

# 基于数据驱动的多智能体规划识别与博弈围捕

王勇<sup>1,2</sup>, 李尚昊<sup>3</sup>, 徐伯辰<sup>3</sup>, 方浩<sup>3</sup>

(1. 江苏自动化研究所, 江苏, 连云港 222061; 2. 西北工业大学 无人技术研究院, 陕西, 西安 710129;

3. 北京理工大学 自动化学院, 北京 100081)

**摘要:** 在进行实时对抗的任务中, 对于敌方的动作识别较为困难, 需要根据对方的移动轨迹或行为来分析对方的意图, 预测其未来目标, 构建规划策略库. 针对此问题, 提出基于数据驱动的多智能体识别算法, 该算法首先采用基于自动机的特征提取方法, 获得规划需要的位置和任务信息; 然后将规划识别问题转换为多分类问题, 并从单智能体角度切入, 给出了一种基于极端梯度提升 (extreme gradient boosting, XGBoost) 的多分类模型; 之后, 对于多智能体之间可能存在的合作行为, 使用无监督学习的一种基于密度对噪声鲁棒的空间聚类算法 (density-based spatial clustering of applications with noise, DBSCAN) 对多智能体进行分簇, 以促进协同合作. 对于同簇智能体, 构建了一种针对多智能体的多分类模型, 完成对多智能体的目标预测. 在获悉敌方目标后, 提出基于博弈的围捕逼停算法, 构建非合作动态博弈模型, 通过求解纳什均衡得到应对敌方的最优策略. 最后, 通过仿真验证了所提出算法的有效性.

**关键词:** 数据驱动; 多智能体; 规划识别; 非合作博弈; 纳什均衡

中图分类号: TP13

文献标志码: A

文章编号: 1001-0645(2025)05-0521-10

DOI: 10.15918/j.tbit1001-0645.2024.167

## Multi-Agent Planning Recognition and Game Encirclement Based on Data Driving

WANG Yong<sup>1,2</sup>, LI Shanghao<sup>3</sup>, XU Bochen<sup>3</sup>, FANG Hao<sup>3</sup>

(1. Jiangsu Automation Research Institute, Lianyungang, Jiangsu 222061, China;

2. Unmanned System Research Institute, Northwestern Polytechnical University, Xi'an, Shaanxi 710129, China;

3. School of Automation, Beijing Institute of Technology, Beijing 100081, China)

**Abstract:** To analyze and predict the opponent's intentions and future targets scientifically in real-time adversarial tasks, a multi-agent recognition algorithm was proposed based on data driving. Firstly, taking a feature extraction method with automata, the arithmetic was arranged to obtain position and task information from samples for planning. Then, transforming the planning recognition into a multi-classification problem, a multi-classification model was developed based on extreme gradient boosting (XGBoost) to start with a single agent. And then, considering collaboration potential among agents, the density-based spatial clustering of applications with noise (DBSCAN) algorithm was employed for multi-agent clustering and collaborating. And a multi-classification model was then developed for clustered multi-agent systems to predict goals within the same cluster. After identifying the opponent's targets, an encirclement and stop strategy was designed based on game theory. Finally, a non-cooperative dynamic game model was constructed, and the Nash equilibrium solution was used to derive the optimal strategy. Simulation results demonstrate the effectiveness of the proposed algorithm.

**Key words:** data driven; multi-agent; planning recognition; non-cooperative game; Nash equilibrium

收稿日期: 2024-11-05

基金项目: 国家重点研发计划资助项目(2022YFA1004703)

作者简介: 王勇(1979—), 男, 研究员, E-mail: wyong\_jari@aliyun.com.

通信作者: 方浩(1973—), 男, 教授, E-mail: fangh@bit.edu.cn.

随着人工智能技术的发展,无人系统逐渐智能化,由多个智能体组成的无人系统将成为多体系统的主要组成,在智能化的多体对抗中至关重要<sup>[1-4]</sup>.在多智能体对抗时,敌我双方都会隐藏本方的攻击意图.为了响应敌方作战计划,多智能体首要的目标任务就是要识别出对手的作战意图,作出预推测和实时响应.针对规划识别问题,KAUTZ等<sup>[5]</sup>提出一种通用的规划识别模型,将每一个观测到的行为看作更高层规划的一个子任务,然后用高层任务规划去解释底层行为,由此,规划识别的过程,也就变成了根据观察到的行为,推测其前向目的的过程.VIL-AIN<sup>[6]</sup>致力于使识别计划更像是解析文本这一概念变得精确,采用了语法分析的方法进行计划识别.PYNADATH等<sup>[7]</sup>提出了基于概率状态独立语法的规划识别方法,该方法引入概率状态相关文法来表示智能体的规划生成过程.KANTHARAJU<sup>[8]</sup>研究从训练数据自动构建分解模型的学习技术,用于分层规划和规划识别.还有人将规划识别过程看作规划,有关这类方法 MENEGUZZI 等<sup>[9]</sup>做了一定总结.ORI-EL等<sup>[10]</sup>提出基于交错时间搜索的探索域规划识别算法.贾镇泽等<sup>[11]</sup>设计了一套敌机战术机动动作的在线识别方法,通过规划识别辅助我方平台进行任务规划与决策.以上这些规划识别方法,大都基于完备的规划库进行层次化描述,无法适应规划库不完备的复杂的多智能体对抗场景.

而且,规划识别也对任务的时序性提出了较高的要求.考虑到多无人平台协同时时间在时间层面的约束,研究者们将研究重点放在了时序逻辑约束下的多无人平台协同规划识别问题上.其主要思想是,利用如线性时序逻辑语言(LTL)<sup>[12]</sup>、信号时序逻辑语言(STL)<sup>[13]</sup>等形式化语言表述需求,以及自动机模型表示方案,通过图搜索方式生成相关策略.丁明等<sup>[14]</sup>针对业务流程规划问题,使用线性时序逻辑表示需求,利用形式化设计和自动机模型求得明确的执行路径.聂兆伟等<sup>[15]</sup>针对武器装备系统安全性问题,也利用形式化建模语言进行分析.因此,为了使得目标规划识别的效果更好,可以将规划识别需求表述为LTL定义的形式化描述,进而生成有限状态转移系统(FTS)来支持后续的围捕工作.

