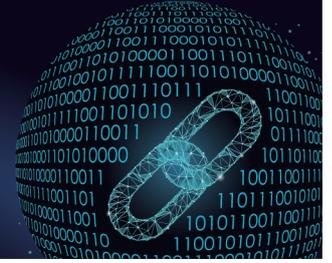




中国科技期刊卓越行动计划项目入选期刊

# 控制与决策

CONTROL AND DECISION



## 基于纳什均衡博弈的多无人机对地攻击目标分配方法

王浩丞, 罗贺, 马滢滢, 王国强

引用本文:

王浩丞, 罗贺, 马滢滢, 王国强. 基于纳什均衡博弈的多无人机对地攻击目标分配方法[J]. *控制与决策*, 2024, 39(4): 1361–1369.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2022.1244>

## 您可能感兴趣的其他文章

### Articles you may be interested in

#### 多无人机协同直播场景下自适应任务卸载决策

Adaptive task offloading decision of multi-UAVs cooperation in live broadcasting scenario

*控制与决策*. 2021, 36(4): 974–982 <https://doi.org/10.13195/j.kzyjc.2019.1104>

#### 多无人机协同直播场景下自适应任务卸载决策

Adaptive task offloading decision of multi-UAVs cooperation in live broadcasting scenario

*控制与决策*. 2021, 36(4): 974–982 <https://doi.org/10.13195/j.kzyjc.2019.1104>

#### 分布式无人机的时变编队非线性控制设计

Time-varying formation nonlinear control of distributed multiple UAVs

*控制与决策*. 2021, 36(10): 2490–2496 <https://doi.org/10.13195/j.kzyjc.2020.0136>

#### 城市低空环境中多旋翼无人机在线航线规划方法

An online route planning method for multi-rotor drone in urban environments

*控制与决策*. 2021, 36(12): 2851–2860 <https://doi.org/10.13195/j.kzyjc.2020.0557>

#### 面向多目标侦察任务的无人机航线规划

UAV trajectory planning for multi-target reconnaissance missions

*控制与决策*. 2021, 36(5): 1191–1198 <https://doi.org/10.13195/j.kzyjc.2019.1284>

# 基于纳什均衡博弈的多无人机对地攻击目标分配方法

王浩丞<sup>1,2</sup>, 罗贺<sup>1,2,3</sup>, 马滢滢<sup>1,2</sup>, 王国强<sup>1,2,3†</sup>

(1. 合肥工业大学管理学院, 合肥 230009; 2. 过程优化与智能决策教育部重点实验室, 合肥 230009;  
3. 安徽省空天系统智能管理工程研究中心, 合肥 230009)

**摘要:** 研究一种新的多无人机对地攻击目标分配问题. 该问题中攻击方试图通过无人机击毁防御方的高价值目标, 防御方试图通过发射拦截导弹对无人机进行拦截, 但攻防双方无法事先观察到对方实际采取的目标分配方案. 通过分析防御方的拦截导弹目标分配方案对攻击方收益的影响, 将问题构建为一个零和矩阵博弈模型, 模型的策略空间随无人机、高价值目标、拦截导弹数量的增加呈爆炸式增长. 鉴于此, 现有算法难以在有效时间内对其进行求解, 提出一种基于两阶段邻域搜索的改进 Double Oracle (DO-TSNS) 算法. 实验结果表明, 相较于 DO、UWMA 和 DO-NS 算法, DO-TSNS 算法能够更有效地求解考虑防御方具有拦截行为的多无人机对地攻击目标分配问题.

**关键词:** 对地攻击; 无人机; 目标分配; 零和矩阵博弈; Double Oracle 算法

中图分类号: TP273

文献标志码: A

DOI: 10.13195/j.kzyjc.2022.1244

引用格式: 王浩丞, 罗贺, 马滢滢, 等. 基于纳什均衡博弈的多无人机对地攻击目标分配方法[J]. 控制与决策, 2024, 39(4): 1361-1369.

## A target assignment method based on Nash equilibrium game for multi UAV ground attack

WANG Hao-cheng<sup>1,2</sup>, LUO He<sup>1,2,3</sup>, MA Ying-ying<sup>1,2</sup>, WANG Guo-qiang<sup>1,2,3†</sup>

(1. School of Management, Hefei University of Technology, Hefei 230009, China; 2. Key Laboratory of Process Optimization and Intelligent Decision-Making of Ministry of Education, Hefei 230009, China; 3. Engineering Research Center for Intelligent Management of Aerospace System of Anhui Province, Hefei 230009, China)

**Abstract:** A new target assignment problem for multi UAV ground attack is studied. In this issue, the attacker tries to destroy the high-value target of the defender through multiple UAVs, and the defender tries to intercept the UAVs by launching interceptor missiles, but the attacker and defender can not observe the target allocation scheme actually adopted by the other side in advance. By analysing the impact of the defender's interceptor missile allocation scheme on the profit of the attacker, this paper constructs the problem as a zero-sum matrix game model. The strategy space of the model increases explosively with the increase of the number of UAVs, high-value targets and interceptor missiles, and the existing algorithms are difficult to solve it in an effective time. Therefore, this paper proposes an improved Double Oracle algorithm based on two-stage neighbourhood search (DO-TSNS). The results of experiments show that, compared with the DO, UWMA and DO-NS algorithm, the DO-TSNS algorithm can solve the target assignment problem of multi UAV ground attack considering the interception behaviour of the defender more effectively.

**Keywords:** ground attack; UAV; target assignment; zero-num matrix game; Double Oracle algorithm

## 0 引言

在现代化战争中,对地面的重要目标如公路、机场、桥梁、军事指挥所和弹药仓库等进行打击具有重要的战略意义<sup>[1]</sup>. 无人机凭借其成本低、体积小、机动性强等特点<sup>[2-4]</sup>, 现已被逐渐应用于对地攻击作

战之中<sup>[5]</sup>. 其中,目标分配作为无人机对地攻击的一个重要环节,攻击方需要决策合理的目标分配方案,确定每架无人机所要打击的目标,最大化对地攻击收益<sup>[6]</sup>.

无人机对地攻击目标分配问题属于武器目标分

收稿日期: 2022-07-12; 录用日期: 2022-12-08.

基金项目: 国家自然科学基金项目(71871079, 71971075, 72271076).

责任编委: 刘宝碇.

†通讯作者. E-mail: gqwang2017@hfut.edu.cn.

配问题. 武器目标分配问题是军事领域中一类经典的运筹学问题. 简单来说, 武器目标分配问题寻求将固定数量的武器最优分配给固定数量的目标, 以最大限度地增加对这组目标造成的总毁伤<sup>[7]</sup>. Manne<sup>[8]</sup>在1958年首次提出了武器目标分配问题, 该问题后被证明为NP-Complete问题<sup>[9]</sup>, 其策略空间规模随武器、目标数量的增加呈指数级增长.

对于对地攻击中的武器目标分配问题, 在现有研究中大部分学者考虑目标的价值、武器对目标的毁伤概率等因素, 以最大化对目标的期望毁伤价值或最小化目标的生存价值, 将其建模为非线性整数规划模型, 并设计启发式算法进行求解. 如: 改进差分进化算法<sup>[10]</sup>、改进粒子群算法<sup>[11]</sup>、人工鱼群算法<sup>[12]</sup>和文化基因算法<sup>[13]</sup>等. 也有部分学者开始运用强化学习的方法求解武器目标分配问题<sup>[14-15]</sup>. 然而以上方法仅适用于防御方没有拦截行为的对地攻击场景.

