

4-15-2024

Simulation Verification and Decision-making Key Technologies of Unmanned Swarm Game Confrontation: A Survey

Xiaolong Liang

Air Traffic Control and Navigation College, Air Force Engineering University, Xi'an 710051, China; Shaanxi Key Laboratory of Meta-Synthesis for Electronic and Information System, Xi'an 710051, China, afeu_lxl@sina.com

Aiwu Yang

Air Traffic Control and Navigation College, Air Force Engineering University, Xi'an 710051, China; Shaanxi Key Laboratory of Meta-Synthesis for Electronic and Information System, Xi'an 710051, China

Jiaqiang Zhang

Air Traffic Control and Navigation College, Air Force Engineering University, Xi'an 710051, China; Shaanxi Key Laboratory of Meta-Synthesis for Electronic and Information System, Xi'an 710051, China

Yueqi Hou

Air Traffic Control and Navigation College, Air Force Engineering University, Xi'an 710051, China; Shaanxi Key Laboratory of Meta-Synthesis for Electronic and Information System, Xi'an 710051, China

See next page for additional authors

Follow this and additional works at: <https://dc-china-simulation.researchcommons.org/journal>



Part of the [Artificial Intelligence and Robotics Commons](#), [Computer Engineering Commons](#), [Numerical Analysis and Scientific Computing Commons](#), [Operations Research, Systems Engineering and Industrial Engineering Commons](#), and the [Systems Science Commons](#)

This Special Column is brought to you for free and open access by Journal of System Simulation. It has been accepted for inclusion in Journal of System Simulation by an authorized editor of Journal of System Simulation. For more information, please contact xtfzxb@126.com.

Simulation Verification and Decision-making Key Technologies of Unmanned Swarm Game Confrontation: A Survey

Abstract

Abstract: Unmanned swarm game confrontation is a new combat mode and plays a crucial role in intelligent warfare. Its core is the autonomous generation of a series of game confrontation decision sequences to "empower" the swarm. The progress of system simulation verification for the unmanned swarm game confrontation is analyzed. The key technologies of the autonomous decision-making are discussed from three aspects, technology based on expert systems and game theory, technology based on swarm intelligence and optimization theory, and technology based on neural networks and reinforcement learning. The key technology research conducted by the author's team on the autonomous decisionmaking is presented. The development directions of the unmanned swarm game confrontation are proposed.

Keywords

unmanned swarm, game confrontation, autonomous decision-making, system simulation validation

Authors

Xiaolong Liang, Aiwu Yang, Jiaqiang Zhang, Yueqi Hou, Ning Wang, Xiao Huang, and Junbin Gong

Recommended Citation

Liang Xiaolong, Yang Aiwu, Zhang Jiaqiang, et al. Simulation Verification and Decision-making Key Technologies of Unmanned Swarm Game Confrontation: A Survey[J]. Journal of System Simulation, 2024, 36(4): 805-816.

无人集群博弈对抗系统仿真验证及决策关键技术综述

梁晓龙^{1,2}, 杨爱武^{1,2}, 张佳强^{1,2}, 侯岳奇^{1,2}, 王宁^{1,2}, 黄骁³, 龚俊斌³

(1. 空军工程大学 空管领航学院, 陕西 西安 710051; 2. 陕西省电子信息系统系统集成重点实验室, 陕西 西安 710051;
3. 中国舰船研究设计中心, 湖北 武汉 430064)

摘要: 无人集群博弈对抗是一种新兴的作战样式, 在智能化战争扮演着至关重要的作用, 其核心是自主生成博弈对抗决策序列, 为集群“赋能”。分析了无人集群博弈对抗系统仿真验证的进展; 从基于专家系统和博弈论的技术、基于群体智能和优化理论的技术, 以及基于神经网络和强化学习的三个技术方面论述了自主决策关键技术, 以及课题组在自主决策上开展的相关工作; 提出了无人集群博弈对抗的发展方向。

关键词: 无人集群; 博弈对抗; 自主决策; 系统仿真验证

中图分类号: V37;TP391 文献标志码: A 文章编号: 1004-731X(2024)04-0805-12

DOI: 10.16182/j.issn1004731x.joss.23-0072

引用格式: 梁晓龙, 杨爱武, 张佳强, 等. 无人集群博弈对抗系统仿真验证及决策关键技术综述[J]. 系统仿真学报, 2024, 36(4): 805-816.

Reference format: Liang Xiaolong, Yang Aiwu, Zhang Jiaqiang, et al. Simulation Verification and Decision-making Key Technologies of Unmanned Swarm Game Confrontation: A Survey[J]. Journal of System Simulation, 2024, 36(4): 805-816.

Simulation Verification and Decision-making Key Technologies of Unmanned Swarm Game Confrontation: A Survey

Liang Xiaolong^{1,2}, Yang Aiwu^{1,2}, Zhang Jiaqiang^{1,2}, Hou Yueqi^{1,2}, Wang Ning^{1,2}, Huang Xiao³, Gong Junbin³

(1. Air Traffic Control and Navigation College, Air Force Engineering University, Xi'an 710051, China;
2. Shaanxi Key Laboratory of Meta-Synthesis for Electronic and Information System, Xi'an 710051, China;
3. China Ship Development and Design Center, Wuhan 430064, China)

Abstract: Unmanned swarm game confrontation is a new combat mode and plays a crucial role in intelligent warfare. Its core is the autonomous generation of a series of game confrontation decision sequences to "empower" the swarm. The progress of system simulation verification for the unmanned swarm game confrontation is analyzed. The key technologies of the autonomous decision-making are discussed from three aspects, technology based on expert systems and game theory, technology based on swarm intelligence and optimization theory, and technology based on neural networks and reinforcement learning. The key technology research conducted by the author's team on the autonomous decision-making is presented. The development directions of the unmanned swarm game confrontation are proposed.

