区块链概论

李斌

武汉大学金融系 2021年9月17日

账本

- 资产 -> 交易
- 交易涉及不同参与方,其商业协议和合同记录于账本(Ledger)

账本Ledger					
账户类型	现金				
TRANSACTION DATE	TRANSACTION DETAIL	REFERENCE	DEBIT	CREDIT	BALANCE
1/1/16	Expenses for Jan	Ref#1	\$100.00		\$100.00
2/1/16	Tax withheld	Ref#2		\$110.00	(\$10.00)

中心化账本面临的问题

- 数据的篡改、抵赖等问题依然无法解决,并且随着互联网的发展,日益突出
- 2010年,一名黑客入侵北京教育考试院的网上证书查询系统,根据受益人的 授意登陆后台数据库,篡改记录,增加了一条关于受益人的证书信息
- 2016年,某公司北京亦庄数据中心系统宕机,73家村镇银行的核心、银行卡、柜面、支付、网银、手机银行等业务全部中断,涉及全国12个省份,并造成部分服务器损坏,银行业务最长恢复时间达到7小时32分钟

.....







中心化账本的弊端

- 无法防止中心内部人篡改信息
- 遭受黑客攻击后无法避免信息丢失与篡改
- 需要建立大规模实名信用系统
- 容易泄露用户隐私
- 防伪成本不断上升, 日常维护要求高
- 覆盖范围有限

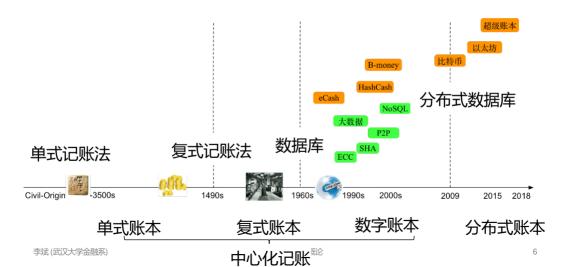
•

李斌(武汉大学金融系) 金融系 金融科技: 区块链概论 4

区块链解决了什么?

- 去中心化
 - 解决了中心化的弊端
- 制造信任
 - 通过机器或者算法实现
 - 巧妙利用所有节点的自利实现了总体的完全可信任
- 创造价值
 - 节约了中心化信息传输过程中的大量成本
 - 可靠的信息共享提高协同效率
 - 区块链+

记账的历史

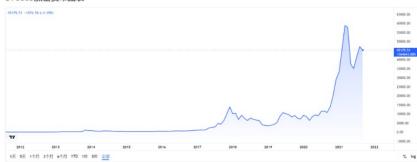


区块链的定义

- 区块链是一种按照时间顺序将数据区块以链式相连的方式组合成的数据 结构,并以密码学方式保证的不可篡改和不可伪造的分布式账本
- 起源:中本聪于2008年发表的论文《比特币: 一种点对点电子现金系统》 09年比特币问世 BTCUSDIN®货币图表



李斌 (武汉大学会融系)



区块链的主要类型

	参与者	记账人	中心化程度	代表项目
公有链	任何人自由进出	所有参与者	去中心化	加密货币
联盟链	联盟成员	联盟成员协商确定	弱中心化	供应链金融、 银行、物流、 电商
私有链	链的所有者	链的所有者	强中心化	大型组织、机 构

区块链的五大特征

- 去中心: 区块链网络中不存在中心化节点,各节点高度自治,具有相等的权利和地位
- 不可篡改: 使用密码学和块链结构来保证区块链上的信息不被篡改
- 可追溯:保存了第一个区块开始的所有历史数据,区块链上任意一条 记录都可以通过链式结构追本溯源
- 开放性: 将对第三方机构的信任转化为对机器代码的信任
- 匿名性: 采取密码学手段, 在实现数据开放的前提下, 保护交易隐私

班聚报名

在一个 50 人的微信群里组织聚餐,如何统计人数?

