Driving Profile Optimization Using a Deep Q-Network to Enhance Electric Vehicle Battery Life

Jihoon Kwon, Manho Kim, Hyeongjun Kim, Minwoo Lee, and Suk Lee

1School of Mechanical Engineering, Pusan National University, Busan 46241, Republic of Korea
2Department of Electric Vehicle, Dong-Eui Institute of Technology, Busan 47230, Republic of Korea
3Department of Future Automotive Engineering, Gyeongsang National University, Jinju 52828, Republic of Korea
4Research and Development Team, Ecoenergy Research Institute Company, Busan 46703, Republic of Korea

Correspondence should be addressed to Suk Lee; slee@pnu.edu

Received 28 March 2023; Revised 18 July 2023; Accepted 10 August 2023; Published 6 September 2023

Copyright © 2023 Jihoon Kwon et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In the COVID-19 era, automobiles with internal combustion engines are being replaced by eco-friendly vehicles. The demand for battery electric vehicles (BEVs) has increased explosively. Treatment of spent batteries has received much attention. Battery life can be extended via both efficient charging and driving. Consideration of the vehicles ahead when driving a BEV effectively prolongs battery life. Several studies have presented eco-friendly driving profiles for BEVs, the cited authors did not develop a BEV driving profile that considered battery life using reinforcement learning. Here, this paper presents a method of driving profile optimization that increases BEV battery life. This paper does not address how to regenerate spent batteries in an eco-friendly manner. The BEV driving profile is optimized employing a deep Q-network (a reinforcement learning method). This paper uses simulations to evaluate the effect of the driving profile on BEV battery life; these verified the applicability of our model. Finally, the speed profile optimization method was limited to improve energy efficiency or battery life in rapid speed change sections.

1. Introduction

In the COVID-19 era, brought about by severe climate change and ecosystem destruction, the automobile industry is replacing vehicles with internal combustion engines, with eco-friendly vehicles. Among such vehicles, the numbers of electric vehicles (EVs) have increased explosively [1, 2]. In 2021, the number of registered vehicles in South Korea was 25,010,000, of which 1,200,000 were eco-friendly. Of these, 240,000 were battery electric vehicles (BEVs), thus 1% of all registered vehicles [3–5]. The rapid increase in BEVs that use large batteries creates demands for charging infrastructure and specialized vehicle maintenance, and creates environmental problems (spent batteries, the treatment of which is a major problem). Many such batteries (used in BEVs and hybrid cars that became popular in the 2010s) must be managed in 2020. However, there is as yet no eco-friendly treatment; the batteries are simply stored [6–8].

Battery life is prolonged by efficient charging and efficient driving [9, 10]. The driver is responsible for efficient charging; this cannot be controlled by the vehicle developer. Driving with consideration of the vehicles ahead can be implemented by the vehicle developer; again, this prolongs battery life. Here, this paper presents a method that extends battery life. This paper creates an efficient driving profile that considers vehicles ahead. In particular, the driving profile optimization method can be a practical alternative for car manufacturers to directly respond to environmental constraints in the process of selling cars.

Several studies have presented eco-friendly driving profiles for BEVs. First of all, the following is a summary of the paper focusing on batteries. Piao et al. [11] published a paper that improves battery efficiency through a battery management system using the cell-balancing algorithm. This paper differs in that it focused on improving the efficiency of electrical energy possessed using the battery management system. Ramkumar et al. [12] published a paper that identifies the function of batteries in EVs and argues for the introduction
of battery management systems to improve battery performance and efficiency. This paper also differs in that it checks the function of the battery in an EV and argues for the introduction of a battery management system to efficiently use the battery in hand. Wang et al. [13] optimized driving of hybrid electric vehicle (HEV) queues. Our method differs in that this paper predicts battery life on the basis of driving style; this paper does not consider HEVs. Sun et al. [14] predicted the speed of an HEV using an exponentially varying, stochastic Markov chain, and a neural network-based model. Our paper differs in that this paper does not consider the efficiency aspect of battery life.

