

# 无人机集群对抗决策算法研究综述



李澐<sup>1</sup>, 黄诗怡<sup>1</sup>, 刘宏明<sup>1</sup>, 孙张俊<sup>2</sup>

1. 东南大学, 江苏 南京 210096

2. 航空工业西安飞行自动控制研究所, 陕西 西安 710076

**摘要:** 无人机集群博弈对抗已经成为未来战争的发展趋势, 无人机对抗决策算法的选择对提升无人机集群作战能力至关重要。本文深入探讨了基于规则的、基于博弈论的和基于神经网络的三大类无人机集群博弈对抗决策算法, 并对它们的优势和局限性进行了全面分析与总结。在此基础上, 提出将“基于多智能体强化学习的信用分配模型”和“基于角色的多智能体强化学习模型”应用于无人机集群博弈对抗的研究思路。最后, 强调了选择适当的决策算法对于提高无人机集群作战效能的重要性, 并为未来无人机对抗决策的发展提出了有益的建议, 为相关领域的研究和应用提供了深入见解。

**关键词:** 无人机集群; 博弈对抗; 专家系统; 博弈论; 多智能体强化学习

中图分类号: V279

文献标识码: A

DOI: 10.19452/j.issn1007-5453.2024.04.002

现代战争形式呈现多样化特征, 各种武器装备层出不穷, 作为一种替代人类执行高风险、高难度任务的工具, 无人机具备广泛的应用前景<sup>[1-2]</sup>。然而, 独立行动的无人机面临载荷能力有限、战术预测性不足等问题。在这一背景下, 无人机集群作战模式以其独特的规模大、分散度高、作战能力强、战术决策水平高等特点, 逐渐成为一种新兴的作战形式。通过充分发挥无人机集群的规模和分散度, 可以实现内部的快速大量信息传输和共享, 从而显著提高无人机集群的对抗博弈能力。在无人机集群作战中, 决策算法直接决定了无人机集群的作战能力。因此, 无人机决策算法的研究至关重要。

本文旨在全面总结目前主流的无人机集群对抗决策算法, 涵盖了基于规则、博弈论和神经网络的对抗决策算法。通过深入分析这些算法的适用场景和局限性, 揭示了它们在解决复杂军事问题上的差异和局限。针对基于强化学习的对抗决策算法中存在的信用分配和角色同质化问题, 本文提出了两种创新性的基于强化学习的对抗决策算法。最后, 结合当前研究现状, 本文强调了选择适当的决策算法对于提高无人机集群作战效能的紧迫性, 并为未来无人机对抗决策算法的发展指明了可行的研究方向, 为相关领域的研究和实际应用提供启示。

## 1 基于规则的无人机对抗决策算法

基于规则的无人机对抗决策算法通过设置规则来指导无人机在对抗环境中做出决策。首先出现的是基于专家知识的对抗决策算法, 这类算法通过大量的先验专家知识来构建规则, 无人机在实际作战中将战场态势与规则进行匹配来做出决策。基于专家知识的算法具有决策速度快的优点, 但是对未知情况的适应性差, 随后出现了基于贝叶斯网络的对抗决策算法, 这类方法通过贝叶斯网络构建各因素之间的依赖关系, 对未知情况有较好的适应能力。

### 1.1 基于专家系统的方法

如图1所示, 基于专家系统的无人机对抗决策算法将专家知识构建为规则库和综合数据库, 在实际作战中将当前输入的战场态势等信息与专家系统中的规则进行匹配, 进而做出决策<sup>[3]</sup>。

1975年, 美国国家航空航天局(NASA)开发了自适应机动逻辑(AML), 首次使用专家系统进行机动决策<sup>[4]</sup>。在面对形式瞬息万变的战场态势时, 基于专家系统的方法能够快速准确地进行决策。这一特点使得基于专家系统的方法得到了充分发展和广泛应用。目前基于专家系统的方法已经成为无人机博弈对抗领域发展最成熟的技术之一。

赵威等<sup>[5]</sup>利用专家系统来判断敌机的威胁等级并对攻

收稿日期: 2023-09-19; 退修日期: 2024-01-15; 录用日期: 2024-02-27

基金项目: 航空科学基金(20200058069001)

引用格式: Li Wei, Huang Shiyi, Liu Hongming, et al. Review of UAV swarm air-combat decision-making algorithms[J]. Aeronautical Science & Technology, 2024, 35(04): 9-17. 李澐, 黄诗怡, 刘宏明, 等. 无人机集群对抗决策算法研究综述[J]. 航空科学技术, 2024, 35(04): 9-17.

