

考虑源荷随机性的跨区互联电网直流联络线调度学习优化

张 延¹, 唐 昊^{1†}, 王 珂², 潘 毅³, 李怡瑾¹

(1. 合肥工业大学 电气与自动化工程学院, 安徽 合肥 230009;

2. 中国电力科学研究院(南京), 江苏 南京 210003; 3. 中国电力科学研究院(北京), 北京 100192)

摘要: 在跨区互联电网中, 充分利用直流联络线调度能力可以有效地平衡电力资源的配置, 促进新能源的消纳. 本文针对源荷不确定性的跨区互联电网直流联络线调度问题, 首先用连续马尔科夫过程模型描述互联电网中风电出力与负荷需求随机动态特性; 然后在功率平衡及联络线日交易电量约束等实际运行要求前提下, 将直流联络线调度优化问题建立成离散马尔科夫决策过程模型. 在该模型下, 调度机构根据互联电网系统各时段源荷的功率情况, 动态调整联络线输电计划和配套的柔性负荷调节方案, 以达到提升系统运行效益的优化目标; 最后引入强化学习方法对调度策略进行优化求解. 通过学习优化, 系统平均日运行代价显著下降且最终收敛. 实验结果表明考虑源荷随机性的直流联络线动态调整方法可有效地提高互联电网发输电系统的运行效益.

关键词: 联络线调度; 新能源消纳; 柔性负荷; 随机性; 强化学习

引用格式: 张延, 唐昊, 王珂, 等. 考虑源荷随机性的跨区互联电网直流联络线调度学习优化, 2019, 36(7): 1047 – 1056

DOI: 10.7641/CTA.2018.80336

Learning-based optimization of direct current tie-line dispatch for inter-regional power grid considering the stochasticity of source-load

ZHANG Yan¹, TANG Hao^{1†}, WANG Ke², PAN Yi³, LI Yi-jin¹

(1. School of Electrical Engineering and Automation, Hefei University of Technology, Hefei Anhui 230009, China;

2. China Electric Power Research Institute (Nanjing), Nanjing Jiangsu 210003, China;

3. China Electric Power Research Institute (Beijing), Beijing 100192, China)

Abstract: In inter-regional power grid, the power resource can be allotted effectively by direct current tie-line to promote utilization ratio of renewable energy. The dispatch problem for direct current tie-line in inter-regional power grid with uncertain renewable sources and demands was researched in this paper. Firstly, the random dynamic characteristics of wind power output and load demand was described as continuous Markov process. Secondly, based on practical operation requirements including the power balance constraint and the limit of tie-line power, the optimal dispatch problem for direct current tie-line was described as a discrete Markov decision process. According to the power of renewable energy output and load demand, the optimized strategy for the plan of tie-line and flexible load in each period was established to promote the running benefit of the system in this model. Finally, a reinforcement learning method was adopted to obtain the optimal policy. The daily average cost of system operation decreases significantly and eventually converges by reinforcement learning. Simulation results show that the operational efficiency of inter-regional power grid is significantly enhanced by the proposed dynamic adjustment method for tie-line.

Key words: tie-line dispatch; accommodation of new energy; flexible load; stochasticity; reinforcement learning

Citation: ZHANG Yan, TANG Hao, WANG Ke, et al. Learning-based optimization of direct current tie-line dispatch for inter-regional power grid considering the stochasticity of source-load. *Control Theory & Applications*, 2019, 36(7): 1047 – 1056

1 引言

新能源的大规模开发利用, 保障了能源供应, 有效地降低了传统化石能源的消耗, 减少了环境污染^[1]. 然而, 由于不同区域的负荷需求特性及新能源分布情

况存在的巨大差异阻碍了新能源的消纳, 并且在风光电等新能源上网提供电能的同时, 其自身出力的不确定性也给电力系统运行带来了巨大的挑战^[2]. 跨区电网互联是实现资源在全国范围内优化配置和提高利

收稿日期: 2018-05-08; 录用日期: 2018-08-18.

[†]通信作者. E-mail: htang@hfut.edu.cn; Tel.: +86 551-62901416.

本文责任编辑: 岳东.

国家重点研发计划项目(2017YFB0902600), 国家电网公司科技项目(SGJS0000DKJS1700840)资助.

Supported by the National Key R&D of China (2017YFB0902600) and the State Grid Corporation of China Project (SGJS0000DKJS1700840).

用效率的重要手段之一,构建跨省、跨区互联电网,能够充分发挥大电网余缺调剂、资源优配、备用共享、事故支援等多方面效益,并大幅提高新能源的消纳水平^[3]。

直流输电具有线输送容量大、送电距离远、线路损耗小、调节速度快及运行稳定等优点。随着复奉、锦苏、宾金、天中等特高压跨区直流工程的相继投运,标志着我国以直流大规模跨区送电为特征的互联大电网已初步形成,为促进新能源消纳、实现能源空间上互补提供了高效的物理平台^[4],但同时区域互联电网也给电网调度计划制定带来了极大的挑战。如何匹配各区域中新能源出力的随机性及负荷需求特性,制定合理的联络线输电方案,利用直流跨区联络线的输电能力,完成资源的优化配置以充分消纳新能源,是当前调度计划急需解决的问题^[5]。

针对以上问题,国内外学者已经开展了相关研究。文献[3]及文献[5]以直流联络线功率为协调变量,提出跨区协调的直流跨区互联电网发输电计划模式,在满足联络线功率调整要求的前提下,建立了直流联络线运行优化的数学模型,通过联络线功率的合理规划促进互联电网的新能源消纳。文献[6]建立了考虑联络线约束的多区域经济调度模型,并提出具有时变系数的进化粒子群算法对多区域调度问题进行求解。上述文献均是在预测数据的基础上编制联络线计划的,文献[7]指出目前多区域实时发电计划是基于日前联络线计划,并通过实时调节各区域内机组出力,以应对新能源出力及负荷需求的随机波动的,尽可能保证日前的联络线计划不变。但新能源出力及负荷需求的不确定性是普遍存在且不可避免的,当实际值与日前预测值相差较大时,若以预测曲线确定的日前联络线计划为实时计划的基准值,将难以保证发输电系统的低代价运行。

另外,柔性负荷参与电力调度已引起诸多学者的关注。文献[8]将柔性负荷成本作为有机牵制扩展到调度模型之中,提出了计及大规模风电和柔性负荷的电力系统供需侧联合随机调度方法,并建立了供需侧联合随机调度模型。文献[9]利用高载能柔性负荷消纳风电,并将该类负荷与常规电源考虑到电网优化调度中,形成源荷协调优化运行的调度模式,以风电消纳电量最大和系统运行成本最小为目标建立源荷协调多目标优化模型。

目前,在智能电网背景下调度不断向智能化的方向发展^[10]。针对互联电网调度问题,将人工智能技术和电力调度技术相结合,可以实现电网运行的经济、环保等多目标最优^[11],为调控系统提供智能化的解决方案。

在上述研究的基础上,本文就源荷不确定情况下的联络线调度问题展开研究,将新能源出力与负荷需

求由随机产生的不确定部分建立成连续马尔科夫过程模型:在满足日交易电量、功率平衡等约束的前提下,根据源荷出力情况对联络线功率动态调整,同时将柔性负荷的调度作用从单区域应用到多区域调度问题中,更充分地发挥区域之间的调度潜力。考虑到当前电网调度一般为离散时段决策,故基于风电出力及负荷需求的马尔科夫过程模型,将柔性负荷协同的跨区直流联络线调度问题建立成离散马尔科夫决策过程(discrete time Markov decision processes, DTM-DP)模型。该模型以调度日各时段总运行代价最低为目标寻求优化策略,因此根据策略选取的行动不仅适用于当前时段,并且考虑了对后续时段调度计划的影响,保证了前后时段调度计划的关联性。最后本文使用强化学习方法对该调度问题进行策略求解,实现机组及联络线的联合优化。

