

|  |  |
| --- | --- |
| 姓 名： | 高艺泷 |
| 学 号： | 2425301072 |
| 学科 类别 领域： | 电子信息 |
| 指 导 教 师： | 郭彤颖 |
|  |  |
| 培 养 类 别： | □全日制 □非全日制 |
| 学 位 类 型： | □学术学位 □专业学位 |
| 开 题 日 期： |  |
| 所在培养单位： | 沈阳建筑大学 |

**沈阳建筑大学研究生院制**

**填表说明**

1．本表由研究生在导师指导下如实填写；

2．封面“学科类别领域”应严格按照招生录取时的学科类别领域名称填写；

3．封面“培养类别”和“学位类型”在相应框内划“√”；

4．表中所有签字部分，均须本人签字，不得代签；

5．除表中明确注明可附页的部分外，不得更改本表格式；

6．本表一式一份，全部内容填写及签字完成后，由研究生扫描成pdf文件，并将原件及扫描完成的电子版报所在培养单位存档。

|  |  |
| --- | --- |
| 论文题目 | 基于深度强化学习的多智能体通信与编队协同优化研究 |
| 课题来源 |  |
| **报告要求：**  开题报告字数应不少于5000字，阅读的参考文献应不少于20篇，其中外文文献不少于三分之一。  **内容提示：**  1．论文选题的背景、理论意义或应用价值  2．国内外研究现状及发展趋势  3．主要研究内容  4．理论依据或技术方案  5. 创新点和预期结论  6．论文工作计划  7．主要参考文献 一．论文选题的背景，理论意义或应用价值 自动驾驶技术的快速发展推动了自动驾驶车辆在实际生活中应用范围，具有辅助驾驶能力的私家车辆广泛进入大众生活中，并且在武汉运行的“萝卜快跑”项目也已经进行实测阶段，单车自动驾驶的前景一片可观，自动驾驶单车同时也引出了自动驾驶多智能体的想法构建，多智能体编队应用范围十分广泛，如货运汽车编队解决驾驶员长期行驶疲惫问题；公路自动驾驶私家车编队在拥挤路段统一调控，缓解拥堵问题；在搜索救灾方面，无人机+地面机器人编队协同搜索幸存者、投送物资、构建临时通信网；在农业领域多无人机协同播种、施肥、喷药，基于田间异质性动态调整队形；本文就高速公路多车编队进行通信和队列编排方面的讨论。传统人工驾驶编队依赖多名司机协同操作，不仅人力成本高，还面临驾驶员短缺和疲劳驾驶带来的安全隐患。同时，为确保安全，车辆之间需保持较大间距（通常超过50米），导致空气阻力增加，燃油经济性下降10%至15%。更关键的是，人类驾驶员平均1.5秒的反应时间难以实现精准同步，在紧急制动等场景下容易引发连锁事故或“幽灵堵车”，严重制约了运输效率和安全性；另一方面，在车流量较大的高速公路上，发生碰撞事故的可能性往往会大幅增加，在自动驾驶货车编队行驶的过程中，如若其中一条道路发生碰撞事故，往往需要紧急进行队列变换，但往往非编队车辆因其信息局限性，会选择较近或者只适用于自身最好情况的道路进行行驶，往往会对编队的重组进行影响，在理想情况下，假定每一套车辆都具有自动驾驶能力以及信号发射装置，在遇到紧急情况时，非编队车辆可以进入编队内进行统一调度。  传统延迟分为信号延迟和物理延迟[1]，信号延迟指车倒车或者车到基站之间的信号传输，所需要的时间，所带来的时滞性；物理延迟是机械的传输运动所带来的时间延迟。本文主要解决通信信号延迟。为突破这些瓶颈，基于车-车通信（V2V）的自动驾驶编队技术应运而生。通过实时共享车辆状态和协同控制，自动驾驶编队可将车距压缩至10到15米，显著降低风阻并提升道路通行能力。然而，其性能高度依赖V2V通信的稳定性与效率。在实际运行中，车辆频繁进出编队、高速移动导致通信拓扑动态变化，链路容易中断，实测丢包率超过30%。同时，有限的通信带宽在大规模编队中面临资源分配难题[2]：集中式调度计算延迟高，难以满足实时性要求[3]；而分布式方法受限于单个车辆只能获取邻近车辆信息，缺乏全局视野，容易做出局部次优决策[4]。  此外，高速公路环境中的多径效应和同频干扰使无线信道质量剧烈波动，现有基于固定规则的资源分配策略频谱利用率不足40%[5]，难以适应复杂动态场景。