|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | | | | |
| 武汉大学logo | | | | |
|  | | | | |
| 高级数据库技术课程报告 | | | | |
|  | 研究生姓名 | ： | 陈小龙 |  |
| 学号 | ： | 2019282110194 |
| 指导教师姓名、职称 | ： | 梁意文教授 |
| 专业名称 | ： | 计算机科学与技术 |
|  |  |  |
|  | | | | |
| 二〇一九年十二月 | | | | |

# 

# 摘要

区块链是一种全新的去中心化架构，同时也是一种分布式计算范式。在多国央行和政府热议区块链技术的背景下，区块链技术正日益走近公众视野和政策层面。本文通过分析区阐述比特币以及区块链的技术原理和应用，揭示了区块链的由来和本质，以及其发展前景。

关键词：区块链；去中心化；比特币

目录

[**摘要 I**](#_Toc26199050)

[**引言 1**](#_Toc26199051)

[**1比特币与区块链概述 2**](#_Toc26199052)

[**2区块链的架构模型 3**](#_Toc26199053)

[2.1区块链架构模型概述 3](#_Toc26199054)

[2.2数据层 3](#_Toc26199055)

[2.3网络层 4](#_Toc26199056)

[2.4共识层 5](#_Toc26199057)

[2.5激励层 5](#_Toc26199058)

[2.6合约层 6](#_Toc26199059)

[**3区块链的现存问题 6**](#_Toc26199060)

[3.1效率问题 6](#_Toc26199061)

[3.2安全问题 7](#_Toc26199062)

[3.3资源问题 7](#_Toc26199063)

[3.4博弈问题 7](#_Toc26199064)

[**4区块链应用前景 8**](#_Toc26199065)

[**参考文献 9**](#_Toc26199066)

[**致谢 10**](#_Toc26199067)

# 引言

区块链起源于比特币。雷曼兄弟倒闭两周后，2008年11月1日，一位名叫中本聪的人发表了《比特币：一种点对点的电子现金系统》一文，阐述了P2P网络技术、加密技术、时间戳技术、区块链技术等电子现金系统的架构理念，比特币从此诞生。

两个月后理论变成实践，第一个序号为0的比特币诞生。几天后出现序号为1的区块，并与序号为0的区块接续，形成的结果就叫区块链。广义的区块链技术有望彻底重塑人类的社会生活形态，为金融、科技、文化、政治等领域带来深刻的变革。

# 1比特币与区块链概述

区块链是一种分布式去中心化账本，具体来说，它按照时间顺序将数据区块用类似链表的方式组织成的数据结构，并且用密码学的方法让其不可篡改和伪造。此外，区块链可以安全存储简单的、有先后关系的、可在系统内进行验证的数据。

区块链的出现解决了数字货币的两大问题：双重支付问题以及拜占庭将军问题[1]。

双重支付问题是同一笔钱被使用超过了一次，这在原有的以物理实体为基础的传统金融体系中是可以自然避免的问题。在区块链出现之前，数字货币都是通过可信任的中心化第三方机构来保证，例如：银行、支付宝、微信等等。区块链技术则是通过共识机制和分布式账本，来解决双重支付问题，并不需要依赖可信任的第三方机构。

拜占庭将军问题是现实世界问题的模型化，适用于分布式网络的简单抽象为“在缺少可信中心节点的情况下，分布式节点如何达成共识建立互信的问题”。区块链使用“工作量证明”以及“权益证明”或者其他共识机制，结合加密技术，是一个不可信的网络编程可信网络，网络中的节点可以达成一致，无需依赖某个节点或中心节点。与传统的中心机构（如中央银行）的信用背书机制不同，这标志着中心化的国家信用向着去中心化的算法信用的根本性变革。

