

数字货币和区块链

- 共识与区块链

山东大学网络空间安全学院

钱宸 2023年12月13日

区块链与共识协议

- 简介与历史
- 拜占庭广播与Dolev-Strong协议
- 拜占庭广播（无签名）
- 区块链
- 区块链协议

中本聪的区块链

- 区块链技术首先被用于分布式系统以及飞机控制系统
- 在2009年当区块链被用于比特币以后，获得了广泛的关注和发展
- 其中最核心的就是中本聪的区块链系统
- 不仅仅是数字货币的尝试，从理论角度也是突破：
 - 首个在未经许可的环境中（permissionless environment）中达成共识的区块链

未经许可环境中的区块链

- 未经许可环境中的区块链系统很久以来被认为是不存在的
- 难点：
 - 女巫攻击（Sybil Attack）：因为在未经许可环境中，所以
 - 攻击者可以伪装成任意他人
 - 攻击者也可以伪装成很多人
- 突破：
 - 想法：通过工作量证明（Proof of Work）来避免女巫攻击
 - 每一次投票都要通过算力来实现。系统中多数算力是诚实的，则能保证系统的一致性和活跃性

中本聪区块链

- 每个用户都拥有一个链 (chain)
- 区块：
 - 包含首个区块
 - $chain[i] = (h_{-1}, \eta, txs, h)$
 - h_{-1} 是之前一个区块的哈希
 - η 是一个难题的结果
 - txs 是许多交易的集合
 - h 是当前区块的哈希值

一些简单的记号

- $chain[-l]$: 倒数第 l 个区块
- $chain[:l]$: 从第0个到第 l 个区块
- $chain[: -l]$: $chain[0, \dots, k - l]$, 其中 k 是当前区块
- $|chain|$: 链的长度
- $_$: 用来表示任意字符
- 中本聪区块链的第0区块:
 - “The Times 03/Jan/2009 Chancellor on brink of second bailout for banks”
 - “财政大臣正处于第二次银行救助边缘”

区块链

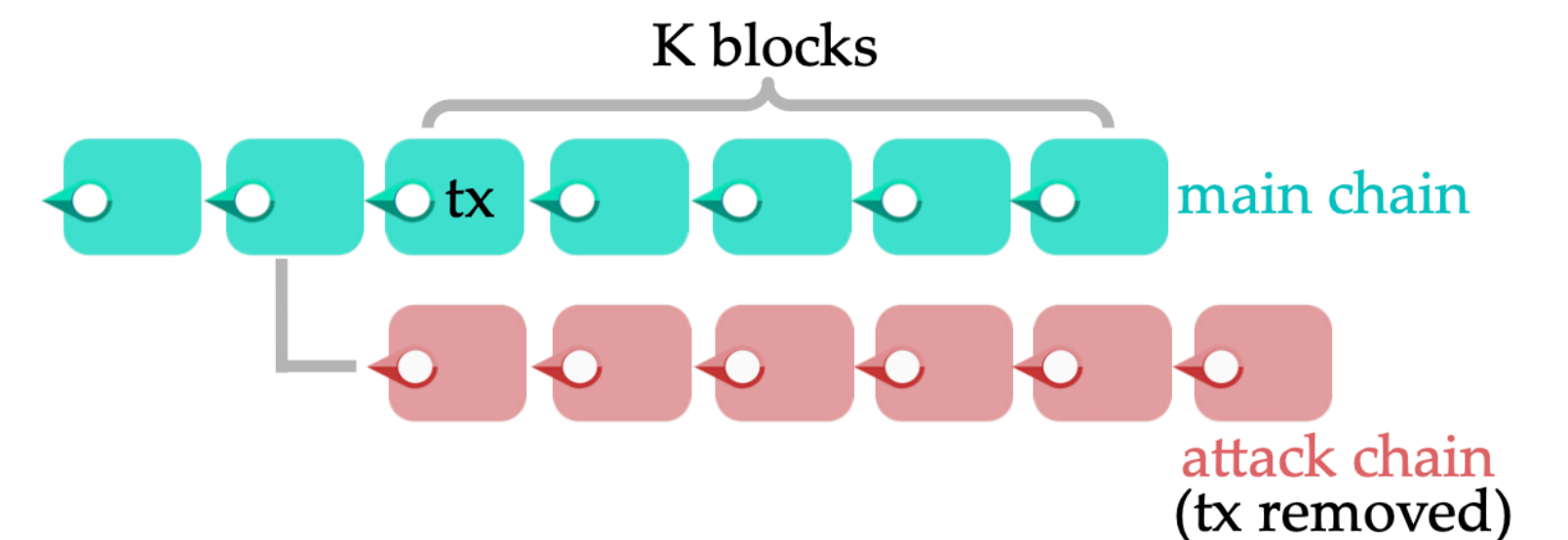
- 挖矿：
 - 对于一个区块链 $chain$ 而言，令最后一个区块是 $(_, _, _, h^\star)$
 - 如果想要“挖”一个新的区块
 - 计算 η 使得 $H(h^\star, \eta, txs) < D_p$
 - 其中 H 为一个工作证明寓言机，通常用一个哈希函数来实现
 - D_p 是一个难度参数， D_p 越小难度越大
 - 如果上述条件满足，则 $(h^\star, \eta, txs, H(h^\star, \eta, txs))$ 为一个新的区块

