中国科学:信息科学 2026 年 第 56 卷 第 6 期: 1420–1436
SCIENTIA SINICA Informationis
论文

《中国科学》杂志社
SCIENCE CHINA PRESS
CrossMark click for updates

航天器轨道威胁博弈中的反馈机制

袁利[1,2], 李婵颖[3,4*], 魏衍[5], 姜甜甜[2,6], 黄盘兴[2,6], 魏春岭[2,6]

1. 中国空间技术研究院, 北京 100094
2. 空间智能控制技术全国重点实验室, 北京 100094
3. 数学科学全国重点实验室, 中国科学院数学与系统科学研究院, 北京 100190
4. 中国科学院大学数学科学学院, 北京 100049
5. 四川大学数学学院, 成都 610065
6. 北京控制工程研究所, 北京 100094
* 通信作者. E-mail: cyli@amss.ac.cn

收稿日期: 2025–04–27; 修回日期: 2025–08–20; 接受日期: 2025–09–12; 网络出版日期: 2026–05–12

国家自然科学基金 (批准号: U21B6001, 11925109) 资助项目

摘要 航天器在轨运行面临碎片碰撞、抵近侦察、恶意袭扰等轨道威胁, 其自主规避是面向不确定威胁行为的动态博弈过程. 受空间环境、感知限制、通信干扰、隐蔽技术等因素影响, 其信息结构不完美且不完全, 使博弈决策陷入双重困境 —— 威胁方不确定性大, 态势变化对各方策略敏感; 策略搜索空间维度随信息复杂度呈指数级增长. 因而在星载计算资源严重受限下, 航天器自主决策系统易生成保守策略或难以确保航天器绝对安全. 为此, 本文引入具有 “垂直分层” 与 “水平调整” 结构的反馈策略设计机制, 在决策过程中基于态势信息在线优化博弈信息结构, 自适应约减搜索空间, 提升策略性能, 以保障航天资产的绝对安全以及满足在轨任务的实际需求.

关键词 轨道博弈, 不完美信息, 不完全信息, 自主决策, 反馈机制

1 引言

航天器在轨运行面临多种形态的轨道威胁, 包括空间碎片/失效卫星碰撞、非合作航天器抵近侦察、恶意袭扰等[1,2], 给航天器的安全和业务连续造成了严重影响. 特别是, 近年来随着空间碎片累积、巨型星座部署, 以及太空竞争加剧, 空间安全态势日益严峻, 在轨危机事件频繁发生, 轨道威胁处置从偶发性事件向常态化转变. 例如, 2019 年欧空局 (European Space Agency, ESA) 为避免 “风神” (Aeolus) 气象卫星与 “星链 -44” (Starlink-44) 卫星的碰撞, 实施了规避机动操作[3]. 2021 年, “一网 -0178” (OneWeb-0178) 卫星在升轨过程中被监测到与 “星链 -1546” (Starlink-1546) 卫星的碰撞概率高达 1.3%, 之后美国太空探索技术公司 (SpaceX) 关闭了 Starlink-1546 卫星的自动碰撞规避系统, 一网卫星在地面指挥控制下完成了规避机动[4]. 同年, 中国空间站因 “星链” 卫星的异常抵近, 被迫两度实施紧急避碰机动[5]. 据统计, 截至 2024 年 11 月, 国际空间站 (International Space Station, ISS) 已累计执行了 39 次碰撞规避机动; 中国空间站自 2021 年入轨以来, 已执行 20 多次应急变轨[1], 消耗大量燃料的同时, 对航天员的在轨安全构成了严峻挑战. 据 SpaceX 提交的《半年度星座状态报告 (SpaceX Starlink

引用格式: 袁利, 李婵颖, 魏衍, 等. 航天器轨道威胁博弈中的反馈机制. 中国科学: 信息科学, 2026, 56: 1420–1436, doi: 10.1360/SSI-2025-0164
Yuan L, Li C Y, Wei Y, et al. Feedback mechanism in spacecraft orbital threat games. Sci Sin Inform, 2026, 56: 1420–1436, doi: 10.1360/SSI-2025-0164

© 2026《中国科学》杂志社 www.scichina.com infocn.scichina.com
https://www.sciengine.com/doi/10.1360/SSI-2025-0164

satellites made 50000 collision-avoidance maneuvers in the past 6 months)》显示, 在 2023 年 12 月至 2024 年 5 月期间, 每颗星链卫星平均执行 14 次碰撞规避机动, 累计执行近 5 万次避碰机动操作, 这一数字约为前六个月的两倍 [6].

与此同时, 随着空间技术向实战化方向演进, 轨道空间逐步演变为新型战略博弈高地. 在轨监视、抵近干扰等新形态的轨道威胁不断涌现, 成为影响空间资产安全的重要因素. 美国 “地球同步轨道空间态势感知计划” (Geosynchronous space situational awareness program, GSSAP) 卫星、空间态势感知实验卫星 Mycroft 等已形成在轨监视和抵近侦察能力 [7,8]. 2014 年以来, GSSAP 卫星累计执行数百次机动操作, 以抵近侦察的方式对中国多颗商业卫星实施干扰 [9]. 美国 “任务延寿飞行器 -1/-2” (Mission Extension Vehicle-1/-2, MEV-1/-2) 分别于 2020 年和 2021 年完成了对失效商业卫星的捕获与轨道服务技术演示验证 [10], 进一步揭示了空间资产面临的非对称威胁形态.

面对多种形态的轨道威胁, 当前主要采取基于 “星地大回路” 的处置方式. 这种方式响应时效性差、运控压力大, 且受测量误差等因素的影响, 处置的时机和力度往往难以精确把握, 严重影响航天器在轨运行的安全性和业务连续性, 难以适应威胁数量加速增长、环境加剧恶化的空间态势. 因此, 航天器轨道威胁自主规避逐渐受到世界范围内的广泛关注 [11,12]. 目前, ESA 的碰撞风险评估和自动缓解 (Collision risk estimation and automated mitigation, CREAM) 项目 [11], 以及 “星链” 卫星上部署的自动防撞软件 [12], 已具备地面支持下的碰撞风险评估和规避机动策略的自主生成能力. 美国太空军计划联合火箭实验室和真近点角公司于 2025∼2027 年间执行战术响应太空 (TacRS) 项目下的 “Victus Haze” 演习, 研制卫星具备高在轨机动能力, 旨在演示验证美太空军对潜在 “在轨侵略” 事件的应急处理能力 [13]. 在空间安全成为国际战略竞争新焦点的背景下, 未来需充分利用航天器的近场感知能力, 构建 “感知 – 决策 – 执行” 一体化星上闭环控制系统 [2], 支撑轨道威胁评估、规避策略制定及动作执行的星上自主完成, 进而优化威胁处置的时机和力度, 确保航天器以最小的代价完成对轨道威胁的适度有效应对. 而航天器自主规避决策则是一体化控制架构的核心智能模块.

航天器轨道威胁自主规避决策可纳入轨道博弈研究范畴, 指发生在轨道上由于双方或者多方航天器之间的任务目标不同甚至冲突而导致的博弈问题. 轨道博弈参与者受到轨道动力学/运动学约束, 在允许的控制能力和可获取的信息支持下, 自主生成博弈策略、机动动作规划并完成控制执行, 实现生存概率、燃料节省以及任务完成度等综合指标的最优. 而微分博弈理论是研究此类具备连续状态与决策空间博弈问题的重要工具. 作为现代控制论的重要组成部分, 微分博弈起源于 20 世纪 50 年代至 60 年代 [14]. 从 20 世纪 50 年代中期起, 美苏等大国将研究重心转移至航天, 面临一系列带目标的控制问题有待解决, 诸如将火箭、卫星等航天器用最少燃料或最短时间准确发射至预定轨道等. 由此建立起最优控制理论, 其大体分为庞特里亚金极大值原理、贝尔曼动态规划方法和线性二次最优控制理论三大类 [15]. 其后, 人们将最优控制与博弈论结合, 至此形成微分博弈论 [16,17].

轨道博弈作为近年来航天领域的研究热点备受关注 [18], 其中单对单航天器轨道追逃场景研究最为广泛, 一般被建模为具有完全且完美信息的零和微分博弈进行研究, 此时双方意图、能力与动力学彼此透明, 且追逃过程中状态信息与历史策略皆动态可观. 部分学者在给定拦截条件下以拦截时间作为核心博弈指标 [19], 并考虑各类工程约束, 如离轨距离约束 [20]、推力方向约束 [21]、燃料消耗约束 [22,23] 等. 此类建模可转化为两点边值问题, 但高维动力学方程与多约束耦合仍导致其解析解难以获取. 因此, 一系列数值解搜索算法被相继提出, 包括混合全局 – 局部优化, 半直接配点与非线性规划相结合的算法及其变体, 以及基于轨道力学特性的降维方法等 [24∼26]. 然而上述数值方法均需对高维终端协态向量进行随机搜索, 其计算复杂度随维度呈指数增长, 难以满足卫星自主规划任务中的时间约束, 严重制约了空间对抗决策的时效性. 考虑到轨道博弈过程中双方实时相对状态及燃料消耗的重要性, 线性二次微分博弈因其对线性动力学及过程二次型指标的兼容性成为轨道追逃博弈的另一理想模型. 此类博弈因其解析可解性而具有显著的计算优势, 对应鞍点策略可通过求解时变黎卡提代数/微分方程显式获得 [27]. 但由于支付矩阵需满足正定约束以保证黎卡提矩阵代数/微分方程解的存在唯一性, 对规避方而言, 该型指标可能不能真实反应危险场景下的生存需求. 此外, 最大推力约束下, 直接引入饱和函数修正策略也会严重破坏其鞍点的最优性 [28]. 现有研究还拓展至多航天器追逃博弈, 借助可达域分析、强化学习、模糊评估等工具考虑了追 – 逃 – 防三方动态博弈 [29,30] 以及多追踪者协同围捕问题 [31∼33]. 上述建模和求解均基于完全

https://www.sciengine.com/doi/10.1360/SSI-2025-0164

完美信息假设, 但实际对抗却不可避免面临信息的缺失与模糊.