在目标规划识别的基础上,对目标进行围捕是多智能体作战的一种重要方式.将博弈思想引入到多智能体攻防对抗中,对于整合、优化集群内部资源,提高多智能体协同任务效率有着重要意义<sup>[16-17]</sup>.而

非合作博弈中,纳什均衡是一个重要概念,赖中安等<sup>[18]</sup>将矩阵博弈的思想应用于水面舰艇的电子对抗中;范衡等<sup>[19]</sup>通过共识主动性完成目标搜索与围捕.以上这些研究内容更多是从任务分配层面研究对抗过程,或研究了无人平台对抗过程中的攻击与逃逸对抗,没有研究基于多智能体协同下的逼停任务.

针对无人装备发展,本文根据无人集群博弈对抗作战需求,针对目标的规划识别首先给出了一种基于XGBoost的多分类模型;之后使用无监督学习方法DBSCAN对多智能体进行聚类分簇,并提出了一种针对多智能体的多分类模型,实现对多智能体的目标预测.在获悉敌方目标后,为执行最优策略,本文提出基于博弈的围捕逼停算法,构建非合作动态博弈模型,通过求解纳什均衡得到应对敌方的最优策略.

由于开放实时对抗中待识别智能体的非合作性,无法直接获得其当前任务信息和目标信息,因此,需要根据对方的移动轨迹或行为来分析对方的意图,预测其未来目标.获知敌方目标后,便可以针对该目标进行保护,并采取博弈算法对敌方进行围捕拦截.

图1为具体行动的示意图,其中 $a$ 、 $d$ 、 $j$ 、 $k$ 、 $p$ 、 $u$ 、 $v$ 等表示我方待保护平台,它们具有不同的防护等级,飞机为敌方入侵单位,线条显示了它们的移动轨迹.

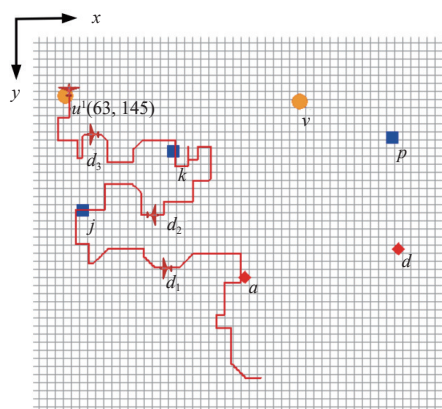


图1 智能体移动轨迹

Fig. 1 Agent track

具体将任务定义如下.

**定义1** 已知多智能体系统 $A$ 的规划库 $P$ 和大量智能体行为轨迹数据 $D$ .规划库 $P$ 中含有多智能体系统可能会执行的、多个相互独立且具有时序逻辑特性的任务 $\varphi_i$ .数据 $D$ 中包含了智能体执行规划库中任务的轨迹信息.根据观测到的多智能体动作轨迹 $M$ ,对每个智能体的未来目标进行预测.得到敌方

单位的未来目标的预测后,我方需要派遣多个智能体对该目标进行保护,同时拦截并围捕敌方单位,阻止其靠近我方平台.

## 1 基于 XGBoost 的多智能体规划识别算法

### 1.1 基于 Büchi 自动机的特征提取

时序逻辑语言和自动机的有机结合为多智能体特征提取提供了新的思路<sup>[20-21]</sup>,为了准确地给多分类模型提供特征输入,需要对智能体历史轨迹数据  $D$  进行特征提取.对于时序逻辑任务下的多智能体轨迹的特征信息,可以分为两部分:位置信息和任务信息.其中,位置信息包括智能体的坐标、姿态;任务信息包括规划库中每个任务的执行状态、已经访问过的目标点.

为了准确地描述多智能体系统的移动轨迹,结合工作环境特点,将其建模成有限状态转移系统 (FTS).

考虑数据  $D$  中的样本  $d_i = \{t, (x, y), f_{\text{arrived}}, f_{\text{goal}}\}$ , 对于位置信息,取  $x$  轴坐标  $x_i$ ,  $y$  轴坐标  $y_i$ , 计算智能体姿态  $\alpha_i$ , 有

$$\alpha_i = \arctan\left(\frac{y_i - y'_i}{x - x'_i}\right) \quad (1)$$

式中  $x'_i$  和  $y'_i$  为上一时刻智能体的  $x$  坐标和  $y$  坐标.因此,可以得到位置信息特征  $f_{\text{PosFeatures}} = [x_i, y_i, \alpha_i]$ .

根据 Büchi 自动机定义可知, Büchi 自动机中的状态  $q$  标志着任务执行的进度.因此,对于任务信息,可以提取样本  $d_i$  中的到达信息  $f_{\text{arrived}}$ , 在自动机中做映射,将得到的自动机状态作为每个任务的状态信息.假设规划库  $P$  含有的任务集合为  $\{\varphi_1, \varphi_2, \dots, \varphi_p\}$ , 则对于每个任务,使用时序逻辑语言进行描述,并转换成对应自动机,可以得到  $P = \{A_{\varphi_1}, A_{\varphi_2}, \dots, A_{\varphi_p}\}$ . 通过算法 1 得到任务信息特征

$$f_{\text{TaskFeatures}} = [q_1, q_2, \dots, q_p]$$

式中  $q_i$  对应任务  $\varphi_i$  的状态.

算法 1 任务信息特征获取

**输入:** 自动机  $P$ , 数据信息  $d_i$

**输出:** 任务信息特征  $f_{\text{TaskFeatures}}$

- 1:  $f_{\text{TaskFeatures}} = \emptyset$
- 2: 取出  $d_i$  中的  $f_{\text{arrived}}$
- 3: for each 自动机  $A_{\varphi_i} \in P$  do
- 4:  $s$  = 自动机  $A_{\varphi_i}$  的初始状态  $Q_0$
- 5:  $m_{\text{imp}} = 0$

- 6: for  $s_{\text{next}} \in \delta(s)$  do
- 7: if  $\delta(s, s_{\text{next}}) \subseteq f_{\text{arrived}}$  then
- 8: if  $|\delta(s, s_{\text{next}})| > m_{\text{imp}}$  then
- 9:  $s = s_{\text{next}}$
- 10:  $m_{\text{imp}} = |\delta(s, s_{\text{next}})|$
- 11:  $f_{\text{TaskFeatures}} \leftarrow s$
- 12: return  $f_{\text{TaskFeatures}}$

因为规划库模型具有不完备性,可能未包含智能体执行的任务.考虑到任务执行的体现方式是已经访问过的目标点序列,因此,可以将  $f_{\text{arrived}}$  作为任务信息特征的补充.为了保持特征维度的固定,需要对  $f_{\text{arrived}}$  做 01 转换,记为  $f_{\text{arrived}01}$ . 根据 FTS 的定义,设定  $f_{\text{arrived}01}$  的长度为  $|I|$ , 已经到达的点记为 1, 没有到的点记为 0.

