

基于智能交通系统的混合动力车队分层控制策略

王浚宇¹, 詹森¹, 肖勇¹, 李宗华², 刘聪¹

(1. 重庆交通大学 机电与车辆工程学院, 重庆 400074; 2. 深蓝汽车科技有限公司, 重庆 401147)

摘要: 为了保证混合动力车队行驶过程的安全性、舒适性并提升燃油经济性, 提出了一种基于智能交通系统的混合动力车队分层控制策略。上层控制器利用车-车 (V2V) 通信技术, 采用非线性分布式模型预测控制 (NDMPC) 进行速度优化控制, 计算最优车速; 下层控制器获取上层车速, 采用深度确定性策略梯度 (DDPG) 强化学习算法进行能量管理, 并嵌入了发动机最佳经济曲线与电池特性曲线作为专家经验提升算法收敛性与稳定性。结果表明, 该策略下车距的最大相对误差为 4.83%, 车队平均加速度为 0.331 m/s^2 , 与下层采用深度 Q 学习 (DQN) 算法对比, 燃油消耗量平均减少了 14.25%, 最大减少 15.30%, 相比于 DP 算法平均只增加了 8.31%, 在保证安全性和舒适性的同时, 有效提升了车队的经济性。

关键词: 混合动力车队; 分层控制; 非线性分布式模型预测控制; 深度确定性策略梯度

中图分类号: U461.8 文献标志码: A DOI: 10.3969/j.issn.2095-1469.2025.04.12

Hierarchical Control Strategy of Hybrid Electric Vehicle Platoon Based on Intelligent Transportation Systems

WANG Junyu¹, ZHAN Sen¹, XIAO Yong¹, LI Zonghua², LIU Cong¹

(1. School of Mechatronics and Vehicle Engineering, Chongqing Jiaotong University, Chongqing 400074, China;
2. Deepal Automobile Technology Co., Ltd., Chongqing 401147, China)

Abstract: In order to ensure the safety, comfort and fuel economy of hybrid electric vehicle platoon, a hierarchical control strategy based on intelligent transportation systems is proposed. The upper controller used vehicle-vehicle(V2V) communication technology and used Nonlinear Distributed Model Predictive Control (NDMPC) to optimize the speed control and calculate the optimal speed. The lower controller obtained the upper vehicle speed, used the Deep Deterministic Policy Gradient (DDPG) reinforcement learning algorithm for energy management, and embedded the engine best economic curve and battery characteristic curve as expert experience to improve the convergence and stability of the algorithm. The results show that under this strategy, the maximum relative error of vehicle distance is 4.83%, and the average acceleration of the platoon is 0.331 m/s^2 . Compared with the Deep Q-Network (DQN) algorithm used in the lower layer, the fuel consumption is reduced by 14.25% on average, and the maximum is reduced

收稿日期: 2024-07-02 改稿日期: 2024-07-17

基金项目: 重庆市教委青年项目 (KJQN202000734); 机械传动国家重点实验室开发基金项目 (SKLMT-MSKFKT-202112); 重庆市技术创新与应用发展重点项目 (cstc2020jscx-dxwtBX0025)

参考文献引用格式:

王浚宇, 詹森, 肖勇, 等. 基于智能交通系统的混合动力车队分层控制策略[J]. 汽车工程学报, 2025, 15(4): 554-566.

WANG Junyu, ZHAN Sen, XIAO Yong, et al. Hierarchical Control Strategy of Hybrid Electric Vehicle Platoon Based on Intelligent Transportation Systems[J]. Chinese Journal of Automotive Engineering, 2025, 15(4): 554-566. (in Chinese)



by 15.30%. Compared with the DP algorithm, the average increase is 8.31%. It not only ensures the safety and comfort, but also effectively improves the economy of the platoon.

Keywords: hybrid electric vehicle platoon; hierarchical control; Nonlinear Distributed Model Predictive Control; Deep Deterministic Policy Gradient

混合动力汽车 (Hybrid Electric Vehicle, HEV) 的能量管理策略能有效提升整车的经济性^[1], 但传统能量管理策略主要针对单车进行研究。随着智能交通系统的发展, 通过车-车通信 (Vehicle-to-Vehicle, V2V) 技术能让车与车之间建立联系形成相互关联的车辆队列^[2], 因此, 目前开展对于 HEV 队列的能量管理策略研究成为热点。HEV 队列的能量管理策略常采用分层控制^[3], 上层控制策略主要是对整个车队进行速度控制, 下层控制根据上层速度来决定动力部件的能量优化分配。

针对上层控制策略, 通常采用分布式模型预测控制 (Distributed Model Predictive Control, DMPC) 进行求解^[4]。DMPC 算法可分为线性分布式模型预测控制 (Linear Distributed Model Predictive Control, LDMPC) 与非线性分布式模型预测控制 (Nonlinear Distributed Model Predictive Control, NDMPC)^[5]。目前主要采用 LDMPC 完成车辆队列速度控制, GRAFFIONE 等^[6] 利用 LDMPC 控制车队间车辆的距离与速度, 以最小化位置及速度偏差等为目标。YUAN Junkai 等^[7] 在无闯红灯与碰撞事故的前提下, 采用 LDMPC 对参考速度进行优化, 得到目标速度与加速序列。WANG Pangwei 等^[8] 针对异常通信的下切换通信拓扑结构的情况, 提出了考虑时延、丢包和中断等的 LDMPC 算法, 该算法能在不同的通信环境下实现通信拓扑的实时切换。LDMPC 通过公式将速度控制问题转化为线性约束的二次规划问题进行求解, 简化了模型结构, 优化了求解速度。但车辆队列系统具有明显的非线性特征, 因此, LDMPC 通常只能获取次优解而非最优解, 无法实现速度最优控制^[9-10]。相比之下, NDMPC 将非线性的问题转化为二阶锥规划问题, 通过内点法能实现毫秒内的求解^[11], 因此在求解速度上与 LDMPC 相差无几, 但

是求解精确性提升明显, 且其稳定性早已得到论证^[12], 然而该算法在车队分层控制中研究相对较少。

在下层车队的能量管理可分为基于规则、基于优化和基于学习的控制策略。针对基于规则的策略, LUO Yugong 等^[13] 设计安全距离控制和紧急停车算法, 在实现安全和乘坐舒适性的基础上, 开发了基于规则的能量管理算法以提高燃油经济性。此方法简单可靠, 但需要根据专家经验制定规则, 算法的适应性与优化效果不足。针对基于优化的策略, QIU Lihong 等^[14] 利用 MPC 框架在上层通过信号灯信息得到目标速度, 下层采用的动态规划算法对能量进行优化分配, 保证了车队的经济性及安全性。虽然基于优化的算法可获取全局最优解, 但必须提前得知全局工况, 难以在线实时控制。针对基于学习的控制策略, YIN Yanli 等^[15] 提出了基于前车跟随式通信拓扑结构的 DMPC 与 DQN 相结合队列分层控制策略。基于 DQN 算法的方法拥有较好的优化效果及实时性, 但其采用的是离散动作空间, 优化效果有限。DDPG 算法采用连续的动作空间, 可以搜索到更合理的动作, 相比于 DQN 算法在保证实时性的基础上, 能获取更优解。

本文提出一种基于智能交通系统的混合动力车队分层控制策略。针对车辆队列协同控制问题, 利用 V2V 通信设计了车队上层 NDMPC 控制算法, 对车队中的各车辆进行速度控制, 提升车辆行驶时的稳定性、安全性以及舒适性; 同时, 为了提升车队燃油经济性, 下层以电池特性与发动机最优工作曲线作为专家经验设计了基于 DDPG 的下层控制算法, 接受上层传递的最优速度来进行能量管理。

1 上层基于 NDMPC 的速度控制

1.1 单车模型

功率分流式 HEV 可以实现无级变速, 拥有较

好的节能能力^[16]。因此，本文选用功率分流式 HEV 均质队列为研究对象，其中，包含 1 辆领航车与 4 辆跟随车，其动力传动系统结构如图 1 所示，包括电池、电机、发动机及行星齿轮组等部件。整车主要参数见表 1，其中特征参数 α 为齿圈齿数与太阳轮齿数之比。

