软件学报 ISSN 1000-9825, CODEN RUXUEW
Journal of Software,2019,30(9):2620−2635 [doi: 10.13328/j.cnki.jos.005770]

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

基于区块链的档案数据保护与共享方法*

谭海波[1], 周 桐[1,2], 赵 赫[1], 赵 哲[1,2], 王卫东[1], 张中贤[1], 盛念祖[1,2], 李晓风[1]

[1](中国科学院 合肥物质科学研究院,安徽 合肥 230031)

[2](中国科学技术大学,安徽 合肥 230026)

通讯作者: 赵赫, E-mail: zhaoh@hfcas.ac.cn

摘 要: 针对现有档案数据管理中普遍存在的数据中心化存储、安全性差和防篡改性弱等问题,提出一种基于区块链的档案数据保护与共享方法:通过智能合约和数字签名技术,实现了数字档案馆的身份认证和档案所有权的确定;通过智能合约和星际文件系统(IPFS)等技术,实现了数字档案的保护、验证、恢复与共享;通过公有链与联盟链结合的方式,在降低经济成本的同时保障了数据的安全性,提高了可扩展性.该方法具有去中心化、安全可信和不可篡改等特点,有望促进档案馆数据存储方式的转型,以满足日益增长的档案数据保护与共享的需求.

关键词: 数字档案;区块链;智能合约;数字签名;IPFS

中图法分类号: TP311

中文引用格式: 谭海波,周桐,赵赫,赵哲,王卫东,张中贤,盛念祖,李晓风.基于区块链的档案数据保护与共享方法.软件学报, 2019,30(9):2620−2635. http://www.jos.org.cn/1000-9825/5770.htm

英文引用格式: Tan HB, Zhou T, Zhao H, Zhao Z, Wang WD, Zhang ZX, Sheng NZ, Li XF. Archival data protection and sharing method based on blockchain. Ruan Jian Xue Bao/Journal of Software, 2019,30(9):2620−2635 (in Chinese). http://www.jos.org.cn/1000-9825/5770.htm

Archival Data Protection and Sharing Method Based on Blockchain

TAN Hai-Bo[1], ZHOU Tong[1,2], ZHAO He[1], ZHAO Zhe[1,2], WANG Wei-Dong[1], ZHANG Zhong-Xian[1], SHENG Nian-Zu[1,2], LI Xiao-Feng[1]

[1](Hefei Institutes of Physical Science, Chinese Academy of Sciences, Hefei 230031, China)

[2](University of Science and Technology of China, Hefei 230026, China)

Abstract: In view of the problems existing in the management of archival data, such as centralized data storage, poor security, and low tamper resistant modification, this study proposes an archival data protection and sharing method based on blockchain technology. The identification of digital archives and the determination of archives ownership have been achieved through smart contracts and digital signature technologies; digital archives files have been protected, verified, restored, and shared through technologies such as smart contracts and the interplanetary file system (IPFS); the economic cost have been reduced meanwhile data security could be guaranteed and data extendibility improved through the combination of permissioned and permissionless blockchains. Featuring the characteristics of decentralization, security, credibility, and tamper-resistant, this method is expected to promote the data storage way transformation in archives so as to meet the increasing demand for the protection and sharing of archival data.

Key words: digital archive; blockchain; smart contract; digital signature; IPFS

* 基金项目: 国家自然科学基金(61602435); 安徽省自然科学基金(1708085QF153); 安徽省科技重大专项(16030901057)

Foundation item: National Natural Science Foundation of China (61602435); Natural Science Foundation of Anhui (1708085QF153); Major Projects of Science and Technology of Anhui (16030901057)

本文由“区块链数据管理”专题特约编辑于戈教授、牛保宁教授、金澈清教授推荐.

收稿时间: 2018-06-05; 修改时间: 2018-08-28; 采用时间: 2018-12-14; jos 在线出版时间: 2019-04-10

CNKI 网络优先出版: 2019-04-11 09:52:37, http://kns.cnki.net/kcms/detail/11.2560.TP.20190411.0952.001.html

档案是一种重要的数据记录,是人们在各种社会活动中直接形成的具有保存价值的原始信息.区别于一般的图书情报资料和电子文档信息,档案的本质属性是原始记录性,它使得档案能够还原真实的历史状况,因而具备重要的保存与参考价值,并且拥有法律效力.

随着 IT 技术的高速发展,人们大规模地运用数字化的手段来提升档案的存储和处理效率.现有的数字档案系统为档案的管理工作带来便利的同时,仍然存在以下主要问题[1]:(1) 相比传统的纸质物理档案,数字档案作为保存在存储介质上的比特字节,具有高度的易变性,并且数字档案在存储、传输和处理等过程中容易被修改;(2) 每个档案馆都是档案信息的“孤岛”,档案馆之间缺乏安全有效的档案共享渠道;(3) 现有的数字档案保护方案大多是通过数字水印和数字签名等技术实现,在档案遭到篡改或破坏后一般难以恢复.

档案数据的真实性和原始性主要依赖于对系统中心或第三方实体的信任,如系统主节点、中心数据库以及系统负责人、数据库管理员等.一旦上述系统中心不再可信(例如系统数据库遭入侵,或管理员被胁迫或收买),档案数据的真实性将荡然无存.

区块链是一种运行于对等网络中各节点遵从特定共识机制的公共账本技术,它具有去中心化、无需信任、防篡改性强等优点[2−4],有望解决现有数字档案馆中普遍存在的数据安全性低、共享性差等问题.本文基于区块链以及 IPFS(inter planetary file system)[5]、数字签名和混合加密等技术,提出了以下解决方法.

1) 数字档案馆之间构建联盟链,共同管理和维护此区块链的稳定运行,并将数据快照定期与公有链进行锚定,保障数据安全的同时,降低了数据保护成本;
2) 通过私有 IPFS 集群加密存储档案原始数据,配合链上的智能合约存储档案指纹等摘要信息,实现了档案信息的保护、验证、恢复与共享;
3) 采用将档案修改记录和档案摘要历史版本存储到智能合约上的方式,实现了对档案修改操作责任人的追责和档案历史版本的回溯;
4) 基于本文提出的数据档案保护和共享方法,研发了档案数据保护与共享系统,该系统与传统的档案管理系统以接口的形式进行对接,为现有的档案数据提供额外的安全保护.

本文第 1 节介绍数字档案馆建设和区块链数据保护领域的相关工作.第 2 节介绍本方法涉及的区块链和 IPFS 等相关技术.第 3 节介绍整体架构,包括系统架构和智能合约架构.第 4 节介绍方法的具体设计.第 5 节介绍方法的具体实现.第 6 节对本方法进行分析与评估.最后在第 7 节对本文工作进行总结.

1 相关工作

自 20 世纪 90 年代起,美国就开始了数字档案馆的相关研究[6],并随后推广到其他发达国家.2004 年,美国华盛顿州数字档案馆完成建设,存储了超过 1 亿份档案文件,其中部分档案通过网络共享至其他国家[7].2009 年,瑞士国家数字档案馆建设完成,通过多地备份等机制,实现档案文件的长期储存[8].上世纪 90 年代末,我国国家档案局提出了构建数字档案馆的规划,以实现纸质档案和音、视频档案的数字化,达到档案的长期存储、高效共享和快捷查询等目标[9].2014 年,国家档案局《数字档案室建设指南》中明确提出了数字档案室的建设原则及内容[10],为推动我国数字档案馆的建设与发展奠定了基础.2015 年,肖敏等人基于大数据分析技术,挖掘各类数字档案之间的关联性和潜在价值,为档案管理人员提供更有针对性的决策建议[11].2016 年,福建省数字档案馆信息平台的建设过程中,利用了云计算等技术来解决档案分散各地的“孤岛效应”问题,提升了档案的共享、处理和存储效率[12].2017 年,王伟等人在数字档案馆建设过程中融合了射频识别技术(RFID),在增加数字档案安全性的同时,提高了档案管理工作的效率[13].

