SCIENCE CHINA
Technological Sciences

• Article •

July 2015 Vol.58 No.7: 1242–1249
doi: 10.1007/s11431-015-5852-x

A learning method for energy optimization of the plug-in hybrid electric bus

SUN Yong[1], CHEN Zheng[2,3*], YAN BingJie[2] & YOU SiXiong[2]

[1] *College of Information Science and Engineering, Ocean University of China, Qingdao 266100, China;*
[2] *State Key Laboratory of Automotive Safe and Energy, Tsinghua University, Beijing 100084, China;*
[3] *School of Sciences, Ningbo University of Technology, Ningbo 315211, China*

Received December 29, 2014; accepted April 2, 2015; published online June 11, 2015

The optimal energy management for a plug-in hybrid electric bus (PHEB) running along the fixed city bus route is an important technique to improve the vehicles' fuel economy and reduce the bus emission. Considering the inherently high regularities of the fixed bus routes, the continuous state Markov decision process (MDP) is adopted to describe a cost function as total gas and electric consumption fee. Then a learning algorithm is proposed to construct such a MDP model without knowing the all parameters of the MDP. Next, fitted value iteration algorithm is given to approximate the cost function, and linear regression is used in this fitted value iteration. Simulation results show that this approach is feasible in searching for the control strategy of PHEB. Simultaneously this method has its own advantage comparing with the CDCS mode. Furthermore, a test based on a real PHEB was carried out to verify the applicable of the proposed method.

plug-in hybrid electric (PHEB), control strategy, dynamic programming (DP), learning algorithm

Citation: Sun Y, Chen Z, Yan B J, et al. A learning method for energy optimization of the plug-in hybrid electric bus. Sci China Tech Sci, 2015, 58: 1242–1249, doi: 10.1007/s11431-015-5852-x

1 Introduction

Nowadays, the plug-in hybrid electric bus (PHEB) has been widely applied as a transportation in many cities of China [1,2]. Compared with conventional bus, more preferable fuel economy might have been achieved, due to the usage of the electric energy from the grid which are relatively more inexpensive than fossil fuels [3]. During the efficiency improvement processes of PHEB, a challenging problem has been proposed to construct an optimal energy management strategy, which might coordinate the distribution of the power demand between the engine and the electric motor (EM) [4,5].

In recent years, a large amount of approaches had been adopted in solving the energy management problem, which described via optimal control theory including dynamic programing [6–8], fuzzy logic control [9,10], Pontryagin Minimum Principle [11,12], and Model Predictive Control [13,14] in the majority of research works. Inherently, if those techniques are attempted to be applied online, it is critical to find a control strategy with some kind of driving cycle prediction [15,16]. For this purpose, some modeling methods proposed to estimate the fuel consumption cost function with a Markov chain which would give the transition probability of a set of torque demand [17,18], meanwhile utilizing the stochastic dynamic programming (SDP) in solving the cost function.

Considering the characteristics of the driving cycles of city buses, the high-confidence regularities, which would well reflect the variations of traffic flow and driving cycles

*Corresponding author (email: chenzheng@nbut.edu.cn)

about the typical routine, might be easily 'extracted' from the collected historical data [19,20]. Obviously, the SDP, which based on statistical regularity, might be the most appropriate algorithm to implement the optimization of the energy management for PHEB [21]. However, utilizing SDP algorithm to design the optimal energy management strategy also faces two challenges. First, the cost function of SDP algorithm is constructed through using the basic discrete method, which takes a constant value over each of the discretization intervals. Nevertheless, this piecewise constant representation might not be represented exactly by many smooth function, which would result in little smoothing over the inputs, and no generalization over the different discretization intervals. Second, the discretization approach owns the problem of "curse of dimensionality". The proper discretization is also necessary to obtain a good approximation. This paper describes an alternative approach for finding control strategy with stochastic Markov model of PHEB energy management, in which the cost function is approximated directly without resorting to discretization. Because the statistical learning method is introduced in this approach, it is not necessary to know all of the parameters in the MDP model. And using the approximate method, it will reduce the burden of the computation in our problem.

The remainder of this paper is organized as follows. In Section 2, a simplified model for the PHEB is proposed. The fitted value iteration method is promoted in Section 3. In Section4, the simulation is proposed to illustrate the advantage of this method. Finally, Section 5 exhibits the conclusion and discussion.

2 Control-oriented model of PHEB

The PHEB structure discussed in this paper is a typical single-shaft parallel configuration shown in Figure 1.

