International Conference on Applied Energy 2019
Aug 12-15, 2019, Västerås, Sweden
Paper ID: 294

A DEEP NEUROEVOLUTION BASED ENERGY MANAGEMENT STRATEGY FOR PLUG-IN HYBRID ELECTRIC VEHILCE

Yuankai Wu[1], Huachun Tan[2], Jiankun Peng[1], Yuecheng Li[1], Hongwen He [1*]

1 School of Mechanical Engineering, Beijing Institute of Technology

2 School of Transportation Engineering, Southeast University

ABSTRACT

Energy management strategy is important for improving fuel economic of hybrid electric vehicles. We present a deep neuroevolution based energy management strategy for hybrid electric vehicles, which learns optimal energy split strategies through evolution of its deep neural networks structure. We define the optimization objective of the deep neural networks by the fuel consumption and properties of target HEV. The deep neural networks controller is learnt through a parallel and evolution way. The simulation results on a standard driving cycles show that the proposed deep neuroevolution method outperforms the DRL based model, and achieves comparative performance to global-optimal method-dynamic programming.

Keywords: deep neuroevolution, plug-in hybrid electric vehicle, energy management strategies

NONMENCLATURE

Abbreviations	
EMS	Energy management strategy
HEV	Hybrid Electric Vehicle
DP	Dynamic Programming
MPC	Model Predictive Control
DRL	Deep Reinforcement Learning
DN	Deep Neuroevolution
DQN	Deep Q Networks
Symbols	
s_t	Vehicle State
v_t	Vehicle Velocity
acc_t	Vehicle Acceleration
SoC_t	State of Charge
P_{eng}	Power of the engine
θ	Neural networks parameters
F	Final return
R_F	Reshaped return
σ	Noise standard deviation
ϵ	Evolve parameters

1. INTRODUCTION

The decline of global oil inventories, as well as the serious concern on air quality has caused several challenges for vehicular industry. These challenges have encouraged the development of hybrid electric vehicles (HEVs), which have been considered as the most feasible and immediate choice by automakers. HEVs structure includes two or more energy sources with their associated energy converters. With multiple power sources, HEVs have greater flexibility to supply power to ensure the power request at the wheels, and are able to decrease the usage of internal combustion engine, thus resulting in Improved fuel economy [1,2].

Crucial to achieving the improved fuel economy is an efficient energy management strategy (EMS) for HEVs. In past, a plethora of EMSs have been proposed. The energy management of HEVs is depicted as an online energy source distribution problem. The solution for the problem can be roughly divided into three categories: rule-based, optimization-based and learning-based.

The rule-based approach can be easily developed and operated in practice. However, its robustness and generality cannot be guaranteed. Compared with optimization-based approach, rule-based HEV control methods produce inferior fuel economy. However, the optimization-based approach also encountered several challenges. The global optimization tools like dynamic programming (DP) require the future driving cycle information, which is not available in practice. As a result, model predictive control (MPC), which calculates the optimal control for the predicted future driving cycle, has become a popular approach. However, the performance of MPC is heavily dependent on the accuracy of driving cycle prediction algorithms. Nevertheless the accurate driving cycle prediction is not an easy task [3].

In recent years, the learning-based approaches have attracted significant attentions as a possible solutions to address the limitations associated with rule and optimization-based approaches. One of the most successful types of learning-based approaches in use is deep reinforcement learning (DRL). Much of the power of DRL comes from its ability to learn optimal control strategies by using deep neural networks [4]. The DRL agent takes actions under numerous environmental states to achieve the highest long-term accumulation of rewards. Successes of this approach in HEVs include energy management agents for a power-split HEV with two planetary gears [5], and series-parallel HEVs [2, 6].

An alternative learning-approach of DRL to solving control problems is using deep neuroevolution (DN) algorithm. It has been shown that DN methods can reliably train neural network policies to many control problems [7]. In this paper, we study the application of DN in EMSs of a series plug-in HEV. We found that DN achieves stronger results than DRL based methods for PHEV. Moreover, it is faster than DRL based model because it is highly parallelizable.

2. HEV POWERTRAIN MODELING AND EMS BASED ON NE

2.1 HEV powertrain model

The powertrain of the series HEV are shown in Fig 1, within which there are two power sources: the battery pack with capacity of 25Ah and the engine-generator set (EGS). The HEV s impelled by two identical electric propulsion systems: one for the front axle and the other for the rear axle. The curb weight of the HEV is 3500kg. The peak power of the engine is 95kW (3600r/min) and 306Nm (1800r/min) for its peak torque. The Generator outputs rated power (41kW) at rotation speed 2400r/min and peak power (93kW) at 4000r/min.

