

# 基于智能合约的数据共享激励机制研究

张伯钧<sup>1,2</sup>, 郭一晨<sup>1,2</sup>, 王子凯<sup>1,2</sup>, 胡凯<sup>1,2</sup>

(1.北京航空航天大学 计算机学院,北京 100191; 2.云南省区块链应用技术重点实验室,昆明 650233)

**摘要:**随着大数据时代的到来,构建数据共享社区成为一种促进数据收集、流通和使用的新模式,但是在传统社区组织方式下,成员由于互不信任导致不愿共享数据,且共享社区普遍缺乏有效的激励机制,从而限制了数据共享社区的发展。区块链的本质是一个分布式账本,能够提升数据共享的透明性,防止信息滥用,所有存在区块链上的数据可溯源且不可篡改,可以保证数据的归属权,这些都为成员之间的信任打下了基础。在基于区块链的数据共享社区平台上,结合演化博弈理论构建一种数据共享激励模型EGDSI,并通过复制动态力学方程与演化稳定策略分析,得出影响数据共享的关键因素与数据共享激励的3种策略。在此基础上,利用区块链智能合约技术实现安全、高效且动态可调节的数据共享模板引擎。对共享策略影响因素和智能合约激励机制进行仿真分析,结果表明,该数据共享激励机制可以有效促进用户参与数据共享,相较EGI模型,EGDSI模型能够更快地达到数据共享饱和状态。

**关键词:**数据共享;演化博弈;激励机制;区块链;智能合约

开放科学(资源服务)标志码(OSID):



中文引用格式:张伯钧,郭一晨,王子凯,等.基于智能合约的数据共享激励机制研究[J].计算机工程,2022,48(8):37-44.

英文引用格式:ZHANG B J, GUO Y C, WANG Z K, et al. Research on data sharing incentive mechanism based on smart contract[J]. Computer Engineering, 2022, 48(8): 37-44.

## Research on Data Sharing Incentive Mechanism Based on Smart Contract

ZHANG Bojun<sup>1,2</sup>, GUO Yichen<sup>1,2</sup>, WANG Zikai<sup>1,2</sup>, HU Kai<sup>1,2</sup>

(1.School of Computer Science and Engineering, Beihang University, Beijing 100191, China;

2.Blockchain Application Technology Key Laboratory of Yunnan Province, Kunming 650233, China)

**[Abstract]** With the advent of big data, the data sharing community has become a new model to promote data collection, circulation, and use. However, under the traditional community organization mode, members may be unwilling to share data due to mutual distrust, and the sharing community generally lacks an effective incentive mechanism, which limits its development. The essence of blockchain is a distributed ledger that improves the transparency of data sharing and prevents information abuse. To ensure data ownership, the blockchain prevents the tracing or tampering of data, which lays a foundation for trust among members. The data sharing incentive model EGDSI is constructed based on blockchain and evolutionary game theory. By copying dynamic mechanical equations and analyzing evolutionary stability strategies, key factors affecting data sharing and three strategies for data sharing incentives are obtained. On this basis, the blockchain's smart contract technology is used to realize a safe, efficient and dynamically adjustable data sharing template engine. A simulation analysis of the aforementioned factors and strategies shows that the data sharing incentive mechanism effectively promotes users participation in data sharing. Compared with the EGI model, EGDSI offers faster performance in reaching the saturation state of data sharing.

**[Key words]** data sharing; evolutionary game; incentive mechanism; blockchain; smart contract

DOI: 10. 19678/j. issn. 1000-3428. 0063155

## 0 概述

在大数据时代,数据流动和信息共享对人类科技进步至关重要,数据共享社区是促进科学数据开放共享的重要载体<sup>[1-2]</sup>。近年来,我国在促进科学数据共享

方面出台了一系列举措,然而,最近的一项调查研究结果表明,我国的研究人员虽有分享数据的意愿,但也担心数据被误用以及版权和许可受到侵犯<sup>[3]</sup>,最终绝大多数学者都不愿意分享自己的数据,即无意愿开放共享,导致数据缺乏流动性,无法有效发挥数据共享社区

**基金项目:**国家重点研发计划(2018YFB1402702);云南省重大科技专项“基于服务智能合约的云南稀贵金属材料基因数据可信交易技术研发”(202002AB080001-8);云南省重大科技专项“生物资源数字化开发应用”(202002AA100007)。

**作者简介:**张伯钧(1997—),男,硕士研究生,主研方向为区块链技术、隐私计算;郭一晨,本科生;王子凯,硕士研究生;胡凯,教授、博士。

**收稿日期:**2021-11-08 **修回日期:**2022-01-11 **E-mail:** zhangbojun@buaa.edu.cn

的功能和优势。共享社区长期缺乏有效且科学的激励机制,普遍面临用户参与积极性低、归属感弱、知识生成动力不足、数据共享程度低等挑战,缓解这一问题的关键是构建一个合理有效的数据共享模型<sup>[4]</sup>。

区块链技术<sup>[5-6]</sup>的出现为解决数据共享所面临的问题提供了新的思路。区块链的本质是一个分布式账本,账本上的每一笔交易都由区块链网络上的各个节点来查看和验证。首先,区块链的分布式特性保证了数据共享流通的透明性,用户可以清晰地看到自己的数据流通过程,防止数据信息被滥用的情况发生;再次,利用区块链智能合约自动执行的特性,可以将数据共享的激励机制写入智能合约,将奖励与惩罚信息公开,规范流程从而实现奖惩透明;最后,区块链也解决了数据共享过程中的确权问题,通过区块链,所有的数据都可以逐一溯源,既可以保证数据的归属感,又可以检验数据信息的真伪。因此,将区块链与数据共享相结合,可以扩大数据规模,规范数据管理流程,实现对共享数据的确权,从而建立一个开放的数据共享生态体系。

