

DOI: 10.19666/j.rlfed.202307118

基于深度强化学习的火电机组制粉系统 自启停智能决策

蔡佳辰¹, 李 军¹, 高 明², 高 林¹, 高耀焜¹, 昌 鹏¹

(1.西安热工研究院有限公司, 陕西 西安 710054;

2.陕西延长石油富县发电有限公司, 陕西 延安 727502)

[摘 要] 针对目前人工决策基础上的制粉系统一键启停技术存在决策主观经验性强、操盘劳动强度大、节能优化潜力难以发掘等问题, 提出了一种综合考虑制粉系统能耗与机组负荷跟踪性能的制粉系统启停决策评价模型, 以安全引入网调负荷计划指令信号作为输入, 研究了基于深度强化学习的制粉系统自启停智能决策方法, 开发了制粉系统自启停决策闭环控制系统。研究结果通过仿真验证, 并已在某超超临界 1 000 MW 机组常用磨煤机上成功应用, 节能降耗效果显著。研究结果可为火电机组少人、无人化运行技术提供有效借鉴。

[关 键 词] 深度强化学习; 制粉系统; 自启停控制; 智能决策

[引用本文格式] 蔡佳辰, 李军, 高明, 等. 基于深度强化学习的火电机组制粉系统自启停智能决策[J]. 热力发电, 2024, 53(3): 146-152. CAI Jiachen, LI Jun, GAO Ming, et al. Intelligent decision-making of start-up and shutdown for coal milling system in thermal power plants based on deep reinforcement learning[J]. Thermal Power Generation, 2024, 53(3): 146-152.

Intelligent decision-making of start-up and shutdown for coal milling system in thermal power plants based on deep reinforcement learning

CAI Jiachen¹, LI Jun¹, GAO Ming², GAO Lin¹, GAO Yaokui¹, CHANG Peng¹

(1.Xi'an Thermal Power Research Institute Co., Ltd., Xi'an 710054, China;

2.Shaanxi Yanchang Petroleum Fuxian Power Generation Co., Ltd., Yan'an 727502, China)

Abstract: A comprehensive evaluation model for the start-up and shutdown decision-making of the milling system, taking into account the energy consumption and tracking performance of the unit load, has been proposed to address issues such as subjective decision-making based on manual experience, high labor intensity in operation, and difficulty in exploring energy-saving optimization potential. This model safely incorporates the grid load scheduling command signal as input. Furthermore, a milling system start-stop intelligent decision-making method based on deep reinforcement learning has been studied, and a closed-loop control system for the automatic start-stop of the milling system has been developed. The research results have been verified through simulation and successfully applied to a commonly used coal milling system in a certain ultra-supercritical 1 000 MW unit, achieving energy savings. The findings of this study can provide effective reference for the development of unmanned or minimally manned operation techniques for thermal power units.

Key words: deep reinforcement learning; milling system; autonomous start-up and shutdown control; intelligent decision-making

随着智能发电概念的提出, 对火电机组运行控制提出了更高的要求^[1-2]。制粉系统作为火电机组的重要辅机, 在机组调峰调频过程中启停频繁^[3-4]、人工操作强度大, 并存在启停决策主观经验性强, 节

收稿日期: 2023-07-25

基金项目: 国家重点研发计划项目(2022YFB4100700); 陕西省重点研发计划项目(2023-YBGY-274)

Supported by: National Key Research and Development Program (2022YFB4100700); Key Research and Development Program in Shaanxi Province (2023-YBGY-274)

第一作者简介: 蔡佳辰(1999), 男, 硕士研究生, 主要研究方向为智慧电站控制技术, 467974006@qq.com。

能优化潜力未深度挖掘等问题。因此,急需研究制粉系统自主决策启停技术,突破制约火电机组运行智能化水平的关键技术瓶颈,为机组少人、无人化运行提供有力支撑。

目前,国内外研究主要聚焦在制粉系统一键启停方面。赵俊杰等^[5]基于 ICS 智能发电运行控制平台,开发了 APS 程序自动启停控制,实现了重要辅机的一键启停功能。於晓博等^[6]设计了暖磨子功能组、磨煤机组(磨组)磨组的自动启子功能组、磨组的自动停运子功能组,实现了制粉系统一键启停管理。王志杰等^[7]建立了制粉系统的启停专家知识库,设计了双进双出制粉系统的自启停步序,实现了该类制粉系统启停过程的全程自动控制。李晓燕等^[8]对原有的顺控启停方案进行优化,解决了模拟量与顺序控制结合的难题,运用柔性顺控技术实现了制粉系统的一键启停功能。以上研究均以人工决策基础上的制粉系统一键启停技术为主,有关制粉系统自主决策启停的相关研究还未见报道。而深度强化学习(deep reinforcement learning, DRL)模拟人类学习的过程,通过不断交互来寻求最优的策略,给学者在智能决策研究上提供了许多启发^[9-10]。王永志^[11]基于深度强化学习算法,研究了机器人智能抓取系统,使得机器人能够获得更高的装配效率。董瑶等^[12]提出了基于竞争网络结构的改进深度双 Q 网络方法,使机器人能够更快地适应未知环境,在未知的复杂环境中可以更好地获取最优路径。王珂等^[13]基于深度强化学习算法,通过构建奖励函数,对电网调度问题进行智能决策,验证了所提算法和技术的有效性。梁涛等^[14]建立了离网型可再生能源大规模制氢系统,基于深度强化学习算法实现在线能量调度,使得可再生能源的消纳能力变强。以上研究均为 DRL 算法在各个领域的应用,其在智能决策方面表现出了优异的性能,但制粉系统启停决策影响因素较多,尚无成熟的决策评价模型可供借鉴。

