

4-19-2022

An Unmanned Swarm Search Method Based on Human-Robot Cooperation

Xin Zhou

National University of Defense Technology, Changsha 410073, China; zhouxin09@nudt.edu.cn

Weiping Wang

National University of Defense Technology, Changsha 410073, China;

Yifan Zhu

National University of Defense Technology, Changsha 410073, China;

Tao Wang

National University of Defense Technology, Changsha 410073, China;

See next page for additional authors

Follow this and additional works at: <https://dc-china-simulation.researchcommons.org/journal>



Part of the [Artificial Intelligence and Robotics Commons](#), [Computer Engineering Commons](#), [Numerical Analysis and Scientific Computing Commons](#), [Operations Research, Systems Engineering and Industrial Engineering Commons](#), and the [Systems Science Commons](#)

This Paper is brought to you for free and open access by Journal of System Simulation. It has been accepted for inclusion in Journal of System Simulation by an authorized editor of Journal of System Simulation.

An Unmanned Swarm Search Method Based on Human-Robot Cooperation

Abstract

Abstract: Human-robot collaboration is a research hotspot and the human and unmanned swarm collaborative search is a typical scenario. It can carry out the more complex tasks by combining the human complex reasoning capabilities with repeated and precise execution capabilities of unmanned swarm. Based on the high-value target search of the uncertain scenarios, *the concept definition for the collaborative search of human and unmanned swarm is given. A multi-agent dynamic programming model under uncertain with unknown prior knowledge is proposed*, established to describe how the multi-agent system carries out the search under human support. *A dynamic programming algorithm based on sequential allocation is proposed and the simulation experiments are carried out.* The experimental results show that the performance of the algorithm is significantly better than that of the basic algorithm.

Keywords

unmanned swarm, human-robot collaboration, dynamic planning, multi-agent system

Authors

Xin Zhou, Weiping Wang, Yifan Zhu, Tao Wang, and Tian Jing

Recommended Citation

Xin Zhou, Weiping Wang, Yifan Zhu, Tao Wang, Tian Jing. An Unmanned Swarm Search Method Based on Human-Robot Cooperation[J]. Journal of System Simulation, 2022, 34(4): 735-744.

基于人机协作的无人集群搜索方法研究

周鑫, 王维平, 朱一凡, 王涛, 井田

(国防科技大学, 湖南 长沙 410073)

摘要: 人与机器人交互是当前一项研究热点, 人与无人集群协作搜索是人机交互的典型场景, 通过结合人的复杂推理能力和无人集群的重复精确执行能力以完成更为复杂的任务。以不确定场景下高价值目标搜索为背景, 给出了人与无人集群协作搜索问题概念定义; 为描述多 Agent 系统如何在人工辅助下进行选择目标区域进行探索, 构建了一种先验态势不确定情况下多 Agent 动态规划问题模型; 提出了一种基于顺次分配的动态规划算法。并对该方法进行了仿真实验, 实验结果表明: 该算法性能明显高于基准算法。

关键词: 无人集群; 人机协作; 动态规划; 多 Agent 系统

中图分类号: TP391.9

文献标志码: A

文章编号: 1004-731X(2022)04-0735-10

DOI: 10.16182/j.issn1004731x.joss.20-0898

An Unmanned Swarm Search Method Based on Human-Robot Cooperation

Zhou Xin, Wang Weiping, Zhu Yifan, Wang Tao, Jing Tian

(National University of Defense Technology, Changsha 410073, China)

Abstract: Human-robot collaboration is a research hotspot and the human and unmanned swarm collaborative search is a typical scenario. It can carry out the more complex tasks by combining the human complex reasoning capabilities with repeated and precise execution capabilities of unmanned swarm. Based on the high-value target search of the uncertain scenarios, the concept definition for the collaborative search of human and unmanned swarm is given. A multi-agent dynamic programming model under uncertain with unknown prior knowledge is proposed, established to describe how the multi-agent system carries out the search under human support. A dynamic programming algorithm based on sequential allocation is proposed and the simulation experiments are carried out. The experimental results show that the performance of the algorithm is significantly better than that of the basic algorithm.

Keywords: unmanned swarm; human-robot collaboration; dynamic planning; multi-agent system

引言

无人技术的快速发展使得无人集群逐渐成为人类的合作伙伴。人机交互系统^[1]是由指挥员与机器人组成的复杂系统, 他们相互协作以实现一个共同的目标。本文聚焦人机交互的一个子类, 即指挥员与无人集群协作搜索问题。无人集群是由

多个无人平台组成的群体, 如无人机集群、无人车集群^[2-4]。在无人集群自动规划领域, 无人平台会根据其当前状态和附近环境, 与相邻无人平台相互交互以及自己的局部控制规则进行自动规划。人机交互系统在未知远程环境和高度非结构化的任务中具有卓越性能, 能够减少人员工作量, 完成单个无人平台无法完成的任务, 且具有更高的

收稿日期: 2020-11-14

修回日期: 2020-12-28

基金项目: 国家自然科学基金(72101263, 62003359)

第一作者: 周鑫(1990-), 男, 博士, 讲师, 研究方向为数字决策与仿真评估、系统仿真、不确定环境下多智能体规划。

E-mail: zhouxin09@nudt.edu.cn

故障鲁棒性和可扩展性。现如今,人机交互的性质也发生了变化,由单人与单无人平台协作逐渐发展为多人与单无人平台交互、多人与多无人平台交互、单人与多无人平台交互^[6]。一般根据自动化层级来描述机器人自主程度^[7],10级表示机器人具有完全自主能力,而1级表示人类需要操控所有行动。现有机器人技术远未达到10级,因此仍需要人为干预来完成复杂任务,例如进行更高级的推理和决策。

