

文章编号: 1007-5321(2022)06-0012-09

DOI: 10.13190/j.jbupt.2022-158

面向智能机器通信的语义信息刻画及度量

马楠, 宋孟书, 刘宜明, 董辰
(北京邮电大学网络与交换技术国家重点实验室, 北京 100876)

摘要: 语义通信中关注信息的内在含义, 而非关注编码符号, 有助于解决智能机器通信中超大规模连接和海量数据传输的难题, 以实现机器间的实时智能通信。针对智能机器的语义通信系统, 基于信息论和人工智能技术, 描述了语义信息高度抽象、智能简约的特性, 归纳了端到端语义通信系统的设计方法、语义信息的刻画及度量方法、不同模态语义信息的编解码方法, 分析了语义通信系统的评价指标, 展望了应用前景。对现有研究存在的不足提出了进一步的研究方向。

关键词: 语义信息; 语义通信; 智能通信; 智能机器

中图分类号: TN911

文献标志码: A

Description and Measurement of Semantic Information for the Intelligent Machine Communication

MA Nan, SONG Mengshu, LIU Yiming, DONG Chen

(State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunications, Beijing 100876, China)

Abstract: Semantic communication focuses on the meaning of information rather than encoded symbols, which is helpful to solve the problem of massive connection and data transmission in intelligent machine communication and realize real-time intelligent communication between machines. The semantic communication system with intelligent machines is studied. First, based on information theory and artificial intelligence technology, the highly abstraction, intelligent and simple characteristics of semantic information are described, the design methods of an end-to-end semantic communication system are summarized, and the characterization and measurement methods of semantic information are presented. Then, the encoding and decoding methods of semantic information in different modes are summarized, and the evaluation indexes of a semantic communication system are analyzed. Furthermore, the research direction is given based on the shortcomings of the existing research.

Key words: semantic information; semantic communication; intelligent communication; intelligent machine

未来的无线通信系统不仅能满足人与人之间通信的需求, 并能支持海量智能机器 (IMs, intelligence

machines) 之间的信息交互及其与人类的智能交互。在语义通信中引入语义信息, 从关注符号传输到关

收稿日期: 2022-06-23

基金项目: 北京邮电大学基本科研业务费项目 (2021RC01); 国家自然科学基金项目 (U21A20448, 62001050)

作者简介: 马楠 (1979—), 男, 教授。

通信作者: 宋孟书 (1995—), 女, 博士生, 邮箱: MSSong@bupt.edu.cn。

注网络达意,将有助于提升通信效率,是未来通信系统的重要演进方向。

1 语义通信的研究现状

1948年,Shannon^[1]提出了经典信息理论,将沟通问题定义为“在接收方精确或近似地再现发送方发送的消息”,认为“通信的语义层面应该被认为与工程问题无关”。然而,用当前的移动通信网络难以处理以指数级增长的数据流量^[2],无法满足未来的通信需求,如 $1 \sim 10 \text{ Gbit} \cdot \text{s}^{-1} \cdot \text{m}^{-3}$ 的流量密度、 $1 \text{ Tbit} \cdot \text{s}^{-1}$ 的上下行链路数据速率、 0.1 ms 的时延等^[3]。此外,在未来无线通信系统中,如B5G (beyond-the fifth generation of mobile communications system) 及第6代移动通信系统(6G, the sixth generation of mobile communications system),将考虑众多智能应用场景,如脑机交互、虚拟现实、增强现实、混合现实等。在此背景下,学术界和产业界开始重视将语义通信理论和技术与智能应用场景相结合,关注信息所表达的内在含义。

Weaver等^[4]首次提出了语义通信的概念,指出语义通信中关注传输信息的含义,同时,定义了一个通信框架,包含语法信息、语义信息、语用信息。其中,关注语法信息即关注传输符号的准确性,关注语义信息即关注传输符号是否准确地表达预期含义。由此,众多学者聚焦语义信息,设计了语义通信系统模型,并在研究语义信息的刻画和度量方面进行了初步探索。

Shannon^[5]研究了文本字母预测的准确度问题,将 N 元熵定义为 F_N 来统计 N 个相邻文字的文本熵,表示当前 $N-1$ 个字母已知时,下一个字母的平均不确定度,并以26个字母的英文字母表为例,计算了一元熵、二元熵、三元熵的值,证明了相邻字母数量越多时,每个字母所占用的比特数越少。同时,Shannon指出:掌握一种语言的任何人都隐含地拥有关于此种语言的大量统计知识,并且熟悉单词、习语、语法和常用语句。所以,在语言校对中,他们能够填补缺失、不正确的字母或在对话中完成未完成的短语,这为自然语言处理(NLP, natural language processing)和语义知识库的建立提供了思路。

研究语义信息论需利用逻辑推理能力理解信息的内在含义^[6]。Zhong^[7]构建了从语法信息和语用信息2个层面获取语义信息的映射模型,但并未提及人工智能(AI, artificial intelligence)技术

