中国科学: 信息科学 2022 年 第 52 卷 第 5 期: 837–855
SCIENTIA SINICA Informationis
论文

中本聪共识安全性质研究方法

周子钰[1], 张宗洋[1, 2*], 刘建伟[1]

1. 北京航空航天大学网络空间安全学院, 北京 100091
2. 武汉大学空天信息安全与可信计算教育部重点实验室, 武汉 430072
* 通信作者. E-mail: zongyangzhang@buaa.edu.cn

收稿日期: 2021–07–07; 修回日期: 2021–09–04; 接受日期: 2021–09–14; 网络出版日期: 2022–05–17

国家自然科学基金 (批准号: 61972017, 61972018, 61932014, 72031001)、北京市自然基金项目 (批准号: M21033)、云南省区块链应用技术重点实验室 (培育) 开放课题 (批准号: YNB202101)、国家密码发展基金 (批准号: MMJJ20180215) 和中央高校基本科研业务费 (批准号: YWF-21-BJ-J-1040) 资助项目

摘要 中本聪共识是区块链共识机制中最基础和研究最广泛的一种共识机制, 其安全性对整个区块链领域的发展具有重要的理论意义和应用价值. 现有大量研究在各种模型假设下对中本聪共识进行了安全性分析和证明. 本文首先详细描述了中本聪共识的执行模型, 包括时间模型、网络模型、敌手模型等. 其次, 系统总结了中本聪共识安全性的形式化定义. 再次, 根据时间模型将中本聪共识研究方法分为基于离散时间模型和连续时间模型两类, 并指出不同方法的优缺点. 最后对中本聪共识的安全性研究进行展望, 旨在为区块链共识机制的研究提供技术支撑.

关键词 中本聪共识, 区块链, 安全性证明, 数字货币, 工作量证明

1 引言

去中心化数字货币首要解决的问题就是数字货币的 “双花”, 即交易双方在完成交易后, 一方通过设法获取记账权改写交易记录使支付失效. 中本聪 [1] 设计了比特币系统以解决去中心化数字货币双花问题并对其进行了安全性证明, 该系统所使用的底层协议被称为 “中本聪共识”. 中本聪共识作为比特币底层共识协议, 使用区块链构造分布式公共账本, 用以维系去中心化支付系统的运作, 使互不信任的参与者就交易账本达成共识. 其中全网参与者通过解决一个工作量证明 (PoW, proof-of-work) 问题获得 “记账权”, 将交易打包成区块并广播, 区块通过引用前一个区块的哈希 (Hash) 值链接成 “区块链”.

中本聪在分析共识安全性时, 将攻击者和诚实参与者看作是竞争双方, 诚实参与者维护了特定交易所在的诚实链, 攻击者希望改写该特定交易而创造了敌手链. 两种链的增长可视为用一定概率朝同一方向进行二项随机游走, 因此敌手链追上诚实链的概率可被刻画为一个赌徒破产问题, 进而该特定

引用格式: 周子钰, 张宗洋, 刘建伟. 中本聪共识安全性质研究方法. 中国科学: 信息科学, 2022, 52: 837–855, doi: 10.1360/SSI-2021-0225
Zhou Z Y, Zhang Z Y, Liu J W. Methods of security analysis for Nakamoto consensus (in Chinese). Sci Sin Inform, 2022, 52: 837–855, doi: 10.1360/SSI-2021-0225

交易被改写的概率随交易在区块链中的深度呈指数级下降. 然而, 中本聪并没有形式化定义区块链及交易账本所需的安全性质, 而且其模型假设比较理想, 无法完美刻画现实网络环境. 中本聪共识作为如今许多区块链的底层共识的基础, 如何更好地建模现实执行环境并证明其安全性, 对区块链技术的发展至关重要.

1.1 中本聪共识安全性研究概述

中本聪 [1] 仅在一个理想的模型下证明中本聪共识能够抵抗双花攻击 (double-spend attack). 在现实环境中, 一方面, 诚实参与者相互之间并不知道对方的情况, 网络通信也存在时间延迟, 另一方面, 敌手可以发动共谋攻击, 通过控制消息传输达到分裂诚实参与者的算力和视图的目的. 因此, 当研究中本聪共识安全性时, 需要合理地建模实际网络环境和敌手攻击策略.

Garay 等 [2] 在同步网络下对中本聪共识进行了安全性研究, 他们剔除共识实现细节, 将中本聪共识抽象成比特币骨干协议, 并提出协议需满足两大基本安全性质 —— 共同前缀性质和链质量性质; 他们还提出使用区块链构建的公共账本协议需满足持续性和活性, 这些都成为后续区块链共识安全性研究的基础. Kiayias 和 Panagiotakos [3] 在共同前缀和链质量性质之上增加了链增长性质, 它们共同构成区块链共识的 3 个基本安全性质. Pass 等 [4] 考虑了通信网络中存在的时间延迟, 并在半同步网络下研究了中本聪共识的安全性, 他们在比特币骨干协议 [2] 的研究基础上, 将一致性 (共同前缀性质) 扩展为普通的一致性 (consistency) 和未来自我一致性 (future self-consistency). Kiffer 等 [5] 继续在半同步网络下开展研究, 使用马尔可夫链 (Markov chain) 描述中本聪共识的执行过程和敌手攻击过程, 简化了比特币骨干协议一致性的证明, 同时给出了一致性满足时更精确的敌手能力约束条件.

上述对于中本聪共识的研究均默认协议参与者数量不变且比特币工作量证明计算的难度调节机制完美, 因此比特币挖矿成功概率不变, 同时出块时间间隔保持稳定. 但当实际执行协议时, 由于比特币的非授权设定 (permissionless setting), 协议参与者数量一直处于动态波动状态, 而相应的工作量证明难度也需随之调整. Garay 等 [6] 在非授权设定和同步网络下, 研究比特币的难度调整机制能否保证中本聪共识的安全性. 他们证明当参与者数量在一定范围内波动时, 中本聪共识仍能满足安全性, 因此中本聪共识可在非授权设定下构建鲁棒的公共账本. 此后, Garay 等 [7] 沿用相似的方法分析了中本聪共识安全性, 证明在时延受限网络中, 即使参与者动态变化, 中本聪共识仍然能满足一致性和活性.

为了便于描述全网参与节点的同步和消息传递, 许多研究使用离散时间模型, 将时间划分为 “轮 (round)” 作为共识执行单位. Ren [8] 为了更好地建模实际共识执行环境, 使用连续时间模型研究半同步网络下中本聪共识的安全性. Li 和 Guo [9] 在 Ren 的研究基础上进一步在连续时间模型下精确计算了中本聪共识一致性和活性得到满足的概率. 针对中本聪共识一致性的最新研究 [10, 11] 使用了两种不同的方法, 在连续时间模型中计算时延受限网络中共识满足安全性所能承受的最大敌手算力.

在一致性和活性之外, Avarikioti 等 [12] 指出大多数针对比特币系统的攻击, 都是在敌手突然拥有超过全网 50% 算力的情况下进行的, 因此需研究出现敌手算力短期突增时, 中本聪共识能否保证安全. Badertscher 等 [13] 形式化定义了新的分布式账本安全性质 —— 自愈性质 (self-healing), 该性质描述了当敌手算力突增时, 分布式账本协议恢复正常状态所需要的时间, 若该时长控制在一定范围之内, 那么该账本协议满足自愈性质.

对于中本聪共识协议安全性的研究, 基本上都遵照 Lynch 在分布式算法 [14] 所提出的步骤进行分析和证明. 首先, 明确共识协议所须满足的安全性质, 抽象并形式化定义具体安全问题; 其次, 明确协议执行环境并提出解决上述问题的分布式系统数学模型; 最后, 在上述模型假设中精确描述分布式共识协议的安全性质, 分析算法复杂度并证明约束条件或不可能性.

1.2 本文贡献

本文沿用上述的安全性证明流程框架, 研究中本聪共识协议, 其中部分研究结果不仅能应用于中本聪共识上, 还能应用于其他区块链共识协议, 包括基于工作量证明、基于权益证明、使用最长链原则等共识协议. 主要研究内容如下所示.

(1) 归纳了中本聪共识的安全性证明基本流程, 将其分为安全性质形式化定义、敌手能力建模、执行环境建模和安全性证明等.

(2) 总结了区块链共识安全性证明的模型, 包括时间模型、网络模型、参与者模型等.

(3) 分类并对比了现有中本聪共识安全性研究方法的基本原理和优缺点.

(4) 总结并展望了中本聪共识协议的安全性研究方向.

本文将中本聪共识的研究模型, 根据是否考虑消息同步传输而以 "轮" 划分共识执行过程, 分为离散时间模型和连续时间模型. 根据如何考虑消息在网络中的传输时延, 将其网络模型分为同步模型、时延受限模型和异步模型. 表 1 总结并对比了各种中本聪共识安全性研究方法, 符号 "–" 代表无需考虑该特性.

