北京大学肖臻老师《区块链技术与应用》公开课笔记

> 比特币共识协议篇，对应肖老师视频：click here 全系列笔记请见：click here **About Me:**点击进入我的Personal Page

数字货币中经常出现的问题

- 双花攻击 数字货币本身为带有签名的数据文件，可以进行复制。即：对用户来说，可以将同一货币花费两次。

> 修改：对货币添加唯一编号（不可篡改），每次支付向货币发行单位查询真伪。 该方法每次交易都需要依赖于第三方机构来判断货币真伪且防止双花攻击。是一个典型的第三方中心化方案。现实中，我们通过支付宝、微信、信用卡等各种支付方式交易时，必然会依赖于第三方机构。由于这些第三方机构具有较高的可信度，有政府进行背书，所以可以采用这种方案。 但是，很多场景下，并不存在这样一个可信赖的第三方机构。基于这个背景，以去中心化思想为核心的比特币系统便吸引了人们的注意力。

去中心化需要解决的问题

- 数字货币的发行由谁执行？如何发行？发行多少？什么时候发行？ 在传统中心化货币体系中，这些问题我们可以交给第三方机构（如：央行）。当引入去中心化思想后，系统中节点平等，交易不通过第三方，那么货币发行权的分配必然是一个需要解决的问题。

> 在比特币系统中由挖矿来决定货币发行权和发行量。

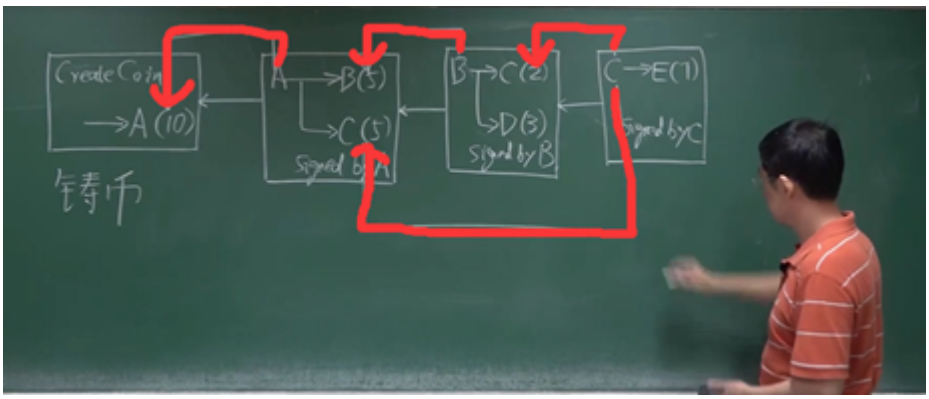
- 如何验证交易是否有效？如何防止双花攻击？ 同样，在传统中心化体系中，该问题的解决由第三方机构来完成。而剔除这一机构后，交易双方如何能够验证交易的有效性？如何防止系统中恶意用户作恶获取收益？这也是去中心化交易系统需要解决的问题。

> 该问题的解决，依赖于系统中维护的一个数据结构，记录货币的使用情况（是否被花过？被谁花过？）。该数据结构由系统中全体用户共同维护，保证了交易的有效性。该数据结构，便是区块链。

案例说明

如下，假定A获得**铸币权**，新新发布了10个比特币（该交易称为铸币交易）。A将10个比特币转给了B(5个)和C(5个)，A对该交易进行签名，同时该交易需要说明所花掉10个比特币来源（来自铸币交易）。之后，B将自己的5个比特币转给C(2个)和D(3个)，该交易需要B的签名，该交易需要说明所花掉的5个比特币来自于第二个交易中。然后，C将自己所拥有的全部7个比特币都转给E，并对该交易签名，可以发现该交易中C的比特币来源于两个交易中。这样，

