区块链赋能可信数据空间安全的研究进展

尚思远[1,2],杜学绘[1,2*],刘敖迪[1,2],王潇涵[1,2],吴翔宇[1,2],王 娜[1,2]
(1. 信息工程大学密码工程学院,河南郑州 450001;2. 河南省信息安全重点实验室,河南郑州 450001)

摘 要: 随着可信数据空间发展行动计划的提出,促进数据要素的交易与流通,解决大数据产业“数据孤岛”式困境,已经上升至国家战略高度. 区块链技术作为一种融合多种现有技术的分布式计算和存储范式,能够通过其不可篡改、去中心化、可溯源等优良特性,为数据可信管理提供有力支撑. 其不同于传统数据库的数据存储,而是通过不可篡改的数据结构与多方共识的信任机制为数据提供者与数据使用者建立信任纽带. 如何将区块链技术优势应用于可信数据空间安全建设,为国家打造可信数据空间的战略目标提供有力支撑,已成为亟需解决的问题. 目前,国内外已经有许多学者对区块链技术进行综述总结,包括具体技术组成,如共识机制、智能合约、网络拓扑等;具体领域应用,如信息安全、系统防护、数据管理等;以及对区块链的技术增强,如分片、跨链、隐私保护等. 然而,仍缺少区块链技术赋能可信数据空间安全的体系化综述研究. 基于此,本文从学术视角对区块链赋能可信数据空间安全研究进行综述分析. 首先,对可信数据空间基础架构与技术需求进行分析,从数据全生命周期角度提炼可信数据空间所面临的安全问题及挑战,提出一种贯穿数据获取、数据验证、数据共享、数据溯源环节的基于区块链的可信数据空间安全技术框架. 随后,将区块链技术与各类主流安全机制结合,从数据可信获取、数据合规验证、数据安全共享、数据联合溯源四个方面,系统梳理归纳区块链赋能可信数据空间安全的研究进展. 最后,对区块链赋能可信数据空间安全的发展趋势进行总结与展望,根据可信数据空间作为数据基础设施的基本需求,整理了不同数据流转阶段现有研究的优势和不足. 当前还需在链上数据检索、数据权属保护、数据法规落实等方面实现进一步技术突破,从而保障可信数据空间可信管控、资源交互、价值共创等核心能力,促进基于区块链的可信数据空间安全建设与技术发展.

关键词: 区块链;可信数据空间;数据共享;数据安全;隐私保护;技术赋能
基金项目: 国家自然科学基金(No.62102449);河南省科技攻关项目(No.252102211080)
中图分类号: TP309 **文献标识码:** A **文章编号:** 0372-2112(2025)12-4833-26
电子学报 URL:http://www.ejournal.org.cn **DOI:**10.12263/DZXB.20250617

Research Progress on Blockchain-Enabled Trusted Data Space Security

SHANG Si-yuan[1,2], DU Xue-hui[1,2*], LIU Ao-di[1,2], WANG Xiao-han[1,2], WU Xiang-yu[1,2], WANG Na[1,2]
(1. *School of Cryptography Engineering*, *Information Engineering University*, *Zhengzhou*, *Henan* 450001, *China*;
2. *Henan Key Laboratory of Information Security*, *Zhengzhou*, *Henan* 450001, *China*)

Abstract: With the 'release of the action plan for trusted data spaces, promoting the trading and circulation of data elements and addressing “data silos” has become a national strategic priority. Blockchain, as a distributed computing and storage paradigm, provides immutability, decentralization, and traceability to support trusted data management. Unlike traditional databases, it uses immutable data structures and multi-party consensus to build trust between data providers and users. An urgent priority is to harness these advantages to strengthen trusted data space security. Currently, numerous domestic and international reviews have surveyed blockchain's foundational technologies (e.g., consensus mechanisms, smart contracts, network topology), application areas (e.g., information security, system protection, data management), and technical enhancements (e.g., sharding, cross-chain, privacy protection). However, a systematic review of how blockchain enables trusted data space security remains lacking. Based on this, this paper offers a comprehensive academic review of blockchain-enabled trusted data space security. First, we analyze the core architecture and system requirements, identify security issues from a full-lifecycle perspective, and propose a blockchain-based security framework for data acquisition, validation, sharing, and provenance. Second, we integrate blockchain with mainstream security mechanisms and synthesize research progress across four domains: trusted acquisition, compliance verification, secure sharing, and federated prove-

收稿日期:2025-07-11;录用日期:2025-12-19;责任编辑:覃怀银
*通讯作者:杜学绘

nance. Third, we survey development trends in blockchain-enabled trusted data space security. Grounded in the foundational requirements of trusted data spaces as data infrastructure, we evaluate the strengths and limitations of existing work across all stages of data circulation. Further breakthroughs are needed in on-chain data retrieval, data ownership protection, and the enforcement of data regulations to safeguard core capabilities, including trustworthy governance, resource interoperability, and value co-creation, and to advance the security architecture and technological development of blockchain-based trusted data spaces.

Key words: blockchain; trusted data space; data sharing; data security; privacy protection; technology enablement

Foundation Item(s): National Natural Science Foundation of China (No.62102449); Henan Province Science and Technology Key Project (No.252102211080)

1 引言

为促进数据要素合规高效流通使用,2024年11月,国家数据局印发《可信数据空间发展行动计划(2024—2028年)》[1]. 该计划提出:“可信数据空间是基于共识规则,连接多方主体,实现数据资源共享共用的基础设施,是数据要素价值共创的应用生态,是支撑构建全国一体化数据市场的重要载体.”可见,可信数据空间如何建设落地、如何实现技术突破,已经上升至国家战略高度并得到充分重视. 推动可信数据空间的建设与发展,有助于打通数据壁垒,实现数据价值充分释放.

然而,随着数据流转速度的增加与流转范围的扩大,人们逐步认识到各类数据安全问题. 可信数据空间针对数据共享流通建立体系架构,赋予数据提供方与数据使用方互联互通手段,但数据在可信数据空间内同样面临众多风险挑战. 例如,数据获取阶段,数据真实性难保证且隐私意图易泄露;数据验证阶段,数据权属认定与责任追溯难实现;数据共享阶段,数据法规落实与权限管控难协同;数据溯源阶段,多方信任建立与可信存储方式仍缺失. 奇安信集团在《2024中国政企机构数据安全风险研究报告》[2]中指出,2024年全球公开报道的201起数据安全事件已造成超过122.7 TB的数据泄露,较2023年的51.8 TB增长136.9%;共计泄露数据471.6亿条,较2023年的103.8亿条增长354.3%. 表1列出了其报告中给出的2024年全球重大数据泄露事件. 面对严峻的数据风险,如何保证可信数据空间数据全生命周期安全已经成为亟需解决的难题.

表1 2024年全球重大数据泄露事件

序号	数据泄露机构	泄露数据类型	泄露数据量(亿条/行)	事件时间	相关报道
1	Twitter、Adobe、LinkedIn	个人信息	260	2024年12月	https://www.secrss.com/articles/74206
2	RockYou2024	账号密码	100	2024年7月	https://www.secrss.com/articles/67823
3	背景调查公司Jerico Pictures Inc.	个人信息	30	2024年8月	https://www.secrss.com/articles/68841
4	美国国家公共数据公司	个人信息	29	2024年12月	https://www.secrss.com/articles/74206
5	ERP软件提供商ClickBalance	商业机密	7.7	2024年7月	https://www.secrss.com/articles/68482

区块链技术是一种新的分布式计算和存储范式,其利用链式数据结构验证数据,使用共识算法更新数据,应用密码学技术保护数据,并通过由自动化脚本代码组成的智能合约操作数据,已广泛应用于金融、政务、医疗等领域[3]. 由于区块链技术具有去中心化、不可篡改、可溯源等技术特性,其能够应用于构建数据流转体系,解决传统数据流转过程对于第三方数据服务的依赖. 越来越多的技术方案以其作为多方互信基础,记录数据全生命周期,实现数据流转全流程治理[4],或用于推进万维网联盟发布的去中心化数字身份[5]与可验证凭证[6]标准. 引导可信数据空间运营者利用区块链技术提升可信数据空间数据管控能力,解决可信数据空间安全问题,具有重要研究意义.

目前,国内外已经有许多学者对区块链技术进行综述总结,包括具体技术组成,如共识机制、智能合约、网络拓扑等[7~9];具体领域应用,如信息安全、系统防护、数据管理等[10~12];以及对区块链的技术增强,如分片、跨链、隐私保护等[13~15]. 然而,可信数据空间作为近年发源于产业界的新方向,仍缺少区块链技术赋能可信数据空间安全的体系化综述研究. 基于此,本文首次从学术角度对区块链赋能可信数据空间安全进行综述分析. 首先,采用映射研究方法,将可信数据空间内数据流转划分为获取、验证、共享、溯源4个关键阶段,分析各阶段安全问题及可应用的安全机制. 随后,在中国知网、Web of Science、EI Compendex等国内外数据库中,以数据空间、区块链、数据获取、数据验证、数据共享、数据溯源作为主关键词,可搜索加密、安全多方计算、联邦学习等具体安全机制作为副关键词,以2010—2025年作为时间窗,对发表的论文研究进行组合检索. 最后,参考中国计算机学会、中国密码学会推荐的期刊

会议目录以及中国科学院分区,优先选取三年内发表的高水平论文以及领域内高被引研究. 经过筛选,数据流转各阶段综述分析的研究类论文共98篇. 其中期刊论文73篇,会议论文25篇,IEEE Trans系列期刊占比58.9%,近三年发表的论文研究占比72.4%. 本文研究类论文年份统计与主题统计如图1所示.

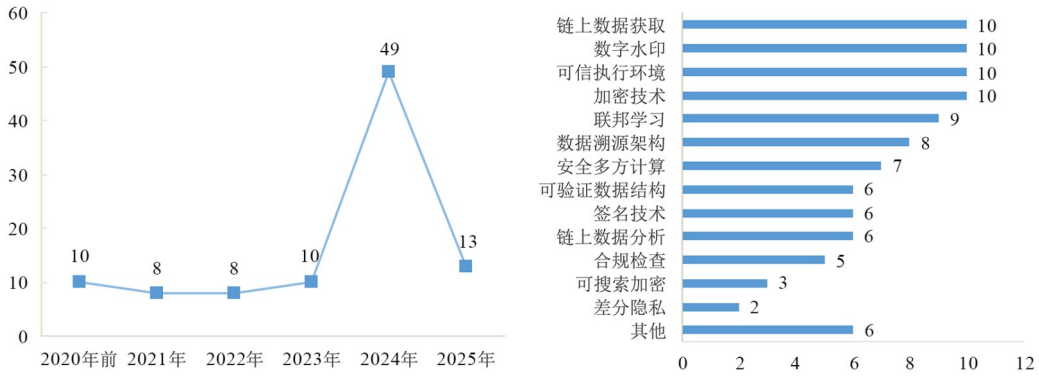

图1　本文研究类论文年份统计与主题统计

总的来说,本文主要贡献如下:

(1)构建了基于区块链的可信数据空间安全技术框架. 通过分析总结国内外可信数据空间基础框架现状,从数据全生命周期角度总结可信数据空间所面临的安全问题及挑战,提出基于区块链的可信数据空间安全技术框架,保障数据空间可信管控、资源交互、价值共创等能力落实.

(2)梳理了区块链赋能可信数据空间安全的研究进展. 依托构建的基于区块链的可信数据空间安全技术框架,从数据可信获取、数据合规验证、数据安全共享、数据联合溯源4个方面系统梳理了区块链赋能可信数据空间安全的研究进展,涵盖了区块链赋能可信数据空间的主要安全机制.

(3)总结了区块链赋能可信数据空间安全的核心问题和发展趋势. 结合可信数据空间作为数据基础设施的基本需求,归纳整理不同数据流转阶段现有研究的优势和不足,并建议结合区块链技术在链上数据检索、数据权属保护、数据法规落实等方面实现进一步技术突破.

2　区块链赋能可信数据空间安全技术框架

梳理可信数据空间基础框架,能够更深入理解其发展理念与建设目标,发现潜在风险挑战与技术需求. 本节首先对国际数据空间现状与计划中提出的可信数据空间基础框架进行介绍;随后在此基础上,提出基于区块链的可信数据空间安全技术框架.

2.1　可信数据空间基础框架

数据作为一种生产要素,需要实现规模化流通利用. 没有流通,就无法进行价值创造,对数据流通的需求促进了数据空间的建立. 数据空间最早起源于欧盟,其于2016年提出建立国际数据空间倡议. 经过发展,美国、日本等国家目前均已形成自己的数据空间体系. 表2对国际数据空间典型示例与核心能力进行总结.

表2　国际数据空间典型示例与核心能力

国家及地区	发展现状	技术架构	核心能力
欧盟	以数据法规为驱动,已形成农业、能源等不同领域数据空间应用	IDS国际数据空间架构[16]	以点对点形式实现可信数据交换,由数据提供商制定数据使用规则,保护数据主权
		Gaia-X数据空间生态[17]	整合云提供商、数据空间与数据基础设施,打造统一的数据生态系统
美国	尚未形成国家战略,但亚马逊、谷歌等知名企业已实现技术应用	NIEM信息交换模型[18]	以业务领域为驱动,采用去中心化治理模式,提升不同来源数据的互操作性
日本	以工业、智能制造等领域的特色项目为指引,推进数据空间建设	Ouranos数据生态系统[19]	采用面向服务的架构,通过松耦合架构与多视角设计增强灵活性
		CIOF互联工业开放框架[20]	针对工业场景建立数据网络,为中小型企业提供接入与管理方案

相比于国际上不同类别的数据空间，我国侧重于打造以“可信”为核心的可信数据空间，构建一个安全可信、可控可管的技术与治理体系，形成一套兼具安全性与灵活性的“中国方案”. 2022 年，中国信通院发布《可信工业数据空间系统架构 1.0 白皮书》[21]，并于 2024 年牵头发布《可信数据空间系统架构》国际标准[22]. 2024 年末，国家数据局印发的《可信数据空间发展行动计划（2024—2028 年）》[1]提出权威界定的可信数据空间基础框架，并迅速得到各界广泛关注与响应. 其明确指出，可信数据空间以三大能力建设作为支撑，即以可信管控能力支持对空间内主体身份、数据资源、产品服务等开展全要素、全过程、全场景的可信认证；以资源交互能力支持统一发布、高效查询、跨主体互认等服务，支撑数据接入、发布、发现、转换、交付全生命周期；以价值共创能力支持多主体共同参与数据开发利用，从企业协同、行业研发、城市建设、金融定制等多方面推动数据资源向数据产品转化. 图 2 展示了该可信数据空间基础框架的核心内容.

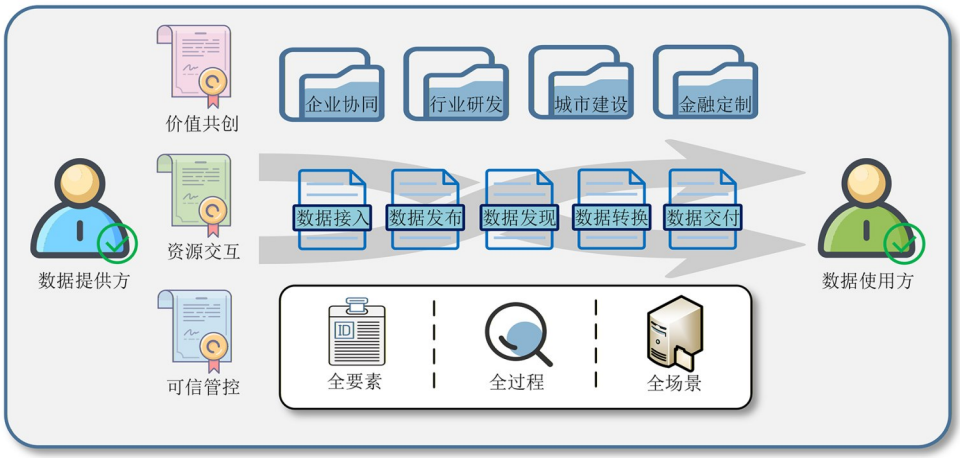

图 2 可信数据空间基础框架核心内容

2.2 基于区块链的可信数据空间安全技术框架

现有可信数据空间基础框架根据不同应用场景设定运营主体并制定规则，实现数据提供方与数据使用方之间的数据流通. 然而，其仅从宏观上为可信数据空间建立进行了概念指导，具体能力建设仍需要安全技术框架赋能. 近年来，数据的隐私保护与治理逐渐成为各国竞争的关键领域，中国、美国、欧盟等主要国家和地区相继制定并实施了一系列政策法规. 例如，我国出台《中华人民共和国数据安全法》[23]，美国发布《促进数据共享与分析中的隐私保护国家战略》[24]，欧盟形成以《通用数据保护条例》[25]为主体的数据法规体系. 这些政策法规引发了对可信数据空间安全的思考，研究人员正逐步探索其具体安全框架、如何建设等问题. 基于此，本文从数据全生命周期安全角度，将可信数据空间中的数据流转过程概括为 4 个阶段，即数据获取、数据验证、数据共享、数据溯源，并在此基础上提出一种基于区块链的可信数据空间安全技术框架，保障可信数据空间可信管控、资源交互、价值共创等能力落实，以期对可信数据空间建设提供启发与借鉴. 图 3 展示了该安全技术框架的具体组成.

