Application of Hash Algorithm on Blockchain System

Jofiandy Leonata Pratama, 13517135
Program Studi Teknik Informatika
Sekolah Teknik Elektro dan Informatika
Institut Teknologi Bandung, Jl. Ganesha 10 Bandung 40132, Indonesia
13517135@std.stei.itb.ac.id

Abstract – In this modern era, trading transaction has been a regular activity in the society through online and offline. With the advancement of technology, this helps changing the way people do transaction. However, many stakes are putted through this transaction. In order to reduce the risks, the trust between parties are put on centralized bodies such as banks or government. Having this third party, the privacy of each individual transactions has been reduced and thus also causing high vulnerability of data breach. With the assistance of modern technology, this privacy issue can be solved with the existence of Blockchain. This paper will discuss regarding the hash function in Blockchain system.

Keywords – Hash functions, blockchain

I. INTRODUCTION

Traditional transaction implements a centralized system where it involves third party in order to reduce risk of being fraud and also to verify the transaction. In this context, the third party can be the bank or the government where they have the rights to access and also to store their citizens' transaction history. However, this situation brings a new problem towards the citizens itself. Having stored every individual's data, it increases the chances of the data to be breached and hacked thus, causing the privacy of each individual to be lost. Besides this issue, some other problems regarding the bank mechanism have also been the drawbacks for the people using this traditional transaction such as a high interest for loans and fees for using the bank facilities. Having agitated with these problems, people start to find new alternatives to revolutionize this traditional transaction. This lead to a decentralized or distributed system.

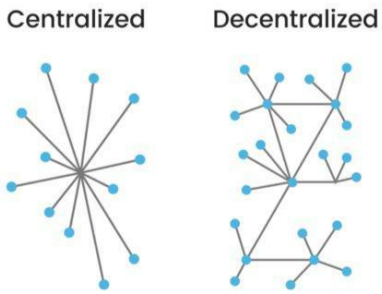

Figure 1.1 Comparisons graphs of Centralized and Decentralized (Distributed) system

This decentralized system or trustless transaction, does not required any third party as the media and also verifier thus, using many distributed ledgers. This distributed ledger records the transactions, such as the exchange of assets or data, among the participants in the network. Besides that, with this ledger, every data is shared, replicated, and synchronized among the members of a decentralized network. One of the solution for this decentralized system for transaction is blockchain system.

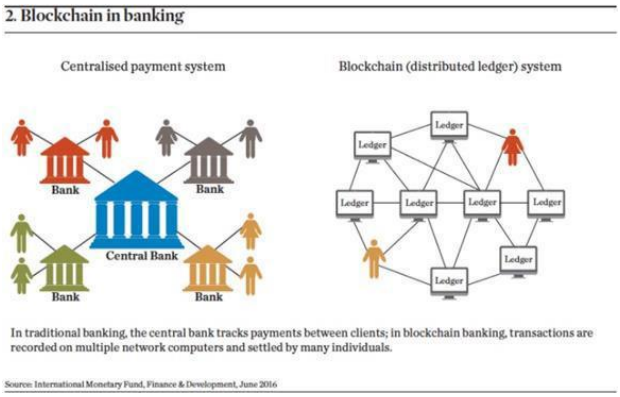

Figure 1.2 The system of Traditional Transaction system and Blockchain System

In Blockchain system, every transaction is validated. Every transaction occured is putted into block and each block is connected to each other and creating irreversible chain or "blockchain". Each individual doing this transaction is given a unique "digital signature" as their identification instead of actual name so that the transaction can be recorded

not only from the date of event but also among which parties it happens.

Every transaction happening in Blockchain is a unique one from its "code" that is stored inside the block. This unique code is generated through 'hash function'. This function ensures that every block is a unique one or its input has its own output. In Blockchain, many hash functions are implemented in the cryptography such as SHA1, SHA256 (Bitcoin), Keccak-256 (Etherum).

In this paper, we will be discussing regarding the hash function algorithm that is used in cryptography and also how actually this algorithm makes the decentralized or distributed system works.

II. HASH AND CRYPTOGRAPHY

In the blockchain system, there are many types of cryptographic hash function namely SHA256, Keccak-256 and many more. This section will focus on explaining both SHA256 and Keccak-256 and how they work.

A. *Hash Function*

Hash functions is a method to convert arbitrary size data into a predefined fixed length of digital strings, called hash.

One of the simplest hash function is modulo operation. It converts any digital string into a number which is divisible by a constant where the remainder of the division is the hash. In this modulo operation, several strings would result in the same modulo value as the output range of this operation is very small.