在实际情形中, 敌方的重要目标通常在防御系统保护之下<sup>[16]</sup>, 防御系统会发射拦截导弹对武器进行拦截. 近年来, 许多学者研究如何制定合理有效的拦截导弹分配方案, 以最大程度地减少目标的毁伤. Guo等<sup>[17]</sup>为了确保每个武器被分配到适当的拦截资源, 提出了固定分组策略和基于惩罚函数的自适应分组策略, 并设计了人工蜂群算法进行求解. Gao等<sup>[18]</sup>通过综合考虑拦截成本和防御方的风险偏好等因素, 提出了一种双目标的导弹拦截策略优化方法. Davis等<sup>[19]</sup>针对动态拦截导弹分配问题, 利用近似动态规划方法进行求解. Shalunov等<sup>[20]</sup>针对目标-导弹-防御装备交战场景, 提出了一种计算效率较高的武器目标分配算法. 然而, 上述研究均是在给定攻击方目标分配方案下找到一个合理有效的拦截导弹分配方案.

针对防御方具有拦截行为的对地攻击武器目标分配问题, 部分学者开展了研究. Han等<sup>[21]</sup>假设防御方能够事先观察到攻击方的目标分配方案, 视其为一个攻击方先决策、防御方后决策的两阶段决策问题, 将该问题建模为斯塔科尔伯格博弈模型, 并提出了一种基于邻域搜索的改进Double Oracle算法进行求解. Boardman等<sup>[22]</sup>在此基础上进一步考虑了防御方拦截导弹的异构性, 同样建模为斯塔科尔伯格博弈模型, 并分别提出基于模拟退火、禁忌搜索的改进Double Oracle算法进行求解. 然而, 在实际的作战场景中, 攻防双方很难准确知道对方要采取的具体目标分配方案, 通常只知道彼此可能采取的目标分配方案集合, 因此双方在决策时均可以视为是同时进行

的. 此时, 攻击方便需要在只给定决策双方目标分配方案集合的情形下, 决策出一个有效的具体目标分配方案, 达到尽可能摧毁防御方目标的效果.

对此, 本文基于纳什均衡博弈思想, 对防御方具有拦截行为的多无人机对地攻击中的目标分配问题进行研究, 主要贡献如下:

1) 考虑防御方无法事先观察到攻击方实际采取的目标分配方案的情形, 提出一个新的防御方具有拦截行为的多无人机对地攻击目标分配问题;

2) 基于纳什均衡博弈思想, 通过分析攻防双方的策略集合, 设计给定攻防双方策略对下的收益函数, 构建该问题的零和矩阵博弈模型;

3) 针对中大规模零和矩阵博弈模型的求解, 提出一种基于两阶段邻域搜索的改进Double Oracle算法, 该算法能够适用于不同规模的该问题实例.

## 1 问题描述与建模

### 1.1 问题描述

从经典的OODA(观察-判断-决策-行动)循环<sup>[23]</sup>角度, 针对多无人机对地攻击场景中的目标分配方案决策环节进行研究, 具体的问题场景是攻击方利用全部无人机资源去尽可能地摧毁防御方的高价值目标, 以最大化对防御方高价值目标的期望毁伤价值, 如图1所示. 其中, 攻击方通过自杀式无人机(后文简称为无人机)对防御方的高价值目标进行打击, 防御方为保护己方的高价值目标, 通过防御系统发射拦截导弹对来袭的无人机进行拦截.

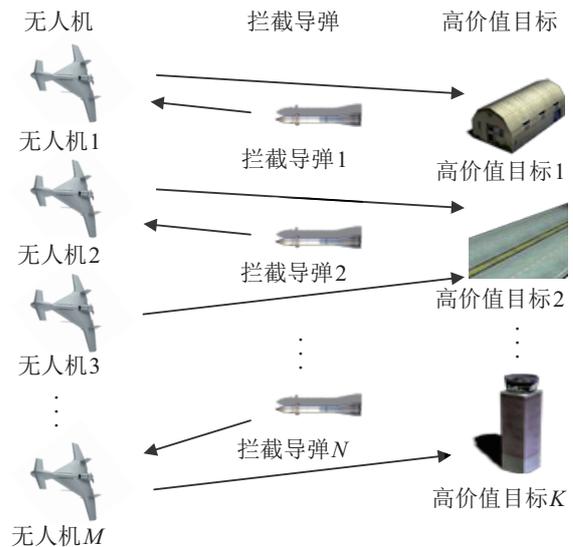


图1 问题描述

我方为攻击方, 记为 $T$ , 有 $M$ 架无人机. 对方为防御方, 记为 $L$ , 有 $K$ 个高价值目标, 防御系统共有 $N$ 枚拦截导弹. 在已知攻击方无人机的数量 $M$ 、防御方的拦截导弹数量 $N$ 、高价值目标数量 $K$ 、无人机对

目标的毁伤概率  $P_{ik}$ 、拦截导弹对无人机的拦截概率  $P_{ji}^{LJ}$ 、目标的值  $v_k$  的前提下,攻击方需要决策一个无人机目标分配方案,分配无人机对高价值目标进行打击,以最大化对高价值目标的期望毁伤价值;防御方需要决策一个拦截导弹目标分配方案,分配拦截导弹对无人机进行拦截,以最小化高价值目标的期望毁伤价值.上述参数可以在OODA循环中的观察和判断环节获得<sup>[24-26]</sup>.

用  $w_T = [a_{mk}]_{m \in \{1,2,\dots,M\}, k \in \{1,2,\dots,K\}}$  表示攻击方的目标分配方案,  $a_{mk} \in \{0, 1\}$ , 其中  $a_{mk} = 1$  表示第  $m$  架无人机攻击第  $k$  个高价值目标, 否则  $a_{mk} = 0$ ; 用  $w_L = [b_{nm}]_{n \in \{1,2,\dots,N\}, m \in \{1,2,\dots,M\}}$  表示防御方的目标分配方案,  $b_{nm} \in \{0, 1\}$ , 其中  $b_{nm} = 1$  表示第  $n$  枚拦截导弹拦截第  $m$  架无人机, 否则  $b_{nm} = 0$ .  $T$  和  $L$  通过决策各自的目标分配方案  $w_T$  和  $w_L$ , 使得己方的优势尽可能大, 劣势尽可能小.

### 1.2 零和矩阵博弈模型构建

针对上述多无人机对地攻击目标分配问题, 由于防御方的防御策略会对攻击方的攻击策略收益产生影响, 本文将其建模为一个零和矩阵博弈模型  $G = (P, S, U)$ . 该模型中使用到的符号及其含义见表1.

表1 参数定义

参数名	含义
$T$	攻击方
$L$	防御方
$M$	无人机数量
$N$	拦截导弹数量
$K$	高价值目标数量
$P_{mk}$	无人机 $m$ 对高价值目标 $k$ 的毁伤概率
$P_{nm}^{LJ}$	拦截导弹 $n$ 对无人机 $m$ 的拦截概率
$v_k$	高价值目标 $k$ 的价值
$w_T$	攻击方的目标分配方案
$w_L$	防御方的目标分配方案
$a_{mk}$	$a_{mk} \in \{0, 1\}$ , $a_{mk} = 1$ 表示第 $m$ 架无人机攻击第 $k$ 个高价值目标, 否则 $a_{mk} = 0$
$b_{nm}$	$b_{nm} \in \{0, 1\}$ , $b_{nm} = 1$ 表示第 $n$ 枚拦截导弹拦截第 $m$ 架无人机目标, 否则 $b_{nm} = 0$
$S_T$	攻击方的纯策略集合
$S_L$	防御方的纯策略集合
$\tau_T$	$S_T$ 中的纯策略总数
$\tau_L$	$S_L$ 中的纯策略总数
$s_T$	$S_T$ 中的一个纯策略
$s_L$	$S_L$ 中的一个纯策略
$u(s_T, s_L)$	$T$ 方在策略对 $(s_T, s_L)$ 下的收益值
$x$	$T$ 方的混合策略
$x_{s_T}$	选择 $s_T$ 的概率
$y$	$L$ 方的混合策略
$y_{s_L}$	选择 $s_L$ 的概率

博弈模型包含参与者、策略集合和收益矩阵3个部分, 具体解释如下:

1) 参与者.

记  $P = \{T, L\}$  为博弈的参与者,  $T$  为攻击方,  $L$

为防御方.