一些注释

- 假设哈希函数H像一个随机寓言机
 - 计算哈希函数H的原象，不能利用代数结构
 - 也就是说计算原象只能够通过穷举的方式
 - 另外，因为哈希函数H满足抗碰撞性：
 - 所以每一个区块都与整条链唯一绑定

最长链

- 中本聪的最重要的思想之一：
 - 总是选取最长链作为确认的链
 - 无论何时 $chain[: - K]$ 被确认
 - 如果一个区块在最长链的倒数K区块以上，那么就确认该区块
 - 如果攻击者想要更改某个区块的数值，则需要伪造K个区块。
- 为什么安全？



中本聪区块链严格定义

- 我们这里简单假设总算力是提前知道的，在实际比特币系统中，挖矿的难度会根据算力进行调整。
- 同样，我们假设在一个同步网络当中。其中诚实用户的信息能够在 Δ 时间后被收到。
- 对算力证明建模：我们使用 $(H, H.ver)$ 其中 H 为算力证明中的工作函数， $H.ver$ 被用来证明计算出来的结果是正确的。
- 我们假设每个计算节点都拥有同样的算力，不同算力的节点可以被拆分成很多的节点。

中本聪区块链

- 第一个区块是 $(0,0, \perp, H(0,0,\perp))$
- 如果一个节点收到 $chain'$ 是一个被验证的链，并且是目前接收到的最长链 $chain$ 的延长。则将 $chain$ 替换成为 $chain'$
- 令 $chain[-1] = (_, _, _, h_{-1})$ ，随机选取 η ，计算 $h = H(h_{-1}, \eta, txs)$ ，如果 $h < D_p$ 则将 $chain \parallel (h_{-1}, \eta, txs, h)$ 发送给所有节点。
- 在任何时间点，每个节点所确认的链都是 $chain[: -K]$

如何选取难度参数

- 比特币中，难度为所有人计算10分钟可以获得下一个区块
- 但是这个时间太长了，通常使用 $K=6$
- 但是这个参数也不能太小：
 - 诚实的节点包含网络延迟 Δ
 - 攻击者可能并不需要受到延迟的困扰

如何选取难度参数

- 简单的估计（不严谨）：
 - p 是每一轮中一个节点计算一个区块的概率（这个概率非常小）
 - 那么所有诚实节点（51%）加在一起能够计算一个节点的概率是
 - $1 - (1 - p)^{0.51n} \approx 0.51pn \ll 1$ ，也就是说所有诚实节点预计 $\frac{1}{0.51pn}$ 产生一个新的区块
 - 假设所有 Δ 轮中所有诚实节点都没有成功，那么打折率是
 - $\frac{\frac{1}{0.51pn}}{\frac{1}{0.51pn} + \Delta} \approx 1 - 0.51pn\Delta$
 - 诚实节点的算力在打折率 $1 - 0.51pn\Delta$ 的情况下，仍然要超过攻击者算力

如何选取难度？

- 选择 $p \in (0,1)$
 - 当 p 固定以后可以选择 $D_p = p \cdot 2^\lambda$
- 假设 q 是攻击者的比例
 - 那么 $(1 - q)(1 - 2pn\Delta) \geq (1 + \varepsilon) \cdot q$
 - 这里2和之前非正式的计算中0.51不一样，但是2才是满足严格证明所需要的参数

中本聪区块链的性质

- 链增长的下界：
 - 在固定时间内，诚实用户的链会增长多少
- 链质量：
 - 对于任意连续 K 个区块，有多少是有诚实用户获得的
- 一致性：
 - 诚实用户的倒数 K 个区块之前一定是别的诚实用户的链的前缀

链增长的下界

- 令 $\alpha = (1 - q)np$ 是每一轮预计有多少诚实用户挖到区块
- 但是，因为有延迟所以我们要乘上打折率
 - $(1 - 2np\Delta)\alpha$
- 为什么我们关心链增长率？
 - 因为活跃性和链增长率高度相关
 - 但是，单独增长率还不够，因为攻击者也可以挖矿，可以制造一些无用的区块。人为降低活跃性。

链质量

- 对于任意连续 K 个区块，有多少是有诚实用户获得的
 - 非常粗略的估计，诚实用户的算力占比是 μ ，那么大约有 $K \cdot \mu$ 个区块是由诚实用户所产生的

一致性与活跃性

- 中本聪区块链的一致性：
 - 只关注最长的一条链的倒数 K 个区块以前的链是否互为前缀
- 活跃性：
 - 如果一个诚实的用户收到一个交易了以后，会在多久以后被所有诚实用户所接收到

数字货币与区块链整个回顾

- 中心化数字货币
- 共识算法
- 区块链