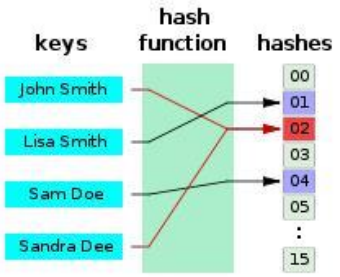

Figure 2.1. Simple Hash Function Collision (Source: Wikimedia Commons)

However, in Cryptography hash function, there are several properties of the function in order to be considered secure as it is used in transaction. First property is Deterministic, which means that everytime you parse a particular input into a hash function, it will generate the same result in order to keep track of every input. Second property is Quick where hash function must return hash quickly so that it is efficient. Third is Pre-Image resistance where it is infeasible to determine A inside H(A) function. Infeasible in this context is that it has a many diverse outputs resulting in a very small chance of colliding making it secure to be operated.

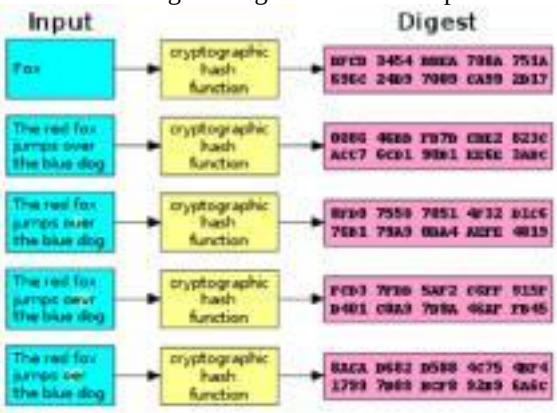

Figure 2.2 Transformation of Input using Cryptographic Hash Function

B. *SHA256 Cryptographic Hash Function*

One of the many Cryptographic hash function is SHA256. This Hash Function is used in one of the biggest cryptocurrency, Bitcoin. This SHA256 Hash Function produces 256-bit hash which means it will produces a fixed 256-bits length output. Here is the example of SHA256 algorithm:

SHA256("Hello World") =
a591a6d40bf420404a011733cfb7b190d62c65b
f0bcda32b57b277d9ad9f146e

In this example, it produces 256-bits string in a hexadecimal form.

Using this SHA256 Hash Function, in order to get the same input on every possibility, it needed an average of 2^256 trial to get the correct answer. Thus, making it impossible to reverse a SHA256 hash value as it takes an enormous time to be done.

C. *Keccak-256 Cryptographic Hash Function*

Another Cryptographic hash function that is used is Keccak-256. This hash function is also used in cryptocurrency which is Etherum. Keccak-256 produces the same output as SHA256 is which is 256-bits length output. Here is the example of Keccak-256 algorithm:

Keccak-256("Hello World") =
592fa743889fc7f92ac2a37bb1f5ba1daf2a5c84
741ca0e0061d243a2e6707ba

However, Keccak-256 uses different type of algorithm compared to SHA256. It uses a permutation from a set of 7 permutations in order to build this string. There are 7 KECCAK-f permutations, indicated by KECCAK-f[b], where = 25 2 and l ranges from 0 to 6. KECCAK-f[b] is a permutation over S ∈ 2 , where the bits

of s are numbered from 0 to b - 1. b is the width of the permutation. These KECCAK-f permutations are iterated constructions consisting of a sequence of almost identical rounds. The number of rounds nr depends on the permutation width, and is given by = 12 + 2l , where 2 = /25 . This gives 24 rounds for KECCAK-f[1600].

D. *Hashchains*

Hashchains is a sequence of homogeneous data-chunks that are linked together by hash function. In result, it creates a link of blocks where each block connects to each other its own 'key'.

Figure 2.3 Example of Hashchains

Each of the block in the hashchains is a unique one but they are connected to each other. Each data block contains hash and payload. This payload consists of arbitrary data. This hashchain has the important property where no data can be modified at any block without affecting the integrity of the subsequent blocks that is connected to it. For example, if the hash of the first block is changed then the hash of the second block will be changed too and the third one as well and so on. Based on this characteristic, nobody can change the data inside a block without changing the other blocks that are connected to it otherwise it will be invalid. Besides that, anyone who wants to add another block to the hashchains must also have the hash of the previous block.

This property adds with the public key cryptography become the basis of single blockchain.

