中国科学: 信息科学 2023 年 第 53 卷 第 8 期: 1608–1628
SCIENTIA SINICA Informationis
论文

无人机系统自主安全: 定义、建模与分级

潘泉*, 郭亚宁, 吕洋, 李扬, 谈政

西北工业大学自动化学院, 西安 710072
* 通信作者. E-mail: quanpan@nwpu.edu.cn

收稿日期: 2022–08–25; 修回日期: 2022–11–09; 接受日期: 2022–12–02; 网络出版日期: 2023–08–17

国家自然科学基金 (批准号: 62203358, 62233014) 资助项目

摘要 自主与安全平衡是无人机技术与产业健康、快速发展的战略需求和核心关键. 在同时面对 "不断提升的无人自主飞行性能" 与 "日趋严峻的物理、信息、智能等多域安全威胁" 的挑战下, 仅借鉴有人机安全体系, 适度拓展适航飞行和感知规避为核心的安全架构已不能满足现在和未来无人机系统的安全需求, 构建无人机自主安全新架构显得十分必要和迫切. 本文从自主性这一无人机固有的本质属性出发, 初步研究了无人机自主安全架构, 重点定义和表征了 "无人机自主安全" 新概念、构建了面向多域安全威胁和系统高度自主的无人机系统自主安全新模型、提出了无人机系统自主安全能力分级, 为我国无人机行业健康快速发展提供基础支撑, 为其他无人系统安全体系建立提供借鉴.

关键词 无人机系统, 多域威胁, 自主安全, 资源配置

1 引言

近年来, 无人系统技术, 特别是无人机系统 (unmanned aerial system, UAS) 技术发展迅猛, 系统类型多样, 涵盖领域广泛, 应用数量巨大. 其显著特征是无人机与环境、人的关系更加复杂多样, 无人机安全性面临的新挑战日益突出, 主要体现在: (1) 复杂的飞行场景: 如临近空间、传统空域、山地、城市、丛林、极低高度, 甚至是与人相伴的室内外场景; 如全天候、全天时、长周期、密集频繁的飞行任务; 如越来越多以编队、集群等合作方式工作等. 飞行器与环境、人深度耦合, 飞行器之间, 飞行器与环境、人之间的交互更加密切. (2) 多域的安全威胁: 飞行器的安全从适航安全 (传统飞行安全) 拓展到信息安全、智能安全等多个领域, 飞行器不仅仅需要考虑对物理障碍 (山、树、塔、楼, 其他飞行器等) 的感知与规避技术 (sense and aviod, SAA) [1], 更需要考虑电磁空间、赛博空间、智能算法空间带来的新安全威胁, 需要考虑保护 – 检测 – 响应 (protection-detection-response, PDR) [2]、策略 – 保护 – 检测 – 响应 (policy-protection-detection-response, PPDR) [3]、防护 – 检测 – 响应 – 恢

引用格式: 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: 1608–1628, doi: 10.1360/SSI-2022-0338
Pan Q, Guo Y N, Lyu Y, et al. Autonomous safety and security of UAV systems: definition, modeling, and gradation (in Chinese). Sci Sin Inform, 2023, 53: 1608–1628, doi: 10.1360/SSI-2022-0338

复 (protection-detection-response-recovery, PDRR) [4,5] 等信息安全模型在无人机系统安全中的表述. (3) 变化的人/机角色: 飞行器在观察 – 判断 – 决策 – 行动 (observation-orientation-decision-action, OODA) [6] 环路中被赋予了极大的独立于人的自主性, 人与机的关系发生了深刻变化. 考虑飞行场景和多域安全的现实状况, 从机的角度看, 飞行器飞行的自主能力提升几乎没有上界; 而从人的角度看, 安全保障却没有得到足够的重视与关切. 由此可见, 当前和未来的飞行器, 已不再是 “高高在上, 不‘视’人间烟火的空中机器人”, 而是可以随时出现在我们身边, 向纵深发展的必然趋势是智能赋能与多域协同. 智能赋能是更广泛地采用人工智能技术不断提升无人系统的自主性, 多域协同则是高效整合和利用 “陆、海、空、天、电、网” 资源, 实现诸如 “马赛克战”、“多域战” 等新战法, 取得非对称优势, 应对大国竞争.

近年来, 围绕无人系统 (unmanned system, UMS) 的感知、规划决策、控制涌现了大量技术成果和验证系统, 但在无人机系统自主性、智能化水平不断提升的同时, 针对环境、任务的变化, UMS 安全性不但没有在自主框架下给予拓展定义和能力提升, 反而涌现了新的系统性风险和脆弱问题. 在物理安全方面以适航安全框架为主, 其核心 SAA 技术得到长足发展, 包括相关政策法规制定、理论技术研究和试验平台研制. 2018 年北约 (North Atlantic Treaty Organization, NATO) 第一次发布了针对无人机感知与规避的标准化操作文件 STANREC 4811, 对无人机感知与规避提出标准化的配置要求和操作指南. 美国更是在 SAA 体系、政策、技术和装备等方面取得了一系列进展. 在信息安全方面, 对有人机而言, 其对信息技术的依赖程度从提升飞机性能与运行效率起步, 逐步进入飞行器系统核心, 但是角色仍然是 “锦上添花”. 对于无人机而言, 其对信息技术的依赖, 从其诞生开始就是 “生存” 的根本保障. 2017 年发布的最新版美国无人系统路线图中, 网络安全被定义为无人系统未来技术研究的四个核心主题之一. 2019 年美国司法部 (United States Department of Justice, DOJ) 发布无人机修正法案, 要求进行网络安全评估. 新一代人工智能技术极大推动了智能无人系统发展, 其中最具代表性的无人机正快速从大中型/中高空向中小型/中低空应用拓展, 由简单开阔环境向城市/山地/丛林的复杂环境延伸, 无人机与人、环境的交互耦合更深更密, 进入的作战空间、时间和领域更加丰富多样. 当前, 大部分人工智能相关模型均缺乏可解释性, 且绝大部分人工智能核心处理模组、开发平台均掌握在少数西方发达国家手中, 自主可控性差. 当面临真实战场环境的各种恶意攻击, 甚至经济外交打压时, 按照目前的发展模式, 新一代人工智能的实战效果将大打折扣. 未来作战对抗模式将发生巨大变化, 由原来信息域对抗演变为以认知域为主的对抗, 以深度学习为主要算法的智能识别系统愈发容易受到欺骗攻击, 智能识别模型更显脆弱. 2018 年欧盟委员会 (European Commission, EC) 提出鲁棒及可解释的人工智能专项, 旨在提升智能技术的透明性、可靠性及数据安全性, 更多国内外学者从可解释性和鲁棒性等方向开始考虑智能算法的安全性问题. 2019 年美国海军研究局 (United States Naval Research Laboratory, NRL) 发展 “新型数字伪装欺骗技术”, 该技术可迷惑敌方的智能侦察与打击系统, 将坦克识别为轿车. 2019 年美国 DAPAR 提出 “确保人工智能抵御欺骗稳健性” (CARD) 项目, 开发新一代对机器学习模型的欺骗防御技术, 主要体现在防御性机器学习理论、构建防御系统、搭建测试平台等. 在空天基平台 (卫星、无人机) 对机场、阵地、航母等高价值目标的光电识别时, 目标成像分辨率低 (目标成像少像素特性), 仅需少量像素扰动就能欺骗智能识别模型, 导致识别任务失败.

上述研究为无人机系统应对新安全挑战提供了可借鉴的理论和技术成果, 但由于各项研究相对独立, 难以真正有效解决无人机系统面临的 “不断提升的无人自主飞行性能” 与 “日趋严峻的多域安全威胁” 所带来的挑战. 无人机在复杂任务达成和环境交互过程中, 面临的物理、信息、智能域的威胁是交互渗透的, 并非简单叠加. 特别是在资源受限、电磁拒止等约束下, 解决当前无人机系统在物理、信息和智能等领域的安全威胁, 简单的序贯处理, 或者功能叠加更是无法有效应对多域并发协同威胁.

而有效应对这些威胁的核心困难源于无人机系统不断提升的自主性, 可谓 “自主必须安全, 安全需要自主” 相互交错. 本文在阐述无人机安全性与自主性研究进展的基础上, 研究和建立面向物理、信息、智能算法等多域安全威胁的无人机系统自主安全体系架构: 概念、表征模型、能力模型与自主安全能力分级, 简述了构建完善的无人机自主安全理论框架和技术体系所面临的挑战.

2 无人机系统自主与安全研究进展

本节将详细介绍当前关于无人机自主性、无人机物理安全、无人机信息安全, 以及无人机智能安全的相关研究现状和发展动态.

2.1 无人机系统自主性

在无人系统的研究过程中, 人们发现它之所以能够在无操作者参与环境下工作, 其中一个非常关键的因素就是它能够进行自我管理, 即具有一定自主能力. 无人机作为一类典型的无人系统, 其在飞行控制、感知规避、攻防博弈、智能协同等方面的自主性更是受到长期关注和日益重视.

在无人机自主性研究方面, 美国联邦航空管理局 (Federal Aviation Administration, FAA)、国际民航组织 (International Civil Aviation Organization, ICAO)、欧洲航空安全局 (European Aviation Safety Agency, EASA), 包括美国国家标准与技术研究院 (National Institute of Standards and Technology, NIST)、美国陆军无人系统路线图研究小组、美国靶场指挥官委员会、美国宇航局 (National Aeronautics and Space Administration, NASA)、美国国防部 (United States Department of Defense, DoD)、英国国防部 (Ministry of Defence, MoD) 等组织和机构长期致力于无人系统 (unmanned system, UMS) 自主能力研究, 将自主性放在无人系统研究的首位 [7], 认为自主性是无人系统固有的、本质的属性 [8~10], 是有人系统所不具备的标志性特征. 其中, NIST 给出了被公认是较为全面规范的定义 [9]: “自主性是指无人系统拥有感知、观察、分析、交流、计划、制定决策和行动的能力, 并完成人类通过人机交互布置给它的任务. 自主性可根据任务的复杂性、环境的困难性, 以及为完成任务进行的人 – 机交互程度等因素来区分等级, 进而表示出无人系统自我管理的状态和质量.”

无人系统的自主性量化评估对于无人系统的研制和设计、无人系统研究政策制定以及无人系统用户都有非常重要的意义. 无人系统自主性量化评估规避了无人系统在自主性能描述上的模糊性, 有利于提高工程科学性和可操作性. 无人系统自主性评价方法主要有以下几种: 等级法、双坐标轴法、三坐标轴法 [2]、查表法和公式法 [3] 等. 其中, 三坐标轴法因其综合考虑决定系统自主性的各个因素而成为主流评价方法, 比较有代表性的是 NIST 智能系统部门 Huang 等 [2] 提出无人系统自主等级 (autonomy levels for unmanned systems, ALFUS). ALFUS 从任务复杂度、环境复杂度、人在回路比重 3 个维度给出自主的 10 级定义, 对无人系统自身的综合感知、认知、分析、交流、规划、决策和行动/执行等能力进行了描述, 如表 1 所示. NASA 则给出了高空长航时无人机的自主性定义 [10]: “可以视为驾驶员虽不完全脱离无人机回路, 但最大限度地减少人员参与无人机操作, 同时保证无人机遂行任务的灵活性和无人机的稳定性”, 对应自主水平分级如表 2 所示. 无人系统自主性研究目前也成为国内关注的热点问题, 比如王越超等 [7] 采用的蛛网模型对 ALFUS 工作组三因素自主评价模型的扩展. 自从 2008 年以后, ALFUS 逐步减少, 之前对于 ALFUS 的研究都是自主优先, 安全往往成为默认满足的前提条件. 但是对于无人机的自主飞行问题而言, 自主和安全应是两个紧密联系的概念, 必须将两者结合起来, 对已有自主等级模型施加一定的安全约束, 从而实现安全自主.

