

BLOCKCHAIN-BASED CONTENT PROCESSING METHOD AND APPARATUS, DEVICE, AND STORAGE MEDIUM

RELATED APPLICATION

[0001] This application claims priority to Chinese Patent Application No.
5 2019109322215, titled "BLOCKCHAIN-BASED CONTENT PROCESSING METHOD
AND APPARATUS, DEVICE, AND STORAGE MEDIUM" filed on September 29, 2019,
which is incorporated herein by reference in its entirety.

FIELD OF THE TECHNOLOGY

[0002] Embodiments of this application relate to the field of blockchain technologies,
10 and in particular, to a content processing method and apparatus, a device, and a storage
medium.

BACKGROUND OF THE DISCLOSURE

[0003] With the development of network technologies and user content, a large
amount of content data is usually generated on a network platform. Therefore, the processing
15 of the content data is increasingly in focus. Generally, a relatively complex data storage
system is adopted for storage of the content data, and consequently, a series of problems will
arise, such as a complex read process and a long read time, resulting in low efficiency of a
content data processing procedure. Therefore, a content processing method of higher
efficiency is urgently needed.

20 SUMMARY

[0004] The embodiments of this application provides a blockchain-based content
processing method, an apparatus thereof, a device, and a storage medium, to provide a content
processing manner of higher efficiency. The technical solutions are as follows.

[0005] According to an aspect, a blockchain-based content processing method is
25 provided, including:

obtaining, by a content server, fingerprint information of newly published
content data;

transmitting, by the content server, the fingerprint information of the content
data to a blockchain system; and

storing, by the content server by using the fingerprint information of the content data as an index, the content data into a target file system in response to receiving a storage success message returned by the blockchain system.

[0006] According to an aspect, a blockchain-based content processing method,

5 applicable to a node device in a blockchain system, is provided, including:

receiving, by the node device, fingerprint information of content data, and broadcasting in the blockchain system;

transmitting, by the node device, a verification success message in response to that no fingerprint information whose similarity with the fingerprint information of the
10 content data is greater than or equal to a target threshold exists in fingerprint information stored in a blockchain of the blockchain system; and

storing, by the node device, a block including the fingerprint information of the content data into the blockchain of the blockchain system in response to reaching, by the blockchain system, a consensus on the fingerprint information of the content data.

15 **[0007]** According to an aspect, a blockchain-based content processing apparatus is provided, including:

a fingerprint obtaining module, configured to obtain fingerprint information of newly published content data;

20 a transmitting module, configured to transmit the fingerprint information of the content data to a blockchain system; and

a storage module, configured to store the content data into a target file system by using the fingerprint information of the content data as an index in response to receiving a storage success message returned by the blockchain system.

[0008] According to an aspect, a blockchain-based content processing apparatus is
25 provided, including:

a receiving module, configured to receive fingerprint information of content data transmitted by a content server;

a transmitting module, configured to broadcast in a blockchain system,

30 the transmitting module being further configured to transmit a verification success message in a case that existing fingerprint information in a blockchain of the

blockchain system includes no fingerprint information whose similarity with the fingerprint information of the content data is greater than a target threshold; and

a storage module, configured to store a block including the fingerprint information of the content data into the blockchain of the blockchain system in a case that a consensus on the fingerprint information of the content data is reached based on the verification success message in the blockchain system.

[0009] According to an aspect, a computer device is provided, including a processor and a memory, the memory storing at least one piece of program code, the at least one piece of program code being loaded and executed by the processor to implement the blockchain-based content processing method according to any one of the foregoing.

[0010] According to an aspect, a computer-readable storage medium is provided, storing at least one piece of program code, the at least one piece of program code being loaded and executed by a processor to implement the blockchain-based content processing method according to any one of the foregoing.

[0011] The technical solutions provided in the embodiments of this application bring at least the following beneficial effects.

[0012] Content data and fingerprint information thereof are stored in combination by providing a blockchain system and a target file system, and the fingerprint information may be persistently stored by using the blockchain system. In addition, by using the fingerprint information as a connection between the blockchain system and the target file system, and by taking advantage of trustworthiness and reliability characteristics of data storage in the blockchain system, a storage manner with low data storage pressure may be provided, so that the data storage pressure is significantly reduced, and data processing efficiency can be improved.

BRIEF DESCRIPTION OF THE DRAWINGS

[0013] To describe the technical solutions of the embodiments of this application more clearly, the following briefly introduces the accompanying drawings for describing the embodiments. Apparently, the accompanying drawings in the following description show merely some embodiments of this application, and a person of ordinary skill in the art may derive other drawings from the accompanying drawings without any creative efforts.

[0014] FIG. 1 is a schematic diagram of a plurality of blocks of a blockchain according to an embodiment of this application.

[0015] FIG. 2 is a schematic diagram of an implementation environment of a blockchain-based content processing method according to an embodiment of this application.

[0016] FIG. 3 is a flowchart of a blockchain-based content processing method according to an embodiment of this application.

5 **[0017]** FIG. 4 is a schematic diagram of a blockchain-based content processing apparatus according to an embodiment of this application.

[0018] FIG. 5 is a schematic diagram of a blockchain-based content processing apparatus according to an embodiment of this application.

10 **[0019]** FIG. 6 is a schematic structural diagram of a computer device according to an embodiment of this application.

DESCRIPTION OF EMBODIMENTS

[0020] To make the objectives, technical solutions, and advantages of this application clearer, the following further describes implementations of this application in detail with reference to the accompanying drawings.

15 **[0021]** First, several terms involved in this application are explained.