• 方案一: 大家各自向组织者发信息报名

• 方案二:群里接龙报名(区块链思维)

• 张三: 1. 张三

• 李四: 1. 张三 2. 李四

• 王五: 1. 张三 2. 李四 3. 王五

• : 1. 张三 2. 李四 3. 王五 4. 5.

班聚报名

• 在一个 40 人的微信群里组织聚餐,如何统计人数?

微信群组饭局	区块链的概念
接龙报名	链式数据结构
规则:复制上一条信息+下一条编号+自己名字	根据严格的规则和公开的协议形成的共 识机制
规则定下来后,大家各自报名	去中心化,无单一用户控制名单
每条报名消息都推送至各个成员的手机	分布式(多点备份),高冗余
每个人群成员都可以查看所有报名信息	共享账本
如果出现分叉,少数服从多数,以跟帖多的分支为准	最长链规则

大纲

- 1. 区块链原理
 - •加密学基础
 - 区块与链式结构
 - 共识机制
 - •智能合约
- 2. 区块链金融

如何设计去中心化的账本/货币?

•解决办法

- 多个用户使用公共账本记录转账记录, 月底结算, 账本公开
- 每个人都可以修改, 比如添加新行: 小明转账给小红10块钱

•核心问题

- 1. 账本数据的安全问题 『加密与签名
- 2. 账本数据的存储问题 『区块与链式结构
- 3. 账本放在哪里 『分布式与共识机制

核心需求1:账本数据的安全问题

- •恶意用户无法仿冒其他用户发起交易,发起交易后无法抵赖
 - 如何不让人看到你的账本数据? -> 加密
 - 如何证明收到的账本数据是你的? -> 签名

李斌(武汉大学金融系) 金融科技:区块链概论 14

数据的安全:对称VS非对称加密

• 对称加密:一个密码,用于加密和解密

• 发信人: 加密 (明文, 密码) = 密文

• 收信人:解密(密文,密码) = 明文

• 非对称加密:公开密钥(公钥)和私有密钥(私钥)

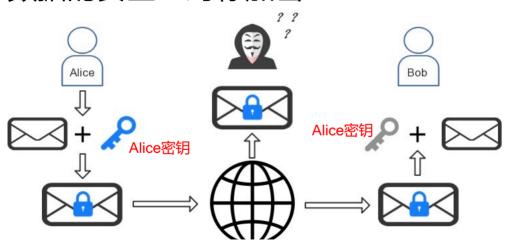
• 用公钥对数据加密,用对应的私钥才能解密

• 用私钥对数据签名,用对应的公钥才能验证

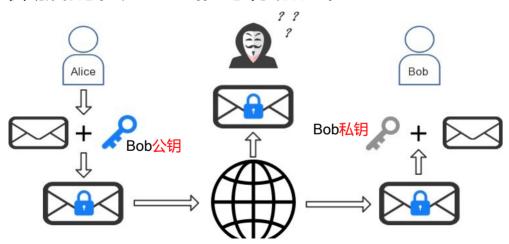
发信人: 加密 (明文, 收件人公钥) = 密文

收信人:解密(密文,收件人私钥)=明文

数据的安全:对称加密



数据的安全:非对称加密



数据的安全:哈希算法

- 哈希算法:将任意长度的二进制原文串映射为固定长度的二进制串(哈希值) 也叫散列算法、摘要算法
 - 哈希是单向函数
 - 哈希不是加密, 因为你不能解密
 - 不管输入值有多大, sha256 的输出均为 256 位
- 展示: SHA256 算法, https://cryptii.com/pipes/hash-function

message	SHA-256
fintech	110599 ccb 008 c6746985524 be 4 ba 99 ff 588 fed 6 bc 5 c4d 65d 2a 15b 6a 92 eda 16a 7a be 2b 6a 92 eda 16b 6a 9
fintecn	79c747c2e0a689714a9d4b75ee14165112670199b9359be683935dac1425fdb9