表 1 无人系统自主等级

Table 1 Autonomy levels for unmanned systems (ALFUS)

参考等级	任务复杂度	环境复杂度	人在回路比重
10	最高适应性、决策空间、团队协作任务; 完全实时规划; 人类智能水平	最低解决方案/可能性比率; 最低的错误余量、可理解性; 最高水平的动态、变化、风险、不确定性限制	独立执行并接近零人在回路互动
9 8 7	高度适应、决策空间、团队协作任务/任务; 高实时性规划	低解决方案/可能性比率; 可理解性高度动态、复杂、对抗性高风险、不确定性限制	UMS 通知人类; 人类提供战略目标, 人在回路比重在 6% ∼ 35% 之间
6 5 4	有限的适应、决策空间、控制任务; 有限的实时规划	中等解决方案/可能性比率; 可理解性; 动态、简单中风险、不确定性限制	人类批准决策, 提供战术目标, 人在回路时间在 36% ∼ 65% 之间
3 2 1	子系统任务/技能	高解决方案/可能性比率; 可理解性静态、简单低风险、不确定性限制	人类决定提供航点, 人在回路时间在 66% ∼ 95% 之间
0	简单任务	静态、简单场景	遥操作

表 2 NASA 高空长航时无人机自主水平分级

Table 2 Autonomous levels of high altitude long endurance UAVs by NASA

等级	级别	描述	自主程度
0	远程控制	在回路中由人操控遥控飞机, 并做所有决策, 操作者处于不间断的控制之中. (100% 人工)	遥控飞机
1	简单自动化	使用一些自动化技术进行远程驾驶, 以减少驾驶员工作量. 人工监控启动/停止任务. (80% 人工)	基础自主驾驶
2	远程操作	操作员允许无人机上系统自主驾驶, 作为外部控制回路的一部分, 人类决定去哪里、何时去、到达后要做什么等, 远程监控、健康监控和有限诊断. 操作员允许无人机执行预编程任务, 只有在无人机未能正确执行任务时才接管. (50% 人工)	飞行器综合健康管理、机载应急管理、路径点导航
3	高度自动化或半自主	无人机自动执行复杂的任务, 系统了解其环境 (态势感知), 并做出日常决策和任务改进, 动态调整飞行和任务变量. 有限的人力监督, 异常管理, 适应不断变化的飞行条件. (20% 人工)	通信中断任务延续、自动起飞/着陆、自适应地形识别
4	全自动化	无人机接收高层次任务目标 (如位置、时间), 将其转化为无需进一步人工干预即可执行的任务. 无人机有能力和权威来做所有的决定, 广泛的态势感知 (内部和外部)、预测和机载飞行重新规划能力, 单机操作. (5% 人工)	自主飞行、实时航路规划、任务传感器导航
5	协同操作	多无人机作为一个集体智能系统共同自主工作, 机群协同, 协同中的单机/系统将拥有最少半自主 (3 级), 以保持协同操作员工作在可管理的等级. 所有无人机的人工操控时间之和不超过操控单个无人机时间	协同飞行、主从协同行动、协同机队领航者概念、机群

2.2 无人机系统物理安全

无人机系统物理安全的感知与规避是指无人机系统通过传感器和数据链路实现对空中交通环境和各类障碍物的有效检测和评估, 预测可能的碰撞威胁, 生成有效规避路径, 并进行应急机动, 从而实

现碰撞规避的技术. FAA 民用无人机路线图定义: “SAA 为无人机提供自分离和碰撞规避功能, 保障与有人机 “看到并规避” 操作类似的等价安全飞行” (SAA capability must provide for self-separation and ultimately for collision avoidance protection between UAS and other aircraft analogous to the “see and avoid” operation of manned aircraft that meets an acceptable level of safety). 感知与规避系统的关键技术包括: 实现空间环境感知, 利用各种传感器对空间环境和各类障碍物进行检测, 获取可能存在碰撞威胁的动态或静态目标的检测环节; 通过有效的估计、跟踪算法实现飞行空间目标的飞行状态估计, 位置、速度的跟踪环节; 根据目标的运动状态和空间分离规则对碰撞进行威胁程度的计算; 对存在多个碰撞威胁目标时, 对目标进行基于威胁程度的等级划分和先后排序等威胁评估等环节; 根据威胁评估结果判定给出飞行管理决策, 根据最小分离点 (closest point of approach, CPA)、碰撞时间 (time to collision, TTC) 等计算规避路径等规避决策与路径规划环节; 通过机动输出, 执行规避决策和规避路径跟踪的规避机动环节等.

在面向物理安全的感知规避技术发展中, 规避是重要的响应阶段 [11], 目前发展了大量感知规避相关方法, 主要包括基于数学规划的方法、基于路标图的方法、基于空间搜索方法、遗传算法等 [12]. 传统的动态规划法是应用较为广泛的一种方法, 通过将规划问题等效为多级决策问题. 例如, 为克服混合整数线性规划 (mixed integer linear programming, MILP) 方法计算量大的缺陷, 文献 [13] 提出基于有限视场的动态 MILP 路径规划算法. 而 Oliver 等 [14] 则将 MILP 与模型预测控制法 (model predictive control, MPC) 相结合, 辅助训练语言决策树 (linguistic decision trees, LDTs) 以提升在线航路规划的实时性. 针对城市环境下无人机自适应目标跟踪与避障问题, Wu 等 [15] 首先利用分布式模型预测控制 (distributed model predictive control, DMPC) 方法进行优化建模, 其次利用自适应蝗虫优化算法 (adaptive grasshopper optimization algorithm, AGOA) 在线优化求解最优控制器. 文献 [16] 提出通过利用滚动时域控制法 (receding horizon control, RHC) 优化人工势场中的附加控制力来实现无人机的在线避障. 针对未知情况下的航路规划问题, 文献 [17, 18] 采用动态规划策略确保飞机实时规划未来路径, 通过引入基于马尔可夫 (Markov) 的生存模型实时评估飞机的生存状态概率. 基于路标图的方法主要包括可视图法 (visibility graph) [19] 和 Voronoi 图法 [20] 等. 为提高 Voronoi 图在突发威胁下无人机动态航迹规划方面的实用性, 文献 [21] 首先通过引入威胁源的不可穿越区域边界提出一种改进型的 Voronoi 图构造模型, 提高了航迹段对威胁的敏感性, 其次在分析突发威胁对于航迹拓扑空间影响的基础上, 提出了一种基于改进型 Voronoi 图的航迹重规划模型, 结合 D* 算法对突发情况下的航迹重规划进行了研究, 进而规划出理想航迹. 基于空间搜索方法是另一个重要的碰撞规避方式, 常见的搜索方法包括 A* 算法 [22]、D* 算法 [21] 以及快速搜索随机树算法 (rapidly-exploring random tree, RRT) 等. 其中, RRT 算法是一种采用增量方式增长的随机采样算法, 用于解决有代数约束 (障碍带来的) 和微分约束 (非完整性和动态环境带来的) 的高维空间问题. 其优势在于无需对系统进行建模, 无需对搜索区域进行几何划分, 覆盖率高, 搜索范围广, 可以尽可能地探索未知区域, 但同时也存在算法计算代价过大的问题. 为解决此问题, 研究者们提出 RRT 的各种改进形式. 如 Lin 等 [23] 通过 Close-Loop RRT, 将闭环路机动控制与 RRT 路径搜索相结合, 通过闭环机动控制飞行轨迹预测实现更小的预测误差和碰撞风险, 路径规划考虑目标机动约束.

随着强化学习的发展, 很多学者将其成功地运用在了无人机的感知与规避过程中, La 等 [24] 提出了一种基于强化学习的障碍物/捕食者规避方法, 通过将高层的 Q-learning 决策过程与低级的无人机集群控制结合, 实现了对障碍物的有效规避. 在该方向进一步的成果中, Hung 等 [25] 实现了完全基于 Q-learning 方法的障碍规避功能, 其中碰撞规避被建模为一项激励函数. Long 等 [26] 基于深度强化学习方法 (deep reinforcement learning, DRL), 构建深度神经网络模仿基于几何关系的障碍规避方法

ORCA, 从而实现在多智能体设置下的碰撞规避. 在文献 [27] 提出的类似的深度强化学习框架中, 通过对学习策略和激励函数的改进, 实现了比 ORCA 更好的规避性能. Lyu 等 [28] 基于集中式的强化学习框架和分布式的控制策略大大提高了强化学习的效率. 除此之外, 由于 Q-learning 算法的状态空间和动作空间均为离散的, 因此其规划航路的可飞性较差, 且难以应付动态威胁. 针对此缺陷, 研究者提出将深度学习和强化学习相结合, 组成深度强化学习算法, 以满足状态空间或动作空间连续化的需求. Yan 等 [29] 在路径规划中引入了不同改进型的深度 Q 网络 (deep-Q network, DQN) [30], 取得了较好的效果, 但由于 DQN 的动作空间仍然是离散形式的, 因此规划的路径质量仍有提升空间.

无人机感知系统主要分为合作式感知系统及非合作式感知系统, 常见的合作式感知系统主要由 ADS-B 和 T-CAS 技术作为支撑. 其中 ADS-B 能够获得目标的绝对经纬高度信息, 有效防止有人机与无人机在空中发生碰撞, 其感知精度取决于目标的导航系统精度 [31]. 早期的 ADS-B 收发器的尺寸和重量限制了它们在机翼上的应用, 潜在的 SAA 应用主要是在模拟环境中进行研究 [32]. 在此之后, 低功率和低成本的 ADS-B 变异型也为无人机进行了开发 [33]. 2018 年, NASA [34] 还在其无人机系统 (UAS) 交通管理 (UTM) 项目中探索了 ADS-B 在无人机中的使用. 随着 ADS-B 用于低空无人机的增加, ADS-B 感知技术的局限性也凸显. 最主要的是频率拥塞, 这将影响无人机机队的增长. 从 Guterres 等 [35] 的研究中可以看出, ADS-B 是一种不适合未来无人机应用增长的解决方案, 因为它会造成 ADS-B 共信道干扰的可测量增加, 这可能会对飞机之间的机对机和机对地性能产生负面影响. 而 T-CAS 系统能够获得高精度的目标的相对距离、相对方位和高度, 是一种在有人机上得到广泛应用的成熟技术 [36]. 1981 年 FAA [37] 决定在 BCAS 的基础之上开发和实施 TCAS. 2006 年 MIT 林肯实验室 [38] 联合 MITRE 公司基于美国空军和 FAA 对全球鹰无人机的感知能力需求完成无人机机载 T-CAS II 的感知与规避测试. 2007 年美国空军研究实验室 [39] 研究分析了使用多个传感器数据驱动一种与配备 TCAS II 的载人飞机兼容的避碰算法的可行性. Lin 等 [40] 为解决小型飞机在低空飞行时的冲突问题, 基于视觉飞行规则 (visual flight rules, VFR) 下的自动依赖监视 (automatic dependent surveillance-broadcast, ADS-B) 概念, 提出了一种交通预警与避障系统 (traffic collision avoidance system, TCAS) 算法, 为低空飞行提供了更广泛的航空安全.

