

航天器轨道追逃动力学与控制问题研究综述

朱彦伟*, 张乘铭, 杨傅云翔, 杨乐平

(国防科技大学 空天科学学院, 湖南 长沙 410073)

摘要:随着航天器交会与接近操作技术的快速发展,轨道追逃问题逐渐成为航天领域的研究热点。从动力学与控制视角,对航天器轨道追逃问题的研究现状进行综述。给出了基于定量微分对策的轨道追逃问题模型的一般形式,系统梳理了各种类型的轨道追逃问题;对于追逃策略求解,分别针对闭环策略和开环策略,分析了各种方法的优缺点;围绕人工智能算法与轨道追逃问题的结合,阐述了基于深度神经网络和强化学习的轨道追逃策略的研究现状。关于未来展望,提出了追逃博弈态势分析、多航天器博弈控制、三体条件下博弈动力学与控制等发展方向。

关键词:航天器追逃博弈;微分对策;深度神经网络;强化学习

中图分类号:V448.2 文献标志码:A 开放科学(资源服务)标识码(OSID):

文章编号:1001-2486(2024)03-001-11



听语音
与作者
聊科研

Survey on dynamics and control problem research in spacecraft orbital pursuit-evasion game

ZHU Yanwei*, ZHANG Chengming, YANG Fuyunxiang, YANG Leping

(College of Aerospace Science and Engineering, National University of Defense Technology, Changsha 410073, China)

Abstract: With the rapid development of spacecraft rendezvous and proximity operation technology, the problem of orbital pursuit-evasion has gradually become a research hotspot in the aerospace field. From the perspective of dynamics and control, the research status of spacecraft orbital pursuit-evasion was reviewed. General form of the orbital pursuit-evasion problem model based on quantitative differential games was given, and various types of orbital pursuit-evasion problems were systematically sorted out. For the solution of pursuit and escape strategies, the advantages and disadvantages of various methods were analyzed for closed-loop strategy and open-loop strategy. Focusing on the combination of artificial intelligence algorithm and orbital pursuit and escape problem, the research status of orbital pursuit and escape strategy based on deep neural network and reinforcement learning was expounded. Regarding future prospects, development directions has been proposed, including the pursuit-evasion game situation analysis, the multi-spacecraft game control, the game dynamics and control under three-body problem.

Keywords: spacecraft pursuit-evasion games; differential games; deep neural networks; reinforcement learning

随着太空技术的快速发展及其在政治、经济、文化和军事等各领域的广泛应用,太空日益成为国际战略竞争新的制高点。根据美国安全世界基金会发布的2023年度 *Global Counterspace Capabilities* 开源评估报告,航天器交会与接近操作 (rendezvous and proximity operations, RPO) 技术已经成为太空技术发展的重点,近年来国际上已进行了多次技术测试与演示验证^[1]。其中,美国于2014年开始实施的地球同步轨道空间态势感知计划 (geosynchronous space situational awareness program, GSSAP) 卫星,漂移运行在地球同步轨道 (geosynchronous orbit, GEO) 上下两侧,可观测运行在 GEO 轨道的卫星。根据俄罗斯数据显示,

GSSAP 卫星已经实施了数百次机动,并对多个国家的几十颗 GEO 卫星实施了秘密靠近或抵近操作。此外,根据美国的相关声明显示,除了 GSSAP 项目以外,美国还正在实施 ANGELS、EAGLE、Mycroft、S5 等交会与接近操作试验验证项目^[2]。

在航天器交会与接近操作过程中,一般将作为操作对象的航天器称为目标航天器,将主动接近操作对象的航天器称为追踪航天器。按照轨迹要求、导航方式以及操作条件的不同,一个完整的航天器交会与接近操作任务可以分为远程引导、近程引导、交会逼近和接近操作等阶段。目标航天器通常可分为合作、非合作非机动和非合作机动三类。当目标航天器为非合作机动目标且具备

决策能力时,交会与接近操作问题就转化为航天器轨道追逃问题。此时,追踪航天器和目标航天器则分别作为追踪方和逃逸方,双方航天器围绕各自目标并依托自身机动能力展开博弈,可见该问题实质上是一个双边控制的连续动态对抗的问题^[3]。

一般地,航天器轨道追逃问题研究以微分对策理论^[4]为基础建立追逃模型,将对手的策略纳入自身策略的制定过程中,更好地描述了具有强对抗性的问题场景。自 20 世纪 Issacs 提出微分对策理论以来,许多围绕航天器追逃问题的研究相继展开^[5-9]。微分对策理论可分为定性微分对策和定量微分对策两类。在定性微分对策中,追逃双方关注博弈结果能否实现,侧重于对当前追逃态势的评估,相关研究工作^[8,10-12]较少,这是因为求解定性微分对策问题的关键在于将状态空间划分为捕获区与逃逸区,需要构造复杂界栅,当其与轨道动力学模型结合时计算难度将大大增加。而在定量微分对策中,追逃双方更加关注追逃策略的求解,相关方面吸引了大量研究人员。但在获得策略过程中往往需要求解边值问题,从而面临初值敏感性强、计算效率低等问题。近年来,人工智能技术的迅猛发展,为求解轨道追逃策略提供了新的可行方案。在“航天+智能”思路的牵引下,深度神经网络和强化学习技术不断地被应用于追逃策略求解,显著提高了问题求解效率,对提高航天器的自主化和智能化水平具有重要意义。

本文在对航天器轨道追逃问题模型及其求解方法进行评述的基础上,重点关注了深度神经网络和强化学习在航天器轨道追逃问题中的应用进展。最后,给出了航天器轨道追逃未来的可能研究方向。

1 轨道追逃的问题模型

基于定量微分对策的轨道追逃问题模型包含动力学模型、局中人个数、信息结构、博弈终止条件、优化目标、均衡策略等诸多要素。在此模型中追逃双方地位是均等的,所求解的均衡策略对追逃双方互为最优,即当任何一方单方面地偏离均衡策略都会导致其处于更加不利的处境,因此适用于对轨道追逃问题中的双方策略进行求解。

目前,绝大多数研究是基于完全信息、连续有限推力的“一对一”追逃问题。完全信息为一种理想状态,即动力学建模、状态测量以及博弈目标和控制量等信息对双方均为已知的和确定的。问

题模型的一般形式^[13]如下

$$\begin{cases} \dot{\mathbf{x}}_i = \mathbf{f}_i(\mathbf{x}_i, \mathbf{u}_i, t) & i = p, e \\ \Psi(\mathbf{x}_{p0}, \mathbf{x}_{e0}, \mathbf{x}_{pf}, \mathbf{x}_{ef}, t_f) = \mathbf{0} \\ J = \phi(\mathbf{x}_{p0}, \mathbf{x}_{e0}, \mathbf{x}_{pf}, \mathbf{x}_{ef}, t_f) \end{cases} \quad (1)$$

其中:p 表示追踪方,e 表示逃逸方, t_f 为博弈终止时间, \mathbf{x} 为状态向量, \mathbf{u} 为控制向量;第一式为动力学模型,第二式为博弈终止条件,第三式为博弈优化目标。问题的核心是求解鞍点控制策略,即双方作为完全理性的博弈者应采取的一对互为最优的策略,任何一方单方面偏离这一策略时将使自身处于更不利的处境,可表示为 $\{\mathbf{u}_p^*, \mathbf{u}_e^*\}$ 。