Keywords: unmanned swarm; game confrontation; autonomous decision-making; system simulation validation

0 引言

从纳卡冲突、俄乌战争等现代化战争来看，无人机在侦察情报、对地攻击、斩首行动等作战行动中发挥着重要作用。可以预见，未来高自主无人机、无人前置射手、无人前置传感器等多类型无人节点组成的无人集群将有潜力成为主战武器，无人集群博弈对抗这种新兴的作战样式，将在智能化战争扮演着至关重要的作用，深刻影响着未来战争的发展走向^[1-4]。

无人集群是一种战术、技术、装备高度融合的作战体系，无人集群体系作战通过开放式体系架构，对体系内节点实施“积木式”编组，实时塑造态势，实时构造杀伤网，快速闭合最优杀伤链，从而夺取决策行动优势。从上述制胜机理来看，无人集群博弈对抗相较于单机对抗，主要区别在于作战思想由“平台中心”向“体系中心”转变，作战结构由“杀伤链”向“杀伤网”转变^[5-6]，核心在于“破体系”，关键在于智能自主决策，发展自主决策关键技术能够有效提升无人集群博弈对抗体系作战能力。

系统仿真验证是推动无人集群博弈对抗从构想到现实的重要环节。通过设定多样化作战想定和任务，在不同输入条件和参数下，测试博弈对抗算法的有效性和适应性。为推动无人集群从实验验证向实战运用发展，亟需开展自主智能决策技术软硬件在环仿真推演，减小“人在回路”影响，通过采集大量博弈对抗数据，反馈算法模型迭代升级，进而验证并提升无人集群博弈对抗能力。

本文从博弈对抗决策系统仿真验证研究进展入手，分析了国内外在系统仿真验证方面的进展，重点从知识、优化和学习三个角度论述了无人集群自主智能决策关键技术以及课题组相关工作，提出了无人集群博弈对抗发展方向。

1 博弈对抗决策系统仿真验证进展

为了推动无人集群作战概念与作战样式的创

新，加快无人集群自主决策技术发展和运用，国内外开展了多项单机/无人集群自主决策系统仿真验证工作。通过设计单机博弈对抗和集群博弈对抗多样化场景，搭建系统仿真验证平台，以仿真形式推动决策关键技术验证与落地，主要形式体现为人机仿真对抗、半实物仿真对抗、虚拟仿真对抗等。

1.1 单机博弈对抗系统仿真验证

2020年，DAPPA开展AlphaDogfight实验验证，以人机仿真对抗、虚拟仿真对抗形式，进行近距空战格斗决策关键技术仿真验证工作，验证人工智能算法在空战场景下的决策能力，如图1所示。AlphaDogfight作战想定为1v1近距空战格斗，交战双方作战能力相等，武器攻击区设定为机头方向锥形区域。苍鹭系统公司(Heron Systems)开发的空战系统“隼”(Falco)采用AdeptRL深度强化学习框架以5:0的战绩击败美空军F-16人类飞行员，同样洛克希德-马丁公司(Lockheed Martin)采用分层强化学习方法来解决近距空战决策问题^[7]。强化学习算法通过强大的探索和试错能力，能够获取超过人类经验知识的策略。

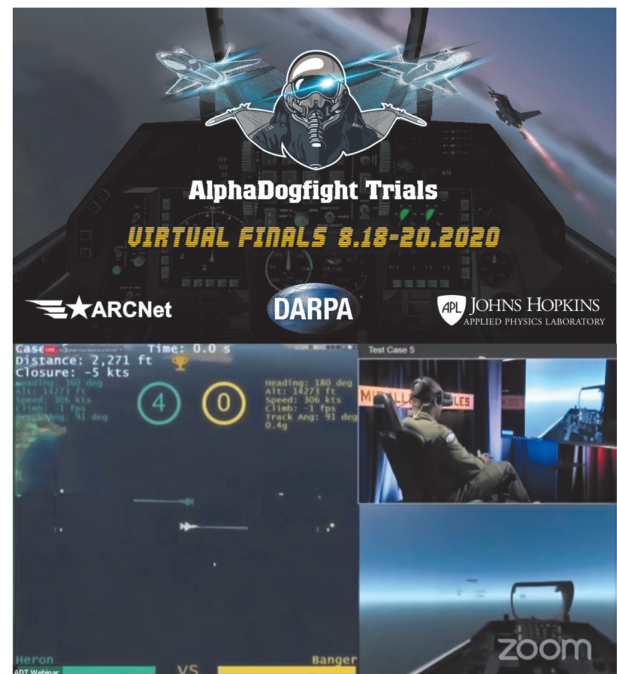


图1 AlphaDogfight空战场景
Fig. 1 Scene of air combat in AlphaDogfight

AlphaDogfight 主要目的在于测试人工智能算法在空战中的应用能力, 作为美军空战演进项目的一部分, 旨在提高人工智能算法的可信度, 促进从飞行员空战到人机协同空战再到自主空战的演进。

2023年, 第一届“天行杯”智能空战算法挑战赛, 采用人机仿真对抗、虚拟仿真对抗形式, 验证单机空战博弈对抗自主决策和控制能力。挑战赛设置 1v1 超视距空战和近距空战两个科目, 红蓝双方以多种初始态势进入, 依托机载传感器获取态势信息, 在智能算法支撑下, 实施战术决策与控制, 其中智能算法需要涵盖飞行控制、信息融合、机动决策等模块。挑战赛通过搭建高保真度无人自主空战博弈对抗环境, 验证了基于规则、基于优化、基于学习, 以及混合决策的智能决策算法有效性和适应性, 推动智能决策算法向实际装备迁移应用。

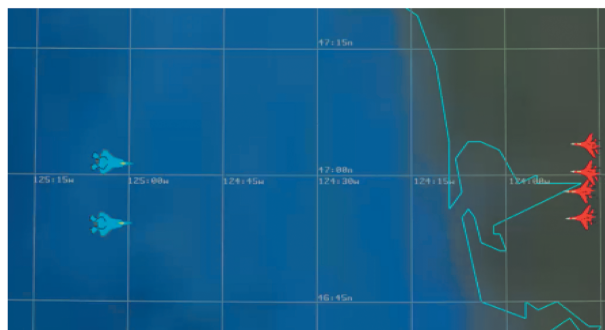
1.2 无人集群博弈对抗系统仿真验证

2016年, 美国辛辛那提大学 Nicholas Ernest 等研发人工智能空战系统 ALPHA, 以人机仿真对抗形式开展无人集群空战博弈对抗系统仿真验证工作, ALPHA 系统在模拟空战中击败美军空军顶级飞行员^[8]。作战想定为飞行员控制 2 架蓝方飞机对抗 ALPHA 的 4 架红方飞机, 蓝方在态势感知、武器性能和武器数量上占优, 如图 2 所示。

