Blockchain-Based Voting System

Jiayang Yao*, Lan Wei, Tianqi Liu
North China University of Science and Technology, China. E-mail: 2912178216@qq.com

Abstract: The advantage of electronic voting is that the voting is convenient and quick and the counting of votes is simple, but it is difficult to meet the requirements of both high security and large-scale application. This paper proposes an electronic voting protocol and system based on blockchain technology. The blockchain layer will use the P2P network to build a distributed database, use digital signature algorithms and encryption technologies to ensure that the data cannot be tampered with, and consensus algorithms to ensure the consistency of the data in the network, and use time stamping technology to connect the data blocks end to end chain structure preservation. And the application of Ethereum fully integrates different functional modules as a whole, achieves consensus through a proof-of-work mechanism, miners mine, and the new network protocol is formulated to achieve synchronous operation of the blockchain. Finally, the smart contract deployed on Ethereum runs on an Ethereum-specific virtual machine, and interacts with the underlying blockchain through the Ethereum virtual machine and RPC interface.

Keywords: Voting System; Blockchain; Smart Contract; Encryption Algorithm

1. Blockchain technology

1.1 Technical principle

Blockchain refers to a technical solution for collectively maintaining a reliable database through decentralization and trust. Figure 1 shows a simplified schematic diagram of the blockchain. A blockchain is a series of data blocks connected in chronological order, which can verify the correctness of each other and use the consensus algorithm between nodes to generate and update the data, and realize the secure transmission of the data through cryptography.

Each data block records a set of tree transaction status information composed of hash algorithms. The pointer is the hash value of the block header generated by processing the block header using a cryptographic hash algorithm. Each data block records a set of tree transaction status information composed of hash algorithms to ensure that the transaction data in each block cannot be tampered with.

Figure 1. Simplified schematic diagram of blockchain.
1.2 Features and classification

1.2.1 Characteristics of blockchain

Blockchain has the characteristics of decentralization, detrust, time series data, collective maintenance and security and credibility.

1) Decentralization. All nodes in the blockchain network jointly provide services and participate in the process of data verification, accounting, storage, maintenance and transmission in the network. There is no central node with special rights in the network, thus forming a decentralized distributed system.

2) Detrust. The trust in the blockchain network is endorsed by mathematical algorithms. When the nodes in the blockchain network exchange data with each other, they do not need to know the credibility of the other party and establish trust through mathematical algorithms. The operation of the system is open and transparent, and the data can be checked.

3) Time series data. The blockchain stores data blocks on the blockchain in a chain structure, with time stamps in the data blocks. There is a block hash value in the head of each block. Through the block hash value, all the data blocks in the blockchain are serialized into a chain, and the longest chain (main chain). Therefore, the time dimension of the data is increased, and the data can be traced and verified conveniently.

4) Safe and reliable. Through the "block ten-chain structure", the blockchain can discover any data tampering in time. The blockchain provides a time-series, checkable record, which can be regarded as a tamper-proof, authentic database.

1.2.1 Classification of blockchain

According to the classification of participants, blockchain can be divided into: Public Blockchain, Private Blockchain and Consortium Blockchain.

Different blockchains can also form a network, and the interconnection of chains and chains in the network creates the concept of Interchain. A summary of the three types of blockchains is shown in Table 1.

| Contrastive terms      | Public chain          | Alliance chain        | Private chain       |
|------------------------|-----------------------|-----------------------|---------------------|
| Degree of decentralization | Complete decentralization | Partial decentralization | Centralization     |
| Participant            | Anybody               | Alliance members only | Designated member of central controller |
| Bookholder             | All participants      | Negotiation decision  | Custom              |
| Trust mechanism        | Proof of work         | Consensus mechanism   | Endorsement on its own initiative |
| Advantage               | Go to the center, go to the trust; any user can access; the application is easy to deploy; | It is easy to limit the right of control and has high scalability; | Low network energy consumption, easy to modify rules, unlimited transaction volume and speed: no 51% attack risk |
| Shortcoming            | Limited trading volume | Can't solve the information problem completely | Access nodes are limited: information problems cannot be solved |

Table 1. Comparison of three types of blockchains

The autonomy of the blockchain determines that the nodes in the blockchain network can independently monitor
the data information occurring in other nodes and share them at any time. The entire process is implemented independently by the blockchain network without manual intervention. Each node in the blockchain network is the maintainer of the entire network, and no node in the network has absolute priority.

2. Construction of a decentralized secure electronic voting system

2.1 Design plan

In order to meet the needs of fairness, work and transparency of the system, this design uses blockchain technology to implement the voting system. The blockchain is a non-tamperable, authentic and reliable database. In this design, blockchain data (voting data) is stored in the blockchain, and ordinary data (such as user names and passwords) is stored in the local database.

