智能天空镜：面向低空经济中安全非正交多址接入的SAC驱动多功能可重构智能表面优化

赵赛、孔凡进、唐栋、吴拓、杨顺兴、黄凯杰，会士，*IEEE*，汤健辉，会士，*IEEE*，以及陆贵文，终身会士，*IEEE*

arXiv:2512.19109v1 [eess.SP] 22 Dec 2025

摘要—**低空经济（LAE）已成为智慧城市建设和经济增长的关键驱动力。为应对低空经济中的频谱效率与通信安全挑战，本文研究利用智能天空镜最大化安全能效（SEE），其中由无人机搭载的多功能可重构智能表面（MF-RIS）辅助非正交多址接入（NOMA）系统。这些空中镜面能够智能放大合法信号，同时对窃听者产生干扰。我们构建了一个联合优化问题，涵盖无人机轨迹、基站功率分配、RIS 相移、放大因子和调度矩阵。针对分式安全能效目标和动态无人机场景，我们提出一种两层优化方案：第一层采用SAC驱动的方法进行轨迹与功率管理，第二层基于信道对齐进行相位优化。仿真结果表明，我们提出的方案显著优于基准方法。**

索引词—**低空经济、智能天空镜、多功能RIS、非正交多址接入（NOMA）、软演员-评论家。**

I. 引言

低空经济（LAE）已成为推动智慧城市发展和经济增长的关键力量。随着物联网（IoT）、城市空中交通和应急救援行动的快速扩展，低空经济对通信网络的覆盖范围和性能提出了严格要求 [1] [2]。这一驱动力推动了通信技术的加速演进，旨在提升频谱效率、优化能源消耗，并确保通信安全和可靠性 [3]–[6]。

在此背景下，搭载于无人机（UAV）上的可重构智能表面（RIS）成为低空通信中的智能空中镜面。这些空中镜面会动态调整信号相位，并策略性地选择位置，以优化复杂环境中的信号传播。现有研究已探索这种协同效应：Yang 等人 [7] 联合优化波束成形、人工噪声、无人机定位和 RIS 相移，以最大化最坏情况保密速率。另一项研究 [8] 通过联合优化虚拟分区和三维空中 RIS 部署，提升了无线安全性。此外， [9] 证明能量管理对于低空经济（LAE）的可持续发展至关重要。

为满足低空经济（LAE）中日益增长的用户接入需求，非正交多址接入（NOMA）技术因能够提升频谱效率并支持大规模连接而受到关注 [10]。然而，NOMA系统面临安全挑战，因为窃听者可能截获叠加用户信号。传统可重构智能表面（RIS）以及具备放大、反射和折射功能的多功能可重构智能表面（MF-RIS）难以应对这一挑战 [11]，[12]。因此，我们采用一种新型MF-RIS，通过为有源单元智能分配双重功能来增强物理层安全性：在生成干扰信号的同时强化合法传输。以往研究 [13], [14] 考察了用于NOMA安全通信的固定式MF-RIS，而 [15] 探索了用于单用户场景的无人机搭载的透射反射可重构智能表面，但忽略了多用户场景和能量因素。因此，将这些智能空中镜与NOMA系统相结合，为实现低空经济中的安全效率权衡提供了一种有前景的解决方案。

本文通过在NOMA系统中利用智能天空镜优化安全能效（SEE），应对低空经济通信所面临的挑战。我们联合优化无人机轨迹、基站功率分配、RIS 放大系数、相移和调度矩阵，以最大化SEE。分式安全能效目标和动态空中镜面移动性要求采用一种新型双层优化框架：由软演员-评论家驱动的第一层用于轨迹与功率管理，基于信道对齐的第二层用于相位优化。该方法能够有效处理问题非凸性和时变信道。仿真结果表明，我们提出的方案显著优于基准方案，验证了软演员-评论家驱动的智能天空镜在实现安全、高能效低空经济通信方面的有效性。

II. 系统模型与问题构建

A. 系统模型

如图 1所示，我们考虑智能天空镜、无人机搭载的透射反射可重构智能表面（MR-UAV）辅助的非正交多址接入下行链路系统。基站和地面用户 $g \in G$ ={窃听者 1、...、K}（包括 K 合法用户和一名窃听者（Eve））均配备单天线。