2) 策略集合.

记  $S \equiv S_T \times S_L$  为博弈的策略空间.  $S_T$  为攻击方的纯策略集合,  $s_T \in S_T$  为  $S_T$  中一个纯策略, 一个纯策略  $s_T$  对应一个目标分配方案  $w_T$ ,  $\tau_T$  为  $S_T$  中纯策略总数.  $S_L$  为防御方的纯策略集合,  $s_L \in S_L$  为  $S_L$  中一个纯策略, 一个纯策略  $s_L$  对应一个目标分配方案  $w_L$ ,  $\tau_L$  为  $S_L$  中纯策略总数.

由于  $T$  方共有  $M$  架无人机, 每架无人机可以攻击  $L$  方任何一个高价值目标, 而  $L$  方共有  $K$  个高价值目标, 则  $T$  方纯策略数量  $\tau_T$  为  $K^M$ .  $L$  方共有  $N$  枚拦截导弹, 每枚拦截导弹可以拦截  $T$  方任何一架无人机, 则  $L$  方纯策略数量  $\tau_L$  为  $M^N$ , 即博弈的策略空间规模为  $|S| = \tau_T \cdot \tau_L = K^M \cdot M^N$ .

3) 收益矩阵.

记  $U = [u(s_T, s_L)]_{s_T \in S_T, s_L \in S_L}$  为博弈的收益矩阵, 其中  $u(s_T, s_L)$  为  $T$  方在策略对  $(s_T, s_L)$  下的收益值, 即  $T$  方在  $L$  方执行  $s_L$  策略的情况下执行  $s_T$  策略时对  $L$  方高价值目标的总期望毁伤值.  $u(s_T, s_L)$  为

$$u(s_T, s_L) = f(w_T, w_L) = \sum_{k=1}^K \left( 1 - \prod_{m=1}^M \left( 1 - \prod_{n=1}^N (1 - P_{nm}^{LJ})^{b_{nm}} \cdot P_{mk} \right)^{a_{mk}} \right) \cdot v_k, \quad a_{mk} \in \{0, 1\}, b_{nm} \in \{0, 1\}. \quad (1)$$

其中:  $\prod_{n=1}^N (1 - P_{nm}^{LJ})^{b_{nm}}$  为无人机  $m$  未被拦截的概率,

$\prod_{n=1}^N (1 - P_{nm}^{LJ})^{b_{nm}} \cdot P_{mk}$  为无人机  $m$  对高价值目标  $k$

的有效毁伤概率,  $1 - \prod_{n=1}^N (1 - P_{nm}^{LJ})^{b_{nm}} \cdot P_{mk}$  为高价值目标  $k$  未被无人机  $m$  毁伤的概率;

$\prod_{m=1}^M \left( 1 - \prod_{n=1}^N (1 - P_{nm}^{LJ})^{b_{nm}} \cdot P_{mk} \right)^{a_{mk}}$  为高价值目标  $k$  实际未被毁伤的概率;

$1 - \prod_{n=1}^N (1 - P_{nm}^{LJ})^{b_{nm}} \cdot P_{mk}$  为高价值目标  $k$  实际被毁伤的概率;

$\sum_{k=1}^K \left( 1 - \prod_{m=1}^M \left( 1 - \prod_{n=1}^N (1 - P_{nm}^{LJ})^{b_{nm}} \cdot P_{mk} \right)^{a_{mk}} \right) \cdot v_k$  为攻击方对高价值目标的总期望毁伤值.

### 1.3 混合策略纳什均衡

记  $x = (x_{s_T})_{s_T \in S_T}, 0 \leq x_{s_T} \leq 1, \sum_{s_T \in S_T} x_{s_T} = 1$  为  $T$  方一个混合策略, 表示  $T$  方以概率  $x_{s_T}$  选择纯策略  $s_T$ . 记  $y = (y_{s_L})_{s_L \in S_L}, 0 \leq y_{s_L} \leq 1, \sum_{s_L \in S_L} y_{s_L} = 1$

1 为  $L$  方一个混合策略,表示  $L$  方以概率  $y_{s_L}$  选择纯策略  $s_L$ . 在混合策略对  $(\mathbf{x}, \mathbf{y})$ ,  $T$  方的期望收益为  $F(\mathbf{x}, \mathbf{y}) = \sum_{s_T \in S_T} \sum_{s_L \in S_L} x_{s_T} \cdot y_{s_L} \cdot u(s_T, s_L)$ , 根据文献[27]零和矩阵博弈的性质,零和矩阵博弈双方的收益值之和为零,因此  $L$  方的期望收益为  $-F(\mathbf{x}, \mathbf{y})$ . 若存在混合策略对  $(\hat{\mathbf{x}}, \hat{\mathbf{y}})$ , 使得  $F(\hat{\mathbf{x}}, \hat{\mathbf{y}})$  满足如下所示的条件,则称  $(\hat{\mathbf{x}}, \hat{\mathbf{y}})$  为该零和矩阵博弈的混合策略纳什均衡[28]:

$$F(\mathbf{x}, \hat{\mathbf{y}}) \leq F(\hat{\mathbf{x}}, \hat{\mathbf{y}}) \leq F(\hat{\mathbf{x}}, \mathbf{y}). \quad (2)$$

## 2 算法设计与实现

针对零和矩阵博弈模型的求解,目前最好的算法是 Double Oracle (DO) 算法[29],其基本思路是:首先在整个策略空间中生成一个小规模的初始严格博弈矩阵,然后计算当前严格博弈矩阵的混合策略纳什均衡,再计算该混合策略纳什均衡下双方最优反应策略,并将其加入到严格博弈矩阵中,重复迭代直到双方最优反应策略在当前严格博弈矩阵中都已存在,此时严格博弈矩阵的混合策略纳什均衡即为整个零和矩阵博弈模型的混合策略纳什均衡[30].

然而,随着无人机数量、高价值目标和拦截导弹数量的增加,本文问题中的策略空间规模急剧增加,DO 算法通过枚举寻找双方最优反应策略的方法难以在有限时间内求解该问题.为此,提出一种基于两阶段邻域搜索的改进 DO 算法(double oracle based on two stage neighbourhood search, DO-TSNS),该算法以迭代搜索的方式寻找双方最优反应策略,可以有效减少策略的搜索空间,提高求解效率.

### 2.1 DO-TSNS 的框架

类似于文献[18]设计的 DO 算法框架,DO-TSNS 使用 3 个子算法对所提出的零和矩阵博弈模型  $G$  进行求解,子算法分别为 Lemke-Howson (LH)、T-Oracle 和 L-Oracle. DO-TSNS 算法伪代码如下所示.

#### 算法 1 DO-TSNS.

输入: 零和矩阵博弈模型  $G = (P, S, U)$ ;

输出: 混合策略纳什均衡  $(\hat{\mathbf{x}}, \hat{\mathbf{y}})$ .