当前,数据共享被越来越多的学者关注,其中的一个研究热点便是激励机制。庄倩等<sup>[7]</sup>建立了一种参与科学数据共享的科研人员的演化博弈模型,通过该模型分析数据共享的演化过程。WEN等<sup>[8]</sup>提出一种基于讨价还价理论的群智感知激励机制,其着重研究数据质量对数据收集者的影响。郝世博等<sup>[9]</sup>使用区块链以及智能合约,从技术维度来解决数据共享中所存在的信任问题。石进等<sup>[10]</sup>在其提出的基于区块链的中小企业竞争情报共享平台模型中,采用信用积分制度激励用户进行共享。张宝等<sup>[11]</sup>也提出了基于智能合约的积分激励体制,当用户上传数据或数据被查看时奖励用户积分,以此来激励用户共享数据。张旺等<sup>[12]</sup>通过双向委托代理分析,验证了在不承担风险的情况下,选择不共享策略能够最大限度地减少共享成本,从而说明了激励机制在数据共享中的重要性。GAN等<sup>[13]</sup>提出一种基于区块链的eHealth系统,并在该系统中根据医疗数据的质量以及患者参与共享的积极性来对患者进行激励。SHEN等<sup>[14]</sup>在研究多云数据共享时,利用Shapley值构建一个动态公平的激励方案,为多云数据共享协作提供了一种收入分配方法。SHRESTHA等<sup>[15]</sup>在建立一个基于区块链的共享平台的基础上,构建一种在以太坊上部署智能合约的用户激励模型。LI等<sup>[16]</sup>在设计一个基于区块链的车载人群感知框架时,提出一种基于用户声誉的激励机制,其在确保公平交易的同时提高了数据的质量。梁小燕<sup>[17]</sup>为促进医疗机构数据共享,提出一种基于信息熵的医疗数据共享激励机制,通过计算信息熵来判断信息的价值,并提出基于Shapley值的利益分配策略。蔡婷等<sup>[18]</sup>提出一个高效的区块链物联网数据激励共享框架ShareBC,并在此框架中设计了一种基于层次数据拍卖模型的激励机制。王倩等<sup>[19]</sup>基于区块链技术与演化博弈论,构建一种面向科技服务数据共享策略的智能合约激励机制。

虽然已有很多学者提出了各种不同的激励机制,但是这些激励机制仍然存在局限性,例如,某些激励机制不能满足用户的需求,导致用户在该机制下没有数据共享的意愿,或激励机制的奖励过多导致社区通货膨胀,从而使奖励贬值,短时间内虽然可以激励用户进行数据共享,但是该激励并不能长期保持。针对上述问题,本文建立一种基于演化博弈论的数据共享激励模型EGDSI,从数据共享社区的角度设计合理的激励机制,以促进数据的收集、流通和使用。

## 1 基于演化博弈论的数据共享激励模型

早期的博弈论认为所有的参与者均是完全理性的<sup>[20]</sup>,他们总是能思考所有情况,最终选择对自己最有利的方案。完全理性的人总是会遵行优势策略,即玩家从不会在决策时考虑其他玩家对他们的决定会如何回应。完全理性的人是完全的个人主义者,但在现实生活中,完全理性是很难达到的<sup>[21]</sup>。

演化博弈论是将达尔文的生物进化论与博弈论相结合,其不再要求参与人完全理性,而是假定参与人有限理性,所谓的有限理性是指博弈方可以对所处的不同环境进行学习、试错和模仿,在这个过程中动态调整自己的策略,从而寻找最优方案。在数据共享社区中,用户具有共享与不共享这2种行为策略,用户之间会相互影响,即利用有限的认知和信息获取能力,通过不断地学习和模仿他人策略来调整自身策略,以此实现自身利益最大化。

### 1.1 模型构建

数据共享是一个动态选择及学习调整的过程,为了构建模型和简化计算,本文进行如下假设:

1)演化博弈双方具有相同的主体,即数据共享社区的用户,他们具有相同的策略空间 $K$ , $K=\{K_1, K_2\}=\{\text{参与数据共享}, \text{不参与数据共享}\}$ 。

2)在数据共享的过程中,用户需要付出一定的成本,此处的成本包括但不限于数据制作的时间成本、数据被滥用的经济成本等。在本文模型中,数据成本记为 $G$ 。

3)假定在博弈的初级阶段,主体 $P_1$ 参与数据共享的概率为 $p$ ,主体 $P_2$ 不参与数据共享的概率为 $1-p$ 。

4)用户上传数据会给他们带来一些效用收益,该效用收益与用户在社区内的等级以及数据成本 $G$ 有关,用户等级越高,说明对社区的贡献度越大,所获得的收益也就越多。该效用收益不一定是金钱收益,可以是上传数据获得的数字身份徽章或个人信誉值的提升等。本文效用收益简称成本收益,用 $B$ 表示。

5)假定 $y$ 为用户从数据共享中获得的收益,该收益包括但不限于社区积分的直接收益,也可以是利用共享数据解决现实问题而带来的间接收益,亦或是帮助科研人员解决某一项科研任务而获得的收益等。上述均属于共享收益,即从数据共享中获得的收益,此收益与数据成本 $G$ 有关,其缩放系数用 $\alpha$ 表示。

6)收益函数是指每个参与人在参与博弈时依据其所属类型和选择的策略可获得的收益。本文采取



对数收益函数  $\ln$  (以  $e$  为底数的对数函数) 来计算收益,该函数特征符合现实社会经济学特征,即收益是有一定增长抑制的,不会随着成本的增加而无限提高,最大收益总会趋近于某一阈值<sup>[22]</sup>。

7) 在数据共享社区中,为了鼓励用户共享数据,需要一定的外部激励奖励  $I$ ,该激励奖励以积分的形式奖励参与数据共享的用户。但是,该激励并不是一成不变的,原因是一成不变会导致激励达到饱和状态,从而引起社区内的积分通货膨胀,这不符合经济学原理,因此,该激励在一定阶段又会变成数据共享成本,简称激励/成本  $I$ 。

本文基于演化博弈论的数据共享激励模型 EGDSI 是一个二元组的形式,表示为  $M=(P,K)$ ,其中: $P$  表示由大量数据共享社区的用户所组成的一个群体; $K$  为策略空间,  $K=\{K_1, K_2\}$ 。用户有 2 种策略:一种是  $K_1$  (参与数据共享),即在数据共享社区中主动贡献出自己的数据,并可以获得其他共享用户的数据;另一种是  $K_2$  (不参与数据共享),该类用户既不会分享数据,也不会从社区中获取共享的数据。

