systems

MDPI

Article

Blockchain Traceability System in Complex Application Scenarios: Image-Based Interactive Traceability Structure

Chunhua Ju [1,2], Zhonghua Shen [2], Fuguang Bao [1,2,3,*], Zhikai Wen [4], Xi Ran [2], Chaoyang Yu [2] and Chonghuan Xu [3,5]

[1] Department of Modern Business Research Center, Zhejiang Gongshang University, Hangzhou 310018, China; jch@zjgsu.edu.cn
[2] School of Management Engineering and E-Business, Zhejiang Gongshang University, Hangzhou 310018, China; aucnm0202@163.com (Z.S.) ranxi169@163.com (X.R.); y3173104782@163.com (C.Y.)
[3] Academy of Zhejiang Culture Industry Innovation and Development, Zhejiang Gongshang University, Hangzhou 310018, China; talentxch@zjgsu.edu.cn
[4] School of Foreign Languages, Zhejiang Gongshang University, Hangzhou 310018, China; wenzhikai163@163.com
[5] School of Business Administration, Zhejiang Gongshang University, Hangzhou 310018, China
* Correspondence: baofuguang@126.com

Abstract: To solve the problems exposed by the application of blockchain technology under complex scenarios, such as fraudulent use of data, hard to store huge amounts of data, and low traceability efficiency under an ultra-huge number of traceability requests, this paper constructs an image-based interactive traceability structure by using images as an enhancement. By adding pointers to raw image files, a specific file structure is formed for traceability, and the traceability process is separated from the verification process, therefore realizing the distributed traceability of "traceability off the chain and verification on the chain". The experimental results show that, compared with the traditional blockchain traceability mode, the interactive traceability structure can reduce the data retrieval pressure and greatly improve the traceability efficiency of a specific transaction chain. With the growth of the span of the transaction chain, the traceability efficiency advantage of the interactive traceability structure becomes more obvious.

Keywords: blockchain; IPFS; interactive traceability; process optimization; traceability off the chain

Citation: Ju, C.; Shen, Z.; Bao, F.; Wen, Z.; Ran, X.; Yu, C.; Xu, C. Blockchain Traceability System in Complex Application Scenarios: Image-Based Interactive Traceability Structure. *Systems* **2022**, *10*, 78. https://doi.org/10.3390/systems10030078

Academic Editors: Sabah Mohammed and Luis Javier Garcia Villalba

Received: 15 May 2022
Accepted: 8 June 2022
Published: 8 June 2022

1. Introduction

Data traceability answers why the data are in this state, where the data come from, and how the data are obtained [1,2]. The traditional data traceability system that relies on a centralized database to manage data has problems such as serious centralization of the traceability system, vulnerability of data being tampered with, lack of open access and transparency to data, and data loss [3]. In contrast, blockchain technology uses a distributed storage mode to store data and is widely regarded as an inherent information storage mechanism. It is very suitable for solving the problems of system centralization, lack of data monitoring [4–6], data tracking [7], and lack of trust in traditional traceability systems [8]. Blockchain is the core supporting technology of the digital cryptocurrency system represented by Bitcoin [9]. Because of its unique immutability and traceability, blockchain technology has begun to be applied to data traceability, such as supply chain traceability and agricultural product traceability [10,11]. The definition of data traceability in this paper is that all recorded historical information should be traceable from the latest transaction, and the goal of traceability in this paper is to obtain the historical data and files of a certain transaction chain.

However, the more mature application of blockchain technology focuses on digital currency and finance, which correspond to the characteristics of blockchain 1.0 mode and blockchain 2.0 mode, respectively [12]. However, in the next stage of the blockchain 3.0 model, facing more complex application scenarios and social problems, the application of blockchain technology still has critical technical challenges in scalability, throughput, access control, and data retrieval [13]. In more complex application scenarios, it is bound to be accompanied by more complex data patterns [14]. It poses new challenges for data storage that data are not necessarily structured and may be in a more complex unstructured pattern such as video, picture, audio, etc. In addition, there is still a problem of data misuse in practical applications, that is, the theft of key tags associated with the data and the entity. The stolen tag may be bar codes, QR codes, RFID tags, or IOT device data. In this context, using unstructured videos, pictures, etc., to enhance the connection between data and entities has become an effective means of anti-theft. It can increase the credibility of the data and increase the cost of data theft.

When building a quality traceability system for high-value agricultural products, our team found that the source tracing of the consumer level is very concentrated in practical applications. For example, in the cultivation stage of agricultural products, although a large amount of data need to be stored, the demand for data retrieval is not high. A small number of source tracing requests are mainly issued by some regulatory agencies. However, once it comes to the sales stage of those products, there will be an ultra-huge number of source tracing requests from consumers, which are very concentrated and large in the sales season. This phenomenon makes the data tracing pressure of the full node surge in the sales season, which we call the consumer-level-traceability request pressure. Instead of focusing on professional data, consumers pay more attention to visual data such as images and videos with a strong presentation [15–17]. This paper aims to solve the pressure of consumer-level traceability and meet the needs of consumers, using the image evidence as a means of enhancement to build a consumer-level-traceability system based on blockchain technology. It is intended to achieve credible distributed storage of image evidence and improve traceability efficiency under consumer-level-traceability requests.

2. Literature Review

In a blockchain system, each full node stores a complete set of data. The highly redundant storage mode enhances the openness and transparency of data and improves the system's overall credibility. However, at the same time, this highly redundant storage method also brings great pressure on data storage. Due to the limitation of the size of a single block, a large amount of data, such as images and videos, are not suitable for storage on the blockchain [18]. At present, the methods to improve the scalability of blockchain storage are divided into collaborative online storage and offline storage. The more feasible storage method for data such as images and videos is to store the original data off-chain and store verification information on-chain. Zhang Xiaodie et al. [19] constructed a multi-chain storage protection model for agricultural product traceability data based on blockchain technology. The "blockchain + database" storage structure is used, and the multi-chain structure is adopted to ensure transaction throughput. Peng Hongyan et al. [20] proposed a verifiable encrypted image retrieval service scheme based on blockchain technology. This scheme adopts the "blockchain + cloud storage" data storage structure, which also stores index or verification data on the blockchain. It then stores a large amount of original data or encrypted data on the offline database or cloud. However, both storage modes have a storage combination of "decentralized + centralized". Limited by centralized storage, problems such as loss of original data and limited access to original data may occur in the actual traceability process.

Therefore, the storage of original data should also use a decentralized distributed storage mode, which is also the "blockchain + IPFS" storage structure mainly adopted by the academic world [21]. IPFS (Inter Planetary File System) is a content-addressed, versionable, peer-to-peer distributed file system. It integrates and utilizes technologies such as DHT (Distributed Hash Table), Git (version control system), P2P (Peer-to-Peer, peer-to-peer transmission), BitTorrent (bit stream content distribution protocol), and cryptography, and it can link all the computer devices deployed to the system [22]. In the IPFS system, the file will be split into 256K unit blocks, and each block will calculate a block hash value. In order to obtain the unique file identifier CID (Content ID), all the block hash values are combined and hashed. The CID and the node ID of the corresponding storage file are recorded on the DHT, and the user can download the corresponding file in the IPFS network by using CID alone, as shown in Figure 1. Since IPFS is a distributed file system based on content addressing and hash algorithm, the file data in IPFS are immutable and have strong tamper-proof ability. At the same time, to track file updates, IPFS introduces the version control model to form a version-based historical file structure.

