

跨链交互的公共数据安全管控架构设计与实现

张晰, 徐海琛, 骆亦飞, 王睿宇, 张琳
北京市大数据中心, 北京 101117

摘要

公共数据的安全管控技术是实现公共数据高质量供给、利用的重要基础,是激发公共数据价值的重要手段。传统公共数据供给存在跨链限制和单一加密等问题,基于区块链编址选址、跨链调度和通用隐私计算技术,设计了完善的从区块链管理、密文流通技术到应用场景输出的公共数据安全管控架构。以此为理论框架,在北京市实践落地了基于“链道+数道”体系的大数据平台,实现了公共数据在政务场景中的可信、高效流通,有效实现了超大城市的公共数据管控,助力公共数据向安全化、移动化、智慧化方向发展,推进城市全域数字化转型。

关键词

公共数据; 安全管控; 跨链交互; 隐私计算

中图分类号: TP181

文献标志码: A

doi: 10.11959/j.issn.2096-0271.2024066

Design and implementation of public data security control framework for cross-chain interaction

ZHANG Xi, XU Haichen, LUO Yifei, WANG Ruiyu, ZHANG Lin
Beijing Bigdata Centre, Beijing 101117, China

Abstract

The security control technology of public data is an important foundation for achieving high-quality supply and utilization of public data, and an important means to stimulate the value of public data. The traditional public data supply has some problems such as cross-chain limitation and single encryption. The public data security control architecture was designed, which included block chain management, clear and ciphertext circulation technology, and application scenario output. Taking this framework as a theoretical framework, a big data platform based on “Chain + Data” system has been put into practice in Beijing to realize the credible and efficient circulation of public data in the government affairs scene. It effectively realizes the control of public data in super-large cities that helps the development of public data in the direction of security, mobility and intelligence, and promotes the digital transformation of the whole city.

Key words

public data, security control, cross-chain interaction, privacy computing

0 引言

数字中国建设明确提出以数据融通、开发利用贯穿城市全域数字化转型建设始终，更好服务城市高质量发展、高效能治理、高品质生活，支撑发展新质生产力，推进中国式现代化城市建设。《中共中央国务院关于构建数据基础制度更好发挥数据要素作用的意见》（简称“数据二十条”）明确强调建立数据可信流通体系，增强数据的可用、可信、可流通、可追溯水平，实现数据流通全过程动态管理，充分激活数据价值。当前，公共数据由于具有良好基础设施条件和管理，成为目前较有挖掘价值和实践基础的数据类型。如何对公共数据进行规范管理和高效利用成为亟待解决的关键问题。

随着算力、算法等基础技术体系的发展，目前全球部分世界级城市已具备全域数据基础，并通过加强公共数据的汇聚共享和开放开发，强化统筹授权和管理，推进互联互通，开展数据多样化的开发利用，旧金山市开展了“DataSF”计划，已发布420个数据集。上海市按照“自上往下”的数据管控路径，以应用场景为牵引、各部门需求为导向，从“一网通办”“一网通管”入手，汇聚开放公共数据资源超20亿条。山东省采取“自下往上”方式，围绕基础库、主题库、专题库等，发布人口、法人、证照等数据约45.6亿条。

数据采集汇聚质量形成了一个巨大的多模态数据库，具有高复杂性、高成长性的特征，单点的效率提升与整体的效率提升在一定程度上是互斥的，数据泄露的风险仍难以消除，对数据安全监管能力提出了更高的要求。超大城市市场下数据安全

管控的关键痛点在于：一是数据巨量、未知，无法实现城市全域数据安全监管；二是数据分散、敏感，数据协同计算难、效率低；三是场景复杂、多变，领域级技术体系在跨领域应用时响应慢、适配性差。

为解决以上问题，受到隐私计算“可用不可见”的启发，在区块链和隐私计算的总体技术框架下，基于跨链交互的公共数据安全监管架构被提出，以实现城市级复杂系统的数据全域连通。通过区块链解决确权和管控的问题，通过隐私计算解决明文传输和信任的问题，并基于区块链跨链、编址寻址、隐私计算等技术来实现城市复杂系统下全域数据的逻辑管控，由物理上的数“聚”转变为逻辑下的汇“通”，实现安全监管下的公共数据有序流通、按需调用和隐私保护。

1 数据安全监管相关技术的发展现状

1.1 数据管控技术方案

数据管控的技术不断发展迭代，逐步由层次数据库、关系型数据库等集中式数据库向分布式数据库演变^[1]。分布式数据库将数据存储多个地点，提高了数据可靠性和处理效率，但是同时对数据的管理提出了更高的要求。区块链和隐私计算技术作为近年发展起来的数据流通关键技术，已经被逐步应用到数据管控架构中^[2]。目前，数据要素市场体系的不断建设完善，数据管控方法和体系也逐步丰富。国际上ISO/IEC38500系列标准及ISO/IEC38505标准提出了数据应用、数据治理和数据监督的通用模型和方法。窦悦等^[3]提出了以数据标识融合、跨区块链及跨隐私计算平台基础支撑体系为依托，以数据流通交易公共服务体系为载体的国家“数联网”根

服务体系总体建设思路。孟凡等^[4]针对传统省级政务大数据平台的3种建设模式进行了总结和问题提炼,针对数据基础工程的发展提出了一定的建议。

已有的数据管控平台主要注重其明文传输能力和一体化架构能力,对管理框架较重视。但是,对政务系统内部的数据安全和可信流通问题的专门技术研究还较少,亟须通过技术安全手段提升相关数据的安全管控能力。

1.2 区块链相关研究与技术方案

区块链具有存证可靠、难以被篡改等去中心化优势,其构建的新型信任机制可以从基础信任模式上改变当前数据权属难界定、信任机制难建立的现状,促进数据安全可信流通^[5]。尤其是在政务服务领域,区块链技术为解决部门间共享难、数据权责不清晰等问题方面具有独特的优势^[6]。目前,国内外学者对区块链技术在数据安全和数据管理领域已有了一定的研究和实践成果。

Jia等^[7]提出了一种基于区块链探针的跨组织数据共享框架,其通过使用复杂网络的星形数据链路,实现了异构数据源的同步。汪靖伟等^[8]研究了基于区块链的数据市场框架,分析了框架中的安全性和隐私性问题及对应的解决方案。王海龙等^[9]提出了基于区块链的数字水印技术,实现对大数据完整性的轻量级审计和溯源。贾晓丰等^[10]提出了一套基于目录生成的分布式数据标识和协同计算方法。刘彦松等^[11]建立了一套基于区块链网络的链上数据共享架构,主要考虑并解决了存在恶意参与方时的隐私泄露问题。针对数据交易中的隐私保护问题,王飞跃等^[12]以区块链技术为支撑,以联邦智能为驱动,借助联邦控制实现数据联邦化,且联邦生态充分发挥