对于样本  $d_i = \{t, (x, y), f_{\text{arrived}}, f_{\text{goal}}\}$ , 经过特征工程得到的特征  $f_{\text{Features}}$  为

$$f_{\text{Features}} = [x, y, f_{\text{TaskFeatures}}, f_{\text{arrived}01}] \quad (2)$$

### 1.2 针对单智能体的多分类模型

由式 (2) 可以得到每个样本的特征, 样本标签为该时刻目标点. 根据 FTS 定义可知, 对于目标预测问题, 标签的类别数为  $|I|$ , 可以将该问题建模为多分类问题. 在实际应用中, 训练样本往往存在一定的噪音偏差, 如果过度的追求模型在训练集上的准确率而采用较复杂的模型, 会使得模型过拟合从而在真实任务中表现较差, 并且增加训练成本. 但过于简单的模型在数据分布较复杂的环境下, 训练过程中会产生欠拟合的情况, 无法准确估计数据分布. 因此, 综合考虑时序逻辑任务下的多智能体规划识别问题的复杂度和实时性要求, 选择的多分类模型为 XGBoost, 模型算法如下:

假设有  $|D|$  个样本, 特征维度为  $|f_{\text{TaskFeatures}}|$ , 标签类别个数为  $|I|$ , 即  $(d_i, l_i)_{i=1}^{|D|}$ . 其中,  $l_i \in \{0, 1, \dots, |I|\}$ . 令  $l_{i,k}$  表示第  $i$  个样本在第  $k$  类上的真实得分, 有

$$l_{i,k} = \begin{cases} 1 & l_i = k \\ 0 & \text{其他} \end{cases} \quad (3)$$

令  $l^{(\wedge)}_{i,k}$  表示预测的第  $i$  个样本在第  $k$  类上的得分. 针对多分类问题, 采用交叉熵损失函数, 定义为

$$L(l, \hat{l}) = - \sum_{i=1}^{|D|} l(d_i) \lg(\hat{l}(d_i)) \quad (4)$$

对于多分类问题, 需要对样本  $d_i$  每个可能的类都训练一个 CART 分类树, 即在每轮的训练的时候需要同时训练  $|I|$  颗树. 在同一轮的训练中, 对于第  $k$  颗树针对样本  $d_i$  是否属于第  $k$  类, 其输入为  $(d_i, l_{i,k})$ . 多

分类情况下的 XGBoost 模型训练算法如算法 2 所示.

算法 2 多分类 XGBoost 训练算法

**输入:** 训练数据  $D$ , 迭代次数  $M$

**输出:** 分类树  $F_k$

```

1:  $F_k^0 = 0, k = 1, 2, \dots, |I|$ 
2: for  $t = 1, 2, \dots, M$  do
3: for  $k = 1, 2, \dots, |I|$  do
4:  $\hat{f}_{i,k} = \hat{f}_{i,k}^{t-1} + f_k^t(d_i), \hat{l}_{i,k} = \hat{l}_{i,k}^{t-1} + f_k^t(d_i)$ 
5:  $f_k^t(d_i)$  根据最小化平方差损失函数  $L_k^t$  确定.
6:  $F_k^t = F_k^{t-1} + f_k^t, F_k^t = F_k^{t-1} + f_k^t$ 
7: return  $F_k, k = 1, 2, \dots, |I|$ 

```

算法 2 中步骤 5 的平方差损失函数  $L_k^t$  为

$$L_k^t = \sum_{i=1}^{|D|} \left( l_{i,k} - (\hat{l}_{i,k} + f_k^t(d_i)) \right)^2 + \Omega(f_k^t) \quad (5)$$

式中:  $\Omega(f_k^t)$  为惩罚项,  $\Omega(f_k^t) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T \omega_j^2 \Omega(f_k^t) = \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T \omega_j^2$ ;  $T$  为叶子个数;  $\omega_j$  为叶子节点输出值.

训练过程中需要找到一个节点划分, 从而得到一个最优结构的树. 采用贪心算法<sup>[22]</sup>寻找划分点.

测试阶段, 根据算法 2 训练得到  $|I|$  个 boost 树, 每一个 boost 树对测试样本  $d_i$  的预测值为  $F_j(d_i)$ , 通过 softmax 公式计算得到测试样本  $d_i$  属于第  $j$  类的概率为

$$p(d_i, j) = \frac{\exp(F_j(d_i))}{\sum_{j=1}^{|I|} \exp(F_j(d_i))} \quad (6)$$

因此, 测试样本  $d_i$  的目标预测为  $\arg\max\{p(d_i, j) | j = 1, 2, \dots, |I|\}$ .

### 1.3 结合 DBSCAN 算法的多智能体多分类模型

基于 XGBoost 的多分类模型的输入为单个智能体, 没有考虑实际应用中智能体协同执行任务的情况. 因此, 需要将多分类模型扩展到多智能体领域. 扩展算法分为两步, 第一步, 对多智能体系统潜在协同关系进行辨识; 第二步, 改进多分类模型, 使其对具有协同关系的智能体进行目标预测.

多智能体规划识别问题的研究对象具有非合作性, 即无法事先知道智能体的数量. 同时, 多智能体协同关系具有多样性. 因此, 针对未知数量的多智能体规划识别中的协同关系辨识问题, 不论是从模型还是数据角度出发, 都具有一定的难度. 本节将从机器学习的角度, 采用无监督学习方法, 对潜在协同关系进行辨识. 由于多智能体系统的组织结构未知, 分组数量不定, 因此, 选择 DBSCAN 聚类算法进行目

标分簇, 具体算法如算法 3 所示.

