无人机语义安全研究综述

邓余婉祺 王 越 杨 超 冯鹏斌
杨嘉慧 王利娟 李兴华 马建峰
（西安电子科技大学网络与信息安全学院 西安 710071）

摘 要 随着无人机技术的迅速迭代和各类智能算法的不断涌现，无人机安全问题面临着持续的挑战。近年来，无人机语义安全问题逐渐引起了国内外学术界的关注，该问题超出了传统的软件安全和传感器安全范畴，重点关注无人机系统的行为是否与其预期语义保持一致。语义安全问题具有高度隐蔽性、易被利用且产生的后果难以预测，因此亟需引起重视。当前，无人机语义安全研究仍处于起步阶段，现有研究成果尚不足以全面解决无人机系统中的语义安全问题，且缺乏相关的综述性研究。本文首先介绍无人机的飞行控制原理、系统功能和任务，以明确无人机语义安全的研究对象，并定义“无人机语义安全”这一新概念。接着，依托现有研究工作，本文总结了研究支撑工具，并提出一种研究分类方法，将无人机语义安全研究问题拆解为安全检测、漏洞修复、实时防护三大类别，并进一步细分出八个关键研究环节，构建完整的语义安全研究框架。最后，本文基于所建立的研究框架对现有研究进行系统性分析，重点探讨了与传统安全相比，语义安全研究所面临的挑战及可能的解决思路，并总结未来研究方向。本研究旨在为该领域研究者提供全新的视角和思路，填补无人机安全领域空白，推动语义安全研究深化发展，以满足日益复杂环境下无人机整体安全需求。

关键词 无人机；信息物理系统；无人系统安全；防护体系；语义安全；功能安全
中图法分类号 TP309 **DOI号** 10.11897/SP.J.1016.2025.01495

A Survey on Semantic Safety of Unmanned Aerial Vehicle

DENG Yu-Wan-Qi WANG Yue YANG Chao FENG Peng-Bin
YANG Jia-Hui WANG Li-Juan LI Xing-Hua MA Jian-Feng
（School of Cyber Engineering, Xidian University, Xi’an 710071）

Abstract With the rapid iteration of drone technology and the continuous emergence of various intelligent algorithms, drone safety issues are being faced with ongoing challenges. In recent years, the issue of semantic safety of drones has gradually attracted attention from the academic community at home and abroad. This problem goes beyond the traditional scope of software security and sensor security, and focuses on whether the behavior of drone systems is consistent with their expected semantics. Semantic safety issues are highly covert, easily exploitable, and have unpredictable consequences, so they urgently need to be taken seriously. At present, the research on semantic safety of unmanned aerial vehicles is still in its infancy, and the existing

收稿日期：2024-04-29；在线发布日期：2025-02-21。本课题得到国家自然科学基金-重点项目（6223000226）、国家自然科学基金-杰出青年科学基金项目（62125205）、国家自然科学基金-青年科学基金项目（62402364）、陕西省技术创新引导计划（2024QCY-KXJ-171）、陕西省自然科学基础研究计划（2023-JC-QN-0759）和中央高校基本科研业务费专项资金（ZYTS24138）资助。**邓余婉祺**，博士研究生，中国计算机学会（CCF）会员，主要研究领域为无人系统安全、智能无人集群安全。E-mail：dywq2021@163.com。**王 越**，博士研究生，中国计算机学会（CCF）会员，主要研究领域为无人系统安全、无人系统智能算法安全。**杨 超**（通信作者），博士，教授，中国计算机学会（CCF）会员，主要研究领域为无人智能系统安全、安全性测试。E-mail：chaoyang@xidian.edu.cn。**冯鹏斌**，博士，讲师，主要研究领域为恶意软件检测、漏洞检测。**杨嘉慧**，硕士研究生，主要研究领域为无人系统安全。**王利娟**，博士，讲师，主要研究领域为机器学习、智能计算、无人机自主智能控制。**李兴华**，博士，教授，中国计算机学会（CCF）会员，主要研究领域为无线网络安全。**马建峰**，博士，长江学者特聘教授，中国计算机学会（CCF）会员，主要研究领域为无线网络安全、移动智能系统安全。

research results are not sufficient to comprehensively solve the semantic safety problems in unmanned aerial vehicle systems, and there is a lack of relevant review studies. This article first introduces the flight control principles, system functions, and tasks of drones, in order to clarify the research object of drone semantic safety and define the new concept of "drone semantic safety". Drone semantic safety refers to ensuring that all expected behaviors are accurately understood and executed, and maintaining system stability and reliability under unforeseeable conditions. The root cause of semantic safety issues usually occurs in the design phase of software and algorithms, requiring protection throughout the entire lifecycle of drone system design, development, testing, and maintenance. Furthermore, based on existing research work, this article summarizes research support tools. Commonly used open-source support tools include ArduPilot, PX4, Gazebo, AirSim, etc. These tools are used to simulate real flight conditions, generate test cases, and validate system behavior. Then, this article proposes a research classification method to break down the semantic safety research problems of unmanned aerial vehicles into three categories: (1) safety detection, systematically identifying and mining abnormal behaviors of unmanned aerial vehicle systems at the semantic level; (2) vulnerability repair, implementing effective repair strategies for clear semantic safety issues to ensure that drone systems follow correct semantic rules; (3) real-time protection, preventing potential safety threats by designing and applying various defense mechanisms. This method is further subdivided into eight key research stages, including semantic target determination, test case generation, abnormal behavior judgment, error attribution localization, repair execution, repair verification, runtime monitoring, and real-time recovery, to construct a complete semantic safety research framework. Finally, based on the established research framework, this article systematically analyzes existing research, focusing on exploring the challenges and possible solutions faced by semantic safety research compared to traditional security, and summarizes future research directions: (1) exploring semantic safety in new scenarios, (2) artificial intelligence semantic safety, (3) semantic safety of drone swarms, (4) construction of unmanned aerial vehicle semantic safety assessment system, (5) strategies for addressing semantic safety threats posed by drones. This article introduces and summarizes existing research work to help researchers in the field of drone safety understand the current research status, conduct exploratory and innovative research in semantic safety, promote the deepening development of semantic safety research, and meet the overall safety needs of drones in increasingly complex environments.

Keywords unmanned aerial vehicle; cyber-physical system; unmanned system safety; protection system; semantic safety; functional safety

1 引 言

无人机(Unmanned Aerial Vehicle,UAV)在民用和军事领域的应用已经广泛覆盖航拍摄影、物流运输、工业巡检、搜索救援和军事打击等多种任务。尤其是在俄乌战争中,数百万架无人机的部署显著提升了战场态势感知与精确打击能力,深刻改变了传统战争模式。随着无人机技术的快速发展,社会对无人机自主性、智能化及其在各种场景中执行复杂任务的需求不断增加[1-4],与此同时给无人机系统安全带来了新挑战。

无人机端系统安全是云-网-端协同安全架构的基础。根据问题源头的不同,端系统安全可以划分为三类:软件安全、传感器安全和语义安全。软件安全主要防范系统软件遭受恶意访问或破坏,例如权限滥用[5]或内存损坏[6]可能导致飞行控制系统崩溃;传感器安全则要求无人机能准确感知环境,防止传感器欺骗[7-8]或物理干扰[9]等外部攻击的影响。语义安全旨在确保无人机能够准确理解和执行预期行

为，其与传统安全的核心区别在于，语义安全问题通常源于无人机自身的缺陷，而非外部恶意攻击，如图1所示。

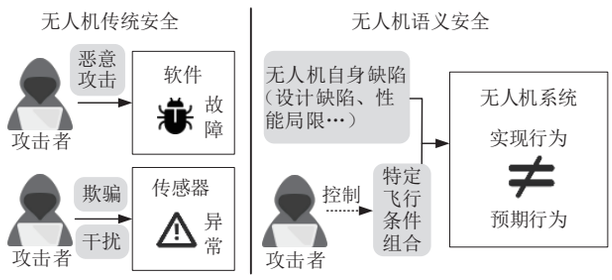


图1　无人机语义安全和传统安全威胁模型对比

长期以来，学术界和工业界主要聚焦于软件安全和传感器安全[10-13]，而对由设计缺陷或性能限制引发的语义安全问题关注不足。然而，这类安全缺陷通常隐蔽性极强，后果难以预测。例如，波音737Max坠机事故[14-15]就是因防失速功能被错误激活，设计缺陷使飞行员无法及时排查异常并解除风险。此外，这些缺陷易被利用，只需发送一条看似合法的地面站指令即可能导致无人机崩溃[16]。近年来，无人机语义安全问题逐渐受到关注。2019年，Taegyu等人在无人机系统中首次发现控制语义漏洞，指出控制变量在特定范围内赋值会导致无人机整体异常[16]。2021年，Hyungsub等人首度聚焦安全策略语义，通过模糊测试揭示无人机违反安全策略的条件[17]。2023年，Yao等人发现无人机群体传播漏洞，并利用该漏洞成功引发事故[18]。此外，也有研究提出了修复和防范语义安全漏洞的方法，如补丁[19-20]和运行时恢复[21]。

综述是研究者全面了解某一领域的重要途径。目前，无人机安全领域的综述已涵盖了信息安全、网络安全、自主安全等多个视角。文献[22]从信息安全角度详细阐述了无人机在传感器、通信、软件和网络四个方面的安全威胁，成为国内无人机安全研究系统化的里程碑。文献[23]从网络安全视角将无人机的安全威胁划分为通信网络安全、软件安全、有效载荷安全、智能安全四大类，系统总结了各类网络攻击与防御技术，为未来无人机网络攻防与安全研究提供重要支持。文献[24]前瞻性地提出了“无人机自主安全”概念，构建无人机系统自主安全模型和能力分级标准，对无人机智能化发展和应对多域安全威胁具有深远意义。文献[25]首次对PX4和ArduPilot等开源平台漏洞进行了大规模实证研究，归纳出八类无人机特有漏洞、常见缺陷模式及修复策略，为理论研究提供了实证依据。文献[26]从控制学科视角，针对干扰估计、故障检测、抗干扰控制、容错控制和任务重构等关键问题提出了解决方案，为构建无人机安全控制系统架构提供了重要理论支持。尽管现有的综述和实证研究从多个角度探讨了无人机安全问题，涵盖了传感器安全、软件安全以及多域安全等多个方面，但尚未系统地从语义安全维度进行探讨。这一空白为无人机安全带来了潜在风险。

为此，本文全面调研了2018年至2024年上半年已发表的无人机语义安全相关研究，并进行系统的梳理和综述。本文将语义安全研究问题拆解为若干阶段和步骤，总结当前研究方法，初步探讨无人机语义安全的研究框架、面临的挑战及可能的解决思路。本文旨在为该领域的研究者提供全新的视角和思路，填补无人机安全领域空白，推动语义安全研究深化发展，以满足日益复杂环境下无人机整体安全需求。

2　无人机语义安全研究架构

2.1　语义研究对象

无人机语义安全错误可以简单理解为无人机未按预期功能和任务执行。无人机任务的成功依赖于各系统功能的有效运行，这些功能由特定算法实现，算法运行则需要飞行控制系统中软硬件组件的协同作用，如图2所示。因此，理解无人机飞行控制系统原理及其功能任务是开展无人机语义安全研究的基础。本节将依次介绍无人机系统组成、飞行控制原理、系统功能及典型任务。

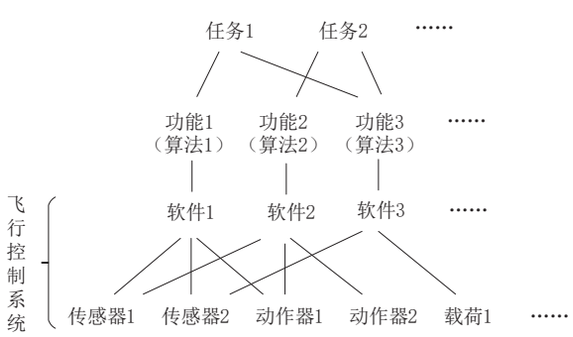


图2　无人机系统多层依赖关系

在自动控制领域，飞行控制系统的研究主要集中于硬件设计和动力系统等底层技术[27-28]，而安全控制研究侧重于新型控制理论，旨在确保无人机在干扰或故障条件下具备抗扰和自愈能力[26,29]。相比之下，语义安全研究更关注功能的正确实现以及安

