Luolin XIONG, Yang TANG, Chensheng LIU, Shuai MAO, Ke MENG, Zhaoyang DONG, Feng QIAN, 2023. A home energy management approach using decoupling value and policy in reinforcement learning. *Frontiers of Information Technology & Electronic Engineering*, 24(9):1261-1272. https://doi.org/10.1631/FITEE.2200667

A home energy management approach using decoupling value and policy in reinforcement learning

Key words: Home energy system; Electric vehicle; Reinforcement learning; Generalization

Corresponding author: Yang TANG, Feng QIAN
E-mail: tangtany@gmail.com; fqian@ecust.edu.cn
ORCID: Yang TANG, https://orcid.org/0000-0002-2750-8029; Feng QIAN, https://orcid.org/0000-0003-2781-332X

Motivation

1. Efficient home energy management (HEM) techniques are promising in optimizing electricity cost, comforting residents, and reducing carbon emissions, which are achieved by intelligently scheduling and controlling electric vehicles (EVs) and household appliances.

2. The scalability of model-driven approaches in different scenarios is poor, and re-established models are a prerequisite for excellent performance in new scenarios of HEM.

3. Standard deep reinforcement learning (DRL) algorithms usually use shared networks for the policy and value functions, which therefore limits the estimation accuracy of the value function.

Method

An improved mathematical model is designed to quantify EVs' energy demand more practically and completely, where driver's experience, unexpected events, and traffic conditions are integrated.

$$Z_{\mathrm{SoC}}(t)=\frac{k_1 k_3\left(\mathrm{e}^{-k_2(t-t_{\mathrm{a}})/(t_{\mathrm{d}}-t_{\mathrm{a}})}-1\right)}{\mathrm{e}^{-k_2}-1}$$

k_1, k_2, and k_3 are the shape parameters changing with driver's experience, unexpected events, and traffic conditions, respectively.

$$E_{t+1}^{\mathrm{SoC}}=\begin{cases}E_t^{\mathrm{SoC}}+\eta_{\mathrm{ch}}P_t^{\mathrm{EV}}\Delta t, & \text{if } P_t^{\mathrm{EV}}\geq 0,\\ E_t^{\mathrm{SoC}}+\eta_{\mathrm{dis}}P_t^{\mathrm{EV}}\Delta t, & \text{otherwise,}\end{cases}$$

$$-P_{t,\max}^{\mathrm{EV}}\leq P_t^{\mathrm{EV}}\leq P_{t,\max}^{\mathrm{EV}},$$

$$0\leq E_t^{\mathrm{SoC}}\leq E_{t,\max}^{\mathrm{SoC}},$$

$$\mathrm{SoC}_t=\frac{E_t^{\mathrm{SoC}}}{C}$$

Different EVs' battery capacities C are considered.

Method

A novel RL-based decoupled actor-critic (DA2C) approach is developed to alleviate the overfitting problem and enhance generalization performance by decoupling the policy and value networks.

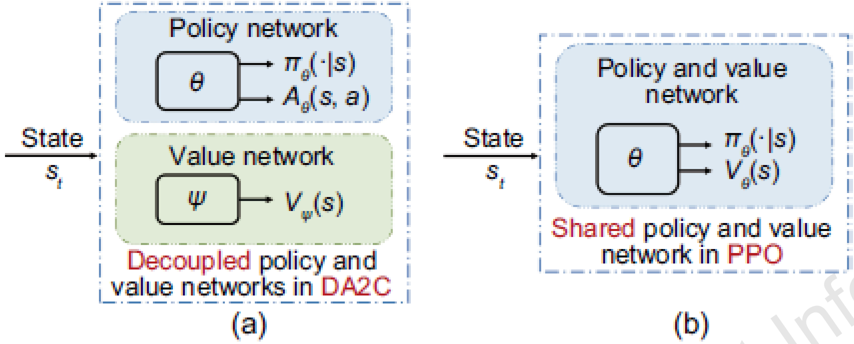


Fig. 2 Comparison of decoupled architecture in DA2C (a) and shared architecture in PPO (b) (DA2C: decoupled advantage actor-critic; PPO: proximal policy optimization)