2 直流跨区互联电网发输电系统模型

2.1 直流跨区电网系统架构

本文以两区域直流发输电系统为研究对象,设区域1为直流送端电网、区域2为直流受端电网,构建直流跨区互联电网发输电系统模型,如图1所示。

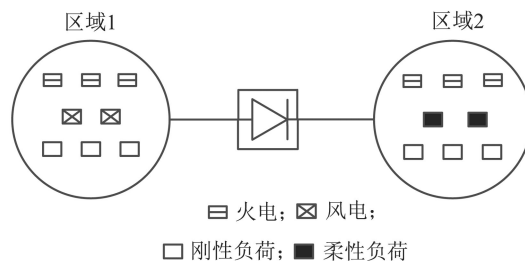


图1 直流跨区发输电系统模型图

Fig. 1 Generation and transmission system of inter-regional power grid via direct current tie-line

区域1代表新能源富足的区域,且假设新能源以风力能源为主。在跨区互联电网实际运行中,联络线两端一般具有明显的送受端特性^[3,5],因此,本文假设电力资源富足的区域1在常规火电机组及风电出力满足自身负荷用电需求的基础上通过直流联络线向电力需求量大的区域2单向传输电能。为简化问题的分析,本文借鉴文献[5]的处理方式,暂不考虑区域电网内部的网络安全约束和网络损耗。由于目前直流联络线外送电量计划主要由跨区电力市场交易确定,为保证交易的执行,调度周期内送出的总电量应在市场交易合同约定范围内,因此联络线日输送电量需满足^[6]:

$$Q_{\text{exp}} = Q/N_{\text{day}}, \quad (1)$$

式中: Q_{exp} 为期望的联络线日输送电量; Q 为两区域电力交易中成交的月输送电量总额; N_{day} 为该月天数。

同时,区域1内部在任意时刻 t 要保证功率平衡:

$$\sum_{n=1}^{N_{\text{wind}}} p_{\text{w}}^{t,n} + \sum_{n=1}^{N_{\text{gs}}} p_{\text{gs}}^{t,n} = \sum_{n=1}^{N_{\text{Ls}}} p_{\text{Ls}}^{t,n} + p_{\text{line}}^t, \quad (2)$$

式中: $p_{\text{w}}^{t,n}$, $p_{\text{gs}}^{t,n}$, $p_{\text{Ls}}^{t,n}$ 分别为 t 时刻区域1中第 n 台风电机组、第 n 台常规火电机组的发电功率及第 n 个负荷的需求功率; N_{wind} , N_{gs} , N_{Ls} 分别为区域1中风电机组与常规火电机组的台数及负荷的数量; p_{line}^t 为 t 时刻的联络线输电功率。

区域2代表负荷需求量大的区域. 由于负荷量大, 将该区域的柔性负荷调节能力考虑到调度优化中. 本文考虑的柔性负荷只限于直接负荷控制 (direct load control, DLC) 类型, DLC负荷是指在用户许可的条件下, 调度中心根据电网需求直接控制用户的一类负荷, 通常这类负荷在短时间停电的情况下不会明显影响用户对供电服务质量的满意度, 并且在每个时段结束后调度中心会根据该时段的中断电量给予这部分用户一定补偿^[12].

区域2内部在任意时刻 t 也要保证功率平衡:

$$p_{\text{line}}^t + \sum_{n=1}^{N_{\text{gr}}} p_{\text{gr}}^{t,n} = \sum_{n=1}^{N_{\text{Lr}}} p_{\text{Lr}}^{t,n}, \quad (3)$$

式中: $p_{\text{gr}}^{t,n}$, $p_{\text{Lr}}^{t,n}$ 分别为 t 时刻区域2中第 n 台常规火电机组的发电功率和第 n 个负荷的电力需求; N_{gr} , N_{Lr} 分别为区域2中常规火电机组台数及负荷数量。

区域2中 t 时刻负荷总需求由该区域的刚性负荷及DLC柔性负荷总需求组成:

$$\sum_{n=1}^{N_{\text{Lr}}} p_{\text{Lr}}^{t,n} = \sum_{n=1}^{N_{\text{rigid}}} p_{\text{rigid}}^{t,n} + \sum_{n=1}^{N_{\text{flex}}} p_{\text{flex}}^{t,n}, \quad (4)$$

式中: $p_{\text{rigid}}^{t,n}$, $p_{\text{flex}}^{t,n}$ 分别为 t 时刻区域2中第 n 个刚性负荷及第 n 个柔性负荷的电力需求; N_{rigid} , N_{flex} 分别为区域2中刚性负荷及柔性负荷的数量。

2.2 系统源荷端模型

2.2.1 风电出力及负荷需求的随机模型

风力发电作为风能的主要利用方式, 也是目前最成熟的新能源发电方式之一. 风电作为一种间歇性电源, 其大规模接入给电力系统的安全稳定运行带来了新挑战^[13-15], 把握新能源出力特性是突破新能源大规模并网瓶颈的前提^[16]. 在实际情况中, 风电出力变化既有预测的确定部分, 同时由于环境等原因, 也具有含一定统计特征的随机不确定部分。

风力发电功率与外界多种因素相关, 其中, 风速的影响较大也最为直接. 马尔科夫过程数学模型可以通过转移函数和逗留时间分布函数或概率密度函数描述随机量变化的逻辑关系和统计特性, 适用于描述实际中一些动态随机变量的序贯变化过程. 国内外已有很多应用马尔科夫理论对风速建模的研究, 文献[17]将风速描述为具有时间序列的随机变量, 用马尔科夫理论对风速建模并取得了较好的效果. 文献[18]提出

了一种基于离散时间马尔可夫链模型的风电功率预测方法, 分析介绍了风电的一阶和二阶马尔科夫链模型。

在实际系统运行过程中, 可以根据气象数据预测得到次日的风速预测曲线, 那么实际风速可视为风速预测确定部分与随机不确定部分的叠加, 表达式为

$$v_{\text{w}}^t = v_{\text{wbase}}^t + \Lambda_{\text{w}}^t, \quad (5)$$

式中: v_{w}^t , v_{wbase}^t 分别表示 t 时刻实际风速和预测风速; Λ_{w}^t 为 t 时刻风速的随机不确定部分。

由于风电出力和风速直接相关^[19], 风电出力也可以看作是确定性部分与不确定部分的叠加. 本文将风电出力与风速的随机不确定部分近似视为成正比关系, 将风速不确定部分的最大范围区间离散为 $-N_{\text{w}}$, $-N_{\text{w}} + 1, \dots, 0, \dots, N_{\text{w}}$ 共 $2N_{\text{w}} + 1$ 个等级, 相应地, 风电出力的不确定部分也因风速的不确定部分离散为相应等级. 将该不确定量的变化建模为马尔科夫过程, 假设状态等级 n_{w}^t 的逗留时间服从参数为 λ_{w}^t 的指数分布, 由于风电出力不确定部分的等级更趋向于向自身及邻近功率等级转移, 因此, 设在各时刻下的状态转移概率服从以自身为中心的离散高斯分布. 在 t 时刻的风电实际出力功率可近似描述为

$$p_{\text{w}}^t = p_{\text{wbase}}^t + n_{\text{w}}^t \Delta_{\text{w}}^t, \quad (6)$$

