动态电磁环境下多功能雷达一体化发射资源管理方案

张 鹏 严俊坤* 高 畅 李 康 刘宏伟
(西安电子科技大学雷达信号处理全国重点实验室 西安 710071)

摘要：传统多功能雷达仅面向目标特性优化发射资源，在动态电磁环境下面临干扰智能时变、优化模型失配的问题。因此，该文提出一种基于数据驱动的一体化发射资源管理方案，旨在通过对动态干扰信息在线感知与利用提升多功能雷达在动态电磁环境下的多目标跟踪(MTT)性能。该方案首先建立马尔可夫决策过程，数学化描述雷达被敌方截获和干扰的风险。而后将该马尔可夫决策过程感知的干扰信息耦合进MTT精度计算，一体化发射资源管理方法被设计为具有约束动作空间的优化问题。最后提出一种贪婪排序回溯算法对其进行求解。仿真结果表明，所提方法在面向动态干扰环境时不仅可以降低敌方截获概率，还能在被干扰时降低干扰对雷达的影响，改善MTT性能。

关键词：一体化发射资源管理；多目标跟踪；动态电磁环境；马尔可夫决策过程；优化问题
中图分类号：TN957.51 **文献标识码：**A **文章编号：**2095-283X(2025)02-0456-14
DOI: 10.12000/JR24230 **CSTR**: 32380.14.JR24230

引用格式：张鹏, 严俊坤, 高畅, 等. 动态电磁环境下多功能雷达一体化发射资源管理方案[J]. 雷达学报(中英文), 2025, 14(2): 456–469. doi: 10.12000/JR24230.

Reference format: ZHANG Peng, YAN Junkun, GAO Chang, *et al*. Integrated transmission resource management scheme for multifunctional radars in dynamic electromagnetic environments[J]. *Journal of Radars*, 2025, 14(2): 456–469. doi: 10.12000/JR24230.

Integrated Transmission Resource Management Scheme for Multifunctional Radars in Dynamic Electromagnetic Environments

ZHANG Peng YAN Junkun* GAO Chang LI Kang LIU Hongwei
(*National Key Laboratory of Radar Signal Processing*, *Xidian University*, *Xi'an* 710071, *China*)

Abstract: Traditional multifunctional radar systems optimize transmission resources solely based on target characteristics. However, this approach poses challenges in dynamic electromagnetic environments owing to the intelligent time-varying nature of jamming and the mismatch between traditional optimization models and real-world scenarios. To address these limitations, this paper proposes a data-driven integrated transmission resource management scheme designed to enhance the Multiple Target Tracking (MTT) performance of multifunctional radars in complex and dynamic electromagnetic environments. The proposed scheme achieves this by enabling online perception and utilization of dynamic jamming information. The scheme initially establishes a Markov Decision Process (MDP) to mathematically model the risks associated with radar interception and adversarial jamming. This MDP provides a structured approach to perceive jamming

收稿日期：2024-11-19；改回日期：2025-03-10；网络出版：2025-03-27
*通信作者： 严俊坤 jkyan@xidian.edu.cn *Corresponding Author: YANJunkun, jkyan@xidian.edu.cn
基金项目：国家自然科学基金(62471356, 62101350, 62192714)，中国航天科工集团第八研究院产学研合作(SAST2023-068)
Foundation Items: The National Natural Science Foundation of China (62471356, 62101350, 62192714), Industry University-Research Cooperation of the 8th Research Institute of China Aerospace Science and Technology Corporation (SAST2023-068)
责任主编：时晨光 Corresponding Editor: SHI Chenguang

information, which is then integrated into the calculation of MTT. The integrated resource management challenge is formulated as an optimization problem with constraints on the action space. To solve this problem effectively, a greedy sorting backtracking algorithm is introduced. Simulation results demonstrate the efficacy of the proposed method, demonstrating its ability to significantly reduce the probability of radar interception in dynamic jamming environments. Furthermore, the method mitigates the impact of jamming on radar performance during adversarial interference, thereby improving MTT performance.

Key words: Integrated transmit resource management; Multiple Target Tracking (MTT); Dynamic electromagnetic environments; Markov Decision Process (MDP); Optimization problem

1　引言

多目标跟踪(Multiple Target Tracking, MTT)问题一直是各种商业和军事应用领域所关注的重点，其要求雷达可以同时估计多个目标的状态[1–4]。如果雷达对目标照射能量越大，对应的目标跟踪精度越高。然而，雷达的探测资源往往是有限的。为了提升MTT性能，实现资源管理是一种有效的方式。

目前，大多资源管理方案通过预设优化模型来提升有效资源的利用效率[5–9]，其大致可以分为两类。第1种在给定资源下最大化MTT性能。其中，文献[5]利用最差情况框架设计一种联合节点选择和功率分配框架来提升MTT性能。文献[6]以提升目标容量为目标建立一种资源管理模型，它的主要目的是通过协调雷达发射功率提升满足目标跟踪精度的数目。文献[7]针对相控阵雷达建立一种联合跟踪序列和停留时间优化模型，并设计两步去耦方法解决该优化问题以提高MTT性能。第2种则是在给定性能需求下最小化资源的消耗。其中，文献[8]在满足跟踪精度需求的情况下最小化时间资源的消耗。文献[9]提出了一种鲁棒的功率分配算法，该方法通过定义任务效用函数灵活评估不同功率分配策略的目标跟踪精度，解决了发射功率不充足时功率分配方案难求解的问题。

尽管上述资源管理模型可以很好地处理有限资源与MTT问题之间存在的矛盾，但大多数工作尝试在纯净的电磁环境中依据目标特性建立资源管理模型，并未考虑干扰带来的影响。随着电子战技术的发展，目标通常利用自身的截获功能拦截雷达信号并针对性地释放干扰来削弱雷达性能。因此，干扰机可以根据雷达发射信号改变自身的干扰信号，这本质上也形成一种动态的电磁环境[10]。由于雷达缺乏这些目标的截获模式和干扰模式的先验信息，因而在这种动态电磁(干扰)环境下大多预设资源管理模型将失配，其跟踪性能常常下降。因此，文献[11]尝试将截获接收机的工作参数作为先验信息，建立优化模型来在保证目标跟踪性能的同时提高低截获性能。文献[12]则利用已知的干扰信息实现功率分配，从而提高目标跟踪性能。事实上，实际电子战中上述截获和干扰信息难以获取。为了处理这种未知且高动态电磁环境中的资源管理问题，文献[13]在目标未知截获模式下利用强化学习逐渐调整跳频时间间隔和平均发射功率来改善脉冲积累效率。文献[14]则利用类似的框架设计抗干扰策略提高雷达的检测性能。文献[15]调整发射频率和功率在减小干扰影响的同时提升MTT性能。

由上述工作来看，利用强化学习实现低截获或者抗干扰都可以直接或间接减少干扰的影响来提升目标跟踪性能。但在MTT挑战下，目标的截获模式和干扰模式未知且存在差异，因而仅实现低截获或者抗干扰的单一设计很难适应复杂电磁环境下的MTT问题。考虑不同目标仍然会对有限系统资源存在竞争，现有资源管理模型失配会进一步加剧MTT性能的恶化，并且多任务的耦合同样会增加资源管理的求解空间，寻找优化发射方案也存在挑战。

为此，本文构建数据驱动的一体化发射资源管理方案感知和利用干扰信息引导多功能雷达优化发射参数配置以提升动态电磁环境下的MTT性能。该方案首先设计一个主动抗干扰的马尔可夫决策过程(Markov Decision Process, MDP)来描述雷达可能遭遇的截获和干扰风险。然后结合该MDP感知的干扰信息，一体化发射资源管理方案被设计为整数规划问题，旨在提升雷达在最差情况下的MTT性能。最后，设计贪婪排序回溯算法快速求解形成的优化问题。仿真结果表明所生成的发射方案可以有效削弱干扰的影响并改善MTT性能。

2　系统模型

考虑一个位于坐标(x,y)的多功能雷达和Q个点目标在二维空间匀速运动。每一个目标可以通过截获功能获取雷达信息，并引导自身携带的干扰机发射针对性的干扰信号降低其跟踪性能。而雷达可以改变自身发射方案降低干扰的影响。目标的初始位置为(x_0,y_0)，初速度为$(\dot{x}_0,\dot{y}_0)$。在kT时刻，目标

的位置和速度分别表示为(x_k, y_k)和$(\dot{x}_k, \dot{y}_k)$，其中T是跟踪时间间隔。

2.1 目标运动模型

考虑目标正在匀速运动[16]：

$$\boldsymbol{x}_{q,k} = \boldsymbol{F}_q \boldsymbol{x}_{q,k-1} + \boldsymbol{u}_{q,k-1} \tag{1}$$

其中，$\boldsymbol{x}_{q,k} = [x_{q,k}, \dot{x}_{q,k}, y_{q,k}, \dot{y}_{q,k}]^{\mathrm{T}}$为目标状态。$\boldsymbol{F}_q$是状态转移矩阵：

$$\boldsymbol{F}_q = \boldsymbol{I}_2 \otimes \begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix} \tag{2}$$

其中，$\otimes$是克罗内克算子，$\boldsymbol{I}_2$表示2阶单位矩阵。项$\boldsymbol{u}_{q,k-1}$为零均值，协方差为$\boldsymbol{Q}_{q,k-1}$的高斯过程噪声[17]：

$$\boldsymbol{Q}_{q,k-1} = q_0 \boldsymbol{I}_2 \otimes \begin{bmatrix} \frac{1}{3}T^3 & \frac{1}{2}T^2 \\ \frac{1}{2}T^2 & T \end{bmatrix} \tag{3}$$

其中，$q_0 \geq 0$描述目标运动的不确定性。

2.2 目标截获和干扰模型

为了有效获取雷达信息，干扰机通常配备接收机来感知其发射信号[13]。如图1所示，截获接收机频率扫描范围为$\left[B^q_{\mathrm{I,lo}}, B^q_{\mathrm{I,up}}\right]$，但接收机每次只覆盖带宽$B^q_{\mathrm{I,min}}$，并停留$N_{\mathrm{P},q}(N_{\mathrm{P},q}>1)$个脉冲时间以确保截获接收机能够接收到完整的脉冲信息。因此，干扰机需要$N_{\mathrm{P},q}(B^q_{\mathrm{I,up}} - B^q_{\mathrm{I,lo}})/B^q_{\mathrm{I,min}}$脉冲时间来覆盖整个带宽。如果目标在某次扫描中雷达的发射频率被覆盖，接收机就会在功率域进一步检测雷达信号，其检测概率为[13]

$$p^k_{\mathrm{d},q} \approx 0.5 \times \mathrm{erfc}\left(\sqrt{-\ln p_{\mathrm{fa},q}} - \sqrt{D^k_{\mathrm{I},q} + 0.5}\right) \tag{4}$$

其中，$p_{\mathrm{fa},q}$为预设的虚警率，$D^k_{\mathrm{I},q}$为接收机信噪比(Signal to Noise Ratio, SNR)[13]：

