

复杂工况下选择性催化还原脱硝系统的迁移强化学习控制

孙小明, 彭晨[†], 程传良

(上海大学 机电工程与自动化学院, 上海 200444)

摘要: 针对复杂工况下选择性催化还原(SCR)系统难以实现精确脱硝控制的问题, 本文提出一种基于迁移强化学习的智能控制方法. 首先根据机组负荷的变化将整体运行过程划分为不同阶段. 然后训练了强化学习控制器以分别学习各个阶段的不同特征, 从而实现了变工况下SCR脱硝系统的精确控制. 此外, 借鉴了迁移学习的思路以应对预料之外的未知工况, 避免了因工况未知所造成的不利影响. 最后将训练好的控制器用于实际SCR脱硝系统的控制中, 实验结果表明所提方法可以有效地控制复杂工况下燃煤机组NO_x的排放量, 为复杂工况下SCR脱硝系统的智能控制提供了借鉴.

关键词: SCR脱硝系统; 变工况; 未知工况; 强化学习; 迁移学习

引用格式: 孙小明, 彭晨, 程传良. 复杂工况下选择性催化还原脱硝系统的迁移强化学习控制. 控制理论与应用, 2024, 41(3): 496 – 501

DOI: 10.7641/CTA.2023.21030

Transfer reinforcement learning control for a selective catalytic reduction denitration system under complex conditions

SUN Xiao-ming, PENG Chen[†], CHENG Chuan-liang

(School of Mechatronic Engineering and Automation, Shanghai University, Shanghai 200444, China)

Abstract: Aiming at the problem that selective catalytic reduction (SCR) system is difficult to achieve precise denitration control performance under complex working conditions, an intelligent control method based on transfer reinforcement learning is proposed in this paper. The overall operation process is firstly divided into different stages according to the changes of unit load. Then a reinforcement learning controller is trained to learn different characteristics of each stage, respectively, so as to realize accurate control of the SCR denitration system under variable working conditions. In addition, the idea of transfer learning is used for reference to deal with unexpected unknown working conditions and avoid adverse effects caused by unknown working conditions. Finally, the trained controller is applied to the control of an actual SCR denitration system. Experimental results show that the proposed method can effectively control NO_x emissions of a coal-fired power unit under complex working conditions, and provide an idea for intelligent control of the SCR denitration system under complex working conditions.

Key words: SCR denitration system; variable working condition; unknown working condition; reinforcement learning; transfer learning

Citation: SUN Xiaoming, PENG Chen, CHENG Chuanliang. Transfer reinforcement learning control for a selective catalytic reduction denitration system under complex conditions. *Control Theory & Applications*, 2024, 41(3): 496 – 501

1 引言

节能减排政策的持续施行一方面对石油、冶金、电力等行业提出了越来越高的要求, 另一方面也促进了这些行业的技术进步, 促使它们在达到预期产量的同时, 不断地升级传统技术, 从而尽可能地降低能耗和减少排放, 电力行业尤为如此^[1]. 据报道, 我国的电力结构以火力发电为主, 并且这一结构在短期内不会

有太大的变化^[2-3]. 对于火力发电来说, 稳定是其最大的优点, 然而, 因其主要以煤、石油、天然气等化石燃料为能量来源, 运行过程中造成的污染物排放不可避免, 而污染物排放带来的环境问题已经十分严峻且亟需解决, 对于污染物的准确及时控制刻不容缓.

我国的火电厂大多以煤为燃料, 在燃煤电厂中, NO_x, SO₂和粉尘是几种主要的大气污染物, 它们极大

收稿日期: 2022-11-22; 录用日期: 2023-10-28.

[†]通信作者. E-mail: c.peng@i.shu.edu.cn.

本文责任编辑: 周平.

国家自然科学基金重点项目(61833011)资助.

Supported by the National Natural Science Foundation of China-Key Program (61833011).

地危害自然环境, 威胁人类健康, 尤其是 NO_x ^[4]. 为此, 传统的燃煤电厂采用了各种技术以降低 NO_x 的排放量, 例如选择性催化还原(selective catalytic reduction, SCR)和选择性非催化还原技术, 二者的原理都是通过化学方法将有害的 NO_x 转化为无害的 N_2 .

