

# 基于强化学习的反集群杀伤网动态重构策略研究

夏绍丰

江南机电设计研究所 贵州 贵阳 550009

**【摘要】**：就无人集群作战面临大规模、强协同、高动态的反制问题，又考虑到传统杀伤链结构固化、响应慢、抗毁性差的不足之处，提出了一种基于强化学习的反集群杀伤网动态重构策略。首先整理出反集群作战和杀伤网重构的主要需求，分析目前静态组网和常规动态调度方法存在的不足，然后建立反集群杀伤网马尔可夫决策模型，确定状态空间、动作空间以及奖励函数的设计思路，将多智能体强化学习算法应用于节点资源的自动分配和链路的动态重组，最后通过仿真实验来比较不同的策略在集群拦截效率、资源利用率、体系抗毁性等各方面的优劣。

**【关键词】**：强化学习；反集群作战；杀伤网；动态重构；智能决策

DOI:10.12417/3041-0630.26.06.060

## 1 引言

杀伤网是网络化作战的主要形式，它冲破了传统杀伤链的线性时序制约，把多域传感器、指控单元、火力平台、通信节点融合成互相联系的网状结构，拥有节点冗余、链路灵活、效能互补的特点，是达成高效反集群作战的理想工具。但现有的杀伤网重构大多依靠人工预先设定规则或者离线规划，不能及时适应敌方集群编队的变化、突防路线和协同方式等动态变化，决策滞后以及方案次优的问题比较明显。强化学习属于一种不需要大量先验数据、依靠试错迭代来改善决策的人工智能技术，在动态不确定的战场环境里可以自主学习最佳策略，达成态势感知、决策规划和行动执行的闭环优化，正好符合杀伤网动态重构的实时性、自主性需求。

## 2 相关研究现状与问题分析

### 2.1 国内外研究现状

目前对于杀伤网和反集群作战相关研究主要是从体系架构设计、资源优化分配、协同决策算法三个方面展开的。国外方面，美军利用智能算法控制跨域作战平台之间的动态协作与匹配，使杀伤链的构建从固定形态升级到可变形态，杀伤链的执行从单一路径转换为网状多路径，由此建立一个韧性抗毁的智能化“杀伤网”；国内学者就杀伤网建模、抗毁性改善、反集群协同拦截等展开研究，提出超图建模、遗传算法、粒子群优化等资源调度办法，在静态组网和离线规划方面取得一些成果。目前的研究大多将强化学习应用于单平台路径规划、单武器目标分配，对于复杂杀伤网下的多节点、多链路、多任务协同重构研究较少，并且普遍存在状态空间建模粗放、奖励函数单一、多智能体协同机制不健全等问题，没有很好地结合反集群作战的特点，无法实现全域资源的一体化动态调配<sup>[1]</sup>。

### 2.2 核心问题与需求分析

反集群杀伤网重构面临两个主要问题：一是战场环境高度

不确定，敌方集群可以实时改变编队密度、突防方向和协同方式，且会受到电磁干扰、节点损毁等各种情况的影响，传统的固定架构不能适应态势的迅速变化；二是作战资源跨域异构性较强，传感器、火力平台、通信设备功能属性、作战半径、响应速度等存在很大差异，资源调度要兼顾毁伤效能、资源消耗、响应速度等诸多约束。

## 3 基于强化学习的反集群杀伤网建模

### 3.1 杀伤网体系架构与作战流程

本文所建立的反集群杀伤网是分布式去中心化的，它由感知层、决策层、执行层、通信层四个部分组成。感知层主要由各种探测设备组成，对战场态势进行实时获取，并对敌方集群参数、节点状态进行识别；决策层为强化学习智能决策核心，接收感知数据，对态势进行评估，重新构建决策；执行层包含弹道光波扰等软硬杀伤平台，执行拦截、干扰任务；通信层建立通信链路，实现各层节点之间信息交互。

完整的作战过程分为三部分，第一部分是感知层感知全域态势，把敌方集群规模、编队形态、威胁等级等信息传送给决策层；第二部分是决策层使用强化学习算法生成最优的重构方案，进行链路切换、目标分配；第三部分就是按照决策层给出的方案执行软硬协同杀伤，瓦解敌方集群，评估作战效果，形成一个决策闭环迭代优化机制。

