash functions used in various blockchain platforms.

***Index Terms*—blockchain, cryptography, digital signatures, hash functions**

I. Introduction

Blockchain is an emerging technology that ensures data integrity by consisting of a immutable ledger in a distributed network [1]. Simply, it can be defined as a distributed ledger managed by a peer to peer network, collectively adhering to a consensus protocol, which lacks a central authority [2]. Blockchain being one of the crypto-intensive creatures, has affected both industrial and academic domains and it has been currently established in many such areas like government, banking and finance, supply chain and e-commerce [3].

The concept of "Distributed Ledger", first came into being in the paper "New Directions in Cryptography" in 1976. With the advancement of cryptography, Stuart Haber and W. Scott Stornetta have envisioned the concept of blockchain, in 1991 [5]. Apart from that, David Chaum introduced the concept of electronic cash, which has always been one of the prominent research topics in the field of cryptography. This system had one major drawback of using patented algorithms, which was considered a major obstacle to acceptance among the cryptographical community [4]. Furthermore, Adam Back in 1997 introduced a concept called "hashcash" which lead for concept of creating money called "b-money", on peer to peer network by Wei Dai [4]. Later, a community of programmers and cryptographers called "Cypherpunks" proposed patent free cryptographic concepts and were based on the paper presented by Adam Back, Wei Dai's "b-money" and Nick Szabo's "Bitgold" proposal [2]. However, it was the paper "Bitcoin: A Peer-to-Peer Electronic Cash System", on the basis of above concepts, published by Satoshi Nakamato in 2008, made the blockchain concept more popular due to the success of the Bitcoin and is considered the inventor of the blockchain technology [5]. The decentralization, double spending resistance, pseudonymity and unforgeability of Bitcoin has generated a new revolution in this domain making it the main framework for many other cryptocurrencies [4].

Based on the implementation design, administration rules, data availability, and access privileges blockchain can be divided into 4 major categories namely permission less public, permissioned public, permission less private and permissioned public [2]. Bitcoin, Ripple, LTO and Hyperledger fabric are some platforms that belongs to these categories respectively [2]. Apart from that a blockchain is composed of 3 main technologies namely cryptography, peer to peer networking and game theory [3]. Cryptography uses public key and hash functions to ensure integrity, transparency and privacy while peer to peer networking is used as a client and a server for storing replicas and for game theory all nodes participating in the system must comply with the rules of consensus and be motivated by economic incentives [3]. However, this paper will only be focused on the cryptographic technology used in the 4 major categories and to provide a perception to the cryptographic applications underlying the blockchain technology by reviewing the previous works and concepts.

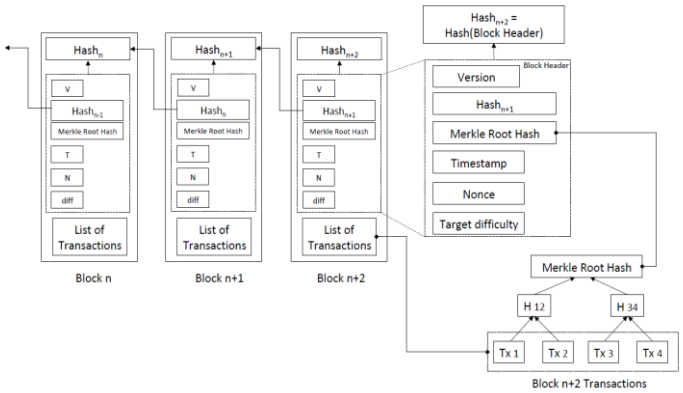

Fig. 1. Basic Architecture of a Blockchain [2]