[0022] Blockchain system refers to a system that stores data by using a block chain technology. Blockchain technology, also referred to as a distributed ledger technology, is an Internet database technology and features decentralization, openness, and transparency, thereby enabling all people to participate in database recording. The blockchain technology
20 may be a distributed infrastructure and computing paradigm in which data is verified and stored by using a block-chain data structure, data is generated and updated by using a distributed node consensus algorithm, safety of data transmission and access is ensured by using cryptology and programming and operation on data are performed by using a smart contract including automation script code. The data stored in the blockchain system have high
25 stability and reliability, and the blockchain system is a database that cannot be tampered with and is reliable. Therefore, compared with a conventional manner in which storage is performed by using a content distribution server or a content server, storing data related to content data by using a blockchain system not only can ensure safety and reliability of the data, but also can provide data query and data comparison functions based on reliable storage.

30 **[0023]** The blockchain system includes a plurality of node devices, and data can be shared between node devices. Each node device can receive input information during normal

operation, and maintain shared data in the blockchain system based on the received input information. To ensure an information exchange in the blockchain system, information connections exist among all the node devices in the blockchain system, and information transmission may be performed among the node devices through the foregoing information connections. For example, when any node device in the blockchain system receives input information, the other node devices in the blockchain system obtain the input information according to a consensus algorithm, and store the input information as data in shared data, so that identical data is stored in all the node devices in the blockchain system.

[0024] For each node device in the blockchain system, there is a node device identifier corresponding to the node device. Each node device in the blockchain system stores node device identifiers of other node devices in the blockchain system, so as to subsequently broadcast a generated block to other node devices in the blockchain system according to the node device identifiers of other node devices. Each node device may maintain a node device identifier list as shown in Table. 1, and node device names and node device identifiers are correspondingly stored in the node device identifier list. Optionally, the node device identifier is an Internet Protocol (IP) address or any other information that can be used for identifying the node device. In Table 1, description is made by using IP addresses only for example.

Table 1

Node device name	Node device identifier
Node device 1	117.114.151.174
Node device 2	117.116.189.145
...	...
Node device N	119.123.789.258

[0025] All node devices in the blockchain system are equipped with a same blockchain. The blockchain includes a plurality of blocks. Referring to FIG. 1, a genesis block includes a block header and a block body. The block header stores an input information eigenvalue, a version number, a timestamp, and a difficulty value, and the block body stores input information. A next block of the genesis block uses the genesis block as a parent block, and also includes a block header and a block body. The block header stores an input information eigenvalue of a current block, a block header eigenvalue of the parent block, a version number, a timestamp, and a difficulty value. By analogy, block data stored in each

block in the blockchain is associated with block data stored in a parent block, thereby ensuring the security of the input information in the blocks.

[0026] Block height refers to the quantity of blocks linked to a main chain, that is, the quantity of blocks connected to a blockchain.

5 **[0027]** Consensus algorithm refers to a mathematical algorithm for building trust and obtaining rights and interests between different node devices in a blockchain system. In a blockchain system, a transaction can be verified and confirmed in a short time through voting of special node devices. For a transaction, if several node devices of no interest to each other can reach a consensus, it can be considered that all node devices in the system can also reach
10 a consensus.

[0028] FIG. 2 is a schematic diagram of an implementation environment of a blockchain-based content processing method according to an embodiment of this application. The implementation environment includes at least one content server 200, a blockchain system 210, and a target file system 220.

15 **[0029]** The content server 200 is a server of any content platform. A user publishes content data on the content platform, for example, publishes a short video or a webpage article. When obtaining newly published content data, the content server 200 stores the content data by using the blockchain system 210 and the target file system 220. Optionally, the content server 200 further has a function of providing users with customized services, for
20 example, recommending content data for the users.

[0030] The blockchain system 210 includes a plurality of node devices. The plurality of node devices includes a management node device used as a processing bridge between the blockchain system and the target file system. The target file system 220 includes a plurality of storage devices. Each storage device can be used for content storage to support content
25 writing and reading of the target file system. The blockchain system 210 is used for storing fingerprint information of content data stored in the target file system 220. Compared with the content data, the fingerprint information has a smaller data volume. Therefore, node storage pressure of a blockchain system can be alleviated, so that each node device can store more information. For example, for the target file system 220, when content data is stored, a
30 key-value pair is used for storage, key being fingerprint information and value being content data, so that fast storage is provided.

[0031] In a possible implementation, a node device in the blockchain system 210 is provided as a server, for example, a server, a server cluster including a plurality of servers, or a cloud computing service center. Each storage device in the target file system 220 is a storage device in any storage format, which is not limited in the embodiments of this application.

[0032] A communication connection can be established between the content server 200, the blockchain system 210, and the target file system 220 by using a wired or wireless network. The wired network is a metropolitan area network, a local area network, a fiber optic network, or the like; and the wireless network may be a mobile communication network or a wireless fidelity (WiFi) network.

[0033] A person skilled in the art is to understand that, the foregoing content server 200, the blockchain system 210, and the target file system 220 are merely examples. Other existing systems or systems possibly appearing in future, if applicable to this application, are also to fall within the protection scope of this application and are included herein by reference.

[0034] Based on the implementation environment shown in FIG. 2, an embodiment of this application provides a blockchain-based content processing method by using an example in which the method is applied to interaction processes between a content server, a blockchain system, and a target file system. As shown in FIG. 3, the method according to this embodiment of this application includes the following steps 301 to 308.

[0035] In step 301, a content server obtains fingerprint information of newly published content data.

[0036] The content data is any type of web-published content, for example, text data or multimedia data. For example, the newly published content data refers to one or more pieces of content data of which a publishing time is nearest to the current time. In another example, the newly published content data refers to one or more pieces of content data of which a publishing time falls within a most recent time period. This is not limited in the embodiments of this application.