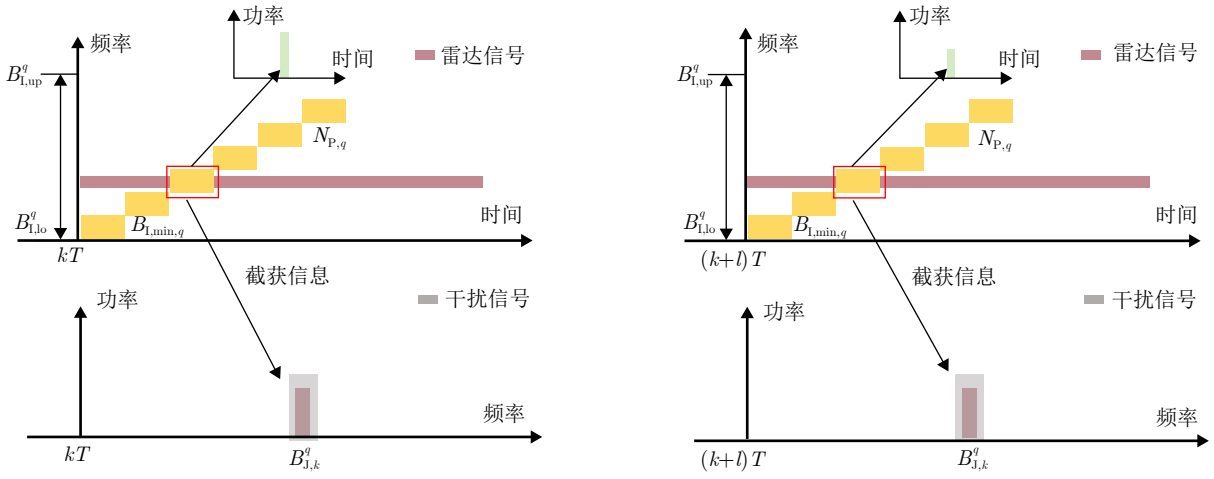
$$D^k_{\mathrm{I},q} = \frac{P^q_{\mathrm{R},k} G^q_{\mathrm{R,I}} \lambda^2_{\mathrm{R},q}}{(4\pi)^2 R^2_{q,k} \eta_{\mathrm{I},q}} \tag{5}$$

其中，$P^q_{\mathrm{R},k}$是雷达的发射功率。$G^q_{\mathrm{R,I}}$是雷达与截获接收机之间的信道增益，其耦合雷达系统的发射增益与截获接收机的接收增益。$\lambda_{\mathrm{R},q}$为雷达信号波长，$R_{q,k}$为目标q与雷达之间的距离，$\eta_{\mathrm{I},q}$为截获接收机的噪声。

在截获雷达信号后，图1表明干扰机通常会针对性地发射干扰信号以降低跟踪性能。考虑干扰机可以采用两种典型干扰模式：阻塞干扰和点频干扰[18]。在点频干扰模式下，干扰机发射窄带信号。对于阻塞干扰模式，干扰机发射宽带信号。由于发射峰值功率有限，阻塞干扰的功率密度通常低于点频干扰信号，对应的干扰效果也被削弱[12]。考虑目标q的干扰功率为$P_{\mathrm{J},q}$，干扰信号的功率密度为

$$P^q_{\mathrm{J,d},k}(B^q_{\mathrm{J},k}) = P_{\mathrm{J},q}/B^q_{\mathrm{J},k} \tag{6}$$

其中，$B^q_{\mathrm{J},k}$为干扰信号的带宽。对于点频干扰模式，$B^q_{\mathrm{J},k}$略大于雷达信号的带宽。对于阻塞干扰，$B^q_{\mathrm{J},k}$取决于干扰机期望覆盖的载频数量。例如，第n个载频$(1 \leq n \leq N)$为$f_n = f_1 + (n-1)\Delta f$，其中$f_1$是雷达可选择的最小载波频率，$\Delta f>0$是两个连续载波之间的步长，$N$表示雷达可用的频率数量。如果干扰机想要干扰载波频率为f_i的雷达，可以使用点频干扰，干扰频率为$B^q_{\mathrm{J},k} \approx \Delta f$。然而，如果



(a) 第k帧
(a) The k frame

(b) 第$k+l$帧($l>0$)
(b) The $k+l$ frame ($l>0$)

图1 不同帧的目标截获与干扰模式

Fig. 1 Target interception and jamming modes in different frames

干扰机想同时干扰f_i和f_j，则需要使用带宽为$B_{\mathrm{J},k}^q = M\Delta f$的阻塞干扰，其中$M = |i-j|+1$表示$f_i$和$f_j$之间的频率数。

此外，为了提升干扰效果，干扰机利用截获的信息和某种规则交替使用点频和阻塞干扰模式，从而形成一种高动态的干扰环境。因而，雷达的跟踪性能将严重下降。

2.3 目标量测模型

在第k帧，目标q的测量如下形式[16]：

$$\boldsymbol{z}_{q,k} = \boldsymbol{h}(\boldsymbol{x}_{q,k}) + \boldsymbol{w}_{q,k} \tag{7}$$

其中，$h(\cdot)$是测量函数：

$$h(\cdot) = \left[h_{R_{q,k}}(\cdot), h_{\theta_{q,k}}(\cdot), h_{v_{q,k}}(\cdot)\right]^{\mathrm{T}} \tag{8}$$

其中

$$\begin{aligned} R_{q,k} &= h_{R_{q,k}}(\boldsymbol{x}_{q,k}) = \sqrt{(x_{q,k}-x)^2 + (y_{q,k}-y)^2} \\ \theta_{q,k} &= h_{\theta_{q,k}}(\boldsymbol{x}_{q,k}) = \arctan\left[(y_{q,k}-y)/(x_{q,k}-x)\right] \\ v_{q,k} &= h_{v_{q,k}}(\boldsymbol{x}_{q,k}) = \frac{\dot{x}_{q,k}(x_{q,k}-x) + \dot{y}_{q,k}(y_{q,k}-y)}{\sqrt{(x_{q,k}-x)^2 + (y_{q,k}-y)^2}} \end{aligned}$$

其中，$\theta_{q,k}$是目标q相对于雷达的角度信息，$v_{q,k}$是对应的多普勒信息。$\boldsymbol{w}_{q,k}$是零均值高斯噪声，其协方差矩阵为$\boldsymbol{\Sigma}$：

$$\boldsymbol{\Sigma} = \mathrm{blkdiag}\left(\sigma_{R_{q,k}}^2, \sigma_{\theta_{q,k}}^2, \sigma_{v_{q,k}}^2\right) \tag{9}$$

其中，$\sigma_{R_{q,k}}^2$，$\sigma_{\theta_{q,k}}^2$和$\sigma_{v_{q,k}}^2$是目标距离、方位和多普勒估计均方误差的克拉美罗下界[19]：

$$\begin{cases} \sigma_{R_{q,k}}^2 \propto \left(D_{q,k}B_{\mathrm{R,r}}^2\right)^{-1} \\ \sigma_{\theta_{q,k}}^2 \propto \left(D_{q,k}/B_{\mathrm{R,NN}}\right)^{-1} \\ \sigma_{v_{q,k}}^2 \propto \left(D_{q,k}T_{\mathrm{R,d}}^2\right)^{-1} \end{cases} \tag{10}$$

其中，$B_{\mathrm{R,r}}$为有效信号带宽，$T_{\mathrm{R,d}}$为驻留时间，$B_{\mathrm{R,NN}}$为 3 dB 接收波束宽度。$D_{q,k}$是目标q在第k帧的信干噪比(Signal to Interference and Noise Ratio, SINR)[14]：

$$D_{q,k} = \frac{N_{\mathrm{R},k}^q P_{\mathrm{R},k}^q G_{\mathrm{R,R}}\sigma_{n,q}\lambda_{\mathrm{R},q}^2}{(4\pi)^3 R_{q,k}^4(\eta_{\mathrm{R}} + P_{\mathrm{J,R},k}^q)} \tag{11}$$

其中，$N_{\mathrm{R},k}^q$是脉冲数，$\sigma_{n,q}$是目标q在载波频率f_n下的雷达散射截面(Radar Cross Section, RCS)，$G_{\mathrm{R,R}}$是雷达回波的信道增益，其包括雷达系统的发射与接收增益。η_{R}是与雷达相关的噪声，$P_{\mathrm{J,R},k}^q$是进入雷达的干扰功率：

$$P_{\mathrm{J,R},k}^q = \frac{P_{\mathrm{J,d},k}^q(B_{\mathrm{J},k}^q)G_{\mathrm{J,R}}^q\lambda_{\mathrm{J},q}^2 B_{\mathrm{R,r}}}{(4\pi)^2 R_{q,k}^2} \tag{12}$$

其中，$G_{\mathrm{J,R}}^q$为干扰机与雷达间的信道增益，其主要包括干扰机的发射增益与雷达系统的接收增益。$\lambda_{\mathrm{J},q}$为干扰信号波长，$P_{\mathrm{J,d},k}^q(B_{\mathrm{J},k}^q)$为与带宽$B_{\mathrm{J},k}^q$相关的干扰功率密度。如果雷达能采取一些对抗措施，迫使接收机引导干扰机增加带宽甚至停止发射干扰信号，进入雷达的干扰功率就会减少，SINR就会增加。值得注意的是SINR可以由发射频率、脉冲数(驻留时间)和功率决定。因此，多功能雷达通过调整这些参数来改变SINR以减少$D_{q,k}$中的测量误差，从而提高目标跟踪性能。

最后，考虑目标RCS起伏服从Swerling I型[20]。对于目标q，其RCS在载波频率f_i处起伏的均值是$\bar{\sigma}_{i,q}$。当相应的波长接近目标尺寸时，目标将处于Mie区域，其中目标在某些方向上的较低频率具有较大的RCS[21]。因此，RCS起伏均值随载频增加而减小同时满足以下关系：$\bar{\sigma}_{1,q}>\bar{\sigma}_{2,q}>\cdots>\bar{\sigma}_{n,q}(f_1>f_2>\cdots>f_n)$。

3 一体化发射资源管理方案

在实际电子战中，目标可以利用截获接收机和干扰机创造未知且动态的干扰环境。对于这种复杂电磁环境下的MTT问题，雷达可以通过调整发射资源来实现低截获以避免干扰，但一味地实现低截获可能会影响目标跟踪性能。此外，雷达有时很难阻止干扰信号进入雷达，抗干扰方法可以减少干扰对雷达的直接影响。然而，现有的抗干扰方法可能会错过完全避免干扰的机会，而有限的系统资源进一步对MTT问题提出挑战。为此，开发一体化资源管理方案不仅可以实现低截获，还可以减少干扰对雷达的影响，同时最大限度提高MTT性能。

3.1 主动抗干扰的MDP

为了解决未知和动态电磁环境中的MTT问题，资源管理方案考虑为一个与发射资源相关的决策问题。作为决策问题的经典形式， 马尔可夫决策过程(MDP)为智能体与环境交互最理想的数学模型。通常，一个MDP由以下4个元素定义[22]。

(1) $\mathcal{S}$ ： 状态$s \in \mathcal{S}$的有限状态集合。

(2) $\mathcal{A}$ ： 动作$a \in \mathcal{A}$的有限动作集合。

(3) $\mathbb{P}$ ： 执行动作a后，从当前状态s到下一状态s'的状态转换概率$\mathbb{P}(s'|s,a)$。

(4) $c(\cdot)$ ： 当智能体在状态s执行动作a对应的成本$c(s,a)$。

通过利用MDP，智能体与环境之间的一般交互描述如下[23]：在每个时间步k，智能体以状态s_k的形式接收环境信息，并根据策略$\pi(\cdot \mid s_k) \in \Delta(\mathcal{A})$

