



基于强化竞争学习鸽群优化的多无人机博弈决策

雷阳琦¹, 段海滨^{1,2*}

1. 北京航空航天大学自动化科学与电气工程学院, 仿生自主飞行系统研究组, 北京 100083;

2. 鹏城实验室, 深圳 518000

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

收稿日期: 2022-01-23; 接受日期: 2022-04-13; 网络版发表日期: 2023-04-14

科技创新2030-“新一代人工智能”重大项目(编号: 2018AAA0100803)和国家自然科学基金项目(批准号: U20B2071, T2121003, 91948204, U1913602)资助

摘要 多无人机(multi-unmanned aerial vehicle, multi-UAV)博弈决策是无人机博弈对抗领域的关键性技术难题。本文提出了一种强化竞争学习鸽群优化(enhanced competitive learning pigeon-inspired optimization, ECPIO)算法, 以解决多无人机博弈决策问题。建立了六自由度无人机模型以及无人机博弈态势评估模型, 据此计算博弈双方无人机的收益矩阵。采用非合作博弈模型, 将多无人机博弈决策问题建模为基于混合策略纳什均衡的目标函数寻优问题, 采用强化竞争学习鸽群优化求取目标函数的最优解, 通过引入强化竞争学习策略可有效降低优化结果陷入局部最优的概率。最后, 通过将所提出的强化竞争学习鸽群优化算法与基本鸽群优化、基本粒子群优化、竞争粒子群优化以及反向学习粒子群优化算法的仿真对比实验, 验证了所提方法解决多无人机博弈决策问题的可行性与有效性。

关键词 多无人机, 博弈决策, 鸽群优化, 强化竞争学习

1 引言

无人机(unmanned aerial vehicle, UAV)执行空中打击任务已经成为未来空中战场的主要作战趋势^[1]。然而, 随着无人机智能化、自主化、网络化、集群化、实战化技术的发展, 单架无人机由于受到探测能力、武器载荷等因素的限制, 难以完成复杂的作战任务。采用多无人机协同作战方式, 能有效提高无人机的战场生存能力以及多无人机整体的作战效能^[2]。

多无人机博弈决策的核心为建立博弈双方无人机收益模型, 使得博弈过程中我方无人机群损失最小, 敌

方无人机群毁伤最大。为解决多无人机博弈决策问题, 文献^[3]提出了一种基于狼群智能的多无人机协同多目标攻防满意决策算法, 该方法兼顾博弈双方态势优势和收益代价, 然而未考虑任务分配后的无人机运动控制问题。文献^[4]提出了一种基于博弈论的多无人机攻击目标分配方法, 采用滑模控制理论实现无人机编队运动控制, 但未考虑无人机运动过程中的博弈情形。多无人机博弈决策是多约束条件下的非确定性多项式完全(non-deterministic complete, NPC)问题, 属于任务分配在无人机博弈对抗领域的应用, 常用的任务分配方法主要包括分布式分配算法与集中式分配算法^[5]。

引用格式: 雷阳琦, 段海滨. 基于强化竞争学习鸽群优化的多无人机博弈决策. 中国科学: 技术科学, 2024, 54: 136–148

Lei Y Q, Duan H B. Decision-making of multi-UAV combat game via enhanced competitive learning pigeon-inspired optimization (in Chinese). Sci Sin Tech, 2024, 54: 136–148, doi: [10.1360/SST-2022-0032](https://doi.org/10.1360/SST-2022-0032)

分布式分配算法通常包括基于市场机制的拍卖算法和合同网算法^[5-8], 基本思想是将任务视作商品, 任务分配视为拍卖活动, 通过“拍卖-竞标-中标”竞争机制实现分布式系统内执行者与任务的最优匹配. Choi等人^[5]提出了一种以资源为导向的一致性拍卖算法, 用于求解多机器人的任务分配问题. Zhen等人^[6]提出了一种改进合同网的分布式分配算法, 解决了异构无人机集群对抗中的协同目标分配问题. Yan等人^[7]结合协同粒子群优化(particle swarm optimization, PSO)算法与合同网方法提出了一种无人机实时任务分配决策算法. 然而上述分布式算法的分配对象均为静止目标, 不适用于多无人机博弈对抗场景. 分布式算法通常应用于具有中小规模数据量的分配决策问题, 且算法不容易收敛至全局最优解^[9]. 考虑到多无人机博弈决策时需要评估双方态势, 处理大规模决策信息, 因此采用集中式分配算法更为合理. 集中式算法将决策问题转化为目标函数的寻优问题, 常用的优化算法分为传统优化算法和启发式优化算法. 传统优化算法包括凸优化^[10]以及梯度下降法^[11], Yang等人^[10]提出了一种基于凸优化的快速最优路径生成算法, Chai等人^[11]设计了一种用于求解路径规划问题的改进梯度下降算法. 然而, 传统优化算法的实时性与优化准确性往往较差, 时间复杂度较高, 因此很难适用于求解多无人机博弈决策问题. 与传统优化算法不同, 启发式优化算法不依赖于问题的数学性质以及初值, 适合求解复杂的高维数学问题. 启发式优化算法主要包括差分进化(differential evolutionary, DE)算法^[12]、粒子群优化算法^[13]以及鸽群优化(pigeon-inspired optimization, PIO)算法^[14]等. 近年来许多学者尝试将启发式优化算法应用于无人机博弈问题, Huang等人^[15]采用改进差分进化算法解决无人机博弈机动轨迹优化问题, Duan等人^[16]提出了一种基于改进鸽群优化的多无人机协同搜索攻击任务分配方法.

鸽群优化是一种模拟自然界中鸽群归巢行为的启发式群体智能优化算法^[17]. 近年来, 鸽群优化因其收敛速度快、控制参数简洁、鲁棒性强等特点已被广泛应用于无人机任务规划、电力电子、图像处理、机器人等诸多领域, 如无人机编队控制器参数整定^[18]、多无人机协同路径规划^[19]以及大规模无人机聚类优化^[20]等. 然而与其他启发式智能优化算法相似, 基本的鸽群优化算法也存在过早收敛、易陷入局部最优解

等局限性. 针对无人机博弈决策难题, 建立了六自由度非线性无人机模型与态势评估模型, 并基于混合策略纳什均衡理论建立了多无人机非合作博弈模型, 将多无人机博弈决策问题转化为集中式优化求解目标函数的全局最优问题. 为了克服基本鸽群优化算法的不足, 本文提出了一种基于强化竞争学习机制(enhanced competitive learning strategy)的改进鸽群优化算法, 该改进策略可增强种群多样性, 以提高全局搜索能力.

2 无人机及态势评估建模

2.1 无人机动力学模型

以F16飞机模型数据为例, 建立无人机六自由度非线性简化模型^[21], 具体如下:

$$\begin{cases} \dot{V} = \frac{uu + vv + ww}{V}, \\ \dot{\alpha} = \frac{uw - wu}{u^2 + w^2}, \\ \dot{\beta} = \frac{vV - vV}{V^2 \cos \beta}, \end{cases} \quad (1)$$

$$\begin{cases} \dot{p} = (c_1 r + c_2 p)q + c_3 \bar{L} + c_4 \bar{N}, \\ \dot{q} = c_5 p r - c_6 (p^2 - r^2) + c_7 \bar{M}, \\ \dot{r} = (c_8 p - c_2 r)q + c_4 \bar{L} + c_9 \bar{N}, \end{cases} \quad (2)$$

$$\begin{cases} \dot{\phi} = p + \tan \theta (r \cos \phi + q \sin \phi), \\ \dot{\theta} = q \cos \phi - r \sin \phi, \\ \dot{\psi} = (r \cos \phi + q \sin \phi) / \cos \theta, \end{cases} \quad (3)$$

$$\begin{cases} \dot{x} = u \cos \theta \cos \psi + v (\sin \phi \sin \theta \cos \psi - \cos \phi \sin \psi) \\ \quad + w (\sin \phi \sin \psi + \cos \phi \sin \theta \sin \psi), \\ \dot{y} = u \cos \theta \sin \psi + v (\sin \phi \sin \theta \sin \psi + \cos \phi \cos \psi) \\ \quad + w (-\sin \phi \cos \psi + \cos \phi \sin \theta \sin \psi), \\ \dot{z} = -u \sin \theta + v \sin \phi \cos \theta + w \cos \phi \cos \theta. \end{cases} \quad (4)$$