1.3 相关工作

共识协议作为区块链应用的基础, 其安全性直接影响上层应用安全性. 中本聪共识作为目前多数区块链共识的基础, 其安全性一向是研究热点. 中本聪 [1] 在白皮书中用赌徒破产问题描述并研究共识在诚实主导设定下的安全性. 后续大量研究 [2~7] 形式化定义并证明了中本聪共识的基本安全性质, 考虑通信时延或非授权设定对共识安全性的影响, 在不同模型下研究中本聪共识安全性. 为了更贴合实际协议执行环境, 部分研究 [8,9] 分析连续时间模型下的共识执行安全性. 然而上述研究均假定敌手资源不超过全网算力的一半, 文献 [12,13] 研究了当敌手在短时间获得超过全网一半算力时, 中本聪共识的安全性与恢复能力. 同时 Gazi 等 [10] 和 Dembo 等 [11] 证明了中本聪共识可容忍算力更高的敌手.

目前有部分研究归纳和总结了中本聪共识或区块链的安全性. Stifter 等 [15] 主要调研了对中本聪共识的形式化定义并对比了不同的分布式计算协议. Garay 和 Kiayias [16] 主要研究了区块链共识的分类方法, 并给出了现有的共识研究模型. Yuan 等 [17] 系统地归纳了各区块链共识协议, 并给出共识算法的演进历程和分类方法. Yang 和 Zhang [18] 介绍并对比分析了代表性区块链共识协议的应用场景和性能. Liu 等 [19] 则进一步将现有区块链共识根据不同模型和应用场景进行分类, 指出各类共识流程的优缺点和潜在攻击. 但这些研究都侧重于安全性质形式化定义和区块链共识本身的研究, 缺乏对其具体研究模型和方法的总结. 而事实上, 为了研究现实环境下共识协议的安全性, 需要准确地描述共识执行的时间延迟和敌手可能进行的各种攻击, 还需要合理地建模共识执行环境并在该模型下证明共识安全性.

2 定义与模型

本节主要介绍本文所使用的定义与模型, 2.1 小节介绍了中本聪共识的执行模型, 包括协议执行模型, 网络模型和时间模型; 2.2 小节介绍了中本聪共识中的参与节点模型; 2.3 小节介绍了工作量证明中使用的哈希模型. 本文使用的参数如表 2 所示.

定义1 (矿工) 矿工 (miner) 是参与共识协议的节点, 共同维系区块链账本. 共识协议执行时收集全网交易, 使用工作量证明计算将交易打包成区块链接到区块链上, 通过创造区块与打包交易获取佣金.

表 1　中本聪共识安全性研究方法

Table 1　Methods of security analysis for Nakamoto consensus

Type	Network	Paper	Model	Unit	Contributions	Limitations
Continuous-time model	Synchronous	[1]	Binomial random walk	–	Uses a binomial random walk to depict the race between the honest chain and the attacker chain which is analogous to a Gambler's ruin problem.	Considers only the private attack and lacks discussions about the communication and execution model.
	Bounded-delay	[8]	Poisson distribution	–	Simplifies the analysis of the Nakamoto consensus in a continuous-time model.	Lacks an execution model and property formalizations that are intuitive enough.
		[11]	Poisson distribution	–	Models all attacks to a race between the adversary and the honest nodes; finds the worst attack and its true security threshold for the longest chain protocols.	Limited to analyzing blockchain protocols that use the longest chain policy.
Discrete-time model	Synchronous	[2]	Binomial distribution	Round	Formalizes the common-prefix property and chain-quality property; formalizes the ledger properties of persistence and liveness.	Lacks consideration of network delays and the permissionless setting.
		[3]	Binomial distribution	Round	Analyzes the trade-off between provable security and transaction processing speed; formalizes the chain growth property.	Lacks consideration of network delays and the permissionless setting.
		[6]	Binomial distribution	Round	Introduces typical executions to analyze the security of Bitcoin PoW difficulty adjustment mechanism.	Lacks consideration of network delays.
	Bounded-delay	[4]	Binomial distribution	Round	Analyzes the security of Nakamoto consensus using the F_{tree} model; introduces the future self-consistency proper.	Does not actually consider the permissionless setting in the security proof.
		[5]	Binomial distribution	Round	Provides a tighter guarantee on Nakamoto consensus's consistency property using Markov-chain.	Lacks consideration of the permissionless setting.
		[7]	Binomial distribution	Round	Analyzes Nakamoto consensus in the setting of bounded communication delays and dynamic participation.	The honest majority assumption could be tighter.
		[12]	Binomial distribution	Round	Proves Bitcoin is secure under the temporary dishonest majority.	Lacks consideration of the adaptive adversary; limited to analyzing PoW protocols.
		[13]	Binomial distribution	Round	Formalizes self-healing properties; quantifies the recover time under the dishonest majority attack.	Adopts a discrete view of protocol time that requires a global clock.
		[10]	Discrete Poisson distribution	Slot	Proves the optimal security threshold for the Bitcoin protocol using characteristic strings.	The discrete approximation requires discussions of several regions.

表 2　参数

Table 2　Parameters

Parameter	Definition
P_i	The ith party that is participating in the protocol
Δ	The upper bound of the network delay
κ	The security parameter
n	The number of parties participating in the protocol
t	The number of parties controlled by the adversary
q	The number of queries per round of a party
p	The probability of success of a single query
ρ	The fraction of the computational power held by the adversary
λ	The total mining rate of all parties

定义2 (工作量证明) 工作量证明是矿工通过暴力破解密码学难题来对其计算能力的证明. 在比特币协议中, 矿工通过暴力计算基于 SHA-256 的哈希不等式以生成区块链的新区块. 矿工进行工作量证明计算的过程称为 "挖矿".

定义3 (授权设定) 在授权设定 (permissioned setting) 中, 参与节点需要进行身份认证, 并在经过授权后才能加入协议.

定义4 (非授权设定) 在非授权设定 (permissionless setting) 中, 参与节点无需授权和身份认证便能自由加入或离开协议, 各节点无法获知其他节点的信息, 而且协议的执行无需考虑参与节点的身份.

定义5 (动态设定) 在动态设定 (dynamic setting) 中, 协议参与节点数量动态变化.

在比特币区块链中, 节点可自由加入或离开协议导致参与节点数变化, 因此是动态设定的.

定义6 (静态设定) 在静态设定 (static setting) 中, 协议参与节点数量不变.

为了简化安全性证明过程, 部分研究会在静态设定下证明中本聪共识的安全性.

定义7 (诚实主导) 在诚实主导 (honest majority) 的共识协议中, 诚实参与者的资源超过全网资源的 50%. 在基于工作量证明和基于权益证明的协议中, 诚实主导分别意味着诚实参与者拥有超过全网 50% 的算力或权益.

定义8 (敌手主导) 在敌手主导 (dishonest majority) 的共识协议中, 敌手的资源超过全网资源的 50%. 在基于工作量证明和基于权益证明的协议中, 敌手主导分别意味着敌手拥有超过全网 50% 的算力或权益.

2.1 执行模型

2.1.1 协议执行模型

中本聪共识协议的安全性分析通常基于通用可组合安全模型 [20], 其中共识协议由环境程序所驱动, 所有的参与节点与敌手均被建模为交互图灵机 (ITM, interactive Turing machines), 每个 ITM 拥有各自的通信带、输出带和输入带. 每个执行特定程序的 ITM 被称为 ITM 的一个实例 (instance) 缩写为 ITI, ITI 可视为区块链共识协议中运行特定程序的节点, 如执行挖矿程序并打包交易的区块链矿工或执行交易程序并发布交易信息的普通节点等. 新 ITI 的生成以及 ITI 之间的通信由控制程序所决定, 该控制程序也是 ITM.

2.1.2 网络模型

大多数区块链协议使用广播 (broadcast) 在参与节点间传输消息. 但是在现实网络环境中, 消息需经过多次转发才能到达所有接收者处, 因此消息传输存在时间延迟. 敌手能利用该时间延迟, 通过恶意控制消息的传输使不同的参与节点在同一时刻的网络视图产生差异, 从而引起区块链分叉. 因此, 区块链共识协议的研究需考虑消息传输时延对安全性的影响.

网络模型根据消息延迟被分为同步网络模型、半同步网络模型、时延受限网络模型和异步网络模型. 异步网络模型中的时间延迟不存在上限, 敌手可任意延迟消息的传输, 从而有足够长的时间延长自己的链使全网接受, 该模型下共识协议的一致性和活性无法同时满足. 同步网络模型属于理想的网络模型, 其中消息以轮为单位传播, 每轮协议执行完成后, 消息一定会被传递到接收者处, 该网络模型一般应用于存在可信第三方的中心化共识协议中. 半同步网络模型是分布式协议分析中常用的网络模型, 由于现实网络环境中, 敌手难以完全控制一定数量参与节点的全部网络通信, 只能尽可能延迟消

息, 因此消息最终会被送至接收者处, 从而存在消息传输延迟上限 Δ, 但 Δ 具体数值并不为参与节点所知.

