

<sup>1</sup>DR MADHUSUDHANA REDDY *et. al.*, / International Journal of Engineering & Science Research

## **A DEEP REINFORCEMENT LEARNING APPROACH TO ENERGY MANAGEMENT CONTROL WITH CONNECTED INFORMATION FOR HYBRID ELECTRIC VEHICLES**

<sup>1</sup>DR MADHUSUDHANA REDDY,<sup>2</sup>J. MARSHITHA,<sup>3</sup>LINGALA BHAGYALAXMI,<sup>4</sup>BANOTH SINDHU,<sup>5</sup>VEENA RADHARAM

<sup>1</sup>ASSOCIATE PROFESSOR, DEPARTMENT OF EEE, MALLA REDDY ENGINEERING COLLEGE FOR WOMEN,  
HYDERABAD

<sup>2,3,4&5</sup>UGSCHOLAR, DEPARTMENT OF IOT, MALLA REDDY ENGINEERING COLLEGE FOR WOMEN, HYDERABAD

### **ABSTRACT**

Considering the importance of the energy management strategy for hybrid electric vehicles, this paper is aiming at addressing the energy optimization control issue using reinforcement learning algorithms. Firstly, this paper establishes a hybrid electric vehicle power system model. Secondly, a hierarchical energy optimization control architecture based on networked information is designed, and a traffic signal timing model is used for vehicle target speed range planning in the upper system. More specifically, the optimal vehicle speed is optimized by a model predictive control algorithm. Thirdly, a mathematical model of vehicle speed variation in connected and unconnected states is established to analyze the effect of vehicle speed planning on fuel economy. Finally, three learning-based energy optimization control strategies, namely q-learning, deep q network (dqN), and deep deterministic policy gradient (ddpg) algorithms, are designed under the hierarchical energy optimization control architecture. It is shown that the q-learning algorithm is able to optimize energy control; however, the agent will meet the “dimension disaster” once it faces a high-dimensional state space issue. Then, a dqN control strategy is introduced to address the problem. Due to the limitation of the discrete output of dqN, the dpdg algorithm is put forward to achieve continuous action control. In the simulation, the superiority of the dpdg algorithm over q-learning and dqN algorithms in hybrid electric vehicles is illustrated in terms of its robustness and faster convergence for better energy management purposes.

### **INTRODUCTION**

Under growing demand for energy and stricter emission standards, developing new energy vehicles is considered a primary strategic measure to ease the global energy crisis and environmental pollution problems (ding and li, 2021, dai et al., 2021, mei et al., 2022b).

Hybrid electric vehicles (hevs), electric vehicles (evs), and fuel cell vehicles are the three primary categories of new energy vehicles (hu et al., 2019, mei et al., 2022a). The benefits of electric vehicles include no emissions, high power efficiency, and low energy use. Still, the current battery technology is not yet developed; achieving an efficient driving range is difficult. Fuel cell vehicles are highly efficient, have low emissions, and can travel long distances without refueling. However, few cases can prove sufficient stability and reliability of fuel cell vehicles. The hev is better than the other two electric vehicles because it can go further and has more flexible working modes (jauch et al., 2018). It is thought that this is the best way to solve environmental and energy problems (delprat et al., 2017, huang et al., 2022, karimi and lu, 2021). As far as hev are concerned, energy management control strategies are a hot topic. On the one hand, the three main energy management strategies for hevs are rule-based, optimization-based, and learning-based (lian et al., 2020). The quick advance of machine learning, in particular deep learning and reinforcement learning (rl), has led to the emergence of alphago and alphago zero in various fields (arulkumaran et al., 2017a, lopez-garcia et al., 2020). As a result, the energy management strategy for hevs was used as a model for the learning-based energy management strategy. The research showed that du et al. (2020a), in terms of fuel cost and calculation speed, the dyna and q-learning algorithms had comparable performance. So that the hybrid power system can be controlled in real-time, lin et al. (2021) combined an online recursive algorithm and q-learning to update the control strategy in real-time. Hu et al. (2018) applies deep rl to the energy management control strategy. The simulation results show its advantages by comparing the rl to the rule-based strategy. Different energy management strategies based on rl and occurring in real-time were made by xiong et al. (2018). The control action can be updated in real-time by integrating the q-learning algorithm into the value function. At present, most reinforcement learning control strategies applied to hev are using the q-learning optimization algorithms. Still, when the q-learning algorithm deals with high-dimensional state space problems, it will cause the “curse of dimensionality” problem in the training process of the agent. The operation of hev is a continuous process, which is more suitable for control strategies such as dqn and ddpq. Although the above literature can optimize the vehicle’s energy management strategy to some extent, the energy management needs to optimize the vehicle controller and consider the impact of traffic flow from a macro perspective. Therefore, designing a hierarchical controller to optimize traffic and vehicle control is necessary, partly motivating us for the current study. On the other hand, the performance of energy optimization relies on predicting the vehicle state. In contrast, the vehicle speed is affected by many factors, such as traffic

