文献引用格式： 杨天，薛质 . 区块链系统中矿池间的博弈问题及优化 [J]. 通信技术，2019，52（05）：1189-1195.

YANG Tian,XUE Zhi.Game Theory among Mining Pools in Blockchain System[J].Communications Technology,2019,52(05):1189-1195.

doi:10.3969/j.issn.1002-0802.2019.05.026

区块链系统中矿池间的博弈问题及优化 *

杨　天，薛　质

（上海交通大学 电子信息与电气工程学院，上海 200240）

摘　要: 区块链中的 PoW（工作量证明）共识算法保证了区块链系统的安全性和可靠性。在挖矿期间，矿工之间可以通过 PoW 算法达成共识。然而除了合作关系，矿工之间是可以互相攻击的，并且矿工可以通过攻击其他矿工增加自己的收益，然而在有矿工选择攻击的情况下，矿池的总收益是会减少的。我们将区块链系统中各矿工选择攻击与合作的问题建立成一个博弈论模型，并以宏观的角度分析，以矿池为单位进行研究，对目前常用的策略进行了比较，并在此基础上提出了一种定值策略，它可以更好地解决各个矿池选择合作还是攻击的困境，从而提高全网的总收益。总的来说，我们从博弈论的角度对区块链系统中各矿池可选择的策略进行了剖析，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；矿池；博弈论；定值策略

中图分类号：TP399　　文献标志码：A　　文章编号：1002-0802(2019)-05-1189-07

Game Theory among Mining Pools in Blockchain System

YANG Tian, XUE Zhi

（School of Electronic Information & Electrical Engineering, SJTU, Shanghai 200240，China）

Abstract: The PoW (workload proof) consensus algorithm in the blockchain guarantees the security and reliability of the blockchain system. During mining period, miners may reach a consensus through the PoW algorithm. However, apart from the cooperative relationship, miners can attack each others, and increase their profits by attacking other miners. However, when miners choose to attack, the total profits of the mining pool would decrease. This paper proposes a game theory model for each miner's choice of attack or cooperation in the blockchain system, and analyzes it from a macropoint of view, discusses the mining pool as a unit, compares the commonly-used strategies at present, and puts forward a setting strategy on this basis, which can be used to fairly solve the dilemma of choosing cooperation or attack by each mining pool, thus improving the total revenue of the whole system. Generally, speaking, from the point of view of game theory, this paper analyzes the strategies that can be chosen by each mining pool in the blockchain system, and provides new ideas and methods for further design of consensus algorithm base on the game theory.

Key words: blockchain; mining pool; game theory; setting strategy

*　收稿日期：2019-01-22；修回日期：2019-04-11　　Received date:2019-01-22;Revised date:2019-04-11

0 引 言

区块链是一种分布式共享总账，系统中的每一个参与者都负责数据的记录与存储，从而实现了去中心化[1]。目前在各种区块链系统中，比特币是区块链最成功的应用之一，它在区块的生成过程中使用了 PoW（Proof of Work，工作量证明）机制。PoW 是一种激励性算法，它的核心概念是通过矿工之间的竞争来保证数据的安全性、连续性与一致性。系统中各节点根据各自的算力争相角逐，共同解决一个求解复杂但是验证容易的 SHA256 数学问题[2]。其中解决该问题的节点会提供一个合理的 Block Hash，同时获得在区块中记录数据的权利，并得到系统自动生成的比特币奖励。一个符合要求的 Block Hash 由 N 个前导零构成，零的个数取决于网络的难度值。要得到合理的 Block Hash 需要经过大量尝试计算，计算时间取决于机器的哈希运算速度。当某个节点提供出一个合理的 Block Hash 值，说明该节点确实经过了大量的尝试计算，当然，并不能得出计算次数的绝对值，因为寻找合理 Hash 是一个概率事件。可见在 PoW 机制下，每个矿工为了获得比特币奖励将会尽其所能地利用其算力解决问题并尝试挖矿，新的数据区块不断生成，从而产生了区块链[3]。

其它基于 PoW 的区块链系统也是以同样的方式运作的，虽然各种系统所提出的数学难题并不相同，但是都有一个共同特征：用算力换收益，并且利用分布网络节点的工作量证明使各个节点达成共识，从而实现交易数据的记录与存储，同时产生了一套有时间先后顺序的，不可篡改的，可信任的数据库。通过复杂的校验机制，区块链系统可以保证数据库中数据的完整性，连续性，和一致性。即使部分参与者作假也无法改变区块链的完整性，也无法篡改区块链中的数据。简而言之，区块链技术涉及的关键点有：去中心化、集体维护、去信任、可靠数据库、时间戳、非对称加密等[4]。