全关键场景的识别。因此，本综述研究重点在于高层任务、功能及算法实现的准确性，而非底层控制原理的细节，并将稳定飞行视为无人机控制层面的基本功能。

(1)无人机飞控系统

无人机是一种无需人员在内部操作，依靠预设程序和遥控指令运行的航空器。具备自主飞行能力的无人机可在无持续人为干预的情况下执行任务。根据任务需求和规模，无人机可分为高空长航时、中高空长航时、战术、小型和微型五类[23]。按飞行平台构型，常见类型包括固定翼、多旋翼、直升机、伞翼、扑翼和飞艇等，其中固定翼与多旋翼最为常见。各类型无人机均包含三个关键组成部分：传感器、决策中心和动作器，如图3所示。

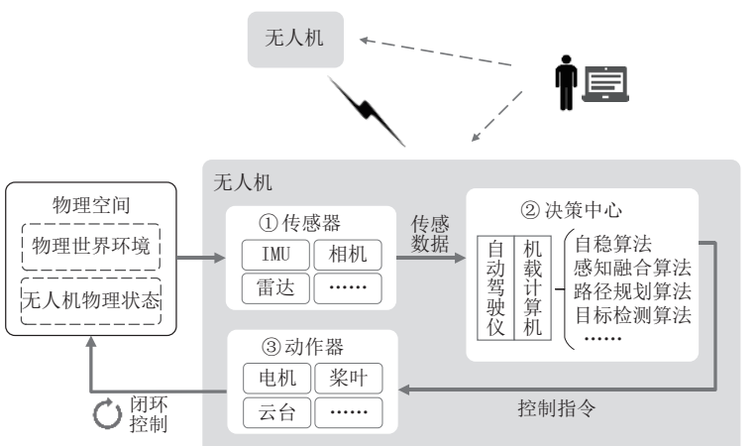


图3 无人机飞行控制系统结构

①传感器：用于获取飞行姿态和感知物理环境，并将数据传输至决策中心。常见传感器包括：惯性测量单元(Inertial Measurement Unit，IMU)，由三轴加速度计、三轴陀螺仪、电子罗盘或磁力计组成，用于获取多旋翼的姿态信息；全球定位系统(Global Positioning System，GPS)模块和磁罗盘用于卫星定位和航向确定；气压计用于测量飞行高度；双目相机和激光雷达等用于构建环境三维模型和获取深度信息。

②决策中心：决策中心也称为控制中心或控制器，是无人机的“大脑”，由自动驾驶仪(Autopilot)提供基础控制，机载计算机(Onboard Computer)提供计算能力。它处理传感器数据和遥控指令，通过一系列算法运算将这些信息转换为控制指令，以维持无人机飞行稳定和执行各项功能。决策中心内集成的算法包括自稳算法、感知融合算法、路径规划算法、目标检测与跟踪算法等。一些高自主化无人机还配备人工智能与深度学习算法以提升智能化水平。

③动作器：动作器负责执行决策中心的指令，实现各种飞行动作。固定翼无人机的动作器主要包括产生推力的电机和用于控制副翼转动的舵机；多旋翼无人机通过调节电机和桨叶的转速与转向来实现六自由度的精确控制。

具备自主飞行能力的单无人机采用闭环控制，持续监测自身状态与外部环境，并与预设目标进行对比，调整控制信号以维持期望状态，如图3所示。在此基础上，无人机操作员和地面站可通过远程无线遥控方式进行适时干预，以应对突发情况或调整飞行参数。地面站通常具备计算能力，为无人机的高效运行提供支持。多个无人机通过通信系统实现互联与合作，并由群控制算法进行协同控制。

(2)无人机系统功能

无人机的基本功能是指为了有效执行特定任务而必须具备的一系列核心功能。这些功能可以划分为组件级功能和系统级功能，依据其作用层次与协作性质进行区分。组件级功能指的是决策中枢以外的各个单一组件或模块的独立运行功能，例如双目相机的深度估计功能。系统级功能则依赖于各组件之间的紧密协作，通过闭环控制系统和群控技术实现更高层次的任务需求，如自动避障和编队飞行。本研究重点关注系统级功能，图4展示了单无人机和多无人机的系统级功能。

无人机的姿态控制、安全约束和故障保护是其核心基础功能，这些功能对无人机的飞行安全和稳定性至关重要。导航功能作为无人机自主执行任务

的关键，通常利用全球定位系统、惯性导航系统、光学流、激光雷达和视觉识别等技术手段，实时获取无人机的位置和环境状态，进行定位建图、航迹规划和自主避障。无人机载荷装置包括摄像机、传感器、投掷设备和云台等，是实现任务的关键部分，通过载荷控制可以执行采集、监控、观测和攻击等操作，以满足多样化和专业化的任务需求。

多无人机系统功能包括编队控制、群任务规划、群路径规划与避障以及群协同。编队控制确保无人机间保持预设的距离和相对方向，实现集群的整体协调飞行。群任务规划涉及高效的任务分配和调度策略设计，确保每架无人机明确其任务目标，并在动态环境中调整任务执行顺序和方式，以最大化整体任务效益。群路径规划与避障关注在保证集群整体效率的同时，使所有无人机遵循最优路径飞行，并能在遇到障碍物时迅速做出反应。群协同是多无人机系统的核心优势，它使各无人机不仅能够独立完成自身任务，还能通过共享信息和协作机制形成集体智慧，共同应对复杂任务，从而展现出远超单个无人机的性能优势。

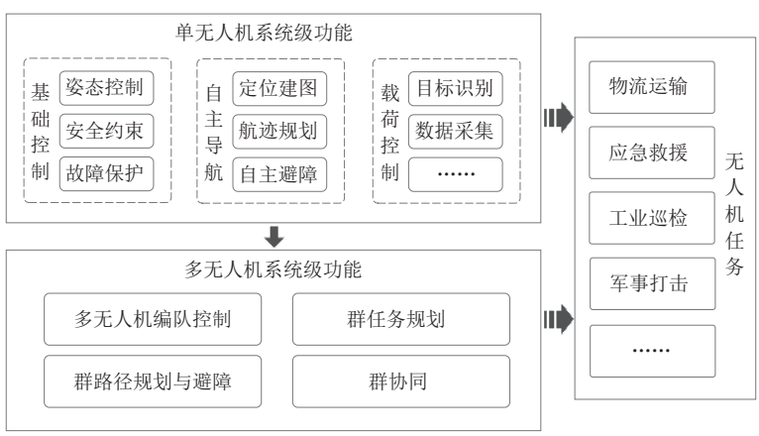


图4　无人机系统级功能和任务

(3)无人机典型任务

无人机各项功能的集成使其能够在多元化的应用场景下执行特定任务，如图4所示，包括但不限于航拍摄影、农业植保、环保监测、地理测绘、物流运输、应急救援、工业巡检、编队表演、军事打击。以下介绍几个典型应用场景中无人机所承担的关键任务及涉及的主要功能。

①物流运输：得益于低空路径规划和自主导航能力，无人机在物流运输领域展现出显著优势。无人机能够绕过交通拥堵区域，实现安全、可靠和高效的货物配送。亚马逊于2013年启动了Prime Air无人机物流配送项目，而美团无人机已经在深圳、上海等城市的商圈、社区写字楼及景区提供配送服务。在偏远地区和紧急情况下，无人机的优势更加明显，例如在地震后，无人机能迅速运送医疗物资，显著提升救援效率。

②工业巡检：在电力设施巡查和石油管道检测等领域，无人机凭借其自主飞行和载荷控制能力，可执行长时间、大范围的精细化巡检任务。中国石油大庆油田已应用无人机进行管道巡检，并推动无人机全自动智能巡检系统的应用。2024年1月，该油田开展了首个无人机智能机场应用实验，通过自动化巡检系统实时判断机采井启停状态，并对人员、牲畜入侵、油水井及设备泄漏等异常情况进行预警，从而减少了90%的人工巡检工作量。

③军事支持：无人机在区域覆盖与监视任务中，通过群任务规划和协同功能，实现了实时战略信息传递与协作。美国的“进攻性蜂群使能战术”(OFFSET)项目利用蜂群无人机执行情报、监视和侦查任务，从而提升地面部队在城市作战中的效率和生存能力。在远程军事打击任务中，无人机的导航和载荷控制功能确保了弹药的精确投送及其他作战行动的能力。例如，中国兵器装备集团开发的蜂群无人机发射车是一种高机动性蜂群武器系统，可用于侦察、区域封控和精确打击等多种任务。

2.2　语义安全定义

语义指语言形式所表达的含义。最初在密码学领域，语义安全(Semantic Security)要求信息的语义不被泄露，即若已知某段密文不会暴露该文段的任何其他信息，则称该密文具备语义安全性。而在无人机领域，语义安全(Semantic Safety)关注的是“行为”的语义，要求无人机系统能够精确理解和执行预

期行为。文献[16]首次探讨了这类问题，并针对无人机系统提出了“语义安全”的概念。该文献指出，这种“语义”问题代表了一个全新的研究方向，与传统的“语法”问题正交。此后，一些研究也采用了这一概念[17,21,24,30-31]。

无人机语义安全的定义如下：无人机语义安全（Semantic Safety of Unmanned Aerial Vehicles）是指在无人机系统的操作过程中，确保所有预期行为能够被准确理解和执行，并在不可预见条件下维持系统的稳定性和可靠性。通过实现语义安全，能够使无人机系统在复杂环境中保持可控，安全高效地执行既定任务，从而降低事故发生的风险。

将无人机系统行为划分为控制层、功能层和任务层，在这三个层面上讨论语义安全目标，不仅便于研究者清晰理解和描述无人机的安全需求，还能从深度和广度上充分覆盖无人机从基础控制到复杂任务执行的关键安全要求，如表1所示。控制层语义是指无人机控制系统所遵循的具体控制目标含义，其语义安全目标在于确保控制系统行为与开发者或操作员的意图保持一致。在该层面上发生语义安全问题，可能直接导致无人机失控，进而引发坠机或其他安全事故。对于无人机群，控制层语义指的是群控算法对群体行为的控制目标的含义。功能层语义指各功能模块预定目标的含义，其语义安全目标在于通过在系统层面组织各个组件和模块协同工作，或通过无人机群内的各个无人机相互协作，确保功能实现的正确性与完整性。任务层语义侧重于具体任务的内容和含义，其语义安全目标是确保无人机能够理解任务要求并适应环境变化，以保障任务顺利完成。

表1 无人机语义安全层面

语义层面	目标	后果	文献
控制层	控制无人机飞行姿态（无人机群体行为）与开发者和操作员的意图一致	飞行不稳定或失控	[16][32][33][34][35]
功能层	各组件（各无人机）有效协同以确保功能实现的正确性和完整性	功能故障	[17][18][19][20][30][35][36][37][38]
任务层	理解任务要求并适应环境变化以确保任务顺利完成	任务失败	[21][31][39]

无人机语义安全问题的根源通常出现在软件（即控制程序）和算法的设计阶段，例如未充考虑软硬件不一致问题[25]，或算法本身存在逻辑缺陷[17]等。当系统存在语义安全缺陷时，攻击者可能通过操控飞行条件绕过现有防御机制，干扰或破坏无人机系统[16]。此外，在恶劣天气等特定条件下，即使没有恶意攻击者，语义安全漏洞也可能被触发[30]。建立无人机语义安全防护体系，需要贯穿无人机系统设计、开发、测试及维护的全生命周期。

