Dean&Francis

ISSN 2959-6157

The Application Status and Prospect of Deep Reinforcement Learning in the Smart Grid

Jiacheng Sui[1, *]

[1]School of Aerospace Engineering, Huazhong University of Science and Technology, Wuhan, China
*Corresponding author: u202211813@hust.edu.cn

Abstract:

Deep reinforcement learning has strong capabilities in data analysis, prediction, and autonomous learning, and it is highly compatible with the demand for big data applications in various aspects of the smart grid. Firstly, this paper summarizes the basic ideas of deep learning, introduces the basic principles and typical algorithms of deep reinforcement learning, and outlines its application characteristics. It reviews the current status of the application of deep reinforcement learning in the smart grid system in aspects such as fault diagnosis, load and new energy power prediction, and power dispatch. In view of the technical characteristics of deep reinforcement learning, combined with the various production links of the power system, an application framework of deep reinforcement learning technology in the power system is established. Finally, the application of deep reinforcement learning is prospected from aspects such as the operation control of multi-energy systems, system security analysis, fault diagnosis of flexible equipment, and privacy information security protection.

Keywords: Artificial Intelligence; Big Data; Deep Reinforcement Learning; Smart Grid.

1. Introduction

With the deep adjustment of energy structure and the large-scale integration of renewable energy, traditional power grids are facing unprecedented challenges, such as increased load volatility, decentralized energy distribution, and increased system operation complexity. In this context, smart grids have emerged. The smart grid integrates advanced sensing technology, communication technology, and information processing capabilities to achieve real-time monitoring, adaptive control, and efficient management of the power system, becoming a key support for achieving a clean, efficient, and reliable power supply system [1].

In order to better cope with the complex decision-making problems and dynamic environment in the smart grid, artificial intelligence technology, especially machine learning (ML), has been widely introduced in recent years. Machine learning can assist power grid systems in achieving various intelligent functions such as state estimation, load forecasting, and equipment fault diagnosis by learning and modeling historical data. However, traditional machine learning models still have certain limitations when dealing with decision optimization problems in high-dimensional dynamic environments.

Deep Reinforcement Learning (DRL), as an emerging research direction that combines deep learning and reinforcement learning, has shown great potential in the fields of intelligent control and optimization in recent years. DRL can learn optimal strategies through interaction with the environment without supervised data labels, especially suitable for complex tasks with high dynamics and uncertainties such as energy scheduling, electric vehicle charging and discharging management, and demand response in smart grids [2]. Therefore, exploring the current application status and future development directions of deep reinforcement learning in smart grids is of great significance for promoting the development of smart grid systems towards a more intelligent, autonomous, and efficient direction.

2. Deep Reinforcement Learning

2.1 Technical Overview

Deep reinforcement learning (DRL) is a method that combines reinforcement learning (RL) and deep learning, aiming to enable agents to learn strategies that maximize long-term rewards in unsupervised contexts through interaction with the environment. Traditional reinforcement learning methods, such as Q-learning and SARSA, rely on table form or shallow function approximation and are difficult to handle high-dimensional state spaces. Deep learning provides the possibility to solve high-dimensional decision problems by approximating complex functions through multi-layer neural networks. DRL approximates the strategy function, value function, or joint distribution of strategy and value through deep neural networks, thereby achieving effective solutions to complex decision-making problems.

The rise of deep reinforcement learning can be traced back to Mnih et al.'s groundbreaking work, Deep Q-Network (DQN), published in the journal Nature in 2015 [3]. This method has reached a level close to or even surpassing that of humans for the first time in multiple Atari games, marking the successful application of deep learning in reinforcement learning. Subsequently, numerous variant methods have been proposed, such as Double DQN, Dueling DQN, Prioritized Experience Replay, etc., aimed at improving the stability and sample efficiency of learning.

Another important type of method is Policy Gradient Methods, with representative algorithms including REINFORCE, Actor Critic, Proximal Policy Optimization (PPO), and Deep Deterministic Policy Gradient (DDPG). These methods are particularly suitable for control problems in continuous action spaces. In the PPO algorithm proposed by Schulman et al., the stability and convergence speed of training were effectively improved by updating the amplitude through constraint strategies [4].