国内区块链应用以联盟链为主，不同于无需授权即可参与的公链，联盟链参与节点需要经过审核与授权，主流联盟链平台包括 Linux Foundation 建立的超级账本 HyperLedger Fabric[26]以及金融区块链合作联盟开发的 FISCO-BCOS[27]. 前者采用通道机制实现数据隐私隔离，具有模块化、可插拔的架构设计，能够作为适用于不同领域的通用框架；后者基于开源区块链平台以太坊[28]的架构进行改进，采用群组机制实现数据隐私隔离，通过并行计算等底层性能优化提升吞吐量，且底层深度集成国密算法. 二者在不同方面的对位比较如表 3 所示. 目前，国内不同领域的区块链应用，如趣链[29]、蚂蚁链[30]、长安链[31]等，已经实现了初步落地.

本文提出的安全技术框架以联盟链作为底层支撑，针对数据流转各阶段问题，与各类安全机制结合进行技术增强，提出以区块链技术为支撑的数据可信获取技术、数据合规验证技术、数据安全共享技术、数据联合溯源技术，以期赋能可信数据空间三大能力建设，促进其支撑金融、政务、医疗等业务领域. 另外，由于公链发展较早，众多领域前沿高水平论文研究针对比特币、以太坊等平台并进行实验验证，本文同样分析借鉴其技术思路，为区块链赋能可信数据空间安全提供参考.

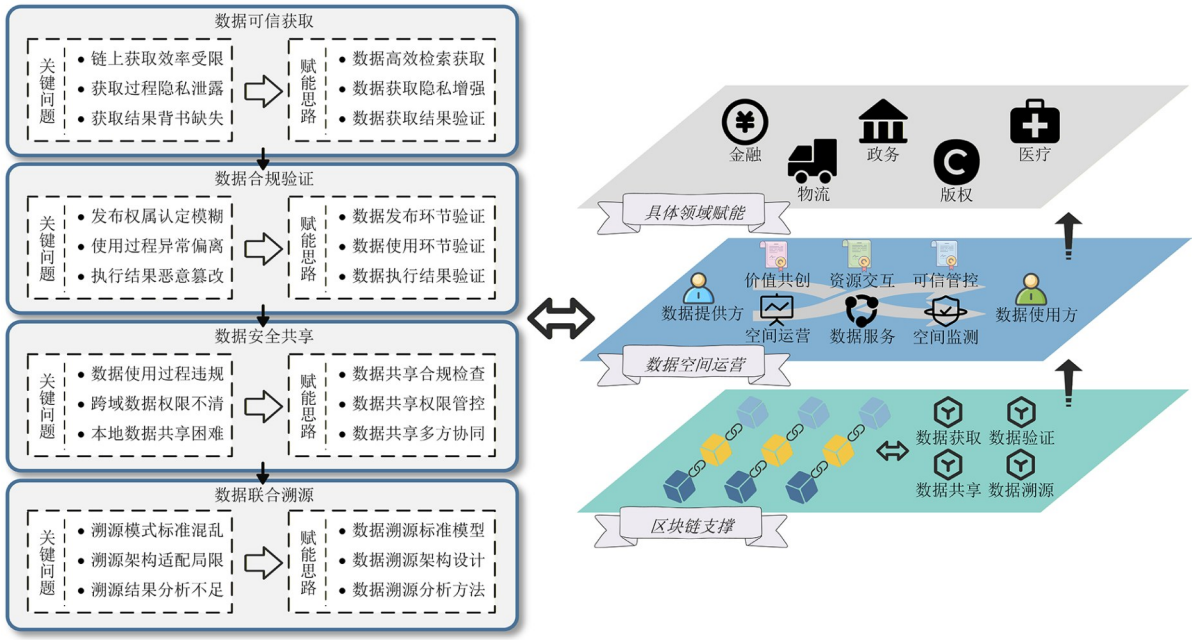

图3　基于区块链的可信数据空间安全技术框架

表3　主流联盟链平台 HyperLedger Fabric 与 FISCO-BCOS 对位比较

名称	权限管控	背书机制	隐私隔离	国密支持	可扩展性
HyperLedger Fabric	基于策略的权限控制	执行、排序、验证三阶段分离,可进行模块化替换	采用通道隔离机制,各通道数据隐私隔离	非原生支持,需要第三方技术改造	通过模块化、可插拔的架构实现扩展,能够作为适用于不同领域的通用框架
FISCO-BCOS	基于角色的权限控制	共识节点同时负责交易执行、交易打包和区块共识	采用群组隔离机制,各群组数据隐私隔离	原生支持,底层深度集成国密算法	通过并行计算等底层性能优化实现扩展,适合吞吐量要求高的业务场景

3　数据可信获取技术

数据获取是数据分析、利用的基础,也是利用区块链技术实现数据管理的前提. 可信数据空间作为国家未来的数据基础设施,需要结合区块链技术不可篡改特性,构建链上高质量数据索引. 假设空间内数据使用者需要与链上数据索引交互,其获取与使用均需要高效检索方案的支撑,并且如何保证过程隐私与结果可验证同样是当前亟需解决的难题. 基于此,本节提出图4所示的包含数据高效检索获取、数据获取隐私增强与数据获取结果验证3个阶段的技术路线,总结分析如何通过现有研究有效赋能可信数据空间,实现数据可信获取.

3.1　数据高效检索获取

链上数据具有不可篡改特性,可信数据空间需要链上数据高效检索技术支撑,帮助实现数据可信获取. 部分研究基于智能合约进行链上数据检索,并通过以太坊开展实验. 例如,Abuhashim 等人[32]分别按直接检索与基于索引构建的检索设计了两种智能合约. 直接检索在500个区块规模下需200 s时间开销,基于索引构建的检索能够通过建立的分类索引将时间开销降低至1 s以下. Chishti 等人[33]同样使用智能合约检索思路,构建支持单参数和多参数检索的数据结构,将区块链数据划分为不相交子集,并将数值型参数使用聚类算法转换为分类数据以提升检索效率. 在1 000个区块规模下,检索时间开销相比于顺序检索从25 s下降至0.11 s. 基于以太坊的智能合约检索同样为联盟链提供了借鉴思路. 在此基础上,Shang 等人[34]基于 HyperLedger Fabric 联盟链,使用局部敏感哈希对链上数据进行组织,通过智能合约逻辑执行检索,为链上数据的组织与使用提供了解决思路. 这些方案基于智能合约检索,与链上数据直接交互,通过智能合约逻辑优化赋能可信数据空间的链上数据获取.

以上检索方案能够与区块链数据直接交互,但智能合约的代码丰富度是核心限制因素,另有部分研究使用外部数据库协助. 例如,针对以太坊平台设计的 EtherQL[35],能够从链上获取并解析数据,转移至外部数据库进行 SQL(Structured Query Language)语句检索,其在真实以太坊数据下进行账户地址与区块范围的检索开销均在毫秒级. 该思路同样启发可信数据空间的

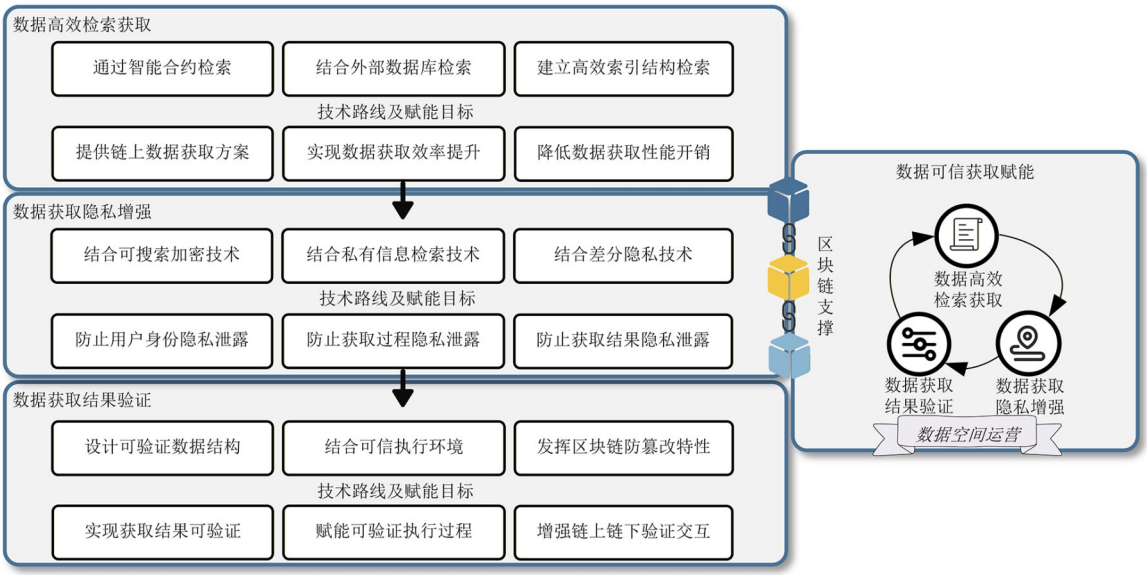

图4 数据可信获取技术路线

链上数据获取,可将链上数据解析至链下提供更丰富的检索方式. Wang等人[36]提出名为ForkBase的区块链存储引擎,支持高效检索及跨数据对象、分支和版本的重复内容检测. 其在本地64个HyperLedger Fabric节点集群开展实验,在确认时延与典型吞吐量配置下测试端到端性能,结果显示其对账本世界状态的读写操作能够达到数万次每秒,不考虑时延下对账本世界状态的写操作时间开销仅为微秒级. 这些方案使用外部数据库来扩展检索功能与性能,但需结合可信数据空间实际需求,考虑外部数据库真实性问题. 预言机是区块链和外部数据交互的纽带,为外部数据库的真实性保证提供了解决思路. Liu等人[37]面向工业互联网数据提出了可信预言机方案,通过数据过滤与节点过滤增强预言机节点的稳定性与获取数据的一致性,其在包含100个预言机节点的网络中连续执行1 000次预言机数据获取任务,时间开销能够稳定在3 s以下. 该技术思路能够保证可信数据空间使用外部数据库协助检索时的数据真实性,但需要预言机预先获取可信数据,面对实时数据检索时需考虑数据同步问题.

另外,链上数据检索需要效率保证,构建高效索引结构或建立检索层是有效解决思路. 例如,Yao等人[38]结合语义信息建立块间索引,存储在区块头,用于快速定位区块,支持语义驱动的关键词检索. 其使用以太坊真实数据开展实验,索引构建的平均时间开销在2 s以下,平均存储开销在5 KB以下. Liu等人[39]加入块内数据聚合索引与块间高效检索索引,将传统区块体分为索引层和数据层,保持原始数据与区块间的关联性. 其在HyperLedger Fabric开展实验,相比于原区块结构,其方案构建100个区块所需时间开销增加了近2 s,但能够通过索引构建将检索复杂度降低至常数级. 这些方案的区块结构设计提升了检索效率,但需考虑区块构建过程的时间与存储开销. 另外,可信数据空间同样需要将链上数据检索扩展至多类型数据. Yao等人[40]将链上数据构建为有向交易图并存储至检索层,在不改变区块存储结构的情况下赋能图数据检索,但其图结构的同步与更新时间呈线性增长,10 000个区块规模下需要200 s以上时间开销. Cui等人[41]设计混合索引进行关键词检索,并基于分区技术与布隆过滤器进行效率优化,其检索开销在毫秒级,但方法的索引存储空间随关键词数量线性增长,同样需考虑索引更新带来的时间开销.

总的来说,以上方案能够结合检索实时性、检索效率、数据同步等不同方面的综合需求,增强可信数据空间链上数据获取能力,充分发挥链上可信数据对链下数据的赋能作用. 数据高效检索获取阶段总结分析如表4所示.

3.2 数据获取隐私增强

以上研究能够实现数据的高效检索获取,但可信数据空间内数据获取过程存在隐私泄露风险,即在检索时泄露身份敏感信息,或者从检索结果中推断检索意图与数据隐私,如何实现可信数据空间的数据获取隐私增强同样至关重要.

研究人员将区块链技术与可搜索加密结合提供数据获取过程隐私增强. 例如,Han等人[42]通过关键词盲化处理对可搜索加密进行增强,确保搜索陷门生成过程无法感知用户的检索隐私,并将检索执行全过程记录至链上,结合区块链不可篡改与可溯源特性追踪非法操作. 在此基础上,Yu等人[43]使用区块链节点代替

表4　数据高效检索获取总结分析

相关工作	核心方法	技术分析	不足之处	适用场景	可信数据空间应用
文献[32~34]	基于智能合约的数据检索获取	通过智能合约直接检索链上数据,结合逻辑优化提升检索效率	需考虑智能合约执行开销与代码丰富度	需与链上数据直接交互,且性能要求不高的场景	通过智能合约帮助空间运营直接与链上数据交互
文献[35,36]	基于外部数据库协助的数据检索获取	通过外部数据库协助数据检索,提升检索效率并扩展检索方式	受限于外部数据库真实性	数据量大且对检索性能要求高的场景	通过外部数据库协助增强空间内数据管理能力
文献[37]	基于区块链预言机增强的数据检索获取	通过预言机技术保证获取数据可信	需考虑数据同步问题	不具有实时检索需求,允许进行数据同步的场景	通过预言机技术保证空间内数据获取可信
文献[38,39]	基于链上索引优化的数据检索获取	通过区块头与区块体的索引结构设计提升检索效率	区块构建时间与存储开销增加	仅考虑链上优化,弱化链下设计的场景	通过区块结构设计增强空间内链上数据检索效率
文献[40,41]	基于多类型检索层设计的数据检索获取	通过检索层设计为链上数据提供多类型检索方式	需考虑检索层构建与同步问题	效率需求高,需要多种检索方式的场景	通过多类型检索层增强空间内链上数据检索能力

中心化的云服务,结合Pedersen承诺生成搜索陷门,增强了可搜索加密执行过程的去中心化. Huang等人[44]提出了基于区块链的数据共享方案,将数据主体存储至云存储服务,构建密态索引存储至链上,通过智能合约执行检索并授权数据使用. 图5展示了基于可搜索加密的数据获取技术思路,其将密态索引存储至链上实施检索,结合可搜索加密的密态检索特性与链上数据的不可篡改特性实现隐私增强,支持关键词检索、精确检索、模糊检索等不同检索类型. 其威胁模型通常假设诚实且好奇的数据服务,正确返回检索结果但被动观察记录检索过程. 该过程面临不可避免的信息泄露,其泄露函数通常由3个阶段的泄露内容组成,即初始化阶段泄露的数据与关键词信息,检索阶段泄露的模式信息,例如访问模式泄露搜索陷门和匹配文件之间的关系,搜索模式泄露搜索陷门中是否包含相同的关键词,以及更新阶段新增数据及关键词的信息. 因此,其使用需要综合考虑不同类别的攻击威胁,例如针对陷门可链接性的攻击等.

