Computer Science and Application 计算机科学与应用, 2022, 12(5), 1292-1300
Published Online May 2022 in Hans. http://www.hanspub.org/journal/csa
https://doi.org/10.12677/csa.2022.125129

Hans汉斯

基于IPFS区块链技术的工业互联网数据可信存储系统

陈泓达，冯云霞，牛云鹤

青岛科技大学信息科学技术学院，山东 青岛

收稿日期：2022年4月13日；录用日期：2022年5月12日；发布日期：2022年5月19日

摘 要

为解决工业互联网设备运行数据的可信存储问题，本文提出了基于IPFS的区块链数据可信存储系统。利用区块链技术的去中心化、可追溯与不可篡改等特性保证存储数据的真实可信。将源文件存入IPFS系统中，只将文件哈希摘要上链存储，在发挥了区块链优势的同时缓解了链上的数据存储压力。为进一步限制对数据的使用权限，系统内加入无证书密码体制对数据哈希摘要进行加解密操作。实验结果表明，本方案能实现工业互联网数据可信存储，且具有良好的系统性能。

关键词

区块链，工业互联网，IPFS，数据存储，无证书密码体制

Industrial Internet Data Trusted Storage System Based on IPFS Blockchain Technology

Hongda Chen, Yunxia Feng, Yunhe Niu

School of Information Science & Technology, Qingdao University of Science and Technology, Qingdao Shandong

Received: Apr. 13th, 2022; accepted: May 12th, 2022; published: May 19th, 2022

Abstract

In order to solve the problem of trusted storage of industrial Internet equipment operating data, this paper proposes a trusted storage system for blockchain data based on IPFS. The decentraliza-

文章引用: 陈泓达, 冯云霞, 牛云鹤. 基于 IPFS 区块链技术的工业互联网数据可信存储系统[J]. 计算机科学与应用, 2022, 12(5): 1292-1300. DOI: 10.12677/csa.2022.125129

tion, traceability and immutability of blockchain technology are used to ensure the authenticity and credibility of stored data. The source file is stored in the IPFS system, and only the hash summary of the file is stored on the chain, which relieves the data storage pressure on the chain while exerting the advantages of the blockchain. In order to further restrict the use authority of the data, a certificateless cryptographic system is added to the system to encrypt and decrypt the data hash digest. The experimental results show that this scheme can realize the trusted storage of industrial Internet data, and has good system performance.

Keywords

Blockchain, Industrial Internet, IPFS, Data Storage, Certificateless Cryptosystem

1. 引言

随着第四次工业革命的快速发展，工业互联网将大力推动工业制造向信息化、智能化发展。工业互联网平台是基于现代制造业的海量数据进行采集、汇聚、分析的服务体系，支撑生产资源实时连接、供给和配置的平台。其本质是在传统云计算平台的基础上加入物联网、大数据、人工智能等新兴技术，利用高效的数据采集体系与面向工业大数据的管理开发环境，实现工业相关经验、技术、知识的模型化、复用化，不断提高设计研发、生产制造等资源配置效率，最终形成资源富集、合作共赢的制造业生态[1]。

工业互联网中的数据具有种类众多、数据量庞大的特点[2]。企业生产数据是工业互联网数据的重要组成部分之一，正确高效地存储与使用企业生产数据是企业生产过程优化与效率提升的前提[3]。但是，在现有的集中式数据存储与管理模式中，已无法实现对企业生产数据实现安全、稳定、可靠的存储要求。在这种存储模式中，一旦服务器在外部攻击下崩溃，必然导致服务器中存储数据的丢失或被破坏，此类数据将不再具有任何使用价值，之前为存储这些数据所使用的资源也被浪费。数据一旦出现被窃取、篡改、删除，甚至虚假捏造的情况，将势必对企业的安全生产带来严重威胁与隐患。

