Medical Chain: 联盟式医疗区块链系统

张 超[1] 李 强[1] 陈子豪[1] 黎祖睿[1] 张 震[1]

摘 要 医疗数据共享、防篡改、防泄漏一直是困扰医疗行业的难题. 一位患者在转诊时, 往往无法提供以往的就诊信息, 原因在于国内各医院医疗信息大多数情况下无法共享, 而仅能通过病历、检验单等极易丢失的纸质信息来实现一部分医疗信息的共享. 同时, 患者的医疗信息极易造成泄露, 在出现医疗纠纷时所提供的医疗信息也无法保证真实性与公正性. 本文设计了一个基于实用拜占庭容错算法 (Practical Byzantine fault tolerance, PBFT) 的联盟式医疗区块链系统, 该系统是一个多节点共同维护与共享的, 并且能够防止医疗数据被篡改、泄露的医疗系统, 可用来解决这些医疗难题. 与现有医疗区块链系统相比, 本系统具有一定的优越性与较好的适用性.

关键词 医疗区块链, 医疗数据, 共识算法, 医疗信息共享

引用格式 张超, 李强, 陈子豪, 黎祖睿, 张震. Medical Chain: 联盟式医疗区块链系统. 自动化学报, 2019, **45**(8): 1495−1510

DOI 10.16383/j.aas.c180131

Medical Chain: Alliance Medical Blockchain System

ZHANG Chao[1] LI Qiang[1] CHEN Zi-Hao[1] LI Zu-Rui[1] ZHANG Zhen[1]

Abstract Medical data sharing, anti-tampering, and anti-leakage have always been difficult problems in medical industry. When a patient is on referral, he or she often can not provide information on previous visits. The reason is that many hospitals in China can not share medical information in most cases, and they can only pass the easily lost paper information such as medical records and checklists to achieve part of the medical information sharing. On the other hand, the patient′s medical information is easy to leak, and medical information provided in a medical dispute can not be guaranteed in terms of authenticity and impartiality. In this research, an alliance medical blockchain system based on the practical Byzantine fault tolerance algorithm (PBFT) is designed, which is a multi-node maintenance and sharing system, and is able to prevent medical data from being tampered with or leaked. This medical system can be used to solve the above medical problems. Compared with the existing medical blockchain systems, this system has some advantages and good applicability.

Key words Medical blockchain, medical data, consensus algorithm, medical information sharing

Citation Zhang Chao, Li Qiang, Chen Zi-Hao, Li Zu-Rui, Zhang Zhen. Medical chain: alliance medical blockchain system. *Acta Automatica Sinica*, 2019, **45**(8): 1495−1510

医疗信息是患者的宝贵资料, 然而在目前国内各医院的医疗系统中, 这些信息大多不能相互通用, 导致患者每到一个医院需要重新办理一张医疗卡, 记录患者的医疗信息. 而患者以往的医疗信息很多时候只能够通过模糊的记忆来获取. 虽然大部分医院都会采用纸质病历, 但纸质病历十分容易损坏或者遗失, 是一种十分不可靠的医疗信息记录方式. 另一方面, 使用传统数据库来实现医疗信息共享往往会因一些没有职业道德的工作人员倒卖泄露, 对患者造成进一步的损失. 因此, 医疗人员与患者迫切需要一种能够在各医院之间实现医疗信息共享, 并能够保证患者信息不会泄露的系统, 而区块链是目前实现这一系统的绝佳方式.

区块链 (Blockchain) 是由多个独立节点参与的分布式数据库系统[1], 能安全地存储比特币[2] 交易或其他数据, 并保证这些数据或信息的安全, 防止被篡改和伪造. 区块链一般部署在 P2P 网络中, 区别于常见的关系型数据库和非关系型数据库, 区块链使用数字签名、哈希算法等加密算法及分布式共识算法, 所存储的数据极难被篡改、销毁或被抹除数据库操作日志. 区块链技术具有去中心化、时序数据、集体维护、可编程和安全可信等特点[3].

按照参与者的不同来划分, 区块链可以分为公有链、联盟链和私有链. 公有链的参与者可以是任何人, 所有想参与到公有链维护的人都可以加入, 服务于比特币的区块链即是一种公有链. 私有链指的

收稿日期 2018-03-07 录用日期 2018-08-14
Manuscript received March 7, 2018; accepted August 14, 2018
四川省基于云模式生态化大型软件平台关键技术研发及应用基金 (2018GZ0105), 四川省基于全 IP 化的新一代音视频生产云平台关键技术研发及应用基金 (2018GZ0104) 资助
Supported by Research and Application of Key Technologies for Large-scale Software Platform Based on Cloud Mode Ecological-ization, Sichuan Province (2018GZ0105) and Development and Application of Key Technologies for New Generation of Audio and Video Production Cloud Platform Based on All-IP, Sichuan Province (2018GZ0104)
本文责任编委 袁勇
Recommended by Associate Editor YUAN Yong
1. 四川大学计算机学院 成都 610065
1. College of Computer Science, Sichuan University, Chengdu 610065

是一个实体内部使用, 信息不公开的区块链. 这里的实体可以是公司、银行、医院等, 目前国内各银行所研究的区块链多为私有链. 联盟链指的是一种由多个实体构成, 并且带有准入限制的区块链. 联盟链相对于公有链, 并不是想加入就可以加入, 而是需要得到一定许可, 才可以准入, 并且其所存储的信息访问权限受到这些实体的约束, 仅在一定条件下才可以向外界公开. 联盟链相对于私有链, 其区别在于参与的实体是多家不同的公司或集团, 这些实体共同维护区块链, 并共享区块链中的信息.

医疗区块链中的实体是一家家医院和医疗机构, 这些实体之间行政、财务等完全独立, 是一个个不同的实体. 同时这些实体接受政府监督与管理, 且有严格的准入和分级制度, 有一定准入限制. 医疗数据既是一位患者的个人隐私, 又涉及到国家机密, 因此其访问是有严格权限限制的. 根据以上特点可以看出, 医疗区块链实际上是一种联盟链.

目前的医疗区块链系统大多采用 POX 系列共识算法来达成分布式共识. 区块链共识算法指的是区块链系统中达成分布式一致的算法, 也就是区块链系统检测数据的合法性, 并将区块加入到区块链系统里的确认机制. POX 系列算法目前主要有工作量证明 (Proof of work, POW)、权益证明 (Proof of stack, POS) 和股份授权证明 (Delegate proof of stack, DPOS). 区块链共识算法所要解决的问题是拜占庭将军问题[4]. 该问题难解的原因在于, 任何时候系统中都可能存在多个提案 (因为提案成本很低), 且要完成最终的一致性确认, 该过程十分困难. POX 系列算法即通过增加提案成本, 放宽最终一致性确认的需求来达成共识. POW 算法一般应用于公有链, 需要较多节点和较大的算力来维护[5], POS 算法生成区块的过程取决于节点所持有的数字货币[5], DPOS 算法需要数字货币持有者选出一定数量 (一般为 101 个) 的区块生成者, 且每隔一段时间会重新选举区块生成者[6]. 它们并不适应医疗区块链的需求, 即不需要大的算力维护, 也不需要电子货币的产生, 且节点数量较少、灵活可变. 本文首次使用实用拜占庭容错算法 (Practical Byzantine fault tolerance, PBFT)[7], 构建了一种既能以较少节点来启动运行, 又不需要大量算力来维护的联盟式医疗区块链 (Medical chain).

1 已有研究

1.1 PBFT 算法

实用拜占庭容错算法 (PBFT)[7] 由 Castro 和 Liskov 于 1999 年提出, 用于解决原始拜占庭容错算法效率不高的问题. 该算法相较于原始拜占庭容错算法, 复杂度由指数级降低到多项式级[7], 使得拜占庭容错算法能用于实际应用中.

PBFT 算法是一种状态机副本复制算法, 每个状态机的副本都保存了服务的状态, 同时也实现了客户端所有合法请求的操作, 能够保证在满足分布式系统活性和安全性的前提下, 允许 $(n-1)/3$ 个节点出错 (数据丢失、不工作等), 其中 n 为分布式系统中所有参与共识过程的节点数量. 即该算法能够保证系统在 $(n-1)/3$ 个节点出现故障或恶意操作的情况下, 依然能正确达成分布式共识.

PBFT 算法中存储副本的节点都在一个视图 (View) 的轮换过程之中. 在编号为 v 的视图中, 一个副本节点是主节点, 其他副本节点是备份节点. 主节点主要用来接收客户端发送的请求消息, 由公式 $p = v \mod |R|$ 计算选出, $|R|$ 表示存储副本节点的个数. 若主节点失效, 则启动视图更换, 更改当前的视图编号 v, 再根据上面的公式选出主节点. PBFT 算法的过程如下:

1) 客户端向主节点发送请求操作消息, 主节点接收到请求操作消息并校验正确后, 保存该消息, 并依据该请求操作消息生成预准备消息, 广播给各备份节点.

2) 各备份节点接收到预准备消息并校验正确后, 保存该消息, 并以该预准备消息为依据, 生成准备消息广播给主节点和其他备份节点.

3) 各存储副本的节点接收到准备消息并校验正确后, 保存该消息, 并以该准备消息为依据, 生成提交消息给客户端、主节点和其他备份节点.

4) 各存储副本的节点接收到 $(2n+1)/3$ 个提交消息并校验正确后, 则执行来自客户端的请求操作消息里的操作.

5) 客户端接收到 $(n+2)/3$ 个提交消息, 验证正确并接受后, 便认为该消息已被副本节点集群所承认与执行. 这里的客户端接受 $(n+2)/3$ 个提交消息而不是 $(2n+1)/3$ 个的原因在于失效的节点数量不超过 $(n-1)/3$, 因此 $(n-1)/3+1$ 个一致响应必定能够保证结果是正确的. PBFT 算法的三阶段过程如图 1 所示.

1.2 医疗区块链

目前区块链的应用主要在金融方面, 在医疗方面的应用相对较少, 原因在于区块链的关注点局限在比特币等数字货币区块链系统上. 国内, 薛腾飞等[8] 利用改进的 DPOS 共识机制提出了一种医疗机构联盟服务器群 (Medical institution federate servers, MIFS) 和审计联盟服务器群 (Auditing federate servers, AFS) 相结合的医疗区块链系统 MDSM. 国外, Azaria 等[9] 利用以太坊区块链, 实现了一个医疗区块链与大数据相结合的医疗信息共享平台 MedRec[10]. Ivan[11] 分析了将区块链作

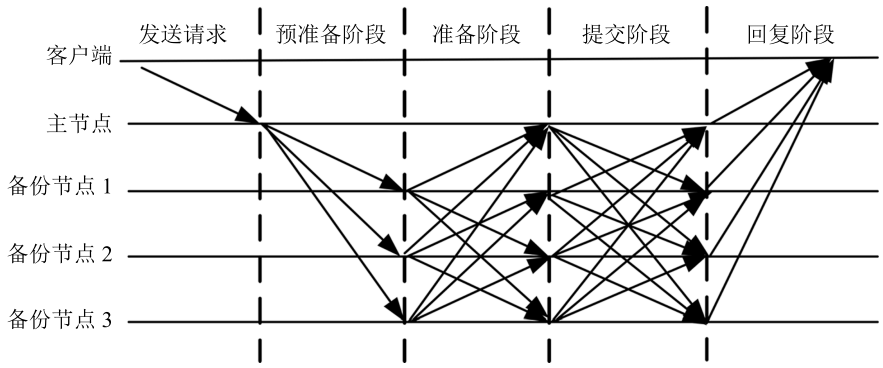


图 1　PBFT 算法三阶段过程

Fig. 1　PBFT algorithm three phases process

为保护医疗健康数据存储的新颖方法、实施障碍以及从当前技术向区块链解决方案逐步过渡的计划. Shrier 等[12] 采用美国麻省理工学院的 OPAL/Enigma 加密平台与区块链技术相结合的方式, 为医疗保健信息的存储和分析创造了一个安全环境. Kuo 等[13] 采用了隐私保护在线机器学习与私有区块链技术相结合的模式. Witchey[14] 介绍了医疗交易单 (Transaction) 验证系统和方法. 可以看出国内关于区块链在医疗方面的应用与研究相对来说较少, 且大多在应用层面上.

