A DRL Unified Interference Suppression Method For MCR-WPT System

Yujun Luo *

School of Electrical Engineering, Xi'an Jiaotong University, Xi'an, 710049, China

* Corresponding Author Email: luoyj6@stu.xjtu.edu.cn

Abstract. Magnetic coupling resonant wireless power transfer (MCR-WPT) technology often experiences performance decline due to compound interference, including coil displacement, metal foreign object intrusion, load mutation, and electromagnetic interference (EMI). Existing control methods struggle to suppress such interference uniformly; thus this paper proposes a unified interference suppression method based on deep reinforcement learning (DRL). First, a simulation environment containing the four types of typical interference is constructed in Simulink. Subsequently, a controller based on the deep deterministic policy gradient (DDPG) algorithm is designed, which regards the inverter's operating frequency and duty cycle as a continuous two-dimensional action space. A reward function is designed, incorporating power tracking, transmission efficiency and electrical safety, guiding the agent to balance multiple objectives in a complex environment. The curriculum learning strategy is also introduced for staged training. The analysis of open-loop characteristics of the system demonstrates nonlinearity and sensitivity under complex disturbances as well as the conflict between optimal power and efficiency goals. These findings emphasize the necessity of the adaptive intelligent control method. The DRL controller designed in this paper provides a complete framework for complex disturbance in MCR-WPT systems, enabling self-adaptation control without relying on an accurate system model, showing strong application potential.

Keywords: MCR-WPT; DRL; Compound interference; Self-adaptation control.

1. Introduction

Magnetic coupling resonant wireless power transfer (MCR-WPT) technology, capable of non-contact power delivery and operation in harsh environments, has been widely adopted in fields such as electric vehicles, smart homes, and implantable medical devices, showing great potential to overcome the constraints of traditional wired power supply [1][2]. However, MCR-WPT systems are often confronted with coil displacement, metal foreign object intrusion, load mutation, electromagnetic interference (EMI) in practical applications, which give rise to a decrease in efficiency and fluctuations in power, severely reducing their reliability [3].

In existing research, the multifrequency superposition methodology achieves power distribution and efficient transmission by modulating the resonant frequencies of the coils [4]. The Q-learning algorithm enhances system compatibility and resistance to interference by optimizing the compensation capacitors. However, these methods primarily deal with single interference sources and struggle to uniformly suppress compound interference [5]. The impact of varying coupling coefficients on power transmission is demonstrated through the phenomena of frequency splitting [6]; optimizing compensation topology and coil structures can improve robustness against misalignment [2]. Nevertheless, neither approach dynamically responds to complex interference.

This paper proposes a unified interference suppression method for MCR-WPT systems based on deep reinforcement learning (DRL). Firstly, a compound interference model including coil displacement, metal foreign object intrusion, load mutation, and EMI is established; then, a deep reinforcement learning model is designed; finally, the model is used to achieve adaptive interference suppression.

2. Multi-Interference Model of MCR-WPT Systems

2.1. Basic Circuit Model of MCR-WPT Systems

A typical MCR-WPT system circuit mainly consists of five parts: a high-frequency power source or inverter circuit, a transmitting end impedance-matching network, a mutual inductor, a receiving end impedance-matching network, and the load. This paper builds the MCR-WPT system circuit with a series/series (S/S) topology on the MATLAB/Simulink platform. This topology is widely used and has high transmission efficiency [7].

DC is transformed into a high-frequency square wave after passing through the Full-Bridge Converter, and then enters the transmission end impedance-matching network. The network, a S/S topology, consists of a compensation capacitor, equivalent resistance, and the equivalent inductance of the transmitting coil. The value of the capacitor is determined in order to achieve series resonance with the inductance at the operating frequency. During resonance, the transmitting end presents purely resistive, significantly reducing reactive power.