本文以中速磨煤机正压直吹式制粉系统为研究对象,首先,安全引入网调负荷计划指令信号,解决了启停决策的核心参考输入难题;然后在此基础上,建立了一种兼顾节能与负荷跟踪能力的制粉系统启停决策评价模型;再基于该评价模型和深度强化学习算法,研究了制粉系统自启停决策方法,开发了制粉系统自启停决策闭环控制系统,并对所提算法与技术仿真模拟;最后,在某 1 000 MW 火电机组成功应用,实现节能效益的同时,为火电机组少人、无人化运行技术提供了有效借鉴。

1 制粉系统启停决策评价数学建模

1.1 制粉系统启停决策影响分析

典型的中速磨煤机正压直吹式制粉系统如图 1 所示。制粉系统主要的耗电设备为磨煤机和一次风机,二者耗电量占制粉系统耗电量 90% 以上^[15]。尤其在磨煤机启动时电流高达 20 A,但此时磨煤机的碾磨出力很低而单耗非常高,因此在满足电网负荷调度要求的前提下,采用尽可能少的磨组是最经济的。

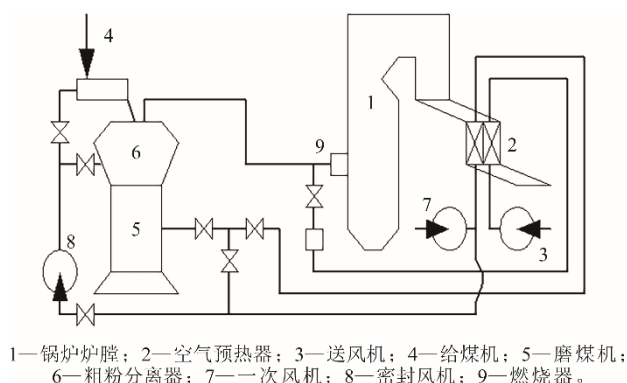


图 1 中速磨煤机正压直吹式制粉系统示意

Fig.1 Schematic of a medium speed mill, positive-pressure, direct-fired milling system

在实际运行过程中,当每台磨煤机都接近满/低负荷运行时,面对升/降负荷指令,磨组调节能力不足,将无法跟踪电网负荷指令。考虑到负荷计划模式下,电网调度计划指令能够提前预知,而又存在安全分区隔离,无法跨区域传输的难题,本文采用视频识别的方式将网调计划指令引入安全可靠 I 区,使制粉系统能根据当前磨组调节能力与目标负荷超前决策。

1.2 制粉系统启停决策建模

为了科学地评价制粉系统启停决策性能,本文首先建立制粉系统启停决策模型,该模型综合考虑了启停过程经济性和负荷跟踪能力 2 个维度,将运行经济性函数与负荷跟踪能力惩罚函数作为评价制粉系统启停决策效果的回报函数。

1.2.1 经济性函数

通常,采用磨煤单耗来定量评估磨煤机运行的经济性。根据理论计算和现场运行经验可知,磨煤机在额定出力范围内运行时,磨煤单耗随磨煤机出力的增加而降低,即磨煤机出力越大时越省电;当给煤量低于最小出力 B_{NCR} 时,磨煤机容易发生部件

碰磨, 从而引起振动和磨损, 对应经济性快速降低; 当给煤量超过磨煤机最大出力 B_{MCR} 时, 磨煤机容易出现堵磨, 同样对应经济性的快速降低; 此外, 磨煤机暖磨及结束后的备用状态将产生额外的热风损失和燃烧损失。因此, 磨煤机经济性函数可表示为:

$$r_i(B_i, C_i) = \begin{cases} Q_0 & \text{备用} \\ Q_1(B_i) & 0 < B_i < B_{NCR} \\ Q_2(B_i \cdot C_i(B_i)) & B_{NCR} \leq B_i < B_{MCR} \\ Q_3(B_i) & B_i \geq B_{MCR} \end{cases} \quad (1)$$

$$R(B, C) = \sum_{i=1}^v r_i(B_i, C_i) \quad (2)$$

式中: B_i 为第 i 台磨煤机的给煤量, t/h; C_i 为第 i 台磨煤机的磨煤单耗函数, 取自磨煤机单耗试验曲线; v 为运行磨煤机的台数; Q 为基于制粉系统能耗的归一化函数。

