基于区块链的敏感数据可信存储系统实现

何　刚[1]，杨元瑾[2]
（1. 中电算力科技应用（宁夏）有限公司，宁夏　银川　750011；
2. 中电长城网际系统应用有限公司，北京　102209）

摘　要：针对敏感数据保密性和可靠性需求，提出了基于区块链的敏感数据可信存储技术，为敏感数据的安全存储和高效管理提供了一种创新性的解决方案。通过去中心化存储和不可篡改账本、区块链安全监管、链外数据可信接入、区块链共识机制等技术，永久记录网络或数据库的状态，有效防止入侵者数据窃取，实时监测数据库，保障数据的完整性。通过成功实现大幅提高并发容量的同时实现实时有效传输和处理数据，使系统具备适应复杂环境的通信能力，推动了敏感数据存储应用的发展，为相关领域的科研和实际应用提供有力支持。

关键词：区块链；可信存储；数据安全
中图分类号：TP393.0　　**文献标识码：**A　　**DOI：**10.19358/j.issn.2097-1788.2024.08.006
引用格式：何刚，杨元瑾. 基于区块链的敏感数据可信存储系统实现［J］. 网络安全与数据治理，2024，43(8)：35-39.

Implementation of trustworthy storage system for sensitive data based on blockchain

He Gang[1], Yang Yuanjin[2]
(1. China Electronics Computing Power (Ningxia) Co., Ltd., Yinchuan 750011, China;
2. China Electronics Cyberspace Great Wall Co., Ltd., Beijing 102209, China)

Abstract: Aiming at the confidentiality and reliability requirements of sensitive data, a trusted storage technology of sensitive data based on blockchain is proposed, which provides an innovative solution for the safe storage and efficient management of sensitive data. Through decentralized storage and tamper-proof account books, blockchain security supervision, trusted access of off-chain data, blockchain consensus mechanism and other technologies, the state of the network or database is permanently recorded, effectively preventing intruders from stealing data, and monitoring whether the database is tampered in real time. By successfully increasing the concurrent capacity, the system can effectively transmit and process data in real time, which makes the system have the communication ability to adapt to complex environment, promotes the development of sensitive data storage applications, and provides strong support for scientific research and practical applications in related fields.

Key words: block chain; trusted storage; data security

0　引言

存储是核心系统的数据底座载体，存储系统的安全可控是保证数据安全的重要前提和基础。由于存储技术本身就有层层的技术壁垒，因此在存储技术的可靠性基础上实现安全保密具有较高的技术门槛。

近年来，随着数据安全和隐私保护需求的不断提升，敏感数据的可信存储成为了一个重要的研究领域。传统的数据存储方式面临着数据篡改、丢失、泄露等问题，难以满足敏感数据的高安全性需求。区块链技术[1-2]的出现，为解决这些问题提供了一种新的思路。区块链具有去中心化、不可篡改、可追溯等特点，在数据存储安全和可信度方面具有天然的优势[3]。因此，基于区块链的敏感数据可信存储技术逐渐成为国内外研究的热点。

在具体实践中，美国国防部正在和一些跨国企业（如 IBM 和微软）积极探索区块链在数据安全和隐私保护中的应用[4-5]，计划创建一个安全可靠的信息获取和提供平台，以确保信息在任何地点的安全传输[6]。国内的研究则更多地关注于区块链技术在具体应用场景中的落地，如金融、医疗、政府和企业管理等领域。

本文针对敏感数据的可信存储需求，通过开展基于

www.pcachina.com

区块链的可信存储技术研究，突破区块链安全监管技术、链外数据可信接入技术、可信安全的分布式存储技术以及区块链共识机制，研制基于区块链的敏感数据可信存储验证平台，实现关键敏感数据的安全性、便利性和可信度，显著提升敏感数据防篡改和安全可靠能力。系统基于飞腾 CPU 体系架构，同时充分融合区块链技术，实现对核心数据的保护和加密。

1 需求分析

随着装备体系建设的跨越式发展，各单位都建立了各自的装备信息数据存储系统，由于数据分属于不同单位和人员，装备数据容易出现堆积、丢失、重复的现象[7]。传统的中心化数据库模式存在很多弊端，可靠性、可用性和可信性均得不到保障，导致敏感数据在存储和管理过程中面临巨大的安全风险。

