



基于合作竞争公共物品博弈的无人机集群动态资源分配

王浩淼¹, 段海滨^{1,2*}, 魏晨¹

1. 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 100191

2. 鹏城实验室, 深圳 518000

* 通信作者. E-mail: hbduan@buaa.edu.cn

收稿日期: 2021-12-12; 修回日期: 2022-02-19; 接受日期: 2022-03-18; 网络出版日期: 2022-09-05

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(批准号: U20B2071, 91948204, U1913602, U19B2033)资助项目

摘要 无人机集群动态资源分配是其任务规划的挑战性关键技术难题. 本文提出了一种基于合作竞争机制的动态资源分配方法, 将无人机集群与网络进化博弈模型进行对应, 建立了基于无人机拓扑网络的博弈模型, 通过设计针对特殊个体的调控规则, 达到提升系统平均收益的目的. 最后, 通过无人机博弈模型的仿真实验, 根据网络结构、模型参数, 以及不同激发机制下系统的稳态结果, 验证了本文所提无人机动态资源分配方法的可行性和有效性.

关键词 无人机集群, 公共物品博弈, 动态资源分配, 合作机制, 竞争机制

1 引言

随着作战环境信息化、网络化和智能化的不断发展, 为实现更为复杂的任务指标, 以无人机集群为基础的多无人机协同控制技术引起了广泛关注^[1]. 无人机集群是由多架无人机组成的复杂系统, 能够利用多种协同控制方法保证集群系统的鲁棒性和稳定性, 具有可回收性高、任务覆盖面广、任务执行效率高等特点^[2]. 未来战场的高动态性和不确定性, 包括作战任务快速调整、无人机个体状态和能力变化、无人机通信拓扑改变, 以及敌方或环境等外界因素的干扰, 对无人机集群协同资源分配方法设计提出了更高的要求^[3,4]. 高效的资源分配方法对提高无人机协同执行任务能力起到重要作用, 通过对无人机执行任务时所需资源进行合理分配, 能够实现以最小全局代价或最大全局收益完成任务. 然而, 无人机作为异质性智能个体, 在进行资源分配和任务规划时存在个体利益与群体利益之间的冲突. 如何设计面向多无人机协同任务的动态资源分配机制, 成为具有现实应用价值的技术挑战^[5,6].

多无人机协同资源分配本质上是一类组合优化问题, 具有自组织性、涌现性、非线性和不可预测性等复杂系统的特征^[6]. 目前常见的研究方法包括数学规划方法^[7]、启发式方法^[8]、群体智能方法^[9],

引用格式: 王浩淼, 段海滨, 魏晨. 基于合作竞争公共物品博弈的无人机集群动态资源分配. 中国科学: 信息科学, 2022, 52: 1598–1609, doi: 10.1360/SSI-2021-0416
Wang H M, Duan H B, Wei C. Dynamic resource allocation of UAV swarms based on cooperative and competitive public goods game (in Chinese). Sci Sin Inform, 2022, 52: 1598–1609, doi: 10.1360/SSI-2021-0416

以及强化学习等^[10]. 然而这些研究方法侧重于对群体的宏观约束而非个体的策略调控, 忽视了个体利益与群体利益之间的冲突, 并且难以适应动态变化的任务场景. 博弈论是一种研究从个体行为到群体智能协同控制的有效方法, 进化博弈论引入了进化生物学中的基本概念, 研究由有限智能个体组成的种群在不断的重复博弈过程中通过自适应地学习来优化收益^[11]. Nowak 等^[12] 提出了一种基于二维晶格网络的囚徒博弈模型, 开启了网络进化博弈研究的先河. 复杂网络用于描述与模拟现实世界中的群体结构, Santos 等^[13] 研究了复杂网络中网络结构异质性对合作行为的影响, 表明复杂网络更接近现实世界的智能个体的交互关系. 为了解决个体利益与群体利益的冲突, 公共物品博弈 (public goods game, PGG) 长期以来被作为典型的范例^[14], 描述智能体通过连续迭代的博弈过程更新策略, 进而不断适应外部环境的过程. Quan 等^[15] 提出了基于惩罚策略的博弈机制, 通过多数合作者对少数背叛者的惩罚实现了群体合作效果. Irlenbusch 等^[16] 针对个体获取信息能力的异质性讨论了透明度对个体策略更新机制的影响. Kawagoe 等^[17] 研究了自愿机制下公共物品博弈的演化机制. Li 等^[18] 研究了基于个体收益的博弈准入机制, 证明了设置门槛有助于提高合作行为. 这些研究说明公共物品博弈提供了一种可靠的集群自主协同控制框架.

将博弈思想应用在无人机集群的资源分配问题中, Luo 等^[19] 从资源分配对博弈策略影响的角度研究了网络进化博弈中资源协同演化与合作, 但没有对复杂网络结构以及个体的连接关系作进一步研究. Liu 等^[20] 提出了一种将无人机弹药资源分配和演化博弈相结合的方法, 推导了合作策略占优的理论条件, 但缺少对分配结果的深入研究, 并且难以推广到更一般的资源分配问题中. Yan 等^[21] 提出一种基于演化博弈的无人机网络接入和资源分配方法, 但重点在于无人机网络的稳定性, 而非具体的资源分配策略. 无人机集群的任务要求并不仅限于提高合作水平, 有时也需要通过促进集群内部的竞争行为, 驱使个体寻找最优策略, 提高群体收益. 为了更好地模拟无人机动态资源分配的过程, 将无人机资源分配问题与公共物品博弈结合, 无人机作为有限理性的个体, 在重复迭代的过程中不断更新策略和状态, 最终达到稳定结果. 演化目标并非一味地提高合作水平, 而是兼顾了合作水平和群体收益的提升, 根据任务的具体要求采取合作或竞争的博弈策略.