算法 3 基于 DBSCAN 的目标分簇算法

**输入:** 多智能体系统  $A$ , 半径  $\varepsilon$ , 密度阈值  $m_{pts}$

**输出:** 分簇结果  $r$

```

1:  $k = 0, r = \emptyset$ 
2: 标记多智能体系统  $A$  中的每个智能体为  $F_{unvisited}$ 
3: while 存在  $F_{unvisited}$  智能体 do
4: 随机选择一个  $F_{unvisited}$  智能体  $a_i$ 
5: 标记为  $a_i$  为  $F_{visited}$ 
6:   if  $a_i$  的  $\varepsilon$  领域内至少有  $m_{pts}$  个智能体 then
7:     创建一个新分簇  $C_k$ , 并将  $a_i$  加入  $C_k$ 
8:     令  $N$  为  $a_i$  的  $\varepsilon$  领域中的智能体集合
9:     for  $a_j \in N$  do
10:      if  $a_j$  为  $F_{unvisited}$  do
11:        标记  $a_j$  为  $F_{visited}$ 
12:      if  $a_j$  的  $\varepsilon$  至少有  $m_{pts}$  个智能体 then
13:        把这些智能体加入  $N$ 
14:      if  $a_j$  不属于任何一个簇 then
15:        将  $a_j$  加入簇  $C_k$ 
16:    $r \leftarrow C_k$ 
17:    $k = k + 1$ 
18: else 将  $a_i$  标记为  $F_{noise}$ 
19: return  $r$ 

```

通过算法 3 得到待识别多智能体系统的分簇结果, 假定同簇的智能体可能存在协同关系, 不同簇智能体不存在协同关系. 针对多智能体目标预测的多分类模型如算法 4 所示. 首先, 通过分簇算法得到多智能体潜在协同关系分组 (步骤 1). 然后, 对于每个智能体  $a_i \in A$ , 根据式 (2) 提取特征, 并采用算法 2 预测目标为  $g_i$  (步骤 2 ~ 步骤 5). 考虑可能发生的协同, 对同簇的智能体, 以智能体  $a_i$  到预测目标  $g_i$  的距离  $D_{dist}(a_i, g_i)$  为索引, 对同簇的元素进行升序排序 (步骤 6). 对于同簇内的智能体, 当智能体  $a_{i-1}$  到达目标点  $g_{i-1}$  时, 对智能体  $a_i$  的输入特征  $f_i$  进行更新, 并重新用算法 2 进行目标预测 (步骤 7 ~ 步骤 10).

算法 4 多智能体多分类模型

**输入:** 已训练好的 XGBoost 模型  $F$ , 待识别多智能体系统  $A$

**输出:** 目标预测结果  $G$

```

1: 根据算法 3 得出多智能体分簇结果  $C$ , 记为  $(a, c)$ ,
 $\forall a_i \in A, c_j \in C$ 
2: for  $a_i \in A$  do
3: 根据式 (2) 提取  $a_i$  特征, 记为  $f_i$ 

```

```

4:  $g_i = F(f_i)$ 
5:  $G \leftarrow (a_i, c_j, g_i)$ 
6: 对同簇的智能体进行排序,  $F_{\text{sorted}}(c_j, \text{by } D_{\text{dist}}(a_i, g_i))$ ,
 $\forall c_j \in C$ 
7: for  $c_j \in C$  do
8:   for  $i = 1, 2, \dots, |c_j|$  do
9:    假设  $a_{i-1}$  已经到达  $g_{i-1}$ , 更新  $a_i$  对应特征  $f_i$ 
10:  $g_i = F(f_i)$ 
11: return  $G$ 

```

综上所述, 数据驱动下的多智能体规划识别算法框架如图2所示. 通过自动机模型对数据进行特征提取, 采用 XGBoost 模型训练得到针对单智能体目标预测的多分类模型. 在考虑多智能体时, 使用无监督学习的 DBSCAN 分簇算法对多智能体潜在协同关系进行辨识. 对于同簇多智能体, 将单智能体多分类模型扩展为针对多智能体目标预测的多分类模型.

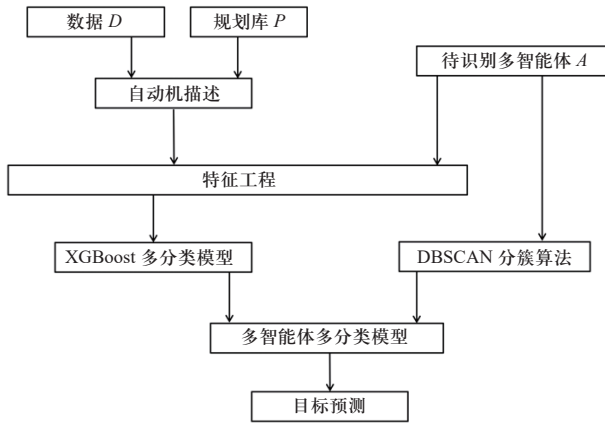


图2 数据驱动下的多智能体规划识别算法框架

Fig. 2 Algorithm framework of multi-agent planning recognition with data driven

## 2 博弈围捕逼停算法

对于已识别的入侵目标, 采用博弈围捕逼停算法, 得到任务规划结果, 从而调用一台或多台无人平台对入侵目标实施截停, 并进行围捕. 为精确执行上述动作, 以应对目标可能出现的对抗策略, 采用基于非合作博弈的无人集群围捕算法, 考虑对手为具有一定智能的理性玩家, 通过合理设计策略空间与收益函数, 将无人集群围捕截停问题建模为双人非合作博弈, 并计算纳什均衡策略. 博弈采用重复博弈, 循环步长设置为固定时间步长, 直至任务完成.

本算法将多智能体围捕的追逃博弈问题建模为多人非零和无限动态非合作博弈问题. 在围捕活动

的前提下, 参与智能体行为策略具体表现为如何进行下一步行动, 即速度方向、大小该如何选取. 策略集的设计中需考虑到实际运动学和动力学限制, 无人平台无法做大角度机动, 在一定角度范围内, 采用离散角度及速度值, 作为双方策略集. 在所有参与智能体都选择各自策略后, 为了评估所形成的博弈局势, 可设计相应的评价函数, 即收益函数  $I$ , 以供参与智能体进行策略选择. 通过上一节的算法可以获得入侵智能体的进攻目标(即我方保护区, 后续采用此种说法), 因此我方应首先派出智能体对其进行拦截, 逼停对方以阻止其靠近我方保护区. 为此我方平台应尽可能处于入侵智能体与我方保护区的中间, 并不断靠近入侵智能体, 以进行下一步围捕动作.

故而可以设计收益函数包括如下三方面:

① 跟踪目标收益.

