



沈阳建筑大学
SHENYANG JIANZHU
UNIVERSITY

基于深度强化学习的多智能体通信与协同编队控制研究



汇报人：高艺泷



指导老师：郭彤颖



目录 CONTENTS

01

研究背景及意义

02

国内外研究现状

03

研究内容和依据

04

创新点

05

论文工作计划

06

参考文献

Part.01

研究背景及意义



1.1 研究背景

研究背景概述

随着交通流量的不断增长，高速公路面临的拥堵问题日益严重，尤其在发生交通事故时，常常引发长时间的“幽灵堵车”，导致交通效率大幅下降。传统人工驾驶难以迅速响应并有效缓解此类突发状况，主要体现在以下几个方面：①反应迟缓 ②资源浪费 ③管理难度大

对于高速公路场景下的通信延迟与堵塞带来的经济损失和能源消耗问题，运用基于深度强化学习的通信-编队协同优化技术可以为编队系统赋予更高水平的协同智能与安全保障。



为降低安全风险，京哈高速沈山段创新启用了“无人机叫醒”服务。



1.2 研究意义

同时本项研究还可以在其他多智能体领域应用如：

①城市无人配送车队



②港口自动化集装箱卡车队列



③农业喷药无人机队列



④仓储物流机器人队列

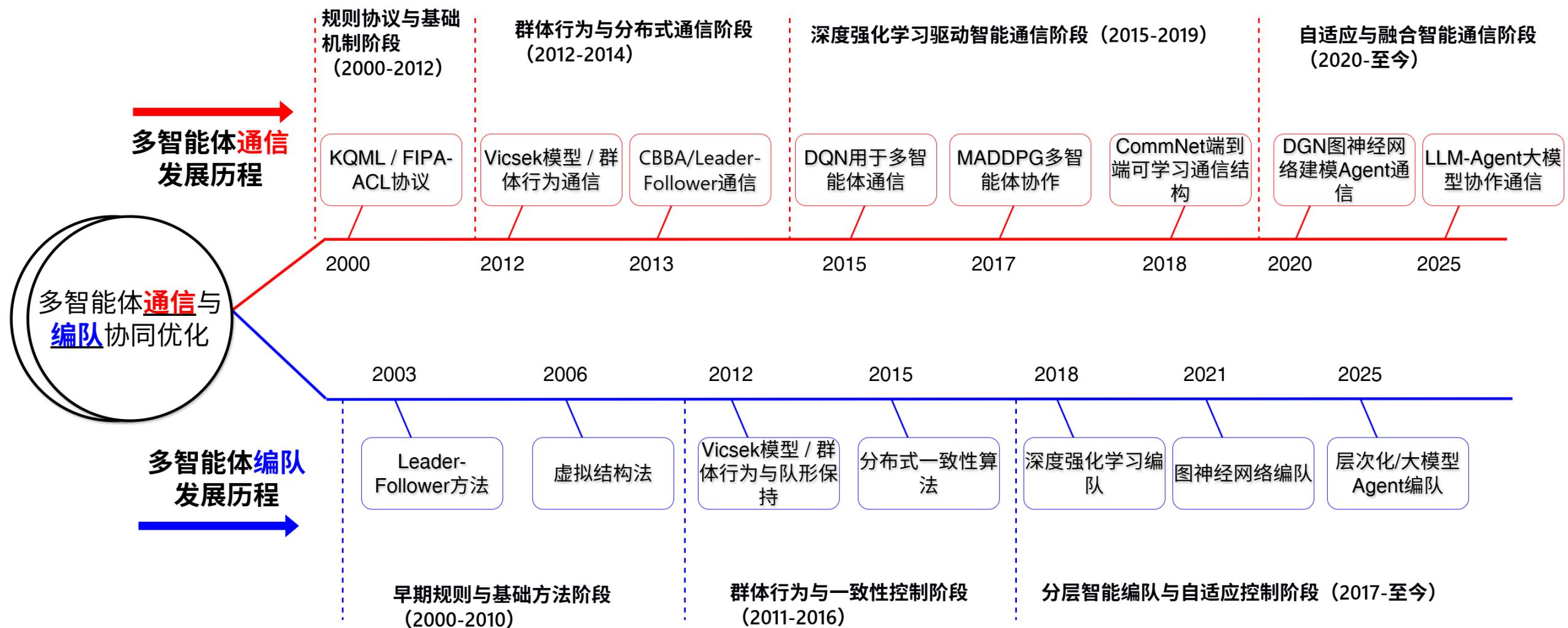


Part.02

国内外研究现状



2.1 国内外研究现状



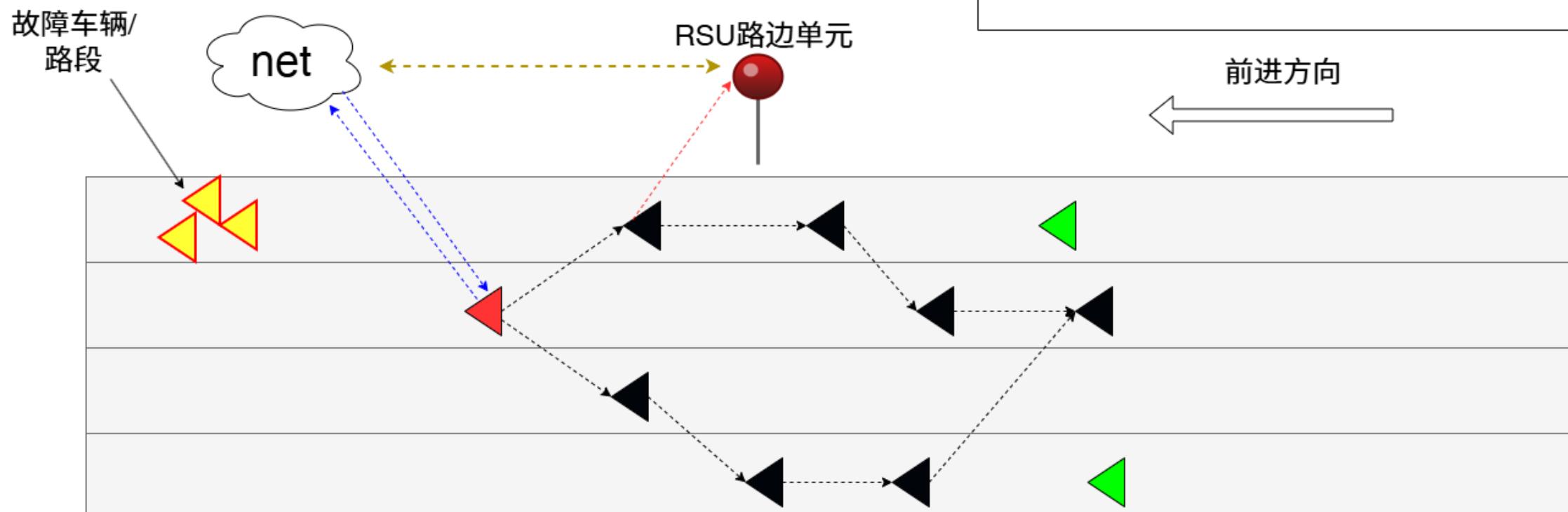
Part.03

研究内容和依据



3.1 研究内容

拟定实验部分环境



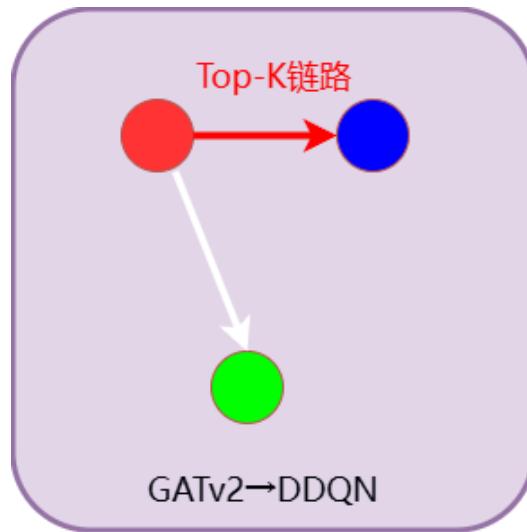
树状拓扑队列



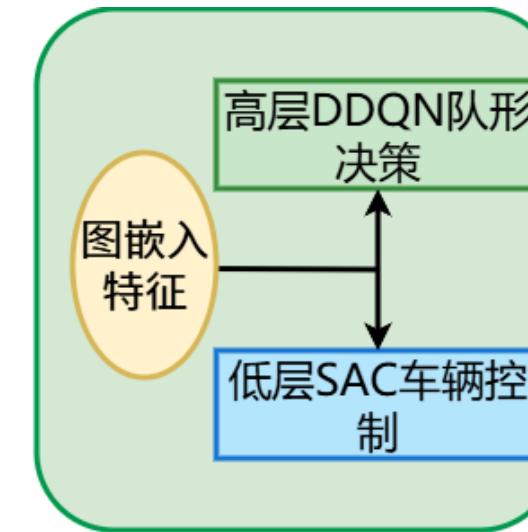
3.1 研究内容

研究内容

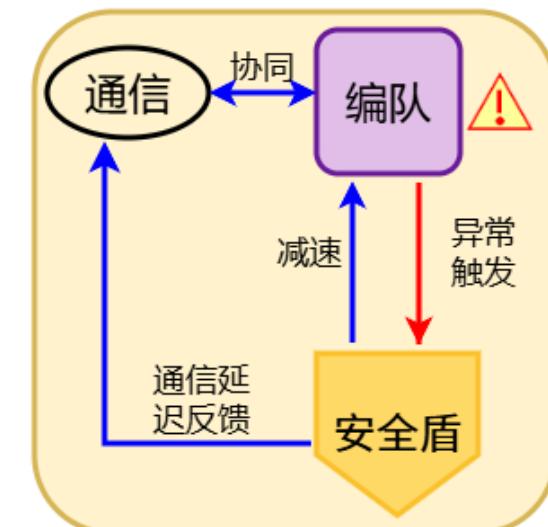
- 通信资源调度流程 (GATv2→DDQN+Top-K)
- 分层优化结构 (高层DDQN←共享嵌入→低层SAC)
- 协同集成逻辑 (通信↔编队→安全盾)



