2021 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA) | 978-1-6654-1867-6/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICAICA52286.2021.9497966

Transaction-based Static Indexing Method to Improve the Efficiency of Query on the Blockchain

Lei Zeng
Dept. of Govt. & Enterprise Business
Hangzhou Qulian Technology Co., Ltd. Hangzhou, China
zenglei@hyperchain.cn

Weiwei Qiu
Research and Development Center
Hangzhou Qulian Technology Co., Ltd. Hangzhou, China
qiuweiwei@hyperchain.cn

Xiaoyi Wang
DataMesh Labs
Hangzhou Qulian Technology Co., Ltd. Hangzhou, China
wangxiaoyi@hyperchain.cn

Hongkai Wang
State Grid Zhejiang Electric Power Corporation Information & Telecommunication Branch Hangzhou, China
isswang@126.com

Yiyang Yao
State Grid Zhejiang Electric Power Corporation Information & Telecommunication Branch Hangzhou, China
1244872862@qq.com

Zhi Yu
School of Software Technology, Zhejiang University
Hangzhou, China
yuzhirenzhe@zju.edu.cn

***Abstract*—In the context of advocating the reduction of greenhouse gases, green energy trade has become the focus of attention. At this stage, green energy is relatively dispersed (such as wind power, hydropower, solar power, etc.), and energy generation is real-time. At present, green energy trading based on blockchain has become a better solution. In the transaction itself, the blockchain has better security and traceability, but in terms of statistical information and interval query, because the basic data structure of the blockchain is a key-value data structure, performance problems are prone to occur. A transaction-based static indexing method is proposed, which can effectively improve the efficiency of statistical information and interval query on the chain.**

Keywords—blockchain, green energy trading, static indexing

I. INTRODUCTION

Electric vehicles have gradually received strong support from China and many other countries due to their low pollution, low emissions and other environmental protection characteristics [1,2,3,4]. The rapid development of electric vehicles has continuously increased the demand for charging services. However, due to the periodicity of people's daily work and rest, the charging time also has certain time characteristics. Specific data on the charging behavior of electric vehicles can guide the formation of effective and reasonable support policies. In order to effectively manage the charging behavior of electric vehicles and realize the traceability of electric vehicle charging data, it is necessary to have a corresponding information system for management. The emerging blockchain technology is the basic technology of Bitcoin. Blockchain is a decentralized and growing distributed database. Smart contracts and consensus mechanisms ensure that all nodes can reach an agreement on the validity of the contained data or data blocks. Therefore, the system built with blockchain technology has the advantages of traceability, security, transparency and anti-tampering [5,6].

At present, there is a green energy traceability system based on blockchain [7], which records and manages green power. Compared with the traditional middleman management model [8], this system uses the anti-tampering feature of the blockchain [9] to ensure data traceability [10].

However, due to the design of the underlying data structure of the blockchain, the efficiency of data statistics, interval query and other operations is low. However, such operations are often more common. For this reason, this article proposes a transaction-based static index method to effectively improve the efficiency of querying on the blockchain.

When the car owner needs to use the charging pile for charging, the shared charging pile platform will record the charging process, and pay and settle the charging fee after confirming the digital identity; the members of the platform can open the charging status and conduct transparent and real-time billing, thereby Resolve the trust conflicts that may arise between multiple parties and form a shared charging alliance, involving electric vehicle time-sharing operators, vehicle owners, charging pile operators, electric vehicle owners and power supply companies.

Finally, the utilization efficiency of the charging pile can be fully stretched and released. The key to solving all problems is the distributed mobile digital identity sharing toll pile platform based on blockchain. This is the latest development of blockchain technology in the field of charging piles. There are always two main problems in the construction of private charging pile sharing platform:

First, the platforms are highly concentrated, and an "information island" has been formed between various platforms. Operator-led centralized platforms have high credit reporting costs and a fragile credit system, and cannot guarantee the credit security of point-to-point direct transactions between charging pile owners and electric vehicle owners. Once the central organization is attacked, data may be lost or tampered with, causing serious consequences. In terms of payment, charging piles of different entities require different payment channels, which makes it difficult for users to pay.

