**区块链共识算法综述**

**1 引言**

区块链首次出现在中本聪发表的《比特币：一种点对点式的电子现金系统》论文中[1]。它的可实现性已由运行至今的比特币所证明。其突出优势在于去中心化设计的分布式数据库，通过运用时间戳、Merkle 树形结构、不对称密钥加密算法、共识算法和奖励机制，使用对等式网络——P2P 网络实现去中心化信用的交易，为解决中心化模式存在的可靠性差、低效率等问题提出新的计算范式。

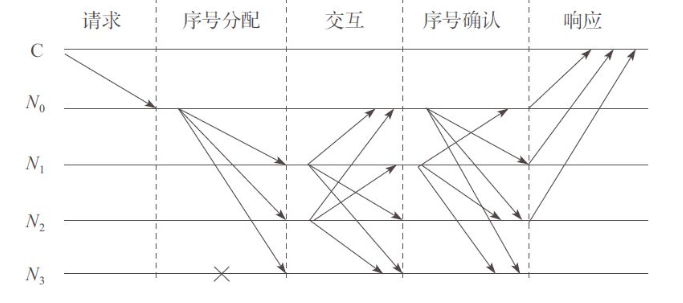
共识问题一直是分布式系统的重要研究课题，在分布式系统中，多个主机通过异步通信方式组成网络集群。在这样的一个异步系统中，需要主机之间进行**状态复制**，以保证每个主机达成一致的状态共识。然而，异步系统中可能出现无法通信的故障主机，主机的性能可能下降，网络可能拥塞，这些会导致错误信息在系统内传播。因此需要**在默认不可靠的异步网络中定义容错协议，以确保各主机达成安全可靠的状态共识。本文对5种代表性的区块链共识算法进行了综述。**

**2 区块链共识算法**

**2.1 PBFT（Practical Byzantine Fault Tolerance）**

实用拜占庭容错系统（PBFT）[2]是一种状态机副本复制算法，即服务作为状态机进行建模，状态机在分布式系统的不同节点进行副本复制。每个状态机的副本都保存了服务的状态，同时也实现了服务的操作。将所有的副本组成的集合使用大写字母R表示，使用0到|R|-1的整数表示每一个副本。为了描述方便，通常假设故障节点数为m个，整个服务节点数为|R|=3m+1个，这里m是有可能失效的副本的最大个数。尽管可以存在多于3m+1个副本，但是额外的副本除了降低性能之外不能提高可靠性。

PBFT要求共同维护一个状态，所有节点采取的行动一致。为此，需要运行三类基本协议，包括一致性协议、检查点协议和视图更换协议。一致性协议至少包含若干个阶段：请求（request）、序号分配（pre-prepare）和响应（reply）。根据协议设计的不同，可能包含相互交互（prepare），序号确认（commit）等阶段。



上图为PBFT协议通信模式，其中C为客户端，N0-N3表示服务节点，特别的，N0为主节点，N3为故障节点。每一个客户端的请求需要经过5个阶段，通过采用两次两两交互的方式在服务器达成一致之后再执行客户端的请求。由于客户端不能从服务器端获得任何服务器运行状态的信息，PBFT中主节点是否发生错误只能由服务器监测。如果服务器在一段时间内都不能完成客户端的请求，则会触发视图更换协议。整个协议的基本过程如下：

1）客户端发送请求，激活主节点的服务操作。

2）当主节点接收请求后，启动三阶段的协议以向各从节点广播请求。

① 序号分配阶段:主节点给请求赋值一个序列号n，广播序号分配消息和客户端的请求消息m，并将构造PRE-PREPARE消息给各从节点；

② 交互阶段:从节点接收PRE-PREPARE消息，向其他服务节点广播PREPARE消息；

③ 序号确认阶段:各节点对视图内的请求和次序进行验证后，广播COMMIT消息，执行收到的客户端的请求并给客户端以响应。

3）客户端等待来自不同节点的响应，若有m+1个响应相同，则该响应即为运算的结果。

PBFT在很多场景都有应用，在区块链场景中，一般适合于对强一致性有要求的私有链和联盟链场景。例如，在IBM主导的区块链超级账本项目中，PBFT是一个可选的共识协议。在Hyperledger的Fabric项目中，共识模块被设计成可插拔的模块，支持像PBFT、Raft等共识算法。

**2.2 Raft**

在分布式系统的实用场景下，其假设条件不需要考虑拜占庭故障，而只是处理一般的死机故障。在这种情况下，采用Paxos等协议会更加高效。Paxos是Lamport设计的保持分布式系统一致性的协议。由于Paxos非常复杂，比较难以理解，因此后来出现了各种不同的实现和变种。Raft是由Stanford提出的一种更易理解的一致性算法，意在取代目前广为使用的Paxos算法。目前，在各种主流语言中都有了一些开源实现，比如本文中将使用的基于JGroups的Raft协议实现[3]。