The system architecture is shown in Figure 2, which mainly includes the application layer and the blockchain layer. The application layer adopts B/S architecture, and the Web application layer mainly implements system functions (registration, login, voting, etc.). The blockchain layer can be subdivided into a network layer and a storage layer. This layer sends voting data to the blockchain network and enables distributed access to the data in the blockchain. In order to meet the requirements of decentralization, data tamper-proof, safe and credible of blockchain, the blockchain layer will use the P2P network to build a distributed database, use digital signature algorithms and encryption technologies to ensure that the data cannot be tampered with, and consensus algorithms to ensure the network consistency of data in the data, and use timestamp technology to save the data blocks in a chain structure connected end to end.

![Figure 2. System solution.](image)

2.2 Design of the system architecture

In order to meet the requirements of decentralization, data tamper-proof, safe and credible of blockchain, the blockchain layer will use the P2P network to build a distributed database, use digital signature algorithms and encryption technologies to ensure that the data cannot be tampered with, and the consensus algorithm to ensure the network consistency of data in the data, and use timestamp technology to save the data blocks in a chain structure connected end to end.

2.2.1 Design of the network layer

The network layer realizes the construction of the network, and the verification mechanism and the consensus mechanism are used to ensure the safety and consistency of the data. The design focus is to build a P2P network to
implement the verification mechanism and the consensus mechanism. The storage layer encapsulates the data blocks and saves the data blocks in a chain end to end. The design focus is to use time stamps, hash functions, Merkle trees, asymmetric encryption and other technologies to build data block structures, using Java as the development language.

2.2.1.1 Node P2P network node construction

Building a P2P network is the initialization process of the blockchain layer. If a new node joins the P2P network for the first time, the IP addresses of other nodes in the network are required. DNSsced can provide a DNS server for the IP address of the node on the P2P network, which helps to discover the node. So DNSsced is adopted to join P2P network, with TCP protocol and 8333 port.

2.2.1.2 Block synchronization

When a node first joined the P2P network, there was no blockchain data or the data was incomplete, and it was necessary to request data blocks from other nodes. This process is called block synchronization and is also the initialization process of the blockchain layer. The node will first load and verify the local data block, which is divided into block verification and data verification. First verify whether the previous block is on the main chain. After completing the block verification, perform block data verification. After verifying the local data block, the node performs block synchronization.

2.2.1.3 Data verification

The verification mechanism ensures that the data cannot be tampered with. Every node in the P2P network continuously receives data. After receiving the data, the node will verify the validity of the data at the first time, by verifying the data structure and digital signature. Only the data that meets all the conditions will be valid. If the data is invalid, it is discarded. If the data is valid, the number of valid data is stored in the data pool.

2.2.1.4 Consensus mechanism

The consensus mechanism provides guarantee for the consistency of data in the blockchain, and is the key to the blockchain's ability to maintain its operation. The general model of the consensus mechanism in the computer field is: in a distributed system with a reliable channel, in the case of malicious nodes, how does the system ensure that other nodes are not affected by malicious nodes, and reach a correct consensus on a certain problem, so that the entire system runs stably and reliably.

2.2.2 Design of storage layer

2.2.2.1 Data block

In order to realize the storage of data in a chain structure based on time stamps and to quickly verify the validity of the data, the data block adopts the structure shown in the figure, and each data block is divided into a block header and a block body. The block header contains the version number, the previous block hash value (Prey-block), timestamp (Timestamp), random number (Nonce), the target block hash value (Bits) and Merkle root (Merkle-root). The valid data and corresponding amount generated during the block creation process are stored in the block body. The valid data generates a unique Merkle root through the hash of the Merkle tree and stores it in the block header.

2.2.2.2 Time stamp and hash function

The application of time stamp technology in digital information can verify the exact time when digital information is generated. Time stamp technology can provide reliable verification of the integrity and existence of digital information before a certain time period or a certain time point. Hash function has the characteristics of unidirectional, timing, fixed length and randomness, so it is very suitable for verifying data. This system will use digital signature algorithm and encryption technology to ensure that the data cannot be tampered with. The choice of hash function will affect the security of the system. This design will use two hash functions SHA256 and RIPEMD160. Among them, SHA256 is used more frequently, which is used for block header hash value, block data, and generation of blockchain
addresses, while RIPEMD160 is only used to generate blockchain addresses.

2.2.2.3 Merkle Tree

When performing block verification, if the data in the current block is verified one by one, the efficiency is very low. To this end, Merkle Tree can be used to quickly check the block data. Merkle Tree is a Hash Tree that can verify the integrity of block data in a short period of time, that is, to ensure that data in the blockchain network is not lost or modified, and nodes do not send fake data blocks.

