**复杂环境下智能网联汽车协同智能增强**

作者：马若轩1，秦成2，刘禹菲1，李亚擘1，李漫琳1

指导教师：杭鹏1

（同济大学 1交通学院，2 计算机科学与技术学院 上海 上海 201804）

摘 要：面对复杂的城市交通场景，现有的自动驾驶技术在应对长尾场景时仍存在挑战，导致接管率较高，安全性有待提升。目前的智能网联技术已实现车辆间的信息共享，但主要应用于辅助驾驶和安全预警，尚未充分发挥协同决策的优势。基于此，本文提出了一种针对复杂环境的智能网联汽车协同智能增强框架，以提升车辆在复杂交通环境中的自主决策能力。本文选取交叉口无信号灯保护环境作为研究场景，将智能网联汽车协同增强问题建模为多智能体协同决策模型。首先，基于“同理心”原则和支持向量机(SVM)，利用路测数据建立交互机理分析模型对车辆交互意图进行分析，实现对交互状态的准确判断；然后构建PPO-OK(Proximal Policy Optimization-Output K)决策模型，采用强化学习与博弈结合的方式寻找交互车辆最佳K值（车辆层级）组合，提高智能网联汽车协同决策效率；最后，基于Level-K博弈思想构建决策执行模型，将PPO-OK决策模型得出的最优K值组合转换为协同车辆最优动作组合，对智能网联汽车进行控制。通过智能网联汽车协同智能增强框架，高效、安全实现智能网联汽车间的协同决策，有效解决自动驾驶面对复杂交通环境单车智能不足的问题，为高等级自动驾驶汽车商业化应用提供技术支持。

关键字：智能网联汽车；复杂环境；交互机理分析；Level-K博弈；强化学习；协同智能增强；

1. 研究背景

面对日益复杂的城市交通环境，目前的自动驾驶车辆难以实现全场景无人化运行。武汉萝卜快跑自动驾驶汽车发生死锁导致交通瘫痪事件，且自动驾驶汽车事故频发、接管率高等问题频现，均反映出单车智能技术难以处理长尾场景，严重制约自动驾驶技术的规模化商用。

智能网联技术迅速发展，让智能网联汽车共享状态和环境信息成为现实。但传统的智能网联技术应用场景大多局限在驾驶辅助和安全预警。基于智能网联技术的协同智能增强框架可以实现智能网联汽车间的协同决策，解决单车智能不足的瓶颈问题，提升交通系统的运行效率，为高等级自动驾驶汽车的落地提供技术支持。

1. 研究内容

2.1 概述

本文针对自动驾驶技术在复杂交通环境中的决策瓶颈，构建了一套“感知-交互判断-博弈决策-执行控制”的智能网联汽车协同增强方法，实现智能网联汽车信息互通及高效合作，优化车辆间的协同决策机制，提升车辆在动态复杂交通环境中的应对能力，确保高效、稳定的协同驾驶。

