

# 面向饱和攻击防空预警的组网雷达异质任务协同调度架构

叶炬航<sup>①</sup> 方愚渊<sup>②</sup> 魏少鹏<sup>③</sup> 段佳<sup>①</sup> 张磊<sup>\*①</sup>

<sup>①</sup>(中山大学电子与通信工程学院 深圳 518107)

<sup>②</sup>(西南交通大学信息科学与技术学院 成都 611756)

<sup>③</sup>(中国石油大学(华东)海洋与空间信息学院 青岛 266580)

**摘要:** 面对无人机集群和智能巡飞弹等大规模、高速度饱和攻击威胁带来的海量、突发且异质预警任务的严峻挑战,如何依托组网雷达设计可拓展的协同调度架构,释放防空预警体系抗饱和和攻击潜力亟待解决。针对现有架构动态协同僵化和异质任务调度能力不足等局限,该文深入分析饱和和攻击场景多雷达多目标多任务调度难题,将非凸的协同调度问题解耦为多维资源动态协同和自适应异质任务调度多阶段决策过程,通过针对性引入派单机制和分层强化学习,以可拓展性为目标提出了一种决策层次上融合了集中式动态目标分配和分布式异质任务调度的混合架构,并具体设计了执行-再派单认知闭环、目标派单算法和分层式指挥调度方法。仿真表明所提架构面对大规模高速度饱和攻击时,能实时动态协同组网雷达以调度搜索、确认和跟踪三类异质预警任务,并在适用处理规模、环境自适应能力、长期调度能力、可拓展性和可解释性等方面具有改进。

**关键词:** 饱和攻击; 组网雷达; 目标分配; 任务调度; 派单机制; 分层强化学习

中图分类号: TN95

文献标识码: A

文章编号: 1009-5896(2026)00-0001-16

DOI: 10.11999/JEIT260373

CSTR: 32379.14.JEIT260373

## 0 引言

组网雷达预警任务协同调度<sup>[1,2]</sup>架构能通过协同空时能多维资源有效释放防空预警体系效能<sup>[3]</sup>。然而,目前无人机集群和智能巡飞弹等大规模、高速度饱和攻击威胁<sup>[4,5]</sup>不断涌现,预警任务呈现海量、突发且异质的特点。一方面,现有架构动态协同僵化,空间上防区划分<sup>[3,6]</sup>难以深度协同多维资源,适用处理规模有限,此外完全分布式或完全集中式结构受限于单一架构局限,而混合结构<sup>[7,8]</sup>在管理模式可拓展性等方面仍存在设计难题。另一方面,现有架构异质任务调度能力不足,时间上贪婪策略<sup>[9,10]</sup>缺乏长期调度能力,环境自适应能力不足,此外同质预警任务假设<sup>[11-13]</sup>与实际中不同资源需求、依赖关系的异质任务特性失配,而异质任务假设<sup>[14]</sup>在架构可拓展性等方面仍存在研究空间。面向饱和攻击场景异质任务调度需求,依托组网雷达体制设计可拓展的协同调度架构,对保障我国空域安全、释放预警体系抗饱和和攻击潜力具有重要价值。

现有研究可从信息融合结构、管理模式和建模方案上予以分类。首先,区别于数据层次上的信息

融合,协同调度架构属于决策层次,旨在根据融合信息实现最优决策,从信息融合结构上可分为三类<sup>[7,8,15]</sup>。(1)完全集中式结构,单一中心节点同时负责原始数据处理和集中决策,信息损失最小且决策唯一,但在饱和攻击场景下受限于单点能力边界,设计复杂且构成要害;(2)完全分布式结构,各节点独立形成局部信息再送入融合中心,通过相互通信形成决策共识,设计难度较低且生存能力强。但缺乏决策中心导致各节点决策可能相互矛盾且无法承担决策责任;(3)混合结构,决策层次上指信息融合方式和决策方式采用不同结构以期突破单一架构局限,但如何设计可扩展的管理模式<sup>[16]</sup>仍是难题。具体而言,组网雷达预警体系可区分组网雷达指挥部和雷达节点两级,从管理模式上可分为两类<sup>[3]</sup>。(1)直通控制模式,指挥部直通各雷达节点实现参数级控制,两级之间形成参数控制强耦合,物理含义清晰、实时性强,但功能相对单一且设计复杂,常见于传统闭环资源管理<sup>[2,12,17,18]</sup>;(2)分级控制模式,指挥部仅生成规划上级任务并下发各雷达节点自主理解执行,两级之间形成任务控制弱耦合,针对复杂任务场景可扩展性强,但如何量化设计兼容人工决策先验信息和多维资源统一管理的可拓展认知闭环仍有研究空间。进一步地,针对需要统一指挥并具体落实指挥责任的饱和攻击防空预警场景,协同调度架构从建模方案上可分为两类。(1)雷达节点主导方案,主要包括非合作博弈<sup>[13]</sup>通过市场机制竞争获取资源和合作博弈<sup>[11,17]</sup>依照Shapley值公

收稿日期: 2026-xx-xx; 改回日期: 2026-06-24; 网络出版: 2026-07-04

\*通信作者: 张磊 zhanglei57@mail.sysu.edu.cn

基金项目: 国家自然科学基金(62301612), 山东省科技重大工程青年人才支持(SDAST2025QTA099), 山东省自然科学基金(ZR2024MF096), 山东省自然科学基金(ZR2023QF004), 广东省自然科学基金(2025A1515010242)

平分配资源等。雷达节点相互通信,通过博弈实现任务调度和资源协同,适用于去中心的完全分布式结构或弱中心的混合结构。然而,针对当前任务场景,各雷达节点一方面无法落实指挥责任,不支持统一指挥;另一方面大规模问题难以保证实时性,且较难求解非线性多约束的多对多分配问题;(2)组网雷达指挥部主导方案,如指挥部已知全局信息建立混合整数规划问题,并利用匈牙利算法<sup>[9]</sup>、拍卖算法<sup>[12,19]</sup>或元启发式算法<sup>[20]</sup>集中求解实现统一指挥或指挥部集中训练、雷达节点间无通信的强化学习如Q学习(Q-Learning, QL)<sup>[9]</sup>。一方面,混合整数规划立足全局信息与集中决策,信息损失小且保证决策唯一,但现有统一指挥缺乏明确的作战分工,多维资源统一管理困难,且NP-Hard问题难以兼顾实时性和最优性,限制了适用处理规模和架构可拓展性;另一方面,多智能体强化学习在场景难以显式建模时端到端表现良好,但单层结构难以保证策略学习一致性,环境奖励稀疏时学习效率低下且可拓展性一般,限制了长期调度能力和环境自适应能力,此外黑盒可解释性差。

综上所述,针对饱和攻击防空预警场景,如何设计一种结构上突破单一架构局限、管理模式上构建支持可扩展设计空间的认知闭环并在建模方案上建立作战分工明确的统一指挥、适用大规模非线性多约束多对多目标分配和自适应长期调度最优的可拓展协同调度架构亟待解决。基于此,本文深入分析饱和攻击场景动态不确定环境多雷达多目标多任务协同调度难题,将非凸的协同调度问题解耦为协同和调度多阶段决策过程,并在组网雷达指挥部协同层和雷达节点调度层分别予以解决,通过引入派单机制<sup>[21]</sup>和分层强化学习<sup>[22]</sup>,以可拓展性为目标提出了一种决策层次上融合了集中式动态目标分配和分布式异质任务调度的混合架构,并具体设计了执行-再派单认知闭环、目标派单算法和分层式指挥调度方法。所提架构优势包括:(1)适用饱和攻击场景,支持大规模、高速度集群目标实时预警,可拓展性较强;(2)支持海量、突发异质预警任务自适应长期调度,可解释性较强。

## 1 组网雷达异质任务协同调度架构设计

### 1.1 多雷达多目标多任务协同调度问题分析

针对饱和攻击场景中 $J$ 个雷达对 $I$ 个目标调度 $M$ 种异质任务的 $O(M^{IJ})$ 协同调度问题,考虑(1)目标状态动态变化,组网雷达信息获取不确定,构成了动态不确定环境;(2)协同调度长期性、空时能资源多维性和任务关系依赖性等约束条件,该问题

具有大规模、长期性、非线性和多约束的特点,构成了多雷达多目标多任务协同调度难题。

为处理该难题,本文(1)从目标威胁度和雷达工作状态联合视角出发,构建动态不确定环境模型,利用跟踪精度和探测概率描述饱和和攻击场景,并构建支持可拓展性设计空间的执行-再派单认知闭环;(2)从解耦非凸的协同调度问题为先协同、再调度的多阶段决策过程,将原问题转化为协同层面动态协同僵化和调度层面异质任务调度能力不足两个子问题,利用可拓展设计空间针对性引入派单机制和分层强化学习分别予以处理。本文将多雷达多目标多任务问题转化为通过认知闭环交互的协同和调度子问题,并对设计以可拓展性为目标的组网雷达异质任务协同调度架构。

### 1.2 动态不确定环境模型

本节定义动态不确定环境模型。一方面,目标威胁度视角仅关注目标状态如位置、速度和角速度等,不考虑压制干扰或欺骗干扰。本文采用贝叶斯卡拉美罗下限(Bayesian Cramer-Rao Lower Bound, BCRLB)描述跟踪不确定性。定义 $k$ 时刻雷达 $j$ 观测目标 $i$ 时的跟踪精度 $\eta_k^{i,j}$ ,单位为米,跟踪精度越小,跟踪不确定性越低,目标威胁度越低。其中, $\mathbf{J}_k$ 为位置估计的费舍信息矩阵。

$$\eta_k^{i,j} = \sqrt{\text{tr}[\mathbf{J}_k]} \quad (1)$$

值得指出的是,所提架构不依赖雷达节点具体的跟踪算法,数据层次上采用分布式信息融合,可根据实际情况选择统计方法或模糊方法在融合中心实现数据关联、形成系统航迹并计算BCRLB。

另一方面,雷达工作状态视角仅关注目标信号和其他雷达因旁瓣泄露造成的友方互扰。本文采用信干噪比估计,利用检测门限和探测概率<sup>[23]</sup>描述检测随机性。定义 $k$ 时刻雷达 $j$ 观测目标 $i$ 时的信干噪比 $\zeta_k^{i,j}$ ,信干噪比越大,探测概率越高,雷达工作状态越好。其中, $k-1$ 时刻 $P_{k-1}^{i,j}$ 为主瓣回波功率, $P_{k-1}^{j,j'}$ 为旁瓣干扰功率, $P_{k-1}^n$ 为噪声功率。

$$\zeta_k^{i,j} = \frac{P_{k-1}^{i,j}}{\sum_{\substack{j,j' \in J \\ j \neq j'}} P_{k-1}^{j,j'} + P_{k-1}^n} \quad (2)$$

根据式(2),本文主瓣照射与协同调度的时序关系为上一轮照射结果作为本轮决策依据,需保证决策时间步长足够小以满足决策依据时效性。

### 1.3 协同调度架构设计

本文将预警体系划分为两级(1)负责全局信息汇总和作战指派统一指挥的组网雷达指挥部;

(2)含外围装备的雷达节点前线，形成指挥部先协同、前线再调度的作战分工，并在分级控制模式的基础上构建支持可扩展设计空间的执行-再派单认知闭环。通过在适用处理规模、环境自适应能力、长期调度能力、可拓展性和可解释性等方面针对性引入适配设计，从而释放防空预警体系抗饱和攻击潜力。如图1所示，所提协同调度架构可划分为协同层和调度层，二者通过执行-再派单认知闭环交互，旨在解决大规模、高速度集群目标饱和攻击场景组网雷达协同调度搜索、确认和跟踪三类异质任务难题。

