

面向6G智能反射面辅助的联邦学习安全速率优化

武影影^{1,2}, 毛伯敏^{1,2}

(1. 西北工业大学网络空间安全学院, 陕西 西安 710129; 2. 西北工业大学太仓长三角研究院, 江苏 太仓 215400)

摘要: 基于非正交多址接入的联邦学习可以实现6G物联网设备的大规模接入、高速率传输及泛在智能。然而, 无线信道的随机性导致模型参数传输时延波动, 而频繁的模型更新可能引发通信-计算资源竞争; 同时, 本地模型上传阶段可能遭遇窃听攻击, 导致用户隐私泄露。为了解决上述问题, 提出了智能反射面辅助的基于非正交多址接入的联邦学习通信系统, 旨在利用智能反射面调控无线传输环境, 增加合法接收者的信号强度并抑制窃听者的接收信号, 为联邦学习构建高可靠、高安全的通信基础。首先建立智能反射面辅助的通信模型; 接着通过联合优化物联网设备发送功率及智能反射面各元件相移, 最大化用户-基站最小安全速率; 最后利用深度确定性策略梯度对上述优化问题求解, 定义发送功率和智能反射面元件相移矩阵为动作空间、信道信息和上一阶段动作为状态空间。仿真结果表明, 与人工噪声方法、块坐标上升方法、不部署智能反射面及不优化智能反射面元件相移相比, 部署并优化智能反射面相移可以大幅提高联邦训练模型更新过程的安全速率, 为后续联邦学习算法在该通信系统中的部署提供物理层保障。

关键词: 非正交多址接入; 6G; 联邦学习; 智能反射面; 安全速率

中图分类号: TN915.08

文献标志码: A

DOI: 10.11959/j.issn.2096-109x.2025021

Federated learning secrecy rate optimization for 6G intelligent reflecting surface assistance

WU Yingying^{1,2}, MAO Bomin^{1,2}

1. College of Cyber Security, Northwestern Polytechnical University, Xi'an 710129, China

2. Yangtze River Delta Research Institute (Taicang), Northwestern Polytechnical University, Taicang 215400, China

Abstract: Non-orthogonal multiple access (NOMA) based federated learning (FL) can achieve the massive connectivity of internet of thing (IoT) devices, high rate transmission, and pervasive intelligence in 6G networks. However, the stochastic channels in NOMA can lead to fluctuations in the transmission delay of model parameters. Fre-

收稿日期: 2024-10-29; 修回日期: 2025-03-02

通信作者: 毛伯敏, maobomin@nwpu.edu.cn

基金项目: 国家重点研发计划 (2024YFF1401305)

Foundation Item: The National Key R&D Program of China(2024YFF1401305)

引用格式: 武影影, 毛伯敏. 面向6G智能反射面辅助的联邦学习安全速率优化[J]. 网络与信息安全学报, 2025, 11(2): 115-124.

Citation Format: WU Y Y, MAO B M. Federated learning secrecy rate optimization for 6G intelligent reflecting surface assistance[J]. Chinese Journal of Network and Information Security, 2025, 11(2): 115-124.

quent model updates may cause competition for communication and computational resources, and the uploaded model can be intercepted, leading to privacy leakage. To address these issues, an intelligent reflecting surface (IRS)-assisted NOMA-based FL communication system was proposed. In this system, IRS was leveraged to reconfigure the wireless communication environment, enhancing the signal strength at the legitimate receivers and suppressing the received signal at eavesdroppers (Eves), thereby providing a high-reliability and high-security communication foundation for federated learning. Firstly, the IRS-assisted wireless communication model was constructed. Subsequently, the minimum user-BS secrecy rate was maximized through the joint optimization of IoT device transmit power and IRS phase shifts. Finally, the deep deterministic policy gradient (DDPG) algorithm was employed to solve the optimization problem, with the action space defined as the transmit power and IRS phase shift matrix, and the channel state information and previous action constituting the state space. Simulation results demonstrated that deploying IRS in the system can greatly enhance the secrecy rate of the FL model update compared to artificial noise (AN), the block coordinate ascent method (BCAM), the scheme without IRS, and the scheme with random phase IRS, providing physical layer support for the subsequent deployment of FL algorithms in this communication system.

Keywords: non-orthogonal multiple access, 6G, federated learning, intelligent reflecting surface, secrecy rate

0 引言

未来 6G 技术有望实现物联网 (internet of things, IoT) 设备的超大规模接入、超高可靠低时延通信及泛在智能^[1]。联邦学习 (federated learning, FL) 作为一种分布式机器学习模型, 通过允许分布式 IoT 设备在只共享模型参数的前提下协同训练^[2], 避免原始敏感数据传输, 同时减少传输时延, 成为推动 6G 实现的使能技术。由于 6G 网络中用户分布不均及数据异构, 需要大量设备共同参与以快速训练可靠机器学习模型^[3]。而密集设备引发的信道接入冲突会导致 FL 模型更新时延呈现显著随机性, 非正交多址接入 (non-orthogonal multiple access, NOMA) 协议通过频谱共享可提升频谱效率^[3], 能支持更多设备参与 FL 迭代。

尽管 FL 中 IoT 设备仅共享本地训练模型参数, 在一定程度上降低了无线传输负载并实现了隐私保护, 但基于 NOMA 的 FL 系统中无线信道的随机性和 FL 训练过程中 IoT 设备和基站 (base station, BS) 之间频繁的数据传输不可避免地增加了训练时延。同时, 模型传输路径上可能存在窃听者 (eavesdropper, Eve) 窃听上传的模型参数^[4]。当 Eve 获取足够多参数时, 可以根据其推断出原始数据, 进而威胁数据隐私。为了解决 NOMA-based FL 中上传模型被窃听问题, Wang 等^[5]和 Qian^[6]等采用了一种基于混沌序列的模型

分发方法。但该方法通过加密和信号扰乱实现安全, 增加了系统开销。Zheng 等^[7]利用低能耗的智能反射面 (intelligent reflecting surface, IRS) 实现模型隐秘传输, 通过联合优化训练设备数量和 IRS 相移矩阵最大化隐秘传输速率。

IRS 作为通信领域的一种新兴前沿技术, 可以有效解决上述速率减小和隐私泄露问题。IRS 是一种由多个低成本无源反射元件组成的可编程表面^[8], 其通过重构无线通信传输环境来提高频谱利用率和能效^[9]。基于此, An 等^[10]将 IRS 应用于临空信息网络, 极大地提高了网络的能效。通过 IRS 控制器, 每个反射元件可以独立地调节相位和 (或) 振幅^[11], 进而为 6G 系统提供可重构的无线传输环境^[12]。通过这种方式, IRS 可以改善信源到合法接收者的信号质量并抑制到 Eve 的信号强度, 进而在提高传输速率的同时保障传输安全。在相关研究方面, Xiu 等^[13]通过在毫米波系统中部署多个 IRS 提升安全速率。

