软件学报 ISSN 1000-9825, CODEN RUXUEW
Journal of Software,2023,34(1):1−32 [doi: 10.13328/j.cnki.jos.006402]

E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel: +86-10-62562563

区块链数据安全服务综述*

王利朋[1,2], 关 志[1], 李青山[1], 陈 钟[1], 胡明生[2]

[1](北京大学 信息科学技术学院, 北京 100871)

[2](郑州师范学院 信息科学与技术学院, 河南 郑州 450044)

通信作者: 关志, E-mail: guan@pku.edu.cn

摘 要: 区块链是由一系列网络节点构建的一种分布式账本, 本身具有不可篡改性、去中心化、去信任化、密码算法安全性和不可否认性等安全属性, 对基于区块链实现的安全服务进行了综述, 这些安全服务包括数据机密性、数据完整性、身份认证、数据隐私、数据可信删除. 首先介绍了区块链和公钥密码学的基础知识, 并围绕上述 5 种安全服务, 给出了用户真实场景中面临的安全问题以及传统的解决方案, 讨论了这些传统实现方案所面临的问题, 之后介绍了使用区块链技术解决相关问题的实现方案, 最后讨论了区块链的价值以及面临的问题.

关键词: 区块链; 机密性; 完整性; 身份认证; 去中心化

中图法分类号: TP311

中文引用格式: 王利朋, 关志, 李青山, 陈钟, 胡明生. 区块链数据安全服务综述. 软件学报, 2023, 34(1): 1–32. http://www.jos.org.cn/1000-9825/6402.htm

英文引用格式: Wang LP, Guan Z, Li QS, Chen Z, Hu MS. Survey on Blockchain-based Security Services. Ruan Jian Xue Bao/Journal of Software, 2023, 34(1): 1–32 (in Chinese). http://www.jos.org.cn/1000-9825/6402.htm

Survey on Blockchain-based Security Services

WANG Li-Peng[1,2], GUAN Zhi[1], LI Qing-Shan[1], CHEN Zhong[1], HU Ming-Sheng[2]

[1](School of Electronics Engineering and Computer Science, Peking University, Beijing 100871, China)

[2](College of Information Science and Technology, Zhengzhou Normal University, Zhengzhou 450044, China)

Abstract: Blockchain is a distributed ledger constructed by a series of network nodes. It owns the following security attributes: unforgeability, decentralization, trustless, provable security based on cryptography and non-repudiation. This paper summarizes those security services, including data confidentiality, data integrity, authentication, data privacy, assured data erasure. This paper first introduces the concept of blockchain and public key cryptography. For the above-mentioned 5 security services, existing security threats faced by users in actual scenarios and their corresponding solutions are analyzed. The drawbacks of those traditional implementations are also discussed, and then countermeasures are introduced based on blockchain. Finally, values and challenges associated with blockchain are discussed as well.

Key words: blockchain; confidentiality; integrity; authentication; decentralization

1 概 述

区块链是一种能够提供数据可信服务的分布式账本技术, 无需第三方可信中心协助, 借助共识算法等策略形成群体决策, 进而实现可信数据服务功能. 区块链本身是一种去中心化的分布式 P2P 网络, 各节点之间地位对等, 通过以交易为载体, 实现数据通信和交换资源功能, 有效抵御针对中心节点的网络攻击行为, 同时有效解决了单点

* 基金项目: 国家重点研发计划 (2020YFB1005404); 国家自然科学基金 (61672060); 河南省科技攻关计划 (202102210359); 河南省高等学校重点科研项目 (20B520040)
收稿时间: 2020-09-17; 修改时间: 2020-11-26, 2021-02-17, 2021-05-05; 采用时间: 2021-06-03; jos 在线出版时间: 2021-10-20
CNKI 网络首发时间: 2022-11-15

故障问题. 虽然区块链可以作为一种数据库来存储数据, 但引入区块链的主要目的还是为了实现数据可信传递. 实际中可以对有价值的资产数字化, 将其存储到区块链中, 然后就可以利用区块链不可篡改、去信任、可信溯源等优良特性, 实现低风险、低成本的安全服务方案[1].

近些年来, 区块链技术得到了快速发展, 并已经成为学术界和工业界的研究热点. 2008 年 11 月 1 日中本聪首次提出了比特币概念, 作为区块链技术的第一个成功应用, 也是迄今为止最为流行的加密数字货币, 比特币常被视为一种数字黄金[2]. 截至 2020 年 7 月, 比特币市值已达到 1 958 亿美元, 占据了整个加密数字货币市场的 65%[3]. 除数字货币领域的应用, 近些年来, 区块链技术同时被广泛应用于监管科技、数字存证和溯源防伪等领域中. 例如摩根大通针对金融监管等领域面临的安全问题, 基于以太坊技术提出了 Quorum 联盟链平台, 很好地解决了隐私和性能等方面的问题[4]. 此外, 为实现对信息流、物流和资金流的全程监管, 由 IBM 主导开发了 Hyperledger Fabric 联盟链平台, 沃尔玛、Aetna 等公司基于该平台开发了一套供应链溯源系统, 在实现监管的同时有效地降低了成本[5].

根据 ISO 7498-2 标准给出的定义, 安全服务旨在为开放互联系统提供相应的数据安全传输服务, 该标准定义了 5 类可选安全服务, 即鉴别、访问控制、数据机密性、数据完整性、不可否认性[6]. 当前基于区块链技术能够有效实现上述安全服务, 但由于区块链平台自带不可否认特性, 可以很方便实现上述功能, 因此本文不再对不可否认安全服务进行重点论述. 本文重点讨论 4 种安全服务: 身份认证、数据隐私、数据完整性、数据机密性, 这几种安全服务事实上也对应到了上述安全标准涉及到的安全服务. 身份认证对应了 ISO 7498-2 标准中的鉴别服务, 包括数据源认证和对等实体身份认证, 数据源认证一般应用于无连接服务场景中, 而对等实体身份认证则一般应用于面向连接的服务场景中. 数据隐私安全服务包含了 ISO 7498-2 标准中的访问控制安全服务, 用于确保用户对自身隐私数据的可控性. 数据机密性用于确保数据不被窃听, 防止数据未授权访问, 一般通过加解密技术进行实现. 数据完整性一般通过数字签名技术进行实现, 用于校验原始数据是否被修改. 需要说明的是, 数据机密性和完整性是信息安全 CIA 三要素其中的两个, 当前出现的签密技术, 能够在一个逻辑步骤中同时实现上述两种功能, 而且要比单独实现上述两种功能的执行效率要高, 本文将对其进行讨论.

2016 年, 欧盟出台了《通用数据保护条例》(General Data Protection Regulation, GDPR), 其中指出数据所有者应该具有对自身数据修改、撤回、遗忘的权利[7]. 我国已于 2018 年 5 月 1 日生效的《信息安全技术个人信息安全规范》同样指出, 数据拥有者应该具有删除、收回数据的能力, 数据可信删除也是组成用户数据安全服务中的重要一环[8]. 数据可信删除可以确保数据所有者对自身数据的删除权, 数据所有者执行数据删除动作后, 任意第三方就不能再在从系统环境中获取被删数据任何有效的信息.

在本文中, 我们主要讨论基于区块链实现上述 5 种安全服务的相关研究, 即数据机密性、数据完整性、身份认证、数据隐私、数据可信删除. 表 1 总结了这些安全服务及其实现技术.

表 1 安全服务及其对应的实现技术

安全服务	定义	实现技术
身份认证	通过一定技术手段, 对用户身份进行确认, 确保只有合法用户才能授权访问相应服务的一种安全机制	加解密技术、数字签名技术等
数据隐私	数据拥有者通过定义相应的数据访问规则, 控制数据资源被其他人访问的行为	加解密技术、访问控制技术等
数据完整性	保证数据在生成、传输和访问时的正确性和一致性, 确保数据在生命周期内不被恶意篡改或破坏	数字签名技术、数据冗余备份技术等
数据机密性	利用加解密等安全技术, 保护数据免遭泄漏, 防止信息被未授权用户获取	加解密技术等
数据可信删除	确保用户在执行删除动作后, 服务提供商或其他授权访问该数据的用户不能从中获取任何有效信息	加解密技术、哈希算法等

传统的安全服务实现方案一般需要借助可信第三方中心节点, 因此可能会出现单点故障问题. 基于区块链的实现方案, 不仅能够提供去中心化的、可证安全的、协同共识的安全服务, 还能增强传统安全服务的效率和安全性.

本文重点讨论基于区块链技术实现安全服务的相关研究工作, 首先介绍了区块链和公钥密码学的基础知识, 然后给出了真实场景中用户面临的安全问题以及传统的解决方案, 并讨论了这些传统方案所面临的安全挑战, 之后介绍了使用区块链技术解决上述问题的实现方案, 文章最后给出了区块链价值以及其面临的问题.

本文第 2 节简要介绍了区块链体系结构、共识机制及其安全特性; 第 3 节简要介绍了公钥密码学相关概念以及 3 种传统的密钥管理技术; 第 4 节讨论了基于加解密和签名方法实现数据机密性和完整性的传统实现方案, 并给出了基于区块链的实现方法, 并在最后介绍了签密算法和基于区块链技术实现的签密算法; 第 5 节讨论了身份认证领域传统的实现方法和基于区块链的实现方法; 第 6 节讨论了隐私保护领域传统的和基于区块链的实现方法; 第 7 节讨论了基于传统方法和基于区块链实现数据可信删除的方法; 第 8 节重点介绍了区块链价值、其所面临的一些问题以及未来展望; 第 9 节对本文进行了总结.

2 区块链背景知识

在本节中, 首先介绍了区块链的系统架构, 然后重点解释了几种常见的区块链共识技术和区块链的安全特性, 同时对比了几种现有区块链系统, 重点讨论了各自不同的功能特性.

2.1 系统架构

区块链是由一系列网络节点构建的一种分布式账本技术, 其结构如图 1 所示. 区块链数据是以交易形式存储到区块中, 然后再由区块构成的一种链式结构. 区块链本身可以作为一种数据库被使用, 确保数据不被篡改, 但由于存储空间限制, 实际中一般存储关键数据信息. 需要说明的是, 当前研究人员已经实现了区块链数据可信编辑功能, 能够确保编辑过程中数据的安全性. 区块中的数据主要包括两个部分: 区块头和区块体. 区块头中存储的信息主要包括时间戳、随机数 (nonce)、哈希值等数据, 其中时间戳代表了区块创建时间, 而随机数主要帮助区块链参与节点达成一致协议或共识. 区块体中则主要存储了详细的交易信息, 也会允许用户附加自定义信息.

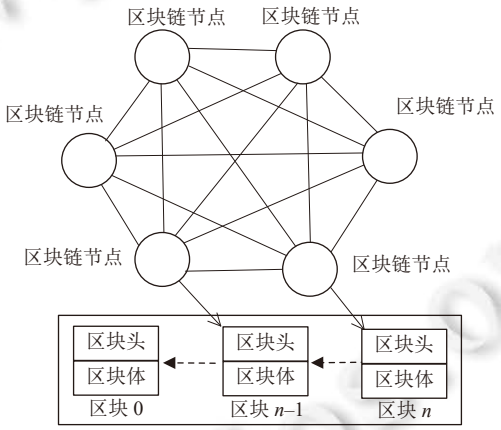


图 1 区块链结构图

区块链架构一般分为 3 层, 从下到上分别为底层服务层、核心层和应用层[9]. 底层服务层包含了运行区块链所需的硬件设施以及存储、P2P 网络服务等软件设施, 本层需要重点关注节点通信问题, 确保数据能够正确无误地传递到对方. 核心层是基于密码学算法、共识算法等技术构建的分布式账本体系, 通过汇聚区块链节点算力资源, 实现共享账本的冗余备份和数据验证, 确保账本数据视图全网一致. 应用层则是基于区块链 API 接口、智能合约等技术构建的各种应用程序案例集, 编程人员无需关注区块链底层实现细节, 只需要聚焦业务逻辑, 实现所需的功能程序.

区块链本身是一种分布式系统应用, 不需要借助第三方中心权威节点, 利用密码学算法、数据冗余备份等技术, 通过节点之间的交互过程, 对外提供一种去中心化的可信数据服务. 当区块链中的用户进行交互时, 它们可以

通过发送交易的形式进行实现. 首先区块链会按照某种共识机制选择某一节点来生成区块, 然后由该节点将交易信息打包进区块, 并打上时间戳, 方便后续实现溯源功能, 再将该区块广播到区块链网络中. 其他区块链节点收到后, 会校验该区块, 如果校验通过, 就可以将该区块添加到区块链中, 如果校验不通过, 区块链系统则会拒绝添加该区块. 存储到区块链中的区块及交易信息都经过了签名, 同时区块链系统节点会备份这些区块信息, 这样可以防止恶意节点未经授权篡改区块数据.

截止到目前, 区块链发展共经历了 3 代, 从原始的仅支持货币交易功能的区块链 1.0, 到后面支持智能合约功能的区块链 2.0, 直到最近提出的支持可编程社会的区块链 3.0[10]. 第 1 代区块链典型代表就是中本聪提出的比特币系统, 其应用场景仅限于数字货币领域. 第 2 代区块链技术的典型代表是以太坊, 该系统将比特币底层实现抽取出来, 形成区块链底层基础设施, 并在上层首次引入了智能合约概念, 能够为用户提供定制化服务, 从而将区块链摆脱了仅限于提供数字货币功能的约束, 极大地拓宽了区块链应用范围. 第 3 代区块链引入了可编程社会这一概念, 其核心目标是为了构建价值互联网, 利用区块链提供的优良特性, 提升社会服务效率. 截止到目前, 区块链技术已经广泛应用于身份认证、物流追踪、民主投票、司法公证、监管科技等领域中[9].

2.2 区块链中的共识机制

区块链利用共识机制来构建区块, 并在构建成功后将其存储到区块链系统中, 后续便可以用来存储相关的交易信息. 现有的区块链系统主要分为 3 类: 第 1 类是私有链, 第 2 类是联盟链, 第 3 类是公有链. 其中私有链和联盟链具有一定的准入机制, 只有授权用户才能接入网络, 而公有链则对外开放, 原则上任何节点都可以接入公有链来访问其中数据. 另外私有链和联盟链一般没有相应的挖矿奖励, 而公有链为了维护节点网络, 会为满足条件的矿工节点分配相应的奖励.

在公有链中, 只有第一个成功构建区块的节点才能获得相应的系统奖励, 这些奖励通常是一些虚拟数字货币, 比如比特币、以太币等, 这一过程称为挖矿过程. 挖矿机制是构成公有链的关键技术之一, 区块链节点利用挖矿机制来创建区块, 参与挖矿的节点称为矿工节点, 矿工节点只有以最快速度构建区块并通过验证, 才能赢得奖励.

区块链利用共识机制来达成系统一致的决策方案, 共识机制一般需要满足一致性、有效性以及终止性. 首先, 对于具有相同初始值的诚实节点, 共识算法需要保证诚实节点最终达成一致的决策, 其次如果所有诚实节点输入相同的数值, 那么共识算法能够保证诚实节点最终会选定该数值, 最后共识算法需要确保系统能够终止. 常见的适用于公有链的共识机制包括工作量证明 (proof of work, PoW)[11]、权益证明 (proof of stake, PoS)[12]、空间证明 (proof of space, PoSpace)[13]、重要性证明 (proof of importance, PoI)[14], 后来一些研究人员又提出了一些新的共识算法, 例如信任度证明 (measure of trust, MoT)[15]、最小哈希值证明[16], 这些共识算法并不常用, 这里不再赘叙. 当前联盟链中常用的共识机制, 主要包含了实用拜占庭容错机制 (practical byzantine fault tolerance, PBFT)[17]等, 表 2 对比了上述这些共识机制.

表 2 常见的几种共识机制对比结果

共识机制	提出时间 (年)	判定指标	适用场景	典型案例
PoW	1999	算力	公有链	比特币
PoS	2011	权益	公有链	以太坊
PoSpace	2014	存储空间	公有链	Permacoin
PoI	2018	节点重要性	公有链	NEM
PBFT	1999	无	联盟链	Hyperledger Fabric

(1) PoW: PoW 在刚提出时, 主要应用于垃圾邮件处理场景中. PoW 核心原理是分布式节点利用算力资源, 来竞争记账权利, 通过增加节点作恶成本, 实现数据一致性以及安全性. PoW 典型应用案例是比特币, 具体在实现时, 该共识算法会要求矿工节点解决一个数学难题, 只有最快解决该问题并通过验证的节点, 才能获得区块记账权并得到相应挖矿奖励. 详细过程如下: 矿工节点会利用校验通过的计算结果构建出一个区块, 然后将其发布到区块链网络中, 其他矿工节点再对其进行校验, 确保结果的正确性. 一旦该区块通过了系统验证, 就会被认为是合法区块,

之后就会被存储到区块链中, 而创建该区块的矿工节点会得到相应的系统奖励. 如果恶意节点要攻击该共识机制, 就需要控制全网超过 50% 的算力资源, 由于实际场景中区块链节点数量众多, 而且异地分布, 很难实现上述目的, 这就意味着该共识机制是安全的. 需要说明的是, 在 PoW 共识机制中, 解决数学难题的过程会耗费大量的计算资源, 绿色共识算法是区块链未来发展的一个重要方向.

(2) PoS: PoW 机制会消耗较多的计算资源, 针对该问题, 相关研究人员提出了 PoS 共识算法. 在 PoS 实现机制中, 区块链会依赖节点拥有的权益多少, 例如代币多少, 并采用一种伪随机策略选择对应的矿工节点来构建区块. 如果某一节点拥有的权益越多, 被选中的机会就越高. PoS 能够有效缩短区块链节点达成共识的时间, 提升系统吞吐率. 然而 PoS 机制也存在一定的问题, 如果每次都选择权益最多的节点来创建新区块, 那么就有可能会出现收益分配不均衡问题, 甚至会出现某些节点垄断挖矿收益的现象, 因此 PoS 后续版本对此进行了一些改进, 在执行相应的奖励或惩罚操作时, 会将节点日常表现考虑进去[12].