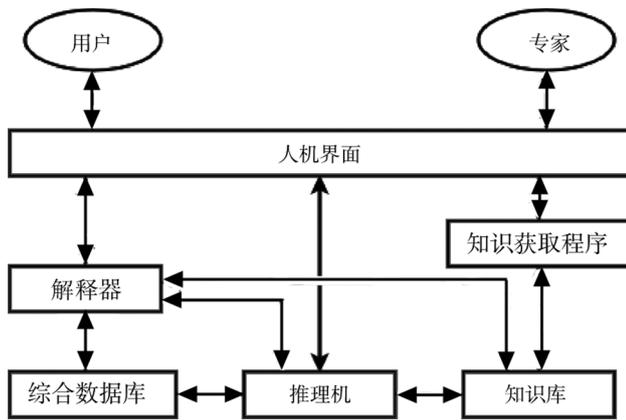


图1 专家系统示意图

Fig.1 Expert system diagram

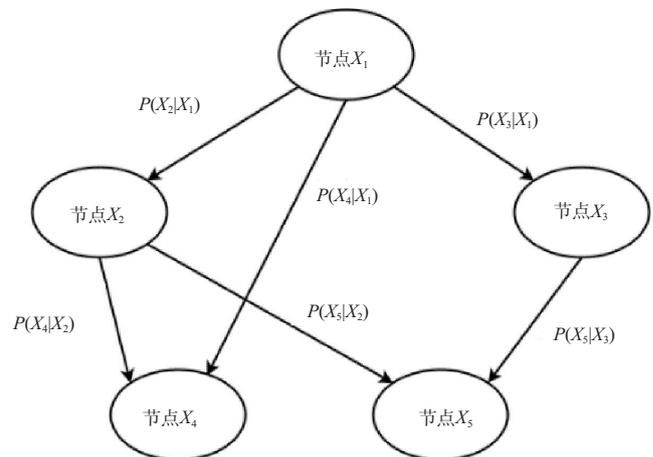


图2 贝叶斯网络示意图

Fig.2 Bayesian network diagram

击目标进行排序,提出了基于空战任务和目标编队组合的协同攻击排序方法。随着深度学习的发展,基于深度学习的专家系统已经涌现出来,这类方法采用深度神经网络进行决策,能够处理更加复杂的情况,并且能够通过学习自主提取特征。Mao Yiming等<sup>[5]</sup>采用基于深度学习的方法来构建专家系统,利用Transformer网络提出了一种能够从历史数据中学习知识的战术状态决策过程建模方法。该方法能够随着实战经验的增加自动更新专家系统以适应不断变化的战斗场景。

快速决策的特点使得基于专家系统的方法得到了广泛应用。但同时,专家系统也受限于知识库和推理机制的局限性,针对未知情况可能表现不佳。将专家系统与其他方法(如贝叶斯网络、深度学习)结合实现共同决策,以弥补其局限性是未来基于专家系统的无人机决策方法的发展趋势。

## 1.2 基于贝叶斯网络的方法

如图2所示,贝叶斯网络是一个有向无环图,它通过节点和由父节点指向子节点的有向边来表示变量之间的依赖关系,并使用条件概率表描述这些依赖关系。令 $G$ 为定义在节点 $\{X_1, X_2, \dots, X_N\}$ 上的一个贝叶斯网络,其联合概率分布表示为各个节点的条件概率分布的乘积

$$P(X) = \prod_i P_i(X_i | \text{Par}_G(X_i)) \quad (1)$$

式中, $\text{Par}_G(X_i)$ 表示节点 $X_i$ 的父节点, $P_i(X_i | \text{Par}_G(X_i))$ 表示节点条件概率,节点条件概率定义了给定父节点值时,该节点取值的概率分布。贝叶斯网络利用这些条件概率来计算给定观察到的证据的情况下,其他节点取值的概率分布。动态贝叶斯网络在时间维度上拓展了贝叶斯网络,不仅包含了各节点之间的依赖关系,还包含了描述了各节点的值

如何随时间变化。在无人机集群对战中,贝叶斯网络通常用于整合各方信息,提供更准确的战场态势认知,进而协助决策。

孟光磊等<sup>[6]</sup>构建了一个4层动态贝叶斯模型,能根据不确定信息来推理敌方轨迹和战略意图,并利用从飞行员经验中获取的先验知识确定对抗策略,能较好地适应战场的形势变化。Ren Zhi等<sup>[7]</sup>提出了一种基于不完全信息动态博弈的协同决策方法,使用动态贝叶斯网络来推断敌方无人机的战术意图,最后结合强化学习框架来求解空战博弈模型。针对无人机空战适应性差、决策模式不合理等问题,Meng Guanglei等<sup>[8]</sup>提出了一种基于目标轨迹预测的最优空战决策方法。首先根据空战情况,设置不同的性能指标函数,然后利用贝叶斯网络识别敌方意图,根据实时情况分析结果,选择不同的性能指标函数进行最优控制。

通过概率分布,贝叶斯网络可以量化不确定性,并提供对不确定性的推理能力,这使得其在面对不完整数据时具有鲁棒性。但目前基于贝叶斯网络的方法通常需要与其他方法结合起来选择对抗策略。

## 2 基于博弈论的无人机集群博弈对抗决策算法

博弈论研究博弈双方如何基于对方的策略来优化自己的策略,基于博弈论的方法也就是研究双方无人机集群之间策略交互的方法<sup>[9]</sup>。不同于依赖大量先验知识库的方法,基于博弈论的策略交互方法通过逐步自学习达到最优策略。在应对无人机集群空战的强对抗性和强动态性方面,博弈论的应用能够生成更适用于实战的决策计划。本

节聚焦于博弈论在无人机集群博弈对抗中的应用,首先介绍博弈对抗决策模型,包括确定策略集合、态势优势函数以及求解纳什均衡的过程。随后,探讨博弈对抗决策与控制方法,包括粒子群算法、蚁群算法、人工鱼群算法和狼群算法等不同群体智能算法的应用。

## 2.1 博弈对抗决策模型

基于纳什均衡的决策模型在无人机集群博弈对抗中被广泛研究<sup>[9-10]</sup>。该方法首先需要确定对抗双方无人机的策略集合和态势优势函数,然后求解纳什均衡。

