

## SWARM-LLM:基于大语言模型的无人集群任务规划系统

李婷婷, 王琪, 王嘉康, 徐勇军

引用本文

李婷婷, 王琪, 王嘉康, 徐勇军. SWARM-LLM:基于大语言模型的无人集群任务规划系统[J]. 计算机科学, 2025, 52(1): 72-79.

LI Tingting, WANG Qi, WANG Jiakang, XU Yongjun. SWARM-LLM:An Unmanned Swarm Task Planning System Based on Large Language Models [J]. Computer Science, 2025, 52(1): 72-79.

---

### 相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

#### [基于混合模仿学习的多智能体追捕决策方法](#)

Multi-agent Pursuit Decision-making Method Based on Hybrid Imitation Learning

计算机科学, 2025, 52(1): 323-330. <https://doi.org/10.11896/jsjcx.240800072>

#### [大语言模型驱动的多元关系知识图谱补全方法](#)

Large Language Model Driven Multi-relational Knowledge Graph Completion Method

计算机科学, 2025, 52(1): 94-101. <https://doi.org/10.11896/jsjcx.240600170>

#### [一种基于知识图谱的检索增强生成情报问答技术](#)

Retrieval-augmented Generative Intelligence Question Answering Technology Based on Knowledge Graph

计算机科学, 2025, 52(1): 87-93. <https://doi.org/10.11896/jsjcx.240900064>

#### [基于预训练大模型的行动方案生成方法](#)

COA Generation Based on Pre-trained Large Language Models

计算机科学, 2025, 52(1): 80-86. <https://doi.org/10.11896/jsjcx.240900075>

#### [提示学习中思维链生成和增强方法综述](#)

Survey of Chain-of-Thought Generation and Enhancement Methods in Prompt Learning

计算机科学, 2025, 52(1): 56-64. <https://doi.org/10.11896/jsjcx.240700172>

# SWARM-LLM:基于大语言模型的无人集群任务规划系统

李婷婷<sup>1</sup> 王琪<sup>1,2</sup> 王嘉康<sup>1,2</sup> 徐勇军<sup>1,2</sup>

1 中国科学院计算技术研究所 北京 100190

2 中国科学院大学 北京 100049

(litingting@ict.ac.cn)

**摘要** 针对无人集群系统自主智能性不足、异构无人集群协同效率低、任务分配不均衡等问题,文中面向无人集群自主规划、高效协作、智能决策的需求,首先提出了一种新的基于大语言模型的无人集群任务规划系统框架(SWARM-LLM)。该框架利用大语言模型将高层次的任务指令转化为具体的智能无人集群任务规划方案,通过任务分解、任务分配、任务执行等多个阶段来实现无人集群协同任务。进一步地,设计了一套适用于无人集群规划的提示工程方法-规划链(Planning Chain,PC),用来指导和优化上述各阶段的实施。最终,在无人集群仿真环境(AirSim)中构建了不同类别和复杂度的任务,并进行了评估实验。与其他基于优化算法和机器学习的算法相比,实验结果证明了SWARM-LLM框架的有效性,并在任务成功率上展现出了显著的优势,平均性能提升了47.8%。

**关键词**:任务规划;无人集群;大语言模型;协同策略;智能决策

**中图分类号** TP181

## SWARM-LLM: An Unmanned Swarm Task Planning System Based on Large Language Models

LI Tingting<sup>1</sup>, WANG Qi<sup>1,2</sup>, WANG Jiakang<sup>1,2</sup> and XU Yongjun<sup>1,2</sup>

1 Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China

2 University of Chinese Academy of Sciences, Beijing 100049, China

**Abstract** In response to the issues of insufficient autonomous intelligence in unmanned cluster systems, low collaborative efficiency of heterogeneous unmanned clusters, and unbalanced task allocation, this paper first proposes a new unmanned cluster task planning framework (SWARM-LLM) based on large language models to meet the needs of unmanned swarm systems for autonomous planning, efficient collaboration, and intelligent decision-making. This framework leverages large language models to transform high-level task instructions into specific intelligent unmanned cluster task planning solutions, achieving collaborative tasks of unmanned clusters through multiple stages such as task decomposition, task allocation, and task execution. Furthermore, this paper designs a prompt engineering method specifically suited for unmanned swarm planning, called the planning chain (PC), to guide and optimize the implementation of these stages. Finally, we construct tasks of various categories and complexities in an unmanned swarm simulation environment (AirSim) and conduct evaluation experiments. Compared with other algorithms based on optimization and machine learning, experimental results demonstrate the effectiveness of the SWARM-LLM framework, showing a significant advantage in task success rates, with an average performance improvement of 47.8%.

**Keywords** Task planning, Unmanned swarms, Large language models, Collaborative strategies, Intelligent decision-making

## 1 引言

随着人工智能技术的迅猛发展,无人集群系统在军事防御、城市管理、工业生产等多个关键领域的应用变得日益普遍和关键<sup>[1]</sup>。随着无人系统技术的不断发展与进步,单个无人系统已经难以通过自身完成复杂繁琐的工作任务。为了应对这一挑战,人们开始探索多无人系统协同工作的方式。多无人系统由众多自主无人单元组成,无人单元之间通过协作可以显著提升任务执行的效率、扩展性和适应性。在多无人系统中,各个无人单元在类型和功能上往往存在差异,这种异质性增加了系统的复杂性,并突显了在这些单元之间进行高效

任务分配的重要性。无人集群系统的核心挑战之一是如何在自主行动的无人系统中高效地进行任务分配和协调工作,这个问题已成为当前研究的焦点和难题<sup>[2]</sup>。