2.2.2.4 Asymmetric encryption technology

In order to achieve data security and ownership verification, this system uses asymmetric encryption. Asymmetric encryption technology will play an important role in application scenarios such as digital signatures and information encryption at the blockchain layer. In the digital signature scenario, the sender in the blockchain uses their own private key to encrypt the information, and sends the public key and private key together to the blockchain network. The node uses the public key to decrypt the information, thereby verifying the information ownership. In the information encryption scenario, the sender in the blockchain will encrypt the information with the receiver's public key, and the receiver will decrypt the information with the private key. RSA relies on the prime factorization of large numbers, which makes it impossible to theoretically measure the confidentiality of RSA. The RSA key is too long, and the large number operation causes its encryption and decryption process to be slow. The elliptic curve encryption algorithm provides a shorter key than RSA, and has the characteristics of higher security and faster processing speed. Therefore, this design will use elliptic curve encryption algorithm, the generation mechanism of asymmetric encryption public key and private key is shown in the figure.

![Figure 3. Asymmetric encryption generation mechanism.](image-url)

3. Solutions and results

First of all, this paper focused on the design of the P2P network in the network layer, node block synchronization, data and block verification mechanism, and consensus mechanism to ensure data consistency. Then it introduced the use of timestamp, Merkle Tree, asymmetric encryption and other technologies to design data blocks, and finally the use of chain structure to store data blocks was introduced.

The network layer realizes the construction of the network, and a verification mechanism and a consensus mechanism are used to ensure the security and consistency of the data. The blockchain node network is a P2P network, and all nodes in the blockchain network jointly participate in the process of data verification, accounting, storage, maintenance, and transmission in the network.

The storage layer stores the data in a chained data structure based on timestamps. The blocks cannot be changed once they have been verified and saved in the blockchain. Through the "block ten-chain structure", the blockchain can discover any data tampering in time. The blockchain provides a time-series, checkable record, which can be regarded as a tamper-proof, authentic database.
4. Application of Ethereum

In order to develop applications more conveniently and flexibly, while allowing applications to share a viable voting environment and reliable blockchain security, the most representative Ethereum platform in blockchain 2.0 is used to achieve the perfect combination of blockchain and smart contracts, providing a trusted execution environment for smart contracts.

4.1 Ethereum

Ethereum is a Turing complete open blockchain technology platform that implements powerful functions through smart contracts and decentralized applications.

An important innovation of Ethereum is the comprehensive integration of different functional modules as a whole, which is a comprehensive platform for creating and deploying decentralized applications. At the same time, Ethereum is also a perfect combination of blockchain and smart contracts, a complete solution for smart contracts with a complete set of tools that can expand functions.

4.2 Public and private keys and addresses

There are two main types of accounts in Ethereum: one is an external account and the other is a contract account. The private key can control the external account, and the code associated with the external account is usually absent, mainly used to change the number of votes. The contract code can manipulate the contract account and can be used for the content of voting information.

4.3 Ethereum Virtual Machine

The Ethereum Virtual Machine is the operating environment of smart contracts in Ethereum. One of its major innovations in the Ethereum project is the selection of a Turing complete computing environment. It is composed of many computers connected to each other, allowing users to create and upload complex applications according to their own wishes, allowing these operations to be performed automatically, while ensuring that the status of all current and previous programs is always publicly visible.

Every node of the Ethereum network runs an EVM and executes contract code, so Ethereum is like a "world computer" running in parallel, and the state transition of accounts is performed on all nodes at the same time.

4.4 Blockchain network

The blockchain network uses the P2P network protocol, which ensures that all nodes in the same network are equal to each other; there is no central node, and all nodes provide the same network services. The P2P network protocol is of open source and decentralization. Blockchain data is not recorded and stored on a few centralized computers or servers, but all nodes participating in this network can equally store and record all data information, as shown in Figure 4.
4.5 Ethereum smart contract

Ethereum's smart contract is a piece of code that can be executed by the Ethereum virtual machine. Ethereum's unique binary form is stored in the Ethereum blockchain with a smart contract account address and is interpreted by the Ethereum virtual machine. It is called Ethereum virtual machine code. A smart contract is like a trustworthy organization. It keeps the vote information for users and will always perform the corresponding operation according to the rules designed in advance. It responds to the received information and can accept and store the information. Figure 5 shows the smart contract operation model.