- 1:  $S_T^0 \leftarrow \text{random}(S_T, 2)$
- 2:  $S_L^0 \leftarrow \text{random}(S_L, 2)$
- 3:  $i = 0$
- 4: repeat
- 5:  $(\hat{\mathbf{x}}, \hat{\mathbf{y}}) \leftarrow \text{LH}(S_T^i, S_L^i)$
- 6:  $\bar{s}_T \leftarrow \text{T-Oracle}(\hat{\mathbf{y}})$
- 7:  $S_T^{i+1} \leftarrow S_T^i \cup \bar{s}_T$
- 8:  $\bar{s}_L \leftarrow \text{L-Oracle}(\hat{\mathbf{x}})$

$$9: S_L^{i+1} \leftarrow S_L^i \cup \bar{s}_L$$

$$10: i \leftarrow i + 1$$

$$11: \text{until } S_T^i = S_T^{i-1} \text{ and } S_L^i = S_L^{i-1}$$

$$12: \text{return } (\hat{\mathbf{x}}, \hat{\mathbf{y}})$$

DO-TSNS 算法计算过程如下:分别从  $T$  方的纯策略集合  $S_T$ 、 $L$  方的纯策略集合  $S_L$  中随机选出两个纯策略,作为初始的严格博弈矩阵的策略集合  $S^0 = (S_T^0, S_L^0)$  (line 1 和 line 2); 在第  $i$  次迭代时,使用 Lemke-Howson (LH) 算法[31]对当前的严格博弈矩阵进行求解,从而得到当前严格博弈矩阵的混合策略纳什均衡  $(\hat{\mathbf{x}}, \hat{\mathbf{y}})$  (line 5); 使用 T-Oracle 子算法计算  $L$  方使用混合策略  $\hat{\mathbf{y}}$  时  $T$  方的最优反应策略  $\bar{s}_T$ , 并将其加入下一阶段严格博弈矩阵的  $T$  方策略集合  $S_T^{i+1}$  中 (line 6 和 line 7); 使用 L-Oracle 子算法计算  $T$  方使用混合策略  $\hat{\mathbf{x}}$  时  $L$  方的最优反应策略  $\bar{s}_L$ , 并将其加入到下一阶段严格博弈矩阵的  $L$  方策略集合  $S_L^{i+1}$  中 (line 8 和 line 9); 若双方的最优反应策略都已在当前严格博弈矩阵的策略集合  $S^i = (S_T^i, S_L^i)$  中,则停止迭代,得到当前严格博弈矩阵的混合策略纳什均衡,否则进入下一次循环 (line 10 ~ line 12).

### 2.2 T-Oracle 算法

在 DO-TSNS 算法第  $i$  次迭代时,给定  $L$  方当前混合策略  $\hat{\mathbf{y}} = (\hat{y}_{s_L})_{s_L \in S_L^i}$  时,需要在  $T$  方的策略空间中寻找一个使得  $T$  方收益最大化的策略作为  $T$  方的最优反应策略  $\text{BRS}_T$  (line 6). 该问题可以被建模为如下非线性整数规划问题:

$$\text{maximize}_{s_T} \sum_{s_L \in S_L^i} u(s_T, s_L) \cdot \hat{y}_{s_L}; \quad (3)$$

$$\text{s.t.} \quad \sum_{m \in \{1, \dots, M\}, k \in \{1, \dots, K\}} a_{mk} = M, \quad (4)$$

$$a_{mk} \in \{0, 1\}. \quad (5)$$

其中: 式 (3) 表示优化目标为找到使  $T$  方收益最大化的策略  $s_T$ ; 式 (4) 表示策略  $s_T = [a_{mk}]_{m \in \{1, \dots, M\}, k \in \{1, \dots, K\}}$  中  $T$  方每个无人机都会攻击某一个高价值目标; 式 (5) 表示策略  $s_T$  中  $a_{mk}$  的取值为 0 或 1.

针对此问题,枚举的方法难以在短时间内求解,为此提出一种基于两阶段邻域搜索的启发式算法 T-Oracle,其算法伪代码如下所示.

#### 算法 2 T-Oracle.

$$1: \bar{s}_T \leftarrow \text{maximize}_{s_T \in S_T^i} \hat{x}_{s_T}$$

$$2: \bar{u} = -\infty$$

$$3: \text{repeat}$$

$$4: N_e \leftarrow \text{neighbors}(\bar{s}_T)$$

```

5: for  $j = 1$  to  $|N_e|$  do
6:   if  $\sum_{s_L \in S_L^i} u(N_e(j), s_L) \cdot \hat{y}_{s_L} \geq \bar{u}$  then
7:      $\bar{s}_T \leftarrow N_e(j)$ 
8:      $\bar{u} \leftarrow \sum_{s_L \in S_L^i} u(N_e(j), s_L) \cdot \hat{y}_{s_L}$ 
9:   end if
10: end for
11: until  $\bar{s}_T = N_e(1)$ 
12: return  $\bar{s}_T$ 

```

将  $T$  方均衡策略  $\hat{x} = (\hat{x}_{s_T})_{s_T \in S_T^i}$  中概率最大的策略  $\hat{x}_{s_T}$  作为初始策略, 即  $\bar{s}_T = \underset{s_T \in S_T^i}{\text{maximize}} \hat{x}_{s_T}$  (line 1); 采用基于两阶段邻域搜索的邻居策略生成方法生成当前策略  $\bar{s}_T$  的所有邻居策略, 记邻居策略集合为  $\bar{s}_T$  (line 4), 其中  $N_e(j)$  为  $\bar{s}_T$  中第  $j$  个策略, 并将策略  $\bar{s}_T$  作为  $\bar{s}_T$  中的第 1 个策略, 记为  $N_e(1)$ ; 寻找  $\bar{s}_T$  中使  $T$  方期望收益最大的策略替换当前策略  $\bar{s}_T$  (line 5 ~ line 10), 并判断  $\bar{s}_T$  是否为  $N_e(1)$  (line 11), 若否则进入下一次循环, 重新生成新的当前策略  $\bar{s}_T$  及其邻居策略集合, 继续寻找使  $T$  方期望收益最大的策略, 若是则返回当前策略  $\bar{s}_T$  作为  $T$  方的近似最优反应策略 (line 12).

算法 2 中基于两阶段邻域搜索的邻居策略集合生成方法 (line 4) 具体分为以下两个步骤:

1) 第 1 阶段邻域搜索.

参考武器目标分配问题求解算法的典型邻域结构<sup>[31]</sup>, 以一维数组的形式表示当前策略  $\bar{s}_T$ , 通过将  $\bar{s}_T$  中某个无人机攻击的高价值目标  $k \in \{1, 2, \dots, K\}$  变换为相邻的高价值目标  $k - 1$  或  $k + 1$ , 生成初始邻居策略集合. 特别地, 当  $k = 1$  时, 令  $k - 1 = K$ ; 当  $k = K$  时, 令  $k + 1 = 1$ . 将策略  $\bar{s}_T$  作为其初始邻居策略集合中的第 1 个策略.

例如, 当  $T$  方有 3 架无人机,  $L$  方有 5 个高价值目标时, 策略  $[1, 3, 5]$  表示  $T$  方的第 1、第 2、第 3 架无人机分别攻击  $L$  方的第 1、第 3、第 5 个高价值目标. 通过第 1 阶段邻域搜索, 生成其初始邻居策略集合为

$$\text{neighbor}([1, 3, 5]) = \left\{ \begin{array}{l} [1, 3, 5][2, 3, 5] \\ [5, 3, 5][1, 4, 5] \\ [1, 2, 5][1, 3, 1][1, 3, 4] \end{array} \right\}.$$

2) 第 2 阶段邻域搜索.

在  $\bar{s}_T$  初始邻居策略集合的基础上, 通过 2-opt 变换, 即对初始邻居策略集合中每个策略中的任意两个元素进行对换, 生成最终的邻居策略集合.

例如, 在策略  $[1, 3, 5]$  的上述初始邻居策略集合的基础上, 通过第 2 阶段邻域搜索, 生成策略  $[1, 3, 5]$

的最终邻居策略集合为

$$\text{neighbors}([1, 3, 5]) = \left\{ \begin{array}{l} [1, 3, 5][3, 1, 5][5, 3, 1][1, 5, 3] \\ [2, 3, 5][3, 2, 5][5, 3, 2][2, 5, 3] \\ [5, 3, 5][3, 5, 5][5, 5, 3] \\ [1, 4, 5][4, 1, 5][5, 4, 1][1, 5, 4] \\ [1, 2, 5][2, 1, 5][5, 2, 1][1, 5, 2] \\ [1, 3, 1][3, 1, 1][1, 1, 3] \\ [1, 3, 4][3, 1, 4][4, 3, 1][1, 4, 3] \end{array} \right\}.$$

2.3 L-Oracle 算法

同理, 在 DO-TSNS 算法第  $i$  次迭代, 给定  $T$  方当前的混合策略  $\hat{x} = (\hat{x}_{s_T})_{s_T \in S_T^i}$  时, 需要在  $L$  方的策略空间中寻找一个使得  $L$  方收益最大化的策略作为  $L$  方的最优反应策略  $\text{BRS}_L$ . 该问题可以被建模为如下非线性整数规划问题:

$$\text{maximize}_{s_L} \sum_{s_T \in S_T^i} u(s_T, s_L) \cdot \hat{x}_{s_T}; \quad (6)$$

$$\text{s.t.} \quad \sum_{n \in \{1, 2, \dots, N\}, m \in \{1, 2, \dots, M\}} b_{nm} = N, \quad (7)$$

$$b_{nm} \in \{0, 1\}. \quad (8)$$

可以看出, 该问题与  $T$  方最优反应策略问题只在参数输入上具有差别, 因此本文同样采用基于两阶段邻域搜索的启发式算法对其进行求解, 并将相应算法命名为 L-Oracle.

