

# 基于强化学习和路况信息的燃料电池汽车 能量管理策略

宋 震<sup>1</sup>, 闵德豪<sup>1</sup>, 陈会翠<sup>1</sup>, 潘 越<sup>2</sup>, 章 桐<sup>1</sup>

(1. 同济大学 汽车学院, 上海 201804; 2. 上海捷氢科技有限公司 动力系统部, 上海 201804)

**摘要:** 为提升整车经济性和耐久性,提出了一种基于强化学习和路况信息的燃料电池汽车能量管理策略。首先,根据关键部件参数搭建了动力系统模型,并根据城市道路工况特征在 VISSIM 软件中搭建交通模型并提取了车辆行驶数据及路况数据。其次,将路况数据作为输入,利用长短期记忆神经网络对车速进行预测。最后,基于强化学习算法,将预测车速、加速度以及动力电池荷电状态作为输入,燃料电池系统功率作为输出进行能量管理策略的设计。仿真结果表明,所提策略的百公里氢耗量与动态规划策略相比仅相差 1.27%,且燃料电池系统的平均功率波动降低了 5.01%,因此可有效提升整车的经济性和耐久性。

**关键词:** 燃料电池汽车;能量管理策略;强化学习;路况信息  
**中图分类号:** U461 **文献标志码:** A

## Energy Management Strategy of Fuel Cell Vehicles Based on Reinforcement Learning and Traffic Information

SONG Zhen<sup>1</sup>, MIN Dehao<sup>1</sup>, CHEN Huicui<sup>1</sup>, PAN Yue<sup>2</sup>, ZHANG Tong<sup>1</sup>

(1. School of Automotive Studies, Tongji University, Shanghai 201804, China; 2. Powertrain System Department, Shanghai Hydrogen Propulsion Technology Co., Ltd., Shanghai 201804, China)

**Abstract:** In order to improve the fuel economy and durability of fuel cell vehicles (FCVs), an energy management strategy (EMS) for FCVs based on reinforcement learning (RL) and traffic information was proposed. First, a powertrain model was built based on parameters of key components. Then, based on the characteristics of urban road conditions, a traffic model was built in VISSIM and the vehicle driving data and traffic data were extracted. The traffic data was then used as input to predict velocity by using the long short-term

memory neural network. Finally, based on the RL algorithm and using the predicted velocity, acceleration, and the state of charge of battery as inputs, the fuel cell system power was used as the output to design the EMS. The simulation results show that the hydrogen consumption per hundred kilometers of the proposed strategy is only 1.27% different from that of the dynamic programming strategy, and fuel cell system average power fluctuation is reduced by 5.01%, effectively improving the fuel economy and durability of the vehicles.

**Key words:** fuel cell vehicle (FCV); energy management strategy (EMS); reinforcement learning; traffic information

燃料电池汽车因其零排放、能量转换效率高、续航里程长等优点近年来备受关注<sup>[1]</sup>,但其制造成本和使用成本高、耐久性较差、冷启动时间长等问题极大限制了燃料电池汽车的商业化进程。目前,燃料电池汽车通常由燃料电池和辅助电源混合驱动,不同动力源输出功率的分配方式是影响整车经济性、耐久性等性能的重要因素之一。良好的能量管理策略能有效解决动力源间的功率分配问题,有利于提升整车性能<sup>[2]</sup>。

燃料电池汽车能量管理策略可分为规则型、优化型和学习型 3 大类<sup>[3]</sup>。规则型策略基于专家经验或模糊逻辑设计,实时性好但功率分配结果达不到全局最优。优化型策略基于最优控制原理设计,功率分配结果可达到近似最优或全局最优,但实时性一般较差。近年来,基于人工智能算法的学习型能量管理策略发展迅速,包括支持向量机、人工神经网络、强化学习等。通过人工智能算法对车辆行驶数据的分析与学习,学习型策略在优化性方面可以达

收稿日期: 2021-07-20

基金项目: 国家自然科学基金青年基金项目(21805210)