传统方法使用加密或人工噪声 (artificial noise, AN) 来实现系统的安全。在文献[14-15]中, 终端设备和无人机被用作干扰器以避免被窃听。但加密会增加系统的复杂性, 进而增加传输延迟。而在系统中引入干扰, 一方面会增加资源有限 IoT 设备的计算负载, 另一方面会损害接收信号的信噪比, 进而降低 FL 训练性能。Zheng

等^[7]在基于空中通信 (over-the-air communication, OTA) 的 FL (OTA-based FL) 系统中利用智能反射面实现隐秘传输。OTA-based FL 系统虽然与 NOMA-based FL 系统有着相同的无线通信特征, 但它在 6G 的多样化服务提供中缺乏灵活性。

基于此, 本文提出 IRS 辅助的 NOMA-based FL 通信系统, 旨在优化 FL 训练过程, 为 FL 提供高可靠性、高安全性的训练环境。

本文工作的主要贡献如下。

1) 在 NOMA-based FL 通信架构中部署 IRS, 调节 IRS 元件相移, 改善 FL 训练过程中 IoT 设备 BS 的信号质量, 同时抑制 IoT 设备 Eve 的信号强度。

2) 定义安全速率, 并建立最大化最小安全速率优化模型。该模型约束于 IoT 设备发送功率和 IRS 相移矩阵。

3) 利用深度确定性策略梯度 (deep deterministic policy gradient, DDPG) 方法对上述优化问题求解。与 AN 和块坐标上升方法 (block coordinate ascent method, BCAM) 对比, 本文算法具有明显优势。

1 IRS 辅助的 NOMA-based FL 通信系统

1.1 场景设计

IRS 辅助的 NOMA-based FL 通信系统架构如图 1 所示。该系统包含一个有 N_T 根天线的基站、 N_d 台单天线的 IoT 设备, 以及一个有 N_I 个无源反射元件的 IRS。基于 Wyner 窃听信道模型的假设^[6], 假设系统中有 K 个单天线的 Eve, 评估最坏情况下本文方案的有效性。IoT 设备、IRS 反射元件和 Eve 的集合分别表示为 $\mathcal{N}_d = \{1, 2, \dots, N_d\}$ 、 $\mathcal{N}_I = \{1, 2, \dots, N_I\}$ 和 $\mathcal{K} = \{1, 2, \dots, K\}$ 。假设系统位于二维笛卡尔坐标系, 其中 BS(0,0) m, IRS 中心(50,0) m。IoT 设备和 Eve 分别随机分布在以 (0,0) m 和 (250,0) m 为圆心、半径 50 m 的圆内。假设系统内包括 IRS 在内的所有节点的信道状态信息 (channel state information, CSI) 都是可知的^[17], 同时假设信道系数是准静态的^[18], 在联邦学习迭代过程中可以被视作常数。

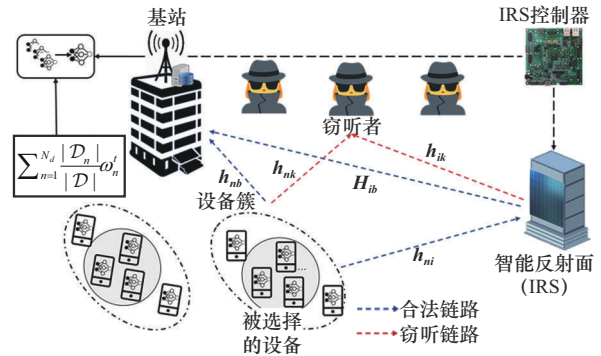


图1 IRS 辅助的 NOMA-based FL 通信系统架构
Figure 1 Architecture of IRS-assisted NOMA-based FL communication system

1.2 联邦学习模型

本文采用联邦平均算法 FedAvg 作为联邦学习模型。假设 N_d 台单天线的 IoT 设备参与本地训练。设备 n 的数据集表示为 $\mathcal{D}_n = \{(\mathbf{x}_i, y_i)\}_{i=1}^{|\mathcal{D}_n|}$, 其中 $(\mathbf{x}_i, y_i)$ 为设备 n 的第 i 个数据对, $\mathbf{x}_i$ 为输入, y_i 为其对应的标签, $|\mathcal{D}_n|$ 为数据集 $\mathcal{D}_n$ 对应的基数。

在第 t 轮训练中, BS 首先将全局模型参数 w^t 分发给每台参与训练的 IoT 设备。IoT 设备根据分发下来的参数基于各自的本地数据执行本地训练, 并更新本地模型参数。本地训练完成后, 参与训练的本地设备将更新的参数模型 w_n^t 上传给 BS, BS 使用加权平均法根据每台 IoT 设备的数据量大小为每台设备分配权重, 并将之与模型参数加权。

本地训练通过最小化本地损失函数实现, 表示为

$$\min_w F_n(w) = \frac{1}{|\mathcal{D}_n|} \sum_{j=1}^{|\mathcal{D}_n|} f(\mathbf{x}_j, y_j; w_j) \quad (1)$$

其中, $F_n(w)$ 为设备 n 的本地损失函数, w 为全局训练参数, $f(\mathbf{x}_i, y_j; w_j)$ 为使用模型参数 w_j 计算的对于数据对 $(\mathbf{x}_i, y_j)$ 的预测损失。本地参数更新定义为 $w_n^t = w_n^{t-1} - \eta \nabla F_n(w)$, 其中 η 为参数更新的学习率, $\nabla F_n(w)$ 为本地损失函数的梯度。每轮训练中, 设备 n 的本地更新会执行 $|\mathcal{D}_n|/M$ 次, M 为批大小。

全局更新通过最小化全局损失函数实现, 全局损失函数通过加权每台设备的损失函数得到。

$$\min_w f(w) = \sum_{n=1}^{N_d} \frac{|\mathcal{D}_n|}{|\mathcal{D}|} F_n(w) \quad (2)$$

其中, $|\mathcal{D}| = \sum_{n=1}^{N_d} |\mathcal{D}_n|$ 。全局更新表示为 $w^{t+1} = \sum_{n=1}^{N_d} \frac{|\mathcal{D}_n|}{|\mathcal{D}|} w_n^t$ 。

1.3 IRS辅助的基于NOMA的通信模型

假设系统中信道经历大尺度路径损耗和小尺度衰落。其中, 大尺度路径损耗遵循对数距离路径损耗模型, 表示为

$$PL = PL_0 - 10\zeta \lg \frac{d}{d_0} \quad (3)$$

其中, PL_0 为参考距离 d_0 处的路径损耗; ζ 为路径损耗指数, 取值与环境 and 信号频带相关; d 为信号发送方和接收方之间的距离。小尺度衰落采用瑞利衰落模型, 表示为 $h_{\text{Rayleigh}} = [h_I - jh_Q]$, h_I 和 h_Q 分别遵循复高斯分布, j 为复虚数单位。