数字档案的相关规范标准已日趋完善,大数据、云计算、物联网等新技术正逐步应用到数字档案馆的建设中.但受限于中心化的存储技术与管理方式,档案的原始性、真实性和安全性等问题尚未得到妥善解决,档案被伪造、窃取和篡改等恶性事件屡有发生[14].

近年来,国内外许多研究人员和机构运用区块链技术在数据保护与共享领域进行一些探索与实践[1].2013 年,Araoz 等人创立了 Proof of Existence 项目,通过将哈希值存储到区块链交易 OP_RETURN 字段的方法,实现

了电子文件的真实性保护[15].Vaughan 等人提出了 Chainpoint 项目,基于区块链实现了一种通用的文件保护框架,该方案计算文件哈希值并以此构建默克尔树[16],以降低数据保护的成本[17].2016 年,Azaria 等人利用智能合约构建了去中心化的医疗数据访问与权限管理系统,实现患者对其医疗数据的所有权,使其自主地对医疗记录进行分享与管理[18].同年 11 月,蔡维德等人提出了基于区块链的应用系统开发方法,包含了账户链、交易链双链设计模型,以及链上代码并行执行模型应用原则[19].2017 年,Rifi 等人[20]通过 IPFS 存储物联网(IoT)设备数据、智能合约控制访问权限的方式,实现了 IoT 设备数据的访问保护.但该方案基于公有区块链和公有 IPFS 实现,仍然具有数据存储成本高和隐私性较差等问题.同年,蚂蚁金服公司基于区块链技术开展捐赠善款流向的追踪管理,提升其系统和数据的透明性、可追溯性和不可篡改性[21].2018 年,百度将区块链技术应用于百度百科的数据保护,将百科词条每次更新的历史版本和作者、编辑时间等信息记录到区块链上,达到数据保护与存证的目的[22].2017 年 9 月,薛腾飞等人综合利用了区块链、MIFS、AFS 和 DDBS 等技术,提出了基于区块链的医疗数据共享模型,实现了各医疗机构的数据共享/但是该方案依然将医疗数据存储在中心化数据库中,没有很好地解决医疗数据篡改和恢复等问题[23].同年 10 月,章宁等人利用区块链、数据库和非对称加密等技术,实现了一种基于区块链的个人隐私数据保护的解决方案框架,初步实现了在互联网租车场景中个人隐私的保护/但是该方案设计的数据交互审计平台运行在中心化服务器上,仍然可能存在安全性和稳定性方面的隐患[24].自 2014 年起,该实验室的研究团队也开展了区块链技术的相关研究,并在物联网和医疗健康数据[25−28]等多个应用场景进行了实践[1].

2 相关技术

一般而言,区块链结构是以数据区块为单位、按照时间顺序链接而成;数据区块则是由分布式节点通过共识算法产生、并以一定的经济激励确保所有节点都有动力参与到区块链的活动中来.分布式系统中的所有节点地位均等,不存在任何中心化的特殊节点,且每个节点均会验证区块数据、传播区块数据,从而保证少量节点的作恶不会影响到整个区块链系统的运行.区块链可分为公有链、私有链、联盟链,其中,公有链被称为非许可链,任何组织或个人都可以参与共识,并具有数据的读写权限;私有链适用于单位或组织的内部系统使用[29],其数据的读写权限是由该组织控制的,不能够完全解决信任问题;联盟链也被称为许可链,其共识由联盟成员参与,数据读写权限按联盟规则制定,节点的加入需要联盟其他节点的同意.主流的公有链和联盟链实例比较[30]见表 1.

Table 1 Comparision between mainstream permission and permissionless blockchains

表 1 主流的公有链和联盟链实例比较

名称	类型	共识算法	智能合约语言	使用开销	安全性
Bitcoin	公有链	PoW	基于栈的脚本	极高	极高
Ethereum	公有链	PoW/PoS	Solidity	高	高
Hyperledger Fabric	联盟链	PBFT/SBFT	Go/Java	低	较高
本方法	公有链+联盟链	PoW+PoA	Solidity	低	高

共识算法致力于解决在去中心化的分布式互连网络中所有的节点如何达成一致的问题[32].目前,区块链系统中的共识算法主要包括工作量证明(proof of work,简称 PoW)、股权证明(proof of stake,简称 PoS)和权威证明(proof of authority,简称 PoA)等[32−34],其中,比特币采用 PoW 共识算法,要求系统的各节点基于自身算力共同求解一个计算复杂但验证容易的数学难题,最快解决该难题的节点将获得一个区块打包的权利;在 PoS 系统中,持有最多数字货币而非最高算力的节点具有最大的概率打包下一个区块,从而解决了 PoW 中算力资源浪费的问题;PoA 则是一种由指定的权威节点产生区块的共识机制,所有的权威节点地位相同,可以通过投票的方式踢出或加入权威节点[35].相比于 PoW 和 PoS 共识机制,虽然 PoA 的去中心化程度以及节点的竞争公平性和匿名性较弱,但是可靠的权威节点认证机制和高效的共识效率更加适用于联盟链的使用场景,并且避免了算力浪费和 51%攻击的问题.现实场景中的每个数字档案馆维护一个权威节点,由组织机构的声誉担保,每个档案馆都会积极遵守共识算法,维护联盟链运行,不会轻易做出损害自身名誉的行为.

智能合约是可以在区块链上自动执行的特殊程序,其特点是程序代码以及数据均存储于链上,因此拥有防

篡改性强、去中心化程度高等特点[36,37].智能合约以交易的形式被创建和调用,合约程序在分布式网络中的所有节点被执行,因此不存在中心节点,且任何节点发生故障都不会影响合约程序的运行[1].以太坊是主流的支持智能合约的区块链,通过以太坊虚拟机实现了智能合约的功能,允许开发人员使用类似JavaScript语法的高级语言 Solidity 进行智能合约的开发[38].

IPFS 是全球互联的分布式文件系统,它综合了包括分布式哈希表、块交换、版本控制系统和自我认证文件等系统的优点,具有内容可寻址、不可篡改、去中心化等特点[39].在存储文件时,IPFS 会根据文件内容计算得出文件指纹,在获取文件时,IPFS 根据文件指纹从存储节点中取出文件并验证之后返回给用户[1].IPFS 分为私有集群和公有集群:公有集群是指全网 IPFS 节点构成的分布式网络,任何人均可作为一个节点加入到该网络中;私有集群仅限于某个团体或组织内部使用,具有相同 swarm-key 的节点才可以参与到该网络中[40].

3 整体架构

3.1 系统架构

如图 1 所示,本文提出的数据档案保护和共享方法由数字档案馆、联盟区块链、公有区块链、私有 IPFS 集群、系统服务(RESTful service)这 5 部分构成系统协同完成:数字档案馆(digital archive,简称 DA)作为数字档案联盟的权威节点参与到联盟区块链中,享有档案保护、验证、共享等服务;系统服务是一种去中心化应用(DApp),本身不存储任何档案数据和身份信息,以 RESTful 接口的形式为数字档案馆系统提供智能合约和 IPFS 接口调用;公有区块链采用基于 PoW 共识算法的以太坊区块链(ethereum),通过定期存储联盟链区块快照信息,实现了对联盟链上数据的保护;联盟区块链采用基于 PoA 共识算法的以太坊联盟链,通过智能合约存储档案馆的数字身份和档案的摘要信息,实现了身份的注册与恢复和档案的保护与共享等业务逻辑,并通过定期与公有链锚定的方式增强数据的原始性和真实性保护;私有 IPFS 集群存储了加密档案的原始信息,并通过 swarm.key 进行节点的身份认证,通过分布式哈希表(DHT)[41]、块交换(BitTorrent)[42]等技术保障了数据的安全性.基于公有区块链、联盟区块链和私有 IPFS 集群结合的数字档案数据流如图 2 所示.