通过以上分析, 合作式感知方式的前提是本机与空域其他飞行器能够建立基于应答机制或广播的合作式信息交互通道, 难以保证在真实飞行环境中非合作目标存在情况下的感知与规避, 因此, 单一的合作式感知与规避技术难以真正保证无人机的空域飞行安全. 为了应对以上问题, 诸多研究者将目光转向非合作式感知. 在非合作环境感知过程中, 无人机搭载的常用非合作目标感知设备包括雷达、激光雷达、超声波等主动感知设备以及光电、红外等被动感知设备. Accardo 等 [41] 描述了一种基于雷达的 SAA 系统, 通过 Kalman 滤波方法实现对目标飞行器的状态估计, 并通过有人机的飞行测试证明了其算法的有效性. Owen 等 [42] 设计了一种适用于大、中型无人机感知与规避任务的雷达系统, 并通过提取信号特征和误差精度分析验证了该雷达的空域目标感知能力. Newmeyer 等 [43] 设计了一种适用于小型无人机的毫米波雷达空域感知系统, 通过一系列的信号处理方法获得目标的距离和方位角度后, 用 RANSAC 方法实现对目标的跟踪. 在检测中通常使用单静态脉冲调制雷达有时也会用到连续波 (continuous wave, CW) 技术 [44, 45].

针对大中型无人机系统, 美国率先启动了地基感知与规避系统的研究, 并在 2014 年前完成了地基感知与规避系统的验证与列装 [46]. 2018 年 Sahawneh 等 [47] 提出了一个完整的、概念证明的小型无人驾驶飞机系统的感知和回避解决方案, 包括一个小型低成本的地面雷达系统、多目标跟踪和估计、碰撞检测和回避计划. 2019 年 Meer 等 [48] 研究了无人机通过蜂窝技术将其位置信息传输到云, 而 GBSAA 基站通过 ADS-B 技术访问聚合信息并将其发送到启用 ADS-B 的飞行器, 同时进行了分析和仿真研

究, 研究了不同网络参数 (如无人机数量和 ADS-B 启用飞行器) 下 ADS-B 信息碰撞概率. Fasano 等 [49] 研究了一种基于雷达/光电数据融合的多传感器障碍物检测与跟踪系统的硬件/软件实现和飞行结果. Chen 等 [50] 研究了一种新的基于主动感知的飞行机器人在动态环境中的避障模式. 他们没有融合多个传感器来扩大视场 (field of view, FOV), 而是引入了一种替代方法, 利用具有独立旋转自由度的立体摄像机来主动感知障碍物.

针对小微型无人机, Alvarez 等 [51] 提出了一种基于单目视觉的四旋翼无人机感知与规避系统, 通过融合单目传感器感知信息和自身运动测量基于机载实时处理系统构建地图的方式, 实现在室内场景中反应式的碰撞规避. Lyu 等 [52] 设计了一种单目视觉感知与规避系统, 包括单目相机以及机载处理系统, 不依赖深度信息, 仅基于角度感知信息实现了小型无人机对空中目标的有效障碍规避. 为解决单目相机在深度估计方面的不足, 基于双目相机 [53]、深度相机 [54]、激光雷达 [55] 等感知方法和系统得到广泛的研究. Gageik 等 [56] 提出一种低成本多传感器障碍规避系统, 通过搭载红外、视觉、激光等多种类型传感器和机载处理系统, 能实现在复杂火灾场景中的搜救任务并保证障碍规避.

2.3 无人机系统信息安全

随着电子信息与无人驾驶技术的迅猛发展, 无人机在各领域的应用规模不断扩展, 其信息安全问题得到了越来越广泛的关注. 早在 2011 年, 伊朗军方就曾通过 GPS 欺骗技术成功捕获了一架美国 RQ-170 无人机 [57], 同年, 一种名为 “keylogger” [58] 的病毒被传播至美国内华达州空军基地中, 窃取了大量无人机的遥测与实时侦察情报数据. 这些案例迫使美国政府与军方开始高度重视无人机信息安全. 2019 年, 美国司法部发布无人机修正法案, 要求对无人机系统进行网络安全评估. 同年, 由美国航空院校、企业与军方专家联合编写的无人机信息安全专著 *Unmanned Aerial System in the Cyber Domain* 发布, 从工业控制系统的角度切入, 对信息对抗反无人机技术与无人机的信息安全防护技术做了全面的分析与总结.

无人机系统一般可以分为地面站、飞控导航系统、通讯网络、传感器系统与指挥控制系统五部分. 由于其飞控软件、通信协议、传感器、操作系统等软硬件相关设备多基于通用芯片、开源系统、通用协议与软件架构进行设计开发, 在设计时也以易用性、轻量化等原则为主, 而忽略安全性 [59]. 这样的设计导致无人机系统会存在明显的安全漏洞, 使得系统架构的每一个层次都可能受到针对其脆弱性特点的恶意攻击. 据调研, 美国普渡大学 (Purdue University System)、密歇根大学 (University of Michigan)、加利福尼亚大学尔湾分校 (University of California, Irvine)、得克萨斯 A&M 大学 (Texas A&M University, TAMU) 等国外知名高校, 以及西安电子科技大学、西北工业大学、华东师范大学、浙江大学等国内院校正在开展相关的研究工作, 具体包括传感器安全、软件安全、通讯网络安全和任务系统安全.

(1) 无人机机载传感器系统信息安全威胁. 无人机系统的传感器由于技术与成本的限制. 在实际的运行中缺乏必要的认证机制, 往往面临着可用性的风险, 与此同时, 为了强调数据传输的实时性, 数据传输也缺乏加密手段, 因此还面临着数据完整性及保密性的风险. 除此之外, 无人机载荷也不具备对异常攻击的主动检测与响应能力. 这些脆弱性往往可以被用来进行干扰攻击和欺骗攻击. 攻击对象一般包括视觉传感器、惯性传感器、GPS 传感器以及激光雷达等. 针对惯性传感器, 常见的攻击包括声波欺骗 [60]、声信号注入 [61]、带外信号注入 [62] 等. 这些攻击通常基于传感器的物理特性进行干扰, 攻击距离短, 作用范围有限, 很难对中大型无人机产生有效的干扰. 针对 GPS 传感器, 攻击者往往进行转移攻击 [63], 即使用注入虚假的数据来欺骗 GPS 接收器. 然而, 由于无人机受到攻击时传感器数据会显著波动, 这种攻击往往可以通过攻击检测器 [64~67] 进行检测, 即将量测的残

差向量作为随机变量, 使用单个或多个样本进行攻击检测. 然而无论是单个还是多个样本, 这种方法都存在明显的缺陷. 不同于上述针对特定攻击进行检测和防护的方法, 多传感器数据融合技术基于惯性测量单元 (inertial measurement unit, IMU)、全球导航卫星系统 (Global Navigation Satellite System, GNSS)、相机、雷达或 LiDAR 等 [68,69] 传感器数据进行融合估计, 能够提高无人机组合导航的精确度、可靠性和鲁棒性, 增强数据的可信度, 在发生欺骗攻击时, 提供了更大的冗余度来检测和防御注入到传感器中的攻击, 以提高系统安全性 [70]. 如加利福尼亚大学尔湾分校 (University of California, Irvine) 的 Shen 等 [67] 证明 GNSS/INS/LiDAR 融合的组合导航系统中, 恒定偏移的 GPS 欺骗攻击有效性会大大降低, 但是该方法仍然无法防御针对融合算法本身存在的不确定性所构造的攻击 [67].

(2) 无人机机载和地面系统软件安全威胁. 对于地面站、飞控导航以及指挥控制系统中普遍使用的软件系统, 攻击者通常通过漏洞利用的方式对目标软件或系统进行攻击. 如安装恶意软件进行后门攻击, 对相关的软件和工具进行逆向破解来获取指令格式 [71], 以及通过缓冲区溢出攻击导致无人机坠毁等 [72], 相应的防护通常分为漏洞发现和漏洞防护两个部分, 在漏洞发现方面, 研究人员通常通过模糊测试技术 [73] 分析程序运行异常来发现漏洞. 如西安电子科技大学的叶向豪使用模糊测试技术发现无人机系统中的拒绝服务漏洞. 而在具体的防护方面, 通过控制流完整性技术来实现对控制流劫持攻击的防御 [74]. 机载操作系统除了保证不同安全级别的软件互不影响外, 还要避免系统遭受来自网络连接的恶意攻击. 面对这种情况, 国外开始研究多级别安全性 (multiple independent levels of security, MILS) 机载嵌入式操作系统, 该标准将操作系统进行层次划分, 采用分而治之的思想可以指数级地减少安全隐患. 如通过内存隔离技术将进程所在内存空间虚拟化, 使攻击者难以进行内存损坏攻击 [75].

(3) 无人机通信系统安全威胁. 无人机系统面临着网络传输协议多样, 缺乏统一安全传输标准的问题, 同时, 无人机系统通讯网络不仅包括传统的互联网通信, 某些链路同时需要电磁通信, 而这些开放且标准不一的通讯环境, 使得无人机系统的通讯网络面临着巨大的脆弱性. 在国内, 2020 年国家标准化管理委员会下达了《民用无人机产品安全要求》的强制性国家标准制修订计划, 其中整机安全中的无线电抗干扰要求提出了无人机无线电发射功率、频段、频率应满足国家规定的要求, 不影响现有国家的通信安全, 也不对操作者、附近相关人员造成电磁辐射伤害. 对于通讯网络, 攻击者可以基于物理层硬件的脆弱性进行攻击, 包括窃听 [76]、篡改 [77] 与干扰 [78] 攻击. 也可以针对其协议中加密认证机制普遍缺失的问题进行重放攻击以及拒绝服务攻击 [79~81]. 对于内部各组件之间的通讯, 无人机往往需要使用 CAN 总线对各个部件进行连接. 但由于缺乏安全机制, 攻击者很容易从受损的 ECU 或安装在总线上的恶意设备发起攻击导致中断传输. 针对 CAN 的攻击可以分为伪装 CAN 帧的欺骗攻击和以高发送速率注入帧的拒绝服务攻击 [82]. 第一种攻击中, 攻击者往往首先渗透正常节点或将恶意节点附加到总线上, 然后控制受害者向 CAN 总线发送伪造的帧. 例如节点发送具有其不允许发送的 ID 的帧 [83], 或者通过注入虚假数据或重放旧数据帧来欺骗 ECU [84]. 在纠错机制中, 攻击者可以进行 “Bus-Off” 攻击 [85], 即使用精心构建的数据帧来主动制造传输冲突, 从而迫使受害者进入总线关闭状态.

针对无线网络存在的安全问题, 防护方法一般包括密码技术和物理层安全技术. 密码技术主要用来对通讯链路和关键数据进行加密认证 [86,87], 但会产生额外的计算开销, 且需要对硬件进行改造. 对于计算资源和能源有限的小型无人机来说应用有限. 另外, 对于多机系统, 无人机数量的增加也会给密钥的管理和传输带来问题. 物理层安全技术是根据实际物理场景建立信道的保密模型, 并通过优化方法来增强模型的保密容量来对抗窃听攻击和干扰攻击 [88~90]. 与应用于上层的加密技术相比, 无线通信物理层安全技术可以在不需要密钥和复杂算法的情况下保护无线数据传输, 因此更适用于大规模分

布式的无线网络. 无论是通讯网络还是内部通讯系统, 大部分防护技术的本质还属于攻击产生后静态的被动防护措施, 而缺乏对潜在安全威胁主动发现与响应技术.