根据上述假设,成本收益  $B$  的表达式如 (1) 所示:

$$B = l \cdot \ln(1 + G), l \in \{1, 2, 3, 4, 5\}, G > 0 \quad (1)$$

其中: $l$  为该用户在社区内的等级; $G$  为数据成本。根据现实中的合理性,成本收益  $B > 0$ , 即  $l \cdot \ln(1 + G) > 0$ ,  $l \in \{1, 2, 3, 4, 5\}$ ,  $\ln(1 + G)$  可保证收益为正。

共享收益  $y$  与数据成本  $G$  有关,缩放系数  $\alpha > 0$ ,  $y$  的表达式如 (2) 所示:

$$y = \alpha \cdot \ln(1 + G), \alpha > 0 \quad (2)$$

博弈双方的收益矩阵如表 1 所示,其中,上方为  $P_2$  的收益,下方为  $P_1$  的收益。

表 1 数据共享博弈收益矩阵

Table 1 Data sharing game revenue matrix

主体 $P_2$ 采取的 策略	主体 $P_1$ 采取的策略	
	共享	不共享
共享	$l \cdot \ln(1 + G) + \alpha \cdot \ln(1 + G) - I$ $l \cdot \ln(1 + G) + \alpha \cdot \ln(1 + G) - I$	$l \cdot \ln(1 + G) - I$ $l \cdot \ln(1 + G)$
不共享	$l \cdot \ln(1 + G)$ $l \cdot \ln(1 + G) - I$	$l \cdot \ln(1 + G)$ $l \cdot \ln(1 + G)$

收益矩阵的含义如下:

1) 情况 1。进行博弈的双方主体都选择  $K_1$  参与数据共享,这种是理想情况,也是共享社区最终的目标。社区内所有成员均共享数据,博弈双方不仅具有原本的成本收益  $B$ ,还会得到由数据共享所带来的共享收益  $y$ 。社区在初始阶段会给每位共享者相应的激励奖励,但是这种奖励不能一直扁平化激励下去,这样会导致数据共享平台中数据的积分价格上涨,引起通货膨胀。因此,需要引入一个激励/成本  $I$ ,此时应是数据共享的成本,即参与共享的用户需要支付少量的积分来进行数据共享,原因是此时已经达到激励饱和状态。

2) 情况 2。进行博弈的双方主体一个选择  $K_1$  参与数据共享,另一个选择  $K_2$  不参与数据共享。博弈双方均上传了数据,不同的是,一个选择共享数据,另一个选择不共享数据。首先,两者都会有数据本身带来的成本收益  $B$ ;其次,由于共享数据的主体  $P_1$  获取不到不共享主体  $P_2$  的数据,因此没有共享收益  $y$ ,需要给予共享主体一定的激励/成本  $I$ ,此时应是外部的奖励激励,即一定的积分奖励。同理,不共享主体由于未共享数据,因此获取不到共享主体所共享的数据,同样也没有共享收益  $y$ 。

3) 情况 3。进行博弈的双方主体均选择  $K_2$  不参与数据共享,此时,数据共享社区所面临的场景是没有人愿意共享数据,因此,主体双方都只能获得成本收益  $B$ 。

在 EGDSI 模型中,由于参与主体均是有限理性的,因此在博弈开始阶段不可能直接达到纳什均衡状态,需要一个演化模拟学习他人策略的过程,以进行动态调节并反复尝试,最终达到一种最优的平衡状态,实现整体利益的最大化。

## 1.2 演化稳定策略分析

在演化博弈论中,假设使用该策略的个体在群体中所占比例的增长率等价于群体中使用某个策略的个体适应度,只要一个策略的个体适应度比群体的平均适应度高,则使用该策略的个体数就会增加,且不会被其他策略入侵<sup>[23]</sup>。

由前文假设可知,在博弈的初级阶段  $t$ ,主体  $P_1$  参与数据共享的概率为  $p$ ,主体  $P_2$  不参与数据共享的概率为  $1 - p$ ,可等效为在  $t$  阶段博弈中,选择共享策略的人数比例为  $p_{(t)}$ ,选择不共享的人数比例为  $1 - p_{(t)}$ 。则在  $t$  阶段选择共享数据策略的期望收益为:

$$U_{(t)}^1 = p_{(t)} [l \cdot \ln(1 + G) + \alpha \cdot \ln(1 + G) - I] + (1 - p_{(t)}) [l \cdot \ln(1 + G) - I] \quad (3)$$

选择不共享数据策略的期望收益为:

$$U_{(t)}^2 = p_{(t)} [l \cdot \ln(1 + G)] + (1 - p_{(t)}) [l \cdot \ln(1 + G)] \quad (4)$$

数据共享社区的期望收益为:

$$\bar{U}_{(t)} = p_{(t)} \cdot U_{(t)}^1 + (1 - p_{(t)}) \cdot U_{(t)}^2 \quad (5)$$

因此,计算出复制动力学方程<sup>[24]</sup>为:

$$Y_{(p_{(t)})} = p_{(t)} \cdot (U_{(t)}^1 - \bar{U}_{(t)}) \quad (6)$$

化简得:

$$Y_{(p_{(t)})} = p_{(t)} \cdot (1 - p_{(t)}) \cdot ((\alpha \cdot \ln(1 + G)) p_{(t)} - I) \quad (7)$$

令  $Y_{(p_{(t)})} = 0$ , 可以求出动力学方程有 3 个演化稳态点,分别为:

$$p_1^* = 0 \quad (8)$$

$$p_2^* = 1 \quad (9)$$

$$p_3^* = \frac{I}{\alpha \cdot \ln(1 + G)} \quad (10)$$

上述演化稳定点在条件成立的情况下就是演化稳定策略 (Evolutionarily Stable Strategy, ESS) 的稳定状态。一种演化稳定策略一旦确立,就会稳定下来,偏离演化稳定策略的行为或个体将会被自然选择所淘汰。

如果一个点为稳定状态,那么一定满足  $Y_{(p^*)}$  的导数小于  $0^{[25]}$ ,即  $Y'_{(p^*)} < 0$ 。 $Y_{(p_0)}$  的导数  $Y'_{(p_0)}$  为:

$$Y'_{(p_0)} = -3(\alpha \cdot \ln(1+G)) \cdot (p_0)^2 + 2 \cdot [I + \alpha \cdot \ln(1+G)] p_0 - I \quad (11)$$