本文建立了一种基于异质性策略公共物品博弈的无人机资源分配模型, 用复杂网络模拟无人机集群的拓扑结构, 从中选择具有较大影响力的特殊个体, 并设计了基于合作机制与竞争机制的两种博弈策略, 利用蒙特卡罗 (Monte Carlo) 方法与统计分析工具进行仿真分析, 验证了激发机制与模型参数选择对集群动态资源分配任务的影响.

2 基于个体属性的无人机网络博弈模型

2.1 无人机模型与网络结构

考虑配有速度、航向角和高度 3 个回路的自动驾驶仪的固定翼无人机简化模型^[5]:

$$\begin{cases} \dot{x}^i = V^i \cos \psi^i, \\ \dot{y}^i = V^i \sin \psi^i, \\ \dot{h}^i = \lambda^i, \\ \dot{V}^i = \frac{1}{\tau_V} (V_c^i - V^i), \\ \dot{\psi}^i = \frac{1}{\tau_\psi} (\psi_c^i - \psi^i), \\ \dot{\lambda}^i = -\frac{1}{\tau_\lambda} + \frac{1}{\tau_h} (h_c^i - h^i), \end{cases} \quad (1)$$

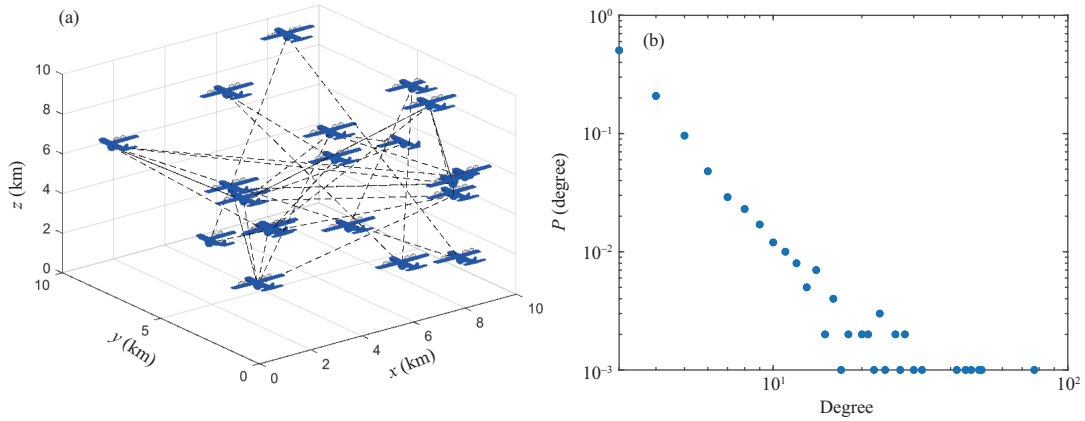


图 1 (网络版彩图) 无人机拓扑网络

Figure 1 (Color online) A UAV topological network. (a) Scale-free network structure; (b) scale-free network degree distribution

其中, (x^i, y^i, h^i) 为无人机个体 i 的位置, V^i, ψ^i 和 λ^i 分别为水平速度、航向角和高度变化率; V_c^i, ψ_c^i 和 λ_c^i 分别为无人机自驾仪 3 个回路上的控制输入指令, τ_V, τ_ψ 和 (τ_λ, τ_h) 分别为 3 个回路的时间常数, 即自驾仪的控制参数.

无人机由于受到通信半径和负载的限制, 仅与有限邻居个体进行交互, 并且由于个体能力、空间位置的差异, 无人机拓扑网络结构具有如图 1 所示的无标度性质. 可以看出在无标度网络中, 节点度分布服从幂律机制, 大多数节点具有很小的度, 少数节点具有较大的连接度. 根据 Barabási 与 Albert^[22] 提出的一种构建无标度网络方法, 网络从最初的 m_0 个节点开始, 每一步增加的节点与已经存在的 m 个旧点建立连接, 并且新节点加入时优先选择连接度大的个体. 连接度大的节点对应无人机集群中影响力较大的个体, 仅对这些少数个体施加控制就能起到干预集体行为的作用. 根据网络节点的异质性选择具有较大连接度的节点作为特殊个体, 在网络传播的过程中起到关键作用.

2.2 基于异质性策略的公共物品博弈模型

公共物品博弈作为一种典型的博弈模型, 在解决多智能体参与的群体利益冲突问题中有着广泛应用^[9]. 将无人机抽象为 N 个智能个体, 个体具有合作 (C) 和背叛 (D) 两种基本策略. 在每一轮中, 个体与其所有邻居分别博弈, 采取合作策略的个体向公共池投入成本 c , 不失一般性, 取单位成本 $c = 1$, 采取非合作策略的个体不投入. 考虑采用合作策略的个体具有异质性, 对应着无人机个体在执行任务时的能力和意愿具有差异. 因此个体 i 的博弈策略 x_i 集合为

$$x_i = 0, \frac{1}{s}, \frac{2}{s}, \dots, \frac{s-1}{s}, 1, \quad (2)$$

其中, s 为合作策略的数量, 通常情况下策略异质性程度与交互邻居个数有关, 因此取 $s = d + 1$, d 为网络平均连接度. 当 $x_i = 0$ 时, 个体的策略为背叛; 当 $x_i \neq 0$ 时, 个体的博弈策略为合作, 且合作策略具有异质性. 一轮博弈的总收益为公共池中总投入乘以增益系数 r , 并被所有博弈参与者均分. 每架无人机参与包括自身以及所有邻居的 $d_i + 1$ 次公共物品博弈, 单次博弈中此个体 i 采取合作与非合作

策略时的收益函数如下所示:

$$\begin{aligned}\pi_{i,C} &= \frac{r}{d_i + 1} \sum_{j \in \Omega_i} \frac{x_j}{d + 1} - \frac{x_i}{d + 1}, \\ \pi_{i,D} &= \frac{r}{d_i + 1} \sum_{j \in \Omega_i} \frac{x_j}{d + 1},\end{aligned}\quad (3)$$