The single-shaft parallel configuration is widely used as the hybrid powertrain. The main difference between this configuration and the traditional ICE vehicle with automatic mechanical transmission (AMT) is that a EM is joined coaxially between the automatic clutch and AMT. In this paper, A Hengtong CKZ6116PHEV quick-charge plug-in gas/electric hybrid bus is studied, and the key parameters of this PHEB model are given in Table 1.

In [1], a backward simulation model of PHEB had been established. The torque of wheel can be expressed through

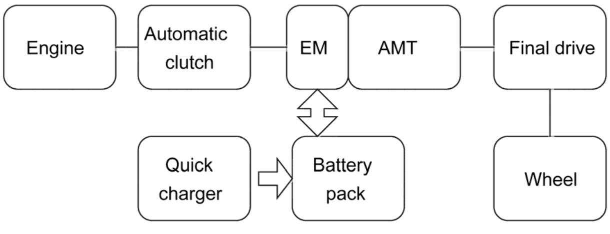

Figure 1 Configuration of the PHEB structure.

Table 1 Key parameters of the HEB

Component	Key parameter
Engine	YC6G230N, 6.454 L, nominal power: 170 kW
EM	Permanent magnet (PM), max torque: 750 Nm, nominal power: 94 kW, peak power: 121 kW
Battery	Lithium manganate, capacity: 60 A h 359 V
Transmission	6-speed AMT, gear ratio: 6.39/3.97/2.4/1.48/1/0.73
Final Drive	Ratio: 5.571

the vehicle longitudinal dynamics equation as follows

$$\begin{aligned} T_w &= \eta_T i_g i_f (T_e + T_m) + T_b \\ &= \left(mgf_r \cos\theta + \frac{1}{2} C_D \rho_d A V^2 + mg \sin\theta + \delta m \frac{\mathrm{d}V}{\mathrm{d}t} \right) r, \end{aligned} \tag{1}$$

where T_w is the torque of wheel, η_T is the transmission. efficiency, i_g and i_f denote the gear ratio of AMT and the final drive ratio, respectively. T_e is the engine torque, and T_m is the EM torque. T_b is for the braking torque acting on the wheel. m is the total mass of the vehicle which equals to the sum of the vehicle mass m_v and passenger mass m_p. The values of the above parameters are shown in the following Table 2.

The consumption rate of compressed natural gas (CNG) (mL/s) for a CNG engine can be expressed as follows:

$$Q_g = \frac{P_e b}{367.1 \rho_g g}, \tag{2}$$

where P_e=$T_e\omega_e$ is the engine power. b is the compressed natural gas consumption rate, corresponding to the current engine torque and rotational speed, which can be obtained through calibration test. ρ_g is the density of CNG.

The EM power P_{EM} can be calculated as follows

$$P_{EM} = \begin{cases} T_m\omega_m / \eta_{EM}, & \text{motor}, \\ T_m\omega_m\eta_{EM}, & \text{generator}, \end{cases} \tag{3}$$

where η_{EM} is the EM efficiency.

A simplified physical model of battery is given by the dynamics of battery SOC, internal current I, and battery power $P_{\mathrm{ess}}(W)$, which can be written as follows:

$$\dot{\mathrm{SOC}} = \frac{V_{OC} - \sqrt{V_{OC}^2 - 4R_{\mathrm{int}}P_{EM}}}{2R_{\mathrm{int}}Q_B}, \tag{4}$$

Table 2 Values of the parameters

Parameter	Value	Parameter	Value
m_v (kg)	12500	C_D	0.51
g (m/s^2)	9.8	ρ_d (kg/m^3)	1.2258
V (m/s)	-	A (m^2)	8.25
f_r	0.0076+0.0002016 V	Wheel radius r (m) C_D	0.48
θ	-	δ	1.1

$$I = -\dot{\mathrm{SOC}}\, Q_B, \tag{5}$$

$$P_{\mathrm{ess}} = V_{OC} I, \tag{6}$$

where V_{OC}, R_{int} and Q_B are the open-circuit voltage, internal resistance, and the capacity of battery, respectively.

3 Energy management strategy of PHEB based on MDP

In this section, a cost function of fuel consumption and electric consumption based on Markov decision process will be presented. Then a learning method is proposed to search for a minimal value of this cost function, and obtain the optimal control strategy simultaneously.

3.1 Markov decision process model

A Markov decision process [22] (MDP) is a tuple (X, A, $\{P_{xa}\}$, γ, R), where

(1) X is a set of states.

(2) A is a set of actions.

(3) P_{xa} are the state transition probabilities, where $x \in X$. For each $x \in X$ and action $a \in A$, P_{xa} is a distribution over the state space. More in detail, P_{xa} gives the distribution over what states we will transition to if we take action a in state x.

(4) $\gamma \in [0,1)$ is called the discount factor.

