

一种基于智能电磁干扰策略的空中平台主动隐身方法

刘志鹏^{*①} 胡宝洁^① 吴耿^① 张樱豪^① 孙震宇^①

^①(电磁空间安全试验评估全国重点实验室 洛阳 471000)

摘要: 当前空中平台主要采取外形优化设计、吸波材料覆盖等被动隐身技术进行雷达隐身,但由于诸多技术瓶颈导致其在多方向、宽频谱上的隐身性能受到制约,主动隐身技术作为一种补充手段逐渐成为研究重点。为提升分布式雷达组网探测系统探测下空中平台隐身性能,该文结合认知电子战特点,以大幅降低雷达接收机对电磁干扰与空中目标的感知能力为目标,提出一种基于自卫/伴随式智能电磁干扰策略的空中平台主动隐身方法,在灵活干扰波束指向、多干扰频段覆盖基础上,通过构建自适应电磁干扰策略,实现多方向、宽频段上的目标雷达散射截面(RCS)等效缩减。具体来说,引入强化学习机制构建电子战策略生成框架,首先依托平台自身或伴随平台认知电子战系统,对外部雷达组网探测系统电磁辐射信号实时感知,并结合先期信息等综合构建高完备性观测空间,然后基于干扰样式带宽、功率、辐射方向等参数构建行为空间,并从影响雷达工作状态、降低电磁干扰暴露风险等角度设计多级奖励函数,最后采用强化学习算法完成智能体的引导训练与智能干扰策略优选。仿真实验结果表明,与传统被动隐身技术与固定干扰策略相比,所提方法有效降低了雷达组网对空中平台的探测范围与电磁干扰感知能力,对多频段雷达站的平均RCS等效缩减量最大可达9.4 dB,电磁干扰隐蔽占比不低于97.83%。生成策略下的干扰参数可根据外部电磁环境变化动态调整,有效提升了空中平台的雷达隐身性能,可为后续主动隐身技术思路发展提供参考。

关键词: 主动隐身; 雷达组网探测; 认知电子战; 强化学习; 智能电磁干扰策略生成

中图分类号: TN971

文献标识码: A

文章编号: 2095-283X(2026)x-0001-21

DOI: 10.12000/JR26079

CSTR: 32380.14.JR26079

引用格式: 刘志鹏, 胡宝洁, 吴耿, 等. 一种基于智能电磁干扰策略的空中平台主动隐身方法[J]. 雷达学报(中英文), 待出版. doi: 10.12000/JR26079.

Reference format: LIU Zhipeng, HU Baojie, WU Geng, *et al.* Active stealth method based on an intelligent electromagnetic jamming strategy for airborne platforms[J]. *Journal of Radars*, in press. doi: 10.12000/JR26079.

Active Stealth Method based on an Intelligent Electromagnetic Jamming Strategy for Airborne Platforms

LIU Zhipeng^{*①} HU Baojie^① WU Geng^① ZHANG Yinghao^① SUN Zhenyu^①

^①(State Key Laboratory of Electromagnetic Space Security Testing and Evaluation, Luoyang 471000, China)

Abstract: Current airborne platforms rely primarily on passive stealth techniques, such as shape optimization and radar-absorbing material coatings, to reduce their radar signatures. However, due to several technical bottlenecks, their stealth performance remains constrained in terms of multidirectional and wideband effectiveness. As a complementary approach, active stealth has gradually become a research focus. To enhance the stealth performance of airborne platforms against distributed radar network detection systems, this paper proposes an active stealth method based on a self-defense/escort intelligent electromagnetic jamming strategy inspired by the principles of cognitive electronic warfare. The proposed method aims to reduce radar receivers' perception of both electromagnetic interference and airborne targets. Through flexible jamming beam steering

收稿日期: 2026-04-28; 改回日期: 2026-06-20; 网络出版: 2026-xx-xx

*通信作者: 刘志鹏 liuzhip1994@163.com *Corresponding Author: LIU Zhipeng, liuzhip1994@163.com

责任编辑: 严俊坤 Corresponding Editor: YAN Junkun

©The Author(s) 2026. This is an open access article under the CC-BY 4.0 License
(<https://creativecommons.org/licenses/by/4.0/>)

and multiband jamming coverage, it achieves an equivalent reduction of the target Radar Cross-Section (RCS) over multiple directions and wide frequency bands via an adaptive electromagnetic jamming strategy. Specifically, a reinforcement learning mechanism is introduced to construct an electronic warfare strategy generation framework. First, the platform's onboard or escort cognitive electronic warfare system is used to sense, in real time, the electromagnetic radiation signals of external radar networked detection systems, and a comprehensive observation space is established by integrating prior intelligence and other relevant data. Then, an action space is formulated based on jamming parameters, such as bandwidth, power, and radiation direction. In addition, a multilevel reward function is designed to influence radar working states and reduce the risk of electromagnetic jamming exposure. Finally, a reinforcement learning algorithm is employed to train the agent and optimize the intelligent jamming strategies. Simulation results show that, compared with conventional passive stealth techniques and fixed jamming strategies, the proposed method effectively reduces both the detection range of radar networks and their perception of electromagnetic interference. The maximum average equivalent RCS reduction achieved for multiband radar stations is 9.4 dB, while the concealment rate of electromagnetic interference remains above 97.83%. Moreover, the general jamming parameters can be dynamically adjusted in response to changes in the external electromagnetic environment, substantially improving the radar stealth performance of airborne platforms and providing a reference for the development of future active stealth technologies.

Key words: Active stealth; Radar network detection; Cognitive electronic warfare; Reinforcement learning; Intelligent electromagnetic jamming strategy generation

1 引言

雷达隐身技术一般是指通过改变己方平台、系统等目标的雷达可探测信号特征,即雷达回波信号特征或在雷达接收处理中与其相关联的信号处理特征,使对方雷达难以在既定威力范围内对目标及时发现、稳定探测的综合技术,用于提升目标对雷达的隐蔽性。从雷达探测机理角度分析,如何缩减目标自身雷达散射截面积(Radar Cross Section, RCS)是雷达隐身的核心关键^[1,2]。

根据目标信号特征处理方式,当前雷达隐身技术可以分为主动和被动两类^[3,4]。被动隐身技术主要包括外形优化设计、隐身吸波材料覆盖、阻抗加载等,通过改变入射电磁波散射方向、将入射电磁波转换为热能、使用电路结构设计使入射电磁波与反射电磁波相互抵消等方式减少雷达回波信号能量,特点是不主动对外辐射或消耗能量,可在特定频段、特定方向上大幅缩减目标RCS,是目前空中飞行平台实现隐身的主流手段,但缺点是隐身频段主要集中在C/X/Ku等高频段,且RCS缩减性能对电磁波入射角度较为敏感,相关器件与结构不能根据来波信号自适应调整,导致隐身区域主要分布在平台前侧窄扇区,无法有效应对包含低频段、多方向的雷达组网探测威胁,实际应用场景受到较大限制^[5-9]。主动隐身技术作为一种补充手段,通过主动采取辐射措施以弥补对低频段、多方向雷达组网探测系统的隐身能力不足,主要包括等离子体、有源对消、智能蒙皮等,通过在材料表面形成等离子体

与入射电磁波碰撞吸收、添加天线辐射源与散射回波对消、在平台蒙皮表面植入微型处理器对入射电磁波截获对消等方式,根据截获来波信号进行参数自适应调整,以综合降低目标在不同方向、不同频段电磁波照射下所产生的电磁散射场,实现目标RCS缩减^[10-12]。王家兴等人^[13]提出一种有源对消隐身系统框架,利用测向阵列、主动对消电路模块等在时域同步的理想条件下实现了集中式与分布式对消,并在L/S/C波段上验证了对RCS缩减的有效性。Sengupta等人^[14]研究了使用微带天线进行主动对消的RCS缩减方法,通过控制微带贴片辐射与散射回波相干对消,并在矩形调制脉冲、高斯脉冲等雷达信号类型下验证了可行性。冯清帅等人^[15]利用矩量法计算得到目标表面电流反相后作为对消电流源,对目标回波散射场进行对消,并证明了所提方法在全角度对消的有效性。综合来看,当前主动隐身技术仍存在需大功率支持、截获分析实时性强以及系统控制精度要求高等应用难题,工程化应用难度较大。

随着雷达隐身研究的不断深入和应用场景拓展,考虑工程可行性和效果可控性,通过直接缩减目标RCS已经难以满足复杂对抗环境下的隐身能力需求。当前人工智能技术与硬件算力迅速发展,认知电子战逐渐成为电磁空间对抗领域的主流研究方向之一,基于智能的认知电子战系统可通过对外部电磁环境实时感知,根据分析结果与对抗效果评估不断优化对抗策略,获取动态电磁环境下的自主学

习能力与多威胁应对能力，可指导系统在高强度对抗环境下进行决策^[16-19]。从隐身需求来看，认知电子战思想高度契合主动隐身技术在自适应、敏捷性等方面的任务需求，因此，如何运用电子战技术实现目标RCS等效缩减、发挥“隐身+电子战”一体化运用效益具有重要研究意义，可为主动隐身提供新思路。

鉴于此，本文从认知电子战出发，提出一种基于智能电磁干扰策略的空中平台主动隐身方法，核心思想是对己方空中平台特性、对方雷达系统特性双向匹配后实施“隐蔽式”干扰，同时降低雷达对己方平台和电磁干扰的感知能力，以提升对多方向、多频段分布式雷达组网探测系统的隐身能力。整体按照“外部环境观测空间构建—智能电磁干扰策略生成—电磁干扰信号辐射—电磁干扰策略优化调整”的技术路径形成闭环，具体来说，如图1所示，通过引入强化学习机制构建策略生成框架，以对雷达组网探测系统电磁辐射信号观测、先验信息与信号处理结果等为实时输入，从支撑电子战系统“基础参数确定+二次精细微调”角度，构建高完备性的观测空间；在专家经验约束下，以电子战系统的干扰样式、带宽、功率等参数构建有限维度的行为空间，降低智能体探索收敛难度，并从雷达工作状态影响、电磁干扰隐蔽性等方面设计基于分级的奖励函数作为行为的即时与长期奖惩，评估主动

隐身性能，引导智能体自主学习更新、生成自适应的电磁干扰策略，以期大幅降低雷达接收机对电磁干扰与空中平台的感知能力，最后结合灵活指向的多干扰波束、多干扰频段覆盖进行干扰信号辐射，通过精准降低雷达探测威力范围实现目标RCS等效缩减，并根据主动隐身性能实时评估持续优化电磁干扰策略。

2 问题分析

2.1 场景设计

典型场景下，红方(我方)运用空中飞机编队，依靠外形设计、材料涂层、电磁掩护等雷达被动、主动隐身技术，对蓝方(对方)目标区域进行抵近；蓝方运用地面雷达组网探测系统对红方空中飞机编队进行探测，具体如图2所示。

过程中，红方飞机编队在前期主要依靠被动隐身技术、地球曲率影响等规避雷达探测。在抵近后，飞机运用基于智能电磁干扰策略的主动隐身方法提升对多方向、宽频段的雷达隐身性能，降低蓝方地面雷达组网探测系统对电磁干扰与目标感知能力。具体来说，飞机编队开启机载认知电子战系统对外部雷达信号实时感知，并综合运用数据链、机载雷达、红外探测系统等持续获取外部环境信息，获取蓝方地面雷达的部署位置、典型技术参数等，形成观测空间；对电磁干扰样式、带宽、功率等参

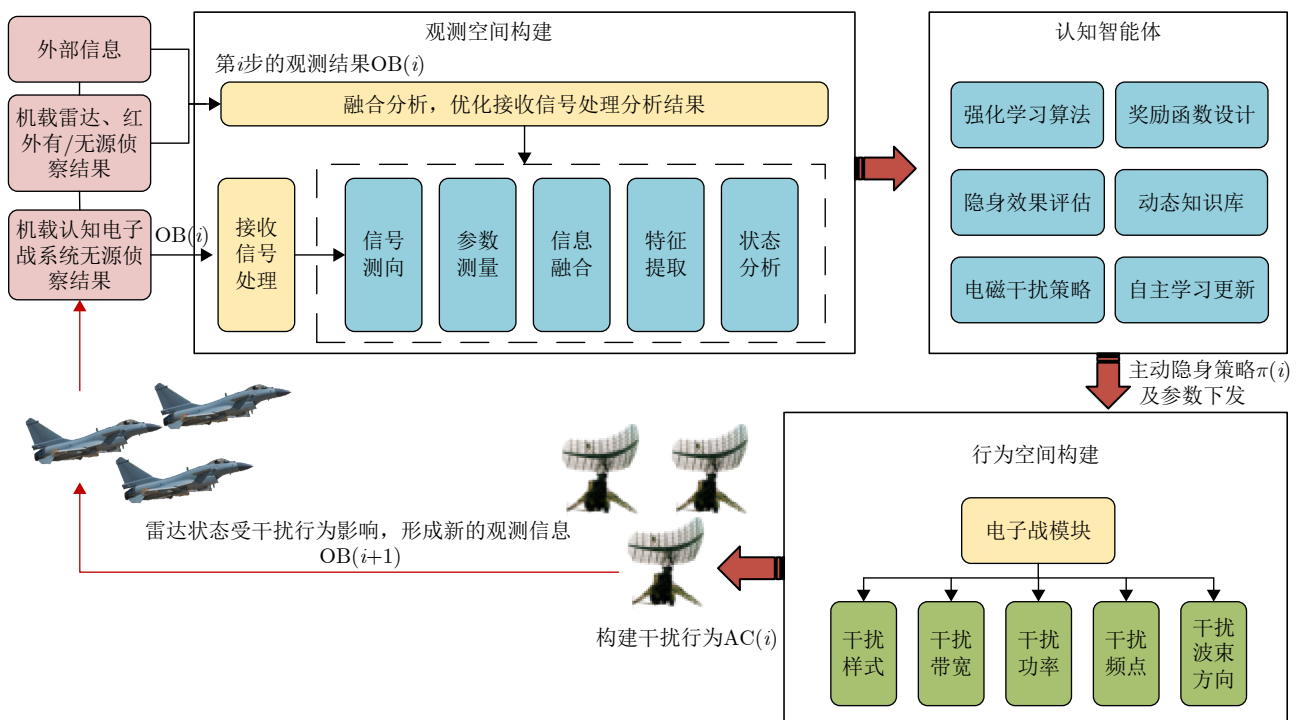


图 1 基于智能电磁干扰策略的空中平台主动隐身方法技术路径

Fig. 1 The pipeline of active stealth method based on an intelligent electromagnetic jamming strategy for airborne platforms

