Survey of blockchain: principle, progress and application

ZENG Shiqin[1], HUO Ru[2,3], HUANG Tao[1,3], LIU Jiang[1,3], WANG Shuo[1,3], FENG Wei[4]

1. State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China

2. Beijing Advanced Innovation Center for Future Internet Technology, Information Department, Beijing University of Technology, Beijing 100124, China

3. Purple Mountain Laboratories, Nanjing 211111, China

4. Department of Information Technology Application and Software Services, Beijing 100846, China

Abstract

Blockchain is a kind of distributed ledger technology that upgrades to a complete storage system by adding logic control functions such as intelligent contracts. With the changes of its classification, service mode and application requirements, the core technology forms of Blockchain show diversified development. In order to understand the Blockchain ecosystem thoroughly, a hierarchical technology architecture of Blockchain was proposed. Furthermore, each layer of blockchain was analyzed from the perspectives of basic principle, related technologies and research progress in-depth. Moreover, the technology selections and characteristics of typical Blockchain projects were summarized systematically. Finally, some application directions of blockchain frontiers, technology challenges and research prospects including Smart Cities and Industrial Internet were given.

Key words: blockchain, cryptocurrency, decentralization, hierarchical technology architecture, technology diversity, industrial blockchain

1. Introduction

In 2008, Nakamoto came up with an idea for decentralized cryptocurrency, which is called Bitcoin. It marked the official birth of Bitcoin since the Bitcoin system was put into operation in 2009, and then Bitcoin has gradually entered the public view in the next few years. With the public statements of various countries on Bitcoin and the raising uncertainties of the world’s mainstream economy from 2016 to 2018, Bitcoin received a surge of attention and rapidly expanded demand. In fact, Bitcoin is one of the most successful applications of blockchain technology. With the emergence of open source blockchain platforms such as Ethereum and other decentralized applications, blockchain has been applied in more industries.

Due to the two characteristics of process credibility and decentralization, blockchain can build the foundation of trust in a low-cost way under the scenario involving multiple stakeholders, aiming to reshape the social credit system. In the past two years, blockchain has developed rapidly, and people try to apply it to finance, education, medical care, logistics, and other fields. However, problems such as waste of resources and inefficient operation restrict the development of blockchain. These factors have caused rapid changes in classification methods, service mode and application requirements of blockchain, and further led to the diversification of core technologies. Therefore, it is necessary to adopt a common structure to analyze the technology roadmap and characteristics of blockchain projects, so as to sort out and clarify the research direction of blockchain.

Blockchain covers a variety of technologies, so the related concepts are easy to be confused and there are numerous application scenarios. Thus, some relevant reviews have analyzed the latest progresses, technological differences and relationships of blockchain from the perspectives of technical architecture, technical challenges, and application scenarios. Meanwhile, they also summarized the technical form and application value. Yuan, et al. presented the basic model of blockchain, and divided the permissionless blockchain into data layer, network layer, consensus layer, incentive layer, contract layer, and application layer, taking Bitcoin as an example[1]. Shao, et al. compared the technical characteristics of various enterprise-level blockchain (permissioned blockchain) based on the details of open source projects[2]. Yang, et al. summarized the characteristics, challenges, and development trends of network service architecture based on blockchain[3]. Han, et al. systematically summarized the research status of blockchain security issues[4]. Ali, et al. summarized the research progresses and trends of the applications of blockchain in the Internet of Things. However, firstly, there is no general hierarchical structure analysis from the perspective of generic technologies of permissioned blockchain and permissionless blockchain. Secondly, these researches lacked of the relationship analysis between blockchain technology and existing technologies, such as networking and routing, data structure, and synchronization mechanism. And lastly, the difference analysis among blockchain projects is relatively scarce. In this paper, the relevant concepts are distinguished, the general hierarchical technical structure and its relevance with existing technologies are discussed, and the relevant academic research progress is analyzed horizontally. Meanwhile, we also compare the technology selection of part of the blockchain projects, according to the hierarchical structure. Finally, the application research status of blockchain is introduced on behalf of smart city, edge computing, and artificial intelligence technology, and the technical challenges and research prospects of blockchain are given.

2. Related concepts

With the in-depth study of blockchain technology, many related terms have been derived, such as "centralization", "decentralization", "public blockchain", and "consortium blockchain", et al. In order to have a comprehensive understanding of blockchain and a systematic understanding of the key terms involved in blockchain, this section will give the definitions of blockchain and its related concepts, as well as their relationships, thus we can distinguish the confusing terms much better.

2.1 Centralization and Decentralization

Centralization and decentralization were first used to describe the distribution characteristics of social governance power. From the perspective of blockchain application, centralization refers to the characteristics of the scenario that building the trust relationship with a single organization as the hub. For example, a user must complete the authentication, credit check, and transaction tracing through the information system of the bank in the electronic payment scenario. Or the verification of terminal identity must be completed by the digital certificate issued by the authoritative organizations in the e-commerce scenario. On the contrary, decentralization refers to the characteristics of the scenario that does not rely on a single organization to build trust. In this scenario, the importance of each organization is basically the same.

2.2 Cryptocurrency

Cryptocurrency is a type of digital currency technology. It uses a variety of cryptographic methods to process currency data, so as to ensure the anonymity of users and the effectiveness of value. Using trusted facilities to issue and check currency data ensures the controllability of the amount of money and the auditability of asset records. Therefore, cryptocurrency can make the monetary data become value exchange medium with circulation attributes, and protect the privacy of users.

The concept of cryptocurrency is originated from an anonymous trading technology based on blind signature[6]. The electronic cash[7] which is the earliest cryptocurrency trading model is shown in Figure 1.

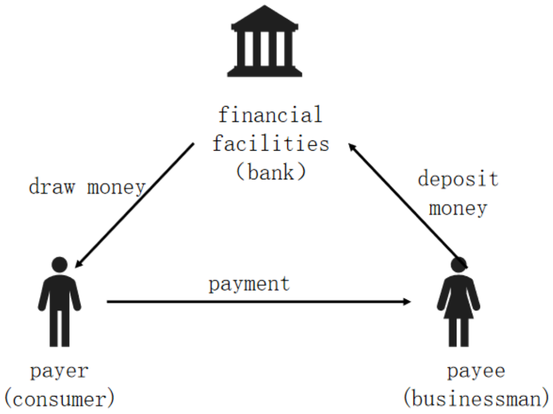

Figure 1: "electronic cash" trading model

Before the transaction starts, the payer uses the bank account to exchange the cryptocurrency, and then sends the currency data to the payee, who initiates a verification request to the bank. If the data is a legitimate currency data issued by the bank, the bank will charge an equal amount to the payee's account. Through the blind signature technology, the bank completes the authentication of the currency data, but cannot obtain the correlation between the issued currency and the received currency, furthermore, ensuring the anonymity of users and the effectiveness of value. Banks naturally have the ability to issue currency and account records, ensuring the controllability of the amount of money and the auditability of asset records.

The earliest ideas of cryptocurrency were centralized, with banks as the basis for building trust. Since then, cryptocurrencies have moved towards decentralization, with attempts to define value in terms of poof of work (PoW)[8] or its improved approaches. Based on the above work, bitcoin adopts a new distributed ledger technology to ensure that the data maintained by all nodes cannot be tampered with, thus successfully building a foundation of trust and becoming a decentralized cryptocurrency.

Blockchain has evolved from the decentralized cryptocurrency. With the further development of blockchain, decentralized cryptocurrency has become one of the main applications of blockchain.

2.3 Blockchain and Workflow

It is generally believed that blockchain is a new distributed computing and storage paradigm that integrates various existing technologies. It uses distributed consensus algorithm to generate and update data, uses peer to peer network for data transmission between nodes, combines cryptography and timestamp to ensure that the stored data cannot be tampered with, and then uses automated script code or intelligent contract to realize the upper application logic. If the traditional database realizes the unilateral maintenance of data, then the blockchain realizes the multi-party maintenance of the same data to ensure the security of data and fairness of business. The workflow of blockchain mainly includes three steps: block generation, consensus verification and ledger maintenance.

1) Block generation. Nodes collect transactions, namely the data items that need to be recorded, which broadcast over the blockchain network. Then package those transactions into blocks, which are the data sets with specific structure.

2) Consensus verification. The nodes broadcast the blocks to the network. After receiving a large number of blocks, the nodes of the whole network reach consensus on the sequence and verify the content to form a ledger, which is a set of blocks with specific structure.

3) Ledger maintenance. The nodes store the verified ledger data for a long time and provide functions such as backtracking verification and access interface for upper layer applications.

2.4 Types of Blockchain

According to the trust building methods in different scenarios, blockchain can be divided into permissionless blockchain and permissioned blockchain.

The permissionless blockchain, also known as a public blockchain, is a completely open blockchain in which anyone can join the network and participate in the full consensual ledger updating process. And there is no need to trust each other. The public blockchain establishes the trust relationship among the nodes of the whole network by means of consuming computing power, which has the characteristics of complete decentralization. While it also brings about problems such as waste of resources and low efficiency. Public blockchain is mostly used in unregulated, anonymous and free cryptocurrency scenarios like bitcoin.

The permissioned blockchain is a semi-open blockchain in which only designated members can join the network, and each member has a different right to participate. The permissioned blockchain usually establishes the trust relationship in advance by issuing the identity certificate, and has the characteristics of partial decentralization, which is more efficient than the permissionless blockchain. Furthermore, the permissioned blockchain is divided into the consortium blockchain and the fully private blockchain. The consortium blockchain is constructed by the consortium composed of many institutions, in which, the generation, consensus, and maintenance of the ledgers are respectively completed by the members designated by the consortium. When combining blockchain with other technologies to carry out scene innovation, the full openness and decentralization of the public blockchain is not necessary, since its low efficiency cannot meet the demand. Therefore, consortium blockchain becomes a blockchain model with stronger practical applicability in some scenes. Compared with the consortium blockchain, the fully private blockchain has a higher degree of centralization. The process of data generation, consensus and maintenance is completely controlled by a single organization, and the members designated by the organization only have the right to access the ledgers.

3. The Architecture of Blockchain

According to the current status of blockchain development, this section will summarize the general hierarchical technical structure, basic principles, and research progresses of blockchain.

Most of the technology selections of existing projects have evolved from Bitcoin. Therefore, blockchain is mainly based on peer-to-peer network communication and has a new basic data structure. It can realize the unification of public ledger data through the consensus of all nodes in the network. However, the blockchain also has problems such as low efficiency, high power consumption, and poor scalability. Therefore, some innovative consensus algorithms, processing models, and transaction mode are used to improve technical solutions, which enriches the logic control functions and applications of blockchain to make it a new type of computing mode. This article presents a general hierarchical technical structure of the blockchain as shown in Figure 2. From bottom to top, it includes network layer, data layer, consensus layer, control layer, and application layer. The network layer is the basis of information interaction for blockchain, undertaking the consensus process and data transmission between nodes, which mainly include the peer-to-peer network based on the basic network and its security mechanism. The data layer includes the basic data structure and principles of the blockchain. The consensus layer guarantees the consistency of node data, encapsulating various consensus algorithms and the reward and punishment mechanism that drives the consensus behavior of nodes. The control layer includes sandbox environment, automated scripts, smart contracts, and permissions management, etc., providing the programmable features of the blockchain to realize the control of the block data, the business data, and the organization structure. The application layer includes relevant application scenarios and practical cases of the blockchain, in which, the data interaction is carried out by calling the interface provided by the control contract. Since the application layer doesn't involve the principle of the blockchain, it is introduced separately in Section 5.