选择动作a_k。执行动作a_k后，状态转换到新状态s_{k+1}，智能体对应的成本为$c_k(s_k,a_k)$。

对于上述MDP，雷达作为智能体将与具备截获和干扰功能的目标所构成的动态电磁环境交互以学习对抗干扰。然而，目标是运动的。通过对比图1(a)与图1(b)，目标距离的变化使得其接收到的雷达发射功率有所差异，这使得雷达可能难以清晰理解当前的干扰环境。为了实现更有效的交互，我们从雷达视角定义主动抗干扰MDP的关键元素。

(1) 动作：第k帧，目标q雷达将采取的动作为

$$\boldsymbol{a}_k^q=\left[t_{d_k^q},f_{c_k^q},p_{e_k^q}\right]\in\mathcal{A}=\{\boldsymbol{a}_1,\boldsymbol{a}_2,\cdots,\boldsymbol{a}_{N_{\mathcal{A}}}\}\tag{13}$$

其中，$N_{\mathcal{A}}$是可用的发射方案数量，每个发射方案$\boldsymbol{a}_k^q$由3个发射变量组成：

$$\left.\begin{aligned}&t_{d_k^q}=t_1+(d_k^q-1)\Delta t,\ \ (d_k^q\in\mathcal{D}=\{1,2,\cdots,N_{\mathcal{D}}\})\\&f_{c_k^q}=f_1+(c_k^q-1)\Delta f,\ \ (c_k^q\in\mathcal{C}=\{1,2,\cdots,N_{\mathcal{C}}\})\\&p_{e_k^q}^q=p_1+(e_k^q-1)\Delta p,\ \ (e_k^q\in\mathcal{E}=\{1,2,\cdots,N_{\mathcal{E}}\})\\&P_{\mathrm{R},k}^q(e_k^q)=p_{e_k^q}^qR_{q,k|k-1}^2/R_{\mathrm{ref},q}^q\end{aligned}\right\}\tag{14}$$

其中，$t_{d_k^q}$表示发射脉冲数(驻留时间)，t_1是可供选择的最小脉冲数，Δt是两个连续脉冲数方案之间的步长，$N_{\mathcal{D}}$则为对应发射脉冲方案的个数。$f_{c_k^q}$表示第c_k^q个载频，f_1为初始载频，Δf是两个连续载频之间的步长，$N_{\mathcal{C}}$为可用载频数。p_1是最小的脉冲发射功率，$p_{e_k^q}^q$为第e_k^q个发射功率，Δp为两个连续发射功率之间的步长，$N_{\mathcal{E}}$为可用脉冲功率数目。特别地，$R_{q,k|k-1}$是雷达在第k帧预测的距离信息，$P_{\mathrm{R},k}^q(e_k^q)$为实际发射功率以保证目标截获概率可以在同一发射功率选择e_k^q下仅与参考距离信息R_{ref}^q相关。因此，R_{ref}^q为常数保证目标的截获概率不随跟踪时间变化，有利于截获模式的感知与学习。

(2) 状态：一旦雷达信号被截获，干扰机将生成针对性的干扰信号。通常，雷达可以通过检测[24]或频谱感知[25]获得对应的干扰信息。因此，由雷达观测到的干扰动作定义为

$$j_{\mathrm{U},k}^q,j_{\mathrm{I},k}^q\in\mathcal{J}=\{0,1,\cdots,N_{\mathcal{J}}\}\tag{15}$$

其中，$j_{\mathrm{U},k}^q$和$j_{\mathrm{I},k}^q$是雷达观测到发射信号被截获前后干扰机采用的干扰动作，由于干扰动作可能发生改变，因此$j_{\mathrm{U},k}^q\neq j_{\mathrm{I},k}^q$，否则$j_{\mathrm{U},k}^q=j_{\mathrm{I},k}^q$。$N_{\mathcal{J}}+1$表示可用的干扰动作数量($N_{\mathcal{J}}=N_{\mathcal{C}}+1$)。关于干扰模式，$j_{\mathrm{U},k}^q=0$或$j_{\mathrm{I},k}^q=0$表示雷达观测到自身未被干扰。$j_{\mathrm{U},k}^q,j_{\mathrm{I},k}^q\in\mathcal{J}/\{0,N_{\mathcal{J}}\}$表示雷达观测到自身受到点频干扰，$N_{\mathcal{J}}$表示雷达观测到自身受到阻塞干扰。此外，干扰机可以利用截获模式提供的雷达历史动作来确定干扰动作，其与雷达之间相互作用可以描述为一个交替序列[14]：

$$\boldsymbol{Y}_k^q=\boldsymbol{a}_1^q,j_{\mathrm{U},1}^q,j_{\mathrm{I},1}^q,\boldsymbol{a}_2^q,j_{\mathrm{U},2}^q,j_{\mathrm{I},2}^q,\cdots,\boldsymbol{a}_k^q,j_{\mathrm{U},k}^q,j_{\mathrm{I},k}^q,\cdots\tag{16}$$

其中，$\boldsymbol{Y}_k^q$是雷达与目标q在第k帧观测到的交互历史。由于序列$\boldsymbol{Y}_k^q$的大小随时间增长。因此，交互信息$\boldsymbol{Y}_k^q$不能直接用作状态。因此，我们利用以下状态来近似交互历史：

$$\begin{aligned}\boldsymbol{s}_k^q&=(\boldsymbol{a}_{k-n}^q,j_{\mathrm{U},k-n}^q,j_{\mathrm{I},k-n}^q,\cdots,\boldsymbol{a}_{k-1}^q,j_{\mathrm{U},k-1}^q,j_{\mathrm{I},k-1}^q)\in\mathcal{S}\\&=\{s_1,s_2,\cdots,s_{N_{\mathcal{S}}}\}\end{aligned}\tag{17}$$

其中，$\boldsymbol{s}_k^q$包括最近n个观测历史，$N_{\mathcal{S}}=(N_{\mathcal{A}}N_{\mathcal{J}}^2)^n$表示状态的数目。

(3) 状态转移函数：通常，雷达最初对干扰环境没有先验知识。因此，雷达首先需要通过与环境的交互来学习潜在的环境动态[26]。作为参数估计问题，利用最大似然法来估计该转移概率[27]：

$$\mathbb{P}_q(\boldsymbol{s}_{k+1}^q|\boldsymbol{s}_k^q,\boldsymbol{a}_k^q)=\frac{N_q(\boldsymbol{s}_k^q,\boldsymbol{a}_k^q,\boldsymbol{s}_{k+1}^q)}{N_q(\boldsymbol{s}_k^q,\boldsymbol{a}_k^q)}\tag{18}$$

其中，$N_q(\boldsymbol{s}_k^q,\boldsymbol{a}_k^q)=\sum_{\boldsymbol{s}_{k+1}^q\in\mathcal{S}}N_q(\boldsymbol{s}_k^q,\boldsymbol{a}_k^q,\boldsymbol{s}_{k+1}^q)$，并且$N_q(\boldsymbol{s}_k^q,\boldsymbol{a}_k^q,\boldsymbol{s}_{k+1}^q)$表示雷达采取行动$\boldsymbol{a}_k^q$时，其状态由$\boldsymbol{s}_k^q$变为$\boldsymbol{s}_{k+1}^q$的次数。若$N_q(\boldsymbol{s}_k^q,\boldsymbol{a}_k^q)=0$，则$\mathbb{P}_q(\boldsymbol{s}_{k+1}^q|\boldsymbol{s}_k^q,\boldsymbol{a}_k^q)=1/|\mathcal{S}|$。

(4) 成本：在MDP中，一个成本函数可以量化在给定状态和动作下智能体消耗的成本。对于雷达，SINR通常可以有效反映干扰对雷达的影响[12]。然而，SINR并不能准确反映跟踪精度的变化，而预测的BCRLB (Bayesian Cramer-Rao Lower Bound)可以表征目标跟踪精度。为此，首先给出预测贝叶斯信息矩阵[28]：

$$\begin{aligned}\boldsymbol{J}(\xi_{q,k})=&\left[\boldsymbol{Q}_{q,k-1}+\boldsymbol{F}_q\boldsymbol{J}^{-1}(\xi_{q,k-1})(\boldsymbol{F}_q)^{\mathrm{T}}\right]^{-1}\\&+\boldsymbol{H}_{q,k}^{\mathrm{T}}\boldsymbol{\Sigma}_{q,k}^{-1}\boldsymbol{H}_{q,k}\Big|_{\xi_{q,k|k}}\end{aligned}\tag{19}$$

其中，$\boldsymbol{H}_{q,k}$表示雅可比矩阵，而协方差矩阵$\boldsymbol{\Sigma}_{q,k}$与SINR相关：

$$\begin{aligned}&D_{q,k}(\boldsymbol{s}_k^q,\boldsymbol{a}_k^q,\boldsymbol{s}_{k+1}^q)\\&=\frac{t_{d_k^q}p_{e_k^q}^qG_{\mathrm{R,R}}\sigma_{c_k^q,q}\lambda_{\mathrm{R},q}^2R_{q,k|k-1}^2}{(4\pi)^3R_{q,k}^4R_{\mathrm{ref},q}^2(\eta_{\mathrm{R}}+P_{\mathrm{J,R},k}^q(\boldsymbol{s}_k^q,\boldsymbol{a}_k^q,\boldsymbol{s}_{k+1}^q))}\end{aligned}\tag{20}$$

其中，$P_{\mathrm{J,R},k}^q(\boldsymbol{s}_k^q,\boldsymbol{a}_k^q,\boldsymbol{s}_{k+1}^q)$是雷达接收到的干扰功率。事实上，目标运动SINR和RCS起伏使得雷达难以直接量化$D_{q,k}$。因此，我们从回波中估计出一个稳定的成本函数。根据引起SINR动态变化不同因素，$D_{q,k}$可以被分解为3个变量：

$$D_{q,k}(\boldsymbol{s}_k^q,\boldsymbol{a}_k^q,\boldsymbol{s}_{k+1}^q)=\hat{D}_{q,k}(\boldsymbol{s}_k^q,\boldsymbol{a}_k^q,\boldsymbol{s}_{k+1}^q)\sigma_{c_k^q,q}U(\boldsymbol{s}_{k+1}^q)\tag{21}$$

其中，$\sigma_{c_k^q,q}$为载频$f_{c_k^q}$处的RCS，$U(\boldsymbol{s}_{k+1}^q)$是由状态决定且与目标位置相关的指示函数：

$$U(\boldsymbol{s}_{k+1}^q)=\begin{cases}1/R_{q,k|k}^2, & j_{\mathrm{U},k}^q=0\text{且}j_{\mathrm{I},k}^q=0\\ 1, & \text{其他}\end{cases}\tag{22}$$