其中, Ω_i 为无人机个体 i 包括自身在内的邻居集合; d_i 为个体 i 的连接度; c_i 为个体 i 的合作策略, c_j 为邻居 j 的合作策略. 第 t 轮博弈过程中个体 i 的总收益以及系统的平均收益为

$$\begin{aligned}\Pi_i(t) &= \sum_{j=1}^{d_i+1} \pi_{i,j}(t), \\ \bar{\Pi}(t) &= \frac{1}{N} \sum_{i=1}^N \Pi_i(t).\end{aligned}\quad (4)$$

所有个体采用 Fermi 机制进行同步更新^[6], 随机选择一个邻居并同时复制其策略及个体类型, t 时刻博弈个体 i 模仿邻居 j 的概率为

$$p(x_j(t) \rightarrow x_i(t+1)) = \frac{1}{1 + e^{(\Pi_i(t) - \Pi_j(t))/\omega}}, \quad (5)$$

其中, $B \rightarrow A$ 表示 B 对 A 进行复制, $x_i(t)$ 表示个体 i 在当前时刻的策略, $x_j(t+1)$ 表示个体 j 在下一时刻的策略, ω 表示选择噪声强度, $1/\omega$ 是选择强度因子, 它刻画了选择过程的不确定性. 当 $\omega \rightarrow 0$ 时, Fermi 规则转化为一个确定性过程, 在此情况下, 若邻居 j 收益大于个体 i 收益, 即 $\Pi_j(t) > \Pi_i(t)$ 时, 则 i 以概率 1 复制其策略, 否则个体保持其当前策略. 当 $\omega \rightarrow \infty$ 时, 该规则转化为完全随机过程, 即不论邻居 j 的收益优劣, 个体 i 总是以 0.5 的概率选择其策略进行复制本文取 $1/\omega = 0.1$, 此时个体倾向于选择收益高于自身的优势策略进行学习, 但劣势策略也存在传播的可能性.

当某无人机个体 i 策略由合作转化为非合作时, 假设其邻居的策略均不发生改变, 其收益变化为

$$\begin{aligned}\pi_{i,D}(t) - \pi_{i,C}(t) &= \frac{r}{d_i + 1} \sum_{j \in \Omega_i} \frac{x_j(t+1)}{d + 1} - \left(\frac{r}{d_i + 1} \sum_{j \in \Omega_i} \frac{x_j(t)}{d + 1} - \frac{x_i(t)}{d + 1} \right) \\ &= \frac{r}{d_i + 1} \sum_{j \in \Omega_i} \left(\frac{x_j(t)}{d + 1} - \frac{x_i(t)}{d + 1} \right) - \left(\frac{r}{d_i + 1} \sum_{j \in \Omega_i} \frac{x_j(t)}{d + 1} - \frac{x_i(t)}{d + 1} \right) \\ &= \frac{x_i(t)}{d + 1} \left(1 - \frac{r}{d_i + 1} \right).\end{aligned}\quad (6)$$

由此得到个体的合作策略收益大于非合作策略时的条件:

$$r > d_i + 1. \quad (7)$$

将式 (7) 的结论推广到集群系统中. 由于 d_i 表示个体的连接度, 其均值为 d , 在规则网络中, $d_i = d$; 在无标度网络中, 如图 1 所示, 少数无人机节点具有较大的连接度, 而大部分节点的连接度低于平均连接度. 因此可以得到群体中更为一般的合作成为优势策略的条件:

$$r > d + 1. \quad (8)$$

由此可见, 增益系数 r 与网络的平均连接度 d 有关. 当满足式 (8) 的条件时, 系统中的大部分个体满足了式 (7) 的条件, 倾向于选择合作, 此时合作成为系统的优势策略.

3 无人机集群动态资源分配策略

3.1 无人机集群与博弈模型映射关系

如何在动态的网络环境下, 通过集群个体的行为交互与协同控制, 消除无人机个体间以及个体与全局利益的冲突, 从而实现任务分配过程中无人机集群资源的最优投入成为了一个具有重要理论与应用价值的研究课题^[19]. 网络进化博弈与多无人机分布式资源分配在本质上存在如下的相同点.

(1) 无人机集群系统具有固定的拓扑结构, 个体地位存在差异, 这与复杂网络的空间结构和节点连接关系是类似的.

(2) 在执行任务的资源投入时, 无人机个体仅与有限的邻居进行信息交互并做出决策, 个体利益互相耦合, 且与群体的利益存在冲突, 这与进化博弈理论中个体与群体利益的冲突是一致的.

(3) 协同分配任务的最终目标是最大化无人机集群系统的收益, 这与博弈模型中个体通过特定的空间结构与学习机制获得最大全局收益是相同的.

因此可以将无人机集群系统资源分配与网络进化博弈中相关概念进行对应, 从而简化研究方法. 集群系统中的单架无人机作为博弈个体, 无人机集群拓扑网络组成了博弈网络, 集群执行任务时个体是否投入资源对应着合作与非合作博弈策略, 个体跟随集体完成任务后获得的回报即为博弈收益, 集群中个体完成任务时不同策略占比稳定代表了稳定博弈结果.

3.2 基于合作竞争机制的特殊个体策略选择

无人机网络结构具有无标度的性质, 根据 2.2 小节中理论推导, 个体连接度越大越有利于其传播策略, 并且对群体的最终演化结果影响较大, 因此有必要针对特殊个体设计更具智能的博弈策略^[12]. 合作与竞争是无人机群体中最简单的两种交互关系. 一般地, 合作能带来更高的集体收益, 由所有参与合作的无人机个体根据团队贡献分配资源; 而竞争则追求提升个体收益, 不参与到其他个体的交互中, 而获得固定的资源分配; 在集群系统中个体通常会选择合作以依附于集体的力量获得更高的收益. 然而合作并非严格的优势策略, 当群体资源有限时合作收益不足以满足个体需要, 或是在充满背叛的环境中合作行为无法带来收益, 此时个体通过竞争策略拒绝参与集体行为, 提高个体收益的同时也提高了集体的平均收益. 因此设计了合作激发和竞争激发两种策略选择方式, 根据外界条件以及系统内部的演化规律自适应地实现全局收益最大化, 体现了无人机群体智能的特性.

