第 53 卷第 2 期 郑州大学学报(理学版) Vol. 53 No. 2
2021 年 6 月 J. Zhengzhou Univ. (Nat. Sci. Ed.) Jun. 2021

基于区块链的医疗信息隐私保护研究综述

刘 炜[1,2,3], 彭宇飞[3,4], 田 钊[1,4], 盛朝阳[2,3], 李 阳[2,3], 佘 维[1,3,4]

(1. 郑州大学 软件学院 河南 郑州 450002; 2. 郑州大学 汉威物联网研究院 河南 郑州 450001;
3. 郑州大学 互联网医疗与健康服务河南省协同创新中心 河南 郑州 450001;
4. 郑州大学 信息工程学院 河南 郑州 450001)

摘要: 随着医疗信息化的发展,医疗数据在共享与访问过程中的隐私问题引起了研究者的广泛关注。区块链作为去中心化、匿名、不可篡改的分布式账本技术,为解决医疗场景中的隐私保护问题提供了新的思路。首先列举了医疗数据的隐私保护需求,并介绍了区块链的整体架构。然后将医疗信息隐私保护分为面向数据的隐私保护和面向用户的隐私保护,并详细介绍了面向数据及面向用户的区块链隐私保护方法。面向数据的隐私保护即对敏感信息本身的保护,主要采用基于加密的隐私保护方法、基于失真的隐私保护方法和基于限制发布的隐私保护方法;面向用户的隐私保护即对数据使用者的隐私保护,包括基于访问控制的隐私保护和基于交易匿名的隐私保护。最后对比总结了各类方法的特点以及区块链在隐私保护领域的研究发展现状,并展望了区块链在医疗信息隐私保护领域的发展方向。

关键词: 区块链; 医疗信息; 数据隐私; 用户隐私; 去中心化

中图分类号: TU528.1 **文献标志码:** A **文章编号:** 1671-6841(2021)02-0001-18

DOI: 10.13705/j.issn.1671-6841.2020324

0 引言

随着信息技术的快速发展,医疗服务逐渐向数字化、信息化转型。医疗数据作为重要数据资产也在不断被分析和挖掘,极大地推动了医疗领域的研究与进展。不同组织机构之间的信息交换与共享能够使得医疗数据发挥更高价值,因此,跨地区、跨机构的医疗信息共享需求日益增长。

医疗数据的来源和范围多样化,具有特殊的敏感性及重要性,不仅承载数据主体的健康状况及医疗处理过程等信息,还涉及个人的隐私保护、行业的发展,甚至关系国家安全。然而,随着数据爆炸式增长及深度挖掘的应用,医疗数据的泄露风险不断增加[1-2]。医疗数据泄露方式主要分为非交互式泄露和交互式泄露,非交互式泄露是指医院内部系统或人员的泄露,如私自使用权限倒卖或滥用信息;交互式泄露即医疗信息在发布以及在不同机构间共享时的泄露[3]。由于医疗数据的价值,在中心化信息系统中,中心节点存在泄露隐私的动机,各类医疗数据在多方交互及共享过程中也容易遭遇攻击,从而导致隐私泄露。目前,医疗数据存储依赖于医院的信息化部门,许多医院信息系统的安全防御能力欠缺,技术措施不足,一旦医院系统遭受攻击将会造成隐私信息泄露。云计算平台能够为医疗数据用户提供强大的计算能力,但直接共享数据的服务模式带来了一系列隐私安全问题及数据所有权问题。医疗数据可通过统计分析、数据挖掘、深度学习等方法从海量数据中剔除无意义部分,筛选出有价值的信息,经过处理(如属性匹配、信息关联)的数据存在泄露敏感信息的可能。

区块链作为一种公开的去中心化分布式账本,具有多方维护、不可篡改等特性[4],有利于解决医疗数据共享过程中的隐私问题,打通数据孤岛,提供医疗数据安全共享及交易平台,明确数据所有权,有效地防止数

收稿日期: 2020-10-12

基金项目: 国家重点研发计划项目(2018YFB1201403);河南省高校科技创新人才支持计划项目(21HASTIT031);河南省高等学校青年骨干教师培养计划项目(2019GGJS018);河南省高等学校重点科研项目(20A520035);郑州市协同创新重大专项(20XTZX06013);赛尔网络下一代互联网技术创新资助项目(NGII20190707)。

作者简介: 刘炜(1981—),男,副教授,主要从事区块链、智慧医疗等研究,E-mail:wliu@zzu.edu.cn;通信作者:田钊(1985—),男,讲师,主要从事信息安全、人工智能、智能交通等研究,E-mail:tianzhao@zzu.edu.cn。

据被恶意篡改或第三方滥用及倒卖现象。区块链技术的发展和应用,在给医疗信息共享带来便利的同时也带来了新的安全隐私问题。相对于中心化架构,区块链能够避免因单点失效或数据泄露而引发的安全风险,从而保证数据的完整性及不可篡改性。但在区块链系统中,交易记录的透明性将显著增加隐私泄露的风险,例如分析交易记录可获得用户的交易规律[5]。在传统医疗信息系统中,隐私保护的重点是确保中心信息系统的安全,然而在区块链中没有统一管理者,采用的信息传递机制和共识机制也为隐私保护带来了新的机遇和挑战。本文立足于医疗信息领域,对区块链及相关隐私保护技术进行综述。

1 基于区块链的隐私保护背景知识

1.1 医疗信息隐私保护需求

随着医疗信息化建设的不断发展,医疗信息的共享与整合为患者提供了更好的服务质量,有助于医学研究的发展。医疗数据来源众多,数据类型复杂,存储于异构的信息系统中,数据共享可实现不同医疗信息系统之间的数据交互。在医疗服务与医学研究信息化的同时,医疗信息隐私问题也随之而来。由于医疗信息的高价值与隐私性,在信息共享过程中,应结合隐私保护手段,在明确数据所有权、不泄露敏感信息的前提下发挥数据价值。要实现医疗数据的安全共享需满足以下需求。

1) 隐私性需求。医疗数据共享很大程度上涉及数据及相关方的隐私。隐私数据如果非法交易或使用,将带来不可估量的后果,因此数据共享必须限制共享范围,防止隐私数据泄露给无权限方。

2) 完整性需求。医疗数据使用方需要确保共享方数据的可靠性,包括数据来源的可靠性、是否被篡改、是否伪造等。数据不完整或篡改将影响数据共享的效率,而且可能给数据使用方带来严重的问题。

3) 可用性需求。可用性是指数据可在任何时间被任何有权限的用户访问和使用[6]。缺乏可用性医疗数据共享,提供方将无法提供数据服务,使用方将无法得到所需的数据,因此保证医疗共享过程中的数据可用性,才能保证数据共享的效益,保证医疗数据发挥价值。

1.2 区块链技术

区块链是一种以比特币为代表的去中心化共享账本,由数字加密货币衍生而来的新型技术架构[7]。区块链按照时间顺序将数据区块组成链式数据结构,并以密码学技术保证链上数据的不可篡改和不可伪造。每个节点都通过 Hash 算法和 Merkle 树,将一段时间内接收到的交易封装到一个带有时间戳的区块中,并将其链接到最长的主链中,形成最新区块[4]。如图 1 所示,每个区块包括区块头和区块体两部分,区块头中封装了前一个区块的 Hash 值、时间戳、随机数、Merkle 根等;区块体存储交易信息,即区块链记录的数据信息,全部交易基于 Merkle 树的 Hash 过程,生成唯一的 Merkle 树根存储在区块头中,每笔交易均由交易方进行数字签名,并永久存储在区块中,供全体用户查询[8]。

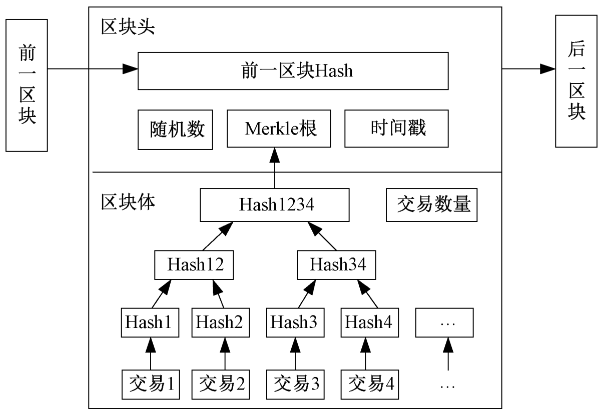

图 1 区块结构

Figure 1 The structure of blocks

尽管区块链系统在整体架构上存在诸多共性,但同时也在不断演变[9]。一般来说,区块链由数据层、网络层、共识层、激励层、合约层和应用层组成,但随着区块链的发展与演变,一些传统的模块被弱化。因此,通过分析区块链的本质特征,结合目前的发展趋势,本文将区块链系统分为 4 层,如图 2 所示。

1) 网络层。网络层封装了区块链网络的组网方式、节点之间的消息传播机制和数据验证机制等,其主要任务是通过传播协议和验证机制使得区块链网络中每个节点都能参与区块的产生与校验过程,仅当验证后的区块才能记入区块链系统[4]。区块链采用对等式网络(peer-to-peer network, P2P)来组织分布式、关系平等、可动态进出的节点。交易节点生成新的交易后,将该交易广播到区块链网络,在 P2P 网络中,每个节点时刻监听网络中广播的数据,当邻居节点接收到交易节点产生的新交易时,首先验证其是否有效,若有效则按照时间顺序记入存储池,同时再转发给自己的邻居节点,由此类推,该交易会逐渐广播到全网;若交易无

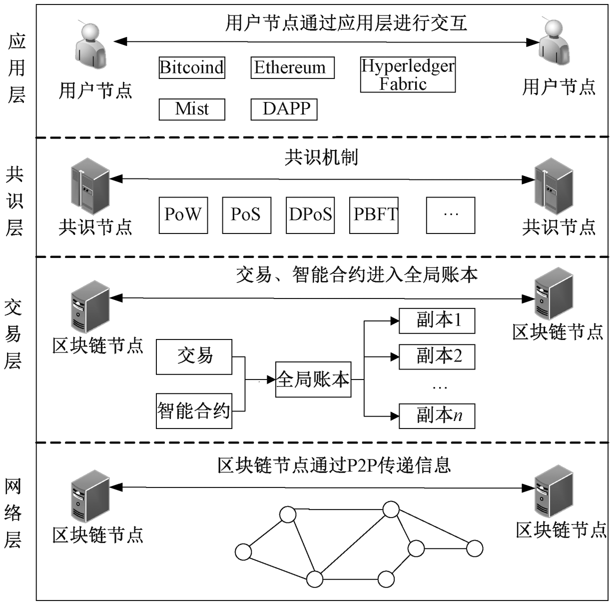

图 2　区块链体系架构

Figure 2　Framework of blockchain technology

效则被废弃，以防止其在网络中继续传播。在此过程中，接收方节点从网络中收集信息，无须直接与发送方进行通信，相比于传统网络，攻击者很难通过监听网络流量来发现信息的来源和去向，因此，可防止通过此类攻击来发现用户之间的通信关系。