(3) PoI: 在 PoI 机制中, 系统会根据个体在群体经济活动中的活跃度, 来为其赋予不同优先级, 系统会选择优先级最高的节点来创建区块, 并获取相应的挖矿收益[14]. PoI 克服了 PoS 机制存在的“富人愈富, 穷人愈穷”的问题, 而会依赖信誉评分机制, 来分配挖矿奖励.

(4) PoSpace: PoSpace 是为了解决 PoW 消耗算力资源较多的问题而提出的一种解决方案, 该方案依赖节点拥有的资源多少进行判定. 所不同的是, PoW 衡量的对象是算力资源, 而 PoSpace 则是存储资源. 在 PoSpace 机制中, 矿工节点是通过证明自身拥有的存储空间的大小, 来竞争挖矿收益. 需要说明的是, 基于 PoSpace 机制实现的区块链系统, 一般会面临存储资源消耗过高的问题.

(5) PBFT: PBFT[17]是一种基于拜占庭容错方法实现的区块链共识机制, 将原始拜占庭容错方法计算复杂度从指数级降低到多项式级别, 有效提升了系统效率, 因此该共识机制常被应用于联盟链系统中. 该机制的实现细节如下, 首先会从系统参与节点中选择一个领导节点, 该领导节点会决定交易的验证方式, 并向其他节点广播区块信息, 区块只有经过 2/3 以上节点验证通过后, 才能被写入到系统. 在某一时刻, 领导节点可视为一种中心化节点, 然而由于领导节点的更换过程是非常频繁的, 因此 PBFT 仍可视为一种去中心化的共识机制. PBFT 机制达成共识速度快, 但由于每个节点都需要与其他节点进行通信, 随着节点数量增加, 性能会急剧下降, 因此 PBFT 共识机制比较适合于节点数量较少的场景中.

2.3 区块链的关键安全特征

区块链的关键安全特征包括不可篡改性、可信溯源、去中心化、去信任化、密码算法安全性和不可否认性, 表 3 简要概括了这些安全特征的定义以及区块链实现机理[18].

表 3 区块链技术的关键特性

特性	含义	区块链实现机理
不可篡改性	保证系统中保存的信息不被恶意修改	基于数据冗余备份和可证安全的密码学算法等策略进行实现
可信溯源	可信追踪数据来源和过程信息	区块链能够将数据流通的各个环节关联起来, 打破数据壁垒, 将数据以及时间戳等过程信息记录到系统中, 结合区块链不可篡改特性, 实现数据可信溯源
去中心化	构建一种扁平化、平等化的分布式节点组织框架. 需要说明的是, 去中心化不是不要中心, 而是任意节点均有可能成为中心, 而且也不是永久固定的	基于共识算法实现群体决策, 减少传统中心化实现策略而引入的安全问题
去信任化	系统参与方无须建立信任关系就可以安全地完成交易任务	区块链技术采用密码学、共识算法等多种机制来实现群体决策, 攻击者很难破坏决策的可信性
密码算法安全性	密码学算法需要通过安全性证明验证, 能够满足相关的安全评价指标要求	区块链一般采用通过安全性验证同时业内公认安全的密码学算法确保系统的安全性
不可否认性	用户不可以否认其与系统的交互行为	基于签名算法, 并结合区块链不可篡改特性, 确保用户不可否认其与区块链系统发生的历史交互行为

2.4 区块链系统的不同实现版本

目前出现了许多开源的区块链系统, 需要根据不同的应用场景来选择对应的实现系统, 表 4 从编程语言、优缺点方面对比了当前几种比较热门的区块链系统.

表 4 不同区块链系统对比结果

平台	编程语言	优点	缺点
以太坊[19]	Solidity, Serpent等	(1) 可扩展性好 (2) PoS不需要消耗大量算力	基于PoS实现的系统会偏好选择具有较多权益的节点来创建新区块
BigchainDB[20]	JavaScript, Python	(1) 查询速度快 (2) 编程难度较低	没有提供智能合约功能
NEM[21]	Java	(1) 引入了命名空间的概念, 可以隔离不同的业务 (2) 达成共识速度较快	只有拥有一定数量虚拟货币的节点才能执行挖矿操作
Hyperledger Fabric[22]	Chaincode	达成共识的速度较快	可扩展性较差
EOS[23]	C++	出块时间快	可扩展性较差

2.5 总 结

区块链是一种分布式账本技术, 起源于比特币, 所具有的关键安全特征包括可信溯源、不可篡改性、去中心化、去信任化、密码算法安全性和不可否认性, 这些安全特征是构成区块链数据安全服务关键所在. 我们讨论了几种不同的区块链共识机制, 例如 PoW、PoS、PoSpace、PoI 和 PBFT 等共识机制, 用于实现区块链去中心化决策, 这些共识机制具有不同的特性和适用范围, 开发人员需要根据不同的需求来选择对应的共识机制.

3 公钥密码学

公钥密码学, 一般也称为非对称密码学, 需要系统为参与用户生成一对公私钥, 同时依赖一套密码学基础设施来创建、撤销、分发、使用和存储相关密钥信息[24]. 对称加密算法中消息发送方和接收方采用相同的密钥来执行加解密运算, 在分发密钥时一般仍需要依赖公钥密码学工具进行实现. 本文讨论的几种安全服务大部分都是基于非对称密码学进行实现, 因此重点对其进行讨论. 公钥密码学中需要重点解决密钥管理问题, 这也是当前基于公钥密码学实现安全服务所要考虑的核心问题, 本节将讨论当前常见的几种密钥管理技术. 本节最后介绍了当前区块链应用中比较常用的两种密码学算法, 即零知识证明和同态加密算法, 被广泛应用于密态内容审计、用户隐私保护等领域中.

3.1 公钥密码学

1975 年, Diffie 和 Hellman 首次提出了公钥密码学的概念, 该密码学理论需要借助一对公私钥来完成相关操作, 其中公钥信息可以对外发布, 而私钥信息则需要秘密保存. 公钥密码学的基本思想是使用其中一个密钥执行加密或签名任务, 然后使用另一个密钥执行解密或签名校验任务. 用户可以利用该机制校验某一信息是否被篡改, 也可以利用该机制加密待发送的数据内容, 数字签名和加解密服务的实现架构如图 2 所示. 以数字签名为例, 用户在发送消息时, 首先利用自己的私钥对消息进行签名, 接收方收到签名信息后, 利用发送方公钥来校验数据和签名信息是否一致, 进而判断数据内容的完整性[25].

当前已经出现了多种公钥密码学算法, 例如 RSA 算法[26]、ElGamal 算法[27]、椭圆曲线密码学算法[28], 包括现在比较流行的零知识证明和同态加密算法等, 这些公钥密码学算法共同组成了当今应用系统的安全基础. 需要说明的是, 这些算法需要依赖相应的密码学基础设施来管理用户的公私钥信息, 其中包括基于公钥基础设施 (public key infrastructure, PKI) 密钥管理机制、基于身份密码学 (identity-based cryptography, IBC) 密钥管理机制、基于无证书公钥密码学 (certificate-less public key cryptography, CL-PKC) 密钥管理机制, 我们将会在后面分别讨论这些方案.

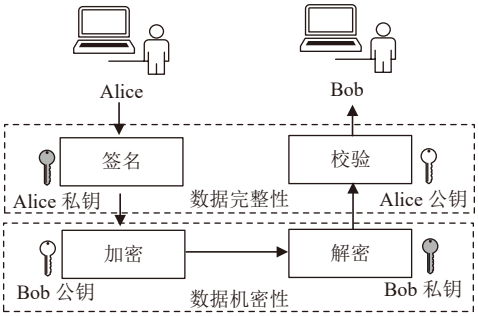


图 2 基于公钥密码学实现数据机密性和完整性原理图

3.2 基于公钥基础设施密码学

基于公钥基础设施密码学, 也就是 PKI, 是公钥密码学引入的一种密钥管理机制, 它是一个包括软件、硬件、人员、规则的集合, 用来管理密钥以及证书的生成、存储、分发、撤销等过程. PKI 出现较早, 发展的也较为成熟, 已经为当前大部分应用提供安全服务, 例如电子商务、Web 服务等. 一个典型的 PKI 系统, 主要包括软硬件设施、证书管理机构、应用系统、注册机构等, 其中证书管理机构负责证书的签发, 是 PKI 核心, 一旦不可用, 会导致整个 PKI 系统瘫痪.

截止到目前, 主要存在 3 种实现 PKI 机制的方法: 第 1 种是中心化证书颁发机构 (certificate authority, CA) 方法, 另一种是分布式的 WoT (Web of trust) 方法, 第 3 种是 SPKI (simple public key infrastructure) 方法. 基于 CA 的 PKI 实现机制较为常见, 而且成熟方案较多[29]. 具体实现时, 首先选择一个系统所有成员都信任的中心化 CA 节点, 它会为系统用户颁发证书, 该证书绑定了用户公钥等信息. 然而基于 CA 的 PKI 实现机制, 存在单点故障等问题, 后来在 1992 年, 相关研究人员提出了 WoT 密钥管理技术, 这项技术采用了一种分布式策略来管理用户密钥信息. 其执行流程如下: 首先用户在本地生成自身的密钥信息, 然后再将密钥信息发送至系统进行校验, 如果系统中至少一个可信节点校验通过, 则意味着密钥信息正确无误, 同时会被整个系统所信任[30]. SPKI 是一种适用于分布式场景中的公钥证书管理机制, 它不需要依赖中心化节点来管理密钥信息, 同时允许用户灵活地定义和管理权限信息.

3.3 基于身份密码学

近些年来, 相关研究人员提出了基于身份密码学的概念. 基于身份的密码学, 也就是 IBC, 同样属于公钥密码学体系. 在 IBC 中, 系统会利用能唯一标注用户身份信息的数据来生成公钥信息, 这些数据包括用户姓名、手机号等, 这样就可以简化证书认证过程. IBC 实现数据机密性的流程如图 3 所示, 主要包括了 4 个步骤: 初始化、密钥提取、加密和解密.

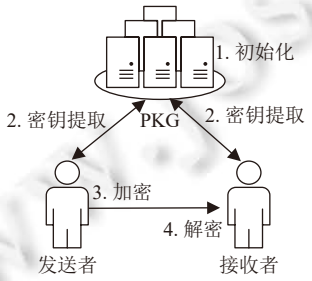


图 3 基于 IBC 实现数据机密性架构图

在 IBC 中, 需要依赖私钥生成器 (private key generation, PKG) 来生成并管理用户密钥信息. 在初始化阶段, PKG 首先生成系统主密钥以及系统公开参数, 并秘密保存系统主密钥. 当用户需要生成自己的密钥信息时, 系

统执行密钥提取步骤, PKG 会利用系统主密钥、用户身份信息以及系统公开参数来生成用户公私钥, 并将其发送给用户. 当用户需要发送消息时, 会利用用户身份信息等数据对明文信息进行加密, 并 1 将密文发送出去. 接收方收到密文信息后, 再使用其自身私钥来解密消息. IBC 除适用于加解密领域, 还能适用于签名领域中. 相关研究人员基于 IBC 机制实现了一种签名方案[31], 数据发送方利用自身密钥生成消息对应的签名, 后续用户便可以利用身份信息等数据完成签名校验操作. 后来相关研究人员提出了一种新的树形 IBC 实现模型, 将用户身份信息层次化, 并基于该层次化的身份信息生成一种树形公钥结构, 该方案允许用户为其孩子节点生成相应的密钥信息[32].

3.4 无证书公钥密码学

无证书公钥密码学, 也就是 CL-PKC, 是由文献 [33] 提出的一种公钥密码学机制. 在 CL-PKC 中, 用户密钥信息由密钥生成中心 (key generation center, KGC) 和用户协作生成, 任意一方均不能单独生成成功, 这样可以避免 KGC 或用户任意一方作恶. CL-PKC 本身并不需要可信中心节点来管理用户密钥信息, 因此有效解决了密钥托管问题, 进而提升了系统安全性.

以基于 CL-PKC 实现数据加解密功能为例, 首先 KGC 执行初始化步骤, 生成系统主密钥和公开参数信息, 系统主密钥一般由 KGC 秘密保管, 而系统公开参数信息则需要公开发布, 恶意用户并不能够从中获取任何有价值的信息. 然后 KGC 与用户协作生成部分私钥以及部分公钥信息, 其中部分私钥信息经由安全通道发送至用户. 用户会在本地端生成秘密信息, 然后再结合 KGC 发送过来的部分私钥以及部分公钥, 合成用户公私钥信息, 后续用户就可以利用公私钥对消息进行加解密. 需要说明的是, 实际中同样可以采用 CL-PKC 实现数字签名功能, 实现步骤同加解密步骤类似, 这里不再赘述.

3.5 零知识证明

零知识证明是一种保护用户隐私的密码学概念, 证明者可以在不出示秘密的前提下, 能够向验证者证明自己知道该秘密信息. 零知识证明本质上是一种两方或更多方, 通过协作来完成某一任务所要求的一系列步骤, 其中验证者将需要验证的信息, 通过计算转化为一系列的挑战信息, 证明者利用秘密以及其他一些信息, 来完成这些挑战. 若证明者能够完成全部挑战信息, 则验证者可以相信证明者的确拥有该秘密信息.

设定两个通信实体 P 和 V, 分别代表证明者和验证者, 如果 P 成功向 V 证明了其拥有的信息后, V 却不能推断出秘密信息, 则意味着 P 实现了最小泄露证明, 这也体现了零知识性. 此外零知识证明还需要满足正确性, 即若 P 无法掌握一种定理的证明方法, 使 V 信任 P 掌握了秘密信息的概率是很低的. 最后零知识证明需要满足完备性, 即 P 掌握了一种定理证明方法, 能够以绝对优势概率使 V 相信 P 能够完成证明.

3.6 同态加密算法

同态加密算法能够保证对密文执行运算的结果, 跟对明文进行运算再加密得到的结果是相同的, 也即同态性. 对于一个加密函数 $enc()$, M 为明文集合, 若对于 $\forall m_1, m_2 \in M$ 和有效算法 $\odot$, 如果能够满足 $enc(m_1) \odot enc(m_2) = enc(m_1 \odot m_2)$, 则意味着该加密函数 enc 具有同态性. 如果存在有效算法 $\oplus$, 能够满足 $enc(m_1 + m_2) = enc(m_1) \oplus enc(m_2)$, 则意味着该加密函数 enc 具有加法同态性. 如果存在有效算法, 能够满足 $enc(m_1 \times m_2) = enc(m_1)enc(m_2)$, 则该加密函数满足乘法同态. 如果函数 enc 同时满足加法同态和乘法同态, 则称其为全同态加密.

同态加密算法能够有效应用于内容审计、数据共享等领域中. 当前出现的 Paillier 算法, 是典型的加法同态算法, RSA 算法属于乘法同态算法, 而 BGN 算法则属于全同态加密算法. 然而, 在将同态加密算法应用于实际领域中时, 性能是其主要瓶颈之一, 如何提升同态加密算法的性能, 是需要重点关注的因素.

3.7 总 结

公钥密码学, 又称为非对称密码学, 是密码学领域一个重要分支, 已广泛应用于信息安全领域中. 基于公钥密码学, 研究人员提出了相应的密钥管理技术, 用于生成、分发和撤销用户密钥信息, 这些技术包括 PKI、IBC 以及 CL-PKC. 在本节中, 我们首先讨论了公钥密码学相关概念, 并分别介绍了 PKI、IBC、CL-PKC 概念和一般执行流程, 表 5 对这些技术进行了总结. 由于零知识证明和同态加密算法在当前区块链应用中愈加广泛, 本节对其进行了介绍.

表 5 公钥密码学几种实现机制对比

实现机制	实现特性	公钥产生方	主要存在的问题
PKI	基于可信CA节点来管理用户公钥信息.	CA	1. 公钥管理过程较为复杂, 证书发布、验证、存储等需要消耗较多资源 2. 基于CA实现方法存在单点故障问题, 一旦CA被攻破, 会破坏整个系统安全性 3. 基于WoT的实现方法, 用户节点很难实现统一的安全防护措施, 而且维系节点信任关系会消耗较多的计算和带宽资源
IBC	利用用户身份信息来生成用户公钥信息.	PKG	1. 存在单点故障问题 2. 密钥托管问题 3. 撤销证书较为困难
CL-PKC	KGC和用户任意一方均不能独立生成用户公私钥信息, 必须协作生成.	KGC和用户	KGC和用户之间需要构建秘密通道协作生成用户公私钥, 会消耗一定的计算和带宽资源

4 数据机密性和完整性

数据机密性和完整性是构成信息安全 CIA 三要素中的两个要素, 也是一个应用系统所必需的两个安全服务. 以智能医疗为例, 该场景需要实现传输数据的机密性, 以保护患者隐私信息不被泄露. 此外, 还需要保证数据在传输过程中不被恶意篡改, 确保接收方能够正确获得医生、患者、医院等节点发送的信息, 因此需要实现数据的完整性[34]. 在现实场景中, 这两种安全服务一般可以通过公钥密码学算法或对称加密算法进行实现, 其中公钥密码学算法最为常见. 传统实现数据机密性和数据完整性的操作是相互独立的, 当前出现的签密技术, 能够在一个逻辑步骤中同时实现上述两种安全服务, 具有较高的执行效率, 本节将重点对其进行讨论.

4.1 基于 PKI 实现数据机密性方案

基于 PKI 实现数据机密性相关研究的时间较长, 当前已经出现了诸多成熟的实现方案, 本节则主要介绍一种应用于 Web 服务上的 PKI 实现方案, 用来加密网络流量数据.

