

基于强化学习的新型电力系统优化策略应用综述

闫正义, 赵康, 王凯*

(青岛大学电气工程学院, 山东省 青岛市 266071)

Review of Application on Optimization Strategies for New-Type Power System Based on Reinforcement Learning

YAN Zhengyi, ZHAO Kang, WANG Kai*

(College of Electrical Engineering, Qingdao University, Qingdao 266071, Shandong Province, China)

摘要:【目的】随着电力系统向更高层次的智能化和自动化演进, 强化学习(reinforcement learning, RL)作为人工智能领域的一项关键技术, 在电力领域的智能化发展方向上展现出广阔前景。完善RL在电力领域的应用研究方案, 对于深入挖掘其在电力系统运行、控制和优化等方面的潜力至关重要。为此, 分析了RL在实际电气应用中的效能表现, 并展望了未来可能的研究方向, 以期为电力系统智能化进程提供助力。【方法】对RL在各类电气领域的实际应用进行了综述。系统性地介绍了RL的基本原理和标志性算法, 详细探讨这些算法如何被应用于新型电力系统领域的实际问题中。对各研究中主流的RL算法进行归类, 并对在这些算法中进行的结构化改进进行优缺点分析。【结果】相比于传统算法, RL显著提升了新型电力系统的智能化水平, 并在多个应用场景中取得了显著成效, 特别是在应对系统复杂性和不确定性方面表现出色。然而, 尽管有诸多成功案例, 但目前该领域仍存在一些亟待解决的问题, 比如计算成本高、训练时间长、泛化能力不足等。【结论】RL为新型电力系统的智能化提供了新的解决方案, 然而, 要实现大规模应用, 还需要克服一系列技术和实践上的挑战。研究成果可为电气工程领域的研究者和实践者提供参考和启示。

关键词: 新型电力系统; 强化学习(RL); 深度强化学习(DRL); 智能电网; 优化策略; 能源管理; 态势感知; 优化调度; 人工智能(AI)

ABSTRACT: [Objectives] As power systems evolve toward higher levels of intelligence and automation,

reinforcement learning (RL), a key technology in artificial intelligence, shows great potential in the intelligent development of the power sector. Enhancing research methods for RL applications is crucial for fully exploring its potential in power system operation, control, and optimization. Therefore, the performance of RL in practical electrical applications is analyzed, and the possible research directions in the future are prospected, so as to provide assistance for the intelligent transformation of power systems. [Methods] This study provides a systematic review of RL applications across diverse fields of electrical engineering. It systematically introduces the fundamental principles and landmark algorithms of RL, detailing how these algorithms are applied to address practical problems in new-type power system. The study categorizes mainstream RL algorithms in current research and analyzes the advantages and disadvantages of structural improvements made to these algorithms. [Results] Compared to traditional algorithms, RL significantly enhances the intelligence level of new-type power system. It achieves remarkable success in various application scenarios, particularly in addressing system complexity and uncertainty. However, despite many successful cases, several urgent issues still exist in this sector, such as high computational costs, long training times, and limited generalization abilities. [Conclusions] Reinforcement learning provides novel solutions for the intelligent development of new-type power system. However, achieving large-scale application still needs to overcome a series of technical and practical challenges. This study provides references and insights for researchers and practitioners in electrical engineering.

KEY WORDS: new-type power system; reinforcement learning (RL); deep reinforcement learning (DRL); intelligent grid; strategy optimization; energy management; situational awareness; optimized scheduling; artificial intelligence (AI)

基金项目: 国家自然科学基金项目(12374088, 51877113); 山东省高等学校青年创新技术项目(2022KJ139)。

Project Supported by National Natural Science Foundation of China (12374088, 51877113); Youth Innovation Technology Project of Higher School in Shandong Province (2022KJ139).

0 引言

随着电力系统逐渐融合更多的可再生能源(如太阳能和风能),其管理的复杂性与日俱增^[1-2]。传统电力系统在应对能源供应的波动性和间歇性方面存在不足,这不仅增加了运行难度,也对电网稳定性提出了新的挑战^[3-4]。

近年来,虚拟电厂^[5-8]作为一种创新概念,通过整合分布式能源资源,如分布式电源^[9-12]、储能系统^[13-17]、可控负荷^[18]及电动汽车^[19-23]等,为上述问题提供了一种有效的解决方案。它利用先进的数字通信技术,实现了资源的统一调度与优化配置,并在功能上模拟传统电厂,用以参与电力市场和电网运行^[24-26]。然而,虚拟电厂的发展也带来了新的不确定性,尤其是在电能协调管理和灵活适应变化方面提出了更为严格的技术要求。为确保可持续且稳定的电力供应,探索智能化的管理方法成为当务之急^[27]。

在此背景下,强化学习(reinforcement learning, RL)作为人工智能(artificial intelligence, AI)领域的一种高效决策机制,因其能够自适应环境变化和优化长期决策而备受关注^[28]。不同于依赖预设规则的传统方法,RL通过与环境互动学习最优策略。这种基于大量经验优化自身模型,而不需要具体计算模型的学习方法使得强化学习在应对电力系统管理中的复杂性和不确定性方面具有独特的优势^[29]。

深度强化学习(deep reinforcement learning, DRL)结合了强化学习的决策能力与深度学习的寻优能力,在新型电力系统的策略优化中扮演着关键角色。它不仅增强了系统的稳定性和可靠性,使电网能够更好地应对可再生能源带来的波动,还实现了能源的有效分配和利用。此外,它促进了分布式能源资源的整合,实现了更多样化的能源形式的高效接入和协同管理,推动了电力系统的智能化转型,提高了系统整体运行效率和可持续发展能力。

本文旨在探讨RL、DRL应用于新型电力系统领域的潜力,特别是在解决虚拟电厂所面临的挑战方面。通过评估当前的研究进展,分析DRL在

该领域应用的优势与局限性,提出未来的研究方向。此外,本文还展现了研究的理论基础和实验方法,以及预期成果对于提升电力系统灵活性和效率的意义,以期促进新型电力系统的发展。

1 强化学习原理

强化学习是机器学习的分支领域,通过智能体(Agent)与环境的交互,学习最优行为策略。作为一种自主学习的方法,智能体能够在没有明确的正确行动指导的前提下,从其历史经验数据中进行学习。

1.1 马尔可夫决策过程

在强化学习中,通常将环境建模为马尔可夫决策过程(Markov decision process, MDP)。MDP在强化学习领域中非常重要,它是强化学习的理论基础,通过形式化智能体与环境的交互过程,为问题建模和算法设计提供了基本的框架。

MDP流程如图1所示。MDP中环境与智能体交互流程核心要素可明确为状态、行动、状态转移概率、奖励函数、策略5个部分。具体如下:

- 1) 状态为系统可以存在的所有可能的情况或配置。
- 2) 行动为在给定状态下决策者(通常称为智能体)可以选择的所有可能行动。
- 3) 状态转移概率为行动被执行后系统从当前状态转移到另一个状态的概率。
- 4) 奖励函数为智能体执行某个行动并由此引起状态转变时所获得的即时回报。
- 5) 策略为从状态到动作的映射,可以通过调整策略来进行智能体的动作控制。