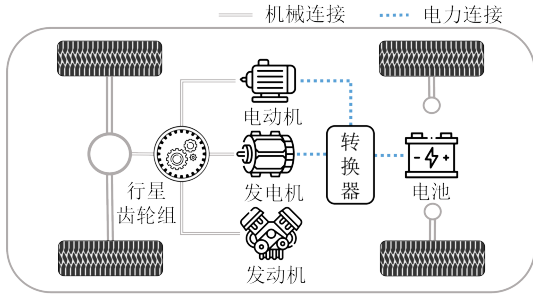


图 1 功率分流式 HEV 动力传动系统结构

表 1 整车主要参数

	参数	值
车身	质量/kg	1 449
	滚动阻力系数	0.013
	空气阻力系数	0.26
	迎风面积/m ²	2.23
电动机	最大功率/kW	50
	最大转矩/Nm	400
	最大转速/(r/min)	6 000
发电机	最大功率/kW	37.8
	最大转矩/Nm	75
	最大转速/(r/min)	10 000
发动机	最大功率/kW	56
	最大转矩/Nm	120
	最大转速/(r/min)	4 000
电池	容量/kWh	1.54
	电压/V	237
变速器	主减速器传动比	3.93
	特征参数 α	2.6

1.1.1 行星齿轮组建模

当行星齿轮机构稳定运行时，各部件之间转速与转矩有着固定的比例关系，关系式为：

$$\begin{cases} \omega_{MG1} + \alpha\omega_{MG2} - (1 + \alpha)\omega_{eng} = 0, \\ T_{eng}: T_{MG1}: T_{MG2} = (\alpha + 1):(-1):(-\alpha). \end{cases} \quad (1)$$

式中： ω_{MG1} 为发电机的转速； ω_{MG2} 为电动机的转速； ω_{eng} 为发动机转速； T_{eng} 为发动机的转矩； T_{MG1} 为发电机的转矩； T_{MG2} 为电动机的转矩。

1.1.2 发动机建模

建立发动机模型时，忽略发动机内部热交换过程，因此瞬时燃油消耗率可以近似为与发动机转速与转矩相关的函数，所建立的发动机模型为：

$$\begin{cases} b_c = f(T_{eng}, \omega_{eng}), \\ P_c = \frac{T_{eng}\omega_{eng}}{1\,000}, \\ Q = \frac{P_c b_c}{367.1\rho g}. \end{cases} \quad (2)$$

式中： b_c 为瞬时燃油消耗率； P_c 为发动机功率； ρ 为燃油密度； Q 为单位时间的燃油消耗量。

1.1.3 电机建模

本文不考虑温度、热传递等对电机效率的影响，构建电动机与发电机的效率模型，如图 2~3 所示。电机的瞬时效率与转矩转速的函数为：

$$\eta_m = g_m(T_m, \omega_m). \quad (3)$$

式中： η_m 为电机的瞬时效率； T_m 为电机的转矩； ω_m 为电机的转速。

电机功率与转矩转速的函数为：

$$P_m = \begin{cases} \frac{T_m\omega_m}{1\,000}\eta_m & T_m > 0, \\ \frac{T_m\omega_m}{1\,000\eta_m} & T_m < 0. \end{cases} \quad (4)$$

式中： P_m 为电机的输出功率。

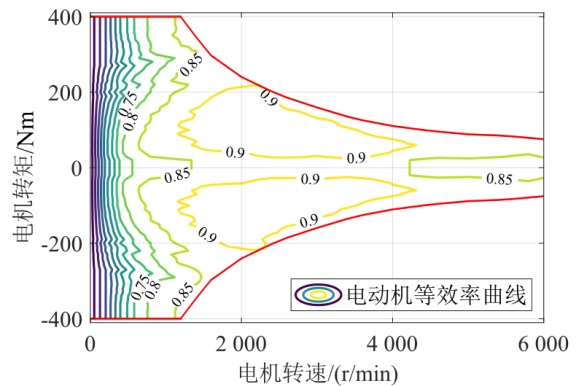


图 2 电动机效率模型

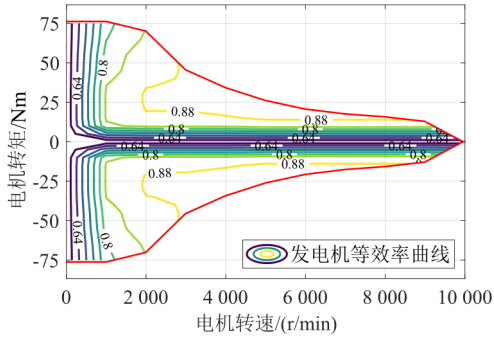


图3 发电机效率模型

1.1.4 电池建模

SOC为在HEV行驶中的关键状态变量，本文忽略电池温度及其老化的影响，建立基于内阻的电池模型，如图4所示，SOC的数学模型为：

$$\begin{cases} P_{\text{batt}}(t) = V_{\text{oc}}(t)I(t) - R_0 \times I^2(t), \\ I(t) = \frac{V_{\text{oc}}(t) - \sqrt{V_{\text{oc}}^2(t) - 4 \times R_0 \times P_{\text{batt}}(t)}}{2 \times R_0}, \\ \text{SOC}(t) = \frac{Q_0 - \int_0^t I(t) dt}{Q}. \end{cases} \quad (5)$$

式中： $V_{\text{oc}}(t)$ 为 t 时刻的开路电压； R_0 为内阻； $I(t)$ 为 t 时刻的电池电流； $P_{\text{batt}}(t)$ 为 t 时刻的充放电循环输出功率； $\text{SOC}(t)$ 为 t 时刻的电池荷电状态； Q_0 为初始电池容量； Q 为电池标称容量。

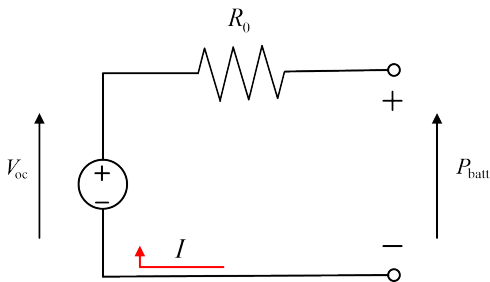


图4 基于内阻的电池模型

1.2 非线性车辆队列建模

如图5所示，本文研究对象为功率分流式HEV均质队列，包含1辆领航车与4辆跟随车，采用前车-领航者跟随式（Predecessor-Leader Following, PLF）通信拓扑方式进行信息交换。

本文只考虑传动系、空气阻力、轮胎摩擦力、

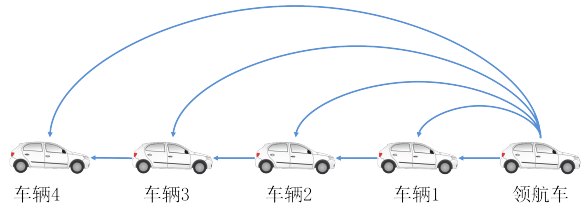


图5 采用PLF通信拓扑结构的均质车辆队列

滚动阻力、重力等因素，采用非线性动力学对队列控制模型进行建模，如式（6）所示。

$$\begin{cases} s_i(t+1) = s_i(t) + v_i(t)\Delta t, \\ v_i(t+1) = \\ v_i(t) + \frac{1}{m_i} \left(\frac{\eta_i}{r_i} T_i(t) - \frac{C_D A_i v_i^2(t)}{21.15} - m_i g f_i \right) \Delta t, \\ T_i(t+1) = T_i(t) - \frac{1}{\tau_i} T_i(t)\Delta t + \frac{1}{\tau_i} u_i(t)\Delta t. \end{cases} \quad (6)$$

式中： $s_i(t)$ 、 $v_i(t)$ 分别为车辆 i 在 t 时刻的位移与速度； Δt 为时间间隔； m_i 为车辆 i 的质量； η_i 为车辆 i 的传动系统机械效率； τ_i 为车辆 i 时滞常数； C_D 为空气阻力系数； A_i 为车辆 i 迎风面积； g 为重力加速度； $T_i(t)$ 为车辆 i 在 t 时刻的实际转矩； $u_i(t)$ 为车辆 i 在 t 时刻的期望需求转矩。

