基于身份授权的可信去中心化存储网络

张靖宇, 刘　奇, 彭竑睿, 杨增辉, 袁开国, 李小勇

北京邮电大学 北京 中国 100876

摘要　去中心化存储网络具有可靠性强、成本低和速度快等优点, 越来越广泛地使用在各个行业中, 但在去中心化存储网络中存在节点好奇和不诚实两大特点, 分别导致去中心化网络节点可能窃取用户信息和用户检索时不能获得正确结果两种风险。本文通过在去中心化存储网络中引入属性基可搜索加密算法和多参数动态激励模型, 构建了一种基于身份授权的可信去中心化存储网络: 为了规避节点好奇的风险, 在去中心化存储网络中融入一种创新的属性基可搜索加密算法, 基于可搜索加密算法保护存储在节点中的数据; 为了规避节点不诚实的风险, 提出了基于声誉值的多参数动态激励模型, 用于奖励执行诚实检索的节点, 同时惩罚不诚实的节点。方案分析表明, 本文提出的基于身份授权的可信去中心化存储网络在多对多数据共享过程中可以避免被好奇节点窃取信息, 在检索过程中可以避免被不诚实节点返回虚假结果, 与现有成果相比, 本文在属性基可搜索加密算法方面提升了公钥和主密钥的生成效率, 将计算集中在 Setup 阶段, 保持后续检索的高效性, 此外, 本文也对创新的属性基可搜索加密算法进行了安全性分析; 在激励模型方面, 充分考虑节点的声誉值情况, 对去中心化存储网络中的搜索节点和验证节点进行合适的奖惩, 在维持去中心化网络节点活跃性的同时, 鼓励节点进行诚实检索, 总体而言, 基于身份授权的可信去中心化存储网络相比于目前的去中心化存储网络具有较为明显的优势。

关键词　去中心化存储网络; 属性基可搜索加密; 激励模型

中图法分类号　TP311.5　DOI 号 10.19363/J.cnki.cn10-1380/tn.2025.05.04

Trusted Decentralized Storage Network Based on Identity Authorization

ZHANG Jingyu, LIU Qi, PENG Hongrui, YANG Zenghui, YUAN Kaiguo, LI Xiaoyong

Beijing University of Posts and Telecommunications, Beijing 100876, China

Abstract　With the advantages of high reliability, low cost, and high speed, decentralized storage network is increasingly used in various industries, but there are two major characteristics of node curiosity and dishonesty in decentralized storage network, which lead to two risks that decentralized network nodes may steal user information and users cannot get correct results when searching, respectively. In this paper, a trusted decentralized storage network based on identity authorization is constructed by introducing an attribute-based searchable encryption algorithm and a multi-parameter dynamic incentive model in the decentralized storage network: to avoid the risk of node curiosity, an innovative attribute-based searchable encryption algorithm is incorporated into the decentralized storage network. The searchable encryption algorithm could protect the data stored in the nodes; to avoid the risk of node dishonesty, a multi-parameter dynamic incentive model based on reputation value is proposed to reward nodes that perform honest retrieval while punishing dishonest nodes. The scheme analysis shows that the trusted decentralized storage network based on identity authorization proposed in this paper can avoid information theft by curious nodes in the process of multi-owner to multi-user data sharing, and can avoid false results returned by dishonest nodes in the retrieval process; compared with existing results, this paper improves the efficiency of public key and master key generation in terms of attribute-based searchable encryption algorithm, and the calculation is concentrated in the Setup stage to maintain the high efficiency of subsequent retrieval. In addition, this paper also conducts a security analysis on the innovative attribute-based searchable encryption algorithm; in terms of incentive model, the nodes' reputation value situation is fully considered, and the search nodes and verification nodes in the decentralized storage network are appropriately rewarded and punished to encourage the nodes to perform honest retrieval while maintaining the activity of the decentralized network nodes. Overall, the trusted decentralized storage network based on identity authorization has more obvious advantages over the current decentralized storage network.

Key words　decentralized storage network; attribute-based searchable encryption; incentive model

通讯作者: 袁开国, 博士, 讲师, Email: flyingdreaming@bupt.edu.cn。

本课题得到国家自然基金项目(No. U1836215)、北京邮电大学大学生创新创业项目(No. 202213012)资助

收稿日期: 2023-04-18; 修改日期: 2023-10-09; 定稿日期: 2025-02-28

1 引言

1.1 目的

目前基于数据中心和云存储的集中式数据存储方案存在数据权限确认和中心信任的问题[1]。去中心化存储网络可以在一定程度上解决这些问题, 其主要特征之一是共享资源, 包括共享内容、存储和 CPU 能力等。同时, 去中心化存储网络具有容错性好、可扩展性高和性能高等优点[2]。因为去中心化存储网络是建立在区块链技术之上, 所以没有一个中心节点来控制网络, 相比其他类型的数据存储更安全。然而, 在去中心化存储网络中的节点可能存在好奇的和不诚实的问题。

针对节点的好奇问题, 本文将属性基可搜索加密引入到去中心化存储网络中。目前, 大多数属性基可搜索加密方案都是在云环境中使用的[3-6]。可搜索加密提供了一种有效的机制, 实现了对加密数据的安全搜索。可搜索加密的一个常见应用模型是: 数据所有者将加密数据存储在服务器上, 服务器可以根据数据用户提交的查询陷门对加密数据进行有效的基于关键字的搜索, 其中所有者的数据和用户的查询被秘密保存在服务器中[7]。在去中心化存储网络中融入属性基可搜索加密, 使去中心化网络具有身份授权功能, 只有满足访问策略的用户才能够检索到密文并进行解密。

针对节点不诚实问题, 本文将激励模型引入到去中心化存储网络中[8]。传统的去中心化存储网络存在一些不可避免的缺陷, 如系统不稳定、缺乏审计和激励机制等[9], 而区块链具有一系列优良特性, 包括去中心化、不变性、透明性和可审计性等, 恰好可以弥补去中心化网络在这方面的不足。E Daniel 等人[10]提供了下一代对等数据网络的相关定性比较。R Kumar 等人[11]将星际文件系统(InterPlanetary File System, IPFS)[12]和区块链结合起来实现了一个去中心化文件存储网络和访问框架。S Xuan 等人[13]提出了一种基于进化博弈论的数据共享激励模型, 并引入带有智能合约的区块链, 其中, 智能合约机制可以动态控制激励参数, 持续鼓励用户参与数据共享。Kamboj P 等人[14]提出在基于角色的访问控制中使用基于区块链的智能合约进行用户身份验证。通过上述研究, 本文为去中心化存储网络引入一种基于声誉值的多参数动态激励模型, 鼓励去中心化存储网络中的搜索节点积极参与检索倒排索引表, 并通过多个验证节点进行验证, 去中心化网络奖励验证通过的搜索节点, 并惩罚不诚实的搜索节点。

1.2 本文贡献

为了在多所有者对多用户模型下的去中心化存储网络中实现安全可信的搜索, 本文提出了一种基于身份授权的可信去中心化存储网络。贡献如下:

(1) 将属性基可搜索加密(Attribute-Based Searchable Encryption, ABSE)与密文策略属性基加密(Ciphertext-Policy Attribute-Based Encryption, CP-ABE)相结合, 解决了节点的好奇问题, 实现了隐私保护和安全的数据共享, 适用于多对多搜索场景。在方案构造中, 对倒排文件索引进行属性基可搜索加密, 当节点对去中心化存储网络维护的倒排索引表进行搜索时, 满足了实际场景的安全需求。节点上传的数据文件采用对称加密算法进行加密, 并使用 CP-ABE 对对称密钥进行加密, 将两者结合起来组成数据文件密文存储在 IPFS 中。将属性基可搜索加密和 CP-ABE 配合使用, 保证了多所有者与多用户之间的安全数据共享。

(2) 在去中心化存储网络中融入基于声誉值的多参数动态激励模型。只要搜索节点按照激励模型诚实地执行检索操作, 就可以获得相应的奖励, 反之, 则会受到惩罚, 这样用户就可以得到正确的检索结果。通过激励模型, 鼓励节点参与去中心化网络中的检索活动和验证活动, 积极执行检索任务并返回正确的结果, 有效地保证了去中心化网络中节点的诚实性。

(3) 改进了属性基可搜索加密算法。本文改进了公钥和主密钥的结构, 使算法在 Setup 阶段的速度大大加快, 并且节约了存储空间。改进后的方案将计算集中在 Setup 阶段, 在索引生成、搜索令牌生成、检索效率和验证正确性等方面具有一定优势。对属性基可搜索加密算法进行的正确性分析验证了新方案的可行性, 算法的性能与效率分析验证了新方案的高效性。

