

Reinforcement Learning-Based Optimization of Fast Charging Strategy for Electric Vehicle Batteries to Extend Cycle Life

Yu Gao ^a, Dingyi Yuan ^b, Yichi Wang ^c

China Auto Information Technology (Tianjin) Co., Ltd., Tianjin, China

^agaoyu@catarc.ac.cn, ^byuandingyi@catarc.ac.cn, ^cwangyichi@catarc.ac.cn

Abstract

To address the problems of intensified polarization and excessively rapid cycle life degradation of electric vehicle batteries caused by the traditional constant current-constant voltage (CC-CV) fast charging strategy, an optimized reinforcement learning fast charging strategy based on Deep Q-Network (DQN) is proposed. Firstly, a second-order RC equivalent circuit model and an SEI film growth aging model are constructed to accurately characterize the electrochemical properties and life degradation mechanism of batteries during the charging process. Secondly, a reinforcement learning framework is designed, with the battery state-of-charge (SOC), voltage, temperature, and polarization voltage as the state space, the charging current step as the action space, and a multi-objective reward function considering both fast charging efficiency and life protection is established. Finally, experimental verification is carried out on 18650 ternary lithium batteries, comparing with the traditional CC-CV strategy and PID adaptive strategy. The results show that the proposed strategy controls the fast charging time at 42.8 minutes (SOC from 20% to 80%), which is 17.2% shorter than the CC-CV strategy; the capacity retention rate reaches 86.3% after 1000 cycles, which is 21.5% higher than the CC-CV strategy and 10.8% higher than the PID strategy; the peak polarization voltage is reduced by 34.7%, effectively suppressing the battery aging rate. This method realizes the dynamic balance between fast charging speed and cycle life, providing a new path for the engineering application of electric vehicle battery fast charging technology.

Keywords

Reinforcement Learning; Electric Vehicle Battery; Fast Charging Strategy; Cycle Life; Deep Q-Network (DQN); Polarization Suppression.

1. Introduction

With the rapid development of the new energy vehicle industry, the fast charging performance of power batteries has become a core bottleneck restricting user experience [1]. Although the current mainstream CC-CV fast charging strategy is simple to control, the current drops sharply in the constant voltage phase in the middle and late stages of charging, resulting in low charging efficiency. Moreover, the accumulation of concentration polarization and electrochemical polarization caused by high-current charging will accelerate SEI film growth and lithium dendrite deposition, significantly shortening the battery cycle life [2]. Data shows that ternary lithium batteries adopting the traditional CC-CV strategy have a capacity retention rate of less than 70% after 800 cycles, which is difficult to meet the service life requirement of 5 years/100,000 kilometers for electric vehicles [3].

To solve the above contradictions, scholars at home and abroad have carried out relevant research. Reference [4] designed a multi-stage constant current fast charging strategy based on the battery thermal-electric coupling mechanism, which reduced the polarization degree but

did not consider the dynamic change of battery aging state. Reference [5] adopted PID control to realize adaptive adjustment of charging current, improving the stability of fast charging, but it relied on an accurate mathematical model and had insufficient robustness. Reference [6] applied reinforcement learning to battery charging control and optimized the current curve through the Q-learning algorithm, but traditional Q-learning has the problems of slow convergence speed and poor generalization ability in high-dimensional state space.

Reinforcement learning has the advantages of not requiring an accurate mechanism model and being able to independently optimize strategies through interaction, making it suitable for complex nonlinear systems such as batteries [7]. This paper proposes a reinforcement learning fast charging strategy based on DQN, with the core innovations as follows: (1) Integrate the equivalent circuit model and aging model to construct a reinforcement learning environment more in line with reality; (2) Design a multi-objective reward function to balance fast charging efficiency and life protection; (3) Verify the superiority of the strategy in improving fast charging speed and cycle life through experiments.