We use the backward approach to simulate the powertrain. The desired speed and acceleration are used as direct inputs to calculate the required power of axis. The powertrain can only work under series mode, which means that the axis power are provided by the motors and the engine and generator can be only used to recharge the battery. The torque balance equation and the rotational speed equation of the powertrain are given by

$$T_{eng} = T_{gen}, T_{mot} = T_{axle},$$
$$W_{eng} = W_{gen}, W_{mot} = W_{axle}. \quad (1)$$

T_{eng}, T_{gen} and T_{axle} are torques of the engine, generator and axle respectively. W_{eng}, W_{gen} and W_{axle} are rotational speed of the engine, generator and axle respectively. Regenerative braking is allowed when the HEV is slowdown. When the PHEB is performing regenerative braking, the engine is shutdown ($T_{eng} = T_{gen} = 0$, $W_{eng} = W_{gen} = 0$).

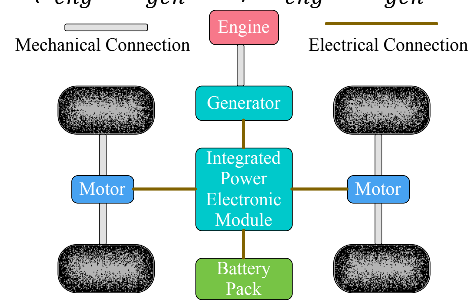

Fig 1 Powertrain of the series HEV

The battery model is established as an internal resistance model:

$$\frac{dSoC}{dt} = -\frac{(V_{oc} - \sqrt{V_{oc}^2 - 4R_{batt}P_{batt}})}{2R_{batt}C_{batt}},$$
$$P_{batt} = P_{mot} + P_{gen} \quad (2)$$

V_{oc} denotes the open-circuit voltage. R_{batt} denotes the internal resistance of battery. The charge and discharge curves of R_{batt} are different. P_{batt} denotes the battery load power, which equals the sum of motor power P_{mot} and generator power P_{gen}. The vehicle's parameters and efficiency maps of the engine and motors are calibrated from bench tests.

2.2 EMS based on NE

The deep neuroevolution based EMS is built upon deep neural networks. The energy management problem is modeled as a temporal decision process. In each time step, the network takes an action according to the state of the HEV. The neural network is a mapping between the power assignments and vehicle state. The aim of the neural network is to find the optimal strategy gives the maximum return. In this study the state s_t of the HEV are defined as a 3-dimentional vector (v_t, acc_t, SoC_t), in which v_t, acc_t and SoC_t denotes velocity, acceleration and state of charge respectively.

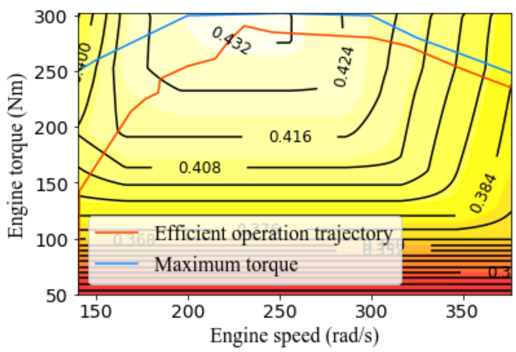

Fig 2 Efficient operation trajectory and efficiency map of the engine

For EMS of a series plug-in HEV given in Fig 2, the primary control parameters are torque and rotation speed of the engine. The work point of the engine determines the recharge power and efficiency of the battery. With a well-calibrated engine map, we can obtain the efficient operation trajectory of the engine. The action of the neural network can be defined as the output power of the engine P_{eng}. The torque and speed of the engine can then be obtained through interpolation on the operation trajectory.

The neural networks are therefore defined as

$$P_{eng,t} = NN_{\theta}(v_t, acc_t, SoC_t) \tag{3}$$

NN_{θ} is a neural network parameterized by θ. The neural network outputs the required engine power according to the velocity, acceleration and state of charge of battery in each step.

Instead of using gradient descent to optimize θ, the DN approach uses evolutionary algorithms to generate optimal parameters θ. The main benefit is that DN can be applied more widely than DRL algorithms, which require a sequence of state-action-reward pairs. In contrast, DN requires only a final return F at a task to update θ. The objective of the evolutionary strategy is to minimize the final return F.