真实轨道博弈中追逃方均难以事先获知彼此的策略意图, 因此双方很难在博弈指标上达成共识, 即为不完全信息. 而空间非合作目标的精确测量难题则导致轨道博弈往往呈现不完美信息特性. 两者在博弈中引入显著不确定性. 不难理解, 不确定性的处理是博弈实现高效决策的关键基础, 目前主流思路是对具有已知结构的指标泛函 (主要为二次型) 进行参数估计, 但这依赖历史信息的获取. 即, 信息是否完美至关重要. 文献 [34] 考虑了一类信息不完全但完美的线性二次微分博弈, 假设对手方二次型指标函数中的矩阵权重未知, 提出了基于强跟踪无迹卡尔曼滤波的估计方法, 并给出规避策略. 同样在指标为线性二次型的轨道追逃博弈中, 文献 [35,36] 基于交互式多模态方法与滤波估计, 在不完全信息条件下对对手模态进行推测, 并融合概率信息以生成其近似策略, 以此求出一种高效拦截策略. 文献 [37] 在线性二次微分博弈框架下, 基于广义卡尔曼滤波实现对对手策略的实时估计与相应最优控制, 并证明策略满足 ϵ 纳什均衡. 文献 [38] 将追逃分为远距抵近和近距拦截两个阶段, 提出基于深度确定性策略梯度算法的两段式拦截策略 (远距阶段脉冲控制、近距阶段连续推力), 实现对目标航天器的拦截. 至于双方博弈指标彼此完全未知的场景, 则导致无穷维信息量的估计, 即使完美信息下亦难以实现. 解决方案为对手意图或策略建模. 例如, 文献 [39, 40] 通过把对手策略建模为干扰噪声, 将完美但不完全信息博弈转化为不完美但完全信息微分博弈进行研究. 除经典轨道追逃博弈外, 轨道博弈还包括轨道侦察与反侦察 [41~43]、轨道伴飞与尾随监视 [44, 45]、通信中继与干扰对抗 [46, 47] 等, 其场景皆具极大不确定性, 均有推断对手意图与策略的需求.

值得一提的是, 强对抗环境下, 对对手策略的错误或无效建模可能导致灾难性后果. 这主要是由该类博弈的复杂性导致. 航天器涉及重要的太空资产, 博弈中需采取纯策略以避免混合策略引入的随机性, 否则必然带来一定概率的损失. 而在二次型以外的指标中, 纯策略的纳什均衡存在性极难保证. 即使纳什均衡存在, 而非零和博弈可能具有不同均衡解, 对应的各方收益往往不等且不存在偏序, 以致各方在均衡策略间无从选择 (显然, 此时只有知道对手的选择, 才能做出正确决策). 在没有可容许 (admissible) 纳什均衡的可靠保障下, 强对抗博弈对对手的策略往往极其敏感. 同一策略面对对手的不同策略, 可能出现成、败两种极端结果. 但不完全不完美信息结构为对手策略的建模与估计造成极大阻碍, 若不能解决该难点, 基于此而设计的博弈策略将存在严重的安全风险. 信息不完备带来的另一个难点是策略解搜索空间庞大. 航天器威胁规避是动态演化博弈, 复杂的非线性演化结合动态不确定性与各类约束, 甚至导致策略空间结构的非常规性, 难以快速优化求解, 只能进行大规模搜索, 这对计算资源和存储空间等提出极高要求, 与在轨资源严重受限的现状产生矛盾.

航天器轨道威胁自主规避是面向不确定威胁行为的动态博弈过程. 事实上, 由于时空尺度大、空间环境复杂, 轨道威胁精准感知具有相当困难性, 致使威胁目标运动/形态测量信息不完备, 碰撞、抵近、绕飞等轨道威胁行为特征不明显等. 各类信息模糊性造成航天器在轨博弈信息结构不完全且不完美. 此外, 航天器轨道速度大, 横向机动能力有限, 规避机动燃料消耗严重影响在轨寿命, 通信、导航等任务对航天器业务连续性要求苛刻. 因此, 如果考虑到航天器博弈的真实场景与诸多任务约束, 当下主流零和博弈与二次型泛函指标的设计方案则可能难以胜任. 但更精细的博弈建模将直面上文所述的困境, 即纳什均衡的缺失以及策略空间的搜索困难. 目前博弈决策主要面向特定场景和相对确定的任务模式, 虽然具有一定的泛化能力和鲁棒性, 但在应对未知动态场景和不确定威胁行为时依然面临挑战. 众所周知, 反馈是应对不确定性的有力工具. 为此, 本文在不完全不完美信息下的动态博弈框架中引入基于态势的反馈机制, 以信息结构实时优化为核心功能, 以在线决策自适应约减为关键技术, 实现航天器轨道威胁规避的适应性 “动态均衡” 高效求解.

2　不完美不完全信息结构下的轨道博弈反馈机制设计

本文考虑航天器面向伴飞、绕飞、跟踪及风险交会等威胁行为时的自主规避. 当威胁方具备一定的策略选择, 本问题可以数学建模为轨道博弈. 博弈论以数学化的方式揭示了决策行为的深层逻辑, 用以理解竞争与合作现象, 构建策略互动的威胁或协作方式. 当威胁方仅具备单一策略, 则博弈求解可退化为最优控制设计, 如空间碎片碰撞、侦察卫星自然绕飞场景. 故本文将基于博弈论的统一框架研究轨道威胁 (无机动威胁与机动威胁) 的

https://www.sciengine.com/doi/10.1360/SSI-2025-0164

自主规避策略设计. 由于感知技术限制以及复杂空间环境 (包括电子对抗) 等因素, 态势模糊性势所难免, 博弈信息结构不完美/不完全, 导致可能的策略搜索空间过大, 甚至 “最优策略” 不存在现象, 这为航天器自主威胁规避决策设计带来极大难度. 下文假设航天器在既定任务轨道上运行时遭遇其他航天器的抵近威胁, 为以极小代价提升对抗中的生存能力和任务连续性, 我们在不完全不完美信息下的动态博弈框架中引入反馈机制, 实现航天器在线决策的自适应约减及性能提升.

2.1　航天器动态博弈建模

航天器动态博弈具有自身特色, 如轨道约束强, 航天器横向机动能力弱; 燃料珍贵, 轨道规避机动代价高昂; 航天器任务连续性不可忽视; 星上计算资源严重受限等. 传统博弈论中常用以策略求解的极大极小值原理或 Hamilton-Jacobi-Isaacs 方程此时无法直接应用, 甚至大部分真实博弈不存在纯策略纳什均衡. 为进一步阐明本问题, 我们首先引入严格的数学描述.

2.1.1　航天器动态博弈基本要素

(1) 参与者: 参与者为我方航天器与威胁方 N 个航天器, 可用非负整数上的子集合表示

$$\mathscr{N}_0 = \{0, 1, 2, ..., N\}.$$

在本文设置下, 0 代表规避方, $1, ..., N$ 代表威胁方.

(2) 演化方程: 航天器动态博弈演化通常基于相关动力学/运动学方程, 可以由如下微分方程组刻画:

$$\begin{cases} \dot{\boldsymbol{x}}_0 = \boldsymbol{f}_0(\boldsymbol{x}_0, \boldsymbol{u}_0, t), \\ \dot{\boldsymbol{x}}_1 = \boldsymbol{f}_1(\boldsymbol{x}_1, \boldsymbol{u}_1, t), \\ \vdots \\ \dot{\boldsymbol{x}}_N = \boldsymbol{f}_N(\boldsymbol{x}_N, \boldsymbol{u}_N, t), \end{cases} \tag{1}$$

其中下标 $i \in \mathscr{N}_0$ 对应航天器 i, 博弈时间 $t \in [0, t_f]$, $t_f \in [0, \infty)$ 为终端时刻. 航天器 i 状态变量为 $\boldsymbol{x}_i$, 轨道与姿态控制变量为 $\boldsymbol{u}_i$, 时变函数 $\boldsymbol{f}_i$ 则根据具体场景考虑地球引力场、大气密度等各种因素的影响.

注释1　尽管航天器实际动态博弈演化规律由姿轨运动学/动力学方程决定, 但为节省计算资源, 在行为决策层 (见 2.2.1 小节), 可结合态势特征与演化方程 (1), 将动态博弈粗粒化为有限策略式博弈 (如矩阵博弈)、有限扩展式博弈 (如博弈树) 或者两者结合体, 并指定行动顺序. 其后, 在动作规划层 (见 2.2.1 小节) 考虑演化方程 (1). 此时两层演化模型为同时包含连续时间动态与离散事件动态的混杂系统.

(3) 策略集: 航天器 $i \in \mathscr{N}_0$ 的策略是从信息集到行动集的映射, 其中信息集为该航天器无法区分的决策时间节点信息 (如所有航天器的历史状态、行动子集). 对航天器而言, 其策略受到轨道动力学以及底层控制限制, 其容许控制函数类是由有界函数组成的, 即

$$\mathcal{U}_i([0, t_f]) = \left\{ \boldsymbol{u}_i(\cdot) : [0, t_f] \to \mathbf{U}_i | \mathbf{U}_i \text{ 是有界集} \right\}.$$

航天器轨道控制形式可包括脉冲推力和有限连续推力, 姿态控制一般为姿态力矩设计.

(4) 目标集: 存在容许控制 $\boldsymbol{u}_i \in \mathcal{U}_i([0, t_f]), i \in \mathscr{N}_0$, 使系统 (1) 的初始状态在 $\boldsymbol{u}_i$ 作用下于 t_f 时刻到达给定的终端状态约束集 (目标集)

$$X_i = \{\boldsymbol{x}_i \in \mathbb{R}^{n_i} | \boldsymbol{g}_i(\boldsymbol{x}_i(t_f), t_f) \leqslant 0\}.$$

博弈终端时间 t_f 可以事先给定, 也可以是待优化的自由变量.

https://www.sciengine.com/doi/10.1360/SSI-2025-0164

(5) 收益函数: 航天器 $i, i \in \mathcal{N}_0$ 的收益函数表示在策略组合下的效用. 为刻画各方意图和任务, 引入描述控制过程或结果品质的有界指标泛函[1)]

$$J_i(\boldsymbol{u}_0, \boldsymbol{u}_1, ..., \boldsymbol{u}_N) = \varphi_i(\boldsymbol{x}_0, \boldsymbol{x}_1, ..., \boldsymbol{x}_N, t_f) + \int_0^{t_f} L_i(\boldsymbol{x}_0, \boldsymbol{x}_1, ..., \boldsymbol{x}_N, \boldsymbol{u}_0, \boldsymbol{u}_1, ..., \boldsymbol{u}_N, t)\mathrm{d}t. \tag{2}$$

(6) 信息结构: 轨道动态博弈中的信息结构决定了航天器在不同时间决策节点所掌握的信息类型和程度, 直接影响策略选择和均衡结果. 一般分为如下两大类:

完美信息与不完美信息: 若每个航天器 $i \in \mathcal{N}_0$ 在任意决策时间节点能够获取整个对策空间的完整历史状态信息, 则称该博弈具有完美信息, 否则为不完美信息博弈. 由于敏感器观测误差或感知设备物理限制造成的信息缺失, 航天器轨道博弈过程通常为不完美信息结构, 根据其反馈形式具体还可包括开环、闭环、无记忆、时滞等不完美信息结构.

