PoW 共识算法中的博弈困境分析与优化

唐长兵[1] 杨 珍[1] 郑忠龙[1] 陈中育[1] 李 翔[2,3]

摘 要 区块链是随着比特币等数字加密货币逐渐兴起而盛行的一种新型去中心化分布式系统, 具有去中心化、时序数据、集体维护、可编程和安全可信等特点. 目前, 区块链已引起政府部门、金融机构、科技企业和资本市场的高度重视与广泛关注. 如何在一个去中心化的分布式系统中高效地达成共识是区块链技术研究的重要问题. 本文从工作量证明 (Proof of work, PoW) 共识算法的挖矿困境入手, 分析 PoW 共识过程中矿工策略选择的纳什均衡存在条件. 利用零行列式 (Zero determinant, ZD) 策略对矿工策略选择进行优化, 并通过数值仿真来验证优化算法的有效性. 概括来说, 本文从博弈论角度来理解和剖析 PoW 共识算法, 为进一步设计基于博弈论的共识算法提供新的思路和方法.

关键词 区块链, 工作量证明, 共识算法, 区块截留攻击, 纳什均衡, 零行列式策略

引用格式 唐长兵, 杨珍, 郑忠龙, 陈中育, 李翔. PoW 共识算法中的博弈困境分析与优化. 自动化学报, 2017, **43**(9): 1520−1531

DOI 10.16383/j.aas.2017.c160672

Game Dilemma Analysis and Optimization of PoW Consensus Algorithm

TANG Chang-Bing[1] YANG Zhen[1] ZHENG Zhong-Long[1] CHEN Zhong-Yu[1] LI Xiang[2,3]

Abstract Blockchain is a new decentralized distributed system with the prevalence of Bitcoin and other cryptocurrencies, whose characteristics include decentralization, time-series data, collective maintenance, programmability, security, and so on. Currently, blockchain has attracted intensive attention from governments, financial institutions, high-tech enterprises, and capital markets. Under the framework of this decentralized distributed system, one of the key research issues is how to reach a consensus effectively. In this paper, we analyze the Nash equilibria existence of the miner′s strategy choice in the process of proof of work (PoW) consensus algorithm. Besides, we apply the zero determinant (ZD) strategy to optimize the strategy choosing of the miner, and verify the effectiveness of the optimization algorithm through numerical simulation. In brief, this work contributes understanding and analyzeing the PoW consensus algorithm, and provides a new idea and method for the design of consensus algorithm based on the game theory.

Key words Blockchain, proof of work (PoW), consensus algorithm, block withholding attack, Nash equilibrium, zero determinant (ZD) strategy

Citation Tang Chang-Bing, Yang Zhen, Zheng Zhong-Long, Chen Zhong-Yu, Li Xiang. Game dilemma analysis and optimization of PoW consensus algorithm. *Acta Automatica Sinica*, 2017, **43**(9): 1520−1531

区块链技术最初是为比特币设计的一种特殊数据库技术, 它基于密码学中的椭圆曲线数字签名算法来实现去中心化的 P2P 系统设计[1−3]. 但区块链的作用不仅仅局限于比特币. 现在人们在使用区块链这个词时, 有时是指数据结构, 有时是指数据库, 有时则是指数据库技术. 从数据的角度来看, 区块链是一种分布式数据库 (或称为分布式共享总账[4], Distributed shared ledger), 这里的 “分布式” 不仅体现为数据的分布式存储, 也体现为数据的分布式记录 (即由系统参与者集体维护); 从记录效果的角度来看, 区块链可以生成一套记录时间先后、不可篡改、可信任的数据库, 这套数据库是去中心化存储且数据安全能够得到有效保证. 具体地说, 区块链技术就是一种大家共同参与记录信息和存储信息的技术. 过去, 人们将数据记录和存储的工作交给中心化的机构来完成, 而区块链技术则让系统中的每一个人都可以参与数据的记录和存储. 区块链

收稿日期 2016-09-15 录用日期 2017-02-21
Manuscript received September 15, 2016; accepted February 21, 2017
国家自然科学基金 (61272007, 61503342, 61672467), 国家自然科学基金重点项目 (71731004), 国家杰出青年基金 (61425019), 浙江省自然科学基金 (LY16F030002) 资助
Supported by National Natural Science Foundation of China (61272007, 61503342, 61672467), Key Projects of National Natural Science Foundation of China (71731004), National Science Foundation for Distinguished Young Scholar of China (61425019), and Natural Science Foundation of Zhejiang Province (LY16F030002)
本文责任编委 袁勇
Recommended by Associate Editor YUAN Yong
1. 浙江师范大学数理与信息工程学院 金华 321004 2. 复旦大学电子工程系自适应网络与控制实验室 上海 200433 3. 复旦大学智慧网络与系统研究中心 上海 200433
1. College of Mathematics, Physics and Information Engineering, Zhejiang Normal University, Jinhua 321004 2. Adaptive Networks and Control Laboratory, Department of Electronic Engineering, Fudan University, Shanghai 200433 3. Research Center of Smart Networks and Systems, Fudan University, Shanghai 200433

技术在没有中央控制点的分布式对等网络下, 使用分布式集体运作的方法, 构建了一个 P2P 的自组织网络[3]. 通过复杂的校验机制, 区块链数据库能够保持完整性、连续性和一致性, 即使部分参与人作假也无法改变区块链的完整性, 更无法篡改区块链中的数据. 区块链技术涉及的关键点包括: 去中心化 (Decentralized)、去信任 (Trustless)、集体维护 (Collectively maintain)、可靠数据库 (Reliable data base)、时间戳 (Time stamp)、非对称加密 (Asymmetric cryptography) 等[1].

区块链技术原理的来源可归纳为数学上的拜占庭将军问题[5−6]. 将拜占庭将军问题延伸到互联网生活中来, 其内涵可概括为: 在互联网大背景下, 当需要与不熟悉的对手进行价值交换活动时, 人们如何才能防止不会被其中的恶意破坏者欺骗和迷惑, 从而做出错误的决策. 而如果进一步将拜占庭将军问题延伸到技术领域中来, 其内涵可概括为: 在缺少可信任的中央节点和可信任通道的情况下, 分布在网络中的各个节点应如何达成共识. 从这些角度来看, 区块链技术解决了闻名已久的拜占庭将军问题, 它提供了一种无需信任单个节点, 还能创建共识网络的方法.

作为区块链技术最成功的应用, 比特币系统应用工作量证明 (Proof of work, PoW)[7] 的共识机制实现交易的不可篡改性和不可伪造性[8]. PoW 共识机制的核心思想是通过引入分布式节点的算力竞争来保证数据的一致性和共识的安全性. 比特币系统中, 各节点基于各自的算力相互竞争, 共同解决一个求解复杂但验证容易的 SHA256 数学难题, 最快解决该难题的节点将获得区块记账权和系统自动生成的比特币奖励. 具体过程如下: 如果想产生一个区块并写入到区块链中, 需要找到一个小于系统规定难度值的随机数, 这样才可能被其他节点认可, 并写入到区块链中. 而找到随机数需要输出密码散列函数家族 SHA256 的哈希算法. 其中, 一个符合要求的输出值由 N 个前导零构成. 零的个数取决于网络的难度值, 挖矿难度越高, 零的个数会越多. 当输出值不满足要求时, 这个随机数就会增加一个单位, 直到找到为止. 找到合适随机数后, 节点获得记账权和相应比特币奖励, 并将该过程中产生的所有交易记录在区块上, 所有区块按时间顺序连接则构成区块链. 一般地, 比特币系统通过灵活调整随机数搜索的难度值来控制区块的平均生成时间.

在比特币系统中, 产生区块的过程称为挖矿, 进行挖矿的参与者称为矿工[9]. 由于比特币系统大约每 10 分钟产生一个区块, 这意味着大部分矿工在一定时间内很难产生区块. 为了增加获得稳定收益的可能性, 矿工会选择加入开放矿池进行合作挖矿. 具体地, 矿池中的矿工需要耗费资源尝试产生区块, 即发送完整工作量证明给管理者. 但完整工作量很难产生, 矿工也可以选择发送部分工作量证明获得相应收益. 无论哪个矿工产生区块, 获得的收益将按贡献比例分配给每个矿工. 参与者注册为矿工很简单, 只需要提供一个公共的网络接口就可以加入开放矿池, 因此开放矿池很容易受到攻击. 有些注册矿工只发送部分工作量证明, 当产生完整工作量证明时就会将其抛弃, 这种攻击方式被称为区块截留攻击[10−11]. 在这种情形下, 攻击者发送部分工作量证明, 但不会对矿池产生有效收益, 这也导致攻击者与其他矿工共同分享矿池收益, 从而减少其矿池的收益[12−13].