The final return of the model is divided into three parts ($F = F_T + F_C + F_S$). For a series plug-in HEV given in Fig 1, the EMS controller should be able to recharge the battery before the battery is very low to make sure the HEV could finish the trip. For a simulation trip during training, the simulation will terminated if the state of charge is below 0.1 and the trip return is set to -1000. The cost return is set as the RMB cost (￥) of the HEV during the full trip. The resistances are relatively high when the state of charge is higher than 0.85 or lower than 0.2. A SoC return F_S is added to the final return to avoid the battery working under high resistance. $F_S = -10n(SoC_o) - f(SoC_{final})$. $n(SoC_o)$ denotes the number of time that the SoC_t is lower than 0.15 or higher than 0.85. We also restrict the final SoC of the HEV. $f(SoC_{final})$ is defined as

$$f(SoC_{final}) = \begin{cases} 20(0.3 - SoC_{final}), & if\ SoC_{final} < 0.3 \\ 0, & if\ SoC_{final} \geq 0.3 \end{cases} \tag{4}$$

In order to search for a parameter set θ that achieves the optimal final return F for HEVs, the slightly modified DN algorithm proposed in [] is used. At every iteration, a population of parameter vectors is perturbed by $\theta + \sigma\epsilon$. ϵ is a random variable. σ is a fixed value. Then we evaluate the final returns $F(\theta + \sigma\epsilon)$ of the populations on a same trip.

We only use the information provided by the population that achieved higher return ($F(\theta_g + \sigma\epsilon_i) > F(\theta_g)$) to evolve the neural networks parameter θ_g. The return $F(\theta_g + \sigma\epsilon_i), i = 1, \dots, n$ may exhibit very large variance. We reshape the return $F(\theta_g + \sigma\epsilon_i)$ according to the rank r_i of $F(\theta_g + \sigma\epsilon_i)$ in the population. The reshape function is defined as

$$R_F(\epsilon_i) = \frac{\max(0, \log_2(\frac{n_g}{2}+1) - \log_2(r_i+1))}{\sum_i^{n_g} \max(0, \log_2(\frac{n_g}{2}+1) - \log_2(r_i+1))} + \frac{1}{n_g} \tag{5}$$

Eq (5) maps all the return into the range $(\frac{1}{n_g}, 1 + \frac{1}{n_g})$. Only the population ranked in the first half outputs reshaped returns above $\frac{1}{n_g}$.

Different from the stochastic gradient descent in [], we use Adam optimizer to [] evolve neural networks parameter θ. We found that Adam optimizer could help the neural networks to avoid local optimum that never uses the engine. See algorithm 1 for pseudo-code of our proposed algorithm for NE based EMS for HEV. The process 4 in Algorithm 1 can run in parallel by adopting multithread technology, which can reduce the computation time.

Algorithm 1 Deep neuroevolutionary for energy management of hybrid electric vehicle

1: **Input:** Driving cycle D, noise standard deviation σ, initial policy parameters θ_0, number of population n.
2: **for** $g = 0,1,2, \dots$ **do**
3: Sample $\epsilon_0, \epsilon_1, \dots, \epsilon_n \sim N(0, I)$
4: Calculate final returns $F(\theta_g + \sigma\epsilon_i)$ on D using simulated HEV models
5: Reshaping final return $R_F(\epsilon_i)$ using Eq (5)
6: Updating θ_g using Adam on gradient $\frac{1}{\sigma} R_F(\epsilon_i)\epsilon_i$
7: **end for**

3. SIMULATION EXPERIMENTS

The DN based EMS model are then simulated for the input of five consecutive ChinaCity standard drive cycle since 1 drive cycle would not show desirable comparison and maximum capability of PHEV vehicle. The total distance traveled by vehicles is 35.3km. The DN based model is compared with DRL based approach deep Q networks (DQN) proposed in [8] and global-optimal method DP. In order to guarantee the comparison between DN and DQN is fair, the neural networks of DN and DQN are set with a similar structure and same hidden activation. The inputs of DN and DQN are (v_t, acc_t, SoC_t). They all have two hidden layers with 80 neurons and SELU activation. The outputs of DN is the power of the engine, whereas the outputs of DQN is the estimated Q value of action to set engine power to {−5kW,−1kW,0kW,+1kW,+5kW,+25kW,engine_off}. The initial SOC of the PHEV is set as 0.5.