(1)执行-再派单认知闭环。区别于传统资源管理闭环，本文所提架构聚焦于构建支持可扩展设计空间的认知闭环，并采用含自适应权重的目标派单模型这一形式构建执行-再派单认知闭环<sup>[24]</sup>。具体而言，首先定义目标派单要素；接着，采用自适应权重形式量化设计兼容人工决策先验信息和多维资源统一管理的可拓展认知闭环；紧接着，针对空时能三个正交的资源维度，从目标威胁度和雷达工作状态联合视角出发，利用可拓展设计空间从单站波束指向稳定性、资源利用率和工作状态，以及多站视角协同、负载均衡和高威胁目标响应六个方面量化设计综合作战效能函数；最后，利用信噪比估计建模探测概率模型以期加强认知真实程度。

(2)协同层。注意到资源消耗与目标处理规模直接相关，本文将多维资源深度协同视为动态目标分配，先将原问题化归为 $J$ 个单雷达多目标多任务调度问题。此时动态协同僵化具体表现为静态协同易响应滞后且多站资源协同程度浅，本文引入派单机理，以动态目标分配为核心显式建模协同为 $O(2^J)$ 广义二分图匹配问题。派单机理属于组网雷

达指挥部主导方案，一方面该机制适配数据层次上分布式信息融合以获取全局信息，并支持决策层次上指挥部统一指挥实现全局最优目标分配，从而动态深度协同空间、时间和功率多维资源；另一方面，该机制支持将饱和攻击过程参数化封装为海量、突发的目标订单，并通过事件驱动实现即时响应。具体而言，针对大规模非线性多约束广义二分图匹配难题，采用蚁群算法(Ant Colony Optimization, ACO)<sup>[20]</sup>结合可行性规则约束处理构造性求解，利用次优解逼近最优解。

(3)调度层。继续考虑单雷达多目标多任务调度问题，注意到当前任务场景异质预警任务具有海量、突发且任务依赖关系交织的特点，本文隐式建模 $J$ 个 $O(M^{N_j})$ 车间调度问题。此时异质任务调度能力不足具体表现为仅支持同质任务调度和长期调度能力不足，本文引入元控制器-控制器分层设计<sup>[22]</sup>，考虑策略稳定收敛采用QL框架设计一种分层Q学习(Hierarchical Q-Learning, HQL)以期验证分层设计有效性。一方面，分层设计将长期调度过程转化为宏观-微观混合时间步决策，有效改善了决策时间跨度长和环境奖励稀疏时智能体学习效率，从而改善长期策略质量；另一方面，该分层设计无关具体实现方法，可拓展性强，且针对当前任务场景可映射清晰的物理含义，相比于其他分层强化学习方法可解释性更强。值得指出的是，HQL保证稳定收敛，但离线训练且难以处理高维状态空间，后续可推广分层深度Q网络，采用神经网络作为值函数拟合形式解决维度灾难。具体而言，本文基于任务比例控制这一子目标，构建上层作战意图动态指挥控制下层具体波束驻留的分层式指挥调度方法。

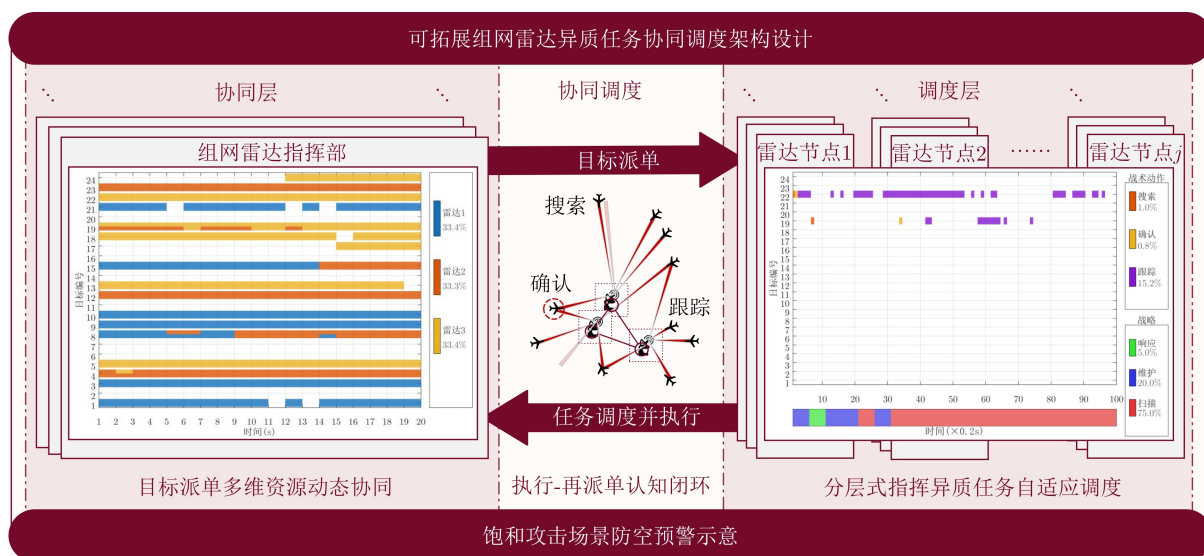


图1 可拓展组网雷达异质任务协同调度架构设计示意图

所提架构存在数据层次上的融合中心，也是决策层次上的决策中心。故融合中心损坏同时意味着决策中心的损坏，此时动态协同能力失效。但考虑到协同和调度解耦，失去中心后各节点某种意义上仍能根据局部信息自主调度，异质任务调度能力仍有效。

## 2 执行-再派单认知闭环

本文具体设计执行-再派单认知闭环，旨在利用可扩展设计空间构建目标派单、任务调度并执行后，通过持续反馈任务执行效果从而动态调整目标派单的闭环过程。首先定义目标派单要素如下：(1)订单为目标，目标被组网雷达检测与否构成事件，综合作战效能作为定价；(2)派单方为指挥部，统一指挥体现为统一派单，派单与否构成二元变量，允许目标被复数派单；(3)接单方为雷达节点，允许雷达复数接单。

基于此，如式(3)所示定义含自适应权重的目标派单模型。其中， $J$ 为目标函数， $U_k^{i,j}$ 为目标 $i$ 派单给雷达 $j$ 的综合作战效能， $x_k^{i,j}$ 为派单二元变量，派单为1否则为0。考虑实际中雷达威力、检测门限和无漏警要求，定义硬约束包括(1)接单状态仅为接受或拒绝；(2)目标必须处于雷达最大威力范围或视距内；(3)回波的信干噪比必须过最小可检测门限；(4)每个目标至少被派单给一个雷达。

$$\begin{aligned} \max_{x_k^{i,j}} J &= \sum_{j \in J} \sum_{i \in I} U_k^{i,j} \cdot x_k^{i,j} \\ \text{s.t. } \forall j \in K_J, \forall i \in K_I, x_k^{i,j} &\in \{0, 1\} \\ \forall j \in K_J, \forall i \in K_I, \left| p_k^j - p_k^i \right| &> R_{j,\max} \Rightarrow x_k^{i,j} = 0 \\ \forall j \in K_J, \forall i \in K_I, \zeta_k^{i,j} < \zeta_{\text{th}} &\Rightarrow x_k^{i,j} = 0 \\ \forall i \in K_I, \sum_{j \in K_J} x_k^{i,j} &\geq 1 \end{aligned} \quad (3)$$

其中，如式(4)所示综合作战效能 $U_k^{i,j}$ 包括(1)描述目标威胁度的 $k$ 时刻下 $t$ 个目标作战效能函数 $U_{k,t}^i(j)$ 及其权重 $w_t$ ；(2)描述雷达工作状态的 $r$ 个雷达作战效能函数 $U_{k,r}^j(i)$ 及其权重 $w_r$ ；(3)调整二者侧重的权重因子 $\alpha_U$ ；(4)构成执行-再派单认知闭环的目标综合作战效能自适应权重 $\beta_{T,k}^i$ 和雷达综合作战效能自适应 $\beta_{R,k}^j$ 权重；(5)表征动态不确定环境的探测概率 $P_d(\zeta_k^{i,j})$ 。

$$\begin{aligned} U_k^{i,j} &= P_d(\zeta_k^{i,j}) \left[ \alpha_U \beta_T \left[ \sum_t w_t U_{k,t}^i(j) \right] \right. \\ &\quad \left. + (1 - \alpha_U) \beta_R \left[ \sum_r w_r U_{k,r}^j(i) \right] \right] \end{aligned} \quad (4)$$

第一，定义自适应权重。注意到单轮目标派

单、任务调度并执行的效果直接反映目标威胁度与雷达工作状态的变化，据此定义二元变量 $y_k^i$ 为 $k$ 时刻目标 $i$ 的任务执行效果，仅信干噪比过门限且跟踪精度小于阈值时取1，否则取0。区别于传统资源管理闭环直接反馈跟踪精度等性能指标，如式(5)所示本文以任务执行效果为性能指标，采用自适应权重量化设计兼容人工决策先验信息和多维资源统一管理的可拓展认知闭

$$\begin{cases} y_k^i = (\zeta_k^{i,j} > \zeta_{\text{th}}) \cap (\eta_k^{i,j} < \eta_{\text{th}}) \\ \beta_{R,k+1}^j = \text{clip} \left[ \beta_{R,k}^j - \gamma_R (7\beta_{R,\text{th}} - \frac{1}{W} \sum_{k'=k-W+1}^k \sum_i y_{k'}^i x_{k'}^{i,j}) \right. \\ \quad \left. \cdot \frac{\sum_i y_{k'}^i x_{k'}^{i,j}}{\sum_i x_{k'}^{i,j}}, \beta_{\min}, \beta_{\max} \right] \\ \beta_{T,k+1}^i = \text{clip} \left[ \beta_{T,k}^i + \gamma_T \left( \beta_{T,\text{th}} - \frac{1}{W} \sum_{k'=k-W+1}^k \max_j y_{k'}^i \right) \right. \\ \quad \left. \cdot \frac{\sum_j y_{k'}^i x_{k'}^{i,j}}{\sum_j x_{k'}^{i,j}}, \beta_{\min}, \beta_{\max} \right] \end{cases} \quad (5)$$

其中， $\gamma_R$ 与 $\gamma_T$ 分别为二者取值范围 $[0, 1]$ 的更新系数，更新系数越大，自适应权重越敏感； $\beta_{R,\text{th}}$ 与 $\beta_{T,\text{th}}$ 分别为二者取值范围 $[0, 1]$ 的基准系数，由先验信息和人工决策设定，用于控制自适应权重动态平衡点和反馈调整方向。一方面，遍历每个雷达计算接单的历史任务执行效果，其值越高表示雷达工作状态越好，高于基准系数时权重调整方向为正，相应增大雷达自适应权重，从而增大雷达视角下全部目标的订单定价以倾向于更多接单，反之亦然；另一方面，遍历每个目标计算其在全局雷达范围内的历史任务执行效果，其值越低表示目标威胁度越大，低于基准系数时权重调整方向为正，相应增大目标自适应权重，从而增大目标视角下全部雷达的潜在定价以增加被派单倾向，反之亦然。

第二，定义归一化作战效能函数。为实现多维资源统一管理，聚焦作战效能函数及其权重设计具备可拓展性这一思路，本文从空间、时间和功率三个正交的资源维度出发，对应考虑单站波束指向稳定性、资源利用率和工作状态和多站视角协同、负载均衡和高威胁目标响应，延续联合视角，建立如表1所示的归一化作战效能函数定义。其中， $x_k^{i,j}$ 为 $k$ 时刻目标 $i$ 和雷达 $j$ 的派单二元变量，1为派单、0为未派单， $N_J$ 为雷达负载上限， $p_k^i$ 和 $\dot{p}_k^i$ 分别是目标在 $k$ 时刻的位置和速度， $P_k^i$ 为如式(14)定义的目