研究表明, 在一个开放的矿池中, 矿工可以通过攻击其他矿工增加自己的收益. 如果所有矿工都选择攻击对方, 那么他们获得的收益将少于他们互不攻击时获得的收益. 这就是 PoW 共识算法中的挖矿困境, 而这种困境也对应到博弈论中经典的囚徒困境 (Prisoner′s dilemma)[14−16], 即攻击对个体而言是最优策略, 但却不是系统最优的. 如何理解和分析挖矿过程中的博弈困境无疑给比特币的发展和技术开发乃至投入使用提供了理论基础. 例如 Eyal 基于博弈理论, 定性地分析了挖矿过程中的困境[17], 但并没有给出纯策略存在条件以及相应证明. 本文在文献 [17] 的基础上进一步分析矿工博弈困境的纯策略和混合策略均衡, 并给出两种均衡存在的条件.

更为重要的是, PoW 共识机制存在着显著的缺陷, 其强大算力造成的资源浪费 (例如算力) 历来为研究者所诟病, 而且长达 10 分钟的交易确认时间使其相对不适合小额交易的商业应用. 与此同时, 随着区块链技术的发展和各种数字币的相继涌现, 研究者提出多种不依赖算力而能够达成共识的机制, 例如权益证明 (Proof of stake, PoS) 共识[18], 授权股份证明机制 (Delegated proof of stake, DPoS) 共识[19], 缠结 (Tangle)[20] 以及 Tendermint[21] 机制等. 而最理想的共识算法是系统中的节点达成的共识是一个纳什均衡[21], 即单方面改变自己的策略都不会提高自身的收益. 这为基于博弈论构建共识机制提供了新的思路. 另一方面, PoW 共识过程中的挖矿困境对应经典的囚徒困境模型, 其纳什均衡为互相攻击, 此时的系统收益并不能达到最优. 为提高系统的整体效益, 有必要建立相关机制, 使矿工趋向于合作, 以获得较高的系统收益, 从而为实现高效的共识算法提供依据.

零行列式 (Zero determinant, ZD) 策略是近几年在博弈论中兴起的一种新方法, 它能够打破传统的纳什均衡理论. 如 Press 和 Dyson[22] 用 ZD 策略优化囚徒困境模型, 一方面可以解决系统收益低问

题, 另一方面, 无论对手采取何种策略, 都可以强迫对手与自己收益之间满足线性关系[23−24]. 此外, ZD 策略被应用到无线网络中的资源管理[25] 和频谱共享[26] 等问题. 这些都为本文运用 ZD 策略对矿工的策略选择进行优化提供了参考和借鉴.

本文组织结构为: 第 1 节介绍区块截留攻击和博弈理论中的纳什均衡及囚徒困境模型; 第 2 节利用博弈均衡理论对矿工算力相同和不相同两种情形的挖矿困境进行分析, 给出纯策略均衡以及混合策略均衡存在条件; 第 3 节运用 ZD 策略对区块截留攻击博弈进行优化, 得到获得较高系统收益时矿工策略选择的优化条件; 第 4 节给出数值仿真; 第 5 节总结本文内容, 并对今后工作进行展望.

1 预备知识

1.1 区块截留攻击

区块截留攻击是指在一个开放矿池中矿工与矿工之间的攻击. 攻击者只发送部分工作量证明给矿池管理员, 当发现完整工作量证明时就将其抛弃. 而工作量证明只能被任务的创建者使用, 攻击者不能将区块截留攻击的算力用于其他用途, 也不能从这部分算力中获得任何其他的收益. 因此, 这种攻击一方面会造成算力浪费, 另一方面也会使整个矿池收益降低. 此外, 少量部分工作量证明不会在很大程度上影响矿池的有效算力和有效收益, 但矿工进行攻击后, 整个矿池的有效算力和有效收益将低于所有矿工正常挖矿时所获得的收益. 虽然矿池管理员检测到整个矿池的总收益降低, 发现正在遭受区块截留攻击, 但并不能判断哪个矿工正在发起攻击.

除了区块截留攻击, 还有其他几种类型的攻击, 例如矿池间的区块截留攻击、自私挖矿攻击、劫持攻击以及服务拒绝攻击等.

1.2 博弈理论

博弈论被誉为 “社会科学中的数学”, 是研究具有斗争或竞争条件下最优决策问题的数学理论和方法. 更确切地说, 是指在双方相互竞争对立的环境条件下, 参与者依靠所掌握的信息, 遵循一定的规则约束, 各自选择策略并取得相应的结果 (或收益) 的过程.

1.2.1 纳什均衡

通常认为 Neumann 与 Morgenstern 在 1944 年合著的《博弈论与经济行为》标志着现代博弈理论的初步形成. 由此延续, 在上世纪 50 年代, 纳什提出了纳什均衡 (Nash equilibrium, NE) 理论, 刻画了所有博弈者策略构成的一种最优情势 (Profile), 即任何博弈者单方面试图改变自己的策略, 则在该情势下该博弈者的收益将受到损害 (至少不会改善). 换言之, 这种情势下所有博弈者的策略都是所有其他对手策略的最优反应 (Best response)[27].

考虑 n 人博弈模型, 其策略空间为 $S=\prod S_i$, S_i 是策略 x_i 的集合. 定义 $X=(x_1,x_2,\cdots,x_n)$, $x_i\in S_i$, $x_{-i}=(x_1,x_2,\cdots,x_{i-1},x_{i+1},\cdots,x_n)$. 设 $U_i(X)=U_i(x_i,x_{-i})$ 是个体 i 采取策略 x_i 时的收益函数. n 个参与者在相互作用过程中可以达到纳什均衡 $X^*=(x_1^*,x_2^*,\cdots,x_n^*)$. 这种均衡是指没有个体可以通过单方面改变自己的策略而增加收益, 即

$$\forall i,\ U_i(x_i^*,x_{-1}^*)\geq U_i(x_i',x_{-i}^*),\ \ x_i^*\in X^*;\ x_i'\notin X^* \tag{1}$$

如果式 (1) 对每个 $x_i'\notin x_i^*$ 都严格成立, 称该均衡为严格纳什均衡. 如果 x_i^* 是一个纯策略, 称这个均衡为纯策略纳什均衡; 否则, 称这个均衡为混合纳什均衡.

1.2.2 囚徒困境博弈模型

考虑两个策略 A 和 B 的博弈模型, 当两个个体都采用 A 时, 各自获得奖励 R; 当两者都采用 B 时, 各自获得惩罚 P; 当 A 策略个体遇到 B 策略个体时, 前者收益为 S, 后者收益为 T. 其收益矩阵表示如下:

$$\begin{array}{c} \\ A \\ B \end{array}\begin{array}{c} \begin{array}{cc} A & B \end{array} \\ \begin{pmatrix} R & S \\ T & P \end{pmatrix} \end{array} \tag{2}$$

其中, R 表示对 “双方合作的奖励” (Reward for mutual cooperation), P 表示对 “双方背叛的惩罚” (Punishment for mutual defection), 当一方合作而另一方背叛时, S 表示合作者获得 “傻瓜的报酬” (Sucker's payoff), T 表示背叛者获得 “背叛的诱惑” (Temptation to defect).

囚徒困境模型是博弈论中最为经典的博弈模型. 其背景如下: 两个嫌疑犯 $P1$ 和 $P2$ 作案后被捕, 接受隔离审讯; 如果两人都坦白则各判 8 年, 如果一人坦白另一人不坦白, 坦白一方获释, 另一方判 10 年; 如果两人同时抗拒则因证据不足各判 1 年. 对于 $P1$ 和 $P2$ 来说, 坦白意味着背叛 (Defect), 不坦白意味着合作 (Cooperate), 用收益矩阵表示如式 (2), 满足条件 $T>R>P>S$, 且 $2R>T+S$. 对于行参与者而言, 如果列参与者选择合作, 则他的最优选择为背叛; 如果列参与者选择背叛, 背叛对手的收益高于合作时获得的收益. 因此, 无论对手采用何种策略, 选择背叛策略都是最优的. 理性个体最终会处于互相背叛状态, 即 (D,D) 是囚徒困境博弈的纳什均衡状态. 但是, 根据收益间的数值关系可知, 该状态

的收益将低于两者选择合作时的收益, 理性参与者将面临选择合作或背叛的困境.

参与者进行一次交互, 会面临囚徒困境, 选择纳什均衡点 — 背叛. 当进行多次博弈时, 就构成迭代的囚徒困境 (Iterated prison′s dilemma, IPD), 这时, 参与者最优策略依赖于对手策略选择, 从而改变原先的均衡状态, 以达到系统较好的 “均衡”. 两个最经典的迭代策略是 “针锋相对” (Tit-for-tat, TFT) 和 “赢存输变” (Win-stay, lose-shift, WSFS). TFT 策略: 首先参与者不会背叛对方, 如果对手选择背叛, 在下一次博弈中他将选择背叛来惩罚对手; 如果下一次对手选择合作, 他将会和对手再次合作. WSFS 策略: 对收益设一个阈值, 当第一轮收益高于该值时, 在下一轮博弈时, 参与者将继续保持上一轮采取的策略; 如果上一轮所得收益低于该阈值, 则在下一轮博弈时, 参与者将采用与上一轮相反的策略. 若 1 表示合作, 0 表示背叛, TFT 策略可表示为 $[1,0,1,0]$, WSFS 策略表示为 $[1,0,0,1]$.

