MASTER THESIS XIDIAN UNIVERSITY

西安电子科技大学

硕士学位论文

一种基于区块链的医疗信息隐私保护和共享方案

作者姓名 任延辉

学校导师姓名、职称 樊凯 副教授

企业导师姓名、职称 耿航 高工

申请学位类别 工程硕士

学校代码 10701　　学　号 1501120351
分 类 号 TN918　　密　级 公开

西安电子科技大学

硕士学位论文

一种基于区块链的医疗信息隐私保护和共享方案

作者姓名：任延辉

领　　域：电子与通信工程

学位类别：工程硕士

学校导师姓名、职称：樊凯副教授

企业导师姓名、职称：耿航高工

学　　院：通信工程学院

提交日期：2018 年 6 月

Blockchain-based EHR Privacy Preserving and Sharing Scheme

A thesis submitted to

XIDIAN UNIVERSITY

in partial fulfillment of the requirements

for the degree of Master

in Electronics and Communications Engineering

By

Ren Yanhui

Supervisor: Fan Kai Title:Associate Professor

Supervisor:Geng Hang Title:Senior Engineer

June 2018

西安电子科技大学
学位论文独创性（或创新性）声明

秉承学校严谨的学风和优良的科学道德，本人声明所呈交的论文是我个人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢中所罗列的内容以外，论文中不包含其他人已经发表或撰写过的研究成果；也不包含为获得西安电子科技大学或其它教育机构的学位或证书而使用过的材料。与我一同工作的同事对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

学位论文若有不实之处，本人承担一切法律责任。

本人签名：位延辉　　日期：2018.6.18

西安电子科技大学
关于论文使用授权的说明

本人完全了解西安电子科技大学有关保留和使用学位论文的规定，即：研究生在校攻读学位期间论文工作的知识产权属于西安电子科技大学。学校有权保留送交论文的复印件，允许查阅、借阅论文；学校可以公布论文的全部或部分内容，允许采用影印、缩印或其它复制手段保存论文。同时本人保证，结合学位论文研究成果完成的论文、发明专利等成果，署名单位为西安电子科技大学。

本人签名：位延辉　　导师签名：

日　　期：2018.6.18　　日　　期：2018.6.18

摘要

随着信息与网络技术的发展，数据呈现爆发性增长的态势。数据为国民经济、生活、军事等领域的量化进步提供重要的支撑，可以说数据已经深入我们生活的方方面面。在我们享受数据资源带来的便利的同时，也在承受着隐私泄露的风险。个人信息暴露带来的可能不仅是财务风险，严重时甚至有可能危及个人生命。电子健康档案记录着居民健康管理过程产生的所有信息，包括疾病防治、健康保护、健康促进等。其中有大量个人敏感信息，它的安全保护更是不容忽视。

目前，对于数据隐私和安全保护，访问控制是核心。区块链作为一种分布式架构，账本上记录的信息具有公开可验证、不可篡改等特点，为隐私保护和数据共享提供了新的解决思路，是现在异常火热的研究方向。

本文提出了基于区块链的医疗信息隐私保护和共享方案。针对现有的健康信息管理系统存在的信息碎片化、信息所有权错位等问题，方案利用区块链账本的公开、不可篡改的特性以及用户定制访问策略的方法实现了系统中数据拥有者对数据的完全控制以及共享者对数据的安全获取。针对区块链系统中共识效率的问题，所提方案选择改进的 POS（股权证明）算法提高账本的同步速度，进一步提升共享效率。对基于区块链的医疗信息存储和共享方案从安全性和效率两个方面进行系统分析，结果表明，基于区块链的信息存储和共享系统解决了用户信息碎片化存储的问题，实现健康信息的跨机构安全共享，且相较于现有的文件存储系统具有高效、安全的特点。

提出了基于区块链的权威时间分发和同步方案。在分析传统时间同步协议的基础上，设计了一种基于区块链的时间同步系统，解决了原有系统容错率低、同步效率低等问题。利用时间同步方主动请求的方式，降低时间节点的开销，提升系统的效率。经过分析和仿真说明该方案能够抵抗恶意节点对时间节点的攻击，具有安全、高效的优势。

本文主要研究基于区块链的隐私保护和共享方案。主要提出了医疗信息的隐私保护与共享方案以及权威时间分发和同步方案。利用区块链的分布式架构，以及区块链账本公开、不可篡改的特性，在账本上记录价值信息，前一个方案解决了医疗信息碎片化分布、信息所有权错位、信息细粒度访问控制的问题；后一个方案则是实现了权威时间的分发与同步，解决物联网中伪造时间信息扩散、节点时间不同导致的异常联动工作等问题。

关 键 词：区块链，电子健康档案，隐私保护，共享，访问控制，时间同步

ABSTRACT

With the development of information and network technologies, an increasingly large volume of data has been generated and collected. And data provides important support for quantitative progress in the national economy, life and military affairs. And it can be said that data has penetrated into all aspects of our lives. While enjoying the convenience of data resources, we are also exposed to the risk of privacy leakage. The exposure of personal information may not only cause financial risk, but it may endanger personal life in serious cases. The Electronic Health Record (EHR), as a scientific record of the residents' health management (disease control, health protection, health promotion, etc) process, stores a large amount of personal sensitive information. Thus its safety protection cannot be ignored.

At present, access control is the core of data privacy and security protection. Blockchain is a distributed architecture. The information recorded on the ledger is publicly verifiable and cannot be modified, providing a new solution for privacy protection and data sharing. The idea is now an extremely hot research direction.

This article proposes a blockchain-based privacy preserving and sharing scheme for medical information. In view of information fragmentation and misalignment of information ownership in the existing EHR management system, the characteristic of openness and tamper-resistance of blockchain ledger and user-defined access policy are used to realize the data owner's control and other's secure access to data. Aiming at the problem of consensus efficiency in blockchain, an advanced PoS consensus algorithm improves the synchronization speed of the ledgers and enhances the sharing efficiency. Finally, a systematic analysis of security and efficiency of the blockchain-based medical information storage and sharing system is conducted. The results show that the blockchain-based information storage and sharing system solves the problem of user information fragmentation storage, and achieves cross-institutional sharing of EHR, having advantages of safety and efficiency.

This article proposes an authoritative time distribution and synchronization scheme based on blockchain. Based on the analysis of traditional time synchronization

protocol, this paper designs a time synchronization system based on blockchain, which solves the problems of low error-tolerance rate and low synchronization efficiency of the original system. The method that time synchronization party actively requests is utilized to reduce the overhead of consensus nodes. Finally, the results of analyzing and simulating reflects this scheme can resist the attacks of malicious nodes to the time nodes, which has advantages of security and high efficiency.

This article focuses on blockchain-based privacy preserving and sharing schemes. It mainly proposes privacy protection and sharing schemes for medical information and authoritative time distribution and synchronization schemes. Using the distributed architecture of blockchain and its open and non-derogable nature of blockchain ledger, valuable information is recorded on ledger. The former solution addresses the problems of distribution of medical information fragmentation, misplaced user ownership of information, and fine-grained access to information. The latter solution is to achieve the distribution and synchronization of authoritative time, solving the problem of proliferation of forgery time information in the Internet of Things, and abnormal collaboration caused by different node time.

Keywords: blockchain, EHR, privacy preserving, sharing, access control, time synchronization.

插图索引

表格索引

符号对照表

符号	符号名称
acp	访问控制策略
$Role=(r_1,r_2,r_3,r_4,...r_n)$	身份角色集合
key_{pub},key_{priv}	公私钥对
key_{mas}	主密钥
seq_i	文件索引号
F	文件明文
C	文件密文
t_s,t_e	访问起始时间
$hash$	文件密文哈希
P_i	文件指针
$identity_{pro}$	数据拥有者
R	随机数
loc	文件在存储结构中的位置

缩略语对照表

缩略语	英文全称	中文对照
EHR	Electronic health record	电子健康档案
ABE	Attribute-Based Encryption	属性加密
PBAC	Purpose-Based Access Control	基于目的的访问控制
IBE	Identity-based Encrytion	身份加密
RBAC	Role-Based Access Control	基于角色的访问控制
P2P	Peer-to-Peer	点对点
PoW	Proof of Work	工作量证明
PoS	Proof of Stake	权益证明
DPoS	Delegated Proof of Stake	股份授权证明机制
PBFT	Byzantine fault tolerance	拜占庭容错
NTP	Network Time Protocol	网络时间协议
EID	Electronic Identity	电子身份标识
DoS	Denial of Service	拒绝服务

目录

第一章　绪论

1.1 研究背景和意义

随着社会信息化与网络技术的发展，大数据时代已经来临。数据，开始深入到当今每一个行业和领域，并成为重要的生产要素。人们关于大数据的挖掘和使用，预示着新一波生产率增长和消费者盈余浪潮的到来。确实，各种移动终端、社交媒体和物联网络时刻都在产生数据，通过数据挖掘，人们可以获取以前未知但潜在的有用信息，这些海量数据的使用使得各个领域步入量化进程。各个互联网行业巨头甚至是各国政府都已经开始将发展大数据提升至战略高度。可见，数据，正在成为一种堪比石油与黄金的资源[1]。

而大数据的研究面临着许多的问题，安全与隐私就是一个不得不重视的发展瓶颈。由于各种异构网络的跨域互联，各种数据遍布终端设备、网络和云端，加上黑色经济利益的驱使，数据隐私暴露的风险日益加剧。近些年，国内外网站数据和个人信息泄露事件频发，对政治、经济、社会的影响逐步加深，甚至威胁到个人生命安全。在国外，社交网络巨头 Facebook 被指泄露 5000 万用户信息，用于美国总统大选进行精准推送，甚至有可能影响到了美国大选结果；雅虎两次账户信息泄露涉及约 15 亿的个人账户，致使其被收购计划搁置甚至可能取消。在国内，某知名网站用户简历信息发生泄露，在第三方平台被公开售卖；信息泄露导致电信诈骗案件频发，威胁人民群众财产安全。

信息泄露导致的安全事件层出不穷，由于关乎个体生命健康，医疗领域的隐私保护问题格外受到公众的关注[2-5]。电子健康档案（EHR）作为信息技术与网络技术发展后在医疗领域的必然应用，承载着大量的个人信息。所谓的电子健康档案，不仅仅是将以往的纸质病历信息电子化，同样包含有病人的健康信息、病案记录如检验检查记录、影像结果等[6]。EHR 记录的信息中，包含有姓名、住址、联系方式、工作单位；个人疾病史；甚至是财务状况等，这些敏感数据一旦篡改或被泄露，有可能使个体遭受社会的负面评价，甚至对该个体产生歧视或者侮辱，影响个体的生活或者心理状态[7]。在如今这个越来越重视个人隐私的时代，电子健康档案的隐私保护，成为人们关注且疑虑的问题[3]。

然而，EHR 本身又具有重要意义，不能弃之不用。对国家而言，可以通过健康档案的管理，掌握居民的整体健康水平以及公众的疾病情况，有利于实现公共卫生服务的均等化；对于个人而言，健康档案的存在便于公民得到更好的医疗服务，提高生命质量的同时降低医疗花费；对于医疗机构而言，健康档案便于医务人员快速、准确、

直观的了解病人的病情以及病史，缩短确诊时间，而且大量的电子健康记录为医疗机构进行医学研究提供了基础[8,9]。

因此，如何在保证用户个人隐私的基础上实现安全的数据共享成为关键。健康档案在存储过程中涉及到诸多安全问题，个人数据可能会遭到外来攻击，导致数据被非法用户窃取或者篡改；用户不希望存储服务提供商私下获取自己的个人数据；作为用户的个人健康记录，用户希望能够控制自己的个人数据，在就诊时能够被医务人员安全的读取，共享给他人时能够设置读取的权限，即健康档案的所属权问题。医疗机构对于健康档案的使用必须尊重个人的意愿，个人应该拥有对健康档案的绝对控制权。针对以上问题，本文将根据电子健康档案的特点以及其自身的敏感性，研究基于区块链的电子健康档案隐私保护及共享方案。区块链技术能够传递信任，利用用户自定义的访问策略以及 EHR 的密态存储，将使得信息能够得到有效的保护；同时为了使电子健康档案能够被医疗研究机构使用以及系统的健壮运行，提出了可行的区块链共识机制。实现数据安全高效的存储以及共享。

1.2 国内外研究现状

在数据呈现出爆发性增长的今天，数据的安全存储与隐私保护是人们不得不面临的重要课题，所幸，云计算的发展为人们提供了海量数据存储的方案。云存储能够以较低的开销为用户提供高质量的存储服务，但是单一的云存储服务是无法满足隐私保护与数据共享要求的。首先，数据存储于云端，用户会丧失数据控制权，云服务提供商可以轻易获取服务器中的隐私数据；其次，云存储无法提供细粒度的数据共享服务，即访问控制问题。

目前，关于数据隐私保护与访问控制主要从以下两个方面进行研究：访问控制模型以及基于加密体制的访问控制，其中后者的研究重点是基于属性的加密机制（ABE）。访问控制模型多用于为静态用户分配权限，而在动态系统用户的控制方面存在不足；属性加密在保证数据安全的条件下，具有对数据的细粒度控制、动态访问等优点，但是也有缺陷，例如在执行属性撤销操作时，需要承担巨大的计算开销，导致效率低下。通过研究发现，近两年兴起的区块链技术为数据的隐私保护与共享提供了新的思路，利用区块链技术实现云端电子健康档案的安全存储与共享是本文的主要工作。

1.2.1 国外研究现状

早在 2007 年，Hung 等人就提出了一种基于角色的隐私扩展控制方案。该方案在 EHR 隐私保护的前提下，展示了一个基于角色的访问控制模型与聚合决策层模型

交互的系统，从而实现保护个人健康信息的目的[10]。

Akinyele 等人提供了一种在移动设备上使用属性加密以实现数字信息自我保护的方案。该系统旨在提供细粒度的加密方案，并能够保护记录中的单个项目，使得每个加密项目都可以拥有自己的访问控制策略以实现细粒度的信息共享[11]。

Hur 等人在 2011 年提出关于属性撤销的访问控制方案，该方案描述了用户撤销及属性撤销，但是在属性撤销的执行过程中存在失效的问题，同时授权机构需要面临大量的密钥再生成以及密钥再分发问题，甚至还会涉及重新加密的问题，这对于授权机构来说是庞大的负担[12]。

虽然提出的很多方案都可以实现个人健康档案隐私保护与共享功能，但是都存在着各种各样的问题。访问控制模型虽然可以在某种程度上阻止非法用户获取数据，但是在细粒度控制方面存在缺陷，没能与医疗信息管理系统实现有机的结合；基于属性的加密方案在动态访问以及细粒度访问控制方面具有优势，但是在执行属性撤销时，会产生额外的开销，严重降低系统效率。虽然区块链技术还不够成熟，但是已经有许多基于区块链技术的医疗信息隐私保护研究成果，提供了可行的解决方案。

2016 年，美国学者 Kevin Peterson 等人利用区块链技术来实现医疗信息的共享，提出了一种新颖的共识机制，利用语义的准确性作为证明来挖矿，产生区块[13]。

2016 年，麻省理工媒体实验室和以色列的研究人员利用区块链技术，以以太坊的智能合约为基础，提出了一种分布式的信息管理系统 MedRec，实现了敏感医疗信息的身份认证、加密、追责、共享以及数据的轻量级分布式存储，保证了信息的安全。同时提出了一种概念证明机制，确保系统的正常运行维护[14]。

2016 年，加州大学圣迭戈分校的研究人员提出基于区块链网络的去中心医疗信息隐私保护及评估框架 ModelChain。在私链网络中应用隐私保护在线机器学习，不显示病人健康信息的情况下让各个机构贡献出模型参数并设计了一套信息证明算法来决定在线机器学习的处理顺序，实现尊重隐私的医疗卫生预测建模，增加各机构间可获操作性的框架[15]。

区块链的公开可验证、不可篡改的特性使得它在医疗领域拥有广阔的发展前景，但是目前也存在着匿名性、系统效率等问题需要进一步改善。

1.2.2 国内研究现状

2016 年，国务院下发关于大数据在医疗健康领域应用发展的指导意见，意见指出要关注公民的个人隐私，在没有公民许可的条件下，不能公开或泄露公民健康记录。同时要实现健康医疗信息的共享、挖掘，为改善民生和经济发展服务做出贡献[16]。

2015 年，张艺婷等人将基于目的的访问控制模型与基于身份的加密机制相结合，提出了一种存储密态数据的访问控制方案。其主要方法是根据用户目的、身份以及条

件访问位来构造公钥，而只有通过验证的用户才可以获取对应的私钥解密数据。通过实验仿真，验证方案实现里细粒度的访问控制以及隐私保护的目的[17]。

2016 年，华中科技大学的陈敏等人提出了一种新颖的基于碎片云的隐私保护、数据共享以及入侵检测的医疗系统。利用数论研究组算法加密从病人可穿戴设备中收集到的信息并上传至临近的碎片云，而将远端的医院云当中的数据切割加密加以保护；提出了一种新的信任模型帮助病人相互共享信息；并设计出了一种新的协作入侵检测系统，能够有效的避免远端云当中的数据受到攻击[18]。

