



非完备策略集下人机对抗空战决策方法

李守义¹, 陈谋^{1*}, 王玉惠¹, 吴庆宪¹, 贺建良²

1. 南京航空航天大学自动化学院, 南京 210016

2. 光电控制技术重点实验室, 洛阳 471000

* 通信作者. E-mail: chenmou@nuaa.edu.cn

收稿日期: 2022-06-13; 修回日期: 2022-09-06; 接受日期: 2022-09-19; 网络出版日期: 2022-12-06

科技创新 2030—“新一代人工智能”重大项目 (批准号: 2018AAA0100805) 资助

摘要 在真实的空战对抗环境中, 不确定性广泛存在. 本文考虑人机对抗场景中的空战决策问题, 提出了一种非完备策略集下的博弈决策方法. 首先, 将对抗的无人机群和有人机群的动态博弈过程分解为不同的决策阶段, 并将当前的决策阶段建模为一个零和博弈模型. 所建模型考虑了有人机群的策略不完全可知的情形, 并用一个不确定效用矩阵来描述这种非完备策略. 随后, 基于矩阵博弈中的最大最小策略概念, 提出了非完备策略下博弈模型的求解方法, 从而给出了无人机群的决策序列. 最后通过仿真实验验证了所提方法的有效性.

关键词 无人机, 空战决策, 人机对抗, 非完备策略集, 博弈论, 最大最小策略

1 引言

随着科学技术的进步, 军事装备正加速革新. 先进的军事装备必然引起未来空战形态的变化, 无人化、系统化和自主化将是未来空战最主要的典型特征^[1~3]. 无人机 (unmanned air vehicle, UAV) 因其成本低、操作灵活、零伤亡和机动性强等特点, 被越来越多地应用到各种军事场景中. 经历了从靶机、侦察机, 到攻击机的角色演变, 无人机正逐渐成为具有侦察、打击一体化功能的作战装备^[4,5], 其作战形式也有从辅助作战发展到完全自主协同作战的趋势. 由于单架无人机无法胜任复杂性的军事任务, 无人机往往以集群的方式参与作战. 但是由于空战信息具有高度复杂且非完备的特点, 这给无人机群协同作战带来了极大的挑战^[6,7]. 因此, 在高动态、强对抗、信息非完备的空战环境中, 提高无人机群的认知决策和协同作战能力, 是一个亟待解决的问题.

近年来, 许多学者在无人机空战对抗领域取得了不少突出的研究成果. Ruan 等^[8] 针对无人机混战中的机动决策问题, 提出了一种受鸽群启发的转移学习优化算法, 该算法采用基于 Kullback-Leibler 散度的迁移学习方法来初始化种群, 提高了优化算法的搜索精度. Zhang 等^[9] 基于动态博弈策略研究了

引用格式: 李守义, 陈谋, 王玉惠, 等. 非完备策略集下人机对抗空战决策方法. 中国科学: 信息科学, 2022, 52: 2239–2253, doi: 10.1360/SSI-2022-0222

Li S Y, Chen M, Wang Y H, et al. Human-computer gaming decision-making method in air combat under an incomplete strategy set (in Chinese). Sci Sin Inform, 2022, 52: 2239–2253, doi: 10.1360/SSI-2022-0222

无人机空战中的任务管理和控制问题, 并设计了一种受图论和 Kuhn-Munkres 算法启发的快速最优搜索算法来给出最优决策. Fan 等^[10] 考虑使用机器学习中的进化算法来训练大型无人机编队的策略, 随着训练数量的增加, 编队显示出越来越高的智能. 博弈论是一个研究多个理性的决策者之间策略交互的有效工具^[11]. 自从 Morgenstern 和 von Neumann^[12] 开启了博弈论研究以来, 博弈论已在许多领域得到了广泛的应用, 如经济、政治、军事、工程、管理等^[13~15]. 近年来, 博弈论因其在多决策者之间的强大推理能力而在多无人机空战决策中受到了广泛关注^[9]. Duan 等^[16] 研究了一种基于捕食者-猎物粒子群优化的动态博弈论方法, 并利用该算法获得每一步的混合策略纳什均衡. Li 等^[2] 考虑了多无人机空战中的时敏信息, 并将其表述为矩阵博弈中的线性约束. 随后给出了一种有效方法来求解所提出的受限策略博弈, 进而给出了时敏信息约束下的多无人机空战决策.

毫无疑问, 上述研究^[2, 8~10, 16] 大大推动了多机协同对抗决策的发展, 但通过分析发现, 现有成果大多是在假设对方策略完全已知的情形下进行决策的, 即彼此都完全了解对方的所有可能的策略, 这在空战决策中是一个相当理想的假设. 在真实的空战中, 对抗方飞机的机动性能、电子设备的侦察能力、所携带导弹的打击效果等不一定是事前可知的, 从而导致对抗方的策略集在对抗决策时是不完全可知的. 我们将这种带有不可知策略的策略集称为非完备策略集. 策略集的非完备性会影响决策的制定, 而此类问题也不能直接应用传统的博弈决策方法进行求解, 因而研究非完备策略集下的决策方法具有深远意义. 本文主要针对部分有人机的机动性能和导弹打击效果不确定导致的策略集未知情况展开研究, 以提高无人机群在这类非完备策略集下的自主决策能力. 在经典的博弈模型中, 纳什均衡解已应用到很多的博弈场景中. 纳什在 20 世纪 50 年代证明了 N 人博弈至少存在一个纳什均衡点^[17]. 然而, 当博弈中带有非完备策略时, 纳什均衡的概念很难被平行推广到这类博弈问题中. 并且, 纳什均衡往往是混合策略的形式, 这种以概率方式描述的策略在空战中难以实际应用. 基于经典的矩阵博弈中纳什均衡策略和最大最小策略之间的内在联系^[18], 本文给出了非完备策略集下博弈模型的一种决策方法, 提高了无人机群处理非完备信息的能力.

本文的后续章节组织如下. 第 2 节阐述了人机对抗场景中的机群空战决策问题, 然后给出了机群对抗的相对态势评估方法, 并给出了矩阵博弈的相关表述. 第 3 节给出了非完备策略集下的人机对抗空战博弈决策方法. 首先定义了所研究问题中的非完备策略集, 然后建立了非完备策略集下的人机对抗空战博弈模型, 随后给出了该博弈模型的求解方法. 第 4 节给出了数值仿真, 验证了所提方法的有效性. 第 5 节为结论, 总结了本文的主要内容.

2 问题描述与预备知识

2.1 问题描述

本文考虑人机对抗场景中的如下空战决策问题: 空战中的两个机群, 无人机群 (计算机, 记为 C) 和有人机群 (人类, 记为 H), 分别包含 N^c 架无人机和 N^h 架有人机. 双方都希望通过一系列的空中机动和发射导弹等操纵, 逐步消灭对方的主要作战力量, 从而获取空战的胜利. 在对抗过程中, 双方基于当前的空战相对态势采取各自的空中操纵 (机动操纵或攻击操纵), 对抗后幸存的飞机数量和相对态势都已发生变化, 随后双方再基于变化后的相对态势, 进行下一轮的决策, 如此重复进行下去, 直至某一方获得胜利. 这是一个涉及多个决策者的序贯决策问题, 其本质上是 C 和 H 的一个动态博弈过程. 假设每经历 Δt 时间间隔, 双方完成一次空中操纵, 从而 C 和 H 的动态博弈决策过程可以分解为多个决策阶段, 如图 1 所示, 其中 t_k 为第 k 个时刻, 满足 $t_{k+1} - t_k = \Delta t$ ($k = 1, 2, \dots, T-1$), 时刻 t_k 到 t_{k+1}