一些学者对人与无人集群协作问题进行过研究。Aamodh提出了一种协作框架,在框架中指挥员能够控制集群的形状和编成^[8]。Shishir研究表明适当的人与无人集群协作可以使无人集群受益,从而能够更好地达成目标^[9]。指挥员与无人集群协作方式很多,其中一种是指挥员向无人集群提供辅助消息。在很多情况下指挥员能够根据经验或其他消息途径为无人集群提供消息,消息可能降低信息的不确定性,从而降低无人集群决策的复杂度,进而提高系统整体性能。事实上,从指挥员获得所需信息的最直接方法是遇到情况直接询问。假设无人集群具备较高自主能力,能够自己完成任务,偶尔会得到指挥员的支持。指挥员作为消息提供者,主要提供信息状态和决策行动2种方式。作为信息状态提供者,常见的问题框架是HOP-POMDP(human observation provider partially observable markov decision process)^[10]。HOP-POMDP考虑了Agent在规划决策时能够咨询指挥员,通过咨询获得必要信息,但指挥员可能提供非完整信息,且不一定能实时响应。作为决策行动提供者,常见的框架为MU-POMDP(model uncertainty partially observable markov decision process)^[11]。MU-POMDP考虑了转移函数、观测概率和回报值不确定的规划过程,Agent向指挥员询问一系列“是/否”问题或者描述行动列表,然后由指挥员确定采用何种行动。上述问题模型与本文提出的问题框架有所不同,因而其求解算法无法有效地应用于提出的搜索问题。

人与无人集群协作决策可以抽象为一种方案空间探索问题。假设存在很多备选目标区域,无人集群选择其中若干区域进行搜索,方案可以认为是无人平台搜索一个目标区域的过程,所有方案的集合称为方案空间。方案空间越大,越有可能包含最有价值的目标。对于方案空间搜索问题,通常将其抽象为一种目标优化问题。根据目标的数量,可以将优化问题分为单目标优化和多目标优化。为了便于分析,多个目标可以整合为单个目标。现有研究中,每个方案的回报值是确定的,可以直接得到而无需经过复杂计算。目前普遍采用帕累托最优方法,如NSGA-II(non dominated sorting genetic algorithm II)和NSGA-III(non dominated sorting genetic algorithm III)^[12-13],以及这些算法的扩展算法。这些算法适用于输入输出关系确定的问题中,但难以有效解决本文的问题框架,即无人集群的行动可能产生大量的、甚至无穷多的不同数值的搜索问题。指挥员首先需要确定拟探索方案的先后顺序,然后在已搜索的方案中选择最有价值的若干方案。换句话说,问题的目标是寻找到最优搜索路径,以保证在有限时间和资源内寻找到最优方案。该问题类似于旅行商问题^[14-15],最优搜索路径的计算复杂度为 $O(n^2 2^n)$,其中 n 表示方案空间中的方案数;但又不同于旅行商问题,本文考虑的路径优选问题中每种方案的价值服从某种分布,准确的价值在方案被探索之前是未知的。考虑到搜索目标需要耗费一定代价,搜索路径根据需求长度可调。Blanco等^[16]提出了一种搜索算法,决策者通过花费一定代价以获取未知方案的信息,并提出了停止搜索的条件。Weitzman^[17]提出了一种基于评判指标的序贯搜索问题,不过该方法适用于选择一个最优方案的情景。Chen等^[18]提出了一种用于人工辅助机器人决策的最优搜索算法,同样该方法只能选取一种最优方案,无法解决本文提出的问题。

本文考虑的场景是假设具有高度自主性的无人集群在人工干预下对敌方目标区域进行探测,以求在大范围环境中寻找到若干最有价值的区域。由于先期情报不准,无法确定高价值目标具体位置。此外,派遣无人集群执行任务时需要一些代价,一方面需要耗费燃料,另一方面有可能被对手击毁。因此,需要无人集群通过搜索判断不同区域的价值,以便为下一步行动提供支撑。因指挥员能够获取额外的信息以弥补无人集群的不足,即指挥员可以减少无人集群搜索过程的不确定因素,所以在很多情况下,人工干预的无人集群规划是十分必要的^[5]。但是指挥员的精力和能力是有限的,难以实时操控每个无人平台,因而无法同时为所有无人平台分析和处理数据。

鉴于此,本文聚焦于如何以有限的代价和资源寻找到最有价值的目标。高价值目标可以指指挥所、火炮阵地、机场等指挥员感兴趣的事物。本文弱化了实际应用背景,以便提出的问题框架以及求解算法具有一般性。指挥员与无人集群协作搜索问题面临三方面挑战:①各个区域具有不确定性。在派遣无人集群前往目标区域探测之前需要收集先验知识,如卫星图像。先验知识一方面具有不确定性,另一方面能够作为启发式信息辅助决策;②派遣无人集群前往目标区域探测需要花费代价,如油耗、损坏等。因而指挥员需要决定是继续探测下一个地区,还是停止搜索并选择已探测的最有价值的区域;③进行探测的手段很多,例如,无人机实地抵近侦察,或者回传高空侦察数据以寻求人工分析,因此指挥员应权衡不同行动的效用和成本。

1 人机协作搜索问题框架

本节给出了人与无人集群协作搜索问题的概念定义,并将该问题抽象为一种动态规划问题。

1.1 概念定义

把地理环境划分为若干区域,无人集群需要

在这些区域中找到高价值目标,如果包含的高价值目标越多,则这片区域就越有价值。无人集群的目标是以尽可能少的代价找到价值尽可能高的若干区域。假设每个区域的价值是独立的,在被无人集群探测之前具体数值是未知的。

定义1(搜索价值): 搜索价值是指无人集群探索某个区域的必要性综合度量。

搜索价值由目标的重要程度决定,目标越重要,该目标所具备的价值就越大,探索这片区域的价值也就越高。考虑到先验态势信息的不准确性,每片区域的价值事先是不确定的。假设不同区域的价值是相互独立的。对于区域 $u \in U$,其价值 x_u 服从概率分布 $F_u(x_u)$,其中 U 表示所有备选区域,或称为方案空间。无人集群可以采用抵近侦察或者高空侦察的方式获取信息。其中,抵近侦察能够获取区域的确切价值,高空侦察数据需要成功咨询指挥员后才可能获得确切价值。另外,不同行动耗费的成本是不一致的,比如抵近侦察就存在被击落的风险,而高空侦察则不易被击落。

