基于特殊个体的网络进化博弈群智合作激发调控

王浩淼, 魏晨*, 邓亦敏

北京航空航天大学自动化科学与电气工程学院, 飞行器控制一体化技术重点实验室仿生自主飞行系统研究组, 北京 100083

* E-mail: weichen@buaa.edu.cn

收稿日期: 2021-07-01; 接受日期: 2021-10-14; 网络版发表日期: 2022-08-09

科技创新2030-“新一代人工智能”重大项目(批准号: 2018AAA0102403)和国家自然科学基金项目(批准号: U20B2071, 91948204, U1913602, U19B2033)资助

摘要 引导群体中智能行为的产生与传播是群体智能领域的重要研究方向. 相互合作是自然界中常见的群体智能行为. 本文建立了一种基于特殊个体的网络进化博弈模型, 分析了群体激发合作水平的进化特性以及调控手段. 首先通过在规则网络进化博弈模型中建立有效的干预控制体系结构, 实现对群集系统的调控作用; 然后通过设计基于群体智能的特殊个体控制规则, 特殊个体能起到促进合作行为涌现的群智激发作用; 最后利用蒙特卡罗仿真方法分析了特殊个体对群体进化特性的影响, 并讨论了不同参数条件对合作水平的影响.

关键词 特殊个体, 网络进化博弈, 软控制, 群智合作激发, 蒙特卡罗仿真

1 引言

近年来, 群体智能的有关研究得到迅速发展, 有关合作现象产生与维持的研究一直以来都受到了不同领域学者的普遍关注. 合作与竞争是自然界中常见的博弈关系, 与完全不合作者“坐享其成”相比, 合作者需要承担额外的代价, 是一种对个体生存繁衍不利的行为. 然而事实上, 合作行为在绝大多数群体中都得到广泛的应用, 如自然界中狼群围捕猎物、工蚁寻找食物以及人类社会中的种种合作行为等, 这表明合作行为具有其独特的优势机制, 体现了群体层面涌现出的智能性^[1,2].

博弈论是研究群体智能中多个利益主体间的冲突和决策而产生的一门学科, 将生物进化理论与经典博

弈论相结合, 得到了蕴含生物学中自然选择及进化机制的网络进化博弈. 与传统进化博弈相比, 网络进化博弈中的个体是不完全理性的, 通过不断试错和调整策略达到群体的博弈平衡. 1992年, Nowak与May^[3]发表在*Nature*上的文章开启了网络进化博弈研究的先河, 为揭示群体智能中合作行为涌现与维持机制提供了新的研究框架. 以此为起点, 学者们对博弈论中影响合作行为的因素进行了广泛而深入的研究. 在网络结构的研究中, Realpe-Gómez等人^[4]在正则网络以及具有固定连接关系的复杂网络上研究了进化动力学和社会规则对合作行为的影响. 在重复博弈机制的研究中, Hong等人^[5]研究了囚徒博弈中互惠机制对合作行为的促进作用. Mcavoy等人^[6]研究了重复博弈机制下不同策略的收益及其对群体行为的影响. Press和Dyson^[7]

引用格式: 王浩淼, 魏晨, 邓亦敏. 基于特殊个体的网络进化博弈群智合作激发调控. 中国科学: 技术科学, 2023, 53: 167–176

Wang H M, Wei C, Deng Y M. Excitation and control of group intelligence cooperation in network evolutionary game based on special individuals (in Chinese). Sci Sin Tech, 2023, 53: 167–176, doi: 10.1360/SST-2021-0265

提出了零行列式(zero-determinant, ZD)策略, 能够在重复博弈中单方面控制博弈结果. 在策略更新规则的研究中, Szolnoki等人^[8]提出了个体策略更新时的渐近学习规则, 并与传统的模仿规则进行对比. Nadini等人^[9]研究了累计收益对个体合作倾向的影响. Shu等人^[10]研究了具有记忆机制的囚徒博弈和雪堆博弈模型, 个体根据历史记忆信息和邻居信息通过异步算法进行合作演化. 大量研究表明, 空间结构、博弈规则、个体策略选择对合作行为具有较大影响. 然而大部分研究工作主要集中于如何构建博弈模型、何种因素更有利于合作现象的涌现与维持这一问题, 对群体演化效果的调控机制研究较少. 在博弈模型的基础上, 设计一种具有普适性的群智调控方法, 具有一定的研究价值.