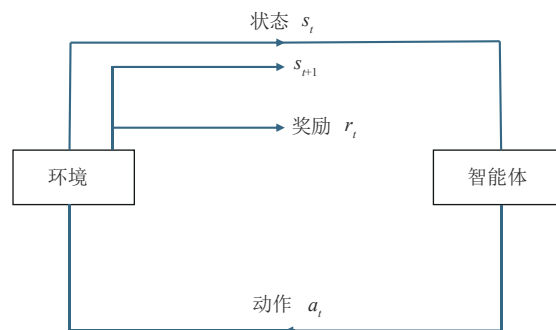


图1 MDP流程示意

Fig.1 Schematic diagram of MDP process

包含奖励的有限 MDP 可以由五元组 $\{S, A, P, R, \pi\}$ 来描述, 其中, S 为该过程的状态集合; A 为该过程中所有动作的集合; P 为状态转移概率, 即由当前状态通过智能体执行动作进入到下一状态的概率; R 为智能体在决策过程中进行状态转移时获得的即时奖励; 策略 π 为智能体所有动作的概率密度函数。

决策的迭代过程如图 2 所示。通过智能体状态与不同的动作组合可形成不同的结果, 该行为可描述为状态-动作对, 即 (s_t, a_t) , 其中 s_t 为时隙 t 的状态, $s_t \in S$, a_t 为时隙 t 的动作, $a_t \in A$, $t=1, 2, \dots$ 。由于不同的状态-动作会对环境产生不同的影响, 也就是对应于不同的状态转移概率, 从而导致下一时隙 $t+1$ 的状态 s_{t+1} 的变化。同时, 智能体获得迭代本次时隙 t 的奖励 r_t 。

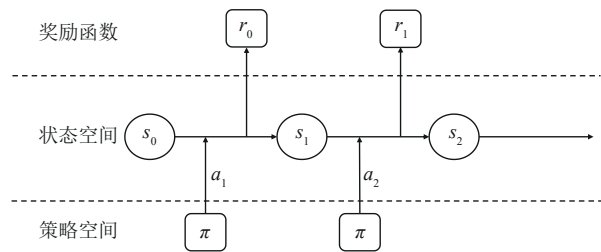


图2 MDP迭代信息示意图

Fig.2 Schematic diagram of MDP iterative information

强化学习的目标就是通过优化决策模型, 使得MDP获得最大累积折扣回报 U_t' 。

$$U_t' = \sum_{i=0}^{\infty} \gamma^i r_{t+i} \quad (1)$$

式中: i 为当前的时间步; r_{t+i} 表示时隙为 $t+i$ 时实际所获得的奖励; γ 表示折扣因子, 为超参数, $\gamma \in (0, 1)$, 当 γ 越接近于 0 时, 决策过程中越重视短期收益, 反之, 则越重视长期回报。

在决策过程中, 由于策略函数 π 带来的动作 a 与状态转移概率 P 带来状态 s 的不确定性概率结果, 实际上无法获得累积折扣回报 U_t' , 因此, 引入动作价值函数 Q , 通过对未来时刻的不确定性进行期望化, 以便于进行回报的具体量化分析。

为了求得当前可获得的最大回报策略, 将 Q_π 作为在当前状态、给定策略 π 时的动作价值函数, 令 Q^* 为对比所有策略后的最优动作价值函数, 则

有 $Q^* = \max_{\pi} Q_\pi$ 。

1.2 优化策略

在强化学习中, 优化策略是指通过调整智能体的行为规则(即策略 π)以最大化长期累积奖励的过程。它定义了智能体在给定状态下选择行动的概率分布, 旨在找到一个能够在所有可能状态下所能获得的最大累计奖励, 因此智能体需要在探索新动作与利用已知最佳动作之间找到平衡。最终, 优化策略确保智能体能够在复杂环境中做出最优决策, 实现最大化的累积奖励。

1.3 学习机制

学习机制是强化学习原理的核心部分, 它决定了智能体如何通过与环境的交互来优化策略。

1.3.1 价值学习与策略学习

RL 算法根据学习目标的不同, 可以简单划分为价值学习(value-based reinforcement learning, VRL)算法与策略学习(policy-based reinforcement learning, PRL)算法 2 类。

VRL 可直接通过学习价值函数 Q , 来获得当前策略执行后的价值函数 Q_π 。然而, 对于难以直接求得价值函数 Q 的高维复杂环境, 可通过在深度 Q 网络(deep- Q -network, DQN)中引入深度神经网络, 使得超参数权重 w 能够进行自适应调整, 从而将与 w 、 s (状态)、 a (动作)参数相关的价值函数简化为近似于与 s 、 a 相关的函数, 以实现复杂环境下的强化学习决策。

相比于 VRL 学习价值函数 Q 的做法, PRL 选择对策略函数 π 进行优化改进, 通过对策略函数进行梯度上升迭代, 从而获得极大策略函数, 最终获得最大回报, 而当面临复杂情况时, 可以通过引入网络偏置与权重参数 θ 直接调整策略函数, 从而获得近似策略函数 π 。

总的来说, PRL 可直接优化策略函数(如神经网络), 然后输出动作或动作概率分布, 而无需依赖价值函数, 而 VRL 通过估计状态或动作的长期价值(如 Q), 间接优化策略(通常通过贪心选择动作)。

1.3.2 混合学习

除上述 2 类基本类别外, 还有将 2 类方法相互

结合的混合学习方法,如演员-评论家(actor-critic, AC)算法。该算法结合了VRL与PRL的学习机制,分别使用2个神经网络对策略函数 π (actor)与价值函数 Q (critic)进行模拟,其中actor(策略网络)生成动作, critic(价值网络)评估动作优劣,这种算法模式提供了低方差梯度信号,使得学习过程更加稳定和高效。混合学习能够有效地处理连续动作空间的决策问题,并且相对于单纯的策略梯度或者价值学习方法,混合学习方法收敛速度更快,方差更小,因此,被广泛应用于各种复杂的环境中,如机器人控制、游戏AI、自动驾驶等。

1.3.3 多智能体强化学习算法

多智能体强化学习(multi-agent reinforcement learning, MARL)是一种在多个智能体互动的环境中应用RL的框架。在MARL中,每个智能体根据自己的观察,采取行动并接受个体奖励,大多数情况下,智能体的目标是最大化自己的累积奖励。由于多智能体环境的状态由所有智能体共同决定,即每个智能体的动作不仅会影响环境的状态,还会间接影响其他智能体的策略和决策过程。这些智能体可以是合作的、竞争的或者这两者的混合,根据智能体之间的不同关系, MARL又派生出了多种不同目的的学习机制,从而衍生出不同的多智能体强化学习算法。

