Information Propagation Methods Based on Blockchain

GECHANG FENG
StudentID:515030910601
Shanghai Jiao Tong University
May 27, 2018

Abstract

Blockchain technology, as a decentralized and non-hierarchical platform, has the potential to replace centralized systems. As the information propagation methods are growing a P2P tendency, the information propagation method based on blockchain is meaningful to study with.

In this paper, I introduce the Blockchain technology and the exsiting information sharing methods and analyze how Bitcoin propagate the information of transactions using blocks through the network to update the ledger replicas. Then I assume the method of information propagation based on the blockchain based on the study of Bitcoin propagation, including the information resource chain model and the information index chain model and then show the advantages and limitations of these methods.

Keywords: block, blockchian, information propagation, Bitcoin

1 Introduction

With the development of Internet, there are increasingly varied number of methods of accessing information. From the traditional centralized media such as newspaper and television news broadcast to the social network which is popular now, the information propagation methods are growing a P2P tendency.

1.1 Blockchain

Blockchain was invented as a decentralized public data ledger. Every data committed in blockchain should be verified by all users.

A blockchain is a continuously growing list of records, called blocks, which are linked and secured using cryptography.[1] Each block typically contains a cryptographic hash of the previous block,[3] a timestamp and transaction data. By design, a blockchain is inherently resistant to modification of the data. For use as a distributed ledger, a blockchain is typically managed by a peer-to-peer network collectively adhering to a protocol for inter-node communication and validating new blocks. Once recorded, the data in any given block cannot be altered retroactively without the alteration of all subsequent blocks, which requires collusion of the network majority.

The information technology involved in blockchain includes peer to peer (P2P), asymmetric encryption algorithm, database technology, digital currency, and Merkle tree, Hash algorithm, timestamp technology, and multiple consensus mechanism (Consensus Mechanism),etc..

As a new IT infrastructure, blockchain has 5 basic characteristics: Decentralization, Autonomy, Security and trustworthiness, Open and transparent, Anonymity.

According to the different range of service oriented groups, there are three types of blockchain networks - public blockchains, private blockchains and consortium blockchains. A public blockchain has absolutely no access restrictions. Anyone with an internet connection can send transactions to it as well as become a validator (i.e., participate in the execution of a consensus protocol). Usually, such networks offer economic incentives for those who secure them and utilize some type of a Proof of Stake or Proof of Work algorithm. Some of the largest, most known public blockchains are Bitcoin and Ethereum. A private blockchain is permissioned. One cannot join it unless invited by the network administrators. Participant and validator access is restricted. This type of blockchains can be considered a middle-ground for companies that are interested in the blockchain technology in general but are not comfortable with a level of control offered by public networks. Typically, they seek to incorporate blockchain into their accounting and record-keeping procedures without sacrificing autonomy and running the risk of exposing sensitive data to the public internet. A consortium blockchain is often said to be semi-decentralized. It, too, is permissioned but instead of a single organization controlling it, a number of companies might each operate a node on such a network. The administrators of a consortium chain restrict users' reading rights as they see fit and only allow a limited set of trusted nodes to execute a consensus protocol.

According to the hierarchical structure of block chain, we can build

various applications in the application layer of block chain by using different versions and types of block chains.

1.2 Methods of Information Propagation

According to the different ways of information dissemination, there are two basic modes of information resource sharing, that is, the central mode and the non-central mode.The typical representative of the central mode is C/S and B/S application, in which there are two fixed roles of the server and the user. The typical representatives of non-central models are P2P applications, which including only equal participants.

The prerequisite of information sharing is the accumulation of information resources, that is, organizing information resources in some way. Under the central mode, all information resources are centrally allocated to the center, and will gather to form an information resource center while With the non-central mode, all information resources are dispersed to each node, so that no information resource center will be formed.

