

语义通信研究现状与展望

◆ 张雪菲 古 静 李潇娴 崔琪楣 陶小峰

北京邮电大学信息与通信工程学院, 北京 100876

摘 要 随着6G超级无线宽带业务的涌现, 数据速率需求达到新的高峰。语义通信通过收发双方的知识共享和语义理解, 能够减少传输信息量, 降低带宽需求, 具有良好的应用前景。聚焦语义通信的编解码技术和端到端传输技术, 结合其业务特点和安全性需求, 重点阐述语义通信的高效编解码和传输技术。此外, 结合已有研究成果, 探讨语义通信存在的挑战并展望其未来发展方向。

关键词: 语义通信; 高效性; 鲁棒性

中图分类号: TN92 **文献标识码:** A

文章编号: 1009-2412(2023)04-0001-07

DOI: 10.3969/j.issn.1009-2412.2023.04.001

1 研究背景

在6G时代, 沉浸式无线扩展现实(extended reality, XR)、全息通信、感官互联和数字孪生等新技术与新需求不断涌现。上述主要技术场景具有沉浸化和智能化两大共同特征。预计2040年, 基于XR设备、全息设备等新型终端设备的沉浸式业务将快速发展, 月均流量将突破3万亿GB, 贡献超过一半的6G月均总流量^[1]。基于语法的传统通信通过增加传输维度、改变通信资源利用方式等, 已接近系统容量极限。而语义通信采用计算转换通信的思路, 通过收发双方的知识共享和语义理解, 减少传输信息量, 降低带宽需求, 可有效支持新型超级无线宽带业务的传输。

近些年, 语义通信作为6G移动通信的潜在关键技术之一, 得到了学术界和产业界的广泛关注。中

国国际移动通信IMT-2030(6G)推进组、欧盟奏鸣曲(SONATA)计划等都将语义通信视作一种打破跨域通信壁垒、提高传输效率的6G关键技术^[1-2]。华为发布的“后香农时代面向数学的十大挑战问题”中将语义通信的基础理论研究作为首要问题^[3]。在标准化方面, 国际电信联盟电信标准化部门(ITU-T)的第20研究组(SG20)和第13研究组(SG13)分别提出语义通信的国际标准技术报告《面向物联网和智慧城市/社区的语义通信架构》和《面向未来网络的语义认知网络需求》。中国通信标准化协会(CCSA)的无线通信技术工作委员会(TC5)也设立了基于语义通信的“智简通信关键技术研究”项目。IMT-2030(6G)推进组针对语义通信及语义认知网络架构研究进行了立项。

当前, 语义通信还处于较为初级的研究阶段, 其研究成果主要集中在利用深度学习完成语义的编解码与传输方面。图1为一个典型的端到端语义通信模型, 具体包括信源、发送端的语义编码器和信道编码器, 接收端的信道解码器和语义解码器, 以及公共知识库和私有知识库。在一些工作中, 语义编码与信道编码被深度网络建模为信源信道联合编码器, 同理信道解码与语义解码构成了联合解码器。

本文聚焦语义通信的编解码技术和端到端传输技术, 从业务特点角度, 探讨了面向单模态业务(如文本、音频和图像等)和多模态业务(如图文转换业务、沉浸式XR业务等)的端到端语义通信模型。从安全性角度, 讨论了针对深度学习和信道开放性的语义攻击或语义噪声, 以及其对语义通信性能的影响。此外, 基于现有研究进展, 探讨语义通信存在的挑战并展望其未来发展方向。