2）交易层。交易层实现两个区块链地址之间的数据传输，主要内容包括地址格式、交易格式、全局账本和智能合约。区块链地址是用户在区块链系统中的假名，通常由公钥加密算法产生，公钥用于交易的输入或输出地址，私钥由用户保存，用于对交易进行签名[5]。交易记录了地址之间的数据交互记录，主要有基于交易的模型和基于账户的模型，其中：比特币采用基于交易的模型；以太坊及 Hyperledger Fabric 采用基于账户的模型[10]。基于账户的模型在可编程性、灵活性等方面更有优势，基于交易的模型中计算负担由钱包来承担，一定程度上减少了链上压力。对于选择何种模型，要从具体的业务场景出发。全局账本存储所有交易信息、合约及相关参数，通常由区块构成，每个区块中包含一定的交易信息。智能合约是一种能够实现自我执行和验证的计算机协议，其本质上是将对象程序化并部署在区块链上，再由外部事件来触发合约的自动生成与执行，进而改变对象的状态与数值，同交易类似，智能合约内容也将写入全局账本。

3）共识层。共识机制是区块链系统的核心要素，共识安全对区块链的数据安全起到重要的支撑作用[6]。去中心化的区块链网络由多方进行维护，保证系统中所有节点高效达成一致，维护相同的全局账本是区块链共识机制所要解决的核心问题，常见的共识机制包括工作量证明机制（proof of work，PoW）、权益证明机制（proof of stake，PoS）、股份授权证明机制（delegated proof of stake，DPoS）以及实用拜占庭容错共识算法（practical byzantine fault tolerance，PBFT）等[11]。

比特币区块链采用 PoW 作为共识机制，其思想是通过算力竞争来保证数据一致性。各节点（矿工）同时竞争挖矿，挖矿过程中付出最大算力的节点将被选为记账节点并由该节点生成下一区块。PoS 采用权益证明来代替算力，使用币龄（一定数量的币与最后一次交易时间长度的乘积）来选择记账节点，解决了 PoW 机制的资源浪费问题。DPoS 是基于 PoS 衍生的解决方案，每个节点能够自主选择其信任的代理节点来轮流记账生成新区块，因而大量减少验证记账节点，使交易效率更高。由于硬件错误、网络阻塞、恶意攻击等原因，分布式网络中存在不可信节点，因而需要支持拜占庭容错（byzantine fault-tolerant，BTF）。PBFT 是基于 BTF 的共识算法，解决了共识算法容错率不高的问题，并且将算法复杂度由指数级降低到多项式级，使得拜占庭容错在实际系统中得到应用。除了以上共识机制，实际应用中衍生出多种组合共识机制以及现有共识机制的变种。这些共识机制从效率、安全性等角度进行改进，各有优劣，适用于不同的业务场景。

4）应用层。应用层封装了各种应用场景和案例，如医疗数据共享平台等，为用户提供各种应用场景的程序和接口，以实现各方的交互。本层类似于各种软件程序，也是去中心化应用（decentralized application， DAPP),包括智能合约以及调用合约的接口。从当前的区块链应用发展来看,应用层在兼顾隐私的同时面临监管缺失问题。监管技术的目的是对非法行为进行检测、追踪和追责,从而保证区块链平台的内容安全。然而,区块链去中心化、不可篡改、匿名等特点增加了监管机制设置的难度[6]。

1.3 区块链隐私保护

数据隐私保护的核心是保护数据隐私性与完整性,隐私性指防止数据被未授权用户访问,完整性是指保证数据真实、未被篡改[8]。由于区块链的存储容量有限,将区块链技术用于数据隐私保护,通常使用存储与管理分离的方式,大量的原始数据存储于链下服务器,并由这些服务器保证数据的隐私性,数据索引及权限由区块链进行管理,由区块链的公开账本保证数据完整性。访问控制技术用于实现用户权限的管理,由于区块链数据对所有用户可见,并且不可篡改,相比于基于单一节点的授权策略,基于区块链对用户权限进行管理,可实现公开透明的授权和访问控制,不存在第三方的越权行为。区块链的不可篡改特性可用于对数据的流向进行全程监控,实现对数据存储、传输、计算等过程中的可追溯、可审计。然而,区块链全局账本的公开透明会给链上数据隐私带来一定的不利因素。区块链中的隐私分为身份隐私和交易隐私[12]。

身份隐私是指用户身份信息与区块链地址的关联。区块链中的地址由用户公钥自行产生,并与用户身份信息无关联,可作为用户在区块链系统中的假名。然而,区块链的匿名性并不能保证绝对隐私性,在用户使用区块链地址参与交易时,有可能泄露用户身份信息。交易隐私是指区块链账本中存储的交易数据与交易数据背后的知识。在比特币系统中,交易是公开记录在链上的,没有采取隐私保护措施。但随着区块链技术在医疗领域的应用,交易记录中通常包含敏感信息。此外,随着数据分析及挖掘技术在医疗领域的应用,医疗交易记录背后通常能够反映一系列知识。身份隐私和交易隐私是区块链隐私保护的重点内容,一旦用户身份与区块链地址之间的映射关系泄露,就可能对用户隐私造成严重危害,不同于传统的中心化系统,在区块链中无法通过删除存储在全局账本中的数据来限制泄露隐私信息的传播,即使采用硬分叉手段,形成一条不同的新链,也无法挽回数据泄露产生的后果。因此,在区块链中更应注重隐私信息的保护。

医疗信息隐私保护包括数据在发布、存储、交换、分析等过程中的保护。本文将医疗信息隐私保护方法分为面向数据和面向用户的隐私保护。面向数据的隐私保护是指对医疗数据中敏感信息本身的保护[5];面向用户的隐私保护是指用户的数据访问授权以及在数据使用、交易等过程中,对用户信息的隐私保护。

2 面向数据的隐私保护

面向医疗数据的隐私保护即对敏感信息本身进行保护,主要分为3类:基于数据加密的隐私保护;基于数据失真的隐私保护以及基于数据限制发布的隐私保护。

2.1 基于数据加密的隐私保护

区块链中使用基于哈希算法的交易存储机制以及数字签名保证链上数据的真实、不可抵赖性,然而在复杂的医疗业务场景中,交易数据以及智能合约的隐私性需要结合其他的密码学技术,如同态加密、安全多方计算等。此外,在基于区块链的医疗数据共享平台中,区块链地址一旦产生,就只能通过配套的私钥对该地址的数字资产进行转移和支付,因此拥有私钥即拥有该地址数字资产的掌控权,钱包的安全性及密钥存储方式对区块链系统中数据隐私安全带来重要影响。

2.1.1 同态加密 在医疗大数据时代,对数据进行挖掘、学习、分析等操作可为患者、医院及相关机构带来更多价值,如深入了解疾病、推动实现个性化服务等。而在数据的计算、处理过程中,保护数据隐私是医疗数据共享的重要前提。同态加密(homomorphic encryption, HE)由 Rivest 等在20世纪70年代首次提出,是一类具有特殊属性的加密算法。与一般的加密算法相比,同态加密除了基本的加密操作,还能实现密文之间的多种计算操作,即先计算后解密可等价于先解密后计算,该特性对于信息安全具有重要意义。

在医疗场景中,不具备计算资源的数据所有者可委托第三方对数据进行计算,并将计算结果返回,然而不可信第三方存在泄露隐私的风险。同态加密可实现第三方在不解密数据的情况下对医疗数据进行计算与验证,如图3所示,第三方(如用户、云服务方等)或其他计算框架对加密的隐私数据进行相关操作,而不影响其保密性。在文献[13]中,改进的同态加密算法能够保证 MapReduce 框架中的计算安全,使得用户在自主控制数据的同时保证数据计算和共享的安全性。

区块链交易和智能合约的透明性容易泄露数据隐私，同态加密可保证交易数据的隐私验证及智能合约的计算安全，避免在数据验证或操作过程中泄露隐私。文献[14]提出了一种基于同态加密的区块链隐私保护方法，验证节点对加密的用户交易信息进行操作和验证，提高了区块链技术的数据安全性和隐私性。针对医疗保险理赔过程中患者明文数据需被保险公司查看的问题，文献[15]将同态加密与区块链智能合约相结合解决医疗保险理赔过程中的隐私泄露问题。智能合约对密文数据进行操作，无须查看患者明文数据即可判断是否符合理赔条件，并实现自动理赔功能，避免患者在与其他角色交互时泄露敏感数据。

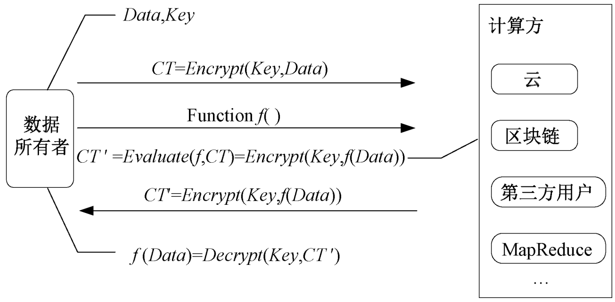

图 3　同态加密过程

Figure 3　Homomorphic encryption process

云服务可为医疗数据提供强大的计算能力与存储空间，区块链可将医疗数据的相关操作记录保存到公开账本，以便用户随时查看，并保证数据不被篡改。区块链与云技术的融合，既有利于解决区块链的存储与性能瓶颈，也能够有效防止医疗数据被随意伪造篡改。然而云服务及区块链在实际场景中均面临如何保证数据隐私的问题，同态加密可以在一定程度上解决数据在云端计算过程中的隐私问题，数据所有者可以将加密数据委托给不受信任的云服务方对数据进行处理而不泄露隐私信息。在区块链中，同态加密技术不仅提供了隐私保护，同时允许随时访问链上的加密数据进行审计或其他操作。使用同态加密在区块链上存储数据将能够保留公有链和私有链的优点，同时完整地保留以太坊的优点。

2.1.2　安全多方计算　医疗数据的挖掘分析可使数据产生增值价值，尤其是相关数据的统计分析及反映集体特征的汇总信息具有更高的价值。然而由于医疗数据存储的碎片化，单个医疗节点的私有数据存在数据量不足，数据维数不够丰富等问题，医疗数据的隐私性及高价值导致数据所有方不愿共享原始数据，影响数据产生增值价值。因此，在多方联合计算过程中保护各相关方的数据隐私对医疗数据共享具有重要意义。

安全多方计算（secure multi-party computation, SMPC）是一种通用的密码原语，可表示为

$$f(x_1, x_2, \cdots, x_n) = (y_1, y_2, \cdots, y_n)。\tag{1}$$

使互不信任的分布式环境下多个数据所有者联合计算同一函数，而不泄露自己的输入数据 x_i，计算结束后，每个参与方除计算结果及自己的输入值，无法了解其他的任何信息。SMPC 在保证参与方输入隐私的前提下完成协同计算任务，其框架涉及混淆电路、秘密共享、同态加密等密码学技术[16]。SMPC 技术的主要特点如下。

