

# *Optimization of Charging Strategies for New Energy Vehicles Based on Reinforcement Learning Algorithms*

Lei Yao<sup>1,\*</sup>

<sup>1</sup>*Zeekr Intelligent Technology Holding Limited, Hangzhou City, China*

*\*Corresponding author*

**Keywords:** Reinforcement learning algorithms, New energy vehicles, Optimization of charging strategy

**Abstract:** With the popularization of new energy vehicles (NEVs) and the increasing severity of traffic congestion, charging difficulties have become a concern for people. Effective management and optimization of the NEV charging process have become increasingly important. This not only concerns the safety and stable operation of the power grid, but also directly affects the efficiency of road traffic, the utilization of renewable energy, and the charging experience and cost for users. The charging behavior of NEVs is a complex process that involves multiple considerations, such as grid load balancing, availability and efficiency of charging stations, user charging needs, and electricity prices. To address these issues, this paper proposes a NEV charging optimization strategy based on reinforcement learning (RL) algorithm, which can handle high-dimensional and complex environments and effectively deal with randomness and uncertainty factors. This strategy can not only reduce the load fluctuation of the power grid, improve the safety and stability of the power grid, but also reduce the charging time and cost of users, improve charging efficiency and user satisfaction. Meanwhile, by combining renewable energy, this strategy can also promote sustainable development, reduce reliance on traditional energy, and improve the utilization rate of renewable energy.

## **1. Introduction**

With the vigorous development of the global economy, energy demand has sharply increased, and the environmental problems caused by it have become increasingly prominent [1]. Especially automobile exhaust emissions have become one of the main culprits leading to global warming and air pollution [2]. In this context, NEV has received widespread attention as an important means to address climate change and promote the transition to green energy [3]. With the continuous progress and maturity of technology, the market share of NEV is rapidly increasing [4]. At the national level, there is also a strong promotion of the use of clean energy to reduce environmental pollution, which has led to a rapid increase in the number of NEVs. The natural advantage of NEV lies in its ability to significantly reduce greenhouse gas emissions and reduce dependence on fossil fuels, which is of great significance for achieving the country's "dual carbon" goals. Moreover, as an important demand response resource, NEV is playing an increasingly important role in the power system (EPS). They can provide flexible support for EPS, help balance grid load, and improve the

stability and efficiency of EPS [5].

However, with the rapid growth of NEV users, it has also brought a series of new challenges [6]. The dynamic driving behavior and random charging behavior of NEVs have complex coupling interactions with urban power grids and transportation networks [7]. This interaction may not only exacerbate traffic congestion, but also lead to an unstable increase in grid load. Therefore, it is necessary to design a reasonable charging scheduling strategy and carry out collaborative control of the charging process of NEV. At the same time, issues such as insufficient battery power and poor user charging experience are increasingly becoming bottlenecks that constrain the further development of NEVs. To solve these problems, it is necessary to start from multiple aspects. Firstly, it is necessary to further improve the charging infrastructure of NEV, increase the coverage and charging efficiency of charging stations, in order to meet the growing charging needs of users. Secondly, it is necessary to optimize the driving and charging behavior of NEVs through technological innovation and intelligent management methods, and reduce their impact on urban power grids and transportation networks [8].

RL is an important branch in the field of machine learning (ML), which focuses on enabling agents to continuously optimize their behavioral strategies through interactive learning with the environment, in order to achieve the goal of maximizing cumulative rewards or minimizing certain losses. In RL, agents interact with the environment by executing a series of actions and evaluate the effectiveness of their actions based on the reward signals returned by the environment. The goal of an intelligent agent is to learn a strategy that can guide it to choose the best action in different states to maximize long-term cumulative rewards. In the context of optimizing NEV charging strategies, RL can be used to train agents to learn how to make optimal charging decisions based on factors such as power grid status, charging station availability, user demand, and electricity prices. By interacting with the environment, intelligent agents can learn a charging strategy that maximizes long-term benefits, thereby improving the safety and stability of the power grid, reducing user charging time and costs, and improving the utilization rate of renewable energy.

## 2. Feasibility Analysis of NEV Charging Load Control

### 2.1 Feasibility

The charging load of NEV has randomness in time and space, which is one of the key reasons why disorderly charging brings pressure and challenges to the power grid. Disorderly charging may lead to increased fluctuations in power grid load, and may even cause local power grid overload, which has adverse effects on the stability and safe operation of EPS. Therefore, when studying the NEV charging load prediction method, it is necessary to fully consider this random distribution feature. In order to effectively manage the charging load of NEV, achieving orderly charging control has become a necessary means. Ordered charging uses intelligent scheduling and control technology to adjust the charging time of NEVs reasonably and optimize their distribution over time. This can not only prevent a large number of NEVs from charging during peak hours, thereby preventing the accumulation of load peaks, but also effectively alleviate the adverse impact of NEV charging load on the power grid. It is very important to distinguish between office and residential areas in the assessment of the number of NEVs in a region. By collecting and analyzing these data, the number of NEVs can be more accurately evaluated or determined, providing strong support for subsequent charging facility planning and grid scheduling [9].