Second, it is difficult to share private toll booths. Due to parking spaces, electricity bills and safety issues, private charging piles (about 60% of the total number of charging piles) cannot form a sharing economy. On the one hand, it hinders the development of the electric vehicle industry, on the other hand it also causes a lot of waste of social resources.

At the beginning of the year, the "2019-2020 Annual Report on China's Charging Infrastructure Development" released by the China Electric Vehicle Charging Infrastructure Promotion Alliance showed that as of December 2019, the cumulative number of charging infrastructure across the country was 1.219 million, of which 516,000 were public piles. . . , There are 703,000 private piles, and private piles account for 57.6%.

Public data shows that by the end of 2019, the number of new energy vehicles reached 3.81 million, of which the number of pure electric vehicles reached 3.1 million, accounting for 81.19% of the total number of new energy vehicles. If calculated on the basis of 3.1 million pure electric vehicles, the current car-to-car ratio in my country is only 2.6:1, which is far lower than the 1:1 proposed in the "Electric Vehicle Charging Infrastructure Development Guide". (2015-2020)".

It is estimated that by 2030, China's electric vehicle sales are expected to exceed 15 million, and the inventory is expected to exceed 80 million. By then, pure electric vehicles may reach 64.8 million. If the calculation is strictly based on the pile-to-pile ratio of 1:1, more than 63 million new piles will need to be built in the 10 years from 2021 to 2030. During this period, shared private piles will continue to cautiously explore the profit model of the charging pile industry. Among them, blockchain technology will be everywhere, incarnate as an accelerator, and further demonstrate its own value:

The multi-centralization of blockchain technology, the characteristics of being difficult to tamper with and traceable, combined with typical blockchain-based application scenarios such as digital identity systems, self-incentive credit systems, and blockchain-based automobile networking systems, can build an A set of practical application scenarios. Effective charging piles share the operating platform. The digital identity system can standardize and customize the identity information of people and the equipment information of vehicles and stakes. On the one hand, it can use cryptography to share and share identity information while ensuring the security of personal information. On the other hand, it can be used by identity. The structure of the system ensures the safety of personal charges. Pile up in the sharing process.

Establish a self-motivating trust mechanism to solve the sharing problem. The car owner confirms the right of charging behavior and records it on the blockchain to establish a credit system that integrates "people and cars". On the one hand, it improves the efficiency of multi-party collaboration and solves the problem of payment difficulties for car owners. On the other hand, it has established a credit system based on consumption scenarios (such as charging and parking), which is an effective supplement to the existing credit system.

Energy point-to-point real-time autonomous transactions based on blockchain technology can effectively reduce the electricity transaction costs of distributed energy point-to-point real-time autonomous transactions. The widespread application of blockchain technology has greatly changed the production and trading methods of clean energy, that is, real-time autonomous transaction entities can achieve efficient production and effective consumption of clean energy through peer-to-peer transactions.

Energy point-to-point real-time autonomous transactions based on blockchain technology mainly have the following functions: The transaction price can be coordinated effectively, the transaction costs can be reduced effectively, and the transaction efficiency can be improved effectively. Also the clean energy can be consumed and used effectively

II. Related Work

Data interval query operations on the blockchain are more common. Currently, there are usually the following three common methods:

The first method is to index part of the data in the database with the help of an external relational database. First, find the relevant transaction in the database, and then read the relevant transaction data on the blockchain for calculation. The advantage of this method is speed, but due to the introduction of a database, considering the consideration of ensuring the consistency of the data on and off the chain, and the relational database may be tampered with by a third party, this method will affect the calculation of reputation. data.

The second method is to use the smart contract of the blockchain, but the performance of the contract itself is limited. After the data reaches a certain scale, performance will drop sharply, and may be restricted by GAS, and query and insert operations cannot be completed within the specified time.