通信资源调度流程



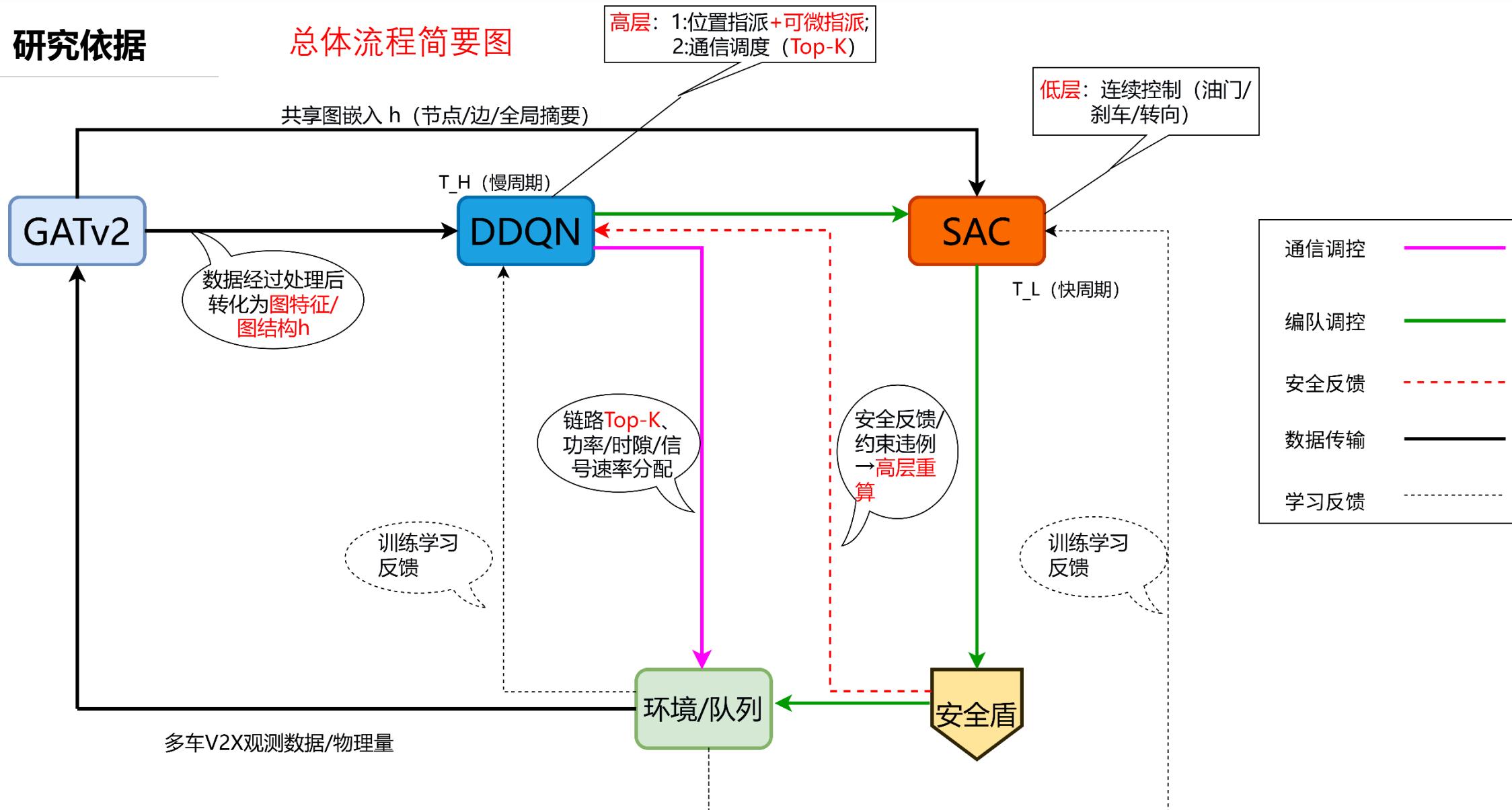
分层优化结构



协同集成逻辑

3.1 研究依据

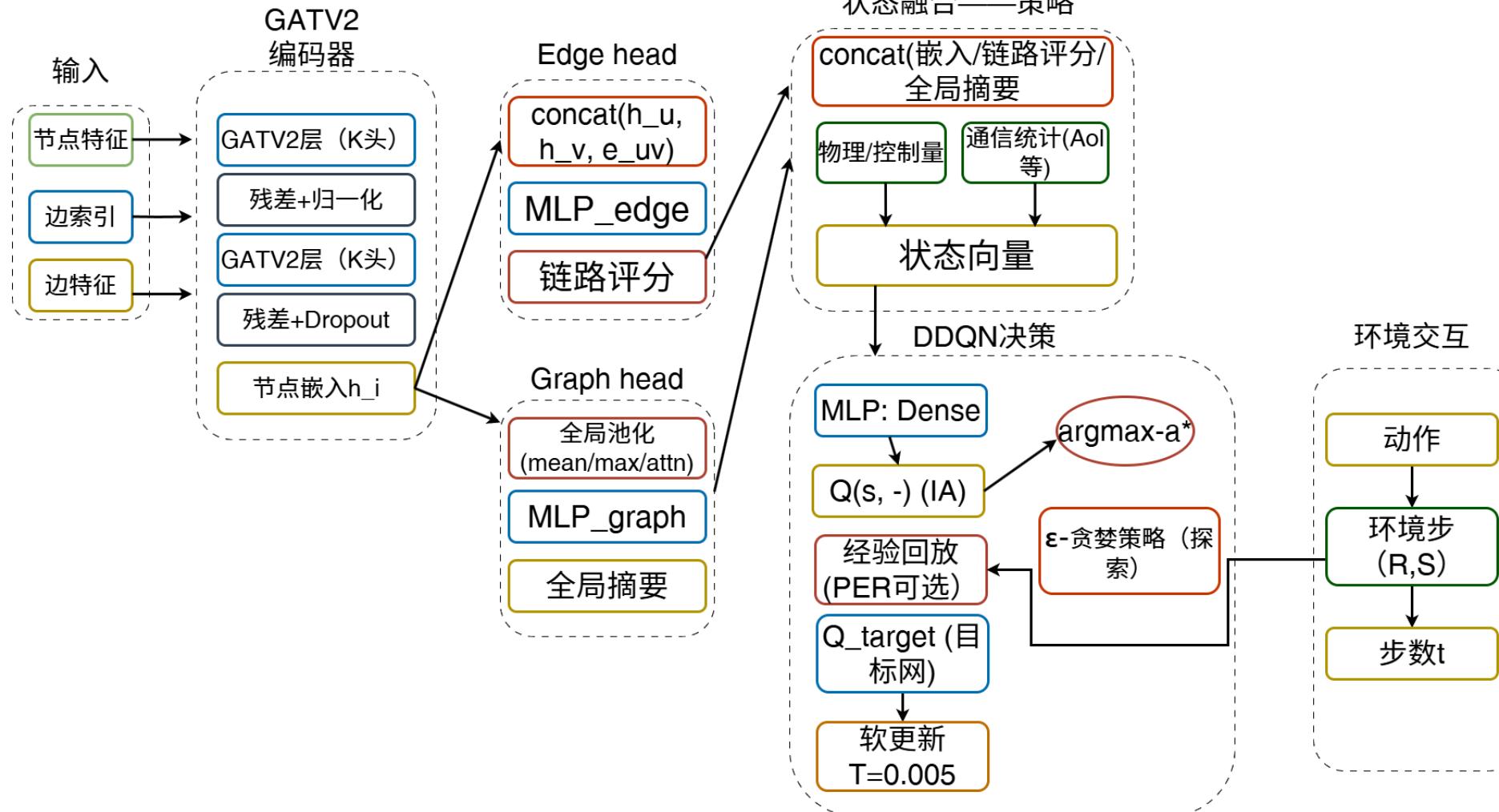
总体流程简要图





3.1 研究依据

GATV2图注意机制应用结构框图



GATv2+DDQN 算法流程

- 输入：节点与边特征
- GATv2编码，提取节点嵌入
- 边/图头处理特征，汇总为状态
- DDQN决策，选最优动作
- 环境交互与回放，持续更新策略



3.1 研究依据

第一步特征嵌入变换

$$(\tilde{h}_i = Wx_i)$$

将原始节点信息转换为高维特征空间，便于后续处理和融合。

第二步邻居-目标融合打分

$$(e_{ij} = \frac{\mathbf{a}^\top \sigma(W_a[\tilde{h}_i \| \tilde{h}_j])}{\tau})$$