(通信作者：吴拓。) S. Zhao、F.Kong和D.Tang任职于中国广州510006的广州大学电子与通信工程学院（电子邮箱： {zhaosai,tangdong}@gzhu.edu.cn, fanjin@e.gzhu.edu.cn）。T. Wu和K.-M. Luk任职于中国香港的香港城市大学电子工程系太赫兹与毫米波国家重点实验室（电子邮箱： {tuowu2,eekmluk}@cityu.edu.hk）。S. Yang任职于中国广州510000的中国电信研究院（电子邮箱：yangsx@chinatelecom.cn）。K.-K. Wong任职于英国伦敦WC1E 6BT的伦敦大学学院电子与电气工程系，同时任职于韩国首尔03722的延世大学延世前沿实验室和综合技术学院（电子邮箱：kai-kit.wong@ucl.ac.uk）。K. F. Tong任职于中国香港特别行政区的香港都会大学科学与技术学院。（电子邮箱： {byliu,ktong}@hkmu.edu.hk）。吴拓开展的这项研究工作获得香港研究资助局卓越学科领域计划项目AoE/E-101/23-N的资助。

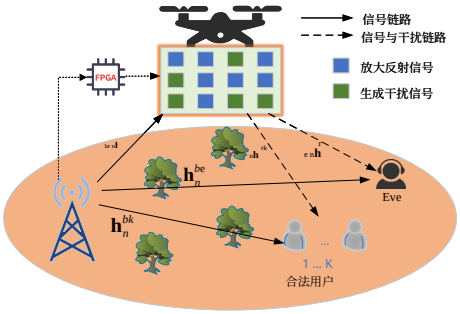

图 1. 智能天空镜辅助的 NOMA 下行链路系统。

总时长 N 由 n 个时长为 δ的时隙组成。无人机以固定高度 H 飞行，以避免与障碍物发生碰撞， [16], 其在时隙 n 的水平坐标表示为 $L_n^u = (x_n, y_n)$。在每个时隙内，无人机的移动策略包括水平飞行方向 $\theta_n \in [0, 2\pi)$ 和水平飞行距离$m_n \in [0, m_{\max}]$。相应地，无人机在时隙 n内的飞行能耗表示为

$$P_n^{\text{uav}} = \delta\Bigg[P_0\left(1+\frac{3v_n^2}{U_{\text{tip}}^2}\right)+\frac{1}{2}d_0\rho s G v_n^3 + P_1\left(\sqrt{1+\frac{v_n^4}{4v_0^4}}-\frac{v_n^2}{2v_0^2}\right)^{1/2}\Bigg] \tag{1}$$

其中， $v_n = m_n/\delta$ 为水平飞行速度； P_0 和P_1 分别为悬停时的桨叶型阻功率和诱导功率； U_{tip} 和 v_0 分别表示悬停时的旋翼桨尖速度和平均旋翼诱导速度； d_0 为机身阻力比； s 为旋翼实度； ρ和 G 分别表示空气密度和旋翼盘面积[17]。

基站与 RIS、RIS 与地面用户以及基站与地面用户之间的信道系数分别记为 $\mathbf{h}_n^{br} = [h_1^{br}, \ldots, h_m^{br}, \ldots, h_M^{br}]^T \in \mathbb{C}^{M\times 1}$、$\mathbf{h}_n^{rg} = [h_1^{rg}, \ldots, h_m^{rg}, \ldots, h_M^{rg}] \in \mathbb{C}^{1\times M}$ 和 $h_n^{bg} \in \mathbb{C}^{1\times 1}$。这些信道建模如下 [18]

$$\begin{aligned} h_m^{br} &= \sqrt{\frac{h_0}{d^{\kappa_1}}}h_{m,\text{LoS}}^{br}, \quad m \in M, \\ h_m^{rg} &= \sqrt{\frac{h_0}{d^{\kappa_2}}}h_{m,\text{LoS}}^{rg}, \quad m \in M, g \in G, \\ h_n^{bg} &= \sqrt{\frac{h_0}{d^{\kappa_3}}}h_{\text{NLoS}}, \quad g \in G. \end{aligned} \tag{2}$$

其中， h_0、 d、 κ分别表示路径损耗、欧几里得距离和路径损耗指数。$h_{m,\text{LoS}}^{br}$ 和$h_{m,\text{视距}}^{rg}$ 为视距（LoS）分量。 h_{NLoS} 为非视距（NLoS）分量。

RIS单元相移 $\theta^{m,n,g}$旨在通过补偿级联信道相位，并使级联信道相位与直达信道相位对齐（即 $\angle\left(h_m^{rg}\theta^{m,n,g}h_m^{br}\right) = \angle h_n^{bg}$，其中 $\angle\cdot$表示复数的相位） [19], [20], 从而增强合成信道 $h_n^g = h_m^{rg}\theta^{m,n,g}h_m^{br} + h_n^{bg}$的增益。本文考虑能够灵活地使RIS单元与地面用户对齐的最优RIS相位。