定义2(搜索代价): 搜索代价是指无人集群探索某个区域时的成本综合度量。

派遣无人集群前往目标地点搜索是需要耗费成本的,记为 $c \in \mathbb{R}^+$ 。假设采用不同行动搜索不同区域的代价是已知且相互独立的。那么,无人集群在考虑进一步探索时,需要权衡搜索价值和搜索代价的关系,即回报率。回报率可以认为是净效益,本文无人集群的目标是使回报率最大化。

定义3(回报率): 回报率是搜索价值与探索成本之差。

每个区域的状态分为状态未知、状态已知和状态分析3类,如图1所示。状态未知是指该区域还未被无人集群探索,确切价值尚不可知;状态已知是指该区域已经被无人集群探测过,已经查明了确切价值;状态分析是指该地区正在被无人集群探测,无人集群正向指挥员请求数据分析。可以采用多种手段进行探测,不失一般性,无人集

群主要有2种手段：①无人集群前往目标区域 u 进行抵近侦察，采用多个视角和多种传感器收集和处理数据，记为 a'_u ；②无人集群前往目标区域 u 进行高空侦察，并请求指挥员进行数据分析，记为 a''_u 。所有行动集合记为 $A=\{a'_u, a''_u|u \in U\}$ 。

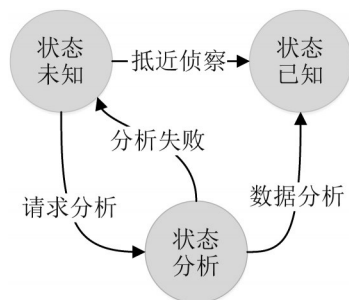


图1 方案状态转移关系
Fig. 1 Scheme state transition relationship

1.2 动态规划问题

人与无人集群协作搜索被形式化建模为一类多Agent系统动态规划问题，其中无人平台被抽象为Agent，每个区域被称为一个方案。将备选方案集合记为 $u \in U$ ，将其分为2个互斥集合：①不断增加的已探索方案集合 $I \in U$ ；②未探索方案集合 $\bar{I} \in U$ ， $I \cup \bar{I} = U$ ， $I \cap \bar{I} = \emptyset$ 。每个Agent在决策时，可以选择是否从集合 $\bar{I}$ 中选择一个未探索方案进行探索，此时可以采用抵近侦察和咨询指挥员两种方式获得确切回报值；或者该Agent停止探索并从集合 I 中选择一个已探索方案作为最终方案。设定整个集群一共选取 $|K|$ 个方案，每个Agent只能选择一个最有价值的方案，使探索价值最大化：

$$y^* = \max_{i \in I} x_i \quad (1)$$

求解问题的充分统计量记为 $(\bar{I}, y)$ ，称为系统状态。系统状态分为2部分：未探索方案信息 $\bar{I}$ 和已探索方案集合 $y = \{y_1, y_2, \dots, y_{|I|}\}$ 。将Agent系统多动态规划问题转化为系统状态评估函数。

定义4(系统状态评估函数)。系统状态评估函数为当已知价值集合为 y 且未探索方案集合为 $\bar{I}$

时，执行最优策略能够获得的期望回报值，记为 $\psi(\bar{I}, y)$ 。

由于每个方案探索价值和探索成本是独立的，因此集群的期望回报值可以分解为每个Agent的期望回报值。

$$\psi(\bar{I}, y) = \sum_{k \in K} \psi^k(\bar{I}^k, y^k) \quad (2)$$

式中： $\psi^k(\bar{I}^k, y^k)$ 为第 k 个Agent的子系统状态评估函数。对于任意2个Agent $w, v \in K$ 所选择的方案 $\hat{y}^w, \hat{y}^v$ ，有 $\hat{y}^w \neq \hat{y}^v$ ，表示每个Agent选择不同的方案。每个Agent都有局部方案空间，不同Agent的方案空间可能交叉 $U^w \cap U^v \neq \emptyset$ ，也可能不交叉 $U^w \cap U^v = \emptyset$ 。对于每个子集 $\bar{I}^k$ 和价值集合 y^k ，子系统状态评估函数 $\psi^k(\bar{I}^k, y^k)$ 需要满足式(3)所示的迭代关系。不失一般性，本文以Agent $k \in K$ 为例进行说明，其中 K 表示无人集群索引集合。

$$\psi^k(\bar{I}^k, y^k) = \max \{ \hat{y}^k, \psi^{*k}(\bar{I}^k, y^k), \psi^{**k}(\bar{I}^k, y^k) \} \quad (3)$$

子系统状态评估函数 $\psi^{*k}(\bar{I}^k, y^k)$ 为

$$\max_{i \in \bar{I}^k} \left\{ -c'_i + \int_{-\infty}^{\infty} \psi^k(\bar{I}^k - \{i\}, y^k \cup \{x_i\}) dF_i(x_i) \right\} \quad (4)$$

子系统状态评估函数 $\psi^{**k}(\bar{I}^k, y^k)$ 为

$$\max_{i \in \bar{I}^k} \left\{ -\frac{c''_i}{p} + \int_{-\infty}^{\infty} \psi^k(\bar{I}^k - \{i\}, y^k \cup \{x_i\}) dF_i(x_i) \right\} \quad (5)$$

式中： $\psi^{*k}(\bar{I}^k, y^k)$ 为Agent k 执行抵近侦察行动后的系统状态评估函数； $\psi^{**k}(\bar{I}^k, y^k)$ 为Agent k 执行询问指挥员行动(记为 a'')后的系统状态评估函数； p 为指挥员响应并成功分析出该方案价值(记为 a'')的概率； c'_i, c''_i 为对于方案 i ，Agent执行行动 a' 和行动 a'' 所耗费的成本。当系统状态为 $(\bar{I}^k, y^k)$ 时，Agent k 需要确定是停止搜索并从集合 y^k 选择最大价值 $\hat{y}^k$ ，或是在比较行动 a'_i 和行动 a''_i 效果后继续探索未知方案 i 。