标优先级。值得指出的是，作战效能函数的设计应保证在完备性的前提下尽可能相互独立。

具体而言，(1)波束指向调整代价与雷达视线与目标速度方向的夹角有关，夹角越大，雷达接单该目标后可能产生的波束指向调整越大，该项代价越大，从而要求派单不造成雷达节点波束指向大幅度调整；(2)资源充分利用收益与全体雷达节点当前负载水平有关，通过找到资源最空闲的节点并倾向增加派单，从而鼓励每个节点的资源都被充分利用；(3)跟踪精度收益定义为BCRLB的单调归一化形式，在保留信息论工具的同时，体现了目标派单应综合考虑观测几何和处理负载等全局信息下，通过统一指挥鼓励目标被派单给能提供有效跟踪的节点的决策逻辑；(4)多站协同收益与多部雷达关于同一目标的雷达视线夹角有关，相互间雷达视线夹角越大，观测几何差异越大，协同获取的信息量越

丰富，该项收益越大，从而鼓励不同观测视角的雷达节点相互协同；(5)超负荷代价与全体雷达节点当前负载水平有关，通过惩罚超负荷现象要求雷达节点接单尽可能不超过其处理能力；(6)高威胁目标战术倾斜收益与目标优先级有关，目标优先级越高，该项收益越大，越倾向于被派单，从而体现战术倾斜。

第三，考虑雷达探测概率特性直接影响目标派单过程，建立以信干噪比为核心的探测概率 $P_d(\zeta_k^{i,j})$ ，模拟杂波和信干噪比闪烁等非线性影响。其中， $P_f$ 为虚警率， $n_{\text{pulse}}$ 为脉冲压缩数。

$$P_d(\zeta_k^{i,j}) = \left(1 + \frac{(P_f)^{-\frac{1}{n_{\text{pulse}}}} - 1}{1 + \zeta_k^{i,j}}\right)^{-n_{\text{pulse}}} \quad (6)$$

### 3 协同层目标派单算法

如表2所示，所提目标派单算法(Target Dispatching Algorithm, TDA)在最大最小蚂蚁系统

表 1 归一化作战效能函数定义

序号	命名	定义
1	波束指向调整代价	$U_{k,1}^j(j) = \frac{(\mathbf{p}_k^j - \mathbf{p}_k^i) \cdot \mathbf{p}_k^i}{\ \mathbf{p}_k^j - \mathbf{p}_k^i\  \cdot \ \mathbf{p}_k^i\ } - 1$
2	资源充分利用收益	$U_{k,2}^j(i) = \max \left( \min \left( \frac{\sum_i x_k^{i,j}}{N_j} \right) \right)$
3	跟踪精度收益	$U_{k,3}^j(i) = 1 - \frac{\eta_k^{i,j}}{\max_i \eta_k^{i,j}}$
4	多站协同受益	$U_{k,1}^i(j) = \frac{1}{\pi} \max_{j \neq j'} \left( \arccos \left( \frac{\mathbf{p}_k^j - \mathbf{p}_k^i}{\ \mathbf{p}_k^j - \mathbf{p}_k^i\ } \cdot \frac{\mathbf{p}_k^{j'} - \mathbf{p}_k^i}{\ \mathbf{p}_k^{j'} - \mathbf{p}_k^i\ } \right) \right)$
5	超负荷代价	$U_{k,4}^j(j) = -\max \left( 0, \frac{\sum_i x_k^{i,j} - N_j}{N_j} \right)$
6	高威胁目标战术倾斜收益	$U_{k,2}^i(j) = \frac{P_k^i}{\max_i P_k^i}$

表 2 目标派单算法表

输入：综合作战效能矩阵 $\mathbf{U}_k$ ，信息素权重 $\alpha$ 、启发式信息权重 $\beta$ 、挥发率 $\rho$ 、信息素常数 $Q$ 、最大信息素 $\tau_{\min}$ 和最小信息素 $\tau_{\max}$ 、雷达负载上限 $N_j$ 和目标覆盖上限 $N_I$
输出：全局最优目标派单方案 $\hat{X}_k^n$
(1)初始化最大迭代次数 $K$ 、种群规模 $N$ 、信息素矩阵 $\tau_0^{i,j} = \tau_{\max}$ 和综合作战效能 $J(X_0^n) = 0$
(2)目标覆盖层求解，随机顺序遍历目标，筛选未满载、处于威力范围内且过门限的雷达并计算路径转移概率 $P_{n,k}^{i,j}$ ，确保满足零漏警硬约束
(3)多站协同层求解，再次随机顺序遍历目标，剔除已派单雷达，筛选剩余资源可用雷达并派单，计算观测几何增益
(4)针对每个蚂蚁的解 $X_k^n$ 评估综合作战效能 $J(X_k^n)$ ，更新最优解 $\hat{X}_k^n$ 并找到该蚂蚁
(5)所有信息素进行挥发
(6)仅允许找到最优解 $\hat{X}_k^n$ 的蚂蚁更新信息素
(7)所有信息素限幅 $[\tau_{\min}, \tau_{\max}]$
(8)若达到最大迭代次数则输出 $\hat{X}_k^n$ ，否则转至(2)

(Max-Min Ant System, MMAS)的基础上通过构造可行性规则约束处理方法,从而在动态不确定环境中实时求解最优目标派单方案,实现了空时能资源深度协同。

考虑到所提动态、多对多目标派单模型大规模、非线性和多约束特点,传统方法如匈牙利算法和拍卖算法不适用。出于实时性和目标派单质量考虑,本文选取ACO作为求解算法,并选择MMAS作为工程实现形式。但针对多约束问题,MMAS一般需引入约束处理方法以免陷入无效搜索。然而,传统罚函数法针对所提模型难以设计惩罚项并调参,欠惩罚时难以避免MMAS浪费时间搜索不可行解,而过惩罚时则导致早熟收敛于次优解。

注意到所提模型中每个目标至少被派单一个雷达的硬约束和鼓励目标复数派单的软约束之间存在主次关系且可以顺序处理,启发了本文采用可行性规则约束处理方法,通过构造无需额外参数的分层求解机制,从而保证MMAS直接寻优可行解,并在保证零漏警的前提下尽可能地实现多站协同。具体而言(1)目标覆盖层,算法首先随机遍历目标,在满足雷达不过载、处于威力范围及过门限的前提下确保目标完全被派单,特别是将低概率但物理可行的派单作为强制派单的可选集合;(2)多站协同层,算法则再次随机遍历目标,剔除已派单雷达并尝试再派单,充分利用雷达剩余资源但不会为了追求多站协同强行派单。

由单个蚂蚁单次遍历每个雷达 $O(N \cdot J)$ 可得TDA计算复杂度为 $O(K \cdot N \cdot I \cdot J)$ ,实际中雷达节点数目、迭代次数和种群规模已知,计算量与目标规模呈线性关系,一定程度上保证了实时性。此外,TDA在MMAS信息素限幅等机制的基础上,依靠可行性规则保证合法解,进一步降低了初值敏感性。

#### 4 调度层分层式指挥调度方法

如图2所示,所提分层式指挥调度方法采取元

控制器-控制器分层设计<sup>[22]</sup>并基于QL框架完整实现为HQL,通过在长期调度过程中保证策略学习一致性,从而实现搜索、确认和跟踪三种异质任务自适应调度长期最优策略,实际中宏观时间步 $k^{\text{up}}$ 对齐派单时刻 $k$ 。

本文定义分层式指挥调度要素如下:(1)元控制器层为指挥规划层。对应指挥员,根据指挥部的目标派单,综合考虑目标优先级、威胁度和节点资源紧张度,决策节点当前一个阶段的指挥规划指令如响应威胁、维持预警或搜索空域。(2)控制器层为战术执行层。对应战位操作员,根据当前的指挥部目标派单和指挥员指挥规划指令,综合考虑任务紧急度、依赖关系和节点资源余量,具体调度异质任务并执行。值得指出的是,雷达节点直接采用跟踪误差协方差矩阵的迹作为实际跟踪精度指标,表征前线在实时调度过程中真实的防空预警性能;而组网雷达指挥部采用BCRLB作为理论跟踪精度指标,表征指挥部不关注雷达节点具体的跟踪算法,而旨在通过目标派单实现全局防空预警性能长期最优。

##### (1)指挥规划层

在宏观时间步 $k^{\text{up}}$ 的状态空间 $s_{k^{\text{up}}}^{\text{up}}$ 由不同优先级目标占比 $P_{\text{pri}}$ 和被派单目标中包含位置、速度在内的最大不确定性 $\text{tr}(\mathbf{P}_{k^{\text{up}}}^i)_{\text{max}}$ 构成,分箱组合确定状态索引;动作空间 $a_{k^{\text{up}}}^{\text{up}}$ 基于 $s_{k^{\text{up}}}^{\text{up}}$ 及策略参数 $\theta_{k^{\text{up}}}^{\text{up}}$ 按照习得策略 $\pi^{\text{up}}$ 决策,采用任务池构建传递奖励内部交互;外部环境奖励 $R^{\text{up}}$ 由若干分量奖励 $R_{k^{\text{up}},h}^{\text{up}}$ 按 $\beta_{k^{\text{up}}}^{\text{up}}$ 加权获得,引导指挥规划层在异质任务环境中学习长期最优策略。

$$\begin{cases} s_{k^{\text{up}}}^{\text{up}} = \{P_{\text{pri}}, \text{tr}(\mathbf{P}_{k^{\text{up}}}^i)_{\text{max}}\} \\ a_{k^{\text{up}}}^{\text{up}} \sim \pi^{\text{up}}(s_{k^{\text{up}}}^{\text{up}}, \theta_{k^{\text{up}}}^{\text{up}}) \\ R^{\text{up}}(s_{k^{\text{up}}}^{\text{up}}, a_{k^{\text{up}}}^{\text{up}}) = \sum_h \beta_h^{\text{up}} R_{k^{\text{up}},h}^{\text{up}} \end{cases} \quad (7)$$