影响无人机态势优势函数的因素主要由双方对抗关系得到,如图3所示。由图3可知,影响因素可包括双方无人机之间的角度、速度、距离等。 $v_A$ 和 $v_B$ 分别表示无人机A和B的速度, $\alpha_A$ 和 $\alpha_B$ 分别表示无人机A和B的方位角, $d_{AB}$ 表示无人机A和B之间的距离。综合多种优势函数,根据权重不同,可以得到无人机间总的态势优势函数<sup>[11]</sup>。

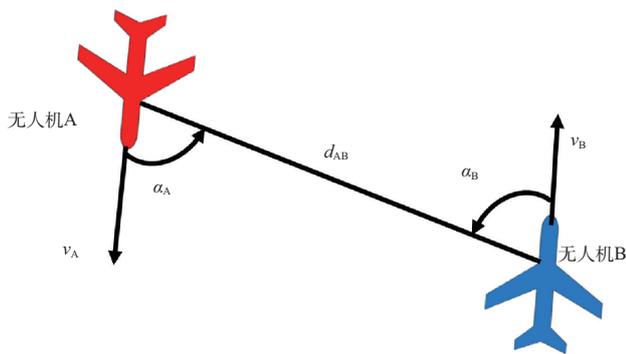


图3 无人机对抗方位图

Fig.3 Azimuth map of unmanned aerial vehicle confrontation

在确定对抗双方的策略集合和态势优势函数后,进一步结合集群中每个无人机的态势优势,得到对抗双方各自的总体态势优势函数矩阵,并根据具体场景中的收益函数,得到双方各自的收益矩阵。基于由双方策略集合组成的混合策略以及双方的收益矩阵,可定义博弈对抗策略模型。混合策略的纳什均衡解即为无人机集群的最优策略。为了求得纳什均衡解,常将最优策略的约束转化为优化问题,最优值对应的混合策略就是博弈问题的纳什均衡点<sup>[12]</sup>。

上述过程给出了单位时间内对抗双方最优策略的求解方法,由于无人机集群的对抗过程是动态的,因此需要将上述过程应用于博弈对抗时间段内的每个时间点。也就是从初始时间起,在每个单点时间内,双方无人机集群根据求得的最优策略,得出下一时刻的状态,包括无人机姿态和位置等,从而进行下一步的最优策略求解。

姚宗信等<sup>[10]</sup>将对抗双方可能的互相攻击方式作为策略集合,根据双方的攻击有效性和攻击代价建立优势态势函数和收益矩阵,并基于过往战术经验和模拟仿真结果对混合策略的分布进行预测。盛磊等<sup>[11]</sup>为对抗的攻守双方建立了各自的动态攻防场景和策略,根据攻守方的不同假设构建收益矩阵。

部分研究通过构建除纳什均衡模型外的其他模型来对无人机集群对抗问题进行建模,其中也包括对纳什均衡模型的变体。惠一楠等<sup>[13]</sup>提出,在实际空战场景中,无人机之间只能知道他人的决策分布,而难以观察到他人的真实决策,所以无人机集群的博弈对抗并不是完全信息博弈,而是非完全信息博弈。因此,他们基于非完全信息动态博弈理论,构建贝叶斯纳什均衡模型并求解。黄宇铭等<sup>[14]</sup>提出只研究纳什均衡条件意味着模型只有单步预见的能力,因此他们综合考虑了包括纳什稳定性、一般元理性、对称元理性和序贯稳定性在内的4种稳定性,使所得决策拥有更高的预见水平。陈侠等<sup>[15]</sup>提出在实际空战环境中,由于各种因素的干扰(如传感器精度和环境干扰等),无人机所能获得的信息往往不是精确的,因此他们建立了模糊信息下的态势优势函数,并构建了基于模糊信息的收益矩阵,提出了模糊信息下的博弈对抗决策模型。

## 2.2 博弈对抗决策与控制方法

基于2.1节提出的博弈对抗决策模型,无人机集群通过控制个体无人机决策,在群体水平上表现出复杂行为,来达到不同的作战能力。在这一过程中,模拟自然界各种群体行为而产生的群体智能算法,因为具有较好的可扩展性、并行性、实现简单等特点,被广泛应用于无人机集群博弈对抗的决策。

粒子群算法是一种模拟鸟群行为的随机搜索算法,具有能够在保持多样性的同时快速收敛的特点,其基本思想如图4所示。图4中,A为粒子的当前位置,B为粒子的个体极值位置,C为群体的全局最优位置。绿色为个体当前速度(矢量),蓝色为向个体极值位置的速度(矢量),红色为向全局最优位置的速度(矢量),黄色为基于粒子群算法得到的目标速度(矢量),受算法参数影响。粒子群算法首先初始化一群粒子表示可行的解,每个粒子具有位置、速度、适应度三个特征。随后粒子在问题的搜索空间中运动,其运动方向同时跟踪该粒子的个体极值位置和群体的全局最优位置。粒子每运动一次,即每更新一次位置,就计算一次适应度值,并与自己的个体极值和群体的全局最优值比较,从而更新各自极值位置和群体最优位置。粒子

群算法的标准形式如式(2)所示

$$\begin{cases} v_i = v_i + c_1 \cdot (pbest_i - x_i) + c_2 \cdot (gbest_i - x_i) \\ x_i = x_i + v_i \end{cases} \quad (2)$$

式中,  $v_i$  为粒子的速度;  $x_i$  为粒子的当前位置;  $c_1$  和  $c_2$  为学习因子;  $pbest_i$  和  $gbest_i$  分别为个体极值位置和群体的全局最优位置。

