Blockchain Beyond Cryptocurrency: An Overview

Hong Wan, NCSU.

In the past 30 months, "blockchain" has evolved from a computer science terminology to a buzz word in the still ongoing cryptocurrency hype. People start to show enormous interests of the blockchain, which is the foundation of cryptocurrency. Several articles have praised the blockchain as "the most important invention after the internet." [1]

Among all the discussions, one essential topic is how to apply blockchains beyond cryptocurrency. Like all emerging technology, there have been many confusions and misunderstanding of the blockchain concepts. Among them, a popular one is that "Blockchains are all like the ones used in bitcoin”. This is wrong. In this paper, we will introduce various kinds of blockchains, and why more centralized blockchain structures are more appropriate for business usage.

A blockchain is a **digital, append-only, timestamped ledger.** In details, a blockchain is a consensus-based, peer-to-peer distributed network with "a growing list (chain) of records, called blocks, that are linked using cryptography. Each block contains a unique hash value of the previous block, a timestamp, and transaction data." [2] Here consensus refers to a set of rules that users follow to agree on the states of the system. It makes the blockchain self-auditing ecosystem[3]. We will discuss more of the consensus later in the article. Hash is a cryptographic function that converts a string into a non-meaningful, fixed length output. It is non-reversible since the hash value is highly sensitive to the input. A small change of input will lead to a completely different hash value. The mathematical structure of the blockchain implies two essential properties of the blockchains. First, the data (in block) is immutable. Specifically, if you change a block, all blocks BEFORE this one become invalid since all hash values in these blocks will become invalid. Besides, a distributed network with consensus allows users to communicate directly with each other for broadcasting a new block and synchronize the blockchain status. All users can download a copy of the current ledger and add blocks, which means that there is redundancy of the data in the network. Therefore, the blockchain is more tolerable to nodes' failures. Combined the two points, the longer the chain and more users (nodes) in the network, the harder to hack into the chain and change blocks without detection, and the blockchain is more reliable.[4]

The first work on building a cryptographically secured chain of blocks was proposed by Stuart Haber and W. Scott Stornetta in 1991[5]. The concept is formalized by Satoshi

[1] For example, https://medium.com/@markymetry/blockchain-technology-is-the-most-significant-invention-since-the-internet-and-electricity-f2d44a631ef6

[2] https://en.wikipedia.org/wiki/Blockchain

[3] https://lisk.io/academy/blockchain-basics/how-does-blockchain-work/consensus-protocols

[4] https://bitcoin.org/bitcoin.pdf

[5] Haber, S. & Stornetta, W.S. J. Cryptology (1991) 3: 99. https://doi.org/10.1007/BF00196791

Nakamoto[6] in 2008, who proposed "A purely peer-to-peer version of electronic cash that would allow online payments to be sent directly from one party to another without going through a financial institution… **(a system that is) Based on cryptographic proof instead of trust...**"[3]. In Jan 2009, he mined the first *bitcoin* and started the era of cryptocurrency. In recent years, big tech companies like IBM and Microsoft as well as many emerging startups put significant effort to extend the blockchain systems to various industry. Many of them use blockchains that with significant different structural compared to cryptocurrency, which we elaborate below.

From the governance point of view, the blockchain can *be* ***public (open) or private (closed)***. In a public chain, anyone can initiate transactions, generate and broadcast blocks, as well as download a copy of the whole ledger. In a private chain, only **authorized** users can access the network. The blockchain can also be permissioned ***and permissionless***. The permissioned chain means the rights/authorization of users can be different. Some have more authorities (for example, validate the block) than the others (read-only). The permissionless chain means that all users have the same rights. Anyone in the network can download the ledger, generate blocks, and validate transactions. While many consider permissioned and private as the same, they are, in fact, two concepts. A blockchain is public or private is independent of whether it is permissioned or not. As discussed in Chris Jaikaran's testimony to congress: "Discussing a blockchain as public or private refers to the level of freedom users have to create identities and read data on that blockchain. Discussing a blockchain as permissioned or permissionless refers to the level of access the user would have on that blockchain." [7]Specifically, the public and permissionless chain has the most decentralized structure and assumes no trust among users. The private and permissioned chain, on the other hand, has the most centralized architecture and the highest level of trust. Besides, the private and permissioned chain usually do not need tokens/coins. The other two combinations are in between as hybrid. Each kind of blockchain has unique properties that are fit for different applications, which we elaborate below.

A majority of cryptocurrencies use public and permissionless chains. This is also the most well-known blockchain type. To understand how it works, we compare how financial transactions are handled by a traditional bank and by a blockchain. Suppose Alice want to transfer $10 to Bob. After she initiates the transfer, someone needs to check that her account has enough money, then deduct $10 from it and add $10 to Bob's account. In traditional banking, this is done in a centralized database controlled by the bank. In a cryptocurrency wallet (we use bitcoin wallet here for illustration), there is no trusted third

[6] We still do not know the identity of Satoshi Nakamoto, which is considered the biggest mystery in bitcoin society.