两人两策略的博弈除了囚徒困境, 当收益矩阵 (2) 中的参数 R, S, T, P 满足不同条件时, 还有其他博弈类型. 例如雪堆博弈 (Snowdrift game, SG)、鹰鸽博弈 (Hawk-dove game, HDG)、胆小鬼博弈 (Chicken game, CG) 等.

2　挖矿困境的博弈均衡分析

矿工正常挖矿, 会获得与其付出算力成正比的收益, 付出的算力会耗费大量的电力、人力、物力等资源; 矿工也可以通过只发送部分工作量证明进行区块截留攻击, 获得高于实际应获得的收益. 攻击是最优策略, 但当所有矿工都选择这种策略时, 整个矿池的有效收益几乎为零, 所以矿工最终获得的收益少于不攻击时获得的收益. 因此对于矿工而言, 攻击与否是一个困境.

2.1　相同算力的情形

在开放矿池中, 忠实矿工进行正常挖矿, 会耗费一定的算力, 假设耗费的资源为 c $(0 < c < 0)$. 矿工之间合作挖矿, 在一定程度上会增加挖到区块的概率, 并且各个矿工的期望收益也将大于单独进行挖矿时获得的收益. 假设矿工合作挖矿时收益会扩大 r $(r > 1)$ 倍, 系统将扩大后的收益再按算力进行公平分配. 为了简单化模型, 假设每个矿工的算力相同, 扩大后的收益会进行平分. 当矿工不忠实挖矿, 对该矿池进行区块截留攻击时, 矿池管理员仍会按其贡献算力分配收益, 这样的行为不仅使该矿池总收益减少, 所有矿工的收益也将降低, 这就是挖矿过程中的 “搭便车者” (Free rider). 这里只考虑矿池中有两个矿工的情况, 每个矿工有两种策略选择: 合作 (Cooperation, C) 和攻击 (Attack, A). 假设每个矿工正常挖矿时的收益为 1, 他们的收益情况如下所示:

$$\begin{array}{c} \\ \mathrm{C} \\ \mathrm{A} \end{array}\begin{array}{c} \begin{array}{cc} \mathrm{C} & \mathrm{A} \end{array} \\ \begin{pmatrix} r(1-c), & r(1-c) & \frac{1}{2}-c, & \frac{1}{2} \\ \frac{1}{2}, & \frac{1}{2}-c & 0, & 0 \end{pmatrix} \end{array}$$

当策略选择为 (C, C) 时, 两个矿工合作挖矿会扩大 r 倍收益, 也会耗费一定的资源 c, 因此对应的收益为 $(r(1-c), r(1-c))$. 当一个矿工选择合作, 另一个选择攻击时, 即 (C, A) 或 (A, C), 对于合作者而言, 只有一个矿工进行挖矿, 收益不会被扩大 r 倍, 矿池收益为 1, 忠实矿工和攻击者获得相同的收益 1/2, 但忠实矿工正常挖矿会耗费其资源 c, 所以其最终收益为 $1/2-c$, 攻击者收益为 1/2. 当两个矿工都选择攻击 (A, A) 时, 整个矿池的有效收益为 0, 则这两个矿工的收益为 0.

2.1.1　纯策略纳什均衡

当 $r(1-c) > 1/2$ 且 $c > 1/2$, 一个矿工选择 C 策略时, 另一个矿工选择 C 策略比选择 A 策略获得收益大, 为了使自己的收益达到最大, 另一个矿工也会选择 C 策略; 当一个矿工选择 A 策略时, 另一个矿工选择 C 策略比选择 A 策略损失更多, 为了使自己的收益不损失太多, 另一个矿工只能选择 A 策略. 因此, 在这种情况下, 矿工的纳什均衡点为 (C, C), (A, A). 在图 1 中, 区域 (c) 满足该条件.

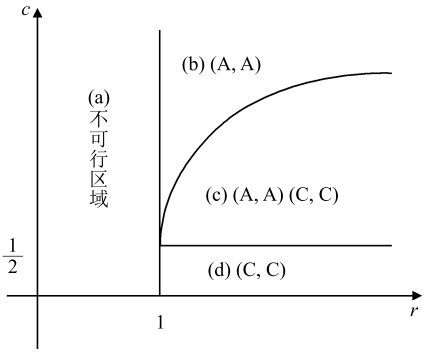

图 1　矿工纯策略纳什均衡分布图

Fig. 1　The distribution of pure strategies Nash equilibrium for miners

当 $r(1-c) < 1/2$ 且 $c > 1/2$, 一个矿工选择合作时, 另一个矿工会选择攻击来增加自己的收益; 当一个矿工选择攻击时, 为了不降低自己的收益, 另一个矿工只能选择攻击. 也就是说, 无论对手选择何种策略, 矿工的最优策略是攻击, 这时矿工将面临矿工式的囚徒困境. 此时的纳什均衡点为 (A, A), 在图 1 中, 表现为区域 (b).

当 $r(1-c)>1/2$ 且 $c<1/2$, 一个矿工选择合作时, 为了使自己的收益达到最高, 另一个矿工会选择合作; 当一个矿工选择攻击时, 为了使自己的收益不会损失很多, 另一个矿工仍旧会选择合作. 即无论对手矿工选择何种策略, 该矿工选择合作对其自身利益是最大的. 因此, 这时矿工的纳什均衡点为 (C, C), 这种情况在图 1 中为区域 (d).

当 $r(1-c)<1/2$ 且 $c<1/2$ 时, 有 $r<1$, 而 r 为收益扩大的倍数, 这与模型假设 $r>1$ 矛盾, 因此这种情况为图 1 中的区域 (a).

2.1.2 混合策略纳什均衡

矿工博弈有一个唯一的混合均衡点. 设矿工 1 选择合作的概率为 x $(0\le x\le 1)$, 矿工 2 选择合作的概率为 y $(0\le y\le 1)$, 则矿工 1 合作的期望收益为 $y/2$, 矿工 1 攻击的期望收益为 $yr(1-c)+(1-y)(1/2-c)=ry-y/2-ryc+1/2-c+cy$. 在混合策略均衡点下, 矿工 1 合作与攻击的收益相等, 即

$$ry-\frac{y}{2}-ryc+\frac{1}{2}-c+cy=\frac{y}{2}$$

化简后有

$$y=\frac{c-\dfrac{1}{2}}{(1-c)(r-1)} \tag{3}$$

同样地, 矿工 2 选择合作时的期望收益为 $xr(1-c)+(1-x)(1/2-c)=rx-x/2-rxc+1/2-c+cx$. 矿工 2 选择攻击时的期望收益为 $x/2$. 在混合策略均衡下, 矿工 2 选择合作与攻击对应的收益也应是相等的, 即

$$rx-\frac{x}{2}-rxc+\frac{1}{2}-c+cx=\frac{x}{2}$$

化简后有

$$x=\frac{c-\dfrac{1}{2}}{(1-c)(r-1)} \tag{4}$$

根据式 (3) 和式 (4) 可以得到混合均衡存在的条件

$$\frac{1}{2}\le c\le 1-\frac{1}{2r} \tag{5}$$

定理 1. 设算力相同的两个矿工合作的概率分别为 x, y, 则 x, y 与挖矿资源耗费 c 成正比, 与合作后收益扩大倍数 r 成反比, 且混合策略均衡存在的条件为

$$\frac{1}{2}\le c\le 1-\frac{1}{2r}$$

注 1. 图 1 中, 区域 (b) 和区域 (d) 只存在纯策略纳什均衡, 不满足不等式 (5) 的条件约束. 只有区域 (c) 满足不等式 (5) 的条件约束, 即存在混合纳什均衡.