完全信息与不完全信息: 若每个航天器 $i \in \mathcal{N}_0$ 均已知其余所有航天器的指标泛函、策略空间及对抗能力等信息, 则称该博弈具有完全信息, 否则为不完全信息博弈.

2.1.2　航天器轨道博弈纳什均衡

针对航天器轨道动态博弈问题, 如果存在一组控制策略 $(\boldsymbol{u}_0^*, \boldsymbol{u}_1^*, \ldots, \boldsymbol{u}_i^*, \ldots, \boldsymbol{u}_N^*)$, 其中 $\boldsymbol{u}_i^*(\cdot) \in \mathcal{U}_i([0, t_f])$, 满足

$$J_i(\boldsymbol{u}_0^*, \boldsymbol{u}_1^*, \ldots, \boldsymbol{u}_i^*, \ldots, \boldsymbol{u}_N^*) = \min_{\boldsymbol{u}_i \in \mathcal{U}_i([0,t_f])} J_i(\boldsymbol{u}_0^*, \boldsymbol{u}_1^*, \ldots, \boldsymbol{u}_i, \ldots, \boldsymbol{u}_N^*), \quad \forall i \in \mathcal{N}_0,$$

则称上述策略为航天器轨道博弈 (1) 和 (2) 的纳什均衡, 相应指标记为 $J_i(\boldsymbol{u}_0^*, \boldsymbol{u}_1^*, \ldots, \boldsymbol{u}_N^*)$. 其意义在于对于每一个航天器, 采取纳什均衡之外的策略总存在收益降低的风险.

注释2　虽然动态博弈的随机策略 (含混合策略) 具有纳什均衡, 且博弈论发展至今对此建立了相当成熟的理论及分析工具, 但对于纯策略来说, 其纳什均衡的存在唯一条件却较为苛刻, 且无一般性理论保证. 而航天器轨道博弈需要求解纯策略以保障其绝对安全性. 纳什均衡不存在意味着威胁不确定性将导致航天器在该框架下无论求解出何种纯策略, 均存在一定概率的安全风险. 因此, 面对真实威胁, 上述经典动态博弈建模可能难以完全支撑各类基于纯策略的现实应用.

为此, 我们引入基于态势信息的反馈机制进行在线动态博弈建模.

2.2　基于态势信息的反馈机制在线动态博弈建模

无纳什均衡/纳什均衡不唯一的动态博弈正体现了 “兵无常势, 水无常形” 的战略思想. 以不变应万变的纳什均衡不存在时, 随机应变、因敌制胜才是破解之道. 本文将引入基于态势信息的反馈机制, 构建 “垂直分层, 水平调整” 结构的自适应博弈在线建模方案, 以实现合理应对威胁并兼顾既定任务的轻量化策略实时求解.

2.2.1　用于策略设计的初始博弈模型选定

本文轨道威胁规避策略设计基于自适应博弈建模动态实时生成, 首先需要选定初始模型.

(A) 初始模型架构. 用于策略求解的初始模型选用非合作演化博弈模型, 为上下两层垂直分层结构, 通过闭环交互实现动态决策与执行. 每层根据 2.1 小节提供的经典建模方式建模, 完整包含博弈所需的六大关键要素, 可涵盖威胁规避、燃料节省、任务连续性、轨道约束等诸多考虑, 并以不同时间尺度演化, 如图 1 所示. 其中用于闭环反馈的态势信息包括目标形态特征、运动特征、异动特征、威胁行为、威胁等级等. 形态特征为航天器外

1) 指标泛函设计通常因场景而异. 大体而言, 式 (2) 中终端型指标 φ_i 用以评估航天器在终端时刻 t_f 的安全性与业务完成性, 可衡量其博弈结束时的终端轨道状态 (如相对距离)[24]、终端姿态 (如对地定向观测指向) 或其他特定终端约束需求; 积分型过程指标 L_i 则旨在评估在 $[0, t_f]$ 博弈期间系统状态和控制输入对航天器形成的累积代价, 使航天器既可兼顾终端目标 φ_i, 亦可智能地权衡燃料经济性、全程安全性和业务连续性等多重过程目标[27], 以寻求具备全局性能的威胁规避策略.

https://www.sciengine.com/doi/10.1360/SSI-2025-0164

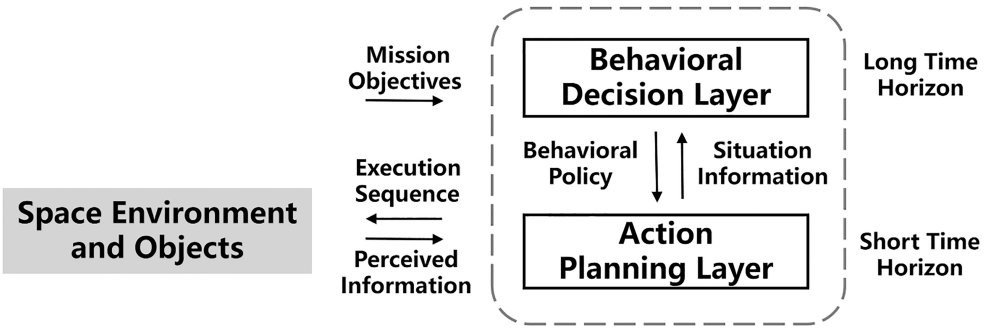


图 1　两层垂直分层结构.

Figure 1　Two-layer vertical hierarchical architecture.

形大小、结构以及所携带载荷等形态属性的参数化表达; 运动特征为轨道六根数等描述正常轨道运动的参数集合, 而位置、速度、加速度则描述抵近、绕飞等异常轨道运动的参数集合; 异动特征为区别常规轨道行为的运动特征; 威胁行为为碎片碰撞、抵近详查、激光照射、电子干扰等; 威胁等级则根据不同行为的威胁程度定义.

行为决策层负责全局性、长期性的行为策略生成, 为动作规划提供建模要素. 其模型为基于行为模型的非合作演化博弈; 输入态势特征 (由下层传递)、航天器整体任务; 输出上层策略意图, 如下层目标设定、条件约束、搜索空间等.

动作规划层负责上层行为意图的执行, 生成具体操作指令. 基于动力学/运动学模型的非合作演化博弈 (2.1 小节), 以小时间尺度动态演化; 输入环境/目标等感知信息 (由星载传感器或数据监测结果提供)、上层行为指令; 输出可执行动作序列.

层级间交互机制 (闭环反馈) 描述如下.

环境与目标 → 动作规划层: 感知传递;

动作规划层 → 行为决策层: 态势传递;

行为决策层 → 动作规划层: 策略传递;

动作规划层 → 环境与目标: 执行/扰动.

(B) 初始模型建模效能与局限. 该初始模型以两层协同通过 “决策 – 规划 – 反馈” 闭环实现动态博弈的持续演进. 其结构为内环动作规划层嵌套入外环行为决策层, 具有不同的时间演化尺度. 该结构中, 以行为层博弈实现 “大局观”, 把控整体收益; 在行为层调控下, 以下层动作层博弈实现轻量化的具体执行策略. 上下层闭环结构 (行为决策层 + 动作规划层 + 反馈机制) 通过分层抽象与动态协同, 显著提升了系统的性能与适应性. 例如, 行为决策层低频更新, 动作规划层高频响应, 可优化计算负载分配; 实时反馈调整, 策略迭代, 快速响应动态变化以及提升长期博弈收益等.

另一方面, 该初始模型双层架构除面临时间异步与实时性、跨层协同失效、长期策略退化等传统挑战外, 基于真实轨道博弈建模的嵌入带来两大核心难点: (1) 纯策略纳什均衡可能不存在/不唯一, 威胁方不确定性大; (2) 策略搜索空间大/求解复杂, 星载资源难以支撑在线计算.

从本质上讲, 上述两点均根源于博弈信息的不完备与不透明, 导致威胁不确定性大, 甚至基于反馈闭环结构求该博弈解亦可能不存在立于不败之地的航天器威胁规避策略. 即使该类策略存在, 由于建模考虑元素较多、不确定性大, 其求解/搜索复杂度也将超出星载计算机的实时处理能力. 我们将在下文中解决该问题.

2.2.2　基于态势信息反馈的自适应信息结构优化

本文选择上下两层垂直分层结构的非合作博弈模型作为规避策略设计框架, 并以基于态势信息的反馈机制自适应调节模型要素 (见图 2), 最终在线生成兼顾各项任务与约束的轻量化策略. 对于动态博弈而言, 信息结构不确定性越小, 策略搜索空间也越小, 其均衡解存在唯一的可能性却更大. 因此, 本节中反馈机制设计的根本目的是博弈信息结构的实时优化. 具体包括面向威胁方的不完美信息结构优化与不完全信息结构优化, 分行为决

https://www.sciengine.com/doi/10.1360/SSI-2025-0164

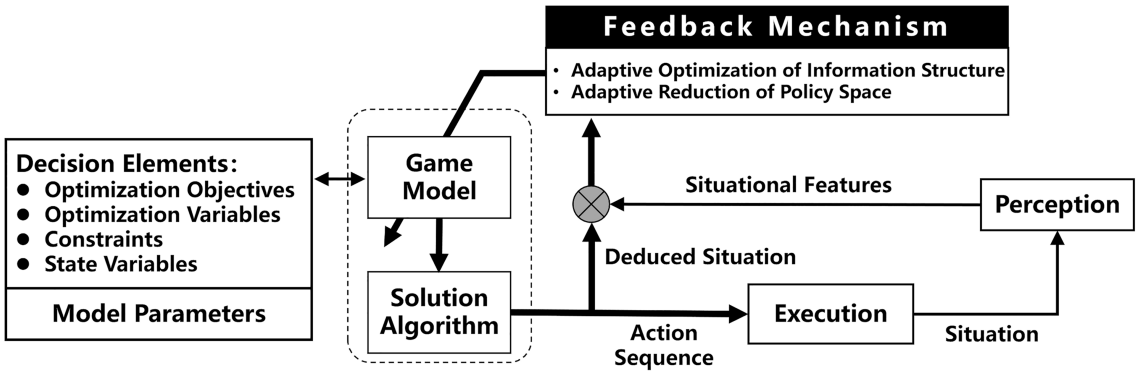


图 2　引入反馈机制实现决策空间自适应约减.

