基于RIS和对抗网络辅助的语义隐蔽通信方案

易印雪[1,2]，沙婵[1]，唐睿[1]，张祖凡[1]
（1.重庆邮电大学通信与信息工程学院，重庆 400065；2.重庆邮电大学网络空间安全与信息法学院，重庆 400065）

摘　要：针对传统隐蔽通信在低信噪比（SNR）条件下语义信息重构性能受限的问题，提出一种融合可重构智能表面（RIS）和三方生成对抗网络（GAN）的语义隐蔽通信方案。首先，联合物理层与语义层，构建语义隐蔽保真度最大化的优化问题。然后，提出一种分层优化框架。在物理层，在满足Kullback－Leibler（KL）散度的隐蔽约束下，联合优化发射功率与RIS相移，以构造有利于信号传输的信道环境；在此基础上，语义层引入面向语义信息的三方对抗训练机制，通过端到端学习，实现语义恢复性能与隐蔽性的协同优化。实验表明，所提方案相较最优基准方案的BLEU分数提升29.8%，同时窃听者检测准确率接近0.5。

关键词：隐蔽通信；语义通信；可重构智能表面；生成对抗网络；KL散度
中图分类号：TN92
文献标志码：A
DOI：10.11959/j.issn.1000

Semantic Covert Communication Scheme Assisted by RIS and Adversarial Networks

YI Yinxue[1,2], SHA Chan[1], TANG Rui[1], ZHANG Zufan[1]
1. School of Communications and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
2. School of Cyber Security and Information Law, Chongqing University of Posts and Telecommunications, Chongqing 400065, P. R. China

Abstract: To address the limitation of semantic reconstruction performance in traditional covert communications under low signal-to-noise ratio (SNR) conditions, a semantic covert communication scheme integrating reconfigurable intelligent surfaces (RIS) and a tripartite generative adversarial network (GAN) was proposed in this work. First, the physical layer and the semantic layer were jointly modeled, and an optimization problem was formulated to maximize semantic covert fidelity. Then, a hierarchical optimization framework was developed. At the physical layer, the transmit power and RIS phase shifts were jointly optimized under Kullback–Leibler (KL) divergence-based covert constraints to construct a channel environment favorable for signal transmission. Based on this design, a semantic-oriented tripartite adversarial training mechanism was introduced at the semantic layer, and end-to-end learning was enabled to jointly optimize semantic recovery performance and covertness. Experimental results show that, the proposed scheme achieves a 29.8% improvement in BLEU score compared with the best benchmark scheme, while the eavesdropper's detection accuracy remains close to 0.5.

Keywords: covert communication, semantic communication, reconfigurable intelligent surface, generative adversarial learning, KL divergence

收稿日期：2026–01–21；**修回日期：**2026–04–15
通信作者：沙婵，18195556232@163.com
基金项目：重庆市自然科学基金项目资助（NO.CSTB2025NSCQ-LZX0051）
Foundation Items：Natural Science Foundation of Chongqing，China(NO.CSTB2025NSCQ-LZX0051)

0 引言

随着无线通信技术的演进，数据传输规模和内容丰富度持续提升，使得传输数据中涉及的隐私与敏感信息大幅增加，通信安全面临严峻挑战。传统的防御体系主要依赖加密认证技术[1]和物理层安全技术[2]，其核心目标在于防止窃听者解码机密内容。然而，上述方法难以掩盖通信行为本身的存在性。一旦信号被检测到，窃听者仍可能利用隐私重建攻击[3]等手段恢复敏感信息。为克服这一局限，隐蔽通信[4-5]作为一种更高层级的防御机制应运而生。该技术通过优化设计传输信号特性，将通信信号特征淹没于背景噪声之中，从而显著降低窃听者的检测概率[6-[8]，实现了比传统加密和物理层安全更为彻底的安全保障。

近年来，语义通信[9]凭借其高效的语义表征与卓越的抗干扰能力，成为通信领域的研究热点。不同于关注比特精确传输的传统范式，语义通信通过提取并传输数据的高维语义特征，即使在低信噪比（signal-to-noise ratio, SNR）条件下仍能保持显著的鲁棒性[10-11]。这一特性使其天然适用于隐蔽通信所要求的低功率、低可检测性场景，由此催生了语义隐蔽通信这一新兴研究方向。语义隐蔽通信的核心目标是在确保语义信息有效传递的同时，将传输行为彻底隐匿。现有研究已对此展开初步探索，文献[12]引入生成式人工智能与协作干扰技术以增强隐蔽性；文献[13]构建隐蔽可靠语义通信框架，通过联合优化语义编码与传输机制实现隐匿传输；文献[14]针对文本语义传输提出有限长隐蔽方案，并联合优化发射功率、人工噪声和语义符号映射率，提升了语义谱效率。

然而，现有的语义隐蔽通信研究多基于理想信道建模来展开[14-[16]，这使得系统在实际信道中面临多径衰落与复杂环境干扰时，难以同时保障语义传输的有效性和隐蔽性。在此背景下，可重构智能表面（reconfigurable intelligent surface, RIS）为解决该问题提供了新的可能。RIS 可利用方向性增益与零陷效应重构无线传播环境[17-18]，从而增强合法链路增益并抑制窃听信道质量。已有研究进一步探索了 RIS 在隐蔽通信中的应用。例如，文献[19]利用大规模同时透射与反射可重构智能表面产生的不可预测性构建环境不确定性，显著增强隐蔽通信性能；文献[20]提出一种 RIS 辅助的分组合作隐蔽通信方案，通过解决非凸优化问题，实现了复杂多用户下隐蔽通信的性能提升；文献[21]设计了双功能 RIS，通过切换中继与反向散射模式以混淆检测者。与此同时，RIS 也被引入语义通信研究中。文献[22]通过联合设计 RIS 波束赋形与语义压缩机制，有效提升工业物联网系统的语义感知传输速率与频谱效率。文献[23]利用 RIS 构建优化信道并联合优化语义编码，在功耗受限场景下，通过端到端训练实现了比传统系统更优的语义接收质量与鲁棒性。

为提升隐蔽性能，部分研究引入生成对抗网络（generative adversarial network, GAN）实现隐蔽通信。其优势在于 GAN 的生成器与判别器之间的极小极大博弈机制，和隐蔽通信中发射端与窃听端之间的信号拟合与判别模式具有天然的同构性。通过对抗训练，生成器能够迫使发射信号的概率分布逐渐逼近信道噪声，以实现统计意义上的不可区分性。相关研究中，文献[24]提出了一种基于 GAN 的功率分配策略，实现了隐蔽速率与检测错误概率之间的有效折衷；文献[25]构建了三方对抗模型，通过发射机与检测机之间的动态竞争，生成具备欺骗性的隐蔽波形；文献[26]开发了模型驱动的对抗优化框架，通过联合设计无人机轨迹与发射功率来提升隐蔽性能。然而，现有基于对抗学习的隐蔽通信方案多依赖于信号特征的经验性统计拟合，缺乏与信息论隐蔽性度量之间的联系；另一方面，这些研究大多停留在传统通信层面，还没有实现基于语义级别的对抗训练。

针对上述挑战，本文提出一种 RIS 和三方对抗网络辅助的语义隐蔽通信优化方案。该方案利用面向语义信息的三方对抗机制，在发射端提取高维语义特征并映射为具备隐蔽特性的波形信号。在物理传输层，联合优化发射功率与 RIS 相移矩阵，实现波束的智能调控，保障接收端高保真语义重构的同时，显著降低窃听者的检测概率。特别地，不同于已有方案[25]，本文提出的语义隐蔽生成器由语义编码器与信号生成器级联构成，是一种能够融入语义特征的网络结构。该生成器一方面用于生成满足隐蔽性要求的物理信号，另一方面保留能够支持语义恢复的高层语义特征。由此，本文的三方 GAN 优化策略从传统的单一物理层分布匹配，扩展为物理层与语义层的跨层联合优化。本文的主要贡献如下。

1）提出了一种基于 RIS 和对抗网络的语义隐蔽通信方案。该方案将语义通信引入隐蔽通信场景，在 RIS 的辅助下联合设计物理信道调控与语义信息表征，并利用生成对抗学习将发射方、接收方与窃听方建模为三方博弈系统，实现语义可靠性与隐蔽性的协同优化；

2）构建并求解了隐蔽语义保真度最大化的跨层资源分配问题。该问题同时涉及物理层与语义层变量，是一个带有发射功率约束、RIS 单元相移约束、物理层隐蔽性约束以及语义层网络参数约束的非凸耦合优化问题。为此提出一种跨层优化方法，将物理层资源分配与语义层对抗训练进行解耦；其中，物理层通过变量替换与等价变换重构优化问题，并结合交替优化与半定松弛（semidefinite relaxation, SDR）技术求解满足隐蔽约束的最优解，语义层部分则在三方对抗网络框架下进行联合训练。

3）仿真结果验证了所提方案的有效性。在相同参数配置下，相较于基准方案与消融实验，所提方案在低 SNR 条件下的语义恢复性能和隐蔽传输性能均表现出显著优势。

1　系统模型

1.1　场景描述

本文考虑一个由 RIS 和三方 GAN 联合设计的语义隐蔽通信框架，该系统由发射端 Alice、RIS 辅助信道、合法接收端 Bob 以及窃听端 Willie 四部分组成，如图 1 所示。Alice 的目标是在 RIS 的辅助下，向 Bob 发送携带语义信息的隐蔽信号，同时规避 Willie 的检测。其中，Alice、Bob 和 Willie 均配备单天线，RIS 由 N 个无源反射单元组成，通过调节各单元的相移来重构无线传输环境。为保证研究的普适性与公平性，Bob 和 Willie 置于相同的无线传播环境中。此外，假设文中所有信道均服从准静态平坦衰落模型。为便于理论分析，本文假设发送端可获得完美信道状态信息（channel state information, CSI）。在实际系统中，RIS 相关 CSI 通常通过导频训练与级联信道估计获取，但当 RIS 单元规模较大时，会引入额外的训练与反馈开销。鉴于本文重点在于语义隐蔽通信方案设计，故在系统模型中采用理想 CSI 假设以突出方法本身性能。