鞍点控制策略通常基于定量微分对策求解,引入协态变量 λ 和哈密顿函数 H ,则有

$$\begin{cases} H = \lambda_p^T \mathbf{f}_p + \lambda_e^T \mathbf{f}_e \\ \dot{\lambda}_i = -\frac{\partial H}{\partial \mathbf{x}_i} \\ \Phi = \phi + \nu^T \Psi \\ \lambda_i(t_f) = \frac{\partial \Phi}{\partial \mathbf{x}_{if}} \\ H(t_f) = -\frac{\partial \Phi}{\partial t_f} \end{cases} \quad (2)$$

其中, ν 为拉格朗日乘子,第二式为协态变量微分方程,第三式为广义终端条件,第四式为协态变量终端条件,第五式为横截条件。

鞍点控制策略需满足

$$\begin{cases} \mathbf{u}_p^* = \arg \min_{\mathbf{u}_p} H \\ \mathbf{u}_e^* = \arg \min_{\mathbf{u}_e} H \end{cases} \quad (3)$$

若 t_f 不定,则为生存型或自由时域型追逃问题;若 t_f 固定,则为固定时间追逃问题。特别地,当动力学模型为线性系统且博弈目标函数为二次型函数时,固定时间追逃问题为线性二次型微分对策问题^[3]。下面将根据博弈模型中的要素对当前的研究工作进行总结和分类。

1.1 按动力学模型分类

考虑追逃任务场景中追逃双方的相对距离对动力学模型以及相应求解算法的设计具有显著影响,因此可以将追逃问题分为近距离追逃和远距离追逃两种情况。

近距离追逃是指追逃双方初始距离远小于双方轨道半长轴的追逃场景,基于相对坐标系(如 LVLH 坐标系)进行动力学建模更便于问题研究,航天器的状态量由 $\mathbf{X} = (x, y, z, \dot{x}, \dot{y}, \dot{z})$ 进行描述,其中 (x, y, z) 和 $(\dot{x}, \dot{y}, \dot{z})$ 分别为航天器在相对坐标系中的位置和速度。当追逃双方位于近圆轨道附近且偏心率较小、追逃时间较短时,其运动过程通

常由 C-W 方程描述,祝海^[12]、Stupik 等^[14]、Woodbury 等^[15]、Ye 等^[16-17]均基于 C-W 方程开展了相关研究工作。但当追逃时间长、逃逸方轨道偏心率较大时,需要更为精确的相对运动动力学模型。在此方面,孙松涛^[18]基于精确相对运动方程研究了航天器追逃问题;当参考轨道为椭圆时,Prince 等^[19]采用 T-H 方程进行了研究工作,Zhang 等^[20]应用基于 C-W 方程的近似相对运动方程^[21]研究了此类问题。

在远距离追逃问题中,追逃双方的初始相对距离较远,相对运动动力学模型不再适用,需要基于绝对运动动力学模型对两航天器的动力学过程进行描述。与相对运动不同,绝对运动模型通常基于地心惯性坐标系进行建模,航天器状态量的选择不再拘泥于位置和速度,而是根据问题研究需要进行选择,主要有两种方法:一种是利用球坐标对航天器状态进行描述,另一种是在笛卡儿坐标系中进行描述。在基于球坐标的研究工作中,Pontani 等^[13]选择状态量 $\mathbf{X} = (r, v, \gamma, \xi, \phi, \zeta)$ 对追逃航天器进行描述,其中, r 和 v 分别为航天器的瞬时位置大小和瞬时速度大小, γ 表示航天器的航迹角, ξ 表示航天器的经度, ϕ 为航天器的纬度, ζ 为速度方位角;王强等^[22]在此基础上进一步研究了含 J2 摄动力的追逃问题。在应用笛卡儿坐标系进行建模方面,Shen 等^[23]在惯性系下将 $\mathbf{X} = (x, y, z, \dot{x}, \dot{y}, \dot{z})$ 作为航天器的状态量进行描述,其中 (x, y, z) 和 $(\dot{x}, \dot{y}, \dot{z})$ 分别为航天器在惯性系中的位置和速度。Zeng 等^[24]对比了两种建模方法,其数值计算结果说明:在采用球坐标求解时间问题维度更低,在计算效率上具有一定优势;在采用笛卡儿坐标系进行建模时,收敛速度较慢但求解精度更高。

1.2 按博弈终止条件分类

博弈终止条件又称为博弈终端曲面,决定了追逃双方在达到何种情况时结束博弈过程并进行胜负判定。在基于线性二次型微分对策理论进行博弈建模时无法施加明确的终止条件,一般仅以时间作为终止条件。在时间自由的博弈问题中,根据任务不同可以分为位置匹配型和位置速度约束型。一般认为追踪航天器成功拦截逃逸方时博弈终止(即追逃双方位置重合而对速度无要求),但此类约束条件易导致追逃双方速度相差较大从而发生在轨碰撞。在一对一追逃问题中,除上述以拦截作为问题终止条件外,常燕等^[25]运用微分对策对共面机动目标交会问题进行研究,并采用了非线性规划算法得到了追逃双方的最优轨迹。

Innocenti 等^[26]基于状态相关系数参数化和状态相关黎卡提方程(state-dependent Riccati equation, SDRE)方法求解了非线性动力学条件下航天器交会过程中非零和博弈的纳什均衡解。Prince 等^[19]基于微分对策理论分别研究了拦截、交会、沿太阳矢量方向、能量匹配等六种终止条件下均衡策略的推导,并应用遗传算法(genetic algorithm, GA)和粒子群优化(particle swarm optimization, PSO)算法进行了策略求解。Venigalla 等^[27]以经典轨道六要素视角对脉冲推力方式的交会型问题进行了研究,使得航天器的无动力飞行阶段得以高效利用。

与一对一问题不同,在多对一追逃问题中,由于存在多个追踪航天器,因此博弈终端条件需要进行特殊设计。史帅科^[28]分析了追捕卫星数量和初始位置分布对围捕任务的影响,明确了完成围捕任务的必要条件。刘彦昊等^[29]设计了虚拟中心距离项、围捕半径项、卫星避碰项等函数,共同组成了适应度函数以评估最终包围态势。可以看出在多对一问题中,通常需要基于几何视角对博弈终止条件进行设计。

1.3 按博弈优化目标分类

在博弈过程中,追逃双方分别有自身的目标函数,当两者目标函数相同而仅优化方向相反时即为零和博弈问题,否则为非零和博弈问题^[30]。根据目标函数的不同可以分为时间最优、能量最优、速度冲量最优、终端时刻相对距离最短等不同类型的问题。孙松涛等^[18,31]和 He 等^[32]均以二次型的终端相对距离为目标函数给出了控制策略。Ye 等^[16]通过实时计算博弈剩余时间,以终端脱靶量作为目标函数。Jagat 等^[33]针对固定时间博弈问题设计了二次型目标函数,综合考虑相对位置和燃料消耗等因素。此外,还有许多研究工作选择时间作为目标函数,Pontani 等^[13]、Shen 等^[23]、Prince 等^[19]的研究工作中仅对博弈终止条件进行了明确,其博弈目标为尽可能缩短或延长达成终止条件的时间。当航天器为脉冲推力方式时,还需要考虑总速度冲量这一因素,Venigalla 等^[27]在研究远距离以交会为目标的航天器博弈问题时,基于消耗的速度冲量设计了目标函数。