(5) $R: X \times A \to \mathbb{R}$ is the reward function. (The reward function is sometimes also written as a function of the state X only, in which case we would have $R: X \to \mathbb{R}$). Then, a cost function of fuel consumption will be implemented by the MDP.

The dynamics of an MDP proceeds as follows: it starts in some state x_0, and get to choose some action $a_0 \in A$ to take in the MDP. As a result of the choice, the state of the MDP randomly transitions to some successor state x_1, drawn according to $x_1 \sim P_{x_0 a_0}$. Meanwhile, the reward $R(x_0, a_0)$ is obtained. Then, another a_1 is picked. As a result of this action, the state x_1 transitions to some $x_2 \sim P_{x_1 a_1}$. At the same time $R(x_1, a_1)$ is obtained again. This process will be continued.

Here, let $x = [\mathrm{soc}, T_r, n] \in X$, where soc, T_r, n are the SOC of battery, torque demand, and transmission input-speed , respectively. Let $a = T_e \in A$, where T_e is the engine torque. Therefore, the optimal cost function of fuel consumption can be described as follows:

$$J(x) = R(x) + \gamma \min_a \int_{x'} P_{xa}(x') J(x') \mathrm{d}x', \tag{7}$$

where $x' \sim P_{xa}$, where x' is the next state of state x', and the reward function R is defined as the cost of fuel and electric consumption.

$$R(x,a) = \alpha_1 W_{\mathrm{fuel}} + \alpha_2 W_{\mathrm{ele}}, \tag{8}$$

where α_1 and α_2 are price of fuel and electric power respectively. W_{fuel} and W_{ele} are the consumption of fuel and electric power respectively. In addition, the $P_{xa}(x')$ in eq. (7) should be expressed.

Because of the constraints of the hybrid powertrain, this optimization process might satisfy these constrains.

$$\begin{gathered} T_e(k) + T_m(k) = T_r(k), \\ n_{e,\min} \leqslant n_e(k) \leqslant n_{e,\max}, \\ n_{m,\min} \leqslant n_m(k) \leqslant n_{m,\max}, \\ T_{e,\min} \leqslant T_e(k) \leqslant T_{e,\max}, \\ T_{m,\min} \leqslant T_m(k) \leqslant T_{m,\max}, \\ \mathrm{soc}_{\min} \leqslant \mathrm{soc}(k) \leqslant \mathrm{soc}_{\max}, \end{gathered} \tag{9}$$

where $n_{*,\min}$, $n_{*,\max}$, $T_{*,\min}$ and $T_{*,\max}$ are permitted lower and upper bound of engine or motor torque and speed, respectively. $\mathrm{soc}_{\min}$ and $\mathrm{soc}_{\max}$ are bounds of battery admissible sets. In this paper, the gear shifting strategy could be mapped as a lookup table of single parameter (vehicle speed) [23].

Then, a learning method is proposed to find the $P_{xa}(x')$. Suppose extracting m trial in which one repeatedly take actions in the above MDP as follows:

$$\begin{gathered} x_0^{(1)} \xrightarrow{a_0^{(1)}} x_1^{(1)} \xrightarrow{a_1^{(1)}} x_2^{(1)} \xrightarrow{a_2^{(1)}} \cdots \xrightarrow{a_{n-1}^{(1)}} x_n^{(1)}, \\ x_0^{(2)} \xrightarrow{a_0^{(2)}} x_1^{(2)} \xrightarrow{a_1^{(2)}} x_2^{(2)} \xrightarrow{a_2^{(2)}} \cdots \xrightarrow{a_{n-1}^{(2)}} x_n^{(2)}, \\ \cdots \\ x_0^{(m)} \xrightarrow{a_0^{(m)}} x_1^{(m)} \xrightarrow{a_1^{(m)}} x_2^{(m)} \xrightarrow{a_2^{(m)}} \cdots \xrightarrow{a_{n-1}^{(m)}} x_n^{(m)}, \end{gathered} \tag{10}$$

In the eq. (10), the action $a=T_e$ can be picked at random.

One can be chosen to learn a linear model as follows:

$$x_{t+1} = Ax_t + Ba_t + \varepsilon_t, \tag{11}$$

where $\varepsilon_t \sim N(0, \Sigma)$, which is a normal distribution. The covariance matrix Σ can be computed from the data in eq. (10). The parameters A and B in the model (10) can be obtained from the following equation:

$$\arg\min_{A,B} \sum_{i=1}^{m} \sum_{t=0}^{n} \left\| x_{t+1}^{(i)} - (Ax_t^{(i)} + Ba_t^{(i)}) \right\|^2. \tag{12}$$

Thus, a stochastic model is built, in which x_{t+1} is a random function of the x_t and a_t. The $P_{xa}(x')$ also could be given in this way.