燃煤机组的运行是一个十分复杂的动态过程, 有着强非线性、大惯性和大时延等特性, 这些特性导致传统的比例积分微分(proportional-integral-derivative, PID)控制方法难以实现脱硝系统的精确快速控制. 为此, 国内外许多研究人员致力于SCR脱硝系统及电厂其他控制系统的研究^[5-8]. 马增辉等人^[9]设计了一种带有鲁棒时滞滤波器的PID控制器, 并将其用于SCR脱硝系统的控制中, 实验结果表明所设计的控制器结构简单且鲁棒性好. 为兼顾SCR脱硝系统的效率和经济性, 杨婷婷等人^[10]将基于神经网络的预测、基于遗传算法的多目标寻优与预测控制相结合, 实现了SCR脱硝系统的高效经济控制. 针对SCR脱硝系统的大惯性、大时延特性, 杜鸣等人^[11]提出一种改进的模糊线性自抗扰控制器, 所提方法不仅很好地追踪了设定值, 并且在发生模型失配时, 仍能保持良好的鲁棒性. 此外, 近年来备受关注的强化学习(reinforcement learning, RL)算法也被应用于SCR脱硝系统及其他工业过程的智能控制中. 陈皓炜等^[12]将出口 NO_x 浓度设定值与测量值之间的误差求积分与微分, 并存储于经验池中, 接着设计了用于前馈补偿的扰动观测器以降低控制过程的扰动, 所提方法有效地提升了原始强化学习算法的控制性能. 为提升脱硝效率, 林康威等人^[13]建立了Stacking融合框架预测模型, 实现了 NO_x 排放量的精确预测, 并将预测模型作为强化学习算法的环境, 达到了预期的控制目标.

然而, 燃煤机组在运行过程中, 大量的过程参数不断变化, 当负荷发生变化时, 机组的其他参数也会发生变化, 脱硝系统的模型也因此改变, 单一的控制策略难以在机组整个运行过程中保证良好的控制效果, 进而造成 NO_x 排放量超标, 导致电厂必须注入大于预期值的还原剂以降低 NO_x 排放量, 而还原剂的价格通常较为昂贵, 尽管可以达到预期的控制目标, 过量注入的还原剂仍会给电厂带来经济损失. 更重要的是, 过量投入的还原剂与水发生反应后可能会阻塞甚至腐蚀设备, 从而引发严重的后果^[14]. 因此, 在机组负荷发生变化时如何实现良好的控制具有十分重要的研究意义. 侯鹏飞等人^[15]根据SCR系统在不同工况下的特性, 将一段时间内系统的运行划分为4个子区间, 并设计了适应于各个线性子系统模型的多模型切换动态矩阵-PID串级预测控制器. 尹贵豪等人^[16]以不同工况下脱硝系统入口 NO_x 质量浓度的预测值为参考, 引入多模型预测控制器, 实现了喷氨量的准确控制. 上述研究均考虑了变工况下SCR系统的控制问题, 然而, 并没有考虑某些预料之外的未知工况.

考虑到未知工况与已知工况间存在部分相近的模式, 具备特征迁移能力的迁移学习方法可以为未知工况下的脱硝控制提供参考^[17]. 为更好地利用历史故障数据, Sun等人^[18]通过权重迁移、隐含特征迁移和权重更新的方式实现了从历史信息到在线信息的迁移, 并在实际故障数据集中验证了所提方法的有效性. 陈佳鲜等人^[19]融合了迁移学习与支持向量回归方法, 将离线工况下的时序信息迁移到在线工况中, 实现了滚动轴承剩余寿命的准确在线评估.