RIS单元调度矩阵表示为$\mathbf{A}_n^g = \text{diag}(\alpha^{1,n,g}, \ldots, \alpha^{M,n,g})$，其中 $\alpha^{m,n,g} \in \{0,1\}$是第 m个单元与第 g个地面用户之间的信道对齐指示符。此外， $\vartheta_n^g = \text{diag}(\theta^{1,n,g}, \ldots, \theta^{M,n,g})$和 $\boldsymbol{\beta}_n = \overline{\text{diag}}(\sqrt{\beta_n^1}, \ldots, \sqrt{\beta_n^M})$ 分别表示与第 g个地面用户对齐的RIS相移矩阵和RIS放大系数矩阵。

我们假设多反射面无人机具备同时放大反射信号和生成干扰信号的功能 [13], [14]。具体而言，可重构智能表面处的整体相移矩阵$\boldsymbol{\Theta}_n$ 由两部分组成： $\boldsymbol{\Theta}_n = \boldsymbol{\Theta}_n^R + \boldsymbol{\Theta}_n^J$, 其中 $\boldsymbol{\Theta}_n^R = \sum_{k=1}^K \mathbf{A}_n^k\vartheta_n^k\boldsymbol{\beta}_n$ 和 $\boldsymbol{\Theta}_n^J = \mathbf{A}_n^e\vartheta_n^e\boldsymbol{\beta}_n\boldsymbol{\Phi}$。这里，调制矩阵 $\boldsymbol{\Phi}$ 用于在不同符号周期内不规则且动态地改变入射信号的相位，以迷惑窃听者。由于 $\boldsymbol{\Phi}$ 仅影响干扰生成过程，而 ϑ_n^g 与系统性能的关系更为密切，因此此处我们不考虑 $\boldsymbol{\Phi}$的设计，而是重点关注 ϑ_n^g 对性能的影响 [13], [14]。

B. 信号与信道模型

基于上述信道建模和可重构智能表面设计，基站发射的信号表示为

$$\mathbf{x}_n = \sum_{k=1}^K p_n^k s_n^k. \tag{3}$$

其中， p_n^k 是时隙 n内用户 k 的发射功率，且$\mathbb{E}\{|s_n^k|^2\} = 1$。因此，时隙 n 内第 g个地面用户接收的信号可表示为

$$y_n^g = h_n^g\mathbf{x}_n + \mathbf{h}_n^{rg}\mathbf{z}_n + \mathbf{h}_n^{rg}\boldsymbol{\Theta}_n n_s + n_0. \tag{4}$$

其中， $h_n^g = \mathbf{h}_n^{rg}\boldsymbol{\Theta}_n^R\mathbf{h}_n^{br} + h_n^{bg}$表示合成信道， $\mathbf{z}_n = \boldsymbol{\Theta}_n^J\mathbf{h}_n^{br}$ $\mathbf{x}_n$ 表示在可重构智能表面处生成的干扰信号。此外， n_s 和 n_0 分别表示功率为 σ_1^2 的热噪声和功率为 σ_0^2的加性高斯白噪声。

C. 非正交多址接入与安全性考量

在不失一般性的前提下，我们根据信道增益 [10], 确定解码顺序，即 $|h_n^i|^2 \geq \ldots \geq |h_n^j|^2$。因此，用户 i解码用户 j 时的信干噪比表示为

$$\gamma_n^{i,j} = \frac{|h_n^i|^2 p_n^j}{\sum_{l=j+1}^K p_n^l|h_n^i|^2 + |\mathbf{h}_n^{ri}\mathbf{z}_n|^2 + \sigma_1^2\|\mathbf{h}_n^{ri}\boldsymbol{\Theta}_n\|^2 + \sigma_0^2}, \quad j = \{1, \ldots, K\}, i = \{j+1, \ldots, K\}. \tag{5}$$

根据非正交多址接入原理，用户 j的目标信干噪比必须小于或等于其他更强用户解码得到的信干噪比，即应满足约束 $\gamma_n^{j,j} \leq \gamma_n^{i,j}$。用户 j 的目标速率表示为 [21]

$$R_n^j = \log_2(1 + \min(\gamma_n^{j,j}, \gamma_n^{j+1,j}, \ldots, \gamma_n^{K,j})), \quad j = \{1, \ldots, K\}. \tag{6}$$

在最坏情况下，假设窃听者能够消除用户间干扰，窃听者窃听用户 j 的窃听速率可表示为

$$R_n^{e,j} = \log_2\left(1 + \frac{|h_n^e|^2 p_n^j}{|\mathbf{h}_n^{re}\mathbf{z}_n|^2 + \sigma_1^2\|\mathbf{h}_n^{re}\boldsymbol{\Theta}_n\|^2 + \sigma_0^2}\right). \tag{7}$$