It is worth while to note that here the next state x_{i+1} is a linear function of the current state x_t and action a_t. However, of course nonlinear functions are also available in this problem. Mathematically, one can use a model

$$x_{t+1} = Af(x_t) + Bg(a_t) + \varepsilon_t, \tag{13}$$

where f and g are some nonlinear mapping of the states and actions.

3.2 Fitted value iteration algorithm

As the above mentioned, a proper approach could be implemented to solve the optimal problem (7). In this section, a learning method is given as following. This method is called the fitted valued iteration algorithm, which approximates the value function of a MDP. The main idea of this learning method is that we are going to approximately carry out this step, over a finite sample of states $x^{(1)}, x^{(2)}, \cdots, x^{(m)}$. In detail, a supervised learning algorithm can be used as linear regression in our description below to approximate the cost function as a linear or non-linear function of the states:

$$J(x) = \boldsymbol{\theta}^{\mathrm{T}} \phi(x), \tag{14}$$

where ϕ is certain appropriate feature mapping of the states.

For each state x in finite sample of m states, fitted valued iteration will first compute a quantity $y^{(i)}$, which will be approximation to $R(x) + \gamma \max_x E_{x' \sim P_{xa}}[J(x')]$. Then, it will use a supervised learning method to get $J(x)$ close to $R(x) + \gamma \max_x E_{x' \sim P_{xa}}[J(x')]$. In other words, it means that trying to get $J(x)$ close to $y^{(i)}$.

The processes of this learning algorithm are as follows.

(1) Randomly sample m states $x^{(1)}, x^{(2)}, \cdots, x^{(m)} \in S$.

(2) Initialize $\theta = 0$.

(3) Repeat {

For $i = 1, \cdots, m$

{

For each action $a \in A$ {

Sample $x'_1, \cdots, x'_k \sim P_{x^{(i)}a}$ (Using a model of the MDP)

Set $q(a) = \dfrac{1}{k} \sum_{j=1}^{k} R(x^{(i)}) + \gamma J(x'_j)$

}

Set $y^{(i)} = \max_a q(a)$

}

Set $\theta = \arg\min_\theta \dfrac{1}{2} \sum_{i=1}^{m} (\theta^T \phi(x^{(i)}) - y^{(i)})^2$

}

As mentioned above, fitted value iteration using linear regression as the algorithm to make $J(x^{(i)})$ close to $y^{(i)}$ had written out. The step of this algorithm is analogous to a standard regression problem in which the vectors $(x^{(1)}, y^{(1)})$, $(x^{(2)}, y^{(2)}), \cdots, (x^{(m)}, y^{(m)})$, are training set, and a function mapping from x to y should be learned. The main difference is that [soc, T_r, n] plays the role of x in above algorithm. Obviously, other regression algorithms can also used in this problem.

At last, fitted valued iteration outputs J, which is an approximation to optimal J^*. If the system is in some state x, and we need to choose a control strategy, we would like to choose the strategy as follows:

$$\arg\max_a E_{x' \sim P_{xa}}[J(x')]. \tag{15}$$

Therefore, the optimal problem (7) can be solved by the approximate method and learning method. As the process of the above algorithm, the optimal control strategy is given by eq. (15).

4 Simulation validation

The driving cycle for simulation in this section starts from Yudong station (Point A in Figure 2) to Nanping station (Point B in Figure 2) in Chongqing city, and then returns to Station A, including 32 bus stops.

Next the performance of the proposed approach will be compared with a baseline control policy. Here the baseline control strategy is the CDCS strategy, and a certain test driving cycle is shown in Figure 3.

The simulation results show in details as following. As we know, the CDCS strategy tends to deplete battery energy in the first place, while the MDP makes the optimal decision based on the solving a global optimization problem. So the SOC-time curve of MDP exhibits better performance than the CDCS shown in Figure 4. In detail, the SOC comparison curves show that the proposed MDP strategy might

