

Review on Reinforcement Learning-Based Energy Management Strategies for Hybrid Electric Vehicles

Zhengyu Yao and Hwan-Sik Yoon*

The University of Alabama, USA

ISSN: 2640-9690



*Corresponding author: Hwan-Sik Yoon,
The University of Alabama, Box 870276,
Tuscaloosa, AL 35487-0276, USA

Submission: 📅 December 08, 2021

Published: 📅 February 07, 2022

Volume 4 - Issue 1

How to cite this article: Zhengyu Yao and Hwan-Sik Yoon*. Review on Reinforcement Learning-Based Energy Management Strategies for Hybrid Electric Vehicles. *Evolutions in Mechanical Engineering*. 4(1). EME.000579. 2022.
DOI: [10.31031/EME.2022.04.000579](https://doi.org/10.31031/EME.2022.04.000579)

Copyright@ Hwan-Sik Yoon, This article is distributed under the terms of the Creative Commons Attribution 4.0 International License, which permits unrestricted use and redistribution provided that the original author and source are credited.

Abstract

Hybrid Electric Vehicles (HEVs) achieve better fuel economy than conventional vehicles by employing two different power sources: a mechanical engine and an electrical motor. These power sources have conventionally been controlled by a rule-based algorithm or optimization-based control. Besides these conventional approaches, reinforcement learning-based control algorithms have actively been studied recently. Reinforcement learning, which is one of three machine learning paradigms, has the capability of determining optimal control actions to maximize a vehicle's fuel economy without the vehicle model nor a priori driving route information. To provide a useful reference to researchers interested in this technology, this article reviews reinforcement learning-based energy management strategies for HEVs with their advantages and disadvantages.

Introduction

Due to increased environmental and economic considerations, energy-efficient vehicles have become the focus of automotive research community. Hybrid Electrical Vehicle (HEVs) is regarded as an alternative solution to further reduce the fuel consumption and emission by employing two different power sources: a mechanical engine and an electrical motor [1]. However, controlling the two power sources requires a more involved method than conventional vehicles to determine how to split the power optimally between the engine and the motor. A systematic approach for this task is known as the Energy Management Strategies (EMS) [2].

EMS for HEVs can be classified into two general categories: a rule-based algorithm and optimization-based control. Rule-based control strategies are robust, simple, and easy to understand [3]. However, a fixed rule-based control algorithm may not be optimal for changing driving conditions. On the other hand, an optimization-based controller utilizes a cost or objective function to calculate an optimal solution with a given set of constraints. Various methods exist for solving the optimization problem including Dynamic Programming (DP) [4], Equivalent Consumption Minimization Strategy (ECMS) [5], Sequential Quadratic Programming (SQP) and Model Predictive Control (MPC) [6,7]. In general, these algorithms can determine the optimal power split between the engine and the motor for a given driving cycle. However, these optimization-based strategies require either a priori knowledge on the drive cycle or high computation power, which prevents their wide adoption in real-time applications.

Reinforcement Learning (RL) [8], which is one of three machine learning paradigms, has the capability of determining optimal control actions to maximize a vehicle's fuel economy without the vehicle model nor a priori driving route information. The foundation of RL is Dynamic Programming (DP) which finds an optimal policy for a Markov Decision Process (MDP) [9]. In RL, an agent learns an optimal policy that returns maximum rewards from a series of actions that the agent takes in each given state. Different approaches exist for finding the optimal policy, such as value-based methods and policy gradient methods. While value-based methods find the corresponding policy by estimating an optimal value function, policy gradient methods compute the optimal policy directly from samples. This article reviews

various RL-based energy management strategies that have been developed for HEVs with detailed discussions on their advantages and disadvantages.