Figure 2　Adaptive policy space reduction via feedback mechanism.

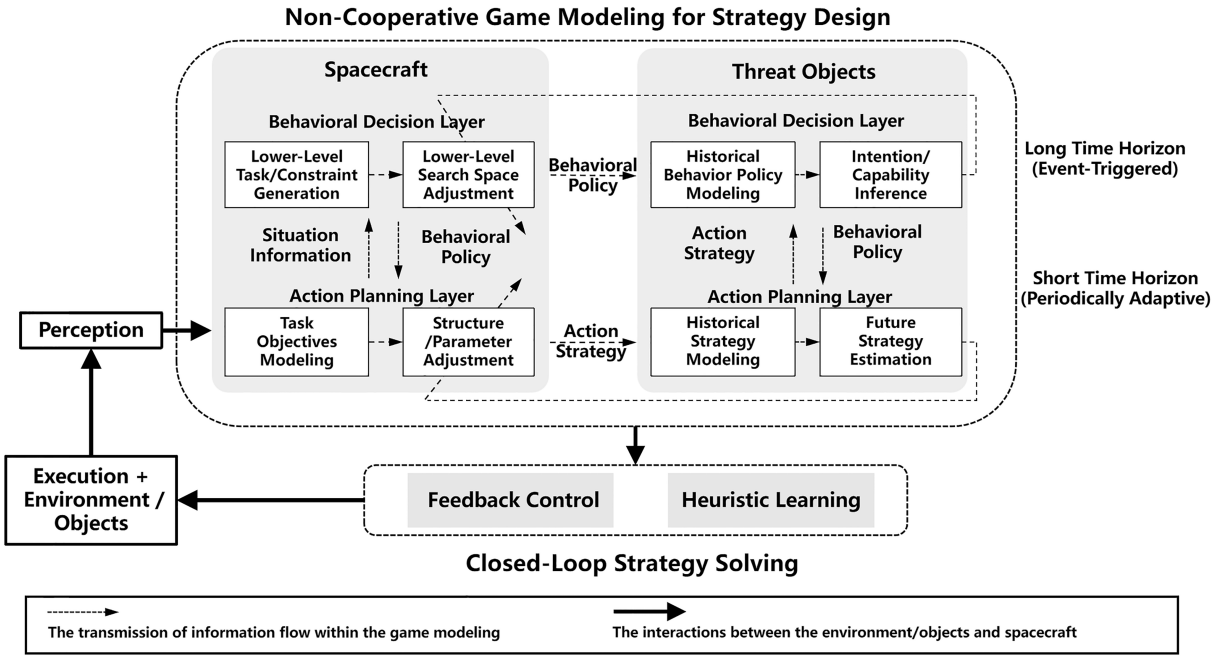


图 3　垂直分层、水平调整反馈架构.

Figure 3　An integrated feedback architecture with vertical hierarchy and horizontal adaptation.

策层与动作规划层两层结构优化, 形成纵向分工、横向适应的工作模式, 详见图 3.

(A) 不完美信息结构自适应优化. 航天器轨道博弈中, 不完美信息主要包括以下方面.

(1) 数据误差: 航天器依赖雷达、光学设备等传感器探测目标, 但受限于测量精度、噪声、数据延迟, 探测数据存在误差, 并且信息战干扰可误导航天器态势感知. 此外, 非均匀引力场、大气密度变化、太阳辐射压力等摄动也将影响编目目标的轨道预报精度.

(2) 信息延迟与丢失: 航天器通信受距离、干扰等因素影响, 信息获取不及时或不完整.

(3) 历史策略不可知: 双方历史策略因客观/主观原因使对手无法获知.

不完美信息结构优化先后在动作规划层与行为决策层进行. 在威胁方动作规划层, 通过数据融合、滤波算法、轨道外推、物理模型优化等方法处理数据误差、信息延迟/丢失, 并基于此对威胁方动作规划层历史策略建模或估计; 在威胁方行为决策层, 结合历史态势特征与威胁方动作规划层的历史策略, 学习其历史行为决策. 不完美信息结构自适应优化可为不完全信息结构自适应优化提供技术基础.

https://www.sciengine.com/doi/10.1360/SSI-2025-0164

(B) 不完全信息结构自适应优化. 轨道动态博弈结果对双方意图、能力以及具体执行策略非常敏感, 但彼此都不可能提前知悉所有信息. 此外, 航天器类型繁多, 在功能、性能、结构等方面存在显著差异. 轨道博弈初期, 受空间环境、感知限制、隐蔽技术等因素影响, 袭扰航天器/空间碎片被探测和识别的概率较低, 加之威胁方初始阶段可能采取欺骗性行为, 这皆增大了博弈态势的模糊性. 如何利用动态博弈过程中变化的态势信息准确判断/推测威胁方的真实意图、威胁能力以及具体策略逻辑, 是航天器因敌制胜、精准决策的关键所在. 因此, 规避策略设计将在行为决策层及动作规划层同时优化博弈模型的不完全信息结构. 具体而言, 在威胁方行为决策层, 基于威胁方历史行为决策与当前态势特征估计其意图、能力与联盟逻辑; 在威胁方动作规划层, 基于威胁方历史动作策略与当前态势信息估计其当前/未来策略. 信息结构自适应优化由反馈机制实现.

(C) 架构内嵌反馈机制. 航天器反馈变量包括下列四种: 环境状态信息, 如环境扰动参数、轨道参数偏差及威胁目标位置与运动轨迹等; 自身状态信息, 如燃料余量、能源状态及载荷状态等; 目标行为信息, 如威胁目标轨道机动特征、任务偏好及通信干扰等; 任务效能信息, 如任务完成度指标、风险等级评估及资源竞争状态等. 综合本节与 2.2.1 小节, 航天器规避策略设计架构中的内嵌反馈机制可实现跨层级之间与同层级各模块之间的信息流动, 由垂直反馈与横向反馈构成, 具体阐述如下.

(1) 垂直反馈是连接决策层、规划层与执行机构的核心纽带, 通过实时反馈态势变量, 自适应优化资源分配与不完美信息结构. 其路径为自下而上与自上而下, 具体包括四种情况.

威胁方动作规划层 → 威胁方行为决策层: 传递威胁目标下层历史策略.

威胁方行为决策层 → 威胁方动作规划层: 传递威胁目标协同合作策略.

航天器动作规划层 → 航天器行为决策层: 实时反馈执行偏差与态势变化.

航天器行为决策层 → 航天器动作规划层: 动态调整意图/约束/搜索空间.

(2) 横向反馈是实现同层级模块间协同的核心机制, 基于态势信息反馈变量自适应优化不完全信息结构. 其路径自左而右与自右而左, 具体包括四种情况.

航天器行为决策层 → 威胁方行为决策层: 传递航天器行为决策层策略.

威胁方行为决策层 → 航天器行为决策层: 调整航天器动作规划层策略搜索空间.

航天器动作规划层 → 威胁方动作规划层: 传递航天器动作规划层策略.

威胁方动作规划层 → 航天器动作规划层: 调整航天器动作规划层任务指标结构/参数.

(D) “垂直 + 横向” 结构反馈机制性能评价. 本策略设计架构内的反馈机制主要功能是优化信息结构. 信息结构作为动态博弈的核心要素, 直接影响参与者的策略选择与博弈结果. 传统博弈论研究中心是寻找纳什均衡, 其成立依赖于所有参与者策略建模的相互可知性. 不完全不完美信息结构打破其完全理性与共同知识假设, 导致博弈模型、决策过程和可能收益发生根本性变化. 事实上, 信息结构粗糙将破坏纳什均衡的存在性, 增加决策者风险与策略搜索难度. 常见解决方案包括采用最小化后悔准则, 即权衡不同决策可能导致的后悔程度, 选择最小后悔策略, 偏于保守防御型; 对手指标泛函参数估计, 仅适于全体参与者指标泛函结构已知场景; 基于贝叶斯框架/部分可观察马尔可夫决策过程对不完全信息建模并优化博弈决策过程, 然而不能实现航天器的绝对安全; 在深度强化学习中学习威胁方策略, 但计算量大、奖励稀疏、泛化能力难以保障.

本文通过基于态势信息的反馈机制自适应在线优化博弈信息结构, 降低态势模糊性, 如根据威胁方行动观测与态势变化感知等, 反推其隐藏信息、物理模型及执行策略, 在决策过程中逐步消除信息不完全不完美带来的不确定性, 使策略设计中的博弈模型越来越 “知己知彼”, 并最终在一定程度上逼近与模拟真实威胁场景, 从而做到有针对性的精准决策. 该机制设计可有效回应初始模型架构选定中的两大难点.

(1) 从理论需求的纳什均衡转向实际需求的适应性生存策略, 即最大化动态环境下的生存能力与任务效能. 机制合理设计下, 即使信息不对称不完备, 某种 “动态均衡” 依然存在.

(2) 策略搜索空间随信息复杂度呈指数级或更快速度增长. 信息结构合理优化可极大降低每个信息集内的不可区分状态量, 从而大幅降维闭环策略的实时搜索空间, 减小计算复杂度.

最后我们指出, 以非合作演化博弈的垂直分层结构为框架, 以态势反馈信息优化的水平动态适应为调整的航天器自主规避策略设计具备资源优化、实时适应、快速响应以及任务兼顾等优势, 但其决策性能亦取决于具

https://www.sciengine.com/doi/10.1360/SSI-2025-0164

体优化实现方式. 其反馈效能评级应综合如下关键维度.

反馈精准度: 反馈反映威胁态势的真实性, 如传感数据量测误差、轨道威胁预测误差及规避决策与真实威胁匹配度等.

反馈实时度: 反馈在规避决策中的实时性, 如反馈更新频率、在线决策时间及轨道修正窗口等.

反馈多样度: 反馈涵盖信息类型的多样性, 如威胁目标信息、燃料消耗信息、任务影响信息等.

反馈品质度: 反馈自主规避的决策效率性, 如规避成功率、规避成本及任务连续性等.

反馈稳定度: 反馈数据的可靠性, 如反馈数据波动性、反馈一致性及决策收敛性等.

基于加权融合并分类评级等方式可为轨道威胁自主规避的反馈系统提供量化评估, 确保规避决策的可靠性与最优性.

3　反馈机制信息结构优化应用实例

本节以空间碎片威胁自主规避与航天器反侦察载荷自主保护为例, 说明轨道博弈反馈机制的实现过程, 核心功能为信息结构优化.