鉴于上述的问题, 本文设计了一种迁移强化学习(transfer RL, TRL)控制器, 并将其用于SCR脱硝系统的控制中. 首先, 根据负荷大小将运行过程划分为3个阶段, 然后以各阶段系统模型为目标, 训练了强化学习控制器以学习不同阶段下系统模型的特征, 并用于变工况下SCR脱硝系统的控制中. 此外, 采用迁移学习方法以应对未知工况. 最后, 在某燃煤机组SCR脱硝系统的控制中充分验证了本文所提方法的有效性和优越性. 本文的主要贡献总结如下:

1) 在训练过程中, RL控制器充分学习了不同工况下SCR系统的模式, 因此, 在变工况下仍能保持良好的控制性能;

2) 将RL控制器在已知工况中学习到的模式迁移到未知工况, 较好地应对了可能出现的未知工况.

2 研究目标

为实现低 NO_x 排放, 需要深入了解 NO_x 的形成机理, 进而有针对性地做出控制. 在机组运行过程中, NO_x 主要由燃料中的含氮化合物经过分解和氧化形成. 为此, 电厂通常从燃料燃烧前、中、后3个阶段实现对 NO_x 排放量的控制, 其中, 最有效的方法是SCR脱硝技术.

SCR脱硝技术是一种烟气后处理技术, NO_x 在炉膛内形成后, 沿着烟气管道进入锅炉尾部烟道, 而烟道中布置了催化剂, 在烟气经过烟道时注入一定量的还原剂(一般为氨气), 烟气中的 NO_x 在催化剂和还原剂的作用下, 转化为无害的 N_2 ^[20]. SCR脱硝技术效率高、可靠性好, 现已广泛应用于燃煤电厂中. 本文以某燃煤电厂SCR脱硝系统的外回路为研究对象^[21].

3 研究方法

3.1 深度强化学习

强化学习的灵感来源于人脑的工作方式, 即不断试错并积累经验, 最终习得应对环境的策略. 在2013年, Google团队以强化学习为基础, 结合了深度学习方法^[22], 并成功用于游戏控制中^[23]. 自此, 深度强化学习受到了科研领域和工程领域的广泛关注^[24], 深度学习强大的特征映射能力与强化学习的自学习能力为各类不同的任务提供了无限的可能性. 深度强化学习算法的详细流程如下:

1) 给定一个环境 E , 其中包含了有限个描述环境特征的状态, 所有状态的集合定义为 S ;

2) 针对 t 时刻的环境状态 s_t , 智能体做出相应的动作 a_t , 得到奖励 r_t ;

3) 环境状态也由 s_t 以概率 p_t 转变为 s_{t+1} , 其中 p_t 为式(1)所示的状态转移概率, 即

$$p_t = p(s_{t+1}|s_t, a_t); \quad (1)$$

4) 重复流程2)–3), 直到累积奖励 R 或训练周期Epochs达到设定值, 最终得到最优的策略。

其中, 智能体所有动作的集合记为 A , 所有环境状态转移概率的集合记为 P , 同时, 设计了折扣系数 γ 以求取累积奖励 R , 即

$$R = \sum_{i=1}^{T-1} \gamma^i r_{t+i}, \quad \gamma \in [0, 1]. \quad (2)$$

上述过程可以总结为图1的形式, 在整个过程中, 每一时刻的环境状态仅与上一时刻的环境状态相关, 元组 $\langle S, P, A, R, \gamma \rangle$ 即代表了马尔可夫决策过程, 该过程也是深度强化学习算法的基础。

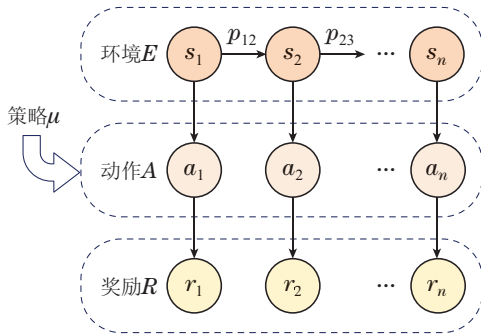


图1 强化学习算法示意图

Fig. 1 Schematic diagram of reinforcement learning algorithm

3.2 DDPG算法

随着深度强化学习算法在不同场景中的成功应用, 算法本身也得到了许多的改进, 考虑到实际电厂的连续运行特性, 本文使用了擅于处理连续控制问题的深度确定性策略梯度(deep deterministic policy gradient, DDPG)算法。DDPG是一种基于策略的无模型算法, 其主体框架为“Actor-Critic”结构, 其中, Actor网络和Critic网络分别用于产生和评估动作, 基于这一结构, 在Actor网络和Critic网络中均设置了主网络和目标网络^[25]。该算法的训练过程如下:

1) Actor主网络针对当前的环境状态 s 根据确定的策略 μ 输出动作 a , 即

$$a = \mu_{\theta}(s); \quad (3)$$

2) Critic主网络接收并评估 a , 评估结果 $Q_w(s, a)$ 进而用于计算损失 $Loss_a$, 随后以最小化该损失为目标优化Actor主网络(策略)的参数 θ , 即

$$Loss_a = -Q_w(s, a); \quad (4)$$

3) 采用指数移动平均(软更新)方法以系数 τ 更新Actor目标网络的参数 θ' , 即

$$\theta'_{t+1} = \tau\theta + (1 - \tau)\theta'_t; \quad (5)$$

4) Actor目标网络同样针对下一环境状态 s' 根据确定的策略 μ' 输出动作 a' , 即

$$a' = \mu'_{\theta'}(s'); \quad (6)$$

5) Critic目标网络接收 a' 与下一环境状态 s' , 输出评估结果 $Q_{w'}(s', a')$ 。这一结果连同Critic主网络的评估结果 $Q_w(s, a)$ 共同计算损失 $Loss_c$, 随后以最小化该损失为目标优化Critic主网络的参数 w , 即

$$Loss_c = \text{MSE}(Q_w(s, a), r + \gamma Q_{w'}(s', a')); \quad (7)$$

6) 同样采用指数移动平均方法以系数 τ 更新Critic目标网络的参数 w' , 即

$$w'_{t+1} = \tau w + (1 - \tau)w'_t. \quad (8)$$

DDPG采用了经验缓存机制, 将训练过程中的经验数据 (s_t, a_t, r_t, s_{t+1}) 存储在缓存中, 每次选取定量的经验数据用于网络参数更新。此外, 为更好地寻找最优策略, 在Actor网络中添加了随机噪声。DDPG算法的结构如图2所示, 本文使用的DDPG控制器中Actor与Critic网络均采用了全连接神经网络, 详细的参数设置如表1所示。

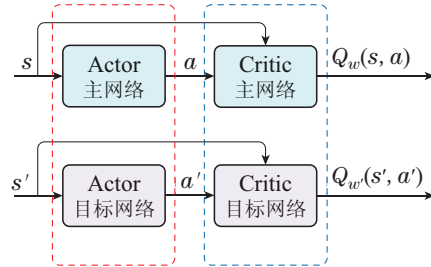


图2 DDPG算法详细结构

Fig. 2 Detailed structure of DDPG algorithm

表1 DDPG算法的参数设置

Table 1 Parameter setting of DDPG algorithm

参数	值
Actor网络神经元个数	3; 36; 1
Critic网络神经元个数	3; [50], [100, 50]; 1
学习率	1e-5
批处理大小	256
折扣因子	0.9
噪声方差	0.3

3.3 迁移学习

同样类似于人类的学习过程, 在处理多个任务时, 人类总是将旧任务中学到的知识套用到新的任务中, 这使得新任务并不是从零开始, 而是以旧任务中积累的经验为先验知识, 并在此基础上继续学习。迁移学

习的思路也是如此,在旧任务中所学习到的模式完全可以借鉴到新任务中,且新旧任务间的模式越接近,最终的效果越好^[26]。尽管存在着差异,不同工况下SCR脱硝系统的模型之间也会有上述的模式相近现象。因此,借助迁移学习,本文提出了针对未知或异常工况的控制方法,将算法在已知模型中学习到的经验作为未知模型的先验知识,从而更好地针对未知模型做出控制。

4 实验结果

本文设计了多组实验以充分验证所提方法的有效性。在实验1中,对比了RL及PID两种控制器在系统遭受干扰时的性能;在实验2中,验证了RL控制器在机组变工况时的控制效果;此外,设计了实验3以研究TRL控制器在机组未知工况发生时的控制效果。