在该框架中，RIS 与三方 GAN 协同优化系统性能。RIS 通过调节反射单元相位改变等效信道，提高合法接收端的信噪比，同时抑制窃听端信号强度，从而稳定语义特征表示，降低恢复损失并提升三方 GAN 训练梯度稳定性。与此同时，三方 GAN 训练过程中，发送端根据监听者判别反馈调整信号统计特性，使其接近背景噪声分布。这会影响合法信道的接收性能及窃听端信号分布，因此 RIS 相移矩阵需随信号更新同步优化，以进一步提升系统性能。

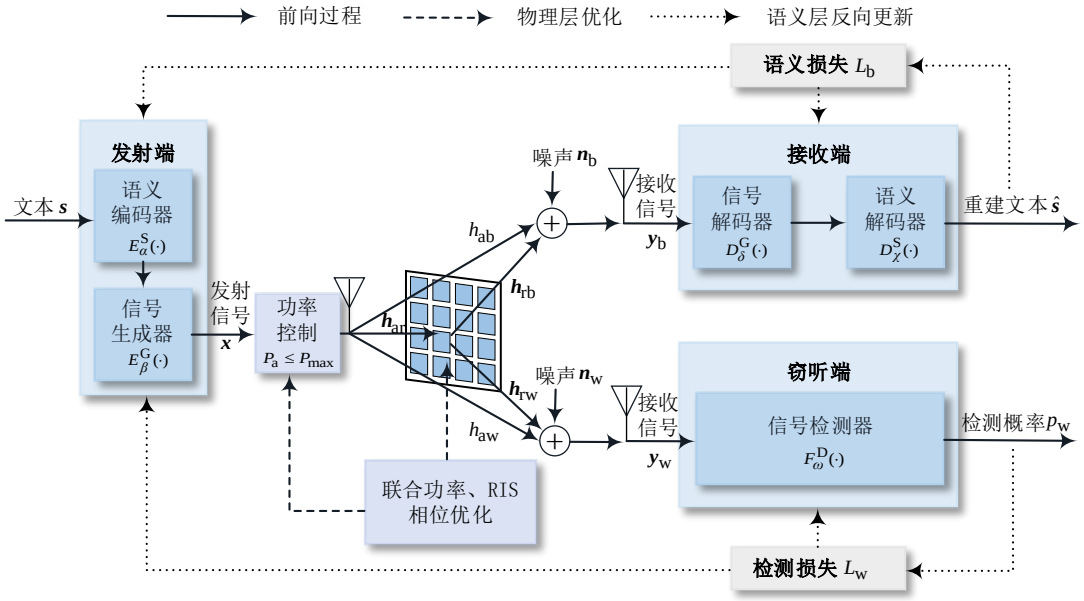


图1　系统架构

1.2　隐蔽发收发模型

发送端 Alice 的目的是实现语义特征的提取与隐蔽映射。发送端将处理后的文本数据 $\boldsymbol{s}$ 分批次输入语义编码器网络 $E_{\alpha}^{\mathrm{S}}(\cdot)$，提取出高维语义特征向量 $\boldsymbol{m}$，可表示为

$$\boldsymbol{m} = E_{\alpha}^{\mathrm{S}}(\boldsymbol{s}) \tag{1}$$

随后，$\boldsymbol{m}$ 被送入信号生成器网络 $E_{\beta}^{\mathrm{G}}(\cdot)$。该网络负责将语义特征映射为适合无线传输的波形，同时通过对抗训练使输出信号在统计特性上逼近背景噪声。生成的隐蔽发射信号 $\boldsymbol{x}$ 可表示为

$$\boldsymbol{x} = E_{\beta}^{\mathrm{G}}(\boldsymbol{m}) \tag{2}$$

合法接收端 Bob 的目的是从含噪接收信号中恢复原始文本。接收信号 $\boldsymbol{y}_{\mathrm{b}}$ 先经过信号解码器 $D_{\delta}^{\mathrm{G}}(\cdot)$，得到估计的语义特征向量 $\hat{\boldsymbol{m}}$，可表示为

$$\hat{\boldsymbol{m}} = D_{\delta}^{\mathrm{G}}(\boldsymbol{y}_{\mathrm{b}}) \tag{3}$$

接着，语义解码器 $D_{\chi}^{\mathrm{S}}(\cdot)$ 根据语义特征向量重构文本 $\hat{\boldsymbol{s}}$，可表示为

$$\hat{\boldsymbol{s}} = D_{\chi}^{\mathrm{S}}(\hat{\boldsymbol{m}}) \tag{4}$$

1.3　RIS 辅助信道与信号传输

为有效对抗信道衰落的影响，本文构建了 RIS 辅助的通信系统模型。Alice 与接收端之间的通信链路包含直连链路与 RIS 级联链路。具体地，定义 Alice 到 RIS、Bob 以及 Willie 的等效基带信道向量分别为 $\boldsymbol{h}_{\mathrm{ar}} \in \mathbb{C}^{N\times 1}$，$h_{\mathrm{ab}}$，$h_{\mathrm{aw}}$；RIS 到 Bob 和 Willie 的信道向量分别为 $\boldsymbol{h}_{\mathrm{rb}} \in \mathbb{C}^{N\times 1}$，$\boldsymbol{h}_{\mathrm{rw}} \in \mathbb{C}^{N\times 1}$。RIS 反射系数矩阵 $\boldsymbol{\Theta} = \mathrm{diag}(\beta_1 \mathrm{e}^{\mathrm{j}\theta_1}, \beta_2 \mathrm{e}^{\mathrm{j}\theta_2}, \cdots, \beta_N \mathrm{e}^{\mathrm{j}\theta_N})$，其中，$\beta_n = 1$ 和 $\theta_n \in [0, 2\pi)$ 分别表示第 n 个 RIS 单元的振幅和相移，$\forall n \in \{1,2,\cdots,N\}$。经过直连链路和 RIS 反射链路，Bob 的接收信号为

$$\boldsymbol{y}_{\mathrm{b}} = \sqrt{P_{\mathrm{a}}}\,(\boldsymbol{h}_{\mathrm{rb}}^{\mathrm{H}}\boldsymbol{\Theta}\boldsymbol{h}_{\mathrm{ar}} + h_{\mathrm{ab}})\boldsymbol{x} + \boldsymbol{n}_{\mathrm{b}} \tag{5}$$

其中，P_{a} 是 Alice 的发射功率，$\boldsymbol{n}_{\mathrm{b}} \sim \mathcal{CN}(0, \sigma_{\mathrm{b}}^2)$ 表示均值为 0，方差为 σ_b^2 的复加性高斯白噪声。Bob 的接收 SNR 可表示为

$$\gamma_{\mathrm{b}} = \frac{P_{\mathrm{a}}\left|\boldsymbol{h}_{\mathrm{rb}}^{\mathrm{H}}\boldsymbol{\Theta}\boldsymbol{h}_{\mathrm{ar}} + h_{\mathrm{ab}}\right|^2}{\sigma_{\mathrm{b}}^2} \tag{6}$$

1.4　窃听者模型与隐蔽性分析

Willie 的目标是通过其接收信号检测 Alice 是否正在进行通信。Alice 的发射信号经由直连链路及 RIS 级联链路传播后到达 Willie 处，Willie 面临一个典型的二元假设检验问题，其接收信号可表示为

$$\begin{cases} \mathcal{H}_0 : \boldsymbol{y}_{\mathrm{w}} = \boldsymbol{n}_{\mathrm{w}} \\ \mathcal{H}_1 : \boldsymbol{y}_{\mathrm{w}} = \sqrt{P_{\mathrm{a}}}\,(\boldsymbol{h}_{\mathrm{rw}}^{\mathrm{H}}\boldsymbol{\Theta}\boldsymbol{h}_{\mathrm{ar}} + h_{\mathrm{aw}})\boldsymbol{x} + \boldsymbol{n}_{\mathrm{w}} \end{cases} \tag{7}$$

其中，$\mathcal{H}_0$ 表示 Alice 与 Bob 未进行通信的原假设，$\mathcal{H}_1$ 表示 Alice 与 Bob 进行通信的备择假设。$\boldsymbol{n}_{\mathrm{w}} \sim \mathcal{CN}(0, \sigma_{\mathrm{w}}^2)$ 表示均值为 0，方差为 σ_{w}^2 的复加性高斯白噪声。

Willie 检测性能通常由总检测错误概率（DEP, detection error probability）衡量，表达式为

$$\xi = \pi_0 P_{\mathrm{FA}} + \pi_1 P_{\mathrm{MD}} \tag{8}$$

其中，π_0 和 π_1 分别表示假设 $\mathcal{H}_0$ 和 $\mathcal{H}_1$ 的先验概率，$P_{\mathrm{FA}} = \Pr\{\mathcal{D}_1 | \mathcal{H}_0\}$ 和 $P_{\mathrm{MD}} = \Pr\{\mathcal{D}_0 | \mathcal{H}_1\}$ 分别对应虚警概率和漏检概率。本文假设先验等概，即 $\pi_0 = \pi_1 = 0.5$，该设定在隐蔽通信相关研究[27-28]中已被广泛采用。

文献[5]给出了 Willie 的最优检测阈值及相应的最小检测错误概率 ξ^*。然而，ξ^* 表达式中包含不完全 Gamma 函数，直接处理 ξ^* 非常困难[6]。根据文献[29]，对于最优检测，最小检测错误概率 ξ^* 与总变差距离 $\mathcal{V}_T(\mathbb{P}_0, \mathbb{P}_1)$ 存在如下关系

$$\xi^* = 1 - \mathcal{V}_T(\mathbb{P}_0, \mathbb{P}_1) \tag{9}$$

其中，$\mathbb{P}_0$ 和 $\mathbb{P}_1$ 分别表示假设 $\mathcal{H}_0$ 和假设 $\mathcal{H}_1$ 下窃听者观测向量的概率分布，总变差距离定义为

$$\mathcal{V}_T(\mathbb{P}_0, \mathbb{P}_1) = \frac{1}{2}\left\| p_0(\boldsymbol{y}_{\mathrm{w}}) - p_1(\boldsymbol{y}_{\mathrm{w}}) \right\|_1 \tag{10}$$