尽管物理安全、信息安全[22]和网络安全[23]等领域的研究各具重要意义，但这些领域相对独立，难以在应对跨领域复杂威胁时提供全面的解决方案。无人机语义安全作为一个高层次的综合概念，旨在分析和管理多个安全因素之间的相互作用及复合效应，以增强无人机系统的整体安全性。

2.3 研究支撑工具

仿真技术为研究者提供可控且可重复的环境，用以模拟真实的飞行条件。目前，针对无人机语义安全的研究工作普遍依赖于仿真技术，其相关开源支撑工具列于表2中。

飞控系统是控制无人机飞行姿态和运动方向的核心组件，负责无人机的起飞、空中飞行、任务执行和返航等整个飞行过程，也被称为自动驾驶仪。其中，ArduPilot和PX4是主流的开源飞控软件，支持多种无人平台，包括多旋翼、固定翼、直升机、地面机器人和无人艇，并提供基础的控制和导航功能。仿真平台用于模拟真实飞行条件。其中，Gazebo是最常用的物理仿真平台，可模拟障碍物、天气等物理世界环境。AirSim是一款基于游戏引擎开发的仿真平台，它不仅提供了物理仿真功能，还拥有非常出色的视觉仿真效果。地面站是管理和控制无人机的重要设备，用于上传任务、发送指令和监控飞行状态等。目前的主流开源飞控系统基于MAVLink协议通信，Pymavlink是该协议的python实现库，而MAVProxy和QGroundControl是支持该协议的地面站软件。高层算法是指运行在机载计算机上以实现各种功能或任务的算法，在语义安全领域中它们是应用层面的研究对象。由于此类算法种类繁多且复杂，我们选取了3个代表性算法列于表2中。EGO-Planner是当前开源社区中最知名的无人机路径规划器之一，适用于ROS[40]环境，并可直接部署到实际无人机上。EGO-Swarm是基于EGO-Planner扩展的分布式自主导航算法，专门用于无人机群。Vicsek Swarm是一种高度抽象化的无人机群算法，在SwarmLab[41]模拟器中可进行仿真。此外，ROS（Robot Operating System）是当前主流的开

表2　无人机语义安全研究相关开源支撑工具

工具类型	名称	描述
飞控系统	ArduPilot[44]	支持多类型无人机、无人车、无人艇的多功能自动驾驶仪
	PX4[45]	支持多类型无人机、无人车、无人艇的多功能自动驾驶仪
	Paparazzi[46]	用于固定翼、多旋翼的自动驾驶仪
仿真平台	Gazebo[47]	机器人仿真环境
	AirSim[48]	高保真视觉和物理仿真环境
	SITL[49]	ArduPilot内置的模拟器，可以模拟各种飞行条件和场景
	SwarmLab[41]	Matlab编写的无人机群模拟器
	JSBSim[50]	飞行器动力学仿真软件库
地面站	Pymavlink[51]	MAVLink协议python实现库
	PPRZLINK[52]	Paparazzi使用的通信库
	MAVProxy[53]	地面控制站软件
	QGroundControl[54]	地面控制站软件
高层算法	EGO-Planner[39]	轻量级四旋翼飞行器局部规划器
	EGO-Swarm[36]	支持多障碍物场景的无人机群分布式自主导航算法
	Vicsek Swarm[55]	受限环境中的自主优化无人机群
其他	ROS[40]	机器人操作系统
	LLVM[56]	模块化、可重用的编译器

源机器人操作系统框架，为各种类型机器人的构建提供灵活、可扩展和分布式的环境。LLVM是一个模块化、可重用的编译器，在文献[17]、[19-20]、[32]、[42-43]中用于对源代码进行静态分析。

2.4　研究问题分类

PDRR（Protect-Detect-Response-Recovery）、IPDRR（Identify-PDRR）和WPDRRC（Warn-PDRR-Counterattack）等安全防护模型的核心特质在于依据安全事件的生命周期进行阶段划分。这种阶段划分方法为安全防护提供了系统化的框架，有助于在不同的阶段采取相应的措施。基于这一原则，我们将无人机语义安全研究领域分为三类：无人机语义安全检测、漏洞修复及安全实时防护。

（1）语义安全检测：通过一系列测试技术与方法，系统性地检测和分析无人机系统在语义层面的脆弱性和异常行为。此阶段的研究主要集中在开发和应用各种检测工具和技术，以有效发现系统中的语义缺陷，为后续的安全改进提供基础数据。

（2）语义漏洞修复：针对已识别的语义安全问题进行深入分析，并制定出能解决这些问题的具体措施。此阶段不仅包括对已知问题的修补，还包括优化系统设计，以确保无人机系统在运行过程中能遵循正确的语义规则，提升系统的整体安全性。

（3）安全实时防护：采取预防性的安全措施，包括从系统设计阶段开始考虑语义安全，并在运行时实施实时监控和防护策略。此阶段研究重点在于完善系统防御机制，使其具备对新型威胁的适应能力，确保无人机系统在复杂环境中持续安全运行。

该分类方法依循安全生命周期的不同阶段，覆盖了发现、解决及长期防御的全过程。基于这三大类研究问题，进一步细分出八个关键研究环节，如图5所示，详见第3～5节。

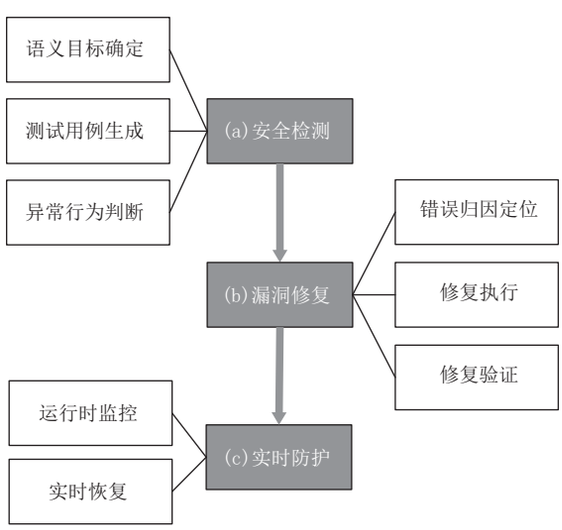


图5　无人机语义安全研究分类

3　无人机语义安全检测

无人机语义安全问题涉及数据解释错误、控制指令混淆等隐患，具有隐蔽性和根因复杂性，单纯依

赖人工检测手段往往难以全面排查。因此，语义安全检测通常借助自动化测试技术，通过设计并执行大量测试用例，高效地发现并验证语义漏洞。

在传统软件的自动化测试中，被测软件的输入与预设输出之间存在明确的对应关系，可以通过比较实际输出和预期输出的一致性来识别潜在错误。然而，传统软件的自动化测试技术难以直接应用于无人机系统，原因有三：其一，无人机系统运行在高度动态和复杂的环境中，输入空间庞大，包括控制参数、环境条件和飞行轨迹等多种因素；其二，无人机系统输出的正确性评估标准更为复杂和模糊，可能涉及飞行稳定性、安全性、任务完成度等多个维度；其三，无人机系统的输入与输出之间的关系存在很大的不确定性，两次看似相同的输入可能产生完全不同的输出。因此，针对无人机系统的测试，需要寻求更适合其特性的测试方法。

总体来说，无人机语义安全检测（图5(a)）包括三个关键环节：环节一为确定语义安全目标；环节二为制定生成测试用例的策略；环节三为设定异常行为的判断标准。首先根据语义安全目标确定策略来生成测试用例，随后运行这些测试用例并判断是否存在异常行为，以确定语义安全目标是否达成，异常判断结果可用于指导下一次测试用例生成，如图6所示。本节将基于这一研究顺序，系统性分析现有研究方法，初步探讨无人机语义安全检测面临的挑战及可能的解决思路。表3总结了现有无人机语义安全检测的相关工作。

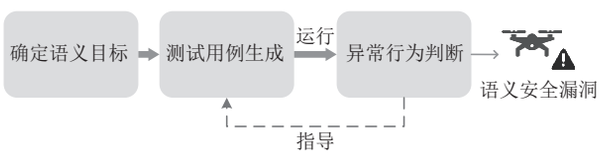


图6 安全检测关键环节

表3 无人机语义安全检测现有工作

测试框架	语义安全目标	测试用例生成	错误判断标准	测试平台		
				被测对象	模拟器	地面站（干预）
RVFUZZER[16]	输入验证错误	基于反馈	非形式化	文献[44,45]	文献[47,49]	文献[51,53,54]
CPI[30]	信息物理不一致	基于反馈	非形式化	文献[44,45]	文献[47]	文献[51]
PGFUZZ[17]	安全策略违反	基于反馈	形式化	文献[44,45,46]	文献[47,49,50]	文献[51,52]
AITester[57]	系统功能故障	AI生成	形式化	文献[44]	文献[49]	文献[51]
LGDFUZZER[33]	范围规范错误	AI生成	非形式化	文献[44]	文献[48]	文献[51]
SWARMFLAWFINDER[58]	群任务失败	基于反馈	非形式化	文献[59]	文献[32]	文献[32]
SwarmFuzz[18]	群体传播漏洞	基于反馈	非形式化	文献[55]	文献[41]	文献[41]

3.1 语义目标确定

在无人机系统中，由于涉及控制程序深层次的逻辑问题，语义安全问题往往不易察觉。当前已知的无人机语义安全漏洞几乎都是在大规模测试或实际飞行过程中暴露出来的，这些特定的语义漏洞通常出现在极端情况或边缘场景中。通过深入分析这些已显现的安全漏洞的本质共性，可以提取并定义明确的语义安全目标，从而有针对性地开展系统化测试。

现有工作中的语义安全目标及其对应的语义层次如图7所示。其中，RVFUZZER提出的输入验证错误包括范围规范错误和范围实现错误两类。

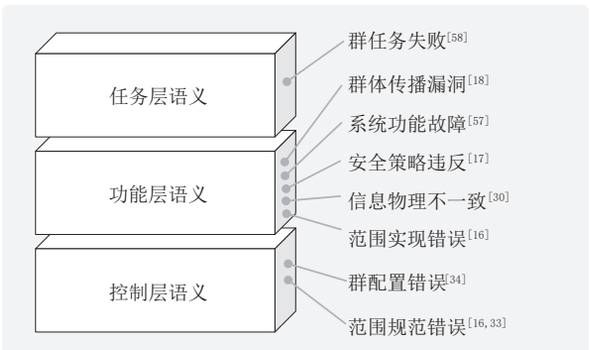


图7 现有工作语义目标对应的语义层次

（1）控制层语义目标

RVFUZZER和LGDFUZZER关注无人机控制程序中的参数范围规范错误。开发人员在指定参数范围时，允许了导致无人机控制错误的危险值存在。这些漏洞被攻击者利用只需构造并发出一个看似无害但实际上恶意的参数变更命令并发给目标无人机，使无人机姿态控制功能发生错误，进而导致无人机不稳定运动甚至崩溃。SWARMBUG[34]关注群配置错误，该错误由群控算法或无人机参数（即配置变量）的不合理设置导致。

