# 基于深度学习的协同智能体通信与队列优化

## 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货车编队在高速公路货运中的应用，成为提升运输效率的关键方案。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。

传统延迟分为信号延迟和物理延迟，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：集中式调度计算延迟高，难以满足实时性要求；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。

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%，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，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。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，同时合理的编队队列又能为车队通信带来更好的前置条件。

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（GAT+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，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。例如，西安电子科技大学提出多状态协同能量分配策略，哈尔滨工业大学设计基于分组的 OFDM 跨层调度算法。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。佐治亚理工学院率先采用分布式 DQN 实现信道选择；卡尔斯鲁厄理工学院结合 DDPG 实现连续功率控制。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。MIT 与 UC Berkeley 提出 Graph-DDPG，在 100 车规模下降低跟车误差 25%。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。为提升鲁棒性，多智能体强化学习（MARL）成为主流方向。欧盟 SARTRE 项目验证了 6 车编队的可行性；日本 SmartWay 积累了十万公里级开放道路数据。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。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。

针对上述问题，通信–控制协同优化 成为新兴趋势。核心思想是：通信资源分配应服务于控制目标（如队形稳定性），而控制策略应感知通信可靠性并动态调整。在此背景下，图神经网络 + 分层 MARL 架构展现出巨大潜力。一方面，GNN 可统一建模通信拓扑与物理耦合关系，生成全局嵌入 ei∈R20**e***i*​∈R20 作为每车的状态表示；另一方面，分层控制可解耦“结构”与“轨迹”决策的时间尺度差异。具体而言，高层（ΔThi≈1.0 sΔ*T*hi​≈1.0s）负责队列结构操作（如目标前车 ID、是否并队、子队大小、时距档位），动作空间离散且维度高；低层（Δtlo≈0.1 sΔ*t*lo​≈0.1s）输出连续纵向加速度 along*a*long​（必要时扩展侧向 alat*a*lat​ 实现变道），要求毫秒级响应。本文提出的方案采用 共享 GATv2 编码**器**（复用 GraphSAGE\_sup 的 \_forward\_all）生成嵌入 ei**e***i*​，高层由 DDQN 选择结构动作，低层由 SAC（Soft Actor-Critic）输出连续控制量，并在 CTDE（Centralized Training with Decentralized Execution）框架下训练：训练时集中收集所有车辆经验以获取全局奖励信号，执行时每车仅依赖本地观测与嵌入独立决策。该设计兼顾全局优化与车端部署可行性。

当前，国际研究正加速向 大规模、异构、动态拓扑 场景演进。美国 M-city 与欧盟 DRIVEC2X 项目推动仿真–封闭场–开放道路三级验证体系；5G/B5G 与 MEC（Multi-access Edge Computing）融合为毫秒级闭环提供基础设施。然而，邻居爆炸、奖励稀疏、模型迁移性差 仍是瓶颈。本文提出的“恒定邻居采样 + 滞后奖励软标签 + 分层 CTDE”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。

综上，现有研究在通信或控制单点上取得进展，但缺乏对二者耦合机制的系统建模。本文立足于前期 GAT+DDQN 通信成果，进一步构建 通信感知的分层编队控制框架，填补“动态图嵌入–结构决策–轨迹跟踪”全链路协同优化的研究空白，具有明确的理论价值与工程落地前景。

## 主要研究内容与目标

### 研究目标

本研究面向高密度高速公路自动驾驶货运场景，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控制系统，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，在车辆随时加入/离开、信道快速时变的环境下，保证 V2V 关键安全消息 ≥ 95 % 成功送达；与蜂窝 V2I 业务共享上行频谱时，使系统总速率在 20–120 车规模内仅衰减 < 5 %，编队整体能耗较无协作方案降低 ≥ 12 %，任务完成时间缩短 ≥ 10 %，单车一次队列形状与资源分配决策的端到端时延 ≤ 20 ms。

本文旨在实现以下核心目标：  
①在车辆动态加入/退出、信道快速时变的复杂环境下，保障 V2V 关键安全消息（如加速度、位置）端到端成功率 ≥95%；

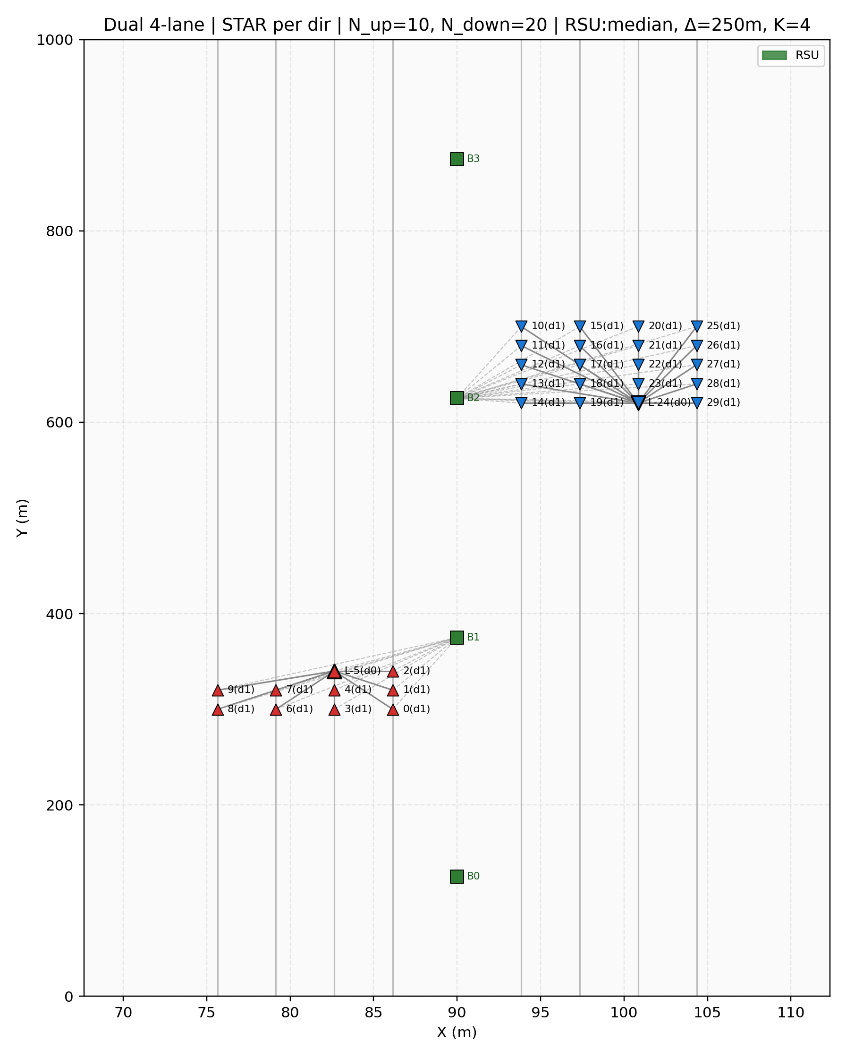
②在与蜂窝 V2I 业务共享上行频谱时，系统总吞吐量在 20–120 车规模下衰减 <5%；