为了得到更简洁的形式，将车辆 i 的状态变量定义为 $x_i(t) = [s_i(t), v_i(t), T_i(t)]^T \in \mathbf{R}^{3 \times 1}$ ，输出变量定义为 $y_i(t) = [s_i(t), v_i(t)]^T \in \mathbf{R}^{2 \times 1}$ ，如式（7）所示。

$$\begin{cases} x_i(t+1) = \phi_i(x_i(t)) + \psi_i \times u_i(t), \\ y_i(t) = \gamma x_i(t). \end{cases} \quad (7)$$

式中： $\phi_i =$

$$\begin{bmatrix} s_i(t) + v_i(t)\Delta t \\ v_i(t) + \frac{1}{m_i} \left(\frac{\eta_i}{r_i} T_i(t) - \frac{C_D A_i v_i^2(t)}{21.15} - m_i g f_i \right) \Delta t \\ T_i(t) - \frac{1}{\tau_i} T_i(t)\Delta t \end{bmatrix}$$

$$\in \mathbf{R}^{1 \times 3}, \psi_i = \left[0, 0, \frac{1}{\tau_i} \Delta t \right]^T \in \mathbf{R}^{3 \times 1}, \gamma = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$$\in \mathbf{R}^{2 \times 3}.$$

对于由 N 辆车组成的队列，其状态矩阵 $\mathbf{X}(t)$ 、控制矩阵 $\mathbf{U}(t)$ 以及输出矩阵 $\mathbf{Y}(t)$ 可以定义为：

$$\begin{cases} \mathbf{X}(t) = [x_1^T(t), x_2^T(t), \dots, x_N^T(t)]^T, \\ \mathbf{U}(t) = [u_1(t), u_2(t), \dots, u_N(t)]^T, \\ \mathbf{Y}(t) = [y_1^T(t), y_2^T(t), \dots, y_N^T(t)]^T. \end{cases} \quad (8)$$

因此, 队列的纵向动力学模型可以被表示为:

$$\begin{cases} \mathbf{X}(t+1) = \Phi \cdot \mathbf{X}(t) + \Psi \cdot \mathbf{U}(t), \\ \mathbf{Y}(t) = \Gamma \cdot \mathbf{X}(t). \end{cases} \quad (9)$$

式中: $\Phi = [\Phi_1^T(x_1), \Phi_2^T(x_2), \dots, \Phi_N^T(x_N)]^T \in \mathbf{R}^{3N \times 1}$;

$$\Psi = \text{diag}\{\varphi_1, \varphi_2, \dots, \varphi_N\} \in \mathbf{R}^{3N \times N};$$

$$\Gamma = \text{diag}\{\gamma, \gamma, \dots, \gamma\} \in \mathbf{R}^{2N \times 3N}.$$

1.3 基于NDMPC控制算法设计

对于非线性的车辆队列模型, 本文采用NDMPC控制算法进行控制。在控制过程中, 每个车辆都可以通过V2V技术按照PLF的通信拓扑方式获取领航车与前车的信息, 并结合自身状态在每个控制时域中构建优化问题。通过对这些优化问题滚动求解, 从而实现队列的纵向控制。

本文采用固定车间距, 即每个相邻车辆之间的期望距离为定值 d_0 ^[17]。领航车的位移与速度分别用 $s_0(t)$ 和 $v_0(t)$ 表示, 因此, 队列中任意一辆车的期望状态与期望输入可以被表示为:

$$\begin{cases} x_{i, \text{des}}(t) = [s_{i, \text{des}}(t), v_{i, \text{des}}(t), T_{i, \text{des}}(t)]^T, \\ u_{i, \text{des}}(t) = T_{i, \text{des}}(t). \end{cases} \quad (10)$$

式中: $s_{i, \text{des}}(t) = s_0(t) - i \times d_0$; $v_{i, \text{des}}(t) = v_0$; $T_{i, \text{des}}(t) = h_i(v_0)$; $h_i(v_0) = \frac{R_i}{\eta_i} (C_{A,i} v_0^2 + m_i g f_i)$ 。其中, $h_i(v_0)$ 为维持车辆匀速行驶所需的转矩。

车辆队列中的信息拓扑结构可以采用有向图 $G_N = \{V_N, \varepsilon_N\}$ 和增广有向图 $G_{N+1} = \{V_{N+1}, \varepsilon_{N+1}\}$ 表示。定义队列中任意车辆 $i \in N$ 能获取信息的车辆集为 $N_i = \{i_1, i_2, \dots, i_m\}$ 以及领航车的可达集为:

$$P_i = \begin{cases} \{0\}, & p_i = 1, \\ \emptyset, & p_i = 0. \end{cases} \quad (11)$$

式中: $p_i = 1$ 为车辆 i 可以获得领航车的信息, 反之, $p_i = 0$ 。

最后对于任意车辆 i 的每个预测优化子问题, 预测的步长都为 N_p , 在此中定义 $x_i^p(k|t)$ 与 $y_i^a(k|t)$

分别为预测状态与假设状态, $u_i^p(k|t)$ 、 $u_i^*(k|t)$ 与 $u_i^a(k|t)$ 分别为预测控制输入序列、最优预测控制输入序列与假设控制输入序列, $y_i^p(k|t)$ 与 $y_i^a(k|t)$ 为预测输出序列与假设输出序列, 其中, $k \in [0, 1, 2, \dots, N_p - 1, N_p]$, 以队列行驶的舒适性与安全性为目标, 结合队列模型可以给出NDMPC中的优化问题的定义:

$$\begin{aligned} & \min_{U_i} J_i(y_i^p(:|t), u_i^p(:|t), y_i^a(:|t), y_i^a(:|t)) \\ & = \sum_{k=0}^{N_p-1} l_i(y_i^p(k|t), u_i^p(k|t), y_i^a(k|t), y_i^a(k|t)) \quad \circ \\ & \text{s.t.} \end{aligned} \quad (12a)$$

$$\begin{aligned} \dot{x}_i^p(k+1|t) &= \phi_i(x_i^p(k|t)) + \psi_i \times u_i^p(k|t) \\ y_i^p(k|t) &= \gamma x_i^p(k|t) \\ k &= 0, \dots, N_p - 1 \\ x_i^p(0|t) &= x_i(t). \end{aligned} \quad (12b)$$

$$u_i^p(k|t) \in U. \quad (12c)$$

$$y_i^p(N_p|t) = \frac{1}{|I_i|} \sum_{j \in I_i} (y_j^a(N_p|t) - \tilde{d}_{j,i}). \quad (12d)$$

$$T_i^p(N_p|t) = h_i(v_i^p(N_p|t)). \quad (12e)$$

式中: 式(12b)为预测时的动力学模型; 式(12c)为控制变量的幅值约束, U 为其取值范围; 式(12d~e)为预测时域终端约束, 其中, I_i 为集合 $I_i = N_i \cup P_i$ 的基数, 即车辆 i 能接收到信息的车辆数目, $\tilde{d}_{i,j} = [d_{ij}, 0]^T$ 为车辆 i 与车辆 j 之间的距离; 式(12d)保证预测时域终端输出为 I_i 中所有车辆输出的均值; 式(12e)确保预测时域终端处于稳态, 即匀速行驶; 式(12a)中的目标函数 l_i 定义为:

$$\begin{aligned} l_i(y_i^p(k|t), u_i^p(k|t), y_i^a(k|t), y_i^a(k|t)) &= \\ & \| \mathbf{Q}_i (y_i^p(k|t) - y_{i, \text{des}}(k|t)) \|_2 + \\ & \| \mathbf{R}_i (u_i^p(k|t) - h_i(v_i^p(k|t))) \|_2 + \\ & \| \mathbf{F}_i (y_i^p(k|t) - y_i^a(k|t)) \|_2 + \\ & \sum_{j \in N_i} \| \mathbf{G}_i (y_i^p(k|t) - y_j^a(k|t) - \tilde{d}_{i,j}) \|_2. \end{aligned} \quad (13)$$

式中: \mathbf{Q}_i 、 \mathbf{R}_i 、 \mathbf{F}_i 、 \mathbf{G}_i 为权重矩阵, 且皆为半正定矩阵, 即 $0 \leq \mathbf{Q}_i$ 、 \mathbf{F}_i 、 $\mathbf{G}_i \in \mathbf{R}^{2 \times 2}$, $0 \leq \mathbf{R}_i \in \mathbf{R}$ 。