2 研究现状

语义通信的概念由Weaver和Shannon首次提出, 其旨在准确地传输与重建语义信息, 关注点是符号之上的语义信息内涵^[4]。近些年, 随着人工智能技

收稿日期: 2023-06-02 修回日期: 2023-06-20

项目资助: 国家重点研发计划青年科学家项目(2022YFB2902900)。

通信作者: 张雪菲, 副教授, zhangxuefei@bupt.edu.cn。

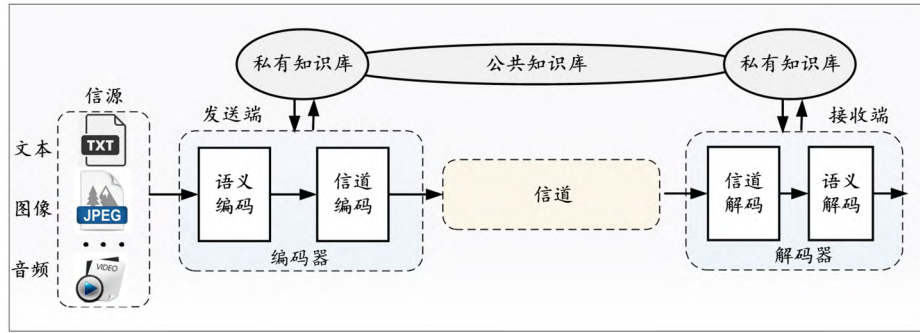


图1 端到端语义通信模型

术的快速发展以及算力的普及，语义通信再一次受到学术界和工业界的广泛关注。

2.1 单模态语义通信技术研究

语义通信中涉及的信源包括文本、音频、图像、视频和全息点云等各种模态。由于不同模态之间存在较大差异，目前学术界和工业界尚未出现广泛认可的语义信息归一化表征方式。因此，现有的大多数单模态语义通信系统仍侧重基于深度学习技术对特定模态展开相应的研究和探索。本文所述单模态语义通信的关键技术及性能特点如表1所示。

2.1.1 文本语义通信

随着基于深度学习的自然语言处理技术的快速发展，文本语义通信旨在恢复句子含义而不是传统通信中的抽象符号。Xie等^[5]提出了基于Transformer模型的联合信源信道编码方法来应对信道噪声和语义失真，在低信噪比下所提出的基于深度学习的语义通信系统（deep learning enabled semantic communication, DeepSC）比传统通信系统更具鲁棒性，且在BLEU分数与句子相似性上能实现更好的性能。为了使物联网设备能够负担得起DeepSC计算所需资源，Xie等^[6]

表1 单模态语义通信关键技术及性能特点

模态	关键技术	性能特点
文本	长短期记忆网络（long short-term memory, LSTM）、Transformer	成熟的语言模型；高鲁棒性
	知识图谱	复杂度高；精度要求更低
音频	循环神经网络（recurrent neural network, RNN）、基于注意力的SE-ResNet	高鲁棒性
	语音识别、语音合成	面向任务减少数据传输量
图像	深度学习信源信道联合编码（deep joint source-channel coding, DeepJSCC）	高鲁棒性
	非线性变换源信道编码（nonlinear transform source-channel coding, NTSCC）	可变长编码
	生成对抗网络（generative adversarial networks, GAN）	节省带宽
视频	基于深度学习的可变带宽无线视频传输（deep-learning-aided wireless video, DeepWiVe）、无线深度视频语义传输（wireless deep video semantic transmission, DVST）	减少数据传输量；高感知保真度
全息点云	AITransfer	显著减少数据传输量

提出了一种基于深度学习的轻量级分布式语义通信系统,命名为L-DeepSC。通过修剪模型冗余和降低权重分辨率,L-DeepSC显著降低了物联网设备与云/边缘之间传输模型权重所需的带宽。此外,为了减少衰落信道对传输的影响,Xie等^[6]还设计了一种利用信道状态信息辅助训练的方法。为了应对不同句子在不断变化的通信信道和信噪比上的语义差异,Zhou等^[7]基于通用Transformer引入自适应循环机制,打破了原始的固定结构。该方法通过不同的循环对不同的语义信息做出响应,从而使语义通信系统在处理不同的通信情况时更具适应性。为了充分理解传输数据背后的含义,Wang等^[8]设计了一种新的语义通信框架。该框架根据知识图谱将语义信息建模为“实体—关系—实体”的语义三元组形式,同时提出了一种语义相似性的数学度量指标,可从文本语义恢复的准确性和完整性两个方面进行衡量。不同于现有研究对显性语义信息的关注,Xiao等^[9]提出了一种新颖的基于推理的隐式语义感知通信网络架构。该架构允许用户学习语义推理机制,使其能够根据收到的有限信息自动生成复杂的隐式语义信息。

2.1.2 音频语义通信

对于音频信号的传输,Weng等^[10]设计了一个支持深度学习的语义通信系统DeepSC-S。该系统在特征提取的过程中,将基于注意力的SE-ResNet作为更重要的语音信息赋予更高权重,从而提高了语音信号恢复的准确性。此外,Weng等^[11]进一步开发了面向语音识别任务的语义通信系统DeepSC-SR,通过在发送端提取与文本相关的语义特征,并在接收端恢复文本转录,实现了在保持性能的同时减少数据传输量。为了同时考虑源信号恢复和智能任务,Han等^[12]提出了一种高度语义聚焦的支持语音到文本传输和语音到语音传输的通信系统,引入了基于注意力的软对齐模块和冗余去除模块,仅从语音频谱中提取与语义相关的信息进行传输。为了保证语音信号的准确恢复,对纹理化的语音特征相关信息进行了额外的提取。与现有方法相比,该方法在提高预测转录准确率和保持语音信号重建质量的同时,显著提高了传输效率。为了解决直接对脉冲编码调制信号进行编码时带来的带宽浪费问题,Xiao等^[13]整合了经典传输方案和深度学习的优势,通过引入非线性变化将语音信号源映射到语义潜在空间,并利用变分建模思想设置了一个熵模型来估计语义特征的重要性,根据重要性分配不同的编码速率,以此增大系统的编码增益。与现有的语音语