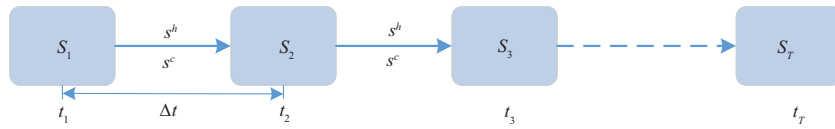


图 1 (网络版彩图) C 和 H 的动态博弈过程示意图
 Figure 1 (Color online) Dynamic game process of C and H

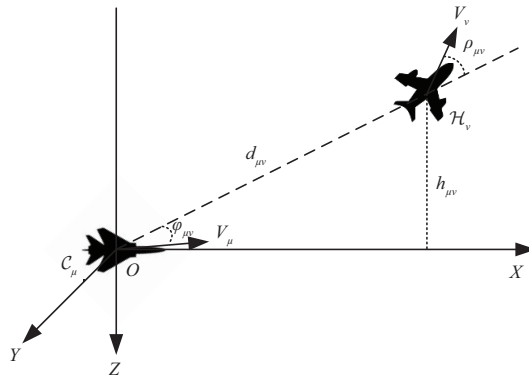


图 2 C_μ 和 H_ν 的空战相对态势
 Figure 2 Air combat situation of C_μ and H_ν

之间的过程称为第 k 个决策阶段, S_k 为时刻 t_k 时的相对态势, s^c 和 s^h 分别为 C 和 H 在该决策阶段的空战行动.

2.2 机群相对态势评估

相对态势评估刻画了 C 和 H 在对抗过程中, 双方飞机方位、距离、高度、速度以及性能等相对态势评估因子的相对占优程度. 空战相对态势评估是后续的博弈决策的依据. 本小节基于单机对单机的相对态势评估^[19], 给出了机群对机群的相对态势评估方法.

若记 C 的第 μ 架无人机为 C_μ ($\mu = 1, 2, \dots, N^c$), H 的第 ν 架有人机为 H_ν ($\nu = 1, 2, \dots, N^h$), C_μ 和 H_ν 的空战相对态势如图 2 所示, 其中 V_μ 和 V_ν 分别表示两机的速度, $d_{\mu\nu}$ 为两机之间的距离, $h_{\mu\nu}$ 表示两机之间的高度差, $\varphi_{\mu\nu}$ 和 $\rho_{\mu\nu}$ 分别为 H_ν 相对于 C_μ 的方位角和进入角. 上述参数均可通过两机的位置信息和速度信息计算得到. 当给出 C_μ 和 H_ν 的导弹攻击距离、雷达探测距离和导弹携带数量这些参数时, 通过文献 [19] 中的相对态势评估方法, 可以计算出 C_μ 对于 H_ν 的角度优势、距离优势、高度优势、速度优势和性能优势, 并最终给出 C_μ 相对于 H_ν 的综合相对态势优势 $W_{\mu\nu}$. 对于无人机群或有人机群, 假设同一机群内飞机的性能相差不大. 在单机对单机的相对态势评估的基础上, 我们定义无人机群 C 对有人机群 H 的总体相对态势优势如下:

$$W = \sum_{\mu=1}^{N^c} \sum_{\nu=1}^{N^h} W_{\mu\nu}. \tag{1}$$

事实上, 这是 C 的所有无人机对 H 的所有有人机的累加态势优势.

2.3 矩阵博弈的基本概念

本小节主要回顾矩阵博弈的基本概念及其纳什均衡解的定义^[18, 20].

考虑矩阵博弈 $G_0 = \langle S^1, S^2, u_0 \rangle$, 其中

- S^1 和 S^2 分别表示参与人 1 和参与人 2 的策略集, 元素 $s^1 \in S^1$ 和 $s^2 \in S^2$ 分别为参与人 1 和参与人 2 的一个策略. 策略对 (s^1, s^2) 为一个策略组合, 所有策略组合的全体记为 $S^1 \times S^2$.

- $u_0 : S^1 \times S^2 \rightarrow \mathbb{R}$ 表示参与人 1 的效用函数. 对于任意一个策略组合 $(s^1, s^2) \in S^1 \times S^2$, u_0 给出参与人 1 的效用值 $u_0(s^1, s^2)$. 参与人 2 的效用函数为 $-u_0$, 即这是一个零和博弈.

若策略集 S^1 和 S^2 中分别有 m_0 和 n_0 个策略, 并记 $S^1 = \{s_1^1, s_2^1, \dots, s_{m_0}^1\}$, $S^2 = \{s_1^2, s_2^2, \dots, s_{n_0}^2\}$, 那么矩阵博弈 G_0 可由如下效用矩阵 A_0 来表示:

$$A_0 = \begin{matrix} & s_1^2 & s_2^2 & \cdots & s_{n_0}^2 \\ \begin{matrix} s_1^1 \\ s_2^1 \\ \vdots \\ s_{m_0}^1 \end{matrix} & \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n_0} \\ a_{21} & a_{22} & \cdots & a_{2n_0} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m_01} & a_{m_02} & \cdots & a_{m_0n_0} \end{pmatrix} \end{matrix}, \quad (2)$$

其中, a_{ij} ($i = 1, 2, \dots, m_0; j = 1, 2, \dots, n_0$) 为策略组合 (s_i^1, s_j^2) 下参与人 1 的效用值 $u(s_i^1, s_j^2)$, 即 $a_{ij} = u(s_i^1, s_j^2)$.

在博弈论中, 纳什均衡通常被用来作为博弈决策的解. 一个策略组合 (s^{1*}, s^{2*}) 称为 G_0 的一个纯策略纳什均衡, 如果对于任意的 $s^1 \in S^1, s^2 \in S^2$, 有下面的不等式成立:

$$u_0(s^1, s^{2*}) \leq u_0(s^{1*}, s^{2*}) \leq u_0(s^{1*}, s^2). \quad (3)$$

在纳什均衡下, 所有的参与人都不能单方面地改变自身的策略来增加其效用值. 纳什均衡是博弈的稳定状态. 然而, 并不是所有的矩阵博弈都存在纯策略纳什均衡^[18]. 下面的引理给出纯策略纳什均衡的存在性刻画.

引理 1 ([20]) 矩阵博弈 $G_0 = \langle S^1, S^2, u \rangle$ 存在纯策略纳什均衡 (s^{1*}, s^{2*}) 的充要条件是下面的等式成立:

$$\max_{s^1 \in S^1} \min_{s^2 \in S^2} u_0(s^1, s^2) = \min_{s^2 \in S^2} \max_{s^1 \in S^1} u_0(s^1, s^2). \quad (4)$$

作为一个零和博弈, 一方的收益就是另一方的损失. 对于参与人 1 的任意一个策略 $s^1 \in S^1$, 参与人 2 总是从其策略集 S^2 中选取一个策略 $s^2 \in S^2$, 使得参与人 1 的效用值 $u_0(s^1, s^2)$ 最小. 因而, 参与人 1 选择策略 $s^1 \in S^1$, 可以得到的最小效用值为 $\min_{s^2 \in S^2} u_0(s^1, s^2)$. 参与人 1 将会从其策略集 S^1 中选择策略 \hat{s}^1 来最大化 $\min_{s^2 \in S^2} u_0(s^1, s^2)$, 即

$$\min_{s^2 \in S^2} u_0(\hat{s}^1, s^2) = \max_{s^1 \in S^1} \min_{s^2 \in S^2} u_0(s^1, s^2). \quad (5)$$

该值称为 G_0 的最大最小值, \hat{s}^1 称为参与人 1 的最大最小策略. 同样地, 参与人 2 也可以从其策略集中选择策略 \hat{s}^2 , 使得

$$\max_{s^1 \in S^1} u_0(s^1, \hat{s}^2) = \min_{s^2 \in S^2} \max_{s^1 \in S^1} u_0(s^1, s^2). \quad (6)$$

该值称为 G_0 的最小最大值, \hat{s}^2 称为参与人 2 的最小最大策略. 显然, 最大最小策略和最小最大策略一定存在.