开始时我方智能体与入侵智能体较远, 这时越接近敌方, 越有利于逼停, 以进行下一步围捕, 因此距离入侵智能体越近收益越高.

$$I_i^1 = -\frac{L_1}{d_{\text{goal}}} \quad (7)$$

式中  $d_{\text{goal}}$  为智能体到入侵智能体的距离.

② 团队避障收益.

在实现逼停的过程中, 团队内部智能体之间也应该避免碰撞. 可以设置安全距离  $L_2$ , 当超过该安全距离时, 不需要考虑彼此间的安全问题, 进入该范围后, 则应当彼此之间保持远离, 以避免碰撞.

$$I_i^2 = \begin{cases} 1 - \frac{L_2}{d_m} & d_m < L_2 \\ 0 & \text{其他} \end{cases} \quad (8)$$

式中:  $d_m$  表示智能体  $i$  到团队内部任一智能体的距离, 对团队内部的每一个智能体都应计算该收益.

③ 逼停收益.

为了有效阻止入侵智能体靠近我方保护区, 我方的智能体必须尽可能位于入侵智能体与我方保护区中间, 以入侵智能体为顶点, 分别向我方智能体和对方保护区做连线, 所形成的角度尽可能小, 以形成对入侵智能体的包夹之势.

$$I_i^3 = \begin{cases} \frac{\theta_{\text{threshold}}}{d_m} & \theta_{\text{threshold}} < \theta_m \\ 0 & \text{其他} \end{cases} \quad (9)$$

式中:  $\theta_{\text{threshold}}$  为夹角阈值, 在一定角度内对敌人的包夹达到最优效果, 小于该值所带来的提升则不明显;  $\theta_m$  为以入侵智能体为顶点分别向我方智能体和我方

保护区做连线所形成的角度, 其中我方保护区即上一节中所识别出的入侵智能体的目标.

故可设第  $i$  个参与智能体的收益函数为

$$I_i = w_1 I_i^1 + w_2 I_i^2 + w_3 I_i^3 \quad (10)$$

式中:  $w_1, w_2, w_3$  为分收益的权重参数, 通过调整参数大小, 可调整相应收益所占比重, 进而调整相应策略倾向的比重, 最终实现较理想的逼停效果. 结合多智能体围捕的实际背景, 为了实现对非合作目标的有效围捕, 同时协调围捕智能体之间的关系, 设计各项收益如下:

#### ① 智能体围捕收益.

围捕智能体和目标之间在不发生碰撞的前提下, 距离目标越近, 围捕越有效, 同时设置安全距离为  $L_1$ , 为避免围捕智能体进入碰撞危险区域, 当距离小于安全距离时, 距离越近收益越小.

$$I_i^1 = \begin{cases} \frac{L_2}{d_{\text{goal}}} & d_{\text{goal}} > L_1 \\ -\frac{L_2}{d_{\text{goal}}} & \text{其他} \end{cases} \quad (11)$$

式中  $d_{\text{goal}}$  为智能体到目标的距离.

#### ② 团队避障收益.

此部分收益与逼停过程一致.

逃跑智能体反围捕收益函数的设计分为两部分, 避碰收益和突围收益, 与上述过程类似.

由博弈模型可知, 各个局中人选择的策略对其他局中人的收益影响很大, 因为所有局中人同时做出决策, 且只能确定自己所选的策略, 对其他局中人要执行的策略并不知情, 故局中人的最优策略是一种平衡, 对应于博弈论中的纳什均衡解, 即无论各个局中人怎么改变自己的策略, 都不能增加自身收益的策略组合.

考虑到围捕环境的诸多因素, 我方最终采取的策略应该是一个确定性策略, 而不是策略的概率值, 因此本算法没有计算混合策略, 而是求解我方多智能体的纯纳什均衡策略. 根据上述分析, 围捕策略的求解可以考虑为一个多层的优化问题, 每个个体在计算自身策略时受到他人策略的影响, 同时应保证他人策略的最优性. 这样的优化问题在实际求解时具有很大难度, 尤其当平台数增加时, 问题规模陡增. 并且当策略集的动作考虑为连续变量时, 传统的矩阵形式博弈求解方法也无法适用. 为了满足实时性需求, 本算法使用迭代最优响应法(算法 5), 以求得一个近似纳什均衡策略. 在求得纳什均衡策略后,

我方智能体应采用该策略, 同时更新环境信息, 并重复上述策略求解过程, 直至完成跟踪围捕任务.

算法 5 迭代最优响应

输入: 初始解  $u$

输出: 纳什均衡策略  $u$

- 1: 初始化  $u, u_{\text{prev}}$
- 2: while  $\|u - u_{\text{prev}}\|$  do
- 3: for  $i \leftarrow 1$  to  $m$  do
- 4:  $u_i^* = \max_{u_i} G_i(x, u_i, u_{-i})$
- 5:  $u_{\text{prev}} = u, u = u^*$
- 6: Return  $u$

## 3 仿真实验

### 3.1 多智能体规划识别仿真

为了验证所设计算法在目标预测方面的有效性, 针对数据驱动下的多智能体规划识别算法进行仿真实验验证.

本节采用实时对抗的仿真环境如图 3 所示, 多智能体规划识别算法应用于识别敌方智能体的攻击目标. 本节仿真实验的前提是已有大量智能体历史任务执行轨迹, 通过提取轨迹中的位置点并记录, 得到原始数据  $D$ .

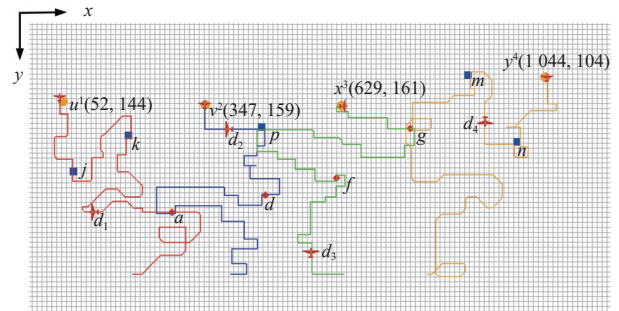


图 3 多智能体规划识别仿真场景

Fig. 3 Simulation scenario for multi-agent planning recognition

在敌方智能体进攻时, 完备的规划库  $P$  为

$$\varphi_1 = \Diamond u \wedge (\neg u \cup (j \wedge k)) \wedge (\neg (j \vee k)) \cup a)$$

$$\varphi_2 = \Diamond v \wedge (\neg v \cup (a \wedge d \wedge p))$$

$$\varphi_3 = \Diamond z \wedge (\neg z \cup (f \wedge p \wedge g))$$

$$\varphi_4 = \Diamond y \wedge (\neg y \cup (m \wedge n)) \wedge (\neg (m \vee n)) \cup g)$$

根据式 (2) 对数据  $D$  中的样本点进行特征提取, 对于目标预测问题, 样本标签的类别数为 13 个, 分别为  $\{a, d, f, g, j, k, m, n, p, u, v, y, z\}$ .