2.4 DO-TSNS 输出的混合策略纳什均衡解分析

本文算法输出的混合策略纳什均衡解为算法停止迭代时严格博弈矩阵的混合纳什均衡, 即算法停止迭代时严格博弈矩阵的策略集合  $S^i = (S_T^i, S_L^i)$  中每个策略被选择的概率分布. 例如: 以无人机数量  $M = 5$ 、拦截导弹数量  $N = 5$ 、高价值目标数量  $K = 5$  为例, 若停止迭代时严格博弈矩阵的策略集合  $S^i = (S_T^i, S_L^i)$ , 其中  $S_T^i$  为  $\{[3, 5, 4, 2, 1], [1, 4, 5, 2, 3]\}$ ,  $S_L^i$  为  $\{[2, 4, 5, 1, 3], [2, 4, 1, 5, 3]\}$ , 混合策略纳什均衡  $(\hat{x}, \hat{y}) = ((0.775, 0.225), (0.395, 0.605))$ , 则此时混合策略纳什均衡的含义为  $T$  方以 0.775 的概率选择  $[3, 5, 4, 2, 1]$  这一策略, 以 0.225 的概率选择  $[1, 4, 5, 2, 3]$  这一策略;  $L$  方以 0.395 的概率选择  $[2, 4, 5, 1, 3]$  这一策略, 以 0.605 的概率选择  $[2, 4, 1, 5, 3]$  这一策略.

由于在求解双方的最优反应策略时, 本文算法使用两阶段邻域搜索方法, 得到的是最优反应策略的近似解, 而不是精确的混合纳什均衡解. 然而, 针对大规模问题实例, 现有算法均无法在有限时间内得到精确的混合纳什均衡解, 而本文算法可以快速求得高质

量近似的混合策略纳什均衡解,从而能够为多无人机对地攻击目标分配决策环节提供决策支持.在实践应用过程中,基于本文算法输出的混合策略纳什均衡解,攻击方通常可以采用轮盘赌方法从  $S_T^i$  中选取具体的目标分配方案去执行,这种方式不容易被防御方掌握规律,防御方难以猜中攻击方实际采取的具体目标分配方案,从而确保了攻击方决策方案的有效性.

### 3 数值实验与结果分析

#### 3.1 实验设置

为验证本文算法的有效性,参照文献[21]设计小规模和中大规模两类实验,并选取3种算法分别从求解时间和求解质量两个方面进行对比.3种对比算法为现有求解零和矩阵博弈模型的主流算法,分别为DO、UWMA<sup>[32]</sup>和DO-NS算法<sup>[31]</sup>.其中:DO和UWMA算法都可以求得精确的纳什均衡解,但是求解时间随着问题规模的增加迅速增长,难以有效求解中大规模的问题实例;DO-NS算法可以快速求得近似的纳什均衡解,但是与精确纳什均衡解的期望收益差距较大.实验采取计算最差期望收益的方法对算法的求解质量进行度量,其中最差期望收益是指  $T$  方执行算法计算得到的均衡策略  $\hat{x}$  以及  $L$  方执行对  $\hat{x}$  的最优反应策略  $\text{maximize}_{s_L} \sum_{\hat{x}_{s_T} \in \hat{x}} u(s_T, s_L) \hat{x}_{s_T}$  时  $T$

方的期望收益值<sup>[31]</sup>.用  $E[u(\text{DO})]$  表示使用DO算法计算得到的解对应的最差期望收益,用  $E[u(Q)]$  表示使用  $Q$  算法计算得到的解对应的最差期望收益,则  $Q$  算法与DO算法之间gap的计算公式为

$$(E[u(\text{DO})] - E[u(Q)])/E[u(\text{DO})]. \quad (9)$$

参考“萨德”反导系统的拦截概率<sup>[33]</sup>,实验中拦截导弹对无人机的拦截概率  $P_{ji}^{LJ}$  在(0.6, 0.9)内随机生成.另外,实验中无人机对高价值目标的毁伤概率  $P_{ik}$  在(0.6, 0.9)内随机生成,高价值目标的价值系数  $v_k$  在(6, 10)的范围内随机生成.

本文实验环境为: Intel(R) Core(TM) i7-7700 CPU@3.60 GHz 4核处理器、32 GB RAM、Windows 10操作系统,所有算法均使用C++语言编程实现.

#### 3.2 小规模实验及其结果分析

将高价值目标数量  $K$  设为3、5、8,不同  $K$  下将无人机数量  $M$  和拦截导弹数量  $N$  分别设为3、5、8,共得到27种组合.每种组合下随机生成5组实验数据,每组实验数据下分别运行DO-TSNS和3种对比算法5次,记录每次运行算法的计算时间和得到的最差期望收益.针对每种组合分别统计各算法平均计算时间和得到的平均最差期望收益,并按照策略空间规模  $|S|$  的大小升序排列,实验结果如表2所示,其中“/”表示该算法在有限时间(72小时)内无法求解.