1.4 不完全信息追逃问题

对于不完全信息追逃问题,Aures-Cavalieri^[34]提出了行为学习技术估计对手的策略,并将追逃问题转化为一个单边最优控制问题。该方法在终端时间固定、无限时域和终端时间自由的问题中

都得到了验证。Li 等^[35]基于线性二次型微分对策模型,在假设追踪航天器和目标航天器支付函数未知的情况下,采用在线参数估计的方法提出了最优逃逸策略。Tang 等^[36]应用平滑变结构滤波器研究了不完全信息条件下航天器博弈策略切换问题,实现了对目标状态的估计。周俊峰^[37]针对不完善信息追逃和不完整信息追逃问题均进行了深入的研究工作,分别基于不确定集理论和多模型行为估计方法进行了策略求解。Zhang 等^[38]提出了一种基于多项式逼近的半解析航天器目标状态估计方法,可大大减少估计逃逸目标状态时的计算量。Wang 等^[39]设计了一种自适应感知和交互自主博弈控制方法,从而可以放宽对逃逸航天器目标函数这一信息的要求。Zheng 等^[30]研究了不完全信息情况下的非零和追逃博弈控制问题,首先设计了一种控制增益估计器,并求解离散时间黎卡提方程得到了追踪方的控制策略。当前关于不完全信息追逃博弈的研究工作主要对基于线性动力学问题进行了研究,一般思路为受限设计观测估计器对手相关信息进行估计,进一步基于估计结果设计己方的反馈控制律。

1.5 其他追逃问题

以上工作中多假设航天器机动方式为连续有限推力,此种推力模式下的研究工作较为广泛,而对于脉冲机动追逃问题,无法应用微分对策理论建立对策模型进行分析求解,相关研究工作仍较少。但脉冲推力可以有效利用推力间隔时间进行充分的无动力飞行,对减少燃料消耗和提高航天器在轨寿命具有显著意义。在此方面,Venigalla 等^[27,40]利用可达域理论对采用特定机动方式的航天器远距离追逃问题进行了研究,分别对逃逸方已知和未知追踪方机动能力的情况进行了分析。于大腾^[41]从能量和测量两个方面设计了航天器规避机动策略的生成方法,提出了潜在威胁区和完全不可观测机动等规避机动指标,具有较好的现实参考价值。

有学者还开展了多航天器追逃的相关研究工作。其中较为典型的有追-逃-防问题、多对一追逃问题等。Liu 等^[42]基于模糊评价和纳什均衡理论提出了一种分布式在线任务规划算法对含有多个航天器的追-逃-防问题进行了研究。进一步,Liu 等^[43]继续研究三航天器追-逃-防问题,其中目标航天器无机动力且追踪航天器和防御航天器仅进行一次机动,提出了一种基于粒子群

算法和牛顿插值法的混合算法。Zhou 等^[44]考虑了连续小推力下的三航天器追-逃-防问题,并采用多重打靶法和遗传算法对其进行求解分析。许旭升等^[45]研究了集群航天器共同围捕单个目标的问题,提出了几种典型博弈行为。当前关于脉冲推力追逃和多航天器追逃问题的研究通常仅针对特定的问题场景和约束条件,如何建立更为成熟的问题模型仍需进一步探索。

综上所述,航天器轨道追逃的问题模型依据研究的侧重点不同有多种形式,微分对策理论是建模的基础,鞍点策略求解是问题的核心。

2 追逃策略的求解方法

轨道追逃问题中动力学规律不直观,这与传统的无人机追逃、小车追逃和弹目追逃等问题不同,需要将其作为一种特殊的追逃问题进行研究。当前许多研究工作围绕鞍点策略求解展开,根据求解得到的追逃策略是否具有反馈结构可以分为闭环策略和开环策略^[12]。

闭环策略是在求解固定时域问题或无限时域问题时获得的,其求解过程一般要求动力学模型为线性模型,且追逃目标函数为二次型形式,称为线性二次型微分对策,其状态方程^[46]可以表示为

$$\dot{X}(t) = AX(t) + B_p U_p(t) - B_e U_e(t) \quad (4)$$

式中, A 和 B 为相对运动方程中的系数矩阵, U 为航天器的输入控制变量。其中对于固定时域问题,目标函数可以表示为

$$J = \frac{1}{2} X(t_f) Q_{pf} X(t_f) + \frac{1}{2} \int_{t_0}^{t_f} [X^T(t) Q_p X(t) + U_p^T(t) R_p U_p(t) - U_e^T(t) R_e U_e(t)] dt \quad (5)$$

式中, Q 和 R 为支付函数的系数矩阵。

此类问题通常需要将其转化为黎卡提微分方程进行求解。Jagat 等^[47]针对线性动力学近距离追逃问题进行了研究,推导了相应的反馈控制策略。李振瑜^[48]以线性 C-W 方程为基础,针对含有观测噪声、观测时延和未知支付信息的问题分别进行了研究。为了进一步将闭环策略应用于求解非线性微分对策问题,Jagat 等^[33]通过应用 SDRE 方法对标准线性黎卡提微分对策理论进行拓展,其仿真结果表明该方法相较于文献[47]中的线性解更具优势。闭环策略与博弈过程中的瞬时状态相关,所以便于通过实时状态测量进行在线更新,更适用于信息结构不完整的追逃场景。但由于缺少追逃双方终端状态的约束,通常只能实现对逃逸航天器的尽可能接近,而无法确保终

端时刻位置的匹配。

开环策略是在求解自由时域问题中获得的,此类问题中将追逃时间作为目标函数展开博弈,且要求终端时刻双方满足约束条件(如拦截条件等),目标函数^[13]表示为

$$J = t_f \quad (6)$$

此类问题类似轨迹优化,其求解方法可分为间接法、半直接法和直接法等。比较而言,间接法和半直接法的相关研究较为丰富,而直接法研究相对较少。

间接法求解需要基于庞德里亚金极大值原理推导得到最优必要性条件,将原问题转化为两点边值问题并应用计算密集型算法进行求解,问题维度和初值敏感性较高,难以快速收敛。Stupik等^[14]针对近距离追逃场景,利用基于内插法和外插法的Kriging软件进行求解,实现了基于粒子群优化算法的实时控制,但仅对共面问题进行了验证,说明此方法仍具有局限性。Hafer等^[49]在求解远距离追逃问题中,提出了一种敏感性分析方法,得到状态敏感性矩阵和约束条件敏感性矩阵,再利用同伦法从无重力问题求解延拓到对真实问题的求解,有效提高了计算效率。王强等^[22]研究了航天器在受到J2非球形摄动情况下的远距离轨道追逃问题,应用基于序列二次规划的混合方法对问题进行了求解。Shen等^[23]研究了考虑质量变化的远距离追逃问题,提出了以忽略逃逸方推力的单边优化问题结果作为追逃初始解的方法;Shi等^[50]借鉴这一思路进一步研究了含J2摄动且推力方向存在角度限制的问题,并分析了推力方向约束对博弈终止时间的影响。Zeng等^[24]通过引入权值向量和扩展边界条件将两点边值问题转换为优化问题,并设计了混合优化算法进行求解,求解效率相比差分进化算法有明显提高。