已有的执行任务轨迹为 4 000 条, 提取样本点个数为 1 280 000, 按照 7 : 3 的比例随机将总的样本划

分为训练集和测试集两部分. 利用训练集进行 XGBoost 模型的训练, 在测试集上评估模型的效果. 训练过程中采用交叉熵作为损失函数, 测试集的评估指标为准确率. 为验证数据中噪声对模型效果的影响, 对样本中的  $f_{\text{arrived}}$  进行随机更改, 噪声的占比分别选为 5%、20% 和 40%. 同时, 为验证规划库不完备对所设计的多智能体规划识别算法的影响, 选取多种不完备规划库进行实验, 分别为:

25% 缺失:  $P_{234} = \{\varphi_2, \varphi_3, \varphi_4\}$ ,  $P_{134} = \{\varphi_1, \varphi_3, \varphi_4\}$

50% 缺失:  $P_{34} = \{\varphi_3, \varphi_4\}$ ,  $P_{14} = \{\varphi_1, \varphi_4\}$ ,  $P_{12} = \{\varphi_1, \varphi_2\}$

75% 缺失:  $P_4 = \{\varphi_4\}$ ,  $P_2 = \{\varphi_2\}$

选取不同特征工程下的 XGBoost 模型做对比实验, 对比实验的测试集准确率结果如表 1 所示.

表 1 不同模型对比实验测试集准确率

Tab. 1 Comparison of experimental test set accuracy for different models

| 模型              | 噪声    |       |       |
|-----------------|-------|-------|-------|
|                 | 5%    | 20%   | 40%   |
| xgb             | 73.72 | 72.88 | 72.17 |
| xgb + arr       | 85.86 | 81.15 | 75.31 |
| xgb + $P$       | 91.25 | 90.66 | 88.54 |
| xgb + $P_{234}$ | 90.56 | 89.46 | 86.83 |
| xgb + $P_{134}$ | 89.62 | 87.46 | 84.04 |
| xgb + $P_{34}$  | 88.97 | 85.94 | 82.34 |
| xgb + $P_{14}$  | 88.01 | 83.58 | 78.50 |
| xgb + $P_{34}$  | 88.58 | 85.55 | 81.48 |
| xgb + $P_4$     | 86.90 | 81.98 | 76.86 |
| xgb + $P_2$     | 88.08 | 83.32 | 79.09 |

其中, xgb 表示输入特征仅有位置信息的 XGBoost 多分类模型, 作为基准; xgb + arr 表示没有结合自动机的 XGBoost 模型, 即特征包含位置信息和  $f_{\text{arrived}}$ ; xgb +  $P$  表示结合完备规划库  $P$  对应自动机的 XGBoost 模型, 即本章提出的多分类模型; xgb +  $P_i$  表示结合不完备规划库  $P_i$  的 XGBoost 模型, 用于测试模型在不完备规划库下的表现.

分析可以得出以下结论:

① 对于目标预测问题, 结合自动机的 XGBoost 多分类模型具有更高的准确率;

② 对于数据噪声, 结合自动机的 XGBoost 多分类模型具有更强的鲁棒性;

③ 结合自动机的 XGBoost 多分类模型的准确率与规划库的完备程度正相关;

④ 针对多智能体的多分类模型能够考虑多智能体之间的协同任务.

### 3.2 围捕逼停仿真

为验证所设计算法在围捕逼停方面的有效性, 针对该算法进行仿真实验验证.

本节所采用的仿真环境与上一节一致, 但是由于已知敌方的入侵目标, 所以对仿真环境做了简化, 更侧重于展示多智能体围捕截停的过程. 首先是对敌方目标的拦截逼停, 当对方有一个智能体朝我方保护区行进时, 我方派遣两个无人平台对其进行拦截逼停, 示意如图 4 所示.

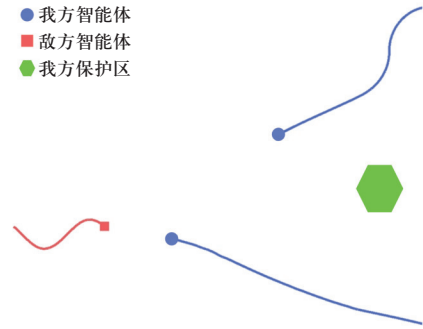


图 4 多智能体拦截逼停仿真结果

Fig. 4 Simulation results of multi-agent intercepting and forced stopping

其中红色矩形智能体为敌方入侵者, 绿色六边形为我方保护区, 两个蓝色圆形智能体为我方平台, 图中显示出各方行驶的轨迹, 其中入侵者以变速曲线运动向我方保护区靠近, 我方智能体根据算法实施最优决策, 在不断靠近敌方的同时尽可能地靠近入侵者与保护区的连线, 以形成对入侵者的包夹之势.

当靠近敌方智能体一定范围后, 敌方智能体意识到自身的危机, 往往会采取掉头逃逸的举措, 此时, 我方应当加派智能体, 对其进行围追堵截, 仿真结果如图 5 所示. 图中蓝色圆形表示我方智能体, 红色矩形表示敌方智能体, 通过颜色由浅变深表示其随时间变化的运动轨迹. 这里设置敌方智能体每时刻改变运动的速度大小和方向, 以躲避我方的围捕.

从图中可以看出, 我方选择加派一个智能体参与围捕任务, 不断接近目标并对其形成合围. 同时由于考虑了智能体本身的动力学约束, 无法做出大角度机动, 所以部分轨线呈现出并没有直接朝向入侵智能体的情况.

图 5(a) 展示了刚开始进行围捕的情况, 此前通过我方智能体的逼停策略, 使得敌方智能体不敢再

接近我方保护区,转而驶离.从图中可以看出我方两个智能体在初始对敌方智能体展示出包夹之势,而我方派出的第三个智能体则距离敌方较远,并且由于团队间的避障要求,一开始其并没有直接靠近,而是根据博弈算法逐渐与我方另两个智能体对敌方形成合围.说明该算法可以较好实现智能体内部的避碰.