注释1 事实上,由引理 1 的证明过程^[20]可知,在纯策略纳什均衡存在的前提下,矩阵博弈的最大最小策略 s^1 和最小最大策略 s^2 构成该博弈的纯策略纳什均衡 (s^1, s^2) . 这揭示了矩阵博弈中最大最小策略,最小最大策略和纳什均衡之间的关系.

混合策略表示以概率的方式来选择策略.若记参与人 1 和参与人 2 的混合策略集分别为 X^1 和 X^2 ,那么它们有如下形式:

$$X^1 = \left\{ x^1 \in \mathbb{R}^{m_0} \mid \sum_{i=1}^{m_0} x_i^1 = 1, x_i^1 \geq 0 \right\}, \quad (7)$$

$$X^2 = \left\{ x^2 \in \mathbb{R}^{n_0} \mid \sum_{j=1}^{n_0} x_j^2 = 1, x_j^2 \geq 0 \right\}, \quad (8)$$

其中, $x^1 = (x_1^1, x_2^1, \dots, x_{m_0}^1)$, $x^2 = (x_1^2, x_2^2, \dots, x_{n_0}^2)$, x_i^1 和 x_j^2 分别表示参与人 1 和参与人 2 选择其策略 s_i^1 和 s_j^2 的概率.纳什证明了任何一个 N 人博弈存在混合策略纳什均衡^[17].特别地,矩阵博弈 G_0 的混合策略纳什均衡可通过求解一对对偶线性规划得到^[21].

3 非完备策略集下人机对抗博弈决策

在实际的空战对抗过程中,对方的策略集通常是非完备的,这使得博弈决策过程存在不确定性.本节针对 \mathcal{H} 的策略非完备时进行 \mathcal{C} 的自主决策方法研究.非完备策略集下人机对抗博弈决策方法的总体流程如图 3 所示.由图 3 可知,首先对当前相对态势进行评估;然后构造 \mathcal{C} 和 \mathcal{H} 的策略集,其中 \mathcal{H} 的策略集由已知策略集和未知策略集组成;其次将当前决策阶段建模为 \mathcal{C} 和 \mathcal{H} 的非完备策略下博弈决策模型;最后基于本文提出的 Max-Min-Optimization (MMO) 算法,给出博弈模型的决策结果,并执行决策;当一方获得胜利时对抗结束,否则更新相对态势信息后继续进行博弈决策.

3.1 非完备策略集

在构造 \mathcal{C} 和 \mathcal{H} 的策略集之前,首先给出单架飞机的运动学模型和机动库的设计.在惯性坐标系下,单架飞机的运动学模型由如下 3 自由度质点模型来描述^[22]:

$$\begin{cases} \dot{x} = v \cos \gamma \cos \chi, \\ \dot{y} = v \cos \gamma \sin \chi, \\ \dot{z} = v \sin \gamma, \\ \dot{v} = g (n_x - \sin \gamma), \\ \dot{\gamma} = \frac{g}{v} (n_z \cos \theta - \cos \gamma), \\ \dot{\chi} = \frac{g}{v \cos \gamma} n_z \sin \theta, \end{cases} \quad (9)$$

其中, x, y, z 表示 3 个坐标方向的位置坐标, v 表示速度, γ 为航迹倾角, χ 为航向角, g 是重力加速度, θ 表示滚转角, n_x 和 n_z 分别为切向过载和法向过载.控制量 $[n_x, n_z, \theta]$ 控制飞机进行机动,一组控制量代表一个机动决策.

自 NASA^[23] 给出基本机动动作库以来,机动库有了长远的发展.本文使用包含 15 种机动动作组成的机动库^[24,25].文献^[25]给出了这 15 种机动动作的控制量 $[n_x, n_z, \theta]$,如表 1 所示.考虑到人机对抗场景中无人机和有人机的机动能力存在较大差异(如纵向和侧向过载能力),在上述控制量的基础

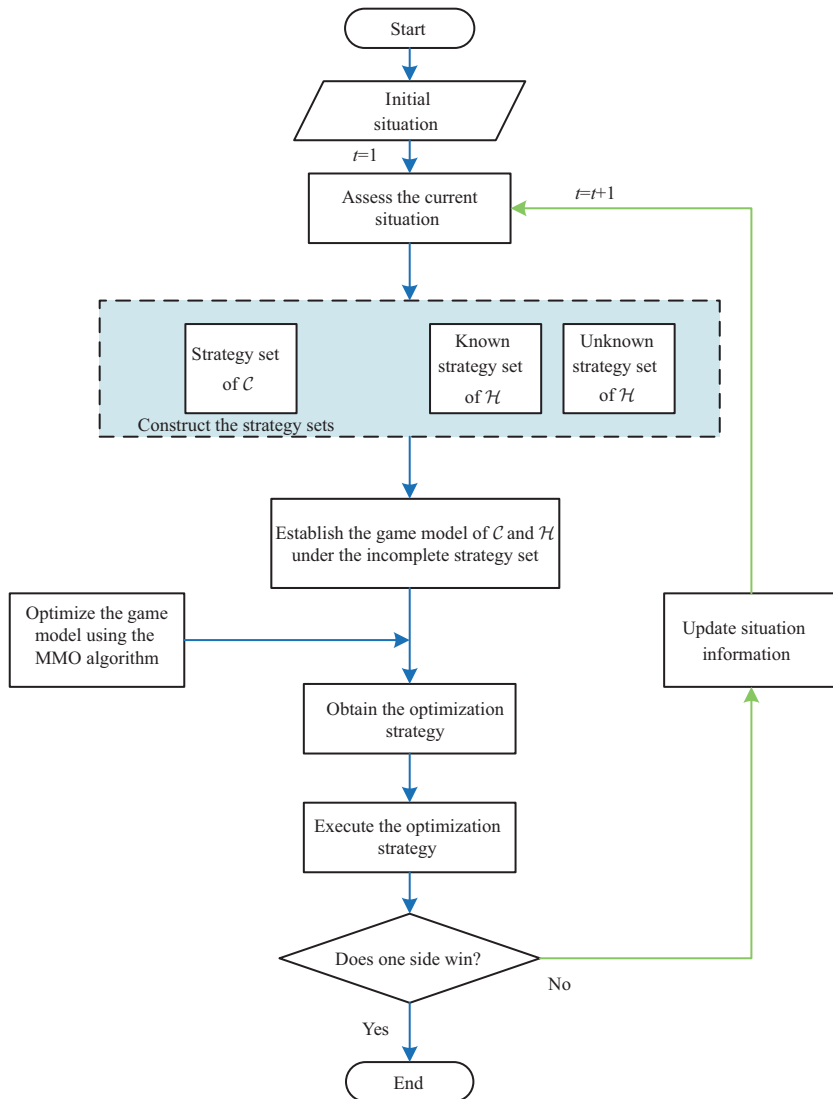


图 3 (网络版彩图) 非完备策略集下人机对抗博弈决策总体流程图

Figure 3 (Color online) Flowchart of human-computer gaming decision-making under the incomplete strategy set

上, 我们给出了机动能力参数 λ ($\lambda > 0$), 用来刻画无人机和有人机不同的机动能力. 具体地, 若分别记无人机和有人机的机动能力参数为 λ^c 和 λ^h , 则它们的控制量分别为 $[\lambda^c n_x, \lambda^c n_z, \theta]$ 和 $[\lambda^h n_x, \lambda^h n_z, \theta]$, 其中 $[n_x, n_z, \theta]$ 为表 1 中机动动作的控制量, $\lambda^c n_x$ 和 $\lambda^h n_x$ 分别为无人机和有人机的切向过载, $\lambda^c n_z$ 和 $\lambda^h n_z$ 分别为它们的法向过载. 在上述控制量下, 无人机和有人机的机动库分别为 \mathcal{M}^c 和 \mathcal{M}^h , 它们均含有 15 个机动动作, 但是由于控制量的不同, 同样的动作将有不同的机动效果.