为了避免求解间接法中的高维两点边值问题,有学者提出了半直接法的思想^[51-52]。在半直接法中,首先基于推导得到的最优必要性条件将原双边问题转化为单边最优控制问题,然后应用非线性规划算法进行求解。Pontani等^[13]利用半直接法有效降低了远距离追逃问题的维度,并应用配点法和遗传算法进行求解,具有良好的收敛性。孙松涛^[18]进一步研究了半直接控制量参数法和混合法,减少了计算复杂度,具有更好的计算效率。Carr等^[53]在研究航空器追逃博弈时,同样应用半直接法的思想得到了一方为比例导引律时的近似最优策略,仿真结果说明该方法具有良好

的准确性和最优性。

基于半直接法的思路,进一步诞生了直接法。在直接法中,不再需要推导问题的最优必要性条件,而是通过迭代求解两个单边的最优控制问题直至收敛从而得到原问题的解,因此当前该方法仍存在理论上的不足,即所获得的解无法证明为原问题的鞍点解。Schoenwetter^[54]通过在每次迭代求解过程中固定一方策略并求解另一方的最优策略,直到双方的策略趋于收敛。张乘铭^[55]在研究含J2摄动的轨道追逃问题时应用伪谱法进行迭代求解,其仿真结果说明应用直接法得到的鞍点策略与精确解仍存在差距,这也是直接法求解时普遍面临的最优性问题。

随着人工智能的不断发展,不少学者试图将人工智能应用于航天器轨道追逃问题,大体可以分为深度神经网络(deep neural network, DNN)和强化学习两种思路。其中,有关深度神经网络的研究仍基于微分对策理论展开,重点关注策略求解算法的设计;有关强化学习的研究则基于马尔可夫决策过程,通过应用深度神经网络完成对控制器的设计,从而使智能体在与环境交互中不断学习博弈策略。

综上所述,当前闭环策略无法处理控制约束和终端约束,应用场景较为受限;而开环策略求解计算量大,难以在线应用。人工智能算法的“训练—测试—应用”特点,使得离线训练、在线应用成为可能,逐渐成为研究热点。

3 基于DNN的轨道追逃策略研究

由于DNN具备很强的数据函数映射关系的拟合能力,近几年已有越来越多的学者关注DNN在航天器轨迹优化等方面的应用^[56-59],可有效提高问题求解效率。航天器轨道追逃的传统求解方法面临航天器平台计算能力不足和实时性要求高的矛盾,应用DNN代替开环鞍点策略中的计算密集算法具有广阔的应用前景。当前应用DNN直接求解航天器追逃问题的研究工作仍处于起步阶段,一般可总结为数据驱动和物理驱动两种框架。

3.1 数据驱动框架

数据驱动框架使用传统监督学习模式,基于神经网络强映射能力,使用数据集学习特定映射,分为数据集生成、数据预处理、网络设计、模型训练验证和网络应用五个阶段,如图1所示。

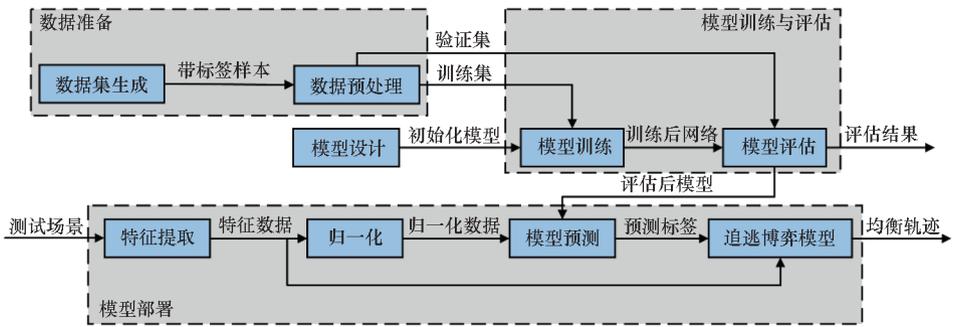


图 1 DNN 数据驱动框架

Fig. 1 DNN data-driven framework

数据集生成阶段负责网络训练必需的原始数据集生成。在存在可用真实数据的情况下,该部分只需要将数据按照特定格式进行整理;真实数据不存在时,则需要仿真生成数据再进行整理。航天器追逃问题属于后者,其使用的数据格式如下

$$data = [input\ data; label]^T$$

其中: *input data* 表示网络的输入变量,通常使用追逃双方的状态信息;而 *label* 则表示网络输出,需要与情况进行匹配。数据预处理阶段对于提高学习效果有较高影响,一般分为归一化与数据分类两步,前者是为了控制数据范围,后者则是将样本数据随机分为训练数据与验证数据。网络设计阶段用于实现“状态—轨迹”映射的网络,目前该部分包含大量的人工干预并且需要相应的机器学习背景知识作为支撑。模型训练验证阶段先训练网络学习中输入与输出之间的映射关系,再通过验证数据检验,最后输出一个完成学习的网络,至此网络才掌握解决该问题的相关知识,可以实现快速计算。网络应用阶段直接使用训练后网络解决实际问题,主要分为三步:①对测试数据进行归一化;②输入网络计算结果;③将网络输出输入动力学模型得到最终轨迹。

在数据驱动框架下,吴其昌等^[60]针对无限时域问题,将微分对策模型的求解结果作为训练集,使用神经网络学习“状态—协态”映射关系,并通过仿真结果证明该方式可以极大程度地提高求解效率,但轨迹误差随时间的累积效应明显。Zhang等^[20]针对椭圆参考轨道附近的自由时域轨道追逃问题,设计了相应的基于DNN的鞍点策略求解方法,在其网络应用部分,将神经网络预测输出作为数值优化算法的初始条件以快速获得精确结果,从仿真结果可以看出其算法在计算稳定性和效率方面均具有显著优势。

数据驱动框架的核心在于大量高质量训练数据,样本数据的采样范围与质量决定了框架的泛

化性能。对于超出采样范围的场景,该框架并不能有效处理。

3.2 物理驱动框架

为了降低网络的学习难度并且使训练过程满足物理约束,可将轨道动力学等物理约束作为先验知识事先赋予网络,由此得到物理驱动框架,如图2所示^[61]。这是一种新的求解思路,目前相关研究尚不多见。

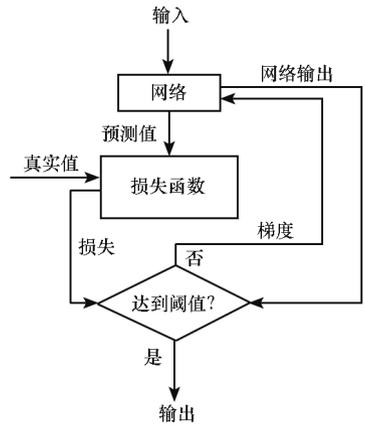


图 2 物理驱动框架

Fig. 2 Physics-driven framework

Yang等^[61]针对近距离轨道追逃问题,提出了一种基于神经重参数化^[62]的求解思路:基于微分对策理论将其建立为一个两点边值问题,进一步设计了一种新型网络,包含参数化、物理模型以及约束三个部分,如图3所示。其中被称为物理模型的次级结构仅用于记忆物理约束,而相关的权重与偏倚并不会因训练过程而更新。最后通过仿真算例证明了该方法的有效性,且该网络模型可以在没有训练数据的情况下使用。

物理驱动框架通过向网络中添加物理约束等先验信息,成功实现了小样本或无样本情况下的训练。但嵌入物理约束的适用范围也影响了框架的泛化性能。网络结构和损失函数都需要根据具

