



研究与开发

基于区块链与可信度函数的电子加密信息上链存储方法

周婷

(国网山西省电力公司营销服务中心, 山西 太原 030032)

摘要: 为优化加密信息在区块链的存储空间占比, 提高上链存储速度, 提出了基于区块链与可信度函数设计电子加密信息上链存储方法。建立基于区块链和可信度函数的信息加密分类模型, 获取溯源凭证判定函数, 对电子信息进行代理重加密; 计算集群存储空间容量, 使用加密密文直接匹配访问结构树; 区分主体可信度与客体可信度, 基于可信度函数分类电子加密信息, 设计电子加密信息上链存储方法。实验结果表明, 应用所提方法后, 当文件数量在 1 000~10 000 个时, 数据存储速度最快, 存储空间占比以及存储速度上均有较大优化。

关键词: 区块链; 可信度函数; 电子加密信息; 上链存储; 多级分类; 存储空间

中图分类号: TP309

文献标志码: A

doi: 10.11959/j.issn.1000-0801.2023147

Blockchain and credibility function based chain storage method of electronic encrypted information

ZHOU Ting

State Grid Shanxi Marketing Service Center, Taiyuan 030032, China

Abstract: To optimize the proportion of encrypted information storage space in the blockchain and improve the speed of online storage, a design method for electronic encrypted information online storage based on blockchain and credibility function was proposed. An information encryption classification model based on blockchain and credibility function was established, and a traceability voucher judgment function was obtained. Proxy re-encryption was performed on electronic information. The storage capacity of the cluster was calculated, and encrypted ciphertext was used to directly match the access structure tree. Subject credibility and object credibility were distinguished, electronic encrypted information was classified based on the credibility function, and an electronic encrypted information uplink storage method was designed. The experimental results show that after applying the proposed method, when the number of files is between 1 000 and 10 000, the data storage speed is the fastest, and the storage space proportion and storage speed are greatly optimized.

Key words: blockchain, credibility function, electronically encrypted information, uplink storage, multi level classification, storage space

0 引言

通常将区块链作为一个数据库,共享存储在其中的信息,且基本所有被存储在区块链中的数据都具备可追溯性。在去中心化的管理方法下,任何人都可以通过区块链技术上传数据,也同样可以依据区块链技术下载其中被公开的内容。因此,用户信息共享存储过程中的灵活性,以及大量数据在区块链中占据的内存容量成为亟待解决的问题。

为保证数据存储以及共享业务高效安全,在基本存储的方案之上,文献[1]提出了一种以实现用户共享数据完整性的存储方法,通过双线性映射对以及布隆过滤器,建立了动态数据的存储结构,并实时对其进行动态更新、删除、验证,使新的存储方法具备更好的运行效率。文献[2]以电力物联网中的数据为核心,设计了一种基于区块链和数据湖的信息存储方法,建立共享访问请求和智能合约的公共模型,通过发送指定的数据,减少数据湖在存储控制模型中的时延,并进一步提高了数据存储的安全性。文献[3]在传统的存储方法基础上,重新计算了通信区域内区块链的传送帧量,在保证最大物理输出值的基础上,生成对应的加密密钥,并设计了一种基于排序优先算法的数据存储方法。该方法可以有效减少数据存储的时间,保证数据存储效率。文献[4]为了补偿受攻击的残缺数据,融合了区块链技术与基于属性的访问控制的优势,提出基于数据共享系统的可审计加密方案。通过对云用户进行授权数据访问控制,为使用区块链技术的用户提供数据属性安全审计。该方法可以有效保留存储数据的完整性和一致性,提高数据存储的安全性。

综合分析上述研究方法,本文通过区块链与可信度函数,设计了一种电子加密信息上链存储的方法。

1 电子加密信息分类模型

1.1 电子信息代理重加密

在一个基于区块链技术的数据存储过程中,想要保证存储信息的安全,就需要对电子信息进行加密。在不可信的开放云环境下,以不解密信息原文为前提,可以将信息以代理加密的方式处理,此时最主要的代理对象就是分片信息。对电子信息 H_p 进行数字化处理,将其转换为十进制格式,得到信息 H_n 。根据区块链的节点,获取交易分片的数量,构造多项式时,需要满足以下条件。

$$\begin{cases} H_n = \frac{f_i(t_{\max} - t_i)}{P_k} \\ P_k > \max(t_{\max}, f_i) \\ t_{\max} = \max(t_i) \\ 1 \leq i \leq k-1 \end{cases} \quad (1)$$