The mutual inductor is composed of transmitting coils and a receiving counterpart that are spatially separated but magnetically coupled. The efficiency of energy exchange between the coils is determined by the coupling coefficient k and the respective quality factor Q [8]. The k is related to the shape, size, turns, relative position and posture of the coils, and the Q indicates the ratio of the energy storage to loss in the resonant circuit. When resonant and impedance-matched, the system efficiency \eta can be approximately expressed as below.

$$\eta \approx \frac{k^2 Q_P Q_S}{1+k^2 Q_p Q_s} \tag{1}$$

When two coils with high quality factors resonate, even if the k is small, the system can still achieve high efficiency [9]. Similar to the transmitting end, the receiving part also has a compensation capacitor, whose value is determined to resonate with the inductance of the receiving coil, maximizing the received inductive energy. In Simulink, all the devices are constructed with modules in the Simscape Electrical Library, and the circuit model is shown in figure 1.

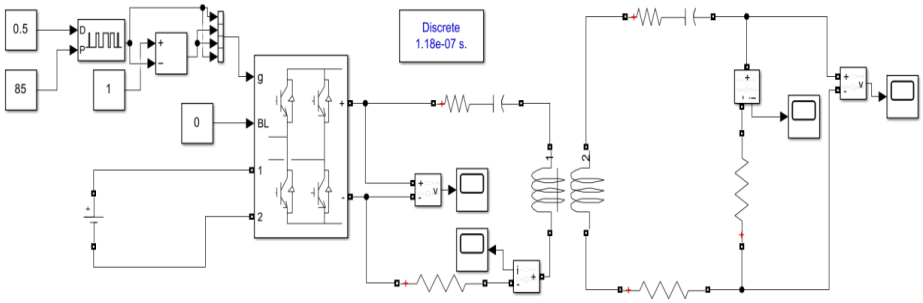


Fig. 1 Basic circuit model of the MCR-WPT system

2.2. Compound Interference Modeling

In practical applications, the system inevitably encounters various interferences. The ideal stable resonant state is rather difficult to maintain, leading to impedance mismatch and significant fluctuations in power transmission and efficiency. To simulate the complex environment, this paper adopts the 'Separate Modeling Method'. Models for the coil displacement, metal foreign objects, load mutation and EMI are established separately, subsequently integrated into a unified compound interference environment.

2.2.1 Model for coil displacement

Coil displacement includes relative position changes in the axial and radial directions as well as angular deflection. It is modeled in this paper as continuous and periodic low-frequency oscillations, simulating minor vibrations during the driving of an electric vehicle. Coil displacement primarily

causes a change in the mutual inductance M, which in turn affects the coupling coefficient k. The change in the k affects the transmission power and efficiency of the system. This paper employs a sine wave to model the disturbance, and the expression for the coupling coefficient $k_{displacement}(t)$ is

$$k_{displacement}(t) = A_k \sin(\omega_k t + \phi_k) \tag{2}$$

where A_k is the amplitude of the displacement; ω_k is the disturbance frequency, modeling the rate of the oscillation. Both of them are randomized, which can be achieved through the Sine Wave module in Simulink.

2.2.2 Model for metal foreign objects

Since the charging environment is often open, metal foreign objects such as coins, cans, and vehicles may accidentally enter the magnetic field between or near the coils, which can cause performance degradation. In an alternating magnetic field, conductive metal objects induce eddy currents, whose magnetic field is opposite to the original one, equivalent to weakening the mutual inductance M between the coils according to Faraday's law. Thereby, the coupling coefficient k is reduced. The effect is similar to coil displacement. Additionally, eddy currents in a metal object generate heat, which is a loss of energy that comes from the WPT system, equivalent to increasing the resistance R of the resonant circuit.

In this paper, the interference is randomly triggered or not by a step signal. When the interference occurs, both k and R will change simultaneously. The nominal value of the coupling coefficient decreases Δk_{metal} and the equivalent resistance of the coil increases ΔR_{metal}, both of which are randomized. In Simulink, both Gain modules are driven simultaneously by the Step module so that the k and R of the mutual inductor can be modified respectively.