引导智能行为产生是群体智能的重要研究领域^[11]. 在多智能体系统中, 难以通过设计所有个体规则以促进群体智能行为涌现. 对于这一问题, Han等人^[12]提出了群体行为干预的框架——软控制(soft control), 旨在通过向原始系统中引入一定数量的可控特殊个体, 达到使复杂系统朝着期望状态演化的目的. Han等人^[13,14]进一步研究了DeGroot模型中引入可控个体对群体演化效果的影响, 并在囚徒博弈模型中进行了验证. Sartoretto^[15]研究了在群体中引入领导者的作用, 并进行动力学分析. 软控制的核心思想在于通过引入可控个体的方法以干预系统, 在不破坏群体现有规则的前提下控制整个系统. 由于群集系统中个体的同质性, 增加或部分个体不会影响整个系统的性质. 因此对群集系统施加软控制可以使系统向着预定的目标演化. 软控制的提出为群集系统群智合作行为的研究提供了新方法, 也为群体智能的激发提供了新思路.

本文在规则网络进化博弈中建立了一种有效的干预控制体系结构, 通过引入可控个体并设计相应的控制规则, 利用蒙特卡罗仿真与统计分析工具, 研究基于群体智能更新规则的特殊个体对群体进化特性的影响. 研究结果不仅可揭示特殊个体对群体合作水平的激发效果, 也可对复杂系统中的合作行为干预控制提供理论基础.

2 网络进化博弈模型

2.1 空间网络模型

在二维正则晶格网络中, 个体间具有稳定、规

则、固定的连接关系, 是一种简单实用的空间网络结构, 因此选取具有周期性边界限制条件的 $L \times L$ 二维正则晶格作为基本网络结构. 晶格网络每一个节点上都有一个博弈参与者占据, 个体按照二维晶格结构与最近的四个邻居分别进行交互, 网络结构具有周期重复性. 利用二维坐标表示每一个节点的位置, 对网络中全体节点进行编号, 任意节点 (x, y) , 其编号 $n_{x,y}$ 及邻居集合neigh表示如下:

$$\begin{cases} n_{x,y} = x + (y-1)L, \\ \text{neigh} = \{(x, y+1), (x+1, y), (x, y-1), (x-1, y)\}, \end{cases} \quad (1)$$

其中, 节点编号 $n_{x,y}$ 取值范围为 $1 \sim L^2$, 是对晶格网络上所有个体进行排序操作; 当邻居坐标计算结果超出周期边界限制条件时, 需要加减一个周期长度, 即

$$\begin{aligned} x \pm 1 &= \begin{cases} x+1, & x+1 \leq L, \\ x+1-L, & x+1 > L, \end{cases} \\ x \mp 1 &= \begin{cases} x-1, & x-1 > 0, \\ x-1+L, & x-1 \leq 0. \end{cases} \end{aligned} \quad (2)$$

2.2 博弈策略与博弈收益

选取重复囚徒困境博弈(prisoner's dilemma game, PDG)作为基本博弈模型. 每个博弈参与者与邻居分别进行 P 次囚徒困境博弈, P 为重复博弈长度. 单次博弈的支付矩阵如表1所示, 其中甲、乙表示两个博弈参与者, b 表示博弈强度.

定义个体的博弈策略为平稳马尔可夫策略 $x=(p_0, p_c, p_d)$ ^[16]. 其中 p_0, p_c, p_d 表示每一次博弈中的合作概率, 均在 $[0,1]$ 区间内随机取值. p_0 为个体在一轮重复博弈中的初次交互时选择合作的概率, p_c 与 p_d 分别表示在上一次博弈中对手选择合作策略 C 或背叛策略 D 时个体在当前轮次中选择合作的条件概率, 马尔可夫策略表明双方的当前决策仅取决于对手在上次博弈中的表现, 而与双方历史表现无关. 一轮博弈的流程如图1所示.

表1 单次囚徒困境博弈支付函数

Table 1 Payment function of a single prisoner's dilemma game

甲/乙	C	D
C	(1, 1)	(-b, 1+b)
D	(1+b, -b)	(0, 0)

图 1 单轮重复博弈流程

Figure 1 Flow of a single round repeated game.