了大数据和人工智能的潜力,进而为实现联邦智能提供了一种新的思路。

当前不同区块链间的数据流通还存在一定的限制,不同的区块链底层协议的数据格式、协议流程等都有差别,在实际的应用中,会导致不同链的数据不能互通,从而产生“信息孤岛”^[13]。区块链跨链互通不仅涉及数据的可信交互,还需实现身份互认。当前不同系统的实现方案不同,增加了跨链互通的难度,导致“链级孤岛”问题日益突出,影响区块链网络的互操作性^[14]。现有的数据管控系统作为多种技术的集成系统,区块链在自身的可扩展性性能、安全性底层技术平台等方面仍存在技术屏障。尤其是在跨链交互方面,存在一定的瓶颈,对数据管控流通造成了较大的困难^[15]。跨链技术是突破区块链上数据迁移的一种高效方式,其技术分类包括公证人机制、侧链、中继等多种方式,其技术特点见表1。跨链技术按照参与双方区块链底层平台技术,可分为同构链跨链技术和异构链跨链技术^[16]。同构链跨链技术相对简单,如Polkadot的跨链型公链,通过提供给开发者SDK (software development kit)来构建自己的区块链,然后通过中继链/Hub解决跨链的消息传递和验证问题^[17]。而异构链跨链技术相对复杂,如Polkadot和Cosmos之间跨链交互,还需进行额外的匹配互认。

1.3 隐私计算相关研究与技术方案

现有数据管控技术、方法和手段大多以访问控制、安全防护为主,获取数据明文成为数据使用方或计算设备开展数据应用的前提,存在各种潜在的数据泄露和滥用风险。随着密码应用的发展,数据脱敏、可信计算环境、差分隐私等技术相继出现,

表1 主流跨链技术

分类	技术特点
公证人机制	公证人模式是一种中心化的结构,通过选定一批见证人并在见证人之间采用拜占庭容错结构,监听目标链上的事件和状态,并签名进行资产的转移
侧链	侧链是以锚定某种原链上的代币为基础的新型区块链,它是主链外的另一个区块链,通过主链上的计算力来维护侧链的真实性,实现公共区块链上价值与其他账簿上价值在多个区块链间的转移
中继	中继技术通过在两个链中加入一个数据结构,使两个链可以通过该数据结构进行数据交互,并通过在一个链上调用数据结构的API,实现监听并验证另一个链上的交易,而若该数据结构是一个链式结构,并具备侧链的形式则其可以被称作中继链
哈希锁定	交易双方在区块链上预先设有一个可扩展的比特币微支付子链网络,用于提升链外的交易处理能力,使用哈希锁定将发起方的交易代币进行锁定,并通过时间锁定让接收方在某个约定的时刻前生成支付的密码学证明,并与先前约定的哈希值一致,则可完成交易
分布式私钥控制	加密资产通过分布式私钥生成与控制技术被映射到基于协议的内置资产模板的区块链上,根据跨链交易信息部署新的智能合约、创建新的资产

在保证数据隐私安全、促进数据融合应用方面发挥了一定作用,但明文仍存在于数据共享和计算过程中,存在隐私泄露的隐患。隐私计算也被称为隐私保护技术、隐私保护计算,是面向隐私信息全生命周期的隐私保护计算理论和方法^[18],目前主流的隐私计算技术可分为以多方安全计算(multi-party computation, MPC)为代表的密码学技术和以可信硬件为载体的可信执行环境(trusted execution environment, TEE)技术等多种技术体系^[19]。

理论研究层面,符芳诚等^[20]梳理总结了近年来隐私计算在隐私集合求交技术、联邦学习、异步并行计算、可信执行环境等关键技术方面的创新成果。Arachchige等^[21]通过研究联邦学习、差分隐私、区块链等技术的特点,建立了一种融合上述技术的新型隐私计算模型,能够进一步加强数据隐私保护的安全性和可靠性。

Huang等^[22]结合差分隐私与联邦学习技术,设计出一种包含自适应梯度下降算法的差分隐私专用卷积神经网络框架,能够有效提升隐私计算在不平衡数据集中的表现性能。李懿等^[23]提出了一种基于区块

链和加密技术的数据安全共享的隐私数据流通管控方法,将隐私计算的零知识证明技术同区块链进行了融合应用。目前主流隐私计算平台主要包括百度MesaTEE安全计算平台、华控清交PrivPy多方安全计算平台、Microsoft SEAL等。但是,由于明确的激励机制、收益分配机制与平台收费机制等商业化落地模式尚未形成,隐私计算技术仍处于快速发展的早期阶段^[24]。

针对已有的区块链和隐私计算在数据管控中还存在的跨链难、模式多、体系缺等问题,本文着重通过现有的区块链技术和隐私计算基础,创新性地利用区块链跨链技术和通用隐私计算能力,结合明文计算、大数据基础设施、公共数据应用场景,构建了一套基于区块链和隐私计算的可信、高效数据安全管控架构。

笔者的研究与已有的研究和应用相比,主要存在3个方面的主要差异。一是完成了区块链不同链之间数据格式的互认互通,并保证跨链的可验证性,即跨链的消息从源头到传输过程是真实的和可验证的。二是综合运用各类隐私计算技术和明文传输方法,把数据可见的“具体信息”和不用

看见即可进行计算的“使用价值”分开，做到隐私数据“可用不可见”，防止数据使用过程中的信息泄露，并且规定数据的具体用途和用量以杜绝数据滥用，实现既能支持多方数据融合应用，又能保护各方数据隐私的目标。三是结合相关突破性技术，设计和实现了一整套具有“链道+数道”的数据安全管控系统，保证了超大城市数据流通场景下的政务数据流通的高效性和安全性，为数据要素的价值发现提供了坚实的理论和工程基础。

2 公共数据安全管控架构设计

2.1 总体架构

本文提出的公共数据安全管控架构采

用“链道+数道”架构设计，如图1所示。一方面基于区块链技术，通过“链道”承载控制流，实现复杂多模态大数据的统一授权管控，实现公共数据在不同节点之间有序流通的事前驱动、事中管控和事后记录。另一方面通过“数道”承载数据流，提升公共数据流通能力，实现明文数据和密文数据的汇聚处理，解决敏感数据的流通使用问题，保障数据的互信和安全。

(1) 基于目录区块链实现数据跨域管控

目录区块链系统通过数据发现、数据申请、合约管理、审计留痕模块实现对数据的顶层管理。具体来说，数据发现对数据目录跨域权限进行分级，并提供跨域数据发现服务。数据申请根据算法权限分析提供数据跨域申请服务。合约