分别用 $\mathbf{h}_{nb} \in \mathbb{C}^{1 \times N_T}$, $\mathbf{h}_{nk} \in \mathbb{C}^{1 \times 1}$, $\mathbf{h}_{ni} \in \mathbb{C}^{1 \times N_I}$ 表示设备 n 到 BS、Eve 及 IRS 的信道, 其中 $\mathbb{C}^{1 \times N_T}$ 表示大小为 $1 \times N_T$ 的复数矩阵; $\mathbf{H}_{ib} \in \mathbb{C}^{N_I \times N_T}$ 和 $\mathbf{h}_{ik} \in \mathbb{C}^{N_I \times 1}$ 分别表示 IRS 到 BS、IRS 到 Eve 的反射信道。IRS 反射元件的相移 Φ 表示为 $\Phi = \{e^{j\varphi_1}, \dots, e^{j\varphi_{N_I}}\}^T \in \mathbb{C}^{N_I \times 1}$, 其中 $\varphi \in [0, 2\pi]$ 。假设 IRS 的反射过程是理想的, 没有能量损耗^[19], 即对于每一个元件 $\Phi_l = e^{j\varphi_l}$, 满足 $|\Phi_l|^2 = 1, l \in \mathcal{N}_I$ 。设备到 BS 的信号是设备直达 BS 和通过 IRS 再到达 BS 两路信号的叠加, 因此设备 n 到 BS 处的有效信道表示为

$$\mathbf{h}_{nb}^{\text{eff}} = \mathbf{h}_{ni} \text{diag}(\Phi) \mathbf{H}_{ib} + \mathbf{h}_{nb}, i \in \mathcal{N}_d \quad (4)$$

同理, 设备 n 到 Eve 的有效信道可以表示为

$$\mathbf{h}_{nk}^{\text{eff}} = \mathbf{h}_{ni} \text{diag}(\Phi) \mathbf{h}_{ik} + \mathbf{h}_{nk}, \forall n \in \mathcal{N}_d, \forall k \in \mathcal{K} \quad (5)$$

在第 t 轮训练中, 设备 n 按照 NOMA 协议将本地训练模型参数 w_n^t 上传给 BS。参数在上传之前需要先被封装为传输符, 用 s_n^t 表示。由于发送给 BS 的信号能够被 Eve 窃听, 假设 BS 和 Eve 遵循相同的通信机制, 则 BS 和 Eve 处接收到的信号分别表示为

$$y_b = \mathbf{H}_{nb} \mathbf{P} \mathbf{s} + \beta_b, \forall n \in \mathcal{N}_d, \forall k \in \mathcal{K} \quad (6)$$

$$y_k = \mathbf{H}_{nk} \mathbf{P} \mathbf{s} + \beta_k, \forall n \in \mathcal{N}_d, \forall k \in \mathcal{K} \quad (7)$$

其中, $\mathbf{H}_{nb} = \{\mathbf{h}_{1b}^{\text{eff}}, \dots, \mathbf{h}_{N_d b}^{\text{eff}}\} \in \mathbb{C}^{N_d \times N_T}$ 和 $\mathbf{H}_{nk} =$

$\{\mathbf{h}_{1k}^{\text{eff}}, \dots, \mathbf{h}_{N_d k}^{\text{eff}}\} \in \mathbb{C}^{N_d \times 1}$ 分别为设备到 BS 和设备到 Eve 的信道矩阵, $\mathbf{P}$ 为设备发送功率矩阵, 满足 $\mathbf{P} \leq \mathbf{P}_{\text{Max}}$, $\mathbf{P}_{\text{Max}}$ 为设备的最大发送功率, $\mathbf{s} = [s_1^t, s_2^t, \dots, s_{N_d}^t]^T$ 为 IoT 设备发送的信号向量, β_b 和 β_k 分别是 BS 和 Eve 处的高斯白噪声。

在采用 NOMA 协议的无线传输系统中, 接收方采用串行干扰删除方法对接收的数据进行解码, 接收方对接收到的信号按信道质量从高到低依次进行解码。假设备 1 到设备 N_d 的信道质量逐渐提高, 则来自第 n 台设备的信号被第 $(N_d - n + 1)$ 设备解码。因此, 来自较差质量信道的信号会成为来自较好质量信道信号的干扰^[20]。设备 n 到 BS 和设备 n 到 Eve 处的传输速率分别如下。

$$R_{nb} = \text{lb} \left(1 + \frac{\mathbf{p}_n |\mathbf{h}_{nb}^{\text{eff}}|^2}{\sum_{j=1}^{n-1} p_j |\mathbf{h}_{jb}^{\text{eff}}|^2 + 1} \right), \forall n \in \mathcal{N}_d \quad (8)$$

$$R_{nk} = \text{lb} \left(1 + \frac{\mathbf{p}_n |\mathbf{h}_{nk}^{\text{eff}}|^2}{\sum_{j=1}^{n-1} p_j |\mathbf{h}_{jk}^{\text{eff}}|^2 + 1} \right), \forall n \in \mathcal{N}_d, \forall k \in \mathcal{K} \quad (9)$$

$\mathbf{p}_n$ 是矩阵 $\mathbf{P}$ 的第 n 个列向量。

窃听能力最强的 Eve 处信噪比最大, 系统安全速率取决于该窃听者。因此, 安全速率可以定义为设备 BS 速率与最大设备 Eve 速率之差, 即

$$R_{sn} = [R_{nb} - \max_k R_{nk}]^+, \forall n \in \mathcal{N}_d, \forall k \in \mathcal{K} \quad (10)$$

其中 $[x]^+ = \max\{x, 0\}$ 。

1.4 问题建模

在采用 NOMA 协议的无线通信系统中, BS 在所有本地设备上传完本地模型参数后才开始进行全局模型聚合。因此, 优化目标为最大化终端设备中最小的安全速率, 目标函数为

$$\begin{aligned} & \text{maximize min } R_{sn} \\ & \text{s.t.: } \mathbf{P} \leq \mathbf{P}_{\text{Max}}, \\ & |\Phi_l|^2 = 1, \forall l \in \mathcal{N}_I \end{aligned} \quad (11)$$

观察可以发现, 由于 IRS 相移约束的存在, 以及传输功率 $\mathbf{P}$ 和 IRS 相移 Φ 的耦合性, 上述优化目标为非凸问题。而传统算法 (如闭式解) 并

不适用于本优化问题。本优化问题求解需要联合优化 $\mathbf{P}$ 和 Φ ，而这两个变量的取值只与当前环境状态相关，因此可以被建模为具有状态空间、行为空间和奖励函数的马尔可夫决策过程。同时，考虑到此优化问题涉及大规模连续状态和动作空间，本文采用基于深度强化学习的DPG优化方法进行求解。