区块链技术的出现为工业数据可信存储提供了可行的解决方案。区块链技术本身具有去中心化、点对点传输、不可篡改、可追溯、分布式存储等技术优点，经过多方共同维护可以实现在不需要第三方信任机构的情况下，将数据存储达到去中心化、可信、难以篡改的目的[4] [5]。所以，通过区块链技术对数据进行可信存储，同时对数据添加严格的访问权限条件，足以充分保证所保存数据的真实可信。但是，区块链技术目前因自身区块结构限制，导致系统效率与吞吐量不足，很难满足在数据规模庞大的场景中使用。

因此，本文针对工业互联网中工业数据如何实现可信存储的问题，提出了一个基于 IPFS 优化的联盟链解决方案。以联盟链自带的用户准入机制，以及安全、可溯源、不可篡改的特性，结合 IPFS 星际文件系统，添加基于无证书密码体制对数据哈希摘要进行加解密操作，共同构建一个工业数据可信存储系统，以实现在工业互联网场景下工业数据的安全可信存储。

2. 相关工作

学术界研究人员在运用 IPFS、区块链技术在分布式存储、数据保护及共享等领域积极探索并取得了

一定近展。周方明等人提出将利用物联网设备自动采集的焊接相关数据存入 IPFS 分布式存储系统中，大幅降低了原有数据存储成本[6]；高文涛等人基于 IPFS 分布式存储系统与区块链技术实现了去中心化的音乐数据共享平台[7]；范贤丽使用 IPFS 系统与区块链相结合，对粮食供应链隐私信息进行存储，保护了重要数据的安全不可篡改[8]；朱彦霞等人采用 IPFS 系统与区块链技术融合进行融媒体数据安全存储，降低了存储成本并节省了网络带宽[9]。Zheng 等人提出一种基于 IPFS 的区块链存储模型，以拓展比特币存储空间，旷工将交易数据存入 IPFS，并将返回的 IPFS 哈希值写入区块链[10]。Kumar 等人提出一种基于区块链和 IPFS 的图像视频版权保护模型，使用 IPFS 存储图像和视频元数据，区块链仅存储图像指纹[11]。谭海波等人提出一种基于区块链的档案数据保护与共享方法，结合 IPFS 存储档案数据，实现了数字档案的保护、验证、恢复与共享功能，该方法可拓展区块链数据存储能力、减轻链上数据高频访问压力[12]。许丽等人以区块链技术在云存储平台的应用研究现状作为切入点，配合 IPFS 网络构建更具稳定性和安全性的云储存平台模型[13]。但目前并没有使用 IPFS 结合区块链技术解决工业互联网中工业数据的安全、可信存储方面的研究。

3. 关键技术

3.1. 联盟链

联盟链，是区块链发展至今的核心产品，链内可存在多个组织机构进行共同维护[14]。链内节点可分群组管理，不同群组由不同组织机构负责，实现不同功能与应用，多个群组在链中共同完成记账功能。通过对应设置，群组间可将特定信息在链内进行分享。如今，应用最多也是最广泛的就是联盟链，相较私链联盟链具有更多操作空间，相较公链联盟链具有更好的安全机制，联盟链因此具有最大的商业用价值。

联盟链具备多个技术优势：1) 去中心化。在联盟链内，同时存在多组织机构，多方共同管理维护整个联盟链，数据由不同组织机构分别管理，在内部实现共享，易实现共识；2) 可管理性强。公链中节点非常多，在区块发布时便会立即被全部节点共识，并接着这个区块继续出块打包记录交易，区块链链状结构一直延续下去。由于出块快且多，可以说只要上链便再无更改可能，但在联盟链当中，因为节点数量有限，交易数量有限，如果发布的交易有更改需求，只要大多数节点同意，即可再发布一次交易并在交易中添加说明，让之前上传的交易作废，重新记录在区块链上，这样间接完成了对已上链数据更改的要求，能满足部分特定场景的业务需求；3) 有节点准入权限控制。链内区块中的数据只有被认证的节点才能进行调用与获取，没有相关认证的节点是没有相关访问权限的；4) 交易速度快。由于联盟链具备严格的节点准入机制，所以联盟链中存在的节点数量往往比较有限，因为节点数量少所以在有交易时可实现快速同步并实现共识。

