

姚昌华<sup>1</sup> 安蕾<sup>1</sup> 刘鑫<sup>2</sup> 韩贵真<sup>1</sup> 高泽邻<sup>1</sup>

# 面向多点侦察和通信服务的多无人机协同任务分配

## 摘要

针对多无人机对多个异构任务目标进行侦察和通信服务的协同优化问题,通过考虑不同目标的任务要求和价值,以及多机协同增益与任务行为制约关系,构建斯坦伯格博弈模型,将上层无人机建立为博弈领导者,下层无人机建立为博弈的跟随者,并提出一种分布式策略更新迭代算法,实现了多无人机任务分配方案的稳定收敛以及系统任务收益优化。仿真结果显示,所提方法能有效提升多无人机系统同时完成多个任务的效益,并能在不同环境下实现面向异构任务价值的高效协同。

## 关键词

多无人机系统;任务分配;斯坦伯格博弈;迭代算法

中图分类号 TN929.5;V279

文献标志码 A

收稿日期 2021-08-16

资助项目 国家自然科学基金(61971439,61961010);江苏省自然科学基金(BK20191329);中国博士后科学基金(2019T120987);南京信息工程大学人才启动经费(2020r100)

## 作者简介

姚昌华,男,博士,教授,研究方向为智能无人集群、智能无线通信.ych2347@163.com

安蕾(通信作者),女,硕士生,研究方向为智能无人集群.1178535838@qq.com

1 南京信息工程大学 电子与信息工程学院,南京,210044

2 桂林理工大学 信息科学与工程学院,桂林,541006

## 0 引言

随着人工智能技术的快速发展,无人机的智能化水平也越来越高,数量众多的无人机组成无人机群以其高度的灵活性、广泛的适应性、可控的经济性,拥有越来越广泛的应用潜力。无人机可以对地面目标近距离地实施选择性和针对性的观测和通信<sup>[1]</sup>。多无人机系统具有容错性强、自适应性好等优势,更适合在复杂环境下执行任务<sup>[2]</sup>。

多无人机执行任务时,必须对其进行任务分配,以提高任务执行效率。无人机集群的任务规划是指根据任务需求、自身特性等对目标任务进行综合调度,从而建立无人机与任务目标之间合理的映射协同关系<sup>[3-4]</sup>。对于异构主体之间的协同问题,已有了一些相关研究,比如:文献[5]采用多架无人机协同辅助实施任务分配和路径优化的分层优化方案,能够解决降低时间和能耗的优化问题;文献[6]针对多无人机作战飞机协同任务分配问题建立了一种扩展的多目标整数规划模型,并采用改进的量子粒子群算法求解最优方案;文献[7]针对任务随机下发场景中由于任务完成时间约束带来的任务完成度低的问题,通过强化学习方法进行无人机的行为决策,达到新任务与正在执行任务的动态分配以提高任务完成度;文献[8]考虑飞行航程和任务分配均衡性,在自适应遗传算法运行过程中对交叉率和变异率进行实时动态调整,以克服标准遗传算法陷入局部最优的缺点;文献[9]针对多目标跟踪任务分配中传感器之间竞争与合作的关系,提出了一种基于博弈理论的多目标跟踪任务传感器资源分配方法;文献[10]考虑在通信带宽有限条件下多无人机组队的任务分配问题,通过协调侦察、执行和评估任务,改进基于一致性的竞拍算法来减少通信负担;文献[11]基于相邻局部通信的分布式拍卖算法,实现了多无人机任务协同分配的优化求解问题;文献[12]采用多无人机辅助移动边缘计算系统联合优化一个有限周期内的无人机轨迹和用户调度;文献[13]考虑任务分配中联盟的构建和无人机资源管理方法,使联盟中各无人机能够以更加平衡的方式消耗资源,提升系统性能;文献[14]综合考虑无人机的物理性能约束,应用基于模拟退火的混合粒子群算法进行任务分配求解;文献[15]建立多任务的分配问题模型,采用多余负载竞拍方案减少非法劣解,通过实数编码建立粒子和实际分配方案之间的映射关系,解决实际分配问题;文献[16]研究了MEC辅助无人机群中所有成员的总延迟的最小化问题,提出卸载模

型,缩短了任务的完成时间;文献[17]提出一种动态分散任务分配算法,用于任务分配问题中在线新任务的分配。

上述文献所研究的任务目标大多是同构的<sup>[9,12,16-17]</sup>,未考虑异构的任务价值<sup>[10-11,13]</sup>,也未考虑同时存在多点侦察任务和通信服务.从方法上看,现有多数研究基于集中式分配算法<sup>[5-8,14-15]</sup>,即需要一个中心控制实体来为集群内的所有成员分配任务.这种模式不利于提高无人机集群的鲁棒性和环境适应能力.在无人机集群执行任务过程中,大部分的环境状态都是动态变化的,其任务分配的方案也应该随时优化调整,以有效应对动态环境的变化.集中式任务分配算法存在计算复杂度高、依赖中心节点等问题.研究自适应强的多无人机分布式动态任务分配方法,是多无人机协同任务分配的现实需要,也是难点问题。

本文研究多无人机网络中的任务目标调度问题,构建面向异构任务类型和价值的斯坦伯格(Stackelberg)博弈模型,设计通信和侦察的任务效用函数,并提出分层用户侦察和通信任务调度以及功率控制算法,实现基于无人机自主任务选择的多无人机系统任务分配稳定收敛以及系统任务总收益优化。

## 1 系统模型和问题建模

### 1.1 系统模型

无人机通信侦察任务分配系统模型如图1所示,由一个领头无人机(Leader Drone, LD)和随机分

布在其周围的  $N$  个协同无人机(Collaborative Drone, CD)构成,共同完成多个任务.每一架无人机监测或服务范围内均有随机分布的不同数量的通信和侦察两类任务.每个任务的重要程度或属性不同,其对应的任务价值也不同.各架无人机根据自身位置与目标之间的距离、目标价值、侦察或通信服务所获得的期望收益,以及与其他无人机通信链路之间可能的干扰关系,来自主决定选择任务对象。

令  $0$  表示领头无人机,则分布在周围的协同无人机集合表示为  $A = \{CD_1, CD_2, \dots, CD_N\}$ ,领头及协同无人机可调度通信任务目标集合表示为  $T_{u,i} = \{1, 2, \dots, m\}, i \in A \cup \{0\}$ ,可调度的侦察任务目标集合表示为  $T_{z,i} = \{1, 2, \dots, n\}$ .信道增益为  $g_{i,j}, j \in T_{u,i} \cup T_{z,i} \cup \{0\}$ ,并且假定信道增益在任务目标调度和功率调整时期稳定不变.LD 和 CD 的发射功率向量为  $\mathbf{p} = [p_0, p_1, \dots, p_N]$ ,背景干扰噪声功率为  $\sigma^2$ .CD 执行通信侦察任务后,需要将信息上传汇报给领头无人机.侦察任务中,每一架无人机对各个任务目标对象的分辨率  $r$  为定值,构建分辨率矩阵.通信任务调度中,当给定 LD 以及其他 CD 的调度策略后,  $CD_i$  服务第  $k$  个通信任务目标的下行信噪比为

$$\gamma_i^k(p_i, \mathbf{p}_{-i}) = \frac{p_i g_{i,k}}{I_i^k(\mathbf{p}_{-i})}, \quad (1)$$

通信信息上传时的信噪比为

$$\tilde{\gamma}_i^k(p_i, \mathbf{p}_{-i}) = \frac{p_{i,0} g_{i,0}}{I_i^k(\mathbf{p}_{-i,0})}, \quad (2)$$