1）隐私性。任何参与方无法获得其他参与方的输入信息；

2）正确性。每个参与方均能获得正确计算结果；

3）终止性。保证在有限时间内有输出；

4）忠诚性。大部分参与者按照规定执行计算。

区块链和 SMPC 均为参与方按照特定规则（协议）进行交互，区块链强调计算结果的可验证性，并防止结果被篡改，而 SMPC 强调的是计算过程中输入数据的隐私性，并不能确保数据的可验证性，其目的是在输入保密的情况下得到正确的计算结果。区块链能够为 SMPC 做存证，SMPC 也能够应用于区块链智能合约，在保证各方隐私的同时对多方数据进行联合计算。针对 SMPC 中参与方的诚实问题，可通过区块链来公开参与方的信誉，并通过激励机制鼓励诚实的参与方。BFR-MPC 是一种基于区块链的多方计算方案，区块链公共账本为各方维护了一个公开的信誉系统，并通过激励机制鼓励各方执行既定的计算协议，更好地保证了 SMPC 的公平性和鲁棒性[17]。文献[18]基于区块链智能合约构建了惩罚机制，提出了安全的 SMPC 协议，协议通过超时机制来判断恶意参与方的提前终止行为，并对其进行经济惩罚。文献[19]构建了基于 SMPC 的智能合约框架，保证了智能合约执行中的输入隐私和计算正确性。

云计算的引入使得 SMPC 执行的外部环境变得多样和复杂，云服务器可看作不被完全信任的第三方，参与方将各自输入利用同态加密后，上传至云服务器，由云服务器对同态密文进行计算并返回结果，从而可以保证数据机密性[20]。同态加密可作为 SMPC 框架中的基础协议，如文献[21]和[22]将 SMPC 与同态加密

相结合,为各参与方在SMPC协议中的输入数据进行保密。

SMPC能够为医疗数据提供一种隐私计算方案,旨在保证多方输入数据不被泄露给其他参与方;区块链技术旨在记录医疗数据流通日志,为隐私计算提供可验证性及可追溯性,激励参与方提供真实数据并诚实履行既定计算协议。区块链与SMPC的结合,可实现增值医疗数据的安全共享,并确保数据交易的可验证、可追溯,有利于对各参与方的监管及相关责任方的定位。

2.1.3 密钥管理 区块链结合加密技术可实现安全的医疗数据共享,由于现代密码体制要求加密算法公开,密码系统的安全性并不取决于算法的保密性而取决于密钥的保密性。因此,对于加密的医疗数据,合理的保存私钥是保证数据隐私的关键。在区块链系统中,私钥即资产,钱包即存储私钥的载体,掌握私钥即拥有一个区块链地址背后的数据资产,因此私钥安全对数据隐私尤为关键。按照是否联网访问,私钥的存储方式可以分为冷钱包和热钱包;按照是否由第三方代理钱包可分为托管类型钱包和非托管类型钱包。

冷钱包由于不能被网络访问,避免了黑客盗取私钥的风险,比热钱包会更安全;热钱包在联网状态下可以随时进行交易,对于频繁交易的投资者来说,热钱包会更方便快捷。托管钱包是指第三方代表用户控制加密货币的钱包,该钱包不提供用户私钥或将私钥存储在本地服务器中。如果用户丢失私钥,可以在服务器端恢复账户,因此更倾向于便捷使用钱包的用户。然而服务提供商由于系统维护或实名制(know your customer, KYC)等原因冻结用户资产时,托管类型钱包具有遭遇资金被黑客盗走的风险。此外,用户的资产实际被交易所掌控,交易所可能会出于实名制或者反洗钱(anti money laundering, AML)等原因冻结用户资产。非托管钱包是指用户完全控制并持有自己的加密货币的钱包,私钥会被加密存储在用户的设备中。在非托管类型钱包中,如果用户丢失助记词,将无法恢复资产。文献[23]提出了一个基于区块链技术的分布式公钥方案,通过区块链网络中的节点共同承担密钥存储的职责。相比传统的公钥系统,将现实的存储系统拆分到区块链分布式节点的方案具有较好的响应性能和抗干扰能力。

为解决分布式网络密钥管理困难、通讯开销过大等问题,文献[24]提出了一种基于区块链的密钥管理方案(key management schemes based on blockchain, KMSBoB),设计了分布式群组网络下密钥管理和传输过程,其核心是分布式网络环境下基于区块链的密钥管理策略,将MTI/CO协议与区块链多节点挖矿相结合,形成一种分布式环境下多节点会话密钥生成协议。KMSBoB把传统的密钥管理模型转变为基于区块链的运行模式,将验证密钥分量合法性、密钥分量传输和密钥分量共享的3个过程统一于全体安全管理节点(secure management node, SMN)的挖矿过程,通过全体SMN成员验证区块数据合法性的过程,实现密钥数据的动态安全管理,降低了传统自治域之间因密钥传输的通信开销和密钥泄露的安全风险。

2.2 基于数据失真的隐私保护

医疗数据集中通常包含着敏感信息,如患者身份、诊断结果等,若不采取隐私保护技术而直接将数据进行分析或发布,会造成隐私的泄露。如何从医疗数据集中提取有价值的信息是医疗隐私保护的关键问题,研究人员试图找到能够抵御最大背景知识攻击的隐私保护方法,差分隐私在一定程度上满足了上述要求。差分隐私是Dwork在2006年提出隐私定义,是一种独立于数据集的强隐私概念,能够防止攻击者在拥有任何背景知识下的攻击。设有随机算法M,P_M为M所有可能的输出构成的集合。对于任意两个邻近数据集D和D',以及P_M的任何子集S_M,若算法M满足式

$$Pr[M(D)\in S_M]\leqslant \exp(\varepsilon)*Pr[M(D')\in S_M],\tag{2}$$

则称算法M提供参数ε的差分隐私保护。其中ε为隐私保护预算,用来控制差分隐私算法在两个相邻数据集上获得相同输出的概率比值,体现算法所能提供的隐私保护水平[25]。差分隐私的基本思想是对原始数据的查询添加噪声,从而使得单独数据的插入或删除不会对任何计算的输出造成影响,攻击者因此难以根据多次查询结果的差异性反推出数据集中某条数据。

差分隐私为保证隐私信息在数据发布共享中不被披露提供了思路。TDPS(transaction data publish strategy)是一种有效地满足差分隐私约束的事务数据发布策略[26]。首先构建事务数据库的完整Trie树T_D^l,然后基于压缩感知技术对T_D^l添加满足差分隐私的噪声得到含噪Trie树NT_D^l,最后在NT_D^l上进行频繁项集挖掘任务。对TDPS-LP-Singal、TDPS-LP-Result、TDPS-EP三种加噪机制的效用性进行评估,证明该策略在保护个人隐私的同时提供较高的数据效用性。文献[27]提出一种基于区块链的去中心化数据共享模型,数据需求方根据需求编写用于隐私计算的智能合约,只有满足差分隐私的计算结果才能作为智能合约的输出并共

享给数据请求方,保证了智能合约的计算隐私。

基于差分隐私的数据发布策略需要在保护隐私的同时兼顾数据的效用性及隐私预算 ε 的分配策略。文献[28]结合差分隐私约束和数据效用性优化构建了分布式非线性规划模型,设计了两种解决方案安全地求解该模型,即全局解决方案(GS)和局部解决方案(LS),并在实验中基于不同参数、执行时间、计算复杂度评价该差分隐私发布策略。理论分析与实验结果均表明,该方案满足差分隐私要求且具有很好的实用性。合理的预算分配策略要尽可能使 ε 的生命周期长一些,常用的分配策略包括线性分配、均匀分配、指数分配、自适用性分配以及混合策略分配等[29]。文献[30]通过区块链来验证隐私预算 ε 的使用,并根据数据所有者提供的隐私要求通过智能合约自适应地更改其分配。

差分隐私是目前信息安全研究领域中的热点之一,其严格的数学定义与实现机制为医疗信息隐私保护提供了一个有效并可靠的解决方案,但从实际应用上看,将差分隐私保护投入医疗领域实际应用中还需要更深入的研究。医疗实际应用中存在许多复杂数据集,数据之间存在关联性,而传统的差分隐私方法并未考虑数据之间的联系,医疗图像数据以及动态医疗数据的发布也是差分隐私保护方法有待解决的问题。

2.3　基于数据限制发布的隐私保护

在区块链技术中存在“不可能三角”,所谓的“不可能三角”是指在区块链公有链中,很难同时做到去中心化、安全性及很高的交易处理性能。区块链的链上链下系统都有一定的局限性,链下系统具有可扩展计算及存储性能,但难以验证数据确保数据不被篡改;区块链链上系统可保证信息的公开透明及不可篡改性,但其扩展性及吞吐量一直是瓶颈。医疗数据通常信息量大、类型复杂,对隐私保护的需求不同,因此不宜将大量原始数据或对隐私要求高的数据直接存储在公开账本中。在实际医疗场景中,可将链下信息系统与区块链相结合,一方面利用区块链保证链下原始医疗数据的完整性,另一方面可通过链下系统实现海量数据的存储与计算。此外,结合链下系统,通过限制链上信息的发布,可直接将敏感医疗信息从公开账本中移除,从而保护隐私信息的安全。

链外存储是将原始数据存储在链下数据库,对应的摘要哈希值或索引信息存储在分布式账本中。哈希算法可将任意长度的字符串映射成固定长度的二进制值串,通过原始数据映射之后得到的二进制值串即哈希值。文献[13]的数据共享方案中,由于原始数据的数据量较大,且可能包含隐私信息,不适合存储在链上,因此采用链上索引表信息与链下数据库相结合的方式进行存储,一方面能够释放区块链上的大量存储空间,另一方面亦可提高信息共享的效率,以此保证数据的安全性。对于数据量较大的医疗信息,不适合在共识过程中传输,因此,现有的医疗区块链系统大多采用此类存储方式,如文献[31]提出的基于区块链的医疗数据分享模型(medical data share model, MDSM)。链外存储提供了较强的隐私性,但是由于原始数据存储在链下,仍然需要各方自行维护链下数据库安全。此外,随着量子计算的发展,哈希算法面临安全性降低甚至被攻破的危险,因此从长远来看,势必要发展适用于区块链的抗量子攻击的密码技术[10]。