其中, P_k 表示消息分片的大素数, $t_{\max}$ 表示多项式分片数量的最大值, f_i 表示交易分片信息的数量, t_i 表示第 i 组多项式的数值, k 表示区块链节点数量。结合式(1),可以得到多组分片信息 (x_p, y_p) , 并组合成多个参数,在这些参数内, P_k 作为公共参数,此时公共参数的输入值可以表示为:

$$\text{arams}(p_k) = (p_f, q_f, e_f) \quad (2)$$

其中, $\text{arams}(p_k)$ 表示公共参数的输入值,一般具备一定程度的安全性; p_f 和 q_f 为满足该公共参数的大素数; e_f 表示乘法循环的阶数^[5-6]。

在分片消息以及随机数的碰撞下,以 $\text{arams}(p_k)$ 作为排序的溯源凭证,并将所有加密数据经过主节点,记录在区块链上。此时,通过随机交易获取的电子加密信息以区块链作为溯源凭证的判断函数:

$$f(d) = L \sum_{i=1}^2 f(d_i)^{e_f} \prod_{j=1, j \neq i}^2 \frac{d_k - d_j}{d_i - d_j} \text{arams}(p_k) \quad (3)$$

其中, $f(d)$ 表示区块链溯源凭证的判定函数,当



$f(d)$ 大于或等于 0 时, 可以断定其在区块链上溯源成功; L 表示重加密密文; $f(d_i)$ 表示第 i 个循环下的溯源结果; d_k 表示 k 组分片信息; d_i 和 d_j 分别表示第一组分片信息与最后一组分片信息^[7]。在给定的安全参数下, 系统属性已经生成了假定的加密属性集合, 由此可以得到电子信息属性代理的 L 。

1.2 直接匹配访问结构树

获取了 L 后, 如果数据拥有者想要正确解密密文, 则其生成的私钥必须与上述加密属性集合中的加密策略互相匹配, 符合 $f(d)$ 的判定要求。在这样的架构内, 可以将文件大致分为 3 份, 其矩阵格式可以表示为:

$$\mathbf{F}_{\text{SC}} = \begin{bmatrix} f_{11} & f_{12} & \cdots & f_{1m} \\ f_{21} & f_{22} & \cdots & f_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ f_{n1} & f_{n2} & \cdots & f_{nm} \end{bmatrix} \quad (4)$$

其中, $\mathbf{F}_{\text{SC}}$ 表示存储服务集群矩阵, f_{nm} 表示第 n 个存储节点中第 m 个序号的文件备份^[8-9]。

在同组文件正常工作的前提下, 对 $\mathbf{F}_{\text{SC}}$ 中的备份数据进行分组, 并决定每一个分组的最大存储容量, 节点标号可以表示为:

$$G_{\text{cap}} = \mathbf{F}_{\text{SC}} \min(Z(s_1), Z(s_2), Z(s_3)) \quad (5)$$

其中, G_{cap} 表示任意分组下存储容量最大的节点序号; $Z(s_1)$ 、 $Z(s_2)$ 和 $Z(s_3)$ 分别表示同一集群在不同分组内的存储容量。计算该集群下存储空间总容量为:

$$G_{\text{SC}} = \frac{G_{\text{cap}} \sum_{i=1}^2 f_i}{N_s} \quad (6)$$

其中, G_{SC} 表示该集群下所有分组内存储空间的容量之和, f_i 表示通过第 i 组多项式所获得的文件备份, N_s 表示服务器的单位负载压力^[10]。在这样的存储压力下, 可以通过访问结构树, 提高算法的运行效率, 访问结构树如图 1 所示。