II. CRYPTOGRAPHIC CONCEPTS IN BLOCKCHAIN

A. Cryptographic Hash functions

Hash functions are among the cryptographic primitives which typically doesn't encrypt or decrypt messages and can be used to ensure the data integrity [6]. Simply, hash functions are capable of mapping arbitrary sized inputs to a fixed size output [2] and the resultant output is called hash or digest [6]. Hash functions are generated using the concept of Trapdoor One-Way Functions (TOWF) and hash function is a unidirectional function [6. When it is applied to a message m of variable size, where the message belongs to a certain set of messages, M; a fixed, predetermined resultant output of bit size n can be obtained [7]. The number of possible hashes is much smaller than the number of different input messages because, a hash function converts a message of any length to a collection of n bits [7]. Thus, a hash function is described as follows.

$$h: M \rightarrow \{0,1\}^n, with\ h(m) = \hat{m} \quad (1)$$

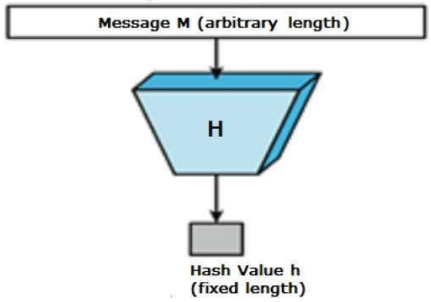

Fig. 2. Image of a typical Hash Function

One-wayness and collision resistance are two security requirements of hash functions and the former ensures, the underlying function is not invertible, while latter implies that it is impossible to find two inputs having the same hash value [4].The usage of hash functions in a block chain can be classified into 6 main categories namely proof of work (POW), address generation, message digest in signatures (MDS), pseudorandom number generation (PNG) and bridge components (known as Fiat-Shamir mechanism) [4]. The last four stages were quite popular even before the origin of blockchain, however, it was the uprise of cryptocurrencies and blockchain that made POW and address generation categories popular [4]. POW protocol is an expensive computer computational involving hashing, merkel tree and peer to peer networking for broadcasting, creating and validating a block in the blockchain network [1].

1) Comparison of Cryptographic hash functions used in different blockchain platforms

SHA-2 are the most popular hash functions which are being used in the blockchain industry and it is one of the algorithms from a family of cryptographic hash functions named Secure Hash Algorithms (SHA) [2]. United States National Security Agency (NSA) and States Federal Information Processing Standard originally designed SHA for verifying message integrity, data identification and password verification [8]. Hash functions of the SHA-2 family is the successor of SHA-1 function. SHA-224, SHA-256, SHA-384 and SHA-512 includes the specifications of SHA-2 which provides hashes of 224,256,384 and 512 respectively [6]. SHA-1, SHA-2 and SHA-256 are the most adopted algorithms that are associated with the blockchain technology, because of their ability of creating unique outputs when different inputs are given [8]. Furthermore, the latest member of the SHA family is SHA-3 which was released by National Institute of Standards and Technology (NIST) in 2012 as the winning function of their Cryptographic Hash Algorithm Competition [8]. This was initially known as Keccak which became the kernel of SHA-3 family and later its' 256 and 512 versions were adopted as the hash function of Ethereum blockchain [8].

A typical application of blockchain is cryptocurrency. Bitcoin being the most popular and first established crypto currency, uses a SHA-256 hash function, along with the other cryptocurrencies like Counterparty, Namecoin and Bitcoin cash [1]. Furthermore, there are many alternative hash functions which are being used by different cryptocurrencies. Apart from the cryptocurrencies, SHA-256 is also being used in Hyper Leger Fabric, which is a modular blockchain framework that acts as a foundation for blockchain-based product development, solutions, and applications using plug-and-play components that use within private enterprises [9].

SCrypt is one such hash function which is being used in Bit connect, Bitcoin gold and Litecoin. SCrypt is also considered as a memory-hard hash function and it requires more memory than SHA-256 [4]. However, the hash rate shown by SCrypt is 620MHash/s, which is at a lower level when compared with the SHA-256 function [1]. Here the hash rate is a metric for comparing the mining power of a computer and the processing power of a blockchain such as Bitcoin. The mining difficulty is adjusted according to the amount of hash rate that is available to the network, which means if the hash rate is high, the mining difficulty too increase and vice versa [9].