3.1 Network Layer

The network layer focuses on the basic communication method of blockchain network called peer-to-peer (P2P) network, which is a communicating and storage architecture of computer and different from the service mode of client-server. Each node in the network is both a data provider and a data user. The nodes share computing resources and information through the direct exchange. Therefore, the status of each node is equal. This layer is composed of networking structure, communication mechanism, and security mechanism. The networking structure describes the routing and topological relationships between nodes, the communication mechanism is used to realize the information exchange between nodes, and the security mechanism covers the peer security and the transmission security.

1) Networking Structure

The architecture of the peer-to-peer network can be divided into unstructured peer-to-peer networks, structured peer-to-peer networks, and hybrid peer-to-peer networks[9]. According to the logical topology of nodes, the structure of blockchain network can also be classified into three kinds as shown in Figure 3.

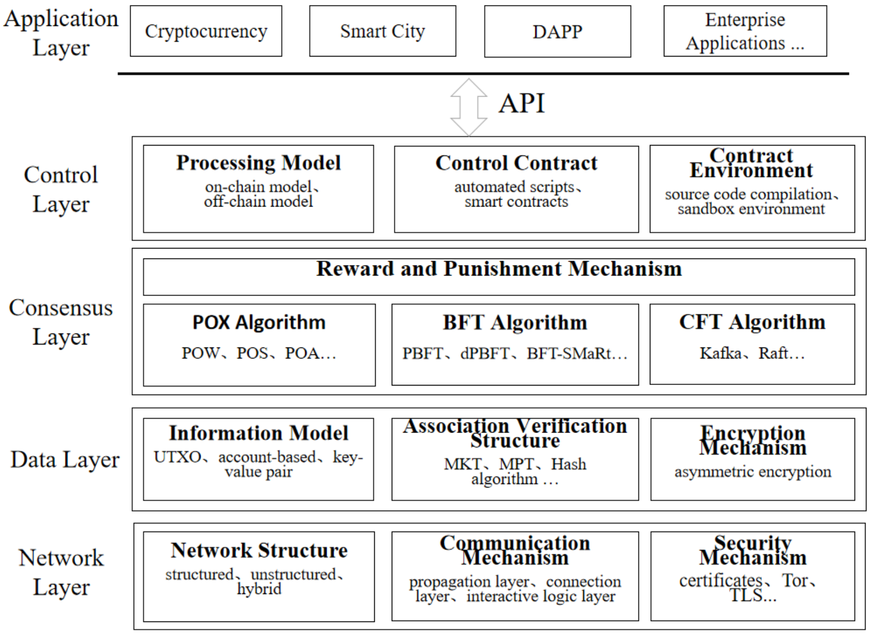

Figure 2: The hierarchical technical structure of blockchain

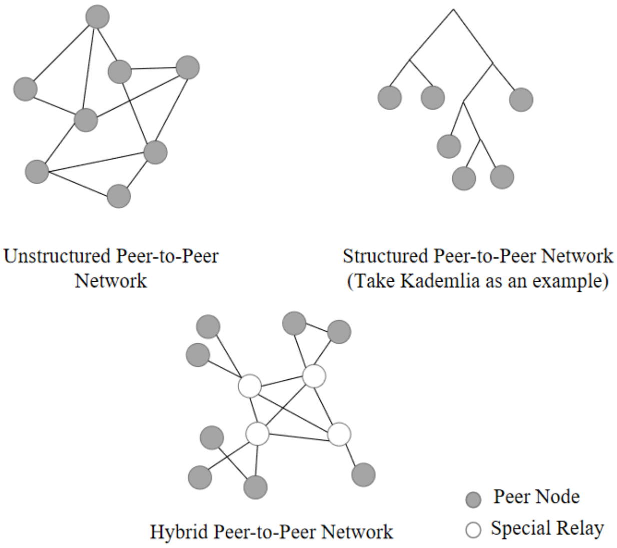
Unstructured Peer-to-Peer Network
Structured Peer-to-Peer Network (Take Kademlia as an example)
Hybrid Peer-to-Peer Network
Peer Node
Special Relay

Figure 3: The networking structure of blockchain

The unstructured peer-to-peer network means there are no special relay nodes, no definite rules for routing table generation of nodes, and random network topologies. This type of network has a loose, but simple design structure, with good fault tolerance and anonymity. However, its scalability is poor due to the flooding mechanism, such as Gnutella which is a typical protocol.

The structured peer-to-peer network means there are no special relay nodes, but the routing tables are generated between nodes according to specific algorithms, and the network topology has strict rules. This type of network is complex to implement but has good scalability, and can accurately locate nodes through structured addressing to achieve diversified functions. DHT (distributed hash table) network is the most common structured network, and the typical algorithms includes Chord and Kademlia.

The hybrid peer-to-peer network refers to a type of peer-to-peer network in which nodes implement message routing throughout the network through distributed relay nodes. Each relay node maintains part of the network node address, file index, and other tasks to jointly realize the function of data relay, such as Kazza which is a typical protocol.

2) Communication Mechanism

The communication mechanism refers to the peer-to-peer communication protocol between nodes in the blockchain network, which is based on TCP/UDP and is located in the application layer of the computer network protocol stack, as shown in Figure 4. This mechanism carries the specific interaction logic of the peer-to-peer network, such as node handshake, heartbeat detection, transaction and block propagation, etc. Due to the different protocol functions included (such as basic link and extended interaction), this article subdivides the communication mechanism into three levels: the propagation layer, the connection layer, and the interaction logic layer.

The propagation layer realizes the basic transmission of data between peer nodes, including two kinds of data dissemination modes, which are single-point dissemination and multi-point dissemination. The single-point dissemination mode means that data transfer directly between two known nodes, without going through other nodes for forwarding. And the multi-point dissemination mode means that the node who receives data forwards data to neighbor nodes by broadcasting mode. Blockchain network is generally based on the Gossip protocol[10] to achieve flood propagation. The connection layer is used to obtain node information, monitor and change the connection status between nodes, ensuring the availability of links between nodes. Specifically, the protocols of the connection layer help the newly added nodes to obtain data in the routing table, and maintain a stable connection for the node through timing heartbeat monitoring, and disconnects for the node in the case of a neighbor node failure. The interaction logic layer is the core of the blockchain network. From the perspective of the main process, the protocol of this layer carries the information interaction logic of the ledger data synchronization between peer nodes, the transaction and block data transmission, and the results feedback of data verification. In addition, it also provides message channels for complex operations and extended applications, such as node election and implementation of consensus algorithm.

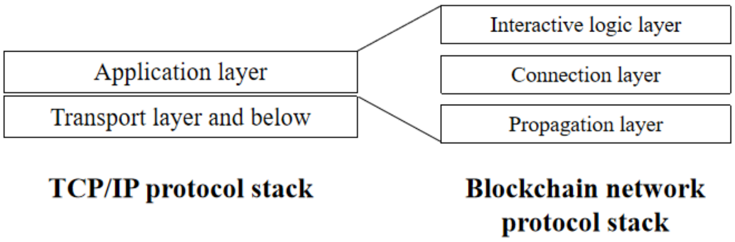

Figure 4: The network communication mechanism of blockchain

3) Security Mechanism

Security is an essential element of every system. The permissionless blockchain represented by Bitcoin uses the mechanisms of its data layer and consensus layer to ensure the consistency and validity of data by consuming computing power, without considering the data transmission process. Instead, it is built on an untrusted transparent P2P network. With the need of privacy protection, the permissionless blockchain also uses some anonymous communication methods. For example, the anonymous network Tor (the onion router) protects the peer identity through layers of data encryption along the path. The permissioned blockchain has higher requirements for the credibility of members, and appropriate security mechanisms are adopted at the network level, which mainly include identity security and transmission security. The identity security is the main security requirement of the permissioned blockchain to ensure end-to-end credibility. It is generally realized by digital signature technology to sign the full life cycle of the node (such as node interaction, voting, synchronization, etc.), so as to realize the admission of the permissioned blockchain. The transmission security prevents data from being tampered with or monitored during transmission. TLS-based point-to-point transmission and Hash algorithm-based data verification technology are often used.

4) Related Researches

At present, researches on the blockchain network layer is mainly concentrated in three directions: measurement optimization, anonymous analysis and privacy protection, and security protection.

With the explosive development of blockchain networks in recent years and the characteristics of open source, the academic community has begun to pay attention to the network status of large public blockchain projects, monitoring and studying their characteristics, and the research object mainly focuses on the Bitcoin network. Decker et al.[11] designed and implemented measurement tools, and analyzed the propagation delay data, the protocol data, and the address data. They also modeled and analyzed the factors of the network layer that affect the performance of Bitcoin network, based on which proposing their own optimization methods. Fadhil et al.[12] proposed a Bitcoin network simulation model based on event simulation, using real measurement data to verify the effectiveness of the model, and finally proposed an optimization mechanism named BCBSN, which aimed to set up super nodes to reduce network fluctuations. Kaneko et al.[13] divided the blockchain nodes into the consensus nodes and the verification nodes. The consensus nodes used the unstructured networking mode, and the verification nodes used the structured networking mode, which use the advantages of different networking methods to achieve network load balancing.

Anonymity is one of the important characteristics of cryptocurrency, but from the perspective of the network layer, the anonymity of the blockchain cannot be effectively guaranteed. Because attackers can infer the relationship between transactions and addresses by listening to and tracking IP addresses, and they can actively explore security risks and avoid potential hazards through anonymous privacy research. Koshy et al.[14] proposed a heuristic algorithm to identify the mapping relationship between Bitcoin addresses and IP addresses, and learned nearly 1,000 pairs of possible mapping relationships. Biryukov et al.[15] identified node identities by monitoring the address propagation information of the Bitcoin network, and then proposed a de-anonymization method for client. Venkatakrishnan et al.[16-17] modeled the Bitcoin network from three aspects: network topology, propagation layer protocol, and malicious model. Through theoretical analysis and simulation experiments, it is proved that the Bitcoin network protocol only had weak anonymity under the tree network structure. On this basis, the Dandelion network strategy was proposed to optimize anonymity with lower network overhead, and then the Dandelion++ principle is proposed to resist large-scale de-anonymity attacks with the guarantee of optimal information theory.

Blockchain focuses on its data layer and consensus layer mechanisms, and builds an open interconnected environment based on ordinary networks, which is extremely vulnerable to attacks. In order to improve the security of the blockchain network, academia has carried out research and provided corresponding solutions. Heilman et al.[18] implemented an eclipse attack on Bitcoin and Ethereum networks, which the feasibility of the attack was proved by shielding the correct node to completely control the information source of a specific node. Apostolaki et al.[19] proposed a BGP (border gateway protocal) hijacking attack against the Bitcoin network, which created node communication congestion by manipulating routes between autonomous domains or intercepting inter-domain traffic, indicating that attacks on key data along the way could greatly reduce block propagation performance.