账本隔离是指将不同隐私需求的账本分别放在不同的分布式账本上,如图 4 所示。文献[32]提出的 Fabric 支持多通道,在增强隐私性方面做了很大改进,实现账本隔离保护隐私性。共识服务由排序节点 Orderers 提供,账本由 Peer 节点管理,Channel 代表私有广播通道,保证消息的隔离性和私密性,不同的智能合约关联主体仅执行并验证相关交易,Peers 可订阅多个 Channel,并且只能访问订阅通道上的交易。共识服务接收所有交易,虽然外部节点无法看到通道内交易数据,但 Orderers 可以看到所有通道的数据,为防止 Orderers 节点泄露交易的内容,必须利用其他技术来隐藏敏感数据,例如哈希散列或同态加密。然而,使用多通道来保护数据安全也存在不足,为了保护任意多方之间的隐私,需要构建 $O(n^2)$ 量级的通道,从而消耗大量资源,此外多通道机制无法满足动态的数据访问控制。

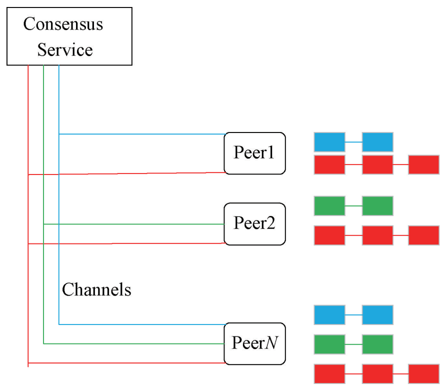

图 4　账本隔离

Figure 4　Ledger isolation

公有链是指任何节点都可以随时进入系统中读取数据、发送可确认交易、竞争记账的区块链,典型的公有链如比特币、以太坊。基于公有链的系统具有很高的公信力,但隐私性及交易性能较低。私有链的写入权

限是由某个组织或机构控制,其参与节点是有限和可控的,因此私有链往往有更快的交易速度、更好的隐私保护、更低的交易成本。联盟链是指由若干个机构共同参与管理的区块链,每个机构都运行着一个或多个节点,其中的数据只允许系统内不同的机构进行读写和发送交易,并且共同来记录交易数据。与私有链相比,联盟链中的权限设计要求会更加复杂。私有链、联盟链架构从根本上关闭了非授权节点接触数据的渠道,降低了隐私泄露的风险[5],但与公有链相比,这两种架构是不具备或是不完全中心化的。

基于链外存储的区块链医疗信息共享平台不仅能够为存储在各医疗机构服务器中的数据提供完整性保障,又能够将数据量大、对隐私需求高的医疗信息从公开账本中移除,既提高了区块链的交易性能,又保证了数据的隐私。与公有链相比,联盟链和私有链限制了区块链节点的加入,具有更强的隐私性,结合医疗数据共享的特点,可将不同类型的区块链相结合,通过“以链制链”的方式,实现隐私保护与监管机制的平衡。

在面向数据的隐私保护中,区块链结合基于数据加密、数据失真及限制发布的隐私保护方法实现了对敏感数据本身的保护,表1介绍了相关方法的特点,并分析了各种隐私保护技术的优势与不足。

表1 面向数据的隐私保护机制对比

Table 1 Comparison of data privacy protection mechanisms

类别	方法	方法描述	优势	挑战及不足
数据加密	同态加密	在不解密数据的情况下可对密文数据进行计算	安全性高; 可对隐私数据执行计算操作	效率低,计算开销大; 对计算结果难以验证
	安全多方计算	在不信任网络环境中,多用户能够在不泄露各自私有输入的同时合作执行某项计算任务		
数据失真	差分隐私	在数据中加入噪声提供隐私保护	有效防御背景知识攻击; 建立在严格的数学基础上	复杂、连续数据发布难点; 分布式差分计算难点
限制发布	链外存储	完整数据存储链外,只在公开账本上存储部分信息	隐私性强;节省链上计算、存储开销	原数据需各方自行维护
	账本隔离	将不同隐私需求的账本,放在不同分布式账本上		对区块链结构的改变可能引入新的安全问题
	联盟、私有链	区块链中加入节点准入机制	根本上防止无权限用户的访问	去中心化程度降低

基于数据加密的隐私保护方法通常具有较高的安全性,同态加密、SMPC技术可实现数据的隐私计算,结合区块链能够存储隐私计算过程中的相关凭证,为验证计算结果正确性提供可信日志,加密效率及计算开销问题是本类方法需要重点考虑的问题。差分隐私能够有效防御基于知识背景的攻击,然而对于复杂、连续的医疗数据,以及分布式差分计算需要进一步深入研究。基于限制发布的隐私方案能够在保证数据安全的同时减轻链上开销,但是对区块链底层的改进可能带来新的安全隐患。此外,由于原始数据需各方维护,链上链下的对接安全也需深入研究。

3 面向用户的隐私保护

面向用户的隐私保护是指用户对数据的访问控制权限及其在区块链系统中的匿名性。访问控制策略是保护数据隐私的重要方式之一,其作用是限制数据的共享范围,保证信息不被非法获取。虽然用户在区块链系统中的身份与现实世界中的身份无关联,然而通过分析、聚类等技术处理链上交易记录能够获得不同账户间的交易关系图谱,推测交易的输入输出关系,对交易者的隐私造成威胁。由于区块链的公开透明,当医疗隐私信息保存在区块链上时,匿名性就成为了首要重视问题。面向用户的隐私保护方法主要基于访问控制和交易匿名。访问控制包括基于智能合约的访问控制策略和基于属性加密的访问控制策略;交易匿名方法包括混合协议、零知识证明、数字签名、安全通道协议及K匿名技术。

3.1 基于访问控制的隐私保护

访问控制技术用于对用户权限进行管理,允许合法用户依照其所拥有的权限访问系统内的相应资源,禁止非法用户对系统的访问,从而保证信息的安全[7]。EMR为医疗机构记录和保存患者就诊记录提供了极大的便利,而当前的EMR访问控制机制存在授权效率低、灵活性差、权力中心化等问题,基于区块链智能合约

及基于属性加密的访问控制机制对解决上述问题提供了应对方案。

3.1.1　智能合约　基于第三方的访问控制策略由可信中心实体进行构建，然而，在实际医疗应用场景中不存在绝对安全的第三方实体，中心化 EMR 访问控制机制依赖主节点的稳定性，主节点作恶或发生故障会严重威胁医疗机构及患者的隐私和利益。区块链技术具有分布式去中心化不可篡改的优势，因此将区块链技术应用在医疗数据访问控制中可以很好地解决问题，目前国内外研究中基于区块链智能合约可实现对数据的访问控制，如 MedRec[33]、Ancile[34]、MeDShare[35]等医疗区块链平台。智能合约的本质是一种能够实现自我执行和验证的计算机协议，提供了一种先进的思想，使得在区块链网络上实施更多的访问控制模型。MedRec 通过以太坊平台对电子医疗数据进行访问控制，医疗组织在获得病人授权后方可使用医疗数据，使得医疗数据真正属于患者所有，实现了对病人数据的隐私保护及不同医疗机构之间的数据整合[33]。文献[34]使用以太坊智能合约实现医疗区块链系统中不同角色之间的交互，适应不同用户的需求。

区块链能够为数据进行审计跟踪，对敏感信息的访问保留行为日志，对恶意行为进行惩罚。MeDShare 采用智能合约和访问控制机制有效地追踪数据的行为，为区块链中的实体提供数据来源、审计和控制，对恶意实体取消访问控制权限[35]。文献[36]针对云服务提供商的数据安全共享，提出一种基于区块链的数据共享框架，利用区块链实现云中敏感数据的访问控制，并保留访问节点的行为日志，保证后续的责任认证。医疗机构联盟服务器群(medical institution federate servers, MIFS)配合改进的股份授权机制，通过信用积分策略对诚实节点增加信用积分，对恶意节点进行惩罚并限制其访问数据的权限，实现数据的访问控制[31]。

3.1.2　属性加密　属性加密(attribute-based encryption, ABE)是解决医疗数据隐私和细粒度访问控制问题的重要技术。基于区块链的 EMR 共享方案(blockchain based privacy preserving data sharing for EMPS, BPDS)采用基于密文策略属性基加密(ciphertext policy attribute based encryption, CP-ABE)的访问控制机制让患者预先设置 EMR 的访问结构，并利用访问结构对 EMR 进行加密，使得患者可以完全控制自己的电子病历，当且仅当访问者的属性能够满足访问结构时才能解密 EMR[37]。文献[38]提出的区块链数据共享模型基于 ABE 对企业数据进行访问控制与共享，实现了更安全的细粒度访问控制。一个数据所有者将一份明文数据加密发送给 N 个不同共享方，若使用传统公钥加密算法，如图 5(a)所示，数据所有者需要使用 N 个不同公钥将明文加密 N 次，形成 N 份不同密文，分别发送给 N 个共享方。若使用 ABE，如图 5(b)所示，加密时只需要根据成员属性生成用于加密的公钥 PK，而不需要关心群体中成员的数量和身份，降低了数据加密开销，也保护了成员隐私，能够实现更加灵活的访问控制策略。

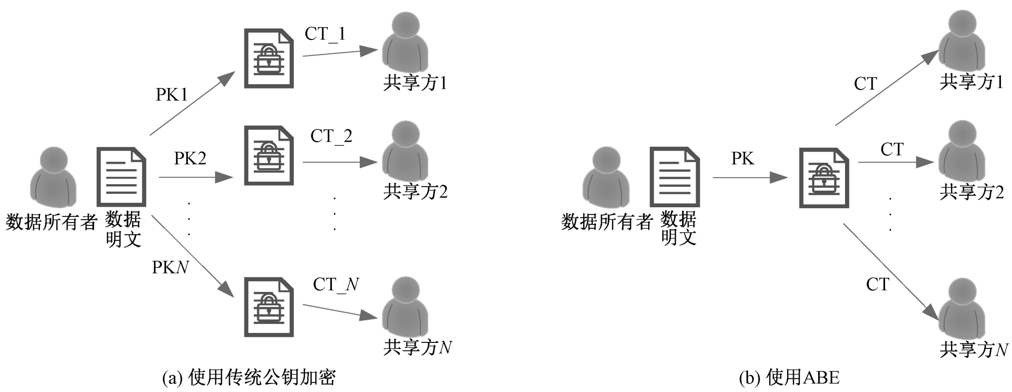

图 5　ABE 算法与传统非对称加密的对比

Figure 5　Comparison between ABE and traditional encryption

在中心化系统的 ABE 方案中，单点故障或第三方不诚实可能会使系统的安全受到威胁。文献[39]提出 Etherenm 与 ABE 相结合的框架，解决集中式云存储系统中密钥滥用、隐私数据泄露等问题。文献[40]提出了一种基于 Ethereum 区块链技术的访问控制安全云存储框架，数据所有者可以通过链上智能合约设置数据的有效访问期限，并对访问控制实现跟踪。DAHB 框架(distributed architecture based on hierarchical blockchain for internet of things)以基于属性下访问控制(attribute-based access control, ABAC)模型为基础，采用智能合约实现对物联网设备基于属性的域内和跨域的灵活、动态、自动化的访问控制[41]。智能合约实现 ABAC 模型中策略执行点、属性权威、策略决策点部分，在属性中引入信任度量，利用信任值与诚实度反映实