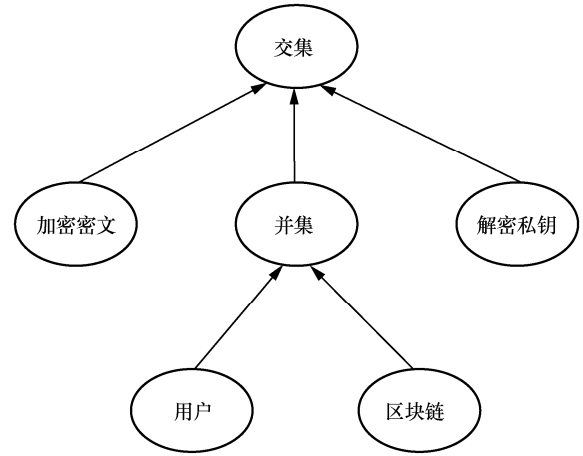


图 1 访问结构树

图 1 中, 在直接与访问结构树相匹配时, 访问结构树最下方的叶节点可以分为用户和区块链两类。在该类节点下, 需要首先设置门限阈值, 控制 G_{SC} 的数值, 避免存储空间过度浪费。该门限阈值在访问结构树中, 可以表现为交集与并集两种关联方式, 其约束条件分别为:

$$0 \leq y_k < N_m^{G_{\text{SC}}} \quad (7)$$

$$y_k = N_m \quad (8)$$

其中, y_k 表示门限阈值的系数, 一般通过访问结构树获取; N_m 表示结构树中叶节点的数量^[11-12]。在这些节点下, 可以获取用户与区块链的关联函数, 并同步推导得到加密密文与解密私钥, 并实现访问结构树与电子加密信息的匹配, 该方法可以节省电子加密信息转换成密文结构的时间。

1.3 分类电子加密信息

在电子信息中, 需要根据 y_k 分别区分主体特征与客体特征, 并建立多级安全访问模型, 实现对主体的加密操作。分别在函数中设置主体集合与客体集合, 保证可信度集合具备安全标记的控制模式。此时, 安全标记的函数集合可以表示为:

$$S_f = \{(S_1, S_2, S_3)\} \quad (9)$$

其中, S_f 表示安全标记的函数集合, S_1 表示主体内的最大安全标记函数, S_2 表示客体的安全标记函数, S_3 表示当前的安全标记函数^[13]。结合以上

标记函数, 获取可信度评估函数的边界条件。在线性递减的形式中, 得到主体可信度与客体可信度的关系式:

$$\begin{cases} K_x = \max \left(s_1 \frac{u_i + u_j}{2} - \Delta f_w \right) \\ Z_x = \max \left(s_2 \frac{u_j}{u_i} - \Delta f_w \right) \end{cases} \quad (10)$$

其中, K_x 和 Z_x 分别表示主体可信度与客体可信度, u_i 和 u_j 分别表示可信度函数的违规等级评估系数, Δf_w 表示安全等级标记的惩罚函数。

通过计算 K_x 与 Z_x , 可以得到电子信息多级加密模型, 用于对不同信用等级客户进行分类。

$$H_i(\beta) = \frac{\text{Hash}(p_s, f_i)}{(\sqrt{L_d} \times N_m) K_x} + Z_x d_p \quad (11)$$

其中, $H_i(\beta)$ 表示电子信息在 β 级分类中的加密密钥, $\text{Hash}(\cdot)$ 表示哈希递归函数, p_s 表示密钥的加密种子数量, L_d 表示电子加密信息安全存储的长度, d_p 表示加密信息的数量^[14-15]。结合式(11), 可以在可信度函数下对电子信息进行多级分类, 从而减少了区块链内大数据的查找工序。

2 电子加密信息上链存储方法

在保证电子信息安全性的同时, 完成其在区块链内的上链存储, 综合数据目标的接收者以及数据的拥有者, 可以在一定程度上保证加密信息不会被信用等级不足的客户获取。电子加密信息上链存储方法用伪代码形式可表示为:

```
...
public_key, private_key = generate_key_pair()
...
encrypted_data = encrypt(data, public_key)
...
storage_capacity = calculate_storage_capacity
(cluster)
...
re_encrypted_data = proxy_re_encrypt (en-
```

```
rypted_data, private_key, storage_capacity)
...
access_structure_tree = match_access_structure_
tree(re_encrypted_data)
...
subject_trustworthiness, object_trustworthiness
= distinguish_trustworthiness(access_structure_tree)
...
classified_data = classify_data(re_encrypted_
data, subject_trustworthiness, object_trustworthiness)
...
blockchain.store(classified_data)
...
```