尽管强化学习和图神经网络等方法被引入解决这一问题，但仍存在明显局限。例如，强化学习常依赖全局信息假设，在实际局部观测条件下性能下降；图神经网络虽能建模车辆间关系[6]，但传统静态图结构无法反映编队的动态重组过程，也忽略了通信链路的方向性和质量差异。  因此，亟需构建一个通信感知的协同编队系统：一方面，通信层需根据当前编队结构与干扰环境，动态分配信道与功率，保障安全消息高可靠传输；另一方面，控制层需感知通信链路的可靠性，自适应调整队形策略（如切换为更鲁棒的串行结构）或控制参数（如增大安全间距），在通信受限条件下仍维持编队稳定性与能效优势。  本研究正是在此背景下，提出基于深度学习的通信–控制协同优化框架。在已有 V2V 通信资源分配工作（GATV2+DDQN）基础上，进一步引入多智能体深度强化学习与动态图建模，将车辆编队视为一个通信–物理耦合的智能体网络，通过统一的状态表示与联合优化目标，实现“通信为控制服务，控制反哺通信需求”的闭环协同。该框架不仅适用于封闭高速公路货运场景，还可扩展至城市无人配送、港口自动集卡等高密度、高动态的智能物流系统，为车路云一体化协同提供可落地的技术范式。  在蜂窝车联网侧，将 GAT 与 DDQN 结合可在百车以内把 V2I 总速率再抬升 3–5 %，同时维持 V2V 安全消息成功率 95 % 以上，但高密度场景仍需 GAT+DDQN 保持 20 ms 级实时决策；换到自动驾驶编队，GAT提供每辆车的速度，位置，相对距离，高层为结构层，DDQN 选择队列结构操作[7]，例如目标前车ID/换位/是否并队/子队大小/目标时距档位；低层轨迹层，SAC 输出连续纵向加速度；CTDE（集中训练、分布执行），训练时可以集中收集所有车的经验；执行时每车独立决策。 二．国内外研究现状及发展趋势 自动驾驶货车编队（Autonomous Truck Platooning）作为智能交通系统（ITS）与绿色物流的关键使能技术，近年来受到学术界与工业界的广泛关注。其核心目标是通过车–车协同控制，将多辆货车以紧凑间距（通常 10–15 米）高速行驶，从而显著降低空气阻力、提升燃油效率（10%–15%）并增强道路通行能力。然而，该技术的规模化落地面临两大核心挑战：高动态环境下 V2V 通信的可靠性保障 与 复杂交通场景中队列结构的自适应优化。二者相互耦合，共同决定了编队系统的安全性、能效性与鲁棒性。  在 V2V 通信资源分配 方面，传统方法多基于博弈论、凸优化或启发式调度，依赖精确的信道状态信息（CSI）与静态网络假设[8]。然而，在高速公路场景中，车辆高速移动（>80 km/h）导致多普勒频移与多径效应剧烈，CSI 难以准确获取，且通信拓扑随车辆进出编队动态重构，使得传统方法性能急剧下降。  近年来，深度强化学习（DRL）被引入解决该问题。Zhengda Huan1, Jian Sun等人采用分布式 DQN 实现信道选择[9]；蒋立伟提出结合 DDPG 实现连续功率控制[10]。然而，纯 DRL 方法在局部观测下易陷入次优解，且对环境噪声敏感。为克服此局限，图神经网络（GNN）被用于建模车辆间干扰关系。HW与YW 提出 Graph-DDPG，在 100 车规模下降低跟车误差 25%[11]。本文前期工作进一步提出 GAT + DDQN 框架，通过构建“基于通信关系的动态图”，将邻居数恒定在 ≈12，显著降低计算复杂度，并在 20 ms 内完成资源分配决策，保障 V2V 安全消息成功率 ≥95%。该成果为后续编队控制提供了可靠的通信底座。  与此同时，车队编队控制研究正从固定结构向动态自适应演进。早期工作多基于集中式或 leader–follower 架构的 CACC（Cooperative Adaptive Cruise Control），依赖理想通信假设，难以应对链路中断或拓扑变化[12]。为提升鲁棒性，多智能体强化学习（MARL）成为主流方向。杨等人开发了一种考虑不同驾驶策略的基于强化学习的决策方法徐等人提出了一种确保安全的安全强化学习算法，该算法将强化学习算法与人工势场（APF）及轨迹跟踪方法相结合，通过加权输出动作，但该方法仅在低障碍物场景中表现良好，未针对高复杂度场景进行验证[13]。