单架飞机的基本空中操纵包含机动操纵和攻击操纵, 其中机动操纵的目的是占据有利的攻击位置, 在达到导弹攻击区时, 触发导弹发射条件. 导弹攻击区的解算请参见文献 [26]. 假设在进行攻击时, 飞机沿着当前的速度飞行并发射导弹. 具体地, 无人机 C_μ 的策略集构造为 $\mathcal{M}^c \cup \mathcal{A}_\mu^c$, 其中 \mathcal{A}_μ^c 表示当前相对态势下进入该机导弹攻击区所有可以攻击的有人机. 若当前 C_μ 的导弹攻击区内没有对方有人机,

表 1 机动库中的基本机动动作及其控制量
 Table 1 Basic maneuvers and their control variables in the maneuver library

Symbol	Maneuver	n_x	n_z	θ ($^\circ$)
e_1	Steady flight	0	1	0
e_2	Accelerated flight	1	1	0
e_3	Decelerated flight	-1	1	0
e_4	Steady turn left	0	8	-82.8
e_5	Accelerated turn left	2	8	-82.8
e_6	Decelerated turn left	-2	8	-82.8
e_7	Steady turn right	0	8	82.8
e_8	Accelerated turn right	2	8	82.8
e_9	Decelerated turn right	-2	8	82.8
e_{10}	Steady push over	0	8	180
e_{11}	Accelerated push over	2	8	180
e_{12}	Decelerated push over	-1	8	180
e_{13}	Steady pull up	0	8	0
e_{14}	Accelerated pull up	2	8	0
e_{15}	Decelerated pull up	-1	8	0

那么 $\mathcal{A}_\mu^c = \emptyset$. 无人机群 \mathcal{C} 的策略集 S^c 有如下的形式:

$$S^c = \prod_{\mu=1}^{N^c} (\mathcal{M}^c \cup \mathcal{A}_\mu^c), \quad (10)$$

其中的每一个策略 s^c 指定了 \mathcal{C} 的每一架无人机采取何种操纵, 策略集 S^c 给出了所有可能的操纵组合. 类似地, 有人机群 \mathcal{H} 的策略集 S^h 给出如下:

$$S^h = \prod_{\nu=1}^{N^h} (\mathcal{M}^h \cup \mathcal{A}_\nu^h). \quad (11)$$

如第 1 节中所述, 在人机对抗空战决策问题中, 对无人机群 \mathcal{C} 来说, 有人机群 \mathcal{H} 的策略集通常是非完备的. 本文考虑有人机群的机动行动和攻击行动不确定这种非完备策略集. 具体地, 假设 \mathcal{C} 对 \mathcal{H} 中部分有人机的机动性能不确定, 根据经验, 这些有人机的机动能力参数 λ^h 可用一个区间来表示; 同样地, 假设 \mathcal{C} 对 \mathcal{H} 中某些有人机的导弹性能不确定, 根据经验, 这些有人机的毁伤概率 σ 也可用一个毁伤区间来表示. 本文将带有机动或攻击不确定的策略称为未知策略, 所有未知策略构成未知策略集, 记为 \tilde{S}^h . S^h 中其余策略称为已知策略集, 记为 \bar{S}^h . 显然, 未知策略集和已知策略集满足 $S^h = \bar{S}^h \cup \tilde{S}^h$. 由于策略集 S^h 中含有未知策略, 因而称之为非完备策略集.

注释 2 尽管我们只考虑了 \mathcal{H} 的非完备策略, 事实上, 对于 \mathcal{H} 来说, \mathcal{C} 的策略也是非完备的. 这里我们只是从 \mathcal{C} 的视角考虑当 \mathcal{H} 的策略非完备时如何进行决策, 因而我们省去了 \mathcal{C} 的策略非完备的描述, 若需要, 可进行类似的分析. 此外, 当机动能力参数 λ^h 或毁伤概率 σ 不确定时, 可通过先验知识获得它们的上下界, 即不确定参数的上下界是事先已知的.

3.2 人机对抗博弈决策模型

在 \mathcal{C} 和 \mathcal{H} 的动态对抗中, 我们将当前决策阶段建模为 \mathcal{C} 和 \mathcal{H} 的不确定博弈模型, $G = \langle S^c, S^h, u \rangle$, 其中参数如下:

- S^c 表示 \mathcal{C} 的策略集, 其中的元素 s^c 为 \mathcal{C} 的一个策略. 假设 S^c 中有 m 个策略, 那么 $S^c = \{s_1^c, s_2^c, \dots, s_m^c\}$.

- $S^h = \bar{S}^h \cup \tilde{S}^h$ 表示 \mathcal{H} 的非完备策略集, 其中 \bar{S}^h 和 \tilde{S}^h 分别为 \mathcal{H} 的已知策略集和未知策略集 (从 \mathcal{C} 的视角来说). 策略集 S^h 中的元素 s^h 为 \mathcal{H} 的一个策略. 若已知策略集 \bar{S}^h 中有 n 个已知策略, 未知策略集 \tilde{S}^h 中有 q 个未知策略, 那么 $\bar{S}^h = \{\bar{s}_1^h, \bar{s}_2^h, \dots, \bar{s}_n^h\}$, $\tilde{S}^h = \{\tilde{s}_1^h, \tilde{s}_2^h, \dots, \tilde{s}_q^h\}$.

- $u: S^c \times S^h \rightarrow \mathbb{R}$ 为 \mathcal{C} 的效用函数. 对于每一个策略组合 $(s^c, s^h) \in S^c \times S^h$, 效用函数 u 给出 \mathcal{C} 的效用值 $u(s^c, s^h)$. 由于 \mathcal{C} 和 \mathcal{H} 的对抗属性, \mathcal{H} 的效用函数定义为 $-u$, 即这是一个零和博弈.

对于 \mathcal{C} 的任意一个给定的策略 s_i^c , 当 \mathcal{H} 的策略是已知策略 $\bar{s}_j^h \in \bar{S}^h$ 时, \mathcal{C} 的效用值 $u(s_i^c, \bar{s}_j^h)$ 是一个确定的数值, 并记 $a_{ij} = u(s_i^c, \bar{s}_j^h)$; 而当 \mathcal{H} 的策略 $\tilde{s}_l^h \in \tilde{S}^h$ 时, \mathcal{C} 并不能确定效用值 $u(s_i^c, \tilde{s}_l^h)$ 的具体数值, 但是可以估计出 $u(s_i^c, \tilde{s}_l^h)$ 的取值, 用区间 $I_{i,l} = [I_{i,l}^-, I_{i,l}^+]$ 来表示. 即给定 $s_i^c \in S^c$, 对任意的 $\tilde{s}_l^h \in \tilde{S}^h$, 我们有 $u(s_i^c, \tilde{s}_l^h) \in I_{i,l}$. 综上, 对于任意的 $s_i^c \in S^c$, $s^h \in S^h$ ($s^h = \bar{s}_j^h \in \bar{S}^h$ 或 $s^h = \tilde{s}_l^h \in \tilde{S}^h$), 我们有

$$\begin{cases} u(s_i^c, \bar{s}_j^h) = a_{ij}, & \bar{s}_j^h \in \bar{S}^h, \\ u(s_i^c, \tilde{s}_l^h) \in I_{i,l}, & \tilde{s}_l^h \in \tilde{S}^h. \end{cases} \quad (12)$$

类似于矩阵博弈, 不确定博弈 G 也可用如下的一个不确定效用矩阵 \mathbf{A} 来表示:

$$\mathbf{A} = \begin{matrix} & \begin{matrix} \bar{s}_1^h & \bar{s}_2^h & \cdots & \bar{s}_n^h & \tilde{s}_1^h & \tilde{s}_2^h & \cdots & \tilde{s}_q^h \end{matrix} \\ \begin{matrix} s_1^c \\ s_2^c \\ \vdots \\ s_m^c \end{matrix} & \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & I_{1,1} & I_{1,2} & \cdots & I_{1,q} \\ a_{21} & a_{22} & \cdots & a_{2n} & I_{2,1} & I_{2,2} & \cdots & I_{2,q} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & I_{m,1} & I_{m,2} & \cdots & I_{m,q} \end{pmatrix} \end{matrix}, \quad (13)$$