5. Solutions and results

5.1 Solutions

The experimental platform runs a 64-bit Windows 10 operating system and uses Ethereum as a blockchain framework. Through the genesis block file, the operating environment of the two nodes of the Ethereum private chain is built locally. One node runs the vote initiator client, and the other node runs multiple voter clients to simulate multiple voters for voting tests.
After connecting the Ethereum wallet to the local private chain network, deploy the smart contract of the voting scheme to the local private chain through the Ethereum wallet, obtain the contract address and the application binary interface (ABI), and send it to the main contract to call the two contracts. Table 2 shows the deployed contract addresses of these three smart contracts.

The total amount of natural gas consumed by each operation during the voting process is tested and calculated. In the voting process, the simulated vote included 40 voters, and each operation communicates data through transactions and voting smart contracts. Therefore, every operation in the voting process needs to be consumed, and the cost incurred during the voting process is the basis. It is determined by consumption and natural gas price, \( \text{cost} = \text{natural gas price} \times \text{consumption} \). The revenue generated is determined by the calculation of the transaction (or execution contract code), and the price of natural gas is set by the originator of the transaction.

| Participant | Operation | Gas consumption | Total gas consumption |
|-------------|-----------|-----------------|-----------------------|
| Voting initiator | Deploy Vote Contract Registry.sol | 883,497 | 16,618,022 |
| | Deploy Vote.sol | 3,779,936 | |
| | Deploy Local Crypto.sol | 4,171,615 | |
| | Set voting items | 254,794 | |
| | Verify voters | 2,052,714 | |
| | Start voting | 2,975,343 | |
| | Vote counting | 2,500,123 | |
| Voter | register | 663,241 | 3,153,673 |
| | vote | 2,490,432 | |

Table 3. Gas consumption statistics

5.2 Results analysis

According to the gasoline consumption required by the voting initiator and the gasoline consumption required by each voter in Table 3 above, the deployment of these three smart contracts is within the maximum gasoline consumption limit of a single transaction in the public chain Ethereum (8 million gas), which proves that the voting scheme in this paper can run in the local private chain or in the public chain. This shows that the multi-candidate electronic voting system is safe and reliable.

6. Conclusion

The main work of this paper is to realize the relevant functions of voting on the basis of building the Ethereum blockchain. Blockchain technology and smart contracts have decentralized features, which can enhance the verifiability of voting data, reduce the cost of the entire system, and meanwhile keep the voting open and transparent. In the model architecture of the blockchain, the consensus mechanism and the security mechanism are the core technical
components of the blockchain. According to the application of the electronic voting system, this paper puts forward the asymmetric digital password, the distributed structure, and the consensus mechanism, respectively.

Of course, due to the limited personal ability, the entire system still has certain defects. For example, in the voting process, voters need to complete more steps, which is not conducive to the enthusiasm of voting. In the subsequent work, it still needs to be improved. At the same time, this system will conduct in-depth discussion and research in the following aspects in the follow-up research process.

A blockchain platform is used to record all other information including the voter's personal information, and the encryption and immutability of the blockchain ensure that the voter's personal information is not leaked or tampered with. In the voting system, it is possible to reduce the number of third parties such as signature agencies, and implement signature verification through smart contracts to achieve further decentralization.

When a specific group of people vote (such as a group of specific age, job, etc.), the blockchain voting platform should be able to set up a relevant screening mechanism, so that only those who meet the conditions can vote.

References

1. Yan C. Research and design of block chain-based secure voting system (in Chinese). Hangzhou Dianzi University; 2018.
2. Huang Z, Li X, Lai X, et al. Block chain technology and its application (in Chinese). Information Security Research.
3. Chen H. Design and implementation of IMIX transmission system based on block chain (in Chinese). 2016.
4. Ma A, Pan X, Wu L, et al. Overview of research on the basis and application of block chain technology (in Chinese). Information Security Research 2017; 3(11).
5. Dong Y, Zhai J, Sun T, et al. Electronic voting system based on privacy protection. Computer Application 2016; (s2): 73-76.
6. Hu H. Design and development of blockchain-based voting system. Nanjing University of Posts and Telecommunications.
7. Zhang H. Research on trust mechanism based on node information exchange in P2P network environment. Network Security Technology and Application 2016; (11). doi: 10.3969/j.issn.1009-6833.2016.11.035.
8. Han X, Liu Y. Research on consensus mechanism in block chain technology. Information Network Security 2017; (9). doi: 10.3969/j.issn.1671-1122.2017.09.034.
9. Gervais A, Karame GO, Wüst K, et al. On the security and performance of proof of work blockchains. The 2016 ACM SIGSAC Conference; ACM 2016. doi: 10.1145/2976749.2978341.
10. Buldas A, Niitsoo M. Can we construct unbounded time-stamping schemes from collision-free hash functions? ProvSec 2008; 2008 Oct 30-Nov 1; Proceedings. Springer-Verlag, 2008. doi: 10.1007/978-3-540-88733-1_18.