2.2 不同算力的情形

本节分析矿工算力不同时的系统分配收益情况. 假设在开放矿池中有两个矿工挖矿, 整个矿池的总算力为 1, 矿工 1 的算力为 t $(0<t<1)$, 则另一个矿工的算力为 $(1-t)$, 与上一个模型相同的是, c 表示各种资源支出, r 表示矿工合作后的收益扩大的倍数; 不同的是, 假设矿工挖矿后, 矿池的收益根据矿工挖矿情况而定, 资源支出 c 需满足 $c<\min(t, 1-t)$. 不失一般性, 设 $t<1-t$, 这里仍然存在 "搭便车" 者: 通过区块截留攻击, 增加自己的收益. 矿工的策略选择仍然为 C 和 A, 他们的收益情况如下:

$$\begin{array}{c} \\ \mathrm{C} \\ \mathrm{A} \end{array}\begin{array}{c} \begin{array}{cc} \mathrm{C} & \mathrm{A} \end{array} \\ \begin{pmatrix} r(t-c), & r(1-t-c) \\ t(1-t), & (1-t)^2-c \end{pmatrix}\begin{array}{cc} t^2-c, & t(1-t) \\ 0, & 0 \end{array}\end{array}\Bigg)$$

当两个矿工策略选择为 (C, C) 时, 合作挖矿时整个矿池收益为 1, 矿工获得收益为自己的算力, 这样会耗费 c, 但收益会扩大 r 倍, 因此这时的收益为 $(r(t-c), r(1-t-c))$, 当两个矿工策略选择为 (C, A) 时, 矿工 1 选择合作, 矿工 2 选择攻击, 这时整个矿池的收益为 t, 矿工 1 获得收益为它所占整个矿池算力的比例与耗费资源之差, 即 (t^2-c). 同样的, 矿工 2 的收益为 $t(1-t)$; 当两个矿工都选择攻击时, 系统有效收益为 0, 此时每个矿工的收益也为 0.

2.2.1 纯策略纳什均衡

1) 当 $r(t-c)>(1-t)t$, $(1-t)^2-c<0$, 矿工 1 选择合作时, 矿工 2 为了获得更多的收益会选择合作: 矿工 1 选择攻击时, 矿工 2 为了使自己的收益不会降低的太多也会选择攻击. 因此这时矿工的纳什均衡点有两个: (C, C) 和 (A, A).

2) 当 $r(t-c)<(1-t)t$, $(1-t)^2-c>0$, 矿工 1 选择攻击时, 矿工 2 为了有收益必须选择合作 (当选择攻击时, 收益将为 0), 此时的纳什均衡点为 (A, C).

3) 当 $r(t-c)>(1-t)t$, $(1-t)^2-c>0$ 或者 $t^2-c>0$ 时, 矿工 1 选择攻击, 另一个矿工选择合作时获得收益将高于选择攻击时获得的收益: 而当矿工 1 选择合作时, 另一个矿工选择合作时的收益同样高于选择攻击时获得的. 因此, 矿工的纳什均衡点为 (C, C).

4) 当 $r(t-c)<(1-t)t$, $(1-t)^2-c<0$ 时, 对手选择攻击, 该矿工选择合作时, 会获得较少收益,

当选择攻击时, 会获得相对多的收益. 此时矿工的纳什均衡点为 (A, A), 这就是矿工版的囚徒困境.

2.2.2　混合策略纳什均衡

矿工以某种概率进行策略选择, 当概率为 0 或 1 时, 存在纯策略纳什均衡, 当概率不为 0 或 1 时, 为混合策略均衡问题. 假设矿工 1 合作的概率为 x ($0 \le x \le 1$), 矿工 2 合作的概率为 y ($0 \le y \le 1$), 根据均衡的性质可以知, 矿工 1 选择是否合作, 收益应该是相等的, 即

$$yr(t-c)+(1-y)(t^2-c)=yt(1-t)$$

化简后为

$$y=\frac{c-t^2}{(r-1)(t-c)} \tag{6}$$

同样地, 均衡条件下, 矿工 2 选择攻击或合作的收益应该是相等的. 即

$$xr(1-t-c)+(1-x)[(1-t)^2-c]=xt(1-t)$$

化简后为

$$x=\frac{c-(1-t)^2}{(r-1)(1-t-c)} \tag{7}$$

根据式 (6) 和式 (7), 可以得到在矿工算力不相同时混合均衡存在的条件

$$c>(1-t)^2,\ \ r\ge 1+\frac{c-t^2}{t-c} \tag{8}$$

定理 2. 设算力不相同的两个矿工合作的概率分别为 x, y, 则 x, y 与资源耗费 c 成正比, 与收益扩大倍数 r 成反比, 矿工的合作概率与其本身算力成反比, 且混合策略均衡存在的条件为

$$c>(1-t)^2,\ \ r\ge 1+\frac{c-t^2}{t-c}$$

注 2. 根据 x, y 存在条件, 可以得出, 只有第 1 种和第 4 种情形的纯策略纳什均衡存在混合均衡. 由此发现, 矿工算力相同是算力不同情况下的特殊情况.

3　挖矿困境的博弈优化

由第 2 节可知, 区域 (c) 的纳什均衡为 (A, A), (C, C). 当两个矿工相互攻击时, 将获得较低的系统收益, 甚至为零. 为了提高矿池收益, 我们将 ZD 策略应用到该挖矿困境中, 对系统收益进行优化, 最终得到较高的收益.

具体模型如下: 假设在一个开放矿池中, 有两类矿工, 一类忠实矿工 (LM, 正常挖矿, 维护整个矿池的利益), 另一类自私矿工 (SM, 自私挖矿, 只考虑自己的收益). 对于 (C, C)、(C, A)、(A, C)、(A, A) 四种情况, LM 和 SM 的混合策略概率分别为 $lm=[a_1, a_2, a_3, a_4]$ 和 $sm=[b_1, b_2, b_3, b_4]$, lm, sm 分别是下一状态下选择合作的转移概率向量. 例如, 上一状态两个矿工都选择 C 时, 下一状态 LM 选择 C 的概率为 a_1, 选择 A 的概率为 $(1-a_1)$; SM 选择 C 的概率为 b_1, 选择 A 的概率为 $(1-b_1)$. LM 的收益向量为

$$\boldsymbol{W}^L=\left[R^L,S^L,T^L,P^L\right]^{\mathrm{T}}=\left[r(1-c),\frac{1}{2}-c,\frac{1}{2},0\right]^{\mathrm{T}}$$

SM 的收益向量为

$$\boldsymbol{W}^S=\left[R^S,T^S,S^S,P^S\right]^{\mathrm{T}}=\left[r(1-c),\frac{1}{2},\frac{1}{2}-c,0\right]^{\mathrm{T}}$$

因此, 两个矿工的策略选择转移情况可由马尔科夫状态转移矩阵 M 来表示. 其中, 马尔科夫转移矩阵的稳态向量 $s=[s_1,s_2,s_3,s_4]^{\mathrm{T}}$, 且 $s_1+s_2+s_3+s_4=1$.

$$M=\begin{bmatrix} a_1b_1 & a_1(1-b_1) & (1-a_1)b_1 & (1-a_1)(1-b_1)\\ a_2b_3 & a_2(1-b_3) & (1-a_2)b_3 & (1-a_2)(1-b_3)\\ a_3b_2 & a_3(1-b_2) & (1-a_3)b_2 & (1-a_3)(1-b_2)\\ a_4b_4 & a_4(1-b_4) & (1-a_4)b_4 & (1-a_4)(1-b_4)\end{bmatrix}$$

$$\boldsymbol{s}^{\mathrm{T}}M=\boldsymbol{s}^{\mathrm{T}} \tag{9}$$

则 LM 和 SM 的稳态期望收益分别为

$$U^L=\boldsymbol{s}^{\mathrm{T}}W^{\mathrm{T}}$$
$$U^S=\boldsymbol{s}^{\mathrm{T}}W^S$$

定义 $M'=M-I$, 这里 I 是单位矩阵. 因此式 (9) 等价为

$$\boldsymbol{s}^{\mathrm{T}}M'=0$$

根据克拉默法则, 可以得到

$$\mathrm{Adj}(M')M'=\det(M')I$$

其中, $\mathrm{Adj}(M')$ 为 M' 的伴随矩阵, 它的每一行都与 $\boldsymbol{s}$ 成比例. 选取 $\mathrm{Adj}(M')$ 的最后一行, 经过行列变换可以得到, 对于任意向量 $\boldsymbol{v}=[v_1,v_2,v_3,v_4]^{\mathrm{T}}$, 与 $\boldsymbol{s}$ 的点积是一个行列式

$$\boldsymbol{s}^{\mathrm{T}}\boldsymbol{v}\equiv D(lm,sm,\boldsymbol{v})=$$

$$\det\begin{bmatrix} -1+a_1b_1 & -1+a_1 & -1+b_1 & v_1 \\ a_2b_3 & -1+a_2 & b_3 & v_2 \\ a_3b_2 & a_3 & -1+b_2 & v_3 \\ a_4b_4 & a_4 & b_4 & v_4 \end{bmatrix}$$

假设 $\boldsymbol{v}=\alpha\boldsymbol{W}^L+\beta W^S-\gamma I$, 这里 α, β 是不为零的系统参数, 令 $lm'=[-1+a_1,-1+a_2,a_3,a_4]$, 如果 LM 的混合策略 $lm'=\phi\boldsymbol{v}=\phi(\alpha\boldsymbol{W}^L+\beta\boldsymbol{W}^S-\gamma I)$, 其中 ϕ 是不为零的次数, 则

$$\alpha U^L+\beta U^S-\gamma=0 \tag{10}$$

定理 3. 在一个开放矿池中, 当 LM 采取 ZD 策略时, LM 可以控制 SM 的收益与自己的收益保持线性关系: $\gamma=\alpha U^L+\beta U^S$.