Raft最初是一个用于管理复制日志的共识算法，它是一个为真实世界应用建立的协议，主要注重协议的落地性和可理解性。Raft是在非拜占庭故障下达成共识的强一致协议。

在区块链系统中，使用Raft实现记账共识的过程可以描述如下：首先选举一个leader，接着赋予leader完全的权力管理记账。leader从客户端接收记账请求，完成记账操作，生成区块，并复制到其他记账节点。leader简化了记账操作的管理。例如，leader能够决定是否接受新的交易记录项而无需考虑其他的记账节点，leader可能失效或与其他节点失去联系，这时，系统就会选出新的leader。

在Raft中，每个结点会处于下面三种状态中的一种：

* **follower**：所有结点都以follower的状态开始。如果没收到leader消息则会变成candidate状态
* **candidate**：会向其他结点“拉选票”，如果得到大部分的票则成为leader。这个过程就叫做Leader选举(Leader Election)
* **leader**：所有对系统的修改都会先经过leader。每个修改都会写一条日志(log entry)。leader收到修改请求后的过程如下，这个过程叫做日志复制(Log Replication)：
* 复制日志到所有follower结点(replicate entry)
* 大部分结点响应时才提交日志
* 通知所有follower结点日志已提交
* 所有follower也提交日志
* 现在整个系统处于一致的状态

Raft阶段主要分为两个，首先是leader选举过程，然后在选举出来的leader基础上进行正常操作，比如日志复制、记账等。

**2.3 PoW（Proof of Work）**

PoW是以每个节点的计算能力即“算力”来竞争记账权的机制。在比特币系统中，大约每10分钟进行一轮算力竞赛，竞赛的胜利者，就获得一次记账的权力，并向其他节点同步新增账本信息。比特币系统是通过一个称为“工作量证明”（Proof of Work，PoW）的机制来判定竞争的结果的[1]。

PoW系统的主要特征是计算的不对称性。工作端需要做一定难度的工作得出一个结果，验证方却很容易通过结果来检查工作端是不是做了相应的工作。

​在PoW中，对于特定字符串后接随机nonce值所构成的串，要找到这样的nonce值，满足前n位均为0的SHA256值，需要多次进行哈希值的计算。一般来说，n值越大，需要完成的哈希计算量也越大。由于哈希值的伪随机特性，要寻找4个前导0的哈希值，预期大概要进行216次尝试，这个数学期望的计算次数，就是所要求的“工作量”。

关于比特币PoW共识机制能否解决拜占庭将军问题一直在业界有争议。2015年，Juan Garay对比特币的PoW共识算法进行了正式的分析，得出的结论是比特币的PoW共识算法是一种概率性的拜占庭协议（Probabilistic BA）。Garay对比特币共识协议的两个重要属性分析如下。

1）一致性（Agreement）

在不诚实节点总算力小于50%的情况下，同时每轮同步区块生成的几率很少的情况下，诚实的节点具有相同的区块的概率很高。用数学的严格语言说应该是：当任意两个诚实节点的本地链条截取K个节点，两条剩下的链条的头区块不相同的概率随着K的增加呈指数型递减。

2）正确性（Validity）

大多数的区块必须由诚实节点提供。严格来说，当不诚实算力非常小的时候，才能使大多数区块由诚实节点提供。

因此可以看到，当不诚实的算力小于网络总算力的50%时，同时挖矿难度比较高，在大约10分钟出一个区块情况下，比特币网络达到一致性的概念会随确认区块的数目增多而呈指数型增加。但当不诚实算力具一定规模，甚至不用接近50%的时候，比特币的共识算法并不能保证正确性，也就是，不能保证大多数的区块由诚实节点来提供。

PoW机制存在明显的弊端。一方面，PoW的前提是，节点和算力是均匀分布的，因为通过CPU的计算能力来进行投票，拥有钱包（节点）数和算力值应该是大致匹配的，然而随着人们将CPU挖矿逐渐升级到GPU、FPGA，直至ASIC矿机挖矿，节点数和算力值也渐渐失配。另一方面，比特币网络每秒可完成数百万亿次SHA256计算，但这些计算除了使恶意攻击者不能轻易地伪装成几百万个节点和打垮比特币网络，并没有更多实际或科学价值。当然，相对于允许世界上任何一个人在瞬间就能通过去中心化和半匿名的全球货币网络，给其他人几乎没有手续费地转账所带来的巨大好处，它的浪费也许只算是很小的代价。