其中  $I_i^k(\mathbf{p}_{-i}) = p_0 g_{0,k} + \sum_{j \neq i, j \in A} p_j g_{j,k} + \sigma^2, I_i^k(\mathbf{p}_{-i,0}) =$

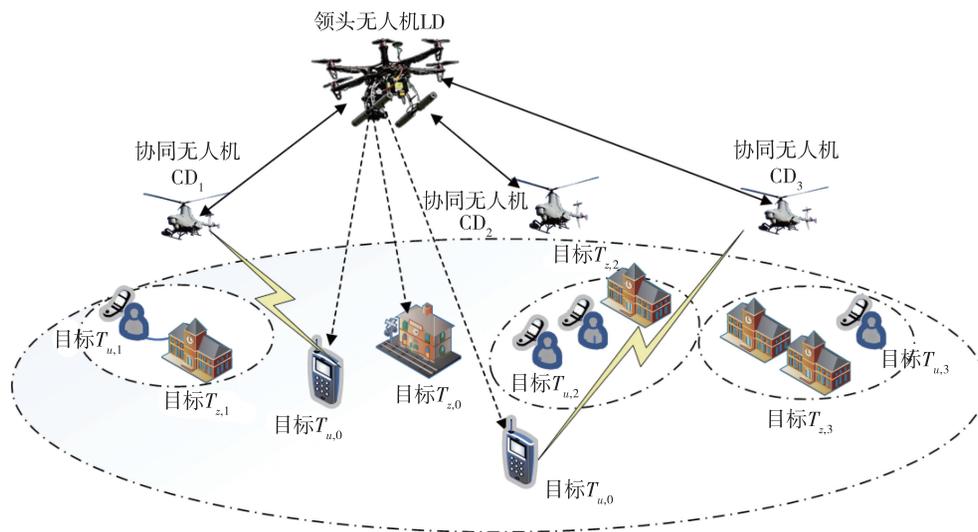


图1 无人机通信侦察任务分配系统模型

Fig. 1 Task assignment of multi-UAV communication reconnaissance

$\sum_{j \neq i, j \in A} p_{j,0} g_{j,0} + \sigma^2, \mathbf{p}_{-i} = [p_0, p_1, \dots, p_{i-1}, p_{i+1}, \dots, p_N]$  表示除  $CD_i$  以外的所有无人机的功率分配向量。 $CD_i$  下的通信任务目标同时收到来自邻居 CD 的同层干扰以及来自 LD 的跨层干扰。 $p_{i,0}$  为通信上传功率值, 假定 CD 到 LD 通信上传速率为  $R_i$ , 分配带宽为  $\bar{B}_i$ , 由  $R_i = \bar{B}_i \log_2 \left( 1 + \frac{p_{i,0} g_{i,0}}{\sigma^2} \right)$  可求得  $p_{i,0}$ 。另外, LD 服务的第  $l$  个通信任务目标的下行信噪比可以表示为

$$\gamma_0^l(p_0, \mathbf{p}_{-0}) = \frac{p_0 g_{0,l}}{I_0^l(\mathbf{p}_{-0})} = \frac{p_0 g_{0,l}}{\sum_{j \in A} p_j g_{j,l} + \sigma^2}. \quad (3)$$

## 1.2 斯坦伯格分层博弈模型

在本文多无人机系统协同模型中, 任务类型包含通信任务和侦察任务两类。因目标任务的重要程度不同, 需要对 LD 和 CD 任务执行进行合理规划, 任务目标重要程度较高的服务质量 (QoS) 需要首先得到保障。本文基于任务优先级以及任务目标需求的差异性, 采用分层博弈模型来刻画领头无人机和协同无人机之间的任务目标调度和功率分配问题。在该博弈中, 先做出决策的一方为领导者 (leader), 其余观测领导者的决策从而做出行动的一方称为跟随者 (follower)。本文将上层领头无人机视为领导者, 下层协同无人机视为跟随者, 利用分层斯坦伯格 (Stackelberg) 博弈模型刻画 LD-CD 之间的分层竞争交互关系。目标任务调度中, Stackelberg 定义为  $\mathcal{G} = \{A, \{P_0, C_0\}, \{P_i, C_i\}_{i \in A}, \{U_0\}, \{U_i\}_{i \in A}\}$ , (4) 其中  $\{P_0, C_0\}$  和  $\{P_i, C_i\}_{i \in A}$  分别表示 LD 和 CD 的策略空间,  $\{U_0\}$  和  $\{U_i\}_{i \in A}$  分别表示 LD 和 CD 执行目标任务效用函数。

对于执行通信服务任务的无人机, 无人机效用函数的设计同时考虑了目标任务的满意度和功率消耗。对于给定的通信目标  $k$ , LD 的效用函数可以表示为

$$\mathcal{U}_0^k(p_0, \mathbf{p}_{-0}) = U_0^k(p_0, \mathbf{p}_{-0}) - C_0^k(p_0, \mathbf{p}_{-0}) = \frac{1}{1 + \exp(-\alpha_0^k(\gamma_0^k - \beta_0^k))} \cdot v_k - \mu_0 p_0. \quad (5)$$

该效用函数包含两部分: 第一部分为服务通信任务目标对象的收益  $U_0^k(p_0, \mathbf{p}_{-0})$ , 被建模为 S 型函数, 代表所服务通信目标任务的满意度; 第二部分是代价函数  $C_0^k(p_0, \mathbf{p}_{-0})$ , 表示动态的功率开销, 其中的参数  $\alpha_0$  和  $\beta_0$  分别为 S 型函数的陡度和中心值。 $v_k$  代表通信任务目标  $k$  的价值。 $\mu_0$  为常数, 用来权衡服务目标的满意度和功率能量消耗。当给定  $CD_i$  服务第  $k$

个通信任务目标时, 其效用函数可以表示为

$$\mathcal{U}_i^k(p_i, \mathbf{p}_{-i}) = U_i^k(p_i, \mathbf{p}_{-i}) - C_i^k(p_i, \mathbf{p}_{-i}) = \frac{1}{1 + \exp(-\alpha_i^k(\gamma_i^k + \theta \bar{\gamma}_i^k - \beta_i^k))} \cdot v_k - \mu_i p_i - \lambda_i g_{i,0} p_i - \kappa_i p_{i,0}, \quad (6)$$

其中,  $CD_i$  的收益函数部分同时考虑了服务通信任务目标对象的满意度和通信上传的满意度,  $\theta$  为常数, 用于折中通信下行信噪比和上传信噪比。此外  $\alpha_i$  和  $\beta_i$  分别为函数陡度和中心值。 $CD_i$  代价函数部分同时考虑了执行目标任务的功率消耗、上传通信信息的功率消耗和下层  $CD_i$  对上层 LD 通信服务的干扰惩罚。 $\kappa_i$  表示上传通信信息功率消耗系数,  $\lambda_i$  表示干扰惩罚参数, 用于调节跨层干扰对上层服务目标的影响。当  $CD_i$  增加发射功率时, 服务任务目标对象的满意度增加, 同时将会对上层 LD 带来更高的跨层干扰, 影响 LD 服务任务目标的 QoS, 因此  $CD_i$  需要进行折中优化。

对于执行侦察服务任务的无人机, 无人机侦察效用函数的设计同样包括目标任务的满意度和功率消耗两部分。侦察目标任务调度中, 每一架无人机对各个任务目标的分辨率为定值, 构建分辨率矩阵。给定 LD 服务侦察任务目标时, 其侦察效用函数可以表示为  $\mathcal{U}_0^x = U_0^x - C_0^x =$

$$\frac{v_x}{1 + \exp\left(-\alpha_0^x \left(\frac{r_0^x}{d_0^x} \cdot p_0' - \beta_0^x\right)\right)} - \delta_0 p_0'. \quad (7)$$