2.2 技术路线

单车智能失效多发生在交叉口场景，故主要对交叉口场景的协同增强进行研究。本文提出的智能网联汽车协同增强模型主要由三部分组成：

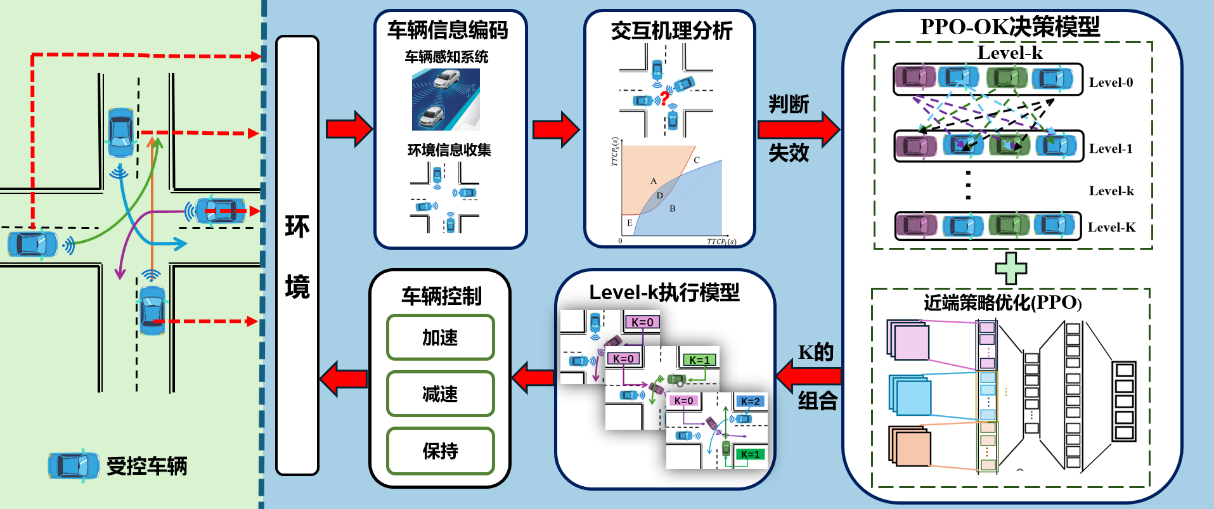


图1 智能网联汽车协同智能增强技术路线图

智能网联汽车周围环境与状态信息通过传感器等方式输入到模型，并进行编码构成环境信息。环境信息输入到交互机理分析模型，对车车交互状态进行判断，若处于交互失效状态，则智能网联汽车进入协同增强。

将协同增强问题建模为可观测的马尔科夫决策过程，使用基于Actor-Critic架构的强化学习算法训练得到PPO-OK模型，输入环境中各智能网联汽车的特征信息，得到各车辆最优K值（车辆层级）的组合。

通过基于Level-K博弈思想的决策执行模型，根据最优K值组合得到各车的最优动作并进行执行，实现智能网联汽车协同决策。

1. 交互机理动态分析模型

在自动驾驶路测数据中，安全员接管往往意味着单车智能已经无法应对当前场景。为更好地识别车辆间交互失效，本文基于支持向量机（SVM）提出一套交互机理动态分析模型，并引入“同理心”原则作为建模基础，使车辆在决策时充分考虑社会性交互。

以车辆视角为例，协作加速度为：

其中，是车辆到达冲突点的距离；是车辆的速度；是车辆到达冲突点的时间。

基于协同加速度模型，车辆能够当前运动状态和对方运动状态做出决策。引入SVM处理非线性分类问题，以区分不同的决策类型。通过对数据集的学习，SVM找到最优超平面将不同的决策类别分开。

SVM的输入特征向量为：

X = [, , ]

其中，是车辆自身到达冲突点的时间；是对方车辆到达冲突点的时间；是对方车辆的协作加速度。

输出标签Y定义为：

Y =

在交互过程中，车辆根据当前的运动状态评估对方的优先权。将分界面在TTCP时间域中进行映射，变形得到公式：

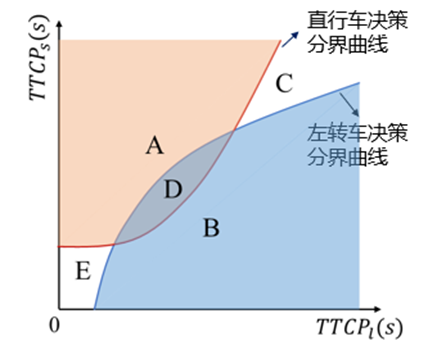
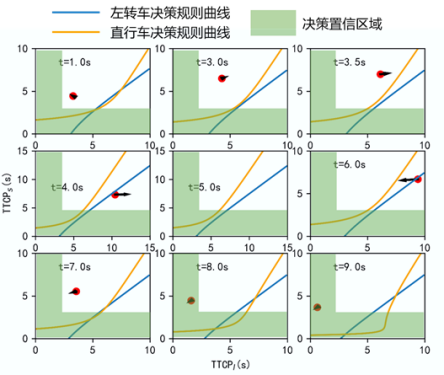
 

图2 决策时域图 图3 交互机理动态分析

引入 “决策时域图”以进一步描述车辆在不同时间点的决策过程，将交互过程分为不同区域。基于车辆运动状态与周围环境信息，利用失效场景分析模型进行分析，识别交互失效状态，及时开启协同增强以保证智能网联汽车交互过程的高效、安全。