证明. 通过式 (10) 可知, 无论 SM 采取哪种策略, LM 都可以通过 ZD 策略使 SM 的收益与其自己收益保持线性关系. 并且, 由于 LM 考虑系统的整体收益, 提高自身收益, 也将提高对手收益. □

引理 1. LM 采取 ZD 策略时, 系统参数 α, β 应满足

$$-1<\frac{\alpha}{\beta}<0 \tag{11}$$

γ 满足

$$\gamma\le\min(\alpha R^L+\beta R^S,\alpha S^L+\beta T^S)$$
$$\gamma\ge\max(\alpha T^L+\beta S^S,\alpha P^L+\beta P^S)$$

或者

$$\gamma\ge\max(\alpha R^L+\beta R^S,\alpha S^L+\beta T^S)$$
$$\gamma\le\min(\alpha T^L+\beta T^S,\alpha P^L+\beta P^S)$$

证明. 根据定理 3 可知, $\alpha/\beta<0$, 由式 (10) 可得, LM 的策略 lm 应满足

$$\begin{bmatrix} a_1-1 \\ a_2-1 \\ a_3 \\ a_4 \end{bmatrix}=\phi\begin{bmatrix} \alpha R^L+\beta R^S-\gamma \\ \alpha S^L+\beta T^S-\gamma \\ \alpha T^L+\beta T^S-\gamma \\ \alpha P^L+\beta P^S-\gamma] \end{bmatrix} \tag{12}$$

混合策略概率 a_i $(i=1,2,3,4)$ 应该满足 $0\le a_i\le 1$, 有

$$\begin{aligned} -1\le\phi(\alpha R^L+\beta R^S-\gamma)\le 0 \\ -1\le\phi(\alpha S^L+\beta T^S-\gamma)\le 0 \\ 0\le\phi(\alpha T^L+\beta S^S-\gamma)\le 1 \\ 0\le\phi(\alpha P^L+\beta P^S-\gamma)\le 1 \end{aligned} \tag{13}$$

根据不等式组 (13), 当 $\phi<0$ 时, γ 需满足

$$\gamma\le\min(\alpha R^L+\beta R^S,\alpha S^L+\beta T^S)$$
$$\gamma\ge\max(\alpha T^L+\beta S^S,\alpha P^L+\beta P^S) \tag{14}$$

在挖矿过程中, 很明显 $S^L=S^S<T^S=T^L$, 对于这种情况不失一般性, 假设 $\alpha>0$, $\beta<0$, 则

$$\alpha S^L+\beta T^S\le\alpha T^L+\beta S^S \tag{15}$$

很明显, 式 (15) 与式 (14) 矛盾.

根据式 (13), 当 $\phi>0$ 时, γ 需满足:

$$\gamma\ge\max(\alpha R^L+\beta R^S,\ \alpha S^L+\beta T^S)$$
$$\gamma\le\min(\alpha T^L+\beta S^S,\ \alpha P^L+\beta P^S) \tag{16}$$

由式 (16) 可以得到系数 α 和 β 应满足:

$$\alpha R^L+\beta R^S\le\alpha P^L+\beta P^S \tag{17}$$

即

$$\frac{\alpha}{\beta}>\frac{P^S-R^S}{R^L-P^L}=-1 \tag{18}$$

当 $\alpha<0$, $\beta<0$ 时, 可以得到条件与式 (16) 矛盾, 通过式 (14) 仍然可以得到 $\alpha/\beta>-1$.

综上分析, 当 LM 采取 ZD 策略时, 即 $lm'=\phi\boldsymbol{v}=\phi(\alpha\boldsymbol{W}^L+\beta\boldsymbol{W}^S-\gamma I)$. 可以得到系统参数 α, β 应满足式 (11). □

定理 4. 在挖矿过程中, 当 LM 采取 ZD 策略, 满足 $\gamma=\alpha U^L+\beta U^S$ 和 $\alpha/\beta<0$ 时, 两个矿工的最终收益可以为图 2 上线段 AC 和 AF 上的任意一点, 并且 LM 可以用 ZD 策略单方面控制线段 AC. 在 A 点可以使该模型达到一个新的纳什均衡 — 这个均衡具有较高的系统收益, 即: $U^L(lm^*,sm^*)+U^S(lm^*,sm^*)$, 并且 LM 和 SM 不会偏离这个均衡, 即:

$$\begin{aligned} U^L(lm^*,sm^*)+U^S(lm^*,sm^*)\ge \\ U^L(lm,sm^*)+U^S(lm,sm^*) \\ U^L(lm^*,sm^*)+U^S(lm^*,sm^*)\ge \\ U^L(lm^*,sm)+U^S(lm^*,sm) \end{aligned}$$

LM 的 ZD 策略 lm^* 为

$$\begin{cases} a_1=1 \\ a_2=1+\phi(\alpha(S^L-T^L)+\beta(T^S-R^S)) \\ a_3=\phi(\alpha(T^L-R^L)+\beta(S^S-R^S)) \\ a_4=\phi(\alpha(P^L-R^L)+\beta(P^S-R^S)) \end{cases} \tag{19}$$

证明. 下面对矿工算力不相同时, 对纯策略纳什均衡中的第 4 种情况进行证明, 即满足 $r(t-c)<$

$(1-t)t$, $(1-t)^2-c<0$ 条件, 纳什均衡为 (A, A). 则该情形下, 两个矿工的收益 $(U^{L'}, U^{S'})$ 分布为图 2 中四边形 $ABEC$, 不失一般性, 假设 $\alpha>0$, $\beta<0$. 当 $\gamma=\alpha U^{L'}+\beta U^{S'}$ 时, 则 LM 可以通过采取 ZD 策略控制线性关系 $\alpha(R^L-U^{L'})+\beta(R^S-U^{S'})=0$.

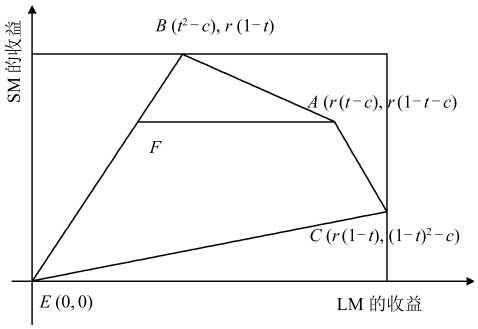

图 2　算力不同时矿工的收益分布情况

Fig. 2　The payoff distribution of miners with different powers

下面分析当矿工收益为四边形 $ABEC$ 各区域时, 矿工挖矿博弈的不同结果, 以及最优区域.

当 $U^{L'}<R^L$, $U^{S'}<R^S$ 时, 也就是图 2 中的三角形 ABF 区域, 有 $\gamma=\alpha U^{L'}+\beta U^{S'}>\alpha R^L+\beta R^S$, 这与式 (16) 矛盾. 因此, 当 $U^{L'}<R^L$, $U^{S'}<R^S$ 时, LM 不能采取 ZD 策略满足两个收益的线性关系式 $\alpha(R^L-U^{L'})+\beta(R^S-U^{S'})=0$. 也就是说当 LM 采取 ZD 策略时, 矿工挖矿博弈结果是不会优于矿工收益为四边形 $AFEC$ 时所得到的收益.

收益分布在线段 FA. 当 LM 的 ZD 策略 $lm'=\phi(\alpha\boldsymbol{W}^L+\beta\boldsymbol{W}^S-\gamma I)$ 中的 $\alpha=0$ 时, 可以满足 $\gamma=\beta U^S$, 并且满足式 (15). 这时, SM 的收益都在线段 FA 上. 即, 当 LM 采取 ZD 策略时, 不管 SM 采取何种策略, LM 都可以控制 SM 的收益为线段 FA 上的任意一点, 并且可以保持两个矿工收益的线性关系值不变.

同样的, 可以得到在线段 AC 上, LM 采取 ZD 策略时, 即可以保持两个矿工收益的线性关系. 而在线段 CE 上并不能保持该线性关系.

综上分析, 可以得到当两个矿工最终受益可以为 FA 和 AC 上的任意一点. 对于线段 AC, 不管 SM 采取何种策略, LM 都可以通过调整 α/β 的值单方面控制两个矿工的收益.

不失一般性, 假设 $\alpha>0$, $\beta<0$, 当 $\phi>0$ 时, 对于任意的 α 和 β, 只有满足式 (11), 并且 $\gamma=R^L+R^S$, 则 LM 采取的 ZD 策略 lm^* (式 (19)) 可以由式 (12) 得到, 即, 不管 SM 采取何种策略, LM 可以单方面控制两个矿工收益的线性关系, 并且系统收益可以达到较高值. □

注 3. 定理 4 从博弈论的角度对 PoW 共识算法中矿工的策略选择进行优化, 给出系统达到较高收益时的均衡条件. 即采用 ZD 策略方法使系统丢弃原来挖矿困境相互攻击的纳什均衡, 形成新的均衡, 从而提高系统的整体收益. 所以, 基于 ZD 策略的博弈论方法为设计高效的共识算法提供新的思路和研究方法.