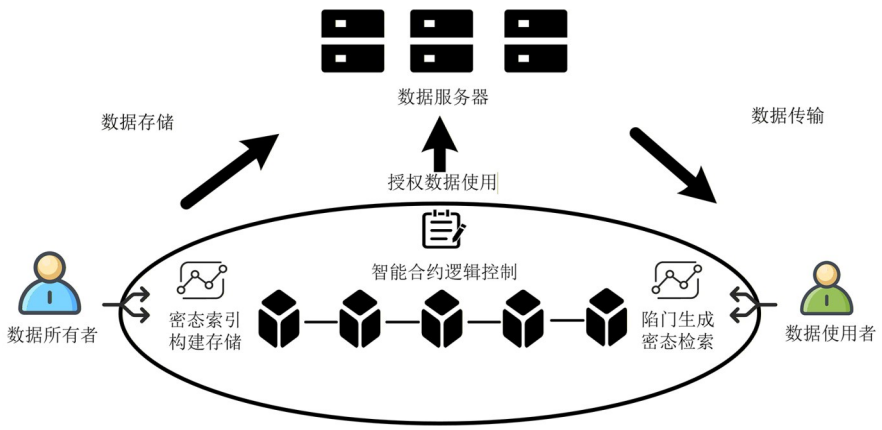

图5　基于可搜索加密的数据获取隐私增强技术思路

可信数据空间内的数据获取隐私增强同样能够结合其他安全技术. 例如,Ge等人[45]基于布隆过滤器进行数据隐私获取,在不暴露数据内容下执行检索. 其方案能够根据检索历史和隐私泄漏容忍阈值动态调整隐私保护性能,实现了基于混合策略纳什均衡的检索隐私保护,但受限于布隆过滤器带来的假阳性结果,在误报率可接受的情况下隐私保护性能下降. Hou等人[46]提出了一种基于变分自编码器的数据转换模型,能够防止在数据从终端设备传输的过程中,攻击者使用机器学习技术针对特定模式进行隐私提取,可用于可信数据空间内数据传输过程的隐私增强,但其隐私保护效果受限于模型训练的数据质量. Xie等人[47]通过私有信息检索技术进行隐私增强,由客户端构造双检索份额,对两个存储索引与文件的数据服务实施隐私检索,其威胁模型假设数据服务间不会合谋共享它们各自收到的检索份额,在该条件下能够有效保证检索过程隐私. 然而,该思路在可信数据空间内使用需考虑索引维护与通信开销问题. Xu等人[48]结合差分隐私技术,将噪声算法与数据获取过程结合,在数据获取结果的可用性与隐私性间进行权衡,但需考虑差分隐私使用过

程的参数选择,且差分隐私的使用仅保护数据获取结果.类似地,Zhang等人[49]针对跨链交易互操作过程的隐私性建立第三方公证机制,统一不同节点之间的隐私偏好,在回复外部数据获取请求前进行数据扰动,并提供了效率优先与隐私优先方案.

总的来说,现有技术对数据获取不同阶段与内容具有不同的隐私增强效果,可根据不同场景需求赋能可信数据空间建设,保护空间内数据提供者数据隐私与数据使用者使用隐私.数据获取隐私增强阶段总结分析如表5所示.

表5 数据获取隐私增强总结分析

相关工作	核心方法	技术分析	不足之处	隐私保护级别	效率	可信数据空间应用
文献[42~44]	基于可搜索加密的数据获取隐私增强	通过链上存储密态索引,利用可搜索加密保护检索过程隐私	受限于链上存储与检索执行开销	较高,执行密态检索,但可能泄露模式信息隐私	较低,密态索引的检索与构建存在较大开销	帮助构建空间内链上密态索引并执行检索
文献[45]	基于布隆过滤器的数据获取隐私增强	通过布隆过滤器执行隐私检索并根据检索历史和容忍阈值权衡	检索结果易出现假阳性	低,误报率可接受情况下隐私保护下降	高,布隆过滤器技术提供良好检索性能	为空间提供支持隐私权衡的高效检索手段
文献[46]	基于数据转换模型的数据获取隐私增强	通过变分自编码器防止攻击者针对数据特定模式的隐私提取	仅适用于数据传输过程	中,保护数据隐私,但受数据质量影响	中,具有模型训练开销,训练后可直接使用	提供空间内数据获取传输过程隐私保护
文献[47]	基于私有信息检索的数据获取隐私增强	通过私有信息检索技术保护检索过程隐私,建立双数据服务存储文件与索引	需考虑索引存储与维护问题	高,能够有效保护检索过程隐私	低,双数据服务设计增加通信与检索执行开销	提供空间内检索意图与检索过程隐私保护
文献[48,49]	基于差分隐私的数据获取隐私增强	通过差分隐私技术提供数据获取结果可用性与隐私性权衡	需考虑使用过程的参数选择	较低,仅保护数据获取结果,需权衡可用性	较高,噪声添加过程开销较小	提供空间内数据结果可用性与隐私性权衡

3.3 数据获取结果验证

以上研究对数据获取过程提供隐私保护,但可信数据空间内数据获取结果的可验证性同样关键.基于此,研究人员开展了针对性研究.

可信数据空间内每个区块链节点均维护完整链上数据具有较大存储开销,轻节点如何向全节点发起检索请求并验证检索结果成为一大难题,部分方案通过设计可验证数据结构进行技术增强.例如,Xu等人[50]提出了vChain方案,通过重构数据所在的默克尔树实现数据获取结果正确性与完整性验证,并设计与之适配的块内和块间可验证索引提升验证效率.在此基础上,Wang等人[51]提出了vChain+方案,通过窗口滑动实现快速检索验证,显著提升了验证效率.Xu等人[52]针对时空数据检索进行可验证数据结构设计,在区块内增加字段以支持关键词与近似最近邻检索,并提出了相应检索算法,以相同思路进行检索验证.针对图数据,Wu等人[53]通过RSA(Rivest-Shamir-Adleman)累加器高效验证元素存在性,在区块链辅助的云环境下实现可验证图查询.为解决可验证检索过程的兼容性,Wang等人[54]与现有数据库引擎集成,其方案支持多种检索类型,并能够适用于多种区块链系统.为增强可验证检索过程效率,Sun等人[55]将布隆过滤器与默克尔树结合提出一种链上数据结构,能够加速验证过程.这些方案通过可验证数据结构设计为多场景下数据获取结果验证赋能,但可信数据空间具体应用需考虑区块结构变化后的链上存储开销增加.

可信数据空间同样能够结合可信执行环境增强数据获取结果可验证性.例如,Cai等人[56]将可信执行环境与链上数据检索结合,将关键逻辑置于可信执行环境执行,利用其隔离特性为数据获取结果提供验证.Zhou等人[57]设计了基于可信执行环境的可验证数据库,支持SQL检索.其在存储层通过验证算法进行后台持续验证,检索过程同样在可信执行环境内执行.可信执行环境的使用保证了检索过程,增强了获取结果的可验证性,但在可信数据空间内引入仍需考虑现有技术的具体性能,例如执行内存限制.

可信数据空间内不仅存储链上数据,同样需要增强链下数据获取的可验证性,研究人员同样进行了针对性研究.例如,Wu等人[58]提出了一种名为VQL的可验证检索层,提取链上数据并高效重组至数据库增强检索功能性能,生成加密的数据库指纹上链存储实现交叉验证.Jiang等人[59]基于以太坊平台提出了可验证数据获取框架,将加密文件存储至IPFS(InterPlanetary File System)系统或云服务器,将加密元数据索引存储至区块链,在数据获取后利用链上数据验证链下数据完整性.这些方案建立链上链下协同机制,通过不可篡改的链上数据增强链下数据可验证性,但需考虑数据同步问题.

总的来说,现有方案利用链上数据结构对数据获取结果进行验证,对可信数据空间内区块链与其他存储服务的结合使用具有重要借鉴意义. 数据获取结果验证阶段总结分析如表6所示.

表6　数据获取结果验证总结分析

相关工作	核心方法	技术分析	不足之处	验证开销	可信数据空间应用
文献[50~55]	基于可验证数据结构的数据获取结果验证	通过链上可验证数据结构设计提供数据获取结果的可验证性	链上可验证设计增加了存储开销	低,通过可验证数据结构设计提供验证效率	提供空间内链上可验证数据结构设计
文献[56,57]	基于可信执行环境的数据获取结果验证	通过可信执行环境保证获取过程,实现获取结果可验证	需考虑可信执行环境执行内存限制	高,可信执行环境安全增强性能开销大	保证空间内数据获取过程与结果可验证
文献[58,59]	基于链上链下协同的数据获取结果验证	通过链上不可篡改数据验证链下数据主体	需考虑数据同步问题	中,需要建立链上数据与链下数据映射关系	建立空间内链上链下数据协同验证机制

4　数据合规验证技术

数据验证是确保数据质量与可信度的关键环节,也是利用区块链技术实现数据管理的保障. 近年来,数据流转速度的增加与流转范围的扩大引发了数据盗版、数据篡改等一系列问题,数据所有权难以得到有效保护. 可信数据空间作为数据基础设施,假设空间内数据提供者需要共享自身数据,如何结合区块链技术进行数据发布前的权属保障、数据使用过程与结果的可信验证是当前亟需解决的难题. 基于此,本节提出图6所示的包含数据发布环节验证、数据使用环节验证与数据执行结果验证3个阶段的技术路线,总结分析如何通过现有研究有效赋能可信数据空间,实现数据合规验证.

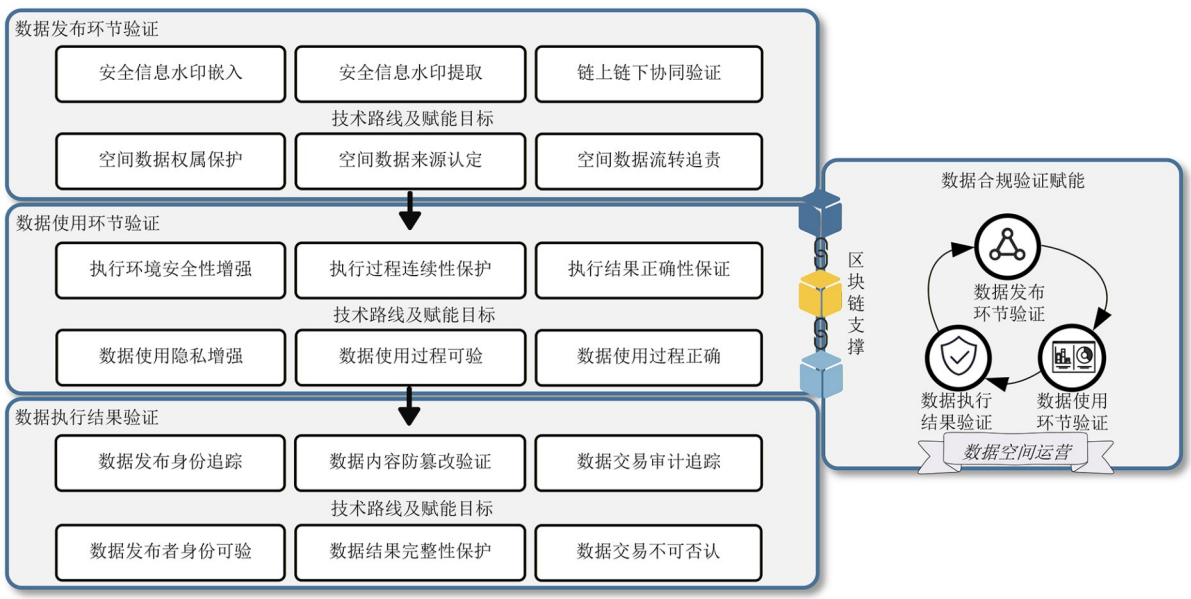

图6　数据合规验证技术路线

4.1　数据发布环节验证

可信数据空间数据获取阶段得到的可信数据需要在发布与流转后实现有效追责. 随着大语言模型的普及,亟需有效手段在可信数据空间内对其产出的图像、文本及音频等数字资产进行权属保护与来源认定. 在数据发布前进行验证,将版权、权属、安全等级等信息以水印形式嵌入数据主体是该问题的有效解决思路. 研究人员已发现该技术思路优势并与区块链结合解决实际问题.

针对图像数据,Xu等人[60]利用智能合约实现图像水印嵌入与提取的自动化,引入同态加密增强水印内容的安全性,帮助可信数据空间内数据提供者安全地共享嵌入水印信息的数据. Xiao等人[61]提出了名为FingerChain的去中心化多媒体共享方案,以区块链为支撑提升运营透明度,通过非对称数字水印支持用户端的水印嵌入与提取. 其水印算法在裁剪、压缩、噪声

等鲁棒性测试下仍具有高识别率,能够应用于可信数据空间,落实数据价值创造需求. 针对传统图像水印导致的图像质量下降问题,Wang等人[62]提出了一种基于零水印的图像版权保护框架. 其通过获取稳定的图像特征生成零水印图像,不对原始图像进行修改. Li等人[63]将水印信息转换为二值图像,采用置乱加密生成含水印图像,将身份标识与水印参数存储上链,在水印提取阶段验证身份合规性与水印有效性.

以上方案针对图像数据,其他类型数据同样能够进行关键信息嵌入与验证. 例如,针对文本数据,Rizzo等人[64]提出一种基于同形字符替换的文本水印思路,使用替代的Unicode符号保证视觉不可区分. Yoo等人[65]将文本数据转换为二进制表示,并通过同义词替换实现比特信息的嵌入. 另有部分研究针对大模型输出内容进行研究,Gu等人[66]针对大模型训练过程,对其能否直接生成含水印信息的文本展开研究. Wang等人[67]在大模型输出层采用可编码水印,将多比特信息隐蔽嵌入到大模型输出. 针对音频数据,Liu等人[68]提出Groot音频水印,水印生成和音频合成过程同时进行,通过轻量级解码器检索嵌入音频中的水印,在多种攻击手段干扰下仍能保持平均95%以上精度. 针对训练后模型参数,Wang等人[69]提出了一种模型所有权验证框架,根据模型关键参数生成水印并通过预先保留的后门验证. 这些方案能够根据不同数据类型进行水印设计,实现信息的有效嵌入与提取.

总的来说,在可信数据空间内建立数据发布前的验证阶段,实现数据发布前的水印嵌入与数据使用后的水印提取,是解决可信数据空间内数据权属问题的有效技术思路. 现有研究主要将关键水印参数存储至链上用于后续验证,利用链上不可篡改性保证水印信息的完整性,但仍缺少针对链上数据结构的水印设计. 由于水印技术的特殊性,关键信息常被转化为比特表示,如何实现链上链下协同的数据验证,将是未来的有效改进思路之一. 数据发布环节验证总结分析如表7所示.

表7 数据发布环节验证总结分析

验证数据类型	相关工作	技术分析	适用场景
图像数据	文献[60]	通过同态加密保护水印内容,使用智能合约执行	适用于对水印安全性要求高的场景
	文献[61]	通过非对称数字水印进行信息嵌入与提取	适用于提供用户端可交互性的场景
	文献[62]	通过零水印技术解决信息嵌入后图像质量下降问题	适用于对原图像质量要求高的场景
	文献[63]	通过链上存储身份标识与水印参数实现协同验证	适用于需提供不可篡改凭证的场景
文本数据	文献[64]	通过同形字符替换在文本中嵌入信息	适用于需保持文本信息并实现隐蔽嵌入的场景
	文献[65]	通过同义词替换在文本中嵌入信息	适用于在保持语义通顺下嵌入安全信息的场景
大模型输出	文献[66,67]	针对大模型训练及推理过程,将信息嵌入大模型输出	适用于防止模型滥用并提供溯源审计的场景
音频数据	文献[68]	音频合成过程生成水印,通过轻量级解码器检索音频水印内容	适用于需快速检测和验证音频数据版权的场景
模型参数	文献[69]	根据模型关键参数生成水印,通过预先保留的后门验证	适用于需保护机器学习模型知识产权的场景

4.2 数据使用环节验证

可信数据空间内的数据执行使用需要提供可验证性. 可信执行环境是一种安全的计算环境,通过硬件和操作系统的隔离机制,在不受外部干扰的情况下执行代码和处理数据. 在可信数据空间内建立可信执行环境,利用其特性增强特定数据使用环节是解决该问题的有效思路.