以上12个微分方程为无人机六自由度非线性动力学方程, 其中

$$c_1 = \frac{(I_y - I_z)I_z - I_{xz}^2}{I_x I_z - I_{xz}^2}, \quad c_2 = \frac{(I_x - I_y + I_z)I_{xz}}{I_x I_z - I_{xz}^2}, \quad c_3 = \frac{I_z}{I_x I_z - I_{xz}^2},$$

$$c_4 = \frac{I_{xz}}{I_x I_z - I_{xz}^2}, \quad c_5 = \frac{I_z - I_x}{I_y}, \quad c_6 = \frac{I_{xz}}{I_y}, \quad c_7 = \frac{1}{I_y},$$

$$c_8 = \frac{(I_x - I_y)I_x - I_{xz}^2}{I_x I_z - I_{xz}^2} \text{ 以及 } c_9 = \frac{I_x}{I_x I_z - I_{xz}^2}, \quad I_x, I_y, I_z \text{ 为无人机}$$

绕机体坐标系下x轴、y轴以及z轴的转动惯量, I_{xz} 为惯性积. $[u, v, w]^T$ 为无人机飞行速度 V 在机体坐标系下各

轴上的分量, 描述无人机模型的状态向量为 $[V, \alpha, \beta, p, q, r, \phi, \theta, \psi, x, y, z]$, 其中 V 为空速, α 为迎角, β 为侧滑角, p, q, r 分别表示滚转、俯仰、偏航速度, ϕ, θ, ψ 分别表示滚转角、俯仰角和偏航角, x, y, z 表示无人机相对于惯性坐标系的位置信息, L, M, N 分别表示滚转力矩、俯仰力矩以及偏航力矩.

2.2 无人机态势评估建模

考虑参与博弈的敌我双方无人机均配备有探测雷达和武器导弹, 当敌方无人机进入我方无人机探测范围后, 博弈双方无人机基于态势评估结果进行目标分配决策, 态势评估主要考虑因素包括无人机攻击优势以及作战效能优势, 其中攻击优势有角度优势、距离优势以及能量优势, 作战效能优势与无人机的机动能力、打击能力、探测能力等因素密切相关^[22].

(1) 角度优势

假设我方无人机导弹的最大发射角为 $\varphi_{M\max}$, 当敌方无人机在我方无人机最大发射角之外时, 我方无人机角度优势为零. 定义我方无人机导弹不可逃逸角为 $\varphi_{Mk\max}$, 当敌方无人机位于我方无人机最大不可逃逸角度内时, 我方无人机角度优势最大. 角度优势函数定义如下:

$$T_\varphi = \begin{cases} 0, & |\varphi| > \varphi_{M\max}, \\ e^{\frac{|\varphi| - \varphi_{Mk\max}}{\varphi_{M\max} - \varphi_{Mk\max}}}, & \varphi_{Mk\max} \leq |\varphi| \leq \varphi_{M\max}, \\ 1, & 0 \leq |\varphi| \leq \varphi_{Mk\max}. \end{cases} \quad (5)$$

(2) 距离优势

假设我方无人机导弹最大攻击距离为 $D_{FM\max}$, 当敌方无人机位于我方无人机最大攻击距离外时, 我方无人机距离优势为零. 当敌方无人机间位于我方无人机导弹的最小不可逃逸距离 $D_{FMk\min}$ 内与最大不可逃逸距离 $D_{FMk\max}$ 之间时, 我方无人机距离优势最大. 距离优势函数定义如下:

$$T_d = \begin{cases} 0, & d > D_{FM\max}, \\ e^{\frac{d - D_{FMk\max}}{D_{FM\max} - D_{FMk\max}}}, & D_{FMk\max} \leq d \leq D_{FM\max}, \\ 1, & D_{FMk\min} \leq d \leq D_{FMk\max}, \\ e^{\frac{d - D_{FMk\min}}{D_{FMk\max} - D_{FMk\min}}}, & 0 \leq d \leq D_{FMk\min}. \end{cases} \quad (6)$$

(3) 能量优势

博弈过程中无人机单位能量表示如下:

$$E = H + \frac{V^2}{2g}. \quad (7)$$

定义我方无人机能量为 E_R , 敌方无人机能量为 E_B , 能量威胁指数为

$$T_E = \begin{cases} 1, & 2E_B \leq E_R, \\ 1 + \frac{0.9(E_B - E_R)}{1.5E_B}, & 0.5E_B \leq E_R, \\ 0.1, & E_R < 0.5E_B. \end{cases} \quad (8)$$

综上, 无人机综合态势优势表达式如下:

$$T_s = w_1(T_\varphi)^{\gamma_1}(T_d)^{\gamma_2} + w_2T_E, \quad (9)$$

式中, $0 \leq \gamma_1 \leq 1$, $0 \leq \gamma_2 \leq 1$, $\gamma_1 + \gamma_2 = 1$, $0 \leq w_1 \leq 1$, $0 \leq w_2 \leq 1$, $w_1 + w_2 = 1$.

(4) 作战效能

无人机目标分配决策不仅需要考虑双方态势优势, 还需要考虑无人机作战效能. 无人机作战效能评估表达式为

$$C = [\ln B + \ln(A_1) + \ln(A_2)] \cdot k_1 \cdot k_2 \cdot k_3 \cdot k_4, \quad (10)$$

式中, B 为无人机的机动性, A_1 为无人机的火力参数, A_2 为探测能力参数, k_1 为操纵性能参数, k_2 为生存能力参数, k_3 为航程系数, k_4 为电子对抗能力系数. 无人机效能优势函数表达式如下:

$$T_C = \begin{cases} 0, & C_R < 0.3C_B, \\ 0.25, & 0.3C_B \leq C_R < C_B, \\ 0.5, & C_R = C_B, \\ 0.75, & C_B < C_R \leq 1.5C_B, \\ 1, & C_R \geq 1.5C_B. \end{cases} \quad (11)$$

2.3 无人机机动作控制指令

建立基于离散动作空间的无人机机动作库, 无人机博弈本质为选择当前态势下收益最大的机动作. 为生成六自由度模型下的机动作控制指令, 首先将六自由度模型简化为三自由度模型如下:

$$\begin{cases} \frac{dx}{dt} = v \cos \mu \cos \varphi, \\ \frac{dy}{dt} = v \cos \mu \sin \varphi, \\ \frac{dz}{dt} = v \sin \mu, \\ \frac{dv}{dt} = g(n_x - \sin \mu), \\ \frac{d\mu}{dt} = \frac{g}{v}(n_f \cos \gamma - \cos \mu), \\ \frac{d\varphi}{dt} = \frac{g}{v \cos \mu} n_f \sin \gamma, \end{cases} \quad (12)$$

式中, $[x, y, z]$ 为无人机位置矢量, v 为无人机速度, μ 为无人机航迹角, φ 为无人机航迹倾角, n_x 为控制无人机速度的切向过载, n_f 为法向过载, γ 为速度滚转角. 三自由度无人机模型控制量为 $[n_x, n_f, \gamma]$, $[x, y, z, v, \mu, \varphi]$ 为状态变量.

基于法向过载 n_f 和滚转角 γ 设计离散基本机动动作库, 基本机动动作定义如图1所示, 法向过载指令取值 $n_f = [n_{f1}, n_{f2}, \dots, n_{fu}]$, 滚转角指令取值 $\gamma = [\gamma_1, \gamma_2, \dots, \gamma_w]$, 组合生成的基本动作库如下:

$$L = \left\{ (\gamma_1 \times n_{f1}), (\gamma_1 \times n_{f2}), \dots, (\gamma_1 \times n_{fu}), \dots, (\gamma_w \times n_{f1}), (\gamma_w \times n_{f2}), \dots, (\gamma_w \times n_{fu}) \right\}. \quad (13)$$

这里, 将三自由度模型指令转化为包含迎角 α 和滚转角 ϕ 的六自由度无人机模型控制指令, 其中 $\phi = \gamma$, $\alpha = kn_f$, 即可生成用于无人机决策运动控制的六自由度指令.

