

基于深度强化学习的智能电网优化调度策略研究

张梓扬¹ 胡晓玥²

(1. 新西兰奥克兰大学 新西兰 奥克兰 1052;

2. 北京智芯微电子科技有限公司 北京 102299)

摘要 文中探讨了深度强化学习在智能电网优化调度策略中的应用。通过引入深度Q网络(DQN)算法,验证了该技术在提高供电可靠性、降低发电成本和减少网损方面的有效性。深度强化学习调度策略在多种电网场景下均展现出显著的优势。实验结果表明,DRL调度策略能提高近3个百分点的供电可靠性,并明显降低发电成本和网损。

关键词: 深度强化学习;智能电网;优化调度

中图分类号 TM732

Research on optimal scheduling strategy of smart grid based on deep reinforcement learning

ZHANG Ziyang¹ and HU Xiaoyue²

(1. University of Auckland, Auckland 1052, New Zealand;

2. Beijing Smartchip Microelectronics Technology Co., Ltd., Beijing 102299, China)

Abstract This paper explores the application of deep reinforcement learning in optimizing scheduling strategies for smart grids. By introducing the Deep Q-Network (DQN) algorithm, the effectiveness of this technology in improving power supply reliability, reducing power generation costs, and minimizing network losses is verified. The deep reinforcement learning scheduling strategy has shown significant advantages in various power grid scenarios. The experimental results show that the DRL scheduling strategy can improve power supply reliability by nearly 3 percentage points and significantly reduce power generation costs and network losses.

Key words Deep reinforcement learning, Smart grid, Optimal scheduling

0 引言

在面对日益复杂的电网结构和多变的电力需求时,传统的电网调度方法逐渐显露出局限性。为了克服这些挑战,并推动智能电网技术的进一步发展,本文引入深度强化学习方法来探索智能电网的优化调度策略。深度强化学习结合深度学习的感知能力和强化学习的决策能力,使得智能系统能在与环境的交互中自我学习和优化,从而实现更智能和高效的调度决策。

1 方法与实验

1.1 方法论述

深度强化学习是人工智能领域中的重要技术,它结合了深度学习的表征学习能力和强化学习的序贯决策能力。在深度强化学习中,智能体通过与环境的交互来学习如何做出最优决策。这种交互基于一个奖励信号,该信号反映智能体在某个状态下采取某个动作的好坏。深度强化学习的基本原理涉及到一个关键概念——Q值函数,它表示在

给定状态下采取特定动作所能获得的预期回报。Q-learning是一种经典的强化学习算法,它通过不断更新Q值函数来学习最优策略。在Q-learning中,智能体根据以下规则更新Q值,如式(1)所示:

$$Q_{new}(s, a) = Q(s, a) + \alpha \cdot (r + \gamma \cdot \max_{a'} Q(s', a') - Q(s, a)) \quad (1)$$

其中, $Q_{new}(s, a)$ 是更新后的Q值, s 和 a 分别是当前的状态和动作, r 是即时奖励, s' 是下一个状态, a' 是在下一个状态可能采取的动作, α 是学习率, γ 是折扣因子。

然而,当状态空间和动作空间非常大时,传统的Q-learning方法会因需要存储和更新大量的Q值而变得不切实际,这时深度Q网络(DQN)算法就显得尤为重要。DQN使用深度神经网络来近似Q值函数,从而能处理高维的状态空间。DQN还引入经验回放和目标网络等技巧来提高学习过程的稳定性和效率。在智能电网的调度策略中应用深度强化学习算法,首先需要定义状态空间^[1]。状态空间可以包括电网的实时负荷、发电机的出力、线路的潮流等关键运行参数。这些参数反映了电网的当前状态,是智能体进行决策的基础。此外,动作空间的设计也是至关重要的。在智能电网调度中,动作包括调整发电机的出力、切换线路

作者简介:张梓扬(2000—),硕士,研究方向为计算机科学与技术;胡晓玥(1988—),硕士,工程师,研究方向为计算机科学与技术。

的开关状态等,旨在优化电网的运行状态,如最小化网损、最大化供电可靠性等。最后,构建合适的奖励函数是深度强化学习在智能电网调度中应用的关键。奖励函数应能反映电网运行的目标和约束,如供电可靠性、经济性、环保性等。通过合理设计奖励函数,可以引导智能体学习到符合电网运行需求的优化调度策略。综上所述,深度强化学习在智能电网调度策略中的应用具有巨大的潜力。通过合理定义状态空间、设计动作空间以及构建奖励函数,可以训练出能自适应调整的智能电网调度策略,从而提高电网的运行效率和稳定性。