2 基于DDPG的安全速率优化

2.1 DDPG模型设计

DDPG算法基于actor-critic架构，包含4个深度神经网络：actor网络 $\mu(s|\theta^\mu)$ ，critic网络 $Q(s,a|\theta^Q)$ ，以及对应的目标actor网络 $\mu'(s|\theta^{\mu'})$ 和目标critic网络 $Q'(s,a|\theta^{Q'})$ 。其中， s 和 a 分别为状态和动作， θ^μ 、 θ^Q 、 $\theta^{\mu'}$ 和 $\theta^{Q'}$ 分别为对应网络的网络参数。本文所提DDPG算法中的actor网络和critic网络结构相同，都包含1个输入层，2个隐藏层和1个输出层，具体结构如图2所示。

每进行一次迭代训练，critic和actor网络都会相应更新。critic网络以时序差分误差作为损失函数，通过梯度下降法最小化损失函数更新网络参数 θ^Q ，即

$$\begin{cases} \theta^{Q^{t+1}} = \theta^{Q^t} - \mu_c \nabla_{\theta^Q} L_c \\ L_c = (y - q)^2 \end{cases} \quad (12)$$

其中， μ_c 为critic网络更新的学习率， $\nabla_{\theta^Q} L_c$ 为损失函数 L_c 关于网络参数 θ^Q 的梯度， y 为目标 Q 值， q 为当前 Q 值。

$$\begin{cases} y = r^{t+1} + \gamma(1 - \text{done})Q'(s^{t+1}, a^{t+1}|\theta^{Q'}) \\ q = Q(s^t, a^t|\theta^Q) \end{cases} \quad (13)$$

其中， r^{t+1} 为智能体在给定状态 s^t 下执行动作 a^t ，环境给予的奖励； γ 为衰减因子， $\gamma \in [0, 1]$ ，决定未来的奖励所占比重；done为终止符，当done = 1时，表示训练完成。

actor网络通过最大化critic对actor网络的策略梯度更新，具体如下。

$$\theta^{\mu^{t+1}} = \theta^{\mu^t} - \mu_a \nabla_{a'} Q(s^t, a'|\theta^Q) \nabla_{\theta^\mu} \mu(s^t|\theta^\mu) \quad (14)$$

其中， μ_a 为actor网络更新的学习率， $\nabla_{a'} Q(s^t, a'|\theta^Q)$ 为critic网络关于 a' 的梯度， $\nabla_{\theta^\mu} \mu(s^t|\theta^\mu)$ 为actor网络关于 θ^μ 的梯度。

目标网络通过指数滑动平均进行软更新，具体如下。

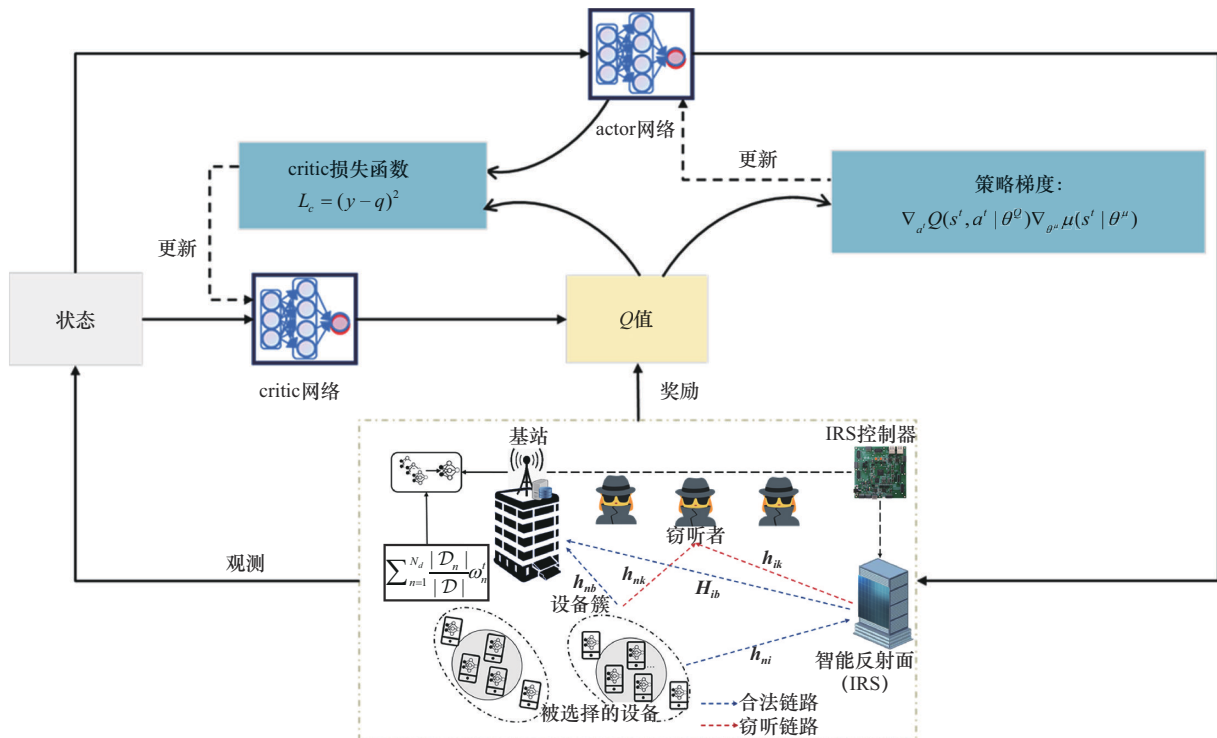


图2 安全速率优化的DDPG算法结构

Figure 2 DDPG algorithm structure for secrecy rate optimization

$$\begin{cases} \theta^{Q'} = \tau_c \theta^{Q'} + (1 - \tau_c) \theta^Q \\ \theta^{\mu'} = \tau_a \theta^{\mu'} + (1 - \tau_a) \theta^\mu \end{cases} \quad (15)$$

其中, τ_c 、 τ_a 分别是目标 critic 网络和目标 actor 网络更新的学习率。 $\tau_c \in (0,1)$, $\tau_a \in (0,1)$ 。

本文采用 DDPG 算法实现安全速率优化策略, 即最大化最小安全速率, 为此需要设计设备发送功率及 IRS 相移矩阵, 并满足系统功率和 IRS 相移约束。

1) 状态空间设计: t 时刻的状态包括当前时刻设备的发送功率 $\mathbf{P}^t$, 信道信息 $\mathbf{h}_{nb}^t$ 、 $\mathbf{h}_{nk}^t$ 、 $\mathbf{h}_{ni}^t$ 、 $\mathbf{H}_{ib}^t$ 和 $\mathbf{h}_{ik}^t$, 以及 $t-1$ 时刻应用的行为 a^{t-1} , 则 t 时刻状态空间 s^t 为

$$s^t = \{ \mathbf{P}^t, \mathbf{h}_{nb}^t, \mathbf{h}_{nk}^t, \mathbf{h}_{ni}^t, \mathbf{H}_{ib}^t, \mathbf{h}_{ik}^t, a^{t-1} \} \quad (16)$$