3 多无人机博弈决策

3.1 问题描述

如图2所示, 考虑我方 M 架无人机对敌方 N 架无人机的博弈决策, 基于非合作博弈模型, 我方视为博弈参与者的红方, 敌方视为蓝方, 问题可描述如下:

$$G = \langle \text{UAVs}, \{\text{strategy}_{\text{red}}, \text{strategy}_{\text{blue}}\}, \{\text{payoff}_{\text{red}}, \text{payoff}_{\text{blue}}\} \rangle, \quad (14)$$

式中, $\text{UAVs} = \{\text{red}, \text{blue}\}$ 为博弈参与者, $\text{strategy}_{\text{red}}$ 和 $\text{strategy}_{\text{blue}}$ 为红蓝双方博弈的纯策略集, 红蓝双方博弈

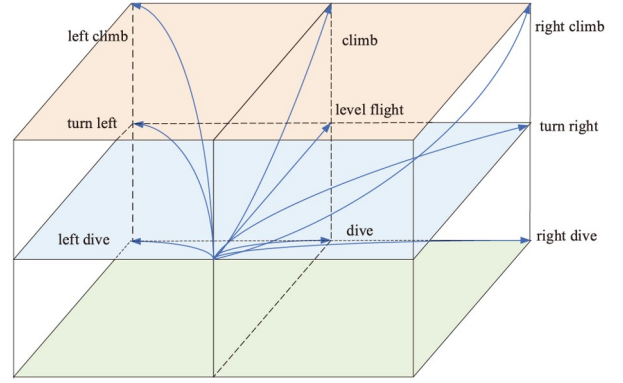


图1 (网络版彩图)无人机基本机动动作
Figure 1 (Color online) The basic maneuvers of UAV.

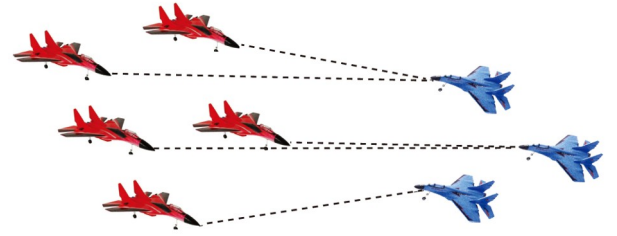


图2 (网络版彩图)多无人机攻击目标分配
Figure 2 (Color online) Attack target allocation of multiple UAVs.

矩阵分别为 $\text{payoff}_{\text{red}}$ 和 $\text{payoff}_{\text{blue}}$. 红方纯策略集数目为 num_{red} , 蓝方纯策略集数目为 num_{blue} , 红方博弈矩阵表达式如下, 其中 $J_{i,j}$ 表示红方采用 i 纯策略与蓝方采用 j 纯策略博弈所得收益值:

$$\text{payoff}_{\text{red}} = \begin{bmatrix} J_{1,1} & J_{1,2} & \cdots & J_{1,\text{num}_{\text{blue}}} \\ J_{2,1} & J_{2,2} & \cdots & J_{2,\text{num}_{\text{blue}}} \\ \vdots & \vdots & \ddots & \vdots \\ J_{\text{num}_{\text{red}}+1} & J_{\text{num}_{\text{red}}+2} & \cdots & J_{\text{num}_{\text{red}},\text{num}_{\text{blue}}} \end{bmatrix}. \quad (15)$$

无人机博弈过程中, 博弈双方均尽可能减小己方损伤概率并增大对敌方的毁伤概率. 因此, 定义无人机博弈收益值为

$$J_{i,j} = \frac{pb_i}{pr_j}, \quad (16)$$

式中, pb_i 为蓝方被红方采用 i 策略攻击后的损伤概率, pr_j 为红方被蓝方采用 j 策略攻击后的损伤概率. pr_j 的表达式如下:

$$pr_j = \sum_{m=1}^M p_attack_{nm} \cdot p_damage_{nm} \cdot value_m, \quad (17)$$

式中, p_attack_{nm} 为红方第 m 架无人机被蓝方第 n 架无人机攻击的概率, p_damage_{nm} 为红方第 m 架无人机被蓝方第 n 架无人机毁伤概率, $value_m$ 为第 m 架无人机的价值. 无人机攻击概率与毁伤概率的计算公式如下:

$$p_attack_{nm} = \begin{cases} 1, & T_{nm} > T_{\text{threshold}}, \\ \exp(T_{nm} - T_{\text{threshold}}), & T_{nm} \leq T_{\text{threshold}}, \end{cases} \quad (18)$$

式中, T_{nm} 为第 2.2 节计算所得蓝方第 n 架无人机相对于红方第 m 架无人机的态势优势值, $T_{\text{threshold}}$ 为态势函数阈值, 当态势优势高于该阈值时, 认为无人机发起攻击概率为 1.

$$p_damage_{nm} = \rho \cdot p_{\text{kill}}, \quad (19)$$

式中, ρ ($0 < \rho < 1$) 表示环境影响因子, p_{kill} 为蓝方第 n 架无人机对红方第 m 架无人机的损伤概率.

3.2 目标函数

在非合作博弈问题中, 纯策略纳什均衡问题不一定存在平衡点, 然而混合策略纳什均衡问题一定存在平衡解. 假设 S_{red} 和 S_{blue} 为红蓝双方混合策略为

$$S_{\text{red}} = \left\{ (s_1^r, s_2^r, \dots, s_N^r) \mid s_i^r \geq 0, \sum_{i=1}^N s_i^r = 1 \right\}, \quad (20)$$

$$S_{\text{blue}} = \left\{ (s_1^b, s_2^b, \dots, s_M^b) \mid s_j^b \geq 0, \sum_{j=1}^M s_j^b = 1 \right\}, \quad (21)$$

式中, s_i^r 为红方无人机选择 i 纯策略的概率, s_j^b 表示蓝方无人机选择 j 纯策略的概率. 红蓝双方在混合策略下的收益期望值为

$$E_{\text{red}} = S_{\text{red}} \cdot \text{payoff}_{\text{red}} \cdot S_{\text{blue}}^T, \quad (22)$$

$$E_{\text{blue}} = S_{\text{blue}} \cdot \text{payoff}_{\text{blue}} \cdot S_{\text{red}}^T. \quad (23)$$

定义 n 个参与者的混合策略表示为 $x = \{x_1, x_2, \dots, x_n\}$, 对于任何一个博弈参与者 i , 若存在混合策略 $x^* = \{x_1^*, x_2^*, \dots, x_n^*\}$ 满足^[23]:

$$f_i(x^*) \geq f_i(x_i, x_{-i}^*), \quad (24)$$

则称混合策略 x^* 为纳什均衡. f_i 为参与者 i 的收益值,

$x_{-i}^* = \{x_1, x_2, \dots, x_{i-1}, x_{i+1}, \dots, x_n\}$ 为所有参与者除去参与者 i 的混合策略. 因此混合策略纳什均衡存在的充分必要条件为

$$f_i(x^*) \geq f_i(s^i, x_{-i}^*). \quad (25)$$

对于红蓝双方双矩阵对策情形, 混合策略记作 $\{S_{\text{red}}^*, S_{\text{blue}}^*\}$, 纳什均衡的充分必要条件为

$$E_{\text{red}}(S_{\text{red}}^*, S_{\text{blue}}^*) \geq \text{payoff}_i^{\text{red}} \cdot (S_{\text{blue}}^*)^T, \quad (26)$$

$$E_{\text{blue}}(S_{\text{red}}^*, S_{\text{blue}}^*) \geq S_{\text{red}}^* \cdot \text{payoff}_j^{\text{blue}}, \quad (27)$$

式中, $\text{payoff}_i^{\text{red}}$ 为红方收益矩阵的第 i 行向量, $\text{payoff}_j^{\text{blue}}$ 为蓝方收益矩阵第 j 列向量.