2.4 无人机系统智能安全

新一代人工智能极大推动了下一代无人机任务系统的发展, 无人机系统架构和软硬件结构将大幅革新, 然而大规模的智能化应用将会给无人机系统带来较大的风险. 首先, 多数人工智能模型是黑盒模型, 具有不可解释性. 这些模型可能本身存在缺陷, 可靠性不足, 这极有可能导致其在动态复杂的任务环境中出现错误的决策行为. 其次, 人工智能模型的线性特性缺乏对扰动的抵抗能力 [91]. 即使决策 "可靠" 的模型也同样容易受到对抗攻击, 攻击者只需要在输入样本中添加精心构造的、隐蔽的扰动就可以很容易地让模型决策出错, 如数据污染攻击、后门攻击等. 在无人机系统中, 相机和 LiDAR 的感知数据往往会作为人工智能模型的数据输入, 因此极易被当成攻击的切入点. 如 Cao 等 [92] 通过恶意构造的点云数据成功通过 LiDAR 欺骗目标跟踪检测器. 在此基础上, Cao 等 [93] 提出同时攻击相机和 LiDAR 的方式会带来更佳的攻击效果. 对于无人机系统来说, 这样的攻击方式可以用来干扰控制算法 [94] 和攻击检测模型 [95], 可用于无人机侦察人工智能算法的对抗攻击样本生成, 来降低分类算法的准确率, 也可以用来骗取无人机在侦察任务中寻找敌人的活动. 这些因素会增加支持无人机智能化的软硬件、算法和大规模数据应用的脆弱性, 给无人机自主智能的实时性、精确性和完备性带来智能安全, 也将带来更为严峻的智能安全挑战.

同样, 人工智能技术还可用于信息安全防护中. 从技术角度来看, 传统的安全防护措施可以防御许多已知安全威胁, 但随着无人机的应用领域不断拓宽, 应用场景趋于复杂, 以及各种攻击手段的推陈出新, 传统的安全防御技术已经难以应对. 人工智能技术具有较强的自动特征提取与分析能力, 可用于无人机系统中多个层次的安全防范, 成为防护新型攻击的可行技术方向. 在无人机的安全问题中, 监督学习的应用广泛, 如使用 LSTM 对合法 ADS-B 报文序列进行航路识别来计算飞行偏差 [96], 使用 LSTM, SVM 等算法检测传感器欺骗攻击 [95, 97~99], 基于随机森林算法对无人机操作员进行身份认证来检测恶意重放攻击 [100]. 在多无人机系统中, 监督学习算法可以对异常状态进行学习来保护无人机免受拒绝服务攻击 [101].

强化学习在无人机信息安全中也具有广泛的应用场景. 如 Johansson 等 [102] 使用强化学习方法学习策略函数来获取最佳的路径参数, 使无人机在受到干扰失去控制时快速找到最佳飞行路径, 前往可以与地面站进行正常通信的位置. Lu 等 [103] 提出了一种基于强化学习的无人机电机温度异常检测方法, 通过学习电机的温度阈值并在不同阈值下执行相应的飞行策略来防止其在异常温度下运行, 可在电机过热时进行迫降, 保证飞行安全. Lin 等 [104] 设计了一种无人机轨迹和功率控制方案, 通过强化学习的方法对用户感知信道轨迹和发射功率进行选择, 用以对抗恶意干扰攻击. 如 Yang 等 [105] 在一对互相通信的无人机和一个攻击者之间建立一个连接博弈, 并用生成对抗网络这个博弈进行优化, 使无人机可以相应地调整连接策略, 以防御恶意攻击, 提高攻击者对无人机的干扰能力. Xiao 等 [106] 研究了主观攻击者和无人机传输之间基于展望理论的静态博弈, 并提出了一种基于 Q-learning 的功率分配方案, 有效抵御欺骗与干扰攻击.

综上, 在无人机系统自主性、智能化水平不断提升的同时, 针对环境、任务的变化, 无人机系统安全性不但没有在自主框架下给予拓展定义和能力提升, 反而涌现了新的系统性风险和脆弱问题. 上述研究为无人机系统应对新安全挑战提供了可借鉴的理论和技术成果, 但由于各项研究相对独立, 难以真正有效解决无人机系统面临的 "不断提升的无人自主飞行性能" 与 "日趋严峻的多域安全威胁" 所带来的挑战. 考虑到自主性是无人系统固有的本质属性, "自主安全" 更适于无人机系统, 也符合未来

智能无人机系统技术的发展趋势. 在第 3 节中, 我们率先从自主与多域安全威胁的角度提出 “无人机自主安全架构”.

3 无人机系统自主安全

本节首先定义了 “自主安全” 新概念; 其次, 基于 IPDRR 模型与 OODA 环构建了自主安全要素表征模型; 然后, 建立了无人机系统自主安全模型, 并提出无人机系统自主安全能力分级.

3.1 自主安全概念

无人机自主要求其具有自主感知、信息处理、决策以及执行自主决策的能力. 通过自主控制, 无人机可以根据环境和任务的变化和不确定性自适应地、自主地调整控制目标. 鉴于自主性是无人机系统中人、机、环境的核心关键, 我们提出 “无人机自主安全” 新概念. 所谓 “自主安全” 是指面向系统高度自主或智能自主时, 无人机本体所具备的自主提供一种高效可靠的任务达成和安全保障的能力 (能够自主执行并完成任务、自主达到并保持系统安全的能力). 这种能力不仅包含无人机本体在执行任务时具备感知、规划、决策、学习、行动以及人 – 机交互的能力, 也包括在面临物理、信息、智能等多域威胁时, 无人机本体具备自主安全威胁感知、检测、识别、规避与防御等能力. 其中, 安全是保障功能, 自主是内核, 自主安全资源配置、外部资源以及人工干预程度为衡量指标. 则其内涵即为在特定任务场景下, 无人机本体不依赖或很少依赖外部资源和人工干预, 以最小的自主安全资源配置, 最高程度地自主完成任务, 最大限度地自主避免不安全状态和任务环境中的不安全因素, 确保本体不遭受突发或者恶意的破坏或损毁, 以及不形成对人、环境的威胁.

从上述定义可以看出, 无人机系统自主安全是一种能力, 传递的是一种信任, 而这种信任来自于无人机本体所具备的高度自主和安全保障能力. 新一代无人机系统安全观应实现 “以无人系统为本、以自主为核心、以资源配置优化为支撑” 的面向物理、信息、智能等多域威胁的安全保障能力. 无人机系统本体是安全能力的载体, 安全体系建设要重点考虑无人机系统本体的自主能动因素. 自主性是无人机系统安全能力的核心, 自主安全技术未来将更加深入细分到无人机系统产业与应用领域, 安全产品和应用将更加多样化. 值得注意的是这里所提出的 “无人机自主安全” 是面向系统高度自主或智能自主的安全框架, 是一种理想框架, 或者说是理论框架, 具体技术实现可以分阶段、分级, 或按照应用需求具体定义与实施. 我国是各类无人机系统的技术和产业大国, 自主与安全协调平衡是无人机技术与产业发展的战略需求和核心关键. 研究和建立我国无人机系统自主安全体系和生态 (理论、技术、标准), 推进示范性和产业化应用, 将为我国无人机行业的健康快速发展奠定基础和提供支撑, 也为其他无人系统安全体系建设提供借鉴.

3.2 自主安全要素表征模型

IPDRR[107] 模型是由美国国家标准与技术研究院 (National Institute of Standards and Technology, NIST) 所提出的一种网络安全框架, 主要包括识别 (identify)、防护 (protection)、检测 (detection)、响应 (response) 和恢复 (recovery) 5 个重要环节. IPDRR 实现了 “事前、事中、事后” 的全过程覆盖, 从以防护为核心的模型, 转向以检测能力为核心的模型, 变被动为主动, 最终达成自适应的安全能力. OODA 也被称为 “博伊德循环” 或 “决策周期环”, 是美国空军军官 John Boyd 在 20 世纪 70 年代基于对抗性决策提出的作战理论. 其本质是被动式, 没有观察 O 就不会有判断 O、决策 D 和行动 A. 在复杂环境中, 人机环境系统的 OODA 需要 “零信任” 模式, 尤其是在充满 “诡、诈” 的博弈智能中, 从

表 3 R^3 安全模型要素对照表

Table 3 Element comparison of the R^3 safety model

IPDRR	OODA	R^3
Identify (I) Protection (P)	Observation (O)	Representation (R_1)
Detection (D)	Orientation (O)	Reasoning (R_2)
Response (R) & Recovery (R)	Decision (D) & Action (A)	Response (R_3)

“是什么” 到 “应是什么” 之间完成 “为什么” 的过程跳跃. 融入各种智能技术的 OODA 环将呈现出全维感知、快速准确、优化高效、实时精准的特点, 为战场决策的科学化, 进而夺取主动权提供了新的视角和手段.

在信息安全领域, IPDRR 模型中的最后响应 (或者修复) 一般可以做到 “药到病除”, 立竿见影. 但是往往针对静态系统, 甚至是离线系统. 在指挥控制领域, OODA 模型是环状结构, 4 个节点始终处于循环状态, 一般具有反馈机制, 可以不断优化. 就无人机这类运动平台, 对安全威胁的快速响应十分重要, 但是考虑多域威胁, 还必须兼顾全状态安全目标. 比如系统躲过了电子干扰, 但是核心数据可能被窃取, 同样是不能接受的. 受此启发, 我们以面向无人自主的物理、信息、智能等多域威胁为切入点, 基于 IPDRR 模型与 OODA 循环的共性属性特征, 初步建立了无人机自主安全要素表征模型 (R^3 模型), 如表 3 所示. 其中, R_1, R_2, R_3 分别代表表征 (representation)、推理 (reasoning) 与响应 (response).

• **表征 (R_1): 对无人机系统面临物理、信息、智能等多域威胁的属性和态势有效认知与表征.** 无人机的安全表征从传统的物理域逐步拓展到信息域和智能域, 需要获得物理、信息、智能域交互渗透场景下精准的安全态势. 如何有效将多域、异类、非结构化的海量信息映射为无人机自主安全态势, 需要考虑各域内安全威胁描述载体、数据组织形式等的差异性, 研究数据呈现的异源、异构、异步、跨模态、跨语义等特征, 找到表示、关联对齐和挖掘分析的方法, 解决由多域威胁导致域间数据共生特性难以挖掘的难题, 实现无人机自主安全的多域安全表征, 这是实现多域威胁无人机自主安全研究的首要问题.

• **推理 (R_2): 对无人机系统高自主运行、强约束环境的多威胁可靠预测推理与预案准备.** 伴随无人机系统自主性能的不断提升, 对外部资源、人为干预的依赖愈加减少, 这可以认为是 “主动” 的资源受限、电磁拒止行为; 而实际战场环境, 或者特殊应用场景下 “被动” 的资源受限、电磁拒止将是更加严峻的挑战. 在这类强约束条件下, 无人机系统面临信息不完备、不对称造成的长周期不确定性挑战, 需要寻找鲁棒可信的安全评估方法, 研究高可靠安全风险预测机制, 基于多域威胁模式辨识、细粒度分级安全态势感知和多域安全知识图谱 (先验信息), 以最小化期望自由能为目标生成自主推理模型, 实现基于威胁属性和安全等级的多域威胁长周期决策预案 (后验信息), 这是无人机系统实现多域威胁自主安全的关键指标.

• **响应 (R_3): 对无人机系统多域威胁信息域快速响应和物理域迭代控制的策略与行动.** 传统单域的威胁消减方案、防护手段和控制策略很难有效处理多域并发、状态多维度、任务多约束的复杂无人机系统安全威胁. 当某一域存在安全威胁, 会导致无人机系统其他域的不稳定甚至崩溃.

提出此表征模型的目标是构建无人机自主安全架构 (模型、分级等), 探索无人系统应对多域安全威胁问题, 提出在资源受限、电磁拒止等复杂约束下无人系统自主安全的新理念和新范式.