4.1 实验1

燃煤机组锅炉内的燃烧过程十分复杂,涉及到许多质量、能量及形态的转化,各类情况所引发的干扰不可避免。例如:锅炉内高温高压的环境会加速传感器老化,烟气中的一些物质会在锅炉烟道内发生反应生成强腐蚀性液体,从而腐蚀设备,此外,锅炉内部大量的强电设备也会带来电磁干扰及辐射干扰。这些情况都不利于实际的控制过程。因此,研究控制器在干扰情况下的性能是十分有必要的。

如图3所示,在没有干扰时,PID控制器与RL控制器均可以将被控量控制在设定值附近,且PID控制器的调节时间略短,而RL控制器的超调量较小,这是由于RL控制器的奖励函数是根据参考值与输出值间的误差设定的,在训练过程中以最大化奖励为目标不断完善自身的策略,在实验2和实验3中也可以观察到类似的结果。在600 s时加入干扰,并且参考轨迹也发生变化。此时,两种控制器经过一段时间的调节后均可以将被控量控制在设定值附近,并且PID控制器的超调量较大,而RL控制器的超调量几乎可以忽略不计,二者的详细对比总结在表2中。实验结果表明本文所提的RL控制器有着良好的抗干扰能力,适用于存在大量干扰的实际电厂中。

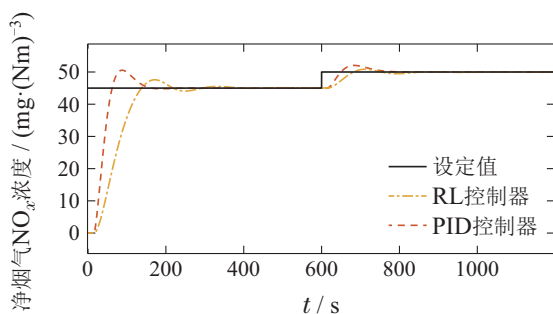


图3 遭受干扰时RL与PID的控制性能

Fig. 3 Control performance of RL and PID under disturbance

表2 遭受干扰时RL与PID控制器的性能对比

Table 2 Performance comparison of RL and PID controllers under disturbance

		无干扰/%	遭受干扰/%
超调量	RL	5.74	1.70
	PID	12.34	4.08
提升		6.60	2.38

4.2 实验2

本文所研究的机组在实际的运行过程中,大量的参数会随着机组负荷的变化而变化,SCR脱硝系统的模型也因此变化,而单一的控制策略难以针对不同的系统模型实现稳定的控制。

类似于人脑的学习过程,强大的学习能力使得强化学习算法有了无限的可能性,只要训练过程充足,训练方法合适,算法可以习得应对环境各种变化的最优策略,从而做出使得奖励最大的动作。因此,设计了随机负荷变化的系统模型切换机制,即600 s前后分别为机组高负荷与中负荷运行阶段,也分别对应高负荷与中负荷运行下的SCR系统模型。在训练过程,RL控制器与不同的系统模型交互,从而充分学习他们的特征,在学习过程结束之后测试了控制器的性能,并将其与电厂常用的PID控制器做了比较。比较结果如图4所示。

在第1阶段(1~600 s),PID控制器在经过一个较大的超调之后,迅速将被控量控制在设定值附近,与之相反的RL控制器经过了一个很小的超调,最终也将被控量控制在设定值附近;在第2阶段(601~1200 s),二者的控制效果相近。综合来看,PID控制器在调节时间方面略有优势,而RL控制器的超调量比PID控制器小很多,二者的详细对比总结在表3中。在实际的电厂中,对于一个控制器来说,其稳定性是最重要的,较大的超调量会引起控制阀门等设备的剧烈振荡,加速设备的老化甚至损坏设备,给电厂带来经济损失和安全威胁。因此,有着较小超调量的RL控制器可以更好地满足电厂实际的控制需求。