表2 小规模实验结果

K	M	N	S	平均计算时间/s				平均最差期望收益			
				UWMA	DO	DO-NS	DO-TSNS	UWMA	DO	DO-NS	DO-TSNS
3	3	3	$3^3 \times 3^3$	3.931	0.083	<b>0.028</b>	0.054	<b>3.928</b>	<b>3.928</b>	<b>3.928</b>	<b>3.928</b>
	3	5	$3^3 \times 3^5$	11 953.37	0.137	<b>0.062</b>	0.078	<b>1.337</b>	<b>1.337</b>	<b>1.337</b>	<b>1.337</b>
	5	3	$3^5 \times 3^3$	52 849.68	0.732	<b>0.071</b>	0.434	<b>9.280</b>	<b>9.280</b>	8.359	8.960
	3	8	$3^3 \times 3^8$	/	0.762	<b>0.204</b>	0.549	/	<b>0.265</b>	0.261	<b>0.265</b>
	5	5	$3^5 \times 3^5$	/	0.930	<b>0.306</b>	0.578	/	<b>5.077</b>	4.905	5.056
	8	3	$3^8 \times 3^3$	/	2.224	<b>0.381</b>	0.669	/	<b>16.291</b>	15.054	15.772
	5	8	$3^5 \times 3^8$	/	45.208	<b>0.409</b>	0.987	/	<b>2.203</b>	2.139	2.201
	8	5	$3^8 \times 3^5$	/	53.070	<b>0.446</b>	1.261	/	<b>11.697</b>	10.708	11.420
5	8	8	$3^8 \times 3^8$	/	2 562.59	<b>0.715</b>	1.544	/	<b>6.960</b>	6.753	6.871
	3	3	$5^3 \times 3^3$	5 524.63	0.126	<b>0.048</b>	0.080	<b>3.958</b>	<b>3.958</b>	<b>3.958</b>	<b>3.958</b>
	3	5	$5^3 \times 3^5$	40 587.39	0.317	<b>0.095</b>	0.235	<b>1.383</b>	<b>1.383</b>	1.249	<b>1.383</b>
	5	3	$5^5 \times 3^3$	/	0.916	<b>0.081</b>	0.296	/	<b>13.094</b>	11.672	12.660
	3	8	$5^3 \times 3^8$	/	0.789	<b>0.159</b>	0.449	/	<b>0.260</b>	0.254	0.255
	5	5	$5^5 \times 3^5$	/	1.034	<b>0.303</b>	0.496	/	<b>5.146</b>	4.596	5.124
	8	3	$5^8 \times 3^3$	/	43.724	<b>0.428</b>	0.794	/	<b>21.892</b>	20.693	21.211
	5	8	$5^5 \times 3^8$	/	121.340	<b>0.561</b>	1.037	/	<b>2.196</b>	2.089	2.152
8	8	5	$5^8 \times 3^5$	/	130.827	<b>0.686</b>	1.512	/	<b>13.129</b>	12.483	12.889
	8	8	$5^8 \times 3^8$	/	3 145.76	<b>0.881</b>	1.690	/	<b>7.604</b>	7.059	7.489
	3	3	$8^3 \times 3^3$	18 635.53	0.465	<b>0.069</b>	0.198	<b>4.406</b>	<b>4.046</b>	3.585	3.974
	3	5	$8^3 \times 3^5$	/	0.768	<b>0.083</b>	0.291	/	<b>1.435</b>	1.403	1.411
	3	8	$8^3 \times 3^8$	/	0.790	<b>0.233</b>	0.503	/	<b>0.270</b>	0.249	0.268
	5	3	$8^5 \times 3^3$	/	6.636	<b>0.264</b>	0.649	/	<b>15.449</b>	14.662	15.135
	8	5	$8^5 \times 3^5$	/	4.134	<b>0.393</b>	0.567	/	<b>5.720</b>	5.174	5.572
	8	3	$8^8 \times 3^3$	/	57.588	<b>0.442</b>	0.612	/	<b>31.024</b>	28.534	29.951
8	5	8	$8^5 \times 3^8$	/	4 650.48	<b>0.648</b>	1.073	/	<b>2.390</b>	2.279	2.343
	8	5	$8^8 \times 3^5$	/	5 239.498	<b>0.798</b>	1.131	/	<b>19.638</b>	18.807	19.176
	8	8	$8^8 \times 3^8$	/	5 841.210	<b>0.963</b>	2.671	/	<b>8.693</b>	8.018	8.617

从表2可以看出,随着规模的增大,每个算法的平均计算时间都在增大.其中:DO算法每次都能够得到纳什均衡解,但是其计算时间随着规模的增大而急剧增加;UWMA算法也能得到纳什均衡解,但其可以求解的规模较为有限,在规模达到 $M = 3, N = 5$ 和 $M = 5, N = 3$ 时,计算时间就已经达到10 000 s以上;DO-NS算法的计算时间略小于DO-TSNS算法,但其得到的平均最差期望收益比DO-TSNS算法的差.

同时,当 $M$ 和 $N$ 不变时,随着 $K$ 的增加,每个算法得到的平均最差期望收益都呈上升趋势.其中,当 $K$ 从小于 $M$ 增长到等于 $M$ 时,无人机可以攻击更多的高价值目标,平均最差期望收益增长幅度较大.例如当 $M = 8, N = 3, K$ 分别为3、5、8时,DO-TSNS算法求得的平均最差期望收益分别为15.772、21.211和29.951.但当 $K$ 从等于 $M$ 增长到大于 $M$ 时,由于无法分配出更多的无人机去攻击高价值目标,只能为无人机选择更高价值的目标进行打击,平均最差期望收益增长幅度较小.又如 $M = 3, N = 3, K$ 分别为3、5、8时,DO-TSNS算法求得的最差期望收益分别为3.928、3.958和3.974.当 $M$ 一定时,随着 $N$ 的增大,每个算法的平均最差期望收益相应减小.这是因为防御方拦截导弹越多,攻击方无人机被拦截的概率越大,攻击方的平均最差期望收益越小.当 $N$ 一定时,随着 $M$ 的增大,每种算法平均最差期望收益相应增加.这是因为攻击方的无人机数量越多,防御方高价值目标被毁伤的概率越大,攻击方的平均最差期望收益也越大.

图2对比了 $K = 3$ 时DO-TSNS算法与DO-NS算法得到的平均最差期望收益和最大值的gap.

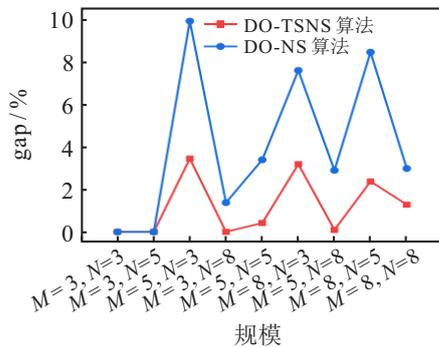


图2  $K = 3$ 时DO-TSNS算法与DO-NS算法得到的平均最差期望收益与最大值的gap对比曲线

从图2中可以看出,DO-TSNS算法得到的平均最差期望收益与最大值的gap始终低于或等于DO-NS算法对应的gap.其中:DO-TSNS算法对应的gap最大值为3.448%,gap平均值为1.198%;而DO-NS算法对应的gap最大值为9.924%.