比特币作为一种新兴的数字货币，以区块链技术为支撑，凭借其先发优势，目前已经形成体系完备的涵盖发行、流通和金融衍生市场的生态圈和产业链。这也是比特币长期占据绝大多数数字加密货币市场份额的主要原因。比特币的开源特性吸引了大量开发者持续性地贡献其创新技术、方法和机制。比特币的各网络节点（矿工）提供算力以保证比特币的稳定性与安全性，其算力大多来自于设备商销售的专门用于PoW共识算法的专业设备（矿机）。比特币网络为每一个新发现的区块发行一定数量的比特币以奖励矿工，同时部分矿工可能会相互合作建立收益共享的矿池，以便汇集算力来提高获得比特币的概率。比特币经过发行进图流通环节后，持币人可以通过特定的软件平台来购买商品或者服务，这体现了比特币的货币属性。同时，由于比特币的涨跌机制使其完全具备金融衍生产品的所有属性，因此出现了比特币交易平台以方便持币人投资或者投机比特币。在流通环节和金融市场中，每一笔比特币交易都会由比特币网络的全体矿工验证并记入区块链。

基于上述情况，比特币可以成为区块链技术的典型应用。迄今为止，区块链的核心技术和人才资源大多数仍在比特币研发领域。然而，区块链作为未来新一代的底层基础技术，其应用范围势必会超越数字加密货币进而延伸到金融、经济、科技和政治领域；并且区块链已经进入政府的视野。我们有理由相信，基于比特币开发出来的技术、模式和机制，将会为区块链在其他领域的应用提供有益的借鉴；反过来其他领域的发展也会回馈解决比特币现存的问题。因此，比特币和区块链技术必定会相互依存、协同发展。

# 

# 2区块链的架构模型

## 2.1区块链架构模型概述

一般来说，区块链系统由数据层、网络层、共识层、激励层、合约层和应用层组成。其中，数据层封装了底层数据块以及相关的数据加密和时间戳技术。网络层则包括分布式组网机制、数据传播机制和数据验证机制等。共识层主要封装网络节点的各类共识算法。激励层将经济因素集成到区块链技术中来，分为经济激励的发行机制和分配机制等。合约层主要封装各类脚本、算法和只能合约，这些是区块链可编程特性的基础。用用层则封装了区块链的各种应用场景和案例。

上述模型中基于分布式节点的共识机制、基于共识算力的经济激励和可编程的智能合约是区块链技术的最大创新点。

## 2.2数据层

狭义的区块链就是去中心化系统各个节点之间共享的数据账本。系统中的每一个节点通过哈希算法和Merkle树数据结构，把一段时间内的交易数据和代码封装成一个带有时间戳的数据区块，并链接到当前最长的主区块链上，形成新的区块。

数据区块一般包含区块头(Header)和区块体(Body)两部分。区块头封装了当前版本号(Version)、前一区块地址(Prev-block)、当前区块的目标哈希值[3](Bits)、当前区块PoW共识过程的解随机数(Nonce)、Merkle根(Merkle-root)以及时间戳(Timestamp)等信息。比特币网络可以动态调整PoW共识过程的难度值，最先找到正确的解随机数Nonce并经过全体矿工验证的矿工将会获得当前区块的记账权。区块体则包括当前区块的交易数量以及经过验证的、区块创建过程中生成的所有交易记录。这些记录通过Merkle树的哈希过程生成唯一的Merkle根并记入区块头。

链式结构是指：取得记账权的矿工将当前区块链接到前一区块，形成最新的区块主链；各个区块依次环环相接，形成从创世区块到当前区块的一条最长主链；从而记录了区块链数据的完整历史，能够提供区块链数据的溯源和定位功能，任意数据都可以通过此链式结构顺藤摸瓜、追本溯源。需要说明的是，如果短时间内有两个矿工同时“挖出”两个新的区块加以链接的话，区块主链可能会出现暂时的“分叉”现象，其解决方法是约定矿工总是选择延长累计工、作量证明最大的区块链。因此，当主链分叉后，后续区块的矿工将通过计算和比较，将其区块链接到当前累计工作量证明最大化的备选链上，形成更长的新主链，从而解决分叉问题[2]。