区块链技术拥有去中心化、去信任化、可扩展和安全可靠等技术特点，有效利用区块链的技术特征，可以解决集中式存储系统的可靠性风险，通过区块链技术构建的可信存储系统，可以实现多方参与数据管理，不仅能保持数据的完整性，还能及时识别数据篡改的风险。同时，区块链技术具有的全系统对等的分散化部署特性，保证系统在任何节点损失的情况下仍能正常运行。其高度系统安全性和不可篡改的特性[8]，为解决当前敏感信息可信存储的多个难题提供了有效手段。

区块链技术结合加密存储技术可以在装备的全生命周期跟踪方面发挥重要作用[9]。通过建立分布式的记录系统，准确记录每一次装备的维护、更新、使用和修复信息，确保装备系统运行的可靠性和持久性。这种高度透明和可追溯的管理系统，不仅提升了装备的操作效能，还有助于及时发现和处理潜在的安全隐患，保障作战安全。

2 总体技术方案

结合加密存储和区块链在数据存储安全和数据完整性方面的优势，系统主要包含两层：数据存储层和区块链层，如图 1 所示。

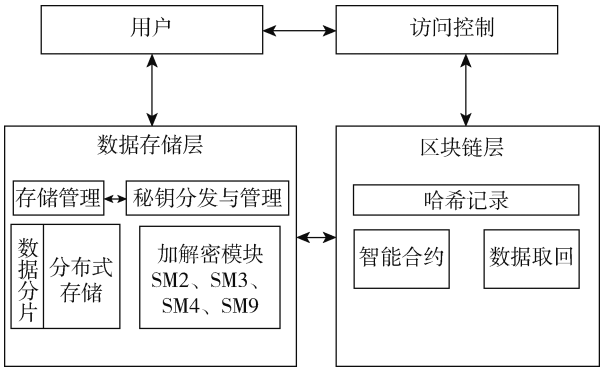

图 1 总体架构图

2.1 数据存储层

数据存储层是整个系统架构中负责数据存储、安全性和高效访问的关键部分。它主要包括数据加密、数据分片和分布式存储三个核心组件，每个组件都在确保数据的安全性和可用性方面发挥重要作用。

（1）数据加密

当用户提交数据时，首先通过加密算法对数据进行加密处理。加密的目的是确保数据在存储和传输过程中始终保持机密，防止未经授权的访问和泄露。通过加密技术，如非对称公私钥数据加密技术和时间戳等，覆盖底层数据区块的链式结构。系统兼容国际标准如 AES、ECDSA、SHA2[10]，同时也考虑了国内机构采用的国密算法，使用 SM2、SM3、SM4、SM9 加密技术进行加密[11]。数据在存储前进行加密处理，使用强大的加密算法如 SM9，保障数据即使被截获也无法解密和读取。

（2）数据分片

加密后的数据被分割成若干小片段。这种数据分片技术不仅提高了存储和处理数据的效率，还增加了数据的安全性。即使某个片段被攻破，由于其仅为整个数据的一部分，攻击者无法获取完整信息。每个片段在存储前进行进一步加密处理，并生成唯一的哈希值，用于后续的数据完整性验证。

（3）分布式存储

加密且分片后的数据片段存储在分布式存储系统中。这意味着数据被分散存储在多个节点上，而不是集中在单一的存储介质上，从而提高了系统的容错能力和数据的可用性。分布式存储系统不仅能有效防止单点故障，还能通过多副本存储技术确保数据的高可用性和容错能力。每个数据片段在多个节点上存储多个副本，这样即使某些节点发生故障，数据仍能被恢复。

（4）数据写入和读取流程

数据写入流程包括：用户提交数据，加密数据，分片，加密片段生成哈希值，并将哈希值记录在区块链上。数据读取流程则包括：用户请求数据，从区块链上获取哈希值，从分布式存储系统中读取相应的数据片段，重组和解密数据，并验证数据的完整性。这确保了数据在整个生命周期内的安全性、完整性和可用性。

通过以上步骤，数据存储层确保了数据在存储和传输过程中的安全性和可靠性，同时通过分布式存储和数据分片技术，提高了数据的访问效率和系统的整体性能。

2.2 区块链层

区块链层是整个系统架构中确保数据完整性、安全性和可追溯性的关键部分。它主要包括哈希记录和智能合约两个核心组件，通过这些组件实现数据操作的透明

www.pcachina.com

化和自动化管理。

（1）哈希记录

数据经数据存储层处理并分片后，生成的每个数据片段都会有一个唯一的哈希值。哈希值是一种数字指纹，用于唯一标识和验证数据的完整性。区块链层负责将这些哈希值记录在区块链上。区块链作为一种分布式账本技术，其不可篡改和分布式存储的特性确保了数据哈希值的安全存储和验证。记录在区块链上的哈希值为数据提供了一个可靠的验证机制，确保数据在存储和传输过程中不被篡改。