[0037] The content server can obtain content data published by a user in a login state or as a visitor. The content server stores the content data, and can further perform publishing based on the stored content data, to ensure normal and quick spreading of the content data.

After storage, the content server may generate fingerprint information of the content data based on a target eigenvalue algorithm and the content data.

[0038] In a possible implementation, step 301 may include: obtaining, according to a publishing source of the content data, a target eigenvalue algorithm matching the publishing source; and generating the fingerprint information of the content data based on the target eigenvalue algorithm and the content data. The publishing source refers to a publishing platform, a publishing sector on a platform, or the like of the content data. Different publishing platforms or publishing sectors may have different characteristics in terms of content and form. Therefore, for the different characteristics, corresponding eigenvalue algorithms are provided, to ensure that the generated fingerprint information can be distinguished more clearly.

[0039] In a possible implementation, the content server generates fingerprint information of any content data by using a same eigenvalue algorithm. In another possible implementation, the content server generates, by using different eigenvalue algorithms, fingerprint information of pieces of content data having different sources. This is not limited in the embodiments of this application. For example, the eigenvalue algorithm is a hash algorithm.

[0040] In step 302, the content server obtains a remainder between the fingerprint information and the quantity of node devices in the blockchain system.

[0041] The content server and the blockchain system cooperate with each other. Therefore, the content server can learn about the quantity of current node devices in the blockchain system. To disperse storage pressure, the content server can determine, according to the fingerprint information and the quantity of node devices in the blockchain system, a node device to which the fingerprint information of the content data is to be transmitted. For example, a processing manner of step 302 is adopted to determine a target node device by means of a remainder.

[0042] In step 303, the content server determines a target node device, and transmits the fingerprint information of the content data to the target node device in the blockchain system, a node device identifier of the target node device being the remainder.

[0043] For a blockchain system, there is a unique node device identifier corresponding to each node device. For example, the node device identifier is a numerical number. In this case, after determining the remainder between the fingerprint information and the quantity of

node devices, the content server transmits the fingerprint information by using the remainder as a selected node device identifier.

[0044] For example, the fingerprint information is a series of 64-bit harsh integers calculated and generated based on the Message-Digest Algorithm 5 (MD5) encryption algorithm. A remainder thereof is obtained through calculation according to the following formula (1):

$$\text{id} = \text{hash}(\text{content}) \% N \quad (1),$$

where id is the remainder obtained by calculation, that is, the node device identifier of the target node device, hash() represents a hash function, content represents the content data, % represents an remainder arithmetic, N represents the quantity of node devices in the blockchain system, and N is a positive integer greater than 3.

[0045] It should be noted that in this embodiment of this application, determining the target node device by the content server is used as an example for description. In some possible implementations, the content server directly transmits fingerprint information to the blockchain system. After the fingerprint information is received, a management node or a gateway node device of the blockchain system determines the foregoing target node device based on the fingerprint information, and further, the management node device or the gateway node device transmits the fingerprint information to the target node device. A specific implementation of the foregoing process is not limited in the embodiments of this application.

[0046] In step 304, the target node device in the blockchain system receives the fingerprint information of the content data, and broadcasts in the blockchain system.

[0047] For any node device in the blockchain system, when receiving the fingerprint information of the content data, the node device can spread the fingerprint information to each node device in the blockchain system by broadcasting in the blockchain system.

[0048] In step 305, a verification success message is transmitted in response to that the target node device determines that existing fingerprint information in a blockchain of the blockchain system includes no fingerprint information whose similarity with the fingerprint information is greater than a target threshold.

[0049] Any node device in the blockchain system can query existing fingerprint information in the blockchain after receiving the fingerprint information. The existing fingerprint information is fingerprint information already stored in the blockchain, and is fingerprint information of published content data on the content server. A similarity is

calculated based on the existing fingerprint information and the received fingerprint information. When a similarity between the received fingerprint information and any piece of existing fingerprint information is greater than the target threshold, it is determined that the fingerprint information is successfully verified, and a verification success message is
5 broadcast in the blockchain system.

[0050] In the foregoing process, because fingerprint information is used for uniquely identifying a piece of content data, a similarity between fingerprint information can reflect whether pieces of content data are the same or similar, thereby determining whether same or similar content data has been published. For pieces of content data that are the same or
10 similar, it is very likely that there are cases such as plagiarism, that is, current intellectual property thefts such as article spinning. Therefore, through the query and comparison, it is determined whether such behaviors as the foregoing plagiarism have occurred, to ensure interests of originators.

[0051] In step 306, the target node device stores a block including the fingerprint
15 information of the content data into the blockchain of the blockchain system in response to reaching a consensus on the fingerprint information of the content data based on the verification success message in the blockchain system.

[0052] The reaching a consensus refers to that a target quantity of node devices in the blockchain system determines, through verification, that the fingerprint information is not
20 stored on the node devices. For the blockchain system, consensus is performed on the foregoing fingerprint information based on any consensus mechanism. For example, consensus is performed by using a representative node device. That is, a representative node device in the blockchain system performs verification and transmits a verification success message. When an enough quantity of representative node devices transmit a verification
25 success message, it is determined that a consensus on the fingerprint information is reached. In another example, all node devices participate in consensus. That is, when receiving the fingerprint information, each node device in the blockchain system performs verification and transmits a verification success message. When an enough quantity of node devices transmit a verification success message, it is determined that a consensus on the fingerprint information
30 is reached.