本文以自主决策启停方法及其应用为核心目标, 对该经济性评价函数进行了简化。各归一化经济性函数的表达式为:

$$\begin{aligned} Q_0 &= 0 \\ Q_1(B_i) &= \left(\frac{E_{\max}}{v} - B_{NCR} \cdot C_i(B_{NCR}) \right) / E_{\max} \\ Q_2(B_i \cdot C_i(B_i)) &= \left(\frac{E_{\max}}{v} - B_i \cdot C_i(B_i) \right) / E_{\max} \quad (3) \\ Q_3(B_i) &= \left(\frac{E_{\max}}{v} - k_M (B_i - B_{MCR}) - B_{MCR} \cdot C_i(B_{MCR}) \right) / E_{\max} \end{aligned}$$

式中: E_{\max} 为 6 台磨煤机均运行时的最大能耗, 暖磨及备用状态的热风损失和燃烧损失被忽略; 当 B_i 低于最小出力 B_{NCR} 时, 磨煤机的运行能耗假设为最小出力 B_{NCR} 时的能耗; 当 B_i 高于最大出力 B_{MCR} 时, 以最大出力时的能耗斜率 k_M 进行线性扩展。此外, 由于实际磨煤机最大出力和最小出力受燃煤可磨性等煤质参数影响而并非常数, 因此, 本文实际采用的最大和最小连续出力煤量均保留一定的安全裕度 ε 。

1.2.2 负荷跟踪能力惩罚函数

从前述磨煤机经济性分析可知, 理论上最经济运行点应该为磨煤机最大出力点, 但磨煤机在最大出力点运行时, 其潜在的升负荷能力受到的限制越大, 这样在机组突然升负荷时, 将导致该磨煤机无法进一步提高出力而限制机组负荷的进一步提升, 从而表现为负荷跟踪能力不足; 反之在磨煤机接近最低煤量运行时, 降负荷能力将受到限制。因此制粉系统启停决策需要根据机组负荷升降的预期进

行一定的预判, 从而尽可能及时地进行制粉系统的启停操作, 既能满足负荷响应的需求, 又能兼顾磨煤机的整体运行经济性。

综上所述, 需要对制粉系统启停决策引起的机组负荷跟踪能力的不足进行惩罚, 从而最终建立启停决策的评价体系。

然而, 实际机组的负荷跟踪能力受超温、超压等诸多因素限制。为了避免其它因素的干扰, 本文以磨煤机出力指令接近甚至超过最大出力或低于最小出力的程度作为惩罚函数自变量。

$$p_i(D_i) = \begin{cases} G_1(D_i - B_{NCR}) & , D_i < B_{NCR} \\ 0 & , B_{NCR} \leq D_i < B_{MCR} \\ G_3(D_i - B_{MCR}) & , D_i \geq B_{MCR} \end{cases} \quad (4)$$

$$P(D) = \sum_{i=1}^v p_i(D_i) \quad (5)$$

式中: D_i 为按照负荷指令计算的磨煤机 i 的给煤量指令; G_i 为根据归一化经济性函数设置的惩罚曲线, 本文为简化问题在控制逻辑中实施的复杂程度, 采用线性表达式, $G_j(x) = k_j(x)$, 对于试验机组采用的中速磨煤机, 取 $k_1 = k_3 = 0.5$ 。

为方便对实际评价结果进行反馈和调整, 采用加权的方式形成最终的回报函数:

$$\text{return} = c_1 R(B, C) - c_2 P(D) \quad (6)$$

式中: c_1 、 c_2 为加权调整系数, 取 $c_1 = c_2 = 1$ 。

这样不同操作过程的效果可使用给定时间窗口内该函数的积分进行决策过程的优劣评价。

2 基于深度强化学习的制粉系统自启停决策方法

2.1 决策架构

图 2 给出了基于深度强化学习的制粉系统自启停决策架构, 整体架构由离线训练、在线决策与用户反馈 3 部分组成。

离线训练模块 首先, 基于设备模型、机组运行状态、制粉系统结构等信息构建仿真环境, 模拟机组运行的真实环境; 其次, 结合仿真环境与历史数据, 从中提取电网调度指令、总给煤量、运行磨煤机台数等数据做数据预处理, 输入离线策略网络 (离线训练智能体), 基于 DRL 算法进行训练, 生成制粉系统启停计划; 最终, 将启停计划送入仿真环境, 结合制粉系统启停决策评价模型计算运行代价, 并反馈更新离线训练智能体的相关参数。离线训练智能体与仿真环境迭代交互, 直到其参数趋于稳定。