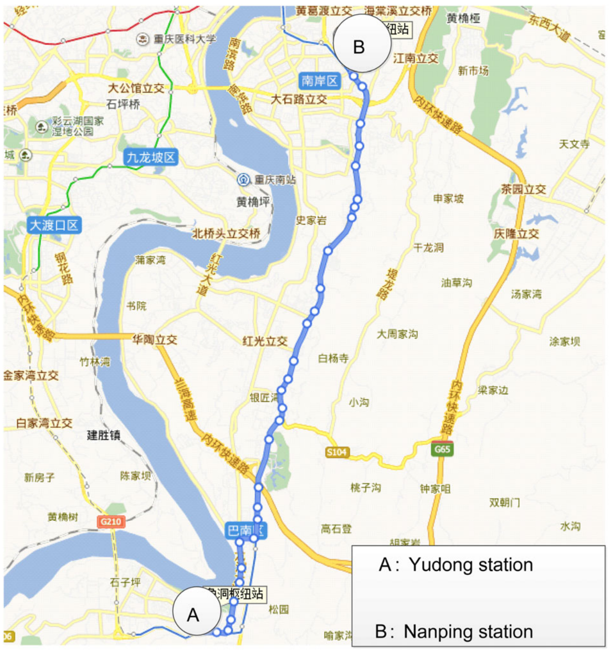

Figure 2 (Color online) Bus route 303 in Chongqing.

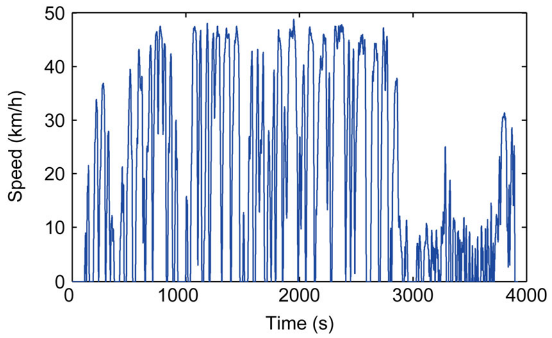

Figure 3 (Color online) Velocity-time curve of the bus route 303.

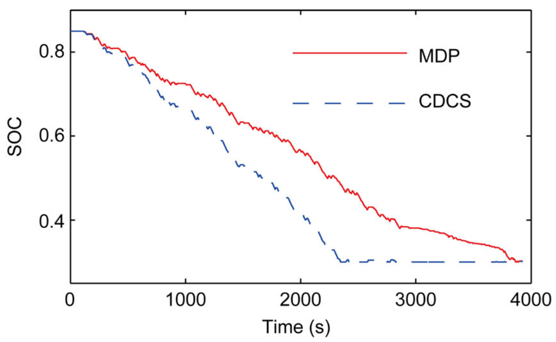

Figure 4 (Color online) SOC-time curve of the CDCS and MDP.

be very close to the global optimal electric energy distribution in the given driving cycle.

In Figure 5, the engine operating points for CDCS and MDP are shown respectively. The engine might operate in higher effective area in the MDP model compared with CDCS model.

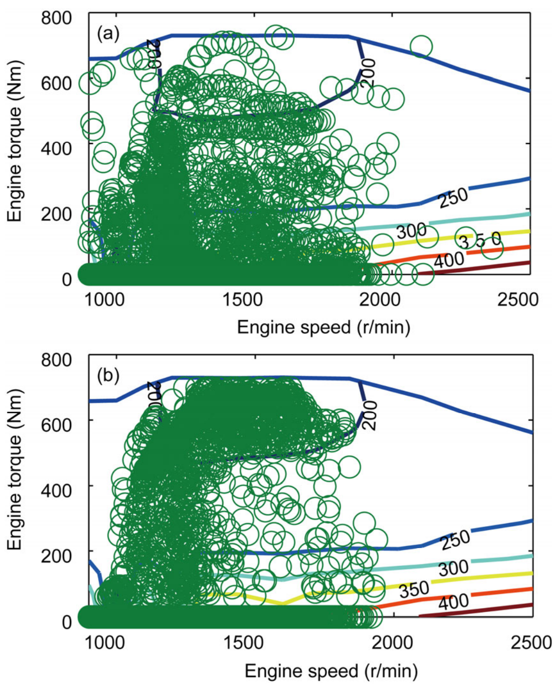

Figure 5 (Color online) (a) Engine operating points of CDCS; (b) engine operating points of MDP.

For motor working points, it can be seen from the Figure 6 that, the most of engine working points are located nearby the low-load area, which reflects lower engine efficiency. While more engine working points lie in the higher effective area due to the optimal distribution of MDP.

As shown in the Figure 7, simulation results under the bus route 303 reflect the basic curves of the PHEB in the given driving cycle. Obviously, in the MDP method the engine could well participate in the driving process almost on the whole driving cycle compared with the CDCS method. Therefore the energy of the battery could be well utilized among the whole cycle to avoid the low-efficiency areas of the engine. And the total energy consumption of the two methods are shown in Figure 8, the energy consumption of the CDCS method increase rapidly when it works on CS mode after 2500s, due to its weaker ability to keep the engine working on the high efficient areas.