计算任意两节点之间的信息重要性（注意力分数），识别关键邻居，帮助我们自动甄别哪些通信链路或协同关系需要重点关注，比如谁是队列的关键节点，谁的信号最重要。

第三步注意力权重归一化分配

$$(\alpha_{ij} = \frac{e^{e_{ij}}}{\sum_{k \in \mathcal{N}(j)} e^{e_{kj}}})$$

通过权重归一化，将关注度分配给各邻居节点，突出重点信息让每个车在通信或队列协同时，可以“自适应”从关键邻居那里优先接收信息，实现智能筛选。

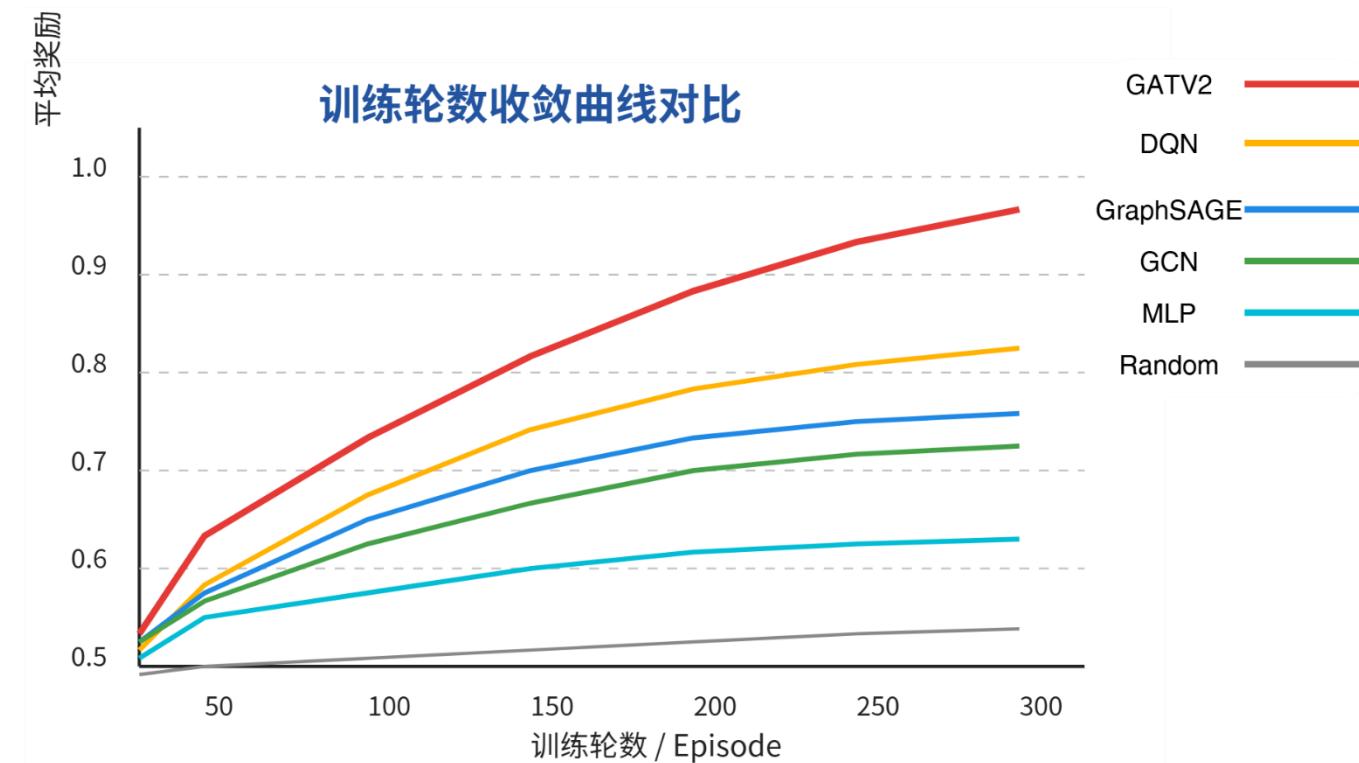
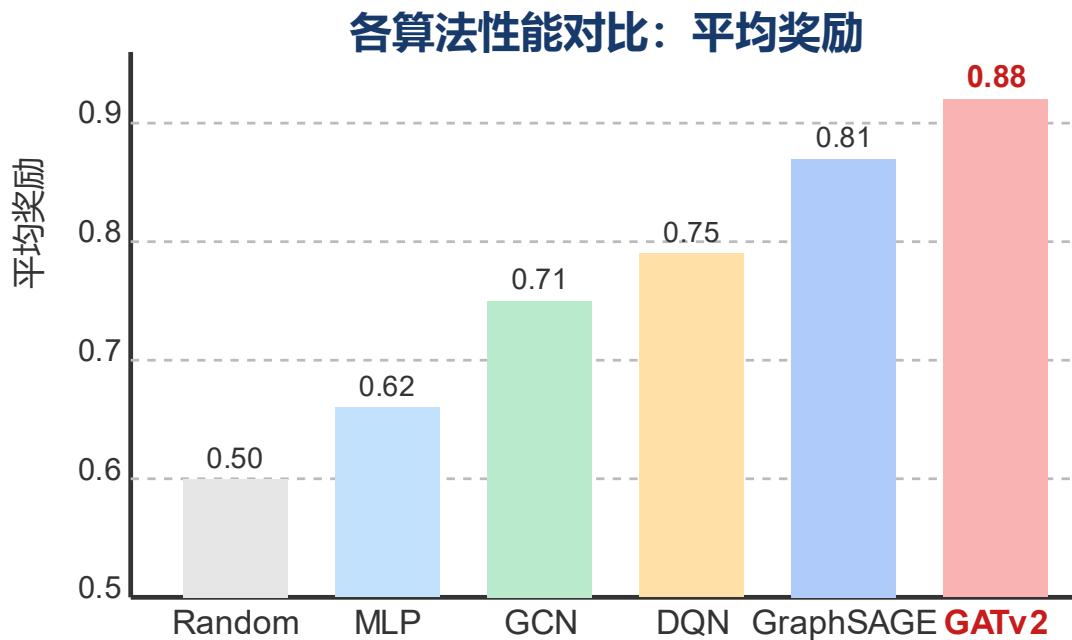
第四步加权信息聚合融合