[0053] It should be noted that the foregoing block including the fingerprint information of the content data may be generated by any node device and broadcast in the blockchain system, or may be generated by the foregoing representative node device and

broadcast in the blockchain system, so that each node device can receive the block and stores the block, thereby ensuring consistency of data in the blockchain system.

[0054] Optionally, the block including the fingerprint information of the content data further includes link address information and a content update record of the content data. The link address information is used for uniquely identifying the fingerprint information of the content data in the blockchain system. When a user terminal accesses the content data through the link address information, the content server transmits an access request to the blockchain system based on the link address information. The blockchain system finds through query, based on the link address information, the fingerprint information stored in the block, and searches the target file system by using the fingerprint information as an index, to obtain corresponding content data stored in the target file system. The content update record is used for indicating a publishing time, updated content, and the like of the content data, which is not limited in the embodiments of this application.

[0055] Through the foregoing consensus mechanism of the blockchain system, publication of content data and any modification to the content data can both be recorded on the blockchain, so that the content data can be effectively protected from being illegally tampered with in a spreading process. Even if the content data is tampered with, the tampering can be traced, where a time of the tampering and a person making the tampering can be tracked by using the data on the blockchain, thereby achieving the purpose of maintaining the content data.

[0056] Further, when consensus in the blockchain system fails, a storage failure message is transmitted to the content server. The storage failure message is used for indicating that the content data may be plagiarized content.

[0057] Optionally, the block including the fingerprint information of the content data may further include user feedback information. For the user feedback information, when any node device in the blockchain system receives user feedback information of any user for second content data, second fingerprint information of the second content data and the user feedback information for the second content data are correspondingly stored in the blockchain. By storing the user feedback information in the blockchain, the user feedback information is recorded in a block and spread to the entire blockchain system.

[0058] The user feedback information refers to feedback information collected based on the published content data and is of any feedback type such as evaluation feedback, scoring feedback, or reporting feedback. The evaluation feedback is comment information,

reply information, or the like. The scoring feedback is specific scoring information, and the reporting feedback is complaint information, reporting information, or the like against the published content data. In a possible implementation, the content server provides, through an access page of the content data, a page for accessing a questionnaire or the like, a feedback menu, providing a feedback entry for users to submit feedback information to the content server, for example, "Thumb-up", "Report" or "Dislike". The feedback mechanism can be used as an effective supplement to the consensus mechanism, and provide an important data reference for protecting content data, promoting positive feedback and the like.

[0059] Based on the foregoing user feedback information, a blockchain-based user incentive mechanism is provided, which can stimulate users to participate in performing feedback for content data, so that more accurate and credible feedback can be based on, and users can further be stimulated to provide content data of higher quality. Specifically, the user feedback information-based processing includes at least one of (1) to (3) described below.

[0060] (1) Accumulate, based on the user feedback information, a participation score for a first target user submitting the user feedback information, and reward the first target user based on the accumulated participation score.

[0061] A corresponding reward is given for participation of a user, to stimulate the user to participate. A user identifier of the first target user submitting the user feedback information is extracted from the user feedback information, and a participation score is accumulated for the first target user. For example, each time the first target user submits user feedback information, a target participation score is accumulated for the first target user. When the participation score of the first target user reaches a reward score, the first target user is rewarded.

[0062] (2) When it is determined that third content data fails to conform to a target rule, accumulate, based on user feedback information for the third content data, a contribution score for a second target user submitting positive user feedback information, and reward the second target user based on the accumulated contribution score.

[0063] A corresponding reward is given for effective feedback of a user, to stimulate the user to participate and increase effective feedbacks. A user identifier of the second target user submitting the user feedback information is extracted from the user feedback information, and a contribution score is accumulated for the second target user. For example, each time the second target user submits user feedback information, a target contribution

score is accumulated for the second target user. When the contribution score of the first target user reaches a reward score, the second target user is rewarded.

[0064] Effective user feedback information refers to user feedback information that has positive effect on identifying illegal content or providing a reference for another user.

5 Therefore, a contribution score is set for quantifying such positive effect. For example, if any content data is determined as illegal information, user feedback information corresponding to the content data is obtained, and a user submitting illegality reporting information in the user feedback information is determined and rewarded. Understandably, in a case that content data is determined as legal information, user feedback information corresponding to the content
10 data is obtained, and a user submitting illegality reporting information in the user feedback information is determined and punished. For example, an account of the user is banned.

[0065] In practice, the foregoing two incentive manners for participation and rewarding can be combined into an overall incentive manner. For example, an incentive function is defined as $G(U) = f(\text{accumulated participation score, accumulated positive contribution score})$, where $G(U)$ represents a right allocation index score of a user U . A
15 criterion for setting a value of the foregoing incentive function includes: the higher the accumulated participation score is, the larger the value of $G(U)$ is; and the higher the accumulated contribution score is, the larger the value of $G(U)$ is. Through such a combined incentive manner, when a user is stimulated to participate, the user is also stimulated to
20 provide more accurate feedbacks, thereby avoiding cases such as frame-up. In practice, the foregoing two factors, namely, the participation score and the contribution score, are both positively correlated to $G(U)$, and a final value of $G(U)$ is obtained by means of weighted summation.

[0066] It should be noted that the foregoing reward is a virtual reward or a material
25 reward given for a user and is flexibly specified according to a platform, for example, a point coupon reward or a VIP user qualification reward. In addition, for a blockchain system, the foregoing scores of each user can be periodically counted, to achieve the purpose of regular incentive.

[0067] (3) Accumulate, according to received user feedback information, a content
30 quality score for a third target user submitting the content data, and when the content quality score of the content data is lower than a target quality score, process at least one of the content data and the third target user.