通过上述定义，本文将人与无人集群协作搜索问题建模为一种动态规划问题。该问题特点是递归计算，因此难以直接求解状态评估函数。

2 基于顺次分配的动态规划算法

本节提出了一种基于顺次分配的动态规划 (sequential allocation based dynamic planning, SADP) 算法, 通过决策指标与采样值之间的大小关系来选择行动。

2.1 决策指标

对于每个方案 i , 执行行动 a'_i 和行动 a''_i 的决策指标分别记为 d'_i 和 d''_i 。决策指标是继续探索与停止探索的综合度量。

$$\begin{cases} d'_i = -c'_i + d'_i \int_{-\infty}^{d'_i} dF_i(x_i) + \int_{d'_i}^{\infty} x_i dF_i(x_i) \\ d''_i = -\frac{c''_i}{p} + d''_i \int_{-\infty}^{d''_i} dF_i(x_i) + \int_{d''_i}^{\infty} x_i dF_i(x_i) \end{cases} \quad (6)$$

参考文献[18], 可以解算式(6)得到:

$$\begin{cases} c'_i = \int_{d'_i}^{\infty} (x_i - d'_i) dF_i(x_i) \\ \frac{c''_i}{p} = \int_{d''_i}^{\infty} (x_i - d''_i) dF_i(x_i) \end{cases} \quad (7)$$

根据系统状态 $\psi(\bar{I}, y)$ 和指标集合 $\{c'_i, c''_i/p | i \in \bar{I}\}$, 可以设计一种简单但最优的搜索规则, 分为判断规则和选择规则。判断规则是指如果 Agent 要进一步探索未知方案, 则选择决策指标最大的未知方案, 同时选择对应的行动; 停止规则是指如果当前最大价值大于所有未知方案的决策指标时, 则停止搜索并选择具有最大价值的方案作为最终方案。

通过该规则, 复杂的系统状态函数计算问题就转化为简单的指标判断问题, 且每个指标的计算都是独立的, 不受其他方案价值概率分布的影响。

2.2 规划算法

在探索程序中, 首先根据式(6)和(7)计算所有方案所有行动的决策指标, 并保存至集合 π 中。其次嵌套遍历所有 Agent $k \in K$ 和每个 Agent 的未知方案空间 U^k 。在嵌套遍历中, 应先找出集合 π 中最大指标 d^* , 以及对应的方案 i^* 、行动 a^* 和代价 c^* 。为了便于描述, 指标集合是以全集为基础计数, 而不以每个 Agent 的局部集合计数, 即 $\pi =$

$\{d'_i, d''_i | i \in \bar{I}\}$ 。如果最大采样价值大于最大决策指标 $\hat{y}^k \geq d^*$, 则停止搜索并返回最大采用价值 $\hat{y}^k$ 、最终方案 i^k 和累积搜索代价; 反之则进入仿真程序 G, 获取是否咨询成功 $s \in \{\text{True}, \text{False}\}$ 以及采样价值 x 。如果获得确切采样价值, 那么更新集合 I^k 、 $\bar{I}^k$ 、 π^k 和累积搜索代价 c 。

算法 1: 探索程序

```

Procure: Exploring( $U$ )
 $\bar{I} \leftarrow U$ 
 $I, y, \pi \leftarrow \emptyset$ 
 $\hat{y}^k, \hat{i}^k, c \leftarrow 0, k \in K$ 
for  $i = 1$  to  $|\bar{I}|$  do
     $d'_i \leftarrow \text{Solving} \left( c'_i = \int_{d'_i}^{\infty} (x_i - d'_i) dF_i(x_i) \right)$ 
     $d''_i \leftarrow \text{Solving} \left( \frac{c''_i}{p} = \int_{d''_i}^{\infty} (x_i - d''_i) dF_i(x_i) \right)$ 
end for
 $\pi \leftarrow \{d'_i, d''_i | i \in \bar{I}\}$ 
for each  $k \in K$  do
    for  $j = 1$  to  $|U^k|$  do
         $d^* \leftarrow \max(\pi^k)$ 
         $i^* \leftarrow \underset{i \in I^k}{\text{argmax}}(\pi^k), a^* \leftarrow \underset{a \in A}{\text{argmax}}(\pi^k)$ 
         $c^* \leftarrow \text{Parsing}(i^*, a^*)$ 
        if  $\hat{y}^k \geq d^*$  then
            return  $(\hat{y}^k, \hat{i}^k, c)$ 
        else
             $(s, x) \leftarrow G(i^*, a^*)$ 
            end if
            if  $s = \text{true}$  then
                 $I^k \leftarrow I^k \cup \{i\}, k \in K$ 
                 $\bar{I}^k \leftarrow \bar{I}^k \setminus \{i\}, k \in K$ 
                 $\pi^k \leftarrow \pi^k \setminus \{d'_i, d''_i\}, k \in K$ 
                 $c \leftarrow c + c^*$ 
                if  $\hat{y}^k < x$  then
                     $\hat{y}^k \leftarrow x$ 
                     $\hat{i}^k \leftarrow i$ 
                end if
            else
                 $j \leftarrow j - 1$ 
            end if
        end for
    end for
end Procedure

```

在仿真程序 G 中, 如果采取行动 a'' , 则将生成的随机数 r 与咨询成功率 p 对比, 以判断能否咨询成功。另外, 符号 “ $\sim$ ” 表示采样, $x \sim F_i(x_i)$ 表示对概率分布 $F_i(x_i)$ 进行一次采样。

算法2: 仿真程序

```

Procedure G( $i, a$ )
 $s \leftarrow \text{false}$ 
 $x \leftarrow 0$ 
if  $a = a''$  then
 $r \leftarrow \text{Rand}(0, 1)$ 
if  $r \leq p$  then
 $s \leftarrow \text{true}$ 
 $x \sim F_i(x_i)$ 
end if
else
 $s \leftarrow \text{true}$ 
 $x \sim F_i(x_i)$ 
end if
return( $s, x$ )
end Procedure