### 3.3 中规模、大规模实验及其结果分析

与小规模实验类似,中规模实验的 $K、M、N$ 分别设置为10、15、20,大规模实验的 $K、M、N$ 分别设

表3 中大规模实验结果

K	M	N	S	平均计算时间/s		平均最差期望收益	
				DO-NS	DO-TSNS	DO-NS	DO-TSNS
10	10	10	$10^{10} \times 10^{10}$	<b>0.98</b>	3.44	8.230	<b>9.694</b>
	10	15	$10^{10} \times 10^{15}$	<b>1.03</b>	3.69	3.552	<b>4.073</b>
	10	20	$10^{10} \times 10^{20}$	<b>1.29</b>	5.05	0.841	<b>1.117</b>
	15	10	$10^{15} \times 15^{10}$	<b>1.06</b>	4.96	17.703	<b>20.140</b>
	15	15	$10^{15} \times 15^{15}$	<b>1.81</b>	9.01	11.441	<b>15.375</b>
	15	20	$10^{15} \times 15^{20}$	<b>2.65</b>	13.82	7.058	<b>7.806</b>
	20	10	$10^{15} \times 20^{10}$	<b>1.97</b>	11.34	26.103	<b>31.186</b>
	20	15	$10^{15} \times 20^{15}$	<b>2.83</b>	16.62	21.363	<b>24.272</b>
	20	20	$10^{15} \times 20^{20}$	<b>3.06</b>	20.94	14.686	<b>17.300</b>
	15	10	10	$15^{10} \times 10^{10}$	<b>0.86</b>	2.85	8.325
10		15	$15^{10} \times 10^{15}$	<b>1.20</b>	4.51	3.961	<b>4.368</b>
10		20	$15^{10} \times 10^{20}$	<b>1.63</b>	7.30	1.021	<b>1.250</b>
15		10	$15^{15} \times 15^{10}$	<b>1.42</b>	5.17	23.144	<b>27.553</b>
15		15	$15^{15} \times 15^{15}$	<b>1.84</b>	9.44	12.616	<b>15.399</b>
15		20	$15^{15} \times 15^{20}$	<b>2.78</b>	14.10	7.485	<b>8.557</b>
20		10	$15^{20} \times 20^{10}$	<b>1.74</b>	11.40	34.458	<b>42.583</b>
20		15	$15^{20} \times 20^{15}$	<b>3.08</b>	17.17	25.012	<b>29.475</b>
20		20	$15^{20} \times 20^{20}$	<b>3.59</b>	26.03	19.941	<b>22.765</b>
20		10	10	$20^{10} \times 10^{10}$	<b>0.94</b>	3.14	9.025
	10	15	$20^{10} \times 10^{15}$	<b>1.55</b>	6.85	4.196	<b>4.411</b>
	10	20	$20^{10} \times 10^{20}$	<b>1.89</b>	9.56	1.172	<b>1.560</b>
	15	10	$20^{15} \times 15^{10}$	<b>1.56</b>	6.67	25.722	<b>29.893</b>
	15	15	$20^{15} \times 15^{15}$	<b>1.77</b>	11.60	11.669	<b>16.359</b>
	15	20	$20^{15} \times 15^{20}$	<b>2.96</b>	19.20	7.939	<b>8.828</b>
	20	10	$20^{20} \times 20^{10}$	<b>2.05</b>	12.25	45.013	<b>52.649</b>
	20	15	$20^{20} \times 20^{15}$	<b>3.03</b>	24.04	27.309	<b>35.286</b>
	20	20	$20^{20} \times 20^{20}$	<b>3.89</b>	30.12	18.475	<b>26.998</b>
	30	30	30	$30^{30} \times 30^{30}$	<b>35.12</b>	191.52	21.211
30		40	$30^{30} \times 30^{40}$	<b>42.97</b>	233.19	13.223	<b>15.953</b>
30		50	$30^{30} \times 30^{50}$	<b>49.79</b>	394.25	6.125	<b>7.452</b>
40		30	$30^{40} \times 40^{30}$	<b>47.50</b>	304.84	54.509	<b>62.138</b>
40		40	$30^{40} \times 40^{40}$	<b>58.66</b>	491.30	32.099	<b>37.272</b>
40		50	$30^{40} \times 40^{50}$	<b>105.22</b>	889.57	19.215	<b>22.313</b>
50		30	$30^{50} \times 50^{30}$	<b>71.08</b>	589.72	70.163	<b>79.925</b>
50		40	$30^{50} \times 50^{40}$	<b>121.65</b>	916.87	46.359	<b>55.865</b>
50		50	$30^{50} \times 50^{50}$	<b>137.14</b>	997.26	28.585	<b>36.607</b>
40		30	30	$40^{30} \times 30^{30}$	<b>45.04</b>	323.92	24.504
	30	40	$40^{30} \times 30^{40}$	<b>47.78</b>	358.80	13.357	<b>16.975</b>
	30	50	$40^{30} \times 30^{50}$	<b>53.36</b>	457.33	6.068	<b>7.690</b>
	40	30	$40^{40} \times 40^{30}$	<b>74.67</b>	607.92	69.207	<b>78.162</b>
	40	40	$40^{40} \times 40^{40}$	<b>92.97</b>	856.94	37.090	<b>44.905</b>
	40	50	$40^{40} \times 40^{50}$	<b>156.38</b>	1178.76	22.638	<b>25.336</b>
	50	30	$40^{50} \times 50^{30}$	<b>86.71</b>	743.69	80.155	<b>96.192</b>
	50	40	$40^{50} \times 50^{40}$	<b>133.78</b>	960.73	50.334	<b>60.618</b>
	50	50	$40^{50} \times 50^{50}$	<b>168.10</b>	1298.20	38.897	<b>43.708</b>
	50	30	30	$50^{30} \times 30^{30}$	<b>47.66</b>	355.87	25.269
30		40	$50^{30} \times 30^{40}$	<b>62.66</b>	505.68	15.234	<b>18.486</b>
30		50	$50^{30} \times 30^{50}$	<b>122.98</b>	993.71	7.584	<b>8.850</b>
40		30	$50^{40} \times 40^{30}$	<b>89.59</b>	740.61	59.737	<b>71.198</b>
40		40	$50^{40} \times 40^{40}$	<b>140.99</b>	1022.00	40.597	<b>46.178</b>
40		50	$50^{40} \times 40^{50}$	<b>171.39</b>	1306.70	22.965	<b>26.366</b>
50		30	$50^{50} \times 50^{30}$	<b>91.38</b>	780.62	93.774	<b>109.363</b>
50		40	$50^{50} \times 50^{40}$	<b>144.15</b>	1013.61	58.169	<b>66.562</b>
50		50	$50^{50} \times 50^{50}$	<b>185.11</b>	1422.29	40.569	<b>49.299</b>

置为30、40、50,各自得到27种组合.

由于UWMA和DO算法在 $K = 8$ 、 $M = 8$ 、 $N = 8$ 以上规模实验时,72小时内无法得到运行结果,本节只针对DO-TSNS和DO-NS算法进行中大规模实验,对其平均计算时间和平均最差期望收益进行分析.中大规模实验结果如表3所示.

从表3可以看出,随着策略空间规模指数级增长,DO-TSNS算法仍能在有限时间内得出结果,其最长计算时间为1422.29s,短于DO算法在 $K = 8$ 、 $M = 8$ 、 $N = 8$ 规模下5841.21s的计算时间.

当 $M$ 和 $N$ 不变时,随着 $K$ 的增加,平均最差期望收益相应增大;当 $K$ 和 $N$ 不变时,随着 $M$ 的增加,平均最差期望收益相应增大;当 $K$ 和 $M$ 不变时,随着 $N$ 的增加,平均最差期望收益相应减少.可以看出,中大规模实验结果的变化趋势与小规模实验结果的变化趋势相类似,且在 $M$ 、 $N$ 、 $K$ 的值均相同的情况下,DO-TSNS算法所取得的平均最差期望收益均高于DO-NS算法所取得的平均最差期望收益.

综上所述,通过小中大3种规模的实验结果可以看出,本文算法能够适用于不同规模的多无人机对地攻击目标分配问题实例.在小规模问题实例下,相较于DO算法和UWMA算法,本文算法可以在小幅度降低解的质量的情况下大幅减少求解时间;相较于DO-NS算法,本文算法能够求得具有更好质量的解.在中大规模问题实例下,相较于DO-NS算法,本文算法也能够求得具有更好质量的解.

## 4 结论

针对多无人机对地攻击中的武器目标分配问题,本文基于纳什均衡博弈的思想将其建模为零和矩阵博弈模型,并提出了一种基于两阶段邻域搜索的改进Double Oracle算法(DO-TSNS),以寻找该模型的混合策略纳什均衡.该混合策略纳什均衡能够为攻击方提供一种均衡的策略,且基于轮盘赌的策略选择方式不容易被防御方掌握规律,防御方难以猜中攻击方实际采取的具体目标分配方案,从而确保攻击方决策方案的有效性.实验结果表明,相比于现有的DO、UWMA、DO-NS等算法,本文算法可以适用于不同规模的问题实例.

在后续的研究工作中,一方面将在保证解的质量的前提下,研究更合适的邻居策略集生成方法,以进一步提升算法的求解效率;另一方面,将在最大化对防御方高价值目标的期望毁伤价值的基础上,进一步考虑既最大化防御方所消耗拦截导弹的总价值,又最小化攻击方所消耗无人机的总价值,研究相应的多

目标博弈模型和求解算法.