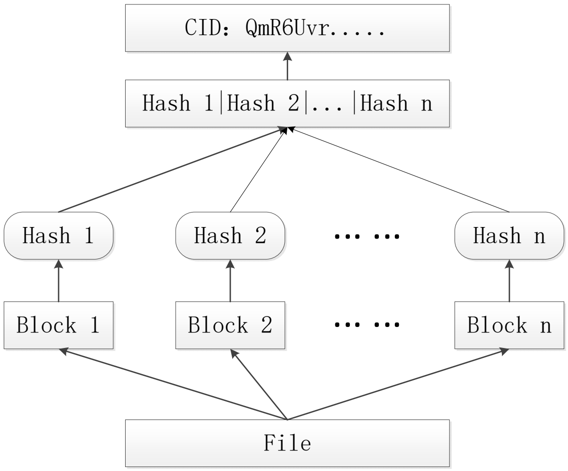

Figure 1. Generation principle of CID in IPFS.

Casino et al. [23] proposed a food supply chain FSC traceability model based on blockchain and smart contracts. Yuan Jian et al. [24] proposed an art blockchain certificate traceability model based on three chains. Wang Keke et al. [25] proposed an efficient solution based on the alliance chain for the safety and efficiency problems in the agricultural product traceability system. Gao Qijuan et al. [26] designed a tea quality safety traceability system using blockchain technology. You Yao et al. [27] proposed a digital asset transaction model based on blockchain and a hybrid indexing mechanism based on transaction chain. In addition, Hyoeun et al. [28], Randhir et al. [29], Tao et al. [30], and Yaseen et al. [31] also proposed corresponding "blockchain + IPFS" data management in the fields of automobile, engineering, construction, and medical care, respectively, for safety, traceability, data sharing, and other fields. The application scenarios in these fields are relatively more complex, and there are large-scale and unstructured data storage requirements such as video data, industrial images, BIM models, and medical images. In this case, these researchers have adopted the solution of "storing verification and indexing data on the chain, and storing gross raw data off-chain". The data storage structure of "storing verification and index data on the chain, and storing gross raw data off-chain" is shown in Figure 2.

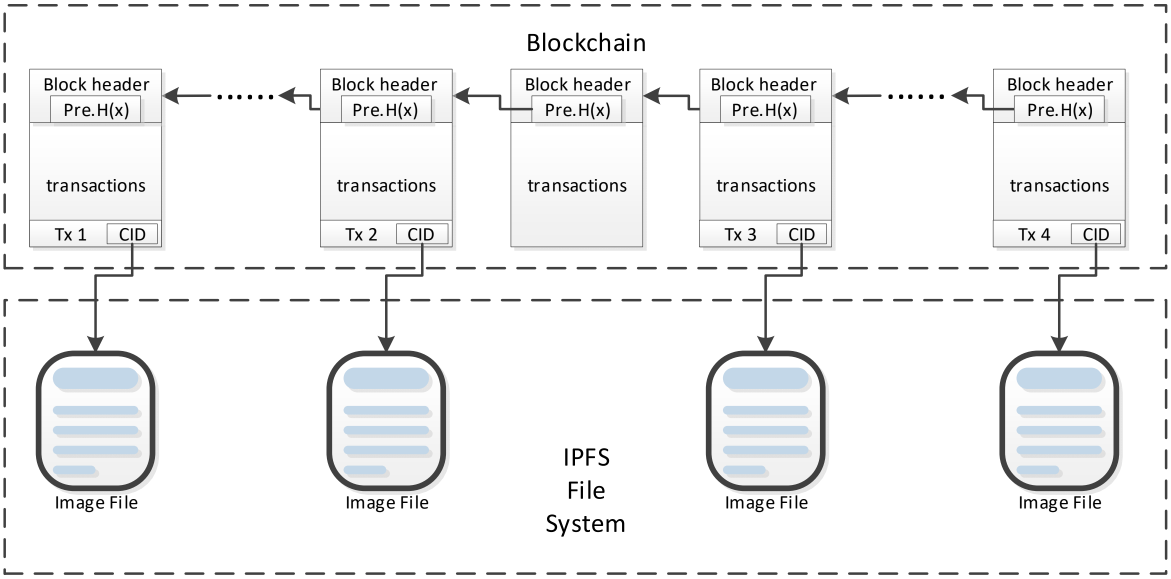

Figure 2. Traditional distributed storage structure based on IPFS system.

Since the decentralized distributed storage solution is adopted for verification, indexing, and raw data, this idea can solve the distributed storage problem of large amounts of raw data and eliminates centralized problems such as data loss and limited access in the traceability process. The complexity of the storage structure and the introduction of IPFS technology have solved the problem of storing gross raw data. Moreover, both blockchain technology and IPFS technology are based on the hash algorithm, making the content of files immutable and the entire traceability system more secure and credible [24,28].

However, the solution of "storing verification and index data on the chain and storing large amounts of raw data off-chain" has its drawbacks. Most of the current studies are carried out on the premise that the magnitude of stored data is small and only the efficiency of the single traceability request is taken into account. Moreover, considering the consensus problem in complex application scenarios, most existing studies also adopt the block of the alliance chain mode. For the practical application scenarios, especially in the consumer-grade agricultural product traceability system, the storage level of data may be millions or even tens of millions. In the consumption season, the traceability requests to be processed will be tens of millions, which poses a serious challenge to the traceability efficiency of the whole system. In the storage structure of "storing verification and index data on the chain, and storing large amounts of data off-chain", data retrieval and verification in the traceability process are completed on the blockchain [23–31]. In other words, the data retrieval and verification must be performed locally by full nodes. Especially in the blockchain mode that adopts the alliance chain, the number of full nodes is far less than the number of full nodes in the public chain mode, which makes the authoritative total nodes in the alliance chain process massive traceability requests in the consumption season. In order to bear such tremendous data indexing pressure, the data traceability system needs to be optimized to cope with higher-level data traceability requests.

Based on the review and analysis, this paper constructs an interactive traceability structure based on image evidence and uses images as an enhancement method to prevent fraudulent use of data. Furthermore, the "blockchain + IPFS" fully distributed data storage solution is adopted, and the data storage strategy of "storing verification and indexing data on the chain, and storing large amounts of unstructured data off-chain" is adopted. A specific storage structure for image evidence is formed by injecting pointers into the original image files. After receiving the traceability request, the traceability process will be carried out in the distributed storage certificate file structure rather than on the blockchain. The user obtains the transaction information and then verifies it on the blockchain to realize "traceability off-chain and verification on-chain". The separation of the traceability and verification processes is realized, thereby improving traceability efficiency and reducing the data retrieval pressure of full nodes in the data traceability process. Table 1 shows the difference between the proposed interactive traceability structure and related works.

Table 1. Comparison between the existing and proposed interactive traceability structure.

Ref.	Year	Application Scenario	Storage Structure	Centralized	Storage Capacity	Data Retrieval Mode	Traceability Request Processing Capability
[19]	2021	Agricultural products	blockchain + database	Y	medium	On-chain	low
[20]	2022	Image retrieval service	blockchain + cloud	Y	high	On-chain	low
[23]	2019	Food supply chain	blockchain + IPFS	N	high	On-chain	medium
[24]	2021	Artworks	multi-chain + IPFS	N	high	On-chain	medium
[25]	2019	Agricultural products	multi-chain + IPFS	N	high	On-chain	high
[26]	2021	Tea quality	blockchain + IPFS	N	high	On-chain	medium
[27]	2019	Digital asset	blockchain + index	N	low	On-chain	high
[28]	2021	Vehicle Data storage	blockchain + IPFS	N	high	On-chain	medium
[29]	2021	Industrial image	blockchain + IPFS	N	high	On-chain	medium
[30]	2021	BIM storage	blockchain + IPFS	N	high	On-chain	medium
[31]	2021	Medical images	blockchain + IPFS	N	high	On-chain	medium
us	2022	Agricultural products	Interactive traceability structure	N	high	Off-chain	ultra-high

3. Interactive Traceability Structure Design and Image Evidence Generation

In order to illustrate the interactive traceability structure based on image evidence proposed in this paper, we present the special storage structure design in Section 3.1. This part explains how the data are stored in the blockchain and IPFS system and how the relationship between the transaction and the image evidence is constructed. Then, in Section 3.2, we show the image evidence generation method used in this article. Of course, this method is not irreplaceable. As long as readability and tamper-proofness are ensured, they can be changed to meet the needs of different scenarios. Table 2 explains the attributes and abbreviations in this Section.