其中，$p_0(\boldsymbol{y}_{\mathrm{w}})$ 和 $p_1(\boldsymbol{y}_{\mathrm{w}})$ 分别是 $\mathbb{P}_0$ 和 $\mathbb{P}_1$ 的概率质量函数。可见总变差距离为 Willie 的假设检验误差提供了下界，然而分析观测变量时需要用到概率分布的乘积形式，且计算总变差距离需要在高维空间上积分，往往难以得到闭式表达式，给后续求解带来复杂性。为此，引入 Pinsker 不等式，其表达式为

$$\mathcal{V}_T(\mathbb{P}_0, \mathbb{P}_1) \leqslant \sqrt{\frac{1}{2}\mathcal{D}(\mathbb{P}_0 \| \mathbb{P}_1)} \tag{11}$$

其中，$\mathcal{D}(\mathbb{P}_0 \| \mathbb{P}_1)$ 是 $\mathbb{P}_0$ 到 $\mathbb{P}_1$ 的 KL 散度，在准静态平坦衰落信道条件下，$\mathcal{D}(\mathbb{P}_0 \| \mathbb{P}_1)$ 可表示为

$$\mathcal{D}(\mathbb{P}_0 \| \mathbb{P}_1) = L \ln\left(1 + \frac{P_{\mathrm{a}}\left|\boldsymbol{h}_{\mathrm{rw}}^{\mathrm{H}}\boldsymbol{\Theta}\boldsymbol{h}_{\mathrm{ar}} + h_{\mathrm{aw}}\right|^2}{\sigma_{\mathrm{w}}^2}\right) - \frac{LP_{\mathrm{a}}\left|\boldsymbol{h}_{\mathrm{rw}}^{\mathrm{H}}\boldsymbol{\Theta}\boldsymbol{h}_{\mathrm{ar}} + h_{\mathrm{aw}}\right|^2}{P_{\mathrm{a}}\left|\boldsymbol{h}_{\mathrm{rw}}^{\mathrm{H}}\boldsymbol{\Theta}\boldsymbol{h}_{\mathrm{ar}} + h_{\mathrm{aw}}\right|^2 + \sigma_{\mathrm{w}}^2} \tag{12}$$

其中，L 是信道使用次数。

将 Pinsker 不等式带入 ξ^* 的表达式，可得到最小检测错误概率 ξ^* 的下界

$$\xi^* \geqslant 1 - \sqrt{\frac{1}{2}\mathcal{D}(\mathbb{P}_0 \| \mathbb{P}_1)} \tag{13}$$

为使隐蔽约束 $\xi^* \geqslant 1 - \epsilon$ 成立，其中 ϵ 为允许泄露的隐蔽性能等级，仅需约束 KL 散度满足

$$\mathcal{D}(\mathbb{P}_0 \| \mathbb{P}_1) \leqslant 2\epsilon^2 \tag{14}$$

在实际对抗场景中，窃听者往往并不局限于传统的能量检测器，而可能采用一些智能检测策略。为此，本文引入基于深度神经网络（deep neural network, DNN）的检测器来模拟 Willie 的检测行为。该检测器以 Willie 接收到的信号序列为输入，输出通信行为存在的概率，其映射关系可表示为

$$p_{\mathrm{w}} = F_{\omega}^{\mathrm{D}}(\boldsymbol{y}_{\mathrm{w}}) \tag{15}$$

其中，$F_{\omega}^{\mathrm{D}}(\cdot)$ 表示参数为 ω 的神经网络模型。设定判决阈值为 $V = 0.5$，当 $p_{\mathrm{w}} \geqslant V$ 时，Willie 判定存在通信行为，否则判定仅存在噪声。

本文在系统设计中同时引入物理层与语义层的双重隐蔽性约束。物理层由 KL 散度来约束隐蔽性，限制最优似然比检测下窃听者的最坏情况检测能力，从而在信息论意义上保证信号统计不可区分。语义层将基于 DNN 的检测器视为复杂检测策略的可学习近似，通过三方对抗训练逼近上述不可区分状态。因此，在 KL 散度约束下，即便窃听者采用更复杂的学习模型，其检测性能仍受物理层统计不可区分性的根本限制。

2　方案设计

2.1　基于三方 GAN 的端到端网络设计

为缓解隐蔽通信中语义传输高效性与隐蔽性之间的内在矛盾，本文构建了一种基于三方对抗机制的端到端网络架构。该系统由三类功能网络组成：语义隐蔽发射机（Alice）、语义恢复接收机（Bob）以及隐蔽信号检测器（Willie）。在联合训练过程中，Alice 与 Bob 通过最小化语义重构误差实现语义信息的有效传输；Willie 作为对抗方，从统计判别角度对 Alice 生成的隐蔽信号进行检测约束。各模块的具体网络结构设计见图 2。

如图 2(a)所示，语义隐蔽发射机 Alice 由语义编码器与信号生成器级联构成，用于提取文本的语义特征，并将其映射为适合信道传输的隐蔽信号。

语义编码器采用 Transformer 编码器结构，通过自注意力机制捕获长距离语义依赖并生成全局语义表示。输入文本序列经词嵌入与位置编码后映射为 128 维特征矩阵，再通过 3 层由多头自注意力、前馈网络、残差连接和层归一化组成的编码器提取深层语义特征。信号生成器采用四层全连接网络结构，将语义特征映射为低检测性的连续信道符号。其中前三层引入 ReLU 激活函数以增强非线性表示能力，最后一层保持线性输出以生成连续信道符号。随后对生成信号进行功率归一化处理，以消除深度网络输出幅度不确定性对训练稳定性的影响。同时，由于信号生成器输入为语义特征的连续表示且输出信号保持可微分性，接收端的信号解码器能通过端到端学习构建近似逆映射，从而保持通信可靠性。定义语义隐蔽发射机的损失函数为

$$L_{\mathrm{a}} = \lambda_1 L_{\mathrm{b}} + \lambda_2 L_{\mathrm{cov}} \tag{16}$$

其中，λ_1 是 Bob 的语义重构损失权重，λ_2 是 Willie 的检测损失权重，L_{b} 是 Bob 的语义重构损失，L_{cov} 是隐蔽对抗损失，表达式为 $L_{\mathrm{cov}} = -\log(1 - F_{\omega}^{\mathrm{D}}(\boldsymbol{y}_{\mathrm{w}}))$。可以看出网络不是基于简单的 GAN，而是基于语义驱动的信号生成机制。

如图 2(b)所示，语义恢复接收机 Bob 由信号解码器和语义解码器组成，对应于 Alice 的逆过程，负责在接收端从含噪信号中恢复原始文本，保证高可靠的语义重构。

信号解码器采用四层全连接网络，实现物理层接收信号到语义特征空间的逆映射。该模块以经过信道衰落与噪声扰动的接收信号为输入，通过全连接网络将信号进行特征升维，并使用 ReLU 激活函数以增强非线性拟合能力，随后通过残差连接和层归一化抑制信道噪声和非线性变换带来的幅度波动，使恢复后的语义特征在统计分布上与 Transformer 解码器期望的输入空间保持一致。语义解码器采用 3 层 Transformer 解码器结构，每层由带掩码的自注意力、编码器-解码器注意力和前馈网络组成，用于融合语义上下文并逐步重建原始文本序列。最后通过线性预测层将 128 维特征矩阵映射到目标词表空间，以计算预测词的概率分布。语义恢复接收机使用交叉熵损失函数来最小化发送文本数据 $\boldsymbol{s}$ 和估计文本数据 $\hat{\boldsymbol{s}}$ 之间的差异，可以表示为

$$\begin{aligned} L_{\mathrm{b}}(\boldsymbol{s},\hat{\boldsymbol{s}}) = &-\sum_{k=1} q(s_k)\log(p(s_k)) \\ &+(1 - q(s_k))\log(1 - p(s_k)) \end{aligned} \tag{17}$$

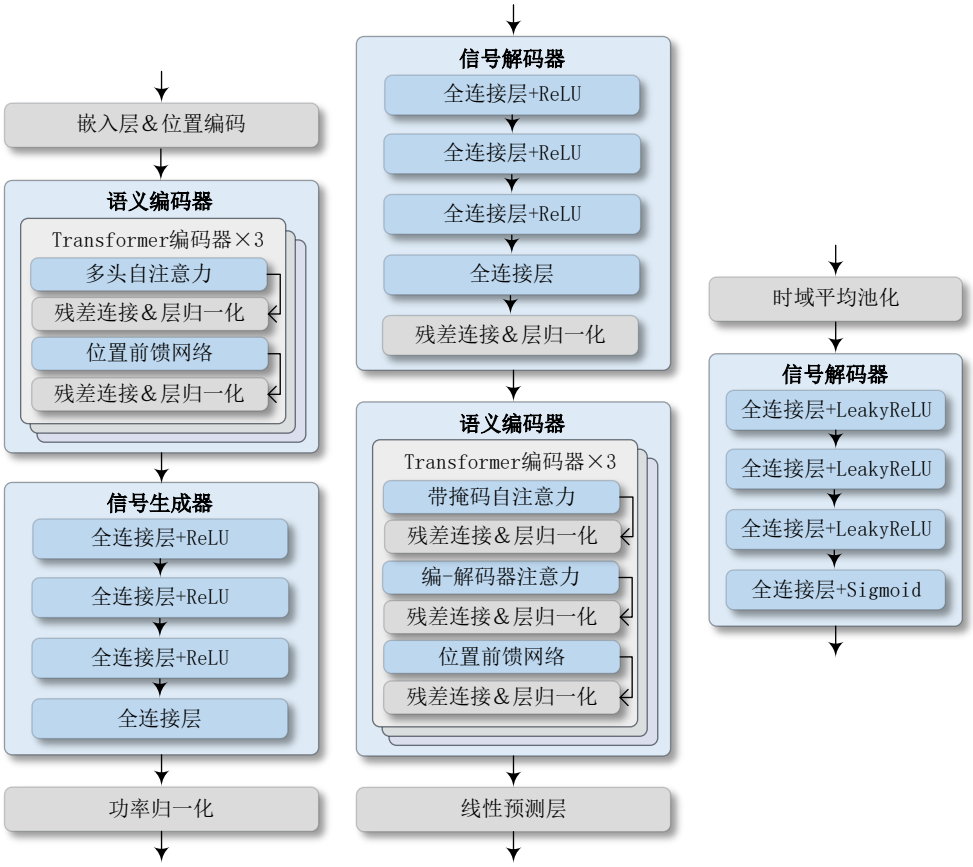


图2　三方GAN的内部结构设计

(a)语义隐蔽发射机 (b)语义恢复接收机 (c)隐蔽信号检测器

其中，$q(s_k)$是输入文本$\boldsymbol{s}$中第k个单词s_k出现的真实概率，$p(s_k)$是估计文本$\hat{s}$中第k个单词s_k出现的预测概率。