[0068] Both of (1) and (2) described above are about incentive measures for users submitting user feedback information. For the third target user submitting the content data, quality evaluation is performed on the content published by the third target user based on feedback information thereof. For any piece of content data, user feedback information for the content data in a blockchain system is obtained. According to a feedback type corresponding to the user feedback information, a score accumulation criterion corresponding to the feedback type is obtained, so that a content quality score can be accumulated based on the user feedback information and the score accumulation criterion. For example, when the feedback type is evaluation feedback, a quantity of pieces of evaluation information is counted. The score accumulation criterion is a correspondence relationship between the quantity and a score, and a current accumulated content quality score is determined based on the quantity of pieces. When the user feedback information is a positive feedback, for example, Like or Thumb-up, the accumulated content quality score is a positive number. When the user feedback information is a negative feedback, for example, Dislike or Thumb-down, the accumulated content quality score is a negative number, or the accumulated content quality score is a score to be subtracted this time. In practice, another calculation manner can alternatively be used. That is, for any type of user feedback information, accumulation is performed, to obtain a positive-number content quality score. Content quality is measured by using a specific score as a cutoff score. When the feedback type is reporting feedback, a quantity of times of reporting is counted. The score accumulation criterion is a correspondence relationship between the quantity of times of reporting and a score, and a current accumulated content quality score is determined based on the quantity of times of reporting. When the feedback type is evaluation feedback, a quantity of pieces of evaluation information is counted. The score accumulation criterion may be a correspondence relationship between the quantity of pieces and a score, and a current to-be-subtracted content quality score is determined based on the quantity of pieces. When the feedback type is interaction feedback, for example, the user feedback information is Like, Forward, or the like, a quantity of times of interaction is counted. The score accumulation criterion is a correspondence relationship between the quantity of times of interaction and a score, and a current accumulated content quality score is determined based on the quantity of times of interaction.

[0069] In a possible implementation, to ensure quantity of content data presented on the platform, content data can further be processed based on the foregoing content quality score. For example, in a case that the content quality score is lower than a target quality score,

the content data is deleted from the target file system. The process specifically includes:
periodically obtaining a content quality score corresponding to fingerprint information of each
piece of content data in the blockchain system; transmitting a deleting request to the target file
system in a case that a content quality score of any piece of content data is lower than a target
5 quality score, the deleting request carrying fingerprint information of the content data; and
deleting, by the target file system after receiving the deleting request, the content data.

[0070] In step 307, the target node device transmits a storage success message to the
content server.

[0071] After the block is added to the blockchain, the target node device transmits a
10 storage success message to the content server, to inform the content server of successful
storage.

[0072] In step 308, the content server stores the content data into the target file system
by using the fingerprint information of the content data as an index in response to receiving
the storage success message returned by the blockchain system.

[0073] After determining that the storage succeeds, the content server stores the
15 content data into the target file system. For the target file system, a relatively fast and
convenient storage system is provided by using a Key-Value mechanism involved in the
foregoing embodiments. In this case, the fingerprint information of the content data is used as
Key, and the content data is used as Value, thereby implementing the storage of the content
20 data and providing a basis for subsequent quick query.

[0074] In a possible implementation, when receiving a storage failure message
returned by the blockchain system, the content server can learn about that the content data is
similar content, and process at least one of the content data and a publishing user of the
content data. For plagiarized content, the content data is deleted, while for a publishing user
25 of the plagiarized content, punishment, such as account banning, is performed. In practice, to
avoid false processing, target fingerprint information can further be obtained from the
blockchain system based on fingerprint information whose similarity is greater than the target
threshold, the target fingerprint information being fingerprint information whose similarity
with the fingerprint information is greater than the target threshold. Content data
30 corresponding to the target fingerprint information is obtained from the target file system by
using the target fingerprint information as an index, and is submitted to an administrator user
for manual review. If it is determined, through the manual review, that the content data is
plagiarized content, punishment is performed thereon, and if it is determined, through the

manual review, that the content data is not plagiarized content, the content data is not processed, but is marked and transmitted to the blockchain system again.

[0075] In a possible implementation, the foregoing content quality score may further play a crucial role in recommendation, that is, the method further includes: determining, by
5 using a content quality score of candidate content data as a ranking factor and at least one piece of correlation information, content data satisfying recommendation requirements from the candidate content data in content data recommendation, the content quality score being determined based on the user feedback information.

[0076] In content data recommendation, any one or a combination of the following
10 recommendation methods is used, for example, 1) user interest tag mining; 2) content tag mining; 3) a user-interest-and-content matching algorithm; and 4) user-interest collaborative filtering. The foregoing recommendation methods are implemented by mining user interests, mining content tags that a user has browsed or watched, interest-content matching, or prediction based on interests of similar users. That is, recommendation is performed by using
15 correlation information of a user. Because a content quality score of the content data can be determined based on the foregoing user feedback information, the content quality score is used as a ranking factor. When a candidate recommendation list is determined during a push process, with reference to other correlation information (tag correlation, keyword correlation, category correlation, or the like), a recommendation score of the candidate content data and a
20 recommendation score of each user are respectively calculated, and content data recommendation is performed for a user according to a specified recommendation rule in descending order of scores.

[0077] According to technical solutions provided by the embodiments of this application, content data and fingerprint information thereof are stored in combination by
25 providing a blockchain system and a target file system, and the fingerprint information can be persistently stored by using the blockchain system. By using the fingerprint information as a connection between the blockchain system and the target file system, and by taking advantage of trustworthiness and reliability characteristics of data storage in the blockchain system, a storage manner with low data storage pressure is provided, so that the data storage pressure is
30 significantly reduced, and data processing efficiency can be improved.