李斌 (武汉大学金融系) 金融科技:区块链概论 18

数据的安全:哈希算法示例

哈希1:除留余数法

键	哈希值	哈希值	 键	键平方	散列值
琏	(M=100)	(M=97)	阺	(進十万)	HXYJIE

哈希2:平方取中法

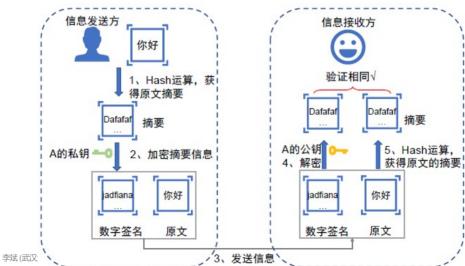
键	哈希值	哈希值	键	键平方	散列值	
₩ E	(M=100)	(M=97)	延	延十刀	FIX 7 JIEL	
212	12	18	0100	0 <u>010</u> 000	010	
618	18	36	1100	1 <u>210</u> 000	210	
302	2	11	1200	1 <u>440</u> 000	440	
940	40	67	2061	4 <u>247</u> 721	247	

数据的安全:哈希算法的特点

- 正向快速
- 逆向困难: 从输入计算输出比较容易, 但反之则几乎是不可能的任务
- 输入敏感:原始输入信息发生任何改变,新产生的哈希值会发生较大变化
- 碰撞避免:很难找到两段内容不同的原文,使他们的哈希值一致,即发生碰撞。有限资源内,找到一个碰撞是不可行的。

•猜测:SHA256 哈希中有多少组合?

数字签名:纸上手写签名的数字模拟



比特币:地址

- 通过私钥可以计算出公钥
- 通过公钥可以计算出比特币地址(双哈希 +Base58Check 编码)
- 上述过程均不可逆



比特币:匿名性

- 在比特币中,公钥是你的用户名(也叫 ID,身份)
- 确定性
 - 私钥和公钥成对产生
 - · 如果要使用某个身份/ID, 就必须用到配对的私钥
 - 私钥丢了无法确认身份
- 匿名性
 - 私钥和公钥是随机产生的
 - 你可以有多个用户名,没人知道你是谁(不可逆)
- 去中心化的用户管理
 - 无需向中心注册, 只需要随机生成一对私钥和公钥即可

- Stefan Thomas曾制作过一个加密货币的解释视频,并获得了7002枚比特币(今天超过3亿美元)的打赏,而Thomas也在此后丢失了他的私钥。
- 据报道,他已经尝试过8次 密码,而在第10次尝试失败 后,这些财富将永远"加密"

李斌 (武汉大学金融系) 金融系) 金融系技:区块链概论

课后作业:比特币

- •注册一个比特币钱包
- 查看一笔比特币交易,结构是怎么样的?
- https://www.blockchain.com/

Summary 0

BTC

This transaction was first broadcast to the Bitcoin network on September 14, 2021 at 3:45 PM GMT+8. The transaction is currently unconfirmed by the network. At the time of this transaction, 0.34886765 BTC was sent with a value of \$15,929.04. The current value of this transaction is now \$15,922.02. Learn more about how transactions work.

Hash

86e08696b821d4246514b7779d0d082d16ecef6429c5edd9151...

2021-09-14 15:45

bc1q8p825fm4tkye3k5d5d6lcfkqwzkthkeui89rsq 0.34886911 BTC #



35utFvrR4Sui3rMdqkK4dwQTmYnSG8qaWk 1DK1LDtzVH88HWVxqktkEzwXqAZczQGFhP 0.34870765 BTC (#)

0.00016000 BTC (III)

Fee

0.00000146 BTC (0.646 sat/B - 0.253 sat/WU - 226 bytes) (1.007 sat/vBvte - 145 virtual bvtes)



核心需求2:账本数据如何存储

- •如何防止篡改账本数据? [区块与链式结构
- 账本数据太大如何存储? 『梅克尔树

区块与链式结构

• 区块: 区块链的基本构成单位



李斌 (武汉大学金融

区块与链式结构

- 区块链:将区块链接构建一个链表
- 每个区块既告诉我们上一个区块的值在哪里,还包括了该值的哈希值, 使我们能够验证值有没有改变



27

区块与链式结构:防篡改特性

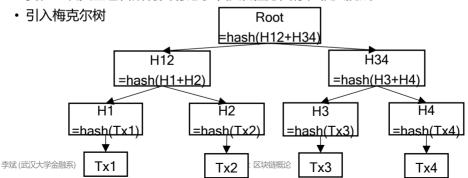
- 如有人修改了区块链中的任意数据,必须重置后续区块的所有哈希值
- 如果我们锁定区块链的头部,就可以检测到篡改行为