Table 2. Abbreviations and attributes explanation.

Attributes/Abbreviations	Explanation of Values
Tx	Tx is the abbreviation of transaction. It is the information that needs to be recorded at every stage of the product. In this paper, it includes the main content, the initiator of the transaction, the hash pointer to the previous stage block, a set of CIDs pointing to the files stored in the IPFS, and a hash pointer to the previous transaction. The Tx 1, Tx 2,... Tx n−1, Tx n, etc., appeared in the figures to indicate the sequence of transactions in the transaction chain. If those transactions in the figures, such as transcation1, transaction2, etc., are without a link in a certain block body, this means different transactions in one block may not be relevant.
Content	It is the text information of the transaction, including the current status of the product, the owner account, the buyer account, and the verification information.
Author	It is the initiator of this transaction, namely the editor of the transaction content and the uploader of the image evidence.
Pre.H(x)	It represents the hash value of the previous block.
Pre.TxID	It is a hash pointer, which is the hash value of the previous transaction.

CIDs	CIDs is a list of different image evidence CID, and there can be one or multiple CIDs in it. CID is the hash pointer pointing to the IPFS system, which the Author obtains after uploading the image evidence.
TxID	It is the current transaction's hash value.
Pointer	It is a set of pointers pointing to the previous stage of the transaction, and the block where the previous transaction is in.
Pre.setp Block Height	It is the block height of the block where the previous transaction is located.
Pre.Image evidence CIDs	It is the CID of the previous transaction's image evidence.
Raw file	It is the original image file, provided by the initiator of the current transaction.
Image evidence	After CID is obtained by uploading an image with a pointer, and is written to the blockchain, the file corresponding to the CID is called the image evidence.
Tag	It is a series of numbers determined by the encoding rule and used to distinguish the categories of information. In this paper, Tag = [11,21,31,255] 11 means the block height of the block where the previous transaction is located; 21 means the previous transaction's hash value; 31 means the CID of the previous transaction's image evidence; 255 is a terminator. When decoding the pointer, if the R value becomes 255, then the process is terminated.

3.1. Interactive Traceability Structure Design

Aiming at the problem that a huge amount of original data are difficult to store on-chain, this paper adopts the underlying data storage mode of "blockchain + IPFS" for the original data. In the existing on-chain and off-chain storage method, after the user uploads the image to IPFS, the CID returned by IPFS will be written into the transaction request. After the blockchain consensus is reached, the CID will be recorded on the blockchain. When a traceability request is generated or a user requests verification, the full node of the blockchain returns transaction information containing CID, whereby the user sends a request to IPFS to obtain the original file. This data storage method stores only the CIDs pointing to the file stored in IPFS on the blockchain for retrieval. The CIDs here refer to a CID set. How many files need to be recorded for a transaction, then how many CIDS are in this set. The proposed interactive traceability structure based on image evidence in this paper assumes that each transaction has at least one image evidence to enhance the relationship between transaction data and real entities. Therefore, each transaction will record at least one CID, as shown in Figure 3.

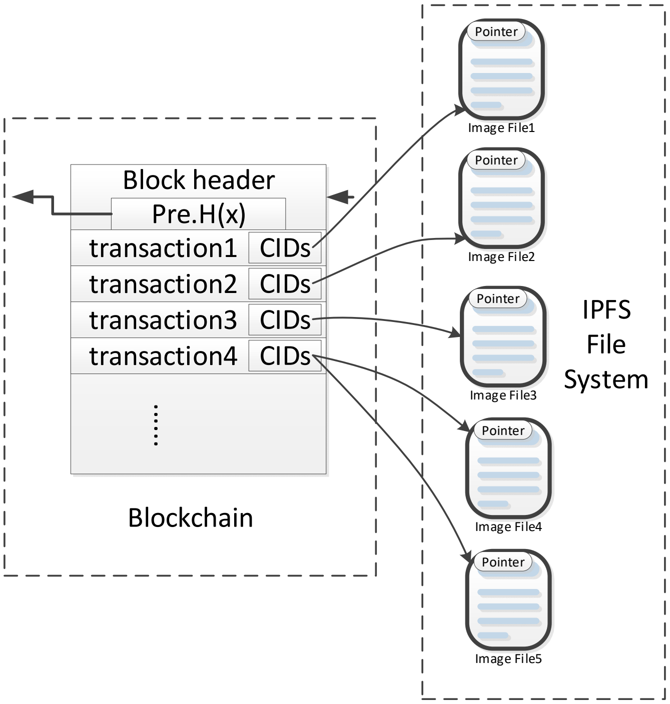

Figure 3. Pointer design from blockchain to IPFS.

Transaction (Tx) in this article refers to the information that needs to be recorded when the asset status information, such as its attribute, form, and owner, changes in the production cycle. Tx is defined as:

Tx = [Content, Author, Pre.TxID, CIDs, TxID]

Among them, Content represents the text information of the transaction, which records the current status of a specific product or asset, the owner account (Owner) and the buyer account (To), and related verification information (Input, Output), namely Content = [Status, Owner, To, Input, Output]; Author refers to the initiator of this transaction, namely the editor of the transaction content and the uploader of the transaction file; Pre.TxID is the hash pointer, which is the hash value of the previous transaction; CIDs is a list of different image evidence CID; there can be one or multiple CIDs in it. CID is the hash pointer pointing to the IPFS system, which the Author obtains after uploading the image evidence, that is, the CID = upload (Image evidence); TxID represents hashing the full content of the current transaction, namely TxID = Hash (Content, Author, Pre.TxID, CIDs).

Each transaction in the block records the corresponding CIDs. A transaction may point to one or multiple files stored in IPFS. The interactive traceability structure based on image evidence proposed in this paper assumes that each transaction at least storage one image evidence to enhance the connection between transaction data and entities in reality, so each transaction will be recorded at least with one CID.

The image-based interactive traceability structure proposed in this paper is different from the traditional IPFS-based off-chain storage structure. The former preprocesses the image evidence stored in the IPFS system. Moreover, it adds a pointer to the image pointing to the previous transaction, its block, and the corresponding image evidence, as shown in Figure 4.

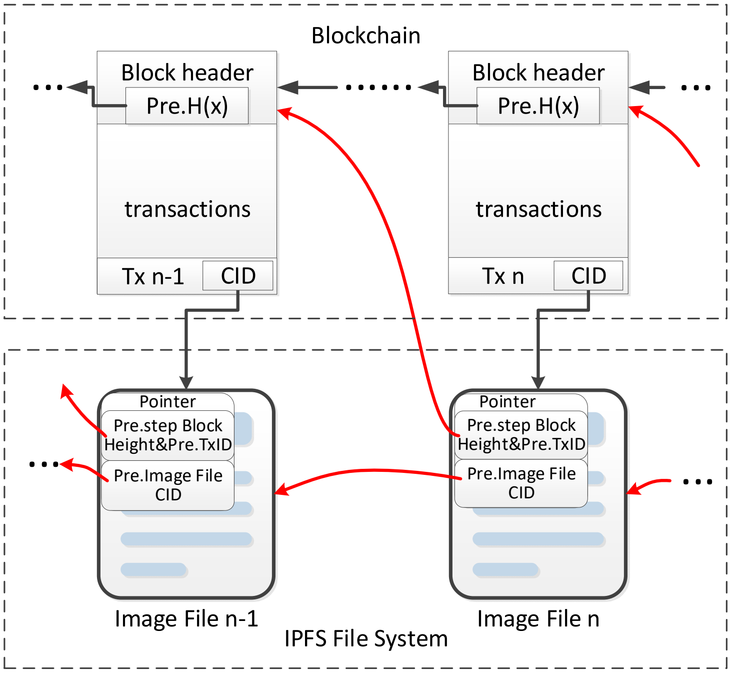

Figure 4. Pointer design of image evidence.