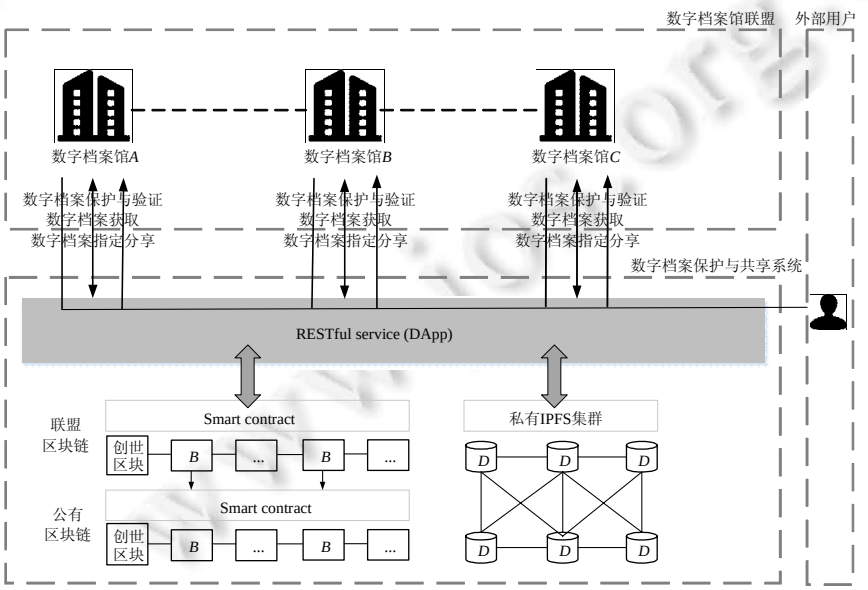


Fig.1 System artchitecture diagram

图 1 系统架构图

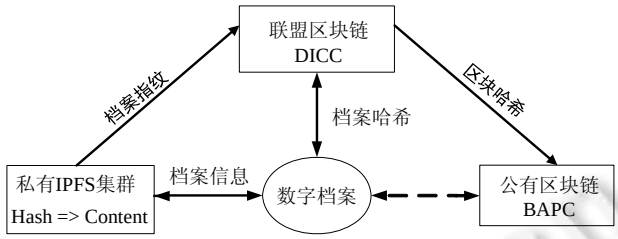


Fig.2 Data flow chart of digital archives

图 2 数字档案数据流图

3.2 合约架构

本方法主要利用智能合约一致性和不可篡改的特性,在区块链原始安全功能的基础上,还具有灵活性和可编程性.本方法的智能合约体系(如图 3 所示)由公有链中区块数据保护合约(block data protection contract,简称 BAPC)和联盟链合约共同构成:BAPC 合约部署在公有链中,并存储联盟链的数据快照信息;联盟链中的智能合约包括数字身份控制合约(digital identity controller contract,简称 DICC)、数字身份管理合约(digital identity manage contract,简称 DIMC)、数字档案管理合约(digital archive manage contract,简称 DAMC)等.

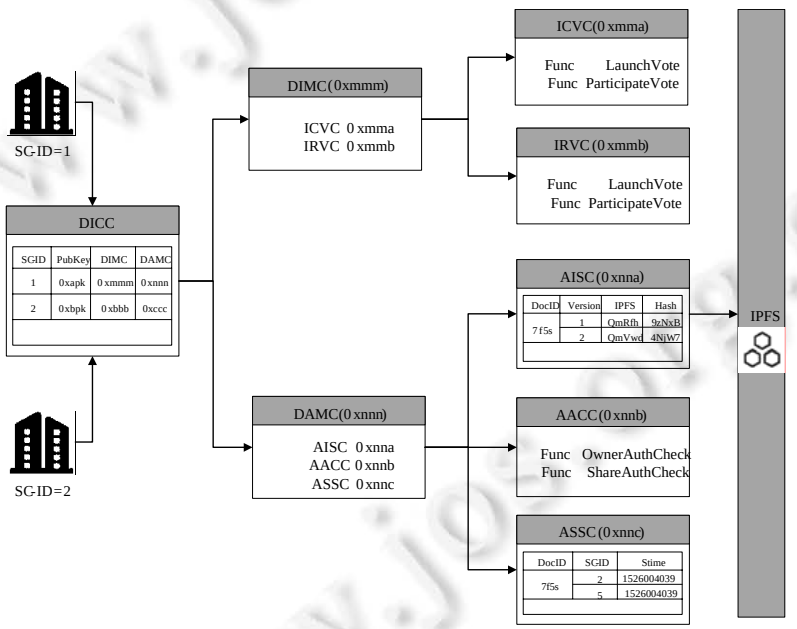


Fig.3 Structure chart of smart contracts

图 3 智能合约架构图

- DICC 作为全局合约记录该联盟链中所有档案馆的数字身份标识(SC-ID)、对应的公钥(PubKey)和与其相关联的 DIMC 和数字档案管理合约(DMAC).在创建 DICC 合约时,首个档案馆的数字身份以及相关合约被一同创建;
- DIMC 通过民主投票的方式实现档案馆联盟数字身份的内部自治,包括身份创建投票合约(identity creation vote contract,简称 ICVC)和身份重置合约(identity reset vote contract,简称 IRVC):ICVC 合约用于为新加入联盟的档案馆在 DICC 合约中创建投票请求并为其投票;IRVC 合约用于在档案馆私钥泄露时,通过民主投票的方式重置其公钥.

- DAMC 用于实现数字档案保护、验证、恢复和共享等业务逻辑,包括档案信息存储合约(archive information storage contract,简称 AISC)、档案共享(信息)存储合约(archive sharing storage contract,简称 ASSC)和档案权限控制合约(archive authority control contract,简称 AACC): AACC 合约用于数字档案的权限控制;AISC 合约用于存储档案的摘要信息,包括数字档案对象(DocJSON)的 IPFS 地址、哈希值、版本号、创建时间和最后修改时间等;ASSC 合约用于存储档案馆分享的档案信息,包括档案编号(DocID)、档案馆身份标识(SC-ID)和分享时间等.一个 DocID 可以共享给多个 SC-ID 访问.ASSC 中的 SC-ID 的分类如表 2 所示:若 DocID 对应的 SC-ID 的值为−1,则该档案是完全对外部开放的,外部用户或档案馆都可以直接从 ASSC 合约中获取档案 IPFS 地址;若为 0,则联盟内部的所有成员都访问;若 SC-ID 的值≥1 时,则只有指定 SC-ID 的档案馆有权访问该档案.

Table 2 SC-ID classification

表 2 SC-ID 分类

SC-ID	优先级	含义
−1	高	联盟内部档案馆和外部用户都可通过 RESTful Service 访问
0	中	联盟内部的所有档案馆都可访问
≥1	低	仅联盟内的指定档案馆可访问

4 方法设计

4.1 数字档案馆身份的注册与找回

数字档案的数字身份代表此档案馆参与到联盟链中,是档案保护、共享等活动的基础.每个希望加入联盟的数字档案馆都需要获得半数以上联盟成员的同意,方可为其在联盟中完成数字身份的注册.具体流程如下.

- 首先,加入联盟的数字档案馆基于 ECDSA 椭圆曲线算法[43]在本地秘密地生成公私密钥对$\langle PK,SK\rangle$,其中,SK 本地秘密存储;
- 然后,通过可靠信道,将 PK 及其身份信息发送给联盟内的所有成员,并委托某个成员通过其 ICVC 合约创建数字身份的投票请求,其他成员则通过其 ICVC 合约参与投票;
- 最后,当票数超过半数时,DICC 合约保存该档案馆的公钥信息,然后为其生成 SC-ID 并创建 DIMC 和 DAMC 等一系列合约,从而完成数字档案馆身份的注册.

由于私钥由档案馆秘密存储,一旦内部人员或黑客非法窃取到私钥,就可以伪造该档案馆的身份进行档案的查看、修改和分享等操作.因此,一方面需要档案馆妥善保管 SK,避免因管理不善导致 SK 的泄露;另一方面,本文设计了基于投票机制的密钥重置方案.