SSLP (secure sockets layer protocol) 协议最早由 Netscape 公司提出[35], 其中采用了基于 PKI 机制实现的数据加解密技术, 来保证数据在 Internet 上传输的安全性, 同时确保数据不为恶意第三方截取或侦听. 当前最新的版本是 3.0, 已经被广泛应用于 Web 浏览器和网络服务器之间数据传输场景中, 实现原理如图 4 所示. 首先 Web 服务器需要向 CA 服务器申请 SSL 证书, CA 会对申请信息进行验证, 校验通过后, 会将证书颁发给 Web 服务器. 当用户需要访问 Web 服务器时, 服务器会将自己公钥信息发送给用户. 用户收到上述信息后, 首先生成一个对称加密密钥, 再利用收到的公钥信息对其进行加密, 并将生成的密文发送给 Web 服务器. 服务器收到密文后, 利用自己私钥解密上述信息, 进而得到对称加密密钥, 后续 Web 服务器和用户就可以利用该对称密钥加密消息.

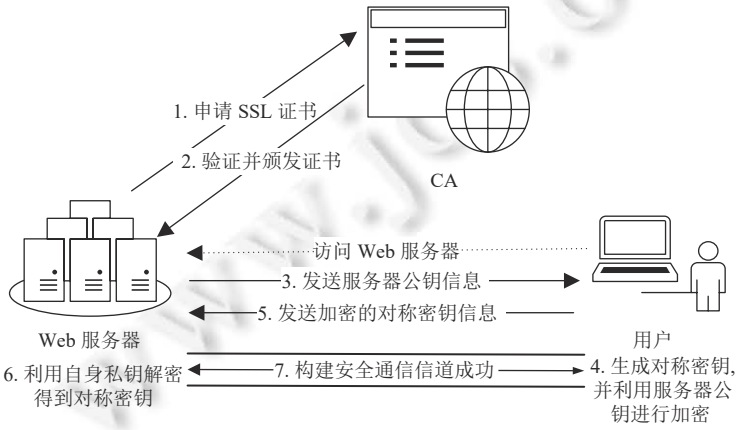


图 4 SSLP 实现机制

4.2 基于区块链和 PKI 实现数据机密性

区块链本身具有不可篡改特性, 利用该安全属性, 可以直接将数据存储到区块链系统中, 进而增强数据的可信性. 受限于区块链有限的存储空间, 上传的数据需要严格控制大小. 由于区块链中的数据对系统节点可见, 一般需要对敏感数据进行加密, 然后再将密文数据或元数据信息上传到区块链系统中, 进而实现数据安全共享.

IKP (instant Karma PKI)[36]是一种能够报告 CA 行为的自动化平台, 能够激励 CA 正确分发证书, 同时帮助系统成员快速反馈未授权的证书分发行为. IKP 将 CA 的行为日志记录到区块链系统中, 基于区块链不可篡改特性, 确保日志信息的可信性. 一旦系统检测到 CA 的异常行为, IKP 便会启动相应的安全防御措施, 并基于日志信息溯源 CA 历史行为. 该方案解决了传统 PKI 实现机制过于依赖第三方中心节点的问题, 同时能够实现可信溯源功能. 需要说明的是, 日志信息中可能会包含一些敏感信息, 因此在将 IKP 系统应用到实际场景时, 需要对其中的敏感数据进行脱敏处理, 例如可以采用同态加密算法进行实现, 这样可以在不丢失统计特征的前提下, 实现数据机密性.

IKP 项目最早基于以太坊进行实现, 具有较低的构建成本, 兼具公开可验证性等特性. 由于 IKP 系统需要利用 CA 来分发证书, 故仍存在单点故障问题.

4.3 基于区块链和 IBC 实现数据机密性

基于区块链的 IBC 实现方案可以将用户公钥存储到区块链中, 基于区块链不可篡改特性, 确保用户公钥信息的真实可信. 由区块链承担传统 PKG 角色, 能够基于区块链去中心化特性, 很好地解决单点故障问题, 还可以对密钥生成过程进行审计, 实现数据可信溯源. 下面介绍基于 IBC 和区块链技术实现数据机密性的方案, 并讨论了其中可能存在的问题.

适用于信息中心网络 (information-centric networking, ICN) 的实现方案迎合了用户获取内容或服务的需求, 将互联网使用模式从以位置中心转为以内容或服务为中心. 由于 ICN 对数据内容与终端位置进行了剥离, 用户可以基于内容来查询和访问相应的信息, 而不是利用传统 IP 地址来访问数据, 因此 ICN 中安全问题主要集中于内容安全这一块. 利用区块链和 IBC 技术, 可以保护 ICN 中的关键数据, 确保数据不被恶意篡改或者未经授权访问.

ICN 中的内容要素被设计为一种分层结构, 基于该特点, 文献 [32] 提出了一种 HIBC 方案来增强 ICN 的安全性. 文献 [37] 利用区块链技术和 IBC 机制为 ICN 网络设计了一种分布式数据共享系统, 允许数据所有者授权其他用户来访问自身数据. 用户历史操作都会被记录在区块链中, 确保用户无法否认自己的行为信息.

4.4 传统的数据完整性实现技术

数据完整性能够保证数据在生成、传输和访问时的正确性和一致性, 确保数据在生命周期内不被恶意篡改或破坏[38]. 实现数据完整性是十分有必要的, 当前涉及数据存储、处理和检索的应用系统, 会存在类型众多的通信实体, 各实体之间存在复杂的交互逻辑, 此外在通信过程中可能还会涉及第三方监管机构. 数据流通途径多样化, 使得用户数据面临的攻击面更为宽泛和隐蔽, 更容易遭受各种各样的网络攻击, 这将会造成严重的后果. 例如, 在智能医疗系统中, 如果攻击者篡改了传感器数据, 会导致出现严重的医疗事故. 此外对于工业物联网、智能武器装备系统, 一旦数据遭受恶意篡改, 会对国家安全造成严重危害[39], 因此如何保障信息的完整性就变得至关重要.

数据完整性主要分为两类, 即物理完整性和逻辑完整性[40]. 物理完整性面临的挑战主要包括自然灾害、断电、设备故障、化学腐蚀等其他极端环境因素, 实际中可以采用冗余设备、UPS (uninterruptible power supply)、可修复芯片、可纠错存储器、带容错功能的文件系统等措施来实现数据完整性. 逻辑完整性则主要保护数据结构的正确性和合理性, 在关系数据库中, 则主要体现在引用完整性、实体完整性、域完整性以及其他用户自定义完整性.

传统的两种实现数据完整性技术是密码学技术和数据副本策略, 其中密码学技术主要基于 PKI、IBC 等密码学理论进行实现, 具体实现技术包括消息认证码 (message authentication code, MAC)、哈希树等. 系统可以利用签名技术检测数据是否被篡改, 如果攻击者想要绕过检测技术, 就必须借助密码学技术来伪造数据的签名信息, 对于设计良好的签名算法, 攻击者采用的一种手段就是利用公开信息计算用户的私钥信息, 然后再对修改后的数据进行签名, 而这在实际中是很难实现的. 数据副本策略是另外一种保障数据完整性的手段, 攻击者想要篡改数据, 则必须要查找整个系统, 修改所有的副本信息, 实际中也是很难实现的. 数据副本策略本身是一种以空间换安全的实

现手段, 分布式存储系统为了节省存储空间, 一般会采用纠删码技术来实现副本策略[41]. 在实际应用时, 一般会综合利用密码学技术和数据副本策略来实现数据完整性.

传统的数据完整性校验技术, 一般存在下面一些问题. 首先传统的实现方案一般是一种事后处理策略, 只有数据篡改动作发生之后, 系统才能检测出相关修改信息. 此外对于一些设计较为复杂的系统, 数据篡改行为可能发生在数据采集、传输、处理或者存储中任何一个或几个阶段中, 在构建完整性检测模型时, 需要覆盖数据流通的各个路径, 实现过程较为复杂. 存在的另一个问题就是, 负责执行完整性校验的节点, 有可能被黑客攻击并控制, 进而篡改校验结果. 此外如果上述节点设计为一种中心化架构, 这会引入单点故障问题. 最后数据完整性校验操作会消耗一定的计算资源, 同时会增加系统设计时的复杂度, 甚至会影响业务处理效率.

4.5 基于区块链实现数据完整性技术

区块链技术内嵌数据完整性校验机制, 能够有效解决前文所述问题. 在某种意义上来讲, 区块链技术也是通过密码学技术和数据副本策略来实现数据完整性. 用户在发起交易时, 区块链会为每笔交易进行签名, 然后再将这些信息发送给矿工节点. 矿工节点收到交易数据后, 会对其进行校验, 确保数据的完整性. 区块链系统只会存储通过校验的交易数据, 而且数据一旦被区块链接收后, 将很难再被修改. 表 6 对比了传统的以及基于区块链实现数据完整性的技术原理和存在的问题, 后面再讨论几种使用区块链技术实现数据完整性的应用系统.

表 6 传统数据完整性实现技术以及基于区块链的解决方案

实现技术	实现原理	存在的问题
密码学技术	可证安全的签名机制	1. 需要消耗额外的计算资源 2. 校验节点作恶 3. 密钥信息泄露 4. 基于中心化的实现方案存在单点故障问题
数据副本策略	数据冗余机制	1. 需要耗费额外的存储空间 2. 无法溯源攻击行为
区块链技术	内嵌数据完整性验证功能, 确保数据不可篡改, 同时可以溯源攻击者身份	1. 需要进一步提升效率

(1) PPDP: PPDP 是 Wang 等人利用区块链技术提出的一种可证数据完整性校验方案, 允许校验方不需要下载完整数据即可完成校验操作, 主要用于云计算场景中[42]. 云服务器在存储用户文件时, 一般需要对文件进行切片, 然后再将切片数据上传到服务器端. 随着切片数据增加, 传统的数据校验方案性能会随之下降. 为了解决上述问题, PPDP 提出了一种减少切片数量的策略, 以减少带宽和计算资源消耗, 系统实现架构如图 5 所示.

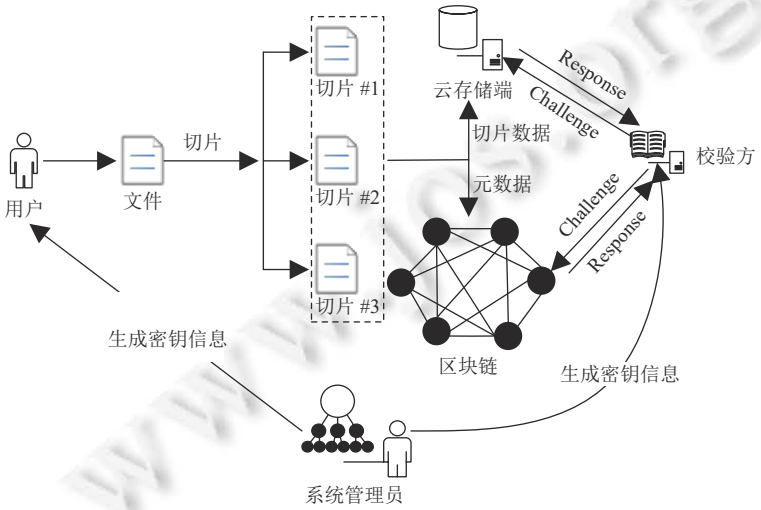


图 5 PPDP 方案框架图

下面介绍 PPDP 执行流程, 首先执行初始化步骤, 系统管理员会为用户和数据校验方生成相应的密钥信息. 当用户存储文件时, 系统会自动将文件切片, 并将切片连同哈希值存储到云端, 然后再将哈希值、签名信息、元数据等存储到区块链中. 当执行数据完整性校验操作时, 校验方发送 Challenge 信息给云端和区块链端, 云端和区块链端再计算 Response 信息发送至校验方进行信息验证. PPDP 不需要下载所有的原生数据, 而只需对数据量较小的 Response 信息校验即可, 执行速度较快.

(2) Zeppar: Zikratov 等人利用区块链技术提出了一种适用于分布式环境中数据共享系统, 能够在保证数据完整性的同时, 实现数据访问控制、数据审计等功能[43]. 当前存在的一些数据校验系统, 引入了第三方审计者 (third party auditor, TPA) 的概念, TPA 能够校验云端数据的完整性, 一旦发现数据被修改, 就会发出预警, 以便其采取相应的安全防御措施. 这种实现策略存在的问题就是系统需要维护一个安全信道, 来保证 TPA 可信访问数据.

Zeppar 方案将系统分为前端和后端, 前端是一个 Web 应用, 允许用户上传或下载数据, 后端则提供了可信数据存储服务, 用户只有通过身份认证并获得会话密钥后, 才能访问数据. 后端会监听操作系统中文件操作动作, 包括创建、修改、删除和重命名等, 然后创建并发送交易到区块链系统中. 交易信息包括了文件哈希值、时间戳、文件路径、文件所有者签名等信息, 用户可以利用这些信息, 对文件完整性进行校验. 该系统基于私有链进行实现, 支持动态文件完整性校验操作.

(3) Liu 方案[44]: Liu 等人[44] 基于区块链提出了一种适用于物联网场景中的数据完整性校验框架, 利用智能合约来实现数据安全共享. 在物联网场景中, 数据所有者共享数据时, 需要将数据发布到云上, 然后其他用户就可以从云上访问相应的数据. 在该框架中, 数据所有者会将加密后的数据哈希值以智能合约或者交易的形式发布到区块链系统中, 再将物联网数据上传到云端, 之后便可以授权其他用户来访问上述数据. 当需要校验数据完整性时, 用户会向云端发出请求, 云端获取到请求后, 再将物联网数据发送给用户. 当用户获取到上述数据后, 会计算数据对应的哈希值, 再向区块链查询上述数据对应的哈希值, 如果区块链上存储的哈希值与本地计算得到的哈希值相匹配, 则意味着数据未被篡改.

该系统基于私有链进行实现, 不需要借助 TPA 即可实现数据完整性校验. 随着客户端数量增多, 该系统的执行效率比传统实现方案要高, 而且数据拥有者能够获得相应的报酬. 需要说明的是, 用户可能仅是一些结构简单的物联网设备, 计算能力较弱, 无法执行耗时的数据校验操作, 一种可行的方法就是将计算哈希值操作和完整性校验操作全部委托给区块链, 由区块链在空闲时再执行上述操作.

(4) Storj: Storj 是一个基于区块链实现的 P2P 数据存储系统, 同时兼容 S3 存储协议, 允许用户共享自己的硬盘空间和带宽资源, 并可以从中获得一定报酬[45]. 与前面所述系统类似, Storj 同样是将数据的哈希值存储到区块链系统中, 利用区块链不可篡改特性确保哈希值的可信性, 然后再基于该哈希值执行数据完整性校验操作, 其架构如图 6 所示.

Storj 系统的执行步骤如下, 首先数据拥有者在本地对数据进行加密, 确保数据内容不会被其他人知悉. 系统再对加密后的文件进行切片, 并将切片数据离散存储到云端数据库中, 同时用户需要为此支付一定的费用. 系统采用纠删码算法来管理数据切片, 用户可以调整算法参数以便在成本和可用性中进行权衡. 系统会将元数据信息存储在区块链中, 其中元数据包含文件哈希值、切片存储位置以及 Merkle 根值等信息. 当用户想要访问数据时, 区块链系统会利用元数据信息来校验云端数据完整性, 校验通过后, 区块链再将数据位置信息发送给用户, 用户就可以从云端获取到相应的数据.

Storj 是一个开源系统, 并基于数据副本策略和密码学算法来保证数据完整性. 需要说明的是, 该系统已经商用, 并在实际业务场景中取得了良好的效果.

(5) Ericsson: Ericsson 是另外一种提供数据完整性服务的系统, 能够保存用户的数字资产[46]. 该系统会利用 KSI 为数据生成无密钥签名信息, 其中 KSI 是在 2006 年被提出的一种签名技术, 该技术会利用哈希树和时间戳为多个文档进行签名. 当需要生成签名信息时, 系统会利用 SHA-256 算法生成固定长度的哈希值, 再结合其他信息生成唯一签名信息. 当需要校验签名时, 用户端会将签名发送给服务端来获取数据对应的哈希值, 并与本地计算得到的哈希值进行匹配. Ericsson 提供了 RESTful API, 方便用户调用系统功能.

KSI 最初的实现版本, 需要依赖可信中心节点来构造哈希树和签名信息, 会存在单点故障问题. 后来 Guardtime 基于区块链技术提出了一种去中心化的 KSI 方法[47], Ericsson 方案采用了这种改进后的 KSI 方法, 实现了一种去中心化的分布式实现架构, 克服了单点故障问题[48]. Ericsson 系统代码已经开源, 开发人员能够免费利用该代码来快速构建系统.

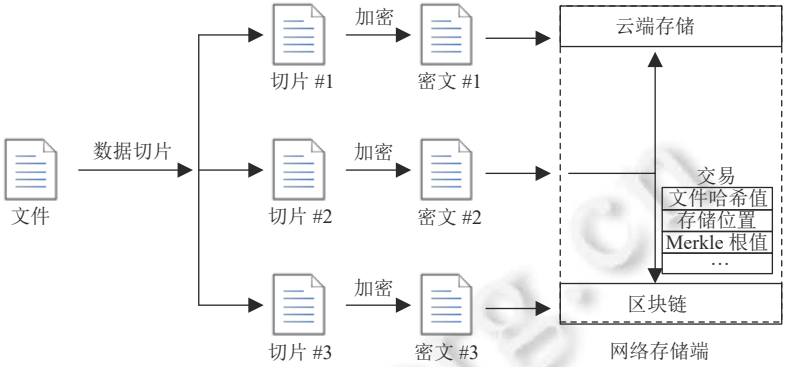


图 6 Storj 方案框架图

4.6 传统的无证书签密技术

近些年来, 无证书公钥密码学, 也就是 CL-PKC, 已经被广泛应用到无证书签名、无证书密钥协商、无证书身份认证等领域中. 文献 [49] 则首次将 CL-PKC 的概念引入到签密领域中, 提出了无证书签密算法. 当前存在的一些无证书签密算法, 按照操作类型主要分为 3 种, 即基于离散对数实现方案、基于椭圆曲线实现方案以及基于双线性对实现方案. 一般来讲, 这 3 种签密实现方案, 其对应的攻击难度依次上升, 而其对应的入门难度也随之上升. 后量子时代, 相关研究人员又提出了一些能够抗量子攻击的签密方案, 主要包括基于哈希的签密方案[50]、基于编码的签密方案[51]、基于格的签密方案[52]以及基于多变量的签密方案[53]. 抗量子攻击的签密方案不是本文的重点, 这里不再对其进行论述.