In addition, understanding the battery power characteristics of NEV is also crucial for charging load prediction and orderly control. This includes specific data on the charging power requirements and battery capacity of different vehicle models. In terms of NEV driving patterns, user decisions have a significant impact on the selection of charging locations. By collecting and analyzing user

driving and charging data, future charging needs can be more accurately predicted, providing more accurate basis for power grid scheduling and charging facility planning. The comprehensive consideration of the randomness of NEV charging load, orderly charging control, evaluation of regional NEV ownership, battery power characteristics, and driving patterns is the key to effectively addressing the challenges of NEV charging load. By comprehensively utilizing advanced prediction technology, intelligent scheduling, and control methods, the sustainable development of the NEV industry can be promoted, while ensuring the safe and stable operation of EPS [10].

## 2.2 Real Time Optimization Scheduling Scenario Modeling

To improve the real-time efficiency of NEV charging management, adopting a layered management architecture based on NEV Aggregator (NEVA) is an effective solution. This architecture can optimize the flow of information and the exchange of energy, thereby improving the operational efficiency of the entire system. As shown in Figure 1, NEVA plays a crucial role in the hierarchical management architecture, serving as a "bridge" between the power grid and NEV, achieving the upload and release of information and the transfer of energy. Specifically, NEVA is responsible for transmitting real-time information of NEVs to the power grid, including but not limited to vehicle location, battery status, charging needs, etc., enabling the power grid to make more accurate and timely decisions.

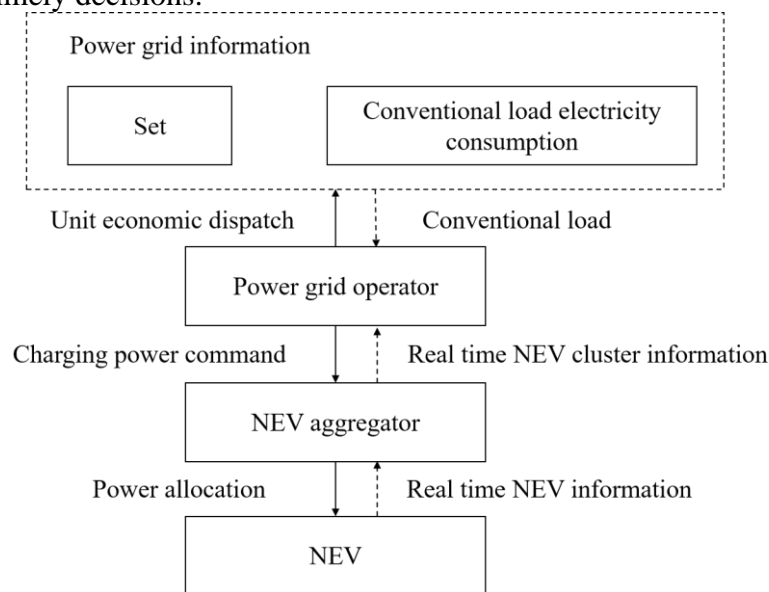


Figure 1: Real time phase architecture diagram

At the same time, NEVA also transmits subsidy price information to the power grid and guides and controls the charging process to ensure that charging behavior meets the scheduling requirements of the power grid. Through NEVA's hierarchical management, NEVA can collect and process NEV charging information in real time, and transmit this information to the power grid in a timely manner, enabling the power grid to grasp NEV charging needs in real time and make more accurate scheduling. NEVA can flexibly adjust the charging behavior of NEVs according to the scheduling requirements of the power grid, including charging time, charging power, etc., to meet the load balancing needs of the power grid. Through the subsidy price transmission mechanism of NEVA, NEV users can be incentivized to charge during periods of low grid load, thereby reducing charging costs and improving the economic benefits of the grid. The hierarchical management architecture of NEVA can ensure accurate information transmission and stable energy exchange, thereby improving the operational reliability of the entire system.