1 背景知识

1.1 矿池概念

区块链系统中，生成区块的过程被称为挖矿，所有参与挖矿的节点被称为矿工。由于全网算力非常巨大，而区块产出有限，单个设备或少量的算力都很难在比特币网络上获取到比特币网络提供的区块奖励。为追求稳定收益，人们自发地将算力联合起来形成矿池。一个矿池有一位管理员，当矿池的一位参与者挖到区块时，比特币奖励会发送到矿池管理员。然后，管理员根据每个参与者对矿池算力的贡献，向参与者发放比特币奖励。在此机制中，不论个人矿工所能使用的运算力多寡，只要是通过加入矿池来参与挖矿活动，无论是否有成功挖掘出有效资料块，皆可经由对矿池的贡献来获得少量比特币奖励，亦即多人合作挖矿，获得的比特币奖励也由多人依照贡献度分享。

1.2 区块截留攻击

由于每一个矿工只要提供一个网络 ID 数字就可以加入矿池，一个开放的矿池很容易遭受攻击。具体形式为：一些矿工会发送部分工作量证明（Partial Proof of Work，对产出几乎无帮助，只是证明干了活），抛弃完整工作量证明（Full Proofs of Work，收益来源）。这就是所谓的区块截留攻击，这个行为看似损人不利己，那么选择攻击的矿工的目的是什么呢？主要原因是，区块链协议具有难度自适应的特征，会根据当前区块生成速度调整前导 0 的个数，从而改变难度，控制区块生成速度保持不变。有矿工选择攻击会导致矿池有效算力减少，协议为了保持区块生成速度不变，自会降低挖矿难度，这样滥竽充数的矿工就会得到更多收益。其次是因为，得到奖励的矿池会根据工作量证明按照矿池中每个矿工贡献算力的比例将所获奖励分配给每一个矿工。而完整的工作量证明很难生成，偷奸耍滑的矿工可以选择向开放矿池发送部分工作量证明来获得本该贡献实际算力才能得到的奖励。

在一个开放矿池中，一个矿工可以选择与其他矿工合作或是对其他矿工发动区块截留攻击，两种选择都会获得相应的收益。当所有矿工都选择攻击时，它们获得的收益比所有矿工都不选择攻击时获得的收益要小。这种情况被称为 PoW 共识算法中的挖矿困境[5]，对于一个矿工来说攻击是最好的选择，而这个选择会使系统收益减少。所以，为了促使矿池中的所有矿工合作挖矿提高系统收益，需要开发一种激励性机制来促进矿工合作从而优化区块链 PoW 共识算法。

1.3 初步建模及策略

为了避免区块链中的矿工陷入挖矿困境，选择合作的矿工可以采取一些策略来解决攻击矿工“拿钱不干事”的问题或者尽量减小损失。有了相应的

策略，矿池中一个矿工不管与它竞争的对手矿工采取什么策略，它都能单方面地控制对手从自己这里得到的期望收益并分给对手一定比例的期望收益从而促进对手更倾向于合作。以更宏观的视角分析，不妨尝试将两个矿工之间的博弈以类似的思路放大为两个矿池间的博弈。考虑到全网的其中两个矿池 A 和矿池 B，B 矿池派出总算力为 b 的矿工，在 A 矿池注册，这些矿工在 A 矿池进行区块截留攻击。这样一来总算力降低，根据协议，区块生成难度降低，B 矿池获得正收益，而 A 矿池通过不断减少的收益中能发现它正在遭受攻击，但很难发现究竟是那些矿工进行攻击。实际情况中，攻击往往是双向的，设 A 矿池派出的渗透算力为 a，B 矿池的渗透算力为，则对于 A 矿池来说，面对 B 矿池可以采取多种策略来应对。