独立 Q 学习(independent Q -learning, IQL)方法是常见的有合作关系的MARL方法,该方法将单智能体强化学习方法直接套用在多智能体系统中,即每个智能体把其他智能体都当作环境中的因素(无视其他智能体的存在),而自身仍然按照单智能体学习方式并通过与环境交互来更新策略。虽然算法逻辑简单,但计算效果通常受限于环境的非静态性。

minimax Q -learning方法可用于有完全竞争关系的多智能体中,此时每个智能体需要考虑在其他智能体做出最坏动作(使自身回报最差的动作)的情况下,自身做出的动作仍能获得最大的回报。其表达式为

$$V^*(s) = \max_{\pi} \{ \min_{a'} [\sum_a Q^*(s, a, a') \cdot \pi(s, a)] \} \quad (2)$$

式中: $V^*(s)$ 表示在状态 s 下,智能体通过最优策略能获得的最大最小化保障值(即对抗最坏情况下的最优收益); $Q^*(s, a, a')$ 表示在状态 s 下,当前智能体采取动作 a ,其他智能体采取动作 a' 时的预期长期回报。

在多智能体混合关系中,可以将minimax Q -learning方法扩展为Nash Q -learning方法。该方法通常用于多人一般和博弈模型中, Nash Q -learning的核心是通过对每个智能体的 Q 值进行适当更新,从而使得多智能体系统趋于一个Nash均衡。与标准 Q -learning只涉及单一智能体的情况不同, Nash Q -learning需要处理多个智能体的交互和可能产生的策略冲突。但该算法的缺点是,当智能体增加时,存储信息量会以指数形式增长,因此,该算法适用于小规模智能体环境。

此外,多智能体深度确定性策略梯度(multi-agent deep deterministic policy gradient, MADDPG)也是一种用于解决多智能体强化学习问题的有效算法。它是对单智能体深度确定性策略梯度(deep deterministic policy gradient, DDPG)算法的扩展,能够有效解决多智能体并行学习时产生的策略交互问题。MADDPG采用集中式训练与分散式执行的框架,使每个智能体在学习过程中都能充分考量其他智能体的策略行为。此外, MADDPG的critic网络会综合其他智能体的动作信息,以更好地评估当前智能体的动作对整个系统的影响。总体而言, MADDPG可以实现多智能体之间的协同学习,从而使得整个系统达到更好的性能。

2 强化学习在新型电力领域的应用

2.1 总体应用情况

RL作为一种能够通过环境交互实现自主决策的机器学习方法,近年来在新型电力系统中的应用日益广泛,覆盖了调度优化、能源交易、能源转换、需求响应、电压控制等多种应用领域。因此,对其主要应用方向及典型案例进行分析,能够将RL的理论优势与新型电力系统的实际需求相结合。此外,明确RL在电力系统中的应用场景,

还有助于挖掘其在复杂动态环境下的优化潜力，推动电力系统控制、调度、交易等领域的技术创新，为电力系统的智能化发展提供科学依据和实践参考。

强化学习算法在各个新型电力系统领域的应用情况^[30-42]如表1所示。

表1 强化学习算法在各新型电力系统领域应用情况

Tab. 1 Applications of reinforcement learning algorithms in new-type power system

来源	解决问题	系统形式	求解算法	优化目标
文献[30]	能源交易	风电	Rainbow	提高收敛速度,减少所需采样次数,提高平均收益
文献[31]		微电网	MATD3	降低微电网的运行成本
文献[32]		微电网	DA-MAPPO	降低微电网的峰值负荷,同时降低各用户的用电成本
文献[33]		火电	MADDPG	提高市场的整体效率,最佳化整体碳排放收益率
文献[34]		电力市场	DDPG	降低市场风险溢价
文献[35]	多目标经济调度	热-电耦合	孪生延迟DDPG/多目标奖励函数	大幅降低优化时间,提高鲁棒性
文献[36]	需求响应	智能电网	Q -learning	减少总能耗,提高买卖双方收益
文献[37]		用户侧	PPO2	提高收敛性和数据利用率
文献[38]	电压控制	配电网	MATD3PG	降低网损
文献[39]		配电网	hyper Q -network	最小化长期预期电压偏差
文献[40]	频率控制	光伏-微电网	Q -learning	具有更好的频率调控作用
文献[41]		孤岛微电网	Q -learning	具有较好的调频效果与适应性
文献[42]	能量管理	微电网	DQN	高负载率(高于50%)

注:MATD3表示多智能体双延迟深度确定性策略梯度算法;DA-MAPPO表示分布式注意力多智能体近端策略优化算法;PPO2表示近端策略优化算法第2版;MATD3PG表示多智能体双延迟深度确定性策略梯度;hyper Q -network表示超 Q 学习算法。

2.2 能源管理

对于能源管理领域，优化能源生产和分配，以实现更高效的能源利用和供应链管理是强化学习的广泛应用方式。文献[43]引入DDPG算法，结合低保真分析模型的知识辅助，提出了一种协同风电场控制的知识辅助深度确定性策略梯度(knowledge-assisted deep deterministic policy gradient, KADDPG)算法。KADDPG将知识辅助的强化学习框架与DDPG相结合，通过协调平衡风电场上下游风力发电机的输出功率，生成优化风电场控制策略，以实现风电场总发电量的最大化。文献[44]提出了一种结合联邦学习和DRL的超短期风电预测方案(federated deep reinforcement learning, FedDRL)，该方案采用DDPG算法作为基本预测模型，提高了预测精度；通过引入联邦学习算法，将DDPG预测模型集成到联邦学习框架中，保护了敏感的个体隐私数据。文献[45]提出了一种基于复杂深度强化学习的最优策略学习算法，旨在降低可再生能源的不确定性，实现储能系统的实时最优规划。仿真结果表明，采用该算法不仅成功实现了净利润最大化和系统稳定性提升的目标，还展现了深度强化学习在应对可再

生能源高度不确定性时的显著抑制能力。

综上，对于风力发电等新能源系统，强化学习多应用于能源管理，通过帮助优化发电设备的运行策略，实现能源产量最大化、减少系统的不稳定性与优化出力；而其在出力预测等研究方向的应用相对较少。这是因为，强化学习的高度决策动态性与环境适应性使得其能够更好地应对能源形式更加灵活多变的新能源环境，处理好新能源系统管理的高复杂度和不确定性。

2.3 优化调度

强化学习可以用于优化电力网的运行和调度。电力系统的运行涉及复杂的调度和控制问题，而强化学习可以通过自主学习和智能决策，提高电力系统的运行效率和稳定性。文献[46]结合双层深度 Q 网络(double-DQN, DDQN)算法，以最小化24 h微电网电力成本为目标，在满足微电网系统电压偏差等约束条件下，训练得到能量调控优化策略。文献[47]采用DDPG算法搭建了一种同时将微电网的经济性和低碳性作为优化目标的调度模型。文献[48]在时序差分 Q -learning算法基础上引入 n 步自举法，提出基于 n 步 Q -learning算法的风电抽水蓄能日随机优化调度方

法。相比于单步 Q 学习方法,该算法收敛速度更快、求解时间更少。综上,目前强化学习多成熟应用于规模较小、内部结构和运行模式更为简单的微电网系统。

对于大电网系统而言,由于其通常包含大量的发电站、输电线路和负荷节点,且这些组件之间的相互作用关系非常复杂,使用强化学习进行系统建模往往面临计算量庞大、复杂度高的挑战。因此,如何克服强化学习在大电网系统中的应用难题,将成为优化调度领域未来发展的一个重要方向。