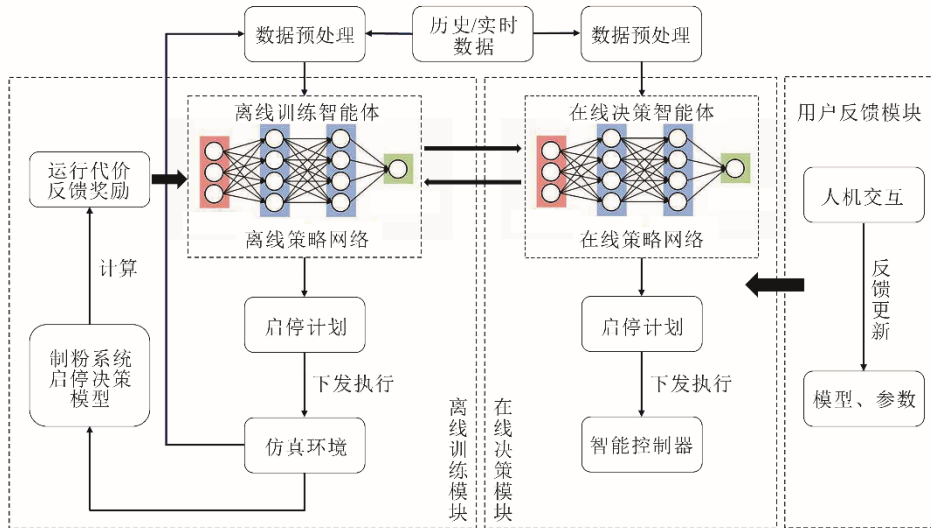


图2 基于深度强化学习的制粉系统自启停决策架构

Fig.2 Deep reinforcement learning-based self-start-stop decision architecture for milling system

在线决策模块 会从实时数据中提取当前电网调度指令、总给煤量、运行磨煤机台数等数据做数据预处理，输入在线策略网络（在线决策智能体），生成当前状态下制粉系统的启停计划，经过确认后送入智能控制器执行计划。离线训练智能体与在线决策智能体交互更新。

用户反馈模块 通过人机交互^[16]方式传达于运行人员，运行人员结合历史经验与现场条件，对决策结果做出评价、提出修改意见，并反馈给智能体，从而更新模型与参数。

2.2 DRL 算法

DRL3 1 算法是将深度学习的卷积神经网络对数据的处理感知能力与强化学习的决策能力结合的一种方法。DQN（deep Q network）算法^[17-18]是一种代表性的 DRL 算法，其主要特点包括经验回放机制和目标网络。通过经验回放机制，可以减少样本之间的相关性，提高训练的稳定性。而目标网络则是在原有神经网络的基础上建立一个结构完全相同的网络，以增加智能体学习的稳定性。然而，DQN 算法不能适用于连续动作空间，并且其训练收敛速度较慢。

深度确定性策略梯度（deep deterministic policy gradient, DDPG）算法^[19-21]是在 DQN 算法的基础上进行了重大突破。其将动作空间升级为连续变量，并引入了演员-评论家（actor-critic）网络^[22-23]，具体如图3所示，使智能体具备一定的探索能力。DDPG 算法也继承了 DQN 算法的经验回放和独立目标网络等优点，这有助于减少数据的相关性，并降低训练的难度。

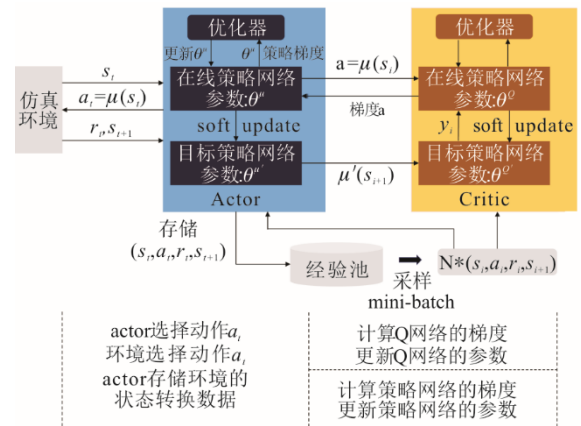


图3 DDPG 算法

Fig.3 Deep deterministic policy gradient algorithm

将 DDPG 算法用于制粉系统启停决策有 2 个难点，即确定状态之间的交互程度与更新策略网络。以确定状态之间的交互程度为出发点，建立状态-动作矩阵，状态 S_t 主要为可控的机组状态参数，动作 A_t 主要为制粉系统的出力与负荷的变化。以此构建状态-动作矩阵为：

$$F = \begin{bmatrix} f_{11} & \dots & f_{1n} \\ \dots & \dots & \dots \\ f_{n1} & \dots & f_{nm} \end{bmatrix} \quad (7)$$