目前常见的策略有如下 9 种[6](分别用 P_n 表示)：

P_1--ALLC：All Cooperate，永远合作策略，无论对手采取何种策略，都选择合作，即令渗透算力 a 恒为 0。

P_2--ALLD：All Defect，永远背叛策略，无论对手采取何种策略，都选择背叛，即令渗透算力 a 恒为最大。

P_3--TFT：Tit For Tat，一报还一报策略，第一次选择合作 (即 a=0) 此后，如果对手的渗透算力 b 大于某一阈值，则下一轮背叛 (即 a=max)，否则合作 (即 a=0)。

P_4--Grim：冷酷策略，第一次选择合作 (即 a=0), 只要对手背叛一次，就不再合作，令 a=max。

P_5--WSFS：Win Stay Fail Shift，赢存输变策略，第一次选择合作 (即 a=0)，之后每一轮如果收益高于某一阈值，就保持策略不变，否则采取相反的策略 (即 a=max)。

P_6--Random：离散型随机取值策略，a 以等概率取 0 或 max。

P_7--TFT_D：TFT_Defect，类似于 TFT，区别在于第一次选择背叛 (即 a=max)。

P_8—Grim_D：Grim_Defect，同上，类似于 Grim。

P_9—WSFS_D：WSFS_Defect，同上，类似于 WSFS。

2　矿池博弈模型与 IPD

2.1　数学建模

根据 1.3 所描述的情况，全网中的两个矿池 A 和 B 互相派出渗透矿工攻击对方。为方便计算，设全网总算力为 1，A 矿池算力为 s，B 矿池算力为 t，显然 $s+t<1$。A 矿池派出的渗透算力为 a，B 矿池的渗透算力为 b，派出的间谍矿工将进行区块截留攻击，因此不贡献任何算力。能派出的算力有限，故可设 $0 \leqslant a \leqslant 0.1s$。全网实际算力从 1 降低为 $1-a-b$，故单位算力的产出速度提升到原来的 $\frac{1}{1-a-b}$ 倍。

设 $\lambda=\frac{1}{1-a-b}$，则 A 池产出为：

$$A(a,b)=(s-a)\lambda\frac{s-a}{s-a+b}+(t-b)\lambda\frac{a}{t-b+a} \tag{1}$$

同理，B 池产出为：

$$B(a,b)=(s-a)\lambda\frac{b}{s-a+b}+(t-b)\lambda\frac{t-b}{t-b+a} \tag{2}$$

由于 $0 \leqslant a \leqslant 0.1s$，$0 \leqslant b \leqslant 0.1t$，$s$ 与 t 的值相差不大，故可将 a、b 视为小量，则 $\frac{a}{s}$、$\frac{b}{s}$、$\frac{a}{t}$、$\frac{b}{t}$、a^2、b^2、ab 趋近于 0。所以等式 (1) 可以化简为：

$$A(a,b)=\frac{s^2-2as}{s-as-bs-a-b}+\frac{at}{t-at-bt+a-b}=\frac{s-a}{1-a-b} \tag{3}$$