（2）功能层语义目标

RVFUZZER还关注控制程序中的范围实现错

误，即参数范围检查不正确或缺失导致控制程序接受了规定范围外的控制参数值设置，这种错误属于安全约束功能未能有效实施的错误类型。CPI一文提出了名为信息物理不一致（Cyber-Physical Inconsistency，CPI）的漏洞类型，这类漏洞主要针对无人机的安全检查机制，包括碰撞检测和推力损失检测等。信息物理不一致性漏洞被利用时，可能会出现两种情况：一是真实物理事故未能被安全检查机制报告，即存在漏报风险，二是无人机正常运行时，安全检查机制却产生虚假警报，即存在误报风险。该错误属于安全检查机制的故障保护功能存在的异常。攻击者利用这些漏洞仅需操纵环境条件，例如在无人机飞行轨迹上放置具有特定重量和角度的障碍物。这类问题的根源在于现有的安全检查大多依赖于通用编程语言实现的简单范围检查，这些检查不足以描述复杂和微妙的物理世界。事实上，除了安全检查机制，信息物理不一致还可能存在于其他任何功能和算法中。此外，无人机中存在类似于自动驾驶交通规则[60]的安全策略，例如“无人机高度必须高于100 m才能打开降落伞”或“无人机降落时必须在10 m内以1 m/s的速度下降”，这些条件可以被视为无人机的安全约束功能。PGFUZZ的目标正是检测出无人机违反这类安全策略的场景，此工作从被测飞行控制系统的官方文档和源代码注释中提取了这些安全策略。AITester[57]的语义目标是系统级的功能实现错误，重点关注无人机整体行为是否满足其预定的行为约束。这类语义目标在范围上包含了上述范围实现错误、信息物理不一致和违反安全策略三类语义目标，但更侧重于从高层视角评估整个系统的正确性，而不过度关注各个内部组件或子功能的具体细节。SwarmFuzz的关注的是无人机群中存在的群体传播漏洞（Swarm Propagation Vulnerabilities，SPVs），当无人机群中的单个无人机发生轻微偏差时，可能间接导致其他群成员偏离它们的路线，引发与障碍物相撞等灾难性后果，该错误涉及多无人机编队控制功能错误。SwarmFuzz在漏洞利用时，通过向蜂群中的单个无人机实施GPS欺骗干扰，导致群控算法产生错误的控制命令，从而引发群体行为异常。

（3）任务层语义目标

SWARMFLAWFINDER是现有的关注任务层语义安全问题的典型工作。具体而言，该工作关注分布式无人机群控算法的逻辑缺陷，这些缺陷可能导致各种群任务失败，例如军事任务中的搜索或运送失败。这些逻辑缺陷可以通过引入攻击无人机以及物理障碍物等方式加以利用。

3.2　测试用例生成

测试用例生成阶段面临的主要挑战在于输入空间的庞大、复杂与高度动态性。无人机控制程序涉及的可配置参数数量众多，可能高达数百个，每一个参数的选择都可能显著影响系统的运行。此外，环境输入因素如气候条件、地理特征以及与其他物体的交互情况等，都是高度动态和复杂的，这使得无人机的实际状态呈现出丰富的变化。各种输入组合可能性几乎是无穷的，这使得全面地探索所有潜在问题变得极其困难。因此，采用科学的输入选择策略和测试用例生成方法，以节省测试和验证所需的时间、成本和资源，提高测试效率和覆盖率，显得至关重要。

在制定测试用例生成策略之前，需要先确定输入空间。输入空间根据具体问题确定，可以通过一些缩减策略来优化，如排除与特定问题无关的参数，或按照影响力对输入组合进行优先级排序等。

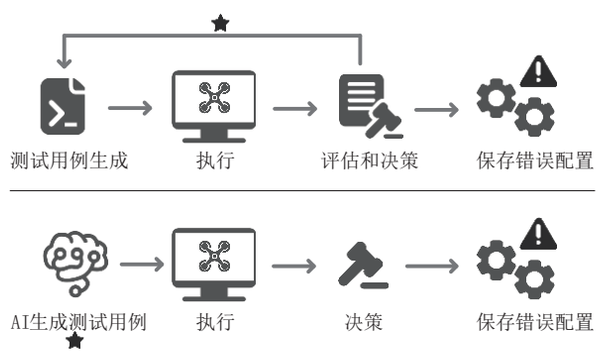


图8　基于反馈（上）和基于AI（下）的测试框架对比

自动化测试要求自动生成测试用例，现有测试用例生成方法主要有两种：基于反馈引导的测试用例生成和基于AI的测试用例生成，图8是运用这两种方法的测试框架对比。基于反馈的测试方法在每次执行测试用例后，对结果进行评估和决策，保存错误配置并利用评估结果引导下一次测试用例的生成；基于AI的测试方法则由AI模型直接生成可能导致错误的测试用例，并在模拟器执行后进行决策，保存错误配置。对于生成的测试用例的有效性和充分性，前者侧重于设置有效的“变异-反馈”策略，而后者侧重于训练AI模型。

现有工作的输入空间选择和自动化测试用例生成方法如下。

（1）输入空间选择方法

RVFUZZER和LGDFUZZER的输入空间仅包

括控制参数，由控制程序的动态可调整参数列表、每个参数的所有可能值范围以及每个参数的默认值定义。CPI的输入空间包括物理世界天气和障碍物，根据安全检查程序的特点和漏洞检测目标（漏报或误报），人为设置初始测试用例。PGFUZZ考虑了较为完整的输入空间，包括配置参数、用户命令和环境因素，利用程序分析技术，建立“策略-输入项”映射表，消除与目标安全策略无关的输入，由此缩减庞大的输入空间。SWARMFLAWFINDER引入了一台攻击者无人机，从外部对目标无人机群实时干扰，输入空间为该攻击者无人机的初始姿态和四种干扰策略的组合。该输入空间本质上还是物理世界中的障碍物组合。SwarmFuzz的输入空间包括GPS偏差和障碍物坐标。GPS偏差是对单个无人机施加的，该偏差会导致群控算法产生错误的控制命令。因此，在蜂群层面上，该输入空间可以抽象为群控命令与外部障碍物的组合。

（2）测试用例自动生成方法

①反馈引导输入变异

RVFUZZER的“变异-反馈”策略基于以下观点：当控制参数值增加（减少）并开始出现控制不稳定性时，进一步增加（减少）参数值只会维持或加剧这种不稳定性。因此，该测试框架采用控制不稳定性检测器的结果作为反馈，通过一种类似二分搜索的方法有效地进行输入变异，有助于快速触发漏洞，并发现错误的参数范围。CPI测试技术定义了两种成本函数：网络成本函数和物理成本函数。网络成本用于衡量安全检查接近于检测到碰撞的程度，物理成本用于量化速度和角度等物理状态的异常程度，通过计算预测值与真实值的差来得到。为了暴露出信息物理不一致漏洞，该工作采用了基于多目标优化的遗传算法（Genetic Algorithm，GA）去突变环境配置，以最小化网络成本并最大化物理成本来暴露误报漏洞，相反，以最大化网络成本并最小化物理成本来暴露漏报漏洞。PGFUZZ基于“策略-输入项”映射表和由MTL公式表示的安全策略，计算命题距离，量化无人机当前状态与策略违规之间的接近程度，从而引导输入变异。SWARMFLAWFINDER用因果贡献度（Degree of Causal Contribution，DCC）抽象群体行为，并通过计算DCC序列之间的归一化交叉相关度（Normalized Cross Correlation，NCC）来衡量群体行为的相似度。基于该相似度反馈引导测试用例变异：测试用例产生新行为时将进行轻微突变，测试用例产生与之前同类行为时将进行显著突变，从而发现多样化漏洞。SwarmFuzz在对选中的无人机对进行测试时，将目标视为一个凸优化问题，其目标函数是带有时间约束的“受害者-障碍物”距离，并利用梯度引导的优化方法来搜索能导致碰撞的GPS欺骗参数。

②AI生成关键输入

AITester[57]采用深度强化学习算法深度Q网络（Deep Q-network，DQN）与长短期记忆（Long Short-Term Memory，LSTM）网络相结合的方法，在测试过程中基于环境上下文动态地生成并执行可能导致无人机行为不满足预期功能的测试场景。LGDFUZZER利用收集到的飞行日志数据进行特征提取并训练一个基于LSTM的状态预测模型，该模型能根据给定的参数配置以及无人机当前的状态准确估算下一时刻无人机应达到的状态。随后，LGDFUZZER采用遗传算法进行搜索，检测可能存在的错误配置。与RVFUZZER相比，LGDFUZZER使用机器学习模型代替了实际运行的飞控程序，能够更高效地生成可能导致错误的测试用例。

3.3 异常行为判断

异常行为判断的核心在于设定“预期结果”，即我们期望系统表现出的行为。通过将实际测试结果与这一预设的期望结果进行比较，从而确定测试用例是否导致异常。

自动化测试中的异常行为判断需要将“预期结果”抽象成数学语言描述。对于具有准确定义和严密逻辑的语义目标，可以采用线性时序逻辑（Linear Temporary Logic，LTL）[61]、计算树逻辑（Computation Tree Logic，CTL）[62]等形式化语言。然而，无人机系统存在许多边界条件模糊且解释依赖于人类理解的语义目标，难以准确描述。例如，无人机姿态控制功能旨在确保飞行稳定，但稳定性本身是个定量标准，其错误研判往往依赖于人工定性；路径规划算法在同一场景下运行多次，可能产生多种飞行轨迹方案，究竟哪种轨迹最优或具有更强鲁棒性难以准确评估；在无人机的运输任务中，如果无人机完全停止，将导致其无法到达目的地，此外，显著降低的飞行速度也可能造成战斗失败或经济损失等不良后果。

当前无人机语义安全检测工作的预期结果选择、抽象以及判定方法如下。其中，大部分采用非形式化语言描述预期结果，通常是在自动化测试发现可能的漏洞后，再手动进行测试和分析。

(1)非形式化描述

RVFUZZER通过观测两种类型的控制状态偏差——观测状态偏差和参考状态偏差，对姿态控制功能进行异常判断。前者指控制器无法根据参考状态稳定其观测状态的情况，利用积分绝对误差(Integral Absolute Error，IAE)公式来量化，当这种偏差持续增加并超过某个预设阈值时，会判断存在观测状态偏差。后者是指无人机偏离既定任务的情况，即控制器未能调整其参考状态以跟踪任务目标，直接使用检查参考状态与任务目标之间差异来量化，并检查该偏差是否持续大于一个阈值。CPI为了判定网络物理不一致错误，分别需要对程序报告结果和真实物理事故进行判定。其中安全检查程序的结果是确切的布尔值。而要判定真实物理事故，通过系统识别[63](System Identification，SI)模型将无人机的动力学特性和控制逻辑建模为一系列方程，使用该模型预测无人机状态，并通过传感器测量真实状态。当预测状态与真实状态之间的差异超过设定的阈值，则判断为发生了真实的事故。LGDFUZZER执行验证时，通过观测飞行冻结和偏差情况来确定无人机是否处飞行不稳定。在执行AVC2013[64]飞行任务过程中，如果无人机在15秒内移动的距离总是小于0.5米，则被视为飞行冻结；如果无人机的飞行偏差持续高于1.5米的15倍，则被认为是严重偏差。在SWARMFLAWFINDER中，对于任务失败的判定标准与任务目标有关。例如，Adaptive Swarm运输任务[59]的目标是投送一个物体，需要四架无人机合作，每架无人机都附有一根绳子来固定物体，正常情况下该任务完成需要189.4(±5.8)个单位时间。若该无人机群未能在两倍正常时间，也就是400个单位时间内到达目的地，则判定该群任务失败。SwarmFuzz仅考虑受害者无人机与障碍物的碰撞，以此判断目标群功能错误。

(2)形式化描述

PGFUZZ利用带有时间约束的时态逻辑(Metric Temporal Logic，MTL)公式来抽象表示安全策略，并在运行时检查这些策略，通过计算全局距离来判断策略的违反。AITester[57]使用对象约束语言(Object Constraint Language，OCL)表示被测系统的预期行为。违反OCL约束被视为系统预期功能的偏离，即系统规范的违反。