数组组合构建干扰行为,形成动作空间。当经观测空间信息分析确认雷达行为后,下发指令对单/多部蓝方雷达生成对应干扰策略,释放对应干扰信号,并根据观测空间变化实时调整干扰参数。蓝方地面雷达全程开机对空保持搜索状态,探测到目标后立即转入跟踪状态;当雷达接收机感知到较强电磁干扰时,则采取相关抗干扰措施且保持对空搜索。

2.2 机理分析

2.2.1 雷达工作状态分析

实现主动隐身关键在于通过电磁干扰降低对方雷达接收机对干扰和目标的感知程度,进而影响雷达工作状态,因此,实时分析雷达当前可能所处工作状态是基础。雷达主要工作状态通常可划分为搜索、跟踪与识别共3种典型状态。雷达一般带有干扰感知功能,其主要通过测量接收机通道内信号强度、脉冲重复周期特征及航迹连续性等,判断当前是否遭受干扰及干扰强度,因此可进一步细化为:无干扰/干扰下搜索、无干扰/干扰下跟踪及无干扰/干扰下识别共6种工作状态。

从雷达发射角度,不同工作状态下发射信号参数不同,主要包括脉宽、重周、频率、波形等,虽然雷达体制、性能各异,但依然可以采取定性、定量相结合的方式对其进行区分。假设雷达在*i*状态下发射信号参数集,表示为 $\mathbf{A}(i) = [a_1(i), a_2(i), \dots, a_n(i)]$,其中*i*包括无干扰/干扰下搜索、无干扰/干扰下跟踪以及无干扰/干扰下识别共6种状态,*n*代表雷达发射信号参数的个数, $a_n(i)$ 代表*i*状态下第*n*个雷达发射信号参数取值;认知电子战系统在*t*时刻接收到雷达信号,根据参数测量结果其参数集可表示为 $\mathbf{B}^t = [b_1^t, b_2^t, \dots, b_n^t]$, b_n^t 是在*t*时刻对 a_n 的测量结果。在有先验信息情况下,可将 \mathbf{B}^t 与 $\mathbf{A}(i)$ 进行匹配,以快速获取当前雷达可能所处的工作状态;当先验信息不足时,可基于专家经验规则对多个采集周期内的雷达信号进行定性分析,如表1所示,不同工作状态之间的主要区别如图3所示。

例如,雷达一般在搜索状态下倾向采用大脉宽、低重频信号和宽天线波束,以提升搜索区域覆盖与远距离目标发现概率;跟踪状态一般采用小脉

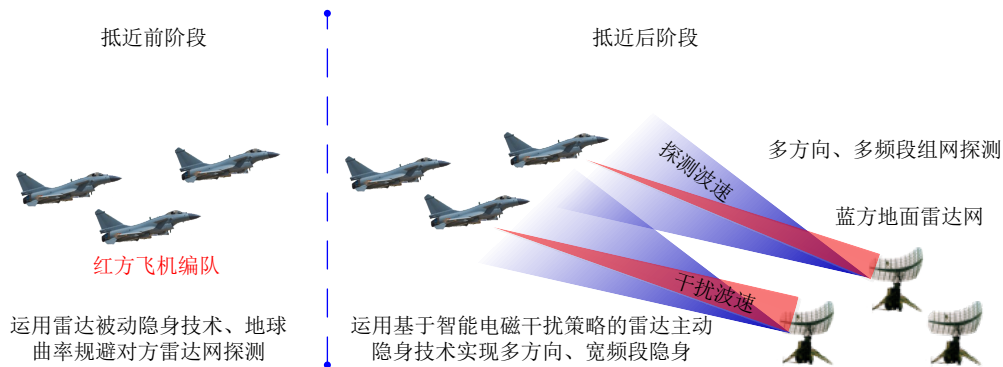


图2 场景阶段划分

Fig. 2 Phase division of the scenario

表1 主要雷达工作状态特点

Tab. 1 The characteristics of the main radar working states

雷达工作状态	发射信号特点
无干扰下搜索	主空域内搜索信号波形种类相对固定,一般为低重频信号,脉宽和重周可能存在多种组合以满足不同波位的能量需求,波束驻留时间可随区域搜索任务灵活调整,可利用多波束并行搜索提升搜索效率,波束驻留时间整体较短,单个驻留周期内的脉冲数目较为固定,工作频点固定,波束扫描顺序固定。
干扰下搜索	相比于无干扰下搜索,倾向采用中高重频信号、增加波束驻留时间以提高回波信号能量积累,单个驻留周期内的脉冲数目增多,并使用频率捷变、波形捷变、PRI抖动等措施抗干扰。
无干扰下跟踪	相比于搜索,信号波形种类更多(例如相位编码、频率编码等),具备更高的距离和速度分辨率,以满足不同目标类型的跟踪需求;一般采用中高重频信号,为保证目标跟踪稳定性,波束驻留时间整体较长,单个驻留周期内的脉冲数目较多,工作频点固定,且可根据目标机动特性实时调整波束指向。
干扰下跟踪	相比于无干扰下跟踪,增加波束驻留时间以提高回波信号能量积累,单个驻留周期内的脉冲数目增多,并使用频率捷变、波形捷变等措施抗干扰,并利用多波束合成技术增强目标跟踪稳定性。
无干扰下识别	采用大带宽线性调频或特定编码信号以求最大分辨力,采用长相干处理间隔,波束驻留时间整体较长且稳定,呈现“凝视”特征,单个驻留周期内的脉冲数目多。
干扰下识别	相比于无干扰下识别,为维持目标识别所需的信干噪比,使用频率捷变、波形捷变等措施抗干扰。

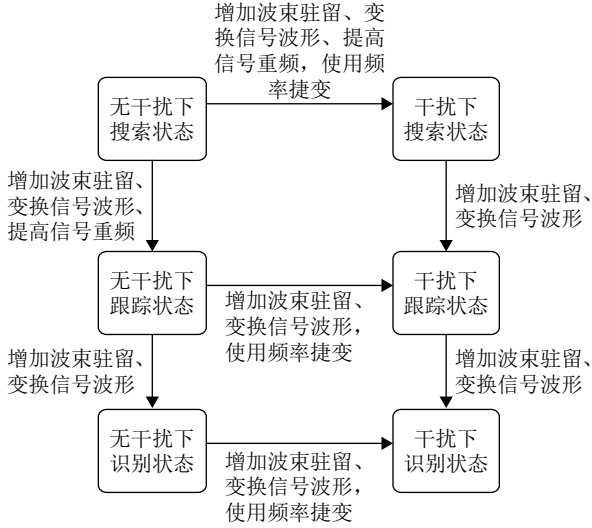


图 3 不同雷达工作状态之间转换分析

Fig. 3 Analysis of transitions between different radar working states

宽、高重频信号和窄天线波束，以提升目标跟踪稳定性与抗干扰能力等，目标识别状态一般采用大带宽信号和长时间驻留，在窄波束高采样状态下获取目标精细特征，提升目标识别能力，因此可通过分析脉宽、重周、波形、波束驻留时间等参数及变化，对雷达搜索、跟踪、识别等状态进行区分。通过分析信号频率、波形、波束驻留时间等参数以及变化，对无干扰与干扰状态进行区分。此外，雷达信号参数的设置与抵近意图紧密相关。例如在远距离跟踪场景下，为克服路径损耗并保持足够的雷达回波信噪比，也可能采用大脉宽信号，同时采用具备脉冲压缩、相参累积等信号处理增益的信号波形，以在兼顾距离分辨率的同时确保目标的持续稳定探测。因此，在雷达工作状态分析中，需要对接收雷达信号的测量分析结果和变化状态进行综合考虑。

2.2.2 主动隐身机理分析

结合雷达工作状态分析，主动隐身的核心需求可进一步具体为：通过电磁干扰策略影响雷达接收机内的干扰与目标检测判定，让雷达持续处于无干扰下搜索状态，实现空中平台主动隐身。

假设 t 时刻的电磁干扰策略表示为 $\pi(t)$ ， $Pb_{ji}(t)$ 为任意雷达工作状态 j 向无干扰下搜索状态 i 的转移概率，则有

$$\begin{aligned} \max_{\pi(t)} Pb_{ji}(t) &= P \{ \mathbf{B}^{t+1} = \mathbf{A}(i) | \mathbf{B}^t = \mathbf{A}(j) \} \\ \text{s.t. } \pi(t) &\in \Phi, \pi(t) = [x_1(t), x_2(t), \dots, x_m(t)] \end{aligned} \quad (1)$$

其中， Φ 代表有限的干扰资源，即 $\pi(t)$ 在有限的干扰资源下生成， $x_1 \sim x_m$ 为 m 个干扰参数。

在降低雷达接收机对目标检测判定上，选取用于目标检测的信干噪比为控制参数。假设雷达接收机内经信号处理后的回波信号功率为 P_e ，具体为

$$P_e = \frac{P_t G_t G_r \lambda^2 \delta D}{(4\pi)^3 R^4 L_1} \quad (2)$$

其中， P_t 为雷达发射峰值功率， G_t 、 G_r 分别为雷达天线的发射、接收增益， λ 为雷达发射信号波长， δ 为目标相对于雷达的 RCS， D 为信号处理增益， R 为目标与雷达的径向距离， L_1 为雷达发射接收综合损耗。

进入接收机的噪声干扰信号功率为 P_{ij} ，具体为

$$P_{ij} = \frac{P_j G_r \lambda^2 \gamma M}{(4\pi)^2 R_j^2 L_2} \quad (3)$$

其中， P_j 为干扰机等效辐射功率， γ 为极化损耗系数， M 为信号处理增益，噪声干扰一般为 0， R_j 为干扰机与雷达的径向距离，由于为自卫或伴随干扰，这里 $R_j \approx R$ ， L_2 为干扰接收综合损耗。

则接收机内用于目标检测的信干噪比 SJNR 表示为

$$\text{SJNR} = \frac{P_e}{P_{ij} + P_n} \quad (4)$$

其中， P_n 为雷达接收机内部噪声功率。假设目标符合 Swerling I 型起伏特性，则雷达对其单脉冲检测概率 P_d 可以表示为

$$P_d = \exp\left(-\frac{V_T^2}{2(1 + \text{SJNR})}\right) = \exp\left(\frac{\ln P_f}{(1 + \text{SJNR})}\right) \quad (5)$$

其中， V_T^2 为目标检测门限， P_f 为虚警率。因此，当确定雷达对某类型目标的 P_d 、 P_f 等技术参数时，使得 $P_{ij} = 0$ ，就可以计算出接收机在无干扰下用于目标检测所需的信噪比 SNR_T 。当实际的 SJNR 小于 SNR_T 时，可认为此时雷达无法对目标实现稳定探测与跟踪，实现接收机对目标感知能力下降。

在降低雷达接收机对干扰检测判定上，雷达对干扰检测一般区分为射频/中频阶段和数字信号处理阶段。射频/中频阶段，为降低杂波带来的虚警影响，通常对一段时间内射频/中频层输入信号功率的统计平均值作为干扰能量检测阈值；数字信号处理阶段，则根据频谱平坦度、恒虚警检测门限等参数进行干扰检测。因此，针对雷达接收机的干扰检测机理，对电磁干扰进行设计，这里区分噪声压制干扰与转发欺骗干扰 2 种类型。

噪声压制干扰属于连续随机信号，当干扰机以大功率持续辐射时，会导致雷达接收机在主瓣波束内接收到的干扰信号幅度远大于雷达回波信号，甚至在副瓣波束内也检测到较强的信号幅度，为降低

雷达接收机对噪声压制干扰的感知,需要同时满足:(1)干扰时机与窗口控制,只在雷达天线主瓣波束驻留期间辐射,可以通过信号包络测量、能量变化检测等手段进行主瓣波束驻留周期测量,且辐射窗口只覆盖返回目标回波所在及相邻单元,可有效降低射频/中频层的干扰能量检测阈值;(2)干扰波形设计,将噪声信号对地杂波、海杂波、气象杂波等进行拟合,统计分布符合韦伯分布、K分布等常规杂波分布,且信号频谱品坦度较好;(3)干扰功率控制,通过匹配空中平台及运动特性计算所需的干扰辐射功率,目的是使接收机内的SJNR略小于 SNR_T (一般降低幅度不超过3 dB),结合干扰时机与窗口控制,可进一步降低恒虚警检测门限,降低干扰感知能力,结合式(4)和式(5),可表示为

$$\begin{aligned} \frac{P_e}{P_{ij} + P_n} &\leq \ln(P_f - P_d) - 1 \\ \Rightarrow P_{ij} &\geq \frac{P_e}{\ln(P_f - P_d) - 1} - P_n \end{aligned} \quad (6)$$

转发欺骗干扰主要通过截获雷达信号并调制转发以实现距离、速度维度上的欺骗,且由于获取了部分信号处理增益,会在雷达接收机内形成较强的假目标,不利于降低干扰感知。为了增加转发欺骗干扰的隐蔽性,这里主要采用“快采快发、随机噪声调制”的切片干扰信号^[20,21],具体是:对接收到的雷达信号进行部分采集,复制多份进行拼接,形成时间上连续的干扰信号,然后再与噪声信号调制后形成具备噪声随机特性的欺骗干扰信号进行转发,如图4所示。最终干扰信号可在雷达接收机内形成近似连续的假目标群,获取带有部分信号处理增益(如脉压、相参累积等)的噪声干扰效果,且干扰功率可进一步降低,综合降低雷达对转发欺骗干扰的感知程度。

2.3 目标RCS等效缩减计算

为了定量评估电磁干扰策略下主动隐身性能的提升,便于将干扰效果与传统被动隐身的RCS缩

减效果统一到雷达探测方程框架下进行比较,这里使用RCS等效缩减这一概念进行计算。RCS等效缩减的本质在于,在不改变平台物理结构和目标散射电磁场情况下,通过释放精准电磁干扰降低雷达接收机端的信干噪比,影响雷达对目标探测概率、工作状态等,在雷达探测威力范围计算中,从检测意义上等效于目标RCS缩减,因此这里RCS等效缩减并非指空中平台在真实物理空间中的RCS缩减。

