An Overview of Blockchain Technology: Architecture, Consensus, and Future Trends

Zibin Zheng[1], Shaoan Xie[1], Hongning Dai[2], Xiangping Chen[4], and Huaimin Wang[3]

[1]School of Data and Computer Science, Sun Yat-sen University Guangzhou, China
[2]Faculty of Information Technology, Macau University of Science and Technology, Macau, SAR
[3]National Laboratory for Parallel & Distributed Processing
National University of Defense Technology, Changsha 410073 China
[4]Institute of Advanced Technology,National Engineering Research Center of Digital Life
Sun Yat-sen University, Guangzhou, China
Email: zhzibin@mail.sysu.edu.cn

***Abstract*—Blockchain, the foundation of Bitcoin, has received extensive attentions recently. Blockchain serves as an immutable ledger which allows transactions take place in a decentralized manner. Blockchain-based applications are springing up, covering numerous fields including financial services, reputation system and Internet of Things (IoT), and so on. However, there are still many challenges of blockchain technology such as scalability and security problems waiting to be overcome. This paper presents a comprehensive overview on blockchain technology. We provide an overview of blockchain architechture firstly and compare some typical consensus algorithms used in different blockchains. Furthermore, technical challenges and recent advances are briefly listed. We also lay out possible future trends for blockchain.**

***Index Terms*—Blockchain, decentralization, consensus, scalability**

I. Introduction

Nowadays *cryptocurrency* has become a buzzword in both industry and academia. As one of the most successful cryptocurrency, Bitcoin has enjoyed a huge success with its capital market reaching 10 billion dollars in 2016 [1]. With a specially designed data storage structure, transactions in Bitcoin network could happen without any third party and the core technology to build Bitcoin is *blockchain*, which was first proposed in 2008 and implemented in 2009 [2]. Blockchain could be regarded as a public ledger and all committed transactions are stored in a list of blocks. This chain grows as new blocks are appended to it continuously. Asymmetric cryptography and distributed consensus algorithms have been implemented for user security and ledger consistency. The blockchain technology generally has key characteristics of decentralization, persistency, anonymity and auditability. With these traits, blockchain can greatly save the cost and improve the efficiency.

Since it allows payment to be finished without any bank or any intermediary, blockchain can be used in various financial services such as digital assets, remittance and online payment [3], [4]. Additionally, it can also be applied into other fields including smart contracts [5], public services [6], Internet of Things (IoT) [7], reputation systems [8] and security services [9]. Those fields favor blockchain in multiple ways. First of all, blockchain is immutable. Transaction cannot be tampered once it is packed into the blockchain. Businesses that require high reliability and honesty can use blockchain to attract customers. Besides, blockchain is distributed and can avoid the single point of failure situation. As for smart contracts, the contract could be executed by miners automatically once the contract has been deployed on the blockchain.

Although the blockchain technology has great potential for the construction of the future Internet systems, it is facing a number of technical challenges. Firstly, scalability is a huge concern. Bitcoin block size is limited to 1 MB now while a block is mined about every ten minutes. Subsequently, the Bitcoin network is restricted to a rate of 7 transactions per second, which is incapable of dealing with high frequency trading. However, larger blocks means larger storage space and slower propagation in the network. This will lead to centralization gradually as less users would like to maintain such a large blockchain. Therefore the tradeoff between block size and security has been a tough challenge. Secondly, it has been proved that miners could achieve larger revenue than their fair share through selfish mining strategy [10]. Miners hide their mined blocks for more revenue in the future. In that way, branches could take place frequently, which hinders blockchain development. Hence some solutions need to be put forward to fix this problem. Moreover, it has been shown that privacy leakage could also happen in blockchain even users only make transactions with their public key and private key [11]. Furthermore, current consensus algorithms like *proof of work* or *proof of stake* are facing some serious problems. For example, proof of work wastes too much electricity energy while the phenomenon that the rich get richer could appear in the proof of stake consensus process.