研究人员将可信执行环境与区块链技术结合,提出了众多赋能数据使用环节验证的技术思路. 例如,Fei等人[70]结合区块链设计了基于可信执行环境的密钥生成与分发方案,以区块链存储身份凭证,以可信执行环境验证凭证真实性并执行密钥生成,提供了硬件级的安全信任锚点,可并行处理多达16个密钥请求,能够为可信数据空间内身份验证、密钥协商等关键环节提供保障. 类似地,Jie等人[71]基于区块链设计安全灵活的离线支付协议,利用链上智能合约与离线钱包交互,结合可信执行环境维护离线账户状态并提供交易的完整性验证,实验结果显示该方案各阶段时间开销在毫秒级,在可信数据空间内具有可行性.

云服务是可信数据空间内数据使用的重要组成部分,其应用程序状态的连续性保护至关重要. 针对此问题,Peng等人[72]使用区块链初始化分布式系统,结合可信执行环境为云服务进行快速且安全的状态更新. 其方案创新地使用区块链赋能可信执行环境安全,将区块链作为去中心化信任锚点,构建分布式可信执行环境,利用区块链不可篡改记录存储配置信息与状态信息,解决了可信执行环境面临的回滚攻击与分叉攻击,即将状态回退到旧版本或创建同一应用的多个并行实例产生状态分歧. 该方案实现了区块链与可信执行环境的有机结合,二者互为技术增强,为可信数据空间内的应用部署提供了技术思路.

部分研究结合可信执行环境增强区块链中的共识协议与信任管理. 例如,Zhao等人[73]提出了一种可信执行环境辅助的无领导拜占庭共识协议,结合可信执行环境内计数器设计相关协议原语,实现了广播过程

身份与内容的可验证.Gu等人[74]基于联盟链建立车辆信任管理系统,通过数据驱动和事件驱动双模式评估车辆行为可信度.其通过可信执行环境安全存储车辆身份证书,保证参数生成过程可验证,并建立激励模型促进评分者提供真实信任评级.

另外,可信数据空间以智能合约实现链上数据交互与逻辑执行,是数据使用环节的关键支撑.智能合约的直接使用面临数据泄露风险,可信执行环境能够为智能合约安全赋能.Li等人[75]通过富有表现力的语法和安全属性对可信执行环境赋能的智能合约进行定义,并建立安全模型进行分析,揭示了在可信执行环境中执行复杂合约逻辑的性能挑战.Lu等人[76]针对物联网场景下数据交易合约的使用,提出了基于可信执行环境的数据交易隐私保护方案,通过正确性与完整性证明解决智能合约在可信执行环境中面临的状态回滚攻击.对于智能合约的全生命周期安全,Li等人[77]使用分布式可信执行环境集群构建基于投票的智能合约保护机制.一旦发现已部署合约的漏洞,进行合约销毁与重部署.总的来说,可信数据空间内能够借鉴以上技术思路建立可信执行环境集群,将密钥管理、数据交易、共识过程等数据使用的关键环节转移至可信执行环境,提供数据使用可验证性,但该过程仍需综合考虑可信执行环境自身的性能开销与安全漏洞.数据使用环节验证总结分析如表8所示.

表8　数据使用环节验证总结分析

相关工作	赋能环节	技术分析	可落地性分析
文献[70]	密钥生成与分发	通过可信执行环境验证凭证真实性并执行密钥生成,提供密钥生成与分发环节可验证性	可落地性较高,可通过建立密钥存储池减少通信开销与计算开销
文献[71]	数据交易验证	通过可信执行环境维护离线账户状态,提供数据交易过程可验证性	可落地性高,仅将验证计算的过程转移至可信执行环境执行
文献[72]	云服务执行过程	通过分布式可信执行环境增强云服务执行过程可验证性,以区块链解决回滚攻击与分叉攻击	可落地性低,需要重构云服务集群使其与可信执行环境结合
文献[73,74]	共识与信任管理	通过可信执行环境增强共识协议与信任管理系统内信任评估过程可验证性	可落地性较低,需要改造共识协议,综合考虑节点数量与通信开销
文献[75~77]	智能合约安全执行	通过可信执行环境增强智能合约全生命周期安全,提供执行过程可验证性	可落地性适中,复杂计算任务可转移至链下处理

4.3　数据执行结果验证

可信数据空间需要支持对数据执行结果进行验证,追踪数据提供者身份,验证数据内容完整性,并对数据交易进行审计追踪.签名技术是区块链技术的支撑技术之一,对可信数据空间内数据执行结果进行签名验证是在数据发布、使用后的必需环节.基于以上需求,研究人员改进各类签名技术,将其与区块链结合以解决实际问题.

部分研究针对执行结果的高效验证开展研究.例如,Jiang等人[78]设计了一种基于Schnorr签名的多重签名,与智能合约集成以实现区块链中的多参与方协同签名.传统思路将多个签名简单串联,需逐个验证每个签名.其方案将一种识别机制结合进多重签名,通过多方问题到两方问题的简化,实现了签名大小与签名用户数量的无关.该技术思路能够优化可信数据空间内数据交易验证,增强了可扩展性.针对物联网数据,Burgos等人[79]基于ECDSA签名提出了一种创新的物联网数据验证框架.传统ECDSA签名验证对物联网设备计算能力要求过高,其方案通过部署边缘服务器增强签名验证效率并结合零知识证明保证其行为正确性,解决了物联网设备和区块链间持续、直接交互的需求,为可信数据空间针对物联网数据的高效验证提供了技术思路.

近年来,门限签名作为区块链中实现多方协同和信任管理的关键技术,得到了广泛的研究.例如,Xie等人[80]提出了一种可问责门限签名,能够精确追踪恶意签名方,并通过主动刷新来增强密钥安全性.其通过签名份额生成的并行化提升吞吐量,签名速度达到毫秒级,促进了其在区块链系统中的应用.Duan等人[81]针对典型隐私保护协议RingCT设计了可链接门限环签名,使多个付款人能够在不泄露其密钥的情况下共同构建匿名数据交易,并解决了交易规模随参与用户线性增加的问题.两种方案均促进了门限签名在可信数据空间内的实际应用.图7展示了基于门限签名的数据验证技术思路.

同样,部分签名技术综合考虑抗量子需求.例如,Yu等人[82]为实现量子安全,提出了无证书抗量子盲签名和无证书抗量子聚合签名,应用于物联网场景与智能发票系统.其签名具有签名密钥短、执行效率高等优势,为区块链在后量子时代的持续发展提供了重要支撑,提供了可信数据空间内抗量子攻击的数据验证技术方案.另外,签名技术与区块链中智能合约的结合同样能够为数据验证提供有力支撑.Wu等人[83]设计了一种轻量级无配对属性基签名,实现对智能合约灵活高

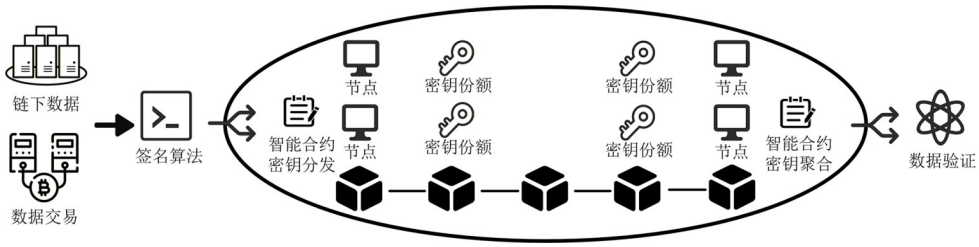

图7 基于门限签名的数据验证技术思路

效的访问控制,解决了资源受限的区块链节点无法适配访问控制轻量级要求的问题. 其策略变更无需重部署合约,策略更新延迟从分钟级降至秒级,为可信数据空间内基于智能合约实现数据权限验证提供了新思路.

总的来说,签名技术特性决定了其能够对可信数据空间的数据结果进行来源、完整性等多方面验证,空间内的数据交易同样能够结合签名技术进行审计追责. 现有研究结合区块链实际改进签名技术,众多方案能够通过功能增加与性能增强提供可验证性支撑,赋能可信数据空间建设. 数据执行结果验证总结分析如表9所示.

表9 数据执行结果验证总结分析

相关工作	核心方法	技术分析	技术特点
文献[78]	基于Schnorr签名的执行结果验证	通过基于Schnorr的多重签名进行多用户协同验证,实现了签名大小与签名用户数量无关	具有批量验证与线性特性,签名大小固定,链上开销小
文献[79]	基于ECDSA签名的执行结果验证	通过部署边缘服务器增强ECDSA签名验证效率,并结合零知识证明保证其行为正确性	区块链广泛使用的签名技术方案,技术生态强
文献[80,81]	基于门限签名的执行结果验证	通过主动更新等机制增强门限签名,实现多参与方数据交易匿名性与可验证性	私钥由多参与者持有,通过多方协同解决单点故障
文献[82]	基于无证书抗量子签名的执行结果验证	通过无证书抗量子盲签名和无证书抗量子聚合签名增强数据验证过程	无需复杂的证书管理,能够抵御未来量子计算机攻击
文献[83]	基于属性基签名的执行结果验证	通过属性基签名赋能智能合约访问控制实现可验证权限管控,策略变更无需重部署合约	能够结合复杂业务,具有细粒度权限控制与高度灵活性

5 数据安全共享技术

在保证提供数据权属认定及过程与结果的合规验证后,数据能够以共享为目的在可信数据空间内流通. 可信数据空间需要避免对云服务器、数据中心等第三方服务的过度依赖,区块链去中心化的技术特性提供了解决方案. 当前,越来越多的技术方案通过区块链技术代替第三方服务,将其与各类技术结合以达成数据安全共享. 假设在可信数据空间内发起多方协同的数据合作训练,如何结合区块链技术,保证该过程数据使用的合规性、数据权限的可控性及多方协同的安全性是当前亟需解决的难题. 基于此,本节提出如图8所示的包含数据共享合规检查、数据共享权限管控与数据共享多方协同3个阶段的技术路线,总结分析如何通过现有研究有效赋能可信数据空间,实现数据安全共享.

5.1 数据共享合规检查

数据共享前首先需要进行数据分析使用的合规性检查. 随着欧盟《通用数据保护条例》[25]的提出,各国对数据权利保护的重视程度进一步提升,可信数据空间需要在数据共享前检查数据分析程序是否满足政策法规要求. 然而,随着数据政策法规复杂度的增加,亟需针对数据合规检查过程的有效技术方案.

研究人员针对此问题展开研究,Wang等人[84]结合政策法规中的数据隐私要求,提供了一种新的技术思路. 其参照访问控制策略结构定义了检查数据分析程序合规性的分析策略,结合通过访问控制判决进行数据权限授予的思路,对分析策略同样进行判决,但并没有给出具体的开源工具. 随后,其在原有工作基础上提出了名为PrivGuard的系统设计[85]. 他们对常用的Python库函数进行分析,对程序代码静态分析以执行数据分析程序的合规性检查. 同样,Ferreira等人[86]针对Web开发场景,提出了名为RuleKeeper的数据分析程序检查器. 其允许开发人员提交程序开发标准并自动转化为分析策略,通过静态分析检查开发程序代码,帮助验证开发内容是否合规. 这些方案针对程序代码是否满足数据政策法规进行分析,能够应用于可信数据空间作为数据共享前的合规检查.

另有部分研究将数据分析程序合规性检查过程结合区块链技术. 例如,Zhang等人[87]针对物联网场景的海量数据,利用智能合约实现自动化的程序代码合规

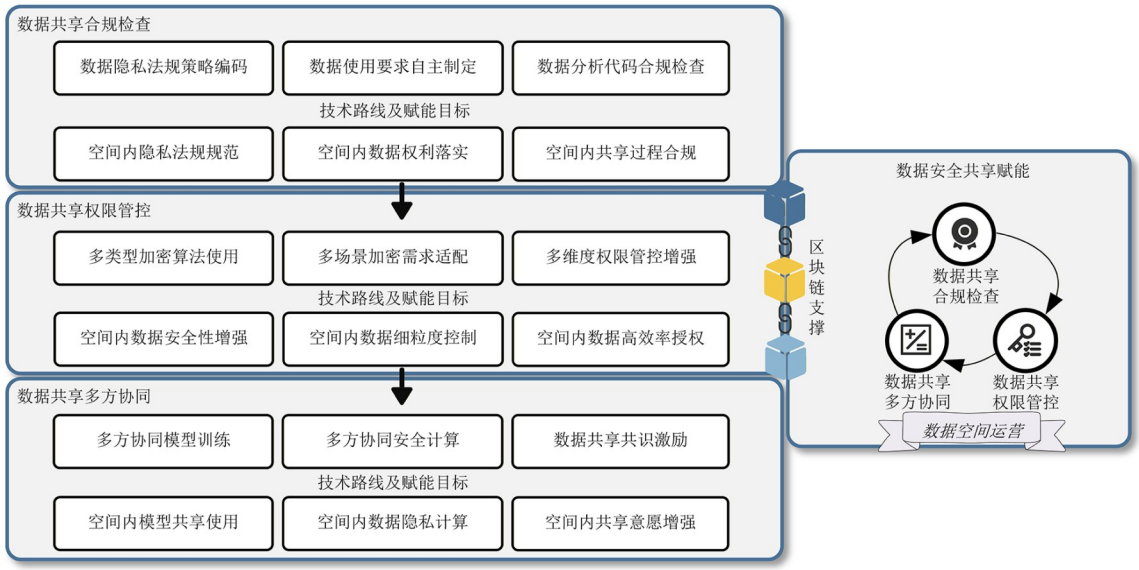

图 8　数据安全共享技术路线

性检查,并将智能合约关键逻辑运行于可信执行环境以增强该过程的可验证性. Lu 等人[88]通过数据封装将数据主体与程序的合规要求封装为规范数据结构并上链存储,数据分析使用前需要提交程序并通过智能合约执行检查,链上不可篡改数据提供了合规性检查凭证.

总的来说,可信数据空间作为数据基础设施,需要在数据共享前进行合规性检查,充分落实数据提供者权利. 现有技术方案转化数据政策法规要求并对常用代码库进行合规检查,为可信数据空间内数据权利的落实提供了有效技术思路. 数据共享合规检查阶段总结分析如表 10 所示.

表 10　数据共享合规检查总结分析

相关工作	核心方法	技术分析	适用场景	可信数据空间应用
文献[84]	结合数据政策法规的合规检查方案	通过将数据政策法规定义为标准策略格式,检查数据分析程序合规性	适用于需将政策法规转化为策略的场景,如将《中华人民共和国数据安全法》关于重要数据出境条款转化为系统标准策略	帮助空间内数据政策法规执行落实
文献[85]	基于程序静态分析的合规检查方案	通过对常用 Python 库函数进行分析,静态检查数据分析程序合规性	适用于数据分析脚本开发阶段的合规检查场景,如数据科学家在编写数据分析脚本时检查库函数使用合规性	进行空间内数据分析程序合规检查
文献[86]	面向 Web 开发人员的合规检查方案	通过转化程序开发标准为程序合规检查的输入,帮助开发人员检查代码合规性	适用于检查 Web 应用开发过程合规的场景,如检查开发人员编写的数据收集、存储和处理流程是否合规	帮助开发人员检查自身代码合规性
文献[87]	基于智能合约的合规检查方案	通过智能合约执行与可信执行环境安全增强,保证程序合规检查过程可验证	适用于要求高可信度和可验证性的业务场景,如金融数据、医疗数据等高风险场景下的合规检查	增强空间内合规检查过程安全性
文献[88]	基于链上封装存证的合规检查方案	通过数据封装机制将数据政策法规上链存储,提供程序合规检查存证	适用于需要审计追踪和不可否认证明的场景,如涉及法律纠纷的数据合规审计	提供空间内合规检查过程可信存证

5.2　数据共享权限管控

在数据共享过程中,需要对多类型、多来源数据进行权限管控. 加密技术与区块链技术的结合,能够在去中心化系统中建立细粒度访问控制机制,帮助解决数据安全问题. 可信数据空间需要结合多种加密技术并进行有效选择,为不同场景下数据共享提供数据权限管控.