- 首先,档案馆秘密地重新生成一对新的公私密钥$\langle PK_{NEW},SK_{NEW}\rangle$,通过可靠信道,将新的 PK_{NEW} 和 SC-ID 等信息发送给联盟内的其他成员,并请求某一成员为其创建重置投票;
- 其他联盟成员为其进行重置投票,当票数超过半数时,DICC 合约重置该档案馆的公钥.

本方法在设计过程中充分考虑了档案馆密钥重置的可能,通过身份编号 SC-ID 实现了公钥与业务逻辑的解耦.档案查询与共享等操作都基于 SC-ID 进行权限判断,即使重置了 DICC 合约中的公钥,只要签名对应的公钥与 DICC 中的 SC-ID 保持一致,依然可以验证档案馆的数字身份.

4.2 数字档案的保护与验证

数字档案保护是指通过联盟链和公有链的链上智能合约,配合 IPFS 私有集群将数字档案存储在区块链上,防止其内容被非法篡改和破坏,并提供验证和恢复等操作,从而达到保护数字档案的目的.新增档案是将数字档案对象 DocJSON 保护在链下的私有 IPFS 集群和其档案指纹保存在链上的 AISC 合约中;更新档案则将每次更新生成的 DocJSON 档案对象和更新档案时产生的数据进行保护.如图 4 所示,数字档案对象是 JSON 形式的档案信息组织结构,包含档案编号、版本号、创建时间、操作管理员和档案附件等信息.

Fig.4 Structure diagram of JSON of digital archives

图 4 数字档案 JSON 对象结构图

档案新增流程见算法 1,DA 首先生成一对随机的密钥 *edk*(*key*,*iv*)用于档案附件和档案对象的加密,然后,先用 edk 对档案附件加密后存储到 IPFS 集群,并将附件哈希值($Hash_{Files}^{sha256}$)、加密附件的指纹($Ipfs_{EncryptedFiles}^{add}$)和其他档案属性整合为档案对象(DocJSON)加密后存入 IPFS 集群,并对档案馆身份(SC-ID)、档案编号(DocID)、档案对象哈希值($Hash_{DocJson}^{sha256}$)和加密档案对象的档案指纹($Ipfs_{ciphertext}^{add}$)等信息进行签名,通过 RESTful Service 发送到智能合约进行处理.AISC 合约收到新增档案请求后,调用 AACC 合约从签名中恢复公钥信息,并与 DICC 合约中登记的密钥进行对比:若身份检查通过,则在合约中添加档案编号与摘要等信息的映射.

算法 1. Saving of Archive.

1: **Procedure** *SaveArchive*(*DocID*,*DocAttrs*,*Files*)
2: system executes:
3: generate a random *keyPair*→*edk*⟨*key*,*iv*⟩
4: $AES_{encrypt}(Files,key,iv)\rightarrow EncryptedFiles$
5: $extract(Hash_{Files}^{sha256},Ipfs_{EncryptedFiles}^{add})\rightarrow Files_{attrs}$
6: $combine(DocAttrs,Files_{attrs})\rightarrow DocJSON_{plaintext}$
7: $AES_{encrypt}(DocJSON_{plaintext},key,iv)\rightarrow DocJSON_{ciphertext}$
8: $ECDSA_{sign}(SC\text{-}ID,DocID,Hash_{DocJson}^{sha256},Ipfs_{ciphertext}^{add})\rightarrow signature$
9: contract executes:
10: **if** $AACC_{OwnerAuthCheck}(signature,SC\text{-}ID)$=true **then**
11: $AISC_{mapping}^{add}(DocId,[version,IpfsAddr,Hash,Time])$
12: **endif**
13: **end Procedure**

档案更新操作流程与新增类似,不同之处在于 DA 不会重新生成 *edk*,而是使用在新增档案时创建的 *edk*;DA 会根据 DocID 从 AISC 合约和 IPFS 中取出 DocJSON 并解密,然后根据更新的档案信息生成新的 DocJSON 并在加密后存储至 IPFS 和 AISC 合约中.

档案验证操作包括公有链对联盟链上的数据验证、联盟链对 IPFS 中档案数据的验证和 DocJSON 对档案馆本地数据库中档案信息的验证这 3 部分,具体流程见算法 2.

算法 2. Validation of Archives.

1: **Procedure** *ValidateArchive*(*DocID*)
2: system executes:
3: $signature\rightarrow ECDSA_{sign}(PrivateKey_A,SC\text{-}ID,DocID)$

4: service executes:
5: **if** $validate(PrivateChain_{getBlock},BDPC_{getlastblock})$=false **then**
6: return $error_{invalidChain}(DocID,timestamp)$
7: **endif**
8: contract executes:
9: **if** $AACC_{OwnerAuthCheck}(signature,SC\text{-}ID)$=true **then**
10: **return** $AISC_{DocID}^{getlastversion} \rightarrow List\langle DocJSON_{ipfs},DocJSON_{hash}\rangle$
11: **endif**
12: system executes:
13: $AES_{decrypt}(Ipfs_{cipherjson}^{get}(DocJSON_{ipfs}),edk) \rightarrow DocJSON_{plaintext}$
14: **if** $validate(DocJSON_{hash},sha256(DocJSON_{plaintext}))$=false **then**
15: **return** $error_{invalidIpfs}(DocID,DocJSON_{ipfs},timestamp)$
16: **endif**
17: **if** $validate(DocJSON_{plaintext}, LocalDB_{DocID}^{select})$=false **then**
18: **return** $error_{invalidDB}(DocID,timestamp)$
19: **endif**
20: **end Procedure**

DA 对 SC-ID,DocID 等信息进行签名发送到 RESTful Service 处理,Service 收到请求后,从公有链上的 BDPC 合约中获取最新的联盟链区块快照信息,并与联盟链中的区块信息进行比对验证:若验证不通过,则返回联盟链数据异常错误;若验证通过,则将签名发送到智能合约处理.AISC 合约收到请求后,先通过 AACC 合约对档案馆身份进行检查,然后,根据 DocID 从合约中查询该档案的摘要信息 $List\langle DocJSON_{ipfs},DocJSON_{hash}\rangle$并返回.DA 从 AISC 合约获取信息后,先根据 $DocJSON_{ipfs}$ 从 IPFS 集群中获取 $DocJSON_{ciphertext}$,然后根据本地的 edk 信息解密得到 $DocJSON_{plaintext}$,并验证其哈希值是否与 $DocJSON_{hash}$ 一致:若验证不通过,则返回 IPFS 数据异常错误.最后,将可信任的 $DocJSON_{plaintext}$ 与本地数据库中的档案信息进行比对验证:若验证不通过,则返回本地数据异常的错误.

在档案验证过程中产生的数据异常错误,本文提供与之对应的恢复手段.

- 联盟区块链数据异常:通过联盟链和公有链上的 BDPC 合约存储的区块信息进行比对,发现异常后可以继续与 BDPC 合约之前存储的区块信息进行对比,定位出异常区块高度,并在联盟链中基于此区块高度重新开始创建新的区块;
- IPFS 集群数据异常:由于 AISC 合约中保存有该档案各历史版本的数字指纹和哈希值,一旦系统侦测到当前档案信息被篡改,可恢复前期已保存的正确版本;
- 本地数据库数据异常:依照链上保存的可信档案文件来重置本地数据库被攻击者篡改的档案.相较于去中心化的区块链及 IPFS,本地数据库中的档案数据更易被篡改,因此这是本方案中需要重点考虑的保护内容.

4.3 数字档案的共享与获取

数字档案的共享是指在数字档案联盟成员内部或数字档案联盟与外部用户之间,通过智能合约、IPFS 和混合加密机制实现了安全可靠的档案数据的共享,并使得传统的档案系统可以安全高效地获取区块链系统,从而保护档案数据,具体流程如图 5 所示.