2.2.3 Model for load mutation

WPT systems often encounter load variations, such as sudden changes caused by the connection and disconnection of devices on a shared power supply platform. When the load changes suddenly, the output voltage and current jump, giving rise to a drastic fluctuation in output power and charging efficiency.

To simulate the dynamic connection of the load, a parallel switching circuit is designed at the load end of the model, randomly triggered or not. The switch is closed or opened at a specific moment, causing the total load to suddenly change from R_{load} to $R_{load}//R_{mutation}$ or the opposite.

2.2.4 Model for EMI

MCR-WPT systems operate at high frequencies, typically ranging from several kHz to several MHz [10][11]. It is not only a potential EMI emission source but also susceptible to interference from other electromagnetic signals in the environment, inducing noise voltages. The frequency of the noise voltage is often unrelated to the system's operating frequency, contaminating the voltage and current in the resonant circuit.

This model equivalently represents external EMI as a high-frequency sinusoidal AC voltage source in series with the transmitting coil, which is established through the Controlled Voltage Source module in Simulink. The complex electromagnetic environment is simplified by introducing energy of different frequencies independent of the system's operating state without losing generality.

2.3. Unified Model for Compound Interference

The above interference models are integrated together. Among them, the coupling coefficient $k(t)$ is composed of three parts

$$k(t) = k_{offset} + k_{displacement}(t) + \Delta k_{metal}(t) \tag{3}$$

where the k_{offset} is the base offset. The model for compound interference is demonstrated in figure 2.

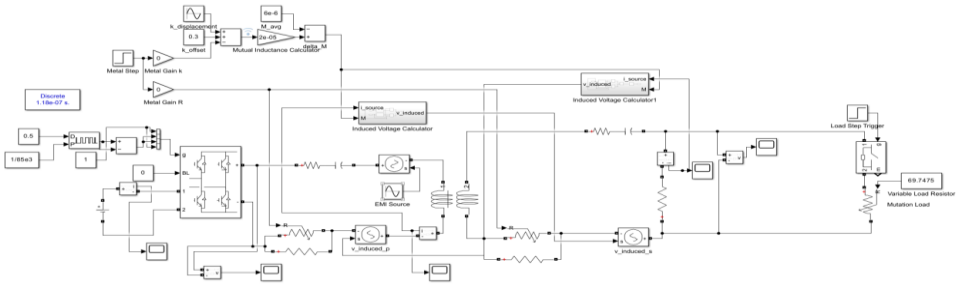


Fig. 2 Model for compound interference

3. Unified Interference Suppression Based on DDPG

This research employs DRL to address the complex, dynamic, and non-linear compound interference. In DRL, the agent interacts directly with the environment, conducts trial-and-error learning, and finds the optimal control strategy without an accurate system model. The Deep Deterministic Policy Gradient (DDPG) algorithm is effective for control problems with continuous state and action spaces, which makes it suitable for uniformly suppressing interference precisely in the MCR-WPT system.

3.1. Markov Decision Process

This research formulates the interference suppression problem of the MCR-WPT system as a Markov Decision Process (MDP).

A state space is a set of all system states and is the input domain of the agent. The state space $s_t \in S$ in this paper is a three-dimensional vector

$$s_t = \left[e_P(t), \int e_p(t), \eta(t)\right]^T \tag{4}$$

where the power error $e_p(t) = P_{desired} - P_{out}(t)$ reflects the track performance of the desired power; $\int e_p(t)$ is the integral of the power error in order to eliminate the static error of the system; the transmission efficiency $\eta(t) = \frac{P_{out}(t)}{P_{in}(t)}$ considers energy efficiency.

An action space is a set of operations that an agent can perform. In this paper, both the operating frequency and the duty cycle are the control means of the agent, forming a two-dimensional continuous action space. The frequency range is constrained to [70kHz, 100kHz], covering the resonant point and sufficient to adjust; the range of the duty cycle is [0.1, 0.9].