（2）智能合约

智能合约是运行在区块链上的自执行代码，负责自动管理和执行与数据操作相关的逻辑。它们通过预先定义的规则来管理数据的访问控制、验证和审计功能。例如，智能合约可以规定哪些用户有权限访问特定数据片段，或者自动验证请求数据的哈希值是否匹配，从而确保数据的完整性和安全性。智能合约的自动化执行特性减少了人工干预的风险，增强了系统的安全性和可靠性。

（3）数据操作流程

当用户提交数据时，经过数据存储层的加密和分片处理后，生成的哈希值被记录在区块链上。智能合约自动管理这些哈希值的存储和验证，确保每个数据片段都有对应的哈希值记录。当用户请求读取数据时，系统首先从区块链上获取相应数据片段的哈希值。然后，数据存储层从分布式存储系统中读取加密的各个数据片段并进行重组。解密后的数据会重新计算哈希值，并与区块链上的哈希值进行比对，确保数据未被篡改。

区块链层通过其不可篡改的特性确保了数据的透明性和安全性。所有的数据操作记录（如数据写入、读取和修改）都通过智能合约记录在区块链上，形成不可篡改的操作日志。这些日志提供了详细的审计轨迹，使得系统可以随时追溯和审计数据操作，识别潜在的安全威胁。

3　系统功能与关键技术

系统的核心功能包括：

（1）敏感数据可信存储平台

通过区块链技术实现分布式、不可篡改、可追溯的敏感数据存储，解决数据堆积、丢失和重复等问题。区块链的去中心化特性提供了比传统数据库更高的可靠性、可用性和可信性。

（2）应用区块链共识机制

应用高吞吐量、可扩展的共识算法（如 PoW、PoS 或 PBFT），采用分片技术以提高系统性能，实现同步处理不同交易，从而提高整体网络吞吐量，克服了传统共识算法的一些限制。

系统的关键技术包括：

（1）区块链共识机制

区块链共识机制整合了网络节点的多种共识算法，是区块链技术的核心部分。这些算法决定了记账节点的选取方式[12]，直接影响系统的安全性和可靠性。针对敏感数据存储要求，有效利用现有的共识算法，包括：工作量证明（PoW），适用于需要高安全性但不太在意能源消耗的场景；权益证明（PoS），适用于需要节能且有一定安全性的场景；授权股权证明（DPoS），适用于需要高效率和快速共识的场景；拜占庭容错（BFT）和实用拜占庭容错（PBFT），适用于需要高容错性和快速达成共识的场景；PoS 与 BFT 结合的算法（PoS + BFT），适用于需要高安全性和高效率的场景；零知识证明（ZKP），适用于需要高隐私保护的数据存储和验证；以及骨干权益证明（Ouroboros PoS），适用于公链和需要高安全性的场景。通过结合这些算法的特点，根据具体的装备信息管理需求，选择最适合的共识算法，并不断探索和改进新的算法，以确保敏感数据存储的安全性和可靠性。

（2）区块链安全监管技术

基于区块链的敏感数据可信存储验证平台上存储了装备交易与供应链的全流程信息，使得监管机构能够依据相关法规设定智能监管合约[13]，对信息系统上的数据流进行实时监控，自动验证交易和用户的合规性，实现装备供应链的智能监管。智能监管能帮助监管部门实时监管装备的生产、运输安全、交付服役到退役报废等问题，其监管重心也从传统的事后追溯逐渐转向事前预警和事中控制，将各方损失降到最低。由于验证平台具有多方共识、节点共同维护等特点，平台使用方能够通过平台实现自我监管并与各节点用户进行信息交互，最终实现共同监管。

（3）链外数据可信接入技术

区块链主要采用非对称加密算法来构建可信接入，实现节点间信任、数据所有权验证和访问权限控制，并结合其他加密算法来实现数据的真实性、私密性和完整性保护。非对称加密是指使用公钥和私钥对存储和传输的数据进行加密和解密、签名（对数据摘要加密）和验签（验证摘要一致性）。公钥可公开发布，用于发送方加密要发送的信息，私钥用于接收方解密接收到的加密内容。常见的非对称加密算法有 RSA 和 ECC（即椭圆曲线加密算法，如国密 SM2、SM9 算法），可采用公钥基础设施（Public Key Infrastructure，PKI）和基于标识的密码技术（Identity-Based Cryptograph，IBC）两种公钥密码算法技术体制。

www.pcachina.com

（4）可信安全的分布式存储技术

基于区块链的敏感数据可信存储采用分布式数据存储模式[14]，每个网络节点采用互为备份方式存储数据。在每次存储新数据时，均是以之前已存储的加密数据作为基础，结合新数据进行再次加密。通过以上一系列方式，使得存储的数据无法被篡改，所记录的所有数据及数据变化均可被追踪和查询。