合作激发的意义在于对合作行为的鼓励, 特别是对合作策略 $c = 1$ 的完全合作者, 在传播合作策略的过程中起到积极的作用, 然而由于其他合作策略带有的部分背叛性质, 导致完全合作者的收益往往低于其他合作者以及非合作者. 引入对合作者的奖励机制是合作激发模式下的必要策略, 奖励机制下完全合作者的博弈收益如下:

$$\pi_{i,c=1} = \frac{r}{d_i + 1} \sum_{j \in \Omega_i} \frac{c_j}{d + 1} - \frac{c_i}{d + 1} + \beta \frac{N_D}{N_{c=1}}, \quad (9)$$

其中 N_D 表示无人机邻居中非合作者的数量, $N_{c=1}$ 表示邻居中包括自身在内合作者的数量 β 为奖励系数. 式 (9) 表明奖励程度与非合作者数量有关, 奖励收益由完全合作者平分, 当博弈中不存在非合作者时, 合作激发机制退化为一般博弈策略. 合作激发机制随着个体策略更新同步进行传播.

竞争激发的意义在于应对背叛的风险以及合作无法满足期望收益的损失, 一般地, 个体会选择非合作策略克服上述问题, 然而当群体中全部个体均为非合作状态时, 系统将会陷入演化博弈的“公地困境”^[19]. 需要对非合作者设计自愿机制作为竞争激发机制, 使其成为孤立者脱离博弈群体, 孤立者

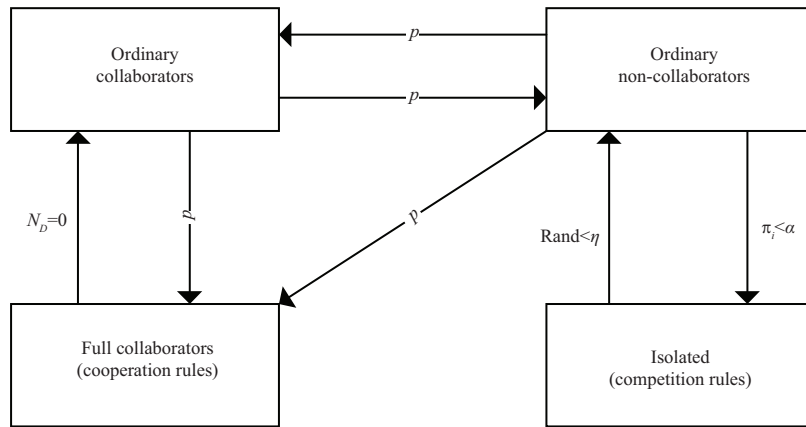


图 2 集群个体状态转化机制
Figure 2 State transition rules of cluster individuals

获得与博弈无关的固定收益 α , 竞争机制下非合作者的收益如果低于期望的固定收益, 则会转变成为孤立者. 非合作者成为孤立者一方面使自身得到稳定收益, 另一方面消除了博弈中的背叛行为, 保障了合作者的利益. 个体处于孤立状态会维持一段时间 τ , 随后重新作为非合作者进入博弈中.

引入合作与竞争激发机制后, 系统中个体的状态转换机制如图 2 所示. 其中, p 为普通个体根据 Fermi 规则计算模仿邻居策略的概率, 而特殊个体转换需要满足特定的条件.

综上, 基于特殊个体的网络进化博弈流程如下.

- Step 1. 随机分布所有个体的初始状态, 即合作或非合作策略.
- Step 2. 根据网络度分布的异质性选择具有特殊策略的个体, 特殊策略包括合作或竞争激发机制.
- Step 3. 博弈参与者根据式 (3) 和 (9) 计算收益, 孤立者获得固定收益 α .
- Step 4. 博弈参与者根据式 (5) 计算策略更新概率.
- Step 5. 所有个体按照图 2 所示的状态转化机制更新个体的策略和状态.
- Step 6. 重复 Step 3~5, 直至达到仿真终止条件.
- Step 7. 统计稳态收益、合作频率等作为评价标准, 进行结果分析.

4 仿真实验分析

4.1 网络结构与博弈模型参数的影响

无人机集群规模为 1000, 进行 1000 次迭代过程, 采用仿真结果中的稳态值, 包括合作频率、平均收益、不同策略比例等变量在迭代过程最后 10% 的平均结果作为演化博弈模型的评价指标. 各个变量定义如下.

合作频率 (合作水平) f_c 定义为 t 时刻所有个体在一轮公共物品博弈中累计合作行为的权重比例. 对于 N 个个体共存在 $(d+1)N$ 个单独的行动, 所以合作频率计算公式如下:

$$f_c = \frac{1}{(d+1)N} \sum_{i=1}^N (d_i + 1)w_i, \tag{10}$$

其中, N 为系统中无人机个体数量, d_i 表示个体 i 的连接度, 即邻居数量, w_i 表示个体 i 在与其所有邻居进行博弈时采取的行为. 若个体采取合作行为, 则 $w_i = 1$; 若个体采取非合作行为, 则 $w_i = 0$.