在指标计算上,假设雷达工作在固定 P_d , P_f , SNR_T 下,采用某频段对空中平台进行探测,在未使用主动隐身技术时,雷达对距离为 R_1 处的目标进行探测,当雷达接收机内检测到的回波信号功率使得 $SNR \geq SNR_T$ 时,视为检测到目标,此时目标RCS为 σ_1 ;使用主动隐身技术后,雷达进入无干扰下搜索状态,有 $SJNR < SNR_T$,雷达无法检测到目标与干扰,此时在 R_1 处的目标等效RCS为 σ_2 ,可以计算得到目标RCS缩减量 $\Delta\sigma$ 为

$$\Delta\sigma = \sigma_1 - \sigma_2 = \delta_1 \left(1 - \frac{SJNR}{SNR}\right) \quad (7)$$

其中,

$$SNR = \frac{P_t G_t G_r \lambda^2 \delta_1 D}{(4\pi)^3 R_1^4 L_1 P_n}, SJNR = \frac{P_t G_t G_r \lambda^2 \delta_2 D}{(4\pi)^3 R_1^4 L_1 P_n} \quad (8)$$

3 基于强化学习的智能电磁干扰策略生成

3.1 智能体定义

基于强化学习的电磁干扰策略生成智能体,本质是一个具备“观察-分析-决策-行动-评估”闭环能力的认知电子战系统,其组成结构、功能定位、边界限制主要如下:

(1) 组成结构:智能体主要由感知模块(基于观测空间的雷达工作状态分析)、决策模块(基于神经网络的自适应决策)、动作执行模块(电子战系统干扰参数生成)以及效果评估模块(主动隐身性能评估)共4部分组成。

(2) 功能定位:将获得的外部信号/信息侦察数据与先验信息等转化为雷达工作状态判定(例如搜索/跟踪、有/无干扰),并针对判定结果动态实时输出包含合适的干扰样式、带宽与功率等在内的干扰策略。

(3) 边界限制:智能体的输入边界为观测空间,涵盖信号测量、先验信息与实时分析结果,输出边界为行为空间,包括干扰样式、带宽、功率、频点以及干扰波束方向等。智能体的工作边界限制在机载电子战系统的物理性能指标之内,例如最大辐射功率、最小频率步进、最大动态范围等,且干扰策

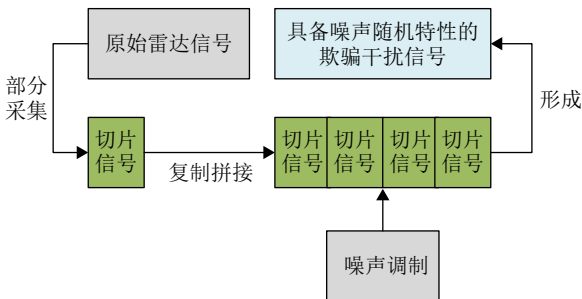


图4 切片干扰信号形成示意

Fig. 4 Schematic of sliced jamming signal formation

略在专家经验约束规则下生成，以保证动作的合理性和有效性。

3.2 观测空间构建

观测空间是智能体进行干扰决策的基础，除了需要电磁信号实测参数等基础信息外，还需要包含用于雷达工作状态分析的信号/信息处理结果等，有助于干扰策略快速收敛与动态调整。假设观测空间集合表示为 \mathbf{OB} ，具体为 $\mathbf{OB} = [\mathbf{ob}_1, \mathbf{ob}_2, \mathbf{ob}_3, \mathbf{ob}_4]$ ， $\mathbf{ob}_1 \sim \mathbf{ob}_4$ 为观测源，主要包括接收信号处理结果、先验信息分析结果、有无源探测信息分析结果、计算分析结果等。

接收信号处理结果观测源 \mathbf{ob}_1 ，主要包括信号来波方向 DOA、信号波形 Wf、接收信号功率 P_e 、信号频率 f 、信号带宽 Bw、信号脉宽 Pw、信号重周 PRI、信号数据率 Df、单位时间内信号频率变化方差 $D(f)$ 等基础要素。其中， P_e 为机载无源侦察设备接收到的信号功率， Df 为单次采样时间 T_s 内的脉冲信号数目，雷达在不同工作状态下的驻留周期、脉宽、重周可能不同，因此单次采样时间内发射信号数目也会随之变化，一般跟踪状态下大于搜索状态，干扰状态下大于无干扰状态； $D(f)$ 是对 T_s 内信号频率变化的方差统计，用于分析雷达信号的跳频带宽。先验信息分析结果 \mathbf{ob}_2 与有无源探测信息分析结果 \mathbf{ob}_3 ，主要包括己方平台 RCS 特性，即不同信号到达方位、不同频段下己方平台的 RCS 测量值，表示为 $\text{RCS}_{\text{DOA}-f}$ ；对方地面/机载雷达大致位置 PO_r ，用于优化 DOA， $\text{RCS}_{\text{DOA}-f}$ 等参数测量分析结果；对方地面/机载雷达接收机用于目标检测的信噪比阈值 SNR_T ，可基于对方雷达技战术性能先验信息与专家经验分析结果获取。

计算分析结果观测源 \mathbf{ob}_4 ，主要包括对方雷达接收机内实际信噪比的估计结果 SJNR^{est} ，以及 SJNR^{est} 与理论信噪比阈值 SNR_T 的差异值 J ，即 $J = |\text{SJNR}^{\text{est}} - \text{SNR}_T|$ 。假设对方雷达接收机内接收到的己方空中平台的回波信号功率为 $P_{\text{rs}}^{\text{est}}$ ，即

$$P_{\text{rs}}^{\text{est}} = \frac{P_s L_{\text{rec}} \sigma_{\text{DOA}-f} G_r \lambda_f^2 D}{(4\pi)^2 R_1^2 L_r} \quad (9)$$

其中， P_s 为机载无源侦察设备接收到的信号功率， L_{rec} 为侦察设备的接收损耗， $\sigma_{\text{DOA}-f}$ 为根据侦察设备实时测向与信号测频结果计算的空中平台 RCS 值，因此 $P_s L_{\text{rec}} \sigma_{\text{DOA}-f}$ 则为空中平台侧反射的雷达回波信号功率。 G_r 为雷达天线主瓣增益， λ_f 为雷达信号波长， D 为脉压、非相参/相参积累等带来的雷达回波信号处理增益， R_1 为空中平台与雷达的径向距离， L_r 为雷达的接收综合损耗。

假设空中平台侧的干扰机等效辐射功率为 P_j ，干扰信号带宽为 Bw_j ，雷达接收机带宽为 Bw_s ，且 Bw_s 一般略大于雷达信号带宽，在计算上可直接使用雷达信号带宽的测量结果，或通过外部先验信息分析获取。则接收机内干扰信号功率为 $P_{\text{rj}}^{\text{est}}$ 可表示为

$$P_{\text{rj}}^{\text{est}} = \frac{P_j G_r \lambda_f^2 \gamma M}{(4\pi)^2 R_1^2 L_r} \times \frac{\text{Bw}_s}{\text{Bw}_j} \quad (10)$$

其中， γ 为极化损耗系数， M 为干扰信号的信号处理增益。当干扰信号样式为噪声压制干扰时，一般有 $M=1$ ，为转发欺骗干扰时，一般有 $M = k_1 D$ ， k_1 为比例系数。由于雷达接收机内部噪声功率远小于干扰信号功率，则雷达接收机内实际信噪比估计结果 SJNR^{est} 可简化表示为

$$\text{SJNR}^{\text{est}} = \frac{P_{\text{rs}}^{\text{est}}}{P_{\text{rj}}^{\text{est}}} = \frac{P_s L_{\text{rec}} \sigma_{\text{DOA}-f} D}{P_j \gamma M} \times \frac{\text{Bw}_j}{\text{Bw}_s} \quad (11)$$

通过式(11)可知， SJNR^{est} 主要和雷达发射峰值功率，回波信号与干扰信号的信号处理增益等有关。在非合作场景中，对方雷达的信号处理增益往往是未知的，因此需要结合先期的已有信息与目标信号库开展大样本仿真，以分析推断对方目标雷达的信号样式、信号处理方式及信号处理增益等，提升观测空间内信息精确性。这里针对三种典型雷达信号类型进行了分析，具体见附录。

由于观测空间内各要素的取值会随着己方空中平台的机动实时变化，需要设计采样间隔对观测空间的要素进行持续提取与融合。因此，可将观测空间表示为

$$\begin{aligned} \mathbf{OB} &= [\mathbf{ob}_1, \mathbf{ob}_2, \mathbf{ob}_3, \mathbf{ob}_4] \\ \mathbf{ob}_1 &= [\text{DOA}, \text{Wf}, P_e, f, \text{Bw}, \text{Pw}, \text{PRI}, \text{Df}, D(f)] \\ \mathbf{ob}_2 &= [\text{RCS}_{\text{DOA}-f}, \text{SNR}_T] \\ \mathbf{ob}_3 &= [\text{PO}_r] \\ \mathbf{ob}_4 &= [\text{SJNR}^{\text{est}}, J] \end{aligned} \quad (12)$$

综合分析来看，观测空间内所含要素可有效支撑电子战系统进行“基础参数确定+二次精细微调”，具体为：对理论计算结果 \mathbf{ob}_4 观测，有助于智能体在自主学习过程中快速形成下一步电磁干扰基础参数；在非合作式观测场景下，当雷达技战术指标先验信息不足、参数测量与估计值误差较大时，例如，估计值 SJNR^{est} 与理论值 SNR_T ，以及差异值 J 可能与实际值存在差异，可通过观测 \mathbf{ob}_1 细化分析当前雷达实际可能所处的工作状态，并根据分析结果对电磁干扰基础参数持续进行二次精细调整，进一步降低估计值与实际值的差异，有助于提升干扰策略有效性。

3.3 行为空间构建

干扰行为由电子战系统的具体干扰参数构建,主要包括干扰样式 I_t 、干扰带宽 I_b 、干扰功率 P_j 、干扰频点 I_f 、干扰波束方向 I_a 等,参数的取值范围与组合共同约束行为空间 AC 的边界,表示为 $AC = [I_t, I_b, P_j, I_f, I_a]$ 。干扰时机和窗口默认只覆盖雷达波束主瓣驻留时间,干扰波形包括采用自然杂波拟合调制生成的噪声信号、切片信号。

I_f 和 I_a 可直接采用接收信号处理的实时测量结果,保证干扰与对方雷达回波在频域和空间上对准, I_t, I_b, P_j 由智能体自主设计。为保证干扰行为的有效性和参数快速收敛,采取“专家经验约束+智能体自主选择”相结合方式,先基于专家经验对部分干扰参数的取值范围与取值规则进行约束,压缩行为空间边界,再运用智能体在约束空间内自主选择,具体如表2所示。

干扰样式方面,主要包括噪声压制干扰与转发欺骗干扰共2类,约束规则上:(1)噪声压制干扰一般为不同带宽的宽带噪声,转发欺骗干扰为切片干扰(结合2.2.2主动隐身机理分析),以应对不同类型的雷达信号。(2)干扰样式选择,对大脉宽、高功率雷达信号,当在最大功率约束下,噪声压制干扰无法满足雷达接收内信噪比要求时,一般使用转发欺骗干扰以获取高信号处理增益;对小脉宽、低功率雷达信号,一般使用主动释放的噪声压制干扰以避免侦察漏脉冲、干扰时域覆盖不完全问题,但考虑到频谱占用度,干扰时长会受到限制。(3)当对方雷达处于搜索状态,倾向使用噪声压制干扰,避免假目标信号提前暴露空中飞机抵近目的;当对方雷达处于跟踪状态,若信号为大范围跳频情况,倾向使用噪声压制干扰,以降低侦察漏脉冲概率,当信号为定频或小范围跳频情况,此时空中飞机可能已被对方雷达锁定,需要通过假目标干扰降低被对方雷达持续锁定概率,同时获取部分信号处理增益,因此倾向使用转发欺骗干扰。

干扰带宽方面,包括宽带与窄带共2类。约束规则上^[22]:(1)噪声压制干扰一般为宽带,用于应

对雷达信号的频率捷变;转发欺骗类干扰一般为窄带,用于模拟雷达回波信号特性,获取高信号处理增益。(2)干扰带宽选取应综合考虑实时侦察结果、威胁信号匹配关系以及先验信息等因素,并结合具体场景进行合理设置。对于噪声压制干扰,为提高对方频率捷变雷达信号的频域覆盖能力,通常采用宽带干扰方式,其工程上可选带宽一般处于几十MHz至数百MHz范围;对于转发欺骗干扰,由于该类干扰通过对截获雷达信号进行延时、调制与转发以模拟目标回波,其有效带宽通常应与目标雷达信号带宽保持一致,以保证假目标信号在匹配滤波和后续处理中的相容性。干扰功率方面,受限于电子战系统的最大发射功率、动态范围及可调精度等工程因素,其调节方式通常表现为有界离散控制而非理想连续控制,且存在最小步进值限制。

3.4 奖励函数分层设计

奖励函数设计是引导智能体生成干扰策略的核心,其目的是通过对观测空间分析开展主动隐身性能评估,从而指导干扰行为调整。

从理论上讲,可以将主动隐身问题建模为有限时域马尔可夫决策过程(Markov Decision Process, MDP),定义为元组 $\{S, A, P, R, \gamma\}$ 。其中,状态空间 S 为观测空间,动作空间 A 为行动空间内的具体动作,转移概率 P 代表在环境状态 s 下采取策略 π 后,状态演化至环境状态 s' 的概率; R 为奖励函数, γ 为折扣因子,用于权衡智能体的即时奖励与长期奖励。智能体的优化目标是最大化累积期望奖励 $J(\pi) = E_{\pi}[\sum_{t=0}^T \gamma^t R_t]$ 。在该框架下,通过对 R 设计,将雷达探测概率 P_d 最小化问题转化为强化学习可处理的约束条件下的梯度优化问题,即基于MDP框架的优化问题,旨在保证收敛性前提下实现对无干扰下搜索状态转移概率的最大化。在具体设计上,评估要素主要包括雷达工作状态影响、干扰功率控制、干扰带宽匹配以及电磁暴露风险等。