式中: p_{w}^t , p_{wbase}^t 分别表示 t 时刻风电的实际出力功率和预测出力功率; Δ_{w}^t 表示 t 时刻风电出力不确定部分离散后的最小单位; n_{w}^t 表示在 t 时刻风电出力不确定部分的功率等级。

类似地, 系统中两区域负荷的实际需求量也可视为日前预测确定部分与由随机不确定部分的叠加, 并将其随机不确定部分变化用连续马尔科夫过程进行描述^[20]. 通过对环境(天气、温度等)以及用户历史行为预测得到日负荷曲线, 则各时刻下的负荷需求实际功率可表示为(以区域1为例)

$$p_{\text{Ls}}^t = p_{\text{Lsbase}}^t + \Lambda_{\text{Ls}}^t, \quad (7)$$

式中: p_{Ls}^t , p_{Lsbase}^t 分别表示 t 时刻负荷的实际需求功率和预测需求功率; Λ_{Ls}^t 为负荷需求功率的不确定部分。

将负荷需求不确定部分的最大范围区间离散为 $-N_{\text{L}}$, $-N_{\text{L}} + 1, \dots, 0, \dots, N_{\text{L}}$ 共 $2N_{\text{L}} + 1$ 个等级, 状态等级变化也可用连续马尔科夫过程进行描述. 假设状态 n_{L}^t 逗留时间服从参数为 λ_{L}^t 的指数分布, 在各时刻下的状态转移概率服从以自身为中心的离散高斯分布. 负荷在 t 时刻的实际需求功率可描述为(以区域1为例)

$$p_{\text{Ls}}^t = p_{\text{Lsbase}}^t + n_{\text{Ls}}^t \Delta_{\text{Ls}}^t, \quad (8)$$

式中: Δ_{Ls}^t 表示 t 时刻负荷需求不确定部分离散后的最小单位; n_{Ls}^t 表示在 t 时刻负荷需求的功率等级。

2.2.2 DLC柔性负荷模型

区域2中 t 时刻下的DLC柔性负荷的电力需求量与该区域当前时刻的负荷总需求量有着一定的统计关系,假设两者成一定比例的线性关系,如下:

$$\sum_{n=1}^{N_{\text{flex}}} p_{\text{flex}}^{t,n} = \mu \sum_{n=1}^{N_{\text{Lr}}} p_{\text{Lr}}^{t,n}, \quad (9)$$

式中 μ 为DLC柔性负荷占该区域总负荷需求的比例。

决策者可在时刻 t_k 中断下一时段 k 中部分DLC柔性负荷的电力需求,以达到配合区域内机组出力及降低日负荷峰谷差的目的。

2.2.3 等效火电机组模型

在直流跨区互联电网发输电计划编制过程中,下级的机组计划和上级的联络线计划独立开展,在联络线计划为已知的条件下才开展机组计划编制^[3]。由于本文主要研究上级的联络线动态调整问题,其重点不针对区域内部的机组组合和经济调度问题,因此根据参考文献[5]采取等效火电机组来近似代替两区域的常规火电机组运行情况。假设区域1的发电机组由等效火电机组和等效新能源机组组成,区域2的发电机组仅认为有等效火电机组。等效火电机组的出力表达式为(以区域1为例)

$$p_{\text{gs}}^t = \sum_{n=1}^{N_{\text{gs}}} p_{\text{gs}}^{t,n}, \quad (10)$$

式中 p_{gs}^t 为区域1等效火电机组在 t 时刻的总出力。

若联络线输电功率给定时, t 时刻的等效火电机组所需出力可由两区域该时刻的功率平衡约束得到。同时等效火电机组也要满足常规火电机组自身的约束(以区域1为例):

1) 机组出力上下限约束:

$$p_{\text{gs}}^{\min} \leq p_{\text{gs}}^t \leq p_{\text{gs}}^{\max}, \quad (11)$$

式中 $p_{\text{gs}}^{\min}$ 和 $p_{\text{gs}}^{\max}$ 分别表示区域1等效火力发电机的最低和最高输出功率。

2) 机组爬坡约束:

$$-p_{\text{gs}}^{\text{dn}} \leq \bar{p}_{\text{gs}}^k - \bar{p}_{\text{gs}}^{k-1} \leq p_{\text{gs}}^{\text{up}}, \quad (12)$$

式中: $p_{\text{gs}}^{\text{up}}$ 和 $-p_{\text{gs}}^{\text{dn}}$ 分别为区域1等效火电机组的爬坡增出力上限和爬坡减出力下限; $\bar{p}_{\text{gs}}^k$ 和 $\bar{p}_{\text{gs}}^{k-1}$ 分别代表两区域等效发电机组 k 时段和 $k-1$ 时段等效火电机组的平均出力功率。

由于机组出力下限的约束,在区域1风电出力较大的部分时刻,可能会出现风电和火电机组总出力超过联络线输电和该区域负荷总需求的情况,表示为

$$\sum_{n=1}^{N_{\text{w}}} p_{\text{w}}^{t,n} + p_{\text{gs}}^{\min} > p_{\text{line}}^t + \sum_{n=1}^{N_{\text{Ls}}} p_{\text{Ls}}^{t,n}, \quad (13)$$

此时为保证功率平衡,将会出现弃风现象。

3 DTMDP模型与求解方法

为实现联络线的动态调整,所求调度策略需要对任一决策时刻所有可能出现的新能源出力及负荷需求情况,给出优化行动,属于具有序贯特性的动态决策问题。因此,本文将系统决策过程建模为相应的DTMDP模型。该模型包括状态、行动、转移代价及优化目标函数,目前已应用于电力系统领域相关建模与优化问题中^[21]。

3.1 状态空间及行动集

将一天时间分为 $K+1$ 个决策周期,采用等周期决策的方式,决策时刻为 t_k ,其中 $k \in \{0, 1, \dots, K\}$;由公式(6)和公式(8)可知,决策时刻 t_k 的风电实际功率 $p_{\text{w}}^{t_k}$ 和两区域负荷实际功率 $p_{\text{Ls}}^{t_k}$, $p_{\text{Lr}}^{t_k}$ 的集合如公式(14)–(16)所示:

$$p_{\text{w}}^{t_k} = p_{\text{wbase}}^{t_k} + n_{\text{w}} \Delta_{\text{w}}^{t_k},$$

$$n_{\text{w}} \in \{-N_{\text{w}}, \dots, 0, 1, \dots, N_{\text{w}}\}, \quad (14)$$

$$p_{\text{Ls}}^{t_k} = p_{\text{Lsbase}}^{t_k} + n_{\text{Ls}} \Delta_{\text{Ls}}^{t_k},$$

$$n_{\text{Ls}} \in \{-N_{\text{Ls}}, \dots, 0, 1, \dots, N_{\text{Ls}}\}, \quad (15)$$

$$p_{\text{Lr}}^{t_k} = p_{\text{Lrbase}}^{t_k} + n_{\text{Lr}} \Delta_{\text{Lr}}^{t_k},$$

$$n_{\text{Lr}} \in \{-N_{\text{Lr}}, \dots, 0, 1, \dots, N_{\text{Lr}}\}. \quad (16)$$