With the deepening of research, DRL has gradually been applied to multiple complex systems such as energy systems, autonomous driving, and robot control, demonstrating good generalization ability and robustness. In the smart grid scenario, DRL is widely used in tasks such as energy storage optimization scheduling, electric vehicle charging strategies, and demand response management, and has extremely high application value.

2.2 Typical Algorithm

2.2.1 Discrete intelligent control strategy based on DQN

DQN (Deep Q-Network) combines traditional Q-learning with convolutional neural networks to output efficient optimal action strategies in complex state spaces, making it suitable for optimization tasks in discrete action spaces. Zhang et al. proposed a distributed electric vehicle charging control method based on DQN. By constructing a multi-agent environment and combining local state features for strategy learning, they achieved time of use charging optimization under dynamic electricity price changes, significantly reducing system peak load and user electricity costs [5]. This type of method has positive significance for the safe operation of the power grid during peak periods of electric travel.

2.2.2 Energy management under continuous control based on DDPG

DDPG (Deep Deterministic Policy Gradient) is a typical algorithm for handling continuous action space problems, which achieves efficient learning of high-dimensional control variables through the Actor Critic architecture and policy gradient method. Wang et al. introduced the DDPG algorithm into a multi-source system containing photovoltaic, energy storage, and electric vehicle loads when studying the energy management problem of household microgrids. They considered real-time electricity prices and load prediction errors, and obtained the optimal charging and discharging strategy through end-to-end training, effectively improving energy utilization efficiency and economy [6]. The advantage of DDPG in energy storage system scheduling is that it can naturally handle

ISSN 2959-6157

continuous variables such as charging and discharging power, avoiding the accuracy loss caused by action discretization.

2.2.3 Robust distributed scheduling optimization based on PPO

Proximal Policy Optimization (PPO) is a policy optimization method that balances stability and efficiency. Its introduced Trust Region update policy can avoid the problem of performance degradation caused by excessive policy updates. Xu et al. applied the PPO algorithm to the intelligent scheduling problem of multi energy systems, constructing a comprehensive energy system that considers the coordinated operation of wind power, photovoltaic power, and thermal energy. They achieved fast convergence and good generalization ability of scheduling strategies through a parallel policy update mechanism, verifying the strong adaptability of the PPO algorithm in multi scenario power grid environments [7]. This method is particularly suitable for practical power grid scenarios with uncertain inputs.

2.2.4 High entropy intelligent scheduling strategy based on SAC

Soft Actor Critic (SAC) encourages agents to fully explore in the early stages of policy learning by maximizing policy entropy, thereby enhancing robustness in complex environments. Han et al. introduced the SAC algorithm in the joint scheduling problem of wind power and energy storage. By constructing a Markov decision process and considering the uncertainty of wind power output and capacity constraints of the energy storage system, they designed an entropy regularization term to enhance the diversity and response flexibility of the strategy. The experimental results showed that the SAC strategy has better convergence and robustness compared to methods such as DDPG, and has good application prospects in scenarios with significant wind wave dynamics [8].

2.2.5 Collaborative application of multi-agent deep reinforcement learning in multi node systems

In actual power grid operation, there are multiple independent and autonomous energy entities (such as residential users, regional microgrids, etc.), and their interactive behavior is often difficult to capture by modeling with a single agent. Multi-Agent Deep Reinforcement Learning(MADRL) Collaborative optimization control between system levels can be achieved through multi-agent parallel learning and game coordination. Zhang et al. applied the MADRL method to user side intelligent load control, learning response strategies of different users through multiple agents to achieve refined management of electricity consumption behavior and optimization of the overall load curve of the power grid [9]. The MADRL method is applicable to multi-agent interaction problems such as demand response and regional power grid collaborative scheduling, and is currently one of the hot research directions in smart grid.