Reinforcement Learning-Based Energy Management Strategies

Value-based approach

A value-based approach, such as Temporal Difference (TD) Learning [10], State Action Reward State Action (SARSA) [11] and Q-Learning [12], estimates the optimal value function and derives the corresponding policy based on the value function. This approach has proven to be very effective for small and low-dimensional control tasks. For HEV powertrain control and energy management, Lin X et al. [13] applied the TD(λ) Learning to a parallel HEV powertrain. Simulation results over real-world and test-driving cycles demonstrated the proposed parallel HEV power management policy can improve fuel economy by 42% compared with a rule-based policy. Xu B et al. [14] implemented a model-free off-policy Q-Learning strategy on a 48V mild HEV simulation model and improved the fuel economy (MPG) by 0.88% compared with the ECMS. With the same algorithm, Rui X et al. [15] obtained the optimal power distribution between the battery and ultracapacitor of a plug-in HEV, thereby significantly decreasing the energy loss by 16.8%. Liu T et al. [16] proposed a Dyna-Q algorithm for a hybrid electric tracked vehicle. Their results showed that, compared with stochastic dynamic programming, the Dyna-Q algorithm has strong adaptability, optimality and learning ability, and can effectively reduce computational time. Shahrzad et al. [17] applied a SARSA learning algorithm to a HEV for battery State of Charge (SOC) stability while achieving the optimal fuel consumption. The simulation results showed that the SARSA algorithm outperforms Q-learning in battery SOC preservation. However, all these methods suffer from the "curse of dimensionality," which makes it difficult to be implemented in a production vehicle.

Deep Reinforcement Learning (DRL), which is comprised of an offline deep neural network construction phase and an online Deep Q-Learning (DQL) phase, showed the capability of handling high-dimensional state and action spaces in an actual decision-making process. Zhao P et al. [18] proposed a HEV power management framework based on a Deep Q Network (DQN) [19] for optimizing the fuel economy. Simulation results based on an actual vehicle setup over real-world and test-driving cycles demonstrated the effectiveness of the proposed framework in optimizing HEV fuel economy. Hu Y et al. [20] presented an online learning architecture for DQN-based EMS, where the online learning approach can learn from different driving conditions. Simulation results demonstrated that the DQN-based EMS can obtain better fuel economy than the rule-based EMS. Wu J et al. [21] proposed a DQN-based EMS for a power split Hybrid Electric Bus (HEB). Fuel economy of the proposed DQL-based strategy was 5.6% better than Q-learning and achieved nearly 90% of that of the DP benchmark over an unseen driving cycle. This study also indicated that the main limitation

of the Q-learning is discretization of the state variables. However, DQL-based control algorithms do not generate continuous actions because the control variables are discretized. Even though DQL has proven to be very effective for handling high-dimensional control tasks, the need for discrete state and action spaces limits its applicability in real-world optimization problems.

Policy gradient approach

Policy gradient approaches directly update the policy without using value estimations, exhibiting better performance in terms of solving for deterministic policies and problems in continuous state space. Li Y et al. [22] presented a power management strategy for a plug-in HEV based on an Actor-Critic (AC) method [23]. The simulation results demonstrated that the performance of a well-trained AC-based power management system can be close to that of a DP-based method, while saving considerable amount of computation time than DP. With the same algorithm, Tan H et al. [24] proposed a self-learning EMS for a plug-in hybrid electric bus. Experimental results showed that AC outperforms DP with an acceptable range of discretization precision, lower energy cost and less computation time. This study also illustrated that the reason for the better performance of AC than DP is that AC searches the optimal strategy in continuous state and action spaces, thus can fundamentally avoid the discretization error and the curse of dimensionality. Shota et al. [25] applied a new policy-based approach named Proximal Policy Optimization (PPO) to a real-time HEV energy management problem. They tried to improve the vehicle performance by predicting the future behavior of the vehicle in an informed traffic situation. They showed that the engine torque as a continuous value and the gear number as a discrete value can be learned by using the PPO algorithm together with V2V and V2I information.