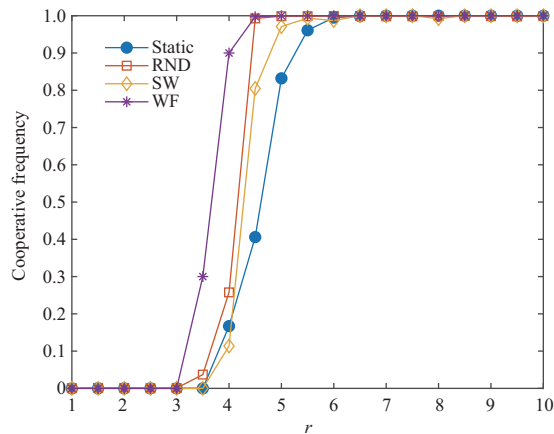


图 3 (网络版彩图) 网络结构对合作频率影响

Figure 3 (Color online) The influence of the network structure on cooperative frequency

平均收益定义为 t 时刻所有个体收益的平均值:

$$\bar{\pi}(t) = \frac{1}{N} \sum_{i=1}^N \pi_i(t). \quad (11)$$

个体的初始策略随机生成, 为了消除随机策略的影响, 实验进行了 100 次独立的蒙特卡罗仿真. 在引入合作竞争激发机制时, 合作机制的奖励系数 $\beta = (r - 1)/2$, 竞争机制中个体的固定收益 $\alpha = (r - 1)/r$, 孤立者的潜伏时间 $\tau = 10/r$, 其中 r 为增益系数.

图 3 给出了平均连接度 $d = 4$, 不存在特殊个体时, 规则网络、随机网络、小世界网络和无标度网络的合作频率与增益系数的关系. 可以看出, 当达到合作策略占优的临界条件 $r = d + 1$ 时, 4 种网络均达到了较高的合作水平, 验证了理论推导的正确性. 相较于规则网络, 复杂网络的网络结构能促进合作行为的提前出现. 特别地, 无标度网络在增益系数较低时仍然有合作行为出现, 这是由于无标度网络中存在少数连接度较大的节点, 当合作策略占据这些节点时更容易扩散给与之相连的节点, 使得合作行为在充满背叛的环境中进行传播.

图 4 给出了博弈模型参数对演化结果的影响. 图 4(a) 表明在不同的增益系数下, 合作频率遵循典型的进化模式, 初始合作者比例因背叛者的入侵有所速降低, 当增益系数较低时, 合作行为受到抑制; 当增益系数较高时, 少量的合作者通过形成团簇维持了集体的利益, 而非合作者无法享受到合作系数增加的收益, 转而向邻居中合作者学习, 合作行为得到传播. 图 4(b) 给出了无标度网络中不同平均度分度下合作频率的变化, 随着连接度的增加, 不仅导致合作成为优势策略时的增益系数增加, 也导致稳态时的合作频率降低. 当 $r > d + 1$ 时, 合作策略理论上已经成为群体优势策略. 增益系数 r 表征了无人机投入资源执行任务时的回报比例, 平均连接度 d 表征了无人机在集群中交互个体的数量, 参数选择需要在合理范围内, 且满足 $r < d + 1$ 的基本条件.

4.2 合作竞争激发机制的影响

结合 2.1 小节给出的无人机运动模型, 可根据无人机的初始位置和速度构建拓扑网络. 选取网络中度分度最大, 即邻居数量最多的前 10% 无人机个体作为特殊个体, 初始状态下合作机制和竞争机制的个体各占据一半. 分别验证单独合作机制、单独竞争机制、合作竞争机制作用下系统的演化结果, 并未引入特殊个体时的结果作为对照.

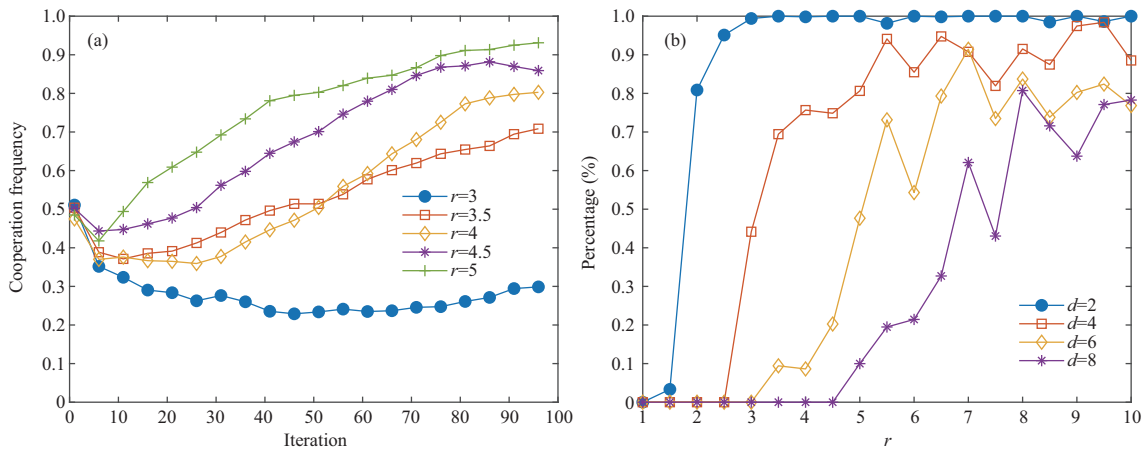


图 4 (网络版彩图) 网络参数对合作频率的影响

Figure 4 (Color online) The influence of network parameters on cooperative frequency. (a) Gain coefficient; (b) average degree of connection