The reason for pointing to the previous image evidence is to form a specific file structure so that the traceability process can be carried out on the file structure. In order to locate the information on the blockchain after traceability in the file structure, a pointer from IPFS to the blockchain system is required. However, once the user uploads the image evidence, the pointer cannot be changed anymore; otherwise, the CID will change. The image upload is prior to block mining. When uploading the image, we are not sure which block the current transaction will be packed into. Therefore, the pointer to the blockchain system cannot point to the latest block but only to the block that has been added to the chain in the previous stage.

3.2. Image Evidence Generation

In our design, the pointer and the image are a complete file called image evidence, and the CID is also obtained based on this complete file (image evidence). Once the pointer content is changed, the CID of the image evidence will also be changed. Therefore, each image evidence also has strong tamper resistance, making the overall file structure credible.

Define the Image evidence as:

$$\text{Image evidence} = \text{Image(Pointer, Raw File)}$$

Raw File is the original image file; Pointer is a set of pointers pointing to the previous stage of the transaction, including the pointer to the previous transaction (Pre.TxID) and its block (Pre.setp Block Height) and the pointer to the corresponding file (Pre.Image evidence CIDs), namely Pointer = [Pre.setp Block Height, pre.TxID, Pre.Image evidence CIDs]; Finally, Pointer is written into the original image file (Raw File) to obtain the final image evidence.

The pointer writing process named Algorithm 1 is as follows:

Algorithm 1 encode the pointer into image file

```
Require: ImageFile, Pre_TxID, Tag
//When a node initiates a new transaction, the original image file of the transaction should be provided by the initiator, and the previous transaction should be known. Tag is a series of numbers determined by the encoding rule. As shown in Table 3, Tag = [11,21,31,255]
[block height, TxID, CIDs] ← search_pre_Tx(Pre_TxID)
//The storage node receives the request, locates the previous transaction in the blockchain in the local database, then records the height of the block, the CIDs, and the full content of the particular transaction.
for text in [block height, TxID, CIDs]:
    for i in text:
        index ← ord(i)
        rgb ← (Tag, (index & 0xFF00 >> 8, index & 0xFF))
        new_ImageFile ← ImageFile.chageRGB_sequentially(rgb)
//The 'block height', 'TxID', 'CIDs' of the previous transaction are needed to be encoded. Each letter of these strings is needed to be converted to Unicode and recorded separately by the G value and B value in each pixel. Then use the R value as a tag to identify whether the content is 'block height' or others.
api ← ipfshttpclient.connect()
// initiate the InterPlanetary File System
CIDs ← api.add(new_ImageFile)
// The initiator uploads the encoded ImageFile to the Inter Planetary File System and obtains the file's CID.
new_transaction ← generate_new_Tx(content,author,Pre_TxID,TxID,CIDs)
// Complete transaction information includes Content, Author, Pre_TxID, TxID, and CIDs. After the new_ImageFile is uploaded to the IPFS and obtains the CIDs, the initiator could send a request for the new transaction.
waiting for confirm
// Waiting for the miner node to verify, package the transaction, make consensus, and add it to the blockchain.
```

The process of reading the pointer in the image evidence named Algorithm 2 is as follows:

Algorithm 2 decode the pointer from image evidence

```
Require: CIDs,Tag
//When a user node requests a tracing source, the node may not have the raw image file but just some basic transaction information. After the full node returns the CIDs, the user first needs to use it to request the image evidence from the IPFS, and then decode the file for the previous transaction's information.
api ← ipfshttpclient. connect()
// Initiate the InterPlanetary File System
image evidence ← api.cat(CIDs)
for pixel in image evidence:
    R,G,B ← get_pixel_RGBvalue()
    if R != 255:
        compare R with Tag:
            index = chr((G<<8)+B)
            block height, TxID, CIDs ← content.append(index)
    else:
            break
// Decoding is the inverse operation of encoding. The G value and B value are spliced to form a letter. The pixels in the image are segmented according to the tag list and R value. Each segment is a complete piece of information. Tag is a series of numbers determined by the encoding rule. As shown in Table 3, Tag = [11,21,31,255] In tag, we define a terminator. In Table 3, the terminator's value is 255. If the R value equals the terminator, the algorithm will be stopped. Eventually, we obtained the pointer of the image evidence, including block height, TxID, and CIDs.
```

In this paper, the block height (Height) of the previous block, the TxID of the previous transaction, and the CIDs of the corresponding image stored in the IPFS are written into the file pointer. Using the method that converts each letter of the information text into Unicode encoding and splits it into two parts, then write them into each pixel's G and B values separately. Finally, to facilitate subsequent decoding, the corresponding tag is written in the R value of each pixel according to the different types of information, where the tag is an artificially defined value, as shown in Table 3. Because the pointer data are short and maybe only tens of bits, this method will only affect the first tens of pixels of the image file. Relative to the size of the HD image itself, we believe that the pointer will not affect the key information of the original image, as shown in Figure 5.

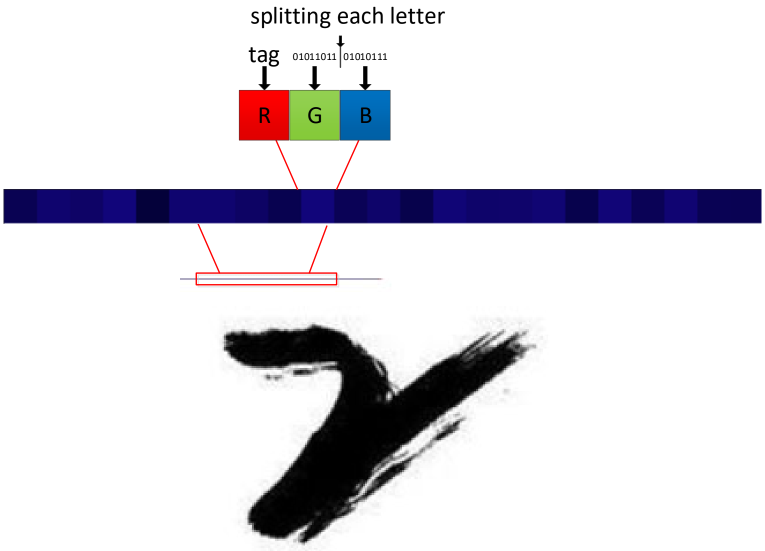

Figure 5. An example of pointer information writing.

In a certain image evidence, the details of pointer information are shown in Table 3.

Table 3. Example of image evidence pointer information.

Tag	Text Message	Text Category
11	QmSEZ7v71MbAqwDxsMWgGwd7CDuojfja4ENwMHfsqJ7QGx	CID
21	f7bf1ffd5650ff6732a066ba212df64e9f5eef1099eeeef553358e3ee69b0b0e	Pre.TxID
31	64	height
255	none	Terminator

It can be seen that this image evidence points to one historical image evidence, and the historical image evidence is stored in a block with a height of 64. Through the pointer information in the image evidence, the image evidence of the previous historical stage can be quickly located in the IPFS system. Furthermore, in the blockchain system, the previous transaction can also be quickly located according to the block height and the TxID so as to realize the rapid positioning of files, blocks, and transactions and facilitate the follow-up traceability work.