这些功能和关键技术使得系统能够有效解决敏感数据管理中的安全性、隐私性和可信性等问题，为各单位提供安全可靠的解决方案。

4　成果验证

（1）实验环境

本实验在一个包含 200 个高配置节点的区块链网络上进行。所有敏感数据通过区块保存在可信存储环境中。在这 200 个共识节点上，使用 PBFT 或 PoS 共识算法，并采用国密加密算法（如 SM2、SM3、SM4、SM9），实现数据的加密存储和共识验证。实验平台可以选择使用 Hyperledger Fabric 或以太坊平台进行搭建。系统部署如图 2 所示。

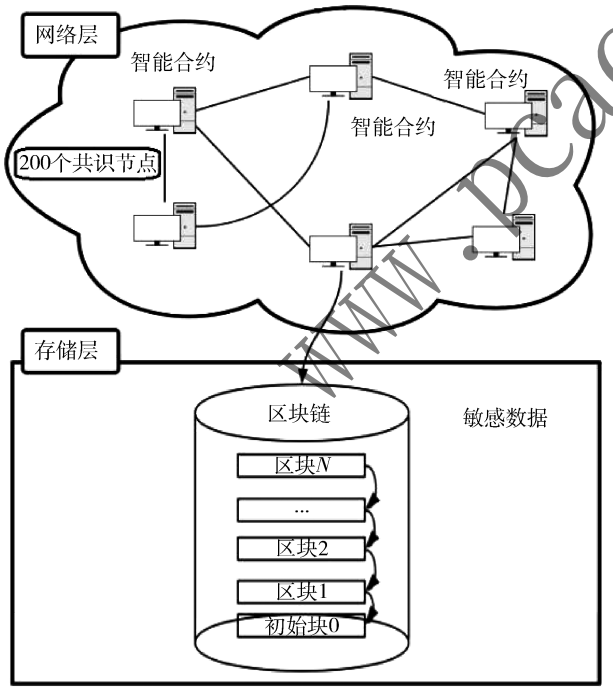

图 2　测试部署图

（2）实验数据

实验数据包括金融交易记录、医疗记录和个人身份信息，每笔记录的大小为 1 KB ~ 10 KB，总数据量为 1 000 000笔。数据种类的多样性和数据量的规模性，旨在模拟真实应用环境下的敏感数据存储需求。

（3）实验步骤

数据写入：首先将数据加密，生成哈希值，并通过分布式存储系统进行存储。目标是达到每秒 6 000 笔交易的写入吞吐量。

数据读取：从区块链上获取哈希值，并从分布式存储系统中读取相应的数据片段，进行重组和解密。目标是验证系统的读取吞吐量和数据完整性。

篡改检测：通过对某个节点的数据进行篡改，验证平台是否能够检测并拒绝错误信息，确保区块链上数据的正确性。

（4）实验结果

实验结果表明，基于区块链的敏感数据可信存储系统在高负载下表现良好，达到了设计的性能指标。在数据写入实验中，系统成功实现了每秒 6 000 笔交易的写入吞吐量。在数据读取实验中，系统能够快速准确地读取和验证数据。在篡改检测实验中，系统成功检测并拒绝了篡改的数据，验证了区块链技术在数据完整性和安全性方面的优势。详细的测试项目以及测试结果如表 1 所示。

表 1　测试系统实验结果

测试项目	测试方式及测试结果
功能测试	用例执行覆盖率 100%，测试通过率 95%
安全测试	用例执行率 100%，通过率 100%
UI 测试	用例执行率 100%，通过率 100%
模拟测试	用例执行率 100%，通过率 100%
稳定性测试	6 套设备稳定运行 7 × 24 小时未出现死机或无响应的问题，测试通过
高可用测试	用例执行率 100%，通过率 100%
性能测试	达到 45 万 IOPS

在 FC（Final Test）测试中，对各个用例的 IOPS（Input/Output Per Second）、带宽、平均响应时间等性能指标进行测试，结果如表 2 所示。