第一作者: 宋震(1994—),男,博士生,主要研究方向为燃料电池混合动力汽车能量管理策略。E-mail: songzhen0621@foxmail.com

通信作者: 陈会翠(1987—),女,硕士生导师,工学博士,主要研究方向为车用燃料电池系统。E-mail: chenhuicui@tongji.edu.cn

到近似最优或全局最优,且实时性较优化策略有大幅提升。Lin等<sup>[4]</sup>提出了一种基于反向传播神经网络的速度预测方法,并基于此方法设计了燃料电池汽车的能量管理策略。结果表明,所提出的策略能有效提升经济性,整车氢耗量较规则型策略减少了17.07%。Muñoz等<sup>[5]</sup>开发了一种基于非线性自回归模型神经网络的燃料电池汽车在线能量管理策略,通过不同工况下燃料电池和动力电池间的最优功率分配结果训练神经网络。仿真结果显示,所提出的策略在高速和城际工况下可节省等效氢耗2%,在城市道路工况下可节省等效氢耗18%。林歆悠等<sup>[6]</sup>采用强化学习Q-learning算法设计了燃料电池汽车的能量管理策略,并通过硬件在环试验对策略进行验证,结果显示城市行驶工况下整车综合能耗可降低20.8%。此外,车速预测对于能量管理策略的设计也尤为重要,通过一定时域内的车速预测,能量管理策略能根据预测车速提前进行功率分配,从而优化分配结果,提升策略性能<sup>[7]</sup>。

通过以上文献分析可见,燃料电池汽车能量管理策略的目标是在行驶过程中更好地分配不同动力源间的需求功率以获得较好的整车经济性和耐久性,强化学习算法是一种能够让机器是通过不断尝试和试错,累加和学习经验,最终在环境中拿到高分的智能算法,非常适合用来解决功率分配这类问题。基于此,本文构建了城市道路交通模型并提取路况信息进行了车速预测研究,同时基于预测车速结合强化学习深度确定性策略梯度算法(deep deterministic policy gradient,DDPG)进行能量管理策略的设计。仿真结果显示,整车经济性和耐久性均得到了较好的提升。

## 1 燃料电池汽车动力系统建模

本文所研究的燃料电池汽车的动力系统型式为并联式,即燃料电池和锂离子电池分别通过DC/DC变换器和DC/AC变换器与功率总线相连,两动力源均可单独向驱动电机供电。动力系统由燃料电池系统、锂离子电池、驱动电机、DC/DC变换器、DC/AC变换器、主减速器等部件组成。具体结构如图1所示,主要参数如表1所示。

### 1.1 燃料电池模型

燃料电池建模采用功率效率模型,其功率与效率关系如图2所示。燃料电池系统峰值功率为60 kW,对应效率43%。高效率区功率为5~15 kW,最高效率为54%。

表1 动力系统主要部件参数

Tab.1 Parameters of powertrain components

部件	参数	参数值
燃料电池系统	额定功率/kW	50
	峰值功率/kW	60
锂离子电池	容量/(A·h)	37
	总线额定电压/V	350
	最大电流/A	100
	持续电流/A	72
驱动电机	荷电状态允许范围/%	5—95
	峰值功率/kW	160
	最大转速/(r/min)	12000