该效用函数包括两个部分,  $U_0^x$  表示服务侦察任务目标的收益,  $C_0^x$  表示服务侦察目标的代价, 即 LD 图像识别的功率消耗。其中,  $r_0^x$  为 LD 对任务目标的分辨率,  $d_0^x$  为 LD 与任务目标距离, 侦察收益建模为 S 型函数,  $\delta_0$  为图像识别功率消耗比例常数。当给定  $CD_i$  服务侦察任务目标  $x$  时, 其侦察效用函数可以表示为

$$\mathcal{U}_i^x = U_i^x - C_i^x = \frac{v_x}{1 + \exp\left(-\alpha_i^x \left(\frac{r_i^x}{d_i^x} \tau_i^x p_i' - \beta_i^x\right)\right)} - \hat{\lambda}_i g_{i,0} (1 - \tau_i) p_i' - \delta_i \tau_i p_i', \quad (8)$$

其中代价函数部分同时考虑了图像上传的功率消耗和 LD 识别的功率消耗,  $p_i'$  为每个  $CD_i$  进行侦察任务的总功率,  $\tau_i$  用于识别计算处理的功率比例,  $1 - \tau_i$  表示识别完成后侦察信息上传功率消耗比例,  $\hat{\lambda}_i$  为侦察信息上传干扰惩罚参数, 用来权衡侦察信息上

传对领头无人机产生的干扰,  $\delta_i$  为常数, 权衡  $CD_i$  用于拍照的功率消耗, 默认  $\delta_0 = \delta_i$ .

## 2 斯坦伯格均衡求解

定义符号  $\Phi_i = \{p_i, c_i\}$ ,  $\Phi_{-i} = \{\Phi_0, \Phi_1, \dots, \Phi_{i-1}, \Phi_{i+1}, \dots, \Phi_N\}$ ,  $\Phi_m = \{\Phi_1, \Phi_2, \dots, \Phi_N\}$ .

**定义 1** (斯坦伯格均衡, Stackelberg Equilibrium, SE)<sup>[18]</sup>  $\Phi_0^*$  表示上层博弈最大化效用函数的最佳相应策略,  $\Phi_m^*$  表示下层博弈的最佳响应策略. 对于任意的策略组合, 均满足以下条件:

$$\mathcal{U}_0(\Phi_0^*, \Phi_m^*) \geq \mathcal{U}_0(\Phi_0, \Phi_m^*), \quad (9)$$

$$\mathcal{U}_i(\Phi_i^*, \Phi_{-i}^*) \geq \mathcal{U}_i(\Phi_i, \Phi_{-i}^*). \quad (10)$$

$(\Phi_0^*, \Phi_m^*)$  称为斯坦伯格均衡. LD 最优策略由下层博弈最佳响应策略给定, 最大化自身效用函数求解. 同理, CD 的最优策略是由给定上层博弈的最佳相应策略, 最大化自身效用函数求解. 通过逆向递推法寻求上下两层子博弈的均衡, 如图 2 所示.

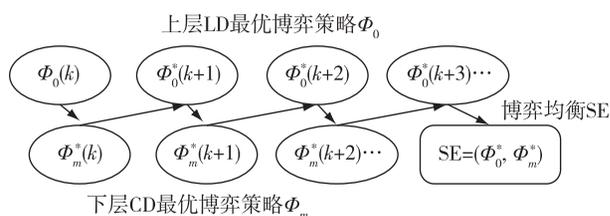


图 2 斯坦伯格均衡求解

Fig. 2 Stackelberg equilibrium solution

### 2.1 下层均衡求解

给定上层 LD 的任务目标选择和发射功率, 每一个 CD 独立地选择最佳策略来最大化自身效用函数, 因此, 下层子博弈定义为

$$\mathcal{G} = \{A, \{\Phi_i\}_{i \in A}, \{U_i\}_{i \in A}\}. \quad (11)$$

**定理 1** 给定其他无人机的策略  $\Phi_{-i}$ ,  $CD_i$  最优的通信任务目标选择:

$$t_i^* = \arg \max \left[ \left( \frac{g_{i,k}}{I_i^k(\mathbf{p}_{-i})} + \theta \frac{\varepsilon_i g_{i,0}}{I_i^k(\mathbf{p}_{-i,0})} \right) v_k \right]. \quad (12)$$

**证明** 当给定其他无人机策略  $\Phi_{-i}$ , 令  $p_{i,0} = \varepsilon_i p_i$ ,  $A_i = \frac{g_{i,k}}{I_i^k(\mathbf{p}_{-i})} + \theta \frac{\varepsilon_i g_{i,0}}{I_i^k(\mathbf{p}_{-i,0})}$ ,  $CD_i$  服务于任意两个通信任务目标  $k$  和  $l$ , 假定  $A_i^k \geq A_i^l, v_k \geq v_l$ , 显然,  $\gamma_i^k + \theta \tilde{\gamma}_i^k \geq \gamma_i^l + \theta \tilde{\gamma}_i^l$ , 然后计算服务不同任务目标的效用差值为

$$\mathcal{U}_i^k - \mathcal{U}_i^l = U_i^k - U_i^l = \frac{v_k}{1 + \exp(-\alpha_i(\gamma_i^k + \theta \tilde{\gamma}_i^k - \beta_i))} -$$

$$\frac{v_l}{1 + \exp(-\alpha_i(\gamma_i^l + \theta \tilde{\gamma}_i^l - \beta_i))} \geq 0. \quad (13)$$

因此,  $CD_i$  最优通信任务目标  $t_i^* = \arg \max \left[ \left( \frac{g_{i,k}}{I_i^k(\mathbf{p}_{-i})} + \theta \frac{\varepsilon_i g_{i,0}}{I_i^k(\mathbf{p}_{-i,0})} \right) v_k \right]$ , 定理 1 得证.

$CD_i$  通过定理 1 确定最优通信任务目标, 然后进一步优化发射功率最大化效用函数. 为方便求解, 令  $\Gamma_i = \gamma_i^k + \theta \tilde{\gamma}_i^k, p_{i,0}^k = \varepsilon_i p_i^k$ , 其  $\mathcal{U}_i^k$  对  $p_i$  求偏导得:

$$\frac{d\mathcal{U}_i^k}{dp_i} = \mathcal{U}_i^k{}'(\Gamma_i) \frac{d(\Gamma_i)}{dp_i} - \frac{dC_i}{dp_i}. \quad (14)$$

令  $\frac{d\mathcal{U}_i^k}{dp_i} = 0$ , 即得:

$$\mathcal{U}_i^k{}'(\Gamma_i) \frac{d(\Gamma_i)}{dp_i} = \frac{dC_i}{dp_i}. \quad (15)$$

令  $\frac{g_{i,i}}{I_i^i(\Phi_{-i})} + \theta \frac{\varepsilon_i g_{i,0}}{I_i^i(\Phi_{-i,0})} = \frac{1}{B_i}$ , 即得:

$$\mathcal{U}_i^k{}'(\Gamma_i) = (\mu_i + \kappa_i \varepsilon_i + \lambda_i g_{i,0}) \times B_i. \quad (16)$$

令  $\Psi(\Gamma_i) = \mathcal{U}_i^k{}'(\Gamma_i)$ , 可得:

$$\Gamma_i = \Psi^{-1}[\mathcal{U}_i^k{}'(\Gamma_i)] = \Psi^{-1}[(\mu_i + \kappa_i \varepsilon_i + \lambda_i g_{i,0}) \times B_i]. \quad (17)$$