Based on the above, deep reinforcement learning technology is gradually moving from theoretical research to engineering practice in smart grids, demonstrating significant flexibility, scalability, and data-driven modeling capabilities. In the future, DRL can be combined with emerging artificial intelligence methods such as Graph Neural Networks (GNN), Federated Learning, and Transfer Learning to further enhance its generalization ability and safety reliability in multi-source heterogeneous, strong time-varying large-scale power systems, and promote the development of smart grids towards a more intelligent, autonomous, and low-carbon direction.

3. The Current Application Status of Deep Reinforcement Learning in Smart Grid

3.1 Fault Diagnosis of Power Equipment and Systems

With the increasing complexity of the power grid structure, traditional rule-based diagnostic methods have shown certain limitations in dealing with fault diversity, speed, and concealment. Deep reinforcement learning provides a dynamic and adaptive alternative that continuously optimizes decision strategies during training, achieving real-time and accurate fault recognition and response.

In the fields of transmission line protection, distribution network fault location, and circuit breaker anomaly detection, researchers have attempted to construct strategy networks based on Deep Q-Network (DQN) or Near End Policy Optimization (PPO) to replace fixed expert systems.

In higher dimensional systems, modeling based on graph structure information is gradually becoming a trend. Research has shown that transforming the power grid topology into a graph structure input graph convolutional network (GCN) and embedding it into a DRL framework can achieve stronger perception ability and broader fault generalization performance. This type of method is particularly suitable for scenarios with flexible structures but strong information dependence, such as complex distribution systems and flexible transmission systems.

3.2 Load and New Energy Power Forecasting

The operation of the power system relies on accurate prediction of future load demand and renewable energy output. Especially in the context of the increasing penetration rate of new energy, external variables such as wind speed and solar irradiance lead to strong fluctuations and high nonlinearity in output, seriously affecting system scheduling and stable operation.

Traditional prediction methods, such as regression analysis, ARIMA, SVR, etc., rely on a large number of historical samples for training and are difficult to cope with the evolution of system behavior after sudden events or policy adjustments. Deep reinforcement learning provides a new path for optimizing long-term prediction accuracy. Unlike static prediction models, DRL models can continuously update strategies from real-time feedback during the prediction process, achieving rolling prediction and dynamic correction.

In terms of joint prediction, for multi regional and multi energy interconnected systems, DRL can autonomously determine the prediction weights of key regions based on real-time data streams, forming an overall prediction strategy. Such methods have important potential in the development of regional energy internet and new power system.

3.3 Power Dispatch

As one of the core tasks of power grid operation, power dispatching is often based on deterministic or stochastic optimization algorithms, such as linear programming (LP), dynamic programming (DP), genetic algorithm (GA), etc. However, these methods generally rely on precise modeling of the system and are difficult to adapt to real-time changes and large-scale variable optimization in complex situations.

DRL solves the optimal decision-making problem of "state action benefit" in power dispatch problems through strategy learning, and has strong model adaptability. This type of method is applicable to multiple sub fields such as load response, microgrid energy management, and optimization of electric vehicle (EV) charging and discharging.

Some studies focus on using reinforcement learning for dynamic pricing strategy formulation, thereby guiding user behavior to adjust the load curve and achieve proactive demand response. For example, the pricing model using the Actor Critic framework can adjust the optimal electricity price based on real-time load changes and user elasticity, guide load to shift to the valley, and effectively alleviate peak load pressure.

Reinforcement learning has also been successfully applied in the charging scheduling of electric vehicles. By learning user travel patterns, electricity price trends, and grid carrying capacity, a reasonable charging schedule is constructed to achieve resource sharing and system load balancing.

3.4 Summary of Application Status

In recent years, the research on deep reinforcement learning (DRL) in smart grids has continued to heat up, and related applications have gradually been launched in multiple key fields, showing good prospects and practical value. Currently, DRL has gradually achieved practical results in power equipment status diagnosis, load and new energy power prediction, power dispatch optimization, and is becoming an important technical support for promoting the intelligent transformation of the power grid. Compared with traditional control or prediction models, DRL has high adaptability and strong generalization ability, and can continuously optimize strategies in complex state spaces through continuous interaction with the environment. This makes it particularly suitable for problem scenarios with strong dynamics, uncertainty, and nonlinear characteristics in smart grids, such as distributed energy scheduling, electric vehicle group charging and discharging management, multi regional collaborative operation, and other tasks.