ALPHA 核心技术为遗传模糊逻辑^[9-10], 通过遗传算法对模糊树推理系统进行优化训练从而实现智能化决策, 决策时间达到 ms 级别, 并且小型单片机就能满足计算硬件需求。ALPHA 决策本质上属于规则驱动, 相较于固定规则, 实现了空战决策规则可进化, 充分展现了人工智能技术对于无人集群自主空战博弈对抗能力的推动。2018年, 美国佐治亚理工学院, 开展基于模型的强化学习方法在无人机集群要地防空和集群空战中应用研究, 旨在开发博弈对抗环境下, 无人机集群有效的策略生成方法^[11]。



(a) 空战模拟器中人与ALPHA对战



(b) 空战场对抗态势

图2 Alpha 空战对抗场景

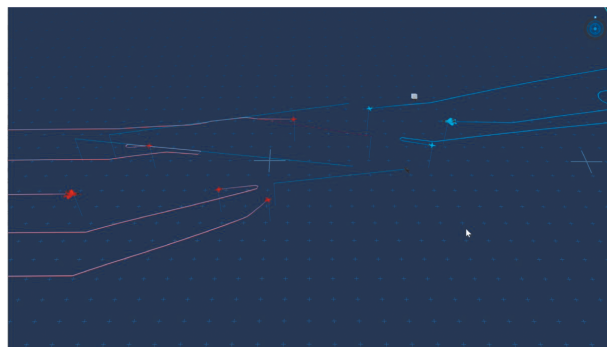
Fig. 2 Scene of air combat in Alpha

2021年, 中国指挥控制学会举办首届空中智能博弈大赛, 开展异构无人集群空战系统虚拟仿真验证工作, 作战想定设置为异构、透明多机协同空战场景, 1 架有人机和 4 架无人机进行自由空战博弈, 模拟雷达探测功能和空空导弹攻击过程, 支持规则驱动、机器学习、强化学习等建模仿真验证方式, 如图 3 所示。

由于有人机与无人机均可独立程序控制, 有人机在一定程度上仅为高价值和高性能无人节点, 上述想定可视为异构无人集群空战博弈对抗。比赛通过虚拟仿真对抗形式验证异构无人集群系统在编队行动、目标分配、协同制导、目标打击等关键技术上的自主决策能力。除此之外, 第二届“无人争锋”中开展的无人集群突防与拦截场景下半实物自主决策系统仿真验证, 海上无人平台立体跨域协同虚拟仿真验证等。



(a) 模拟雷达探测范围



(b) 虚拟空战对抗过程

图3 空中智能博弈大赛场景
Fig. 3 Scene of Air Intelligence Game

国内外正在加速进行无人集群人机混合、半实物以及虚拟仿真博弈对抗验证工作，目的在于研究无人集群博弈对抗关键技术，积累和沉淀博弈对抗算法、数据、知识，促进自主决策能力提升，为无人集群作战指挥决策“赋能”。从决策技术角度来看，基于专家知识的系统(ALPHA系统等)通过提高决策速度来调整最优战术；基于智能优化的系统通过探索决策空间来提高单步决策精度并获取近似最优解；基于策略学习的系统(Falco系统等)则通过海量模拟训练来获取对人类飞行员的决策优势。尽管上述场景属于小规模集群或单机间的对抗，但不可否认以进化算法、深度强化学习为代表的人工智能算法，具备解决复杂无人集群博弈对抗决策问题的潜力。

1.3 博弈对抗决策系统仿真验证平台

无人集群系统仿真验证正在如火如荼展开，依托开发的系统验证平台和仿真对抗形式，开展无人

集群系统决策关键技术验证。无人集群博弈对抗系统仿真验证平台整体架构如图4所示，分为仿真平台层、交互接口层和智能体开发层。仿真平台层承担核心推演功能，负责载入作战场景和数据，通过交互接口层，获取态势并接收任务指令，进行实时态势显示并记录对抗日志。在智能体开发层，红蓝对抗双方决策算法以单独子进程形式存在，相互隔离，主要包括两个模块：观测构建模块将观测数据生成与决策算法需求对应的观测数据；决策算法模块接收观测数据，输出指令动作数据。

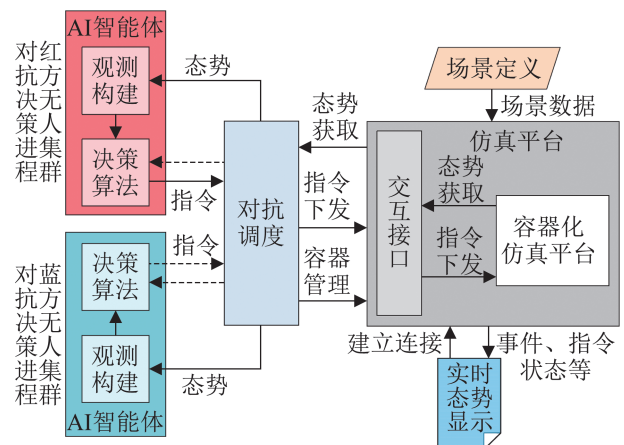


图4 系统仿真验证平台整体架构

Fig. 4 Architecture of platform for system simulation and verification

关于无人集群系统仿真验证平台研制，需要考虑以下几个方面：一是平台的实际性和模型粒度，平台与真实作战环境的差异将直接影响到集群系统向实际运用的迁移，环境模型中需要考虑作战全要素、全流程。此外，无人平台、传感器、武器模型粒度需要与实际装备贴合；二是平台的扩展性和兼容性，验证平台需要支持任务场景、无人平台性能、作战条令条例可根据任务需求配置，参数可定义，算法与平台柔性集成；三是平台仿真推演速度，支持并行分布式训练，平行仿真，满足算法探索与交互需求；四是注重实况、虚拟和构造的融合，利用数字孪生、LVC技术，支撑无人集群博弈软/硬件在环仿真推演，减小“人在回路”影响。

面向未来无人集群作战体系复杂、态势复杂、协同交互关系复杂问题需求, 以体系架构设计为牵引、仿真验证平台为支撑、决策算法为核心, 将促进无人集群自主作战从仿真环境走向复杂开放现实作战环境。

2 无人集群博弈对抗自主决策关键技术

无人集群博弈对抗是一项复杂的作战任务, 其面临的决策问题主要是博弈对抗策略生成, 即以博弈对抗智能决策技术为支撑, 自主生成拦截集群的对抗决策序列。需要解决的实际问题主要包括机动决策、目标分配、协同攻击决策等, 解决上述问题的关键技术从知识、优化、学习角度可以归纳为基于专家系统和博弈论、群体智能和优化理论, 以及神经网络和强化学习的技术。