As quantitative perspective, the simulation results with three strategies: CDCS, MDP, DP are shown in Table 3. The results demonstrate that the energy consumption generated by the proposed MDP strategy is higher than that generated by the standard DP algorithm, but significantly lower than that of CDCS strategy. As the results shows that the proposed MDP strategy could reduce the energy consumption through moving the operating points of two power sources into their own high efficiency areas.

In theory, the proposed learning algorithm can reduce the computational burden due to it is constructed based on a large number of statistical data about the PHEB. Compared to the proposed algorithm in this paper, classical methods, such as DP, are of limited utility in solving MDP because of

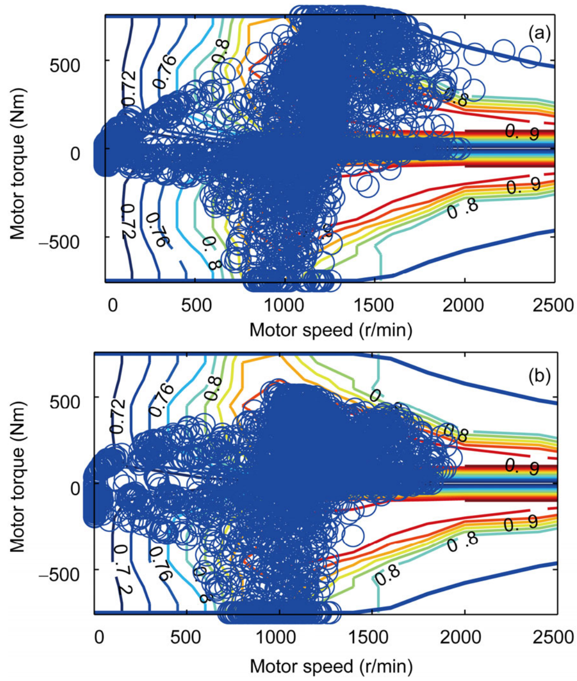

Figure 6 (Color online) (a) Motor operating points of CDCS; (b) motor operating points of MDP.

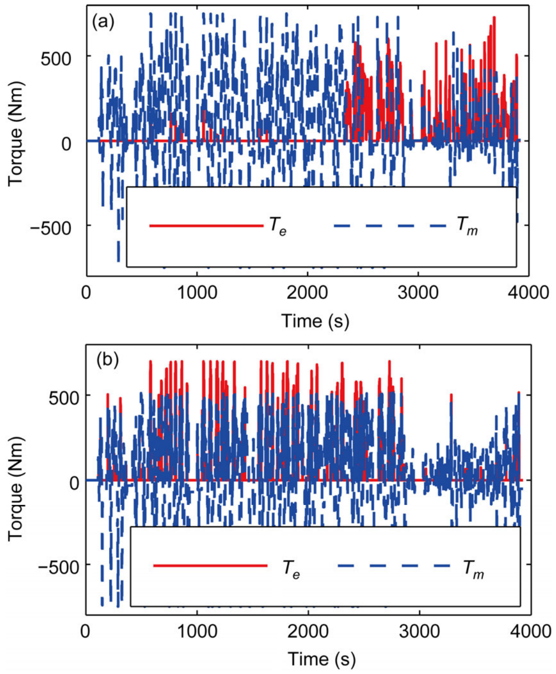

Figure 7 (Color online) (a) Torque split of CDCS; (b) torque split of MDP.

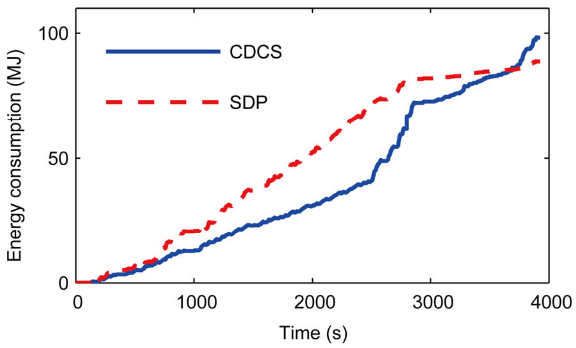

Figure 8 (Color online) The total energy consumption of MDP and CDCS.

Table 3 The fuel economy of CDCS and MDP

	CDCS	MDP	DP
Fuel consumption (m^3/(100 km))	29.45	24.12	21.53
Electric consumption (kWh/(100 km))	31.6	30.25	31.6
Cost (yuan(RMB)	149.40	126.73	117.7
Improvement (%)	-	15.17	21.2

their assumption of a perfect model and their great computational expense [24]. In addition, although the energy consumption generated by the proposed MDP strategy is higher than the standard DP algorithm, the MDP strategy is to optimize total energy consumption over a family of random driving cycles in an average sense instead of optimizing over a given driving cycle by DP. Therefore the optimized results are valid for several driving cycles rather than for a certain driving cycle, and it has the potential to be applied in actual.

