中国科学 : 信息科学 2026 年 第 56 卷 第 3 期: 741–742
SCIENTIA SINICA Informationis
《中国科学》杂志社
SCIENCE CHINA PRESS
精选文章导读

基于多模态大模型的个性化速率拆分在车联网中的应用

张圣羽[1], 张世尧[2], 袁伟杰[3*], 石嘉[1], 李赞[1], Tony Q.S. QUEK[4]

1. 西安电子科技大学通信工程学院, 空天地一体化综合业务网全国重点实验室, 西安 710071, 中国
2. 大湾区大学先进工程学院, 东莞 523000, 中国
3. 南方科技大学系统设计与智能制造学院, 深圳 518055, 中国
4. Information Systems Technology and Design Pillar, Singapore University of Technology and Design, Singapore 487372, Singapore
* 通信作者. E-mail: yuanwj@sustech.edu.cn

随着智能交通系统 (intelligent transportation system, ITS) 和车联网 (vehicle-to-everything, V2X) 技术的快速演进, 未来交通系统正逐步迈向高效、智能与低时延的深度融合阶段. V2X 的核心在于车辆、基础设施与环境的实时信息交互, 但车辆高速移动、环境复杂多变, 使得获取精确的信道状态信息在实际场景中几乎难以实现 [1,2]. 特别是在多车协同、城市道路交叉口、高速公路合流等复杂环境下, 信道快速时变、反馈延迟以及遮挡效应使传统波束赋形与多址技术面临明显的性能瓶颈.

为应对上述挑战, 速率拆分多址接入 (rate-splitting multiple access, RSMA) 技术被提出 [3,4]. 因其对不完美信道的天然鲁棒性而被认为是后 5G/6G 时代的重要多址方案之一. 然而, 现有 RSMA 波束赋形设计大多依赖迭代式优化求解或基于相对静态的信道环境的假设, 无法满足高速 V2X 场景的实时性与复杂性需求. 同时, 近年来研究者提出了预测式波束赋形技术. 其通过深度学习实现 “单步” 信道状态到波束赋形的范式, 为解决动态场景下的波束预测提供了新思路. 但因其大多仍停留在单模态输入 (如历史信道信息), 无法充分利用道路几何信息、交通状态语义等异构数据的潜在价值, 导致多址接入技术的性能难以被充分利用 [5,6].

SCIENCE CHINA Information Sciences 在 2025 年第 68 卷第 7 期发表了论文 “Personalizing rate-splitting in vehicular communication via large multi-modal model”. 针对上述痛点, 该文系统性提出了一种基于多模态大模型 (large multi-modal model, LMM) 的个性化 RSMA V2X 通信框架. 旨在引入道路拓扑、交通语义及历史信道信息, 综合性构建一个具备理解与推理能力的 GTR (GCN-transformer-reasoner) 网络, 实现对未来波束的精准预测与频谱资源合理分配.

具体而言, 在高度动态且结构复杂的 V2X 环境中, 传统依赖单一信道历史信息的波束预测方法难以全面刻画车辆移动、道路拓扑与场景语义之间的多维耦合关系. 为突破这一限制, 本文提出的 GTR 框架首先从多模态视角重新构建 V2X 场景描述方法, 将道路几何、交通语义与历史信道信息统一纳入端到端学习体系. 如图 1 所示, 在该框架中, 道路中心线、交叉口与车道连接关系被抽象为带有拓扑结构的车道图, 经由图卷积网络提取重要的空间依赖特征; 与此同时, 历史信道信息的时序关联通过变压器 (Transformer) 模型编码器进行建模, 使模型能够捕捉信道随时间变化的潜在规律. 与之配合的还有语义信息输入模块, 它将路段名称、交通条件等高层语义加入推理链条, 使模型能够对道路上下文形成更具人类知识特征的理解. 多模态特征在统一表示空间中的融合, 使 GTR 能够在复杂城市道路下准确把握交通环境结构信息对信道演化的深度影响, 为后续的波束与资源分配奠定了数据基础.

在此多模态信息融合的基础上, GTR 构建了一个兼具空间解析能力、时间依赖建模能力与上下文理解能力的统一决策架构. 该架构以图卷积网络 (GCN) 模块提供的几何特征和变压器模型编码器提取的历史信

英文原文: Zhang S Y, Zhang S Y, Yuan W J, et al. Personalizing rate-splitting in vehicular communication via large multi-modal model. Sci China Inf Sci, 2025, 68: 170305, https://doi.org/10.1007/s11432-024-4423-4