为了能对区块链共识安全性进行更准确的数值分析, 大多数研究使用存在确定延迟上限的时延受限网络模型, 其中消息的时间延迟上限 Δ 已知, 敌手能在 $0 \sim \Delta$ 之间任意延迟不同节点的消息, 其中 Δ 一般根据实际网络带宽估计. 由于时延受限网络模型的消息传输延迟上限 Δ 和不同处理器之间的相对速度上限确定, 且在计算时作为已知参数, 因此时延受限网络模型有时被认为属于同步网络模型.

定义9 (同步网络) 在同步网络 (synchronous network) 中, 诚实节点之间的消息按照轮来传输, 每轮节点发出的消息一定会在该轮结束前到达相应接收者处, 即 $\Delta = 0$. 基于轮的同步网络模型又称为锁步同步 (lock-step synchrony) 模型, 一般用于存在同步时钟或中央管理者的区块链共识协议的安全性研究, 但同步网络对于中本聪共识等去中心化协议的研究则过于理想.

定义10 (半同步网络) 半同步网络 (semi-synchronous network) 也被称为部分同步网络 (partially synchronous network), 其中存在消息传输延迟上限 Δ 和不同处理器之间的相对速度上限. 诚实节点的消息在 Δ 时间后, 一定会到达相应接收节点处, 然而 Δ 并不预先确知.

定义11 (时延受限网络) 在时延受限网络 (bounded-delay network) 中, 诚实节点之间的消息传输存在确定延迟上限 Δ, 诚实节点的消息在 Δ 时间后, 一定会到达相应接收节点处. 与半同步网络的区别在于, 时延受限网络的 Δ 确定, 可以在安全分析中直接作为参数计算.

定义12 (异步网络) 在异步网络 (asynchronous network) 中, 敌手能够任意延迟传输到诚实节点的消息, 但消息最终能到达相应接收节点处.

2.1.3 时间模型

区块链共识协议的执行时间模型主要分为离散时间模型 (discrete-time model) 和连续时间 (continuous-time model) 模型. Garay 等 [2] 使用 Katz 等 [21] 定义的同步、延迟、轮执行, 以及多方计算等概念, 在基于轮的离散时间模型下研究共识安全性. 此后的大量工作 [3~7] 也基于该模型研究中本聪共识的不同性质. 基于轮的离散时间模型能很好地描述区块链网络中节点间的消息同步和传输时延, 并刻画参与者的计算能力, 同时利于共识协议与其他算法的安全组合. 然而, 以轮为单位的执行模型过于理想, 难以建模现实世界中的网络.

Ren [8] 认为需使用更简单的方法研究中本聪共识, 在连续时间模型下形式化定义中本聪共识安全性质, 并给出安全性质满足的条件. Li 和 Guo [9] 基于 Ren 的研究进一步精确计算了连续时间模型下中本聪共识满足一致性和活性的概率. Gazi 等 [10] 和 Dembo 等 [11] 分别在连续时间模型和基于时段 (slot) 的离散时间模型下证明时延受限网络中, 中本聪共识安全性能在更高的敌手能力约束条件下满足安全性, 上述两个研究 [10,11] 使用不同方法获得了相同结论.

离散时间模型. 基于轮的离散时间模型便于描述网络中的同步概念, 它将时间划分为 "轮", 并假设存在一个所有节点都能访问的世界时钟 F_{CLOCK}, 每轮节点完成一次消息传递, 该模型也适用于描述半同步网络和时延受限网络 [22].

为了更好地描述现实连续时间下的协议执行, 部分研究 [10,13,23] 使用基于时段的离散时间模型, 该模型将时间划分为长度相等的时段, 以时段作为区块链协议执行的时间单位. 时段是足够小的时间间隔, 当时段趋近于 0 时, 该模型能近似为连续时间模型. 当只需考虑在一定时间内中本聪共识参与者所生成的区块数而无需考虑底层通信细节时, 使用基于时段的离散时间模型可简化分析和安全性证明过程.

连续时间模型. 由于基于轮的离散时间模型需假设世界同步时钟的存在, 不符合实际区块链协议去中心化的特性, 因此大量研究在连续时间模型下分析区块链协议的安全性. 在连续时间模型下, 如何描述参与节点算力、敌手能力, 以及消息传输时延是共识研究面临的挑战, 一般使用泊松参数刻画单位时间内参与者的计算能力. 但该模型将参与者分为诚实参与者和敌手两个阵营, 忽视敌手可能使用特定策略分裂网络等方式影响诚实参与者有效挖矿成功率的情况, 难以研究适应性敌手的攻击.

2.2 参与节点模型

2.2.1 计算能力模型

在基于轮的离散时间模型的研究 [2~4] 中, 参与者被简化为 "扁平模型" (flat-model). 其中所有的诚实参与者均被视为独立的节点, 每个节点算力相同, 分别证明各自的工作量, 每轮均执行 q 次哈希计算, 而控制 t 个节点的敌手每轮能执行 $t \cdot q$ 次哈希计算. 现实中算力较大的参与者, 如矿池, 则被视为多个节点的集合. 在连续时间模型中, 参与者算力无法简单地用执行多少次工作量证明来描述, 大多数研究用泊松参数定义单位时间内参与者找到有效 PoW 解的个数.

2.2.2 敌手能力模型

共识安全性研究往往考虑能力最强的敌手. 基于轮的离散时间模型的研究应用通用可组合安全执行模型 [20], 其中敌手每轮决定腐化策略且腐化节点总数不超过 t, 同时敌手能获取每轮所有节点发送的消息并决定各节点所接收的消息, 适应性敌手则能在获取全部消息之后决定腐化策略.

扁平模型假设敌手能完全共谋并汇聚算力计算工作量证明, 然而它没有考虑实际诚实参与者每轮可完成大于 q 次哈希计算的事实, 进而低估了诚实参与者的实际算力. 在考虑消息时延时, 一般假设敌手会最大程度延迟诚实节点之间的消息传输, 且敌手内部可共谋而不存在时延. 在上述比实际环境更加严格的假设下, 若能证明共识安全, 那么实际协议执行的安全性便能得到保障.

2.2.3 动态加入与难度调整

在非授权设定的动态协议中, 参与者能随时加入并离开协议的执行, 随着参与人数的增加、矿机的使用与矿池的出现等, 全网总算力也随之改变. 在基于工作量证明的区块链协议中, 参与者通过进行一定难度的工作量证明计算来争取记账权 (出块权), 工作量证明难度需要根据全网总算力的变化动态调整以保证稳定的出块时间间隔, 因此工作量证明难度调节机制的安全性也是区块链共识研究中的重点.

部分研究为了简化分析, 假设协议中的难度调整机制是完美的, 即单位时间内找到的 PoW 解数量确定, 但这种假设过于理想. Pass 等 [4] 指出需考虑区块链网络参与人数动态变化的事实, 应把每轮参与人数和工作量证明难度视为可变量 $n(\cdot)$ 和 $p(\cdot)$, 但他们没有更深入地分析该设定下的中本聪共识安全性. Garay 等 [6] 在动态设定的同步网络模型下研究中本聪共识工作量证明难度调整机制的安全性, 并在时延受限网络模型下 [7], 形式化证明中本聪共识使用该机制时的安全性. 他们指出当每轮的节点数量在一定范围内波动时, 中本聪共识能满足一致性和活性.

2.3 工作量证明与哈希模型

在 PoW 共识中, 参与者通过计算出满足目标难度的哈希值来生成新区块, 该计算在大多数研究中被建模为理想的随机预言机 (random oracle, RO), 使用 RO 模型能简化共识执行的描述和安全性证

明. 在扁平模型中, 每个节点每轮能进行 q 次哈希计算, 相当于进行 q 次 RO 质询, 哈希计算满足目标难度的概率为 p, 那么 n 个节点每轮生成区块数的期望为 npq.

3 中本聪共识协议

本节基于第 2 节的定义和模型介绍中本聪共识安全性研究中所涉及的概念. 其中, 3.1 小节简单介绍了中本聪共识的概念; 3.2 小节介绍区块链协议中最常见的双花攻击; 3.3~3.5 小节介绍基于 PoW 的区块链协议的区块生成、区块验证和区块链视图.

3.1 中本聪共识

中本聪 [1] 提出了首个去中心化数字货币系统, 该系统是一个基于 PoW 与最长链原则的匿名区块链协议, 区块链中每个区块包含数字货币的交易记录, 并用密码学方式以一定顺序链接. 参与者之间的交易在广播后, 被矿工打包成区块并链接到当前区块链. 上述协议能构造不可篡改的公共账本, 实现参与者在无需第三方身份认证的前提下, 就交易内容达成共识. Garay 等 [2] 提出的比特币骨干协议基于 PoW 和最长链原则, 忽略比特币区块链协议中的输入验证、内容写入和链读取等细节, 即为一般认为的中本聪共识或中本聪协议.