其中, $X_i(t)$ 表示个体 i 在第 t 轮重复博弈中采取的策略, $W_i(n)$ 表示个体 i 在同一轮博弈中第 n 次交互时采取的行为. 在同一轮博弈中, 个体的博弈策略不变, 每一次交互的行为依概率选取合作或背叛, 并且每一次的交互结果将对下一次博弈产生影响. 一轮博弈结束后, 个体根据自身策略更新规则更新策略.

在博弈模型中, 每个个体占据网络中的一个节点, 在博弈过程中与周围相邻的四个邻居进行博弈, 每次博弈的收益进行累计, 并以收益之和作为博弈过程中的支付值, 博弈收益的定义如下:

$$\Pi_i(t) = \sum_{j \in \Omega_i} \pi_{ij}(t), \quad (3)$$

其中, Ω_i 为个体 i 的邻居集合; $\pi_{ij}(t)$ 为个体 i 与邻居 j 进行 P 次重复囚徒博弈后的累计收益, 而 $\Pi_i(t)$ 则为在第 t 轮博弈过程中个体 i 的总收益值. 定义个体单次博弈行为收益函数如下:

$$R = \Pi_i(t) / 4P. \quad (4)$$

合作频率(合作水平) f_c 定义为 t 时刻所有个体在一轮重复博弈中累计合作行为所占的比例. 对于 N 个个体共存在 $4PN$ 个单独的行动, 所以合作频率计算公式

如下:

$$f_c = \frac{1}{4PN} \sum_{i=1}^N w(i), \quad (5)$$

其中, $w(i)$ 表示个体 i 在与其四个邻居进行一轮重复博弈的过程中累计采取合作行为的次数.

值得注意的是, 平均合作比例与平均单次收益具有相同的进化特性, 这一结果可作如下证明: 对于单次囚徒困境, 令博弈双方均为混合策略 (x, y) , 且 $x+y=1$, 其中 x 为合作概率, y 为背叛概率, 则个体的期望收益:

$$\begin{aligned} E(R) &= x(1 \times x - by) + y[(1+b)x + 0 \times y^2] \\ &= x^2 - bxy + bxy + xy \\ &= x(x+y) \\ &= x. \end{aligned} \quad (6)$$

因此在后续仿真中, 用平均合作频率 f_c 代替平均单次合作收益 $\langle R \rangle$.

2.3 个体策略更新规则

在网络进化博弈模型中, 普通个体以“无条件模仿”规则进行策略更新, 即在每一轮策略更新时刻 t , 普通个体选择其邻域中当前单轮博弈收益最大的个体所

采用的马尔可夫策略作为自身下一时刻 $t+1$ 的博弈策略, 即

$$\begin{aligned} x_i(t+1) &= x_{j^*}(t), \\ j^* &= \operatorname{argmax}_{j \in \Omega_i} \Pi_j, \end{aligned} \quad (7)$$

其中, $\operatorname{argmax}_{j \in \Omega_i} \Pi_j$ 为邻域中收益最高个体. 如果最优个体数量大于1, 选择其中编号最小的邻居个体作为模仿对象.

在基本博弈模型的基础上, 还需填入能实现调控作用的特殊个体. 普通个体与特殊个体占比分别为 p_n 与 p_s , 且 $p_n+p_s=1$, $p_n > p_s$. 引入特殊个体能够体现出普通个体不具备的群体智能的特性, 从而促进系统合作水平的提高. 文献[17]中提出了一种基于粒子群优化 (particle swarm optimization, PSO) 的特殊个体设计规则, 揭示了特殊个体存在时的合作进化特性. 在此基础上有学者提出了基于PSO的群智激发策略更新规则, 并探讨了其对混合策略、合作行为的影响^[18,19]. 众多机制下的策略更新规则体现了群体智能的激发作用, 但对群体智能的设计规则较为简单, 并且缺少对特殊个体产生机理的研究. 本文在文献[17]的基础上提出了一种基于邻居信息反馈的优化策略更新规则, 能够实现对群体进化特性的调控, 体现群智合作激发作用.

3 群智激发下的特殊个体控制策略

与传统的分布式控制不同, 基于特殊个体的干预控制仅对群体中的少数个体施加控制作用, 而非对所有个体的行为进行直接干预^[20]. 并且不需要针对特殊个体设计特定的选取规则, 特殊个体从完全同质的普通个体中产生, 这体现出了由部分个体行为扩散到群体智能的特性, 有助于调控群体朝着期望的目标演化. 特殊个体与普通个体均遵守相同的博弈规则, 但也存在着显著的差别: 普通个体仅能模仿上一轮对手博弈策略, 特殊个体具有全局记忆的能力, 能够实时掌握邻居反馈信息, 并综合邻居信息和历史信息产生新的策略. 为了体现特殊个体调控效果不依赖于初始状态的影响, 特殊个体与普通个体具有完全相同的空间地位, 在网络中随机分布.