义传输方法相比,该方案在达到相同恢复质量的情况下,可节省高达75%的信道带宽成本。

2.1.3 图像语义通信

对于图像信源,Bourtsoulatz等^[14]提出了DeepJSCC方案。该方案不再使用传统通信中基于分离的显式信源和信道编码,而是将图像像素直接映射为复值信道的输入符号,从而有效避免了数字信号接收的“悬崖效应”。Yang等^[15]引入了策略网络,该网络可根据通道条件和图像内容动态控制编码速率。对于高信噪比场景或简单图像内容,能有效降低信道带宽利用率。然而,当图像编码向量的维数增加时,如传输大规模图像,DeepJSCC无法实现可变长编码,会使得系统性能迅速下降。因此,Dai等^[16]提出了NTSCC方案来突破上述限制。该方案运用非线性变换来提取图像源的语义特征,并结合自适应速率传输机制和超先验辅助编解码器细化机制进行编码传输。仿真结果表明,该方案可在各种感知指标上实现更好的性能,如峰值信噪比(peak signal-to-noise ratio, PSNR)、多尺度结构相似性指数(multi-scale structural similarity, MS-SSIM)、学习感知图像块相似度(learned perceptual image patch similarity, LPIPS)等。Huang等^[17]基于生成对抗网络设计了从粗到细的图像语义编码模型,通过生成器在基础层中生成保留语义信息,并使用BPG编码的残差在增强层中恢复图像的细节。此外,Huang等^[18]还提出了基于强化学习的自适应语义编码方法,定义类别、空间排列和视觉特征为图像数据的语义表示单元,接收端通过注意力模块融合局部和全局特征以实现图像重建。Lokumarambale等^[19]利用语义分割图像代替真实图像传输,并在接收端使用预训练的GAN网络生成预期图像,与传统通信相比节省了大量带宽。Dong等^[20]设计了基于层的图像语义通信系统。该系统可根据不同语义、精度和信道条件的要求,使用不同的语义切片模型(semantic slice models, SeSM)实现图像语义提取、传输和图像重建。

2.1.4 视频语义通信

除了文本、音频和图像模态,视频也逐渐成为多媒体业务中关键的组成部分。Tung等^[21]首次提出了DeepWiVe方案。该方案运用强化学习带宽分配策略来优化视频帧之间的带宽。仿真结果表明,DeepWiVe的多尺度结构相似性指标度量(multi-scale structural similarity index measure, MS-SSIM)性能在所有测试的通道条件下均优于H.264+LDPC编解码。

Wang 等^[22]设计了一种高效率信源信道联合编解码方案,实现了视频端到端的 DVST。DVST 利用非线性变换和条件编码架构进行视频语义特征提取,基于语义特征上的时间自适应熵模型提供的先验信息,为特征的潜在表示引入了可变长度传输机制。实验结果显示,与传统的无线视频编码传输方案相比,该方案可以更好地支持机器视觉任务,同时对人类视觉具有更高的感知保真度。Jiang 等^[23]提出了一个基于关键点传输的语义视频会议方案,人物的外貌图像作为知识被共享,视频仅利用关键点传输来表示面部表情的运动并进行恢复。为了应对视频会议系统中不同的信道环境,还设计了语义视频会议的增量冗余混合自动重复请求机制。该机制可以适应不同的误码率,并显著降低了对无线传输资源的要求。

2.1.5 全息点云语义通信

全息点云视频与传统视频服务相比,能够提供更沉浸式的观看体验,逐渐受到了越来越多的关注。Huang 等^[24]针对全息点云视频提出了一种 AI 驱动的带宽感知自适应传输技术,通过提取和传输关键点云特征,显著减少了数据传输量。此外,该技术将动态网络引入端到端网络的设计和训练中,并提供自适应传输控制算法来平衡时延和体验。Huang 等^[25]提出了一种面向全息点云视频的高效压缩和高精度解码重建技术。该技术不同于以往压缩点云的方法,而是提取了点云的关键特征并优化了点云下的采样过程以减少计算复杂度。同时,以用户体验质量为优化目标,又提出一种基于深度强化学习的自适应传输方法,根据移动设备和网络情况动态匹配最优云编码器—解码器模型以实现最佳的传输体验。