图5(b)展示了围捕中期的情况,图中可以看出在接近坐标(600,280)的位置,我方智能体与敌方智能体遭遇,而我方智能体因为考虑了和敌方的避障

收益,所以提前降速,避免了碰撞,说明该博弈算法可以较好实现与敌方智能体的避碰.随着进程逐渐后移,可以看出我方三个智能体已经对敌方智能体实现了较好的包围,均匀地分布在敌方智能体的周围,说明该算法可以实现对敌方智能体的围捕策略.

图5(c)展示了围捕后期的情况,从图中可以看出我方智能体由于运动学约束无法进行大角度机动,而敌方虽然速度不如我方却可以进行相对大角度的转向.但是从始至终都对敌方智能体形成了较好的包围,并最终驱赶其远离我方区域.

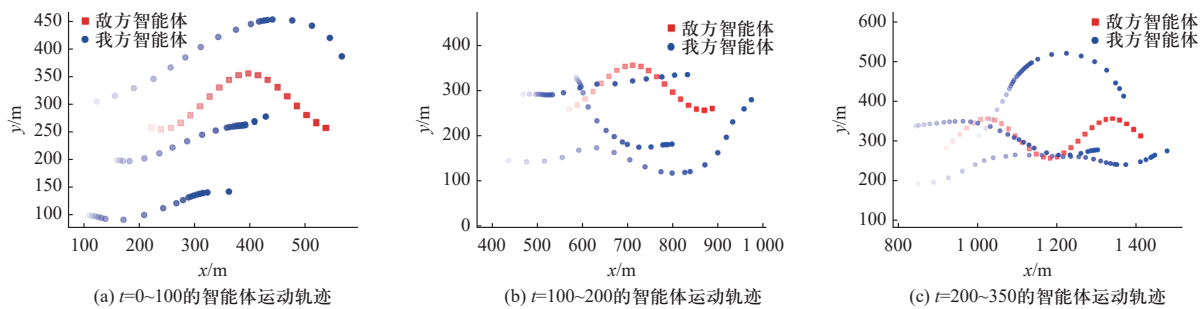


图5 多智能体围捕仿真结果

Fig. 5 Simulation results of multi-agent encirclement

通过上述仿真可以看出三个智能体的相对位置关系是有所变化的,这也说明了所提出的算法并不是根据一定规则制定逻辑来指定围捕智能体的位置,而是基于环境信息提取参与者的收益,并应用博弈理论所做出的智能决策.

为了进一步验证所提出算法的适应性与有效性,对敌方智能体的运动策略进行调整,设置三种反围捕策略,分别是:匀速直线运动,随机扰动的匀速运动,以及智能博弈决策运动,其结果如图6所示.

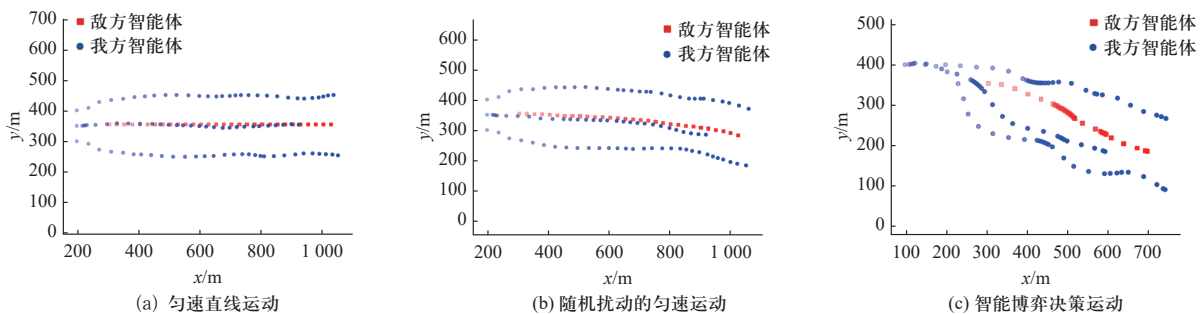


图6 应对不同敌方策略的多智能体围捕仿真结果

Fig. 6 Simulation results of multi-agent encirclement against different enemy strategies

图6(a)中敌方智能体以匀速直线运动进行逃离,从图中可以看出根据博弈算法结果,我方上下两个智能体立刻加速对敌方形形成合围,而另一个智能体从后方接近敌方,与另外两个智能体一起共同形成包围圈,快速实现了对敌方智能体的围捕,同时在后续过程中维持了这种合围队形且没有发生碰撞.图6(b)中敌方智能体仍然匀速运动,但被施加了一个转角

的随机扰动,可以看出根据博弈算法我方智能体仍然很好完成了围捕任务,并且整个过程中无碰撞发生.图6(c)中随机改变了我方智能体的初始位置,并且敌方智能体也采用博弈算法进行智能反围捕的逃离运动,我方智能体依旧成功实现了对敌方的围捕.以上测试的算法时间复杂度均为多项式级别,且在仿真测试中能以毫秒级的速度实现实时结果输出.

综上所述, 所提出的多智能体博弈围捕逼停算法, 在逼停拦截以及围捕敌方智能体等应用上都表现出了良好的性能, 在面对敌方的不同策略时, 算法也展现出了较好的适应性与有效性.

## 4 结 论

本文针对多智能体对抗过程中两大关键问题进行了研究. 针对规划识别中规则库不完备问题, 提出了基于数据驱动的多智能体识别算法: 首先采用结合自动机的特征提取方法, 获得样本的位置信息和任务信息; 然后将规划识别问题转换为多分类问题, 并以单智能体为切入点, 构建一种基于 XGBoost 的多分类模型; 最后, 对于多智能体之间可能存在的协同, 使用无监督学习方法 DBSCAN 对多智能体进行分簇, 对于同簇智能体, 构建针对多智能体的多分类模型, 实现对多智能体的目标预测. 在获悉敌方目标后, 为执行最优策略, 提出基于博弈的围捕逼停算法, 构建非合作动态博弈模型, 通过求解纳什均衡得到应对的最优策略. 在仿真阶段, 本文定义了一个对抗任务场景, 对全过程关键算法进行了仿真验证, 验证了算法的有效性.