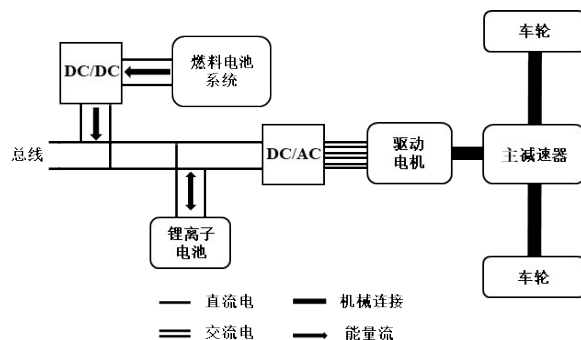


图1 燃料电池汽车动力系统构型

Fig.1 Architecture of FCV powertrain

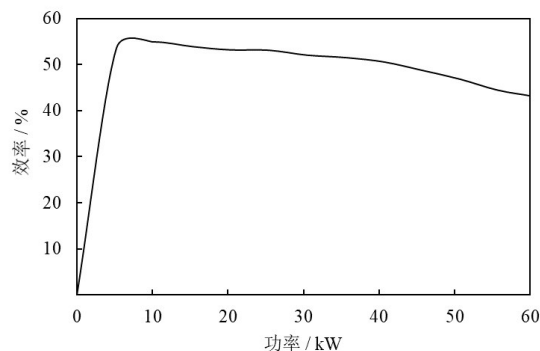


图2 燃料电池系统功率与效率关系曲线

Fig2 Fuel cell system power versus efficiency

### 1.2 锂离子电池模型

锂离子电池建模采用Rint内阻等效电路模型,其电流 $I$ 和荷电状态(SOC)的变化率 $\dot{\text{SOC}}$ 为

$$I = \frac{U_{\text{OCV}} - \sqrt{U_{\text{OCV}}^2 - 4R_{\text{ess}}P_{\text{ess}}}}{2R_{\text{ess}}} \quad (1)$$

$$\dot{\text{SOC}} = -\frac{I}{Q} = -\frac{U_{\text{OCV}} - \sqrt{U_{\text{OCV}}^2 - 4R_{\text{ess}}P_{\text{ess}}}}{2QR_{\text{ess}}} \quad (2)$$

式中: $U_{\text{OCV}}$ 为锂离子电池的开路电压; $R_{\text{ess}}$ 为锂离子电池内阻; $P_{\text{ess}}$ 为锂离子电池输出功率; $Q$ 为锂离子电池的额定容量。

## 2 基于路况信息的车速预测

### 2.1 交通道路模型搭建及路况信息提取

本文选择 VISSIM 软件进行城市道路交通模型的搭建,具体如图 3 所示。为了较为真实地还原城市实际交通状况,模型中包含道路、人行道、十字路口、交通信号灯、乘用车、商用车。其中,东西方向道路为双向 6 车道,南北方向道路为双向 4 车道,乘用车与商用车的比例为 80% 和 20%。

车辆在城市道路行驶过程中,其车速受前车的运行状态影响较大,因此在搭建的模型中我们除了提取目标车辆的车速与加速度,还提取前车的车速与加速度,这四部分将共同作为车辆预测算法的输入。

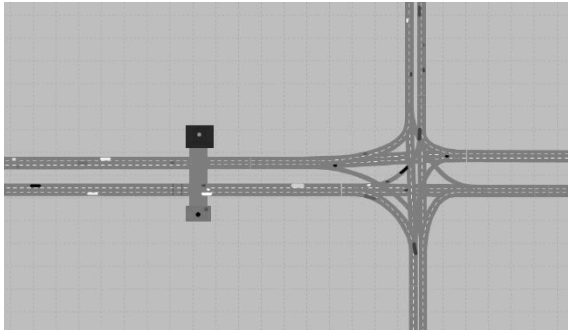


图 3 道路交通模型示意图

Fig3 Schematic diagram of traffic modell

### 2.2 基于长短期记忆神经网络的车速预测算法

车速数据通常可以看作是一种随时间变化的物理量,那么车速预测问题则可以被当成时间序列问题进行处理。基于长短期记忆神经网络(long short-term memory, LSTM)作为一种特殊的循环神经网络,可以通过对比当前与过去输入信息来优化网络参数,使之更适合处理和预测事件序列,因而能有效处理车速预测问题,故本节将采用 LSTM 进行车速预测算法的设计。