Current research indicates that deep reinforcement learning has the ability to integrate multiple deep network structures, such as embedding graph neural networks (GNNs), convolutional neural networks (CNNs), or long short-term memory networks (LSTMs) into policy networks, which can effectively enhance the modeling accuracy of power grid topology, temporal sequence features, and device relationships. In addition, Multi Agent DRL (Multi Agent DRL) has shown significant advantages in scenarios such as regional autonomous control, energy storage system management, and virtual power plant game strategies as a way to address the multi-source distributed characteristics of power grid systems. Its distributed collaborative learning mechanism can gradually approach the global optimal solution of the system while ensuring local optima, thus meeting the operational requirements of modern power grids for the coexistence of "autonomy, collaboration, and optimization".

However, the application of deep reinforcement learning still faces many challenges. On the one hand, it has a high dependence on the number of training samples and the number of environmental interactions, and in actual deployment, there is often a lack of sufficient real running data. Relying on simulation environment training can introduce bias and affect the transferability of the strategy. On the other hand, the "black box" attribute of DRL models still results in insufficient interpretability and security, lacking clear decision logic outputs and behavior

ISSN 2959-6157

boundary control mechanisms, which pose significant obstacles in the critical infrastructure field of smart grids. In addition, most existing research focuses on single objective optimization, and further theoretical and algorithmic breakthroughs are needed to achieve multi-objective coordination between ensuring operational economy, safety, and carbon reduction goals.

Looking towards the future, research on deep reinforcement learning in smart grids is expected to further advance in the following directions. Firstly, at the model level, combining DRL with advanced structures such as graph neural networks, attention mechanisms, and meta learning can enhance its modeling capability and policy robustness in large-scale power grid systems. Secondly, at the data level, introducing a federated learning framework to achieve cross regional collaborative training while protecting data privacy and improving model generalization performance is also a feasible path. In addition, for the implementation of the project, it is necessary to strengthen research in constraint learning, safety learning, strategy verification mechanisms, and other aspects to ensure its stable operation and safety guarantee in the real power grid. At the policy level, promoting the development of standards and industry norms for AI technology in smart grids will also provide institutional guarantees for its comprehensive implementation.

The application of deep reinforcement learning in smart grids is gradually advancing from theoretical verification to engineering practice. Through its ability to learn strategies and optimize decisions, it is expected to play an important role in multiple aspects such as power grid operation scheduling, load regulation, and resource optimization. In the future, with the continuous maturity of relevant algorithms and the increasingly perfect system architecture, DRL is expected to become a key technical pillar for achieving efficient, intelligent, and autonomous operation of smart grids.

4. Conclusion

With the continuous optimization of energy structure and the gradual construction of intelligent power grid system, the application prospects of Deep Reinforcement Learning (DRL) in smart grid are becoming increasingly broad. As an intelligent technology that integrates perception, decision-making, and control, DRL has demonstrated significant advantages in dealing with complex, high-dimensional, and nonlinear environmental problems. However, current research mostly focuses on theoretical verification and simulation, and its engineering deployment in real power systems still faces many challenges. In order to further promote the practical and large-scale development of this technology, future research can be carried out from the following aspects.

Firstly, the security and robustness of the algorithm should be enhanced to ensure its reliability in critical tasks of the power system. The smart grid has extremely high requirements for operational safety, and DRL models must have clear safety boundaries and fault-tolerant mechanisms in practical deployment. Therefore, introducing frameworks such as Safe RL and Constrained RL can help meet the physical and security constraints of system operation during policy learning, thereby ensuring the controllability of decision-making behavior and the stability of the system.