表 4 要素特征先验数据集

Table 4 Prior dataset of elemental feature prior distribution

Elemental feature sets		Autonomous safety & security resource sets		
		Sensing resources (image, navigation, communication, environmental sensors, etc.)	Computing resources (algorithms, examples, platform, etc.)	Execution resources (dynamic system, attitude system, mission system, etc.)
		$\mathcal{G}_1=\{g_1^k\|k=1,\ldots,m_1\}$	$\mathcal{G}_2=\{g_2^l\|l=1,\ldots,m_2\}$	$\mathcal{G}_3=\{g_3^m\|m=1,\ldots,m_3\}$
Case I	$\mathcal{F}_1=\{f_1^u\|u=1,\ldots,n_1\}$	$\{\delta_{11}(i,j)\|i\in\mathcal{F}_1,j\in\mathcal{G}_1\}$	$\{\delta_{12}(i,j)\|i\in\mathcal{F}_1,j\in\mathcal{G}_2\}$	$\{\delta_{13}(i,j)\|i\in\mathcal{F}_1,j\in\mathcal{G}_3\}$
Case II	$\mathcal{F}_2=\{f_2^v\|v=1,\ldots,n_2\}$	$\{\delta_{21}(i,j)\|i\in\mathcal{F}_2,j\in\mathcal{G}_1\}$	$\{\delta_{22}(i,j)\|i\in\mathcal{F}_2,j\in\mathcal{G}_2\}$	$\{\delta_{23}(i,j)\|i\in\mathcal{F}_2,j\in\mathcal{G}_3\}$
Case III	$\mathcal{F}_3=\{f_3^w\|w=1,\ldots,n_3\}$	$\{\delta_{31}(i,j)\|i\in\mathcal{F}_3,j\in\mathcal{G}_1\}$	$\{\delta_{32}(i,j)\|i\in\mathcal{F}_3,j\in\mathcal{G}_2\}$	$\{\delta_{33}(i,j)\|i\in\mathcal{F}_3,j\in\mathcal{G}_3\}$
⋮	⋮	⋮	⋮	⋮
Case N	$\mathcal{F}_N=\{f_N^r\|r=1,\ldots,n_N\}$	$\{\delta_{N1}(i,j)\|i\in\mathcal{F}_N,j\in\mathcal{G}_1\}$	$\{\delta_{N2}(i,j)\|i\in\mathcal{F}_N,j\in\mathcal{G}_2\}$	$\{\delta_{N3}(i,j)\|i\in\mathcal{F}_N,j\in\mathcal{G}_3\}$

3.3 面向多域威胁和系统高度自主的自主安全能力模型

根据自主安全定义可知, 在面向多域威胁和系统高度自主的无人机系统自主安全中, 无机本体是关键, 安全是保障, 自主性是内核, 安全资源配置、外部资源以及人工干预依赖程度优化是支撑. 本小节我们采用 PASS (power of autonomous safety & security) 表示无人机在执行特定任务时, 对人和环境产生威胁时, 以及无人机系统遭遇物理、信息、智能等多域安全威胁时, 无人机本体所具备的自主与安全保障能力, 其中 safety & security 表征面向多域威胁的安全性.

记 $\mathcal{N}=\{$情形 I, 情形 II, $\ldots$, 情形$N\}$ 为 N 种情形集合, 其中情形 p, $p=$ I, II, $\ldots$, N 表示第 p 种无人机本体执行任务时所处的环境情景形式, 主要包括威胁类型 (如物理域、信息域, 或智能域威胁) 和无人机本体的自主水平. 记 $\mathcal{F}_p=\{f_p^\ell=(s_p^\ell,a_p^\ell),\ell=1,2,\ldots,n_p\}$ 为第 $p,p=1,2,\ldots,N$ 种情形中所包含的所有要素特征集合, 其中 (s_ℓ,a_ℓ) 表示情形 p 中第 ℓ 个要素特征元素, s_ℓ 为威胁特征, a_ℓ 为无人机自主等级, n_p 为威胁特征总个数. 根据属性和功能类型, 无人机系统机载自主安全资源可分为传感资源 (影像、导航、通讯、各类环境传感器等)、各类计算资源 (算法、算例、平台等)、各类执行资源 (动力姿态、姿态系统、任务系统等). 记 $\mathcal{G}_1=\{g_1^k|k=1,\ldots,m_1\}$ 为各类传感资源集合, $\mathcal{G}_2=\{g_2^l|l=1,\ldots,m_2\}$ 为各类计算资源集合, 及 $\mathcal{G}_3=\{g_3^m|m=1,\ldots,m_3\}$ 为各类执行资源集合, 其中 m_1,m_2,m_3 分别为各类传感资源、各类计算资源以及各类执行资源总类型数. 假设在有限自主安全资源条件下, 则我们可得要素特征先验数据集如表 4 所示.

记 $D_{pq}=\{\delta_{pq}(i,j)|i\in\mathcal{F}_p,j\in\mathcal{G}_q\}$, 其中 $p=1,2,\ldots,N,q=1,2,3$, 则要素特征先验数据集可表示为 $\mathcal{D}=\{D_{pq}|p=1,2,\ldots,N,q=1,2,3\}$. 此数据集通常可通过以下两种方式获取:

- 分析无人机飞行记录, 确定威胁类型与无人机本体的自主水平, 找到此类型威胁及无人机自主水平与无人机安全状态之间的关系, 进而构造出所需要素特征先验数据集;

- 通过设计专门的实验飞行, 获取不同威胁和自主水平下的要素特征先验数据集.

则根据上述方式所获取的数据集 $\mathcal{D}$, 可得多域要素特征先验分布矩阵, 记为

$$\Theta := \begin{bmatrix} \vartheta_{11}(D_{11}) & \vartheta_{12}(D_{12}) & \vartheta_{13}(D_{13}) \\ \vartheta_{21}(D_{21}) & \vartheta_{22}(D_{22}) & \vartheta_{23}(D_{23}) \\ \vartheta_{31}(D_{31}) & \vartheta_{32}(D_{32}) & \vartheta_{33}(D_{33}) \\ \vdots & \vdots & \vdots \\ \vartheta_{N1}(D_{N1}) & \vartheta_{N2}(D_{N2}) & \vartheta_{N3}(D_{N3}) \end{bmatrix} \triangleq [\vartheta_{pq}]_{p=1,2,\ldots,N,q=1,2,3}, \tag{1}$$

其中, $\vartheta_{pq} = \vartheta_{pq}(D_{pq})$, $p = 1, 2, \ldots, N, q = 1, 2, 3$, 表示当前第 q 种资源下第 p 类要素特征先验分布, 是基于历史数据集而提炼出的自主安全资源配置与要素特征之间的关系.

假设当前可配置自主安全资源集分别为: 传感资源 $\mathcal{O} \subseteq \mathcal{G}_1$、各类计算资源 $\mathcal{A} \subseteq \mathcal{G}_2$、各类执行资源 $\mathcal{C} \subseteq \mathcal{G}_3$, 则多域要素特征先验分布条件下的自主安全资源配置后验概率矩阵可表示为

$$\Omega = \begin{bmatrix} P\left(\mu_{11}\left(\xi_{11}\right)|\vartheta_{11}\right) & P\left(\mu_{12}\left(\xi_{12}\right)|\vartheta_{12}\right) & P\left(\mu_{13}\left(\xi_{13}\right)|\vartheta_{13}\right) \\ P\left(\mu_{21}\left(\xi_{21}\right)|\vartheta_{21}\right) & P\left(\mu_{22}\left(\xi_{22}\right)|\vartheta_{22}\right) & P\left(\mu_{23}\left(\xi_{23}\right)|\vartheta_{23}\right) \\ P\left(\mu_{31}\left(\xi_{31}\right)|\vartheta_{31}\right) & P\left(\mu_{32}\left(\xi_{32}\right)|\vartheta_{32}\right) & P\left(\mu_{33}\left(\xi_{33}\right)|\vartheta_{33}\right) \\ \vdots & \vdots & \vdots \\ P\left(\mu_{N1}\left(\xi_{N1}\right)|\vartheta_{N1}\right) & P\left(\mu_{N2}\left(\xi_{N2}\right)|\vartheta_{N2}\right) & P\left(\mu_{N3}\left(\xi_{N3}\right)|\vartheta_{N3}\right) \end{bmatrix} \triangleq [P_{pq}]_{p=1,2,\ldots,N,q=1,2,3}, \tag{2}$$

其中, $\xi_{pq} = [\xi_{pq}^{ij}]$, $p = 1, 2, \ldots, N, q = 1, 2, 3$, 表示第 p 种要素特征下的第 q 类自主安全资源配置方案, μ_{pq} 表示第 p 种要素特征下的第 q 类自主安全资源配置方案的评价函数. 值得注意的是 ξ_{pq}^{ij} 中 i 与 j 的取值如表 5 所示, 其中 $\mathrm{card}\,(\mathcal{M})$ 表示的是集合 $\mathcal{M}$ 的基数.

为达到对人工干预和外部资源依赖程度最低下的最小自主安全资源配置, 采用多域要素特征先验分布下的自主安全资源配置代价的最小期望估计构建无人机系统自主安全优化函数, 如下所示:

$$\begin{aligned} \mathrm{PASS} = \min_{\xi_{pq}} & \sum_{p=1}^{N}\sum_{q=1}^{3}\left\{\left\|\mathbb{E}\left[\mu_{pq}\left(\xi_{pq}\right)|\vartheta_{pq}\right]\right\| + \frac{1}{\varrho}\hbar_{pq}\right\} \\ \text{s.t.} \quad & \cup_{p=1}^{N}\chi_1(\xi_{p1}) \subseteq \mathcal{O}, \\ & \cup_{p=1}^{N}\chi_2(\xi_{p2}) \subseteq \mathcal{A}, \\ & \cup_{p=1}^{N}\chi_3(\xi_{p3}) \subseteq \mathcal{C}, \end{aligned} \tag{3}$$

其中, PASS 表示无人机在执行特定任务时, 对人和环境产生威胁时, 以及无人机系统遭遇物理、信息、智能等多域安全威胁时, 无人机本体所具备的自主与安全保障能力, $\hbar_{pq}$ 为惩罚函数用于强化人工干预和外部资源依赖程度对无人机自主安全资源配置的影响, ϱ 为折扣因子, $\chi_q(\xi_{pq})$ 是量测安全资源占用的函数, 表示对第 p 种情形下安全资源配置方案 ξ_{pq} 中所用到的资源的集合, 是对应可配置自主安全资源集的子集. 以传感器资源为例, $\mathcal{G}_1 = \{g_1^1, g_1^2, g_1^3\}$. 不失一般性, 考虑情形 I 下的一种自主安全资

表 5 自主与安全资源配置后验分布

Table 5 Posteriori distribution of autonomous safety & security resource configuration