2.1 基于专家系统和博弈论的技术

随着无人集群作战概念的演进, 专家对无人集群作战理解更加深刻并积累大量经验知识, 专家决策系统则通过建立知识规则库, 实施推理决策。典型的有美军 ALPHA 空战系统中遗传模糊树^[9-10], 专家系统以快速推理决策占据博弈对抗主动, 但博弈对抗策略相对固定, 适用于小规模简单场景下的无人集群博弈对抗问题, 在大规模、战场突发事件多场景下, 需要解决专家系统中规则的扩展性、规则间冲突以及参数的适应性问题。

基于博弈论的技术是在专家经验的基础上, 建立无人集群博弈对抗模型, 求解纳什均衡策略。例如, 针对无人机集群空战机动占位决策问题, 将敌我双方的博弈对抗建立为零和博弈问题, 通过运用带邻域搜索的改进 DO(double oracle)算法求解大规模零和矩阵^[12]; 考虑时间敏感信息, 对约束策略占位问题进行建模, 并利用线性规划和线性不等式的约束策略求解^[13]; 针对无人集群博弈对抗兵力分配问题, 通过将大规模对抗问题进行分组, 基于 DO 算法搜索兵力分配纳什均衡解^[14];

针对无人机集群在区域防御博弈场景攻防决策问题, 将敌我博弈对抗建立为 N-联盟非合作博弈纳什均衡解搜索问题, 通过极值搜索方法估计由动态平均一致性协议得到的代价函数梯度^[15]。

博弈论技术以其精确解占据策略优势, 但是对于复杂无人集群博弈对抗问题, 在不满足限制性条件时, 难以得到精确的全局纳什均衡解, 此时需要简化问题模型或者放宽纳什均衡要求, 会导致决策问题模型实际贴合度和纳什均衡策略求解效率以及质量之间的矛盾点。

在基于知识和规则的决策技术方面, 课题组针对无人集群博弈对抗动态运转和演化决策问题, 提出了一种基于专家系统的状态-事件-条件-动作 (state-event-condition-action, SECA) 的演化决策机制, 可作为无人集群博弈对抗决策框架。通过专家知识定义无人集群博弈对抗场景下的状态集、事件集、条件集、动作集, 以及状态转移函数。状态转移函数根据观测信息匹配不同状态下规则进行决策, 进而驱动整个决策系统运转, 如图 5 所示。

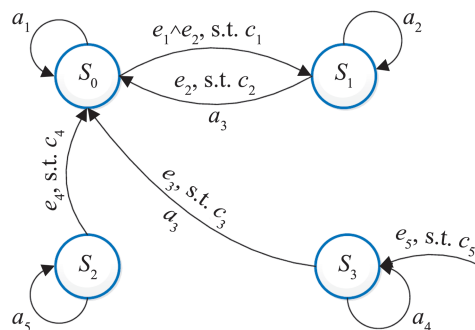


图5 SECA模型中状态转移示意图

Fig. 5 State transition in SECA model

SECA模型综合运用了有限状态机 (finite state machine, FSM) 的状态转移^[16]和ECA规则的条件拆分^[17], 将无人集群决策问题进行层次化划分, 当满足一定条件时可以自动触发演化阶段/任务进程/系统功能向另一个阶段/进程/功能转换, 主要分为状态决策和动作决策。

状态决策是指以战场态势信息为输入, 通过状态转移函数匹配相应的作战条件, 决定无人集

群将转移到下一状态。动作决策是指无人集群在当前状态下，相应触发事件和作战条件激活时，决定无人集群应执行相应动作，从而将高维复杂决策问题进行降维解耦，降低决策维度、复杂度，以及系统响应频率，具备良好的可解释性和扩展性。图6为无人集群空战博弈中单架无人机相关状态、事件监控的时序图，可作为空战策略仿真验证的可视化呈现。

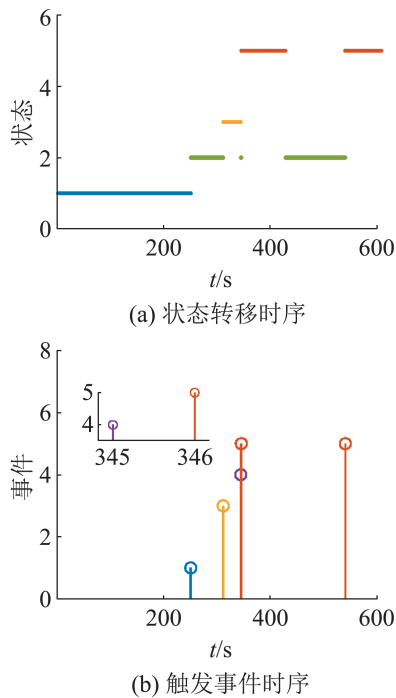


图6 空战策略仿真验证中状态和事件时序

Fig. 6 States and events sequence in simulation verification of an air combat strategy

2.2 基于群体智能和优化理论的技术

群体智能技术通过模拟自然界中生物种群行为，群内个体通过交互和合作来解决复杂的作战问题。相较于博弈论技术，模型限制性条件相对宽松，算法能够在解空间搜索寻优问题的最优解。群体智能技术可以分为集中式架构和分布式架构。

在集中式决策架构方面，针对任务分配/武器-分配问题，建立全局优化目标函数和约束条件，通过优化算法迭代寻优最优解，包括基于捕食者粒子群算法的博弈论方法^[18]以及改进的非支

配排序遗传算法^[19]。集中式架构需要中心节点汇聚全局信息，再根据任务要求设置目标函数和约束条件后，能够迭代收敛至优化问题的最优解。

在分布式决策架构方面，针对无人机集群攻防决策中目标分配和运动决策问题，通过建立环境感知、协同决策和执行等协同模型来实现群体协同决策，基于分布式自组织攻防对抗决策算法求解每一步决策^[20]；针对异构大规模集群目标分配问题，通过设置任务投标-竞标机制来获得任务，自适应调整策略和行动^[21]。分布式架构无需中心节点，复杂作战决策问题分散至群内个体协商解决，对通信要求高，收敛速度较慢。基于群体智能的决策技术面临的主要问题在于算法的实时性差、结果不稳定以及自适应参数设定，优化求解过程会导致系统运行的“空转”。