conditions, road type, weather, and driving style (lois et al., 2019). It is, therefore, necessary to incorporate future driving information into the energy management system to improve fuel economy. In addition, recent studies have shown that sensors and information technology are critical for fuel economy improvement as the constant changing of driving and road conditions makes it difficult to predict (wu et al., 2019). Thus, combining vehicle-specific signals with eco-driving to optimize the energy management system will receive continuous attention in the future. It is shown that predictive information to design energy management systems can help improve vehicle fuel economy and enhance real-time practicability. Energy optimization can be divided into control strategies based on path preview, ecological driving, and real-time vehicle speed prediction based on the obtained information. Tang et al. (2021) pointed out that using the traffic preview information to optimize the hev's power distribution and speed can further improve fuel economy. In order to improve energy efficiency, du et al. (2022) achieved sufficient speed tracking through previewing road gradient information and speed allocation within a short field of view, thereby achieving potential energy savings. Xie et al. (2019) considered the influence of road gradient on battery charge and discharge. They put forward an hev energy management approach utilizing a stochastic model predictive controller under the premise of road gradient preview. The result demonstrates that the battery soc can be kept within a specific range, resulting in improved fuel efficiency. Nie and farzaneh (2022) used a neural network model to predict the following vehicle's velocity, considering the vehicle-to-vehicle and vehicle-to-infrastructure communication. Moreover, the authors applied the predicted speed to the equivalent consumption minimization strategy to achieve power splitting and improve fuel economy. Although many scholars have researched speed planning, only some link speed planning with the energy management strategy for more detailed analysis. This article has the following innovations: (i) a hierarchical energy optimization control architecture based on network information is designed. Aiming at the optimal target speed planning problem of hev in the upper system based on the network information obtained by v2i and v2v communication, the traffic signal timing model is used to plan the vehicle target speed range. Then the vehicle speed is optimized through the mpc algorithm. (ii) in the lower layer system of the hierarchical energy optimization control architecture, various energy optimization control strategies for hybrid electric vehicles based on reinforcement learning are designed to improve the vehicle's fuel

economy. Regarding the q-learning algorithm dealing with high-dimensional state space problems will cause the agent to have a “dimension disaster” problem during training. Also, considering dqn can only output discrete actions, an energy-optimized control strategy based on ddpq deep deterministic policy gradient algorithm is designed to realize continuous action control.

### **Hybrid power system modeling**

The powertrain architecture of hevs, shown in fig. 1, includes an engine, motor 1 (mg1), motor 2 (mg2), power battery, and planetary gear mechanism. The specific powertrain configuration is shown in table 1. The engine is inactive when the hev starts, and mg2 is the power source. If the hev cruise under low load, the engine torque is split into two parts via planetary gears. Part of the torque is used to drive the hev; the other part is utilized to charge the battery via mg1. The motor and

### **Speed planning control**

Based on the hierarchical optimization control architecture of network information, this paper focuses on the energy optimization control problem of single-hev in the road environment. Fig. 4 illustrates that the upper control system can use v2v and v2i communication to obtain information such as the vehicle’s position, speed, and traffic light status in the road environment. The signal phase and timing (spat) model can calculate an hev passes in a blue light state (huang et al., 2021). The

### **Deep reinforcement learning**

In the networked environment, the upper control system of hev uses the network information obtained by v2i and v2v communication to plan the optimal target speed. The primary object of the lower control system is to design a reasonable and efficient energy optimization controller for the hev.