3.2 Data Layer

The words "block" and "chain" are both used to describe the characteristics of the data structure of blockchain. It can be seen that the data layer is the core of the blockchain technology. The data layer defines the connection and structure of data in each node, and uses a variety of algorithms and mechanisms to ensure the strong correlation of data and the efficiency of verification, so that the blockchain has practical data tamper-proof characteristics. In addition, the behavior of each node to store entire data increases the risk of information leakage. Privacy protection has become an urgent demand, and the data layer realizes the anonymous protection of the application information through the cryptography principles such as the asymmetric encryption, which promotes the application popularization and the ecosystem construction of blockchain. As a result, the key technologies of the data layer can be divided into three categories: the information model, the association verification structure, and the encryption mechanism, considering the requirements of data relevance, validation efficiency and information anonymity.

1) Information Model

The blockchain carries the data of different applications (such as payment records, audit data, supply chain information, etc.), while the information model refers to the logical structure of the node recording application information, mainly including UTXO (unspent transaction output), the account-based model, and the key-value model. It should be noted that in most blockchain networks, each user is assigned a transaction address, which is generated by a pair of public and private keys, using the address to identify the user and verifing the validity of the transaction by means of digital signatures.

UTXO is the core concept in Bitcoin transactions and has gradually evolved into the main information model for blockchain applications in the financial field, as shown in Figure 5. Each transaction (Tx) is composed of the input data (Input) and the output data (Output). The output data includes the transaction amount (Num) and the address (Adr) that calculated from the user’s public key. The input data includes a pointer to the output data of the previous transaction (Pointer), which leads to the initial transaction of the Bitcoin that is issued to the node by the blockchain network.

The account-based information model stores data in the form of key-value pairs, maintains the current effective balance of the account, and continuously updates account data by executing transactions. Compared with UTXO, the account-based information model is similar to the bank savings accounts, which is more intuitive and efficient.

Both UTXO and the account-based information models are built on a more general key-value pair model. Therefore, in order to adapt to a wider range of application scenarios, the key-value pair model can be directly used to store business data in the form of a sheet or a collection. This model facilitates data access and supports more complex business logic, but it also has the problem of high complexity.

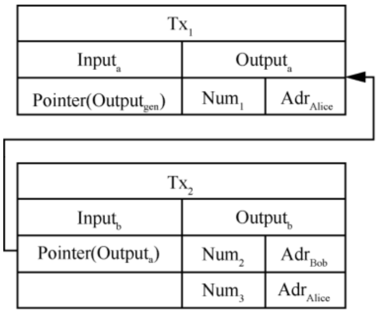

Figure 5: the information model of UTXO

2) Association Verification Structure

The anti-tampering feature of the blockchain benefits from the strong relevance of the data structure of the chain. This structure determines the binding relationship between data. When a certain data is tampered with, the relationship will be destroyed. Since the cost of forging this relationship is much higher than that of verifying it, the probability of tampering is reduced to an extremely low level. The basic data unit of the chain structure is "block", as shown in Figure 6.

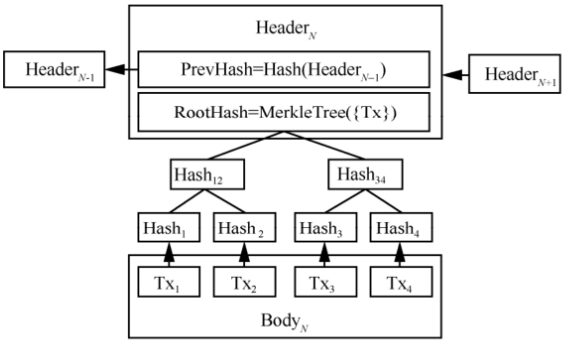

Figure 6: Basic block structure

A block is composed of two parts: a header and a body. The block body contains a certain number of transaction sets. The block header maintains its association with the previous block through prevhash to form a chain structure. It uses the root hash (RootHash) generated by MKT (MerkleTree) to quickly verify the integrity of the transaction set of the block body. Therefore, the hash algorithm and MKT are the key of the associated verification structure, which will be introduced in details below.

Hash algorithm is also called hash function, which realizes the irreversible mapping from the plaintext to the ciphertext. Meanwhile, the hash algorithm can change the input of any length to obtain a fixed-length output. Finally, even if the metadata has subtle gap, the output after the change will also make a significant difference. Utilizing the one-way, fixed-length and differential amplification characteristics of the hash algorithm, the node can determine the correctness of the previous block content by comparing the previous hash of the current block header, so that the chain structure of the block can be maintained. Commonly used hash algorithms in blockchains include SHA256 and so on.

MKT includes root hash, hash branch, and transaction data. MKT firstly hashes the transaction, then hashes these hash values in groups, and finally recursively step by step until the root hash. MKT brings many benefits. On one hand, the completeness of the root hash is determined, that is, the completeness of the transaction is confirmed indirectly, which improves efficiency. On the other hand, according to the hash path of the transaction (for example, Tx 1: Hash 2, Hash 34), the complexity of verifying the existence of a transaction can be reduced. In addition, there are other data structures used in conjunction with it, for example, Ethereum uses MPT (Merkle Patricia tree), which is combined with PatriciaTrie and MerkleTree structure, to efficiently verify its account-based information model data.

In addition, other information can be flexibly added to the block header according to different project requirements. For example, adding dimension of time by timestamp could form a time sequential record. Adding a identifier for accounting node could maintain the rights and interests. Adding the number of transactions could further improve the security of block data.

3) Encryption Mechanism

It can be seen from the above-mentioned encryption currency principle that the blockchain technology evolved by Bitcoin has inherent anonymity. Through asymmetric encryption and other technologies, both the user's privacy and the user's identity are verified. Asymmetric encryption technology refers to an encryption mechanism in which the encryptor and the decryptor use two different secret keys to complete encryption and decryption, and the secret keys cannot be deduced from each other. Commonly used asymmetric encryption algorithms include RSA, Elgamal, knapsack algorithm, Rabin, D-H, ECC (elliptic curve encryption algorithm), etc. As shown in Figure 5, Alice initiates a transaction Tx 2 to Bob. Alice uses Bob's public key to sign the transaction. Only when Bob uses the private key to verify the digital signature, he could have the right to create another transaction to make the currency he owns valid. This mechanism uses the public key as the basic identification of the user, making the user's identity unreadable and protecting privacy to a certain extent.

4) Related Research

The research direction at the data layer focuses on three aspects: the efficient verification, the anonymous analysis, and the privacy protection.

The academic problem of efficient verification stems from the authenticated data structure (ADS), that is, the use of a specific data structure to quickly verify the integrity of the data. In fact, MKT is also one of them. In order to adapt to the dynamical of blockchain data and maintain good performance, academia has carried out related researches. Reyzin et al.[20] proposed AVL+ based on the AVL tree structure, and simplified the block header verification process of lightweight nodes by balancing verification paths and default stack transaction sets. Zhang et al.[21] proposed the GEM2-tree structure and optimized it into the GEM2*-tree structure, which reduced the computing overhead of Ethereum nodes by decomposing the single tree structure, dynamically adjusting the node calculation speed, and expanding the data index.

Block data directly carries business information, so the anonymous correlation analysis of block data is more direct. Reid et al.[22] modeled block data as a transaction network and a user network, and successfully reduced network complexity by using user-directed analysis of multiple transaction data. Meiklejohn et al.[23] used heuristic clustering methods to analyze the flow characteristics of the transaction data and the group users, identifying the Bitcoin addresses of major institutions through interaction with these services. Awan et al.[24] used the dominant set method to automatically classify blockchain transactions to improve the accuracy of analysis.

In terms of privacy protection, Saxena et al.[25] proposed composite signature technology to weaken the relevance of data, based on the Diffie-Hellman assumption in the two-line mapping to ensure computational difficulty, thereby protecting user privacy. Miers et al.[26] and Sasson et al.[27] proposed Zerocoin and Zerocash, which broke the connection between transactions without adding a trusted party. Those works were the first that used zero-knowledge proof technology to hide the input and output of transactions, and amount information to improve the anonymity of Bitcoin. Asymmetric encryption is the core of blockchain data security, but it seems ordinary in the face of quantum computing. For this reason, Yin et al.[28] used the bonsai tree model to improve the lattice-based signature technology to ensure the randomness and security of public and private keys, and made the anti-quantum encryption technology suitable for the generation of user addresses on blockchain.

3.3 Consensus Layer

Each node in the blockchain network must maintain exactly the same ledger data. However, each node generates data at different times, and the source of data is unknown. There is a possibility that the node deliberately broadcasts the wrong data, which will lead to security risks such as Sybil Attack[29] and Double-Spending attacks[30]. In addition, data anomalies caused by node failures and network congestion are also unpredictable. Therefore, how to achieve network-wide unification of ledger data in an untrusted environment is a key issue that the consensus layer solves. In fact, the above-mentioned error of blockchain caused by the Byzantine Generals Problem[31], namely byzantine error. It means that independent components can perform arbitrary or malicious behaviors, and may cooperate with other error components. Such errors are widely studied in the field of trusted distributed computing.

State-machine replication is a commonly used theory to solve the fault tolerance problem of distributed systems. The basic idea is that any calculation is expressed as a state machine, which changes its state by receiving messages. Assuming that a group of replicas start with the same initial state and can agree on the order of a group of common messages, they can independently perform state evolution calculations, so as to correctly maintain the consistency between their respective replicas. Similarly, the blockchain also uses the state machine replication theory to solve the Byzantine fault tolerance problem. If the data of each node is regarded as a copy of the ledger data, then the transactions and blocks received by the node are the messages that cause the state of the copy to change. The realization of state machine replication theory and the maintenance of the consistency of replicas mainly include two elements, which are a deterministic state machine that correctly executes the calculation logic and a consensus protocol that propagates the same sequence of messages. Among them, the consensus protocol is the key to influence the effect of fault tolerance, the throughput, and the complexity. Besides, the consensus protocol required by systems with different security and scalability requirements is different. Academia generally distinguishes consensus protocols according to communication models and fault tolerance types[32]. Therefore, strictly speaking, the consensus protocol used by blockchain needs to solve the Byzantine fault tolerance problem under the partial synchrony model[33].

The blockchain network mainly includes the basic consensus protocols such as PoX (poof of X)[34], BFT (byzantine-fault tolerant) and CFT (crash-faulttolerant). The PoX protocol is a new consensus protocol driven by the reward and punishment mechanism represented by PoW (proof of work). In order to meet the needs of data throughput, resource utilization and security, people have proposed PoS (proof of stake) and PoST (proof of space-time) and other improvements. Their basic feature is to design proof basis so that honest nodes can prove their legitimacy, achieving Byzantine fault tolerance. BFT-type protocols refer to traditional consensus protocols and their improved protocols that solve the Byzantine fault tolerance problem, including PBFT, BFT-SMaRt, Tendermint, etc. CFT-type protocols are used to achieve crash tolerance, and avoid node evil through means such as identity certification, only considering the node or network crash (crash) failure, such as Raft, Paxos, and Kafka, etc.

