

互联电网自动发电控制功率分配的改进逼近于理想解的 排序-Q多目标优化算法

张孝顺[†], 余涛

(华南理工大学 电力学院, 广东 广州 510640; 广东省绿色能源技术重点实验室, 广东 广州 510640)

摘要: 本文提出了一种多目标决策与强化学习相结合的改进的逼近于理想解的排序(technique for order preference similar to an ideal solution, TOPSIS)-Q算法, 有效解决了自动发电控制(automatic generation control, AGC)总功率指令分配到风电、水电、火电等各类AGC机组的动态随机多目标优化问题. 算法采用3个不同的目标Q值矩阵进行迭代更新, 然后利用改进TOPSIS方法对Q值矩阵进行多目标决策处理, 客观地给出各目标的动态最优权重系数, 从而得到各状态-动作对的综合评价判据. IEEE标准两区域模型仿真研究验证了改进TOPSIS-Q算法在AGC机组功率多目标动态优化分配过程应用的可行性和有效性, 在复杂随机扰动的环境中提高系统CPS性能的同时, 有效降低了AGC机组调节成本和碳排放.

关键词: 改进TOPSIS-Q; 强化学习; 多目标优化; AGC功率分配

中图分类号: TP71 文献标识码: A

Stochastic optimal generation command dispatch of interconnected power grids based on improved multi-objective technique for order preference similar to an ideal solution-Q algorithm

ZHANG Xiao-shun[†], YU Tao

(College of Electric Power, South China University of Technology, Guangzhou Guangdong 510640, China;
Guangdong Key Laboratory of Clean Energy Technology, Guangzhou Guangdong 510640, China)

Abstract: This paper proposes an improved technique for order preference similar to an ideal solution (TOPSIS)-Q learning approach to solve the dynamic optimization of generation command dispatch (GCD) for automatic generation control (AGC) in a multi-energy power system. Three optimization objectives are simultaneously optimized by three different Q-value matrixes. Then dynamic optimal weight of each objective is calculated by improved TOPSIS method such that the evaluation criterion of each state-action is obtained. Case studies are carried out to evaluate the optimization performance of the proposed algorithm in the two-area load frequency control (LFC) power system model. Simulation results indicate that the proposed method is feasible and effective for dynamic optimization of GCD problem, and can reduce the regulating cost and carbon emissions while improving the control performance of AGC systems in the complex stochastic power system.

Key words: improved TOPSIS-Q; reinforcement learning; multi-objective optimization; AGC generation command dispatch

1 引言(Introduction)

作为电网调度中心的重要控制系统之一, 自动发电控制(automatic generation control, AGC)控制器的主要任务是通过实时调整AGC机组的发电功率来应对随机的负荷扰动, 使得互联电网频率和联络线交换功率保持在额定值^[1-2]. 当AGC控制器根据电网负荷

扰动得到一个总发电功率指令后, 如何将总发电功率指令实时最优分配到各台AGC机组就成为一个研究难点.

目前, 国内外关于AGC机组功率指令分配的研究还较少. 在工程应用上, 调度人员往往根据工作经验或机组的可调容量、调节速率等某一特性来进行AGC

收稿日期: 2014-07-11; 录用日期: 2015-01-07.

[†]通信作者. E-mail: xszhang1990@sina.cn; Tel.: +86 15017527246.

国家重点基础研究发展计划(“973”计划)项目(2013CB228205), 国家自然科学基金项目(51177051, 51477055), 中国南方电网科技项目资助.

Supported by National Key Basic Research Program of China (“973” Program (2013CB228205)), National Natural Science Foundation of China (51177051, 51477055) and Science & Technology Project of China Southern Power Grid Company.

功率分配^[3],并没有考虑到各机组间的其他非线性约束,难以满足复杂工况下的控制性能评价要求(control performance standard, CPS)性能要求.文献[4]采用了粒子群优化(particle swarm optimization, PSO)算法进行AGC功率分配,有效解决了CPS指令的最优分配问题.针对AGC功率最优分配的几个难点,笔者采用了具有很好随机优化能力的强化学习Q方法开展了系列研究^[5-8].文献[6]在算法迭代更新中引入了资格迹,有效解决了火电机组大时延问题,文献[7-8]则对AGC机组进行聚类分层,有效避免了功率分配的维数灾难问题.然而,上述方法都只是针对某一个控制目标,或在Q算法的奖励函数采用线性加权的方式进行多目标优化,并没有考虑各个目标之间量纲的差异,同时采用了较为主观的固定目标权重系数,从而难以满足AGC机组功率分配的多目标动态随机优化需要.