2016 年，中南大学的研究人员提出了一种基于区块链技术的 App 架构 HGD（Healthcare Data Gateway）。其中的中心模式显示器能够收集各种类型的医疗信息，中心用途的访问模型确保病人能够完全拥有并控制他们的个人信息，同时在不泄露病人隐私的前提下提升医疗系统的智能性以便于对病人隐私数据的保护与共享[19]。

2017 年，电子科技大学的夏琦等人提出了一种基于区块链的数据共享方案 MeDShare，实现了医疗数据的验证、审计以及共享信息的控制。方案利用智能合约和访问控制机制来追踪数据的行为，将数据的转移过程全部记录下来，一旦发现违规情况发生，将会撤销非法实体的访问权限。最后，通过实验得出 MeDShare 在性能开销方面可以与现有的基于云的前沿方案相提并论[20]。

2017 年，Fu 等人在麻省理工媒体实验室研究人员基于区块链技术的研究基础上，使用了一种更好的加密算法来执行分布式隐私，并用可信性证明机制来代替工作量证明，从而改进了系统，并分析了攻击情景[21]。

1.3 论文主要研究内容

论文首先对数据隐私安全的现状做出简要概述，指明数据隐私保护与共享的重要意义。然后结合研究背景和国内外研究现状，介绍关于隐私保护的研究现状，常见的针对 EHR 的隐私保护与共享方案有两类：访问控制模型和属性加密，之后分析其优劣。鉴于本文方案中的数据是以密态存储在云端，所以会简要介绍数据加密算法。文章基于区块链技术设计方案，因此会详细的介绍区块链技术，包括其结构、账本的组成、以及核心的共识机制，还有区块链技术的瓶颈。

本文提出一种基于区块链的安全高效的数据隐私保护与共享方案；区块链作为一条时序链条，时间是一个重要参数，基于此，本文提出了基于区块链的权威时间分发和同步方案。

1)针对如今用户 EHR 呈现出碎片化分布，信息所有权错位的情况，提出基于区块链的数据保护方案，该方案能够满足用户完整收集并细粒度控制医疗信息的要求，在保证隐私安全的前提下，实现信息的共享，满足用户的就诊需求、医疗机构的学术

研究需要以及政府提供医疗均等化服务的要求。针对区块链技术发展所面临的瓶颈，比如在效率与安全性之间存在矛盾，提出改进的共识机制，在高效的同时能够实现数据的安全共享。

2)研究现有的时间同步协议，针对同步过程中恶意节点攻击导致错误时间在系统中扩散的问题，提出了基于区块链的权威时间分发与同步方案，该方案在时间发布时有中心，在时间同步时无中心，因此可以最小化恶意节点攻击所造成的影响；由同步请求方主动发起服务请求，可以降低系统时间节点的开销，减少网络拥塞；针对性的共识机制可以保证系统的高效同步。该方案作为一条独立的区块链，不仅可以为医疗信息保护系统提供时间戳服务，还可以跨链为更多的智能设备或者物联网系统提供准确高效的时间服务。

1.4 本文章节安排

整篇论文共分为五章，我们在本节对各章节内容安排做简要概述，具体如下：

第一章：简要介绍论文的研究背景及意义，概述当前隐私保护与共享在国内外的研究现状，然后对整个论文的研究内容作简要概括，给出论文章节安排。

第二章：介绍本文用到的一些密码学知识以及区块链的相关内容。首先介绍对称、非对称加密以及哈希函数。其次介绍区块链的定义、区块链的分类，然后简要叙述了共识机制、区块链发展所遇到的问题。

第三章：首先介绍访问控制的发展，其次，针对现在电子健康档案分布碎片化，所属权模糊，以及存储不安全等问题，提出一种基于区块链的医疗信息隐私保护与共享方案。保证在健康档案密态存储的基础上，用户能够实现对每条记录细粒度的访问控制；通过改进区块链的共识算法以及激励机制，实现系统的高效安全运行。同时在本章对方案进行安全性证明，并给出仿真结果，验证方案的可行性以及性能开销。

第四章：首先介绍网络时间协议的相关知识以及常用的时间同步方法，给出简单的分析。针对现有方案中存在的伪造信息易于传染等缺陷，提出一种基于区块链的权威时间分发与同步方案。从系统模型、算法的具体实现以及功能上进行分析，验证方案的可行性，并给出方案的各项性能以及开销的仿真结果。

第五章：对论文研究内容做出总结，对需要改进的地方作出说明，阐明下一步研究重点并给出展望。

第二章　相关知识

本章主要介绍一些方案中用到的密码学基础以及区块链的相关知识。包括对称与非对称加密、哈希函数、区块链的定义、区块链的分类，然后简要叙述共识机制等内容。

2.1 密码学基础

2.1.1　对称加密体制

对称加密是使用较为广泛的密码算法，技术成熟。在对称加密算法中，数据发送方使用密钥通过加密算法将数据明文转换成密文，然后将密文经公共信道发送给接收方，将密钥经安全信道发送给接收方，接收方收到密钥和密文后，使用密钥经加密算法的逆算法解密，从而得到数据明文。对称加密算法的一个典型特征就是加解密密钥相同，因此要求能够安全传递密钥或者收发双方事先知道密钥，这样才可以保证数据的机密性和完整性。该算法的优点是计算量小，速度快，效率高。缺点也很明显，因为加解密使用相同的密钥，安全性无法得到保证，而且在加密大量数据时，密钥管理成为负担，同时对称加密算法无法进行签名验签[38]。如图 2.1，展示了一个基本的对称加密的过程。

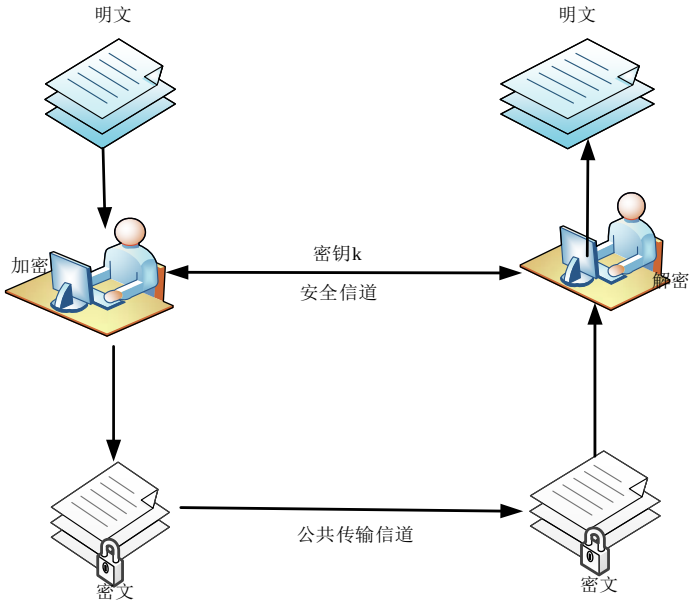

图 2.1 对称加密模型

对称加密算法按照密钥的长度可以分为：流密码与分组密码。流加密中，密钥的长度和明文的长度是一致的，即如果明文的长度为 n bit，那么加密时使用的密钥也应为 n bit，可选择的密钥数量为 2^n 个；分组加密也叫块加密，是将信息流切割为相同长度的分组，分别加密，然后再拼接成密态信息流。对称加密算法共有多种加密模式：ECB、CBC、CFB 等。

2.1.2 公钥加密体制

区别于对称加密体制加解密使用同一个密钥，公钥加密体制拥有一个公私钥对，在加解密时使用不同的密钥。具体来说就是有一个公钥 P_k，对外公开，用于进行加密操作；另一个密钥为私钥 S_k，对外保密，用于解密操作。在加解密的过程中，公私钥是一一对应的，一个密钥加密的数据只能由对应的一个密钥解开。相比于对称密码体制，公钥加密体制的优点明显：用户增加新的数据进行加密只需要掌握一个公私钥对，便于管理；从模型图中可以看出，公钥加密的密钥可以在公共信道传输而不惧被攻击者看到；最重要的是可以利用公钥密码体制进行数字签名，验证身份。缺点则是加解密的速度不够快。常用的公钥加密算法有：RSA 算法、Rabin 算法、椭圆曲线加密算法、ElGamal 公钥加密算法等[38]。

如图 2.2 所示，为一般的公钥加密模型，发送方 Alice 利用公钥加密数据得到密文，经过网络传输后到达接收方 Bob，此时 Bob 利用私钥来解密收到的密文，最后得到原来的明文数据。

图 2.2 公钥加密模型

2.1.3 散列函数

散列函数，又称为哈希函数，杂凑函数。就是将任意长度的输入，通过散列算法，变换成固定长度的输出，输出的值就是散列值[38]。实际上，这种转换是一种压缩映射，即将长度较长的输入值 M 压缩为较短的散列值 h。

数据完整性：输入的数据 M 中有任何一个 bit 的内容发生了变化，都会导致输出发生很大的变化。

单向性：已知输入 M 计算出 h 非常容易。但是相反，已知输出 h 要得到 M，在计算上是不可行的。

无碰撞性：存在一个优良的哈希函数，对于已知的两个消息 M_1 和 M_2 且 $M_1 \neq M_2$，经过该哈希函数后，理论上对于对应的输出值 h_1 和 h_2，有 $h_1 \neq h_2$。

$$h = Hash(M) \tag{2-1}$$

散列函数的这些优良特性使得它在信息安全领域有很好的应用。其中，在文件完整性校验过程中，可以将文件进行哈希，得到对应的散列值，如果攻击者更改了文件，更改后的文件的散列值将迥异于原文件的散列值；同样由于公钥加密的效率问题，可以选择对文件的散列值进行数字签名，因为在统计上对原文件的签名是等价于对原文件的散列值进行签名的，这既提高了签名验签的效率也保证了验证的安全性。常用的散列函数有：SHA-1、SHA-256、SHA-512、MD5 等。

2.2 区块链相关知识

区块链作为比特币的底层应用技术，它的出现始于 2008 年中本聪发表的比特币白皮书[22]。作为一种完全分布式无中心的基础架构，其账本中的数据具有不可篡改、公开可验证的特性，实现了全网任意陌生节点间的信任，因此吸引了众多科技企业甚至是各国政府的关注。

2.2.1 区块链的定义

区块链技术最早出现在比特币中，是一个分布式架构，具有去中心、去信任的特性，是一种通过集体维护分布式账本实现信任的技术方案。从数据的角度来看，区块链中账本上记录的数据经过全网验证，几乎不可能更改。区块链的分布式不仅体现为数据的分布式存储，也体现为数据的分布式记录，由系统参与者共同维护[24]。区块链本身并不是一种新型创新性技术，它是由已经存在的技术通过整合形成的新的数据记录、存储和表达的方式。区块链网络传递的是信任，记录的是价值。在缺少可信任的中心节点与可信通道的条件下，区块链实现了网络的共识，即无需单个信任节点也能实现网络节点的互信[23-27]。

2.2.2 区块链的分类

按照准入机制来划分，区块链可以分为三类：公有链、联盟链与私有链，后两者也被称为广义上的私链[28]。

公有链，顾名思义，是一种完全对外公开的，全世界任何网络用户都可以读取、发送交易且交易能够被有效确认、任何网络用户都能参与竞争实现共识的区块链形式[28]。公有链适用于因为公众参与而需要保证数据公开透明的系统，如公共管理、福利

分配、支付交易等领域。共识指的是确定哪个区块可以被添加到链上以及明确当前状态的过程。由于公有链的开放特性，安全就显得尤为重要，因此公有链的共识多选用需要算力竞争的工作量证明机制或权益证明机制，将激励与加密数字验证相结合。遵循的原则是：每个参与共识的节点所获经济奖励与其对共识所作贡献成正比。系统公开带来的问题还有：节点数量不固定、节点的在线数量无法控制、甚至节点是否恶意也无法保证。

私有链，一个与公有链相对的概念，指不对外开放，仅仅在组织内部使用的区块链形式，比如企业的票据管理、账务审计、供应链管理等[29]。区别于公有链，私有链需要进行身份认证，权限管理。只有合法用户才可以拥有读取权限，读取账本的权限可以进行任意程度的限制。在封闭的系统环境中，节点的数量与状态可控，这保证私有链拥有更高的安全性。因此私有链不需要采用竞争的方式来筛选共识节点，可以采用更加高效环保的方式，如权益证明 PoS、委托权益证明 DPoS、实用拜占庭容错算法 PBFT 等。此外封闭的系统环境也保证了系统不容易被恶意攻击，即使遭受了攻击也可以迅速的追踪溯源。

联盟链，是一种介于公有链与私有链之间的区块链形式，通常是使用在多个成员角色的环境中，比如银行间的支付结算、企业间的物流等。在由众多参与者组成的行业联盟当中预选出部分节点作为记账节点，这些记账节点共同决定了区块的生成与添加。与私有链一样，联盟链系统同样具备身份认证和权限设置的能力，但节点的数量却是动态变化的，这意味着联盟链在节点可扩展的场景中有应用前景。同样的，由于节点的数量和状态可控，因此通常可以选择更加高效低耗的共识机制。

按照对接类型来划分，区块链可以分为单链、侧链和互联链。

单链，即可以单独运行的区块链系统，比如各种数字货币的主链；超级账本项目中使用的联盟链等，目前的区块链系统多属于单链。

侧链，起源于比特币侧链，目的是实现跨区块链信息交互。实际上，随着越来越多的区块链系统的出现，且各个不同的系统各有优缺点，侧链技术能够将不同的链结合起来，将多条链打通，实现优势互补。虽然被叫做侧链，但是它拥有一个区块链系统应该具有的所有功能，是一个完整的区块链系统，实际上，侧链通过主链提供的接口将其锚定在主链上，二者按照一定的协议进行数据通信。在跨链数据交互的过程中，得益于侧链的辅助，主链的功能得到扩展，即主链中无法实现的功能可以在侧链实现；而侧链则增强了自己的可靠性。

由于区块链的应用场景各不相同，因此实现的功能各不相同。但是区块链系统本身又有共性，比如数据的不可篡改、完整性证明以及智能合约等。试想将这些链彼此互联会产生的化学反应，互联链由此诞生。一旦功能不同的区块链互联，且能够实现功能上的互补，将有可能为我们的社会发展带来更高层次的智能化。另外，区块链之

间的互联，不仅仅有助于进行链与链之间的多重验证，还能够改善系统可靠性以及提高系统扩展性。

2.2.3 区块链的结构

如图 2.3 所示，区块链是由区块构成的逐渐延伸的链式数据结构[29]，是一种数据容器，其中包含有区块头与由交易信息构成的区块体。区块头包含有父哈希、Merkle 根、时间戳、难度值以及 Nonce 等内容，父哈希用于将该区块与上一区块钩联，实质上父哈希是将上一个区块的区块头做哈希处理得到；Merkle 根则是一种将区块体中的数据进行有效总结的数据，相当于整个区块体的数字指纹，具有唯一性，在区块当

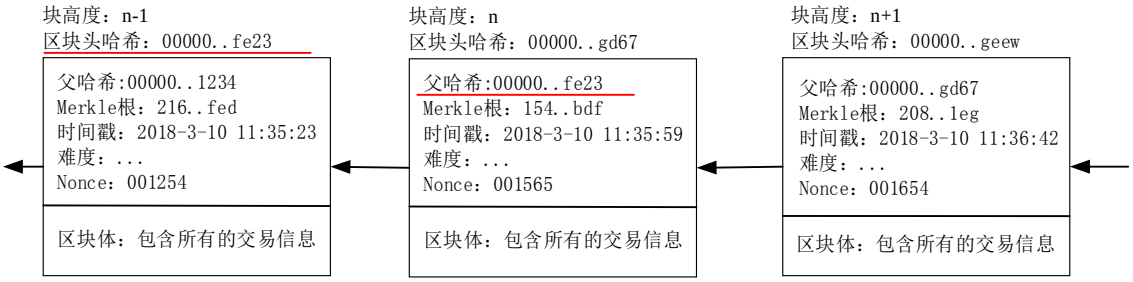

图 2.3 区块链的结构

中，包含有许多的交易，这些交易在逻辑上以一颗树的形式从叶子节点开始递归的进行哈希，直到最后得到根节点的哈希值就是 Merkle 根，用于快速归纳并验证大量数据的完整性；至于后三者则是与挖矿竞争有关，时间戳保证了区块以时序的方式添加，是一种存在性证明，难度值则是用来说明添加区块的难度，Nonce 用来判断是否挖矿成功，挖矿相当于让矿工做算术题，他们利用计算机的算力，经过反复的试错，计算出一个正确的哈希解。而随机数 Nonce 就代表正解，挖矿过程就是寻找这个随机数的过程。