$$(h_j = \sum_{i \in \mathcal{N}(j)} \alpha_{ij} \tilde{h}_i)$$

根据分配的权重，聚合邻居节点信息，实现智能体的信息融合与协同感知，让编队每辆车都能自主融合来自周边的最有效信息，用以后续通信调度和队列协同优化。



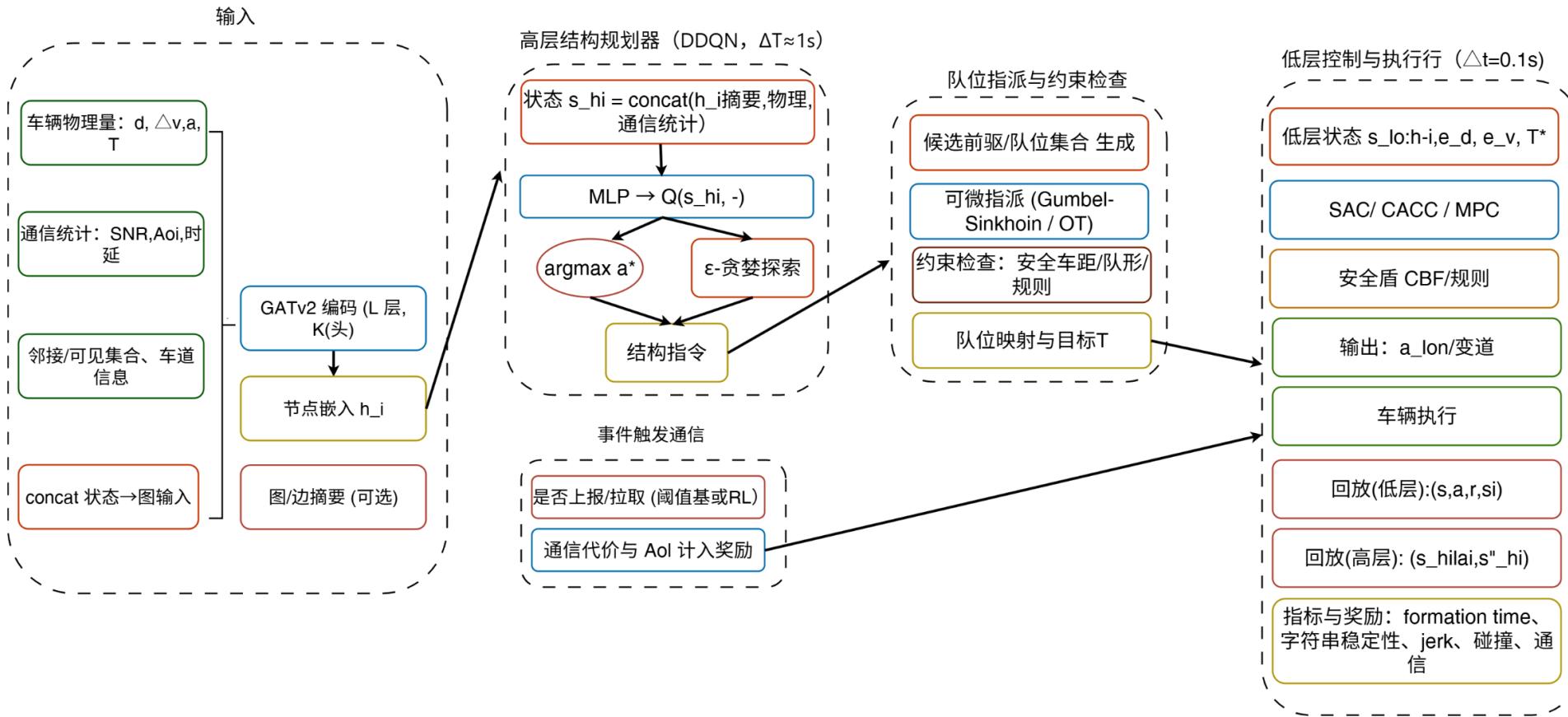
3.1 研究依据



- GATv2 能动态分配“注意力权重”给每个邻居节点，精细刻画每个车辆之间的关系
- GATv2在多节点、多边复杂图结构下，能更快收敛到优解，具有更高的泛化能力和鲁棒性。
- GATv2 的注意力机制是全参数化和非线性的，能学习复杂的关系和特征交互，有利于处理交通网络这种高度动态、异质的环境。

3.1 研究依据

决策 - 控制机制框图



分层决策控制流程

- 1 车辆物理和通信信息 → 图神经网络编码
- 2 高层 (DDQN) : 队形 / 分组方案决策
- 3 约束检查和队伍分派
- 4 低层 (SAC等) : 具体动作控制和执行



$$\mathbf{1}_{LC} = \underbrace{\mathbf{1}\{a_f^{\text{new}} \geq -b_{\text{safe}} \wedge a_b^{\text{new}} \geq -b_{\text{safe}}\}}_{\text{安全门控}} \cdot \underbrace{\mathbf{1}\{\Delta a_i + p(\Delta a_f + \Delta a_b) - c_{lc} - a_{thr} > 0\}}_{\text{激励/收益判据}}$$