3.4 小 结

无人机语义安全检测的目标是系统地识别和挖掘无人机系统在语义层面的异常行为。现有方法依赖自动化测试技术，通过设计并执行大量测试用例，来发现无人机系统中的语义安全漏洞。在语义目标确定环节，目前主要依赖人工分析，通常需要在大规模测试中识别具有本质共性特征的角落问题。在测试用例生成环节，目前主要有两种方法：基于反馈的生成方法和基于AI的生成方法。基于反馈的方法根据已有测试结果和反馈生成用例，能够有效覆盖已知问题区域，易于理解调试，但可能忽略新类型的问题。基于AI的方法具有较强的自适应性，能够自动学习和适应新的测试需求，生成更多样化的用例，有助于发现潜在的边缘案例，并能在短时间内大量生成，从而加快测试过程。然而，这种方法的初期成本高，建立有效的AI模型需要大量的训练数据和计算资源，并且模型解释性差，这给调试工作带来困难。因此，基于反馈的生成方法更适用于有测试历史的项目，特别是在针对特定问题进行深入测试时；而基于AI的生成方法则更适合需要大规模数据处理和需求快速变化的场景，如开发阶段的探索性测试。异常行为判断阶段的核心在于设定“预期结果”。对于明确的语义目标，如PGFUZZ[17]中的“违反安全策略”，可以使用形式化语言来定义；而对于模糊的目标，如“飞行不稳定”[16]，则高度依赖人工设定。具体采用哪种方法，需要根据具体的语义安全目标进行确定。

4 无人机语义漏洞修复

无人机语义安全问题的根源通常在于控制程序中的逻辑缺陷，例如开发者对无人机预期功能语义理解不足而导致的错误编程决策。因此，对于确切已知的无人机语义安全漏洞，通常可以通过更新或修改控制程序源代码来进行修复。这种修复过程通常在无人机离线时进行，通过应用新代码来提供一种相对永久性的解决方案，从而防止因该特定漏洞导致的安全问题再次发生。

无人机语义漏洞修复(图5(b))包括三个关键环节：环节一为错误归因和定位；环节二为修复策略执行；环节三为修复正确性验证。首先对发现语义安全漏洞进行归因分析和错误定位，然后根据分析结果设定并执行相应的修复策略，最后验证修复的完整性和可用性，如图9所示。本节将基于这一研究顺序，系统性分析现有研究方法，初步探讨无人机语义漏洞修复面临的挑战及可能的解决思路。表4总结了现有语义漏洞修复的相关工作。

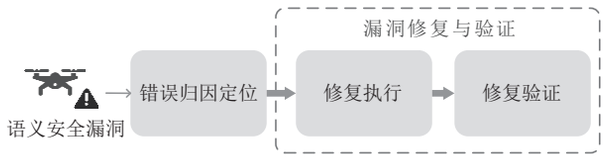


图9 漏洞修复关键环节

4.1 错误归因定位

在软件修复工作中，错误归因定位是基础且关键的环节。归因涉及确定导致错误发生的根本原因，而定位是找出错误在程序源代码或执行流程中的具体位置，只有准确地识别问题所在，才能有针对性地进行修复，从而解决存在的问题。然而，无人机系统及其行为语义的复杂性给错误归因定位工作带来挑战。

错误归因方法主要有两种：基于日志调查的归因与基于重放分析的归因，如图10所示。基于日志调查的归因方法依赖于详尽的日志记录，通过对日志数据的调查分析来查找错误的根本原因。该方法的关键在于设置有效的日志收集策略和控制程序插桩方法。基于重放分析的归因方法则通过重现错误发生的场景，在模拟环境中重复那些导致错误的配置来识别具体的错误原因。此方法的侧重点在于制定重放分析策略，如基于反事实因果的重放分析[34-35]。这两种归因方法各有优势，可以结合使用。语义安全错误定位通常采用程序分析方法[19,32]。

现有工作的错误归因和定位方法如下。

(1)基于日志调查的错误归因方法

MAYDAY使用控制变量依赖图(Control Variable Dependency Graph，CVDG)来捕获无人机控制系统中控制变量间的依赖关系，通过将CVDG映射到控制程序，实现针对控制程序的选择性插桩，进而在飞行过程中生成控制级和程序级的日志。在事故发生后，MAYDAY执行两个阶段的调查：控制级调查，分析控制级日志以识别出在事故期间所有原语控制器中首个发生故障的控制器，并推断导致该控制器失效的控制变量污染路径；程序级调查，利用第一步的结果以及CVDG模型与程序间的映射来缩小需要分析的程序级日志范围，进而精确地定位到引发事故的根本原因所在的基本块，从而实现控制器到程序的追溯。RVPLAYER提出了一种需求驱动的自适应日志记录方法，对控制程序插桩并记录三部分运行时信息：自适应记录环境扰动，在车辆行为偏离动力学模型预测时采用高采样频率，其他情况下则降低采样频率；以较低但固定的频率记录状态信息，包括位置、姿态和传感器测量值；完整记录离散事件，例如参数更新。这种方法能够以低开销记录关键信息，对于因更新错误参数并在一定条件下触发的漏洞的诊断非常有效。

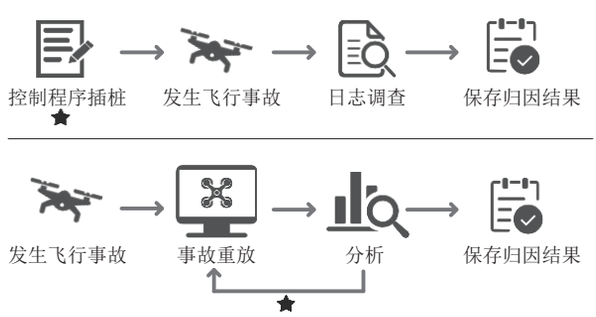


图10 基于日志(上)和基于重放(下)的归因方法对比

(2)基于重放的错误归因方法

SWARMBUG对特定的群配置错误进行基于反事实因果的重放分析，在重放过程中依次删除每个环境变量的影响，并通过记录无人机姿态的偏差来计算得到因果贡献度(Degree of Causal Contribution，DCC)，这一概念抽象了环境配置对集群中无人机行为决策的影响程度。RVPLAYER基于因果推理对事故进行重放分析，在重放过程中使用遗传算法系统地禁用离散事件并更改参数，来寻找能够避免事故发生的最小变化集，通过对这些关键变化的分析追溯事故根本原因。

(3)基于程序分析的错误定位方法

PGPATCH通过名称匹配的方法将每个术语与源代码中的特定变量/函数进行映射，以确定补丁的位置。同时，PGPATCH还创建了一个访问模式映射表，显示了变量或函数在源代码中的位置及其

表4 无人机语义漏洞修复现有工作

现有工作	语义安全漏洞	错误归因定位	修复执行	修复验证
MAYDAY[32]	控制语义错误	程序分析＋日志调查	无	无
SWARMBUG[34]	群配置错误	重放分析	参数调优	模糊测试
RVPLAYER[35]	控制语义错误	日志收集＋重放分析	无	无
PGPATCH[19]	安全策略违反	程序分析	自动补丁	模糊测试
PATCHVERIF[20]	逻辑错误/语义错误	无	无	模糊测试

访问方式，以便于生成补丁。

4.2 漏洞修复验证

当前，无人机语义错误归因与定位技术存在局限性，难以精准地识别错误的根本原因和具体位置，通常只是缩小了错误根源范围。因此，在漏洞修复工作中，传统软件的漏洞修复方法无法直接应用，而是需要根据已知的有限信息，制定适合的修复策略，通过更新或修改源代码等方式执行修复。执行具体修复后，必须进行修复验证，以确保修复工作的完整性以及修复后控制程序的可用性。验证工作可以通过自动化测试的方法来检查被修复漏洞是否仍存在，以及是否引入了新漏洞。此外，还需通过对比修复前后的系统行为差异，评估修复是否对原有系统的功能和性能产生不利影响。无人机语义安全漏洞的修复和验证环节对于保证无人机系统的安全性与可靠性至关重要。

现有工作的修复执行和验证方法如下。

(1)修复执行策略

SWARMBUG基于DCC趋势，通过调节关键参数，应用强化、倒置、破坏平衡和建立平衡四种策略，获得候选修复集，后续再进一步验证候选修复的正确性。PGPATCH为给定的逻辑错误生成补丁。它利用自动程序修复技术，通过接收现有或新的用于发现逻辑错误的逻辑公式作为输入，对先前已知的逻辑错误进行分类并确定其所属的补丁类型。接着，根据识别出的补丁类型、违反的公式以及补丁位置，PGPATCH使用定制算法生成源代码补丁作为输出。然而，该工作具有局限性，只能自动生成简单补丁，对于需要重头实现新功能、定义新变量或添加数学公式等复杂技术来修复的错误，PGPATCH无法自动创建修补程序。此外，虽然MAYDAY和RVPLAYER两个工作并未提及具体的修复执行方法，但MAYDAY已经将错误定位到约几十行代码的基本块中，调查人员能够轻松检查这些基本块的源代码并进行手动修复，而RVPLAYER在错误归因定位阶段找出了能抑制事故发生且对初始配置改变最小的变化集，调查人员可以利用该最小变化集进行手动修复。

(2)修复验证方法

SWARMBUG对调整参数后的群算法执行反馈引导的模糊测试，在飞行区域内生成其他机器人和障碍物，测试期间未发生碰撞或超时则认为修复有效。当存在多个有效修复时，SWARMBUG会测量这些修复与原始执行的DCC之间的均方误差(Mean Squared Error，MSE)。具有更小MSE值的修复被认为是更好的修复方案，因为它与原始行为更接近，更能保持群体的原始行为特征。PGPATCH通过在不同的测试场景下运行应用补丁后的控制程序并运行PGFUZZ来检查补丁是否成功修复了给定的逻辑错误以及确保该错误不会在其他任务和环境条件下重新出现。在验证过程中，PGPATCH会利用PGFUZZ的分析引擎自动查找与给定公式相关联的状态，如果检测到任何未预期的状态变化，且这些变化未包含在PGFUZZ找出的相关状态之中，则认为该补丁可能影响程序的功能或者降低其性能。PATCHVERIF的目标是评估给定补丁是否会在无人机控制软件中引入错误，结合了静态和动态分析方法来度量分析的补丁对无人机物理状态的影响。具体来说，PATCHVERIF首先识别出与所分析补丁相关的无人机输入，包括用户命令、配置参数和环境因素。然后，通过一种输入突变算法对这些识别出的输入进行变异，使原始代码与修复后的代码在物理空间中的行为差异最大化，从而触发可能由错误补丁引起的错误行为。为了判断这些错误行为，PATCHVERIF确定了控制软件中的五个物理关键特性作为物理不变量，并使用支持向量机(Support Vector Machines，SVMs)判断错误行为是否由该补丁引起，从而推断补丁是否为故障补丁。

4.3 小 结

无人机语义漏洞修复的目标是针对明确的语义安全问题，实施有效的修复策略，以确保无人机系统遵循正确的语义规则。在错误归因和定位环节，目前主要采用日志调查和重放分析两种方法。日志调查要求在事故发生前确定测量指标并进行预先插桩，这种方法能够提供准确且实时的数据，但在飞行过程中会额外消耗存储空间。此外，由于日志记录的数据可能不全面，往往难以还原复杂故障场景，因此，仅凭日志调查难以准确找出根本原因。重放分析方法则是在事故发生后，通过重现飞行过程及环境变量变化来分析错误原因。此方法适合深层次的故障分析，但需要较高的计算资源来模拟飞行过程。由于无人机飞行存在不确定性，此方法无法完全还原事故情境，也不能立即提供结果。基于日志调查的方法适用于需要快速响应的初步故障排查，能够迅速定位简单故障场景中的错误原因；而基于重放分析的方法更适合处理复杂故障或难以直接通过日志定位的问题，尤其适合用于研发阶段的详尽