在已有的医疗区块链系统中, 薛腾飞等以改进的 DPOS 算法作为共识算法的区块链系统, 以及 Azaria 等利用以太坊, 即 POW 算法作为共识算法的区块链系统, 所采用的共识算法都属于 POX 系列算法. 其中, 薛腾飞等的医疗区块链启动时需要拥有 101 个节点的医疗机构联盟服务器群 (Medical institution federate servers, MIFS) 以及拥有 20 个节点的审计联盟服务器群 (Auditing federate servers, AFS), 也就是说该医疗区块链的启动需要 121 所医院或医疗机构同时参与到区块链维护中, 因此启动代价较大, 不适合从早期探索到后期大规模成熟应用的渐进研究过程. Azaria 等使用的以太坊区块链所采用的共识机制为 POW 算法, 其维护过程类似于比特币, 即互联网中所有人都可以随时参与或退出维护过程, 算力浪费较大, 且每次操作都需要支付一定的代币作为报酬, 并不适合医疗区块链的使用. ModelChain[13] 并不是一个专为医疗所设计的区块链, 其共识算法 — 信息证明 (Proof of information, POI) 将机器学习与工作量证明算法相结合, 所需算力会更加庞大, 因此也不适合医疗区块链.

PBFT 算法仅需要 4 台以上的节点即可启动, 相对于基于 POX 算法的区块链系统, 基于 PBFT 算法的区块链系统的启动代价小, 适合早期探索与后期扩展, 且不需要大量算力来维护. 因此, 本文将采用 PBFT 共识算法来实现一种适合医疗系统的区块链.

1.3　密码学基础

密码学是一个区块链系统最重要的组成部分, 是实现区块链功能的基础. 区块链中涉及到的密码学基础主要包括加解密算法、哈希算法与数字摘要、Merkle 树、数字签名、数字证书和 PKI 体系等.

加解密算法是密码学中的核心技术之一, 主要可以分为对称加密和非对称加密两种. 在对称加密中, 加密与解密的密钥相同, 其计算的效率较高, 但密钥在传输过程中容易泄露. 在非对称加密中, 加密密钥与解密密钥不同. 上面两种密钥一般公开一把, 保密一把. 被公开的密钥称为公开密钥, 简称公钥, 因此非对称加密又称为公开密钥加密. 被保密的密钥称为私有密钥, 简称私钥. 在本系统中主要采用非对称加密技术对一些数据进行加密. 其加密过程通过加密密钥与加密算法, 对这些数据进行加密, 获得密文. 在解密时, 通过解密密钥与解密算法, 重新获得这些数据. 常见的非对称加密算法有 RSA 和 ECC 等.

哈希算法是把可变长度输入串转换成固定长度输出串的一种算法, 具有单向性、抗第二原像攻击、抗强碰撞攻击等性质. 常见的哈希算法有 MD5[15]、SHA1[16]、SHA2[16] 等. 而数字摘要即为明文经哈希算法计算后得到的固定长度输出串, 同样的明文经同样的哈希算法计算后, 得到的摘要都是相同的, 否则, 其摘要必定不一致. 在区块链系统中, 通过哈希算法, 将区块或交易单相关内容作为输入, 计算出摘要, 来生成区块或交易单的 ID.

Merkle 树[17] 是一种存储哈希值的树, 其叶子节点存储的是每条数据的哈希值, 非叶子节点存储的是其所有孩子节点的哈希值. 其优势在于可以快速定位某条数据是否被篡改过. 在本区块链系统中, 将各交易单 ID 作为叶子节点构建 Merkle 树, 用来快速定位被篡改过的交易单.

数字签名是附加于被签名数据的一个大整数, 用以确定被签名数据的完整性与签署者的身份. 在生成时, 使用签名算法、签名者的私钥, 对被签名数据进行计算, 得到的结果即为数字签名. 在验证时, 利用签名者的公钥和数字签名, 来验证被签名数据的完整性. 常见的数字签名方案有 RSA[18] 签名方案, 数字签名算法 (Digital signature algorithm, DSA)[19] 以及椭圆曲线签名方案[20]. 在区块链系统中, 使用数字签名技术对区块、交易单相关信息进行签名, 来保证区块、交易单的完整性与不可抵赖性, 进而保护整条区块链的安全.

数字证书由可信的证书颁发机构 (Certificate authority, CA) 生成, CA 的主要作用是使用自己的私钥, 对已经验证身份的证书申请者的个人资料和公钥进行签名, 生成证书. 发送者 A 在发送原始数据、数字签名时, 也发送自己的数字证书, 这样接收者 B 在接收数据后, 就可以根据数字证书验明发送者 A 的数据.

证书颁发机构 CA 属于 PKI 体系的一部分, 它的功能是绑定证书持有者与密钥, 实现身份认证, 为用户提供证书申请、获取、查询、撤销等功能, 具有完整性、不可抵赖性和保密性[21] 的特点. 除 CA 以外, PKI 体系还主要包括密钥管理机构 (Key management center, KMC) 和管理证书撤销列表 (Certificate revocation list, CRL) 等. KMC 主要用于对密钥的生命周期进行管理, CRL 主要用于管理失效的证书清单. 在 Medical chain 中引入 PKI 体系, 用以证明各区块链参与者的身份信息, 保证数据的完整性与不可抵赖性. PKI 体系相关机构可以部署在卫生管理部门, 这虽然引入了一定的中心化, 但可以通过将证书申请过程中的各种文件、图片、视频等信息存入区块链, 来防止这些中心化的操作可能会出现的恶意行为.

2 基于 PBFT 算法的联盟式医疗区块链系统

Medical chain 中的区块链系统主要包括存储管理、节点管理和用户管理三个部分. 存储管理指的是逻辑上如何将医疗数据存储到区块链上, 以及区块链如何实际地存储在各种存储设备上. 节点管理是对运行区块链系统的各个节点的管理. 用户管理指的是对 Medical chain 中的参与者的认证与权限管理.

2.1 区块链存储管理

区块链存储管理主要包括区块、交易单和医疗数据存储方面管理, 是医疗区块链最基本的构成部分.

2.1.1 医疗区块链与医疗区块

医疗区块链主要由两部分构成: 区块 (Block) 和交易单 (Transaction). 一条区块链由一个个记录着前一个区块 ID 的区块组成, 而每个区块又包含了若干交易单. 这些交易单是实际存储区块链 (Blockchain) 数据的载体. 举例来讲, 一条区块链可以看作是一个数据库, 构成区块链的每一个区块可以看作是数据库中的一张表, 交易单可以看作是每张表上的一条记录 (Record). 一条区块链的构成如图 2 所示.

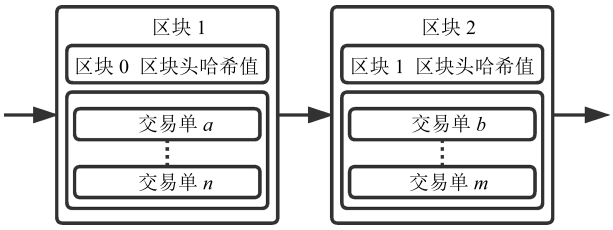


图 2 一条区块链的构成

Fig. 2 Composition of a blockchain

每个区块的具体构成如图 3 所示, 一个区块主要由区块头和区块头以外的内容构成. 区块头表示的是需要进行数字签名的部分. 区块头中包含上一个区块的 ID, 区块生成者的公钥, 由交易单 ID 生成的 Merkle 树根哈希值和生成区块的时间戳. 区块头以外的内容包括区块生成者对于区块头的数字签名, 交易单 ID 的个数, 和保存在此区块中所有的交易单 ID. 数字签名是为了保证区块内容不被篡改, 并且确保区块生成者在生成恶意区块后无法抵赖. 另外, 区块中仅保存交易单的 ID, 即仅保存指向某个交易单的索引, 而不保存交易单本身, 这样便可使每个区块容量降低, 便于同步与备份. 区块、交易单物理上都是保存在数据库里的, 在逻辑上以区块链的形式来存储. 在交易单设计存储上, 实际上只是在正常存储于数据库的数据上添加交易单 ID, 交易单类型、时间戳、公钥、数字签名等交易单字段信息, 将所要存储的信息作为交易单内容, 形成逻辑上的交易单, 其物理存储上与一般数据存储并无太大区别.

每个交易单中的内容如图 3 所示. 交易单类型表示这个交易的类型, 如增加、删除、查询和修改, 以指示验证器集群进行相应的操作, 关于验证器集群的介绍参见第 2.2.3 节. 之所以在增删查改的操作时使用交易单而不是直接访问某一拥有区块链的节点, 有两个原因. 首先某一节点可能进行恶意操作, 违规暴露患者的信息, 或是篡改患者的信息等. 另一方面是为了将操作者的操作记录在区块链中, 操作者在进行操作时需要使用自己的私钥对交易单进行数字签名, 使得操作者对于自己进行过的操作无法抵赖. 交易单内容是该交易单中所存储的内容, 如

患者的医疗信息等. 时间戳表示该交易单生成的时间, 公钥为交易单生成者的公钥, 交易单 ID 为对交易单类型, 交易单内容, 时间戳和公钥进行哈希运算后生成的哈希值. 哈希算法和编码算法可以选择 SHA-256[16] 哈希算法或 BASE64[22] 编码算法, 其可靠性在各种区块链系统中得到了检验. 数字签名是交易单生成者对交易单 ID 的签名, 防止交易单被篡改.