由于状态空间的变量都为复数, 不能直接作为神经网络的输入, 需要将其拆分为实数和虚数部分, 分别作为神经网络的输入。以发送功率为例, 设备 n 的发送功率可以表示为 $p_n = \Re(p_n) + \Im(p_n)$, 其贡献的输入维度为 $2N_T$ 。同理, 可得 t 时刻信道信息贡献的输入维度为 $2(N_T + N_d N_I + N_I N_T + N_d + N_I)$ 。上一时刻行为维度为 $2N_T + 2N_I$ 。通过式(17)可以得到状态空间的总维度 D_s 。

$$D_s = 2(2N_T + N_d + N_d N_T + N_d N_I + N_I N_T + 2N_I) \quad (17)$$

2) 动作空间设计: 本文通过设计设备发送功率 $\mathbf{P}$ 和 IRS 相移矩阵 $\text{diag}(\Phi)$ 最大化最小安全速率, 因此, 定义 t 时刻的动作空间 $a^t = \{ \mathbf{P}^t, \text{diag}(\Phi^t) \}$ 。通过式(18)可以得到动作空间维度 D_a 。

$$D_a = 2N_T + 2N_I \quad (18)$$

3) 奖励设计: 奖励函数应与系统目标一致。因此, 定义奖励函数为 $\min R_{sn}$, 表示在选定终端设备发送功率和 IRS 相移后最小安全速率的大小。

4) 约束条件: 设备发送功率 $\mathbf{P}$ 和 IRS 相移 Φ 需要分别满足 $\mathbf{P} \leq P_{\text{Max}}$ 和 $|\Phi_l|^2 = 1, \forall l \in \mathcal{N}_I$ 。为了实现上述约束, 在 actor 网络中加入标准化步骤。发送功率的约束通过式(19)实现。

$$\mathbf{P} = \frac{\mathbf{P}}{\sqrt{P_{\text{Max}}}} \quad (19)$$

由于 IRS 相移为复数, 因此其实数部分 $\Re(\Phi_l)$ 和虚数部分 $\Im(\Phi_l)$ 分别需要满足

$$\begin{cases} \Re(\Phi_l) = \frac{\Re(\Phi_l)}{\sqrt{\Re(\Phi_l)^2 + \Im(\Phi_l)^2}}, \\ \Im(\Phi_l) = \frac{\Im(\Phi_l)}{\sqrt{\Re(\Phi_l)^2 + \Im(\Phi_l)^2}} \end{cases} \quad (20)$$

2.2 安全速率优化的 DDPG 学习

本文目标是通过 DDPG 算法选择最优发送功率 $\mathbf{P}$ 和 IRS 相移 Φ , 以最大化最小安全速率。在本文中, DDPG 算法共执行 N 轮, 每轮迭代 T 次。算法首先初始化 t 时刻网络参数: actor 网络 θ^μ 、critic 网络 θ^Q 、目标 actor 网络 $\theta^{\mu'}$ 和目标 critic 网络 $\theta^{Q'}$, 设备发送功率 $\mathbf{P}$ 、IRS 相移 Φ 及经验回放缓冲区 R 。智能体与环境交互, 获取当前环境状态 s^t 。将 s^t 作为 actor 网络输入, 计算 s^t 下的动作 a^t 。执行动作, 并观察奖励 r^{t+1} 及下一时刻状态 s^{t+1} 。

若奖励与预设最优奖励 $R = \log_2(\frac{1 + p_n |h_{nb}^{\text{eff}}|^2}{N_d - 1})$ 相等, 则令终止状态 $\text{done} = 1$, 否则 $\text{done} = 0$, 并将 $(s^t, a^t, r^{t+1}, s^{t+1}, \text{done})$ 存入经验缓存区 R 。接着从经验缓存区 R 随机抽取大小为 M 的样本 $(s^i, a^i, r^{i+1}, s^{i+1}, \text{done})$, 用 (s^i, a^i) 作为 critic 网络输入计算当前 Q 值 q , (s^{i+1}, a^{i+1}) 作为目标 critic 网络输入计算目标 Q 值 y , 构造损失函数, 并通过最小化损失函数更新 critic 网络。同时根据式(14)更新 actor 网络, 并根据式(15)更新目标网络。DDPG 算法伪码如算法 1 所示。

算法 1 优化安全速率的 DDPG 算法

输入 信道信息 $\mathbf{h}_{nb}$ 、 $\mathbf{h}_{nk}$ 、 $\mathbf{h}_{ni}$ 、 $\mathbf{H}_{ib}$ 、 $\mathbf{h}_{ik}$

输出 最优行为 $a^* = \{ \mathbf{P}^*, \Phi^* \}$

初始化: actor 网络 θ^μ 、critic 网络 θ^Q 、目标 actor 网络 $\theta^{\mu'}$ 和 critic 网络 $\theta^{Q'}$, 设备发送功率 $\mathbf{P}$ 、IRS 相移 Φ 及经验回放缓冲区 R

1) while True do

2) for $i = 0$ to $N - 1$ do

3) 收集信道信息获得初始状态 s^0 ;

4) for $j = 0$ to $T - 1$

5) 根据当前策略选择动作 a^t , 即发送功率 $\mathbf{P}$ 和 IRS 相移 Φ 的值;

6) 执行动作 a^t , 并观察系统奖励 r^{t+1} (即安全速率的值) 及下一状态 s^{t+1} ;

7) 将 $(s^t, a^t, r^{t+1}, s^{t+1}, \text{done})$ 存入 R ;

8) 从 R 随机抽取大小为 M 的样本数据;

- 9) 根据式(10)计算 Q 值；
- 10) 根据式(9)构造损失函数，利用随机梯度下降最小化损失函数更新 critic 网络；
- 11) 根据式(11)更新 actor 网络；
- 12) 根据式(12)更新目标网络；
- 13) **end**
- 14) **end**
- 15) **end**

2.3 算法复杂度分析

DDPG 算法的复杂度主要取决于 actor 网络的层数及其输入输出维度。由于 actor 网络为全连接网络，其一个线性层的复杂度可以表示为 $O(D_s D_a)$ ，其中 D_s 为该层输入维度， D_a 为该层输出维度。因此，DDPG 算法的复杂度可以表示为 $O\left(\sum_{l=1}^L D_s^l D_a^l\right)$ ， L 为网络层数，本算法中 $L = 4$ 。 D_s^l 和 D_a^l 分别为 l 层的输入维度、输出维度。

3 仿真实验

仿真的硬件环境配置如表1所示。软件环境为 Python 3.11.3、Pycharm。仿真参数设置如表2所示。为了验证本文算法的有效性，将本文算法与以下4种机制进行对比：①AN；②BCAM；③没有 IRS 辅助（Without IRS）的 NOMA-based FL；④随机相移 IRS（Random-phase IRS）辅助的 NOMA-based FL。