The interactive traceability structure based on image evidence proposed in this paper, through the pointer design and writing of the image evidence stored in IPFS distributed storage, forms a historical file structure of the image and a two-way pointer between the two systems of blockchain and IPFS. As a result, to "trace the source off the chain and verify on the chain" becomes possible. The pointer-designed "blockchain + IPFS" interactive traceability structure is shown in Figure 6.

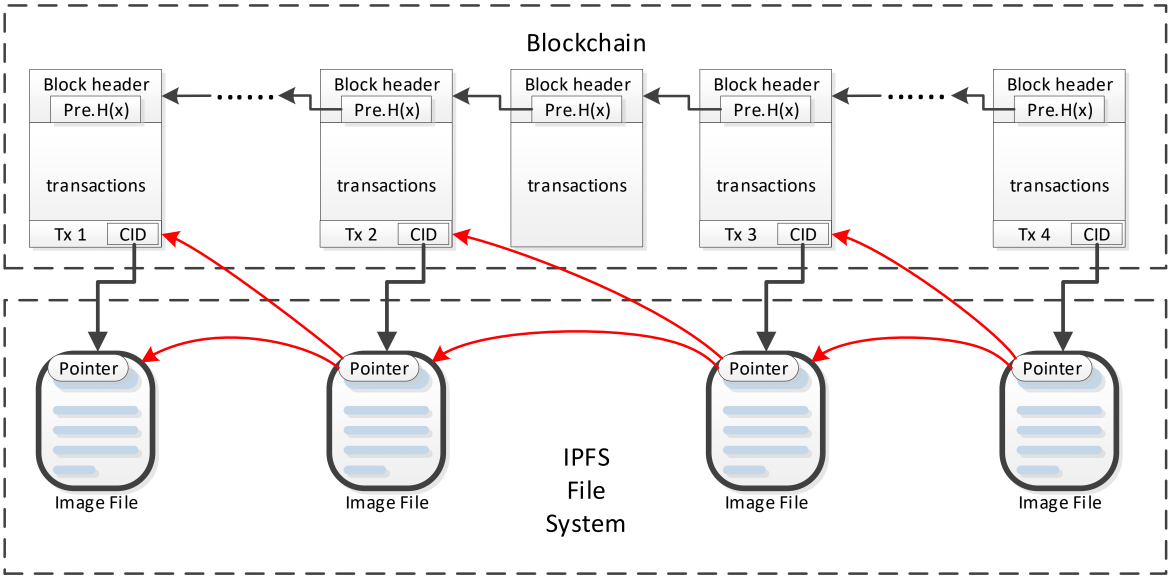

Figure 6. The interactive traceability structure based on image evidence.

4. Comparison with the Traditional Structure of Blockchain + IPFS

In the traditional structure of blockchain + IPFS, there is only a CID pointer from the blockchain to the IPFS system. After the full node of the blockchain receives the user's traceability request, the full node will query and verify the historical transaction information on the chain until all of the transactions are traced back and then return all the transaction verification information to the user. At the same time, the full node retrieves the CIDs in the transaction information and sends file requests to the IPFS system. It will query in the IPFS system according to the CIDs and return the image evidence to the user. Users can verify the credibility of the corresponding transaction and image evidence locally as required. Figure 7 shows the traditional structure traceability process of blockchain + IPFS (Traditional on-chain traceability process).

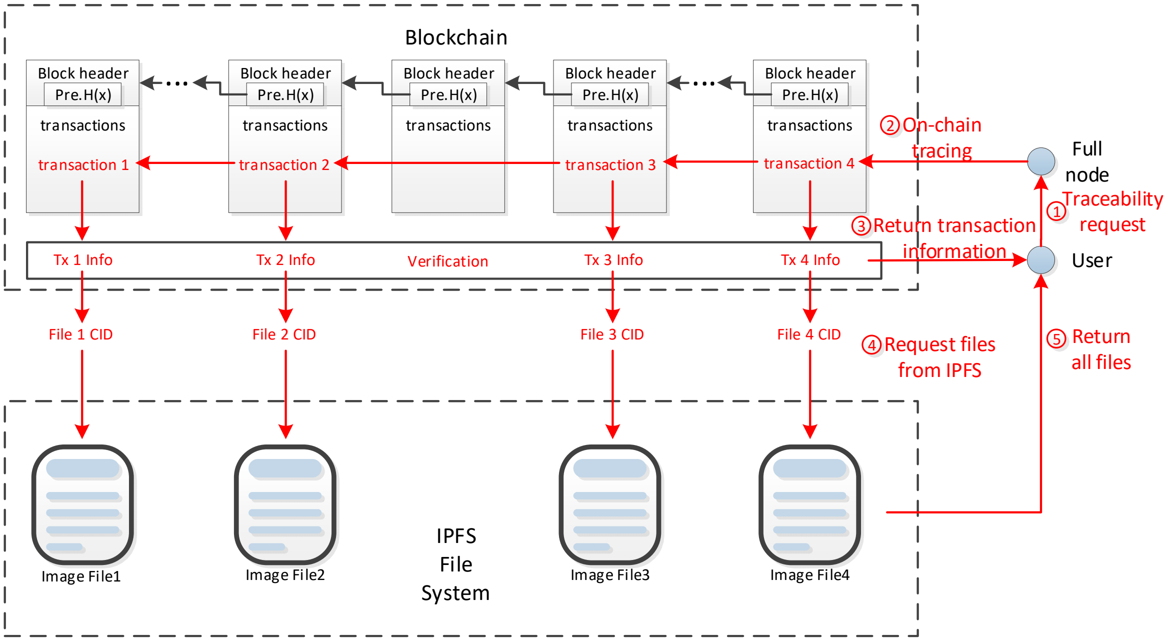

Figure 7. The traditional traceability process is completed by full nodes.

Moreover, the image-based interactive traceability structure proposed in this paper improves the pointers in the whole system and forms a blockchain + IPFS interactive traceability structure. In this structure, after the blockchain full node receives the user's traceability request, it only needs to retrieve the last transaction information on the blockchain and send the file acquisition request to the IPFS system based on the CIDs in the transaction information and return the image evidence. After that, the user can trace the historical image evidence according to the pointer in the current image evidence, obtain all the historical image evidence, and then send the verification request to full nodes according to the block information recorded in the file pointer. The full node quickly locates the transaction location according to block height, and TxID in the verification request retrieves the corresponding transaction information and verification data and delivers them back to the user. The user can verify the credibility of the corresponding transaction and image evidence as wished locally. Figure 8 illustrates this traceability process.