时间戳：区块链技术要求获得记账权的节点必须在当前数据区块头中加盖时间戳，表明区块数据的写入时间。因此，主链上各区块是按照时间顺序依次排列的。时间戳技术本身并不复杂，但其在区块链技术中的应用是具有重要意义的创新。时间戳可以作为区块数据的存在性证明(Proofofexistence)，有助于形成不可篡改和不可伪造的区块链数据库，从而为区块链应用于公证、知识产权注册等时间敏感的领域奠定了基础。更为重要的是，时间戳为未来基于区块链的互联网和大数据增加了时间维度，使得通过区块数据和时间戳来重现历史成为可能。

哈希函数：区块链通常并不直接保存原始数据或交易记录，而是保存其哈希函数值，即将原始数据编码为特定长度的由数字和字母组成的字符串后记入区块链。哈希函数(也称散列函数)具有诸多优良特点，因而特别适合用于存储区块链数据。

Merkle[3]树：Merkle树是区块链的重要数据结构，其作用是快速归纳和校验区块数据的存在性和完整性。Merkle树通常包含区块体的底层(交易)数据库，区块头的根哈希值(即Merkle根)以及所有沿底层区块数据到根哈希的分支。Merkle树运算过程一般是将区块体的数据进行分组哈希，并将生成的新哈希值插入到Merkle树中，如此递归直到只剩最后一个根哈希值并记为区块头的Merkle根。最常见的Merkle树是比特币采用的二叉Merkle树，其每个哈希节点总是包含两个相邻的数据块或其哈希值。

非对称加密：非对称加密是为满足安全性需求和所有权验证需求而集成到区块链中的加密技术，常见算法包括RSA、Elgamal、Rabin、D-H、ECC(即椭圆曲线加密算法)等。非对称加密通常在加密和解密过程中使用两个非对称的密码，分别称为公钥和私钥。非对称密钥对具有两个特点，首先是用其中一个密钥(公钥或私钥)加密信息后，只有另一个对应的密钥才能解开；其次是公钥可向其他人公开、私钥则保密，其他人无法通过该公钥推算出相应的私钥。

## 2.3网络层

网络层封装了区块链系统的组网方式、消息传播协议和数据验证机制等要素。网络层通过设计特定的传播协议和数据验证机制，可使得区块链系统中每一个节点都能参与区块数据的校验和记账过程。其核心在于仅当区块数据通过全网大部分节点验证后，才能记入区块链。

组网方式：区块链系统的节点具有分布式、自治性、开放可自由进出等特性，因而一般采用对等式网络(Peer-to-peernetwork，P2P网络)来组织散布全球的参与数据验证和记账的节点。P2P网络中的每个节点均地位对等且以扁平式拓扑结构相互连通和交互，不存在任何中心化的特殊节点和层级结构，每个节点均会承担网络路由、验证区块数据、传播区块数据、发现新节点等功能。按照节点存储数据量的不同，可以分为全节点和轻量级节点。前者保存有从创世区块到当前最新区块为止的完整区块链数据，并通过实时参与区块数据的校验和记账来动态更新主链。全节点的优势在于不依赖任何其他节点而能够独立地实现任意区块数据的校验、查询和更新，劣势则是维护全节点的空间成本较高。

数据传播协议：任一区块数据生成后，将由生成该数据的节点广播到全网其他所有的节点来加以验证。现有的区块链系统一般根据实际应用需求设计比特币传播协议的变种，例如以太坊区块链集成了所谓的“幽灵协议”以解决因区块数据确认速度快而导致的高区块作废率和随之而来的安全性风险。

数据验证机制：P2P网络中的每个节点都时刻监听比特币网络中广播的数据与新区块。节点接收到邻近节点发来的数据后，将首先验证该数据的有效性。如果数据有效，则按照接收顺序为新数据建立存储池以暂存尚未记入区块的有效数据，同时继续向邻近节点转发；如果数据无效，则立即废弃该数据，从而保证无效数据不会在区块链网络继续传播。