分别将  $p_0=0$ 、 $p_0=1$ 、 $p_0=\frac{I}{\alpha \cdot \ln(1+G)}$  代入式(11)可得:

$$Y'_{(0)} = -I \quad (12)$$

$$Y'_{(1)} = I - \alpha \cdot \ln(1+G) \quad (13)$$

$$Y'_{\left(\frac{I}{\alpha \cdot \ln(1+G)}\right)} = \frac{I(\alpha \cdot \ln(1+G) - I)}{\alpha \cdot \ln(1+G)} \quad (14)$$

演化稳态点  $p_3^*$  可能与  $p_1^*$ 、 $p_2^*$  值相同,最终会趋向于不参与数据共享  $p_1^*$  以及参与数据共享  $p_2^*$ 。因此,  $p_3^*$  需要分情况具体讨论:

1) 条件1。

假设  $0 < I < \alpha \cdot \ln(1+G)$ , 即激励/成本  $I$  小于共享收益  $y$ :

(1) 对于  $Y'_{(0)}$ , 由于  $I > 0$ , 则有  $-I < 0$ , 即  $Y'_{(0)} < 0$ , 因此  $p_1^*$  是数据共享演化博弈的稳定策略。

(2) 对于  $Y'_{(1)}$ , 由于  $I < \alpha \cdot \ln(1+G)$ , 则有  $I - \alpha \cdot \ln(1+G) < 0$ , 即  $Y'_{(1)} < 0$ , 因此  $p_2^*$  同样也是数据共享演化博弈的稳定策略。

(3) 对于  $Y'_{\left(\frac{I}{\alpha \cdot \ln(1+G)}\right)}$ , 由于  $0 < I < \alpha \cdot \ln(1+G)$ , 则  $\alpha \cdot \ln(1+G) - I > 0$ , 即  $Y'_{\left(\frac{I}{\alpha \cdot \ln(1+G)}\right)} > 0$ , 说明  $p_3^*$  不是演化稳定策略。

条件1说明:在  $0 < I < \alpha \cdot \ln(1+G)$  的初始条件下,演化稳定策略取决于初始主体参与共享的比例  $p$ 。

当  $0 < p < \frac{I}{\alpha \cdot \ln(1+G)}$  时,随着时间  $t$  的推移,种群不断模拟学习他人的策略,使自己的利益达到最大化,最终会导致种群趋于不进行数据共享,即达到  $p_1^*$  的演化稳定策略,整体系统的增益不足以支持种群转向数据共享策略。当主体参与数据共享的比例  $p$  满足  $\frac{I}{\alpha \cdot \ln(1+G)} < p < 1$  时,随着时间  $t$  的推移,进行演化模拟,最终种群会趋向于数据共享,即达到  $p_2^*$  的演化稳定策略。

2) 条件2。

假设  $0 < \alpha \cdot \ln(1+G) < I$ , 即激励/成本  $I$  大于共享收益  $y$ :

(1) 对于  $Y'_{(0)}$ , 由于  $I > 0$ , 则有  $-I < 0$ , 即  $Y'_{(0)} < 0$ , 因此  $p_1^*$  是数据共享演化博弈的稳定策略。

(2) 对于  $Y'_{(1)}$ , 由于  $\alpha \cdot \ln(1+G) < I$ , 则有  $I - \alpha \cdot \ln(1+G) > 0$ , 即  $Y'_{(1)} > 0$ , 因此  $p_2^*$  不是数据共享演化博弈的稳定策略。

(3) 对于  $Y'_{\left(\frac{I}{\alpha \cdot \ln(1+G)}\right)}$ , 由于  $0 < \alpha \cdot \ln(1+G) < I$ , 则  $\alpha \cdot \ln(1+G) - I < 0$ , 即  $Y'_{\left(\frac{I}{\alpha \cdot \ln(1+G)}\right)} < 0$ , 在现实场景中,要求进行数据共享的比例  $p$  应满足  $0 < p < 1$ , 而  $p_3^* =$

$\frac{I}{\alpha \cdot \ln(1+G)}$ , 显然已经大于1, 因此, 不符合社会学规律。

条件2说明:  $p_1^* = 0$  是唯一的演化稳定策略, 即无论初始种群比例  $p$  为多少, 随着时间的不断推移, 最终种群的演化稳定策略都会趋向于  $p_1^*$  不进行数据共享, 而这是在实际应用中应该避免出现的情况。

3) 条件3。

假设  $I < 0$  且  $\alpha \cdot \ln(1+G) > 0$ , 即激励/成本  $I$  小于0时, 此时应为数据共享激励奖励:

(1) 对于  $Y'_{(0)}$ , 由于  $I < 0$ , 则有  $-I > 0$ , 即  $Y'_{(0)} > 0$ , 因此  $p_1^*$  不是数据共享演化博弈的稳定策略。

(2) 对于  $Y'_{(1)}$ , 由于  $\alpha \cdot \ln(1+G) > 0$  且  $I < 0$ , 则有  $I - \alpha \cdot \ln(1+G) < 0$ , 即  $Y'_{(1)} < 0$ , 因此  $p_2^*$  是数据共享演化博弈的稳定策略。

(3) 对于  $Y'_{\left(\frac{I}{\alpha \cdot \ln(1+G)}\right)}$ , 由于  $\alpha \cdot \ln(1+G) > 0$ ,  $I < 0$ , 可得  $\alpha \cdot \ln(1+G) - I < 0$ , 即  $Y'_{\left(\frac{I}{\alpha \cdot \ln(1+G)}\right)} < 0$ , 但是在现实场景中, 要求进行数据共享的比例  $p$  应满足  $0 < p < 1$ , 而  $p_3^* = \frac{I}{\alpha \cdot \ln(1+G)}$ , 显然已经小于0, 因此, 不符合社会学规律。

条件3说明:  $p_2^* = 1$  是唯一的演化稳定策略, 即无论初始种群比例  $p$  为多少, 随着时间的不断推移, 最终种群的演化稳定策略都会趋向于  $p_2^*$  进行数据共享。