As advanced RL algorithms, Deep Deterministic Policy Gradients (DDPG) [26] have been applied to EMS of HEVs. As a combination of Deterministic Policy Gradient (DPG) and DQN, DDPG is one of the most powerful model-free off-policy RL algorithms. DQN stabilizes the learning of the Q-function by experience replay and the frozen target network. Whereas the original DQN works in a discrete space, DDPG extends it to a continuous space with the actor-critic framework learning a deterministic policy. Wu Y et al. [27] addressed the energy management problem of a Plug-in HEB using the DDPG algorithm. The simulation results over driving cycles showed that the proposed approach outperforms One Q-learning approach and exhibits performance close to that of DP. Roman et al. [28] implemented the DDPG on a mild HEV and achieved a nearly optimal fuel consumption result with a locally trained strategy. They also used a stochastic driver model for improved state generalization and prevented the strategy from overfitting. Ma Z et al. [29] applied DDPG with a time-varying weighting factor to further improve economic performance of Hybrid Electric Tracked Vehicle and reduce computational burden. The results showed that DDPG-based EMS with an online updating mechanism can achieve nearly 90% fuel economy of that of DP while computation time is

greatly reduced. Moreover, a hardware-in-loop experiment result proved that the proposed algorithm can be applied in real-time. Tang X et al. [30] developed a novel Double Deep Reinforcement Learning (DDRL)-based EMS, which uses DQN to learn the gear-shifting strategy and DDPG to control the engine throttle. After offline training, an online simulation test of the proposed DDRL-based EMS showed an improvement of 2.33% in fuel efficiency over the Deterministic Dynamic Programming (DDP)-based EMS by overcoming some inherent drawbacks of DDP.

Recently, Yao Z et al. [31] applied a relatively new approach called Twin-Delayed Deep Deterministic Policy Gradient (TD3) [32] to maximize the fuel economy of a mild HEV. As an extension of the DDPG algorithm, TD3 can prevent the overestimation of the value function, and further improve performance. Using the same algorithm, Zhou J et al. [33] embedded a heuristic rule-based local controller in the DRL loop to eliminate irrational exploration. The results showed that an improved TD3 based EMS can produce the best fuel efficiency, fastest convergence speed and highest robustness in comparison to typical value-based and policy-based DRL EMSs over various driving cycles. However, they all suffer from a significant drawback, which is a very slow and resource intensive training process.

Model-free DRL relies on a large number of real samples from the environment, which is often hindered by low sampling efficiency [34]. In many cases, human experiences can provide optimal training samples or preferences in guiding the learning agent in exploration during the training process [35,36]. Lian R et al. [37] applied a rule-interposing DRL-based EMS to a Prius model. By embedding expert knowledge such as an optimal curve in the engine Brake-Specific Fuel Consumption (BSFC) map into the DRL-based EMS, the engine can work along the optimal BSFC curve. Li Y [38] proposed a DDPG-based EMS for a series HEV enabled by historically cumulative trip information. Simulation results showed that without a priori knowledge of the future trip, the original DDPG-based EMS achieved an average improvement of 3.5% over a benchmark. After further applying output frequency adjustment, an average improvement of 8.7% was obtained, which is comparable to an MPC-based EMS.

Conclusion

This article reviewed various RL-based energy management strategies for HEVs. It has been shown that RL-based EMS can effectively control HEVs without requiring the vehicle model nor a priori driving route information. Among various RL-based EMS algorithms, policy gradient-based EMS show better performance in solving high-dimensional problems in continuous state space. Since RL-based EMS show great potential in many aspects compared to the conventional methods, it is expected that RL-based EMS will continue to be studied in future especially for the control of battery electric vehicles.