如图2(c)所示，隐蔽信号检测器Willie被设计为轻量级二分类器，根据接收到信号判断是否存在通信行为。考虑到隐蔽信号可能隐藏在长序列噪声中，首先对接收信号在时间维度上执行平均池化，以消除时间序列的瞬时波动，提取全局统计特征。随后通过四层全连接网络，其中前三层使用Leaky-ReLU激活函数，以避免梯度消失并增强对微弱信号特征的捕捉能力；输出层使用Sigmoid激活函数，最后输出当前信号存在通信行为的概率值$p_{\mathrm{w}}\in[0,1]$。隐蔽信号检测器选择二元交叉熵作为其损失函数，该损失函数驱使判别器准确区分含通信信号的接收信号和纯噪声信号。其表达式为

$$L_{\mathrm{w}}=-\log F_{\omega}^{\mathrm{D}}(\boldsymbol{y}_1)-\log\left(1-F_{\omega}^{\mathrm{D}}(\boldsymbol{y}_0)\right) \tag{18}$$

其中，$F_{\omega}^{\mathrm{D}}(\boldsymbol{y}_1)$是含通信信号的接收信号的判别器输出结果，$F_{\omega}^{\mathrm{D}}(\boldsymbol{y}_0)$是仅含纯噪声的接收信号的判别器输出结果。隐蔽信号检测器的目标是最小化此损失函数，这意味着在有隐蔽通信时最大化检测概率，同时在无通信时最小化虚警概率。通过这种对抗性训练，检测器能够不断进化，迫使发送端Alice必须生成更具伪装性的信号以维持通信的隐蔽性。

文献[24]指出，当窃听方采用基于DNN的检测器时，检测统计量与判决边界可通过训练过程隐式学习得到，因此在检测阶段无需显式计算最优检测门限，仅需将Sigmoid输出结果与固定阈值0.5进行比较即可完成判决。

2.2　物理层和语义层的联合优化问题设计

为降低信道衰落对接收恢复和窃听检测的影响，本文联合设计物理层与对抗语义层。物理层通过优化发射功率和RIS相移，在满足隐蔽约束、功率约束和RIS反射系数约束的前提下最大化Bob的接收信噪比；语义层通过对抗训练提升合法接收方的数据重构质量，并抑制窃听方的检测能力。

为表征该方案在物理层和语义层的整体性能，

本文引入隐蔽语义保真度（covert semantic fidelity, CSF）这一跨层性能指标。需要说明的是，CSF在本文中仅作为一种概念性的性能度量指标，用于统一描述系统的跨层优化目标，而不限定为某一固定形式的数学函数。在后续优化问题建模过程中，本文将进一步分解为物理层与语义层两个可优化子目标，并分别进行求解。

因此，所建立的联合优化问题可以描述为

$$\begin{aligned}\text{P1:}\quad &\max_{P_a,\boldsymbol{\Theta},\alpha,\beta,\delta,\chi,\omega} && \mathrm{CSF}(\gamma_\mathrm{b},L_\mathrm{a},L_\mathrm{b},L_\mathrm{w})\\ &\text{s.t.} && \mathrm{C1}{:}\mathcal{D}(\mathbb{P}_0\|\mathbb{P}_1)\leqslant 2\epsilon^2\\ & && \mathrm{C2}{:}0\leqslant P_\mathrm{a}\leqslant P_\mathrm{max}\\ & && \mathrm{C3}{:}0\leqslant\theta_n\leqslant 2\pi,\forall n=1,2,\cdots,N\\ & && \mathrm{C4}{:}\alpha,\beta,\delta,\chi,\omega\text{都是超参数}\end{aligned}\tag{19}$$

可以看出，求解P1需要对语义层网络参数和物理层发射功率、RIS相移进行联合设计。语义层和物理层的跨层优化问题在一定程度上具有解耦特性，即物理层侧重在隐蔽约束下最大化接收用户的信号强度；语义层侧重于最小化损失函数以优化网络参数。约束条件中，C1是隐蔽约束，涉及发射功率和RIS相移，C2是发射功率约束，C3是RIS相移约束，C4是网络超参数的约束。其中C1、C2和C3是物理层约束，C4是语义层约束，物理层约束和语义层约束分别作用于不同参数空间，在约束层面可分离。因此所构建的跨层优化问题P1可以分解为两个子问题，即物理层C1、C2和C3约束下，最大化隐蔽接收信号功率γ_b；语义层C4约束下，最小化损失函数。在此基础上，提出一个分层训练框架，将在后续小结具体描述。

尽管如此，物理层与语义层仍存在相互作用影响，即物理层资源配置影响语义传输性能，语义模型训练效果反过来影响物理信道的输入分布。为降低问题求解复杂度，本文采用一种分层迭代近似求解的跨层设计方法。具体而言，在固定语义模型参数下优化物理层资源，随后在更新后的信道条件下训练语义层网络参数。通过多轮迭代可实现物理层与语义层的协同优化，近似解决原始跨层联合优化问题。

2.3　RIS辅助物理层信道优化

为了增强物理层信道传输，需要在隐蔽约束下，联合设计发送端发射功率和RIS被动相移，以最大化隐蔽接收端的接收SNR。将物理层优化从全局联合函数中解耦得到子问题$\mathrm{CSF}(\gamma_\mathrm{b},L_\mathrm{a},L_\mathrm{b},L_\mathrm{w})=\gamma_\mathrm{b}$，具体描述为

$$\begin{aligned}\text{P2:}\quad &\max_{P_\mathrm{a},\boldsymbol{\Theta}} && \frac{P_\mathrm{a}\left|\boldsymbol{h}_\mathrm{rb}^\mathrm{H}\boldsymbol{\Theta}\boldsymbol{h}_\mathrm{ar}+h_\mathrm{ab}\right|^2}{\sigma_\mathrm{b}^2}\\ &\text{s.t.} && \mathrm{C1}{:}\mathcal{D}(\mathbb{P}_0\|\mathbb{P}_1)\leqslant 2\epsilon^2\\ & && \mathrm{C2}{:}0\leqslant P_\mathrm{a}\leqslant P_\mathrm{max}\\ & && \mathrm{C3}{:}0\leqslant\theta_n\leqslant 2\pi,\forall n=1,2,\cdots,N\end{aligned}\tag{20}$$

目标函数中的隐蔽接收端噪声功率为常数，因此优化过程中可以省略。此外，目标函数和隐蔽约束的优化变量相互耦合，且RIS反射系数为单位模约束，导致该问题难以直接求解，因此需要对其进行化简。

窃听方Willie处的接收SNR为

$$\gamma_\mathrm{w}=\frac{P_\mathrm{a}\left|\boldsymbol{h}_\mathrm{rw}^\mathrm{H}\boldsymbol{\Theta}\boldsymbol{h}_\mathrm{ar}+h_\mathrm{aw}\right|^2}{\sigma_\mathrm{w}^2}\tag{21}$$

对于隐蔽约束C1，由式(12)可知，$\mathcal{D}(\mathbb{P}_0\|\mathbb{P}_1)$是$\gamma_\mathrm{w}$的单调递增函数。因此，隐蔽约束C1可等效重写为$\gamma_\mathrm{w}\leqslant\bar{\gamma}_\mathrm{w}$，其中$\bar{\gamma}_\mathrm{w}$是方程

$$L\left[\ln\left(1+\bar{\gamma}_\mathrm{w}\right)-\frac{\bar{\gamma}_\mathrm{w}}{1+\bar{\gamma}_\mathrm{w}}\right]=2\epsilon^2\tag{22}$$

的解。令$f(\gamma_\mathrm{w})=L\left[\ln(1+\gamma_\mathrm{w})-\frac{\gamma_\mathrm{w}}{1+\gamma_\mathrm{w}}\right]-2\epsilon^2$，其中$\gamma_\mathrm{w}\geqslant 0$。对于任意给定的$\epsilon\geqslant 0$和正整数$L$，有$f(0)=-2\epsilon^2\leqslant 0$，$\lim_{\gamma_\mathrm{w}\to\infty}f(\gamma_\mathrm{w})=\infty$，由连续函数介值定理，存在某个$\bar{\gamma}_\mathrm{w}\geqslant 0$使得$f(\bar{\gamma}_\mathrm{w})=0$，存在性得证。进一步，求导得$f'(\gamma_\mathrm{w})=\frac{L\gamma_\mathrm{w}}{(1+\gamma_\mathrm{w})^2}$，当$\gamma_\mathrm{w}>0$时，$f'(\gamma_\mathrm{w})>0$，且$f'(0)=0$，因此$f(\gamma_\mathrm{w})$在$[0,\infty)$上严格单调递增。结合存在性与单调性可知，方程$f(\gamma_\mathrm{w})=0$存在唯一解$\bar{\gamma}_\mathrm{w}\geqslant 0$。由于$f(\gamma_\mathrm{w})$无闭式反函数，本文采用数值方法求解。具体地，使用二分法在区间$[0,\gamma_\mathrm{max}]$内搜索$\bar{\gamma}_\mathrm{w}$，其中γ_max满足$f(\gamma_\mathrm{max})>0$。