3.1　空间碎片碰撞威胁自主规避

当前航天领域空间碎片威胁规避面临几大技术瓶颈. 碎片监测与跟踪技术上, 微小碎片存在探测盲区, 高轨碎片定位误差较大, 碎片轨道数据更新频率低, 无法支撑实时预警等. 碰撞风险评估与预测技术上, 轨道动力学模型因各类摄动误差导致长期轨迹预测不准, 碰撞概率计算难度大、耗时长, 多碎片环境下碰撞虚警率攀升, 航天器频繁执行无效规避等. 规避机动技术上, 受航天器机动能力、燃料与任务限制, 最优机动时间、最优规避轨迹计算代价大, 并且不易应对突发事件, 尤其在星座卫星群或密集轨道区域中, 级联碰撞风险逐渐显现, 即单次规避可能导致后续与其他碎片/卫星的碰撞风险增大等.

综上, 规避任务的一个主要难点在于信息的不完美性, 因此无法准确评估碰撞风险、判断机动时间与规划最优轨迹. 空间碎片无机动能力, 可视为博弈退化情形. 另一方面, 尽管碎片真实轨道固定, 但由于测量的不确定性, 其预测轨道以一定估计概率具有多种可能性, 因此我们可以认为碎片威胁规避存在不完全信息 (复杂碎片规避场景也可考虑双方均具多策略的博弈框架). 该不完全信息由不完美信息引起, 优化不完美信息结构的同时亦即优化了不完全信息结构. 两者的优化包括数据收集与预处理、轨道预报、潜在威胁筛选、碰撞概率计算等. 此外, 航天器自主规避系统可基于星载传感器近距观测提高测量精度和实时性, 降低虚警率, 并不断更新轨道信息. 规避策略除采取脉冲机动提前变轨, 还可通过基于自主感知的自适应规避方式应对突发情况或调整规划误差, 实现近距微调或紧急规避. 但航天器自主感知精准测量具有一定范围, 考虑到大多数空间碎片为掠飞抵近, 自适应近距机动意味着短期规避, 与长期规避策略相比存在如下问题.

(1) 航天器横向机动能力弱, 当空间碎片进入其自主感知系统视场后, 航天器推进系统是否具备支撑其自主成功规避的能力? 如高速掠飞场景相对速度可达 7.8∼15 km/s, 而航天器典型加速度为 0.01∼1 m/s^2, 此时无法实现短期规避.

(2) 早期规避更节省燃料, 但根据地基/天基观测数据的轨道长期预报可能存在较大误差导致虚警, 航天器执行无效机动反而造成不必要的燃料消耗. 另一方面, 等待碎片接近, 以星上测量系统获取可靠信息可降低虚警率, 在相对速度较低时航天器能够实现自主成功规避, 但需消耗更多燃料. 如何权衡两者?

第一个问题可根据预报数据与基于自主感知的自适应规避快速算法, 计算某个空间碎片进入航天器视场时间 t_F、两者碰撞时间 t_C 以及近距规避所需速度增量大小 δv 进行判断. 假设航天器最大轨道加速度大小为 $u_{\max}$, 如果

$$\hat{P}\{\delta v > u_{\max}(t_C - t_F)\} > \epsilon_{Inev}, \tag{3}$$

https://www.sciengine.com/doi/10.1360/SSI-2025-0164

表 1　例 1 中的航天器初始轨道根数.

Table 1　Initial orbital elements of the spacecraft in Example 1.

Symbol	Description	Value
a_{ref}	Semi-major axis	7155.46 km
e_{ref}	Eccentricity	1.17×10^{-3}
i_{ref}	Inclination	1.29 rad
Ω_{ref}	Right ascension of the ascending node	0.30 rad
ω_{ref}	Argument of periapsis	1.47 rad
ϕ_{ref}	True anomaly	0.23 rad

其中 $\hat{P}$ 为估计概率, ϵ_{Inev} 为设定阈值[2)], 那么执行该算法进行短期规避存在碎片碰撞风险, 必须提前规划最优机动轨迹, 在航天器自主感知范围外便采取轨道机动.

针对第二个问题, 我们引入等效燃料成本 (expected fuel consumption) 概念. 首先假设式 (3) 对所有关注的空间碎片均不成立, 即在航天器视场范围内, 基于自主感知的自适应快速算法可进行成功规避. 令 $\hat{P}_c$ 为某个时刻基于早期观测数据计算的一组碎片碰撞概率, 超过需要机动规避的给定阈值. 此时机动的速度增量大小设为 δV, 为其等效燃料成本. 我们定义自适应规避快速算法的等效燃料成本为 $\hat{P}_c\cdot\delta v$. 如果

$$\delta V\geqslant\hat{P}_c\cdot\delta v, \tag{4}$$

那么自适应短期规避则更具优势. 这主要是由于近距自主感知精度较高, 可降低虚警率, 即使在早期预测准确的情形中燃料相对消耗较高, 但由于其不可预知, 平均期望上自适应规避成本更低.

我们将早期脉冲机动规避 (提前规划生成) 与短期自适应规避相结合, 根据判别准则 (3) 和 (4) 集成于航天器自主规避决策模块内 [48], 在节省燃料的同时可应对机动规划误差与突发威胁状况.

航天器运行于近圆轨道, 初始轨道参数见表 1, 其最大加速度为 0.03 m/s^2, 星载传感器最大探测距离为 60 km, 位置、速度观测最大标准偏差为 10 m 和 1 m/s. 假设地面观测站对非合作目标的位置、速度测量标准偏差为 600 m 和 20 m/s. 初始时刻航天器收到地面预警, 预测 4 小时后将有 4 个空间碎片与其发生碰撞, 一者为高速碰撞, 碰撞速度为 10.52 km/s, 其余三者则为低速碰撞, 相应初始时刻轨道根数及最近接近时刻 (TCA) 如表 2 所示. 本例将比较基于地面观测站的传统规避方法与基于星载传感器的自主规避方法在碎片威胁中的规避效能.

工况 1: 采取传统脉冲机动规避算法. 基于地面观测数据生成针对碎片包络的脉冲机动策略, 安全距离设置为 8 km, 相应速度增量大小为 0.28 m/s, 仿真显示卫星避开 4 个威胁碎片.

工况 2: 采取自主规避系统决策算法. 根据表 1 与判别式 (3), 航天器首先执行脉冲机动以规避高速碰撞碎片, 其速度增量沿原速度方向增加 0.07 m/s, 变轨后与低速碎片碰撞概率为 $\hat{P}_c=10^{-3}$. 其次, 若基于视场内自主感知进行规避, 速度增量大小粗略估计满足 $\delta v<7.51$ m/s, 其中相应安全距离设置为 100 m. 显然, 0.07 m/s + $10^{-3}\times7.51$ m/s <0.28 m/s. 由判别式 (4), 航天器将在视场内自适应调整规避轨道以躲避三个低速碰撞碎片, 其执行的自适应算法源自文献 [48]. 假设在航天器自主感知下, 其中一低速碎片被排除风险, 其余两碎片在 TCA 时刻相对运动状态如表 3 所示. 仿真表明航天器成功规避所有碎片, 实际总速度增量大小为 0.12 m/s. 与前者比较, 自主规避系统决策算法燃料消耗可节省约 57.142%, 这主要归功于反馈机制对不确定性的大幅度消除.

2) ϵ_{Inev} 对应航天器实施自适应规避机动后需满足的碰撞概率阈值, 该阈值通常比碰撞概率预警阈值低 1∼2 个数量级. 碰撞概率预警阈值的选定由轨道预报精度、安全冗余度、航天器价值以及规避燃料成本等因素共同决定, 阈值过小会导致频繁虚警燃料消耗过大, 过大则可能导致漏警. 一般而言, 碰撞概率预警阈值的选定范围在 10^{-3}∼10^{-4}, 对于空间站、载人航天器等高价值空间资产采用最严格阈值 10^{-4}. 本文例 1 中碰撞概率预警阈值选定为 10^{-3}, 其对应机动后阈值低 2 个数量级, 即 $\epsilon_{Inev}=10^{-5}$.

https://www.sciengine.com/doi/10.1360/SSI-2025-0164

表 2　四空间碎片初始轨道根数.

Table 2　Initial orbital elements of four pieces of space debris.

Symbol	Description	Value
t_C	Time of closest approach (TCA)	14400/14525.22/15413.68/16096.38 s
a	Semi-major axis	7160.40/7238.40/7348.71/7346.47 km
e	Eccentricity	$2.08\times10^{-3}/9.87\times10^{-3}/2.41\times10^{-2}/2.55\times10^{-2}$
i	Inclination	2.14/1.30/1.29/1.29 rad
Ω	Right ascension of the ascending node	1.67/0.31/0.30/0.30 rad
ω	Argument of periapsis	1.64/4.00/4.95/6.20 rad
ϕ	True anomaly	$0.31/-2.07/-2.68/2.45$ rad

表 3　LVLH 坐标系下低速碎片与航天器 TCA 时刻相对运动状态.

Table 3　Relative motion state of the low-velocity debris and the spacecraft at TCA in LVLH coordinate system.

Symbol	Description	Value
t_C	Time of closest approach (TCA)	15412.24/16090.75 s
r_{Cx}	TCA distance in x-axis direction	$-1.68/-260.68$ m
r_{Cy}	TCA distance in y-axis direction	$-3.70/-371.51$ m
r_{Cz}	TCA distance in z-axis direction	$-12.95/38.27$ m
v_{Cx}	TCA velocity in x-axis direction	$32.21/-29.46$ m/s
v_{Cy}	TCA velocity in y-axis direction	89.92/93.35 m/s
v_{Cz}	TCA velocity in z-axis direction	$-7.15/44.07$ m/s

3.2　航天器反侦察载荷自主保护

侦察卫星抵近侦察时, 航天器可通过隐藏载荷等方式来规避其侦察. 载荷侦察与反侦察为典型的不完美不完全信息非合作博弈. 与空间碎片威胁不同, 侦察卫星具备明显的策略性. 但高动态环境感知中微动特征解译困难, 可能无法完全破解其历史策略与行为, 因此博弈信息不完美. 此外, 由于空间光照条件复杂, 星载传感器成像条件苛刻、连续性差等因素, 威胁目标特征不易识别, 对侦察卫星载荷种类/能力信息获取不完备; 伴飞、掠飞、绕飞等轨道威胁行为特征不明显, 侦察行为持续时间可能较长且多变, 编队侦察星群内部协作机制不清, 因此即使综合利用威胁目标形态、运动甚至历史行为等信息也难以对其意图与策略作出精准预测, 对应信息不完全性. 另一方面, 航天器需综合载荷保护、燃料消耗、任务连续性、光照约束等要素, 进行多目标多约束优化问题在线求解. 考虑到航天器在轨长期运行期间设备难以更换升级, 星载计算、存储等资源严重受限, 而对抗博弈场景对策略时效性要求高, 因此面临时间和资源上的严重冲突.