近期，研究聚焦于混合动作空间与分层决策架构。例如，QMIX 被用于协同变道，MADDPG 学习连续加速度控制。ML,BDY提出了使用LLM agent来辅助深度学习进行模型训练[14]。然而，现有 MARL 方法多假设完美通信，未显式建模通信质量对策略性能的影响。更关键的是，队列结构本身（如串行、双列、子队划分）作为高层决策变量，其优化常被忽略或简化为规则驱动，缺乏数据驱动的自适应能力。  当前，国际研究正加速向大规模、异构、动态拓扑 场景演进。本文提出的“恒定邻居采样 + 滞后奖励软标签 + 分层 ”方案，有望在 50–300 车规模下实现 ≤20 ms 端到端延迟，同时保障通信成功率与编队稳定性，为智慧高速与无人物流提供可复制的技术路径。 三.主要研究内容与目标1.研究目标 本研究面向高密度高速公路自动驾驶多车编队场景，聚焦 V2V 通信可靠性与编队结构自适应性的双重挑战，提出一种 基于深度学习的通信–控制协同优化框架。在前期 GAT+DDQN 通信资源分配工作的基础上，进一步构建图神经网络驱动的分层多智能体编队系统，实现“通信为控制服务，控制反哺通信需求”的闭环协同针对多车编队中的编队在紧急情况下的队列结构稳定性问题提出了解决，在车车通信加入/离开、信道快速时变的环境下，在车辆动态加入/退出、信道快速时变的复杂环境下，V2X成功率高且时延降低，在复杂情况下车队编队结果最优，排队长度较短，总通过时间较短，实现通信与编队的双重优化。 2.研究内容 本研究围绕自动驾驶货车编队中通信可靠性与队列结构自适应性的耦合问题，构建一种基于深度学习的通信–控制协同优化框架，具体研究内容包括以下几个方面：   * 动态图建模与嵌入表示   针对智能体高速移动导致拓扑频繁变化的特点，构建以智能体为节点、干扰或通信关系为边的稀疏动态图，边权重由距离或信道状态动态确定。在此基础上，采用GATv2进行图嵌入学习，其动态注意力机制能自适应区分邻居重要性，有效聚合局部观测与全局结构信息，生成低维、判别性强的节点表示，支持归纳式推理与规模泛化。   * 通信资源智能分配   将GATv2输出的嵌入向量与本地状态（如信道增益、队列信息）融合，作为DDQN的状态输入。每个智能体通过独立DDQN网络选择信道与功率等资源动作，奖励函数综合考虑通信成功率、对其他链路的干扰及能耗，实现分布式、高鲁棒的决策。该框架在保留完全分布式架构的同时，借助图神经网络弥补局部观测局限，提升协作效率。   * 强化学习队列变换与通信—变道协同调整   为实现队列结构自适应优化与实时变道，构建分层强化学习架构：高层基于DDQN每秒进行离散结构决策（如目标前车选择、并/拆队、子队规模及时距档位），低层采用SAC算法以0.1秒间隔输出连续加速度指令，最小化间距与速度误差；两层共享GATv2生成的嵌入向量，实现协同。同时，建立通信与控制的双向耦合机制——通信调度优先保障控制关键链路，控制决策显式感知通信状态（时延、AoI等），并在信息受限时保持鲁棒性；引入“安全盾”机制，在临近安全边界时自动修正动作，确保防撞与队列稳定性。    图一 仿真虚拟环境图 理论依据或技术方案1通信资源智能调度（基于图神经网络与深度强化学习） 本系统采取基于图神经网络的联合决策架构，整体流程如图所示。  首先，输入端将每辆车的节点特征（如位置、速度、加速度等）、边索引（邻接关系）、边特征（如相对距离、信道质量等）统一输入到GATv2编码器。编码器由多层多头图注意力模块组成，每层配备残差连接和归一化/Dropout机制，使其能高效聚合邻域信息，突出关键节点和链路，最终输出各节点的高维嵌入。  在GATv2编码器后，系统通过两个分支进一步提取结构信息。Edge head分支将不同节点的嵌入和边特征拼接，经过MLP网络，输出链路的综合评分，为后续通信资源调度提供依据。Graph head分支则采用池化和多层感知机，将全队车辆状态聚合为全局摘要，用于描述编队整体的稳定性和交通密度。  随后，状态融合模块将节点嵌入、链路评分、全局摘要与物理控制量（如速度、加速度）、通信统计（如AoI等）合并，形成完整的状态向量，作为DDQN决策网络的输入。