在雷达工作状态影响评估方面,选取观测空间内相关要素建立雷达工作状态映射关系,如表3所示。

表2 行为空间约束规则

Tab. 2 The constraints rules of action space

干扰参数	基于专家经验的约束限制
干扰样式	①对大脉宽、高功率雷达信号,噪声压制干扰无法满足雷达接收内信噪比要求时,一般使用转发欺骗干扰以获取高信号处理增益;对小脉宽、低功率雷达信号,一般使用噪声压制干扰避免侦察漏脉冲、干扰时域覆盖不完全。②在对方雷达搜索状态下,倾向使用宽带压制干扰信号,避免假目标信号提前暴露自身;在跟踪状态下,若对方使用大范围跳频信号,倾向使用宽带压制信号,降低侦察漏脉冲概率;否则倾向使用转发欺骗干扰信号,以获取信号处理增益。
干扰带宽	噪声压制干扰:宽带;转发欺骗干扰:窄带,与雷达信号带宽相同。
干扰功率	存在最大干扰功率与动态范围限制,最小步进值限制。

表 3 观测空间-雷达工作状态关系映射

Tab. 3 Mapping between observation space and radar working states

观测要素	对应雷达工作状态分析
信号数据率 Df	①当单位采样时间内采集到的雷达脉冲信号数量(来自同一雷达)上升, 即 Df 增大, 代表此时雷达可能由搜索状态转为跟踪、识别, 或干扰下搜索状态; ②当 Df 减小甚至降为 0, 代表雷达对平台照射行为减少, 可能转为无干扰下搜索状态。
信号频率变化方差 $D(f)$	①当单位采样时间内采集到的雷达信号频点变化增大, 即 $D(f)$ 增大, 代表此时雷达可能由于感受到明显干扰而采取频率捷变措施, 处于干扰下搜索、跟踪或识别状态; ②当 $D(f)$ 减小甚至降为 0, 代表雷达可能处于无干扰下搜索、跟踪或识别状态。
信号重周 PRI	与信号数据率相关联, ①当 PRI 减小时, 代表此时雷达可能为了增加单个相关处理间隔内的信号累积数目、提升信号数据率, 可能由搜索状态转为跟踪、识别, 或干扰下搜索状态; ②当 PRI 增大时, 代表雷达可能重新转为无干扰下搜索状态。

在干扰功率控制、干扰带宽匹配以及电磁暴露风险评估方面, 主要考虑:

(1) 基于对空中平台与雷达双向匹配的干扰信号功率控制, 在空中平台运动过程中, 雷达接收机内所形成的实际信干噪比与目标检测所需信噪比的差异 J 需一直维持在可接受范围内, 降低雷达在既定距离上对目标的检测概率, 以使空中平台一直处于雷达最大探测范围边界, 雷达始终处于无干扰下搜索状态, 实现“低-零”干扰。

(2) 基于雷达工作状态分析的干扰信号带宽匹配, 结合行为空间构建中干扰样式、带宽约束规则分析, 当雷达采用大范围跳频时, 即 $D(f)$ 增大, 为降低电子战系统的侦察漏脉冲概率, 此时宽带压制干

扰的匹配奖励更高; 当单位时间内侦察到的雷达信号数目增多, 即 Df 增大, 雷达可能运用窄波束开展针对性搜索, 或已经转入跟踪状态, 为迅速增大雷达信号处理误差, 转发欺骗干扰的匹配奖励更高。

(3) 降低电磁暴露风险, 长期使用大带宽干扰信号、干扰功率突发式变化可能会明显增加电磁干扰信号被感知的概率, 需要对干扰过程中的平均干扰带宽, 以及相邻干扰步长的干扰功率变化差异进行约束。

由于智能体的干扰行为调整与观测步长有关, 因此干扰参数取值设计与干扰效果评估是时域离散的, 结合上述考虑, 第 i 步行为的奖励函数 $Rw(i)$ 具体设计为

$$\left\{ \begin{array}{l} Rw(i) = \lambda_1 e^{-|SJNR^{est}(i) - SNR_T|} + \lambda_2 e^{-E\left(\sum_i |SJNR^{est}(i) - SNR_T|\right)} - \lambda_3 D(f(i)) I_b(i)^{-1} - \lambda_4 Df(i) I_b(i) \\ \quad - \lambda_5 i^{-1} \sum_i I_b(i) + \lambda_6 e^{-D(SJNR^{est}(i))} \\ D(SJNR^{est}(i)) = SJNR^{est}(i) / SJNR^{est}(i-1) \\ E\left(\sum_i |SJNR^{est}(i) - SNR_T|\right) = i^{-1} \sum_i |SJNR^{est}(i) - SNR_T| \\ D(f(i)) = n^{-1} \sum_n |f(i)^n - \overline{f(i)}| \end{array} \right. \quad (13)$$

其中, $\lambda_1 \sim \lambda_6$ 为比例系数, $I_b(i)$ 为第 i 步采用的干扰信号带宽, $Df(i)$ 为第 i 步观测到的雷达信号数据率, $D(SJNR^{est}(i))$ 代表第 i 步形成的实际信干噪比与第 $i-1$ 步的比值, $E\left(\sum_i |SJNR^{est}(i) - SNR_T|\right)$ 代表 $1 \sim i$ 步内所形成的平均信干噪比差异值, $D(f(i))$ 为第 i 步观测时间内测量到的雷达信号频率变化方差, $f(i)^n$ 为第 n 个脉冲的频率测量结果, $\overline{f(i)}$ 为观测时间内的频率均值, 当雷达信号跳频范围越大, $D(f(i))$ 的值越大。

奖励函数 Rw 中各要素具体考虑如下:

第 1 项是对信干噪比差异 $J = |SJNR^{est} - SNR_T|$ 的指数奖励, 是干扰效果评估的直观奖励。对应于式(5)所构建的雷达检测概率-信干噪比曲线, 信干噪比差异值越小, 智能体的即时奖励越大, 目的是

引导智能体快速将雷达接收机内信干噪比快速推向信噪比检测阈值下方, 采用指数形式是为了当智能体形成可接受的差异值后可获取大量奖励, 强化对智能体引导作用。第 2 项是对信干噪比差异均值 $E\left(\sum_i |SJNR^{est}(i) - SNR_T|\right)$ 的指数奖励, 通过收集多个差异值 Dif_i 做平均处理, 随着飞机抵近, 平均误差可能逐步增大或减小, 反映干扰效果的持续变化, 体现对智能体的长期奖励。因此, 第 1 项、第 2 项本质上是针对雷达检测概率的梯度优化引导项。

第 3 项是对干扰带宽的匹配惩罚, 若雷达处于干扰下搜索或跟踪状态, 观测时间内的检测到的雷达信号跳频范围可能增大, 即雷达信号频率变化方差 $D(f(i))$ 增大, 此时采用窄带转发欺骗干扰可能存在频率测量误差、漏脉冲等问题, 采用宽带噪声

压制干扰可减少惩罚。第4项是对当前所处工作状态的惩罚,根据表3分析可知,当单位采样时间内雷达脉冲信号数量 $Df(i)$ 上升,此时雷达可能由搜索状态转为跟踪、识别状态或抗干扰状态,采用宽带噪声压制干扰可能无法满足信噪比要求,进而导致雷达可能转入稳定跟踪状态,采用窄带转发欺骗干扰可减少惩罚;当 $Df(i)$ 下降甚至降为0,代表雷达对平台照射行为减少,干扰行为有效。

第5项是对非必要干扰带宽进行惩罚,统计过程中干扰信号带宽的平均值,目的是对干扰信号带宽进行约束,长期使用大带宽可能导致多波束干扰收发之间产生互扰,且干扰信号被雷达检测到的概率将提升,电磁干扰频谱暴露风险较大。第6项是对干扰功率的波动进行约束,对突发式的功率变化行为进行惩罚。通过上述奖励函数的分层设计,可引导智能体在训练过程中综合考虑降低雷达接收机内对干扰和目标的检测能力,使得雷达系统尽可能处于无干扰下搜索状态,实现主动隐身。

3.5 智能电磁干扰策略生成

3.5.1 强化学习算法分析与选取

由于电磁干扰行为属于连续控制类问题,为实现主动隐身的智能干扰策略动态生成,结合雷达对抗场景的连续性、干扰参数调整的动态性以及电磁环境的噪声干扰特性,采用基于策略梯度优化的无模型方法对智能体进行训练,这里对双延迟深度确定性策略梯度算法(Twin Delayed Deep Deterministic Policy Gradient, TD3)、深度确定性策略梯度算法(Deep Deterministic Policy Gradient, DDPG)、近端策略优化算法(Proximal Policy Optimization, PPO)等相关主流算法进行对比分析^[23-25]。具体分析见附录。

量化对比分析方面,基于相同观测空间、行为空间和奖励函数,对DDPG,PPO和TD3算法进行部分步长与轮次(50个仿真步长与1000个训练回合数)的预训练对比分析,训练平台采用Intel(R) Core(TM) i9-14900HX (2.20 GHz)处理器,并使用NVIDIA GeForce RTX 4060 Laptop GPU进行并行加速训练。需要说明的是,该对比仅用于智能体

训练算法选型,并选取相同轮次训练时间,最后 K 个回合($K=200$)的平均回合奖励、奖励标准差等指标评估算法性能。智能体在不同算法下的训练效果曲线如图5所示,评估结果如表4所示。

在相同训练条件下,3种算法均能通过与环境交互逐步提升训练奖励,说明其均具备智能干扰策略生成能力,但在收敛速度、训练时间和策略稳定性方面存在明显差异。根据图5看出,TD3在训练后期的平均回合奖励最大,奖励波动幅度较小,策略稳定性最优。从表4的量化结果可以看出,TD3在收敛轮次、平均回合奖励和奖励标准差等指标上具有较好的综合表现,能够在较短训练时间内获得更稳定的策略输出。因此,综合考虑连续动作优化能力、训练收敛速度、最终策略性能和训练后期稳定性,本文选用TD3算法作为主动隐身策略智能体的训练算法。

3.5.2 智能体训练流程

基于TD3算法的智能体具体训练流程如算法1所示^[26]。

在具体应用中,结合典型场景生成新的环境状态 (s_t, \mathbf{ob}_t) 输入到 π_ϕ ,根据参数 ϕ_t ,完成干扰参数设置,发射干扰信号。当收集到足够新数据后,可对奖励行数进行重新设计,重复表5进行训练,对 π_ϕ 进行更新。

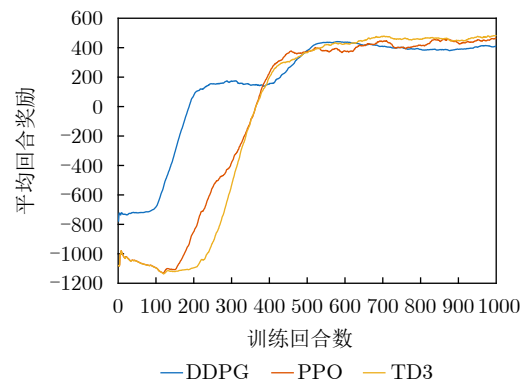


图5 主流算法预训练效果分析对比

Fig. 5 Comparison of pre-training performance among mainstream algorithms

表4 各算法性能评估结果对比

Tab. 4 Comparison of evaluation results for mainstream algorithm

算法	训练时间(s)	收敛轮次/回合数	最后 K 个回合的平均回合奖励($K=200$)	最后 K 个回合的平均回合奖励标准差($K=200$)
DDPG	564	532	394.3	12.5
PPO	342	457	443.5	10.52
TD3	436	436	471.5	10.18

算法 1 基于TD3的主动隐身策略智能体训练伪代码附录1典型雷达信号类型的信号处理增益计算分析

Alg. 1 Training pseudocode of TD3-based active stealth strategy agent

输入：随机初始化环境参数，主要包括地面雷达和飞机编队的位置信息，以及地面雷达的工作参数、飞机编队的RCS参数特性等先验信息；Actor网络 π_ϕ 、Critic_1网络 Q_{θ_1} 、Critic_2网络 Q_{θ_2} 、目标Actor网络 $\pi_{\phi'}$ 、目标Critic_1网络 $Q_{\theta_1'}$ 、目标Critic_2网络 $Q_{\theta_2'}$ ，经验回放池D等。

输出：主动隐身策略模型 π_ϕ 。

For 训练轮次 Episode = 1 to M 执行

初始化空中平台状态 s_0 和观测空间特征向量 \mathbf{ob}_0 ，其中 \mathbf{ob}_0 为OB转换后的一维向量， $\mathbf{ob}_0 \in R^{14 \times 1}$

For 时间步 $t = 1$ to T 执行

π_ϕ 根据当前状态 (s, \mathbf{ob}_0) 输出动作 \mathbf{ac} ： $\mathbf{ac} \leftarrow \pi_\phi(s) + \varepsilon$ ， $\varepsilon \sim N(0, \delta)$ 。其中， $\mathbf{ac} \in R^{3 \times 1}$ ， ε 表示服从均值为0、方差为 σ 的高斯噪声

在环境中执行动作 \mathbf{ac} ，观测奖励 Rw 和下一时刻状态 (s', \mathbf{ob}') ，并将 (s, \mathbf{ac}, Rw, s') 存入经验回放池D

从D中随机抽取N个样本进行小批量训练(由于RCS随空中平台运动变化较大，这里不采用优先经验回放，有助于训练收敛)，对 π_ϕ ， Q_{θ_1} ， Q_{θ_2} ， $\pi_{\phi'}$ ， $Q_{\theta_1'}$ ， $Q_{\theta_2'}$ 进行更新

计算目标Actor网络 $\pi_{\phi'}$ 在状态 s' 下的输出动作 $\overline{\mathbf{ac}} \leftarrow \pi_{\phi'}(s') + \varepsilon'$ ， $\varepsilon' \sim clip(N(0, \delta), -l, l)$ 。其中， l 为截断噪声边界

运用目标Critic网络 $Q_{\theta_1'}$ ， $Q_{\theta_2'}$ 计算目标Q值 y ：

$y \leftarrow Rw + \gamma \min(Q_{\theta_1'}(s', \overline{\mathbf{ac}}), Q_{\theta_2'}(s', \overline{\mathbf{ac}}))$ 。其中， $Q_{\theta_1'}(s', \overline{\mathbf{ac}})$ ， $Q_{\theta_2'}(s', \overline{\mathbf{ac}})$ 为 $Q_{\theta_1'}$ ， $Q_{\theta_2'}$ 对状态 s' 下 $\overline{\mathbf{ac}}$ 的动作价值估计

更新Critic网络 Q_{θ_1} ， Q_{θ_2} ，有 $\theta_i \leftarrow \arg_{i=1,2} \min [N^{-1} \cdot \sum (y - Q_{\theta_i}(s, \mathbf{ac}))^2]$ 。其中， $Q_{\theta_i}(s, \mathbf{ac})$ 为 Q_{θ_1} ， Q_{θ_2} 对状态 s 下 \mathbf{ac} 的动作价值估计