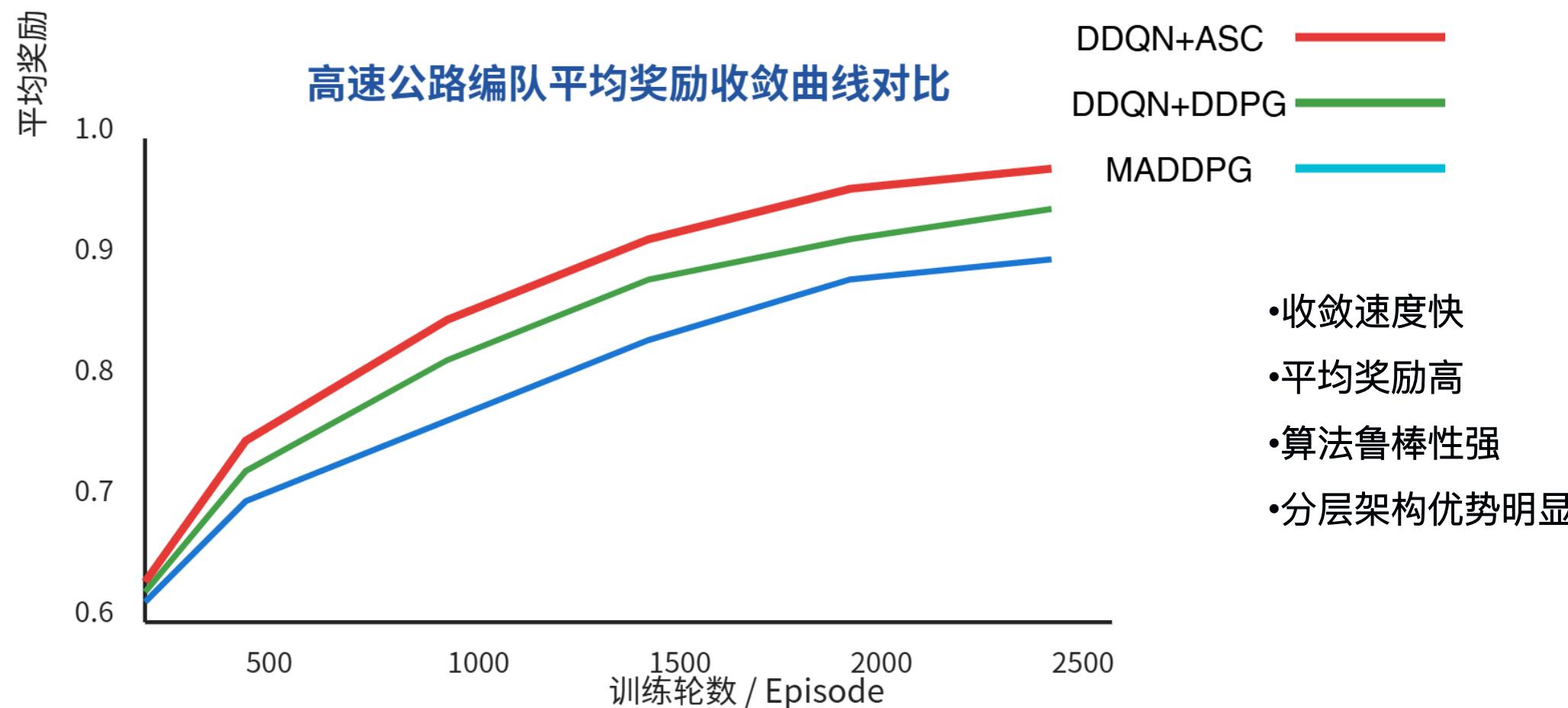
只有安全门控和激励判据都为1时（即都满足条件）， $(\mathbf{1}_{LC})$ 才为1，车队/编队中的目标车辆才执行变道。

$\mathbf{1}\{a_f^{\text{new}} \geq -b_{\text{safe}} \wedge a_b^{\text{new}} \geq -b_{\text{safe}}\}$ 这是编队变道的**安全门控**，只有满足安全约束车辆才允许考虑变道，防止事故风险。

$\mathbf{1}\{\Delta a_i + p(\Delta a_f + \Delta a_b) - c_{lc} - a_{thr} > 0\}$ 这是编队变道的**激励和收益判据**，只有当整体收益大于成本和阈值时，变道才被认为是有益的决策。



3.1 研究依据



Part.04

创新点



创新点

(1)

针对受限带宽与强干扰环境下车辆编队通信质量下降与高延迟问题，提出融合图注意力网络（GATv2）与深度强化学习（DDQN+Top-K）的新范式，构建车联网图表示，实现关键车辆与链路的差异化资源调度，保障数据新鲜度与时延性能。

(2)