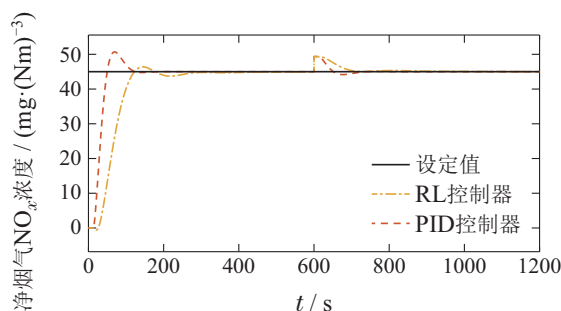


图4 变工况下RL与PID的控制性能

Fig. 4 Control performance of RL and PID under variable operating conditions

表3 变工况下RL与PID控制器的性能对比

Table 3 Performance comparison of RL and PID controllers under variable operating conditions

		1~600 s/%	601~1200 s/%
超调量	RL	3.08	9.58
	PID	12.63	9.73
提升		9.55	0.15

4.3 实验3

燃煤机组绝大多数情况运行都十分稳定,然而,一旦出现预料之外的未知或异常工况,机组必须检修甚至停机,从而影响机组的安全性和经济性.针对机组未知或异常工况的检测及控制一直是电厂的研究难点之一.因此,在上述实验的基础上,设计了针对SCR脱硝系统未知或异常工况的基于TRL算法的控制器,以机组高负荷、中负荷运行阶段为已知工况,低负荷运行阶段为未知工况.在训练过程中,控制器学习到了已知工况下系统的模型,随后,将学习到的经验作为先验知识,以应对未知的工况.

由图5可知,在未知工况发生时,由于有已知工况的先验知识,且未知工况与已知工况之间有部分相近的模式,TRL控制器的控制结果超调量较小,调节时间略长,而PID控制器的控制结果超调量较大,调节时间较短,二者的详细对比总结在表4中.上述结果表明,在先验知识的帮助下,TRL控制器可以较好地应对未知工况.

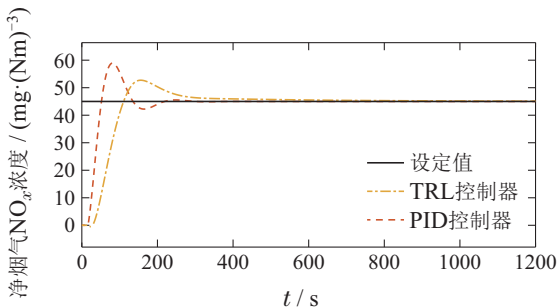


图5 未知工况下TRL与PID的控制性能

Fig. 5 Control performance of TRL and PID under unknown conditions

表4 未知工况下TRL与PID控制器的性能对比

Table 4 Performance comparison of TRL and PID controllers under unknown conditions

		值/%
超调量	TRL	17.15
	PID	30.94
提升		13.79

5 结论

复杂的运行工况给SCR脱硝系统的精确控制带来极大的困扰,为此,本文提出一种迁移强化学习控制

器,分别针对干扰、变工况及未知工况实现了准确的 NO_x 排放量控制,并为实际电厂的SCR脱硝控制及其他工业现场的类似场景提供了良好的参考.本文的研究结果如下:

1) 相比于电厂常用的PID控制器,所提出的RL控制器有着更强的抗干扰能力,可以较好地应对实际电厂中的强干扰环境;

2) 强大的学习能力使得RL控制器实现了变工况下SCR脱硝系统的准确控制,并且优于PID控制器;

3) 将已知工况的训练结果作为先验知识,使得TRL控制器在应对未知工况时表现出比PID控制器更好的性能.

然而,实际电厂中的场景远比本文的实验环境复杂得多,需考虑的因素也更为繁杂,这些情况都不可避免地给本文所提方法的实际应用提出了挑战.为此,后续的研究将致力于更加全面地考虑电厂的实际情况,并尝试将本文所提方法用于电厂的其他子系统及整体系统的控制中.