	上个区块的哈希值	区块2的哈希值	区块3的哈希值	区块4的哈希值		上个区
略	记录: 甲转给乙6个币 丁转给丙6个币 刘转给丁6个币	记录: 内转给丁6个币 乙转给李6个币 丁转给丙6个币	记录: 李转给陈6个币 乙转给丙6个币 丙转给丁6个所 区块4	记录: 大转给李6个币 丁转给丙6个币 乙转给陈6个币	略	刘转 陈转 丁转

上个区块的哈希值 记录: 刘转给李6个币 陈转给丙6个币 丁转给刘6个币

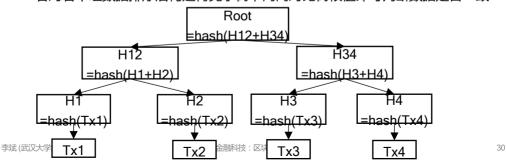
数据的存储:梅克尔树

- 交易记录的累积将需要越来越多的存储空间
 - 2021年3月比特币的交易记录需要334GB的存储空间
 - 要在区块头里包含所有交易记录,扩展性方面存在很大挑战



数据的存储:梅克尔树

- 作用1: 快速比较大量数据
 - Tx1的数据有变化,则H1->H12>Root 都会有变化
 - 若两颗梅克尔树的树根值相同,则所有数据相同;否则必然不同
 - 若对若干组数据排序后构建梅克尔树,两两对比树根值即可判断数据是否一致

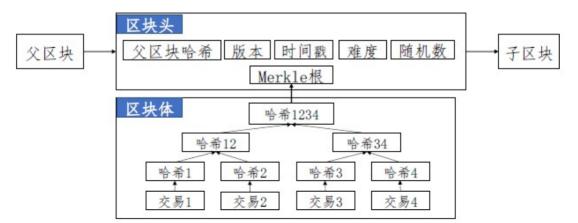


数据的存储:梅克尔树

- 作用2: 简易支付验证
 - 轻节点希望验证交易数据Tx2的真实性,向其他全节点发送请求
 - 全节点发送log₂ N个哈希值(H1, H12; N 为交易总数)
- 计算得到Root哈希,与自身保存的Root哈希比较 Root hash(H12+H34) H₁₂ **H**34 =hash(H1+H2) =hash(H3+H4) H2 H3 H4 =hash(Tx1) =hash(Tx2) =hash(Tx3) =hash(Tx4) 区块链概论 李斌(武汉大学金融系) T_x1 Tx2 Tx3 Tx4

数据的存储:比特币的区块链结构

• 全节点保存所有交易;轻节点保留梅克尔根与自己相关的交易



去中心化

- •核心需求3:账本放在哪里才能去中心化?
 - 简单的做法:每个用户保存账本,分布记账;用户产生一笔交易就广播到网络所有的节点上(P2P)
 - 问题:如何让所有人都统一这个账本?如何保持这些账本的同步?一笔交易发生时,如何让其他人都听到并相信这笔交易?



李斌(武汉大学金融系)

图2-2a) P2P网络模式

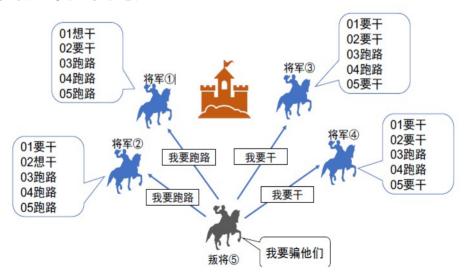
图2-2b) 中心化网络模式

拜占庭将军问题 (Byzantine Generals Problems)