```

3 仿真实验

3.1 实验设置

假设需要对目标区域进行探测, 建立战场高价值目标态势图, 为后续行动提供最新支撑。为了完成一系列使命, 拟派遣高智能无人集群前往目标区域, 并在人工辅助下以尽可能低的探索成本寻找若干高价值目标区域。那么无人集群如何自动规划自身行动是本实验需要解决的问题。

为了评估基于顺次分配的动态规划(sequential allocation based dynamic planning, SADP)算法的性能, 定义统计指标: ①平均回报值, 指多次仿真中 $|K|$ 个最高价值减去累积搜索代价的平均值; ②平均运行时间, 指运行多次仿真的平均时间; ③平均探索次数, 指每轮无人集群探索未知方案的平均次数; ④平均咨询次数, 指每轮中无人集群向人类咨询的平均次数。

此外, 为了对比 SADP 算法的性能, 设计了在问题框架下 2 种基准算法:

(1) 快速探索算法(fast exploration algorithm, FEA), 其评判指标是最高期望价值与代价之差, 即每轮决策选择方案:

$$i^* = \operatorname{arccmax}_{i \in I} \left\{ E(x_i) - c'_i, E(x_i) - c''_i/p \mid i \in \bar{I} \right\} \quad (8)$$

当已探索方案中的最高回报值大于该指标集合中的所有指标, 则停止探索并选择具有价值最高的方案。

(2) 试触探索算法(touch exploration algorithm, TEA), 其核心思想是比较当前最大回报值与下一

时刻期望最大回报值相比, 如果小于 0, 则继续探索, 反之则停止探索。每轮选择方案:

$$i^*_{t+1} = \operatorname{arccmax}_{i \in I} \left\{ E(x_i) - \left(\sum_{t=1}^T c(t) + c'_i \right), E(x_i) - \left(\sum_{t=1}^T c(t) + c''_i/p \right) \mid i \in \bar{I} \right\} \quad (9)$$

式中: $c(t)$ 为第 t 次决策时花费的代价值。

3.2 方案空间扩展性实验

本节设计了 3 个典型场景用以评估算法对于方案空间可扩展性的适应程度。一般而言, 咨询代价小于抵近侦察代价 $c'_i < c''_i, i \in I$, 但是咨询次数未知。设定所有搜索价值函数服从独立的均匀分布, 即 $F_i(x_i) \sim U(a_i, b_i)$, $a_i \sim U(1\,000, 1\,500)$, $b_i \sim U(1\,500, 2\,000)$, $i \in I$; 设定搜索代价也服从独立的均匀分布, 即 $c'_i \sim U(10, 20)$, $c''_i \sim U(3, 4)$, $i \in I$; 咨询成功概率为 $p = \{0, 0.1, 0.2, \dots, 1\}$ 时, 无人集群选取 $|K| = 9$ 个方案, 每个场景独立重复运行 1 000 次仿真。评估 3 个场景下算法的性能指标:

场景 A1: 方案空间大小为 $|I| = 50$ 。

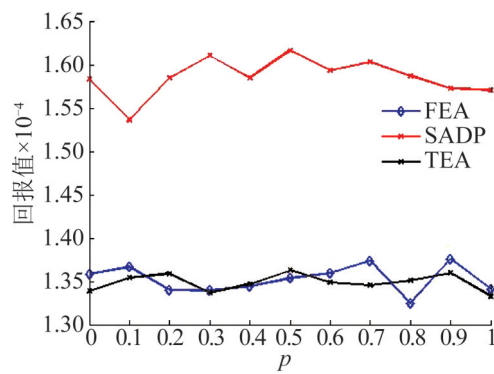
场景 A2: 方案空间大小为 $|I| = 500$ 。

场景 A3: 方案空间大小为 $|I| = 5\,000$ 。

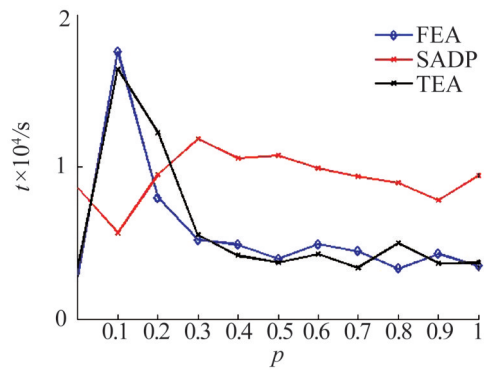
场景 A1 仿真结果如图 2 所示。实验结果表明在这些场景中 SADP 回报值远高于 FEA 和 TEA 的回报值, SADP 运行时间略高于 FEA 和 TEA, FEA 和 TEA 探索次数非常接近, 随着 p 的增加, SADP 探索次数波动增长。当 $p \geq 0.3$ 时, SADP 咨询次数大于 FEA 和 TEA 咨询次数。

场景 A2 仿真结果如图 3 所示。实验结果表明在这些场景中 SADP 回报值远高于 TEA, 而 TEA 回报值高于 FEA。SADP 运行时高于 FEA 和 TEA。SADP 探索次数明显大于 FEA 和 TEA 的探索次数。当 $p > 0.2$ 时, SADP 咨询次数大于 FEA 和 TEA 咨询次数。

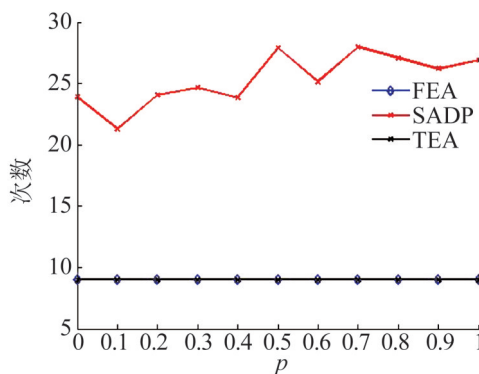
场景 A3 仿真结果如图 4 所示。实验结果表明在大部分场景下, SADP 的回报值、运行时间、探索次数和咨询次数大多高于 TEA, 而 TEA 的回报值、运行时间、探索次数和咨询次数又大多高于 FEA。