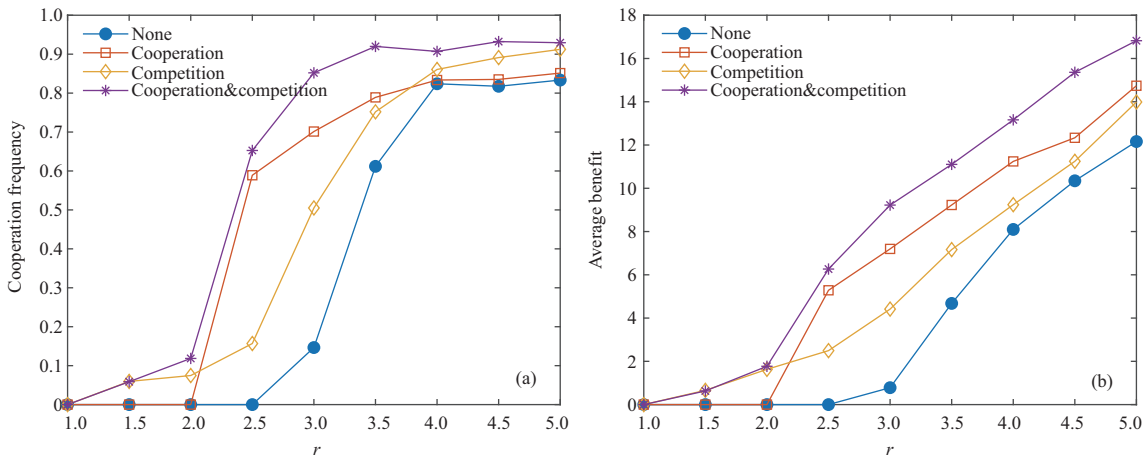


图 5 (网络版彩图) 不同激发机制下集群演化结果

Figure 5 (Color online) Cluster evolution results under different excitation rules. (a) Cooperation frequency; (b) average benefit

图 5 给出了不同激发机制下系统合作频率和平均收益的结果. 由仿真曲线可见, 与未引入任何特殊个体的传统模式相比, 引入不同机制的特殊个体可显著提高系统的平均收益. 具体地, 合作激发机制下, 合作成为优势策略的临界条件明显降低, 此时完全合作者的利益得到保障, 对完全合作者的奖励变相地惩罚了系统中的非合作者, 促进了合作行为的涌现; 竞争激发机制下, 孤立者策略为非合作者在增益系数较低的环境中提供了新选择, 个体成为孤立者的过程表现了对邻居中非合作者的不信任以及对合作者的保护; 在合作竞争机制共同作用下, 系统的合作水平与平均收益进一步提升, 说明两种机制互相协调, 能够共同发挥作用.

图 6 展示了不同状态个体的稳态演化比例. 增益系数较低 ($1 \leq r \leq 2$) 时, 个体倾向于模仿竞争机制, 个体在充满背叛的环境中倾向于成为远离博弈的孤立者, 孤立者又以一定概率成为普通个体, 以非合作者的身份重新参与博弈, 竞争机制个体与普通个体的比例维持动态稳定, 几乎不存在采用合作

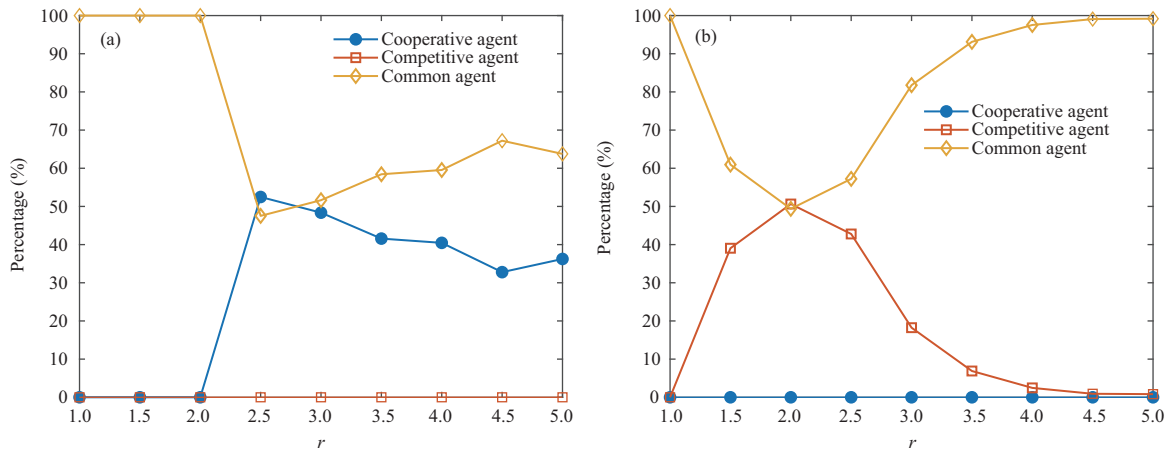


图 6 (网络版彩图) 不同激发机制下各状态个体比例

Figure 6 (Color online) Individual proportion of each state under different excitation rules. (a) Cooperative motivation mechanism; (b) competitive motivation mechanism

机制的个体; 随着增益系数的增加 ($2 \leq r \leq 4$), 对连接度较大的特殊个体, 竞争机制带来的固定收益逐渐低于参与博弈得到的收益, 因此转而开始学习合作机制, 竞争机制个体比例不断降低, 合作机制个体逐渐增加; 增益系数较高 ($4 \leq r \leq 5$) 时, 对完全合作者的奖励系数进一步增加, 采用合作机制成为主导机制, 合作机制个体的邻居中不存在非合作者时, 退化为采用完全合作者策略的普通个体, 合作机制个体与普通个体的比例维持动态稳定, 而采用竞争机制的个体基本消失. 可以看出, 合作机制和竞争机制的主要作用区间不同, 进一步验证了两种机制的合理性.

图 7 给出了合作者中不同策略个体的稳态演化比例. 可以看出, 在增益系数较低时, 投资比例较低的策略占比较高, 随着增益系数的增加, 投资比例较高的策略逐渐增加, 这进一步说明了异质性策略中某些合作者具有非合作者的性质: 相较于完全合作者, 其他投资策略的合作者的投入更低, 而收益相同, 具有部分背叛的性质. 特别地, 在合作机制作用下, 完全合作者 ($c = 1$) 在合作策略中始终占据主导地位, 这是由于对完全合作者的额外奖励制度, 使其收益高于其他投资策略合作者, 可见合作机制能抑制异质性策略中部分合作者的非合作行为产生.