此外,强化学习还可以用于优化电力市场的交易和定价,从而实现经济效益的最大化。文献[33]构建了基于MADDPG算法的竞价策略模型,分析了不同火力发电商组合的竞价差异化策略,优化了多主体报价报量策略,引导了火力发电商采取合理的竞价方式以提高市场效率。文献[34]提出一种基于深度强化学习的电力市场量价组合竞价策略分析方法,并采用DDPG方法求解,最后,通过算例验证了所提方法的鲁棒性和有效性。

2.4 电网电压控制与频率控制

电压控制可确保电力系统中各节点电压在规定的范围内,以保障系统的安全和稳定运行。传统的电压控制方法包括PID控制、有功和无功功率调节等,但这些方法在面对电网规模增大和电网结构变化时,可能调节不够灵活或响应不够快速。强化学习可以通过智能算法学习电网的动态行为,自动调整控制策略以应对复杂变化的电网环境。通过强化学习算法,可以自动调节变压器的档位、调整发电机的无功输出或者协调分布式能源的接入和调控,实现更精确和高效的电压控制。

为解决自主电压控制(autonomous voltage control, AVC)问题,文献[49]提出了基于DQN和基于DDPG两种最先进的DRL算法,制定了AVC策略,以支持电网运营商做出有效和及时的控制行动。文献[38]在单智能体算法的基础上,提出一种基于MATD3PG的有源配电网实时电压控制策略,该策略将参与决策的各个光伏逆变器看作独立的智能体,实现了电压控制过程的最优出力

调控。文献[50]将实时电压控制问题转换为马尔科夫博弈过程(Markov game process, MGP),并构建了多智能体模型,采用离线训练-在线运行的方法进行模型训练,以双延迟DDPG算法的电压控制策略来控制电压,结果表明,该策略减少了累计误差,并使得目标函数平滑。

频率控制可以保障电力系统频率稳定,对电力系统的安全和可靠运行至关重要。传统的频率控制方法依赖于既定的频率恢复策略和备用发电容量调节,这在新兴的电网结构中,尤其是在可再生能源大量接入的情况下,可能造成响应不够及时或控制不精确的问题。通过强化学习进行频率控制可以帮助系统在不确定性和变化性较大的环境下作出更优的决策。强化学习模型能够实时学习和更新策略,优化频率控制的响应速度和效果,控制储能系统快速释放或吸收能量以补偿瞬时的频率波动,或者调节可再生能源的输出以适应负载变化。

文献[40]采用强化学习控制策略进行功率分配并动态调整各光储虚拟同步发电机(virtual synchronous generator, VSG)的输出功率,实现了多光储VSG的频率协调控制。文献[41]设计了一种利用 Q 学习算法进行微电网频率调节的控制方法,此方法对信息精确度的要求低,且可以针对不同场景进行学习,具有较好的适应性。文献[51]提出一种基于协同奖励函数的多目标强化学习(top- Q multi-objective reinforcement learning, TOPQ-MORL)智能频率控制策略,该策略构建了计及多维度频率控制性能评价标准的协同奖励函数,实现了多维度频率控制性能标准在时间尺度上的配合评价。

综上,在电压控制与频率控制领域,RL、DRL多应用于接入大量光伏电源的配电网中,为电压和频率控制提供智能解决方案。通过动态调整无功功率补偿、优化光伏逆变器操作及智能管理储能系统,可实现VSG控制策略与分布式能源资源的协同优化调度,提升系统的等效惯量支撑与动态阻尼特性,确保电压与频率的稳定。结合电压频率控制领域的应用形式和能源系统类型可以看出,通过强化学习可以提升配电网应对光伏

电源间歇性挑战的能力，实现更加灵活、可靠的电力供应。

2.5 态势感知

态势感知不仅需要对当前的网络状态有准确的理解，还需要预测未来的变化趋势和潜在的风险。电网态势感知涉及到对电力系统整体安全状态的实时监控和评估，包括对潜在的故障、攻击或其他威胁的识别。强化学习能够建立动态仿真模型和实时预测机制，显著提升对电网运行状态的认知与把控能力。通过与传感器网络和实时监控系统的集成，强化学习模型可以持续学习和更新，实现复杂环境中潜在安全威胁的快速识别和响应，例如，可以通过基于实时数据构建的模型预测电网可能的超载情况、自动调整电网运行参数或预警操作人员进行干预。

文献[52]提出了一种基于深度强化学习的虚假数据注入攻击(false data injection attacks, FDIAs)检测方法，解决了现有强化学习检测方法中状态特征提取不够精确的问题。该方法在无模型深度强化学习检测算法的基础上增加了一种注意力机制，用以提取状态特征，使状态更具代表性和可区分性，从而改进了强化学习框架，使其能更快更准确地检测攻击。文献[53]基于图神经网络易于处理电网拓扑结构的优势，引入了强化学习自适应学习的特点，提出了基于强化学习的图神经网络算法，实现了智能体能够自主、动态地选择合适的风险指标的目标，提高了新型电力系统的风险态势感知能力。文献[54]将强化学习引入智能电网领域，提出了一种基于MADDPG边缘计算范式的深度强化学习算法，用于智能电网的安全态势感知(security situational awareness, SSA)。该算法充分发挥了边缘智能的优势，有效缓解了电力云主站模式下因大量异构电力终端而导致的SSA复杂化，以及故障信息传递延迟的问题。其结构设计使系统能够在运行过程中实时监控故障信息，在攻击发生前自动识别并切断潜在威胁，同时自动修复故障，这确保了电网系统的稳定运行，为智能电网提供了可靠的实时保护。

综上，强化学习能够有效应对复杂动态环境

中的不确定性和对抗性威胁，十分适用于智能电网的安全态势感知方面。强化学习支持分布式与局部化决策，能够在多个相对独立的子系统中实现最优保护策略，同时促进整体系统的协同工作。此外，它在复杂环境中可以高效探索不同策略，提前识别潜在安全威胁并增强系统的恢复能力，减少了人为干预带来的风险。这些特性使强化学习成为提升智能电网安全性能、确保其稳定可靠运行的关键工具。

3 展望

尽管强化学习在电力系统管理中具有巨大的潜力，但其在实际应用中仍面临着一些挑战和限制。新型电力系统是一个高度复杂和动态的系统，其中涉及大量的变量和不确定性因素，因此，设计和训练适用于电力系统管理的强化学习模型需要考虑到系统的复杂性和不确定性，以确保模型的鲁棒性和可靠性。针对上述挑战，当前强化学习在电力系统管理中的改进主要集中在以下4个方向：

1) 多智能体系统方向。多智能体强化学习为解决电力系统内多个参与方之间的协作提供了新思路，它不仅提高了系统的响应速度和灵活性，还促进了资源的有效分配。

2) 模型预测控制与强化学习结合方向。这种组合充分利用了两者的优势，既保障了短期最优性的计算效率，又增强了长期规划能力，对动态调度等研究方向尤为适用。

3) 提高样本效率方向。强化学习算法通过经验回放、优先级经验回放及迁移学习等手段，显著降低了系统对新数据的需求，加速了学习过程，这对于数据获取成本高的电力行业尤为重要。