3.2 双花攻击

对区块链协议所能进行的最直观攻击是双花攻击, 敌手在完成交易并获取相应服务后, 秘密创造一条不包括该交易的新链, 并通过各种手段使得新链长度超过原有包含该交易的链, 使得全网节点转移到新链上, 造成区块链的分叉, 达成颠覆交易的目的. 而敌手在原区块链的旧有交易中使用的数字货币能被他再次用于其他交易, 实现数字货币双花. 在比特币区块链中, 敌手双花攻击成功需要有足够强大的算力, 理论上只要他能掌控超过全网 50% 的总算力, 就能成功颠覆交易, 因此双花攻击也被称为 50% 攻击.

在区块链协议的实际应用中, 敌手可在网络层发动日蚀攻击, 通过覆盖节点的连接地址使节点连接到敌手控制的节点, 从而控制节点消息传输使其与网络分裂. 敌手还能利用现实网络环境存在的消息时延发动自私挖矿攻击, 在找到新区块后不马上公布, 而是接着该区块私下进行工作量证明, 构造仅自己所知的 “私链”, 在网络中存在其他节点公布新区块后, 选择性释放私链中的区块, 使诚实参与者视图产生分歧. 自私挖矿攻击结合日蚀攻击能分散诚实参与者算力, 让大部分节点接受敌手的私链, 使敌手可利用少于 50% 的算力成功实现数字货币双花.

实际上, 各衍生攻击最终的目的都是破坏协议一致性而改写区块链实现双花, 因此安全的区块链共识协议首先需要防止双花攻击, 即证明控制一定计算能力的敌手在完成交易后所创造的新链, 无法使用任何手段来超过并取代包含旧有交易的原区块链, 从而双花攻击无法成功.

3.3 工作量证明

工作量证明是区块链协议中矿工生成区块的一种方式, 实际上是对哈希函数的暴力计算. 协议定期指定一个工作量难度 D, 矿工在进行工作量证明前, 首先更新本地有效链并确定将写入区块链中的信息 x (比特币区块链中, 该信息为全网新产生的交易, 上一个区块哈希和时间戳等), 接着初始化计数器 η 并执行哈希计算. 矿工根据协议指定算法 $H(\cdot)$, 不断改变 η 的值, 直到 $H(\eta, x) < D$, 此时的 η 是满足难度要求的 PoW 有效解, 代表成功生成一个有效区块.

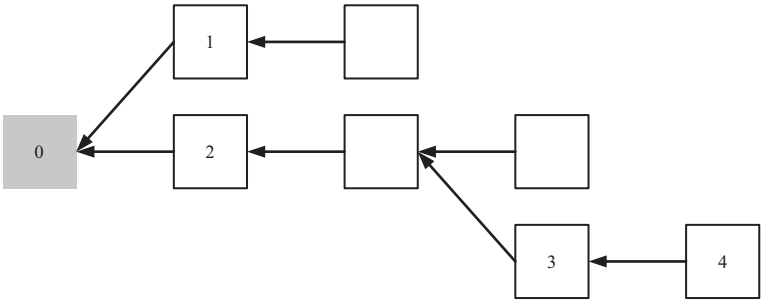

图 1 区块树

Figure 1 Blocktree

每当参与者生成一个区块, 他将获取相应区块奖励和交易佣金, 其计算能力越强, 单位时间可计算的哈希次数越多, 那么成功找到的 PoW 有效解越多, 得到的奖励也越多, 区块链协议通过这种方式激励参与者进行工作量证明计算.

3.4 区块创造与验证

区块链中每个区块可视为一个数组 (h_{-1}, η, m, h), 其中前一个区块的哈希值 h_{-1} 是区块指向前一个区块的指针, 用以保证区块链的顺序; η 为计数器常数, 矿工通过改变该常数计算工作量证明; m 是当前区块记录, 包括交易内容、交易双方地址、交易时间等相关信息; 当前区块指针 h 是该区块数据的哈希值. 协议存在挖矿难度函数 $p(\cdot)$, 随着协议的执行定期调整使区块能以一定的速率被全网生成, 它定义工作量证明难度 $D_p = p(\kappa) \cdot 2^{\kappa}$, 使矿工对任意 η, 挖矿成功概率 $Pr_{\eta}[H(h_{-1}, m, \eta) < D_p] = p(\kappa)$, 其中 κ 为安全参数 (比特币区块链中的 $\kappa = 128$).

矿工挖矿前, 需验证所更新的区块链是否有效, 区块链有效需满足其中的每个区块 $b=(h_{-1}, \eta, m, h)$ 对于上一个区块 $b_{-1} = (h'_{-1}, \eta', m', h')$ 有效, 区块有效需满足的 3 个条件如下:

(1) 指针指向上一个区块: $h_{-1} = h'$;

(2) 哈希值计算正确: $h = H(h_{-1}, \eta, m)$;

(3) 达到工作量难度要求: $h < D_p$.

作为公共账本的区块链, 其区块中的内容也需有效, 如交易支付方所花费金额不能多于其账户余额, 数字货币不能重复交易 (双花) 等, 因此矿工打包交易时还需验证交易内容是否合法.

3.5 比特币区块链

比特币区块链中的矿工更新本地链时, 首先检查网络中是否存在比自己本地链长的区块链, 选择其中最长的有效链替换原有本地链, 在新链的基础上挖矿以生成新的区块, 若成功找到 PoW 有效解, 则将新的区块添加到链上并广播给全网节点.

由于在短时间内, 可能同时有多个矿工找到有效区块, 并且网络时延将导致矿工视图中的最长链不一定是真正的全网最长链, 同时敌手可能故意在非最长链上挖矿以造成区块链分叉. 因此比特币区块链的真正视图是一个存在许多分叉的区块树, 如图 1 所示. 其中, 区块树中区块 0 为所有参与者视图中一致的创世区块; 区块 1 可能是敌手故意制造的分叉, 试图分裂全网节点; 区块 4 的生成者可能同时看到两条一样长度的区块链, 最终选择在区块 3 后面进行计算.

4 安全性质的形式化定义

比特币骨干协议及其后续研究 [2,3] 形式化定义了区块链所需满足的安全性质, 以及区块链作为分布式公共账本所需具备的安全性质. 将区块链用作交易账本的协议直观上需要满足一致性 (consistency) 和活性 (liveness). 作为公共账本之外, 区块链还被应用于其他如投票、众筹、拍卖等领域中, 为了研究这些应用的安全性, Garay 等 [2] 从比特币协议中抽象出底层比特币骨干协议并提出区块链协议需满足的 3 个安全性质: 共同前缀 (common prefix) 性质、链质量 (chain quality) 性质和链增长 (chain growth) 性质. 他们证明满足上述安全性质的区块链所构造的账本协议将满足一致性和活性. 4.1 小节介绍主要的区块链安全性质; 4.2 小节介绍公共账本的安全性质; 4.3 小节给出区块链安全性质的其他定义.

4.1 区块链安全性质

4.1.1 共同前缀性质

共同前缀性质 [2] 是区块链协议最重要的安全性质, 它保证协议参与者能就区块链顺序和内容达成一致. 满足共同前缀性质的区块链协议里, 两个诚实参与者的区块链有很大一部分都是相同的, 仅末尾最新的几个区块可能不同. Pass 等 [4] 认为共同前缀性质并不严谨, 不能避免全网参与者在奇数轮接受一条链, 而在偶数轮接受另一条链的特殊情况. 因此他们将共同前缀性质扩展为一致性, 分为普通一致性和未来自我一致性. 普通一致性指在任意时刻, 两个诚实参与者的区块链最多只有末尾的一定数量的区块不同. 未来自我一致性指任意诚实参与者在两个不同的时刻, 其区块链只有末尾一定数量的区块不同.

定义13 (共同前缀性质 Q_{cp} [2]) 任意两个诚实参与者 P_1, P_2 的区块链 C_1, C_2, 满足 $C_1^{\lceil k} \preceq C_2$, 即链 C_1 截去 k 个块后剩下的区块链属于链 C_2.

定义14 (一致性 [4]) 如果在轮 r_1 时存在链 C_1, 在轮 r_2 时存在链 C_2, 且 $r_1 \leqslant r_2$, 那么 $C_1^{\lceil k} \preceq C_2$, 即 C_1 截去最末尾 k 个区块后剩下的链属于链 C_2. 需注意的是, C_1 和 C_2 可能为同一个参与者在不同时刻的区块链.

4.1.2 链质量性质

链质量性质 [2] 意味着区块链中的大部分区块由诚实参与者所生成, 进而限制自私挖矿及其衍生攻击攻击等.

定义15 (链质量性质 Q_{cq} [2]) 在诚实参与者区块链的任意连续 ℓ 个区块中, 由敌手生成的区块所占比例最多为 μ. 理想的链质量性质中, 参与者贡献的区块比例与其算力成正比.

4.1.3 链增长性质

链增长性质 [3] 描述在一定时间内区块链至少增长的区块数, 该性质反映了区块链协议处理数据的速度, 保证协议能正常持续执行.

定义16 (链增长性质 Q_{cg} [3]) 区块链协议执行 s 轮后, 任意诚实参与者的区块链, 至少增加 $\tau \cdot s$ 个区块.