2. Battery Model Construction

2.1. Second-Order RC Equivalent Circuit Model

To accurately characterize the voltage response and polarization characteristics of the battery during the charging process, this paper abandons the first-order RC model with insufficient characterization accuracy and the high-order model with redundant structure, and adopts the second-order RC equivalent circuit model which achieves the optimal balance between accuracy and complexity in engineering applications. This model can effectively distinguish the polarization processes with different response speeds inside the battery, providing a reliable state feedback basis for the optimization of fast charging strategies. The model structure is shown in Figure 1. From the perspective of circuit composition, the model reproduces the battery electrochemical behavior through the structure of "ohmic internal resistance + dual polarization branches": the ohmic internal resistance R_0 corresponds to the sum of the internal electrolyte resistance, electrode material resistance, and contact resistance of the battery, which is the core factor causing the instantaneous voltage drop when the charging current changes abruptly; the electrochemical polarization resistance R_1 and polarization capacitor C_1 are connected in series to form the first polarization branch, which is used to describe the polarization effect of the charge transfer process in the electrochemical reaction. This process has a fast response speed, and the time constant is usually on the order of seconds; the concentration polarization resistance R_2 and polarization capacitor C_2 constitute the second polarization branch, corresponding to the concentration gradient polarization formed by the diffusion of lithium ions inside the electrode material. Its response speed is slow, and the time constant can reach the order of minutes. Through the synergistic effect of the dual polarization branches, the model can fully capture the dynamic change law of polarization voltage during the charging process. Its state equations are as follows:

$$\begin{cases} \dot{U}_{C1} = -\frac{1}{R_1 C_1} U_{C1} + \frac{1}{C_1} I \\ \dot{U}_{C2} = -\frac{1}{R_2 C_2} U_{C2} + \frac{1}{C_2} I \\ U_{OCV} = f(SOC, T) \\ U = U_{OCV} - U_{C1} - U_{C2} - IR_0 \end{cases}$$

Where, I is the charging current (positive for charging), U_{OCV} is the open-circuit voltage (related to SOC and temperature T), and U is the battery terminal voltage. The model parameters are identified through HPPC experiments, as shown in Table 1.

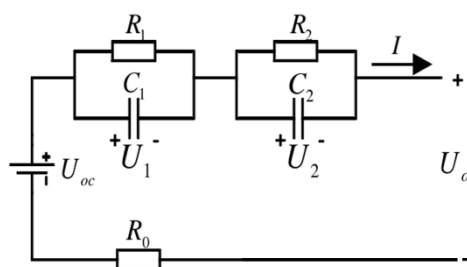


Figure 1. Second-order RC equivalent circuit model

Table 1. 18650 Ternary Lithium Battery Model Parameters (SOC=50%, T=25°C)

| Parameter | Value | Unit |
|---|-------|----------|
| Ohmic internal resistance R_0 | 0.018 | Ω |
| Electrochemical polarization resistance R_1 | 0.025 | Ω |
| Concentration polarization resistance R_2 | 0.032 | Ω |
| Polarization capacitor C_1 | 800 | F |
| Polarization capacitor C_2 | 2000 | F |

2.2. Battery Aging Model

From the perspective of modeling logic, battery aging models are mainly divided into three categories: the first is mechanism-based models, which reveal the essence of aging based on electrochemical principles. By analyzing the core aging mechanisms such as active material loss, electrolyte decomposition, SEI film growth, and lithium dendrite deposition during the lithium ion intercalation/deintercalation process, multi-physics field coupled mathematical equations are established (such as the lithium ion diffusion model based on Fick's law and the internal resistance evolution equation considering SEI film growth kinetics). Typical representatives include the PNGV model, bipolar electrode model, and pseudo-two-dimensional (P2D) model based on porous electrode theory. Such models have clear physical significance and strong interpretability, but they rely on accurate electrochemical parameters (such as diffusion coefficient and exchange current density) and have high computational complexity for multi-field coupling. The second is empirical/semi-empirical models, which obtain data through a large number of accelerated aging tests (such as different temperatures, charge-discharge rates, and SOC cycle conditions) and fit the statistical relationship between performance degradation and influencing factors. Common models include the Arrhenius equation (describing the effect of temperature on aging rate), power function model (characterizing the correlation between cycle number and capacity degradation), Peukert equation (correcting the rate effect), and combined models (such as the superposition model of capacity degradation = cycle aging + calendar aging). Such models are simple in calculation and strong in engineering practicality, but their generalization ability is limited by the coverage of test conditions, making it difficult to adapt to complex dynamic conditions. The third is data-driven models, which are based on machine learning, deep learning and other algorithms (such as support vector machines, LSTM, Transformer, Gaussian process regression). By mining the implicit correlation between battery

operating data (such as voltage, current, temperature) and aging state, they do not rely on clear electrochemical mechanisms and only realize aging prediction through data feature extraction and model training. They are especially suitable for complex scenarios with multi-factor coupling and strong nonlinearity. In recent years, with the accumulation of battery monitoring data and the improvement of algorithm computing power, the fusion of data-driven models and mechanism-based models (such as deep learning models under mechanism constraints) has become a research hotspot, which not only retains the physical boundaries of the mechanism but also optimizes parameter adaptability through data-driven methods, balancing interpretability and accuracy.