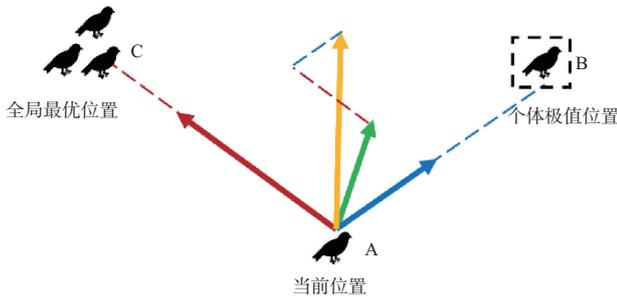


图4 粒子群算法的基本思想

Fig.4 Basic idea of particle swarm optimization algorithm

陈侠等<sup>[16]</sup>将粒子群算法和区间数多属性方案排序方法相结合,求解不完全信息下无人机攻防博弈中的纳什均衡问题。Duan Haibin等<sup>[17]</sup>引入生物界的捕食者-猎物机制,提出混合捕食者-猎物粒子群算法,缓解了粒子群算法容易陷入局部最优的问题,在无人机博弈对抗的任务分配中得到有效应用。该方法通过将任务分配方案表示为双方的可选策略集,将粒子群算法应用于无人飞行器的任务分配问题,并通过使用粒子群算法求解混合纳什均衡来获得博弈双方的协同任务分配结果。在每个决策步骤,针对敌方决策,捕食者-猎物粒子群算法通过计算矩阵对策的混合纳什均衡来最大化其自身收益。在此过程中,通过将粒子分为捕食者和猎物两类,并相应地调整它们的速度,捕食者-猎物算法获得了较高的寻优性能。多种粒子群算法变体也被提出来求解纳什均衡。

蚁群算法受到蚂蚁觅食路径的启发,是一种寻找优化路径的概率算法<sup>[18]</sup>,如图5所示。蚁群算法的标准形式如式(3)所示

$$\begin{cases} P_{ij}^k = \frac{(c_{ij}(t))^{\alpha} \cdot n_{ij}(t)}{\sum((c_{ij}(t))^{\alpha} \cdot n_{ij}(t))} \\ c_{ij}(t+1) = (1 - \alpha) \cdot c_{ij}(t) + \Delta c_{ij} \\ \Delta c_{ij} = \sum \Delta c_{ij}^k \end{cases} \quad (3)$$

式中,  $P_{ij}^k$  是  $t$  时刻蚂蚁  $k$  从节点  $i$  转移到节点  $j$  的概率;  $n_{ij}(t)$  为启发函数;  $\alpha$  是信息素重要程度因子,  $\alpha$  表示信息素挥发程度;  $c_{ij}^k$  是第  $k$  只蚂蚁在连接节点  $i$  与节点  $k$  的路径上释放信息素而增加的信息素浓度;  $\Delta c_{ij}^k$  为所有蚂蚁在连接节点  $i$

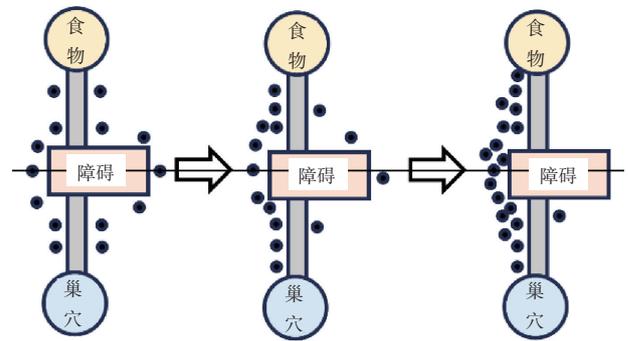


图5 蚁群算法的基本思想

Fig.5 Basic idea of ant colony optimization algorithm

与结点  $k$  的路径上释放信息素而增加的信息素浓度。

Zhen Ziyang等<sup>[19]</sup>针对无人机集群的协同搜索和攻击任务规划问题,提出改进的分布式蚁群算法对无人机进行任务分配。Yang Fan等<sup>[20]</sup>针对无人机策略中完全未知的搜索空间,提出了改进的蚁群算法行为准则和信息素图的更新原则,达到了更高的覆盖率和搜索效率。

人工鱼群算法利用鱼群典型的觅食行为、聚群行为、追尾行为和随机游动行为,解决复杂非线性优化问题<sup>[21]</sup>,如图6所示。Li Zhanwu等<sup>[22]</sup>提出改进的鱼群算法,以保证无人机生存能力和最小化武器消耗为目标,应用于无人机集群空战中的武器分配问题。Jiang Haobo等<sup>[23]</sup>针对无人机协同作战的目标分配问题,将蚁群算法中的拥挤度引入鱼群算法,所得到的混合算法在搜索前期不易陷入局部最优,在搜索后期收敛速度快。

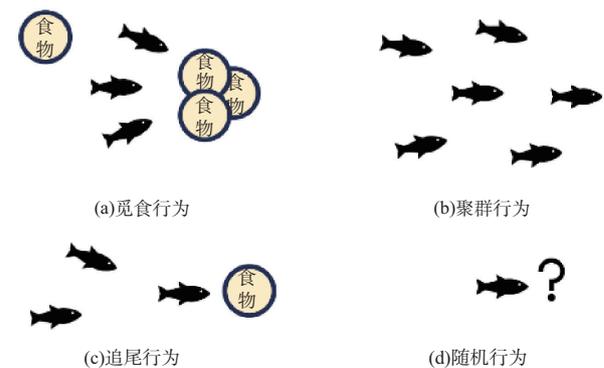


图6 鱼群算法的基本思想

Fig.6 Basic idea of fish swarm optimization algorithm

狼群算法模拟狼群的社会等级层次和群体捕食行为来达到优化的目的,并平衡局部搜索和全局搜索<sup>[24]</sup>。Hua Xiang等<sup>[25]</sup>将无人机集群视为多个并行的狼群,提出具有多种群机制的狼群算法,通过采用多层优化策略,有效地解决