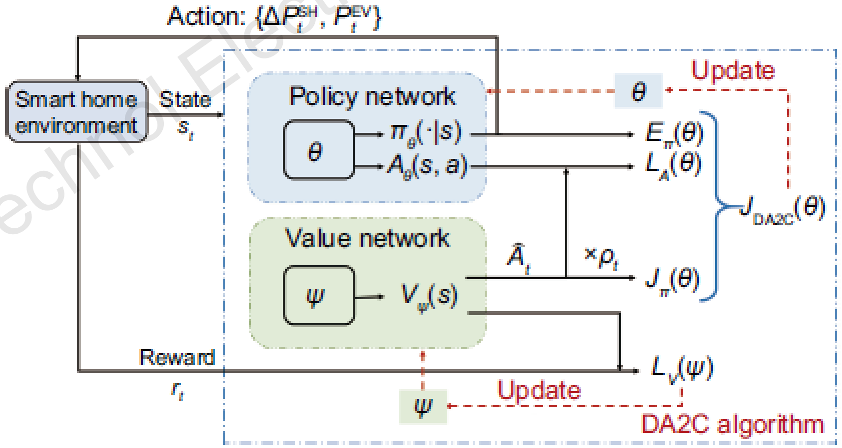


Fig. 3 Decoupled advantage actor-critic (DA2C) algorithm

Method

Table 2 Comparison of different methods and different Markov decision processes (MDPs) in terms of training and test performances on tasks from different residents

Method	Training cost ($\times 10^2$ A\$)			Test cost (A\$)		
	MDP	MDP_1	MDP_2	MDP	MDP_1	MDP_2
PPO	15.45±7.08	**9.31±6.40**	17.98±10.53	59.87±24.65	60.67±34.72	65.56±21.27
IDAAC	14.93±9.35	10.01±7.80	14.12 ± 7.67	62.57±25.38	51.70±23.76	**58.92±22.38**
APDAC	14.58±8.86	13.51±8.32	15.17 ± 6.63	64.13±21.25	43.33±18.01	68.32±23.20
DA2C	**13.31±7.40**	9.90±4.88	**13.90±7.72**	**58.33±25.72**	**42.76±18.80**	60.87±20.64

Results in bold and cells colored gray denote the best and the second best, respectively

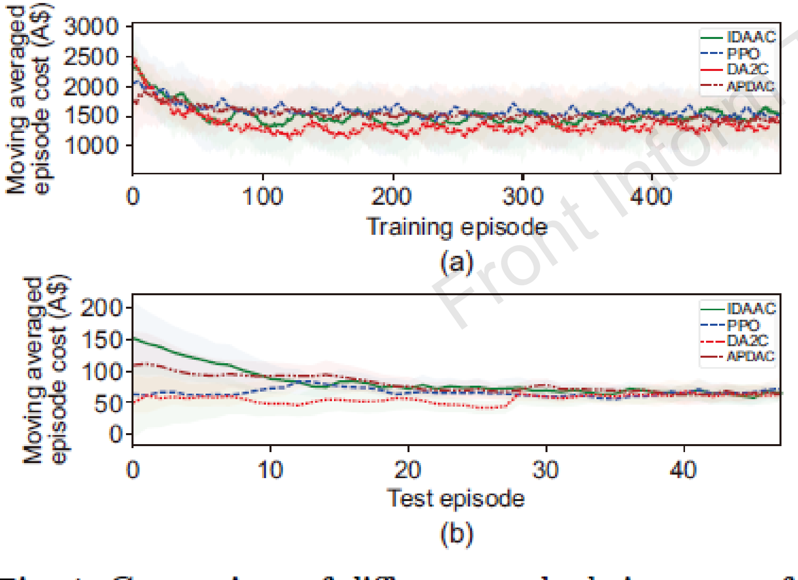


Fig. 4 Comparison of different methods in terms of training (a) and test (b) performances on tasks from different residents