对提取语义信息的重要性。Zhang等^[8]提出了语义基的概念,将AI技术和无线网络技术相结合,建立了智能高效的语义通信网络体系架构。该架构包括语义智能平面、语义物理传输层、语义网络协议层、语义应用意图层和语义信息流层,以支持6G通信场景中各类通信对象间的智能交互。Shi等^[9]引入语义转换获得语义符号,并对语音传输系统的语义保真度进行了研究。然而,智能通信网络中不仅要准确地传输数据,也要能降低设备的响应延迟和功耗。Xie等^[10]针对这2个关键问题,研究了分布式神经网络训练中的语义表达,展现了语义表达的前景。

2 语义通信与万物智联

移动通信技术与AI技术相结合,有利于IMs理解信息的内在意义,评估信息的真实价值。具有先进传感、快速计算和学习能力的IMs可以模仿人类的多种能力。同时,IMs间的互联也将产生大量数据。此外,由于通信系统往往具有特定的通信目标和时效性要求,所以IMs也需具备实时感知、实时通信、自主交互和自动决策的能力。

现代通信中,IMs的广泛应用产生了3种基本类型的交互通信模式:人—人通信(H2H Com, human-to-human communication),人—机通信(H2M Com, human-to-machine communication),机—机通信(M2M Com, machine-to-machine communication)^[11]。有IMs参与的通信(H2M Com和M2M Com)与H2H Com的本质区别在于所传输的信息不仅能被人理解,也能被机器理解。具体来讲,H2M Com通常设有智能人机界面,支持人与IMs之间的通信。IMs能够理解一个人发送的消息,并且做出对人有意义的行动或响应。H2M Com典型的应用领域包括人与AI共生系统、智慧医疗等。M2M Com指在没有人类参与的情况下,有效连接和协调海量物联网设备,执行无线网络中的感知、分析、学习、推理、决策等特定任务。M2M Com典型的应用领域包括自动驾驶、智能制造等。

3 语义通信模型

构建语义通信模型的基本思想是关注语义特征新维度,以准确传输信息的内在含义或特征为目的,在传统通信系统的基础上添加语义编码和语义解码,设计语义信源信道联合编解码模块,提升传输

效率。

3.1 语义编解码器

在传统通信系统中,发送端和接收端设备分别完成信源编码(解码)和信道编码(解码),信号的发送和接收不涉及任何智能。而在语义通信系统中,收发端是人、机器或其他具有智能的设备,人们关注的是传递消息的含义,而不是准确的比特流。语义通信系统的终端需要能感知环境和使用各种高度智能的算法,并能实现自主操作。

语义通信系统中引入了语义编码和语义解码模块。在语义发送端,识别并提取源信号的语义特征,采用信源信道联合编码(JSCC, joint source-channel coding)方案将语义特征进行编码并通过有噪信道传输。在语义接收端,利用语义解码器对接收信号进行解码,提取语义特征,经过逻辑推理后,将接收信号恢复成接收端用户可理解的信号。

在语义通信中,提高语义编码的效率需关注双重目标:

- 1) 检测并提取源信号的含义和语义特征;
- 2) 根据传输目标和实时信道反馈,在源头尽可能地压缩数据。

由于不同场景具有不同模态(文本、图像、语音或视频)信息和不同类型的收发实体,语义编码器还应具备以下功能:

- 1) 逻辑推理;
- 2) 检测发送端和接收端的知识差异;

- 3) 多模态信息处理。
- 语义解码器还应具备以下功能:
- 1) 逻辑推理;
 - 2) 将重建信息实时地反馈给编码器;
 - 3) 重建接收器无法识别或理解的符号^[9];
 - 4) 评估接收端的用户满意度。

3.2 语义知识库

语义通信系统是一个基于知识的系统^[3],发送端和接收端的智能机器可以感知现实环境中的信息,并通过自我学习建立不同模态信息和不同类型的知识库。发送端的IMs根据自身的知识库提取消息的语义信息,接收端的IMs根据自身的知识库解释和推断接收到消息的含义,以理解发送端IMs发送消息的含义^[12]。

建立背景知识库的本质是一个学习过程,类似于建立知识图谱,复杂且耗时。Zhong^[7]指出:可以通过将人类思维进行抽象来建立知识库,即将大量的知识组成一个大规模多级知识体系,如图1所示。同一级别的知识具有相同(或相似)的结构特征和实用特性,知识级别越低越具体,知识级别越高越抽象。例如,圆珠笔、毛笔和铅笔都具有各自的结构特征和实用特性,分别为结构特征1和实用特性1、结构特征2和实用特性2、结构特征3和实用特性3。因为它们有相似的结构特征,故都将“笔”作为它们的类名。显然,名称“笔”比名称“圆珠笔、毛笔、铅笔”更为抽象。