多无人系统协同任务规划带来的新挑战包括无人系统的自主智能性不足、异构无人系统的协同效率低、任务分配不均衡、各个任务之间的相互依赖性强等,而且任务规划难度随着无人机数量的增加而增加。传统的多无人机协同规划方法往往依赖固定算法,难以应对复杂环境和多样化任务。鉴于指令的不完整性和模糊性,利用广泛的先验知识来解释任务并帮助进行有效的任务规划至关重要。大语言模型<sup>[3]</sup> (Large Language Models, LLM) 具有优秀的自然语言理解、逻辑推理

和通用性,其凭借强大的能力和无限的潜力引领了新一轮技术革命,也为实现多无人系统的任务规划提供了一个潜在路径。自2020年GPT-3模型推出以来,LLM已成为自然语言处理<sup>[4]</sup>领域的一个新兴研究领域,当前基于LLM进行机器人规划决策的研究成果大多还是针对机械手臂<sup>[5]</sup>这些封闭世界简单任务中的尝试,对于军事作战问题,需要结合其开放、动态、安全等特性进行进一步研究。

基于以上背景,为提升多无人系统的任务规划能力,本文提出了一种基于大语言模型的无人集群任务规划框架(SWARM-Large Language Model, SWARM-LLM)。该方法利用大语言模型的强大处理能力,将复杂任务分解为与特定对象或动作相关的子任务,并根据各无人单元的技能合理分配这些子任务,从而实现任务的有效规划和执行。本文的主要贡献包括:

- 1)提出了一种基于大语言模型的无人集群任务规划框架,为解决多无人系统中的任务分配问题提供了新的视角。
- 2)设计了一套适用于无人集群规划的提示工程方法——规划链(Planning Chain, PC),引导LLM生成高质量的任务规划方案,能够有效地将任务分解并分配给最适合执行的无人单元。
- 3)通过实验验证了SWARM-LLM方法在多无人系统任务规划中的有效性,展示了其在处理多种任务时的潜力和优势。

## 2 相关工作

多无人系统任务规划<sup>[6]</sup>是无人系统领域中的一个重要研究方向,它关注的是如何在—组无人单元之间有效地分配和协调任务,以实现共同的目标。这一过程要求无人单元之间实现有效协作,并智能应对任务的复杂性及环境的动态变化<sup>[7]</sup>。多无人系统任务规划算法需要全面考虑无人机自身约束、环境约束、任务约束等多种因素,以实现在不完全战场信息条件下的无冲突完满解与最优解,从而最大化任务执行的收益。

现阶段无人系统自主规划的主要技术思路分为优化算法和机器学习。优化算法的思路为通过模拟大自然中的某种自然现象来实现无人系统的自主规划问题。机器学习通过模拟人类的学习行为,实现在样本下的自主学习,解决复杂环境下的无人系统自主规划问题。已经有许多的学者将优化算法应用到了无人集群规划问题中。文献[8]提出了一种基于蚁群的启发式算法,以最小化任务的时间消耗为导向,解决异构无人机群多区域的合作搜索规划问题;文献[9]设计了一种三阶段的启发式算法,通过地面车辆与多无人机协同来解决多区域覆盖侦查路径规划问题;文献[10]提供了一个多无人机作战场景的模型,该模型使用非合作博弈来模拟作战场景,并使用匈牙利匹配算法(Kuhn-Munkres)的变体来解决决策过程中的任务分配问题。作为当前研究的热点,机器学习在多无人系统自主规划领域的应用也相当广泛。文献[11]通过改进的深度强化学习方法实现了无人机在动态场景下的路径规划;文献[12]进一步提出了一种创新的基于深度强化学习的框架,该框架通过分布式控制策略,优化了无人机群的协同工作,显著提高了整体任务执行的效率和效果。此外,文献[13]

结合了深度强化学习和可解释人工智能(Explainable Artificial Intelligence, XAI)的技术,为小型无人机在未知环境中的自主路径规划提供了新的解决方案,这一研究不仅增强了无人机的自主决策能力,还通过可解释性提高了操作者对系统行为的信任度。智能优化和机器学习方法在多无人系统任务规划中虽有显著贡献,但它们仍面临计算效率低、适应性差、数据依赖性强和模型解释性差等挑战。

作为新兴的领域,大语言模型在无人系统的自主规划领域内也有许多的成果。文献[14]提出的“Drive like a human”框架,通过精心设计的提示工程,实现了利用大型语言模型进行汽车的自主导航。文献[15]提出的Driving with LLMs则是通过微调预训练模型,让大语言模型成为“专职司机”。Microsoft的研究<sup>[16]</sup>则提出了一种创新策略,将提示工程原则与高级功能库相结合,通过将ChatGPT与AirSim<sup>[17]</sup>模拟器相融合,展示了该方法在执行多样化无人任务方面的有效性。Tazir等<sup>[18]</sup>成功地将ChatGPT与PX4/Gazebo无人机模拟器相结合,创建了一个能够响应自然语言命令的系统,该系统能够执行起飞、着陆等无人机操作,显著降低了非专家用户操作无人机的难度。文献[19]提出了一个名为TypeFly的系统,该系统通过结合边缘计算的视觉效果与大型语言模型(LLMs)的提示工程,实现了对无人机的高效控制和任务规划。尽管这些研究成果展示了LLMs在无人系统自主规划方面的潜力,但目前的研究方向多集中在单无人机系统上,需要进一步探索如何将这些技术应用于多无人系统的任务规划中,以实现更高效、协同的无人任务执行。

## 3 问题描述

任务规划可表述为元组 $\langle R, O, P, U, A, S, T, E \rangle$ ,在无人系统 $R$ 框架下, $O$ 是环境中所有可用对象的集合, $P$ 是对象的一组属性, $A$ 为一组可执行的动作,每个无人系统都有可执行的动作 $a \in A$ ,根据定义为 $s \in S$ 的当前环境状态而变化。状态 $S$ 是所有对象属性的具体状态, $T$ 表示转移模型 $T: S \times A \rightarrow S$ 。