为求解混合策略纳什均衡, 设计混合策略纳什均衡的目标函数为

$$f(S_{\text{red}}, S_{\text{blue}}) = \max \left\{ \max_{1 \leq i \leq n} \left\{ \text{payoff}_i^{\text{red}} \cdot (S_{\text{blue}})^T - E_{\text{red}}(S_{\text{red}}, S_{\text{blue}}) \right\}, 0 \right\} + \max \left\{ \max_{1 \leq j \leq m} \left\{ S_{\text{red}} \cdot \text{payoff}_j^{\text{blue}} - E_{\text{blue}}(S_{\text{red}}, S_{\text{blue}}) \right\}, 0 \right\}. \quad (28)$$

由此, 求解无人机博弈决策的混合策略纳什均衡问题可转化为目标函数优化问题, 使用强化竞争学习鸽群优化算法求得最优解. 鸽群优化所得目标函数值越接近于 0, 则混合策略越接近纳什均衡点.

3.3 强化竞争学习鸽群优化

3.3.1 鸽群优化

在鸽群优化算法中, 鸽群在归巢的第一个阶段依靠太阳与磁场信息进行导航, 当鸽群逐渐靠近目的地时依靠地标信息进行导航. 受鸽群归巢行为的启发, 鸽群优化算法分别引入了地图、指南针算子与地标算子作为更新鸽子位置的导航工具. 每只鸽子的位置代表优化问题的潜在解, 鸽群的巢穴代表最优解.

假设 M 只鸽子处在 D 维搜索空间中, 第 i 只鸽子在第 t 次迭代时的位置和速度分别记作 $x_i(t) = (x_{i1}(t), x_{i2}(t), \dots, x_{iD}(t))$ 和 $v_i(t) = (v_{i1}(t), v_{i2}(t), \dots, v_{iD}(t))$. 鸽子 i 基于全局最优鸽子的位置 x_{gbest} 更新速度和位置信息, 直到当前迭代次数 t 达到地图、指南针算子阶段最大迭代次数 NC_{max}^1 .

$$v_i(t+1) = v_i(t) \cdot e^{-R(t+1)} + rand \cdot (x_{gbest}(t) - x_i(t)), \quad (29)$$

$$x_i(t+1) = x_i(t) + v_i(t+1), \quad (30)$$

式中, R 为地图、指南针因子, 表示前一时刻的速度对当前时刻速度的影响, $rand$ 为0到1之间均匀分布的随机数. 随着鸽群逐渐接近目的地, 太阳与指南针的影响逐渐降低, 鸽群依赖目的地附近的地标信息进行导航.

3.3.2 强化竞争学习鸽群优化

在鸽群优化的地图、指南针算子中, 鸽子 i 基于全局最优鸽子位置 x_{gbest} 进行速度与位置的更新, 然而这将导致算法过早收敛以及陷入局部最优解. 受启发于竞争学习机制^[23], 为平衡算法的快速收敛性与全局搜索能力, 在鸽群优化的基础上引入了一种竞争强化学习机制用于地图、指南针算子阶段.

(1) 地图、指南针算子

竞争学习机制使种群中的个体两两竞争, 按照适应度值的大小将个体分为获胜者winner与失败者loser, 失败者通过向获胜者 x_w 学习更新其位置和速度信息, 获胜者在本轮迭代中不再进行速度与位置的更新. 竞争学习机制使得每次迭代过程中仅有50%个体学习更新, 因此算法的搜索效率较低. 为增强种群的多样性与算法的搜索能力, 采用强化竞争的策略对鸽群中的每个个体进行速度与位置的更新. 在每代鸽群的生成过程中, 首先将鸽群随机分为数量相等的三个子群, 每轮竞争学习时, 从三个子群中各随机选出一个个体参与竞争. 竞争结果由适应度值确定, 适应度值最高的鸽子称为获胜者winner, 适应度值中等的个体为失败者loser₁, 适应度值最低的个体为失败者loser₂. loser₁根据式(31)和(32)更新速度和位置, 学习对象包括全局最优 x_{gbest} 和所在子群最优 x_{sbest} , 该更新方式保持了鸽群优化快速收敛的性质. loser₂向winner学习, 并依据式(33)和(34)更新. winner向 x_{gbest} 学习, 与鸽群优化更新方式相同, 如式(35)和(36)所示.

$$v_{i1}(t+1) = v_{i1}(t) \cdot e^{-R(t+1)} + rand_1 \cdot (x_{sbest}(t) - x_{i1}(t)) + rand_2 \cdot (x_{gbest}(t) - x_{i1}(t)), \quad (31)$$

$$x_{i1}(t+1) = x_{i1}(t) + v_{i1}(t+1), \quad (32)$$

$$v_{i2}(t+1) = v_{i2}(t) \cdot e^{-R(t+1)} + rand_3 \cdot (x_w(t) - x_{i2}(t)), \quad (33)$$

$$x_{i2}(t+1) = x_{i2}(t) + v_{i2}(t+1), \quad (34)$$

$$v_w(t+1) = v_w(t) \cdot e^{-R(t+1)} + rand_4 \cdot (x_{gbest}(t) - x_w(t)), \quad (35)$$

$$x_w(t+1) = x_w(t) + v_w(t+1), \quad (36)$$

式中, 参数 R 为地图、指南针因子, $rand_1$, $rand_2$, $rand_3$ 和 $rand_4$ 均为0到1之间均匀分布的随机数, 地图、指南针算子阶段最大迭代次数为 NC_{max}^1 .

(2) 地标算子

随着鸽群抵近目标, 太阳和磁场信息不断减弱鸽群依赖地标信息进行导航. 在更新过程中, 部分鸽子会因为不熟悉地标而被抛弃. 因此, 每次迭代中鸽群的数量会减少50%, 鸽子 i 根据地标算子 x_{center} 更新位置, 其含义为鸽群在 D 维搜索空间的平均位置. 鸽子更新后位置 x_i 的计算公式如下:

$$x_i(t+1) = x_i(t) + rand \cdot (x_{center}(t) - x_i(t)), \quad (37)$$

$$x_{center}(t) = \frac{\sum_{i=1}^{M(t)} x_i(t) \cdot fitness(x_i(t))}{M(t) \cdot \sum_{i=1}^{M(t)} fitness(x_i(t))}, \quad (38)$$

$$M(t) = \text{ceil}\left(\frac{M(t-1)}{2}\right), \quad (39)$$

式中, t 为当前迭代次数, 两个算子总最大迭代次数为 NC_{max} , $rand$ 为 $[0, 1]$ 之间均匀分布的随机数, $fitness(x_i(t))$ 为个体适应度值, ceil 函数为取整函数.

在强化竞争学习机制中, 参与竞争的三只鸽子均来自于不同的鸽子子群, 向鸽子子群中最优个体学习可增强鸽群中各鸽子之间的交互. 每轮竞争中无论获胜者还是失败者, 所有鸽子的位置速度都将得到更新, 极大地提升了算法的搜索效率. 且由于分组机制降低了每次迭代需要更新计算的次数, 算法有利于计算大规模优化问题, 因此适用于求解多纯策略集时的混合策略纳什均衡问题.

定义目标函数的复杂度为 L_f , 总迭代次数为 N_c , 则鸽群优化地图、指南针算子与地标算子阶段的时间复杂度分别为 $O(M(D + L_f))$ 和 $O(M \log M + D \log M + L_f \log M)$. 与鸽群优化相比, 强化竞争学习鸽群优化两个阶段算子的时间复杂度均未发生变化, 因此ECPIO算法的时间复杂度为 $O(N_c(M \log M + DM + L_f M))$.

3.4 基于ECPIO的多无人机博弈决策

ECPIO算法优化得到的最优解为 $x_{gbest} = \{x_1, x_2, \dots, x_D\}$, 记 x_j 为混合策略中概率最大的纯策略, 表示红蓝双方多无人机攻击目标决策的结果. 假设决策结果为红方第 m 架无人机与蓝方第 n 架无人机博弈.