Elemental feature sets		Configurable autonomous safety & security resource sets		
		Sensing resources (image, navigation, communication, environmental sensors, etc.)	Computing resources (algorithms, examples, platform, etc.)	Execution resources (dynamic system, attitude system, mission system, etc.)
		$\mathcal{O} \subseteq \mathcal{G}_1$	$\mathcal{A} \subseteq \mathcal{G}_2$	$\mathcal{C} \subseteq \mathcal{G}_3$
Case I	$\mathcal{F}_1 = \{f_1^u \mid u = 1, \ldots, n_1\}$	$P\left(\mu_{11}\left(\xi_{11}\right)\mid\vartheta_{11}\right)$ $\xi_{11} = [\xi_{11}^{ij}]$ $i \leqslant \mathrm{card}\left(\mathcal{F}_1\right)$ $j \leqslant \mathrm{card}\left(\mathcal{O}\right)$	$P\left(\mu_{12}\left(\xi_{12}\right)\mid\vartheta_{12}\right)$ $\xi_{12} = [\xi_{12}^{ij}]$ $i \leqslant \mathrm{card}\left(\mathcal{F}_1\right)$ $j \leqslant \mathrm{card}\left(\mathcal{A}\right)$	$P\left(\mu_{13}\left(\xi_{13}\right)\mid\vartheta_{13}\right)$ $\xi_{13} = [\xi_{13}^{ij}]$ $i \leqslant \mathrm{card}\left(\mathcal{F}_1\right)$ $j \leqslant \mathrm{card}\left(\mathcal{C}\right)$
Case II	$\mathcal{F}_2 = \{f_2^v \mid v = 1, \ldots, n_2\}$	$P\left(\mu_{21}\left(\xi_{21}\right)\mid\vartheta_{21}\right)$ $\xi_{21} = [\xi_{21}^{ij}]$ $i \leqslant \mathrm{card}\left(\mathcal{F}_2\right)$ $j \leqslant \mathrm{card}\left(\mathcal{O}\right)$	$P\left(\mu_{22}\left(\xi_{22}\right)\mid\vartheta_{22}\right)$ $\xi_{22} = [\xi_{22}^{ij}]$ $i \leqslant \mathrm{card}\left(\mathcal{F}_2\right)$ $j \leqslant \mathrm{card}\left(\mathcal{A}\right)$	$P\left(\mu_{23}\left(\xi_{23}\right)\mid\vartheta_{23}\right)$ $\xi_{23} = [\xi_{23}^{ij}]$ $i \leqslant \mathrm{card}\left(\mathcal{F}_2\right)$ $j \leqslant \mathrm{card}\left(\mathcal{C}\right)$
Case III	$\mathcal{F}_3 = \{f_3^w \mid w = 1, \ldots, n_3\}$	$P\left(\mu_{31}\left(\xi_{31}\right)\mid\vartheta_{31}\right)$ $\xi_{31} = [\xi_{31}^{ij}]$ $i \leqslant \mathrm{card}\left(\mathcal{F}_3\right)$ $j \leqslant \mathrm{card}\left(\mathcal{O}\right)$	$P\left(\mu_{32}\left(\xi_{32}\right)\mid\vartheta_{32}\right)$ $\xi_{32} = [\xi_{32}^{ij}]$ $i \leqslant \mathrm{card}\left(\mathcal{F}_3\right)$ $j \leqslant \mathrm{card}\left(\mathcal{A}\right)$	$P\left(\mu_{33}\left(\xi_{33}\right)\mid\vartheta_{33}\right)$ $\xi_{33} = [\xi_{33}^{ij}]$ $i \leqslant \mathrm{card}\left(\mathcal{F}_3\right)$ $j \leqslant \mathrm{card}\left(\mathcal{C}\right)$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
Case N	$\mathcal{F}_N = \{f_N^r \mid r = 1, \ldots, n_N\}$	$P\left(\mu_{N1}\left(\xi_{N1}\right)\mid\vartheta_{N1}\right)$ $\xi_{N1} = [\xi_{N1}^{ij}]$ $i \leqslant \mathrm{card}\left(\mathcal{F}_N\right)$ $j \leqslant \mathrm{card}\left(\mathcal{O}\right)$	$P\left(\mu_{N2}\left(\xi_{N2}\right)\mid\vartheta_{N2}\right)$ $\xi_{N2} = [\xi_{N2}^{ij}]$ $i \leqslant \mathrm{card}\left(\mathcal{F}_N\right)$ $j \leqslant \mathrm{card}\left(\mathcal{A}\right)$	$P\left(\mu_{N3}\left(\xi_{N3}\right)\mid\vartheta_{N3}\right)$ $\xi_{N3} = [\xi_{N3}^{ij}]$ $i \leqslant \mathrm{card}\left(\mathcal{F}_N\right)$ $j \leqslant \mathrm{card}\left(\mathcal{C}\right)$

源配置方案如下:

$$\hat{\xi}_{11} = \begin{bmatrix} \hat{\xi}_{11}^{11} & 0 & 0 \\ 0 & \hat{\xi}_{22} & \hat{\xi}_{23} \\ 0 & 0 & \hat{\xi}_{33} \\ \vdots & \vdots & \vdots \\ \hat{\xi}_{n_1 1} & 0 & \hat{\xi}_{n_1 3} \end{bmatrix}, \tag{4}$$

则安全资源占用的量测函数为

$$\chi_1(\xi_{11}) = \{g_1^1\} \cup \{g_1^2, g_1^3\} \cup \{g_1^3\} \cup \{g_1^1, g_1^3\} = \{g_1^1, g_1^2, g_1^3\}. \tag{5}$$

此例中, 为规避和防御情形 I 中的威胁, 并提升和保障无人机自主水平, 自主安全资源配置方案 (4) 需要使用所有的传感资源, 这显然并不是最优配置方案. 因此, 我们需要求解优化问题 (3), 以期获得最优的资源配置方案 ξ_{pq}^*. 通过对 PASS 取整可得到此最优配置方案所对应的自主安全能力分级.

3.4 自主安全能力分级

在本文所提出的无人机自主安全定义和面向多域威胁自主安全模型下, 无人机系统自主安全是无人机本体以自动的方式持续地执行部分或全部的安全威胁规避与防御任务. 即通过各种机载传感器

表 6 无人机系统自主安全能力分级 [a)]

Table 6 Autonomous safety & security power levels for UAV systems[a)]

等级	自主可控与可恢复性	自主安全描述	自主程度
A1	全程遥控	人在回路中全程遥控飞机, 来规避/防御/对抗无人机所遭遇的物理、信息及智能等安全威胁, 以此避免或减少对无人机本体、人/环境造成威胁和危害.	100% (人工干预 + 外部资源)
A2	局部自主可控与可恢复	无人机本体通过个别预设的避障/防撞路径规划算法或攻击防护技术, 自主规避/防御少量的物理、信息和智能安全威胁, 避免或减少这些对无人机本体、人/环境造成威胁和危害, 人工监控规避/防护算法启动/停止任务.	80% (人工干预 + 外部资源)
A3	半自主可控与可恢复	操作员远程监控、健康监控和有限诊断, 确定物理、信息或智能安全威胁的类型和等级, 设计规避和防御算法以确保无人机安全飞行和任务执行, 避免或减少无人机本体、人/环境损伤. 由无人机本体自主执行预编程算法进行自主规避和防御, 只有在无人机未能正确规避时操作员才接管.	50% (人工干预 + 外部资源)
A4	高度自主可控与可恢复	无人机自动识别、评估所遭遇的物理、信息或智能安全威胁, 做出自主规避、自主防御或博弈对抗决策, 动态调整安全变量和指标, 最大可能避免或减少各类安全威胁对无人机、人/环境造成威胁和危害. 有限外界资源支持安全威胁监测与防御, 以适应不断变化的飞行条件与任务类型.	20% (人工干预 + 外部资源)
A5	完全自主可控与可恢复	无人机预载物理、信息和人工智能威胁自主规避、自主防御或博弈对抗任务目标, 无需人工干预或外界资源即可执行的任务, 无人机本体具备广泛的态势感知 (内部和外部)、威胁预测与评估, 以及机载自主威胁规避、防御、对抗与决策能力, 能够独立高效地保障无人机、人/环境的安全.	5% (人工干预 + 外部资源)

a) 自主可控与可恢复性: 描述风险出现时, 无人机依靠自身具备的能力, 不依赖或很少依赖人工干预和外部资源, 自主诊断、决策和控制, 实现主动危险规避和主动安全防御, 减少/避免或完全恢复无人机本体遭受突发或者恶意的破坏或损毁, 亦确保不对人和环境构成威胁.

物理安全威胁: 静/动态障碍物、电子围栏、电磁干扰等.

信息安全威胁: 通信链路窃听、木马攻击、系统/芯片漏洞、GPS 欺骗、传感器攻击、无线电劫持等.

智能安全威胁: 复杂动态环境认知、多任务规划、多威胁判别与评估、对抗攻击、数据污染/投毒、模型篡改、算法黑箱等.

(如相机、雷达、红外、GPS、惯性传感器等) 来识别无人机所处环境和状态, 并根据所获得的环境和状态信息, 自主分析判断所面临的系统异常或物理、信息、智能等多域威胁, 从而自主地规避和防御威胁, 尽量减少/避免或完全恢复无人机本体遭受突发或者恶意的破坏或损毁, 亦确保不对人和环境构成威胁, 最终实现自主安全.

本文将 3.3 小节所建立的无人机自主安全模型中的要素特征 (威胁的类型、威胁的数量、威胁的危害/损伤程度、无人机自主水平等)、自主可控与可恢复程度 (传感资源、各类计算资源、各类执行资源等机载自主安全资源配置方案)、人工干预和外界资源依赖程度 (人工决策和操作, 地面资源辅助) 作为评测无人机系统自主安全能力的三要素, 通过分析三要素所占比例将无人机自主安全能力分为 A1∼A5 五个级别, 如表 6 所示.

基于 A1∼A5 五个级别的评价, 无人机系统的自主安全能力差异可以直观地从等级划分中体现出来. 需要指出的是 A1 级别自主安全可感知环境, 并提供报警、辅助或短暂介入以辅助操作员 (如航迹偏离预警、紧急制动等应急辅助功能), 但并未主动 “驱动” 无人机, 物理、信息、智能等多域安全威胁

探测与响应完全由人进行操作来完成, 故无自主安全能力.

4 总结与展望

自主安全更契合无人机系统安全理念, 也符合未来智能无人系统的发展趋势. 本文提出了无人机自主安全概念, 并给出初步定义、模型和分级. 但是要从理论内涵到技术关键, 更加准确、完整地开展无人机自主安全系统性研究, 必须要更多、更广泛地结合实际工程, 以真正解决工程实际安全问题关切作为评价标准, 逐步建立科学完善的无人机系统自主安全理论框架和技术体系还有大量的挑战需要面对. 例如如下挑战:

(1) 有效的多域自主安全表征问题. 美军在多域战中首先关注的就是态势表征问题, 因为各域内安全威胁描述的载体、数据组织形式等方面存在很大的差异, 要想在时间、空间、属性、语义等维度得到表示、关联对齐, 以便实现高效挖掘分析是十分困难的. 对无人机系统, 多域安全特征表述首先需要考虑资源、环境、任务、成本的约束, 同时还需结合具体应用解决用户的实际需求. 多源、异构、异步、跨模态、跨语义等数据呈现特征, 由多域威胁导致域间数据共生特性难以挖掘等难题, 是实现多域威胁无人机自主安全研究首先要面对的问题.

(2) 信息不完备、不对称导致安全态势不确定问题. 无人机系统自主性能的不断提升, 对外部资源、人为干预的依赖愈加减少, 这可看作 “主动” 资源受限、电磁拒止行为; 实际战场环境或特殊应用场景下 “被动” 资源受限、电磁拒止将是更加严峻的挑战. 在资源受限、电磁拒止条件下, 甚至加上战场的对抗、博弈情形, 信息不完备、不对称必然造成的无人机系统安全认知的不确定, 系统更多只能依靠自身能力进行安全态势评估, 鲁棒的可信安全评估和长周期可靠的自主推理是十分困难的, 考虑到飞行器的运动体特点, 这又是必须实时面对的.

(3) 无人机系统全状态安全的快速响应与迭代控制. 考虑多域威胁, 必须兼顾全状态安全目标, 比如系统躲过了电子干扰, 但是核心数据可能被窃取, 同样是不能接受的. 传统单域威胁消减方案、防护手段和控制策略很难有效处理多域并发、状态多维度、任务多约束的问题, 当某一域存在威胁, 会导致无人机系统其他域的不稳定甚至崩溃. 如何融合 IPDRR 与 OODA 两者的要素, 解决优化指标和优化策略设计, 既有数学上的挑战, 更要结合实际系统需求精心处理, 以解决对安全威胁的快速响应与全状态优化问题.