在给定指令 $U$ 后,LLM理解该指令,计算完成任务所需的步骤,并制定可以使多智能系统执行的任务计划。通过在可行的情况下并行执行任务,最大限度地利用无人系统的方式执行任务。

基于无人系统可执行动作 $A$ 和环境 $E$ ,指令 $U$ 可以分解为一组临时排序的 $K$ 个子任务 $T = \{T_1^i, T_2^i, \dots, T_K^i\}$ ,其中 $t_j$ 表示子任务的时间顺序,且 $j \leq K$ 。值得注意的是,某些子任务可以并行执行,具有相同的时间优先级。令 $T_j^s$ 为无人机完成子任务 $T_j^i$ 所需的动作集合,其中 $T_j^s \subseteq A$ 且 $T_j^s \subseteq T$ 。基于 $T_j^s$ ,可以将子任务分配给无人系统 $r$ ,如果 $T_j^s \subseteq A$ ,则具有完成部分任务的能力,其中 $r \in R$ 且 $A_r \subseteq A$ 。在没有单无人系统满足此约束的情况下,需要两个或多无人系统组成的团队来执行子任务。在这种情况下便组成了一个由 $Q$ 个无人系统组成的团队, $R = \{r_1, r_1, r_1, \dots, r_Q\}$ ,满足 $T_j^s \subseteq A$ 且 $T_j^s \subseteq T$ 。

## 4 本文方法

本文基于大语言模型的强大处理能力,将高层次的任务指令转化为具体的、可执行的智能无人集群任务规划方案。

整体框架如图1所示,SWARM-LLM 主要由4个主要阶段组成。1) 构建提示工程:创建能够有效引导大型语言模型处理和生成任务规划的提示。这包括定义任务的高层次指令,以及如何将这些指令转化为具体的操作步骤。2) 任务分解:结合任务需求、无人机能力和任务分解样本,形成指令输入,并

通过 LLM 模型,以实现任务的细化分解。3) 任务分配:基于任务分解的结果,制定任务指派策略并将其提示给 LLM, LLM 根据这些信息生成针对特定任务的详细指派方案。4) 任务执行:无人机根据 LLM 生成的指派计划执行具体任务。下面详细介绍4个阶段。

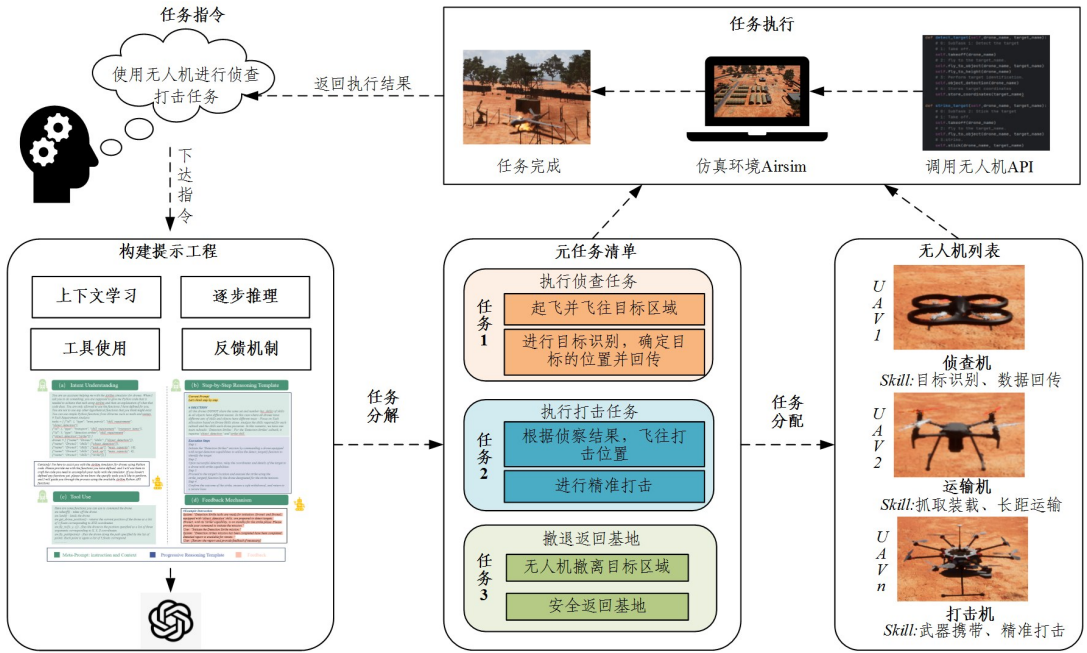


图1 SWARM-LLM 的总体框架图

Fig. 1 Overall framework of SWARM-LLM

#### 4.1 构建提示工程

提示(Prompt)指向大语言模型提供输入以引导其生成特定输出的文本或指令<sup>[20]</sup>。Prompt 的目的是引导模型产生所需的回应,以便更好地控制生成的输出,具体构建流程如

图2所示。本文通过构建适应于无人集群任务规划的 Prompt——规划链来优化任务分解、任务分配和任务执行等关键阶段的实施,从而提高整个无人集群系统的运行效率和任务完成质量,具体步骤如下。