传统的签密方案主要是一对一的通信模式, 也就是发送者一次只能将一条消息发送至指定的一个接收者. 随着广播通信技术的发展, 发送者往往需要一次向多个接收者发送多条不同的消息, 同时只有授权用户才能获取对应的消息, 进而出现了多接收者多消息签密算法. 在这种方案中, 发送者只需要执行一次签密过程, 即可生成密文信息, 并将其广播出去. 接收者在收到消息后, 会利用自身私钥等信息来解密消息内容.

目前衡量签密方案的安全指标主要包括了机密性、不可篡改性、用户身份匿名性以及公开可验证性等. 当前主要存在两种安全模型来形式化验证这些安全指标, 第 1 种是标准模型, 第 2 种是随机预言机模型. 在真实应用场景中, 基于标准模型证明的实现方案其安全性能一般能够得到保证, 但其计算效率往往较低. 当前绝大部分签密方案均是基于随机预言机模型进行安全分析, 但其安全性在真实场景中并不一定能够得到保证, 研究基于标准模型的签密方案是一个重要研究方向. 本文总结了无证书签密方案的研究框架, 如图 7 所示.

4.7 基于区块链实现的签密技术

在数据共享场景中, 一方面需要实现数据机密性和完整性, 保证用户提供的数据不被恶意嗅探或篡改, 另一方面, 也需要系统提供一种激励制度, 促使用户积极贡献自己的数据. 区块链能够在系统不同节点间同步交易数据, 形成全网一致的数据视图, 因此能够适配到具有数据同步需求的场景. 此外公有链本身自带一定的奖励属性, 能够量化用户的贡献, 并通过群体共识形成的奖励方案, 为用户提供了较为公平的价值分配策略[54]. 另外, 当前一些区块链系统提供了延时机制, 可以实现可信的时效管理功能. 基于区块链实现的签密技术, 能够增强传统签密算法的安全性, 使其能够适配到更为广泛的应用场景中.

(1) 适用于群智感知场景中的数据采集系统

文献 [55] 基于区块链技术提出了一种群智感知场景中的数据采集系统, 该系统主要利用区块链的去中心化

以及不可篡改特性, 实现感知数据的可信采集和存储, 能够有效适配到智慧城市、工业控制等领域中.

该系统结构如图 8 所示, 其执行步骤如下. 首先系统发送相应的数据采集任务, 这些任务会附带相应的奖励. 用户节点收到任务通知后, 会上传相应的感知数据, 系统中其他用户节点会对这些数据进行校验, 并将校验得到的数据质量报告、元数据信息以及经签密处理的密文数据以交易的形式发送至区块链系统中. 矿工节点收到上述数据后, 再对信息进行校验和评估, 并根据一定的规则为数据上传者分配相应的奖励. 该系统对用户进行了分组, 组内成员之间是相互信任的, 因此可以对敏感数据直接进行校验. 当要将数据上传到区块链系统时, 就必须要将敏感数据隔离到用户组内. 利用签密技术对敏感数据进行处理, 一方面能够实现数据安全性, 另一方面能够提升算法执行效率.

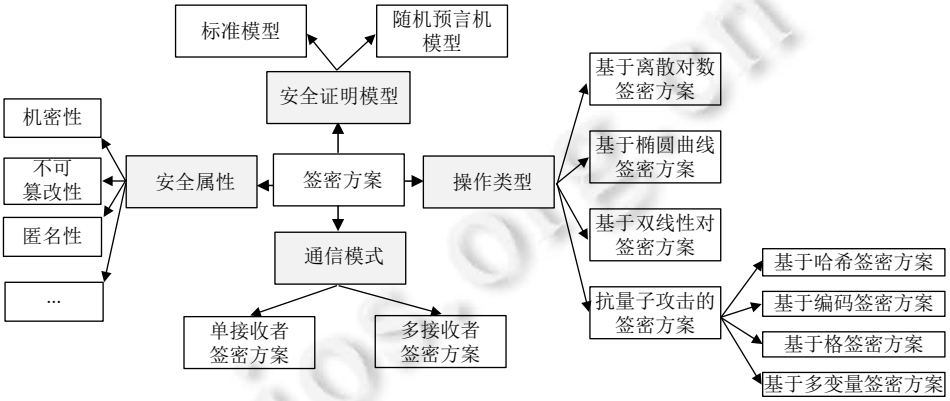


图 7 无证书签密方案研究框架图

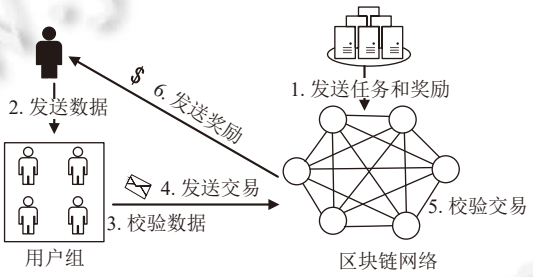


图 8 适用于群智感知场景中的数据采集系统

(2) 带时效的数据签密系统

文献 [56] 基于 EOS 系统引入的“Delayed Transaction”机制[57], 实现了一种基于区块链的带时效签密方案, 确保系统用户按照时效配置策略退出, 中间不需要人工参与.

利用区块链技术实现节点退出功能时, 系统会为用户公钥附属相应的时效期, 并将上述信息一起存储在区块链中, 并由区块链管理公钥信息. 一旦节点时效期已过, 区块链自动更新用户公钥信息. 由于区块链数据的公开性, 其他用户可以查询区块链来判定用户公钥是否失效.

(3) 基于区块链实现的多 KGC 签密系统

传统的签密系统均需要依赖 KGC 来协助生成用户的公私钥信息, 一旦 KGC 出现故障, 会导致整个系统不可用, 而且 KGC 中保存的系统主密钥一旦为恶意用户所掌握, 会泄露用户隐私数据. 传统的分布式 KGC 实现方案, 虽然可以解决单点故障问题, 但却会增加系统实现复杂度和运行成本.

针对上述问题, 文献 [58] 提出了一种多 KGC 实现策略来管理系统主密钥, 各 KGC 能够定期更新自己的子秘密信息, 即使恶意用户掌握了一定数量的子秘密信息, 但若超过规定时间这些子秘密信息均会失效. KGC 之间需要依赖一种同步机制来触发子秘密更新操作, 该系统则主要利用区块时间戳和区块高度等信息来完成. 需要说明

的是, 上述整个过程系统并不会发起新的交易, 只需要侦听区块链网络状态即可, 因此不需要消耗相应的区块链虚拟数字货币.

4.8 总 结

在本节中, 我们介绍了数据机密性和完整性概念及其重要性, 并介绍了传统的实现方案及其面临的一些问题, 同时介绍了基于区块链的解决方案, 最后介绍了传统的签密算法以及对应的基于区块链安全增强算法. 表 7 对比了本节介绍的基于区块链实现方案, 分别从安全特性、区块链平台、公钥产生方、密钥管理机制以及利用的区块链特征等方面进行考察.

表 7 基于区块链实现数据机密性和完整性方案

方法	安全特性	区块链平台	公钥产生方	密钥管理机制	主要利用的区块链特性
IKP	机密性	以太坊	CA	PKI	去中心化
文献[32]	机密性	Namecoin	PKG	IBC	所有特性
文献[37]	机密性	Namecoin	PKG	IBC	所有特性
PPDP	完整性	私有区块链	CA	PKI	不可篡改性
Zeppar	完整性	私有区块链	CA	PKI	所有特性
Liu方案[44]	完整性	私有区块链	CA	PKI	所有特性
Storj	完整性	Florincoin[48]	CA	PKI	所有特性
Ericsson	完整性	Guardtime	用户	KSI	所有特性
文献[55]	机密性和完整性	超级账本	KGC和用户	CL-PKC	所有特性
文献[56]	机密性和完整性	EOS	KGC和用户	CL-PKC	延迟交易
文献[58]	机密性和完整性	以太坊	KGC和用户	CL-PKC	所有特征

5 身份认证

身份认证是通过一定技术手段, 对用户身份进行确认, 确保只有合法用户才能访问相应服务的一种安全机制[59]. 当前已经出现了一些基于公钥密码学机制实现的身份认证方案, 例如基于 PKI 机制和基于 IBC 机制的实现方案. 在本节中, 我们首先讨论了传统实现方案所存在的问题, 然后介绍了利用区块链技术实现身份认证的方案, 并对其进行了总结.

5.1 身份认证技术

认证技术是信息系统确认用户身份的一种手段, 所有现实世界中的物理实体都会进行数字化, 形成信息世界中一串数字符号, 用户身份同样也是这样, 用户授权动作针对的对象也是用户数字身份. 构建可信网络的前提就是确保参与实体身份信息可信, 认证是建立可信的前提.

人们很早之前就已经对身份认证技术进行了研究, 古代战争士兵晚上巡夜所用的口令就是一种最简单的身份认证技术. 现有的身份认证技术, 按照判定条件数量分类, 可分为单因子认证和双因子认证. 按照是否依赖物理介质, 可分为软件认证和硬件认证. 当前常用的几种身份认证技术, 包括静态口令、动态口令、IC 卡、生物特征认证、USB Key 认证等技术, 发展的均比较成熟, 表 8 总结了这些技术以及其存在的问题. 本文主要关注基于密码学手段实现身份认证相关技术, 其中重点论述基于 PKI 机制和 IBC 机制实现身份认证的技术手段.

5.2 传统基于 PKI 机制实现身份认证面临的问题

当前基于 PKI 机制实现身份认证方案, 主要基于两种方案, 即 CA 和 WoT, 然而基于这两种技术实现的认证方案, 会面临一些安全挑战, 下面将对其进行详细论述.

基于 CA 实现身份认证方案主要面临 3 种问题: 第三方信任、单点故障和构建成本. 首先对于第三方信任, 由于 CA 一般是通过中心化节点进行实现, 系统会依赖 CA 来生成、分发和管理用户密钥信息, 因此必须要保证可信. 如果 CA 被第三方恶意节点攻破, 会泄露用户密钥信息. 其次系统还会面临单点故障问题, 如果 CA 不可用, 会

导致整个系统失去可用性. 最后如果由 CA 节点来管理所有用户的密钥信息, 会占用大量的存储和带宽资源, 同时由于密钥管理过程涉及大量的插入、修改、查询等操作, 会导致系统实现既昂贵又低效[60].

表 8 身份认证技术原理以及面临的问题

身份认证技术	技术原理	面临的问题
静态口令	用户输入密码是否与系统存储的密码匹配进行身份认证	静态口令容易泄露, 需要用户定期更改口令
动态口令	系统按照时间或使用次数生成用户密码, 并确保该密码只能被使用一次	动态口令一般会依赖专用硬件进行实现, 会增加一定成本, 而且硬件容易丢失或损坏
IC卡	内置用户身份信息的集成电路卡片, 而且不可复制	攻击者可以通过内存扫描技术来获取卡片中存储的信息
生物特征认证	采用指纹、虹膜等每个人独一无二的生物特征来进行身份认证	生物特征仍可以被伪造
USB Key认证	基于USB设备内置密码学算法实现身份认证, 轻便易携带	存在丢失或损坏的风险

基于 WoT 实现身份认证的方案, 需要依赖一些节点事先建立信任关系, 新用户想要加入系统网络, 必须通过至少一个可信成员节点校验后, 才能成功加入系统. 如果系统中被信任的某些用户作恶, 会破坏后加入用户的信任关系, 进而破坏整个系统的安全性.

基于 CA 和 WoT 技术实现的身份认证方案, 除了上述所列问题之外, 还存在身份保持问题, 也就是恶意用户冒名加入系统并获取系统服务的问题. 针对身份保持问题, 相关研究人员提出了一些基于日志的解决方案, 但也存在存储资源消耗过多的问题.

5.3 传统基于 IBC 机制实现身份认证面临的问题

基于 IBC 机制实现的身份认证方案, 所存在的问题就是需要依赖中心化的 PKG 节点来生成和管理用户的私钥信息, 这种实现架构存在单点故障问题, 一旦 PKG 被破坏, 会导致整个系统不可用. 此外, 由于 PKG 掌握了所有用户的私钥信息, 一旦被恶意第三方攻破, 会泄露用户的隐私信息.

5.4 基于区块链实现身份认证方案

区块链具有去中心化、不可篡改和可信溯源的能力, 可以为实现多种安全服务提供相应的支撑技术. 更重要的是, 当前大部分区块链系统已经开源, 开发人员可以基于这些免费系统来构建相应的解决方案. 本节主要介绍传统的几种基于 PKI 机制和基于 IBC 机制实现身份认证的方案, 表 9 总结了传统实现方案所存在的问题以及基于区块链的解决方案.

表 9 传统身份认证实现方案存在的问题以及基于区块链的解决方案

传统方法	面临的问题	基于区块链的解决方案
基于CA的身份认证方案	1. CA节点可信问题 2. 单点故障问题 3. 身份保持问题 4. 部署成本高	1. 基于分布式共识机制形成群体决策 2. 去中心化实现方案解决了单点故障问题 3. 基于区块链溯源功能, 通过记录用户访问日志来判定用户是否重复注册 4. 开源实现, 免费获取
基于WoT的身份认证方案	1. 必须通过系统可信节点的校验 2. 身份保持问题	1. 依赖群体投票方式进行信任校验 2. 基于区块链实现的可信溯源功能, 校验该用户是否已经注册过

(1) Pomcor: 当 CA 为用户颁发证书时, Pomcor[61]会计算出证书对应的哈希值, 并将其存储到区块链中, 基于区块链不可篡改特性, 确保存储数据的可信性. Pomcor 系统需要构建两条区块链, 一条用于存储证书以及其哈希值, 记为 A 区块链, 另一条用于存储被撤销的证书及其哈希值, 记为 B 区块链, 证书是否有效需要依赖上述两种信息进行确认. 在校验证书有效性时, 系统会首先计算证书的哈希值, 然后检查上述两条区块链, 如果证书哈希值存在于 A 区块链中, 但并不存在于 B 区块链中, 则说明证书有效, 否则就意味着该证书失效.

Pomcor 采用了一种轻量级的证书校验算法, 设计架构简单, 但十分有效, 系统具有较低的时间延迟, 能够有效应用于身份认证领域中.

(2) Gan 方案: Gan 等人[62]基于区块链技术提出了一种适用于 IoT 场景的节点身份认证系统. 该系统采用私有链来存储节点公钥信息, 同时允许成员节点访问上述信息. 该系统事先设定中央 CA 节点 (CCA) 是完全可信的, 并且设备制造商验证器 (DMV) 只有通过 CCA 认证后, 才能连接到 CCA 节点. 系统假设 DMV 节点具有足够的算力和存储资源, 能够满足系统所需资源要求. DMV 之间会构建一条区块链, 用于提供数据安全服务. 当新的 DMV 申请加入区块链网络时, 需要通过 CCA 身份认证, 认证通过后, CCA 会将该 DMV 的公钥信息、地址信息以及签名信息以交易形式存储到区块链系统中, 这样其他节点就会知晓新加入 DMV 节点的相关信息.

(3) Nguyen 方案: Nguyen 等人基于区块链技术实现了一种符合 GDPR 标准的个人数据管理方案[63], 从而将数据管理权赋予数据拥有者, 并使服务提供商在管理个人数据时满足 GDPR 标准. 该方案利用区块链智能合约技术实现了一种去中心化架构, 保证身份管理过程的透明和可溯源.

该方案会为区块链系统中每一个参与实体构建一个唯一的身份标识, 并利用非对称密码学算法实现身份认证功能, 之后系统为通过认证的实体赋予不同的文件访问权限, 并利用区块链系统记录数据访问日志. 身份认证过程采用了一种基于 RSA/DSA 实现的签名算法, 并利用智能合约进行实现. 身份认证规则代表了区块链的群体共识, 这样可以防止出现传统中心化身份认证系统存在的过程不透明以及单点故障等问题.

(4) 分布式 PKI (DPKI): Allen 等人[64]基于区块链技术实现了一种分布式 PKI (DPKI) 系统, 能够实现用户身份认证. DPKI 将公钥绑定到用户身份信息中, 本身为一种分布式 Key-Value 数据库. 系统会将用户公钥信息直接存储到区块链系统中, 也允许只存储公钥指针, 以减少存储空间消耗. 为了保护系统主密钥信息, 系统采用了门限签名算法, 将主密钥切分为多个子秘密, 并将其存储到多个可信节点中, 即使部分节点被攻破, 系统仍能从剩下满足一定数量要求的可信节点中恢复系统主密钥. 用户公钥信息是由系统主密钥生成的, 系统还可以通过主密钥来撤销用户公钥信息.

系统在设计时, 主要是利用了区块链可信存储能力, 确保用户公钥信息的真实可信. 系统同时消除了对 CA 的依赖, 解决了单点故障问题. DPKI 能够保护用户身份信息的完整性, 能够有效适配于资源有限的物联网场景中.

(5) Li 方案: Li 等人基于区块链技术提出了一种适用于 IoT 场景的身份认证技术[65], 利用公钥密码学技术对 IoT 设备进行身份认证, 防止非法设备访问 IoT 网络. 该方案主要包括 3 个步骤, 分别是设备注册、身份认证以及完整性校验. 设备注册阶段, 待加入 IoT 网络的新设备需要将设备 ID、公钥、关键数据哈希值等数据存储到区块链系统中. 当新设备需要访问 IoT 网络时, 此时需执行身份认证步骤, 其他设备通过访问区块链网络校验新设备的身份信息. 在完整性校验阶段, IoT 设备需要定期向邻居设备请求关键数据, 然后再利用存储在区块链中的关键数据哈希值对其进行校验.