### 参考文献:

- [1] WANG X M, XU Y H, CHEN C H, et al. Machine learning empowered spectrum sharing in intelligent unmanned swarm communication systems: challenges, requirements and solutions[J]. *IEEE Access*, 2020, 8: 89839 – 89849.
- [2] KIM H, MOKDAD L, BEN-OTMAN J. Designing UAV surveillance frameworks for smart city and extensive ocean with differential perspectives[J]. *IEEE Communications Magazine*, 2018, 56(4): 98 – 104.
- [3] GU J J, TAO S, WANG Q H, et al. Multiple moving targets surveillance based on a cooperative network for multi-UAV[J]. *IEEE Communications Magazine*, 2018, 56(4): 82 – 89.
- [4] WANG J J, JIANG C X, HAN Z, et al. Taking drones to the next level: cooperative distributed unmanned-aerial-vehicular networks for small and mini drones[J]. *IEEE Vehicular Technology Magazine*, 2017, 12(3): 73 – 82.
- [5] KAUTZ H A, ALLEN J F. Generalized plan recognition[C]// *Proceedings of National Conference on Artificial Intelligence*. Philadelphia: AAAI, 1986.
- [6] VILAIN M B. Getting serious about parsing plans: a grammatical analysis of plan recognition[C]// *Proceedings of the 8th National Conference on Artificial Intelligence*. Boston, Massachusetts: AAAI, 1990.
- [7] PYNADATH D V, WELLMAN M P. Probabilistic state-dependent grammars for plan recognition[C]// *Proceedings of the Sixteenth Conference on Uncertainty in Artificial Intelligence*. [S.l.]: ACM, 2013.
- [8] KANTHARAJU P. Learning decomposition models for hierarchical planning and plan recognition[D]. Philadelphia: Drexel University, 2021.
- [9] MENEGUZZI F R, PEREIRA R F. A survey on goal recognition as planning[C]// *Proceedings of the 30th International Joint Conference on Artificial Intelligence (IJCAI)*. Canada: [s.n.], 2021.
- [10] ORIEL U, REUTH D, YA'AKOV G. Plan recognition for exploratory domains using interleaved temporal search[J]. *AAAI Workshop on Plan, Activity, and Intent Recognition*, 2013: 61 – 66.
- [11] 贾镇泽, 樊晓光, 薛明浩, 等. 基于机动动作元的敌机战术机动在线识别方法[J]. *北京理工大学学报*, 2018, 38(8): 820 – 827.  
JIA Zhenze, FAN Xiaoguang, XUE Minghao, et al. Online identification method for tactical maneuver of target based on air combat maneuver element[J]. *Transactions of Beijing Institute of Technology*, 2018, 38(8): 820 – 827. (in Chinese)
- [12] DONG W, YIN X and LI S. A uniform framework for diagnosis of discrete-event systems with unreliable sensors using linear temporal logic[J]. *IEEE Transactions on Automatic Control*, 2024, 69(1): 145 – 160.
- [13] YU X, DONG W, Li S, et al. Model predictive monitoring of dynamical systems for signal temporal logic specifications[J]. *Automatica*, 2024, 160: 111445.
- [14] 丁明, 张书玲, 张琛. 业务流程的形式化设计与验证[J]. *北京理工大学学报*, 2016, 36(11): 1147 – 1153.  
DING Ming, ZHANG Shuling, ZHANG Chen. Formal design and verification of business processes[J]. *Transactions of Beijing Institute of Technology*, 2016, 36(11): 1147 – 1153. (in Chinese)
- [15] 聂兆伟, 陈志伟, 马晓东, 等. 基于多视图建模的武器装备系统安全性分析[J]. *北京理工大学学报*, 2022, 42(4): 437 – 446.  
NIE Zhaowei, CHEN Zhiwei, MA Xiaodong, et al. Safety analysis of weapon system based on multi-view modeling[J]. *Transactions of Beijing Institute of Technology*, 2022, 42(4): 437 – 446. (in Chinese)
- [16] 张鹏程. 基于博弈的空中目标航迹预测及攻防对抗研究[D]. 杭州: 浙江大学, 2023.  
ZHANG Pengcheng. Research on game-based aerial target track prediction and confrontation[D]. Hangzhou: Zhejiang University, 2003. (in Chinese)
- [17] 房肖, 温广辉, 付俊杰, 等. 基于博弈的水面无人艇集群对抗问题研究[J]. *控制工程*, 2022, 29(3): 492 – 497.  
FANG Xiao, WEN Guanghui, FU Junjie, et al. Study on

- group confrontation of unmanned surface vessels based on game theory[J]. *Control Engineering of China*, 2022, 29(3): 492 – 497. (in Chinese)
- [ 18 ] 赖中安, 余波, 周涛. 矩阵博弈在水面舰艇电子对抗中的应用 [J]. *舰船电子对抗*, 2010, 33(4): 21 – 23.  
LAI Zhongan, YU bo, ZHOU Tao. Application of matrix game to the electronic counter measure in surface warships[J]. *Shipboard Electronic Counter Measure*, 2010, 33(4): 21 – 23. (in Chinese)
- [ 19 ] 范衡, 孙福赞, 马培立, 等. 基于共识主动性的群体机器人目标搜索与围捕 [J]. *北京理工大学学报*, 2022, 42(2): 158 – 167.  
FAN Zhun, SUN Fuzan, MA Peili, et al. Stigmergy-based swarm robots for target search and trapping[J]. *Transactions of Beijing Institute of Technology*, 2022, 42(2): 158 – 167. (in Chinese)
- [ 20 ] TIAN D, FANG H, YANG Q, et al. Decentralized motion planning for multiagent collaboration under coupled LTL task specifications[J]. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2021, 52(6): 3602 – 3611.
- [ 21 ] TIAN D, FANG H, YANG Q, et al. Two-phase motion planning under signal temporal logic specifications in partially unknown environments[J]. *IEEE Transactions on Industrial Electronics*, 2023, 70(7): 7113 – 7121.
- [ 22 ] 贺嘉, 肖英杰. 基于贪心蚁群算法的无人水面艇全局路径规划 [J]. *上海海事大学学报*, 2021, 42(1): 7 – 12.  
HE Jia, XIAO Yingjie. Global path planning for unmanned surface vessels based on greedy ant colony algorithm[J]. *Journal of Shanghai Maritime University*, 2021, 42(1): 7 – 12. (in Chinese)

(责任编辑: 李兵)