If $t \pmod T == 0$ 则

更新 π_ϕ ，有 $\nabla_\phi J(\phi) = N^{-1} \sum \nabla_{\mathbf{ac}} Q_{\theta_1}(s, \mathbf{ac}) \Big|_{\mathbf{ac}=\pi_\phi(s)} \nabla_\phi \pi_\phi(s)$

软更新目标网络 $\pi_{\phi'}$ ， $Q_{\theta_1'}$ ， $Q_{\theta_2'}$ 参数， $\theta_i' \leftarrow \tau \cdot \theta_i + (1 - \tau) \theta_i'$ ， $\phi' \leftarrow \tau \cdot \phi + (1 - \tau) \phi'$ 。其中 τ 为软更新参数

End if

End for

End for

输出智能干扰策略模型 π_ϕ

表 5 雷达站主要典型参数

Tab. 5 The main typical parameters of radar stations

参数类型	雷达1站	雷达2站	雷达3站	雷达4站
发射机峰值功率	40 kW	100 kW	100 kW	50 kW
天线主瓣增益	20 dB	30 dB	30 dB	20 dB
工作频段	L频段	C频段	C频段	S频段
接收机目标检测所需信噪比	12 dB	21 dB	21 dB	13 dB
发射/接收损耗	3 dB/3 dB	3 dB/3 dB	3 dB/3 dB	3 dB/3 dB
搜索信号脉宽/带宽/重周组合	150 us/5 MHz/ 2000 us	30 us/5 MHz/500 us	30 us/5 MHz/500 us	100 us/5 MHz/ 1000 us
跟踪信号脉宽/带宽/重周组合	组合1: 100 us/ 5 MHz/1500 us	组合1: 15 us/ 5 MHz/200 us	组合1: 15 us/ 5 MHz/200 us	组合1: 50 us/ 5 MHz/334 us
	组合2: 50 us/ 5 MHz/1000 us	组合2: 7 us/ 5 MHz/100 us	组合2: 7 us/ 5 MHz/100 us	组合2: 25 us/ 5 MHz/167 us
搜索/跟踪状态下单个相干处理间隔内的脉冲累积数目	32/64	16/32	16/32	8/16

4 仿真验证

4.1 场景构建

假设我方单机抵近场景下，对方4部地面雷达对我飞机进行组网探测跟踪，如图6所示。其中，

飞机出发点为(70000 m, 60000 m, 500 m)，终点为(0 m, 0 m, 1000 m)，雷达1站位置为(20000 m, 40000 m, 0 m)，雷达2站位置为(50000 m, 8000 m, 0 m)，雷达3站位置为(10000 m, 30000 m, 0 m)，雷达4站位置为(60000 m, 30000 m, 0 m)，飞机沿

设计航线抵近,与终点之间的径向速度为250 m/s,仿真采样间隔为 $\Delta t=1$ s。

4.2 模型参数及工作策略设计

4.2.1 模型参数设计

对方雷达组网探测系统主要包括1个L波段、1个S波段、2个C波段雷达站,主要典型参数如表5所示,包括发射机峰值功率、天线主瓣增益、工作频段、接收机目标检测所需信噪比、信号组合等。假设我方机载电子战系统设备具备多波束、多目标干扰能力,可在方位向上开展全向侦察、测向与干扰波束发射,且采用窄波束实时精确干扰,各干扰波束之间相互独立,主要典型参数如表6所示,包括等效辐射功率、动态范围、接收损耗、步进精度等。

假设飞机在地面雷达探测视角下为点目标,且RCS随地面雷达工作频段、探测视角变化而变化,参考典型飞机RCS“侧向最大、前后向最小”变化特点,具体设计变化曲线如图7(a)所示。其中,L频段下的RCS最大,S频段次之,C频段最小,机头向(方位角为 0°)L频段下的RCS为 -5.2 dBsm,

S频段下为 -11.1 dBsm,C频段下 -13.6 dBsm;正侧向(方位角为 90°)L频段下的RCS为22.0 dBsm,S频段下为20.9 dBsm,C频段下为19.8 dBsm;机尾向(方位角为 180°)L频段下的RCS为 -1.7 dBsm,S频段下为 -12.6 dBsm,C频段下为 -14.0 dBsm。随着探测视角变化,飞机相对于各雷达站的RCS变化如图7(b)所示。

4.2.2 工作策略设计

为逼真模拟对方雷达对抗过程中的工作状态变化,设计对抗规则与策略主要如下:

(1) 工作状态约束与转换规则:为便于效果分析与对抗规则设计,这里将雷达站的工作状态类别简化为无干扰/干扰下搜索、无干扰/干扰下跟踪共4种状态。各雷达站初始为无干扰下搜索状态,在连续 N 个仿真步长内不少于 M 次(这里设置 $N=4, M=3$)接收机内信干噪比大于目标检测所需信噪比时(差异值超过3 dB),视为检测到目标,转为跟踪状态。在连续 N 个仿真步长内超过 M 次接收机内信干噪比小于目标检测所需信噪比时(差异值超过3 dB),视为检测到干扰,转为干扰下搜索或跟踪状态。在

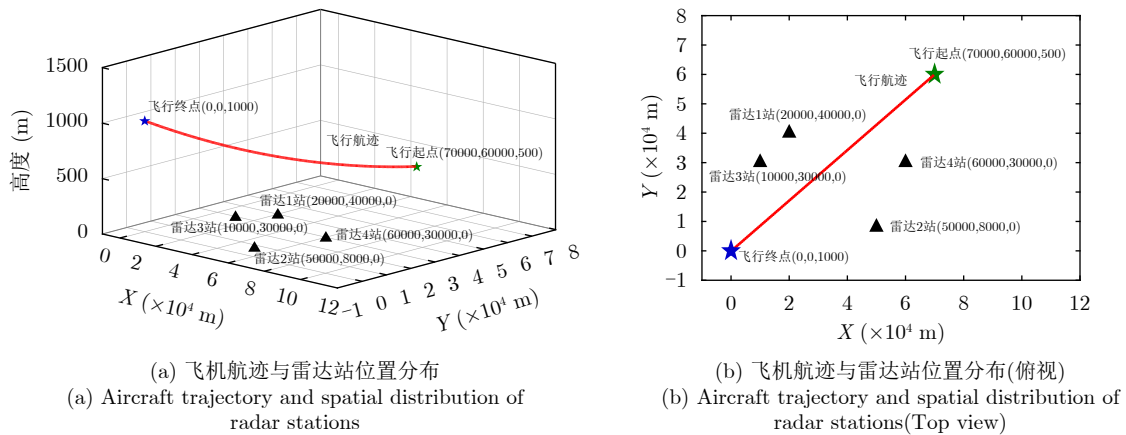


图6 仿真对抗场景示意

Fig. 6 Schematic of spatial distribution of aircraft and radar Stations

表6 机载电子战系统主要典型参数

Tab. 6 The main typical parameters of airborne electronic warfare systems

参数类型	具体取值指标
等效辐射功率	100 W
动态范围	70 dB
步进精度	1 dB
干扰样式	宽带压制干扰(噪声干扰)、转发欺骗干扰(切片干扰)
干扰带宽	宽带压制干扰: [50 MHz, 100 MHz, 150 MHz, 200 MHz, 250 MHz, 300 MHz]共6种带宽; 转发欺骗干扰: 与雷达信号带宽相同
接收损耗	4 dB
极化失配损失	3 dB

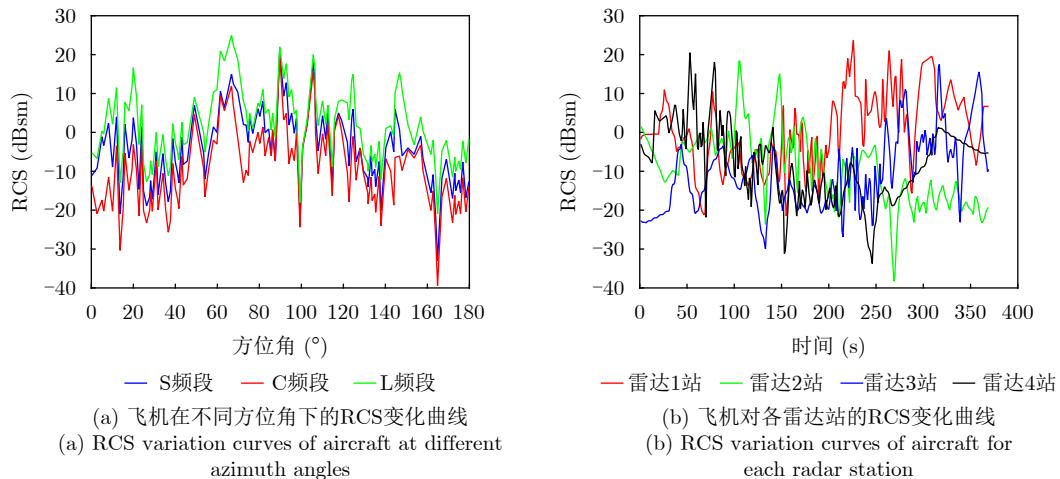


图7 飞机RCS变化曲线

Fig. 7 RCS variation curves of aircraft

连续 N 个仿真步长内不少于 M 次接收机内信干噪比小于目标检测所需信噪比时(差异值不超过3 dB),视为未检测到目标和干扰,转为无干扰下搜索状态。

(2) 驻留时间策略。雷达在无干扰下搜索状态时采用初始驻留时间,转为干扰下搜索以及跟踪状态后,单个仿真步长内的驻留时间变长,雷达信号处理中脉冲累积数目增多。

(3) 工作频点策略。当雷达未遭受到明显电磁干扰时,即接收机内信干噪比略低于目标检测所需信噪比(差异值不超过3 dB),采用单频点进行搜索或跟踪;当遭受明显电磁干扰时,采用频率捷变措施进行抗干扰,且受扰时间越长,雷达工作频点使用范围逐步增大,使得机载无源侦察设备在单位观测时间内测频结果的方差增大。

(4) 信号波形切换策略。雷达信号波形区分搜索信号组合和跟踪信号组合,在无干扰下搜索状态时使用搜索信号组合,在跟踪状态时使用跟踪信号组合,且目标抵近至近距离后会切换至重频更高的跟踪信号组合,以适应不同探测距离与目标跟踪精度需求;在干扰下搜索状态时,此时会增大驻留时间、采用高重频信号,即采取跟踪信号组合以提升目标发现概率。

(5) 雷达组网工作规则:各雷达全程开机对飞机航线方向进行对空搜索,探测到目标后统一上报到集中式节点进行信息融合;当任一雷达可实现对目标的搜索或跟踪,雷达组网探测系统的工作状态即为搜索状态或跟踪状态。

4.3 模型训练及结果分析

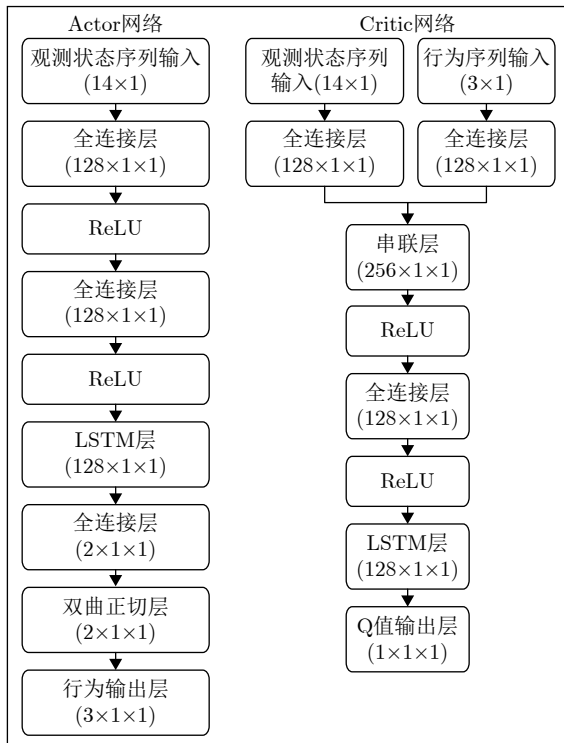
根据仿真场景与参数设置对智能体进行训练,主要训练参数设置:最大训练回合数为10000,经

验回放池大小为10000,训练批量尺寸Batchsize=256,回放样本采样连续个数为16,模型软更新参数 $\tau=0.1$,奖励衰减因子为0.9,初始探索噪声方差为1;奖励函数中 $\lambda_1 \sim \lambda_6$ 比例系数按照不同奖励项的重要程度,信干噪比差异控制优先级最高、干扰带宽控制优先级次之,分别对其取值为2, 1, 0.5, 0.01, 0.01, 0.001(不同取值可能会影响网络训练的收敛速度和最终的策略效果,这里主要关注方法的可行性,对参数的最优选取不展开分析)。TD3中的Actor网络与目标Actor网络结构相同,Critic_1网络与Critic_2网络、目标Critic_1网络、目标Critic_2网络结构相同,具体网络结构如图8(a)所示。为保证训练收敛效果,对观察空间、行为空间内的要素设置变化范围以实现归一化,减小不同要素之间的取值起伏。训练结果如图8(b)所示。

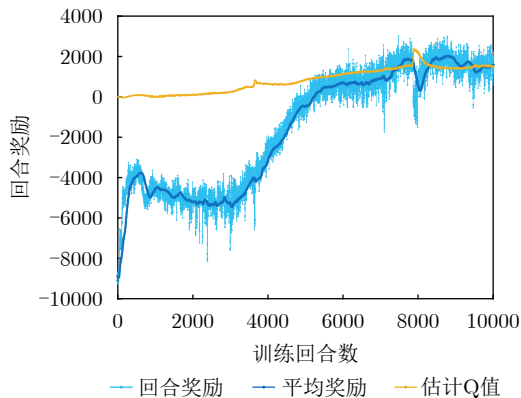
4.3.1 空中平台隐身性能分析

从经验回放池中随机选取500个样本进行测试,由于各雷达站的接收机检测因子分别设置为12 dB, 21 dB, 21 dB, 13 dB,实际对抗中,接收机的检测因子可以根据目标雷达对典型目标在既定范围内的检测概率、虚警率进行估计确定,这里设计对各雷达站的信干噪比有效控制范围分别为9~12 dB, 18~21 dB, 18~21 dB, 10~13 dB。对各雷达站的信干噪比控制、干扰策略选择测试结果如图9、图10所示。