事实上, 航天器反侦察载荷保护为强对抗非对称博弈, 即由于信息不对称导致博弈双方指标设置非零和, 通常不存在纯策略纳什均衡. 为此, 我们引入侦察卫星协同累积观测区域

$$S(t)=\bigcup\nolimits_{i\in\mathcal{N}_0\setminus\{0\}} S_i(t),$$

其中 $S_i(t)$ 为第 i 个侦察卫星 t 时刻对航天器的观测区域, 由侦察卫星载荷成像要素 (包括相机视场角与逆光成像能力)、当前时刻光照条件以及我方航天器外型等因素建模刻画. 相关侦察能力参数将在博弈过程中基于自主感知或天/地基上注威胁目标身份/能力信息得以完善. 我们认为侦察卫星编队的目标在于极大化 $\bigcup_t S(t)$ 以期实现对载荷的潜在覆盖 (侦察卫星没有航天器载荷的先验知识), 而航天器则旨在极小化载荷落入 $S(t)$ 的时间.

下面首先进行不完美不完全信息结构优化. 具体而言, 在侦察方行为决策层, 完善侦察能力参数信息, 并根据侦察卫星历史行为与当前态势特征, 估计其编队协同策略, 优化不完全信息. 其中态势特征包括目标轨道根数、相对位置与相对速度、轨道演化趋势、目标姿态变化与对准方向、威胁方编队或分布方式及周边威胁目标密度

https://www.sciengine.com/doi/10.1360/SSI-2025-0164

表 4　例 2 中的航天器初始轨道根数.

Table 4　Initial orbital elements of the spacecraft in Example 2.

Symbol	Description	Value
a_{ref}	Semi-major axis	6778.14 km
e_{ref}	Eccentricity	0
i_{ref}	Inclination	1.71 rad
Ω_{ref}	Right ascension of the ascending node	0
θ_{ref}	Argument of latitude	1.57 rad

表 5　太阳方位矢量、目标指向矢量及载荷分布矢量.

Table 5　Sun vector, mission orientation vector and payload distribution vector.

Vector type	Symbol	Value	Coordinate system
Sun vector	$\hat{\boldsymbol{\alpha}}_{\mathrm{sun}}$	$[1,0,0]$	ECI coordinate system
Mission orientation vector	$\hat{\boldsymbol{q}}_{\mathrm{mission}}$	$[-1,0,0]$	LVLH coordinate system
Payload 1 initial orientation vector	$\hat{\boldsymbol{q}}_1$	$[-1,0,0]$	Body coordinate system
Payload 2 initial orientation vector	$\hat{\boldsymbol{q}}_2$	$\left[-\frac{\sqrt{3}}{2},\frac{1}{2},0\right]$	Body coordinate system
Payload 3 initial orientation vector	$\hat{\boldsymbol{q}}_3$	$\left[-\frac{\sqrt{3}}{2},-\frac{1}{2},0\right]$	Body coordinate system

等. 在侦察方动作规划层, 引入侦察卫星投影在航天器模型上的星下点历史轨迹, 并结合光照条件估计其历史侦察观测区域 $\hat{S}_t$, 优化不完美信息; 结合编队协同策略, 以最大化 $S(t)\setminus\hat{S}_t$ 为目标生成侦察卫星机动策略, 优化不完全信息.

之后航天器进行决策空间反馈调节. 具体而言, 在航天器行为决策层, 根据侦察卫星数量、相机成像能力、双方轨控能力差异等态势因素 (由信息结构优化所得) 判断是否需要姿轨协同机动, 并调整轨控指标结构/参数. 在航天器动作规划层, 基于侦察卫星机动策略 (由信息结构优化所得), 实时调整姿控指标结构/参数.

考虑航天器遭遇三颗侦察卫星编队侦察的高威胁场景. 航天器需保护 3 个重要载荷以避免落入威胁目标视场范围内, 同时尚需考虑燃料节省与任务连续性. 设定航天器最大轨控加速度为 0.01 $\mathrm{m/s^2}$, 姿控最大力矩为 10 N·m; 侦察卫星最大轨控加速度均为 0.01 $\mathrm{m/s^2}$, 视场半角为 $\frac{29\pi}{60}$ rad[3]. 航天器轨道参数、太阳方位矢量、目标指向矢量以及载荷分布矢量见表 4 和 5. 此外, 侦察卫星在 LVLH 坐标系下的初始相对位置分别为 $[0.02,0,-2500]\,\mathrm{m}$, $[0.02,0,-2450]\,\mathrm{m}$ 和 $[0.02,0,-2400]\,\mathrm{m}$, 对应初始相对速度为 $[2.83,0.50,0]\,\mathrm{m/s^2}$, $[2.83,0,0]\,\mathrm{m/s^2}$ 和 $[2.83,-0.50,0]\,\mathrm{m/s^2}$. 航天器初始姿态向量为 $[0,0,0,0,0,0]$. 设置弈时间为 10000 s. 下文将比较基于动态博弈建模设计的反馈自适应算法与无策略空间反馈自适应约减的载荷保护算法在反侦察场景中的策略性能.

工况 1: 航天器基于在线动态博弈建模设计策略. 假定侦察卫星执行固定协同绕飞侦察, 这是实际空间侦察中广泛采用的一类策略. 此处轨迹跟踪具体实现细节可参见文献 [45], 其中我们预设三颗侦察卫星绕飞轨道长、短轴分别为 5000 m 与 2500 m. 图 4 展示了该固定绕飞策略下, 侦察卫星协同累积观测区域增长过程, 表明侦察星群在持续获取关于航天器的新观测. 图 5 与 6 分别量化了航天器载荷指向损失[4]与任务连续性[5], 图 7 则展示了航天器的燃料消耗. 仿真表明航天器可顺利保护 3 个载荷并能维持一定的任务连续性, 总速度增量为 7.876 m/s. 我们指出, 由于航天器借助反馈机制实时优化博弈态势信息并以此调整策略, 即使侦察卫星真实策略

3) 一般而言, 视场半角取值因应用场景与光学设计目标而异, 从 0 rad 到接近 $\frac{\pi}{2}$ rad 均有覆盖, 视场半角越小通常对应分辨率越高. 本例未考虑分辨率对侦察任务的影响, 即假设侦察星群具有较强的成像能力, 适配较大视场半角 ($\frac{29\pi}{60}$ rad).

4) 载荷指向损失定义为航天器的载荷指向 $\hat{\boldsymbol{q}}_j$ 与侦察卫星相机指向方向 $\hat{\boldsymbol{r}}_i, i=1,2,3$ 之间的最大夹角, 如下式所示:

$$\max_{i\in\mathcal{N}} \hat{\boldsymbol{q}}_j^T\hat{\boldsymbol{r}}_i,\quad j=1,2,3.$$

若载荷指向损失小于某一给定阈值 (由侦察星群相机成像能力等因素决定), 则认为该载荷暴露, 航天器博弈失败.

5) 任务连续性定义为目标指向方向 $\hat{\boldsymbol{q}}_{\mathrm{mission}}$ 与航天器载荷 1 指向方向 $\hat{\boldsymbol{q}}_1$ 之间的夹角, 其在博弈过程中应保持尽可能小.

https://www.sciengine.com/doi/10.1360/SSI-2025-0164

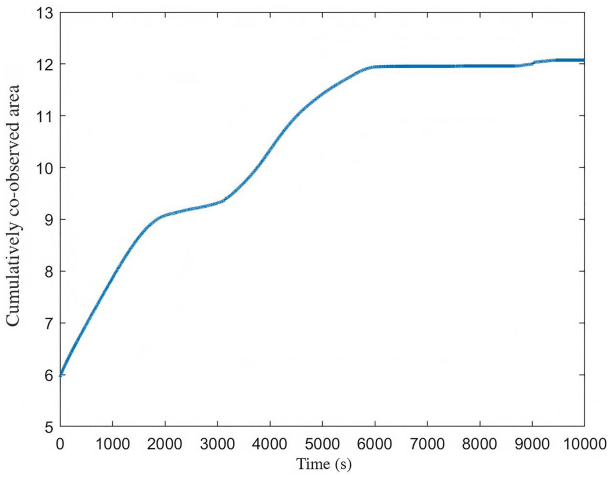


图 4　(网络版彩图) 工况 1 侦察方协同累积观测区域面积.

Figure 4 (Color online) Cumulatively co-observed area coverage of the reconnaissance team under Scenario 1.

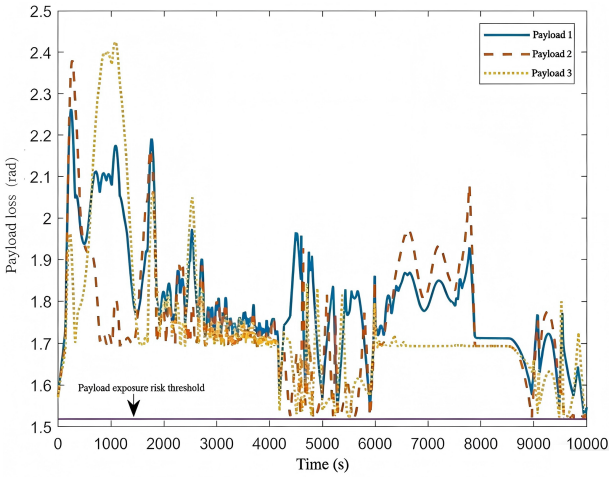


图 5　(网络版彩图) 工况 1 航天器载荷指向损失.

Figure 5 (Color online) Payload pointing loss of the spacecraft under Scenario 1.

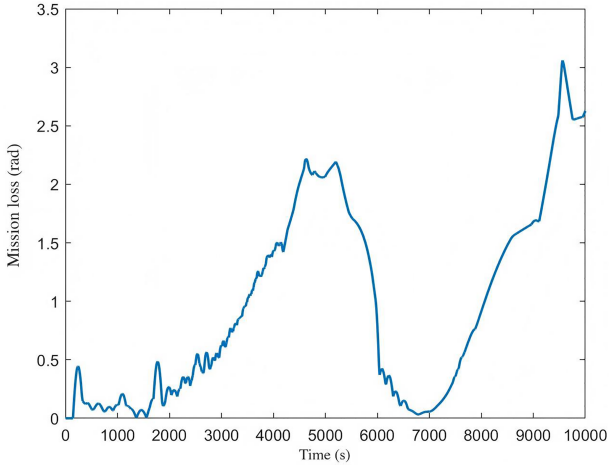


图 6　(网络版彩图) 工况 1 航天器任务连续性.