其中，$U(\boldsymbol{s}_{k+1}^q)$保证$\hat{D}_{q,k}(\boldsymbol{s}_k^q,\boldsymbol{a}_k^q,\boldsymbol{s}_{k+1}^q)$仅由固定的参考距离$R_{\mathrm{ref},q}$决定，使干扰动态性与目标位置变化实现去耦，更易于评估干扰对雷达的影响以及计算后续不同目标位置下的成本函数。具体而言，当雷达没有被干扰时，$j_{\mathrm{U},k}^q=0$且$j_{\mathrm{I},k}^q=0$，$U(\boldsymbol{s}_{k+1}^q)=1/R_{q,k|k}^2$：

$$\begin{aligned}&\hat{D}_{q,k}(\boldsymbol{s}_k^q,\boldsymbol{a}_k^q,\boldsymbol{s}_{k+1}^q)\\&=\frac{D_{q,k}(\boldsymbol{s}_k^q,\boldsymbol{a}_k^q,\boldsymbol{s}_{k+1}^q)}{\sigma_{c_k^q,q}U(\boldsymbol{s}_{k+1}^q)}\\&=\frac{t_{d_k^q}p_{e_k^q}^qG_{\mathrm{R,R}}\sigma_{c_k^q,q}\lambda_{\mathrm{R},q}^2R_{q,k|k-1}^2R_{q,k|k}^2}{(4\pi)^3R_{q,k}^4R_{\mathrm{ref},q}^2\eta_{\mathrm{R}}\sigma_{c_k^q,q}}\\&=\frac{t_{d_k^q}p_{e_k^q}^qG_{\mathrm{R,R}}\sigma_{c_k^q,q}\lambda_{\mathrm{R},q}^2}{(4\pi)^3R_{\mathrm{ref},q}^2\eta_{\mathrm{R}}\sigma_{c_k^q,q}}\frac{R_{q,k|k-1}^2R_{q,k|k}^2}{R_{q,k}^4}\\&\approx\frac{t_{d_k^q}p_{e_k^q}^qG_{\mathrm{R,R}}\sigma_{c_k^q,q}\lambda_{\mathrm{R},q}^2}{(4\pi)^3R_{\mathrm{ref},q}^2\eta_{\mathrm{R}}\sigma_{c_k^q,q}}\end{aligned}\tag{23}$$

其中，$R_{q,k|k}$是估计的距离信息，$\hat{D}_{q,k}(\boldsymbol{s}_k^q,\boldsymbol{a}_k^q,\boldsymbol{s}_{k+1}^q)$与雷达参数相关。当雷达被干扰时，$j_{\mathrm{U},k}^q\neq0$或$j_{\mathrm{I},k}^q\neq0$，$U(\boldsymbol{s}_{k+1}^q)=1$：

$$\begin{aligned}&\hat{D}_{q,k}(\boldsymbol{s}_k^q,\boldsymbol{a}_k^q,\boldsymbol{s}_{k+1}^q)\\&=\frac{D_{q,k}(\boldsymbol{s}_k^q,\boldsymbol{a}_k^q,\boldsymbol{s}_{k+1}^q)}{\sigma_{c_k^q,q}U(\boldsymbol{s}_{k+1}^q)}\\&=\frac{t_{d_k^q}p_{e_k^q}^qG_{\mathrm{R,R}}\sigma_{c_k^q,q}\lambda_{\mathrm{R},q}^2R_{q,k|k-1}^2}{(4\pi)^3R_{q,k}^4R_{\mathrm{ref},q}^2P_{\mathrm{J,d},k}^q(\boldsymbol{s}_k^q,\boldsymbol{a}_k^q,\boldsymbol{s}_{k+1}^q)\sigma_{c_k^q,q}}\\&=\frac{t_{d_k^q}p_{e_k^q}^qG_{\mathrm{R,R}}\lambda_{\mathrm{R},q}^2(4\pi)^2}{(4\pi)^3R_{\mathrm{ref},q}^2P_{\mathrm{J,d},k}^q(\boldsymbol{s}_k^q,\boldsymbol{a}_k^q,\boldsymbol{s}_{k+1}^q)G_{\mathrm{J,R}}^q\lambda_{\mathrm{J},q}^2B_{\mathrm{R,r}}\sigma_{c_k^q,q}}\\&\quad\cdot\frac{R_{q,k|k-1}^2R_{q,k}^2}{R_{q,k}^4}\\&\approx\frac{t_{d_k^q}p_{e_k^q}^qG_{\mathrm{R,R}}\lambda_{\mathrm{R},q}^2}{4\pi R_{\mathrm{ref},q}^2P_{\mathrm{J,d},k}^q(\boldsymbol{s}_k^q,\boldsymbol{a}_k^q,\boldsymbol{s}_{k+1}^q)G_{\mathrm{J,R}}^q\lambda_{\mathrm{J},q}^2B_{\mathrm{R,r}}}\end{aligned}\tag{24}$$

其中，$\hat{D}_{q,k}(\boldsymbol{s}_k^q,\boldsymbol{a}_k^q,\boldsymbol{s}_{k+1}^q)$与干扰功率密度$P_{\mathrm{J,d},k}^q(\boldsymbol{s}_k^q,\boldsymbol{a}_k^q,\boldsymbol{s}_{k+1}^q)$相关。从式(23)和式(24)可以看出$\hat{D}_{q,k}(\boldsymbol{s}_k^q,\boldsymbol{a}_k^q,\boldsymbol{s}_{k+1}^q)$主要由干扰特性决定，而$\sigma_{c_k^q,q}$与目标自身的起伏特性相关，$U(\boldsymbol{s}_{k+1}^q)$代表目标位置信息的变化。因此，相关变量可以被当作相互独立的随机变量，在线感知可以进一步转换为3个平行的估计过程：

$$\begin{aligned}&\bar{D}_{q,k}(\boldsymbol{s}_k^q,\boldsymbol{a}_k^q,\boldsymbol{s}_{k+1}^q)\\&=\mathbb{E}(D_{q,k}\mid\boldsymbol{s}_k^q,\boldsymbol{a}_k^q,\boldsymbol{s}_{k+1}^q)\\&\approx\mathbb{E}(\hat{D}_{q,k}\mid\boldsymbol{s}_k^q,\boldsymbol{a}_k^q,\boldsymbol{s}_{k+1}^q)\\&\quad\cdot\mathbb{E}(\sigma_{c_k^q,q}\mid\boldsymbol{a}_k^q)U(\boldsymbol{s}_{k+1}^q)\end{aligned}\tag{25}$$

其中，$\mathbb{E}(\sigma_{c_k^q,q}\mid\cdot)$为载频$f_{c_k^q}$处RCS起伏的均值$\bar{\sigma}_{c_k^q,q}$，$\mathbb{E}(\hat{D}_{q,k}\mid\cdot)$用于估计干扰对雷达的影响：

$$\mathbb{E}(\hat{D}_{q,k}\mid\cdot)=\frac{\sum_{n=1}^{N(\boldsymbol{s}_k^q,\boldsymbol{a}_k^q,\boldsymbol{s}_{k+1}^q)}\hat{D}_{q,k}^n(\boldsymbol{s}_k^q,\boldsymbol{a}_k^q,\boldsymbol{s}_{k+1}^q)}{N(\boldsymbol{s}_k^q,\boldsymbol{a}_k^q,\boldsymbol{s}_{k+1}^q)}\tag{26}$$

通过将上述估计的$\overline{D}_{q,k}(\cdot)$代入式(19)，预测贝叶斯信息矩阵将是$(\boldsymbol{s}_k^q,\boldsymbol{a}_k^q,\boldsymbol{s}_{k+1}^q)$的函数，对应的BCRLB是贝叶斯信息矩阵$\boldsymbol{J}(\boldsymbol{s}_k^q,\boldsymbol{a}_k^q,\boldsymbol{s}_{k+1}^q)$的逆[16]：

$$\boldsymbol{B}_q(\boldsymbol{s}_k^q,\boldsymbol{a}_k^q,\boldsymbol{s}_{k+1}^q)=(\boldsymbol{J}(\boldsymbol{s}_k^q,\boldsymbol{a}_k^q,\boldsymbol{s}_{k+1}^q))^{-1}\tag{27}$$

其目标跟踪精度与对角线元素相关[11]：

$$G_q(\boldsymbol{s}_k^q,\boldsymbol{a}_k^q,\boldsymbol{s}_{k+1}^q)=\sqrt{\boldsymbol{B}_q(1,1)+\boldsymbol{B}_q(3,3)}\tag{28}$$

其中，$\boldsymbol{B}_q(1,1)$和$\boldsymbol{B}_q(3,3)$表示矩阵$B_q(\cdot)$的第1个和第3个对角元素，分别表示预测目标位置在x轴和y轴方向上的误差下界。根据跟踪性能指标，最终的成本函数定义如下：

$$c(\boldsymbol{s}_k^q,\boldsymbol{a}_k^q)=\sum_{\boldsymbol{s}_{k+1}^q\in\mathcal{S}}\mathbb{P}_q(\boldsymbol{s}_{k+1}^q|\boldsymbol{s}_k^q,\boldsymbol{a}_k^q)G_q(\boldsymbol{s}_k^q,\boldsymbol{a}_k^q,\boldsymbol{s}_{k+1}^q)\tag{29}$$

简而言之，所构建的主动抗干扰MDP使雷达能够捕获未知且动态的干扰环境信息。在解决模型估计问题后，如何利用这些感知信息来为MTT问题选择合适的方案仍不清楚。下面将介绍一体化发射资源管理方案。

3.2　一体化发射资源管理方案的最优化形式

给定一个MDP，成本函数$c(\boldsymbol{s}_k^q,\boldsymbol{a}_k^q)$隐含干扰的影响，可以用于评估动态干扰环境下雷达对目标q的跟踪性能。对于MTT，雷达的目的则是最大化MTT性能。

通过利用该主动抗干扰的MDP，一体化发射资源管理问题本质是一个针对多目标的决策问题。首先，假设所有目标处于联合状态$\boldsymbol{s}_k=[\boldsymbol{s}_k^1,\boldsymbol{s}_k^2,\cdots,\boldsymbol{s}_k^q,\cdots,\boldsymbol{s}_k^Q]$，则雷达选择联合发射方案$\boldsymbol{a}_k=[\boldsymbol{a}_k^1,\boldsymbol{a}_k^2,\cdots,\boldsymbol{a}_k^q,\cdots,\boldsymbol{a}_k^Q]$。最差目标情况下的目标跟踪精度可以评估MTT性能[13]：

$$\boldsymbol{c}(\boldsymbol{s}_k,\boldsymbol{a}_k)=\max_q\{c(\boldsymbol{s}_k^q,\boldsymbol{a}_k^q)\}\tag{30}$$

进一步考虑雷达系统有限的系统资源，通过利用上述MTT评估准则，一体化发射资源管理方案被建立为一个具有动作空间约束的整数优化问题：

$$\begin{aligned}&\min_{\boldsymbol{a}_k}\boldsymbol{c}(\boldsymbol{s}_k,\boldsymbol{a}_k)\\&\text{s.t. }\boldsymbol{a}_k\in\boldsymbol{\Omega}\end{aligned}\tag{31}$$

其中

$$\boldsymbol{\Omega}=\left\{\boldsymbol{a}_k\left|\sum_{q=1}^Qt_{q,k}=t_{\mathrm{total}},\ t_{\min}\le t_{q,k}\le t_{\max}\right.\right\}$$

表明MTT的总时间资源受限于t_{total}，$t_{\min}$和$t_{\max}$分别表示驻留时间资源的上界和下界。

3.3 贪婪排序回溯算法

对于所提优化问题，目标函数与状态$\boldsymbol{s}_k$和动作$\boldsymbol{a}_k$相关，其计算复杂度同时取决于状态空间与动作空间的维度。通常，状态$\boldsymbol{s}_k$包括多个观测历史。当观测历史增加时，状态空间的尺寸呈指数增长，对应的状态转移函数以及成本函数的维度也急剧增长，从而导致优化问题的复杂度显著上升。然而，所提优化问题核心目标在于最大化即时的MTT性能。因此，在每一帧状态$\boldsymbol{s}_k$已被给定，相应的状态转移函数和成本函数的维度得以大幅缩减。因此，该优化问题的复杂度主要取决于动作空间$\boldsymbol{\Omega}$。此外，由于动作空间$\boldsymbol{\Omega}$由离散的发射方案构成，一体化发射资源管理本质是一个整数优化问题。