### 3.2 马尔可夫决策过程建模

将反集群杀伤网动态重构的过程转化成马尔可夫决策过程，利用状态、动作、奖励的闭环迭代实现策略优化的目的，核心要素定义如下。

(1) 状态空间：由敌方目标、我方资源、战场态势等因素组成的三维向量。第一维是敌方集群状态，包括集群规模、编队密度、移动速度、威胁等级；第二维是杀伤网状态，即可

用节点数、各个节点的作战能力、链路连通性、资源剩余量；第三维是战场环境，即电磁干扰强度、地形约束、时间窗口。

(2) 动作空间：对杀伤网重构的可执行动作进行定义，分为节点调度、链路重组、目标分配三种主要的动作。节点调度包含激活闲置节点，关停故障节点，调节节点的工作方式，链路重组包含新建作战链路，更换备用链路，中断敌方协同链路，目标分配包含火力分配，侦察任务分配，干扰任务分配。

(3) 奖励函数：使用多维复合奖励函数，既考虑即时作战效能也考虑长期体系韧性，防止因为单一维度的考虑造成战略失衡。奖励值由毁伤效能奖励、资源消耗惩罚、抗毁性奖励、响应速度奖励四部分加权组成，毁伤效能奖励和敌方集群毁伤率、协同链路切断率正相关，资源消耗惩罚和火力、能源消耗负相关，抗毁性奖励和体系连通度正相关，响应速度奖励和决策耗时负相关。

### 3.3 多智能体强化学习算法设计

考虑到反集群杀伤网节点数量多、协同关系复杂，单智能体强化学习不能实现全域协同优化，本文使用多智能体深度确定性策略梯度算法，把各个作战节点当作独立智能体，设置全局协调智能体，实现局部自主决策和全局协同优化的结合。全局协调智能体对全域态势进行统筹，确定整体重构的目标，协调各个节点智能体的动作冲突；各个节点智能体根据自身的局部态势以及全局指令，自主优化自己的动作策略，采用集中训练、分散执行的方式，减少决策计算的复杂程度，提高实时响应的能力。算法用经验回放池保存历史决策数据，定时更新网络参数，慢慢逼近最优重构策略，达成战场态势改变和杀伤网重塑的即时对接<sup>[2]</sup>。

## 4 反集群杀伤网动态重构策略实现

### 4.1 态势感知与威胁评估模块

态势感知模块依靠多源传感器数据融合技术，对敌方集群和我方杀伤网的全部数据实施实时采集，用目标识别算法精确找到敌方集群的种类、编队形式、重要协同节点，威胁评价模块把集群突防几率、打击价值、距离参数融合起来创建威胁评价模型，对各个集群目标的威胁等级加以量化，给强化学习决策赋予优先级参照，先将优质资源投向高威胁目标。

### 4.2 动态重构核心策略

本文设计的动态重构策略有常态自适应重构、受损应急重构、任务优先级重构三种模式，适应不同的战场环境。

(1) 常态自适应重构，对于敌方集群常规机动、编队微调等平缓的态势变化，依靠强化学习智能体对实时状态的反应，对节点的工作模式以及链路参数进行微小的调整，从而维持杀伤网的最佳作战状态，减少资源消耗，保证持续稳定的拦

截。

(2) 受损应急重构，对由于我方节点损坏、通信中断、电磁干扰等突发状况而触发的应急重构进行快速响应，智能体先切换备用节点和链路，隔离故障单元，然后改变目标分配方案来弥补局部作战缺口，保证体系作战效能不受影响。

(3) 任务优先级安排，根据威胁评价结果把任务进行分级，重新安排杀伤网链路，集中优势兵力对付高威胁集群，调配闲置资源对付低威胁目标，达成全域协同拦截<sup>[3]</sup>。

## 5 仿真实验与结果分析

### 5.1 实验环境与参数设置

为检验提出方法的有效性，建立反集群作战仿真环境，设置我方杀伤网有15个异构作战节点，包括传感器、火力平台、电子干扰设备三种类型；模拟敌方3组不同的无人集群，分别为小规模精锐集群、中规模协同集群、大规模饱和攻击集群，集群具有自主编队、动态规避、电磁干扰的能力。设置对照组，分别使用传统静态杀伤网策略、常规动态调度策略，和本文提出的强化学习动态重构策略进行比较，实验指标为集群毁伤率<sup>[4]</sup>、资源利用率、决策耗时、体系抗毁性<sup>[5]</sup>这四个主要参数，每组实验做10次，取平均值作为结果。

### 5.2 实验结果与对比分析

从实验数据可知，本文提出的方法各项指标都比传统的要好。集群毁伤率由原来的82.1%提高到现在的89.5%，比传统的静态策略高27.2个百分点，比一般的动态策略高13.7个百分点，主要是因为强化学习能够准确找到重要的协同节点和链路，从而提高了集群的协同效果；资源利用率可以达到82.1%，避免了资源的浪费以及重复使用，使得各个异构节点可以高效地合作；平均决策耗时为180ms，满足战场实时决策的需求，克服了传统方法决策滞后的问题；体系抗毁性评分大幅提高，在局部节点遭到损坏之后，短时间内就能建立起新的链路并维持住稳定的作战效能。