状态-动作矩阵的每行表示智能体的状态，每列表示智能体的动作，矩阵中的每个元素表示执行动作后，状态与状态之间的关联度。

DDPG 算法的最终优化目标是通过策略梯度方法来更新策略函数。通过训练，持续更新策略函数中神经网络的参数，从而指导 Actor 选择最佳动作，

即最优的启停磨计划。

$$\nabla \theta^\mu J(\theta^\mu) \cong \frac{1}{N} \sum_i \nabla \theta^\mu \mu(s_i | \theta^\mu) \nabla_a Q(s_i, a | \theta^Q) |_{a=\mu(s_i)}$$

$$\theta^{Q'} \leftarrow \tau \theta^Q + (1 - \tau) \theta^{Q'}$$

$$\theta^{\mu'} \leftarrow \tau \theta^\mu + (1 - \tau) \theta^{\mu'}$$
(8)

式中： θ^μ 为策略网络参数； θ^Q 为 Q 网络参数； μ 为策略； N 为采样出的序列集个数； τ 为学习率。

3 系统开发与仿真验证

3.1 系统开发

本文从制粉系统自启停决策的实际应用需求出发，开发了制粉系统自启停决策闭环控制系统如图 4 所示。系统由网络摄像头、超融合服务器、智能控制器等设备组成。

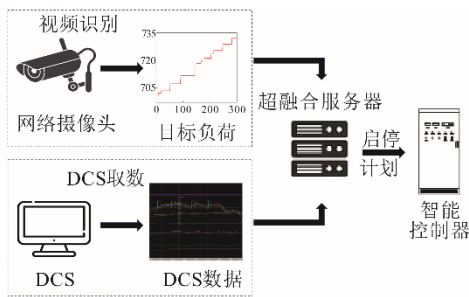


图 4 制粉系统自启停决策闭环控制系统

Fig.4 Reinforcement learning-based closed-loop control system for self-start-stop decision in milling system

首先，通过网络摄像头，对电网调度指令等曲线视频识别；其次，采用串口通信的方式将识别信号安全引入超融合服务器^[24]，通过在线决策智能体生成启停磨煤机的计划；再次，超融合服务器将决策结果送入智能控制器；最后，智能控制器发出智能启停磨指令，并调节相关开关量、模拟量参数^[25-26]，实现制粉系统自启停决策、控制的智能化和无人化。

3.2 实验设置

本文以某超超临界 1 000 MW 机组为对象，锅炉制粉系统采用中速磨煤机正压直吹式制粉系统，共配置 6 台中速磨煤机，每台磨的常规出力范围 30 ~75 t/h，算法超参数见表 1。

表 1 算法超参数

Tab.1 Hyperparameters of the algorithm

| 项目 | 含义 | 数值 |
|----------------|-------|--------------------|
| learning_rate | 学习率 | 1×10^{-3} |
| gamma | 折扣因子 | 0.99 |
| train_episodes | 训练次数 | 10 000 |
| capacity | 经验池大小 | 50 000 |

3.3 仿真结果

3.3.1 离线训练结果

训练数据采用历史数据中 90 日制粉系统的启停数据作为数据集，选取 70% 的数据作为训练集、15% 的数据作为验证集、15% 的数据作为测试集，训练的均方误差 (mean-square error, MSE) 曲线如图 5 所示。在第 7 000 次迭代时逐渐稳定并趋于最佳验证性能，并在第 10 000 次迭代时达到最佳验证性能。

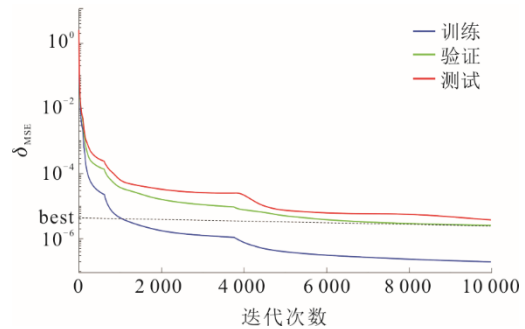
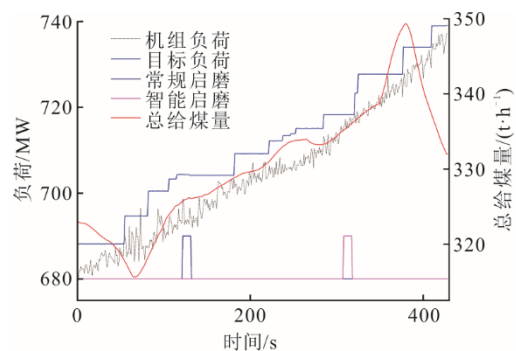