步骤 1:数字档案馆 A(DA-A)使用私钥 SK_A 对待分享档案编号 DocID、分享目标档案馆 SC-ID-B 等信息进行签名,并通过 Service 发送到智能合约.ASSC 合约在收到请求后,先调用 AACC 合约,通过签名对档案馆身份进行检查,检查通过后,将 SC-ID-B 写入合约中 DocID 对应的分享列表;

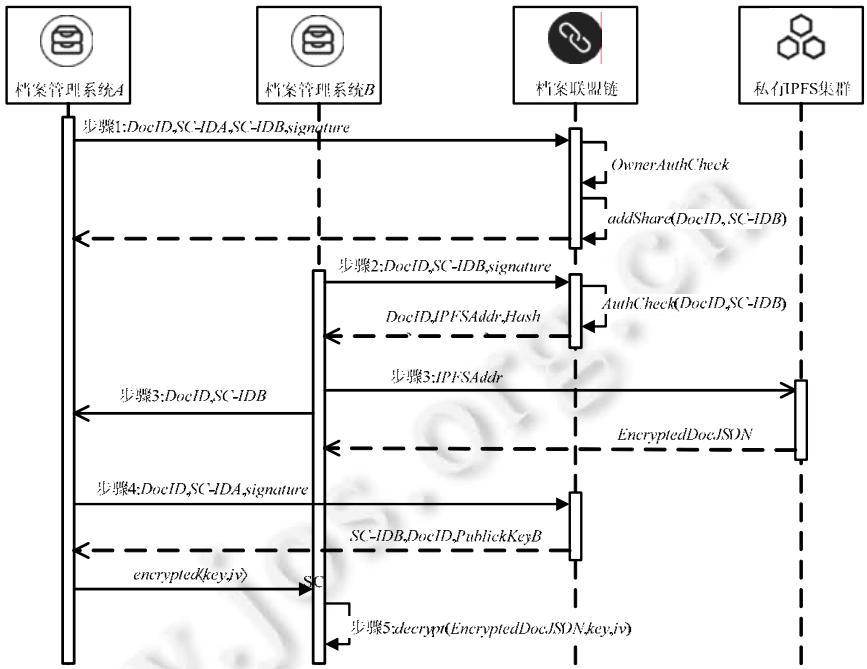


Fig.5 Flow chart of archievs sharing and acquisition

图 5 档案共享与获取流程图

步骤 2:数字档案馆 *B*(DA-B)使用私钥 SK_B 对待分享档案编号 DocID 和身份标识 SC-ID-B 等信息进行签名,并发送到智能合约.ASSC 合约在收到请求后,先调用 AACC 合约对 DA-B 进行见算法 3 的权限检查,检查通过后,返回档案指纹和对应的哈希值;

算法 3. Authority Check.

1: **Procedure** *AuthorityCheck*(*DocID*,*SC-ID*,*hash*,*signature*)
2: **if** $ASSC_{DocID}^{isPublicShare}$ =true **then**
3: **return** true
4: **endif**
5: *ecrecover*(*hash*,*signature*)→*PublicKey*
6: **if** $DICC_{checkPublicKey}$(*SC-ID*,*PublicKey*)=false **then**
7: **return** false
8: **endif**
9: **if** $ASSC_{DocID}^{isInnerShare}$ =true||$AACC_{OwnerAuthCheck}$(*SC-ID*)=true **then**
10: **return** true
11: **endif**
12: **return** $ASSC_{checkShareList}$(*DocID*,*SC-ID*)
13: **end Procedure**

步骤 3:DA-B 根据从合约中获取的档案指纹,从 IPFS 集群中异步获取加密的档案对象 $DocJSON_{ciphertext}$,同时,DA-B 通过异步 https 请求将身份标识 SC-ID-B 和 DocID 发送给 DA-A,获取 DocJSON 的解密密钥;

步骤 4:DA-A 收到 DA-B 的请求后,根据 DocID 和 SC-ID-B 等参数,通过 ASSC 合约检查共享记录的真实性,并从 DIMC 合约中获取 SC-ID-B 对应的 PK_B,然后使用 PK_B 对解密密钥 *edk* 进行非对称加密,并返回给 DA-B,见公式(1).

$$ECDSA_{encrypt}(PK_B,edk\langle key,iv\rangle)\rightarrow encrypted_{\langle key,iv\rangle} \tag{1}$$

步骤 5:DA-B 收到 DA-A 的返回数据后,使用私钥 SK_B 进行解密得到原始的 $edk\langle key,iv\rangle$,然后使用 edk 对 $DocJSON_{ciphertext}$ 进行解密,得到原始的档案对象 $DocJSON_{plaintext}$;还可以根据 JSON 结构中的档案附件指纹从 IPFS 中获取附件密文,并通过 edk 进行解密查看,见公式(2)~公式(4).

$$ECDSA_{encrypt}(SK_B,encrypted\langle key,iv\rangle)\rightarrow edk_{\langle key,iv\rangle} \tag{2}$$

$$AES_{encrypt}(DocJSON_{ciphertext},edk)\rightarrow DocJSON_{plaintext} \tag{3}$$

$$AES_{decrypt}(Ipfs^{get}_{EncryptedFiles},edk)\rightarrow Files \tag{4}$$

5 方法实现

基于本文提出的数据档案保护和共享方法研发了档案数据保护与共享系统,其实现主要分为区块链智能合约开发、私有 IPFS 集群搭建、RESTful Service(DApp)开发、数字档案系统研发这 4 部分.其中,

- 智能合约部分,首先使用 puppeth 程序生成创世区块的相关信息 Creation.json,并为数字档案馆联盟链中的初始档案馆分配权威节点;然后使用 go-ethereum[44]客户端,基于 Creation.json 创建联盟链中的节点;最后,使用 go-ethereum 客户端进行以太坊公有链的同步,以便实现进行公有链 BAPC 合约的调用.本文使用 Solidity 语言进行 BAPC,DICC 等合约的开发,并使用 truffle 框架进行合约的管理、编译、调试和部署;
- 对于 IPFS 部分,本文使用了 go-ipfs[45]客户端进行本地私有 IPFS 集群的搭建,通过环境变量 LIBP2P_FORCE_PNET 的配置启用私有集群模式,限制仅具有相同 swarm.key 文件的节点可进行访问;
- 对于 RESTful Service(DApp),本文采用基于 Node.js 的 Web 框架 Express 进行开发,并通过 web3.js 实现了对智能合约的调用,通过 js-ipfs-api 实现了对私有 IPFS 的接口调用.

本系统已在中国科学院合肥物质科学研究院档案馆(下属多个研究所,形成联盟)进行初步试用,对研究院的基建档案和大科学工程等档案进行保护,在现有的传统档案管理系统的基础上,将录入和更新操作的档案信息通过 RESTful Service 接口同步到区块链上,并为传统档案管理系统增加了档案的验证、共享和恢复功能.

数字档案验证操作的界面如图 6 所示,档案的修改历史追溯和分享界面如图 7 所示;每个档案馆运行了一个权威节点和多个数据同步节点,自 2018 年 5 月至今档案联盟链已创建 1 057 459 区块,保护了 10 342 条档案信息.