4) 考虑实时性和可扩展性方向。轻量化模型和联邦学习的应用，既满足了分布式能源资源管理的快速决策需求，也提升了系统的整体可扩展性。

表2为应用强化学习的经典算法特点及改进方向分析。

表 2 算法特点及相关分析

Tab. 2 Algorithm characteristics and related analysis

代表性算法	DQN	A-C	DDPG	Rainbow	MADDPG
特点	具有高维输入空间的决策能力,但可能会出现 Q 值过高的梯度爆炸情况	适用于处理连续动作空间和大型状态空间;能够处理部分可观测环境	适用于连续动作空间问题;训练时相比离散动作空间算法收敛稳定性更好;环境探索能力弱;超参数复杂	集成了改进的强化学习、优先经验回放、DDQN、分布式 Q 函数等多种技术	解决多智能体环境下的协作与竞争问题;集中式训练;分布式执行
架构优化改进方向	优先级经验回放机制、DDQN、多步学习等	采用更复杂的拓扑结构,或者引入注意力机制	引入噪声增加探索能力;策略延迟更新;使用自适应学习率算法调整步长	在网络上引入注意力机制,将算法扩展到多任务学习或迁移学习场景	多层次策略;自适应混合策略;分层式经验回放等
应用场景	智能电网管理、经济调度、新能源并网优化运行	电力市场交易、能源管理、智能电网控制	电能调度、电动车充电管理、电池储能系统的充放电控制	电力设备维护、电力系统规划	智能电网调度、安全态势感知等
相关文献	文献[55-59]	文献[60-63]	文献[35, 64-66]	文献[30, 67]	文献[33, 54]

4 结论

通过对比分析不同的案例研究和应用模型,可以看出,强化学习技术在电网优化、需求响应、可再生能源整合及电力系统自动化控制等方面显示出了显著优势。

1) 电力系统优化控制的突破性实现。强化学习在电力系统优化方面展现出了极大的潜力,尤其是在电网的实时调度和负荷平衡中。强化学习能够有效地处理不确定性和动态变化,提高电网的稳定性和效率。深度强化学习更是能够处理高维状态-动作空间问题,实现对动态负荷和分布式能源的协同优化,且其决策精度较传统强化学习更具优势。

2) 需求响应的智能化升级。强化学习对于需求响应管理也有重要贡献,它可以帮助电力公司预测和调整电力消费模式,实现更高效的能源分配和使用。深度强化学习则可以通过端到端学习机制,直接从海量用户用电数据中挖掘隐含模式,实现社区级需求响应的动态博弈优化。

3) 电力市场交易的策略创新。强化学习在电力市场竞争策略优化中展现出独特优势。基于强化学习框架的模型可有效模拟市场主体博弈,增强市场竞争力和灵活性。值得一提的是,深度分层强化学习还被用于处理中长期合约交易与现货市场的耦合优化问题。

4) 可再生能源整合的技术跃迁。随着可再生能源比重的逐渐增加,强化学习在整合和管理这些不稳定能源方面显示了不可小觑的价值。它可

以帮助预测能源产出,并实时调整电网运行,以适应可再生能源的波动性。此外,针对风光功率预测的不确定性构建的基于深度强化学习的模型,可以同步优化预测与调度,并在应对极端天气导致的能源波动方面表现出强泛化能力。

强化学习技术在电力领域的应用前景广阔,但要实现其全面的潜力,还需要更多的研究和技术创新。未来的工作应当集中在改进算法效率、增强系统的可扩展性及开发更为精确的预测工具上。此外,加强与业界的合作,推动实验研究向实际应用的转化,也是推动该领域发展的关键。通过这些努力,强化学习将在电力领域扮演更加重要的角色,有助于构建更智能、更高效的电力系统。

参考文献

[1] 许洪华,邵桂萍,鄂春良,等.我国未来能源系统及能源转型现实路径研究[J].发电技术,2023,44(4):484-491.
XU H H, SHAO G P, E C L, et al. Research on China's future energy system and the realistic path of energy transformation[J]. Power Generation Technology, 2023, 44(4): 484-491.

[2] 马晓伟,王文倬,薛晨,等.西北新型电力系统先行示范体系探究[J].电网与清洁能源,2024,40(1):1-7.
MA X W, WANG W Z, XUE C, et al. Research on the leading demonstration system of new-type power system in northwest China[J]. Power System and Clean Energy, 2024, 40(1): 1-7.