所采用的 LSTM 结构如图 4 所示,包括 1 个输入层、3 个隐藏层、1 个输出层。输入层节点数为 4 个,4 个节点分别对应神经网络的 4 个输入变量,即本车车速和加速度、前车车速和加速度。输出层为节点数为 1 个,对应下一时刻的预测车速。隐藏层节点数均为 100 个。为了达到更好的神经网络训练效果,使每一个神经元都学习到足够的信息,在第 1 和第 3 个隐藏层后进行神经元丢失(dropout)处理,丢失率为 0.2。

### 2.3 车速预测结果分析

将 VISSIM 中提取的数据分为训练集和测试集,比例分别为 80% 和 20%。为了提高网络训练的收敛速度和输出的准确性,训练前先对车速和加速度数据进行 0-1 标准化处理。基于训练集数据对 LSTM 网络进行训练优化,再通过测试集进行预测精度的验证,得出的车速预测结果和误差如图 5 所示。从图 5a 可以看出,预测车速能够较好地跟随实际车速;图 5b 显示训练集和测试集上的车速预测误差能够达到  $10^{-3}$  级,表明所设计的预测算法精度满足预期效果,可以较好地进行车速预测。

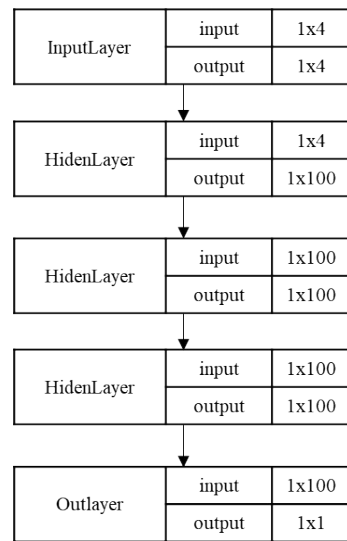


图 4 LSTM 结构图

Fig.4 Architecture of LSTM

## 3 基于强化学习的能量管理策略设计

车辆功率分配是一个连续的控制过程,而强化学习中的 DDPG 算法适合处理优化控制问题中控制变量和状态变量都连续的问题,故本节将采用 DDPG 算法进行能量管理策略的设计。

DDPG 算法由策略神经网络和评价神经网络两部分构成,策略神经网络和评价神经网络又各包含了两个神经网络,即当前网络和目标网络。其中策略当前网络和策略目标网络的输入分别为研究对象的当前和下一时刻的状态,输出分别为当前状态和下一时刻状态下所采取的动作。评价当前网络和目标网络的输入分别为研究对象当前的状态和动作与下一时刻的状态和动作,输出为当前状态和动作的价值以及下一时刻状态和动作的价值,算法的具体流程如表 2 所示。