3.2. IPFS

IPFS (InterPlanetary File System)，星际文件传输系统，是一个利用点对点传输建立的分布式超媒体分发协议。IPFS 的设计中集成了 DHT、BitTorrent、自认证文件系统 SFS 和 Git 的优点，IPFS 被认为是最有可能取代 HTTP 的新一代互联网协议，它提供了永久的去中心化存储文件的方法[15]。IPFS 基于内容寻址，将信息保存到 IPFS 节点中，IPFS 系统将会返回基于该信息计算得出的唯一哈希值。哈希值与信息内容一一对应，即使只对信息做轻微修改，也会得到完全不同的哈希值。当 IPFS 被请求一个文件哈希时，它会使用一个分布式哈希表找到文件所在的节点，取回文件并验证文件数据。

IPFS 可存入文件格式不受限制，存储空间也不受限制，具有通用性。若在系统中存入很大的文件，该文件会被分割成块，获取该文件时只需要从多个节点一起下载即可。IPFS 是具有灵活性、细粒度、分布式的网络，能够胜任在网络中内容分发的任务。IPFS 架构方便用户可以在上面做任何文件的分享与传播。

3.3. 智能合约

智能合约是区块链的关键组成部分。智能合约在区块链中可以比作一张电子合同，该电子合同已经提前写好了验证条件与执行条件，一旦触发要求就将自动按合同中写明的开始事务，过程中不可中断且自动执行[16] [17]。该合同一旦部署于系统当中就不能在对其中的相关规定细则进行更改与删除，只要启动开始执行，不需要第三方的辅助，双方即可完成真实可信的交易，执行结果也不可逆转。智能合约完美契合了区块链的去中心化思想，是区块链在具体应用时重要技术手段，为区块链在金融、医疗、物联网领域的应用提供了重大的辅助作用。

3.4. 无证书密码体制

无证书公钥密码体制由 Al-Riyami 和 Paterson 提出[18]。在无证书公钥密码体制中，密钥生成中心通过与用户身份相绑定，为用户生成公私钥对，不存在传统公钥证书机制与秘钥托管问题，且密钥生成中心不能获知任何用户的私钥[19] [20]。鉴于以上优势，无证书公钥密码体制从刚刚提出便得到了相关业内人士的关注与肯定，发展至今也诞生了无证书公钥密码体制、无证书签名方法、无证书代理重加密方案等相关学术成就[21]。

4. 系统设计

在本文提出的工业数据可信存储的系统中，通过把需要存储的源文件存入 IPFS 系统中，将系统返回的文件哈希摘要存在区块链上，在最大限度减轻链上存储压力的情况下实现基于区块链的数据可信存储。其次，通过无证书密码体制对文件哈希摘要上链存储前进行加密，只有特定用户可通过自身私钥对哈希密文进行解密，只有使用解密后的哈希明文才可在 IPFS 系统中下载到源文件，以此保护数据的使用权限。