There are differences in the degree of openness and the fault tolerance requirements of the permissionless blockchain and the permissioned blockchain, where the technologies on the consensus layer are extremely different. Specifically, the permissionless blockchain is completely open, and serious Byzantine risks need to be resisted, often using the protocols like PoX and BFT to achieve consensus in cooperation with the reward and punishment mechanisms. The permissioned blockchain has an access mechanism, and the identities of the nodes in the network are known, which reduces the Byzantine risk to a certain extent. Therefore, BFT-type protocols and CFT-type protocols can be used to build the same trust model[35].

Due to space limitations, this section only analyzes three types of protocols, which are PoW, PBFT, and Raft.

1) The Protocol of PoX class

PoW is also called the Nakamoto protocol, which is the core consensus protocol used by Bitcoin and its derivatives, as shown in Figure 7.

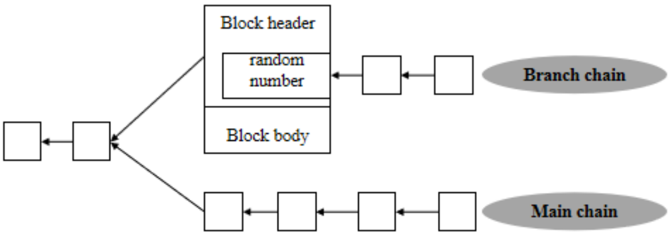

Figure 7: Schematic diagram of PoW protocol

The protocol adds a random number Nonce to the blockchain header structure, and designs proof basis as follows. In order to generate a new block, the node must calculate an appropriate Nonce value so that the newly generated block header is less than a certain threshold after double SHA256 operations. The overall process of the protocol is that each node in the entire network calculates the proof basis, and then the successfully solved node determines the legal block and broadcasts it. At the same time, the remaining nodes verify the legal block header, and if the verification is correct, it will form a chain structure with the local block and forward it. Finally, the nodes reached the consensus of the whole network. PoW is a random protocol. Any node may find a basis. The non-unique legal block will lead to the generation of a branch chain. At this time, the node selects the longest chain generated within a certain period of time as the main chain according to the "longest chain principle", and abandons the remaining branch chains, so that the data of each node finally converges.

The PoW protocol uses a random computing power election mechanism. The key achieving Byzantine fault tolerance lies in the struggle for the right of updating the ledger. The current method for finding proof is brute force searching, and its speed depends entirely on the performance of the computing chip. Therefore, when the number of honest nodes is more than half, that is when the "honest computing power" is over half, PoW can keep the legal branch chain at the fastest growth rate, which means that the main chain is always legal. PoW is a consensus protocol that relies on the saturation of computing power to correct Byzantine errors. It focuses on Byzantine fault tolerance in the process of block generation and propagation. While ensuring the prevention of double-spending attacks, it also has problems such as waste of resources and poor scalability.

2) The Protocol of BFT class

PBFT is a classic BFT consensus protocol, and its main process is shown in Figure 8. PBFT divides nodes into primary nodes and secondary nodes. The primary node is responsible for packaging transactions into blocks, and the secondary nodes participate in verification and forwarding. Assume that the number of malicious nodes is *f*. The PBFT consensus is mainly divided into three stages, which are the pre-preparation stage, the preparation stage, and the acceptance stage. The primary node firstly collects transactions and then sorts and proposes a legal block proposal. The remaining nodes verify the legality of the proposal, and then execute the results according to the order of transactions in the block, and finally multicast the digest of the result. Each node receives 2*f* digests that are the same as its own, and then multicasts to accept votes. When the node receives more than 2*f*+1 votes, it stores the block and its new state[36].

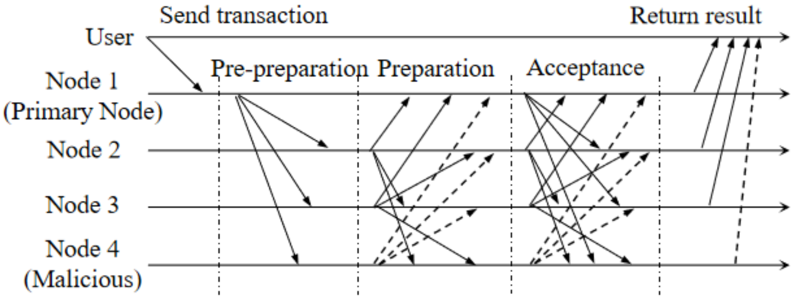

Figure 8: A schematic diagram of PBFT protocol

The PBFT protocol solves the Byzantine fault tolerance of the message propagation process. Because the algorithm complexity is $O(n^2)$ and the procedure of electing the primary node is deterministic, PBFT is only suitable for permissioned blockchain systems with a small number of nodes.

3) The Protocol of CFT class

Raft[37] is a typical crash-tolerant consensus protocol, known for its strong availability. Raft divides nodes into the follower nodes, the candidate nodes, and the leader nodes. The leader node is responsible for packaging transactions into blocks, the follower node responds to the synchronization instructions of the leader node, and the candidate node completes the election of the leader node. When the network is running stable, there are only the leader node and the follower node, and the leader node pushes block data to the follower node to achieve synchronization. The life time of those nodes is set to determine the changing cycle of the role. The survival time of the follower node is constantly reset by the heartbeat information of the leader node. When the leader node collapses, the follower nodes are automatically converted into the candidate nodes and enter the election process to realize self-recovery for the network.

The main idea of Raft protocol to achieve crash tolerance lies in the self-election mechanism of the leader node. Some permissioned blockchain chooses to reduce credibility requirements, and converts Byzantine fault tolerance to crash tolerance, thereby increasing the speed of consensus.

4) Reward and Punishment Mechanism

The reward and punishment mechanism includes incentive mechanism and punishment strategy. The incentive mechanism is a measure to make up for the node's computing power consumption and balance the profit ratio of the protocol operation. When the node can obtain the profit in the consensus process, it will compete for the accounting right, so the incentive mechanism uses economic benefits to drive the sustainable operation of consensus agreements. Incentive mechanisms are generally designed based on the theory of value equilibrium, and representative mechanisms include PPLNS, PPS, etc. In order to maximize revenue, nodes may adopt dishonest operating strategies (such as block deduction attacks, selfish mining, etc.), which damages the interests of honest nodes. Punishment strategies are based on game theory and other theories to punish nodes to correct misbehaving nodes to maintain consensus and sustainability.

5) Related Research

In addition to the traditional BFT, CFT protocol and PoX protocol derivatived researches, hybrid protocols (Hybrid) have also been proposed, with the increasing demands of the scalability and the performance. The hybrid protocol mainly includes the PoXs hybrid and the PoX-BFT hybrid. Therefore, this section summarizes the research progress of the consensus layer, represented by PoX, BFT and Hybrid protocols.

As mentioned above, the basic feature of PoX-type protocols is to design proof basis so that honest nodes can prove their legitimacy, thereby achieving Byzantine fault tolerance. uPoW[38] proves the legitimacy of nodes by calculating meaningful orthogonal vector problems, so that computing power is not wasted. PoI (proof-of-importance)[39] uses the principle of graph theory to assign importance weights to each node. The higher the weight, the more likely the node to calculate the block. PoS (poof-of-stake) defines the "token age" for nodes. Nodes with higher token age will be allocated more stake, and the stake are used as proof for the election of block nodes. Ouroboros[40] increased the randomness of elections by introducing a multi-party token toss protocol, and introduced a nearly Nash equilibrium incentive mechanism to further improve the security of PoS. PoRep (proof-of-replication)[41] is applied to a decentralized storage network, using the proof basis as a reward for contributing storage space to promote the reuse of storage resources.

The BFT protocol has a long history of development and has been given new vitality in blockchain research. SCP[42] and Ripple[43] are based on the federal byzantine consensus[44] that is a multi-pool (determined scale federation) consensus with intersection, which allows nodes to independently select or form consensus federations with designated nodes, respectively, to reach network wide consensus through federal intersection. Tendermint[45] uses the Gossip communication protocol to basically implement asynchronous Byzantine consensus, which not only simplifies the process but also improves usability. HotStuff[46] combines BFT with chain structure data, so that the primary node can push the agreement to reach consensus with actual network delay and $O(n)$ communication complexity. LibraBFT[47] adds a reward and punishment mechanism and a node replacement mechanism on the basis of HotStuff, thereby optimizing its performance.

Hybrid protocol is one of the research trends. PoA[48] uses PoW to generate empty block headers, and uses PoS to determine which nodes are responsible for accounting and endorsement.

The rewards are shared by endorsing nodes and block producers. PeerCensus[49] is implemented by a group of nodes to achieve consensus through the Byzantine agreement, while they must be based on Bitcoin network and obtain voting rights by generating blocks through PoW. ByzCoin[50] uses the computing power of PoW to construct dynamic membership, and introduces a joint signature scheme to reduce the round communication overhead of PBFT, and increase transaction throughput and reduce confirmation delay. Casper[51] uses PoS stakes to determine the formation of nodes which conduct a BFT consensus, while the right of vote also depends on stakes.

3.4 Control Layer

Blockchain nodes conduct block interaction based on the peer-to-peer communication network and the basic data structure, and achieve data consistency through consensus protocols, thus forming a unified ledger for the entire network. The control layer is the center of interaction between various applications and the ledger. If the ledger is compared to the database, then the control layer provides the database model, as well as the corresponding methods of encapsulation and operation. Specifically, the control layer consists of the processing model, the controlling contract, and the execution environment. The processing model analyzes and describes the differences in business/transaction processing methods from the perspective of the blockchain system. The controlling contract transforms the business logic into the specific operations of transactions, blocks, and ledgers. The execution environment encapsulates the common operating resources for nodes, so that the blockchain has stable portability.

1) Processing Model

The ledger is used to store all or part of the business data, so according to the distribution characteristics of the data, the processing model can be divided into two types that are the on-chain model and the off-chain model.

The on-chain model means that business data is completely stored in the ledger, and the business logic realizes data interaction through direct access to the ledger. The trust foundation of this model is established in a strongly related ledger structure, which not only realizes tamper-proofing but also simplifies the upper-level control logic. However, excessive resource consumption and huge data growth make the scalability of the system a bottleneck, so this model is suitable for business with small amount of data, strong security, decentralization and high degree of transparency.

The off-chain model means that business data is partially or completely stored outside the ledger, and only pointers or other data that prove the existence, authenticity and validity of the business data are stored in the ledger. The model takes "minimizing the cost of trust" as the criterion, and builds the trust foundation in the proof mechanism of the ledger and off-chain data to reduce the cost of ledger construction. Due to its decoupling from the public ledger, this model has good privacy and scalability, and is suitable for businesses with low decentralization, strong privacy, and high throughput.