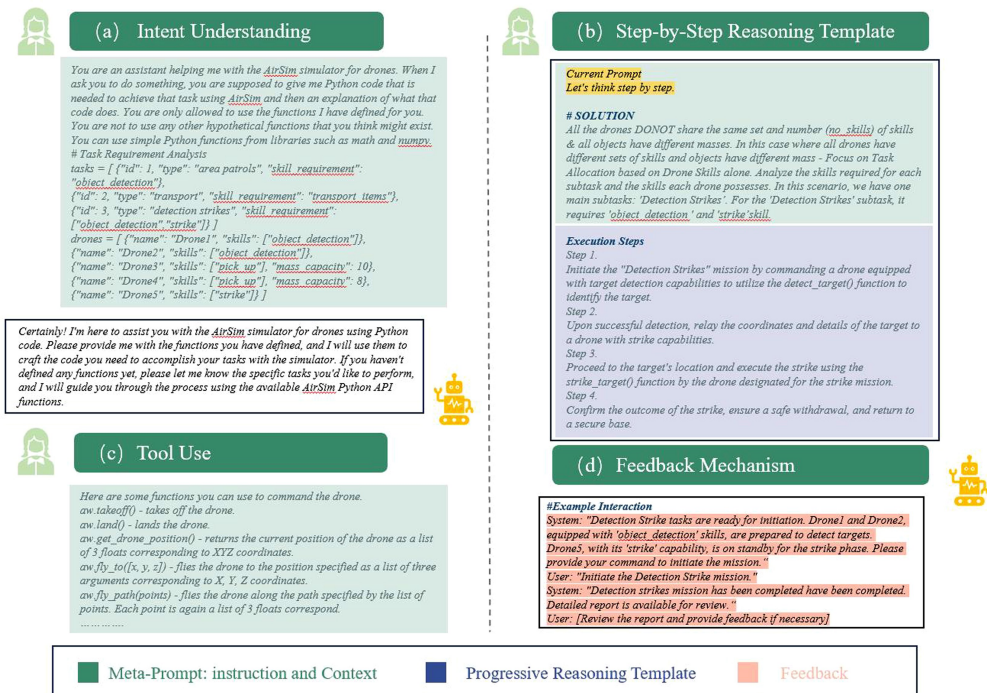


图2 提示工程构建框图

Fig. 2 Construction diagram of prompt engineering

#### 4.1.1 上下文学习

上下文学习涉及系统对任务的理解,以便在任务分解、任务分配和执行中提供适当的支持和指导。在这一过程中,系统首先进行任务需求分析,准确把握任务的具体需求,包括任务类型、技能要求以及无人机的特定能力和限制。接着,系统分析任务说明与输入文本之间的关联性,确定如何有效地引导大型语言模型生成符合任务要求的文本提示。此外,系统还需要具备上下文感知能力,考虑不同任务的特殊上下文要求,以在任务分配过程中提供准确的指导和支持。

#### 4.1.2 逐步推理

逐步推理模板允许系统按照逻辑顺序处理任务分解、任务分配和执行过程,确保任务流程的合理性和有效性。具体可以包括以下内容。

1)任务分配逻辑:提供了解决任务分配问题的思路和方法。指出无人机的技能集合各不相同,任务所需的技能也不相同,强调任务分配应该基于无人机自身的技能,以匹配任务的技能要求。

2)任务执行步骤设计:制定逻辑清晰的任务执行步骤,从任务开始,到系统根据任务要求分配无人机,再到无人机执行具体任务。

3)推理过程管理:引导系统按照步骤逐步推理和执行任务,确保整个过程的顺利进行和任务目标的达成。

#### 4.1.3 工具使用

工具使用部分的核心在于介绍无人机可调用的 API 工具集,以确保能够有效理解并执行各项任务。我们对无人机的 API 工具集进行详尽的功能介绍,如起飞、拍照、目标识别等动作,确保智能体在任务规划时能够明确需调用的工具,从而在执行过程中实现精确的 API 调用,以完成相应的动作。

#### 4.1.4 反馈机制

反馈机制是实现系统与用户之间有效互动的关键。在任务执行过程中,系统及时向用户提供反馈,包括任务启动通知、任务分配结果和执行情况更新。此外,还要能够响应用户的反馈和指令,调整任务分配或执行计划以确保任务执行的准确性。任务执行后,系统利用结果优化任务分配算法,从而提高系统在未来任务中的决策能力。

### 4.2 任务分解

任务分解主要是将高层次的任务指令分解为无人机集群可以执行的具体任务,这一过程的关键在于综合考虑任务特征、无人机的技能集以及既定目标,同时借鉴历史任务分解的样本,引导大语言模型对任务进行深入的分析 and 拆解。具体而言,任务特征详尽地描述了任务的类别以及任务所需技能,为任务分解提供了关键背景。同时,无人机的技能集涵盖了无人机的名称、基本技能、可携带载荷重量等多方面信息。此外,历史任务分解样本的引入展示了过去任务是如何被分解、规划和执行的,为当前任务的分解提供了具体的参考框架,从而提高了新任务分解的准确性和效率。

在这个阶段,SWARM-LLM 根据输入的高层次任务指令和样本提示,生成一系列可执行的子任务。当接收到用户

请求时,LLM 将用户请求转换为机器可处理的格式,分析用户请求,理解其意图,并将其分解为多个子任务,为后续的任务分配打下基础。

### 4.3 任务分配

任务分配阶段是整个规划流程的核心环节,其主要任务是依据无人机列表,对经过前期任务分解阶段细化的子任务进行合理分配。这一阶段的关键挑战在于确保任务分配的有效性,同时考虑到任务间的依赖关系、无人机的执行能力以及可用资源的限制。

首先,通过任务分解阶段确定了任务列表,该列表包括所有经过分解的子任务,每个子任务都明确了具体的执行要求和优先级。随后,构建了详尽的无人机列表,其中包含了每架无人机的名称、技能以及当前状态等相关信息。在此基础上,制定了任务分配策略,旨在根据无人机的技能和任务需求,合理分配无人机资源,确保每个子任务都能得到最适合的无人机来执行。为了实现这一目标,在 LLM 的提示中包含了分解任务的样本和相应的任务分配策略。这些策略描述了如何根据无人机的技能和任务需求,将可用的无人机分配给各个子任务。任务分配策略考虑了多种情况,包括:

1)单个无人机拥有执行子任务所需的所有技能,从而实现一对一的任务分配。

2)没有单个无人机具备完成子任务所需的所有技能,需要多个无人机协作完成同一任务。

3)无人机具备执行子任务的必要技能,但存在某些限制(例如重量限制),需要通过匹配限制来克服这些约束。

基于这些信息,大语言模型输出一个详尽的任务分配计划并产生相应的可执行代码,该计划详细描述了每个无人机的任务列表、执行顺序和资源需求。任务分配策略不仅能确保每个无人机都能高效地执行其分配的任务,而且还考虑到了任务间的相互依赖性和无人机的实际操作能力,从而最大化整个无人机团队的执行效率。

### 4.4 任务执行

任务执行阶段,无人机集群依据前期生成的任务分配计划开始执行各项任务。这一阶段的核心在于实时监控无人机的任务执行情况,以确保所有任务能够顺利完成。

在执行指令方面,无人机接收由大语言模型生成的具体可执行代码,并调用 API 完成相应动作。同时,反馈机制在整个任务执行过程中发挥着至关重要的作用。系统通过实时监控无人机的状态和任务进展情况,能够及时发现并反馈出现的问题。通过这种实时的监控和灵活的调整,SWARM-LLM 框架不仅确保了任务的顺利执行,还使得无人机团队能够高效地协同工作,完成一系列复杂的任务。

## 5 实验设置及任务设计

### 5.1 实验环境

为了评估 SWARM-LLM 框架在无人机集群任务规划中的性能,并促进与其他基线方法的定量比较,我们使用 Microsoft AirSim 模拟环境来评估基于 LLM 的任务规划框架在多无人机场景中的有效性。AirSim 是一个专为无人系统模拟而设计的仿真平台,它提供了一个极为逼真的仿真环境,可

利用 Unreal Engine 强大的物理引擎来精确模拟真实世界的飞行动力学;同时支持多架无人机的仿真,能够测试和研究无人机群的协同作业能力。此外, AirSim 提供了一套全面的 API 接口,能够通过编程方式控制无人机的行为。

## 5.2 任务设置

为了全面评估 SWARM-LLM 框架在不同任务复杂性下的性能,我们设计了一系列任务,以覆盖从单一无人机执行的基础任务到需要异构无人机团队协作的复杂任务。这些任务被分为 4 个类别,即元任务、简单任务、复杂任务和异构任务,旨在测试 SWARM-LLM 框架在多样化任务场景中的适应性和效率。

1)元任务。元任务是为单一无人机设计的,假设该无人机具备所有必要的技能和能力来独立完成任务。这类任务排除了多无人机协调的复杂性,专注于评估 SWARM-LLM 框架在指导单个无人机执行简单、明确任务时的效率和准确性,例如单个无人机执行目标识别任务。

2)简单任务。简单任务涉及多个目标或对象,可以分解为顺序或并行的子任务,但每个子任务仅由一个无人机执行。在这类任务中,所有无人机都具备执行任务所需的全部技能,任务规划的挑战在于如何合理地安排子任务的执行顺序和分配资源。简单任务旨在评估 SWARM-LLM 框架在处理多无人机执行多步骤顺序或并行任务时的性能,例如多个区域需要就近安排多个无人机执行区域巡逻任务。

3)复杂任务。复杂任务在执行策略上类似于简单任务,但任务的执行更加灵活,可以根据实际情况选择顺序、并行或混合的执行方式。这类任务中的无人机具有不同的专门技能和属性,能够处理与自身技能相匹配的子任务。复杂任务用于评估 SWARM-LLM 框架在处理需要考虑无人机特定技能和属性的任务时的适应性和智能性,例如物品运输任务,各个无人机的可携带重量不同,需要匹配相应的无人机执行相应的运输任务。

4)异构任务。异构任务是设计用来测试异构无人机团队的协作能力的,它在任务分解、多无人机参与和多个对象的存在等方面与复杂任务相似。不同的是,由于技能或属性的限制,单个无人机无法独立执行所有子任务,需要通过战略性的团队分配和协作来共同完成任务。异构任务是对 SWARM-LLM 框架在处理高度复杂、需要团队协作的任务时的性能的考验,例如,侦查打击任务需要派遣具备目标识别能力的无人机进行侦查并将数据传回,以协调其他无人机进行打击,实现团队协作。

## 5.3 评估指标

为了全面评估 SWARM-LLM 框架在无人机集群任务规划中的表现,本文参考了文献[21]提出的方法,采用了 3 个评估指标:任务成功率(Success Rate, SR)、动作执行性(Executability, Exe)和无人机利用率(Unmanned Aerial Vehicle Utilization Rate, UR)。

1)SR(任务成功率)。其表示在一定时间内,成功完成的符合预定目标和要求的任务数量与总任务数量的比例。在无人集群中,任务成功不仅意味着任务目标的达成,还需要考虑执行任务的无人机的利用率。具体的计算式为:

$$SR = \frac{T_{\text{success}}}{T_{\text{total}}}$$

其中,  $T_{\text{success}}$  是在满足无人机利用率条件下成功完成的任务数量,  $T_{\text{total}}$  是总任务数量。

2)Exe(动作执行性)。其表示任务计划中可以成功执行的动作比例。因为每个任务都可以被拆分为一系列具体的执行动作,所以这是一个衡量任务计划可执行性的指标,即使任务未能完全成功,可执行性也可能很高。具体计算式为:

$$Exe = \frac{A_{\text{success}}}{A_{\text{total}}}$$

其中,  $A_{\text{success}}$  是成功执行的动作数量,  $A_{\text{total}}$  是任务中预定的总动作数量。

3)UR(无人机利用率)。其表示在任务分配过程中,成功分配给具备必要技能的无人机的任务数与可分配任务总数的比例。具体计算式为:

$$UR = \frac{S}{N}$$

其中,  $S$  是成功分配给具备必要技能的无人机的任务数量,  $N$  是需要执行的所有任务的总数。

## 6 实验结果与分析

### 6.1 不同任务规划模型性能比较与分析

本文通过比较不同模型在无人机任务规划场景中的表现,评估了多种方法在元任务、简单任务、复杂任务和异构任务上的效果。具体模型包括 SWARM-LLM 系列(GPT-3.5-Turbo, GPT-4, GPT-4o, Claude3)、优化算法(遗传算法)和机器学习方法(决策树)。下文将详细分析每种方法的性能表现,其中 SWARM-LLM 系列为了测试不同大语言模型的能力,采用统一的提示策略来评估它们在不同任务下的性能(见表 1)。