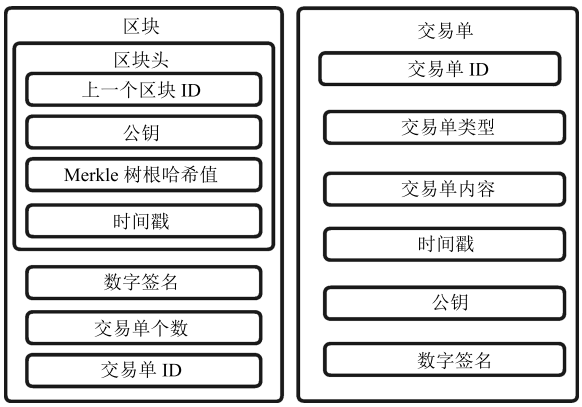


图 3　区块与交易单构成

Fig. 3　Composition of block and transaction

2.1.2　医疗信息交易单

交易单中存储的内容, 包括患者信息、医生信息、医疗记录信息、各节点的信息等. 即交易单实际上是传统数据库每张表里每条数据记录的载体, 而交易单内容相当于每条记录. 交易单内容主要有如下几类:

1) 实体信息类. 主要用于记录患者、医疗人员等实体的详细信息, 如患者的身份证号、姓名、性别、年龄、婚姻状况, 联系方式等个人信息, 以及患者所拥有的密钥中的公钥信息. 对于医护人员, 与患者较为相似, 另外还要记录其所在的医院、科室、级别等. 对于实体中的隐私信息, 如姓名、联系方式等, 在存储到区块链之前, 会采用非对称加密技术, 使用实体公钥对这些信息进行加密, 但不加密其公钥信息, 方便实体检索自己的信息. 在需要查看这些信息时, 经实体授权, 使用实体自己的私钥进行解密, 获取这部分信息.

2) 医疗信息类. 主要用于记录患者的相关医疗信息, 如某患者 P 在某一时间到医院 H 接受医生 D 的门诊时, 则生成一条门诊记录, 主要包括: 就诊时间, 就诊地点, 就诊的具体情况等. 若患者进行了类似于 B 超之类存在图片或视频的检查, 则对产生的图片或视频进行哈希运算获得哈希值, 存入到交易单里. 这样, 当图片或视频被篡改时, 其哈希值就会发生变化, 与存储在区块链中的哈希值不符, 从而保证多媒体资料的防篡改. 该条医疗信息所在交易单 ID 即为该条医疗信息的 ID. 图片或视频等较大的数据存储在生成节点, 并备份到医疗管理部门. 医疗信息类交易单中的公钥为产生该医疗信息交易单的医疗人员的公钥与其所在医疗机构的公钥, 数字签名为医疗人员的数字签名和医疗机构的数字签名. 这样在出现医疗纠纷等情况时, 可以根据医疗信息交易单中的公钥和数字签名, 来保证医疗人员和医疗机构的诊断或检验信息的不可抵赖性. 对于不属于医疗机构的有资质的医疗人员, 则仅需医疗人员的公钥和数字签名即可. 在医疗信息提交与修改方面, 若医疗人员属于某医疗机构, 则当进行提交或修改操作时, 需要通过其所属的医疗机构. 若有资质的医疗人员不属于医疗机构, 则可以直接进行. 医疗人员与医疗机构的从属关系也保存在区块链当中. 同时患者、法院、公安部门等可以通过向医疗管理部门申请提交或修改医疗信息, 用来更正有问题的医疗信息.

3) 实体 – 信息关联信息类. 该类信息主要用于关联实体与医疗信息或其他敏感信息, 因为该类信息需要进行加密操作, 防止实体的隐私泄露. 例如, 当患者 P 就诊后产生一条医疗信息 R 的时候, 同时会产生患者 P 与医疗信息所在的交易单 ID 的关联信息, 该交易单 ID 由患者的公钥进行加密. 这样患者 P 即可通过自己公钥, 找到该关联信息, 用自己的私钥进行解密后, 即可查询到自己的相关医疗信息. 这样做的原因是:

a) 包括患者的信息、就诊产生的医疗信息都是存储在医疗机构或者相关机构的数据库中, 这样就不可避免存在数据库管理员倒卖患者相关数据的情况.

b) 通过加密患者的隐私数据和患者与自己就诊记录的关联数据, 即使信息被泄露, 患者的隐私也不会被暴露.

4) 增加、删除、更新、查询类.

a) 进行增加操作时, 所生成的交易单中的交易类型为 “增加”, 交易单内容为要增加的数据, 如病历信息等.

b) 当进行删除操作时, 所生成交易单中的交易单类型为 “删除”, 交易单内容为被 “删除” 的交易单的 ID. 这里所说的 “删除” 并不是将该交易单从区块链中直接删除, 而是在进行查询操作时, 同时检索删除类型的交易单, 排除该删除交易单中所存储交易单 ID 所对应的交易单, 从而实现区块链的 “删除” 操作.

c) 当进行更新操作时, 所生成交易单中的交易单类型为 “更新”, 交易单内容存储更新前的交易单的 ID、更新后的交易单 ID、更新的时间和更新原因

等内容. 这里的 “更新” 并不是用新的交易单去覆盖旧的交易单, 而是更新的时候产生两个交易单, 一个是拥有修改过内容的交易单, 另一个是记录该交易单 ID 和被更新的交易单的 ID. 在查询时, 排除被更新的交易单, 从而实现区块链的 “更新” 操作.

d) 进行查询操作时, 所生成交易单中的交易类型为 “查询”, 交易单内容为查询条件. 验证器在接收到查询交易单时, 同时会检索权限交易单, 判断查询发起者是否拥有相应的查询权限, 以此防止患者信息被泄露. 医疗区块链中的数据主要有区块链相关数据和医疗相关数据, 区块链相关数据在逻辑上构成一条链表, 即通过一个区块记录上一个区块的 ID 来构成一条无法篡改的链, 最后一个区块的安全由参与到区块链维护的节点通过共识机制来保证. 这里的无法篡改指的是, 若某一区块被区块链系统接受后, 则更改该区块之前任意区块中的内容, 都会被检测出来. 而区块链相关数据和医疗相关数据在物理存储上是存储在数据库等相关存储设备上的, 例如在检索医疗信息时, 以医疗信息交易单的 ID 为主键, 查询存储在数据库中该医疗信息所在的交易单, 而不是通过遍历区块链. 得到医疗信息后, 客户端通过交易单 ID、交易单中的公钥和数字签名判断该条医疗数据是否被篡改过.

在检索信息时, 并不需要去完整验证区块链, 而是通过共识机制来保证区块链、交易单和医疗信息的安全. 需要校验的情况主要包括: i) 系统每隔一段时间校验一遍各区块链节点上的区块链, 检测节点是否存在恶意行为或同步出错等问题. 新节点加入后, 获取到区块链后, 也需要对区块链进行校验. ii) 区块链中的区块直接保存的是交易单的 ID, 因此可以将区块链基本信息、区块基本信息和区块中保存的交易单 ID 公开, 外界拿到后, 可以去校验整条区块链, 起到外部监督作用. 法院、公安部门需要强有力的证据时, 需要通过卫生管理部门进行完整验证. 一般情况下, 仅需通过共识过程去验证即可. 其他的一些需要完整验证区块链的情况可以根据需要, 通过卫生管理部门来验证.

5) 权限类. 其交易单中的交易类型为 “权限”, 交易单内容为具体的权限信息, 例如患者可以授权医生在某段时间内查看自己的之前的就诊信息, 过了该段时间后, 医生便不能再查看该患者的之前的就诊信息了.

通过以上五大类型的交易单内容, 可以保证患者的隐私不受到侵犯, 患者就诊中产生的各种医疗数据被篡改后可以被及时发现. 同时, 还可以实现将以前的医疗数据加入到区块链中. 例如, 某医院的数据库中已保存了若干年的医疗数据, 这些数据只要添加交易类型为 “添加”, 数据本身作为交易单内容, 数据的时间戳作为交易单的时间戳, 并加入医院的公钥, 哈希运算后得到交易单 ID, 再用医院的私钥进行数字签名, 就可以产生一条完整的交易单. 这些交易单再发送给验证节点集群, 经过验证后就可以打包进区块, 并加入到区块链当中.

2.1.3 医疗数据存储

医疗区块链的存储需要结合医疗信息系统, 即数字化医院来统筹安排存储措施. 数字化医院是指利用计算机、网络、数据库等信息技术, 有机结合医院业务信息和管理信息, 实现文字、图像、语音、数据、图表等信息数字化采集、存储、阅读、检索的医院信息体系. 其主要组成部分包括、医院信息系统 (Hospital information system, HIS), 临床管理信息系统 (Clinic information system, CIS), 医学影像归档和通信系统 (Picture archiving and communication systems, PACS), 实验室检验信息系统 (Laboratory information system, LIS) 和电子病历 (Electronic medical record, EMR) 等.

HIS 是利用计算机及其网络通信设备和技术, 对医院内外的相关信息进行自动收集、处理、存储、传输和利用, 为临床、教学、科研和管理服务的应用信息系统[23]. CIS 是应用于临床治疗过程中的信息系统, 主要包括医生工作站系统、护士工作站系统、输血管理系统、手术麻醉管系统和临床决策支持系统. PACS 是用来管理医疗图像 (如心电图、脑电图、超声图像) 的系统. LIS 是指利用计算机技术、网络技术、实现临床实验室的信息采集、存储、处理、传输、查询, 并提供分析诊断支持的软件系统. EMR 电子病历是由医疗机构以电子化方式创建、保存和使用的, 重点针对门诊、住院患者 (或保健对象) 临床诊疗和指导干预信息的数据集成系统, 是居民个人在医疗机构就诊过程中产生和被记录的完整、详细的临床信息资源[24]. 它们之间的关系如图 4 所示.

根据卫生部《电子病历基本架构与数据标准 (试行)》[24], 电子病历的基本内容由: 病历概要、门 (急) 诊诊疗记录、住院诊疗记录、健康体检记录、转诊 (院) 记录、法定医学证明及报告、医疗机构信息等七个业务域的临床信息记录构成. 这些记录可以分为两类, 即结构化数据和非结构化数据. 结构化数据即一般的数值或文字性数据, 这些数据通过添加交易单信息, 作为交易单的内容, 形成交易单存储在一般数据库中, 如 Oracle Database、DB2、MySQL 等. 通过收集交易 ID 生成的区块也存储在这些数据库中. 通过数据库分片将海量的结构化数据分布式存储在节点的各个数据库服务器上, 通过数据库备份技术备份这些数据. 非结构化数据指的是医疗过程中产生的图片与音频数据, 如 CT、B 超、心音