S 型函数倒数满足以下关系:

$$\mathcal{U}_i^k{}'(\Gamma_i) = \alpha_i \mathcal{U}_i^k(\Gamma_i) [1 - \mathcal{U}_i^k(\Gamma_i)]. \quad (18)$$

根据式 (16) 和式 (18) 可得:

$$p_i^* = B_i \beta_i - \frac{B_i}{\alpha_i} \times \ln \left[ \frac{A_i}{2} - 1 - \sqrt{\left( \frac{A_i}{2} - 1 \right)^2 - 1} \right], \quad (19)$$

其中  $A_i = \frac{\alpha_i \times v_i}{(\mu_i + \kappa_i \varepsilon_i + \lambda_i g_{i,0}) \times B_i}$ . 下层调度通信任务目标和上传信息总信噪比和为

$$\Gamma_i = \beta_i - \frac{1}{\alpha_i} \times \ln \left[ \frac{A_i}{2} - 1 - \sqrt{\left( \frac{A_i}{2} - 1 \right)^2 - 1} \right]. \quad (20)$$

CD 的通信效用函数即可转化为

$$\mathcal{U}_i^k(\Gamma_i) = \frac{v_i}{1 + \exp(-\alpha_i(\Gamma_i - \beta_i))} - (\mu_i + \kappa_i \varepsilon_i + \lambda_i g_{i,0}) \cdot B_i \cdot \Gamma_i. \quad (21)$$

通过式 (18) 和代价直线是一条过原点的切线<sup>[19]</sup> 可得:

$$\mathcal{U}_i^k(\Gamma_i) \Gamma_i = \mathcal{U}_i^k(\Gamma_i). \quad (22)$$

经过运算参数  $\beta_i$  可以设置如下:

$$\beta_i = \Gamma_i - \frac{\ln(\alpha_i \Gamma_i - 1)}{\alpha_i}. \quad (23)$$

同理, 令  $p_i^l = p_i$ , 侦察效用值由式 (8) 求得, 比较

所有侦察任务目标效用值,选择最优侦察任务目标  $x_i^*$ .若效用值为负,此时  $p_i^* = 0$ ,选择放弃该任务目标的侦察任务.分析通信侦察效用值,确定最优的任务目标选择为  $c_i^*$ .假设对于任意的  $i, B_i \cdot \Psi^{-1}[(\mu_i + \kappa_i \varepsilon_i + \lambda_i g_{i,0}) \times B_i]$  在区间  $[L_i, H_i]$  上为增函数,其中  $\frac{g_{i,t_i}}{I_i^i(\Phi_{-i})} + \theta \frac{\varepsilon_i g_{i,0}}{I_i^i(\Phi_{-i,0})} = \frac{1}{B_i}, L_i = \min[L_i^1, L_i^2, \dots, L_i^m], H_i = \max[H_i^1, H_i^2, \dots, H_i^m]$ .下层迭代中 CD 的功率控制形式为  $p_i^{t_i^*}(k+1) = \min[p_i^1(k+1), p_i^2(k+1), \dots, p_i^m(k+1)]$ ,其中  $p_i^{t_i^*}(k+1) = F_i^{t_i^*}(p(k))$ .令  $F_i(p(k)) = \min[F_i^1(p(k)), F_i^2(p(k)), \dots, F_i^m(p(k))]$ ,最终功率更新迭代和服务任务目标迭代策略为

$$p_i^*(k+1) = F_i(p(k)), \quad (24)$$

$$c_i^*(k+1) = \arg \max(\mathcal{U}_i^{i^*}, \mathcal{U}_i^{i^*}). \quad (25)$$

下层博弈中,上述任务目标选择和策略更新迭代过程最终收敛在唯一的纳什均衡点.下面引入标准干扰函数相关内容辅助证明.

**定义 2** (唯一性证明)<sup>[20]</sup> 如果函数  $F(p)$  满足以下条件,则该函数为标准干扰函数:

- 1) 非负性:  $F(p) > 0$ ;
- 2) 单调性:对于任意  $p^1 > p^2$ ,则  $F(p^1) > F(p^2)$ ;
- 3) 伸缩性:对于任意  $\tau > 1$ ,则  $\tau F(p) > F(\tau p)$ .

如果函数  $F^1(p)$  和  $F^2(p)$  为标准干扰函数,其组合函数也必然满足标准函数特性;如果博弈参与者最佳响应函数为标准干扰函数,则该博弈必然存在唯一的纳什均衡解.

**定理 2** 下层博弈中 CD 服务通信任务目标的功率迭代更新为标准干扰函数,即  $p_i^*(k+1) = F_i(p(k))$ ,将从任意可行的初始值收敛到唯一的纳什均衡点.

**证明** 对于  $CD_i$  中任意一个任务目标  $t_i$ ,即  $p_i^{t_i} = F_i^{t_i}(p)$ ,则有

$$F_i^{t_i}(p) = B_i \cdot \Gamma_i = B_i \cdot \Psi^{-1}[(\mu_i + \kappa_i \varepsilon_i + \lambda_i g_{i,0}) \times B_i],$$

其中  $\frac{g_{i,t_i}}{I_i^i(\Phi_{-i})} + \theta \frac{\varepsilon_i g_{i,0}}{I_i^i(\Phi_{-i,0})} = \frac{1}{B_i}$ .下层 CD 策略调整时,由于上层 LD 的策略是固定不变的,下层其他 CD 的策略也不会影响当前 CD 策略,只受其余 CD 发射功率的影响,因此  $I_i^i(\Phi_{-i}) = I_i^i(p_{-i})$ .

1) 非负性:由于背景噪声非零,因此  $B_i > 0$ ,非负性满足条件.

2) 单调性:如果  $p^1 > p^2$ ,则  $I_i^i(p_{-i}^1) > I_i^i(p_{-i}^2)$ ,  $I_i^i(p_{-i,0}^1) > I_i^i(p_{-i,0}^2)$ ,因为函数  $B_i \cdot \Psi^{-1}[(\mu_i +$

$\kappa_i \varepsilon_i + \lambda_i g_{i,0}) \times B_i]$  在区间  $[L_i, H_i]$  上为增函数,其中  $L_i = \min[L_i^1, L_i^2, \dots, L_i^m], H_i = \max[H_i^1, H_i^2, \dots, H_i^m]$ ,因此  $F_i^{t_i}(p^1) > F_i^{t_i}(p^2)$ ,单调性满足条件.

3) 伸缩性:对于任意  $\tau > 1$ ,则有

$$F_i^m(\tau p) - \tau F_i^m(p) = \frac{1}{\eta_1} \cdot \Psi^{-1}\left[\omega \times \frac{1}{\eta_1}\right] - \tau \frac{1}{\eta_2} \cdot \Psi^{-1}\left[\omega \times \frac{1}{\eta_2}\right] < \frac{1}{\eta_1} \cdot \Psi^{-1}\left[\omega \times \frac{1}{\eta_2}\right] - \tau \frac{1}{\eta_2} \cdot \Psi^{-1}\left[\omega \times \frac{1}{\eta_2}\right] < 0,$$

其中  $\eta_1 = \frac{g_{i,m}}{I_i^m(\tau p_{-i})} + \theta \frac{\varepsilon_i g_{i,0}}{I_i^m(\tau p_{-i,0})}, \eta_2 = \frac{g_{i,m}}{I_i^m(p_{-i})} +$

$\theta \frac{\varepsilon_i g_{i,0}}{I_i^m(p_{-i,0})}, \omega = \mu_i + \kappa_i \varepsilon_i + \lambda_i g_{i,0}$ .因此,伸缩性满足条件.故  $p_i^{t_i^*}(k+1) = F_i^{t_i^*}(p(k))$  为标准干扰函数.假定第  $k+1$  次迭代,服务的通信任务目标为  $t_i^*(k+1)$ ,功率更新迭代可得:

$$p_i^{t_i^*}(k+1) = \min[p_i^1(k+1), p_i^2(k+1), \dots, p_i^m(k+1)] = \min[F_i^1(p(k)), F_i^2(p(k)), \dots, F_i^m(p(k))].$$

因  $F_i^1(p(k)), F_i^2(p(k)), \dots, F_i^m(p(k))$  均为标准干扰函数,由此可得  $p_i^{t_i^*}(k+1) = F_i^{t_i^*}(p(k))$  为标准干扰函数,下层子博弈迭代更新函数必然存在唯一的纳什均衡解,得证.