**2.4 PoS（Proof of Stake）**

鉴于上述PoW的弊端，人们提出了权益证明（Proof of Stake，PoS）。

PoS[4]模式会根据持有数字货币的量和时间，分配相应的利息。点点币（Peercoin）是首先采用权益证明的货币，点点币在SHA256的哈希运算的难度方面引入了币龄的概念，使得难度与交易输入的币龄成反比。在点点币中，币龄被定义为币的数量与币所拥有的天数的乘积，这使得币龄能够反映交易时刻用户所拥有的货币数量。实际上，点点币的权益证明机制结合了随机化与币龄的概念，未使用至少30天的币可以参与竞争下一区块，越久和越大的币集有更大的可能去签名下一区块。

然而，一旦币的权益被用于签名一个区块，则币龄将清为零，这样必须等待至少30日才能签署另一区块。同时，为防止非常老或非常大的权益控制区块链，寻找下一区块的最大概率在90天后达到最大值，这一过程保护了网络，并随着时间逐渐生成新的币而无需消耗大量的计算能力。点点币的开发者声称这将使得恶意攻击变得困难，因为没有中心化的挖矿池需求，而且购买半数以上的币的开销似乎超过获得51%的工作量证明的哈希计算能力。

权益证明必须采用某种方法定义任意区块链中的下一合法区块，依据账户结余来选择将导致中心化，例如单个首富成员可能会拥有长久的优势。为此，人们还设计了其他不同的方法来选择下一合法区块。

PoS机制虽然考虑到了PoW的不足，但依据权益结余来选择，会导致首富账户的权力更大，有可能支配记账权。股份授权证明机制（Delegated Proof of Stake，DPoS）的出现正是基于解决PoW机制和PoS机制的这类不足。

**2.5 DPoS(Delegated Proof of Stake)**

授权股份证明（Delegated Proof of Stake，DPoS）[5] 共识算法尝试解决PoW和PoS存在的问题，通过实施去中心化的民主方式，每个币相当于一张选票，持有币的人可以根据自己持有币的数量来将自己的若干选票投给自己信任的受托人。系统会选出获得投票数量最多的N个人作为系统受托人，他们的工作是签署（生产）区块，且在每个区块被签署之前，必须先验证前一个区块已经被受信任节点签署。

比特股（Bit Shares）是采用DPoS算法的区块链，它期望通过引入一个技术民主层来减少中心化的负面影响。比特股引入了见证人（witness）这个概念，见 证人可以生成区块，每一个持有比特股的人都可以投票选举见证人。得到总同意票数中的前N个候选者可以当选为见证人，当选见证人的个数N需满足至少一半的参与投票者相信N已经充分地去中心化。见证人的候选名单每个维护周期更新一次，见证人随机排列，每个见证人按序有既定的权限时间生成区块，若见证人在给定的时间片不能生成区块，区块生成权限交给下一个时间片对应的见证人。

**3 总结**

共识机制解决了区块链如何在分布式场景下达成一致性

的问题。但目前的共识算法还都不完善，有很多需要改进的空间。根据当前研究进展，未来区块链共识算法的研究方向主要集中在以下3个方面。

1. 提高联盟链交易处理性能。 设计具有高吞吐量、低时延性能的共识算法是区块链技术发展的重点，而联盟链作为目前备受企业等组织机构青睐的应用场景，在溯源、供应链等领域已落地，其交易处理能力亟需提高。
2. 共识算法的性能评估标准与测试方法。 区块链共识算法评估标准可以从一致性、安全 性、扩展性、资源消耗等维度综合考量[79]。在测试方法上，可借助网络中间件和交易回放工具对拜占庭节点、复杂网络、可扩展性、压力等方面进行测试，其中压力测试的测试场景可包括常规交易、远程处 理性能、复杂交易、不同网络时延等。另外，可以降级共识算法的复杂度，进一步完善和使其透明简单化。
3. 基于隐私保护的共识算法。 区块链的数据隐私和访问控制可以借助设计基于隐私保护的共识算法来改善，例如，联盟链中可以借助基于节点权限管理的共识算法来保障区块链数据和网络安全。

**参考文献**

1. NAKAMOTO S. Bitcoin: A Peer-to-Peer Electronic Cash System [EB/OL]. https://bitcoin.org/bitcoin.pdf, 2019-1-7.
2. M. Castro and B. Liskov. Practical Byzantine Fault Tolerance. Proceedings of the Third Symposium on Operating Systems Design and Implementation, New Orleans, USA, February 1999
3. Ongaro D, Ousterhout J K. In search of an understandable consensus algorithm[C]//USENIX Annual Technical Conference. 2014: 305-319.
4. SUNNY K, SCOTT N. PPcoin: Peer-to-Peer Crypto-Currency with Proof-of-Stake[EB/OL]. https://decred.org/research/king2012. pdf,2012-8-19.
5. LARIMER D. Delegated Proof of Stake [EB/OL]. https://how. bitshares.works/en/master/technology/dpos.html, 2018-8-29.