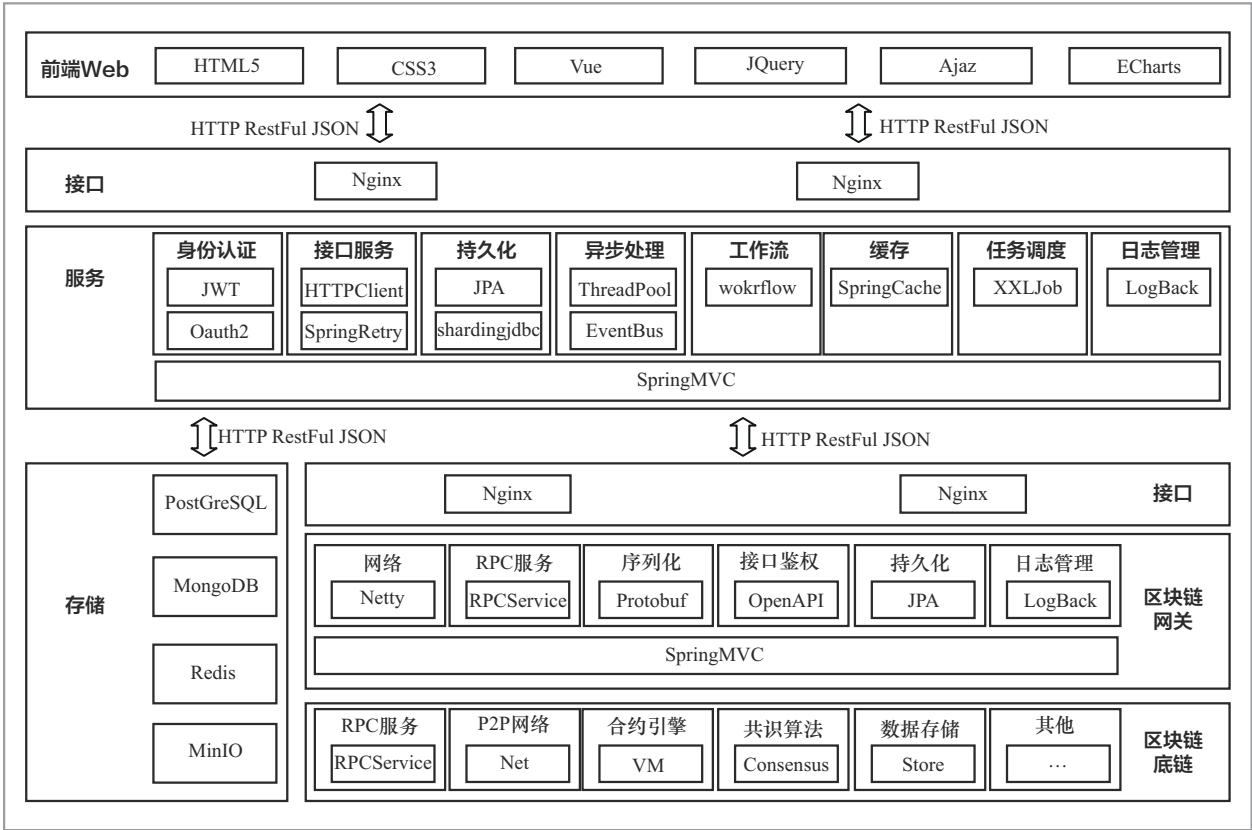


图1 目录区块链技术架构

管理对跨域数据订立计算合约，发起计算合约流程。审计留痕是对跨域数据的操作全程记录。

（2）基于隐私计算基础平台实现数据安全流通

“数道”通过明文交换和隐私计算基础平台实现数据的安全流通，其中明文能力包括常规的数据交换、API、数据探针等功能，隐私计算基础平台提供了多种密文处理能力，主要包括数据路由实现跨域数据的查询和定位、应用开发实现跨域数据的样例测试和应用开发、算力分配通过对跨域的算力进行自适应分配实现最佳算力效能、合约执行模块根据跨域数据的计算合约和场景需求选择适当的算力分配模式。

2.2 “链道”区块链技术架构

为加强公共数据统筹、建立“秩序”，在数据目录体系的基础之上，引入区块链的公开、分权、难以篡改、合约机制等理念和技术，提出了“目录区块链”，实现数据与职责的强关联、数据变化的实时探知及数据访问的全程留痕，保证目录的可见、可用、可考核。

目录区块链中承载了数据的确权和分权管控机制，其核心作用体现在数据流转的所有动作，已发生的都由“链”来记录、在发生的都由“链”来管控、将发生的都由“链”来驱动。目录区块链技术架构分为前端 Web 层、接口层、服务层和存储层 4 层，技术架构如图 1 所示。

具体来说，前端 Web 层采用 Vue 框架，构建用户界面的渐进式框架。接口层主要实现后台接口代理、负载均衡、服务层健康状态检查等。服务层主要实现身份认证、接口服务、异步处理等功能。存储层主要分为区块链网关和区块链底链两部

分。区块链网关主要实现与区块链网络通信。区块链底链主要实现 P2P（peer-to-peer）网络、合约引擎、共识算法等功能，共识算法模块目前支持 Solo、Raft、TBFT、HotStuff 共 4 种共识类型。

2.3 “数道”明密文技术架构

明文数据存在易被无限复制、随意使用、隐私泄露等风险。笔者利用多方安全计算、联邦学习等隐私计算技术，提出了隐私计算基础平台，可将明文数据的“具体信息”和“计算价值”分开，在数据流通过程中保护具体信息、释放计算价值，解决敏感数据的“可用不可见”问题，实现数据在不被浏览、复制和修改的前提下按需计算的目的，并对计算的用途和用量进行有效控制。

建设创新性、领先性和实用性的隐私计算基础平台，可以促进多源数据之间的安全融合，解决数据融合中隐私保护和促进交互的核心矛盾。其多方安全计算技术的特性，使数据使用方和计算平台管理方都无法接触到明文数据，并且能够保证计算结果与明文结果完全一致，真正实现数据的可用而不可见，从而进一步提升数据管理效率效能，促进数据的安全有序流动。

隐私计算基础平台通过任务调度和计算执行引擎，让数据提供方接入数据、算法提供方开发程序、数据获取方得到计算结果，能够快速定制不同的业务规则（计算合约），支持快速开发不同的应用场景。整个业务逻辑的核心在于计算合约中规定了数据的来源、用法和用量以及参与计算和收益分配的规则，计算合约执行的每一步均被存证，保证整个过程的清晰、可监管和可审计。隐私计算基础平台功能架构如图 2 所示。

隐私计算基础平台主要分为计算引擎层、算法模型层和运行监管层3层结构设计。

(1) 计算引擎层

具体来说，计算引擎层包含数据服务和计算引擎两部分模块。数据服务模块的主要功能为数据进行密文接入和上传、计算结果解密，在明密文协同计算中接收本地计算任务，并进行执行。计算引擎模块的主要功能在充分保护数据隐私的基础上，进行纯密文计算和明密文协同计算，并进行计算资源的调度，其支持多种隐私计算引擎，包括秘密分享引擎、同态加密引擎、混淆电路引擎、联邦学习引擎和TEE计算引擎。