[3] 周翔,王继业,陈盛,等.基于深度强化学习的微网

- 优化运行综述[J]. 全球能源互联网, 2023, 6(3): 240-257.
- ZHOU X, WANG J Y, CHEN S, et al. Review of microgrid optimization operation based on deep reinforcement learning[J]. Journal of Global Energy Interconnection, 2023, 6(3): 240-257.
- [4] 冯斌, 胡轶婕, 黄刚, 等. 基于深度强化学习的新型电力系统调度优化方法综述[J]. 电力系统自动化, 2023, 47(17): 187-199.
- FENG B, HU Y J, HUANG G, et al. Review on optimization methods for new power system dispatch based on deep reinforcement learning[J]. Automation of Electric Power Systems, 2023, 47(17): 187-199.
- [5] 彭道刚, 税纪钧, 王丹豪, 等. “双碳”背景下虚拟电厂研究综述[J]. 发电技术, 2023, 44(5): 602-615.
- PENG D G, SHUI J J, WANG D H, et al. Review of virtual power plant under the background of “dual carbon” [J]. Power Generation Technology, 2023, 44(5): 602-615.
- [6] 吕小红, 刘维, 刘克恒, 等. 虚拟电厂供需侧双层协调自适应鲁棒优化调度[J]. 全球能源互联网, 2024, 7(4): 431-442.
- LYU X H, LIU W, LIU K H, et al. Two-layer coordinated adaptive robust optimal scheduling on supply and demand side of virtual power plant[J]. Journal of Global Energy Interconnection, 2024, 7(4): 431-442.
- [7] 吴垠, 牛文娟, 诸晓骏. 考虑不同场景的虚拟电厂建模与优化分析[J]. 电测与仪表, 2024, 61(11): 40-45.
- WU Y, NIU W J, ZHU X J. Modeling and optimization analysis of virtual power plant considering different scenarios[J]. Electrical Measurement & Instrumentation, 2024, 61(11): 40-45.
- [8] 张艺, 刘蕊. 考虑不确定性风险的虚拟电厂优化调度模型研究[J]. 智慧电力, 2024, 52(8): 9-18.
- ZHANG Y, LIU R. Virtual power plant optimal scheduling model considering uncertain risks[J]. Smart Power, 2024, 52(8): 9-18.
- [9] 于艾清, 濮梦燕, 王育飞, 等. 基于改进鲸鱼算法的分布式电源规划方法[J]. 电测与仪表, 2024, 61(8): 63-69.
- YU A Q, PU M Y, WANG Y F, et al. Planning method for distributed generation based on improved WOA[J]. Electrical Measurement & Instrumentation, 2024, 61(8): 63-69.
- [10] 凡鹏飞, 李宝琴, 侯江伟, 等. 配电网分布式电源经济可承载力评估[J]. 中国电力, 2024, 57(7): 196-202.
- FAN P F, LI B Q, HOU J W, et al. Economic capacity assessment of renewables in distribution networks[J]. Electric Power, 2024, 57(7): 196-202.
- [11] 王辉, 王勇, 张晓滨, 等. 构网型分布式电源渗透率高的微电网潮流计算及优化控制[J]. 电网与清洁能源, 2024, 40(4): 35-43.
- WANG H, WANG Y, ZHANG X B, et al. Power flow calculation and optimization control of the microgrid with high penetration of grid-forming distributed generators[J]. Power System and Clean Energy, 2024, 40(4): 35-43.
- [12] 李杰, 王杰, 梁文腾, 等. 含逆转型分布式电源的负荷模型构建及参数辨识方法[J]. 南方能源建设, 2024, 11(6): 164-173.
- LI J, WANG J, LIANG W T, et al. Load model construction and parameter identification method of inverter-interfaced distributed generator[J]. Southern Energy Construction, 2024, 11(6): 164-173.
- [13] 冯弋舟, 吴志, 李新煜, 等. 光-蓄-储混合储能系统灵活性提升与容量规划[J]. 电力工程技术, 2024, 43(5): 27-36.
- FENG Y Z, WU Z, LI X Y, et al. Capacity optimization and flexibility enhancement of photovoltaic-battery-pumped hybrid storage system[J]. Electric Power Engineering Technology, 2024, 43(5): 27-36.
- [14] 孙婉玉, 蒋天茁, 付强, 等. 谷电储能供暖系统优化设计方法[J]. 电力科技与环保, 2024, 40(3): 286-295.
- SUN W Y, JIANG T Z, FU Q, et al. Optimization design method of valley power energy storage heating system[J]. Electric Power Technology and Environmental Protection, 2024, 40(3): 286-295.
- [15] 王育飞, 张新宇, 张文韬, 等. 考虑调频死区的电池储能系统自适应频率控制策略[J]. 智慧电力, 2024, 52(8): 33-41.
- WANG Y F, ZHANG X Y, ZHANG W T, et al. Adaptive frequency control strategy of battery energy storage system considering frequency regulation dead band[J]. Smart Power, 2024, 52(8): 33-41.
- [16] 史宏思, 孙新伟, 王凯. 基于电化学阻抗谱的锂离子电池健康状态估计[J/OL]. 发电技术, 2024: 1-15. (2024-07-19)[2024-08-22]. <https://kns.cnki.net/KCMS/detail/detail.aspx?filename=SLJX20240716004&dbname=CJFD&dbcode=CJFQ>.
- SHI H S, SUN X W, WANG K. Estimation of health state of lithium ion battery based on electrochemical

- impedance spectroscopy[J/OL]. China Industrial Economics, 2024: 1-15. (2024-07-19)[2024-08-22]. <https://kns.cnki.net/KCMS/detail/detail.aspx?filename=SLJX20240716004&dbname=CJFD&dbcode=CJFQ>.
- [17] 张效伟, 衣振晓, 王凯. 基于改进自适应蜜獾优化算法优化时间卷积网络的车载锂离子电池健康状态估计[J/OL]. 发电技术, 2024: 1-13. (2024-07-22)[2024-08-22]. <https://kns.cnki.net/KCMS/detail/detail.aspx?filename=SLJX20240719001&dbname=CJFD&dbcode=CJFQ>.
- ZHANG X W, YI Z X, WANG K. Health state estimation of vehicle-mounted lithium-ion battery based on improved adaptive honey badger optimization algorithm to optimize time convolution network[J/OL]. China Industrial Economics, 2024: 1-13. (2024-07-22)[2024-08-22]. <https://kns.cnki.net/KCMS/detail/detail.aspx?filename=SLJX20240719001&dbname=CJFD&dbcode=CJFQ>.
- [18] 黄河, 王燕, 姜念, 等. 考虑用户诉求差异的居民可控负荷资源优化控制[J]. 发电技术, 2023, 44(6): 896-908.
- HUANG H, WANG Y, JIANG N, et al. Optimal control of residents' controllable load resources considering different demands of users[J]. Power Generation Technology, 2023, 44(6): 896-908.
- [19] 何云华, 程宇航, 袁晓冬, 等. 面向V2G调度的可信联邦学习方法[J]. 分布式能源, 2024, 9(6): 65-74.
- HE Y H, CHENG Y H, YUAN X D, et al. Trustworthy federated learning approach for V2G scheduling[J]. Distributed Energy, 2024, 9(6): 65-74.
- [20] 马继洋, 蔡永翔, 唐巍, 等. 考虑电动汽车参与韧性提升的配电网状态平滑切换控制策略[J]. 电力建设, 2024, 45(5): 29-36.
- MA J Y, CAI Y X, TANG W, et al. Enhancing distribution network resilience: electric vehicle integration in seamless state switching strategy[J]. Electric Power Construction, 2024, 45(5): 29-36.
- [21] 朱永胜, 常稳, 武东亚, 等. 考虑充放储一体站与电动汽车互动的主从博弈优化调度策略[J]. 电力系统保护与控制, 2024, 52(7): 157-167.
- ZHU Y S, CHANG W, WU D Y, et al. A Stackelberg game optimization scheduling strategy considering the interaction between a charging-discharging-storage integrated station and an electric vehicle[J]. Power System Protection and Control, 2024, 52(7): 157-167.
- [22] 吕志鹏, 宋振浩, 李立生, 等. 含电动汽车的工业园区综合能源系统优化调度[J]. 中国电力, 2024, 57(4): 25-31.
- LV Z P, SONG Z H, LI L S, et al. Optimization scheduling of integrated energy system scheduling in industrial park containing electric vehicles[J]. Electric Power, 2024, 57(4): 25-31.
- [23] 高源, 王凯. 新能源汽车电驱逆变器故障锁存保护电路研究[J]. 青岛大学学报(工程技术版), 2023, 38(4): 105-110.
- GAO Y, WANG K. Study on electric vehicle inverter system fault latch and protection circuit[J]. Journal of Qingdao University (Engineering & Technology Edition), 2023, 38(4): 105-110.
- [24] PERERA A T D, KAMALARUBAN P. Applications of reinforcement learning in energy systems[J]. Renewable and Sustainable Energy Reviews, 2021, 137: 110618.
- [25] CHEN X, QU G N, TANG Y J, et al. Reinforcement learning for selective key applications in power systems: recent advances and future challenges[J]. IEEE Transactions on Smart Grid, 2022, 13(4): 2935-2958.
- [26] 尚玉朝, 刘春豪, 王凯. 基于融合优化算法的超级电容器健康状态预测模型[J/OL]. 发电技术, 1-11 [2025-02-19]. <http://kns.cnki.net/kcms/detail/33.1405.TK.20250117.1141.002.html>.
- SHANG Y Z, LIU C H, WANG K, et al. A state of health prediction model for supercapacitors based on an integrated optimization algorithm[J/OL]. Power Generation Technology, 1-11[2025-02-19]. <http://kns.cnki.net/kcms/detail/33.1405.TK.20250117.1141.002.html>.
- [27] ZHANG Z D, ZHANG D X, QIU R C. Deep reinforcement learning for power system applications: an overview[J]. CSEE Journal of Power and Energy Systems, 2020, 6(1): 213-225.
- [28] VÁZQUEZ-CANTELI J R, NAGY Z. Reinforcement learning for demand response: a review of algorithms and modeling techniques[J]. Applied Energy, 2019, 235: 1072-1089.
- [29] JORDEHI A R. Optimisation of demand response in electric power systems: a review[J]. Renewable and Sustainable Energy Reviews, 2019, 103: 308-319.
- [30] YANG J J, YANG M, WANG M X, et al. A deep reinforcement learning method for managing wind farm uncertainties through energy storage system control and external reserve purchasing[J]. International Journal of