Table 1 The performace of DN, DQN and DP on 6×ChinaCity

Method	RMB cost(￥)	Fuel Cost(￥)	Electricity Cost(￥)	Final SoC	Training time (s)	Evaluation time(s)
DN	10.53	8.520	2.010	0.246	612.21	0.571
DQN	12.26	12.09	0.176	0.462	1430.32	2.221
DP	10.40	8.751	1.649	0.300	\	755.85

As shown in Table 1, fuel economy of DN based energy management strategy is more close to global optimal DP compared with DQN. It only consumes 10.53 ￥ to finish the trip, which is 0.13￥higher than DP. DQN uses 12.22￥. Another advantage of DN is that it is faster in both training and evaluation process. In our experiments, DN takes around 1500 episodes to reach a stable performance. DQN takes 1000 episodes. DN is faster because it could run a large amount of episodes in parallel. DN is also faster during evaluation. Although the neural networks of DN and DQN are set with same hidden layer structure and activation function, the neural networks output of DN is a value corresponding to the engine power, the one of DQN is a vector of 7 dimensions. This small difference makes the neural networks structure of DN simpler and faster.

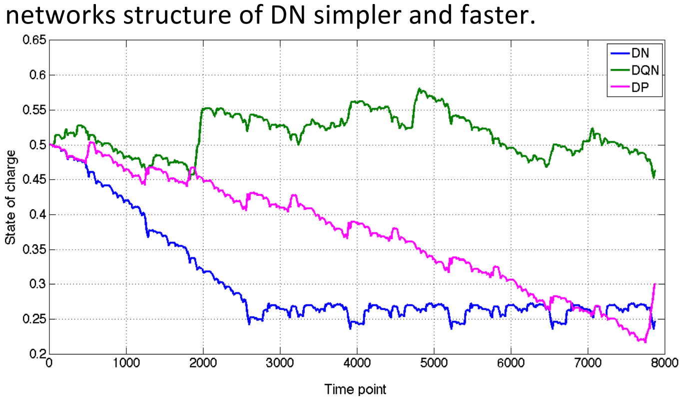

Fig 3 Battery SOC trajectories of DN, DQN and DP

Fig 3 shows SOC trajectories of DN, DQN and DP on 6×ChinaCity. Notably, DP is able to determine the best recharge timing with global driving cycle information. DN starts recharging when battery SOC is below 0.26. The reason is that the future driving cycle is unknown for DN, therefore it is not able to learn the optimal recharging behavior when SOC is high. Its energy management strategy is similar to the EMS rule that designed by expertise. DQN does not produce satisfactory SOC trajectory. It starts recharging when SOC is lower than 0.46, which resulted in higher fuel consumption. The SOC trajectories indicate that DN is more powerful for EMS of HEV compared with DRL based approach DQN.

4. CONCLUSION

In this paper, we propose a deep neuroevolution based energy management strategies for a plug-in hybrid electric vehicle. The comprehensive experiments on 6 × ChinaCity driving cycle demonstrate deep neuroevolution outperforms deep reinforcement learning approach. This work introduces a potentially alternative EMS method for HEVs.

ACKNOWLEDGEMENT

This project is supported by National Key R&D Program of China (Grand No.2018YFB0105900), and National Natural Science Foundation of China (Grant No.51705020 & No.61620106002).

REFERENCE

[1] Martinez CM, Hu X, Cao D, et al. Energy management in plug-in hybrid electric vehicles: Recent progress and a connected vehicles perspective. IEEE Transactions on Vehicular Technology. 2017;66(6):4534-49.
[2] Wu Y, Tan H, Peng J, et al.. Deep reinforcement learning of energy management with continuous control strategy and traffic information for a series-parallel plug-in hybrid electric bus. Applied Energy. 2019;247:454-66.
[3] Wu Y, Tan H, Qin L, et al. A hybrid deep learning based traffic flow prediction method and its understanding. Transportation Research Part C: Emerging Technologies. 2018; 90:166-80.
[4] Mnih V, Kavukcuoglu K, Silver et al. Human-level control through deep reinforcement learning. Nature. 2015;518(7540):529.
[5] Wu J, He H, Peng J, et al. Continuous reinforcement learning of energy management with deep Q network for a power split hybrid electric bus. Applied energy. 2018;222:799-811.
[6] Qi X, Luo Y, Wu G, et al. Deep reinforcement learning enabled self-learning control for energy efficient driving. Transportation Research Part C: Emerging Technologies. 2019;99:67-81.
[7] Salimans T, Ho J, Chen X, et al. Evolution strategies as a scalable alternative to reinforcement learning. arXiv preprint arXiv:1703.03864. 2017.
[8] Li Y, He H, Peng J, et al. Energy Management Strategy for a Series Hybrid Electric Vehicle Using Improved Deep Q-network Learning Algorithm with Prioritized Replay. DEStech Transactions on Environment, Energy and Earth Sciences. 2018.