References

1. Enang W, Bannister C (2017) Modelling and control of hybrid electric vehicles (a comprehensive review). *Renewable and Sustainable Energy Reviews* 74: 1210-1239.
2. Becerra G, Alvarez-Icaza L, Pantoja Vázquez A (2016) Power flow control strategies in parallel hybrid electric vehicles. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering* 230(14):1925-1941.
3. Jalil N, Kheir NA, Salman M (1997) A rule-based energy management strategy for a series hybrid vehicle. *Proceedings of the 1997 American Control Conference, Cat. No. 97CH36041, 1997, IEEE, Albuquerque, New Mexico, USA.*
4. Bellman R (2013) *Dynamic Programming* (Courier Corporation, 2013).
5. Onori S, Serrao L, Rizzoni G (2016) *Hybrid electric vehicle energy management systems*. Springer, New York, USA, p.112.
6. Boggs Paul T, Jon Tolle W (2008) Sequential quadratic programming. *Acta numerica* 4(1995): 1-51.
7. Camacho Eduardo F, Carlos Bordons A (2013) *Model predictive control*. Springer science & business media, Berlin/Heidelberg, Germany, pp.13-30.
8. Sutton RS, Barto AG, Bach F (1998) *Reinforcement learning: An introduction*, MIT Press, USA.
9. Puterman ML (1994) *Markov Decision processes: Discrete stochastic dynamic programming*. (1st edn), John Wiley & Sons, Inc. New York, USA.
10. Tesauro Gerald (1995) Temporal difference learning and TD-Gammon. *Communications of the ACM* 38(3): 58-68.
11. Watkins CJCH (1989) *Learning from delayed rewards*. King's College, Cambridge, Oxford, England, UK.
12. Rummery GA, Niranjan M (1994) On-Line Q-Learning using connectionist systems.
13. Lin Xue, Yanzhi Wang, Paul Bogdan, Naehyuck Chang, Massoud Pedram (2014) Reinforcement learning based power management for hybrid electric vehicles. *IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, IEEE, San Jose, CA, USA.
14. Xu B, Malmir F, Rathod D, Filipi Z (2019) Real-Time reinforcement learning optimized energy management for a 48V mild hybrid electric vehicle. *SAE Technical Paper* pp. 01-09.
15. Xiong Rui, Jiayi Cao, Quanqing Yu (2018) Reinforcement learning-based real-time power management for hybrid energy storage system in the plug-in hybrid electric vehicle. *Applied energy* 211: 538-548.
16. Liu Teng, Yuan Zou, Dexing Liu, Fengchun Sun (2015) Reinforcement learning of adaptive energy management with transition probability for a hybrid electric tracked vehicle. *IEEE Transactions on Industrial Electronics* 62(12): 7837-7846.
17. Kouche Biyouki, Seyed Mohammad ANJ, Amin Noori, Fatemeh J (2018) Power management strategy of hybrid vehicles using sarsa method. *Electrical Engineering (ICEE)*, Iranian Conference on IEEE, Mashhad, Iran.
18. Zhao Pu, Yanzhi Wang, Naehyuck Chang, Qi Zhu, Xue Lin, et al. (2018) A deep reinforcement learning framework for optimizing fuel economy of hybrid electric vehicles. *23rd Asia and South Pacific design automation conference (ASP-DAC)*. IEEE, Jeju, Korea.
19. Mnih V, Kavukcuoglu K, Silver D, Rusu A, Veness J, et al. (2015) Human-level control through deep reinforcement learning. *Nature* 518: 529-533.