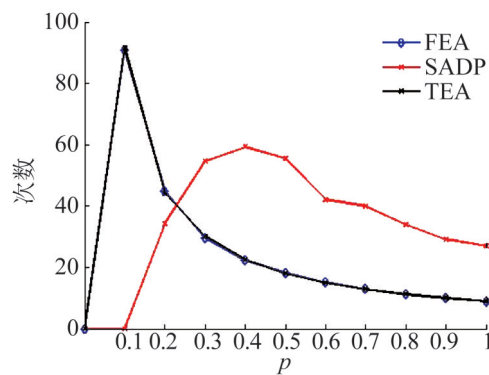
(a) 回报值



(b) 运行时间



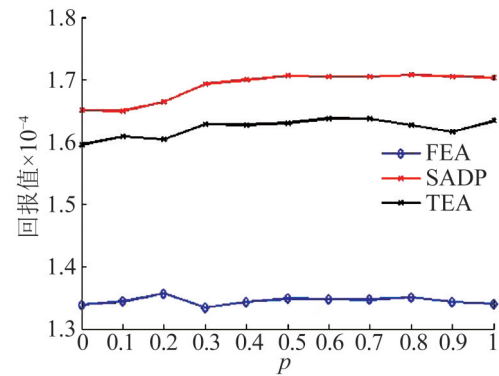
(c) 探索次数



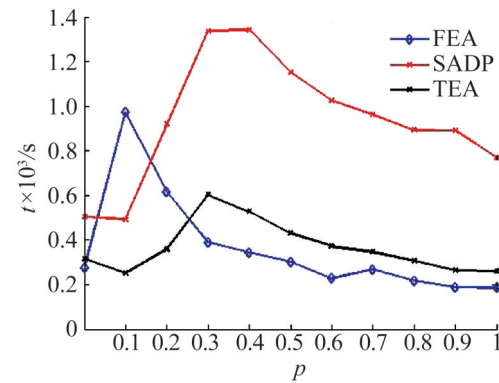
(d) 咨询次数

图2 场景A1仿真结果

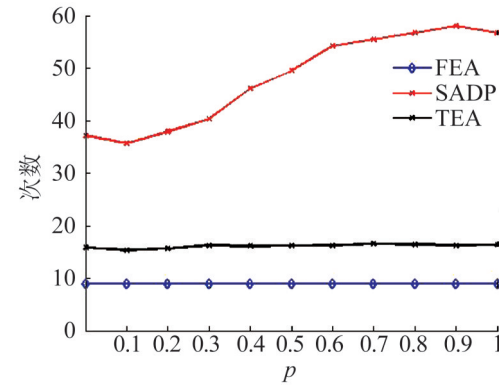
Fig. 2 Simulation results in scenario A1



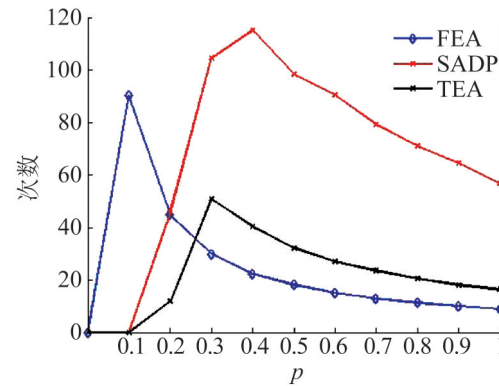
(a) 回报值



(b) 运行时间



(c) 探索次数



(d) 咨询次数

图3 场景A2仿真结果

Fig. 3 Simulation results in scenario A2

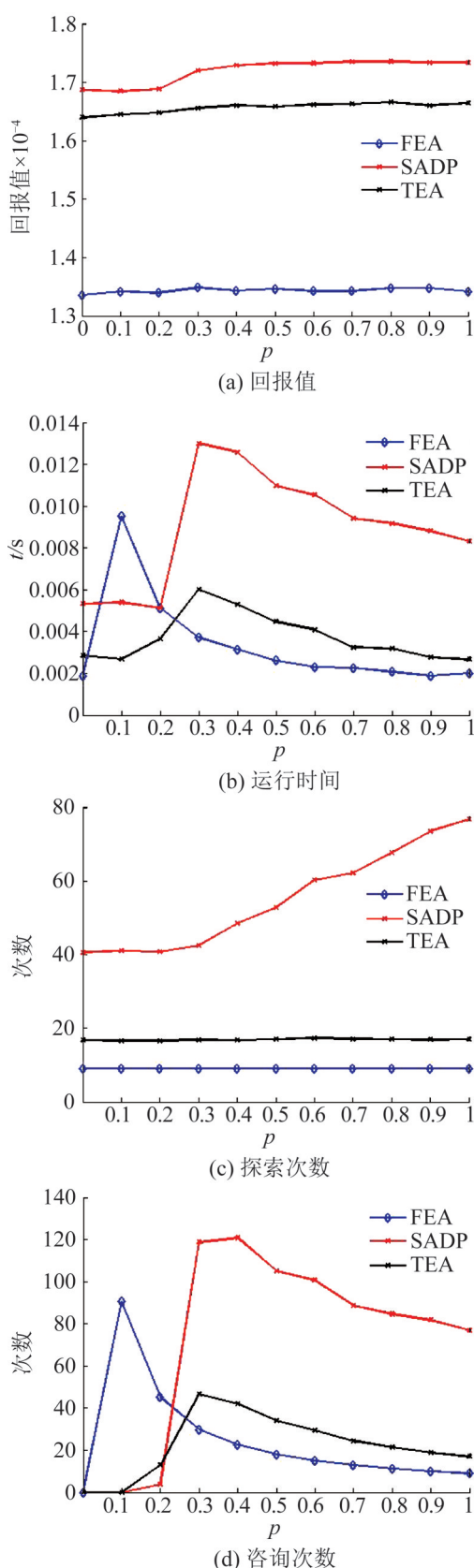


图4 场景A3仿真结果

Fig. 4 Simulation results in scenario A3

3.3 参数敏感性实验

设定所有搜索价值函数服从独立的均匀分布，即 $F_i(x_i) \sim U(a_i, b_i)$, $a_i \sim U(1\,000, 1\,500)$, $b_i \sim U(1\,500, 2\,000)$, $i \in I$; 设定搜索代价也服从独立的均匀分布，即 $c'_i \sim U(10, 20)$, $c''_i \sim U(3, 4)$, $i \in I$ 。此外，咨询成功概率设为 $p=0.3$ ，方案空间大小设为 5 000，无人集群选取 $|K| = \{1, 3, \dots, 19\}$ 个方案时的回报值、运行时间、搜索次数、咨询次数 4 项性能指标。每个场景独立重复运行 1 000 次仿真。