- 拜占庭帝国派出5位将军围攻一座城市,这5位将军之间只能通过信使 联系,只有一起进攻才会胜利
- 于是将军们在出发前商量了进攻的策略,少数服从多数,当超过3位及以上同意进攻,则5个人一起进攻,否则就一起撤退
- 但是,如果5个人之中出现一个叛徒,就可能导致最终的行动不一致

拜占庭将军问题

李斌(



35

拜占庭将军问题

- 拜占庭将军问题:几个相互协同的人(系统),如果其中有个人作恶的话,可能不同的成员得出的最终结论不一致,从而破坏了系统的一致性。
- 区块链是一套去中心化的分布式系统,既然是分布式就意味着全网多台机器节点进行协同合作。
- 每个节点相当于拜占庭将军,最终要共同维护一套数据。但是:
 - 无法保证节点诚实(单点一致性)
 - 无法保证系统内部信息统一

拜占庭将军问题:中本聪的解决方案

- •加入工作量证明(Proof of Work, PoW),提高做叛徒的成本。
- •比如,拜占庭将军们可以约定工作量:
 - 收到进攻信息后,大家共同完成一项需要花一点时间的任务
 - •如:脱鞋,右脚持毛笔用楷书抄一首诗(4份,需要抄一天)
 - 第一个抄完的人可以发布进攻时间
 - •没有抄完的人--作废
 - 右脚抄写的诗可以被验证(非对称加密技术)
 - 时间无法被篡改(时间戳)

李斌 (武汉大学金融系) 金融系 金融系 (武汉大学金融系) 金融系 (武汉大学金融系)

比特币的工作量证明

•比特币的工作量证明:对"Hello?"(?为随机数)进行哈希,要求哈希值前6位为0。通过变化前置0的个数,可以调整工作量的大小

```
\label{eq:hash256} $$ \ hash256("Hello1") = ffb7a43d629d363026b3309586233ab7ffc1054c4f56f43a92f0054870e7ddc9 $$ \ hash256("Hello2") = e085bf19353eb3bd1021661a17cee97181b0b369d8e16c10ffb7b01287a77173 $$ \ hash256("Hello3") = c5061965d37b8ed989529bf42eaf8a90c28fa00c3853c7eec586aa8b3922d404 $$ \ hash256("Hello4") = 42c3104987afc18677179a4a1a984dbfc77e183b414bc6efb00c43b41b213537 $$ \ hash256("Hello4") = 42c3104987afc18677179a4a1a984dbfc77e183b414bc6efb00c43b41b213537 $$
```

....

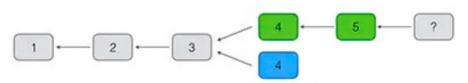
hash256("Hello15583041") = 0000009becc5cf8c9e6ba81b1968575a1d15a93112d3bd67f4546f6172ef7e76

拜占庭将军问题:单点一致性

- 单点一致性问题:一个节点可能同时向不同的服务器发送不一致的消息,导致节点之间存储的信息不一致
 - 工作量证明 PoW 可以保证在一个时间只有一个节点(或很少)在进行广播,同时在广播时会附上自己的签名
 - 这就降低了节点发送消息的速率,保证一段时间内,大部分节点收到的是一条一样的消息

拜占庭将军问题:最长链原则

- 系统一致性问题: 如果同时接收到两条不同的区块链怎么办?
 - 最长链原则: 当网络中同时达到出现分叉时,根据全体参与者(算力) 的选择,主网会默认累积难度值最大的为主链
 - 意味着: 少数人永远服从于多数人的共识



拜占庭将军问题:六次确认

比特币网络的六次确认要求大大提高系统的不可篡改性:

- 为了防止在比特币网络中出现信息篡改情况的发生,需要每条消息经过6次确认没有更改之后才认为有效。
- 因为如果想更改一条已经确认的消息,需要正确算出6道难题,然后还需要保证自己的更改之后的消息链条为最长链。
- 这会消耗大量的成本,导致篡改数据的成本高到无法承受,最大程度 地保证了系统不会出现确认过的消息前后不一致的问题