(2) 算法模型层

算法模型层主要功能是为算法提供方提供一整套基于密文计算应用研发、调试、测试、部署和发布的环境和工具，支持应用从研发到发布的整个流程，对算法脚本进行指令、编译解析和优化，并预置常用

的密文计算函数库、机器学习和深度学习算法库，用于业务应用模型算法开发，降低通用场景的研发成本。

(3) 运行监管层

运行监管层包含平台运行和平台监管两个功能模块。平台运行模块主要提供供需对接、计算合约管理、运行服务、计量计费功能。平台监管模块主要提供供方监管、需方监管、计算特征功能，其支持区块链存证，可以实现问题溯源和倒查。

3 “数链”管控体系关键技术

3.1 技术框架

本文提出的公共数据安全管控架构拟通过基于区块链和隐私计算技术，开展“数链”管控体系的工程化实践，形成“链道”（目录区块链）+“数道”（明密文传输）的公共数据安全管控的技术框架，如图3所示。

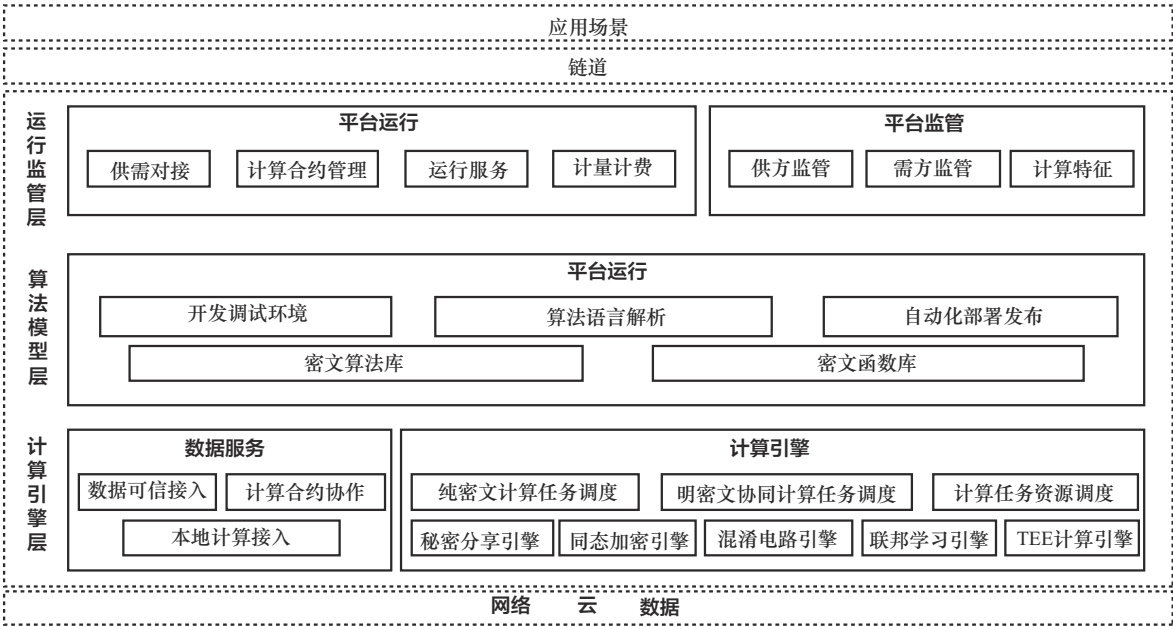


图2 隐私计算基础平台功能架构

笔者基于共识协议、智能合约、编址寻址、跨链交互的区块链核心技术，设计了“链道”技术体系，其具备全流程数据管控的能力；同时，基于隐私计算等密文流通能力和数据探针等多种明文传输能力，构建了“数道”技术体系，并综合利用多方安全计算、联邦学习、同态加密、可信执行环境等多种隐私计算技术，设计了隐私计算基础平台，可以执行多方数据的流转，保证计算过程的安全和只有确定的结果接收方能够获得计算结果。在跨域数据共享流通的过程中，“链道”承载控制流，实现复杂多模态大数据的统一授权管控，“数道”承载数据流，实现敏感数据的安全流通。“链道”和“数道”的高效协同，可以利用智能合约对计算任务执行和数据流转全过程在“链道”进行存证，同时，利用“数道”的隐私计算技术保护数据的隐私性，保障了数据流通的强监管、互信和安全，支撑数据在不同应用场景的汇聚和

流动。

为了实现公共数据安全监管技术框架规模大、结构多、隐私重等特点，笔者创新整合了跨链调度、区块链编址与寻址、明密文协同计算等关键技术，以促进总体架构的实现。具体来说，跨链调度主要解决了公共数据区块链结构多样的兼容性问题，实现了多类结构的联通管理。区块链编址与寻址解决了公共数据规模巨大、场景多样的区块链构建问题。明密文协同计算融合了明文数据采集流通能力和隐私计算安全流通能力，从理论上实现了数据可用不可见。

3.2 跨链调度

跨链调度是解决同构或异构区块链网络之间互操作、价值转移的一种重要的技术。针对公共数据流通的需求，为兼容多种异构链，跨链方案采用双向操作的合约实现方式，当开展基于区块链技术的不同