区块链是典型的分布式大数据技术。全网数据同时存储于去中心化系统的所有节点上，即使部分节点失效，只要仍存在一个正常运行的节点，区块链主链数据就可完全恢复而不会影响后续区块数据的记录与更新。这种高度分散化的区块存储模式比基于中心化结构基础上的多重存储和多重数据备份模式具有更高的数据安全性。

## 2.4共识层

在分布式系统中高效地达成共识是分布式计算领域的重要研究问题。区块链技术的核心优势之一就是能够在决策权高度分散的去中心化系统中使得各节点高效地针对区块数据的有效性达成共识。

区块链共识层封装了多种共识机制[4]，比如：PoW共识机制、PoS共识机制、DPoS共识机制。除了这三种主流的共识机制外，在区块链的应用中也衍生出了PoW+PoS、行动证明等多个变种机制。这些共识机制各有优势，比特币的PoW共识机制凭借其先发优势已经形成了成熟的挖矿产业链；PoS和DPoS等新兴的机制则更为安全、高效。

## 2.5激励层

区块链共识过程通过汇聚大规模共识节点的算力资源来实现共享区块链账本的数据验证和记账工作，因而其本质上是一种共识节点间的任务众包过程。去中心化系统中的共识节点本身是自利的，最大化自身收益是其参与数据验证和记账的根本目标。因此,必须设计激励相容的合理众包机制，使得共识节点最大化自身收益的行为与保障区块链系统安全、有效的整体目标相吻合。区块链系统通过设计适度的经济激励机制并与共识过程相集成，从而汇聚大规模的节点参与并形成了对区块链历史的稳定共识。

比特币系统中每个区块发行比特币的数量是随着时间阶梯性递减的。创世区块起的每个区块将发行50个比特币奖励给该区块的记账者，此后每隔约4年(21万个区块)每区块发行比特币的数量降低一半，依此类推，一直到比特币的数量稳定在上限2100万为止。比特币交易过程中会产生手续费，目前默认手续费是万分之一个比特币，这部分费用也会记入区块并奖励给记账者。这两部分费用将会封装在每个区块的第一个交易(称为Coinbase交易)中。虽然现在每个区块的总手续费相对于新发行比特币来说规模很小(通常不会超过1个比特币)，但随着未来比特币发行数量的逐步减少甚至停止发行，手续费将逐渐成为驱动节点共识和记账的主要动力。同时,手续费还可以防止大量微额交易对比特币网络发起的“粉尘”攻击,起到保障安全的作用。

分配机制：比特币系统中,大量的小算力节点通常会选择加入矿池,通过相互合作汇集算力来提高“挖”到新区块的概率,并共享该区块的比特币和手续费奖励。据Bitcoinmining.com统计，目前已经存在13种不同的分配机制。主流矿池通常采用PPLNS(PayperlastNshares)、PPS(Paypershare)和PROP(PROPortionately)等机制。矿池将各节点贡献的算力按比例划分成不同的股份(Share)，其中PPLNS机制是指发现区块后，各合作节点根据其在最后N个股份内贡献的实际股份比例来分配区块中的比特币；PPS则直接根据股份比例为各节点估算和支付一个固定的理论收益，采用此方式的矿池将会适度收取手续费来弥补其为各节点承担的收益不确定性风险；PROP机制则根据节点贡献的股份按比例地分配比特币。矿池的出现是对比特币和区块链去中心化趋势的潜在威胁,如何设计合理的分配机制引导各节点合理地合作、避免出现因算力过度集中而导致的安全性问题是亟待解决的研究问题。