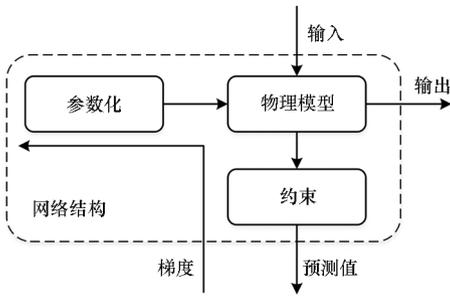


图3 网络结构

Fig.3 Network structure

体问题进行专门设计。

综上,数据驱动框架具有极高的通用性,但网络需要同时学习样本数据中包含的映射关系以及轨道动力学等物理约束。这极大增加了网络的训练负担,并导致网络复杂度的提升;另外,即使在完成训练之后,其计算结果也难以完全符合物理学约束。物理驱动框架虽然降低了网络的学习负担,满足了物理约束,但缺乏通用性,需要针对特定问题专门设计相应的网络结构以嵌入先验知识。如何将数据驱动与物理驱动相结合,让模型在满足物理约束的前提下具有一定的通用性,可以作为未来此类方法的一个研究方向。

需要说明的是,当前基于 DNN 的轨道追逃策略基本上都是针对连续有限推力假设下的追逃博

弈问题,结合微分对策理论和 DNN 开展研究,无法应用到脉冲推力假设下的追逃博弈问题。

4 基于强化学习的轨道追逃策略研究

强化学习是智能体在与其所处环境的不断交互中进行学习的一种方法,并以最大化长期收益为目标获得最优策略。在强化学习的过程中,智能体通过与环境进行频繁的交互不断学习更新自己的策略,智能体必须根据当前的状态,通过不断尝试选择动作并发现哪种动作能获得更高的奖励,然后以此为依据来更新自己的策略^[63]。上述过程通常采用马尔可夫决策过程进行建模,并将其作为强化学习的基础。近年来,强化学习在航天器动力学与控制中的研究吸引了不少学者的关注^[57-58]。

在当前针对航天器轨道追逃问题的研究工作中,通常仅关注一方的控制策略,而将另一方航天器作为“环境”的一部分,所以其研究的问题对象更接近于一个面向机动非合作目标的接近控制问题,其基本求解框架如图4所示。图中“生成先验信息”为基于微分对策理论生成先验信息的过程,这一过程有助于加速神经网络的收敛,但同时可能造成网络收敛至局部最优解,因此这一步并非必要的,可根据问题需要灵活选择。

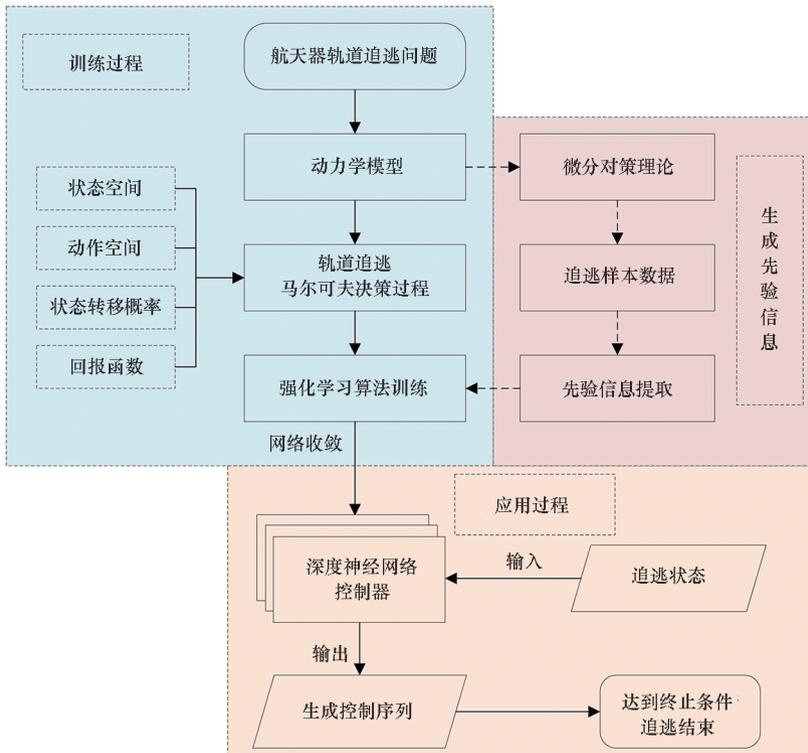


图4 强化学习算法求解追逃问题基本框架

Fig.4 Basic framework of reinforcement learning algorithm to solve the pursuit-evasion problem

Zeng 等^[64] 针对共面轨道追逃问题, 建立了一种马尔可夫决策过程模型, 如图 5 所示。在图 5 中, 在第 k ($0 \leq k < n$) 步时的当前状态为 x_k , 动作为 u_k , 转移到下一状态为 x_{k+1} , 所获得的奖励值为 r_k 。当追踪航天器与目标航天器达到博弈终止条件时, 判定马尔可夫链结束。此时求解该追逃问题, 即为训练智能体搜索得到一个最优确定性策略 π^* , 使得初始状态值函数 $V_0^*(x_0)$ 达到最小。为此, 文献[64]设计了一种“经验深度确定性策略梯度搜索”的强化学习算法, 该算法相较于传统确定性策略梯度搜索算法的区别在于: 在正式训练之前将先验信息提前注入了深度神经网络, 从而一定程度上解决了“维数灾难”^[65] 的问题。

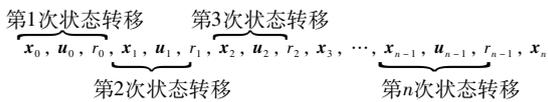


图 5 马尔可夫链

Fig. 5 Markov chain

此外, 刘冰雁等^[66] 针对非合作目标的空间交会问题, 构建了空间行为模糊推理模型并引入了多组并行的网络分支和共享行为决策模块, 从而提出了分支深度强化学习的方法, 对此类问题求解具有一定的借鉴意义。许旭升等^[45] 探索建立了包含多智能体的强化学习追逃博弈环境, 并给出了多种典型的协同追逃博弈场景, 但并未给出详细的数值研究结果。Yang 等^[67] 将追逃过程划分为远距离交会阶段和近距离博弈阶段, 并采用深度强化学习方法对近距离博弈问题中的不完全信息脉冲机动博弈问题进行了求解。Zhang 等^[68] 针对二维平面内的轨道追逃博弈场景, 应用深度强化学习算法解决了初始状态位于捕获区外的问题。袁利等^[69] 设计了一种多阶段强化学习算法, 在第一阶段基于逻辑规则对双方的追逃策略进行设计, 并预训练网络与逻辑规则进行博弈, 第二阶段中双方预训练后的网络模型互相博弈直至稳定, 有效改善了算法的收敛性和可解释性。耿远卓等^[70] 在近端策略优化 (proximal policy optimization, PPO) 算法框架的基础上, 通过预测博弈双方在博弈结束时刻的相对位置改进了奖励函数的设计方法, 有效解决了脉冲推力条件下的追逃博弈问题。Zhao 等^[71] 提出了一种设计预测—奖励—检测训练框架对多智能体强化学习算法进行了改进, 有效获得了脉冲推力条件下的轨道博弈机动策略。