A reward function is the goal for the agent, guiding its actions. The reward function $r_t = R(s_t, a_t)$ in this paper aims for power matching, high efficiency, and safe operation, expressed as below.

$$r_t = r_{power} + r_{efficiency} + r_{safety} + p_{voltage} + p_{current} \tag{5}$$

It is divided into reward and penalty parts. The power tracking reward $r_{power} = w_{track} \cdot e^{-\frac{e_p(t)^2}{2\sigma^2}}$ is in the form of a Gaussian function, peaking when the error is 0, and is the most weighted reward term; the efficiency reward $r_{efficiency} = w_{efficiency} \cdot \eta(t)$ encourages the agent to reach an operating point of high-efficiency; the safety reward r_{safety} gives the agent a small reward when the voltage and current are within the safe range. Conversely, penalty terms, p_{voltage} and p_{current}, are designed to prevent dangerous operations. when the system crosses the safety boundary, extremely severe penalties are imposed through the tanh function and the magnitude of the penalty increases as the deviation rises.

State transition probability $P(s_{t+1}|s_t, a_t)$ describes the changing pattern of the state. The states of this project are modeled by Simulink, which is extremely complex and nonlinear, and it is difficult to write in an analytical form. This is why the model-independent DRL algorithm DDPG is adopted.

3.2. DDPG Implementation

DDPG adopts an Actor-Critic architecture, which consists of a primary actor network, a primary critic network, and their two corresponding target networks, respectively [12].

The Actor network $\mu(s_t|\theta^{\mu})$ is composed of multiple fully connected layers, and finally maps its output to the action space, such as $[70,100]$ kHz in this paper, through a tan activation function and a scaling layer. It outputs the action a_t based on the current state s_t .The Critic network $Q(s_t, a_t|\theta^{Q})$ evaluates the quality of the Actor's actions and outputs the Q-value based on the s_t and the a_t, estimating the reward that can be obtained in the future.

The training process of the agent relies on an experience replay mechanism. After interacting with the Simulink environment, it generates experience tuples (s_t, a_t, r_t, s_{t+1}), which are stored in the experience replay buffer. During each update, a mini-batch of data is randomly selected from it. The Actor network is updated by the policy gradient method, adjusting the strategy to enable the Critic network to provide a higher Q-value. The Critic network is updated by minimizing the temporal difference error to make the output Q-value close to the more stable target one calculated by the target network.

4. Simulation Analysis

4.1. Simulation Platform and Parameters

This research employs MATLAB R2024b and Simulink for simulations. The main parameters of the MCR-WPT system are shown in table 1.

Table 1. Parameters of the MCR-WPT system

Parameter	Symbol	Value
DC input voltage	V_{in}	$48V$
Coil Self-Inductance (Primary, Secondary)	L_p, L_s	$20\mu H$
Coil Internal Resistance (Primary, Secondary)	R_p, R_s	0.1Ω
Compensation Capacitance (Primary, Secondary)	C_p, C_s	$175nF$
Nominal Load Resistance	R_{load}	30Ω
Nominal Resonant Frequency	f_0	$\sim 85kHz$

The main training hyperparameters of DDPG are shown in table 2.

Table 2. Training hyperparameters of DDPG

Hyperparameters	Value	Description
Agent Sample Time, $\boldsymbol{T_s}$	$1.18 \times 10^{-4}s$	Time interval between agent decisions.
Episode Duration, $\boldsymbol{T_f}$	$0.1s$	Duration of a single training episode.
Mini-Batch Size	128	Number of experience samples drawn from the buffer for each update.
Replay Buffer Size	1×10^6	Capacity for storing experience tuples
Number of Episodes	2000	Total number of training episodes.
Actor, Critic Learning Rate	$1 \times 10^{-4}, 1 \times 10^{-3}$	Learning rate for the actor, critic network
Discount Factor, γ	0.99	Factor discounting the value of future rewards.