从条件1~条件3可以看出, 数据共享的演化稳定策略并不唯一。参与数据共享和不参与数据共享均具有演化潜力。EGDSI模型演化稳定策略的趋势如图1所示。

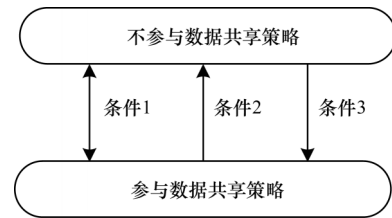


图1 EGDSI模型演化稳定策略趋势

Fig.1 Trend of EGDSI model evolutionarily stable strategy

为了促使更多人参与到数据共享中, 有必要根据数据共享程度对激励/成本  $I$  进行动态控制。本文设定参与数据共享的比例小于阈值  $p_3^*$  时为低数据共享程度, 此时需要使用条件3进行外部激励, 给予数据共享者一些积分奖励, 从而刺激社区用户选择共享策略; 当参与数据共享的比例大于阈值  $p_3^*$  时为高数据共享程度, 此时需使用条件1收取相应的共享成本, 防止通货膨胀, 并且由上述条件1的分析可知, 这时仍可以确保数据共享社区在参与共享策略方面能够维持而无需外部激励。

社区用户在每一个博弈阶段都对应一个参与成本, 这会导致每一阶段的阈值  $p_3^*$  均不相同, 当共享人数比例小于阈值  $p_3^*$  时使用条件3进行激励, 当共享

人数比例大于阈值  $p_3^*$  时使用条件 1 收取成本,直到社区用户共享程度达到饱和,进行下一阶段博弈,重新计算阈值  $p_3^*$ 。如此循环往复,体现了社区成员相互博弈的过程,并从中获得由数据共享所带来的收益。博弈阶段流程如图 2 所示。

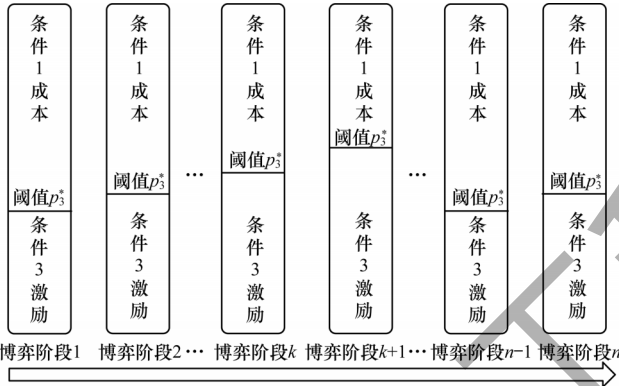


图 2 社区成员博弈阶段流程  
Fig.2 Community members game stage process

2 基于智能合约的数据共享模板引擎

数据共享模板引擎(Data Sharing Incentive Template Engine, DSITE)将EGDSI模型以区块链智能合约的形式在数据共享社区中进行应用。智能合约是可以在区块链上自动执行的特殊程序,其程序代码及数据均存储在链上,具有防篡改等特性,所有节点必须按照智能合约预置的触发条件和相应规则进行计算。基于上述特性,可以在数据共享社区中将数据共享激励机制写入智能合约,用户在社区内进行数据共享时,会自动触发执行数据共享模板引擎进行共享激励,从而为数据共享激励提供可信、安全的保障。

由于EGDSI模型假定所有参与博弈方的本质都是相同的,因此与所有博弈方相关的参数也应相同。参数设定如表 2 所示。

表 2 参数设定		
Table 2 Parameters setting		
DSITE 变量名称	EGDSI 对应含义	解释
DataCost	$G$	数据成本
ZoomFactor	$\alpha$	数据共享收益缩放系数
Cost[]	激励/成本 $I$ , 这里指成本	数据共享成本参数序列, 是一个一维数组, 数组中每一个值即为不同博弈阶段所对应的共享成本
Incentive Reward	激励/成本 $I$ , 这里指激励	数据共享积分奖励
SaturationRate	$Q_y^*$	数据共享饱和率

数据共享激励的计算过程如下:

**步骤 1** 初始化全局参数信息,包括数据成本 Data Cost、数据共享收益缩放系数 ZoomFactor、数据共享成本参数序列 Cost[]、数据共享的积分激励

Incentive Reward。

**步骤 2** 根据当前博弈阶段  $t$  得到数据共享成本参数 Cost[ $t\%m$ ],其中,  $m$  为数据共享成本参数序列数组长度,计算出该成本下的阈值  $p_3^*$ 。

**步骤 3** 计算出当前时间参与数据共享人数占数据共享社区总人数的比例  $p_t$ ,将  $p_t$  与阈值  $p_3^*$  进行对比,若  $p_t < p_3^*$ ,则返回数据共享积分奖励 Incentive Reward;反之,进入步骤 4。

**步骤 4** 返回相应的数据共享成本 Cost[ $t\%m$ ],同时判断当前共享比例  $p_t$  与数据共享饱和率 SaturationRate 大小,若  $p_t > \text{SaturationRate}$ ,则进入下一阶段博弈,博弈阶段  $t$  自增,数据共享参与人数重置,此阶段博弈结束;反之,则重复步骤 1,继续进行当前阶段的演化博弈。

```
算法 1 DSITE 智能合约-数据共享激励算法
输入
Data cost of uploaded data, DataCost//数据成本
Data sharing cost parameter sequence, Cost[]//数据共享
//成本序列
Data sharing incentives, Incentive Reward//数据共享积分
//奖励
Data sharing saturation rate, SaturationRate//数据共享饱
//和率
Cost index used in the current stage, t//博弈轮次
输出
Cost or incentive, res//成本或激励
1. p ← ParticipateDataSharingCnt/UsersCnt //计算数据共
//享人数比例
2. m ← GetTheLengthOfTheCostArray(Cost[])//计算 Cost 数
//据共享成本序列长度
3. p3 ← Cost[t%m]/(ZoomFactor · ln(1 + DataCost))//计算阈
//值 p3
4. if p < p3 then
5. res ← Incentive Reward//返回激励奖励
6. else
7. if p > SaturationRate then //数据共享人数比例大于饱和率
8. t ← t + 1//博弈轮次+1
9. ParticipateDataSharingCnt ← 0//重置共享人数
10. end if
11. res ← Cost[t%m]//返回当前博弈阶段的共享成本
12. end if
13. return res
```