表 1 不同 LLM 在不同任务场景下的性能评估

Table 1 Performance evaluation of different large language models(LLMs) in various task scenarios

Methods	元任务			简单任务			复杂任务			异构任务		
	SR	Exe	UR	SR	Exe	UR	SR	Exe	UR	SR	Exe	UR
SWARM-LLM (GPT3.5-Turbo)	1	1	1	0.916	0.944	1	<b>0.833</b>	0.900	0.833	0.750	0.795	0.833
SWARM-LLM(GPT-4)	1	1	1	<b>1</b>	<b>1</b>	<b>1</b>	<b>0.916</b>	0.950	0.916	<b>0.833</b>	<b>0.924</b>	<b>0.833</b>
SWARM-LLM(GPT-4o)	1	1	1	<b>1</b>	<b>1</b>	<b>1</b>	<b>0.916</b>	<b>0.966</b>	<b>1</b>	<b>0.833</b>	<b>0.924</b>	<b>0.833</b>
SWARM-LLM(Claude3)	1	1	1	0.833	0.861	0.916	0.750	0.750	0.750	0.583	0.757	0.666
优化算法(遗传算法)	0.583	1	0.583	0.250	0.291	0.500	0.166	0.233	0.333	0.083	0.378	0.250
机器学习(决策树)	1	1	1	0.500	0.694	0.500	0.416	0.533	0.500	0.250	0.500	0.416

总体来说,SWARM-LLM 系列模型(包括 GPT-3.5-Turbo,GPT-4,GPT-4o 和 Claude3)在无人集群任务规划中的表现均明显优于传统的优化算法和机器学习方法。传统的优化算法如遗传算法在元任务上的成功率较低,随着任务复杂性的增加,成功率进一步下降,特别是在复杂任务上的成功率仅为 0.083,显示出其在多无人机任务分配问题上的适用性有限。而机器学习方法如决策树在元任务上的成功率为 1,但在处理简单任务、复杂任务和异构任务时,成功率分别降至 0.500,0.416 和 0.250。尽管决策树能找到合适的无人机完成相应任务,但无法协调多个无人机同步或按顺序执行任务。这意味着决策树方法在单无人机任务规划中有一定效果,但在多无人机协作任务中表现不佳。相比之下,SWARM-LLM 系列模型不仅在任务规划的准确性和执行效率上表现出色,还能较好地处理多步骤任务和多无人机协作任务,和上述传统方法相比任务成功率平均提升了 47.8%。

另外,在 SWARM-LLM 框架下,本文一共使用了 4 种大模型方法:GPT-3.5-Turbo,GPT-4,GPT-4o 和 Claude3。其中 GPT3.5-Turbo 在处理基础任务时表现良好,能够生成合理的任务分配方案。然而,当任务复杂性增加,需要更深层次的推理和协调时,GPT3.5-Turbo 模型的性能有所下降。另外,GPT3.5-Turbo 模型由于上下文窗口较小,在处理长距离依赖和复杂任务规划时受限。相比之下,GPT-4/GPT-4o 在任务规划的准确性和复杂任务的处理上优于 GPT-3.5 模型,它在多步骤任务规划和多无人机协作任务中的表现更为稳定。GPT4 与 GPT-4o 的性能表现相差无几,两者在任务规划的成功率和执行率上均达到了令人满意的水平。然而,GPT-4o 在执行速度方面具有明显优势,这在需要快速决策和实时响应的任务场景中尤为重要。Claude3 则明显落后于其他模型,在任务规划的准确性和执行效率方面未能达到同等水平,特别是在复杂任务的处理上表现不佳。此外,Claude3 在 4 个模型中的反应速度也最慢。

综上,GPT-4 和 GPT-4o 在综合性能上表现最佳,尤其是在复杂和多步骤异构任务中的出色表现使其成为无人机任务规划的理想选择。尽管如此,GPT4 系列在面对极端复杂的任务或需要广泛领域知识的任务时,仍然面临着一定的挑战。

## 6.2 消融实验分析

本文通过消融实验评估不同提示方法对无人集群任务规划性能的影响。选择综合性能最好的 GPT-4o 作为基础大语言模型,针对 3 种提示策略,即少样本提示(Few-shot Prompting)、思维链(Chain of Thought, CoT)以及规划链(Planning Chain, PC)进行消融实验,以评估它们在不同任务难度级别下的成功率。其中,Few-shot Prompting 是文献[16]中的少样本提示方法,包含了基本的无人机指令;思维链(Chain of Thought, CoT)通过逐步推理和解释来解决复杂问题;规划链(Planning Chain, PC)是我们设计的一种适用于无人集群任务规划的提示工程方法,涵盖了从上下文理解到反馈机制的全过程。不同提示方法在各任务类型下的成功率如表 2 所列。

1) Few-shot Prompting: 采用少量样本提示来启动无人机的基本任务规划流程,使 SWARM-LLM 能够对无人机

执行的任务进行基本的理解和分析。实验证明在执行单一任务,如派遣单个无人机进行目标检测等基础任务时,此策略展现出了不错的任务成功率。当面临复杂的任务协调或需要多无人机协作的场景时,其成功率则显著下降。这表明基础任务分配策略在处理任务场景的深入规划或协同多个无人机方面存在局限。