相反，优化理论在解决无人集群博弈问题时，实时性高并且能够得到稳定解。例如，针对无人集群拦截部署点优化以及集结点优化问题，通过建立最优部署模型和最优集结模型，采用凸优化理论求解部署点解析表达式^[22]以及分布式优化理论实现指定时间最优集结^[23]。上述优化理论的运用都能得到模型解析解，决策实时性高，但都涉及到目标函数凸性等严格限制性条件，限制了技术的适用范围。但是不可否认优化理论潜力，通过适当简化问题模型达到实时决策，通过决策速度优势占据博弈态势主动。

在基于优化的决策技术方面，课题组针对无人集群博弈对抗群体智能汇聚问题，提出一种基于分布式决策方法的群智汇聚框架，来解决无人集群博弈对抗作战复杂全局决策问题，如图7所示。将复杂作战全局问题分散至个体协商解决，充分发挥群内个体的智慧，通过分布式决策方法，包括分布式优化、分布式协商、分布式纳什均衡，群内个体求解局部问题形成高层级决策依赖信息，基于分布式决策交互模型，在群内分享交互高层级决策依赖信息，实现决策方案的一致最优汇聚，并表现为全部初级信息下集中式决策的最优形态。

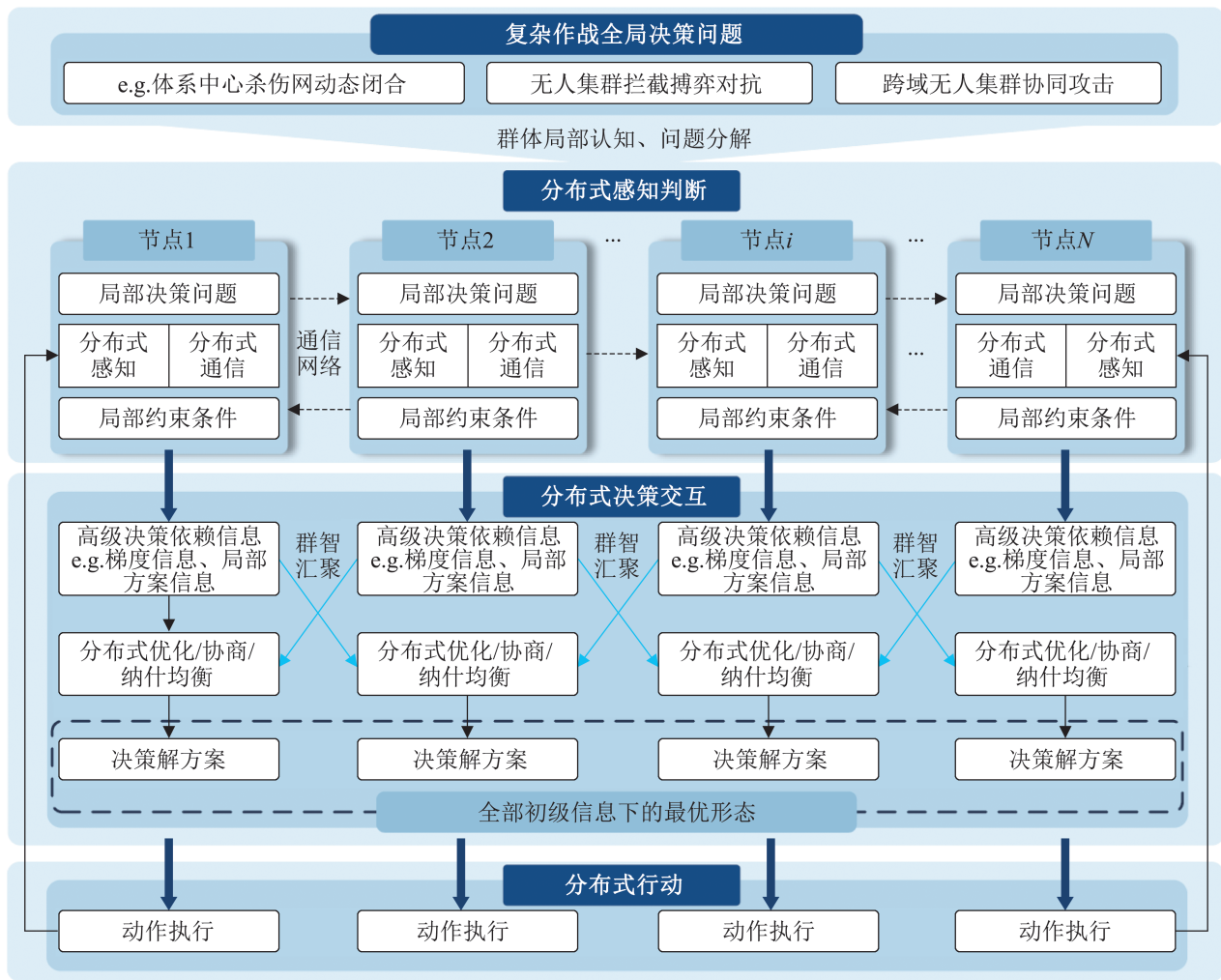


图7 分布式“群智汇聚”框架

Fig. 7 Distributed crowd intelligence aggregation framework

2.3 基于神经网络和强化学习的技术

基于神经网络的决策技术，通过训练数据的拟合和挖掘，优化网络参数，以网络参数形式存储空战规则和知识，对训练数据的数量和质量要求较高。在无人集群博弈对抗场景下，如果缺乏足够的训练样本，仅靠神经网络训练难以学习到有效的集群博弈策略，基于神经网络的决策技术目前在单机空战决策上应用较多^[24]。

强化学习方法无需学习样本，通过与战场环境的交互迭代值函数，更新自身策略，与深度学习结合形成的深度强化学习技术已在视频游戏、围棋、机器人领域得到成功应用。深度学习解决

“是什么”问题，强化学习解决“怎么做”问题，逐步成为探索无人集群博弈对抗策略的研究热点。根据有无模型区分，可以将深度强化学习技术分为有模型强化学习和无模型强化学习。

在有模型方向上，针对无人机集群要地防空和自主空战场景决策问题，基于经验知识设计空战优势模型，有3种辅助训练思路：①采用基于模型随机选择决策方案^[11]，并逐步改进优化目标函数；②将模型作为前端，将信息经过模型处理后作为强化学习算法的输入^[25]；③基于模型来设计奖励函数^[26]。