体能够履行合约的信用能力和稳定性,作为授权决策的依据。文献[42]针对大数据资源来源广、动态性强的特点,提出基于ABAC的区块链大数据访问控制机制,保证访问控制信息的透明公开、可审计、可验证及不可篡改,实现数据的自动化灵活访问。

访问控制是EMR常用的隐私保护策略,基于角色的访问控制(role-based access control, RBAC)着重于有权访问信息的策略,被授权的用户可以访问全部数据,因此很难控制数据原始区域的隐私性,然而在实际应用中,用户仅需要获得部分数据即可。文献[43]提出了一种EMR差分隐私访问控制机制,授权不受信任的第三方用户仅拥有有限的数据视图,该方法不仅考虑了时间、用户特征等变量,也有效地控制了需要提供的信息量。针对包含敏感信息的分析数据(例如流行病分析需要分析身份来追踪疾病传播),文献[44]将访问控制机制与区块链及差分隐私相集成以实现数据的保护。

基于属性加密可实现对医疗数据更灵活、更细粒度的访问控制,同时可实现一对多的授权,降低数据加密开销。借助区块链去中心化、不可篡改特点实现数据所有者(患者、医疗机构等)对数据的控制,避免了不可信第三方对数据的泄露,同时实现了公开透明、可审计的访问控制,有助于制定更加灵活的访问策略。

3.2 基于交易匿名的隐私保护

保证区块链交易匿名性是医疗区块链数据共享平台隐私性的关键。各种数字签名是实现匿名交易的常用手段。随着医疗大数据挖掘与分析技术的发展,攻击者可能通过对链上交易分析得到交易地址之间的关联,降低区块链匿名性,对此可将交易进行混淆,或限制链上交易信息的发布来增大交易分析的难度。

3.2.1 *混币机制* 混币机制能够模糊交易发送方与接收方之间的关系,增加攻击者通过公开账本分析用户交易规律的难度,是一种有效的区块链隐私保护机制。混币机制基本思想可表达为

$$C_M(Z_1, C_A(Z_0, m), A) \rightarrow C_A(Z_0, m), A, \tag{3}$$

式(3)左侧代表中间方从发送方接收到的信息,右侧代表中间方将信息处理后发送给接收方的消息。发送方首先使用接收方公钥 C_A 对消息 Z_0 和 m 进行加密,再将加密信息、接收方地址 A、验证消息 Z_1 打包并使用中间方公钥加密,中间方对消息解密并验证无误后将 $C_A(Z_0, m)$ 发送到地址 A,接收方可对 $C_A(Z_0, m)$ 进行解密[45]。在混币过程中,中间方仅能验证加密的消息 $C_A(Z_0, m)$ 而无法得知具体消息内容。

在中心化混币机制中,混币过程由第三方节点执行,通过第三方节点处理,改变交易传递过程中流动轨迹,提高攻击者发现交易流向的难度。此类混币机制中提供混币服务的第三方掌握输入、输出地址之间的联系,存在泄露混币过程的可能性。Mixcoin是一种改进的中心化混币机制,为提高第三方节点的可信度,Mixcoin为用户设置了审计功能,混币服务节点的违规行为将造成信誉的损失,但Mixcoin并未在根本上解决第三方对混币过程的泄露问题[46]。去中心化混币机制不需要第三方节点参与,最早的去中心化混币方案是比特币论坛上提出的CoinJoin机制,核心思想是通过将多笔单输入-单输出交易合并为一笔多输入-多输出交易[45]。CoinJoin能够避免第三方泄露混币信息的风险,然而无法保护所有参与方的诚信,CoinShuffle的输出地址洗牌机制能够使参与者无法得到除自己以外的交易地址关联,但该方案在混币过程中需要参与者同时在线,攻击者可趁机发动拒绝服务攻击[47]。CoinParty由安全多方计算模拟可信第三方,在网络中存在恶意节点的情况下,能够保证混币过程的有效性[48]。

混币机制实现简单,能够提高区块链交易隐私性,并且不会对区块链原有共识机制造成影响,但单独使用混币机制的效果十分有限,因此需要在未来的研究中结合安全高效的加密方案保证混币过程的隐私性。目前的混币机制主要应用于数字货币领域,对于数据量大、类型复杂的医疗信息,需要选择有代表性的轻量数据作为交易信息,降低混币过程难度。在密码学机制保护下的混币机制需要充分考虑医疗服务器的计算和存储性能,同时尽量避免对区块链底层协议的修改。

3.2.2 *零知识证明* 零知识证明(zero-knowledge proof)由S. Micali和C. Rackoff提出,实现验证者在不需要任何有效信息的情况下验证消息的有效性。在区块链中,零知识证明能够隐藏交易双方的地址、金额等细节,并保证验证节点验证交易有效性。Zerocoin协议可将比特币铸造成Zerocoin,验证者通过零知识证明来验证Zerocoin是否被花费,而无法获取Zerocoin的交易信息[49]。Zerocash采用简洁非交互性零知识证明技术(zero-knowledge succinct non-interactive argument of knowledge, zk-SNARK),实现去中心化匿名支付,是目前区块链UTXO模型中隐私性最强的加密货币,与Zerocoin相比,Zerocash对交易金额保密并支持任意面值的交易[50]。Hawk是一种基于零知识证明的隐私智能合约框架,由编译器使用零知识证明等加密原语自动

编译,实现参与者、执行者和区块链三方之间的密码学协议。Hawk 的安全保障包括链上隐私和合约隐私;链上隐私指不向未参与合约的第三方公开交易细节;合约隐私则保护的是合约参与者之间的合约共识[51]。

在医疗应用场景中,用户身份与其他信息通常具有关联性,攻击者一旦通过交易分析手段将交易地址与用户真实身份关联起来,交易信息将完全暴露。零知识证明是保证隐私数据可验证性的有效方法,但由于效率问题,该方法一般适用于数据量较小的数据。对于海量、结构复杂的医疗数据,一方面需要选择具有代表性的数据作为验证信息打包成交易,另一方面零知识证明的效率是需要考虑的关键问题。

3.2.3　数字签名　群签名是 1991 年由 Chaum 提出的签名概念,一个群体中的任意成员可以以匿名的方式代表整个群体对消息进行签名,而且可以仅用单个群公钥公开验证。群签名由群管理员管理,群管理员的存在保证了签名的可追踪性。环签名是一种简化的群签名,签名者只有选择一个可能的签名者集合,获得其公钥,然后公布这个集合即可,所有成员平等,不需要管理者。在群签名中,群管理员可以撤销签名,揭露真正的签名者,而环签名在不添加额外信息的前提下无法揭示签名者。盲签名除了满足一般的数字签名条件外,还必须满足下面的两条性质。

1）签名者对其所签署的消息是不可见的,即签名者不知道他所签署消息的具体内容;

2）签名消息不可追踪,即当签名消息被公布后,签名者无法知道这是他哪次签署的。

数字签名常被用于保护区块链交易匿名性,文献[52]结合混币及聚合签名等技术,实现保护收付款者身份和交易金额隐私的全匿名区块链系统。文献[53]在 Mixcoin 的基础上采用盲签名进行改进,使第三方节点在混币过程中无法获得交易双方信息,避免中心化混币机制中交易双方在混币过程中对第三方节点透露信息。文献[54]采用环签名,使参与方在混币过程中无须与其他用户进行交互,为去中心化混币机制中的拒绝服务攻击提供了有效的防御措施。

交易匿名是区块链应用场景中保护用户隐私的关键,在实际医疗应用场景中,不仅需要在保护隐私的前提下对节点的真实身份进行认证,还需要考虑可监管问题,通过对非法交易的溯源,实现对恶意行为的追责。文献[55]提出一种基于身份认证的多密钥生成中心(key generation center, KGC)群签名方案,通过多 KGC 群签名保护交易双方的用户身份,实现在节点间验证身份的同时保护节点的隐私,该方案具有签名不可伪造性,符合联盟链区块链部分去中心化和节点隐私要求。针对匿名认证的监管问题,文献[56]在匿名认证过程中加入监管机制,通过匿名证书来确定用户的权限,同时在出示证书时可选择属性,以确保用户身份隐私信息不会泄露,可信中心(CA)在匿名认证的过程,若用户出现不诚信行为,审计人员可随时恢复出资产内容和交易方身份信息对其进行追责。文献[56]结合匿名认证技术、群签名和零知识证明技术,使 CA 能够对匿名证书进行身份追踪。文献[57]通过追踪比特币系统中交易信息在网络层的传播路径,将交易中的匿名地址和发起交易节点的 IP 地址相关联,实现比特币交易溯源。通过基于主动嗅探的邻居节点识别方法,对特定节点发送探测信息以推测邻居节点,实现轻量级交易溯源,该溯源机制有较强的实践意义和使用价值,但很难适用于非交易型数据区块链。

在医疗数据隐私保护的同时,还应当推动各医疗机构和监管部门的数据分享和连通,实现监管全覆盖,提高监管效率。监管机制有助于拓宽区块链在医疗领域的应用范围,提供健康的信息共享环境,预防及遏制攻击者利用区块链进行非法活动,使区块链技术在医疗领域发挥更大的价值。

3.2.4　安全通道协议　链下支付网络将原本链上大量的交易细节放在链下处理仅将最终结果上链,区块链作为仲裁平台以确保交易安全性,对支付过程中的异常进行处理,保护交易细节隐私性的同时,间接提高系统的交易吞吐量[58]。

闪电网络是最早的链下支付通道,主要包括 RSMC(recoverable sequence maturity contract)协议和 HTLC(hashed timelock contract)协议,使得系统内任意两个节点都可以通过支付通道实现转账,并在交易不上链的情况下确认交易[59]。文献[60]中的支付通道是一个不受信任的中介,用来发行匿名凭证(vouchers)。发送方使用比特币在支付通道中交换匿名凭证并发送收款方,收款方凭借匿名凭证兑换比特币,区块链作为仲裁平台,通过链上智能合约来确保凭证交易期间的公平性,可防止恶意中介的抵赖,确保 vouchers→BTC 及 BTC→vouchers 过程的公平性。

安全通道通常需要结合数据加密技术实现交易过程中的安全性。文献[60]与中心化混币机制不同的是,盲签名使得支付通道无法将凭证的发行与兑换相关联,即无法得知交易双方的关系,保证交换期间的隐

私性。文献[61]针对链下第三方支付通道的隐私保护问题,运用盲签名技术及零知识证明技术,使第三方不能获取用户的交易信息,从而防止第三方从中作恶,保证用户的隐私性。TumbleBit是一种兼容比特币系统的链下交易通道,在无信任的中介Tumbler中实现快速匿名的链下支付[62]。通过RSA和ECDSA密码学技术,Tumbler能够验证用户交易的真实性,而无法获取用户的交易信息,实现用户交易的不可链接性,从而保证用户隐私性。