Figure 6 (Color online) Mission continuity of the spacecraft under Scenario 1.

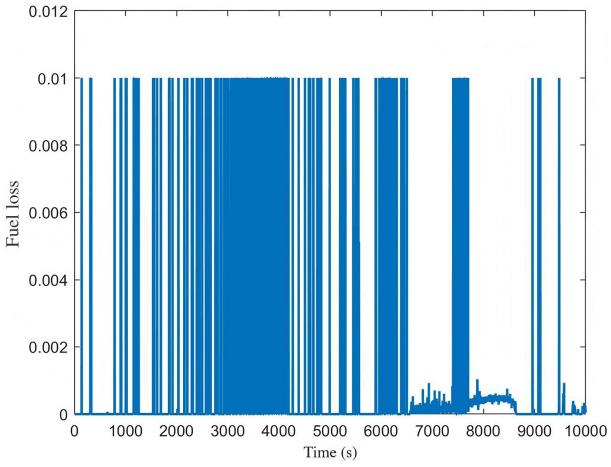


图 7　(网络版彩图) 工况 1 航天器燃料消耗量.

Figure 7 (Color online) Fuel consumption of the spacecraft under Scenario 1.

与我们建模存在一定偏差, 也可保证航天器基于态势反馈策略的相对高效性, 即该策略设计鲁棒性较好.

工况 2: 航天器不基于本文反馈机制设计约减决策空间. 本工况中侦察卫星采取与工况 1 相同轨道配置的固定协同绕飞策略, 而航天器则不利用该反馈机制约减搜索空间. 但由于计算量过大, 无法实现全局寻优, 为保证策略实时性我们将对其执行滚动优化这一传统方案, 每步决策时间为 1 s, 与工况 1 中反馈决策算法保持一致. 此时, 侦察卫星协同累积观测区域与航天器载荷指向损失分别由图 8 与 9 刻画. 图 10 与 11 则分别量化了航天器任务连续性及燃料消耗, 其中总速度增量为 6.514 m/s. 对比工况 1 可知, 尽管协同累积观测区域更小, 但载荷多次暴露. 不难看出航天器此处未把载荷保护作为第一目标, 而是优先考虑燃料节省与业务连续, 导致决策归于失败. 这主要是由于传统滚动优化虽能减轻计算负担, 但航天器决策却因此过于 “短视”, 策略搜索陷入局部最优. 而工况 1 引入前述反馈机制自适应调整, 在约减决策空间的同时亦兼顾长期整体利益, 具有较好的全局性. 最后, 表 6 对比了相同硬件环境下工况 1 与工况 2 的计算性能, 可见即使工况 2 采用滚动优化来降低计算负荷, 但其计算性能仍不及工况 1 中反馈机制带来的轻量化优势.

https://www.sciengine.com/doi/10.1360/SSI-2025-0164

表 6　性能对比 (CPU @ 2.1 GHz): 内存消耗和 CPU 占用.

Table 6　Performance comparison (CPU @ 2.1 GHz): memory consumption and CPU usage.

Scenario 1		Scenario 2	
Peak memory usage	Peak CPU usage	Peak memory usage	Peak CPU usage
560.79 MB	42.31%	834.50 MB	51.48%

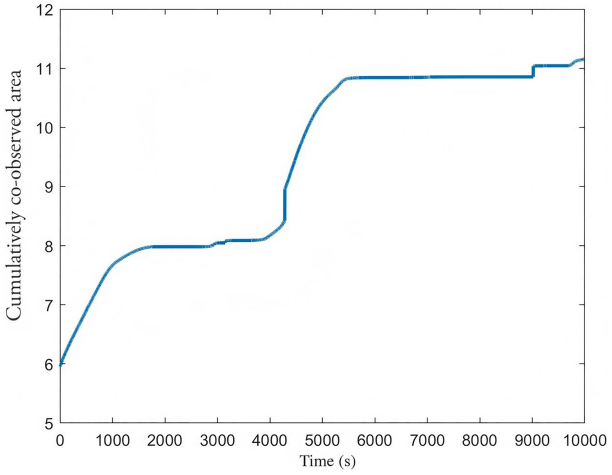


图 8　(网络版彩图) 工况 2 侦察方协同累积观测区域面积.

Figure 8　(Color online) Cumulatively co-observed area coverage of the reconnaissance team under Scenario 2.

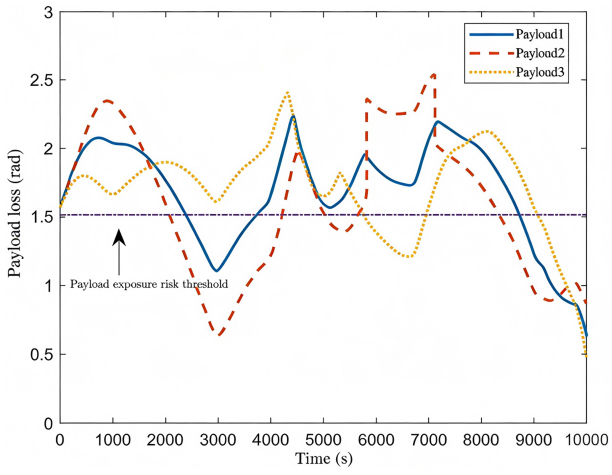


图 9　(网络版彩图) 工况 2 航天器载荷指向损失.

Figure 9　(Color online) Payload pointing loss of the spacecraft under Scenario 2.

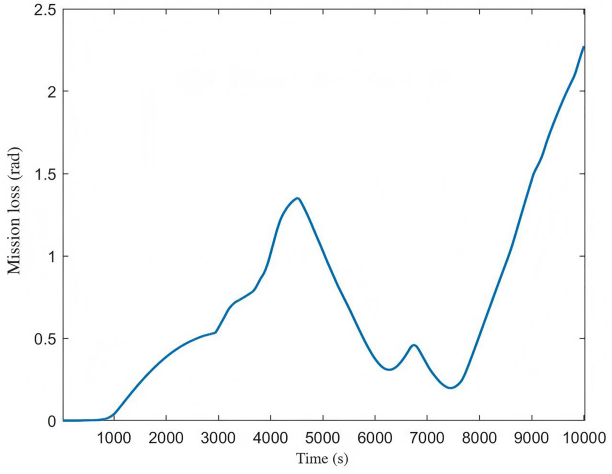


图 10　(网络版彩图) 工况 2 航天器任务连续性.

Figure 10　(Color online) Mission continuity of the spacecraft under Scenario 2.

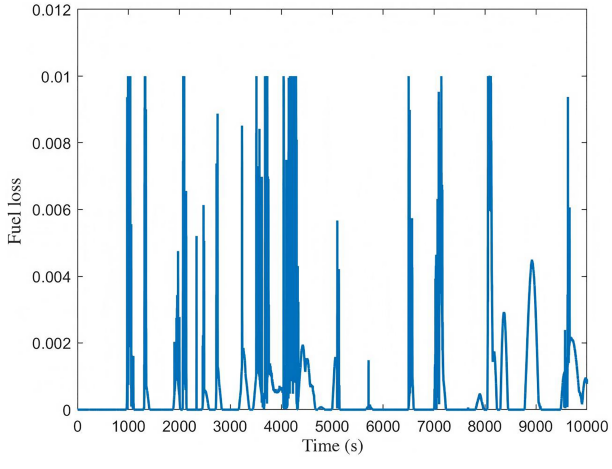


图 11　(网络版彩图) 工况 2 航天器燃料消耗量.

Figure 11　(Color online) Fuel consumption of the spacecraft under Scenario 2.

4　总结与展望

面向航天器轨道威胁自主规避, 本文提出一种具有 “垂直分层” 与 “水平调整” 结构的非合作博弈在线建模与优化架构. 通过引入基于态势信息的反馈机制, 进行模型结构与要素的自适应优化与约减, 可有效实现不完美不完全信息下的多任务多约束决策轻量化求解. 本文以空间碎片规避与反侦察载荷保护两个典型轨道威胁场景

https://www.sciengine.com/doi/10.1360/SSI-2025-0164

为例, 说明该架构下博弈策略设计具备资源优化、实时适应、快速响应以及任务兼顾等优势. 但反馈机制应对不完备信息结构的能力极限, 即能够应对多大的意图策略模糊性、模型不确定性、信息更新滞后以及机动执行偏差等, 尚需进一步定量研究.

参考文献

1 Gao W Y, Wu J F, Wei C L. Review on spacecraft autonomous decision-making and planning for orbital threat avoidance. Chin Space Sci Technol, 2024, 44: 71–89 [高婉莹, 吴健发, 魏春岭. 航天器威胁规避自主决策规划方法研究综述. 中国空间科学技术, 2024, 44: 71–89]

2 Yuan L, Jiang T T. Review on intelligent autonomous control for spacecraft confronting orbital threats. Acta Autom Sin, 2023, 49: 229–245 [袁利, 姜甜甜. 航天器威胁规避智能自主控制技术研究综述. 自动化学报, 2023, 49: 229–245]

3 Moranta S, Hrozensky T, Dvoracek M. Towards a European approach to space traffic management. ESPI Report 71, European Space Policy Institute, Vienna, 2020

4 Roulette J. OneWeb, SpaceX satellites dodged a potential collision in orbit. 2021. https://www.theverge.com/2021/4/9/22374262/oneweb-spacex-satellites-dodged-potentialcollision-orbit-space-force

5 中国常驻联合国 (维也纳) 代表团. 2021 年 12 月 3 日致秘书长的普通照会 [AAC.1051262]. 2021-12-03. https://www.unoosa.org/res/oosadoc/data/documents/2021/aac_105/aac_1051262_0_html/AAC105_1262C.pdf

6 Pultarova T. SpaceX Starlink satellites made 50000 collision-avoidance maneuvers in the past 6 months. What does that mean for space safety. 2022. https://www.space.com/spacex-starlink-50000-collision-avoidance-maneuvers-space-safety

7 Gong J G, Ning Y, Lv N. Development and enlightenment of space based situational awareness technology for high orbit in the United States. Aerospace Control Appl, 2021, 47: 1–7 [宫经刚, 宁宇, 吕楠. 美国高轨天基态势感知技术发展与启示. 空间控制技术与应用, 2021, 47: 1–7]