将联络线日输送电量在最大输电区间离散为相应的 $0 \sim N_{\text{Q}}$ 个状态等级,在每个决策时刻下已完成的日交易电量等级 $n_{\text{Q}} \in \{0, 1, \dots, N_{\text{Q}}\}$,在末状态时根据所处状态等级与交易规定电量 Q_{exp} 对应状态等级 $n_{\text{Q}}^{\text{exp}}$ 的差值对所处末状态进行评价。定义系统状态 s 由决策时刻 t_k ,两区域负荷实际功率 $p_{\text{Ls}}^{t_k}$, $p_{\text{Lr}}^{t_k}$,风电实际功率 $p_{\text{w}}^{t_k}$ 以及当前已完成的日交易电量等级 n_{Q} 组成,即 $s = (t_k, p_{\text{Ls}}^{t_k}, p_{\text{Lr}}^{t_k}, p_{\text{w}}^{t_k}, n_{\text{Q}}) \in \Phi$, Φ 为系统状态空间,状态总数为

$$N = (K+1)(2N_{\text{Ls}}+1)(2N_{\text{Lr}}+1)(2N_{\text{w}}+1) \cdot (N_{\text{Q}}+1). \quad (17)$$

各时刻 t 下联络线输送功率必须满足联络线安全输送功率的约束,不能大于其安全功率,表达式如下:

$$p_{\text{line}}^t \leq p_{\text{line}}^{\lim}, \quad (18)$$

式中 $p_{\text{line}}^{\lim}$ 为联络线安全输送功率。

将联络线安全输送功率区间 $[0, p_{\text{line}}^{\lim}]$ 和各个时段中断的DLC柔性负荷量占该时段总DLC柔性负荷量的比例区间 $[0, 1]$ 也离散为相应的 $0 \sim N_{\text{tl}}$, $0 \sim N_{\text{f}}$ 个状态等级。系统行动包括决策周期内的联络线传输功率等级 $a_{\text{tl}} \in D_{\text{tl}} = \{0, 1, \dots, N_{\text{tl}}\}$ 和中断的DLC柔性负荷功率占比的等级 $a_{\text{f}} \in D_{\text{f}} = \{0, 1, \dots, N_{\text{f}}\}$ 。系统行动集为 $D = D_{\text{tl}} \times D_{\text{f}}$,系统行动向量为 $a = (a_{\text{tl}}, a_{\text{f}}) \in D$ 。

系统的策略可表示为

$$\pi = \begin{pmatrix} a_{tl}^1 & a_f^1 \\ \vdots & \vdots \\ a_{tl}^N & a_f^N \end{pmatrix}, \quad (19)$$

式中: a_{tl}^N, a_f^N 分别代表在状态编号 N 下采取的联络线传输功率和DLC柔性负荷切除行动。

对这 N 个状态进行排序, 给定状态 $\hat{s} = (\hat{t}_k, \hat{p}_{Ls}^{t_k}, \hat{p}_{Lr}^{t_k}, \hat{p}_w^{t_k}, \hat{n}_Q)$ 的状态编号 $\hat{n}_s$, 即该状态在排序中所处的位置序号 $\hat{n}_s$, 可由下式确定:

$$\begin{aligned} \hat{n}_s &= \hat{k}(N_{Ls} + 1)(N_{Lr} + 1)(N_w + 1)(N_Q + 1) + \\ &\quad \hat{n}_{Ls}(N_{Lr} + 1)(N_w + 1)(N_Q + 1) + \\ &\quad \hat{n}_{Lr}(N_w + 1)(N_Q + 1) + \\ &\quad \hat{n}_w(N_Q + 1) + \hat{n}_Q + 1. \end{aligned} \quad (20)$$

类似地, 给定行动 $\hat{a} = (\hat{a}_{tl}, \hat{a}_f)$ 的行动编号 $\hat{n}_a$, 即该行动在行动集排序中所处的序号 $\hat{n}_a$, 可由下式确定:

$$\hat{n}_a = \hat{a}_{tl}(N_f + 1) + \hat{a}_f + 1. \quad (21)$$

3.2 代价设置及目标函数

本文中每个决策周期产生的代价包括等效火电机运行代价、弃风代价及对柔性负荷进行补偿代价等。另外, 为一定程度降低日负荷需求峰谷差, 使各时段负荷需求量尽量趋向于预测的日平均功率, 将决策后的各时段负荷需求偏离日负荷均值的量也设置成代价。

发电机运行代价可近似为功率的二次函数关系^[3], 因此, 决策周期 k 内产生的等效火电机运行代价 c_g^k 可表示为

$$c_g^k = c_{gs}^k + c_{gr}^k, \quad (22)$$

$$c_{gs}^k = \int_{t_k}^{t_{k+1}} (a_s p_{gs}^t{}^2 + b_s p_{gs}^t + c_s) dt, \quad (23)$$

式中: c_g^k 由该时段区域1等效机组运行代价 c_{gs}^k 和区域2等效机组运行代价 c_{gr}^k 组成; a_s, b_s, c_s 为区域1中机组运行的代价系数; 区域2机组运行代价 c_{gr}^k 与区域1形式类似。

决策周期 k 内产生的弃风代价 c_{aw}^k 可表示为

$$c_{aw}^k = \alpha \int_{t_k}^{t_{k+1}} \xi_w^t dt, \quad (24)$$

$$\xi_w^t = \sum_{n=1}^{N_w} p_w^{t,n} + p_{gs}^{\min} - p_{line}^t - \sum_{n=1}^{N_{Ls}} p_{Ls}^{t,n}, \quad (25)$$

式中: α 为弃风的代价系数, ξ_w^t 为 t 时刻的弃风功率。

决策周期 k 内对DLC柔性负荷的补偿代价 c_c^k 可表示为

$$c_c^k = \beta \int_{t_k}^{t_{k+1}} p_f^t dt, \quad (26)$$

$$p_f^t = (a_f / N_f) \sum_{n=1}^{N_{flex}} p_{flex}^{t,n}, \quad (27)$$

式中: β 为对柔性负荷中断的补偿系数; p_f^t 为 t 时刻的DLC柔性负荷切除功率。

决策周期 k 内的负荷偏离日负荷均值的代价表示为

$$c_{pvd}^k = \omega \int_{t_k}^{t_{k+1}} (p_{Lr}^t - \bar{p}_{Lr})^2 dt, \quad (28)$$

$$\bar{p}_{Lr} = (\int_0^T p_{Lrbase}^t) / T, \quad (29)$$

式中: c_{pvd}^k 表示在决策周期 k 内的负荷峰谷差衡量代价; ω 为其代价系数; $\bar{p}_{Lr}$ 为区域2基于日负荷预测曲线的平均负荷需求。

约束条件包括联络线日电量约束, 联络线安全输送功率约束, 联络线阶梯化约束, 功率平衡约束以及等效火电机上下限限制约束和爬坡约束。其中, 决策周期 k 内若违反爬坡约束可用代价的形式来体现:

$$c_{pun}^k = \gamma \int_{t_k}^{t_{k+1}} (\bar{p}_{difs}^k + \bar{p}_{difr}^k) dt, \quad (30)$$

式中: c_{pun}^k 表示决策周期 k 产生的爬坡惩罚代价; γ 为爬坡惩罚系数, 为使优化后的结果尽可能不违反爬坡约束, 一般选取较大的爬坡惩罚系数; $\bar{p}_{difs}^k$ 和 $\bar{p}_{difr}^k$ 分别为两区域中等效火电机前后时段平均功率变化量超出爬坡限制的功率值, 以 $\bar{p}_{difs}^k$ 为例, 表示为

$$\bar{p}_{difs}^k = \begin{cases} \bar{p}_{gs}^{k-1} - \bar{p}_{gs}^k - p_{gs}^{dn}, & \bar{p}_{gs}^k < \bar{p}_{gs}^{k-1}, \\ \bar{p}_{gs}^k - \bar{p}_{gs}^{k-1} - p_{gs}^{up}, & \bar{p}_{gs}^k > \bar{p}_{gs}^{k-1}. \end{cases} \quad (31)$$