方案所用的区块链是 Hyperledger Fabric. 该方案能够及时检测出固件后门程序, 有效抵御恶意节点入侵攻击, 同时允许 IoT 设备将日志信息定期存储到区块链系统中, 从而实现设备行为审计功能.

(6) Guardtime 方案: Guardtime 是一种基于物理不可克隆函数 (physical unclonable function, PUF), 并结合区块链技术, 实现对 IoT 设备进行身份认证的方案[66]. PUF 是一种依赖设备特征生成设备唯一标识符的数字指纹硬件, 本身具有随机性和唯一性, 该指纹信息代表了设备独一无二的身份特征, 实际应用时, 可利用 PUF 这种特性, 生成相应的密钥信息, Guardtime 就是利用上述原理进行实现. Guardtime 生成公钥信息后, 会将其存储到区块链系统中, 同时利用 KSI 来确保固件以及会话数据的完整性. 由于 KSI 不需要存储相关的密钥信息, 因此可以节省相应的存储空间. 受限于物联网设备有限的计算能力和存储空间, 其相应的安全防护措施较弱, Guardtime 则可以利用有限的资源, 在保证系统安全的前提下, 实现对设备行为的追踪和溯源.

5.5 总　结

身份认证是实现安全服务所必需的一种技术, 本文首先概述了身份认证概念, 并介绍了传统实现方案及其存在的问题, 并给出了基于区块链的实现方案. 表 10 对比了基于区块链实现身份认证的方案, 需要说明的是, 其中大部分都没有开源代码, 实际中很少被应用.

表 10 基于区块链的身份认证方案对比结果

方法	区块链平台	公钥产生方	密钥管理机制
Pomcor	以太坊	CA	PKI
Gan方案	—	CA和DMV	PKI
Nguyen方案	以太坊	CA	PKI
DPKI	—	用户	PKI
Li方案	Hyperledger Fabric	用户	PKI
Guardtime	—	用户	KSI

6 隐私保护

根据经济合作与发展组织给出的定义, 隐私是指与已识别的或可识别的个人有关的任何信息, 涉及数据搜集、使用、存储和销毁整个过程[67]. 数据隐私保护是数据拥有者通过定义相应的数据访问规则, 控制数据资源被其他人访问的行为, 实际中可以通过访问控制列表 (access control lists, ACL) 来实现保护用户隐私的功能. 在本节中, 我们首先讨论用户隐私相关概念以及隐私保护的重要性, 然后讨论了传统的隐私保护技术及其面临的挑战, 最后介绍了基于区块链实现隐私保护功能的相关方案.

6.1 隐私保护简介

在互联网应用无处不在的今天, 越来越多的用户需要共享自己的数据资源, 而云计算能够提供相应的基础设施. 根据 Gartner 发布的报告显示, 2019 年全球云计算市场规模已达到 445 亿美元, 同比增长 37.3%[68]. 然而随着业务量增加, 云数据在不同安全域内的流动更为频繁, 用户隐私泄露问题变得日益严重. 首先云计算提供的用于隔离物理资源的虚拟化功能, 一旦出现安全漏洞, 会导致用户数据泄露. 其次云平台可能会将虚拟资源分配给多个实体, 而在这些实体间构建统一的安全防护措施非常困难. 最后由于云平台可能需要实时处理大量的数据访问请求, 受限于成本问题, 在设计系统时需要为性能分配较多软硬件设施, 因此可能会减少安全防护设施的计算资源, 从而导致系统安全性下降.

在数据共享场景中, 数据拥有者在将数据共享出去后, 隐私保护机制要求数据拥有者还能够拥有对自身数据的控制能力. ACL 是实现隐私保护功能的一种重要工具, 它通过定义一组规则来控制数据被访问的行为, 例如可以规定什么人可以在何时以何种途径来访问何种数据. 一般来讲, ACL 需要为数据的生成、使用、共享、存储、销毁等过程, 利用加密、签名等技术构建统一的安全防护措施.

6.2 传统的隐私保护技术

隐私保护问题已研究多年, 但相关的安全事件仍层出不穷, 目前仍是安全领域的一个研究热点. 当前已经出现了多种隐私保护技术, 例如 ACL、同态加密[69]、基于属性加密[70]等, 用来构筑隐私保护系统相关支撑技术. 本文重点论述的 ACL 实现机制, 要求数据流通中的各个关键环节, 均应严格遵守规则要求, 即便是数据存储节点未经授权也无法从中获取用户的隐私信息.

传统的隐私保护方案, 主要包括了数据匿名化和差分隐私, 通过采用上述策略, 使得攻击者难以将数据对应到数据所有者. 对于数据匿名化, 相关研究人员提出了 K-匿名算法, 能够对数据集执行匿名操作, 确保匿名信息至少与其他 $K-1$ 条记录相似[71]. Roberto 等人在 K-匿名算法基础上提出了一种改进的数据管理策略, 剔除了比较耗时的排序操作, 且能够在满足匿名前提下减少待扰动的数据量[72]. 后面相关研究人员分别提出了 L-diversity 方法[73]和 T-closeness 方法[74], 确保敏感数据存储在足够分散的位置中. 差分隐私则是采用数据扰动、添加噪声等操作, 最大化减少外界识别用户隐私信息的概率, 同时提升从统计数据库查询结果的准确性[75]. 例如医保中心记录有居民的身份证信息、社保卡号等敏感信息, 国家依赖这些数据建立了一个包含全局统计信息的数据服务中心, 来供政府部门进行决策. 如果攻击者从医院拿到了用户匿名化后的住院信息, 同时拿到了前述全局统计数据, 就可以通过链接相关数据识别出个人身份信息, 使得上述匿名功能失效. 针对上述问题, 相关研究人员提出的差分隐私则能

够有效防止数据库脱匿名.

6.3 传统隐私保护方案所面临的问题

尽管当前已经出现了多项隐私保护方案, 但这些方案仍面临一些安全或性能上的挑战, 主要包括数据安全壁垒、共享和隐私之间的权衡、效率、数据所有权判定以及系统性的数据生命周期管理方法[76], 我们将在下面对其进行详细讨论.

数据安全壁垒导致数据完整性缺乏保障: 在传统的数据共享场景中, 尤其是在政务监管领域, 各个机构存在数据壁垒. 每个机构都有自己的安全利益诉求, 会按照自己预先定义的隐私保护规则, 对数据进行过滤, 导致汇总后的数据是割裂的、不完整的, 数据缺乏在时间或空间上完整性保障.

数据共享和隐私保护之间的权衡: 隐私保护问题诞生于数据共享场景中, 在制订数据安全防护措施时, 需要在充分考虑数据使用场景以及数据主体知情权前提下, 充分发挥数据使用价值. 例如病人的医疗病历属于个人隐私信息, 但是如果能够在保护病人隐私前提下, 充分挖掘病历中有用信息, 帮助攻克癌症等疾病, 对病人来讲也是一种收益, 甚至很多人可能会降低对隐私的要求.

效率: 大部分隐私保护技术需要借助密码学算法进行实现, 而且需要依赖中心化节点来生成密钥和管理证书. 一方面系统实现较为复杂, 难以适配到对实时性要求较高的应用场景中. 另一方面由于其中心化的设计架构, 可能会引入单点故障等问题. 虽然研究人员已经提出了一些改进措施, 但其中大部分仍停留于理论层面, 实际应用效果有待验证[77].

数据所有权判定: 实际应用场景中, 判定数据归属以及数据是否被指定用户所使用, 属于数据所有权中的重要内容. 当前大部分隐私保护技术, 一般会由数据拥有者设定相关的数据访问规则, 然而数据接收方一旦获得数据, 数据拥有者就失去了对这部分数据的控制权.

系统性的数据生命周期管理方案: 数据在流动过程中, 需要为数据生命周期各个阶段定义细粒度的安全防护措施, 形成一套统一的系统化用户隐私保护框架. 该框架应该允许数据拥有者为数据生命周期各个阶段定义或调整相应的访问规则, 还可以撤销相应的访问授权. 构建这样的隐私保护框架, 有助于更全面地保护用户隐私, 但目前还缺乏相关的研究.

6.4 基于区块链的隐私保护方案

基于区块链实现的隐私保护方案, 能够提供一种扩展性良好的分布式框架, 可以解决传统实现方案所面临的一些问题. 首先可以解决数据壁垒问题, 区块链能够提供统一的数据访问接口和安全规则, 能够有效地促进数据共享; 其次关于数据共享和隐私保护权衡问题, 用户可以利用智能合约及其他密码学算法, 适时调整相关的隐私保护规则; 然后关于数据所有权判定问题, 可以利用区块链提供的可信溯源功能进行实现; 最后区块链可以对数据生命周期各个阶段进行追踪, 通过发送新的交易或者利用智能合约功能, 促进各实体协同参与数据管理过程. 由于区块链本身也依赖密码学技术进行实现, 同时也引入了比较耗时的共识机制, 因此基于区块链构建的方案仍存在效率较低的问题. 表 11 列出了传统隐私保护方案面临的问题以及基于区块链的实现策略.

表 11 传统隐私保护方案存在的问题以及基于区块链的解决方案

传统隐私保护方案面临的问题	基于区块链的解决方案
数据安全壁垒导致数据完整性缺乏保障	基于区块链去中心化特征, 形成统一的数据过滤准则
权衡数据共享和隐私保护较为困难	利用智能合约和密码学算法, 适时调整数据共享和隐私保护之间平衡点
大部分需要依赖密码学算法, 执行效率较低	尚没有提升效率的实现方案
数据确权困难	基于区块链不可篡改性和溯源特性实现数据确权
缺乏系统性的数据生命周期管理方案	利用智能合约来调整数据生命周期内各个阶段的访问规则

当前大部分利用区块链解决隐私保护问题的实现方案主要是利用了区块链可信数据存储能力以及智能合约功能, 首先数据所有者利用智能合约来定义数据访问规则, 并将智能合约、元数据等其他信息以交易的形式发布到区块链中, 同时可以将原生数据存储到云端, 系统保证只有符合规则要求的用户, 才能访问上述数据. 系统一般

会采用密码学算法对共享数据进行加密, 然后再将密文数据上传到系统中, 确保数据未经授权不能被访问. 除此之外, 当前相关研究人员提出了利用混币[78]、零知识证明[79]、环签名[80]、安全多方计算[81]、基于硬件隔离的安全执行环境[82]、同态加密[83]等技术来保护用户隐私的方案, 表 12 对比了这些技术. 在下文中, 我们将讨论几种利用区块链技术保护用户隐私的实现方案.

表 12 区块链隐私保护方案对比

方案名称	实现机理	优点	缺点
混币	混淆交易, 隐藏敏感信息	有效保护用户隐私	需要结合盲化等技术增强效果
零知识证明	不提供有用信息的前提下证明某个论断正确性	有效保护用户隐私	执行效率较低
环签名	基于哈希函数、椭圆曲线等密码学算法实现交易双方身份匿名	隐藏交易方身份信息	增加了监管难度
安全多方计算	交易方在不泄漏自己信息前提下实现协同合作, 完成交易	去中心化实现	执行效率较低
基于硬件隔离的安全执行环境	基于TEE (trusted execution environment)等技术保护用户隐私	执行效率较高	升级困难, 同时受限于设备存储空间, 导致TEE无法执行较为复杂的运算
同态加密	利用同态加密算法对用户敏感数据进行加密, 然后区块链再对密文数据进行校验	有效保护用户隐私	执行效率较低

(1) Hawk: Hawk 是 Ahmed 等人[84]提出的一种基于零知识证明技术实现隐私保护功能的智能合约编程模型, 该模型并不需要在区块链中存储包含转账金额等敏感信息的交易数据, 以此来减少隐私泄露的风险. Hawk 提供了一种较易入门的编程框架, 非专家人员也能够很快写出 Hawk 程序. 此外, Hawk 主要利用密码学手段确保链上数据的隐私安全, 确保未参与合约计算的实体不能获知任何有用的信息, 同时还实现一种合约安全机制, 确保参与计算的实体不能获知其他参与实体的隐私信息.

Hawk 框架中的智能合约分为两部分, 即 private 部分和 public 部分, private 部分主要涉及到交易金额等隐私数据, 标记为 ϕ_{priv} , 而 public 部分则代表可以公开的数据, 标记为 ϕ_{pub} . ϕ_{priv} 利用 C 语言进行编写, 而 ϕ_{pub} 利用 Serpent 进行编写. Hawk 执行效率较高, 后续将开源相关代码, 为系统开发人员提供一种廉价高效的区块链隐私保护框架.

(2) Zerocash: Zerocash 是 Eli 等人在 2014 年基于比特币模型, 利用零知识证明技术提出的一种去中心化匿名支付数字货币[85]. 相比较原生比特币, Zerocash 能够对交易方身份信息和交易金额进行脱敏处理, 进而有效保护用户隐私信息.

Zerocash 主要采用非交互式零知识证明技术 zk-SNARKs 来对资产归属权进行判定, 例如, 用户无需展示自己的资产数目, 就可以向其他人证明自己能够支付应缴税款, 而其中第三方不能从中侦知应付税款、交易金额等敏感信息. 用户在发起交易的时候, 首先要执行铸币过程, 向资金池中注入资产, 并在全局列表中写入相应的承诺信息, 再利用 zk-SNARKs 证明自己持有承诺所需的私钥信息, 以此来表明自己持有承诺中资产. 然后用户执行浇筑操作, 将承诺中的资产切分为不同价值的份额, 方便用户在保护自身隐私的前提下灵活地进行交易. 实验结果显示, Zerocash 中的每笔交易记录数据小于 1 KB, 校验时间低于 6 ms, 比原生比特币系统更具竞争力.

(3) HCB-SDPP: Wei 等人在 2019 年基于同态加密算法提出了一种智能家居控制系统 HCB-SDPP (homomorphic consortium blockchain for sensitive data privacy preserving), 利用联盟链 Hyperledger Fabric 来保护用户敏感信息[83]. 用户在上传自身数据时, 首先利用 Paillier 加密算法对数据进行盲化处理, 再将其上传到联盟链中, 联盟链再对上述密文数据进行校验.

当前出现的区块链系统利用了 SHA256、MD5、SHA512、Paillier、Goldwasser-Micali 等加密算法, Sharath 等人[86]则对上述算法在适配到区块链场景时进行了考察, 实验结果显示, 同态加密算法 Paillier 和 Goldwasser-Micali 能够在安全性和性能方面取得良好效果. HCB-SDPP 系统则是利用 Paillier 算法重新设计了一套新的区块结构, 用来存储交易信息, 能够有效保护用户隐私信息.

(4) Guan 方案: Guan 等人[87]提出了一种适用于智能电网领域的用户信息收集方案, 能够有效保护隐私信息.

在智能电网领域, 为了优化用电配置, 电网公司会搜集住户的用电信息, 而耗电量多少会泄露住户隐私信息. 针对上述问题, Guan 方案首先对住户进行分组, 每个组维护一条私有链, 用来存储住户身份、用电信息等数据. 系统会为组内每个住户生成多个不同的匿名身份, 这样住户上传的数据会绑定不同的身份信息, 外界很难从中推断出住户的真实耗电量. 其次, 当住户需要加入区块链网络时, 系统采用布隆过滤器来加快身份认证效率. 通过身份认证后, 住户就可以将自己的用电信息存储到区块链中. 每隔一定时间, 区块链系统会随机选择一个住户作为矿工节点来聚合信息, 并将信息以交易的形式发送到区块链中, 其他成员节点再对其进行校验, 如果校验成功, 聚合信息会被上传到智能电网控制中心. 密钥管理中心负责为住户以及控制中心生成相应的密钥信息, 也可以撤销已分发的密钥信息.

实验结果显示, 随着住户数量增加, 系统执行时间并没有出现显著增加的趋势. 但是 Guan 方案并没有开源代码, 实际运行结果仍有待验证.

(5) Shafagh 方案: 当前物联网场景中的大多数 ACL 实现机制, 都需要一个可信中心节点来管理访问授权功能, 然而这种实现方案会面临单点故障问题. Shafagh 等人[88]使用区块链技术设计了一种可审计的数据安全共享系统, 主要解决 3 个方面的问题: 第 1 是实现去中心化的可审计访问控制机制, 主要解决数据确权以及数据安全共享问题; 第 2 是实现数据安全存储机制, 确保数据的机密性、完整性、可用性; 第 3 是确保兼容 IoT 数据格式, IoT 采集的数据主要是 append-only 类型的数据流, 具有一次写入多次读取的特性, 新的安全机制需要兼容这种数据格式.

Shafagh 方案基于区块链交易来管理 IoT 数据访问控制过程, 为此设计了一种新的区块结构, 用以存储数据所有权和相应的访问权限. 客户端为 IoT 数据流附属相应的所有权以及访问权限, 打包后再进行加密, 之后将其存储到系统中. 当需要共享数据给其他用户时, 数据所有者需要提交包含授权内容的交易信息, 也可以提交撤销授权交易来取消数据访问权限. 当某一用户想要访问数据时, 存储节点会向区块链查询相应的数据访问权限, 并根据查询结果来决定是否返回用户数据. 由于区块链和存储节点中的数据都经过了加密处理, 其他未授权节点并不能从中嗅探出用户的隐私数据. 为了增加安全性, 系统采用了一种低成本的密钥更新算法, 能够频繁更新加密密钥.

(6) BBDS: BBDS 是一种基于区块链技术实现的数据共享系统, 能够为医疗数据提供相应的隐私保护功能[89]. BBDS 基于联盟链技术实现用户身份认证功能, 认证通过后, 系统会将用户后续的行为信息存储起来, 以实现可信审计功能. 该系统包括了 3 层架构: 用户层、管理层和存储层, 如图 9 所示. 用户层主要包括了需要共享或访问数据的个人或组织, 例如医院、科研机构、大学、政府部门等. 管理层主要负责管理数据以及构建相应的数据安全机制, 主要包括 3 种节点: 认证节点、验证节点和共识节点. 认证节点主要负责用户注册和身份认证; 验证节点主要负责管理用户的密钥信息, 同时负责校验用户身份信息; 共识节点主要用于构建区块链网络. 存储层包括了云基础设施和相应的安全防护设施, 主要用来提供安全高效的数据存储服务.