4.2 公共账本安全性质

区块链普遍应用于数字货币的公共账本, 基于共同前缀、链质量和链增长这 3 个区块链安全性质,

Garay 等 [2] 提出使用区块链作为公共账本的鲁棒协议所需满足的两个安全性质 —— 一致性和活性, 这两个性质能由区块链的安全性质推出, 即安全的区块链能构造安全的公共账本. 一致性意味着一个交易若存在于某诚实参与者所承认的区块链账本中且有一定深度, 那么该交易必须也同时存在于其他诚实参与者的区块链中, 同时内容和所在位置均一致, 它保证所有参与者需就交易顺序和内容达成共识, 即他们有统一的区块链账本视图, 防止双花攻击. 活性意味着合法交易会被区块链及时处理并在所有诚实参与者的区块链上实现稳定, 且内容和所在位置均相同, 它保证账本协议能持续运行, 防止拒绝服务攻击.

定义17 (稳定交易) 若交易 tx 所在的区块后有超过 k 个区块, 那么该交易为稳定交易, 其中 $k \in \mathbb{N}$ 为 "深度 (depth)" 参数.

定义18 (一致性 [2]) 若交易 tx 在时刻 t 是某诚实参与者区块链中的稳定交易, 那么在时刻 t 后, 所有诚实参与者将认为 tx 是稳定交易, 且 tx 在他们区块链上的位置相同.

定义19 (活性 [2]) 若交易 tx 合法生成且不与过去的稳定交易冲突, 当 tx 连续 μ 轮作为输入广播给所有诚实参与者后, 所有诚实参与者将认为 tx 是稳定交易, 其中 μ 为 "等待时间".

Avarikioti 等 [12] 研究中本聪共识是否能在短暂的敌手主导攻击后, 恢复账本原有的安全性质. Badertscher 等 [13] 在其基础上引入适应性敌手, 形式化定义公共账本所需满足的第 3 个安全性质 —— 自愈性质, 用以描述基于区块链的账本协议受到上述攻击后, 恢复安全性所需的时间与敌手算力增量之间的函数关系.

定义20 (自愈性质 [13]) 若敌手在 (ρ_a, ρ_b) 期间用超过全网 50% 的算力发动攻击, 那么满足自愈性质的账本协议在 $(\rho_a - \tau_l, \rho_b + \tau_h)$ 之外, 一致性和活性不被满足的概率可忽略, 其中 τ_l, τ_h 决定协议脆弱期 (vulnerability period).

4.3 其他安全性质

区块链安全性研究基于不同的假设和研究方法, 提出了许多不同的安全性质定义. 比如 Ren [8] 认为区块链协议需满足安全性和活性, 其中安全性意味着诚实参与者不承认与自己本地链上的某区块高度相同但内容不同的其他区块, 活性意味着协议中的合法交易最终会被所有的诚实参与者承认. 他所定义的这两个性质能和 Garay 等 [2] 提出的共同前缀、链质量和链增长性质互相推导.

基于不同的模型, 假设和所研究的安全问题, 不同研究的协议描述和安全性定义也会随之变化. 目前针对区块链和账本协议的安全性研究大多证明文献 [2] 提出的区块链安全性质和账本协议安全性质能否在一定假设下被满足.

5 中本聪共识安全性证明模型和关键

在实际复杂的网络条件和执行环境下, 如何抽象底层区块链逻辑, 并使用数学模型进行形式化安全证明, 是共识安全性研究面临的最大挑战. 本节归纳了中本聪共识协议在不同模型下的形式化描述和安全性证明方法, 这些方法不仅能应用于中本聪共识的研究, 还能应用于其他区块链协议的安全性研究. 研究区块链协议安全性需首先形式化定义协议执行模型和相关安全性质, 接着描述敌手能力, 最终得出共识安全性满足时的约束条件.

由于区块链协议的执行环境存在时延, 如何准确描述协议执行一段时间后区块链的增长情况是其安全性研究的一大挑战. 基于轮的离散时间模型中, 参与节点在扁平模型下每轮执行固定次数的 PoW

计算, 区块以一定的速度生成. 基于时段的离散时间模型对协议执行的描述更符合实际, 当所选取的时段长度足够小时, 它可近似为连续时间模型, 使用该模型的方法将参与者 PoW 计算建模为离散近似的泊松过程, 泊松参数确定每个时段参与者找到区块数. 连续时间模型将 PoW 计算建模为泊松过程, 参与者在单位时间内找到的 PoW 有效解数根据其算力由相应泊松期望确定. 不同区块链协议的研究, 需根据其执行环境和具体安全问题选择不同的模型.

共识安全性研究需准确描述敌手能力和攻击方式, 并确定敌手在什么情况下攻击失败. 由于诚实参与者之间存在竞争而敌手之间能共谋, 敌手可通过控制消息传输等手段, 使得诚实参与者之间无法就区块链内容和顺序达成一致从而被分裂. 因此, 为了判定敌手是否攻击成功, 需研究诚实参与者视图达成一致的条件. 若敌手链长度无法超过诚实参与者视图达成一致时的区块链长度, 那么敌手攻击失败. Garay 等 [2] 在同步网络下提出独立成功轮 (uniquely successful round) 的概念, 证明在独立成功轮生成的区块会被所有诚实参与者所接受. Pass 等 [4] 在半同步网络模型中提出汇聚机会 (convergence opportunity) 的概念, 证明汇聚机会的出现能使诚实参与者区块链视图达成一致. Ren [8] 提出孤立 (loners) 的概念描述连续时间模型中的汇聚机会.

5.1 离散时间模型下的研究

在基于轮的离散时间模型中, 5.1.1 小节以 Garay 等 [2] 在同步网络下的中本聪共识研究为例, 介绍使用典型执行对区块链协议执行过程建模的安全性研究方法; 5.1.2 小节以 Pass 等 [4] 在时延受限网络下的研究为例, 介绍使用 F_{tree} 模型对区块链协议执行过程建模的安全性研究方法; 5.1.3 小节以 Kiffer 等 [5] 在时延受限网络下的研究为例, 介绍使用马尔可夫链建模区块链协议中敌手链与诚实链竞争的安全性研究方法. 在基于时段的离散时间模型中, 5.1.4 小节以 Gazi 等 [10] 的研究为例, 介绍使用特征字符串 (characteristic strings) 描述每个时段的协议执行情况的安全性研究方法.

5.1.1 典型执行方法

用典型执行描述区块链协议的区块生成情况能排除特殊事件给协议执行带来的偶然性. 在典型执行的协议中, 参与节点每次 RO 哈希计算视为一次伯努利试验, 成功找到区块的概率为 p, 因此可用二项分布期望描述每轮参与者区块生成情况.

比特币骨干协议 [2] 证明在同步网络中的诚实主导设定下, 中本聪共识安全性遭到破坏的可能性可忽略. 在基于轮的离散时间模型中, 每轮执行情况受 3 个事件影响, 即存在诚实节点找到 PoW 有效解, 有且仅有一个诚实节点找到一个 PoW 有效解和敌手控制的第 k 个节点的第 j 次哈希计算找到 PoW 有效解. 用布尔变量 X_i, Y_i, Z_{ijk} 表示这些事件在第 i 轮是否发生, 若事件发生, 则布尔变量等于 1, 否则等于 0, 其中, 若 $X_i = 1$, 则轮 i 为成功轮, 当 $Y_i = 1$, 轮 i 为独立成功轮. 因为敌手能同时控制多个节点寻找 PoW 有效解, 且会尽可能多地找到不同的区块以分裂诚实参与者视图, 所以敌手在第 i 轮找到的区块数为 $Z_i = \sum_{k=1}^{t}\sum_{j=1}^{q} Z_{ijk}$. 协议执行时每轮概率事件发生的期望如下:

$$
\begin{aligned}
&\mathrm{E}[X_i] = 1 - (1-p)^{q(n-t)},\\
&\mathrm{E}[Y_i] = q(n-t)p(1-p)^{q(n-t)-1},\\
&\mathrm{E}[Z_i] = pqt.
\end{aligned}
$$

若协议在一段时间内执行了 S 轮, $X(S) = \sum_{r\in S} X_r$ 表示该期间成功轮数, 同理, $Y(S)$ 和 $Z(S)$ 分别表示该期间的独立成功轮数和敌手所找到的全部区块数. 根据成功轮和独立成功轮的数量与敌手找到

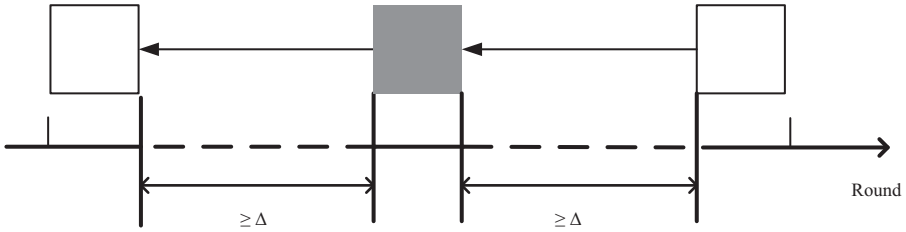

图 2 汇聚机会

Figure 2 Convergence opportunity

的有效解数之间的关系, 可证明共识是否满足安全性质, 如当 $Z(S) \leqslant Y(S)$ 时, 比特币区块链满足共同前缀性质.

