

# 基于深度强化学习的火电厂机组负荷优化分配研究

乔仙保 王刚 马瑞

北方联合电力有限责任公司临河热电厂 内蒙古 巴彦淖尔 015000

**摘要：**本文聚焦基于深度强化学习的火电厂机组负荷优化分配研究，阐述了深度强化学习基本原理，包括强化学习概述、深度学习与强化学习的结合及常见算法。分析了火电厂机组负荷分配的目标、约束条件以及传统方法的局限性。随后构建基于深度强化学习的负荷优化分配模型，涵盖总体架构、状态空间、动作空间、奖励函数设计及算法选择与训练。通过仿真实验，将基于DDPG算法的模型与传统算法对比，结果表明深度强化学习算法在实时性和应对动态工况等方面优势显著，能实现机组负荷高效合理分配。

**关键词：**深度强化学习；火电厂；负荷优化分配

## 1 深度强化学习基本原理

### 1.1 强化学习概述

强化学习作为机器学习领域中极具活力与挑战性的分支，是一种依靠智能体与环境交互来习得最优决策策略的独特方法。在强化学习的特定框架里，智能体扮演着学习主体的关键角色。当它在所处环境中执行某一动作后，环境会迅速做出响应，给予智能体一个奖励信号，这一信号犹如“裁判的评分”，精准反映了该动作的优劣程度；智能体的核心目标十分明确，即凭借持续不断地尝试与探索，挖掘出一个最优策略，从而在长期的运行进程中，使自身获得的累计奖励达到最大化。强化学习的基本要素涵盖智能体、环境、状态、动作和奖励。智能体依据环境当前的状态，审慎地选择动作；环境则依据智能体的动作，生成新的状态并给予相应奖励。状态是对环境当下状况的细致描述，动作是智能体能够采取的行为集合，而奖励作为环境对智能体动作的反馈，如同“导航灯”，为智能体的学习过程指引方向。

### 1.2 深度学习与强化学习的结合

深度学习具有强大的特征提取和表示学习能力，能够自动从大量数据中学习到高层次的特征表示。将深度学习与强化学习相结合，形成了深度强化学习。深度强化学习利用深度神经网络来近似强化学习中的价值函数或策略函数，从而能够处理高维状态空间和复杂的环境问题。在深度强化学习中，常用的深度神经网络结构包括卷积神经网络（CNN）、循环神经网络（RNN）及其变体（如长短时记忆网络LSTM、门控循环单元GRU）等。CNN适用于处理具有网格结构的数据，如图像；RNN及其变体则擅长处理序列数据，能够捕捉数据中的时序信息。通过合理选择和设计深度神经网络结构，深度强化学习算法能够在复杂环境中实现高效的决策和学习<sup>[1]</sup>。

## 1.3 常见的深度强化学习算法

### 1.3.1 深度Q网络（DQN）

DQN是深度强化学习领域的经典算法之一。它使用深度神经网络来近似Q函数，即状态-动作值函数。Q函数表示在给定状态下采取某个动作后能够获得的未来累计奖励的期望值。DQN通过不断更新神经网络的参数来最小化预测Q值与目标Q值之间的误差，从而实现对Q函数的近似。为了提高学习的稳定性和收敛性，DQN引入了经验回放机制和目标网络等技术。

### 1.3.2 深度确定性策略梯度（DDPG）

DDPG是一种基于演员-评论家架构的深度强化学习算法，适用于连续动作空间的问题。在DDPG中，演员网络负责输出确定性的动作，评论家网络则用于评估演员网络输出的动作的价值。通过同时训练演员网络和评论家网络，DDPG能够学习到最优的策略。DDPG结合了深度学习和策略梯度方法的优点，在处理连续控制问题上具有较好的性能。

## 2 火电厂机组负荷优化分配问题分析

### 2.1 火电厂机组负荷分配的目标与约束条件

火电厂机组负荷分配的主要目标是在满足电力系统对电厂总负荷需求的前提下，优化各机组的负荷分配，使得电厂的运行成本最低、效率最高。运行成本主要包括燃料成本、维护成本等，其中燃料成本占据主导地位。因此，通常以最小化燃料成本作为负荷分配的主要优化目标。同时，机组负荷分配需要满足一系列约束条件，包括机组的技术特性约束，如机组的出力上下限、爬坡速率限制等；电力系统的平衡约束，即电厂总出力要等于电力系统对电厂的负荷需求；以及其他约束，如环保约束、设备运行安全约束等。