结合（1）、（6）和（7），SEE 表示为

$$\mathrm{SEE}_n = \frac{\sum_{k=1}^{K}[R_n^k - R_n^{e,k}]^+}{P_n^{\mathrm{sum}} = P_n^{\mathrm{ris}} + P_n^{\mathrm{uav}}}. \tag{8}$$

其中，$P_n^{\mathrm{ris}} = \sum_{k=1}^{K} p_n^k \|\boldsymbol{\Theta}_n \mathbf{h}_n^{br}\|^2 + \sigma_1^2 \|\boldsymbol{\Theta}_n\|_F^2$ 表示可重构智能表面的功耗。

D. 问题构建

在新兴的低空经济中，无人机的大规模应用对通信安全和无人机能源消耗管理提出了严峻要求。因此，我们的目标是通过联合优化功率分配$\mathbf{P_n} = \{p_n^1, \ldots, p_n^K\}$、可重构智能表面单元调度矩阵 $\mathbf{A}_n^g$、可重构智能表面相移矩阵 ϑ_n^g、可重构智能表面放大因子矩阵 $\boldsymbol{\beta}_n$、无人机飞行距离 m_n 和无人机水平飞行角度 θ_n，最大化多反射无人机辅助非正交多址下行系统的 SEE。该优化问题可构建为

$$(\mathbf{P}_0) \quad \max_{\{\mathbf{P}_n, \mathbf{A}_n^g, \vartheta_n^g, \boldsymbol{\beta}_n, m_n, \theta_n\}} \quad \sum_{n=1}^{N} \mathrm{SEE}_n \tag{9a}$$

$$\text{s.t.} \quad \sum_{k=1}^{K} p_n^k \le P_{b,\max}, \quad \forall n \in \mathcal{N}, \tag{9b}$$

$$P_n^{\mathrm{ris}} \le P_{r,\max}, \quad \forall n \in \mathcal{N}, \tag{9c}$$

$$R_n^k \ge Q_{\min}, \quad \forall k \in \mathcal{K}, \forall n \in \mathcal{N}, \tag{9d}$$

$$\theta^{m}{}_{n}{}^{g} \in [0, 2\pi), \ \forall m \in \mathcal{M}, \forall g \in G, \forall n \in \mathcal{N}, \tag{9e}$$

$$L_n^u \in \mathcal{L}^{\mathrm{fly}}, \quad \forall n \in \mathcal{N}, \tag{9f}$$

其中，$\mathcal{N} = \{1, 2, \ldots, N\}$、$\mathcal{K} = \{1, 2, \ldots, K\}$、$\mathcal{M} = \{1, 2, \ldots, M\}$和 $G = \{$窃听者 1，……，$K\}$ 分别表示时隙、用户、可重构智能表面单元和地面节点的集合。具体而言，（9b）和（9c）分别表示基站发射功率约束和可重构智能表面输出功率约束。（9d）表示用户的服务质量（QoS）约束。（9e）为可重构智能表面相移约束，（9f）为飞行空间约束。由于 SEE 具有分式形式，且无人机场景具有动态性，传统优化方法难以有效求解该问题。因此，我们提出一种基于软演员-评论家算法和信道对齐方法的低复杂度两层优化方案，以实现更快收敛。

III. 所提出的两层优化方案

由于 SEE 的分式形式以及无人机辅助系统的动态特性，问题 $\mathbf{P}_0$ 是一个非凸混合整数优化问题，采用传统优化方法难以求解。为应对这一挑战，我们提出了一种两层优化方案，将问题分解为易于处理的子问题，同时保持计算效率。

在所提出的方案中，第一层采用SAC算法联合优化无人机轨迹、基站发射功率分配、RIS放大因子矩阵和RIS单元调度矩阵。第二层采用信道对齐方法计算RIS单元的最优相移。这种分解策略在确保快速收敛的同时降低了计算复杂度。

A. 基于软演员-评论家的第一层优化

软演员-评论家（SAC）算法是一种离策略演员-评论家方法，它通过同时最大化期望奖励和策略熵 [22]，有效解决深度强化学习中的探索与利用权衡。这一双重目标确保学习到的策略不仅高效，而且具有充分的探索性，因此特别适用于 $\mathbf{P}_0$ 这类复杂优化问题。

SAC 非常适合我们的问题，因为它能够：（1）通过离策略学习有效利用收集到的经验数据；（2）处理高维连续动作空间；（3）在动态环境中保持稳定学习。该算法基于马尔可夫决策过程（MDP）框架运行，并通过智能体-环境交互学习最优策略。