## 2.6合约层

合约层封装区块链系统的各类脚本代码、算法以及由此生成的更为复杂的智能合约。合约层则是建立在区块链虚拟机之上的商业逻辑和算法，是实现区块链系统灵活编程和操作数据的基础。包括比特币在内的数字加密货币大多采用非图灵完备的简单脚本代码来编程控制交易过程，这也是智能合约的雏形。随着技术的发展，目前已经出现以太坊等图灵完备的可实现更为复杂和灵活的智能合约的脚本语言，使得区块链能够支持宏观金融和社会系统的诸多应用。

# 3区块链的现存问题

尽管近年来区块链快速发展，但是仍然存在着制约其发展的若干难题有待解决。这些难题存在于安全、效率、资源、和博弈等各个方面。

## 3.1效率问题

目前效率问题主要表现在以下几个方面。

分布式记账本数据量问题。分布式记账本记录了整个区块链网络从诞生到当前时间节点的一切交易记录，在保证区块链数据不可篡改的同时，带来了存储和同步的问题。目前比特币的数据量已经超过了60GB，数据量巨大，按照比特币愈发活跃的走势来看，账本过大是一个急需解决的问题。

同步时间问题。截至目前为止，比特币网络已经有几十万个区块被开采出来，新添加进网络的节点同步账本所花费的时间就长达几天。如果没有改进方案，随着区块链的发展，新的节点代价会越来越大，甚至会阻碍区块链网络的发展。

## 3.2安全问题

安全性威胁是区块链迄今为止所面临的最重要的问题.其中,基于PoW共识过程的区块链主要面临的是51%攻击问题，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。以比特币为例，据统计中国大型矿池的算力已占全网总算力的60%以上，理论上这些矿池可以通过合作实施51%攻击，从而实现比特币的双重支付。虽然实际系统中为掌握全网51%算力所需的成本投入远超成功实施攻击后的收益，但51%攻击的安全性威胁始终存在。基于PoS共识过程在一定程度上解决了51%攻击问题，但同时也引入了区块分叉时的N@S(Nothingatstake)攻击问题。研究者已经提出通过构造同时依赖高算力和高内存的PoW共识算法来部分解决51%攻击问题，更为安全和有效的共识机制尚有待于更加深入的研究和设计。

区块链的非对称加密机制也将随着数学、密码学和计算技术的发展而变的越来越脆弱。据估计，以目前天河二号的算力来说，产生比特币SHA256哈希算法的一个哈希碰撞大约需要2年，但随着量子计算机等新计算技术的发展，未来非对称加密算法具有一定的破解可能性，这也是区块链技术面临的潜在安全威胁。

区块链的隐私保护也存在安全性风险。区块链系统内各节点并非完全匿名，而是通过类似电子邮件地址的地址标识来实现数据传输。虽然地址标识并未直接与真实世界的人物身份相关联，但区块链数据是完全公开透明的，随着各类反匿名身份甄别技术的发展，实现部分重点目标的定位和识别仍有可能。

## 3.3资源问题

PoW共识过程高度依赖区块链网络节点贡献的算力，这些算力主要用于解决SHA256哈希和随机数搜索，除此之外并不产生任何实际社会价值，因而一般意义上认为这些算力资源是被“浪费”掉了，同时被浪费掉的还有大量的电力资源。随着比特币的日益普及和专业挖矿设备的出现，比特币生态圈已经在资本和设备方面呈现出明显的竞赛态势，逐渐成为高耗能的资本密集型行业，进一步凸显了资源消耗问题的重要性。因此，有效汇集分布式节点的网络算力来解决实际问题，是区块链技术需要解决的重要问题。未来的潜在发展趋势是设计行之有效的交互机制来汇聚和利用分布式共识节点的群体智能，以辅助解决大规模的实际问题。