E. *Public Key Cryptography*

Public Key Cryptography has the same idea as hash function which is one-way computation. The idea of one-way computation is that only a specific key value can access or encrypted a particular data. For example: given a key of *'m'*, where it can encrypt data *E(m)* while a key of *'n'* cannot access that data.

Having acknowledge that method, it can be said that Public Key Cryptography involves a pair of keys which is public and private keys (public key pairs). By having these keys, they will be associated with the entity that authenticate with the identity of the user electronically to encrypt or decrypt the data. The Public key is used to encrypt the data whereas the private key where it is kept secret is used to decrypt the encrypted data. Besides enabling the decryption and encryption of data, Public Key Cryptography also enables us to do Nonrepudiation which prevent the sender of data to claim the data was never sent and also preventing the data to be altered.

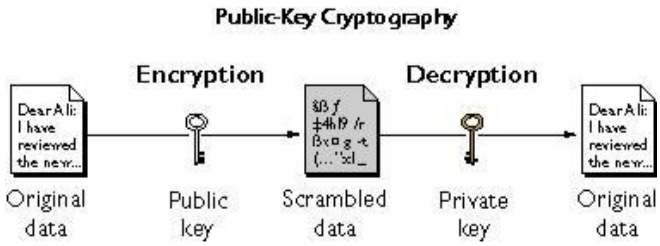

Figure 2.4 Public Key Cryptography Process (Source: IBM)

However, this Public Key Cryptography is not limited to store our data. It also enables us to send an encrypted message so that the data stay secret. This process consists of two scenarios. For example: Bobby and Rudy wants to send an encrypted message so that the content will only be acknowledged by them.

i) Bobby wants to send a secret message so he will encrypt the message using Rudy Public Key (Let's say r) because it can only be accessed using Rudy Private Key based on the one-way computation method. The only vulnerability of this process is that everyone using this Blockchain must know and trust this Public key in the first place.

ii) Rudy will receive the encrypted message. By having the private key where it suits according to one-way computation (the private key can decrypt *E(r))* then Rudy will be able to access and decrypt the secret message. However, Rudy needs to publish both original and encrypted data so that anyone who owns the public key can verify that the data is actually decrypted with the matching pair of keys. This verification process is called digital signing and the decrypted data is taken as the digital signature.

III. BLOCKCHAIN AND DIGITAL STRUCTURE

In the previous sections, it has been explained regarding the concept of hash function algorithm, list of cryptography hash function, hashchains and also public key cryptography. This section will further discuss about how cryptography hash function acts as the key in enabling the distributed and trustless transactions which by making digital signature in blockchain.

A. *Blockchain*

Blockchain is a ledger where it tracks every transaction that is happening. It groups the transaction entries by chaining them together to build a complete ledger. The transactions is grouped and chained together to enable the distributed transactions methods.

The concept of blockchain is the same almost the same as hashchains. We can say that blockchain consists of many hashchains inside. The structure of the transaction entries is almost the same as hashchain as it contains many blocks with its own unique properties. Therefore, we cannot change only one particular block inside the blockchain without changing other blocks inside of it as they are connected and chained to each other.

Having understand the concept of blockchain, it can be said that the structure of blockchain consists of many linked lists where each block contains data and also hash of previous block header.

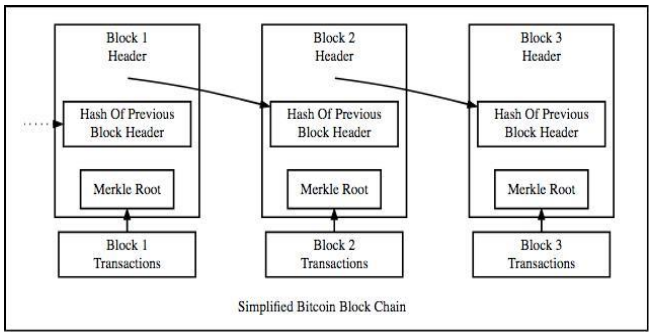

Figure 3.1. Blockheader
(Source: Wikimedia Commons)

Inside this Block header, it stores the data regarding the timestamp of transaction, the version of the block, the target block and also the previous block and also Merkle root. This blockchain can also be described into a binary-hashchain or Merkle Tree