1. PPO-OK决策模型

将协同增强问题建模为可观测的马尔科夫决策过程，使用路测数据集与仿真环境训练PPO-OK决策模型。该模型将Level-K与强化学习PPO算法结合，优化协同决策过程，采用选择Level-K模型中的k值。

|  |  |
| --- | --- |
| Algorithm 1: PPO-OK决策模型 | |
| 参数: 全局状态 s  输出: 各自动驾驶车辆的 k 值 | |
| 训练PPO-OK决策模型 | |
| 1 | 初始化：以为初始参数的策略网络 ;  以为初始参数的价值网络;  学习率，裁剪参数. |
| 2 | **for** t = 0, 1, 2, … **do** |
| 3 | 收集轨迹集合，即在环境中运行策略 获得数据. |
| 4 | 计算奖励回报 (折扣累计奖励) . |
| 5 | 计算优势函数 **.** |
| 6 | 最大化目标函数更新策略网络参数 : |
| 7 | 均方误差回归更新价值网络参数 : |
| 8 | **end** |
| 调用PPO-OK决策模型 | |
| 9 | 将参数中的全局状态 s 输入策略网络 |
| 10 | 得到各自动驾驶车辆的 k 值 |

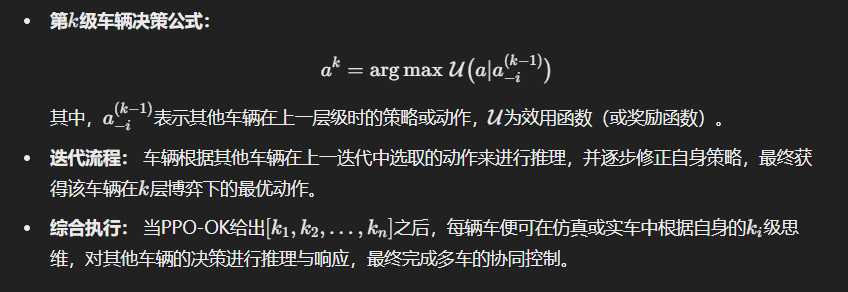
在PPO算法中，当前策略与旧策略之间的比率通过重要性采样进行计算，以衡量当前策略在特定状态下选择某一动作的相对重要性。基于此比率，PPO算法评估当前策略更新与旧策略的差异，并据此调整策略。

使用强化学习训练得到的PPO-OK模型，将车辆状态信息作为输入，输入为环境中各车的特征向量，特征向量包含；输出为各智能网联汽车k值的排列。

1. Level-K执行模型

在获得PPO-OK模型输出的最优K值后，需要进一步将其映射为各车辆的具体驾驶动作。为此，本文基于Level-K博弈思想构建了执行模型，使不同层级的车辆可逐层迭代推理，进而逼近纳什均衡或博弈最优解。

|  |  |
| --- | --- |
| Algorithm 2: Level-K执行模型 | |
| 参数: 各自动驾驶车辆的 k 值、规划步数 frames, 环境 env | |
| 输出: 各自动驾驶车辆的动作 | |
| 1 | 初始化：设定仿真步数 frames = 30;  初始化动作表 ActionTable,大小为 count\_vehicle × count\_k. |
| 2 | **for** k = 0, 1, 2, …, count\_k **do** |
| 3 | **for** vehicle\_id = 0, 1,…, count\_vehicle **do** |
| 4 | **for** action **in** action\_type **do** |
| 5 | **if** k = 0 **then** |
| 6 | 假设其他自动驾驶车辆静止，当前车辆执行action. |
| 7 | **end** |
| 8 | **if** k > 0 **then** |
| 9 | 假设其他自动驾驶车辆执行 k – 1 级别时的最优动作，当前车辆执行action. |
| 10 | **end** |
| 11 | **for** i = 0, 1,…, frames − 1 **do** |
| 12 | 在上述执行动作的基础上使用PID控制车辆更新位置，累计奖励. |
| 13 | **end** |
| 14 | **if** 累计奖励大于当前累计奖励 |
| 15 | 更新 ActionTable 中的数据为 action |
| 16 | **end** |
| 17 | **end** |
| 18 | **end** |
| 19 | **end** |

对于任何车辆，通过顺序迭代的计算过程，在每一轮迭代中，车辆会根据其他车辆在上一轮迭代中的决策，即。通过这种逐步迭代，车辆可以逐渐逼近更高层次的博弈均衡。对于每个，通过这种顺序迭代的计算，最终获得 级推理下的最优决策，其中是预定的最大推理层级。