集群对抗中的多目标分配问题。

综上所述,博弈论在无人机集群博弈对抗中提供了一种框架,用于研究无人机之间的策略交互和最优决策计划,能够确定对抗双方无人机的策略集合,构建态势优势函数,并求解纳什均衡。而群体智能算法在无人机集群博弈对抗中可以帮助实现复杂的群体行为和优化目标,这类具有可拓展性和并行性特点的算法可用于控制个体无人机的决策,使无人机集群在博弈对抗中表现出协同性和适应性,从而提高无人机集群的性能,帮助其应对动态和复杂的对抗环境。

### 3 基于神经网络的无人机集群对抗决策算法

基于神经网络的无人机对抗决策算法通过神经网络来实现对对抗态势的感知并做出决策,能动态适应战场态势的变化。由于训练神经网络需要大量数据,而实战数据较难获取,所以目前训练神经网络的数据大多来自仿真平台。根据训练方式的不同,可以分为基于强化学习和基于深度学习的算法,基于强化学习的方法通过与仿真环境的交互来不断地直接优化策略网络,而基于深度学习的方法往往关注如何通过神经网络提取战场态势信息,从而做出决策。本节首先介绍这两类方法的工作方式,然后介绍该方向的代表性工作,最后总结这类方法的使用场景和限制。

#### 3.1 基于强化学习的对抗决策算法

强化学习无须依赖模型或先验信息,而是通过不断试错,根据从环境中得到的奖励来优化自身策略。如图7所示,强化学习包含智能体、环境、状态、动作和奖励5个主要组成部分。当智能体执行动作时,环境进入一个新的状态,并发出相应的正奖励信号或负奖励信号。在接收到这些信息后,智能体会根据策略选择新的动作以适应新的状态和奖励信号。

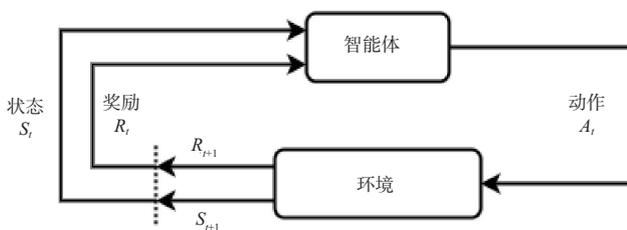


图7 强化学习中智能体与环境交互过程

Fig.7 The process of interaction between agent and environment in reinforcement learning

深度神经网络与强化学习结合后,智能体可以在仿真环境中迭代优化无人机集群对抗算法,适用于解决那些建模困难、决策复杂且多变的无人机集群决策控制问题。

王尔申等<sup>[26]</sup>从无人机集群实战特性入手,通过对多智能体的奖励函数进行创新,来搭建基于非完全信息的多智能体柔性行动器-评判器(MASAC),并通过试验证明该方法可以在无人机集群对抗博弈的场景取得很好的性能。但该方法还不能很好地应用于存在异质无人机集群的对抗博弈中。符小卫等<sup>[27]</sup>针对主流算法多智能体深度确定性策略梯度(MADDPG)中存在的值函数高估问题,引入TD3算法并将其扩展到多智能体领域中,提出了优先经验回放多智能体双延迟深度确定性策略算法(PER-MATD3),该算法采用优先经验回放机制,提高了算法的收敛速度和稳定性,并在障碍物随机分布的复杂环境下取得了很好的对抗效果。文永明等<sup>[28]</sup>针对复杂场景中无人机集群对抗的突防轨迹和目标分配等决策问题,提出了集群对抗多耦合任务智能决策方法。该方法解决了无人机对抗中决策空间大、场景不确定和在线决策耦合任务多的问题,并在多种随机场景的红蓝博弈中验证了该方法的有效性和先进性。

基于强化学习的无人机集群对抗决策算法适用于建模困难的集群对抗环境,但是基于强化学习的方法需要在试错的基础上优化自身,因此往往需要先在仿真环境下训练再投入实际战场环境。此外,奖励函数直接决定了强化学习算法的性能,如何设置奖励函数也是基于强化学习的方法当前面临的重要问题。

针对多智能体强化学习中的信用分配问题,李淮团队<sup>[29]</sup>提出了“基于多智能体强化学习的信用分配模型”,并在此基础上设计了“基于注意力机制的内部奖励网络”用于提高智能体之间团队协作的有效性,以及“内部奖励和外部奖励混合网络”使得网络输出的总奖励值更加适应环境的变化。该模型从建立精准内部奖励的角度,根据每个智能体在团队协作任务中的贡献度,生成对应的内部奖励,并动态结合来自多智能体环境中的全局奖励。针对无人机集群博弈对抗场景,相较于现有代表性算法,该模型有望实现更快收敛速度和更高获胜率。

针对多智能体强化学习算法中的策略同质化问题,李淮团队<sup>[30]</sup>提出了“基于角色的多智能体强化学习模型”。该模型包含了一种新的角色定义方法以及角色分配策略,使得智能体能够根据实际环境依照动态的角色分配从不同的角度做出决策,避免了决策的狭隘性。同时,为了解决引入