(a) 数字档案验证结果一致界面 (b) 数字档案验证结果不一致界面

Fig.6 Interface of verification results of digital archives

图 6 数字档案验证界面

(a) 数字档案历史追溯界面 (b) 数字档案分享界面

Fig.7 Interface of digital archives history and sharing

图 7 数字档案历史和分享界面

6 分析与评估

6.1 运行成本分析

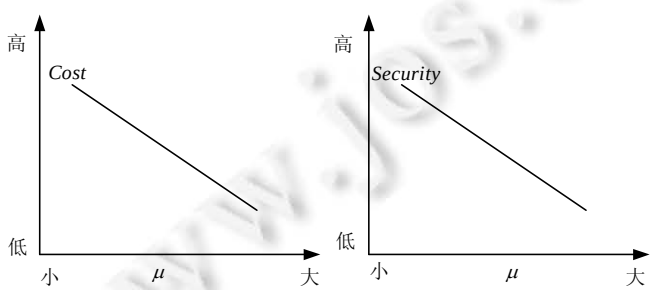
假设系统中每月需进行档案保护操作的次数为α,档案分享操作次数为β,身份管理操作次数为γ,则在以太坊公有链上部署智能合约的每月成本为 $Cost_{ethereum}$,见公式(5).

$$Cost_{ethereum}=\alpha\times Cost_{protection}+\beta\times Cost_{share}+\gamma\times Cost_{identity} \tag{5}$$

本文提出的联盟链与公有链相结合的方案,可以将大多数智能合约部署于联盟链中,在联盟链中部署和调用合约的成本可以忽略不计,仅需考虑联盟链网络每月的运行成本记为ω.联盟链每进行μ次数据存储操作时调用公有链 BDPC 合约锚定联盟链中当前的区块快照信息,锚定频率可以动态调整,见公式(6).

$$Cost_{privateChain}^{ethereum}=\frac{\alpha+\beta+\gamma}{\mu}Cost_{protection}+\omega,\mu>1 \tag{6}$$

在联盟链规模较小时,可以频繁锚定;随着联盟链的规模不断扩大,系统的安全性、稳定性随之提升,锚定频率也可以逐渐减小.μ参数的值越大,该方案的经济成本越低,BDPC 合约存储的区块高度间隔越大,对联盟区块链的保护和数据可恢复程度也会降低(如图 8 所示)[1].



(a) μ与*Cost*关系图 (b) μ与*Security*关系图

注:μ表示公有链保护的区块高度间隔,*Cost* 表示系统运行的经济成本,*Security* 表示系统的安全性

Fig.8 Relations diagram between parameter μ and systemative cost and security

图 8 参数μ与系统成本和安全性关系图

根据 EthGasStation 的数据统计[46],在撰写本文时(2018 年 5 月),以太坊调用智能合约的平均交易费用为 8Gwei,折合成人民币约为 0.7 元.调用智能合约的交易费用与存储数据的字节数的有关,为简化计算,我们忽略

该差异,统一使用平均交易费用.当联盟链的运行成本ω(电力成本)忽略不计时,公有链和联盟链结合的方案可以将经济成本大约缩减为原来的 1/μ,见公式(7).

$$\mu \approx \frac{Cost_{ethereum}}{Cost_{privateChain}^{ethereum}}, \mu > 1 \tag{7}$$

6.2 安全性评估

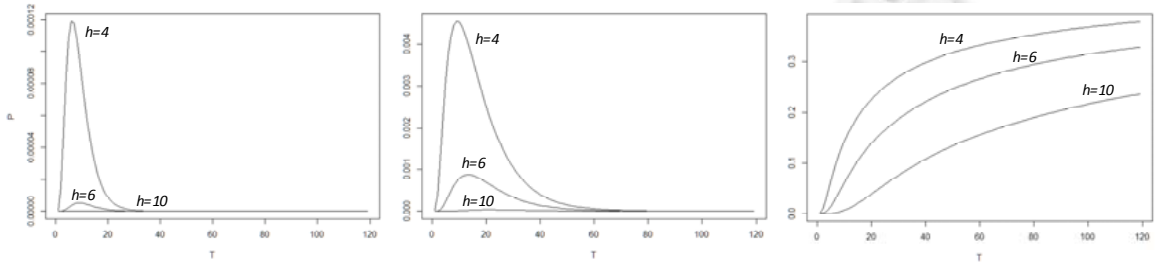
假设联盟链中的区块高度为α的快照信息存储在公有链中高度为β的区块中,距离当前区块的高度差为 h,若攻击者想要通过分叉的方式替代该区块,则需要重新计算并生成从第β个区块到当前区块高度的所有区块,并通过全网节点的验证.假设当前公有区块链中诚实节点的算力为 p 次哈希每秒,攻击者控制的节点算力为 q 次哈希每秒.只要没有大量节点的加入或退出,新区块的计算难度不会有明显增加.为了方便计算,假设没有新节点的加入,且每秒中诚实节点产生新区块的概率为 w,攻击节点获得新区块的概率为 u.假设 h 表示诚实节点与攻击节点区块高度差,则每秒高度差 h 有 3 种可能的结果[28],即高度差缩小、高度差变大和高度差不变,每个结果出现的概率分别为 P_1,P_2,P_3.假设在 t 秒内会出现 t 种结果,每个结果出现的次数用随机变量 X_1,X_2,X_3 表示,其中,X_1 表示发生的次数为 n,X_2 表示发生的次数为 m,X_3 表示发生的次数为 $t-m-n$.诚实节点与攻击节点的高度差 h 的变化概率符合多项分布[28].

如在 t 秒内,攻击节点想要追上诚实节点,则需要满足 $n\in[0,(t-h-1)/2]$.$m=n+h+k$,且 $1\leqslant k\leqslant t-2n-h$,其概率见公式(8).

$$P_h(t)=\sum_{n=0}^{(t-h-1)/2}\sum_{k=1}^{t-2n-h}\frac{t!}{m!n!(t-m-n)}P_1^nP_2^mP_3^{t-m-n} \tag{8}$$

其中,$P_1=w(1-u)$表示诚实节点产生新区块,而攻击节点没有,即高度差变大的概率;$P_2=u(1-w)$表示攻击节点产生新区块,而诚实节点没有,即高度差缩小的概率;$P_3=1-P_1-P_2$ 表示两者都产生或都没有产生新区块的概率,即高度差不变的概率.

根据公式(8)得到攻击节点成功篡改区块数据的概率.如图 9 所示,h 为攻击者实现数据篡改需要替代的区块数量,横坐标 T 为时间(单位:诚实节点发现一个区块的平均时间),纵坐标 P 为攻击者篡改数据成功的概率.



(a) 攻击节点算力为诚实节点的 25% (b) 攻击节点算力为诚实节点的 50% (c) 攻击节点算力与诚实节点相等

Fig.9 Probability distribution when attackers successfully tampered with block data

图 9 攻击者成功篡改区块数据的概率分布图

由图可知:攻击者成功篡改区块数据的概率会随着区块高度差 h 的增大而减小.当攻击节点的算力小于诚实节点时,由于攻击节点需要预先生成新的 h 个区块,然后再与诚实节点竞争产生区块,在刚开始的一段时间内,P 会先上升,然后逐渐减小并最终趋于 0;当攻击节点算力为诚实节点的 25%、高度差 h=4 时,攻击成功的最高概率不到 0.01%,高度差为 6 和 10 时,攻击者成功的概率基本为 0;当攻击节点算力为诚实节点的 50%、区块高度差 h=4 时,攻击者成功的最高概率不到 0.05%;当攻击节点算力与诚实节点相当,在一段时间后,成功概率逼近 35%[28].

在数字档案保护的应用场景中,篡改档案的操作往往是发生在与利益相关的一些考核过程中,往往距离档案录入日期较为久远,因此区块高度差 x 会为千或万的数量级,即使攻击节点与诚实节点相当时,完成这种规模

的区块替换操作也是几乎不可能的.由于以太坊激励机制的存在,假如攻击者拥有如此巨大的算力,其作为诚实节点所产生区块的奖励也远远大于数据篡改的收益[47].