8 Martin M, Pfrang K, Weeden B. U.S. military and intelligence rendezvous and proximity operations in space. Technical Report, Secure World Foundation, Washington, 2021. 1–4

9 Yuan L. Spacecraft intelligent autonomous control technology toward uncertain environment. J Astronaut, 2021, 42: 839–849 [袁利. 面向不确定环境的航天器智能自主控制技术. 宇航学报, 2021, 42: 839–849]

10 Pyrak M, Anderson J. Performance of Northrop Grumman's mission extension vehicle (MEV) RPO imagers at GEO. In: Proceedings of the Advanced Maui Optical and Space Surveillance Technologies Conference, Maui, 2020. 6031

11 Fernholz T. SpaceX's new satellites will dodge collisions autonomously. 2021. https://qz.com/1627570/howautonomous-are-spacexs-starlink-satellites/

12 Virgili B B, Flohrer T, Krag H, et al. CREAM ESA's proposal for collision risk estimation and automated mitigation. In: Proceedings of the 1st International Orbital Debris Conference, Sugar Land, 2019. 6031

13 Garofalo M. US Space Force picks Rocket Lab for 2025 Victus Haze space domain awareness mission. 2024-04-27. https://www.space.com/space-exploration/satellites/us-space-force-picks-rocket-lab-for-2025-victus-haze-space-domain-awareness-mission

14 Isaacs R. Differential Games: A Mathematical Theory With Applications to Warfare and Pursuit, Control and Optimization. New York: Dover Publications, 1999

15 Li Q, Chen L, Tai C, et al. Maximum principle based algorithms for deep learning. J Mach Learn Res, 2018, 18: 5998–6026

16 Evans L, Souganidis P. Differential-games and representation formulas for solutions of Hamilton-Jacobi-Isaacs equations. Inda Univ Math J, 1984, 33: 773–797

17 Limebeer D J N, Anderson B D O, Khargonekar P P, et al. A game theoretic approach to $\mathcal{H}^\infty$ control for time-varying systems. SIAM J Control Optim, 1992, 30: 262–283

18 Luo Y Z, Li Z Y, Zhu H. Survey on spacecraft orbital pursuit-evasion differential games. Sci Sin Tech, 2020, 50: 1533–1545 [罗亚中, 李振瑜, 祝海. 航天器轨道追逃微分对策研究综述. 中国科学技术科学, 2020, 50: 1533–1545]

19 Shen H X, Casalino L. Revisit of the three-dimensional orbital pursuit-evasion game. J Guidance Control Dyn, 2018, 41: 1823–1831

20 Zeng X, Yang L, Zhu Y, et al. Comparison of two optimal guidance methods for the long-distance orbital pursuit-evasion game. IEEE Trans Aerosp Electron Syst, 2021, 57: 521–539

21 Shi M, Ye D, Sun Z, et al. Spacecraft orbital pursuit-evasion games with J2 perturbations and direction-constrained thrust. Acta Astronaut, 2023, 202: 139–150

22 Venigalla C, Scheeres D J. Delta-v-based analysis of spacecraft pursuit-evasion games. J Guidance Control Dyn, 2021, 44: 1961–1971

23 Carr R W, Cobb R G, Pachter M, et al. Solution of a pursuit-evasion game using a near-optimal strategy. J Guidance Control Dyn, 2018, 41: 841–850

24 Pontani M, Conway B A. Numerical solution of the three-dimensional orbital pursuit-evasion game. J Guidance Control Dyn,

https://www.sciengine.com/doi/10.1360/SSI-2025-0164

2009, 32: 474–487
25 Hafer W T, Reed H L, Turner J D, et al. Sensitivity methods applied to orbital pursuit evasion. J Guidance Control Dyn, 2015, 38: 1118–1126
26 Li Z, Zhu H, Yang Z, et al. A dimension-reduction solution of free-time differential games for spacecraft pursuit-evasion. Acta Astronaut, 2019, 163: 201–210
27 Jagat A, Sinclair A J. Optimization of spacecraft pursuit-evasion game trajectories in the Euler-Hill reference frame. In: Proceedings of the AIAA/AAS Astrodynamics Specialist Conference, 2014. 1–10
28 Pannocchia G, Rawlings J B, Mayne D Q, et al. On computing solutions to the continuous time constrained linear quadratic regulator. IEEE Trans Automat Contr, 2010, 55: 2192–2198
29 Liu Y, Li R, Hu L, et al. Optimal solution to orbital three-player defense problems using impulsive transfer. Soft Comput, 2018, 22: 2921–2934
30 Li Z Y. Orbital pursuit-evasion-defense linear-quadratic differential game. Aerospace, 2024, 11: 443
31 Liu Y, Ye D, Hao Y. Distributed online mission planning for multi-player space pursuit and evasion. Chin J Aeronautics, 2016, 29: 1709–1720
32 Xu X S, Dang Z H, Song B, et al. Method for cluster satellite orbit pursuit-evasion game based on multi-agent deep deterministic policy gradient algorithm. Aerospace Shanghai, 2022, 39: 24–31 [许旭升, 党朝辉, 宋斌, 等. 基于多智能体强化学习的轨道追逃博弈方法. 上海航天 (中英文), 2022, 39: 24–31]
33 Li R F, Sun C, Wang Z, et al. Capture configuration design for space non-cooperative target based on reachable domain. Chin Space Sci Tech, 2024, 44: 83–94 [李瑞峰, 孙冲, 王政, 等. 基于可达域分析的空间非合作围捕构型设计. 中国空间科学技术, 2024, 44: 83–94]
34 Li Z Y, Zhu H, Luo Y Z. An escape strategy in orbital pursuit-evasion games with incomplete information. Sci China Tech Sci, 2021, 64: 559–570
35 Ye D, Tang X, Sun Z, et al. Multiple model adaptive intercept strategy of spacecraft for an incomplete-information game. Acta Astronaut, 2021, 180: 340–349
36 Tang X, Ye D, Huang L, et al. Pursuit-evasion game switching strategies for spacecraft with incomplete-information. Aerospace Sci Tech, 2021, 119: 107112
37 Tang X, Ye D, Xiao Y, et al. Epsilon Nash equilibrium differential game strategy for spacecraft terminal pursuit-evasion under incomplete information. J Astron, 2024, 45: 63–73 [汤旭, 叶东, 肖岩, 等. 不完全信息 Epsilon 纳什均衡的航天器末端追逃博弈策略. 宇航学报, 2024, 45: 63–73]
38 Yang B, Liu P, Feng J, et al. Two-stage pursuit strategy for incomplete-information impulsive space pursuit-evasion mission using reinforcement learning. Aerospace, 2021, 8: 1–15
39 Wang Z, Gong B, Yuan Y, et al. Incomplete information pursuit-evasion game control for a space non-cooperative target. Aerospace, 2021, 8: 1–12
40 Tang X, Ye D, Luo S, et al. A hybrid game strategy for the pursuit of out-of-control spacecraft under incomplete-information. Aerospace, 2022, 9: 1–15
41 Plante D, Levis A. On the reconnaissance game. IEEE Trans Automat Contr, 1972, 17: 831–833
42 Wang C, Chen D, Liao W. Research on maneuver strategy in satellite observation and counter-observation game. Adv Space Res, 2024, 74: 3170–3185
43 Dong C, Zhu M, Guo J, et al. Spacecraft anti-reconnaissance game based on particle swarm optimization algorithm. Adv Astronaut Sci Technol, 2024, 7: 121–131
44 Yao Y, Xie R, He F. Flyaround Orbit Design for Autonomous Rendezvous Based on Relative Orbit Elements. J Guidance Control Dyn, 2010, 33: 1687–1692
45 Lee X, Yang B, Yao Y. Autonomous approach and fly around a target satellite with input constraints. In: Proceedings of the AIAA Guidance, Navigation, and Control (GNC) Conference, 2013. 1–12
46 Zhuang H X, Yao Z J, Xu C, et al. Research on space countermeasures system. Aerospace Electron Warfare, 2013, 29: 11–13, 30 [庄海孝, 姚志健, 徐超, 等. 空间电子对抗系统研究. 航天电子对抗, 2013, 29: 11–13, 30]
47 Zhou Y J, Zhou N, Qi W K. Present condition and development trend of foreign space electronic countermeasure techniques. Electron. Countermeas, 2013, 1: 6–11 [周业军, 周钠, 齐维孔. 国外空间电子对抗技术现状与发展趋势展望. 电子对抗, 2013, 1: 6–11]
48 Li Z, Li H, Li C. ELVO-based autonomous satellite collision avoidance with multiple debris. Aerospace, 2025, 12: 402

https://www.sciengine.com/doi/10.1360/SSI-2025-0164

Feedback mechanism in spacecraft orbital threat games

Li YUAN[1,2], Chanying LI[3,4*], Yan WEI[5], Tiantian JIANG[2,6], Panxing HUANG[2,6] & Chunling WEI[2,6]

1. *China Academy of Space Technology, Beijing 100094, China*
2. *National Key Laboratory of Space Intelligent Control, Beijing 100094, China*
3. *State Key Laboratory of Mathematical Sciences, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China*
4. *School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China*
5. *School of Mathematics, Sichuan University, Chengdu 610065, China*
6. *Beijing Institute of Control Engineering, Beijing 100094, China*
* Corresponding author. E-mail: cyli@amss.ac.cn

Abstract Autonomous spacecraft operations in orbit are increasingly challenged by orbital threats such as debris collisions, close-range reconnaissance, and malicious interference. The avoidance process in such scenarios constitutes a dynamic game under uncertain adversarial behaviors. Influenced by environmental constraints, sensing limitations, communication disturbances, and stealth technologies, the information structure is both imperfect and incomplete, resulting in a dual dilemma for game-theoretic decision-making: significant uncertainty in adversarial actions and high sensitivity of strategies to situational dynamics. Additionally, the dimensionality of the strategy search space increases exponentially with the complexity of information, which, under the stringent constraints of onboard computational resources, often forces autonomous decision-making systems to adopt overly conservative strategies or compromises that may jeopardize spacecraft safety. To address these challenges, this study proposes a feedback-based strategy design framework featuring a vertically layered and horizontally adjustable architecture. By dynamically optimizing the game information structure in real time based on situational features, the method adaptively reduces the search space and improves strategy performance, thereby ensuring the security of space assets and satisfying the operational demands of on-orbit missions.

Keywords orbital game, imperfect information, incomplete information, autonomous decision-making, feedback mechanism

https://www.sciengine.com/doi/10.1360/SSI-2025-0164