This paper establishes an aging model focusing on capacity degradation caused by SEI film growth. The growth rate of SEI film thickness is positively correlated with charging current and temperature. The capacity degradation model is as follows:

$$Q(N) = Q_0 \exp \left(-k \cdot I^m \cdot \exp \left(\frac{-E_a}{RT} \right) dt \right)$$

Where, $Q_0 = 2.6 \text{ Ah}$ is the initial capacity, N is the number of cycles, $k = 1.2 \times 10^{-6}$ is the degradation coefficient, $m = 1.5$ is the current influence factor, $E_a = 30 \text{ kJ/mol}$ is the activation energy, and $R = 8.314 \text{ J/(mol}\cdot\text{K)}$ is the gas constant.

3. Design of Reinforcement Learning Fast Charging Strategy

3.1. Reinforcement Learning Framework

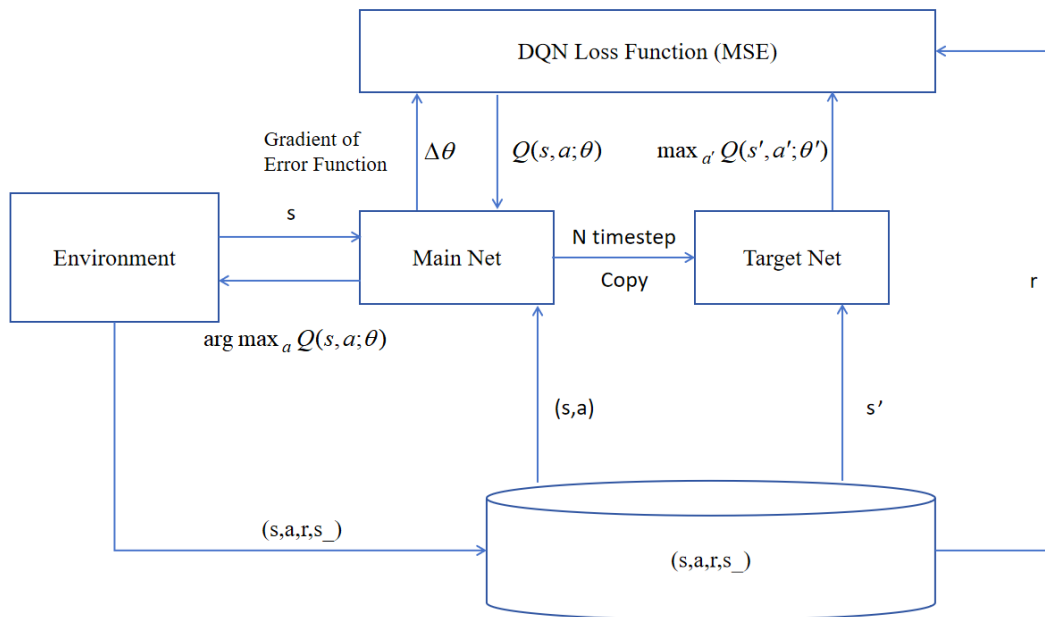


Figure 2. Fast Charging Strategy Optimization Framework Constructed by DQN Algorithm

In the scenario of battery fast charging optimization, the dynamic adjustment of charging current needs to be made based on the real-time SOC, temperature, internal resistance, and state of health (SOH) of the battery. However, there is a highly nonlinear correlation between these states and battery life degradation. Traditional static charging curves or model-based optimization methods often fail to achieve the global optimization of fast charging performance and battery life due to model simplification or insufficient condition matching. Deep Q-Network (DQN) combines the feature extraction capability of deep learning with the reinforcement learning framework of Q-learning, which can effectively handle the perception and decision-

making problems in high-dimensional state space, providing a feasible path for solving this complex sequential decision-making problem. Based on this, this paper adopts the DQN algorithm to construct a fast charging strategy optimization framework, as shown in Figure 2. The agent perceives the real-time state of the battery, outputs charging current adjustment instructions, the environment (battery model) feeds back state changes and reward values, and the agent optimizes the strategy by updating the Q network to achieve the balance between fast charging and life.