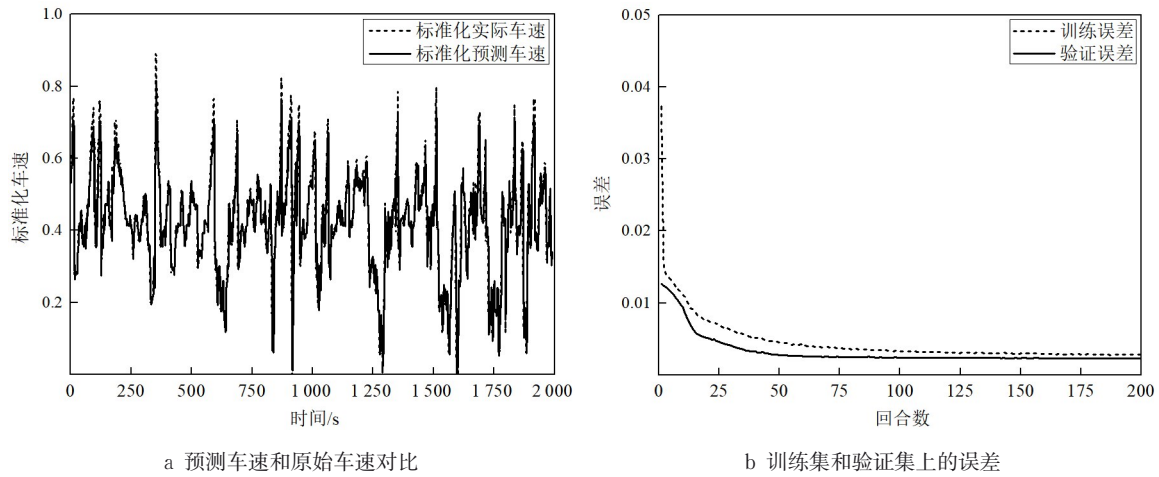


图5 车速预测结果

Fig.5 Results of velocity prediction

表2 DDPG 算法流程

Tab.2 Procedures of DDPG algorithm

DDPG算法流程
<p><b>输入:</b> 策略当前网络参数 <math>\theta^\mu</math>, 策略目标网络参数 <math>\theta^{\mu'}</math>, 评价当前网络参数 <math>\theta^Q</math>, 评价目标网络参数 <math>\theta^{Q'}</math>, 衰减因子 <math>\gamma</math>, 软更新系数 <math>\tau</math>, 经验回放采样数 <math>n</math>, 目标网络参数更新频率 <math>f</math>, 迭代回合数 <math>M</math>, 单回合迭代次数 <math>T</math>, 随机过程 <math>N</math>.</p> <p><b>输出:</b> 最优的策略当前网络参数 <math>\theta^\mu</math> 和评价当前网络参数 <math>\theta^Q</math>.</p> <ol style="list-style-type: none"> <li>1. 随机初始化 <math>\theta^\mu, \theta^Q</math>, 并令 <math>\theta^{\mu'} = \theta^\mu, \theta^{Q'} = \theta^Q</math>. 初始化经验回放池 <math>P</math></li> <li>2. for 回合数 <math>= 1:M</math> do             获得初始状态 <math>S_1</math>             for <math>t = 1:T</math> do             a) 基于策略当前网络, 输入状态 <math>S_t</math> 得到动作 <math>A_t = \mu(S_t) + N_t</math>             b) 执行动作 <math>A_t</math>, 得到新状态 <math>S_{t+1}</math> 和奖励 <math>R_t</math>             c) 将 <math>(S_t, A_t, R_t, S_{t+1})</math> 存入 <math>P</math>             d) 令 <math>S_t = S_{t+1}</math>             e) 从 <math>P</math> 中随机采样 <math>n</math> 个样本 <math>(S_i, A_i, R_i, S_{i+1}), i=1, 2, \dots, n</math>, 计算当前目标 <math>Q</math> 值 <math>y_i</math> <math display="block">y_i = R_i + \gamma Q'(S_{i+1}, \mu'(S_{i+1}), \theta^{Q'})</math>             f) 基于均方差损失函数 <math>L = \frac{1}{n} \sum_{i=1}^n (y_i - Q(S_i, A_i, \theta^Q))^2</math>, 通过梯度反向传播更新评价当前网络的参数 <math>\theta^Q</math>             g) 基于损失函数 <math>J(\theta^\mu) = -\frac{1}{n} \sum_{i=1}^n Q(S_i, A_i, \theta^Q)</math>, 通过梯度反向传播更新策略当前网络的参数 <math>\theta^\mu</math>             h) 当 <math>M \% f = 1</math>, 更新评价目标网络和策略目标网络参数: <math display="block">\theta^{Q'} = \tau \theta^Q + (1 - \tau) \theta^{Q'}</math> <math display="block">\theta^{\mu'} = \tau \theta^\mu + (1 - \tau) \theta^{\mu'}</math>             end for           end for </li> </ol>

