比特币区块结构Merkle树及简单支付验证分析

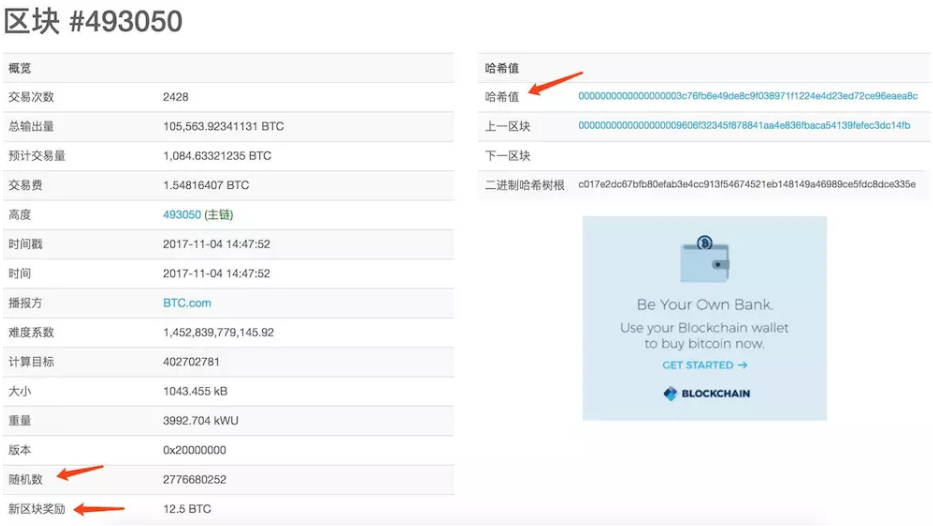
汇新云区块链学习社区 (/u/ff045813a1c3) + 关注
2018.01.19 10:47* 字数 1597 阅读 466 评论 0 喜欢 0
(/u/ff045813a1c3)

【汇新云 (https://link.jianshu.com?t=http%3A%2F%2Fwww.huixinyun.com%2F%3Fr%3D201722j00lj1222)】为大家定期更新文章，【汇新云 (https://link.jianshu.com?t=http%3A%2F%2Fwww.huixinyun.com%2F%3Fr%3D201722j00lj1222)】IT人的产业链平台

在比特币网络中，不是每个节点都有能力储存完整的区块链数据，受限于存储空间的的限制，很多节点是以SPV（Simplified Payment Verification简单支付验证）钱包接入比特币网络，通过简单支付验证可以在不必存储完整区块链下对交易进行验证，本文将分析区块结构Merkle树及如何进行交易验证。

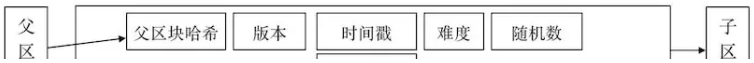
区块结构

在工作量证明 (https://www.jianshu.com/p/32061a9e7d53)中出现过一个区块信息截图：



细心的同学一定已经在里面发现了很多未讲的其他信息，如：时间戳，版本号，交易次数，二进制哈希树根(Merkle根)等。

我们来看看一个区块结构到底是怎样的：



如上图（下文称：区块结构图）所示：每个数据区块包含区块头和区块体。

区块头封装了当前版本号、前一区块哈希值、当前区块PoW要求的随机数(Nonce)、时间戳、以及Merkle根信息。

区块体则包括当前区块经过验证的、 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.

区块哈希值实际上并不包含在区块的数据结构里，其实区块打包时只有区块头被用于计算哈希（从网络被接收时由每个节点计算出来），常说的区块哈希值实际是区块头哈希值，它可以用来唯一、明确地标识一个区块。

区块头是80字节，而平均每个交易至少是250字节，而且平均每个区块包含2000个交易。因此，包含完整交易的区块比区块头的4千倍还要大。

SPV节点只下载区块头，不下载包含在每个区块中的交易信息。这样的不含交易信息的区块链，大小只有完整区块链的几千分之1，那SPV节点是如何验证交易的呢?

哈希验证

上面先留一个引子，先来回顾下哈希函数，记账原理(https://www.jianshu.com/p/0939c72e0760)我们知道原始信息任何微小的变化都会哈希完全不同的哈希值。

简单文件验证

我们通常用哈希来检验下载的文件是否完整，我经常看到这样的下载页面:

Files

Version	Operating System	Description	MD5 Sum	File Size	GPG
Gzipped source tarball	Source release		e9180c69ed9a878a4a8a3ab221e32fa9	22673115	SIG
XZ compressed source tarball	Source release		b9c2c36c33fb89bda1fefd37ad5af9be	16974296	SIG
Mac OS X 64-bit/32-bit installer	Mac OS X	for Mac OS X 10.6 and later	ce31f17c952c657244a5cd0cccae34ad	27696231	SIG

可以看到下载链接后面提供了一个MD5（MD5也是一种Hash算法），这样我们可以在下载之后对文件计算MD5，如果MD5与提供的MD5相等，说明文件有没有被损坏，这个验证过程相信大家都能理解。

多点文件验证(哈希列表)

现在复杂度提高一点，在P2P网络中下载时，会把大文件切成小文件，同时从多个机器上下载数据，这个时候怎么验证数据呢?

以BT下载为例，在下载真正的数据之前，我们会先下载一个哈希列表的（每个下小块计算出一个哈希），如果有一个小块数据在传输过程中损坏了，那我只要重新下载这一个数据块就行了，这时有一个问题就出现了，那么多的哈希，怎么保证它们本身(哈希列表中的哈希值)都是正确地呢?

答案是把每个小块数据的哈希值拼到一起，然后对这个长字符串在作一次哈希运算，得到哈希列表的根哈希。只要根哈希校对比一样就说明验哈希列表是正确的，再通过哈希列表校验小数据块，如果所有的小数据块验证通过则说明大文件没有被损坏。

Merkle树

验证交易的过程和文件验证很相似，可以人为每个交易是一个小数据块，但比特币使用Merkle树的方式进行验证，相对于哈希列表，Merkle树是一种哈希二叉树，它的明显的一个好处是可以单独拿出一个分支来（作为一个小树）对部分数据进行校验，更加高效。

我们回看下上面的区块结构图，区块体就包含这样一个Merkle树,Merkle树被用来归纳一个区块中的所有交易。

每个叶子节点是每个交易信息的哈希，往上对相邻的两个哈希合并成字符串再哈希，继续类似的操作直到只剩下顶部的一个节点，即Merkle根，存入区块头。

因为Merkle树是二叉树，所以它需要偶数个叶子节点。如果仅有奇数个交易需要归纳，那最后的交易就会被复制一份以构成偶数个叶子节点，这种偶数个叶子节点的树也被称为平衡树。

简化支付验证

SPV节点不保存所有交易也不会下载整个区块，仅仅保存区块头，我们来看看它是如何对交易数据进行验证的。

假如要验证区块结构图中交易6，SPV节点会通过向相邻节点索要（通过Merkleblock消息）包括从交易6哈希值沿Merkle树上溯至区块头根哈希处的哈希序列 (即哈希节点6, 5, 56, 78, 5678, 1234 1~8 - 称为认证路径) 来确认交易的存在性和正确性。（在N个交易组成的区块中确认任一交易只需要计算log2(N)个字节的哈希值，非常快速高效）

大家明白了吗?

汇新云区块链学习社区 (/u/ff045813a1c3)
写了 108286 字，被 180 人关注，获得了 69 个喜欢
(/u/ff045813a1c3)

+关注

行业十大区块链产品经理都在这里：http://www.huixinyun.com?r=201722j00lj1222 来这里一展拳脚： http://…