In addition, Etherum based coins use Ethash function uses Keccak function as mentioned above. It is also known as an Application specific integrated circuit (ASIC) resistant hash function and the highest average hash rate recorded is 140THash/s [4].

Moreover, X11 hash function proposed by Duffield is another memory hard hash function which is used in Darkcoin, having an average hash rate of 32.5GHash/s [4]. Furthermore, blockchain designs such as IOTA has their own hash function called "curl-p", which was negatively reviewed by the crypto community [2].

2) Role of hash functions in blockchain and its security

Hash functions play a major role in blockchain security. If an attacker wants to crack a 256-bit length private key, then it must exhaust 2256 key possibilities and if a typical super computer performing 1018 keys per second is used for hacking such system, it will take 3x1051 years to find the key [3]. Even in the worst-case scenario, where an attacker is equipped with an extremely powerful computer and find the

key within a day; blockchain network can withstand the attack by linking blocks together using a cryptographic hash function [3].

Consider a block A that has a hash value called HA. When a new block B is added to the blockchain network, the miners calculate the hash of the newly added block, which is called HB. This process is the solution to the POW problem. The link between the block A and B is created by using HA for the generation of HB. The hash value of block B can be computed by, "HB = hash (HA + info_block_B + nonce) < target" , where info_block_B is transaction information in block B, nonce is the value to look for to solve the POW problem, and target is the threshold to find the nonce value to satisfy the current difficulty [3]. Due to this linking, if an attacker tries to alter a transaction of block A, then he has to solve the POW problem at block A and find one or more new blocks with a valid hash value to replace block B. This is called a double spending attack, where the probability of this attack gets arbitrarily small as the blockchain network grows [3]. Since the hashes of 2 blocks are connected, any change in the blockchain information will result in a different hash string and will affect all involved blocks [8]. Due to this method of hash creation, the data integrity of blockchain is secured.

For a hash function to be secure it should satisfy some properties like Preimage resistance, Second preimage resistance and collision resistance [10]. Preimage resistance defines that for any output it is infeasible to find any input that hashes to that output, Second preimage resistance means that it is computationally infeasible to find any second input that has the same output as a given input and collision resistance means that it is hard to find two inputs that hash to the same output [10]. Brute force attacks on SHA-256 are being launched by attackers in correspond to these three properties. Most modern attacks on hash functions are focused on finding collisions and the most popular generic attack for finding collisions is birthday attack [11]. In addition, Keccak 256 in Ethereum, which are also known as sponge functions are subjected to attacks called inner collisions [9].

B. Digital Signature Algorithms

Digital signature is yet another ineludible cryptographic primitive in blockchains beside the hash functions. A Digital signature is an authentication mechanism that allows the creator of a message to attach a secret code along with the message which act as a signature and it guarantees the source and integrity of the message [12]. Furthermore, absence of a digital signature can create disputes among different parties. It was Diffie and Hellman that put forward the concept of digital signature when they opened the gate of public key cryptography in 1976 and today it is being used widely in different areas ensuring authentication, integrity and non-repudiation. A digital signature is composed using two types of algorithms namely Hash Function Algorithms and Digital Signature Algorithms (DSA). Since hash functions were discussed in detail above, this section will be focused on DSA only. Signature schemes are used in almost every blockchain and is used in blockchain to sign the transaction, by, authenticating the intended sender and providing transaction integrity as well as non-repudiation of the sender. The fig 3. given below explain the functioning process of Digital Signature [12].

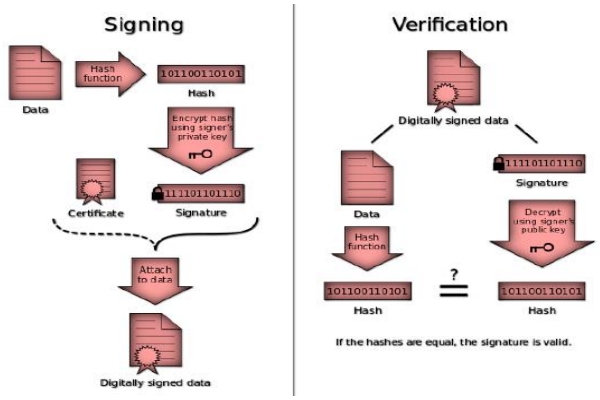

Fig. 3. Digital Signature Process [12]