为在软演员-评论家框架下构建优化问题，我们将马尔可夫决策过程定义为四元组 $\langle \mathbb{S}, \mathbb{A}, \mathbb{R}, \mathbb{P} \rangle$，其中：

1）状态空间：状态空间 $\mathbb{S}$ 涵盖决策所需的所有相关系统信息。在时隙 n，状态包括：信道状态信息：$\mathbf{h}_n = \{\mathbf{h}_n^{br}, \mathbf{h}_n^{rg}, h_n^{bg}\}$；无人机和地面用户位置：$L_n = \{L_n^u, L_n^g\}$ 以及可重构智能表面相移矩阵：ϑ_n^g。

因此，状态空间形式化定义如下：

$$\mathbb{S} = \{\mathbf{h}_n, L_n, \vartheta_n^g\}. \tag{10}$$

2）动作空间：动作空间 $\mathbb{A}$包含智能体在每个时隙可调整的所有控制变量：发射功率分配 $\mathbf{P}_n$、可重构智能表面单元调度矩阵$\mathbf{A}_n$、可重构智能表面放大系数矩阵 $\boldsymbol{\beta}_n$、无人机水平飞行距离m_n以及无人机水平飞行角度 θ_n。

完整的动作空间定义如下：

$$\mathbb{A} = \{\mathbf{P}_n, \mathbf{A}_n, \boldsymbol{\beta}_n, m_n, \theta_n\}. \tag{11}$$

3）动作映射：由于策略网络的输出使用值域为 $[-1, 1]$ 的 $\tanh(\cdot)$ 激活函数，因此需要通过动作映射过程将连续输出转换为离散调度决策。对于可重构智能表面单元调度矩阵 $\mathbf{A}_n = [w_\alpha^1, \ldots, w_\alpha^M]$，将每个单元映射到相应的用户分配 $\mathbf{A}_n^g$。

例如，对于 $K = 2$ 个用户，映射规则为：

$$\alpha_n^{m,g} = \begin{cases} 1, & \text{if } \frac{2g'-K-3}{K+1} \le w_\alpha^m < \frac{2g'-K-1}{K+1}, \\ 0, & \text{otherwise}, \end{cases} \tag{12}$$

其中，$g' \in \{1, 2, 3\}$ 表示地面节点 $g \in \{$窃听者 $\mathbf{1}, 2\}$的索引。

此外，发射功率动作 $\mathbf{P}_n$ 被定义为功率分配比。为确保满足约束，当 RIS放大因子 $\boldsymbol{\beta}_n$ 违反功率约束时，我们采用以下调整：

$$\hat{\boldsymbol{\beta}}_n = \sqrt{\frac{P_{r,\max}}{P_n^{\mathrm{ris}}}} \boldsymbol{\beta}_n, \quad \text{if } P_n^{\mathrm{ris}} > P_{r,\max}. \tag{13}$$

4) 奖励函数： 奖励函数的设计对于引导智能体找到最优解至关重要。我们的奖励函数平衡了两个主要目标：(1) 最大化 SEE；以及 (2) 惩罚约束违反。奖励函数表示为：

$$r_n = R_n^{\text{sum}}(1 - w_1\widehat{P}_n^{\text{sum}} - w_2 U_n) - p, \tag{14}$$

其中， $\widehat{P}_{\text{总和}_n} = \frac{P_n^{\text{sum}} - P_{\min}}{P_{\max} - P_{\min}}$ 表示归一化总功耗， $U_n = \frac{1}{K}\sum_{k=1}^{K}[Q_{\min} - R_n^k]^+$ 表示QoS约束违反程度， w_1， w_2 是功耗和QoS惩罚的加权参数， p 是无人机飞出限制区域 $\mathcal{L}^{\text{fly}}$的惩罚。

5）软演员-评论家训练过程：软演员-评论家算法通过两个交替阶段运行：策略评估和策略改进。在策略评估阶段，智能体通过观察当前状态s_n、根据当前策略选择并执行动作 a_n 、接收奖励 r_n，并转移到下一个状态s_{n+1}，与环境交互。每个经验元组 $\{s_n, a_n, r_n, s_{n+1}\}$ 都会存储在经验回放缓冲区$\mathbf{D}$中，以供后续训练。在策略改进阶段，使用从$\mathbf{D}$中随机抽取的 B 经验转移小批次来更新神经网络。每个回合重复此过程 I 次，以确保稳定收敛。