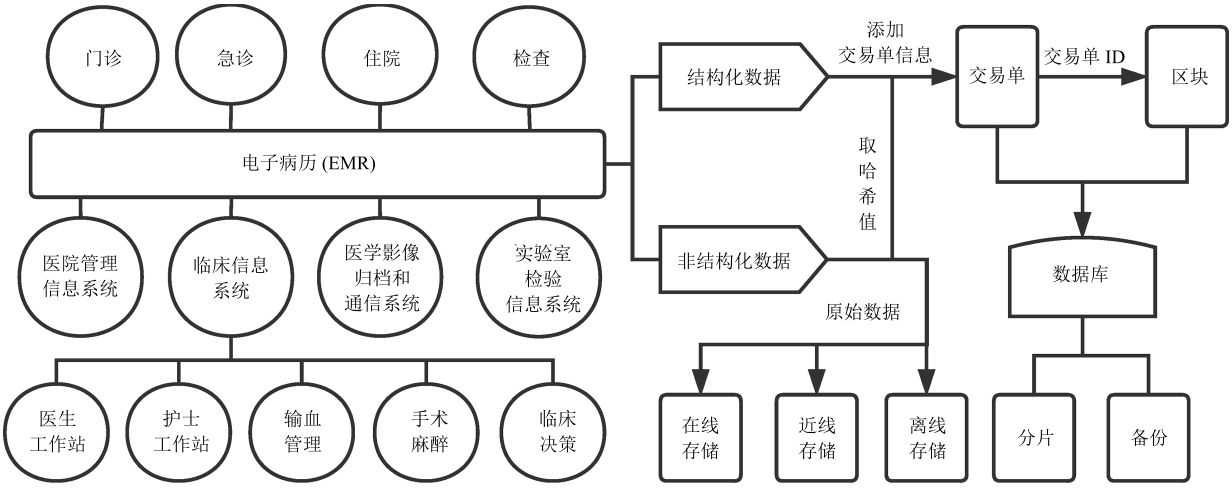


图 4　Medical chain 数据存储架构

Fig. 4　Medical chain data storage architecture

等. 这些数据本身数据量较大, 不适合在共识过程中进行传输, 因此仅将这些数据的哈希值上链, 同时根据医院所处地区与级别, 在市、省一级的卫生管理部门进行备份, 防止这些数据的丢失. 节点在存储这些数据时, 首先取其哈希值, 作为结构化数据存储起来. 其次是对原始数据的存储, 主要分为三级: 在线存储, 近线存储和离线存储. 在线存储和近线存储采用存储区域网络 (Storage area network, SAN) 文件系统, SAN 文件系统是指通过 SAN 存储区域网络将文件数据直接传输到存储设备, 或从存储设备传输到 SAN 文件系统[25]. SAN 网络使用高速光纤作为传输媒介, 利用光纤通道 (Fiber channel, FC) 和小型计算机系统接口 (Small computer system interface, SCSI) 协议来实现高速共享存储. 在存储介质上, 在线存储采用磁盘存储时间较近 (如半年) 的数据, 近线储存采用磁带库 (半年至两年) 存储时间较远的数据. 离线存储使用磁带, 实现以较低的费用长期保存时间久远 (大于两年) 的数据. 与磁盘相比, 磁带存储能够以更低的成本实现存储数据的耐久性与安全性. 因为医院节点存储能力有限, 各节点每隔一段时间就将这段时间内校验过的来自其他节点的数据删除.

2.2　节点管理

一个区块链系统最主要的节点是用来验证交易单 (Transaction) 和区块 (Block) 正确性的验证器 (Validator)、用来生成交易单的交易单生成器和用来生成区块的区块生成器 (Blocker), 它们在共识算法的规范下共同协作运行. 以比特币为例, 其挖矿客户端会接收互联网上的所有交易单, 进行校验后, 计算随机数, 生成区块并广播至全网. 与比特币区块链组件类似, 本区块链系统中的节点也主要有客户端、验证器 (Validator) 和打包器 (Blocker) 三个组件, 如图 5 所示.

2.2.1　共识算法

采用 PBFT 算法作为医疗区块链中的共识算法, 是因为 PBFT 算法是一种适用于联盟链的共识算法, 其优势与优点在于:

1) PBFT 算法不需要像 POW 算法那样靠大量算力来避免 “51 % 攻击” 的发生, 也不用像 POS 算法或 DPOS 算法那样需要靠代币作为衡量投票权的标准, 就可以允许系统中少于 $(n-1)/3$ 个节点出错 (数据丢失、不工作等) 的情况.

2) PBFT 算法作为一种拜占庭容错算法 (Byzantine fault tolerance, BFT) 在系统中存在小于或等于 $(n-1)/3$ 个故障或恶意节点的情况下, 才能保证一次分布式共识过程正常执行[26], 这就要求采用 PBFT 算法的系统中的节点, 在每次共识过程中至少有 $(2n+1)/3$ 个正常节点, 因此这些节点所运行的环境必须是相对安全、稳定的.

3) 医疗区块链是一种联盟链, 参与到医疗区块链中的实体有政府背书, 具有一定公信力, 并由卫生管理部门严格监管, 出现恶意行为的情况远远少于比特币等区块链系统. 同时经过多年的信息化发展, 各医院具有较为完备的网络、服务器和数据库系统. 因此, 现有医疗系统可以提供一个相对安全、稳定的运行环境供 PBFT 算法正常运行. 同时, 因为运行 PBFT 算法的集群中各个节点地位平等, 不存在投票权高低的情况, 避免医疗区块链系统验证交易单或区块链时的中心化. 因此 PBFT 算法十分适合医疗区块链.

目前还没有采用 PBFT 算法的医疗区块链系

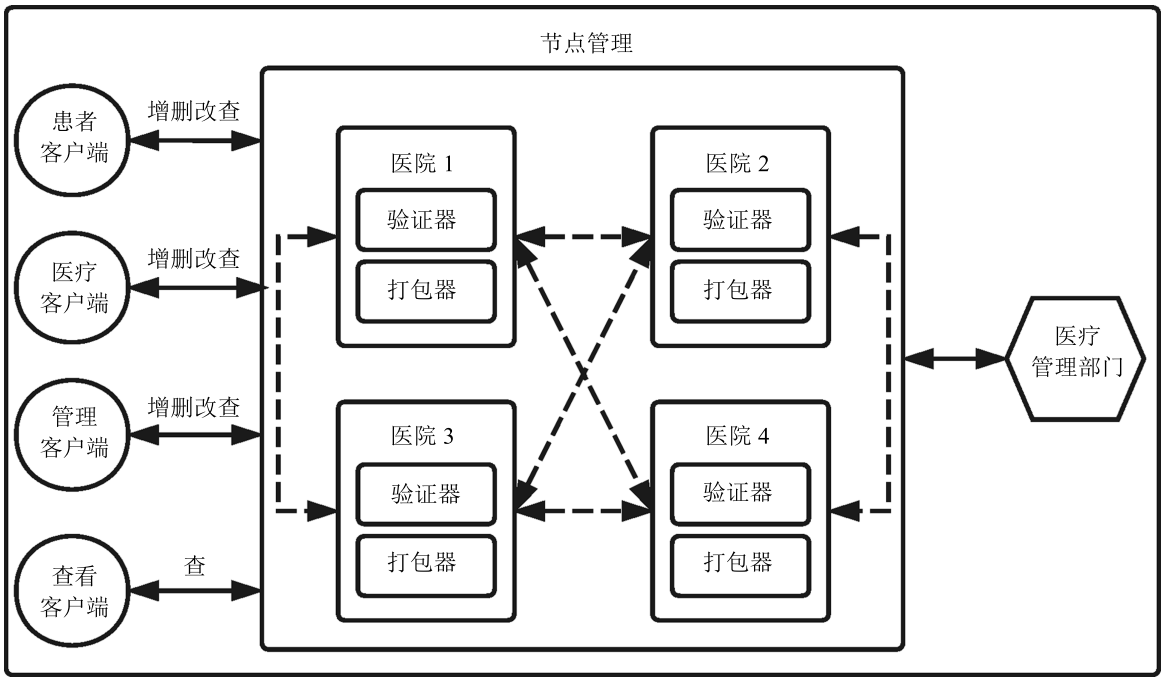


图 5 Medical chain 节点管理

Fig. 5 Medical chain node management

统, 本文创新性地使用了该算法, 提出了一个十分适合医疗领域的联盟式医疗区块链系统.

2.2.2 客户端

客户端是用来生成交易单的组件, 其主要的功能是增删改查. 区别与常见数据库, 如 MySQL, 这里的删除和修改操作并不是直接从区块链里删除相应的交易单, 而是重新生成一个新的交易单, 覆盖掉原来的. 这是因为区块链通过数字签名与一个区块记录上一个区块的 ID, 来保证区块链中已有的内容不会被修改与删除, 并且覆盖过的内容可以被追溯. 客户端主要分为三类:

1) 患者客户端: 患者客户端主要拥有的功能是增删改查. 在增加方面, 患者可以增加授权信息, 授权医生或他人增删改查自己的医疗信息. 在删除方面, 患者可以删除相关的授权信息. 在修改方面, 患者可以修改授权信息, 修改自己的相关信息、如家庭住址等. 在查看方面, 患者可以查询自己以往的医疗记录, 包括被修改或删除前记录.

2) 医生客户端: 医生客户端主要拥有的功能是增删改查. 在增加方面, 医生可以为患者增加医疗记录, 如门诊记录等. 在删除和修改方面, 医生可以在患者授权期限内, 删除有错误的医疗记录, 或者修改患者医疗记录中存在问题的地方. 同时, 医生也可以修改自己的一些信息, 如自己的所就职的医院、科室、专长等信息. 在查询方面, 为保护患者的隐私、医生可以在患者授权的前提下, 查看相关患者的医疗信息, 为患者做出更合适的诊断与治疗. 这些对患者的增加、修改、删除和查看等操作需要在患者的授权情况下, 且一定期限内进行. 未经授权或授权过期情况下, 系统将拒绝执行这些操作, 并根据医生的公钥与数字签名, 将其加入到黑名单中. 若在授权期间, 医生出现恶意行为, 这些恶意行为是极难在区块链中抹除的, 此时患者可以将系统中所存储的就诊信息作为证据, 通过医学专家评审, 追究医生的相关责任.

3) 查询客户端: 查询客户端仅具有查询的功能. 一些机构可能需要查看一下患者的相关信息. 如当医院与患者出现纠纷时, 公安机关或法院可能需要查看该患者的医疗记录, 以此为依据做出合理的判断, 这时候就需要授与公安机关或法院查看权限.

2.2.3 验证器 (Validator)

验证器即为 PBFT 算法中的副本节点. 验证器主要负责接收客户端发来的交易单消息和打包器发来的区块消息并进行验证, 两种消息结构如图 6 所示. 验证器接收到客户端发来的交易单消息或区块消息后, 首先校验交易单消息和交易单本身的数字签名是否正确, 然后校验交易单是否符合规定, 例如一个患者在没有获得另外一个患者的授权情况下是无法获得另外一个患者的医疗记录的. 校验结束后, 主节点验证器则生成预准备消息, 将该交易单消息加入到预准备消息中, 广播给各备份节点, 经过 PBFT 三阶段过程后, 各节点接受该交易单, 并将该