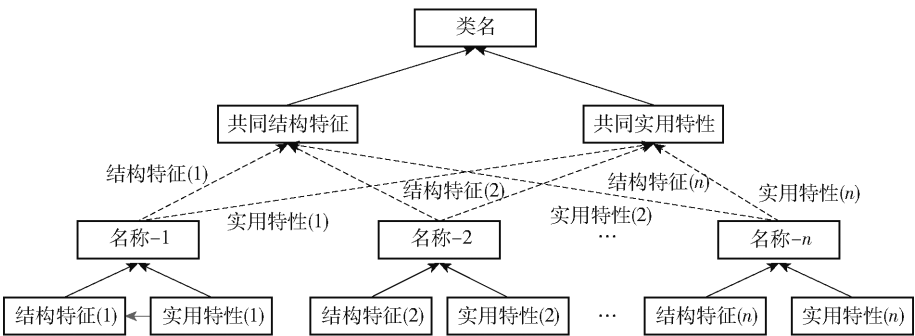


图1 基于语义信息的知识组织原理^[7]

在图像研究领域,可以利用全卷积神经网络、U-Net神经网络、点云神经网络、掩膜区域卷积神经网络等进行图像语义分割,输入整幅图像,输出图像每个像素的类别标签。利用目标检测的方法可预测图像中每个对象的位置和类别,利用语义分割方法可以对每个像素进行分类,利用实例分割方法可以在已检测目标物的范围内提取目标边缘。假设使用

IMs 对一场篮球赛进行自动打分,如果评价标准为“投进球加1分,未投进球不计分”,则只需对图像中篮框部分进行标注、提取、训练、分析;如果评价标准为“站在三分线之外投中的球加3分,站在三分线线上及线内投中的球计2分”,则需要标注篮框的基础上标注运动员的投球位置,并对这2个位置进行区分;进一步地,在上述2条规则的基础上加

入“运动员出现犯规行为时,投进的球不计分”的规定,则不仅需要标注篮框位置及投球地点,还需要对运动员的动作和行为进行标注及分析,从而使评分规则不断增加,像素分类不断细化,调用的目标提取模型不断增多。

在构建知识库时,可能会出现以下问题:

- 1) 消息具有多种模态;
- 2) 消息属于同一模态,但不在背景知识库中;
- 3) 收发端知识库存在差异;
- 4) 动态通信环境;
- 5) 信息私密性。

利用深度神经网络(DNNs, deep neural networks)的泛化能力及灵活结构^[13],可以通过以下方法处理上述问题:

- 1) 建立具有分级结构的多模态公共语义知识库;
- 2) 知识库扩增,或用现有知识库中具有相同语义含义的符号代替;
- 3) 在通信开始前进行收发端知识库协同共享;
- 4) 迁移学习方法,扩展、更新知识库;
- 5) 用户建立私有知识库,存储自己或仅与部分用户共享的私有语义知识和信息。

知识库可以通过训练、交流、共享,不断扩展和更新。语义知识库的建立基于系统的背景知识,IMs可以快速利用系统的背景知识从源信息中提取语义信息,并利用感知的新信息对背景知识库进行动态更新。同时,与收发端IMs的知识库协同,接收端IMs基于语义知识库可以更快速地理解接收到的信息,避免了数据采集过程中产生不必要的冗余,提高了IMs的信息提取和理解速度,增强了信息通信的时效性和可靠性。

3.3 语义噪声

语义通信系统中存在2种不同类型的噪声:物理信道噪声和语义噪声^[13]。物理信道噪声在无线通信中普遍存在,信道传播引起的误差可以通过信道解码来纠正。由于发送消息中的单词、句子或符号存在歧义,在消息内部的解释过程中^[14]存在语义噪声,使得接收端对接收到的语义信息产生了错误理解。语义噪声主要来源于以下2个方面。

1) 发送端和接收端的背景知识库不匹配,导致语义错误。例如:发送端用的是中文,但接收端只能识别英文。

2) 一词多义和同义词导致语义错误。例如:

“cherry”是一种水果,但“c”大写时,“Cherry”表示一个人的名字;“car”和“automobile”尽管在句法上不同,但在语义上是相同的。

在语义通信系统中,很难辨别信号传输错误是由语义噪声还是信道噪声引起的,所以为了在接收端成功解释语义信息,需要同时克服物理信道噪声和语义噪声。

4 语义编码

在香农经典信息论中考虑了信息的客观不确定性。在此基础上,Luca等^[15]进一步考虑信息的主观不确定性,刻画语义信息的熵。对涉及IMs的端到端通信场景,如H2M和M2M,引入深度学习(DL, deep learning)方法的自动编码器结构和去除分块结构,将发射机和接收机进行联合优化,将信息作为指令或待解决的问题传递给接收端IMs,以便IMs能够执行发送端的指令或做出适当响应。

传统通信系统为分离式信源信道模块化编译码。常见信源编码技术采用Huffman编码、算术编码、L-Z编码3种无损编码方式和脉冲编码调制、差分脉冲编码调制等有损信源编码技术。传统信道编码包括分组码、卷积码、Turbo码和Polar码等,语义通信系统中将信源编码和信道编码模块相结合,在香农熵的基础上研究语义熵,并引入机器学习方法设计信源信道联合编码方案。