According to previous research, reinforcement learning is suitable for solving the following problems (du et al., 2020a, liu et al., 2017, zou et al., 2016): (1) the reinforcement learning agent will take

### **Optimal control strategy verification analysis**

The HEV model can be discretized into a dynamic system, given by (Lois et al., 2019):  
 $x(t+1) = f(x(t), u(t))$ ,  $t = 0, 1, \dots, N-1$  where  $x(t)$  denotes the system state;  
 $u(t)$  refers to the control variable;  $f(x(t), u(t))$  is the state transition equation;  $t$   
means the current discrete time;  $N$  is the duration of control.

As the HEV travel speed can be discretized, the state space  $X = \{x_1, x_2, \dots, x_N\}$  and  
the control space  $U = \{u_1, u_2, \dots, u_{N2}\}$  are also limited.  $N1$  and  $N2$  denote the state  
and control amounts, respectively. The key ...

## Conclusion

This paper designed a hierarchical optimization control architecture based on network information. The upper control system proposes network information such as traffic light status information, the distance of the driving section, road speed limit, and other network information based on V2I and V2V communication. The Model Predictive Control algorithm calculates the optimal speed of the vehicle. The lower layer system uses the reinforcement learning algorithm to study the energy optimization. This paper has proposed a novel design approach to obtain an autonomous longitudinal vehicle controller. To achieve this condition, a vehicle architecture with its ACC subsystem has been presented. With this architecture, we have also described the specific requirements for an efficient autonomous vehicle control policy through RL and the simulator in which the learning engine is embedded. A policy-gradient algorithm estimation has been introduced and has used a backpropagation neural network for achieving the longitudinal control. Then, experimental results, through simulation, have shown that this design approach can result in efficient behavior for CACC. Much work can still be done to improve the vehicle controller proposed in this paper. First, it is clear that some modifications to the learning process should be made to improve the resulting vehicle-following behavior. Issues related to the oscillatory behavior of our vehicle control policy can be addressed by using continuous actions. This case would require further study to efficiently implement this approach, because it brings additional complexity to the learning process. Once the oscillatory behavior of the RL approach has been addressed, it would be profitable to compare it with a control obtained by the traditional proportional–integral–derivative (PID) controllers. Some elements with regard to our simulation framework can also be improved, with the ultimate goal of having an even more realistic environment through which we can make our learning experiments. In fact, an important aspect to consider would be to integrate a more accurate simulator for sensory and communication systems. This way, we can eliminate some of our current assumptions, e.g.,

the absence of sensor and communication noise. This condition would make the learning process more complex, but the resulting environment would be much closer to real-life conditions. Our controller can also be completed by an autonomous lateral control system. Again, this approach can be done using RL, and a potential solution is to use a reward function in the form of a potential function over the width of a lane, similar to the current force feedback given by the existing lane-keeping assistance system. This reward function will surely direct the driving agent toward learning an adequate lane-change policy. The lateral control system may be completed by a situation assessment for automatic lane-change maneuvers, as proposed by Schubert et al. Finally, the integration of an intelligent vehicle coordination system for collaborative decision making can transform our system into a complete DAS.