从图9可以看出,智能体具备较好的信干噪比控制精度,在信号功率、工作频点、信号形式等动态变化的雷达信号环境下,约87.2%的仿真时间内可将雷达1站接收机内信干噪比控制在9~12 dB范围内,约85.8%, 85.6%时间内可将雷达2站、雷达3站接收机内信干噪比控制在18~21 dB范围内,约



(a) Actor与Critic网络结构图
(a) Network structure of Actor and Critic



(b) 智能体训练迭代结果图
(b) Agent training iteration result

图8 网络结构与智能体训练结果

Fig. 8 Network structure and agent training result

92.4%时间内可将雷达4站接收机内信干噪比控制在10~13 dB范围内, 总体具备较好的控制精度。从图10中可以看出, 智能体具备高灵活度的干扰策略, 可通过推理当前雷达工作状态与受扰程度, 阶段性变换干扰信号样式、功率与带宽, 且为提升干扰的隐蔽性, 干扰样式以不同带宽的噪声干扰为主, 以提升对不同雷达信号的干扰有效性和持续性。

为对比智能干扰策略对各雷达站的工作状态影响, 分别给出未使用干扰策略、使用固定干扰策略下各雷达站工作状态对飞机的工作状态变化。

(1) 未使用干扰策略

飞机在未使用干扰策略下进行抵近, 各雷达站工作状态变化如图11和表7所示。可以看出, 飞机凭借在S,C频段的低RCS可对雷达2站、3站、4站实现一定的雷达被动隐身(无干扰下搜索段), 但在L频段的RCS较大, 基本可被雷达1站全程跟踪。综合各雷达站的探测情况, 飞机在雷达组网探测系统下可被全程跟踪, 即无干扰下跟踪状态占比100%, 雷达被动隐身性能总体较差。

(2) 使用固定干扰策略

飞机在固定干扰策略下进行抵近, 具体策略设计为: 侦察到雷达信号后, 全程释放电磁干扰, 首先使用最大干扰功率率(100 W)、最大干扰带宽(300 MHz)的宽带噪声对特定方向的雷达站进行压制, 感知到雷达信号数据率上升后, 判断雷达可能由搜索转跟踪状态, 使用最大干扰功率(100 W)的转发欺骗干扰迅速降低雷达站对其检测概率。各雷达站及雷达组网探测系统对飞机的工作状态变化如图12所示, 对应工作状态具体占比情况如表8所示。可以看出, 飞机在固定电磁干扰策略下可有效压缩各雷达及组网探测系统的探测范围, 使其全程无法跟踪目标, 但干扰下搜索状态占比98.92%, 各雷达站持续检测到电磁干扰, 飞机抵近意图暴露, 未实现雷达隐身。

(3) 使用智能干扰策略

飞机在使用智能干扰策略后, 各雷达站及雷达组网探测系统对飞机的工作状态变化如图13所示, 各雷达工作状态具体占比情况如表9所示。可以看出, 所生成的干扰策略对飞机掩护效果良好, 可大幅降低各雷达站对飞机与干扰的检测性能, 对各雷达站形成的隐身段占比, 即无干扰下搜索状态占比, 不小于88.0%, 飞机在雷达组网下的隐身段占比提升至79.95%。其中, 在L波段(雷达1站)且探测视角始终处于飞机侧向(RCS极大值区)时, 未使用干扰策略时雷达1站对飞机的稳定跟踪占比高达98.37%, 通过干扰功率和样式的动态优选, 雷达对飞机稳定跟踪占比降至4.07%。图14给出了雷达组网下对飞机的探测态分布, 可以看出当飞机抵近时隐身性能较好, 远离时在部分段上可被雷达组网探测系统实现跟踪, 结合图6分析, 由于飞机在远离时RCS变化幅度较大, 当雷达回波信号强度起伏明显时, 导致干扰参数自动调整滞后, 飞机主动隐身性能下降, 因此干扰策略适应性还存在优化空间。

综合上述分析表明, 飞机在智能干扰策略下的主动隐身性能显著提升, 有效弥补了传统被动隐身技术在低频段、大角度下的性能缺陷, 也大幅降低

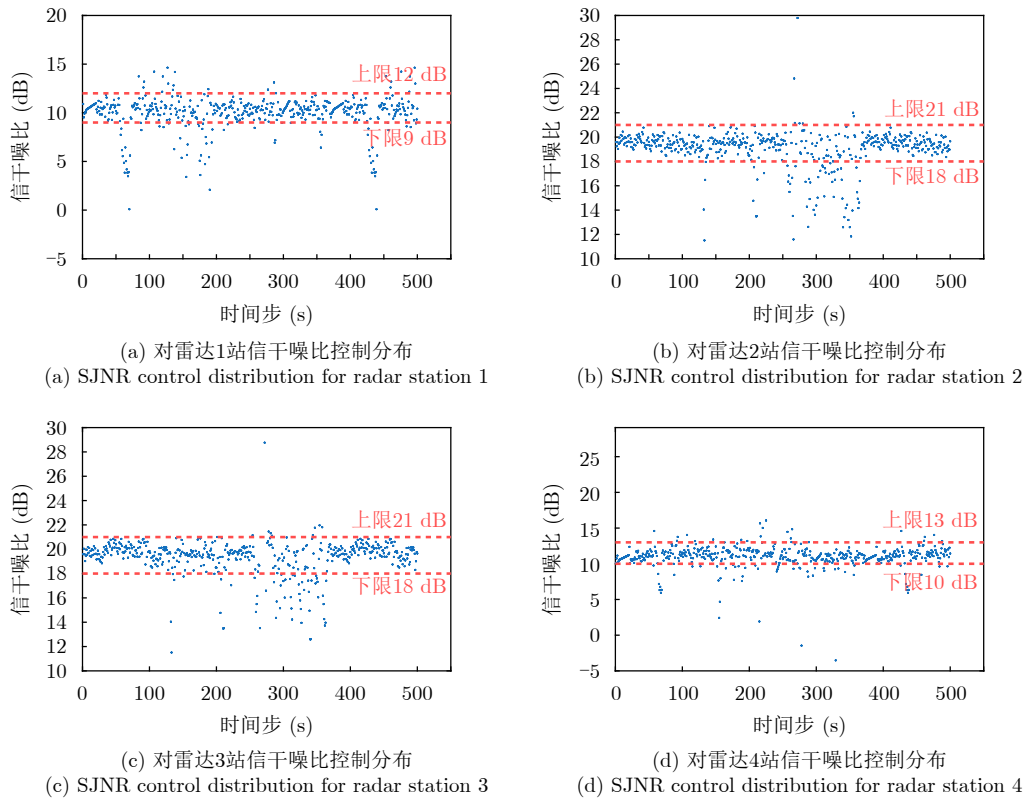


图 9 各雷达站信干噪比控制结果
Fig. 9 SJNR control results for each radar station

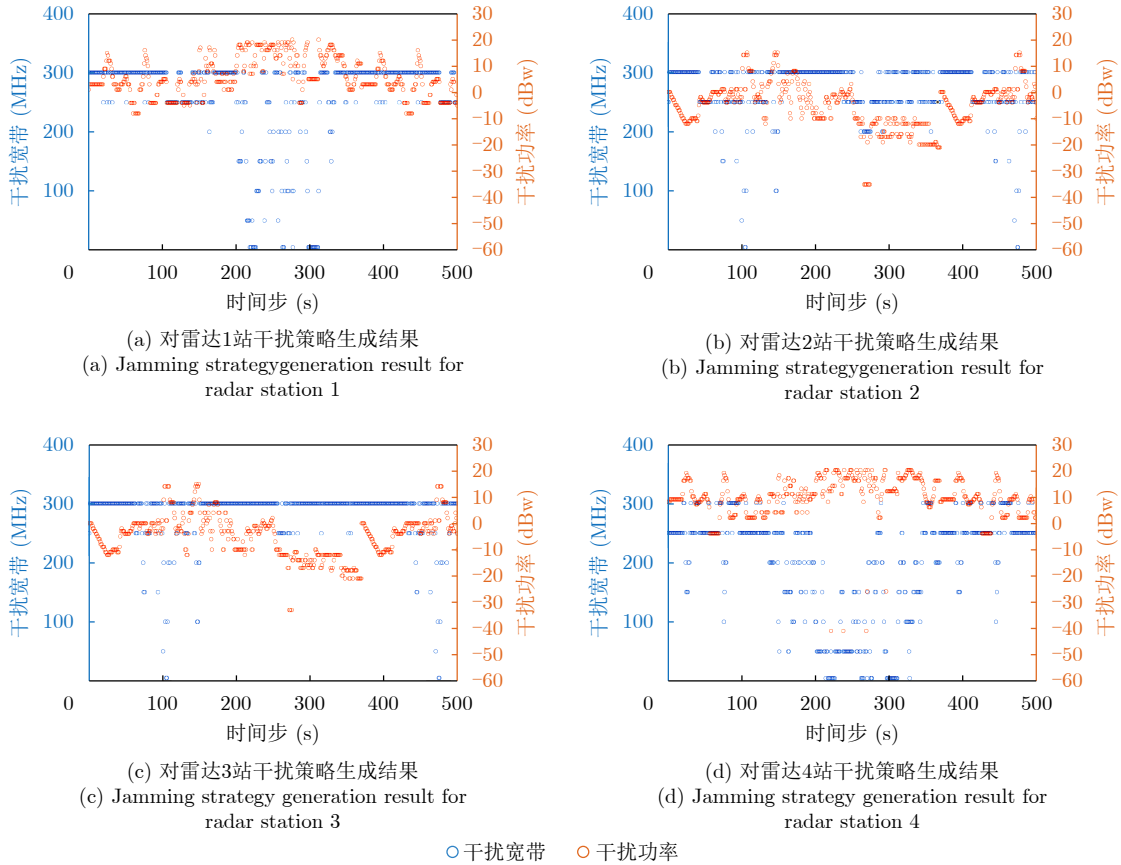


图 10 对各雷达站的干扰策略结果
Fig. 10 Jamming strategy generation result for each radar station

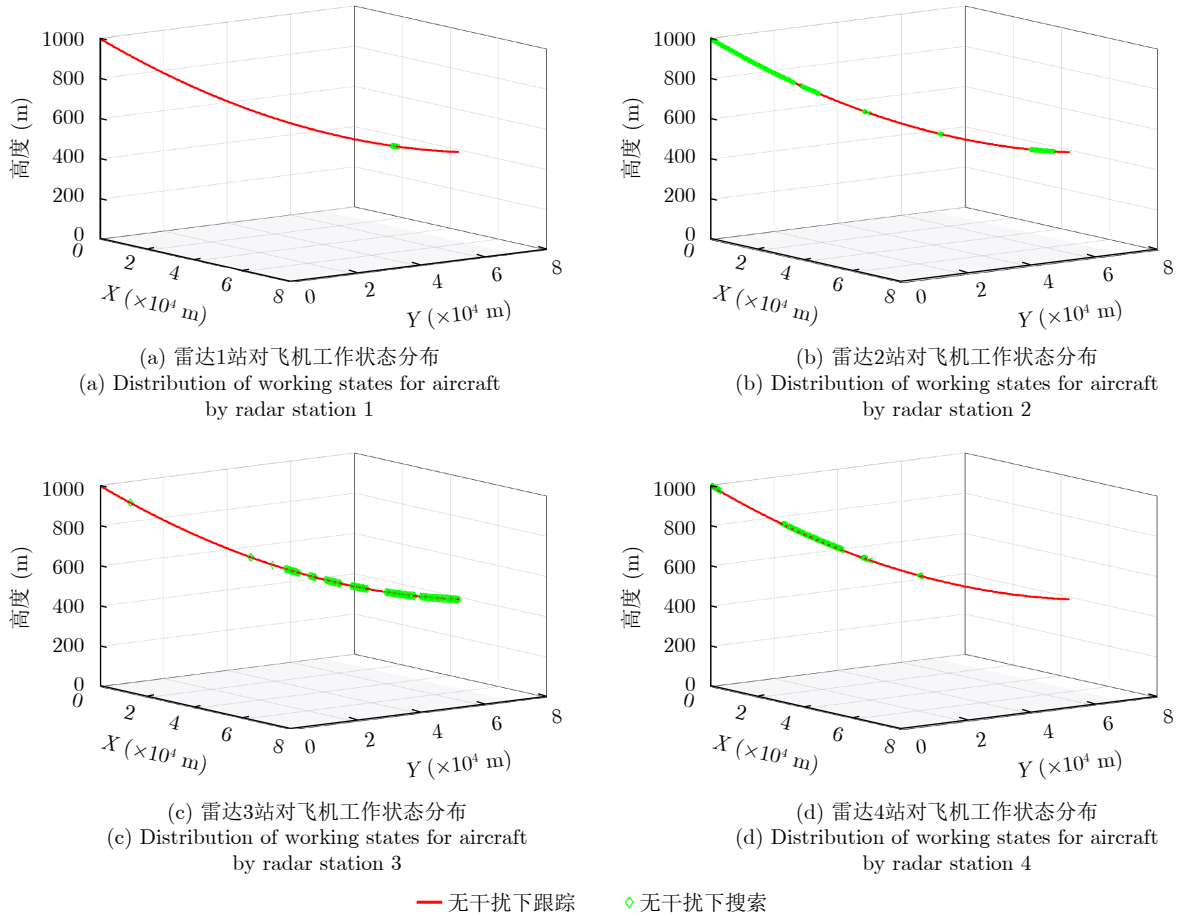


图 11 未使用干扰策略下各雷达站工作状态情况

Fig. 11 The working states of each radar station without jamming strategy

表 7 未使用干扰策略下各雷达站及雷达组网探测系统工作状态占比情况(%)

Tab. 7 The working state proportions of each radar station and radar network detection system without jamming strategy (%)

类型	无干扰下搜索	无干扰下跟踪
雷达1站	1.63	98.37
雷达2站	37.13	62.87
雷达3站	35.77	64.23
雷达4站	22.22	77.78
雷达组网探测系统	0	100

了固定干扰策略下抵近意图暴露概率，雷达隐身性能总体良好。

进一步分析飞机在使用智能干扰策略前后的 RCS 等效缩减量。对未使用干扰策略下飞机暴露段与使用智能干扰策略下各雷达站的无干扰下搜索段进行提取，并进一步提取两者的重叠段。分别统计重叠段内，未使用干扰策略下雷达接收机内 SNR、使用智能干扰策略下的 SJNR，依据式(7)计算对各雷达站的 RCS 平均等效缩减量，如表 10 所示。其中，对 L 频段的雷达 1 站平均 RCS 等效缩减量最大，