\mathbf{Q}_i 的值由车辆 i 是否得到领航车的信息决定,

若得到领航车信息, 则 $Q_i > 0$, 否则, $Q_i = 0$ 。 R_i 用来限制控制序列的过度变化, 确保车辆能尽可能地平稳行驶, 以满足舒适性要求。 F_i 用来限制预测输出与假设输出之间的差值, 确保车辆能按照假设的输出行驶。 G_i 用来限制车辆 i 与根据所获取信息的车辆所假设的输出之间的差距, 确保车辆能按照队列中的相对位置行驶, 不会发生车辆队形的改变, 以满足安全性要求。

对于该优化问题, 整个车队的目标函数单步迭代下降如式 (14) 所示 [18]。

$$J_{\Sigma}^*(t+1) - J_{\Sigma}^*(t) \leq - \sum_{i=1}^N L_i(y_i^*(1|t), u_i^*(0|t), y_i^a(1|t), y_i^s(1|t)) + \sum_{k=1}^{N_s-1} \varepsilon_{\Sigma}(k) \quad (14)$$

式中: $J_{\Sigma}^*(t) = \sum_{i=1}^N J_i^*(t)$ 为车辆 1 到 N 在时刻 t 的最优目标函数值之和; $\varepsilon_{\Sigma}(k)$ 为:

$$\varepsilon_{\Sigma}(k) = \sum_{i=1}^N \left[\sum_{j \in O_i} \|G_j(y_i^*(k|t) - y_i^a(k|t))\|_2 \right] - \left[\sum_{i=1}^N \|F_i(y_i^*(k|t) - y_i^s(k|t))\|_2 \right] \quad (15)$$

式中: O_i 为能接收车辆 i 信息的车辆集合。

因为信息拓扑结构 G_{N+1} 中存在以领航车为根节点的有向生成树, 所以预测终端会渐近收敛于其期望状态 [19], 即预测最优输出序列 $y_i^*(k|t)$ 与假设输出序列 $y_i^a(k|t)$ 会逐渐接近, 这时 $\varepsilon_{\Sigma}(k)$ 将会趋近于 0, 即 $J_{\Sigma}^*(t+1) \leq J_{\Sigma}^*(t)$ 。根据 Lyapunov 稳定性分析可知, NDMPC 控制下的车队是渐近稳定的。

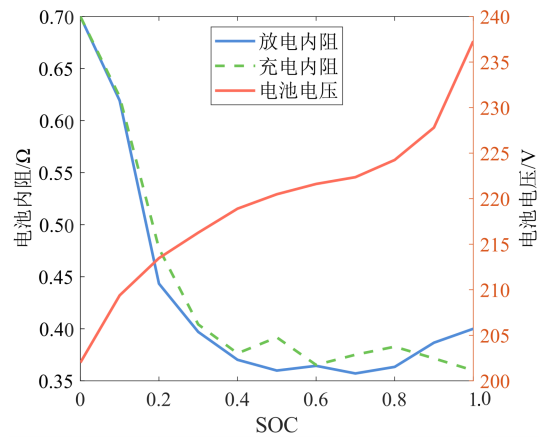
2 下层基于 DDPG 的能量管理策略

本文在下层设计了一种基于专家经验的 DDPG 控制算法, 将 HEV 的电池特性以及发动机的燃油最佳经济曲线嵌入到强化学习之中, 提升算法的优化效果。

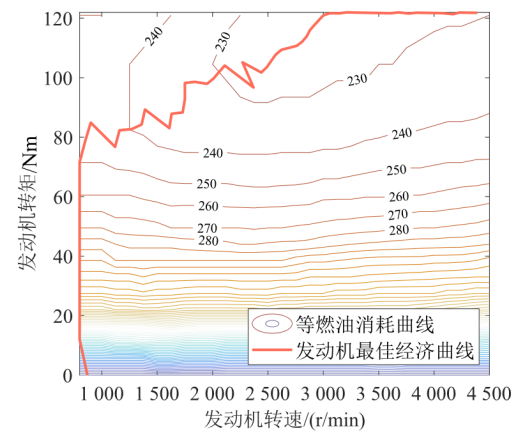
2.1 HEV 的专家经验

在 HEV 中, 电池内阻会受到 SOC 影响, 如图 6a 所示。当 SOC 在 0.4~0.8 范围内, 电池内阻相对

较小, 电池效率较高。发动机油耗 map 图如图 6b 所示, 当发动机沿着最佳经济曲线运行时, 发动机的效率达到最优, 因此对于特定发动机功率, 可通过查表得到油耗最低点的转矩与转速, 使功率分流式 HEV 系统的控制量由转矩与转速变为功率, 减少了控制变量的数量, 大幅减少了发动机的动作空间, 提升了算法的运算速度。



(a) 电池特性曲线



(b) 发动机 map 图

图 6 专家经验

2.2 基于 DDPG 的能量管理策略设计

DDPG 是基于连续动作的强化学习算法, 它是 Actor-Critic 结构的扩展, 主要的思想是将 Actor 与 Critic 都替换为神经网络, 是对 DQN 算法的一种优化, 具体结构如图 7 所示。

为了确保收敛性与稳定性, 采用金字塔结构的神经网络 [20], 网络规模逐渐减小, 同时 Actor 网络

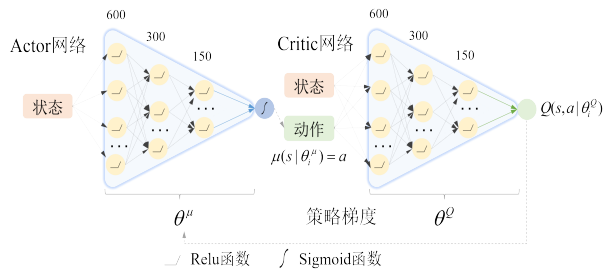


图 7 Actor-Critic 结构

与 Critic 网络结构相似，有 3 个隐藏层，每层依次包含 600、300、150 个神经元。虽然 DDPG 算法仍存在训练效率低、未知环境下随着搜索空间的增加收敛不稳定等缺点，但是专家经验的嵌入可以在一定程度上缓解上述问题。

队列下层能量管理的主要目的是在满足所有约束条件下，在确保 SOC 均衡的同时降低整车油耗，所以用来 DDPG 训练的奖励函数可以设置为：

$$\text{Reward}_i = -c(\text{SOC}_i(t) - \text{SOC}_i(t_0))^2 - b \int_{t_0}^t \hat{m}_i(t) dt \quad (16)$$

式中：Reward_{*i*} 为车辆 *i* 的奖励值；*b* 为燃油消耗的权重系数；*t*₀ 为起始时间；*c* 为 SOC 维持均衡的权重系数。

在 DDPG 算法中，状态变量包括了速度、加速度以及 SOC，可以表示为：

$$\text{State}_i = \{v_i(t), a_i(t), \text{SOC}_i(t)\} \quad (17)$$

式中：State_{*i*} 为车辆 *i* 的状态空间；*a_i(t)* 为车辆 *i* 在 *t* 时刻的加速度。

本文的研究对象为功率分流式汽车，由式 (1) 可知，确定发动机的转矩转速，即确定电动机与发电机的转矩转速，而发动机的转矩转速可以通过功率查表可得，所以算法中的动作可以只有发动机的功率，表示为：

$$\text{Action}_i = \{P_{ei}(t)\} \quad (18)$$

式中：Action_{*i*} 为车辆 *i* 的动作空间；*P_{ei}(t)* 为车辆 *i* 在 *t* 时刻的发动机功率。

在 DDPG 算法中，限制条件来源于车中的发动机、电机以及电池，即发动机与电机的转矩转速，

电池的 SOC 及充放电功率，可以表示为：

$$\begin{cases} T_{\text{eng, min}} \leq T_{\text{eng}}(t) \leq T_{\text{eng, max}}, \\ \omega_{\text{eng, min}} \leq \omega_{\text{eng}}(t) \leq \omega_{\text{eng, max}}, \\ T_{\text{MG1, min}} \leq T_{\text{MG1}}(t) \leq T_{\text{MG1, max}}, \\ \omega_{\text{MG1, min}} \leq \omega_{\text{MG1}}(t) \leq \omega_{\text{MG1, max}}, \\ T_{\text{MG2, min}} \leq T_{\text{MG2}}(t) \leq T_{\text{MG2, max}}, \\ \omega_{\text{MG2, min}} \leq \omega_{\text{MG2}}(t) \leq \omega_{\text{MG2, max}}, \\ \text{SOC}_{\text{min}} \leq \text{SOC}(t) \leq \text{SOC}_{\text{max}}, \\ P_{\text{batt, min}} \leq P_{\text{batt}}(t) \leq P_{\text{batt, max}} \circ \end{cases} \quad (19)$$