个体 i 在第 t 轮重复博弈中的马尔可夫策略为 $x_i(t) = (p_0, p_c, p_d)$, 每一特殊个体等效为在三维空间 $R[0,1]^3$ 进行随机搜索的粒子, 其在自身历史中得到最大收益时

采取的策略为 $x_i^L(t)$, 定义为历史最优策略; 当前邻居中收益最高的策略为 $x_i^U(t)$, 定义为邻居最优策略. 在前半部分迭代过程中, 特殊个体 i 在综合考虑自身历史最优信息以及邻居信息的基础上, 按照下式更新当前策略:

$$\begin{cases} x_i(t+1) = x_i(t) + v_i(t+1), \\ v_i(t+1) = v_i(t)e^{-Rt} + \omega_1(x_i^L(t) - x_i(t)) \\ \quad + (1 - \omega_1)(x_i^U(t) - x_i(t)), \end{cases} \quad (8)$$

其中, ω_1 为调节个体自身历史信息与邻居反馈信息的权值, 定义为信息权重. 当 $\omega_1 \rightarrow 0$ 时, 个体倾向于跟随邻居中最优个体的策略; 当 $\omega_1 \rightarrow 1$ 时, 个体在主要从自身历史信息中选取最优策略, ω_1 越大, 邻居反馈作用对特殊个体的影响越强. R 为衰减系数, 表明随着迭代过程的进行, 特殊个体的策略调整幅度逐渐降低, 这里取 $R=0.3$.

在后半部分迭代过程中, 普通个体基本达到了最优策略, 而特殊个体由于继续搜索新策略, 反而会影响整体合作水平. 此时邻居反馈信息成为特殊个体策略更新的主要影响因素, 调整特殊个体的策略更新规则, 选取邻居中收益高于自身的个体, 定义为优势邻居个体, 综合优势邻居平均策略和自身当前策略, 按照下式更新策略:

$$\begin{cases} x_i^{U-\text{avg}}(t) = \frac{\sum_{j=1}^{N_\Omega^*} x_j^{N_\Omega^*}(t)}{N_\Omega^*}, \\ x_i(t+1) = \omega_2 \cdot x_i(t) + (1 - \omega_2)(x_i^{U-\text{avg}}(t) - x_i(t)), \end{cases} \quad (9)$$

其中, ω_2 为调节邻域平均策略与个体当前策略的权值, $x_i^{U-\text{avg}}(t)$ 为优势邻居个体的平均策略, N_Ω^* 为优势邻居个体的数量.

通过引入具有全局记忆机制的特殊个体策略更新规则, 特殊个体在一轮策略更新中可以产生基于邻居最优策略和自身最优策略的混合策略, 激发合作行为的产生, 普通个体充当着策略传播者的角色, 模仿特殊个体产生的混合策略, 从而促进整个集群提高合作行为. 当普通个体的策略基本达到最优策略时, 特殊个体受到邻居反馈作用的调节, 进一步提高系统合作水平. 特殊个体搜索新策略并学习邻居优势策略的过

程体现了群体智能激发的过程。

综上, 基于特殊个体的网络进化博弈流程如下:

Step 1 对所有博弈个体的初始策略随机设置;

Step 2 根据图1中的流程完成一轮重复博弈;

Step 3 由式(3)和(4)计算个体的博弈收益, 由式(5)计算整体合作频率;

Step 4 普通个体按照式(8)更新博弈策略, 特殊个体按照式(8)和(9)更新博弈策略;

Step 5 重复Step 2~Step 4, 直至达到仿真终止条件;

Step 6 绘制稳态合作频率、平均马尔可夫策略等曲线, 进行结果分析。

4 仿真实验与结果分析

采用仿真结果中的稳态值, 包括合作频率 f_c 与平均策略 $\langle x \rangle$ 等参数在迭代过程最后10%的平均结果作为演化博弈模型的评价指标。晶格网络规模为 50×50 , 具有2500个节点, 每个博弈参与者占据一个节点, 根据规则方格网络结构与四个邻居进行交互。个体的初始策略随机生成, 为了消除随机策略的影响, 该过程进行了100次独立的蒙特卡罗仿真。仿真实验中的全部参数设置如表2所示。