Next, the papers written by incorporating technology rather than machine learning are as follows. Krasopoulos et al. [15] developed a multiobjective optimization method for the speed and torque trajectories of a light EV traveling on a predefined route. Our work differs in that this paper, search for driving profiles that optimize battery life on an arbitrary road. Bozorgi et al. [16] proposed a paper that generates a speed profile of an EV using two-choice routing algorithms: reducing driving time through data mining or improving battery energy efficiency. This paper has similarities in that it generates a velocity profile, but it differs from this paper in that it combines data mining technology and considers energy efficiency. Zhang et al. [17] developed a cloud-based, velocity profile optimizer that determined the driving profile and charge status using a genetic algorithm, and presented a dynamic programming targeting plug-in hybrid buses. Our work differs in that this paper estimates battery life using reinforcement learning (a form of artificial intelligence [AI]). Finally, Song et al. [18] used machine-learning methods for HEV energy efficiency management. However, the cited authors did not develop a BEV driving profile that considered battery life using reinforcement learning.

In addition, research on applying reinforcement learning to vehicles has been presented. Terapontomakol et al. [19] proposed a deep Q-network (DQN) method to develop an autonomous vehicle control system to achieve trajectory design and collision avoidance with regard to obstacles on the road in a virtual environment. Mohammed et al. [20] employed deep reinforcement learning to assist unmanned aerial vehicles to find air pollution plumes in an equal-sized grid space. Zheng et al. [21] proposed to model the dynamic scheduling of automatic guided vehicles as a Markov decision process (MDP) with mixed decision rules based on a DQN to generate the optimal policy. Although the cited authors conducted about reinforcement learning, but did not develop a BEV driving profile that considered battery life.

Here, this paper presents a method of driving profile optimization that increases BEV battery life. Among the methods that can increase the lifespan of the battery of a BEV, the most reliable method possible from the developer’s point of view is a method of optimizing the driving profile. At the same time, optimizing the driving profile is contributing in terms of providing an alternative to automakers that are constrained by environmental constraints. The profile employs a DQN (a reinforcement learning method). Presenting the DQN method rather than the existing optimization algorithm as a method of optimizing the driving profile of a BEV expands the applicability of reinforcement learning such as DQN to the automotive field. This paper uses simulation to evaluate the method; these verify its applicability.

The present paper is organized into five sections. In Section 2, machine-learning methods including reinforcement learning are described. In Section 3, the proposed, reinforcement learning-based driving profile model is explained. In Section 4, the environment used for performance evaluation of the model, and the results, are described. Finally, in Section 5, conclusions are presented.

2. Machine Learning

Machine learning can solve problems effectively using data-based experiences generated in a specific field. Thus, a machine-learning method automatically learns the rules from data and makes decisions based on these rules; no human programming is required. AI renders computers intelligent; they learn and infer as do humans. Thus, AI includes machine learning. For example, AI systems for autonomous vehicle driving follow learned rules when driving.

Machine learning is broadly divided into supervised, unsupervised, and reinforcement learning depending on the signals received and the feedback required for learning [22–24]. Supervised learning allows of predictions, estimations, and classifications using training data. Supervised learning features both independent and dependent variables, and (under supervision) generalizes relationships between them. On the other hand, unsupervised learning searches for hidden patterns or rules in observed data. Such learning generalizes a hidden pattern in a large number of data. The only variables are input variables; there are no dependent variables and no need for supervision. There is often no obvious “correct” way to solve a problem and no way to check whether learning is appropriate. Reinforcement learning determines the actions that are optimal under the current conditions (Figure 1) [25, 26]. A reward is given (in an external environment) whenever an agent takes an action; learning proceeds in directions that maximize the reward. A reward may not be given immediately an action is taken. Thus, a credit assignment problem may occur. The reward stays the same even if the difficulty of the current problem suddenly exceeds the difficulties of the previous two problems.