式中：*P_{batt}* 为电池功率；min 与 max 分别代表车辆参数的最小与最大值。

在 DDPG 算法中，Critic 网络用于估计状态-动作对应的价值函数，而 Actor 网络则为用于生成连续动作的策略，两者皆采用神经网络来拟合。对于 Critic 网络，它的参数通过贝尔曼方程进行更新，而 Actor 网络采用策略梯度来更新网络参数，如式 (20) 所示。

$$\begin{cases} y_i = r(s_i, a_i) + \gamma Q'(s_{(t+1)_i}, \mu'(s_{(t+1)_i} | \theta_i^{\mu'}) | \theta_i^{\rho'}), \\ L_i(\theta_i^{\rho}) = E \left[\left(Q(s_i, a_i | \theta_i^{\rho}) - y_i \right)^2 \right], \\ \nabla_{\theta_i^{\rho}} L_i(\theta_i^{\rho}) = E \left[\begin{matrix} r + \gamma Q'(s_{(t+1)_i}, \\ a_{(t+1)_i} | \theta_i^{\rho'}) - \\ Q(s_i, a_i | \theta_i^{\rho}) \end{matrix} \right] \nabla_{\theta_i^{\rho}} Q(s_i, a_i | \theta_i^{\rho}), \\ \nabla_{\theta_i^{\mu}} J_i \approx E \left[\nabla_{\theta_i^{\mu}} Q(s, a | \theta_i^{\rho}) \Big|_{s=s_i, a=u(s_i | \theta_i^{\mu'})} \right] = \\ E \left[\nabla_a Q(s, a | \theta_i^{\rho}) \Big|_{s=s_i, a=\mu(s_i)} \nabla_{\theta_i^{\mu}} \mu(s | \theta_i^{\mu'}) \Big|_{s=s_i} \right]. \end{cases} \quad (20)$$

式中：*r(s_{*i*}, a_{*i*})* 为动作车辆 *i* 在状态 *s_{*i*}* 选择动作 *a_{*i*}* 的奖励；*θ_{*i*}^ρ* 为车辆 *i* 的 Critic 网络的参数；*E*(·) 为数学期望；*θ_{*i*}^μ* 为车辆 *i* 的 Actor 网络的参数。

为了训练过程的稳定性、收敛性以及减少振荡，本文采用软更新，即在经过一定训练回合次数后，逐渐地将目标网络的参数向当前网络的参数平滑更新，具体的方式为：

$$\begin{cases} \theta_i^{\rho'} \leftarrow \tau \theta_i^{\rho} + (1 - \tau) \theta_i^{\rho'} \\ \theta_i^{\mu'} \leftarrow \tau \theta_i^{\mu} + (1 - \tau) \theta_i^{\mu'} \end{cases} \quad (21)$$

式中：*θ_{*i*}^{ρ'}* 为目标 Critic 网络的参数；*θ_{*i*}^{μ'}* 为目标 Actor 网络的参数。

车队整体的控制策略结构如图 8 所示。表 2 为下层 DDPG 控制算法的伪代码。

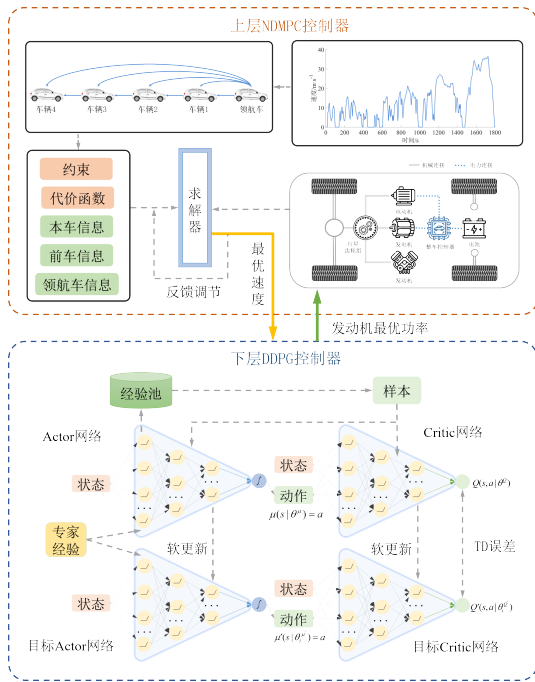


图 8 车队整体的控制策略结构

表 2 下层 DDPG 控制算法的伪代码

DDPG 算法
1 初始化: Actor 网络的参数 θ_i^μ 、目标 Actor 网络参数 $\theta_i^{\mu'}$ 、Critic 网络的参数 θ_i^Q 、目标 Critic 网络的参数 $\theta_i^{Q'}$ 以及经验回放缓冲区 R
2. 对于每个训练回合:
3 设置车辆的初始状态 S_0 ；将此时的车辆 i 的车速、加速度、SOC 传递给下层 DDPG 控制算法
4 对于每个步骤:
5. 车辆 i 通过策略网络、噪声以及专家经验选择出动作 a_{it}
6. 车辆 i 执行动作，并观察下一个状态 $s_{(t+1)i}$ 以及奖励 r_{it}
7. 将得到的经验储存在经验回放缓冲区 R 中
8. 随机从经验池 R 中抽出一小批量样本 $B=(s_{it}, a_{it}, r_{it}, s_{(t+1)i})$
9. 设置: $y_{it} = r(s_{it}, a_{it}) + \gamma Q(s_{(t+1)i}, \mu'(s_{(t+1)i} \theta_i^{\mu'}) \theta_i^Q)$
10. 通过最小化损失函数更新 Critic 网络, 即: $L_i(\theta_i^Q) = E \left[\left(Q(s_{it}, a_{it} \theta_i^Q) - y_{it} \right)^2 \right]$
11. 通过策略梯度更新 Actor 网络, 即: $\nabla_{\theta_i^\mu} J_i \approx E \left[\nabla_a Q(s, a \theta_i^Q) \Big _{s=s_{it}, a=\mu(s_{it} \theta_i^\mu)} \nabla_{\theta_i^\mu} \mu(s \theta_i^\mu) \Big _{s=s_{it}} \right]$
12. 利用软更新更新两个目标网络的, 即: $\theta_i^{Q'} \leftarrow \tau \theta_i^Q + (1 - \tau) \theta_i^{Q'}$ $\theta_i^{\mu'} \leftarrow \tau \theta_i^\mu + (1 - \tau) \theta_i^{\mu'}$
13. 将当前状态 $s_{(t+1)i}$ 设置为下一个状态 s_{it}
14. 回合结束
15. 训练结束, 将选择的动作转变为控制变量, 并将控制变量传递给车辆的相关执行器

上层 NDMPC 控制算法中，每辆车结合领航车、前车的信息以及车辆结构，求解出最优速度，下层控制算法接受上层控制算法预测的每辆车的速度，使 Actor 与 Critic 神经网络进行训练，最后通过训练好的神经网络输出发动机的最优功率，实现车队的能量管理。

3 仿真结果与分析

为了验证所提出算法的可行性与有效性，选用 WLTC 工况进行仿真验证，设定每辆车之间的初始距离为 20 m，期望的安全距离为 20 m，每辆车的初始 SOC 为 0.6。

3.1 算法训练

本文选择的训练工况如图 9 所示，由 NEDC、UDDS 以及 HWFET 等 8 种标准工况组成，覆盖了城市、郊区和高速公路等不同驾驶条件。

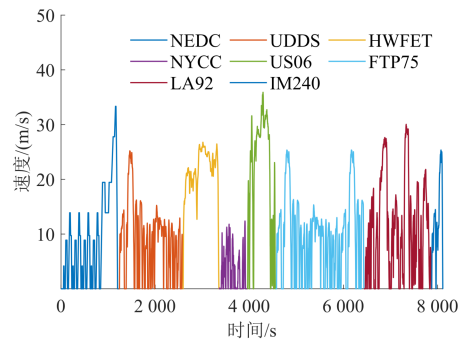


图 9 不同算法的收敛曲线

由图 10 可知，DDPG 算法收敛快，收敛后稳定，但无专家知识的 DDPG 算法收敛慢且振荡较大，同时 DQN 算法收敛较快但后续波动仍然较大，证明了所采用的算法对于提升学习效率与稳定性方面的有效性。