通过level-k的迭代过程最终将不同车辆最优K值排列转化为不同车辆的最优决策动作，在复杂的交通环境中实现高效、安全决策。

1. 结果分析

本章节从训练效果、有无协同对比、模型效果对比多个角度对模型表现进行评估，证明协同智能增强的可行性。

6.1训练效果

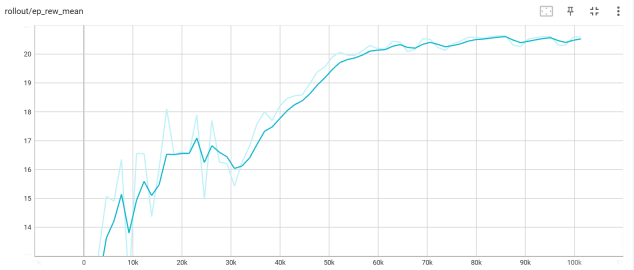
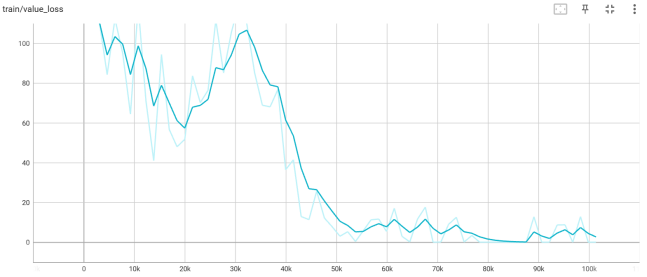
 

图4 训练奖励变化趋势图 图5 Loss曲线图

上图展示了PPO-OK在训练过程中的奖励值变化趋势（左）与Loss曲线（右）。在初始阶段，模型对环境的认知不足，奖励波动大，Loss也大且波动频繁；随着训练的深入，策略逐渐稳定并收敛；当训练步数达到40k步后，奖励值趋于稳定高值，Loss显著降低并波动减小，表明模型能够在复杂交通环境中实现稳定的协同决策，并有效提升协同效率。

6.2有无协同对比

为验证协同智能的提升效果，本文选取三类典型的无信号灯交叉口场景：四车同时左转；四车同时直行；四车直行与左转混合交叉。无协同车辆采用PPO训练得出的单车智能算法，协同车辆采用智能网联协同智能增强框架算法，得到协同过程如下：

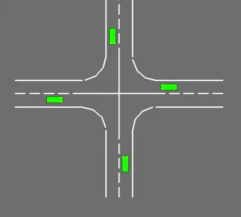
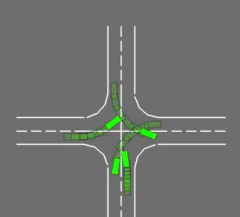
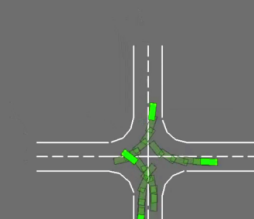
   

图6 交叉口四车左转场景协同示意图

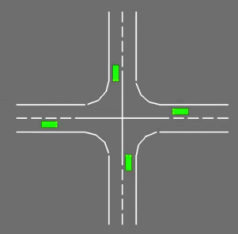
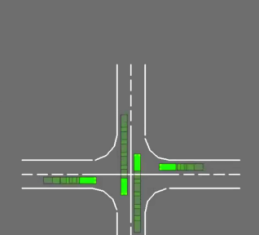
   

图7 交叉口四车直行场景协同示意图

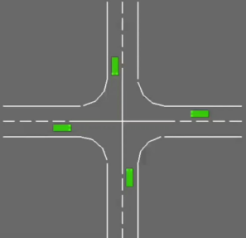
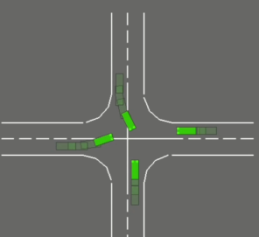
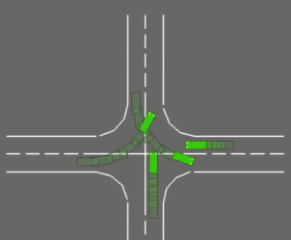
   

图8 交叉口左转直行混合场景协同示意图

由三种场景可以看出，智能网联协同增强框架能够在交叉口典型冲突场景中，合理分配K值，输出安全、高效的决策，受控车辆得以有序通过。

6.3模型效果对比

选取四车左转、四车直行和左转直行混合三种典型交叉口复杂场景，从平均车速、完全通过时间和通过率三个维度，对单车智能、Level-K（基线模型）和PPO-OK三个模型进行比较。