考虑到优化变量$\boldsymbol{\Theta}$为对角矩阵，不易处理，定义$\boldsymbol{u}=[u_1,u_2,\cdots,u_N]^\mathrm{H}$，其中$u_n=\mathrm{e}^{\mathrm{j}\theta_n},\forall n=1,2,\cdots,N$，$\boldsymbol{a}=\mathrm{diag}(\boldsymbol{h}_\mathrm{rw}^\mathrm{H})h_\mathrm{ar}$，$\boldsymbol{b}=\mathrm{diag}(\boldsymbol{h}_\mathrm{rb}^H)h_\mathrm{ar}$。因此，优化问题P2可重写为

$$\begin{aligned}\text{P3:}\quad &\max_{P_{\mathrm{a}},\boldsymbol{u}} \quad P_{\mathrm{a}}\left|\boldsymbol{u}^{\mathrm{H}}\boldsymbol{b}+h_{\mathrm{ab}}\right|^2\\ &\text{s.t.}\quad \text{C1:}P_{\mathrm{a}}\left|\boldsymbol{u}^{\mathrm{H}}\boldsymbol{w}+h_{\mathrm{aw}}\right|^2\leqslant\bar{\gamma}_{\mathrm{w}}\sigma_{\mathrm{w}}^2\\ &\quad\ \ \text{C2:}0\leqslant P_{\mathrm{a}}\leqslant P_{\max}\\ &\quad\ \ \text{C3:}\left|u_n\right|=1,\forall n=1,2,\cdots,N\end{aligned}\tag{23}$$

尽管原问题已简化为P3形式，但该问题的目标函数包含非线性表达式，直接求解较困难。为了便于处理，引入松弛变量$t\in\mathbb{R}$，满足$|t|^2=1$，用于将非线性目标函数转换为一个二次型目标函数[30]。该二次型结构能够通过矩阵表示转化为标准的优化问题形式，从而便于采用SDR方法进行求解。同时，令

$$\boldsymbol{v}=\begin{bmatrix}\boldsymbol{u}\\ t\end{bmatrix},\boldsymbol{B}=\begin{bmatrix}\boldsymbol{b}\boldsymbol{b}^{\mathrm{H}} & \boldsymbol{b}h_{\mathrm{ab}}^{*}\\ h_{\mathrm{ab}}\boldsymbol{b}^{\mathrm{H}} & \left|h_{\mathrm{ab}}\right|^2\end{bmatrix},\boldsymbol{A}=\begin{bmatrix}\boldsymbol{a}\boldsymbol{a}^{\mathrm{H}} & \boldsymbol{a}h_{\mathrm{aw}}^{*}\\ h_{\mathrm{aw}}\boldsymbol{a}^{\mathrm{H}} & \left|h_{\mathrm{aw}}\right|^2\end{bmatrix}\tag{24}$$

则优化问题P3可重写为

$$\begin{aligned}\text{P4:}\quad &\max_{P_{\mathrm{a}},\boldsymbol{v}} \quad P_{\mathrm{a}}\boldsymbol{v}^{\mathrm{H}}\,\mathbf{B}\mathbf{v}\\ &\text{s.t.}\quad \text{C1:}P_{\mathrm{a}}\boldsymbol{v}^{\mathrm{H}}\,\mathbf{A}\mathbf{v}\leqslant\bar{\gamma}_{\mathrm{w}}\sigma_{\mathrm{w}}^2\\ &\quad\ \ \text{C2:}0\leqslant P_{\mathrm{a}}\leqslant P_{\max}\\ &\quad\ \ \text{C3:}\left|v_n\right|=1,\forall n=1,2,\cdots,N+1\end{aligned}\tag{25}$$

若$\boldsymbol{v}$是问题P4的解，则$\boldsymbol{v}/t$是原问题的最优解。因此，原优化问题的解可通过求解P4得到。然而，P4中仍无法处理单位模约束。为了解决单位模约束，令$\boldsymbol{M}=\boldsymbol{v}\boldsymbol{v}^{\mathrm{H}}$，且满足$\boldsymbol{M}\succcurlyeq 0$，$\mathrm{rank}(\boldsymbol{M})=1$，则优化问题P4被转换为

$$\begin{aligned}P5:\quad &\max_{P_a,\boldsymbol{M}} \quad P_a\mathrm{Tr}(\mathbf{B}\mathbf{M})\\ &\text{s.t.}\quad C1:P_a\mathrm{Tr}(\mathbf{A}\mathbf{M})\leqslant\bar{\gamma}_w\sigma_w^2\\ &\quad\ \ C2:0\leqslant P_a\leqslant P_{\max}\\ &\quad\ \ C3:\boldsymbol{M}\succcurlyeq 0\\ &\quad\ \ C4:\mathrm{rank}(\boldsymbol{M})=1\\ &\quad\ \ C5:M_{n,n}=1,\forall n=1,2,\cdots,N+1\end{aligned}\tag{26}$$

优化问题P5可以采用交替优化算法求解。首先给定P_a，松弛秩1约束，优化问题P5可转化为

$$\begin{aligned}P6:\quad &\max_{\boldsymbol{M}} \quad P_a\mathrm{Tr}(\mathbf{B}\mathbf{M})\\ &\text{s.t.}\quad C1:P_a\mathrm{Tr}(\mathbf{A}\mathbf{M})-\bar{\gamma}_w\sigma_w^2\leqslant 0\\ &\quad\ \ C2:\boldsymbol{M}\succcurlyeq 0\\ &\quad\ \ C3:M_{n,n}=1,\forall n=1,2,\cdots,N+1\end{aligned}\tag{27}$$

优化问题P6中，目标函数是线性目标函数，约束是凸约束集，则P6是一个SDR问题，可以采用CVX工具箱获得其最优解。

接下来给定$\boldsymbol{M}$，优化问题化简为

$$\begin{aligned}\text{P7:}\quad &\max_{P_{\mathrm{a}}} \quad P_{\mathrm{a}}\mathrm{Tr}(\mathbf{B}\mathbf{M})\\ &\text{s.t.}\quad \text{C1:}P_{\mathrm{a}}\mathrm{Tr}(\mathbf{A}\mathbf{M})-\bar{\gamma}_{\mathrm{w}}\sigma_{\mathrm{w}}^2\leqslant 0\\ &\quad\ \ \text{C2:}0\leqslant P_{\mathrm{a}}\leqslant P_{\max}\end{aligned}\tag{28}$$

上述优化问题为线性规划问题，可以得到最优解为$P_{\mathrm{a}}{}^{*}=\min\left\{\dfrac{\bar{\gamma}_{\mathrm{w}}\sigma_{\mathrm{w}}^2}{\mathrm{Tr}(\mathbf{A}\mathbf{M})},P_{\max}\right\}$。

接下来对上述交替优化算法进行收敛性分析。令$f(P_{\mathrm{a}},\boldsymbol{M})=P_{\mathrm{a}}\mathrm{Tr}(\mathbf{B}\mathbf{M})$表示目标函数。在每次迭代中，先固定$P_{\mathrm{a}}{}^{(t)}$优化$\boldsymbol{M}$，即求解式(27)的SDR问题得到$\boldsymbol{M}^{(t+1)}$，由于该子问题为凸问题，其最优解满足$f(P_{\mathrm{a}}{}^{(t)},M^{(t+1)})\geqslant f(P_{\mathrm{a}}{}^{(t)},M^{(t)})$；接着固定$M^{(t+1)}$优化$P_{\mathrm{a}}$，即求解式(28)的线性规划问题得到$P_{\mathrm{a}}{}^{(t+1)}$，其最优解满足$f(P_{\mathrm{a}}{}^{(t+1)},M^{(t+1)})\geqslant f(P_{\mathrm{a}}{}^{(t)},M^{(t+1)})$。因此满足$f(P_{\mathrm{a}}{}^{(t+1)},M^{(t+1)})\geqslant f(P_{\mathrm{a}}{}^{(t)},M^{(t)})$，即一轮迭代后目标函数序列单调不减。同时，受限于发射功率上限$P_{\max}$和RIS最大反射增益，接收功率存在上界。由单调有界收敛定理，序列必收敛至一个驻点，即满足KKT条件的局部最优解。

2.4 基于三方GAN的语义层优化训练

为了优化语义层的网络训练，需要最小化损失函数来更新网络参数。其优化子问题可以表述为

$$\begin{aligned}\text{P8:}\quad &\min_{\alpha,\beta,\delta,\chi,\omega} \quad \mathrm{CSF}(L_{\mathrm{a}},L_{\mathrm{b}},L_{\mathrm{w}})\\ &\text{s.t.}\quad \text{C4:}\alpha,\beta,\delta,\chi,\omega\text{都是超参数}\end{aligned}\tag{29}$$

优化问题P8旨在通过调整一组超参数$\{\alpha,\beta,\delta,\chi,\omega\}$来最小化三方对抗神经网络的三个损失函数。三方对抗网络被建模为一个非合作博弈，其目标是寻找系统的纳什均衡解。

对于接收机，利用交叉熵与KL散度之间的关系，其损失函数的最小化问题可重写为

$$\begin{aligned}&\min L_{\mathrm{b}}=\min H(p(s)\|q(\hat{s}))=\\ &\min(\mathrm{KL}(p(s)\|q(\hat{s})))+H(p(s))\end{aligned}\tag{30}$$

其中，$p(s)$表示原始数据的分布，$q(\hat{s})$表示重构数据的分布，$H(p(s)\|q(\hat{s}))$是$p(s)$和$q(\hat{s})$之间的交叉熵，$\mathrm{KL}(p(s)\|q(\hat{s}))$是$p(s)$和$q(\hat{s})$之间的的KL散度。信号检测器的损失函数最小化问题可表示为

$$\begin{aligned}&\min L_{\mathrm{w}}=\\&\min\left(-\mathrm{E}_{y\sim\mathbb{P}_0}(\log(F_\omega^{\mathrm{D}}(\boldsymbol{y})))-\mathrm{E}_{y\sim\mathbb{P}_0}(\log(1-F_\omega^{\mathrm{D}}(\boldsymbol{y})))\right)=\\&\max\mathrm{E}_{y\sim\mathbb{P}_1}(\log(F_\omega^{\mathrm{D}}(\boldsymbol{y})))+\mathrm{E}_{y\sim\mathbb{P}_1}(\log(1-F_\omega^{\mathrm{D}}(\boldsymbol{y})))]\end{aligned}\tag{31}$$