3.2 队列上层速度控制

图 11a 为车队速度随时间变化，从整体看，各个车辆与领航车轨迹相似，跟随情况良好，跟随车辆对加减速的反应比领航车更快，反映了上层 DNMPC 控制算法预测的特性。由局部放大图可知，车辆之间车速差距较小，跟车效果较好。

图 11b 为车队位移随时间变化，可以看出车队行驶过程中，车辆之间未发生超车和车辆退出，队

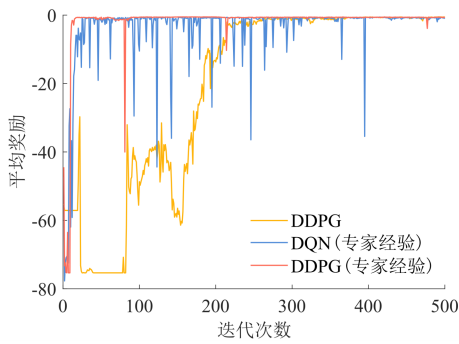
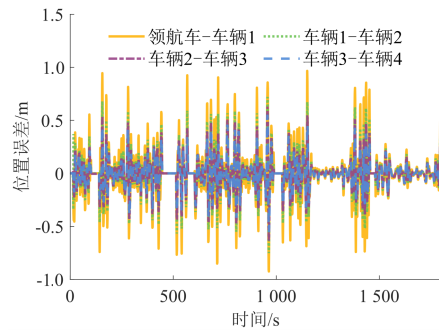


图 10 不同算法的收敛曲线

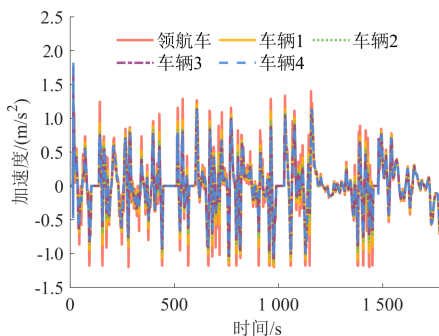
形很稳定。由局部放大图可知，在长时间的行驶后车辆间距与初始设置相比保持良好。

同时由图 11c 可知，整个行驶过程中，平均误差为 0.12 m，最大位移误差为 0.96 m，与标准值的相对误差为 4.83%，安全性较好。同时随着车辆数的增加，位移误差也越小，验证了算法的弦稳定性。

最后由图 11d 可知，加速度的变化范围绝大部分都在 $-1.5 \sim 1.5 \text{ m/s}^2$ 之间。同时随着车辆数的增多，加速度的变化越小，即当领航车出现急加减速，对于跟随车辆的影响也会减弱，使车队平均加速度仅为 0.331 m/s^2 ，车辆行驶时的舒适性也得到了保障。综上所述，上层 DN MPC 控制算法在实现车辆队列跟随性、稳定性和安全性的同时，也在一定程度上保证了行驶舒适性。



(c) 位移误差随时间变化



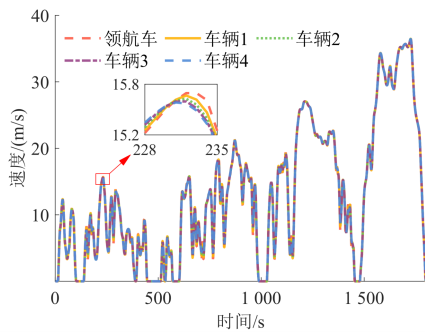
(d) 加速度随时间变化

图 11 队列上层速度跟随情况

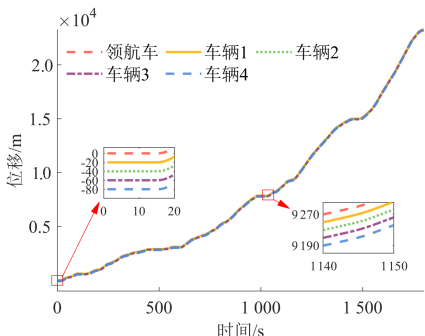
3.3 队列下层能量管理仿真

图 12a 为基于下层 DQN 控制算法的车辆队列 SOC 变化。可以看出在 DQN 控制算法下，每辆车的 SOC 都处于低电阻的范围，但是 SOC 变化范围为 $0.48 \sim 0.66$ ，跨度为 0.18，波动较明显。

但如图 12b 所示，在下层 DDPG 控制算法中，SOC 在 $0.54 \sim 0.61$ 之间，因此，电池工作在高效率区间内，且 SOC 波动较小，结合图 11c，其 SOC 的整体趋势与范围相对于 DQN 控制算法更接近于 DP 控制算法。同时由表 3 可知，采用 DDPG 控制算法的末端 SOC 仍然在初始 SOC 附近，并且相较于下层 DQN 控制算法更加接近初始值，说明所提出的算法 SOC 维持能力更强。



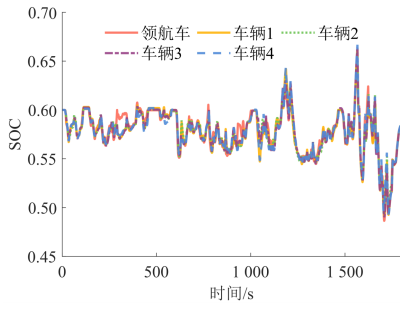
(a) 速度随时间变化



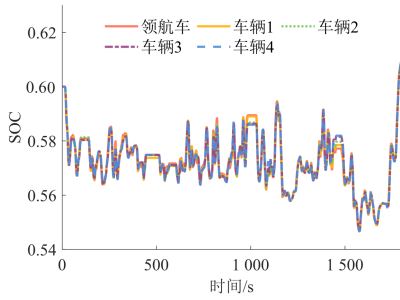
(b) 位移随时间变化

表 3 两种控制算法 SOC 末端值对比

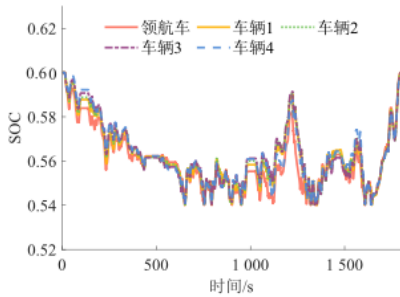
	领航车	车辆 1	车辆 2	车辆 3	车辆 4
DQN	0.589 6	0.584 7	0.583 9	0.584 1	0.585 2
对比初始值/%	1.73	2.55	2.68	2.65	2.47
DDPG	0.608 7	0.609 0	0.609 1	0.608 4	0.608 5
对比初始值/%	1.45	1.50	1.52	1.40	1.42