According to the different gathering process, the central mode of information resource sharing can be divided into two types: information resource sharing model based on Cloud and information resource sharing model based on Web. The model based on Cloud is commonly used in information, intelligence and knowledge sharing, in which information gathering is generally done automatically by the Cloud platform, and mainly based on behavioral data, interactive data, information and knowledge of Cloud users. The model based on Web is commonly used for digital content sharing, in which information aggregation is usually done by producers and added to the platform. Both Web1.0 and Web2.0 follow this resource aggregation model. However, by removing the centralized audit and publishing mechanism of Web1.0, the Web2.0 liberalized the writing and publishing rights to the users, and the users became the content producers, thus realizing the deprivilegization, non-elites, the user-centered, and the people-centered.

According to the different ways of organization of information resources, the non central mode of information resources sharing can be divided into three types: centerless unstructured P2P mode (pure P2P mode), centerless structured P2P mode (mixed P2P mode), and central unstructured P2P mode (centralized P2P mode). The centerless unstructured P2P mode, where participants query or find information resources shared in a peer-to-peer network based on the Flooding technology by broadcast, characterized by disordered and decentralized information and also scattered content. In the centerless structured P2P model of information resource sharing, the Overlay Network such ans Chord ring or Binary Tree are established based on DHT technology to query and find information resources shared in the

P2P network. The exchange or dissemination of information resources among participants is not covered by the Overlay Network, characterized by the orderly and decentralized information and also scattered content. In the central unstructured P2P model of information resource sharing, participants rely on the index center to query or find information resources shared in the P2P network. The exchange or dissemination of information resources among participants does not pass through the index center, which is characterized by orderly and centralized information, but the content is separated. These three models are mostly used for digital content sharing in file form.

2 Information Propagation on the Bitcoin Network

In order to learn the information propagation methods on blockchain, I study the information propagation method on the Bitcoin network. Bitcoin is the first truly decentralized global currency system which is based on blockchian. Like any other currency, its main purpose is to facilitate the exchange of goods and services by offering a commonly accepted good. Bitcoin does not rely on a centralized authority to control the supply, distribution and verification of the validity of transactions. Bitcoin relies on a network of volunteers, to collectively implement a replicated ledger. The ledger tracks the balance of all accounts in the system. Each node keeps a complete replica of the ledger. It is crucial for the replicas of the ledger to be in a consistent state across all nodes at all times as the validity of transactions is verified against them.

2.1 Transactions and Blocks

In Bitcoin two distinct types of information are disseminated: *transactions* and *blocks*. Transactions are the primitives that allow the transfer of value, whereas blocks are used to synchronize state across all nodes in the network.[4]

At an abstract level a transaction transfers bitcoins from one or more *source accounts* to one of more *destination accounts*. An account is in essence a public-/private-keypair. An address derived from the public key is used to identify the account. To transfer bitcoins to an account a transaction is created with the address of the account as destination. To send bitcoins from an account, the transaction has to be signed with the private key associated with the sending account. Instead of aggregating the balance of each account, the ledger tracks *outputs* that transferred the bitcoins to the ac- count. An output is a tuple of a numeric value in bitcoins and a condition to claim or

spend that output. As transactions are broadcast through the network the state of the ledger replicas changes. When a node receives a new transaction, it is verified and committed to the local replica.

In order for the ledger replicas to remain consistent a common order over the transactions has to be agreed among the nodes. Agreeing on a common order of transactions in a distributed system is not trivial. Bitcoin solves this problem by tentatively committing transactions and then synchronizing at regular intervals by broadcasting a block created by one of the nodes. A block b contains the set of transactions T_b that the node which created the block has committed since the previous block. The block is then distributed to all the nodes in the network and each node receiving it will roll back the tentatively committed transactions since the last block and apply the transactions from the current block. At this point all the nodes have agreed on the validity of all the transactions in the block. Transactions that were committed as part of the block are confirmed and do not have to be reapplied. The transactions that have been rolled back will then be validated again and reapplied on top of the new base state. Transactions that are now invalid because they conflict with transactions committed as part of the block are discarded. The node cannot forge any transactions as long as the underlying public- /private-key cryptosystem is secure. The block creator may only decide in which order transactions arrived and whether to include transactions in its block.