第一,指挥规划层承接目标派单后实时计算并划分高、中和低目标优先级占比以实时评估当前战场态势。本文选取表征紧迫性的迎击距离优先级

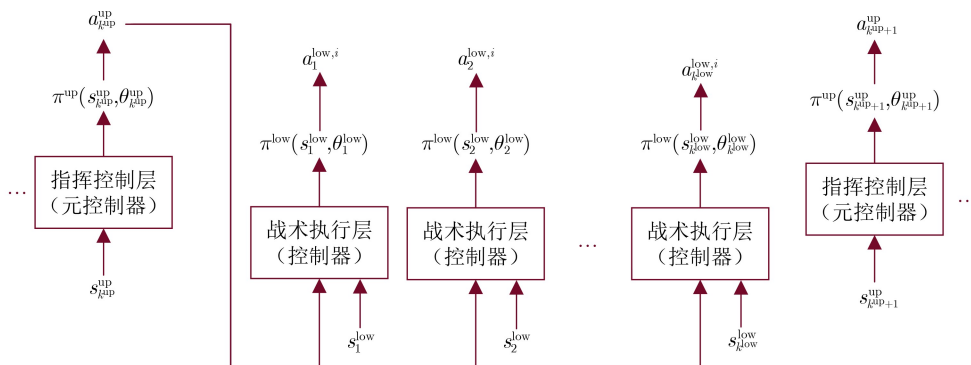


图2 分层式指挥调度方法马尔可夫决策过程

$P_{\text{dist},k}^i$  及权重  $w_{\text{dist}}$ ，表征机动性的角速度优先级  $P_{\text{mea},k}^i$  及权重  $w_{\text{mea}}$  和表征不确定性的跟踪精度优先级  $P_{\text{unc},k}^i$  及权重  $w_{\text{unc}}$  三个维度融合构建当前雷达  $j$  视角下目标  $i$  当前的优先级  $P_k^i$ 。其中， $k$  时刻下  $d_k^{i,j}$  为目标当前距离， $\omega_k^i$  为目标当前角速度， $\eta_k^{i,j}$  为当前跟踪精度， $(d_{\text{max}}, \omega_{\text{max}}, \eta_{\text{max}})$  和  $(d_{\text{min}}, \omega_{\text{min}}, \eta_{\text{min}})$  分别为相应用于归一化的最大值和最小值。实际中可根据具体场景特点，通过调整权重系数设置距离、速度和跟踪精度等因素对目标优先级的影响程度，如包围来袭过程中，目标距离一定呈减小趋势，可在平权设计的基础上略增大该项权重。

$$\begin{cases} P_k^i = w_{\text{dist}} P_{\text{dist},k}^i + w_{\text{mea}} P_{\text{mea},k}^i + w_{\text{unc}} P_{\text{unc},k}^i \\ P_{\text{dist},k}^i = 1 - \frac{d_k^{i,j} - d_{\text{min}}}{d_{\text{max}} - d_{\text{min}}} \\ P_{\text{mea},k}^i = \frac{\omega_k^i - \omega_{\text{min}}}{\omega_{\text{max}} - \omega_{\text{min}}} \\ P_{\text{unc},k}^i = \frac{\eta_k^{i,j} - \eta_{\text{min}}}{\eta_{\text{max}} - \eta_{\text{min}}} \\ w_{\text{dist}} + w_{\text{mea}} + w_{\text{unc}} = 1 \end{cases} \quad (8)$$

第二，本文依据跟踪精度和信干噪比指标，定义指挥规划层外部环境奖励如式(9)所示，包含(1)表征检测性能的信噪比奖励  $\text{SNR}_{k^{\text{up}}}^i$  及其系数  $\beta_{\text{SNR}}$ ；(2)表征防空预警性能的跟踪精度奖励及其系数  $\beta_{\text{P}}$ ；(3)基于信噪比  $\text{SNR}_{k^{\text{up}}}^i$  定义表征资源公平用于每个目标的公平性惩罚。

$$\begin{aligned} R^{\text{up}} = & \beta_{\text{SNR}} \sum_i P_{k^{\text{up}}}^i \frac{\text{SNR}_{k^{\text{up}}}^i}{(\text{SNR}_{k^{\text{up}}}^i)_{\text{max}}} \\ & + \beta_{\text{P}} \sum_i P_{k^{\text{up}}}^i \frac{\text{tr}(\mathbf{P}_{k^{\text{up}}}^i)}{\text{tr}(\mathbf{P}_{k^{\text{up}}}^i)_{\text{max}}} \\ & - \min \left( \beta_{\text{fair}} \text{Var} \left( P_{k^{\text{up}}}^i \cdot \frac{\text{SNR}_{k^{\text{up}}}^i}{(\text{SNR}_{k^{\text{up}}}^i)_{\text{max}}} \right), \right. \\ & \left. \beta_{\text{SNR}} \sum_i P_{k^{\text{up}}}^i \frac{\text{SNR}_{k^{\text{up}}}^i}{(\text{SNR}_{k^{\text{up}}}^i)_{\text{max}}} \right) \end{aligned} \quad (9)$$

第三，定义指挥规划层任务池如表3所示，考虑到上层动作须具备异质任务自适应能力、表征节点动态响应过程且作为作战意图可解释，本文选取控制具体调度方案任务比例作为指挥规划层指挥形式。针对可扩展的参数化封装的任务池，本文在保证上层QL状态动作空间规模可控的前提下经验地将连续指挥规划离散化为三种对应不同基础传递奖励的指挥动作。

## (2) 战术执行层

在微观时间步  $k^{\text{low}}$  的状态空间  $s_{k^{\text{low}}}^{\text{low}}$  由当前接收到的指挥指令  $a_{k^{\text{low}}}^{\text{up}}$ 、每个目标的优先级  $P_{k^{\text{low}}}^i$  和不确

定性  $\text{tr}(\mathbf{P}_{k^{\text{low}}}^i)$  构成，分箱组合确定状态索引；动作空间  $a_k^i$  基于  $s_k^{\text{low}}$  及策略参数  $\theta_k^{\text{low}}$  按照习得策略  $\pi^{\text{low}}$  决策，采用任务池建立任务依赖关系；内部传递奖励  $R^{\text{low}}$  由若干分量奖励  $R_{k,l}^{\text{low}}$  按  $\beta_l^{\text{low}}$  加权获得，引导战术执行层在服从指挥的前提下学习最优任务调度方法。

$$\begin{cases} s_{k^{\text{low}}}^{\text{low}} = \{a_{k^{\text{low}}}^{\text{up}}, P_{k^{\text{low}}}^i, \text{tr}(\mathbf{P}_{k^{\text{low}}}^i)\} \\ a_k^{\text{low},i} \sim \pi^{\text{low}}(s_{k^{\text{low}}}^{\text{low}}, \theta_{k^{\text{low}}}^{\text{low}}) \\ R^{\text{low}}(s_{k^{\text{low}}}^{\text{low}}, a_{k^{\text{low}}}^{\text{low},i}) = \sum_l \beta_l^{\text{low}} R_{k^{\text{low}},l}^{\text{low}} \end{cases} \quad (10)$$

第一，定义战术执行层内部传递奖励。在表3基础上构建比例惩罚机制，通过惩罚各类任务超出的不当比例，引导战术执行层学会服从指挥，调度异质任务符合预期比例。其中， $R_{\text{base}}$  为基础传递奖励， $w_{\text{radio}}$  为比例惩罚权重， $C_{\text{act}}$  为预期任务比例和  $C_{\text{goal}}$  为实际任务比例。

$$R^{\text{low}} = R_{\text{base}} - w_{\text{radio}} [\max(0, C_{\text{act}} - C_{\text{goal}})]^2 \quad (11)$$

第二，本文重点关注时间预算、功率预算和波束预算约束，要求具体调度任务时，波束驻留总用时不超过调度周期上限  $T_{\text{max}}$ ，总消耗功率不超过雷达总功率上限  $P_{\text{max}}$ ，每个目标至多分配一个波束且分配总数不超过  $B_{\text{max}}$ 。其中， $k$  时刻针对目标  $i$  的最终战术执行  $a_k^i$  定义驻留时间  $t_{a_k^i}$ 、消耗功率  $P_{a_k^i}$  和波束覆盖  $B_{a_k^i}$ 。此外，本文采用资源仲裁器审查调度方案，通过驳回并要求重新生成从而保证调度方案满足资源约束。

$$\begin{cases} \sum_i t_{a_k^i} \leq T_{\text{max}} \\ \sum_i P_{a_k^i} \leq P_{\text{max}} \\ \sum_i B_{a_k^i} \leq B_{\text{max}} \end{cases} \quad (12)$$

表 3 指挥规划层任务池定义

编号	命名	预期任务比例	基础传递奖励设计及说明
1	响应	搜索10%	跟踪+8鼓励跟踪
		确认10%	失跟-8更重视失跟惩罚
		跟踪80%	搜索+1允许新增个别搜索 确认+1允许新增个别确认 其他情况-0.1防止怠工 跟踪+6维护现有跟踪
2	维护	搜索20%	失跟-6失跟予以惩罚
		确认20%	搜索+2允许新的搜索
		跟踪60%	确认+2允许新的确认 其他情况-0.1防止怠工 跟踪+2允许继续跟踪
3	扫描	搜索40%	失跟-2允许取消跟踪
		确认40%	确认+4鼓励确认
		跟踪20%	搜索+4鼓励搜索 其他情况-0.1防止怠工

第三, 战术执行层根据战场态势计算任务紧急度作为调度依据, 排序建立任务队列。依据式(13)计算 $k$ 时刻第 $i$ 个目标的任务紧急度 $u_k^i$ , 其中 $u_0$ 是基础任务紧急度。举例而言, 优先级0.9的目标A的搜索任务紧急度9.5可以低于优先级0.1目标的跟踪任务紧急度22.0, 因而应优先维持现有目标B的跟踪以免失跟而不是转而搜索高威胁目标A。

$$u_{k,low}^i = u_0 (1 + P_{k,low}^i) \quad (13)$$

第四, 定义战术执行的任务池如表4所示。考虑雷达节点搜索责任空域、检测到目标后转跟踪的预警流程。首先, 相较于搜索, 跟踪是实现目标状态估计的主要环节, 据此本文设置跟踪任务相比于搜索任务更紧急。在此基础上, 考虑检测随机性, 搜索后到跟踪前往存在一个中间环节, 本文将该环节定义为确认任务, 并将任务紧急度设置为介于二者之间, 以期引导分层式指挥习得搜索后尽快确认, 确认成功后转跟踪的决策逻辑。此外, 考虑到单个雷达资源有限, 有可能出现接单后剩余资源不足的情形, 此时会出现部分目标未被实质处理、如同雷达静默的现象, 本文对此不单独建立任务类型。综上所述, 考虑到下层动作需标准化封装异质预警任务, 本文基于经验选取时间、功率、基础任务紧急度和任务依赖关系参数, 为保证任务抽象与具体装备参数无关, 本文定义任务的时间或功率资源消耗为该任务消耗在总可用时间或功率的占比, 而基础任务紧急度参数基于经验设置。此外, 本文定义任务状态机检查任务依赖关系, 保证预警任务执行符合实际规则, 并采用失跟计数器实现跨宏观步的失跟维护。

## 5 实验结果与分析

### 5.1 协同层与调度层消融及对比实验设计

本文设计一种饱和和攻击要地防空仿真场景<sup>[4,5]</sup>。(1)模拟我方针对半径20 km圆形防区, 沿前线部署3部同构S波段中程相控阵雷达组网预警, 其中雷达工作频率4GHz, 工作带宽10 MHz, 发射功率

15 kW, 天线增益30 dB, 系统损耗约3 dB, 脉冲积累数100, 虚警率 $10^{-6}$ , 检测门限0 dB, 旁瓣干扰功率5 w, 干扰旁瓣抑制-30 dB, 噪声温度190 K, 噪声系数2, 采用扩展卡尔曼滤波算法<sup>[25]</sup>, 单站处理上限50个目标, 时间预算100%和功率预算100%均不留裕量, 目标派单按宏观时间步1 s, 包含任务调度5个微观时间步, 不考虑通信延迟;(2)模拟敌方利用200个高速机动目标, 随机分布半径25 km到60 km环形区域包围来袭, 划分25 km、35 km和45 km三组包围圈半径模拟来袭过程, 速度范围100至300 m/s模拟从中速无人机至高亚音速巡飞弹, RCS服从Strling模型I/II采用逆变换采样法近似为随机取值0.1到 $1.5m^2$ 模拟高速巡飞弹至中型攻击无人机。

实验旨在验证(1)仅考虑目标派单时, 所提TDA算法从收敛速度、解质量和实时性方面体现所提架构在适用处理规模方面的改进;(2)仅考虑异质任务调度时, 所提调度方法从学习效率、策略质量、任务执行分布、目标失跟率和平均跟踪误差等方面体现所提架构在环境自适应能力、长期调度能力和可解释性方面的改进。如表5所示, 首先给出本文中权重系数及其他必要仿真参数的具体设定, 长度单位为米, 角度单位为弧度, 需要指出的是参数皆通过实验调优。