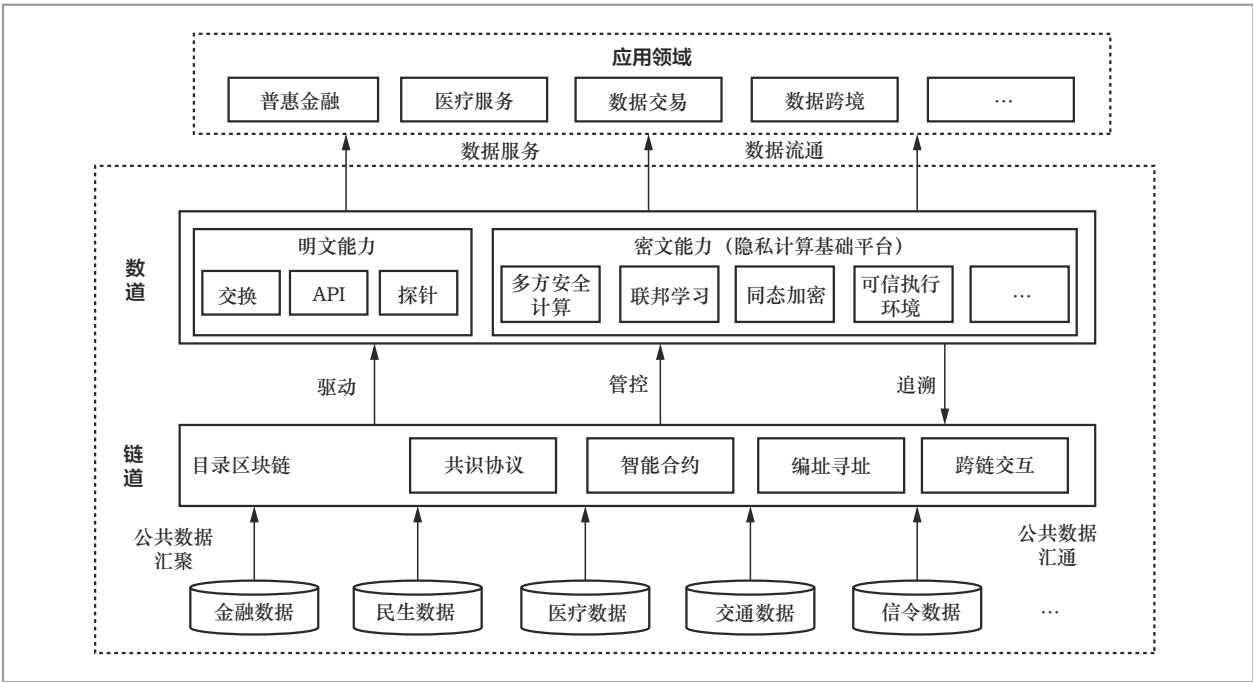


图3 公共数据安全监管技术框架

数据链应用，或市级和区级目录区块链对接时，通过跨链调度将数据链或区级目录区块链上的数据产品发布记录、数据共享申请记录、数据共享授权记录、数据使用记录以及用户其他行为记录等通过跨链技术交互到目录区块链，具体实现方式如图4所示。

在目录区块链端部署本链SDK和数据链SPV (simplified payment verification) 节点，在数据链端部署本链SDK和目录区块链SPV节点，经由跨链核心组件，实现数据由数据链到目录区块链的交互。目录区块链根据数据专区的数据授权，监控专区的数据使用情况。如果出现数据被非授权使用、非法获取或其他不当应用情况，平台发出预警，相关人员及时排查问题并采取得当的处理办法。具体来说，本链SDK用于将待跨链交易提交至本链上，需在本链上登记并配置合法的数字证书。对端链（目录区块链或数据链）SPV节点用于对端链交易实时订阅和交易有效性校验，确保跨链数据的有效性和交易跨链的时效性。

跨链技术考虑事务一致性、交易可验证性和交互方式3个核心问题。事务一致性方面，跨链网关需支持回滚方式的事务处理方式，当某一条链的智能合约执行失败，支持其在该链上发起回滚操作，以确保所有链的事务一致性。交易可验证性方面，采用SPV节点的交易验证逻辑，基于Merkle树存在证明和加密算法实现交易的验证。交互方式方面，采用网关的方式直连两端的区块链节点，支持高效的跨链交互需求，并以简化的实施方式降低运行维护成本。

3.3 链上编址与寻址

BNS (blockchain name service) 通过服务所有权和使用权分离的运营模式划

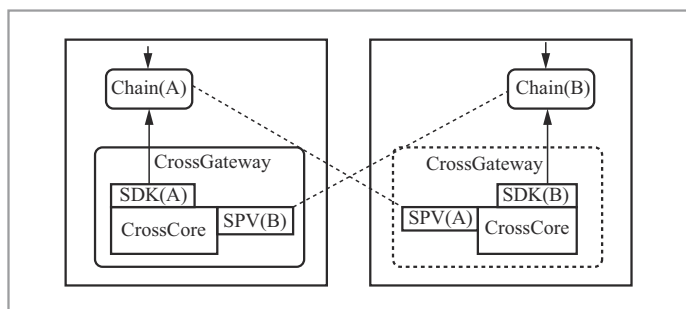


图4 跨链交互实现方式

分服务提供方和使用方角色，在两者之间建立协同运行的快速通道，实现基于区块链的服务资源共享以及应用生态构建。BNS实现方式如图5所示。

BNS在技术架构上分别从服务访问层、服务能力层、服务引擎层、合约管理层，逐层进行系统的实施建设，为社会主体各类区块链应用建立统一的资源命名和访问机制，提供开放的区块链资源访问服务，实现可信、可靠、高效的区块链资源命名、寻址和访问服务。

具体来说，该技术包含域名注册、编址、寻址3个阶段。域名注册阶段，该系统提供注册域名请求、生成签名密钥 (secret key, SK) 功能，并将对应的详细信息上链存储；编址阶段，该技术会将不同信息自动生成地址，并将地址及其对应的信息上链存储；寻址阶段，该技术提供域名解析、寻址SDK下载等功能，用户可通过寻址SDK访问业务链上的资源数据。