The third method is to use the original transaction, embed the last "transaction hash" into the original transaction, and associate related transactions in the form of a linked list, as shown in Figure 1 below. When querying, it needs to traverse the header transaction. This method is not limited by the performance of the smart contract, and the insertion efficiency is high. However, due to the chain structure, the query efficiency may be very low.

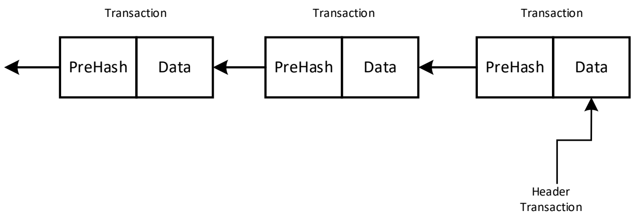

Fig. 1. Example of a original transaction linked list

III. Static indexing

Based on the idea of transaction linked list, a transaction-based static indexing scheme is designed. By embedding index data (*IndexData*) in each transaction, the query efficiency in the blockchain can be improved, as shown in Figure 2.

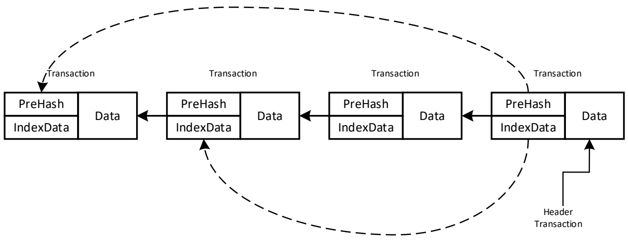

Fig. 2. Example of a static index linked list based on transaction

Assuming that data is inserted in order, time interval query and data statistics need to be performed according to the inserted timestamp. The specific design is as follows:

IndexData in each transaction contains an index field table, which maintains the timestamp and hash value of the first 2^i transactions of the current transaction, as shown in Table I below.

TABLE I. EXAMPLE OF TRANSACTION-BASED STATIC INDEX TABLE

No.	***Timestamp of the 2^i-th TxHash***	***Hash value of the 2^i-th TxHash***
0	$Time_0$	T_{x0}
1	$Time_1$	T_{x1}
...	...	...

This article will introduce the maintenance and function of static indexes through inserting and querying two operations.

When inserting the k-th piece of data into the chain, the *Data* field in the transaction records the business data $Data_k$, and the *PreHash* field records the *Transaction Hash* of the previous transaction. The following details the data update logic of *IndexData*.

A. Insert

Assuming that the *IndexData* of all the data before the k-th data has been updated, then the current k-th *IndexData* data update logic is as follows.

- When i=0 (the 1st data), *IndexData* is *PreHash*.
- When i=1 (the 2nd data), *IndexData* is the data stored in ***No***. 0 of *IndexTable* of *PreHash*.
- When i=2 (the 4th data), *IndexData* is the data stored in ***No***. 1 of *IndexTable* of *PreHash*.
- When i=3 (the 5th dat), *IndexData* is the data stored in ***No***. 2 of *IndexTable* of *PreHash*.
- And so on, until the value of 2^i is updated to be greater than or equal to k.
- If the value of k-2^i is less than 1, the last record in the index table of *IndexData* points to the first transaction data in the chain.

Update complexity: For the case where the data size is N, according to this update method, each row of data only needs to read at most one piece of data, so the update complexity is $O(\log_2 N)$, .

B. Inquire

Suppose you need to query or calculate the interval [a, b], you can quickly query the *TxHash* corresponding to endpoints a and b through the *IndexTable* of each transaction. After querying the two endpoints, more data and statistical information can be obtained.

For a certain endpoint x, the specific query logic starts from *HeaderTransaction*, finds the largest i in the index table, so that it meets the records of the first 2^i *TxHash* timestamps less than x, and then jumps to the corresponding *IndexTable* transaction. Continue to search for the record corresponding to the largest i that satisfies the condition in the new index table until a specific transaction is found, and the *0*th record in the *IndexTable* does not satisfy the search condition.