表1 实验配置

硬件环境	配置
CPU	英特尔 i5-12490F 3.00 GHz
GPU	英伟达 GeForce RTX 4060
RAM	32 GB, 2 666 MHz

本文中不同 P_{Max} 值下 DDPG 算法的收敛曲线如图3所示。从图中可以看出，在两种取值下，DDPG 算法在初始阶段都随机进行发送功率和 IRS 相移探索，通过与环境交互获得经验。随着迭代次数增加，DDPG 算法利用这些经验逐步优化策略，总奖励值不断增大，最后学到最优策略，累积期望奖励值在小范围内波动。不同的是，在 P_{Max} 值较小时，DDPG 算法的收敛速度更

快。这是因为增大发送功率会相应地增加系统干扰，进而影响模型更新，导致收敛减慢。

表2 仿真参数

符号	意义	取值
P_{Max}	设备最大发送功率	30 dBW
d_0	参考距离	1 m
PL_0	参考位置路径损耗	-30 dB
μ_a	actor 网络更新学习率	0.001
μ_c	critic 网络更新学习率	0.001
τ_a	目标 actor 网络更新学习率	0.001
τ_c	目标 critic 网络更新学习率	0.001
γ	衰减因子	0.99
D	经验缓存区大小	100 000
M	批大小	16
N	最大轮数	5 000
T	每轮最大迭代次数	10 000

图4展示了本文算法、AN及BCAM在不同参数下的运算时间。AN、BCAM及本文算法的运算时间均随设备数量和窃听者数量的增加而增长。此外，IRS 反射元件数量的增加也会进一步提升本文算法的计算开销。由图中可以看出，本文算法运算时间显著高于 AN 和 BCAM，但仍处于可接受范围内，且随着硬件性能的提升，运算时间将进一步减少。

尽管 AN 和 BCAM 在运算时间上优于本文算法，但其性能表现远不及本文算法。如图5所示，当 $K = 1$ ， $P = 30$ 时，本文算法的安全速率可达到 $1.7389 \text{ bit}/(\text{s}\cdot\text{Hz}^{-1})$ ，而 AN 和 BCAM 的最大安全速率分别为 $0.2323 \text{ bit}/(\text{s}\cdot\text{Hz}^{-1})$ 和 $2.97 \times 10^{-5} \text{ bit}/(\text{s}\cdot\text{Hz}^{-1})$ 。此外，没有 IRS 辅助的 NOMA-based FL 系统的安全速率为 $1.0048 \text{ bit}/(\text{s}\cdot\text{Hz}^{-1})$ 。可以看出，在 IRS 的辅助下，NOMA-based FL 系统的安全速率提升了 73%。

不同机制下 IRS 反射元件数量对安全速率的影响如图6所示，IRS 反射元件数量增加，安全速率也相应增加。图中蓝色柱状图表示本文算法，可以看出，与另外两种机制相比，不管 IRS 反射元件数量如何变化，本文算法的安全速率均是最高。随着 IRS 反射元件数量从 4 增加到 36，本文算法的安全速率从 $1.1714 \text{ bit}/(\text{s}\cdot\text{Hz}^{-1})$ 提高到

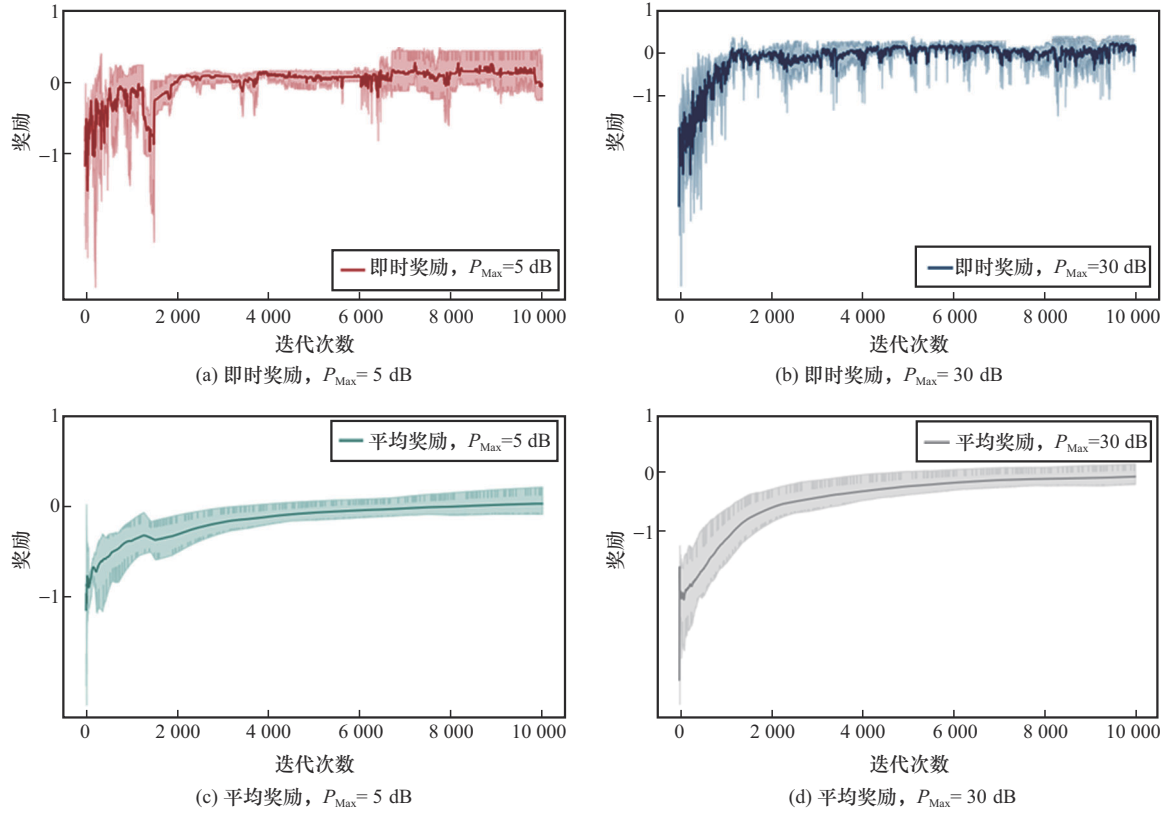
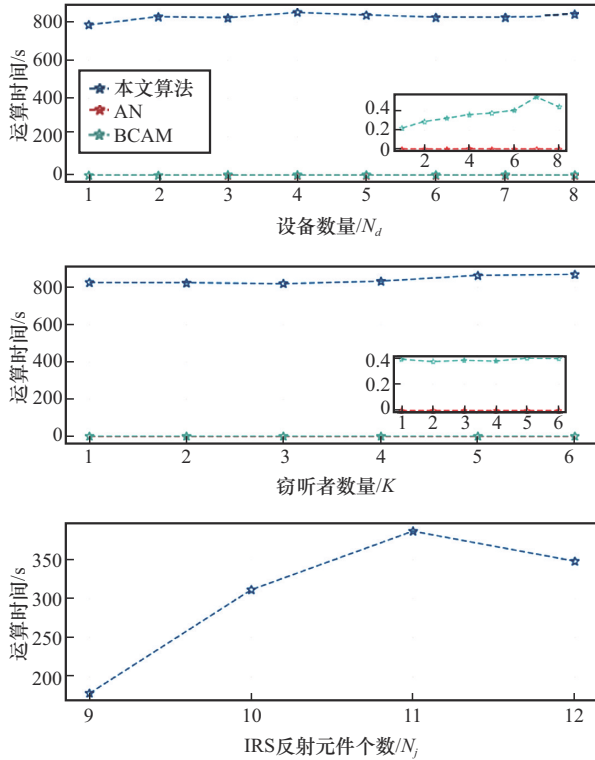
图3 不同 P_{Max} 值下DDPG算法的收敛曲线Figure 3 Convergence curve of DDPG algorithm under different P_{Max} values