1) Comparison of Digital Signatures in Different blockchain platforms

Blockchain technology applies different signature schemes in order to provide privacy, anonymity and unlikability as additional features. The most widely used signature schemes used in blockchain are based on Elliptic curves. Elliptical Curve Digital Signature Algorithm (ECDSA) and Edwards-curve Digital Signature (EdDSA) are 2 prominent digital signature schemes which are frequently being used in blockchains.

In principle, both ECDSA and EdDSA are based on the elliptic curve version of discrete logarithm problem [4].ECDSA was first introduced by Neals Koblitz and Victor Miller in 1991 from independent works and it is a modification of the Digital Signature Algorithm (DSA) [9]. Moreover, it requires smaller numbers in comparison to DSA and even with smaller numbers it is safer as DSA or Rivest–Shamir–Adleman (RSA) algorithm [9].

Blockchain platforms like Bitcoin, Ethereum and Hyperledger use ECDSA and it is used over a general elliptic curve, however, EdDSA works over a (twisted) Edward curve and now is used in Naivecoin and Monero [4]. Edward curve is a plane model of an elliptic curve and provides a better security and efficiency than a general elliptic curve and has been already chosen as the next elliptic curve generation of the Transport Layer Security (TLS) by research professionals [4]. Some advanced signature primitives such as ring signature, multi-signature, blind signature and threshold signature are also widely applied in blockchains to enhance the privacy and anonymity of transactions.

Ring signature uses a protocol where a signature is created on a message by any member of a group, on behalf of the group while preserving the identity of the individual signer of the signature [2]. Anonymity of the signing party in a blockchain can be achieved using ring signatures. Even though it provides privacy protection towards individual signing behavior, it could be abused for some illegal purpose such as wash trading [4]. Cryptocoin and Ringcoin are some applications of Ring Signature scheme.

Threshold signature, allows n parties receive a share of the secret key to create the signature and t out of n parties, create a

signature over any message [2]. The key is never disclosed in the entire scheme as the parties directly generate the signature from the shares [2]. Threshold signature too provides anonymity. CoinParty which is a Secure Multi-Party Mixing of Bitcoins, uses threshold signature scheme [2].

Multi-Signature permits a single signature to function as several ordinary signatures on the same message and the size of both the single signature must have the same size as one regular signature [4]. It might be advantageous to use a multi-signature scheme, when a transaction requires a signature from a group of participants in a blockchain [2]. Openchain and MultiChain are some blockchain platforms that use Multi-Signature scheme.

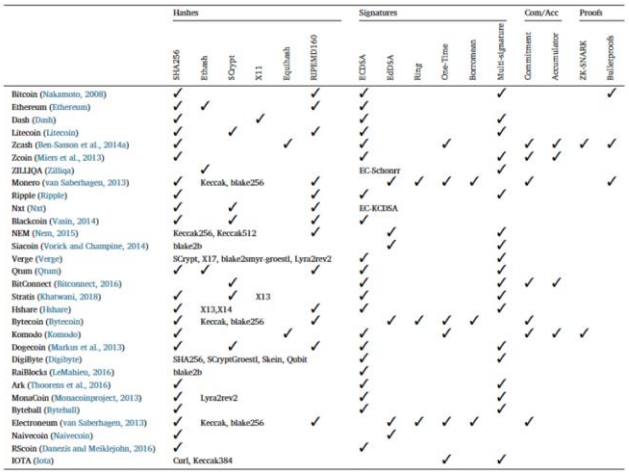

Fig. 4. Summary of different cryptographic algorithms in Blockchain [4]

2) Achieving Data Authentication, Data Integrity and Non-Repudiation through Digital Signatures

Blockchains use digital signatures to achieve data authentication, data integrity and Non-repudiation. Data authentication in a digital signature is achieved in the following way. Before an upcoming message is used, a verifier validates the Digital signature attached with the actual message using Public-Key of a sender and assure that signature has been created just by sender who has the corresponding secret Private-Key and no one else like adversary [12].