其中, $\text{generate_key_pair}()$ 函数用于生成公私钥对, $\text{encrypt}()$ 函数用于对数据进行加密, $\text{calculate_storage_capacity}()$ 函数用于计算存储空间容量, $\text{proxy_re_encrypt}()$ 函数用于进行代理重加密, $\text{match_access_structure_tree}()$ 函数用于匹配访问结构树, $\text{distinguish_trustworthiness}()$ 函数用于区分主体信任度和客体信任度, $\text{classify_data}()$ 函数用于根据信任度函数对加密信息进行分类, $\text{blockchain.store}()$ 函数用于将加密信息上链存储。

由上述伪代码内容可知, 公私钥对生成过程需要保证公私钥对的安全性和唯一性, 以确保数据在加密和解密过程中的安全性; 电子信息加密需要保证加密后的数据无法被破解, 只能由拥有私钥的用户进行解密, 以确保数据在传输和存储过程中的安全性; 存储空间容量计算需要考虑实际需求和存储介质的尺寸等因素, 以确保存储空间能够满足实际需求; 代理重加密是指将密文数据通过代理节点进行加密, 以确保数据的安全性^[16]。在此过程中需要保证代理节点的可信度和安全性; 访问结构树匹配是指将代理重加密后的数据与访问结构树进行匹配, 以确定哪些用户可以访问该数据; 在区分主、客体信任度的过程中, 首先需要通过访问结构树匹配判断哪些用户可以访



问该数据。然后基于历史记录、用户评价、机器学习等方法使用可信度函数来区分主体信任度和客体信任度，并对加密信息进行分类。

在获取电子加密信息上链存储方法伪代码形式的基础上，电子加密信息上链存储示意图如图2所示。

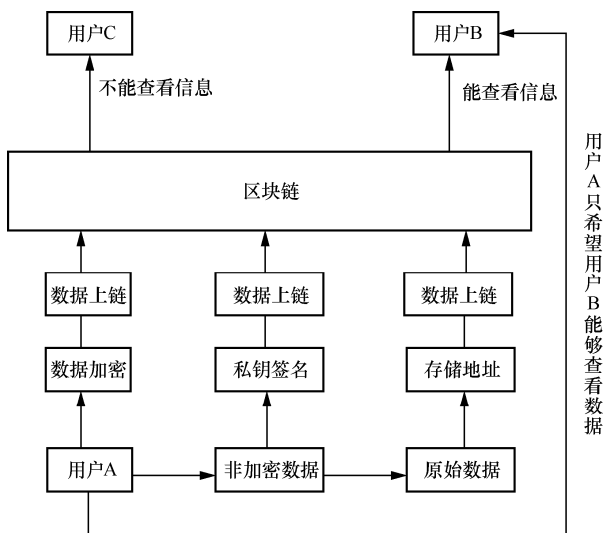


图2 电子加密信息上链存储示意图

图2中，共有3个用户，分别是用户A、用户B和用户C。其中，用户A为向区块链中添加数据的用户，用户B和用户C则为可以在区块链中查看信息的用户。建立初始设定：用户A只希望用户B能够看见区块链中的该项数据，而用户C不具备观看该项数据的权限。则此时只有用户B可以获取相应的信用等级，实现客户的等级分类^[17]。用户A通过将数据分为加密与非加密两个部分，加密数据直接上链存储，而非加密数据则需要通过私钥签名等方式，获取原始数据与存储地址，最后进行上链存储。此时区块链中的电子加密信息既可以实现用户上传数据的过程，也可以实现符合等级的用户下载数据的过程，同时还会拒绝不符合等级的用户对数据的加载。