2.2 多模态语义通信技术研究

在缺乏相关上下文信息和背景知识的情况下,单模态信源的真实语义往往难以准确识别。此外,在无线信道中进行传输时,很难避免噪声的干扰,传输错误可能导致语义失真。为了应对单模态信源的多义性和歧义性,多模态语义通信系统应运而生。多模态语义通信是一项涉及多种数据模态的通信技术,通过将不同模态的数据(如图像、音频和文本等)相互关联和融合,能够充分利用语义信息,提供更丰富、准确和智能化的数据传输和应用。

2.2.1 图文语义通信

Xie 等^[26]建立了面向视觉问答任务的多模态语义通信系统,其中部分用户传输图像,部分用户传输文本以查询有关图像的信息。通过提取来自不同

发送端的图像和文本语义信息,接收端采用注意力机制和合成网络来处理接收到的语义信息,最后生成问题的预测答案。Xie 等^[27]还研究了基于深度学习的多用户语义通信系统,以图像检索、机器翻译和视觉问答 3 个智能任务作为语义通信系统的传输目标,采用基于 Transformer 的图像和文本语义编码结构,有效提取不同任务的语义信息。仿真结果表明,该系统在低信噪比和不完善的信道状态信息(channel state information, CSI)下性能优于传统通信系统。尽管如此,上述系统仅适用于特定任务,一旦任务发生改变,系统则需要随之更新。为了解决这个问题,Zhang 等^[28]开发了服务于各种传输任务的统一语义通信系统,采用领域自适应将不同任务的特征投影到对应的特征域中,以此降低传输开销。同时,为适应每个任务的不同难度要求,其提出了多出口架构,为相对简单的任务提供了早期出口结果。为了充分利用属性和关系知识对问题的不同响应,Chen 等^[29]提出了一种新的双语义知识组成的多模态对话系统,通过多层次的知识组合模块获得潜在的组合响应表示,并分别从组合响应表示和基本真值响应表示中提取语义信息,从而进行表示级语义正则化。

2.2.2 XR 语义通信

XR 是提高用户体验质量的前沿技术,可以分为增强现实(augmented reality, AR)、混合现实(mixed reality, MR)和虚拟现实(virtual reality, VR)。在 XR 系统中,需要捕获高分辨率的图像、视频和音频等模态,帮助远程服务器将用户环境中的对象和场景映射到虚拟显示器中。Wang 等^[30]构建了面向 AR/VR 多模态语义通信赋能的多用户边缘计算(mobile edge computing, MEC)系统。该系统采用双向缓存任务模型,由基于内容流行度的深度 Q 网络缓存算法来预测缓存内容,在缓存大小和通道容量受限的情况下,提出了缓存计算协调算法来实现两者的权衡。针对 XR 超高的数据速率要求,Zhang 等^[31]提出了通用的变长语义—信道编码方案。该方案使用速率分配网络来估计语义信息的最佳码长,并引入代理函数,以端到端的方式学习速率分配。人体网格重建和图像传输实验表明, XR 语义通信系统优于传统通信系统,变长编码方案优于定长编码方案。此外,Zhang 等^[32]还为 XR 开发了联合语义感知、渲染和通信的框架,通过结合语义信息的时空分布提高了语义感知效果。仿真结果显示,该框架可以降低头戴设备(head-mounted devices, HMDS)的感

知成本, 加快服务器的渲染过程, 减少 HMDS 与服务器之间的数据流量。

2.2.3 跨模态语义通信

随着无线通信、多媒体通信和信号处理的快速发展, 高清视频等多媒体应用已经不能满足用户的视听需求, 因此展开了将触觉模态与传统视听信号进行跨模态融合处理的语义通信模式研究。Li 等^[33]提出了一种跨模态语义通信范式, 包括 3 个模块: 跨模态知识图谱, 负责给编解码器提供基本背景知识作为辅助信息; 跨模态语义编码器, 负责推理潜在的隐式语义来减少多义性; 跨模态语义解码器, 支持强化学习的信号恢复, 保证语义一致性。为了支持多模态式服务, Wei 等^[34]提出了基于 AI 的跨模态语义通信, 其中联邦学习负责解决多模态服务中稀疏的数据收集和隐私保护问题; 强化学习负责构建缓存、通信和计算的联合优化框架以实现多流协同传输; 迁移学习负责从不同模态中提取、传输和融合语义特征, 恢复损坏的信号。仿真结果显示, 与其他无 AI 和语义转移的方法相比, 该方法具有最高的触觉重建性能。为了探究音频、视频和触觉信号之间的协作潜力, Zhou 等^[35]提出了内容驱动的协同通信机制, 通过研究跨模态信号的时间、空间和语义相关性, 对跨模态信息进行融合和共享, 开发了完整的信号恢复、重建和渲染方案。