综上, 基于强化学习的航天器轨道追逃策略

求解算法在计算效率上具有明显优势, 但由于问题维度较高, 往往面临训练过程中网络较难收敛的问题。此外, 由于缺少严格的数学证明, 基于强化学习算法得到的策略无法保证其最优性, 通常只能将其作为近似最优解。

因此, 基于强化学习的算法在解决轨道面内追逃问题时求解稳定性较好, 而当问题维度拓展至三维空间, 维度较高的环境空间容易导致训练过程中神经网络无法收敛。为了使训练的网络更好收敛, 要求设计的回报值函数可以充分反映追逃博弈过程中双方动作的收益, 这一步是将轨道动力学知识深入嵌入强化学习算法的关键, 也是当前应用强化学习算法求解航天器追逃问题的主要难点。虽然在应用强化学习算法进行求解时仍存在上述困难, 但考虑到其基于马尔可夫决策理论进行问题建模的特点, 因此其在解决更多类型航天器追逃博弈问题中具有更广泛的应用潜力。

5 总结与展望

航天器轨道追逃问题是近年来航天领域的动力学与控制研究热点。本文针对近地空间航天器轨道追逃策略, 尤其是深度神经网络和强化学习算法在其中的应用研究进行了综述。从中可以看出, 人工智能技术有效改善了轨道追逃问题的求解效率, 在提高航天器主动安全的自主化和智能化方面具有良好的应用前景。但是, 目前的研究大多是对追逃策略的定量求解, 且“一对一”追逃为主, 与实际轨道追逃应用还有一定差距。此外, 人类航天活动正从近地空间向地月空间拓展, 地月空间轨道追逃将逐渐走向现实, 值得专门研究。当前, 尚未查到有关地月空间轨道追逃问题的研究成果。

因此, 关于航天器轨道追逃问题下一步可着重从以下三个方面进行深入研究:

1) 追逃博弈态势分析与评估。随着航天器轨道追逃问题研究的不断深入, 追逃态势分析作为具体任务规划的前置任务与可行性支撑, 迫切需要展开相应研究。一方面可以从博弈局中人的个体可达集出发, 通过状态空间中可达集的相互关系对博弈状态进行研判; 另一方面可以从博弈场景出发, 基于定性微分对策的相关理论建立态势评估模型, 通过求解评估模型获取局中人各自在状态空间的优势域。

2) 多航天器博弈控制。随着在轨航天器机动能力和集群航天器技术的发展, 包含多个航天器甚至大规模航天器集群的追逃场景逐渐成为可

能,迫切需要针对此类问题展开研究。一方面可以尝试基于微分对策理论建立此类追逃问题模型并应用传统方法进行求解;另一方面可以探索利用多智能体强化学习算法,通过与虚拟动力学环境交互训练得到相应控制策略。

3) 三体条件下博弈动力学与控制。随着人类探索太空的步伐逐渐从近地空间拓展到地月空间,二体假设下的追逃问题也逐渐拓展至地月三体空间中。与近地空间不同,地月空间轨道普遍存在周期长、混沌性强,对各类扰动高度敏感,周期性与拟周期性融合等特点,目标跟踪难度大。与二体轨道相比,动力学模型更加复杂,由“易攻难守”变为“易守难攻”,使得追逃问题控制策略难度大大增加。

参考文献 (References)

- [1] Secure World Foundation. Global counterspace capabilities[R/OL]. [2022-08-18]. https://swfound.org/media/207541/swf_global_counterspace_capabilities_2023_es_en.pdf.
- [2] 刁华飞, 张雅声. 美国高轨态势感知卫星能力分析[J]. 航天电子对抗, 2019, 35(4): 48-51.
DIAO H F, ZHANG Y S. Analysis of US geosynchronous situational awareness satellite capability [J]. Aerospace Electronic Warfare, 2019, 35(4): 48-51. (in Chinese)
- [3] 罗亚中, 李振瑜, 祝海. 航天器轨道追逃微分对策研究综述[J]. 中国科学: 技术科学, 2020, 50(12): 1533-1545.
LUO Y Z, LI Z Y, ZHU H. Survey on spacecraft orbital pursuit-evasion differential games [J]. Scientia Sinica (Technologica), 2020, 50(12): 1533-1545. (in Chinese)
- [4] 查文中. 单个优势逃跑者的多人定性微分对策研究[D]. 北京: 北京理工大学, 2016.
ZHA W Z. Multi-player qualitative differential games with single superior evader [D]. Beijing: Beijing Institute of Technology, 2016. (in Chinese)
- [5] MENON P K A, CALISE A J. Guidance laws for spacecraft pursuit-evasion and rendezvous [C]//Proceedings of the Guidance, Navigation and Control Conference, 1988.
- [6] GUTMAN S, ESH M, GEFEN M. Simple linear pursuit-evasion games [J]. Computers & Mathematics with Applications, 1987, 13(1/2/3): 83-95.
- [7] KELLEY H J, CLIFF E M, LUTZE F H. Pursuit/evasion in orbit[J]. Journal of the Astronautical Sciences, 1981, 29: 277-288.
- [8] ANDERSON G M, GRAZIER V W. Barrier in pursuit-evasion problems between two low-thrust orbital spacecraft [J]. AIAA Journal, 1976, 14(2): 158-163.
- [9] REPPERGER D W, KOIVO A J. Optimal terminal rendezvous as a stochastic differential game problem [J]. IEEE Transactions on Aerospace and Electronic Systems, 1972, AES-8(3): 319-326.
- [10] 张秋华, 赵小津, 孙毅. 空间飞行器在视线坐标系中的追逃界栅[J]. 航天控制, 2007, 25(1): 26-30.
ZHANG Q H, ZHAO X J, SUN Y. Pursuit-evasion barrier of two spacecrafts based on the sightline coordinate system [J]. Aerospace Control, 2007, 25(1): 26-30. (in Chinese)
- [11] 张秋华, 孙毅, 黄明明, 等. 近地共面轨道上两飞行器在径向连续小推力下的追逃界栅[J]. 控制与决策, 2007, 22(5): 530-534.
ZHANG Q H, SUN Y, HUANG M M, et al. Pursuit-evasion barrier of two spacecrafts under minute continuous radial thrust in coplanar orbit [J]. Control and Decision, 2007, 22(5): 530-534. (in Chinese)
- [12] 祝海. 基于微分对策的航天器轨道追逃最优控制策略[D]. 长沙: 国防科学技术大学, 2017.
ZHU H. Optimal control of spacecraft orbital pursuit-evasion based on differential game [D]. Changsha: National University of Defense Technology, 2017. (in Chinese)
- [13] PONTANI M, CONWAY B A. Numerical solution of the three-dimensional orbital pursuit-evasion game [J]. Journal of Guidance, Control, and Dynamics, 2009, 32(2): 474-487.
- [14] STUPIK J, PONTANI M, CONWAY B. Optimal pursuit/evasion spacecraft trajectories in the hill reference frame [C]// Proceedings of AIAA/AAS Astrodynamics Specialist Conference, 2012.
- [15] WOODBURY T D, HURTADO J E. Adaptive play via estimation in uncertain nonzero-sum orbital pursuit evasion games [C]//Proceedings of the AIAA SPACE and Astronautics Forum and Exposition, 2017.
- [16] YE D, SHI M M, SUN Z W. Satellite proximate interception vector guidance based on differential games [J]. Chinese Journal of Aeronautics, 2018, 31(6): 1352-1361.
- [17] YE D, SHI M M, SUN Z W. Satellite proximate pursuit-evasion game with different thrust configurations [J]. Aerospace Science and Technology, 2020, 99: 105715.
- [18] 孙松涛. 近地轨道上两航天器追逃对策及数值求解方法研究[D]. 哈尔滨: 哈尔滨工业大学, 2015.
SUN S T. Two spacecraft pursuit-evasion strategies on low earth orbit and numerical solution [D]. Harbin: Harbin Institute of Technology, 2015. (in Chinese)
- [19] PRINCE E R, HESS J A, COBB R G, et al. Elliptical orbit proximity operations differential games [J]. Journal of Guidance, Control, and Dynamics, 2019, 42(7): 1-15.
- [20] ZHANG C M, ZHU Y W, YANG L P, et al. Numerical solution for elliptical orbit pursuit-evasion game via deep neural networks and pseudospectral method [J]. Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering, 2023, 237(4): 796-808.
- [21] LEE S J, PARK S Y. Approximate analytical solutions to optimal reconfiguration problems in perturbed satellite relative motion [J]. Journal of Guidance, Control, and Dynamics, 2011, 34(4): 1097-1111.
- [22] 王强, 叶东, 范宁军, 等. 含有J2项摄动的卫星追逃轨道优化[J]. 北京理工大学学报, 2017, 37(4): 418-423.
WANG Q, YE D, FAN N J, et al. Pursuit evasion game with J2 perturbation [J]. Transactions of Beijing Institute of Technology, 2017, 37(4): 418-423. (in Chinese)
- [23] SHEN H X, CASALINO L. Revisit of the three-dimensional orbital pursuit-evasion game [J]. Journal of Guidance, Control, and Dynamics, 2018, 41(8): 1820-1828.
- [24] ZENG X, YANG L P, ZHU Y W, et al. Comparison of two optimal guidance methods for the long-distance orbital pursuit-evasion game [J]. IEEE Transactions on Aerospace and