场景B仿真结果如图5所示。实验结果表明随着 $|K|$ 的增加，所有算法的回报值、运行时间、搜索次数、咨询次数都逐渐增加。此外在这些场景中，SADP 回报值、运行时间、搜索次数、咨询次数高于 TEA，而 TEA 又高于 FEA。

3.4 实验分析

本节设置若干场景，但实验结果具有一些共性。比如在场景A中，随着变量 p 的增加，回报值略有增加，而运行时间会逐渐减少。这是因为 p 的增加会减少咨询次数并降低期望咨询代价，从而提高期望回报值。值得说明的是， $p=0.3$ 是 SADP 算法的拐点，当 $p < 0.3$ 时，其回报值、运行时间、探索次数、咨询次数都处于低位，而当 $p=0.3$ 时有明显的提高。这与实验参数设置有很大关系，当 $p=0.3$ 时，行动 a' 的代价 c' 基本上大于行动 a'' 的代价 c/p ，因此 Agent 将采用咨询指挥员而不是采用抵近侦察行动。此时即便搜索路径不变，代价值也会降低，回报值增加，咨询次数增加。进一步，代价 c 的变化会影响决策指标 d 的变化，从而影响搜索顺序以及何时停止搜索。决策指标 d 具有期望回报值类似的意义， c 与 d 呈负相关，如果 c 越小，则 d 越大。对于场景B，随着变量 $|K|$ 增加各项指标随之增加。很显然毕竟选择的方案多了，而每个方案回报值为正，所以回报值肯定增加。同样，运行时间、搜索次数、咨询次数也会增加。

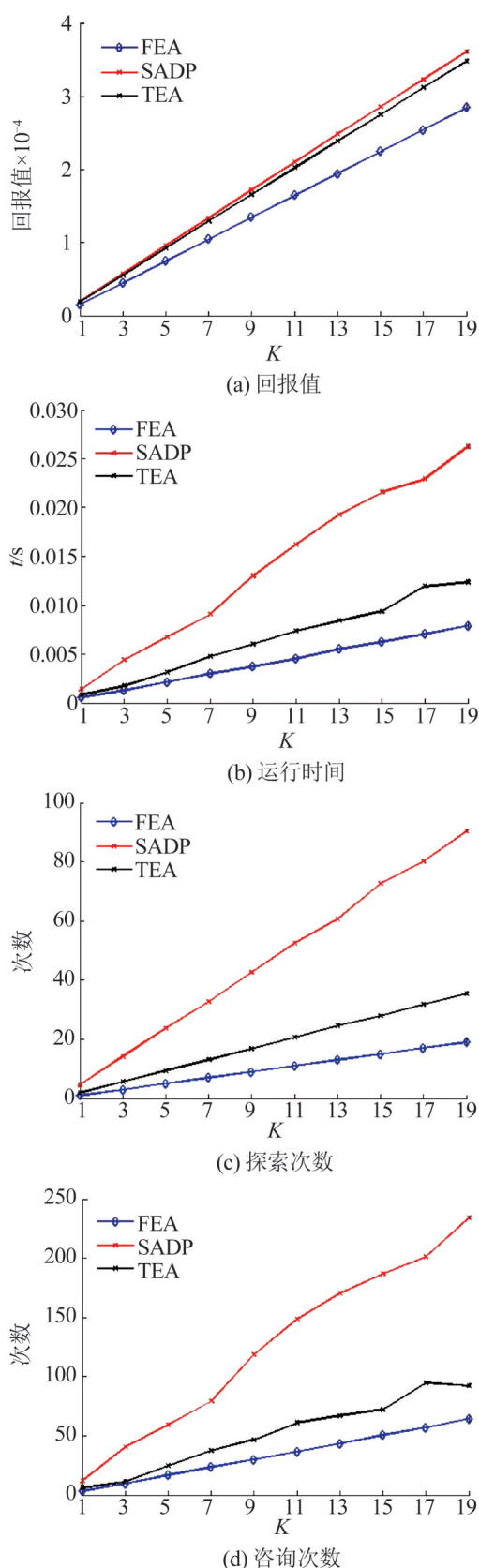


图5 场景B仿真结果

Fig. 5 Simulation results in scenario B

4 结论

人与无人集群协作搜索是具有广泛的应用价值和重要的研究意义。关于人机交互协作搜索,许多学者将其抽象为一种动态规划问题。其中,大多考虑单个最优方案如何生成,而对于多个最优方案并没有考虑。本文的工作是对之前工作的拓展,主要研究如何选择若干最优的方案。针对人与无人集群交互协作搜索问题,提出了一种多项式事件最优的方案空间探索算法。由于算法采用一种基于判定指标的贪婪搜索方法,使得该算法探索的回报值高于其他基准算法。然而该算法也存在一些不足,主要体现在运行时间比基准算法略长,这可能与判定指标计算的复杂度有关。下一步,可以考虑更为复杂的协作搜索问题,以及如何并行分布计算,以提高问题的适用性和计算效率。

参考文献:

- [1] Patel J, Xu Y, Pinciroli C. Mixed-Granularity Human-Swarm Interaction[C]// 2019 International Conference on Robotics and Automation. MA USA: Worcester Polytechnic Institute, 2019: 1059-1065.
- [2] 吕震华, 高亢. 美国无人集群城市作战应用发展综述[J]. 中国电子科学研究院学报, 2020, 15(8): 738-745.
Lü Zhenhua, Gao Kang. Review of the Development of Drone Swarm Urban Combat Applications in the USA [J]. Journal of CAEIT, 2020, 15(8): 738-745.
- [3] 王涛, 王维平, 李小波, 等. 一种多无人机集群持续侦察分层控制框架及关键算法[J]. 系统仿真学报, 2018, 30(4): 1221-1228.
Wang Tao, Wang Weiping, Li Xiaobo, et al. A Hierarchical Control Framework and Key Algorithms of Multi-Swarm Persistent Surveillance[J]. Journal of System Simulation, 2018, 30(4): 1221-1228.
- [4] 王维平, 王涛, 李小波, 等. 一种基于模体的无人机集群智能任务规划方法[J]. 系统仿真学报, 2018, 30(4): 1211-1220.
Wang Weiping, Wang Tao, Li Xiaobo, et al. Motif-based Intelligent Mission Planning Method for UAV Swarms[J]. Journal of System Simulation, 2018, 30(4): 1211-1220.
- [5] 陶伟. 基于马赛克战的水下有人-无人集群控制结构[J]. 指挥与控制学报, 2020, 6(3): 264-270.

- Tao Wei. Control Structure of Underwater Manned and Unmanned Swarm Based on Mosaic Warfare[J]. Journal of Command and Control, 2020, 6(3): 264-270.
- [6] Musić S, Hirche S. Control Sharing in Human-Robot Team Interaction[J]. Annual Reviews in Control (S1367-5788), 2017, 44: 342-354.
- [7] 王越超, 刘金国. 无人系统的自主性评价方法[J]. 科学通报, 2012, 57(15): 1290-1299.
- Wang Yuechao, Liu Jinguo. Evaluation Methods for the Autonomy of Unmanned Systems[J]. Chinese Science Bulletin, 2012, 57(15): 1290-1299.
- [8] Suresh A, Martínez S. Gesture Based Human-Swarm Interactions for Formation Control Using Interpreters[J]. IFAC-PapersOnLine (S2405-8963), 2019, 51(34): 83-88.
- [9] Bashyal S, Venayagamoorthy G K. Human Swarm Interaction for Radiation Source Search and Localization [C]// 2008 IEEE Swarm Intelligence Symposium. Piscataway, NJ, USA: IEEE Computer Society, 2008: 1-8.
- [10] Vanhée L, Jeanpierre L, Mouaddib A I. Augmenting Markov Decision Processes with Advising[C]// AAAI Conference on Artificial Intelligence. Menlo Park, CA, USA: AAAI, 2019: 2531-2538.
- [11] Doshi F, Pineau J, Roy N. Reinforcement Learning with Limited Reinforcement: Using Bayes Risk for Active Learning in POMDPs[J]. Artificial Intelligence (S0004-3702), 2012, 187: 115-132.
- [12] 曹桓, 周慧芳, 许勇, 等. 一种基于NSGA-II的网络规划资源配置优化处理方法[J]. 电信工程技术与标准化, 2020, 33(4): 30-37.
- Cao Heng, Zhou Huifang, Xu Yong, et al. An Optimization Method of Resource Allocation in Network Planning Based on NSGA-II[J]. Telecom Engineering Technics and Standardization, 2020, 33(4): 30-37.
- [13] 刘彬, 王卫涛, 武尤, 等. 改进自适应多种群NSGA-III算法的研究[J/OL]. 计算机应用研究[2020-11-11]. <https://doi-org-s.nudtproxy.yitlink.com/10.19734/j.issn.1001-3695.2019.10.0595>.
- Liu Bin, Wang Weitao, Wu You, et al. Research on Improved Adaptive Multi-Population Nsga-iii[J]. Application Research of Computers[2020-11-11]. <https://doi-org-s.nudtproxy.yitlink.com/10.19734/j.issn.1001-3695.2019.10.0595>.
- [14] 陈科胜, 鲜思东, 郭鹏. 求解旅行商问题的自适应升温模拟退火算法[J/OL]. 控制理论与应用[2020-11-11]. <https://kns-cnki-net-s.nudtproxy.yitlink.com:443/kcms/detail/44.1240.TP.20201015.1122.006.html>.
- Chen Kesheng, Xian Sidong, Guo Peng. Adaptive Temperature Rising Simulated Annealing Algorithm for Traveling Salesman Problem[J]. Control Theory & Applications [2020-11-11]. <https://kns-cnki-net-s.nudtproxy.yitlink.com:443/kcms/detail/44.1240.TP.20201015.1122.006.html>.
- [15] 张松灿, 普杰信, 司彦娜, 等. 基于种群相似度的自适应改进蚁群算法及应用[J/OL]. 计算机工程与应用[2020-11-11]. <https://kns-cnki-net-s.nudtproxy.yitlink.com:443/kcms/detail/11.2127.TP.20200820.1042.016.html>.
- Zhang Songcan, Pu Jiexin, Si Yanna, et al. An Adaptive Improved Ant Colony Algorithm Based on Population Similarity and Its Application[J]. Computer Engineering and Applications [2020-11-11]. <https://kns-cnki-net-s.nudtproxy.yitlink.com:443/kcms/detail/11.2127.TP.20200820.1042.016.html>.
- [16] Branco F, Sun M, Villas-Boas J M. Optimal Search for Product Information[J]. Management Science (S0025-1909), 2012, 58(11): 2037-2056.
- [17] Weitzman M L. Optimal Search for the Best Alternative [J]. Econometrica: Journal of the Econometric Society (S0012-9682), 1979, 47(3): 641-654.
- [18] Chen S, Baarslag T, Zhao D, et al. A Polynomial Time Optimal Algorithm for Robot-Human Search Under Uncertainty[C]// The 25th International Joint Conference on Artificial Intelligence. San Mateo, CA, USA: Morgan Kaufmann Publishers Inc., 2016: 819-825.