同理：

$$B(a,b)=\frac{t-b}{1-a-b} \tag{4}$$

设 s=0.2，t=0.2，b=0，$A(a,0)$ 的图像如图 1 所示。

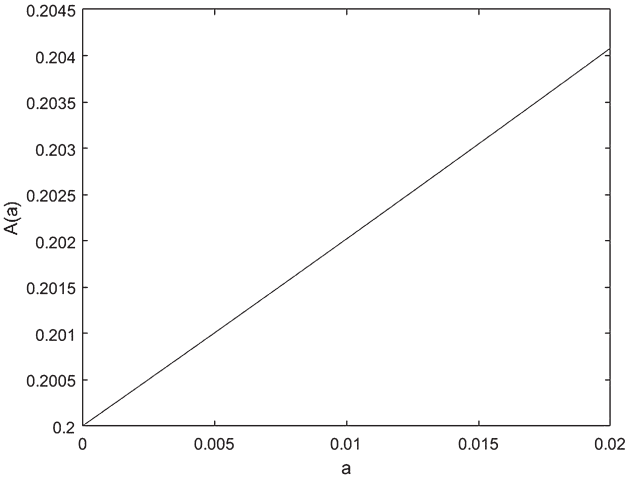

图 1　矿池 B 选择合作时，A 矿池收益与渗透算力 a 的关系

这是近似线性的单调递增函数，即 A 矿池攻击

越强（$a\uparrow$），其收益越高（$A(a,b)\uparrow$）。

将（3）与（4）相加得到 A 矿池与 B 矿池的总收益：

$$G(a,b)=\frac{s+t-(a+b)}{1-(a+b)} \quad (5)$$

将 $a+b$ 视作一个变量，则 $G(a,b)$ 与 $a+b$ 关系如图 2 所示。

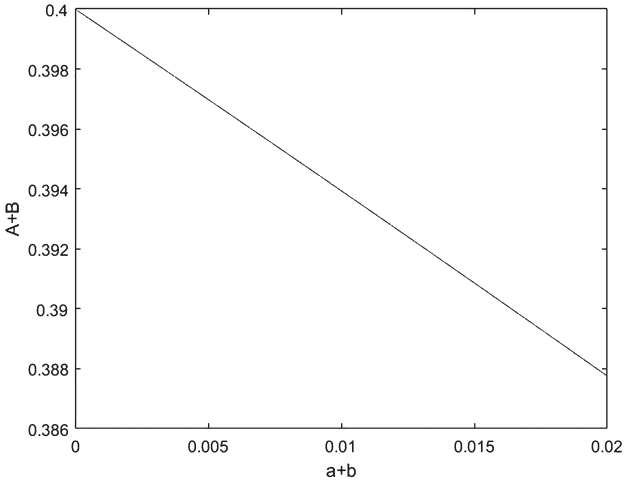

图 2 A、B 两矿池收益总和与渗透算力 a、b 之和的关系

这是近似线性的单调递减函数，即两者的攻击越强（$(a+b)\uparrow$），总收益越低（$G(a,b)\downarrow$）。

2.2 囚徒困境与 IPD

以上所述的特征与经典的囚徒困境十分相似。即两个共谋嫌疑犯被捕后单独审讯，如果两个人都不坦白罪行，则由于证据不足各判一年；如果其中一人坦白，另一人不坦白，则坦白的人因立功而立即获释，不坦白的人因不合作而被判十年；两个人都坦白罪行，则证据确凿，各判八年。由于囚徒无法信任对方，因此倾向于互相揭发，而不是同守沉默。最终导致纳什均衡仅落在非合作点上的博弈模型。其收益矩阵如表 1 所示。

表 1 经典囚徒困境收益矩阵

囚徒选择	坦白	沉默
坦白	（−1，−1）	（0，−10）
沉默	（−10，0）	（−8，−8）

这与挖矿困境的相同点在于，坦白（攻击）对个体而言是最优的选择，而对于总体（全网）而言，每个个体都选择坦白（攻击）就不是最优选择了。不同点在于，经典囚徒困境中策略集只有（坦白，沉默）两种取值，是一个离散问题；而矿池间的博弈问题中，策略集是渗透算力（a,b），是连续的控制量。

囚徒困境指出，如果只进行一次博弈，那么双方都会毫无疑问地选择背叛（在矿池博弈中，即令 a=0.1s，b=0.1t）。但在实际情况中，双方往往会进行多次，甚至海量的相互博弈，为此引入 IPD（Iterated Prisoner’s Dilemma，重复囚徒困境）模型。IPD 中，博弈被反复地进行。因而每个参与者都有机会去“惩罚”另一个参与者前一回合的不合作行为。这时，合作可能会作为均衡的结果出现。欺骗的动机这时可能被受到惩罚的威胁所克服，从而可能导向一个较好的、合作的结果。作为反复接近无限的数量，纳什均衡趋向于帕累托最优（没有再进行帕累托优化的余地）[7]。

囚徒困境的主旨为，囚徒们虽然彼此合作，坚不吐实，可为全体带来最佳利益，但在无法沟通的情况下，因为出卖同伙可为自己带来利益，也因为同伙把自己招出来可为他带来利益，因此彼此出卖虽违反最佳共同利益，反而是自己最大利益所在。但实际上，执法机构不可能设立如此情境来诱使所有囚徒招供，因为囚徒们必须考虑刑期以外之因素（出卖同伙会受到报复），而无法完全以执法者所设立之利益（刑期）作考量。

同样的，对于两个矿池 A 和 B 在无限次重复博弈中，博弈者可以采用的策略有无穷多种，采用什么样的策略才可以实现相对更高的收益呢？收益较高的策略之间存在共性吗？ Axelrod 实验可以解决这个问题。