## 2.2 上层均衡求解

上层子博弈定义为

$$\mathcal{G} = \{\{0\}, \{\Phi_0\}, \{\mathcal{U}_0\}\}, \quad (26)$$

则其 LD 最优的通信任务目标选择为

$$t_0^* = \arg \max \left[ \left( \frac{g_{0,k}}{I_0^k(p_{-0})} \right) \cdot v_k \right]. \quad (27)$$

同理,进一步优化发射功率最大化效用函数可得:

$$p_0^* = B_0 \gamma_0 = B_0 \beta_0 - \frac{B_0}{\alpha_0} \times \ln \left[ \frac{A_0}{2} - 1 - \sqrt{\left( \frac{A_0}{2} - 1 \right)^2 - 1} \right], \quad (28)$$

$$\mathcal{U}_0^0(\Gamma_0) = \frac{v_{t_0}}{1 + \exp(-\alpha_0(\gamma_0 - \beta_0))} - \mu_0 \cdot B_0 \cdot \gamma_0, \quad (29)$$

其中  $B_0 = \frac{I_0^k(p_{-0})}{g_{0,k}}, A_0 = \frac{\alpha_0 \cdot v_{t_0}}{\mu_0 \cdot B_0}$ .上层侦察任务目标效用值由式(7)求得,比较所有侦察任务效用值,选择最优侦察任务目标  $x_0^*$ .类似于下层子博弈的迭代更新,LD 的功率迭代更新为

$$p_0^{\delta}(k+1) = \min[p_0^1(k+1), p_0^2(k+1), \dots, p_0^m(k+1)] = \min[F_0^1(\mathbf{p}(k)), F_0^2(\mathbf{p}(k)), \dots, F_0^m(\mathbf{p}(k))]. \quad (30)$$

因此,LD 服务任务目标和功率更新策略为

$$p_0^*(k+1) = F_0(\mathbf{p}(k)), \quad (31)$$

$$c_0^*(k+1) = \arg \max(\mathcal{U}_0^{\delta}, \mathcal{U}_0^{\delta^*}). \quad (32)$$

### 3 算法流程

子博弈循环采用一般迭代算法求解,达到斯坦伯格均衡迭代结束,上下层目标任务分配不再改变.具体流程如图3所示.

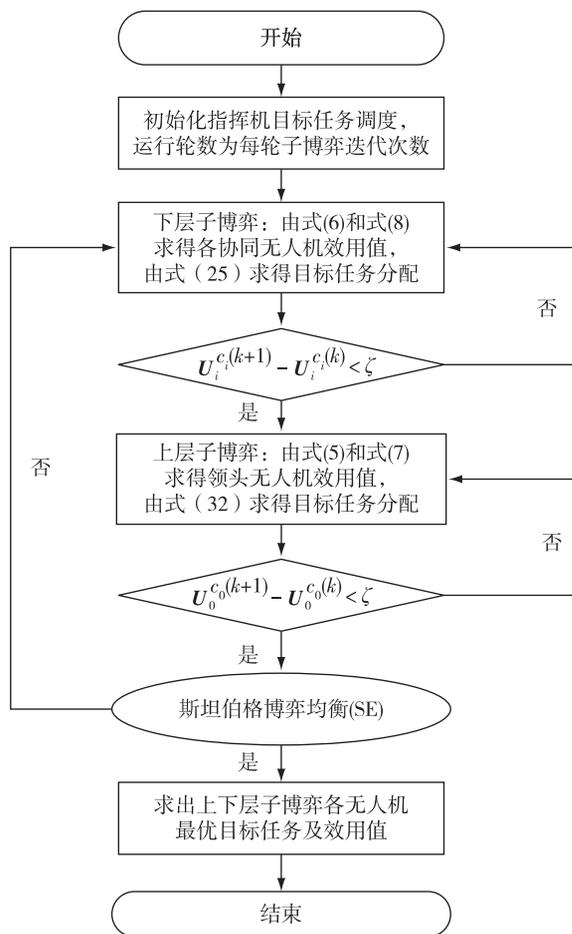


图3 算法流程

Fig.3 Algorithm flowchart

### 4 仿真分析

对于无人机位置、无人机所服务的通信和侦察任务目标个数、通信和侦察任务目标价值等信息构建场景进行仿真分析,同时通过调整无人机位置,设定不同场景对上下层博弈交互进行迭代更新,验证算法的收敛性.

#### 4.1 参数设置

场景设置参数如下:LD 可服务任务目标区域半径为 500 m,10 个 CD 随机分布在 LD 调度范围内,其可服务任务目标半径为 80 m,通信和侦察任务随机分布在 LD 和 CD 服务范围内.LD 可服务的通信任务和侦察任务目标均为 3 个. CD 可服务的通信和侦察任务目标个数分别依次为 4、5、4、5、4、4、5、3、4、3 和 3、4、2、4、3、2、4、1、2、2.LD 所服务的通信和侦察任务目标价值  $v=1$ ,CD 所服务的通信和侦察任务目标价值相对较低,取值在  $[0.9, 0.95]$  内随机生成.其中  $CD_i$  到任务目标  $j$  的信道增益  $g_{i,j} = d_{i,j}^{-2}$ ,  $d_{i,j}^{-2}$  表示对应的距离,信号衰减为 25 dB.LD 所服务通信任务目标信噪比为  $\gamma_0 = 30$  dB,CD 所服务通信任务目标的信噪比和上传信噪比均在  $[10, 20]$  dB 内随机生成.噪声功率  $\sigma^2 = 10^{-8}$  mW.参数  $\alpha_i = 0.2, \theta = 1, \beta_i$  由式(23)确定.通信干扰惩罚和干扰代价参数设置为  $\lambda_i = 10^8, \mu_i = 1$ ,上传功率消耗参数为  $\kappa_i = 1$ .侦察任务中 LD 识别图像功率消耗  $\delta_0 = 1$ ,CD 识别图像功率消耗  $\delta_i = 1$ ,上传干扰惩罚和上传功率比例参数设置为  $\hat{\lambda}_i = 10^8, \tau_i = 0.6$ .表 1 中给出了 LD 和 CD 对侦察任务目标对象的分辨率大小.上述参数在后面仿真中均保持不变.图 4 中给出两种无人机不同位置下的仿真场景.