其中, 效用值 a_{ij} 和 $I_{i,l} = [I_{i,l}^-, I_{i,l}^+]$ 均由效用函数 u 计算得到. 具体地, 效用函数 u 构造如下:

$$u(s^c, s^h) = W(s^c, s^h) + AT(s^c, s^h), \quad (s^c, s^h) \in S^c \times S^h, \quad (14)$$

其中 $W(s^c, s^h)$ 为机动效用函数, 表示双方分别采取策略 s^c 和 s^h 后, \mathcal{C} 相对于 \mathcal{H} 的态势优势值. 若在当前态势下, 无人机群采取策略 s^c 时, 存在至少两架无人机发生碰撞, 则该机动效用 $W(s^c, s^h)$ 的数值设置为 $-\infty$, 否则 $W(s^c, s^h)$ 可由式 (1) 计算得到; $AT(s^c, s^h)$ 为攻击效用函数, 表示攻击行动带来的攻击收益, 可由如下公式计算:

$$AT(s^c, s^h) = \sum \sigma_c V^h - \sum \sigma_h V^c, \quad (15)$$

其中 σ_c 和 σ_h 分别表示策略 (s^c, s^h) 下无人机和有人机的毁伤概率, V^c 和 V^h 表示进入导弹攻击区内的无人机和有人机的价值. 在上述效用函数的计算中, 若某些有人机的机动能力参数为一个区间数, 则机动后的态势不确定, 此时机动收益 $W(s^c, s^h)$ 为一个区间数, 同样当某些有人机的毁伤概率 σ_h 是一个区间数时, 通过区间数的运算规则 [27], 可以得到的攻击效用 $AT(s^c, s^h)$ 也为一个区间数. 总之, 当机动能力参数或毁伤概率为一个区间数时, 对应的收益为一个区间收益 $I_{i,l}$, 其中区间数 $I_{i,l}$ 的上下界, $I_{i,l}^+$ 和 $I_{i,l}^-$, 本质上是由区间机动能力参数和区间攻击概率的上下界计算得到的.

3.3 人机对抗博弈决策方法

我们知道,传统的矩阵博弈 G_0 可通过转化为一对对偶线性规划来求解^[21],其中线性规划的约束系数对应于收益值,目标函数为优化变量的求和.而在所建立的非完备策略集下的博弈决策模型 $G = \langle S^c, S^h, u \rangle$ 中,效用矩阵 A 中存在着不确定数据 I_i ($i = 1, 2, \dots, m$).借鉴传统矩阵博弈的处理方法,我们将博弈 G 转化为一个不确定对偶线性规划,不确定对偶线性规划的约束系数对应于不确定收益值,目标函数为优化变量的求和.具体形式如下:

$$\left\{ \begin{array}{l} \min \sum_{i=1}^m x_i \\ \text{s.t.} \sum_{i=1}^m x_i a_{ij} \geq 1, \quad j = 1, 2, \dots, n, \\ \sum_{i=1}^m x_i I_{i,l} \geq 1, \quad l = 1, 2, \dots, q, \\ x_i \geq 0, \quad i = 1, 2, \dots, m, \end{array} \right. \quad (16)$$

$$\left\{ \begin{array}{l} \max \sum_{j=1}^{n+q} y_j \\ \text{s.t.} \sum_{j=1}^n y_j a_{ij} + \sum_{l=1}^q y_{n+l} I_{i,l} \leq 1, \quad i = 1, 2, \dots, m, \\ y_j \geq 0, \quad j = 1, 2, \dots, n+q, \end{array} \right. \quad (17)$$

其中, x_i, y_j ($i = 1, 2, \dots, m; j = 1, 2, \dots, n$) 为优化变量,向量 (x_1, x_2, \dots, x_m) 和 (y_1, y_2, \dots, y_n) 分别构成 C 和 H 没有归一化的混合策略; a_{ij} 和 $I_{i,l} = [I_{i,l}^-, I_{i,l}^+]$ ($l = 1, 2, \dots, q$) 为效用矩阵 A 中效用值,如式 (13) 所示.注意到 A 是一个 $m \times (n+q)$ 矩阵,根据线性规划法^[21],效用矩阵的每一列对应一个不等式约束,因而式 (16) 中共有 $n+q$ 个不等式约束.同样地,式 (17) 中,效用矩阵的每一行对应于一个不等式约束,因而有 m 个不等式约束.通过求解不确定对偶线性规划 (16) 和 (17) 的最优解 \bar{x} 和 \bar{y} , 即得不确定博弈 G 的纳什均衡策略 $(v\bar{x}, v\bar{y})$, 其中 $v = (\sum_{i=1}^m \bar{x}_i)^{-1}$.

对于不确定线性规划 (16), 其可行域中的任一元素 x 对 $I_{i,l}$ ($i = 1, 2, \dots, m; l = 1, 2, \dots, q$) 中的任意一组数据, 都有式 (16) 中的约束条件成立. 注意到 x 是非负向量, 因而若对区间 $I_{i,l}$ 的下界 $I_{i,l}^-$ 成立上述约束, 则对任意 $I_{i,l}$ 中的数据成立. 类似地, 不确定线性规划 (17) 的可行域也可简化为对 $I_{i,l}$ 的上界 $I_{i,l}^+$ 满足不等式约束即可. 因而, 上述不确定规划可以等价于如下的确定规划:

$$\left\{ \begin{array}{l} \min \sum_{i=1}^m x_i \\ \text{s.t.} \sum_{i=1}^m x_i a_{ij} \geq 1, \quad j = 1, 2, \dots, n, \\ \sum_{i=1}^m x_i I_{i,l}^- \geq 1, \quad l = 1, 2, \dots, q, \\ x_i \geq 0, \quad i = 1, 2, \dots, m, \end{array} \right. \quad (18)$$

$$\begin{cases} \max & \sum_{j=1}^{n+q} y_j, \\ \text{s.t.} & \sum_{j=1}^n y_j a_{ij} + \sum_{l=1}^q y_{n+l} I_{i,l}^+ \leq 1, \quad i = 1, 2, \dots, m, \\ & y_j \geq 0, \quad j = 1, 2, \dots, n+q. \end{cases} \quad (19)$$

式 (18) 和 (19) 的最优解分别对应于 \mathcal{C} 和 \mathcal{H} 的纳什均衡解。

上述分析表明, 通过求解对偶线性规划 (18) 和式 (19), 可以得到 $G = \langle S^c, S^h, u \rangle$ 的纳什均衡. 事实上, 由式 (18) 和 (19) 所得到的纳什均衡通常为混合策略形式的纳什均衡, 而这种以概率描述的策略在真实空战中难于实施. 为此, 我们给出了无人机群的纯策略形式下的决策定义, 具体如下.

定义1 考虑非完备策略集下的空战博弈模型 $G = \langle S^c, S^h, u \rangle$, 无人机群 \mathcal{C} 在 G 中的决策定义为

$$s^{c*} = \arg \max_{s_i^c \in S^c} \left(\min \left\{ \min_{\bar{s}_j^h \in \bar{S}^h} a_{ij}, \min_{s_l^h \in \tilde{S}^h} I_{i,l}^- \right\} \right), \quad (20)$$

其中, \bar{S}^h 为已知策略集, \tilde{S}^h 为未知策略集, $a_{ij} = u(s_i^c, \bar{s}_j^h)$ 为确定效用, $I_{i,l}^-$ 为不确定效用 $I_{i,l}$ 的下界, 如式 (13) 所示.

为保证上述定义的无人机群的决策 s^{c*} 在理论上的合理性, 我们给出如下定理.