3 实验研究

3.1 实验平台与数据

为测试本文设计的基于区块链与可信度函数

的电子加密信息上链存储方法的优越性，设计以下实验。选择7台同样配置的计算机作为实验的平台，用于用户信息的输出。设定协议网关为IEEE 165.127.3.0，将两台计算机分别与区块链存储加密系统以及信息适配器相连，另一台计算机可以作为名称节点（name node），最后4台计算机是数据节点（data node）。计算机内，采用GNU/Linux作为操作系统，硬盘容量为500 GB。分别设置不同文件数量下数据的容量，当文件数为1 000个、2 000个、3 000个、…、10 000个时，数据容量的大小分别为10.65 GB、21.43 GB、31.81 GB、42.41 GB、52.99 GB、63.34 GB、73.88 GB、84.52 GB、95.78 GB和106.34 GB。分别将以上不同文件数据作为实验变量，测试对不同大小的加密信息进行上链存储时，实验测试指标的变化情况。

3.2 存储空间优化分析

在区块链的存储空间中，每一个区块均代表至少500个存储内容，随着区块链内存储的数据越来越多，其生成的冗余数据量也在不断增加，占用的存储空间也在随之增加。想要保证信息上链存储的优化性能，就需要尽量减少并优化区块链内的冗余数据容量。使用数据存储空间，可以测试信息存储方案的优化情况。

$$U_{cp} = \frac{N_H + \sqrt{P_{Hash} \times N_Y}}{N_H + N_S} \quad (12)$$

其中， U_{cp} 表示数据存储空间的大小，存储单位数量的数据所需空间越多，证明数据冗余量越大，所需空间越少，其数据存储的优化程度越高； N_H 表示区块链内区块数量； P_{Hash} 表示哈希函数内每一个存储散列值的系数； N_Y 表示数据库内完成存储的数量； N_S 表示有存储意图的原始存储数量。分别测试区块链在2022年某月内的存储空间占比的变化情况，以时间为单位，分别获取每5日存储空间的优化结果，

得到不同时段内的存储空间优化测试数据如图 3 所示。

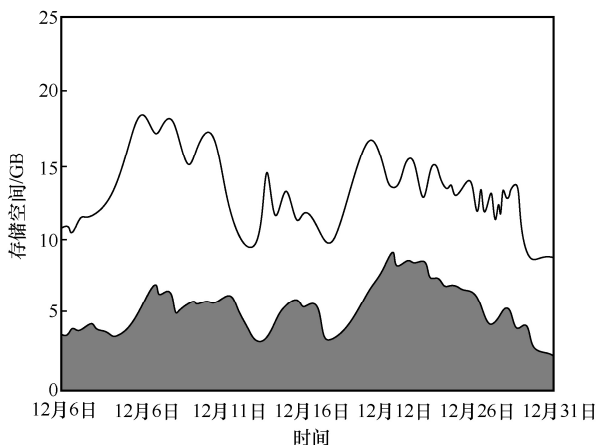


图3 存储空间优化测试数据

在图 3 中，白色部分为原始存储空间的占比，其在 12 月 6 日时达到本月内存储空间占比的最高峰，约为 18.2 GB，在 12 月 31 日达到本月存储空间占比的最低谷，约为 9.6 GB。而使用本文的电子加密信息上链存储方法，可以使存储所需空间如图 3 所示的灰色部分所示。在 12 月 21 日达到本月存储空间占比的最大值，仅为 9.5 GB，同样在 12 月 31 日达到其占比的最小值，约为 3.2 GB。无论是计算存储空间占比的平均值，还是对比其最大值与最小值，均能够实现存储结果的优化。

3.3 信息上链存储速度测试

在信息存储的过程中，除存储空间占比外，还需要考虑电子加密信息存储速度，因此对其进行测试。其计算式为：

$$v_{i,j} = \frac{\sum_{\eta=1}^2 v_{i,j}^{\eta}}{t} \quad (13)$$

其中， $v_{i,j}$ 表示文件被分为多个不同的存储分片后的存储速度， $v_{i,j}^{\eta}$ 表示第 η 次测试时表现出的存储速度， t 表示速度测试的总次数。对比本文的信息上链存储方法以及现有的几种数据存储方法，再结合式 (13)，可以得到电子加密信息上链存储速度测试结果如图 4 所示。