表 1 LD 和 CD 对侦察目标分辨率

Table 1 Resolutions of LD and CDs to reconnaissance targets

UAV	目标 1	目标 2	目标 3	目标 4
LD	1 920×1 080	1 680×1 050	1 280×800	
CD <sub>1</sub>	1 280×960	1 400×1 050	1 920×1 080	
CD <sub>2</sub>	1 600×900	1 360×768	1 400×1 050	1 920×1 080
CD <sub>3</sub>	1 280×960	1 400×1 050		
CD <sub>4</sub>	1 360×768	1 200×800	1 680×1 050	1 920×1 080
CD <sub>5</sub>	1 600×900	1 360×768	1 280×960	
CD <sub>6</sub>	1 600×900	1 920×1 080		
CD <sub>7</sub>	1 600×900	1 400×1 050	1 920×1 080	1 280×960
CD <sub>8</sub>	1 920×1 080			
CD <sub>9</sub>	1 280×960	1 280×800		
CD <sub>10</sub>	1 280×800	1 920×1 080		

#### 4.2 结果分析

图 5 给出了场景 1 和场景 2 中相应的 LD 和 CD 的效用迭代更新曲线,每一轮迭代 15 次,共设置 20 轮,横坐标为迭代轮数,纵坐标为各无人机效用值.给定上层的目标任务选择后,下层为满足通信任务目标的最小信噪比需求进而优化最小的发射功率,

在每一轮的 15 次迭代的过程中,由于功率在增大,各协同无人机之间的同层干扰逐渐增大,需要不断地增大功率满足信噪比需求,不断迭代直至下层稳定.上层迭代过程中,由于协同无人机的功率在增大,跨层干扰在增大,领头无人机需要增大发射功率满足自身的信噪比需求,直至上层稳定.结合图 6 中场景 1 和场景 2 中相应的 LD 和 CD 的各轮目标的迭代变化曲线,分析得出效用曲线中的转折点为无

人机依据效用值的变化选择侦察任务目标.从效用的更新曲线来看,上下层博弈交互迭代后最终均能够达到收敛状态,验证了所提算法的收敛性能.

表 2 中给出了上下层子博弈达到斯坦伯格均衡时无人机最优通信和侦察任务分配结果.各无人机能够自主进行服务通信或侦察目标优化分配.其中,  $T_{u,0}$  为领头无人机通信任务目标,  $T_{z,0}$  为领头无人机侦察任务目标,  $T_{u,1} \sim T_{u,10}$  为各协同无人机的通信任

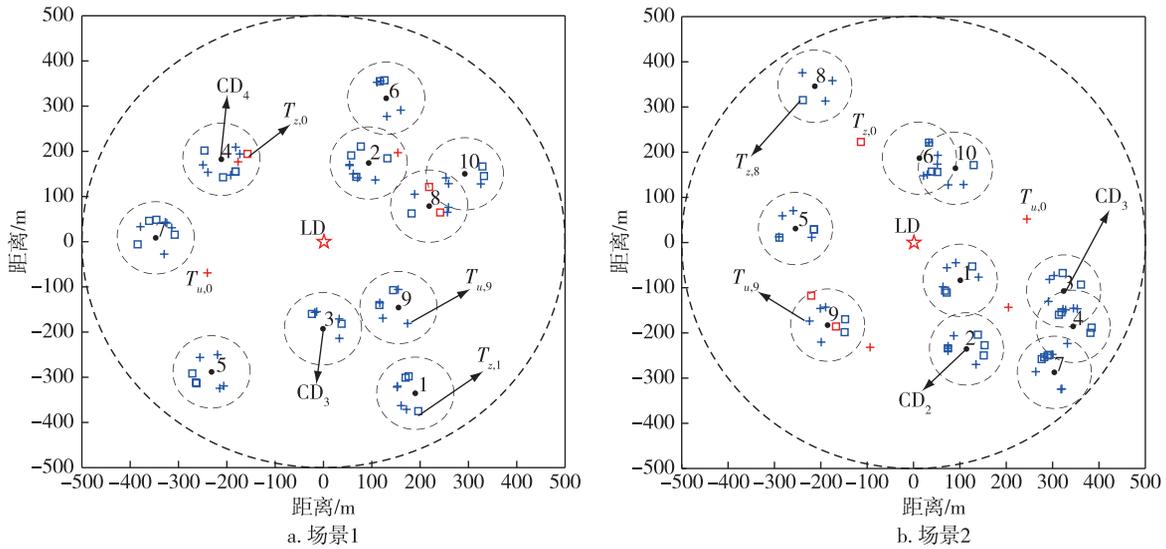


图 4 仿真场景

Fig. 4 Simulation scenes

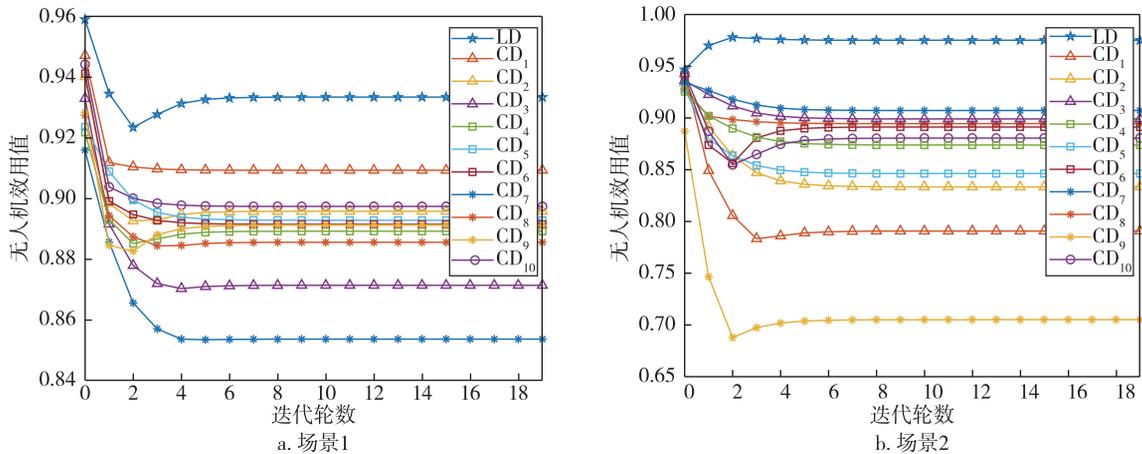


图 5 各无人机效用迭代更新曲线

Fig. 5 The iterative update curves of each UAV utility

表 2 斯坦伯格均衡下无人机通信和侦察任务分配

Table 2 UAV communication and reconnaissance task allocation under SE

场景	LD	CD <sub>1</sub>	CD <sub>2</sub>	CD <sub>3</sub>	CD <sub>4</sub>	CD <sub>5</sub>	CD <sub>6</sub>	CD <sub>7</sub>	CD <sub>8</sub>	CD <sub>9</sub>	CD <sub>10</sub>
1	$T_{u,0}=1$	$T_{z,1}=3$	$T_{u,2}=2$	$T_{u,3}=3$	$T_{u,4}=4$	$T_{u,5}=3$	$T_{z,6}=2$	$T_{u,7}=4$	$T_{u,8}=2$	$T_{z,9}=1$	$T_{z,10}=3$
2	$T_{u,0}=1$	$T_{u,1}=1$	$T_{z,2}=3$	$T_{z,3}=2$	$T_{z,4}=4$	$T_{u,5}=1$	$T_{u,6}=2$	$T_{z,7}=4$	$T_{z,8}=1$	$T_{z,9}=1$	$T_{u,10}=1$

务目标,  $T_{z,1} \sim T_{z,10}$  为各协同无人机的侦察任务目标 (以  $T_{u,0}=1$  为例, 表示领头无人机均衡下选择 1 号通信任务目标)。