算法 1 解释说明:当数据共享社区用户参与数据共享时,会自动触发上述智能合约,调用其方法。首先,计算参与共享的用户比例,通过数组寻找当前博弈阶段下的成本并计算该成本下的阈值,如果参与共享的用户比例小于阈值,则返回相应的激励;如果大于阈值,再判断当前共享用户比例是否大于饱和率,本文设定饱和率为 98%,如果达到了当前的社区数据共享饱和率,则进入下一阶段的演化博弈,重置相应的参数,最后返回相应的数据共享成本。值得注意的是,计算激励/成本是一个取模运算,这是因为需要动态调整激励/



成本,以重置社区用户博弈阶段,而数组本身不可能无穷大,因此采用取模运算,这也体现了社区成员数据共享相互博弈的一个过程。

3 实验验证

本节通过控制变量法设计不同的横向与纵向对比实验,分析EGDSI数据共享模型中的3个演化稳定策略条件,验证数据共享激励模型对社区演化稳定所起的作用,并将EGDSI模型与其他数据共享演化博弈模型进行对比。

3.1 演化稳定策略验证分析

1.2节通过演化博弈论求解演化稳定策略,推导出条件1~条件3。本节通过模拟实验验证其结论的正确性。为满足条件1~条件3,对条件参数进行设置,具体如表3所示。

表3 条件参数设置			
Table 3 Condition parameters setting			
条件	缩放系数 $\alpha$	数据成本 $G$	激励/成本 $I$
条件1	4	10	4
条件2	4	10	12
条件3	4	10	-4

通过Matlab进行演化模拟仿真,针对条件1,其演化曲线如图3所示。

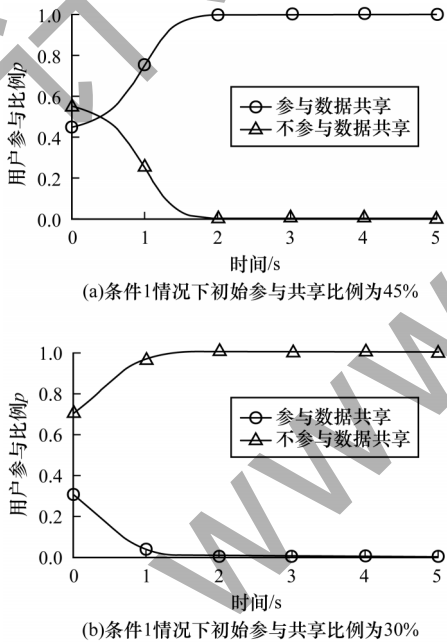


图3 条件1的演化曲线  
Fig.3 Evolution curves of condition 1

共享收益缩放系数 $\alpha=4$ ,数据成本 $G=10$ ,激励/成本 $I=4$ ,可得演化稳定点 $p_3^* = \frac{I}{\alpha \cdot \ln(1+G)} \approx 0.42$ 。如果社区成员进行数据共享的初始比例 $p$ 满足 $0.42 < p < 1$ ,随着时间的推移,数据共享社区的演化稳定策略会趋向于参与数据共享;反之,如果 $p$ 满足

$0 < p < 0.42$ ,那么数据共享社区的演化稳定策略最终会趋向于不参与数据共享。由图3(a)可知,社区成员进行数据共享的初始比例 $p=0.45$ (大于0.42),即45%的社区成员选择了初始参与数据共享,随着时间的推移,最终会达到 $p_2^*=1$ 的演化稳定策略。同理,由图3(b)可知,社区成员进行数据共享的初始比例为 $p=0.30$ (小于0.42),即30%的社区成员选择了初始参与数据共享,随着时间的推移,整个数据共享社区最终会达到 $p_1^*=0$ 的演化稳定策略,进行数据共享的人数比例会越来越小,最终整个数据共享社区趋向于不进行数据共享。上述实验结果表明,在条件1的情况下, $p_1^*=0$ 和 $p_2^*=1$ 是数据共享社区的演化稳定策略。

针对条件2与条件3,可得演化曲线如图4所示。当设定共享收益缩放系数 $\alpha=4$ 、数据成本 $G=10$ 、激励/成本 $I=12$ 时,满足条件2,即 $0 < \alpha \cdot \ln(1+G) < I$ 。由图4(a)可知,当社区成员进行数据共享的初始比例 $p=0.80$ 时,即80%的社区成员选择了初始参与数据共享,但最终整个系统的演化稳定策略仍趋向于 $p_1^*=0$ ,即不进行数据共享。实验结果表明,在条件2的情况下, $p_1^*=0$ 是数据共享社区的演化稳定策略。当设定共享收益缩放系数 $\alpha=4$ 、数据成本 $G=10$ 、激励/成本 $I=-4$ 时,满足条件3,即 $I < 0$ 且 $\alpha \cdot \ln(1+G) > 0$ ,由图4(b)可知,当社区成员进行数据共享的初始比例 $p=0.05$ 时,即仅5%的社区成员选择了初始参与数据共享,但随着时间的推移,最终整个系统的演化稳定策略仍趋向于 $p_2^*=1$ ,即参与数据共享。实验结果表明,在条件3的情况下, $p_2^*=1$ 是数据共享社区的演化稳定策略。