若实际协议执行时, 敌手和诚实节点实际找到的区块数不偏离上述期望太多, 且没有恶性事件的发生 (PoW 哈希计算引起的区块插入、重复、预测等事件), 那么协议是典型执行. 使用 Martingale 约束可证明几乎所有实际协议为典型执行, 因此可使用上述期望描述实际协议的区块生成情况. 因为典型执行中诚实节点区块链的时间戳精确, 即可忽略时间戳和实际产生时间相差很远的区块存在, 且比特币骨干协议 [2] 假设协议难度调整机制完美, 所以在诚实主导条件下, 比特币区块链满足共同前缀和链质量性质, 因此可使用中本聪共识构造鲁棒的公共账本.

使用典型执行能对存在网络时延和动态设定下的区块链协议建模, 只需针对不同的执行环境和假设来调整典型执行的定义. 因此该方法适用于不同条件下的共识研究, 如同步网络和时延受限网络模型下动态设定的中本聪共识安全性研究 [6,7], 敌手进行短期敌手主导攻击后的中本聪共识安全性恢复研究 [12] 等. 但使用典型执行对 PoW 之外协议的建模比较复杂, 且该方法基于轮的离散时间模型, 需假设世界时钟存在而过于理想.

5.1.2 F_{tree} 模型

Pass 等 [4] 在时延受限网络模型下研究中本聪共识的安全性, 使用 F_{tree} 模型排除哈希计算带来的偶然性, 并排除敌手延迟消息反而使得诚实节点更快找到 PoW 有效解的特殊情况. 此模型下的参与节点不直接挖矿, 而是调用理想挖矿函数 F_{tree} 生成区块. 他们使用 Canetti [20] 提出的通用可组合安全 “模拟技术” 证明存在理想挖矿函数 F_{tree}, 使得调用此函数寻找 PoW 的理想协议和现实调用哈希函数挖矿的协议不可区分. F_{tree} 可随机决定诚实节点是否挖矿成功, 且成功概率与该节点所延长的当前链无关. 敌手在现实世界进行的攻击均能模拟为 F_{tree} 模型下的攻击, 因此若使用 F_{tree} 的共识协议满足区块链安全性质, 那么实际协议也满足.

时延受限网络模型存在延迟上限 Δ, 考虑将所有消息均延迟 Δ 的最强能力敌手, 协议中诚实节点由于时延可能导致视图分裂产生链分叉, Pass 等 [4] 用 “汇聚机会” 描述诚实节点就区块链视图达成共识的情况.

定义21 (汇聚机会 [4]) 在以轮为单位的共识执行过程中:

(1) 首先存在一个 Δ 轮的 “‘沉默” 期, 期间没有诚实节点挖出区块;

(2) 沉默期之后的那一轮为独立成功轮, 即有且只有一个诚实节点挖出区块;

(3) 该轮后, 又紧跟一个 Δ 轮的沉默期.

汇聚机会如图 2 所示, 其中深色区块是当前轮诚实节点找到的唯一一个区块, 其前后区块的挖出轮数间隔至少为 Δ. 它的出现代表所有诚实节点均汇聚于同一条链上. 区块的传播最多被延迟 Δ 轮,

因此第 1 个沉默期后, 诚实节点就区块链长度达成一致 (但可能是不同链). 接着当唯一的诚实节点找到新区块并广播后, 第 2 个沉默期结束时他的链将传播给所有诚实节点, 因为沉默期不存在其他节点找到区块, 所以该链最长, 其他节点均会承认并接受这条链.

F_{tree} 模型下, 诚实节点在某轮成功挖出区块的概率为 $\alpha = 1-(1-p(\kappa))^{(1-\rho)n}$, 敌手一轮挖出区块数的期望值为 $\beta(\kappa,n,\rho,\Delta) = \rho np(\kappa)$, 其中 $p(\kappa)$ 为每次哈希计算成功的概率. 若区块链协议执行一段时间后, 有 L 轮诚实节点找到区块, 之前有超过 Δ 轮沉默期的独立成功轮数为 X, 文献 [4] 证明汇聚机会数至少为 $2X-L \geqslant (1-2\alpha(\Delta+1))\alpha t$. 在长为 t 的执行时期内, 敌手攻击成功的唯一方式是在这段时间内挖出足够多能被诚实节点接受的区块, 因此若敌手挖出的区块数 β 少于汇聚机会数, 那么协议安全性能得到保障.

Pass 等 [4] 使用 F_{tree} 模型描述共识协议执行中诚实节点和敌手之间的竞争, 便于在通用可组合模型中描述共识协议, 利于中本聪共识与其他协议模块化安全组合, 并提出汇聚机会以描述诚实节点达成共识的情况, 但他们对汇聚机会次数的计算较复杂并存在精确空间.

5.1.3 马尔可夫链模型

Kiffer 等 [5] 使用马尔可夫链描述区块链协议执行过程以及敌手与诚实节点区块链之间的竞争, 更精确地计算了 Pass 等 [4] 提出的汇聚机会, 并分析中本聪共识的安全性. 在马尔可夫链模型下汇聚机会的状态变化参考文献 [5] 的图 2, 其中每轮敌手链和诚实链的竞争情况都对应于马尔可夫链上的一个状态, 状态之间的转换取决于 (1) 诚实节点是否挖到新区块; (2) 敌手是否挖到新区快; (3) 是否有足够长的沉默时间.

马尔可夫链中描述的汇聚机会存在两种状态, 其中状态 S_0 表示任意两个诚实节点找到区块的时间间隔小于 Δ, 状态 S_1 表示至少存在 Δ 的沉默期. 令 E_{ij} 表示由状态 S_i 转到 S_j 的期望次数, 那么 Pass 等 [4] 计算的汇聚机会次数在马尔可夫链中的表述为 $2X-L = 2(E_{01}+E_{11})-(E_{00}+E_{11}+2E_{10})$. Kiffer 等 [5] 指出在马尔可夫链中呈现的汇聚机会次数实际上等于到达状态 S_1 的次数 E_{11}, 因此上述计算低估了真实的汇聚机会数.

在马尔可夫链模型下, 共识执行过程被描述为一个均匀时间, 不可约与遍历的 (time-homogeneous, irreducible, ergodic) 马尔可夫链, 敌手链和诚实链的竞争情况对应于马尔可夫链上不同的状态, 协议每执行一轮都可能引起状态变化. 马尔可夫链的静态分布能描述特定状态 (如汇聚机会) 出现的期望值, 由 Chernoff-Hoeffding 约束可证明在协议执行足够长时间后, 特定事件偏离其期望的概率随协议执行的轮数呈指数级下降, 因此可使用马尔可夫链的静态分布证明中本聪共识的安全性. 在该模型下, 若敌手生成的区块数少于汇聚机会数, 那么共识一致性得到保障.

上述方法能更精确地计算汇聚机会, 并能更直观地描述共识执行时的状态变化与敌手的攻击. 除了中本聪共识外, 这种方法还适用于 Cliquechain 共识和 GHOST 协议等树状区块链协议的安全分析上. 但马尔可夫链的状态变化仍以轮为单位, 需假设世界时钟的存在.

5.1.4 泊松过程离散近似与特征字符串

因为现实工作量证明是在连续时间中的无记忆性哈希计算, 所以区块的生成能建模为泊松过程, 敌手和诚实节点是否成功找到 PoW 解满足泊松分布, 其泊松参数分别为 r_a 和 r_h. 与上述 3 种以轮为执行单位的方法不同, Gazi 等 [10] 在时延受限网络中, 通过将时间分为长度为 s 的执行单位 – 时段, 对 PoW 泊松过程进行标准离散近似, 进而简化了敌手与诚实节点的竞争过程.