### 4.3 系统效用对比

图 7a 与 7b 中给出了任务目标调度过程中联合考虑通信和侦察任务、只考虑通信任务和只考虑侦察

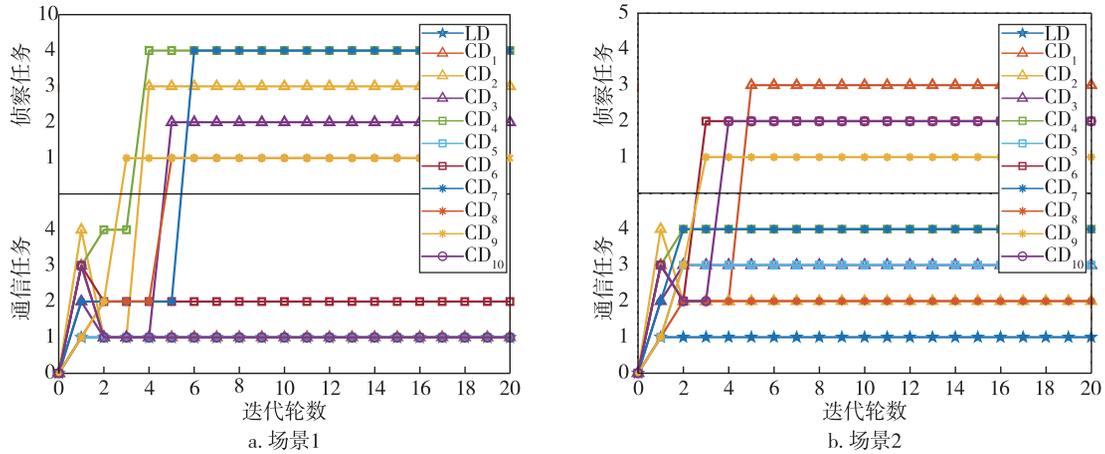


图 6 目标选择随迭代次数变化

Fig. 6 Target selections varied with the number of iterations

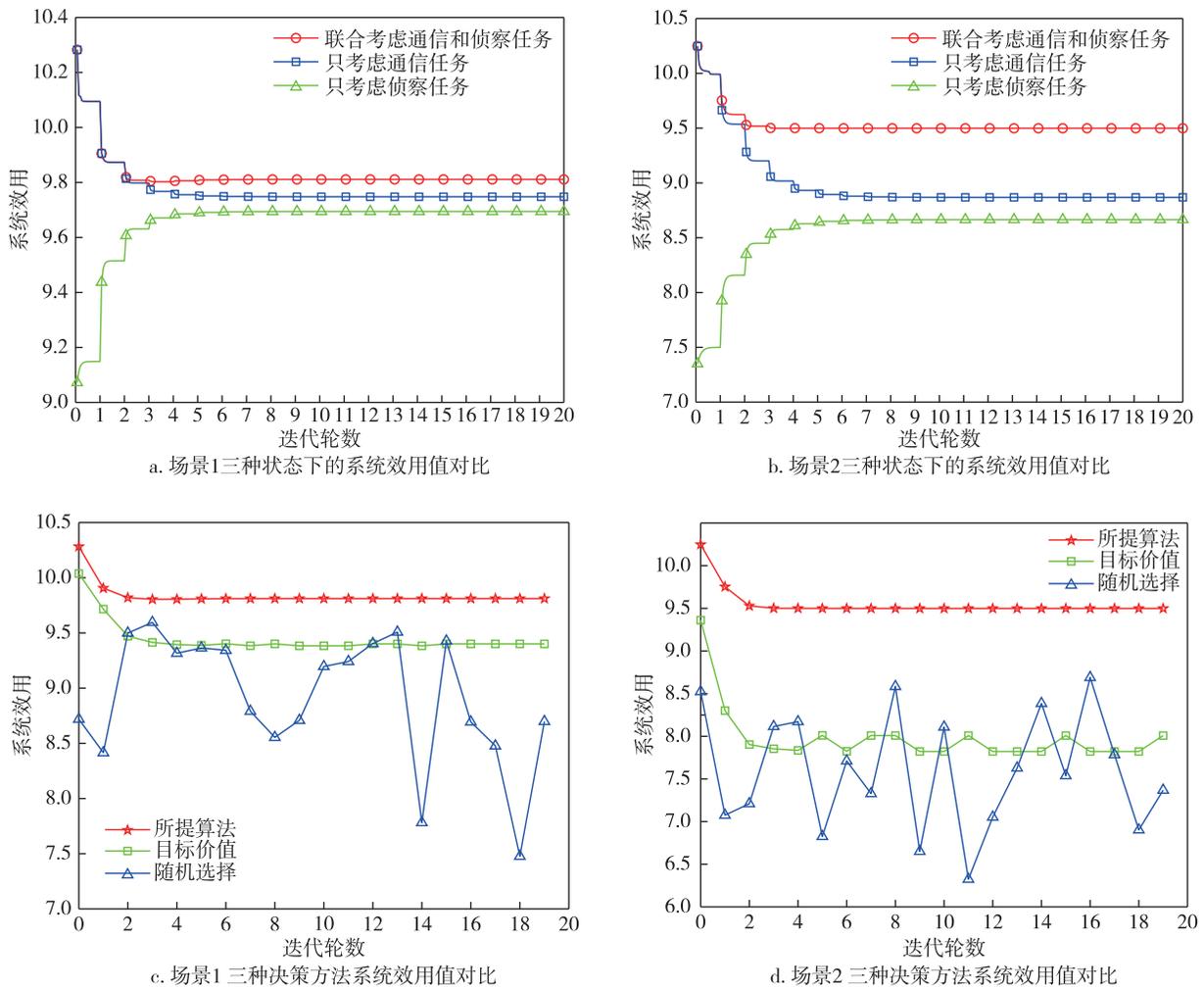


图 7 系统效用对比

Fig. 7 Comparison of system utility

察任务三种状态下的系统效用值变化.在所用算法下,考虑通信任务的收益由于需要满足信噪比需求,功率增大过程中同层和跨层干扰在增大,使得在决策过程中效用值呈现下降的趋势.可以看出,上下层博弈最终收敛至均衡点,所用算法的系统效用值均大于只考虑通信或者侦察单个指标.图7c与7d中联合考虑通信侦察任务下依据所用算法、最大任务目标价值和随机决策三种方法进行任务目标选择,横坐标为迭代轮数,纵坐标为系统效用值,可以看出,所用算法系统效用值均大于依据最大价值选择和随机决策两种方法,并能够有效提高多无人机系统的整体效用.

图8分别给出了8、9、10、11、12、13架协同无人机情况下系统的效用值对比.协同无人机数量的改变构成不同的场景,可以看出,上下层博弈交互迭代达到斯坦伯格均衡后,联合考虑通信侦察任务的效用值均大于考虑通信或者侦察单个指标情况.

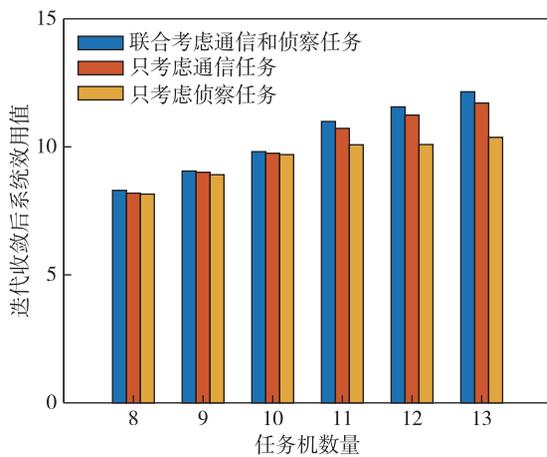


图8 协同无人机数量改变系统效用对比

Fig. 8 Comparison of system utility with different number of CDs

## 5 结束语

多无人机通信和侦察任务分配在无人集群网络优化中具有重要的研究意义.本文聚焦于无人机网络中目标调度和功率控制的联合优化,利用分层博弈框架分析领头无人机和协同无人机的决策行为,采用分布式策略迭代更新算法求解 Stackelberg 均衡,实现无人机最优目标任务调度.对多个场景进行仿真分析,验证了所提算法能够在多无人机系统中实现分布式任务分配的收敛和系统稳定,并有效提高了多无人机系统遂行任务的整体效用.