4.1. 系统架构

本文设计的系统架构如图 1 所示，包含应用层、中间层和存储层。

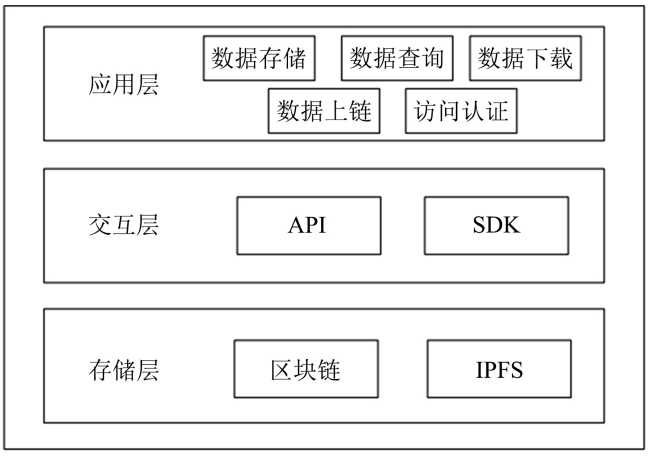

Figure 1. System structure
图 1. 系统架构

1) 应用层：应用层以接口形式向用户提供完整的系统服务。应用层可实现的功能包括：数据存储、数据上链、数据查询、数据下载、访问控制。

- 数据存储：由管理人员将企业内设备运行数据进行整理，数据实时汇总后存入已部署的 IPFS 系统节

点内，IPFS 系统将数据进行分片存储，返回给用户根据数据内容生成的唯一哈希值结果，管理员对所得到的哈希值进行加密，将加密后得到哈希密文上传至区块链系统。IPFS 分布式存储系统承担了数据存储压力，弥补了区块链系统中链上存储空间有限、事务处理效率低下的不足，二者深度融合既保留了区块链自身优势也提高了数据存储效率。

- 数据上链：将被加密的哈希值上传至区块链账本。由管理员发起数据上链智能合约，将工业数据集哈希值写入交易并将该笔交易上链。该笔交易上链后会生成一个唯一标识符 Number，只有在链内有身份授权的用户才可以通过唯一标识符 Number 在链上找到并获取这笔交易信息。
- 链上查询：系统中所部署使用的区块链类型为本身具有准入机制的联盟链，只有具有链上权限的用户才可以调用交易查询智能合约，通过在合约中输入交易唯一标识符 Number，在链上找到对应的交易获得哈希摘要密文。
- 访问控制：通过无证书加密机制对数据哈希值进行加密与解密，充分保证用户对数据使用权限控制，保证数据的隐私性与安全性，避免数据的滥用。假设管理员 A 希望自己上传的数据只能被特定的人群所使用。A 首先基于无证书加密体系生成自己的公钥私钥，在获取数据哈希上传至区块链系统之前，将哈希摘要利用相应加密函数进行加密，即使用具有查看权限的人的公钥进行加密，之后再上传至区块链账本。假设 B 具有 A 的许可且身份认证在与 A 同一条联盟链中，B 通过智能合约，利用唯一标识 Number 找到对应交易并获取交易内记录的哈希摘要密文，之后使用自己的私钥对获得的被加密的哈希摘要进行解密，获取哈希摘要原文，通过哈希摘要原文可以在 IPFS 系统中下载到 A 所上传的源数据文件。系统业务逻辑如图 2 所示。