针对车辆编队动态重构中多层级决策复杂性问题，设计分层强化学习（HRL）控制框架，高层DDQN负责离散结构决策，低层SAC输出连续转向控制，结合共享图嵌入与时间尺度解耦机制，提升训练稳定性与实时响应能力。

(3)

针对队位分配与插队决策中局部贪心导致的次优解与序列抖动问题，引入可微分指派（*Gumbel–Sinkhorn / OT*）优化方法，结合候选前驱约束，实现全局一致性位置分配与低代价平滑调整，有效抑制编队抖动现象。

Part.05

论文工作计划

第①阶段：阅读自动驾驶编队、V2X通信、图神经网络、多智能体强化学习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tensorflow框架，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。（2025.3——2025.5）

第②阶段：深入学习图注意力网络、Double DQN、SAC等算法，完成通信资源调度模块与分层强化学习控制模块的设计与实现。实现CTDE框架的集成，完成动态队建模、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。（2025.5—2025.12）

第③阶段：在Tensorflow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。对比主流方法（优化模型超参数与网络结构），完成消融实验与分析各模块贡献。撰写小论文。预计6个月完成。（2025.12—2026.6）

第④阶段：对论文进行撰写，并进行整体文稿修改，查阅资料、补充细节，最终形成书面材料，装订成册，准备参加毕业答辩。预计6个月完成。（2026.6—2027.1）

Part.06

参考文献

- [1] Hu W, Deng Z, Cao D, et al. Probabilistic Lane-Change Decision-Making and Planning for Autonomous Heavy Vehicles[J]. IEEE/CAA Journal of Automatica Sinica, 2022, 9(12): 2161-2173.
- [2] Bi Z, Guo X, Wang J, 等. Truck-Drone Delivery Optimization Based on Multi-Agent Reinforcement Learning[J]. Drones, 2024, 8(1): 27.
- [3] 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.
- [4] 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.
- [5] 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.
- [6] 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.
- [7] 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.
- [8] 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.
- [9] 邓辉, 张学艳, 胡金玲, 等. 基于车联网的车辆编队标准现状及展望[J]. 移动通信, 2022, 46(8): 86-91.
- [10] Huang Y. Multimedia Tasks-Oriented Edge Computing Offloading Scheme Based on Graph Neural Network in Vehicular Networks[J]. IEEE Access, 2025, 13: 9780-9791.

- [11] 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.
- [12] Lv K, Pei X, Chen C, 等. A Safe and Efficient Lane Change Decision-Making Strategy of Autonomous Driving Based on Deep Reinforcement Learning[J]. Mathematics, 2022, 10(9): 1551.
- [13] 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.
- [14] Learning Automata-Based AODV to Improve V2V Communication in A Dynamic Traffic Simulation[J]. International Journal of Intelligent Engineering and Systems, 2024, 17(1): 666-678.
- [15] 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.
- [16] 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.
- [17] 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.
- [18] 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.
- [19] Ni H, Yu G, Chen P, et al. An Integrated Framework of Lateral and Longitudinal Behavior Decision-Making for Autonomous Driving Using Reinforcement Learning[J]. IEEE Transactions on Vehicular Technology, 2024, 73(7): 9706-9720.
- [20] 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.



- [21] Zhang S, Zhuang W, Li B, et al. Integration of Planning and Deep Reinforcement Learning in Speed and Lane Change Decision-Making for Highway Autonomous Driving[J]. IEEE Transactions on Transportation Electrification, 2025, 11(1): 521-535.
- [22] Yildirim M, Dagda B, Fallah S. HighwayLLM: Decision-Making and Navigation in Highway Driving with RL-Informed Language Model[A]. arXiv, 2024.
- [23] 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.
- [24] Khanh Quy V, Chehri A, Hoai Nam V, et al. Strategic Data Offloading for 5G and Beyond for Internet of Vehicles Networks: Current Trends and Future Directions[J]. IEEE Open Journal of the Communications Society, 2025, 6: 8606-8624.
- [25] Ji M, Wu Q, Fan P, et al. Graph Neural Networks and Deep Reinforcement Learning Based Resource Allocation for V2X Communications[A]. arXiv, 2025.
- [26] Ren Y, Wu C, So D K C, 等. 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.
- [27] Qu N, Wang C, Li Z, 等. A transmission design in dynamic heterogeneous V2V networks through multi-agent deep reinforcement learning[J]. China Communications, 2023, 20(7): 273-289.
- [28] 彭江文. 多智能体协同编队的自适应学习控制方法研究[D]. 天津: 天津大学, 2021.
- [29] Bian F. Lightweight convolutional neural network for multiscale temporal modeling in ultra-reliable low-latency communication[C]/2025 4th International Symposium on Computer Applications and Information Technology (ISCAIT). Xi'an, China: IEEE, 2025: 1364-1368.
- [30] Li Z N, Huang X H, Mu T, et al. Attention-Based Lane Change and Crash Risk Prediction Model in Highways[J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(12): 22909-22922.

请各位老师指正



汇报人：高艺泷



指导老师：郭彤颖