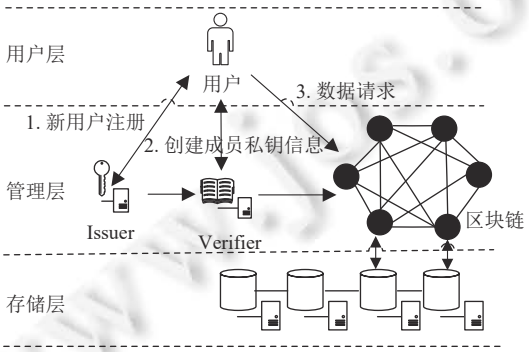


图 9 BBDS 系统架构图

系统采用的区块链是 Bitcoin 的一种改进版本, 通过修改交易和区块头信息, 对外提供轻量级可信数据存储服务. BBDS 采用身份认证技术和数据加密技术来保护数据存储以及交互过程的安全性, 提供了一套安全高效的医

疗数据服务框架.

6.5 总 结

近些年来, 用户隐私数据保护问题遇到了严峻挑战, 如何构建高效的数据安全防护方案已经成为系统开发人员必须要关注的问题. 隐私保护要求用户在共享数据时确保对自身数据的控制权, 防止未经授权的访问. 本节首先讨论了隐私保护相关概念以及传统的隐私保护技术, 同时讨论了几种基于区块链实现的隐私保护方案, 这些方案均需要依赖区块链提供的智能合约或是交易功能进行实现, 表 13 总结了这些方案. 需要说明的是, 尽管区块链提供了伪匿名功能, 第三方恶意用户仍能从中推断出用户的身份信息, 表 13 中一些系统采用了更换身份标识或者生成多个匿名身份的方法来隐藏用户身份.

表 13 基于区块链实现隐私保护方案的对比结果

方法名称	区块链平台	基于智能合约还是交易实现隐私保护
Hawk	以太坊	智能合约
Zerocash	比特币	交易
HCB-SDPP	Hyperledger Fabric	交易
Guan方案	—	交易
Shafagh	比特币	交易
BBDS	比特币	交易

7 数据可信删除

当前大部分应用场景中, 例如在云计算、智慧城市、医疗保健等领域, 用户会将数据托管在运营商服务器中, 然而用户却不能直接控制这些数据, 在执行删除动作时也不能确定远程服务器端是否真正删除掉这些数据. 数据拥有者所享有的数据删除权, 需要确保数据能够真正被删除, 防止任意实体从被删除的数据中提取或恢复任何有效的信息[90]. 本节首先讨论了可信数据删除相关的知识背景, 并介绍了传统的实现方案和其面临的一些挑战, 然后重点讨论了利用区块链技术实现数据可信删除的几种系统实现, 并在最后对其进行了总结.

7.1 数据可信删除

当前云计算应用日渐广泛, 用户会采用云基础设施提供的数据访问接口来存取自身数据, 然而这种服务模式在给用户提供极大便利的同时, 也会使用户丧失对数据的控制权. 数据可信删除, 可以确保用户在执行删除动作后, 云服务提供商或其他授权访问该数据的用户便不能从中获取任何有效信息. 数据可信删除是数据拥有者应该享有的一项权利, 能够确保数据拥有者对自身数据撤回、遗忘的能力. 数据可信删除, 与前面所述的几种安全服务一样, 也是组成用户数据安全机制中的重要一环.

由于云基础设施普遍采用虚拟化、数据控制相分离的实现架构, 导致传统的数据删除技术变得不可用. 例如为了实现系统高可用性, 云基础设施会采用副本策略对数据进行异地备份, 用户执行删除数据动作时, 云基础设施可能仅仅是将对应数据的引用次数减少, 并不会真正地删除数据. 因此, 如何确保数据的可信删除是实现数据安全服务中的一个重要研究问题.

7.2 传统云环境中的可信数据删除方案

传统云环境中的数据删除方案中, 用户一般需要先对数据进行加密处理, 再将密文数据存储到云端. 用户需要删除数据时, 只需要将密钥销毁即可, 即使第三方拿到密文信息, 也无法从中获取任何有价值的信息, 这样可以间接实现数据可信删除功能. 这种实现方案可行的前提就是采用的加解密算法必须是安全的, 因此一般会选用可证安全的密码学算法进行实现. 另一个需要关注的问题就是密钥管理问题, 确保密钥销毁后就无法再恢复出来, 这是设计数据删除系统的难点所在.

按照密钥管理方式不同, 传统的数据可信删除技术主要分为 3 种: 基于可信权威机构的数据删除技术、基于分布式哈希表 (distributed hash table, DHT) 的数据删除技术以及基于用户的数据删除技术. Perlman 等人提出了利

用可信权威机构来管理密钥信息的实现方案, 并基于此来实现数据可信删除功能[91], 然而该方案仅实现了基于时间的数据删除技术, 功能较为单一. 针对上述问题, Tang 等人提出了 FADE 系统[92], 该系统将数据与相应的删除规则对应起来, 当外界条件满足规则要求时, 会自动触发数据删除动作, 中间无需人工干预. 基于可信权威机构实现的数据删除方案, 普遍存在单点故障的问题, 因此后来 Geambasu 等人基于 DHT 提出了一种分布式密钥删除技术, 确保密钥销毁过程由 DHT 自动完成[93]. 虽然该方案克服了单点故障问题, 但是由于 DHT 网络是按照固定的时间间隔来更新数据, 无法提供细粒度的数据删除操作, 此外 DHT 本身也存在一定的安全问题. 针对上述问题, Mo 等人提出了一种基于用户的数据删除技术[94], 数据销毁操作完全在用户端执行, 云端只负责数据存储. 用户只需要保管一个主密钥即可, 后续可以利用主密钥和密钥调制因子链来生成每个数据块对应的加密密钥. 该方案不需要依赖中心化的可信权威机构, 用户只需要管理少量的密钥信息就可以实现数据可信删除功能. 然而, 这种实现机制需要为每个加密密钥保存相应的密钥调整因子链, 会增加存储开销, 此外该方案缺乏数据删除操作审计流程.

传统的数据可信删除技术, 一般存在单点故障、缺乏可信审计、计算或存储资源消耗较多等问题, 基于区块链实现的去中心化数据删除方案, 有助于克服上述问题, 是当前的一个研究热点, 表 14 总结了传统实现方案存在的问题以及基于区块链的解决方案.

表 14 传统数据删除方法存在的问题以及基于区块链的实现方案

传统的方法	存在的问题	基于区块链的解决方案
基于可信权威机构的数据删除技术	1. 单点故障 2. 数据删除过程不透明	1. 基于区块链去中心化特征, 能够有效解决单点故障问题 2. 利用区块链智能合约, 可以为数据删除动作提供访问控制功能, 确保删除动作自动执行, 中间无需人工参与
基于DHT的数据删除技术	1. 执行效率较低 2. DHT本身存在安全问题	1. 基于区块链的实现技术尚不能解决低效问题 2. 区块链采用可证安全的密码学算法、共识机制来保证系统的安全性
基于用户的数据删除技术	1. 保存加密密钥需要额外的存储空间 2. 缺乏可信审计流程	1. 基于区块链实现数据可信删除功能, 用户无需保管密钥信息 2. 基于区块链的去中心化特征和可信溯源能力实现数据删除动作的可信审计

7.3 基于区块链的可信数据删除方案

当前基于区块链实现的一些应用系统, 可能会存在用户滥用行为, 例如非法交易、盗用版权、冗余存储等问题, 因此有必要删除链上存储的非法、错误、冗余数据. 当前已经出现了一些基于私有链和联盟链技术实现的系统, 能够删除上链数据. 文献 [95] 总结了可编辑区块链研究现状, 主要侧重链上数据修改操作. 本文从编辑类型、编辑模态、控制策略和编辑对象 4 个方面, 总结了当前区块链可信数据删除方案相关的研究框架, 如图 10 所示. 基于区块链实现数据可信删除方案, 其难点就是如何在确保原有对数据合法性验证方法不变的前提下, 实现对指定交易精准删除操作, 同时防止链上数据遭受恶意破坏, 下面对现有删除方案进行总结.

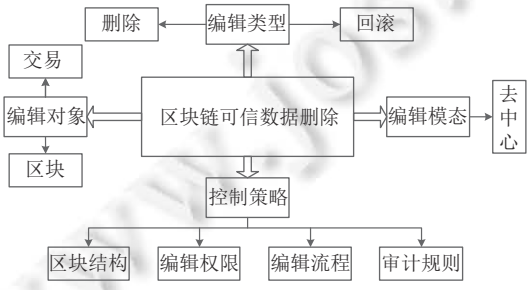


图 10 区块链可信数据删除研究框架

(1) Ateniese 方案: Ateniese 等人[96]提出了一种可重写的区块链实现方案, 主要使用变色龙哈希函数来修改链上信息. 当需要修改区块中的交易信息时, 系统使用变色龙陷门计算碰撞值, 进而实现删除指定区块中相关数据的

功能, 并确保修改后的区块链状态与修改之前保持一致. 然而该方案并不能扩展到公有链场景中, 这是由于该方案需要依赖可信节点秘密共享陷门信息. 后来埃森哲公司基于上述方案分别提出了两种区块链数据编辑方案[97,98], 同时给出了具体的实现细节.

该方案的数据编辑粒度是区块级, 如果要删除指定的数据, 需要对整个区块进行操作. 此外该方案基于多方计算协议来共享陷门信息, 实际运行效率较低.

(2) Derler 方案: Ateniese 方案存在数据修改粒度较大的问题, 即使用户只想删除一个 Byte 的数据, Ateniese 方案也要修改整个区块信息. 针对上述问题, Derler 等人[99]对 Ateniese 方案进行了扩展, 提出了一种基于访问策略的变色龙哈希算法, 将访问策略与哈希计算过程关联起来, 只要满足访问策略要求, 用户就可以计算出相应的哈希碰撞. 该方案将区块链数据编辑粒度下放到交易级别, 减少了需要修改的数据量. 该方案依赖变色龙哈希函数进行实现, 由于需要依赖可信节点来管理密钥信息, 因此也不能扩展到公有链场景中.

(3) Ren 方案: Ren 等人[100]基于 SpaceMint 区块链[101], 使用陷门单向函数提出一种可修改的区块链架构. 只要超过一定阈值数量的节点同意, 用户就可以对失效或错误的链上交易数据进行合法修改. 修改前后, 区块链结构保持不变, 也不会影响区块链上其他数据的正常使用. 数据修改操作需要系统节点达成共识, 因此数据删除操作代表了系统群体意志, 不合法的修改操作无法得到执行.

Ren 方案采用了一种新型的区块结构, 主要包含了 3 个部分: 证明子块、签名子块和交易子块, 如图 11 所示. 证明子块包含了区块的元数据信息, 交易子块则包含了交易信息, 签名子块主要隔离了证明子块和交易子块, 使挖矿挑战与交易本身无关, 矿工无法通过拆分交易来执行自私挖矿操作, 能够更好地对抗“Grinding Blocks”攻击. Ren 方案能够实现交易级别的数据修改功能, 与前面所述实现方案类似, 同样需要依赖陷门单向函数进行实现.

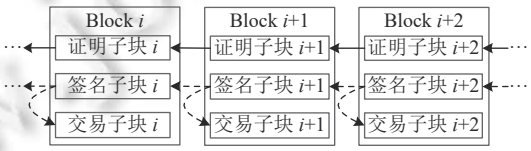


图 11 Ren 方案的区块链结构

(4) Marsalek 方案: Marsalek 等人[102]提出了采用平行的两条链来实现数据编辑功能, 其中双区块链主要由一对共生链组成, 分别是数据链和修正链, 数据链存储的是原始区块链数据, 而修正链则存储修改后的数据, 并且修正链是以区块为粒度进行存储. 当需要读取链上数据时, 系统会查找数据链, 每当发现区块被修改时, 系统就会切换到修正链上读取修改后的数据. 由于制订数据修改权限和授权编辑动作是验证者之间通过投票机制达成的群体共识, 能够利用区块链自带的 PoW 等共识机制完成, 因此该方案能够适配到公有链场景中.

Marsalek 方案是在区块链底层架构中实现的一种数据编辑方案, 能够实现区块级数据删除功能, 然而维护两条区块链会耗费一定的计算和存储资源.

7.4 总 结

表 15 总结对比了前文所述的几种基于区块链实现的数据可信删除系统, 其中大部分基于陷门单向函数进行实现, 需要依赖中心化节点来管理相应的陷门信息, 可能存在掌握编辑权的中心节点篡改链上数据的风险. 虽然可以基于安全多方协议秘密共享陷门信息, 防止某一节点作恶, 但无法避免合谋攻击, 同时也会存在一定的性能损失, 因此如何兼顾系统安全和性能是一个重要研究问题[103]. 此外这些实现方案并没有考虑数据删除操作可能导致出现的内容冲突问题, 虽然数据删除操作都需要进行授权, 但这些数据授权动作一般都只是经过语法层面的简单验证, 并没有考虑上下文数据之间的关联性, 可能会导致链上数据修改前后逻辑不一致, 后续仍需要对其做进一步的研究.

8 区块链价值及其面临的挑战

自人类社会出现以后, 人与人之间可信问题就相伴而生. 传统的基于中心化信任构建方案, 存在单点故障以及

第三方信任问题. 之后出现的区块链技术, 通过依赖机器代码作为信任中介来构建人与人之间信任关系, 提供了一种去信任化、去中心化的解决方案. 然而这在某种程度上来讲, 也显示了人们对现实世界的一种不信任. 本节将重点论述区块链所带来的价值、其面临的挑战以及未来展望.

表 15　基于区块链的可信数据删除系统对比结果

方法	修改粒度	实现原理
Ateniese	区块级别	陷门单向函数
Derler	交易级别	陷门单向函数
Ren	交易级别	陷门单向函数
Marsalek	区块级别	双链架构

8.1 区块链价值

区块链提供的价值包括数据安全、可信资源管理以及数据价值提升, 下面将对其进行重点论述.

首先对于数据安全, 基于区块链提供的不可篡改性, 可以确保区块链上存储数据的可信性. 由于区块链尤其是公有链中存储的数据对于系统参与节点是可见的, 可能会导致用户隐私信息泄露. 针对这些问题, 当前出现了在区块链中引入可信执行环境 (trusted execution environment, TEE) 的解决方案. TEE 本身能够提供一种与操作系统相隔离的计算环境, 外界未经授权很难获取 TEE 环境中的计算数据, 因此可以利用 TEE 来生成、分发和管理密钥信息, 确保区块链中存储数据的机密性和完整性[82]. 此外区块链智能合约允许用户配置和管理链上数据, 确定数据资源访问方式, 能够实现相关的数据访问控制功能.

区块链能够确保数据管理过程可信, 智能合约能够提供原子化操作功能, 计算过程中生成的临时数据和中间结果并不会上链, 只有执行成功的最终结果, 才可能被持久存储到区块链系统中. 当前一些基于区块链实现的系统, 需要结合云端存储和区块链来协同管理数据. 然而区块链在与链下数据交互过程中, 由于断电、网络故障、黑客攻击等原因, 会破坏计算的原子性, 攻击者可以从中嗅探出受保护的数据信息, 破坏数据的安全性. 当前出现了一些解决方案, 对需要与链下数据交互的智能合约, 按照一定原则划分为安全区间和不安全区间. 不安全区间就是需要与链下数据交互的代码部分, 而安全区间则保证了代码执行过程的原子性. 通过引入检查点管理功能, 对不安全区间代码进行安全管控, 保证链上数据与链下数据在流动过程中的数据安全.

对于数据价值提升, 区块链技术非常适合为不同利益主体间构建相应的分布式数据共享环境, 帮助打破各机构之间的数据壁垒, 实现数据安全共享, 充分发挥数据的价值优势. 区块链节点在共享数据时, 需要将数据以交易形式发布到区块链中, 各利益实体通过共识机制对交易进行核查和认证, 只有通过校验, 交易数据才能被写入到区块链中. 实际运行场景中, 可以将监管角色引入到系统中, 如果数据内容不实, 或者数据质量不高, 监管节点可以拒绝该数据上链, 确保数据的真实性和完整性. 数据一旦发布到区块链中, 参与节点就可以访问这些数据, 从而减少了数据跨部门流动中的审批流程, 提升了业务处理效率. 任意试图篡改数据的行为都会被区块链所拒绝, 从而增强了数据流动过程的安全性. 基于区块链提供的时间戳功能, 还可以对区块链的写入、广播、审计活动进行溯源, 提升服务质量.

8.2 区块链面临的挑战

区块链作为一种新型技术, 提出了一种颠覆式的社会信任构建体系, 极大地推动了社会发展和行业进步, 然而区块链本身也存在一定的问题, 下面对其进行详细分析.

首先是数据安全问题, 用户在将数据上链后, 区块链会对其进行冗余备份, 以保证数据的可用性, 然而这在某种程度上, 会增加用户隐私信息泄露的风险. 虽然联盟链和私有链引入了用户准入机制, 数据访问操作需要进行授权, 但相对于传统中心化实现方案, 区块链数据隐私安全强度仍是大打折扣, 因此如何保护链上数据隐私性是需要解决的一个重要问题.

其次是数据监管问题, 区块链作为一种多主体协同治理的组织架构, 能够帮助建立数据可信管理框架, 实现各

机构间信息互动. 区块链需要监管机构或者参与节点进行共同监督, 帮助打击洗钱、诈骗等违法行为. 但是由于当前区块链仍处于发展初期, 相应的监管技术并不成熟. 虽然当前一些联盟链技术引入了监管对象和相应的数据监管技术, 但由于缺少统一的行业规则和法律支撑, 在确定由谁来监管和治理上, 权责界线比较模糊, 缺乏责任划分以及行为免责等方面的研究.