参考文献:

- [1] CHENG C L, PENG C, ZHANG T F. Fuzzy K-means cluster based generalized predictive control of ultra supercritical power plant. *IEEE Transactions on Industrial Informatics*, 2021, 17(7): 4574 – 4583.
- [2] HU Yaxu. *Multiple model predictive control of SNCR denitrification system based on subspace identification*. Taiyuan: Shanxi University, 2019.
(斛亚旭. 基于子空间辨识的SNCR脱硝系统的多模型预测控制. 太原: 山西大学, 2019.)
- [3] CHENG Chuanliang, PENG Chen, ZENG Deliang, et al. Modeling and simulation of ultra supercritical unit using a composite weighted human learning network. *Journal of System Simulation*, 2022, 34(7): 1430 – 1438.
(程传良, 彭晨, 曾德良, 等. 基于复合加权人类学习网络的超超临界机组建模与仿真. 系统仿真学报, 2022, 34(7): 1430 – 1438.)
- [4] 国务院. 国务院关于印发“十四五”节能减排综合工作方案的通知. 中华人民共和国国务院公报, 2022(5): 46 – 52.
- [5] LIU X F, REN M F, YANG Z L, et al. A multi-step predictive deep reinforcement learning algorithm for HVAC control systems in smart buildings. *Energy*, 2022, 259: 124857.
- [6] CHENG C L, PENG C, ZENG D L, et al. An improved neuro-fuzzy generalized predictive control of ultra-supercritical power plant. *Cognitive Computation*, 2021, 13: 1556 – 1663.
- [7] MA L Y, LEE K Y, WANG Z Y. Intelligent coordinated controller design for a 600 MW supercritical boiler unit based on expanded-structure neural network inverse models. *Control Engineering Practice*, 2016, 53: 194 – 201.
- [8] WANG Guoliang, YAN Weiwu, CHEN Shihe, et al. Multivariable constrained predictive control of 1000 MW ultra supercritical once-through boiler-turbine system. *Control Theory & Applications*, 2012, 29(12): 1573 – 1578.
(王国良, 阎威武, 陈世和, 等. 1000 MW超超临界机组的多变量约束预测控制的研究. 控制理论与应用, 2012, 29(12): 1573 – 1578.)
- [9] MA Zenghui, XU Huiyi, ZHU Runchao. Robust disturbance rejection control of SCR denitration system. *Control Engineering of China*, 2020, 27(1): 114 – 120.
(马增辉, 徐慧仪, 朱润潮. SCR烟气脱硝系统鲁棒抗干扰控制研究. 控制工程, 2020, 27(1): 114 – 120.)