3.4 明密文协同计算和通用隐私计算框架

在实际应用场景中平台兼顾各参与方对数据安全性和计算效率的需求，根据数据的隐私性要求和计算任务的具体逻辑，灵活进行明密文计算任务的编排、调度与执

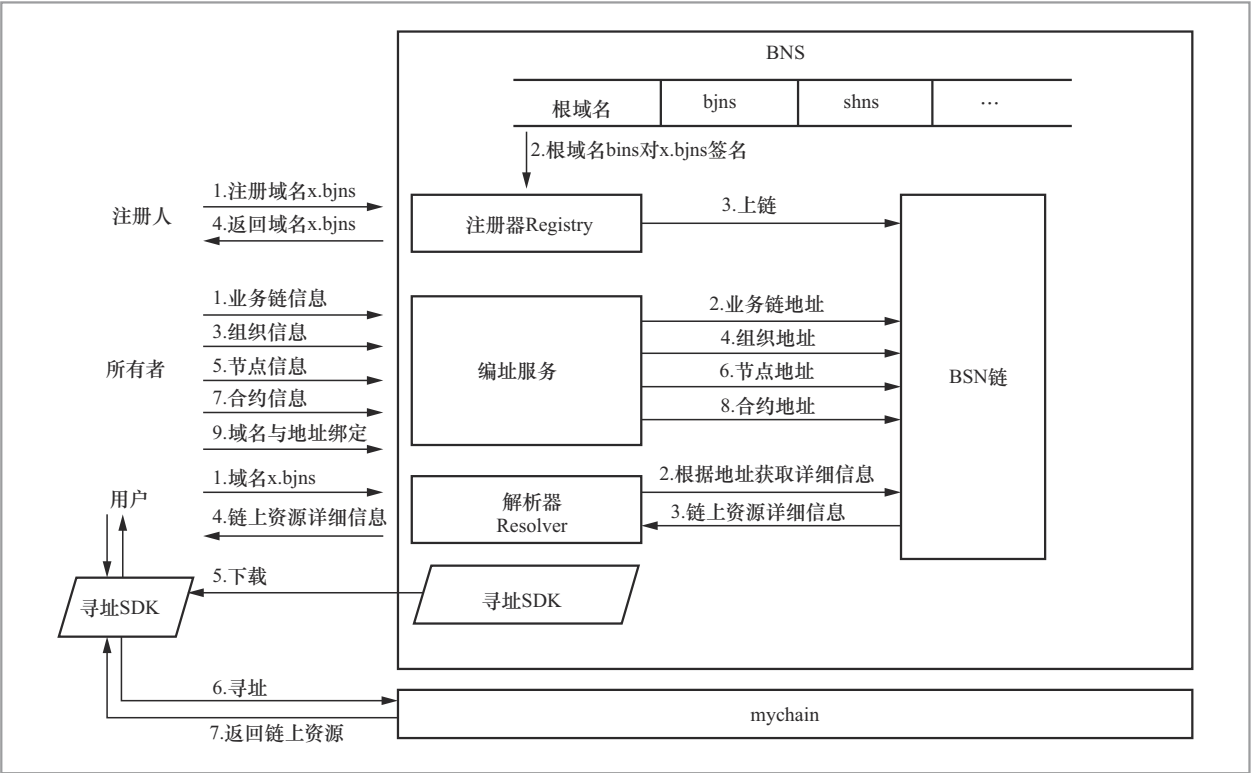


图5 BNS实现方式

行，在保证数据隐私性的前提下最大限度提升计算效率。明密文协同计算框架如图6所示。

明文计算在数据提供方本地的明文计

算引擎开展，密文计算将数据从密文接入模块加密后进入密文计算模块开展，密文计算结果通过密文接入模块解密后，可再次进行明文计算。在一次计算任务中，可

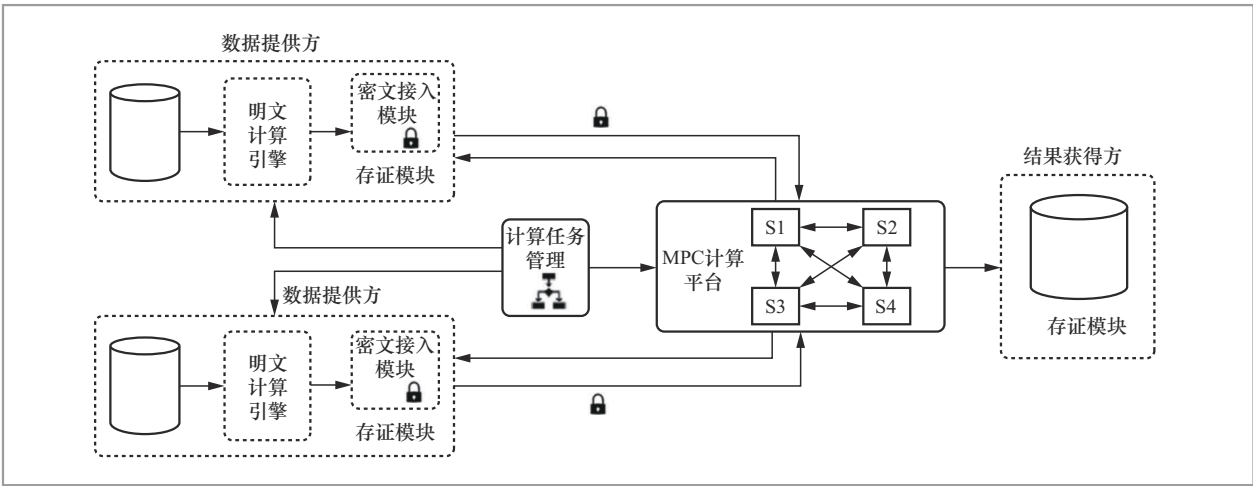


图6 明密文协同计算框架

以进行明文计算和密文计算的多轮交互。

通用隐私计算框架综合运用密码学等各类隐私计算技术的基础理论,在计算机指令集和编译器层面用密文运算代替明文运算,建立通用的隐私计算体系。针对特殊的算法逻辑,进行算法性能优化,实现从密码学原理到应用算法的构建。通用隐私计算框架如图7所示。

考虑到隐私计算体系中源数据和中间结果对计算方不可见的特性,许多明文算法逻辑不再成立(如循环语法),算法复杂度大幅提升(如排序算法)。构建通用隐私计算框架主要包括完成隐私计算体系下基础运算符(加法、乘法、比较)的重载,在基础运算符的基础上实现衍生运算(除法、逻辑函数、ReLU函数等),再进一步实现对科学计算和机器学习常用算法库(Numpy、PyTorch等)的支持。通过该框架形成的隐私计算替代方案和优化策略可重新实现各类算法,保证其正确性和高效性。

4 案例实验

北京市经过多年的信息化建设,积累

了大量的公共数据。为了实现数据的有效流通和价值释放,基于以上关键技术和公共数据安全管控架构的提出,北京市构建了大数据平台。北京市大数据平台通过对区块链的技术应用,创新性地通过跨链调度、链上编址寻址和隐私计算等关键技术,实现了明密文数据流通的管控,完成了在安全性和功能性上的多维度提升,有效解决了传统数据会聚系统中的权责不明和动机缺乏等问题。

具体来说,其基于目录区块链实现数据跨域管控。目录区块链系统通过数据发现、数据申请、合约管理、审计留痕模块实现对数据的顶层管理,其系统页面如图8所示。

通过隐私计算、数据交换等功能模块,北京市大数据平台实现了明密文两方面的数据流通。明文流通由于其安全性难以被保证,多用于局域网络中数据流通。该平台在其基础上,基于隐私计算的相关关键技术,实现了数据安全流通。“数道”通过明文交换与隐私计算基础平台实现数据的安全流通,其中明文能力包括常规的数据交换、API、数据探针等功能,隐私计算基础平台提供了多种密文处理能力,其隐私计算系统和算法环境如图9所示。

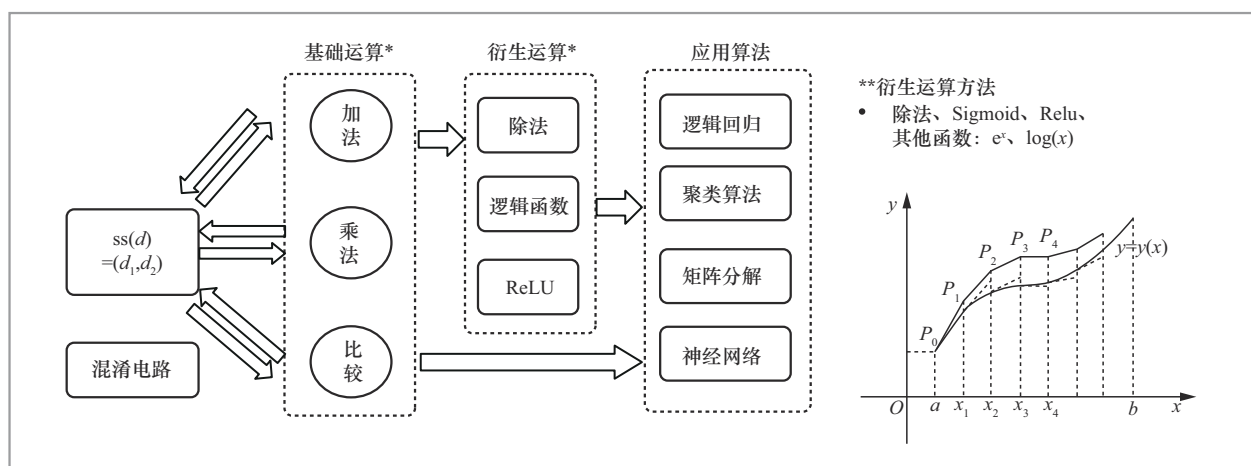


图7 通用隐私计算框架



图8 北京市目录区块链系统页面



图9 隐私计算平台算法编辑页面

4.1 跨链交互性能实验

为了对大数据平台的跨链调度能力和隐私计算等相关关键技术创新进行核实，该平台利用其仿真能力对跨链交互和隐私计算能力进行测试和比较。仿真系统部署在5个连接的虚拟服务器上，每一个服务器都