依据建模部分第2.3节红方无人机机动策略集为 $R = \{r_1, r_2, \dots, r_m\}$, 蓝方无人机机动策略集为 $B = \{b_1, b_2, \dots, b_n\}$. 红蓝双方无人机分别计算每种策略集下可达的状态向量. 假设红蓝双方无人机间距离为 D , 红方无人机目标视线角为 α , 蓝方无人机目标方位角为 β , 定义角度优势得分函数为

$$S_A = 1 - \frac{\alpha + \beta}{\pi}. \quad (40)$$

定义距离优势得分函数为

$$S_R = e^{-\frac{(D-d)^2}{2\sigma^2}}, \quad (41)$$

式中, d 为双方无人机武器可攻击的平均距离, σ 为标

准差, 则优势得分函数为

$$S = S_A \cdot S_R. \quad (42)$$

分别计算每种策略下红方状态与蓝方状态的得分, 得到博弈矩阵如下:

$$S = \begin{bmatrix} S_{11} & S_{12} & \cdots & S_{1n} \\ S_{21} & S_{22} & \cdots & S_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ S_{m1} & S_{m2} & \cdots & S_{mn} \end{bmatrix}. \quad (43)$$

蓝方无人机选择使其获得优势分数最大的策略 b_{\max} , 红方无人机按照最大最小博弈原则选取蓝方最不利于红方情况下, 红方得分最高的策略 r_{\max} .

$$b_{\max} = \max_j(S_{ij}), \quad (44)$$

$$r_{\max} = \max_i(\min_j(S_{ij})). \quad (45)$$

综上, 基于强化竞争学习鸽群优化的多无人机博弈决策算法具体实现流程如图3所示.

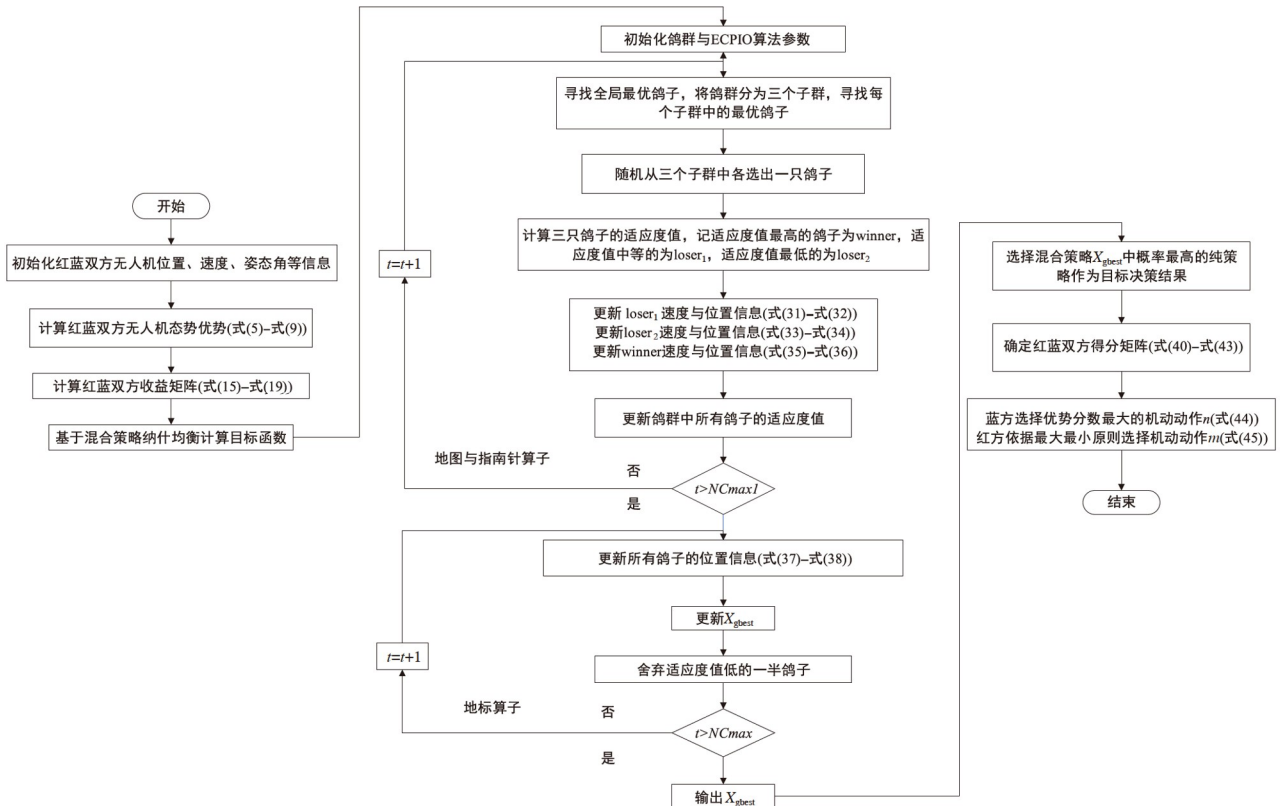


图3 基于ECPIO算法的多无人机博弈决策具体实现流程

Figure 3 Flowchart of multiple UAVs combat game decision-making based on ECPIO algorithm.

4 仿真实验对比分析

4.1 参数设置

本节通过两个场景下的仿真实验分别验证所提出的博弈目标分配算法, 以及基于目标分配结果的多无人机博弈决策算法的可行性与有效性。

场景1: 红蓝双方各四架无人机参与博弈, 红方无人机进行攻击目标分配, 实验设置如表1所示, 包含红蓝双方所有无人机初始时刻的位置、速度和姿态角信息, 假设红方每一架无人机仅可攻击一架蓝方无人机。

场景2: 红方四架无人机与蓝方两架无人机进行博弈, 实验设置如表2所示, 包括红蓝双方无人机初始时刻的位置、速度以及姿态角信息。红方无人机通过目标分配与蓝方无人机两组形成“二对一”的追逃博弈形式。

4.2 实验对比结果分析

场景1: 假设红方每一架无人机仅可攻击一架蓝方无人机, 根据纳什均衡理论红蓝双方各有24种分配纯

策略, 使用ECPIO算法求解混合策略纳什均衡选取红方收益最大的攻击目标分配策略。为选取合适的ECPIO算法参数, 本节设计仿真实验研究了鸽群数量 N_p 以及地图、指南针因子 R 对算法优化性能的影响, 实验参数设计如表3所示, 仿真总迭代次数 $NC_{\max} = 100$, 地图、指南针算子阶段迭代次数 $NC_{\max}^1 = 70$ 。

鸽群数量对ECPIO算法优化性能的影响如图4所示, N_p 依次取值为90, 120, 150, 210, 300, 450, 种群数量的增加丰富了算法的解空间使得算法的优化性能得到增强。鸽群数量 N_p 取值为300时优化效果最好, 随着鸽群数量进一步增加, ECPIO优化能力减弱。证明鸽群数量参数存在极限, 当鸽群数量值超过极限时, 优化能力减弱, 同时过大的种群数量会导致耗费较大的计算资源。图5描述了地图、指南针因子 R 取值对ECPIO优化性能的影响, 当 R 取值为0.02时算法优化性能最佳。理论上地图、指南针因子越小算法的搜索性能越高, 然而过小的 R 值将会导致算法陷入局部最优解。

为验证ECPIO算法的有效性与优越性, 将ECPIO

表1 场景1红蓝双方无人机初始信息

Table 1 Initial conditions of red and blue sides in case 1

无人机	x (m)	y (m)	z (m)	θ (rad)	φ (rad)	ψ (rad)	v (m/s)
red ₁	1000	1000	10000	0	0	0	240
red ₂	1000	2000	8000	0	0	0	250
red ₃	1000	3000	6000	0	0	0	280
red ₄	1000	4000	4000	0	0	0	270
blue ₁	5000	1000	10000	0	0	π	270
blue ₂	5000	2000	8000	0	0	π	280
blue ₃	5000	3000	6000	0	0	π	290
blue ₄	5000	4000	4000	0	0	π	280

表2 场景2红蓝双方无人机初始信息

Table 2 Initial conditions of red and blue sides in case 2

无人机	x (m)	y (m)	z (m)	θ (rad)	φ (rad)	ψ (rad)	v (m/s)
red ₁	5000	1500	6000	0	0	0	240
red ₂	6500	1200	6000	0	0	0	250
red ₃	6700	1200	6000	0	0	0	280
red ₄	8000	1500	6000	0	0	0	270
blue ₁	5800	2000	6000	0	0	0	270
blue ₂	7500	2000	6000	0	0	0	280