实际上，该整数优化问题很难找到对应的松弛问题。因此，一些借助松弛原始优化问题的整数优化方法很难直接用于求解该问题。相比之下，穷举搜索是一种有效方法。而分支定界方法可以通过分割整个动作空间来减少优化时间。然而，该方法仍然需要对整个动作空间进行一定的搜索。为了更快生成发射方案，提出一种贪婪排序回溯方法来求解该优化问题。从数学上讲，该方法首先根据成本函数对每个目标的候选动作进行排序。然后，雷达将从具有最小成本函数的动作开始寻找联合发射方案，确保每一次迭代的成本函数增长最小，最终收获一个可以满足有限资源约束的发射方案，并且考虑算法生成的发射方案可能跳过部分更优发射方案，进一步引入回溯技巧提升搜索效率。具体算法流程如算法1所示。

此外，**引理1**和**引理2**从理论上保证该搜索算法可以更快地协助雷达找到更好的发射方案。

引理1：对于$i<j$，贪婪排序回溯方法使目标函数值递增：$\boldsymbol{c}(\boldsymbol{s}_k,\boldsymbol{a}_i)<\boldsymbol{c}(\boldsymbol{s}_k,\boldsymbol{a}_j)$。

证明：在排序搜索过程中，假设第i次迭代形成的联合发射方案为$\boldsymbol{a}_i=\left[\mathcal{A}_1(\mathbf{IX}_1(n_1^i)),\mathcal{A}_2(\mathbf{IX}_2(n_2^i)),\cdots,\mathcal{A}_Q(\mathbf{IX}_Q(n_Q^i))\right]$，其对应的目标函数为

$$\boldsymbol{c}(\boldsymbol{s}_k,\boldsymbol{a}_i)=\max\boldsymbol{c}_i=\boldsymbol{C}_{q_1}(\mathbf{IX}_{q_1}(n_{q_1}^i))$$

其中，q_1为方案$\boldsymbol{a}_i$下跟踪精度最差的目标指标。在第j次迭代中，发射方案为$\boldsymbol{a}_j=\left[\mathcal{A}_1(\mathbf{IX}_1(n_1^j)),\mathcal{A}_2(\mathbf{IX}_2(n_2^j)),\cdots,\mathcal{A}_Q(\mathbf{IX}_Q(n_Q^j))\right]$，对应的目标函数为

$$\boldsymbol{c}(\boldsymbol{s}_k,\boldsymbol{a}_j)=\max\boldsymbol{c}_j=\boldsymbol{C}_{q_2}(\mathbf{IX}_{q_2}(n_{q_2}^j))$$

其中，q_2为方案$\boldsymbol{a}_j$下跟踪精度最差的目标指标。对于整个搜索过程，若$i<j$，则$n_{q_1}^i\le n_{q_1}^j$，目标函数满足如下关系：

$$\boldsymbol{c}(\boldsymbol{s}_k,\boldsymbol{a}_i)=\boldsymbol{C}_{q_1}(\mathbf{IX}_{q_1}(n_{q_1}^i))\le\boldsymbol{C}_{q_1}(\mathbf{IX}_{q_1}(n_{q_1}^j))\le\boldsymbol{C}_{q_2}(\mathbf{IX}_{q_2}(n_{q_2}^j))=\boldsymbol{c}(\boldsymbol{s}_k,\boldsymbol{a}_j)$$

算法 1 贪婪排序回溯算法流程

Alg. 1 The flow of greedy sort backtracking algorithm

步骤1 输入在第k帧状态$\boldsymbol{s}_k$、动作空间$\mathcal{A}$。初始化$\mathcal{D}_q=\varnothing$、迭代次数$j=1$以及动作索引$n_q=1,\forall q$。

步骤2 评估每一个目标与动作空间$\mathcal{A}$相关的成本函数：

$$\boldsymbol{C}_q=\left[c(\boldsymbol{s}_k^q,\boldsymbol{a}_1),c(\boldsymbol{s}_k^q,\boldsymbol{a}_2),\cdots,c(\boldsymbol{s}_k^q,\boldsymbol{a}_{N_{\mathcal{A}}})\right],\forall q$$

步骤3 将成本函数$\boldsymbol{C}_q$按升序排序，形成索引$\mathbf{IX}_q$：

$$\boldsymbol{C}_q(\mathbf{IX}_q(1))<\boldsymbol{C}_q(\mathbf{IX}_q(2))<\cdots<\boldsymbol{C}_q(\mathbf{IX}_q(N_{\mathcal{A}})),\forall q$$

步骤4 当至少存在一个目标的动作索引$n_q<N_{\mathcal{A}}$，进入步骤5。

步骤5 形成一个联合发射方案并提取其中的驻留时间信息，进入步骤6。

$$\boldsymbol{a}_j=\left[\mathcal{A}_1(\mathbf{IX}_1(n_1)),\mathcal{A}_2(\mathbf{IX}_2(n_2)),\cdots,\mathcal{A}_Q(\mathbf{IX}_Q(n_Q))\right]\Rightarrow\boldsymbol{t}_j=\left[t_1,t_2,\cdots,t_Q\right]$$

步骤6 计算联合发射方案$\boldsymbol{a}_j$对应的成本函数$\boldsymbol{c}_j=[c_j(1),c_j(2),\cdots,c_j(Q)]$：

$$c_j(q)=\begin{cases}\boldsymbol{C}_q(\mathbf{IX}_q(n_q)), & t_q>t_{\min}\text{ 或者}n_q<N_{\mathcal{A}}\\ +\infty, & \text{其他}\end{cases}$$

步骤7 如果$\sum_{q=1}^{Q}t_q\le t_{\text{total}}$，进入到步骤9；否则进入到步骤8。

步骤8 获得$\boldsymbol{c}_j$中具有最小成本函数的目标索引$\text{IX}_{\min}$，将对应的驻留时间方案$t_{\text{IX}_{\min}}$存储进$\mathcal{D}_{\text{IX}_{\min}}$(目标$\text{IX}_{\min}$已遍历过的动作，$\left|\mathcal{D}_{\text{IX}_{\min}}\right|=n_{\text{IX}_{\min}}$)。然后执行$j=j+1$，$n_{\text{IX}_{\min}}=n_{\text{IX}_{\min}}+1$，进入到步骤4。

步骤9 在$\mathcal{D}_q$中回溯寻找具有相同驻留时间参数($t_q\in\mathcal{D}_q$)的最小索引$\text{IX}_{\mathcal{D}_q}$，同时更新动作索引：

$$n_q^*=\begin{cases}\text{IX}_{\mathcal{D}_q}, & \text{IX}_{\mathcal{D}_q}\ne\varnothing\\ n_q, & \text{其他}\end{cases}$$

然后形成最优发射方案$\boldsymbol{a}_k^*=\left[\mathcal{A}_1(\mathbf{IX}_1(n_1^*)),\mathcal{A}_2(\mathbf{IX}_2(n_2^*)),\cdots,\mathcal{A}_Q(\mathbf{IX}_Q(n_Q^*))\right]$，进入步骤10。

步骤10 输出最终的发射方案$\boldsymbol{a}_k^*$。

因此，目标的函数值是递增的。　　证毕

引理2：当为目标分配了最小驻留时间资源$t_{\min}$时，停止搜索其剩余行动可以提高贪婪排序搜索回溯算法的效率。

证明：假设第i次迭代的发射方案$\boldsymbol{a}_i=[\mathcal{A}_1(\mathbf{IX}_1(n_1^i)),\mathcal{A}_2(\mathbf{IX}_2(n_2^i)),\cdots,\mathcal{A}_Q(\mathbf{IX}_Q(n_Q^i))]$不满足约束条件，且目标$q$获得最少时间资源($t_q^i=t_{\min}$)。雷达将继续进行搜索过程，并在第$j$次形成满足约束的发射方案$\boldsymbol{a}_j=[\mathcal{A}_1(\mathbf{IX}_{\boldsymbol{C}_1}(n_1^j)),\mathcal{A}_2(\mathbf{IX}_{\boldsymbol{C}_2}(n_2^j)),\cdots,\mathcal{A}_Q(\mathbf{IX}_{\boldsymbol{C}_Q}(n_Q^j))]$ ($i<j$)。同时，混合方案$\boldsymbol{a}_h$也满足约束：$\boldsymbol{a}_h=[\mathcal{A}_1(\mathbf{IX}_1(n_1^j)),\mathcal{A}_2(\mathbf{IX}_2(n_2^j)),\cdots,\mathcal{A}_q(\mathbf{IX}_q(n_q^i)),\cdots,\mathcal{A}_Q(\mathbf{IX}_Q(n_Q^j))]$。　　证毕

基于搜索规则，我们可以得到$n_q^i\le n_q^j$。因此，贪婪排序搜索法可以在更短的时间内找到$\boldsymbol{a}_h$，即$h\le j$。根据**引理1**，$\boldsymbol{c}(\boldsymbol{s}_k,\boldsymbol{a}_h)\le\boldsymbol{c}(\boldsymbol{s}_k,\boldsymbol{a}_j)$。因此当$t_q^i=t_{\min}$时，至少有一个选项$\boldsymbol{a}_h$比$\boldsymbol{a}_j$更好，对剩余动作的探索不仅增加了目标函数值，而且延长了搜索时间。因此，当为某个目标分配了最小驻留时间资源$t_{\min}$时，停止搜索其剩余行动可以提高贪婪排序搜索算法的效率。

3.4 探索与利用的权衡

雷达与环境的交互通常是在线的。在该过程中，雷达不仅需要通过与环境的交互来收集新的交互经验，还需要利用获取的知识形成一体化发射资源管理方案来提升MTT性能。因此，雷达需要面临在线过程中探索和利用的平衡问题[23]。假设雷达在第k帧处于状态$\boldsymbol{s}_k$。

(1) $\boldsymbol{s}_k$为第1次观测到的状态：至少存在一个目标的干扰环境状态$\boldsymbol{s}_k^q$是被雷达首次观测到的。在新的干扰状态下，其状态转移是未知的，即$\mathbb{P}(\boldsymbol{s}_{k+1}^q|\boldsymbol{s}_k^q,\boldsymbol{a}_k^q)=1/|\mathcal{S}|,\forall\boldsymbol{a}_k^q\in\mathcal{A},\exists q$。因此，雷达需要选择随机发射方案$\boldsymbol{a}_k^{\mathrm{rnd}}\in\boldsymbol{\Omega}$来探索当前动态的电磁环境。

(2) $\boldsymbol{s}_k$为已访问过的状态：每一个目标的干扰环境状态$\boldsymbol{s}_k^q$在过去的交互过程中至少被雷达观测到一次，即$\exists\boldsymbol{a}_k^q\in\mathcal{A},\mathbb{P}(\boldsymbol{s}_{k+1}^q|\boldsymbol{s}_k^q,\boldsymbol{a}_k^q)\ne1/|\mathcal{S}|,\forall q$。此时，雷达需要考虑最大化收益还是继续通过选择随机发射方案来探索环境。因此，尝试引入ε-贪婪策略[29]。根据贪婪方法，雷达将以概率ε采用随机方案$\boldsymbol{a}_k^{\mathrm{rnd}}\in\boldsymbol{\Omega}$来探索环境，并以概率$1-\varepsilon$采取与一体化发射资源管理方案相关的贪婪方案$\boldsymbol{a}_k^*\in\boldsymbol{\Omega}$来改善MTT性能：

$$\boldsymbol{a}_k=\begin{cases}\boldsymbol{a}_k^{\mathrm{rnd}}, & \varepsilon\\ \boldsymbol{a}_k^*, & 1-\varepsilon\end{cases}\tag{32}$$