本方法通过公有链与联盟链的结合实现了一种链式保护机制,通过公有链 BDPC 合约锚定联盟链的区块快照信息,实现对联盟链上数据的保护与验证;通过联盟链 AISC 合约存储档案指纹和哈希值的方式,实现对IPFS 中档案对象的保护与验证;通过 IPFS 存储加密的原始档案对象,实现对本地数据库中的档案信息的保护与验证.但若仅采用以太坊公有链的方案,诚然此方案安全性最高,但是档案的操作成本也随之变高,并且操作效率较低(尤其是改变合约状态等操作,需要公有链的矿工打包数据).此外,一旦公有链被攻击(虽然概率很低),档案系统也会受影响;若仅采用联盟链的方案,运行的节点相对较少(尤其在运行初期),系统的安全性和稳定性不如公有链,还会出现联盟中部分成员共谋造假的风险;而采用基于联盟链和公有链的方式,通过第 6.1 节和第 6.2 节的分析可以看出:运行 PoA 共识算法的以太坊联盟链,本身具备较强的防篡改能力,此外,使用高安全性公有链进行额外保护,还能够防范部分成员共谋造假对联盟链进行攻击的风险,并且不仅限于单一公有链,还可以拓展为多条公有链进行协同保护.3 种方案相互对比结果见表 3.

Table 3 Comparision between different protection schemes

表 3 不同保护方案对比

方案	数据安全性	经济成本	方案可扩展性
基于联盟链	较高	低	一般
基于公有链	高	高	一般
基于联盟链+公有链	高	较低	较好

7 结 语

本文基于智能合约、IPFS、数字签名和混合加密等技术设计了一种基于区块链的档案数据保护与共享方法,并以此为基础研发了档案数据保护与共享系统.该方法通过智能合约和数字签名技术实现了档案馆联盟的内部自治;通过公有链与联盟链的结合,实现了档案数据的保护、验证与恢复;利用智能合约和 IPFS 技术实现了档案数据的共享与获取.未来工作中,我们将就档案数据管理的区块链共识机制开展进一步深入研究.总的来说,本文中的方案综合考虑了经济成本、安全性、易用性和扩展性等问题,系统能以较低的经济开销和较少的系统修改,实现对现有档案管理系统中的档案数据进行高效安全的保护与共享,为数字档案馆的建设和数字资源的保护提供一些有益的启发.

References:

[1] Zhao Z. Research and design of digital archive management system based on blockchain [M.S. Thesis]. Hefei: University of Science and Technology of China, 2018 (in Chinese with English abstract).

[2] He P, Yu Y, Zhang YF, Bao YK. Survey on blockchain technology and its application prospect. Computer Science, 2017,44(4):1−7 (in Chinese with English abstract). [doi: 10.11896/j.issn.1002-137X.2017.04.001]

[3] Tschorsch F, Scheuermann B. Bitcoin and beyond: A technical survey on decentralized digital currencies. IEEE Communications Surveys & Tutorials, 2016,18(3):2084−2123.

[4] Nakamoto S. Bitcoin: A peer-to-peer electronic cash system. 2009. https://bitcoin.org/bitcoin.pdf

[5] Benet J. IPFS—Content addressed, versioned, P2P file system (DRAFT 3). https://raw.githubusercontent.com/ipfs/papers/master/ipfs-cap2pfs/ipfs-p2p-file-system.pdf

[6] Wang DM. The theory and practice of the construction of digital archives in colleges and universities. Journal of Dongbei University of Finance and Economics, 2014,(6):85−88 (in Chinese with English abstract). [doi: 10.3969/j.issn.1008-4096.2014.06.014]

[7] Cheng YY. Analysis of Int'l digital archives construction and core operational mechanism. Archives & Construction, 2014,12:31−34 (in Chinese with English abstract).

[8] Cunningham A, Millar L, Reed B, Peter J. Scott and the Australian ‘Series’ system: Its origins, features, rationale, impact and continuing relevance. In: Proc. of the International Congress on Archives, 2013. 121−144. [doi: 10.3828/comma.2013.1.13]

[9] Shi ZW. Gradual progress and steady progress—On the construction of Shandong digital archives (room). Shandong Archives, 2013,3:10−11 (in Chinese with English abstract).

[10] Liu YJ. Research on the construction of digital archive office in China [MS. Thesis]. Hefei: Anhui University, 2017 (in Chinese with English abstract).

[11] Xiao M. Research on the construction of archives services utilization system under the big data environments [MS. Thesis]. Xiangtan: Xiangtan University, 2015 (in Chinese with English abstract).

[12] You SS. The research on the service of Fujian province digital archives information on the basic of cloud computing [MS. Thesis]. Fuzhou: Fujian Agriculture and Forestry University, 2016 (in Chinese with English abstract).

[13] Wang W. Design of file management system based on RFID and study of anti-collision algorithm [MS. Thesis]. Hangzhou: Hangzhou Dianzi University, 2017 (in Chinese with English abstract).

[14] Wang ZD. Precaution measures against archive frauds should be taken at basic levels. China Youth Daily, 20170526 (in Chinese). http://zqb.cyol.com/html/2017-05/26/nw.D110000zgqnb_20170526_4-01.htm

[15] Proof of existence—An online service to prove the existence of documents. 2018. https://docs.proofofexistence.com/

[16] Merkle RC. A digital signature based on a conventional encryption function. In: Proc. of the Int’l Conf. on the Theory and Applications of Cryptographic Techniques on Advances in Cryptology. Springer-Verlag, 1987. 369−378.

[17] Verify a chainpoint proof directly using Bitcoin. 2017. https://runkit.com/tierion/verify-a-chainpoint-proof-directly-using-bitcoin

[18] Azaria A, Ekblaw A, Vieira T, *et al*. MedRec: Using blockchain for medical data access and permission management. In: Proc. of the Int’l Conf. on Open and Big Data. IEEE, 2016. 25−30. [doi: 10.1109/OBD.2016.11]

[19] Tsai WT, Yu L, Wang R, Liu N, Deng EY. Blockchain application development techniques. Ruan Jian Xue Bao/Journal of Software, 2017,28(6):1474−1487 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/5232.html [doi: 10.13328/j.cnki.jos.005232]

[20] Rifi N, Rachkidi E, Agoulmine N, *et al*. Towards using blockchain technology for IoT data access protection. In: Proc. of the Int’l Conf. on Ubiquitous Wireless Broadband. IEEE, 2018. 1−5. [doi: 10.1109/ICUWB.2017.8251003]

[21] Blockchain+public welfare, concept or trend. 2017. http://www.xinhuanet.com/gongyi/2016-12/21/c_129414848.htm

[22] Baidu’s ‘Wikipedia’ now logs revisions on a blockchain. 2018. https://www.coindesk.com/baidus-wikipedia-now-logs-revisions-on-a-blockchain

[23] Xue TF, Fu CQ, Wang Z, Wang XY. A medical data sharing model via blockchain. Acta Automatica Sinica, 2017,43(9):1555−1562 (in Chinese with English abstract). [doi: 10.16383/j.aas.2017.c160661]

[24] Zhang N, Zhong S. Mechanism of personal privacy protection based on blockchain. Journal of Computer Applications, 2017,37(10): 2787−2793 (in Chinese with English abstract). [doi: 10.11772/j.issn.1001-9081.2017.10.2787]

[25] Zhao H, Li XF, Zhan LQ, Wu ZC. Data integrity protection method for microorganism sampling robots based on blockchain technology. Journal of Huazhong University of Science and Technology (Natural Science Edition), 2015,43(s1):216−219 (in Chinese with English abstract). [doi: 10.13245/j.hust.15s1052]

[26] Li XF, Zhao H, Li F, Tan HB, Sun YN, Liu B. Electronic document anti-tampering method. Chinese Patent ZL201410436231.7, 2014-08-29 (in Chinese).

[27] Zhao H, Zhu XY, Li XF, Tan HB, Wang WD, Zhang ZX, Lv B, Zhou T, Zhao Z, Wang L, Sheng NZ. A digital archives management method and system based on blockchain technology. Chinese Patent CN201711226383.4, 2017-11-29 (in Chinese).

[28] Sheng NZ, Li F, Li XF, Zhao H, Zhou T. Data capitalization method based on blockchain smart contract for internet of things. Journal of Zhejiang University (Engineering Science), 2018,52(11):1−10 (in Chinese with English abstract). [doi: 10.3785/j.issn.1008-973X.2018.11.000].