表 3 ECPIO算法参数

Table 3 Parameter setting of ECPIO algorithm

参数	方案1	方案2
鸽群数量 N_p	90~450	300
地图、指南针因子 R	0.02	0.005~0.05

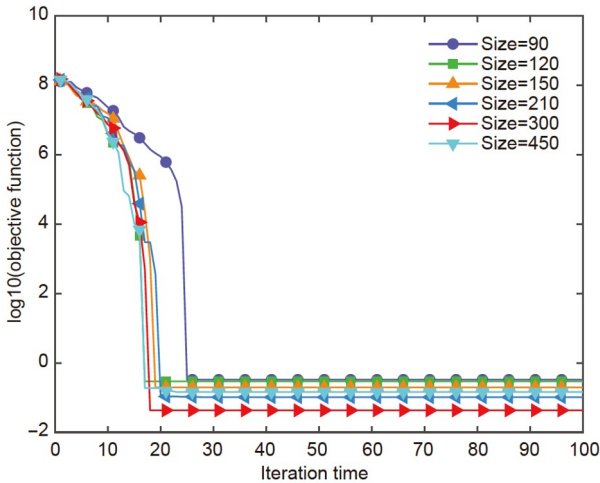


图 4 (网络版彩图) 鸽群数量 N_p 对ECPIO优化性能影响
Figure 4 (Color online) Effect of population size N_p on ECPIO algorithm.

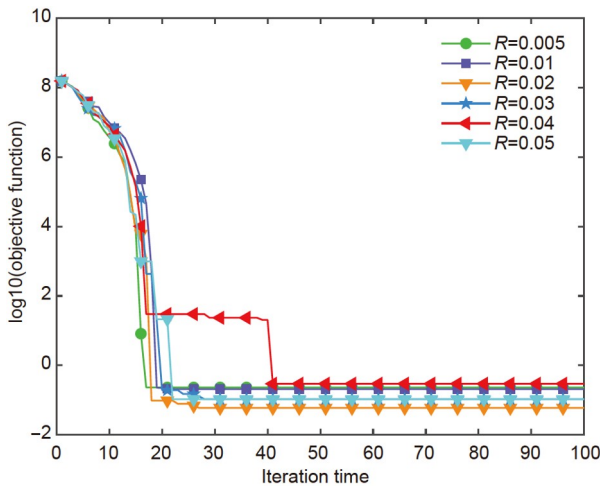


图 5 (网络版彩图) 地图、指南针因子 R 对ECPIO算法优化性能影响
Figure 5 (Color online) Effect of map and compass operator R on the ECPIO algorithm.

算法与现有的四种经典算法进行对比实验, 包括鸽群优化(PIO)算法、粒子群优化(PSO)算法、竞争粒子群

优化(CSO)算法^[23]以及反向学习粒子群优化(OBPSO)算法^[24]. 实验参数设置如表4所示, 设置种群数量 $N_p=300$, 解空间维度 $D=24$, 仿真步长 $NC_{max}=100$, PIO算法地图、指南针算子阶段仿真步长 $NC_{max}^1=70$.

图6给出了ECPIO算法与其余四种优化算法迭代次数与目标函数值的关系, 表5为优化后目标函数取值. ECPIO算法的进化曲线相比于其余四种算法的进化曲线使目标函数达到了更小的值, 证明了强化竞争学习机制改进鸽群优化算法的可行性与有效性. OBPSO算法与CSO算法优化性能明显弱于PIO算法与PSO算法, 证明反向学习策略与竞争学习策略改进的优化算法不适于求解混合策略纳什均衡问题以及数据量较大的大规模优化问题. ECPIO算法在求解混合策略纳什均衡问题时, 收敛速度较快, 其强化竞争的学习机制将鸽群分组竞争, 减少迭代次数适应于解决大规模

表 4 优化算法参数设置

Table 4 Parameter setting of the ECPIO algorithm

算法	参数	取值
PIO	地图、指南针因子 R	0.02
PSO	加速度因子 c_1	1.3
	加速度因子 c_2	1.7
ECPIO	地图、指南针因子 R	1
OBPSO	常数 ϕ	0.3
CSO	常数 ϕ	0.3

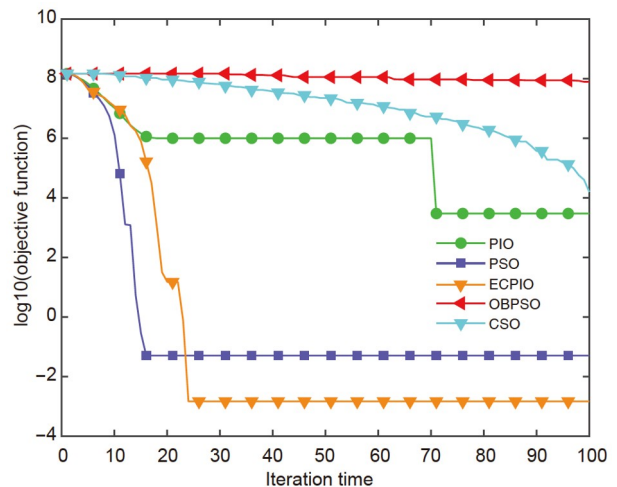


图 6 (网络版彩图) 目标函数进化曲线对比
Figure 6 (Color online) Comparison evolution curves of the objective function.

表5 目标函数最小值

Table 5 Minimum value of objective function

优化算法	最小值
PIO	0.1316
PSO	0.0428
ECPIO	9.664×10^{-4}
OBPSO	7.4315×10^7
CSO	5.2410×10^3

优化问题. 在竞争过程中不同于竞争粒子群算法, 采用三只鸽子对抗的方式, 对获胜者和失败者同时进行更新学习. ECPIO算法使得目标函数收敛至 10^{-4} 级别, 已经非常接近于0即达到纳什均衡点.

使用ECPIO求解混合策略纳什均衡的结果如图7所示, 水平轴从左至右分别表示红蓝双方的24种纯策略, 纵轴表示混合策略中选取每一种纯策略的概率. 根据非合作博弈原则, 红蓝双方均会选择概率最大的纯策略. 因此红方采用第16纯策略, 即红方第一架无人机攻击蓝方第三架无人机, 红方第二架无人机攻击蓝方第二架无人机, 红方第三架无人机攻击蓝方第四架无人机, 红方第四架无人机攻击蓝方第一架无人机. 蓝方采用第9纯策略, 即蓝方第一架无人机攻击红方第二架无人机, 红方第二架无人机攻击蓝方第一架无人机, 红方第三架无人机攻击蓝方第四架无人机, 红方第四架无人机攻击蓝方第三架无人机. 红蓝双方执行混合策略, 红方收益值为1.0455, 蓝方收益值为0.9653. 红方收益值高于蓝方, 因此依据攻击目标分配结果红方获胜的概率较大.

场景2: 根据“二对一”分配原则, 红蓝双方各有6种纯策略可供选择, 采用ECPIO算法优化目标函数求纳什均衡解. 图8为ECPIO算法与PIO算法、PSO算法、CSO算法以及OBPSO算法优化结果对比图, OBPSO算法优化结果依然表明其不适合求解纳什均衡问题, CSO算法优化效果有所改善表明其相对于ECPIO算法不适用于大规模优化问题, 但优化效果仍然不及ECPIO算法.