其中，ε是贪婪率。如果ε设置为0，则雷达将始终选择最优方案。如果ε为1，则雷达将一直探索环境。

事实上，雷达起初并不具备任何环境信息，贪婪率的设置需保证雷达收集环境信息。随着交互时间提升，雷达逐步收集到充足的干扰信息。此时，贪婪率的设置需提高感知信息的利用率改进MTT性能。因此，贪婪率ε将随时间变化。具体而言，ε初始化为1，并随着跟踪时间的增加逐渐减小到0来确保雷达在跟踪过程开始时主动探索环境以获得足够的经验。随着ε的减小，雷达将逐渐增加一体化方案的利用率以提高MTT性能。此外，由于所提方法为在线过程，当发现环境信息不充足导致多功能雷达不能选择合适的发射方案时，适当提升贪婪率可保证雷达收集更多的交互经验以补充不充足的环境信息。

4 仿真实验及结果分析

4.1 仿真参数设计

为了证明所提出算法的有效性，考虑多功能雷达位于坐标$(0,0)$ km处，其与驻留时间相关的方案个数为$N_{\mathcal{D}}=4$，最小脉冲数为$t_1=32$，步长Δd为32，有效总脉冲数为192。可采用的发射频率的数量为$N_{\mathcal{C}}=3$，初始载频被设置为$f_1=1.2$ GHz。两个连续载频之间的步长Δf设置为400 MHz以降低目标回波之间的相关性，同时减小阻塞干扰的功率密度。有效脉冲功率的数量为$N_{\mathcal{E}}=2$，初始发射功率为$p_1=100$ W，步长Δp为1.9 kW。此外，每个目标的参考距离信息R_{ref}^q由初始化的距离信息$R_{q,1|1}$决定，其他与雷达相关的参数可以在表1中找到。

整个雷达的探测区域共有$Q=3$个目标，并且每个目标的RCS随发射频率而变化，表2显示目标

表 1　雷达参数

Tab. 1　Radar parameters

参数	设定值
$G_{\mathrm{R,R}}$	80 dB
$B_{\mathrm{R,r}}$	1 MHz
η_{R}	–141 dBW
$B_{\mathrm{R,r}}$	0.5°

表 2　目标参数

Tab. 2　Target parameters

目标索引	位置(km)	速度(m/s)	$\bar{\sigma}_q$(m^2)
1	(–8, –10)	(10, 20)	[9, 6, 3]
2	(0, 25)	(–25, 10)	[7, 4, 2]
3	(15, 12)	(–25, 15)	[8, 5, 2]

的初始信息以及不同发射频率下RCS起伏的均值。所有目标相对于雷达的位置如图2所示。考虑在动态电磁环境中每个目标都具备不同的截获模式：

(1) 强截获模式：截获接收机具有快速扫频速率和强信号检测能力，可以在频域和功率域上成功拦截雷达信号。

(2) 弱截获模式：截获接收机扫频速率慢，信号检测能力弱。因此，其可能无法截获雷达信号。

当雷达信号被截获后，获得的信息将作为干扰机的感知信息，形成不同的干扰策略：

(1) 干扰策略1：当目标在第k帧处获取频率信息c_k^q时，干扰机利用感知信息c_k^q和上一帧载频信息c_{k-1}^q共同确定当前干扰信号。否则，干扰机将不会释放任何干扰信号。当$c_k^q = c_{k-1}^q$时，干扰机以载频c_k^q发送点频干扰信号；当$c_k^q \neq c_{k-1}^q$时，干扰机释放阻塞干扰，其中干扰信号的带宽为$(\left|c_k^q - c_{k-1}^q\right| + 1)\Delta f$。特别地，如果$c_{k-1}^q = \varnothing$，干扰机释放载频为$c_k^q$的点频干扰信号。

(2) 干扰策略2：当目标在第k帧截获载频为c_k^q的雷达信号时，干扰机将释放具有相同载波频率的点频干扰信号。否则，干扰机释放固定带宽的阻塞干扰信号来覆盖整个扫频范围。

为了测试一体化发射资源管理方案的适应性，仿真实验还结合不同的拦截模式和干扰策略设置一个动态电磁环境，其参数如表3所示。此外，该方案也将与不同基准方案进行比较：

(1) 随机发射资源管理方案：雷达随机选择发射方案。

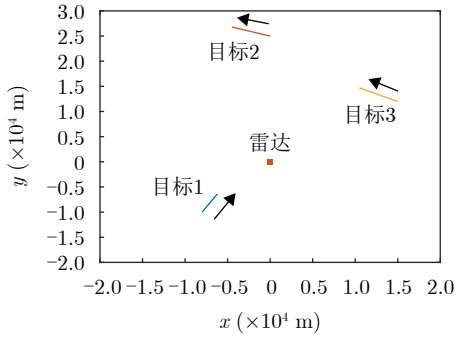


图 2 雷达与目标的空间位置分布

Fig. 2 Spatial position distribution of radar and targets

(2) 传统发射资源管理方案：雷达利用目标状态信息来调整驻留时间和发射功率，发射频率被随机选择[5]。

(3) 低截获发射资源管理方案：截获接收机的参数作为先验信息指导雷达选择合适的驻留时间和发射功率以实现低截获，发射频率被随机选择[11]。

(4) 抗干扰发射资源管理方案：雷达逐渐学习各种干扰策略选择合适的发射频率对抗干扰，同时调整驻留时间和发射功率以提高MTT性能[15]。

在线学习过程中，雷达通过与环境交互获取足够的知识至关重要。因此，仿真实验中雷达以$T = 0.001$ s为跟踪时间间隔与环境进行总共1.8×10^5次交互以获取充足的知识，对应的过程噪声强度设置为$q_0 = 10^{-4}$。然而，展示这种在线学习过程的结果极具挑战性。因此，从在线学习过程中均匀采样100帧数据以评估和展示不同方法的性能。通常，初始采样数据包含的环境信息不足，交互时间的增加使感知信息更加准确，随着在线学习过程的进行，雷达应更多地利用环境信息来生成发射方案。为了保持这种在线学习的特性，所提权衡框架可用于上述具有学习特征的一体化发射资源管理方案和抗干扰发射资源管理方案。通常，贪婪率初始化为1以感知环境，然后每16帧降低0.2，直到达到0，从而提高发射资源管理方案利用率。值得注意的是，随机发射资源管理方案的贪婪率为1，而传统发射资源管理和低截获发射资源管理方案仅依赖跟踪器提供的目标状态信息，其贪婪率设置为0，以确保这两个方案在整个在线过程中始终被执行。

4.2 实验结果及分析

对于动态的电磁环境，雷达利用最近两次观测来近似交互历史。为了更好地展示一体化发射资源管理方案的优势，将所提方法所获得的最差情况均方根误差(Root Mean Square Error, RMSE)与其他基准进行比较。

图3展示了不同方法的最差情况RMSE，可以发现低截获发射资源管理方案的MTT精度最差。这表明仅仅关注低截获反而使得MTT性能变得糟糕。而传统发射资源管理发射方案基于多目标状态分配发射资源。相较于抗干扰发射资源管理方案和

表 3 动态电磁环境参数

Tab. 3 Parameters of dynamic electromagnetic environments

目标索引	截获模式	$\left[B_{\mathrm{I,lo}}^q, B_{\mathrm{I,up}}^q\right]$	$G_{\mathrm{R,I}}^q$	$N_{\mathrm{P},q}$	$B_{\mathrm{I,min}}^q$	干扰策略	$P_{\mathrm{J},q}$	$G_{\mathrm{J,R}}^q$	$B_{\mathrm{J,k}}^q$
1	弱	[0.5, 6.5] GHz	–3 dB	5	0.2 GHz	2	30 W	43 dB	2 MHz
2	弱	[0.5, 3.5] GHz	0 dB	7	0.1 GHz	1	30 W	43 dB	2 MHz
3	强	[0.5, 4.5] GHz	14 dB	5	0.3 GHz	1	30 W	43 dB	2 MHz

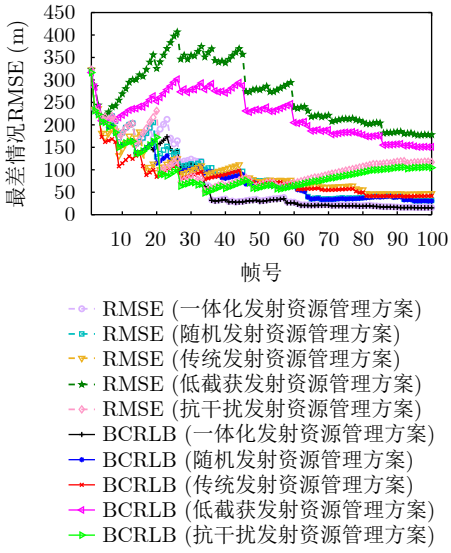


图 3 不同方案的多目标跟踪性能

Fig. 3 MTT performance of different schemes

一体化发射资源管理在初始阶段设置贪婪率为1来随机设置发射方案来感知环境信息，传统发射资源管理方法一直利用目标状态信息引导雷达分配发射资源。因而，利用一定先验信息的传统发射资源管理方法在初始阶段会具备优势。随着交互时间的增加，所提方法的优势更加明显，而仅仅利用干扰信息的抗干扰发射资源管理方案和利用目标状态信息的传统发射资源管理方案由于缺乏完整的环境信息而导致跟踪性能恶化，甚至不如随机发射资源管理方案。

为了研究一体化发射资源管理方案可以最大限度提升MTT精度的原因，图4和图5分别展示一体化和抗干扰两种方法生成的发射方案(脉冲数d_k^q、实际发射功率$P_{\mathrm{R},k}^q$以及载频c_k^q)和目标生成的干扰动作$j_{\mathrm{I},k}^q$。结合MTT性能以及发射方案，可以看出两者起初都随机选择发射方案来探索环境。因此，这两种方法的跟踪性能几乎相同。随着贪婪率降低，一体化方法和抗干扰都提高对动态环境信息的利用率，最终在贪婪率为0时收敛到较为规律的发射模式。但具体资源分配结果存在显著差异。

图4指出目标1在功率域应该具有相对较弱的拦截能力。因而图4(b)的功率选择结果表明一体化发射资源管理方案最终为雷达选择了较低的发射功率($e_k^q=1$, $p_{e_k^q}=100$ W)，可以避免目标1在功率域拦截雷达信号，从而使目标1尽可能释放阻塞干扰信号$j_{\mathrm{I},k}^q=4$。特别地，图4(b)展示同步的实际发射功率$P_{\mathrm{R},k}^q$。由于目标1整体是靠近雷达的。因此，目标1在功率域的截获概率增加。然而，参考距离R_{ref}^q与初始的目标距离$R_{q,1|1}$相关，同步的实际发射功率$P_{\mathrm{R},k}^q$随着目标距离的靠近而降低，使得实际发射功率$P_{\mathrm{R},k}^q$<100 W来保证目标1在功率域的截获能力保持较弱的水平。而对于截获能力弱的目标2，图4(a)表明所提方法仍然选择最小的驻留时间($d_k^q=1$，32个脉冲)和频率1($c_k^q=1$)避免目标2在频域截获雷达信号并释放干扰($j_{\mathrm{I},k}^q=0$)，同时提供最大RCS。在频域完全避开干扰后，雷达为目标2进一步选择最大的发射功率($e_k^q=2$, $p_{e_k^q}=2000$ W)。实际上，目标2在远离雷达，其实际发射功率在$P_{\mathrm{R},k}^q$>2000 W，这本质上可以在利用感知环境信息完全规避干扰的基础上变相增加发射功率，对于提升目标跟踪性能带来好处。类似地，由于频率1和频率2可以提供更大的RCS均值，图4(c)表明一体化发射资源管理方案对于截获能力强的目标3，选择了最大的发射功率($e_k^q=2$, $p_{e_k^q}=2000$ W)，使雷达最终在频率1和频率2之间跳变在降低干扰影响的同时，最大限度地提高当前帧的MTT跟踪性能。而且目标3实际也在靠近雷达，其实际功率也在缩小来保证目标3在功率域的截获能力不会因目标位置改变而产生剧烈变化。