发射机的对抗损失函数最小化问题可表示为

$$\begin{aligned}&\min L_{\mathrm{cov}}=\min\left(-\mathrm{E}_{y\sim\mathbb{P}_0}(\log(1-F_\omega^{\mathrm{D}}(\boldsymbol{y})))\right)=\\&\max\left(-\mathrm{E}_{y\sim\mathbb{P}_1}(\log(F_\omega^{\mathrm{D}}(\boldsymbol{y})))\right)\end{aligned}\tag{32}$$

结合式(30)、式(31)和式(32)，三方 GAN 的优化过程等效于求解一个极大极小值问题[25]，即

$$\begin{aligned}&\min_A\min_B\max_W V(A,B,W)=\\&\min_A\min_B\max_W\lambda_1(\mathrm{KL}(p(s)\|q(\hat{s})))+H(p(s))+\\&\lambda_2(\mathrm{E}_{y\sim\mathbb{P}_1}(\log(F_\omega^{\mathrm{D}}(\boldsymbol{y})))+\mathrm{E}_{y\sim\mathbb{P}_0}(\log(1-F_\omega^{\mathrm{D}}(\boldsymbol{y}))))\end{aligned}\tag{33}$$

当发射机与接收机网络参数固定时，三方 GAN 中信号检测网络的最优解等价于经典 GAN 中的判别器最优形式，即

$$F_\omega^{\mathrm{D}^*}(\boldsymbol{y})=\frac{\mathbb{P}_1}{\mathbb{P}_0+\mathbb{P}_1}\tag{34}$$

将式(34)带入式(33)可得

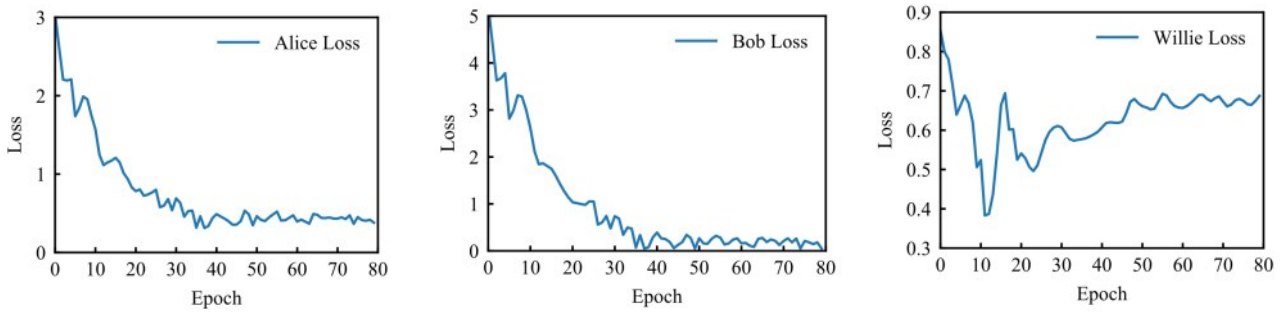
$$\begin{aligned}&V(A,B)=\max_W V(A,B,W)=\lambda_1(\mathrm{KL}(p(s)\|q(\hat{s})))+\\&H(p(s))+2\lambda_2(\mathrm{JS}(\mathbb{P}_0\|\mathbb{P}_1)-\log 2)\end{aligned}\tag{35}$$

其中，$\mathrm{JS}(\mathbb{P}_0\|\mathbb{P}_1)=\frac{1}{2}\mathrm{KL}(\mathbb{P}_0\|m)+\frac{1}{2}\mathrm{KL}(\mathbb{P}_1\|m)$，$m(y)=\frac{1}{2}\mathbb{P}_0+\frac{1}{2}\mathbb{P}_1$。式(35)中 $H(p(s))$ 和 $\log 2$ 是常数，JS 散度和 KL 散度都大于或等于 0，因此最小化 JS 散度和 KL 散度可以得到 $V(A,B)$ 的最小值。当且仅当 JS 散度和 KL 散度等于 0 时取等号，即隐蔽信号分布与噪声信号分布一致，同时接收端能够无失真恢复原始语义信息。在神经网络容量足够大时，上述优化过程可收敛至全局最优解。实验结果如图 3 所示，训练初期损失函数不断波动，但随着迭代进行逐渐平稳，证实了算法的收敛性。

算法 1 跨层端到端训练步骤

1) **输入：**文本数据集 K，衰落信道系数 $\boldsymbol{h}_{\mathrm{ar}},h_{\mathrm{ab}},h_{\mathrm{aw}},\boldsymbol{h}_{\mathrm{rb}},\boldsymbol{h}_{\mathrm{rw}}$，高斯噪声 $\boldsymbol{n}_{\mathrm{b}},\boldsymbol{n}_{\mathrm{w}}$
2) **初始化：**$\boldsymbol{\Theta}$，P_{a} 以及网络参数 $\alpha,\beta,\delta,\chi,\omega$
3) **while 停止条件不满足时，do**
4) 交替优化算法解决 P_2，得到 P_{a} 和 $\boldsymbol{\Theta}$
5) **前向传播：**
6) Alice 生成隐蔽发射信号：$E_\alpha^{\mathrm{S}}(\boldsymbol{s})\rightarrow\boldsymbol{m}$，$E_\beta^{\mathrm{G}}(\boldsymbol{m})\rightarrow\boldsymbol{x}$
7) RIS 辅助信道传输，得到 Bob 接收信号 $\boldsymbol{y}_{\mathrm{b}}$ 和 Willie 接收信号 $\boldsymbol{y}_{\mathrm{w}}$
8) **更新三方网络超参数：**
9) Willie 计算检测概率：$F_\omega^{\mathrm{D}}(\boldsymbol{y}_{\mathrm{w}})\rightarrow p_{\mathrm{w}}$
10) 式(18)计算判别损失 L_{w}，梯度下降更新 ω
11) Bob 解码信号：$D_\delta^{\mathrm{G}}(\boldsymbol{y}_{\mathrm{b}})\rightarrow\hat{\boldsymbol{m}}$，$D_\chi^{\mathrm{S}}(\hat{\boldsymbol{m}})\rightarrow\hat{\boldsymbol{s}}$
12) 式(17)计算语义重构损失 L_{b}，梯度下降更新 χ，δ
13) 式(16)计算语义隐蔽损失 L_{a}，梯度下降更新 β，α
14) **end while**
15) **输出：**神经网络 $E_\alpha^{\mathrm{S}}(\cdot),E_\beta^{\mathrm{G}}(\cdot),D_\delta^{\mathrm{G}}(\cdot),D_\chi^{\mathrm{S}}(\cdot),F_\omega^{\mathrm{D}}(\cdot)$

所提出的跨层端到端训练流程如算法 1 所示。其复杂度主要来自 RIS 辅助的波束形成优化和融合语义编码的三方 GAN 训练。因此，算法 1 的整体实现复杂度为 $O(I(N+1)^{4.5}+L_{seq}^2\cdot K)$[23]，其中，$I$ 为物理层迭代次数，L_{seq} 为句子长度，K 为每个单词所需要的符号数量。实验结果表明，在 NVIDIA A100 GPU 上训练模型时，通常可在 50 轮次达到收敛，训练时间约为 39.16 分钟。完成训练后，在线通信阶段仅需进行一次前向推理，其延迟约为 4.01



(a) Alice　(b) Bob　(c) Willie

图 3　三方 GAN 收敛图

毫秒。

3 仿真分析

3.1 仿真设计

为验证本文提出的基于RIS和三方GAN的语义隐蔽通信方案的性能，设计以下仿真内容。

实验采用欧洲议会会议记录的英文文本数据集，将其划分为训练集与测试集。为保证训练稳定性与计算效率，对输入文本序列长度进行处理，限制在[4,30]词区间，超出部分进行截断，不足部分采用<PAD>标识符填充。训练与测试阶段均采用相同的数据预处理策略。语义通信网络主体包含3层Transformer编码器与解码器，每层配置8个注意力头；信号生成器、解码器及检测器均由全连接层构建。经多次不同权重比例的调试实验后，发现等权重配置能够获得最优的综合性能，因此将语义损失权重系数λ_1与隐蔽损失权重系数λ_2均设定为0.5。构建三维笛卡尔坐标系，Alice、RIS、Bob和Willie的空间坐标分别设定为(0，0，0)m、(50，0，10)m、(55，0，0)m和(50，10，0)m，RIS单元数量默认设为128个，其余信道仿真参数参考文献[6]设置。

为验证所提方案的有效性，本文选取了采用/不采用RIS辅助的传统分离式编解码方案（Huffman+ Turbo）作为基准对比。同时为了量化本文三方对抗网络带来的隐蔽性能提升，实验还引入了未部署对抗机制的传统通信方案和语义通信方案作为消融对比。各基准方案的详细设定如表1所示。实验分别在AWGN信道和瑞利衰落信道环境下进行性能评估。最后，本文采用BLEU分数衡量接收端的语义恢复质量，并采用信号检测器的检测准确率评估系统隐蔽性能。

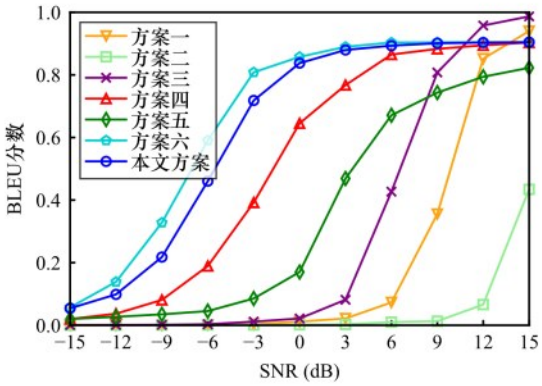
3.2 仿真结果与分析

图4是各方案在AWGN信道和瑞利信道中的BLEU分数对比情况。结果表明，本文方案在全SNR区间内的性能仅次于方案六。这是因为所引入的对抗隐蔽机制为了满足严格的隐蔽性要求，会对语义特征施加扰动并进行非线性映射，从而在一定程度上影响语义信息的完整恢复。除此之外，本文方案均优于其他方案，特别是在低SNR区域下，当传统比特级通信方案（方案一、方案二和方案三）几乎无法工作时，本文方案仍能保持较高的鲁棒性。这得益于RIS提供的波束赋形增益对受限发射功率的补偿，同时对抗机制驱动的语义通信模型能够在保证隐蔽的同时，利用上下文冗余在强噪声干扰下提取核心语义特征。值得注意的是，由于瑞利信道的多径效应导致信号产生随机的幅度和相位畸变，增加了语义重构的难度，使得各方案在瑞利信道下的性能均低于AWGN信道。