联络线日电量约束是由大区间交易计划确定的, 要求每天联络线的输送总电量尽可能接近 Q_{exp} , 因此, 在处于末状态时, 根据该状态下的输送电量完成度对末状态进行评估, 末状态的评估代价可表示为

$$c_Q(s_{K+1}) = \delta (n_Q' - n_Q^{exp})^2, \quad (32)$$

式中: n_Q' 为末状态下输送电量完成状态; n_Q^{exp} 为 Q_{exp} 对应的输送电量完成状态; δ 为末状态代价的系数。

由于联络线的输电功率是以其安全输送功率为最大值进行离散化的, 故得到的联络线输电计划已满足联络线安全输送功率约束, 并且一个决策周期内的联络线输电功率不改变, 因此也满足联络线功率阶梯化约束。

由公式(22)–(31)可得决策周期 k 产生的总代价 c^k , 表达式为

$$c^k = c_g^k + c_{aw}^k + c_c^k + c_{pvd}^k + c_{pun}^k. \quad (33)$$

日代价 V 可表示为有限时段内的运行代价与末状态交易电量完成度代价之和形式:

$$V = c_Q(s_{K+1}) + \sum_{k=0}^K c^k. \quad (34)$$

假设系统在策略 π 下以初始状态为 s_0 的有限时段内优化性能准则为 $V^\pi(s_0)$, 可表示为

$$V^\pi(s_0) = E_\pi[c_Q(s_{K+1}) + \sum_{k=0}^K c^k]. \quad (35)$$

系统的优化目标即找到一个最优策略 π^* , 在该策略下整个系统的平均日运行总代价最低^[22-23], 表示为

$$\pi^* = \arg \min_{\pi \in \Omega} V^\pi, \quad (36)$$

式中 Ω 为策略集, 所有满足式(19)形式的策略构成的策略集合即是策略集 Ω .

本文根据新能源、负荷预测的确定部分以及由随机性产生的不确定部分模拟实际的功率, 以此不断产生训练样本, 提供状态数据. 通过在线学习的方式, 得到优化策略. 系统在各决策时刻 t_k 根据源荷功率值以及已完成的电力输送量判断当前情况所处的状态, 再通过优化策略执行该状态下对应的联络线输送功率及柔性负荷调节方案.

3.3 强化学习方法

针对该调度问题, 本文利用强化学习的方法进行策略求解. 强化学习算法把学习看作试探评价过程, 可以通过学习得到从环境状态到行为的映射, 使得调度体选择的行为能够保证长期范围内整个系统的运行性能最佳^[24-25]. 其中, Q 学习算法是由Watkins提出的一种模型无关的强化学习方法, Q 学习迭代时采用状态-行动对的奖赏作为估计函数, 因此系统每一步学习迭代都需要考察每一个行为, 确保学习过程收敛, 利用随机动态规划的思想寻找决策过程的优化策略, 在解决模型未知和具有随机环境的问题时具有很好的效果^[26-27].

参照第2.1节, 以各决策时刻下的新能源出力及负荷需求功率为状态, 以决策周期内的联络线输电功率及柔性负荷切除方案为行动. 在时段 k 时, 通过观测可知系统处于当前状态 s_k , 根据相关策略采取行动 a_k , 系统状态转移到下一时段的状态 s_{k+1} , 并产生转移代价 c^k , 由此得到一个状态转移为 $\langle s_k, a_k, s_{k+1}, c^k \rangle$, 该过程形成一个完整的样本. 本文采用的基于模拟退火的有限时段 Q 学习算法就是通过对大量样本的知识进行学习, 利用 Q 值更新公式:

$$Q(s_k, a_k) = Q(s_k, a_k) + \alpha_m(c^k - Q(s_k, a_k) + \min_{a_{k+1}} Q(s_{k+1}, a_{k+1})). \quad (37)$$

不断更新各状态行动对应的 Q 值, 并选择各状态下最小 Q 值对应的行动作为该状态下新的动作. 通过 Q 值的反复更新, 最终得到优化策略. 具体求解过程如下所示:

Step 1 初始化 Q 值表, 样本轨道总数 M , 单条样本轨道决策周期数 K , 学习率 α_m , 学习率更新系数 η_α , 模拟退火温度 T_{temp} 和模拟退火系数 η_{temp} ; 令 $m = 0$;

Step 2 $k = 0$, 随机初始化系统状态 $s = s_k$, 即风

电、负荷实际功率及完成的输送电量;

Step 3 根据 Q 值表和贪心策略, 选取当前行动 a_k : 令 $a_{\text{greedy}} = \arg \min_{a_k} Q(s_k, a_k)$, 即当前 Q 值表内 s_k 对应最优的联络线输送功率及配套的柔性负荷切除方案; 令产生的随机行动为 a_{rand} , 即随机产生的可行方案. 若 $e^{(Q(s_k, a_{\text{greedy}}) - Q(s_k, a_{\text{rand}}))/T_{\text{temp}}} < \text{random}(0, 1)$, 则选择 a_{greedy} 为当前行动, 否则选择 a_{rand} 为当前行动. 若 $k < K$, 跳转至Step 4; 若 $k = K$, 跳转至Step 5;

Step 4 执行当前所选行动 a_k , 根据所建系统模型确定下一时段对应的系统状态 s_{k+1} , 并对 k 时段内系统运行效益进行计算评估, 进而产生该时段对应的代价 c^k , 根据更新公式 $Q(s_k, a_k) = Q(s_k, a_k) + \alpha_m(c^k - Q(s_k, a_k) + \min_{a_{k+1}} Q(s_{k+1}, a_{k+1}))$ 更新 Q 值表, $k := k + 1$, 返回Step 3;

Step 5 执行当前行动 a_K , 一个决策周期后观察产生代价 c^K , 及末状态代价 $c_Q(s_{K+1})$, 根据 $Q(s_K, a_K) = Q(s_K, a_K) + \alpha_m(c^K + c_Q(s_{K+1}) - Q(s_K, a_K))$ 更新 Q 值. 令 $m := m + 1$;

Step 6 若 $m \bmod 10000 = 0$, 根据学习率更新系数 η_α 更新学习率 $\alpha_m := \eta_\alpha \alpha_m$;

Step 7 若 $m < M$, 根据退火系数 η_{temp} 更新温度 $T_{\text{temp}} := \eta_{\text{temp}} T_{\text{temp}}$, 返回Step 2; 否则结束程序.

根据更新后的 Q 值表生成策略 π' , 策略中各状态 s 对应的行动由贪心策略得到.

4 仿真结果

本节参照文献[5]中跨区电网典型日的实际运行数据对所提方法开展分析. 对前一部分建立的DTMDP模型用模拟退火 Q 学习进行求解, 并根据结果验证优化方法的有效性.

4.1 系统基础数据

在本文算例中, 设直流联络线的安全传输功率区间为 $[0, 500]$ MW, 防止功率反转; 直流联络线的日交易传输电量 $Q_{\text{exp}} = 7200$ MW; 区域2中的DLC控制的柔性负荷占该区域整体负荷需求的比例 $\mu = 0.15$; 决策周期为24个时段, 每个时段的时间间隔为1 h.