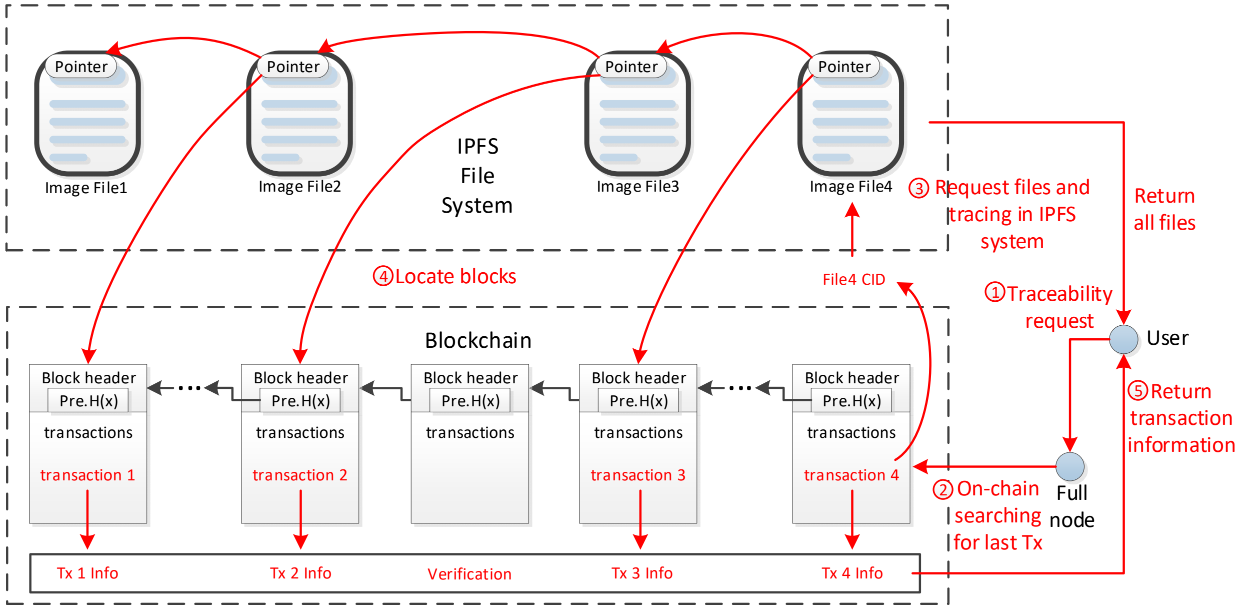

Figure 8. The proposed traceability process is completed by users themselves.

The traceability process under the interactive traceability structure named Algorithm 3 is as follows:

Algorithm 3 source tracing of a certain transaction

```
Require: Last TxID
//A user node asks to trace the source of a specific transaction. The user node provides the TxID of a specific transaction.
1: block height, CIDs, transactions chain ← search_last_Tx(Last TxID)
//The storage node receives the request, locates the last transaction on the blockchain in the local database, then records the height of the block, the CIDs, and the full content of the particular transaction.
2: api ← ipfshttpclient. connect()
// Initiate the Inter Planetary File System
3: while CIDs[−1]:
       image evidence ← api.cat(CIDs[−1])
       pre_height, pre_TxID, preCIDs ← decode(imagefiles)
       block height, CIDs, TxIDs ← list.append(pre_height, pre_TxID, pre_CIDs)
// Use CIDs to download image evidence, then decode the pointer in the image and renew the list of block height, CIDs transaction chain. There is no pre_CIDs in the image evidence's pointer until the original transaction is found.
4:transaction chain ← search_content(block height, TxIDs)
// Obtain full content of the transactions in specific blocks through TxIDs and block height.
5: return the full content
// The storage node sends the full content of the transaction chain and relative image evidence to the user node; the user node can verify the content as usual.
```

Obviously, in the scheme proposed in this paper, the main traceability and retrieval processes are performed by the users themselves in the IPFS system. The full node only retrieves and locates the last transaction when the traceability request is initially received, and all subsequent transaction locating is performed by the pointers of the image evidence in the IPFS system. The process of retrieval and traceability is transferred from the blockchain system to the IPFS system, but the blockchain system still retains the function of verifying legitimacy. The scheme realizes the separation of traceability and verification functions, reduces the pressure of full-node data retrieval, and effectively improves the traceability efficiency of the whole system, especially suitable for the application scenarios with intensive traceability requests, large data volume, and long time span.

5. Simulation Experiment

In this paper, the traceability efficiency of the proposed interactive traceability structure is verified by simulation experiments. The experimental environment is as follows: the operating system is window10, the memory size is 16 GB, the CPU is Intel(R) Core(TM) i7-10875H CPU @ 2.30 GHz, the IPFS version is 0.8.0, network downlink is 28.9 Mbps, and uplink is 29.19 Mbps. The blockchain simulation data set is generated by simulation of local nodes, and the traceability target transaction chain is a 4-level transaction chain. The transaction chain and the corresponding file structure are shown in Figure 9.

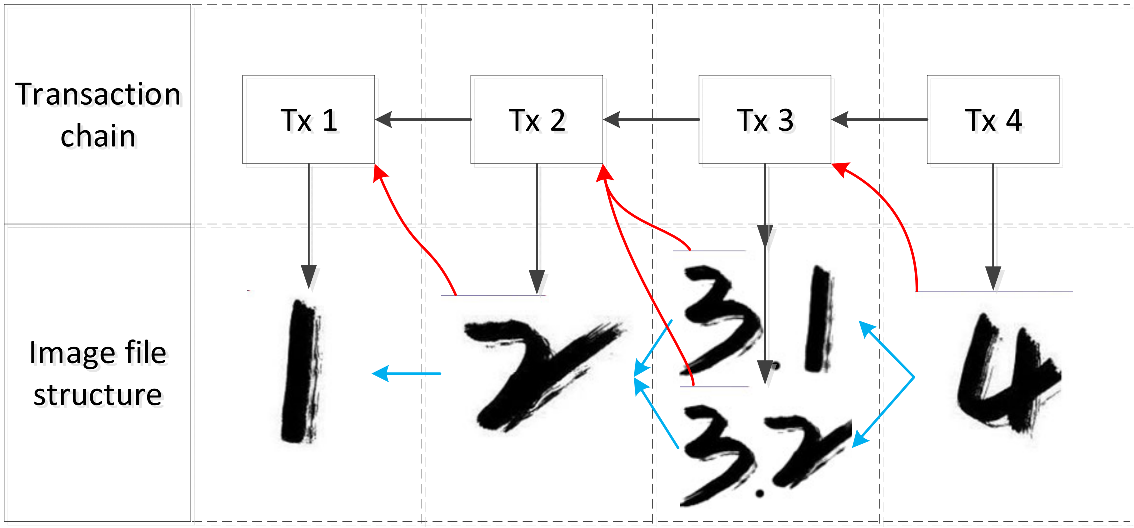

Figure 9. Target transaction chain and corresponding file structure.

We generate ten groups of simulation data with a total trading volume between 0.2 million and 2 million through simulation nodes, pack 200 transactions in each block, and randomly insert a target transaction in every quarter of the trading volume to ensure the step-by-step increment of the span between the first and the last transaction. In different amounts of simulated data, the span between the first and final transaction varies. The traceability effect of the traditional on-chain traceability method is compared with that of the interactive traceability structure. The total traceability time of the corresponding algorithms is shown in Figure 10.

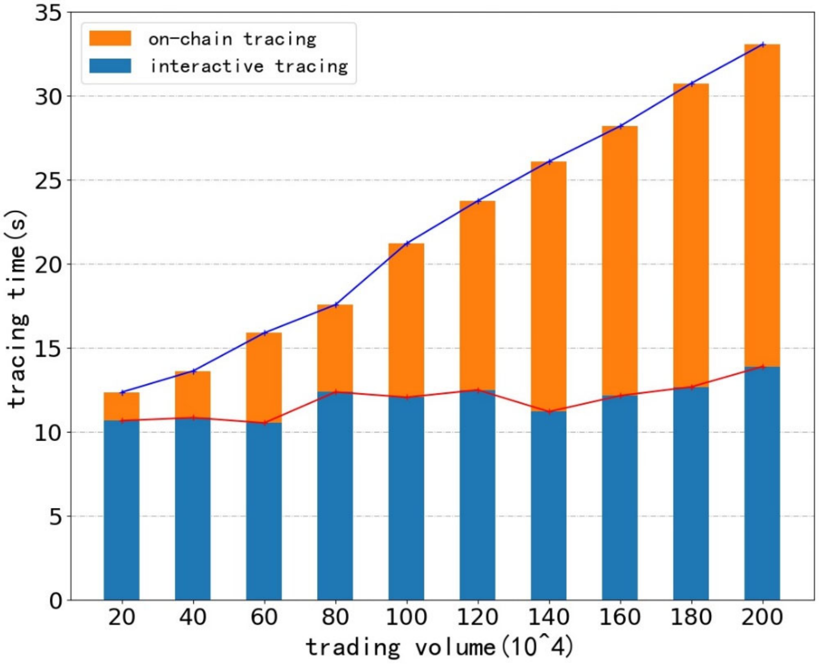

Figure 10. Comparison of efficiency between on-chain tracing and interactive tracing.