按照结构来看，区块链中的 Merkle 树是一种二叉树形结构，存储着数据的 Hash 值。如下图 2.4 所示，Merkle 树的主要作用是快速递归和校验区块数据的完整性。它会将区块体中的交易信息分别进行哈希运算，结果存储于叶子节点，并向上递归产生新的哈希节点，最终产生一个根值存入区块头中。Merkle 树的使用，极大地提高了区块链的运行效率和扩展性，使得区块头只需包含哈希值而不必封装所有交易信息，这大大降低了哈希运算的开销。

以上结构保证了区块账本信息的不可篡改、不可伪造、不可虚构。

不可篡改性：重写区块链上的区块或者区块上的某些信息几乎是不可能完成的。区块的添加需要经过全网绝大多数的节点确认，如果企图改变区块链的已有区块，那么必然的结果是之后的所有区块全部都将被改变。这就要求试图篡改数据的攻击者必

须具有全网算力的一半以上，这极其难以实现。

不可伪造性：区块链上新区块的添加确认，必须由所有参与共识的节点共同验证该区块上交易记录的正确性。由于区块链上的交易信息，全网所有节点可见。因此，一旦出现某个节点记录的信息与其他节点不同，其他节点就不会承认，该区块也就无法成功上链。

不可虚构性：收到广播交易信息后，区块链中参与记录的节点需要做的是通过历史记录进行验证。以时间为轴，向上追溯，验证数据的真实性，从而建立信任链，保证信息的不可虚构。

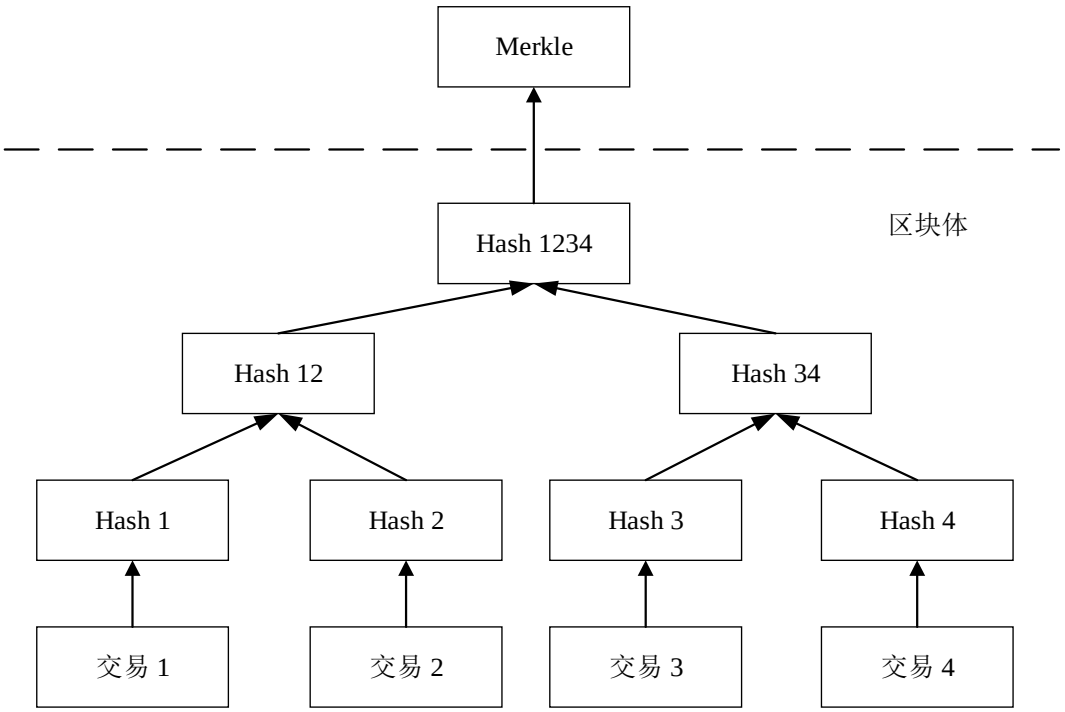

图 2.4 区块中的 Merkle 树

2.2.4 共识机制

区块链作为一种分布式系统，要面临与其他分布式系统一样的难题：如何高效的达成共识。传统的 P2P 系统中都存在一个或者多个中心认证节点，用来进行身份验证、处理交易。但是区块链作为完全去中心的分布式系统，是不存在这个机构的，可以说，中本聪在发布的比特币白皮书中的最主要贡献就是去中心的自发共识机制。自发是指所有节点共同参与竞争以获取激励的过程中，竞争者没有被明确选举，竞争时间不固定，也无法准确预测。即数量不定的独立的网络节点根据系统定义的规则通过异步交互自发竞争实现了共识。常见的共识机制包括工作量证明机制（PoW）、权益证明机制（PoS）、股份授权证明机制（DPoS）以及拜占庭容错算法(PBFT)[30]。

工作量证明机制是比特币中使用的共识机制，也是目前许多数字货币系统中用到的共识算法。简单的说，就是在区块链当中利用算力来进行竞争，竞争获胜方可以获取记账权，添加区块获得系统奖励。以比特币系统为例，每一个区块的区块头中有一个 Nonce，它是一个随机数，只有通过计算哈希得到这个值的节点才可以生成新的区块。由于哈希函数是一个单向函数，因此只能通过不断的试错才能得到正确的结果。工作量证明机制实现了完全的去中心化，节点可以自由进出，同时也利用网络当中的庞大算力保证了自身的安全。但是，在记账权的竞争过程中造成了资源浪费，共识达成的周期也相对较长，不适合商业应用。此外，矿工不一定就是代币持有人，这容易造成利益错位的问题。

权益证明机制最核心的思想是：让在系统中占据大量利益的节点来保障区块链的运行。在 PoS 中，根据每个节点在系统中所占有的权益来分配记账权，从而加速新区块的生成，提高系统运行的效率。也就是说记账权是由系统中占有最高权益的节点获得，而不是拥有最高算力的节点。权益证明极大地降低了工作量证明造成的资源浪费，减少了达成共识的时间，而只有在系统中占有权益的人才能够参与竞争记账权，这解决了工作量证明中的利益错位问题。本文采用的正是 PoS 共识机制，主要原因就是它在安全性与资源利用方面取得了相对的平衡，在尽可能降低资源的条件下实现了一定程度上的安全。

股份授权证明机制，实质是投票机制。系统中的代币持有者投票选出共识节点进行验证和记账。作为一种高效灵活的一致性算法。DPoS 协议利用权益相关者投票的权力，以公平民主的方式解决共识问题，这表明 DPoS 并没有完全去除对于信任的要求。所有的网络参数，从共识时间到块大小和交易大小，都可以通过选举动态调整。共识节点的确定性选取使得交易在一秒内得到确认成为可能。但是 DPoS 仍然存在的缺陷是对于代币的依赖。

拜占庭容错算法，是一种基于消息传递的一致性算法。PBFT 算法改进了原始 BFT 面临的效率问题，降低了算法复杂度。可以用于构建容忍拜占庭容错的高可用系统并在异步环境中具有优异的性能。PBFT 共识机制中，服务器节点分为主节点和从节点两类，主节点轮流担任。主节点收到共识请求后会对共识请求进行排序，从节点则按照排序执行请求。假设系统中共有 n 个节点，该算法在保证活性和安全性的前提下提供了 $(n-1)/3$ 的容错性。

2.2.5 区块链安全问题

在这一部分，我们主要介绍区块链中存在的安全问题，并通过区块在现实中最成功的应用比特币系统来进行描述[31]。

- 挖矿与分叉问题

挖矿，是比特币系统中产生的词汇。在网络中产生大量交易之后，共识节点会执行打包交易、制作区块、算力竞争、全网广播并获得经济奖励的整个过程。而挖矿就代表着算力竞争的过程，实质是穷举随机数，通过不断的试错，得到正确的结果从而获取区块的记账权。新产生的区块会被该节点广播出去以便于其他节点验证。区块头中保存有上一个区块的哈希值，便于沿着链条向上追溯。比特币系统中进行记账权竞争的机制被称作工作量证明。这种机制需要共识节点付出大量的时间以及资源开销，但也保证了付出多的节点能够以较高的概率获取经济奖励。

需要说明的是，区块链的增长不可避免的会带来这样一种问题：同一时刻产生多个区块，即区块链的分叉问题。当出现分叉的时候，系统并不需要判定哪个区块合理。而是根据链的增长情况，总是将后续产生的区块添加到累计工作量证明最大的链上，形成主链。其他的分支链则会被舍弃。这种模式带来的不仅是大量的资源被白白浪费，还带来了安全问题，一旦某一个组织拥有超过全网一半的算力，就能够控制区块链的增长，自己形成一个主链，使原来主链上的交易完全作废，从而为自己谋取私利。

- 时间戳问题

时间戳，用于唯一的标识某一时刻的时间。在区块链系统当中，当共识节点打包区块后会在区块头中添加时间戳，用于记录区块数据的写入时间。之后的每一个区块中的时间戳都是对前一个时间戳的增强，形成一个按时间顺序连接的链条。时间戳形成的时间链为区块链网络增加了一个时间维度，使得数据追溯、重现历史成为可能。同时，时间戳也是一种存在性证明，它能够证明网络中的某些数据在某刻是确实存在的，这种存在性证明为区块链技术应用于公证、知识产权注册等时间敏感领域提供了理论依据。

但是，正因为时间戳的重要性，一旦区块链系统当中的时间戳是不准确的，那么将会带来无法想象的恶劣后果。一旦系统时间被恶意修改、伪造，就有可能使得共识过程无法正常进行。即使区块被成功添加，在利用时间进行追溯验证的时候，也会得到错误的结果。时间戳出现问题，就无法保证链上信息的权威性，因此，可信时间戳的生成也应该引起关注。

以上安全问题仅仅是区块链安全的部分问题，并不代表全部，其他还包括：隐私保护、智能合约漏洞等，这些问题都需要进一步研究解决。

2.3 本章小结

本章主要介绍了一些预备知识，重点介绍了区块链的相关知识，区块链作为一种分布式系统，实现了去中心的互信，区块链账本公开可验、不开篡改的特性为接下来文章中的方案提供了切实可行的基础。但区块链技术在高效低能耗、安全与去中心化

等主要设计追求上存在问题，如果选择去中心化和低能耗，那么在安全性方面有缺陷；如果选择去中心化和安全，能源消耗会使得应用前景受限。因此需要通过分析不同类型区块链的优缺点来为方案选择合适的区块链；之后又介绍区块链系统中常见的几种共识机制，分析他们的优劣，展示不同共识算法应用的场景，为文章下面的工作提供合理的依据。最后，我们分析了区块链当中的一些安全问题，这些安全问题需要我们在文章中解决或者尽量减小他们产生的影响。

第三章　基于区块链的医疗信息隐私保护和共享方案

3.1 方案背景

电子健康档案作为个人医疗记录的信息化载体，不仅可以提高医疗照护的有效性和准确性，降低误诊的概率，而且有助于提升公共医疗服务水平。即 EHR 作为一种数据资源，对电子健康档案进行挖掘分析，不仅利于个人，也有助于提高卫生管理和规划水平，为卫生部门的决策提供依据。但是医疗信息的隐私问题必须引起重视。因此如何在保护个人隐私的基础上实现数据共享成为研究人员关注的重点。

本章针对上述问题，提出了基于区块链技术的医疗信息隐私保护和共享方案。电子健康档案在用户端进行加密，考虑到加密效率问题，采用对称加密算法，密钥存储于用户端，由用户掌握，实现了用户对于个人信息的自主控制；之后将密文信息上传存储到云端，密文存储可以避免云服务提供商侵犯个人隐私；区块链账本上记录着用户健康档案的校验信息，能够保证用户个人信息的完整性以及不被恶意篡改；同时用户可以根据个人需要从多个维度定制访问策略，访问控制策略以明文的形式存放在区块链账本上，只有满足访问策略的用户可以获取解密密钥从而解密文件，此举实现了用户对个人信息的细粒度访问控制；在实现区块链账本的共识方面采用的是 PoS 机制，利用医疗机构如医院或研究所的计算资源来保证区块链的有序生长以及账本的安全；考虑到医疗机构的安全性问题，采用联盟区块链，联盟成员需要经过身份认证；在此基础上，将用户的电子健康档案作为共识激励奖励给成功“挖矿”的联盟成员，以鼓励医疗机构积极维护系统，这种模式既保证了用户安全又可以实现文件的共享，为科学研究提供了便利。

3.2 访问控制技术

在信息安全领域，访问控制技术是安全的核心问题之一。访问控制，顾名思义，即赋予一部分用户访问权限，使其可以得到数据，而限制其他的用户访问数据的能力、范围或者时间，保证信息不被非法用户获取或者篡改。目前，关于访问控制的研究主要集中在访问控制模型以及加密体制两个方面，而关于加密体制最常用的方法是基于属性的加密[39,40]。

访问控制技术经过了半个世纪的发展。在发展的过程当中，Lampson 最先对访问控制进行形式化的描述，并引入主客体以及访问矩阵的概念[41]。之后经过长时间的发展，出现了多种访问控制技术，如表 3.1 所示：

表 3.1 访问控制技术发展表

名称	代表模型	作用
自主访问控制	HRU	由主体自主决定客体的访问，即由主体决定是否将客体的访问权限分发给其他主体
强制访问控制	BLP	保护系统信息的机密性，强制存取控制，多用于军事系统
基于角色的访问控制	RBAC96、RBAC2000	访问权限与角色相关联，用户通过成为特定的角色来获取访问权限

在实际应用中，关于用户对数据的获取是被区分并被限制的，即用户自身必须满足数据的访问权限。因此，当数据上传到云端，安全、高效、可靠的访问控制就显得尤为重要。在云存储兴起之前，用户使用服务器存储数据，并借助服务器验证请求用户的访问权限以实现访问控制。而在云环境中，上述模式并不可靠，因为用户与服务器并不总是处于同一个信用域内。当用户以客体的形式对服务器进行访问时，并不能确认服务器是可信任的角色，一旦服务器被恶意攻击就可能暴露用户的隐私。综上所述，对于隐私保护与共享问题，需要使用合理的技术，才能实现细粒度的访问控制，保证云中的数据不被篡改窃取。目前学术界关于保护数据安全所使用的访问控制手段有以下三种，如表 3.2 所示：

表 3.2 实现访问控制的手段

名称	举例	作用
访问控制规则	访问控制列表、访问控制矩阵	网络端口和准入控制
访问控制模型	RBAC 等	静态分配访问权限
加密机制	ABE 属性加密	保护云中存储的数据和主客体交互

访问控制规则研究重点在于如何制定高效合理的策略以降低访问控制的开销、提高访问效率。传统的访问控制应用于服务器节点，但在云计算成为主流的今天，更关注于对以前模型的改进优化，使得模型能够更好的匹配云环境。通过加密实现访问控制实际是通过控制密钥实现访问控制。云环境下，ABE 成为研究重点。接下来会介绍关于访问控制模型与基于属性的加密体制。

(1)访问控制模型

访问控制模型，即按照一定的访问控制策略构造的访问权限控制模型。访问控制

模型可以为用户分配权限，进而访问云中存储的数据，所以访问控制模型适用于静态分配用户的权限。云环境中的访问控制多是在传统访问控制的基础上进行改进，以适应云计算环境。而根据访问控制功能的不同，研究的侧重点和方式也不同。现有的访问控制模型多是基于角色的访问控制(RBAC)，或者在 RBAC 的基础上使用多种技术进行扩展，如基于任务的访问控制 T-RBAC 是将任务与 RBAC 相结合、基于属性的访问控制将属性与 RBAC 结合。

基于属性的访问控制模型优点在于能够解决存储系统中的细粒度访问控制以及系统扩展问题.用户在接入系统后被分配属性，属性标签用于在数据访问过程中进行身份验证。通过关联数据拥有者、数据请求者、权限和属性集合，描述授权和访问控制约束，使其具有高灵活性和可扩展性。此外，时间也是一个需要考虑的属性约束，时间因素可以限定用户只有在特定的时间段拥有特定的权限，因此在现有的基于角色的访问控制模型中，需要将时间作为一种属性进行研究。

(2)加密机制

加密机制是通过密钥来控制用户的访问，即将数据进行加密操作，只有具有相应密钥的用户才可以解密密文。随着云计算的发展，现在的个人数据多是存储在云端，在云服务器不可信的情况下，用户数据的机密性、完整性以及可用性无法得到保证。用户必须具备自己生成并管理密钥的能力，只有这样，数据拥有者才可以在将数据上传云服务器之前，对数据进行加密，进而通过限制访问者对密钥的获取来实现访问控制。