#### REFERENCES:

- [1] J. Piao and M. McDonald, "Advanced driver assistance systems from autonomous to cooperative approach," *Transp. Rev.*, vol. 28, no. 5, pp. 659–684, Sep. 2008.
- [2] Standard Specification for Telecommunications and Information Exchange Between Roadside and Vehicle Systems—5-GHz Band Dedicated Short-Range Communications (DSRC) Medium Access Control (MAC) and Physical Layer (PHY) Specifications, ASTM E2213-03, 2003.
- [3] J. Peters and S. Schaal, "Reinforcement learning of motor skills with policy gradients," *Neural Netw.*, vol. 21, no. 4, pp. 682–697, May 2008.
- [4] P. Fancher, Z. Bareket, and R. Ervin, "Human-centered design of an ACC with braking and forward-crash-warning system," *Vehicle Syst. Dyn.*, vol. 36, no. 2, pp. 203–223, 2001.
- [5] D. Augello, "Description of three PROMETHEUS demonstrators having potential safety effects," in *Proc. 13th Int. Tech. Conf. Exp. Safety Vehicles*, 1993, pp. 4–7.
- [6] C. Bonnet and H. Fritz, "Fuel consumption reduction experienced by two promote chauffeur trucks in electronic tow bar operation," in *Proc. 7th World Congr. Intell. Transp. Syst.*, 2000.
- [7] E. Commission, CVIS: Cooperative Vehicle Infrastructure Systems. [Online]. Available: <http://www.evisproject.org/en/home.htm>
- [8] E. Commission, COOPERS: Cooperative Systems for Intelligent Road Safety. [Online]. Available: <http://www.coopers-ip.eu/>
- [9] S. Tsugawa, "An introduction to Demo 2000: The cooperative driving scenario," *IEEE Intell. Syst.*, vol. 15, no. 4, pp. 78–79, Jul. 2000.
- [10] K. Petty and W. Mahoney, "Enhancing road weather information through vehicle infrastructure integration," *Transp. Res. Rec.*, no. 2015, pp. 132–140, 2007.

- [11] U. Hofmann, A. Rieder, and E. Dickmanns, "Radar and vision data fusion for hybrid adaptive cruise control on highways," *Mach. Vis. Appl.*, vol. 14, no. 1, pp. 42–49, Apr. 2003.
- [12] L. Davis, "Effect of adaptive cruise control systems on traffic flow," *Phys. Rev. E: Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 69, no. 6, p. 066 110, Jun. 2004.
- [13] J. E. Naranjo, C. González, J. Reviejo, R. Garcia, and T. de Pedro, "Adaptive fuzzy control for intervehicle gap keeping," *IEEE Trans. Intell. Transp. Syst.*, vol. 4, no. 2, pp. 132–142, Sep. 2003.
- [14] J. E. Naranjo, C. González, T. de Pedro, R. Garcia, J. Alonso, M. A. Sotelo, and D. Fernandez, "Autopia architecture for automatic driving and maneuvering," in *Proc. IEEE ITSC*, Toronto, ON, Canada, Sep. 2006, pp. 1220–1225.
- [15] J. E. Naranjo, C. González, R. Garcia, and T. de Pedro, "ACC +stop&go maneuvers with throttle and brake fuzzy control," *IEEE Trans. Intell. Transp. Syst.*, vol. 7, no. 2, pp. 213–225, Jun. 2006.
- [16] L. Xiao and F. Gao, "A comprehensive review of the development of adaptive cruise control systems," *Vehicle Syst. Dyn.*, vol. 48, no. 10, pp. 1167–1192, 2010.
- [17] D. De Bruin, J. Kroon, R. Van Klaveren, and M. Nelisse, "Design and test of a cooperative adaptive cruise control system," in *Proc. IEEE Intell. Vehicles Symp.*, 2004, pp. 392–396.
- [18] G. Naus, R. Vugts, J. Ploeg, M. Van de Molengraft, and M. Steinbuch, "Towards on-the-road implementation of cooperative adaptive cruise control," in *Proc. 16th World Congr. Exhib. Intell. Transp. Syst. Serv.*, Stockholm, Sweden, 2009.
- [19] M. van Eenennaam, W. Wolterink, G. Karagiannis, and G. Heijenk, "Exploring the solution space of beaconing in VANETs," in *Proc. IEEE VNC*, 2010, pp. 1–8.
- [20] B. Van Arem, C. Van Driel, and R. Visser, "The impact of cooperative adaptive cruise control on traffic-flow characteristics," *IEEE Trans. Intell. Transp. Syst.*, vol. 7, no. 4, pp. 429–436, Dec. 2006.