测试以及大型事故后的深入分析。目前已有研究[35]将这两种方法结合使用，但它们之间的关联度仍不高。如何更有效地将二者有机结合，以最大化发挥其优势，是一个值得深入探讨的问题。当前的语义错误定位技术由于系统复杂性，其精确度仍然有限，通常需人工进一步分析，尚未实现真正的“自动程序修复”。在修复执行环节，大多数工作依赖人工干预。尽管PGPATCH[19]开发了自动识别补丁的技术，但其有效处理的漏洞类型仅限于简单情况。至于修复验证环节，目前主要依赖模糊测试，这种方法虽然直观，但效率较低，因此探索更有效的验证方法成为一个值得关注的问题。

5 无人机安全实时防护

无人机在飞行过程中面临着不可预测的飞行状态变化和物理环境变化，其安全检测（图5(a)）和漏洞修复（图5(b)）技术存在局限性，无法保证发现和规避所有安全问题。在此情况下，安全实时防护机制成为确保无人机飞行安全的最后一道防线。该机制通过实时检测无人机的运行状态，及时检测异常行为，并采取应对措施，如记录事件日志、动态调整飞行参数[43]以及执行强制行为[21]等。

与漏洞修复不同，无人机安全防护并非针对已知的安全漏洞，而是旨在实时检测和应对未知威胁和异常。安全防御策略的设计基于无人机的攻击面，而非某一类特定的安全问题。

无人机语义安全实时防护（图5(c)）包括两个关键环节：环节一为运行时监控；环节二为从错误实时恢复。安全实时防护策略主要分为基于偏差的方法和基于形式化的方法两类。在实施过程中，防御系统在无人机启动前预先部署，以便在无人机运行期间进行持续监控，及时发现异常并采取相应的实时恢复措施，如图11所示。基于偏差的防护方法监控相关状态偏差程度，当检测到偏差超过特定阈值时，采取如接管[65-66]等方式进行恢复；而基于形式化的方法则在运行时进行形式化验证并强制执行[21]。本节按照这两个环节顺序，探讨无人机安全实时防护方法。表5总结了现有相关工作。

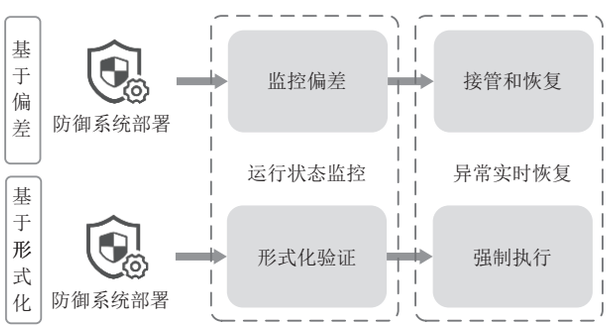


图11 实时防护关键环节

表5 无人机语义安全实时防护相关工作

分类	防御框架	关注的攻击面	运行时监控	实时恢复
基于偏差的方法	CI[67]	物理攻击	状态空间模型	无
	SAVIOR[68]	物理攻击	状态空间模型	无
	SSR[65]	传感器攻击	状态空间模型	接管传感器
	Mini-Me[42]	面向数据的攻击	ML模型	无
	PID-Piper[66]	物理攻击	ML模型	接管控制器
	SCVMON[43]	面向数据的攻击	安全关键变量	恢复异常变量
基于形式化的方法	DLMON[37]	行为隐患和外部攻击	UAS-DL	无
	RMoM[31]	对ROS节点的恶意攻击	HSL-RMoM	无
	DRE[21]	宏观任务与微观行为	DT-MTL	强制执行
	FMMON[38]	无人机故障	LTL	无

注：表中基于偏差的方法不针对语义安全，但语义安全防护能借鉴此类方法。

5.1 运行状态监控

在无人机安全实时防护中，运行时监控可以根据语义目标是否能够被形式化表示采取不同的策略。对于不能形式化表示的语义目标，可以借鉴现有针对传感器攻击及其他异常情况的运行时监控与恢复技术。这些技术通过将预期值与实际感知值之间的偏差与预先设定的阈值进行比较，来检测关键不变量、控制命令输出和传感器测量的完整性，从而识别异常。而对于可以形式化表示的语义目标，则可以通过严格的规约表示，并在系统运行时动态检查这些规约是否被遵循。这两种方法各有侧重：基于偏差的方法主要关注于设定观测指标和阈值；而

基于形式化的方法则侧重于如何使用形式化语言或逻辑语言来表示需要监控的目标，两种方法的根本区别如图12所示。

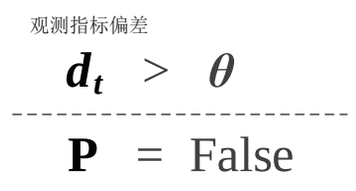


图12　基于偏差和基于形式化方法的根本区别

现有无人机运行时监控相关工作如下。

(1)基于偏差的运行时监控

CI通过派生和监控控制不变量来检测针对无人机的外部物理攻击，包括传感器攻击、控制信号攻击和参数攻击。控制不变量基于无人机的物理属性、其控制算法和物理定律进行联合建模，以状态空间的形式表示，并采用系统识别[63](System Identification，SI)方法从带有未知系数的线性方程中提取无人机的具体系数。将控制不变量的检查代码插入到主控制循环中，在运行时周期性地监测当前系统状态，并利用控制不变量方程独立计算预期状态。如果在设定的监控窗口内，计算状态与实际观测状态之间累积的差异超过了预设阈值，则会被判定为异常并触发报警。SAVIOR与CI的区别在于使用非线性物理不变量构建更精确的状态空间模型。SSR同样建立了一个基于通用系统识别技术的预测状态空间模型。基于该状态空间模型，传感器测量的预测值被用作相应物理传感器的软件备份运行。当检测到物理传感器读数与预期的预测值有显著偏差时，则认为物理传感器可能受到了攻击。Mini-Me是一个数据驱动的在线监测框架，它利用无人机控制程序的内部数据流信息和控制变量依赖关系，训练一个基于神经网络的近似模型作为原控制程序的轻量级副本，以检测面向数据的攻击。在运行时，Mini-Me执行这个最小近似模型，并通过对比估算出的输出与原始控制程序实际计算的输出，来检测是否存在恶意控制状态偏差，以此来监控面向数据的攻击。由于获取并分析了控制流信息，该检测器可以有效地检测到控制语义错误。PID-Piper利用机器学习(Machine Learning，ML)模型设计了一个与无人机主控制器(PID控制器)并行运行的前馈控制器(Feed-Forward Controller，FFC)。它实时监控两者之间的输出偏差，当偏差超过预定义的安全阈值，则判定存在攻击。SCVMON框架采用控制不稳定检测方法识别安全关键变量(Safety-Critical Variables，SCVs)，并对SCVs执行安全临界性和恢复适宜性评估来筛选适合监控的子集mSCVs，通过实时监控这些mSCV，能够有效检测那些试图改变传感器测量值和配置参数的各种面向数据的攻击。

(2)基于形式化方法的运行时验证

DLMON设计了一种名为UAS-DL的语言描述无人机系统的安全监控规约，并利用将运行时验证和贝叶斯网络推断相结合的方法检测控制程序可能出现的安全故障。RMoM提出了一种基于三值语义模型MTL3-RMoM的高级规范语言HSL-RMoM，用于描述ROS上的机器人群体中的各种具有时态和参数化特性的约束和命题。通过一系列规范解析算法，RMoM能够自动生成针对MTL3-RMoM属性的分布式监控器，并在ROS平台上实现对无人机群的运行时验证。这种方法可以有效地为复杂和多层属性的机器人系统自动生成监视器，并且易于定制。DRE基于离散时间度量时态逻辑DT-MTL，设计了一种用形式化规范语言，用于描述和监控无人机群的宏观与微观属性。FMMON是一种利用故障模式数据库和运行时验证技术进行故障检测的方法。它首先通过分析大量历史故障信息来总结故障模式，接着分析并提取无人机在运行过程中的系统级关键安全特性，最后实现了一个监控生成算法和代码检测框架，能够在无人机实际运行过程中实时监测其是否违反了特定的安全属性，从而能够及时准确地发现异常情况。

5.2　异常实时恢复

在无人机飞行过程中，一旦检测到任何异常情况，其内置的防御系统能够迅速响应并采取相应的应对措施，确保无人机能够持续稳定飞行。具体而言，针对两种主要的监控方法——基于偏差的监控与基于形式化的监控，采取了不同的策略来处理异常情况。对于基于偏差的监控方法，通常采用“由外到内”的实时恢复策略，在检测到可能影响飞行的突发情况后，通过及时调整来避免事态恶化。而对于基于形式化的监控方法，则采取更为精确的“由内到外”的强制执行策略，这种方法针对性强，侧重于根据预先定义的形式化规则和模型来指导系统的响应行为，从而确保无人机的行为符合预期的安全标准。现有工作的恢复方法如下。

(1)基于偏差的实时恢复

SSR检测到某个物理传感器受到攻击,对应的软件传感器可以隔离并替换受损的物理传感器,从而从受攻击导致的内部损坏状态中恢复系统,有效避免严重的攻击后果。攻击结束后,系统通过持续监控物理传感器与软件传感器之间的差异来判断是否可以重新切换回物理传感器。PID-Piper在检测到攻击时,会触发恢复机制,将自主导航从PID控制器切换到ML模型的输出,ML模型接管控制器并帮助限制由于攻击引起的行驶轨迹偏差,防止碰撞等不良后果。一旦攻击停止,PID-Piper会将导航控制切换回PID控制器的输出,以确保系统的稳定性和任务的成功执行。SCVMON在检测到攻击时仅恢复异常的mSCV。对于动态mSCV,恢复模式会使用攻击检测前80至100个周期内平均值的预测值来替换其值。对于静态mSCV,恢复模式则将其恢复到更改前的初始值。当预测值接近实际值时,SCVMON会退出恢复模式。通过这种灵活的模式切换机制,SCVMON能够在发生误报或瞬态攻击的情况下确保任务的连续性。

(2)基于形式化方法的强制执行

DRE针对蜂群系统的连续时间状态变化特点,提出了离散时间(Discrete-time,D-time)强制机制。在任务执行期间,当监控器监测到属性违反时,D-time强制机制自动触发强制行为。此外,在首次检测到属性违反的离散时间间隔内,该机制仅执行一次分布式防护措施。这样的设计不仅有助于确保蜂群在危险环境下的全局任务得以有效执行,还能有效减少因频繁而可能干预引发的系统性能问题。

5.3 小 结

无人机安全实时防护的目标是通过设计和应用各种防御机制来消除潜在的安全威胁。现有的防护方法主要分为两类:基于偏差的方法和基于形式化的方法。基于偏差的方法适用于不明确的目标或预防性方案,能够应对多种异常情况,包括未知的攻击类型。这种方法实施成本较低,但依赖于预设阈值,可能导致较高的误报率。此外,对于复杂的语义安全需求,它可能无法提供足够的保障。相对而言,基于形式化的方法则更适用于明确的安全目标。它通过严格的数学模型定义安全要求,提供了较高的可验证性,从而确保更精确的安全保障。然而,该方法需要进行全面的形式化分析,实施成本较高,并且要求具备专业知识进行设计和维护。基于偏差的方法适合应用于应对复杂环境变化和外部干扰,特别是在快速部署的初期阶段;而基于形式化的方法则更适合高安全等级的应用,如军事或关键基础设施的保护。将这两种方法结合,可以在实际应用中为无人机安全提供更全面、有效的保护,增强系统的防御能力。

6 研究方向与展望

语义安全是无人机安全范畴中的关键要素,是确保无人机精准执行功能任务时必须着重考量的问题。本文旨在通过对已有研究工作的梳理和总结,帮助无人机安全领域研究人员了解研究现状,从而在语义安全方面开展具有探索性的创新研究。近年来无人机语义安全在国际上研究热度高涨,如何通过安全检测和安全强化来提高无人机语义安全性已成为热点问题。目前,我们研究团队正针对智能无人机、无人机群开展语义安全防护的探索工作[69]。然而,该领域仍处于起步阶段,存在诸多不足与挑战。表6列出了现有研究面临的挑战和机遇。鉴于当前无人机语义安全研究面临的挑战,我们对未来研究方向进行了总结与展望。