1.2 实验设计

本实验旨在通过模拟平台验证深度强化学习在智能电网调度策略中的有效性。为此,首先需搭建一个包含电网模型和负荷模型的实验环境。在电网模型方面,采用具有多个节点和线路的复杂网络拓扑结构,以模拟实际电网的运行情况。该模型不仅包括传统的发电机、变压器、线路等电力设备,还整合可再生能源发电单元,如风电场和太阳能发电站,以反映现代智能电网的多元化能源结构^[2]。同时,模型中的电力设备均配备了详细的参数设置,如发电机的出力范围、线路的传输容量等,以确保模拟的准确性和真实性。负荷模型则基于历史数据和统计分布进行构建,以模拟电网中的实时负荷变化。通过引入时序变化和随机波动因素,负荷模型能更贴近实际电网的运行状态,为调度策略的制定提供更真实的背景。在实验数据和参数设置方面,采用多种场景数据进行模拟,包括正常负荷、高峰负荷以及异常情况下的负荷变化等。同时,为了全面评估调度策略的性能,设置多个关键参数,如调度周期、发电成本、网损、供电可靠性等。这些参数将在实验过程中进行动态调整,以观察不同参数设置对调度策略的影响。在实验过程中,需要观察和记录的关键指标主要包括电网的运行状态、调度决策的执行情况以及各项性能指标的变化。具体而言,需关注电网的电压稳定性、频率稳定性等运行指标,以及调度过程中发电机的出力分配、线路的潮流分布等决策指标。同时,还需详细记录实验过程中的发电成本、网损、供电可靠性等性能指标,以便后续对调度策略进行综合评估。

1.3 实验过程

1.3.1 智能体的训练

智能体的训练是深度强化学习在智能电网调度中应用的核心步骤。训练开始前,需要构建一个与真实电网环境相近的模拟器,并在其中实现状态转移逻辑和奖励函数。状态转移逻辑负责根据当前状态和智能体选择的动作来更新电网状态,而奖励函数则根据新状态给出奖励值,以指导智能体的学习。智能体的训练采用深度Q网络(DQN)算法,该算法结合了Q-learning和深度学习。在训练过程中,智能体通过 ϵ -greedy策略进行动作选择,即在大部分情况下选择Q值最高的动作,但有一定概率 ϵ 选择随机动作,以探索新的可能^[3],这种探索与利用的平衡是强化学习中的

关键。神经网络的结构设计对于训练效果至关重要。在本实验中,采用包含多个隐藏层的全连接神经网络来近似Q值函数。网络的输入是电网的状态信息,如各节点的电压、电流、功率等,输出则是对应每个可能动作的Q值。在训练过程中,智能体通过与环境不断交互来积累经验,即状态-动作-奖励-新状态(SARS)四元组。这些经验被存储在回放缓冲区中,并用于后续的学习过程。通过从缓冲区中随机抽取经验样本,智能体能打破数据之间的相关性,提高学习的稳定性。每经过一定数量的交互步骤,智能体会从回放缓冲区中抽取一批经验样本,并使用梯度下降算法来更新神经网络的权重。这个过程旨在最小化预测Q值与实际Q值之间的差异,从而优化智能体的决策策略。

1.3.2 与环境的交互方式

智能体与电网环境的交互是通过模拟器实现的。在每个时间步,智能体会观察当前电网状态,并根据其策略选择一个动作。这个动作被传递给模拟器,模拟器根据动作和当前状态更新电网状态,并计算出一个奖励值^[4]。这个奖励值反映执行该动作后电网性能的改善程度,如减少网损、提高供电可靠性等。智能体接收到新状态和奖励值后,会将其与之前的状态、动作和奖励一起存储为一条经验。这条经验随后被添加到回放缓冲区中,用于后续的学习过程。

1.3.3 实验结果展示

为了全面评估智能体的性能,在多种电网场景下进行测试,并将结果与传统的调度策略进行对比,对比的指标包括供电可靠性、发电成本和网损等。如表1所列,在供电可靠性方面,DRL调度策略相较于传统调度策略提高近3个百分点。同时,在发电成本和网损方面,DRL调度策略也表现出明显的优势。这些结果表明,深度强化学习在智能电网优化调度中具有巨大的潜力和应用价值。通过对DRL调度策略在不同电网场景下的性能进行分析可以发现,无论是在正常负荷、高峰负荷还是异常情况下的负荷变化场景中,DRL调度策略都能实现更智能和高效的调度决策,这进一步证明了深度强化学习方法在智能电网调度中的有效性和适应性。