© 2026《中国科学》杂志社

https://www.sciengine.com/doi/10.1360/SSI-2025-0539

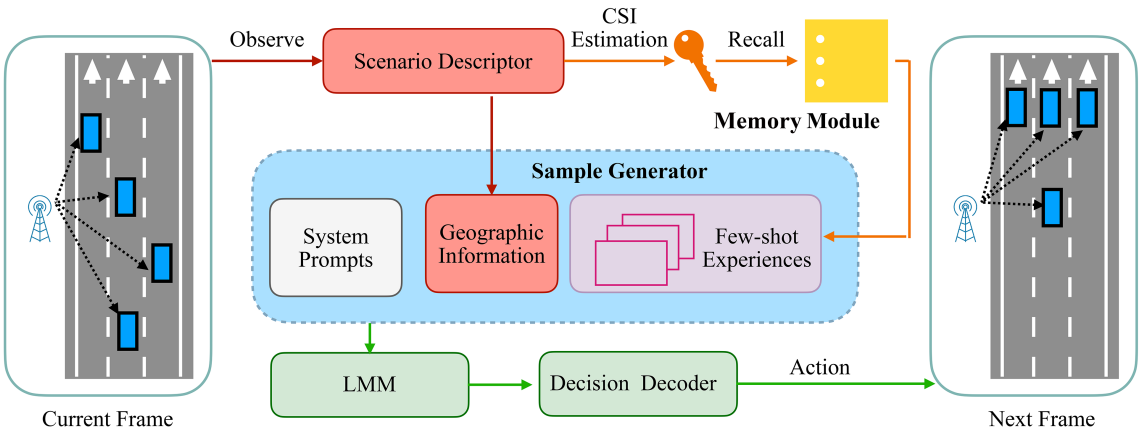


图 1　(网络版彩图) 多模态大模型架构.

Figure 1　(Color online) The proposed LMM framework in V2X.

道特征为输入, 在时间位置编码的辅助下形成可供深度注意力机制处理的序列表示. 随后, 变压器模型解码器通过结合道路语义信息, 对这些空间 – 时间联合特征进行条件化推理, 生成与当前道路场景一致的波束设计方案. 最终决策由生成模块完成, 将模型输出映射为 RSMA 所需的预编码矩阵及公共速率分配, 使整个系统实现从场景理解到通信策略生成的完全自治闭环. 通过这一结构化、层次化的设计, GTR 能够在面对道路环境急剧变化、车辆高速运动及信道快速衰落等情况下依然保持强大的鲁棒性和泛化能力, 为高动态车联网提供了可落地的未来智能通信新范式.

文章还系统总结了当前多模态大模型在未来面向 6G 场景中面临的多方面挑战与潜在突破点, 包括在大规模交通系统中提升模型的可解释性与部署效率, 以及在异构网络中实现跨域自治决策. 随着空天地海一体化网络的发展, V2X 通信将进一步向卫星与高空平台延展, 也使得多模态模型成为具备跨场景联推能力的关键技术方向之一 [7,8]. 此外, 非常规链路, 如流星余迹通信 (meteor burst communication, MBC), 将在应急与偏远场景中发挥补充作用, 其随机可用性为多链路智能调度提出新的研究要求. 未来研究可从更加高效的语义增强机制、更稳健的环境先验建模以及更强的生成式推理能力入手, 迈向更加智能、可靠与韧性的 6G 网络 [9].

参考文献

1　Ding Z G, Schober R, Fan P Z, et al. Next generation multiple access for IMT towards 2030 and beyond. Sci China Inf Sci, 2024, 67: 166301

2　Huang Y Z, Zhang Z Y, Che J Z, et al. Self-attention reinforcement learning for multi-beam combining in mmWave 3D-MIMO systems. Sci China Inf Sci, 2023, 66: 162304

3　Zhang S, Zhang S, Yuan W, et al. Efficient rate-splitting multiple access for the internet of vehicles: federated edge learning and latency minimization. IEEE J Sel Areas Commun, 2023, 41: 1468–1483

4　Zhang S, Zhang S, Mao Y, et al. Transformer-based channel prediction for rate-splitting multiple access-enabled vehicle-to-everything communication. IEEE Trans Wireless Commun, 2024, 23: 12717–12730

5　Shen H, Zhu P C, Liu P C, et al. Finite blocklength MIMO transmission for URLLC: theory and technique. Sci Sin Inform, 2024, 54: 1055–1077 [沈弘, 朱鹏程, 刘平川, 等. 面向超可靠低时延通信的有限块长 MIMO 传输理论与技术. 中国科学: 信息科学, 2024, 54: 1055–1077]

6　Yang G H, He G N, Chen R R, et al. Progress and prospect of 6G wireless air-interface transmission technology research. Sci Sin Inform, 2024, 54: 1078–1113 [杨刚华, 何高宁, 陈睿荣, 等. 6G 无线空口传输技术研究进展与展望. 中国科学: 信息科学, 2024, 54: 1078–1113]

7　Wang F, Zhang S, Yang H, et al. Non-terrestrial networking for 6G: evolution, opportunities, and future directions. Engineering, 2025, 54: 56–68

8　Li Z, Shi J, Si J, et al. Intelligent covert communication: recent advances and future research trends. Engineering, 2025, 44: 101–111

9　Cheng N, Chen F J, Chen W, et al. 6G omni-scenario on-demand services provisioning: vision, technology and prospect. Sci Sin Inform, 2024, 54: 1025–1054 [承楠, 陈芳炯, 陈文, 等. 6G 全场景按需服务: 愿景、技术与展望. 中国科学: 信息科学, 2024, 54: 1025–1054]

https://www.sciengine.com/doi/10.1360/SSI-2025-0539