- [10] YANG Tingting, BAI Yang, LÜ You, et al. Study on multi-objective optimal control of SCR denitrification system. *Proceedings of the CSEE*, 2021, 41(14): 4905 – 4911.
(杨婷婷, 白杨, 吕游, 等. SCR脱硝系统多目标优化控制研究. 中国电机工程学报, 2021, 41(14): 4905 – 4911.)
- [11] DU Ming, NIU Yuguang, PAN Xiangfeng, et al. Optimal control of SCR denitrification system based on improved fuzzy linear active disturbance rejection controller. *Journal of Chinese Society of Power Engineering*, 2021, 41(9): 743 – 748, 757.
(杜鸣, 牛玉广, 潘翔峰, 等. 基于改进型模糊线性自抗扰控制器的SCR脱硝系统优化控制. 动力工程学报, 2021, 41(9): 743 – 748, 757.)
- [12] CHEN Haowei, JIA Xinchun, SUN Xiaoming, et al. Composite cascade control based on reinforcement learning for a SCR denitration system. *Journal of Chinese Society of Power Engineering*, 2022, 42(5): 421 – 428.
(陈皓炜, 贾新春, 孙小明, 等. SCR脱硝系统的强化学习复合串级控制. 动力工程学报, 2022, 42(5): 421 – 428.)
- [13] LIN Kangwei, XIAO Hong, JIANG Wenchao, et al. Optimal control of denitrification processes in coal-fired power plants based on deterministic policy gradients with deep reinforcement learning. *Computer Measurement & Control*, 2022, 30(10): 132 – 139.
(林康威, 肖红, 姜文超, 等. 基于DDPG深度强化学习的电站脱硝过程优化控制. 计算机测量与控制, 2022, 30(10): 132 – 139.)
- [14] TAN P, HE B, ZHANG C, et al. Dynamic modeling of NO_x emission in a 660 MW coal-fired boiler with long short-term memory. *Energy*. 2019, 176: 429 – 436.
- [15] HOU Pengfei, JIA Xinchun, BAI Jianyun, et al. Multi-model switching DMC-PID cascade predictive control for SCR flue gas denitrification system. *Control Engineering of China*, 2020, 27(2): 355 – 360.
(侯鹏飞, 贾新春, 白建云, 等. SCR烟气脱硝系统多模切换DMC-PID串级预测控制. 控制工程, 2020, 27(2): 355 – 360.)
- [16] YIN Guihao, ZHAO Zhongyang, LI Qinwu, et al. Multi-model predictive control of denitrification system based on modification of inlet NO_x mass concentration and its application case. *Chinese Journal of Environmental Engineering*, 2022, 16(6): 1879 – 1891.
(尹贵豪, 赵中阳, 李钦武, 等. 基于入口NO_x质量浓度修正的脱硝系统多模型预测控制策略及其应用案例. 环境工程学报, 2022, 16(6): 1879 – 1891.)
- [17] ZHUANG F, QI Z, DUAN K, et al. A comprehensive survey on transfer learning. *Proceedings of the IEEE*, 2020, 109(1): 43 – 76.
- [18] SUN C, MA M, ZHAO Z, et al. Deep transfer learning based on sparse autoencoder for remaining useful life prediction of tool in manufacturing. *IEEE Transactions on Industrial Informatics*, 2018, 15(4): 2416 – 2425.
- [19] CHEN Jiaxian, MAO Wentao, LIU Jing, et al. Online remaining useful life estimation of bearing under unknown working conditions based on time series transfer recursive prediction. *Control and Decision*, 2023, 38(1): 112 – 122.
(陈佳鲜, 毛文涛, 刘京, 等. 基于时间序列迁移递归预测的未知工况下滚动轴承在线剩余寿命评估. 控制与决策, 2023, 38(1): 112 – 122.)
- [20] NIU Yuguang, PAN Yan, CHEN Xi. Depth structure control and depth control of selective catalytic reduction flue gas denitration system. *Control Theory & Applications*, 2019, 36(1): 65 – 72.
(牛玉广, 潘岩, 陈曦. 选择性催化还原烟气脱硝深度结构和深度控制. 控制理论与应用, 2019, 36(1): 65 – 72.)
- [21] DONG Chen. *Modeling and control of SCR denitration process on coal-fired power plant*. Nanjing: Southeast University, 2013.
(董宸. 燃煤电厂SCR脱硝过程的建模与控制. 南京: 东南大学, 2013.)
- [22] LECUN Y, BENGIO Y, HINTON G. Deep learning. *Nature*, 2015, 521: 436 – 444.
- [23] MNH V, KAVUKCUOGLU K, SILVER D, et al. Playing atari with deep reinforcement learning. *ArXiv Preprint*, 2013, arXiv: 1312.5602.
- [24] TANG Zhentao, SHAO Kun, ZHAO Dongbin, et al. Recent progress of deep reinforcement learning: From AlphaGo to AlphaGo Zero. *Control Theory & Applications*, 2017, 34(12): 1529 – 1546.
(唐振韬, 邵坤, 赵冬斌, 等. 深度强化学习进展: 从AlphaGo到AlphaGo Zero. 控制理论与应用, 2017, 34(12): 1529 – 1546.)
- [25] LILLICRAP T P, HUNT Y, PRITZEL A, et al. Continuous control with deep reinforcement learning. *ArXiv Preprint*, 2015, arXiv: 1509.02971.
- [26] PAN S J, YANG Q. A survey on transfer learning. *IEEE Transactions on Knowledge & Data Engineering*, 2010, 22(10): 1345 – 1359.

作者简介:

孙小明 博士研究生, 目前研究方向为发电过程建模、优化及智能控制, E-mail: xmsun1996@shu.edu.cn;

彭晨 博士, 教授, 目前研究方向为网络化控制、安全控制、工业炉窑能效优化等, E-mail: c.peng@i.shu.edu.cn;

程传良 博士后, 目前研究方向为火电厂建模及优化控制等, E-mail: clcheng@shu.edu.cn.