4.2. Open-Loop System Analysis

An open-loop analysis of the MCR-WPT system was conducted, where the duty cycle and frequency of the inverter were manually adjusted and no agent was connected. When the interference was absent, the system power presented a single sharp peak pattern, as shown in figure 3.

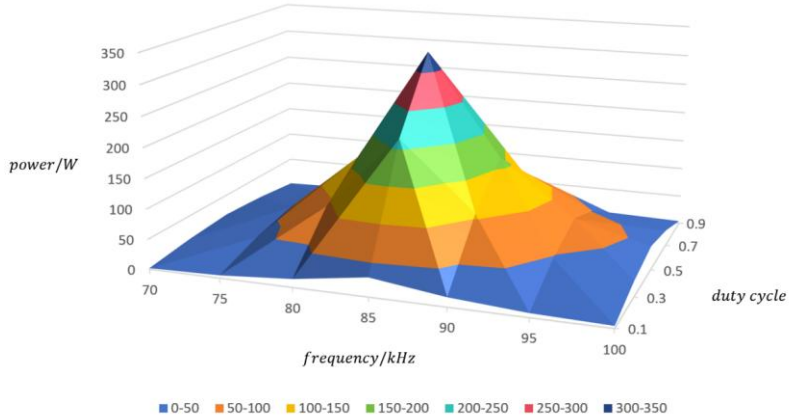


Fig. 3 Relationship between output power versus frequency and duty cycle under interference-free conditions.

The peak value reaches 329W at (85kHz, 0.5). Any slight deviation in frequency or duty cycle causes a sudden drop in power. That is to say, the system is highly nonlinear and sensitive.

After introducing compound interference, the system's response surface becomes distorted, as shown in figure 4.

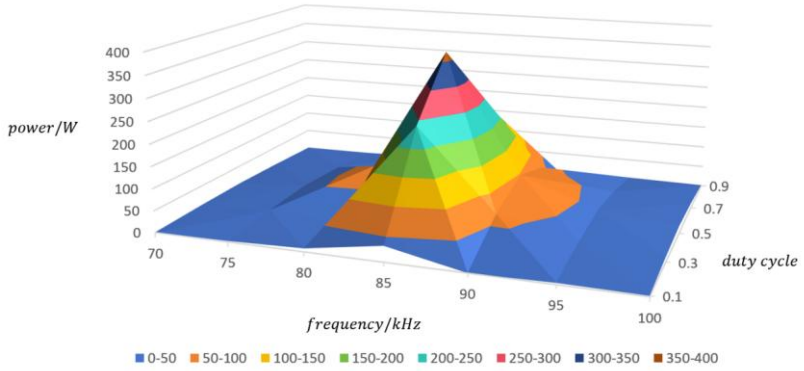


Fig. 4 Relationship between output power versus frequency and duty cycle under interference conditions.

The peak power still occurs at (85kHz, 0.5), which might be due to the relatively coarse discrete grid during parameter scanning, failing to capture the possible small peak shift in the real continuous surface. Nonetheless, the values of each data point have changed significantly, and even deteriorated sharply. In an ideal scenario, deviating from (85 kHz, 0.5) to (90 kHz, 0.5) causes the power to drop from 329 W to 162 W, a reduction of approximately 50%. However, under the same deviation with interference, the power drops from 368.4W to 97W, a decrease of nearly 74%. The interference makes the system more sensitive to parameter changes, greatly compressing the control's fault tolerance margin. Even a slight deviation can lead to a huge error, resulting in catastrophic performance degradation.

In terms of system efficiency, the compound interference has a more severe impact. As shown in figure 5, in the absence of interference, the efficiency can maintain above 80% over a wide range of operations.