DDQN部分采用多层MLP结构，结合经验回放和目标网络软更新机制，生成每一动作的Q值，利用ε-贪婪策略在探索与利用之间动态权衡，输出最优动作，如信道选择，发射功率控制，消息是否上报，功率带宽的选择。    图2 GATV2框架图，以及核心公式     | 符号 | 维度/类型 | 含义 | | --- | --- | --- | |  | 索引 | 图节点（链路/链路节点）索引 | |  | 集合 | 节点 的入邻居集合（常为 Top-K） | |  |  | 节点原始特征（本地信道、干扰、负载/时限、历史 RB 等） | |  |  | 节点投影矩阵 | |  |  | 投影后的节点表示 | |  | - | 向量按特征维拼接 | |  |  | 注意力网络权重 | |  | - | 非线性激活（如 ELU、LeakyReLU） | |  |  | 注意力读出向量 | |  | 标量 | 邻居 对目标 的未归一化打分（logit） | |  | 标量 | 温度参数，越小越聚焦，越大越平均 | |  | 标量 | 归一化注意力权重， | |  |  | 目标节点的新嵌入 | |  | 整数 | 注意力头数（多头时） |   在车—链路图上自适应分配邻居权重，得到判别力更强的节点嵌入 ，作为 DQN 的增强状态。  节点线性映射： ，把链路节点的原始特征 （如本地信道、干扰、剩余负载/时限等）投影到统一的隐空间，便于后续邻居信息融合  注意力打分：，用“中心节点 j 与候选邻居 i 的拼接向量”做一次非线性变换后，再用向量取内积得到边权原始分数。温度参数 τ>0 控制分布“尖/平”：τ 越小，分布越尖锐（更聚焦少数邻居）；越大越平均。这一步使“与谁相似/与谁更相关”能被网络自动学习，而不是用固定的距离或度数。  归一化：，对同一个目标节点 j 的所有入邻居 i 做 softmax，得到权重，并保证求和为 1。  聚合:，用学到的权重加权求和邻居消息，形成 j 的新表示（可再叠层、做多头、加残差和层归一化）  自适应“邻居选择与加权”  选择权重更大的邻居智能体通常是那些对本车有强干扰、竞争相同 RB、或有协同信息（同车/同群）的节点。注意力将这些关系从数据中学出来，用更大的 强化其影响。  将复杂拓扑压缩为“可用的低维状态”  把拓扑、信道、干扰、时限等上下文揉合成一个定长向量，比人工拼接邻居特征更稳健；作为 DQN 的状态增强量，有利于在大规模/动态网络下快速决策。  抗噪与泛化  softmax 归一化抑制噪声邻居；多头注意力可在不同“关系子空间”（如几何接近/干扰强度/业务紧急度）上分别建模，提升泛化到新场景、新车流。  可解释性  可直接读出“决策时谁影响了车辆自身”，便于分析策略是否合理（例如，确实更关注同 RB、近距离、强干扰的邻居） 2车队队列结构优化（分层强化学习 + 图编码 + 安全约束） 本系统针对多车编队场景，采用了图神经网络与分层强化学习相结合的结构优化方法。首先，通过采集车辆物理状态、通信统计以及邻接关系等信息，利用GATv2对全队进行图结构编码，提取关键节点嵌入及全局摘要，为后续队列结构优化和动态编队调整提供全面的数据基础。  在高层决策部分，系统以融合后的节点信息、物理量和通信统计为输入，利用DDQN网络评估结构调整方案，并通过可微分指派与安全约束机制，实现队位分配和目标时距设定。低层控制环节结合安全盾机制，对加速度、变道等指令进行实时过滤，确保车辆执行过程中的队形稳定与安全性。回放与奖励机制则根据编队形成效率、队列稳定性、通信性能等指标持续优化整体系统性能。    图三 DDQN 分层HRL控制图  描述车队/编队中单车是否应当执行变道（Lane Change, LC）。它同时考虑安全可行与收益是否值得，便于作为上层策略的动作掩码与奖励先验。    