20. Hu Yue, Weimin Li, Kun Xu, Taimoor Zahid, Feiyan Qin, et al. (2018) Energy management strategy for a hybrid electric vehicle based on deep reinforcement learning. *Applied Sciences* 8(2): 187.
21. Wu Jingda, Hongwen He, Jiankun Peng, Yuecheng Li, Zhanjiang Li (2018) Continuous reinforcement learning of energy management with deep Q network for a power split hybrid electric bus. *Applied energy* 222: 799-811.
22. Li Yuecheng, Hongwen He, Jiankun Peng, Hailong Zhang (2017) Power management for a plug-in hybrid electric vehicle based on reinforcement learning with continuous state and action spaces. *Energy Procedia* 142: 2270-2275.
23. Konda Vijay, John Tsitsiklis (2019) Actor-critic algorithms. *Advances in neural information processing systems* 12 (NIPS 1999).
24. Tan Huachun, Hailong Zhang, Jiankun Peng, Zhuxi Jiang, Yuankai Wu (2019) Energy management of hybrid electric bus based on deep reinforcement learning in continuous state and action space. *Energy Conversion and Management* 195: 548-560.
25. Inuzuka Shota, Fuguo Xu, Bo Zhang, Tielong Shen (2019) Reinforcement learning based on energy management strategy for HEVs. 2019 IEEE Vehicle Power and Propulsion Conference (VPPC). IEEE, Hanoi, Vietnam, Asia.
26. Lillicrap TP, Hunt JJ, Pritzel A, Heess N, Erez T, et al. (2016) Continuous control with deep reinforcement learning, *International Conference on Learning Representations*.
27. Wu Yuankai, Huachun Tan, Jiankun Peng, Hailong Zhang, Hongwen He (2019) Deep reinforcement learning of energy management with continuous control strategy and traffic information for a series-parallel plug-in hybrid electric bus. *Applied energy* 247: 454-466.
28. Liessner Roman, Christian Schroer, Ansgar Dietermann, Bernard Baker (2018) Deep reinforcement learning for advanced energy management of hybrid electric vehicles. *ICAART* 2:61-72.
29. Ma Zhikai, Qian Huo, Tao Zhang, Jianjun Hao, Wei Wang (2021) Deep deterministic policy gradient based energy management strategy for hybrid electric tracked vehicle with online updating mechanism. *IEEE Access* 9: 7280-7292.
30. Tang Xiaolin, Jiaxin Chen, Huayan Pu, Teng Liu, Amir Khajepour (2021) Double deep reinforcement learning-based energy management for a parallel hybrid electric vehicle with engine start-stop strategy. *IEEE Transactions on Transportation Electrification*.
31. Yao Z, Yoon H (2022) Hybrid electric vehicle powertrain control based on reinforcement learning. *SAE Int J Elec Veh* 11(2): 2022.
32. Fujimoto Scott, Herke van Hoof, David Meger (2018) Addressing function approximation error in actor-critic methods *Proceedings of the 35th International Conference on Machine Learning, Stockholm, Sweden, PMLR* 80.
33. Zhou Jianhao, Siwu Xue, Yuan Xue, Yuhui Liao, JunLiu, et al. (2021) A novel energy management strategy of hybrid electric vehicle via an improved TD3 deep reinforcement learning. *Energy* 224: 120118.
34. Schulman John, Sergey Levine, Philipp Moritz, Michael Jordan, Pieter Abbeel (2015) Trust region policy optimization. *International conference on machine learning. PMLR* 37.
35. Brys Tim, Anna Harutyunyan, Halit Bener Suay, Sonia Chernova, Matthew E, et al. (2015) Reinforcement learning from demonstration through shaping. *Twenty-fourth International Joint Conference on Artificial Intelligence*.
36. Christiano Paul, Jan Leike, Tom B. Brown, Miljan Martic, Shane Legg, et al. (2017) Deep reinforcement learning from human preferences.
37. Lian Renzong, Jiankun Peng, Yuankai Wu, Huachun Tan, Hailong Zhang, et al. (2020) Rule-interposing deep reinforcement learning based energy management strategy for power-split hybrid electric vehicle. *Energy* 197 (2020): 117297.
38. Li Yuecheng, Hongwen He, Jiankun Peng, Hong Wang (2019) Deep reinforcement learning-based energy management for a series hybrid electric vehicle enabled by history cumulative trip information. *IEEE Transactions on Vehicular Technology* 68(8): 7416-7430.

For possible submissions Click below:

[Submit Article](#)