就构成了一个简单的区块链。**【红色部分为比特币来源】**



> 需要注意的是，这里面有两种哈希指针。第一种为指向前面的区块（白色），使得各个区块形成链，第二种则是为了说明比特币的来源（红色）。说明比特币的来源并非凭空捏造，可以防止双花攻击。 在进行交易时，需要付款人的签名和收款人的地址，在比特币系统中，该地址即为收款人的公钥的哈希。可以将其视为银行账户，根据此进行转账交易。（虽然公钥可以公开，但实际中更多公开的是公钥的哈希） 在交易中，收款方需要知道付款方的公钥，从而验证A签名是否有效。即A需要提供自己的公钥，如果所提供公钥与铸币交易中。（实际上其他节点都需要知道付款方公钥，验证交易合法性）实际中A转账时候提供的公钥需要和铸币交易中公钥对的上，这样就防止了恶意节点伪造A的公钥来"偷"走A的比特币。 在比特币系统中，通过执行脚本实现上述验证过程。将当前交易输入脚本与前一个交易输出脚本（说明币的来源的交易）拼接执行，如果可以正确执行，说明交易合法。 在该图中，一个区块仅含有一个交易，实际中一个区块中包含多个交易，交易通过Markle Tree（详见比特币数据结构篇中）组织起来，在区块中存储。

比特币区块信息

block Header（区块宏观信息）	block body(略)
Version(版本协议)	...
Hash of previous block header（指向前一个区块指针）	...
Merkle root hash（默克尔树根哈希值）	...
target（挖矿难度目标阈值）	...
nonce（随机数）	...

> 1. 挖矿求解问题：Hash（block header）<=target
> 2. Hash of previous block header只计算区块块头部分的哈希（ Merkle root hash保证了block body内容不被篡改，所以只需要计算block header即可保证整个区块内容不会被篡改）
> 3. 区块链系统中，轻节点（只存储区块block header信息）只利用区块链，但并不参与区块链系统维护和构造。

分布式共识

> 可否各个节点独立完成区块链构建？ 很明显不行，各个节点独立打包交易，形成区块链，必然无法避免区块链内容不一致。从分布式系统角度来说，**账本内容需要取得分布式共识**，从而保证区块链内容在不同节点上的一致性。

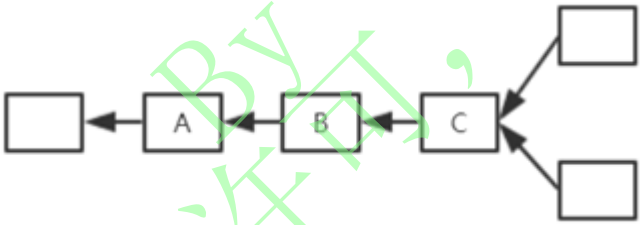
根据FLP不可能结论，在一个异步系统中，网络时延无上限，即使只有一个成员是有问题的，也不可能达成共识。 根据CAP Theorem（Consistency一致性、Availability可靠性、Partition tolerance容错性），任何一个分布式系统中，最多只能满足其中两个性质。 分布式共识中协议Paxos 可以保证Consistency（若达成共识必然一致），但在某些情况下，可能会一直无法达成共识。【在这里附上一个Paxos协议详解：https://my.oschina.net/u/150175/blog/2992187】