Assuming that the values of timestamp and *TxHash* are both 1, 2, 3,..., 100 (the intermediate integers are continuous), the following is an example of querying transaction data with a timestamp of 16.

The current *HeaderTransaction* is a transaction with a timestamp of 100, and its internal *IndexTable* is shown in Table II below.

TABLE II. EXAMPLE OF INDEX TABLE WITH TIMESTAMP OF 100

No.	***Timestamp of the 2^i-th TxHash***	***Hash value of the 2^i-th TxHash***
0	Max(1,100-2^0)=99	99
1	Max(1,100-2^1)=98	98
2	Max(1,100-2^2)=96	96
3	Max(1,100-2^3)=92	92
4	Max(1,100-2^4)=84	84
5	Max(1,100-2^5)=68	68
6	Max(1,100-2^6)=36	36
7	Max(1,100-2^7)=1	1

It can be seen that the maximum value of the timestamp that satisfies the first 2^i *TxHash* in the table is less than x is 32, and the corresponding value of i is 6.

Further jump into the entry corresponding to the 6th record, that is, the transaction with a timestamp of 32. Jump into the *IndexTable* of this transaction, as shown in Table III below.

TABLE III. EXAMPLE OF INDEX TABLE WITH TIMESTAMP OF 32

No.	***Timestamp of the 2^i-th TxHash***	***Hash value of the 2^i-th TxHash***
0	Max(1,32-2^0)=31	31
1	Max(1,32-2^1)=30	30
2	Max(1,32-2^2)=28	28
3	Max(1,32-2^3)=24	24
4	Max(1,32-2^4)=16	16
5	Max(1,32-2^7)=1	1

At this time, i=4 satisfies the condition, and jumps to the *IndexTable* of the transaction with the corresponding timestamp 16. It can be seen that when i=0, the timestamp of the first 2^i *TxHash* is less than x, so the timestamp is determined to be 16. which is the transaction queried.

Finding complexity analysis: For the case where the data size is N, if the internal dichotomy can reach $O(\log_2(\log_2 N))$, and the query complexity of *IndexTable* is $\log_2 N$ each time.

For global operations, the worst one in each IndexTable can reduce the search size by half in the previous query, so the global time complexity is $O(\log_2 N)$.

After querying the two endpoints, data collection and data statistics can be performed. In terms of data collection, based on two endpoints, all data segments can be traversed directly.

In terms of statistics, general statistics must be monotonous. Take the statistics of abnormal transaction data as an example. Whenever abnormal transaction data occurs, you can add statistical information of prefix sum to the data segment:

$$ErrorPrefixSum_i = \sum_1^i Error\ a \quad (1)$$

Then, if you want to calculate the abnormal transaction data in the interval [a, b], after finding the endpoints a and b, the abnormal transaction data in the interval is:

$$\sum Error_{[a,b]} = EPS_b - EPS_{a-1} \quad (2)$$

This method satisfies the prefix requirement and can count data quickly.

IV. Experiment

In order to test the performance of the transaction-based static index, this article sets up three sets of experiments to compare the performance of the smart contract method and the method in this article under different conditions:

- Single thread write data.
- Single thread query data.
- Multithread Query data.

For each group of experiments, it is set to write or query 1000, 10000, 100000, 1000000, 1000000 pieces of data respectively, and calculate the average time consumption. The results are shown in Table IV, Table V, and Table VI below, respectively.