2) Control Contract

The control contract in the blockchain has gone through two stages of development. The first one is a non-Turing complete automated script represented by Bitcoin, which is used to lock and unlock transactions based on the UTXO information model. It overcomes double spending problem together with the strongly related ledger, so that the transaction data has the circulation value. The second one is the Turing complete smart contract represented by Ethereum. The smart contract is a digital contract that is automatically executed based on ledger data. It is pre-defined by the developer according to the needs. It is the key to the upper-level application to compile business logic into a set of nodes and ledger operations. Smart contracts allow participants mutually untrusted to reach an agreement on the execution results of complex contracts without a trusted third party. It is conducive to the flexible definition of business logic and the expansion of blockchain application.

3) Execution Environment

The execution environment refers to the conditions required to execute the controlling contract, which is mainly divided into the native environment and the sandbox environment. The native environment means that the contract is tightly coupled with the node system, and is directly executed after source code compilation. In this way, the contract can undergo complete static analysis to improve security. The sandbox environment provides the necessary virtual environment for node operation, including network communication, data storage, and Turing complete computing/control environment, etc. The contract running in the virtual machine is convenient and flexible to update, and the vulnerabilities generated may also cause loss.

4) Related Research

The research direction of the control layer mainly focuses on scalability optimization and safety protection.

The side-chain builds a new classified asset chain outside the main chain of Bitcoin, and enables Bitcoin and other classified assets to be transferred between multiple blockchains, thereby dispersing the load of a single chain. Tschorsch et al.[52] proposed the Two-way Peg mechanism to implement interactive cross-chain asset conversion to prevent double spending problem. Kiayias et al.[53] used the NIPoPoW mechanism to implement non-interactive cross-chain proof of work and reduce the block redundancy caused by cross-chain. Sharding refers to different subsets of nodes processing different parts of the blockchain, thereby reducing the load on each node. ELASTICO[54] divides the transaction set into different shards, and each shard is verified in parallel by a different set of nodes. OmniLedger[55] optimizes node random selection and cross-slice transaction submission protocol on the basis of the former, thereby improving the security and correctness of the slicing consensus. Different from OmniLedger, PolyShard[56] uses Lagrangian polynomial coding fragmentation to add computational redundancy to the fragmentation interaction process, achieving the scalability optimization and the security guarantees. The above researches can be regarded as a scalability optimization solution for the on-chain processing model in the encrypted currency scenario. In fact, the off-chain processing model itself is an extensible optimization idea. The Lightning Network[57] uses the state channel to confirm the final result of the transaction on-chain, thereby realizing high-frequency off-chain payments during the transaction. Plasma[58] expands the blockchain by tree branches under the chain. The parent node in the tree branch completes the confirmation of the child node business until the root node and the blockchain are finally confirmed.

On the one hand, the sandbox environment carries the operating conditions of blockchain nodes, and the attacks against virtual machines are more direct. On the other hand, smart contracts directly operate on the ledger, and its loopholes are more likely to affect business operations. Therefore, the security protection of the control layer has become a hot spot direction. Luu et al.[59] analyzed the security of smart contracts running in the EVM and pointed out the security problems caused by the differences in the distributed semantics of the underlying platforms. Brent et al.[60] proposed Vandal, a smart contract security analysis framework, which converts EVM bytecode into semantic logic to facilitate the analysis of contract security vulnerabilities. Jiang et al.[61] predefined the characteristics for security vulnerabilities, and then simulate the execution of large-scale transactions, and realize vulnerability detection by analyzing the contract behavior in the log.

4. The Analysis of Technology Selection

Different from other technologies, the most striking feature of the blockchain is its close integration with the industry. With the rise of cryptocurrency and distributed applications, many blockchain projects have appeared in the industry. These projects are the concrete realization of blockchain technology, with both similarities and respective characteristics. This section will analyze the Bitcoin, Ethereum and Hyperledger Fabric according to the above hierarchical structure, and then briefly introduce other representative projects, and furthermore summarize and compare the technology selections and characteristics of each project.

4.1 Bitcoin

Bitcoin is by far the largest and most wide-ranging open source project of the permissionless blockchain. It shows the operation mode of Bitcoin project with ledger as the core in figure 9, which is also the prototype of all permissionless blockchain projects. The Bitcoin network provides users with exchange and money transfer services. In order to maintain the stability of the ledger and the authority of the data, the bitcoin formulates the incentive mechanism, that is, the ledger generates new bitcoins or users pay bitcoins for the nodes, so as to drive the nodes to jointly maintain the ledger.

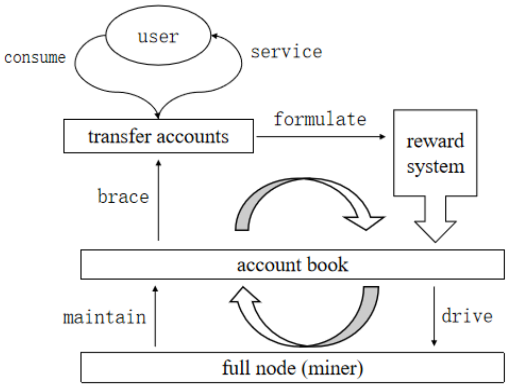

Figure 9: Bitcoin operation mode

The Bitcoin network is mainly composed of two types of nodes, which are full nodes and light nodes. Full nodes are fully functional blockchain nodes, while light nodes do not store complete ledger data and only have verification and forwarding functions. Full nodes are also known as miner nodes. At present, there are nearly 10,000 full nodes in the world. The pool is a pool of miners who rely on a reward allocation strategy to pool the computing power. In addition, there are also clients (wallets) that store private keys and address information and initiate transactions.

1) Network Layer

Bitcoins adopt an unstructured networking way in the network layer, and the routing table presents randomness. The nodes transmit data by means of multi-point propagation, which was implemented based on the gossip protocol, but now is changed to the diffusion protocol in order to improve the anti-anonymous analysis ability of the network[33]. Nodes use a series of control protocols to ensure link availability, including version capture (Vetsion/Verack), address capture (Addr/GetAddr), and heartbeat information (PING/PONG), et al. When a new node joins the network, the initial node list is first requested from the hard-coded DNS node (seed node). Then the routing table is generated by randomly requesting the node information in the routing table from the initial nodes. Finally, nodes establish connections with these nodes through control protocols and update the timestamps of nodes in the routing table according to the frequency of information interaction, so as to ensure that all nodes in the routing table are active. The interaction logic layer provides protocols such as GetBlock, MerkleBlock and CmpctBlock to build consensus interaction channels. Light nodes require only simple block Header validation, so a simple validation path can be established by specifying the block Header that needs to be validated through the Header Validation (GetHeader/Header) protocol and the Filter Setting protocol in the connection layer. In terms of security mechanism, the Bitcoin network can choose to use the anonymous communication network Tor as the data transmission bearer, and protect the identity of the opposite end through layer-by-layer data encryption mechanism along the path.

2) Data Layer

The technical selection of Bitcoin data layer has been extensively studied, using UTXO information model to record transaction data to archive simple and effective proof of ownership, and using MKT, hash function and timestamp to achieve efficient validation of blocks and generate strong correlation. In terms of encryption mechanism, Bitcoin adopts the elliptic curve digital signature algorithm (ECDSA, elliptic curve digital signature algorithm) to generate the user's keypair, and the wallet address is generated by public key through double hash, base58check encoding and other steps, which improves the readability.

3) Consensus Layer

Bitcoin uses PoW algorithm to achieve node consensus, which proves that the calculation difficulty can be changed according to the threshold set in the algorithm. The difficulty of the calculation is determined by the average number of blocks generated per hour, increases if the blocks are generated too quickly. This mechanism is designed to cope with the change of computing power caused by hardware upgrade or attention increase, and keep the proof basis effective all the time. At present, the threshold is set to produce a block in 10 min. In addition, Bitcoin uses a reward and punishment mechanism to ensure the sustainable operation of the consensus, including transfer fees, mining rewards and pool allocation strategies.

4) Control layer

Bitcoin initially use an on-chain processing model, with control statements recorded directly in the transaction, and automated5 locking/unlocking scripts used to validate bitcoin ownership in the UTXO model. Due to the limitations of extensibility and validation delay, Bitcoin generates multiple side-chain items such as Liquid, RSK, Drivechain, and off-chain processing items such as Lightning Network to optimize transaction speed.

4.2 Ethereum

Ethereum is the first open source platform project of programmable permissionless blockchain based on intelligent contract, which supports the use of blockchain network to build distributed applications, including finance, games and other types. When certain conditions are met, these applications will trigger intelligent contracts to interact with the blockchain network, so as to realize its network and storage functions. More importantly, more scenarios and value products will be derived. Such as CryptoKitties which will be used to assign value to the virtual cat with unique identification, Gitcoin and crowdfunding software development platform, etc.

1) Network Layer

The underlying peer-to-peer protocol cluster, called DEVP2P, meets the needs of any networked application associated with Ethereum in addition to the blockchain networking capabilities. DEVP2P uses the node public key as the identity, and uses Kademlia algorithm to calculate the node's xor distance, so as to realize structured networking. DEVP2P is mainly composed of three protocols: node discovery protocol (RLPx), basic communication protocol (Wire) and extended protocol (Wire-Sub). Multi-point propagation between nodes based on Gossip. When a new node is added, it first sends an access request to the hard-coded bootstrap node. Then the bootstrap node calculates and returns the list of nodes logically closest to the new node based on the Kademlia algorithm. Finally, the new node sends a handshake request to the nodes in the list, including network version number, node ID, listening port, etc. After establishing a connection with these nodes, the Ping/Pong mechanism is used to maintain the connection. The Wire sub-protocol constructs transaction acquisition, block synchronization and consensus interaction.Similar to Bitcoin, Ethereum designs light ethereum subprotocol (LES) and its variant PIP for lightweight wallet clients.In terms of security, the nodes use ECIES to generate public and private keys to transmit the Shared symmetric key in the process of establishing the connection with RLPx protocol. After that, the nodes encrypt the hosted data through the Shared key to realize data transmission protection.

2) Data layer

Ethereum maintains the correlation of blocks through hash function, and MPT is adopted to realize efficient verification of account state. The account-based information model records the user's balance and other ERC standard information. The account types are mainly divided into two categories: external account and contract account. External accounts are used to initiate trades and create contracts, and contract accounts are used to create trades during contract execution. The generation of the user's public and private keys is the same as that of Bitcoin, but the public key is calculated by kecak-256 algorithm.

3) Consensus Layer

Ethereum sets the threshold as 15s to produce a block. It is planned to adopt PoS or Casper consensus algorithm in the future. Lower computational complexity leads to frequent branching chains, so Ethereum uses its own reward and punishment mechanism, the GHOST protocol, to encourage consensus among miners. Specifically, the hash values in a block are divided into a parent block hash and a uncle block hash, where the parent block hash points to the precursors, and the uncle block hash points to the precursors of the parent block. When the new block is generated, GHOST calculates the miner reward according to the parent/uncle hash value of the previous 7 generations of blocks, which to some extent makes up for the wasted computing power when the branch chain is discarded.