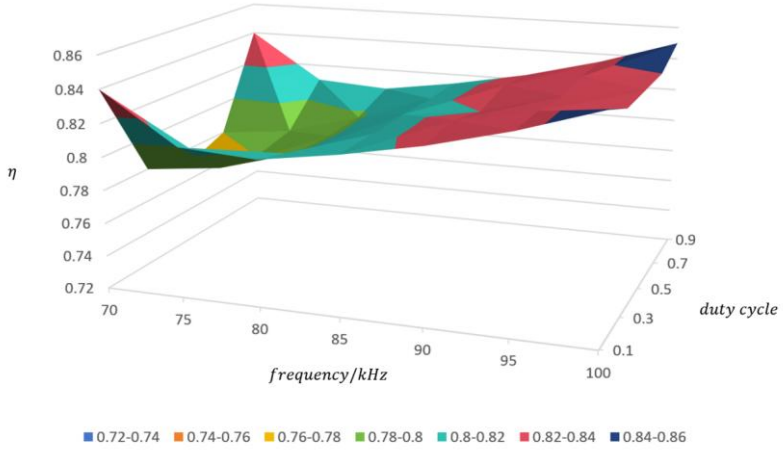


Fig. 5 Relationship between system efficiency versus frequency and duty cycle under interference-free conditions.

However, after introducing the interference, the maximum system efficiency plummets from nearly 85% to only about 62%, and the frequency of the high-efficiency operating point was significantly deviated from the resonant frequency point. It should be noted that data at 70 kHz are not shown due to severe system oscillation causing distortion. The figure 6 is illustrated below.

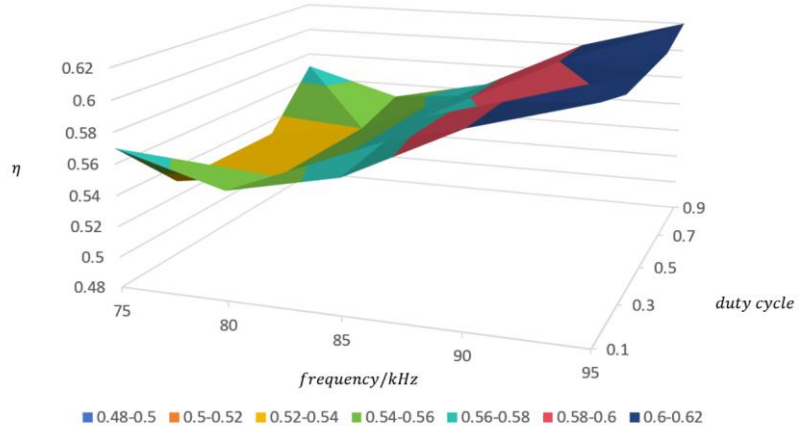


Fig. 6 Relationship between system efficiency versus frequency and duty cycle under interference conditions.

Under interference conditions, the optimal operating points of power and efficiency are separated. If only the maximum power is pursued, the system may be locked in an inefficient and poor state. Therefore, an effective controller must make multi-objective trade-offs, which is precisely the goal that the reward function is designed to guide the agent towards achieving.

4.3. Curriculum Learning Strategy

This research adopts the curriculum learning strategy, where the agent completes simple tasks first, and then gradually adapts to complex compound interference. Before each training episode, the model parameters are randomized and different interference patterns are generated to allow the agent to learn generalization strategies rather than being specific to a particular scenario.

The agent training is divided into four stages. The first stage is basic control without any interference, where the agent learns how to stabilize the power at the target value of 80W in a stable environment; the second stage introduces coil displacement, requiring adaptation to the periodically changing coupling coefficient k(t); the third stage adds load mutations; the fourth stage introduces all compound interference. The training progresses from easy to difficult, significantly reducing the learning difficulty.

4.4. Performance Potential Analysis

When the system encounters compound interference, its performance surface will instantly change from an ideal state (as shown in Figures 3 and 5) to a poor state (as shown in Figures 4 and 6). A trained agent can observe the state change and immediately output new action under the guidance of the reward function, seeking a new balance point that takes into account both power and efficiency. Due to the environmental change, the rapid adaptive control capability from one stable state to another new stable state is precisely the final control effect that this DRL framework aims to achieve, and it is also the advantage over traditional fixed-parameter controllers that require precise modeling.