[0078] Further, by providing the blockchain system and the target file system to store the content data and the fingerprint information thereof in combination, the fingerprint information of the content data can uniquely and independently determine a similarity

between content data, so that a content server can publish content based on the similarity of content data, and determine plagiarized data, thereby greatly reducing data storage pressure while protecting the content data. In addition, due to a characteristic that data can be rapidly spread by using the blockchain system, the problem of large communication latency can be resolved. By storing only fingerprint information in the blockchain system, and storing the content data by using the target file system, the problem of small storage space can also avoided. Further, the embodiments of this application pertain to the field of machine learning and artificial intelligence technologies, and in particular, relate to content distribution in a blockchain environment. In addition, a feedback mechanism is also established, which can resolve quality problems and copyright dispute problems of content services. Further, a corresponding recommendation mechanism is also provided, to solve the problem of effective matching between user interests and content. This technical solution provides an effective method for suppressing plagiarism and spreading.

[0079] Based on the same technical conception, referring to FIG. 4, an embodiment of this application provides a blockchain-based content processing apparatus, including:

a fingerprint obtaining module 401, configured to obtain fingerprint information of newly published content data;

a transmitting module 402, configured to transmit the fingerprint information of the content data to a blockchain system; and

a storage module 403, configured to store the content data into a target file system by using the fingerprint information of the content data as an index in response to receiving a storage success message returned by the blockchain system.

[0080] In a possible implementation, the fingerprint obtaining module 401 is configured to:

obtain, according to a publishing source of the content data, a target eigenvalue algorithm matching the publishing source; and

generate the fingerprint information of the content data based on the target eigenvalue algorithm and the content data.

[0081] In a possible implementation, the transmitting module 402 is configured to:

obtain a remainder between the fingerprint information and the quantity of node devices in the blockchain system; and

determine a target node device, and transmit the fingerprint information of the content data to the target node device in the blockchain system. A node device identifier of the target node device is the remainder.

[0082] In a possible implementation, the storage success message is transmitted in a case that the blockchain system stores no content data whose similarity is greater than a target threshold and a consensus is reached.

[0083] The apparatus further includes: a plagiarism processing module, configured to process at least one of the content data and a publishing user of the content data in response to receiving a storage failure message returned by the blockchain system.

[0084] In a possible implementation, the apparatus further includes:
a content data obtaining module, configured to obtain target fingerprint information from the blockchain system, the target fingerprint information being fingerprint information whose similarity with the fingerprint information is greater than the target threshold; and obtain content data corresponding to the target fingerprint information from the target file system by using the target fingerprint information as an index.

[0085] In a possible implementation, the apparatus further includes: a recommendation module, configured to determine, by using a content quality score of candidate content data as a ranking factor and at least one piece of correlation information, content data satisfying recommendation requirements from the candidate content data in content data recommendation, the content quality score being determined based on the user feedback information.

[0086] It should be noted that when the apparatus provided in the foregoing embodiments implements functions thereof, the division of the functional modules is only an example for purpose of description. In the practical application, the functions may be distributed to different functional modules as needed, that is, the internal structure of the device is divided into different functional modules, to implement all or some of the functions described above. In addition, the apparatus according to the foregoing embodiments and the method embodiments belong to one conception, and for the specific implementation process thereof, reference may be made to the method embodiments, which are not described herein again.

[0087] Based on the same technical conception, referring to FIG. 5, an embodiment of this application provides a blockchain-based content processing apparatus, including:

a receiving module 501, configured to receive fingerprint information of content data transmitted by a content server;

a transmitting module 502, configured to broadcast in a blockchain system,

the transmitting module 502 being further configured to transmit a verification success message in response to successfully verifying the fingerprint information; and

a storage module 503, configured to store a block including the fingerprint information of the content data into the blockchain of the blockchain system in response to reaching, by the blockchain system, a consensus on the fingerprint information of the content data,

the transmitting module 502 being further configured to transmit a storage success message to the content server.

[0088] In a possible implementation, fingerprint information verification refers to determining whether fingerprint information stored in blocks of the blockchain system includes fingerprint information whose similarity with the fingerprint information is greater than a target threshold. If yes, the verification fails; and if not, the verification succeeds.

[0089] In a possible implementation, the successfully verifying refers to that no fingerprint information whose similarity with the fingerprint information of the content data is greater than or equal to the target threshold exists in existing fingerprint information in a node device.

[0090] In a possible implementation, the reaching a consensus refers to that a target quantity of node devices in the blockchain system successfully verify the fingerprint information.

[0091] In a possible implementation, the transmitting module is further configured to transmit the storage success message to the content server in response to a verification success.

[0092] In a possible implementation, the transmitting module is further configured to transmit a storage failure message to the content server in response to a verification failure.

[0093] In a possible implementation, the apparatus further includes:

a determining module, configured to determine, according to first link address information carried in a read instruction, a first block corresponding to the first link address information on the blockchain in response to receiving the read instruction; and

an extraction module, configured to extract first fingerprint information corresponding to the first link address information from the first block,

the transmitting module 502 being further configured to return the first fingerprint information, the first fingerprint information being used for reading first content data corresponding to the first fingerprint information from a target file system.

[0094] In a possible implementation, the block further includes link address information and a content update record of the content data.

[0095] In a possible implementation, the block further includes user feedback information.

[0096] In a possible implementation, the storage module is further configured to correspondingly store, in response to receiving user feedback information of any user for second content data, second fingerprint information of the second content data and the user feedback information for the second content data into the blockchain.