多模态语义通信的相关研究仍处于初级阶段, 但语义通信“先理解后传输”的特点使其具备充分挖掘不同模态数据之间相关性的潜力。通过充分利用多模态数据之间的语义相关性, 语义通信能够大幅减少传输所需的数据量, 并提高多模态语义通信系统的可靠性。尽管面临许多挑战, 但其仍为实现用户沉浸式多模态服务提供了可行的思路。

2.3 鲁棒语义通信技术研究

语义通信可以通过减少传输数据量来提高通信效率, 但神经网络的引入也带来了新的威胁和风险。由于神经网络的脆弱性, 其容易受到对抗攻击的影响, 这使得语义通信在安全性方面存在一定的挑战。与传统通信不同, 语义通信系统中存在语义噪声(或语义攻击), 其可能会出现在编码、数据传输和解码过程中, 导致语义恢复出现错误。

2.3.1 图像语义噪声

不同模态的语义噪声特点不同, 对于图像模态, 可以通过给原始样本添加特定的扰动来欺骗神经网络, 且该扰动是人类的视觉无法感知的。因

此, 本文所关注的图像领域中的语义噪声可以基于对抗样本进行建模。目前已经提出了许多在图像领域生成对抗性样本的方法, 如快速梯度符号法(fast gradient sign method, FGSM)^[36]、投影梯度下降(projected gradient descent, PGD)^[37]以及通用对抗攻击(universal adversarial perturbations, UAP)^[38]等。这些方法都可以用来构建对抗性样本, 并以此作为基础对图像中的语义噪声建模。

针对上述语义噪声, 目前主要采取的防御措施包括对抗训练、修改网络结构(如防御蒸馏和正则化)、附加网络结构(如基于 GAN 的防御和对抗样本检测)等方式。Hu 等^[39]提出一种具有权重扰动的对抗性训练方法来对抗语义噪声的影响, 并设计了具有 Transformer 块的掩码自动编码器作为鲁棒的语义通信系统架构, 通过发送端和接收端共享离散码本进一步提高了系统的鲁棒性。Hu 等^[39-40]分别在发送端和接收端添加了与样本相关和与样本无关的语义噪声, 并且通过功能重要性映射来抑制与噪声相关和与任务无关的特征, 从而进一步降低了传输开销。Nan 等^[41]考虑无线通信的特性, 利用对抗训练生成了具有面向语义、不可察觉性、输入不可知性和可控性等特点的语义噪声, 在此基础上提出了一个适用于无线环境的鲁棒语义通信系统框架, 通过实验证明了其攻击和防御策略的有效性。He 等^[42]通过 GAN 训练的方法, 使接收端能够利用带有噪声的信息进行图像重构, 并在不同传输条件下最大限度地还原原始图像。Li 等^[43]考虑了语义通信过程中不同阶段出现的语义噪声对系统性能的影响。具体包括: 针对信源处的语义噪声, 在语义编码前利用 GAN 网络对其进行输入清理的操作, 去除语义噪声的影响; 针对信道处的语义噪声, 通过调整接收端的网络结构, 在信道解码器之前添加一个去噪自动编码器, 消除信道处语义噪声产生的影响。

2.3.2 文本语义噪声

由于图像具有连续性, 因而图像的对抗样本有很好的人类不可感知性。但对于文本而言, 即使是微小的改动也能明显被人类察觉到。因此, 面向文本的对抗攻击通常采用插入、删除、交换、字符替换以及单词替换等方式。同时, 被修改的部分需要在语法和语义上与原文本有高度相似性, 且修改比例不能过高。Peng 等^[44]全面地探索了文本传输中的语义噪声, 将语义噪声分为字面语义噪声和对抗性语义噪声两类, 并采用校准的自注意力机制和对抗

训练来减小语义噪声对文本恢复的影响。Hu 等^[45]提出了一种由知识图谱驱动的鲁棒语义通信系统,设计了一个由发送端和接收端共享的知识图谱,其中发送端只需传输符号在知识图谱中的索引,接收端通过知识图谱来恢复语义。与仅考虑信道噪声的文本传输相比,该系统在不同信噪比下均有效避免了语义噪声的影响。

目前,鲁棒语义通信技术的研究主要集中在图像和文本领域,而在视频和音频领域鲁棒语义通信尚未得到充分的关注和深入探索。这是因为视频和音频数据具有更高维度和更复杂的特征表示。因此,未来的研究应该继续深入探讨不同模态下的鲁棒语义通信技术,以提供更有效和可靠的语义信息传输方案。