In reinforcement learning, an agent consists of a policy, a value function, and a model [27]. The policy is an action pattern that determines what to do in a given environment. It thus links the environment to an action. The policy may be
3. The Reinforcement Learning-Based EV Driving Profile Model

3.1. Electric Vehicle Model. When a vehicle moves, it experiences resistance in the direction opposite to the direction of travel, including rolling, air, grade, and inertial resistances; all cause energy loss [28]. The rolling resistance is the energy loss attributable to repeated tire rolls (associated with tire deformation and recovery; Equation (1)):

\[ F_R = Cmg. \]  

where the rolling resistance coefficient \( C \) is calculated using Equation (2) [29]:

\[ C = 0.005 + 1/p [0.01 + 0.0095(v/100)^2]. \]  

The air resistance may be drag, lift, or a lateral force. The drag force is the principal cause of energy loss. The drag force is horizontal (in the direction opposite of travel) and is caused by shear stress and pressure generated by the vehicle body because of the viscosity of air (Equation (3)) [28, 29]:

\[ F_d = 1/2 \rho C_d A v^2. \]  

The grade resistance is a force acting in the slope-descending direction, thus the horizontal component of a certain force \( N_d \) expressed using Equation (4):

\[ F_g = mg \times \tan \theta. \]  

The inertial resistance is the force required to increase the vehicle’s speed. All rotating parts in engines and the drive shafts and wheels, and the vehicle per se, experience different rotatory accelerations in the travel direction. The equivalent mass of the rotating parts must thus be considered. The inertial resistance is Equation (5):

\[ F_i = (m + \Delta m)a/g. \]  

The total running resistance of the vehicle is the sum of the rolling, air, grade, and inertial resistance (Equation (6)):

\[ F_T = F_R + F_A + F_G + F_I. \]  

The experiments that derive total running resistance are performed on roads that are not sloped. The total running resistance (except the grade resistance, \( F_G \)) can be expressed as Equation (8):

\[ F_T = F_R + F_A + F_I. \]

\[ F_T = (m + \Delta m)a/g + 1/2 \rho C_d A v^2 + Cmg. \]  

3.2. The Battery Life Model. The equation for the battery life model is that of Meng [30]. The model can be expressed using (Equation (9)) [31, 32]:

\[ Q_{loss} = B \cdot \exp((-E_a)/RT)(A_h)^z, \]

where \( Q_{loss} \) is the percentage of battery life loss, \( B \) is the proportional factor, and \( E_a \) is the activation energy (the minimum energy required for a reaction). \( R \) is the gas constant, \( T \) is the absolute temperature, \( z \) is the power law factor, and \( A_h \) is the total battery capacity.

In earlier studies on battery design, the discharge rates were set to 0.5, 2, 6, and 10 C when estimating variables. However, our battery life model equation considers changing discharge rates. Thus, our model interpolates battery life model equations at various discharge rates (0.1–10 C). The model is expressed by Equation (10) after substituting \( B, E_a, \) and \( z \) into the functions of the discharge rate; the gas constant and absolute temperature are fixed.

\[ Q_{loss} = B(c) \cdot \exp(\frac{(E_a(c))/RT}{A_h})^{z(c)}, \]

where \( c \) is the discharge rate. The three functions of discharge rates are Equations (11)–(13):
Figure 2 shows the battery capacity loss over time at various discharge rates using the interpolated battery model. As this estimates the capacity loss at various discharge rates (C-rates), it reflects different EV driving patterns.

3.3. Reinforcement-Based Driving Profile Model. To optimize the driving profile via reinforcement learning, the problem is viewed as a sequential decision-making problem; a MDP model is appropriate. The MDP model is Equation (14):

\[ S = [v, v_r, d], \]

where \( v \) is the current speed, \( v_r \) is the relative speed of two vehicles, and \( d \) is the distance between the test vehicle and the vehicle ahead.

The action taken by the agent is acceleration, thus a change in vehicle speed. The agent accelerates or decelerates within the physically possible range (0–100 km/hr) with consideration of the current speed and the vehicle specifications. The reward function optimizes the energy efficiency of the vehicle, battery life, and the distance to the vehicle ahead, and is Equation (15):

\[ R_{t+1} = \alpha E(v_t, a_t) + \beta d_t + \gamma Q(v_t, a_t), \]

where \( E(v_t, a_t) \) is the energy efficiency of the vehicle, \( Q(v_t, a_t) \) is the battery life function, \( d_t \) is the distance to the vehicle ahead, and \( \alpha, \beta, \) and \( \gamma \) are weights that sum to 1 (initial value \( \alpha = 0.45, \beta = 0.1, \) and \( \gamma = 0.45 \)). Finally, the state transition probability and discount factor are set to 1 and 0.9, assuming that the MDP environment is deterministic.