表1 实验结果对比表

| 调度策略 | 供电可靠性/% | 平均发电成本/ 元/千瓦时 | 平均网损/% |
|-------|---------|------------------|--------|
| 传统调度 | 96.5 | 0.42 | 2.8 |
| DRL调度 | 99.3 | 0.35 | 1.6 |

2 结果与讨论

2.1 结果分析

深度强化学习在智能电网调度中展现出了显著的优势,但同时也存在一定的局限性。根据实验结果,DRL调度策略相较于传统调度方法,在供电可靠性、网损减少和发电成本降低等多个方面均表现出更优的性能。这主要归功于

DRL能从与环境的交互中自主学习并优化决策策略,而无需依赖预设的规则或模型。通过对比实验数据和图表可以发现,在复杂多变的电网环境下,DRL调度策略能更快地适应并作出优化决策。例如,在供电可靠性方面,DRL策略通过智能地调整电网资源的分配,有效减少停电事件的发生,从而提高系统的可靠性。在经济效益方面,DRL策略通过优化发电机的出力分配和负荷的调度,降低发电成本和网损,提高电网运行的经济性。DRL调度策略的局限性也不可忽视^[5]。首先,深度强化学习算法的训练过程需要大量的数据和计算资源,这在一定程度上限制了其在实时调度系统中的应用。其次,DRL策略的性能高度依赖于训练数据的质量和数量,如果训练数据不足或存在偏差,则可能会导致策略的性能下降。

2.2 策略评估

在实际应用中,不同调度策略的可行性和效果因场景而异。传统调度方法通常基于预设的规则和模型进行决策,这种方法在简单且稳定的电网环境下可能表现出较好的性能。然而,在复杂多变的电网环境下,传统方法的适应性较差,难以应对突发情况和非线性变化。相比之下,DRL调度策略在实际应用中展现出更强的适应性和优化能力。通过自主学习和不断优化决策策略,DRL策略能在各种电网场景下实现更高效的资源分配和负荷调度。这不仅能提高智能电网的运行效率,还能增强电网的稳定性和经济性。深度强化学习调度策略在提高智能电网运行效率、稳定性和经济性方面做出显著贡献。然而,其在实际应用中的效果仍受到训练数据、计算资源等因素的限制。

2.3 未来展望

为了进一步优化和提升智能电网调度的性能,可以引入更先进的深度强化学习算法。例如,使用基于模型的强化学习(Model-Based Reinforcement Learning)来加速学习过程并提高样本效率;或采用分布式强化学习(Distributed Reinforcement Learning),并利用多个智能体之间的协作来

共同优化电网调度策略。此外,还可以结合其他优化技术。例如,可以将深度强化学习与传统的优化算法(如线性规划、动态规划等)相结合,以充分利用两者的优势并弥补各自的不足。这种混合方法有望在保持深度强化学习灵活性和自适应性的同时,提高调度策略的稳定性和效率。未来,智能电网调度在能源互联网和分布式发电等领域的应用前景广阔。随着可再生能源的大规模接入和电动汽车等新型负荷的快速增长,智能电网调度将面临更加复杂和多变的挑战。因此,研究和开发更加智能、高效和稳定的调度策略对于保障电力系统的安全稳定运行具有重要意义。

3 结语

本文通过深度强化学习技术,为智能电网调度策略的优化提供了一种新的方法。实验结果充分证明,该方法在提升电网运行效率、稳定性和经济性方面的潜力。随着智能电网技术的不断发展,未来仍需在算法改进、数据效率提升以及与其他优化技术的结合等方面进行深入探索。随着相关研究的不断推进,深度强化学习将在智能电网调度中发挥更大的作用,为保障电力系统的安全稳定运行贡献重要力量。

参考文献

- [1] 季飞. 基于深度强化学习的可信变电站电力调度系统设计[J]. 自动化应用, 2024, 65(13): 246-248.
- [2] 胡维昊, 曹迪, 黄琦, 等. 深度强化学习在配电网优化运行中的应用[J]. 电力系统自动化, 2023, 47(14): 174-191.
- [3] 张小花, 张中超, 黄晓怡, 等. 基于深度学习的电网实时调度数据可靠度评估算法研究[J]. 自动化应用, 2023, 64(3): 16-19.
- [4] 巨云涛, 陈希. 基于双层多智能体强化学习的微网群分布式有功无功协调优化调度[J]. 中国电机工程学报, 2022, 42(23): 8534-8548.
- [5] 朱永东. 专题: 面向5G/6G的智能边缘计算网络技术特邀策划人[J]. 无线电通信技术, 2022, 48(1): 32-33.