角色选择模块所导致的动作价值估计效率低下问题,该模型在双流估计架构的基础上,采用了一种新的动作价值估计双流架构。针对无人机集群博弈对抗场景,相较于现有代表性算法,该模型有望达到更快收敛速度、更佳鲁棒性及更高获胜率。

### 3.2 基于深度学习的对抗决策算法

神经网络是模仿生物神经系统的计算模型,具有强大的非线性拟合能力和自学习能力,可以适应无人机集群对抗决策中的环境变化和目標变化。

Schvaneveldt等<sup>[31]</sup>根据在模拟空战中测量的各种参数的值,用人工神经网络预测交战结果。证实了人工神经网络模型在空战机动决策领域的可行性和有效性,并实现了具有强鲁棒性的自适应无人机决策控制方法。张宏鹏等<sup>[32]</sup>利用含有36种机动动作的飞行仿真数据构建样本,并利用该样本训练深度神经网络,对所有动作所对应的未来态势进行预测,使无人机可以在较短的时间内做出更具优势的决策。

为了进一步改善决策算法的不足,目前很多研究聚焦于改进神经网络的结构或将其他决策算法与神经网络相结合,来提升无人机对抗决策性能。李锋等<sup>[33]</sup>提出了一种利用模糊神经网络进行决策的方法,将超视距空战分为进入、攻击和脱离三个阶段,并根据不同阶段的特点,设计了相应的模糊神经网络结构和规则库,实现了对空战态势信息和目标预测信息的快速处理和决策输出。Li Bo等<sup>[34]</sup>提出了一种基于卷积神经网络(CNN)的智能机动决策模型。将空战态势数据输入CNN中,得到机动决策变量,然后依照仿真实验的评价指标调整CNN的参数和结构。结果表明该方法比基于人工神经网络的方法具有更好的决策性能。

利用深度神经网络,可以有效地处理空战中的不确定性和非线性性,充分地数据特征和知识进行自动提取和表达,实现对空战态势的快速判断和适应性决策。但基于神经网络的方法需要大量的训练数据和合理的网络结构,并且缺乏可解释性。

## 4 结论与展望

无人机集群已成为未来战争中的重要一环。无人机集群对抗决策技术对无人机集群的智能化至关重要,本文通过梳理现有文献,对无人机集群对抗决策方法及各方法的优势和适用场景进行归纳总结,包含了基于规则的方法、基于博弈论的方法和基于神经网络的方法。三种方法各有优

势,适应不同的应用场景。基于规则的方法在可解释性上表现出色,但在处理复杂和动态的对抗态势时,可能无法灵活适应快速变化的战场环境。基于博弈论的方法在理论建模和策略优化上具有优势,但在处理实际对抗环境时受制于对完全信息的依赖,对于不完全信息和快速变化的战局表现不佳。基于神经网络的方法在处理非线性和不确定性方面表现出色,但需要大量实战数据进行训练,且其性能受训练数据质量和奖励函数设计的影响较大。

综合当前的研究现状,认为无人机集群对抗决策领域以下几个问题值得深入研究。

#### (1) 单一种类的决策算法的限制

无人机集群对抗中,采用单一种类的决策算法会受到一定限制。例如,基于规则的方法在未知场景下决策可信度较低;基于博弈论的方法对环境建模要求较高;基于神经网络的方法需大量数据进行训练。未来决策算法的发展趋势是将不同方法相互结合,突破单一方法的限制,提高决策的准确性和可靠性。

#### (2) 深度学习缺乏可解释性

当前许多无人机决策算法借助深度学习来提高模型的表征能力。但由于无人机集群对抗系统面对的是严肃残酷的战场环境,对于可靠性要求极高。一旦出现难以预测的错误,或由于无法解释决策而产生战术配合失误,将会造成巨大损失。因此决策算法的发展依赖于深度学习的理论突破。

#### (3) 仿真与实际作战环境差距大

由于实际作战环境中各种因素错综复杂,在仿真环境中能顺利完成任务的算法不一定能适应实际作战环境。一方面,完善仿真平台,使其能更好地模拟实际战场;另一方面,提高决策算法对不确定性的适应能力,是未来无人机决策的重要发展方向。

AST

## 参考文献

- [1] 张炯,牛欢,陈雪.基于多传感器融合的无人机应急着陆功能研究[J].航空科学技术,2022,33(11):11-20.  
Zhang Jiong, Niu Huan, Chen Xue. Research on UAV emergency landing function based on multi-sensor fusion[J]. Aeronautical Science & Technology, 2022, 33(11): 11-20. (in Chinese)
- [2] 郭强,何胜杰,程家林,等.一种无人机自主规避导弹的威胁度评估方法[J].航空科学技术,2022,33(7):8-14.  
Guo Qiang, He Shengjie, Cheng Jialin, et al. A threat