The DQN algorithm is used to construct the fast charging strategy optimization framework, as shown in Figure 2. The agent perceives the real-time state of the battery, outputs charging current adjustment instructions, the environment (battery model) feeds back state changes and reward values, and the agent optimizes the strategy by updating the Q network to achieve the balance between fast charging and life.

3.2. Definition of State Space and Action Space

State space S: A high-dimensional state vector is selected to fully characterize the battery state, $S = [\text{SOC}, U, T, U_p]$, where $\text{SOC} \in [0.2, 0.8]$ (core fast charging interval), $U \in [3.0, 4.2]$ V, $T \in [20, 45]$ °C, and $U_p = U_{C1} + U_{C2}$ is the total polarization voltage.

Action space A: A discrete action space is adopted, the charging current step is set to 0.1C, and the action set $A = 1.0C, 1.1C, \dots, 2.0C$, which not only ensures fast charging efficiency but also avoids polarization impact caused by sudden current changes.

3.3. Reward Function Design

A multi-objective reward function R is constructed to balance fast charging speed, polarization suppression, and life protection:

$$R = \omega_1 R_{\text{SOC}} + \omega_2 R_{U_p} + \omega_3 R_T + \omega_4 R_{\text{time}}$$

The definitions of each sub-reward are as follows:

R_{SOC} : SOC improvement reward, $R_{\text{SOC}} = 0.1 \times (\text{SOC}_{t+1} - \text{SOC}_t)$, encouraging rapid SOC growth;

R_{U_p} : Polarization suppression penalty, $R_{U_p} = -0.5 \times \min(U_p - 0.05, 0)$, penalizing when the polarization voltage exceeds 0.05V;

R_T : Temperature protection penalty, $R_T = -1.0 \times \min(T - 40, 0)$, increasing the penalty when the temperature exceeds 40°C;

R_{time} : Time penalty, $R_{\text{time}} = -0.01$, avoiding the strategy being too conservative to extend charging time.

The weight coefficients are determined through optimization: $\omega_1 = 0.4$, $\omega_2 = 0.3$, $\omega_3 = 0.2$, $\omega_4 = 0.1$.

3.4. Core Formulas of DQN Algorithm

DQN improves training stability through experience replay and target network. The Q network update formula is as follows:

$$L(\theta) = E_{(s,a,r,s') \sim D} \left[\left(y_j - Q(s, a; \theta) \right)^2 \right]$$

Where, the target Q value $y_j = r + \gamma \max_{a'} Q(s', a'; \theta^-)$, $\gamma = 0.95$ is the discount factor, θ is the current Q network parameter, θ^- is the target network parameter, which is synchronized every 100 steps.

4. Experimental Verification and Result Analysis

4.1. Experimental Setup

Experimental object: 18650 ternary lithium battery (capacity 2.6Ah, nominal voltage 3.6V, cut-off voltage 4.2V); Experimental equipment: Battery test system (CT-4008-5V10A), high and low temperature environmental chamber (THI-408); Comparison strategies: (1) Traditional CC-CV strategy (1.5C constant current + 4.2V constant voltage); (2) PID adaptive strategy; (3) Proposed DQN strategy; Evaluation indicators: Fast charging time (SOC 20% → 80%), cycle life (number of cycles when capacity retention rate \geq 80%), peak polarization voltage, capacity retention rate (after 1000 cycles).

4.2. Comparison of Fast Charging Performance

Table 2 shows the quantitative data of fast charging performance. The results show that the fast charging time of the proposed DQN strategy is 42.8 minutes, which is 17.2% shorter than that of the CC-CV strategy (51.7 minutes) and slightly longer than that of the PID strategy (40.3 minutes). However, the peak polarization voltage is only 0.042V, which is 34.7% lower than that of the CC-CV strategy (0.064V) and 20.8% lower than that of the PID strategy (0.053V), effectively suppressing the polarization accumulation during the charging process.