4) Control Layer

Each node will have a sandbox environment EVM for the execution of an intelligent contract written in Solidity. Instead of trying to keep down the body of water inside the body, the language is meant to allow users to easily define their own business logic, which is a prerequisite for the development of many distributed applications. In order to optimize the extensibility, Ethereum owns the side-chain project Loom and the off-chain project Plasma. The sharding technology has joined the source code of Ethereum in 2018.

4.3 Hyperledger Fabric

Hyperledger is an open source blockchain project under the Linux Foundation, which aims to provide cross-industry blockchain solutions. Fabric is one of the sub-projects of super ledger, and also the enterprise-level programmable permissioned blockchain project with the most extensive influence. Within known solutions, Fabric is applied in a variety of scenarios such as supply chain, healthcare and financial services.

1) Network layer

Fabric network constructs the node cluster based on the organization unit and adopts the hybrid peer-to-peer network to form the network. Each organization includes common peer, which completes message routing within the organization, and an anchor peer, which is responsible for node discovery and message routing across the organization. The Fabric network propagation layer is based on gossip, which needs to be initialized by configuration file. After the network is generated, each node will regularly broadcast the survival information, and other nodes will update the routing table according to the information to maintain the connection. The interaction logic layer adopts multi-channel mechanism, that is, only nodes in the same channel can conduct state information interaction and block synchronization. Fabric belongs to permissioned blockchain thus it has a stringent security mechanism at the network layer. Such as PKI-ID for authentication, optional TLS bidirectional encrypted communication and so on.

2) Data Layer

The read-write set in the block of Fabric describes the process of reading and writing when a transaction is executed. The read-write set is used to update the state database, which records key-value pairs of versions and values, and therefore belongs to the key-value pair information model. On the one hand, hash function and Merkle Tree are used as efficient implementation techniques for associative structures. On the other hand, the node also needs to verify that the state database is consistent with the latest version in the read-write set based on the key value. Permissioned blockchain scenario require less anonymity, but higher privacy requirements for business data, so Fabric1.2 version began to provide the private data set (PDC, private data collection) function.

3) Consensus Layer

Fabric using PBFT consensus agreement before version 0.6. However, in order to improve the throughput, Fabric1.0 choice to reduce the security, in which the consensus process is decomposed into ordering and validation services. Ordering service is completed by CFT protocol Kafka and Raft (after v1.4), while the verification service is further decomposed into read-write set verification and multi-signature verification to maximize the speed of consensus. Because Fabric is specific to the permissioned blockchain scenario, participants are often known and have the same intention to cooperate, so the assumption of node sabotage and malicious is avoided, and no adjustment of reward and punishment mechanism is required.

4) Control Layer

Fabric has less demand for scalability optimization, which mainly benefits from the optimization of consensus layer and the premise that the permissioned blockchain itself has fewer participating nodes. Therefore, it mainly adopts the chain processing model to facilitate the access of business data. However, PDC only stores private data hash on chains so it belongs to the off-chain processing model, and the smart contract can access the data locally. Fabric node adopts modular design and builds module execution environment based on Docker. Smart contract, known as chaincode in Fabric, are written in GO, Javascript and Java languages, which are also Turing-complete.

4.4 Other Projects

In addition to the three blockchain platforms mentioned above, there are many representative projects in the industry, as shown in Table 1.

5. Application Research of Blockchain

Blockchain can help reduce the audit cost between financial institutions, significantly improve the processing speed and efficiency of payment business, and can be applied to financial scenarios such as cross-border payment. In addition, blockchain is also used in non-financial scenarios such as property rights protection, credit system construction, ecological optimization of education, food safety supervision and network security.

According to the application mode of these scenarios and the features of blockchain, the characteristics of blockchain can be summarized as follows. 1) Decentralization. A node establishes communication and trust endorsement based on a peer-to-peer network, and the failure of a single node won’t affect the overall situation. 2) Non-repudiation. The ledger is maintained by all nodes. Moreover, the consensus process of group cooperation and the strongly correlated data structure ensure that the node data is consistent and basically cannot be tampered with, further making the data verifiable and traceable. 3) Openness and transparency. Except for private data, the data on chain is exposed to every node to facilitate the verification of existence and authenticity of data. 4) Anonymity. Multiple privacy mechanisms allow users to hide their identities, even as they build a foundation of trust. 5) Contract autonomy. Predefined business logic enables nodes to be autonomous based on highly trusted ledger data, automating the execution of business between person-person, person-machine and machine-machine interactions.

Since the applications of above fields have been described in detail in previous studies, this paper will mainly introduce current research of the frontier application of blockchain in the fields of smart city, edge computing and artificial intelligence.

Table 1: Representative Blockchain Projects

Technical roadmap	Corda	Quorum	Libra	Blockstack	Filecoin	Zcash
Control contract	Kotlin, Java	GO	Move	Clarity non-turing complete	Non-Turing complete	Non-Turing complete
Execution environment	JVM	EVM	MVM	Source code compilation	Source code compilation	Source code compilation
Processing model	on-chain	on-chain/off-chain（private data）	on-chain	off-chain（virtual chain）	off-chain（IPFS）	on-chain
Incentive mechanisms	—	—	Libra coins	Stacks token	Filecoin	Zcash/Turnstiles
Consensus algorithm	Notary/RAFT，BFT-SMaRt	Quorum-Chain, RAFT	LibraBFT	Tunable Proofs, proof-of-burn	PoRep, PoET	PoW
Information model	UTXO	Account-based	Account-based	Account-based	Account-based	UTXO
Association verification	Hash algorithm	Hash algorithm	Hash algorithm	Hash algorithm	Hash algorithm	Hash algorithm

Structure	MKT	MPT	MKT	Merklized Adaptive Radix Forest (MARF)	MKT	MKT
Encryption mechanism	Tear-offs, Mixed key	Based on Enclave	SHA3-256/ EdDSA	Based on Gaia/ BlockstackAuth	SECP256K1/BLS	zk-SNARK
Networking Method	Hybrid	Structured	Hybrid	Unstructured	Structured / Unstructured	Unstructured
Communication Mechanism	AMQP1.0/ Unicast	Wire/ Gossip	Noise-Protocol-Framework/Gossip	Atlas/Gossip	Libp2p/Gossip	Bitcoin-Core/Gossip
Security Mechanism	Corda encryption suite /TLS	certificate /HTTPS	Diffie-Hellman	Secure Backbone	TLS	Tor
Blockchain type	Permissioned Blockchain	Permissioned Blockchain	Permissioned Blockchain	Permissionless Blockchain	Permissionless Blockchain	Permissionless Blockchain
Features	Only allow information access and	Based on the Ethereum network to	Stable and fast trading network	Excluding the central service provider's scalable	Incentive-driven storage resource sharing ecology	Privacy protection based on
Application scenarios	Financial business platform	Distributed application	Cryptocurrency	Internet infrastructure	File storage and sharing	Cryptocurrency

5.1 Smart City

Smart city refers to the research field that utilizes ICT to optimize the utilization effect of public resources, improve the quality of life of residents and enrich the informatization capacity of facilities. This field includes specific scenarios such as personal information management, smart medical treatment, smart transportation and supply chain management. Smart cities emphasize the collection, analysis and enabling of all kinds of data, such as residents and facilities. However, demands of data reliability, transparent management and incentive sharing have brought many technical challenges to smart cities. The decentralized interaction mode of blockchain avoids a single point of failure, and also improves the fairness of management. Furthermore, open and transparent ledgers ensure the reliability and traceability of data and various anonymous mechanisms are conducive to the protection of residents' privacy.

Therefore, blockchain is conducive to the solution of the problem. Hashemi et al in [62] used blockchain for permission data storage to build a decentralized personal data access control model. Bao et al in [63] utilize the blockchain for efficient authentication and user identity management to protect the owner's identity, location, vehicle information and other personal data.

5.2 Edge Computing

Edge computing is a distributed information service architecture that migrates computing, storage and network resources from cloud platform to network edge. It attempts to deeply integrate traditional mobile communication network, Internet and Internet of Things services, reduce end-to-end delay of business delivery, and improve user experience. Security is a major technical challenge faced by edge computing. On the one hand, a large number of heterogeneous terminal devices are used in edge computing hierarchy to provide user services, which may generate malicious behaviors. On the other hand, data integrity and authenticity need to be guaranteed during service migration. Blockchain may play a key role in this complex work environment and open service architecture. First, blockchain can build a tamper-resistant ledger in a loose network of devices at the bottom of edge computing, providing a basis for verifying device identity and service data. Second, devices can achieve a high degree of autonomy with the help of intelligent contracts, providing a foundation for trusted interoperability of devices for edge computing. Samaniego in [64] proposed a virtual Internet of Things resource migration architecture based on blockchain to share resource data so as to ensure security. Stanciu in [65] proposed a distributed security cloud architecture combining software defined network (SDN), fog computing and blockchain technology to solve the security distribution problem of flow table strategy of SDN controller in fog nodes. Ziegler in [66] proposed an extensible application scheme of blockchain in fog computing scenario based on Plasma framework to improve the security of fog computing gateway.

5.3 Artificial Intelligence

Artificial intelligence is the research of a kind of intelligent agent, which enables machines to perceive the environment/information and make correct decisions, which means to achieve certain goals predetermined by human beings. The key to artificial intelligence lies in algorithms, while most of the machine learning and deep learning algorithms are built on large data sets and centralized training models. This method is easy to be attacked or maliciously manipulated, which results in the unreliability of the model and the waste of computing power. In addition, the safety of downstream devices and the authenticity and integrity of data sources cannot be ensured during the data acquisition process. Moreover, these consequences will be magnified in scenarios such as automatic driving. Non-repudiation of blockchain can realize the credibility of perception and training process. In addition, decentralization and contract autonomy lay a foundation for the decomposition and decentralization of artificial intelligence training work, and improve computing efficiency on the basis of security. Kim in [67] used blockchain to verify the integrity of the distribution model under the joint learning framework, and provided corresponding incentives according to the calculation cost to optimize the overall learning effect. Bravo-marquez in [68] proposed the consensus mechanism "learning proof" to reduce the computational waste of PoX consensus and build a public verifiable learning model and experimental database.

6. Technical Challenges and Research Prospects

6.1 Hierarchical Optimization and Deep Integration

A paradox of ternary exists in blockchain, that is, security, extensibility and decentralization cannot be achieved at the same time, and the effects of one can only be sacrificed to satisfy the needs of the other two. The public blockchain represented by Bitcoin is characterized by high security and complete decentralization, but problems such as resource waste become the bottleneck of expansion optimization. Although some common protocol optimization schemes, such as PoS and BFT, or on-chain processing models, such as side chain and splicing, or down-chain extension schemes, such as Plasma and lightning network, have emerged successively, they are all at the cost of partial security or de-centralization. Therefore, how to push the blockchain to practical application depends largely on the solution of the paradox of ternary, among which there are mainly two kinds of ideas.