4.1 熵编码

虽然在香农经典信息论中只考虑了信息统计概率的客观因素,但在此基础上通过扩展熵和互信息2个关键概念得到语义熵和语义互信息^[4]。香农信息论中

将消息空间 X 的熵定义为 $H(X) = -\sum_{i=1}^n p(x_i) \lg p(x_i)$,其中 x_1, x_2, \dots, x_n 为 X 的所有可能样本, $p(x_1), p(x_2), \dots, p(x_n)$ 为样本出现的概率。

然而,由于日常所用信息的含义往往具有模糊性,如高、矮、轻、重、老、少、大概、几乎、差不多等,只用统计概率无法描述,需要借助模糊集合论研究消息本身含义的不确定性所产生的模糊信息。为此,Zadeh等^[16]首次提出模糊集理论,并定义了模糊集合。在此基础上,Luca等^[15]首次利用香农熵的概念定义了模糊集合的模糊熵,并定义了考虑信息随机性和模糊性的双重不定性总熵。在此基础上,吴伟陵^[17]将语义信号看作源信号的模糊映射,通过研究广义信源,给出了广义熵的概念,得到形式类似于总

熵的结论。

熵是事件不确定性的度量,信息熵越小所含信息量越小,事件的不确定性越小,可向有序方向演进。协同理论创始人 Haken^[18]指出,信号活动是对行动有指导意义的规律性总结,研究信息的意义内涵有望达到熵减。开放系统耗散理论表明,远离平衡态的开放系统不断与外界交换物质和能量,可从无序状态转变为有序状态。因此,将背景知识看做传统通信系统的系统外干预能使系统向有序方向演进,达到熵减。基于此,Bao 等^[19]证明在通信系统中加入背景知识能得到熵减,同时证明了语义熵 $H(W)$ 和语法熵 $H(X)$ 的关系满足 $H(W) = H(X) + H(W|X) - H(X|W)$ 。具体地,Bao 等^[19]考虑背景知识对刻画语义信息的影响,将语义信息源建模为四元组 (X, K, I, M) ,其中 X 为源信息, K 为源信息的背景知识, I 为与背景知识相关源信息的推断, M 为信息编码器。类似地,Shi 等^[20]认为语义通信的关键问题是要找到一个合适的语义解释 W 和编码方案 $P(W|X)$,使发送消息的语义歧义 $H(X|W)$ 和编码冗余 $H(W|X)$ 之间的差值接近于 0。

4.2 不同模态信源编码方案

4.2.1 文本语义编码

Guler 等^[21]提出语义索引分配法,其中语义相似度越大的词对应的汉明距离越小,并且通过在接收端进行反向索引赋值重建,有效提高了收发端词语的语义相似度。然而,由于编码长度与单词数量呈指数关系,故此方法仅适用于词汇数量有限的情况。为理解长句,Kneser 等^[22]捕捉单词之间的关系定义词语间的空间距离,但这种方法并不能描述语义信息。

Mikolov 等^[23]提出用深度语境化的单词表示方法模拟单词的语义特征。Peters 等^[24]开发了通用预训练语言表征模型——BERT (bidirectional encoder representation from transformers),生成深度双向语言表征,不需要重新设计词的表示便可为各种 NLP 任务提供词向量。Farsad 等^[25]建立了基于 DL 方法的 JSCC 方案,忽略标点错误对传达语义信息含义造成的影响,相对于传统编码方案有更好的性能。Xie 等^[26]基于 Transformer 自注意力机制设计了智能端到端 JSCC 方案,通过学习句子间的潜在关系,提取句子语义,提升传输效率。进一步地,Xie 等^[14]采用多种 DL 方法表征文本的隐含意义,以提取长句的语义信息,利用迁移学习联合训练编解码器,建立了

适用于具有不同背景知识的动态通信环境的 JSCC 方案。

4.2.2 图像语义编码

图像编码的解决方案和文本处理类似,最大的区别在于 DNNs 的结构不同。因为视频文件由不同的图像逐帧组成,所以在视频语义通信系统中基本能通过提取每帧图像的特征来实现视频的传输。Zhai 等^[27]证明,在视频通信中,JSCC 方案的性能优于块编码方案的性能。

Bourtsoulatze 等^[28]提出基于卷积神经网络的端到端 JSCC 方案,实现了加性高斯白噪声和 Rayleigh 信道下的高分辨率图像传输,得到了比传统主流压缩算法,如 JPEG 和 JPEG2000 性能更优的结果,且没有悬崖效应。Kurka 等^[29]充分利用信道的反馈信息,设计了 deepJSCC-f 方案,有效提高了图像传输系统中固定长度传输端到端的重构质量,降低了变长传输的平均延迟,并实现了图像的连续细化^[30]。进一步地,Kurka 等^[31]提出一种基于 DL 方法的图像信源无线信道自适应带宽传输方法。数值仿真结果表明,与单次传输相比,所提方法有效提升了图像信源的自适应传输能力。Ding 等^[32]针对噪声信道下多用户图像传输的 JSCC 方案,设计了一种基于自编码器的新型深度 JSCC 方案。实验结果表明,该方案可应用于多用户场景,具有较强的适应性和鲁棒性。