表1 实验结果与对比分析

作战策略	集群毁伤率(%)	资源利用率(%)	平均决策耗时(ms)	体系抗毁性评分
传统静态杀伤网策略	62.3	51.7	850	65.2
常规动态调度策略	75.8	68.4	420	78.6
本文强化学习动态重构策略	89.5	82.1	180	91.3

根据不同的拦截实验结果可知,本文策略对于中规模协同集群的瓦解效果最好,毁伤率达到了92.7%,可以很快切断集群协同链路;面对大规模饱和攻击集群,可以采用多节点协同调度、分层拦截的方法来降低集群突防的概率,具有较好的场景适应性。

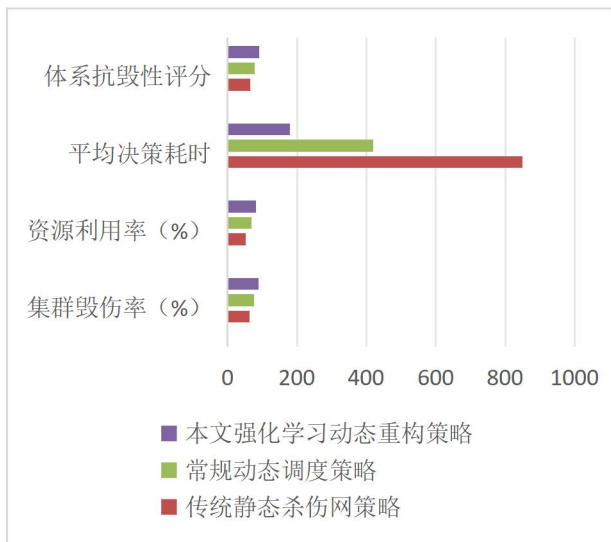


图1 实验数据对比

### 5.3 算法性能分析

经过2000次迭代后,本文所用到的多智能体强化学习算法就可以收敛了,比传统的深度强化学习算法收敛速度更快,并且决策更加稳定,没有明显的波动。算法可以避开局部最优解,从而达到全局效能的最大化,具有较好的泛化性,在未知

的战场环境中也可以得到较好的决策效果。

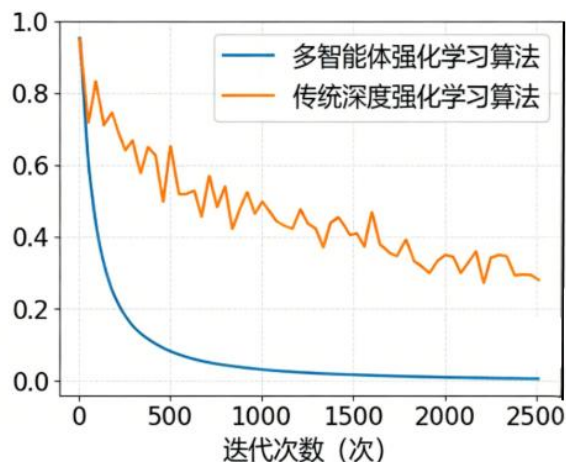


图2 算法收敛图

## 6 结论

本文针对反集群作战杀伤网动态重构的主要问题,把强化学习技术同网络化作战体系深度融合起来,创建多智能体协同推动的动态重构策略,用马尔可夫决策建模,复合奖励函数设计,多模式重构机制改良等手段,较好地解决传统杀伤网结构固定,反应迟缓,抗毁性差的问题,明显加强了反集群作战的毁伤效能,资源利用效率和体系韧性。仿真实验结果证明所提策略是可行的、有效的,可以给智能化反集群作战体系的研发和应用提供理论依据和技术支持。

### 参考文献:

[1] 万莉,李洋,包云霞,等.分布式防御杀伤网密度的定义与评估方法应用[J/OL].哈尔滨工业大学学报,1-11[2026-03-30].  
 [2] 王岳,陶西贵,翟长海,等.蓄意攻击下关键基础设施韧性评估研究进展[J/OL].哈尔滨工业大学学报,1-15[2026-03-30].  
 [3] 张曙光,杜晓明,陈朝阳,等.杀伤网建模方法及其发展趋势[J/OL].计算机工程与应用,1-19[2026-03-30].  
 [4] 高志刚,刘艳彬,陈长远,等.基于排队论的反无人机集群武器部署优化方法[J].装备环境工程,2022,19(06):68-75.  
 [5] 李腾达,王刚,郭相科,等.基于复杂网络的防空反导体系抗毁性评估[J].军事运筹与评估,2023,38(05):34-39.