[7] Chris Jaikaran - Senate Banking Committee - Senate.gov, https://www.banking.senate.gov/download/jaikaran-testimony-10-17-17pdf

party. The transaction instead is broadcasted to the whole network in the following format[8] (the address is generic):

"15N3yGu3UFHeyUNdzQ5sS3aRFRzu5Ae7EZ sent 0.00086 bitcoin to 1JHG2qjdk5Khiq7X5xQrr1wfigepJEK3t on August 8th, 2019, between 11:10 and 11:20 a.m".

Users will compete to validate the transaction for rewards through the mining process. More specifically, users will group transactions happening in a specific period together to form a block. Whoever wants to post the block will need to solve a computationally-intensive problem. For each period (now is 10 minutes for bitcoin), only the first one solving the problem can publish the block and claim all the rewards. There are variations of these rules in other chains. But the concept is similar. This is called "Proof of Work" consensus, which causes the majority of the confusion and criticizing of the blockchain. Due to the difficulty of the non-meaningful problems solved, the mining process significantly slows down the transaction speed and consumes enormous computational power and energy. For the first one, the Bitcoin blockchain can currently guarantee only 4.6 transactions per second (10 minutes per block), compared to Visa that is around 1,736 transactions per second (TPS). The bitcoin is too slow for everyday use. This is also called "poor scalability. For the second point, it is estimated that the global Bitcoin network is consuming more electricity than the country of Switzerland uses over the same time period" (ref). This brings us the question: why is proof of work is necessary for a public and permissionless chain.

The key point is that there is no trust among users. Therefore, the underlying assumption is that there will be malicious users trying to spam the network with fraudulent blocks and modify the existing chain to add self-benefiting transactions. The proof of work consensus makes sure that **creating blocks are expensive** (both timewise and energy cost-wise). So it is difficult and expensive to generate many blocks to flood the system. Also, anyone generating fake blocks, since the 10 minutes waiting period for each block, is likely to be caught and rejected by other users. Therefore, the users are only incentivized to spend their computational power on legit blocks so their work can be compensated. Besides, **the transactions within blocks are guarded by all the energy spent on mining blocks before them[9].** Therefore, the longer the chain, the more expensive for hackers to replace the current chain with a fake one. The Proof of Work consensus allows the truly distributed network, maximum pseudonymous of users, fairness among nodes, and none-tamperable of the chain, with the price of efficiency and cost.

Another popular, more recent consensus for the public and permissionless chain is Proof of State, which attributing mining power to the proportion of coins held by a miner[10]. This consensus significantly reduces the computational power and time laps required to add blocks. The logic is clear: those with more coins have less intention to sabotage the chain. On the other hand, a system where the major stakeholder enjoys extensive control and authority over both technical and economic aspects of the network creates a significant

[8] https://medium.com/@akadiyala/nuances-between-permissionless-and-permissioned-blockchains-f5b566f5d483

[9] https://www.devteam.space/blog/public-vs-private-permissioned-blockchain-comparison/

[10] http://www.coinfox.info/news/reviews/6417-proof-of-work-vs-proof-of-stake-merits-and-disadvantages

monopoly problem. In addition, block generators lose nothing by voting for multiple versions of the chains (nothing-at-stake" problem), Because of this, “some cryptocurrencies are vulnerable to Fake Stake attacks, where an attacker uses no or minimal stake to crash an affected node. " In addition, even with all the improvement, the transaction validation speeds for cryptocurrencies based on POS are still not on par with the traditional system.

For many business and organizations, blockchain is attractive as a record-keeping and sharing system. However, the slow transaction validating speed, the poor scalability, and lack of privacy of the data make the public and permissioned chain a poor choice. Thinking about healthcare record, for example, the network should neither be public (HIPPA protected), or permissionless (only authorized people are allowed to add record), In these cases, the private and permissioned chain is a good option.

The private and permissioned chain allows only authorized users, and users have different authorities in the chain. We will use the IBM Hyperledger Fabric as an example[11]. Hyperledger Fabric is a blockchain network that gets set by various collaborative organizations. The organizations that take part in the network are called "members." Each member will further decide the peers from her organization to involve in the blockchain. Peers have different authorizations. The Endorser peers can validate transactions and decide to approve or disapprove the transactions; the Anker peers are in charge of broadcasting updates; and the Ordered peers is in charge of creating the blocks and delivers to all the peers. You can see that in this network, there are intrinsic trust among peers (since they are from the same organization) and also trusts among members since they are collaborating. In this network, there is usually no incentive for malicious behaviors and only dedicated peers that organizations trust can validate transactions and add blocks. This consensus is called Proof of Authority (PoA), where a number of nodes are "authorities" in charge of validating transactions. The other **popular consensuses include** round-robin scheme, where users on the network take turns adding new blocks; and proof of Elapsed time (PoET), where each node is assigned a random waiting period and the first node to reach the elapsed time gets to create the next node.