相比之下，图5(b)表示抗干扰发射资源管理方案为所有目标选择较高的发射功率($e_k^q=2$, $p_{e_k^q}=2000$ W)。由于缺乏同步手段，所有目标的实际发射功率为$P_{\mathrm{R},k}^q=2000$ W，这使得目标已经很难在功率域避免干扰。因此，图5(d)显示抗干扰发射资源管理方案选择使雷达持续接收到目标1的点频干扰($j_{\mathrm{I},k}^q=1,2,3$)。对于目标2，抗干扰发射资源管理方案同样使雷达选择最小的驻留时间($d_k^q=1$，32个脉冲)以避免干扰的进入，但相比于一体化发射资源管理方案选择频率1($c_k^q=1$)，抗干扰资源管理方案最终选择频率3($c_k^q=3$)并不能提供最大的RCS。而对于一直可以产生干扰的目标3，抗干扰资源管理方案可以感知到干扰策略，并做出与一体化发射资源管理方案相似的跳频方案。

此外，考虑目标的发射脉冲数是有限的，目标1的阻塞干扰需要覆盖整个扫频范围，因而其干扰功率密度低于目标3，并且其距离雷达更近。因此，从图4(a)展示的整体资源分配结果来看，目标3大多数时刻在$d_k^q=3$(96个脉冲)和$d_k^q=4$(128个脉冲)中跳变以得到更多的驻留时间资源。相应地，一体化发射资源管理方案对目标1在大部分时刻会选择$d_k^q=1$(32个脉冲)和$d_k^q=2$(64个脉冲)，偶尔选择$d_k^q=3$(96个脉冲)提升照射目标1的脉冲数来平衡整体的MTT性能。特别地，图4(d)表明目标1在第96帧形成点频干扰。这是因为雷达可以通过选择较低的发射功率来降低拦截的概率。然而，它并不能完全消除雷达信号被截获的风险。同样96帧的点频干扰使得97帧的预测MTT性能下降。相较于96帧

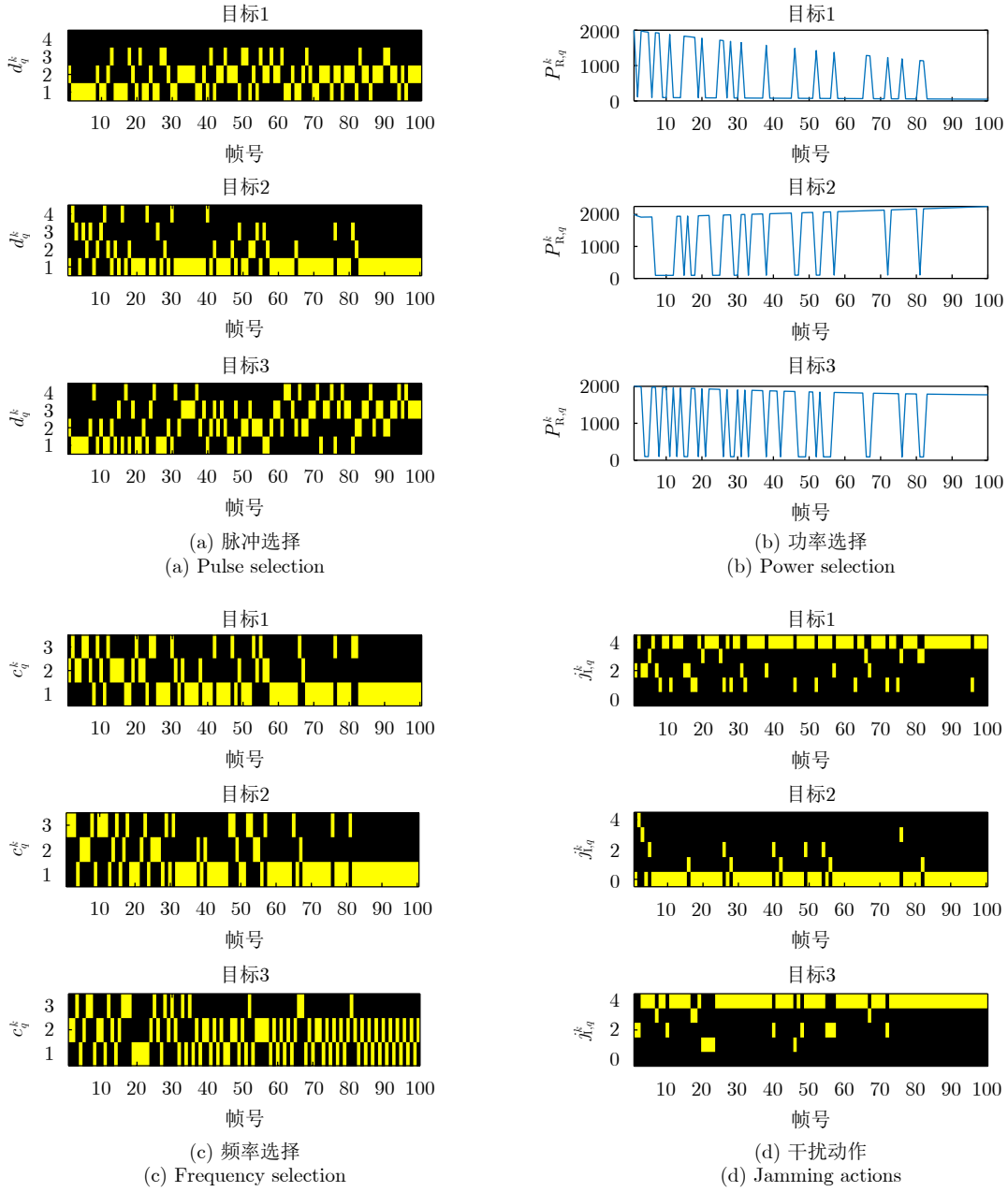


图 4 一体化发射资源管理方案的发射方案以及目标干扰动作

Fig. 4 Transmit scheme of the integrated transmit resource management scheme and jamming actions of targets

选择提供$d_k^q=1$(32个脉冲)。一体化发射资源管理方案在97帧之后一直选择$d_k^q=2$(64个脉冲)为目标1分配更多的驻留时间资源，这也证明了所提方案对于动态干扰环境的适应性。

相比较之下，抗干扰发射资源管理方案使目标1始终受到点频干扰。考虑目标1受到的点频干扰明显强于目标3的阻塞干扰。为了平衡MTT性能，目标1理应需要被分配更多的时间资源。但图5(a)表明抗干扰发射资源管理方案最终使目标1一直得到最少的驻留时间资源($d_k^q=1$，32个脉冲)。由此可以推断抗干扰资源管理方案不能感知完整的环境信息，生成不合理的资源分配结果。

总的来说，通过上述仿真结果可以到一体化发射资源管理方案相较于其他基准表现出更好的MTT性能和适应性。因此，一体化发射资源管理方案是动态电磁环境中处理MTT问题的首选。

4.3 计算复杂度分析

本文构建的优化问题本质为纯整数规划问题，并且很难找到对应的松弛问题。因此，一些借助松弛原始优化问题的整数优化方法很难直接用于求解该问题。相应的穷举搜索方法需要对$(N_{\mathcal{A}})^Q$个联合动作进行计算与判断，进而决定最优发射方案。

此外，分支定界方法具有较强的适应性，可以

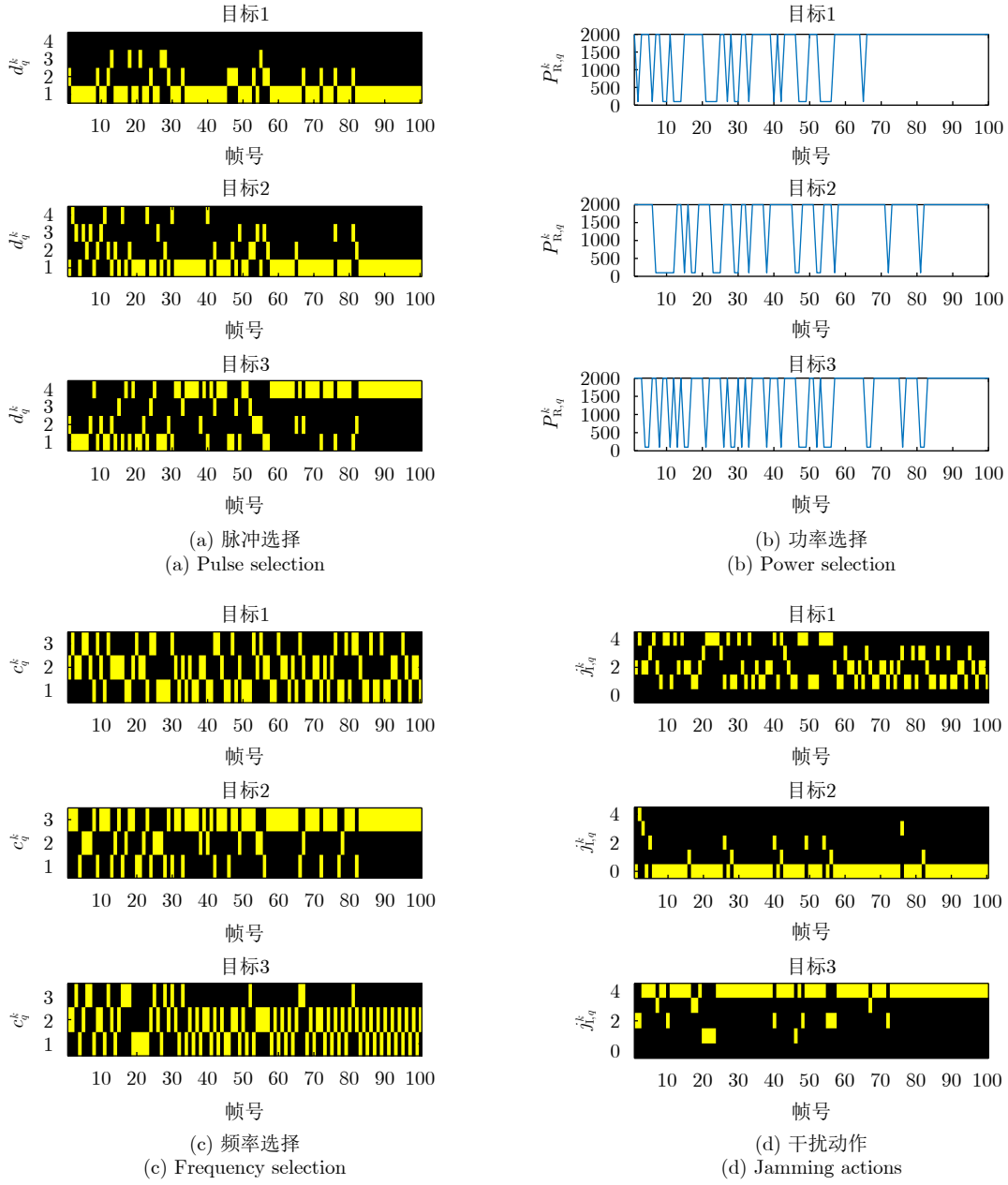


图 5 抗干扰发射资源管理方案的发射方案以及目标干扰动作

Fig. 5 Transmit scheme of the anti-jamming transmit resource management scheme and jamming actions of targets