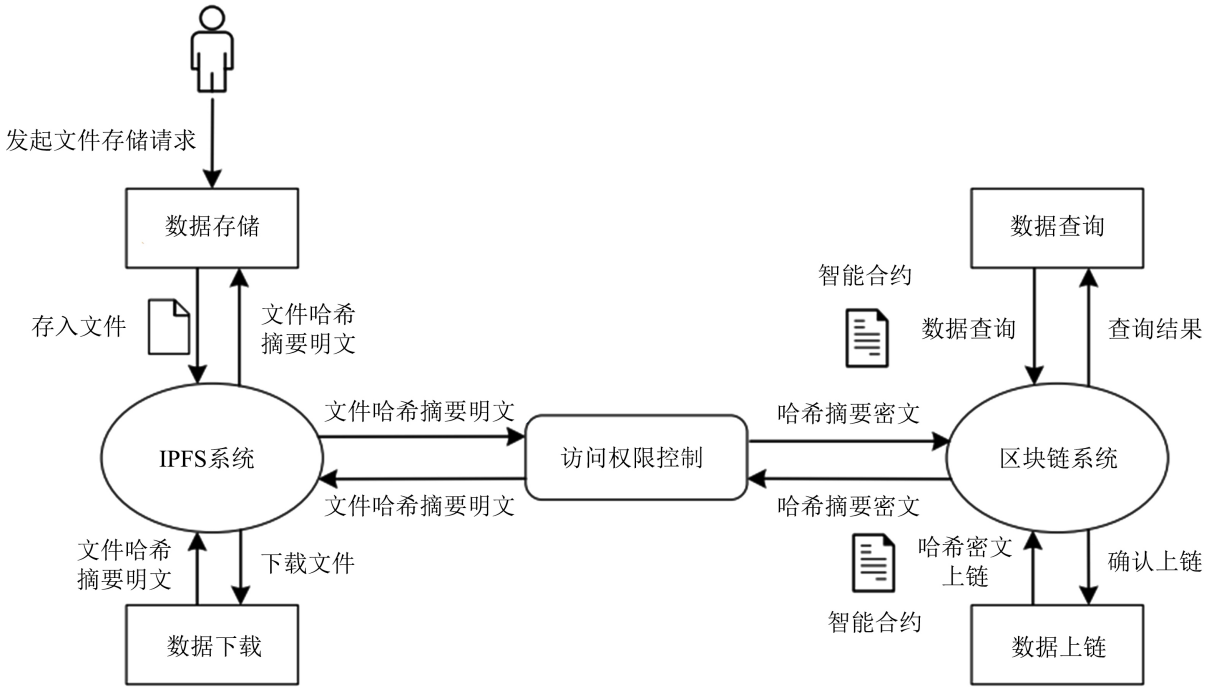

Figure 2. System business logic
图 2. 系统业务逻辑

2) 中间层：中间层作为过渡层，连接上层应用层与下层存储层。主要负责相关功能接口的实现与底层加密方式的调用。

3) 存储层：存储层包含部署的 IPFS 星际文件系统及联盟链平台。IPFS 星际文件系统负责两个功能：a) 完整数据的存储并返回给用户数据哈希摘要；b) 根据哈希摘要对存储的数据进行下载。联盟链平台主要负责用户的权限管理、调用智能合约将交易上链共识以及后续交易查询功能。

4.2. 数据访问控制模块设计

在基于 IPFS 的联盟链工业数据可信存储中，主要使用对关键数据加密解密的方式来实现用户对数据的访问权限控制。在企业管理员获得 IPFS 系统返回的数据哈希值之后，首先利用无证书密码体制中的公私钥对生成方法建立与自己身份相关的公钥和私钥，将数据哈希值通过特定公钥完成加密处理生成哈希密文。在具有权限的用户在链上获取交易信息后，可以使用自己的私钥对哈希密文进行解密，使用解密后的哈希明文即可在 IPFS 系统中下载指定文件。

密钥生成中心 KGC 通过将秘钥与用户身份绑定，为用户产生公钥和私钥。秘钥产生过程如下：第一步，随机产生保密数 λ，KGC 使用保密数 λ 作为函数参数生成系统参数 K、密钥 SK。第二步，KGC 调用 PSKGen 方法，传入参数 K、用户 ID 和 SK 产生不完整私钥 PSK，此时把不完整私钥 PSK 告知用户 A。用户 A 使用 K 和个人身份 ID，利用 USKGen (K, ID)方法产生第二个不完整私钥 X。第三步，用户 A 调用参数 K、PSK 和 X，使用 SKGen (K, X, PSK)方法产生最终用户私钥。第四步，用户 A 执行 PKGen (K, X)方法，通过传入参数 K、X 产生完整用户公钥。详细算法步骤如下：