4.2.3 语音语义编码

由于基于语义信息的语音通信系统的传输质量不仅涉及语音信号的保真度、响亮度、频率、音调,还涉及发送者的情绪,即同一文本可能表达不同的情感。例如:“你怎么了?”如果用柔和的语调来说,是表达对一个人的关心;如果用焦虑的语调来说,便表达一种抱怨之意。此外,从语音信号中识别方言也非常困难。所以,目前对端到端语音语义信息传输的研究相对较少。

Weng 等^[33]建立了注意力机制挤压-激励网络,证明了即使在低信噪比环境下,此网络也具有很好的鲁棒性。Tong 等^[34]引入联邦学习方法,进一步提高了语义信息提取的准确性,仿真结果表明,与传统的编码方案相比,该算法有效降低了恢复后的音频信号与源音频信号之间的均方误差 (MSE, mean square error)。

4.2.4 跨模态语义编码

对于含有 IMs 的端到端语义通信系统,接收端

IMs可能从一个发送端接收到不同模态的消息,所以研究IMs的语义通信系统不仅需要研究单模态语义信息,也不应忽略多模态语义信息之间的关系和转化。

针对多模态信息的IMs的端到端语义通信系统,首要应探析来自多个智能用户多模态数据之间的相关性。例如:将语音翻译成文本,再将文本转化为语义信息进行通信。或发送端IMs通过从不同信号发射机提取异构消息的语义信息;然后根据目的地用户的需求在接收端的IMs直接将接收到的语义信息转化为目的地用户可以理解的模态信息。

Xie等^[35]针对面向任务的具有多模态数据的多用户语义通信系统,联合设计优化了收发器,实现了面向智能任务的信息传输。Wang等^[36]使用了具有层次语义结构的深度跨模态哈希检索方法,通过保留层内和层间的语义相似关系刻画了图像与文本之间的语义相关性。Chen等^[37]提出了多尺度视听融合端到端的通信模型,研究了音频与视觉之间的语义联系,有效提高了声源地及其周围区域重建视频信息的准确性。Tan等^[38]利用生成对抗网络建立了跨模态语义匹配生成对抗网络,提高了文本描述和合成图像之间的语义一致性,实现了细粒度的文本-图像生成。

4.3 信源信道联合编码

考虑到信源的类型,端到端通信的JSCC方案在接收端直接恢复源信息而不是比特位^[39]。

由于DNNs具有强大的表示能力,所以业界提出将DL方法用于JSCC方案。Xie等^[14]提出了一种基于DL方法的JSCC方案,在编码和解码时考虑语义相似度和数据传输速率。Li等^[40]利用神经网络研究了JSCC方案。Xie等^[26]在JSCC方案的基础上提出了一种可以实现系统最大容量的语义通信框架。在此框架中,接收端的IMs基于传感器和自我注意机制可以更容易地理解信息的含义。Luo等^[41]建立了一个通用的基于DL方法的端到端语义通信模型。对于静态通信环境下的端到端语义通信场景,源信息和语义库保持不变,使用随机梯度下降算法对语义编码器和信道编码器进行联合训练;对于动态通信环境下的端到端语义通信场景,利用迁移学习训练编码器和解码器对知识库进行更新。

语义信息的应用使得IMs具备多模态信源感知能力,通过语义编码对感知到的源信息进行有效压缩,可以有效减少通信资源的占用,提高通信的效率

和信息传输的准确性,有助于实现面向目标的感知-通信-计算一体化的高速通信。

5 评价指标

在语义通信中,接收方以最小语义歧义为目标提取源信息的语义信息,并考虑多种影响因素,如对信息的理解与提取能力、语义知识库、系统外在干预、收端用户满意度等对系统性能的影响,确保语义信息的正确传输。同时,语义信息的评价指标根据源信息模态类型的不同而有所不同。

5.1 语义信息距离

香农信息论中利用互信息表达2个事件集合之间的相关性,即接收到消息Y后获得的源信息X的信息量。吴伟陵^[17]利用广义互信息量度量收发端信息的相关性,研究了广义信息源与广义熵,并证明了广义互信息量与香农互信息量具有相似的性质。

针对外界影响对通信系统的干扰,Guler等^[42]考虑外部影响实体,研究了有噪声信道上的语义通信系统,建立了语义错误度量,得到了最小化端到端平均语义错误的最优传输策略。

5.2 各模态语义信息距离

5.2.1 文本语义信息距离

刻画2个单词的语义相似度的一种直观方法是为单词指定相似的索引。Guler等^[21]提出了语义索引分配法,用汉明距离衡量两词之间的语义差异。利用语义分配索引、反向索引分配重建,进行语义信息编码,能够很好地区分语义相似的单词和语义不相关的单词。

在机器翻译中,除了通过误码率来建立句子之间的相似度外,通常使用双语替换评测(BLEU, bilingual evaluation understudy)来评价2个句子中单词的相似程度^[22]。由于BLEU不能评价信息语义的相似程度,Xie等^[14]提出了一种新的度量指标——句子相似度,以度量发送端发送的原句和接收端恢复句子之间的句子相似程度。