2 相关工作

2.1 密文策略属性基加密

CP-ABE 使用属性刻画用户的资格, 由数据加密方制定密文访问策略, 决定谁可以解密密文。在 CP-ABE 中, 用户的私钥与一组属性相关联, 只有用户的属性符合密文的访问结构, 密文才能被解密。CP-ABE 更适用于云存储访问控制类场景, 在 CP-ABE 最初的研究中, Michalass 等人[15]主要以单授权机构模式为主, 缺乏对数据的安全保障。Vaanchig 等人[16]以用户全局标识符的方式增强方案的抗合谋攻击能力, 但该方案还面临着用户计算开销大的问

题。Fa 等人[17]将加密以及解密的计算开销外包给云服务商, 以此提升了用户加解密的效率, 但其安全性还有待提高。

2.2 属性基可搜索加密

自文献[18]提出对称可搜索加密机制以来, 各种搜索方案相继提出。可搜索加密技术解决了基于远程存储网络的密文搜索问题, 提高了云环境和去中心化存储的实用性。为了去除基于身份的加密方案中的可信机构, Sahai 和 Waters[19]提出了 FIBE 方法作为属性基加密的原型。在现实环境中, 属性基加密作为解决多个数据所有者和多个数据用户之间安全共享数据的有效策略, 受到了广泛的关注和讨论。随后, Guo 等人[20]提出了一种 ABE 和 SE 技术融合的方案, 他们设计了一种高效的属性基可搜索加密(Efficient Attribute-Based Searchable Encryption, EABSE)方案来实现云存储方案上的安全关键词搜索, 允许云服务器对密文执行搜索操作, 并禁止云获得除搜索结果之外的更多信息。

2.3 去中心化存储

去中心化存储是安全数据存储的一种解决方案。一方面, 去中心化存储网络具有高度的抗毁性, 一旦存储节点受到攻击, 数据丢失风险相对较低; 另一方面, 去中心化的存储网络可以减少数据垄断的出现。

目前成熟的去中心化存储网络有很多, 如 IPFS、基于以太坊的集群存储网络[21]、Arweave、Storj、Sia[22]等方案。但上述方案中, 有些不使用加密的明文存储, 易产生未经授权的数据访问; 有些采用的加密方式只能适用于一对一或一对多的场景。

当去中心化存储网络引入属性基可搜索加密算法和激励模型时, 既可以有效解决去中心化节点既好奇又不诚实的问题, 也可以使整个存储网络适用于多所有者对多用户的搜索场景。

2.4 激励模型

激励模型是保证去中心化存储网络中每个节点都能提供诚实可信服务的重要手段。它通过经济平衡, 鼓励网络的每个节点参与整个网络的安全运行, 防止其篡改总账本。是长期维持区块链网络运行的原动力, 更好地实现数据的持久化存储。

现有的激励模型大多是基于区块链的加密货币[23], 包括比特币、ETH、Lite-coin、XRP、Zerocoin、FIL 等, 这些不同的加密货币的共识机制、效率、特征都有很大的不同[24]。以其中最著名的比特币和 ETH 为例分析: 比特币使用工作量证明的 PoW 共识协议[25], 但 PoW 共识协议成本较高, 消耗大量计算资源。ETH 是基于以太坊的数字令牌, 支持可编程智能合约, 正在从 PoW 共识机制向 PoS 共识机制转变。然而, PoS 机制的安全性完全依赖于高资产的矿工节点, 因此不适用于高安全性的应用场景。

上述区块链技术的应用研究侧重于保证交易数据的全局一致性, 对于改进共识层与激励层相关机制, 进而提升共识效率方面的研究较少, 为此, 本文采用了一种基于声誉值的链上动态激励模型[26], 根据节点的行为建立动态的声誉值模型, 并根据节点声誉值为诚实节点提供量化奖励, 同时对恶意节点实施惩罚措施。

3 先验知识

3.1 对称双线性组[27]

设 G_1 和 G_2 是 q 阶的两个乘法循环群, 其中 q 是一个大素数。如果一个映射 $e: G_1 \times G_2 \to G_T$ (G_T 是 q 阶的乘法循环群)满足下列性质, 则它是双线性对。

(1) 双线性(Bilinear): 对于任意的 $g_1 \in G_1$, 和任何 $a, b \in Z_q$, 有 $e(g_1^a, g_2^b) = e(g_1, g_2)^{ab}$, 其中 Z_q 为 q 阶整数循环群。

(2) 非简并性(Non-degeneracy): 如果 g_1 是一个 G_1 的生成元, g_2 是一个 G_2 的生成元, 那么, $e(g_1, g_2)$ 是一个 G_T 的生成元。

(3) 可计算性(Computability): 对于任意的 $g_1 \in G_1$, $g_2 \in G_2$ 有高效算法计算 $e(g_1, g_2)$。

假设 $g^a, g^b, g^c \in G_1$,

$$e(g^a g^b, g^c) = e(g^{a+b}, g^c) = e(g, g^c)^{a+b}$$

$$e(g^a, g^c)e(g^b, g^c) = e(g, g^c)^a e(g, g^c)^b = e(g, g^c)^{a+b}$$

得到结论: $e(g^a g^b, g^c) = e(g^a, g^c)e(g^b, g^c)$。

3.2 区块链激励机制

区块链的激励层主要包括经济激励的发行制度和分配制度, 其功能是提供激励措施, 鼓励节点参与区块链中安全验证工作, 并将经济因素纳入到区块链技术体系中, 激励遵守规则参与记账的节点并惩罚不遵守规则的节点。通常, 基于信任度的激励机制依赖于网络节点之间历史时间内的交互行为来建立其可信度, 并根据其信任度对区块链网络中的节点进行分类; 另一方面, 激励措施侧重于针对节点是否具有合作行为而采取的相应措施。可以把区块链激励机制理解为是在参与者之间建立信任的一种方式, 激励机制是通过经济平衡的手段, 鼓励节点参与到维护区块链系统安全运行中来, 防止对总帐

本进行篡改、是长期维持区块链网络运行的动力。

3.3 IPFS

IPFS 是一种点对点的分布式文件系统, 使用了分布式哈希表、BitTorrent、版本管理系统、自验证文件系统和 Merkle DAG 等关键技术。在 IPFS 系统中, 各节点都是平等的, 没有统一的中央服务器, 数据被分割后冗余的存储在 IPFS 网络的各节点中, 任何节点都没有特权, 用户也不需要担心数据被垄断。用户可以通过文件的哈希值来寻找文件的存储位置。这种存储方式的优势是只要文件内容稍有改变, 就会导致数据存储的哈希值产生变化, 有效地避免了数据被人为恶意篡改的风险。此外, 这种分布式的数据管理增强了存储系统的抗攻击能力, 单一节点的故障并不会影响到整个存储系统。使用 IPFS 系统可以降低存储和带宽的成本, 适合运用到大量数据收集的场景。

4 模型与算法

4.1 去中心化存储网络模型

基于身份授权的可信去中心化存储网络模型如图 1 所示, 适用于多所有者对多用户的关键词搜索。它包括 8 个实体: 属性机构(Attribute Authority, AA)、数据所有者(Data Owner, DO)、数据用户(Data User, DU)、搜索节点(Search Node, SN)、验证节点(Verify Node, VN)、去中心化网络(Decentralized Network, DN)、IPFS 和激励模型(Incentive Model, IM)。

(1) 属性机构 AA: AA 是一个可信的组织, 主要作用是生成系统安全参数 λ, 系统主密钥 *msk*、系统公钥 *pk*, 并定义属性集 *U*(包括用户的所有属性)。当一个节点申请加入去中心化存储网络时, 属性授权机构 AA 为其分配一个唯一的身份, 并对该节点的属性 *Attr* 进行认证。

(2) 数据所有者 DO: DO 提取关键词字典 $WD=\{w_1,w_2,\cdots,w_m\}$ 来自多个数据文件, 并使用属性基可搜索加密算法为 *WD* 生成加密的索引 *cphW*。DO 用随机对称密钥 *Key*, 对数据文件进行加密, 得到数据文件的密文, 并用 CP-ABE 对 *Key* 进行加密, 将数据文件的密文和对称密钥的密文统一保存为密文 *cphf*, 并上传 *cphf* 到 IPFS, 得到与密文对应的标识符 ID 为 *cphfid*。多个 *cphfid* 共同构成了倒排文件列表 *cphFID*, 对的加密索引为于 *cphW*。

(3) 数据用户 DU: DU 使用自己拥有的属性 *Attr*, *msk* 和 *pk* 生成私钥 *sk*。使用 *sk*, *pk* 和搜索关键词 *w* 生成 *token*, 将 *token* 上传至去中心化网络 DN, 进而搜索节点进行检索活动。

(4) 搜索节点 SN: SN 是区块链网络中的节点, 负责执行检索操作, 使用 DU 上传的 *token* 去匹配倒排索引表中的加密索引。

(5) 验证节点 VN: VN 也是区块链网络中的节点, 多个验证节点对 SN 返回的检索结果进行验证, 如果验证成功, 则将验证成功结果返回给 DU, 让 DU 对密文进行解密, 获取明文。