5 Experiment validation

A plug-in hybrid electric bus with the single-shaft parallel configuration was used to test the proposed control method. And the segment of the Yun Longxi highway in Wujiang city, Jiangsu province, was chosen as the test cycle, the vehicle speed vs. time curve is shown in Figure 10. Note that the gearshift logic and the engine control are quite different with that in simulation even though the similar configurations are equipped in the bus (e.g. No start-stop system is working on the tested bus while in simulation it does). The tested bus and the devices are shown in Figure 9 and the tested results are shown in Figure 10.

As can be seen in Figures 10(b) and (c), the torque could be well distributed between engine and motor. However, the fuel and electric consumption are 30.07 (m^3/100 km) and 56.75 (kWh/100 km) correspondingly as shown in Figure 10(e), which are not as well as that in simulation. The differences might be caused by several reasons. Firstly, the tested bus keeps the engine in idle speed during the stop time along the tested cycles while in simulation it won't. Therefore, the extra consumption could be added. Secondly, the gearshift logic could lead to some changes for engine work points, which would have some effects on the vehicle economy. Moreover, the simplification of the model could also influence the performance of the vehicle. Even through, the tested results in actual bus are not as well as that in simulation, the applicable of the proposed control method could also be well verified. One points should be explained is that the minimum resolution ratio of SOC is 5‰, therefore, the change of the SOC is stair-step curve, as shown in Figure 10(d).

6 Conclusions

A novel learning algorithm is presented to solve the global energy optimization problem in the real-time controller of PHEB. A MDP model for the cost function of the fuel and

Figure 9 (Color online) The tested bus and devices.

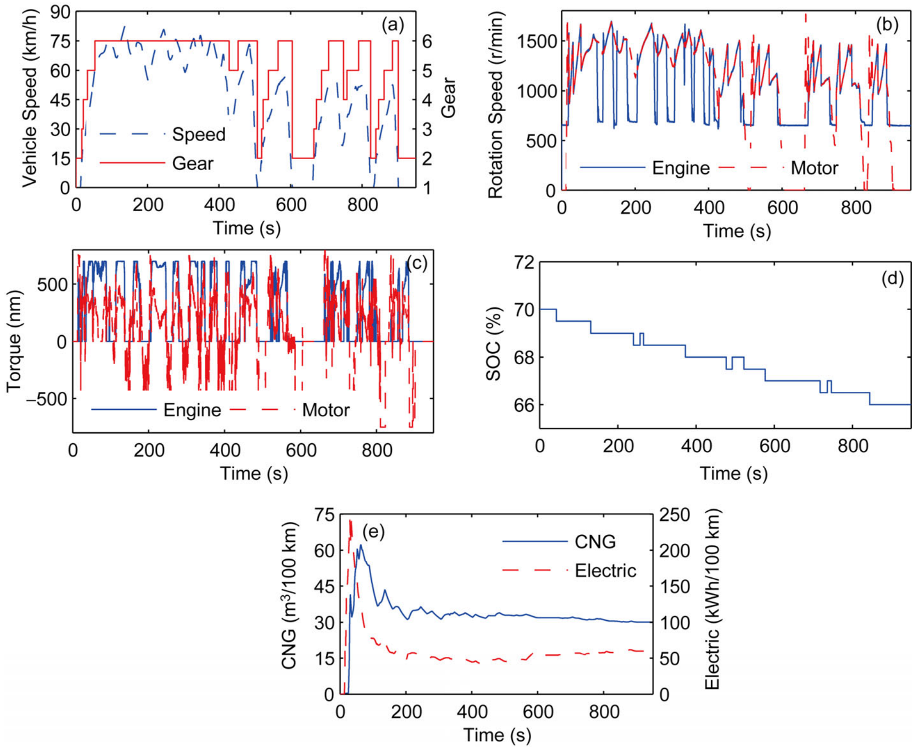

Figure 10 The results of actual experiment. (a) Vehicle speed and gear; (b) the engine and motor speed; (c) the engine and motor torque; (d) battery SOC; (e) fuel and electric consumption.

electric consumption in the PHEB is constructed by the statistical method. Then a simpler function is used for approximating the cost function, and in the process of this method, a linear regression method is adopted which make the problem much easier to solve. Moreover, sample data is easy to be obtained because PHEBs always run on a fixed route many times. At last, simulation and experiment is given to illustrate the effectiveness of this approach. Compared with the CDCS control strategy, the MDP based control strategy can get a lot better results.