4　数值仿真及结果分析

基于上述模型, 用数值仿真来说明 LM 用 ZD 策略 $(lm'=\phi(\alpha\boldsymbol{W}^L+\beta\boldsymbol{W}^S-\gamma I), \gamma=\alpha U^L+\beta U^S)$ 可以使 LM 和 SM 的收益保持线性关系, 并且维持系统收益稳定并达到一定的高度. 当矿工算力相同时, 假设系统参数资源耗费 $c=5/8$, 合作后矿工收益扩大倍数 $r=5/3$, $\phi=1/20$, $\alpha=8$, $\beta=-18$. LM 和 SM 的收益向量分别为 $\boldsymbol{W}^L=[5/8,-1/8,1/2,0]^{\mathrm{T}}$, $\boldsymbol{W}^S=[5/8,1/2,-1/8,0]^{\mathrm{T}}$.

下面是当 LM 和 SM 分别采取不同策略时, 对系统收益的分析. 其中, WSFS 策略的混合策略概率为 $[1,0,0,1]$, TFT 策略的混合策略概率为 $[1,0,1,0]$.

在图 3 中, 当 SM 采取 WSFS 策略时, LM 选择 ZD 或者 WSFS 策略时, 最终系统收益能达到最优值 1.25, 而 LM 采取 TFT 策略或一个随机策略 $[0.1,0.2,0.3,0.4]$ 时, 系统收益不能达到最优. 通过比较可以发现, 在 LM 选择的这 4 种策略中, WSFS 策略是最佳选择, ZD 策略经过迭代一定次数后也能达到系统最优, 而 TFT 策略和随机策略 $[0.1,0.2,0.3,0.4]$ 都只能获得较低的系统收益.

在图 4 中, SM 采用 TFT 策略, LM 分别采取 ZD、WSFS、TFT、$[0.1,0.2,0.3,0.4]$ 四种策略. 观察图 3, 通过对比发现这 4 种博弈都不能使系统达到最优, 而 ZD 策略虽然不能系统收益最优, 但相对于其他策略有明显的优势, 在迭代一定次数后可以使系统收益趋向于稳定, 并且可以得到高于其他策略的收益. LM 的最差策略为随机策略 $[0.1,0.2,0.3,0.4]$.

在图 5 中, SM 始终选择 $[0.1,0.2,0.3,0.4]$ 策略, LM 仍然采取 ZD、WSFS、TFT、$[0.1,0.2,0.3,0.4]$ 四种策略. 这 4 种博弈同样不能达到最优, 当 LM 采取 ZD 策略时, 经过一定次数的迭代博弈后, 系统收益逐渐稳定并且明显高于其他 3 个策略. 这时, SM 最不可取的策略为 TFT 策略.

根据图 6 可以得到, 不论 SM 选择何种策略, 只要 LM 采用 ZD 策略, 两个矿工收益的线性组 $\gamma=\alpha U^L+\beta U^S$, 随着迭代次数的增加, 会逐渐趋向稳定, 并且都为同一个值. 也就是说, 当矿工算力相同时, LM 可以用 ZD 策略控制 SM 的收益与自己收益保持线性关系, 不论 SM 采取何种策略.

当矿工算力不同时, 参数满足不同的条件, 也会

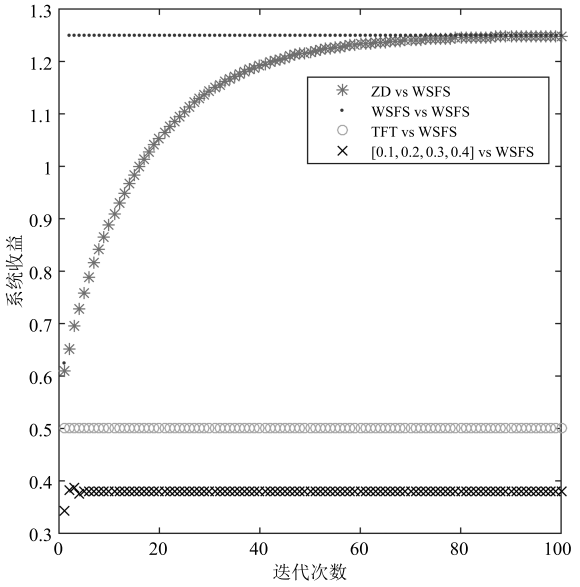

图 3 当矿工算力相同时, LM 采用不同策略, SM 始终采用 WSFS 策略后, 系统收益分布情况

Fig. 3 The distribution of systematic revenue, when LM takes different strategies and SM all uses WSFS strategy in the condition of same power

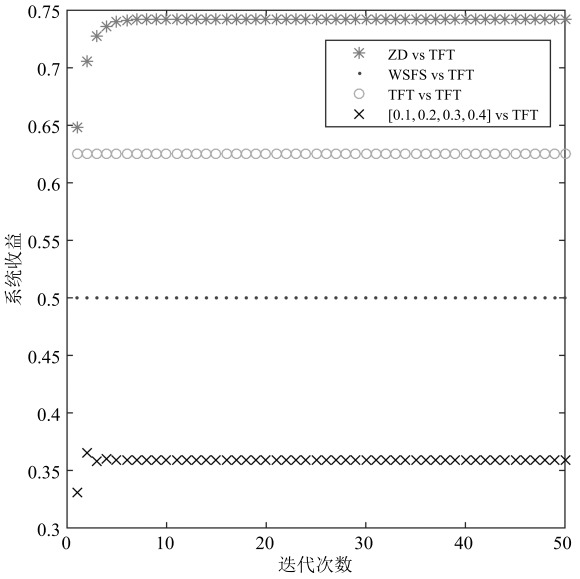

图 4 当矿工算力相同时, LM 采用不同策略, SM 始终采用 TFT 策略后, 系统收益分布情况

Fig. 4 The distribution of systematic revenue, when LM takes different strategies and SM all uses TFT strategy in the condition of same power

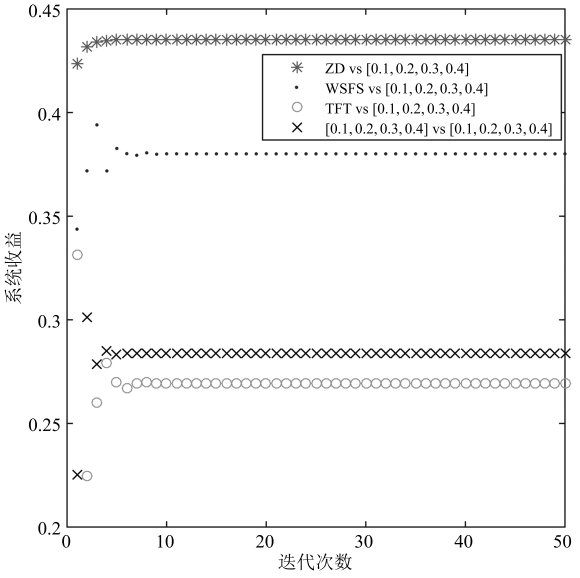

图 5 当矿工算力相同时, LM 采用不同策略, SM 始终采用 $[0.1, 0.2, 0.3, 0.4]$ 策略后, 系统收益分布情况

Fig. 5 The distribution of systematic revenue, when LM takes different strategies and SM all uses $[0.1, 0.2, 0.3, 0.4]$ strategy in the condition of same power

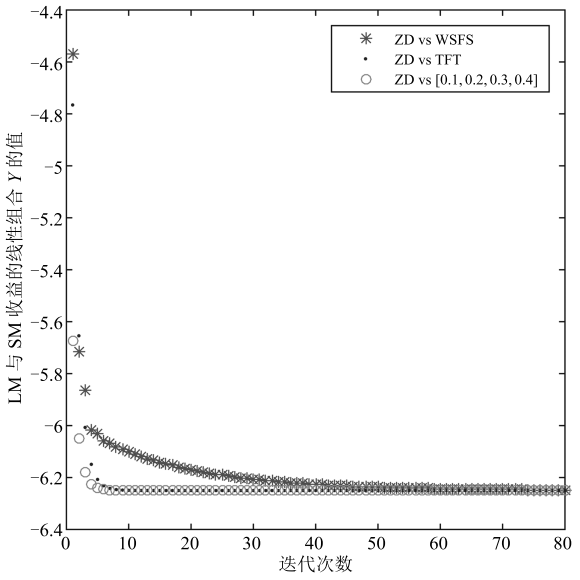

图 6 当矿工算力相同时, LM 采用 ZD 策略, SM 采用不同的三种策略, LM 和 SM 收益的线性关系

Fig. 6 The linear relationship between LM′ payoff and SM′ payoff, when LM takes a ZD strategies and SM uses three different strategy in the condition of same power

有不同的困境. 当系统条件满足纯策略中的第 4 种 (均衡点为双攻击) 情形时, 我们对该情形进行建模, 不妨假设参数 $\phi = 1/20$, $\alpha = 1$, $\beta = -10$, 算力耗费 $c = 0.3$, 收益扩大倍数 $r = 1.1$, LM 的算力 $t = 0.48$, 则 SM 的算力 $(1 - t) = 0.52$.