表6 无人机语义安全面临的挑战与机遇

挑战	机遇
无人机应用场景多元化	识别新的语义安全漏洞、开发新的漏洞挖掘方法
无人机智能化发展	智能无人机上的语义安全、人工智能赋能语义安全
无人机蜂群语义复杂性	无人机蜂群语义安全研究
安全评估标准缺乏	构建安全评估体系、语义安全边界划分
不可预知环境下安全威胁	运行时语义安全强化

(1)新场景下的语义安全探索

随着无人机技术的飞速发展,尤其是制造成本的降低,无人机应用场景日益丰富和多元化,如在GPS信号受阻的隧道或山洞等环境中,无人机需要依赖视觉感知[7,70]和无线测距[71]等技术实现定位和避障。在这些新兴场景中,如何识别新的语义安全漏洞,并开发适应这些环境的安全解决方案,是一个需要持续研究的课题。

(2)人工智能语义安全

人工智能的广泛应用提升了无人机执行复杂任务的能力,同时也带来了新的安全隐患。由于人工智能模型的不可解释性,智能无人机上的语义安全

问题尤为突出，错误发生时难以溯源，且容易遭受对抗性攻击[72]。因此，无人机智能语义安全将是未来研究的一个重要方向，不仅要研究智能无人机语义安全，还要探索如何利用人工智能技术加强无人机语义安全。

(3)无人机蜂群语义安全

在可预见的未来，无人机蜂群技术将在军事和其他领域产生变革性的影响，其安全性不容忽视。当前的研究多聚焦于单个无人机，对多无人机系统的语义安全问题研究尚显不足，而多无人机面临的语义问题更为复杂[73-74]，攻击面也更广。因此，无人机蜂群语义安全是一个迫切需要研究的领域。

(4)无人机语义安全评估体系构建

面对无人机任务的复杂化，当前的语义安全研究仍处在初级阶段，缺乏系统性和完整性。建立统一的无人机语义安全评估标准[75]，以及依据具体任务动态划定无人机语义安全边界，将为后续无人机语义安全研究工作的展开提供有力指导，并对无人机系统的研发与应用起到推动作用。

(5)应对无人机语义安全威胁的策略

无人机运行在不可预知的环境中，需要动态保障其安全性以增强语义鲁棒性。例如，自主无人机集群在执行任务时应能灵活调整任务驱动、避障和队形保持等多目标优先级，并在高度动态和不确定环境下也能保持稳定和可靠。此外，在资源受限的无人机上部署防御机制时[21,76]，需要兼顾效能与资源消耗。因此，如何动态优化无人机的目标优先级和参数等以强化语义安全，以及如何根据不同防御情境选择最佳防御策略，是未来研究的重要方向。

7 总 结

随着无人机技术的不断发展，越来越多的研究开始将模糊测试、程序分析等技术应用于无人机语义安全漏洞的挖掘、修复和防护中。这些工作在很大程度上推动了无人机语义安全领域的研究进展，但总体而言，该领域仍处于起步阶段，现有研究成果尚未能全面解决无人机系统中的语义安全问题。本文在充分调研国际顶尖学术会议/期刊上发表的无人机语义安全研究论文的基础上，首次对该领域进行了系统性的综述，初步探讨了语义安全研究面临的挑战及可能的解决思路，并总结了未来研究发展方向，为本领域研究者提供参考与指导。

作者贡献说明 邓余婉祺、王越为共同第一作者。

参 考 文 献

[1] Papaioannidis C, Mademlis I, Pitas I. Autonomous UAV safety by visual human crowd detection using multi-task deep neural networks//Proceedings of the IEEE International Conference on Robotics and Automation. Xi′an, China, 2021: 11074-11080

[2] Rakotonarivo B H, Drougard N, Conversy S, et al. Cleared for safe take-off? improving the usability of mission preparation to mitigate the safety risks of drone operations//Proceedings of the CHI Conference on Human Factors in Computing Systems. Hamburg, Germany, 2023: 1-17

[3] Alami R, Hacid H, Bellone L, et al. SOREO: A system for safe and autonomous drones fleet navigation with reinforcement learning//Proceedings of the 37th AAAI Conference on Artificial Intelligence. Washington , USA, 2023: 16398-16400

[4] Chambers T, Vierhauser M, Agrawal A, et al. HIFuzz: Human interaction fuzzing for small unmanned aerial vehicles//Proceedings of the CHI Conference on Human Factors in Computing Systems. Honolulu, USA, 2024: 1-14

[5] Koo K, Lee W, Cho S, et al. A secure operating system architecture based on Linux against communication offense with root exploit for unmanned aerial vehicles. The Journal of Information Processing Systems, 2020, 16(1): 42-48

[6] Ding A, Chan M, Hass A, et al. Get your cyber-physical tests done! data-driven vulnerability assessment of robotic aerial vehicles//Proceedings of the 53rd Annual IEEE/IFIP International Conference on Dependable Systems and Network. Porto, Portugal, 2023: 67-80

[7] Zhou Ce, Yan Qi-Ben, Shi Yan, et al. DoubleStar: Long-range attack towards depth estimation based obstacle avoidance in autonomous systems//Proceedings of the 31st USENIX Security Symposium. Boston, USA, 2022: 1885-1902

[8] Sathaye H, Strohmeier M, Lenders V, et al. An experimental study of GPS spoofing and takeover attacks on UAVs//Proceedings of the 31st USENIX Security Symposium. Boston, USA, 2022: 3503-3520

[9] Jang J, Cho M, Kim J, et al. Paralyzing drones via EMI signal injection on sensory communication channels//Proceedings of the 30th Annual Network and Distributed System Security Symposium. San Diego, USA, 2023: 1-18

[10] Sun Cong, Zeng Hui Ming, Song Huan-Dong, et al. Machine learning based runtime detection and recovery method against UAV sensor attacks. Journal of Computer Research and Development, 2023, 60(10): 2291-2303 (in Chinese)
(孙聪，曾荟铭，宋焕东等. 基于机器学习的无人机传感器攻击在线检测和恢复方法. 计算机研究与发展，2023，60(10)：2291-2303)

[11] Schiller N, Chlosta M, Schloegel M. Drone security and the mysterious case of DJI's DroneID//Proceedings of the 30th Annual Network and Distributed System Security Symposium. San Diego, USA, 2023: 1-17

[12] Wei Xiao-Min, Li Xing-Hua, Sun Cong, et al. MagDet: UAV GPS spoofing detection based on the geomagnetic field. Chinese Journal of Computers, 2024, 47(4): 877-891 (in Chinese)
(魏晓敏, 李兴华, 孙聪等. MagDet:基于地磁的无人机GPS欺骗检测方法. 计算机学报, 2024, 47(4): 877-891)

[13] Bi Si-Guo, Li Kai, Hu Shu-Yan, et al. Detection and mitigation of position spoofing attacks on cooperative UAV swarm formations. IEEE Transactions on Information Forensics and Security, 2024, 19: 1883-1895

[14] Lion Air: how could a brand new plane crash? https://www.bbc.com/news/world-asia-46014260

[15] Ethiopian Airlines: 'no survivors' on crashed boeing 737. https://www.bbc.com/news/world-africa-47513508

[16] Kim T, Kim C H, Rhee J, et al. RVFuzzer: Finding input validation bugs in robotic vehicles through control-guided testing//Proceedings of the 28th USENIX Security Symposium. Santa Clara, USA, 2019: 425-442

[17] Kim H, Ozmen M O, Bianchi A, et al. PGFuzz: Policy-guided fuzzing for robotic vehicles//Proceedings of the 28th Annual Network and Distributed System Security Symposium. Virtual , 2021: 1-18

[18] Yao Y E, Dash P, Pattabiraman K. SwarmFuzz: Discovering GPS spoofing attacks in drone swarms//Proceedings of the 53rd Annual IEEE/IFIP International Conference on Dependable Systems and Network. Porto, Portugal, 2023: 366-375

[19] Kim H, Ozmen M O, Celik Z, et al. PGPatch: Policy-guided logic bug patching for robotic vehicles//Proceedings of the 43rd IEEE Symposium on Security and Privacy. San Francisco, USA, 2022: 1826-1844

[20] Kim H, Ozmen M O, Celik Z, et al. PatchVerif: Discovering faulty patches in robotic vehicles//Proceedings of the 32nd USENIX Security Symposium, Anaheim. USA, 2023: 3011-3028

[21] Hu Chi, Dong Wei, Yang Yong-Hui, et al. Decentralized runtime enforcement for robotic swarms. Frontiers of Information Technology & Electronic Engineering, 2020, 21(11): 1591-1606

[22] He Dao-Jing, Du Xiao, Qiao Yin-Rong, et al. A survey on cyber security of unmanned aerial vehicles. Chinese Journal of Computers, 2019, 42(5): 1076-1094 (in Chinese)
(何道敬, 杜晓, 乔银荣等. 无人机信息安全研究综述. 计算机学报, 2019, 42(5): 1076-1094)

[23] Wang Zhao-Xuan, Li Yang, Wu Shi-Hao, et al. A survey on cybersecurity attacks and defenses for unmanned aerial systems. Journal of Systems Architecture, 2023, 138: 102870

[24] Pan Quan, Guo Ya-Ning, YangLyu, et al. Autonomous safety and security of UAV systems: definition, modeling, and gradation. SCIENTIA SINICA Informationis, 2023, 53(08): 1608-1628 (in Chinese)
(潘泉, 郭亚宁, 吕洋等. 无人机系统自主安全:定义、建模与分级. 中国科学:信息科学, 2023, 53(08): 1608-1628)

[25] Wang Ding-Hua, Li Shu-Qing, Xiao Guan-Ping, et al. An exploratory study of autopilot software bugs in unmanned aerial vehicle//Proceedings of the 29th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering. Athens, Greece, 2021: 20-31

[26] Guo Lei, Yu Xiang, Zhang Xiao, et al. Safety control system technologies for UAVs: Review and prospect. SCIENTIA SINICA Informationis, 2020, 50(02): 184-194 (in Chinese)
(郭雷, 余翔, 张霄等. 无人机安全控制系统技术:进展与展望. 中国科学:信息科学, 2020, 50(02): 184-194)

[27] Simon M, Ren A Z, Piqué, et al. FlowDrone: Wind estimation and gust rejection on UAVs using fast-response hot-wire flow sensors//Proceedings of the IEEE International Conference on Robotics and Automation. London, UK, 2023: 5393-5399

[28] Chen Liang-Ming, Xiao Jia-Ping, Zheng Yu-Min, et al. Design, modeling, and control of a coaxial drone. IEEE Transactions on Robotics, 2024, 40: 1650-1663

[29] Chen Mou, Ma Hao-Xiang, Yong Ke-Nan, et al. Safety flight control of UAV: A survey. Robot, 2023, 45(03): 345-366 (in Chinese)
(陈谋, 马浩翔, 雍可南等. 无人机安全飞行控制综述. 机器人, 2023, 45(03): 345-366)

[30] Choi H, Kate S, Aafer Y, et al. Cyber-physical inconsistency vulnerability identification for safety checks in robotic vehicles//Proceedings of the ACM SIGSAC Conference on Computer and Communications Security. Virtual, USA, 2020: 263-278

[31] Hu Chi, Dong Wei, Yang Yong-Hui, et al. Runtime verification on hierarchical properties of ROS-based robot swarms. IEEE Transactions on Reliability, 2020, 69(2): 674-689

[32] Kim T, Kim C H, Ozen A, et al. From control model to program: investigating robotic aerial vehicle accidents with MAYDAY//Proceedings of the 29th USENIX Security Symposium. Boston, USA, 2020: 913-930