1) 密钥生成中心 KGC 随机产生保密数 λ，传入 λ 作为函数参数，执行函数获得系统参数 K 和密钥 SK。

2) KGC 调用 PSKGen (K, ID, SK)方法产生不完整用户私钥 PSK，将 PSK 告知用户。

3) 用户调用 USKGen (K, ID)方法，通过传入系统参数 K 和自身 ID，产生第二个不完整私钥 X。

4) 用户使用 KGC 产生的第一部分不完整私钥 PSK 和自身产生的第二部分不完整私钥 X，以及系统参数 K，使用 SKGen (K, X, PSK)方法合成用户完整私钥。

5) 用户调用系统参数 K 和自身产生的不完整私钥 X，使用 PKGen (K, X)方法，产生用户最终完整公钥 PK。

在权限验证设计中，包括用户签名和认证流程。详细过程介绍如下：

1) 用户 A 通过调用系统参数 K 和用户完整私钥对哈希摘要密文进行签名，签名后返回签名结果。

2) 具有查看权限用户通过传入签名用户 A 的身份 ID、公钥 PK 和哈希摘要密文 C_Hash 参数，来验证签名，对签名有效性进行确认。

在哈希摘要明文加密过程中，详细过程介绍如下：

用户通过系统参数 K、用户身份 ID 和用户公钥 PK，执行 Encrypt (K, PK, ID, Hash)方法把文件哈希摘要明文加密为文件哈希摘要密文 C_Hash，并将加密结果返回。

在哈希摘要密文解密过程中，详细过程介绍如下：

具有权限的用户在区块链系统中查询并获取到对应的交易信息，通过自身私钥，执行解密函数 Decrypt 把得到的哈希密文还原为哈希明文，使用哈希明文可以到 IPFS 系统中下载源文件。

5. 系统实现

5.1. 系统实现环境

区块链平台选择 Fisco Bcos 联盟链。编译部署 Solidity 智能合约，将数据哈希上链，提供交易上链、存储、交易查询等功能，合约基于 Fisco Bcos 的 EVM 运行。分布式存储系统选择 IPFS 星际文件系统并部署在 Linux 中，利用 JPBC 库实现基于 CL-PKC 的无证书密码体制，测试工具使用 Caliper。

5.2. 系统功能测试

将风力发电厂作为数据存入方，提供设备运行数据，将发电机运行数据存入 IPFS 系统中，把系统返回的哈希经过无证书密码体制加密后发送交易上链；发电机制造厂是运行数据哈希密文的获取者。

风力发电厂将数据哈希密文上链，上链请求回执信息如图 3 所示。

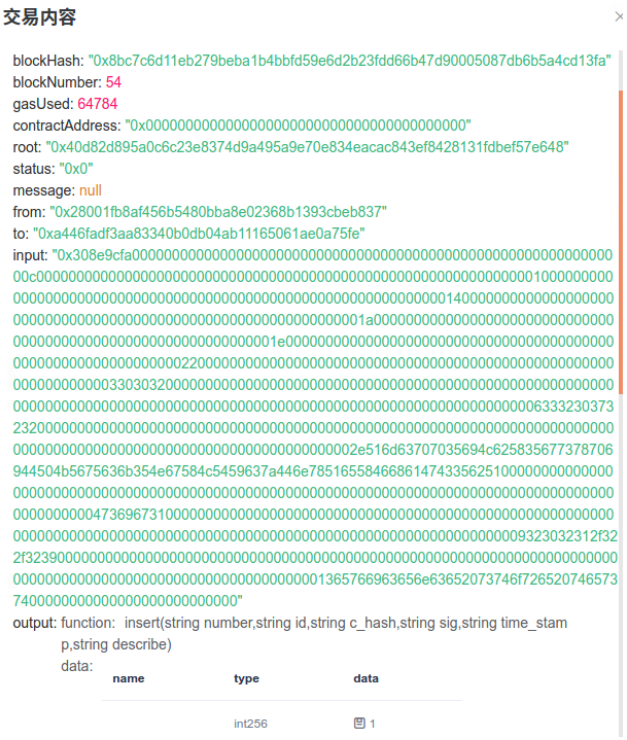

Figure 3. Hash on-chain request transaction receipt
图 3. 上链请求交易回执

发电机制造厂作为数据获取者利用智能合约查询链上数据，请求过程如图 4 所示，成功获取信息如图 5 所示。

Figure 4. Send transaction query request
图 4. 发送交易查询请求

交易内容


```
[
  [
    "002",
    "320722",
    "Qmcpp5iLbX5gsxpiDPKVuck5NgXLTYczDnxQeXFhaGC5bQ",
    "sig1",
    "evicence store test"
  ]
]
```