## 参考文献(References)

- [1] Wang J, Luo P C, Zhou J L, et al. Optimizing weapon target assignment in air to ground strike based on adaptive immune genetic algorithm[C]. The 4th International Conference on Information Science and Control Engineering. Changsha, 2017: 520-526.
- [2] Liu W J, Pei Y, Ge Y X, et al. System-of-systems contribution evaluation of ground-attack UCAV based on ABMS[J]. Acta Aeronautica et Astronautica Sinica, 2022, 43(9): 436-449.
- [3] Zhao Y L, Song Y X, Zhang J J, et al. Fuzzy game decision-making of unmanned aerial vehicles air-to-ground attack based on the particle swarm optimization integrating multiply strategies[J]. Control Theory & Applications, 2019, 36(10): 1644-1652.
- [4] Xian B, Li H T. Time-varying formation nonlinear control of distributed multiple UAVs[J]. Control and Decision, 2021, 36(10): 2490-2496.
- [5] 余婧, 雍恩米, 陈汉洋, 等. 面向多无人机协同对地攻击的双层任务规划方法[J]. 系统工程与电子技术, 2022, 44(9): 2849-2857.  
(Yu J, Yong E M, Chen H Y, et al. Bi-level mission planning method for multi-cooperative UAV air-to-ground attack[J]. Systems Engineering and Electronics, 2022, 44(9): 2849-2857.)
- [6] Wang C, Fu G Y, Zhang D Q, et al. Genetic algorithm-based variable value control method for solving the ground target attacking weapon-target allocation problem[J]. Mathematical Problems in Engineering, 2019, 2019: 1-9.
- [7] 陆一平, 李慧慧. 静态武器目标分配问题的攻击界整数规划求解方法[J]. 系统工程理论与实践, 2019, 39(3): 783-789.  
(Lu Y P, Li H H. An attack-number bounded integer programming method for the static WTA problem[J]. Systems Engineering — Theory & Practice, 2019, 39(3): 783-789.)
- [8] Manne A S. A target-assignment problem[J]. Operations Research, 1958, 6(3): 346-351.
- [9] Lloyd S P, Witsenhausen H S. Weapons allocation is NP-complete[C]. Proceedings of the 1986 Summer Computer Simulation Conference. Piscataway: IEEE, 1986: 1054-1058.
- [10] Wu W H, Guo X F, Zhou S Y. Path planning algorithm based on NURBS and GOBL-ACDE[J]. Systems Engineering and Electronics, 2020, 42(5): 1073-1082.
- [11] 夏维, 刘新学, 范阳涛, 等. 基于改进型多目标粒子群优化算法的武器-目标分配[J]. 兵工学报, 2016, 37(11): 2085-2093.  
(Xia W, Liu X X, Fan Y T, et al. Weapon-target assignment with an improved multi-objective particle swarm optimization algorithm[J]. Acta Armamentarii,

- 2016, 37(11): 2085-2093.)
- [12] Li Z W, Chang Y Z, Kou Y X, et al. Approach to WTA in air combat using IAFA-IHS algorithm[J]. *Journal of Systems Engineering and Electronics*, 2018, 29(3): 519-529.
- [13] 王晓光, 章卫国, 陈伟. 无人机编队超视距空战决策及作战仿真[J]. *控制与决策*, 2015, 30(2): 328-334. (Wang X G, Zhang W G, Chen W. BVR air combat decision making and simulation for UAV formation[J]. *Control and Decision*, 2015, 30(2): 328-334.)
- [14] Luo W L, Lv J H, Liu K X, et al. Learning-based policy optimization for adversarial missile-target assignment[J]. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2022, 52(7): 4426-4437.
- [15] 朱建文, 赵长见, 李小平, 等. 基于强化学习的集群多目标分配与智能决策方法[J]. *兵工学报*, 2021, 42(9): 2040-2048. (Zhu J W, Zhao C J, Li X P, et al. Multi-target assignment and intelligent decision based on reinforcement learning[J]. *Acta Armamentarii*, 2021, 42(9): 2040-2048.)
- [16] 王玮, 刘兴林, 王军, 等. 信息化条件下海上编队区域防空目标分配方法[J]. *系统工程理论与实践*, 2015, 35(4): 1011-1018. (Wang W, Liu X L, Wang J, et al. Method of area anti-aircraft weapon target assignment for the warship formation under informationized conditions[J]. *Systems Engineering—Theory & Practice*, 2015, 35(4): 1011-1018.)
- [17] Guo D, Liang Z X, Jiang P, et al. Weapon-target assignment for multi-to-multi interception with grouping constraint[J]. *IEEE Access*, 2019, 7: 34838-34849.
- [18] Gao K Y, Yan X B. Study on the optimal strategy of missile interception[J]. *IEEE Access*, 2021, 9: 22239-22252.
- [19] Davis M T, Robbins M J, Lunday B J. Approximate dynamic programming for missile defense interceptor fire control[J]. *European Journal of Operational Research*, 2017, 259(3): 873-886.
- [20] Shalumov V, Shima T. Weapon-target-allocation strategies in multiagent target-missile-defender engagement[J]. *Journal of Guidance, Control, and Dynamics*, 2017, 40(10): 2452-2464.
- [21] Han C Y, Lunday B J, Robbins M J. A game theoretic model for the optimal location of integrated air defense system missile batteries[J]. *Informatics Journal on Computing*, 2016, 28(3): 405-416.
- [22] Boardman N T, Lunday B J, Robbins M J. Heterogeneous surface-to-air missile defense battery location: A game theoretic approach[J]. *Journal of Heuristics*, 2017, 23(6): 417-447.
- [23] Xu W Q, Chen C, Ding S X, et al. A bi-objective dynamic collaborative task assignment under uncertainty using modified MOEA/D with heuristic initialization[J]. *Expert Systems with Applications*, 2020, 140: 112844.
- [24] Naseem A, Shah S T H, Khan S A, et al. Decision support system for optimum decision making process in threat evaluation and weapon assignment: Current status, challenges and future directions[J]. *Annual Reviews in Control*, 2017, 43: 169-187.
- [25] 杜永浩, 王凌, 邢立宁. 空天无人系统智能规划技术综述[J]. *系统工程学报*, 2020, 35(3): 416-432. (Du Y H, Wang L, Xing L N. Intelligent planning technologies for unmanned aerospace system: A literature review[J]. *Journal of Systems Engineering*, 2020, 35(3): 416-432.)
- [26] 任敏, 王克波, 沈林成. 多UAV协同突防规划与仿真[J]. *控制与决策*, 2011, 26(1): 157-160. (Ren M, Wang K B, Shen L C. Planning and simulation for multi-UAV cooperative penetration mission[J]. *Control and Decision*, 2011, 26(1): 157-160.)
- [27] Wang X R, Guo Q Y, An B. Stop nuclear smuggling through efficient container inspection[C]. *Proceedings of the 16th Conference on Autonomous Agents and Multi Agent Systems*. New York: ACM, 2017: 669-677.
- [28] Yang Z, Zhu J W, Teng L, et al. A double oracle algorithm for allocating resources on nodes in graph-based security games[J]. *Multimedia Tools and Applications*, 2018, 77(9): 10961-10977.
- [29] Bosansky B, Kiekintveld C, Lisy V, et al. An exact double-oracle algorithm for zero-sum extensive-form games with imperfect information[J]. *Journal of Artificial Intelligence Research*, 2014, 51: 829-866.
- [30] Rosenmüller J. On a generalization of the lemke-howson algorithm to noncooperative N-person games[J]. *SIAM Journal on Applied Mathematics*, 1971, 21(1): 73-79.
- [31] Ma Y Y, Wang G Q, Hu X X, et al. Cooperative occupancy decision making of multi-UAV in beyond-visual-range air combat: A game theory approach[J]. *IEEE Access*, 2019, 8: 11624-11634.
- [32] Ma Y Y, Wang G Q, Hu X X, et al. Weapon target assignment method for multiple UAVs in beyond-visual-range air combat[J]. *Chinese Journal of Management Science*, 2022, 30(3): 248-257.
- [33] 李杰. “萨德”能对中程弹道导弹形成巨大威胁[EB/OL]. [2023-07-23]. [https://www.kepuchina.cn/mil/expert/201710/t20171026\\_238707.shtml](https://www.kepuchina.cn/mil/expert/201710/t20171026_238707.shtml).

## 作者简介

王浩丞(1997—), 男, 硕士生, 从事武器目标分配、零和博弈等研究, E-mail: 2020110827@mail.hfut.edu.cn;

罗贺(1982—), 男, 教授, 博士生导师, 从事多无人平台协同优化与智能决策等研究, E-mail: luohe@hfut.edu.cn;

马滢滢(1994—), 女, 博士生, 从事武器目标分配、零和博弈等研究, E-mail: mayingying@mail.hfut.edu.cn;

王国强(1982—), 男, 副教授, 硕士生导师, 从事多无人机协同优化与智能决策等研究, E-mail: gqwang2017@hfut.edu.cn.