现有研究主要结合属性基加密进行数据权限管控,利用区块链特性进行技术增强. 例如,Ren 等人[89]使用密文策略属性基加密,由区块链节点进行属性密钥管理和分发,并通过非交互式零知识证明与智能合约进行验证. 在此基础上,Xu 等人[90]采用遗传算法实现不同规模区块链网络的最优属性设置和策略动态生成. 杨小东等人[91]采用多属性授权机构联合分发密钥,有效抵抗用户和属性授权机构的合谋攻击,将原始数

据哈希和验证参数上传至区块链以验证完整性. 另外，Cheng等人[92]将属性基加密中的属性中心配置为区块链节点，设计了基于属性管理贡献、数据解密贡献的激励机制. Li等人[93]结合属性基加密与变色龙哈希函数赋予链上数据可控修改的特性. 另有部分研究将属性基加密与访问控制结合并建立完整的权限管控方案. 例如，Dai等人[94]提出了一种基于属性基加密的访问控制方案，将区块链与IPFS结合，设计一种交叉存储架构以释放链上存储压力，并利用属性基加密提供高灵活性与扩展性的数据访问控制. 类似地，Cui等人[95]结合属性基加密管控物联网数据全生命周期，建立了名为DSChain的区块链系统，在单个节点上实现每秒1 000次以上的交易. 这些方案能够为属性基加密进行多维度技术增强，促进在可信数据空间内使用属性基加密进行数据权限管控的实际应用.

除属性基加密外，代理重加密等不同加密技术同样能够赋能数据权限管控. 例如，Wang等人[96]针对物联网设备跨域数据共享困难问题，以区块链存储智能设备元数据，通过代理重加密的重加密密钥提高权限管控效率，实现不同数据域之间的跨域协同. 在此基础上，郭庆等人[97]通过区块链节点对代理重加密密钥进行分割管理，将国密算法SM2与代理重加密结合，并支持数据权限的有效更新. 另外，Liu等人[98]针对零信任场景，提出了一种基于明文可检查加密的数据权限管控方案，数据验证者能够快速判断给定密文是否由特定明文和公钥生成而无需解密. 其方案引入基于区块链的域委员会架构，解决了零信任场景下实体之间的信任问题，并利用分片技术并行处理请求以提高效率，能够帮助可信数据空间适配高权限管控性能需求的零信任架构.

总的来说，以上研究结合区块链去中心化、不可篡改等特性，将加密技术应用于物联网、零信任等实际应用场景实施有效数据权限管控. 其能够对模型训练、数据计算前的数据收集和执行后的数据结果进行有效权限控制，有效赋能可信数据空间内的数据共享. 数据共享权限管控阶段总结分析如表11所示.

表11 数据共享权限管控总结分析

相关工作	核心方法	技术分析	适用场景	可信数据空间应用
文献[89~91]	基于初始化增强属性基加密的数据权限管控	通过区块链特性增强密钥分发、属性设置、策略生成等初始化步骤,促进属性基加密应用	适用于需快速部署属性基加密的联盟链场景,如新增成员后初始化其属性密钥	增强空间内属性基加密应用的初始化过程
文献[92,93]	基于激励增强属性基加密的数据权限管控	通过激励机制提供权限管控正反馈,通过变色龙哈希提供可控修改,促进属性基加密应用	适用于需动态调整数据权限的场景,如根据贡献度调整各方权限的长期协作项目	增强空间内属性基加密与区块链技术结合
文献[94,95]	基于属性基加密与访问控制的数据权限管控	通过属性基加密提供数据细粒度访问控制,管控数据全生命周期	适用于需精确授权数据使用的场景,如医疗数据的细粒度权限配置	提供空间内基于属性基加密的数据细粒度权限管控
文献[96,97]	基于代理重加密的数据权限管控	通过代理重加密实现跨域数据权限管控,利用重加密密钥提高权限管控效率,支持权限更新	适用于数据外包与跨数据域的安全共享场景,如金融机构将加密数据交由云服务商分析	解决空间内跨域场景下数据权限协同问题
文献[98]	基于明文可检查加密的数据权限管控	通过明文可检查加密快速判断密文与明文间生成关系,结合分片技术实现高效权限判决	适用于海量加密数据快速检索与权限验证的场景,如网盘服务快速定位具有访问权限的目标文件	提供空间内高性能数据权限管控解决方案

5.3 数据共享多方协同

在合规检查与权限管控后，可信数据空间需要实现多方协同的数据共享使用. 隐私保护是多方数据协同的关键需求，可信数据空间需要以区块链为基础建立去中心化平台，进行数据模型的隐私训练与数据计算的隐私执行，实现数据价值的有效释放. 联邦学习是实现可信数据空间内数据共享多方协同的生动实例，其通过多个设备协同构建机器学习模型，同时防止数据泄露. 将联邦学习与区块链技术结合，可以实现模型的分布式训练与使用，通过智能合约执行模型更新与聚合过程，并在链上记录模型参数以供审计溯源. 图9展示了基于区块链的联邦学习模型训练技术思路.

众多研究人员将联邦学习与区块链技术结合，建立激励机制以促进模型训练. 例如，Cai等人[99]结合贝叶斯博弈进行激励驱动，鼓励区块链中诚实节点提供有效模型并阻止恶意训练干扰. Luo等人[100]结合泊松博弈模型解决参与者数量不确定问题，并尝试利用智能合约进行贡献计算和收入分配. Zhao等人[101]提出了一种创新的共识算法，尝试提供长期激励保证. 还有部分研究尝试解决联邦学习与区块链结合的效率问题. 例如，Yuan等人[102]将区块链网络层拆分为多个分片，有效减少了信息交换规模. Xu等人[103]将异步聚合方案与基于符号的梯度压缩结合，提高通信和聚合效率.

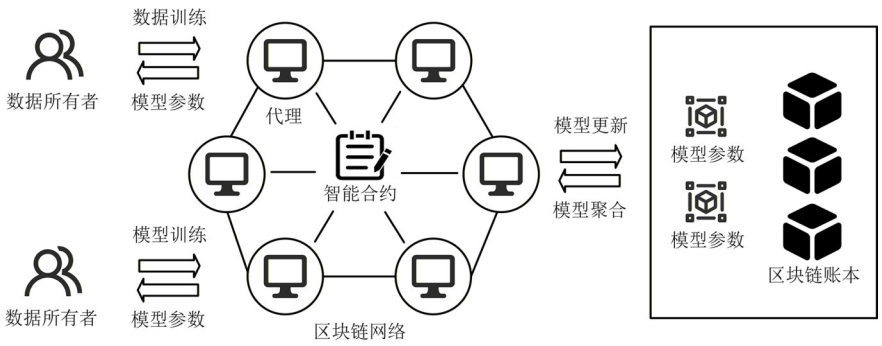

图9　基于区块链的联邦学习模型训练技术思路

Witt等人[104]结合联邦知识蒸馏将模型预测压缩为单比特表示以增强可扩展性. 另外,众多研究人员针对联邦学习与区块链结合过程的安全性开展研究. 例如,Singh等人[105]在方案中集成差分隐私,通过在更新模型权重前添加噪声来防止信息泄露. Wang等人[106]将联邦学习聚合计算加载至可信执行环境,以确保聚合结果的真实性. Yang等人[107]采用多密钥同态加密确保数据、标签和训练模型的机密性. 以上方案从不同角度尝试解决当前联邦学习与区块链耦合的各类问题,但可信数据空间内的实际部署仍面临多重瓶颈. 例如,区块链网络中的模型更新、模型聚合及梯度验证均需多次节点广播过程,与联邦学习显著的通信开销叠加,形成双重负担. 尽管引入分片技术并行处理以提升吞吐量,并通过异步机制缓解延迟,但这些优化可能以牺牲共识过程为代价,仍是实际应用中不可避免的难题. 另外,不同行业数据空间场景具有不同的侧重性问题. 对于医疗场景下的数据空间,数据异质性,即各医院病历格式与分布差异显著;对于金融场景下的数据空间,数据投毒与隐私泄露风险更高,且数据漂移带来的统计变化对模型学习与自适应能力提出了更高要求. 对于可信数据空间内的实际部署,仍需针对不同场景设计差异化模型聚合策略.

相比于联邦学习实现数据模型的训练共享,安全多方计算以秘密共享等技术为支撑,更侧重于数据隐私计算与数据交易保护,确保各方数据在计算过程中始终保持隐私性. 众多研究人员将安全多方计算与区块链技术结合,针对数据共享各阶段提供隐私保护. 例如,Ismayilov等人[108]提出了一种数据聚合协议,结合非对称加密与零知识证明等多个密码原语为数据聚合过程提供隐私保护. Zhang等人[109]利用基于牛顿插值公式的秘密共享安全地恢复管控数据权限的访问控制策略,实现多方协同的细粒度数据治理. Pei等人[110]基于加法同态加密验证数据共享结束后生成的Pedersen承诺,并通过相应智能合约逻辑执行. Yang等人[111]提出了一种非交互式零知识证明委托框架,通过将证明计算过程委托给多个工作程序来优化安全多方计算过程. 还有部分研究通过安全多方计算赋能区块链中的数据交易. 例如,Huang等人[112]针对以隐私为中心的门罗币,使用多个输入和输出地址为数据交易提供隐私保护. Abla等人[113]结合同态加密设计了一种安全多方计算协议,要求数据交易方提供数据有效性证明. 针对跨链数据交易,Han等人[114]通过安全多方计算让各参与方秘密共享数据资产地址的公钥,指导其创建相同额度的链内交易,防止观察者根据资产地址和交易额度攻击数据交易隐私. 这些方案能够提供数据共享隐私保护,使多参与方在不信任的环境中执行计算任务而无需公开其私有输入和输出.

总的来说,不同于数据权限管控后以数据传输进行数据共享,现有研究同样将区块链去中心化特性与联邦学习、安全多方计算结合,实现数据可用不可见的多方协同数据共享,有效促进可信数据空间内的数据价值释放. 数据共享多方协同阶段总结分析如表12所示.

6　数据联合溯源技术

可信数据空间内的数据流通需要全流程数据溯源,发现潜在风险并及时处置. 传统的数据溯源通常是集中式存储,存在中心化依赖问题,区块链去中心化与可溯源特性能够有效赋能可信数据空间内数据溯源. 假设空间运营者发起空间内恶意行为的追踪溯源,如何结合区块链技术提供规范的溯源流程是当前亟需解决的难题. 基于此,本节提出如图10所示的包含数据溯源标准模型、数据溯源架构设计与数据溯源分析方法3个模块的技术路线,总结分析如何通过现有研究有效赋能可信数据空间,实现数据联合溯源.

6.1　数据溯源标准模型

数据溯源标准模型是数据联合溯源的基础,通过标准溯源模型的使用,能够更准确直观地展示溯源结构与溯源过程,为可信数据空间内的数据溯源提供多维度支撑.

如何开发一个通用的领域无关模型来表示溯源信

表 12 数据共享多方协同总结分析

相关工作	核心方法	技术分析	适用场景	可信数据空间应用
文献[99~101]	基于激励增强联邦学习的多方协同数据训练	通过区块链多方共识与激励机制实现去中心化联邦学习,促进多方参与模型协同训练	适用于需按数据贡献度量化激励的场景,如各银行间按贡献分配收益	提供空间内多方协同模型训练激励
文献[102~104]	基于轻量级联邦学习的多方协同数据训练	通过分片、异步等技术提升联邦学习与区块链结合效率,提升多方协同模型训练效率	适用于参与方对训练实时性要求高的场景,如跨医疗机构快速训练疾病预测模型	提升空间内多方协同模型训练效率
文献[105~107]	基于安全增强联邦学习的多方协同数据训练	通过差分隐私、可信执行环境、同态加密等技术,保证多方协同模型训练过程安全	适用于处理高度敏感数据的场景,如政府部门保证公民隐私下进行反诈分析	增强空间内多方协同模型训练安全
文献[108~111]	基于安全多方计算的多方协同数据计算	通过安全多方计算保护数据聚合等数据共享具体阶段隐私	适用于需聚合多方数据且不泄露原始数据的场景,如企业间联合计算行业平均薪资	实现空间内多方协同隐私计算
文献[112~114]	基于安全多方计算的多方协同数据交易	通过安全多方计算为数据交易提供隐私保护	适用于实现数据交易隐私保护的场景,如提供数据授权后使用过程隐私验证	实现空间内多方协同数据交易

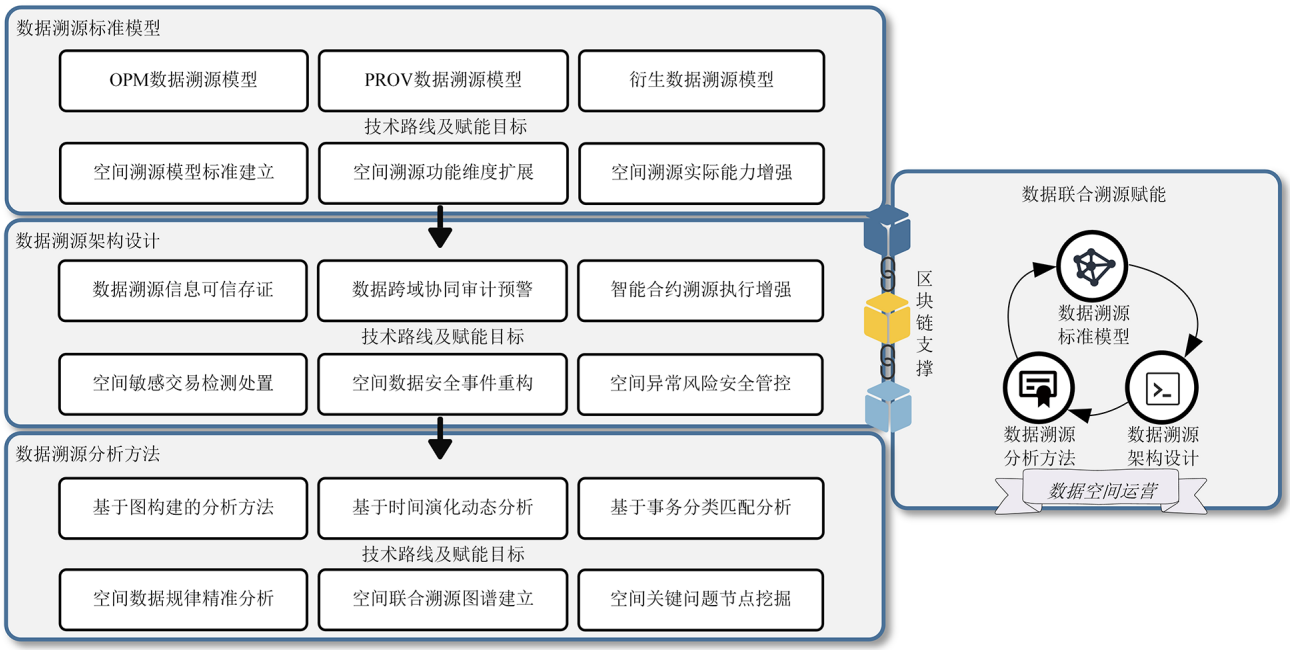

图 10 数据联合溯源技术路线

息是研究界的持续工作. 2006 年国际溯源与标注会议上的研究人员开始关注数据溯源问题. 与会者提出了形式化开放来源模型(Open Provenance Model, OPM),用于在系统之间交换信息[115]. 随后,研究人员引入了新的构造与关系,对 OPM 模型进行扩展提出了 PROV 模型[116],旨在描述网络应用中的数据表示、查询和交换. 以其为基础的模型在不同领域均有应用与扩展. 例如,ProvONE 模型[117]扩展 PROV 模型以支持 DataOne 科学社区,支持对静态结构与执行过程溯源,旨在提供工作流计算过程的更多信息. 全国信息技术标准化技术委员会提出 ProVOC 模型[118],旨在规范数据在采集、发布、分析和处理过程中的设计与应用.