在图 7 中, SM 采用 WSFS 策略, LM 分别采用 ZD、WSFS、TFT、$[0.1, 0.2, 0.3, 0.4]$ 策略. LM 采用 WSFS 能立即使系统收益达到最优, 而采用 ZD 策略后, 在迭代一定次数后也能使系统收益得到满足. 另外两种策略只能使系统收益处于较低值. 在这 4 种策略中, 最差的为随机策略 $[0.1, 0.2, 0.3, 0.4]$.

图 8 与图 7 不同的是, 这里的 SM 采取的是

TFT 策略. 从图 8 可以发现, 虽然 ZD 策略不能使系统收益达到最优, 但它仍然是 LM 的最佳选择, 在重复一定次数后, 系统收益可以达到稳定, 并优于其他 3 个策略, 其他 3 个策略只能使系统收益更低. 这里的劣势策略仍然为 $[0.1, 0.2, 0.3, 0.4]$.

在图 9 中, LM 同样采用 4 种不同的策略, 这 4 种策略都不能使系统收益达到最高, 但比较这 4 种策略, ZD 策略对应的系统收益要明显高于其他 3 种策略, 这说明 ZD 是 LM 的最佳选择, 而 TFT 策略则是最不可取的策略.

图 10 为 LM 采用 ZD 策略, SM 分别采取其他 3 种策略. 从图 10 可以看出, 随着迭代次数的增加,

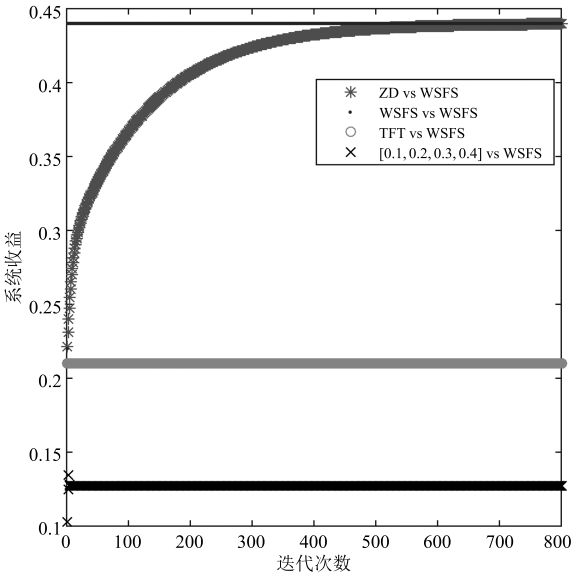

图 7　当矿工算力不相同时, SM 采用 WSFS 策略, LM 采用不同的四种策略, 系统收益情况

Fig. 7　The distribution of systematic revenue, when SM all uses WSFS strategy and LM takes different four strategies and in the condition of different power

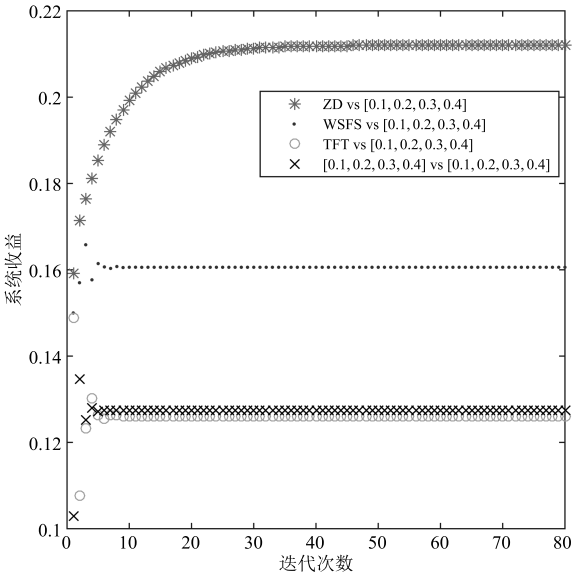

图 9　当矿工算力不相同时, SM 采用 $[0.1, 0.2, 0.3, 0.4]$ 策略, LM 采用不同的四种策略, 系统收益情况

Fig. 9　The distribution of systematic revenue, when SM all uses $[0.1, 0.2, 0.3, 0.4]$ strategy and LM takes different four strategies and in the condition of different power

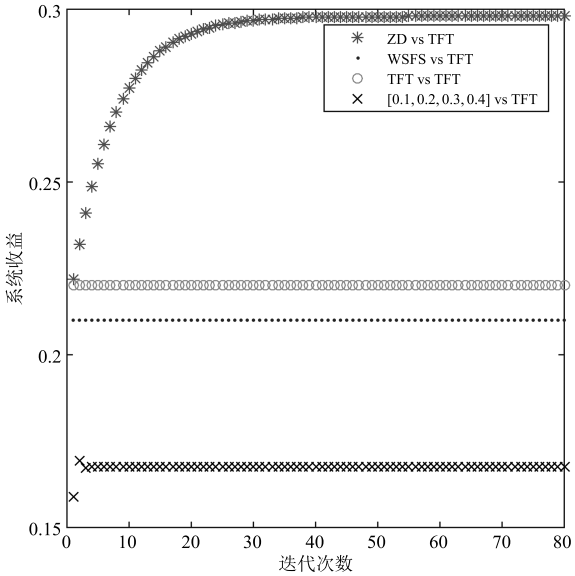

图 8　当矿工算力不相同时, SM 采用 TFT 策略, LM 采用不同的四种策略, 系统收益情况

Fig. 8　The distribution of systematic revenue, when SM all uses TFT strategy and LM takes different four strategies and in the condition of different power

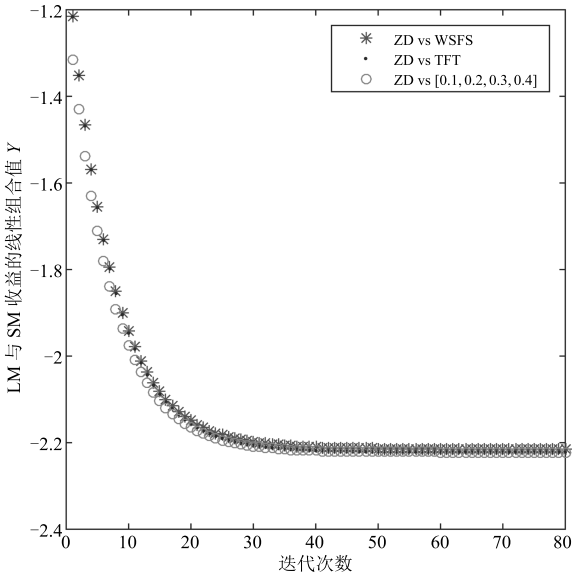

图 10　当矿工算力不相同时, LM 采用 ZD 策略, SM 采用不同的三种策略, LM 和 SM 收益的线性关系

Fig. 10　The linear relationship between LM′ payoff and SM′ payoff, when LM takes a ZD strategies and SM uses three different strategy in the condition of different power

LM 和 SM 收益的线性组合 $\gamma = \alpha U^L + \beta U^S$, 会逐渐稳定到同一值, 由此验证了我们的结论: 当矿工算力不同时, 不论 SM 采取何种策略, LM 都可以通过采用 ZD 策略, 使两个矿工的收益保持线性关系.

综合以上分析, 可以得到下面的结论: 无论矿工算力是否相同, 无论 SM 采取何种策略, 当 LM 采用 ZD 策略时, 经过一定次数的重复博弈, 可以获得相对高的系统收益, 甚至能达到最优. 并且无论 SM 选择哪种策略, 经过一定次数的迭代, LM 都可以使 SM 的收益与自己的收益保持线性关系, 这使得设计高效的博弈共识算法成为可能.

5 总结与展望

一方面, 均衡分析是博弈论中的一块重要内容. 本文对共识算法挖矿困境的分析, 对理解和剖析 PoW 共识算法本身具有一定的理论参考价值. 具体地, 当矿工算力相同时, 3 种情况的纯策略纳什均衡分为 (A, A), (C, C) 和 (A, A), (C, C). 对于混合策略均衡, 给出均衡存在条件, 并得到两个重要结论: 1) 当收益扩大倍数 r 不变时, 资源耗费 c 与合作概率成正比; 2) 当 c 不变时, 收益扩大倍数 r 与合作概率成反比. 当矿工算力不相同时, 根据参数满足不同的条件, 可以将此时的困境分为 4 种, 纯策略纳什均衡分别为 (A, A), (C, C), (A, A) 和 (C, C), (A, C). 通过分析得到混合策略均衡存在条件, 以及一个重要结论: 当收益扩大倍数 r 和费用支出 c 保持不变时, 矿工的算力与合作的概率成反比. 这些性质, 对矿工挖矿的均衡选择起到理论指导意义.