## 参考文献

### References

- [1] 宗群,王丹丹,邵士凯,等.多无人机协同编队飞行控制研究现状及发展[J].哈尔滨工业大学学报,2017,49(3):1-14  
ZONG Qun, WANG Dandan, SHAO Shikai, et al. Research status and development of multi UAV coordinated formation flight control[J].Journal of Harbin Institute of Technology,2017,49(3):1-14
- [2] 张可为,赵晓林,李宗哲,等.多无人机侦察任务分配方法研究综述[J].电光与控制,2021,28(7):68-72,82  
ZHANG Kewei, ZHAO Xiaolin, LI Zongzhe, et al. A review of multi-UAV reconnaissance mission assignment methods[J].Electronics Optics & Control,2021,28(7):68-72,82
- [3] Zhang L, Zhu Y A, Shi X C. A hierarchical decision-making method with a fuzzy ant colony algorithm for mission planning of multiple UAVs[J].Information,2020,11(4):226
- [4] Huang T Y, Wang Y, Cao X W, et al. Multi-UAV mission planning method[C]//2020 3rd International Conference on Unmanned Systems (ICUS).November 27-28, 2020, Harbin, China.IEEE,2020:325-330
- [5] Du X Y, Guo Q C, Li H, et al. Multi-UAVs cooperative task assignment and path planning scheme[J].Journal of Physics:Conference Series,2021,1856(1):012016
- [6] 赵雪森,王社伟,邵校.基于改进量子粒子群优化算法的多UCAV协同任务分配研究[J].四川兵工学报,2015,36(10):120-124  
ZHAO Xuesen, WANG Shewei, SHAO Xiao. Cooperative task allocation for multipleUCAV based on improved quantum-behaved particle swarm optimization algorithm [J]. Journal of Sichuan Ordnance, 2015, 36 ( 10 ): 120-124
- [7] 唐峯竹,唐欣,李春海,等.基于深度强化学习的多无人机任务动态分配[J].广西师范大学学报(自然科学版),2021,39(6):63-71  
TANG Fengzhu, TANG Xin, LI Chunhai, et al. Dynamic task allocation method for UAVs based on deep reinforcement learning [J].Journal of Guangxi Normal University (Natural Science Edition),2021,39(6):63-71
- [8] 王树朋,徐旺,刘湘德,等.基于自适应遗传算法的多无人机协同任务分配[J].电子信息对抗技术,2021,36(1):59-64  
WANG Shupeng, XU Wang, LIU Xiangde, et al. Cooperative task assignment for multi-UAV based on adaptive genetic algorithm[J].Electronic Warfare Technology,2021,36(1):59-64
- [9] Quan B, Lu X M, Zhang Y M, et al. A multi-objective tracking task assignment algorithm based on game theory [J]. Journal of Physics: Conference Series, 2021, 1802 (3):032115
- [10] Fu X W, Pan J, Gao X G, et al. Task allocation method for multi-UAV teams with limited communication bandwidth [C]//2018 15th International Conference on Control,

- Automation, Robotics and Vision (ICARCV). November 18-21, 2018, Singapore. IEEE, 2018; 1874-1878
- [11] 邸斌,周锐,丁全心.多无人机分布式协同异构任务分配[J].控制与决策,2013,28(2):274-278  
DI Bin, ZHOU Rui, DING Quanxin. Distributed coordinated heterogeneous task allocation for unmanned aerial vehicles[J]. Control and Decision, 2013, 28(2): 274-278
- [12] 嵇介曲,朱琨,易畅言,等.多无人机辅助移动边缘计算中的任务卸载和轨迹优化[J].物联网学报,2021,5(1):27-35  
JI Jiequ, ZHU Kun, YI Changyan, et al. Joint task offloading and trajectory optimization for multi-UAV assisted mobile edge computing[J]. Chinese Journal on Internet of Things, 2021, 5(1): 27-35
- [13] 陈璞,严飞,刘钊,等.通信约束下异构多无人机任务分配方法[J].航空学报,2021,42(8):306-319  
CHEN Pu, YAN Fei, LIU Zhao, et al. Communication-constrained task allocation of heterogeneous UAVs[J]. Acta Aeronautica et Astronautica Sinica, 2021, 42(8): 306-319
- [14] 潘楠,刘海石,陈启用,等.多基地多目标无人机协同任务规划算法研究[J].现代防御技术,2021,49(2):49-56  
PAN Nan, LIU Haishi, CHEN Qiyong, et al. Study on cooperative mission planning algorithm for multi-base and multi-target UAV[J]. Modern Defense Technology, 2021, 49(2): 49-56
- [15] 蒋硕,袁小平.改进 PSO 算法在多无人机协同任务分配中的应用[J].计算机应用研究,2019,36(11):3344-3347,3360  
JIANG Shuo, YUAN Xiaoping. Application of improved PSO algorithm in multi UAV cooperative task allocation [J]. Application Research of Computers, 2019, 36(11): 3344-3347, 3360
- [16] Chen R F, Cui L, Zhang Y L, et al. Delay optimization with FCFS queuing model in mobile edge computing-assisted UAV swarms; a game-theoretic learning approach [C]//2020 International Conference on Wireless Communications and Signal Processing(WCSP). October 21-23, 2020, Nanjing, China. IEEE, 2020; 245-250
- [17] Buckman N, Choi H L, How J P. Partial replanning for decentralized dynamic task allocation [C]//AIAA Scitech 2019 Forum. San Diego, California. Reston, Virginia: AIAA, 2019. DOI: 10.2514/6.2019-0915
- [18] 杨婷婷,宋绯,孙有铭,等.面向异构无人机中继网络的负载均衡:一种分层博弈方法[J].通信技术,2018,51(11):2619-2626  
YANG Tingting, SONG Fei, SUN Youming, et al. Load balancing in heterogeneous UAV relay network: a Stackelberg game method [J]. Communications Technology, 2018, 51(11): 2619-2626
- [19] Xiao M B, Shroff N B, Chong E K P. A utility-based power-control scheme in wireless cellular systems [J]. IEEE/ACM Transactions on Networking, 2003, 11(2): 210-221
- [20] Sun Y M, Wang J L, Sun F G, et al. Energy-aware joint user scheduling and power control for two-tier femtocell networks: a hierarchical game approach [J]. IEEE Systems Journal, 2018, 12(3): 2533-2544

## Multi-UAV cooperative task allocation for multi-point reconnaissance and communication service

YAO Changhua<sup>1</sup> AN Lei<sup>1</sup> LIU Xin<sup>2</sup> HAN Guizhen<sup>1</sup> GAO Zehe<sup>1</sup>

<sup>1</sup> School of Electronics & Information Engineering, Nanjing University of Information Science & Technology, Nanjing 210044

<sup>2</sup> College of Information Science and Engineering, Guilin University of Technology, Guilin 541006

**Abstract** Aiming at the collaborative optimization of multi-UAV reconnaissance and communication service for multiple heterogeneous targets, the Stackelberg game model is constructed by considering the mission requirements and target values, as well as the restriction between multi-UAV coordination gain and task behavior. The upper-level drone is established as the leader of the game, while the lower-level drones are established as the followers of the game, and then a distributed strategy update iterative algorithm is proposed, which realizes the stable convergence of the multi-UAV task allocation scheme and the optimization of the task revenue. Simulation results show that the proposed approach can effectively improve the efficiency of multi-UAV systems to complete multiple tasks at the same time, and can achieve efficient collaboration for the values of heterogeneous tasks in different environments.

**Key words** multi-UAV system; task allocation; Stackelberg game; iterative algorithm