典型的共识机制

- 权益证明 (PoS, Proof of Stake): 各记账人根据自己的权益按算法竞争记账权。
 - 将 PoW 中的计算能力替换为权益证明,即节点所拥有的的币龄(币数量乘以天数)
 - 币龄越长,获得记账的概率就越大
- 共识机制中的激励:挖矿成功者有激励

传统合约

- 传统合约面临的问题:捐款募资?
- 其有效执行显然受多方面因素影响:
 - 执行条件的主客观性、执行时间、自动化程度、惩罚力度等

• 如果把执行合约的代码放到区块链上,会如何?

智能合约 (Smart Contract)

- 智能合约是一套以数字形式定义的承诺,包括合约参与方可以在上面执行这些承诺的协议 (Nick Szabo)
- 智能合约是一种设定好的计算机程序,能够在没有第三方中间机构参与的情况下,自动执行可信的合约内容,所有操作公开可查且不可逆转
- 比特币是区块链1.0时代的开端,智能合约则是区块链2.0的代表性产物
- 最早 1994 年由 Nick Szabo 提出的理念,但由于缺少可信的执行环境,智能合约并没有被应用到实际产业中

```
<code style="font-family:Menlo, Courier, monospace, monospace, sans-serif;for 复制
      address public organizer:
      mapping (address => uint) public registrantsPaid:
      uint public numRegistrants:
      uint public quota;
      event Deposit(address from, uint amount); // so you can log these events
      event Refund(address to, uint amount);
                                                                                      过过
      function Conference() { // Constructor
        organizer = msg.sender:
        quota = 500;
        numRegistrants = 0;
      function buyTicket() public returns (bool success) {
        if (numRegistrants >= quota) { return false; }
                                                                                      智能
        registrantsPaid[msg.sender] = msg.value;
        numRegistrants++;
        Deposit(msg.sender, msg.value);
21
22
      function changeQuota(uint newquota) public {
        if (msg.sender != organizer) { return; }
                                                                                           45
        quota = newquota;
```

智能合约的特点

- 1. 智能合约被自动执行,很难作弊;
- 2. 交易的结果将是all-or-nothing,当然随着智能合约的进化,当更加智能、复杂的合约出现后,可能会改变all-or-nothing的结果;
- 3. 合约有可被审计的历史记录—所有的数据被存储在区块(共享分布式账簿)中,不会丢失。

与传统合约进行的比较

	传统合约	智能合约
格式	特定语言+法律术语	代码
确认和同意	签字、盖章	数字签名
争议解决	法官、仲裁员	仍在探索中,比如EOS就设立了仲裁论 坛和仲裁小组
效力	法院或仲裁机构	可通过法院或仲裁机构
执行效率	低	高
付款和执行	根据合约约定,可能依赖可信赖的第三方	合约约定并自动执行

李斌(武汉大学金融系)

古

金融料技:区块链概论

/工

智能合约:智能捐赠/数字慈善

- 我希望向周围的朋友募资1万元捐给希望工程
- 定义两个状态变量: 当前募集总量和被捐款处的地址
- 再定义两个函数:
 - ① 接受募捐:每次发起捐款请求时,先检查捐款人的账户余额是否足够,接着判断当前募集总量加上这笔捐款额是否达到1万元。达到就运行捐款函数,否则就更新募集总量。
 - ② 捐款:将所有款项转账给收款人地址,并清空当前募集总量。
- 合约执行条件是募集资金达到1万元,执行内容是全部转账给灾区的收款人地址
- 将该合约部署到智能合约服务器, 把合约发到区块链上生效。

智能合约:智能捐赠/数字慈善

- 1. 区块链信息可追溯,捐赠物资分发和配送过程公开透明
- 2. 智能合约自动执行,有效解决传统公益项目中繁琐的流程和"人工"操作
- 3.利用区块链是共享式账本这一属性,打破数据信息孤岛,公开捐赠数据及分配数据,在每个流程都记录相关数据信息,做到实时可追查数据信息