总的来说,这些溯源模型能够提供实用功能支撑. 区块链技术能够结合以上模型,利用链上数据实现可信数据空间内数据对象历史和演变的不可篡改存储,通过涉及实体、用户和进程的基本元数据与链下数据建立映射,为可信数据空间内多维度数据溯源提供有力支撑.

6.2 数据溯源架构设计

区块链不可篡改特性提供了溯源数据的完整性保证,将关键数据存储上链,并与链下数据分析过程建立映射关系,是当前的主流研究思路. 图 11 展示了基于区块链的数据溯源架构设计.

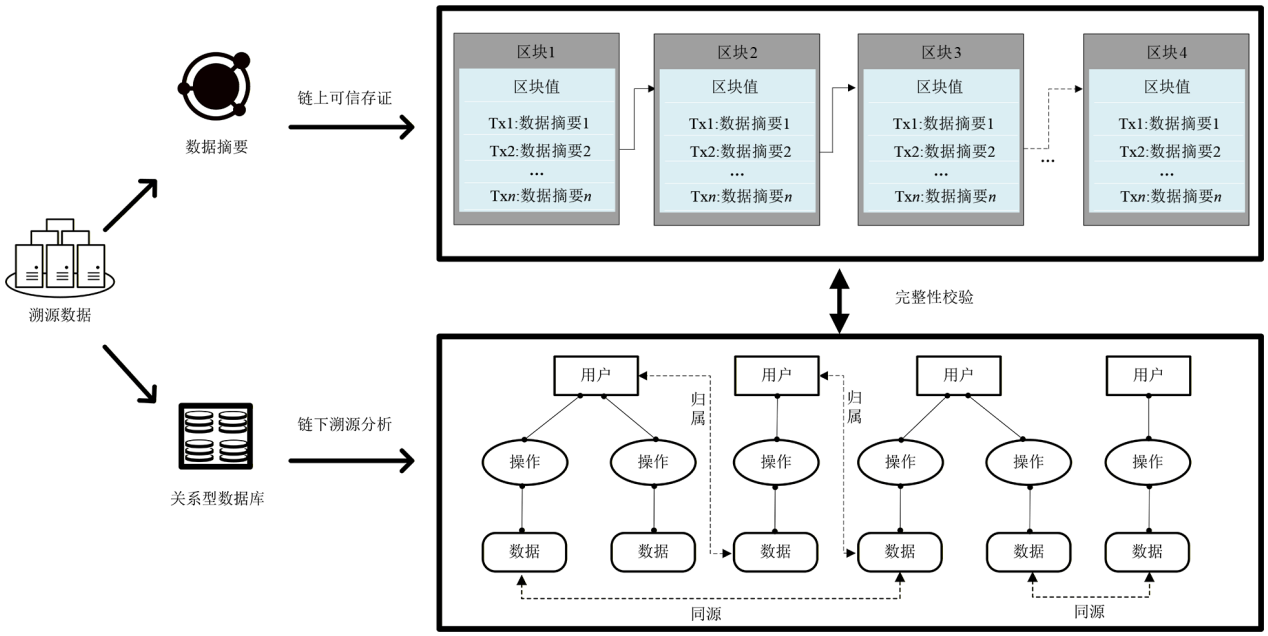

图11　基于区块链的数据溯源架构设计

众多研究人员进行了针对性研究,Liang等人[119]针对云数据提出名为ProvChain的数据溯源架构,本地存储完整溯源数据,链上存储不可篡改的元数据验证溯源有效性. 与ProvChain类似,Ruan等人[120]提出名为LineageChain的数据溯源架构,通过智能合约提供灵活的溯源数据访问与获取,溯源操作同样形成日志记录并上传至链上存储. 在此基础上,Ramachandran等人[121]提出名为SmartProvenance的数据溯源架构,通过智能合约逻辑实现数据有效性验证,防止外部环境与用户之间的潜在勾结. 这些方案在链上存储不可篡改数据,使用智能合约执行溯源过程,其技术思路能够启发可信数据空间建设.

另有部分研究针对数据交易溯源进行架构设计. 例如,Gong等人[122]提出一种基于群签名的可溯源数据交易架构,在确保匿名性的同时提供数据交易者身份溯源. Geng等人[123]使用图注意力机制识别交易者,通过构建三元组数据结构记录细粒度数据交易信息,提升了现有区块链数据交易溯源架构的精确度. 针对跨链数据交换过程中的溯源问题,Guo等人[124]提出了兼顾隐私保护与溯源功能的跨链协议. 实验表明,其有效解决了多条链间的数据孤岛问题,同时显著提升了溯源效率. 另外,部分研究结合特定领域进行数据溯源架构设计. 例如,Zeng等人[125]和Porkodi等人[126]分别针对无线传感器网络数据与物联网设备数据提出基于区块链的数据溯源架构,为可信数据空间应用提供有力支撑.

总的来说,现有的数据溯源架构设计主要结合区块链不可篡改特性,不同场景下的溯源数据均能够通过链上存储保证完整性,并通过智能合约逻辑执行溯源流程,为可信数据空间全流程溯源建设提供有力支撑. 数据溯源架构设计总结分析如表13所示.

6.3　数据溯源分析方法

可信数据空间内数据溯源能够使用标准溯源模型增加溯源维度,设计数据溯源架构提供全流程支撑,溯源过程同样需要数据溯源分析方法赋能. 基于区块链的数据溯源需要结合链上数据特点进行分析,通过数据溯源分析方法精准定位链上异常记录,为可信数据空间内数据溯源提供高精度支撑.

链上数据的图构建与表示能够通过计算分析节点特征与联动关系,探索链上数据的隐含属性. 例如,Liu等人[127]针对区块链中的异常检测问题,提出了一种基于图注意力机制的异常检测方案. 其采用动态属性图网络对每笔交易进行建模,充分提取链上数据深层次特征,动态更新不同时间戳图节点的学习权重. 在此基础上,Qi等人[128]进行该类方案的分析总结,将链上数据图构建过程分为基于交易和基于账户两种,为赋能可信数据空间内链上数据分析提供了技术方案总结.

链上数据分析同样需要具备处理复杂数据能力,Song等人[129]针对链上大量交易数据,通过构建累积网络和时间切片,动态分析交易随时间的变化规律,并总结交易量和交易关系的特定分布,在此基础上预测区块链发展进度. 类似地,Liang等人[130]提出了一种数据驱动的交易数据分析方案,自动从交易记录中提取相关特征,应用节点嵌入技术将交易信息映射到向量空间并生成交易表示,实现复杂数据的简化处理. 这些方

表 13 数据溯源架构设计总结分析

相关工作	架构特点	技术分析	适用场景	可信数据空间应用
文献[119~121]	结合区块链进行全流程溯源架构设计	通过链上存储维护溯源数据,结合智能合约执行的溯源架构	适用于需要全生命周期追溯的场景,如药品溯源、高端制造品质量追溯	帮助建设空间内数据全流程溯源架构
文献[122]	结合群签名的数据交易溯源架构设计	通过群签名提供匿名性,并提供支持数据交易溯源的溯源架构	适用于需保护交易方隐私且可审计的交易场景,如竞标报价、敏感数据交易	提供空间内数据交易匿名性与可溯源性
文献[123]	结合图注意力机制的溯源架构设计	通过构建三元组数据结构记录细粒度数据交易信息,具备细粒度与高精度的溯源架构	适用于需分析复杂数据路径与依赖关系的场景,如数据产品版权溯源、学术数据滥用追溯	提升空间内数据溯源架构粒度与精度
文献[124]	结合跨链协议的溯源架构设计	通过兼顾隐私保护与溯源功能的跨链协议,增强溯源架构在跨链数据交换场景适用性	适用于跨链或跨多个数据空间的数据交换场景,如跨境供应链溯源、多联盟链协同审计	针对空间内多链协同场景提供溯源架构
文献[125,126]	结合领域数据特点的溯源架构设计	通过结合无线传感器网络与物联网等领域数据特点,针对性提供溯源架构	适用于传感器网络、物联网等特定领域数据溯源,如物联网设备数据采集与全流程溯源	针对不同数据空间数据特点提供溯源架构

案均能够有效支撑可信数据空间内复杂链上数据的溯源分析.

链上数据的时间索引能够更好地实现空间内溯源图谱构建与安全事件还原. Zhang 等人[131]针对链上数据进行事务分类匹配并构建了历史数据的时间索引,为可信数据空间内基于时序的链上数据分析提供了技术思路. 另外,链上数据分析仍缺少高质量数据集,Zheng 等人[132]及时发现了区块链应用产生的海量数据所带来的数据分析技术挑战. 他们收集并处理链上数据,构建包含区块、合约、账户创建等数据组成部分的高质量数据集,填补了该领域的空白.

总的来说,现有的链上数据溯源分析方法以常用的数据分析技术为支撑,能够对链上存储的溯源数据进行高精度分析并尝试挖掘潜在风险,为可信数据空间内高精度数据溯源提供有力支撑. 数据溯源分析方法总结分析如表 14 所示.

表 14 数据溯源分析方法总结分析

相关工作	核心方法	技术分析	适用场景	可信数据空间应用
文献[127,128]	基于图构建与表示的数据溯源分析方法	通过动态属性图网络对每笔交易进行建模,充分提取链上数据深层次特征	适用于需揭示复杂交易模式和隐藏关联的场景,如追踪虚拟货币流向	帮助构建并提取空间内链上数据深层次特征
文献[129,130]	基于复杂交易分析的数据溯源分析方法	通过时间切片、向量映射等方式针对链上复杂数据交易分析	适用于分析跨合约、账户等复杂交互行为的场景,如追踪数字资产流转链条	针对空间内复杂链上数据交易进行分析
文献[131]	基于时间索引构建的数据溯源分析方法	通过链上数据事务分类并构建历史数据时间索引,提供基于时序的数据溯源分析	适用于需按时间线还原行为模式的场景,如审计合约调用历史与周期性行为	实现空间内基于时序的链上数据分类
文献[132]	链上数据溯源分析高质量数据集构建	通过区块、合约、账户创建等数据组成部分收集建立高质量链上数据溯源分析数据集	适用于需标准化数据集以评估不同溯源算法的场景,如溯源工具的基准测试	提供空间内数据溯源效果量化的高质量链上数据集

7 总结与展望

可信数据空间作为数据基础设施,存在空间运营、空间监测、数据服务等第三方功能模块,但为适应多方接入、跨域协同、数据交互等亟需解决的实际需求,需要耦合区块链技术以解决分布式信任问题. 随着可信数据空间战略地位的不断提升,区块链技术具有了更广阔的应用前景. 其不同于传统数据库的数据存储,而是通过不可篡改的数据结构与多方共识的信任建立为空间内数据提供者与数据使用者建立信任纽带. 然而,针对不同实际应用场景,地方、行业、企业等不同级别的可信数据空间同样对区块链赋能过程的功能、性能等方面树立了更高的要求. 本节基于对数据获取、数据验证、数据共享、数据溯源 4 个阶段研究的总结分析,以本文提出的基于区块链的可信数据空间安全技术框架

为主线,从数据全生命周期角度对区块链赋能未来可信数据空间的落地应用进行展望,主要体现在以下4个方面.

(1)当前,区块链技术对于可信数据空间数据获取阶段的支撑需要持续突破链上数据检索技术.区块链技术能够赋能可信数据空间实现数据可信获取,但链上数据存储具有资源有限、公开可见等特点,现有链上数据检索方案仍难以满足区块链作为可信数据空间基础设施日益提升的功能需求.当前链上数据检索面临三大难题,即如何实现高效的数据检索、如何提供检索过程中的隐私保护、如何提升检索结果的精确度.由于可信数据空间并非完全去中心化,未来工作中建立高效可信的数据检索层对于提升检索效率与可扩展性具有重要意义.在此基础上,需要精准建立链上数据对链下多类型数据资源的刻画关系,在保证检索精度下提供检索过程隐私保护,实现可信数据空间内的数据可信获取.

(2)当前,区块链技术对于可信数据空间数据验证阶段的支撑需要重点关注数据权属保护.可信数据空间是多参与方联合的数据基础设施,各方数据权属的落实至关重要.数据发布前的权属信息水印嵌入增强了数据全生命周期验证的可行性.现有方案在链下数据嵌入权属信息,关键信息存储至链上以保证其完整性.然而,现有的链上水印信息存储仍缺少针对性设计,在存储开销与协同验证上仍有优化空间.综合考虑水印信息的二进制存储特点与链上区块数据结构设计,能够赋能可信数据空间链上索引功能,进一步发掘链上索引可能性并促进链上链下协同,具有一定的潜在研究意义.未来工作需要在此基础上实现数据安全标签、安全策略、安全等级等更完备信息的链上嵌入,建立链上不可篡改的水印信息链条,对数据全生命周期进行更细粒度的验证与追责,实现可信数据空间内的数据合规验证.

(3)当前,区块链技术对于可信数据空间数据共享阶段的支撑需要有效落实数据政策法规.现有数据共享过程的权限管控主要从整体角度提升系统管控能力,并没有将数据提供者的数据权利按照政策法规有效落实.随着数据价值释放需求增加,数据权限管控已由静态存储与加密保护转变为动态共享与合规使用.如何在可信数据空间内提供数据权限控制,按照数据政策法规要求释放数据价值,已经成为未来研究的重点.未来工作需要转变传统数据共享思路,将数据政策法规转化为直观简洁表达,帮助数据提供者理解并制定数据使用要求上传至可信数据空间链上存储,形成不可篡改凭证.随后,通过智能合约或可信执行环境进行数据分析程序的合规性检查,保证数据分析程序严格按照数据政策法规执行,实现可信数据空间内的数据安全共享.

(4)当前,区块链技术对于可信数据空间数据溯源阶段的支撑需要继续深入链上链下协同.现有数据溯源方案基于各类溯源标准模型,为溯源数据提供直观的表示方式,但仍无法以数据为主体进行细粒度、多功能的溯源分析.其能够根据链上存储的日志信息进行分析取证,但无法解决数据位置、数据变化、变化原因等数据共享过程的动态细节.并且,基于智能合约的链上溯源分析需考虑共识过程的时间开销.未来工作应将关键信息存证链上,大规模数据存储链下,建立链上链下协同机制,并通过聚类等数据分析技术进行可信数据空间内的异常行为发现,通过时空数据交叉验证等技术针对空间内不可信来源、不可靠环境的产生进行安全事件重构,动态呈现数据在空间内的流转过程与风险源头,实现可信数据空间内的数据联合溯源.

8　结束语

近年来,区块链技术已经成为支撑数据流通的关键基础设施,其为可信数据空间安全运营与价值释放提供了理想土壤.二者的结合不仅在技术层面充分适配,更已经上升至国家战略高度.本文致力于分析总结区块链技术在可信数据空间中的赋能作用,通过对可信数据空间架构与需求的分析,明确区块链技术的核心作用与功能优势,并将其与主流安全机制结合,构建了贯穿数据获取、数据验证、数据共享、数据溯源全生命周期的赋能框架,以期对可信数据空间发展提供启发与借鉴,促进其建设与应用.