Improving the generalization and transfer learning abilities of DRL models is the key to achieving their engineering generalization. The existing methods generally rely on data and simulation environments in specific scenarios, lacking sufficient transferability and adaptability. In the future, cross task learning mechanisms such as federated learning and meta learning can be used to achieve knowledge sharing between different power grid regions while ensuring data privacy, significantly improving the generality and training efficiency of the model.

Interpretability is another core issue in the application of DRL to practical smart grid systems. Due to the complex structure of its deep neural network and the difficulty in intuitively analyzing the policy generation process, it lacks sufficient transparency and trustworthiness. Therefore, constructing interpretable DRL models, introducing attention mechanisms, and conducting strategy visualization research will help improve the understanding and acceptance of system intelligent behavior by operation and maintenance personnel, and enhance the controllability and auditability of the model in practical engineering.

In addition, multi-agent deep reinforcement learning (Multi Agent DRL), as an effective means to address the strong characteristics of distributed and heterogeneous power grids, will play a more important role in the future. By building a regional autonomous intelligent agent group, the unity of local control and global coordination can be achieved, which can significantly improve the response speed and overall operational efficiency of the power grid system. However, multi-agent systems still face complex technical challenges in communication coordination, collaborative decision-making, and policy convergence, and there is an urgent need for further in-depth research in system architecture, interaction protocols, and distributed optimization algorithms.

We should strengthen the standardization and interdisciplinary integration of deep reinforcement learning in the field of smart grids. Future research not only needs to focus on algorithm level optimization, but also needs to collaborate with experts from multiple fields such as power systems, artificial intelligence, systems engineering, and control theory to establish unified data interfaces, evaluation standards, and deployment specifications. In addition, policy guidance and industry collaboration will also become important guarantees for promoting technology implementation, especially in areas such as data security, network communication, and equipment collaboration, corresponding standards should be formulated to ensure the trustworthy, controllable, and sustainable development of artificial intelligence technology in the smart grid.

In summary, deep reinforcement learning technology, as an algorithm framework with highly intelligent and adaptive capabilities, is expected to play a key role in the development process of smart grids. In the future, through the collaborative promotion of algorithm innovation, platform construction, standard construction, and engineering verification, DRL will play an irreplaceable role in promoting the development of smart grids towards higher-level autonomous operation and collaborative optimization.

References

[1] X. Fang, S. Misra, G. Xue, and D. Yang, “Smart grid—The new and improved power grid: A survey,” IEEE Communications Surveys & Tutorials, vol. 14, no. 4, pp. 944–980, 2012.

[2] T. Wei, Y. Wang, and Q. Zhu, “Deep reinforcement learning for building HVAC control,” in Proc. 54th Annu. Design Autom. Conf. (DAC), 2017, pp. 1–6.

[3] V. Mnih *et al.*, “Human-level control through deep reinforcement learning,” Nature, vol. 518, no. 7540, pp. 529–533, 2015.

[4] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, “Proximal policy optimization algorithms,” arXiv preprint, arXiv:1707.06347, 2017.

[5] K. Zhang, A. V. Vasilakos, and Y. Liu, “A decentralized deep reinforcement learning approach for charging multiple electric vehicles,” IEEE Trans. Ind. Inform., vol. 15, no. 12, pp. 6520–6529, 2018.

[6] Q. Wang, Y. X. Zhou, and F. S. Li, “Deep reinforcement learning-based energy management for a residential microgrid with renewable energy and electric vehicles,” Energy, vol. 196, p. 117081, 2020.

[7] Z. Xu, J. Zhang, and L. Wang, “Intelligent dispatch of integrated energy system based on PPO reinforcement learning algorithm,” Applied Energy, vol. 310, p. 118492, 2022.

[8] L. Han, Y. J. Chen, and F. Gao, “Soft Actor-Critic based scheduling strategy for wind-storage systems under uncertainty,” Energy Reports, vol. 7, pp. 4123–4134, 2021.

[9] W. Zhang, H. Li, and H. W. Liu, “Multi-agent deep reinforcement learning for smart demand response management in microgrids,” IEEE Trans. Smart Grid, vol. 12, no. 4, pp. 3131–3142, 2021.