因此, 在基于网络进化博弈的无人集群资源分配中, 引入合作和竞争机制, 可以提高资源分配后群体的平均收益. 合作机制相当于促进无人机个体共同投入资源以协同完成任务; 竞争机制鼓励个体独享资源, 单独执行任务, 而不与其他无人机交互. 仿真结果说明, 当增益系数, 即完成任务的收益较低时, 个体倾向于选择竞争激发机制, 表示在资源回报有限的情况下无人机个体之间以竞争为主, 由一部分个体完成任务; 当任务的收益系数增加时, 个体开始选择合作, 向集体贡献资源以协同完成任务, 特别地, 对完全合作者的奖励制度使得个体倾向于提高资源贡献率. 在实际应用场景中, 增益系数由环境以及任务类型决定, 无人机集群拓扑结构以及通信强度与无人机个体能力有关, 一般无法改变, 依据合作竞争激发机制选择自身在资源分配时的策略, 能够提高集群的平均收益水平, 提高任务执行效率.

5 总结

本文将网络进化博弈应用于无人机集群动态资源分配, 提出了一种基于合作竞争激发机制的动态

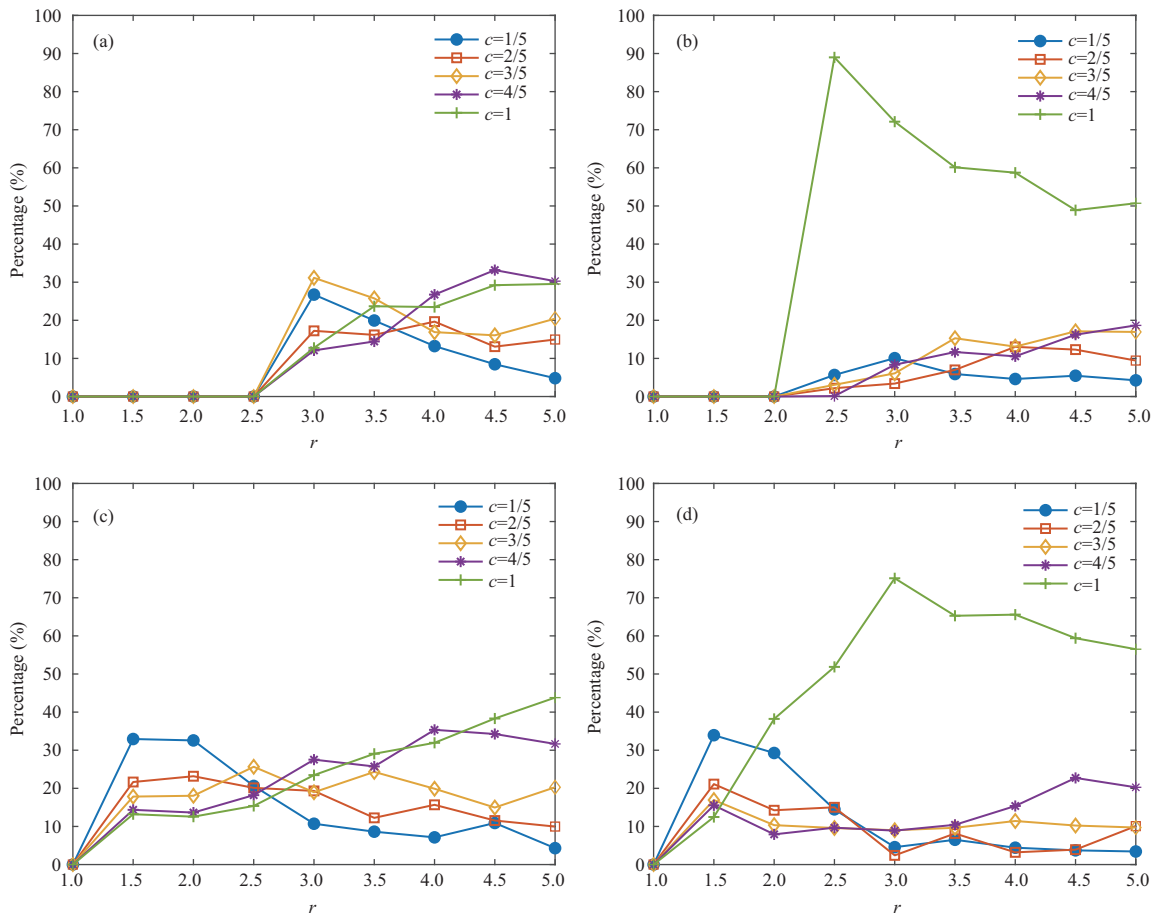


图 7 (网络版彩图) 不同激发机制下合作策略比例

Figure 7 (Color online) Proportion of cooperation strategies under different excitation rules. (a) No particular individual; (b) cooperative motivation mechanism; (c) competitive motivation mechanism; (d) cooperative and competitive motivation mechanism

资源分配方法. 首先根据无人机的速度和位置关系, 将无人机集群抽象为网络进化博弈的数学模型; 分别结合公共物品博弈中的不同机制, 设计了基于合作机制和竞争机制的特殊个体策略, 其中合作机制通过对完全合作者的奖励制度实现, 而竞争机制通过引入孤立状态实现; 然后通过理论推导分析了提高平均收益的可行性, 并进行了仿真分析和验证, 讨论了网络结构、模型参数, 以及不同激发机制对系统的稳态结果的影响, 包括合作频率、平均收益、个体状态比例、合作策略比例等. 研究表明, 引入合作竞争激发机制可以显著提高系统的平均收益, 具体地, 竞争机制在增益系数较低时起到了主要作用; 随着增益系数的增加, 合作机制开始涌现, 采用竞争机制的个体逐渐消失; 增益系数进一步增加, 系统中不同状态个体比例基本保持稳定. 最后将模型演化结果作为无人机动态资源分配的评价依据, 说明了本文所提方法对解决动态资源分配最优解问题的具有一定理论意义.