Axelrod 实验是以计算机程序对弈、竞赛的方式进行的。他要求参与竞赛的编程者充当囚徒困境的局中人，以谋求博弈收益的长期最大化为目标，用计算机程序编成特定的策略，每一策略按一定的规则实施合作或者背叛来对付对手。然后用单循环赛的方式将有所参赛程序两两对弈。[6] 显然，不同的策略对弈会有不同的得分结局，这样就可以通过比较每个策略得分的多少来衡量其优劣。

2.3 不定值策略与定值策略

第一节最后提到的九种策略正是经典 Axelrod 实验中的常见策略，在本文的矿池博弈问题中，策略集被连续化，取值并不确定，在此基础上设计出新的策略：

P_{10}-- 不定值策略：a 的取值在 [0,max] 区间内均匀分布。

在此基础上对不定值策略进行帕累托优化（在

没有使任何矿池收益减少的情况下，至少使一个矿池的收益增加）。单次博弈不可能保证该优化的实现，因为若 a=0，b=0，根据等式（5），总收益 $G(a,b)$ 已达到最大值，则矿池 A,B 其中一个收益增加必然会导致另一个收益减少；同理，若 a、b 不全为 0，说明至少有一个矿池在攻击对方，则无论是攻击矿池还是合作矿池收益增加后必然会导致对方收益减少。所以只有尽可能地在重复博弈中实现帕累托优化。在非合作关系中（a、b 不全为 0），纳什均衡状态即为帕累托最优，即总收益 $G(a,b)$ 不变。为了接近最优状态，根据等式（5），a，b 不全为 0 时，a 在 b 不为 0 时取 0；而在 b 再次为 0 时，设 b 从不为 0 到重新取 0 共经历了 n+1 轮，a 此时取值为 m，且满足本轮 $G(a,b)$ 的 n 倍等于前 n 轮 $G(a,b)$ 之和。这就是基于不定值策略的优化策略 —— 定值策略：

P_{11}—— 定值策略：第一次令 a=0，若此后 b 一直取 0，则 a 继续取 0，若某一轮 b>0，则 a 继续取 0，设本轮 b 取值为 b_1，下一轮若 b>0 则 a 继续取 0，b 取值为 b_2，依此类推。在 b>0 之后的第 n+1 轮 b 再次为 0，此时设 a 取值为 m，且 m 满足：

$$\frac{n(s+t-m)}{1-m}=\sum_{i=1}^{n}\frac{s+t-b_i}{1-b_i} \tag{6}$$

该策略理论上可以使两矿池接近纳什均衡，具体表现还要通过数值仿真来验证。

3　数值仿真

3.1　两两博弈的仿真结果

本文一共提出了 11 种不同的策略（经典策略 9 种，新设计策略 2 种），两两匹配，共有 55 次不同的博弈（每次模拟 100 轮），这里仅选取其中一些具有代表性的结果进行展示和分析。

计算可知，每轮博弈中，一方的最低收益为 0.183 7，最高收益为 0.204 1，差距仅 10% 左右，直接将收益作为纵轴作图，近似为线性，很不直观。为解决这个问题，可将每轮的收益减去 0.183 7，称为“额外收益”，并以此为纵轴作图，使得结果较为直观。