## 3.4博弈问题

区块链网络作为去中心化的分布式系统，其各节点在交互过程中不可避免地会存在相互竞争与合作的博弈关系，这在比特币挖矿过程中尤为明显。一般来说，比特币矿池间可以通过相互合作保持各自稳定的收益。然而,矿池可以通过“区块截留攻击”的方式、通过伪装为对手矿池的矿工、享受对手矿池的收益但不实际贡献完整工作量证明来攻击其他矿池，从而降低对手矿池的收益。如果矿池相互攻击，则双方获得的收益均少于不攻击对方的收益。当矿池收益函数满足特定条件时，这种攻击和竞争将会造成“囚徒困境”博弈结局。设计合理的惩罚函数来抑制非理性竞争、使得合作成为重复性矿池博弈的稳定均衡解，尚需进一步深入研究。此外，正如前文提到的，区块链共识过程本质上是众包过程，如何设计激励相容的共识机制，使得去中心化系统中的自利节点能够自发地实施区块数据的验证和记账工作，并提高系统内非理性行为的成本以抑制安全性攻击和威胁，是区块链有待解决的重要科学问题。

# 4区块链应用前景

从理论上说，围绕区块链这套开源体系可以创造非常丰富的的服务和产品。区块链技术将不仅仅应用在金融支付领域，而会扩展到目前所有的应用范围，例如去中心化的微信、微博、搜索等。因为区块链可以让人们无地域限制的去进行大规模协作。

未来区块链最有可能的应用领域有：一是点对点的交易。如基于P2P的跨境支付和汇款、贸易结算以及证券、期货、金融衍生产品合约的买卖等。二是登记。由于区块链具有可信、可追溯的特点，因此可以作为可靠的数据库来记录各种信息。三是确权。如土地所有权、股权等合约或者财产的真实性验证和转移等等。四是智能管理。也就是利用“智能合约”自动检测是否具备生效的各种环境，一旦满足预先设定的程序，合约会被自动处理，比如自动付息，分红等。

区块链将会驱动平行社会[5]的发展，其主要贡献是为分布式社会系统和分布式人工智能研究提供了一套行之有效的去中心化的数据结构、交互机制和计算模式，并且为平行社会提供了坚实的数据基础和信用基础。

另一方面，区块链与物联网相结合形成的智能资产使得联通现实物理世界和虚拟网络空间成为可能，并可通过真实和人工社会系统的虚实互动和平行调谐实现社会管理和决策的协同优化。可以预见，未来现实物理世界的实体资产都登记为链上智能资产的时候，也就表明区块链驱动的平行社会已经到来。

最后，我们要辩证的看待区块链的未来。区块链的发展前景现阶段还不太明朗，一方面我们应该采取宽容的态度，积极地投身于区块链的研究与应用之中；另一方面，我们也要保持理性和冷静，对于区块链的发展要有自己的独立思考。唯有这样，我们才能更好的理解和把握区块链，从而为区块链乃至社会的发展进步贡献力量。

# 参考文献

[1] The Byzantine Generals [EB/OL].http://research.microsoft.com/en-us/um/people/lamport/pubs/byz.pdf.

[2] Antonopoulos A M. Mastering Bitcoin: Unlocking Digital Cryptocurrencies. USA: O Reilly Media Inc., 2014.

[3] Merkle R C. Protocols for public key cryptosystems. In: Pro- ceedings of the 1980 IEEE Symposium on Security and Pri- vacy. Oakland, CA, USA: IEEE, 1980. 122.

[4] Larimer D. Delegated proof-of-stake white paper [Online], <http://www.bts.hk/dpos-baipishu.html,>2014.

[5]王飞跃，王晓，袁勇，王涛，林懿伦.社会计算与计算社会:智慧社会的基础与必然.科学通报,2015,60(5−6):460−469.

# 

# 致谢

本课程报告是基于梁教授所授的“高级数据库”为基础完成。梁老师丰富的人生阅历，渊博的专业知识，严谨的治学态度，热情洋溢讲课方式，让我获益匪浅。不仅使我对数据库这一门课程有了更深的理解，而且扩展了我的知识视野，更重要的是让我对武大，对做人和求学有了不一样的理解，非常感谢梁老师。