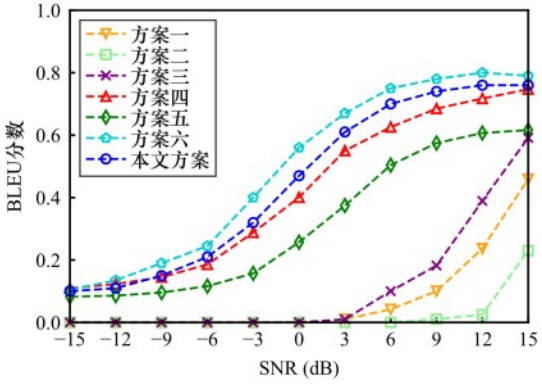
图5是各方案对应信号检测器的检测准确率。可以看出，采用三方GAN机制的方案在全SNR区域内的检测准确率始终保持在0.5左右，接近随机猜测水平。相比之下，仅依赖物理层隐蔽约束的方案六虽能具有一定隐蔽性，但其检测准确率随SNR增长的趋势明显高于采用三方GAN机制的方案；而未引入对抗机制且缺乏物理层隐蔽约束的方案一和方案四上升更为明显。这说明三方GAN生成的隐蔽信号能有效掩盖语义特征，从而降低被监测概率。对比方案三和本文方案可知，语义通信方案下的检测准确率更低，这说明语义通信可以有效降低数据冗余，进而降低信号的可检测性。对比方案五和本文方案可以发现，RIS辅助可通过抑制窃听方向信号进一步提升隐蔽性能。此外，瑞利信道下的检测准确率略低于AWGN信道，这是因为随机衰

表1 基准方案说明

方案名称	三方对抗机制	RIS	通信类型	备注
方案一	-	-	传统通信	传统通信
方案二	√	-	传统通信	传统通信+对抗隐蔽
方案三	√	√	传统通信	传统通信+对抗隐蔽+RIS增强
方案四[31]	-	-	语义通信	语义通信
方案五	√	-	语义通信	语义通信+对抗隐蔽
方案六	-	√	语义通信	语义通信+RIS增强
本文方案	√	√	语义通信	语义通信+对抗隐蔽+RIS增强



(a)　AWGN 信道



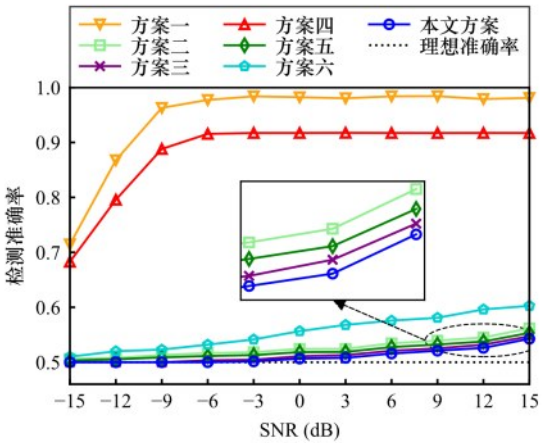
(b)　瑞利信道

图 4　AWGN 信道和瑞利信道下不同方案的隐蔽语义恢复性能对比

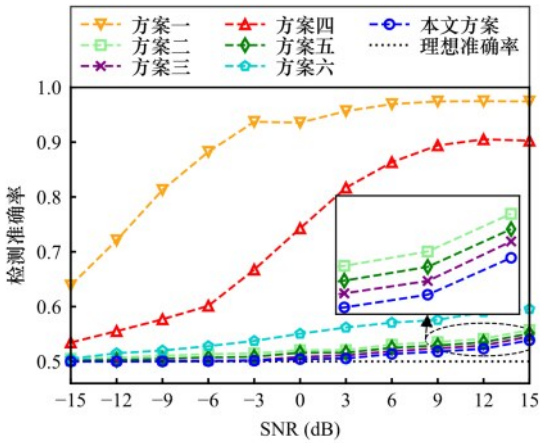
落特性充当了天然的掩护，使得 Willie 难以区分接收功率的波动是由信号传输还是信道变化引起的，从而提升了系统的隐蔽性能。

为进一步评估所提方案的性能，采用基于 BERT 模型的语义相似度（BERT-SS, BERT semantic similarity）作为评价指标，分数越高表明输入句子与恢复句子的语义相似度越高。图 6 是各方案在瑞利信道中的 BERT-SS 评估分数对比情况。结果表明，BERT-SS 指标结果与 BLEU 分数的变化趋势基本一致，进一步验证了本文方案的有效性。

图 7 为 AWGN 信道和瑞利信道下，不同 RIS 反射单元数量对隐蔽语义恢复性能的影响对比。从图 7 可以看出，所提方案的语义恢复质量随着 RIS 反射单元数量 N 增加而增强。以 SNR=0dB 为例，在 BLEU(1-grams)指标下，如图 7(a)所示，AWGN 信道中 N=16 的 RIS 方案对应 BLEU 分数仅为 0.25，而高配 RIS 方案（N=128）的 BLEU 分数提升至 0.83；瑞利信道中则由 0.28 提升至 0.47。这表明大规模反射阵列可提供更高的无源波束成形增益，有效增强



(a)　AWGN 信道



(b)　瑞利信道

图 5　AWGN 信道和瑞利信道下不同方案的窃听者检测准确率对比

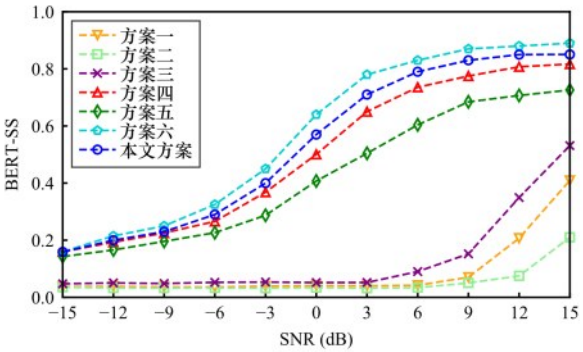
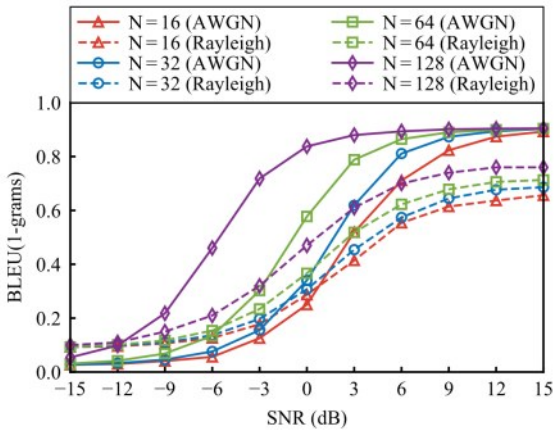


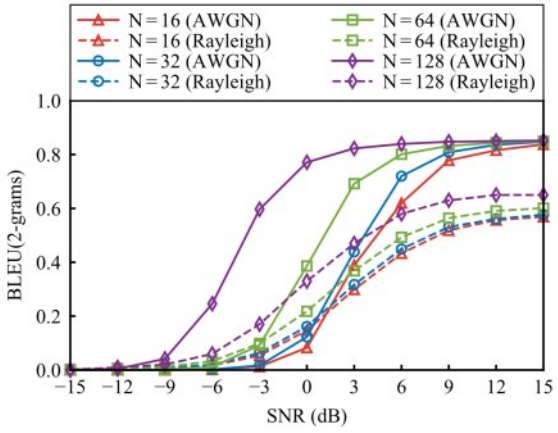
图 6　瑞利信道下不同方案的 BERT-SS 评估分数对比

接收处的信号强度。值得注意的是，在 N=16 时，瑞利信道的结果略优于 AWGN 信道。这主要是因为丰富的多径分量为 RIS 提供了额外的空间多样性增益，从而表现出微弱的优势。但随着 RIS 规模 N 的增大，RIS 在 AWGN 信道下的能量汇聚作用远高于在瑞利信道下对随机多径的补偿作用，因此

BLEU 分数提高幅度更大。



(a)　BLEU(1-grams)

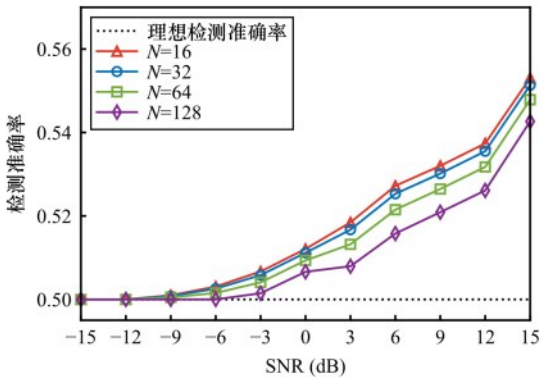


(b)　BLEU(2-grams)