3 未来研究方向

当前,语义通信的研究主要集中在语义通信的编解码技术和端到端传输技术,寻找不同模态/业务下的最佳语义编解码网络结构。本文进一步从语义通信的网络传输和隐私保护等角度探讨语义通信面临的挑战和未来研究方向。

3.1 语义通信传输协议研究

如何传输语义信息是语义通信的关键问题之一,但目前的研究工作还处于较为初级阶段。首先,Zhang 等^[46]从体系架构方面提出了以语义通信赋能的智简网络架构,并设计了全新的抽象协议层,具体包括辨识和分解用户意图的语义赋能应用意图层、语义赋能网络协议层和语义赋能物理承载层,相关研究工作正在开展中。其次,如何基于语义的重要性,提出面向语义信息的可变粒度语义数据帧定义及封包/解包机制,并进一步结合应用层的业务特点,设计动态适应多模态业务的语义数据包传输控制策略,是语义通信亟须解决的问题。最后,充分利用和融合多维网络资源(如通信、计算和存储等),优化多维网络资源感知的多模态业务调度,通过结合“存储换通信”和“存储换计算”等思路,可以进一步实现动态网络环境和复杂业务请求下语义的高效传输。

3.2 语义通信隐私保护技术研究

由于语义通信网络无须传输和通信完整的数据信息,而仅传输相关内容中的语义信息,因此可显著提高网络通信的安全性。但收发双方私有知识库

的多样性和互不可知性,容易造成接收端过度恢复,从而导致发送端隐私泄露问题。但当前对于语义通信的隐私保护机制研究还处于起步阶段。

语义通信依靠深度神经网络和知识库来实现智能通信。与无须理解即可完成编解码的传统通信不同,语义通信的特点是先理解,再进行传输。其中,知识库在语义理解方面发挥着关键作用,能够辅助语义的选择性压缩,并为语义重构提供背景知识。目前,大多数语义通信系统建立在发送端和接收端共享完全相同知识库的假设基础上。然而在实际中,发送端和接收端很难拥有一致的背景知识。在这种情况下,私有知识库的互不可知性和多样性可能会在收发双方间造成知识差距。互不可知性意味着发送端对接收端的私有知识库了解有限,例如,由于 A 与 B 的陌生关系和个人信息的非公开性,导致 A 缺乏对 B 职业、身份的了解。而多样性是指不同的知识库对同一事物可能有不同的理解或是不同的理解程度。例如,就“环球影城”而言,不熟悉的人可能理解为“电影院”,但了解的人知道它是“主题乐园”。由于这种知识差距的存在,语义通信过程中可能会存在隐私泄露问题,即接收端能够恢复出发送端不愿披露的敏感信息。例如,根据“B 和 C 去环球影城团建”的信息和对 C 的公司、职业以及其对“环球影城”的了解, A 可以推断出一些敏感信息,如 B 的地理位置和收入等。上述由于收发双方私有知识库间的知识差距所导致的隐私泄露,是语义通信面临的全新挑战,需要对其进行更深入的研究。

参考文献

- [1] IMT-2030(6G) 推进组.《6G 典型场景和关键能力》白皮书[EB/OL]. (2022-08-22)[2023-09-13]. <https://www.imt2030.org.cn/html/default/zhongwen/xinwendongtai/1561631590004572162.html?index=4>.
- [2] European Research Council. Semantics-empowered wireless connectivity: theoretical and algorithmic foundations[EB/OL]. (2022-09-15)[2023-09-13]. <https://cordis.europa.eu/project/id/101003431>.
- [3] XU W W, ZHANG G, BAI B, et al. Ten key ICT challenges in the post-Shannon era[J]. SCIENTIA SINICA mathematica, 2021, 51(7): 1095.
- [4] ZHANG P, XU W J, GAO H, et al. Toward wisdom-evolutionary and primitive-concise 6G: a new paradigm of semantic communication networks[J]. Engineering, 2022, 8(1): 60-73.
- [5] XIE H, QIN Z, LI G Y, et al. Deep learning enabled semantic