[29] On public and private blockchains. 2015. https://blog.ethereum.org/2015/08/07/on-public-and-private-blockchains/

[30] Shao QF, Jin CQ, Zhang Z, Qan WN, Zhou AY. Blockchain: Architecture and research progress. Chinese Journal of Computers, 2018,41(5):969−988 (in Chinese with English abstract). [doi: 10.11897/SP.J.1016.2018.00969]

[31] He YJ, Gong GC. Research on blockchain technology in security related fields of Internet of things. Telecom Engineering Technics and Standardization, 2017,30(5):12−16 (in Chinese with English abstract). [doi: 10.13992/j.cnki.tetas.2017.05.004]

[32] King S, Nadal S. PPCoin: Peer-to-peer crypto-currency with proof-of-stake. 2012. https://peercoin.net/assets/paper/peercoin-paper.pdf

[33] DPOs consensus algorithm—The missing white paper. 2016. https://steemit.com/dpos/@dantheman/dpos-consensus-algorithm-this-missing-white-paper

[34] POA Network Whitepaper. 2018. https://github.com/poanetwork/wiki/wiki/POA-Network-Whitepaper

[35] Proof of authority: Consensus model with identity at stake. https://medium.com/poa-network/proof-of-authority-consensus-model-with-identity-at-stake-d5bd15463256

[36] Christidis K, Devetsikiotis M. Blockchains and smart contracts for the Internet of things. IEEE Access, 2016. [doi: 10.1109/ACCESS.2016.2566339]

[37] Cruz JP, Kaji Y, Yanai N. RBAC-SC: Role-based access control using smart contract. IEEE Access, 2018. [doi: 10.1109/ACCESS.2018.2812844]

[38] Buterin V. A next-generation smart contract and decentralized application platform. 2014. https://github.com/ethereum/wiki/wiki/White-Paper.

[39] Chen Y, Li H, Li K, *et al*. An improved P2P file system scheme based on IPFS and blockchain. In: Proc. of the Int'l Conf. on Big Data. IEEE, 2017. 2652−2657. [doi: 10.1109/BigData.2017.8258226]

[40] Experimental features of Go-IPFS. https://github.com/ipfs/go-ipfs/blob/master/docs/experimental-features.md#private-networks

[41] Baumgart I, Mies S. S/Kademlia: A practicable approach towards secure key-based routing. In: Proc. of the 2007 Int'l Conf. on Parallel and Distributed Systems, Vol.2. IEEE, 2007. 1−8. [doi: 10.1109/ICPADS.2007.4447808]

[42] Cohen B. Incentives build robustness in Bittorrent. In: Proc. of the Workshop on Economics of Peer-to-Peer Systems, Vol.6. 2007. 68−72.

[43] Johnson D, Menezes A, Vanstone S. The elliptic curve digital signature algorithm (ECDSA). Int'l Journal of Information Security, 2001,1(1):36−63.

[44] Go ethereum. https://geth.ethereum.org/

[45] IPFS documentation. https://docs.ipfs.io/

[46] EthGasStation. https://ethgasstation.info/

[47] Antonopoulos AM. Mastering Bitcoin: Unlocking Digital Cryptocurrencies. 1st ed. Sebastopol: O'Reilly Media, Inc., 2014. 174−177.

附中文参考文献:

[1] 赵哲.基于区块链的档案管理系统的研究与设计[硕士学位论文].合肥:中国科学技术大学,2018.

[2] 何蒲,于戈,张岩峰,鲍玉凯.区块链技术与应用前瞻综述.计算机科学,2017,44(4):1−7. [doi: 10.11896/j.issn.1002-137X.2017.04.001]

[6] 王冬梅.高校数字档案馆建设的理论与实践.东北财经大学学报,2014,(6):85−88. [doi: 10.3969/j.issn.1008-4096.2014.06.014]

[7] 程妍妍.国外数字档案馆建设及核心运行机制分析.档案与建设,2014,12:31−34.

[9] 史志伟.循序渐进稳步进取——谈山东数字档案馆(室)建设.山东档案,2013,3:10−11.

[10] 刘英俊.我国数字档案室建设研究[硕士学位论文].合肥:安徽大学,2017.

[11] 肖敏.大数据环境下档案利用服务体系建设研究[硕士学位论文].湘潭:湘潭大学,2015.

[12] 游姗姗.基于云计算的福建省数字档案馆信息服务研究[硕士学位论文].福州:福建农林大学,2016.

[13] 王伟.基于 RFID 的档案管理系统设计及防碰撞算法研究[硕士学位论文].杭州:杭州电子科技大学,2017.

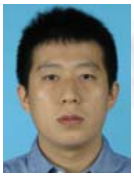
[14] 王钟的.警惕档案造假要从基层抓起.中国青年报,20170526. http://zqb.cyol.com/html/2017-05/26/nw.D110000zgqnb_20170526_4-01.htm

[19] 蔡维德,郁莲,王荣,刘娜,邓恩艳.基于区块链的应用系统开发方法研究.软件学报,2017,28(6):1474−1487. http://www.jos.org.cn/1000-9825/5232.html [doi: 10.13328/j.cnki.jos.005232]

[21] 区块链+公益,概念还是趋势.2017. http://www.xinhuanet.com/gongyi/2016-12/21/c_129414848.htm

[23] 薛腾飞,傅群超,王枞,王新宴.基于区块链的医疗数据共享模型研究.自动化学报,2017,43(9):1555–1562. [doi: 10.16383/j.aas.2017.c160661]
[24] 章宁,钟珊.基于区块链的个人隐私保护机制.计算机应用,2017,37(10):2787–2793. [doi: 10.11772/j.issn.1001-9081.2017.10.2787]
[25] 赵赫,李晓风,占礼葵,吴仲城.基于区块链技术的采样机器人数据保护方法.华中科技大学学报(自然科学版),2015,43(s1):216–219. [doi: 10.13245/j.hust.15s1052]
[26] 李晓风,赵赫,李芳,谭海波,孙怡宁,刘冰.一种电子文件防篡改方法.ZL201410436231.7[发明专利],2014-08-29.
[27] 赵赫,朱晓煜,李晓风,谭海波,王卫东,张中贤,吕波,周桐,赵哲,王丽,盛念祖.一种基于区块链技术的数字档案管理方法及系统.CN201711226383.4[发明专利],2017-11-29.
[28] 盛念祖,李芳,李晓风,赵赫,周桐.基于区块链智能合约的物联网数据资产化方法.浙江大学学报(工学版),2018,52(11):1–10. [doi: 10.3785/j.issn.1008-973X.2018.11.000]
[30] 邵奇峰,金澈清,张召,钱卫宁,周傲英.区块链技术:架构及进展.计算机学报,2018,(5):969–988. [doi: 10.11897/SP.J.1016.2018.00969]
[31] 何渝君,龚国成.区块链技术在物联网安全相关领域的研究.电信工程技术与标准化,2017,(5):12–16. [doi: 10.13992/j.cnki.tetas.2017.05.004]

谭海波(1976－),男,安徽泾县人,博士,正高级工程师,主要研究领域为计算机应用技术,区块链技术,网络传输与测量.



周桐(1991－),男,博士生,主要研究领域为区块链技术,分布式系统,健康信息学.

赵赫(1984－),男,博士,高级工程师,主要研究领域为区块链技术,计算机应用技术,网络安全.

赵哲(1993－),男,硕士生,主要研究领域为计算机应用技术,软件工程.

王卫东(1987－),男,工程师,主要研究领域为区块链技术,运动与健康,网络通信.

张中贤(1988－),男,工程师,主要研究领域为区块链技术,人工智能,网络安全.

盛念祖(1992－),男,硕士生,主要研究领域为区块链技术,物联网,软件架构.

李晓风(1966－),男,博士,教授,博士生导师,主要研究领域为区块链技术,计算机网络,计算机自动控制.