1) Hierarchical Optimization

Each layer in the blockchain hierarchical structure affects the above three characteristics to varying degrees, such as network delay, parallel read-write efficiency, consensus speed and security of on-chain/off-chain interaction mechanism, etc. The optimization of blockchain should be considered as a whole rather than a single layer.

The main defect of network layer is security, while the scalability also needs to be optimized. How to defend against network attacks represented by BGP hijacking will become the security research direction of the underlying network of blockchain [19]. The Information-Centric Networking (ICN) will reshape the basic transmission network of blockchain, reducing redundant traffic in the network and accelerating communication transmission through request aggregation and data caching[69]. Compared with the data layer and the consensus layer, the blockchain network has less attention, but it is the basic factor affecting the security and scalability.

The optimization space of the data layer lies in the high efficiency, mainly for the design of new data verification structures and algorithms. This direction can refer to a variety of data structure theories and complexity optimization methods in the field of computer research to find a structure suitable for blockchain calculation, or even design a new data association structure. In fact, quite a few projects borrow from the idea of chain structure to open up new paths, such as the Segregated Witness (segwit) which compressed block space, or the Tangle structure of parallel associations in a Directed Acyclic Graph (DAG), or the state tree adopted by the Libra project.

Consensus mechanism is the focus of current research, and it is also the most difficult equilibrium layer that affects the ternary property simultaneously. PoW achieves complete decentralization and security at the expense of extensibility. The efficient block-out mode of PoS has extensibility but generates bifurcating problem. PoA combines the two to achieve the balance of the three characteristics. Based on this, Hybrid consensus combined with the dynamic adjustment of the reward and punishment mechanism achieves a good effect and becomes a transitional means of consensus research. However, it remains to be studied how to achieve the real breakthrough of the paradox of ternary.

The control layer is currently the focus of scalability research. Its advantage is that it does not need to change the underlying basic implementation, but can be applied in a short period of time and concentrated in blockchain projects in the industry. The side chain has good flexibility but high operational complexity, the sharding improves the ledger structure, but the security problem of cross-sharding interaction always exists, while the downlink processing model lacks the support of theoretical analysis in security. Therefore, the solution of the ternary paradox has a wide range of research prospects in the control layer.

2) Deep Fusion

If hierarchical optimization is called horizontal optimization, then deep fusion is longitudinal optimization based on scene requirements. On the one hand, the ternary requirements of different scenarios are not the same. For example, access control does not require complete decentralization, and its scalability does not encounter a bottleneck. Therefore, BFT algorithm can be used to build consortium blockchain in a small scope. On the other hand, blockchain application research has changed from simple data on-chain to off-chain storage and on-chain verification, and consensus algorithm changes from POW to scenario based service proof and learning proof. In addition, 5G and edge computing can help blockchain to move the network and computing functions to the edge of the network for saving terminal resources. This means that under the strict scene modeling, the hierarchical technology selection of blockchain will be cross-innovated and deeply integrated with the features of the scene, so it has a relatively broad research prospect.

6.2 Privacy Protection

Cryptocurrencies are known for their anonymity, but blockchain anonymity based on asymmetric encryption is constantly being challenged. Anti-anonymous attack has changed from identity decryption to behavior clustering analysis, including not only IP clustering of network traffic, but also address clustering of transaction data and heuristic model learning of transaction behavior. Therefore, the development of big data analysis technology has changed the idea of blockchain privacy protection. Tor network, mixed currency technology, zero knowledge proof, homomorphic encryption and various kinds of asymmetric encryption algorithms with higher complexity have been proposed, but each method still has limitations, and more efficient methods will be needed in the future. In addition, with the programmable development of blockchain system, the internal complexity will be increasingly high, especially the smart contract needs more strict and effective code detection methods, such as anonymity detection and privacy threat warning, etc.

6.3 Industrial Blockchain

Industrial blockchain refers to the application scenario that uses blockchain to consolidate the basis of data circulation and control in the Industrial Internet and promote value transformation, which has a great research prospect.

The Industrial Internet is an important infrastructure for the digital, networked and intelligent demand of manufacturing, supporting the ubiquitous connection, flexible supply and efficient allocation of manufacturing resources. It is also a service system based on massive data collection, aggregation and analysis. "Industrial Internet Platform" is the core of the Industrial Internet. With the logic closed-loop composes of all-round perception, real-time analysis, scientific decision-making and precise execution, Industrial Internet realizes a new model and a new form of Industrial with the characteristics of whole industrial elements, industrial chain and value chain.It can be seen that the Industrial Internet is intrinsically related to the Internet of Things, Smart City, Consume Internet and other scenario applications, such as ubiquitous connectivity, data sharing and analysis, e-commerce, etc., so Its academic issues and realization of technology are also inevitably related to those fields. Blockchain solves the single point of failure in the central management and control architecture of the Internet of Things, overcomes the security and privacy challenges of ubiquitous perceptive device data, provides solutions for data sharing and access control of smart city scenes, and builds a new Internet value ecology for encouraging resource sharing. Although the industrial Internet, as a new industrial ecosystem, has a more complex technical system and richer connotation, it is not difficult to imagine that blockchain is also conducive to the development of the industrial Internet.

"Platform + Blockchain" can reduce the management cost of data storage, processing and use through distributed data management mode, build a more reliable environment for industrial users to choose and use industrial apps, and realize identity authentication, operation behavior tracing, safe storage and reliable delivery of data. Effective supply chain total factor traceability and collaborative services can be realized by uploaded data to blockchain such as product design parameters, quality test results, order information, etc. It can promote data trading and business collaboration among platforms, realize cross-platform transaction settlement, drive data sharing and knowledge reuse among platforms, and promote the interconnection and interworking among industrial Internet platforms.

Of course, industry is an industry related to the national economy and the people's livelihood. It is not advisable to use the characteristics of the blockchain such as decentralization and anonymity directly on the industrial internet. Therefore, it is necessary to study the management framework of industrial blockchain to realize the manageability and control of industrial blockchain. In this way, blockchain can give full play to its security advantage within a certain range, and provide positive incentives for the operation of Industrial Internet.

7. Conclusion

Based on multiple technologies, blockchain solves the problems of trust construction and privacy protection in the complex production environment involving multiple organizations at a low cost. It has been widely applied in finance, education, entertainment, copyright protection and other scenarios, and has become a research hotspot in the academic community. Bitcoins appeared to reshape the definition of value and blockchain have got rapid development along with industry calls. With the method of hierarchical analysis, we can intuitively distinguish the technical roadmap and characteristics of each project. The method can also provide different perspectives for optimizing blockchain, create conditions for the deep integration of scene application, and promote the follow-up research. In the future, blockchain will become a trust infrastructure and develop in a healthier way in broader fields such as Industrial Internet.

References:

[1] YUAN Y , WANG F Y. Blockchain: the state of the art and future trends[J]. Acta Automatica Sinica, 2016, 42(4): 481-494.

[2] SHAO Q F, ZHANG Z, ZHU Y C, et al. Survey of enterprise blockchains[J]. 2019, 30(9): 2571-2592.

[3] YANG W L, AGHASIAN E, GARG S, et al. A survey on blockchain-based Internet service architecture: requirements, challenges, trends, and future[J]. IEEE Access, 2019, 7: 75845-75872.

[4] HAN X, YUAN Y, WANG F Y. Security problems on blockchain: the state of the art and future trends[J]. Acta Automatica Sinica, 2016, 45(1): 208-227.

[5] ALI M S, VECCHIO M, PINCHEIRA M, et al. Applications of blockchains in the Internet of Things: a comprehensive survey[J]. IEEE Communications Surveys & Tutorials, 2019, 21(2): 1676-1717.

[6] CHAUM D. Blind signature system[M].Berlin: Springer, 1984.

[7] LAW L, SABETR S, SOLINAS J. How to make a mint: the cryptography of anonymous electronic cash[J]. The American University Law Review, 1996, 46(4): 1131-1162.

[8] JAKOBSSON M, JUELS A. Proofs of work and bread pudding protocols(extended abstract)[M]. Berlin: Springer, 1999.

[9] WANG X L, ZHANG J. Survey on peer-to-peer key technologies[J]. Application Research of Computers, 2010, 27(3): 801-805.

[10] DEMERS A, GREENE D, HOUSER C, et al. Epidemic algorithms for replicated database maintenance[J]. ACM SIGOPS Operating Systems Review, 1988, 22(1): 8-32.

[11] DECKER C, WATTENHOFER R. Information propagation in the bitcoin network[C]//IEEE Thirteenth International Conference on Peer-to-Peer Computing. Piscataway: IEEE Press, 2013: 1-10.

[12] FADHIL M, OWENSON G, ADDA M. Locality based approach to improve propagation delay on the Bitcoin peer-to-peer network[C]//2017 IFIP/IEEE Symposium on Integrated Network and Service Management. Piscataway: IEEE Press, 2017: 556-559.

[13] KANEKO Y, ASAKA T. DHT clustering for load balancing considering blockchain data size[C]//2018 Sixth International Symposium on Computing and Networking Workshops. Piscataway: IEEE Press, 2018: 71-74.

[14] KOSHY P, KOSHY D, MCDANIEL P. An analysis of anonymity in bitcoin using P2P network traffic[C]//Financial Cryptography and Data Security: 18th International Conference. Berlin: Springer, 2014: 469-485.

[15] BIRYUKOV A, KHOVRATOVICH D, PUSTOGAROV I. Deanonymisation of clients in bitcoin P2P network[C]//ACM SIGSAC Conference on Computer and Communications Security. New York: ACM Press, 2014: 15-29.

[16] BOJJA VENKATAKRISHNAN S, FANTI G, VISWANATH P. Dandelion[J]. ACM SIGMETRICS Performance Evaluation Review, 2017, 45(1): 57.

[17] FANTI G, VENKATAKRISHNAN S B, BAKSHI S, et al. Dandelion++: lightweight cryptocurrency networking with formal anonymity guarantees[J]. ACM SIGMETRICS Performance Evaluation Review, 2018, 46: 5-7.

[18] HEILMAN E, KENDLER A, ZOHAR A, et al. Eclipse attacks on Bitcoin’s peer-to-peer network[C]//USENIX Conference on Security Symposium. Berkeley: USENIX Association, 2015: 129-144.

[19] APOSTOLAKI M, ZOHAR A, VANBEVER L. Hijacking bitcoin: routing attacks on cryptocurrencies[C]//2017 IEEE Symposium on Security and Privacy. Piscataway: IEEE Press, 2017: 375-392.

[20] REYZIN L, MESHKOV D, CHEPURNOY A, et al. Improving authenticated dynamic dictionaries, with applications to cryptocurrencies[M]//Financial Cryptography and Data Security. Cham: Springer International Publishing, 2017: 376-392.

[21] ZHANG C, XU C, XU J L, et al. GEM^2-tree: a gas-efficient structure for authenticated range queries in blockchain[C]//2019 IEEE 35th International Conference on Data Engineering. Piscataway: IEEE Press, 2019: 842-853.

[22] REID F, HARRIGAN M. An analysis of anonymity in the bitcoin system[C]//2011 IEEE Third International Conference on Privacy, Security, Risk and Trust. Piscataway: IEEE Press, 2011: 1318-1326.