图2(a)和(b)分别给出了5次随机仿真中, 原系统与干预系统的合作频率进化曲线。其中, 蓝色粗线表示5次独立仿真的平均结果。仿真结果显示, 不存在特殊个体干预时, 原系统的进化结果在很大程度上依赖于个

表 2 博弈模型参数

Table 2 Parameters of game model

参数	名称	数值
L	方格网络长度	50
P	重复博弈次数	50
T	蒙特卡罗仿真次数	100
$\langle z \rangle$	交互个体数	4
p_s	特殊个体比例	0.05
ω_1	信息权重	0.95
ω_2	优势邻居权重	0.05

体的初始策略分布, 单次仿真结果存在较大的方差, 而特殊个体($p_s=0.05$)的引入则大幅降低了仿真结果的方差。比较两种情况下系统的进化特性曲线, 可以得出以下结论: 一方面, 由于原系统中个体采取无条件模仿的更新策略, 因此系统中仅存在马尔可夫策略的复制而没有新的策略产生, 导致原系统进化结果依赖于初始状态。在软控制框架下, 特殊个体可同时利用个体历史信息与邻居收益信息, 自适应地寻找新的重复博弈策略, 从而在最大化个体以及群体收益的同时, 也增加了系统策略的多样性, 减轻了进化结果对初始状态的依赖性。综上, 更新机制的软控制不仅可在不利环境下促进合作的产生与维持, 同时也增强了系统对随机因素的鲁棒性。

从图2中可以看出, 在不同的仿真条件下, 合作频率曲线有着相似的进化特性: 初始策略随机分配, 合作频率均为0.5, 每个个体具有相同的选择合作或背叛策

图 2 合作频率进化对比曲线($b=1$). (a) 无干预($p_s=0$); (b) 有干预($p_s=0.05$)

Figure 2 Evolutionary comparison curve of cooperation frequency ($b=1$). (a) Non-intervention ($p_s=0$); (b) under intervention ($p_s=0.05$).

略的概率; 在博弈的初始阶段, 采取背叛策略的个体在面对采取合作策略的个体时能获得较大的收益, 使得背叛行为在系统中迅速传播, 合作频率逐渐降低; 然而随着背叛行为的扩张, 当博弈双方均选择背叛策略时的收益不及双方均选择合作带来的收益, 背叛行为的收益降低而合作行为的收益增加, 这使得合作频率由最低点逐渐上升, 最终达到平衡状态。

图3给出了引入特殊个体比例为 $p_s=0.05$ 时与不存在特殊个体两种情况下的进化特性对比结果。图3(a)给出了稳态时系统的平均合作频率。对于不存在特殊个体的原系统而言, 当博弈强度 b 从0增加至1时, 合作频率从0.95逐渐减小至0.25, 与之相对的是, 在比例为 $p_s=0.05$ 特殊个体的干预下, 系统的平均合作频率得到了极大的提高, 且几乎不受博弈强度 b 的影响。当 b 从0增加至1时, f_c 在0.9~0.95之间波动, 变化幅度较小。导致上述结果的主要原因是特殊个体的更新机制具有全局搜索能力, 可根据自身历史最优信息和邻居最优信息有效地搜索优势策略并对邻居施加影响, 从而影响整个系统的合作水平。

图3(b)给出了系统平均稳态马尔可夫策略 $\langle x \rangle = (\langle p_0 \rangle, \langle p_c \rangle, \langle p_d \rangle)$ 随 b 的变化趋势。当 $p_s=0.05$ 且 $0 < b < 1$ 时, 系统近似收敛于 $\langle x \rangle = (1, 1, 0)$ 的平稳马尔可夫策略, 即“以牙还牙”(TFT)策略^[15]。该策略被认为是重复囚徒困境中最简单有效的策略。其中 $\langle p_0 \rangle$ 和 $\langle p_c \rangle$ 的收敛效果较好, $\langle p_d \rangle$ 并没有严格下降到0, 但博弈过程中背叛行为几乎湮灭, $\langle p_d \rangle$ 对系统的影响较小。此时系统中个体的单次行动平均收益 $\langle R \rangle$ 达到了接近最优值1的状态, 此时系统