利用潜在语义分析(LSA, latent semantic analysis)方法,通过分析文本与词之间的关系构建词-文本矩阵,研究文本与单词之间的基于话题的语义关系,并通过对词-文本矩阵进行奇异值分解,得到低维语义子空间,以此计算任意2个词的相似性。Deerwester等^[43]在LSA方法的基础上设计了概率潜在语义分析(pLSA, probabilistic latent semantic

analysis)方法,将LSA方法的先验分布转为均匀分布并对参数求最大后验估计,得到了优于LSA方法的结果。由于pLSA方法无法生成新的未知文档且随着文档数量的增加容易导致模型过拟合,Blei等^[44]在pLSA方法的基础上提出了潜在狄利克雷分配(LDA, latent Dirichlet allocation)模型。LDA模型是目前文本类模型研究领域中的应用最广泛的模型。

5.2.2 图像语义信息距离

MSE是衡量压缩重构图像质量优劣最普遍、最广泛的指标。Bourtsoulatze等^[28]采用图像像素的最大值和峰值信噪比(PSNR, peak signal-to-noise ratio)计算MSE。MSE越小,PSNR越大,图像的重建质量越好。这种方法易计算、易优化,具有明确的物理意义,但是并不能衡量感知视觉的质量水平。

在人类视觉感知可以从不同实用场景中提取结构信息的假设下,德州大学奥斯丁分校的图像和视频工程实验室研究团队^[45]基于图像的亮度、对比度和结构,提出采用结构相似性指数衡量图像的相似度,为衡量机器的智能感知通信水平提供了很好的方法。Lee等^[46]设计了图像语义通信系统传输-识别的联合方案,衡量了图像的语义识别精度。

5.2.3 语音语义信息距离

采用分段信噪比衡量语音信息可以解决语音信号在不同时间段具有不同信噪比的问题。信号失真比是衡量原始语音向量和重构语音序列之间误差的有效度量方法,较高的SDR(signal to distortion ratio)表示重构语音信号更容易理解。在语音信源的端到端语义通信系统中,Weng等^[33]采用PESQ(perceptual evaluation of speech quality)算法评估各种条件下,如背景噪声、模拟滤波等的语音信号。POLQA(perceptual objective listening quality analysis)是PESQ算法的升级版,用于评估具有嘴唇和耳朵模拟器的人造头部的声学记录语音信号。

6 语义通信的应用

面向IMs的语义通信系统具备语义提取、快速计算、实时传感、有效通信、自主学习等优势,在众多领域中起到至关重要的作用。近来,元宇宙概念受到广泛关注。元宇宙将虚拟世界和现实世界连接起来并在两者间进行实时交互,应用场景包括虚拟面对面办公、虚拟游戏、沉浸式线上购物、智能制造、智慧医疗等。其中,智慧医疗领域中的数字孪生医疗

诊断系统备受关注。

数字孪生医疗诊断系统中通过无人管理、精准感知、实时监控和控制进行医疗诊断,实现智慧管理。实现孪生医疗诊断需要超大规模的数据连接和超高时效以及准确的信息传输,会消耗大量的频谱资源。基于语义通信,医疗诊断系统中在具备多类传感器和计算单元的IMs获取环境状况(障碍物、温度、湿度)和人类状况(体温、心率、血糖)等监测信息后,通过语义编码对原始数据中的重要区域进行目标提取,能够快速准确地理解和捕获所需的医疗信息;通过信源信道联合编码传递语义信息,能大量减少通信开销;同时,接收端IMs通过恢复语义信息并进行重建,保证了低时延通信,有效提高了IMs的运行效率^[47],为实现快速精准的医疗诊断提供了有力保障。此外,建立共享语义知识库和本地知识库,将为解决个人隐私的安全问题提供思路。

7 结束语

尽管对语义通信系统的设计和编解码方案已经开展了一系列的研究,但是实现基于语义信息的IMs通信还面临诸多挑战。

现代无线通信系统对时效性的要求越来越高,虽然引入语义信息减少了大量的通信开销,但在端到端时延方面的研究相对匮乏。处理语义信息需要的时间主要包括语义信息提取、收发端知识库协同、多用户语义信息识别3个方面。如何在有限的网络资源下及时对信息进行语义识别和语义提取以及分析语义信息提取时间和传输时间对提高系统的信号传输效率至关重要。

将发送端IMs接收到源信息至接收端IMs完成对其收到的信息进行解码所经历的时间定义为信息年龄,不同的IMs通信网络对信息年龄有不同的要求。在面向IMs的语义通信网络中引入信息年龄研究语义信息的时间有效性压缩和实时重建,对实现IMs间的高效通信将发挥重要作用。同时,收发端语义知识库的协同需耗费大量的时间和资源,在实际应用中,是否需要收发端知识库进行协同以及利用何种方法进行协同均为值得关注和思考的问题。最后,在多用户传输环境下接收端IMs解释语义信息的难度高,随着用户数量的增多,如何在有效时间内最大化用户数量和网络吞吐量也是急需解决的重要问题。