参考文献

- 1 Duan H B, Zhang D F, Fan Y M, et al. From wolf pack intelligence to UAV swarm cooperative decision-making. *Sci Sin Inform*, 2019, 49: 112–118 [段海滨, 张岱峰, 范彦铭, 等. 从狼群智能到无人机集群协同决策. *中国科学: 信息科学*, 2019, 49: 112–118]

- 2 Duan H B, Yang Q, Deng Y M, et al. Unmanned aerial systems coordinate target allocation based on wolf behaviors. *Sci China Inf Sci*, 2019, 62: 014201
- 3 Wang Y W, Liu X K, Xiao J W, et al. Output formation-containment of interacted heterogeneous linear systems by distributed hybrid active control. *Automatica*, 2018, 93: 26–32
- 4 Cui J J, Liu Y W, Nallanathan A. Multi-agent reinforcement learning-based resource allocation for UAV networks. *IEEE Trans Wirel Commun*, 2020, 19: 729–743
- 5 Duan H B, Qiu H X. *Unmanned Aerial Vehicle Swarm Autonomous Control Based on Swarm Intelligence*. 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]
- 6 Sun C H. Multi-UAV dynamical resource allocation based on network evolutionary game theory. Dissertation for Ph.D. 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]
- 7 Jia Z R, Lu F X, Wang H Y. Multi-stage attack weapon target allocation method based on defense area analysis. *J Syst Eng Electron*, 2020, 31: 539–550
- 8 Dhakal D R, Pan Z, Saito M, et al. Dynamic resource allocation in non-orthogonal multiple access using weighted maximin fairness strategy for a UAV network. *J Sign Process Syst*, 2020, 92: 1397–1406
- 9 Sun C H, Duan H B. Impact of skill intervention on the evolution of cooperation. *Phys A-Stat Mech its Appl*, 2015, 434: 171–180
- 10 Wu C, Yan B, Yu R, et al. An intelligent resource dynamic allocation method for UAV wireless mobile network which supports QoS. *Comput Commun*, 2020, 152: 46–53
- 11 Barreiro-Gomez J, Mas I, Giribet J I, et al. Distributed data-driven UAV formation control via evolutionary games: experimental results. *J Franklin Inst*, 2021, 358: 5334–5352
- 12 Nowak M A, May R M. Evolutionary games and spatial chaos. *Nature*, 1992, 359: 826–829
- 13 Santos F C, Pacheco J M. A new route to the evolution of cooperation. *J Evol Biol*, 2006, 19: 726–733
- 14 Bladon A J, Galla T. Learning dynamics in public goods games. *Phys Rev E*, 2011, 84: 041132
- 15 Quan J, Chu Y Q, Wang X J. Public goods game with punishment and the evolution of cooperation. *Syst Eng-Theory Pract*, 2019, 39: 141–149 [全吉, 储育青, 王先甲. 具有惩罚策略的公共物品博弈与合作演化. 系统工程理论与实践, 2019, 39: 141–149]
- 16 Irlenbusch B, Rilke R M, Walkowitz G. Designing feedback in voluntary contribution games: the role of transparency. *Exp Econ*, 2019, 22: 552–576
- 17 Kawagoe T, Matsubae T, Takizawa H. Quantal response equilibria in a generalized volunteer’s dilemma and step-level public goods games with binary decision. *Evolut Inst Econ Rev*, 2018, 15: 11–23
- 18 Li J, Wang J. Locality based wealth rule favors cooperation in costly public goods games. *Chaos Soliton Fract*, 2018, 116: 1–7
- 19 Luo C, Jiang Z P. Coevolving allocation of resources and cooperation in spatial evolutionary games. *Appl Math Comput*, 2017, 311: 47–57
- 20 Liu X D, He M, Yu M G, et al. Research on UAV swarm ammunition distribution based on public goods game. *Control Dec*, 2021. doi: 10.13195/j.kzyjc.2021.0483 [刘学达, 何明, 禹明刚, 等. 基于公共物品博弈的无人机集群弹药分配研究. 控制与决策, 2021. doi: 10.13195/j.kzyjc.2021.0483]
- 21 Yan S, Peng M, Cao X. A game theory approach for joint access selection and resource allocation in UAV assisted IoT communication networks. *IEEE Int Things J*, 2019, 6: 1663–1674
- 22 Barabási A L, Albert R. Emergence of scaling in random networks. *Science*, 1999, 286: 509–512

Dynamic resource allocation of UAV swarms based on cooperative and competitive public goods game

Haomiao WANG¹, Haibin DUAN^{1,2*} & Chen WEI¹

1. *Bio-inspired Autonomous Flight Systems Research Group, Science and Technology on Aircraft Control Laboratory, School of Automation Science and Electrical Engineering, Beihang University, Beijing 100191, China;*

2. *Peng Cheng Laboratory, Shenzhen 518000, China*

* Corresponding author. E-mail: hbduan@buaa.edu.cn

Abstract The dynamic resource allocation of unmanned aerial vehicle (UAV) swarms is a challenging key technical problem in mission planning. In this paper, a dynamic resource allocation method based on a cooperative and competitive mechanism is proposed. The average income of the system is improved by designing regulation rules for special individuals. Finally, the feasibility and effectiveness of the proposed method for UAV dynamic resource allocation are verified according to the network structure, model parameters, and steady-state results of the system under different excitation mechanisms through simulation experiments of the UAV game model.

Keywords UAV swarm, public goods game, dynamic resource allocation, cooperative mechanism, competition mechanism