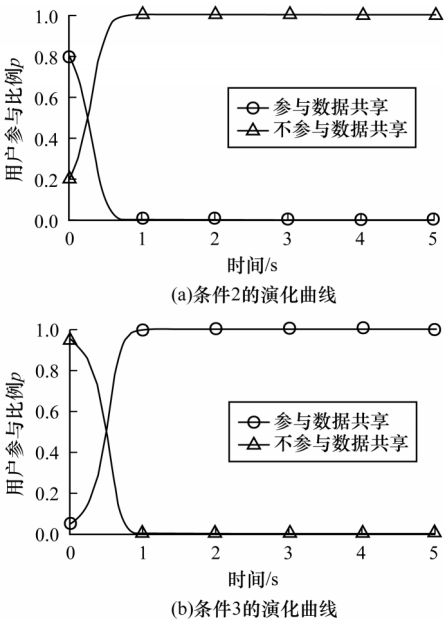


图4 条件2与条件3的演化曲线

Fig.4 Evolution curves of condition 2 and condition 3

3.2 激励机制方案验证分析

为了证明激励机制在用户选择决策时是否发挥促进数据共享的作用,设定共享收益缩放系数 $\alpha=4$ ,数据成本 $G=10$ ,在条件1中激励/成本 $I$ 的取值范围为 $I<\alpha\cdot\ln(1+G)$ ,即 $I<9.6$ ,在条件3中, $I$ 的取值范围为 $I<0$ 。本次实验通过控制变量法,在相同数据共享初始比例的情况下调整不同的激励/成本 $I$ ,以及在相同的激励/成本 $I$ 下调整不同的数据共享初始比例,条件参数设置如表4所示。

表4 激励机制条件参数设置

Table 4 Incentive mechanism condition parameters setting

数据	激励/成本 $I$	共享初始 比例 $p$	数据	激励/成本 $I$	共享初始 比例 $p$
数据1	-8	0.1	数据11	2	0.1
数据2	-8	0.3	数据12	2	0.3
数据3	-8	0.5	数据13	2	0.5
数据4	-8	0.7	数据14	2	0.7
数据5	-8	0.9	数据15	2	0.9
数据6	-3	0.1	数据16	7	0.1
数据7	-3	0.3	数据17	7	0.3
数据8	-3	0.5	数据18	7	0.5
数据9	-3	0.7	数据19	7	0.7
数据10	-3	0.9	数据20	7	0.9

使用上述参数可得EGDSI演化曲线如图5所示。由条件1可知,不同的激励/成本 $I$ 会导致不同的数据共享初始比例阈值 $p_3^*$ ,以 $I=2$ 为例,当激励/成本 $I=2$ 时,数据共享初始比例阈值(即演化稳定点) $p_3^*=\frac{I}{\alpha\cdot\ln(1+G)}\approx 0.21$ ,当数据共享初始比例 $p=0.1$ 时,由图5可知,随着时间的推移,社区成员相互进行模仿学习,不断调整自己的策略,使自己的利益最大化,最终会达到 $p_1^*=0$ 的演化稳定策略,即趋向于不进行数据共享。为了避免此类现象在数据共享社区出现,此时数据共享社区应给予一定的激励奖励,即调整激励/成本 $I$ ,使 $I=-3$ 或 $-8$ 。当激励/成本 $I$ 为负值时,数据共享的比例随着时间推移而增加,直到增加至 $p=0.3$ ,可再次调整激励/成本 $I=2$ ,收取一定的成本,因为此时数据共享社区随着时间的推移一定会趋向于数据共享稳定策略。当 $p$ 达到0.9时,可增加激励/成本 $I$ ,将其增至7,防止进一步通货膨胀,提高共享成本。综上所述,本文所设计的数据共享激励机制可以有效促进更多用户参与数据共享。

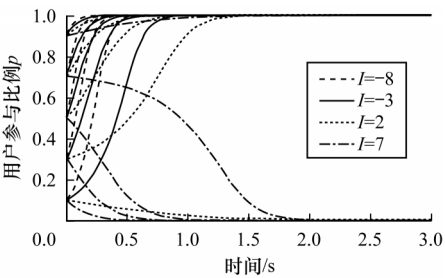


图5 EGDSI激励模型演化曲线

Fig.5 Evolution curves of EGDSI excitation model

3.3 模型对比

本文提出一种基于演化博弈论的数据共享激励模型EGDSI,目的是为了进一步促进数据的收集、共享与流通。文献[19]针对数据共享激励提出一种EGI模型,EGI模型对共享成员的等级未进行划分,没有考虑个体之间奖励的差异性,所有共享成员均采用相同的激励机制对数据共享进行奖励。本节对这2种模型进行对比,设定共享收益缩放系数 $\alpha=4$ ,数据成本 $G=10$ ,2种模型的数据共享演化曲线对比如图6所示。

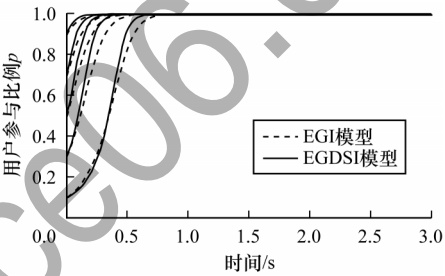


图6 2种模型的数据共享演化曲线对比

Fig.6 Comparison of data sharing evolution curves of two models

由图6可知,在不同的初始共享比例情况下,本文模型达到数据共享饱和状态的速度均优于EGI模型,即本文模型可以在更早的博弈阶段达到数据共享,提升数据共享的效率,其具有更好的激励作用。

4 结束语

数据共享社区普遍面临用户参与积极性低、归属感弱、缺乏相应数据共享激励机制等问题,为此,本文构建一种基于演化博弈论的数据共享激励模型,并在已有的区块链共享平台上利用智能合约技术实现该模型所对应的智能合约数据共享激励引擎。仿真结果表明,该模型可以在数据共享社区中激励更多用户参与数据共享,进一步促进数据的收集、流通和使用。下一步将结合可信数据共享与安全存储相关技术,提升数据在共享过程中的访问权限控制,防止数据非授权访问等风险发生,同时提高数据在共享社区存储时的安全性与隐私性。