#### (1)协同层消融及对比实验

实验分为两组, 第一组为MMAS、人工蜂群算法(Artificial Bee Colony, ABC)<sup>[11]</sup>和遗传算法(Genetic Algorithm, GA)<sup>[16]</sup>构成的横向对比, ABC代表连续域求解优势的群体智能算法, 验证MMAS离散域求解优势。GA代表进化算法, 验证群体智能算法与进化算法在求解所提优化问题上的性能差异, 三者都采用罚函数法处理约束。

第二组为TDA和MMAS组成的消融测试, TDA除分层求解机制外与MMAS保持一致, 验证本文采取可行性规则代替罚函数法所带来的性能改进。具体参数设置如表6所示。其中, 罚函数针对过载和漏警予以严惩, 保证非法解不影响公平对比。

表 4 战术执行层任务池定义

编号	命名	时间消耗	功率消耗	基础任务紧急度	任务依赖关系
1	搜索	5%	5%	5	第一次发现目标时 搜索后未及时确认 确认后未及时跟踪 失跟后未及时确认
2	确认	10%	10%	10	搜索后且处于确认有效期内 失跟后但处于确认有效期内
3	跟踪	20%	20%	20	确认后处于跟踪有效期内 未失跟时

表 5 基本实验参数设置

分量系数	符号	实验设定	分量系数	符号	实验设定
权重因子	$\alpha_U$	0.5	信噪比奖励系数	$\beta_{SNR}$	0.4
雷达自适应权重初值	$\beta_{T,0}$	1.0	跟踪精度奖励系数	$\beta_P$	-0.4
雷达自适应权重初值	$\beta_{R,0}$	1.0	公平性惩罚系数	$\beta_{fair}$	0.2
目标自适应权重基准系数	$\beta_{T,th}$	0.8	比例惩罚权重	$w_{radio}$	200
雷达自适应权重基准系数	$\beta_{R,th}$	0.8	综合作战效能权重	$w_{base}$	10
自适应权重最大值	$\beta_{T,max}/\beta_{R,max}$	1.0	过载惩罚权重	$w_{cap}$	0.5
自适应权重最小值	$\beta_{T,min}/\beta_{R,min}$	0	漏警惩罚系数	$w_{miss}$	0.1
波束指向调整代价权重	$w_1$	0.25	加速度过程噪声	$Q_a$	0.9
资源充分利用收益权重	$w_2$	0.25	转弯过程噪声	$Q_w$	0.2
跟踪精度收益权重	$w_3$	0.25	测距噪声	$R_r$	120
多站协同收益权重	$w_4$	0.5	测角噪声	$R_\theta$	$(0.5 \cdot \frac{\pi}{180})^2$
超负荷代价权重	$w_5$	0.25	测速噪声	$R_v$	5
高威胁目标战术倾斜权重	$w_6$	0.5	PBM资源加权	$\beta_{PBM}$	300
雷达自适应权重更新系数	$\gamma_R$	0.12	窗口大小	$W$	10
目标自适应权重更新系数	$\gamma_T$	0.12	失跟计数阈值	$N_{lost}$	5
迎击距离优先级权重	$w_{dist}$	0.4	蛇形机动幅度	$A_{snake}$	[0.05,0.15]
角速度优先级权重	$w_{mea}$	0.3	蛇形机动频率	$f_{snake}$	[0.1,0.3]
跟踪精度优先级权重	$w_{unc}$	0.3	来袭方向扰动范围	$\theta_{coming}$	$\frac{\pi}{4}$
高优先级阈值	$P_{high}$	0.7	初始协方差	$P_0$	$[500,500,100(10 \cdot \frac{\pi}{180})^2,0.1]$
中优先级阈值	$P_{medium}$	0.4	跟踪精度阈值	$y_{e_{th}}$	50

表 6 求解算法对比实验参数设置

参数类型	TDA	MMAS	ABC	GA
种群规模 $N$	30/100	30/100	30/100	30/100
最大迭代 $K$	100	100	100	100
信息素权重 $\beta_r$	1.0	1.0	-	-
启发式权重 $\beta_U$	2.0	2.0	-	-
挥发率 $\rho$	0.1	0.1	-	-
信息素常数 $Q$	100	100	-	-
特有参数	信息素最大值10 信息素最小值0.1	信息素最大值10 信息素最小值0.1	停滞阈值5 观察蜂0.5	交叉率0.8 变异率0.1

## (2) 调度层消融及对比实验

实验分为两组，第一组为QL和基于优先级(Priority-Based Method, PBM)<sup>[10]</sup>的方法组成的横向对比，QL代表智能调度方法，直接从环境中学习调度三种异质任务；根据式(14)PBM代表基于固定规则的方法，本文设计综合考虑任务紧急度和时间资源成本的加权动态优先级，并采用最大权重调度以验证智能调度相比于固定规则的环境自适应能力优势。

$$a_k^i = \arg \max_{a_k^i \in A(s_k^i)} J(u_k^i, P_k^i, t_a) = u_k^i - \beta_{PBM} \cdot t_{a_k^i} (1 - P_k^i) \quad (14)$$

其中，第  $k$  时刻针对目标  $i$  的预警任务  $a_k^i$  需从当前状态下可选任务集合  $A(s_k^i)$  中选取，在紧急任务优先调度的基础上，考虑时间资源  $t_{a_k^i}$  成本的限制，并动态根据目标优先级的升高相应放宽限制，据此调度任务并仲裁，通过后予以执行。

第二组为HQL和QL构成的消融测试，QL代表单层强化学习方法，验证本文采取分层框架的性能改进。具体参数设置如表7所示，根据式(7)和式(10)，实际有效状态动作空间远小于理论值，本文实验条件下可保证相对固定的状态-动作组合被有效遍历，且学习率按线性衰减，从而保证稳定收敛以验证分层设计有效性。

表7 调度方法对比实验参数设置

参数类型	HQL	QL	PBM
状态空间 $s$	上层1331/下层363	121	-
动作空间 $a$	上层3/下层4	4	3
训练轮数 $e$	1000	1000	-
每轮步数 $n$	100	100	-
学习率 $\alpha$	上层0.05/下层0.1	0.1	-
折扣因子 $\gamma$	0.85	0.85	-
探索率 $\epsilon$	1.1至0.005	1.1至0.005	-
算法类型	指挥规划-战术执行双层Q表	单层Q表	最高优先级优先调度

## 5.2 评价指标设计

为定量分析目标派单求解能力与异质任务调度能力，本文定义评价指标如下。

### (1) 寻优效率与解质量

根据综合作战效能，结合过载与漏警惩罚评估求解算法的寻优效率及解质量。其中， $F_k$ 是第 $k$ 次迭代时的评估值， $w_{\text{base}}$ 、 $w_{\text{cap}}$ 和 $w_{\text{miss}}$ 分别为综合作战效能、过载惩罚和漏警惩罚权重。评估曲线上升速率越大，求解算法的寻优效率越高；稳态收敛值越大，目标派单方案的质量越高。

$$F_k = w_{\text{base}} \sum_{j \in J} \sum_{i \in I} U_k^{i,j} \cdot x_k^{i,j} - w_{\text{cap}} \sum_{j \in J} \left( N_J - \sum_{i \in I} x_k^{i,j} \right) - w_{\text{miss}} \sum_{i \in I} \sum_{j \in J} x_k^{i,j} \quad (15)$$

### (2) 实时性 $T_{\text{cost}}$

记录求解算法从输入综合作战效能矩阵至输出目标派单方案的计算用时。其中， $t_{\text{st}}$ 为启动时刻， $t_{\text{end}}$ 为完成时刻。预警过程中方案求解耗时时必须严格小于目标派单周期，求解算法耗时越短，系统实时响应能力越强。

$$T_{\text{cost}} = t_{\text{end}} - t_{\text{st}} \quad (16)$$

### (3) 学习效率与策略质量

根据外部奖励，结合滑窗历史评估调度方法的学习效率与策略质量。其中， $W$ 为窗口大小， $R(i)$ 为第 $i$ 轮奖励值。平均外部奖励曲线上升速率越大，调度方法的学习效率越高；稳态收敛值越大，异质任务调度策略越优。

$$R_W(\hat{k}) = \frac{1}{W} \sum_{i=\hat{k}-W+1}^{\hat{k}} R(i) \quad (17)$$

### (4) 任务执行分布 $P_{\text{act}}$ 与丢弃率 $P_{\text{drop}}$

统计在遵守资源约束与任务依赖关系的前提

下，调度方法调度并成功执行搜索、确认、跟踪三类异质任务的占比及相应丢弃率。其中， $N_{\text{act}}$ 为特定类型任务成功执行数， $N_{\text{all}}$ 为任务总调度数， $N_{\text{drop}}$ 为因资源不足或违反任务依赖关系，虽调度但未能执行的特定类型任务丢弃数。搜索、确认和跟踪三类任务执行占比越大，异质任务调度策略资源利用率越高；任务丢弃率 $P_{\text{drop}}$ 过高表明长期调度能力弱，过低则表明环境自适应能力弱。

$$\begin{cases} P_{\text{act}} = \frac{N_{\text{act}}}{N_{\text{all}}} \times 100\% \\ P_{\text{drop}} = \frac{N_{\text{drop}}}{N_{\text{all}}} \times 100\% \end{cases} \quad (18)$$

### (5) 目标失跟率 $P_{\text{loss}}$

统计仿真过程中调度方法的目标失跟率 $P_{\text{loss}}$ 作为防空预警性能直接指标。其中， $N_{\text{loss}}$ 为总失跟次数，仿真过程总轮数取训练轮数。目标失跟率越低，调度方法越倾向于维护更多的跟踪任务。失跟率不一定越低越好，需结合平均跟踪误差指标才能综合评价预警性能。

$$P_{\text{loss}} = \frac{N_{\text{loss}}}{e} \times 100\% \quad (19)$$

### (6) 平均跟踪误差 $e_k$

统计 $k$ 时刻有效调度预警任务的目标的平均跟踪误差作为调度后防空预警性能直接指标，平均跟踪误差越小，防空预警性能越好。其中， $N_e$ 为有效调度预警任务的目标数目， $\mathbf{x}_k^i$ 和 $\hat{\mathbf{x}}_k^i$ 分别为目标 $i$ 在 $k$ 时刻的真实坐标和估计坐标。平均跟踪误差越小，调度方法对现有目标的防空预警性能越好。

$$e_k = \frac{1}{N_e} \sum_{i=1}^{N_e} \sqrt{(\mathbf{x}_k^i - \hat{\mathbf{x}}_k^i)^T (\mathbf{x}_k^i - \hat{\mathbf{x}}_k^i)} \quad (20)$$

## 5.3 含消融测试的横向对比

如图3所示，针对饱和和攻击场景，所提算法可有效求解所提大规模非线性多约束目标派单模型。随着包围圈不断缩小，组网雷达探测概率不断增大，目标派单需求上升，二分图从稀疏逐渐稠密，

进而影响不同求解算法目标派单质量。具体而言，ABC作为连续域优势算法，在二分图问题上的寻优效率有限，容易陷入局部最优，而来袭后期二分图稠密程度增大，连续性增强使ABC派单质量相对前期改善，但仍最差。GA二分图搜索效率相比于ABC更高，但缺乏强的正反馈，来袭后期二分图稠密程度增大时难以在有限迭代内收敛至全局最优解。MMAS则可以迅速探索二分图，特别是来袭后期可以在有限迭代内求出相比于ABC和GA质量更高的目标派单方案，验证了针对所提目标派单模型选择蚁群算法求解较为合适。然而，注意到采用罚函数法处理约束时，因所提模型非线性、多约束产生的非凸性，依靠松弛约束为惩罚项较难引导MMAS求出合法解，例如可能因追求多站协同而产生漏警惩罚。而TDA采用可行性规则处理约束，构造引导蚂蚁先完成目标覆盖硬约束，再考虑利用雷达剩余资源增加多站协同，无需额外参数保证求解作战效能最优的合法解，验证了可行性规则约束处理方法通过利用所提模型先验信息，相比于罚函数法有效改进了MMAS寻优效率和解质量。