参考文献

[1] 中华人民共和国国家数据局.可信数据空间发展行动计划(2024—2028年)[EB/OL]. (2024-11-21)[2025-09-18]. https://www.gov.cn/zhengce/zhengceku/202411/content_6996363.htm.
National Data Administration of the People's Republic of China. Trusted data space development action plan (2024—2028)[EB/OL]. (2024-11-21)[2025-09-18]. https://www.gov.cn/zhengce/zhengceku/202411/content_6996363. htm. (in Chinese)

[2] 奇安信.2024中国政企机构数据安全风险研究报告[EB/OL]. (2025-03-11)[2025-09-18]. https://www.qianxin.com/threat/reportdetail?report_id=336.
QI A X. Technology Group Inc. 2024 China government and enterprise data security risk research report[EB/OL]. (2025-03-11) [2025-09-18]. https://www. qianxin. com/threat/reportdetail report_id=336. (in Chinese)

[3] 王利朋, 关志, 李青山, 等. 区块链数据安全服务综述[J]. 软件学报, 2023, 34(1): 1-32.
WANG L P, GUAN Z, LI Q S, et al. Survey on blockchain-based security services[J]. Journal of Software, 2023, 34(1): 1-32. (in Chinese)

[4] XU Y Q, XU G X, LIU Y, et al. A survey of the fusion of traditional data security technology and blockchain[J]. Expert Systems with Applications, 2024, 252: 124151.

[5] Decentralized Identifier Working Group. Decentralized identifiers(DIDs) v1.0[EB/OL]. (2022-07-19)[2025-09-18]. https://www.w3.org/TR/2022/REC-did-core-20220719/.

[6] Verifiable Credentials Working Group. Verifiable credentials data model v2.0[EB/OL]. (2025-05-15)[2025-09-18]. https://www.w3.org/TR/2025/REC-vc-data-model-2.0- 20250515/.

[7] 贺海武, 延安, 陈泽华. 基于区块链的智能合约技术与应用综述[J]. 计算机研究与发展, 2018, 55(11): 2452-2466.
HE H W, YAN A, CHEN Z H. Survey of smart contract technology and application based on blockchain[J]. Journal of Computer Research and Development, 2018, 55(11): 2452-2466. (in Chinese)

[8] 刘懿中, 刘建伟, 张宗洋, 等. 区块链共识机制研究综述[J]. 密码学报, 2019, 6(4): 395-432.
LIU Y Z, LIU J W, ZHANG Z Y, et al. Overview on blockchain consensus mechanisms[J]. Journal of Cryptologic Reseatch, 2019, 6(4): 395-432. (in Chinese)

[9] 司冰茹, 肖江, 刘存扬, 等. 区块链网络综述[J]. 软件学报, 2024, 35(2): 773-799.
SI B R, XIAO J, LIU C Y, et al. Survey on blockchain network[J]. Journal of Software, 2024, 35(2): 773-799. (in Chinese)

[10] 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.
LIU A D, DU X H, WANG N, et al. Research progress of blockchain technology and its application in information security[J]. Journal of Software, 2018, 29(7): 2092-2115. (in Chinese)

[11] 刘敖迪, 杜学绘, 王娜, 等. 区块链系统安全防护技术研究进展[J]. 计算机学报, 2024, 47(3): 608-646.
LIU A D, DU X H, WANG N, et al. Research progress on blockchain system security technology[J]. Chinese Journal of Computers, 2024, 47(3): 608-646. (in Chinese)

[12] 梁秀波, 吴俊涵, 赵昱, 等. 区块链数据安全管理和隐私保护技术研究综述[J]. 浙江大学学报(工学版), 2022, 56(1): 1-15.
LIANG X B, WU J H, ZHAO Y, et al. Review of blockchain data security management and privacy protection technology research[J]. Journal of Zhejiang University (Engineering Science), 2022, 56(1): 1-15. (in Chinese)

[13] 黄华威, 孔伟, 彭肖文, 等. 区块链分片技术综述[J]. 计算机工程, 2022, 48(6): 1-10.
HUANG H W, KONG W, PENG X W, et al. Survey on blockchain sharding technology[J]. Computer Engineering, 2022, 48(6): 1-10. (in Chinese)

[14] 李芳, 李卓然, 赵赫. 区块链跨链技术进展研究[J]. 软件学报, 2019, 30(6): 1649-1660.
LI F, LI Z R, ZHAO H. Research on the progress in cross-chain technology of blockchains[J]. Journal of Software, 2019, 30(6): 1649-1660. (in Chinese)

[15] 白金龙, 曹利峰, 万季玲, 等. 区块链隐私保护技术研究进展[J]. 计算机工程与应用, 2025, 61(2): 19-36.
BAI J L, CAO L F, WAN J L, et al. Research progress of blockchain privacy protection technology[J]. Computer Engineering and Applications, 2025, 61(2): 19-36. (in Chinese)

[16] International Data Spaces Association. International trusted data space architecture[EB/OL]. (2020-05-07) [2025-09-18]. https://internationaldataspaces.org/.

[17] European Union. Gaia-X: A federated secure data infrastructure [EB/OL]. (2022-03-25)[2025-09-18]. https://gaia-x.eu/.

[18] Niemopen. National information exchange model[EB/OL]. (2013-11) [2025-09-18]. https://niem. github. io/niem-releases/.

[19] Information-Technology Promotion Agency. Whitepaper: Ouranos ecosystem dataspaces reference architecture model[EB/OL]. (2025-03-31)[2025-09-18]. https://www.ipa.go.jp/en/digital/architecture-guidelines/ouranos-ecosystem-da - taspaces-ram-white-paper.html.

[20] Ministry of Economy. Trade and industry in Japan connected industries open framework[EB/OL]. (2019-02-15)[2025-09-18]. https://community.ciof-ivi.com/.

[21] 中国信通院. 可信工业数据空间系统架构 1.0 白皮书[EB/OL]. (2022-01)[2025-09-18]. https://www.caict.ac.cn/kxyj/qwfb/ztbg/202201/t20220125_396156.htm.
China Academy of Information and Communications Technology. Trustworthy industrial data space system architecture 1.0 white paper[EB/OL]. (2022-01)[2025-09-18]. https://www.caict.ac.cn/kxyj/qwfb/ztbg/202201/t20220125_396156.htm. (in Chinese)

[22] IEEE Draft Standard for Trusted Data Matrix System Architecture: IEEE P3158/D2 2024[S/OL]. [2025-09-18]. https://ieeexplore.ieee.org/document/10477330.

[23] 中华人民共和国中央人民政府. 中华人民共和国数据安全

法[EB/OL]. (2021-06-10)[2025-09-18]. http://www.npc.gov.cn/c2/c30834/202106/t20210610_311888.html.
The Central People's Government of the People's Republic of China. Data security law of China[EB/OL]. (2021-06-10)[2025-09-18]. http://www.npc.gov.cn/c2/c30834/202106/t20210610_311888.html. (in Chinese)

[24] Office of Science and Technology Policy. National strategy to advance privacy-preserving data sharing and analytics [EB/OL]. (2023-03-31)[2025-09-18]. https://www.nitrd.gov/pubs/National-Strategy-to-Advance-Privacy-Preserving-Data-Sharing-and-Analytics.pdf.

[25] European Union. General data protection regulation[EB/OL]. (2018-05-23)[2025-09-18]. https://gdpr-info.eu/.

[26] ANDROULAKI E, BARGER A, BORTNIKOV V, et al. Hyperledger fabric: A distributed operating system for permissioned blockchains[C]//Proceedings of the Thirteenth EuroSys Conference. New York: ACM, 2018: 1-15.

[27] LI H Z, CHEN Y J, SHI X, et al. FISCO-BCOS: An enterprise-grade permissioned blockchain system with high-performance[C]//Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis. New York: ACM, 2023: 1-17.

[28] WOOD G. Ethereum: A secure decentralized generalized transaction ledger[J]. Ethereum Project Yellow Paper, 2014, 151(2014): 1-32.

[29] 趣链科技. 趣链区块链平台技术文档[EB/OL]. (2019-05-24)[2025-09-18]. https://hyperchain.readthedocs.io/zh-cn/latest/.
Hyperchain. Hyperchain blockchain platform technical documentation[EB/OL]. (2019-05-24)[2025-09-18]. https://hyperchain.readthedocs.io/zh-cn/latest/. (in Chinese)

[30] 蚂蚁集团. 蚂蚁链BaaS平台[EB/OL]. (2018-06-01)[2025-09-18]. https://antdigital.com/products/baas.
Ant Technology Group Co., Ltd. AntChain blockchain as a service platform[EB/OL]. (2018-06-01)[2025-09-18]. https://antdigital.com/products/baas. (in Chinese)

[31] 北京微芯区块链与边缘计算研究院. 长安链平台技术文档[EB/OL]. (2025-07-01)[2025-09-18]. https://docs.chainmaker.org.cn/v2.3.7/html/index.html.
Beijing Academy of Blockchain and Edge Computing. ChainMaker platform technical documentation[EB/OL]. (2025-07-01)[2025-09-18]. https://docs.chainmaker.org.cn/v2.3.7/html/index.html. (in Chinese)

[32] ABUHASHIM A, TAN C C. Smart contract designs on blockchain applications[C]//2020 IEEE Symposium on Computers and Communications. Piscataway: IEEE, 2020: 1-4.

[33] CHISHTI M S, SUFYAN F, BANERJEE A. Decentralized on-chain data access via smart contracts in ethereum blockchain[J]. IEEE Transactions on Network and Service Management, 2022, 19(1): 174-187.

[34] SHANG S Y, DU X H, WANG X H, et al. Private approximate nearest neighbor search for on-chain data based on locality-sensitive hashing[J]. Future Generation Computer Systems, 2025, 164: 107586.

[35] LI Y, ZHENG K, YAN Y, et al. EtherQL: A query layer for blockchain system[C]//Database Systems for Advanced Applications. Cham: Springer, 2017: 556-567.

[36] WANG S, DINH T T A, LIN Q, et al. Forkbase: An efficient storage engine for blockchain and forkable applications[J]. Proceedings of the VLDB Endowment, 2018, 11(10): 1137-1150.

[37] LIU P, XIAN Y Q, YAO C J, et al. A trustworthy and consistent blockchain oracle scheme for industrial Internet of Things[J]. IEEE Transactions on Network and Service Management, 2024, 21(5): 5135-5148.

[38] YAO Z M, XIN J C, HAO K, et al. Learned-index-based semantic keyword query on blockchain[J]. Mathematics, 2023, 11(9): 2055.

[39] LIU M M, WANG H M, YANG F C. An efficient data query method of blockchain based on index[C]//2021 7th International Conference on Computer and Communications. Piscataway: IEEE, 2022: 1539-1544.

[40] YAO Z M, LI T Y, XIN J C, et al. VGQ: Enabling verifiable graph queries on blockchain systems[C]//2025 IEEE 41st International Conference on Data Engineering. Piscataway: IEEE, 2025: 3602-3614.

[41] CUI N N, WANG D, LI J X, et al. Enabling efficient, verifiable, and secure conjunctive keyword search in hybrid-storage blockchains[J]. IEEE Transactions on Knowledge and Data Engineering, 2024, 36(6): 2445-2460.

[42] HAN Y, HAN J G, MENG W Z, et al. Blockchain-based privacy-preserving public key searchable encryption with strong traceability[J]. Journal of Systems Architecture, 2024, 155: 103264.

[43] YU H T, LIU S H, CHEN L Q, et al. Blockchain-enabled one-to-many searchable encryption supporting designated server and multi-keywords for Cloud-IoMT[J]. Journal of Systems Architecture, 2024, 149: 103103.

[44] HUANG W H, CHEN Y, JING D, et al. A multicloud collaborative data security sharing scheme with blockchain indexing in industrial Internet environments[J]. IEEE Internet of Things Journal, 2024, 11(16): 27532-27544.

[45] GE L, JIANG T. A privacy protection method of lightweight nodes in blockchain[J]. Security and Communication Networks, 2021, 2021: 2067137.

[46] HOU J H, LIU D X, HUANG C, et al. Data protection: Privacy-preserving data collection with validation[J]. IEEE Transactions on Dependable and Secure Computing, 2024, 21(4): 3422-3438.

[47] XIE Y K, ZHANG C, WEI L B, et al. Private transaction retrieval for lightweight Bitcoin client[C]//2019 IEEE International Conference on Blockchain and Cryptocurrency. Piscataway: IEEE, 2019: 440-446.

[48] XU L, BAO T, ZHU L H. Blockchain empowered differentially private and auditable data publishing in industrial IoT[J]. IEEE Transactions on Industrial Informatics, 2021, 17(11): 7659-7668.

[49] ZHANG K, TSAI P W, TIAN J, et al. DPNM: A differential private notary mechanism for privacy preservation in cross-chain transactions[J]. IEEE Transactions on Information Forensics and Security, 2025, 20: 2224-2236.

[50] XU C, ZHANG C, XU J L. vChain: Enabling verifiable Boolean range queries over blockchain databases[C]//Proceedings of the 2019 International Conference on Management of Data. New York: ACM, 2019: 141-158.

[51] WANG H X, XU C, ZHANG C, et al. vChain+: Optimizing verifiable blockchain Boolean range queries[C]//2022 IEEE 38th International Conference on Data Engineering. Piscataway: IEEE, 2022: 1927-1940.

[52] XU H, XIAO B, LIU X L, et al. Empowering authenticated and efficient queries for STK transaction-based blockchains[J]. IEEE Transactions on Computers, 2023, 72(8): 2209-2223.

[53] WU H T, LI Z C, SONG R, et al. Enabling privacy-preserving and efficient authenticated graph queries on blockchain-assisted clouds[J]. IEEE Transactions on Knowledge and Data Engineering, 2023, 35(9): 9728-9742.

[54] WANG H X, XU C, CHEN X J, et al. V2FS: A verifiable virtual filesystem for multi-chain query authentication[C]//2024 IEEE 40th International Conference on Data Engineering. Piscataway: IEEE, 2024: 1999-2011.

[55] SUN W J, XU Z H, NI W Z, et al. Authenticated aggregate queries with Boolean range predicates on blockchains[J]. Proceedings of the VLDB Endowment, 2025, 18(10): 3615-3627.

[56] CAI C J, XU L, ZHOU A X, et al. Toward a secure, rich, and fair query service for light clients on public blockchains[J]. IEEE Transactions on Dependable and Secure Computing, 2022, 19(6): 3640-3655.

[57] ZHOU W C, CAI Y F, PENG Y Q, et al. VeriDB: An SGX-based verifiable database[C]//Proceedings of the 2021 International Conference on Management of Data. New York: ACM, 2021: 2182-2194.

[58] WU H T, PENG Z, GUO S T, et al. VQL: Efficient and verifiable cloud query services for blockchain systems[J]. IEEE Transactions on Parallel and Distributed Systems, 2022, 33(6): 1393-1406.

[59] JIANG S R, LIU J Q, CHEN J W, et al. Query integrity meets blockchain: A privacy-preserving verification framework for outsourced encrypted data[J]. IEEE Transactions on Services Computing, 2023, 16(3): 2100-2113.

[60] XU Y, ZHANG C, ZENG Q R, et al. Blockchain-enabled accountability mechanism against information leakage in vertical industry services[J]. IEEE Transactions on Network Science and Engineering, 2021, 8(2): 1202-1213.

[61] XIAO X L, ZHANG Y S, ZHU Y W, et al. FingerChain: Copyrighted multi-owner media sharing by introducing asymmetric fingerprinting into blockchain[J]. IEEE Transactions on Network and Service Management, 2023, 20(3): 2869-2885.

[62] WANG B W, SHI J W, WANG W S, et al. Image copyright protection based on blockchain and zero-watermark[J]. IEEE Transactions on Network Science and Engineering, 2022, 9(4): 2188-2199.

[63] LI M, SHEN Y Z, YE G X, et al. Anonymous, secure, traceable, and efficient decentralized digital forensics[J]. IEEE Transactions on Knowledge and Data Engineering, 2024, 36(5): 1874-1888.

[64] RIZZO S G, BERTINI F, MONTESI D. Content-preserving text watermarking through unicode homoglyph substitution[C]//Proceedings of the 20th International Database Engineering & Applications Symposium on - IDEAS'16. New York: ACM, 2016: 97-104.

[65] YOO K, AHN W, JANG J, et al. Robust multi-bit natural language watermarking through invariant features[C]//Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics. Stroudsburg: ACL, 2023: 2092-2115.

[66] GU C C, LI X L, LIANG P, et al. On the learnability of watermarks for language models[EB/OL]. (2024-05-02)[2025-11-10]. https://arxiv.org/abs/2312.04469.