基于属性的加密自 2005 年开始研究，细化了传统的基于身份的加密体制中身份的概念，将身份定义为一系列属性的集合。图 3.1 是基于属性加密的基本系统模型，在图 3.1 中，我们可以看到共有四个参与者：可信授权中心、云服务器、数据拥有者 data owner 以及用户 user。参与者职责如下：可信授权中心生成主密钥和公开参数，公开系统公钥，同时需要管理用户属性；data owner 制定访问控制策略，使用访问策略与系统公钥加密文件，将加密后的密文与访问策略上传至云端；用户 user 加入系统时需要提交自己的属性集至可信授权中心，可信授权中心会以此为依据利用系统主密钥为用户生成私钥，并下发给用户。user 可以根据文件索引获取密文文件。如果属性满足访问策略，则 user 可以根据属性与访问策略计算出解密密钥从而获取文件明文。云服务器则用于数据存储。

基于属性的加密适合于分布式场景下解密方不定的情况，数据拥有者在加密的时候无需知道解密方的实际身份，只需要解密方能够满足身份属性就可以获取访问权限。ABE 技术利用加密保证数据的安全，利用属性标签实现数据的细粒度访问控制。但是将 ABE 嵌入云存储系统会带来效率与扩展性方面的问题。云端存储的数据是海量的，因此在进行数据加密、用户和密钥管理时需要较高的系统开销，这就导致效率

低下；同时，属性撤销操作后需要对数据重新加密，因此会带来额外的开销，影响系统性能。

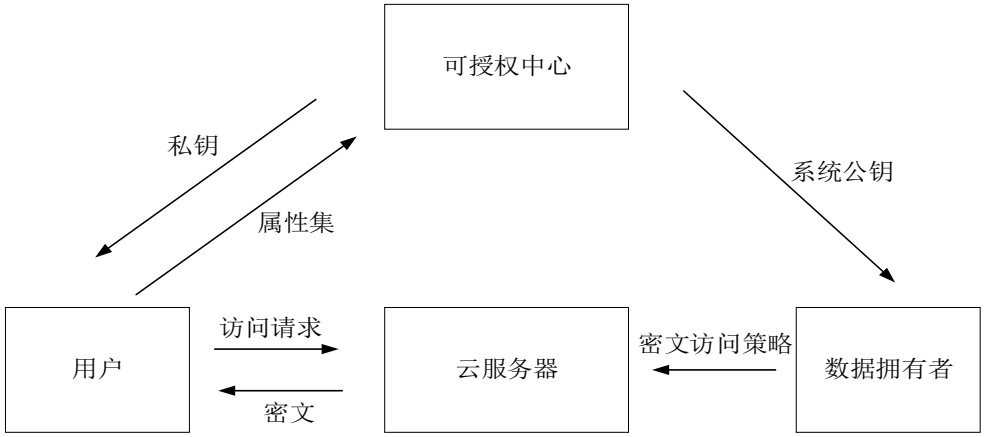

图 3.1 基于属性加密的系统模型

3.3 方案系统模型

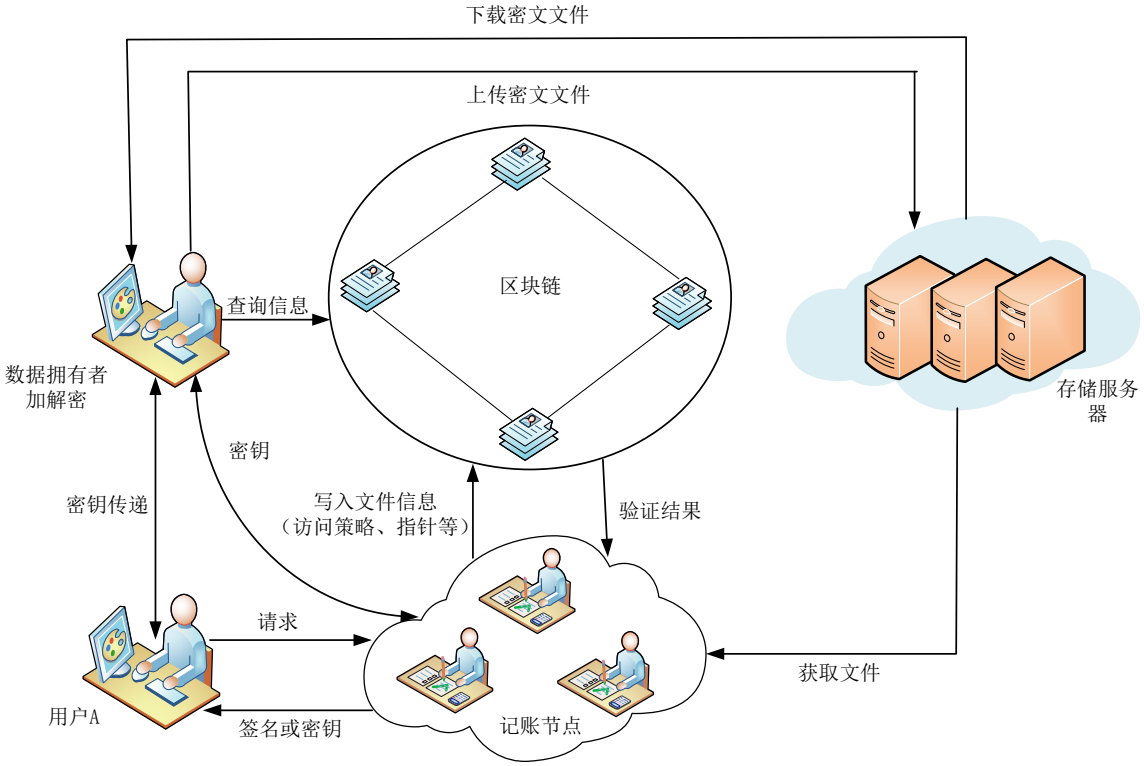

图 3.2 电子健康档案系统模型

下面会具体介绍电子健康档案系统的模型，以及保护 EHR 隐私及共享过程中涉及的算法。如图 3.2 所示，方案中共包括以下五个部分：区块链、云存储服务器、数据拥有者、记账节点以及共享用户。系统中的云是半可信，“诚实但又好奇的”。云存储服务器并不会修改用户的信息，会按照用户的指令或者请求去完成指定的操作，但它有可能去主动试图获取用户的电子健康档案，分析用户存储的数据。方案中的区块链是联盟链，联盟成员经过身份认证，也就是说成员有一定的信任基础。

文件存储时按照一定的层次进行排序，如图 3.3 所示，在叶子节点存储的是具体的用户医疗数据，根节点存储用户信息，中间节点记录的是科室分类信息。这种层次的存储结构便于用户对数据的管理。用户只需掌握主密钥，就可以通过医疗数据所处的位置信息与主密钥一同生成对应的子密钥，便于用户的密钥管理。同样，基于此用户可以细粒度的控制自己的个人健康数据，并在必要时允许自己的医生、朋友或者亲人获取自己的全部健康档案或者健康档案的某个子集。

为了便于生成子密钥，健康档案需要一个唯一的索引号，该索引号根据用户身份 *id* 以及其位置构成。索引号命名规则如下：用户 *id* +中间节点序列号+叶子节点文件号。如图 3.3 所示，阑尾，其索引号 *seq* 为： $id \| 2121$ 。这种存储结构也有利于用户对自己健康数据的归纳整合，用户可以将多个叶子节点代表的同一类型的的数据进行合并，重新加密，并上传到云服务器，更新区块链账本。

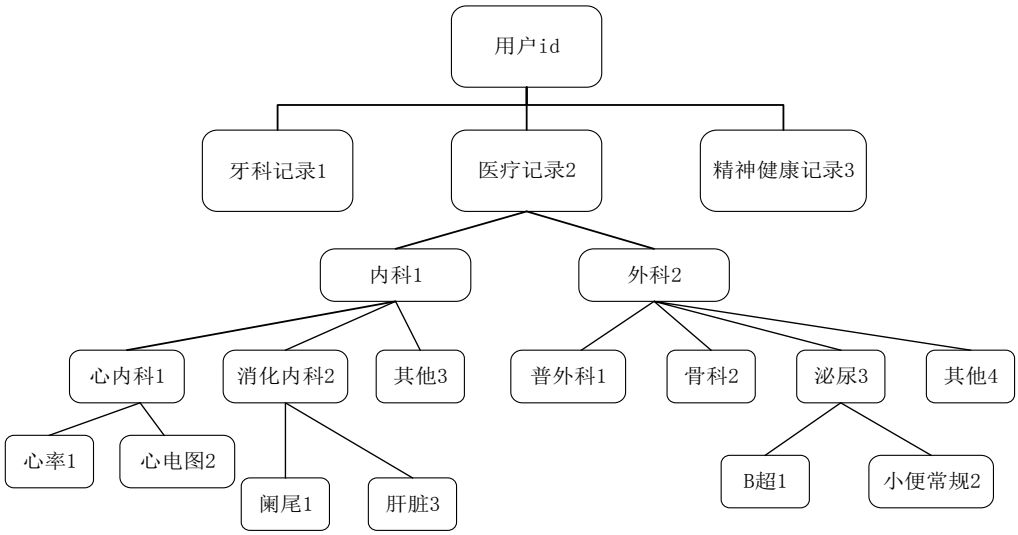

图 3.3 电子健康档案层次存储结构

3.3.1 区块链

方案中的区块链选用联盟链，联盟成员包括普通用户、各种医院或者研究所。相比于普通的用户，大型医疗机构拥有更为突出的计算能力，而计算能力是区块链安全的保证。大型医疗机为了获取更多医疗数据，会进行算力竞争，但是由于医疗机构的

普遍性，最后数据会由众多的机构掌握，避免了中心化的出现。在公开性方面联盟链区别于公有链，联盟链是一个半封闭的环境，可以控制联盟成员的权限。在对成员进行身份认证之后，会为联盟成员颁发身份证书。身份证书中包含有用户的id、用户角色以及公钥key_{pub}。公钥对全网公开，用于进行签名的验证，私钥key_{priv}则用于签名以及生成子密钥。联盟区块链扮演着传统属性加密方案中授权中心的角色。

区块链账本上存储着用户上传的健康档案的相关信息，文件下载地址以及索引号帮助实现信息共享，哈希值可以帮助用户验证信息的完整性，访问控制策略则是既帮助用户实现了对自身信息的细粒度访问控制，又完成了区块链共识的激励。如下面内容所示，区块中包含的信息如下：

-哈希值：用户密文信息的哈希值，虽然无法通过区块账本直接获得用户电子健康档案，但是可以利用账本上的信息验证文件的完整性以及可靠性。

-健康记录的序列号：健康记录对应的唯一索引。

-健康记录的指针：信息的存储地址，用于下载。

-数据拥有者的签名：数字签名，用于判断数据的来源与真实性。

-访问控制策略：以明文形式记录的与健康记录相关的访问控制策略，用于过滤非法访问用户。

以上五个条目构成了一个完整的交易。区块体由一定数量的交易构成，区块由区块头与区块体共同构成，区块头中信息的作用同样不可忽视，下面几项是区块头中包含的内容：

-当前区块哈希：当前区块区块体的哈希值，以默克尔根的形式求出，相当于当前区块的数字指纹。

-父哈希：前一区块头的哈希，用于将当前区块与前一区块连接，从而形成区块链的链式结构。

-时间戳：记录区块生成的时间。每生成一个新的区块，就会在区块头盖上相应的时间戳，保证区块链上的区块以时序进行排列。时间戳的使用有助于降低共识时区块之间产生冲突的可能。

-随机数：成功挖出区块的证明。

-共识节点签名：成功将区块添加到区块链上的共识节点的签名，在进行奖励时用于验证节点的身份。

3.3.2 数据拥有者

数据拥有者，即电子健康档案的主体。在整个系统运转过程中，负责按照图 3.3 所示规则将健康档案进行分类，并根据主密钥key_{mas}以及文件索引号seq_i来构建对应的子密钥；然后将明文文件F_i进行加密，形成密文C_i，并将密文文件C_i上传至云服

务器，获取密文文件的下载地址，文件指针 P_i；在上传文件时，数据拥有者需要计算文件 C_i 的哈希值 $hash_i$，哈希值是文件完整性的证明，用于检测云服务器存储的文件的完整性；数据拥有者可以为每一条 EHR 制定细粒度的访问控制策略 acp_i。完成以上工作之后，将信息 seq_i、P_i、$hash_i$、C_i 制作成一个“交易”，提交后等待共识节点将这些交易打包成区块，添加到区块链上。

访问控制策略可以理解为一个元组，acp_i 是一个通用型的访问控制策略模型，$acp_i = (< identity_{pro}, seq_i >, ro, < t_s, t_e >)$。其中 $identity_{pro}$ 是数据拥有者的身份；seq_i 则是请求访问的健康记录的序列号；ro 是一组角色集合，或者可以说是属性集合，比如某个医院的医生、医疗机构的研究人员或者是某些成功添加区块的共识节点等，代表着数据拥有者允许访问记录 seq_i 的用户应该具有的属性；最后 t_s 和 t_e 则分别代表着允许访问记录的起始和终止时间。从 acp 可以看出，访问控制策略分别从访问者的角色和和时间两个维度来控制用户访问权限，实现细粒度的访问控制。例如：

acp_1： $(< user_2, all >, doc, < t_s, - >)$

acp_2： $(< user_1, seq_3, >, all, < t_s, t_e >)$

acp_3： $(< user_3, default >, miner_{signed}, < t_s, t_e >)$

策略 acp_1 中显示，$user_2$ 是健康数据的所有者，其中在序列号位置填写的是 all，代表 $user_2$ 所有的电子健康档案，doc 则是允许访问者应该具有的角色属性，在时间处填写的是 $< t_s, - >$，整个访问策略可以理解为 $user_2$ 允许属性为 doc 的用户在任意时间访问自己所有的健康档案；策略 acp_2 解释为 $user_1$ 允许所有的用户在时间段 t_s 到 t_e 内访问自己的医疗记录 seq_3；acp_3 是一个比较特殊的访问策略，因为涉及到了区块链的共识激励，代表的含义是 $user_3$ 允许成功达成共识的节点在时间段 t_s 到 t_e 内访问自己的所有健康记录。

3.3.3 记账节点

记账节点，即实现区块链分布式共识的联盟链成员。本方案中，为了尽量保证区块链账本的安全，在共识过程中使用的是改进的权益证明机制，相比于 PoW，PoS 机制的最大优势就是在安全性以及能源消耗方面达成一定的平衡。在使用 PoS 实现共识的过程中，共识节点是通过“代币”的多少来选取，本方案中，电子健康记录就是选取共识节点的“代币”，然后在挖矿的过程中设置不同的挖矿难度，来避免中心化问题。共识过程中涉及计算，拥有较高算力的医疗机构为了获取更多的医疗数据进行研究，会积极地参与挖矿竞争，使得系统能够安全运行。记账节点需要将众多的“交易”打包成区块，进行计算得到正确的随机数后，填写好区块头，然后全网广播，全网节点验证通过后，所有节点更新区块链账本。成功达成共识的节点就可以获取奖励，具体的方式是向普通用户发出请求，用户可以通过区块上的签名验证共识节点的身

份，验证通过后，将自己健康数据的子密钥发送给共识节点，共识节点就可以通过账本上的指针下载密文文件，解密后得到健康记录用以研究。

记账节点除了完成共识之外，还承担着文件共享的职责，在普通用户访问某些数据时，会首先向记账节点发出请求，因为记账节点持有众多的健康记录，记账节点如果拥有数据，在验证过普通用户符合数据要求的访问控制策略，记账节点可以将数据共享给该普通用户，如果记账节点没有请求的数据，记账节点验证符合访问控制策略后，会向普通用户发送带有自身签名的信息 $verifyMes_{signed}$，普通用户向数据拥有者发送 $verifyMes_{signed}$，获取密钥来解密文件。

3.3.4 共享用户

共享用户，系统的合法用户，拥有自己的身份角色集合。向记账节点提出共享请求，提交自己的身份属性完成与记账节点的信息交互。由记账节点完成对自己身份的验证，判断是否符合请求数据的访问控制策略，进而获得密钥，通过区块链账本下载密文解密；或者与数据拥有者进行信息交互，通过验证后获取文件密钥，从而解密数据获得共享数据明文。

3.3.5 云存储服务器

随着云计算的快速发展，云存储服务成为人们存储信息的主要方式。云可以以较低的用户成本为大规模的数据提供存储、管理服务。但是云存储服务同样面临着不容忽视的安全问题，用户将数据提交至云端后，减弱了对数据的控制，因此数据易于被云服务提供商窃取。为了数据的安全，本方案中存储在云端的是数据拥有者加密后的文件密文。

3.4 方案基本思想

本方案基于区块链技术实现，实现用户健康记录的隐私保护与共享。只有符合用户制定的访问控制策略的用户可以共享数据。

3.4.1 方案流程

在本方案中，数据拥有者在用户端加密数据，并定制个性化的访问控制策略，实现细粒度的访问控制；利用区块链账本的的公开性以及不可篡改性，将密文数据的相关信息以及对应的访问控制策略添加到账本中，所有用户可以验证数据的完整性以及可用性；区块链的共识机制实现了账本的安全更新，用户可以将产生的数据添加到链上；当且仅当共享用户的身份属性符合数据拥有者制定的访问控制策略时，才有权限获取文件密钥，解密文件；同样记账节点达成共识后获取数据作为奖励，辅助数据拥