Sprites方案是针对闪电网络方案效率的改进,通过调用全局的Hash原像管理智能合约(preimage manage contract)的状态,近似并行的获知交易是否完成,大大减少了最坏情况下用户的等待时间,减少了时间成本[63]。链下交易方案在交易双方没有直接支付通道时,允许中继节点作为服务提供者完成交易,中继节点能获取交易双方的交易信息,使用户的隐私受到威胁。针对支付通道网络(payment-channel network, PCN)的隐私性与并发性,文献[64]提出了Fulgor和Rayo协议,阻塞协议Fulgor能够为PNC提供隐私性保障,但在支付网络中存在一定死锁概率,非阻塞协议Rayo解决了交易过程中的死锁问题,但与Fulgor相比,Rayo牺牲了部分隐私性能。智能合约Multi-Hop HTLC是Fulgor和Rayo协议的核心,与闪电网络相比,该合约隐藏了交易过程中的支付路径,以保证交易双方的身份隐私。

医疗数据交易场景中可能涉及相关方权限、多方计算等复杂问题,如果将整个交易过程上链,可能会影响区块链性能。链下交易机制既能够为区块链的隐私保护提供新的思路,又在一定程度上提升交易吞吐量,但是这种方法需要对底层协议进行修改,对业务场景的限制较多,目前只能支持数字加密货币领域,将该方法用于医疗数据交易,需要将链上交易机制与链下系统有机结合,实现链上链下的协同。

3.2.5 *K匿名* 随着定位技术的快速发展,基于位置的服务(location-based services, LBS)在生活中日益普及,如医疗领域中的流行病监控,基于位置数据来实现受感染个人的跟踪,可以更好地了解社交距离的有效性,或根据先前的信息向可能接近已知病例的个人发送警报。位置和移动数据为更好地了解和抗击流行病提供了一条有效途径,然而LBS的使用可能伴随用户隐私的泄露,从而阻碍LBS的发展,因此研究者们针对位置隐私保护技术展开了研究。

基于K匿名的隐私保护方法被广泛应用于基于位置的服务中,K匿名是Latanya Sweeney和Pierangela Samarati提出的一种匿名化数据技术,该模型的思想是将属性标识进行泛化压缩处理,使得所有记录被划分到若干等价类,每个等价类具有相同的标识,实现将一个记录隐藏在一组记录中,因此也成为基于分组的隐私保护模型。文献[65]将匿名区的构造视为请求用户与协作用户间的两方博弈,利用区块链账本的不可篡改特性,为博弈双方的真实位置提供证据,惩罚泄露位置和欺骗行为的用户,以约束其自利行为,自利行为多的一方不能构造自己的匿名区。然而K匿名中存在背景知识攻击,攻击者可利用其拥有的背景知识以较高的概率推测出某些记录所对应个体的隐私信息。对此,虚拟位置选择(dummy-location selection, DLS)算法通过熵度量选择虚拟位置,并在此算法的基础上提出一种增强型的DLS,在熵明显增加的情况下扩大隐藏区域,同时保持着原始算法的隐私级别,对于利用背景知识攻击辅助寻找真实位置的攻击者增加了分析难度[66]。相比于K匿名,差分隐私对背景知识攻击可提供较强的隐私保护,在位置数据发布中实现了相关研究与应用。文献[67]根据车辆轨迹的Markov特点计算位置节点的敏感度,并根据位置敏感度,统计阈值和敏感度阈值添加适量Laplace噪音,增加数据的可用性及有效性。文献[68]的CPL算法将地理拓扑关系转化为带权无向图,计算各区域的隐私级别,并提出差分隐私预算模型,该机制不仅能够保护当前位置的隐私,还能够保证之前的位置信息不会因当前位置信息的发布而泄露。

K匿名及差分隐私是位置隐私保护的常用方法,不依赖复杂的密码学技术,因此用户计算开销较小,并可获得精准的查询结果。但此类方法无法提供有效的依据来证明其隐私保护水平,其安全性依赖于攻击者掌握背景知识的多少,因此,面对新型攻击,基于K匿名方法的模型需要不断完善。

面向用户的区块链隐私保护重点是访问权限及交易匿名性,表2介绍了访问控制方法及交易匿名性隐私保护方法,并对各种方法的优缺点进行对比。智能合约作为区块链上能够自动执行的脚本,可实现更灵活、自动化的访问控制策略,智能合约经部署上链后不可修改,因此可保证授权的公平性,避免决策中心化,但智能合约本身存在安全漏洞可能为访问控制带来新的问题。基于属性加密的访问控制能够实现一对多的授权和更细粒度的访问控制策略,但是其使用的双线性加密耗费了较多的时间成本。混币机制可以在一定程度上实现隐私保护,但效果十分有限,因此需要结合其他的隐私保护手段来提高混币过程的隐私性,如数

字签名、零知识证明等。零知识证明具有较强的安全性,但该方法效率较低,只适用于数据量较小的数据。数字签名能够保证交易匿名性,但在具体医疗应用场景中,在保护用户隐私的同时也要兼顾监管问题,对非法行为进行定位追溯,应当结合具体场景实现匿名性和监管机制的平衡。安全通道协议能够为区块链的隐私保护提供新的思路,同时减轻链上存储压力,但是目前这种方法只用于数字加密货币领域,面向医疗领域更加复杂海量的数据,还需要进一步研究。K 匿名是一种基于分组的方法,常用于位置隐私保护,由于其不依赖复杂的密码学技术,通常计算开销较小,但面对新型的攻击,其模型需要不断改进。

表 2 面向用户的隐私保护方法对比

Table 2 Comparison of privacy protection methods for user oriented

类别	方法	方法描述	优势	挑战及不足
访问控制	智能合约	在链上智能合约中设计访问控制策略	避免决策中心化,实现更加灵活、自动化的访问控制	智能合约本身的安全隐私问题
	属性加密	访问控制策略与属性集合相匹配	实现更灵活、更细粒度的访问控制	双线性加密耗费时间
交易匿名	混币机制	混淆交易收发方	提高区块链交易隐私	需要结合其他手段增强隐私保护效果
	零知识证明	不提供有效信息的情况下对消息进行验证	隐私性强	效率低
	数字签名	参与者匿名	保护交易信息	监管难题
	安全通道协议	链下交易链上仲裁	保护隐私的同时提升交易吞吐量	业务场景限制较多;对底层协议修改较多
	K 匿名	基于分组	计算开销小,可获得精准查询结果	背景知识攻击

4 未来研究挑战与方向

区块链技术在隐私保护方面拥有独特优势并取得了诸多研究成果,但区块链在自身的安全性、隐私性、交易性能以及与其他隐私保护技术结合等方面均存在很多需要解决的问题,这些问题也是将区块链应用于医疗信息隐私保护领域时必须解决的关键问题。

1) 链上数据发布问题。区块链通过公开透明的数据账本,促进医疗数据更合理的流通与共享,为打破医院内部及医院之间的数据孤岛提供技术支持。区块链去中心化、不可篡改特性能够避免系统中数据被个人或机构操纵,即使单个节点失效或遭受攻击也不会影响数据的完整性及整个系统的运行。医疗信息具有来源广、数据量大、数据类型复杂及隐私性强等特点,医疗数据集中包含的隐私信息会随着数据的发布和共享而泄露,对于此类数据若不结合其他隐私手段直接发布上链,势必对数据隐私造成影响。

由于区块容量有限,且交易性能与吞吐量也是区块链技术的一大瓶颈,区块链难以直接存储大规模数据。对于大规模的原始医疗数据的存储与计算处理,势必采用链上链下相结合的方式,因此区块链账本中需要发布具有代表性的数据,在保证不泄露隐私的前提下,确保链下存储数据的完整性与实际应用场景中数据交易的可验证性。此外,实现区块链与链下传统信息系统的安全对接是进一步研究的关键问题。

2) 匿名性与监管的平衡。区块链中支持匿名的交易,但从实际情况来看,区块链交易匿名并不具备真正的匿名,随着数据分析技术的发展,攻击者仍然能够通过数据挖掘等技术从链上公开的交易信息中得到地址之间的关联关系,推测真实的用户身份,因此用户隐私性无法得到真正的保障。从消极角度看,这会对交易匿名性造成严重影响,由于交易的透明性,用户身份暴露即链上信息会泄露;但从积极角度看,这有助于监管机制发现非法交易及犯罪痕迹,对相关组织及责任人实现责任的定位与追踪。在实际应用中,应结合具体场景平衡匿名性与可监管性之间的关系,在保证合法用户隐私的同时抵制非法行为。

基于上述问题,我们对区块链在医疗信息隐私保护领域中的研究方向进行展望。

1) 链下隐私计算与链上存证相结合。医疗原始数据只是医疗信息产业的基础,其价值属性远低于作出相关计算分析处理(如大数据挖掘、深度学习等)后的增值价值。单个机构组织通常面临样品数量不充足、数据维度不够丰富等问题,在计算过程中需要多方补充数据,然而,由于数据的高价值及隐私属性,各医疗数

据所有方在联合计算过程中需要对数据进行隐私保护。若将医疗数据直接提供给需求方或进行多方之间的交互,极易造成隐私的泄露,因此许多数据用户不愿提供数据,导致大量数据无法发挥更大的使用价值。

分离数据的所有权和使用权,实现多方医疗数据在无须出库的情况下在本地进行联合隐私计算,并将联合计算产生的增值数据共享给需求方,使得需求方在不接触原始数据即可获取数据的增值价值,降低隐私泄露的风险。同态加密、SMPC等是实现医疗数据隐私计算的有效方法,然而这些方法侧重于计算过程中对数据的保护,无法保证参与方提供数据的真实性以及计算结果的可验证性。区块链作为多方共同维护,能够有效防止抵赖的分布式账本技术,重在强调计算的可验证性,以及防止计算结果的篡改。区块链激励机制也有助于促进各计算参与方履行既定计算协议,增加作恶的成本。因此,未来可尝试将现有的隐私计算技术与区块链技术相结合,通过链下隐私计算保证数据的隐私性与可控性,并由区块链存储计算过程中的相关凭证,以实现计算结果的可验证性。

2）基于智能合约实现对数据的搜索授权及访问控制。由于海量的医疗数据种类复杂,对隐私的需求不同,不适于链上存储,因此在现有的医疗区块链应用中,大多将原始数据加密存储在链下数据库中,链上只保存数据的摘要(如Hash值)以保证数据的完整性。确保医疗数据在共享过程中的隐私安全,需要实现链外数据的安全搜索,设置合理的访问控制权限。智能合约具有自动执行的特点,可对用户进行访问授权,并在公开账本中记录对该用户的授权以及数据搜索日志。基于智能合约的访问控制策略有利于避免权力中心化,使得数据授权过程的透明公开化,实现更灵活、自动化的访问控制。

3）实现"以链制链"的监管机制以及非法交易溯源。监管技术也是隐私保护可持续发展的关键之一,在研究医疗信息隐私保护的同时,还要对非法行为进行监管和追溯,由于区块链的匿名性及去中心化,对监管区块链,尤其是公有链提出了更高的要求。在医疗区块链系统中可采用联盟链+私有链的监管机制,私有链由各机构内部维护,存储数据相关信息,联盟链中引入监管节点,存储数据交易信息以实现对数据的管控与追踪,既可发挥私有链交易速度快、隐私性强的优势,又能通过联盟链确保数据市场交易秩序、约束非法交易行为。此外,监管方需要在保护合法用户隐私的前提下及时发现非法交易,并对非法交易全过程及相关责任人进行定位追踪,为恶意行为提供犯罪证据。

5 总结

医疗信息化为医学领域的研究与服务带来便利的同时也带来了很多新的隐私安全问题。区块链的去中心化、不可篡改、匿名性、可追溯性等特点为保障医疗信息在共享过程中的安全隐私提供了解决思路。本文面向医疗信息领域,系统地梳理了区块链的定义、架构以及数据隐私保护需求,介绍了基于区块链技术的数据和用户隐私保护方法,并对这些方法进行了分析对比,总结了现有方法存在的问题和挑战,最后针对研究现状中的不足与挑战,展望了未来的研究方向。

参考文献:

[1] 王天屹, 刘爱萍. 大数据环境下医疗数据隐私保护对策研究[J]. 信息技术与网络安全, 2019, 38(8): 28-32.
WANG T Y, LIU A P. Research on privacy protection of medical information in big data[J]. Information technology and network security, 2019, 38(8): 28-32.