### 3. Optimization Strategy for NEV Charging Based on RL Algorithm

#### 3.1 RL Algorithm

RL algorithm is an effective algorithm for sequential decision problems, which learns and optimizes decision strategies through the interaction between agents and the environment. Figure 2 shows the basic composition and process of the RL framework. In this framework, the agent observes the current state of the environment and selects an action to execute based on its strategy. After receiving this action, the environment will transition to a new state and generate a reward signal as feedback for the action. The goal of an intelligent agent is to continuously interact with the environment, adjust its strategies, and maximize the accumulated rewards throughout the entire interaction process. When an intelligent agent adopts a certain strategy and receives positive rewards from the environment, it means that the strategy is effective in the current environment, so the agent is more likely to adopt similar strategies in the future. This process of guiding strategy adjustment through reward signals is the core mechanism of RL.

In the optimization of charging strategy for new energy vehicles, the Q-Learning algorithm is used to constantly update the state-action value function;

$$Q(s, a) \leftarrow Q(s, a) + \alpha [r + \gamma \max_{a'} Q(s', a') - Q(s, a)] \quad (1)$$

Here,  $Q(s, a)$  is the expected return of taking action  $a$  in state  $S$ ,  $\alpha$  is the learning rate,  $r$  is the immediate reward,  $\gamma$  is the discount factor,  $s'$  is the next state, and  $a'$  is the possible action in the next state. Every time the vehicle chooses the charging strategy, it will update the action value in the current state according to the instant reward and the expected future value.

The state-action value function satisfies bellman equation, which expresses the relationship between the immediate reward that can be obtained after taking an action in the current state and the expected value after moving to the next state:

$$Q^\pi(s, a) = \sum_{s', r} p(s', r | s, a) \left[ r + \gamma \sum_{a'} \pi(a' | s') Q^\pi(s', a') \right] \quad (2)$$

Under the given strategy  $\pi$ , the value of state-action pair  $(s, a)$  is determined by the value of environmental dynamic  $p(s', r | s, a)$  and subsequent state-action pairs.

Through the strategy gradient method, the parameters of charging strategy can be updated according to the state-action value function, thus optimizing the expected total reward:

$$\nabla_{\theta} J(\theta) \propto \sum_s \mu^\pi(s) \sum_a Q^\pi(s, a) \nabla_{\theta} \pi(a | s) \quad (3)$$

This is the basic formula for updating the strategy parameter  $\theta$  in the strategy gradient method, where  $J(\theta)$  is the expected total reward and  $\mu^\pi(s)$  is the steady distribution of the state  $s$  under the strategy  $\nabla$ . In the process of charging, we will adjust the charging strategy according to the historical data and the current state to maximize the long-term benefits.

The advantage function  $A$  measures the advantage of taking a specific action in a given state compared with other actions:

$$A^\pi(s, a) = Q^\pi(s, a) - V^\pi(s) \quad (4)$$

Where  $V^\pi(s)$  is a state value function.

For the deterministic charging strategy, the deterministic strategy gradient is used to update the

strategy parameters to ensure that the selected charging action in a given state can maximize the long-term reward:

$$\nabla_{\theta} J(\theta) = E_{s \sim \rho^{\mu}} [\nabla_{\theta} \mu\theta(s) \nabla_a Q^{\mu}(s, a) |_{a = \mu\theta(s)}] \quad (5)$$

For deterministic policy  $\mu\theta(s)$ , this is the formula for updating policy parameters, where  $\rho^{\mu}$  is the state access distribution.

In the process of interacting with the environment, intelligent agents usually use a trial and error approach for learning. This means that intelligent agents need to try various different strategies and gradually improve their strategies by observing feedback from the environment. This learning approach does not require clear guidance or labels from the environment, making RL particularly suitable for complex environments that are difficult to model or obtain labels.