本文中假设新能源机组的单位出力成本为零, 等效火电机组成本参数如下表1所示^[5].

区域1中日风电出力预测曲线及两区域的日负荷功率预测曲线采用文献[3]中对应的预测曲线形状, 如图2所示.

表1 等效发电机组成本参数

Table 1 Cost parameters of equivalent generating units

等效机组	$a/(\text{元} \cdot \text{MW}^{-2})$	$b/(\text{元} \cdot \text{MW}^{-1})$	$c/\text{元}$
g_s	0.006	245	200
g_r	0.002	260	200

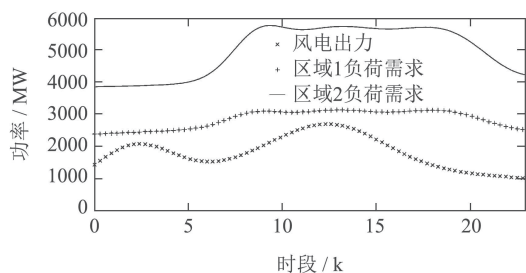


图2 风电与负荷预测曲线图

Fig. 2 Prediction curves of wind power and loads

4.2 结果分析

本文考虑的联络线输电计划的优化问题,在MATLAB R2016a环境下进行算例仿真。

1) 学习曲线分析.

在学习过程中,每经过40000个样本轨道,对策略进行一次评估,策略评估中样本轨道数设为3000,对评估样本轨道结果取均值.图3给出了 Q 学习以及模拟退火 Q 学习的学习优化曲线,可以看出系统运行日代价的统计平均值随着学习步数的增加逐渐减少,最终稳定下来.从图中可以看出,模拟退火 Q 学习相较于传统的 Q 学习,收敛速率更快,且收敛时的日平均代价更低,其收敛区间的均值约为54820万元;图4给出了平均日弃风量的学习曲线,经过学习优化,平均日弃风量由200 MW降低到50 MW以下。

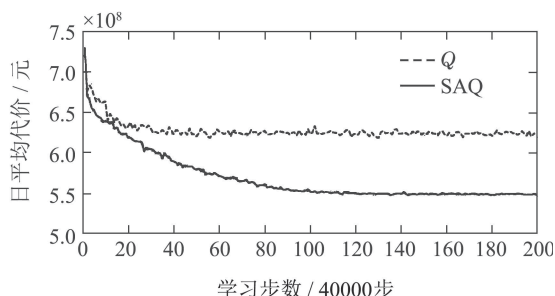


图3 系统总学习优化曲线

Fig. 3 The optimal curve of system

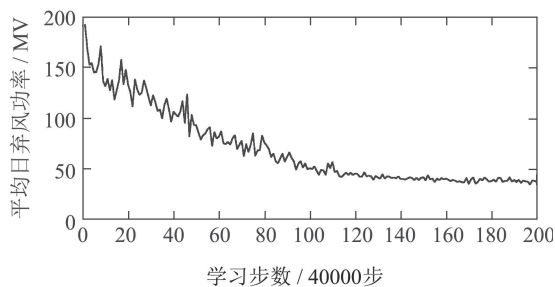


图4 平均日弃风量学习优化曲线

Fig. 4 The optimal curve of average daily wind

从图5可以看出平均每日未完成的日交易电量在学习后最终稳定在0左右,表明了优化后的策略可以保证日交易电量的输送.以上学习曲线说明了该

优化方法可以在满足日电量约束的情况下有效地降低系统运行代价,提升了运行效益。

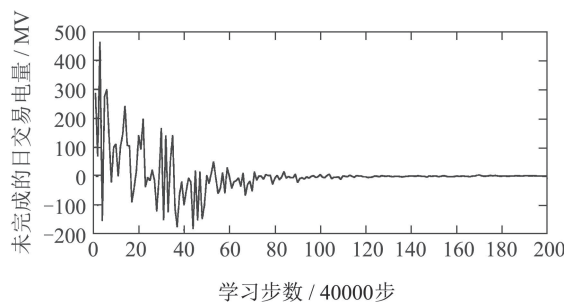


图5 日交易电量学习优化曲线

Fig. 5 The optimal curve of daily transacted electricity quantity

2) 不同模式下的策略比较及相关指标分析.

将文中所述考虑源荷随机性情况下的联络线与柔性负荷协同优化模式记为模式1;将仅考虑柔性负荷协同的优化模式记为模式2;将仅考虑源荷随机性的优化模式记为模式3.各模式下对应的策略分别记为策略1, 2, 3.将各随机量的随机波动的单位 Δ_{L1}^t , Δ_{L2}^t 及 Δ_w^t 都设置为0,使优化处在预测基值的情况下进行,以模拟模式2的优化过程:将DLC柔性负荷占比系数 μ 设置为0,以模拟模式3的优化过程.分别将其优化效果与模式1进行比较。

将模式1与模式3的优化策略在随机环境下进行评估比较.从图6可以看出在出现弃风主要集中在1~4时段和10~14时段,这两个时段风电出力与区域1负荷需求较为接近,由于等效火电机最低出力的限制以及风电出力和负荷需求的不确定性,如公式(13)所述,会导致出现弃风的情况.从图中可以明显看出模式1下的策略各时段产生的弃风量远小于模式2,说明了不考虑随机性制定的联络线计划并不能经济地适用于实际的随机环境,而本文制定各状态下行动策略能很好地适应随机环境,减少弃风量,促进新能源消纳。

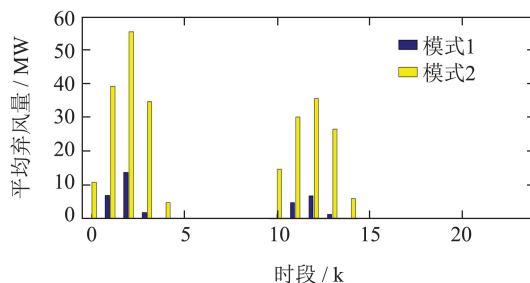


图6 不同模式下的弃风情况比较图

Fig. 6 Comparison diagram of wind power curtailment in different modes

由图7可以看出,模式1与模式3的策略相比,在负荷需求量高的9~18时段,负荷需求有明显差异,说明了策略1在这些时段执行了削负荷行动,在需求量大

较低的0~6时段,两种策略下的负荷需求相差不大,说明未削减负荷需求.从图上可以看出模式1下的策略发挥了柔性负荷的调峰作用,一定程度上降低了负荷峰谷差.

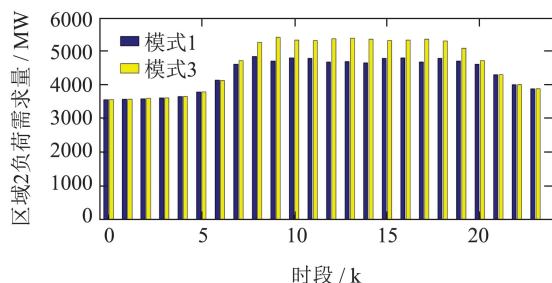


图7 不同模式下的负荷需求比较图

Fig. 7 Comparison diagram of load demand in different modes

从表2呈现的数据可以明显看出,本文模式1下的优化策略相较于另两种策略,更进一步地挖掘了互联电网的调度潜能,降低了每日的弃风量,提升了新能源的消纳.模式1中产生的发电机日运行费用及DLC日补偿金额的总和略大于模式3的发电机日运行费用,说明模式1得到的优化策略是在牺牲了部分经济性的前提下完成降低日负荷峰谷差的目标的.在实际的调度过程中,各优化目标加上更符合实际需求的权重,即可得到适用于不同实际需求的优化结果.