综上所述,利用语义通信技术挖掘信息的语义

特征新维度,研究语义信息的表征方法和信源信道联合编码方法,将会在未来智能通信的发展中发挥重要作用。同时,语义信息理论和语义通信方法也有待进一步研究。

参考文献:

- [1] SHANNON C E. A mathematical theory of communication [J]. The Bell System Technical Journal, 1948, 27(3): 379-423.
- [2] WANG C X, RENZO M D, STANCZAK S, et al. Artificial intelligence enabled wireless networking for 5G and beyond: recent advances and future challenges[J]. IEEE Wireless Communications, 2020, 27(1): 16-23.
- [3] STRINATI E C, BARBAROSSA S. 6G networks: beyond Shannon towards semantic and goal-oriented communications [J/OL]. Computer Networks, 2021, 190: 107930 [2022-06-20]. <http://doi.org/10.1016/j.comnet.2021.107930>.
- [4] WEAVER W, BURKS A W, SHANNON C E. The mathematical theory of communication[J]. The Philosophical Review, 1951, 60(3): 398-410.
- [5] SHANNON C E. Prediction and entropy of printed English [J]. Bell System Technical Journal, 1948, 30(1): 50-64.
- [6] 钟义信, 张瑞. 信息生态学与语义信息论 [J]. 图书情报知识, 2017(6): 4-11.
ZHONG Y X, ZHANG R. Information ecology and semantic information theory [J]. Documentation, Information & Knowledge, 2017(6): 4-11.
- [7] ZHONG Y X. A theory of semantic information [J]. China Communications, 2017, 14(1): 1-17.
- [8] 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: 60-73.
- [9] SHI G, GAO D, SONG X, et al. A new communication paradigm: from bit accuracy to semantic fidelity [EB/OL]. (2021-01-29) [2022-04-29]. <https://ui.adsabs.harvard.edu/abs/2021arXiv210112649S>.
- [10] XIE H Q, QIN Z J. A lite distributed semantic communication system for Internet of things [J]. IEEE Journal on Selected Areas in Communications, 2020, 39: 142-153.
- [11] 石光明, 肖泳, 李莹玉, 等. 面向万物智联的语义通信网络 [J]. 物联网学报, 2021, 5(2): 26-36.
SHI G M, XIAO Y, LI Y Y, et al. Semantic communication networking for the intelligence of everything [J]. Chinese Journal on Internet of Things, 2021, 5(2): 26-36.
- [12] LAN Q, WEN D Z, ZHANG Z H, et al. What is semantic communication? a view on conveying meaning in the era of machine intelligence [J]. Journal of Communications and Information Networks, 2021(4): 336-371.
- [13] SANA M, STRINATI E C. Learning semantics: an opportunity for effective 6G communications [J/OL]. 2022 IEEE 19th Annual Consumer Communications & Networking Conference (CCNC), 2022, 631-636 [2022-06-20]. <https://arxiv.org/abs/2110.08049>.
- [14] 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.
- [15] LUCA A D, TERMINI S. A definition of a nonprobabilistic entropy in the setting of fuzzy sets theory [J]. Information and Control, 1972, 20(4): 301-312.
- [16] ZADEH, LOTFI A. Fuzzy logic-a personal perspective [J]. Fuzzy Sets and Systems, 2015, 281(15): 4-20.
- [17] 吴伟陵. 广义信息源与广义熵 [J]. 北京邮电大学学报, 1982, 5(1): 29-41.
WU W L. Generalized information source and generalized entropy [J]. Journal of Beijing University of Posts and Telecommunications, 1982, 5(1): 29-41.
- [18] HAKEN H. Synergetics [J]. Naturwissenschaften, 1980, 67(3): 121-128.
- [19] BAO J, BASU P, DEAN M, et al. Towards a theory of semantic communication [C] // 2011 IEEE Network Science Workshop. Piscataway, NJ: IEEE Press, 2011: 110-117.
- [20] SHI G M, XIAO Y, LI Y Y, et al. From semantic communication to semantic-aware networking: model, architecture, and open problems [J]. IEEE Communications Magazine, 2021, 59(8): 44-50.
- [21] GULER B, YENER A. Semantic index assignment [C] // IEEE International Conference on Pervasive Computing and Communications Workshops. Piscataway, NJ: IEEE Press, 2014: 431-436.
- [22] KNESER R, NEY H. Improved backing-off for M -gram language modeling [C] // 1995 International Conference on Acoustics, Speech, and Signal Processing. Piscataway, NJ: IEEE Press, 1995: 181-184.
- [23] MIKOLOV T, CHEN K, CORRADO G, et al. Efficient estimation of word representations in vector space [J/OL]. Computer Science, 2013: 114-135 [2022-06-20]. <https://ui.adsabs.harvard.edu/abs/2013arXiv1301.3781M>.
- [24] PETERS M, NEUMANN M, IYYER M, et al. Deep contextualized word representations [J/OL]. 2018 The North American Chapter of the Association for Computational Linguistics, 2018 [2022-06-20]. <https://ui.adsabs.harvard.edu/abs/2018arXiv180205365P>.