[23] MEIKLEJOHN S, POMAROLE M, JORDAN G, et al. A fistful of bitcoins: characterizing payments among men with no names[C]//The 2013 Conference on Internet Measurement Conference. New York: ACM Press, 2013: 127-140.

[24] AWAN M K, CORTESI A. Blockchain transaction analysis using dominant sets[C]//IFIP International Conference on Computer Information Systems and Industrial Management. Geneva: IFIP Newsletter, 2017: 229-239.

[25] SAXENA A, MISRA J, DHAR A. Increasing anonymity in bitcoin[C]//International Conference on Financial Cryptography and Data Security. Berlin: Springer, 2014: 122-139.

[26] MIERS I, GARMAN C, GREEN M, et al. Zerocoin: anonymous distributed e-cash from bitcoin[C]//2013 IEEE Symposium on Security and Privacy. Piscataway: IEEE Press, 2013: 397-411.

[27] BEN S E, CHIESA A, GARMAN C, et al. Zerocash: decentralized anonymous payments from bitcoin[C]//2014 IEEE Symposium on Security and Privacy. Piscataway: IEEE Press, 2014: 459-474.

[28] YIN W, WEN Q Y, LI W M, et al. An anti-quantum transaction authentication approach in blockchain[J]. IEEE Access, 2018, 6: 5393-5401.

[29] DOUCEUR J R. The sybil attack[C]//The First International Workshop on Peer-to-Peer Systems. Berlin: Springer, 2002: 251-260.

[29] DOUCEUR J R. The sybil attack[M]. Berlin: Springer Berlin Heidelberg, 2002: 251-260.

[30] KARAME G O, ANDROULAKI E, CAPKUN S. Double-spending fast payments in bitcoin[C]//The 2012 ACM conference on Computer and communications security. New York: ACM Press, 2012: 906-917.

[31] LAMPORT L, SHOSTAK R, PEASE M. The byzantine generals problem[J]. ACM Transactions on Programming Languages and Systems, 1982, 4: 382-401.

[32] BANO S, SONNINO A, AL-BASSAM M, et al. Consensus in the age of blockchains[J]. arXiv Preprint, arXiv: 1711.03936, 2017.

[33] DWORK C, LYNCH N, STOCKMEYER L. Consensus in the presence of partial synchrony[J]. Journal of the ACM, 1988, 35(2): 288-323.

[34] TSCHORSCH F, SCHEUERMANN B. Bitcoin and beyond: a technical survey on decentralized digital currencies[J]. IEEE Communications Surveys & Tutorials, 2016, 18(3): 2084-2123.

[35] CACHIN C, VUKOLIĆ M. Blockchains consensus protocols in the wild[J]. arXiv Preprint, arXiv: 1707. 01873, 2017.

[36] CASTRO M, LISKOV B. Practical Byzantine fault tolerance and proactive recovery[J]. ACM Transactions on Computer Systems, 2002, 20(4): 398-461.

[37] ONGARO D, OUSTERHOUT J. In search of an understandable consensus algorithm[C]//The 2014 USENIX Conference on USENIX Annual Technical Conference. Berkeley: USENIX Association, 2015: 305-320.

[38] BALL M, ROSEN A, SABIN M, et al. Proofs of useful work[R]. 2017.

[39] BACH L M, MIHALJEVIC B, ZAGAR M. Comparative analysis of blockchain consensus algorithms[C]//2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics. Piscataway: IEEE Press, 2018: 1545-1550.

[40] KIAYIAS A, RUSSELL A, DAVID B, et al. Ouroboros: a provably secure proof-of-stake blockchain protocol [C]//Advances in Cryptology-CRYPTO 2017. Berlin: Springer, 2017: 357-388.

[41] FISCH B. Tight proofs of space and replication[J]. IACR Cryptology ePrint Archive, ePrint-2018-702.

[42] BELOTTI M, BOŽIĆ N, PUJOLLE G, et al. A vademecum on blockchain technologies: when, which, and how[J]. IEEE Communications Surveys & Tutorials, 2019, 21(4): 3796-3838.

[43] WANG W B, HOANG D T, HU P Z, et al. A survey on consensus mechanisms and mining strategy management in blockchain networks[J]. IEEE Access, 2019, 7: 22328-22370.

[44] YOO J, JUNG Y, SHIN D, et al. Formal modeling and verification of a federated Byzantine agreement algorithm for blockchain platforms[C]//2019 IEEE International Workshop on Blockchain Oriented Software Engineering. Piscataway: IEEE Press, 2019: 11-21.

[45] ZHENG Z B, XIE S A, DAI H N, et al. An overview of blockchain technology: architecture, consensus, and future trends[C]//2017 IEEE International Congress on Big Data. Piscataway: IEEE Press, 2017: 557-564.

[46] YIN M, MALKHI D, REITER M K, et al. HotStuff: BFT consensus in the lens of blockchain[C]// ACM Symposium on Principles of Distributed Computing. New York: ACM Press, 2019: 347-356.

[47] ALI S, WANG G J, WHITE B, et al. Libra critique towards global decentralized financial system[M]. Singapore: Springer Singapore, 2019: 661-672.

[48] BENTOV I, LEE C, MIZRAHI A, et al. Proof of activity[J]. ACM SIGMETRICS Performance Evaluation Review, 2014, 42(3): 34-37.

[49] DECKER C, SEIDEL J, WATTENHOFER R. Bitcoin meets strong consistency[J]. arXiv Preprint, arXiv: 1412. 7935, 2014.

[50] KOKORIS-KOGIAS E, JOVANOVIC P , GAILLY N, et al. Enhancing bitcoin security and performance with strong consistency via collective signing[J]. Applied Mathematical Modelling, 2016, 37: 5723-5742.

[51] BUTERIN V, GRIFFITH V. Casper the friendly finality gadget [J]. arXiv Preprint, arXiv: 1710.09437, 2017.

[52] TSCHORSCH F, SCHEUERMANN B. Bitcoin and beyond: a technical survey on decentralized digital currencies[J]. IEEE Communications Surveys & Tutorials, 2016, 18(3): 2084-2123.

[53] KIAYIAS A, MILLER A, ZINDROS D. Non-interactive proofs of proof-of-work [J]. IACR Cryptology ePrint Archive, ePrint-2017-963.

[53] KIAYIAS A, MILLER A, ZINDROS D. Non-interactive proofs of proof-of-work[M]. Cham: Springer International Publishing, 2020: 505-522.

[54] LUU L, NARAY ANAN V, ZHENG C, et al. A secure sharding protocol for open blockchains[C]//The 2016 ACM SIGSAC Conference on Computer and Communications Security. New York: ACM Press, 2016: 17-30.

[55] KOKORIS-KOGIAS E, JOVANOVIC P, GASSER L, et al. OmniLedger: a secure, scale-out, decentralized ledger via sharding[C]//2018 IEEE Symposium on Security and Privacy. Piscataway: IEEE Press, 2018: 583-598.

[56] LI S Z, YU M C, YANG C S, et al. PolyShard: coded sharding achieves linearly scaling efficiency and security simultaneously[J]. IEEE Transactions on Information Forensics and Security, 2021, 16: 249-261.

[57] XIE J F, YU F R, HUANG T, et al. A survey on the scalability of blockchain systems[J]. IEEE Network, 2019, 33(5): 166-173.

[58] BURCHERT C, DECKER C, WATTENHOFER R. Scalable funding of bitcoin micropayment channel networks[C]//Stabilization, Safety, and Security of Distributed Systems. Berlin: Springer, 2017: 361-377.

[59] LUU L, CHU D H, OLICKEL H, et al. Making smart contracts smarter[C]//Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security. New York: ACM Press, 2016: 254-269.

[60] BRENT L, JURISEVIC A, KONG M, et al. Vandal: a scalable security analysis framework for smart contracts[J]. arXiv Preprint, arXiv: 1809.03981, 2018.

[61] JIANG B, LIU Y , CHAN W K. ContractFuzzer: fuzzing smart contracts for vulnerability detection[J]. arXiv Preprint, arXiv: 1807.03932, 2018.

[62] HASHEMI S H, FAGHRI F, CAMPBELL R H. Decentralized user-centric access control using pubsub over blockchain[J]. arXiv Preprint, arXiv: 1710.00110, 2017.

[63] BAO S H, CAO Y, LEI A, et al. Pseudonym management through blockchain: cost-efficient privacy preservation on intelligent transportation systems[J]. IEEE Access, 2019, 7: 80390-80403.

[64] SAMANIEGO M, DETERS R. Hosting virtual IoT resources on edge-hosts with blockchain[C]// IEEE International Conference on Computer & Information Technology. IEEE, 2016: 116-119.

[65] STANCIU A. Blockchain based distributed control system for edge computing[C]// International Conference on Control Systems & Computer Science. Piscataway: IEEE Press, 2017: 667-671.

[66] ZIEGLER M H, GROMANN M, KRIEGER U R. Integration of fog computing and blockchain technology using the plasma frame-work[C]//2019 IEEE International Conference on Blockchain and Cryptocurrency. Piscataway: IEEE Press, 2019: 120-123.

[67] KIM H, PARK J, BENNIS M, et al. Blockchained on-device federated learning[J]. arXiv Preprint, arXiv: 1808.03949, 2018.

[68] BRAVO-MARQUEZ F, REEVES S, UGARTE M. Proof-of-learning: a blockchain consensus mechanism based on machine learning competitions[C]//2019 IEEE International Conference on Decentralized Applications and Infrastructures. Piscataway: IEEE Press, 2019: 119-124.

[69] LIU J, HUO R, LI C C, et al. Information transmission mechanism of Blockchain technology based on named-data networking[J]. Journal on Communications, 2018, 39(1): 24-33.

[About the author]

Zeng Shiqin (1995-), male, born in Nanning, Guangxi, PhD. Candidate, Beijing University of Posts and Telecommunications, majoring in block chain, sign analysis technology, and industrial Internet.

Huo Ru (1988-), female, born in Harbin, Heilongjiang Province, doctor, lecturer of Beijing University of Technology, her main research interests are computer network, information center network, network cache strategy and algorithm, industrial Internet, identification analysis technology, etc.

Huang Tao (1980-), male, born in Chongqing, doctor, professor of Beijing University of Posts and Telecommunications. His research interests include future network architecture, software-defined network, network virtualization, etc.

Liu Jiang (1983-), male, born in Zhengzhou, Henan province, doctor, professor of Beijing University of Posts and Telecommunications. His research interests include future network architecture, software-defined network, network virtualization, information center network, etc.

Wang Shuo (1991-), male, from Lingbao, Henan province, doctor, postdoctoral fellow of Beijing University of Posts and Telecommunications, whose research interests include data center network, software definition network, network traffic scheduling, etc.

Feng Wei (1980-), male, born in Handan, Hebei province, doctor, associate researcher of the Ministry of Industry and Information Technology, his main research interests are industrial Internet platform, digital twin, key technologies for integrated development of informatization and industrialization, etc.