B. 基于信道对齐的第二层优化

我们优化方案的第二层专注于计算RIS单元的最优相移。由于无人机移动性，相位对齐矩阵 ϑ_n^g 必须在每个时隙动态更新，以保持最佳信道对齐。

与第 g个地面用户对齐的第 m个RIS单元的最优相移 $\theta^{m,g}_n$通过最大化合成信道增益确定。这是通过相位对齐补偿级联信道的相位来实现的：

$$\theta_n^{m,g} = e^{j\angle h_n^{bg}} e^{-j\angle h_m^{br}} e^{-j\angle h_m^{rg}}, \tag{15}$$

其中，通过计算相移，使级联信道 $h_m^{br} \cdot h_m^{rg}$ 与直达信道 h_n^{bg}对齐。

相位对齐矩阵 $\vartheta^{m,g}_n$在软演员-评论家智能体执行动作 a_n （即更新无人机位置）后立即更新，并在计算奖励 r_n之前完成。更新后的相位矩阵将成为下一状态s_{n+1} 的一部分，用于后续决策步。完整的两层优化方案如图2所示。

IV. 数值结果与讨论

在本节中，我们将给出全面的数值结果，以评估所提出的智能天空
镜像辅助NOMA下行链路系统。仿真环境由一个 400 m $\times$ 400 m 的正方形区域组成，其中包含基站、多反射面无人机和地面用户。我们考虑 $K = 2$个位于坐标 $[250, 300, 1]$ 和$[250, 315, 1]$, 的位置。基站位于 $[10, 10, 10]$, 而窃听者位于 $[300, 250, 1]$。多反射面无人机从初始位置 $[10, 390, 50]$出发。详细的仿真参数见表 I。

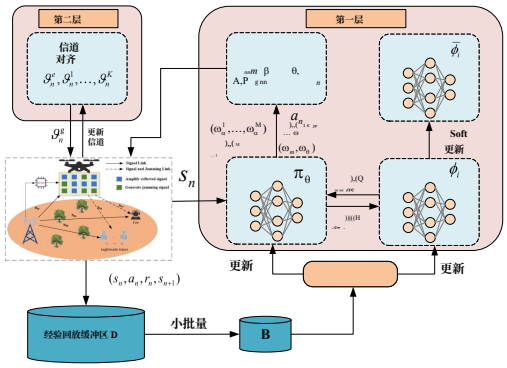

图2. 所提出的两层优化结构。

表 I 仿真参数

参数	设置
U_{tip}, d_0, ρ, s	120, 0.6, 1.225, 0.05
G, v_0, P_0, P_1	0.503, 4.3, 79.86, 88.63
$Q_{\min}$	1 比特/秒/赫兹
σ_0^2, σ_1^2, $P_{b,\max}$, $P_{r,\max}$ (分贝毫瓦)	-105, -105, 30, 0
$\kappa_1, \kappa_2, \kappa_3, h_0$	2, 2, 3.6, 1×10^{-3}
批量大小 B, 经验回放缓冲区大小 $\lvert\mathbf{D}\rvert$	256, 1×10^6
回合 E, 训练迭代次数 I	5000, 70
时隙 N , 时隙持续时间 δ	200, 1s
学习率 ℓ_r	3×10^{-4}
隐藏层	[512, 256]
最大飞行距离 $m_{\max}$	20m

图3 展示了不同强化学习算法在累积奖励方面的收敛性能比较。我们将基于SAC的方法与两种最先进的深度强化学习算法进行比较：近端策略优化（PPO）和双延迟深度确定性策略梯度（TD3）。结果表明，与PPO和TD3相比，SAC实现了更优的收敛性能，具有更快的收敛速度和更高的奖励上限。这种优越性能源于SAC通过离策略学习实现了高效的样本利用，并通过随机策略熵最大化获得了探索优势。

图4 展示了所有对比方案优化后的无人机飞行轨迹，揭示了各自不同的策略行为。所有算法都迅速收敛至各自的最优悬停位置，以最大限度地降低能源消耗。多功能方案（**MF-SAC**、**MF-PPO**和**MF-OMA**）会策略性地将无人机部署在合法用户与窃听者之间，通过智能信号管理同时增强合法通信信道，并削弱窃听能力。相比之下，**AR-SAC**（可重构智能表面仅主动反射信号）会将无人机部署在更靠近合法用户的位置，以最大化信号反射增益，同时尽量减少向窃听者泄漏的信息。相反，**AJ-SAC** （可重构智能表面仅主动生成干扰信号）会将无人机部署在窃听者附近，以最大化干扰有效性，同时尽量减少对合法用户的干扰。值得注意的是，**MF-SAC** 相比**MF-PPO**，实现了效率最高的轨迹，飞行路径距离更短，过渡也更加平滑，这表明了