的总体收益达到最高。与此相反的是, 在不存在特殊个体时, 系统平均策略劣于TFT, 且其在不同情况下合作的概率随 b 的增加而减小, 因此单次平均收益也从0.95降低至0.25。

调整特殊个体的更新规则, 对后半部分迭代过程中特殊个体的策略更新方法按照式(9)进行优化。图4(a)和(b)分别给出了改进特殊个体策略更新规则前后平均合作频率和稳态策略的对比结果。通过调整后半部分迭代过程的特殊个体的策略更新规则, 合作频率从0.95左右提升到0.98左右, 并且具有更高的稳定性。 $\langle p_0 \rangle$, $\langle p_c \rangle$, $\langle p_d \rangle$ 也均有所提高。这是由于当迭代过程达到一定次数时, 普通个体基本保持了收益较高的优势策略, 而特殊个体仍在不断搜索新策略, 新策略相较于普通个体的策略普遍是一种不利于合作行为产生的劣势策略, 反而影响了合作水平的进一步提升。此时特殊个体停止对新策略的探索, 转而向收益高于自身的优势邻居个体学习。从邻居的反馈中学习优势策略体现了特殊个体的群体智能行为。

图5(a)说明了粒子数量对博弈模型的影响。在特殊个体比例 $p_s=0.05$ 、博弈强度 $b=1$ 的条件下, 分别对模型边长 $L=10, 20, 50$ 进行仿真分析, 对应粒子数量为100, 400, 2500。当粒子数量较少时, 不利于群体合作行为的产生, 这说明种群规模较小时对初始条件的依赖性较大, 由于个体数量少, 特殊个体的数量少, 可供特殊个体搜索更新的策略范围较小, 难以搜索得到促进合作行为的优势策略, 并且在进化过程初期抵抗背叛行为扩张的能力较弱, 因此难以形成合作行为; 提高

图3 $p_s=0.05$ 与 $p_s=0$ 时的进化特性对比。(a) 合作频率 f_c ; (b) 平均马尔可夫策略

Figure 3 Comparison of evolutionary characteristics between $p_s=0.05$ and $p_s=0$. (a) Cooperative frequency; (b) average Markov strategy.

图 4 特殊个体策略更新规则改进对比。(a) 合作频率 f_c ; (b) 平均马尔可夫策略

Figure 4 The improvement comparison of special individual policy update rules. (a) Cooperative frequency; (b) average Markov strategy.

图 5 粒子数量(a)和重复博弈长度(b)进化对比曲线($b=1$)

Figure 5 (a) Comparison curve of particle number evolution; (b) length evolution comparison curve of repeated games ($b=1$).

粒子群数量, 合作频率显著提高, 当粒子群数量达到一定水平时, 在特殊个体作用下合作频率基本保持稳定, 继续增加种群规模不会进一步提高群体合作水平, 说明此时种群数量已经足以使特殊个体发挥其促进合作的作用。

图5(b)说明了重复博弈长度对博弈模型的影响。在特殊个体比例 $p_s=0.05$ 、博弈强度 $b=1$ 的条件下, 分别对博弈长度 $P=1, 10, 50$ 进行仿真分析。在单次博弈中引入特殊个体对合作频率的提升没有影响, 这是由于单次博弈中个体决策仅取决于马尔可夫策略中的项, 而特殊个体的优化规则挑选的是累计收益较高的策略, 这说明特殊个体搜索全局最优策略的作用需要

通过重复博弈实现; 分别选取重复博弈长度为10和50, 相较于单次博弈过程合作频率显著提高, 随着博弈长度的增加, 合作频率的变化更加迅速。这说明在一轮重复博弈中, 博弈长度越长, 背叛者考虑到接下来的博弈中遭到对方报复的可能性, 博弈双方在选择背叛行为时会更加谨慎。在长期重复博弈中, 选择互相合作的累计收益更高, 此时合作倾向更高的马尔可夫策略属于优势策略。综上, 提高博弈长度有助于提高群体演化的快速性和激发合作行为的产生。

图6(a)说明了特殊个体比例对博弈模型的影响。固定信息权重 ω_1 , 改变特殊个体的比例, 依次取 $p_s=0.05, 0.15, 0.25, 0.35$, 发现合作频率会随着特殊个体

图 6 合作频率 f_c 与特殊个体比例 p_s ($\omega_1=0.95$) (a)、信息权重 ω_1 ($p_s=0.25$) (b)的关系

Figure 6 The relationship between cooperation frequency f_c vs. the proportion of special individuals p_s ($\omega_1=0.95$) (a), f_c vs. information weight ω_1 ($p_s=0.25$) (b).