(6) 去中心化网络: DN 具有分散的特性, 去中心化存储网络中的所有结点共同维护一张倒排索引表, 方便了搜索节点快速检索与关键词陷门匹配的密文索引。

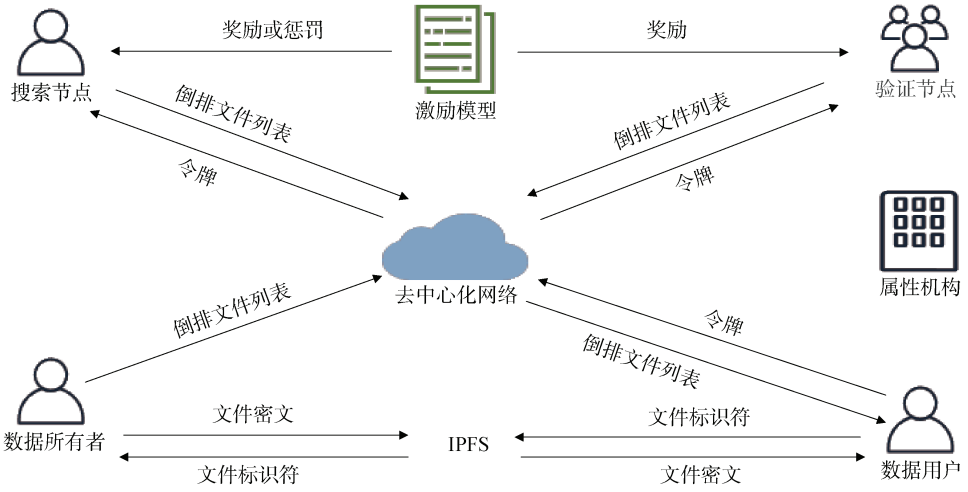

图 1 基于身份授权的可信去中心化存储网络模型

Figure 1 Trusted Decentralized Storage Network Model based on Identity Authorization

(7) IPFS: 负责存储加密后的文件, 并返回文件标识符 ID, 多个文件标识符 ID 形成一个倒排文件列表。

(8) 激励模型: 在去中心化网络中引入 IM 的目

的是让节点保持活跃, 鼓励搜索节点诚实的返回搜索结果, 对不诚实的节点进行惩罚。

4.2 属性基可搜索加密算法

根据去中心化存储网络模型, 可搜索加密算法由 4 个独立的实体执行: AA、DO、DU 和 SN。每个实体的功能如下。

(1) AA: 生成系统安全参数并定义属性集 U(包括用户的所有属性)。并且 AA 认证分布式网络中每个用户的属性并生成公钥和主密钥。

- $(msk, pk) \leftarrow Setup(U, 1^{\lambda})$: 初始化公钥 pk 和主密钥 msk。输入安全参数 λ 和属性集 U(其中包括所有用户的属性), 算法输出 msk 和 pk。

(2) DO: DO 在上传倒排文件列表之前加密文件, 并加密文件索引。

- $(cphW, cphFID) \leftarrow Enc(S, WD, F, pk)$: 使用访问策略 S, 关键字字典 $WD = \{w_1, w_2, \cdots, w_m\}$ 和文件集合 $F = \{f_1, f_2, \cdots, f_n\}$, 加密算法生成加密索引 $cphW$ 和文件的密文。考虑到效率问题, 文件加密采用对称加密, 对称加密密钥采用 CP-ABE 加密。数据密文和密钥密文组成最终的密文 $cphf$。将 $cphf$ 上传到 IPFS 中, 获取文件标识符 $cphfid$, 多个 $cphfid$ 形成倒排文件列表 $cphFID$。

(3) DU: DU 为想要查询的关键字生成令牌。

- $sk \leftarrow KeyGen(Attr, msk, pk)$: 输入用户的属性 $Attr$、主密钥 msk 和公钥 pk, 算法输出用户的私钥 sk;
- $tok \leftarrow TokenGen(pk, sk, w)$: 使用上一步生成的 sk 和要查询的关键字 w、公钥 pk 生成用户搜索令牌 tok。

(4) SN: 网络中的搜索节点执行检索算法, 将 DU 上传的 tok 与倒排索引表中的加密索引 $cphW$ 匹配。如果索引匹配成功, 则将匹配到的加密索引对应的倒排文件列表 $cphFID$ 返回给用户。

- $rslt \leftarrow Search(tok, cphW, cphFID)$: 如果 $cphW$ 可以与 tok 匹配, 则该算法输出 $cphW$ 对应的倒排索引列表 $cphFID$。

4.3 基于声誉的动态激励模型

本文构建起了一种基于节点声誉的存储交易市场的奖惩模型。基于声誉的激励模型核心思想在于量化评估节点的信誉值, 以此来衡量去中心化网络中节点的可靠程度, 以便通过最可靠的节点转发数据包。一般规定, 当一个节点正确地执行去中心化网络所分配的任务时, 它的声誉值就会提高。此类激励模型通过观测其他网络节点的声誉值, 同时结合节点隔离策略(被隔离节点一般为信誉值较低的节点), 旨在实现数据转发的高效可靠。同样, 声誉机制可以优先考虑行为良好的节点的流量。

在去中心化节点网络中, 节点通常主要承担检索的任务, 那么在一次检索任务的委派与执行过程中, 当搜索节点成功执行了检索任务并正确返回了检索任务的结果时, 根据基于声誉的激励模型, 该节点将会得到一定的声誉值奖励与代币奖励; 另一方面, 多个验证节点进行验证, 也会获得相应的奖励。声誉值的高低将会影响着节点在一次任务中获得代币的数量, 同时, 声誉值的评估可由多种动态参数共同影响, 如节点行为等影响因子使声誉值的评估阶段呈现动态性, 从而让整个去中心化网络实现动态的激励。在第 6 节将会详细阐述该激励模型的构建与分析。

5 算法分析

5.1 算法构建

方案采用倒排索引结构, 使用与门作为访问控制, 实现可搜索加密。为了提高计算效率和节省空间开销, 本文改进了公钥 pk 和主密钥 msk 的结构。具体算法构建包括以下五个主要算法。

Init 阶段: 假设所有的属性都包含在属性集合 $U = \{attr_1, attr_2, \cdots, attr_n\}$, 其中 n 是 U 的大小。对于每个属性 $attr_i$ $(1 \leqslant i \leqslant n)$, 有 v_i 和 $\neg v_i$ 两个值。设置一个用户的属性集合 $Attr$, 如果 $attr_i$ $(1 \leqslant i \leqslant n)$ 属于 $Attr$, 则属性 $attr_i$ 的值是 v_i, 否则 $attr_i$ 的值是 $\neg v_i$。为了使属性的描述形式化, 本文采用属性的值来表示用户属性集中是否包含该属性。

Setup 阶段: 给定一个双线性组 $e: G \times G \rightarrow G_T$, p 作为 G 和 G_T 的素数阶, 并且 $H: \{0,1\}^* \rightarrow Z_p$ 作为单向散列函数。随机选择三个数字 $a, b, c \leftarrow Z_p$, 集合 $\{r_1, r_2, \cdots, r_n\} \leftarrow Z_p$ 和集合 $\{x_1, x_2, \cdots, x_n\} \leftarrow G$。计算 $u_i = g^{-r_i}$ 和 $y_i = e(x_i, g)$, 其中 $1 \leqslant i \leqslant n$。然后, 输出公钥 $pk = (g, g^a, g^b, g^c, (u_i, y_i) \mid 1 \leqslant i \leqslant n)$ 和主密钥 $msk = (a, b, c, (r_i, x_i) \mid 1 \leqslant i \leqslant n)$。

Enc 阶段: 随机选择 $t_1, t_2 \in Z_p$。假设当且仅当 $v'_i = v_i$ 或 $\neg v_i$ 时, 访问策略结构为 $S = \bigwedge_{v_i \in U} v'_i$。如果 $v'_i = v_i$, 则设置 $u'_i = u_i$, 否则 $u'_i = \neg u_i$。计算 $u_{gate} = g^{t_2} \prod_{i=1}^{n} u'_i$。为每个关键字 $w \in WD$, 设置 $W' = g^{ct_1}$, $W = g^{a(t_1+t_2)} g^{bH(w)t_1}$。文件内容采用对称加密, 对称加密的密钥采用 CP-ABE 加密。数据密文和

密钥密文共同构成最终文件密文。然后将其上传到IPFS 中, 获取文件标识符。与 WD 相关联的多个文件标识符形成了一个倒排文件列表 $cphFID$。令 $cphW=(W',W,u_{gate})$, 得到 $cph=(cphW,cphFID)$ 作为 Enc 阶段的输出结果。

KeyGen 阶段: 首先, 令 $v=g^{ac}$。对于用户属性集合中的每个属性 v_i^*, 如果 $v_i^*=v_i$, 令 $y_i^*=y_i$, 否则 $y_i^*=e(\neg x_i,g)$。同样, 如果 $v_i^*=v_i$, 那么 $\sigma_i^*=x_iv^{r_i}$, 否则 $\sigma_i^*=\neg x_iv^{\neg r_i}$。令 $\sigma_{user}=\prod_{i=1}^{n}\sigma_i^*$, 最后输出密钥 $sk=(y_{user}=\prod_{i=1}^{n}y_i^*,<v,\sigma_{user}>)$。