### 2.2 传统负荷分配方法的局限性

传统数学规划算法在处理火电厂机组负荷分配问题时，需要将问题抽象为数学模型，然后通过求解数学模型得到最优解。然而，实际火电厂运行工况复杂多变，存在许多不确定因素，如机组性能随时间的变化、燃料质量的波动等，这使得准确建立数学模型变得困难。另外，传统算法的计算复杂度较高，尤其是在处理大规模火电厂和复杂约束条件时，计算时间往往较长，难以满足实时调度的要求。智能算法虽然在一定程度上能够克服传统算法的一些局限性，但在处理动态变化工况和长期优化问题时，仍存在不足。例如，遗传算法等智能算法在搜索过程中容易出现早熟收敛问题，导致陷入局部最优解；粒子群优化算法在处理高维复杂问题时，性能可能会下降<sup>[2]</sup>。

### 3 基于深度强化学习的火电厂机组负荷优化分配模型构建

#### 3.1 模型总体架构

基于深度强化学习的火电厂机组负荷优化分配模型，是一个将智能算法与火电厂实际运行紧密结合的综合性体系。该模型主要由深度强化学习算法模块、火电厂机组仿真环境模块以及交互接口模块这三大部分构成。深度强化学习算法模块是模型的核心“大脑”，它具备强大的学习和优化能力。通过不断地与火电厂机组仿真环境交互，收集数据并分析，持续改进负荷分配策略，以适应火电厂复杂多变的运行状况，实现机组负荷分配的最优化，提升火电厂的运行效率和经济效益；火电厂机组仿真环境模块则是对现实火电厂运行工况的精准复刻，它涵盖了机组的各种性能参数、燃料特性以及电力系统的实际需求等因素，为深度强化学习算法提供了一个逼真且可控的训练和测试场所，让算法能在接近真实的环境中学习和验证，增强算法的实用性和可靠性。交互接口模块作为连接算法模块和仿真环境模块的桥梁，起着至关重要的作用。它负责在两者之间准确、高效地传递信息，包括将仿真环境中的机组运行状态等信息准确无误地传达给算法模块，同时把算法模块根据状态做出的动作决策传递给仿真环境，使环境产生相应反馈，形成完整的学习闭环。

#### 3.2 状态空间设计

状态空间的设计对于深度强化学习模型至关重要，它需全面涵盖能反映火电厂机组运行状态和电力系统需求的关键信息。在本模型里，状态空间包含多方面内容。各机组的当前出力是重要信息之一，它直观呈现了机组当下的实际发电能力，是判断机组运行是否正常以及是否需调整负荷的关键依据。机组的爬坡速率同样不

可忽视，它反映了机组在单位时间内调整负荷的能力，影响着负荷调整的速度和灵活性；电力系统对电厂的总负荷需求是状态空间的核心要素，它决定了火电厂整体的发电任务，是负荷分配的根本出发点。当前燃料价格也是关键因素，燃料成本占火电厂运行成本的大头，价格波动会直接影响经济效益，将其纳入状态空间可让算法在分配负荷时考虑成本因素；机组的运行时间也需纳入考量，随着运行时间增长，机组性能会变化，可能出现效率下降等问题。这些信息共同构成了全面描述机组和电力系统当前状态的状态空间，为深度强化学习算法提供坚实可靠的决策基础。

#### 3.3 动作空间设计

动作空间的设计紧密结合火电厂机组负荷分配的实际情况，在本模型中，将其定义为各机组负荷的调整量。由于机组负荷调整量是连续变化的数值，并非离散的几个固定值，所以本模型采用连续动作空间；这种设计能够更精准地模拟实际机组负荷调整过程，使深度强化学习算法可以根据不同的运行工况和需求，灵活地调整各机组的负荷。例如，当电力系统对电厂的总负荷需求突然增加时，算法可以根据各机组的当前状态，如出力、爬坡速率等，从连续的动作空间中选择合适的负荷调整量，将增加的负荷合理分配到各机组上，避免出现部分机组过载而其他机组负荷不足的情况；同样，当总负荷需求减少时，算法也能通过选择合适的动作，平稳地降低各机组的负荷，确保机组稳定运行和电力系统的平衡。通过这种方式，深度强化学习算法可以实现对机组负荷的动态、精细分配，提高火电厂的运行效率和稳定性<sup>[3]</sup>。