有者实现数据的共享。

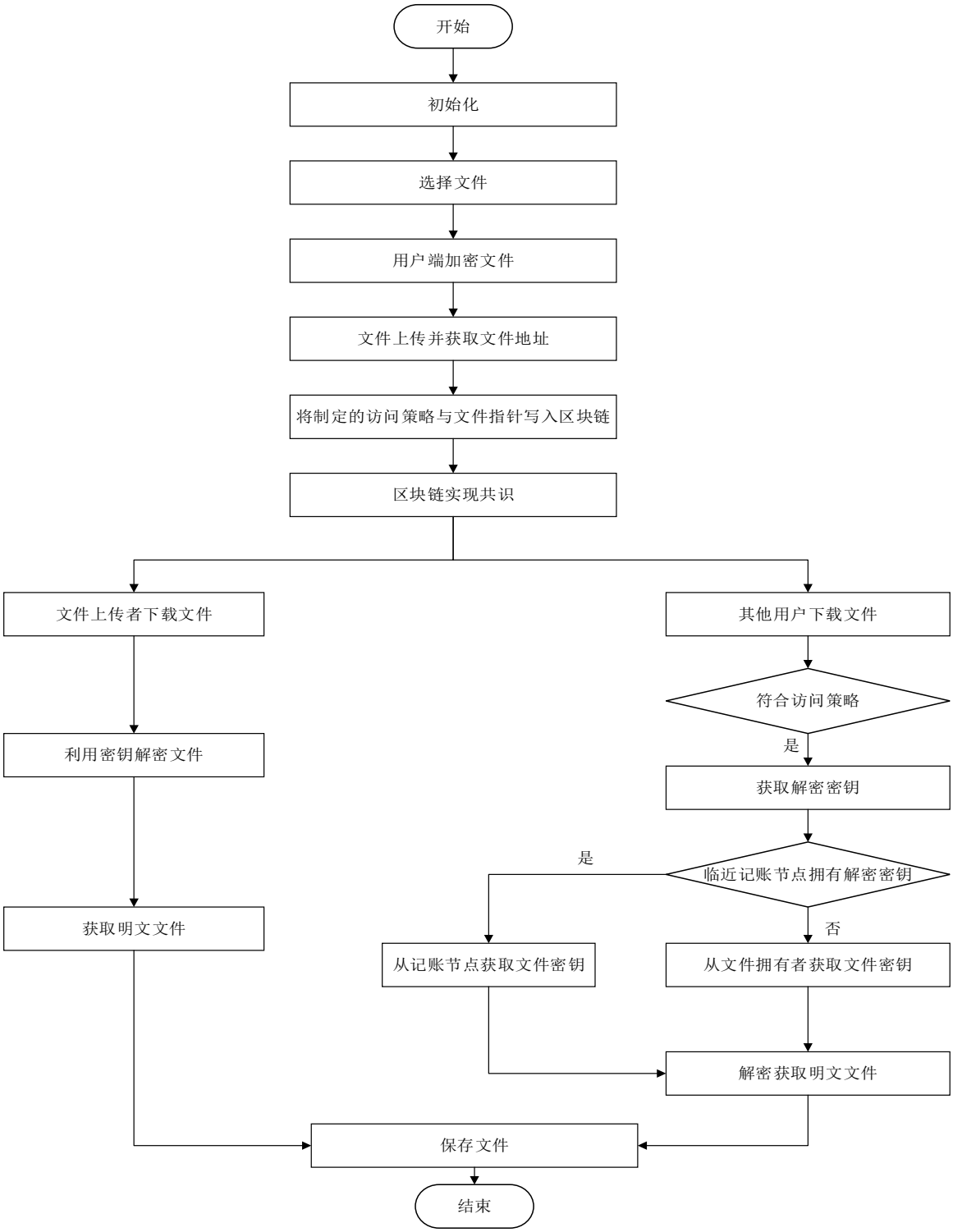

图 3.4 EHR 系统运转流程

本方案具体有以下几个阶段：1)系统初始化、2)文件加密上传、3)区块共识、4)文件下载解密、5)文件共享，系统流程如图 3.4 所示。

3.4.2 算法设计

1)系统初始化

初始化阶段需要区块链验证新用户的身份属性，颁发身份证书，生成用户需要的各项安全参数。身份证书中包含有用户属性集 $Role=(r_1,r_2,r_3,r_4,...r_n)$ 、用户 id 、以及公私钥对 (key_{pub},key_{priv})，公钥全网公开。用户端通过随机数产生器在本地生成随机数 R，将用户私钥与安全随机数作为输入通过哈希函数生成主密钥 $key_{mas}=Hash(key_{priv}\|R)$，并将其保存在用户端。

考虑到系统初始化运行时，所有用户持有的“代币”都相同，此时记账节点的选取是无法根据“代币”量化选取的，此次记账权的获取需要全网所有节点共同参与竞争，随后的共识通过计算用户“代币”数量来选取记账节点。

2)文件加密上传

在对文件进行加密上传之前，首先要生成文件索引、文件加密密钥。其中，在生成文件索引之前用户需要对健康数据分类，判断文件在文件存储结构中所处的位置，以得到位置信息 loc，根据用户 id 及 loc 生成文件索引 $seq=(id\|loc)$；文件加密密钥是根据主密钥以及文件索引生成的，为了保证所有的加密密钥的长度是相同的，使用摘要函数对主密钥以及文件索引进行操作，$key_{seq}=(key_{mas}\|seq)$，这种结构使得用户只需要掌握主密钥以及文件的存储结构就可以计算出所有文件的密钥，方便用户对密钥的管理。

在计算出文件密钥之后，利用对称加密算法对文件进行加密，可采取 AES 或者国密 SM4 算法，加密后的密文 C_{seq} 需要进行哈希处理，得到哈希值 $hash_{seq}$；用户还需要指定文件对应的访问控制策略 acp_{seq}；将密文数据 C_{seq} 上传至云端并得到文件的下载地址，也就是文件指针 P_{seq}；接下来将这些信息封装成一个“交易”，封装后的交易会被记账节点添加到区块链账本上。

3)区块共识

区块共识是通过记账节点实现的，记账节点是通过计算用户的代币持有量选取的。选举出的记账节点会参与竞争记账权，公式(3-1)说明了挖矿成功的规则。左侧为计算出的结果，右侧 nonce 是挖矿的难度，$Target$ 代表目标值，$numOfCoin$ 代表代币数量。由此可知 $Target$ 越大，则挖矿难度越小，反之亦然。算法 3.1 通过伪代码的形式说明了选取共识节点，以及记账的过程。

$$\begin{aligned} proofHash &< nonce \\ &= Target \times numOfCoin \end{aligned} \tag{3-1}$$

算法 3.1：区块共识过程

输入：N，参与竞争记账权的节点

输入：T，选举周期

输入：S，成为公式节点的基准

1：**if** $currentTime()\%T == 0$ **then**

2：　$foreach\ n \in N$

3：　　$numOfCoin \leftarrow getNumOfCoin()$

4：　$consensusNode \leftarrow compare(S)$

5：**end if**

6：$foreach\ i \in consensusNode$

7：　　$t\arg et = setT\arg et(i)$

8：　　$nonce = t\arg et \times numOfCoin$

9：　　$calPofHash()$

10：　**if** $calPofHash() < nonce$ **then**

11：　　$genBlock(trans, hash, timestamp, proofHash, sig)$

12：　　$broadcast()$

13：　**end if**

14：$verifyNewBlock() \leftarrow allUsers$

15：$updateLedger()$

时间 T 用来判断是否开始选取共识节点，系统会统计所有要参与竞争记账权的节点的代币数量，然后通过标准 S 判断并选取记账节点，之后系统会为这些节点分别设置挖矿难度，原则是代币的数量越多，难度越低，保证持有较多代币的节点能够以较大的概率获取记账权。在某个记账节点挖矿成功后，就会向全网广播，系统中的其他节点会验证结果以及签名，最后全网账本同步更新。

4)文件下载解密

文件下载解密分为两种情况：第一种情况为数据拥有者下载文件，利用自己的主密钥计算出子密钥解密文件；第二种情况则是在进行文件共享时，符合访问策略的用户从数据拥有者处获取子密钥，解密文件。

5)文件共享

从算法 3.2 中可以看出，文件共享获取解密密钥有两种途径：一种是从记账节点处获取文件解密密钥；一种是直接从数据拥有者处获取解密密钥。用户在获取共享文件时，需要向记账节点发送请求，提交自己的属性集合 $Role = (r_1, r_2, r_3, r_4, ...r_n)$，记账节点根据请求文件序列号 seq_{share} 验证，搜索区块链账本，验证用户是否符合共享文件

的访问控制策略。假如用户符合，记账节点会检查自己是否拥有该文件，如果拥有，可以直接将文件子密钥发送给请求用户，用户可以下载密文文件并解密，实现文件共享；假如记账节点本地并不拥有序列号为 seq_{share} 的文件，则会发送签名信息 $verifyMes_{signed}$，用户通过该签名信息向数据拥有者表明自己符合访问策略，从数据拥有者获取解密密钥。

算法 3.2：文件共享

输入：*user*，请求共享文件用户

输入：seq_{share}，共享文件索引

输入：*node*，记账节点

输入：*dataOwner*，数据拥有者

1: *send* Re*questToNode*(seq_{share}, *Role*) ← *user*

2: *retrieveLedger*(seq_{share})

3: *getAcp*(seq_{share})

4: *foreach i* ∈ *Role*

5: **if** *verifyRole*(i) == true **then**

6: **break**

7: **else**

8: *refuse*()

9: *flag* = *searchLoaclDatabase*(seq_{share}) ← *node*

10: **if** *flag* == *true* **then**

11: *response*(key_{share}) → *user*

12: **else**

13: *response*(mes_{sign}) → *user*

14: *send* Re*questToDataOwner*(mes_{sign})

15: *verifySig*() ← *dataOwner*

16: *genKey*() ← *dataOwner*

17: *sendKeyToUser*() ← *dataOwner*

18: **end if**

19: *address* = *getPiter*() ← *user*

20: *download*(*address*) ← user

21: *decrypt*(key_{share})

3.5 功能分析

3.5.1 用户对数据的完全控制

现实当中的医疗领域，病人的个人健康数据呈现碎片化分布的特点，对病人在跨地域、跨医疗机构进行就诊时造成不便，也为医护人员为病人提供及时、便捷的医疗服务造成了不利影响。本方案可以帮助病人收集全部的个人健康数据，以密文形式安全的存储在云端，在需要时通过病人掌握的密钥方便快捷的解密得到明文。在使病人获得高水平的医疗服务的同时，实现了病人对数据的完全控制。

3.5.2 细粒度的访问控制

本文提出的方案可以实现对数据的细粒度访问控制，只有满足数据拥有者制定的访问控制策略才是合法的授权用户，并获取密钥解密文件获取明文。文件由数据拥有者完全控制，数据拥有者执行加密上传操作，并自己定制个性化的访问控制策略，数据拥有者并不关心访问者的角色属性信息，也不关心哪些用户会去访问数据，文件是以密文的形式存储在云端，只有获取密钥才可以解密文件，而想要得到密钥，就要通过记账节点对于访问控制策略的验证。

因此，本文方案使得数据拥有者能够在不知道数据访问者的身份的情况下控制数据的权限；避免用户的健康隐私被窃取泄露。数据拥有者在上传数据时，使用对称加密算法，同时制定访问控制策略，只要访问者的角色属性可以满足访问控制策略，都是合法的授权用户，能够对云存储中的密文进行解密，因此更符合用户的需求，有良好的扩展性。

3.5.3 支持多用户共享

基于区块链技术的隐私保护，能够实现文件的安全共享，用户在云端存储的是密态数据。用户将数据上传至云端，在实现文件共享时降低了数据拥有者的负担。数据拥有者掌握制定数据的访问控制策略，决定哪些人可以访问，并且在密文状态下实现文件的共享。

本方案利用区块链技术来实现隐私保护与共享。数据密态存储，保证了数据在云端的安全；利用区块链账本的公开性与不可篡改性，存储密态数据的相关信息，可以验证云端存储信息的可用性以及完整性；区块链账本上存储访问控制策略，使得用户可以便捷的验证访问权限，只要用户符合数据拥有者制定的访问控制策略，就可以获取密钥解密文件，很好的实现了数据的多用户共享。方案使得数据拥有者能够控制自己数据的访问权限，实现了数据的密文共享，提高系统的实用性，更符合我们对云存储的需求。

3.5.4 密钥高效管理

方案中使用对称加密算法对用户健康数据进行加密，随着数据的不断增加，使用的加密密钥也会逐渐增多，为用户管理密钥带来不便。因此，本方案结合主密钥与文件的分类存储结构，来生成文件密钥，用户只需要持有自己的主密钥就可以计算出所有文件对应的解密密钥，实现文件的解密。

3.5.5 访问撤销重加密

区块链账本虽然是不可更改的，但是区块链账本可以通过添加新的信息来实现对已加密文件访问控制策略的撤销。例如数据拥有者 $user_n$ 为索引为 seq_n 的文件制定了访问策略 $acp_n = (< user_n, all >, doc, < t_s, - >)$，在需要对文件重新合并或者撤销的时候，可以重新为索引为 seq_n 的文件制定一条新的访问策略，使得原有的访问策略被撤销，$acp_n = (< user_n, none >, doc, < t_s, - >)$，将新的访问策略添加到账本上后，由于区块是以时序链接到一起的，在查看访问控制策略的时候，总是在检索到最近的访问控制策略后就停止向上追溯。这使得数据拥有者可以实现文件的整理、归并以及访问撤销重加密。

3.6 安全性分析

本方案是基于区块链技术的隐私保护。数据拥有者先将数据加密，之后密文上传至云端，密文可以保证数据的安全性。数据拥有者制定细粒度的访问控制策略，记账节点则会将访问控制策略添加到区块链账本上，使得全网可见、可验证。只有访问者的属性集符合访问控制策略，才可以从数据拥有者或者记账节点处获取密钥，从而解密文件。

3.6.1 数据安全性

本方案能够保证用户上传的数据的安全性，数据在上传前采用对称加密算法加密，上传的文件以密态存储。加密密钥是通过主密钥来生成的，因此只要能确保主密钥的安全，不被泄露，就可以保障用户敏感数据的私密性。用户制定的访问控制策略被添加到区块链账本上，全网可见，不可更改，所有的用户都可以通过验证策略，来判断请求共享用户是否合法。因此，即使攻击者可以通过账本信息得到文件密文，由于无法满足访问控制策略，他们也不能从记账节点或者数据拥有者获得解密密钥，保证了用户上传文档的安全性。如果攻击者伪造用户角色属性，进而获取文件密钥，记账节点可以验证系统颁发的身份证书，身份证书作为合法用户的身份标识，是可以通过证书签名进行验证的。因此，这种伪造攻击也是无法实现的。

本方案在对文件加密时采用对称加密算法，为了保证数据的安全性，每一条健康记录对应一个密钥，而获取密钥就可以解密获取数据明文。因而，密钥的重要性就至关重要，方案中，加密子密钥是通过主密钥产生的，而主密钥只有用户本人掌握，故而密钥是安全的。

3.6.2 区块链安全

方案中的区块链选用的是联盟链，区别于公有链，联盟链具有身份认证以及权限设置的功能。医疗机构这些大型机构，因为具有高出普通用户的算力，一旦参与攻击很可能会影响区块链的安全共识，以及窃取到用户健康隐私，而通过身份认证，可以保证系统用户为合法用户，降低系统风险。

本方案中，区块链安全最重要的就是账本的安全，因为账本中记录着健康数据的地址，访问策略等信息，一旦账本被篡改，或者记账节点伪造区块，就影响到数据的隐私保护与共享。区块链账本是分布式存储的，这种分布式特性保证攻击者无法直接篡改所有的账本；而 PoS 共识机制是需要算力输出正确区块，在众多的记账节点通过计算竞争记账权的情况下，攻击者通过自己的算力竞争过其他所有的记账节点，成本巨大，且攻击成功概率极低。因此，在本方案中区块链的安全也是可以得到有效保证的。

3.7 性能分析

本文基于区块链技术实现健康数据隐私保护与共享。在方案中，数据拥有者对数据加密上传，然后将相关“交易”数据提交，并由记账节点将交易添加到账本中，然后由共享用户下载并解密密文数据。由此可知，区块链系统的主要时间消耗在于用户的加解密操作、区块链账本的同步操作。因为方案中使用的共识机制是 PoS，也就是说区块链的账本同步时间是可控的，在共识难度可以调节的情况下对同步时间进行估算是没有意义的，那么对于系统效率主要的影响就在于加解密是否高效。为了验证方案的可行性，我们在计算机上进行了实验仿真，CPU 处理器参数为 Intel(R) Core(TM) i5-6500 CPU @ 3.20GHz，RAM 为 8GB，Windows7 系统，使用的加密算法是 AES，测算出的数据如图 3.5 所示。