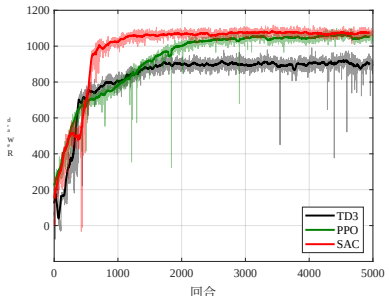

图3. 在 $M = 20$, $\beta_{\max} = 8$下的收敛性能比较。

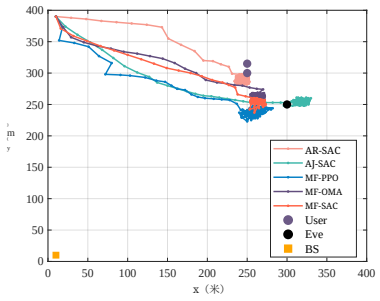

图4. 无人机飞行轨迹。

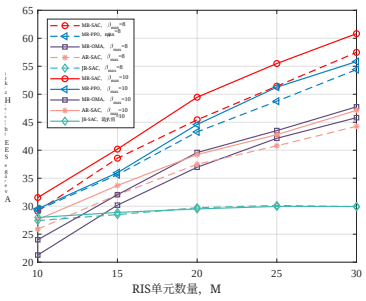

图5. 不同 $\beta_{\max}$下，平均SEE随可重构智能表面单元数量的变化。

SAC算法在面向安全低空经济通信的无人机轨迹优化方面具有更高的学习效率。

图5 展示了不同最大放大因子下，平均SEE性能随可重构智能表面单元数量的变化情况 M 。 $\beta_{\max}$结果清楚地表明，对于不同的 $\beta_{\max}$ 取值，所有方案的SEE均会随着可重构智能表面单元数量的增加而提升，这证实了部署更大规模智能表面的优势。我们提出的**MF-SAC** 方案始终优于所有基准方法，在各种配置下均取得最高SEE值。这一优异表现验证了多功能RIS设计在优化平衡合法信号增强与窃听者干扰方面的有效性，体现了低空经济应用中至关重要的安全性与能效权衡。此外，比较不同放大设置下的结果可知，对于所有方案，较高的 $\beta_{\max} = 10$始终比 $\beta_{\max} = 8$ 带来更高的SEE，这表明放大能力的提升可直接改善系统性能，但能耗也可能随之增加。

V. 结论

本文研究了在低空经济NOMA系统中利用智能天空镜最大化安全能效的问题。我们构建了一个涵盖无人机轨迹、功率分配和RIS配置的联合优化问题，以最大化SEE。我们提出的SAC驱动的双层优化框架有效解决了分式安全能效目标和空中镜面动态移动问题：第一层采用软演员-评论家进行轨迹与功率管理，第二层则采用信道对齐进行相位优化。全面的仿真结果验证了智能天空镜显著优于现有方法，展现了SAC驱动的多功能RIS优化在实现更优安全效率权衡方面的有效性，可应用于未来低空经济通信。

参考文献

[1] W. Ni、R. Luo、X. Zhang、P. Wang、W. Wang和H. Tian，“面向机器人互联网的可重构智能表面”，*IEEE*物联网杂志，第8卷，第2期，第78–86页，2025年。[2] T. Wu、C. Pan、K. Zhi、H. Ren、M. Elkashlan、C.-X. Wang、R. Schober和X.-H. You，“利用高维RIS信息进行定位：故障单元有何影响？”，*IEEE*通信领域精选期刊，第42卷，第10期，第2803–2819页，2024年。