定理1 考虑人机对抗博弈决策模型 $G = \langle S^c, S^h, u \rangle$, 那么式 (20) 所定义的策略 s^{c*} 为 G 的最大最小策略. 并且若 G 存在纯策略纳什均衡, 则 s^{c*} 就是 G 的纯策略纳什均衡.

证明 我们首先定义矩阵博弈 $G_1 = \langle S^c, S^h, u_1 \rangle$, 其效用矩阵 \mathbf{A}_1 给出如下:

$$\mathbf{A}_1 = \begin{matrix} & \begin{matrix} \bar{s}_1^h & \bar{s}_2^h & \cdots & \bar{s}_n^h & \tilde{s}_1^h & \tilde{s}_2^h & \cdots & \tilde{s}_q^h \end{matrix} \\ \begin{matrix} s_1^c \\ s_2^c \\ \vdots \\ s_m^c \end{matrix} & \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} & I_{1,1}^- & I_{1,2}^- & \cdots & I_{1,q}^- \\ a_{21} & a_{22} & \cdots & a_{2n} & I_{2,1}^- & I_{2,2}^- & \cdots & I_{2,q}^- \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & I_{m,1}^- & I_{m,2}^- & \cdots & I_{m,q}^- \end{pmatrix} \end{matrix}. \quad (21)$$

由文献 [21] 可知, G_1 对应的第一个线性规划也为式 (18) 的形式. 因此, \mathcal{C} 在 G 和 G_1 中具有相同的解. 对于矩阵博弈 G_1 , 其最大最小策略 s^{c*} 有如下形式:

$$s^{c*} = \arg \max_{s_i^c \in S^c} \left(\min \left\{ \min_{\bar{s}_j^h \in \bar{S}^h} a_{ij}, \min_{s_l^h \in \tilde{S}^h} I_{i,l}^- \right\} \right),$$

从而 s^{c*} 即为 G 的最大最小策略. 通过注释 1 的分析知, 在纯策略存在时, 最大最小策略 s^{c*} 即为纯策略纳什均衡. 综上, 定理的结论得证.

通过上述的建模与分析, 给出了当前决策阶段下无人机群的决策 s^{c*} . 随后根据变化后的空战态势, 再次应用上述方法给出下一个决策阶段的最佳决策, 如此进行下去直到某一方的飞机被全部击落 (飞机进入对方飞机的导弹攻击区则认为被击落). 这种动态求解整个空战决策过程的方法称为 Max-Min-Optimization (MMO) 算法, 用于给出非完备策略集下无人机群 \mathcal{C} 的决策序列, 具体求解过程如算法 1 所示.

算法 1 Max-Min-Optimization (MMO) 算法

输入: 无人机群 C 和有人机群 H 的飞机数量 N^c 和 N^h , 以及双方机群中每一架飞机的如下信息: 飞机价值、位置、速度、载弹量、导弹攻击距离、雷达探测距离、导弹毁伤概率 (其中 H 的 N_1^h 架有人机的毁伤概率为一个区间);

- 1: **while** 无人机群 C 和有人机群 H 都有幸存飞机, 即 $N^c > 0$ 且 $N^h > 0$ **do**
- 2: 构造无人机 C_μ ($\mu = 1, 2, \dots, N^c$) 当前相对态势下的策略集 $\mathcal{M} \cup \mathcal{A}_\mu^c$, 其中 \mathcal{M} 为机动库, \mathcal{A}_μ^c 为 C_μ 当前相对态势下导弹攻击区内的有人机; 同样地, 构造有人机 H_ν ($\nu = 1, 2, \dots, N^h$) 的策略集 $\mathcal{M} \cup \mathcal{A}_\nu^h$;
- 3: 根据式 (10) 和 (11) 分别计算 C 和 H 的策略集 S^c 和 S^h , 其中 $S^h = \bar{S}^h \cup \tilde{S}^h$, \bar{S}^h 为已知策略集, \tilde{S}^h 为未知策略集;
- 4: 根据效用函数计算公式 (14), 计算每一个策略组合 $(s_i^c, s^h) \in S^c \times S^h$ 下的效用 $u(s_i^c, s^h)$, 其包括机动效用 $W(s_i^c, s^h)$ 和攻击效用 $AT(s_i^c, s^h)$, 其中机动效用 $W(s_i^c, s^h)$ 由相对态势评估式 (1) 给出. 当 $s^h = \bar{s}_j^h \in \bar{S}^h$ 为已知策略时, 效用值 $u(s_i^c, \bar{s}_j^h) = a_{ij}$ 为确定数值; 当 $s^h = \tilde{s}^h \in \tilde{S}^h$ 时, 效用值 $u(s_i^c, \tilde{s}^h)$ 为区间 $I_{i,l}$ 内的不确定数值;
- 5: 将上述效用值表示为式 (13) 所示的 $m \times (n+q)$ 不确定效用矩阵 A , 其中 m, n 和 q 分别为 S^c, \bar{S}^h 和 \tilde{S}^h 中的元素个数;
- 6: 计算矩阵 A 的第 i 行的最小值 p_i^* ($i = 1, 2, \dots, m$), 即 $p_i^* = \min_{s^h \in S^h} u(s_i^c, s^h)$, 其中区间 $I_{i,l}$ 以其下界 $I_{i,l}^-$ 来计 ($l = 1, 2, \dots, q$);
- 7: 选取 S^c 中的策略 s^{c*} 来最大化 p_i^* , 策略 s^{c*} 即为 C 的在当前决策阶段的决策结果, 其中

$$s^{c*} = \arg \max_{s_i^c \in S^c} p_i^*.$$

- 8: 无人机群 C 执行策略 s^{c*} ;
- 9: 观测并更新 C 和 H 现存飞机的数量及其相对态势信息;

10: **end while**

输出: C 的决策序列

注释3 在上述无人机群的决策算法中, 我们没有显式地表述多无人机的协同方法. 事实上, 多无人机的协同蕴含在博弈模型的求解过程中. 这是由于在策略集的构造中, 机群的一个策略 $s^c \in S^c$ 指定了机群中每一架无人机的行动, 这本质上是多无人机的一种协同策略. 通过效用函数给出每一个协同策略的评估值, 并通过求解博弈决策模型, 最终选出最佳的协同方式.

4 仿真实验验证

4.1 实验环境与实验参数

我们通过几组对抗实验来验证所提方法的有效性. 对抗双方, 无人机群 C 和有人机群 H , 均包含两架飞机, 分别记为 C_1, C_2 和 H_1, H_2 . 仿真参数参见表 2, 其中 C_1 和 H_1 的参数, C_2 和 H_2 的参数是基本相同的, 以在相同条件下的仿真对比, 突出所提算法的有效性. 由表 2 可知, 有人机 H_1 的导弹性能不确定, H_2 的机动性能不确定, 进而使得空战对抗策略非完备. 在上述对抗实验中, C 采用所提出的 MMO 算法, 而 H 分别采用不同的算法进行对抗或逃逸. 仿真时, 每个决策阶段的仿真步长 $\Delta t = 4$ s.