图 5 训练数据的 δ_{MSE} 曲线
Fig.5 δ_{MSE} curves for data training

3.3.2 在线决策与结果分析

以磨煤机 A、B、D、E、F 运行，智能启磨磨煤机 C 为例对智能决策效果进行评估。将机组当前运行状态、电网调度指令、总给煤量等参数输入在线决策智能体中，得到智能启磨煤机 (启磨) 指令并决策启磨，与常规启磨对比，效果如图 6 所示。

常规启磨指令在机组负荷为 696.4 MW、总给煤量在 327.43 t/h 时决策启磨，智能启磨指令在机组负荷为 721.4 MW、总给煤量在 338.68 t/h 时决策启磨，整个决策启磨过程无人干预，在 5 min 内完成决策发出智能启磨指令，快速完成启磨、并磨操作。由于智能启磨较常规启磨，每台磨煤机都接近较佳出力运行，在常规启磨启动 6 台磨煤机时，智能启磨用 5 台磨煤机仍能满足机组要求出力且能跟随目标负荷变化，故在 2 种方式的时间差内，智能启磨方式就节省了可观的制粉系统电耗。



a) 负荷变化趋势

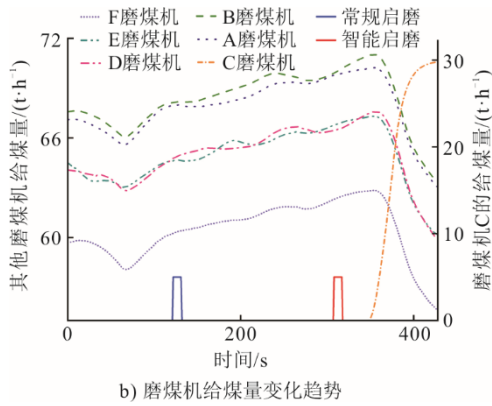


图6 智能启磨与常规启磨效果对比
Fig.6 Comparison between the intelligent start-up grinding and the conventional start-up grinding

以该机组制粉系统运行 24 h, 升负荷 2 次, 降负荷 2 次, 启磨煤机 3 台, 停磨煤机 4 台作为参考, 对比分析了常规启磨与智能启磨, 常规启磨耗电量 $F_C=6.03 \times 10^4$ kW·h, 智能启磨耗电量 $F_I=5.53 \times 10^4$ kW·h, 相比常规启磨每天节约约 7.19%。

3.4 实际应用

图 7 给出了将整套系统应用于某 1 000 MW 火电机组的工程实例, 选取 1 天 24 h 的负荷变化情况作为参考, 整个过程中磨煤机 C、D 作为频繁启停的磨煤机组, 调试并投入了智能决策启停控制, 其余磨煤机以智能启停建议的开环形式投运。

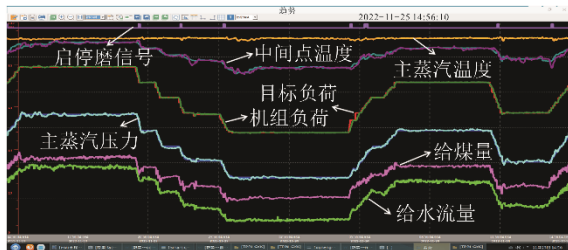


图7 实际应用
Fig.7 Practical application

在第 1 阶段, 机组初始负荷为 825.3 MW, 当负荷降低至 723.1 MW 时智能停磨煤机 C, 负荷降低至 612.2 MW 智能停磨煤机 D, 负荷为 508.6 MW 时, 给出智能停磨建议; 在第 2 阶段, 当机组负荷升高至 495.4 MW 时给出智能启磨建议, 当负荷升高至 613.8 MW 时, 智能启磨煤机 C。

通过智能启停磨煤机指令与建议, 机组负荷 (图中深绿色曲线) 能够跟随目标负荷 (图中红色曲线) 变化, 主蒸汽压力、给煤量、给水流量也能响应机组负荷的变化, 在整个过程中, 主蒸汽温度保持稳定, 控制效果良好。

4 结论

本文以制粉系统自启停决策为目标, 从实际机组出发, 提出了基于 DRL 算法的整套系统架构并进行初步探索应用, 主要研究结果如下。

1) 根据制粉系统启停决策影响分析, 通过安全引入网调指令作为输入, 利用经济性函数与负荷跟踪能力惩罚函数, 构建了一套兼顾节能与负荷跟踪能力的制粉系统启停决策评价模型, 科学地评价启停决策的效果。

2) 提出了制粉系统自启停决策架构, 由离线训练模块、在线决策模块与用户反馈模块 3 部分组成。该决策架构能够基于历史数据进行训练, 并根据在线参数得到最优启停决策。