比特币共识协议

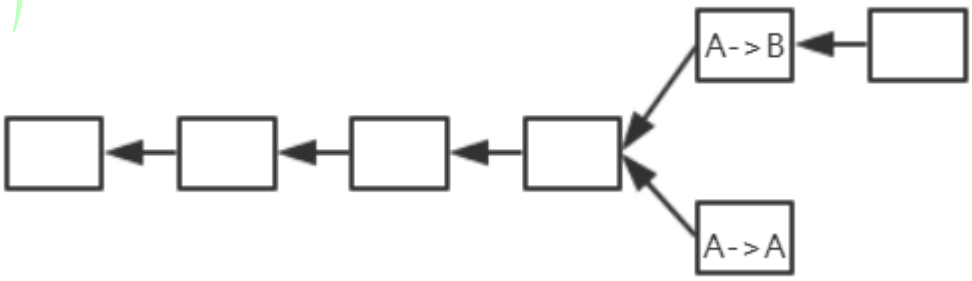
> 背景：假设系统中存在部分节点有恶意，但存在比例较小。大多数节点为"好"的节点，在这种情况下进行共识协议设置。 想法1：直接投票 某个节点打包交易到区块，将其发给其他节点，其他节点检查该候选区块，检查若正确投赞成票，若票数过半数，加入区块链。 存在的问题1——恶意节点不断打包不合法区块，导致一直无法达成共识，时间全花费在投票上。 存在的问题2——无强迫投票手段，某些节点不投票（行政不作为）。 存在的问题3——网络延迟事先未知，投票需要等多久？效率上会产生问题。 更大的一个问题——membership。如果是联盟链，对加入成员有要求，可以基于投票。但比特币系统，任何人都可以加入，且创建账户及其简单，只需要本地产生公私钥对即可。只有转账（交易）时候,比特币系统才能知道该账户的存在。这样，黑客可以使用计算机专门生成大量公私钥对，当其产生大量公私钥对超过系统中一半数目，就可以获得支配地位（**女巫攻击**）。所以，这种简单的投票方案也是不可行的。

比特币系统中采用了很巧妙的方案解决这个问题。虽然仍然是投票，但并非简单的根据账户数目，而是依据计算力进行投票。 在比特币系统中，每个节点都可以自行组装一个候选区块，而后，尝试各种nonce值，这就是挖矿。[H(block header)<=target] 当某个节点找到符合要求的nonce，便获得了**记账权**，从而可以将区块发布到系统中。其他节点受到区块后，验证区块合法性，如果系统中绝大多数节点验证通过，则接收该区块为最新的区块并加入到区块链中。

> 1. 会不会合法区块被拒绝？ 如图所示。发生分叉的情况下，暂时保存分叉情况，但区块链只承认最长合法链，随着时间推移，必然存在某一条链变成最长合法链。这样，也就会导致合法区块被拒绝



> 2. 分叉攻击 如图所示，A用户对上面的A转账给B的记录回滚，从而非法获取利益。在两条链上，发现交易都合法。这是一个典型的双花攻击。A给B转账后，用分叉攻击将钱又转回来，覆盖掉原来的记录。 在比特币系统中，这种情况实际上很难发生。因为大多数矿工认可的是最长的合法链，会沿着上面的链继续挖下去。而A这个攻击者要想回退记录，就必须使得下面的链变得比上面的链还长。理论上来说，攻击者需要达到整个系统中51%的计算力，才能使得这种攻击成功。



> 此外，区块链正常运行场景下，也可能会发生分叉。当两个节点同时获得记账权时，会有两个等长的合法链。在缺省情况下，节点接收最先听到的区块，该节点会沿着该区块继续延续。但随着时间延续，必然有一个链胜出，由此保证了区块链的一致性。（被扔掉的区块称为"孤儿区块"）

可见，依赖于算力竞争，有效的防止了“女巫攻击”。

比特币激励机制

> 为什么系统中节点要竞争记账权？需要提供算力和电力成本，节点为什么要去做？

比特币系统设计之初便考虑到了这个问题，那就是引入激励机制。比特币通过设置**出块奖励**来解决该问题，一个获得合法区块的节点，可以在区块中加入一个特殊交易（铸币交易）。事实上，这种方式也是唯一一个产生新比特币的途径。

> 比特币系统设计规定，起初每个区块可以获得50个比特币，但之后每隔21万个区块，奖励减半。

但是这样就可以了吗？？？ 区块中保存交易记录，那么，会不会存在节点只想发布区块而不想打包交易？中本聪在设计该系统时，引入了交易费。在一个区块中，其输入≥输出，差值便是给区块所属节点的手续费。这些会在后续文章中详细说明。