[0097] In a possible implementation, the apparatus further includes a user processing module, configured to perform at least one of the following:

accumulating, based on the user feedback information, a participation score for a first target user submitting the user feedback information, and rewarding the first target user based on the accumulated participation score;

accumulating, based on user feedback information for third content data in response to that it is determined that the third content data does not conform to a target rule, a contribution score for a second target user submitting positive user feedback information, and rewarding the second target user based on the accumulated contribution score; and

accumulating, according to received user feedback information, a content quality score for the second target user submitting the content data, and processing at least one of the content data and the second target user in response to that the content quality score of the content data is lower than a target quality score.

[0098] It should be noted that when the apparatus provided in the foregoing embodiments implements functions thereof, the division of the functional modules is only an example for purpose of description. In the practical application, the functions may be distributed to different functional modules as needed, that is, the internal structure of the device is divided into different functional modules, to implement all or some of the functions described above. In addition, the apparatus according to the foregoing embodiments and the

method embodiments belong to one conception, and for the specific implementation process thereof, reference may be made to the method embodiments, which are not described herein again.

[0099] FIG. 6 is a schematic structural diagram of a computer device according to an embodiment of this application. The computer device may refer to any node device or content server in the foregoing embodiments. The computer device may vary a lot due to different configurations or performance, and may include one or more central processing units (CPUs) 601 and one or more memories 602. The one or more memories 602 store at least one piece of program code, the at least one piece of program code being loaded and executed by the one or more CPUs 601 to implement the blockchain-based content processing method provided in the foregoing method embodiments. In practice, the computer device may also have components such as a wired or wireless network interface, a keyboard, and an input/output interface for purpose of input and output, and may further include other components for implementing functions of the device, which are not described herein redundantly.

[0100] In a possible implementation, the foregoing computer device is provided as a server, and the server includes a processor and a memory, the memory storing at least one piece of program code, the at least one piece of program code being loaded and executed by the processor to implement the following method:

obtaining fingerprint information of newly published content data;

transmitting the fingerprint information of the content data to a blockchain system; and

storing the content data into a target file system by using the fingerprint information of the content data as an index in response to receiving a storage success message returned by the blockchain system.

[0101] In a possible implementation, the processor is further configured to perform the following steps: obtaining a target eigenvalue algorithm according to a publishing source of the content data, the target eigenvalue algorithm matching the publishing source; and

generating the fingerprint information of the content data based on the target eigenvalue algorithm and the content data.

[0102] In a possible implementation, the processor is further configured to perform the following steps: obtaining a remainder between the fingerprint information and the quantity of node devices in the blockchain system;

determining a target node device, the target node device being a node device of which a node device identifier is the remainder; and

transmitting the fingerprint information of the content data to the target node device in the blockchain system.

5 **[0103]** In a possible implementation, the storage success message is transmitted after the blockchain system reaches a consensus, and

[0104] the processor is further configured to perform the following step: processing at least one of the content data and a publishing user of the content data in a case that a storage failure message returned by the blockchain system is received.

10 **[0105]** In a possible implementation, the processor is further configured to perform the following steps: obtaining target fingerprint information from the blockchain system, the target fingerprint information being fingerprint information whose similarity with the fingerprint information is greater than a target threshold; and

obtaining content data corresponding to the target fingerprint information from
15 the target file system by using the target fingerprint information as an index.

[0106] In a possible implementation, the processor is further configured to perform the following step: transmitting user feedback information for any piece of content data to the blockchain system.

[0107] In a possible implementation, the processor is further configured to perform
20 the following step:

determining, by using a content quality score of candidate content data as a ranking factor and at least one piece of correlation information, content data satisfying recommendation requirements from the candidate content data in content data recommendation, the content quality score being determined based on the user feedback
25 information stored in the blockchain system.

[0108] In a possible implementation, the foregoing computer device is provided as a server, and the server includes a processor and a memory, the memory storing at least one piece of program code, the at least one piece of program code being loaded and executed by the processor to implement the following method.

30 **[0109]** In a possible implementation, the foregoing computer device is provided as a server, and the server is a node device in the blockchain system and includes a processor and

a memory, the memory storing at least one piece of program code, the at least one piece of program code being loaded and executed by the processor to implement the following method:

5 receiving fingerprint information of content data, and broadcasting in the blockchain system;

transmitting a verification success message in response to successfully verifying the fingerprint information; and

10 storing a block including the fingerprint information of the content data into the blockchain of the blockchain system in response to reaching, by the blockchain system, a consensus on the fingerprint information of the content data.

[0110] In a possible implementation, the successfully verifying refers to that no fingerprint information whose similarity with the fingerprint information of the content data is greater than or equal to a target threshold exists in existing fingerprint information in the server.

15 **[0111]** In a possible implementation, the reaching a consensus refers to that a target quantity of node devices in the blockchain system successfully verify the fingerprint information.

[0112] In a possible implementation, the block further includes link address information and a content update record of the content data.

20 **[0113]** In a possible implementation, the processor is further configured to perform the following steps:

determining, according to first link address information carried in a read instruction, a first block corresponding to the first link address information on the blockchain in response to receiving the read instruction; and

25 extracting first fingerprint information corresponding to the first link address information from the first block, and returning the first fingerprint information, the first fingerprint information being used for reading first content data corresponding to the first fingerprint information from the target file system.

[0114] In a possible implementation, the processor is further configured to perform
30 the following step:

correspondingly storing, in response to receiving user feedback information of any user for second content data, second fingerprint information of the second content data and the user feedback information for the second content data into the blockchain.

[0115] In an exemplary embodiment, a computer-readable storage medium is further provided, the storage medium storing at least one piece of program code, the at least one piece of program code being loaded and executed by a processor of a computer device to implement any foregoing blockchain-based content processing method.