部分结果如图 3、图 4、图 5、图 6 所示。

图 3 表明，ALLC 策略平均收益低于离散型随机取值策略。

图 4、图 5 表明，当定值策略、ALLC、TFT、Grim、WSFS 等友善型策略两两互相博弈时，会进行持续合作从而获得不错的收益（约 1.6）。

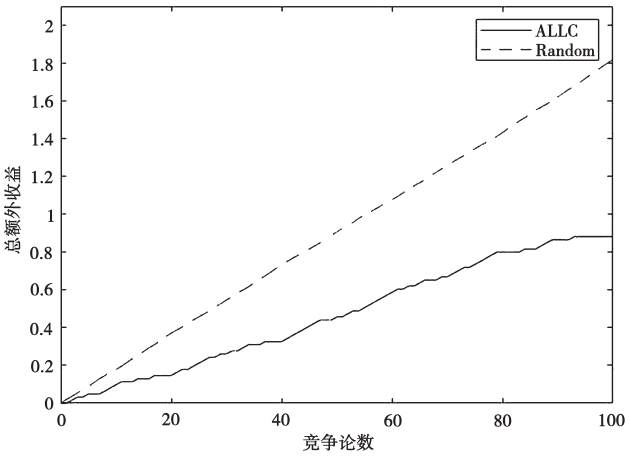

图 3　P_1 与 P_6 的竞争结果

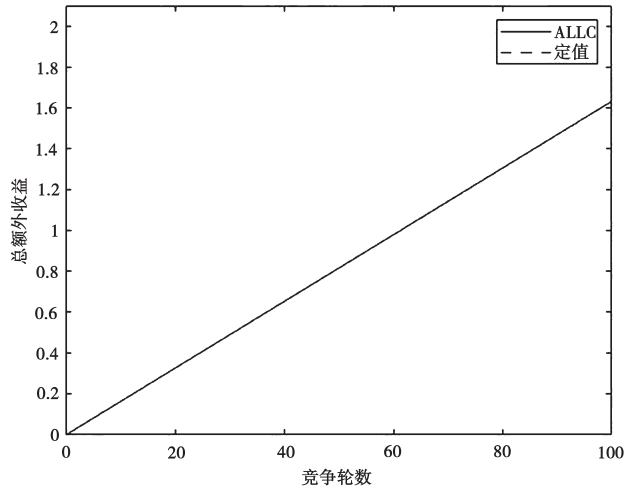

图 4　P_1 与 P_{11} 的竞争结果

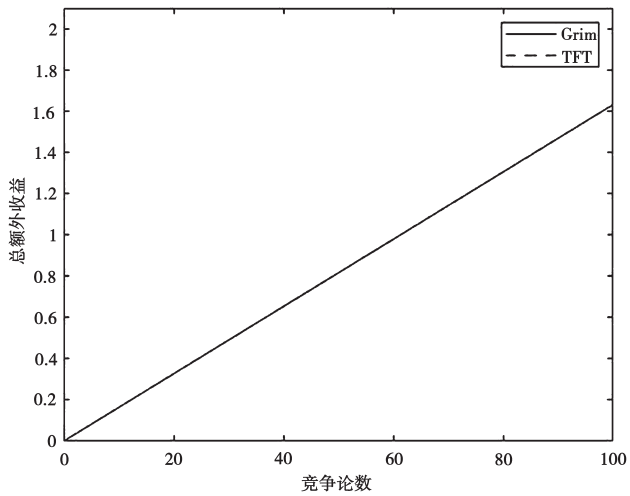

图 5　P_3 与 P_4 的竞争结果

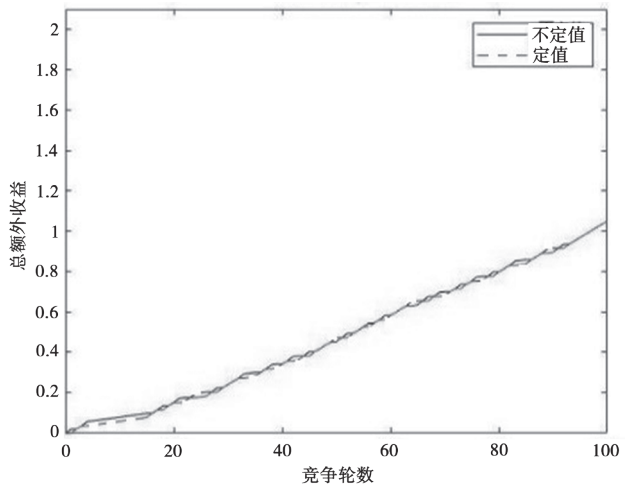

图 6 P_{10} 与 P_{11} 的竞争结果

图 6 表明，定值策略会根据对手策略做出相应的调整，导致不管和谁博弈，其收益都接近，除了 ALLD。定值策略和 ALLD 的博弈与 ALLC 和 ALLD 的博弈是一样的。