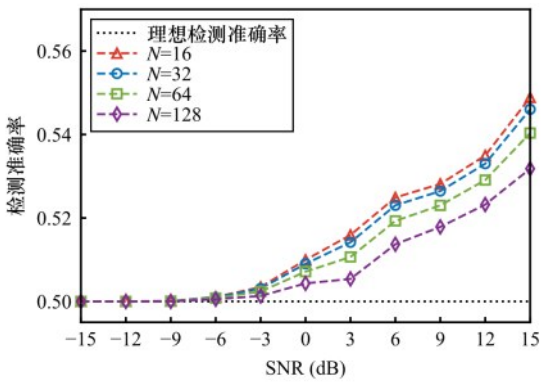
图 7　不同信道和不同 RIS 反射单元数量下的 BLEU 分数

图 7(b)是考虑单词对之间的语义匹配效果。随着 n-grams 中 n 阶数的增大，评价指标对接收信息的准确性与连贯性要求更为严苛，导致 BLEU 分数相较 1-grams 均有所下降，但 RIS 辅助方案始终表现出明显的优势。此外，对比两种信道可知，AWGN 信道的性能在多数情况下优于瑞利信道。这是因为瑞利信道的多径衰落特性引入了信号幅度和相位畸变，破坏了语义特征的结构完整性，进而加剧语义解码的错误率。

图 8 为 AWGN 信道和瑞利信道下，不同 RIS 反射单元数量对窃听者检测准确率的影响对比。整体来看，在全 SNR 范围下均能维持接近 0.5 的检测正确率。具体来看，在低 SNR 区间（-15dB 至 0dB），检测准确率更接近理想隐蔽边界。这是物理层 KL 散度隐蔽约束和语义层对抗生成机制的协同作用结果，迫使窃听者退化至随机猜测状态。随着 SNR 的升高，信号特征逐渐显著，导致检测准确率呈现上升趋势，但增加 RIS 元件数量 N 可有效降低准确率。这得益于 RIS 对窃听者方向的波束零陷作用。此外，受随机衰落影响，瑞利信道下的检测准确率整体略低于 AWGN 信道。



(a)　AWGN 信道



(b)　瑞利信道

图 8　不同信道和不同 RIS 反射单元数量下的检测准确率

4　结束语

考虑到无线环境下语义传输可靠性与物理层隐蔽性之间的权衡关系，即增强隐蔽性可能降低语义恢复性能，而提升语义恢复质量又可能增加检测风险，本文提出了一种基于 RIS 和三方对抗网络辅助的语义隐蔽通信方案。通过联合优化发射功率与 RIS 相移，并采用基于生成对抗网络的跨层训练机制，该系统能有效提升低 SNR 下的语义 BLEU 分数，同时将窃听者的检测准确率抑制在随机猜测水平。然而，所提方案仅基于理想 CSI 与准静态平坦衰落信道假设，未来研究可进一步引入信道估计误差与反馈开销建模，并结合动态时变信道特性设计

自适应资源分配策略，以提高实际系统可部署性。

参考文献：

[1] Lee J W, Kang H C, Lee Y W, et al. Privacy-preserving machine learning with fully homomorphic encryption for deep neural network[J]. IEEE Access, 2022, 10: 30039-30054.

[2] Zhao N, LI Y X, Zhang S, et al. Security enhancement for NOMA-UAV networks[J]. IEEE Transactions on Vehicular Technology, 2020, 69(4): 3994-4005.

[3] Xu X R, Liu P R, Wang W, et al. CGIR: conditional generative instance reconstruction attacks against federated learning[J]. IEEE Transactions on Dependable and Secure Computing, 2023, 20(6): 4551-4563.

[4] Chen X Y, An J P, Xiong Z H, et al. Covert communications: a comprehensive survey[J]. IEEE Communications Surveys & Tutorials, 2023, 25 (2): 1173-1198.

[5] Yan S H, He B, Zhou X Y, et al. Delay-intolerant covert communications with either fixed or random transmit power[J]. IEEE Transactions on Information Forensics and Security, 2019, 14(1): 129-140.

[6] Zhou X B, Yan S H, Wu Q Q, et al. Intelligent reflecting surface (IRS)-aided covert wireless communications with delay constraint[J]. IEEE Transactions on Wireless Communications, 2022, 21(1): 532-547.

[7] Liu Z P, Li X, Ji H, et al. Exploiting STAR-RIS for covert communication in ISAC networks under imperfect CSI[J]. IEEE Transactions on Vehicular Technology, 2025, 74(1): 786-802.

[8] Zhang Z, Yang L, Lei H J, et al. Covert communication in RSMA-assisted ambient backscatter communication systems[J]. IEEE Transactions on Wireless Communications, 2025, 24(8): 6566-6579.

[9] Getu T M, Kaddoum G, and Bennis M. Semantic communication: a survey on research landscape, challenges, and future directions[J]. Proceedings of the IEEE, 2024, 112(11): 1649-1685.

[10] Peng X, Qin Z J, Tao X M, et al. A robust image semantic communication system with multi-scale vision transformer[J]. IEEE Journal on Selected Areas in Communications, 2025, 43(4): 1278-1291.

[11] Hu L L, Yu L S, and Qin Z J. Deep learning-based semantic communication system for wireless image transmission[J]. IEEE Wireless Communications Letters, 2025, 14(8): 2391-2395.

[12] Du H Y, Liu G Y, Niyato D, et al. Generative AI-aided joint training-free secure semantic communications via multi-modal prompts[C]// Proceedings of ICASSP 2024 - 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Piscataway: IEEE Press, 2024: 12896-12900.

[13] Xu R, Li G L, Yang Z H, et al. Covert and reliable semantic communication against cross-layer privacy inference over wireless edge networks[C]//Proceedings of 2024 IEEE Wireless Communications and Networking Conference (WCNC). Piscataway: IEEE Press, 2024: 1-6.

[14] Hu J S, Ye L J, Chen Y G, et al. Covert communications for text semantic with finite blocklength[J]. IEEE Wireless Communications Letters, 2024, 13(10): 2842-2846.

[15] Bounhar A, Sarkiss M, and Wigger M. Unveiling covert semantics: joint source-channel coding under a covertness constraint[C]//Proceedings of GLOBECOM 2024 - 2024 IEEE Global Communications Conference. Piscataway: IEEE Press, 2025: 25-30.

[16] Zhang W J, Hu Y, Luo T, et al. Optimization of private semantic communication performance: an uncooperative covert communication method[J]. IEEE Transactions on Wireless Communications, 2025.

[17] Han Y, Tang W K, Jin S, et al. Large intelligent surface-assisted wireless communication exploiting statistical CSI[J]. IEEE Transactions on Vehicular Technology, 2019, 68(80): 8238 – 8242.

[18] Huang C W, Zappone A, Alexandropoukos G C, et al. Reconfigurable intelligent surfaces for energy efficiency in wireless communication[J]. IEEE Transactions on Wireless Communications, 2019, 18(8): 4157 – 4170.

[19] Li X W, Zhao J W, Chen G J, et al. STAR-RIS-assisted covert wireless communications with randomly distributed blockages[J]. IEEE Transactions on Wireless Communications, 2025, 24(6): 4690-4705.

[20] Lin S B, Ding G R, Wang H C, et al. Grouping enhanced cooperative covert communications in RIS-aided multi-user systems[J]. IEEE Transactions on Communications, 2025, 73(2): 995-1008.

[21] Li X W, Liu M S, Dang S P, et al. Covert communications with enhanced physical layer security in RIS-assisted cooperative networks[J]. IEEE Transactions on Wireless Communications, 2025, 24(7): 5605-5619.

[22] Zhao Z X, Yang Z H, Huang C W, et al. A Joint Communication and Computation Design for Distributed RIS-Assisted Probabilistic Semantic Communication in IIoT[J]. IEEE Internet of Things Journal, 2024, 11(16): 26568-26579.

[23] Ma J Y, Li Q, Liu R H, et al. Enhanced semantic information transfer on RIS-assisted communication systems[J]. IEEE Wireless Communications Letters, 2024, 13(8) 2225-2229.

[24] Liao X M, Si J B, Shi J, et al. Generative adversarial network assisted power allocation for cooperative cognitive covert communication system[J]. IEEE Communications Letters, 2020, 24(7): 1463-1467.

[25] 于季弘, 林子砚, 叶能, 等. 基于三方生成对抗网络的隐蔽通信方法[J]. 通信学报, 2023, 44(11): 225-236.
Yu J H, Lin Z Y, Ye N, et al. Covert communication method based on tripartite generative adversarial network[J]. Journal on Communications, 2023, 44(11): 225-236.

[26] Li Z, Liao X M, Shi J, et al. MD-GAN-based UAV trajectory and power optimization for cognitive covert communications[J]. IEEE Internet of Things Journal, 2022, 9(12): 10187-10199.

[27] Wang H M, Zhang Y, Zhang X, et al. Secrecy and covert communications against UAV surveillance via multi-hop networks[J]. IEEE Transactions on Communications, 2020, 68(1): 389-401.

[28] Zheng T X, Wang H M, Ng D W K, et al. Multi-antenna covert communications in random wireless networks[J]. IEEE Transactions on Wireless Communications, 2019, 18(3): 1974-1987.

[29] Bash B A, Goeckel D, and Towsley D. Limits of reliable communication with low probability of detection on AWGN channels[J]. IEEE Journal on Selected Areas in Communications, 2013, 31(9): 1921 – 1930.

[30] 周小波, 于辉, 彭旭, 等. 智能反射面辅助及人工噪声增强的无线隐蔽通信[J]. 电子与信息学报, 2022, 44(07): 2392-2399.
Zhou X B, Yu H, Peng X, et al. IRS-aided and artificial noise-enhanced wireless covert communications[J]. Journal of Electronics & Information Technology, 2022, 44(07): 2392 2399.

[31] Xie H Q, Qin Z J, Li G Y, et al. Deep learning enabled semantic communication systems[J]. IEEE Transactions on Signal Processing, 2021, 69: 2663-2675.

易印雪（1989– ），女，四川武胜人，博士，重庆邮电大学副教授、博士生导师，主要研究方向为网络信息传播与控制、语义通信、隐蔽通信。

沙婵（2002– ），女，宁夏中卫人，重庆邮电大学硕士生，主要研究方向为语义通信与隐蔽通信。

唐睿（1996– ），女，四川广安人，重庆邮电大学博士生，主要研究方向为无线移动通信理论与技术。

张祖凡（1972- ），男，湖北石首人，博士，重庆邮电大学教授，博士生导师，主要研究方向为无线移动通信理论与技术。