The blockchain chain the blocks to provide an added synchronization on top of the individual transactions.The blocks are organized in a directed tree. Each block contains a reference to a previously found block. The root block in the tree is the *genesis block*, which is hardcoded into the clients. This block is an ancestor of all blocks by definition. The blockchain is defined as the longest path from any block to the genesis block. The chaining is used to assign a chronological order to the transactions: transactions in lower height blocks have been verified before transactions in higher blocks.

2.2 Information Propagation

The Bitcoin network is a network of homogeneous nodes. There are no coordinating roles and each node keeps a complete replica of all the information needed to verify the validity of incoming transactions. Each node verifies information it receives from other nodes independently and there is only minimal trust between the nodes.

In order to avoid sending transaction and block messages to nodes that already received them from other nodes, they are not forwarded directly. Instead their availability is announced to the neighbors by sending them an *inv message* once the transaction or block has been completely verified.

The inv message contains a set of transaction hashes and block hashes that have been received by the sender and are now available to be requested. A node, receiving an inv message for a transaction or block that it does not yet have locally, will issue a *getdata* message to the sender of the inv message containing the hashes of the information it needs. The actual transfer of the block or transaction is done via individual block or transacion messages.

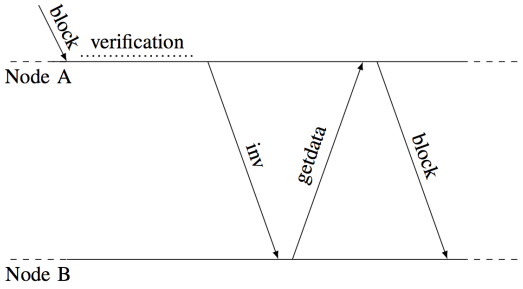

Figure 1: Messages exchanged in order to forward a block message a single hop from Node A to Node B.

Each block or transaction is introduced to the network at one of the nodes, its *origin*, and is then propagated throughout the network using the above broadcast mechanism. At each hop in the broadcast the message incurs in a *propagation delay*. The propagation delay is the combination of transmission time and the local verification of the block or transaction.

The propagation of a data item can be divided into two phases: an initial exponential growth phase in which the most of the nodes receiving inv messages will request the corresponding data item as they do not have it yet, and an exponential shrinking phase in which most of the nodes receiving an announcement already have the corresponding data item.

3 Information Propagation Based on Blockchain

After studying the information propagation method on the Bitcoin network, I found that the kernel of information propagation on blockchain is creating blocks related to every nodes and chain them to assign a chronological order to spread the information. Thus, the transacation message in the Bitcoin network can be replaced by a generalized message contains information to propagates among the users by blockchain.

First, we must solve the problem of information storage to realize the sharing of information resources.

One way is to use information resource chain model, that is, store information directly in blocks and cluster them in blockchain. This method is similar to the storage model in the non-central unstructured P2P model, in which the block chain is used as the query or lookup of the information

resources in the P2P network, as well as the storage of information resources. The characteristic of this method is that the information is out of order and centralized, and is stored in every node in the form of complete replica. Take the Bitcoin blockchain as a example.In the block of the Bitcoin blockchain, few fields can be used for storing information, including the coinbase field of output in Generation TX, and the transaction output address of the general address transaction (Pubkey Hash TX). However, these two are of small capacities and lots of limitation, and will increase the size of the UTXO set in memory.

Thus, I considered the information index chain model. In blocks in this blockchain, only the metadata or index of information resources is stored, and information resources are stored locally, or in servers, Cloud storage or other storage, and information resources are gathered outside the block and block chain. This method is similar to the storage model of the central unstructured P2P mode. The blockchain is used as the information resource query or search in the P2P network, and does not participate in the actual exchange and propagation of information resources. The exchange and propagation of information resources are realized by the transmission mechanism of the underlying peer network, and the Smart Contract may be used to protect the information resources. The evidence sharing behavior is reliably executed. The characteristic of this method is that the information is out of order and centralized, but the content is scattered outside the blockchain.