- evaluation method of autonomous UAV avoidance missile[J]. *Aeronautical Science & Technology*, 2022, 33(7): 8-14. (in Chinese)
- [3] 赵威. 基于专家系统的双机协同攻击决策技术研究[D]. 西安: 西北工业大学, 2007.  
Zhao Wei. Research on decision making technique of two-machine cooperative attack based on expert system[D]. Xi'an: Northwestern Polytechnical University, 2007.(in Chinese)
- [4] Burgin G H, Fogel L J, Phelps J P. An adaptive maneuvering logic computer program for the simulation of one-on-one air-to-air combat: General description[R]. NASA-CR- 2582, 1975.
- [5] Mao Yiming, Xia Zhijie, Li Qingwei, et al. Accurate decision-making method for air combat pilots based on data-driven[C]. *International Conference on Data Mining and Big Data*, 2022: 439-448.
- [6] 孟光磊, 刘德见, 周铭哲, 等. 近距离空战训练中的智能虚拟对手决策与导引方法[J]. *北京航空航天大学学报*, 2022, 48(6): 937-949.  
Meng Guanglei, Liu Dejian, Zhou Mingzhe, et al. Intelligent virtual opponent decision making and guidance method in short-range air combat training[J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2022, 48(6): 937-949.(in Chinese)
- [7] Ren Zhi, Zhang Dong, Tang Shuo, et al. Cooperative maneuver decision making for multi-UAV air combat based on incomplete information dynamic game[J]. *Defence Technology*, 2022, 27:308-317.
- [8] Meng Guanglei, Zhang Cheng, Liu Shouye, et al. UAV Attack and defense optimization guidance method based on target trajectory prediction[C]. *2019 IEEE International Conferences on Ubiquitous Computing & Communications (IUCC) and Data Science and Computational Intelligence (DSCI) and Smart Computing, Networking and Services (SmartCNS)*. IEEE, 2019: 495-499.
- [9] Omidshafiei S, Tuyls K, Czarnecki W M, et al. Navigating the landscape of multiplayer games[J]. *Nature Communications*, 2020, 11(1): 5603.
- [10] 姚宗信, 李明, 陈宗基. 基于博弈论模型的多机协同对抗多目标任务决策方法[J]. *航空计算技术*, 2007(3): 7-11.  
Yao Zongxin, Li Ming, Chen Zongji. Mission decision-making method of multi-aircraft cooperative attack multi-object based on game theory model[J]. *Aeronautical Computing Technique*, 2007(3): 7-11.(in Chinese)
- [11] 盛磊, 时满红, 亓迎川, 等. 基于态势演化博弈的无人机集群动态攻防[EB/OL]. (2023-06-21). <http://kns.cnki.net/kcms/detail/11.2422.TN.20230619.1902.004.html>.  
Sheng Lei, Shi Manhong, Qi Yingchuan, et al. Dynamic offense and defense of UAV swarm based on situation evolution game[EB/OL]. (2023-06-21). <http://kns.cnki.net/kcms/detail/11.2422.TN.20230619.1902.004.html>.(in Chinese)
- [12] Yan Ming, Yuan Huimin, Xu Jie, et al. Task allocation and route planning of multiple UAVs in a marine environment based on an improved particle swarm optimization algorithm [J]. *EURASIP Journal on Advances in Signal Processing*, 2021, 94: 1-23.
- [13] 惠一楠, 朱华勇, 沈林成. 无人机攻防对抗不完全信息动态博弈方法研究[J]. *兵工自动化*, 2009, 28(1): 4-7.  
Hui Yinan, Zhu Huayong, Shen Lincheng. Study on dynamic game method with incomplete information in UAV attack-defends campaign[J]. *Ordnance Industry Automation*, 2009, 28(1): 4-7.(in Chinese)
- [14] 黄宇铭, 葛冰峰, 侯泽强, 等. 基于冲突分析图模型的多无人机协同空战博弈[J]. *系统工程理论与实践*, 2023(9): 2714-2725.  
Huang Yuming, Ge Bingfeng, Hou Zeqiang, et al. Multi-unmanned aerial vehicle cooperative air combat gaming based on graph model for conflict resolution[J]. *Systems Engineering-Theory & Practice*, 2023(9): 2714-2725.(in Chinese)
- [15] 陈侠, 赵明明, 徐光延. 基于模糊动态博弈的多无人机空战策略研究[J]. *电光与控制*, 2014, 21(6): 19-23+34.  
Chen Xia, Zhao Mingming, Xu Guangyan. Fuzzy dynamic game based operation strategy for multiple UAVs[J]. *Electronics Optics & Control*, 2014, 21(6): 19-23+34. (in Chinese)
- [16] 陈侠, 刘敏, 胡永新. 基于不确定信息的无人机攻防博弈策略研究[J]. *兵工学报*, 2012, 33(12): 1510-1515.  
Chen Xia, Liu Min, Hu Yongxin. Study on UAV offensive/defensive game strategy based on uncertain information[J]. *Acta Armamentarii*, 2012, 33(12): 1510-1515.(in Chinese)
- [17] Duan Haibin, Li Pei, Yu Yaxiang. A predator-prey particle

