

# 基于深度强化学习的雷达载频捷变方法

王成 刘婷婷 张恒

中国航空工业集团公司西安航空计算技术研究所 陕西西安 710065

**摘要:** 面对复杂电磁环境下的目标感知挑战, 捷变频雷达亟需突破现有性能瓶颈。尽管通过随机调整载频已初步具备低截获特性, 但在应对复杂频谱干扰时仍显不足。为此, 本文从载波频率的波形特征入手, 系统性提升雷达的抗干扰韧性, 以实现在干扰环境中的可靠探测。针对有源干扰场景, 提出基于强化学习的载频优化方案。仿真实验表明, 相较于传统随机载频方法, 所提方案在多种有源干扰下显著提升了探测稳健性。这一创新不仅增强了捷变频雷达的抗干扰能力, 更为智能雷达系统的设计提供了新思路。

**关键词:** 捷变频雷达; 抗干扰; 强化学习

## 1. 引言:

尽管采用捷变频技术的雷达系统能够借助载波频率的快速跳变降低被截获的可能性, 并展现出一定的抗干扰特性, 然而在真实且复杂的电磁作战环境中, 当干扰设备与通信系统同时存在时, 其抗干扰性能仍受到显著制约。随着数学优化理论的持续发展, 捷变频雷达波形设计领域已形成多种获取最优或近似最优解的有效方法。在这一背景下, 随机优化算法<sup>[1-5]</sup>因其独特的搜索机制, 被广泛视为处理该类问题的关键途径。在结构复杂、约束众多的优化任务中, 这些方法展现出良好的适应性与求解效率。然而, 该类方法亦存在一定局限: 其演化过程带有随机性, 致使收敛速度通常较慢, 因而更适用于离线优化场合。在日益复杂的电磁对抗与动态频谱环境中, 传统跳频方案通常依赖预设的固定频率集, 在面对多样化的电磁干扰时表现出明显的适应不足。为适应现代战场对雷达装备实时感知与自主决策能力的要求, 本文借鉴强化学习理论<sup>[6]</sup>, 提出一种基于强化学习的载波频率动态优化方法, 使雷达能够在变化的电磁条件下实时调整频率选择策略。所提出的方法可使雷达系统自主生成适应频谱动态变化的载频序列, 进而在强干扰条件下仍维持可靠的目标探测性能。

## 2. 基于深度强化学习的有源干扰环境载频优化方法

### 2.1 抗干扰马尔可夫模型

雷达工作模式的动态表述: 假设捷变频雷达的工作频段被划分为  $M$  个相互独立的等宽子带, 系统在每个脉冲周期中仅以随机方式选择其中一个子带对应的载波频率进行

脉冲发射。由于本文所采用的是线性调频信号形式, 每一个载波频率均与一个特定子带唯一对应。为准确描述雷达在当前时刻所使用的载波频率或所属子带, 可将该跳频序列建模为动作变量, 其数学表达如下所示:

$$a_t = c_n, c_n \in [0, M-1] \quad (1)$$

状态空间设计: 在本研究中, 将干扰机在电磁环境中表现出的频谱动态特性定义为环境状态的主要组成部分。考虑到干扰策略的制定通常建立在对雷达已发射信号进行采样与分析的基础上, 因此将雷达自身的历史工作频点序列也纳入状态表征体系。基于上述设计原则, 设定系统在  $t$  时刻所感知的环境状态包含以下信息维度:

$$s_t = [s_{ij}, s_{ta}] \quad (2)$$

其中, 干扰频谱特征向量  $s_{ij}$  采用独热编码 (one-hot encoding) 对干扰的频域分布进行离散化描述。举例来说, 若向量中第三个元素为 1, 即表示雷达当前工作的第三个子频带内存在干扰信号。另一部分为雷达在  $t$  时刻之前若干周期内的动作序列向量  $s_{ta}$ 。

动态奖励机制设计: 针对捷变频雷达的抗干扰任务目标, 本文根据各子频带在给定时间区间内是否感知到干扰存在, 建立了如下所示的奖励函数:

$$r(s_t, a_t, s_{t+1}) = \begin{cases} 1, & s_{t+1,j}(a_t) = 0 \\ 0, & s_{t+1,j}(a_t) = 1 \end{cases} \quad (3)$$