有一个静态固定内存分配64 GB的RAM内存和4核Intel (R) Xeon (R) CPU E5-2620 v4 芯片组，连续运行72小时进行长稳测试，并对其跨链交互主要性能指标TPS (transactions per second) 进行测试。

跨链交互的功能有创新性突破和主要功能：跨链长连接具有保活能力；跨链系

统新增 Merkle 树预加载功能运行正常；跨链系统支持两个独立 Proxy 集群且支持独立配置是否开启 Merkle 树校验。图 10 显示了性能测试结果，平台跨链方式主要同以太坊和 Near 合作公开的彩虹桥 (rainbow bridge, RB) 进行比较。在 Payload 5K 的情况，开启多链验证（两个签名）、双向 Merkle 校验的前提下，RB 方式的 TPS 均值为 263.2，平台跨链交互 TPS 约为 330.1。在 Payload 5K 的情况，开启多链验证（两个签名）、关闭双向 Merkle 校验的前提下，RB 方式的 TPS 均值为 487.9，平台跨链交互 TPS 约为 652.4，其在多状态下的性能均高于现主流跨链方法。其长稳测试方面也表现优异，未发现系统异常。

4.2 隐私计算性能实验

隐私计算功能在现有隐私计算领域内，也有了性能上的突破。在隐匿查询方面，实现了极高计算效率的批量隐匿查询，GB 级数据库万次查询平均响应时间在 1 ms 以内。完全在线计算的批量隐匿查询，10 Mbit/s 带宽下 GB 级数据库万次查询平均响应时间在 100 ms 以内。安全求交方面，两方均为亿级数据量的一次性不分箱、完全在线计算的安全求交，计算耗时小于 1 min，100 Mb/s 带宽下端到端耗时小于 5 min。同态加密方面，基于 LWE、RLWE 的多种全同态加密算法，同态乘法的单线程性能达到毫秒级别。表 2 显示了隐私计算性能测试结果。

该平台不仅在数据管控的功能层面进行了完善和补充，形成了全面立体的公共数据安全管控体系，也在各个关键技术方面不断改进更新，逐步创新并成为更加安全高效的工具，更加能够支撑公共数据的流通应用。

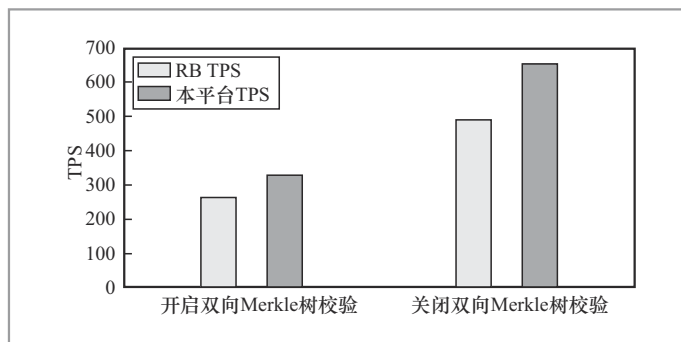


图10 跨链交互性能比较

表2 隐私计算性能测试结果

测试算法	参与方	计算类型	本平台耗时
亿级隐匿查询	两方	高效率(万次)	0.93 ms
		高安全(一次)	2.32 ms
亿级隐匿查询带宽限制	两方	高效率(万次)	89.4 ms
		高安全(一次)	34.6 ms
亿级隐私求交	两方	平衡	54 s
		非平衡	69 s
	三方	平衡	468 s
		非平衡	295 s
同态加密		LWE	0.94 ms
		RLWE	0.75 ms

5 结束语

自2018年开始，公共数据安全管控架构陆续通过筑基工程、区块链先进算力实验平台、目录体系建设等首都重点工程项目完成了技术实现和落地转化，在以区块链为核心的“链道”统一管控下，为北京市数据共享、流通提供全过程、全方位、去隐私化的监管和保护。北京市基于提出的公共数据安全管控架构，建成了北京市目录链，采用跨链方式实现市区两级公共数据对接，促进数据协同计算，通过隐私计算方式实现跨部门明密文数据共享，解决敏感数据共享难题。应用效果方面，通

过公共数据安全监管架构，北京市建立了复杂场景下的数据管控新格局，实现了城市级数据统一管控和共享，支撑北京市上百项核心业务^[25]。目前，北京市已汇聚流通公共数据资源规模上千亿条，涉及城市规划、社会信用、政务服务、交通治理、全民健康、生态环保、公共安全等多个重点领域，并且为普惠金融、城市运行、一网通办等应用场景提供应用支撑，实现了良好的社会效益和经济效益。

以区块链跨链和通用隐私计算技术为核心的公共数据安全监管架构，不仅能够有效解决数据安全流通和数据隐私保护问题，而且随着行业应用的逐渐深入，可以在电子病历管理、融资与信贷、支付与清算、数字身份认证、数字版权保护等对数据安全流通有重大需求的领域，发挥其技术优势，从而推动行业变革和服务创新。

笔者将继续开展核心关键理论的研究和技术的突破，持续致力于数据资源的有效整合、强力监管和开发利用，不断完善创新，优化闭环管理机制和技术体系，更好地服务于数据要素发挥和数字经济建设。

参考文献：

[1] RUAN P C, DINH T T A, LOGHIN D, et al. Blockchains: decentralized and verifiable data systems[M]. Cham: Springer, 2022.

[2] 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.

[3] 窦悦, 易成岐, 黄倩倩, 等. 打造面向全国统一数据要素市场体系的国家数据要素流通共性

基础设施平台: 构建国家“数联网”根服务体系的技术路径与若干思考[J]. 数据分析与知识发现, 2022, 6(1): 2-12.

DOU Y, YI C Q, HUANG Q Q, et al. Constructing a common data circulation infrastructure platform for the national unified data factor market—technical path and policy thinking of constructing the national “data networking” root service system[J]. Data Analysis and Knowledge Discovery, 2022, 6(1): 2-12.

[4] 孟凡, 杨群力, 高阳, 等. 省级政务大数据平台建设模式研究与启示[J]. 大数据, 2024, 10(3): 27-39.