- swarm optimization approach to multiple UCAV air combat modeled by dynamic game theory[J]. *IEEE/CAA Journal of Automatica Sinica*, 2015, 2(1): 11-18.
- [18] Pendharkar P C. An ant colony optimization heuristic for constrained task allocation problem[J]. *Journal of Computational Science*, 2015, 7: 37-47.
- [19] Zhen Ziyang, Xing Dongjing, Gao Chen. Cooperative search-attack mission planning for multi-UAV based on intelligent self-organized algorithm[J]. *Aerospace Science and Technology*, 2018, 76: 402-411.
- [20] Yang Fan, Ji Xiuling, Yang Chengwei, et al. Cooperative search of UAV swarm based on improved ant colony algorithm in uncertain environment[C]. 2017 IEEE International Conference on Unmanned Systems (ICUS). IEEE, 2017: 231-236.
- [21] Pourpanah F, Wang Ran, Lim C P, et al. A review of artificial fish swarm algorithms: Recent advances and applications[J]. *Artificial Intelligence Review*, 2023, 56(3): 1867-1903.
- [22] Li Zhanwu, Chang Yizhe, Kou Yingxin, et al. Approach to WTA in air combat using IAFSA-IHS algorithm[J]. *Journal of Systems Engineering and Electronics*, 2018, 29(3): 519-529.
- [23] Jiang Haobo, Li Song, Lin Chi, et al. Research on target assignment method based on ant colony-fish group algorithm [J]. *Journal of Physics: Conference Series*. IOP Publishing, 2019, 1419(1): 012002.
- [24] Mirjalili S, Mirjalili S M, Lewis A. Grey wolf optimizer[J]. *Advances in Engineering Software*, 2014, 69: 46-61.
- [25] Hua Xiang, Wang Zhao, Yao Hongjuan, et al. Research on many-to-many target assignment for unmanned aerial vehicle swarm in three-dimensional scenarios[J]. *Computers & Electrical Engineering*, 2021, 91: 107067.
- [26] 王尔申, 刘帆, 宏晨, 等. 基于MASAC的无人机集群对抗博弈方法[J]. *中国科学: 信息科学*, 2022, 52(12):2254-2269.
- Wang Ershen, Liu Fan, Hong Chen, et al. A MASAC-based adversarial game approach for UAV clusters[J]. *China Science: Information Science*, 2022, 52(12): 2254-2269. (in Chinese)
- [27] 符小卫, 徐哲, 朱金冬, 等. 基于PER-MATD3的多无人机攻防对抗机动决策[J]. *航空学报*, 2023, 44(7):196-209.
- Fu Xiaowei, Xu Zhe, Zhu Jindong, et al. Multi-UAV attack-defense countermeasure maneuver decision based on PER-MATD3[J]. *Acta Aeronautica et Astronautica Sinica*, 2023, 44(7):196-209. (in Chinese)
- [28] 文永明, 石晓荣, 黄雪梅, 等. 一种无人机集群对抗多耦合任务智能决策方法[J]. *宇航学报*, 2021, 42(4): 504-512.
- Wen Yongming, Shi Xiaorong, Huang Xuemei, et al. An intelligent decision-making method for UAV cluster countermeasures with multiple coupled tasks[J]. *Journal of Astronautics*, 2021, 42(4):504-512.(in Chinese)
- [29] Li Wei, Liu Weiyan, Shao Shitong, et al. Attention-based intrinsic reward mixing network for credit assignment in multi-agent reinforcement learning[J/OL]. *IEEE Transactions on Games*: 1-13. (2023-03-29). <https://doi.org/10.1109/TG.2023.3263013>.
- [30] Li Wei, Qiu Ziming, Shao Shitong, et al. MDDP: Making decisions from different perspectives in multi-agent reinforcement learning[J/OL]. *IEEE Transactions on Games*: 1-14. (2023-11-01). <https://doi.org/10.1109/TG.2023.3329376>.
- [31] Schvaneveldt R W, Goldsmith T E, Benson A E, et al. Neural network models of air combat maneuvering[R]. Armstrong Laboratory, 1992.
- [32] 张宏鹏, 黄长强, 轩永波, 等. 基于深度神经网络的无人作战飞机自主空战机动决策[J]. *兵工学报*, 2020, 41(8):1613-1622.
- Zhang Hongpeng, Huang Changqiang, Xuan Yongbo, et al. Maneuver decision of autonomous air combat of unmanned combat aerial vehicle based on deep neural network[J]. *Acta Armamentarii*, 2020, 41(8):1613-1622.(in Chinese)
- [33] 李锋, 孙隆和, 佟明安. 基于模糊神经网络的超视距空战战术决策研究[J]. *西北工业大学学报*, 2001, 19(2):317-322.
- Li Feng, Sun Longhe, Tong Ming'an. A tactical decision support system for bvr air combat based on neural network[J]. *Journal of Northw estern Polytechnical University*, 2001, 19(2): 317-322.(in Chinese)
- [34] Li Bo, Liang Shiyang, Tian Linyu, et al. Intelligent aircraft maneuvering decision based on CNN[C]// *Proceedings of the 3rd International Conference on Computer Science and Application Engineering*, 2019: 1-5.

## Review of UAV Swarm Air-combat Decision-making Algorithms

Li Wei<sup>1</sup>, Huang Shiyi<sup>1</sup>, Liu Hongming<sup>1</sup>, Sun Zhangjun<sup>2</sup>

1. Southeast University, Nanjing 210096, China

2. AVIC Xi'an Flight Automatic Control Research Institute, Xi'an 710076, China

**Abstract:** UAV swarm air-combat has become the development trend of future warfare, and the selection of UAV swarm air-combat decision-making algorithms is crucial for improving the UAV swarm combat ability. This paper delve into three types of UAV swarm air-combat decision-making algorithms based on rules, game theory, and neural networks, and comprehensively analyze and summarize their advantages and limitations. On this basis, this paper propose to apply the multi-agent reinforcement learning based credit assignment model and role-based multi-agent reinforcement learning model and design for UAV swarm air-combat. Finally, it emphasize the importance of selecting appropriate decision algorithms to improve the combat effectiveness of UAV clusters, and provide useful suggestions for the development of UAV countermeasures decision-making in the future, providing in-depth insights for research and application in related fields.

**Key Words:** UAV swarm; air-combat; expert system; game theory; multi-agent reinforcement learning

---

**Received:** 2023-09-19; **Revised:** 2024-01-15; **Accepted:** 2024-02-27

**Foundation item:** Aeronautical Science Foundation of China(20200058069001)