可以观察到, 当雷达工作频段未受到干扰影响时, 系统将获得正向奖励; 而在其工作频段遭受干扰的情况下, 则会接收到负向反馈。

决策策略优化设计：为达成捷变频雷达在载波频率选择上有效应对干扰的目标，需对现有决策架构实施系统性改进。所提出的决策机制具体设计如下：

$$\pi(a_i | s_t) = \begin{cases} 1 - \varepsilon & , Q\text{-greedy}(Q(s_t, a_i)) \\ \varepsilon & , a = \text{rand}(A) \end{cases} \quad (4)$$

every time:  $\varepsilon = \varepsilon \times \mu$

在捷变频雷达的干扰对抗任务中，深度强化学习网络通过动态调整动作实施概率  $\varepsilon$  与变化因子  $\mu$ ，实现了探索与利用的精准平衡。其中， $\varepsilon$  作为核心调控参数，决定了网络在每次决策时采取探索性动作的概率，其数值直接影响着学习策略的导向性。而  $\mu$  作为概率变化的关键因子，取值范围严格限定在 (0,1) 区间。

### 2.2 基于深度强化学习的载频优化流程

在完成捷变频抗干扰的马尔可夫决策过程建模与策略设计后，本文提出基于深度强化学习的载频优化流程。系统将可用频带均匀划分为  $M$  个子带，发射机随机选择子带发射脉冲以规避干扰。接收端依据回波感知电磁环境状态，判断当前频率是否受扰并生成奖励信号，将其与环境特征一并存入经验回放存储器。估值网络利用存储数据计算各候选载频动作的  $Q$  值（表征长期预期回报），系统据此选择最优发射频率。训练过程中，通过计算估值网络与目标网络的输出差异，以反向传播更新估值网络参数，并周期性同步目标网络，逐步提升频率选择的预测精度。同时，采用  $\varepsilon$ -greedy 策略平衡探索与利用：随训练进行逐步降低随机探索概率，增强策略的收敛性与稳定性。该深度强化学习框架有效提升了捷变频雷达在复杂电磁环境下的自适应抗干扰能力。

### 3. 仿真实验设计与结果分析

通过仿真实验的对比分析，本文所提出的方法在多类型干扰环境中展现出载波频率选择的显著抗干扰特性与优越性能。

雷达与深度强化学习参数：经验池大小 1200；训练集大小 48；相参处理周期数 24；雷达工作载频跳变范围 1.486~1.60225 GHz；线性调频带宽 2 MHz；目标网络更新间隔 4；奖励折扣系数 0.96。

$Q$  值深度网络参数：隐含层 1 采用全连接层 64 个神经元；隐含层 2 采用 ReLU 层 64 个神经元；隐含层 3 采用全连接层 64 个神经元。采用 Adam 优化。

随着现代雷达系统所处电磁环境日益复杂，频谱空间中干扰源类型呈现多样化发展趋势，多种干扰模式并存的复

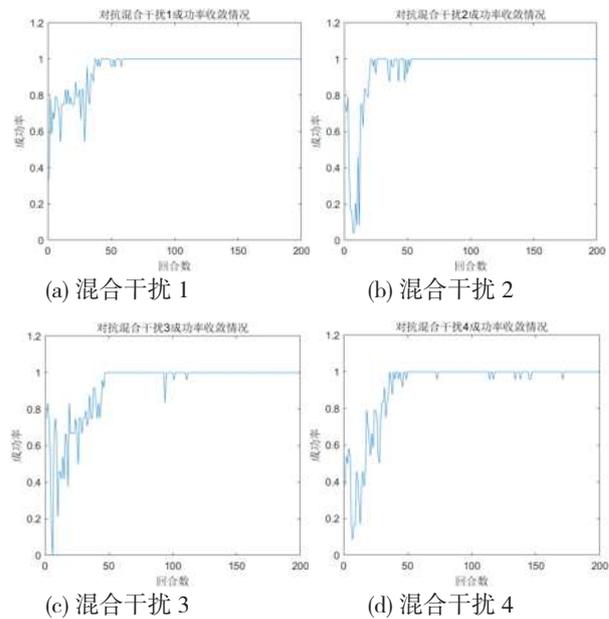
合式干扰场景已成为常态。为全面评估所提方法在复杂电磁对抗条件下的适应能力，本节设计了以下四类典型混合干扰场景进行验证：