表2中的状态  $S$  具体为车速、加速度和锂离子电池的 SOC, 动作  $A$  为燃料电池系统输出功率。奖励函数  $R$  包含4个部分: 燃料电池系统的瞬时氢耗、锂离子电池 SOC 的维持成本、燃料电池系统效率和燃料电池系统功率变化, 其具体形式如下:

$$R = -\{\alpha [\text{fuel}(t)] + \beta [\text{SOC}_{\text{ref}} - \text{SOC}(t)]^2 - \gamma [\text{eff}_{\text{fc}}(t)]^2 + \omega [P_{\text{fcs}}(t) - P_{\text{fcs}}(t-1)]^2\} \quad (3)$$

式中:  $\alpha$  是氢耗量的权重;  $\beta$  是锂离子电池 SOC 维持的权重;  $\text{SOC}_{\text{ref}}$  是 SOC 的参考值 (本文取 0.6);  $\gamma$  是燃料电池系统效率的权重;  $\text{eff}_{\text{fc}}$  是燃料电池系统效率;  $\omega$  是燃料电池系统输出功率变化的权重;  $P_{\text{fcs}}(t)$  和  $P_{\text{fcs}}(t-1)$  是燃料电池系统当前时刻的输出功率和上一时刻的输出功率。奖励函数的负号表示 DDPG 算法将朝着使奖励值更大的方向优化函数参数。



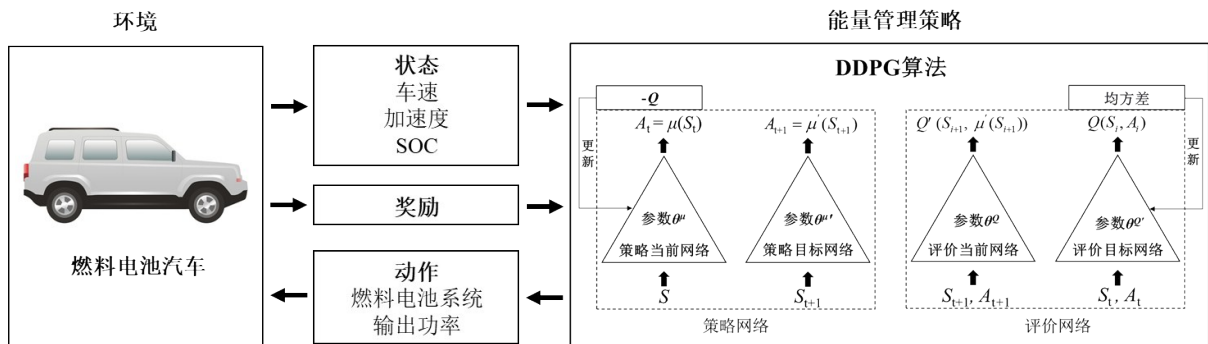


图 6 能量管理策略运作机理

Fig.6 Principle of EMS

图6展示了本文设计的能量管理策略的具体运作机理。燃料电池汽车作为与DDPG算法交互的外部环境,提供车速、加速度和SOC作为DDPG算法的输入,并通过奖励函数的值确定算法的优化的方向,DDPG算法根据车辆状态输出燃料电池系统输出功率。

4 结果与讨论

通过采用预测车速作为训练数据对DDPG算法进行训练,得出训练好的燃料电池系统和锂离子电池的输出功率变化曲线如图7、图8所示,锂离子电池的SOC的变化轨迹如图9所示。此外,本文还采用动态规划算法(dynamic programming, DP)进行相同车速下的仿真,得到了经济性全局最优的功率分配结果,其燃料电池系统和锂离子电池的输出功率结果、锂离子电池的SOC变化结果与DDPG策略的结果分别共同绘制在图7—图9中。两种策略的仿真结果在经济性和耐久性方面的对比如表3所示。