表 2　FC 测试结果

测试粒度及用例	IOPS	带宽/（MB/s）	平均响应时间/ms
512 B 顺序读	445 266. 6	217. 42	2. 298
512 B 顺序写	450 339. 6	219. 9	2. 269
512 B 随机读	375 471	183. 34	2. 726
512 B 随机写	29 836	14. 56	34. 311
1 MB 顺序读	6 149. 5	6 149. 5	166. 564
1 MB 顺序写	3 079. 24	3 079. 24	332. 222
1 MB 随机读	5 161. 3	5 161. 3	198. 397
1 MB 随机写	2 831. 72	2 831. 72	361. 188

www.pcachina.com

5 结论

通过采用基于区块链的敏感数据可信存储技术，本研究在提升敏感数据的保密性和可靠性方面取得了显著成果。在大幅度提高并发容量的同时，成功实现了敏感数据的实时有效传输和处理，使得系统具备了适应复杂环境的通信能力。

建立基于区块链的可信存储技术为最终用户的研发提供了坚实的区块链技术基础，推动了敏感数据存储应用的发展。这不仅在提高敏感数据处理效率方面具有重要意义，同时，也为最终用户的研发提供了区块链技术储备，对推进我国装备信息化的快速发展具有重要的战略意义。

综合而言，本研究为敏感数据的安全存储和高效管理提供了一种创新性的解决方案，充分发挥了区块链技术的优势。未来，可进一步深化研究，不断优化系统性能，以满足不断发展变化的敏感数据管理需求。这将有助于推动区块链技术在敏感数据领域的广泛应用，并为相关领域的科研和实际应用提供有力支持。

参考文献

[1] 刘明达，陈左宁，拾以娟，等. 区块链在数据安全领域的研究进展［J］. 计算机学报，2021，44（1）：1－27.

[2] 曾诗钦，霍如，黄韬，等. 区块链技术研究综述：原理、进展与应用［J］. 通信学报，2020，41（1）：134－151.

[3] 张亮，刘百祥，张如意，等. 区块链技术综述［J］. 计算机工程，2019，45（5）：1－12.

[4] 孙煜飞，杨强，杨朝晖. 区块链军事应用探析［J］. 指挥控制与仿真，2021，43（4）：76－80.

[5] 廉蔺，朱启超，赵炤. 区块链技术及其潜在的军事价值［J］. 国防科技，2016，37（2）：30－34.

[6] 张志威，王国仁，徐建良，等. 区块链的数据管理技术综述［J］. 软件学报，2020，31（9）：2903－2925.

[7] 赵刚，刘涛，李世兴，等. 基于区块链的有无人协同系统可信存储技术［J］. 火力与指挥控制，2021，46（6）：141－144.

[8] 辛伟，刘悦. 部队装备管理信息化建设的系统分析［J］. 中国军转民，2021（14）：55－56.

[9] 赵东波，岳凡. 陆军智能化无人化作战体系构建［J］. 国防科技，2019，42（5）：51－54.

[10] 张诗永，陈恭亮，范磊，等. G－AES 算法［J］. 密码学报，2014，1（2）：187－199.

[11] 刘敖迪，杜学绘，王娜，等. 区块链技术及其在信息安全领域的研究进展［J］. 软件学报，2018，29（7）：2092－2115.

[12] 韩璇，刘亚敏. 区块链技术中的共识机制研究［J］. 信息网络安全，2017（9）：147－152.

[13] 洪学海，汪洋，廖方宇. 区块链安全监管技术研究综述［J］. 中国科学基金，2020，34（1）：18－24.

[14] KEERTI V，SREEJA N. Hybrid encryption based SHA2-256 integration techniques for high security for data stored in cloud environment［J］. International Journal of Computer Applications，2017，168（1）：24－28.

（收稿日期：2024－05－11）

作者简介：

何刚（1976－），男，本科，工程师，主要研究方向：信息安全、网络安全。

杨元瑾（1986－），男，硕士，高级工程师，主要研究方向：人工智能、信息安全、网络安全。

www.pcachina.com

版权声明

凡《网络安全与数据治理》录用的文章，如作者没有关于汇编权、翻译权、印刷权及电子版的复制权、信息网络传播权与发行权等版权的特殊声明，即视作该文章署名作者同意将该文章的汇编权、翻译权、印刷权及电子版的复制权、信息网络传播权与发行权授予本刊，本刊有权授权本刊合作数据库、合作媒体等合作伙伴使用。同时，本刊支付的稿酬已包含上述使用的费用，特此声明。

《网络安全与数据治理》编辑部

www.pcachina.com