混合场景一：兼具稳态阻塞式干扰与动态扫频干扰；

混合场景二：兼具稳态阻塞干扰及相参延迟干扰；

混合场景三：兼具动态扫频干扰与相参延迟干扰；

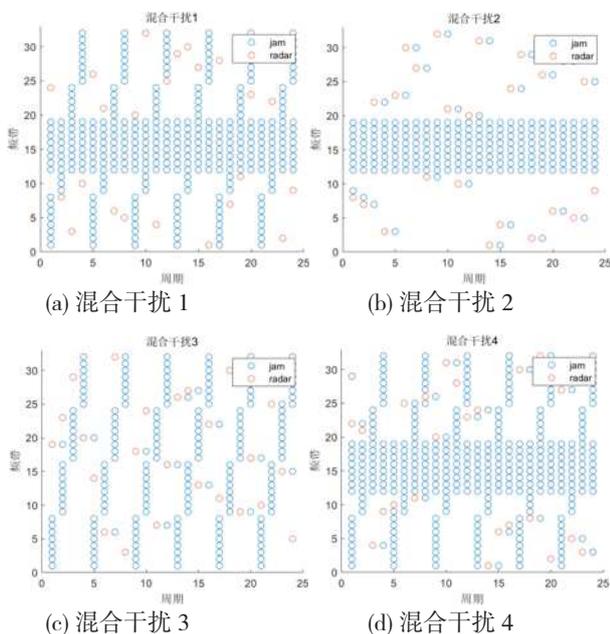
混合场景四：兼具动态稳态阻塞干扰、动态扫频干扰和相参延迟干扰。



各复杂干扰环境的算法收敛情况

根据图 1.1 所示的复杂干扰环境测试结果，本研究提出的深度强化学习载频选择方法经过约 60 次训练迭代后，已基本实现有效的干扰抑制。在复合干扰场景中，虽然算法最终可使抗干扰成功率趋于 100%，但在个别后续回合中仍会出现干扰未能完全消除的现象。这一情况主要源于复杂电磁环境下干扰模式的动态变化，增加了智能体对状态-动作价值函数准确学习的难度，导致  $Q$  网络难以完全逼近理想的价值映射。

进一步分析图 1.2 中系统稳定后的单回合频段使用分布可见，干扰对雷达系统性能的影响仍然较为明显，表明在多重干扰并存条件下，捷变频雷达实现可靠抗干扰的难度进一步增大。尽管如此，从整体性能来看，所提方法在训练收敛后，在绝大多数测试回合中仍能达到 100% 的抗干扰成功率，证明了该策略在复杂电磁环境下的有效性与适用性。



复杂干扰环境训练稳定后的具体决策效果图

#### 4. 总结与展望

本文分析了捷变频雷达采用随机载频策略在抗干扰能力方面存在的不足，指出了开展载频优化以提升干扰抑制效果的必要性。针对存在有源干扰的雷达探测场景，提出了一种基于深度强化学习的自适应频率捷变方法。通过构建捷变频雷达在干扰对抗中的马尔可夫决策过程模型，利用深度 Q 网络在动态变化的频谱干扰条件下开展训练，进而显著提升雷达系统的干扰抑制性能。

需要指出，本研究对干扰信号的建模进行了适当简化，将雷达对抗同频干扰的任务归结为关于载频选择的序列决

策问题。今后的研究应致力于构建更为精确的干扰机行为模型，同时考虑多干扰源协同工作及更先进的干扰形态，如周期性全脉冲转发干扰或全频段压制干扰，从而提升捷变频雷达在未来高度复杂电磁环境下的实战适应能力。

#### 参考文献:

- [1]Miao Z, Yuan X, Zhou F 等。增强层次结构的灰狼优化器及其在无线传感器网络覆盖优化问题中的应用 [J]。应用软计算, 2020, 96:106602。
- [2]Zhong S, Huang X, Wang H 等。基于免疫遗传学的抗间歇采样中继器干扰波形设计 [C]//2019 IEEE 电力、智能计算与系统国际会议 (ICPICS)。中国沈阳, IEEE, 2019:553–559。
- [3]Kumar A, Das S, Mallipeddi R。一种高效的大规模全局优化微分分组算法 [J]。IEEE 进化计算汇刊, 2022 年。557(6928): 513–517。
- [4]Pascarella F, Scarselli C, Novellis G 等。基于遗传算法的同心圆天线阵列设计优化 [C]//2023 IEEE 天线与传播国际研讨会暨 USNC-URSI 无线电科学会议。美国俄勒冈州波特兰, IEEE, 2023:1365–1366。
- [5]Dunn Z, Yeary M, Uysal F 等。通过粒子群优化得到的低旁瓣伪正交码集 [C]//2016 IEEE 雷达会议 (RadarConf)。美国宾夕法尼亚州费城, IEEE, 2016:1–4。
- [6]Mnih V, Kavukcuoglu K, Silver D 等。通过深度强化学习实现人类水平控制 [J]。《自然》, 2015, 518 (7540) : 529–533。