为了维持锂离子电池的SOC的首末值平衡,需

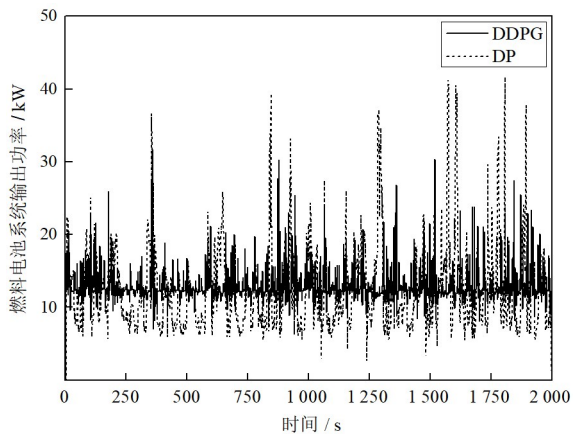


图 7 燃料电池系统输出功率

Fig.7 The output power of fuel cell system

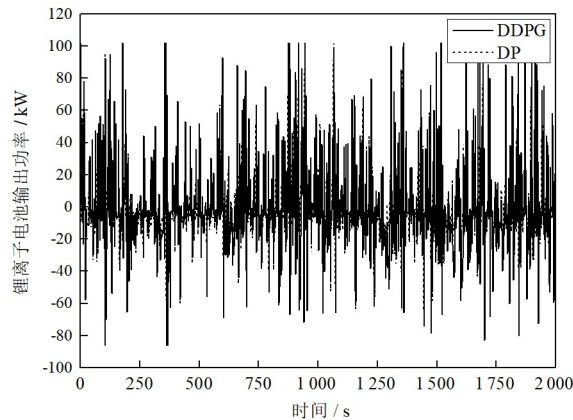


图 8 锂离子电池输出功率

Fig.8 The output power of lithium-ion battery

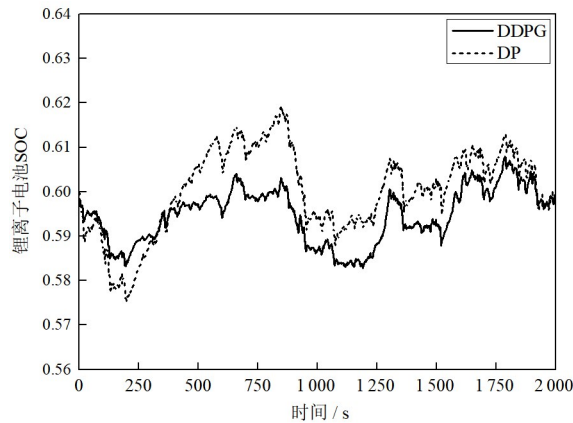


图 9 锂离子电池 SOC

Fig.9 SOC curve of lithium-ion battery

表 3 不同策略下的燃料经济性和功率波动对比

Tab.3 Fuel economy and average power fluctuation under different strategies

策略	SOC 终值	百公里氢耗量/kg	燃料经济性差异/%	平均功率波动/kW	功率波动差异/%
DP	0.6	1.814	—	1.776	—
DDPG	0.6	1.837	1.27	1.687	—5.01

要对DDPG策略奖励函数中参数 $\beta$ 以及DP策略中成本函数的SOC惩罚项参数进行调节。从表3可以看出,通过对相关参数的调节,两种策略的SOC终止均为0.6,与初始值相同,且DDPG策略的SOC轨迹在整个工况区间波动较DP策略更平缓,可有效防止锂离子电池的过充和过放,有利于锂离子电池寿命的延长。

将一个循环工况的氢耗等比例换算成百公里氢耗量,可以看出DDPG策略的百公里氢耗量与DP策略的百公里氢耗量相差仅1.27%,整车经济性达到了近似最优的水平。此外,由于在奖励函数中考虑了燃料电池功率变化这一因素,因而在策略的优化过程中,DDPG算法将会在满足动力性的前提下限制燃料电池功率的剧烈变化,从图6可以看出DDPG策略的燃料电池系统输出功率轨迹较DP策略的更加平缓。此外,从表3的对比结果也可以看出,DDPG策略的平均功率波动相较于DP策略的降低了5.01%,这表明DDPG策略能够有效减少动态加载工况对燃料电池寿命的影响,有利于提升车辆的耐久性。