参考文献

1 Pan Q, Kang T N, Lyv Y, et al. Development and challenge of UAV sense and avoid system. Unmanned Syst Tech, 2018, 1: 51–61 [潘泉, 康童娜, 吕洋, 等. 无人机感知规避技术发展与挑战. 无人系统技术, 2018, 1: 51–61]

2 Liu L, Wang D, Huang M, et al. A multi-dimensional trustworthy reference framework for network. In: Proceedings of the 2nd International Conference on Instrumentation, Measurement, Computer, Communication and Control, 2012. 1632–1636

3 Shen W, Liu Y, Wu Q, et al. Application of dynamic security technology architecture for advanced directional attacks in power system information security. In: Proceedings of International Conference on Power System Technology (POWERCON), 2018. 3042–3047

4 Du J W, Zhang X, Zhou Y, et al. Active defense security model in the application of network deception system design. Appl Mech Mater, 2013, 347-350: 2860–2864

5 Lu T, Guo X, Zhao L, et al. An analysis of sensitive information system security models. In: Proceedings of the 7th International Conference on Security Technology, 2014. 22–25

6 Middelfart M. Improving business intelligence speed and quality through the OODA concept. In: Proceedings of the ACM 10th International Workshop on Data Warehousing and OLAP, 2007. 97–98

7 Wang Y C, Liu J G. Evaluation methods for the autonomy of unmanned systems. Chin Sci Bull, 2012, 57: 1290–1299 [王越超, 刘金国. 无人系统的自主性评价方法. 科学通报, 2012, 57: 1290–1299]

8 Gao J S, Zou Q Y, Chen X D. Study on the concept of autonomy for UAV. Electronics Optics & Control, 2007, 14: 58–61 [高劲松, 邹庆元, 陈哨东. 无人机自主性概念研究. 电光与控制, 2007, 14: 58–61]

9 US Department of Commerce, NIST. Autonomy levels for unmanned systems (ALFUS) framework volume II: framework models initial version. 2005. https://www.nist.gov/publications/autonomy- levels-unmanned-systems-alfus-frameworkvolume-ii-framework-models-initial

10 Young L, Yetter J, Guynn M. System analysis applied to autonomy: application to high-altitude long-endurance remotely operated aircraft. In: Proceedings of AIAA Infotech@Aerospace Conference, 2005. 7103

11 Lyv Y, Kang T N, Pan Q, et al. UAV sense and avoidance: concepts, technologies, and systems. Sci Sin Inform, 2019, 49: 520–537 [吕洋, 康童娜, 潘泉, 等. 无人机感知与规避: 概念、技术与系统. 中国科学: 信息科学, 2019, 49: 520–537]

12 Lyu Y. UAV Environment Sensing and Collision Avoidance. 2019 [吕洋. 无人机飞行环境感知与障碍规避技术研究. 2019]

13 Radmanesh M, Kumar M, Nemati A, et al. Dynamic optimal UAV trajectory planning in the national airspace system via mixed integer linear programming. Proc Institution Mech Engineers Part G-J Aerospace Eng, 2016, 230: 1668–1682

14 Oliver T, Jonathan L, Mark L, et al. A cloned linguistic decision tree controller for real-time path planning in hostile environments. Fuzzy Sets Syst, 2016, 293: 1–29

15 Wu J, Wang H, Li N, et al. Distributed trajectory optimization for multiple solar-powered UAVs target tracking in urban environment by Adaptive Grasshopper Optimization Algorithm. Aerospace Sci Tech, 2017, 70: 497–510

16 Luo G C, Yu J Q, Mei Y S, et al. UAV path planning in mixed-obstacle environment via artificial potential field method improved by additional control force. Asian J Control, 2015, 17: 1600–1610

17 Hu M F, Ning Q, Chen B C. UAV route planning based on RWPSO and Markov chain. J Harbin Instit Tech, 2019, 51: 75–81 [胡美富, 宁芊, 陈炳才, 等. RWPSO 与马尔科夫链的无人机航路规划. 哈尔滨工业大学学报, 2019, 51: 75–81]

18 Cui S T, Zhao C P, Zhou X Z, et al. Online route planning based on Markov survival model and PSO algorithm. J Sichuan Univ (Nat Sci Ed), 2018, 55: 501–506 [崔舒婷, 赵成萍, 周新志, 等. 基于马尔科夫生存模型与粒子群算法的动态航路规划. 四川大学学报 (自然科学版), 2018, 55: 501–506]

19 Lyu T Z, Zhou W, Zhao C X. Improved visibility graph method using particle swarm optimization and B-spline curve for path planning. J Huaqiao Univ (Nat Sci), 2018, 39: 103–108 [吕太之, 周武, 赵春霞. 采用粒子群优化和 B 样条曲线的改进可视图路径规划算法. 华侨大学学报 (自然科学版), 2018, 39: 103–108]

20 Choset H, Burdick J. Sensor-based exploration: the hierarchical generalized Voronoi graph. Int J Robotics Res, 2000, 19: 96–125

21 Zhu J, Lu Y, Zhang H M. Path replanning for UAV in emergent threats. Comput Eng Appl, 2018, 54: 255–259 [朱杰, 鲁艺, 张辉明. 突发威胁情况下的无人机航迹重规划. 计算机工程与应用, 2018, 54: 255–259]

22 Zhan W W, Wang W, Chen N C, et al. Path planning strategies for UAV based on improved A* algorithm. Geomatics Inform Sci Wuhan Univ, 2015, 40: 315–320 [占伟伟, 王伟, 陈能成, 等. 一种利用改进 A* 算法的无人机航迹规划. 武汉大学学报 (信息科学版), 2015, 40: 315–320]

23 Lin Y, Saripalli S. Sense and avoid for unmanned aerial vehicles using ADS-B. In: Proceedings of 2015 IEEE International Conference on Robotics and Automation (ICRA), 2015. 6402–6407

24 La H M, Lim R, Sheng W. Multirobot cooperative learning for predator avoidance. IEEE Trans Contr Syst Technol, 2015, 23: 52–63

25 Hung S M, Givigi S N. A Q-learning approach to flocking with UAVs in a stochastic environment. IEEE Trans Cybern, 2017, 47: 186–197

26 Long P X, Liu W X, Pan J. Deep-learned collision avoidance policy for distributed multiagent navigation. IEEE Robot Autom Lett, 2017, 2: 656–663

27 Chen Y F, Liu M, Everett M, et al. Decentralized non-communicating multiagent collision avoidance with deep reinforcement learning. In: Proceedings of 2017 IEEE International Conference on Robotics and Automation (ICRA), 2017. 285–292

28 Xu Z, Lyu Y, Pan Q, et al. Multi-vehicle flocking control with deep deterministic policy gradient method. In: Proceedings of 2018 IEEE 14th International Conference on Control and Automation (ICCA), 2018. 306–311

29 Yan C, Xiang X, Wang C. Towards real-time path planning through deep reinforcement learning for a UAV in dynamic environments. J Intell Robot Syst, 2020, 98: 297–309

30 Mnih V, Kavukcuoglu K, Silver D, et al. Human-level control through deep reinforcement learning. Nature, 2015, 518: 529–533

31 Zeitlin A, Lacher A, Kuchar J, et al. Collision avoidance for unmanned aircraft: proving the safety case. The MITRE Corporation and MIT Lincoln Laboratory, USA, 2006. https://www.mitre.org/sites/default/files/pdf/06_1396.pdf

32 Martel F, Schultz R, Wang Z M, et al. Unmanned aircraft systems sense and avoid avionics utilizing ADS-B transceiver. In: Proceedings of AIAA Infotech@ Aerospace Conference, 2009

33 Geiver L J U M. Uavionix'ping2020 ADS-B receives FAA certification. UAS Magazine, 2016. https://uasmagazine.com/articles/1494/uavionixundefined-ping2020-ads-b-receives-faa-certification

34 NASA. UAS traffic management (UTM) project. NASA, 2018. https://www.nasa.gov/utm

35 Guterres M, Jones S, Orrell G, et al. ADS-B surveillance system performance with small UAS at low altitudes. In: Proceedings of AIAA Information Systems-AIAA Infotech@ Aerospace, 2017. 1154

36 Kuchar J E, Drumm A C. The traffic alert and collision avoidance system. Lincoln Laboratory J, 2007, 16: 277

37 Administration F A. Introduction to TAS II version 7.1. brochure. 2011. https://www.faa.gov/documentLibrary/media/AdvisoryCircular/TCAS%20II%20V7.1%20Intro%20booklet.pdf

38 Billingsley T B. Safety analysis of TCAS on Global Hawk using airspace encounter models. 2006. https://dspace.mit.edu/handle/1721.1/35294

39 Portilla E, Fung A, Chen W Z, et al. Sense and avoid (SAA) & traffic alert and collision avoidance system (TCAS) integration for unmanned aerial systems (UAS). In: Proceedings of AIAA Infotech@ Aerospace 2007 Conference and Exhibit, 2007

40 Lin C E, Wu Y Y. TCAS solution for low altitude flights. In: Proceedings of 2010 Integrated Communications, Navigation, and Surveillance Conference Proceedings, 2010

41 Accardo D, Fasano G, Forlenza L, et al. Flight test of a radar-based tracking system for UAS sense and avoid. IEEE Trans Aerosp Electron Syst, 2013, 49: 1139–1160

42 Owen M P, Duffy S M, Edwards M W. Unmanned aircraft sense and avoid radar: surrogate flight testing performance evaluation. In: Proceedings of 2014 IEEE Radar Conference, 2014. 0548–0551

43 Newmeyer L, Wilde D, Nelson B, et al. Efficient processing of phased array radar in sense and avoid application using heterogeneous computing. In: Proceedings of the 26th International Conference on Field Programmable Logic and Applications (FPL), 2016. 1–8

44 Scannapieco A F, Renga A, Fasano G, et al. Ultralight radar sensor for autonomous operations by micro-UAS. In: Proceedings of 2016 International Conference on Unmanned Aircraft Systems (ICUAS), 2016. 727–735

45 Scannapieco A F, Renga A, Fasano G, et al. Experimental analysis of radar odometry by commercial ultralight radar sensor for miniaturized UAS. J Intell Robot Syst, 2018, 90: 485–503

46 Wilson M. Ground-based sense and avoid support for unmanned aircraft systems. In: Proceedings of Congress of the International Council of the Aeronautical Sciences (ICAS), 2012

47 Sahawneh L R, Wikle J K, Kaleo R A, et al. Ground-based sense-and-avoid system for small unmanned aircraft. J Aerospace Inf Syst, 2018, 15: 501–517

48 Meer I A, Ozger M, Lundmark M, et al. Ground based sense and avoid system for air traffic management. In: Proceedings of 2019 IEEE 30th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC), 2019. 1–6

49 Fasano G, Accardo D, Tirri A E, et al. Radar/electro-optical data fusion for non-cooperative UAS sense and avoid. Aerospace Sci Tech, 2015, 46: 436–450

50 Chen G, Dong W, Sheng X, et al. An active sense and avoid system for flying robots in dynamic environments. IEEE ASME Trans Mechatron, 2021, 26: 668–678

51 Alvarez H, Paz L M, Sturm J, et al. Collision avoidance for quadrotors with a monocular camera. In: Proceedings of Experimental Robotics, 2015. 195–209