基于混合策略纳什均衡的分配结果如图9所示, 水平轴从左至右分别表示红蓝双方的6种纯策略, 纵轴表示混合策略中选取每一种纯策略的概率. 因此红方选取第1纯策略, 即红方第一、第二架无人机攻击蓝方

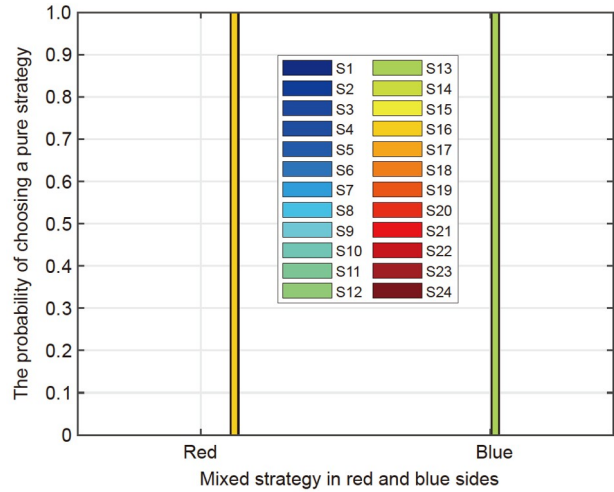


图7 红蓝双方无人机混合策略纳什均衡

Figure 7 Mixed strategies Nash equilibrium of red and blue sides.

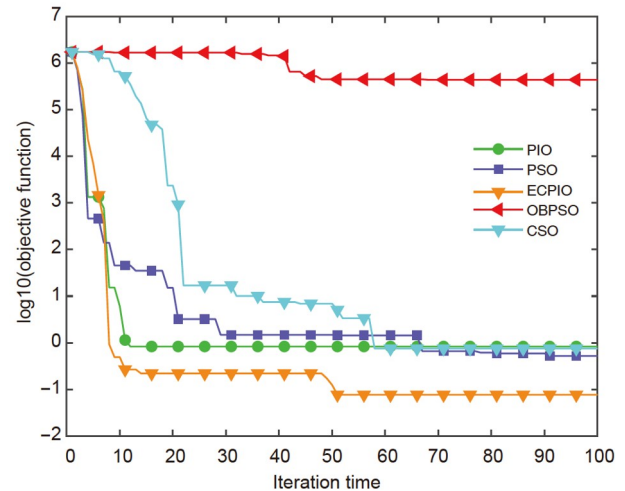


图8 目标函数进化曲线对比

Figure 8 Comparison evolution curves of the objective function.

第一架无人机, 红方第三、四架无人机攻击蓝方第二架无人机; 蓝方采用第3纯策略, 即蓝方第一架无人机与红方第一、三架无人机博弈, 蓝方第二架无人机与红方第二、四架无人机博弈, 红方第三架无人机攻击蓝方第四架无人机, 红方第四架无人机攻击蓝方第三架无人机. 红方收益值为1.2490, 蓝方收益值为0.9175. 红方收益值高于蓝方, 因此依据攻击目标分配结果红方获胜的概率较大.

基于分配结果红方四架无人机分成两个小组对蓝

方两架无人机分别展开追击. 建立红蓝双方离散机动作库, 法向过载 n_f 取值范围为[0.8, 1, 1.2, 1.4, 2], 滚转角 γ 取值范围为[-45°, 0°, 45°], 形成了如表6所示的由15种机动作组成的动作库.

最后, 采用基于Windows 10操作系统的三维视景仿真平台获取无人机实时飞行数据, 再现无人机动态博弈全过程. 三维视景平台由两台计算机构成分别作为人机交互界面与动态展示界面. 通过计算机1的交互界面可获得飞行数据, 同时计算机2可将获取的飞行数据进行再现, 形成动态飞行轨迹. 图10(a)与(b)分别展示了无人机博弈的俯视图以及环绕视角视图. 仿真时长共20 s, 每隔1 s双方无人机选取一次机动作. 以红方一号无人机为例, 为追击蓝方一号无人机红方

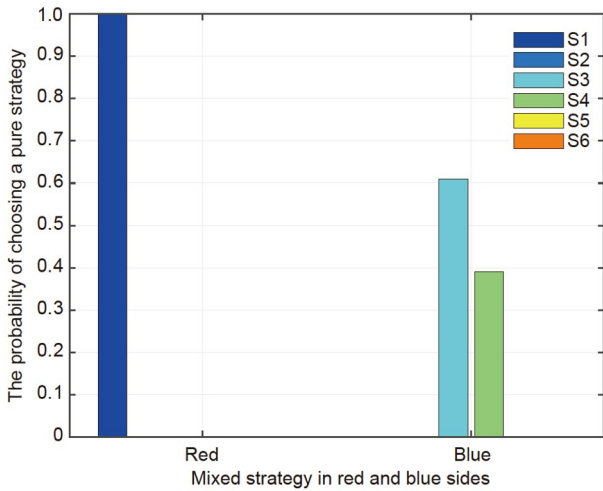


图9 红蓝双方无人机混合策略纳什均衡
Figure 9 Mixed strategies Nash equilibrium of red and blue sides.

表6 机动作库
Table 6 Maneuver library

编号	n_f	γ	编号	n_f	γ
1	0.8	-45°	9	1.4	45°
2	1	0	10	2	-45°
3	1.2	45°	11	0.8	0
4	1.4	-45°	12	1	45°
5	2	0	13	1.2	-45°
6	0.8	45°	14	1.4	0
7	1	-45°	15	2	45°
8	1.2	0	-	-	-

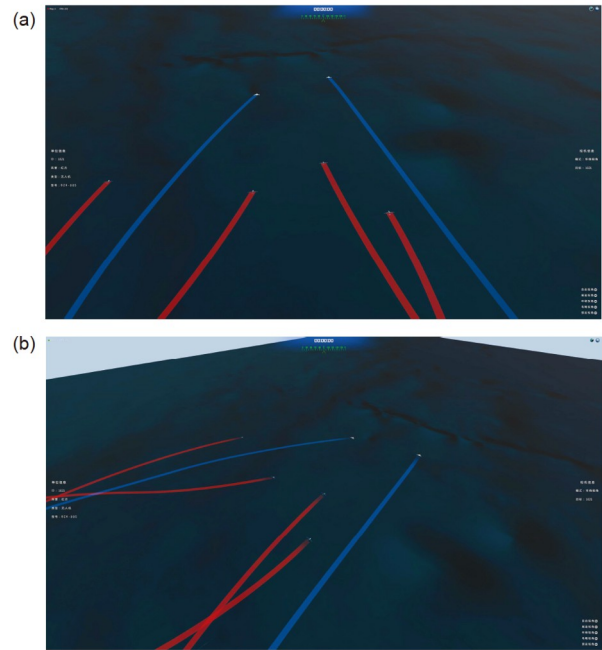


图10 红蓝方无人机博弈飞行轨迹. (a) 俯视图; (b) 环绕视角
Figure 10 Red and blue UAV sides' flight trajectories. (a) Top view; (b) surrounding view.

第一架无人机第1 s选取1号机动作, 2~6 s选择13号机动作, 7~10 s选择3号机动作, 11~20 s选择15号机动作. 蓝方第一架无人机为逃离追击分别采用编号2, 5, 9, 15, 3, 15, 15, 8, 13, 3, 7, 14, 12, 15, 10, 1, 13, 15, 11号机动作. 双方无人机飞行轨迹呈现出由一系列短距转弯组成的剪刀状机动作, 表明参与博弈双方处在中立态势下, 并希望通过机动作选择获得优势态势. 仿真结果实现了结合本文所提方法实现的红方4架无人机对蓝方2架无人机追逃博弈的三维全方位态势集成显示和推演模拟, 进一步验证了所提方法的可行性和有效性.