通过将$(N_{\mathcal{A}})^{Q}$个联合动作按照不同资源类型(脉冲数、频率以及功率)进行分割来减少纯整数优化问题的求解时间。然而该方法需要对整个动作空间进行一定的搜索，其计算复杂度仍然较高。而所提贪婪排序回溯方法仅仅对每个目标的动作进行排序组合，并最终对$QN_{\mathcal{A}}$个动作进行计算和分析，这将有效提高算法的运行效率。

为了进一步展示不同优化算法的计算复杂度，我们给出不同优化算法的跟踪性能，并对每帧的平均优化时间进行统计。正如图6所示，不同优化算法具有类似的跟踪性能。但表4表明所提优化方法可以提高优化时间一个数量级，这充分证明所提方法的优势。

5　结语

本文针对 MTT 问题提出数据驱动的一体化发射资源管理方案，旨在通过在线感知和利用干扰信息引导多功能雷达调整发射方案提高动态电磁环境中的MTT性能。在所提一体化发射资源管理的框架中，我们建立一个主动抗干扰的MDP来描述雷达被敌方拦截和干扰的风险。 然后通过利用该主动抗干扰MDP感知的干扰信息，一体化发射资源

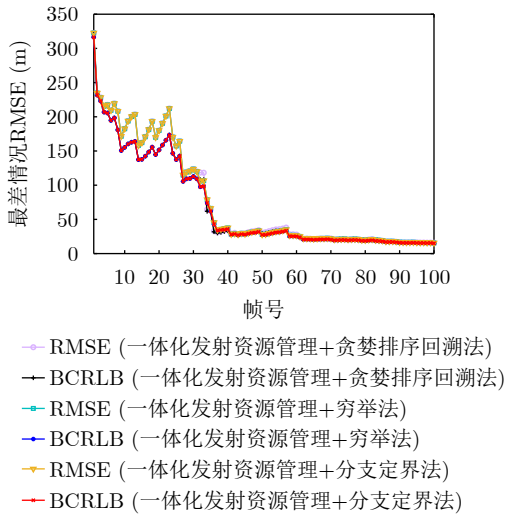


图 6 不同优化算法的多目标跟踪性能

Fig. 6 MTT performance of different optimization algorithms

表 4 不同优化算法运行时间

Tab. 4 Running time of different optimization algorithms

优化算法	运行时间(s)
贪婪排序回溯算法	0.00074
穷举法	0.01790
分支定界法	0.01660

管理方案被设计为一个具有受限动作空间的整数规划问题。最后，一种贪婪排序回溯算法来实现优化雷达发射方案的高效搜索。实验结果也表明相较于其他基准方案，所提的一体化发射资源管理方案可以有效提升MTT性能。

相较于单雷达，多雷达系统展示巨大的潜力。为了提升多雷达系统在动态电磁环境的对抗能力，资源管理方案不仅需要考虑单雷达系统受限的资源，还需探索多雷达探测资源之间存在的潜在联系，这也是未来工作的重点。

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] MORELANDE M R, KREUCHER C M, and KASTELLA K. A Bayesian approach to multiple target detection and tracking[J]. *IEEE Transactions on Signal Processing*, 2007, 55(5): 1589–1604. doi: 10.1109/TSP.2006.889470.

[2] BLACKMAN S S. Multiple-Target Tracking with Radar Applications[M]. Dedham: Artech House, 1986: 1–449.

[3] STONE L D, STREIT R L, CORWIN T L, *et al*. Bayesian Multiple Target Tracking[M]. 2nd ed. Boston: Artech House, 2014: 107–160.

[4] HUE C, LE CADRE J P, and PÉREZ P. Sequential Monte Carlo methods for multiple target tracking and data fusion[J]. *IEEE Transactions on Signal Processing*, 2002, 50(2): 309–325. doi: 10.1109/78.978386.

[5] WANG Xiangli, YI Wei, XIE Mingchi, *et al*. A joint beam and dwell time allocation strategy for multiple target tracking based on phase array radar system[C]. 2017 20th International Conference on Information Fusion (Fusion), Xi'an, China, 2017: 1–5. doi: 10.23919/ICIF.2017.8009856.

[6] 戴金辉, 严俊坤, 王鹏辉, 等. 基于目标容量的网络化雷达功率分配方案[J]. 电子与信息学报, 2021, 43(9): 2688–2694. doi: 10.11999/JEIT200873.

DAI Jinhui, YAN Junkun, WANG Penghui, *et al*. Target capacity based power allocation scheme in radar network[J]. *Journal of Electronics & Information Technology*, 2021, 43(9): 2688–2694. doi: 10.11999/JEIT200873.

[7] YUAN Ye, YI Wei, and KONG Lingjiang. Joint tracking sequence and dwell time allocation for multi-target tracking with phased array radar[J]. *Signal Processing*, 2022, 192: 108374. doi: 10.1016/j.sigpro.2021.108374.

[8] NARYKOV A S, KRASNOV O A, and YAROVOY A. Algorithm for resource management of multiple phased array radars for target tracking[C]. 2013 16th International Conference on Information Fusion, Istanbul, Turkey, 2013: 1258–1264.

[9] YUAN Ye, YI Wei, HOSEINNEZHAD R, *et al*. Robust power allocation for resource-aware multi-target tracking with colocated MIMO radars[J]. *IEEE Transactions on Signal Processing*, 2021, 69: 443–458. doi: 10.1109/TSP.2020.3047519.

[10] SCHLEHER D C. Electronic Warfare in the Information Age[M]. Boston: Artech House, 1999: 1–60.

[11] SHI Chenguang, WANG Yijie, SALOUS S, *et al*. Joint transmit resource management and waveform selection strategy for target tracking in distributed phased array radar network[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2022, 58(4): 2762–2778. doi: 10.1109/TAES.2021.3138869.

[12] ZHANG Haowei, LIU Weijian, ZHANG Qiliang, *et al*. Joint resource optimization for a distributed MIMO radar when tracking multiple targets in the presence of deception jamming[J]. *Signal Processing*, 2022, 200: 108641. doi: 10.1016/j.sigpro.2022.108641.

[13] AILIYA, YI Wei, and VARSHNEY P K. Adaptation of frequency hopping interval for radar anti-jamming based on reinforcement learning[J]. *IEEE Transactions on Vehicular Technology*, 2022, 71(12): 12434–12449. doi: 10.1109/TVT.2022.3197425.

[14] LI Kang, JIU Bo, WANG Penghui, *et al*. Radar active antagonism through deep reinforcement learning: A way to address the challenge of mainlobe jamming[J]. *Signal Processing*, 2021, 186: 108130. doi: 10.1016/j.sigpro.2021.

108130.

[15] ZHANG Peng, YAN Junkun, PU Wenqiang, *et al*. Multi-dimensional resource management scheme for multiple target tracking under dynamic electromagnetic environment[J]. *IEEE Transactions on Signal Processing*, 2024, 72: 2377–2393. doi: 10.1109/TSP.2024.3390119.

[16] YAN Junkun, LIU Hongwei, JIU Bo, *et al*. Simultaneous multibeam resource allocation scheme for multiple target tracking[J]. *IEEE Transactions on Signal Processing*, 2015, 63(12): 3110–3122. doi: 10.1109/TSP.2015.2417504.

[17] YAN Junkun, LIU Hongwei, PU Wenqiang, *et al*. Joint beam selection and power allocation for multiple target tracking in netted colocated MIMO radar system[J]. *IEEE Transactions on Signal Processing*, 2016, 64(24): 6417–6427. doi: 10.1109/TSP.2016.2607147.

[18] LI Nengjing and ZHANG Yiting. A survey of radar ECM and ECCM[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 1995, 31(3): 1110–1120. doi: 10.1109/7.395232.

[19] VAN TREES H L. Detection, Estimation, and Modulation Theory, Part III: Radar-Sonar Signal Processing and Gaussian Signals in Noise[M]. New York: John Wiley & Sons, 2001: 294–307.

[20] SKOLNIK M I. Radar Handbook[M]. New York: McGraw-Hill, 2008: 313–370.

[21] SUKHAREVSKY O I, VASILETS V A, and ZALEVSKY G S. Electromagnetic wave scattering by aerial and ground radar objects[C]. 2015 IEEE Radar Conference (RadarCon), Arlington, USA, 2015: 162–167. DOI: 10.1109/RADAR.2015.7130989.

[22] BERTSEKAS D P. Reinforcement Learning and Optimal Control[M]. Nashua: Athena Scientific, 2019: 1–40.

[23] SUTTON R S and BARTO A G. Reinforcement Learning: An Introduction[M]. 2nd ed. Cambridge: MIT Press, 2018: 37–58.

[24] NERI F. Introduction to Electronic Defense Systems[M]. 2nd ed. Henderson: SciTech Publishing, 2006: 259–368.

[25] STINCO P, GRECO M, GINI F, *et al*. Cognitive radars in spectrally dense environments[J]. *IEEE Aerospace and Electronic Systems Magazine*, 2016, 31(10): 20–27. doi: 10.1109/MAES.2016.150193.

[26] SELVI E, BUEHRER R M, MARTONE A, *et al*. Reinforcement learning for adaptable bandwidth tracking radars[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2020, 56(5): 3904–3921. doi: 10.1109/TAES.2020.2987443.

[27] KOCHENDERFER M J, WHEELER T A, and WRAY K H. Algorithms for Decision Making[M]. Cambridge: MIT Press, 2022: 311–326.

[28] 严俊坤, 纠博, 刘宏伟, 等. 一种针对多目标跟踪的多基雷达系统聚类与功率联合分配算法[J]. 电子与信息学报, 2013, 35(8): 1875–1881. doi: 10.3724/SP.J.1146.2012.01470.
YAN Junkun, JIU Bo, LIU Hongwei, *et al*. Joint cluster and power allocation algorithm for multiple targets tracking in multistatic radar systems[J]. *Journal of Electronics & Information Technology*, 2013, 35(8): 1875–1881. doi: 10.3724/SP.J.1146.2012.01470.

[29] LISI F, FORTUNATI S, GRECO M S, *et al*. Enhancement of a state-of-the-art RL-based detection algorithm for massive MIMO radars[J]. *IEEE Transactions on Aerospace and Electronic Systems*, 2022, 58(6): 5925–5931. doi: 10.1109/TAES.2022.3168033.

作 者 简 介

张　鹏，博士生，主要研究方向为目标检测与跟踪、资源管理。

严俊坤，教授，主要研究方向为雷达智能信号处理、网络化协同探测。

高　畅，讲师，主要研究方向为网络化雷达协同探测、雷达智能化探测。

李　康，副教授，主要研究方向为雷达智能抗干扰、雷达信号处理、强化学习。

刘宏伟，教授，主要研究方向为雷达目标分类与识别、认知网络、网络协同探测。

(责任编辑：于青)