另一方面, 有效共识机制的设计一直是区块链技术的核心问题. 在 PoW 共识算法中正常挖矿的纳什均衡是相互攻击, 给系统本身造成资源浪费. 本文应用 ZD 策略对 PoW 共识算法中矿工的策略选择进行优化, 使得系统收益达到最大化, 为进一步设计基于博弈论的高效共识算法提供了新的研究思路和方法. 具体地, 对 PoW 共识过程中矿工的策略选择进行优化, 发现 LM 采取 ZD 策略, 可以使系统收益达到较高值, 甚至控制系统收益, 使之达到最优, 也可以使 SM 的收益与自己的收益保持线性关系, 并给出数值仿真, 进一步说明结论的正确性.

此外, 以太坊中的 PoS 共识算法中也存在类似的策略选择困境: 权益较大者忠实于矿池是否能获得高收益, 权益较小者忠实于矿池是否一定能维护自己的权益并获得相应收益. 后续工作将对这种情况及区块链中类似模型进行分析, 通过对系统均衡的分析, 设计更合理更有效的共识机制.

References

1 Yuan Yong, Wang Fei-Yue. Blockchain: the state of the art and future trends. *Acta Automatica Sinica*, 2016, **42**(4): 481−494
(袁勇, 王飞跃. 区块链技术发展现状与展望. 自动化学报, 2016, **42**(1): 481−494)

2 Block chain depth report: let the world be your witness (2) [Online], available: http://www.wxrw123.com/rm/201604 20/472846_2.html, April 20, 2016

3 Blockchain introduction [Online], available: http://bravenewcoin.com/assets/Uploads/TransactionsAsProofOfStake-10.pdf, December 14, 2015

4 McConaghy T, Marques R, Müller A, De Jonghe D, McConaghy T T, McMullen G, Henderson R, Bellemare S, Granzotto A. BigchainDB: a scalable Blockchain database [Online], available: https://www.bigchaindb.com/whitepaper/bigchaindb-whitepaper.pdf, June 8, 2016

5 Castro M, Liskov B. Practical byzantine fault tolerance and proactive recovery. *ACM Transactions on Computer Systems*, 2002, **20**(4): 398−461

6 Fan Jie, Yi Le-Tian, Shu Ji-Wu. Research on the technologies of Byzantine system. *Journal of Software*, 2013, **24**(6): 1346−1360
(范捷, 易乐天, 舒继武. 拜占庭系统技术研究综述. 软件学报, 2013, **24**(6): 1346−1360)

7 Bentov I, Lee C, Mizrahi A, Rosenfeld M. Proof of activity: extending Bitcoin′s proof of work via proof of stake. *ACM SIGMETRICS Performance Evaluation Review*, 2014, **42**(3): 34−37

8 Consensus in Bitcoin: one system, many models [Online], available: https://freedom-to-tinker.com/blog/randomwalker/consensus-in-bitcoin-one-system-many-models/, December 26, 2014

9 Rosenfeld M. Analysis of bitcoin pooled mining reward systems. Distributed, parallel, and cluster computing. arXiv preprint arXiv:1112.4980, 2011.

10 Block withholding attacks-recent research [Online], available: http://blog.bettercrypto.com/?p=1131, December 2, 2014

11 Eyal I, Gün Sirer E. It′s time for a hard bitcoin fork [Online], available: http://hackingdistributed.com/2014/06/13/timefor-a-hard-bitcoin-fork/, June 13, 2014

12 Courtois N T. Bitcoin miner optimization [Online], available: http://www.knaw.nl/shared/resources/actueel/bestanden/140212_Bitcoin_presentatie_Nicolas_Courtois.pdf, August 24, 2017

13 Study when be threatened, the bitcoin pool attacking [Online], available: http://www.bitecoin.com/online/2015/01/11102.html, January 4, 2015

14 Tang C B, Li A, Li X. When reputation enforces evolutionary cooperation in unreliable MANETs. *IEEE Transactions on Cybernetics*, 2015, **45**(10): 2190−2201

15 Rong Z H, Wu Z X, Chen G R. Coevolution of strategy-selection time scale and cooperation in spatial prisoner′s dilemma game. *Europhysics Letters*, 2013, **102**(6): 68005

16 Hofbauer J, Sigmund K. *Evolutionary Games and Population Dynamics*. Cambridge: Cambridge University Press, 1998. 50−54

17 Eyal I. The miner′s dilemma. In: Proceedings of the 2015 IEEE Symposium on Security and Privacy (SP). San Jose, CA, USA: IEEE, 2015. 89−103

18 Larimer D. Transactions as proof-of-stake [Online], available: http://bravenewcoin.com/assets/Uploads/TransactionsAsProofOfStake10.pdf, November 28, 2013

19 BitFury Group. Proof of stake versus proof of work [Online], available: http://bitfury.com/content/5-white-papers-research/pos-vs-pow-1.0.2.pdf, September 13, 2015

20 Popov S. The tangle [Online], available: https://iotatoken.com/IOTA_Whitepaper.pdf, December 28, 2015

21 Kwon J. Tendermint: consensus without mining [Online], available: https://iota.org/IOTA_Whitepaper.pdf, September 17, 2015

22 Press W H, Dyson F J. Iterated prisoner′s dilemma contains strategies that dominate any evolutionary opponent. *Proceedings of the National Academy of Sciences of the United States of America*, 2012, **109**(26): 10409−10413

23 Hilbe C, Wu B, Traulsen A, Nowak M A. Evolutionary performance of zero-determinant strategies in multiplayer games. *Journal of Theoretical Biology*, 2015, **374**: 115−124

24 Pan L M, Hao D, Rong Z H, Zhou T. Zero-determinant strategies in iterated public goods game. *Scientific Reports*, 2014, **5**: Article number: 13096

25 Zhang H Q, Niyato D, Song L Y, Jiang T, Han Z. Zero-determinant strategy for resource sharing in wireless cooperations. *IEEE Transactions on Wireless Communications*, 2016, **15**(3): 2179−2192

26 Al Daoud A, Kesidis G, Liebeherr J. Zero-determinant strategies: a game-theoretic approach for sharing licensed spectrum bands. *IEEE Journal on Selected Areas in Communications*, 2014, **32**(11): 2297−2308

27 Nash J F. Equilibrium points in N-person games. *Proceedings of the National Academy of Sciences of the United States of America*, 1950, **36**(1): 48−49

唐长兵 博士, 浙江师范大学数理与信息工程学院讲师. 2015 年获得复旦大学博士学位. 主要研究方向为复杂网络, 博弈理论及其应用, 网络安全与优化控制.
E-mail: tangcb@zjnu.cn
(**TANG Chang-Bing** Ph. D., lecturer at the College of Mathematics, Physics and Information Engineering, Zhejiang Normal University. He received his Ph. D. degree from Fudan University in 2015. His research interest covers complex networks, game theory and application, network security, and optimal control.)

杨 珍 浙江师范大学数理与信息工程学院硕士研究生. 2015 年获得河北科技师范学院学士学位. 主要研究方向为区块链技术, 博弈理论及其应用.
E-mail: y_zh_en@163.com
(**YANG Zhen** Master student at the College of Mathematics, Physics and Information Engineering, Zhejiang Normal University. She received her bachelor degree from Hebei Normal University of Science and Technology in 2015. Her research interest covers blockchain technology, game theory and application.)

郑忠龙 浙江师范大学数理与信息工程学院教授. 2005 年获得上海交通大学博士学位. 主要研究方向为数据科学, 机器学习, 图像处理.
E-mail: zhonglong@zjnu.edu.cn
(**ZHENG Zhong-Long** Professor at the College of Mathematics, Physics and Information Engineering, Zhejiang Normal University. He received his Ph. D. degree from Shanghai Jiao Tong University in 2005. His research interest covers data science, machine learning, and image processing.)

陈中育 浙江师范大学数理与信息工程学院教授. 2011 年获得上海大学博士学位. 主要研究方向为软件形式化方法, 需求建模技术, 区块链应用技术. 本文通信作者. E-mail: czy@zjnu.cn
(**CHEN Zhong-Yu** Professor at the College of Mathematics, Physics and Information Engineering, Zhejiang Normal University. He received his Ph. D. degree from Shanghai University in 2011. His research interest covers formal methods, requirement modeling, and blockchain applications. Corresponding author of this paper.)

李 翔 复旦大学信息科学与工程学院教授. 2002 年获得南开大学博士学位. 主要研究方向为复杂网络与系统控制.
E-mail: lix@fudan.edu.cn
(**LI Xiang** Professor at the School of Information Science and Technology, Fudan University. He received his Ph. D. degree from Nankai University in 2002. His research interest covers complex networks and system control.)