#### 3.4 奖励函数设计

奖励函数在深度强化学习算法中起着引导和激励的关键作用，是促使算法学习到最优策略的核心要素。在本模型中，奖励函数的设计综合考虑了多个重要目标。最小化燃料成本是首要目标之一，燃料成本是火电厂运行的主要开支，降低燃料成本能显著提升经济效益。因此，奖励函数会鼓励算法选择那些能使燃料消耗最低的负荷分配方案。同时，要满足机组出力约束，机组出力必须在规定的上下限范围内，否则可能损坏机组或影响运行稳定性，奖励函数会对违反出力约束的行为给予惩罚；电力系统平衡约束也不容忽视，火电厂的发电量要与电力系统的需求相匹配，以维持系统稳定。奖励函数会引导算法根据总负荷需求合理分配机组负荷。此外，为鼓励算法积极探索有效策略，避免陷入局部最优解，还设置了一个小的正奖励。当算法满足所有约束条件

时，会根据燃料成本、机组出力与经济运行点的接近程度等因素给予综合奖励；当不满足约束条件时，根据违反程度给予相应负奖励，惩罚不合理动作。

### 3.5 深度强化学习算法选择与训练

鉴于火电厂机组负荷分配问题具有连续动作空间的特性，本模型选择深度确定性策略梯度（DDPG）算法作为核心学习算法。DDPG算法结合了深度学习和强化学习的优势，能够在连续动作空间中高效地搜索和学习最优策略，非常适合处理本模型的优化问题。在训练过程中，首先要对DDPG算法的演员网络和评论家网络进行参数初始化。演员网络负责根据当前状态输出动作，即确定各机组的负荷调整量；评论家网络则用于评估演员网络输出动作的价值，为演员网络的学习提供指导。同时，初始化经验回放缓冲区，用于存储智能体与环境交互过程中产生的状态转移信息。接着智能体与火电厂机组仿真环境开始交互。在每个时间步，智能体依据当前状态，利用演员网络选择一个动作并执行。环境根据该动作产生新的状态和相应的奖励，并将这次状态转移信息存入经验回放缓冲区。然后，从缓冲区中随机采样一批状态转移信息，用于更新演员网络和评论家网络的参数。通过不断地迭代训练，智能体在与仿真环境的交互中积累经验，逐渐学习到最优的负荷分配策略，实现火电厂机组负荷的高效、合理分配。

## 4 仿真实验与结果分析

### 4.1 实验环境设置

为了验证基于深度强化学习的火电厂机组负荷优化分配模型的有效性，搭建了仿真实验平台。仿真环境模拟了一个包含多台机组的火电厂，机组的参数根据实际火电厂机组进行设置。实验中，考虑了不同的电力系统负荷需求变化场景，以测试模型在不同工况下的性能。

### 4.2 实验结果与分析

将基于DDPG算法的深度强化学习模型与传统数学规划算法和遗传算法进行对比实验。实验结果表明，在静态负荷需求场景下，三种算法都能够找到近似最优的

负荷分配方案，但深度强化学习算法的计算时间明显短于传统算法，具有更高的实时性；在动态负荷需求变化场景下，深度强化学习算法表现出显著的优势。它能够根据实时负荷需求和机组运行状态，动态调整机组负荷分配，始终保持较低的燃料成本。而传统算法由于计算复杂度高，难以快速响应负荷需求的变化，导致燃料成本较高<sup>[4]</sup>。深度强化学习算法在处理机组故障等突发情况时，也能够迅速调整负荷分配策略，保证电厂的稳定运行；通过进一步分析实验数据发现，深度强化学习算法在训练过程中逐渐学习到了机组的经济运行特性和负荷分配规律，能够根据不同机组的性能优势合理分配负荷，从而实现整体燃料成本的最小化。同时，算法的奖励函数设计有效地引导了学习过程，使得算法在满足约束条件的前提下，不断优化负荷分配策略。

### 结束语

本研究将深度强化学习应用于火电厂机组负荷优化分配，通过构建模型并开展仿真实验，验证了该方法的可行性和有效性。实验显示，深度强化学习算法在处理复杂多变的火电厂运行工况时，展现出强大的自适应和优化能力，不仅能快速响应负荷需求变化，降低燃料成本，还能在机组故障等突发情况下保证电厂稳定运行。未来可进一步优化模型，提升算法性能，同时探索在实际火电厂中的应用，为火电厂的智能化运行和节能减排提供有力支持。

### 参考文献

- [1]翁存兴,王晓宁,刘碧峰.基于深度学习的火电厂机组负荷调度自动控制方法[J].信息与电脑(理论版),2023,35(01):83-85.
- [2]徐亮亮.火电厂集控运行节能降耗技术探讨[J].节能与环保,2023,(09):53-55.
- [3]李越,漆锐,杨盼.燃煤锅炉高温腐蚀及燃烧损伤成因和优化分析[J].电站系统工程,2025,41(02):27-29+32.
- [4]王誉仁.火力发电厂锅炉燃烧效率与环保性能优化措施[J].造纸装备及材料,2025,54(02):97-99.