图4 算法运算时间

Figure 4 Time consumed of the algorithms

1.950 4 bit/(s·Hz⁻¹), 而在没有 IRS 辅助的情况下, 安全速率则维持在 0.9474 bit/(s·Hz⁻¹) 不变。Random phase IRS 辅助的 NOMA-based FL 系统的安全速率不断波动, 导致系统性能不稳定。由此看出, IRS 的部署及对 IRS 相移的优化能够有效提高安全速率。

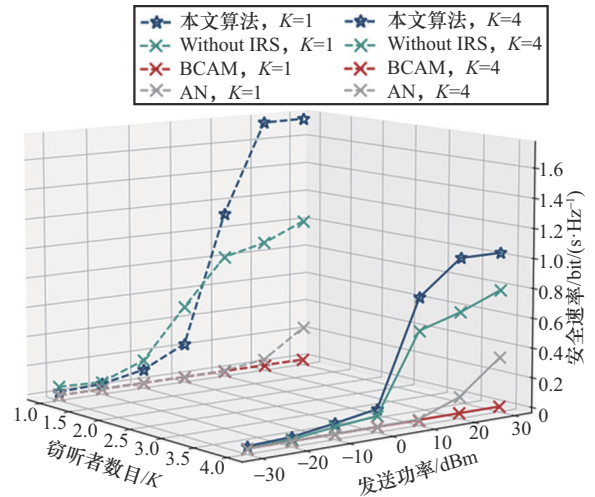


图5 不同Eve数量下发送功率对安全速率的影响

Figure 5 Impact of transmission power on secrecy rate under different number of Eve

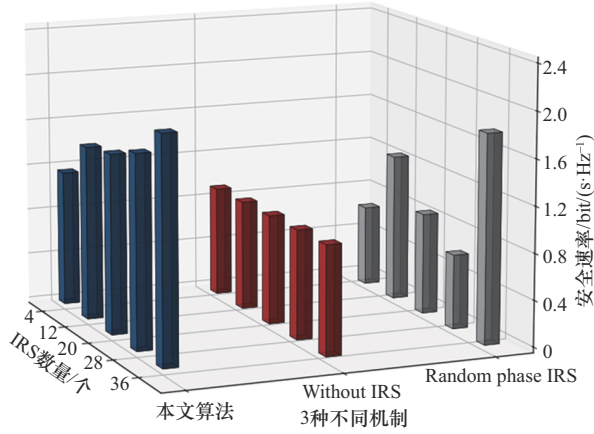


图6 不同机制下IRS反射元件数量对安全速率的影响

Figure 6 Impact of the number of IRS reflective elements on secrecy rate the rate under different operating mechanisms

图7探究了不同系统设定下发送功率对安全速率的影响。从图中可以看出，当系统中BS发送天线和设备较少时，发送功率的改变对安全速率影响较大，而当二者数量增加时，发送功率对安全速率的影响则越来越小。同时可以看出，当发送功率不变时，系统设备数量由4增加到10时，安全速率下降，且发送功率越大，安全速率下降幅度越明显。在 $P=0$ 时，系统设备数量由4增加到10时，安全速率下降73%， $P=30$ 时，安全速率下降95%。

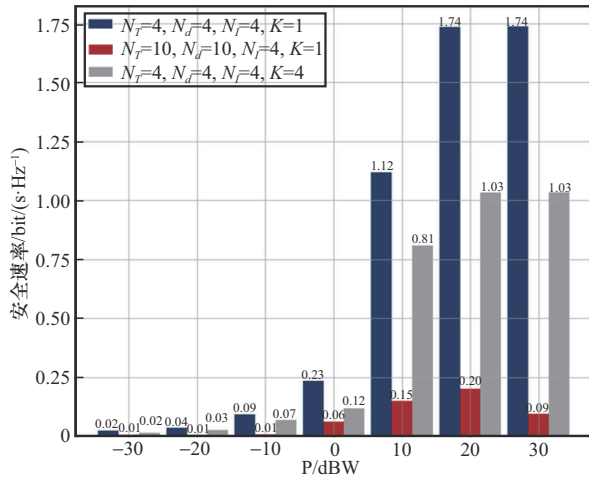


图7 不同系统设定下发送功率对安全速率的影响

Figure 7 Impact of transmission power on secrecy rate under different system settings

图8探讨了不同IoT设备、Eve和IRS反射元件数量对安全速率的影响。结果表明，随着IoT设备和Eve数量增多，安全速率下降。这是因为

随着IoT设备和Eve数量增加，用户之间的干扰也相应增加，使信号干扰变得更加复杂。此外，IRS反射元件数量的增加提升了安全速率，因为更多的IRS反射元件可以使信号更精确地对准IoT设备，从而增加设备处的信噪比，进而提升系统安全速率。