如图4所示，所提算法耗时低于目标派单周期上限，在满足战时实时性要求的前提下以毫秒级的

时间代价有效改进了大规模场景目标派单质量。计算用时主要取决于算法计算复杂度而无关二分图稠密与否，整个来袭过程中基本保持稳定。具体而言，GA在每次迭代中直接并行地交叉、变异智能体，计算复杂度极小，用时最短。ABC虽串行运算雇佣蜂-跟随蜂过程，但只进行领域搜索和贪婪选择，计算复杂度较小，用时较短。MMAS虽并行运算蚂蚁过程，但信息素全局更新和限幅等比较机制属于串行过程，大幅增加了MMAS计算用时，且二分图越稠密，全局比较用时越多。TDA在MMAS基础上引进了串行的分层求解机制，大幅增加了计算复杂度，使TDA计算用时明显高于其他算法。然而，所提架构下预警过程可划分为目标派单和节点调度两个阶段。一方面，指挥部融合全网信息时需综合考虑数据链通信延迟、拥塞与运动目标是否飞出波束等约束，一般设定秒级派单周期；另一方面，各节点可参照独立时钟，按脉冲频率实现微秒级调度，保证指挥部每个目标派单周期内各节点都能并行自主调度多个具体波束驻留序列。协同层算法性能边界可通过测试不同目标规模和雷达数目下派单用时是否超过宏观时间步确定。

如图5所示，随着包围圈不断缩小，目标威胁

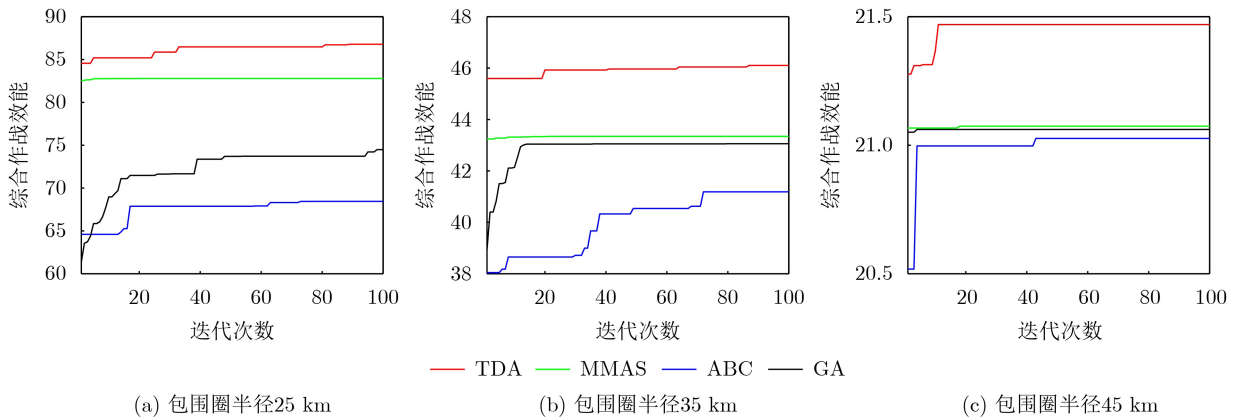


图3 求解算法评估曲线对比

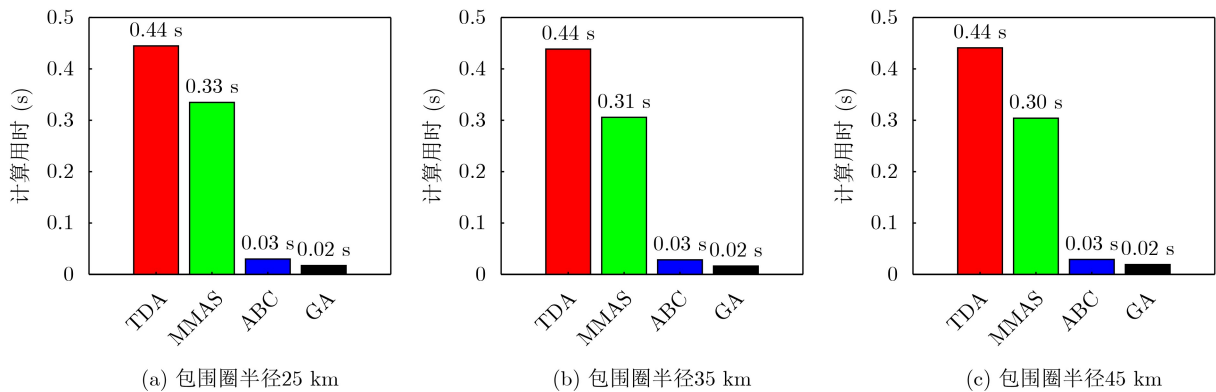


图4 求解算法计算用时对比

度整体呈上升趋势，据此可展现不同调度方法的异质任务调度策略质量。具体而言，PBM最大权重调度属于贪婪策略且动态优先级自适应能力有限，策略质量相比于QL和HQL更低，验证了针对所提调度场景智能调度相比于固定规则长期策略更优、环境自适应能力更强。QL通过在未知环境中自适应学习调度策略改进了策略质量，但面对长期调度过程即时决策难以保证策略稳定，表现为学习曲线震荡，此外环境奖励稀疏使得学习速度较慢，有限训练轮数下仅收敛于次优策略。HQL区分内部和外部奖励曲线，此处选取平均外部奖励曲线进行公平对比，一方面HQL通过宏观步-微观步划分保证了策略一致性，曲线平滑且稳步上升；另一方面内部密集的传递奖励环境有效加速了学习，同等训练轮数下习得策略相比于QL更优，验证了针对所提调度场景指挥规划层-战术执行层分层设计可有效改善智能调度方法的策略质量和学习效率。

如图6所示，所提方法综合考虑资源约束和任务依赖关系的长期影响，通过维持比例均衡的任务调度方案，从而最大化综合作战效能的同时改善了资源利用效率。任务执行分布和丢弃率主要取决于调度方法既有调度策略而无关二分图稠密与否，整

个来袭过程中基本保持稳定。具体而言，PBM考虑资源成本和可选任务集规则本质上形成预仲裁，保证了低任务丢弃率，但过分关注威胁响应和资源约束忽视了跟踪、确认任务平衡，影响整个任务调度链最终导致资源浪费和防空预警性能下降，相比于HQL和QL而言异质任务调度能力不足。QL习得了任务比例平衡的重要性，但短视调度了过多的跟踪任务，容易造成资源过载，进而导致防空预警性能下降，总体上也产生了更高的任务丢弃率。HQL则凭借分层设计习得了更好的调度策略改善了资源利用方式，表现为主动维护更均衡的任务比例，相比于QL调度方案更合理、丢弃率更低。

如图7所示，目标失跟率和平均跟踪误差更具体地刻画了调度方法不同的调度策略各自对防空预警性能产生的影响。注意到失跟率与任务执行比例直接相关，因而距离不敏感；而平均跟踪误差与信号质量等有关，受目标距离直接影响，随着目标逼近而整体改善。以PBM为基准，QL策略尽可能维护更多的目标跟踪任务，从而降低了目标失跟率，符合图6结论，但因资源过载导致平均跟踪误差增大，此外误差曲线在整个训练过程震荡也印证了图5结论。与之相反，HQL主动压减低质量跟踪任