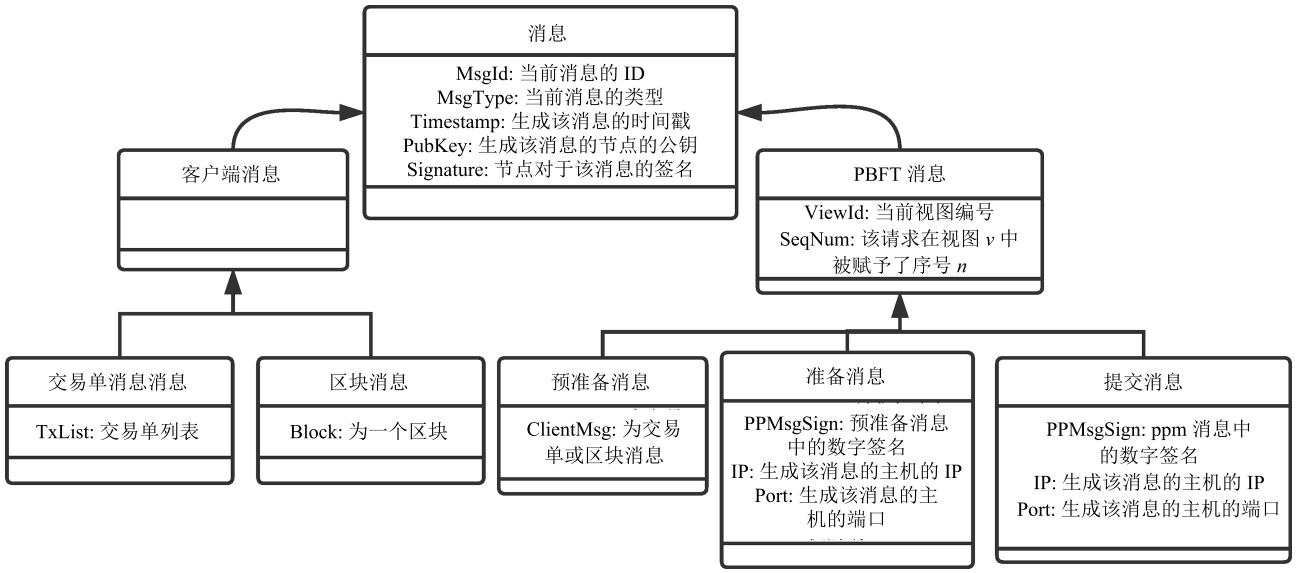


图 6　Medical chain 消息结构

Fig. 6　Medical chain message structure

交易单存入到自己的数据库中. 验证节器集群执行 PBFT 三阶段过程如下:

1) 主节点根据校验过的客户端消息, 分配视图编号 v, 序号 n 给该客户端消息, 生成预准备消息, 生成后广播给各备份节点, 预准备消息的格式如图 6 所示.

2) 各备份节点接收到预准备消息后, 校验预准备消息的数字签名, 视图编号 v 和预准备消息序号 n, 无误后保存, 并根据预准备消息生成准备消息, 准备消息的格式如图 6 所示. 各备份节点生成准备消息后, 广播准备消息至除自己外的所有节点 (即主节点与备份节点).

3) 各节点接收到准备消息后, 校验准备消息的数字签名, 视图编号 v 和预准备消息序号 n, 无误后保存. 当某一副本节点接受了 $(2n+1)/3$ 个视图编号 v, 预准备序号 n 相同, 但是 IP 与端口不同的准备消息后, 根据准备消息生成提交消息, 并广播提交消息至除自己外的所有副本节点. 提交消息的内容如图 6 所示.

4) 各节点接收到提交消息后, 校验提交消息的数字签名, 视图编号 v 和预准备消息序号 n, 无误后保存. 当某一节点接受了 $(2n+1)/3$ 个视图编号 v, 预准备序号 n 相同, 但是 IP 与端口不同的提交消息后, 则接收视图编号为 v, 序号为 n 的预准备消息中所携带的交易单消息或区块消息. 接受后的操作是, 保存交易单或区块, 或执行交易当中所携带的操作.

在一般区块链中, “挖矿” 节点 (即 Medical chain 中的验证器节点) 分为全节点和轻量级节点. 全节点保有一份完整的、最新的区块链及存储在上面的所有数据, 轻量级节点只保留了其中的一部分. 在 Medical chain 中, 验证器节点均为轻量级节点, 每个验证器节点保存所有产生自本医疗机构的数据, 按照第 2.1.3 节中的三级存储方式进行存储. 对于来自其他医疗机构的数据, 每个验证器仅保存半年左右, 并采用在线存储的方式. 同时每个节点完整地保存区块链的信息, 因每个区块仅保存交易单的 ID, 所以整条区块链并不会占用太大的空间. 在这个基础上, 所有的医疗数据均同步至各级医疗管理部门, 以备各节点校验区块链时使用. 采用这种方式在一定程度上依赖于中心化的医疗管理部门, 但大大减少了每个节点的存储压力.

2.2.4　打包器 (Blocker)

打包器主要用来收集交易单的 ID, 生成 Merkle 树[17], 打包成区块, 并发送给验证器 (Validator) 集群, 经过 PBFT 算法三阶段过程后, 加入到区块链里. 打包器生成区块时需要获得当前区块链最后一个区块的 ID, 因此一段时间内仅有一个区块生成. 比特币区块链中每 10 分钟生成一个区块, 区块由 “矿工” 生成, 矿工所做的工作便是挖矿. 所谓的挖矿就是挖矿软件不断生成一个随机数, 与区块头相关内容一起做 SHA-256 哈希运算后得到一个哈希值. 若该哈希值小于一个给定的阈值, 则矿工挖矿成功, 生成一个区块广播至全网, 并得到一笔报酬, 即比特币. 在本区块链系统中, 每生成一个区块都要经过验证器集群校验, 若出现恶意区块, 则验证器集群校验时就会被发现, 并不予接受. 另外, 本区块链部署在各医院, 其服务器和网络环境相对稳定, 不会像比特币那样每时每刻都有主机加入与退出, 并且服务器有专门的管理员管理, 恶意操作的情况相对较

少. 考虑到上面两个情况, 本区块链系统通过以下的方式来生成一个区块:

1) 首先将验证器和打包器部署在一个医院节点上, 一个医院节点包含若干台服务器与数据库, 之所以部署在一起是因为每个医院节点的数据库中保存着验证过的交易单信息, 减少不必要的网络传输. 医院节点如图 5 所示.

2) 每个医院节点在一段时间内校验并接受若干交易单后, 根据 $B = L\,\%\,(N-1)$ 判断当前区块是否由自己来生成. 其中 B 为当前需要生成新区块的节点, L 为当前区块链的长度, % 为取余运算. 考虑到主节点需要接收来自客户端的消息, 因此主节点不参与打包的过程, 以实现各节点任务的负载均衡.

3) 若医院节点根据 2) 中公式检测到当前需要自己生成区块, 则收集一定量的校验过的交易单 ID, 生成 Merkle 树, 打包成区块, 发送给主节点. 经过 PBFT 算法三阶段过程后, 将该区块加入到个节点的区块链链尾. 之所以将交易单 ID 加入到区块中而不是交易单本身, 是因为交易单 ID 是对交易单本身进行哈希运算得来, 是唯一的. 若交易单内容被篡改, 则对交易单中用来生成交易单 ID 的内容进行哈希运算后到的值会与交易单 ID 不同, 因此区块中仅存储交易单 ID 即可.

通过上面三种组件, 可以保证在某一共识过程中出现恶意或故障的节点小于等于 $(n-1)/3$ 时, 共识过程依然可以正常完成. 该正确性的验证过程可参见文献 [26].

2.3 用户管理

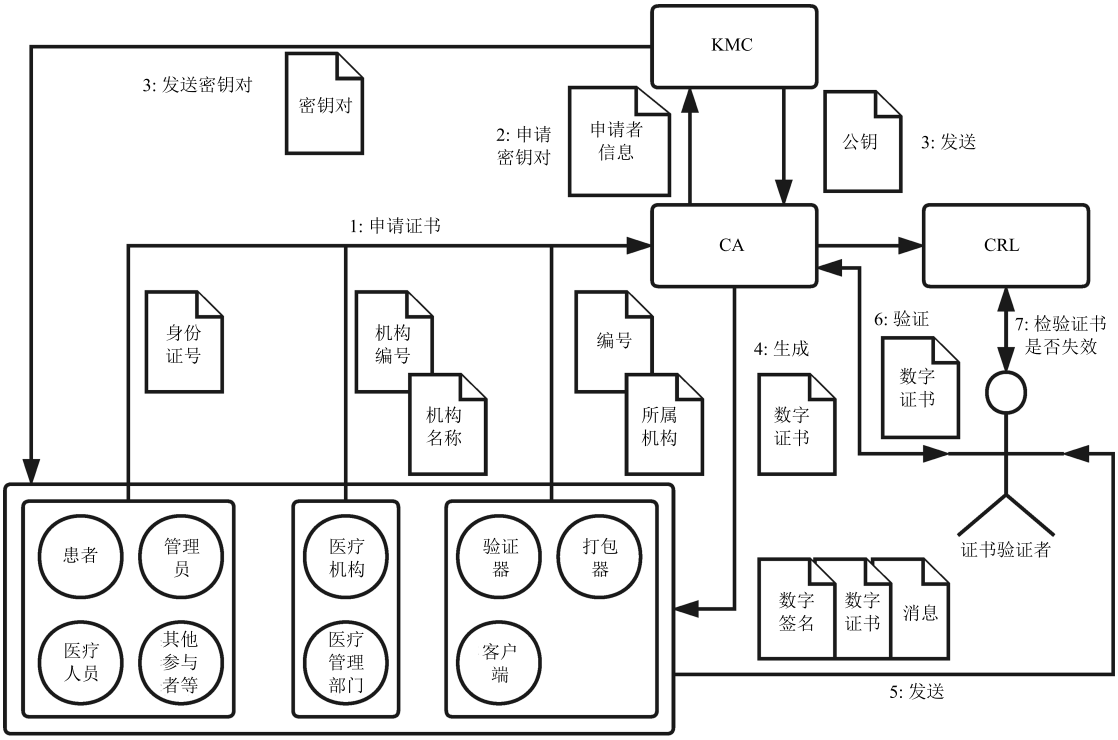
用户管理主要用来管理参与到 Medical chain 中用户的账号、密钥和权限, 是实现身份认证与访问控制的模块.

2.3.1 账户管理

账户管理主要用来管理用户登入、登出、密码找回、公钥绑定等功能. 公钥绑定将用户的登录账户与在密钥与认证架构中用户申请的公钥绑定, 通过实名认证与个人绑定, 用来给用户管理自己的信息.

2.3.2 密钥与认证架构

在 Medical chain 中为了保证区块链系统的机密性、完整性与有效性, 采用了非对称加密、数字签名、公钥基础设施 (Public key infrastructure, PKI) 认证体系等密码学技术. 机密性是指数据传输过程中, 不能被非授权者看到. 完整性是指数据在传输过程中, 不会被篡改. 有效性是指区块链系统的参与者产生的数据不能被否认. 密钥与认证架构提供前面各节所用到的公钥的生成、备份、认证等功能, 其架构如图 7 所示.



证书颁发机构: Certificate authority (CA); 证书吊销列表: Certificate revocation list (CRL); 密钥管理中心: Key management center (KMC).

图 7 Medical chain 密钥与认证架构

Fig. 7 Medical chain key and authentication architecture

1) 所有参与到 Medical chain 系统中的个人、机构与设备均需向证书颁发机构 (Certificate authority, CA) 中心申请属于自己的密钥与证书. 在申请时, 根据申请者的类别, 需要提交相关的资质信息, 如患者的个人身份证件、医疗从业与医院雇佣证明等.

2) CA 中心验证密钥证书申请者的身份资质, 通过后向密钥管理中心 (Key management center, KMC) 提交创建申请者密钥的请求. KMC 接收请求后, 生成密钥对, 添加标志字段等信息之后, 保存在安全的数据库中. 这里的标志字段是指能够标志一个申请者的字段. 如患者可以通过添加证件号来标志, 医疗机构可以通过医疗机构编号来标志.