- [25] FARSAD N, RAO M, GOLDSMITH A. Deep learning for joint source-channel coding of text[C]//2018 IEEE International Conference on Acoustics, Speech and Signal Processing. Piscataway, NJ: IEEE Press, 2018: 2326-2330.
- [26] XIE H Q, QIN Z J, LI G Y, et al. Deep learning based semantic communications: an initial investigation[C]//GLOBECOM 2020-2020 IEEE Global Communications Conference. Piscataway, NJ: IEEE Press, 2020: 1-6.
- [27] ZHAI F, EISENBERG Y, KATSAGGELOS A K. Joint source-channel coding for video communications [J]. Handbook of Image and Video Processing, 2005: 1065-1082.
- [28] BOURTSOULATZE E, KURKA D B, GÜNDÜZ D. Deep joint source-channel coding for wireless image transmission[C]//ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing. Piscataway, NJ: IEEE Press, 2019: 4774-4778.
- [29] KURKA D B, GÜNDÜZ D. Deep joint source-channel coding of images with feedback [C]//ICASSP 2020-2020 IEEE International Conference on Acoustics. Piscataway, NJ: IEEE Press, 2020: 5235-5239.
- [30] KURKA D B, GÜNDÜZ D. Successive refinement of images with deep joint source-channel coding [C]//2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications. Piscataway, NJ: IEEE Press, 2019: 1-5.
- [31] KURKA D B, GÜNDÜZ D. Bandwidth-agile image transmission with deep joint source-channel coding[J]. IEEE Transactions on Wireless Communications, 2021, 20(12): 8081-8095.
- [32] DING M, LI J H, MA M Y, et al. SNR-adaptive deep joint source-channel coding for wireless image transmission [C]//ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Piscataway, NJ: IEEE Press, 2021: 1555-1559.
- [33] WENG Z Z, QIN Z J, LI G Y. Semantic communications for speech signals [C]//ICC 2021-IEEE International Conference on Communications. Piscataway, NJ: IEEE Press, 2020: 1-6.
- [34] TONG H N, YANG Z, WANG S, et al. Federated learning based audio semantic communication over wireless networks [C]//2021 IEEE Global Communications Conference. Piscataway, NJ: IEEE Press, 2021: 1-6.
- [35] XIE H Q, QIN Z J, LI G Y. Task-oriented multi-user semantic communications for VQA task [J]. IEEE Wireless Communications Letters, 2022, 11(3): 553-557.
- [36] WANG D, ZHANG C, WANG Q, et al. Hierarchical semantic structure preserving hashing for cross-modal retrieval [J]. IEEE Transactions on Multimedia, 2022: 1-13.
- [37] CHEN Y X, ZHAO P C, QI M, et al. Audio matters in video super-resolution by implicit semantic guidance [J]. IEEE Transactions on Multimedia, 2022: 4128-4142.
- [38] TAN H C, LIU X P, YIN B C, et al. Cross-modal semantic matching generative adversarial networks for text-to-image synthesis [J]. IEEE Transactions on Multimedia, 2022, 24: 832-845.
- [39] PARK S, SIMEONE O, KANG J. End-to-end fast training of communication links without a channel model via online meta-learning [C]//2020 IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications. Piscataway, NJ: IEEE Press, 2020: 1-5.
- [40] LI R W, WU L N, GUO D L. Joint source/channel coding modulation based on BP neural networks [C]//International Conference on Neural Networks and Signal Processing. Piscataway, NJ: IEEE Press, 2003: 156-159.
- [41] LUO X W, CHEN H H, GUO Q. Semantic communications: overview, open issues, and future research directions [J]. IEEE Wireless Communications, 2022, 29(1): 210-219.
- [42] GULER B, YENER A, SWAMI A. The semantic communication game [J]. IEEE Transactions on Cognitive Communications and Networking, 2018, 4(4): 787-802.
- [43] DEERWESTER S, DUMAIS S T, FURNAS G W, et al. Indexing by latent semantic analysis [J]. Journal of the American Society for Information Science, 1990, 41(6): 391-407.
- [44] BLEI D M, NG A Y, JORDAN M I. Latent dirichlet allocation [J]. Machine Learning Research Archive, 2003, 3: 993-1022.
- [45] WANG Z, BOVIK A C, SHEIKH H R, et al. Image quality assessment: from error visibility to structural similarity [J]. IEEE Transactions on Image Processing, 2004, 13(4): 600-612.
- [46] LEE C H, LIN J W, CHEN P H, et al. Deep learning-constructed joint transmission-recognition for Internet of things [J]. IEEE Access, 2019, 7: 76547-76561.
- [47] KOUNTOURIS M, PAPPAS N. Semantics-empowered communication for networked intelligent systems [J]. IEEE Communications Magazine, 2021, 59(6): 96-102.