It can be seen that with the increase in trading volume, the total time consumption of the traditional on-chain traceability methods is also gradually increasing. In contrast, the overall time consumption of the interactive traceability structure shows no significant change. This phenomenon indicates that in terms of traceability efficiency, compared with interactive traceability, the traditional on-chain traceability method is greatly affected by the block span between the first transaction and the last transaction. The longer the block span is, the more pronounced the advantage of interactive traceability structure is. Figure 11 shows the time consumption of two different traceability processes throughout the traceability process.

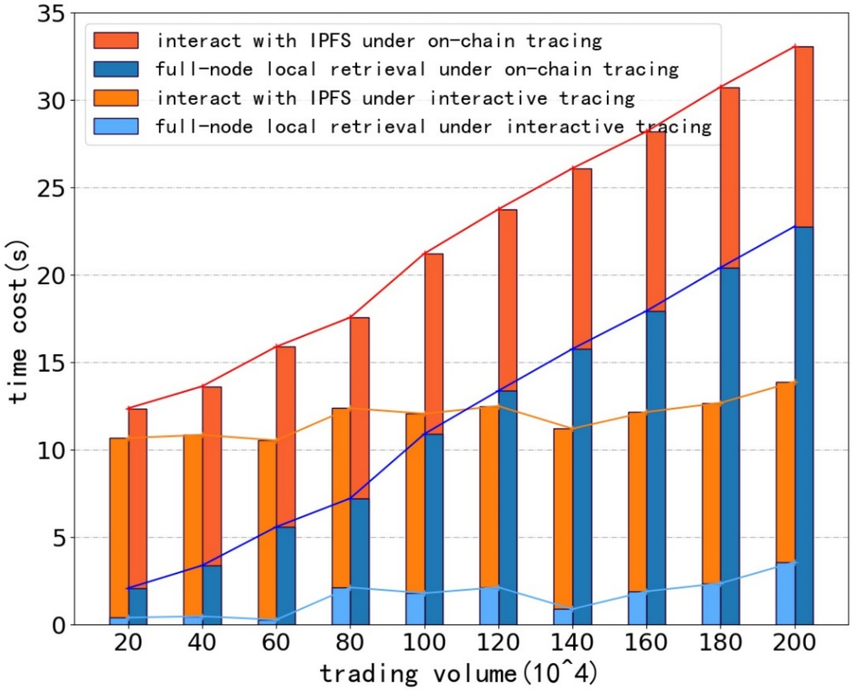

Figure 11. Comparison of traceability efficiency between two different traceability processes.

From Figure 11, it is easy to see that in terms of the interaction time with IPFS, the traditional traceability method and the interactive traceability structure proposed in this paper are the same, both in about 10 seconds. What mainly affects the elapsed time is the file size and the network conditions associated with the transaction chain.

What causes the increase in overall traceability time for traditional methods is the time spent on blockchain full-node local retrieval. Due to the lack of a corresponding traceability structure, traditional traceability methods (on-chain tracing) must traverse all blocks covered by the entire span of the transaction chain, and in the interactive traceability structure, except for the retrieval of the last transaction, the traceability of the rest of the transactions is completed by the user in the IPFS system. Based on the decoding of the file returned by the IPFS system, the block height is acquired, and the transaction is located. There is no need to traverse all the blocks locally in the full node, significantly reducing the local retrieval pressure. In the interactive traceability structure of 0.8 million to 2 million transactions, the fluctuation of the local node retrieval time is mainly caused by the long span between the last transaction block and the latest block, resulting in the fluctuation of retrieval time when locating the last transaction.

6. Conclusions

In order to meet the application requirements in complex scenarios, this paper constructs an interactive traceability structure based on image evidence, proposes the distributed traceability concept of “tracing off the chain and verification on the chain”, and puts forward a new blockchain traceability method so that the traceability system can cope with ultra-high level traceability requests. It provides a feasible scheme for the traceability of blockchain to be applied in the consumer-level field.

The experimental results show that, compared with the traditional blockchain traceability model, the traceability process proposed in this paper can be completed by the user, and the whole node only needs to complete the verification work. This traceability method not only reduces the data retrieval pressure of the full node but also greatly improves traceability efficiency, especially for the traceability efficiency of the long-span transaction chain, and the scheme only improves the file structure of distributed storage and proposes a new traceability process and algorithm, which does not involve changes to the original chain structure and the consensus process of the blockchain. This makes the solution have strong compatibility with the public, consortium, private chains, etc. It can be applied to more consumer-level traceability scenarios, not limited to the agricultural products mentioned in the article.

This study is not free from limitations. The goal of traceability in this paper is to obtain all the files and historical data of a certain transaction chain. However, in practical applications, consumers may not need all the information, and they only care about some key transactions. Moreover, traceability efficiency is also closely related to the interaction frequency. If the interaction is too frequent, the traceability efficiency will be greatly reduced. Therefore, we will focus on supporting traceability with simple semantic qualification and optimization of traceability structure to solve the problem that the interaction frequency may be too high.

Author Contributions: Z.S., F.B. and C.J. designed the study and conceived the manuscript. Z.S. carried out the simulation experiments. Z.S. and F.B. drafted the manuscript. Z.S., X.R., C.Y., C.X., Z.W. and F.B. were involved in revising the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Natural Science Foundation of Zhejiang Province (no. LQ20G010002), the National Natural Science Foundation of China (no. 71571162), the project of China (Hangzhou) cross-border electricity business school (no. 2021KXYJ06), and the Contemporary Business and Trade Research Center of Zhejiang Gongshang University (nos. XT202103 and XT202105).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest: The authors declare that there is no conflict of interest.