3.2 总体表现的仿真结果

单看两两博弈的结果很难分析每一个策略的优劣，为此，我们统计了每个策略与其它所有策略进行博弈时的平均收益，并从高到低进行排序，结果如图 7 所示。

图 7 表明，表现最好的是定值策略，WSFS 仅次于定值策略。表现最差的是 TFT_D、Grim_D。

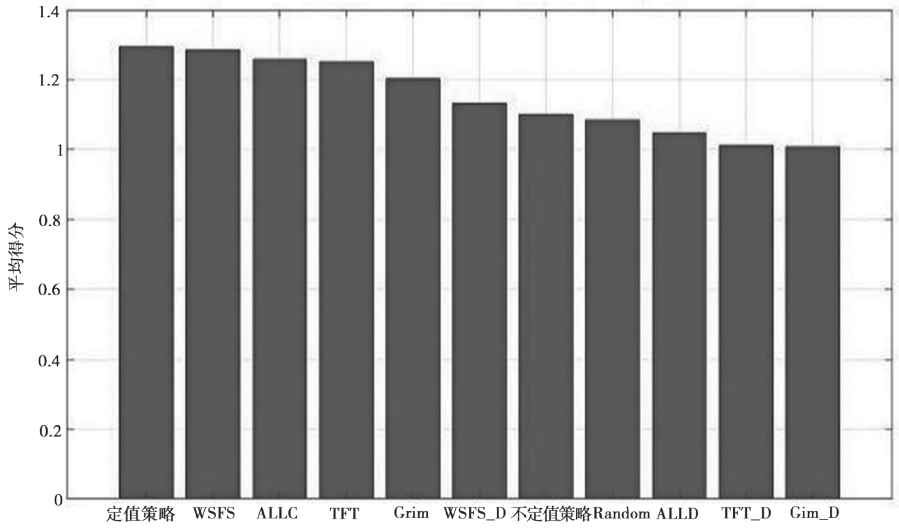

图 7 11 种策略的总体表现排名

3.3 仿真结果分析

对上述仿真结果进行分析，归纳出该博弈模型的以下特点：

（1）ALLD 策略在每一次博弈中都不吃亏，但在总体上却是一种很差的策略；

（2）一般而言，“善意”的策略表现要优于“贪心”的策略；

（3）第一次的选择非常重要，对整个策略的表现影响很大。

传统 IPD 中，TFT 策略和 Grim 策略表现最佳。而在该区块链博弈模型中，定值策略和 WSFS 策略表现最好，且定值策略较传统的 WSFS 更胜一筹。

4 总结与展望

矿池间的相互攻击是区块链中的常见现象，但此前的研究工作大多集中在单个矿池内矿工间的相互博弈。在本文中，我们首次对该问题进行了研究。

在研究方法上，我们选用了数学建模和数值仿真的方法，参考了经典的囚徒困境模型和 IPD 模型，但根据该问题的特点自行建立了具有连续性的收益矩阵，并根据该矩阵设计了两种新的策略——不定值策略、定值策略，并编写代码进行仿真。

在仿真中，我们比较了 11 种不同的策略，结果也与传统 IPD 模型有较大差异。根据结果，我们建议矿池的管理者采取定值策略或 WSFS 策略。

在矿池的博弈模型中一定存在更为优秀的策略等待我们去发掘，也有很多问题摆在我们面前。是否可以引入某些遗传算法，令目前设计的策略进行组合、突变或演化呢？超过两个矿池的多矿池博弈模型中，各种策略及其效果又会如何呢？后续工作将对这些情况进行研究和分析。

参考文献：

[1] Nakamoto S.Bitcoin:a Peer-to-peer Electronic Cash System[J].Consulted,2008(01):2012.

[2] Yang Zhen,Miao Yue,Chen Zhong-yu,et al.Zero-determinant Strategy for the Algorithm Optimize of Blockchain PoW Consensus[C]//Proceedings of the 36th Chinese Control Conference.Dalian:[s.n.],2017:1441-1442.

[3] Gramoli Vincent. From Blockchain Consensus Back to Byzantine Consensus[J].Future Generation Computer Systems,2017:S0167739X17320095.

[4] Yaga D,Mell P,Roby N,et al.Blockchain Technology Overview[EB/OL].(2018-10-05).https://csrc.nist.gov/publications/detail/nistir/8202/draft.

[5] Eyal I.TheMiner'sDiemma[EB/OL].(2014-12-03).http://hackingdistributed.com/2014/12/03/the-miners-dilemma.

[6] Axelrod Robert,William Donald Hamilton.The Evolution of Cooperation[J].Science,1981(211):1390-1396.

[7] Press William,Dyson Freeman.Iterated Prisoner's Dilemma Contains Strategies That Dominate Any Evolutionary Opponent[C].Proceedings of the National Academy of Sciences of the United States of America.USA:[s.n.],2012:587-588.

作者简介：

杨　天（1993—），男，硕士，主要研究方向为网络攻防；

薛　质（1971—），男，博士，教授，主要研究方向为网络攻防。