4.2 对抗实验与结果分析

在每一个决策阶段, 首先构造对抗双方的策略集 S^c 和 S^h , 其中

$$S^c = \prod_{\mu=1}^2 \mathcal{M}^c \cup \mathcal{A}_\mu^c,$$

表 2 C 和 H 的基本参数
Table 2 Basic parameters of C and H

Parameter	C_1	C_2	H_1	H_2
Value of the fighter	5	6	5	6
Speed of the fighter (Ma)	1.6	1.3	1.6	1.3
Track angle ($^\circ$)	0	0	0	0
Course angle ($^\circ$)	0	0	180	180
Number of carried missiles	1	2	1	2
Missile attack distance (km)	7	9	7	9
Radar detection distance (km)	45	56	45	56
Missile damage probability (%)	53~77	61	53~77	61
Maneuverability parameter	1.2	1.3	0.8	0.7~0.9

$$S^h = \prod_{\nu=1}^2 \mathcal{M}^h \cup \mathcal{A}_\nu^h.$$

在对抗的初期阶段, 双方都在进行攻击前的机动, 因而 $\mathcal{A}_\mu^c = \emptyset, \mathcal{A}_\nu^h = \emptyset$ ($\mu = 1, 2; \nu = 1, 2$). 这时, 双方的策略集分别为 \mathcal{M}^c 和 \mathcal{M}^h . 随着对抗过程的推进, 有对方的飞机进入己方的导弹攻击区, 这时策略集中除了有机动操纵, 还有攻击操纵. 从而攻击策略集 \mathcal{A}_μ^c (\mathcal{A}_ν^h) 中含有 1 或 2 个元素. 由于机动库 \mathcal{M}^c (\mathcal{M}^h) 中含有 15 个基本机动动作, 因而 S^c 和 S^h 均包含 $15^2 = 225$ 到 $(15 + 2)^2 = 289$ 个策略. 注意到 C_1 和 H_1 的毁伤概率是一个不确定数值, 含有它们攻击行动的策略称为未知策略. 然后根据效用函数公式 (15) 计算每一个策略组合 $(s^c, s^h) \in S^c \times S^h$ 下的效用 $u(s^c, s^h)$, 其中包含了机动效用和攻击效用. 由于 C 和 H 均有 $225 \sim 289$ 个策略, 因而计算得到一个 $(225 \sim 289) \times (225 \sim 289)$ 的效用矩阵 \mathbf{A} . 注意当 S^h 中含有未知策略时, 效用矩阵 \mathbf{A} 中的元素将出现区间数据. 由于这些矩阵的规模大, 这里省去了具体的数值. 随后, 利用得到的效用矩阵 \mathbf{A} , 基于提出的 MMO 算法, 给出无人机群 C 的决策结果. 若 \mathbf{A} 中含有区间数据 $I_{i,l}$, 那么以 I_i 的下界 $I_{i,l}^-$ 来计. 然后计算效用矩阵 \mathbf{A} 每一行的最小值 p_i^* , 并选取 S^c 中的策略 s^{c*} 来最大化 p_i^* , 那么 C 的策略 s^{c*} 即为其优化策略. 需要指出的是, 尽管效用矩阵规模较大, 但在所提算法中仅用到求最值等简单计算, 其计算复杂度并不高, 因而能够满足空战决策的实时性要求.

图 4 给出了 C 和 H 在不同算法下的空战过程, 在这 4 次对抗实验中, C 均采用所提的 MMO 算法, 而 H 在前两组实验中分别采用传统的矩阵博弈算法 (MG) 和专家经验法 (EE) 进行对抗, H 在后两组实验中分别采用随机策略 (RS) 和分散直线 (SL) 进行逃逸. 我们以图 4(a) 为例分析 C 和 H 的对抗过程, 其中图例中 $C_1 \rightarrow H_1$ 表示 C_1 经过对抗最终击落了到达其攻击区的 H_1 , 同样地, $C_2 \rightarrow H_2$ 表示 C_2 最终击落了 H_2 . 由图 4(a) 可知, C_1 在第 10 个阶段对 H_1 发射了导弹, 并将其击落, 随后 C_1 向着 H_2 飞行, 但是 C_2 先于 C_1 达到攻击位置, 并于第 12 个决策阶段将 H_2 击落, 从而获得了空战的胜利. 其他对抗实验也可类似分析, C 均在一定的决策阶段将 H 全部击落. 图 5 给出了当 H 应用不同对比算法时 C 的相对态势变化. 由图 5 可知, 在 C 和 H 对抗的初期阶段, C 并没有很明显的相对态势优势, 随着空战过程的进行, C 的优势逐渐增大. 由于 H 所采用的这几种算法没有处理非完备策略的能力, 而 C 采用可处理非完备策略的 MMO 算法, 最终获得了明显的态势优势并最终击落所有有人机. 以上仿真验证了所提算法可处理存在非完备策略的决策问题.

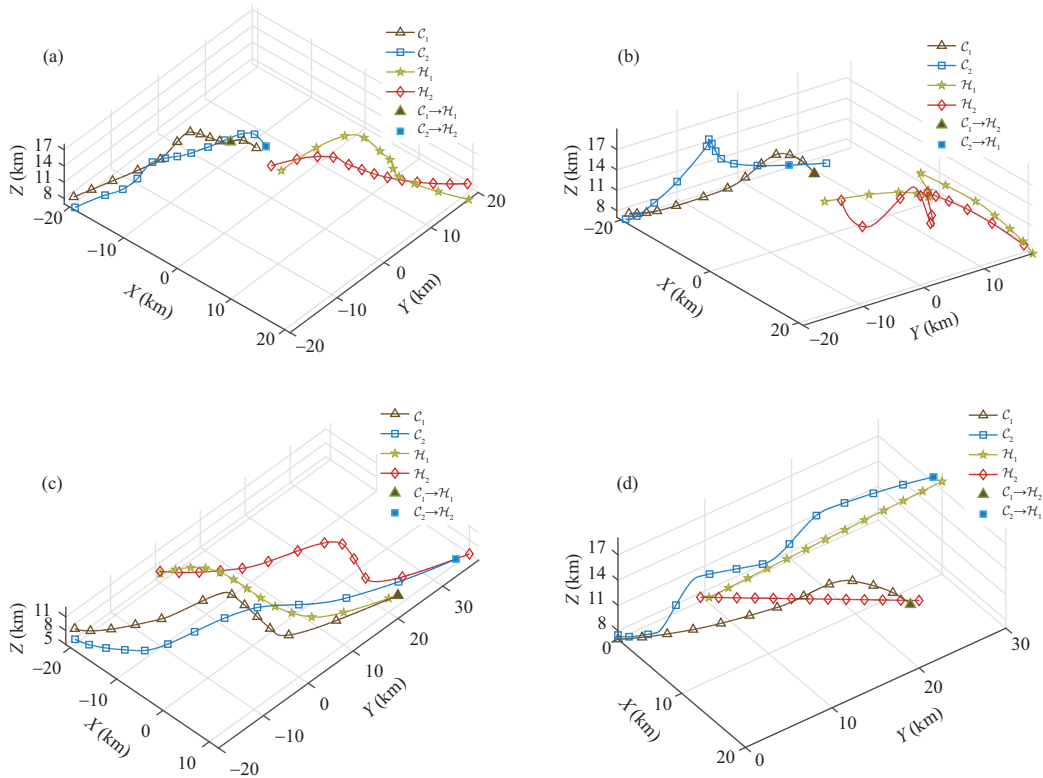


图 4 (网络版彩图) 不同算法下 C 和 H 的对抗过程图

Figure 4 (Color online) The confrontation process of C and H using different algorithms. (a) MG; (b) EE; (c) RS; (d) SL

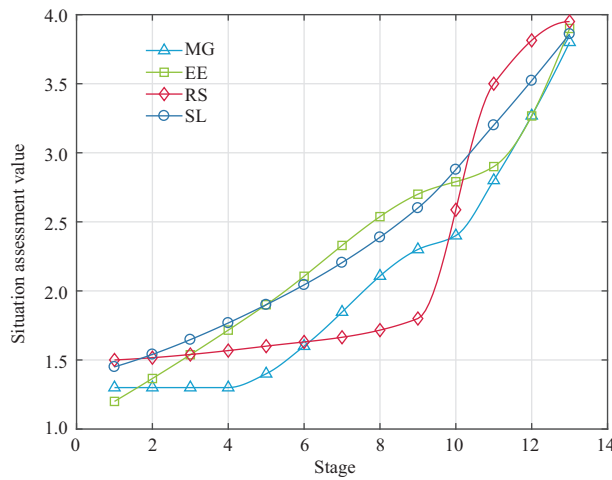


图 5 (网络版彩图) 不同算法下 C 的态势评估

Figure 5 (Color online) The situation evaluations of C using different algorithms

5 结论

针对人机对抗过程中对方策略不完全可知的情形, 本文提出了一种非完备策略集下无人机群的自

主决策方法. 无人机群和有人机群的空战对抗过程被建模为一个不确定博弈模型, 其中考虑了有人机群策略非完备的情形, 并建立了不确定效用矩阵. 基于经典矩阵博弈的最大最小策略, 给出了所建不确定博弈模型的决策方法. 通过对比实验验证了所提方法具有处理非完备信息的能力. 在实际空战中, 无人机群可能面对有人机和无人机协同作战的编队, 未来工作将考虑这种更复杂的空战场景, 以进一步提高无人机群的实战能力.