类比于属性加密方案中一次完整的加密流程，本方案中的流程包括数据拥有者对数据的加密，以及区块链账本实现同步。忽略掉网络当中的时延，我们可以得到本方案与属性加密的效率对比。本方案中用户对医疗信息加密使用的是对称加密算法，通过将加密时间与区块链账本时间相加我们可以得到方案所需的时间，假设本方案的共识时间为 30s，则只需要将文件的加密时间加上 30s 即可得到本方案从加密到成功制

定访问策略所消耗的时间。从图 3.6 可以看出，因为方案中进行区块链账本同步消耗的时间较多，因此本方案在对数据量较小的数据加密的时候消耗的时间较多，而当文件比较大时，就可以凸显出本文提出的方案的优越性。一方面，在医疗领域，由于用户的健康数据中包含有大量的图像，如 CT，X 光片等大数据量的文件，因此，本方案还是有较为明显的优势；另一方面，所谓的同步时间，同步的并不是一条健康记录，而是几十甚至是几百条的健康记录，如果从一个区块上记录的所有健康数据来分析，本方案的效率并不低于属性加密方案。

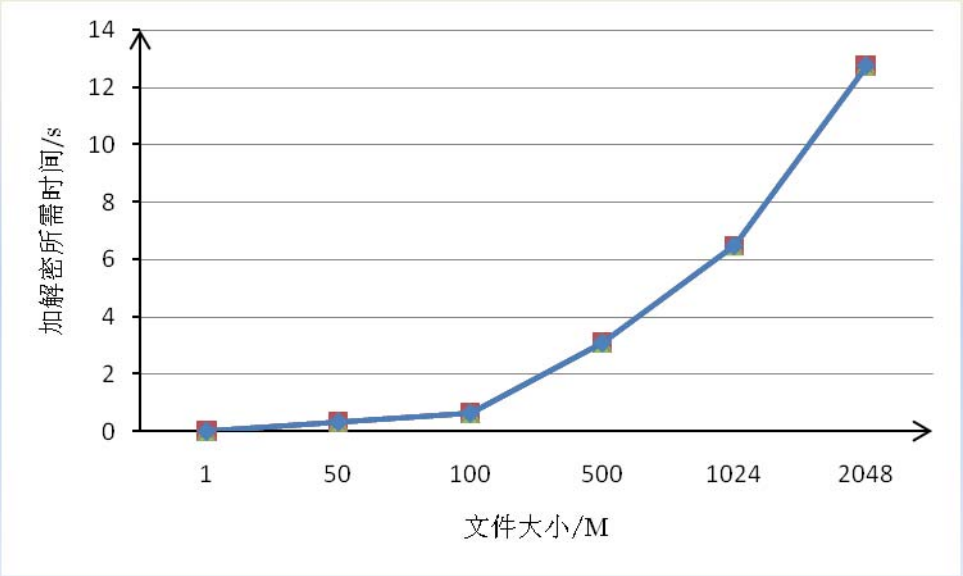

图 3.5 文件加解密时间

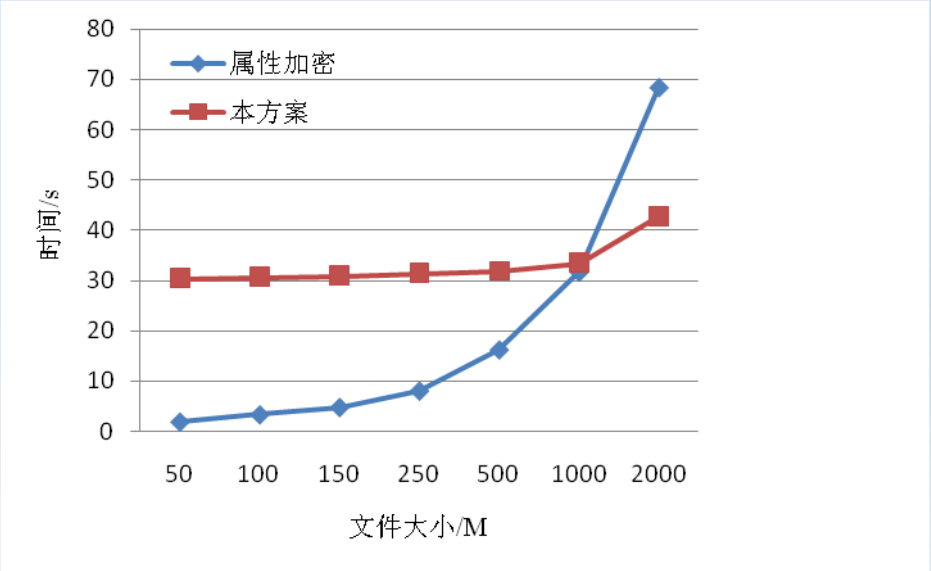

图 3.6 效率对比

在方案中，区块是由区块头与区块体组成。区块头的大小为 80 个字节左右，区块体则是由一定数量的交易组成，一个交易包含着四项信息：文件索引、文件密文哈希值、文件指针以及文件对应的访问控制策略。通过数据分析，得出一个交易的大小在 256 个字节。我们假设一个区块中包含有 20 条交易信息，计算得出一个区块的大小为 $256\times20+80=5200bytes\approx5KB$，在此基础上我们可以测算出网络的吞吐量。

假定一个拥有 1000 名用户的电子健康档案系统，其中每秒产生的交易数量的峰值为 100 名用户在发送交易信息。因此我们可以计算出系统数据的吞吐量，计算过程如下：

- $100(5200)=520000\,\text{bytes/ s}\approx507.8\,\text{Kb/ s}$
- $100(5200\times60)=31200000\,\text{bytes/ min}\approx29.86Mb/\text{min}$
- $100(5200\times60\times60)=1872000000\,\text{bytes/ h}\approx1785.28Mb/\text{h}$
- $100(5200\times60\times60\times24)=44928000000\,\text{bytes/ day}\approx41.84Gb/\text{day}$

根据上面的数据，我们可以推算出系统在一段时间内的数据总量，以及对区块链数据量的增长进行估算，结果如表 4.3 所示：表格中的纵坐标代表的是交易的数量，横坐标则是固定的时间段，表格中的数据是根据网络中的峰值交易数量计算出来的。对于比特币系统，在交易量为 2000 时，一年产生的数据量就能够达到 Pb 级别，而交易量达到 150000 时，就会摧毁整个系统。因此，从数据量来看，我们的方案具有优势。

表 3.3 固定时间内系统数据总量

交易量	1 秒	1 小时	1 天	1 年
1000	4.96Mb	17.43Gb	418.5Gb	149.17Tb
5000	24.79Mb	87.17Gb	2.04Tb	744.6Tb
10000	49.59Mb	174.34Gb	4.09Tb	1.46Pb
50000	247.94Mb	871.72Gb	20.43Tb	7.28Pb
100000	459.90Mb	1.70Tb	40.86Tb	14.56Pb

3.8 本章小结

本章介绍了基于区块链的电子健康档案隐私保护与共享系统，可以使数据拥有者对上传至云端的数据实现细粒度的访问控制，基于区块链账本中记录的访问控制策略能够更好地实现健康数据的共享。本章开头介绍了常见的访问控制方法等内容；之后针对健康数据隐私安全问题，我们提出了本章的方案，并介绍了文件分类存储结构，

系统模型，包括各个部分的功能以及在系统中承担的职责；然后通过流程图介绍整个系统运转的流程，其中以伪代码的形式展示了区块共识以及文件共享的过程。方案实现了文件的细粒度访问控制、多用户共享、密钥管理以及访问策略的撤销等功能。最后通过分析，展示了系统在效率方面的优势，在与属性加密方案的对比中，可以发现在对大文件加密时，方案的效率远远高于属性加密，对于小文件，本文提出的方案，也有不弱于属性加密的效率，最后通过表格评估了系统吞吐量。

第四章　基于区块链的的权威时间分发和同步方案

4.1 方案背景

2005 年，国际电信联盟正式提出物联网的概念，简单的说，物联网就是实现物与物相连的互联网。传统的互联网是以人为主体去获取信息资料，而物联网是通过无线传感器自动读取数据，系统会筛选需要的数据并进行传输从而真正实现无人工干预信息自动捕获目标。在这个过程当中，物联网需要利用众多的分布式单元共同协作完成信息采集、远程监控以及自动化管理等工作，故而时间的准确性和安全性就显得至关重要[32,33]。分布式的物联网中，每一个节点都有自己的本地时钟，这就有可能出现各个节点时间不一致的情况，为了应对这种情况，在系统中会进行时间同步操作。但是物联网本身就是一个相对开放的系统，也就是说恶意节点会对系统发起攻击。此时，这些传感器节点无法辨认接收信息的真假。因而部分节点有可能获取错误时间进行同步，最坏的情况是这些正常的普通节点会传播假的时间信息，导致越来越多的节点受到污染。结果是，少量的恶意节点影响了整个网络的时间同步过程。

区块链是一个时序的链式数据结构，也就是说，区块的添加确认是以时间顺序作为基准的，这种模式既可以避免添加区块时发生区块碰撞，也可以保证系统的安全性。类似于分布式物联网系统，基于区块链的电子健康档案系统也需要高安全性的时间参数，故而为电子健康系统提供安全的时间戳服务必不可少。在本章，我们提出了一种基于区块链的权威时间分发与同步方案。该方案采用联盟链的形式，对时间节点进行身份认证，保证了时间同步过程的安全性，联盟链也可以满足节点扩容方面的需求。在设备同步时，采用设备主动发送请求的方式，既降低系统开销，避免网络拥塞，又提高系统同步效率。我们可以将该系统视为一条锚定在电子健康档案系统上并为其提供时间戳服务的侧链。

4.2 时间同步协议

传统的时间同步方案大体上可以分为两类：基于树的时间同步协议、分布式的时间同步协议。基于树的时间同步协议中，时间源作为树的根节点存在，从根节点一层一层的向下传递时间，以达到时间同步的目的，但是此种方法存在的一个显著问题是当某一个节点发生错误的时候，错误会被传递下去，导致错误的节点以指数形式增长，即健壮性有缺陷。分布式的时间同步协议迥异于树形结构的单向通信，是通过节点向四周的邻居节点广播时间从而实现同步，因此在健壮性方面更占优。类似的，在移除或者增加节点的时候不需要重构树的结构，因此在扩展性方面表现良好[34]。

4.2.1 分布式时间协议

网络时间协议（NTP），用于分布式系统中时间服务器与客户端之间进行时间同步。NTP 协议使得网络中的终端设备具备校正本地时钟以实现时间同步的能力，保持所有设备时间一致，从而使得网络中的各种设备能够基于同一时间执行各种应用，如能源管理，休眠设置等[35]。

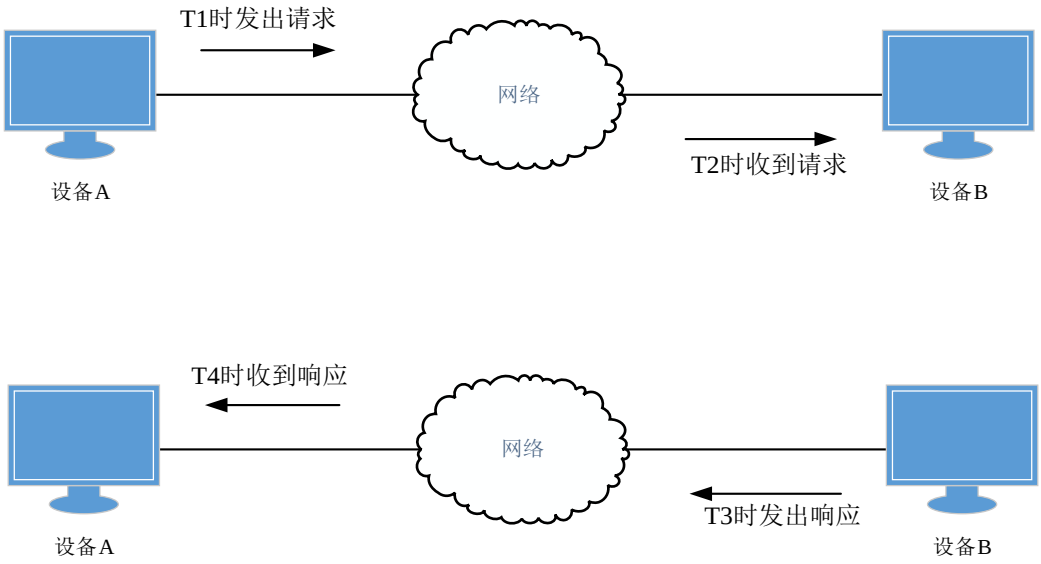

图 4.1 NTP 同步协议

NTP 的基本工作原理如图 4.1 所示，A、B 是两台连接网络的设备，二者拥有独立的本地时钟，需要通过时间同步协议实现时间同步。假设设备 B 为时间服务器，设备 A 需要将自己的时间与 B 同步，首先 A 发送一个 NTP 请求，并附加时间戳 T1，在时间 T2 时，设备 B 收到了 A 的请求，之后 B 向 A 发送响应，并附上时间戳 T3，在 T4 时响应到达设备 A。根据这些时间就可以计算出往返时延，从而计算出二者的时间差：

$$delay = (T4 - T1) - (T3 - T2) \quad (4\text{-}1)$$

$$offset = ((T2 - T1) + (T3 - T4))/2 \quad (4\text{-}2)$$

4.2.2 树形时间模型

SRP 和 RRP 是两种经典的树形时间同步模型，接下来我们将详细介绍他们运行的整个过程。如图 4.2，RRP 时间同步模型在不需要节点 C 广播信息的情况下就可以

实现节点 C 与节点 A 的时间同步，具体过程如下：节点 B 在时间 T1 时向另外两个节点 A、C 广播发送消息 *Mes*1，二者分别在时间 T2 与时间 T3 收到了消息，然后时间节点 A 又发送消息 *Mes*2 至节点 C， *Mes*2 中包含着 A 收到 *Mes*1 的时间 T2，通过公式(4-3)即可计算出 A 与 C 之间的时间差，进而实现二者的时间同步。图 4.3 则展示了 SRP 模型进行时间同步的过程，该模型是通过两个节点之间互相发送消息来计算时间偏移进而实现时间同步，计算公式如(4-4)所示：

$$offset = T2 - T3 \tag{4-3}$$

$$offset = ((T2 - T1) - (T4 - T3)) / 2 \tag{4-4}$$

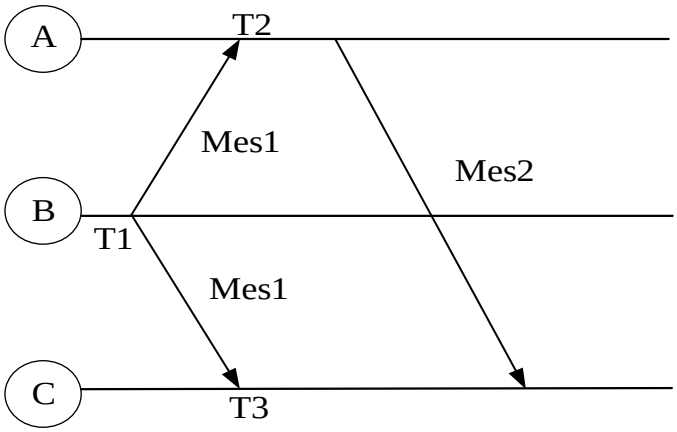

图 4.2 RRP 时间同步模型

在 RRP 模型中，假设节点 A 是恶意节点，可以发现它只能影响到 C 节点；而在 SRP 模型中，假如 A 是恶意节点，错误会被持续传递下去。通过对上述两种树形时间协议的分析，我们可以发现，RRP 模型在对恶意攻击的抵抗方面具有较好的表现，但是错误无法被改正；而 SRP 模型则是有可能引入大规模的错误时间。

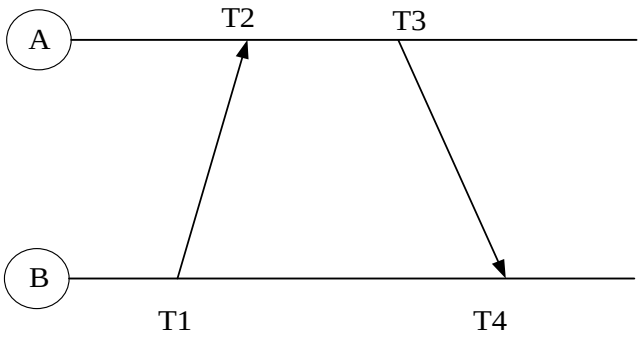

图 4.3 SRP 时间同步模型

基于上述对分布式与树形时间协议的分析，我们引入了区块链技术来实现时间同步。区块链是一种分布式的架构，它能够大幅改进树形时间协议面临的安全性问题；类似于传统的分布式时间同步方案，它在共识方面存在问题，但是通过改进 PoS 共识机制，可以有效提高同步效率。