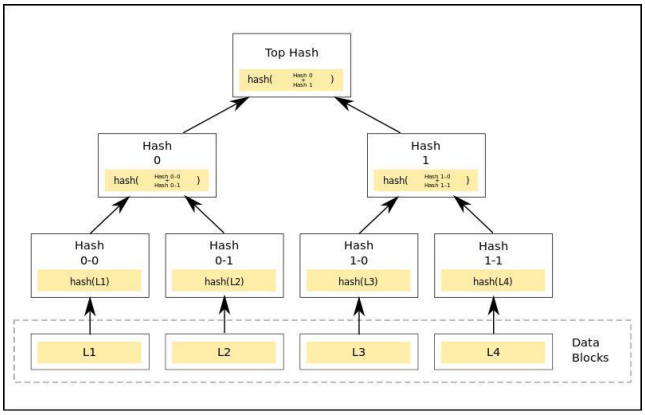

Figure 3.2 Merkle Tree Diagram
(Source: Blockgeeks)

In the Merkle Tree diagram, the Top Hash of the tree is called the Merkle Root where each of the non-leaf node is the hash of the values of their child nodes. By doing this, it will make the structure efficient as each of the block contains its own pointers.

In this header block where it contains these hashes, it also tracks all transactions that are happening between both payee and paid parties. In order to validate whether the transaction is valid or not, there is some authentication that need to be done and that's when the digital signature is used.

B. *DIGITAL SIGNATURE IN CRYPTOGRAPHY HASH FUNCTION*

Digital Signature can be produced through the public key cryptography. The purpose of this digital signature is to validate whether the transaction has happened between two parties and both parties are aware of this transaction. Whenever a transaction wants to be done, the payee party mush sign a validation using his or her private key with the Cryptography Hash Function algorithm.

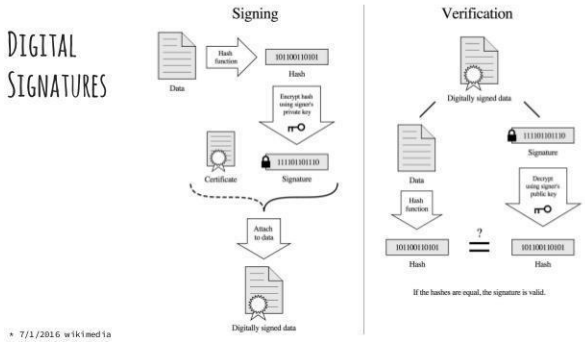

Figure 3.3 Digital Signature in trustless transaction
(Source: Wikimedia Commons)

Whenever anyone wants to validate whether the transaction is valid or not, they can use the payee public key to confirm it. This process needed to be done in order to keep the integrity and also to track every transaction record that is happening.

Another way to think about the digital signature is by visualizing the flow of values – a cryptocurrencies, a coin – in the system. In the original Bitcoin Pape [1] , where the author explains that the coin can be used as digital signature by transferring the coin and digitally signing a hash of the previous transaction and the public key of the next owner and adding these to the end of the coin.

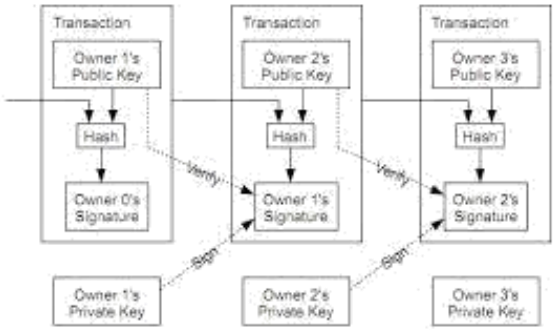

Figure 3.4 Flow of value in Digital Signature (Source: Nakamoto Nakamoto, Satoshi. "Bitcoin: A Peer-to-Peer Electronic Cash System". 2008)

IV. SECURITY AND CONDITION

A. Overview

Traditional transactions have happened since very long ago. Mostly this transaction is based on a trust on a particular medium or third party. The level of security from the third party also keeps on increasing as many methods of frauds keep on coming out. By using the facility from this third party, our data and also savings are secured in their security access.

However, this traditional system has many drawbacks from the implementation. With this centralized transaction is based on trust from the third party in controlling every record of transactions, the people's records or data can be accessed by anybody who is authorized or influenced people so that this data can be manipulated. Therefore, the privacy of each person is still in stake. Besides that, from the facility provided by the third party such as banks, to provide security in our savings, it costs the people a high number of interest. This is also one of the drawbacks from the centralized system of transactions.

While this system has been implemented for many years in this era, it is quite effective even though there are still many drawbacks from it. An even more effective system is proposed which is through distributed system of transaction, or in this context is Blockchain.

Distributed system, on the other hand, does not require third party as the media to store and be the bridge in connecting people to do transaction. The safety and process in this distributed system is also guaranteed through the system of Cryptography hash function. Through this hash function, our transaction is protected with mathematical method where it is irreversible in a way that every transaction that is happening will be a unique one.