2) Chain of Thought: 思维链通过逐步推理和解释来解决问题,使 SWARM-LLM 能够在处理复杂任务时进行逐步推理,逐层深入地解决问题,从而做出合理的任务规划。实验结果表明,CoT 在简单任务和部分复杂任务上的表现优于少样本提示。然而,当任务的复杂性进一步升级,特别是涉及到需要多无人机协同工作的异构任务时,任务成功率显著下降。另外,因为没有反馈机制,SWARM-LLM 在执行任务的过程中无法调整其推理路径,难以适应任务环境的变化和无人机之间的协作需求。

3) Planning Chain: 引入设计的适应于无人集群任务规划提示后,采用任务规划框架融合了对无人集群任务规划的深刻理解,并引导 LLM 通过以下 4 个关键阶段,即上下文理解、逐步推理、工具调用、反馈机制,使 SWARM-LLM 能够执行复杂的任务规划。实验结果表明,在应用了规划链后,无人集群能够更加精准地进行资源分配,灵活编排并执行多样化的任务,复杂任务和异构任务的任务成功率得到了显著提升,从而验证了 SWARM-LLM 在复杂任务规划和协调多无人机行动中的有效性和适应性。

表 2 使用不同提示下任务成功率的比较

Table 2 Comparison of task success rates with different prompts

Methods	元任务	简单任务	复杂任务	异构任务
Few-shot Prompting	1	0.400	0	0
Chain of Thought	1	0.916	0.666	0.416
Planning Chain	1	1	0.916	0.833

## 6.3 无人机任务流程的可视化展示

本节详细展现了 SWARM-LLM 框架下无人机任务规划和执行的全流程,揭示了无人机如何在不同任务场景中被有效分配、协调和执行任务。通过可视化展示,我们能够直观地理解无人机在各类任务中的具体操作和协同工作机制。

如图 3 所示,在任务开始前,不同类型的无人机整齐地停在起点位置,等待指令。每种无人机根据其任务要求和能力,被合理分配到不同的任务队列中。图 4 给出了无人机执行区域巡逻任务的场景。无人机就近飞往指定的目标区域,沿着预设的轨迹进行巡逻,并在这一过程中采集地面照片。图 5 给出了无人机执行物品运输任务的场景,它们根据自身的载重能力拾取目标物品,然后将其运输到指定的位置。图 6 和图 7 分别给出了无人机执行侦查、打击任务的场景,在此任务中,无人机不仅要进行巡逻和目标侦查,还需在识别到目标后迅速与打击机协调,执行精准打击任务。通过实时数据共享和任务协同,无人机能够高效完成复杂的任务目标。

通过可视化展示可以清晰地看出无人机根据各自的特性和能力被合理地指派任务,同时无人机之间通过行动协调,实现了复杂的任务目标。

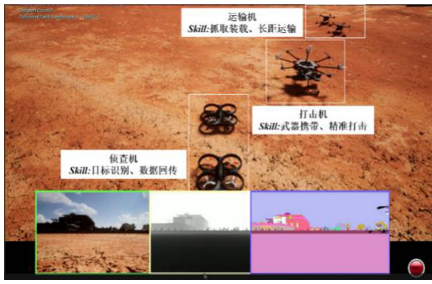


图3 无人机型别图

Fig. 3 Types of unmanned aerial vehicles(UAVs)

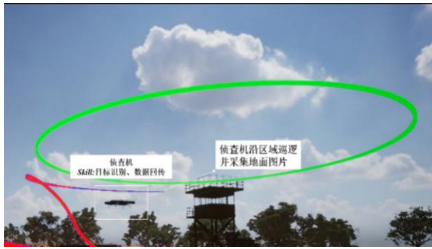


图4 无人机执行区域巡逻任务

Fig. 4 Unmanned aerial vehicle(UAV) executing area patrol mission



图5 无人机执行物品运输任务

Fig. 5 Unmanned aerial vehicle(UAV) executing cargo transportation task

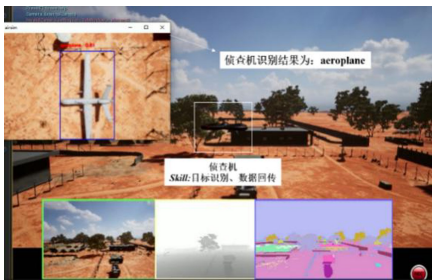


图6 无人机执行侦查任务

Fig. 6 Unmanned aerial vehicle(UAV) executing reconnaissance mission

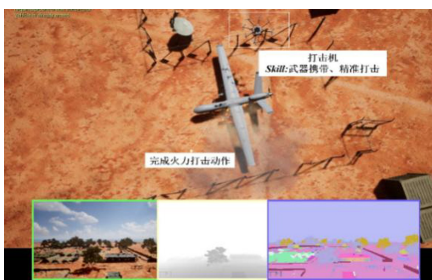


图7 无人机执行打击任务

Fig. 7 Unmanned aerial vehicle(UAV) executing strike mission

**结束语** 本文针对无人集群任务规划的背景,提出了 SWARM-LLM 框架,旨在测试大语言模型在解决异构无人机集群任务规划领域对复杂任务的理解、规划及执行能力。为充分考虑环境特性和无人机个体的能力,我们设计了适应于无人集群规划的提示工程,来生成切实可行的任务计划,并通过一系列模拟实验验证了该方法在处理不同复杂性任务指令方面的有效性。后续我们将继续探索 SWARM-LLM 在无人机之间实现实时动态任务规划的可能性,从而应对实时变化的任务需求和条件。此外,我们计划研究多智能体 LLM 框架,例如借鉴 X-agent 的思想,引入双循环机制从而优化任务规划过程。这将有助于提升 LLM 在复杂、动态环境中的适应性和决策能力,推动无人集群任务规划技术的发展。