本文用于训练策略网络的路况数据基于具有城市实际交通状况特征的道路交通模型所导出,具有较高的复杂性和随机性,故策略在面对其他复杂程度的工况数据时可以通过不断地训练优化以适应新工况,适应性和鲁棒性较强。在实车应用过程中,基于车辆历史工况数据对DDPG策略的参数进行训练优化,再将训练好的策略刷写进整车控制器中,DDPG策略根据预测车速进行功率的实时分配和输出,满足实车应用需求。

## 5 结论

本文设计了一种基于路况信息和强化学习DDPG算法的燃料电池汽车能量管理策略。首先在VISSIM中搭建交通模型并提取了所需路况信息,再基于路况信息和LSTM进行车速预测算法的设计,最后基于DDPG算法开发了能量管理策略。得出的主要结论如下:

(1) 基于前车车速和加速度、本车车速和加速度等信息对车辆速度进行预测,预测误差数量级达到 $10^{-3}$ ,满足预期效果。

(2) 基于DDPG算法的能量管理策略得到整车

百公里氢耗量与DP策略相差仅1.27%,同时限制了运行过程中燃料电池系统输出功率的剧烈变化,平均功率波动较DP策略降低了5.01%,有效提升了整车经济性和耐久性。

在未来的研究工作中,会将更多影响燃料电池汽车耐久性的因素与本文所设计的强化学习策略相结合进行研究,以期进一步提升整车经济性和寿命。

## 参考文献:

- [1] CHEN H, SONG Z, ZHAO X, *et al.* A review of durability test protocols of the proton exchange membrane fuel cells for vehicle[J]. *Applied Energy*, 2018, 224: 289. DOI: 10.1016/j.apenergy.2018.04.050.
- [2] YUE M, JEMEI S, GOURIVEAU R, *et al.* Review on health-conscious energy management strategies for fuel cell hybrid electric vehicles: Degradation models and strategies[J]. *International Journal of Hydrogen Energy*, 2019, 44 (13): 6844. DOI: 10.1016/j.ijhydene.2019.01.190.
- [3] 倪如尧, 刘金玲, 许思传. 燃料电池汽车能量管理控制策略研究[J]. *汽车实用技术*, 2019(1): 34. DOI: 10.16638/j.cnki.1671-7988.2019.01.011.  
NI Ruyao, LIU Jinling, Xu Sichuan. Current research of energy management strategies for FCHEVs: A review [J]. *Automobile Applied Technology*, 2019 (1): 34. DOI: 10.16638/j.cnki.1671-7988.2019.01.011.
- [4] LIN X, WANG Z, WU J. Energy management strategy based on velocity prediction using back propagation neural network for a plug-in fuel cell electric vehicle [J]. *International Journal of Energy Research*, 2020, 45(2): 2629. DOI: 10.1002/er.5956.
- [5] MUNOZ P M, CORREA G, GAUDIANO M E, *et al.* Energy management control design for fuel cell hybrid electric vehicles using neural networks [J]. *International Journal of Hydrogen Energy*, 2017, 42 (48): 28932. DOI: 10.1016/j.ijhydene.2017.09.169.
- [6] 林歆悠, 夏玉田, 魏申申. 基于增强学习算法的插电式燃料电池电动汽车能量管理控制策略[J]. *工程科学学报*, 2019, 41 (10): 1332. DOI: 10.13374 /j.issn2095--9389.2018.10.15.001.  
LIN Xinyou, XIA Yutian, WEI Shenshen. Energy management control strategy for plug-in fuel cell electric vehicle based on reinforcement learning algorithm [J]. *Chinese Journal of Engineering*, 2019, 41(10): 1332. DOI: 10.13374 /j.issn2095--9389.2018.10.15.001.
- [7] 刘秀杰. 基于车速测的FCV能量管理策略研究[D]. 大连: 大连理工大学, 2018.  
LIU Xiujie. Research on FCV energy management strategy based on speed prediction [D]. Dalian: Dalian University of technology, 2018.