Figure 5. Query transaction receipt
图 5. 查询交易回执

5.3. 系统性能测试

针对 IPFS 系统性能进行测试，测试设计 4 个线程组，每组 2000 条查询请求，共查询次数 8000 次，测试结果如图 6 所示。

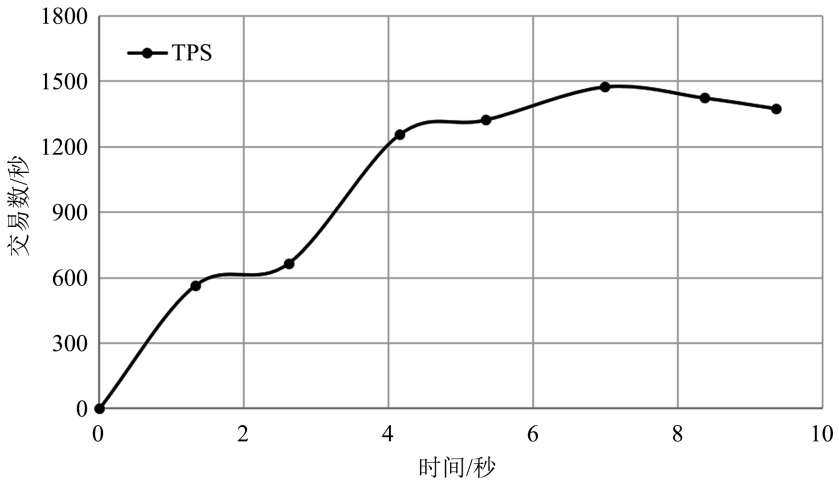

Figure 6. IPFS performance test experimental results
图 6. IPFS 性能测试实验结果

测试结果为 IPFS 系统平均 TPS 为：855.6/sec。从测试结果看，选择 IPFS 作为链下扩容存储方案可充分满足基于 IPFS 的区块链数据可信存储系统需求。

6. 结束语

工业数据的可信存储是工业互联网可以持续发展的重要前提。本文针对工业设备运行数据存储可信度低的问题，提出工业互联网设备运行数据的可信存储方法，提出了基于 IPFS 的区块链数据存储系统架构，系统内加入了无证书密码体制对数据使用权限进行进一步控制，并对方案进行了实现验证。实验结果表明，本文提出的方案可实现工业互联网设备运行数据可信存储。为今后区块链应用于工业互联网的研究提供了理论和实践基础。

基金项目

国家自然科学基金青年科学基金项目(61802217)。

参考文献

[1] 万晓霞, 焦智伟, 刘名轩, 刘段. 工业互联网应用综述[J]. 数字印刷, 2021(2): 1-26.

[2] 吴文君, 姚海鹏, 黄韬, 等. 未来网络与工业互联网发展综述[J]. 北京工业大学学报, 2017, 43(2): 163-172.

[3] 李阳春, 王海龙, 李欲晓, 等. 国外工业互联网安全产业布局及启示研究[J]. 中国工程科学, 2021, 23(2): 112-121.