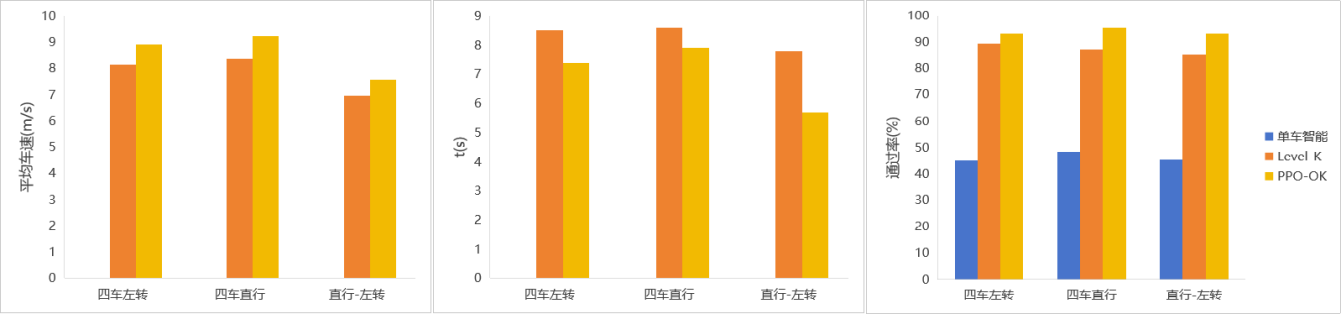


图9 不同场景平均速度、完全通过时间、通过率对比

可以看出，PPO-OK明显提高了车辆的平均通行速度、降低了整体完成通行的时长，并在高复杂度场景中拥有更高的安全通过率，充分证明了协同增强框架在复杂场景中的有效性与优越性。

1. 创新特色

7.1交互机理动态分析模型

本文提出了基于“同理心”原则和SVM的交互机理动态分析模型。相比于传统的概率模型判断，该模型基于交互车辆的状态信息，能够更加准确且实时判断车辆交互状态，识别交互失效状态，为车辆协同增强提供依据。

7.2 Level-K博弈与PPO算法结合

传统的Level-K博弈算法虽然能够实现车辆间的协同，但算法较为复杂，计算效率较慢，应用于实车的难度较大。通过Level-K博弈与PPO算法的结合，由PPO算法给出最优K值组合，将强化学习训练过程前置，在模型应用过程中能够显著提高响应速度和计算效率。

1. 应用前景

在现有自动驾驶尚未完全普及的背景下，智能网联汽车协同智能增强算法可以部署于车载协同控制器，应用于辅助驾驶场景或L3级自动驾驶车辆的多车交互场景，特别是在城市复杂交通环境中，为车辆交互冲突提供更安全和高效的决策支持，减少驾驶员接管频率，提高智能网联汽车的整体运行效率。

在自动驾驶汽车高渗透率的未来，该算法可作为核心协同模块，与端侧智能驾驶决策系统深度融合，推动L4及以上级别自动驾驶的商业化落地。特别是在车联网高度发展的背景下，该算法可广泛应用于共享自动驾驶车队调度，通过多车协同增强整体系统效率，优化道路资源利用。

参考文献

1. Fang, S., Hang, P., Wei, C., Xing, Y., & Sun, J. (2024). Cooperative driving of connected autonomous vehicles in heterogeneous mixed traffic: A game theoretic approach. *IEEE Transactions on Intelligent Vehicles*.
2. Xu, C., Deng, Z., Liu, J., Kong, A., Huang, C., & Hang, P. (2024). Towards Safe and Robust Autonomous Vehicle Platooning: A Self-Organizing Cooperative Control Framework. *arXiv preprint arXiv:2408.09468*.
3. Karimi, S., & Vahidi, A. (2020, July). Receding horizon motion planning for automated lane change and merge using monte carlo tree search and level-k game theory. In *2020 American Control Conference (ACC)* (pp. 1223-1228). IEEE.
4. Fang, S., Zhou, D., Cui, Y., Xu, C., Hang, P., & Sun, J. (2025). Recognize then Resolve: A Hybrid Framework for Understanding Interaction and Cooperative Conflict Resolution in Mixed Traffic. *arXiv preprint arXiv:2501.19025*.
5. Xu, C., Liu, J., Hang, P., & Sun, J. (2025). TeLL-Drive: Enhancing Autonomous Driving with Teacher LLM-Guided Deep Reinforcement Learning. *arXiv preprint arXiv:2502.01387*.