TABLE IV. Resultsle of Single thread write data

Total number of data written	***Average time consumption (ms)***	
	Contract method	***Transaction-based static index***
1000	3.91	5.56
10000	4.51	7.98
100000	6.79	11.55
1000000	11.32	17.96
10000000	90.65	29.67

TABLE V. Resultsle of Single thread query data

Total number of data written	***Average time consumption (ms)***	
	Contract method	***Transaction-based static index***
1000	5.11	4.67
10000	6.87	6.98
100000	13.45	9.69
1000000	92.89	13.13
10000000	417.76	22.35

TABLE VI. Resultsle of Multithread Query data

Total number of data written	***Average time consumption (ms)***	
	Contract method	***Transaction-based static index***
1000	42.67	7.42
10000	63.61	12.83
100000	183.84	17.46
1000000	294.68	21.25
10000000	717.58	39.23

As shown in the table, the average time consumption of this method in various experiments is lower than that of the smart contract method. When the total number of data items reaches tens of millions, the single-threaded write performance is about 3 times that of the contract mode; the single-threaded query performance is about 18 times that of the contract mode; in the case of concurrent operations, the efficiency of the method in this article is also significantly improved. The performance also reaches about 18 times that of the contract method. Compared with the paper method, in the case of multi-threading, the query time has increased, but the corresponding TPS has become higher.

V. Conclusion

In order to improve the query efficiency of the blockchain, this article designs a transaction-based static indexing method. Through this index, the performance defects of existing linked list solutions can be improved, and the efficiency of insertion and query can be realized. Compared with the existing blockchain query method, this method has the following four advantages:

- It does not rely on any external databases, they all happen on the chain, and enjoy a high reputation.
- It does not rely on smart contracts.
- You can create indexes for different fields separately, or you can create indexes for historical data separately.
- Good concurrent scalability, and supports concurrent queries.

However, due to the establishment of the index, the space consumption of this method is relatively high, which is an inevitable shortcoming of this method. Some methods can be used to reduce space consumption, such as further compressing *TxHash* storage to reduce space consumption. And add auxiliary index to prevent unlimited expansion of *IndexData* table.

Acknowledgements

This work was financially supported by the science and technology project of the headquarters of State Grid Corporation of China, "Research on the Application Pattern of Power Blockchain Technology Based on Main-side Chains"(No. SGZJ0000KXJS1900422).

References

[1] P. Ahmadi. Environmental impacts and behavioral drivers of deep decarbonization for transportation through electric vehicles. J. Clean. Prod., vol. 225, pp. 1209-1219, july 2019.

[2] X. Zhang, Y. Zou, J. Fan, H. Guo. Usage pattern analysis of Beijing private electric vehicles based on real-world data. Energy, vol. 167, pp.1074-1085, January 2019.

[3] Y. Zhang, Q. Zhang, A. Farnoosh, S.Y. Chen, Y. Li. GIS-Based MultiObjective Particle Swarm Optimization of charging stations for electric vehicles. Energy, vol. 169, pp. 844-853, February 2019.

[4] Beijing New Energy Vehicle Charging Behavior Report , 2019.

[5] F. Tian. An agri-food supply chain traceability system for China based on RFID & blockchain technology. Proceedings of 2016 IEEE 13th International Conference on Service Systems and Service Management (ICSSSM), pp. 1-6, 2016.

[6] V.G Venkatesh, K. Kang, B. Wang, R.Y. Zhong, A. Zhang. System architecture for blockchain based transparency of supply chain social sustainability. Robotics and Computer-Integrated Manufacturing, vol. 63, June 2020.

[7] Wang H, Yao Y, X Wang, et al. A Green Power Traceability Technology Based on Timestamp in a Main-Side Chain System[C]//

2020 13th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI). 2020.

[8] B. Teufel, A. Sentic, M. Barmet. Blockchain energy: Blockchain in future energy systems. Journal of Electronic Science and Technology, vol. 17, December 2019.

[9] J.X. Li, J.G. Wu, L. Chen. Block-secure: blockchain based scheme for secure P2P cloud storage. Inf. Sci., vol. 465, pp. 219-231, October 2018.

[10] M. Tieman, M.R. Darun. Leveraging blockchain technology for halal supply chains. Islam Civilisational Renewal (ICR), vol. 8, pp. 547-550, 2017.