3) 开发了制粉系统自启停决策闭环控制系统。将视频识别网调计划指令与 DCS 制粉系统运行数据一并接入超融合服务器计算, 决策形成的启停磨煤机指令送入智能控制器执行, 达到整个启停决策过程无人化、智能化。

研究结果通过仿真验证, 在某超超临界 1 000 MW 机组常用磨组上实现了应用, 实现节能降耗的同时, 为火电机组智慧化发展提供宝贵经验。

[参考文献]

[1] 刘吉臻, 胡勇, 曾德良, 等. 智能发电厂的架构及特征[J]. 中国电机工程学报, 2017, 37(22): 6463-6470.
LIU Jizhen, HU Yong, ZENG Deliang, et al. Architecture and feature of smart power generation[J]. Proceedings of the CSEE, 2017, 37(22): 6463-6470.

[2] 华志刚, 郭荣, 崔希, 等. 火电智慧电厂技术路线探讨与研究[J]. 热力发电, 2019, 48(10): 8-14.
HUA Zhigang, GUO Rong, CUI Xi, et al. Discussion and study on technical route of smart thermal power plant[J]. Thermal Power Generation, 2019, 48(10): 8-14.

[3] 方彦军, 贺瑶, 秦晓洁. 直流锅炉快速变负荷磨煤机启停优化[J]. 热力发电, 2013, 42(9): 10-15.
FANG Yanjun, HE Yao, QIN Xiaojie. Optimization of start and stop of coal mills during once through boiler's fast load varying[J]. Thermal Power Generation, 2013, 42(9): 10-15.

[4] HUANG Y, CHEN Q, ZHANG L, et al. Dynamic optimization of variable load process for combined heat and power unit based on sequential quadratic programming and interior point method alternating solution method[J]. Processes, 2023, 11(6): 311-314.

[5] 赵俊杰, 冯树臣, 田景奇, 等. 基于燃煤智能发电 ICS 的主辅机一键启停 APS 技术应用分析[J]. 能源科技, 2021, 19(1): 41-45.
ZHAO Junjie, FENG Shuchen, TIAN Jingqi, et al. Application analysis of one-button start-up & shut-down technology of aps for main and auxiliary machines based on ICS for coal-fired power generation[J]. Energy Technology, 2021, 19(1): 41-45.