大纲

- 一. 区块链原理
- 二.区块链的应用

区块链的应用价值

• 提升协同效率

• 促进数据共享

- 优化业务流程
- 降低运营成本
- 助力穿透监管
- 建设可信体系

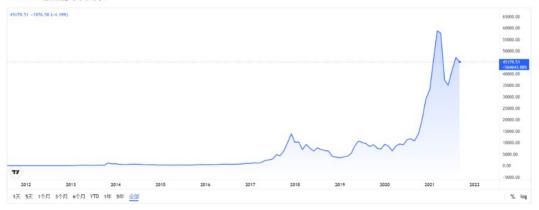


区块链的第一个成功应用:比特币

- 比特币技术上实现了无需第三方中转或仲裁,交易双方可以直接相互 转账的电子现金系统。
- 相比实体货币,数字货币具有易携带存储、低流通成本、使用便利、 易于防伪和管理、打破地域限制,能更好整合等特点。
- 虽然中国对于加密货币本身持慎重和反对态度,但是政府和企业对于加密货币所运用的区块链技术展现出浓厚的兴趣并且支持技术发展。
 - 2014年起,中国央行就开始致力于DCEP (Digital Currency Electronic Payment)的研究,俗称数字人民币。

比特币因其价格过山车般波动而备受瞩目

BTCUSD加密货币图表



比特币 — 一个"大账本"

•问题1:记什么?——交易记录

•问题2:谁来记?——挖到矿的人

•问题3:怎么记?——区块上链

•问题4:怎么保证正确性?——非对称加密

•问题5:账本放在哪?—去中心化,全网可见

问题1:比特币"账本"记录的是什么?

- 比特币的账本采用UTXO(Unspent Transaction Output)模型,即交易 记录本身形成账本,而不是账户信息
- 每个账户对应一个地址(类比银行卡),一旦查询这个地址的余额, UTXO就会跟踪地址前后所有记录给予"余额",这个跟踪计算是由 比特币钱包完成的
- 用户用密钥来签名交易,从而证明他们拥有比特币;比特币的所有权 是通过数字密钥、比特币地址和数字签名来确定的,而不是增加自己 账户里的数字

• 比特币钱包就是储存密钥的数据结构



资料来源: 木蚂蚁论坛

问题2:谁有记账权?—挖到矿的人

- 挖矿: 最快算出SHA-256(上区块哈希值,此区块梅克尔树根值, Nouce随机数)
 - = 前X位为0的节点,便可获得记账权并获得报酬,X用于控制计算难度
 - 初期算法难度低,仅用普通计算机即可获得比特币激励;后来难度逐渐加大,开始出现专门为挖矿而设计的矿机。
 - 记账权奖励:基础产出(最初为50比特币,每四年减半)+该区块所有的手续费 (手续费越高,越优先进入区块)



总量=210000 X (50+25+12.5+6.25+.....) 等比数列, a1=50, q=1/2



金融科技: 区块链概论

问题3:怎么完成"记账"?

- 假设A用户(地址)向B用户(地址)汇款,则该信息形成一个交易记录
- 节点根据自己的策略和记录的手续费选取不同的记录,验证电子签名 为真后将各汇款信息哈希成一个256bit长度的字符串,然后把相应的字 符串整合到一个区块上,采用默克尔根结构得到该区块的哈希值
- 完成工作证明后便可以将这个自己的区块添加到主链上

问题4:如何保证记录的正确性?

数字签名:通过非对称加密算法对签名信息进行处理,私钥处理信息,然后任何人可以使用公钥来核实此记录的真实性



问题5:账本放在哪里?

• 去中心化网络,全网可见



Buy Bitcoin

Blockchain.com

进一步学习

- 专业选修课:区块链技术与应用
- •柴洪峰、马小峰,区块链导论,中国科学技术出版社,2020。
- https://www.bilibili.com/video/av12465079/
- 肖臻老师区块链课程:http://zhenxiao.com/blockchain/

Q & A

【腾讯文档】金融科技课程意见调查



- 李斌, 武汉大学金融系
- binli.whu@whu.edu.cn