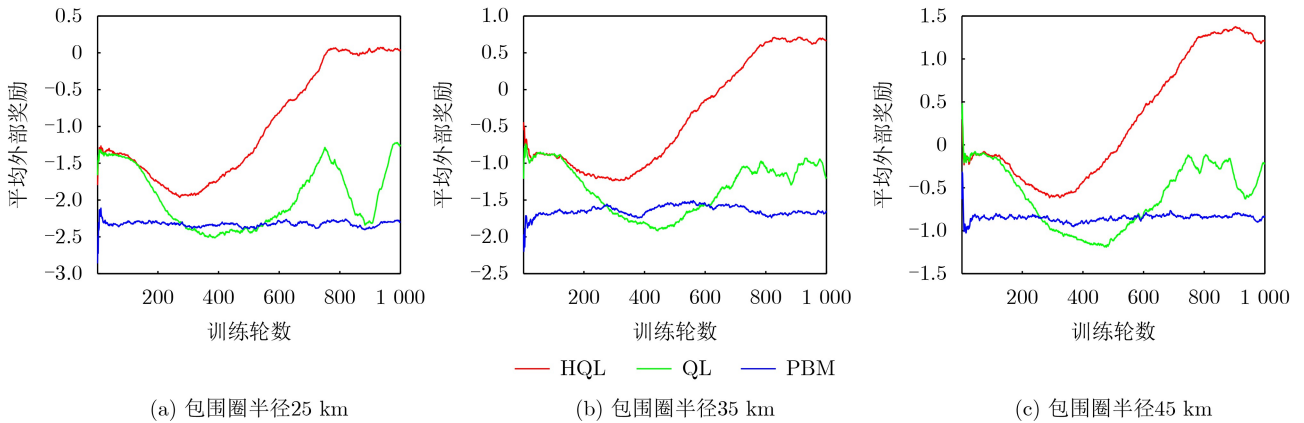


图5 调度方法平均外部奖励曲线对比

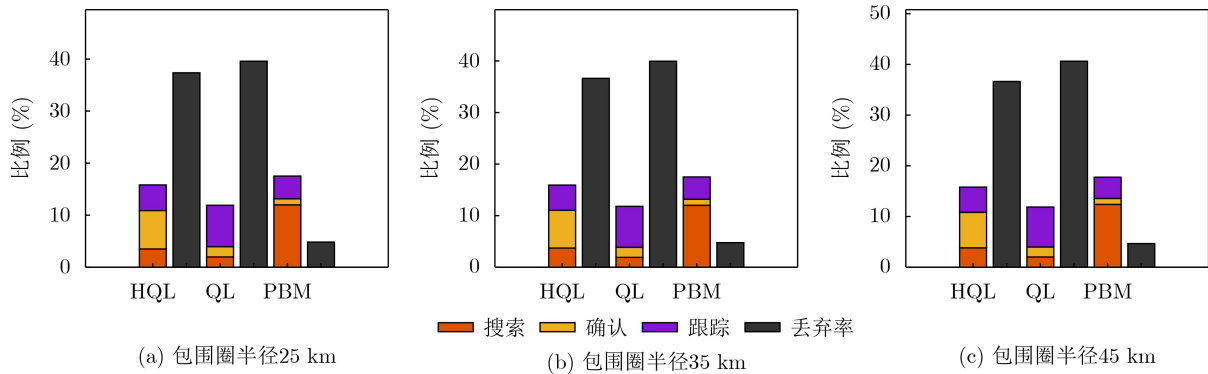


图6 调度方法任务执行分布对比

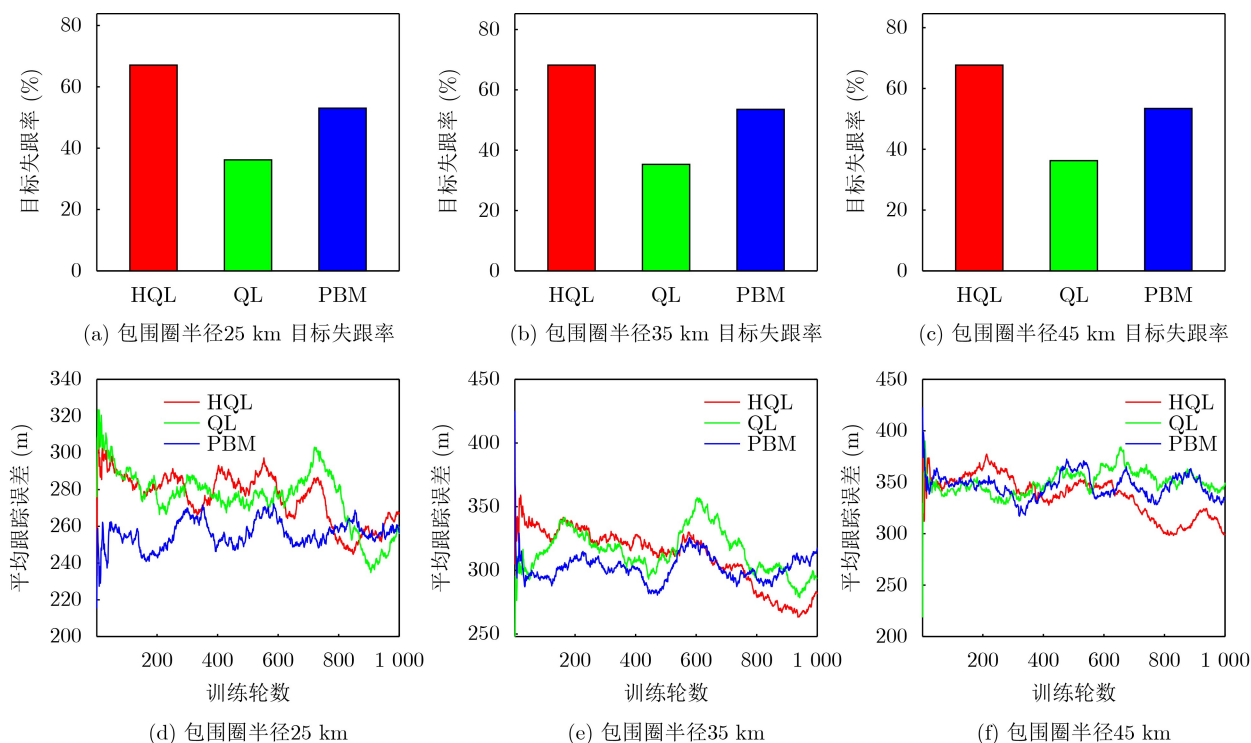


图7 调度方法目标失跟率和平均距离误差对比

务，因而相比PBM和QL目标失跟率都更高，符合图6结论，但从执行搜索、确认和跟踪完整调度链的角度，HQL分层指挥式调度策略聚焦长期调度最优，集中资源在高质量跟踪任务，从而具有最小的平均跟踪误差，且平均跟踪误差随着训练轮数增加呈现稳定下降趋势，印证了图5结论。综合来看，HQL长期策略最优，分层设计有效改善了异质任务调度能力。调度层算法性能边界可通过目标失跟率和平均跟踪精度综合评价确定。

考虑到任务场景、假设等方面的不同，如表8所示本文限定混合结构、分级控制模式和组网雷达指挥部主导这一范围，定性给出所提架构与其他代表性架构的横向对比。所提架构以可拓展性为目标构建了执行-再派单认知闭环，一方面在协同层建立了多维资源统一管理，所提TDA改善了适用处理规模；另一方面，在调度层建立了分层式指挥，所提HQL在改善环境自适应能力和长期调度能力的同时改善了可解释性。

## 6 结束语

本文针对饱和攻击防空预警场景，以可拓展性为目标提出了一种决策层次上融合了集中式动态目标分配和分布式异质任务调度的混合架构，设计了执行-再派单认知闭环、目标派单算法和分层式指挥调度方法，改善了动态协同和异质任务调度能力，实现了大规模、高速度饱和攻击场景下组网雷

表8 协同调度架构横向对比

	拍卖算法 <sup>[11]</sup>	其他分层强化学习 <sup>[26]</sup>	所提架构
任务场景	反导预警	空天网络计算 资源调度	饱和攻击 防空预警
任务假设	异质	同质	异质
适用处理规模	中	大	大
环境自适应能力	中	强	强
长期调度能力	弱	强	强
可拓展性	中	中	强
可解释性	强	弱	强

达协同调度搜索、确认和跟踪三类异质预警任务，释放了预警体系抗饱和攻击潜力。仿真典型场景验证所提架构在适用处理规模、环境自适应能力、长期调度能力、可拓展性和可解释性等方面具有改进，具有一定工程应用价值。未来的研究工作可聚焦于两个方向，一是探索在线学习，以期解决离线训练-在线部署的模型失真问题；二是探索开放式架构，将火力分配等纳入协同调度，推广预警-火控一体化防御体系。

## 参考文献

- [1] YAN Junkun, JIAO Hao, PU Wenqiang, et al. Radar sensor network resource allocation for fused target tracking: A brief review[J]. *Information Fusion*, 2022, 86/87: 104-115. doi: 10.1016/j.inffus.2022.06.009.
- [2] 易伟, 袁野, 刘光宏, 等. 多雷达协同探测技术研究进展: 认知

- 跟踪与资源调度算法[J]. 雷达学报, 2023, 12(3): 471–499. doi: [10.12000/JR23036](https://doi.org/10.12000/JR23036).
- YI Wei, YUAN Ye, LIU Guanghong, *et al.* Recent advances in multi-radar collaborative surveillance: Cognitive tracking and resource scheduling algorithms[J]. *Journal of Radars*, 2023, 12(3): 471–499. doi: [10.12000/JR23036](https://doi.org/10.12000/JR23036).
- [3] 丁建江, 许红波, 周芬. 雷达组网技术[M]. 北京: 国防工业出版社, 2017.
- DING Jianjiang, XU Hongbo, and ZHOU Fen. The Technology of Netted Radar System[M]. Beijing: National Defense Industry Press, 2017. (查阅网上资料, 未找到本条文献页码信息, 请补充).
- [4] 时晨光, 蒋泽宇, 严牧, 等. 针对组网雷达的无人机集群航迹欺骗综合误差分析[J]. 电子与信息学报, 2024, 46(12): 4451–4458. doi: [10.11999/JEIT240289](https://doi.org/10.11999/JEIT240289).
- SHI Chenguang, JIANG Zeyu, YAN Mu, *et al.* Comprehensive error in UAV cluster trajectory deception for networked radar[J]. *Journal of Electronics & Information Technology*, 2024, 46(12): 4451–4458. doi: [10.11999/JEIT240289](https://doi.org/10.11999/JEIT240289).
- [5] 吴桐, 元统帅, 谢伟朋. 无人机集群反制技术研究[J]. 空天防御, 2025, 8(1): 17–23. doi: [10.3969/j.issn.2096-4641.2025.01.003](https://doi.org/10.3969/j.issn.2096-4641.2025.01.003).
- WU Tong, QI Tongshuai, and XIE Weipeng. Research on anti-UAV cluster technology[J]. *Air & Space*, 2025, 8(1): 17–23. doi: [10.3969/j.issn.2096-4641.2025.01.003](https://doi.org/10.3969/j.issn.2096-4641.2025.01.003).
- [6] 何伟. 高机动超远程防空作战中的电火协同拦截决策技术[J]. 指挥控制与仿真, 2025, 47(1): 77–83. doi: [10.3969/j.issn.1673-3819.2025.01.010](https://doi.org/10.3969/j.issn.1673-3819.2025.01.010).
- HE Wei. Electronic and firepower collaborative interception decision technology in high mobility ultra short range air defense operations[J]. *Command Control & Simulation*, 2025, 47(1): 77–83. doi: [10.3969/j.issn.1673-3819.2025.01.010](https://doi.org/10.3969/j.issn.1673-3819.2025.01.010).
- [7] 何友, 王国宏, 陆大焱, 等. 多传感器信息融合及应用[M]. 2版. 北京: 电子工业出版社, 2007.
- HE You, WANG Guohong, LU Dajin, *et al.* Multisensor Information Fusion with Application[M]. 2nd ed. Beijing: Publishing House of Electronics Industry, 2007. (查阅网上资料, 未找到本条文献页码信息, 请补充).
- [8] 杨露菁, 余华. 多源信息融合理论与应用[M]. 北京: 北京邮电大学出版社, 2011.
- YANG Lujing and YU Hua. Theory and Application of Multi-Source Information Fusion[M]. Beijing: Beijing University of Posts and Telecommunications Press, 2011. (查阅网上资料, 未找到本条文献页码信息及英文翻译, 请确认).
- [9] 薛琛衍. 数字阵列雷达资源管理研究[D]. [博士学位论文], 南京航空航天大学, 2023. doi: [10.27239/d.cnki.gnhhu.2023.000104](https://doi.org/10.27239/d.cnki.gnhhu.2023.000104).
- XUE Chenyan. Research on resource management of digital array radar[D]. [Ph. D. dissertation], Nanjing University of Aeronautics and Astronautics, 2023. doi: [10.27239/d.cnki.gnhhu.2023.000104](https://doi.org/10.27239/d.cnki.gnhhu.2023.000104).
- [10] 毕增军, 徐晨曦, 张贤志, 等. 相控阵雷达资源管理技术[M]. 北京: 国防工业出版社, 2016.
- BI Zengjun, XU Chenxi, ZHANG Xianzhi, *et al.* Phased Array Radar Resource Management Technology[M]. Beijing: National Defense Industry Press, 2016. (查阅网上资料, 未找到本条文献页码信息及英文翻译, 请确认).
- [11] 熊奎. 多雷达联盟博弈智能协同跟踪理论与方法研究[D]. [博士学位论文], 电子科技大学, 2024. doi: [10.27005/d.cnki.gdzku.2024.000037](https://doi.org/10.27005/d.cnki.gdzku.2024.000037).
- XIONG Kui. Research on theory and method of multi-radar coalition game intelligent cooperative tracking[D]. [Ph. D. dissertation], University of Electronic Science and Technology of China, 2024. doi: [10.27005/d.cnki.gdzku.2024.000037](https://doi.org/10.27005/d.cnki.gdzku.2024.000037).
- [12] 蒋李兵, 杨庆伟, 郑舒予, 等. 基于拍卖理论的组网雷达多轨道目标ISAR成像资源分配算法[J]. 系统工程与电子技术, 2025, 47(1): 81–93. doi: [10.12305/j.issn.1001-506X.2025.01.09](https://doi.org/10.12305/j.issn.1001-506X.2025.01.09).
- JIANG Libing, YANG Qingwei, ZHENG Shuyu, *et al.* Multi-orbit targets ISAR imaging resource allocation algorithm for netted radar based on auction theory[J]. *Systems Engineering and Electronics*, 2025, 47(1): 81–93. doi: [10.12305/j.issn.1001-506X.2025.01.09](https://doi.org/10.12305/j.issn.1001-506X.2025.01.09).
- [13] PIEZZO M, AUBRY A, BUZZI S, *et al.* Non-cooperative code design in radar networks: A game-theoretic approach[J]. *EURASIP Journal on Advances in Signal Processing*, 2013, 2013(1): 63. doi: [10.1186/1687-6180-2013-63](https://doi.org/10.1186/1687-6180-2013-63).
- [14] 李浩情, 余点, 潘常春, 等. 基于模型知识融合的图神经网络多雷达协同任务调度算法[J]. 雷达学报, 2025, 14(2): 470–485. doi: [10.12000/JR24222](https://doi.org/10.12000/JR24222).
- LI Haoqing, YU Dian, PAN Changchun, *et al.* Multiradar collaborative task scheduling algorithm based on graph neural networks with model knowledge embedding[J]. *Journal of Radars*, 2025, 14(2): 470–485. doi: [10.12000/JR24222](https://doi.org/10.12000/JR24222).
- [15] 杜永浩, 张本奎, 吴健, 等. 大规模遥感卫星智能任务调度方法研究进展[J]. 电子与信息学报, 2025, 47(12): 5033–5047. doi: [10.11999/JEIT251038](https://doi.org/10.11999/JEIT251038).
- DU Yonghao, ZHANG Benkui, WU Jian, *et al.* Survey on intelligent methods for large-scale remote sensing satellite scheduling[J]. *Journal of Electronics & Information Technology*, 2025, 47(12): 5033–5047. doi: [10.11999/JEIT251038](https://doi.org/10.11999/JEIT251038).
- [16] 杜永浩, 黎磊, 徐世龙, 等. 基于智能优化算法引擎的可演进星群智能任务规划[J]. 电子与信息学报, 2025, 47(6): 1645–1657. doi: [10.11999/JEIT240974](https://doi.org/10.11999/JEIT240974).
- DU Yonghao, LI Lei, XU Shilong, *et al.* Evolutionary optimization for satellite constellation task scheduling based on intelligent optimization engine[J]. *Journal of Electronics & Information Technology*, 2025, 47(6): 1645–1657. doi: [10.11999/JEIT240974](https://doi.org/10.11999/JEIT240974).