- Electrical Power & Energy Systems, 2020, 119: 105928.
- [31] CHEN T Y, BU S R, LIU X, et al. Peer-to-peer energy trading and energy conversion in interconnected multi-energy microgrids using multi-agent deep reinforcement learning[J]. IEEE Transactions on Smart Grid, 2022, 13(1): 715-727.
- [32] 韦贵熙, 刘香港, 池明, 等. 基于多智能体强化学习的微电网能源交易[J]. 控制工程, 2023, 30(12): 2274-2279.
- WEI G X, LIU X G, CHI M, et al. Energy trading in microgrid via multi-agent reinforcement learning[J]. Control Engineering of China, 2023, 30(12): 2274-2279.
- [33] 张兴平, 王腾, 张馨月, 等. 基于多智能体深度确定策略梯度算法的火力发电竞价策略[J]. 中国电力, 2024, 57(11): 161-172.
- ZHANG X P, WANG T, ZHANG X Y, et al. Bidding strategy for thermal power generation companies based on multi-agent deep deterministic policy gradient algorithm[J]. Electric Power, 2024, 57(11): 161-172.
- [34] 许丹, 胡晓静, 胡斐, 等. 基于深度强化学习的电力市场量价组合竞价策略[J]. 电网技术, 2024, 48(8): 3278-3286.
- XU D, HU X J, HU F, et al. Strategic bidding of price-quantity pairs in electricity market based on deep reinforcement learning[J]. Power System Technology, 2024, 48(8): 3278-3286.
- [35] 陈思畏, 李建军, 邹信迅, 等. 基于孪生延迟DDPG强化学习的电-热耦合系统低碳经济调度[J]. 现代电力, 2025, 42(2): 314-321.
- CHEN S W, LI J J, ZOU X X, et al. Low-carbon economic dispatch of electro-thermal coupling system based on DDPG reinforcement learning with twin delays [J]. China Industrial Economics, 2025, 42(2): 314-321.
- [36] LU R Z, HONG S H, ZHANG X F. A dynamic pricing demand response algorithm for smart grid: Reinforcement learning approach[J]. Applied Energy, 2018, 220: 220-230.
- [37] 李锦辉, 吴毓峰, 余涛, 等. 数据孤岛下基于联邦学习的用户电价响应刻画及其应用[J]. 电力系统保护与控制, 2024, 52(6): 164-176.
- LI J H, WU Y F, YU T, et al. Characterization of user price response behavior and its application based on federated learning considering a data island[J]. Power System Protection and Control, 2024, 52(6): 164-176.
- [38] 陈潇潇, 周云海, 张泰源, 等. 基于深度强化学习的有源配电网实时电压控制策略[J]. 三峡大学学报(自然科学版), 2024, 46(1): 76-84.
- CHEN X X, ZHOU Y H, ZHANG T Y, et al. Strategy of real-time voltage control in active distribution network based on deep reinforcement learning[J]. Journal of China Three Gorges University (Natural Sciences), 2024, 46(1): 76-84.
- [39] YANG Q L, WANG G, SADEGHI A, et al. Two-timescale voltage control in distribution grids using deep reinforcement learning[J]. IEEE Transactions on Smart Grid, 2020, 11(3): 2313-2323.
- [40] 张华强, 牟晨东, 赵玫, 等. 基于强化学习的多光储虚拟同步机频率协调控制策略[J]. 电气传动, 2021, 51(19): 36-42.
- ZHANG H Q, MU C D, ZHAO M, et al. Frequency coordination control strategy of multiple photovoltaic-battery virtual synchronous generators based on reinforcement learning[J]. Electric Drive, 2021, 51(19): 36-42.
- [41] 姚建华, 胡晟, 王冠, 等. 基于强化学习的孤岛微电网多源协调频率控制方法[J]. 电力建设, 2020, 41(9): 69-75.
- YAO J H, HU S, WANG G, et al. Multi-source coordinated frequency control method based on reinforcement learning for island microgrid[J]. Electric Power Construction, 2020, 41(9): 69-75.
- [42] 刘俊峰, 陈剑龙, 王晓生, 等. 基于深度强化学习的微能源网能量管理与优化策略研究[J]. 电网技术, 2020, 44(10): 3794-3803.
- LIU J F, CHEN J L, WANG X S, et al. Energy management and optimization of multi-energy grid based on deep reinforcement learning[J]. Power System Technology, 2020, 44(10): 3794-3803.
- [43] ZHAO H, ZHAO J H, QIU J, et al. Cooperative wind farm control with deep reinforcement learning and knowledge-assisted learning[J]. IEEE Transactions on Industrial Informatics, 2020, 16(11): 6912-6921.
- [44] LI Y, WANG R N, LI Y Z, et al. Wind power forecasting considering data privacy protection: a federated deep reinforcement learning approach[J]. Applied Energy, 2023, 329: 120291.
- [45] KANG D J, KANG D, HWANGBO S, et al. Optimal planning of hybrid energy storage systems using curtailed renewable energy through deep reinforcement learning[J]. Energy, 2023, 284: 128623.
- [46] 徐钰涵, 季天瑶, 李梦诗. 基于深度强化学习的微电