然后是用户匿名问题, 伪匿名性是区块链的主要特性之一, 可以在一定程度上保护用户的身份信息. 然而区块链的匿名性设计并不完美, 当前大部分区块链系统, 都是直接将用户身份与公钥或其他能够代表身份的数据关联起来, 容易出现隐私泄露的问题. 以比特币为例, 它是将公钥信息作为用户身份标识, 攻击者并不能直接从公钥信息获知用户身份信息, 然而攻击者仍可以从区块链交易列表中, 通过分析交易规律来推断用户身份信息.

再次是资源消耗问题, 区块链在诞生之初, 参与构建网络的大部分节点主要是一些结构简单的异构节点, 本身并不具有较高的算力, 因此区块链设计人员并没有在系统中引入复杂的逻辑计算. 然而随着区块链应用范围的日渐广泛, 区块链需要执行一些较为复杂的运算, 例如远程证明、密文审计、零知识证明等, 而这会增加计算资源消耗. 此外由于业务量和参与节点数量增加, 区块链网络中广播的信息会变多, 而这会增加相应的网络负担.

最后是服务效率问题, 区块链本身设计为一种分布式系统, 具有良好的可扩展性, 然而, 正如文献 [104] 所述, 随着节点数量增加, 区块链系统的性能会明显下降. 在一些业务较重的应用场景中, 系统中会充斥大量交易信息, 而这会进一步降低区块链平台的执行效率. 以 Ethereum 和 Hyperledger Fabric 为例, 根据文献 [105] 给出的测试结果显示, 随着节点数量增加, 这两个平台中数据包数量会快速增加, 导致网络性能下降.

8.3 区块链安全服务未来展望

区块链安全服务未来发展主要聚焦两个方面: 技术层面与制度层面, 其中技术层面主要聚集于区块链支撑技术的安全性问题, 而制度层面则主要关注区块链生态体系中的社会制度、经济属性等非技术层面的安全问题. 区块链技术层面的安全问题, 主要包括区块链网络节点、智能合约、共识算法等方面的问题, 而制度方面安全问题, 则主要包括区块链监管体系建设、行业标准和法律法规是否完善等方面的问题.

8.3.1 区块链技术层面未来展望

区块链网络节点通常都是由监管机构、管理部门或主要发起者承担和维护, 实际运行场景中一般由一些利益相悖的节点共同维护, 实现对事务达成共识. 区块链节点一般具有区块链平台上的全部数据, 能够对其他节点传输过来的数据内容进行合法性验证, 并对数据执行排序、共识和同步等操作, 这些操作都需要消耗一定的时间. 未来区块链节点可以基于可信硬件技术, 例如 SGX 技术, 实现计算共识分离、数据隔离通道等高级功能, 这样能够在保证数据安全性的同时, 提高数据的可用性.

区块链智能合约能够为用户提供定制化服务, 并实现配置管理功能. 智能合约一般运行在区块链的虚拟机中, 支持用户对数据的增加、查询等功能, 并负责将数据和相关状态信息存储到区块链中, 由全网节点各保留一份. 由于区块链虚拟机和区块链存储空间均比较珍贵, 实际运行中一般会要求用户编写较为短小的智能合约, 而这会在一定程度上限制智能合约功能实现. 基于智能合约实现安全服务, 可以利用形式化验证工具检测代码漏洞, 来提升智能合约安全性.

在利用区块链技术实现安全服务时, 随着参与共识节点数量增多, 区块链网络中传输的消息会迅速增加, 导致达成共识的耗时急剧上升, 优化共识算法对提升区块链可用性具有重要作用. 实际中, 可以设置部分区块链节点执行共识计算, 并不需要全部节点参与共识过程, 这样能够降低区块链交易确认时间. 例如在监管科技场景中, 只需要在监管节点间完成共识即可, 这种通过部分区块链节点参与共识的方法, 将共识达成过程仅限于某一节点范围内, 能够提升区块链执行效率. 此外, 还可以利用可信硬件提供的远程证明机制, 只需要某一可信计算节点执行智能合约即可, 其他节点均信任上述可信节点的执行结果, 无需再重复执行智能合约来验证计算结果正确性, 这样能够进一步提升共识算法效率.

8.3.2 区块链制度层面未来展望

加快区块链监管体系建设, 是区块链发展所需要关注的要素之一. 区块链作为一种多主体协同治理的组织架

构, 通过打破各机构之间的数据壁垒, 融合各利益参与方形成群体共识, 实现数据价值可信传递. 在利用区块链提供的服务时, 尤其是加密数字货币领域, 需要建立并完善相关的监督体系, 实现一种包括监管机构、社会群体等在内的上下联动信息反馈系统, 进而构建区块链良性发展生态圈. 基于区块链的不可篡改性、去中心化、可信溯源等优良特性, 实现全过程、自动化数据审计功能, 最大限度实现对政府单位、行业组织、社会团体等对象的监督. 区块链监管技术作为一种新兴产业模式催化剂, 能够利用区块链优良特性再造信任生态圈, 对促进区块链健康发展具有重要意义.

另一个需要重点关注的要素就是区块链相关行业标准和法律法规完善, 这是规范区块链安全服务的依据和标杆. 虽然目前我国已经初步制定了区块链相关的法律和行业规范, 例如《区块链信息服务管理规定》《中国区块链技术和应用发展白皮书》等, 但目前依旧缺乏系统性的法律法规和行业规则, 极大地限制了区块链在各个领域的应用. 未来需要根据各行业实际状况, 制定出完备的区块链应用标准和法律法规, 让基于区块链实现的安全服务做到有章可循、有法可依, 使区块链应用得到有效规制, 减少区块链行业违规违法行为, 促进区块链行业良性发展.

8.4 总 结

在本节中, 我们讨论了区块链技术所带来的价值, 并介绍了区块链在落地时面临的一些挑战, 其中包括数据安全、数据监管、用户匿名、资源消耗和服务效率, 在设计相应的系统时, 需要对这些因素进行综合考虑, 表 16 总结了这些挑战, 最后给出了区块链安全服务未来展望.

表 16 基于区块链系统面临的挑战

挑战	问题	对应用程序的影响
数据安全	区块链节点可以访问系统存储的数据	攻击者可以获取区块链数据, 并从中嗅探用户隐私信息
数据监管	当前区块链监管技术并不成熟, 缺少相关的监管规则和法律支撑	缺乏监管的区块链系统, 容易出现洗钱、诈骗等违法行为, 而且无法对这些行为进行追责
用户匿名	当前大多数区块链系统的匿名性设计是将系统标识与用户身份信息关联在一起	攻击者可以通过分析区块链交易列表, 找到交易规律, 进而推断用户身份信息
资源消耗	区块链本身是一种P2P网络, 需要依赖广播机制来发送交易信息, 带宽资源消耗较多, 同时系统实现依赖较为耗时的密码学运算	随着业务量增加, 计算开销以及带宽开销也会随之增加
服务效率	随着节点数量增加, 区块链系统性能会下降	联盟链或私有链只适用于节点数量有限的场景中

9 结束语

本文对基于区块链实现的相关安全服务进行了综述, 主要包括数据机密性、数据完整性、身份认证、数据隐私、数据可信删除. 首先介绍了区块链的基础知识, 重点介绍了区块链的共识机制和关键安全特征. 由于实现上述安全服务的大部分系统需要依赖公钥密码学, 因此本文介绍了公钥密码学及其在实际场景中的应用. 由于公钥密码学中面临的主要问题就是密钥管理问题, 因此介绍了 3 种常见的密钥管理技术, 即基于 PKI 的实现机制、基于 IBC 的实现机制以及基于 CL-PKC 的实现机制, 并介绍了基于上述密钥管理机制和基于区块链技术实现数据机密性和完整性的方案. 当前出现的数据签密算法, 能够同时实现数据机密性和完整性, 而且要比单独实现上述两种安全特性的执行效率要高, 因此本文介绍了签密算法以及基于区块链的签密算法. 对于身份认证, 本文介绍了传统的基于 PKI 机制和 IBC 机制实现身份认证方案以及其存在的问题, 并介绍了基于区块链实现身份认证的方案. 对于隐私保护, 本文介绍了隐私保护的定义和重要性, 并介绍了传统的隐私保护实现策略, 并分析了其存在的问题, 同时给出了基于区块链的实现方案. 对于数据可信删除服务, 本文介绍了相关定义, 并介绍了传统的 3 种实现策略, 即基于可信权威机构的数据删除技术、基于 DHT 的数据删除技术以及基于用户的数据删除技术, 本文介绍了这 3 种传统方案面临的问题以及基于区块链的实现方案. 最后讨论了区块链的应用价值, 并介绍了基于区块链实现安全服务时所面临的一些问题以及区块链安全服务未来展望.

References:

[1] Pilkington M. Blockchain Technology: Principles and Applications. Cheltenham: Edward Elgar Publishing, 2016.

[2] Nakamoto S. Bitcoin: A peer-to-peer electronic cash system. https://bitcoin.org/bitcoin.pdf

[3] Eyal I, Gencer AE, Sirer EG, Van Renesse R. Bitcoin-NG: A scalable blockchain protocol. In: Proc. of the 13th USENIX Conf. on Networked Systems Design and Implementation. Santa Clara: USENIX Association, 2016. 45–59. [doi: 10.5555/2930611.2930615]

[4] Skinner C. Valueweb: How Fintech Firms are Using Bitcoin Blockchain and Mobile Technologies to Create the Internet of Value. Singapore: Marshall Cavendish International (Asia) Pte Ltd., 2016.

[5] Gafurov AR, Skotarenko OV, Nikitin YA, Plotnikov VA. Digital transformation prospects for the offshore project supply chain in the Russian Arctic. IOP Conf. Series: Earth and Environmental Science, 2020, 554: 012009. [doi: 10.1088/1755-1315/554/1/012009]

[6] Stallings W. Cryptography and Network Security: Principles and Practice. 7th ed., Boston: Pearson, 2016.

[7] Parliament E. General data protection regulation. 2016. https://gdpr-info.eu

[8] General Administration of Quality Supervision, Inspection and Quarantine of the People's Republic of China, Standardization Administration. GB/T 35273-2017 Information security technology—Personal information security specification. Beijing: Standards Press of China, 2018 (in Chinese with English abstract).

[9] Ismail L, Materwala H. A review of blockchain architecture and consensus protocols: Use cases, challenges, and solutions. Symmetry, 2019, 11(10): 1198. [doi: 10.3390/sym11101198]

[10] Xu M, Chen XT, Kou G. A systematic review of blockchain. Financial Innovation, 2019, 5(1): 27. [doi: 10.1186/s40854-019-0147-z]

[11] Bitcoinwiki. Proof of Work. 1999. https://en.bitcoin.it/wiki/Proof_of_work

[12] Bitcoinwiki. Proof of stake. 2013. https://en.bitcoin.it/wiki/Proof_of_Stake

[13] Wikipedia. Proof of space. 2013. https://wikimili.com/en/Proof_of_space

[14] Wikipedia. NEM (Cryptocurrency). 2021. https://en.everybodywiki.com/NEM_(cryptocurrency)#:~:text=NEM%2C%20New%20Economy%20Movement%20is%20a%20cryptocurrency%20written, the%20NXT%20project%20to%20implement%20the%20NEM%20blockchain

[15] Zyskind G, Nathan O, Pentland A. Decentralizing privacy: Using blockchain to protect personal data. In: Proc. of the 2015 IEEE Security and Privacy Workshops. San Jose: IEEE, 2015. 180–184. [doi: 10.1109/SPW.2015.27]

[16] Paul G, Sarkar P, Mukherjee S. Towards a more democratic mining in Bitcoins. In: Proc. of the 10th Int'l Conf. on Information Systems Security. Hyderabad: Springer, 2014. 185-203. [doi: 10.1007/978-3-319-13841-1_11]

[17] Castro M, Liskov B. Practical byzantine fault tolerance. In: Proc. of the 3rd Symp. on Operating Systems Design and Implementation. New Orleans: USENIX Association, 1999. 173–186. [doi: 10.5555/296806.296824]

[18] Li XQ, Jiang P, Chen T, Luo XP, Wen QY. A survey on the security of blockchain systems. Future Generation Computer Systems, 2020, 107: 841–853. [doi: 10.1016/j.future.2017.08.020]

[19] Vitalik Buterin. Ethereum GitHub implementation. 2021. https://github.com/ethereum/go-ethereum

[20] GmbH B. Meet BigchainDB. The blockchain database. 2018. https://github.com/BigchainDB/bigchaindb

[21] Ltd NG. Because together, everything is possible. 2015. https://staging.nem.io

[22] The Linux Foundation. Hyperledger document. 2015. https://www.hyperledger.org

[23] Block.one. Advance your business with EOSIO. 2018. https://eos.io

[24] Cheng YG, Jia ZJ, Gong B, Wang LP, Lei YF. Threshold signature scheme with strong forward security based on Chinese remainder theorem. In: Proc. of the 2nd Security and Privacy in New Computing Environments. Tianjin: Springer, 2019. 15–28. [doi: 10.1007/978-3-030-21373-2_2]

[25] Wang LP, Hu MS, Jia ZJ, Gong B, Lei YF. A signature scheme applying on blockchain voting scene based on the asmuth-bloom algorithm. In: Proc. of the 2018 IEEE 4th Int'l Conf. on Computer and Communications. Chengdu: IEEE, 2018. 2372–2378. [doi: 10.1109/CompComm.2018.8780775]

[26] Rivest RL, Shamir A, Adleman L. A method for obtaining digital signatures and public-key cryptosystems. Communications of the ACM, 1978, 21(2): 120–126. [doi: 10.1145/359340.359342]

[27] ElGamal T. A public key cryptosystem and a signature scheme based on discrete logarithms. IEEE Trans. Information Theory, 1985, 31(4): 469–472. [doi: 10.1109/TIT.1985.1057074]

[28] Koblitz N. Elliptic curve cryptosystems. Mathematics of Computation, 1987, 48(177): 203–209. [doi: 10.1090/S0025-5718-1987-0866109-5]

[29] Cooper D, Santesson S, Farrell S, Boeyen S, Housley R, Polk W. Internet X. 509 public key infrastructure certificate and certificate revocation list (CRL) profile. 2008. https://tools.ietf.org/html/rfc5280

[30] Wikipedia. Public key infrastructure. 2021. https://nl.wikipedia.org/wiki/Public_Key_Infrastructure

[31] Boneh D, Franklin M. Identity-based encryption from the Weil pairing. In: Proc. of the 21st Annual Int'l Cryptology Conf. Santa Barbara: Springer, 2001. 213–229. [doi: 10.1007/3-540-44647-8_13]

[32] Lewko A, Waters B. Unbounded HIBE and attribute-based encryption. In: Proc. of the 30th Annual Int'l Conf. on the Theory and Applications of Cryptographic Techniques. Tallinn: Springer, 2011. 547–567. [doi: 10.1007/978-3-642-20465-4_30]

[33] Al-Riyami SS, Paterson KG. Certificateless public key cryptography. In: Proc. of the 9th Int'l Conf. on the Theory and Application of Cryptology and Information Security. Taipei: Springer, 2003. 452–473. [doi: 10.1007/978-3-540-40061-5_29]

[34] Zhang JL, Hu MS, Gong B, Jia ZJ, Wang LP. A blind signature scheme applying on electronic payment scene based on quantum secret sharing. In: Proc. of the 2nd Int'l Conf. on Security and Privacy in New Computing Environments. Tianjin: Springer, 2019. 3–14. [doi: 10.1007/978-3-030-21373-2_1]

[35] The SSL protocol. 2013. https://library.netapp.com/ecmdocs/ECMP1155684/html/GUID-20073505-6C40-4A9B-85D9-D398C2991102.html

[36] Matsumoto S, Reischuk RM. IKP: Turning a PKI around with decentralized automated incentives. In: Proc. of 2017 IEEE Symp. on Security and Privacy. San Jose: IEEE, 2017. 410–426. [doi: 10.1109/SP.2017.57]

[37] Fotiou N, Polyzos GC. Decentralized name-based security for content distribution using blockchains. In: Proc. of the 2016 IEEE Conf. on Computer Communications Workshops. San Francisco: IEEE, 2016. 415–420. [doi: 10.1109/INFCOMW.2016.7562112]

[38] Wang LP, Hu MS, Jia ZJ, Cheng YG, Fu JJ, Wang YB, Gong B. Identity-based threshold group signature scheme of blockchain verification. In: Proc. of the 13th Chinese Conf. on Trusted Computing and Information Security. Shanghai: Springer, 2019. 144–158. [doi: 10.1007/978-981-15-3418-8_11]

[39] Zhang JL, Hu MS, Jia ZJ, Bei-Gong, Wang LP. A novel e-payment protocol implented by blockchain and quantum signature. Int'l Journal of Theoretical Physics, 2019, 58(4): 1315–1325. [doi: 10.1007/s10773-019-04024-8]

[40] Suárez-Otero P, Suárez-Cabal MJ, Tuya J. Leveraging conceptual data models for keeping cassandra database integrity. In: Proc. of the 14th Int'l Conf. on Web Information Systems and Technologies. Seville: SciTePress, 2018. 398–403. [doi: 10.5220/0007236303980403]

[41] Rodrigues R, Liskov B. High availability in DHTs: Erasure coding vs. replication. In: Proc. of the 4th Int'l Workshop on Peer-to-peer Systems. Ithaca: Springer, 2005. 226–239. [doi: 10.1007/11558989_21]

[42] Wang HQ, Wang QH, He DB. Blockchain-based private provable data possession. IEEE Trans. on Dependable and Secure Computing, 2021, 18(5): 2379–2389. [doi: 10.1109/TDSC.2019.2949809]