在无模型方向上，针对无人机集群对抗策略优化问题，解决无人集群博弈对抗问题的思路主

要有训练架构上创新、算法上创新，以及结合其他学习方法上创新。

(1) 训练架构创新方面，通过提出基于网络化分布式部分可观测马尔可夫决策过程的空战协同策略框架^[27]，改进学习网络结构以及完善奖励机制的集中训练和分散执行的训练框架^[28]，以及将集中训练分布式执行框架推广至近端策略优化算法^[29]等。

(2) 算法创新方面，将注意力机制引入强化学习算法，提出加权平均场强化学习新范式，来应对维数灾难问题和策略泛化性问题^[30]，以及将规则耦合方法引入强化学习训练，解决集群反集群作战中速度控制和攻击角选择问题^[31]。

(3) 结合其他学习方法方面，结合课程学习方法，通过多目标任务策略渐近学习，解决集群空战中状态空间维数大的问题^[32]，以及采用分层强化学习方法，将有限的基本动作序列定义为宏动作，结合专家经验来设计宏动作，解决稀疏和延迟奖励以及状态空间和动作空间维度大问题^[33]。

表1总结了深度强化学习方法应用于无人集群博弈对抗中存在的问题，以及当前文献重点解决问题的现状。

在基于学习训练的决策技术方面，课题组针对无人集群博弈对抗无模型策略学习中存在的稀

疏奖励、直接学习效率低，以及策略网络泛化性不足问题，提出了一种“课程-迁移-强化”学习框架，如图8所示。将复杂问题进行约简，采用课程学习方法^[34]构建由易到难的训练课程，在学习过程中利用迁移学习方法^[35]实现知识和策略的迁移，并搭建自博弈环境，通过由易到难、对抗竞争、共同提升的方式进行渐近学习，提高采样和探索效率，从而获得更加鲁棒的作战策略。

目前，制约无人平台作战模式由单平台、小规模作战向集群对抗博弈转变的主要因素之一在于智能化不足、决策能力不足、过分依赖地面站远程操控、缺乏实时自主决策能力。在激烈的对抗博弈过程中，无人集群必须根据瞬息万变的作战环境实时自主决策出正确的战术动作，从而促进杀伤链的闭环。

表1 存在问题与研究现状

Table 1 Existing problems and research status

存在的问题	文献
训练效率低，如何引入专家知识来提高训练效率	[11, 25-27, 31]
作战目标属于稀疏奖励或延迟奖励问题	[26-28, 33]
状态空间和动作空间维度大，探索采样效率低	[30, 32-33]
策略网络泛化性、鲁棒性不足	[30, 32]
局部观测信息下作战决策	[27-29]

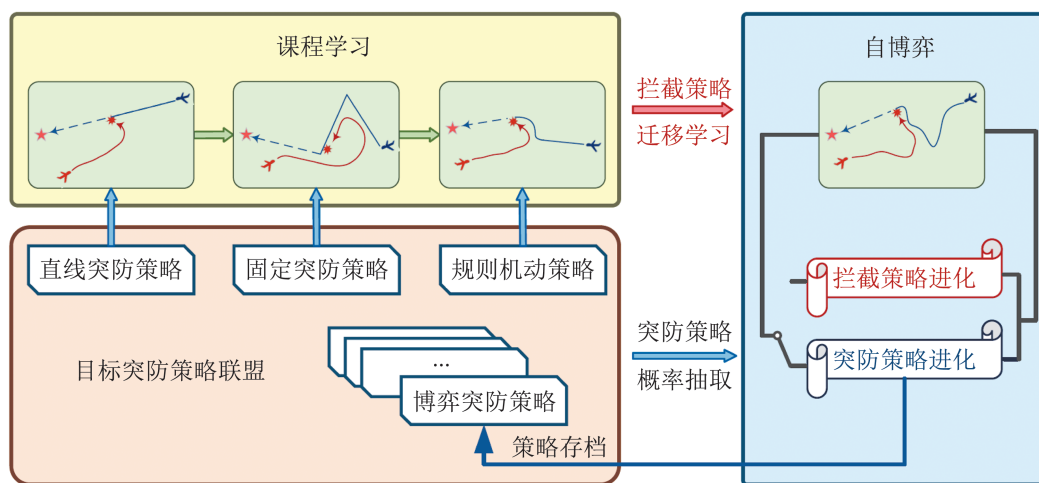


图8 “课程-迁移-强化”学习框架

Fig. 8 "Curriculum-Transfer-Reinforcement" learning framework

3 无人集群博弈对抗发展方向

无人集群博弈对抗决策能力涌现依赖于作战体系结构, 按照从体系结构推演到决策关键技术研究再到系统仿真验证的思路, 能够支撑无人集群体系能力生成、决策能力跃升, 以及行为动态运转。

3.1 体系架构模型物理时空仿真推演

从系统工程角度出发, 体系结构建模技术能够支撑无人集群开展博弈对抗任务逻辑自闭环验证, 逻辑自闭环验证重点面向无人集群博弈对抗作战时序、战场触发事件、平台运行状态和决策序列进行逻辑自检, 确保体系架构模型逻辑完整、各平台状态可达, 以及信息一致性。但体系结构模型往往和系统仿真推演模型相互独立, 导致无人集群系统在多样化场景下适用性不足, 扩展迁移应用受阻。

针对此问题, 需要建立体系架构模型与作战仿真推演模型双向联合交互仿真推演系统。基于无人集群体系架构模型的标准化工程文件, 利用仿真推演系统模型库进行作战平台模型选择与编辑。根据体系架构模型中各平台的运行状态机进行任务级/功能级行为模型开发, 在仿真推演系统中完成包括战场环境设计、作战兵力部署、作战规则设置在内的作战想定编辑。

采用联合仿真接口交互管理组件, 基于双向驱动数据的定义和事件触发发布订阅机制, 完成体系架构模型与作战仿真推演模型的接口配适和协同仿真接口程序自动生成。作战仿真推演系统将战场态势提取为状态转换触发事件, 发送至体系架构模型触发各平台逻辑状态的变迁, 体系架构模型返回逻辑决策结果, 实现双向联合交互仿真推演, 为无人集群体系架构渐进迭代演进提供细粒度设计依据和数据支撑。