[4] Tufail, H., Zafar, K. and Baig, A.R. (2019) Relational Database Security Using Digital Watermarking and Evolutionary Techniques. *Computational Intelligence*, **35**, 693-716. https://doi.org/10.1111/coin.12209

[5] Gazi, P., Kiayias, A. and Zindros, D. (2019) Proof-of-Stake Sidechains. 2019 *IEEE Symposium on Security and Privacy* (*SP*), San Francisco, 19-23 May 2019, 139-156. https://doi.org/10.1109/SP.2019.00040

[6] 周方明, 张泽华. 基于 IPFS 区块链技术的焊接数据存储与共享系统[J]. 徐州工程学院学报(自然科学版), 2021, 36(4): 10-17.

[7] 高文涛, 张桂芸. 基于联盟区块链和 IPFS 的音乐共享模型[J]. 天津师范大学学报(自然科学版), 2020, 40(2): 68-74.

[8] 范贤丽, 范春晓, 吴岳辛. 基于区块链和 IPFS 技术实现粮食供应链隐私信息保护[J]. 应用科学学报, 2019, 37(2): 179-190.

[9] 朱彦霞, 张雪萍, 华南, 罗刘敏. 基于 IPFS 及区块链的互联网融媒中心平台设计[J]. 电子设计工程, 2021, 29(18): 10-16.

[10] Zheng, Q., Yi, L., Ping, C., *et al*. (2018) An Innovative IPFS-Based Storage Model for Blockchain. 2018 *IEEE/WIC/ACM International Conference on Web Intelligence* (*WI*), Santiago, 3-6 December 2018, 704-708. https://doi.org/10.1109/WI.2018.000-8

[11] Kumar, R., Tripathi, R., Marchang, N., *et al*. (2021) A Secured Distributed Detection System Based on IPFS and Blockchain for Industrial Image and Video Data Security. *Journal of Parallel and Distributed Computing*, **152**, 128-143.

[12] 谭海波, 周桐, 赵赫, 等. 基于区块链的档案数据保护与共享方法[J]. 软件学报, 2019, 30(9): 2620-2635.

[13] 许丽, 何泰霖. 基于区块链的加密云存储平台模型研究[J]. 电脑编程技巧与维护, 2022(3): 90-92+102. https://doi.org/10.16184/j.cnki.comprg.2022.03.015

[14] 代闯闯, 栾海晶, 杨雪莹, 过晓冰, 陆忠华, 牛北方. 区块链技术研究综述[J]. 计算机科学, 2021, 48(S2): 500-508.

[15] 倪赛华. 基于分布式文件系统的海量数据存储专利技术综述[J]. 中国新通信, 2019, 21(3): 68-69.

[16] Wang, S., Ouyang, L., Yuan, Y., *et al*. (2019) Blockchain-Enabled Smart Contracts: Architecture, Applications, and Future Trends. *IEEE Transactions on Systems, Man, and Cybernetics*: *Systems*, **49**, 2266-2277.

[17] Wang, Z., Tian, Y. and Zhu, J. (2018) Data Sharing and Tracing Scheme Based on Blockchain. 2018 8*th International Conference on Logistics, Informatics and Service Sciences* (*LISS*), Toronto, 3-6 August 2018, 1-6. https://doi.org/10.1109/LISS.2018.8593225

[18] Al-Riyami, S.S. and Paterson, K.G. (2003) Certificateless Public Key Cryptography. 9*th International Conference on the Theory and Application of Cryptology and Information Security*, Taipei, 30 November-4 December 2003, 452-473. https://doi.org/10.1007/978-3-540-40061-5_29

[19] 杨小东, 麻婷春, 陈春霖, 等. 面向车载自组网的无证书聚合签名方案的安全性分析与改进[J]. 电子与信息学报, 2019, 41(5): 1265-1270.

[20] 孙华, 孟坤. 标准模型下可证安全的有效无证书签密方案[J]. 计算机应用, 2013, 33(7): 1846-1850.

[21] Zhao, Y., Hou, Y., Wang, L., *et al*. (2020) An Efficient Certificateless Aggregate Signature Scheme for the Internet of Vehicles. *Transactions on Emerging Telecommunications Technologies*, **31**, e3708. https://doi.org/10.1002/ett.3708