5. Conclusion

This paper addresses the compound interference problems faced by the MCR-WPT system by designing a unified interference suppression method based on DRL.

Firstly, a compound interference model is constructed, integrating four typical interferences including coil displacement, metal foreign objects intrusion, load mutation and EMI. The dynamic coupling of the coil is simulated with the 'Separate Modeling Method'. Secondly, this research designs a controller with the DDPG algorithm, setting the operating frequency and duty cycle of the inverter as continuous two-dimensional action spaces, allowing the agent to simultaneously manipulate two degrees of freedom, which makes the system achieve enhanced adaptive capabilities without an accurate system model. Finally, this research designs a reward function that takes into account power tracking, transmission efficiency, and electrical constraints, and introduces the curriculum learning strategy, which can reduce the difficulty for the agent to learn. This paper provides a complete framework for DRL in MCR-WPT interference suppression. Moreover, it demonstrates the necessity of adaptive intelligent control and the framework has great potential to deal with compound interference.

This study still has limitations. Currently, the DDPG algorithm requires extensive training to converge when facing complex environments such as the one described in this paper. In the future, it is necessary to fully train the agent to verify the effectiveness of this framework. Alternative DRL algorithms with higher sample efficiency, such as Soft Actor-Critic (SAC) or Twin Delayed DDPG (TD3), may be a good choice to accelerate the training process and improve stability.

References

[1] Buja G, Bertoluzzo M, Mude K N. Design and Experimentation of WPT Charger for Electric City Car[J]. IEEE Transactions on Industrial Electronics, 2015(12).

[2] Abu Houran M, Yang X, Chen W. Magnetically Coupled Resonance WPT: Review of Compensation Topologies, Resonator Structures with Misalignment, and EMI Diagnostics[J]. Electronics, 2018(11).

[3] Li J, Li S, Wang Y, et al. Review on magnetic coupling resonant type wireless power transmission technology[J]. Electrical Measurement & Instrumentation, 2024(8).

[4] Liu F, Yang Y, Ding Z, et al. A Multifrequency Superposition Methodology to Achieve High Efficiency and Targeted Power Distribution for a Multiload MCR WPT System[J]. IEEE Transactions on Power Electronics, 2018, 33(10): 9005-9016.

[5] Liu X, Chao J, Rong C, et al. Compatibility and Performance Improvement of the WPT Systems Based on Q-Learning Algorithm[J]. IEEE Transactions on Power Electronics, 2024(8 Pt.2).

[6] Huang R, Zhang B, Qiu D, et al. Frequency Splitting Phenomena of Magnetic Resonant Coupling Wireless Power Transfer[J]. IEEE Transactions on Magnetics, 2014(11).

[7] Hou J, Chen Q, Wong S C, et al. Analysis and control of series/series-parallel compensated resonant converter for contactless power transfer[J]. IEEE Journal of Emerging and selected topics in Power Electronics, 2014, 3(1): 124-136.

[8] Sample A P, Meyer D T, Smith J R. Analysis, experimental results, and range adaptation of magnetically coupled resonators for wireless power transfer[J]. IEEE Transactions on industrial electronics, 2010, 58(2): 544-554.

[9] Zhang W, Wong S C, Chi K T, et al. Design for efficiency optimization and voltage controllability of series–series compensated inductive power transfer systems[J]. IEEE Transactions on Power Electronics, 2013, 29(1): 191-200.

[10] Hui S Y R, Zhong W, Lee C K. A critical review of recent progress in mid-range wireless power transfer[J]. IEEE transactions on power electronics, 2013, 29(9): 4500-4511.

[11] Puterman M L. Markov decision processes[M]. Handbooks in operations research and management science, 1990, 2: 331-434.

[12] Qiu C, Hu Y, Chen Y, et al. Deep deterministic policy gradient (DDPG)-based energy harvesting wireless communications[J]. IEEE Internet of Things Journal, 2019, 6(5): 8577-8588.