自车及邻车的“期望加速度”（由同一跟驰模型/控制器给出）：  ：自车变道前/后对其“新前车”的期望加速度  ：目标车道“自车后车 follower”的期望加速度（并入前/后）  ：原车道“自车后车”的期望加速度（离开前/后）  收益项：  （自车收益）  （目标车道后车收益，通常≤0）  （原车道后车收益，通常≥0）  常数/超参：  ：允许的最大舒适减速度（正数，2–3 m/s² 常用）  ：礼让（politeness）系数，越大越在意他人收益/损失  ：变道固有成本（稳定性/舒适/风险）  ：触发阈值，抑制小收益导致的频繁变道  ：指示函数；条件为真取 1，否则 0  运算流程  1计算几何可行（车道目标前/后车距 ≥ 安全车距），不满足直接不变道。  2用 计算 ，得到安全门控与收益项。  3评估：先安全、后收益。  4将安全门控作为上层 DDQN 的动作掩码；将“收益项”并入奖励或作为价值先验。  下层连续控制（如 SAC）在执行期再用 CBF/QP 做硬安全投影。 3通信—编队协同与系统集成 本系统针对多车编队场景，集成了图神经网络与分层强化学习，实现队列结构优化与安全控制。首先，系统输入包括车辆物理量（如车距d、相对速度Δv、加速度a、时距档位T），通信统计、以及邻接/可见集合和车道信息。这些状态信息通过concat操作整合为图输入，或直接作为节点/边特征送入GATv2编码器，参数包括编码层数L、注意力头数K等（如L=2, K=4~8），输出每辆车的节点嵌入 。  在高层结构优化部分，系统以节点嵌入、物理量和通信统计融合为高层状态向量，通过MLP网络输出Q值，采用ε-贪婪策略，动作空间如队列操作、目标时距选择等），实现结构指令的智能规划。队位分配采用可微指派机制，结合安全约束（如最小车距=0.6s、最大队长=10）、队形与交通规则，输出目标队位及时距参数T\*。  低层控制模块以节点嵌入、队列误差项和目标为输入，采用SAC实现纵向加速度及变道决策。所有控制指令均通过安全盾（控制屏障函数CBF/规则裁剪）进行实时过滤，确保车辆执行阶段满足安全车距、速度和队形要求。环境执行后，系统记录每步状态、奖励与回放数据，分为高层与低层两类，支持经验回放与多层训练。最终，系统根据编队形成时间（formation time）、字符串稳定性、加加速度jerk、碰撞率、通信性能等指标给予奖励，关键参数如formation time<20s、jerk均值<0.8m/s³、AoI均值<0.5s，持续优化整体性能，实现高效可靠的编队队列结构协同。    图4 通信——编队协同极致图 五．创新点与预期结论1.创新点  (1)以GATv2进行车联网图表示，使用GATV2作为局部信息的处理结合DDQN实现控制感知的通信资源调度，优先保障关键车辆与关键链路的数据新鲜度（AoI）与时延指标。  (2) 提出分层强化学习（HRL）结构对编队队列进行动态重构：高层DDQN负责离散的结构决策，低层SAC输出连续横向，纵向控制，二者共享图嵌入并解耦时间尺度，提升训练稳定性与在线可执行性。  (3) 在队位分配与插队问题上引入可微分指派思想和Top‑K候选前驱约束，实现“全局一致性”的位置分配与低重构代价的序列调整，缓解局部贪心导致的次优与抖动。 2.预期结果 (1) 与基线方法（固定RB/功率、仅通信优化、仅控制优化、传统ACC/CACC/MPC）相比，本方案在同等带宽预算下显著降低端到端时延与丢包率，提高有效邻接率与减小信息年龄（AoI），从而改善控制侧的时距与速度跟踪性能。  (2) 在中高流量与受干扰场景中，编队形成时间缩短，平均时距误差与相对速度误差明显下降，受影响区段车辆全部通过的总时间减小，每小时/每分钟通过事故区或关键路段的车辆数增大。 六．论文工作计划 第一阶段：阅读自动驾驶编队、V2X通信、图神经网络（GNN）、多智能体强化学习（MARL）等相关领域的国内外文献，重点掌握编队控制、通信资源分配、队列结构优化等关键技术。学习Python编程语言及PyTorch深度学习框架，掌握SUMO与Carla仿真平台的基本使用方法，搭建仿真环境，完成系统建模与信道数据采集。预计3个月完成。（2025年3月——2025年5月）  第二阶段：深入学习图注意力网络（GATv2）、Double DQN（DDQN）、Soft Actor-Critic（SAC）等算法，完成通信资源调度模块（GATv2+DDQN）与分层强化学习控制模块（高层DDQN+低层SAC）的设计与实现。