52 Lyu Y, Pan Q, Zhao C, et al. Vision-based UAV collision avoidance with 2D dynamic safety envelope. IEEE Aerosp

Electron Syst Mag, 2016, 31: 16–26
53 Oleynikova H, Honegger D, Pollefeys M. Reactive avoidance using embedded stereo vision for MAV flight. In: Proceedings of IEEE International Conference on Robotics and Automation (ICRA), 2015. 50–56
54 Hu J, Niu Y F, Wang Z C. Obstacle avoidance methods for rotor UAVs using realsense camera. In: Proceedings of Chinese Automation Congress (CAC), 2017. 7151–7155
55 Grzonka S, Grisetti G, Burgard W. A fully autonomous indoor quadrotor. IEEE Trans Robot, 2012, 28: 90–100
56 Gageik N, Benz P, Montenegro S. Obstacle detection and collision avoidance for a UAV with complementary low-cost sensors. IEEE Access, 2015, 3: 599–609
57 Lee Y S, Kang Y J, Lee S G, et al. An overview of unmanned aerial vehicle: cyber security perspective. In: Proceedings of IT Convergence Technology, 2016
58 Shachtman N. Exclusive: computer virus hits U.S. drone fleet. Wired Magazine, 2011. https://www.wired.com/2011/10/virus-hits-drone-fleet/
59 Wang Z X, Li Y, Lyu Y, et al. Frontier technology development trend of UAS cyber security. Software Guide, 2021, 20: 7–12 [王兆轩, 李扬, 吕洋, 等. 无人机系统信息安全前沿技术发展趋势. 软件导刊, 2021, 20: 7–12]
60 Trippel T, Weisse O, Xu W Y, et al. WALNUT: waging doubt on the integrity of MEMS accelerometers with acoustic injection attacks. In: Proceedings of 2017 IEEE European Symposium on Security and Privacy (EuroS&P), 2017. 3–18
61 Jeong J, Kim D, Jang J H, et al. Un-rocking drones: foundations of acoustic injection attacks and recovery thereof. In: Proceedings of Network and Distributed System Security Symposium, 2023
62 Tu Y Z, Lin Z Q, Lee I, et al. Injected and delivered: fabricating implicit control over actuation systems by spoofing inertial sensors. In: Proceedings of the 27th USENIX Security Symposium, 2018
63 Fu K, Xu W. Risks of trusting the physics of sensors. Commun ACM, 2018, 61: 20–23
64 Rana M M. IoT-based electric vehicle state estimation and control algorithms under cyber attacks. IEEE Internet Things J, 2020, 7: 874–881
65 Ju Z, Zhang H, Tan Y. Deception attack detection and estimation for a local vehicle in vehicle platooning based on a modified UFIR estimator. IEEE Internet Things J, 2020, 7: 3693–3705
66 Yang T, Lv C. A secure sensor fusion framework for connected and automated vehicles under sensor attacks. IEEE Internet Things J, 2022, 9: 22357–22365
67 Shen J J, Won J Y, Chen Z Y, et al. Drift with devil: security of multi-sensor fusion based localization in high-level autonomous driving under GPS spoofing. In: Proceedings of the 29th USENIX Security Symposium (USENIX Security 20), 2020. 931–948
68 Kang R, Xiong L, Xu M Y, et al. VINS-vehicle: a tightly-coupled vehicle dynamics extension to visual-inertial state estimator. In: Proceedings of 2019 IEEE Intelligent Transportation Systems Conference (ITSC), 2019. 3593–3600
69 Li J, Zhao J Q, Kang Y C, et al. DL-SLAM: direct 2.5D LIDAR SLAM for autonomous driving. In: Proceedings of 2019 IEEE Intelligent Vehicles Symposium (IV), 2019. 1205–1210
70 Arafin M T, Kornegay K. Attack detection and countermeasures for autonomous navigation. In: Proceedings of the 55th Annual Conference on Information Sciences and Systems (CISS), 2021. 1–6
71 Luo A. Drones hijacking. In: Proceedings of DEF CON, 2016
72 Hooper M, Tian Y F, Zhou R X, et al. Securing commercial WIFI-based UAVs from common security attacks. In: Proceedings of MILCOM 2016-2016 IEEE Military Communications Conference, 2016. 1213–1218
73 Alhawi O M, Mustafa M A, Cordeiro L C. Finding security vulnerabilities in unmanned aerial vehicles using software verification. In: Proceedings of International Workshop on Secure Internet of Things (SIOT), 2019. 1–9
74 Clements A A, Almakhdhub N S, Saab K S, et al. Protecting bare-metal embedded systems with privilege overlays. In: Proceedings of 2017 IEEE Symposium on Security and Privacy (SP), 2017. 289–303
75 Koo K, Lee W Y, Cho S R, et al. A secure operating system architecture based on linux against communication offense with root exploit for unmanned aerial vehicles. J Inform Process Syst, 2020, 16: 42–48
76 Fotouhi A, Qiang H, Ding M, et al. Survey on UAV cellular communications: practical aspects, standardization advancements, regulation, and security challenges. IEEE Commun Surv Tut, 2019, 21: 3417–3442
77 Sampigethaya K. Aircraft cyber security risk assessment: bringing air traffic control and cyber-physical security to the forefront. In: Proceedings of AIAA Scitech 2019 Forum, 2019

78 Robinson M, Mitchell A. Knocking my neighbors kids cruddy drone offline. In: Proceedings of DEF CON 23, 2015
79 Krishna C G L, Murphy R R. A review on cybersecurity vulnerabilities for unmanned aerial vehicles. In: Proceedings of 2017 IEEE International Symposium on Safety, Security and Rescue Robotics (SSRR), 2017
80 Westerlund O, Asif R. Drone hacking with Raspberry-Pi 3 and WiFi pineapple: security and privacy threats for the Internet-of-Things. In: Proceedings of the 1st International Conference on Unmanned Vehicle Systems-Oman (UVS), 2019
81 Pasquazzo A D. Hacking drones with dronesploit. In: Proceedings of Black Hat, 2019
82 Miller C, Valasek C. A survey of remote automotive attack surfaces. Black Hat USA, 2014, 2014: 94
83 Miller C, Valasek C. Remote exploitation of an unaltered passenger vehicle. Black Hat USA, 2015, 2015: 1–91
84 Murvay P S, Groza B. Security shortcomings and countermeasures for the SAE J1939 commercial vehicle bus protocol. IEEE Trans Veh Technol, 2018, 67: 4325–4339
85 Cho K T, Shin K G. Error handling of in-vehicle networks makes them vulnerable. In: Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, 2016. 1044–1055
86 Allouch A, Cheikhrouhou O, Koubaa A, et al. MAVSec: securing the MAVLink protocol for ardupilot/PX4 unmanned aerial systems. In: Proceedings of 2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC), 2019. 621–628
87 Abdallah A, Ali M Z, Mišić J, et al. Efficient security scheme for disaster surveillance UAV communication networks. Information, 2019, 10: 43
88 Zhang G C, Wu Q Q, Cui M, et al. Securing UAV communications via trajectory optimization. In: Proceedings of Globecom IEEE Global Communications Conference, 2017. 1–6
89 Zhang G, Wu Q, Cui M, et al. Securing UAV communications via joint trajectory and power control. IEEE Trans Wireless Commun, 2019, 18: 1376–1389
90 Cui M, Zhang G, Wu Q, et al. Robust trajectory and transmit power design for secure UAV communications. IEEE Trans Veh Technol, 2018, 67: 9042–9046
91 Papernot N, McDaniel P, Jha S, et al. The limitations of deep learning in adversarial settings. In: Proceedings of 2016 IEEE European Symposium on Security and Privacy (EuroS&P), 2016. 372–387
92 Cao Y, Xiao C, Cyr B, et al. Adversarial sensor attack on LiDAR-based perception in autonomous driving. In: Proceedings of the ACM SIGSAC Conference on Computer and Communications Security, 2019. 2267–2281
93 Cao Y, Wang N, Xiao C, et al. Invisible for both camera and lidar: Security of multi-sensor fusion based perception in autonomous driving under physical-world attacks. In: Proceedings of IEEE Symposium on Security and Privacy (SP), 2021. 176–194
94 Davidson D, Wu H, Jellinek R, et al. Controlling UAVs with sensor input spoofing attacks. In: Proceedings of the 10th USENIX Conference on Offensive Technologies, 2016. 221–231
95 Kim K, Nalluri S, Kashinath A, et al. Security analysis against spoofing attacks for distributed UAVs. In: Proceedings of Workshop on Decentralized IoT Systems and Security, 2020
96 Habler E, Shabtai A. Using LSTM encoder-decoder algorithm for detecting anomalous ADS-B messages. Comput Secur, 2018, 78: 155–173
97 Panice G, Luongo S, Gigante G, et al. A SVM-based detection approach for GPS spoofing attacks to UAV. In: Proceedings of 2017 23rd International Conference on Automation and Computing (ICAC), 2017. 1–11
98 Sun Y, Cao C J, Lai J X, et al. AntiGPS spoofing method for UAV based on LSTM-KF model. Chinese J Network Inform Secur, 2020, 6: 80–88 [孙扬, 曹春杰, 赖俊晓, 等. 基于 LSTM-KF 模型的无人机抗 GPS 欺骗方法. 网络与信息安全学报, 2020, 6: 80–88]
99 Abbaspour A, Yen K K, Noei S, et al. Detection of fault data injection attack on UAV using adaptive neural network. Procedia Comput Sci, 2016, 95: 193–200
100 Shoufan A. Continuous authentication of UAV flight command data using behaviometrics. In: Proceedings of 2017 IFIP/IEEE International Conference on Very Large Scale Integration (VLSI-SoC), 2017. 1–6
101 Rani C, Modares H, Sriram R, et al. Security of unmanned aerial vehicle systems against cyber-physical attacks. J Defense Modeling Simul, 2016, 13: 331–342
102 Johansson R, Hammar P, Thorén P. On simulation-based adaptive UAS behavior during jamming. In: Proceedings of 2017 Workshop on Research, Education and Development of Unmanned Aerial Systems (RED-UAS), 2017. 78–83

103 Lu H, Li Y, Mu S, et al. Motor anomaly detection for unmanned aerial vehicles using reinforcement learning. IEEE Internet Things J, 2018, 5: 2315–2322
104 Lin Z H, Lu X Z, Dai C H, et al. Reinforcement learning based UAV trajectory and power control against jamming. In: Proceedings of International Conference on Machine Learning for Cyber Security, 2019. 336–347
105 Yang B, Liu M. Attack-resilient connectivity game for UAV networks using generative adversarial learning. In: Proceedings of the 18th International Conference on Autonomous Agents and MultiAgent Systems, 2019. 1743–1751
106 Xiao L, Xie C, Min M, et al. User-centric view of unmanned aerial vehicle transmission against smart attacks. IEEE Trans Veh Technol, 2018, 67: 3420–3430
107 Lis P, Mendel J. Cyberattacks on critical infrastructure: an economic perspective. Econ Bus Rev, 2019, 5: 24–47

Autonomous safety and security of UAV systems: definition, modeling, and gradation

Quan PAN*, Yaning GUO, Yang LYU, Yang LI & Zheng TAN

School of Automation, Northwestern Polytechnical University, Xi'an 710027*, China*
* Corresponding author. E-mail: quanpan@nwpu.edu.cn

Abstract The balance between autonomy and safety is the strategic demand and key for the rapid development of unmanned aerial vehicle (UAV) technology and industry. In the face of "continuously improving unmanned autonomous flight capability" and "increasingly severe physical, information, intelligence, and other multidomain security threats," the traditional security architecture with fit-to-fly and sense-and-avoid technologies as the core cannot meet the safety and security requirements of current and future UAVs only by referring to the manned aerial vehicle safety system. It is necessary and urgent to construct a new UAV autonomous safety and security architecture in multiple domains. Starting from the inherent essential attribute of autonomy in UAVs, the autonomous safety and security architecture of UAVs is studied in this paper, concentrating on the definition and characterization of the new concept "autonomous safety and security of UAVs," constructing a new autonomous safety and security model facing multidomain security threats and proposing the gradation of the autonomous safety and security power. This study would provide basic support for the rapid development of UAV industry and a reference for establishing security systems for other unmanned systems.

Keywords unmanned aerial system, multi-domain threat, autonomous safety and security, resource configuration