TokenGen 阶段: 选择 $s\leftarrow Z_p$。为了给关键词 w 生成搜索令牌, 计算 $tok1=(g^ag^{bH(w)})^s$, $tok2=g^{cs}$。因此, 搜索令牌 $tok=(y_{user}^s,<v^s,\sigma_{user}^s>,tok1,tok2)$。

Search 阶段: 首先计算

$$E=\frac{e(u_{gate},v^s)e(\sigma_{user}^s,g)}{y_{user}^s}$$

如果用户的属性满足密文的访问策略, 则 $E=e(g,g)^{acst_2}$ 和 $e(W',tok1)E=e(W,tok2)$ 成立。

如果 DU 的属性满足加密关键字所使用的访问策略, 则搜索令牌 tok 可以与倒排索引中的 $cphW$ 匹配。然后将 $cphW$ 对应的 $cphFID$ 返回给 DU。经过验证节点验证后, DU 根据 $cphFID$ 从 IPFS 下载加密后的数据文件, 并在本地解密, 得到明文。

5.2 正确性分析

当SN 接收 DU 上传的 tok, 其使用 tok 来匹配倒排索引表中的加密索引, 如果 DU 的属性满足加密索引所指定的访问策略, $E=e(g,g)^{acst_2}$ 和 $e(W',tok1)E=e(W,tok2)$ 成立, 可以成功匹配, 证明如下:

1) $E=e(g,g)^{acst_2}$ 的证明:

$$\begin{aligned}E&=\frac{e(u_{gate},v^s)e(\sigma_{user}^s,g)}{y_{user}^s}\\&=\frac{e(g^{t_2}\prod_{i=1}^{n}u_i,g^{acs})e((\prod_{i=1}^{n}x_iv^{r_i})^s,g)}{(\prod_{i=1}^{n}e(x_i,g))^s}\\&=\frac{e(g,g)^{acst_2}e(\prod_{i=1}^{n}g^{-r_i},g^{acs})e(\prod_{i=1}^{n}x_i^s,g)e(\prod_{i=1}^{n}g^{acr_is},g)}{(\prod_{i=1}^{n}e(x_i,g))^s}\\&=e(g,g)^{acst_2}\end{aligned}$$

2) $e(W',tok1)E=e(W,tok2)$ 的证明:

$$\begin{aligned}e(W',tok1)E&=e(g^{ct_1},(g^ag^{bH(w)})^s)e(g,g)^{acst_2}\\&=e(g,g)^{acst_1}e(g,g)^{bcst_1H(w)}e(g,g)^{acst_2}\end{aligned}\quad(1)$$

$$\begin{aligned}e(W,tok2)&=e(g^{a(t_1+t_2)}g^{bH(w)t_1},g^{cs})\\&=e(g^{at_1}g^{at_2},g^{cs})e(g^{bH(w)t_1},g^{cs})\\&=e(g,g)^{acst_1}e(g,g)^{acst_2}e(g,g)^{bcst_1H(w)}\end{aligned}\quad(2)$$

因为等式(1)等于等式(2), 所以可证得 $e(W',tok1)E=e(W,tok2)$ 成立。

5.3 性能与效率分析

在去中心化存储的场景中, 节点倾向于更快的响应和更少的带宽成本。在算法方案中, 为了提升计算速度, 减少空间消耗, 在文献[20]的基础上, 本文改进了可搜索加密的 Setup 阶段中公钥和主密钥的结构。通过对素数阶 p 的群的计算, 主要从指数运算、乘法运算、配对运算等方面评估时间复杂度, 从 $|G|$、$|G_T|$、$|Z_p|$等方面评估空间复杂度。

在表 1 中, 详细说明了算法比较过程中每个参数的含义。

在表 2 中, 将本文的算法方案与其他算法方案在时间复杂度方面进行了比较。

在表 3 中, 将本文的算法方案与其他算法方案在空间复杂度方面进行了比较。

考虑到去中心化网络中节点的不同的检索能力, 尽可能地将计算和空间消耗放在算法的 Setup 阶段。与文献[7]和文献[30]相比, 本文算法方案在属性基可搜索加密的 Enc, KeyGen, TokenGen, Search 等阶段的整体时间和空间复杂性相对较低。与文献[20]相比, 本文改进了属性基可搜索加密的 Setup 阶段, 极大地降低了时间和空间复杂度, 同时仍然保持了算法后续步骤的效率, 为去中心化存储网络提供了实时保证。

表 1 各参数含义

Table 1 The Meaning of Each Parameter

参数	描述
E	对群 G 中元素进行指数运算
E_T	对群 G_T 中元素进行指数运算
M	对群 G 中元素进行乘法运算
M_T	对群 G_T 中元素进行乘法运算
P	配对操作
$\|G\|$	群 G 的空间大小
$\|G_T\|$	群 G_T 的空间大小
$\|Z_p\|$	群 Z_p 的空间大小
A	去中心化网络中属性的数量
N	满足访问策略的属性数量
S	数据用户所拥有的属性的数量

表 2 时间复杂度分析

Table 2 Time Complexity Analysis

	本方案	方案[20]	方案[28]	方案[7]	方案[29]
Setup	$(3+A)E+AP$	$(3+2A)E+2AP$	$3E$	$P+2E$	$(A+3)E+E_T+P$
Enc	$4E+(N+2)M$	$4E+(N+2)M$	$(2N+4)E+M$	$(2N+2)E+P+E_T+M_T$	$(A+3)E+M$
KeyGen	$(S+1)E+2SM+SM_T$	$(S+1)E+2SM+SM_T$	$(2S+2)E+SM$	$(2S+4)E+(S+1)M$	$(S+1)E$
TokenGen	$6E+M$	$6E+M$	$(2S+4)E+M$	$E+M$	$(2S+2)E+SM$
Search	$4P+E_T+3M_T$	$4P+E_T+3M_T$	$(2N+3)P+NE_T+(N+2)M_T$	$(2N+1)P+(N+1)M_T+NE_T$	$(N+1)M+(N+2)P$

表 3 空间复杂度分析

Table 3 Space Complexity Analysis

	本方案	方案[20]	方案[28]	方案[7]	方案[29]
Setup	$(4+2A)\|G\|+A\|G_T\|+(3+A)\|Z_p\|$	$(4+4A)\|G\|+2A\|G_T\|+(3+2n)\|Z_p\|$	$5\|G\|+3\|Z_p\|$	$3\|G\|+\|G_T\|+\|Z_p\|$	$(A+5)\|G\|+\|G_T\|$
Enc	$3\|G\|$	$3\|G\|$	$(2N+3)\|G\|$	$(2N+1)\|G\|+\|G_T\|$	$(N+1)\|G\|$
KeyGen	$2\|G\|+\|G_T\|$	$2\|G\|+\|G_T\|$	$(2S+1)\|G\|$	$(2S+2)\|G\|$	$(S+1)\|G\|$
TokenGen	$4\|G\|+\|G_T\|$	$4\|G\|+\|G_T\|$	$(2S+3)\|G\|$	$(2S+1)\|G\|$	$(S+2)\|G\|$

5.4 算法实现

算法基于 JPBC 库实现, 在 Windows 11 中进行仿真实验, 计算机配置为 Intel(R) Core(TM) i7-10875H CPU, 2.30GHz 主频, 16GB 内存, 64 位操作系统。在实验仿真过程中, 将涉及到的属性数量(用户属性数量或访问策略属性数量)从 1 增加到 50, 步长为 10, 并将每个实验运行 10 次以获得平均执行时间。

考虑到去中心化网络节点检索能力差异, 本方案将大量计算都放在 Setup 阶段, 尽量减少 Search 阶段的计算开销, 保证整个去中心化存储网络的高效性。为了证明改进后的算法在去中心化存储网络中的有效性, 本节从属性基可搜索加密的 Setup, Enc, KeyGen, TokenGen, Search 五个阶段分别对比本算法方案、文献[20]、文献[28]、文献[7]的计算开销。

如图 2 所示, 在属性基可搜索加密的 Setup 阶段, 本算法方案相比于改进前的文献[20]节省了大量计算开销。

在 Enc 阶段, 如图 3 所示, 本算法方案、文献[28]、文献[7]的计算开销均随着访问策略属性数量的增加而增加, 本算法方案优于文献[28]的方案, 但逊色于文献[7]的方案。

在 KeyGen 阶段, 图 4 显示了本方案优于文献[28]和文献[7]。

在 TokenGen 阶段, 图 5 显示了本方案同样优于文献[28], 并且与文献[7]在该阶段的计算开销基本相当。

在图 6 的 Search 阶段显示了本方案的优越性, 本方案对于单个加密索引检索时间与属性数量无关, 耗时基本不变, 维持在 43ms 左右, 而文献[28]与文献[7]在该阶段的计算开销均与属性数量呈线性关系, 当属性数量很大时, 计算开销也会很大, 会降低搜