(a) 基于下层 DQN 控制算法的车辆队列的 SOC 变化



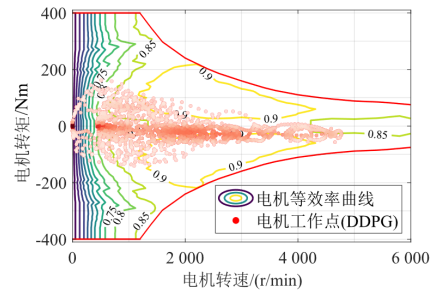
(b) 基于下层 DDPG 控制算法的车辆队列的 SOC 变化



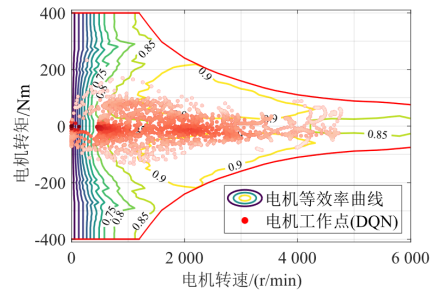
(c) 基于下层 DP 控制算法的车辆队列的 SOC 变化

图 12 三种不同下层控制算法的 SOC 对比

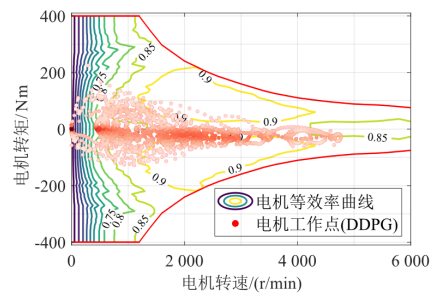
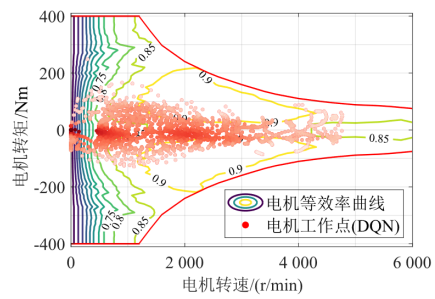
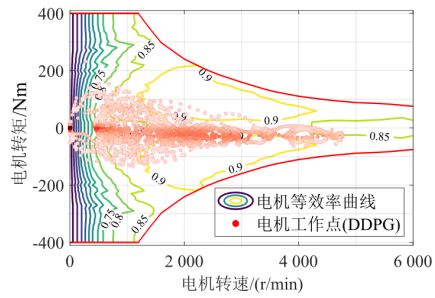
如图 13 所示，各个车辆的电动机工作点分布极为类似，侧面体现了上层算法中速度跟踪的较好效果。同时 DQN 和 DDPG 算法均使电机的工作点分布在高效区域，但是 DDPG 算法能让电机的工作点更集中。结合图 14 可知，DDPG 算法不仅减少电机在低速低效率区的分布，而且减少了电机的启停次数，延长了电机的使用寿命。



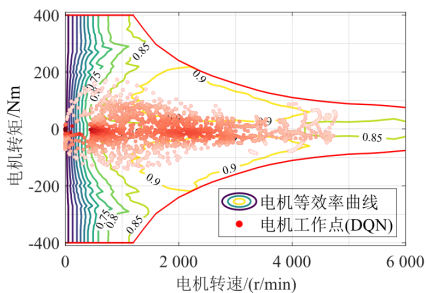
(a) 车辆 1

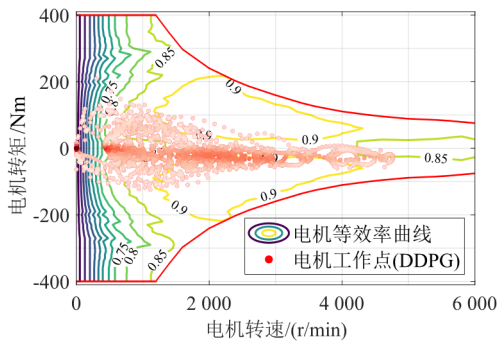
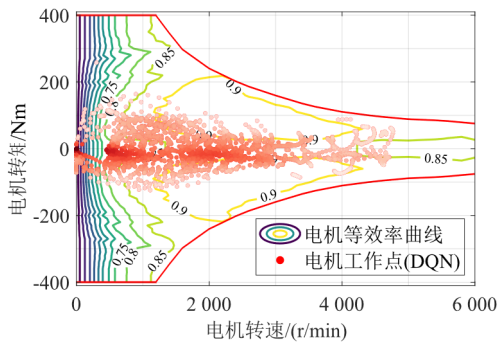


(b) 车辆 2



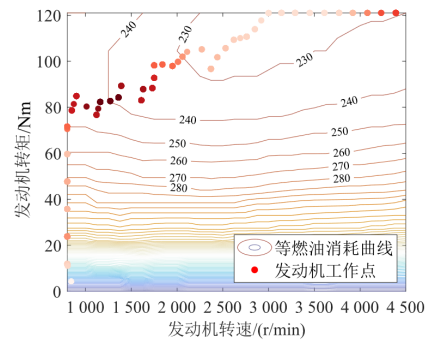
(c) 车辆 3



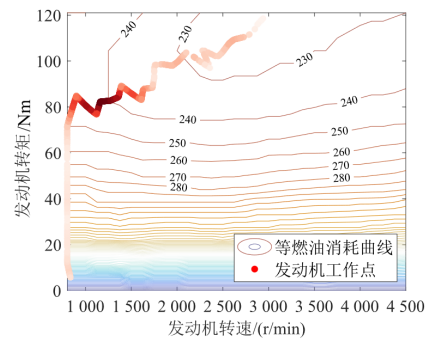


(d) 车辆4

图 13 两种不同下层控制算法的电机工作点对比



(a) DQN



(b) DDPG

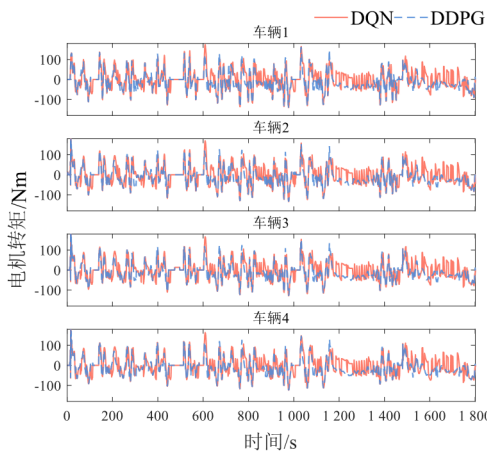
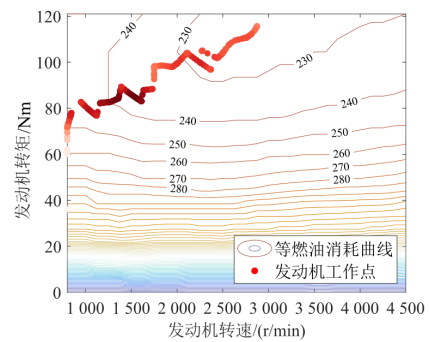
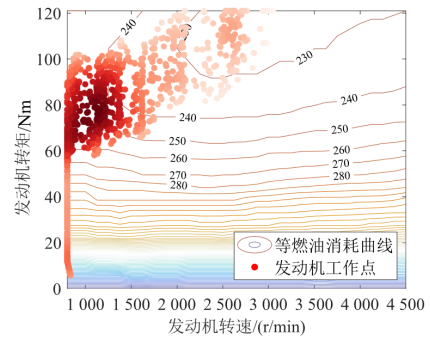


图 14 电动机转矩输出

如图 15a 所示, 发动机工作点主要集中在低油耗区域内, 但仍存在油耗较高的工作点, 同时工作点为离散的点, 体现出 DQN 算法分散动作空间的特性。如图 15b 所示, DDPG 控制算法能使发动机工作点更集中在低油耗区域, 且高油耗区域的工作点分布相对较少, 与图 15c 中 DP 的全局最优的分布更近似, 而图 15d 中没有专家经验的工作点则分布在整个发动机图中。



(c) DP



(d) DDPG(无专家经验)

图 15 不同下层控制算法的发动机工作点对比

同时结合图16可知,DDPG算法的发动机转矩变化更平滑且启停更少,延长了发动机的使用寿命,体现出DDPG算法连续动作空间的特性。

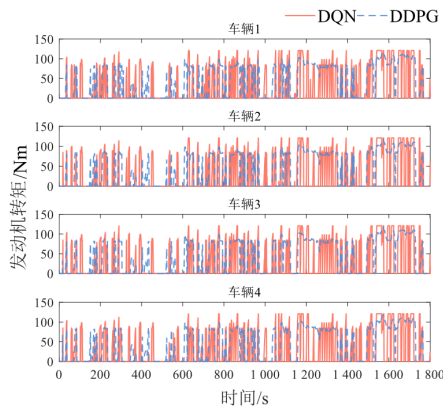


图16 发动机转矩输出

图17为下层控制算法的各个车辆燃油消耗对比,由图可知,使用DDPG控制算法相对于使用DQN算法的油耗平均减少了14.25%,最大减少了15.30%,相比于DP算法平均只增加了8.31%,验证了所提出算法在燃油经济性方面的优势。同时,DDPG算法与DP算法随着车辆数的增加而油耗减少,然而DQN算法却没有呈现这样的趋势,在验证了上层速度优化对下层能量管理提升作用的同时,还证明了DDPG算法相对于DQN算法有更好的稳定性。