[33] Han Rui-Dong, Yang Chao, Ma Si-Qi, et al. Control parameters considered harmful: Detecting range specification bugs in drone configuration modules via learning-guided search//Proceedings of the 44th IEEE/ACM International Conference on Software Engineering. Pittsburgh, USA, 2022: 462-473

[34] Jung C, Ahad A, Jung J, et al. Swarmbug: Debugging configuration bugs in swarm robotics//Proceedings of the 29th ACM Joint European Software Engineering Conference and Symposium on the Foundations of Software Engineering. Athens, Greece, 2021: 868-880

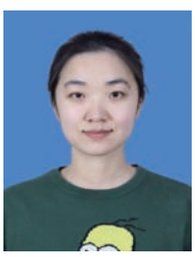
[35] Choi H, Cheng Zhi-Yuan, Zhang Xiang-Yu. RVPLAYER: Robotic vehicle forensics by replay with what-if reasoning//Proceedings of the 29th Annual Network and Distributed System Security Symposium. San Diego, USA, 2022: 1-18

[36] Zhou Xin, Zhu Jiang-Chao, Zhou Hong-Yu, et al. EGO-Swarm: A fully autonomous and decentralized quadrotor swarm system in cluttered environments//Proceedings of the IEEE International Conference on Robotics and Automation. Xi'an, China, 2021: 4101-4107

[37] Yang Dong, Shi Hao, Dong Wei, et al. Security and safety threat detection method for unmanned aerial system based on runtime verification. Journal of Software, 2018, 29(5): 1360-

1378 (in Chinese)
(杨栋，史浩，董威等. 基于运行时验证的无人飞行系统安全威胁检测方法. 软件学报，2018，29(5)：1360-1378)
[38] Hou De-Fei, Su Qing-Ran, Song Yi, et al. Research on drone fault detection based on failure mode databases. Drones, 2023, 7(8): 486
[39] Zhou Xin, Wang Zhe-Pei, Ye Hong-Kai, et al. EGO-Planner: An ESDF-free gradient-based local planner for quadrotors. IEEE Robotics Autom, 2021, 6(1): 478-485
[40] Quigley M, Conley K, Gerkey B P, et al. ROS: An open-source robot operating system//Proceedings of the ICRA Workshop on Open Source Software. Kobe, Japan, 2009, 3(3.2): 5
[41] Soria E, Schiano F, Floreano D. SwarmLab: A Matlab drone swarm simulator//Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems. Las Vegas, USA, 2020: 8005-8011
[42] Ding A, Murthy P, Garcia L, et al. Mini-Me, you complete me! data-driven drone security via DNN-based approximate computing//Proceedings of the 24th International Symposium on Research in Attacks, Intrusions and Defenses. San Sebastian, Spain, 2021: 428-441
[43] Park S, Kim Y, Lee D H. SCVMON: Data-oriented attack recovery for RVs based on safety-critical variable monitoring//Proceedings of the 26th International Symposium on Research in Attacks, Intrusions and Defenses. Hong Kong, China, 2023: 547-563
[44] ArduPilot. https://ardupilot.org/
[45] PX4. https://px4.io/
[46] Hattenberger G, Bronz M, Gorraz M. Using the paparazzi UAV system for scientific research//Proceedings of the International Micro Air Vehicle Conference and Competition, Delft, Netherlands, 2014: 247-252
[47] Koenig N, Howard A. Design and use paradigms for gazebo, an open-source multi-robot simulator//Proceedings of the 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems. Sendai, Japan, 2004: 2149-2154
[48] Shah S, Dey D, Lovett C, et al. AirSim: High-fidelity visual and physical simulation for autonomous vehicles//Proceedings of the 11th Field and Service Robotics. Zurich, Switzerland, 2017: 621-635
[49] SITL simulator (software in the loop). https://ardupilot.org/dev/docs/sitl-simulator-software-in-the-loop.html
[50] Berndt J. JSBSim: An open source flight dynamics model in C++. AIAA Modeling and Simulation Technologies Conference and Exhibit, 2004: 4923
[51] Pymavlink. https://github.com/ArduPilot/pymavlink
[52] PPRZLINK. https://github.com/paparazzi/pprzlink
[53] MAVProxy. https://ardupilot.org/mavproxy/
[54] QGroundControl. https://qgroundcontrol.com/
[55] Vásárhelyi G, Virágh C, Somorjai G, et al. Optimized flocking of autonomous drones in confined environments. Science Robotics, 2018, 3(20): eaat3536
[56] Lattner C, Adve V. LLVM: A compilation framework for lifelong program analysis & transformation//Proceedings of the 2nd IEEE/ACM International Symposium on Code Generation and Optimization. San Jose, USA, 2004: 75-88
[57] Sartaj H. Automated approach for system-level testing of unmanned aerial systems//Proceedings of the 36th IEEE/ACM International Conference on Automated Software Engineering, Melbourne. Australia, 2021: 1069-1073
[58] Jung C, Ahad A, Jeon Y, et al. SWARMFLAWFINDER: Discovering and exploiting logic flaws of swarm algorithms//Proceedings of the 43rd IEEE Symposium on Security and Privacy. San Francisco, USA, 2022: 1808-1825
[59] Timurovich A R. Adaptive control of swarm of drones for obstacle avoidance [Master Thesis]. Skolkovo Institute of Science and Technology, Moscow, 2019
[60] Sun Yang, Poskitt C M, Sun Jun, et al. LawBreaker: An approach for specifying traffic laws and fuzzing autonomous vehicles//Proceedings of the 37th IEEE/ACM International Conference on Automated Software Engineering. Rochester, USA, 2022: 62:1-62:12
[61] Pnueli A. The temporal logic of programs//Proceedings of the 18th Annual Symposium on Foundations of Computer Science. Providence, USA, 1977: 46-57
[62] Clarke E M, Emerson E A. Design and synthesis of synchronization skeletons using branching-time temporal logic. Logic of Programs, 1981: 52-71
[63] Giraldo J, Urbina D, Cárdenas A, et al. A survey of physics-based attack detection in cyber-physical systems. ACM Computing Surveys, 2018, 51(4): 76:1-76:36
[64] SparkFun autonomous vehicle competition. https://avc.sparkfun.com/2013
[65] Choi H, Kate S, Aafer Y, et al. Software-based realtime recovery from sensor attacks on robotic vehicles//Proceedings of the 23rd International Symposium on Research in Attacks, Intrusions and Defenses. San Sebastian, Spain, 2020: 349-364
[66] Dash P, Li Guan-Peng, Chen Zi-Tao, et al. PID-Piper: Recovering robotic vehicles from physical attacks//Proceedings of the 51st Annual IEEE/IFIP International Conference on Dependable Systems and Networks. Taipei, China, 2021: 26-38
[67] Choi H, Lee W, Aafer Y, et al. Detecting attacks against robotic vehicles: a control invariant approach//Proceedings of the ACM SIGSAC Conference on Computer and Communications Security. Toronto, Canada, 2018: 801-816
[68] Quinonez R, Giraldo J, Salazar L, et al. SAVIOR: Securing autonomous vehicles with robust physical invariants//Proceedings of the 29th USENIX Security Symposium. Boston, USA, 2020: 895-912
[69] Wang Yue, Yang Chao, Zhang Xiao-Dong, et al. DPFuzzer: Discovering safety critical vulnerabilities for drone path planners//Proceedings of the 47th International Conference on Software Engineering. Ottawa, Canada, 2025: 1-13
[70] Wei Zhe, Li Cong-Li, Shen Yan-An, et al. Thick cloud region content generation of UAV image based on two-stage model. Chinese Journal of Computers, 2021, 44(11): 2233-2247 (in Chinese)
(韦哲，李从利，沈延安等. 基于两阶段模型的无人机图像厚云区

域内容生成. 计算机学报，2021，44(11)：2233-2247)
[71] Luo Zhong-Jie. Research on the cooperative positioning method of UAV based on the fusion of IMU and wireless ranging and direction finding [Master Thesis]. Beijing University of Posts and Telecommunications , Beijing, 2022 (in Chinese)
(罗中杰. 基于IMU与无线测距测向融合的无人机协同定位方法研究[硕士学位论文]. 北京邮电大学，北京，2022)
[72] Ma Chen, Shen Chao, Lin Chen-Hao, et al. Attacks and defenses for autonomous driving intelligence models. Chinese Journal of Computers, 2024,47(6): 1431-1452 (in Chinese)
(马晨，沈超，蔺琛皓等. 针对自动驾驶智能模型的攻击与防御. 计算机学报，2024,47(6)：1431-1452)
[73] Wang Feng, Zhang Heng, Han Meng-Chen, et al. Co-evolution based mixed-variable multi-objective particle swarm optimization for UAV cooperative multi-task allocation problem. Chinese Journal of Computers, 2021, 44(10): 1967-1983 (in Chinese)
(王峰，张衡，韩孟臣等. 基于协同进化的混合变量多目标粒子群优化算法求解无人机协同多任务分配问题. 计算机学报，2021，44(10)：1967-1983)
[74] Liao Jian, Cheng Jun, Xin Bin, et al. UAV swarm formation reconfiguration control based on variable-stepsize MPC-APCMPIO algorithm. Science China Information Sciences, 2023, 66(11): 212207
[75] Shen Bo, Wu Wen-Liang, Yang Gang, et al. Evaluation models and methods for intelligence of unmanned swarm systems based on collective OODA loop. Acta Aeronautica et Astronautica Sinica, 2023, 44(14): 263-278 (in Chinese)
(沈博，武文亮，杨刚等. 基于群体OODA的无人集群系统智能评价模型及方法. 航空学报，2023，44(14)：263-278)
[76] Qu Yu-Ben, Qin Zhen, Ma Jing-Hao, et al. Service provisioning for air-ground collaborative mobile edge computing. Chinese Journal of Computers, 2022, 45(4): 781-797 (in Chinese)
(屈毓锛，秦蓁，马靖豪等. 面向空地协同移动边缘计算的服务布置策略. 计算机学报，2022，45(4)：781-797)



DENG Yu-Wan-Qi, Ph. D. candidate. Her current research interests include unmanned systems security, intelligent unmanned swarm security.

WANG Yue, Ph. D. candidate. His current research interests include unmanned systems security, unmanned systems intelligent algorithm security.

YANG Chao, Ph.D., professor. His current research interests include unmanned intelligent systems security, security testing.

FENG Peng-Bin, Ph. D., lecturer. His current research interests include malware detection, vulnerability detection.

YANG Jia-Hui, M.S. candidate. Her current research interest is unmanned systems security.

WANG Li-Juan, Ph. D., lecturer. Her current research interests include machine learning, intelligent computing, autonomous intelligent control of drones.

LI Xing-Hua, Ph.D., professor. His current research interest is wireless network security.

MA Jian-Feng, Ph. D., professor. His current research interests include wireless network security, mobile intelligent systems security.

Background

UAV semantic safety is an important branch of UAV security, which is different from the traditional security of UAVs in that it focuses more on the functional defects of UAVs themselves rather than on external malicious attacks. Currently, the problem of semantic safety for drones has gradually gained attention internationally, especially in top academic conferences and journals in the fields of cybersecurity and software engineering, and a number of works have been focused on this issue.

We have investigated a large number of UAV security related papers published in recent years, and found that there is no review article that considers UAV security from the perspective of "semantic safety". In order to fill this gap, this paper provides a comprehensive review and synthesis of the current research status of UAV semantic safety. First, it defines UAV semantic safety and divides the research branches. Subsequently, the challenges of various research issues are analyzed and research methods are summarized. Finally, the future research directions are summarized and prospected.

This work was supported in part by the National Natural Science Foundation of China (6223000226, 62125205, 62402364), the Technology Innovation Leading Program of Shaanxi (2024QCY-KXJ-171), the Natural Science Basic Research Plan in Shaanxi Province of China (2023-JC-QN-0759), and the Fundamental Research Funds for the Central Universities (ZYTS24138).