- Electronic Systems, 2021, 57(1): 521–539.
- [25] 常燕, 陈韵, 鲜勇, 等. 机动目标的空间交会微分对策制导方法[J]. 宇航学报, 2016, 37(7): 795–801.
CHANG Y, CHEN Y, XIAN Y, et al. Differential game guidance for space rendezvous of maneuvering target [J]. Journal of Astronautics, 2016, 37(7): 795–801. (in Chinese)
- [26] INNOCENTI M, TARTAGLIA V. Game theoretic strategies for spacecraft rendezvous and motion synchronization [C]// Proceedings of the AIAA Guidance, Navigation, and Control Conference, 2016: 0873.
- [27] VENIGALLA C, SCHEERES D J. Delta-V-based analysis of spacecraft pursuit-evasion games [J]. Journal of Guidance, Control, and Dynamics, 2021, 44(11): 1961–1971.
- [28] 史帅科. 基于博弈论的多卫星围捕策略研究[D]. 北京: 北京邮电大学, 2020.
SHI S K. Research on multi-satellites pursuit evasion strategy based on game theory [D]. Beijing: Beijing University of Posts and Telecommunications, 2020. (in Chinese)
- [29] 刘彦昊, 余浩平, 蒙波, 等. 基于狼群优化的卫星集群对空间目标围捕方法[J/OL]. 北京航空航天大学学报, 2022: 1–14. <https://kns.cnki.net/kcms/detail//11.2625.V.20221229.1708.001.html>.
LIU Y H, SHE H P, MENG B, et al. Round-up method of space target by satellites swarm based on wolf pack optimization[J/OL]. Journal of Beijing University of Aeronautics and Astronautics, 2022: 1–14. <https://kns.cnki.net/kcms/detail//11.2625.V.20221229.1708.001.html>. (in Chinese)
- [30] ZHENG Z X, ZHANG P, YUAN J P. Nonzero-sum pursuit-evasion game control for spacecraft systems: a Q-learning method[J]. IEEE Transactions on Aerospace and Electronic Systems, 2023, 59(4): 3971–3981.
- [31] 张秋华, 孙松涛, 湛颖, 等. 时间固定的两航天器追逃策略及数值求解[J]. 宇航学报, 2014, 35(5): 537–544.
ZHANG Q H, SUN S T, CHEN Y, et al. Strategy and numerical solution of pursuit-evasion with fixed duration for two spacecraft[J]. Journal of Astronautics, 2014, 35(5): 537–544. (in Chinese)
- [32] HE H Q, SHI P, ZHAO Y S. Hierarchical optimization algorithm and applications of spacecraft trajectory optimization[J]. Aerospace, 2022, 9(2): 81.
- [33] JAGAT A, SINCLAIR A J. Nonlinear control for spacecraft pursuit-evasion game using the state-dependent Riccati equation method[J]. IEEE Transactions on Aerospace and Electronic Systems, 2017, 53(6): 3032–3042.
- [34] AURES-CAVALIERI K D. Incomplete information pursuit-evasion games with applications to spacecraft rendezvous and missile defense[D]. Texas A & M University, 2014.
- [35] LI Z Y, ZHU H, LUO Y Z. An escape strategy in orbital pursuit-evasion games with incomplete information [J]. Science China Technological Sciences, 2021, 64(3): 559–570.
- [36] TANG X, YE D, HUANG L, et al. Pursuit-evasion game switching strategies for spacecraft with incomplete-information[J]. Aerospace Science and Technology, 2021, 119: 107112.
- [37] 周俊峰. 基于微分对策理论的航天器追逃控制方法研究[D]. 哈尔滨: 哈尔滨工程大学, 2021.
ZHOU J F. Research on control method for spacecraft pursuit-evasion based on differential game theory [D]. Harbin: Harbin Engineering University, 2021. (in Chinese)
- [38] ZHANG Z, SUN C, CHEN J L, et al. Prediction analysis of target state in spacecraft pursuit game [C]// Proceedings of the 5th International Symposium on Autonomous Systems (ISAS), 2022: 1–6.
- [39] WANG M, WU H N. Autonomous game control for spacecraft rendezvous via adaptive perception and interaction[J]. IEEE Transactions on Aerospace and Electronic Systems, 2023, 59(3): 3188–3200.
- [40] VENIGALLA C, SCHEERES D. Spacecraft rendezvous and pursuit/evasion analysis using reachable sets [C]// Proceedings of the 2018 Space Flight Mechanics Meeting, 2018.
- [41] 于大腾. 空间飞行器安全防护规避机动方法研究[D]. 长沙: 国防科技大学, 2017.
YU D T. Approaches for the spacecraft security defense and evasion maneuver method [D]. Changsha: National University of Defense Technology, 2017. (in Chinese)
- [42] LIU Y, YE D, HAO Y. Distributed online mission planning for multi-player space pursuit and evasion [J]. Chinese Journal of Aeronautics, 2016, 29(6): 1709–1720.
- [43] LIU Y F, LI R F, HU L, et al. Optimal solution to orbital three-player defense problems using impulsive transfer[J]. Soft Computing, 2018, 22(9): 2921–2934.
- [44] ZHOU J F, ZHAO L, CHENG J H, et al. Pursuer's control strategy for orbital pursuit-evasion-defense game with continuous low thrust propulsion [J]. Applied Sciences, 2019, 9(15): 3190.
- [45] 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[J]. 上海航天(中英文), 2022, 39(2): 24–31.
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 [J]. Aerospace Shanghai (Chinese & English), 2022, 39(2): 24–31. (in Chinese)
- [46] 孙景瑞. 线性二次二人零和随机微分对策[D]. 合肥: 中国科学技术大学, 2014.
SUN J R. Linear quadratic two-person zero-sum stochastic differential games [D]. Hefei: University of Science and Technology of China, 2014. (in Chinese)
- [47] JAGAT A, SINCLAIR A J. Optimization of spacecraft pursuit-evasion game trajectories in the euler-hill reference frame[C]// Proceedings of the AIAA/AAS Astrodynamics Specialist Conference, 2014.
- [48] 李振瑜. 航天器追逃博弈路径规划与控制技术[D]. 长沙: 国防科技大学, 2019.
LI Z Y. Trajectory planning and control technology for spacecraft orbital pursuit-evasion game [D]. Changsha: National University of Defense Technology, 2019. (in Chinese)
- [49] HAFER W T, REED H L, TURNER J D, et al. Sensitivity methods applied to orbital pursuit-evasion [J]. Journal of Guidance, Control, and Dynamics, 2015, 38(6): 1118–1126.
- [50] SHI M M, YE D, SUN Z W, et al. Spacecraft orbital pursuit-evasion games with J_2 perturbations and direction-constrained thrust[J]. Acta Astronautica, 2023, 202: 139–150.
- [51] HORIE K. Collocation with nonlinear programming for two-