Table 2. Comparison of Fast Charging Performance

| Strategy | Fast Charging Time (min) | Peak Polarization Voltage (V) | Temperature at the End of Charging (°C) |
|-------------------------|--------------------------|-------------------------------|---|
| CC-CV | 51.7 | 0.064 | 38.5 |
| PID Adaptive | 40.3 | 0.053 | 41.2 |
| DQN Strategy (Proposed) | 42.8 | 0.042 | 35.8 |

4.3. Comparison of Cycle Life

Table 3. Comparison of Cycle Life

| Strategy | Capacity Retention Rate after 1000 Cycles (%) | Number of Cycles when Capacity Retention Rate is 80% (times) |
|-------------------------|---|--|
| CC-CV | 71.0 | 820 |
| PID Adaptive | 78.0 | 1050 |
| DQN Strategy (Proposed) | 86.3 | 1280 |

Table 3 shows the cycle life evaluation results. After 1000 cycles, the capacity retention rate of the proposed DQN strategy reaches 86.3%, which is 21.5% higher than that of the CC-CV

strategy (71.0%) and 10.8% higher than that of the PID strategy (78.0%); when the capacity retention rate drops to 80%, the number of cycles of the proposed strategy reaches 1280, which is 56.1% higher than that of the CC-CV strategy (820) and 21.9% higher than that of the PID strategy (1050). The results show that the proposed strategy significantly delays battery aging by suppressing polarization and temperature rise.

4.4. Robustness Verification of the Strategy

The robustness of the strategy is verified under different initial temperatures (25°C, 35°C, 45°C), and the results are shown in Table 4. Under extreme temperature (45°C), the fast charging time of the proposed strategy is only extended to 48.5 minutes, and the capacity retention rate (after 1000 cycles) still reaches 81.2%, which is 24.3% higher than that of the CC-CV strategy (65.3%), indicating that the strategy still has a stable optimization effect under complex working conditions.

Table 4. Robustness Test Results under Different Temperatures

| Initial Temperature (°C) | Strategy | Fast Charging Time (min) | Capacity Retention Rate after 1000 Cycles (%) |
|--------------------------|-------------------------|--------------------------|---|
| 25 | DQN Strategy (Proposed) | 42.8 | 86.3 |
| 35 | DQN Strategy (Proposed) | 45.2 | 83.7 |
| 45 | DQN Strategy (Proposed) | 48.5 | 81.2 |
| 45 | CC-CV | 55.3 | 65.3 |

5. Conclusion

This paper proposes an electric vehicle battery fast charging strategy based on DQN reinforcement learning. By constructing an accurate battery model and a multi-objective reward function, the dynamic balance between fast charging efficiency and cycle life is realized. Experimental results show that the proposed strategy controls the fast charging time from 20% to 80% SOC at 42.8 minutes, which is 17.2% shorter than the traditional CC-CV strategy, balancing fast charging speed and user experience; the capacity retention rate reaches 86.3% after 1000 cycles, and the cycle life is increased by 56.1%, effectively suppressing battery aging; the strategy shows good robustness under different temperature conditions and has engineering application potential. Future research can further integrate transfer learning to improve the generalization ability of the strategy on different types of batteries, and consider the inconsistency of battery cells to optimize the fast charging strategy in multi-cell series scenarios.

References

- [1] China Association of Automobile Manufacturers. (2024). *2024 China new energy vehicle industry development report* [Report]. China Automobile Industry Press.

- [2] Wang, Y., Li, J., & Zhang, H. (2022). A review of fast charging technologies for lithium-ion batteries in electric vehicles. *Journal of Power Sources*, 534, 231568. <https://doi.org/10.1016/j.jpowsour.2022.231568>.
- [3] Chen, C., Liu, Y., & Wang, X. (2021). Cycle life optimization of lithium-ion batteries under fast charging via model predictive control. *Applied Energy*, 295, 117089. <https://doi.org/10.1016/j.apenergy.2021.117089>.
- [4] Li, J. Q., Zhang, N., & Hu, D. Y. (2022). Multi-stage fast charging strategy for power batteries based on thermal-electric coupling model. *Journal of Mechanical Engineering*, 58(8), 189–198.
- [5] Zhang, L., Chen, Z., & Li, H. (2020). PID-based adaptive charging strategy for lithium-ion batteries considering state of health. *IEEE Transactions on Transportation Electrification*, 6(4), 1621–1630. <https://doi.org/10.1109/TTE.2020.2990898>.
- [6] Liu, S., Wang, H., & Li, Y. (2021). Q-learning based fast charging strategy for lithium-ion batteries. *Journal of Energy Storage*, 41, 102885. <https://doi.org/10.1016/j.est.2021.102885>.
- [7] Mnih, V., Kavukcuoglu, K., Silver, D., et al. (2015). Human-level control through deep reinforcement learning. *Nature*, 518(7540), 529–533. <https://doi.org/10.1038/nature14236>.