4.3 系统模型

如图 4.4 所示，本章所展示的基于区块链的时间分发与同步系统由四个组件构成，分别为：时间源、共识节点、普通时间节点以及智能设备。

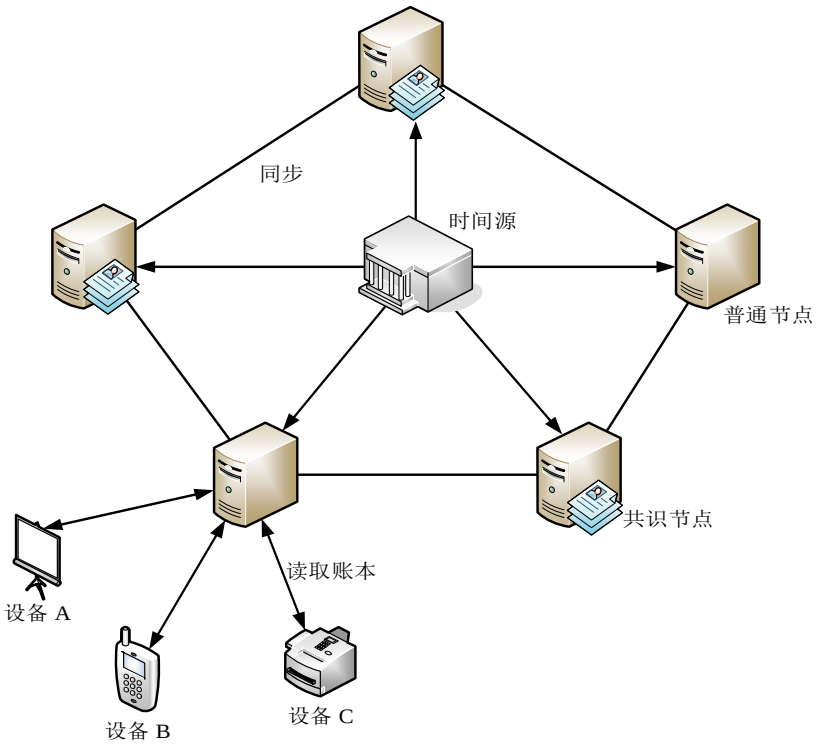

图 4.4 系统模型

4.3.1 时间源

时间源，也就是联盟链的发起者，拥有着最高的系统权限。同时，掌握最精确的全局时间，是时间同步过程的发起点。除此之外，时间源也作为身份认证中心存在，即如果某一节点需要加入到该区块链系统当中，需要时间源对其提交的身份信息进行安全认证，只有通过时间源验证的节点才可以作为时间节点加入该联盟。通过验证的节点会得到由时间源颁发的安全证书，这个证书中包含有一个唯一的序列号，而公私钥对也由时间源产生，序列号代表节点的身份，公私钥对用于签名。在智能设备初次连接时间节点进行初始化的时候，安全证书可以证明节点身份合法。精确的时间由时

间源产生，虽然在产生时间方面是中心化的，但是时间的同步过程却是去中心化的。

4.3.2 普通时间节点

普通节点，可以看做是网络拓扑当中的一个区域中心，因为众多的设备就是通过连接时间节点接入到时间系统来进行时间同步的。在时间节点存储有区块链的账本，账本当中记录的是时间源发布出来的时间信息。同时，该节点自身有一个本地时钟，用于记录当前时间与最后一个区块上记录的时间的时间偏移。而为了确保时钟同步的安全与高效率，在时间节点维持着白名单以及黑名单，黑名单上记录的是恶意设备的电子身份标识（EID）以及 IP 地址，在设备发出时间同步请求时，需要提交身份信息，时间节点就会验证设备是否在黑名单中，如果证明请求设备是恶意的，普通节点就会拒绝提供服务。

可以明确的说，一台设备要实现时间同步，需要三个时间参数：区块链账本上最后一个区块中记录的时间、时间节点本地时钟提供的时间偏移、网络时延。毫无疑问，前两个参数是可以直接由时间节点得到，第三个参数是根据网络情况实时变动的，因此需要设备在发出同步请求时主动去探测网络拥塞情况，根据时间节点的响应进行计算。

4.3.3 共识节点

共识节点，可以看做是一个特殊的普通时间节点，它除了需要完成一个普通时间节点的工作之外，还需要达成共识以实现账本的更新。在比特币系统中，使用的是工作量证明机制，这种机制的优点在于完全的去中心以及全网节点无限制的访问。缺点就是在争夺记账权的时候带来的算力消耗，它迫使节点去升级自己的硬件设施，提升自己的计算能力，而这必然导致能源的消耗，甚至最后出现中心化的趋势。考虑到资源的过度消耗以及时钟同步系统对于效率的高要求，我们的方案采用权益证明机制(PoS)来实现共识。

所谓的权益，代表的是区块链节点所拥有的某种资源，从博弈论的角度来说，节点持有的这种资源保证了节点不会做出有悖于整个系统利益的操作。同时，共识机制利用这种资源来进行共识节点的选举。在我们的方案中，这种资源代表着时间节点所连接的要进行时间同步的设备，在进行共识节点的选取时，时间源会给出一个资源量，所有连接的设备数量大于该标准的时间节点会被选举出来成为共识节点，而这些共识节点会轮流参与新区快的生成、广播与共识。具体的流程就是，时间源将最新的精确时间发送给共识节点，共识节点会打包新区快，添加自己的签名，之后在整个网络广播，网络中所有的节点验证后全网更新账本。这种共识机制有利于减小系统开销、提升系统性能。

区块信息如图 4.5 所示，具体包含的内容解释如下：

-当前 hash: 当前区块的区块体的哈希值。

-前一区块 hash: 前一个区块头的哈希值，用于与前一个区块进行连接。

-共识节点签名: 共识节点的数字签名，用于验证区块信息的合法性、真实性。

-共识节点 ID: 共识节点的独一无二的身份 ID。

-当前时间: 时间源发布的最新的时间信息。

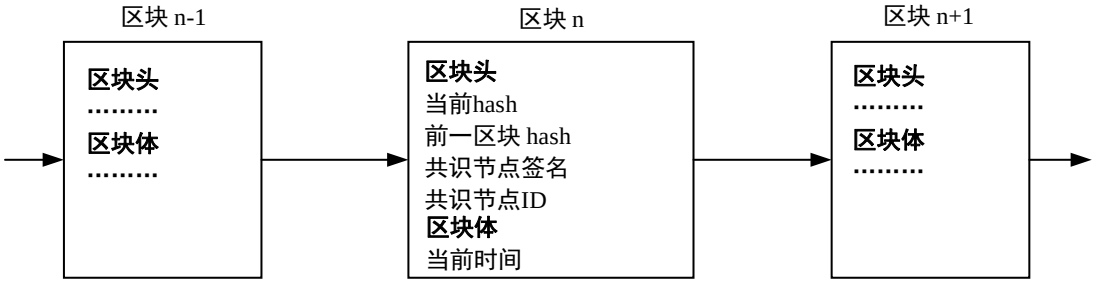

图 4.5 时间区块结构

4.3.4 同步设备

方案中的设备就是网络中需要更新时间的智能设备，例如智能手机、智能家居系统或者无线射频终端等。这些设备在本地维护着一个时钟，用于在时间同步之外的时间段实现时间服务。而如果要减小本地时间与绝对时间的偏移，就需要进行时间同步操作。过程如下，待同步设备主动连接时间节点并在 T_1 发送请求，读取节点账本得到 T_b，时间节点在 T_2 收到请求，在时间 T_3 做出响应，并将时间节点本地时钟信息 T_c 发送给设备，设备在时间 T_4 收到响应。最后计算出该时刻的绝对时间：

$$T = offset + T_b + T_c = ((\mathrm{T}2 - \mathrm{T}1) + (\mathrm{T}3 - \mathrm{T}4))/2 + T_b + T_c \tag{4-5}$$

4.4 算法设计

接下来会介绍方案中设计的具体算法，包括系统初始化、共识节点选取以及同步过程的具体过程。

4.4.1 系统初始化

系统最开始只有时间源一个节点存在，之后通过扩展、准入时间节点形成一个联盟区块链系统。在系统初始化的过程中，时间源负责对申请加入的节点进行安全验证，并颁发身份证书，在这个过程中是一个中心节点。

算法 4.1: 系统初始化

```
1: if the node is time source then
2:     verifyNode()
3:     sendMesToConsNode()
4:else
5:     nodeType ← common time node
6:end if
7:  selectConsNode() ← randomSelect()
```

算法 4.1 利用伪代码展示了系统初始化的过程。从时间源开始，在有节点申请加入区块链时间系统时，区块链会验证节点身份，并颁发证书，此时所有的时间节点都是普通时间节点，初始化过程也保证了系统的扩展性。之后就会选取时间节点，区别于已经开始运行的系统，在初始化阶段，权益证明机制的规则无法选取共识节点。因此，选择使用随机函数来随机的选取固定数量的共识节点来执行系统的共识步骤。

4.4.2 选取共识节点

算法 4.2：选取共识节点

```
输入:网络中 N 个时间节点
输入: T，选举期
输入: q，成为共识节点连接设备的基准
1:if  currentTime() %T == 0 then
2:      foreach n ∈ N
3:            numOfDevConted ← submit()
4:      consensusNodes ← compare(q)
5:end if
6: currentConsNode ← choose()
7: broadcast() ← currentConsNode
8: verifySig() ← normal time node
9: confirmBlock()
10: foreach n ∈ N
11:       setLocalClock(0)
```

由于本方案是基于区块链实现的，因此共识节点的选取就是一个必不可少的重要环节。选出的共识节点完成了添加区块的任务，保证了账本的同步更新，使得设备能

够获取时间源更新的最新绝对时间。算法 4.2 中显示，共识节点的选取是定期举行的。对于时间节点来说，会有大量的设备读取账本来更新时间，这些设备就是他们的资源，资源越多，就越有可能成为共识节点。选取共识节点的具体步骤如下：判断当前时间是否是选举时间，如果是的话，所有的时间节点需要提交自身连接的设备数至时间源，时间源会选择出部分节点成为共识节点。在这些共识节点当中，基于顺序成为当前时间的共识节点，制作时间区块并广播区块，得到全网验证后，制作的区块得到证实通过，全网所有的时间节点将该区块添加到区块链上，最后将所有时间节点本地时钟的时间偏移置零。整个选取共识节点的过程基于权益证明机制，但在进行算力挖矿的步骤有所不同，省去计算的过程是为了提高整个过程的效率，以尽快的完成账本更新，减小设备同步时间的误差。

4.4.3 时间同步

算法 4.3: 时间同步

1: send request to time node nearby
2: **if** connection between device and time node exists
then
3: $verify() \leftarrow time\,node$
4: **if** device is in black list **then**
5: refuse to serve
6: **else**
7: $t_1 \leftarrow readLedger() \leftarrow device$
8: $t_2 \leftarrow sendOffsetToDev() \leftarrow time\,node$
9: $t_3 \leftarrow dectTransTime()$
10: $updateTime(t_1 + t_2 + t_3)$
11: **end if**
12: **else**
13: broadcast to reconnect a new time node
14: **return** line 1
15: **end if**

正是基于时间同步的目的，设计了本文的时间分发和同步方案，算法 4.3 利用伪代码展示了时间同步算法。经过系统初始化以及选取共识节点的过程，系统中的设备可以主动发送请求并查询时间节点存储的账本来完成时间同步的操作。在设备发送请求时，会首先判断是否能够连接到以前使用的时间节点，如果连接不存在，智能设备

需要重新寻找时间节点并建立连接，提交设备 EID 以及 IP 地址，时间节点则需要根据接收到的信息更新节点白名单。如果连接存在，就只需验证设备是否在黑名单中，如果在黑名单中，时间节点拒绝服务，否则节点需要接收到设备发出的请求后作出响应。根据算法 4.3，时间更新需要用到三个参数，即账本最后一个区块记录的时间 t_1、时间节点本地时钟记录的时间偏移 t_2、网络时延 t_3。其中前两个参数可以通过读取账本、时间节点的响应信息获取到，如公式（3-1）所示，第三个参数根据网络时间协议计算得出。+

4.5 性能分析

本方案是基于联盟区块链的时间分发与同步系统，对内可以在物联网场景下同步设备的时间，对外可以跨链为其他的区块链系统提供安全的时间服务。相比于传统的时间同步方案，本方案在面对恶意设备的攻击时，仍然能实现高安全性，并拥有高效率的表现。

4.5.1 安全性分析

(1) 数据安全性

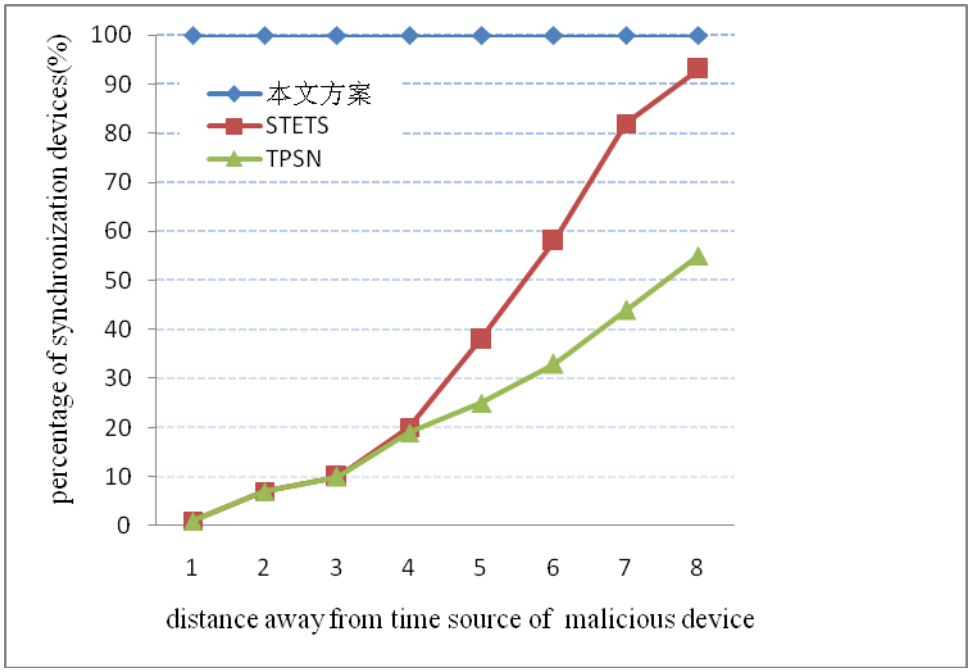

图 4.6 数据安全性分析

所谓的数据安全性，指的是传递的时间信息的安全性。传统的时间同步方案中，同步设备在获取到时间信息后，仍然需要将时间传递下去。在这种模式中，存在着一

个缺陷，即如果同步设备本身是一个恶意节点，它就可以伪造错误的时间信息传递给其他节点，严重的后果是这些收到错误时间信息的设备会继续将错误的信息扩散出去，导致一小部分的恶意攻击者就可以扰乱整个同步系统。在本文的方案中，则不会产生这种情况，因为所有的设备实现时间同步需要自己主动发送请求，读取时间节点账本信息后更新本地时间，而所有的时间节点都是经过时间源的身份认证之后才加入的联盟链系统。图 4.6 展示了时间源附近的恶意设备对几种方案的同步率的影响，之所以选取时间源附近的节点是考虑到时间源附近的节点对于信息传播的影响，越靠近时间源影响越大。为了便于分析，我们使用 Ganeriwal[36]等人提出的 TPSN 和 Qiu[37]等人提出的 STETS 两种方案与本文方案进行对比，从结果可以看出 TPSN 与 STETS 两种方案容错性较差，恶意设备距离时间源越近，同步率越低。而本文设计的方案，排除掉网络中丢包等因素，无论恶意节点距离时间源多近，不会影响设备实现正确的时间同步。

(2)抗 DoS 攻击

DoS 是 Denial of Service 的简称，即拒绝服务，通过发送大量垃圾请求，DoS 攻击可以使计算机或者网络瘫痪，无法正常提供服务。本方案的时间同步操作时是通过设备主动发送请求执行，这虽然降低了系统开销，但也给 DoS 攻击提供了可能。因此，我们设计了黑名单机制，时间节点在本地维护一个黑名单，将请求异常的设备记录在黑名单中，当恶意设备持续发送同步请求时，时间节点会拒绝提供连接，以保护自身的资源，这种机制可以大大降低 DoS 攻击的危害性。