比例的增加而下降. 这说明通过少量特殊个体对大量普通个体的影响, 系统中的合作水平得以大幅度提升. 在这一过程中, 特殊个体和普通个体扮演着不同的角色. 特殊个体通过寻找并产生收益最高的新策略扮演策略激发者的角色, 与特殊个体相邻的普通个体通过无条件模仿的更新规则充当策略传播者的角色, 因此少量的特殊个体能够将自身的优势策略辐射到群体中. 当提高特殊个体比例时, 虽然有助于提高系统的策略多样性, 但普通个体比例的下降会使得优势策略的传播缺少途径, 特殊个体具有搜索优势策略的能力但不具备模仿优势策略的能力, 如果优势策略不能及时传播, 会导致其被特殊个体产生的新策略覆盖. 因此, 当特殊个体的邻居中缺少足够数量的普通个体对优势策略进行有效传播时, 群体达成较高的合作水平变得更加困难. 存在特殊个体的最优比例, 使得群体合作水平达到最佳.

图6(b)说明了邻居反馈信息权重对博弈模型的影响. 固定特殊个体比例 $p_s=0.25$, 改变信息权重, 依次取 $\omega_1=0.95, 0.75, 0.55, 0.35$, 发现合作频率会随着信息权重的减小而略微增加. 信息权重决定了特殊个体在产生新策略时参考自身历史最优解和邻居最优解的比例, 在网络进化博弈系统中, 个体收益不仅取决于其策略选择, 同时也实时受到环境的反馈影响. 换言之, 个体历史最优策略对于当前形势而言并非好的选择. 在这种情况下, 历史信息对于个体收益以及系统合作水平的提高影响较小. 另一方面, 社会信息作为系统的

实时动态反馈, 可为特殊个体的策略搜索提供有效的信息支持. 综上所述, 通过较小的 ω_1 为社会信息分配更高的权值, 可在充满背叛的环境中有效地激励合作行为, 而单纯地依赖个体历史信息则会为系统的合作水平带来负面的影响.

5 总结

本文在进化博弈软控制的框架下, 首先建立了基于二维正则晶格网络的重复囚徒博弈模型, 然后针对特殊个体和普通个体选取了不同的策略更新规则, 普通个体采取无条件模仿, 特殊个体采用基于邻居反馈信息和自身历史信息的综合策略作为更新规则, 并设计了优化策略. 最后进行了仿真验证, 先分析了群体进化特性以及特殊个体对提高合作水平的影响, 随后讨论了模型参数的选取依据和特殊个体比例, 信息权重以及博弈强度对合作行为的影响. 研究结果表明, 基于群智合作激发的特殊个体设计规则可以在高博弈强度的环境中提高种群的合作比例, 同时通过蒙特卡罗仿真可消除初始参数中随机因素带来的不确定性. 其次, 单纯地增加特殊个体的比例并不能提升群体合作水平, 相反地, 过多的特殊个体使得群体合作频率降低. 最后, 降低信息权重会略微提升群体合作频率. 本文的研究结果可以表明群体合作行为中的涌现机制, 并为大规模多智能体系统, 如无人机集群中的合作机制设计提供新的思路.