[67] WANG L A, YANG W K, CHEN D L, et al. Towards codable watermarking for injecting multi-bits information to LLMs[EB/OL]. (2024-04-03)[2025-10-19]. https://arXiv.org/abs/2307.15992.

[68] LIU W Z, LI Y, LIN D D, et al. GROOT: Generating robust watermark for diffusion-model-based audio synthesis[C]// Proceedings of the 32nd ACM International Conference on Multimedia. New York: ACM, 2024: 3294-3302.

[69] WANG R X, ZHU Y J, XIA D X. Cascade ownership verification framework based on invisible watermark for model copyright protection[J]. Concurrency and Computation: Practice and Experience, 2025, 37(4-5): e8394.

[70] FEI S F, YAN Z, XIE H M, et al. Sec-E2E: End-to-end communication security in LS-HetNets based on blockchain[J]. IEEE Transactions on Network Science and Engineering, 2024, 11(1): 761-778.

[71] JIE W Q, QIU W J, VOUNDI KOE A S, et al. A secure and flexible blockchain-based offline payment protocol[J]. IEEE Transactions on Computers, 2024, 73(2): 408-421.

[72] PENG W, LI X, NIU J Y, et al. Ensuring state continuity for confidential computing: A blockchain-based approach[J]. IEEE Transactions on Dependable and Secure Computing, 2024, 21(6): 5635-5649.

[73] ZHAO L R, DECOUCHANT J, LIU J K, et al. Trusted hardware-assisted leaderless Byzantine fault tolerance consensus[J]. IEEE Transactions on Dependable and Secure Computing, 2024, 21(6): 5086-5097.

[74] GU C, MA B S, HU D H. A dependable and efficient decentralized trust management system based on consortium blockchain for intelligent transportation systems[J]. IEEE Transactions on Intelligent Transportation Systems, 2024, 25(12): 19430-19443.

[75] LI R J, WANG Q, LI Y Z, et al. Bringing smart contract confidentiality via trusted hardware: Fact and fiction[J]. IEEE Transactions on Information Forensics and Security, 2025, 20: 159-174.

[76] LU X, ZHANG Z J, MA T, et al. Trusted execution environment with rollback protection for smart contract-based IoT data trading[J]. IEEE Internet of Things Journal, 2024, 11(20): 32901-32909.

[77] LI Z C, XIAO B, GUO S T, et al. Securing deployed smart contracts and DeFi with distributed TEE cluster[J]. IEEE Transactions on Parallel and Distributed Systems, 2023, 34(3): 828-842.

[78] JIANG S Q, ALHADIDI D, KHOJIR H F. Key-and-signature compact multi-signatures for blockchain: A compiler with realizations[J]. IEEE Transactions on Dependable and Secure Computing, 2025, 22(1): 579-596.

[79] BOJIČ BURGOS J, PUSTIŠEK M. Decentralized IoT data authentication with signature aggregation[J]. Sensors, 2024, 24(3): 1037.

[80] XIE Y M, FAN Q, ZHANG C, et al. Accountable and secure threshold EdDSA signature and its applications[J]. IEEE Transactions on Information Forensics and Security, 2024, 19: 7033-7046.

[81] DUAN J K, ZHENG S H, WANG W, et al. Concise RingCT protocol based on linkable threshold ring signature[J]. IEEE Transactions on Dependable and Secure Computing, 2024, 21(5): 5014-5028.

[82] YU H F, LI B N, HUI W X. Signature-based anti-quantum schemes for blockchain-based donation and E-invoice[J]. IEEE Internet of Things Journal, 2024, 11(18): 30245-30259.

[83] WU X Y, DU X H, YANG Q T, et al. Dynamic fine-grained access control for smart contracts based on improved attribute-based signature[J]. The Journal of Supercomputing, 2024, 81(1): 44.

[84] WANG L, NEAR J P, SOMANI N, et al. Data capsule: A new paradigm for automatic compliance with data privacy regulations[C]//Heterogeneous Data Management, Polystores, and Analytics for Healthcare. Cham: Springer, 2019: 3-23.

[85] WANG L, KHAN U, NEAR J P, et al. PrivGuard: Privacy regulation compliance made easier[C]//31st USENIX Security Symposium. Berkeley: USENIX Association, 2022: 3753-3770.

[86] FERREIRA M, BRITO T, SANTOS J F, et al. RuleKeeper: GDPR-aware personal data compliance for web frameworks[C]//2023 IEEE Symposium on Security and Privacy. Piscataway: IEEE, 2023: 2817-2834.

[87] ZHANG Y X, YANG J C, LEI H, et al. PACTA: An IoT data privacy regulation compliance scheme using TEE and blockchain[J]. IEEE Internet of Things Journal, 2024, 11(5): 8882-8893.

[88] LU T P, ZHANG B S, REN K. PrivData network: A privacy-preserving on-chain data factory and trading market[J]. IEEE Transactions on Dependable and Secure Computing, 2024, 21(3): 1424-1436.

[89] REN Z X, YAN E H, CHEN T W, et al. Blockchain-based CP-ABE data sharing and privacy-preserving scheme using distributed KMS and zero-knowledge proof[J]. Journal of King Saud University - Computer and Information Sciences, 2024, 36(3): 101969.

[90] XU C H, QU Y Y, XIANG Y, et al. An optimized privacy-protected blockchain system for supply chain on Internet of Things[J]. IEEE Internet of Things Journal, 2024,

11(5): 9019-9030.

[91] 杨小东, 陈艾佳, 汪志松, 等. 基于区块链的多授权密文策略属性基等值测试加密方案[J]. 电子学报, 2024, 52(3): 898-908.
YANG X D, CHEN A J, WANG Z S, et al. Blockchain-based multi-authority ciphertext-policy attribute-based encryption scheme with equality test[J]. Acta Electronica Sinica, 2024, 52(3): 898-908. (in Chinese)

[92] CHENG H L, LO S L, LU J. A blockchain-enabled decentralized access control scheme using multi-authority attribute-based encryption for edge-assisted Internet of Things[J]. Internet of Things, 2024, 26: 101220.

[93] LI S Y, NIU K L, WU B. A blockchain-based secure data sharing scheme with efficient attribute revocation[J]. Journal of Systems Architecture, 2025, 159: 103309.

[94] DAI Y Y, WU J, MAO S Q, et al. Blockchain empowered access control for digital twin system with attribute-based encryption[J]. Future Generation Computer Systems, 2024, 160: 564-576.

[95] CUI J, LI Y T, ZHANG Q Y, et al. DSChain: A blockchain system for complete lifecycle security of data in Internet of Things[J]. IEEE Transactions on Dependable and Secure Computing, 2024, 21(4): 3977-3993.

[96] WANG F Q, CUI J, ZHANG Q Y, et al. Blockchain-based secure cross-domain data sharing for edge-assisted industrial Internet of Things[J]. IEEE Transactions on Information Forensics and Security, 2024, 19: 3892-3905.

[97] 郭庆, 田有亮, 万良. 基于代理重加密的区块链数据受控共享方案[J]. 电子学报, 2023, 51(2): 477-488.
GUO Q, TIAN Y L, WAN L. Blockchain data controlled sharing scheme based on proxy re-encryption[J]. Acta Electronica Sinica, 2023, 51(2): 477-488. (in Chinese)

[98] LIU Y Z, XING X X, TONG Z H, et al. Secure and scalable cross-domain data sharing in zero-trust cloud-edge-end environment based on sharding blockchain[J]. IEEE Transactions on Dependable and Secure Computing, 2024, 21(4): 2603-2618.

[99] CAI L Y, DAI Y Y, HU Q W, et al. Bayesian game-driven incentive mechanism for blockchain-enabled secure federated learning in 6G wireless networks[J]. IEEE Transactions on Network Science and Engineering, 2024, 11(5): 4951-4964.

[100] LUO M S, HE Y H, YUAN T L, et al. A Poisson game-based incentive mechanism for federated learning in web 3.0[J]. IEEE Transactions on Network Science and Engineering, 2024, 11(6): 5576-5588.

[101] ZHAO Y, QU Y Y, XIANG Y, et al. Long-term proof-of-contribution: An incentivized consensus algorithm for blockchain-enabled federated learning[J]. IEEE Transactions on Services Computing, 2024, 17(5): 2558-2570.

[102] YUAN S, CAO B, SUN Y, et al. Secure and efficient federated learning through layering and sharding blockchain[J]. IEEE Transactions on Network Science and Engineering, 2024, 11(3): 3120-3134.

[103] XU C H, GE J Q, DENG Y, et al. BASS: A blockchain-based asynchronous SignSGD architecture for efficient and secure federated learning[J]. IEEE Transactions on Dependable and Secure Computing, 2024, 21(6): 5388-5402.

[104] WITT L, ZAFAR U, SHEN K, et al. Decentralized and incentivized federated learning: A blockchain-enabled framework utilising compressed soft-labels and peer consistency[J]. IEEE Transactions on Services Computing, 2024, 17(4): 1449-1464.

[105] SINGH M B, SINGH H, PRATAP A. Energy-efficient and privacy-preserving blockchain based federated learning for smart healthcare system[J]. IEEE Transactions on Services Computing, 2024, 17(5): 2392-2403.

[106] WANG H, CAI Y C, WANG J, et al. Voltran: Unlocking trust and confidentiality in decentralized federated learning aggregation[J]. IEEE Transactions on Information Forensics and Security, 2024, 19: 9744-9759.

[107] YANG W T, WANG X D, GUAN Z T, et al. SecureSL: A privacy-preserving vertical cooperative learning scheme for web 3.0[J]. IEEE Transactions on Network Science and Engineering, 2024, 11(5): 3983-3994.

[108] ISMAYILOV G C, ÖZTURAN C. Trustless privacy-preserving data aggregation on Ethereum with hypercube network topology[J]. Computer Communications, 2025, 230: 108009.

[109] ZHANG C, ZHAO M Y, LIANG J W, et al. NANO: Cryptographic enforcement of readability and editability governance in blockchain databases[J]. IEEE Transactions on Dependable and Secure Computing, 2024, 21(4): 3439-3452.

[110] PEI H M, YANG P, DU M, et al. Blockchain-assisted verifiable secure multi-party data computing[J]. Computer Networks, 2024, 253: 110712.

[111] YANG Y B, CHENG Y J, WANG K L, et al. Siniel: Distributed privacy-preserving zkSNARK[EB/OL]. (2024-11-04)[2025-11-11]. https://eprint.iacr.org/2024/1803.

[112] HUANG K, MU Y, REZAEIBAGHA F, et al. Monero with multi-grained redaction[J]. IEEE Transactions on Dependable and Secure Computing, 2024, 21(1): 241-253.

[113] ABLA P, LI T T, HE D B, et al. Fair and privacy-preserved data trading protocol by exploiting blockchain[J]. IEEE Transactions on Information Forensics and Security, 2024, 19: 9012-9025.

[114] HAN P P, YAN Z, YANG L T, et al. P2C2T: Preserving the privacy of cross-chain transfer[C]//2025 IEEE Symposium on Security and Privacy. Piscataway: IEEE, 2025: 1474-1492.

[115] WORLD WIDE WEB CONSORTIUM THE. The OPM provenance model(OPM) [EB/OL]. (2006-05) [2025-09-18]. https://openprovenance.org/opm/.

[116] MISSIER P, BELHAJJAME K, CHENEY J. The W3C PROV family of specifications for modelling provenance metadata[C]//Proceedings of the 16th International Conference on Extending Database Technology. New York: ACM, 2013: 773-776.

[117] VÍCTOR C V, BERTRAM L, PAOLO M, et al. ProvONE: A PROV extension data model for scientific workflow provenance[EB/OL]. (2016-05-01)[2025-09-18]. https://jenkins-1.dataone.org/jenkins/view/Documentation%20Projects/job/ProvONE-Documentation-trunk/ws/provenance/ProvONE/v1/provone.html.

[118] 国家质量监督检验检疫总局, 中国国家标准化管理委员会. 信息技术——数据溯源描述模型: GB/T 34945—2017[S]. 北京: 中国标准出版社, 2017.
General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, Standardization Administration of the People's Republic of China. Information technology—Data provenance descriptive model: GB/T 34945—2017[S]. Beijing: Standards Press of China, 2017. (in Chinese)

[119] LIANG X P, SHETTY S, TOSH D, et al. ProvChain: A blockchain-based data provenance architecture in cloud environment with enhanced privacy and availability[C]//2017 17th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing. Piscataway: IEEE, 2017: 468-477.

[120] RUAN P C, DINH T T A, LIN Q, et al. LineageChain: A fine-grained, secure and efficient data provenance system for blockchains[J]. The VLDB Journal, 2021, 30(1): 3-24.

[121] RAMACHANDRAN A, KANTARCIOGLU M. SmartProvenance: A distributed, blockchain based DataProvenance system[C]//Proceedings of the Eighth ACM Conference on Data and Application Security and Privacy. New York: ACM, 2018: 35-42.

[122] GONG B, CUI C, HU M S, et al. Anonymous traceability protocol based on group signature for blockchain[J]. Future Generation Computer Systems, 2022, 127: 160-167.

[123] GENG Z Q, CAO Y, LI J, et al. Novel blockchain transaction provenance model with graph attention mechanism[J]. Expert Systems with Applications, 2022, 209: 118411.

[124] GUO Y H, XU M H, CHENG X Z, et al. zkCross: A novel architecture for cross-chain privacy-preserving auditing[C]//33rd USENIX Security Symposium. Berkeley: USENIX Association, 2024: 6219-6235.

[125] ZENG Y, ZHANG X, AKHTAR R, et al. A blockchain-based scheme for secure data provenance in wireless sensor networks[C]//2018 14th International Conference on Mobile Ad-Hoc and Sensor Networks. Piscataway: IEEE, 2019: 13-18.

[126] PORKODI S, KESAVARAJA D. Secure data provenance in Internet of Things using hybrid attribute based crypt technique[J]. Wireless Personal Communications, 2021, 118(4): 2821-2842.

[127] LIU C L, XU Y H, SUN Z X. Directed dynamic attribute graph anomaly detection based on evolved graph attention for blockchain[J]. Knowledge and Information Systems, 2024, 66(2): 989-1010.

[128] QI Y X, WU J, XU H S, et al. Blockchain data mining with graph learning: A survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024, 46(2): 729-748.

[129] SONG W S, ZHANG W Y, ZHAI L B, et al. EOS.IO blockchain data analysis[J]. The Journal of Supercomputing, 2022, 78(4): 5974-6005.

[130] LIANG Y Z, WU W J, LIANG R J, et al. A plug-and-play data-driven approach for anti-money laundering in Bitcoin [J]. Expert Systems with Applications, 2025, 266: 126072.

[131] ZHANG M Q, QU Q, NING L, et al. On time-aware cross-blockchain data migration[J]. Tsinghua Science and Technology, 2024, 29(6): 1810-1820.

[132] ZHENG W L, ZHENG Z B, DAI H N, et al. XBlock-EOS: Extracting and exploring blockchain data from EOSIO[J]. Information Processing & Management, 2021, 58(3): 102477.

作者简介

尚思远 男,2000年12月出生于黑龙江省齐齐哈尔市. 现为信息工程大学博士研究生. 主要研究方向为区块链安全共享.
E-mail: a525435400@163.com

王潇涵 男,2000年12月出生于河南省郑州市. 现为信息工程大学博士研究生. 主要研究方向为大数据安全.
E-mail: wang523648@163.com

杜学绘 女,1968年11月出生于河南省新乡市. 现为信息工程大学教授. 主要研究方向为网络信息安全.
E-mail: dxh37139@163.com

吴翔宇 男,1997年11月出生于安徽合肥市. 现为信息工程大学博士研究生. 主要研究方向为区块链安全.
E-mail: 1378406814@mail.nwpu.edu.cn

刘敖迪 男,1992年4月出生于黑龙江省伊春市. 现为信息工程大学讲师. 主要研究方向为区块链安全.
E-mail: ladyexue@163.com

王 娜 女,1980年4月出生于山西省运城市. 现为信息工程大学教授. 主要研究方向为大数据安全.
E-mail: twftina_w@126.com