Besides that, the transparency and also the validity of the transaction is also guaranteed through the Public Key Cryptography method. This ensure that whenever a secret message or transaction wanted to be done, both parties must have the matching keys so that it can be conducted. Moreover, every transaction is also chained to each other so that we can keep record on every transaction that has been done.

Through this distributed system, people do not need to pay the provided facilities meaning, this system has eliminated the high interest and fees issue in the centralized system of transaction. This makes people to choose distributed system rather than the traditional centralized system.

B. Mathematical Calculation

In the previous sections, it is discussed regarding the Cryptographic Hash Function that is implemented. In this context, two of the functions are discussed which is SHA256 and Keccak-256. From these two functions that are discussed, it is known that both of these functions produced 256-bits strings output, with one function uses 2^8 combinations to produce the strings while the other uses permutation from a set of 7 permutations in order to build this string.

In the SHA256 hash function, it gets even complicated when operations are included if people want to crack the hash functions. There is a total of 600 operations in additions operation, 576 in bitwise rotation, 96 bitwise shifts, 320 bitwise AND,640 bitwise EX-OR with total operations of 2232. In result, it will return number of Mod 2^{32}.

Additions (Mod 2^{32})	= (7*64) + (3*48) + 8 = 448 + 144 + 8 = **600**	(message compression) + (message scheduler) + (intermediate/final hash computation)
Bitwise Rotations (ROTR)	= (6*64) + (4*48) = 384 + 192 = **576**	$(\sum_0, \sum_1) + (\sigma_0, \sigma_1)$
Bitwise Shifts (SHR)	= 2*48 = **96**	σ_0, σ_1
Bitwise AND (∧)	= 5*64 = **320**	Maj, Ch
Bitwise EX-OR (⊕)	= (7*64) + (4*48) = 448 + 192 = **640**	(message compression) + (message scheduler)
Total Operations	= 600 + 576 + 96 + 320 + 640 = **2232**	

Figure 4.1 Calculations of Complexity in SHA256 (Source: Rahul P. Naik. Optimising the SHA256 Hashing Algorithm for Faster and More Efficient Bitcoin Mining. 2013)

This calculation proves the complexity so high where the security is guaranteed in using this distributed system. Therefore, this minimize and even eliminating the possibility of being hacked or getting fraud from transactions that is done.

V. IMPACTS

The distributed system of transactions with the assistance of Blockchain has revolutionized the way people do transactions nowadays. Through this system, it gives many impacts towards the society in both economic and social sectors. The advantages of its safety, transparency, the ability to track every transaction record and also minimum to no fee in using the facilities has given many economical and also social impacts towards the society. Not only that, it also helps increasing the effectiveness and efficiency in doing transactions as this system eliminates the third party between the payee and payer parties. Moreover, this system also opens many ways of recording not only in transaction records but also in land ownership and many other things.

This system can only be achieved with the help of Cryptographic Hash Functions algorithm. With this algorithm, there will be many more improvements towards the distributed system so that it can be the alternative towards the traditional transaction system.

VI. ACKNOWLEDGEMENT

I would like to thank all the lecturers of the knowledge that is shared regarding discrete mathematics. This subject given me the chance to explore more regarding several topics in the subject and also its application. I would also like to thank my friends and families who support me in the process of learning and also making this paper

REFERENCES

[1] Nakamoto, Satoshi. *Bitcoin*: *A Peer-to-Peer Electronic Cash System. 2008*

[2] Oleg Mazonka. *Blockchain: Simple Explanation. 2016*

[3] https://blockgeeks.com/guides/what-is-hashing/

[4] https://www.investopedia.com/articles/investing/083115/blockchain-technology-revolutionize-traditional-banking.asp

[5] A. Gholipour and S. Mirzakuchaki. *A Pseudorandom Number Generator with KECCAK Hash Function. 2011*

[6] https://www.ibm.com/support/knowledgecenter/en/SSB23S_1.1.0.13/gtps7/s7pkey.html

[7] Rahul P. Naik. *Optimising the SHA256 Hashing Algorithm for Faster and More Efficient Bitcoin Mining. 2013*

PERNYATAAN

Dengan ini saya menyatakan bahwa makalah yang saya tulis ini adalah tulisan saya sendiri, bukan saduran, atau terjemahan dari makalah orang lain, dan bukan plagiasi

Bandung, 8 Desember 2018

Jofiandy Leonata Pratama
13517135