3.2 无人集群自主决策与行为机理研究

在复杂作战环境下, 无人集群作战体系面临

着决策与控制的挑战。分布式无中心节点已成为无人集群的典型特征, 在分布式设定下探索无人集群博弈对抗决策和控制的增效机理, 进而提升自主决策与行为控制能力。

在自主决策方面, 针对单一节点难以解决复杂的全局决策优化问题, 将复杂问题分解至群内个体解决, 充分发挥群内个体的“智”。在通信拓扑时变、决策信息更新迟延等作战环境下, 分析无人集群“群智汇聚”机理和协商合作机制, 挖掘“群智汇聚”决策方法, 建立分布式决策交互模型。在不依赖中心节点汇聚全局信息的情况下, 实现知识和智能等决策依赖信息的汇聚, 使无人集群系统决策一致汇聚到最优, 进而提升无人集群自组织决策能力, 产生复杂环境下的决策涌现。

在行为控制方面, 针对集群内节点运动行为一致有序需求, 分析无人集群作战体系运动行为涌现机理。从行为控制需求出发, 分析内部自组织机制和外部行为调控机制, 解算无人集群在复杂环境下的可控规模和行为能力, 将分布式控制规则机理融入到无人集群系统自组织机制中, 通过内部控制信息和通信资源的高度共享和利用, 使无人集群系统行为、能力朝既定任务需求演化发展, 进而提升无人集群自组织行为能力, 产生复杂作战环境下的群体行为涌现。

从技术层面来看, 无人集群博弈对抗研究热点从对抗模型与演化机理下的有模型方法向无模型策略学习方法发展。基于学习算法强大的探索和试错能力, 能够充分挖掘无人集群博弈决策和行为机理, 超越人类经验知识。从应用层面来看, 复杂体系对抗问题属于稀疏奖励问题, 并且状态空间和动作空间维度大, 直接应用学习算法学习效率低, 甚至无法学出有效策略。从 AlphaDogfight 等成功案例来看, 以强化学习为核心, 辅以新型训练架构, 结合人类经验知识以及课程学习、迁移学习、元强化学习等方法, 能够有效提升训练效率, 扩展人工智能算法在无人集群博弈场景中的应用。

3.3 无人集群验证平台与体系能力评估

在无人集群系统验证平台方面，国内外研发了诸多无人集群系统验证平台，平台模型粒度、计算模型、仿真模型、评估模型等不尽相同，缺乏有效的平台评估手段以及规范标准，导致平台模型重复建设，资源浪费。需要尽快形成符合无人集群发展趋势的军用标准，研究标准化的评估手段方法，形成“试-验-评”三位一体布局，注重平台间可扩展性、可迁移性、可维护性，支撑无人集群博弈对抗多场景验证。

在无人集群体系能力评估方面，面向无人集群复杂系统表现出的涌现性、动态演化性等复杂性特征，考虑传统体系能力评估指标构建主观化、模型不适用、结论不可信的问题，需要依据复杂系统理论，挖掘体系能力评估关键性指标，构建网络化体系能力评估指标框架，从仿真推演数据学习提炼动态指标网，基于分布式“杀伤网”进行体系指标建模与仿真分析，以实验数据统计分析为主，消除专家经验等主观因素，以数据迭代评估无人集群多样化任务场景下的体系作战能力。

4 结论

“智能+无人+集群”作战是抢占智能化战争战略高地的主要手段，无人集群体系作战具有广泛的应用前景。目前，无人集群尚未形成体系作战能力，仍需群智群力进行关键技术攻关，开展贴合实际作战环境的系统仿真验证，推动无人集群从仿真环境走向装备实际运用，促进无人集群体系作战能力提升。

参考文献：

- [1] 梁晓龙, 胡利平, 张佳强, 等. 航空集群自主空战研究进展[J]. 科技导报, 2020, 38(15): 74-88.
Liang Xiaolong, Hu Liping, Zhang Jiaqiang, et al. Research Status and Prospect of Aircraft Swarms Autonomous Air Combat[J]. Science & Technology Review, 2020, 38(15): 74-88.
- [2] 梁晓龙, 何吕龙, 张佳强, 等. 航空集群构型控制及其演

- 化方法[J]. 中国科学(技术科学), 2019, 49(3): 277-287.
Liang Xiaolong, He Lulong, Zhang Jiaqiang, et al. Configuration Control and Evolutionary Mechanism of Aircraft Swarm[J]. Scientia Sinica(Technologica), 2019, 49(3): 277-287.
- [3] Zhu Xiaoning. Analysis of Military Application of UAV Swarm Technology[C]//2020 3rd International Conference on Unmanned Systems (ICUS). Piscataway, NJ, USA: IEEE, 2020: 1200-1204.
- [4] 刘雷, 刘大卫, 王晓光, 等. 无人机集群与反无人机集群发展现状及展望[J]. 航空学报, 2022, 43(增1): 4-20.
Liu Lei, Liu Dawei, Wang Xiaoguang, et al. Development Status and Outlook of UAV Clusters and Anti-UAV Clusters[J]. Acta Aeronautica et Astronautica Sinica, 2022, 43(S1): 4-20.
- [5] Gulden T R, Lamb J, Hagen J, et al. Modeling Rapidly Composable, Heterogeneous, and Fractionated Forces: Findings on Mosaic Warfare from an Agent-based Model [EB/OL]. [2023-03-20]. https://www.rand.org/pubs/research_reports/RR4396.html.
- [6] O'Donoghue N A, McBirney S, Persons B. Distributed Kill Chains: Drawing Insights for Mosaic Warfare from the Immune System and from the Navy[EB/OL]. [2023-03-20]. <https://doi.org/10.7249/RR4573-1>.
- [7] Pope A P, Ide J S, Daria Mićović, et al. Hierarchical Reinforcement Learning for Air-to-air Combat[C]//2021 International Conference on Unmanned Aircraft Systems (ICUAS). Piscataway, NJ, USA: IEEE, 2021: 275-284.
- [8] 周光霞, 周方. 美军人工智能空战系统阿尔法初探[C]//第六届中国指挥控制大会论文集(上册). 北京: 电子工业出版社, 2018: 66-70.
- [9] Ernest N, Carroll D, Schumacher C, et al. Genetic Fuzzy Based Artificial Intelligence for Unmanned Combat Aerial Vehicle Control in Simulated Air Combat Missions[J]. Journal of Defense Management, 2016, 6(1): 1000144.
- [10] Ernest N, Cohen K, Kivelevitch E, et al. Genetic Fuzzy Trees and Their Application Towards Autonomous Training and Control of a Squadron of Unmanned Combat Aerial Vehicles[J]. Unmanned Systems, 2015, 3(3): 185-204.
- [11] Fan D D, Theodorou E A, Reeder J. Model-based Stochastic Search for Large Scale Optimization of Multi-agent UAV Swarms[C]//2018 IEEE Symposium Series on Computational Intelligence (SSCI). Piscataway, NJ, USA: IEEE, 2018: 2216-2222.
- [12] Ma Yingying, Wang Guoqiang, Hu Xiaoxuan, et al. Cooperative Occupancy Decision Making of Multi-UAV in Beyond-visual-range Air Combat: A Game Theory