If an attacker alters the data before it reaches the destined receiver, the verification will fail since the hash of the altered data and the output provided by the verification algorithm will not match. Therefore, the receiver can safely deny the message with strong hypothesis that data integrity has violated. This will help in protecting integrity.

Since, only the signer of the message has enough information about the signature keys, only the signer can create unique signature on a given message [12]. As a result, the receiver can present data and the digital signature attached along with that message to a third party as evidence if any dispute occurs in the future. So, it helps in achieving non-repudiation.

Apart from that, digital signatures have some pros like better customer experience, business efficiency, cost saving, legal validity, global acceptance, independent verification, security, workflow efficiency and key management and time consumption are some of its' cons.

III. Conclusion

Despite the extensive studies and staggering progress, that the cryptography has achieved in the blockchain technology in the recent years; yet it remains a very challenging domain. This paper focused on offering a systematic review on, how the application of hash functions and digital signatures have affected the blockchain technology. Based on the reviewed concepts and associated properties, it was a prominent factor that hash functions and digital signatures play a major role in ensuring the security of a blockchain. It has also provided an insight to different hash functions and DSA algorithms which are being used in different blockchain platforms by reviewing them. In conclusion, this compressive study has been a vivid eyeopener for the usage of cryptography in blockchain technology.

References

[1] 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.

[2] 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.

[3] K. T. Son, N. T. Thang, L. P. Do, and T. M. Dong, "Application of Blockchain Technology to Guarantee the Integrity and Transparency of Documents," *Int. J. Comput. Sci. Netw. Secur.*, vol. 18, no. 12, pp. 7–15, 2018.

[4] L. Wang, X. Shen, J. Li, J. Shao, and Y. Yang, "Cryptographic primitives in blockchains," *J. Netw. Comput. Appl.*, vol. 127, no. November 2018, pp. 43–58, 2019, doi: 10.1016/j.jnca.2018.11.003.

[5] S. S. Sarmah, "Understanding Blockchain Technology," *Comput. Sci. Eng.*, vol. 8, no. 2, pp. 23–29, 2018, doi: 10.5923/j.computer.20180802.02.

[6] V. G. Martínez, L. Hernández-Álvarez, and L. H. Encinas, "Analysis of the cryptographic tools for blockchain and bitcoin," *Mathematics*, vol. 8, no. 1, 2020, doi: 10.3390/math8010131.

[7] H. Gilbert and H. Handschuh, "Security analysis of SHA-256 and sisters," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 3006, pp. 175–193, 2004, doi: 10.1007/978-3-540-24654-1_13.

[8] L. A. Ajao, J. Agajo, E. A. Adedokun, and L. Karngong, "Crypto Hash Algorithm-Based Blockchain Technology for Managing Decentralized Ledger Database in Oil and Gas Industry," J, vol. 2, no. 3, pp. 300–325, 2019, doi: 10.3390/j2030021.

[9] T. F. I. N. D. E. Grado, "TRABAJO FIN DE GRADO A Comprehensive Survey on Blockchain ' s Technology," 2019.

[10] G. Wang and S. Wang, "Preimage attack on hash function RIPEMD," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 5451 LNCS, no. 112, pp. 274–284, 2009, doi: 10.1007/978-3-642-00843-6_24.

[11] A. K. Sharma, S. K. Mittal, and S. Mittal, "Attacks on Cryptographic Hash Functions and Advances," vol. 5, no. November, pp. 89–96, 2018.

[12] K. P. S, S. Kapoor, K. S. Oza, and R. K. Kamat, "A Comprehensive Study on Sentiment Analysis Using Deep Forest International Journal of Computer Sciences and Engineering Open Access A Comprehensive Study on Sentiment Analysis Using Deep Forest," no. August 2018, 2019, doi: 10.26438/ijcse/v7i4.654658.