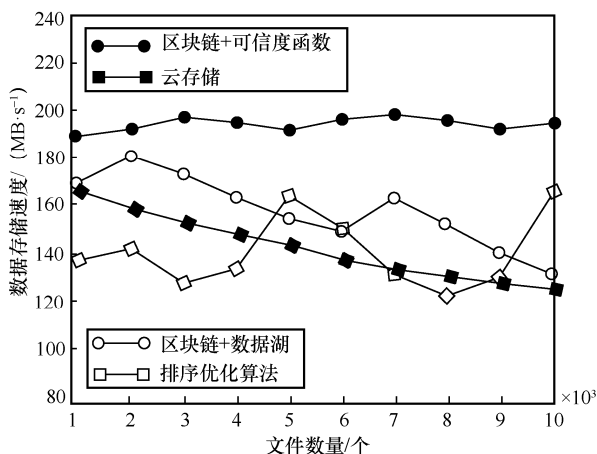


图4 电子加密信息上链存储速度测试结果

在图 4 中共有 4 种方法对加密信息上链存储的速度进行测试，由于 4 种方法差异极大，其存储速度的变化趋势也有很大不同之处。在区块链和可信度函数的测试方法中，随着文件数量由 1 000 增加至 10 000，数据的存储速度没有大的变化，基本均在 190~200 MB/s，可见在该方法下，存储数据量的变化，不会对存储速度造成影响。在云存储方法中，随着文件数量的增加，数据存储速度也在不断增加，由 1 000 个文件数量时的 168 MB/s 降低到 10 000 个文件数量时的 126 MB/s。在区块链与数据湖方法中，每增加 5 000 个文件数量，其数据的存储速度发生一次周期性改变，先大幅度增加，再连续缓慢减小。但是整体而言，在该方法下，数据存储速度仍然会随着文件数量的增加而不断减小。在排序优先算法中，文件数量的变化不会对数据存储速度造成影响。

4 结束语

本文设计了一种基于区块链与可信度函数的电子加密信息上链存储方法，通过随机交易获取的电子加密信息以区块链作为溯源凭证的判断函数，设计电子信息属性代理的重加密密文，基于用户与区块链的关联函数，保证加密密文直接匹配访问结构树，保证上链存储过程的安全性与效



率性,更具优越性。但本文方法在应用过程中,同时存储和处理大量数据时,针对应用场景进行实时交互的共识性能较差。因此在未来研究中,可以将改进共识算法与侧链技术融入该方法,通过划分子网络,优化数据存储结构,提升存储的扩展性。

参考文献:

- [1] 李秀艳, 刘明曦, 史闻博, 等. 基于云存储的动态组共享数据完整性验证方案[J]. 计算机工程与设计, 2022, 43(6): 1510-1519.
LI X Y, LIU M X, SHI W B, et al. Dynamic group shared data integrity verification scheme based on cloud storage[J]. Computer Engineering and Design, 2022, 43(6): 1510-1519.
- [2] 曾飞, 杨雄, 苏伟, 等. 基于区块链与数据湖的电力数据存储与共享方法[J]. 电力工程技术, 2022, 41(3): 48-54.
ZENG F, YANG X, SU W, et al. Power data storage and sharing method based on blockchain and data lake[J]. Electric Power Engineering Technology, 2022, 41(3): 48-54.
- [3] 谢瑞. 基于排序优化算法的通信信息加密存储方法[J]. 科学技术创新, 2021(18): 100-101.
XIE R. Communication information encryption and storage method based on sorting optimization algorithm[J]. Scientific and Technological Innovation, 2021(18): 100-101.
- [4] EZHIL A V, INDRA G K, KULOTHUNGAN K. Auditable attribute-based data access control using blockchain in cloud storage[J]. Journal of Supercomputing, 2022, 78(8): 10772-10798.
- [5] 斯兴瑶, 廖映华, 胥云, 等. 西门子 828D 数控系统数据采集与存储[J]. 制造技术与机床, 2022(4): 82-87.
SI X Y, LIAO Y H, XU Y, et al. Data acquisition and storage of Siemens 828D CNC system[J]. Manufacturing Technology & Machine Tool, 2022(4): 82-87.
- [6] 于简宁, 薛伟, 张华超, 等. 基于区块链的森林火灾档案数据共享方案[J]. 森林工程, 2022, 38(3): 95-105.
YU J N, XUE W, ZHANG H C, et al. Forest fire archives data sharing scheme based on blockchain[J]. Forest Engineering, 2022, 38(3): 95-105.
- [7] 李晗. 新加坡南洋理工大学科研数据管理探究与启示——基于数据生命周期视角[J]. 图书馆学研究, 2022(2): 68-73.
LI H. Study and enlightenment of research data management at Nanyang Technological University, Singapore-based on the perspective of data life cycle[J]. Research on Library Science, 2022(2): 68-73.
- [8] 许娜, 耿恒高, 徐传鹏, 等. 基于 MongoDB 的地震勘探数据管理系统的设计与实现[J]. 实验室研究与探索, 2022, 41(2): 251-260.
XU N, GENG H G, XU C P, et al. Design and implementation of seismic exploration data management system based on MongoDB[J]. Research and Exploration in Laboratory, 2022, 41(2): 251-260.
- [9] 许璐, 刘正军, 陈一铭. 轻小型无人机高光谱遥感原始数据存储与验证[J]. 测绘科学, 2022, 47(2): 95-101.
XU L, LIU Z J, CHEN Y M. Storage and verification of hyper-spectral remote sensing raw data for light and small UAVs[J]. Science of Surveying and Mapping, 2022, 47(2): 95-101.
- [10] 王健, 张蕴嘉, 刘吉强, 等. 基于区块链的司法数据管理及电子证据存储机制[J]. 信息安全, 2022(2): 21-31.
WANG J, ZHANG Y J, LIU J Q, et al. Blockchain-based mechanism for judicial data management and electronic evidence storage[J]. Netinfo Security, 2022(2): 21-31.
- [11] 杨河山, 张世明, 曹小朋, 等. 基于 Hadoop 分布式文件系统的地震勘探大数据样本采集及存储优化[J]. 油气地质与采收率, 2022, 29(1): 121-127.
YANG H S, ZHANG S M, CAO X P, et al. HDFS-based collection and storage optimization of seismic exploration big data samples[J]. Petroleum Geology and Recovery Efficiency, 2022, 29(1): 121-127.
- [12] 牛淑芬, 宋蜜, 方丽芝, 等. 智慧医疗中基于属性加密的云存储数据共享[J]. 电子与信息学报, 2022, 44(1): 107-117.
NIU S F, SONG M, FANG L Z, et al. Cloud storage data sharing based on attribute encryption in smart healthcare[J]. Journal of Electronics & Information Technology, 2022, 44(1): 107-117.
- [13] 王庆乐. 基于区块链技术的高校数字图书馆科研数据安全策略研究[J]. 图书馆工作与研究, 2021(12): 63-69.
WANG Q L. Study on the strategy of scientific research data security management in university digital library based on blockchain technology[J]. Library Work and Study, 2021(12): 63-69.
- [14] 彭柳, 张淼, 高杰欣. 基于区块链技术的电子档案安全存储与可信验证方案[J]. 中南民族大学学报(自然科学版), 2022, 41(6): 728-733.
PENG L, ZHANG M, GAO J X. Blockchain technology-based solution for secure storage and trusted verification of electronic files[J]. Journal of South-Central University for Nationalities

- (Natural Science Edition), 2022, 41(6): 728-733.
- [15] 霍颖瑜, 钟勇. 基于区块链技术的用户信息加密存储系统设计[J]. 现代电子技术, 2021, 44(21): 60-64.
- HUO Y Y, ZHONG Y. Design of user information encryption storage system based on blockchain technology[J]. Modern Electronics Technique, 2021, 44(21): 60-64.
- [16] 章坚武, 安彦军, 邓黄燕. DNS 攻击检测与安全防护研究综述[J]. 电信科学, 2022, 38(9): 1-17.
- ZHANG J W, AN Y J, DENG H Y. A survey on DNS attack detection and security protection[J]. Telecommunications Science, 2022, 38(9): 1-17.
- [17] 符安文. 可见光通信技术支持下的电子密码锁系统与系统加

密技术[J]. 数字技术与应用, 2021, 39(8): 194-196.

FU A W. Electronic code lock system and system encryption technology supported by visible light communication technology[J]. Digital Technology & Application, 2021, 39(8): 194-196.

[作者简介]



周婷(1990—),女,国网山西省电力公司营销服务中心工程师,主要研究方向为电力客户服务、电力市场等。