- Approach[J]. IEEE Access, 2020, 8: 11624-11634.
- [13] Li Shouyi, Chen Mou, Wang Yuhui, et al. Air Combat Decision-making of Multiple UCAVs Based on Constraint Strategy Games[J]. Defence Technology, 2022, 18(3): 368-383.
- [14] Ji Xiang, Zhang Wanpeng, Xiang Fengtao, et al. A Swarm Confrontation Method Based on Lanchester Law and Nash Equilibrium[J]. Electronics, 2022, 11(6): 896.
- [15] Liu Fei, Dong Xiwang, Yu Jianglong, et al. Distributed Nash Equilibrium Seeking of N -coalition Noncooperative Games with Application to UAV Swarms[J]. IEEE Transactions on Network Science and Engineering, 2022, 9(4): 2392-2405.
- [16] Hu Wenfeng, Zhang Qiang, Mao Yaqin. Component-Based Hierarchical State Machine-A Reusable and Flexible Game AI Technology[C]//2011 6th IEEE Joint International Information Technology and Artificial Intelligence Conference. Piscataway, NJ, USA: IEEE, 2011: 319-324.
- [17] Claudia Vannucchi, Diletta Romana Cacciagrano, Flavio Corradini, et al. A Formal Model for Event-condition-action Rules in Intelligent Environments[C]//12th International Conference on Intelligent Environments-5th International Workshop on the Reliability of Intelligent Environments (WoRIE 2016). Amsterdam, Netherlands: IOS Press, 2016: 56-65.
- [18] 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.
- [19] Gao Chunqing, Kou Yingxin, Li You, et al. Multi-objective Weapon Target Assignment Based on D-NSGA-III-A[J]. IEEE Access, 2019, 7: 50240-50254.
- [20] Xing Dongjing, Zhen Ziyang, Gong Huajun. Offense-defense Confrontation Decision Making for Dynamic UAV Swarm Versus UAV Swarm[J]. Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering, 2019, 233(15): 5689-5702.
- [21] Zhen Ziyang, Wen Liangdong, Wang Bolan, et al. Improved Contract Network Protocol Algorithm Based Cooperative Target Allocation of Heterogeneous UAV Swarm[J]. Aerospace Science and Technology, 2021, 119: 107054.
- [22] Wang Chengcai, Wu Ao, Hou Yueqi, et al. Optimal Deployment of Swarm Positions in Cooperative Interception of Multiple UAV Swarms[J]. Digital Communications and Networks, 2023, 9(2): 567-579.
- [23] 丁超, 魏瑞轩, 周凯. 基于时域映射的多无人机系统给定时间分布式最优集结[J]. 北京航空航天大学学报, 2021, 47(2): 315-322.
- Ding Chao, Wei Ruixuan, Zhou Kai. Distributed Optimal Rendezvous of Multi-UAV Systems in Prescribed Time Based on Time-domain Mapping[J]. Journal of Beijing University of Aeronautics and Astronautics, 2021, 47(2): 315-322.
- [24] Teck Hou Teng, Ah Hwee Tan, Yuan Sin Tan, et al. Self-organizing Neural Networks for Learning Air Combat Maneuvers[C]//The 2012 International Joint Conference on Neural Networks (IJCNN). Piscataway, NJ, USA: IEEE, 2012: 1-8.
- [25] Wang Luhe, Hu Jinwen, Xu Zhao, et al. Autonomous Maneuver Strategy of Swarm Air Combat Based on DDPG[J]. Autonomous Intelligent Systems, 2021, 1(1): 15.
- [26] Hu Jinwen, Wang Luhe, Hu Tianmi, et al. Autonomous Maneuver Decision Making of Dual-UAV Cooperative Air Combat Based on Deep Reinforcement Learning[J]. Electronics, 2022, 11(3): 467.
- [27] Zhang Jiandong, Yang Qiming, Shi Guoqing, et al. UAV Cooperative Air Combat Maneuver Decision Based on Multi-agent Reinforcement Learning[J]. Journal of Systems Engineering and Electronics, 2021, 32(6): 1421-1438.
- [28] Xu Dan, Chen Gang. The Research on Intelligent Cooperative Combat of UAV Cluster with Multi-agent Reinforcement Learning[J]. Aerospace Systems, 2022, 5(1): 107-121.
- [29] Xuan Shuzhe, Ke Liangjun. UAV Swarm Attack-defense Confrontation Based on Multi-agent Reinforcement Learning[C]//Advances in Guidance, Navigation and Control. Singapore: Springer Singapore, 2022: 5599-5608.
- [30] Wang Baolai, Li Shengang, Gao Xianzhong, et al. Weighted Mean Field Reinforcement Learning for Large-scale UAV Swarm Confrontation[J]. Applied Intelligence, 2023, 53(5): 5274-5289.
- [31] Xiang Lei, Xie Tao. Research on UAV Swarm Confrontation Task Based on MADDPG Algorithm[C]//2020 5th International Conference on Mechanical, Control and Computer Engineering (ICMCCE). Piscataway, NJ, USA: IEEE, 2020: 1513-1518.
- [32] Johan Källström, Fredrik Heintz. Agent Coordination in Air Combat Simulation Using Multi-agent Deep Reinforcement Learning[C]//2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC). Piscataway, NJ, USA: IEEE, 2020: 2157-2164.
- [33] Wang Baolai, Li Shengang, Gao Xianzhong, et al. UAV Swarm Confrontation Using Hierarchical Multiagent

- Reinforcement Learning[J]. International Journal of Aerospace Engineering, 2021, 2021: 3360116.
- [34] Narvekar S, Peng Bei, Leonetti M, et al. Curriculum Learning for Reinforcement Learning Domains: A Framework and Survey[J]. The Journal of Machine Learning Research, 2020, 21(1): 181.
- [35] Zhu Zhuangdi, Lin Kaixiang, Jain A K, et al. Transfer Learning in Deep Reinforcement Learning: A Survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2023, 45(11): 13344-13362.