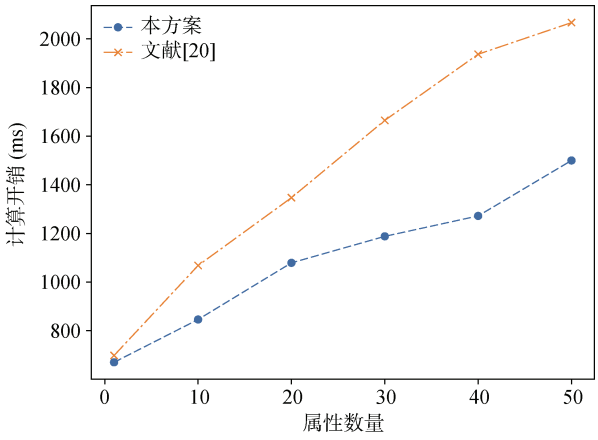

图 2 Setup 阶段计算开销对比

Figure 2 Comparison of calculation cost in Setup phase

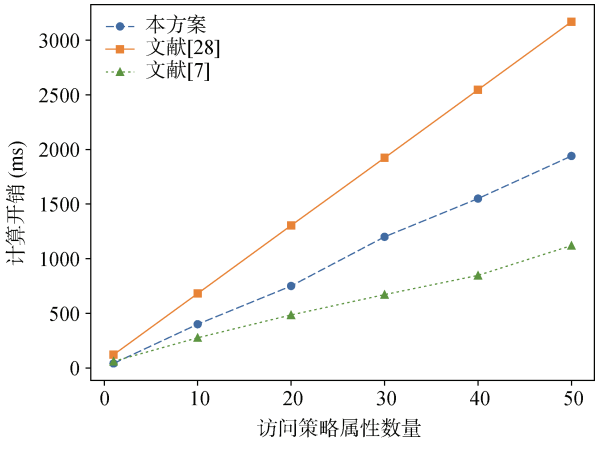

图 3 Enc 阶段计算开销对比

Figure 3 Comparison of calculation cost in Enc phase

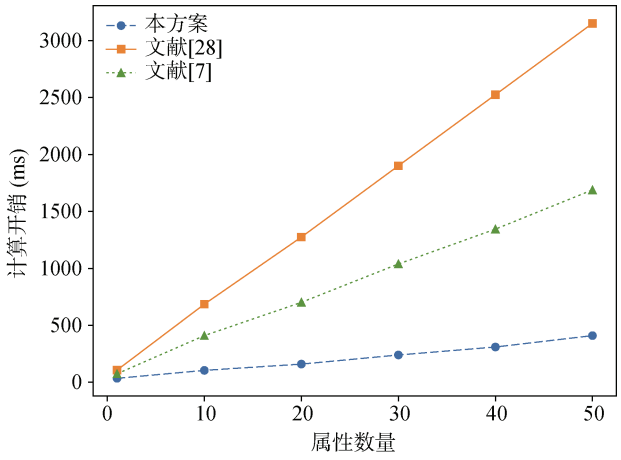

图 4 KeyGen 阶段计算开销对比

Figure 4 Comparison of calculation cost in KeyGen phase

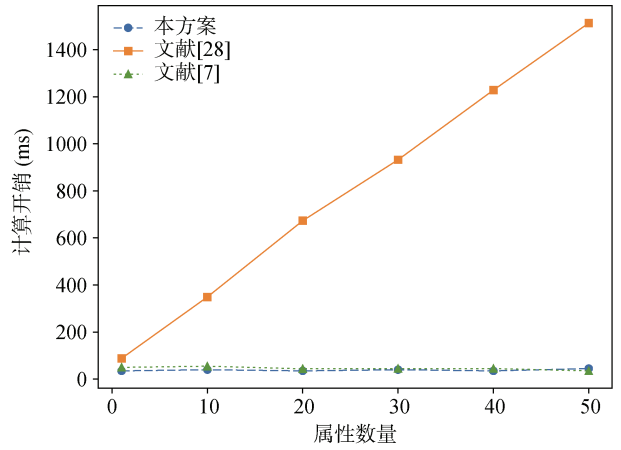

图 5 TokenGen 阶段计算开销对比

Figure 5 Comparison of calculation cost in TokenGen phase

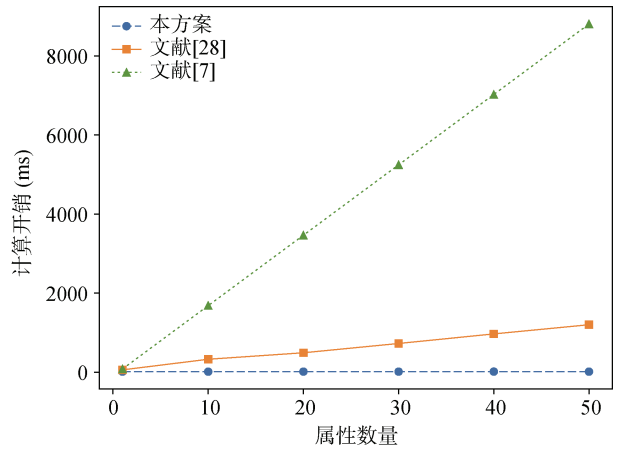

图 6 Search 阶段计算开销对比

Figure 6 Comparison of calculation cost in Search phase

索节点在去中心化存储网络中的检索效率, 不符合本文构建的去中心化存储网络的要求。

5.5 安全性分析

在随机预言模型下, 我们的方案被证明是可以抵抗选择性关键词攻击的。选择性关键词攻击的安全游戏定义如下:

初始化: 首先, 攻击者将访问策略 S 发送给挑战者, 挑战者执行 Setup 算法, 并保留主密钥 msk。

阶段 1: 通过执行 KeyGen 算法, 攻击者向挑战者发送属性集合 $Attr$, 如果属性能够满足访问策略 S, 则从挑战者处获得与属性相关联的秘钥 sk, 否则过程终止;之后根据 TokenGen 算法, 攻击者在输入 sk 和关键字 w 后, 可以获得一个搜索陷门。

挑战: 攻击者随机选择两个关键词 w_0 和 w_1。挑战者随机选择其中一个关键词 $w_b(b\in\{0,1\})$, 然后运行 Enc 算法。挑战者向对手返回 w_b 的密文 cph^*。

阶段 2: 攻击者可以重复阶段 1 的操作对不同关键词的陷门进行查询, 但是不能再查询关键字 w_0 和 w_1。

猜测: 攻击者给出一个对 b 的猜测 b', 如果 $b'=b$, 攻击者赢得游戏的胜利。

上述游戏中, 攻击者的优势定义为 $\Pr[b'=b]-\frac{1}{2}$, 如果没有攻击者能够在多项式时间内以不可忽略的优势赢得上述安全游戏, 则本方案是安全的。

DL 假设: 根据系统安全参数选择阶数为 p 的乘法循环群 G, 随机选取 $g,f,h\in G$ 和 $r_1,r_2\in Z_p$。DL 假设是给定多元组 $(g,f,h,f^{r_1},g^{r_2},Q)$, 其中 $Q=h^{r_1+r_2}$。攻击者可以打破该假设的优势可以定义为: $adv=\left|\Pr[\mathcal{A}(g,f,h,f^{r_1},g^{r_2},h^{r_1+r_2})=1]-\Pr[\mathcal{A}(g,f,h,f^{r_1},g^{r_2},Q)=1]\right|$。

如果不存在一种算法 $\mathcal{A}$ 能够在多项式时间内以不可忽略的概率区分 $(g,f,h,f^{r_1},g^{r_2},Q)$, 即打破该假设的优势可忽略, 则 DL 假设成立。

定理1: 如果攻击者能以不可忽略的优势k赢得上述安全游戏, 则挑战者能以 $\frac{k}{2}$ 的优势解决DL困难问题。

安全性证明:

初始化: 挑战者执行 Setup 算法, 随机选择 $\{r_1,r_2,\cdots,r_n\}\leftarrow Z_p$, $\{x_1,x_2,\cdots,x_n\}\leftarrow G$。令 $g^a=h$, $g^c=f,g^b=f^d$, 其中 $a,b,c,d\in Z_p$。计算 $u_i=g^{-r_i}$ 和 $y_i=e(x_i,g)$。输出公钥 $pm=(h,f^d,f,(r_i,x_i)|1\leqslant i\leqslant n)$ 和主密钥 $mk=(d,(r_i,x_i)|1\leqslant i\leqslant n)$。攻击者将访问策略 S 发送给挑战者。