- sided flight path optimization [D]. Urbana-Champaign: University of Illinois at Urbana-Champaign, 2002.
- [52] HORIE K, CONWAY B A. Optimal fighter pursuit-evasion maneuvers found via two-sided optimization [J]. *Journal of Guidance, Control, and Dynamics*, 2006, 29(1): 105 - 112.
- [53] CARR R W, COBB R G, PACHTER M, et al. Solution of a pursuit-evasion game using a near-optimal strategy [J]. *Journal of Guidance, Control, and Dynamics*, 2017, 41(4): 841 - 850.
- [54] SCHOENWETTER L. Game theory applications in astrodynamics and space domain awareness [D]. Alabama: the University of Alabama, 2021.
- [55] 张乘铭. 航天器追逃博弈制导策略研究[D]. 长沙: 国防科技大学, 2021.
ZHANG C M. Research on guidance strategy for spacecraft pursuit-evasion games [D]. Changsha: National University of Defense Technology, 2021. (in Chinese)
- [56] 程林, 蒋方华, 李俊峰. 深度学习在飞行器动力学与控制中的应用研究综述 [J]. *力学与实践*, 2020, 42(3): 267 - 276.
CHENG L, JIANG F H, LI J F. A review on the applications of deep learning in aircraft dynamics and control [J]. *Mechanics in Engineering*, 2020, 42(3): 267 - 276. (in Chinese)
- [57] 黄旭星, 李爽, 杨彬, 等. 人工智能在航天器制导与控制中的应用综述 [J]. *航空学报*, 2021, 42(4): 524201.
HUANG X X, LI S, YANG B, et al. Spacecraft guidance and control based on artificial intelligence: review [J]. *Acta Aeronautica et Astronautica Sinica*, 2021, 42(4): 524201. (in Chinese)
- [58] IZZO D, MÄRTENS M, PAN B F. A survey on artificial intelligence trends in spacecraft guidance dynamics and control [J]. *Astrodynamics*, 2019, 3(4): 287 - 299.
- [59] SHIROBOKOV M, TROFIMOV S, OVCHINNIKOV M. Survey of machine learning techniques in spacecraft control design [J]. *Acta Astronautica*, 2021, 186: 87 - 97.
- [60] 吴其昌, 李彬, 李君, 等. 基于深度神经网络的无限时域型航天器追逃策略求解 [J]. *航天控制*, 2019, 37(6): 13 - 18, 58.
WU Q C, LI B, LI J, et al. Solution of infinite time domain spacecraft pursuit strategy based on deep neural network [J]. *Aerospace Control*, 2019, 37(6): 13 - 18, 58. (in Chinese)
- [61] YANG F, YANG L P, ZHU Y W, et al. A DNN based trajectory optimization method for intercepting non-cooperative maneuvering spacecraft [J]. *Journal of Systems Engineering and Electronics*, 2022, 33(2): 438 - 446.
- [62] HOYER S, SOHL-DICKSTEIN J, GREYDANUS S. Neural reparameterization improves structural optimization [EB/OL]. (2019 - 09 - 14) [2022 - 05 - 20]. <https://arxiv.org/pdf/1909.04240.pdf>.
- [63] SUTTON R S, BARTO A G. Reinforcement learning: an introduction [M]. Cambridge: the MIT Press, 2018.
- [64] ZENG X, ZHU Y W, YANG L P, et al. A guidance method for coplanar orbital interception based on reinforcement learning [J]. *Journal of Systems Engineering and Electronics*, 2021, 32(4): 927 - 938.
- [65] 赖俊, 魏竞毅, 陈希亮. 分层强化学习综述 [J]. *计算机工程与应用*, 2021, 57(3): 72 - 79.
LAI J, WEI J Y, CHEN X L. Overview of hierarchical reinforcement learning [J]. *Computer Engineering and Applications*, 2021, 57(3): 72 - 79. (in Chinese)
- [66] 刘冰雁, 叶雄兵, 高勇, 等. 基于分支深度强化学习的非合作目标追逃博弈策略求解 [J]. *航空学报*, 2020, 41(10): 324040.
LIU B Y, YE X B, GAO Y, et al. Strategy solution of non-cooperative target pursuit-evasion game based on branching deep reinforcement learning [J]. *Acta Aeronautica et Astronautica Sinica*, 2020, 41(10): 324040. (in Chinese)
- [67] YANG B, LIU P X, FENG J L, et al. Two-stage pursuit strategy for incomplete-information impulsive space pursuit-evasion mission using reinforcement learning [J]. *Aerospace*, 2021, 8(10): 299.
- [68] ZHANG J R, ZHANG K P, ZHANG Y, et al. Near-optimal interception strategy for orbital pursuit-evasion using deep reinforcement learning [J]. *Acta Astronautica*, 2022, 198: 9 - 25.
- [69] 袁利, 耿远卓, 汤亮, 等. 航天器轨道追逃博弈多阶段强化学习训练方法 [J]. *上海航天(中英文)*, 2022, 39(4): 33 - 41.
YUAN L, GENG Y Z, TANG L, et al. Multi-stage reinforcement learning method for orbital pursuit-evasion game of spacecrafts [J]. *Aerospace Shanghai (Chinese & English)*, 2022, 39(4): 33 - 41. (in Chinese)
- [70] 耿远卓, 袁利, 黄煌, 等. 基于终端诱导强化学习的航天器轨道追逃博弈 [J]. *自动化学报*, 2023, 49(5): 974 - 984.
GENG Y Z, YUAN L, HUANG H, et al. Terminal-guidance based reinforcement-learning for orbital pursuit-evasion game of the spacecraft [J]. *Acta Automatica Sinica*, 2023, 49(5): 974 - 984. (in Chinese)
- [71] ZHAO L R, ZHANG Y L, DANG Z H. PRD-MADDPG: an efficient learning-based algorithm for orbital pursuit-evasion game with impulsive maneuvers [J]. *Advances in Space Research*, 2023, 72(2): 211 - 230.