References

1. Cheney, J.; Chiticariu, L.; Tan, W.-C. Provenance in Databases: Why, How, and Where. *DBS* **2009**, *1*, 379–474. https://doi.org/10.1561/1900000006.
2. Qian, W.; Shao, Q.; Zhu, Y.; Jin, C.; Zhou, A. Research Problems and Methods in Blockchain and Trusted Data Management. *J. Softw.* **2018**, *29*, 150–159. https://doi.org/10.13328/j.cnki.jos.005434.
3. Wu, X.; Liu, P.; Wang, Z. Traceability System of Agricultural Products Based on Blockchain. *Comput. Appl. Softw.* **2021**, *38*, 42–48. https://doi.org/10.3969/j.issn.1000-386x.2021.05.007.
4. Dhall, S.; Dwivedi, A.D.; Pal, S.K.; Srivastava, G. Blockchain-Based Framework for Reducing Fake or Vicious News Spread on Social Media/Messaging Platforms. *ACM Trans. Asian Low-Resour. Lang. Inf. Process* **2022**, *21*, 8. https://doi.org/10.1145/3467019.
5. Pereira, F.; Crocker, P.; Leithardt, V.R.Q. PADRES: Tool for PrivAcy, Data REgulation and Security. *SoftwareX* **2022**, *17*, 100895. https://doi.org/10.1016/j.softx.2021.100895.
6. Ju, C.; Jiang, Y.; Bao, F.; Zou, B.; Xu, C. Online Rumor Diffusion Model Based on Variation and Silence Phenomenon in the Context of COVID-19. *Front. Public Health* **2022**, *9*, 788475. https://doi.org/10.3389/fpubh.2021.788475.
7. Kaushik, K.; Dahiya, S.; Singh, R.; Dwivedi, A.D. Role of Blockchain in Forestalling Pandemics. In Proceedings of the 2020 IEEE 17th International Conference on Mobile Ad Hoc and Sensor Systems (MASS), Delhi, India, 10–13 December 2020; pp. 32–37.
8. Xie, C.; Guo, H.-Y.; He, D.-F. Research on the Construction of Traceability System for E-Commerce Agricultural Products Quality and Safety in China Based on Blockchain. In Proceedings of the 4th Annual International Conference on Social Science and Contemporary Humanity Development (SSCHD 2018), Wuhan, China, 14–16 December 2018; Atlantis Press: Wuhan, China, 2019.
9. Nakamoto, S. Bitcoin: A Peer-to-Peer Electronic Cash System. Available online: https://bitcoin.org/bitcoin.pdf (accessed on 31 March 2022).
10. Ding, J.; Wu, Q. Research Review and Development Prospects of Food Supply Chain Traceability system based on blockchain. *Food Mach.* **2021**, *37*, 72–77. https://doi.org/10.13652/j.issn.1003-5788.2021.02.013.
11. Hao, Z.; Mao, D.; Zhang, B.; Zuo, M.; Zhao, Z. A Novel Visual Analysis Method of Food Safety Risk Traceability Based on Blockchain. *Int. J. Environ. Res. Public Health* **2020**, *17*, 2300. https://doi.org/10.3390/ijerph17072300.

12. Yuan, Y.; Wang, F. Blockchain: The State of the Art and Future Trends. *Acta Autom. Sin.* **2016**, *42*, 481–494. https://doi.org/10.16383/j.aas.2016.c160158.
13. Wu, H.; Cao, J.; Yang, Y.; Tung, C.L.; Jiang, S.; Tang, B.; Liu, Y.; Wang, X.; Deng, Y. Data Management in Supply Chain Using Blockchain: Challenges and a Case Study. In Proceedings of the 2019 28th International Conference on Computer Communication and Networks (ICCCN), Valencia, Spain, 29 July–1 August 2019; pp. 1–8.
14. Xu, C.; Ding, A.S.; Zhao, K. A Novel POI Recommendation Method Based on Trust Relationship and Spatial-Temporal Factors. *Electron. Commer. Res. Appl.* **2021**, *48*, 101060. https://doi.org/10.1016/j.elerap.2021.101060.
15. Bao, F.; Xu, W.; Feng, Y.; Xu, C. A Topic-Rank Recommendation Model Based on Microblog Topic Relevance & User Preference Analysis. *Hum.-Cent. Comput. Inf. Sci.* **2022**, *12*, 10. https://doi.org/10.22967/HCIS.2022.12.010.
16. Xu, C.; Liu, D.; Mei, X. Exploring an Efficient POI Recommendation Model Based on User Characteristics and Spatial-Temporal Factors. *Mathematics* **2021**, *9*, 2673. https://doi.org/10.3390/math9212673.
17. Bao, F.; Mao, L.; Zhu, Y.; Xiao, C.; Xu, C. An Improved Evaluation Methodology for Mining Association Rules. *Axioms* **2022**, *11*, 17. https://doi.org/10.3390/axioms11010017.
18. Sun, Z.; Zhang, X.; Xiang, F.; Chen, L. Survey of Storage Scalability on Blockchain. *J. Softw.* **2021**, *32*, 1–20. https://doi.org/10.13328/j.cnki.jos.006111.
19. Zhang, X.; Huang, Z.; Zhao, J.; Zou, H. Application of Agricultural Products Supply Chain Traceability Based on Multi-Blockchain. *J. Chongqing Univ. Technol. (Nat. Sci.)* **2021**, *35*, 172–179.
20. Peng, H.; Li, J.; Shi, Z.; Li, X. A Blockchain-based Verifiable Encrypted Image Retrieval Scheme. *Comput. Eng.* **2022**, *48*, 25–33+39. https://doi.org/10.19678/j.issn.1000-3428.0061419.
21. Huang, H.; Lin, J.; Zheng, B.; Zheng, Z.; Bian, J. When Blockchain Meets Distributed File Systems: An Overview, Challenges, and Open Issues. *IEEE Access* **2020**, *8*, 50574–50586. https://doi.org/10.1109/ACCESS.2020.2979881.
22. Benet, J. IPFS—Content Addressed, Versioned, P2P File System. Available online: https://arxiv.org/pdf/1407.3561v1.pdf (accessed on 31 March 2022).
23. Casino, F.; Kanakaris, V.; Dasaklis, T.K.; Moschuris, S.; Rachaniotis, N.P. Modeling Food Supply Chain Traceability Based on Blockchain Technology. *IFAC-PapersOnLine* **2019**, *52*, 2728–2733. https://doi.org/10.1016/j.ifacol.2019.11.620.
24. Yuan, J.; Wang, X. Art Blockchain Certificate Traceability Model Based on Three Chains. *Appl. Res. Comput.* **2021**, *38*, 2915–2918+2925. https://doi.org/10.19734/j.issn.1001-3695.2021.02.0037.
25. Wang, K.; Chen, Z.; Xu, J. Efficient Traceability System for Quality and Safety of Agriculture Products Based on Consortium Blockchain. *J. Comput. Appl.* **2019**, *39*, 2438–2443. https://doi.org/10.11772/j.issn.1001-9081.2019020235.
26. Gao, Q.; Yang, C.; Wu, X.; Zhao, Y.; Wang, Z.; Wu, Y.; Zhang, Y. Research on the Traceability System of Tea Quality and Safety Based on Blockchain. *J. Anhui Agric. Univ.* **2021**, *48*, 299–303. https://doi.org/10.13610/j.cnki.1672-352x.20210510.001.
27. You, Y.; Kong, L.; Xiao, Z.; Zheng, Y.; Li, Q. Hybrid Indexing Dcheme Supporting Blockchain Transaction Tracing. *Comput. Integr. Manuf. Syst.* **2019**, *25*, 978–984. https://doi.org/10.13196/j.cims.2019.04.021.
28. Ye, H.; Park, S. Reliable Vehicle Data Storage Using Blockchain and IPFS. *Electronics* **2021**, *10*, 1130. https://doi.org/10.3390/electronics10101130.
29. Kumar, R.; Tripathi, R.; Marchang, N.; Srivastava, G.; Gadekallu, T.R.; Xiong, N.N. A Secured Distributed Detection System Based on IPFS and Blockchain for Industrial Image and Video Data Security. *J. Parallel Distrib. Comput.* **2021**, *152*, 128–143. https://doi.org/10.1016/j.jpdc.2021.02.022.
30. Tao, X.; Das, M.; Liu, Y.; Cheng, J.C.P. Distributed Common Data Environment Using Blockchain and Interplanetary File System for Secure BIM-Based Collaborative Design. *Autom. Constr.* **2021**, *130*, 103851. https://doi.org/10.1016/j.autcon.2021.103851.
31. Jabarulla, M.Y.; Lee, H.-N. Blockchain-Based Distributed Patient-Centric Image Management System. *Appl. Sci.* **2021**, *11*, 196. https://doi.org/10.3390/app11010196.