阶段 1: 过执行 KeyGen 算法, 攻击者向挑战者

发送属性集合 $Attr$, 首先, 令 $v=f^{\kappa}$, 其中 $\kappa \in Z_p$。对于用户属性集合中的每个属性, 计算 σ_i 以及 y_i。令 $\sigma_{user}=\prod_{i=1}^{n}\sigma_i$, $y_{user}=\prod_{i=1}^{n}y_i$。最后攻击者得到密钥 $sk=(y_{user},<v,\sigma_{user}>)$。之后根据 TokenGen 算法, 随机选择 $s \in Z_p$, 攻击者在输入 sk 和关键词 w 后可以得到搜索陷门:

$$tok=(tok1=(hf^{dH(w)})^s, tok2=f^s, <v^s, \sigma_{user}^s>, y_{user}^s)。$$

挑战: 攻击者随机选择两个关键词 w_0 和 w_1。挑战者随机选择其中一个关键词 $w_b(b \in \{0,1\})$, 然后运行 Enc 算法, 随机选择 $t_1, t_2 \in Z_p$, 计算 $u_{gate}=g^{t_2}\prod_{i=1}^{n}u_i$, $W'=f^{t_1}$, $W=Qf^{dH(w_b)t_1}$。挑战者向对手返回 w_b 的密文 $cphW^*=(W', W, u_{gate})$。

阶段 2: 攻击者可以重复阶段 1 的操作对不同关键词的陷门进行查询, 但是不能再查询关键字 w_0 和 w_1。

猜测: 攻击者给出一个对 b 的猜测 b', 如果 $b'=b$, 则挑战者输出 $\mu=0$ 来猜测 $Q=h^{r_1+r_2}$, 否则输出 $\mu=1$ 表示 Q 为群 G 中的随机元素。

假设攻击者可以以不可忽略的优势 k 打破该方案, 当 $\mu=0$, 即 $Q=h^{r_1+r_2}$ 时, 有 $\Pr[b'=b \mid \mu=0]=k+\frac{1}{2}$。当 $\mu=1$ 时, 即 Q 为群 G 中的随机元素。因此, 有 $\Pr[b'=b \mid \mu=1]=\frac{1}{2}$。最后, 可以得到 $adv=\frac{1}{2}\cdot\Pr[b'=b]-\frac{1}{2}=\frac{1}{2}\cdot(k+\frac{1}{2}+\frac{1}{2})-\frac{1}{2}=\frac{k}{2}$。所以在多项式时间内, 若攻击者能以不可忽略的优势 k 攻破本方案, 则挑战者可以以 $\frac{k}{2}$ 的优势攻破 DL 问题。因此, 若 DL 假设成立, 则没有多项式攻击者能够以不可忽略的优势选择性地攻破本文方案。

6 激励模型分析

6.1 激励模型构建

本文提出了一种基于声誉的动态激励模型, 该模型会在一次检索任务的委派与执行过程中, 当某节点成功执行了其任务时, 给予该节点一定的声誉值(*reputa*)与相应数量的代币奖励(*reward*)。针对不同的任务阶段与状态, 有以下并行策略:

(1) 初始: 去中心化网络中各节点具有一个初始值为定值 R_{min} 的声誉值。该值在节点随后参与后续任务阶段过程中, 会随着节点自身行为而发生动态变化, 直到声誉值达到阈值 R_{max}, 然后节点声誉值将重置为初始值 R_{min}。

(2) 检索阶段: 当网络中存在检索任务后, 多个搜索节点进行竞争检索, 最先检索出结果的节点作为胜出搜索节点, 其检索出的结果将在验证阶段进行验证。其余执行检索活动的节点中, 选取检索用时较短的几个节点将作为验证节点, 在验证阶段对检索的结果进行验证。

(3) 验证阶段: 多个验证节点同时验证上阶段产生的检索结果, 若多数验证节点验证通过, 则奖励检索胜出节点与验证节点声誉值与代币; 若多数验证节点验证不通过, 则不给予检索胜出节点代币奖励, 并扣除其相应的声誉值, 只对正确验证节点进行奖励。

(4) 代币奖励: 当验证通过后, 按照各节点的声誉值高低的比例, 对正确验证节点进行代币分发, 高声誉节点在一次任务完整通过后获得更多的代币奖励。所有正确验证节点获取的代币奖励总额为 $reward=C$, 参数 C 为一个常数, 为固定值。然后, 各个正确验证节点获取的奖励表示为: $reward_i=\frac{reputa_i}{\sum_i reputa_i}\times reward$; 而未正确验证的节点不予以代币奖励, 并扣除相应的声誉值。同时, 如果搜索节点检索的数据得到大多数验证节点的验证, 其结果正确, 则给予其声誉值奖励, 并依据检索的数据量大小 $dsize(dsize>0)$, 以及检索用时 $t(t>0)$ 来发放相应的代币奖励, 表示为: $reward=k\frac{e^{disze}}{ln(t+2)}$; 否则, 不予以搜索节点代币奖励, 并扣除相应声誉值。

(5) 声誉值奖惩: 当验证通过后, 参与的各节点会根据其检索或者验证的结果的正确与否, 进行声誉值的奖惩行为。各节点获得的声誉值数量按照其声誉值的高低而衰减或增多, 这意味着高声誉值节点与低声誉值节点相比, 做出贡献时获得的声誉值收益较低, 使得低声誉值节点保持积极性; 另一方面对低声誉节点所做出的扣除声誉值处罚会随着该节点声誉值的降低而加剧。其奖惩的声誉值可表示为:

$$acquire\_reputa=\cos(\frac{reputa\times\pi}{2R_{max}})\times R$$

$$deduct\_reputa=-\cos(\frac{reputa\times\pi}{2R_{max}})\times R$$

6.2 可行性分析

本文的激励模型使用节点的声誉值作为其参与共识行为的量化指标, 所以声誉值可以从一定程度上反映节点的工作量以及其行为是否诚实可信。本小节从激励模型、动态激励以及模型可靠性等方面对方案的可行性作出了分析。

(1) **可行性 1:** 在确定激励模型时, 充分考虑了节点的历史累计声誉值和当前轮次声誉值的变化, 这二者分别是对节点声誉值静态累计和动态变化的直观表现。通过声誉值模型结合节点行为进行动态评定, 然后给予其声誉值奖惩, 从而使得整个激励模型能够全面且合理地量化节点的声誉值。

(2) **可行性2:** 在动态激励过程中, 充分考虑了共识算法过程中可能发生的所有情况, 并分析每种情况下是否需要重新计算或分配节点的声誉值, 从而确立了不同情况下的当前共识轮次声誉值模型, 能够形成对诚实或恶意节点的有效奖惩, 实现共识过程中节点声誉值的动态变化。

(3) **可行性 3:** 在当前轮次声誉值模型中, 如果本轮共识成功, 声誉值较高的诚实节点能够获取更多的奖励, 而声誉值较低的诚实节点虽然获取的奖励较少, 但是能够获取更多的声誉值, 这样做就有利于激发节点的积极性, 让低声誉值节点获取更多声誉值, 高声誉值节点获取更多奖励, 从而保持去中心化存储网络的活跃性。

(4) **可行性 4:** 通过设置节点声誉值的边界条件, 规定了 $R_{min} \leqslant reputa < R_{max}$, 当 *reputa* 低于 R_{min} 时, 将对该实施进行一定的惩罚措施: 罚没一定资产, 并限制其节点参与去中心化存储网络中搜索节点的竞争; 当 *reputa* 达到阈值 R_{max} 时, 节点声誉值将重置为 R_{min}。该边界可以有效防止节点连续作恶以及声誉值无限上升的问题。

(5) **可行性 5:** 在对搜索节点进行代币激励时, 充分考虑了其检索的数据量大小 $dsize(dsize>0)$ 和检索用时 $t(t>0)$ 的影响因素, 将其作为搜索节点对 DU 提供搜索服务的评判标准, 并以此为依据来发放相应的代币奖励。对该过程的算法 $reward = k\dfrac{e^{disze}}{ln(t+2)}$ 进行简要分析, 分析过程如下:

1) 一般情况下, 搜索节点的网络传输率是比较稳定的, 假设为 g bit/s, 那么就有 $dsize = g \cdot t$

2) 故代币奖励算法可变式为:

$$reward = k\frac{e^{disze}}{ln(t+2)} \Rightarrow k\frac{e^{g\cdot t}}{ln(t+2)}$$

3) 对其求导, 可得:

$$\begin{aligned} f(t)' &= (k\frac{e^{g\cdot t}}{ln(t+2)})' \\ &= k\cdot\frac{e^{g\cdot t}\cdot g\cdot ln(t+2) - e^{g\cdot t}\cdot\dfrac{1}{t+2}}{(ln(t+2))^2} \\ &= \frac{k\cdot e^{g\cdot t}}{(ln(t+2))^2}(g\cdot ln(t+2) - \frac{1}{t+2}) \end{aligned}$$

4) 显然 $h(t) = g\cdot ln(t+2) - \dfrac{1}{t+2}$ 在定义域 $t>0$ 内是一个单调递增函数, 故有: $h(t) = g\cdot ln(t+2) - \dfrac{1}{t+2} > h(0) = g\cdot ln(2) - \dfrac{1}{2}$