- communication systems[J]. IEEE transactions on signal processing, 2021, 69: 2663–2675.
- [6] XIE H, QIN Z. A lite distributed semantic communication system for internet of things[J]. IEEE journal on selected areas in communications, 2021, 39(1): 142–153.
- [7] ZHOU Q, LI R, ZHAO Z, et al. Semantic communication with adaptive universal transformer[J]. IEEE wireless communications letters, 2022, 11(3): 453–457.
- [8] WANG Y, CHEN M, LUO T, et al. Performance optimization for semantic communications: an attention-based reinforcement learning approach[J]. IEEE journal on selected areas in communications, 2022, 40(9): 2598–2613.
- [9] XIAO Y, SUN Z, SHI G, et al. Imitation learning-based implicit semantic-aware communication networks: multi-layer representation and collaborative reasoning[J]. IEEE journal on selected areas in communications, 2023, 41(3): 639–658.
- [10] WENG Z, QIN Z. Semantic communication systems for speech transmission[J]. IEEE journal on selected areas in communications, 2021, 39(8): 2434–2444.
- [11] WENG Z, QIN Z, LI G Y. Semantic communications for speech signals[C]//ICC 2021-IEEE international conference on communications. Montreal: IEEE, 2021: 1–6.
- [12] HAN T, YANG Q, SHI Z, et al. Semantic-preserved communication system for highly efficient speech transmission[J]. IEEE journal on selected areas in communications, 2023, 41(1): 245–259.
- [13] XIAO Z, YAO S, DAI J, et al. Wireless deep speech semantic transmission[C]//ICASSP 2023 – 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Rhodes Island: IEEE, 2023: 1–5.
- [14] BOURTSOULATZE E, KURKA D B, GÜNDÜZ D. Deep joint source-channel coding for wireless image transmission[J]. IEEE transactions on cognitive communications and networking, 2019, 5(3): 567–579.
- [15] YANG M, KIM H S. Deep joint source-channel coding for wireless image transmission with adaptive rate control[C]//ICASSP 2022 – 2022 IEEE international conference on acoustics, speech and signal processing (ICASSP). Singapore: IEEE, 2022: 5193–5197.
- [16] DAI J, WANG S, TAN K, et al. Nonlinear transform source-channel coding for semantic communications[J]. IEEE journal on selected areas in communications, 2022, 40(8): 2300–2316.
- [17] HUANG D, TAO X, GAO F, et al. Deep learning-based image semantic coding for semantic communications[C]//2021 IEEE global communications conference (GLOBECOM). Madrid: IEEE, 2021: 1–6.
- [18] HUANG D, GAO F, TAO X, et al. Toward semantic communications: deep learning-based image semantic coding[J]. IEEE journal on selected areas in communications, 2023, 41(1): 55–71.
- [19] LOKUMARAMBAGE M U, GOWRISSETTY V S S, REZAEI H, et al. Wireless end-to-end image transmission system using semantic communications[J]. IEEE access, 2023, 11: 37149–37163.
- [20] DONG C, LIANG H, XU X, et al. Semantic communication system based on semantic slice models propagation[J]. IEEE journal on selected areas in communications, 2023, 41(1): 202–213.
- [21] TUNG T Y, GÜNDÜZ D. DeepWiVe: deep-learning-aided wireless video transmission[J]. IEEE journal on selected areas in communications, 2022, 40(9): 2570–2583.
- [22] WANG S, DAI J, LIANG Z, et al. Wireless deep video semantic transmission[J]. IEEE journal on selected areas in communications, 2023, 41(1): 214–229.
- [23] JIANG P, WEN C K, JIN S, et al. Wireless semantic communications for video conferencing[J]. IEEE journal on selected areas in communications, 2023, 41(1): 230–244.
- [24] HUANG Y, ZHU Y, QIAO X, et al. AITransfer: progressive ai-powered transmission for real-time point cloud video streaming[C]//Proceedings of the 29th ACM International Conference on Multimedia. New York: Association for Computing Machinery, 2021: 3989–3997.
- [25] HUANG Y, ZHU Y, QIAO X, et al. Toward holographic video communications: a promising ai-driven solution[J]. IEEE communications magazine, 2022, 60(11): 82–88.
- [26] XIE H, QIN Z, LI G Y. Task-oriented semantic communications for multimodal data[J]. IEEE wireless communications letters, 2021, 11(3): 553–557.
- [27] XIE H, QIN Z, TAO X, et al. Task-oriented multi-user semantic communications[J]. IEEE journal on selected areas in communications, 2022, 40(9): 2584–2597.
- [28] ZHANG G, HU Q, QIN Z, et al. A unified multi-task semantic communication system with domain adaptation[C]//GLOBECOM 2022–2022 IEEE Global Communications Conference. Rio de Janeiro: IEEE, 2022: 3971–3976.
- [29] CHEN X, SONG X, WEI Y, et al. Dual semantic knowledge composed multimodal dialog systems[C]//Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: Association for Computing Machinery, 2023: 1518–1527.
- [30] WANG C, YU X, XU L, et al. Multimodal semantic communication accelerated bidirectional caching for 6G MEC[J]. Future generation computer systems, 2023, 140: 225–237.
- [31] ZHANG B, QIN Z, LI G Y. Semantic communications with variable-length coding for extended reality[EB/OL]. [2023–05–12]. <https://arxiv.org/abs/2302.08645>.