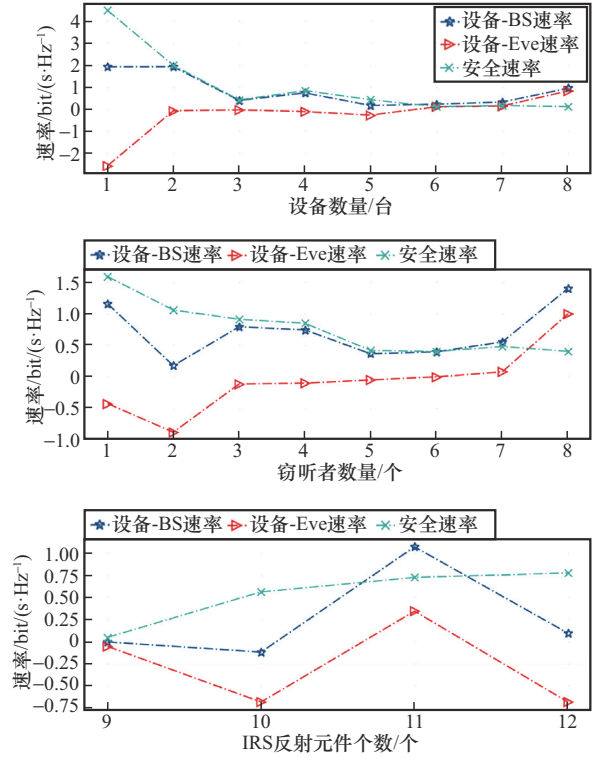


图8 参与者数量变化对安全速率的影响

Figure 8 Impact of changes in the number of participants on secrecy rate

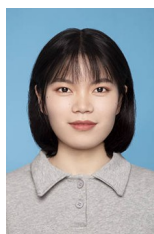
4 结束语

本文提出一种IRS辅助的NOMA-based FL通信系统，借助IRS的智能反射特性优化信道环境，解决模型传输效率与隐私安全问题。IRS被部署在设备与基站间，通过联合优化发送功率与IRS相移矩阵，构建安全速率最大化模型，并采用DDPG算法实现动态策略搜索。仿真结果表明，经优化的IRS配置较传统方案（如AN、BCAM等）可提升安全速率，为联邦学习部署提供物理层安全保障。鉴于本文工作的局限性，未来将研究轻量化算法以降低DDPG算法的时间复杂度，同时在本文所提通信系统中部署FL，进一步验证并优化FL模型的准确率。

参考文献:

- [1] 廖建新, 戚琦, 王敬宇, 等. 6G 智能业务网络: 愿景、架构与关键技术[J]. 中国科学: 信息科学, 2024, 54(5): 991-1024.
LIAO J X, QI Q, WANG J Y, et al. 6G intelligent service networking: vision, architecture, and key technologies[J]. Science in China (Information Sciences), 2024, 54(5): 991-1024.
- [2] DANG S P, AMIN O, SHIHADA B, et al. What should 6G be?[J]. Nature Electronics, 2020, 3(1): 20-29.
- [3] LE T H T, CANTOS L, PANDEY S R, et al. Federated learning with NOMA assisted by multiple intelligent reflecting surfaces: latency minimizing optimization and auction[J]. IEEE Transactions on Vehicular Technology, 2023, 72(9): 11558-11574.
- [4] MAO B M, LIU J J, WU Y Y, et al. Security and privacy on 6G network edge: a survey[J]. IEEE Communications Surveys Tutorials, 2023, 25(2): 1-1.
- [5] WANG T S, HUANG X M, SONG Y X, et al. Energy optimization for NOMA assisted federated learning with secrecy provisioning[C]// 2021 IEEE/CIC International Conference on Communications in China (ICCC). Piscataway: IEEE Press, 2021: 1189-1194.
- [6] QIAN L P, LI M Q, YE P, et al. Secrecy-driven energy minimization in federated-learning-assisted marine digital twin networks[J]. IEEE Internet of Things Journal, 2024, 11(3): 5155-5168.
- [7] ZHENG J, ZHANG H J, KANG J W, et al. Covert federated learning via intelligent reflecting surfaces[J]. IEEE Transactions on Communications, 2023, 71(8): 4591-4604.
- [8] 席云光, 李素月. 空中智能反射面辅助毫米波通信系统能效优化[J]. 电讯技术, 2024, 64(12): 1-11.
XI Y G, LI S Y. Optimizing energy efficiency of aerial reconfigurable intelligent surface-assisted mmWave communication systems[J]. Telecommunication Engineering, 2024, 64(12): 1-11.
- [9] ZHU Y S, MAO B M, KATO N. Intelligent reflecting surface in 6G vehicular communications: a survey[J]. IEEE Open Journal of Vehicular Technology, 2022, 3: 266-277.
- [10] AN P G, YANG P, CAO X B, et al. Energy-efficient URLLC service provision via a near-space information network[J]. IEEE Transactions on Wireless Communications, 2024, 23(8): 9839-9853.
- [11] ZHU Y S, MAO B M, KATO N. A dynamic task scheduling strategy for multi-access edge computing in IRS-aided vehicular networks[J]. IEEE Transactions on Emerging Topics in Computing, 2022, 10(4): 1761-1771.
- [12] 王平, 杨志伟, 李贺举. 智能反射面赋能的联邦边缘学习及其在车联网中的应用[J]. 通信学报, 2023, 44(10): 46-57.
WANG P, YANG Z W, LI H J. Federated edge learning with reconfigurable intelligent surface and its application in Internet of vehicles [J]. Journal of Communications, 2023, 44(10): 46-57.
- [13] XIU Y, ZHAO J, YUEN C, et al. Secure beamforming for multiple intelligent reflecting surfaces aided mmWave systems[J]. IEEE Communications Letters, 2021, 25(2): 417-421.
- [14] WANG T S, LI Y, WU Y, et al. Secrecy driven federated learning via cooperative jamming: an approach of latency minimization[J]. IEEE Transactions on Emerging Topics in Computing, 2022, 10(4): 1687-1703.
- [15] CONSUL P, BUDHIRAJA I, GARG D. A hybrid secure resource allocation and trajectory optimization approach for mobile edge computing using federated learning based on Web 3.0[J]. IEEE Transactions on Consumer Electronics, 2024, 70(1): 1167-1179.
- [16] WYNER A D. The wiretap channel[J]. Bell System Technical Journal, 1975, 54(8): 1355-1387.
- [17] ZHU Y S, MAO B M, KATO N. On a novel high accuracy positioning with intelligent reflecting surface and unscented kalman filter for intelligent transportation systems in B5G[J]. IEEE Journal on Selected Areas in Communications, 2024, 42(1): 68-77.
- [18] ZHU Y S, MAO B M, KATO N. IRS-aided high-accuracy positioning for autonomous driving toward 6G: a tutorial[J]. IEEE Vehicular Technology Magazine, 2024, 19(1): 85-92.
- [19] SALEEM R, NI W, IKRAM M, et al. Deep-reinforcement-learning-driven secrecy design for intelligent-reflecting-surface-based 6G-IoT networks[J]. IEEE Internet of Things Journal, 2023, 10(10): 8812-8824.
- [20] WANG W, NI W L, TIAN H, et al. Intelligent omni-surface enhanced aerial secure offloading[J]. IEEE Transactions on Vehicular Technology, 2022, 71(5): 5007-5022.

[作者简介]



武影影 (1999-), 女, 四川泸州人, 西北工业大学博士生, 主要研究方向为智能反射面、通信系统安全。



毛伯敏 (1989-), 男, 湖北黄冈人, 博士, 西北工业大学教授、博士生导师, 主要研究方向为通信系统安全、空天地一体化网络。