而通常情况下, 网络传输率 g 都远大于 1, 故有:

$$h(t) > h(0) = g\cdot \ln(2) - \frac{1}{2} \gg ln(2) - \frac{1}{2} > 0$$

5) 那么, 由 $f(t)' = \dfrac{k\cdot e^{g\cdot t}}{(ln(t+2))^2}\cdot h(t) > 0$ 可知, 代币奖励算法: $reward = k\dfrac{e^{g\cdot t}}{ln(t+2)}$ 是一个随 t 单调递增的函数。显然, 这是符合现实情况的: 对于搜索节点, 网络传输率是比较稳定的, 基本保持为 g bit/s, 这种情况下随着检索的数据量大小 *dsize* 的增加, 检索用时 t 增加, 代币奖励 *reputa* 也随之增加。并且不同搜索节点的网络传输率 g 不同, 这也意味着性能更好的节点能够在提供更好的搜索服务的同时, 能够获取更多的代币奖励, 从而鼓励更多的搜索节点提供高性能的搜索服务。

(6) **可行性 6:** 在验证阶段, 利用检索阶段竞争失败的节点作为验证节点, 用以验证检索结果的正确性, 可以减少去中心化存储网络的算力的浪费; 其次, 给予验证节点一定的代币奖励, 可以更好地鼓励各节点积极参与去中心化存储网络的任务, 即使无法成为搜索节点, 也可以成为验证节点来获取代币奖励; 最后, 通过声誉值的占比来分配各正确验证节点的代币奖励, 鼓励更多节点进行正确的诚实行为, 以获取更多声誉值和代币奖励, 从而更好地为整个去中心化存储网络服务。

6.3 激励模型仿真

激励模型的仿真是通过以上文所提出的基于声誉的动态激励模型为蓝本, 构建若干节点并使用既定的奖惩机制与算法进行多轮次的模拟所得到。在具体的模拟方案中, 所涉及的实验节点的数量选取为 7 个, 一次实验进行 50 个任务轮次, 每一轮次均完整执行了检索、验证、奖惩多个阶段, 在保留当前

轮次的数据结果的基础上继续进行后续轮次, 直到所有轮次结束, 将本次模拟过程中各个任务轮次下的数据量变化(包括代币数量、声誉值)进行图像绘制以展示该激励模型的可行性。

如图 7 所示, 在各节点无消费行为的情况下, 大部分节点的代币变化曲线主要呈现出两种增长特征, 一是在竞争检索阶段失败并作为验证节点时, 获得小量代币奖励, 或在某轮次竞争胜出但验证未通过而导致呈现小幅度的缓慢增长或增长停滞; 二是作为胜出的检索节点, 在检索成功并验证通过后获得大量代币奖励, 因此呈现大幅度的跳跃式增长。除此之外, 也有极少部分节点的代币变化曲线因为节点的消极态度和不诚实行为, 而出现长期的增长停滞以及大幅度的减少。这几种特征均吻合本模型的代币奖惩机制, 总体上呈现出代币系统的稳定性, 能够证明本模型的代币奖惩机制较为合理。同时, 将节点 node0、node1、node2 的行为进行对比, 最初代币数量相近的三个节点因为其不同的行为, 而出现了明显的差异, 其结果也能很好的契合可行性 1、2、3、4 中的分析, 只有保持活跃, 并进行诚实的行为, 才会不断的获取代币奖励。这将充分保持去中心化存储网络的活跃性, 激励节点诚实并调动其积极性。

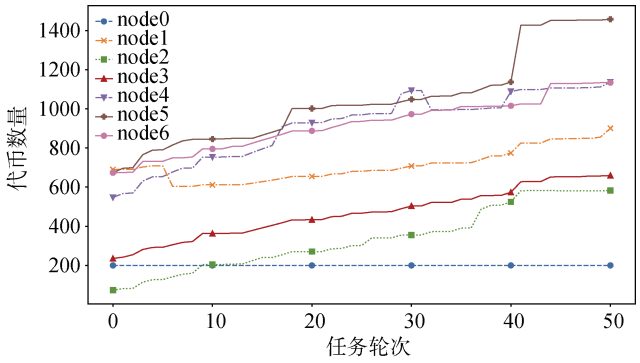

图 7 各节点代币变化情况(无消费)

Figure 7 Changes in coins at each node (no consumption)

如图 8 所示, 大多数节点的声誉值变化曲线都呈现出降速增长与部分小范围、不同幅度的下降以及当声誉值达到阈值后被重置的变化过程; 也存在极少数节点的声誉值变化曲线出现一定程度的增长停滞的情况。这也符合声誉值奖惩机制与算法特征, 契合可行性 1、2、3、4 中的分析, 同时对比节点 node0、node1、node2 的声誉值变化曲线, 对其行为进行分析, 可以知道: 只有保持活跃, 积极地进行诚实的行为, 才会让声誉值增长, 从而能够在以后的任务中获取更多的代币奖励。这也将进一步保持去中心化存储网络的活跃性, 激励节点诚实并调动其积极性。

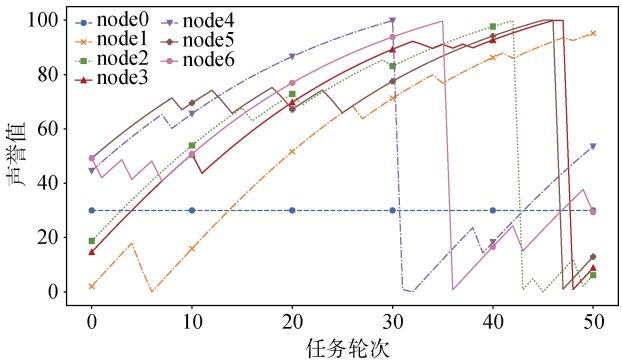

图 8 各节点声誉值变化情况

Figure 8 Changes in reputation at each node

7 系统方案分析

在本节中, 从表 4 所示的四个方面将本文提出的系统方案与近年来其他系统方案进行了对比, 包括倒排索引、激励策略、可搜索加密和去中心化存储。

表 4 方案对比

Table 4 The Comparison of Schemes

	高检索效率	支持动态激励	支持关键词搜索	去中心化
文献[1]	✓	×	×	✓
文献[3]	×	×	✓	✓
文献[20]	✓	×	✓	×
文献[30]	×	✓	×	✓
文献[31]	×	×	✓	×
文献[32]	×	×	✓	✓
本方案	✓	✓	✓	✓

文献[1]提出了一种用于跨组织的可靠存储和安全共享存储数据的系统, 并引入联盟区块链兼容的认证机制用于身份访问和数据验证, 但是没有考虑到数据持久化存储和系统节点可能存在的好奇和不诚实的问题; 文献[3]提出了一种基于 LSSS 访问策略的高效的可搜索算法, 一定程度上改善了计算和存储的效率, 但同样缺少对数据持久化存储和节点可能存在不诚实行为的考虑; 文献[20]提出了一种基于属性加密的多用户可搜索加密方案, 在云中保持数据的安全, 并允许具有适当授权的用户对加密数据进行检索操作, 但是其中心化的存储方式使其更容易遭受数据泄露的威胁; 文献[30]提出了一种新颖的激励兼容的 DSN 方案。其在去中心化的存储网络中引入了一种新颖的博弈论机制, 允许客户端挑战存储提供者, 在该方案中, 只有客户端提交挑战请求

时, 才会执行存储证明(PoS), 这将很好地让 PoS 验证免受拒绝服务攻击, 但该方案缺少对数据机密性的保护, 没有考虑节点好奇的问题; 文献[31]研究了在多用户环境下对加密数据的布尔查询的授权问题, 提出了一种支持布尔查询的动态多客户 SSE (DMSSE)方案, 允许数据所有者授权多个用户对加密数据库进行布尔查询, 但其中心化的存储方式也使得其同文献[20]一样, 更容易遭受数据泄露的威胁, 同时该系统方案缺乏激励策略来保证系统数据的持久化存储; 文献[32]研究了如何为去中心化存储服务中的加密文件带来安全可靠的内容搜索, 将可搜索加密技术应用于去中心化的系统, 并提出利用智能合约在区块链上记录加密搜索的日志, 通过设计一个公平的协议来处理争议和发出公平的付款, 但缺乏对检索高效性的考虑。

本文的方案利用了去中心化存储网络的高可靠性、可扩展性和低成本等优点, 节点存储采用属性基可搜索加密, 保证数据的安全性。为鼓励去中心化节点参与检索并返回正确结果, 在去中心化网络中引入基于声誉的动态激励模型作为激励层, 并利用倒排索引提高节点的检索效率, 更好地构建基于身份授权的可信去中心化存储网络。

8 结论

本文提出了一种基于身份授权的可信去中心化存储网络, 主要解决两个方面的问题: 首先, 去中心化节点具有好奇性, 希望查看存储在其上的数据; 针对这种情况, 本文结合属性基可搜索加密和 CP-ABE 对文件索引和对称密钥进行加密, 使去中心化存储网络能够在多所有者与多用户之间安全的共享数据。其次, 去中心化节点可能是不诚实的, 可能会向文件搜索者返回错误的内容, 本文在去中心化存储网络中引入基于声誉的动态激励模型来鼓励节点执行正确的检索操作, 从而使去中心化存储网络中的节点保持诚实。此外, 改进了属性基可搜索加密算法, 简化了公钥和主密钥的结构, 降低了 Setup 阶段时间复杂度和空间复杂度, 提高了去中心化存储网络中同步公钥和主密钥的效率, 同时在 Search 阶段的检索高效性弥补了网络中不同节点的检索能力差异。

参考文献

[1] Peng S L, Bao W X, Liu H, et al. A Peer-to-Peer File Storage and Sharing System Based on Consortium Blockchain[J]. *Future Generation Computer Systems*, 2023, 141: 197-204.