4.5.2 性能分析

系统的开销影响着系统的工作效率，在这一部分，通过计算在同步过程当中信息的交换量来对系统开销做出分析。我们假设区块链中的时间节点数是 N，需要同步的设备数量是 $10N$，即总数为 $11N$，所有设备完成时间同步需要的信息总量为 M_{blk}。根据本文的方案，区块链中的时间节点进行同步是通过广播发送信息的，因此区块链当中的账本同步需要的信息总量是 N。之后设备会在需要时发送同步请求，根据网络时间协议，一台设备完成同步需要一次信息交互，所以所有的设备完成一次同步需要的交互信息数是 $10N$，最后可以得到：

$$M_{blk} = 10N + N = 11N \tag{4-6}$$

在方案 STETS 中，我们使用同样的方法进行计算。基于同样的参数，我们假设 STETS 中节点数为 $11N$。STETS 中的节点分为两类：骨干节点与被动节点，骨干节点会执行信息交换的操作，作用类似于区块链方案中的时间节点，被动节点只是接收

信息，类似于区块链方案中的同步设备。因此假设二者的数量分别为N、$10N$，信息总量为M_{STETS}，根据 STETS 方案介绍可以得到，两种节点完成时间同步交互信息的数量分别是$3N$、$10N$，最后得到：

$$M_{STETS} = 3N + 10N = 13N \tag{4-7}$$

比较得出，本文提出的方案在系统开销方面有一定的改进。

4.6 本章小结

本章在对分布式以及树形时间同步协议进行分析的基础上提出基于区块链的时间同步方案，区块链作为一种分布式架构，应用到时间分发与同步领域可以有效抵抗恶意节点的攻击。该方案主要包括四个组件：时间源、共识节点、普通时间节点以及同步设备，之后分别介绍各个组件的功能。方案的重要算法以伪代码的形式展现，本方案对传统的 PoS 共识机制进行改进，使其契合了时间同步系统对效率的高要求。最后进行了性能分析，得出相比于传统的方案，本方案在安全性、开销等方面有了显著提高。该方案本身是一个独立的区块链系统，但是亦可以作为侧链为其他的区块链系统提供时间戳服务。

第五章　展望与总结

5.1 论文工作总结

随着互联网信息技术的飞速发展，网络数据量迅猛增长，不可避免的数据隐私保护成为用户关注的重点。在医疗领域，经历了纸质信息到电子健康档案的转变，用户使用云服务器来存储个人数据，但是健康数据包含有众多的个人隐私，云服务提供商可以轻易的获取用户的个人数据、窃取用户隐私。因此，用户为了保障敏感数据的隐私性，往往将数据加密后上传至云端，但是这会导致用户失去数据的操控性，用户无法实现对数据细粒度的访问控制。健康数据作为一种特殊的资源，对于个人而言，健康档案便于用户得到更好的医疗服务，提高生命质量的同时减少大笔的医疗费用；对于医疗机构而言，健康档案便于医务人员快速、准确、直观的了解病人的病情以及病史，缩短了确诊时间，而且大量的 EHR 为医疗机构进行医学研究提供了可能。在此基础上，本文提出了一种基于区块链技术的医疗信息隐私保护与共享方案。

论文首先介绍了电子健康档案隐私保护的研究背景与意义，简要阐述了该领域国内外研究现状。其次介绍了区块链，区块链作为一种最近兴起的热门技术，是一种去中心的分布式架构，实现了网络中任意两个用户之间的信任；区块链账本的公开性、不可篡改性、可验证性为文章的方案提供了可行性；同样介绍了区块链技术在效率、安全与能耗方面存在的问题。首先，介绍健康档案隐私保护系统的系统架构。为保证系统能够获取安全的时间，提出了一种权威时间分发与同步方案。具体工作内容总结如下：

1. 针对云环境中个人电子健康档案存储存在的安全问题，提出一种基于区块链的数据安全存储以及共享方案。在现实的医疗领域，病人的个人健康信息呈现碎片化分布的特点，数据加密存储后无法实现细粒度的访问控制等问题。本方案通过加密机制实现了用户对数据的完全收集与控制，结合区块链的公开账本，将数据拥有者个性化定制的访问控制策略添加到账本上，实现访问请求的公开验证以及用户数据的细粒度访问控制。利用改进的 PoS 共识机制，系统可以根据安全性要求动态的调整共识时间，在安全的基础上，实现了医疗数据的共享。总体来看，本方案在数据安全的基础上，实现系统的高效、低能耗运行。

2. 针对物联网领域中存在的权威时间分发与同步问题，提出了一种可以用于物联网环境的时间分发与同步方案。在传统的树形时间同步模型下，小部分恶意节点的攻击可能导致整个网络大部分时间的错误。文章中提出的同步方案，通过分布式的区块链技术来达成时间的同步，在联盟链环境中，通过身份认证的共识节点，其安全性

可以得到最大程度的保障，时间同步设备通过主动发送请求的方式来更新本地时间，免去了设备之间传递错误时间的通道。通过改变传统的 PoS 共识机制，实现共识效率的提高，并减少授时误差，提供设备同步精度。并且该方案中的授时系统作为一条独立的区块链系统可以与其他链交互，为不同应用场景下的区块链系统提供时间服务，具有广阔的应用场景。

5.2 展望

本文方案主要将区块链技术用于医疗领域的信息保护场景下，实际上方案可以用于普通的数据保护与共享场景，方案的安全性不会有任何的影响。本方案的提出是基于作者本人目前的能力，受限于个人的时间有限以及能力的不足，方案虽然提出基于区块链技术来解决隐私保护过程中所遇到的一些问题，但是仍然有诸多不足之处需要在以后重点关注研究：

1. 区块链技术本身存在着矛盾，比如安全性和能源消耗方面的对立，这是由于目前的共识机制存在着缺陷，如何通过改进现有的共识机制或者提出新的共识方案实现高效与安全之间的平衡仍是一个棘手的问题。

2. 在方案中提出由用户为医疗记录制定访问控制策略，并且提出了一种访问控制策略的模型，但是这种策略稍显粗糙，针对用户需求，制定更加细化的访问控制模型也是改进之处。

3. 方案最后是通过仿真分析，对系统的效率以及存储开销作出评估，因此需要构建实际的基于区块链的隐私保护与共享系统实现对方案的进一步改进。

参考文献

[1] Zhang X, Qi L, Dou W, et al. MRMondrian: Scalable Multidimensional Anonymisation for Big Data Privacy Preservation[J]. IEEE Transactions on Big Data, 2017, PP(99):1-1.

[2] Löhr H, Sadeghi A R, Winandy M. Securing the e-health cloud[C]//Proceedings of the 1st ACM International Health Informatics Symposium. ACM, 2010: 220-229.

[3] Who will keep the public healthy?: educating public health professionals for the 21st century[M]. National Academies Press, 2003.

[4] Lu R, Liang X, Li X, et al. Eppa: An efficient and privacy-preserving aggregation scheme for secure smart grid communications[J]. IEEE Transactions on Parallel and Distributed Systems, 2012, 23(9): 1621-1631.

[5] Chen M, Qian Y, Chen J, et al. Privacy protection and intrusion avoidance for cloudlet-based medical data sharing[J]. IEEE Transactions on Cloud Computing, 2016.

[6] He C, Fan X, Li Y. Toward ubiquitous healthcare services with a novel efficient cloud platform[J]. IEEE Transactions on Biomedical Engineering, 2013, 60(1): 230-234.

[7] 朱晓卓. 论电子健康档案的隐私特性及保护[J]. 中国卫生事业管理, 2014, 31(8):603-604.

[8] Li M, Yu S, Cao N, et al. Authorized private keyword search over encrypted data in cloud computing[C]//Distributed Computing Systems (ICDCS), 2011 31st International Conference on. IEEE, 2011: 383-392.

[9] Lu R, Lin X, Shen X. SPOC: A secure and privacy-preserving opportunistic computing framework for mobile-healthcare emergency[J]. IEEE Transactions on Parallel and Distributed Systems, 2013, 24(3): 614-624.

[10] Hung P C K, Zheng Y. Privacy access control model for aggregated e-health services[C]//EDOC Conference Workshop, 2007. EDOC'07. Eleventh International IEEE. IEEE, 2007: 12-19.

[11] Akinyele J A, Pagano M W, Green M D, et al. Securing electronic medical records using attribute-based encryption on mobile devices[C]//Proceedings of the 1st ACM workshop on Security and privacy in smartphones and mobile devices. ACM, 2011: 75-86.

[12] Hur J, Noh D K. Attribute-based access control with efficient revocation in data outsourcing systems[J]. IEEE Transactions on Parallel and Distributed Systems, 2011, 22(7): 1214-1221.

[13] Peterson K, Deeduvanu R, Kanjamala P, et al. A blockchain-based approach to health information exchange networks[C]//Proc. NIST Workshop Blockchain Healthcare. 2016, 1: 1-10.

[14] Ekblaw A, Azaria A, Halamka J D, et al. A Case Study for Blockchain in Healthcare:“MedRec” prototype for electronic health records and medical research data[C]//Proceedings of IEEE Open

& Big Data Conference. 2016, 13: 13.

[15] Kuo T T, Ohno-Machado L. ModelChain: Decentralized Privacy-Preserving Healthcare Predictive Modeling Framework on Private Blockchain Networks[J]. arXiv preprint arXiv:1802.01746, 2018.

[16] 吴银燕. 国务院办公厅关于促进和规范健康医疗大数据应用发展的指导意见[J]. 甘肃医药, 2016, 35(7):561-561.

[17] 张怡婷, 傅煜川, 杨明,等. 基于 PBAC 模型和 IBE 的医疗数据访问控制方案[J]. 通信学报, 2015, 36(12):200-211.

[18] Chen M, Qian Y, Chen J, et al. Privacy Protection and Intrusion Avoidance for Cloudlet-based Medical Data Sharing[J]. 2016, PP(99):1-1

[19] Yue X, Wang H, Jin D, et al. Healthcare Data Gateways: Found Healthcare Intelligence on Blockchain with Novel Privacy Risk Control.[J]. Journal of Medical Systems, 2016, 40(10):218.

[20] 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.

[21] Fu D, Fang L. Blockchain-based trusted computing in social network[C]// IEEE International Conference on Computer and Communications. IEEE, 2017:19-22.

[22] Nakamoto S. Bitcoin: A peer-to-peer electronic cash system[J]. 2008.

[23] Sasson E B, Chiesa A, Garman C, et al. Zerocash: Decentralized anonymous payments from bitcoin[C]//Security and Privacy (SP), 2014 IEEE Symposium on. IEEE, 2014: 459-474.

[24] Zyskind G, Nathan O. Decentralizing privacy: Using blockchain to protect personal data[C]//Security and Privacy Workshops (SPW), 2015 IEEE. IEEE, 2015: 180-184.

[25] Kosba A, Miller A, Shi E, et al. Hawk: The blockchain model of cryptography and privacy-preserving smart contracts[C]//Security and Privacy (SP), 2016 IEEE Symposium on. IEEE, 2016: 839-858.

[26] Peters G, Panayi E, Chapelle A. Trends in cryptocurrencies and blockchain technologies: a monetary theory and regulation perspective[J]. 2015.

[27] Danezis G, Meiklejohn S. Centrally banked cryptocurrencies[J]. arXiv preprint arXiv:1505.06895, 2015.

[28] 沈鑫, 裴庆祺, 刘雪峰. 区块链技术综述[J]. 网络与信息安全学报, 2016, 2(11).

[29] 何蒲, 于戈, 张岩峰,等. 区块链技术与应用前瞻综述[J]. 计算机科学, 2017, 44(4):1-7.

[30] Zheng Z, Xie S, Dai H N, et al. Blockchain challenges and opportunities: A survey[J]. Work Pap.–2016, 2016.

[31] Lin I C, Liao T C. A Survey of Blockchain Security Issues and Challenges[J]. IJ Network Security, 2017, 19(5): 653-659.

[32] Xu E, Ding Z, Dasgupta S. Target tracking and mobile sensor navigation in wireless sensor networks[J]. IEEE Transactions on mobile computing, 2013, 12(1): 177-186.

[33] Liu Y, Li J, Guizani M. Lightweight secure global time synchronization for wireless sensor networks[C]//Wireless Communications and Networking Conference (WCNC), 2012 IEEE. IEEE, 2012: 2312-2317.

[34] Kadowaki Y, Ishii H. Event-based distributed clock synchronization for wireless sensor networks[J]. IEEE Transactions on Automatic Control, 2015, 60(8): 2266-2271.

[35] 李明国，宋海娜，胡卫东．Internet 网络时间协议原理与实现[J]．计算机工程，2002, 28(2):275-277.

[36] Ganeriwal S, Kumar R, Srivastava M B. Timing-sync protocol for sensor networks[C]//Proceedings of the 1st international conference on Embedded networked sensor systems. ACM, 2003: 138-149.

[37] Qiu T, Chi L, Guo W, et al. STETS: A novel energy-efficient time synchronization scheme based on embedded networking devices[J]. Microprocessors and Microsystems, 2015, 39(8): 1285-1295.

[38] 王育民，刘建伟．高等学校电子信息类规划教材，通信网的安全——理论与技术[M]．西安电子科技大学出版社，1999.

[39] 王于丁，杨家海，徐聪,等．云计算访问控制技术研究综述[J]．软件学报，2015, 26(5):1129-1150.

[40] 牛宇，颜苗苗，郑红,等．云计算环境下医疗数据访问控制研究综述[J]．智慧健康，2016, 2(2):23-28.

[41] Lampson B W. Protection[J]. Acm Sigops Operating Systems Review, 1974, 8(1):18-24.

致谢

值此论文即将完成时刻，三年的硕士生涯即将技术。三年一瞬，欢笑与泪水相伴，但是更多的是收获，内心充满的是感激。在此，我要向所有帮助过我的老师、家人和朋友致以衷心的感谢！

首先要感谢我的导师樊凯副教授，研究生期间，学术上樊老师给我指引方向，行动上树立榜样，在论文的研究和写作的过程中给予了悉心的指导和帮助，在修改过程中提出宝贵的意见。樊老师治学严谨，工作兢兢业业，他一丝不苟的作风，是我一生学习的榜样，衷心祝愿老师家庭幸福、工作顺利、桃李芬芳满天涯。

感谢实验室的各位，无论是研三的同学还是师弟师妹，他们在学习和生活上都给我不断地鼓励和帮助。感谢实验室安静和谐的氛围，让我能够顺利完成论文工作。感谢我亲爱的朋友们，我们共同见证了各自的的改变，是你们让我体会到情同手足的感觉，在我困难疲倦的时候给我安慰和鼓励，为了梦想我们埋头苦干，感谢你们带来的喜悦和感动！感谢我的舍友们，让我感到家的温暖，在这段最美好的青春里我们共同奋斗共同进步，未来的日子我们虽各奔东西，但心依然在一起，祝愿你们都能事业有成，早日找到自己的幸福！

感谢我的父母及家人，感谢他们的默默陪伴，是他们的关心和支持让我充满动力，是他们的宽容和理解让我砥砺前行，是他们的鼓励和付出让我勇往直前。在未来的人生路途中，我会继续努力，不辜负你们的期望。

最后衷心感谢每一位参与论文评审的专家和老师，感谢你们的宝贵时间和辛勤付出，感谢你们对我论文提出的宝贵意见，祝你们身体健康，工作顺利。

作者简介

1. 基本情况

任延辉，男，河南郑州人，1992 年 9 月出生，西安电子科技大学网络与信息安全学院电子与通信工程专业 2015 级硕士研究生。

2. 教育背景

2011.09～2015.07 西安电子科技大学长安学院，本科，专业：通信工程
2015.08～至今西安电子科技大学，硕士研究生，专业：电子与通信工程

3. 攻读硕士学位期间的研究成果

3.1 发表学术论文

[1] Fan K, Ren Y, Wang Y, et al. Blockchain-based efficient privacy preserving and data sharing scheme of content-centric network in 5G[J]. IET Communications, 2017，12(5): 527-532.

[2] Kai Fan, Yanhui Ren, Zheng Yan, Shangyang Wang, Hui Li, Yintang Yang. Secure Time Synchronization Scheme in IoT based on Blockchain. The 2018 IEEE International Conference on Blockchain. (Accepted)

3.2 申请（授权）专利

[1] 樊凯、任延辉、王尚阳等. 专利名称:一种基于区块链的安全文件存储和共享方法. 中国，申请号：201810139906.X.

3.3 参与科研项目及获奖

[1] 国家自然科学基金面上项目，无线体域网中敏感数据的高效隐私保护方法研究(No. 61772403)，2018.1.1-2021.12.31，参与基于区块链的敏感数据高效隐私保护方法研究。

[2] 国家重点研发计划“网络空间安全重点”专项子任务，异构身份联盟与监管基础科学问题研究（2017YFB0802300），2017.7-2020.6，参与异构环境下数

据的隐私保护和安全共享方法研究。

[3] 国家自然科学基金青年基金项目，手机支付环境下财产及隐私保护研究(No. 61303216，2014.1.1-2016.12.31，参与数据及财产隐私保护方法研究。

[4] 中国科学院国家授时中心项目，互联网时间服务系统安全性能测试评估，2017.4-2018.3，参与互联网时间分发和同步安全性能测试。

西安电子科技大学
XIDIAN UNIVERSITY

地址：西安市太白南路2号
邮编：710071
网址：www.xidian.edu.cn