为 9.40 dB，对 C 频段的雷达 3 站最小，为 2.84 dB，进一步体现了智能干扰策略对低频段下被动隐身性能的弥补作用。

4.3.2 电磁干扰隐蔽性分析

为分析电磁干扰的低可探测性，结合表 9 中给出的各雷达站工作状态进行分析。其中，干扰下搜索状态与干扰下跟踪状态代表雷达站感知到了干扰，因此通过计算各雷达站在对抗过程中感知到电磁干扰的时间占比，量化电磁干扰隐蔽性，如表 11 所示。可以看出，由于智能体对电磁干扰的精准控制，飞机抵近过程中电子战系统对各雷达站的电磁干扰隐蔽占比不低于 97.83%，对雷达组网下为 94.58%，可知电磁干扰隐蔽性表现良好。

总的来说，基于强化学习的智能体在典型威胁场景下经过多轮次迭代训练后，初步具备以“观察-分析-决策-行动-评估”环路为核心机理的认知电子战特性，在动态电磁信号环境下可通过前期嵌入基于专家经验的数据分析准则，自动生成高灵活性、强适应性的电磁干扰策略，可有效提升空中目标的雷达隐身性能。

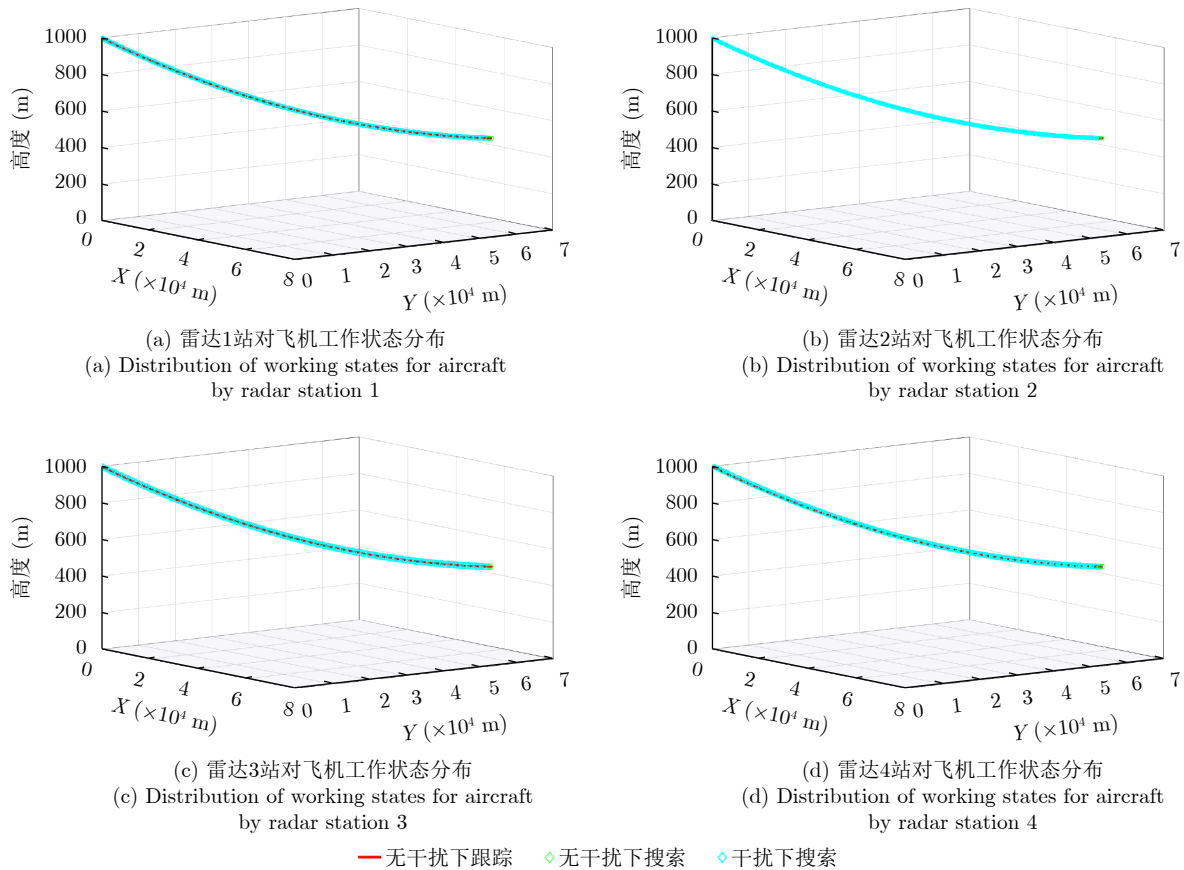


图 12 使用固定干扰策略下各雷达站工作状态情况

Fig. 12 The working states of each radar station with fixed jamming strategy

表 8 使用固定干扰策略下各雷达站及雷达组网探测系统工作状态占比情况(%)

Tab. 8 The working state proportions of each radar station and radar network detection system with fixed jamming strategy(%)

类型	无干扰下搜索	干扰下搜索	无干扰下跟踪	干扰下跟踪
雷达1站	1.08	98.92	0.00	0.00
雷达2站	1.08	98.92	0.00	0.00
雷达3站	1.08	98.92	0.00	0.00
雷达4站	1.08	98.92	0.00	0.00
雷达组网探测系统	1.08	98.92	0.00	0.00

4.3.3 适用边界分析

虽然仿真结果表明所提方法能够有效提升空中平台面对分布式雷达组网探测系统时的雷达隐身性能,但仍受到条件约束,主要包括:(1)雷达信号可侦察性约束。当对方雷达采用低截获概率波形、短时突发辐射、严格静默工作或由被动探测系统辅助完成目标发现时,电子战系统可获得的观测信息可能不完整,从而影响雷达工作状态识别和干扰策略生成效果。(2)电子战系统资源约束。主动隐身策略的实现受到干扰系统的工作频率范围、瞬时带宽、最大辐射功率、波束指向能力、动态响应速度和同时作用目标数量等因素限制。当雷达工作频段

超出电子战系统覆盖范围,或同时威胁雷达数量超过系统资源承载能力时,主动隐身效果将受到影响。(3)雷达抗干扰能力约束。当对方雷达具备较强频率捷变、波形捷变、多站协同处理、自适应波束形成或多源信息融合能力时,单一电子战系统的干扰策略可能难以持续降低对方所有雷达的目标检测信干噪比,需要与编队协同、航迹规划和被动隐身设计等手段结合使用。(4)环境与先验信息约束。智能干扰策略生成需要结合外部环境测量信息、先验信息、平台RCS数据库和雷达信号处理增益估计等构建观测空间,当外部测量及先验信息误差较大,或外部电磁环境显著偏离训练样本分布时,生成的干扰策略可能出现性能下降,在工程应

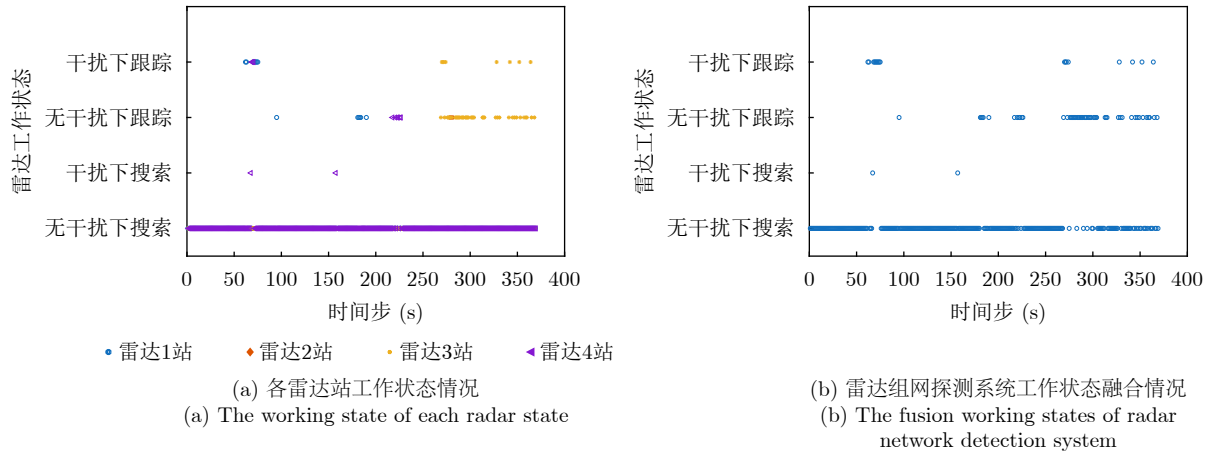


图 13 使用智能干扰策略下各雷达站及雷达组网探测系统工作状态情况

Fig. 13 The working state of each radar station and radar network detection system with intelligent jamming strategy

表 9 使用智能干扰策略下各雷达站及雷达组网探测系统工作状态占比情况(%)

Tab. 9 The working state proportions of each radar station and radar network detection system with intelligent jamming strategy (%)

类型	无干扰下搜索	干扰下搜索	无干扰下跟踪	干扰下跟踪
雷达1站	95.93	0.00	1.90	2.17
雷达2站	99.19	0.00	0.81	0.00
雷达3站	88.00	0.00	10.40	1.60
雷达4站	97.02	0.54	1.63	0.81
雷达组网探测系统	79.95	0.54	14.63	4.88

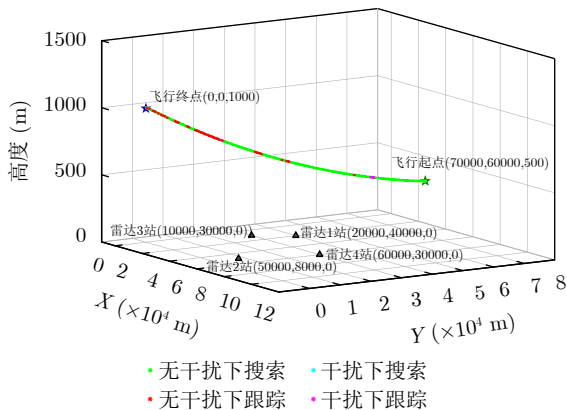


图 14 使用智能干扰策略下雷达组网探测系统工作状态情况

Fig. 14 The working states of radar network detection system with intelligent jamming strategy

表 10 相对各雷达站的 RCS 平均等效缩减量 (dB)

Tab. 10 The average equivalent RCS reduction for each radar station (dB)

类型	最大 RCS 等效缩减量	最小 RCS 等效缩减量	平均 RCS 等效缩减量
雷达1站	23.67	-21.64	9.40
雷达2站	18.43	-33.13	2.91
雷达3站	17.52	-26.78	2.84
雷达4站	20.51	-31.30	3.24

表 11 使用智能干扰策略下对各雷达站与雷达组网的电磁干扰隐蔽性比例 (%)

Tab. 11 The jamming stealth ratio for each radar station and radar network detection system with intelligent jamming strategy (%)

类型	电磁干扰暴露比例	电磁干扰隐蔽比例
雷达1站	2.17	97.83
雷达2站	0.00	100.00
雷达3站	1.60	98.40
雷达4站	1.35	98.65
雷达组网探测系统	5.42	94.58

用中需要通过在线更新、仿真样本扩充和专家规则约束等方式提高策略鲁棒性。

因此, 所提方法更适用于对方雷达存在可侦察电磁辐射信号、电子战系统具备相应频段覆盖与干扰资源基础, 且环境与先验信息具备较高准确度的对抗场景, 可针对传统被动隐身技术在低频段、大角度下的性能缺陷进行补充增强。

5 结语

针对对方分布式多频段雷达组网探测系统对我方空中平台构成的全域探测威胁, 以及传统被动隐身在多方向、宽频段场景下性能受限的问题, 本文

将认知电子战与主动隐身深度融合，以影响雷达工作状态变化为出发点，提出一种基于认知电子战的空中平台主动隐身方法。通过构建融合信号实测、先验信息与实时估计的观测空间，建立专家经验约束的干扰行为空间，设计面向雷达工作状态影响、电磁暴露风险抑制的分层奖励函数，采用TD3强化学习算法实现自适应智能干扰策略自主生成与在线优化，形成“感知-决策-干扰-评估”闭环架构。仿真结果表明，所提方法可将雷达稳定停留在无干扰搜索状态，有效压缩雷达组网探测系统对目标稳定探测范围，实现 RCS 等效缩减与电磁干扰隐蔽性协同提升，干扰参数能够随雷达信号特征与外部环境实时自适应调整，在动态复杂电磁环境下具备优异鲁棒性与泛化能力。本文工作突破了传统隐身仅依赖 RCS 物理缩减的技术路径，为“隐身+电子战”一体化主动隐身方法提供了理论框架与工程可行技术途径，对提升空中平台复杂电磁环境下生存能力具有重要参考价值。后续将面向多平台协同与多威胁并行对抗场景，开展多智能体协同干扰决策、分布式资源协同优化等研究，重点解决协同决策机制设计、容许观测误差下干扰策略泛化等问题，进一步拓展主动隐身的适用边界。

附录A 典型雷达信号类型的信号处理增益计算分析

当雷达信号为线性调频信号(Linear Frequency Modulation, LFM)时，其信号处理增益 D 主要包括脉压增益 D_1 与累积增益 D_2 。 D_1 可以根据信号脉宽、带宽 B 的测量结果进行估算，即 $D_1 = B\tau$ ； D_2 需要结合对单个相干处理间隔(Coherent Processing Interval, CPI)内雷达脉冲数目 N 的测量结果进行估算，即 $D_2 = N^\eta$ ，其中 η 的取值范围一般为 0.7~0.9。比例系数 k_1 需要根据具体的干扰样式确定，一般可以通过充分的静态试验获取经验值。