[2] Zahed Benisi N, Aminian M, Javadi B. Blockchain-Based Decentralized Storage Networks: A Survey[J]. *Journal of Network and Computer Applications*, 2020, 162: 102656.

[3] Wang H Y, Li Y, Susilo W, et al. A Fast and Flexible Attribute-Based Searchable Encryption Scheme Supporting Multi-Search Mechanism in Cloud Computing[J]. *Computer Standards & Interfaces*, 2022, 82: 103635.

[4] Mohd Kamal A A A, Iwamura K. Searchable Encryption Using Secret Sharing Scheme that Realizes Direct Search of Encrypted Documents and Disjunctive Search of Multiple Keywords[J]. *Journal of Information Security and Applications*, 2021, 59: 102824.

[5] Yin H, Zhang Y, Li F, et al. Attribute-Based Secure Keyword Search for Cloud Computing[M]. Cybersecurity and High-Performance Computing Environments. Chapman and Hall/CRC, 2022: 123-150.

[6] Zhang Y H, Deng R H, Xu S M, et al. Attribute-Based Encryption for Cloud Computing Access Control: A Survey[J]. *ACM Computing Surveys*, 2020, 53(4): 1-41.

[7] Yin H, Zhang J X, Xiong Y Q, et al. CP-ABSE: A Ciphertext-Policy Attribute-Based Searchable Encryption Scheme[J]. *IEEE Access*, 2019, 7: 5682-5694.

[8] Han R, Yan Z, Liang X Q, et al. How Can Incentive Mechanisms and Blockchain Benefit with each Other?A Survey[J]. *ACM Computing Surveys*, 2022, 55(7): 1-38.

[9] Huang H W, Lin J R, Zheng B C, et al. When Blockchain Meets Distributed File Systems: An Overview, Challenges, and Open Issues[J]. *IEEE Access*, 2020, 8: 50574-50586.

[10] Daniel E, Tschorsch F. IPFS and Friends: A Qualitative Comparison of Next Generation Peer-to-Peer Data Networks[J]. *IEEE Communications Surveys & Tutorials*, 2022, 24(1): 31-52.

[11] Kumar R, Tripathi R. Implementation of Distributed File Storage and Access Framework Using IPFS and Blockchain[C]. *2019 Fifth International Conference on Image Information Processing*, 2019: 246-251.

[12] Chen Y L, Li H, Li K J, et al. An Improved P2P File System Scheme Based on IPFS and Blockchain[C]. *2017 IEEE International Conference on Big Data*, 2017: 2652-2657.

[13] Xuan S C, Zheng L, Chung I, et al. An Incentive Mechanism for Data Sharing Based on Blockchain with Smart Contracts[J]. *Computers & Electrical Engineering*, 2020, 83: 106587.

[14] Kamboj P, Khare S, Pal S. User Authentication Using Blockchain Based Smart Contract in Role-Based Access Control[J]. *Peer-to-Peer Networking and Applications*, 2021, 14(5): 2961-2976.

[15] Michalas A, Weingarten N. HealthShare: Using Attribute-Based Encryption for Secure Data Sharing between Multiple Clouds[C]. *2017 IEEE 30th International Symposium on Computer-Based Medical Systems* (*CBMS*), 2017: 811-815.

[16] Vaanchig N, Xiong H, Chen W, et al. Achieving Collaborative Cloud Data Storage by Key-Escrow-Free Multi-Authority CP-ABE Scheme with Dual-Revocation[J]. *International Journal of Network Security*, 2018, 20(1): 95-109.

[17] Fan K, Liu T T, Zhang K, et al. A Secure and Efficient Outsourced Computation on Data Sharing Scheme for Privacy Computing[J].

Journal of Parallel and Distributed Computing, 2020, 135: 169-176.

[18] Curtmola R, Garay J, Kamara S, et al. Searchable Symmetric Encryption: Improved Definitions and Efficient Constructions[J]. *Journal of Computer Security*, 2011, 19(5): 895-934.

[19] Sahai A, Waters B. Fuzzy Identity-Based Encryption[C]. *Advances in Cryptology – EUROCRYPT 2005*, 2005: 457-473.

[20] Guo W F, Dong X L, Cao Z F, et al. Efficient Attribute-Based Searchable Encryption on Cloud Storage[J]. *Journal of Physics: Conference Series*, 2018, 1087: 052001.

[21] Ozyilmaz K R, Yurdakul A. Designing a Blockchain-Based IoT with Ethereum, Swarm, and LoRa: The Software Solution to Create High Availability with Minimal Security Risks[J]. *IEEE Consumer Electronics Magazine*, 2019, 8(2): 28-34.

[22] Vorick D, Champine L. Sia: Simple decentralized storage[J]. *Retrieved May*, 2014, 8: 2018.

[23] Kumaresan R, Bentov I, Kumaresan R, et al. How to Use Bitcoin to Incentivize Correct Computations[C]. *The 2014 ACM SIGSAC Conference on Computer and Communications Security*, 2014: 30-41.

[24] Li X Q, Jiang P, Chen T, et al. A Survey on the Security of Blockchain Systems[J]. *Future Generation Computer Systems*, 2020, 107: 841-853.

[25] Squarepants S. Bitcoin: A Peer-to-Peer Electronic Cash System[J]. *SSRN Electronic Journal*, 2008: 21260.

[26] Wang Jiawei. Research on the incentive mechanism of blockchain under the power Internet of things [D]. University of Electronic Science and Technology of China, 2022.

[27] Boneh D, Franklin M. Identity-Based Encryption from the Weil Pairing[C]. *Advances in Cryptology - CRYPTO 2001*, 2001: 213-229.

[28] Zheng Q J, Xu S H, Ateniese G. VABKS: Verifiable Attribute-Based Keyword Search over Outsourced Encrypted Data[C]. *IEEE INFOCOM 2014 - IEEE Conference on Computer Communications*, 2014: 522-530.

[29] Chaudhari P, Das M L. Privacy Preserving Searchable Encryption with Fine-Grained Access Control[J]. *IEEE Transactions on Cloud Computing*, 2021, 9(2): 753-762.

[30] Vakilinia I, Wang W H, Xin J J. An Incentive-Compatible Mechanism for Decentralized Storage Network[J]. *IEEE Transactions on Network Science and Engineering*, 2023, 10(4): 2294-2306.

[31] Du L L, Li K L, Liu Q, et al. Dynamic Multi-Client Searchable Symmetric Encryption with Support for Boolean Queries[J]. *Information Sciences*, 2020, 506: 234-257.

[32] Cai C J, Weng J, Yuan X L, et al. Enabling Reliable Keyword Search in Encrypted Decentralized Storage with Fairness[J]. *IEEE Transactions on Dependable and Secure Computing*, 2021, 18(1): 131-144.

张靖宇 现在北京邮电大学网络空间安全专业攻读学士学位。研究领域为网络空间安全。Email: jingyuzhang@bupt.edu.cn

刘奇 现在北京邮电大学网络空间安全专业攻读学士学位。研究领域为网络空间安全。Email: liuqi666@bupt.edu.cn

彭竑睿 现在北京邮电大学网络空间安全专业攻读学士学位。研究领域为网络空间安全。Email: penghongruif@bupt.edu.cn

杨增辉 现在北京邮电大学网络空间安全专业攻读硕士学位。研究领域为密码学应用。研究兴趣包括: 区块链、属性基加密、可搜索加密。Email: yangzh@bupt.edu.cn

袁开国 于 2009 年在北京邮电大学获得博士学位, 现在北京邮电大学网络空间安全学院讲师。研究领域为网络空间安全。研究兴趣包括: 可搜索加密、人工智能安全和物联网安全。Email: flyingdreaming@bupt.edu.cn

李小勇 于 2009 年在西安交通大学获得博士学位, 现在北京邮电大学网络空间安全学院教授。研究领域为网络空间安全。研究兴趣包括: 信任计算、内部威胁和物联网安全。Email: lixiaoyong@bupt.edu.cn