参考文献

- 1 Sun C, Duan H. Impact of skill intervention on the evolution of cooperation. *Phys A-Stat Mech Its Appl*, 2015, 434: 171–180
- 2 Chen L, Duan H B, Fan Y M. Anatomy of information transfer for ibis using transfer entropy and active information storage (in Chinese). *Sci Sin Tech*, 2020, 50: 380–388 [陈琳, 段海滨, 范彦铭. 基于传递熵和有效信息存储的隐鹈信息传递过程研究. 中国科学: 技术科学, 2020, 50: 380–388]
- 3 Nowak M A, May R M. Evolutionary games and spatial chaos. *Nature*, 1992, 359: 826–829
- 4 Realpe-Gómez J, Vilone D, Andrighetto G, et al. Learning dynamics and norm psychology supports human cooperation in a large-scale Prisoner's Dilemma on networks. *Games*, 2018, 9: 1–14
- 5 Hong L, Geng Y, Du C, et al. Average payoff-driven or imitation? A new evidence from evolutionary game theory in finite populations. *Appl Math Comput*, 2021, 394: 1–12
- 6 McAvoy A, Hauert C. Autocratic strategies for alternating games. *Theor Population Biol*, 2017, 113: 13–22
- 7 Press W H, Dyson F J. Iterated Prisoner's Dilemma contains strategies that dominate any evolutionary opponent. *Proc Natl Acad Sci USA*, 2012, 109: 10409–10413
- 8 Szolnoki A, Chen X. Gradual learning supports cooperation in spatial prisoner's dilemma game. *Chaos Solit Fractals*, 2020, 130: 109447
- 9 Nadini M, Pongsachai P, Spinello C, et al. Empirical evidence of upward social comparison in a Prisoner's Dilemma Game. *IEEE Access*, 2020, 8: 52884–52894
- 10 Shu F, Li M, Liu X. Memory mechanism with weighting promotes cooperation in the evolutionary games. *Chaos Solit Fractals*, 2019, 120: 17–24
- 11 Li P, Duan H B. A flocking model base on selective attention mechanism (in Chinese). *Sci Sin Tech*, 2019, 49: 1040–1050 [李沛, 段海滨. 一种基于注意力机制的群集运动模型. 中国科学: 技术科学, 2019, 49: 1040–1050]
- 12 Han J, Li M, Guo L. Soft control on collective behavior of a group of autonomous agents by a skill agent. *J Syst Sci Complex*, 2006, 19: 54–62
- 13 Han H, Qiang C, Wang C, et al. Soft-control for collective opinion of weighted DeGroot model. *J Syst Sci Complex*, 2017, 30: 550–567
- 14 Han J, Wang C Y. Guiding swarm behavior by soft control. *IET Digital Library*, 2018: 345–384
- 15 Sartoretti G. Leader-based versus soft control of multi-agent swarms. *Artif Life Robotics*, 2016, 21: 302–307
- 16 Duan H, Sun C. Swarm intelligence inspired skills and the evolution of cooperation. *Sci Rep*, 2014, 4: 5210–5218
- 17 Sun C H. Dynamic resource allocation of multi-UAV based on network evolutionary game (in Chinese). Dissertation for Doctoral Degree. Beijing: Beihang University, 2015 [孙昌浩. 基于网络进化博弈的多无人机动态资源分配. 博士学位论文. 北京: 北京航空航天大学, 2015]
- 18 Wang X, Lv S. The roles of particle swarm intelligence in the prisoner's dilemma based on continuous and mixed strategy systems on scale-free networks. *Appl Math Comput*, 2019, 355: 213–220
- 19 Liu Y, Yang C, Huang K, et al. Swarm intelligence inspired cooperation promotion and symmetry breaking in interdependent networked game. *Chaos*, 2019, 29: 043101
- 20 Duan H B, Qiu H X. Unmanned Aerial Vehicle Swarm Autonomous Control Based on Swarm Intelligence (in Chinese). Beijing: Science Press, 2018 [段海滨, 邱华鑫. 基于群体智能的无人机集群自主控制. 北京: 科学出版社, 2018]

Excitation and control of group intelligence cooperation in network evolutionary game based on special individuals

WANG HaoMiao, WEI Chen & DENG YiMin

Bio-inspired Autonomous Flight Systems (BAFS) Research Group, Science and Technology on Aircraft Control Laboratory, School of Automation Science and Electrical Engineering, Beihang University, Beijing 100083, China

Guiding the generation and propagation of intelligent behavior in crowds is an important research direction in the field of swarm intelligence. Cooperation is a common behavior of swarm intelligence in nature. In this study, an evolutionary game model based on special individuals is established to analyze the evolutionary characteristics and regulatory measures of group-inspired cooperation levels. First, an effective intervention control architecture is established in the rule network evolutionary game model to realize the regulation effect on the cluster system. Second, by designing the control rules of special individuals based on swarm intelligence, special individuals can stimulate the emergence of cooperative behaviors. Finally, the Monte Carlo simulation method is used to analyze the influence of special individuals on the evolution characteristics of a group, and the influence of different parameter conditions on the cooperation level is discussed.

special individual, network evolutionary game, soft control, swarm intelligence cooperative excitation, Monte Carlo simulation

doi: [10.1360/SST-2021-0265](https://doi.org/10.1360/SST-2021-0265)