[3] C. Huang、S. Fang、H. Wu、Y. Wang 和 Y. Yang，“低空智能交通：系统架构、基础设施与关键技术”，*J. Ind. Inf. Integr.*，第 42 卷，第 100694 页，2024年。[4] T. Wu、K. Luk、J. Tang、K.-K. Wong、J. Zheng、B. Liu、D. Morales- Jim´enez、M. Elkashlan、K.-F. Tong、C.-B. Chae、F. Adachi 和 G. K. Karagiannidis，“流体天线系统中用于保密性分析的可变分块相关建模与优化”，2025年。[5] T. Wu、Y. Tian、J. Tang、K. Zhi、M. Elkashlan、K.-F. Tong、N. Al-Dhahir、C.-B. Chae、M. Valenti、G. K. Karagiannidis 和 K. Luk，“可扩展流体天线系统：阵列信号处理的新范式”，2025年。[6] X. Lai、T. Wu、L. Mai、M. Elkashlan、N. Al-Dhahir、M. Debbah、G. K. Karagiannidis 和 C. Yuen，“重新审视流体天线系统的空间分块相关模型：从恒定相关到可变相关”，*IEEE J. Sel. Areas Commun.*，第 1–1 页，2025年。[7] H. Yang、S. Liu、L. Xiao、Y. Zhang、Z. Xiong 和 W. Zhuang，“无人机-RIS辅助通信系统中基于学习的可靠安全传输”，*IEEE Trans. Wireless Commun.*，第 23 卷，第 7 期，第 6954–6967 页，2024年。[8] S. Arzykulov、A. Celik、G. Nauryzbayev 和 A. M. Eltawil，“空中RIS辅助物理层安全：最优部署与分区”，*IEEE Trans. Cognit. Commun. Networking*，第 10 卷，第 5 期，第 1867–1882 页，2024年。[9] W. Chen、Y. Zou、J. Zhu 和 L. Zhai，“有源飞行RIS辅助移动边缘计算网络的能效优化：一种深度强化学习方法”，*IEEE Internet Things J.*，第 1 1 页，2025年 – .[10] L. Guo、J. Jia、J. Chen 和 X. Wang，“采用无人机搭载的STAR-RIS的NOMA系统安全通信优化”，*IEEE Trans. Inf. Forensics Secur.*，第 19 卷，第 2300–2314 页，2024年。[11] W. Ni、W. Wang、A. Zheng、P. Wang、C. You、Y. C. Eldar、D. Niyato 和 R. Schober，“一石二鸟：通过多功能RIS增强通信与感知”，*IEEEWireless Commun.*，第 32 卷，第 4 期，第 72–80 页，2025年。[12] W. Ni、A. Zheng、W. Wang、D. Niyato、N. Al-Dhahir 和 M. Debbah，“从单功能RIS到多功能RIS：架构、关键技术、挑战与应用”，*IEEE Netw.*，第 39 卷，46，2025年 1，38第 pp. – .[13] M. Ji、J. Chen、L. Lv、Q. Wu、Z. Ding 和 N. Al-Dhahir，“采用双功能RIS的安全NOMA系统：同时信息中继与干扰”，*IEEE Trans. Commun.*，第 71 卷，6528，2023年 11，6514第 no. pp. – .[14] L. Lv、S. Zhao、Y. Zhou、Y. Chen、G. Huang 和 D. Tang，“双功能IRS辅助认知无线电NOMA网络的安全传输”，*IEEEInternet Things J.*，第 12 卷，10782，2025年 8，10768第 no. pp. – .[15] W. Wang、W. Ni、H. Tian、Y. C. Eldar 和 D. Niyato，“用于对抗无线网络窃听的无人机搭载多功能RIS”，*IEEE Wireless Commun. Lett.*，第 12 卷，第 10 期，第 1667–1671 页，2023年。[16] R. Dong、B. Wang、K. Cao、J. Tian 和 T. Cheng，“利用深度强化学习设计RIS赋能无人机通信网络的安全传输”，*IEEE Trans. Veh. Technol.*，第 73 卷，第 6 期，第 8404–8419 页，2024 年。[17] H. Mei 、 K. Yang 、 Q. Liu 和 K. Wang，“利用深度强化学习设计RIS辅助无人机系统的三维轨迹与相移”，*IEEE Trans. Veh. Technol.*，第 71 卷，3029，2022年 3，3020第 no. pp. – .[18] B. Adhikari、A. S. Khwaja、M. Jaseemuddin、A. Anpalagan 和 A. Nal- lanathan，“采用双延迟的高能效RIS辅助无人机网络。

DDPG技术，”*IEEE Trans. Wireless Commun.*，第23卷，第12期，第18423–18439页，2024年。[19] S. Arzykulov、A. Celik、G. Nauryzbayev和A. M. Eltawil，“人工噪声与可重构智能表面辅助的物理层安全：可重构智能表面最优分区与功率控制”，*IEEE Wireless Commun. Lett.*，第12卷，第6期，第992–996页，2023年。[20] B. Y. D. Rito和K. H. Li，“采用干扰消除和干扰对齐方案的可重构智能表面辅助非正交多址接入”，*IEEE Commun. Lett.*，第27卷，第3期，第1011–1015页，2023年。[21] W. Wang、X. Liu、J. Tang、N. Zhao、Y. Chen、Z. Ding和X. Wang，“智能反射面辅助安全非正交多址网络的波束成形与干扰优化”，*IEEE Trans. Wireless Commun.*，第21卷，第3期，第1557–1569页，2022年。[22] T. Haarnoja、A. Zhou、P. Abbeel和S. Levine，“软演员-评论家：采用随机演员的离策略最大熵深度强化学习”，arXiv:1801.01290，2018年。