[43] Zikratov I, Kuzmin A, Akimenko V, Niculichev V, Yalansky L. Ensuring data integrity using blockchain technology. In: Proc. of the 20th Conf. of Open Innovations Association. St. Petersburg: IEEE, 2017. 534–539. [doi: 10.23919/FRUCT.2017.8071359]

[44] Liu B, Yu XL, Chen SP, Xu XW, Zhu LM. Blockchain based data integrity service framework for IoT data. In: Proc. of the 2017 IEEE Int'l Conf. on Web Services. Honolulu: IEEE, 2017. 468–475. [doi: 10.1109/ICWS.2017.54]

[45] Wilkinson S, Boshevski T, Brandof J, Prestwich J, Hall G, Gerbes P, Hutchins P, Pollard C. Storj a peer-to-peer cloud storage network. 2016. https://www.storj.io/storjv2.pdf

[46] ERICSSON. Data integrity assurance user guide. Implementation of service in predix. 2016. https://docplayer.net/53889118-Data-integrity-assurance-user-guide-implementation-of-service-in-predix.html

[47] Guardtime. Keyless signature infrastructure. 2015. https://m.guardtime.com/files/KSI_data_sheet_201509.pdf

[48] Deshpande S. Implementing blockchain as a microservice for IoT platforms. 2017. https://www.ericsson.com/en/blog/2017/4/implementing-blockchain-as-a-microservice-for-iot-platforms

[49] Barbosa M, Farshim P. Certificateless signcryption. In: Proc. of the 2008 ACM Symp. on Information, Computer and Communications Security. Tokyo: ACM, 2008. 369–372. [doi: 10.1145/1368310.1368364]

[50] Chandrasekhar S, Singhal M. Efficient and scalable aggregate signcryption scheme based on multi-trapdoor hash functions. In: Proc. of the 2015 IEEE Conf. on Communications and Network Security. Florence: IEEE, 2015. 610–618. [doi: 10.1109/CNS.2015.7346875]

[51] Tang JY, Zhang FG. A new code-based encryption scheme and its applications. Int'l Journal of High Performance Computing and Networking, 2017, 10(6): 515–523. [doi: 10.1504/IJHPCN.2017.087469]

[52] Liu ZY, Tso R, Tseng Y, Mambo M. Signcryption from NTRU lattices without random oracles. In: Proc. of the 14th Asia Joint Conf. on Information Security. Kobe: IEEE, 2019. 134–141. [doi: 10.1109/AsiaJCIS.2019.00009]

[53] Wang Z, Han YL, Liu WC, Chen L. Anti-quantum generalized signcryption scheme based on multivariate and coding. In: Proc. of the 2019 Chinese Control and Decision Conf. Nanchang: IEEE, 2019. 3587–3594. [doi: 10.1109/CCDC.2019.8832722]

[54] Cui WJ, Hu MS, Jia ZJ, Wang LP. A new signcryption scheme without hash or redundant functions. In: Proc. of the 6th Int'l Conf. on Behavioral, Economic and Socio-cultural Computing. Beijing: IEEE, 2019. 1–4. [doi: 10.1109/BESC48373.2019.8962991]

[55] Wang JZ, Li MR, He YH, Li H, Xiao K, Wang C. A blockchain based privacy-preserving incentive mechanism in crowdsensing applications. IEEE Access, 2018, 6: 17545–17556. [doi: 10.1109/ACCESS.2018.2805837]

[56] Wang LP, Chen Z, Guan Z, Li QS. Blockchain-based signcryption scheme with aging mechanism in crowdsensing applications. Chinese Journal of Computers, 2021, 44(11): 2216–2232 (in Chinese with English abstract). [doi: 10.11897/SP.J.1016.2021.02216]

[57] Elrom E. EOS. IO wallets and smart contracts. In: Apress E, ed. The Blockchain Developer. Berkeley: Apress, 2019. 213–256. [doi: 10.1007/978-1-4842-4847-8_6]

[58] Wang LP, Gao JB, Li QS, Chen Z. Blockchain-based multi-recipient multi-message signcryption scheme. Ruan Jian Xue Bao/Journal of Software, 2021, 32(11): 3606–3627 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/6034.htm [doi: 10.13328/j.cnki.jos.006034]

[59] Hammi MT, Bellot P, Serhrouchni A. BCTrust: A decentralized authentication blockchain-based mechanism. In: Proc. of the 2018 IEEE Wireless Communications and Networking Conf. Barcelona: IEEE, 2018. 1–6. [doi: 10.1109/WCNC.2018.8376948]

[60] Ellison C, Schneier B. Ten risks of PKI: What you're not being told about public key infrastructure. Computer Security Journal, 2000, 16(1): 1–8.

[61] Corella F. Implementing a PKI on a blockchain. 2016. https://pomcor.com/2016/10/25/implementing-a-pki-on-a-blockchain

[62] Gan S. An IoT simulator in NS3 and a key-based authentication architecture for IoT devices using blockchain. 2017. https://journal-home.s3.ap-northeast-2.amazonaws.com/site/2021s/presentation/0100.pdf

[63] Truong NB, Sun K, Lee GM, Guo YK. GDPR-compliant personal data management: A blockchain-based solution. IEEE Trans. on Information Forensics and Security, 2020, 15: 1746–1761. [doi: 10.1109/TIFS.2019.2948287]

[64] Allen C, Brock A, Buterin V, Callas J, Dorje D, Lundkvist C, Kravchenko P, Nelson J, Reed D, Sabadello M, Slepak G, Thorp N, Wood HT. Decentralized public key infrastructure. 2015. https://danubetech.com/download/dpki.pdf

[65] Li DX, Peng W, Deng WP, Gai FY. A blockchain-based authentication and security mechanism for IoT. In: Proc. of the 27th Int'l Conf. on Computer Communication and Networks. Hangzhou: IEEE, 2018. 1–6. [doi: 10.1109/ICCCN.2018.8487449]

[66] Guardtime. Internet of things authentication: A blockchain solution using SRAM physical unclonable functions. 2017. https://www.intrinsic-id.com/wp-content/uploads/2017/05/gt_KSI-PUF-web-1611.pdf

[67] Chen DY, Zhao H. Data security and privacy protection issues in cloud computing. In: Proc. of the 2012 Int'l Conf. on Computer Science and Electronics Engineering. Hangzhou: IEEE, 2012. 647–651. [doi: 10.1109/ICCSEE.2012.193]

[68] Gartner. Gartner says worldwide IaaS public cloud services market grew 37.3% in 2019. 2020. https://www.gartner.com/en/newsroom/press-releases/2020-08-10-gartner-says-worldwide-iaas-public-cloud-services-market-grew-37-point-3-percent-in-2019

[69] Salavi RR, Math MM, Kulkarni UP. A survey of various cryptographic techniques: From traditional cryptography to fully homomorphic encryption. In: Saini HS, Sayal R, Govardhan A, Buyya R, eds. Innovations in Computer Science and Engineering. Singapore: Springer, 2019. 295–305. [doi: 10.1007/978-981-13-7082-3_34]

[70] Yi X, Paulet R, Bertino E. Homomorphic Encryption and Applications. Cham: Springer, 2014. [doi: 10.1007/978-3-319-12229-8]

[71] Byun JW, Kamra A, Bertino E, Li NH. Efficient *k*-anonymization using clustering techniques. In: Proc. of the 12th Int'l Conf. on Database Systems for Advanced Applications. Bangkok: Springer, 2007. 188–200. [doi: 10.1007/978-3-540-71703-4_18]

[72] Bayardo RJ, Agrawal R. Data privacy through optimal k-anonymization. In: Proc. of the 21st Int'l Conf. on Data Engineering. Tokyo: IEEE, 2005. 217–228. [doi: 10.1109/ICDE.2005.42]

[73] Machanavajjhala A, Kifer D, Johannes G, Gehrke J. *L*-diversity: Privacy beyond *K*-anonymity. ACM Trans. on Knowledge Discovery from Data, 2007, 1(1): 3. [doi: 10.1145/1217299.1217302]

[74] Li NH, Li TC, Venkatasubramanian S. t-closeness: Privacy beyond k-anonymity and l-diversity. In: Proc. of the 23rd IEEE Int'l Conf. on Data Engineering. Istanbul: IEEE, 2007. 106–115. [doi: 10.1109/ICDE.2007.367856]

[75] Dwork C. Differential privacy: A survey of results. In: Proc. of the 5th Int'l Conf. on Theory and Applications of Models of Computation. Xi'an: Springer, 2008. 1–19. [doi: 10.1007/978-3-540-79228-4_1]

[76] Bertino E, Ferrari E. Big data security and privacy. In: Flesca S, Greco S, Masciari E, Saccà D, eds. A Comprehensive Guide Through the Italian Database Research over the Last 25 Years. Cham: Springer, 2017. 425–439. [doi: 10.1007/978-3-319-61893-7_25]

[77] Kreuter B, Mood B, Shelat A, Kevin B. PCF: A portable circuit format for scalable two-party secure computation. In: Proc. of the 22nd USENIX Security Symp. Washington: USENIX Association, 2013. 321–336.

[78] Duffield E, Diaz D. Dash: A privacycentric cryptocurrency. 2015. https://whitepaperdatabase.com/cgi-sys/suspendedpage.cgi

[79] Li WX, Guo H, Nejad M, Shen CC. Privacy-preserving traffic management: A blockchain and zero-knowledge proof inspired approach. IEEE Access, 2020, 8: 181733–181743. [doi: 10.1109/ACCESS.2020.3028189]

[80] Noether S. Ring signature confidential transactions for monero. https://eprint.iacr.org/2015/1098

[81] Zhong HR, Sang YP, Zhang YC, Xi ZC. Secure multi-party computation on blockchain: An overview. In: Proc. of the 10th Int'l Symp. on Parallel Architectures, Algorithms and Programming. Guangzhou: Springer, 2019. 452–460. [doi: 10.1007/978-981-15-2767-8_40]

[82] Lind J, Naor O, Eyal I, Kelbert F, Sirer EG, Pietzuch P. Teechain: A secure payment network with asynchronous blockchain access. In: Proc. of the 27th ACM Symp. on Operating Systems Principles. Huntsville: ACM, 2019. 63–79. [doi: 10.1145/3341301.3359627]

[83] She W, Gu ZH, Lyu XK, Liu Q, Tian Z, Liu W. Homomorphic consortium blockchain for smart home system sensitive data privacy preserving. IEEE Access, 2019, 7: 62058–62070. [doi: 10.1109/ACCESS.2019.2916345]

[84] Kosba A, Miller A, Shi E, Wen ZK, Papamanthou C. Hawk: The blockchain model of cryptography and privacy-preserving smart contracts. In: Proc. of the 2016 IEEE Symp. on Security and Privacy. San Jose: IEEE, 2016. 839–858. [doi: 10.1109/SP.2016.55]

[85] Miers I, Garman C, Green M, Rubin AD. Zerocoin: Anonymous distributed E-cash from bitcoin. In: Proc. of the 2013 IEEE Symp. on Security and Privacy. Berkeley: IEEE, 2013. 397–411. [doi: 10.1109/SP.2013.34]

[86] Yaji S, Bangera K, Neelima B. Privacy preserving in blockchain based on partial homomorphic encryption system for AI applications. In: Proc. of the 25th IEEE Int'l Conf. on High Performance Computing Workshops. Bengaluru, IEEE: 2018. 81–85. [doi: 10.1109/HiPCW.2018.8634280]

[87] Guan ZT, Si GL, Zhang XS, Wu LF, Guizani N, Du XJ, Ma YL. Privacy-preserving and efficient aggregation based on blockchain for power grid communications in smart communities. IEEE Communications Magazine, 2018, 56(7): 82–88. [doi: 10.1109/MCOM.2018.1700401]

[88] Shafagh H, Burkhalter L, Hithnawi A, Duquennoy S. Towards blockchain-based auditable storage and sharing of IoT data. In: Proc. of the 2017 on Cloud Computing Security Workshop. Dallas: ACM, 2017. 45–50. [doi: 10.1145/3140649.3140656]

[89] Xia Q, Sifah EB, Smahi A, Amofa S, Zhang XS. BBDS: Blockchain-based data sharing for electronic medical records in cloud environments. Information, 2017, 8(2): 44. [doi: 10.3390/info8020044]

[90] Reardon J, Basin D, Capkun S. Sok: Secure data deletion. In: Proc. of the 2013 IEEE Symp. on Security and Privacy. Berkeley: IEEE, 2013. 301–315. [doi: 10.1109/SP.2013.28]

[91] Perlman R. File system design with assured delete. In: Proc. of the 3rd IEEE Int'l Security in Storage Workshop. San Francisco: IEEE, 2005. 83–88. [doi: 10.1109/SISW.2005.5]

[92] Tang Y, Lee PPC, Liu JCS, Perlman R. FADE: Secure overlay cloud storage with file assured deletion. In: Proc. of the 6th Int'l Conf. on Security and Privacy in Communication Systems. Singapore: Springer, 2010. 380–397. [doi: 10.1007/978-3-642-16161-2_22]

[93] Geambasu R, Kohno T, Levy AA, Levy HM. Vanish: Increasing data privacy with self-destructing data. In: Proc. of the 18th USENIX Security Sym. Montreal: USENIX Association, 2009. 299–316.

[94] Mo Z, Qiao Y, Chen SG. Two-party fine-grained assured deletion of outsourced data in cloud systems. In: Proc. of the 34th IEEE Int'l Conf. on Distributed Computing Systems. Madrid: IEEE, 2014. 308–317. [doi: 10.1109/ICDCS.2014.39]

[95] Yuan Y, Wang FY. Editable blockchain: Models, techniques and methods. Acta Automatica Sinica, 2020, 46(5): 831–846 (in Chinese with English abstract). [doi: 10.16383/j.aas.2020.y000002]

[96] Ateniese G, Magri B, Venturi D, Andrade E. Redactable blockchain-or-rewriting history in bitcoin and friends. In: Proc. of the 2017 IEEE European Symp. on Security and Privacy. Paris: IEEE, 2017. 111–126. [doi: 10.1109/EuroSP.2017.37]

[97] Giuseppe A, T CM, David T, Magri B, Venturi D. Multiple-link blockchain. 2017. https://patents.google.com/patent/US9785369B1/en

[98] Giuseppe A, T CM, David T, Magri B, Venturi D. Rewritable blockchain. 2018. https://patents.google.com/patent/US10348707B2/en

[99] Derler D, Samelin K, Slamanig D, Striecks C. Fine-grained and controlled rewriting in blockchains: Chameleon-hashing gone attribute-based. In: Proc. of the 26th Annual Network and Distributed System Security Symp. San Diego: The Internet Society, 2019. 18–21.

[100] Ren YL, Xu DT, Zhang XP, Gu DW. Scheme of revisable blockchain. Ruan Jian Xue Bao/Journal of Software, 2020, 31(12): 3909–3922 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/5894.htm [doi: 10.13328/j.cnki.jos.005894]

[101] Park S, Kwon A, Fuchsbauer G, Gaži P, Alwen J, Pietrzak K. SpaceMint: A cryptocurrency based on proofs of space. In: Proc. of the 22nd Int'l Conf. on Financial Cryptography and Data Security. Nieuwpoort: Springer, 2018. 480–499. [doi: 10.1007/978-3-662-58387-6_26]

[102] Marsalek A, Zefferer T. A correctable public blockchain. In: Proc. of the 18th IEEE Int'l Conf. on Trust, Security and Privacy in Computing and Communications/the 13th IEEE Int'l Conf. on Big Data Science and Engineering. Rotorua: IEEE, 2019. 554–561. [doi: 10.1109/TrustCom/BigDataSE.2019.00080]

[103] Deuber D, Magri B, Thyagarajan SAK. Redactable blockchain in the permissionless setting. In: Proc. of the 2019 IEEE Symp. on Security and Privacy. San Francisco: IEEE, 2019. 124–138. [doi: 10.1109/SP.2019.00039]

[104] James-Lubin K. Blockchain Scalability. Sebastopol: O'Reilly Media, 2015.

[105] Dinh TTA, Wang J, Chen G, Liu R, Ooi BC, Tan KL. Blockbench: A framework for analyzing private blockchains. In: Proc. of the 2017

ACM Int'l Conf. on Management of Data. Chicago: ACM, 2017. 1085–1100. [doi: 10.1145/3035918.3064033]

附中文参考文献:

[8] 中华人民共和国国家质量监督检验检疫总局, 中国国家标准化管理委员会. GB/T 35273-2017 信息安全技术 个人信息安全规范. 北京: 中国标准出版社, 2018.

[56] 王利朋, 陈钟, 关志, 李青山. 群智感知中基于区块链的带时效签密方案. 计算机学报, 2021, 44(11): 2216–2232. [doi: 10.11897/SP.J.1016.2021.02216]

[58] 王利朋, 高健博, 李青山, 陈钟. 应用区块链的多接收者多消息签密方案. 软件学报, 2021, 32(11): 3606–3627. http://www.jos.org.cn/1000-9825/6034.htm [doi: 10.13328/j.cnki.jos.006034]

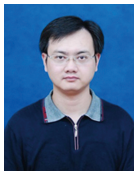
[95] 袁勇, 王飞跃. 可编辑区块链: 模型、技术与方法. 自动化学报, 2020, 46(5): 831–846. [doi: 10.16383/j.aas.2020.y000002]

[100] 任艳丽, 徐丹婷, 张新鹏, 谷大武. 可修改的区块链方案. 软件学报, 2019, 31(12): 3909–3922. http://www.jos.org.cn/1000-9825/5894.htm [doi: 10.13328/j.cnki.jos.005894]

王利朋(1987－), 男, 硕士, 主要研究领域为网络安全.

陈钟(1963－), 男, 博士, 教授, 博士生导师, CCF 会士, 主要研究领域为区块链.

关志(1980－), 男, 博士, 副研究员, CCF 专业会员, 主要研究领域为密码学.



胡明生(1973－), 男, 博士, 教授, CCF 高级会员, 主要研究领域为网络安全.

李青山(1977－), 男, 博士, 主要研究领域为区块链.