[2] 周洋. 医疗大数据的网络安全与隐私保护简析[J]. 信息安全与通信保密, 2017,15 (9): 28-32.
ZHOU Y, Brief analysis of network security and privacy protection of medical big data[J]. Information security and communications privacy, 2017,15 (9): 28-32.

[3] 胡荣磊, 何艳琼, 范晓红. 医疗隐私保护安全性技术研究[J]. 北京电子科技学院学报, 2018, 26(3): 46-54.
HU R L, HE Y Q, FAN X H. Research on safety technology of medical privacy protection[J]. Journal of Beijing electronic science and technology institute, 2018, 26(3): 46-54.

[4] 袁勇, 王飞跃. 区块链技术发展现状与展望[J]. 自动化学报, 2016, 42(4): 481-494.
YUAN Y, WANG F Y. Blockchain: the state of the art and future trends[J]. Acta automatica sinica, 2016, 42(4): 481-494.

[5] 祝烈煌, 高峰, 沈蒙, 等. 区块链隐私保护研究综述[J]. 计算机研究与发展, 2017, 54(10): 2170-2186.
ZHU L H, GAO F, SHEN M, et al. Survey on privacy preserving techniques for blockchain technology[J]. Journal of computer

research and development, 2017, 54(10): 2170-2186.
[6] 韩璇, 袁勇, 王飞跃. 区块链安全问题:研究现状与展望[J]. 自动化学报, 2019, 45(1): 206-225.
HAN X, YUAN Y, WANG F Y. Security problems on blockchain: the state of the art and future trends[J]. Acta automatica sinica, 2019, 45(1): 206-225.
[7] NAKAMOTO S. Bitcoin: a peer-to-peer electronic cash system[EB/OL]. [2020-08-10]. http://bitcoin.org/bitcion.pdf, 2009.
[8] 刘敖迪, 杜学绘, 王娜, 等. 区块链技术及其在信息安全领域的研究进展[J]. 软件学报, 2018, 29(7): 2092-2115.
LIU A D, DU X H, WANG N, et al. Research progress of blockchain technology and its application in information security[J]. Journal of software, 2018, 29(7): 2092-2115.
[9] 斯雪明, 徐蜜雪, 苑超. 区块链安全研究综述[J]. 密码学报, 2018, 5(5): 458-469.
SI X M, XU M X, YUAN C. Survey on security of blockchain[J]. Journal of cryptologic research, 2018, 5(5): 458-469.
[10] 邵奇峰, 金澈清, 张召, 等. 区块链技术: 架构及进展[J]. 计算机学报, 2018, 41(5): 969-988.
SHAO Q F, JIN C Q, ZHANG Z, et al. Blockchain: architecture and research progress[J]. Chinese journal of computers, 2018, 41(5): 969-988.
[11] 袁勇, 倪晓春, 曾帅, 等. 区块链共识算法的发展现状与展望[J]. 自动化学报, 2018, 44(11): 2011-2022.
YUAN Y, NI X C, ZENG S, et al. Blockchain consensus algorithms: the state of the art and future trends[J]. Acta automatica sinica, 2018, 44(11): 2011-2022.
[12] FENG Q, HE D B, ZEADALLY S, et al. A survey on privacy protection in blockchain system[J]. Journal of network and computer applications, 2019, 126: 45-58.
[13] 王童, 马文平, 罗维. 基于区块链的信息共享及安全多方计算模型[J]. 计算机科学, 2019, 46(9): 162-168.
WANG T, MA W P, LUO W. Information sharing and secure multi-party computing model based on blockchain[J]. Computer science, 2019, 46(9): 162-168.
[14] YU P, ZHANG S F, ZHONG J. Block-chain privacy protection based on fully homomorphic encryption[C]//Proceedings of the 2019 3rd International Conference on Innovation in Artificial Intelligence. Suzhou: ACM Press, 2019: 239-242.
[15] 徐文玉, 吴磊, 阎允雪. 基于区块链和同态加密的电子健康记录隐私保护方案[J]. 计算机研究与发展, 2018, 55(10): 2233-2243.
XU W Y, WU L, YAN Y X. Privacy-preserving scheme of electronic health records based on blockchain and homomorphic encryption[J]. Journal of computer research and development, 2018, 55(10): 2233-2243.
[16] ZHAO C, ZHAO S N, ZHAO M H, et al. Secure multi-party computation: theory, practice and applications[J]. Information sciences, 2019, 476: 357-372.
[17] GAO H M, MA Z F, LUO S S, et al. BFR-MPC: a blockchain-based fair and robust multi-party computation scheme[J]. IEEE access, 2019, 7: 110439-110450.
[18] 黄建华, 江亚慧, 李忠诚. 利用区块链构建公平的安全多方计算[J]. 计算机应用研究, 2020, 37(1): 225-230, 244.
HUANG J H, JIANG Y H, LI Z C. Constructing fair secure multi-party computation based on blockchain[J]. Application research of computers, 2020, 37(1): 225-230, 244.
[19] 朱岩, 宋晓旭, 薛显斌, 等. 基于安全多方计算的区块链智能合约执行系统[J]. 密码学报, 2019, 6(2): 246-257.
ZHU Y, SONG X X, XUE X B, et al. Smart contract execution system over blockchain based on secure multi-party computation[J]. Journal of cryptologic research, 2019, 6(2): 246-257.
[20] 蒋瀚, 徐秋亮. 基于云计算服务的安全多方计算[J]. 计算机研究与发展, 2016, 53(10): 2152-2162.
JIANG H, XU Q L. Secure multiparty computation in cloud computing[J]. Journal of computer research and development, 2016, 53(10): 2152-2162.
[21] DAS D. Secure cloud computing algorithm using homomorphic encryption and multi-party computation[C]//2018 International Conference on Information Networking (ICOIN). Chiang Mai: IEEE, 2018: 391-396.
[22] YANG Y H, WEI L J, WU J, et al. Block-SMPC: a blockchain-based secure multi-party computation for privacy-protected data sharing[C]//Proceedings of the 2020 The 2nd International Conference on Blockchain Technology. New York: ACM, 2020: 46-51.
[23] 刘敬浩, 平鉴川, 付晓梅. 一种基于区块链的分布式公钥管理方案研究[J]. 信息网络安全, 2018(8): 25-33.
LIU J H, PING J C, FU X M. Research on A distributed public key system based on blockchain[J]. Netinfo security, 2018(8): 25-33.

[24] 戴千一, 徐开勇, 郭松, 等. 分布式网络环境下基于区块链的密钥管理方案[J]. 网络与信息安全学报, 2018, 4(9): 23-35.

DAI Q Y, XU K Y, GUO S, et al. Blockchain-based key management scheme for distributed networks[J]. Chinese journal of network and information security, 2018, 4(9): 23-35.

[25] 熊平, 朱天清, 王晓峰. 差分隐私保护及其应用[J]. 计算机学报, 2014, 37(1): 101-122.

XIONG P, ZHU T Q, WANG X F. A survey on differential privacy and applications[J]. Chinese journal of computers, 2014, 37(1): 101-122.

[26] 欧阳佳, 印鉴, 刘少鹏, 等. 一种有效的差分隐私事务数据发布策略[J]. 计算机研究与发展, 2014, 51(10): 2195-2205.

OUYANG J, YIN J, LIU S P, et al. An effective differential privacy transaction data publication strategy[J]. Journal of computer research and development, 2014, 51(10): 2195-2205.

[27] 董祥千, 郭兵, 沈艳, 等. 一种高效安全的去中心化数据共享模型[J]. 计算机学报, 2018, 41(5): 1021-1036.

DONG X Q, GUO B, SHEN Y, et al. An efficient and secure decentralizing data sharing model[J]. Chinese journal of computers, 2018, 41(5): 1021-1036.

[28] 欧阳佳, 印鉴, 刘少鹏. 一种分布式事务数据的差分隐私发布策略[J]. 软件学报, 2015, 26(6): 1457-1472.

OUYANG J, YIN J, LIU S P. Differential privacy publishing strategy for distributed transaction data[J]. Journal of software, 2015, 26(6): 1457-1472.

[29] 张啸剑, 孟小峰. 面向数据发布和分析的差分隐私保护[J]. 计算机学报, 2014, 37(4): 927-949.

ZHANG X J, MENG X F. Differential privacy in data publication and analysis[J]. Chinese journal of computers, 2014, 37(4): 927-949.

[30] YANG M, MARGHERI A, HU R S, et al. Differentially private data sharing in a cloud federation with blockchain[J]. IEEE cloud computing, 2018, 5(6): 69-79.

[31] 薛腾飞, 傅群超, 王枞, 等. 基于区块链的医疗数据共享模型研究[J]. 自动化学报, 2017, 43(9): 1555-1562.

XUE T F, FU Q C, WANG C, et al. A medical data sharing model via blockchain[J]. Acta automatica sinica, 2017, 43(9): 1555-1562.

[32] KIM Y, KIM K H, KIM J H. Power trading blockchain using hyperledger fabric[C]//2020 International Conference on Information Networking (ICOIN). Barcelona: IEEE, 2020: 821-824.

[33] AZARIA A, EKBLAW A, VIEIRA T, et al. MedRec: using blockchain for medical data access and permission management [C]//2016 2nd International Conference on Open and Big Data (OBD). Vienna: IEEE, 2016:25-30.

[34] DAGHER G G, MOHLER J, MILOJKOVIC M, et al. Ancile: Privacy-preserving framework for access control and interoperability of electronic health records using blockchain technology[J]. Sustainable cities and society, 2018, 39: 283-297.