MENG F, YANG Q L, GAO Y, et al. Research and enlightenment on the construction mode of provincial government big data platform[J]. Big Data Research, 2024, 10(3): 27-39.

[5] 李源, 高宁, 孙晶, 等. 基于区块链的大数据交易模式研究与探索[J]. 大数据, 2021, 7(4): 37-48.

LI Y, GAO N, SUN J, et al. Research and exploration of big data transaction model based on blockchain[J]. Big Data Research, 2021, 7(4): 37-48.

[6] 陈锐, 贾晓丰, 赵宇. 基于模糊聚类的智慧城市多源信息协同结构测度与优化[J]. 计算机应用研究, 2016, 33(7): 1945-1951.

CHEN R, JIA X F, ZHAO Y. Structure measure and optimization of multisource information coordination of smart city based on fuzzy clustering[J]. Application Research of Computers, 2016, 33(7): 1945-1951.

[7] JIA X F, XU W J, DENG P, et al. Cross-organisational data sharing framework based on blockchain-probes[J]. IET Networks, 2023, 12(2): 77-85.

[8] 汪靖伟, 郑臻哲, 吴帆, 等. 基于区块链的数据市场[J]. 大数据, 2020, 6(3): 21-35.

WANG J W, ZHENG Z Z, WU F, et al. Blockchain based data marketplace[J]. Big Data Research, 2020, 6(3): 21-35.

[9] 王海龙, 田有亮, 尹鑫. 基于区块链的大数据确权方案[J]. 计算机科学, 2018, 45(2): 15-19, 24.

WANG H L, TIAN Y L, YIN X.

- Blockchain-based big data right confirmation scheme[J]. Computer Science, 2018, 45(2): 15-19, 24.
- [10] 贾晓丰, 高嵩, 周琰, 等. 一种面向超大城市治理的数据高效跨域流通技术框架[J]. 数据与计算发展前沿, 2023, 5(5): 35-45.
JIA X F, GAO S, ZHOU Y, et al. A technical framework of efficient cross-domain data circulation for mega-city governance[J]. Frontiers of Data & Computing, 2023, 5(5): 35-45.
- [11] 刘彦松, 夏琦, 李柱, 等. 基于区块链的链上数据安全共享体系研究[J]. 大数据, 2020, 6(5): 92-105.
LIU Y S, XIA Q, LI Z, et al. Research on secure data sharing system based on blockchain[J]. Big Data Research, 2020, 6(5): 92-105.
- [12] 王飞跃, 王艳芬, 陈慧竹, 等. 联邦生态: 从联邦数据到联邦智能[J]. 智能科学与技术学报, 2020, 2(4): 305-311.
WANG F Y, WANG Y F, CHEN Y Z, et al. Federated ecology: from federated data to federated intelligence[J]. Chinese Journal of Intelligent Science and Technology, 2020, 2(4): 305-311.
- [13] 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.
- [14] 王洒洒, 戴炳荣, 朱孟禄, 等. 面向跨链系统的用户身份标识认证模型[J]. 计算机工程与应用, 2022, 58(19): 135-141.
WANG S S, DAI B R, ZHU M L, et al. User identity authentication model for cross-chain system[J]. Computer Engineering and Applications, 2022, 58(19): 135-141.
- [15] HE Y H, ZHANG C, WU B, et al. A cross-chain trusted reputation scheme for a shared charging platform based on blockchain[J]. IEEE Internet of Things Journal, 2022, 9(11): 7989-8000.
- [16] YANG G Z, ZANG C, CHEN J J, et al. Distributed fusion cross-chain model and architecture[J]. IET Blockchain, 2022, 2(2): 29-43.
- [17] WOOD D G. Polkadot: vision for a heterogeneous multi-chain framework[J]. White Paper, 2016, 21(2327): 4662.
- [18] LI F H, LI H, NIU B, et al. Privacy computing: concept, computing framework, and future development trends[J]. Engineering, 2019, 5(6): 1179-1192.
- [19] LI T, WANG H Q, HE D B, et al. Blockchain-based privacy-preserving and rewarding private data sharing for IoT[J]. IEEE Internet of Things Journal, 2022, 9(16): 15138-15149.
- [20] 符芳诚, 侯忱, 程勇, 等. 隐私计算关键技术与创新[J]. 信息通信技术与政策, 2021(6): 27-37.
FU F C, HOU C, CHENG Y, et al. Key technology and innovation of privacy preserving computing[J]. Information and Communications Technology and Policy, 2021(6): 27-37.
- [21] ARACHCHIGE P C M, BERTOK P, KHALIL I, et al. A trustworthy privacy preserving framework for machine learning in industrial IoT systems[J]. IEEE Transactions on Industrial Informatics, 2020, 16(9): 6092-6102.
- [22] HUANG X X, DING Y, JIANG Z L, et al. DP-FL: a novel differentially private federated learning framework for the unbalanced data[J]. World Wide Web, 2020, 23(4): 2529-2545.
- [23] 李懿, 王劲松, 张洪玮. 基于区块链与函数加密的隐私数据安全共享模型研究[J]. 大数据, 2022, 8(5): 33-44.
LI Y, WANG J S, ZHANG H W. Research on privacy data security sharing scheme based on blockchain and function encryption[J]. Big Data Research, 2022, 8(5): 33-44.
- [24] 王鹏, 魏必, 王聪. 区块链技术在政务数据共享中的应用[J]. 大数据, 2020, 6(4): 105-114.
WANG P, WEI B, WANG C. Application of blockchain technology in government data sharing[J]. Big Data Research, 2020, 6(4): 105-114.
- [25] 国家工业信息安全发展研究中心. 区块链助力

数字新基建：2022年区块链典型应用实践精
编[M]. 北京：电子工业出版社，2024.
National Industrial Information Security De -
velopment Research Center. Blockchain helps

new digital infrastructure: a compilation of
typical application practices of blockchain in
2022[M]. Beijing: Publishing House of Elec -
tronics Industry, 2024.

作者简介



张晰（1980-），女，北京市大数据中心高级工程师，主要研究方向为公共数据治理、区块链等。



徐海琛（1971-），男，北京市大数据中心高级工程师、副主任，主要研究方向为数据挖掘、政务大数据等。



骆亦飞（1992-），男，博士，北京市大数据中心工程师，主要研究方向为时空数据、区块链、数据挖掘等。



王睿宇（1996-），男，北京市大数据中心助理工程师，主要研究方向为时空数据、数据挖掘等。



张琳（1974-），男，博士，北京市大数据中心主任，主要研究方向为时空数据、人工智能、信息处理等。

收稿日期: 2024-08-15
通信作者: 张琳, zhanglin@jxj.beijing.gov.cn
基金项目: 国家重点研发计划项目(No.2021YFB2104805)
Foundation Item: The National Key Research and Development Program of China (No.2021YFB2104805)