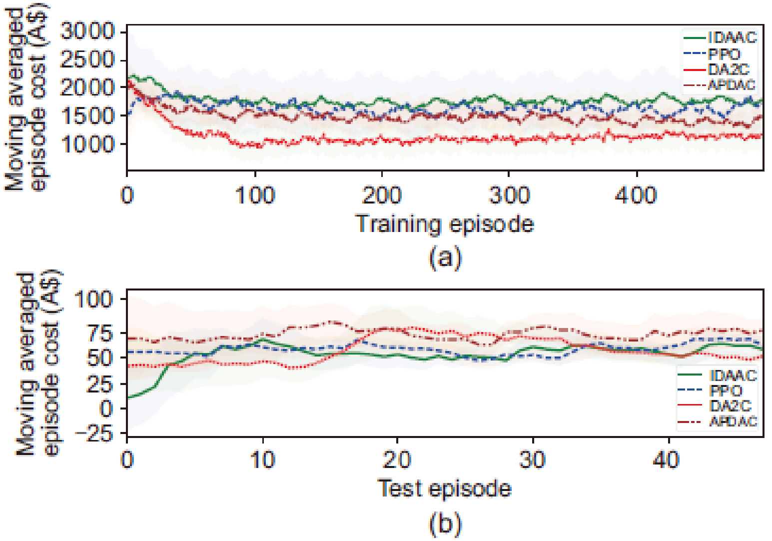
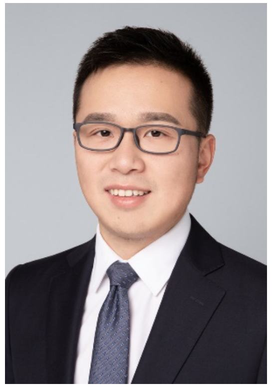


Fig. 6 Comparison of different methods in terms of different seasons: (a) training; (b) test

Conclusions

1. We proposed a data-driven intelligent optimization approach to achieve optimal energy management of smart home systems, by appropriately scheduling the charging operation of EVs and the energy consumption of household appliances.
2. An improved mathematical model is designed to quantify EVs' energy demand more practically and completely, where driver's experience, unexpected events, and traffic conditions are integrated.
3. A novel RL-based DA2C approach is developed to alleviate the overfitting problem and enhance generalization performance by decoupling the policy and value networks.

Luolin XIONG received her BS degree from the School of Control Science and Engineering, East China University of Science and Technology, Shanghai, China, in 2020, where she is currently pursuing her PhD degree. Her research interests include smart grids, microgrids, reinforcement learning, and their applications.



Yang TANG received his BS and PhD degrees in electrical engineering from Donghua University, Shanghai, China, in 2006 and 2010, respectively. From 2008 to 2010, he was a research associate with The Hong Kong Polytechnic University, Hong Kong, China. From 2011 to 2015, he was a post-doctoral researcher with the Humboldt University of Berlin, Berlin, Germany, and with the Potsdam Institute for Climate Impact Research, Potsdam, Germany. He is currently a professor with the East China University of Science and Technology, Shanghai, China. His current research interests include distributed estimation/control/optimization, cyberphysical systems, hybrid dynamical systems, computer vision, reinforcement learning, and their applications. He was a recipient of the Alexander von Humboldt Fellowship and has been the ISI Highly Cited Researchers Award by Clarivate Analytics, since 2017.

Feng QIAN received his BS degree in chemical automation and meters from Nanjing Institute of Chemical Technology, Nanjing, China, in 1982, and his MS and PhD degrees in automation from the East China Institute of Chemical Technology, Shanghai, China, in 1988 and 1995, respectively. He was the director of the Automation Institute, East China University of Science and Technology, from 1999 to 2001, and the head of the Scientific and Technical Department, from 2001 to 2006. He is currently the director of the Key Laboratory of Smart Manufacturing in Energy Chemical Process, Ministry of Education, Shanghai, and the director of the Process System Engineering Research Center, Ministry of Education, Shanghai. His current research interests include modeling, control, optimization, and integration of petrochemical complex industrial processes and their industrial applications, neural network theory, and real-time intelligent control technology. He is a member of the China Instrument and Control Society, the Chinese Association of Higher Education, and the China's PTA Industry Association. He is also an Academician of the Chinese Academy of Engineering.