3) KMC 发送密钥对给申请者, 申请者接收到密钥后, 存储在服务器、个人电脑、手机或其他加密设备上, 并对私钥采用基于口令加密 (Password based encryption, PBE) 方法进行加密, 在使用时再进行解密. PBE 是一种根据口令生成密钥并使用该密钥进行加密和解密的方法[27], 让申请者可以通过口令来安全保存密钥. 同时 KMC 将公钥返回给 CA, 用于生成数字证书.

4) CA 根据申请者的相关信息与 KMC 返回的公钥生成数字证书, 证书可以采用 X.509 标准, 它由 ITU-T 提出, 其详细信息见文献 [28]. 数字证书生成后, 返回给申请者, 申请者通过比对 KMC 返回的密钥与 CA 返回的数字证书判断是否有误, 若出现问题, 则需重新申请.

5) Medical chain 中的参与者在发送消息时均需对消息进行数字签名, 发送给其他参与者的消息包含三部分: 数字证书、数字签名、消息本身. 也可以仅发送公钥而不发送数字证书, 接收者通过公钥和数字签名, 向 CA 中心确认发送者的身份.

6) 当某个参与者接收到其他参与者的消息后, 需要通过 CA、CRL 来验证消息发送者的身份有效性, 证书验证成功后, 接着验证数字签名的有效性. 上述两项验证均成功后, 则该消息便为一条有效的消息.

由于 CA、KMC、CRL 包含了大量安全信息, 因此需要部署在安全可靠的环境下. Medical chain 是一种联盟式的区块链, 带有一定的中心化, 因此密钥与认证体系可以部署在国家卫生和计划生育委员会与各省级卫生和计划生育委员会, 这样既可以保证信息的安全性, 又可以实现负载均衡, 分担各地区的服务器访问压力.

患者的信息是通过患者自己的私钥解密的, 信息加密过程为: 首先患者使用公钥加密除公钥以外的个人敏感信息 (如身份证号、姓名、联系方式、住址), 接着对公钥、非敏感信息和加密后的敏感信息进行数字签名. 当患者的私钥不慎丢失时, 可通过自己的身份证件重新从 KMC 处换取自己的密钥. 当患者的私钥出现安全性问题时, 及时通过个人证件与吊销密钥申请, 重新申请新的密钥, 用旧私钥解密自己的信息后, 再用新申请的公钥加密, 以此来保证安全问题. 另外在一些特殊情况下, 如患者失去自我意识、医疗纠纷的情况时, 医疗机构、公安机关或法院可以通过医疗管理机构向 KMC 来获取, 及时应对. 获取的过程也要生成带有相关操作信息的查询交易单, 作为记录保存.

2.3.3　权限管理

权限管理功能是医疗区块链中十分重要的一部分, 它直接关系到医疗区块链的安全性, 以及是否能妥善保护患者的隐私等方面. 医疗区块链的参与者大致分为三类, 包括: 卫生管理部门、医疗机构和医疗服务接受者. 卫生管理部门主要分为国家级、省级和市级, 在中国则对应于国家卫生和计划生育委员会、省卫生和计划生育委员会和市卫生和计划生育委员会. 医疗机构中主要由医疗人员和管理人员构成, 医疗人员是医疗机构中提供医疗服务的人员, 如医生、护士等, 管理人员是维持医院正常运行的人员, 如人力、财政等. 医疗服务接受者包括患者和患者的家属. 针对这种特点, Medical chain 采用基于角色的访问控制模型 (Role-based access control, RBAC) 来实现权限管理. RBAC 将用户映射到角色, 用户通过角色享有许可. 该模型通过定义不同的角色、角色的继承关系、角色之间的联系以及相应的限制, 动态或静态地规范用户的行为[29]. Medical chain 中的角色分为三大类: 卫生管理部门、医疗机构和医疗服务接受者. 如图 8 所示.

1) 卫生管理部门是 Medical chain 运行的监管部门, 在三类角色中管理权限最高. 其中国家级管理

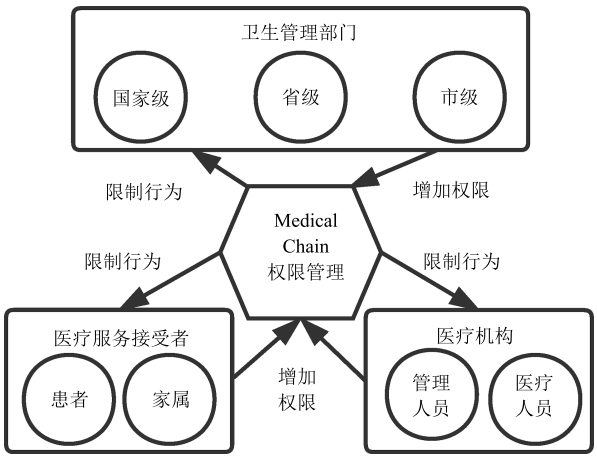


图 8　Medical chain 角色与权限

Fig. 8　Medical chain role and authority

部门拥有最高的管理权限, 可以赋予省级管理部门权限. 省级管理部门可以在自己的权限范围内赋予市级管理部门权限. 这里的卫生管理部门不包括公安、法院等部门, 当它们需要介入到系统中或者需要系统中的信息时, 需要通过卫生管理部门授权.

2) 医疗机构是 Medical chain 中提供医疗服务的部门, 分为医疗人员和管理人员两类角色. 医疗人员角色可以在病人授权的期限内, 查看患者以往在外院或本院的医疗信息. 在就诊过程中, 为患者增删改相关的医疗信息. 在期限以内可以查看自己问诊过的患者的医疗信息. 管理人员是医疗机构中没有直接涉及到医疗过程中的角色, 其可以授权医疗人员查看未在自己处就诊的患者的医疗信息 (不包括隐私信息). 同时可以通过卫生管理部门, 查询其他医疗机构的医疗信息.

3) 医疗服务接受者是接受医疗服务的角色, 主要分为患者和家属两种角色. 患者可以对自己的个人信息和医疗信息进行查询、更改、删除、增加和隐藏, 家属可以在患者的授权之下, 对患者的信息进行以上操作. 当患者不具备行为能力时, 可以通过医疗管理部门获取授权.

与一般授权系统不同的是, Medical chain 中的权限信息存储在区块链中, 极难被篡改. 另外采用智能合约的方式, 执行增删改查操作. 智能合约 (Smart contract) 由 Szabo[30] 于 1995 年提出, 他对智能合约的定义为: 智能合约是一系列以数字形式指定的承诺, 包括各方履行这些承诺的协议. 简单来讲, 智能合约就是发布在区块链上的一段代码, 在某个时间触发合约中的条款时, 代码就会自动执行. 这些代码在 Medical chain 中由卫生管理部门发布, 并在全网公开, 通过这种方式实现透明的权限管理.

2.3.4 隐私与安全

Medical chain 在隐私与安全保护上主要通过人为干预和系统安全模块来保证. 在人为干预方面主要通过: 1) 严格筛选系统管理维护人员. 2) 设置明确的管理维护准则. 3) 对系统管理人员进行权限划分与监督. 通过这些方面保证系统能够安全稳定的运行, 降低恶意行为或其他故障的发生频率. 在系统安全模块方面主要通过: 1) 利用 "密钥与认证架构" 来限制参与人员与标志参与人员的身份. 2) 利用 "权限管理" 模块来保证参与到 Medical chain 中的各方能够在规定权限下, 正常使用其所需功能. 通过卫生管理部门分级, 实现有效的权限监管与分配. 3) 通过对电子病历 (Electronic medical record, EMR) 的分类存储, 实现医疗数据在 Medical chain 中安全高效稳定地存储. 4) 通过 Medical Chain 中的验证器 (Validator) 和打包器 (Blocker) 组件, 采用 PBFT 共识算法, 保证系统可以处理节点的恶意行为与运行故障, 且保证存储在区块链中的医疗数据不被篡改或抵赖.

2.4 一次患者转诊的完整流程

本节通过一个患者 P 从医院 H_A 转诊到医院 H_B 接受医生 D 的门诊治疗的过程, 展示本区块链的基本工作流程. 其中, 患者的私钥与公钥存储在患者的手机当中.

1) 首先患者的 P 信息存储在区块链当中的一个交易单中, 患者的信息有: 身份证号、姓名、性别、年龄、公钥 PUB_KEY_P. 其中身份证号、姓名是经过非对称加密的, 而性别、年龄是没有经过加密的. 医生 D 的信息与患者类似.

2) 患者 P 得有疑难杂症 I, 在得知医院 H_B 在该疾病上的治疗效果较好, 准备从医院 H_A 转诊到医院 H_B.

3) 患者 P 到达医生 D 的诊室, 患者 P 使用手机, 根据公钥 PUB_KEY_P 生成一个查询交易单, 用自己的私钥进行数字签名 SIG_P, 向系统获取含有自己信息的交易单. 系统通过公钥 PUB_KEY_P 和数字签名 SIG_P 在 "密钥与认证架构" 中校验 P 的身份, 验证成功后, 返回含有患者信息的交易单. 该查询交易单会被保存在每个验证节点的本地数据库中, 等待打包器打包进新的区块. 患者通过私钥解密后, 将解密后的信息 (不包括患者的私钥), 发送给医生客户端, 展示在医生的电脑上. 同时根据公钥 PUB_KEY_P 获取患者 P 在医院 H_A 时就诊的交易单 ID, 进而获取患者 P 在医院 H_A 时的就诊信息. 获取患者 P 在医院 H_A 的就诊记录是通过公钥得到实体信息关联交易单, 进而得到就诊记录交易单的过程得到的, 该过程与获取患者信息的过程类似. 医生 D 了解患者 P 在医院 H_A 的就诊记录, 并经过一番问诊, 确认了患者 P 的病情, 在门诊系统中输入患者的目前病情与新的治疗方案, 点击保存. 医生点击保存后, 生成一个 ID 为 TX_1、类型为增加、公钥为医生的公钥 PUB_KEY_D 和医院 H_B 的公钥 PUB_KEY_HB, 数字签名为医生使用自己的私钥 PVT_KEY_D 进行签名的数字签名 SIG_D 和医院 H_B 使用医院私钥 PVT_KEY_HB 进行签名的数字签名 SIG_HB, 内容为 "公钥 = PUB_KEY_P, 病情 = 目前病情, 治疗方案 = 新的治疗方案" 的交易单. 然后生成一个 ID 为 TX_2, 类型为保存、公钥为医生的公钥 PUB_KEY_D 和医院 H_B 的公钥 PUB_KEY_HB, 数字签名为医生使用自己的私钥 PVT_KEY_D 进行签名的数字签名 SIG_D 和医院 H_B 使用医院私钥 PVT_KEY_HB 进行签名的数字签名 SIG_HB, 内容为 "公钥 = PUB_KEY_D, 门诊记录 ID = TX_1" 的交易单. 根据上面两个交易单生成一条交易单消息 TXM, 发送给主验证器节点

VA_P. 其中 $VA_P = v \mod |R|$ 计算得到, v 表示当前的视图编号, $|R|$ 表示存储副本节点的个数. 交易单消息的结构见图 6.