- 网日前日内协调优化调度[J]. 南方电网技术, 2024, 18(9): 106-116.
- XU Y H, JI T Y, LI M S. Day-ahead and intra-day coordinated optimal scheduling of microgrid based on deep reinforcement learning[J]. Southern Power System Technology, 2024, 18(9): 106-116.
- [47] 冯文韬, 李龙胜, 曾愚, 等. 基于深度强化学习的微电网源-荷低碳调度优化研究[J]. 四川电力技术, 2023, 46(6): 75-82.
- FENG W T, LI L S, ZENG Y, et al. Research on source-load low-carbon optimal dispatching for microgrid based on deep reinforcement learning[J]. Sichuan Electric Power Technology, 2023, 46(6): 75-82.
- [48] 李文武, 马浩云, 贺中豪, 等. 基于 n 步Q-learning算法的风电抽水蓄能联合系统日随机优化调度研究[J]. 水电能源科学, 2022, 40(1): 206-210.
- LI W W, MA H Y, HE Z H, et al. Research on daily stochastic optimal scheduling of wind power pumped storage system based on n -step Q-learning algorithm[J]. Water Resources and Power, 2022, 40(1): 206-210.
- [49] DUAN J J, SHI D, DIAO R S, et al. Deep-reinforcement-learning-based autonomous voltage control for power grid operations[J]. IEEE Transactions on Power Systems, 2020, 35(1): 814-817.
- [50] 张涛, 郝正航, 徐玉韬, 等. 基于多智能体强化学习的两阶段电压控制策略[J]. 南方电网技术, 2024, 18(12): 77-86.
- ZHANG T, HAO Z H, XU Y T, et al. Two-stage voltage control strategy based on multi-agent reinforcement learning[J]. Southern Power System Technology, 2024, 18(12): 77-86.
- [51] 韩保军, 高强, 代飞, 等. 基于协同奖励函数多目标强化学习的智能频率控制策略研究[J]. 电力科学与技术学报, 2023, 38(2): 18-29.
- HAN B J, GAO Q, DAI F, et al. Intelligent frequency control strategy based on multi-objective reinforcement learning of cooperative reward function[J]. Journal of Electric Power Science and Technology, 2023, 38(2): 18-29.
- [52] HUANG R, LI Y C, WANG X. Attention-aware deep reinforcement learning for detecting false data injection attacks in smart grids[J]. International Journal of Electrical Power & Energy Systems, 2023, 147: 108815.
- [53] 郭创新, 刘祝平, 刘永刚, 等. 基于图神经网络和强化学习的电网风险态势感知[J]. 电网与清洁能源, 2023, 39(12): 41-49.
- GUO C X, LIU Z P, LIU Y G, et al. GNN and RL based power system risk situation perception[J]. Power System and Clean Energy, 2023, 39(12): 41-49.
- [54] LEI W X, WEN H, WU J S, et al. MADDPG-based security situational awareness for smart grid with intelligent edge[J]. Applied Sciences, 2021, 11(7): 3101.
- [55] 易仕琪, 孔政敏, 王帅, 等. 基于泛化强化学习的变电站巡检机器人路径规划研究[J]. 广东电力, 2023, 36(11): 114-121.
- YI S Q, KONG Z M, WANG S, et al. Study on substation inspection robot path planning based on generalization reinforcement learning[J]. Guangdong Electric Power, 2023, 36(11): 114-121.
- [56] 倪爽, 崔承刚, 杨宁, 等. 基于深度强化学习的配电网多时间尺度在线无功优化[J]. 电力系统自动化, 2021, 45(10): 77-85.
- NI S, CUI C G, YANG N, et al. Multi-time-scale online optimization for reactive power of distribution network based on deep reinforcement learning[J]. Automation of Electric Power Systems, 2021, 45(10): 77-85.
- [57] 张超, 赵冬梅, 季宇, 等. 基于改进深度Q网络的虚拟电厂实时优化调度[J]. 中国电力, 2024, 57(1): 91-100.
- ZHANG C, ZHAO D M, JI Y, et al. Real time optimal dispatch of virtual power plant based on improved deep Q network[J]. Electric Power, 2024, 57(1): 91-100.
- [58] 吴润泽, 霍金鑫, 郭昊博. 基于DQN的电力协同计算与缓存的任务卸载策略[J]. 电力建设, 2024, 45(8): 149-158.
- WU R Z, HUO J X, GUO H B. DQN-based task offloading strategy for power co-computing and caching[J]. Electric Power Construction, 2024, 45(8): 149-158.
- [59] 张一凡. 基于深度学习的水风光短期随机优化调度研究[J]. 水电与新能源, 2024, 38(3): 34-37.
- ZHANG Y F. Short-term stochastic optimization scheduling of hydro-wind-solar power system based on deep learning[J]. Hydropower and New Energy, 2024, 38(3): 34-37.
- [60] XIE J, SUN W. Distributional deep reinforcement learning-based emergency frequency control[J]. IEEE Transactions on Power Systems, 2022, 37(4): 2720-2730.
- [61] 马庆, 邓长虹. 基于单/多智能体简化强化学习的电力系统无功电压控制[J]. 电工技术学报, 2024,

- 39(5): 1300-1312.
- MA Q, DENG C H. Single/multi agent simplified deep reinforcement learning based volt-var control of power system[J]. Transactions of China Electrotechnical Society, 2024, 39(5): 1300-1312.
- [62] YUAN Y X, DEHGHANPOUR K, WANG Z Y, et al. A joint distribution system state estimation framework via deep actor-critic learning method[J]. IEEE Transactions on Power Systems, 2023, 38(1): 796-806.
- [63] YU Q S, WANG X Y, LV D P, et al. Data fusion and situation awareness for smart grid and power communication network based on tensor computing and deep reinforcement learning[J]. Electronics, 2023, 12(12): 2606.
- [64] 陈明昊, 孙毅, 谢志远. 基于双层深度强化学习的园区综合能源系统多时间尺度优化管理[J]. 电工技术学报, 2023, 38(7): 1864-1881.
- CHEN M H, SUN Y, XIE Z Y. The multi-time-scale management optimization method for park integrated energy system based on the bi-layer deep reinforcement learning[J]. Transactions of China Electrotechnical Society, 2023, 38(7): 1864-1881.
- [65] 梁煜东, 陈峦, 张国洲, 等. 基于深度强化学习的多能互补发电系统负荷频率控制策略[J]. 电工技术学报, 2022, 37(7): 1768-1779.
- LIANG Y D, CHEN L, ZHANG G Z, et al. Load frequency control strategy of hybrid power generation system: a deep reinforcement learning-based approach[J]. Transactions of China Electrotechnical Society, 2022, 37(7): 1768-1779.
- [66] 刘建行, 刘方. 基于深度强化学习的梯级水蓄风光互补系统优化调度策略研究[J]. 广东电力, 2024, 37(5): 10-22.
- LIU J H, LIU F. Research on optimized dispatching strategy of cascade hydropower-pumping-storage-wind-photovoltaic multi-energy complementary system based on deep reinforcement learning[J]. Guangdong Electric Power, 2024, 37(5): 10-22.
- [67] 于一潇, 杨佳峻, 杨明, 等. 基于深度强化学习的风电场储能系统预测决策一体化调度[J]. 电力系统自动化, 2021, 45(1): 132-140.
- YU Y X, YANG J J, YANG M, et al. Prediction and decision integrated scheduling of energy storage system in wind farm based on deep reinforcement learning[J]. Automation of Electric Power Systems, 2021, 45(1): 132-140.
-
- 收稿日期: 2024-10-22。
修回日期: 2025-01-19。
作者简介:
-
- 闫正义
- 闫正义(1998), 男, 硕士研究生, 研究方向为锂电池荷电状态估计, zzzx1998111@163.com;
-
- 赵康
- 赵康(2000), 男, 硕士研究生, 研究方向为锂电池剩余寿命估计, mailboxofzk@163.com;
-
- 王凯
- 王凯(1985), 男, 博士, 教授, 研究方向为分布式微电网和储能技术、新能源储能技术、能源互联网等, 本文通信作者, wangkai@qdu.edu.cn。
- (责任编辑 苗雪连)