This work was supported by the National Natural Science Foundation of China (Grant No. 51275557), and the National Science-technology Support Plan Projects of China (Grant No. 2013BAG14B01). The author gratefully acknowledges the support of K.C. Wong Education, Hong Kong.

1 Yang C, Jiao X H, Li L, et al. Electromechanical coupling driving control for single-shaft parallel hybrid powertrain. Sci China Tech Sci, 2014, 57: 541–549
2 Yang C, Jiao X, Li L, et al. Robust coordinated control for hybrid electric bus with single-shaft parallel hybrid powertrain. IET Control Theory Appl, 2014, 9: 270–282
3 Li L, Yang C, Zhang Y H, et al. Correctional DP based energy management strategy of plug-in hybrid electric bus for city bus route. IEEE Trans Veh Tech, 2014, doi: 10.1109/TVT. 2014. 2352357
4 Wirasingha S G, Emadi A. Classification and review of control strategies for plug-in hybrid electric vehicles. IEEE Trans Veh Tech, 2011, 60: 111–122
5 Liu J, Peng H. Modeling and control of a power-split hybrid vehicle. IEEE Trans Control Syst Technol, 2008, 16: 1242–1251
6 Zhang Y H, Jiao X H, Li L, et al. A hybrid dynamic programming-rule based algorithm for real-time energy optimization of plug-in hybrid electric bus. Sci China Tech Sci, 2014, 57: 2542–2550
7 Perez L V, Bossio G R, Moitre D, et al. Optimization of power management in an hybrid electric vehicle using dynamic programming. Math Comput Simul, 2006, 73: 244–254
8 Li L, Zhang Y, Yang C, et al. Hybrid genetic algorithm-based optimization of powertrain and control Parameters of plug-in Hybrid electric bus. J Franklin Inst, 2015, 352: 776–801
9 Abdelsalam A A, Cui S. A fuzzy logic global power management strategy for hybrid electric vehicles based on a permanent magnet electric variable transmission. Energies, 2012, 5: 1175–1198
10 Yi Z, Heping L, Huabin W. Torque control strategy for parallel hybrid electric vehicles using fuzzy logic. Wseas Trans on Syst, 2011, 10: 116–125
11 Kim N, Cha S, Peng H. Optimal control of hybrid electric vehicles based on Pontryagin's minimum principle. IEEE Trans Control Syst Technol, 2011, 19: 1279–1287
12 Hou C, Ouyang M, Xu L, et al. Approximate Pontryagin’s minimum principle applied to the energy management of plug-in hybrid electric vehicles. Appl Energy, 2014, 115: 174–189
13 Borhan H, Vahidi A, Phillips A M, et al. MPC-based energy management of a power-split hybrid electric vehicle. IEEE Trans Control Syst Technol, 2012, 20: 593–603
14 Di Cairano S, Bernardini D, Bemporad A, et al. Stochastic MPC with learning for driver-predictive vehicle control and its application to HEV energy management. IEEE Trans Control Syst Technol, 2014, 22: 1018–1031
15 Johannesson L, Asbogard M, Egardt B. Assessing the potential of predictive control for hybrid vehicle powertrains using stochastic dynamic programming. IEEE Trans Intell Transp Syst, 2007, 8: 71–83
16 Donateo T, Pacella D, Laforgia D. A method for the prediction of future driving conditions and for the energy management optimisation of a hybrid electric vehicle. Int J Veh Des, 2012, 58: 111–133

17 Li W, Lee T K, Filipi Z S, et al. Development of electric machine duty cycles for parallel hybrid electric Beijing city bus based on Markov chain. Int J Veh Des, 2012, 58: 348–366
18 Souffran G, Miegeville L, Guerin P. Simulation of real-world vehicle missions using a stochastic markov model for optimal powertrain sizing. IEEE Trans Veh Tech, 2012, 61: 3454–3465
19 Bin Y, Zhong Z Y, Baozhen Y. Bus arrival time prediction using support vector machines. J Intell Transp Syst, 2006, 10: 151–158
20 Bender F A, Kaszynski M, Sawodny O. Drive cycle prediction and energy management optimization for hybrid hydraulic vehicles. IEEE Trans Veh Tech, 2013, 62: 3581–3592
21 Bashash S, Moura S J, Forman J C, et al. Plug-in hybrid electric ve hicle charge pattern optimization for energy cost and battery longevity. J Power Sources, 2011, 196: 541–549
22 White D J. Markov Decision Processes. New York: John Wiley & Sons, 1993
23 Haj-Fraj A, Pfeiffer F. Optimal control of gear shift operations in automatic transmissions. J Franklin Inst, 2001, 338: 371–390.
24 Barto A G. Reinforcement Learning: An Introduction. Cambridge: MIT Press, 1998