The blockchain on P2P network realizes information resource aggregation by block list, and the information resource sharing is realized by the synchronous copy of blockchain, which has the function of gathering and sharing information resources itself. Therefore, the expansion of blockchain network is a way to achieve wider sharing of information resource on blockchain. That is, more people join the blockchain network to become nodes and participants in the network, so as to share the information resources on the blockchain according to rules. The public chain, private chain and consortium chain can all adopt this mode to share information, but it is more meaningful to the information sharing on public chains.

4 Conclusion

Comparing with the traditional centralized information propagation methods, the information propagation method based on blockchain has some advantages.

1. Different from the information resource sharing model based on Web, the information sharing method based on blockchain is decentralized,

and there is no fixed service center in the system. Each participant in this system is sharing information and obtaining information on an equal basis, which realize the information propagation based on information sharing.

2. The process of information sharing is the process of information gathering, which are recorded in the transaction. The transactions in the block are transparent, credible and traceable, and will not reveal the private information of the participants because of the anonymity of the blockchain.
3. Information propagation based on blockchain does not need intermediary participation, and direct value transfer in the process of information sharing is realized with the aid of virtual currency and end-to-end direct exchange. The information sharing behavior is driven and motivated by benefits, and the information resources are automatically configured to each Full Node, rather than a specific fixed service center, is especially beneficial to the construction of information resources co construction and sharing system.
4. It can effectively resist virus and many kinds of network attacks. As long as there is a complete node, all the information of the whole system can be restored and run again. The security and reliability of the system is obviously superior to the information resource sharing model based on Cloud or Web, which can ensure the sustainable and stable development of the information resource sharing.

There exists limitations in the method as well.

1. Information resource chain model is restricted by the block 1MB size limit. The storage space available in the block is limited, which can not meet the storage requirements of large granularity information, and it is also difficult to meet the variable requirements of the information resource to the storage space, and it also limits the number of transactions that can be accommodated in the block, which brings certain influence to the normal operation of blockchain.
2. As for information index chain model, in order to ensure that the corresponding information can be reliably accessed at any time, it usually needs to store information in secure and persistent online places such as cloud storage or storage cluster, which will produce different degrees of dependence on cloud storage and weaken the real effect of the whole system to centralization.
3. Because of the restriction of the operating mechanism of blockchain network, the propagation speed of the information sharing system is

relatively low, especially the actual propagation speed of the public chain is far lower than the actual propagation speed of the current mainstream centralized system. It is easy to form a bottleneck in large-scale information propagation environment.

4. Due to the unmodifiable features of the block chain, the data and information recorded in block chain blocks can be viewed, but cannot be changed, undeleted, or shielded, even if the information is proved to be wrong and illegal. This poses a great challenge to the management of information.

The disadvantages of information propagation model and corresponding system based on blockchain are mainly due to the constraints and deficiencies of the current blockchain network and the blockchain technology.

The application in the field of information propagation service of blockchain is still in the early stage of trial. The typical block, blockchain and blockchain network are oriented to the virtual currency and financial field, and there are some inappropriate places when they are applied to the field of information propagation service. To overcome these shortcomings, we need to improve the existing blocks, blockchains and their common understanding mechanism, operation mechanism, and even blockchain network.[5] It is a tough work remain to address in the future years. The application of blockchain in scale of information propagation service still needs a long time.

References

1. "Blockchains: The great chain of being sure about things" (31 October 2015). *The Economist.*

2. Brito, Jerry; Castillo, Andrea (2013). Bitcoin: A Primer for Policymakers

3. Narayanan, Arvind; Bonneau, Joseph; Felten, Edward; Miller, Andrew; Goldfeder, Steven (2016). Bitcoin and cryptocurrency technologies: a comprehensive introduction.

4. Decker C, Wattenhofer R. (2013) Information propagation in the Bitcoin network *IEEE Thirteenth International Conference on Peer-To-Peer Computing*, 2013:1-10.

5. Ersoy O, Ren Z, Erkin Z, et al. (2017) Information Propagation on Permissionless Blockchains

6. Nakasumi M. (2017) Information Sharing for Supply Chain Management Based on Block Chain Technology *IEEE, Conference on Business Informatics*, 2017:140-149.