4. Evaluation of the Reinforcement-Based EV Driving Profile Model

The selected vehicle is model A EV of Company H; the specifications are listed in Table 1. A features a permanent-magnet synchronous motor that yields a 204 horsepower ("pferdestarke") at 3,600 RPM or higher, and a maximum torque of 395 N.m from 0 to 3,600 RPM. The torque and RPM scales were modified to reflect the performance of the KONA motor using the holding data of the motor efficiency map. The regenerative mode efficiency map was reversed. The KONA-EV battery is of the lithium-ion polymer type, but the battery used in our battery life model was an LiFePO4 battery, because reference data were available.

The vehicle was modeled using Cruise M vehicle simulation software of Company AVL; the power generated during driving was calculated. The simulations considered the loss caused by the total running resistance, thus including the rolling, air, grade, and inertial resistances, and the power recharged by regenerative braking when decelerating. Figure 3 shows the vehicle model simulated using AVL Cruise M software; this paper derived the energies consumed and powers generated. The simulation was conducted every 10 ms and the results were collected over 100 runs. The motor controller and inverter were included in the motor block, and the efficiency of the inverter was set to 92% by referring to the manufacturer’s data.
The method maintained a safe distance from the vehicle ahead. This optimized energy efficiency and battery life. According to the safety distance standards of KoROAD, when the speed limit is 80 km/hr or above and the vehicle speed is \( A \) (km/hr), the safe distance is \( A m \); when the speed limit is 80 km/hr or less and the vehicle speed is \( B \) (km/hr), the safe distance is \( (B-15) \) m. In addition, the time-to-collision (TTC) was set to 1.6 s to ensure flexibility of the safe distance [33]. In terms of these two safe distance standards, the KoROAD standard was followed if the speed was 25 km/hr or more (maximum safety distance); otherwise the TTC was followed (minimum safety distance). Finally, at least a 2 m safe distance was assumed at low speeds and when stopped.

The hardware used for simulation was a desktop computer with an AMD 3600x processor, 32 GB of main memory, and a GeForce GTX 1080 ti graphics processing unit. Some sections of Federal Test Procedure-75 (FTP-75) were used when establishing the driving profile of the vehicle ahead. The test vehicle was assumed to follow this driving profile during reinforcement learning on how to optimize...
energy efficiency and battery life while maintaining the aforementioned safety distances. Simulations were conducted to compare energy consumption efficiencies and battery lives when driving on some sections of FTP-75. The energy consumption efficiency (km/kWh) were calculated by dividing the electricity energy (kWh) consumed by the distance traveled (km). In addition, as battery life does not change rapidly, simulations were conducted over a 1-year cycle.

The performance of Q-learning and DQN (representative reinforcement learning methods) were compared. The exploration and entire episode Q-learning steps were 15,000 and 11,000, respectively, for a 120 s duration driving profile of FTP-75. An optimal value could not be found; values that did not completely converge were frequently generated although similar driving profiles when the episodes exceeded 10,000 steps. In the case of Q-learning, it has been concluded that it takes a lot of exploration and learning to complete the Q-table to meet the number of too many cases. The DQN was trialed using the same driving profile sample. Q-learning revealed similarity when the episode exceeded 10,000 steps whereas the DQN began to evidence similarity after 400 steps after the sample to be learned was gathered in replay memory. Thus, DQN attained an optimized value faster and more accurately than did Q-learning. In other words, when solving a problem with high complexity, DQN’s replay memory method is more effective than adding depth of Q-learning.

In DQN learning, the learning rate was 0.001, the target update frequency was 3, the maximum episode was 11,000 times, the discount factor $r$ was 0.9, the mini-batch size was 256, and the gradient threshold was 1, which was selected through a tuning process. In addition, ReLU was selected as the activation function.