[35] XIA Q, SIFAH E B, ASAMOAH K O, et al. MeDShare: trust-less medical data sharing among cloud service providers via blockchain[J]. IEEE access, 2017, 5: 14757-14767.

[36] XIA Q, SIFAH E, SMAHI A, et al. BBDS: blockchain-based data sharing for electronic medical records in cloud environments [J]. Information, 2017, 8(2): 44.

[37] LIU J, LI X, YE L,et al. BPDS: A blockchain based privacy-preserving data sharing for electronic medical records[C]//2018 IEEE Global Communications Conference (GLOBECOM). Abu Dhabi: IEEE, 2018: 1-6.

[38] 王秀利, 江晓舟, 李洋. 应用区块链的数据访问控制与共享模型[J]. 软件学报, 2019, 30(6): 1661-1669.

WANG X L, JIANG X Z, LI Y. Model for data access control and sharing based on blockchain[J]. Journal of software, 2019, 30(6): 1661-1669.

[39] WANG S P, ZHANG Y L, ZHANG Y L. A blockchain-based framework for data sharing with fine-grained access control in decentralized storage systems[J]. IEEE access, 2018, 6: 38437-38450.

[40] WANG S P, WANG X, ZHANG Y L. A secure cloud storage framework with access control based on blockchain[J]. IEEE access, 2019, 7: 112713-112725.

[41] 杜瑞忠, 刘妍, 田俊峰. 物联网中基于智能合约的访问控制方法[J]. 计算机研究与发展, 2019, 56(10): 2287-2298.

DU R Z, LIU Y, TIAN J F. An access control method using smart contract for Internet of Things[J]. Journal of computer research and development, 2019, 56(10): 2287-2298.

[42] 刘敖迪, 杜学绘, 王娜, 等. 基于区块链的大数据访问控制机制[J]. 软件学报, 2019, 30(9): 2636-2654.

LIU A D, DU X H, WANG N, et al. Blockchain-based access control mechanism for big data[J]. Journal of software, 2019, 30(9): 2636-2654.

[43] GUTIERREZ O, SAAVEDRA J J, ZURBARAN M, et al. User-centered differential privacy mechanisms for electronic medical records[C]//2018 International Carnahan Conference on Security Technology (ICCST). Montreal: IEEE, 2018: 1-5.

[44] ALNEMARI A, ARODI S, SOSA V R, et al. Protecting infrastructure data via enhanced access control, blockchain and differential privacy[M]. Cham: Springer International Publishing, 2018: 113-125.

[45] 祝烈煌, 董慧, 沈蒙. 区块链交易数据隐私保护机制[J]. 大数据, 2018, 4(1): 46-56.
ZHU L H, DONG H, SHEN M. Privacy protection mechanism for blockchain transaction data[J]. Big data research, 2018, 4(1): 46-56.

[46] BONNEAU J, NARAYANAN A, MILLER A, et al. Mixcoin: anonymity for bitcoin with accountable mixes[M]. Berlin: Springer, 2014: 486-504.

[47] RUFFING T, MORENO-SANCHEZ P, KATE A. CoinShuffle: practical decentralized coin mixing for bitcoin[M]. Cham: Springer International Publishing, 2014: 345-364.

[48] ZIEGELDORF J H, GROSSMANN F, HENZE M, et al. CoinParty: secure multi-party mixing of bitcoins[C]//The Fifth ACM Conference on Data and Application Security and Privacy (CODASPY 2015). New York: ACM, 2015: 75-86.

[49] MIERS I, GARMAN C, GREEN M, et al. Zerocoin: anonymous distributed E-Cash from bitcoin[C]//2013 IEEE Symposium on Security and Privacy (SP). Berkeley: IEEE, 2013: 397-411.

[50] SASSON E B, CHIESA A, GARMAN C, et al. Zerocash: Decentralized anonymous payments from bitcoin[C]//2014 IEEE Symposium on Security and Privacy. San Jose: IEEE, 2014: 459-474.

[51] KOSBA A, MILLER A, SHI E, et al. Hawk: The blockchain model of cryptography and privacy-preserving smart contracts[C]//2016 IEEE Symposium on Security and Privacy (SP). San Jose: IEEE, 2016: 839-858.

[52] 王子钰, 刘建伟, 张宗洋, 等. 基于聚合签名与加密交易的全匿名区块链[J]. 计算机研究与发展, 2018, 55(10): 2185-2198.
WANG Z Y, LIU J W, ZHANG Z Y, et al. Full anonymous blockchain based on aggregate signature and confidential transaction[J]. Journal of computer research and development, 2018, 55(10): 2185-2198.

[53] VALENTA L, ROWAN B. Blindcoin: blinded, accountable mixes for bitcoin[M]//Financial Cryptography and Data Security. Berlin: Springer, 2015: 112-126.

[54] WIJAYA D A, LIU J, STEINFELD R, et al. Monero ring attack: recreating zero mixin transaction effect[C]//2018 17th IEEE International Conference on Trust, Security and Privacy in Computing and Communications/12th IEEE International Conference on Big Data Science and Engineering (TrustCom/BigDataSE). New York: IEEE, 2018: 1196-1201.

[55] 杨亚涛, 蔡居良, 张筱薇, 等. 基于 SM9 算法可证明安全的区块链隐私保护方案[J]. 软件学报, 2019, 30(6): 1692-1704.
YANG Y T, CAI J L, ZHANG X W, et al. Privacy preserving scheme in block chain with provably secure based on SM9 algorithm[J]. Journal of software, 2019, 30(6): 1692-1704.

[56] 王震, 范佳, 成林, 等. 可监管匿名认证方案[J]. 软件学报, 2019, 30(6): 1705-1720.
WANG Z, FAN J, CHENG L, et al. Supervised anonymous authentication scheme[J]. Journal of software, 2019, 30(6): 1705-1720.

[57] 高峰, 毛洪亮, 吴震, 等. 轻量级比特币交易溯源机制[J]. 计算机学报, 2018, 41(5): 989-1004.
GAO F, MAO H L, WU Z, et al. Lightweight transaction tracing technology for bitcoin[J]. Chinese journal of computers, 2018, 41(5): 989-1004.

[58] 潘晨, 刘志强, 刘振, 等. 区块链可扩展性研究: 问题与方法[J]. 计算机研究与发展, 2018, 55(10): 2099-2110.
PAN C, LIU Z Q, LIU Z, et al. Research on scalability of blockchain technology: problems and methods[J]. Journal of computer research and development, 2018, 55(10): 2099-2110.

[59] DECKER C, WATTENHOFER R. A fast and scalable payment network with bitcoin duplex micropayment channels[C]//Symposium on Self-stabilizing Systems. Cham: Springer, 2015: 3-18.

[60] HEILMAN E, BALDIMTSI F, GOLDBERG S. Blindly signed contracts: anonymous on-blockchain and off-blockchain bitcoin transactions[M]. Berlin: Springer, 2016: 43-60.

[61] GREEN M, MIERS I. Bolt: anonymous payment channels for decentralized currencies[C]//Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security. New York: ACM, 2017: 473-489.

[62] HEILMAN E, ALSHENIBR L, BALDIMTSI F, et al. TumbleBit: an untrusted bitcoin-compatible anonymous payment hub[C]//

Proceedings 2017 Network and Distributed System Security Symposium. San Diego: Internet Society, 2017. http://eprint.iacr.org/2016/575.pdf.

[63] MILLER A, BENTOV I, BAKSHI S, et al. Sprites and state channels: payment networks that go faster than lightning[M]. Cham: Springer International Publishing, 2019: 508-526.

[64] MALAVOLTA G, MORENO-SANCHEZ P, KATE A, et al. Concurrency and privacy with payment-channel networks[C]//Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security. New York: ACM, 2017: 455-471.

[65] 刘海, 李兴华, 雒彬, 等. 基于区块链的分布式K匿名位置隐私保护方案[J]. 计算机学报, 2019, 42(5): 942-960.
LIU H, LI X H, LUO B, et al. Distributed K-anonymity location privacy protection scheme based on blockchain[J]. Chinese journal of computers, 2019, 42(5): 942-960.

[66] NIU B, Li Q H, ZHU X Y, et al. Achieving k-anonymity in privacy-aware location-based services[C]//IEEE INFOCOM 2014-IEEE Conference on Computer Communications. Toronto: IEEE, 2014: 754-762.

[67] 朱维军, 游庆光, 杨卫东, 等. 基于统计差分的轨迹隐私保护[J]. 计算机研究与发展, 2017, 54(12): 2825-2832.
ZHU W J, YOU Q G, YANG W D, et al. Trajectory privacy preserving based on statistical differential privacy[J]. Journal of computer research and development, 2017, 54(12): 2825-2832.

[68] 吴云乘, 陈红, 赵素云, 等. 一种基于时空相关性的差分隐私轨迹保护机制[J]. 计算机学报, 2018, 41(2): 309-322.
WU Y C, CHEN H, ZHAO S Y, et al. Differentially private trajectory protection based on spatial and temporal correlation[J]. Chinese journal of computers, 2018, 41(2): 309-322.

A Survey on Medical Information Privacy Protection Based on Blockchain

LIU Wei[1,2,3], PENG Yufei[3,4], TIAN Zhao[1,4], SHENG Zhaoyang[2,3], LI Yang[2,3], SHE Wei[1,3,4]

(1. *School of Software*, *Zhengzhou University*, *Zhengzhou* 450002, *China*; 2. *Hanwei Internet of Things Research Institute*, *Zhengzhou University*, *Zhengzhou* 450001, *China*; 3. *Henan Collaborative Innovation Center for Internet Medical and Health Services*, *Zhengzhou University*, *Zhengzhou* 450001, *China*; 4. *School of Information Engineering*, *Zhengzhou University*, *Zhengzhou* 450001, *China*)

Abstract: With the development of medical information, the privacy of medical data has attracted widespread concern among researchers in the process of sharing and accessing. As a decentralized, anonymous, non-tamperable distributed ledger technology, blockchain provided new ideas for solving privacy protection problems in medical scenarios. Firstly, the privacy protection requirements of medical data were listed, and the overall architecture of the blockchain was introduced. Then, the medical information privacy protection technology was introduced in detail, which was divided into data-oriented privacy protection and user-oriented privacy protection. Data-oriented privacy protection was referred to as the protection of sensitive information itself. Encryption-based privacy protection methods, distortion-based privacy protection methods and privacy protection methods based on restricted release were used. User-oriented privacy protection was the privacy protection of data users. It included privacy protection based on access control and transaction anonymity. Finally, the characteristics of various methods were compared and the research status of blockchain in the field of privacy protection was summarized. We prospected the development direction of blockchain in the field of medical information privacy protection was discussed.

Key words: blockchain; medical information; data privacy; users privacy; decentralization

(责任编辑:方惠敏　孔　薇)