(下转第13页)

physicochemical properties of marine organic aerosols, and the climate effects of marine organic aerosols and their impact on the ocean-atmosphere system. By constructing numerical models of the evolutionary processes and climate effects of marine organic aerosols, the interaction processes and mechanisms between

marine organic aerosols and the ocean-atmosphere system can be elucidated, providing scientific basis for reducing climate prediction uncertainties and addressing climate change.

Keywords: marine organic aerosol; ocean-atmosphere system; earth system model; climate change

(上接第7页)

- [32] ZHANG B, QIN Z, GUO Y, et al. Semantic sensing and communications for ultimate extended reality[EB/OL]. [2023-05-12]. <https://arxiv.org/abs/2212.08533>.
- [33] LI A, WEI X, WU D, et al. Cross-modal semantic communications[J]. IEEE wireless communications, 2022, 29(6): 144-151.
- [34] WEI X, ZHOU L. AI-enabled cross-modal communications[J]. IEEE wireless communications, 2021, 28(4): 182-189.
- [35] ZHOU L, WU D, CHEN J, et al. Cross-modal collaborative communications[J]. IEEE wireless communications, 2020, 27(2): 112-117.
- [36] GOODFELLOW I J, SHLENS J, SZEGEDY C. Explaining and harnessing adversarial examples[EB/OL]. [2023-05-19]. <https://arxiv.org/abs/1412.6572>.
- [37] MADRY A, MAKELOV A, SCHMIDT L, et al. Towards deep learning models resistant to adversarial attacks[EB/OL]. [2023-05-19]. <https://arxiv.org/abs/1706.06083>.
- [38] MOOSAVI-DEZFOOLI S M, FAWZI A, FAWZI O, et al. Universal adversarial perturbations [EB/OL]. [2023-05-19]. <https://arxiv.org/abs/1610.08401>.
- [39] HU Q, ZHANG G, QIN Z, et al. Robust semantic communications against semantic noise[C]//2022 IEEE 96th Vehicular Technology Conference (VTC2022-Fall). London: IEEE, 2022: 1-6.
- [40] HU Q, ZHANG G, QIN Z, et al. Robust semantic communications with masked vq-vae enabled codebook[EB/OL]. [2023-05-19]. <https://ieeexplore.ieee.org/document/10101778>.
- [41] NAN G, LI Z, ZHAI J, et al. Physical-layer adversarial robustness for deep learning-based semantic communications[EB/OL]. [2023-05-19]. <https://dl.acm.org/doi/10.1109/JSAC.2023.3288249>.
- [42] HE Q, YUAN H, FENG D, et al. Robust semantic transmission of images with generative adversarial networks[C]//GLOBECOM 2022-2022 IEEE Global Communications Conference. Rio de Janeiro: IEEE, 2022: 3953-3958.
- [43] LI X, ZHANG X, CUI Q, et al. Robust semantic communication systems based on image transmission[C]// 2023 International Symposium on Communications and Information Technologies(ISCIT).

Sydney: ISCIT, 2023: 1-5.

- [44] PENG X, QIN Z, HUANG D, et al. A robust deep learning enabled semantic communication system for text[C]//GLOBECOM 2022 - 2022 IEEE Global Communications Conference. Rio de Janeiro: IEEE, 2022: 2704-2709.
- [45] HU L, LI Y, ZHANG H, et al. Robust semantic communication driven by knowledge graph[C]//2022 9th International Conference on Internet of Things: Systems, Management, and Security (IOTSMS). Milan: IEEE, 2022: 1-5.
- [46] ZHANG P, XU W, GAO H, et al. Toward wisdom-evolutionary and primitive-concise 6G: a new paradigm of semantic communication networks[J]. Engineering, 2022, 8(1):60-73.

Recent Progress on Semantic Communications

ZHANG Xuefei, GU Jing, LI Xiaoxian, CUI Qimei, TAO Xiaofeng
School of Information and Telecommunications Engineering,
Beijing University of Posts and Telecommunications, Beijing
100876

Abstract: With the emergence of 6G super wireless broadband services, the data rate demand reaches new heights. Semantic communication, depending on knowledge sharing and semantic understanding between the transmitter and the receiver, shows a potential on the bandwidth reduction due to the less amount of data needed to be transmitted. This paper focuses on the semantic coding/decoding technology and end-to-end semantic transmission technology, from the perspectives of the service characteristics and security requirements. In addition, based on the existing research results, it points out the challenges of semantic communication and discusses the perspective of the future work in this research area.

Keywords: semantic communication; high efficiency; high robustness