实现CTDE（集中训练、分布执行）框架的集成，完成动态图建模、嵌入表示、Top-K邻居采样、事件触发机制等核心功能模块的开发。预计7个月完成。（2025年5月——2025年12月）  第三阶段：在tensflow平台中开展稳态与动态交通场景下的实验测试，验证所提方法在通信成功率、编队稳定性、能耗与任务完成时间等方面的性能。对比基线方法（，优化模型超参数与网络结构，完成消融实验分析各模块贡献。撰写小论文。预计6个月完成.（2025年12月——2026年6月）  第四阶段：对大论文进行撰写，并进行整体式修改，查阅资料，补充细节最终形成书面材料，装订成册，并准备参加毕业答辩。预计 6个月完成。（2026年6月——2027年1月） 七．参考文献 [1] 王旭, 张春波, 兰翔, 等. 通信时延对自动驾驶车队稳定性的影响研究[J]. 石家庄铁道大学学报（自然科学版）, 2024, 37(4): 77-83.  [2] Ren Y, Wu C, So D K C, et al. DRL-Based Joint Aggregation Frequency and Edge Association for Energy-Efficient Hierarchical Federated Learning[J]. IEEE Transactions on Wireless Communications, 2025, 24(8): 6841-6856.  [3] Liang Z, Cao J, Jiang S, et al. Hierarchical Reinforcement Learning with Partner Modeling for Distributed Multi-agent Cooperation[J]. IEEE Transactions on Parallel and Distributed Systems, 2024: 1-13.  [4] 刘志鹏. 基于C-V2X的车队协同资源分配方案研究与实现[D]. 北京: 北京邮电大学, 2024.  [5] 况申旭. NR-V2X侧链通信无线资源分配关键技术研究[D]. 上海: 东华大学, 2025.  [6] 勉海荣, 焦小刚, 毕利. 基于图注意力强化学习的电动自动驾驶运营车队实时控制[J]. 交通运输工程与信息学报, 2025.  [7] Fan B, Xie H, Li T. Platoon communication power control under V2V data uncertainty: a robust DRL approach[J]. IEEE Transactions on Intelligent Transportation Systems, 2025: 1-15.  [8] 谢海文. 多智能体分布式协同任务分配策略设计[D]. 北京: 北方工业大学, 2025.  [9] Huan Z, Sun J, Chen Z, et al. Resource allocation in V2X networks: a double deep Q-network approach with graph neural networks[J]. Computers, Materials & Continua, 2025, 84(3): 5427-5443.  [10] 蒋立伟, 王满江, 邱千, 等. 基于封闭场景自动驾驶的DDPG决策规划方法[J]. 专用汽车, 2025(10): 54-58.  [11] Wang Y, Wu H, Li R. Deep graph reinforcement learning for mobile edge computing: challenges and solutions[J]. IEEE Network, 2024, 38(5): 314-323.  [12] Wang Z, Wu G, Barth M J. A review on cooperative adaptive cruise control (CACC) systems: architectures, controls, and applications[C]//2018 21st International Conference on Intelligent Transportation Systems (ITSC). Maui, HI: IEEE, 2018: 2884-2891.  [13] Yang H, He Y, Xu Y, et al. Collision avoidance for autonomous vehicles based on MPC with adaptive APF[J]. IEEE Transactions on Intelligent Vehicles, 2024, 9(1): 1559-1570.  [14] Yildirim M, Dagda B, Fallah S. HighwayLLM: Decision-Making and Navigation in Highway Driving with RL-Informed Language Model[A]. arXiv, 2024.  [15] Garcia M H C, Molina-Galan A, Boban M, et al. A tutorial on 5G NR V2X communications[J]. IEEE Communications Surveys and Tutorials, 2021, 23(3): 1972-2026.  [16] Ji M, Wu Q, Fan P, et al. Graph Neural Networks and Deep Reinforcement Learning Based Resource Allocation for V2X Communications[A]. arXiv, 2025.  [17] Gyawali S, Xu S, Qian Y, et al. Challenges and solutions for cellular based V2X communications[J]. IEEE Communications Surveys and Tutorials, 2021, 23(1): 222-255.  [18] Yuan K, Huang Y, Yang S, et al. Evolutionary Decision-Making and Planning for Autonomous Driving: A Hybrid Augmented Intelligence Framework[J]. IEEE Transactions on Intelligent Transportation Systems, 2024, 25(7): 7339-7351.  [19] Cho S, Lim S, Lee J. DRL-Enabled Hierarchical Federated Learning Optimization for Data Heterogeneity Management in Multi-Access Edge Computing[J]. IEEE Access, 2024, 12: 147209-147219.  [20] Hu Y. Improving the Stability of Multi-hop IoT Communication Paths Using Graph Neural Networks[C]//2025 4th International Conference on Electronics, Integrated Circuits and Communication Technology (EICCT). Chengdu, China: IEEE, 2025: 354-358. | |

注：本表可另附页。

|  |
| --- |
| 本人签字：  年 月 日 |