当雷达信号为频率编码、相位编码等非线性调频信号时，常规的信号处理增益计算方式不再适用，需要综合考虑信号编码长度、编码方式等进行计算。其中，脉内频率编码信号的信号处理增益计算方式一般与 LFM 信号相同，结合对单个脉冲信号的 τ, B 测量结果计算 D_1 ；由于频率编码破坏了信号的相参性，可能导致部分累积损耗，相较于 LFM 信号， η 的取值更小。假设脉内相位编码信号的编码长度为 m ，子脉冲宽度为 τ_c ， D_1 一般为

$$D_1 = B\tau = \frac{1}{\tau_c} \times m\tau_c = m$$

同样由于相位编码破坏了信号的相参性，相较于 LFM 信号， η 的取值更小。

附录B 主流强化学习算法对比分析

对双延迟深度确定性策略梯度算法(Twin Delayed Deep Deterministic Policy Gradient, TD3)、深度确定性策略梯度算法(Deep Deterministic Policy Gradient, DDPG)、近端策略优化算法(Proximal Policy Optimization, PPO)等相关主流算法进行对比分析。

训练收敛速度方面，DDPG 算法可适配连续动作空间优化问题，但存在 Q 值过估计缺陷，评论家网络对动作价值的高估会导致演员网络的策略更新偏差，需通过大量迭代修正，收敛速度较慢。PPO 算法通过裁剪目标函数可避免策略更新幅度过大，解决了策略梯度算法不稳定性问题，收敛速度优于 DDPG，但需要针对连续动作空间需进行离散化处理，会损失参数优化精度，一定程度上会降低收敛效率。TD3 算法在 DDPG 基础上引入双评论家网络和策略延迟更新机制，可有效抑制 Q 值过估计问题，同时针对连续动作空间无需额外离散化处理，在高维度连续动作空间中收敛效率高于 PPO 算法。

生成策略稳定性方面，DDPG 算法因 Q 值过估计和单网络更新的固有缺陷，策略更新过程中易出现震荡发散问题，即使达到收敛状态，受电磁环境微小噪声影响，收敛曲线也会出现大幅波动，生成的干扰策略稳定性较差。PPO 算法通过信任域裁剪和批量更新机制，策略更新过程更平滑，收敛后波动幅度小于 DDPG，策略稳定性中等，但在连续动作空间的适配性不足，导致策略在最优解附近小幅震荡。TD3 算法的双评论家网络偏差抑制、策略延迟更新的误差累积减少，以及动作噪声正则化的引入，使策略网络的更新过程更平稳，收敛曲线的震荡幅度远低于 DDPG，能实现干扰参数的精细化连续优化，策略收敛后可稳定在最优解附近。

抗噪声能力方面，DDPG 算法对环境噪声较为敏感，噪声会进一步放大 Q 值过估计偏差，导致演员网络策略更新方向错误，当电磁环境噪声超过一定阈值时，策略易发散，抗噪声能力较弱。PPO 算法因策略更新的信任域约束，对环境噪声的耐受度高于 DDPG，少量噪声不会导致策略大幅偏离最优解，但在强噪声环境下，离散化的动作空间无法快速适配噪声带来的环境变化，策略调整的灵活性不足，抗噪声能力中等。TD3 算法在动作噪声正则化的基础上，通过目标策略平滑机制对目标动作添加小幅噪声，使评论家网络在训练过程中适应环

境的随机波动,提升了算法对噪声的鲁棒性,同时连续动作空间可快速补偿噪声带来影响。

利益冲突 所有作者均声明不存在利益冲突

Conflict of Interests The authors declare that there is no conflict of interests

参 考 文 献

- [1] XU Fan, LAI Wenhai, and SHEN Kaiming. Intelligent surface assisted radar stealth against unauthorized ISAC[J]. *IEEE Wireless Communications Letters*, 2025, 14(4): 1149–1153. doi: [10.1109/LWC.2025.3535921](https://doi.org/10.1109/LWC.2025.3535921).
- [2] TIWARI P, PATHAK S K, and SIJU V. Design, development and characterization of resistive arm based planar and conformal metasurfaces for RCS reduction[J]. *Scientific Reports*, 2022, 12(1): 14992. doi: [10.1038/s41598-022-19075-x](https://doi.org/10.1038/s41598-022-19075-x).
- [3] SUN Chengtao, LI Dawei, LIAO Wenhe, *et al.* Rigid-flexible interlocked metastructures enable conformal stealth[J]. *Science Advances*, 2025, 11(49): eacb7870. doi: [10.1126/sciadv.aeb7870](https://doi.org/10.1126/sciadv.aeb7870).
- [4] OUYANG Wenchong, DING Chengbiao, LIU Qi, *et al.* Arrayed multiple atmospheric-pressure plasma jet sources for active stealth[J]. *Cell Reports Physical Science*, 2024, 5(1): 101771. doi: [10.1016/j.xcrp.2023.101771](https://doi.org/10.1016/j.xcrp.2023.101771).
- [5] CHANG Qi, JI Jinzu, WU Wenxing, *et al.* An optically transparent metamaterial absorber with tunable absorption bandwidth and low infrared emissivity[J]. *Materials*, 2023, 16(23): 7357. doi: [10.3390/ma16237357](https://doi.org/10.3390/ma16237357).
- [6] ZHANG Haonan, GUO Qingxin, LIU Jinbo, *et al.* Low radar cross section metasurface based on polarization converter[C]. 2024 IEEE International Symposium on Antennas and Propagation and INC/USNC-URSI Radio Science Meeting, Firenze, Italy, 2024: 1321–1322. doi: [10.1109/AP-S/INC-USNC-URSI52054.2024.10686477](https://doi.org/10.1109/AP-S/INC-USNC-URSI52054.2024.10686477).
- [7] LU Yao, SU Jianxun, LIU Jinbo, *et al.* Ultrawideband monostatic and bistatic RCS reductions for both copolarization and cross polarization based on polarization conversion and destructive interference[J]. *IEEE Transactions on Antennas and Propagation*, 2019, 67(7): 4936–4941. doi: [10.1109/TAP.2019.2911185](https://doi.org/10.1109/TAP.2019.2911185).
- [8] RAJABALIPANAH H and ABDOLALI A. Ultrabroadband monostatic/bistatic RCS reduction via high-entropy phase-encoded polarization conversion metasurfaces[J]. *IEEE Antennas and Wireless Propagation Letters*, 2019, 18(6): 1233–1237. doi: [10.1109/LAWP.2019.2913465](https://doi.org/10.1109/LAWP.2019.2913465).
- [9] SU Jianxun, LI Wenyu, QU Meijun, *et al.* Ultrawideband RCS reduction metasurface based on hybrid mechanism of absorption and phase cancellation[J]. *IEEE Transactions on Antennas and Propagation*, 2022, 70(10): 9415–9424. doi: [10.1109/TAP.2022.3184538](https://doi.org/10.1109/TAP.2022.3184538).
- [10] 刘雄. 主动隐身技术中目标的有源对消效果研究[D]. [硕士学位论文], 电子科技大学, 2018.
LIU Xiong. Research on cancellation effect of targets in active stealth technology[D]. [Master dissertation], University of Electronic Science and Technology of China, 2018.
- [11] 边晓臣, 黄沛霖, 姬金祖. 基于线性调频波的有源对消隐身仿真及分析[J]. 北京航空航天大学学报, 2016, 42(8): 1769–1776. doi: [10.13700/j.bh.1001-5965.2015.0492](https://doi.org/10.13700/j.bh.1001-5965.2015.0492).
BIAN Xiaochen, HUANG Peilin, and JI Jinzu. Simulation and analysis of active cancellation stealth based on LFM wave[J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2016, 42(8): 1769–1776. doi: [10.13700/j.bh.1001-5965.2015.0492](https://doi.org/10.13700/j.bh.1001-5965.2015.0492).
- [12] LEE I G, YOON Y J, CHOI K S, *et al.* Design of an optical transparent absorber and defect diagnostics analysis based on near-field measurement[J]. *Sensors*, 2021, 21(9): 3076. doi: [10.3390/s21093076](https://doi.org/10.3390/s21093076).
- [13] 王家兴. 有源对消隐身试验系统关键技术研究[D]. [硕士学位论文], 电子科技大学, 2024. doi: [10.27005/d.cnki.gdzku.2024.000532](https://doi.org/10.27005/d.cnki.gdzku.2024.000532).
WANG Jiaying. Research on active stealth cancellation test system and its key technologies[D]. [Master dissertation], University of Electronic Science and Technology of China, 2024. doi: [10.27005/d.cnki.gdzku.2024.000532](https://doi.org/10.27005/d.cnki.gdzku.2024.000532).
- [14] SENGUPTA S, COUNCIL H, JACKSON D R, *et al.* Active radar cross section reduction of an object using microstrip antennas[J]. *Radio Science*, 2020, 55(2): 1–20. doi: [10.1029/2019RS006939](https://doi.org/10.1029/2019RS006939).
- [15] 冯清帅. 电磁散射调控理论与技术研究[D]. [硕士学位论文], 电子科技大学, 2022. doi: [10.27005/d.cnki.gdzku.2022.003203](https://doi.org/10.27005/d.cnki.gdzku.2022.003203).
FENG Qingshuai. Research on electromagnetic scattering control theory and technology[D]. [Master dissertation], University of Electronic Science and Technology of China, 2022. doi: [10.27005/d.cnki.gdzku.2022.003203](https://doi.org/10.27005/d.cnki.gdzku.2022.003203).
- [16] 刘松涛, 雷震烁, 温镇铭, 等. 认知电子战研究进展[J]. 探测与控制学报, 2020, 42(5): 1–15.
LIU Songtao, LEI Zhenshuo, WEN Zhenming, *et al.* A development review on cognitive electronic warfare[J]. *Journal of Detection & Control*, 2020, 42(5): 1–15.
- [17] BARBOSA M, PRALON L, RAMOS A L L, *et al.* On a closer look of a Doppler tolerant noise radar waveform in surveillance applications[J]. *Sensors*, 2024, 24(8): 2532. doi: [10.3390/s24082532](https://doi.org/10.3390/s24082532).
- [18] CHEN Dong, ZHANG Kaixiang, WANG Yongqiang, *et al.* Communication-efficient decentralized multi-agent reinforcement learning for cooperative adaptive cruise control[J]. *IEEE Transactions on Intelligent Vehicles*, 2024, 9(10): 6436–6449. doi: [10.1109/TIV.2024.3368025](https://doi.org/10.1109/TIV.2024.3368025).

- [19] XU Yuting, WANG Chao, LIANG Jiakai, *et al.* Deep reinforcement learning based decision making for complex jamming waveforms[J]. *Entropy*, 2022, 24(10): 2441. doi: [10.3390/e24101441](https://doi.org/10.3390/e24101441).
- [20] 鲁永为, 张赛楠, 郭慧峰, 等. 基于间歇采样延时叠加的干扰时序研究[J]. *现代雷达*, 2020, 42(5): 52–56. doi: [10.16592/j.cnki.1004-7859.2020.05.010](https://doi.org/10.16592/j.cnki.1004-7859.2020.05.010).
LU Yongwei, ZHANG Sainan, GUO Hui Feng, *et al.* A study on jamming time sequence based on interrupted sampling with time-delay superposition[J]. *Modern Radar*, 2020, 42(5): 52–56. doi: [10.16592/j.cnki.1004-7859.2020.05.010](https://doi.org/10.16592/j.cnki.1004-7859.2020.05.010).
- [21] 潘小义, 刘晓斌, 陈吉源, 等. 间歇采样转发干扰技术研究述评[J]. *系统工程与电子技术*, 2024, 46(9): 2887–2907. doi: [10.12305/j.issn.1001-506X.2024.09.01](https://doi.org/10.12305/j.issn.1001-506X.2024.09.01).
PAN Xiaoyi, LIU Xiaobin, CHEN Jiyuan, *et al.* Overview of intermittent sampling repeater jamming technology[J]. *Systems Engineering and Electronics*, 2024, 46(9): 2887–2907. doi: [10.12305/j.issn.1001-506X.2024.09.01](https://doi.org/10.12305/j.issn.1001-506X.2024.09.01).
- [22] 譙梁, 杨帅, 王鑫, 等. 雷达干扰效果评估与协同干扰策略分配算法研究[J]. *航天电子对抗*, 2019, 35(3): 27–32. doi: [10.3969/j.issn.1673-2421.2019.03.007](https://doi.org/10.3969/j.issn.1673-2421.2019.03.007).
QIAO Liang, YANG Shuai, WANG Xin, *et al.* Effect evaluation of radar interference and cooperative interference strategy allocation algorithm[J]. *Aerospace Electronic Warfare*, 2019, 35(3): 27–32. doi: [10.3969/j.issn.1673-2421.2019.03.007](https://doi.org/10.3969/j.issn.1673-2421.2019.03.007).
- [23] FUJIMOTO S, VAN HOOFF H, and MEGER D. Addressing function approximation error in actor-critic methods[C]. The 35th International Conference on Machine Learning, Stockholm, Sweden, 2018: 1587–1596.
- [24] WANG Di and HU Mengqi. Deep deterministic policy gradient with compatible critic network[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2023, 34(8): 4332–4344. doi: [10.1109/TNNLS.2021.3117790](https://doi.org/10.1109/TNNLS.2021.3117790).
- [25] SAHA U, JAWAD A, SHAHRIA S, *et al.* Proximal policy optimization-based reinforcement learning approach for DC-DC boost converter control: A comparative evaluation against traditional control techniques[J]. *Heliyon*, 2024, 10(18): e37823. doi: [10.1016/j.heliyon.2024.e37823](https://doi.org/10.1016/j.heliyon.2024.e37823).
- [26] 李健涛, 王轲昕, 刘凯, 等. 基于深度强化学习的干扰资源分配方法[J]. *现代雷达*, 2023, 45(10): 44–51. doi: [10.16592/j.cnki.1004-7859.2023.10.005](https://doi.org/10.16592/j.cnki.1004-7859.2023.10.005).
LI Jiantao, WANG Kexin, LIU Kai, *et al.* Jamming resource allocation method based on deep reinforcement learning[J]. *Modern Radar*, 2023, 45(10): 44–51. doi: [10.16592/j.cnki.1004-7859.2023.10.005](https://doi.org/10.16592/j.cnki.1004-7859.2023.10.005).

作者简介

刘志鹏, 工程师, 主要研究方向为雷达信号处理、电磁干扰策略设计和雷达射频隐身技术等。

胡宝洁, 高级工程师, 主要研究方向为雷达及雷达对抗技术。

吴 耿, 助理工程师, 主要研究方向为雷达信号处理、雷达射频隐身技术等。

张樱豪, 助理工程师, 主要研究方向为雷达信号处理、雷达射频隐身技术等。

孙震宇, 助理工程师, 主要研究方向为雷达信号处理、电磁干扰策略设计等。

(责任编辑: 于青)