- [17] JIN Biao, KUANG Xiaofei, LIU Shujin, *et al.* Joint allocation of transmit power and signal bandwidth for distributed cognitive tracking radar network using cooperative game[J]. *Digital Signal Processing*, 2023, 135: 103964. doi: [10.1016/j.dsp.2023.103964](https://doi.org/10.1016/j.dsp.2023.103964).
- [18] 焦浩, 严俊坤, 郝佳, 等. 面向多机动目标的资源分配与精细化跟踪算法[J]. *雷达学报*, 2026, 15(1): 292–306. doi: [10.12000/JR25037](https://doi.org/10.12000/JR25037).  
JIAO Hao, YAN Junkun, HAO Jia, *et al.* Resource allocation and precise tracking algorithm for multiple maneuvering targets[J]. *Journal of Radars*, 2026, 15(1): 292–306. doi: [10.12000/JR25037](https://doi.org/10.12000/JR25037).
- [19] 周琳, 王祎, 朱效洲, 等. 面向反导预警作战的多传感器任务分配拍卖方法[J]. *现代雷达*, 2023, 45(6): 42–50. doi: [10.16592/j.cnki.1004-7859.2023.06.006](https://doi.org/10.16592/j.cnki.1004-7859.2023.06.006).  
ZHOU Lin, WANG Yi, ZHU Xiaozhou, *et al.* Multi-sensor task assignment auction method for anti-missile Early warning campaign[J]. *Modern Radar*, 2023, 45(6): 42–50. doi: [10.16592/j.cnki.1004-7859.2023.06.006](https://doi.org/10.16592/j.cnki.1004-7859.2023.06.006).
- [20] 崔莉莉. 基于蚁群算法的武器—目标分配问题研究[D]. [硕士学位论文], 上海交通大学, 2011.  
CUI Lili. Ant colony algorithm for solving the weapon-target assignment problem[D]. [Master dissertation], Shanghai Jiaotong University, 2011.
- [21] 孙玉洁, 苏波, 段乐毅, 等. 空地协同侦察任务派单算法研究[J]. *兵工学报*, 2024, 45(S2): 317–328. doi: [10.12382/bgxb.2024.0319](https://doi.org/10.12382/bgxb.2024.0319).  
SUN Yujie, SU Bo, DUAN Leyi, *et al.* A reconnaissance task allocation algorithm for air-ground heterogeneous unmanned system[J]. *Acta Armamentarii*, 2024, 45(S2): 317–328. doi: [10.12382/bgxb.2024.0319](https://doi.org/10.12382/bgxb.2024.0319).
- [22] KULKARNI T D, NARASIMHAN K R, SAEEDI A, *et al.* Hierarchical deep reinforcement learning: Integrating temporal abstraction and intrinsic motivation[C]. *Advances in Neural Information Processing Systems 29: Annual Conference on Neural Information Processing Systems 2016*, Barcelona, Spain, 2016.
- [23] 毛滔, 蔺美青, 周建飞. 地面防空预警雷达探测概率仿真模型研究[J]. *空军预警学院学报*, 2013, 27(6): 430–434. doi: [10.3969/j.issn.2095-5839.2013.06.010](https://doi.org/10.3969/j.issn.2095-5839.2013.06.010).  
MAO Tao, LIN Meiqing, and ZHOU Jianfei. Simulation model for detection probability of ground air defense Early-warning radar[J]. *Journal of Air & Space Early Warning Research*, 2013, 27(6): 430–434. doi: [10.3969/j.issn.2095-5839.2013.06.010](https://doi.org/10.3969/j.issn.2095-5839.2013.06.010).
- [24] FANG Yuyuan, WEI Shaopeng, ZHANG Lei, *et al.* Online emission policy selection for radar anti-jamming using bandit-optimized policy search[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2024, 60(3): 3132–3147. doi: [10.1109/TAES.2024.3358793](https://doi.org/10.1109/TAES.2024.3358793).
- [25] MAHAFZA B R. Radar Systems Analysis and Design Using MATLAB[M]. Boca Raton: CRC Press, 2005. (查阅网上资料, 未找到本条文献页码信息, 请补充).
- [26] WEI Yuting, JI Zhe, WU Sheng, *et al.* Hierarchical reinforcement learning for task scheduling in space-air integrated edge computing networks[J]. *IEEE Internet of Things Journal*, 2025, 12(23): 49722–49736. doi: [10.1109/JIOT.2025.3605153](https://doi.org/10.1109/JIOT.2025.3605153).
- 叶炬航: 男, 硕士生, 研究方向为雷达资源管理。  
方愚渊: 男, 副教授(讲师), 研究方向为组网雷达、认知电子战等。  
魏少鹏: 男, 副教授, 研究方向为雷达信号处理、雷达干扰对抗、雷达成像、雷达目标探测。  
段佳: 女, 副教授, 研究方向为雷达成像、电磁特征提取。  
张磊: 男, 教授, 研究方向为SAR、ISAR高分辨成像与运动补偿等。

责任编辑: 余蓉

## Heterogeneous Task Cooperative Dispatch Scheduling Architecture of Networked Radar for Saturation Attack Air Defense Early Warning

YE Juhang<sup>①</sup> FANG Yuyuan<sup>②</sup> WEI Shaopeng<sup>③</sup> DUAN Jia<sup>①</sup> ZHANG Lei<sup>①</sup>

<sup>①</sup>(School of Electronics and Communication Engineering, Sun Yat-sen University, Shenzhen 518107, China)

<sup>②</sup>(School of Information and Technology, Southwest Jiaotong University, Chengtu 611756, China)

<sup>③</sup>(School of Oceanography and Information, China University of Petroleum, Qingdao 266580, China)

### Abstract:

**Objective** To address the severe challenges of UAV swarm and intelligent loitering missile saturation attacks bringing massive, sudden, and heterogeneous early warning tasks, designing a scalable networked-radar-based collaborative scheduling architecture is critical for maximizing air defense early warning potential. Overcoming existing limitations like rigid dynamic collaboration and inadequate heterogeneous task scheduling, this paper

decouples the non-convex collaborative scheduling problem into a multi-stage decision process of multi-dimensional dynamic resource collaboration and adaptive heterogeneous task scheduling, and, by introducing a dispatching mechanism and hierarchical reinforcement learning, proposes a hybrid architecture integrating network-level centralized dynamic target allocation and node-level distributed scheduling with scalability as the design goal. The architecture is specifically realized through an execution-redispaching cognitive closed-loop encompassing environmental perception, network collaboration, node scheduling, task execution, and target re-dispatching, together with a target dispatching algorithm and a hierarchical command-and-scheduling method, resolving multi-radar, multi-target, and multi-task scheduling dilemmas in saturation scenarios.

**Methods** Collaborative scheduling is initially decoupled into dispatch-centric network-level multi-dimensional resource collaboration and hierarchical-command-centric adaptive node-level scheduling. The environmental perception layer establishes a fidelity model utilizing the Bayesian Cramér-Rao Lower Bound (BCRLB) and radar detection probability to jointly evaluate target threats and radar states. Subsequently, the network collaboration layer references a dispatching mechanism, dynamically allocating targets via an adaptively weighted dispatching model based on quantified combat effectiveness and constructing the scalable execution-redispaching cognitive closed-loop. To solve the ensuing large-scale, nonlinear, multi-constraint generalized bipartite graph matching problem, the proposed Target Dispatch Algorithm (TDA), an MMAS variant incorporating feasibility rule constraint handling, provides a constructive solution. The node scheduling layer then references Hierarchical Q-Learning (HQL), utilizing a serial dual Q-table engineering implementation for distributed heterogeneous task scheduling. Addressing the massive multi-constraint job-shop scheduling problem, a hierarchical mechanism uses task proportion control as a sub-goal, translating upper-level operational intent into specific lower-level beam dwells for the adaptive, long-term, and interpretable execution of intertwined heterogeneous tasks.

**Results and Discussions** A simulated point-defense scenario against swarm or missile saturation attacks was designed, deploying three networked isomorphic S-band medium-range phased array radars against 200 high-speed maneuvering targets. For network collaboration, the proposed Target Dispatch Algorithm (TDA) employs a hierarchical solving mechanism with feasibility rule constraints, replacing traditional penalty functions. Leveraging model prior information, TDA demonstrates superior scheme quality and convergence speed compared to the Max-Min Ant System (MMAS), Artificial Bee Colony (ABC), and Genetic Algorithm (GA) (Fig.3). Despite increased computational complexity, the calculation time of TDA remains strictly within the dispatching cycle, enhancing scheme quality at a millisecond-level cost while ensuring wartime real-time compliance (Fig.4). For node scheduling, the proposed Hierarchical Q-Learning (HQL) guarantees policy consistency via macro/micro-step decisions. Aided by an internal dense transfer reward environment, HQL demonstrates superior long-term policy quality and learning efficiency compared to the Q-Learning (QL) and the Priority-Based Method (PBM) (Fig.5). Utilizing a hierarchical, serial dual Q-table structure, HQL maintains balanced task proportions, maximizing combat effectiveness while improving resource utilization (Fig.6). Further, target loss rate and mean tracking error confirm that HQL achieves the lowest mean tracking error by concentrating resources on high-quality tracking tasks, despite a higher target loss rate, validating superior long-term scheduling performance (Fig.7).

**Conclusions** The proposed hybrid architecture integrating network-level centralized dynamic target allocation and node-level distributed heterogeneous task scheduling effectively counters swarm saturation attacks by overcoming limitations of dynamic coordination rigidity and inadequate heterogeneous task scheduling capability. By dynamically coordinating networked radars in real-time for scheduling of search, confirmation, and tracking tasks during large-scale, high-speed saturation attacks, the architecture releases the defensive potential of early warning network. Simulations validate its improvement in applicable processing scale, environmental adaptability, long-term scheduling capability, scalability, and interpretability. Future research will explore online learning to mitigate offline-training-online-deployment model distortion, or extend the architecture to integrate weapon-target assignment for unified early warning and fire control defense systems.

**Key words:** Saturation attack; Networked radar; Target allocation; Task scheduling; Dispatch mechanism; Hierarchical reinforcement learning