4) 主验证节点 VA_P 接收到这条交易单消息 TXM 后, 验证后进行保存. 验证器集群完成三阶段过程后便接受交易单 TX_1, TX_2.

a) 首先主节点根据 TXM 生成预准备消息 PPM, 并保存到本地数据库中, 然后广播给验证节点 VA_1, VA_2, VA_3.

b) VA_1, VA_2, VA_3 接受主节点发来的预准备消息后, 生成准备消息 PM_1, PM_2, PM_3, 并保存这两条消息到本地数据库, 接着广播 PM_n 给除本节点外的其他所有节点.

c) 各节点接受了 3 条来自不同节点的准备消息 PM 后, 便生成提交消息 CM_P, CM_1, CM_2, CM_3, 并保存这两条消息到本地数据库, 接着广播提交消息 CM_n 到除本节点外的所有节点.

d) 各节点接受至少 3 条来自不同节点的提交消息 CM_n 后, 便接受交易单 TX_1, TX_2. 此时, 交易单共识过程结束.

e) 已知目前区块长度为 1, 最后一个区块 ID 为 1, 当前验证器节点数为 4, 经过一段时间后 (如 10 分钟), 各验证器节点开始根据公式 $B = L\%(N-1)$ 来判断当前区块是否需要自己来生成. 此时 $B = 1\%(4-1) = 1$, 故此时各验证器节点知道下个区块的生成为者 VA_1, VA_1 开始收集一定数量, 且包含 TX_1, TX_2 的交易单 ID, 生成区块 Block, 其中保存前一个区块的 ID. 区块生成后, VA_1 将新生成的区块发送给主验证器节点, 主验证器节点接受到该区块消息后, 验证区块生成者的数字签名判断该区块的生成者是否为 VA_1, 之后生成预准备消息广播至其他备份节点, 这些节点也会根据预准备消息中存储的区块消息, 来验证区块是否为 VA_1 生成. 经过 PBFT 三阶段过程后, 将该区块加入到各节点的区块链中.

f) 此时, 患者的就诊信息已被保存在区块链中. 考虑非法操作的情况:

i) 若患者的门诊信息 TX_1 遭到泄露, 因窃取者并不知道该信息属于哪位患者, 故并不会对患者产生影响. 若关联信息 TX_2 遭到泄露, 窃取者没有患者的私钥, 理论上无法解密其中信息, 因此也不会知道患者的具体医疗新信息, 对患者不会产生影响.

ii) 若患者的就诊信息遭到篡改, 原交易单 ID 与重新计算的交易单 ID 不同, 则证明该交易单遭到了篡改; 若交易单 ID 与重新计算的交易单 ID 相同, 则通过到区块链中查看交易单 ID 是否存在便可知道交易单是否被篡改; 若某一验证器节点中的交易单被恶意删除, 因为其他节点还保存该交易单, 因此还可以从其他节点获取.

以上便是 Medical chain 针对一个患者转诊中进行的一系列操作. 其他的过程与本过程相似, 在遇到新的需求时, 可以以此为依据进行实现. 如患者通过查询交易单获取自己的数据时, 其过程与上文中医生添加新的医疗记录类似.

另外, 在上述过程中, 若医院节点出现恶意行为时, 可以根据如下流程来检测和防止恶意行为的发生:

a) 若主节点 VA_P 出现恶意行为:

i) 主节点丢弃患者信息, 不予进行共识, 则医生就诊客户端隔一段时间后发现患者信息符合规范的情况下, 没有入链, 则认为主节点失效, 更换主节点.

ii) 主节点篡改患者信息, 则主节点发送给各备份节点的预准备消息中的客户端消息数字签名验证失败, 备份节点检测出来后, 进行视图更换过程, 更换主节点.

iii) 若主节点 VA_P 在广播准备消息或提交消息时出现恶意行为时, 参见下面 b) 中的解决方式.

b) 若副本节点 VA_i 出现恶意行为:

i) VA_i 丢弃主节点 VA_P 发送的预准备消息或接受预准备消息后不广播准备消息, 接受来自节点 VA_P 或 VA_j $(i \neq j)$ 准备消息后不广播提交消息, 则不影响 PBFT 算法共识过程, 并在检测出 VA_i 的恶意行为后, 将其排除.

ii) VA_i 发送恶意的准备消息或提交消息. 对于验证节点 VA_P 或 VA_j 在接收到 VA_i 发来的恶意准备消息或提交消息后, 会校验该准备消息. 对比提交消息的视图编号 v, 消息序号 n 和预准备消息签名 $PPMsgSign$ 与自己所保存的预准备消息中的视图编号 v, 消息序号 n 和预准备消息签名 $PPMsgSign$ 是否一致, 若不一致则会丢弃该准备消息或提交消息.

本节涉及到认证的过程, 系统都会根据交易单或区块上的公钥与数字签名验证其所对应的身份, 验证成功后, 根据保存在区块链中的权限判断其行为是否合法.

3 评估

采用对比分析的方式来对比已有医疗区块链系统与本医疗区块链系统, 目前主要的医疗区块链系统有 MDSM[8]、MedRec[10] 与 ModelChain[13], 与现有解决方案对比结果如下:

由表 1 对比可知, 使用 PBFT 作为共识算法的 Medical chain:

表 1 Medical chain 与现有医疗区块链对比

Table 1 Medical chain vs. existing medical blockchain

系统	基于区块链	共识机制	算法类型	支付报酬	需要节点数	算力需求	投票权比重设定
MDSM	是	改进 DPOS	POX	否	121 个	小	是
MedRec	是	POW	POX	是	多	大	否
ModelChain	是	POI	POX	是	多	大	否
Medical chain	是	改进 PBFT	BFT	否	少, 至少 4 个	小	否

1) 相对于使用改进 DPOS 算法 MDSM, 所需要的启动节点个数远远少于 MDSM, 且 MDSM 需要人为设定每个医院是否具有投票的权力与投票在决定最终结果中的比例.

2) 相对于使用 POW 算法的 MedRec, 所需维护区块链系统的节点数远远少于 MedRec, 不需要支付给区块链系统共识参与节点报酬, 且不需要大量算力去维护区块链系统.

3) ModelChain 采用了私有区块链的形式, 其所需节点个数不确定. 但由于工作证明共识机制容易受到 "51 % 攻击", 即节点通过掌握全网超过 51 % 的算力就有能力成功篡改和伪造区块链数据[3], 因此需要较多的节点来 "平均" 算力, 防止这种攻击的发生. 所以相对于使用 POI 算法的 ModelChain, 不需要支付给共识参与节点报酬, 需要的节点数也较少, 并且 POI 算法基于 POW 算法, 因此所需的算力也较大.

因此, 可以看出, 采用 PBFT 共识算法更加适合医疗区块链系统. 其不需要支付报酬、所需启动与运行节点少、后期可扩展、不需要进行 "挖矿" 运算, 算力需求小, 且不需要人为设置投票权比重的特点, 对各医院或其他医疗机构公平, 因此与医疗系统中的需求与特点相契合.

医疗区块链系统在算力需求方面的分析是基于其所采用的共识机制. 算力需求的大小判定基于以下方面:

1) 采用改进 DPOS 共识机制的 MDSM 的区块生成方式是由医疗机构联盟服务器群 (Medical institution federate servers, MIFS) 中的 101 个记账节点轮流生成区块链, 之后医疗机构联盟服务器群中其他 100 个节点, 以及审计联盟服务器群 (Auditing federate servers, AFS) 中的 20 个校验节点会对该区块进行校验. 采用改进 DPOS 共识算法的 MDSM 的区块由 101 个节点的医疗机构联盟服务器群轮流生成, 因此每生成一个区块除生成 Merkle 树外, 仅需 2 次哈希运算用来计算前一个区块的 ID 和对新生成的区块进行数字签名. 采用 PBFT 共识算法的 Medical chain 中的区块由备份节点们轮流生成, 与 MDSM 类似, 每生成一个区块, 除生成 Merkle 树外, 仅需 2 次哈希运算来计算本区块的 ID 与对新生成的区块进行数字签名. 但在生成区块后, MDSM 中的需要固定的接受拥有 100 个节点的医疗机构联盟服务器群和拥有 20 个节点的审计联盟服务器群校验, 而 Medical chain 中新生成的区块接受校验的次数为 $n-1$, n 为当前集群节点的个数, 并且医疗区块链中的节点个数通常不会太大, 因此 Medical chain 对于 MDSM 在算力需求上相对灵活, 总体相差不大.

2) 采用 POW 共识机制的区块链大多为公有链, 其维护者参与维护的动力在于赚取虚拟货币, 以谋求较高的利润. POW 共识机制的原理在于 "矿工" 不断寻找一个随机数, 谁先找到便生成一个区块, 将一定数量的比特币奖励给这个 "矿工", 而其他矿工需要在这个新区块后继续挖矿, 之前的哈希运算变为完全无用的工作. 因此 POW 机制会刺激 "矿工" 们提高自己的算力, 尽快找到这样一个随机数, 而当每个 "矿工" 都不断购买硬件增加自己算力, 新的 "矿工" 不断加入时, 总算力便会飞速增长到海量的程度. 这里算力指的是做哈希运算的能力. MedRec 基于以太坊, 以太坊是一个开放式的区块链平台, 任何人都可以在以太坊中建立和运行区块链应用, 但需要支付一定的 "以太币". 以太坊的平台维护方式与比特币类似, 采用工作量证明的共识方式, 任何人都可以随时加入或退出到以太坊的维护过程中, 因而造成大量的算力白白浪费. 因此, 采用工作量证明共识机制的 MedRec 所需的算力是巨大的. 截止到 2018 年 2 月 1 日以太坊全网总算力已达到 264 THash/s[31], 即以太坊全网每秒可以进行 264×10^{12} 次哈希运算. 以 AMD RX580 显卡为例, 该显卡的算力大约为 25 MHash/s, 即以太坊全网总算力大约相当于 10^8 张 AMD RX580 显卡的总算力, 而每张 AMD RX580 显卡约为 2 600 ~ 3 000 元, 因此仅显卡就需要耗费将近 260 ~ 300 亿元, 还有相关的其他设备、电力等, 因此所造成的资源浪费是十分严重的. 而比特币全网算力截止 2018 年 2 月 1 日已达到 19.59 EHash/s[32−33], 相当于此时以太坊全网算力的 8.9 万倍, 所造成的浪费更为严重.

3) 采用 POI 共识机制的 ModelChain 将隐私

保护在线机器学习[34−35] 与私有区块链网络整合, 使用隐私保护在线机器学习来预测患者重新入院风险的模型, 使用工作量证明 (POW) 作为共识机制, 因此与 MedRec 类似, 其所需的算力依然是巨大的. 并且 ModelChain 还采用了线上机器学习, 这对硬件进一步提出了更高的要求.

4) 采用 PBFT 共识机制的 Medical chain 由备份节点轮流负责生成区块, 不需要进行大量无用的 "挖矿" 运算, 因此 Medical chain 所需的算力较小.

经过多年的医疗信息化的发展, 目前中国医疗信息化已经较为成熟, 各大医院已基本实现各种医疗过程的信息化, 医疗过程更加标准、完备. 但医疗信息化依然面临者诸多问题, 急需解决. 医疗信息化所面临的问题, 与 Medical chain 所提供的解决方式如表 2 所示.