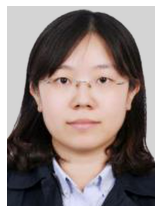
## 参考文献

- [1] JIANG B T, WEN G H, ZHOU J L, et al. Research Status and Prospects of Cross-Domain Collaborative Technology for Intelligent Unmanned Swarm Systems [J]. Chinese Engineering Science, 2024, 26(1): 117-126.
- [2] WANG X, WANG H, SUN Q, et al. Air-ground coordinated unmanned swarm systems: A multitasking framework for control design [J]. ISA Transactions, 2024, 145: 315-329.
- [3] ZHAO W, ZHOU K, TANG T, et al. A Survey of Large Language Models [J]. arXiv: 2303. 18223, 2023.
- [4] RIBEIRO M T, WU T, GUESTRIN C, et al. Beyond Accuracy: Behavioral Testing of NLP models with CheckList [J]. Computing Research Repository 2020(2020): 4902-4912.
- [5] LIANG J, HUANG W, XIA F, et al. Code as policies: Language model programs for embodied control [C] // 2023 IEEE International Conference on Robotics and Automation (ICRA). IEEE, 2023: 9493-9500.
- [6] QI X, LI B, FAN Y, et al. A survey of mission planning on UAVs systems based on multiple constraints [J]. CAAI Transactions on Intelligent Systems, 2020, 15(2): 204-217.
- [7] WANG Q, LI T, XU Y, et al. How to prevent malicious use of intelligent unmanned swarms? [J]. The Innovation, 2023, 4(2): 209-211.
- [8] CHEN J C, LING F Y, ZHANG Y, et al. Coverage path planning of heterogeneous unmanned aerial vehicles based on ant colony system [J]. Swarm and Evolutionary Computation, 2022, 69: 101005.
- [9] XIA Y S, SHI J M. Vehicle-Aircraft Collaborative Multi-Area Coverage Reconnaissance Path Planning Method [J]. Journal of Command and Control, 2020, 6(4): 372-380.
- [10] ZHANG T, LI C, MA D, et al. An optimal task management and control scheme for military operations with dynamic game strategy [J]. Aerospace Science and Technology, 2021, 115: 106815.
- [11] TANG J, LIANG Y, LI K. Dynamic Scene Path Planning of UAVs Based on Deep Reinforcement Learning [J]. Drones, 2024, 8(2): 60.
- [12] LIU C, MA X, GAO X, et al. Distributed energy-efficient multi-UAV navigation for long-term communication coverage by deep reinforcement learning [J]. IEEE Transactions on Mobile Computing, 2020, 19(6): 1274-1285.

- [13] HE L, AOUF N, SONG B. Explainable Deep Reinforcement Learning for UAV autonomous path planning[J]. Aero-space Science and Technology, 2021, 118: 107052.
- [14] FU D, LI X, WEN L, et al. Drive like a human; Rethinking autonomous driving with large language models[C]// Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2024: 910-919.
- [15] CHEN L, SINAVSKI O, HÜNERMANN J, et al. Driving with llms: Fusing object-level vector modality for explainable autonomous driving[J]. arXiv: 2310. 01957, 2023.
- [16] VEMPRALA S, BONATTI R, BUCKER A, et al. Chatgpt for robotics: Design principles and model abilities[J]. arXiv: 2306. 17582, 2023.
- [17] SHAH S, DEY D, LOVETT C, et al. Airsim: High-fidelity visual and physical simulation for autonomous vehicles [C] // Field and Service Robotics; Results of the 11th International Conference. 2017: 621-635.
- [18] TAZIR M L, MANCAS M, DUTOIT T. From Words to Flight: Integrating OpenAI ChatGPT with PX4/Gazebo for Natural Language-Based Drone Control[C]// International Workshop on Computer Science and Engineering. 2023.
- [19] CHEN G, YU X, ZHONG L. TypeFly: Flying Drones with Large Language Model[J]. arXiv: 2312. 14950, 2023.
- [20] YE Q, AXMED M, PRYZANT R, et al. Prompt engineering a prompt engineer[J]. arXiv: 2311. 05661, 2023.
- [21] KANNAN S S, VENKATESH V L N, MIN B C. SWARM-llm: SWARM multi-agent robot task planning using large language models[J]. arXiv: 2309. 10062, 2023.



**LI Tingting**, born in 1997, master. Her main research interests include group decision-making intelligence and large model theory and application.



**WANG Qi**, born in 1985, Ph.D, associate professor, Ph.D supervisor, is a senior member of CCF (No. 77141M). Her main research interests include intelligent wireless networks and LLM assisted decision making.

(责任编辑:喻黎)

## CCF YOCSEF 换届选举完成,贺瑞君当选新一任 AC 主席

2024年12月14日下午,CCF YOCSEF第二十七届学术委员会第二次会议在中国科学院计算技术研究所四层报告厅举行,会议进行了AC委员的选举以及学术委员会主席会议成员的换届选举。

经过差额竞选和无记名投票,陈果(湖南大学)、王少华(中国科学院空天信息创新研究院)当选为学术界AC委员;汪卫国(新华三集团)、刘文懋(绿盟科技)当选为非学术界AC委员。AC委员的选举由YOCSEF现任主席沈华伟主持。

YOCSEF现任AC委员、人民邮电出版社贺瑞君当选新一届YOCSEF(2025-2026)学术委员会主席。中国人民大学范举、北京交通大学金一、北京智谱华章科技有限公司鄢兴雨当选AC副主席。首都师范大学唐晓岚、微软亚洲研究院陈昊当选学术秘书。主席会议成员的换届选举由YOCSEF秘书长谭晓生主持。

AC委员的选举结果已通过CCF YOCSEF指导委员会批准;新一届CCF YOCSEF学术委员会主席会议成员的选举结果经YOCSEF主席和秘书长共同确认后,已通过CCF秘书长批准,他们将于2025年5月上任(AC委员任期三年,主席会议成员任期一年)。



新一届主席会议成员合影

左起:唐晓岚、范举、金一、贺瑞君、鄢兴雨、陈昊