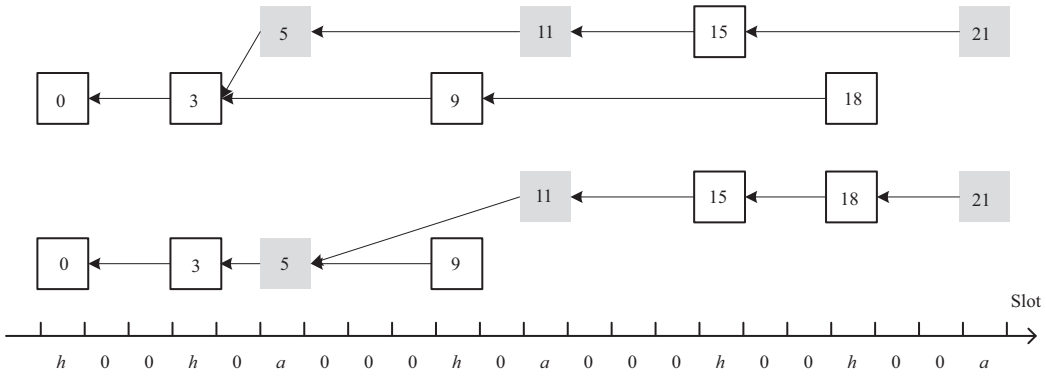

图 3 同一个特征字符串所对应的不同分叉

Figure 3 Forks of the same characteristic string

在长度为 L 的时间间隔中存在 L/s 个时段, 由于区块生成为泊松过程, 根据联合界 (union bound) 可得, 在一个时段内有两个区块生成的概率为 $O(L\lambda^2 s)$. 因此当 s 足够小时, 一个时段内最多能找到一个 PoW 有效解, 进而可使用特征字符串来描述中本聪共识区块生成状态. 假设每个时段存在一个特征字符 w_i, w_i 的取值为 h, a, 0, 分别对应该时段有诚实参与者找到一个有效解, 敌手找到一个有效解和不存在节点找到有效解. 特征字符取值概率分别为 p_h, p_a 和 $1-p_a-p_h$, 其中 $p_a = r_a \cdot s$, $p_h = r_h \cdot s$ 分别为敌手和诚实参与者在一个时段中成功找到 PoW 有效解的概率.

特征字符构成的特征字符串表示一段时间的协议执行中诚实参与者和敌手区块生成的情况. 一条特征字符串能表示的区块生成情况能构成不同的区块树, 又称为区块链分叉, 如图 3 所示. 图 3 中白色和灰色区块分别代表当前时段存在诚实节点或敌手找到有效区块, 区块上的数字代表区块在哪个时段被找到. 一个特征字符串可能生成不同的区块分叉, 在进行安全性分析时, 通常考虑敌手产生最大分叉链的区块分叉. 某特征字符串生成的区块树中存在敌手生成的分叉链和大部分诚实节点接受的主链.

特征字符串模型使用优势描述区块树中某条支链与全网最长诚实链之差, 使用边际差额描述某分叉的最大优势. 特征字符串所能形成的所有区块树的最大边际差额, 能描述敌手能进行的最优攻击, 即敌手能产生的最长分叉链. 特征字符串分为一个个步骤 (step), 每个步骤以 h 结尾且至少有 Δ 个特征字符, 用以描述诚实节点达成一致的视图. 协议每执行一步后的边际差额变化, 反映敌手能产生的最长分叉链的变化. Gazi 等 [10] 证明当敌手和诚实节点的算力满足 $p_a < \frac{1}{\Delta-1+1/p_h}$ 时, 每步边际差额的增长小于负数 $p_a/\theta - 1$, 因此敌手生成的最大分叉链短于大多数诚实节点所接受的主链, 诚实节点不会转移到敌手链上导致交易重写, 中本聪共识满足一致性.

使用泊松过程离散近似模型建模的中本聪共识协议, 当其时段的长度 s 趋近于零时, 它可转化为符合现实连续模型下的泊松过程. 除此之外, 这种建模方法能很好地描述敌手能进行最优攻击, 还能应用于其他基于 PoS 的共识安全性研究中, 但证明过程比较复杂.

5.2 连续时间模型下的协议执行

Ren [8] 指出能用更简洁的方式在连续时间模型下研究中本聪共识协议, 他把 PoW 挖矿过程建模为均匀分布的泊松点过程, 诚实节点和敌手在单位时间找到的区块数用泊松参数表示. Dembo 等 [11] 在连续时间模型中研究比特币协议一致性得到保障时的敌手算力最大限制, 将敌手对诚实节点进行的各式攻击转换为各敌手树和虚拟诚实链的竞争. 5.2.1 和 5.2.2 小节以上述研究为例介绍在连续时间模

型下的区块链协议模型与安全性研究方法.

5.2.1 连续时间的中本聪共识分析

Ren [8] 在时延受限模型下使用泊松参数 α 和 β 分别表示诚实节点和敌手的挖矿率, 提出无后档 (non-tailgaters) 和孤立 (loners) 的概念, 用以描述诚实节点达成共识的条件.

定义22 (无后档和孤立) 假设诚实节点在时刻 t 找到区块 B, 若在 $t-\Delta$ 到 t 的时间内没有诚实节点找到其他区块, 那么 B 是无后档区块. 若在 $t-\Delta$ 到 $t+\Delta$ 的时间内没有其他诚实节点找到区块, 那么 B 是孤立区块. 诚实节点找到的一个区块为无后档区块的概率是 $g=\mathrm{e}^{-\alpha\Delta}$, 为孤立区块的概率是 g^2.

在无后档区块的生成与上一个诚实区块之间时长至少相差 Δ, 因此孤立区块是无后档区块, 同时它后面的诚实区块也是无后档区块. 在时延受限模型下的中本聪共识中, 一个诚实区块如果不是无后档区块, 那么在它到达某节点时, 它上一个区块可能未被该节点接受, 使得它被 “浪费”. 因此区块链中, 无后档区块和孤立区块的数量能反映协议执行安全性.

孤立区块是其区块高度上的唯一区块, 它的出现意味着诚实节点就包含孤立区块的那条链达成共识. 孤立区块可对应 Pass 等 [4] 在离散时间模型中提出的汇聚机会. 在中本聪共识中, 若孤立区块数多于敌手所生成的区块数, 那么共识安全性就能得到保障.

Ren 用一种较简单的方法在连续时间模型研究中本聪共识安全性, 但他缺少对协议执行环境和敌手的建模与形式化描述.

5.2.2 区块树分割

Dembo 等 [11] 针对所有使用最长链原则的区块链协议, 重新排列全网视图中的区块树, 构造由诚实区块组成的虚拟诚实链 (fictitious honest chain) 和由敌手区块构成的敌手树, 将攻击建模成虚拟诚实链和敌手树的竞争, 通过区块树分割 (blocktree partitioning) 方法研究中本聪共识中的敌手最优攻击. 此外, 他们提出的中本聪共识在时延受限模型下满足一致性时的条件, 与 Gazi 等 [10] 在离散时间模型下得出的结论一致.

在中本聪白皮书研究的攻击里, 敌手将全部算力集中于一条敌手链, 并且每生成一个区块便广播给全部节点, 区块树中仅存在一条诚实链和一条敌手链的竞争, 如文献 [11] 的图 2(a) 所示. 而在一般攻击中, 敌手会在不同的支链后添加区块, 试图分裂诚实节点的视图和算力, 区块树包含多个分叉, 如文献 [11] 的图 2(b) 所示.

区块树分割将区块树中所有诚实参与者生成的区块排列成一条虚拟诚实链, 将其他敌手区块放入一个以某诚实区块为根的敌手树中, 把所有攻击转化为某敌手树和虚拟诚实链之间的竞争. 对于所有基于最长链原则的区块链协议, 其区块树均能分割成敌手树和虚拟诚实链的形式, 因此可用他们之间的竞争描述该协议可能出现的任何攻击. 通过区块树分割可看出, 中本聪共识的敌手最优攻击为比特币白皮书研究的基础攻击 [1], 且当敌手树的增长慢于虚拟诚实链的增长时, 敌手攻击失败.

Dembo 等 [11] 提出中本聪块 (Nakamoto blocks) 的概念描述诚实节点达成共识的情况, 如图 4 所示. 图 4 中加粗边框区块为中本聪块, 现实中的所有攻击可看作不同的敌手树以各自增长速度与虚拟诚实链进行竞争. 中本聪块由诚实节点生成, 它的出现意味着此前所有的敌手树都无法追上到达该区块的诚实链, 同时该诚实链也是最长链, 其上所有中本聪块之前的区块都稳定且不可篡改. 现实世界中诚实节点的区块链增长速度不慢于虚拟诚实链的增长, 且虚拟诚实链的增长速率至少为 $\frac{\lambda_h}{1+\lambda_h\Delta}$, 其中 λ_h 为诚实节点区块生成的泊松参数. Dembo 等证明若敌手挖矿率 $\lambda_a<\frac{\lambda_h}{1+\lambda_h\Delta}$, 那么中本聪块将以

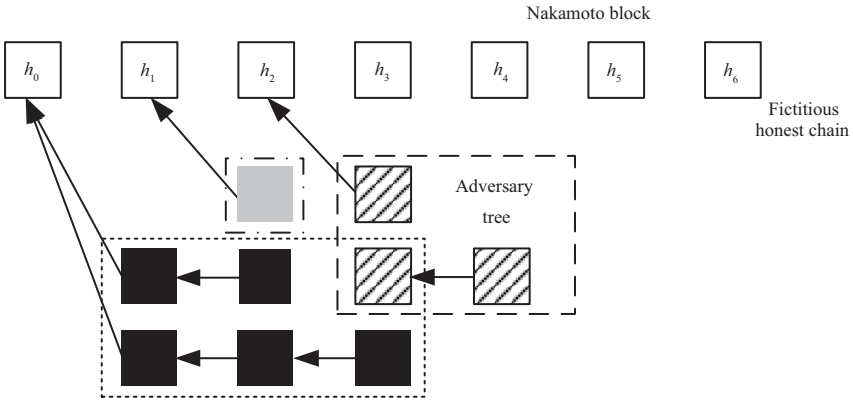

图 4 中本聪块

Figure 4 Nakamoto blocks

高于敌手块的频率出现, 共识安全性得到保障.