表 2 医疗信息化问题与 Medical chain 应对方式

Table 2 Medical informatization issues and medical chain′s solutions

类型	面临问题	应对方式
隐私与安全	恶意攻击医疗数据保护不可抵赖性	将访问控制信息和每次的操作都加入到区块链中, 并且采用非对称加密机制, 能够保证患者的隐私不会受到威胁, 并且无法抵赖
患者参与度	患者难以获取自己的数据每到一家医院需要重新办理就诊卡, 流程繁琐	患者可以在一个平台上查看自己在各家医院的就诊记录, 并且可以隐藏自己相关的医疗信息来保护自己的隐私
数据访问性	医疗研究人员获取医疗数据困难患者获取自己在各医院的就诊记录困难医生获取患者之前的医疗信息困难	采用匿名形式以及访问控制来保证相关研究人员在得到授权后, 可以访问部分医疗信息, 患者可以查看到自己所有的医疗记录, 医生在患者授权下可以访问患者之前的医疗信息
医疗纠纷	在医院与患者发生医疗纠纷时, 医院与患者提供的证据难以确保其真实性	通过数字签名将相关医疗信息及其修改记录存储在区块链中, 确保证据的不可抵赖与真实性

4 总结与展望

目前, 区块链技术越来越受到国内外研究者们的关注, 并在他们的研究探索下, 一步步走向完善与成熟. 医疗区块链作为区块链技术应用的一种, 在实现医疗数据的安全共享与存储方面有着显著的优势, 是区块链技术应用研究中的一个重要的发展方向. 本文提出的联盟式医疗区块链系统采用 PBFT 共识机制, 能够保证以很小的算力来实现系统安全稳定的运行. 同时该系统以较少的节点启动, 有助于区块链技术在医疗信息方面的应用推广. 联盟式医疗区块链系统 Medical chain 在一致性确认 (共识) 及区块生成加入上仍存在效率不高等问题, 未来的研究将采用拜占庭容错算法与非拜占庭容错算法相结合的共识方式, 来提高系统的运行效率.

References

1 Tsai Wei-Tek, Yu Lian, Wang Rong, Liu Na, Deng En-Yan. Blockchain application development techniques. *Journal of Software*, 2017, **28**(6): 1474−1487
(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)

2 Nakamoto S. Bitcoin: a peer-to-peer electronic cash system [Online], available: https://bitcoin.org/bitcoin.pdf, August 19, 2018

3 Yuan Yong, Wang Fei-Yue. Blockchain: the state of the art and future trends. *Acta Automatica Sinica*, 2016, **42**(4): 481−494
(袁勇, 王飞跃. 区块链技术发展现状与展望. 自动化学报, 2016, **42**(4): 481−494)

4 Lamport L, Shostak R, Pease M. The Byzantine generals problem. *ACM Transactions on Programming Languages and Systems (TOPLAS)*, 1982, **4**(3): 382−401

5 Han Xuan, Liu Ya-Min. Research on the consensus mechanisms of blockchain technology. *Netinfo Security*, 2017, (9): 147−152
(韩璇, 刘亚敏. 区块链技术中的共识机制研究. 信息网络安全, 2017, (9): 147−152)

6 Bitshares.org. Delegated proof-of-stake consensus [Online], available: https://bitshares.org/technology/delegated-proof-of-stake-consensus/, July 10, 2018

7 Castro M, Liskov B. Practical Byzantine fault tolerance and proactive recovery. *ACM Transactions on Computer Systems (TOCS)*, 2002, **20**(4): 398−461

8 Xue Teng-Fei, Fu Qun-Chao, Wang Cong, Wang Xin-Yan. A medical data sharing model via blockchain. *Acta Automatica Sinica*, 2017, **43**(9): 1555−1562
(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)

9 Wood G. Ethereum: a secure decentralised generalised transaction ledger [Online], available: http://gavwood.com/paper.pdf, August 19, 2018

10 Azaria A, Ekblaw A, Vieira T, Lippman A. MedRec: using blockchain for medical data access and permission management. In: Proceedings of the 2nd International Conference on Open and Big Data (OBD). Vienna, Austria: IEEE, 2016. 25−30

11 Ivan D. Moving toward a blockchain-based method for the secure storage of patient records [Online], available: https: //www.healthit.gov/sites/default/files/9-16-drew_ivan_20160804_blockchain_for_healthcare_final.pdf, August 19, 2018

12 Shrier A A, Chang A, Diakun-thibault N, et al. Blockchain and health IT: algorithms, privacy, and data [Online], available: http://www.truevaluemetrics.org/DBpdfs/Technology/Blockchain/1-78-blockchainandhealthitalgorithmsprivacydata_whitepaper.pdf, August 19, 2018

13 Kuo T T, Hsu C N, Ohno-Machado L. ModelChain: decentralized privacy-preserving healthcare predictive modeling framework on private blockchain networks [Online], available: https: //www.healthit.gov/sites/default/files/10-30-ucsd-dbmi-onc-blockchain-challenge.pdf, August 19, 2018

14 Witchey N. Healthcare Transaction Validation Via Blockchain Proof-of-Work, Systems and Methods, WIPO Patent Application WO/2015/175722, November 2015.

15 Rivest R. The MD5 message-digest algorithm. *RFC*, 1992, **473**(10): 492

16 Eastlake D 3rd, Hansen T. US secure hash algorithms (SHA and HMAC-SHA) [Online], available: https: //www.rfc-editor.org/rfc/pdfrfc/rfc4634.txt.pdf, August 19, 2018

17 Merkle R C. A digital signature based on a conventional encryption function. In: Proceedings of the 1987 Conference on the Theory and Applications of Cryptographic Techniques. Heidelberg, Berlin, Germany: Springer, 1987. 369−378

18 Rivest R L, Shamir A, Adleman L M. Cryptographic Communications System and Method, USA Patent 4405829, September 1983.

19 Kravitz D W. Digital Signature Algorithm, USA Patent 5231668, July 1993.

20 Johnson D, Menezes A, Vanstone S. The elliptic curve digital signature algorithm (ECDSA). *International Journal of Information Security*, 2001, **1**(1): 36−63

21 Zhang Ming-De, Liu Wei. *PKI/CA and Digital Certificate Technology*. Beijing: Publishing House of Electronics Industry, 2015. 22−25
(张明德, 刘伟. PKI/CA 与数字证书技术大全. 北京: 电子工业出版社, 2015. 22−25)

22 Josefsson S. The base16, base32, and base64 data encodings [Online], available: https://www.rfc-editor.org/rfc/pdfrfc/rfc4648.txt.pdf, August 19, 2018

23 Dong Jian-Cheng. Analysis of status and causes of hospital information system in China. *Chinese Journal of Hospital Administration*, 2003, **19**(4): 228−230
(董建成. 我国医院信息系统现状及原因分析. 中华医院管理杂志, 2003, **19**(4): 228−230)

24 National Health Commission of the People′s Republic of China. Electronic medical record basic architecture and data standards (Trial) [Online], available: http: //www.moh.gov.cn/mohbgt/s6718/200912/45414.shtml, August 19, 2018

25 Goda K. Storage area network. *Encyclopedia of Database Systems*. Boston, MA, USA: Springer, 2009. 335−336

26 Castro M, Liskov B. A Correctness Proof for a Practical Byzantine-Fault-Tolerant Replication Algorithm. Massachusetts Institute of Technology, Cambridge, MA, USA, 1999.

27 Yuki H [Author], Zhou Zi-Heng [Translator]. *Graphical Cryptography*. Beijing: The People′s Posts and Telecommunications Press, 2016.
(结城浩 [著], 周自恒 [译]. 图解密码技术. 第 2 版. 北京: 人民邮电出版社, 2016.)

28 Galperin S, Malpani A, Adams C, Ankney R, Santesson S, Myers M. X.509 Internet Public Key Infrastructure Online Certificate Status Protocol-OCSP, RFC 6960, 1999.

29 Li Feng-Hua, Su Mang, Shi Guo-Zhen, Ma Jian-Feng. Research status and development trends of access control model. *Acta Electronica Sinica*, 2012, **40**(4): 805−813
(李凤华, 苏锦, 史国振, 马建峰. 访问控制模型研究进展及发展趋势. 电子学报, 2012, **40**(4): 805−813)

30 Szabo N. Formalizing and securing relationships on public networks. *First Monday*, 1997, **2**(9): 1−21

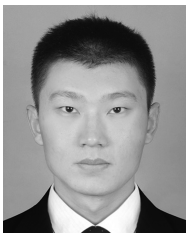
31 Etherscan.io. Ethereum network HashRate growth chart [Online], available: https: //etherscan.io/chart/hashrate, August 19, 2018

32 Bitcoin cash charts [Online], available: https: //charts.bitcoin.com/chart/hash-rate, August 19, 2018

33 Etherscan.io. Ethereum block difficulty growth chart [Online], available: https: //etherscan.io/chart/difficulty, August 19, 2018

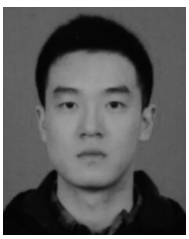
34 Wang S, Jiang X Q, Wu Y, Cui L J, Cheng S, Ohno-Machado L. EXpectation propagation LOgistic REgRession (EXPLORER): distributed privacy-preserving online model learning. *Journal of Biomedical Informatics*, 2013, **46**(3): 480−496

35 Yan F, Sundaram S, Vishwanathan S V N, Qi Y. Distributed autonomous online learning: regrets and intrinsic privacy-preserving properties. *IEEE Transactions on Knowledge and Data Engineering*, 2013, **25**(11): 2483−2493



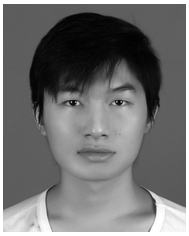
张 超 四川大学计算机学院硕士研究生. 主要研究方向为区块链.
E-mail: zcsd2668@163.com
(**ZHANG Chao** Master student at the College of Computer Science, Sichuan University. His main research interest is blockchain.)

李 强 四川大学计算机学院副教授. 主要研究方向为嵌入式程序设计, 移动云计算, 大数据. 本文通信作者.
E-mail: liq@scu.edu.cn
(**LI Qiang** Associate professor at the College of Computer Science, Sichuan University. His research interest covers embedded program design, mobile cloud computing, and big data. Corresponding author of this paper.)



陈子豪 四川大学计算机学院硕士研究生. 主要研究方向为区块链技术与机器学习. E-mail: chenzihao838@163.com
(**CHEN Zi-Hao** Master student at the College of Computer Science, Sichuan University. His research interest covers blockchain technology and machine learning.)

黎祖睿 四川大学计算机学院硕士研究生. 主要研究方向为区块链.
E-mail: lizurui163@163.com
(**LI Zu-Rui** Master student at the College of Computer Science, Sichuan University. His main research interest is blockchain.)



张 震 四川大学计算机学院硕士研究生. 主要研究方向为区块链.
E-mail: sun_zhangzhen@sina.com
(**ZHANG Zhen** Master student at the College of Computer Science, Sichuan University. His main research interest is blockchain.)