There is a lot of literature on blockchain from various sources, such as blogs, wikis, forum posts, codes, conference proceedings and journal articles. Tschorsch et al. [12] made a technical survey about decentralized digital currencies

978-1-5386-1996-4/17 $31.00 © 2017 IEEE
DOI 10.1109/BigDataCongress.2017.85

TABLE I: Comparisons among *public blockchain*, *consortium blockchain* and *private blockchain*

Property	Public blockchain	Consortium blockchain	Private blockchain
Consensus determination	All miners	Selected set of nodes	One organization
Read permission	Public	Could be public or restricted	Could be public or restricted
Immutability	Nearly impossible to tamper	Could be tampered	Could be tampered
Efficiency	Low	High	High
Centralized	No	Partial	Yes
Consensus process	Permissionless	Permissioned	Permissioned

- *Auditability.* Bitcoin blockchain stores data about user balances based on the Unspent Transaction Output (UTXO) model [2]: Any transaction has to refer to some previous unspent transactions. Once the current transaction is recorded into the blockchain, the state of those referred unspent transactions switch from unspent to spent. So transactions could be easily verified and tracked.

D. Taxonomy of blockchain systems

Current blockchain systems are categorized roughly into three types: public blockchain, private blockchain and consortium blockchain [17]. In public blockchain, all records are visible to the public and everyone could take part in the consensus process. Differently, only a group of pre-selected nodes would participate in the consensus process of a consortium blockchain. As for private blockchain, only those nodes that come from one specific organization would be allowed to join the consensus process.

A private blockchain is regarded as a centralized network since it is fully controlled by one organization. The consortium blockchain constructed by several organizations is partially decentralized since only a small portion of nodes would be selected to determine the consensus. The comparison among the three types of blockchains is listed in Table I.

- *Consensus determination*. In public blockchain, each node could take part in the consensus process. And only a selected set of nodes are responsible for validating the block in consortium blockchain. As for private chain, it is fully controlled by one organization and the organization could determine the final consensus.
- *Read permission*. Transactions in a public blockchain are visible to the public while it depends when it comes to a private blockchain or a consortium blockchain.
- *Immutability*. Since records are stored on a large number of participants, it is nearly impossible to tamper transactions in a public blockchain. Differently, transactions in a private blockchain or a consortium blockchain could be tampered easily as there are only limited number of participants.
- *Efficiency*. It takes plenty of time to propagate transactions and blocks as there are a large number of nodes on public blockchain network. As a result, transaction throughput is limited and the latency is high. With fewer validators, consortium blockchain and private blockchain could be more efficient.
- *Centralized*. The main difference among the three types of blockchains is that public blockchain is decentralized, consortium blockchain is partially centralized and private blockchain is fully centralized as it is controlled by a single group.
- *Consensus process*. Everyone in the world could join the consensus process of the public blockchain. Different from public blockchain, both consortium blockchain and private blockchain are permissioned.

Since public blockchain is open to the world, it can attract many users and communities are active. Many public blockchains emerge day by day. As for consortium blockchain, it could be applied into many business applications. Currently Hyperledger [18] is developing business consortium blockchain frameworks. Ethereum also has provided tools for building consortium blockchains [19].

III. Consensus Algorithms

In blockchain, how to reach consensus among the untrustworthy nodes is a transformation of the Byzantine Generals (BG) Problem, which was raised in [20]. In BG problem, a group of generals who command a portion of Byzantine army circle the city. Some generals prefer to attack while other generals prefer to retreat. However, the attack would fail if only part of the generals attack the city. Thus, they have to reach an agreement to attack or retreat. How to reach a consensus in distributed environment is a challenge. It is also a challenge for blockchain as the blockchain network is distributed. In blockchain, there is no central node that ensures ledgers on distributed nodes are all the same. Some protocols are needed to ensure ledgers in different nodes are consistent. We next present several common approaches to reach a consensus in blockchain.

A. Approaches to consensus

PoW (Proof of work) is a consensus strategy used in the Bitcoin network [2]. In a decentralized network, someone has to be selected to record the transactions. The easiest way is random selection. However, random selection is vulnerable to attacks. So if a node wants to publish a block of transactions, a lot of work has to be done to prove that the node is not likely to attack the network. Generally the work means computer