参考文献

- 1 Zhou W Q, Zhu J H, Kuang M C. An unmanned air combat system based on swarm intelligence. *Sci Sin Inform*, 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]
- 2 Li S Y, Chen M, Wang Y H, et al. Air combat decision-making of multiple UCAVs based on constraint strategy games. *Defence Tech*, 2022, 18: 368–383
- 3 Liu Z D, Chen M, Wu Q X, et al. Prediction of unmanned aerial vehicle target intention under incomplete information. *Sci Sin Inform*, 2020, 50: 704–717 [刘钻东, 陈谋, 吴庆宪, 等. 非完备信息下无人机空战目标意图预测. *中国科学: 信息科学*, 2020, 50: 704–717]
- 4 Yan D, Chen M, Wu Q X, et al. Determining safe flight area of UAVs based on variable weight threat assessment. *Sci Sin Inform*, 2021, 51: 663–677 [闫东, 陈谋, 吴庆宪, 等. 基于变权重威胁评估的无人机安全飞行区域确定方法. *中国科学: 信息科学*, 2021, 51: 663–677]
- 5 Xu G T, Meng Z Y, Long T, et al. Trajectory distributed receding planning for UAV swarms subject to limited communication distance. *Sci Sin Inform*, 2022, 52: 1527–1541 [徐广通, 孟子阳, 龙腾, 等. 通信距离受限下无人机集群轨迹分布式滚动规划. *中国科学: 信息科学*, 2022, 52: 1527–1541]
- 6 Li S Y, Chen M, Wang Y H, et al. A fast algorithm to solve large-scale matrix games based on dimensionality reduction and its application in multiple unmanned combat air vehicles attack-defense decision-making. *Inf Sci*, 2022, 594: 305–321
- 7 Wang M L, Wang L X, Yue T, et al. Influence of unmanned combat aerial vehicle agility on short-range aerial combat effectiveness. *Aerospace Sci Tech*, 2019, 96: 105534
- 8 Ruan W Y, Duan H B, Deng Y M. Autonomous maneuver decisions via transfer learning pigeon-inspired optimization for UCAVs in dogfight engagements. *IEEE/CAA J Autom Sin*, 2022, 9: 1639–1657
- 9 Zhang T, Li C C, Ma D Y, et al. An optimal task management and control scheme for military operations with dynamic game strategy. *Aerospace Sci Tech*, 2021, 115: 106815
- 10 Fan D D, Theodorou E A, Reeder J. Model-based stochastic search for large scale optimization of multi-agent UAV swarms. In: *Proceedings of 2018 IEEE Symposium Series on Computational Intelligence (SSCI)*, 2018. 2216–2222
- 11 Mi X M, Liao H C, Zeng X J, et al. The two-person and zero-sum matrix game with probabilistic linguistic information. *Inf Sci*, 2021, 570: 487–499
- 12 von Neumann J, Morgenstern O. *The Theory of Games and Economic Behavior*. Princeton: Princeton University Press, 1944
- 13 Guo L. Estimation, control, and games of dynamical systems with uncertainty. *Sci Sin Inform*, 2020, 50: 1327–1344 [郭雷. 不确定性动态系统的估计、控制与博弈. *中国科学: 信息科学*, 2020, 50: 1327–1344]
- 14 Churkin A I, Sauma E E, Pozo D, et al. Enhancing the stability of coalitions in cross-border transmission expansion planning. *IEEE Trans Power Syst*, 2022, 37: 2744–2757
- 15 Liu L, Zhang L, Liao S, et al. A generalized approach to solve perfect Bayesian Nash equilibrium for practical network attack and defense. *Inf Sci*, 2021, 577: 245–264
- 16 Duan H B, Li P, Yu Y X. A predator-prey particle swarm optimization approach to multiple UCAV air combat modeled by dynamic game theory. *IEEE/CAA J Automa Sin*, 2015, 2: 11–18
- 17 Nash J F. Equilibrium points in n -person games. *Proc Natl Acad Sci USA*, 1950, 36: 48–49
- 18 Maschler M, Solan E, Zamir S. *Game Theory*. Cambridge: Cambridge University Press, 2013
- 19 Li S Y, Wu Q X, Chen M, et al. Air combat situation assessment of multiple UCAVs with incomplete information. In: *Proceedings of Chinese Intelligent Systems Conference*. Singapore: Springer, 2020. 18–26
- 20 Xie Z. *Introduction to Game Theory*. Beijing: Science Press, 2010 [谢政. 对策论导论. 北京: 科学出版社, 2010]

- 21 Parthasarathy T, Rosenthal R W. Some topics in two-person games. *SIAM Rev*, 1972, 14: 356–357
- 22 Huang C Q, Dong K S, Huang H Q, et al. Autonomous air combat maneuver decision using Bayesian inference and moving horizon optimization. *J Syst Eng Electron*, 2018, 29: 86–97
- 23 Austin F, Carbone G, Falco M, et al. Game theory for automated maneuvering during air-to-air combat. *J Guidance Control Dyn*, 1990, 13: 1143–1149
- 24 Yang Q M, Zhang J D, Shi G Q, et al. Maneuver decision of UAV in short-range air combat based on deep reinforcement learning. *IEEE Access*, 2020, 8: 363–378
- 25 Xie J F, Yang Q M, Dai S L, et al. Air combat maneuver decision based on reinforcement genetic algorithm. *J Northwestern Polytechnical Univ*, 2020, 38: 1330–1338 [谢建峰, 杨啟明, 戴树岭, 等. 基于强化遗传算法的无人机空战机动决策研究. *西北工业大学学报*, 2020, 38: 1330–1338]
- 26 Hui Y L, Nan Y, Chen S D, et al. Dynamic attack zone of air-to-air missile after being launched in random wind field. *Chin J Aeronautics*, 2015, 28: 1519–1528
- 27 Buckley B, Jowers L. *Monte Carlo Methods in Fuzzy Optimization*. Berlin: Springer-Verlag, 2008

Human-computer gaming decision-making method in air combat under an incomplete strategy set

Shouyi LI¹, Mou CHEN^{1*}, Yuhui WANG¹, Qingxian WU¹ & Jianliang HE²

1. *College of Automation Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China;*

2. *Science and Technology on Electro-optic Control Laboratory, Luoyang 471000, China*

* Corresponding author. E-mail: chenmou@nuaa.edu.cn

Abstract In the real air combat environment, uncertainty is widespread. This paper considers the air combat decision-making problem in the human-computer gaming scenario, and a game decision-making method is proposed under an incomplete strategy set. First, the dynamic game process of unmanned and manned air vehicle swarms is decomposed into different decision-making stages, and the current decision-making stage is modeled as a zero-sum game model. In the established model, the situation where the strategy of the manned air vehicle swarm is not completely known is considered, and an uncertain payoff matrix is used to describe the incomplete strategies. Then, a solution method of the established game under an incomplete strategy set is proposed based on the concept of the maximum-minimum strategy in matrix games. Thus, the decision sequence of unmanned air vehicle swarms is given. Finally, the effectiveness of the proposed method is verified through simulation experiments.

Keywords unmanned air vehicle (UAV), air combat decision-making, human-computer gaming, incomplete strategy set, game theory, maximum-minimum strategy