## 参考文献

- [1] MUKHERJEE M, MATAM R, SHU L, et al. Security and privacy in fog computing: challenges[J]. IEEE Access, 2017, 5: 19293-19304.
- [2] 毕达天, 曹冉, 杜小民. 科学数据共享研究现状与展望[J]. 图书情报工作, 2019, 63(24): 69-77.  
BI D T, CAO R, DU X M. A summary of researches on research data sharing and development trends at home and abroad[J]. Library and Information Service, 2019, 63(24): 69-77. (in Chinese)
- [3] 胥婷, 于广军. 国内健康医疗数据开放共享现状研究[J]. 中国数字医学, 2020, 15(9): 64-66, 103.  
XU T, YU G J. Research on the current situation of open access and sharing of health and medical data in China[J]. China Digital Medicine, 2020, 15(9): 64-66, 103. (in Chinese)
- [4] WANG Z, TIAN Y L, ZHU J M. Data sharing and tracing scheme based on blockchain[C]//Proceedings of the 8th International Conference on Logistics, Informatics and Service Sciences. Washington D. C., USA: IEEE Press, 2018: 1-6.
- [5] NOFER M, GOMBER P, HINZ O, et al. Blockchain[J]. Business & Information Systems Engineering, 2017, 59(3): 183-187.
- [6] 张亮, 刘百祥, 张如意, 等. 区块链技术综述[J]. 计算机工程, 2019, 45(5): 1-12.  
ZHANG L, LIU B X, ZHANG R Y, et al. Overview of blockchain technology[J]. Computer Engineering, 2019, 45(5): 1-12. (in Chinese)
- [7] 庄倩, 何琳. 科学数据共享中科研人员共享行为的演化博弈分析[J]. 情报杂志, 2015, 34(8): 152-157, 168.  
ZHUANG Q, HE L. Analysis of scientific researchers sharing behavior based on evolutionary game theory[J]. Journal of Intelligence, 2015, 34(8): 152-157, 168. (in Chinese)
- [8] WEN Y T, SHI J Y, ZHANG Q, et al. Quality-driven auction-based incentive mechanism for mobile crowd sensing[J]. IEEE Transactions on Vehicular Technology, 2015, 64(9): 4203-4214.
- [9] 郝世博, 徐文哲, 唐正韵. 科学数据共享区块链模型及实现机理研究[J]. 情报理论与实践, 2018, 41(11): 57-62.  
HAO S B, XU W Z, TANG Z Y. Research on blockchain model and implementation mechanism of scientific data sharing[J]. Information Studies: Theory & Application, 2018, 41(11): 57-62. (in Chinese)
- [10] 石进, 邵波, 苗杰. 基于区块链的中小企业竞争情报共享平台研究[J]. 图书情报工作, 2019, 63(20): 112-120.  
SHI J, SHAO B, MIAO J. Research on competitive intelligence sharing platform for SME based on blockchain[J]. Library and Information Service, 2019, 63(20): 112-120. (in Chinese)
- [11] 张宝, 王梦寒. 基于区块链技术的数据共享新模式[J]. 当代金融家, 2019(10): 121-123.  
ZHANG B, WANG M H. A new data sharing model based on blockchain technology[J]. Modern Bankers, 2019(10): 121-123. (in Chinese)
- [12] 张旺, 程慧平. 科学数据开放共享策略机制及优化路径研究[J]. 情报杂志, 2020, 39(5): 154-161.  
ZHANG W, CHENG H P. Research on open sharing strategy mechanism and optimization path of scientific data[J]. Journal of Intelligence, 2020, 39(5): 154-161. (in Chinese)
- [13] GAN C Q, SAINI A, ZHU Q Y, et al. Blockchain-based access control scheme with incentive mechanism for eHealth systems: patient as supervisor[J]. Multimedia Tools and Applications, 2021, 80(20): 30605-30621.
- [14] SHEN M, DUAN J X, ZHU L H, et al. Blockchain-based incentives for secure and collaborative data sharing in multiple clouds[J]. IEEE Journal on Selected Areas in Communications, 2020, 38(6): 1229-1241.
- [15] SHRESTHA A K, VASSILEVA J, DETERS R. A blockchain platform for user data sharing ensuring user control and incentives[J]. Frontiers in Blockchain, 2020, 3: 497985.
- [16] LI F, FU Y C, ZHAO P C, et al. An incentive mechanism for nondeterministic vehicular crowdsensing with blockchain[C]//Proceedings of 2020 IEEE/CIC International Conference on Communications in China. Washington D. C., USA: IEEE Press, 2020: 1074-1079.
- [17] 梁小燕. 区块链环境下基于信息熵的医疗数据共享激励机制[D]. 兰州: 西北师范大学, 2020.  
LIANG X Y. Incentive mechanism of medical data sharing based on information entropy in blockchain environment[D]. Lanzhou: Northwest Normal University, 2020. (in Chinese)
- [18] 蔡婷, 林晖, 陈武辉, 等. 区块链赋能的高效物联网数据激励共享方案[J]. 软件学报, 2021, 32(4): 953-972.  
CAI T, LIN H, CHEN W H, et al. Efficient blockchain-empowered data sharing incentive scheme for Internet of Things[J]. Journal of Software, 2021, 32(4): 953-972. (in Chinese)
- [19] 王倩, 徐尚英, 陈冬林, 等. 面向科技服务数据共享策略的智能合约激励机制研究[J]. 情报杂志, 2021, 40(10): 157-165.  
WANG Q, XU S Y, CHEN D L, et al. Research on smart contract incentive mechanism oriented to sci-tech services data sharing strategy[J]. Journal of Intelligence, 2021, 40(10): 157-165. (in Chinese)
- [20] COLMAN A M. Cooperation, psychological game theory, and limitations of rationality in social interaction[J]. The Behavioral and Brain Sciences, 2003, 26(2): 139-153.
- [21] 朱雅敏, 魏永波. 完全理性到有限理性: 博弈论理性基础的变更[J]. 科技视界, 2015(27): 165-166.  
ZHU Y M, WEI Y B. Complete rationality to bounded rationality: the change of the rational basis of game theory[J]. Science and Technology Vision, 2015(27): 165-166. (in Chinese)
- [22] SENGUPTA S, CHATTERJEE M. An economic framework for dynamic spectrum access and service pricing[J]. ACM Transactions on Networking, 2009, 17(4): 1200-1213.
- [23] 张维迎. 博弈论与信息经济学[M]. 上海: 上海人民出版社, 2004.  
ZHANG W Y. Game theory and information economics[M]. Shanghai: Shanghai People's Publishing House, 2004. (in Chinese)
- [24] FUDENBERG D, DREW F, LEVINE D K, et al. The theory of learning in games[M]. Cambridge, USA: MIT press, 1998.
- [25] FRIEDMAN D. Evolutionary games in economics[J]. Econometrica, 1991, 59(3): 637-666.