5 结论

本文针对多无人机博弈决策问题, 在建立无人机动力学模型与博弈态势评估模型的基础上, 基于非合作博弈模型将多无人机博弈决策问题转化为基于混合策略纳什均衡的目标函数寻优问题, 采用本文提出的强化竞争学习鸽群优化算法来求取最优解. 在多无人机博弈决策的两个典型仿真场景想定中, 通过对比强

化竞争学习鸽群优化与基本鸽群优化、基本粒子群优化、竞争粒子群优化和反向学习粒子群优化的在线优化效果, 验证了本文所提强化竞争学习鸽群优化求解多无人机博弈混合策略纳什均衡的可行性和有效性。

参考文献

- Zhou W Q, Zhu J H, Kuang M C. An unmanned air combat system based on swarm intelligence (in Chinese). *Sci Sin Inf*, 2020, 50: 363–374 [周文卿, 朱纪洪, 匡敏驰. 一种基于群体智能的无人空战系统. 中国科学: 信息科学, 2020, 50: 363–374]
- Zhang D F, Duan H B, Fan Y M. UAV swarm containment control inspired by spatial interaction mechanism of wolf-pack foraging (in Chinese). *Sci Sin Tech*, 2022, 52: 1555–1570 [张岱峰, 段海滨, 范彦铭. 仿狼群狩猎空间交互机制的无人机集群包围控制. 中国科学: 技术科学, 2022, 52: 1555–1570]
- Zhou T L, Chen M, Zhu R G, et al. Attack-defense satisficing decision-making of multi-UAVs cooperative multiple targets based on WPS algorithm (in Chinese). *J Command Control*, 2020, 6: 251–256 [周同乐, 陈谋, 朱荣刚, 等. 基于狼群算法的多无人机协同多目标攻防满意决策方法. 指挥与控制学报, 2020, 6: 251–256]
- Liu C, Sun S, Tao C, et al. Sliding mode control of multi-agent system with application to UAV air combat. *Comput Electrical Eng*, 2021, 96: 107491
- Choi H L, Brunet L, How J P. Consensus-based decentralized auctions for robust task allocation. *IEEE Trans Robot*, 2009, 25: 912–926
- Zhen Z, Wen L, Wang B, et al. Improved contract network protocol algorithm based cooperative target allocation of heterogeneous UAV swarm. *Aerosp Sci Tech*, 2021, 119: 107054
- Yan F, Zhu X P, Zhou Z, et al. Real-time task allocation for a heterogeneous multi-UAV simultaneous attack (in Chinese). *Sci Sin Inf*, 2019, 49: 555–569 [严飞, 祝小平, 周洲, 等. 考虑同时攻击约束的多异构无人机实时任务分配. 中国科学: 信息科学, 2019, 49: 555–569]
- Zitouni F, Harous S, Maamri R. A distributed approach to the multi-robot task allocation problem using the consensus-based bundle algorithm and ant colony system. *IEEE Access*, 2020, 8: 27479–27494
- Chen Y, Yang D, Yu J. Multi-UAV task assignment with parameter and time-sensitive uncertainties using modified two-part wolf pack search algorithm. *IEEE Trans Aerosp Electron Syst*, 2018, 54: 2853–2872
- Yang H, Bai X, Baoyin H. Rapid generation of time-optimal trajectories for asteroid landing via convex optimization. *J Guid Control Dyn*, 2017, 40: 628–641
- Chai R, Savvaris A, Tsourdos A, et al. Trajectory optimization of space maneuver vehicle using a hybrid optimal control solver. *IEEE Trans Cybern*, 2019, 49: 467–480
- Chow C K, Yuen S Y. An evolutionary algorithm that makes decision based on the entire previous search history. *IEEE Trans Evol Computat*, 2011, 15: 741–769
- Riwanto B A, Tikka T, Kestila A, et al. Particle swarm optimization with rotation axis fitting for magnetometer calibration. *IEEE Trans Aerosp Electron Syst*, 2017, 53: 1009–1022
- Chen L, Duan H B, Fan Y M, et al. Multi-objective clustering analysis via combinatorial pigeon inspired optimization. *Sci China Tech Sci*, 2020, 63: 1302–1313
- Huang H, Dong K, Yan T, et al. Tactical maneuver trajectory optimization for unmanned combat aerial vehicle using improved differential evolution. *Soft Comput*, 2020, 24: 5959–5970
- Duan H, Zhao J, Deng Y, et al. Dynamic discrete pigeon-inspired optimization for multi-UAV cooperative search-attack mission planning. *IEEE Trans Aerosp Electron Syst*, 2021, 57: 706–720
- Duan H, Qiao P. Pigeon-inspired optimization: A new swarm intelligence optimizer for air robot path planning. *Int J Intell Comput Cybernet*, 2014, 7: 24–37
- Yang Z Y, Duan H B, Fan Y M. Unmanned aerial vehicle formation controller design via the behavior mechanism in wild geese based on Levy flight pigeon-inspired optimization (in Chinese). *Sci Sin Tech*, 2018, 48: 161–169 [杨之元, 段海滨, 范彦铭. 基于莱维飞行鸽群优化的仿雁群无人机编队控制器设计. 中国科学: 技术科学, 2018, 48: 161–169]
- Zhang D, Duan H. Social-class pigeon-inspired optimization and time stamp segmentation for multi-UAV cooperative path planning. *Neurocomputing*, 2018, 313: 229–246

- 20 Huo M Z, Wei C, Yu Y P, et al. Clustering optimization algorithm for large-scale unmanned aerial vehicle based on intelligent behavior of pigeons (in Chinese). *Sci Sin Tech*, 2020, 50: 475–482 [霍梦真, 魏晨, 于月平, 等. 基于鸽群智能行为的大规模无人机集群聚类优化算法. 中国科学: 技术科学, 2020, 50: 475–482]
- 21 Duan H, Huo M, Yang Z, et al. Predator-prey pigeon-inspired optimization for UAV ALS longitudinal parameters tuning. *IEEE Trans Aerosp Electron Syst*, 2019, 55: 2347–2358
- 22 Diao X H, Fang Y W, Xiao B S, et al. Task allocation in cooperative air combat based on multi-agent coalition (in Chinese). *J Beijing Univ Aeronaut Astronaut*, 2014, 40: 1268–1275 [刁兴华, 方洋旺, 肖冰松, 等. 基于多智能体联盟的多机协同空战任务分配. 北京航空航天大学学报, 2014, 40: 1268–1274]
- 23 Wang Y, Zhang W G, Fu L, et al. Nash equilibrium strategies approach for aerial combat based on elite re-election particle swarm optimization (in Chinese). *Control Theory Appl*, 2015, 32: 857–865 [王昱, 章卫国, 傅莉, 等. 基于精英改选机制的粒子群算法的空战纳什均衡策略逼近. 控制理论与应用, 2015, 32: 857–865]
- 24 Cheng R, Jin Y C. A competitive swarm optimizer for large scale optimization. *IEEE Trans Cybern*, 2015, 45: 191–204
- 25 Xu H H, Tang L R. Particle swarm optimization with adaptive elite opposition-based learning for large-scale problems. In: *International Conference on Computational Intelligence and Applications (ICCIA)*. Beijing, 2020. 44–49

Decision-making of multi-UAV combat game via enhanced competitive learning pigeon-inspired optimization

LEI YangQi¹ & DUAN HaiBin^{1,2}

¹ *Bio-inspired Autonomous Flight Systems (BAFS) Research Group, School of Automation Science and Electrical Engineering, Beihang University, Beijing 100083, China;*

² *Peng Cheng Laboratory, Shenzhen 518000, China*

Decision-making of multi-unmanned aerial vehicle (multi-UAV) combat game is a crucial problem in the field of unmanned aerial vehicle combat game. In this study, an enhanced competitive learning pigeon-inspired optimization (ECPIO) algorithm is proposed to handle decision-making of multi-UAV combat game. Firstly, a six degree of freedom UAV model is adopted and situation assessment between UAVs is designed, the payment matrixes corresponding to two players in combat game are calculated. Then, non-cooperative game model is selected, the problem of multi-UAV combat game decision-making is transformed into optimization based on the mixed Nash equilibrium, and ECPIO is adopted to calculate the optimal solution. ECPIO preserves the obvious advantage of pigeon-inspired optimization (PIO), which has fast convergence rate. Our proposed ECPIO can reduce the probability of optimization results trapping into local optimum by introducing enhanced competitive learning strategy. Finally, ECPIO is compared with the basic PIO, basic particle swarm optimization (PSO), competitive particle swarm optimization (CSO) and opposition-based learning particle swarm optimization (OBPSO) by a series of comparative simulation experiments, and the experimental results verify the feasibility and superiority in solving the decision-making of multi-UAV combat game problem.

multi-unmanned aerial vehicle (multi-UAV), combat game decision-making, pigeon-inspired optimization (PIO), enhanced competitive learning

doi: [10.1360/SST-2022-0032](https://doi.org/10.1360/SST-2022-0032)