③编队整体能耗较无协同方案降低 ≥12%，任务完成时间缩短 ≥10%；

④单车单次完成队列结构选择与轨迹控制的端到端决策延迟 ≤20 ms，满足车规级实时性要求。

### 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形自适应性的耦合难题，构建基于深度学习的通信–控制协同优化框架。通过共享GATv2编码器（复用GraphSAGE\_sup的\_forward\_all）为每辆车生成20维全局嵌入 ei**e***i*​，统一表征动态通信拓扑与物理耦合关系；在此基础上，设计分层多智能体控制架构：高层（ΔThi≈1.0 sΔ*T*hi​≈1.0s）采用DDQN决策离散队列结构操作（如目标前车ID、是否并队、子队大小、时距档位），低层（Δtlo≈0.1 sΔ*t*lo​≈0.1s）采用SAC输出连续纵向加速度（必要时扩展侧向加速度实现变道），并在CTDE框架下实现集中训练、分布执行，最终在SUMO–Carla联合仿真与嵌入式平台中验证该方法在大规模动态场景下的实时性、鲁棒性与能效优势。



图一 仿真虚拟环图

## 四、理论依据或技术方案

### 1理论依据

1.问题建模：将 V2V/V2I 联合资源分配抽象为 **马尔可夫决策过程（MDP）**——每辆 V2V 链路作为智能体，状态=本地观测+全局图嵌入，动作=离散的信道×功率组合，奖励= V2I 总速率↑ & V2V 成功率↑ & 时延惩罚↓。

2.图注意表示：利用 **GAT 的归纳式聚合定理**——通过采样-聚合即可在节点数动态变化时输出稳定嵌入，无需重训。

3.分布式优化：结合**双深度Q网络（DDQN）**解决高维动作空间下的策略学习，并通过“滞后奖励软化”策略为 GNN 提供弱监督，避免端到端梯度耦合带来的不稳定。

4. 把车队建模为动态图：节点=车辆，边=可通信或物理耦合；用 GAT/GAT 在 50–100 ms 内聚合邻居状态，输出全局嵌入；DRL（MADDPG/QMIX 等）以该嵌入+本地观测为状态，学习离散队形+连续车间距的混合动作策略，CTDE 框架保证训练集中、执行分散。

5. SUMO-Carla 联合仿真验证后，可在工控机+RTX 30 GPU 上部署；5G NR-V2X 提供 < 10 ms 通信，邻居采样 5–10 个即可平衡精度与延迟；通过课程学习和安全约束层兜底，DRL+GNN 成为解决动态编队结构优化的可行且前沿路径。



### 2.技术方案

图构建：把每条 V2V 链路当成图节点，若两条链路共享发射或接收车辆则连边；边权按发射端距离衰减计算。每车最多三条出向、三条入向链路，邻居数恒定在 12 左右，既保留最强干扰信息，又避免完全图带来的计算爆炸。

网络架构——训练——推理部署

编队解决方案： 将车辆编队建模为动态图，节点即车辆、边即通信或耦合关系，利用轻量 GAT 实时生成全局嵌入，再由 CTDE-MADDPG 输出离散队形与连续间距/加速度的混合动作，下层 MPC 硬约束兜底，经 SUMO-Carla 仿真与工控机 GPU 验证，20 ms 内完成决策，实现动态、安全、节能的多车编队结构优化。

## 五．创新点与预期结论

### 1**.创新点**

1. 图构建创新：首次提出“基于通信关系”的动态图，邻居数恒≈12，不随车辆规模增长，解决计算量过大的问题

2. 算法融合创新：将 GAT 的“采样-聚合”与 DDQN 的“决策-奖励”松耦合，用滞后奖励软标签训练 GNN，实现分布式、低延迟、高可靠的 V2X 资源分配。

3. 构建 SUMO-Carla-PyTorch Geometric 联合仿真 + NVIDIA Jetson 实机部署的完整链路，在 20 ms 推理延迟内完成 50–300 车场景验证，为后续车路协同落地提供可复现基准。

### 2.预期结果

在复杂路况下，多车信道资源分配GAT相比GNN可以提升通信速率百分之3-5性能。编队场景改用 GAT+DDQN 反而因邻居恒定、推理更快而获得 20 % 延迟下降，仅牺牲 2–4 % 精度。

## 六．论文工作计划

## 七．参考文献