[0116] Optionally, the foregoing computer-readable storage medium may be a read-only memory (ROM), a random access memory (RAM), a compact disc ROM (CD-ROM), a magnetic tape, a floppy disk, an optical data storage device, or the like.

[0117] In an exemplary embodiment, a computer program product (or application) is further provided, storing at least one piece of program code, the at least one piece of program code being loaded and executed by a processor of a computer device to implement any foregoing blockchain-based content processing method.

[0118] It should be understood that "plurality of" mentioned in the specification means two or more, "and/or" describes an association relationship between associated objects and represents that three relationships may exist. For example, A and/or B may represent the following three cases: only A exists, both A and B exist, and only B exists. The character "/" generally indicates an "or" relationship between the associated objects.

[0119] The above described are only some exemplary embodiments of this application, but are not intended to limit this application. Any modification, equivalent replacement, or improvement made without departing from the spirit and principle of this application shall fall within the protection scope of this application.

CLAIMS

What is claimed is:

1. A blockchain-based content processing method, comprising:

obtaining, by a content server, fingerprint information of newly published content data;

5 transmitting, by the content server, the fingerprint information of the content data to a blockchain system; and

storing, by the content server by using the fingerprint information of the content data as an index, the content data into a target file system in response to receiving a storage success message returned by the blockchain system,

10 wherein the fingerprint information of the content data is verified in response to determining that no fingerprint information whose similarity with the fingerprint information of the content data is greater than or equal to a target threshold exists in existing fingerprint information of a node device in the blockchain system; and a block including the fingerprint information of the content data is added in the blockchain system in response to determining
15 that the verification succeeds, and the content server receives the storage success message from the blockchain system.

2. The method according to claim 1, wherein the obtaining, by a content server, fingerprint information of newly published content data comprises:

obtaining, by the content server according to a publishing source of the content data, a
20 target eigenvalue algorithm, the target eigenvalue algorithm matching the publishing source; and

generating, by the content server, the fingerprint information of the content data based on the target eigenvalue algorithm and the content data.

3. The method according to claim 1, wherein the transmitting, by the content server, the
25 fingerprint information of the content data to a blockchain system comprises:

obtaining, by the content server, a remainder between the fingerprint information and the

quantity of node devices in the blockchain system;

determining, by the content server, a target node device, a node device identifier of the target node device being the remainder; and

transmitting, by the content server, the fingerprint information of the content data to the
5 target node device in the blockchain system.

4. The method according to claim 1, wherein the storage success message is transmitted after the blockchain system reaches a consensus; and

after the transmitting, by the content server, the fingerprint information of the content data to a blockchain system, the method further comprises:

10 processing, by the content server, at least one of the content data and a publishing user of the content data in a case that a storage failure message returned by the blockchain system is received.

5. The method according to claim 4, further comprising:

obtaining, by the content server, target fingerprint information from the blockchain
15 system, the target fingerprint information being fingerprint information whose similarity with the fingerprint information is greater than the target threshold; and

obtaining, by the content server, content data corresponding to the target fingerprint information from the target file system by using the target fingerprint information as an index.

6. The method according to claim 1, further comprising:

20 transmitting, by the content server, user feedback information for any piece of content data to the blockchain system.

7. The method according to claim 6, further comprising:

determining, by the content server by using a content quality score of candidate content data as a ranking factor and at least one piece of correlation information, content data
25 satisfying recommendation requirements from the candidate content data in content data recommendation, the content quality score being determined based on the user feedback

information stored in the blockchain system.

8. A blockchain-based content processing method, applied to any node device in a blockchain system, the method comprising:

5 receiving, by the node device, fingerprint information of content data, and broadcasting in the blockchain system;

verifying, by the node device, the fingerprint information of the content data, in response to determining that no fingerprint information whose similarity with the fingerprint information of the content data is greater than or equal to a target threshold exists in existing fingerprint information of the node device;

10 adding, by the node device, a block comprising the fingerprint information of the content data into a blockchain of the blockchain system, in response to determining that the fingerprint information is verified successfully; and

transmitting, by the node device, a storage success message to a content server.

15 9. The method according to claim 8, wherein the reaching a consensus refers to that a target quantity of node devices in the blockchain system successfully verify the fingerprint information.

10. The method according to claim 8, wherein the block further comprises link address information and a content update record of the content data.

11. The method according to claim 9, further comprising:

20 determining, by the node device according to first link address information carried in a read instruction, a first block corresponding to the first link address information on the blockchain in response to receiving the read instruction; and

extracting, by the node device, first fingerprint information corresponding to the first link address information from the first block, and returning the first fingerprint information, the first fingerprint information being used for reading first content data corresponding to the first fingerprint information from a target file system.

12. The method according to claim 8, further comprising:

correspondingly storing, by the node device in response to receiving user feedback information of any user for second content data, second fingerprint information of the second content data and the user feedback information for the second content data into the blockchain.

5 13. A blockchain-based content processing apparatus, comprising a plurality of functional modules configured to perform the blockchain-based content processing method according to any one of claims 1 to 7, or the blockchain-based content processing method according to any one of claims 8 to 12.

10 14. A server, comprising a processor and a memory, the memory storing at least one piece of program code, the at least one piece of program code being loaded and executed by the processor to implement the blockchain-based content processing method according to any one of claims 1 to 7, or the blockchain-based content processing method according to any one of claims 8 to 12.

15 15. A computer-readable storage medium, storing at least one piece of program code, the at least one piece of program code being loaded and executed by a processor to implement the blockchain-based content processing method according to any one of claims 1 to 7, or the blockchain-based content processing method according to any one of claims 8 to 12.