4 结论

1) 本文提出一种基于智能交通系统的混合动力车队分层控制策略。上层采用NDMPC算法对车

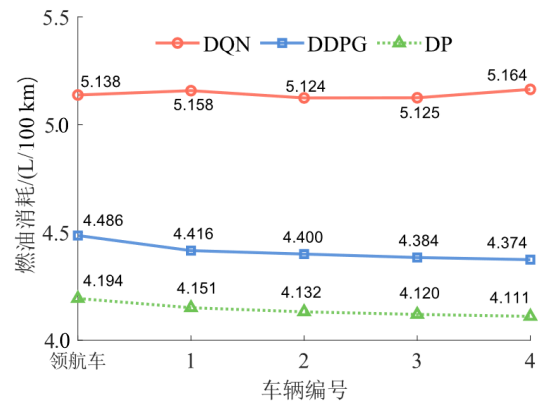


图17 燃油消耗对比

队中的各个车辆进行速度控制,下层以电池特性与发动机最优工作曲线作为专家经验设计了基于DDPG的下层能量管理算法。

2) 通过仿真验证了上层基于NDMPC的速度控制算法跟随性、安全性以及舒适性,在行驶的过程中,车队队形保持良好,其中车辆没有发生脱离以及超车的情况,车距平均误差为0.12 m,最大位移误差为0.96 m,与标准值的相对误差为4.83%,车队平均加速度为 0.331 m/s^2 ,主要集中在 $-1.5 \sim 1.5 \text{ m/s}^2$ 之间,且波动较小。

3) 通过仿真并与下层DQN算法相对比,验证了DDPG算法能更好地维持SOC均衡,且车队燃油消耗量平均减少14.25%,最大减少15.30%,相比于DP算法平均只增加了8.31%,车队的燃油经济性提升明显。

参考文献 (References)

- [1] 金辉,张子豪.基于自适应动态规划的HEV能量管理研究综述[J].汽车工程,2020,42(11):1490-1496.
JIN Hui, ZHANG Zihao. Research Review of HEV Energy Management Based on Adaptive Dynamic Programming [J]. Automotive Engineering, 2020, 42 (11): 1490-1496. (in Chinese)
- [2] 尹燕莉,黄学江,潘小亮,等.基于PID与Q-Learning的混合动力汽车队列分层控制[J].吉林大学学报(工学版),2023,53(5):1481-1489.
YIN Yanli, HUANG Xuejiang, PAN Xiaoliang, et al. Hierarchical Control of Hybrid Electric Vehicle Platooning Based on PID and Q-Learning Algorithm [J]. Journal of Jilin University (Engineering and Technology Edition), 2023, 53 (5): 1481-1489. (in Chinese)
- [3] 郭景华,王班,王靖瑶,等.智能网联混合动力汽车队列模型预测分层控制[J].汽车工程,2020,42(10):1293-1301,1334.
GUO Jinghua, WANG Ban, WANG Jingyao, et al. Hierarchical Model Predictive Control of Intelligent and Connected Hybrid Electric Vehicles Platooning [J]. Automotive Engineering, 2020, 42 (10): 1293-1301, 1334. (in Chinese)
- [4] WANG Xuan, HU Manjiang, BIAN Yougang, et al. DMPC-Based String Stable Platoon Control with Robustness Against Communication Delays [J]. Vehicular Communications, 2023, 43: 100655.1-100655.14.

- [5] SHEN Jinglai, KAMMARA E K H, DU Lili. Nonconvex, Fully Distributed Optimization Based CAV Platooning Control Under Nonlinear Vehicle Dynamics [J]. IEEE Transactions on Intelligent Transportation Systems, 2022, 23(11): 20506–20521.
- [6] GRAFFIONE S, BERSANI C, SACILE R, et al. Model Predictive Control of a Vehicle Platoon [C]//2020 IEEE 15th International Conference of System of Systems Engineering (SoSE), June 2–4, 2020, Budapest, Hungary. Piscataway NJ: IEEE, c2020: 513–518.
- [7] YUAN Junkai, SHAO Jinju, LI Xunyi, et al. Speed Planning and Energy Optimal Control of Hybrid Electric Vehicles Based on Internet of Vehicles [J]. IFAC-PapersOnLine, 2021, 54(10): 169–175.
- [8] WANG Pangwei, DENG Hui, ZHANG Juan, et al. Model Predictive Control for Connected Vehicle Platoon Under Switching Communication Topology [J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 23(7): 7817–7830.
- [9] ZHAO Shuaidong, ZHANG Kuilin. A Distributionally Robust Stochastic Optimization-Based Model Predictive Control with Distributionally Robust Chance Constraints for Cooperative Adaptive Cruise Control Under Uncertain Traffic Conditions [J]. Transportation Research Part B: Methodological, 2020, 138: 144–178.
- [10] PLUM T, WEGENER M, EISENBARTH M, et al. A Simulation-Based Case Study for Powertrain Efficiency Improvement by Automated Driving Functions [J]. Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering, 2019, 233(5): 1320–1330.
- [11] JIA Yanzhao, JIBRIN R, GÖRGES D. Energy-Optimal Adaptive Cruise Control for Electric Vehicles Based on Linear and Nonlinear Model Predictive Control [J]. IEEE Transactions on Vehicular Technology, 2020, 69(12): 14173–14187.
- [12] HE Defeng, QIU Tianxiang, LUO Renshi. Fuel Efficiency-Oriented Platooning Control of Connected Nonlinear Vehicles: A Distributed Economic MPC Approach [J]. Asian Journal of Control, 2020, 22(4): 1628–1638.
- [13] LUO Yugong, CHEN Tao, LI Keqiang. Multi-Objective Decoupling Algorithm for Active Distance Control of Intelligent Hybrid Electric Vehicle [J]. Mechanical Systems and Signal Processing, 2015, 64: 29–45.
- [14] QIU Lihong, QIAN Lijun, ZOMORODI H, et al. Global Optimal Energy Management Control Strategies for Connected Four-Wheel-Drive Hybrid Electric Vehicles [J]. IET Intelligent Transport Systems, 2017, 11(5): 264–272.
- [15] YIN Yanli, HUANG Xuejiang, ZHAN Sen, et al. Hierarchical Energy Management Control Based on Different Communication Topologies for Hybrid Electric Vehicle Platoon [J]. Journal of Cleaner Production, 2023, 412: 137414.1–137414.23.
- [16] 王越. 基于车联网信息的混合动力客车智能能量管理控制策略研究 [D]. 长春: 吉林大学, 2020.
WANG Yue. Research on Intelligent Energy Research on Intelligent Energy Management Strategy for Hybrid Electric Bus Based on Vehicular Network Information [D]. Changchun: Jilin University, 2020. (in Chinese)
- [17] 刘雪琪. 考虑信息流拓扑结构的车辆纵向队列控制研究 [D]. 长沙: 湖南大学, 2021.
LIU Xueqi. Research on Longitudinal Platoon Control Considering Information Flow Topology [D]. Changsha: Hunan University, 2021. (in Chinese)
- [18] ZHENG Yang, LI S E, LI Keqiang, et al. Distributed Model Predictive Control for Heterogeneous Vehicle Platoons Under Unidirectional Topologies [J]. IEEE Transactions on Control Systems Technology, 2017, 25(3): 899–910.
- [19] SHAHAM M H, PADIR T. Distributed Model Predictive Control for Heterogeneous Platoons with Affine Spacing Policies and Arbitrary Communication Topologies [Z]. arXiv Preprint: 2404.12441, 2024: 1–8.
- [20] XU Jingyi, LI Zirui, GAO Li, et al. A Comparative Study of Deep Reinforcement Learning-Based Transferable Energy Management Strategies for Hybrid Electric Vehicles [C]//2022 IEEE Intelligent Vehicles Symposium (IV), June 4–9, 2022, Aachen, Germany. Piscataway NJ: IEEE, c2022: 470–477.

作者简介



王浚宇 (2000–), 男, 四川广安人, 硕士研究生, 主要研究方向为新能源汽车控制策略。

E-mail: wjydx822@163.com

通信作者



詹森 (1986–), 男, 湖北十堰人, 博士, 讲师, 主要研究方向为新能源汽车动力传动系统及控制。

E-mail: fishjason@163.com