表2 不同模式下的相关指标

Table 2 Related indexes under different modes

	总体代 价/万元	弃风情况/ (MW·d ⁻¹)	发电机 运行费用/ (万元·d ⁻¹)	DLC 补偿金额/ (万元·d ⁻¹)
模式1	54820	38.09	34550	2144
模式2	84490	97.68	34870	2924
模式3	70520	50.13	36540	0

3) 不同模式下的策略在单样本中的结果分析.

取其中一条样本轨道进行分析,图8给出的是一条仿真样本轨道中各时段的风电出力及负荷需求的实际情况.由图8与图2比较可看出,在该样本轨道中风电及1区域负荷需求与预测值相比没有较大的变化,然而在0~14时段2区域负荷需求较预测值偏低,在18~23时段明显高于预测值.用上述的3种模式下的优化策略为这条样本轨道提供各时段的行动.

3种模式对应的优化策略下各时段输电情况如图9所示,3种模式下的日输电量和都为7200 MW,均满足日交易电量约束.由模式2的输电方案可知,若不考虑随机性,按照预测曲线制定的输电计划会在1~40及10~13风电出力较高的时段提高输电功率,

而在17~23时段输电量相对较低.由于所取样本轨道中2区负荷在18时段后出现高峰情况,策略1,3在18~23时段相较于策略2提升了输送电量,以满足2区负荷较预测偏高的情况.说明考虑了源荷随机性的策略能有效应对各类不确定情况,降低系统运行代价.

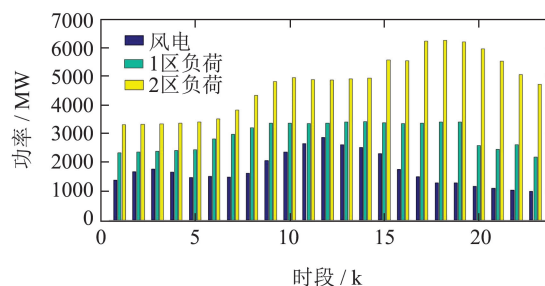


图8 样本中风电出力及负荷需求实际功率

Fig. 8 Practical wind power and load demand in a sample

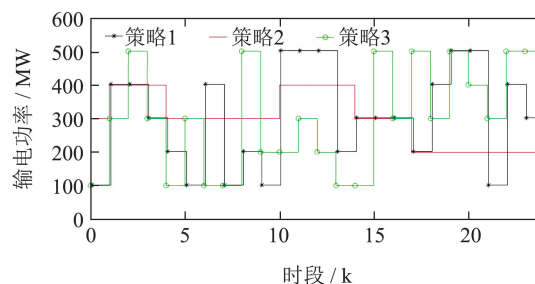


图9 不同模式下各时段的联络线功率

Fig. 9 Tie-line power at different periods in different modes

模式1下样本的各时段柔性负荷切除功率如图10所示.可以看出在负荷高峰期执行了削负荷命令,出现负荷高峰的几个时段,削负荷的等级也明显提升.表3给出了各策略在该样本轨道下产生的各项指标.模式1的日弃风量明显小于模式2与模式3,发电机的日运行费用也略小于模式2,3,说明了在单条样本下,考虑柔性负荷协同及源荷随机性的策略的优化效果更好,进一步证明了其优化方法的有效性.

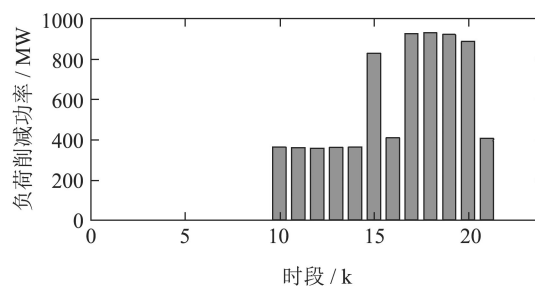


图10 模式1下的削负荷情况

Fig. 10 Reduction of load demand in mode 1

表3 单样本不同模式下的相关指标
Table 3 Related indexes under different modes in a sample

	弃风情况/ MW	发电机运行 费用/万元	DLC补偿金额/ 万元
模式1	1.4343	34774	2058
模式2	21.944	34218	2676
模式3	14.083	36696	0

5 总结

本文研究了源荷双侧随机性的互联电网联络线调度优化问题, 首先采用连续马尔科夫过程模型近似描述新能源与负荷随机不确定部分的变化, 然后将该联络线调度问题建立为相应的DTMDP模型, 最后通过算例仿真表明通过学习优化系统的运行代价显著降低且最终稳定, 说明将新能源出力及负荷需求的随机性和DLC负荷的可调节性考虑进互联电网联络线调度中, 能够有效应对随机情况, 提升电网运行效益。

本文主要针对互联电网上层的联络线调度进行研究, 而对上层联络线以及下层的常规机组进行细致化建模, 实现上下层协同优化也十分值得进一步研究; 同时还可考虑系统状态行动连续化, 并利用深度强化学习的方法解决可能出现的维数灾问题; 另外, 考虑电力市场交易环境下各区域间利益的博弈问题也是有意义的研究方向。

参考文献:

- [1] LIU Jizhen. Basic issues of the utilization of large-scale renewable power with high security and efficiency. *Proceedings of the CSEE*, 2013, 33(16): 1 – 8.
(刘吉臻. 大规模新能源电力安全高效利用基础问题. 中国电机工程学报, 2013, 33(16): 1 – 8.)
- [2] XIE L, CARVALHO P M S, FERREIRA L A F M, et al. Wind integration in power systems: operational challenges and possible solutions. *Proceedings of the IEEE*, 2011, 99(1): 214 – 232.
- [3] WANG Bin, XIA Ye, XIA Qing, et al. Model and methods of generation and transmission scheduling of inter-regional power grid via HVDC tie-line. *Automation of Electric Power Systems*, 2016, 40(3): 8 – 13.
(王斌, 夏叶, 夏清, 等. 直流跨区互联电网发输电计划模型与方法. 电力系统自动化, 2016, 40(3): 8 – 13.)
- [4] LIANG Xuming, CHEN Guoping, LOU Dianqiang, et al. A smart grid oriented architecture of operation and maintenance system based on centralized controlling and monitoring for inter-area HVDC power transmission project. *Power System Technology*, 2011, 35(3): 1 – 7.
(梁旭明, 陈国平, 娄殿强, 等. 面向智能电网的跨区直流输电集控化运维体系. 电网技术, 2011, 35(3): 1 – 7.)
- [5] ZHONG Haiwang, XIA Qing, DING Maosheng, et al. A new mode of HVDC tie-line operation optimization for maximizing renewable energy accommodation. *Automation of Electric Power Systems*, 2015, 39(3): 36 – 42.
(钟海旺, 夏清, 丁茂生, 等. 以直流联络线运行方式优化提升新能源消纳能力的新模式. 电力系统自动化, 2015, 39(3): 36 – 42.)
- [6] SHARMA M, PANDIT M, SRIVASTAVA L. Multi-area economic dispatch with tie-line constraints employing evolutionary approach. *International Journal of Engineering Science and Technology*, 2010, 2(3): 141 – 143.
- [7] ZHANG Yantao, REN Jing, SUN Xiaoqiang, et al. Two stage generation scheduling model based on coordination of regional and provincial power grid generalized tie-line. *Automation of Electric Power Systems*, 2015, 39(6): 128 – 133.
(张彦涛, 任景, 孙晓强, 等. 基于广义联络线网省协调的两阶段实时发电计划模型. 电力系统自动化, 2015, 39(6): 128 – 133.)
- [8] YANG Nan, WANG Bo, LIU Dichen, et al. An integrated supply-demand stochastic optimization method considering large-scale wind power and flexible load. *Proceedings of the CSEE*, 2013, 33(16): 63 – 69.
(杨楠, 王波, 刘涤尘, 等. 计及大规模风电和柔性负荷的电力系统供需侧联合随机调度方法. 中国电机工程学报, 2013, 33(16): 63 – 69.)
- [9] LIU Wenying, WEN Jing, XIE Chang, et al. Multi-objective optimal method considering wind power accommodation based on source-load coordination. *Proceedings of the CSEE*, 2015, 35(5): 1079 – 1088.
(刘文颖, 文晶, 谢昶, 等. 考虑风电消纳的电力系统源荷协调多目标优化方法. 中国电机工程学报, 2015, 35(5): 1079 – 1088.)
- [10] MA Taotao, GUO Chuangxin, CAO Yijia, et al. Current status and development trend of intelligent dispatching automation system of power system. *Automation of Electric Power Systems*, 2010, 34(9): 7 – 10.
(马韬韬, 郭创新, 曹一家, 等. 电网智能调度自动化系统研究现状及发展趋势. 电力系统自动化, 2010, 34(9): 7 – 11.)
- [11] XU Hongqiang, YAO Jianguo, NAN Guilin, et al. New features of application function for future dispatching and contral systems. *Automation of Electric Power Systems*, 2018, 42(1): 1 – 7.
(许洪强, 姚建国, 南贵林, 等. 未来电网调度控制系统应用功能的新特征. 电力系统自动化, 2018, 42(1): 1 – 7.)
- [12] ZHANG Qin, WANG Xifan, WANG Jianxue, et al. Survey of demand response research in deregulated electricity markets. *Automation of Electric Power Systems*, 2008, 32(3): 97 – 106.
(张钦, 王锡凡, 王建学, 等. 电力市场下需求响应研究综述. 电力系统自动化, 2008, 32(3): 97 – 106.)
- [13] LUO Yi, LIU Feng, LIU Xiangjie. Short-term wind power prediction based on principal component analysis and genetic neural network. *Power System Protection and Control*, 2012, 23(40): 47 – 53.
(罗毅, 刘峰, 刘向杰. 基于主成分-遗传神经网络的短期风电功率预测. 电力系统保护与控制, 2012, 23(40): 47 – 53.)
- [14] MU Pengtao, ZHAO Dongmei, WANG Jiacheng. Impact of high penetration wind power integration on small signal stability of regional power grid. *Automation of Electric Power Systems*, 2016, 40(11): 137 – 142.
(牟澎涛, 赵冬梅, 王嘉成. 高渗透率风电接入对区域电网小信号稳定性的影响. 电力系统自动化, 2016, 40(11): 137 – 142.)
- [15] CHEN Haoyong, WANG Yongchao, XUAN Peizheng, et al. Robust economic dispatch method of microgrid containing high propotion of wind power. *Control Theory & Applications*, 2017, 34(8): 1104 – 1111.
(陈皓勇, 王勇超, 禚培正, 等. 含高渗透率风电的微网系统鲁棒经济调度方法. 控制理论与应用, 2017, 34(8): 1104 – 1111.)
- [16] ZHAO Dongmei, YIN Haohan, WANG Jianfeng. Comprehensive analysis system of intermittent energy output characteristics and its application. *Southern Power System Technology*, 2015, 9(5): 7 – 14.
(赵冬梅, 尹颢涵, 王建锋. 间歇性能源出力特性综合分析体系及其应用. 南方电网技术, 2015, 9(5): 7 – 14.)
- [17] KANTZ H, HOLSTEIN D, RAGWITZ M, et al. Markov chain model for turbulent wind speed data. *Physica A Statistical Mechanics and Its Applications*, 2004, 342(1): 315 – 321.

- [18] CARPINONE A, GIORGIO M, LANGELLA R, et al. Markov chain modeling for very-short-term wind power forecasting. *Electric Power Systems Research*, 2015, 122(5): 152 – 158.
- [19] SUN Yuanzhang, WU Jun, LI Guojie, et al. Dynamic economic dispatch considering wind power penetration based on wind speed forecasting and stochastic programming. *Proceedings of the CSEE*, 2009, 29(4): 41 – 47.
(孙元章, 吴俊, 李国杰, 等. 基于风速预测和随机规划的含风电场电力系统动态经济调度. 中国电机工程学报, 2009, 29(4): 41 – 47.)
- [20] BREIPOHL A M, LEE F N, ZHAI D, et al. A gauss-markov load model for application in risk evaluation and production simulation. *Transaction on Power System*, 1992, 4(7): 1493 – 1499.
- [21] ZHOU B, CHAN K W, YU T. Q-learning approach for hierarchical agc scheme of interconnected power grids. *2011 IEEE International Conference on Smart Grid and Clean Energy Technologies (ICSGCE)*. Chengdu: IEEE, 2011, 12: 43 – 52.
- [22] ARLOTTO A, STEELE J M. Finite horizon Markov decision problems and a central limit theorem for total reward. *Mathematics*, 2015, 41(4): 1448 – 1468.
- [23] YIN B Q, LI Y J, ZHOU Y P, et al. Performance optimization of semi-markov decision processes with discounted-cost criteria. *European Journal of Control*, 2008, 14(3): 213 – 222.
- [24] GUO X P, CAO X R. Optimal control of continuous-time Markov chains with average sample path rewards. *SIAM Journal on Control and Optimization*, 2005, 44(1): 29 – 48.
- [25] WANG Y Z, XIE Q, AMMARI A, et al. Deriving a near-optimal power management policy using model-free reinforcement learning and Bayesian classification. *Proceedings of the 48th Design Automation Conference*. New York: IEEE, 2011, DOI: 10.1145/2024724.2024735.
- [26] XU H, JAGANNATHAN S. Model-free Q-learning over finite horizon for uncertain linear continuous-time systems. *2014 IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning (ADPRL)*. New York: IEEE, 2014: 1 – 6.
- [27] YU Tao, ZHOU Bin, ZHEN Weiguo. Application and development of reinforcement learning theory in power systems. *Power System Protection and Control*, 2009, 37(14): 122 – 128.
(余涛, 周斌, 甄卫国. 强化学习理论在电力系统中的应用及展望. 电力系统保护与控制, 2009, 37(14): 122 – 128.)

作者简介:

张 延 硕士, 目前研究方向为源荷不确定电力系统的调度优化,

E-mail: zhangyan@mail.hfut.edu.cn;

唐 昊 教授, 博士生导师, 目前研究方向为随机动态系统的控制与优化、智能电网中的优化与控制方法等, E-mail: htang@hfut.edu.cn;

王 珂 高级工程师, 目前研究方向为交直流电力系统的分析与控制、高级应用软件开发和工程化等, E-mail: wangke@epri.sgcc.com.cn;

潘 毅 高级工程师, 目前研究方向为电力系统EMS, DTS及其标准化等, E-mail: pany@epri.ac.cn;

李怡瑾 博士, 目前研究方向为多类型能量系统的建模与优化, E-mail: yjlee@mail.hfut.edu.cn.