Figure 4 shows the energy efficiency results for Model A of Company H (the test model when DQN reinforcement learning was applied and not). Numbers 1–6 on the x-axis of Figure 4 refer to (arbitrary) Sections 1–6 around 120 s in the FTP-75 profile, and the y-axis is the energy efficiency (km/kWh). All of Cases 1–6 (except Case 2) improved when the DQN was applied (compared to not applied). The least improved case was Case 4 (4.92% in terms of energy efficiency) and the most improved, Case 3 (15.39%). The energy efficiency of the least improved, Case 4, was 12.50 km/kWh when DQN was not applied. However, when DQN was applied, the energy efficiency was 13.11 km/kWh. Furthermore, the energy efficiency of the most improved, Case 3, was 10.71 km/kWh when DQN was not applied. When DQN was applied, the energy efficiency was 12.35 km/kWh. By contrast, Case 2 exhibited a better energy efficiency when DQN was not applied, compared to applied. The energy efficiency was 8.52 km/kWh when DQN was not applied but 8.40 km/kWh when DQN was applied, thus a decrease of 1.48%. The reason for this is that, for Case 2, the speed rapidly increased or decreased; DQN learning did not significantly impact energy efficiency.

Figure 5 shows the battery capacity loss results for Model A of Company H (the test model when DQN reinforcement learning was applied and not). Numbers 1–6 on the x-axis of Figure 5 refer to (arbitrary) Sections 1–6 around 120 s in the FTP-75 profile, and the y-axis is the battery capacity loss (kWh/1,000 km). All of Cases 1–6 (except Case 2) improved when the DQN was applied (compared to not). The least improved case was Case 4 (13.00%) and the most improved, Case 3 (29.14%). The battery capacity loss in the least improved, Case 4, was 0.055 kWh/1,000 km when DQN was not applied, but 0.048 kWh/1,000 km when DQN was applied. The battery capacity loss in the most improved, Case 3, was 0.094 kWh/1,000 km when DQN was not applied but 0.066 kWh/1,000 km when DQN was applied. By contrast, Case 2 evidenced a better energy efficiency when DQN was applied than not. The battery capacity loss was 0.056 kWh/1,000 km when DQN was not applied and 0.059 kWh/1,000 km when DQN was applied, thus a fall of 4.64%. The reason for this is that, for Case 2, the speed rapidly increased or decreased; DQN learning did not significantly impact energy efficiency. In Case 2, if the information about the driving profile is sufficiently learned through long-term driving, the BEV will have improved battery energy consumption efficiency and battery capacity loss. In addition, if the driving profile of the BEV is sufficiently learned, it is judged
that it will have better energy efficiency and battery capacity reduction rate than the result through simulation in other cases.

The method for optimizing the driving profile assuming that there is a vehicle ahead is activated after the driver gets into the vehicle and starts the engine. The driving profile optimization method is deactivated according to the driver’s intention or when the safety distance of 2 m cannot be maintained. The case where the driving profile optimization method cannot be applied is when the vehicle changes lanes, overtakes, or reverses the previous vehicle. When the driving profile optimization method is activated, it can be said to be effective in improving the energy efficiency and battery life of the BEV.

5. Conclusions

This paper presents a method that optimizes the driving profile to increase BEV battery life. The BEV driving profile employed a DQN reinforcement learning method. This paper verified the applicability of the method using simulations. Our conclusions are:

First, the BEV battery life varies by the driving profile. In the simulation results with the proposed optimization method, the battery performance varied from 29.14% to −4.64%. In particular, the method to optimize the driving profile of BEVs is the method to improve battery life from the BEV developer’s viewpoint.

Second, the proposed optimization method of driving profile based on reinforcement learning was effective in improving energy efficiency and battery life. Energy efficiency using the proposed optimization method was improved by 7.99% on average, and battery capacity loss was reduced by 16.84% on average.

However, the method did not improve energy efficiency or battery life when the speed changed rapidly. Such changes negatively impact energy efficiency and BEV battery life. This result verified that rapid speed changes can have negative impacts on the energy efficiency and battery life of BEVs.

Also, we used only some FTP-75 profiles. More profiles should be evaluated to enhance reliability under dynamics and extreme condition. In addition, other reinforcement learning algorithms such as Dueling DQN, Double DQN, and D3QN that feature approximation techniques should be tested. Finally, to apply it to an actual BEV, the driver’s sense of heterogeneity and driving mode change (4WD to 2WD) must be considered.

Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Authors’ Contributions

All the authors contributed significantly to this work.

Acknowledgments

This research was supported by the Ministry of Trade, Industry & Energy (MOTIE), Korea Institute for Advancement of Technology (KIAT) through the Core Technology Development Program for The Industries of Vehicle (N20006869, 2019).

References