The main criticism toward the private and permissioned chain is that it is not a real blockchain. It has multi-centralized control system (members), so it defeats the decentralization property of the blockchain. Many of these chains also do not need a token/coin to incentivize participants, which means that there is no mining. In these cases, the question becomes " why not just use a distributed database." This is a legit question. If a business has critical data that it wants to share internally, a combination of the current database, cloud, and identity management technologies will likely be adequate for its needs. However, if several organizations seek to collaborate and want to have the data to be immutable and auditable to avoid data discrepancy, then a blockchain is probably more appropriate and convenient — for example, a supplier-vendor relationship or an insurance company-clinic-patients relationship.

[11] https://medium.com/coinmonks/how-does-hyperledger-fabric-works-cdb68e6066f5

Between the two ends of the spectrum, there are the private and permissionless chain; and the public and permissioned chain. Neither of them has been widely implemented, and they are both mainly at conceptual levels. The private and permissionless chain would allow anybody in the network to submit and process transactions but would control who can be involved in the chain. It could be used to, for example, handle government records, compile research results from different teams, or scenarios that need to keep data's privacy. The public and permissioned chain would allow anyone to access the ledger, but only verified parties to submit, process, and validate transactions. The public and permissioned chains emphasize who can write in/regulate data. They can be applied to, for example, real estate registries, diploma checking system, or other scenarios that regulation or anti-forgery is needed.

Potential application of the four types of the blockchain is demonstrated in Table 1. We can further expand the categorization with partial permissions and a combination of public and private schemes. Why is this interesting? From the system design point of view, there are three major characters of the blockchains: **scalability, security, and decentralization**. Here scalability refers to the ledger's ability to handle growth, security the attack-resistance, and decentralization the transparent, synchronizing, and fairness among the network. The trade-off among these three is called the **Scalability Trilemma** (Figure 1), which means it is hard to maximize the other two without sacrificing the third. Blockchain mechanism design is to find the specific (usually hybrid) blockchain structure based on the requirement of the application. This goal motivates our lab to explore a better understanding of the blockchain as a complex system, quantifying and modeling the system's key features and performances, and simulating various structures of the blockchains for a structural mechanism design paradigm.

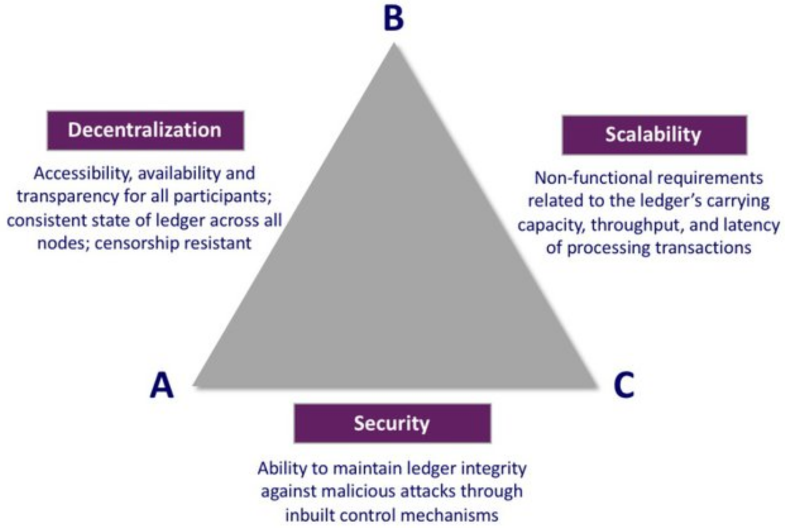
B
Decentralization
Accessibility, availability and transparency for all participants; consistent state of ledger across all nodes; censorship resistant
Scalability
Non-functional requirements related to the ledger's carrying capacity, throughput, and latency of processing transactions
A
C
Security
Ability to maintain ledger integrity against malicious attacks through inbuilt control mechanisms

Figure 1: Scalability Trilemma[12]

[12] https://steemit.com/blockchain/@reverseacid/the-scalability-trilemma

To close the talk, we discussed different kinds of blockchains based on their governance. We do want to point out that blockchain is not for everything. In many (if not most) cases, as we discussed before, a distributed database with access control is more than adequate, and this option is much faster and cheaper. So be careful with the blockchain hype and always ask yourself when you want to "blockchain the system," why is that necessary? What problems does it solve? Are there other options?

On the other hand, we are interested in the blockchain as a tool for big data and AI. It can potentially serve as the *Neural Network* that connects the data and AI by validating, auditing and sharing data safely and, if wanted, anonymously. It is especially promising when combining with IOT and wearable sensors to collect and distribute data automatically, and individuals in these case may not need to worry about the misuse of their data by big companies. We look forward to more solid business and social applications of blockchains in the future.