大多数使用最长链原则的共识能使用区块树分割的方法描述敌手的不同攻击手段, 但对基于最大难度等其他链选择原则的共识或树状结构区块链的共识而言, 该方法存在局限性.

6 总结与展望

本文总结了主流的中本聪共识研究方法, 并给出其网络模型、执行模型和安全模型. 本文归纳的区块链及其账本协议安全性质, 为新的区块链协议提出基础安全要求. 今后, 在研究中本聪共识与其他区块链协议的共识安全性时, 可根据协议特征, 所需的安全性质与协议执行环境, 在适当的网络模型下, 基于本文总结的研究方法, 对协议建模并计算安全性满足的条件. 在尽可能贴近实际的情况下, 简化区块链协议的安全性证明过程.

区块链协议的安全性作为区块链技术发展的基础, 其安全研究趋势如下.

第一, 在网络模型层面, 网络模型需更符合实际网络对参与节点有效性验证的影响, 如需考虑区块链网络拓扑结构带来的节点视图差异, 考虑网络实际存在的通信时延与敌手攻击策略导致的消息, 交易和区块传播顺序的改变.

第二, 在时间模型层面, 对中本聪共识建模所使用的时间模型分为连续时间模型和离散时间模型, 连续模型较符合实际协议的执行, 但是在对攻击的描述上存在局限性, 难以刻画适应性敌手对非授权共识安全性的影响; 离散时间模型一般以轮作为协议执行单位, 在描述参与节点动态变化与敌手攻击行为上存在优势, 然而这种模型需假设同步时钟的存在. 因此研究共识安全性时, 一方面可根据研究重点选择合适的时间模型, 另一方面, 可考虑连续模型的离散近似模型, 如使用特征字符串描述协议执行情况等.

第三, 在敌手层面, 需考虑更灵活的现实攻击者, 目前的研究方法难以描述诚实参与者可能进行的自私挖矿等合理攻击, 这种参与者通常被归为攻击者, 而实际上他们有利于区块链共识达成一致. 因此敌手模型在考虑适应性敌手的同时, 还需考虑 “理智” 的诚实参与者, 并且由于真正的攻击往往是持续一定时间的敌手主导攻击, 研究敌手行为可限制其攻击时间, 并增大协议能承受的敌手能力阈值.

第四, 在参与者层面, 针对中本聪共识非授权设定下的参与人数动态变化, 需考虑 PoW 难度调整机制带来的安全性影响, 以及新加入参与者对网络同步性造成的影响. 此外, 在目前的数字货币协议

中, 真正的交易参与者大多并不是参与共识的矿工, 他们在安全证明模型中通常被忽略, 但他们的区块链视图对共识一致性存在见证作用, 因此考虑该参与者的贡献, 可提高协议安全阈值, 有利于扩展共识应用, 在新的共识协议里, 也可增加相关激励机制使他们能参与到共识安全性维护中.

第五, 在应用层面, 在将本文提及的研究方法应用于其他协议时, 需考虑到不同共识的特点, 如基于 PoS 协议的权益更新计算等. 还需考虑区块链与其他协议组合应用时的安全性, 如将其应用于物联网等工业互联网系统中等. 最后还需考虑共识数据交互带来的安全风险, 如通信层数据传输造成的参与者身份泄露、参与者出块数据带来的能力泄露、区块链之间的数据传输造成的信息泄露等.

参考文献

1 Nakamoto S. Bitcoin: a peer-to-peer electronic cash system. Working Paper, 2008. https://bitcoin.org/bitcoin.pdf
2 Garay J, Kiayias A, Leonardos N. The bitcoin backbone protocol: analysis and applications. In: Proceedings of Annual International Conference on the Theory and Applications of Cryptographic Techniques, 2015. 281–310
3 Kiayias A, Panagiotakos G. Speed-security tradeoffs in blockchain protocols. IACR Cryptol, 2015. https://eprint.iacr.org/2015/1019.pdf
4 Pass R, Seeman L, Shelat A. Analysis of the blockchain protocol in asynchronous networks. In: Proceedings of Annual International Conference on the Theory and Applications of Cryptographic Techniques, 2017. 643–673
5 Kiffer L, Rajaraman R, Shelat A. A better method to analyze blockchain consistency. In: Proceedings of ACM SIGSAC Conference on Computer and Communications Security, 2018. 729–744
6 Garay J, Kiayias A, Leonardos N. The bitcoin backbone protocol with chains of variable difficulty. In: Proceedings of Annual International Cryptology Conference, 2017. 291–323
7 Garay J A, Kiayias A, Leonardos N. Full analysis of Nakamoto consensus in bounded-delay networks. IACR Cryptol, 2020. https://eprint.iacr.org/2020/277.pdf
8 Ren L. Analysis of Nakamoto consensus. IACR Cryptol, 2019. https://eprint.iacr.org/2019/943.pdf
9 Li J, Guo D N. Continuous-time analysis of the bitcoin and prism backbone protocols. 2020. ArXiv:2001.05644
10 Gazi P, Kiayias A, Russell A. Tight consistency bounds for bitcoin. In: Proceedings of ACM SIGSAC Conference on Computer and Communications Security, 2020. 819–838
11 Dembo A, Kannan S, Tas E N, et al. Everything is a race and Nakamoto always wins. In: Proceedings of ACM SIGSAC Conference on Computer and Communications Security, 2020. 859–878
12 Avarikioti G, Käppeli L, Wang Y Y, et al. Bitcoin security under temporary dishonest majority. In: Proceedings of International Conference on Financial Cryptography and Data Security, 2019. 466–483
13 Badertscher C, Gazi P, Kiayias A, et al. Consensus Redux: distributed ledgers in the face of adversarial supremacy. IACR Cryptol, 2020. https://eprint.iacr.org/2020/1021.pdf
14 Lynch N A. Distributed Algorithms. San Francisco: Morgan Kaufmann Publishers Inc., 1996
15 Stifter N, Judmayer A, Schindler P, et al. Agreement with satoshi-on the formalization of Nakamoto consensus. 2018. https://spiral.imperial.ac.uk/handle/10044/1/62946
16 Garay J, Kiayias A. SoK: a consensus taxonomy in the blockchain era. In: Proceedings of the Cryptographers' Track at the RSA Conference, 2020. 284–318
17 Yuan Y, Ni X C, Zeng S, et al. Blockchain consensus algorithms: the state of the art and future trends. Acta Autom Sin, 2018, 44: 93–104
18 Yang Y G, Zhang S X. Review and research for consensus mechanism of block chain. J Inform Secur Res, 2018, 2018: 369–379
19 Liu Y Z, Liu J W, Zhang Z Y, et al. Research on consensus mechanisms of blockchain technology. J Cryptologic Res, 2019, 6: 395
20 Canetti R. Universally composable security: a new paradigm for cryptographic protocols. In: Proceedings the 42nd IEEE Symposium on Foundations of Computer Science, 2001. 136–145
21 Katz J, Maurer U, Tackmann B, et al. Universally composable synchronous computation. In: Proceedings Theory of Cryptography Conference, 2013. 477–498
22 Dwork C, Lynch N, Stockmeyer L. Consensus in the presence of partial synchrony. J ACM, 1988, 35: 288–323

23 Kiayias A, Russell A, David B, et al. Ouroboros: a provably secure proof-of-stake blockchain protocol. In: Proceedings of the 37th Annual International Cryptology Conference, 2017. 357–388

Methods of security analysis for Nakamoto consensus

Ziyu ZHOU[1], Zongyang ZHANG[1, 2*] & Jianwei LIU[1]

1. *School of Cyber Science and Technology, Beihang University, Beijing* 100191*, China*;
2. *Key Laboratory of Aerospace Information Security and Trusted Computing, Ministry of Education, Wuhan University, Wuhan* 430072*, China*
* Corresponding author. E-mail: zongyangzhang@buaa.edu.cn

Abstract Nakamoto consensus is one of the most fundamental consensus used in blockchain protocols. Its security is of significance in the field of blockchain and cryptocurrency applications. Existing studies have analyzed and proved the security of Nakamoto consensus from different aspects under various assumptions and models. This paper systematically summarizes the existing mainstream formalization and research methods for the security of Nakamoto consensus. Firstly, this paper describes the execution model of Nakamoto consensus including the time model, network model and attacker model. Secondly, this paper summarizes the formal definitions of the security properties of Nakamoto consensus. Thirdly, according to the time model, this paper divides different analysis methods into the discrete-time model and continuous-time model, and points out their advantages and disadvantages. Finally, this paper points out the future research directions for blockchain consensus security.

Keywords Nakamoto consensus, blockchain, proof of security, cryptocurrency, proof-of-work