- [6] 於晓博. 大型火电机组一键启停方案设计研究[D]. 保定: 华北电力大学(保定), 2017: 1.
YU Xiaobo. Design and research of one-key start and stop scheme for large thermal power unit[D]. Baoding: North China Electric Power University (Baoding), 2017: 1.
- [7] 王志杰, 彭梁, 朱晓星, 等. 基于专家知识库的双进双出制粉系统自启停技术[J]. 中国电力, 2018, 51(1): 133-138.
WANG Zhijie, PENG Liang, ZHU Xiaoxing, et al. The automatic startup and shutdown technology for double inlet and outlet coal-pulverizing system based on expert knowledge base[J]. Electric Power, 2018, 51(1): 133-138.
- [8] 李晓燕, 郑卫东, 熊志成, 等. 基于柔性顺控的制粉系统一键启停控制技术在1 000 MW超超临界火电机组的实践与应用[J]. 浙江电力, 2019, 38(12): 39-43.
LI Xiaoyan, ZHENG Weidong, XIONG Zhicheng, et al. Application and practice of one-click start-stop control technology for pulverized coal preparation system based on flexible sequential control in 1 000 MW ultra-supercritical thermal power generating units[J]. Zhejiang Electric Power, 2019, 38(12): 39-43.
- [9] CHEN X C, YAO L N, MCAULEY J L, et al. Deep reinforcement learning in recommender systems: a survey and new perspectives[J]. Knowledge-Based Systems, 2023, 264: 1-19.
- [10] 刘全, 翟建伟, 章宗长, 等. 深度强化学习综述[J]. 计算机学报, 2018, 41(1): 1-27.
LIU Quan, ZHAI Jianwei, ZHANG Zongchang, et al. A survey on reinforcement learning[J]. Chinese Journal of Computers, 2018, 41(1): 1-27.
- [11] 王永志. 基于深度强化学习的机器人抓取及智能装配研究[D]. 沈阳: 沈阳工业大学, 2022: 1.
WANG Yongzhi. Research on robotic grasping and intelligent assembly based on deep reinforcement learning[D]. Shenyang: Shenyang University of Technology, 2022: 1.
- [12] 董瑶, 葛莹莹, 郭鸿湧, 等. 基于深度强化学习的移动机器人路径规划[J]. 计算机工程与应用, 2019, 55(13): 15-19.
DONG Yao, GE Yingying, GUO Hongyong, et al. Path planning for mobile robot based on deep reinforcement learning[J]. Computer Engineering and Applications, 2019, 55(13): 15-19.
- [13] 王珂, 姚建国, 余佩遥, 等. 基于深度强化学习的电网前瞻调度智能决策架构及关键技术初探[J]. 中国电机工程学报, 2022, 42(15): 5430-5439.
WANG Ke, YAO Jianguo, YU Peiyao, et al. Architecture and key technologies of intelligent decision-making of power grid look-ahead dispatch based on deep reinforcement learning[J]. Proceedings of the CSEE, 2022, 42(15): 5430-5439.
- [14] 梁涛, 刘伟, 曹欣, 等. 基于深度确定性策略梯度算法的可再生能源大规模制氢系统能量调度[J/OL]. 电网技术: 1-16. (2023-04-11)[2023-07-23]. <https://doi.org/10.13335/j.1000-3673.pst.2023.0211>.
LIANG Tao, LIU Wei, CAO Xin, et al. Research on energy scheduling of renewable energy large-scale hydrogen production system based on deep deterministic strategy gradient algorithm[J/OL]. Power System Technology, 1-16. (2023-04-11)[2023-07-23]. <https://doi.org/10.13335/j.1000-3673.pst.2023.0211>.
- [15] 谭天宇. 计及经济性与快速性的厂级负荷优化分配研究[D]. 北京: 华北电力大学, 2019: 1.
TAN Tianyu. Research on optimal load dispatch of thermal power plant considering economy and speediness[D]. Beijing: North China Electric Power University, 2019: 1.
- [16] 张凤军, 戴国忠, 彭晓兰. 虚拟现实的人机交互综述[J]. 中国科学: 信息科学, 2016, 46(12): 1711-1736.
ZHANG Fengjun, DAI Guozhong, PENG Xiaolan. A survey on human-computer interaction in virtual reality[J]. Scientia Sinica (Informationis), 2016, 46(12): 1711-1736.
- [17] WANG Y H, CHEN L, ZHOU H, et al. Flexible transmission network expansion planning based on DQN algorithm[J]. Energies, 2021, 14(7): 1-21.
- [18] 董永峰, 杨琛, 董瑶, 等. 基于改进的DQN机器人路径规划[J]. 计算机工程与设计, 2021, 42(2): 552-558.
DONG Yongfeng, YANG Chen, DONG Yao, et al. Robot path planning based on improved DQN[J]. Computer Engineering and Design, 2021, 42(2): 552-558.
- [19] 刘建伟, 高峰, 罗雄麟. 基于值函数和策略梯度的深度强化学习综述[J]. 计算机学报, 2019, 42(6): 1406-1438.
LIU Jianwei, GAO Feng, LUO Xionglin. Survey of deep reinforcement learning based on value function and policy gradient[J]. Chinese Journal of Computers, 2019, 42(6): 1406-1438.
- [20] 刘朝阳, 穆朝絮, 孙长银. 深度强化学习算法与应用研究现状综述[J]. 智能科学与技术学报, 2020, 2(4): 314-326.
LIU Chaoyang, MU Chaoxu, SUN Changyin. An overview on algorithms and applications of deep reinforcement learning[J]. Chinese Journal of Intelligent Science and Technology, 2020, 2(4): 314-326.
- [21] HOMAYUN K, LUEN M T, CHOON Y C. Twin delayed DDPG based dynamic power allocation for mobility in IoRT[J]. Journal of Communications Software and Systems, 2023, 19(1): 19-29.
- [22] LI X, WU G, SUN Z, et al. Research on multi-agent D2D communication resource allocation algorithm based on A2C[J]. Electronics, 2023, 12(2): 360-360.
- [23] GERPOTT F, LANG S, REGGELIN T, et al. Integration of the A2C algorithm for production scheduling in a two-stage hybrid flow shop environment[J]. Procedia Computer Science, 2022, 200: 585-594.
- [24] 杨新民, 高海东, 陈丰. 数字化电厂概念的解析及探讨[J]. 热力发电, 2015, 44(5): 98-101.
YANG Xinmin, GAO Haidong, CHEN Feng. Discussions on concept of digital power plant[J]. Thermal Power Generation, 2015, 44(5): 98-101.
- [25] 高耀焄, 王林, 高海东, 等. 火电厂智能控制系统体系架构及关键技术[J]. 热力发电, 2022, 51(3): 166-174.
GAO Yaokui, WANG Lin, GAO Haidong, et al. Architecture and key technology of intelligent control system in thermal power plant[J]. Thermal Power Generation, 2022, 51(3): 166-174.
- [26] 马天霆. 磨煤机模糊与模糊PI复合控制系统[J]. 热力发电, 2015, 44(10): 52-57.
MA Tianting. A coal mill control system combining fuzzy control with fuzzy PID control[J]. Thermal Power Generation, 2015, 44(10): 52-57.

(责任编辑 邓玲惠)