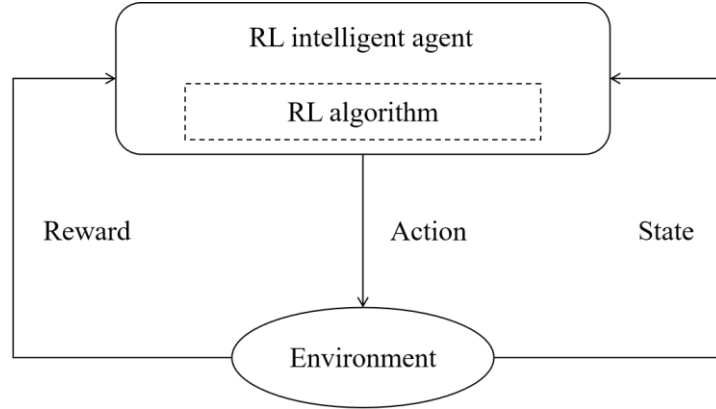


Figure 2: RL algorithm structure

### 3.2 Optimization Strategy

In static charging scheduling scenarios, the optimization objective mainly focuses on minimizing the charging cost for NEV users. With the continuous development of smart grids and intelligent transportation systems, these two systems operate together, providing NEV users with a large amount of data related to the power grid and transportation network. These data can be used for NEV charging navigation, helping users find the best charging site and time to minimize charging costs. Using data-driven methods to model and mine the data of the "NEV Cluster Optimization Energy Storage Cloud Platform", obtain the driving and charging information required for NEV travel, urban charging station information, and dynamic traffic network information. Secondly, the RL method is applied to solve the multi-objective optimization of NEV charging navigation problem. Real time information of "vehicle station network" is mined as state input, and appropriate charging stations and charging paths are recommended for car owners through action execution. In this RL paradigm, intelligent agents (i.e. NEV users or charging navigation systems) gradually optimize their strategy for selecting charging stations through interactive learning with the environment. The vehicle by vehicle recommendation RL framework can flexibly cope with irregular charging requests and high-dimensional environmental characteristics. By continuously interacting and learning from the environment, intelligent agents can gradually improve the accuracy of their recommended charging stations, thereby helping NEV users minimize charging costs.

## 4. Conclusions

To address the economic dispatch problem of power grid units with NEV randomness in the real-time stage, orderly charging has become an effective solution strategy. Ordered charging can dynamically adjust the charging power of each time period by planning and controlling the charging behavior of NEVs in a reasonable manner, thereby achieving optimal resource allocation and efficient energy utilization while meeting the operational requirements of the power grid and user sides. This article first analyzes the impact of NEV randomness on the economic dispatch of power grid units, and proposes a NEV charging optimization strategy based on the RL algorithm. Due to the uncertainty and randomness of NEV charging behavior, it will bring additional load fluctuations and scheduling difficulties to the power grid. Therefore, it is necessary to balance this uncertainty and ensure the stable operation and economy of the power grid through orderly charging. Using the RL algorithm, enable the intelligent agent to learn how to select the optimal action based on the current state in the interaction between the immediate power grid and the NEV system. Through continuous trial and error and learning, intelligent agents can gradually learn a charging strategy that can minimize grid load fluctuations and unit operating costs. Ordered charging is an effective strategy for solving the economic dispatch problem of power grid units with NEV randomness in the real-time stage. By constructing an optimized and orderly charging method, reasonable planning and control of NEV charging behavior can be achieved, thereby balancing grid load fluctuations, improving energy utilization efficiency, and meeting user charging needs.

## References

- [1] Zhang Wenxin, Li Ran, Zang Xiangdi, et al. Real-time scheduling strategy optimization of electric vehicle power station based on reinforcement learning [J]. *Electric Power Automation Equipment*, 2022(010):042.
- [2] Shen Guohui, Zhao Rongsheng, Dong Xiao, et al. Navigation strategy for electric vehicle charging based on multi-information interaction and deep reinforcement learning [J]. *Southern Power Grid Technology*, 2022(001):016.
- [3] Yao Tianhao, Ye Peng, Zhao Siwen. Research on Orderly Charging Strategy of Electric Vehicles Based on Bilevel Optimization Algorithm [J]. *Journal of Shenyang Institute of Engineering: Natural Science Edition*, 2019, 15(2):8.
- [4] Zhang Yanyu, Rao Xinpeng, Zhou Shukui, et al. Research progress of electric vehicle charging scheduling algorithm based on deep reinforcement learning [J]. *Power System Protection and Control*, 2022(016):050.
- [5] Yang Yue, Pan Gang, Zhu Jinghua. Intelligent charging strategy for real-time electric vehicles under real traffic data [J]. *Computer and Digital Engineering*, 2023, 51(1):133-141.
- [6] Su Mingliang, Yao Fang. Energy management strategy of hybrid electric vehicle based on deep reinforcement learning [J]. *Electric Automation*, 2023, 45(4):115-118.
- [7] Wu Jiawu, Qiu Xiaoyan, Pan Yinji, et al. Research on Orderly Charging Strategy of Electric Vehicles Based on Improved Chicken Swarm Algorithm [J]. *Electric Measurement and Instrument*, 2019, 56(9):7.
- [8] Chen Yinkun, Wu Songze, Wu Yujian. Research on an optimization method for orderly charging of electric vehicles [J]. *Integrated circuits and embedded systems*, 2023, 23(11):75-79.
- [9] Chen Guo, Wang Xiuli, Yuan Shengqi, et al. Deep reinforcement learning and orderly charging strategy for large-scale charging stations [J]. *Power System Automation*, 2023, 47(2):8.
- [10] Zhan Hua, Jiang Changxu. A Method of Electric Vehicle Charging Guidance Strategy Based on Hierarchical Deep Reinforcement Learning [J]. *Electric Power Automation Equipment*, 2022, 42(10):9.