在进行多目标决策时,常用的方法有线性加权和法、改进的逼近于理想解的排序法(technique for order preference similar to an ideal solution, TOPSIS)、目标规划法、交互式法等^[9-11],这些方法都需要确定各目标的权重系数,一般采用主观赋权法或客观赋权法.相比主观赋权法,客观赋权法得到的权值系数更加恰当、合理,但其计算方法大多比较繁琐,难以满足快速多目标决策的需求,尤其是在秒级时间尺度的AGC系统应用中.

传统TOPSIS法,又称为理想点法,是一种简便快速的多目标决策方法,通过构造多目标的最理想点,然后以各方案到理想点距离的加权平均和作为综合评价标准,但各目标的权重系数也是主观地确定.与之相比,改进TOPSIS法^[12]能够充分利用决策矩阵的信息,构造出各目标权重的最优化模型,客观地赋以各目标的最优权重系数.为更好解决Q学习算法的快速多目标决策问题,本文提出一种改进TOPSIS-Q算法,在算法中以不同的Q值矩阵代表各个优化目标,然后利用改进TOPSIS法进行多目标决策.同时,为适应间歇式能源的大量并网,本文还在功率分配模型中增加了风电机组.最后,本文利用了IEEE标准两区域模型对改进TOPSIS-Q算法进行了验证,并与其他分配方法进行了分析比较.

2 AGC机组功率分配模型(Model of generation command dispatch)

如图1所示^[8],AGC闭环控制分为两个过程:AGC控制和功率指令分配.在电网发生负荷扰动 ΔP_L 后,AGC控制器计算出一个AGC总发电指令 $\Delta P_{order-\Sigma}$,然后TOPSIS-Q动态优化分配器根据所采集数据状态将总指令 $\Delta P_{order-\Sigma}$ 分配到每台调节机组 $\Delta P_{order-i}$,最后每台机组实际输出功率 ΔP_{Gi} .本文AGC控制器采用PI控制结构^[13],并设置AGC控制周期为8s.

本文在AGC功率分配的过程中考虑了功率偏差、调节成本和碳排放3个目标,具体描述如下:

$$\left\{ \begin{array}{l} \min f_1 = \sum_{k=1}^T \sum_{i=1}^n \Delta P_{error-i}^2(k), \\ \min f_2 = \sum_{k=1}^T \sum_{i=1}^n C_i [\Delta P_{Gi}(k)], \\ \min f_3 = \sum_{k=1}^T \sum_{i=1}^n D_i [\Delta P_{Gi}(k)], \\ \text{s.t. } P_{ratei}^- \leq \Delta P_{order-i}^{ch}(k) \leq P_{ratei}^+, \\ \Delta P_{Gi}^{\min} \leq \Delta P_{Gi}(k) \leq \Delta P_{Gi}^{\max}, \\ \Delta P_{order-\Sigma}(k) = \sum_{i=1}^n \Delta P_{order-i}(k), \end{array} \right. \quad (1)$$

式中: k 为迭代次数; $\Delta P_{error-i}$ 为第 i 台机组接收的功率指令与该机组实际出力的偏差; f_1, f_2, f_3 分别为在总迭代次数 T 内的各个累积目标函数值; C_i 为第 i 台机组的调节成本系数,元·MWh⁻¹; D_i 为第 i 台机组的碳排放强度系数,kg/kWh; $\Delta P_{order-i}^{ch}$ 为第 i 台机组接收的功率指令的变化值; P_{ratei}^+, P_{ratei}^- 分别为第 i 台机组的上升和下降调节速率限制,MW/min; $\Delta P_{Gi}^{\max}, \Delta P_{Gi}^{\min}$ 分别为第 i 台机组的调节容量上下限,MW.

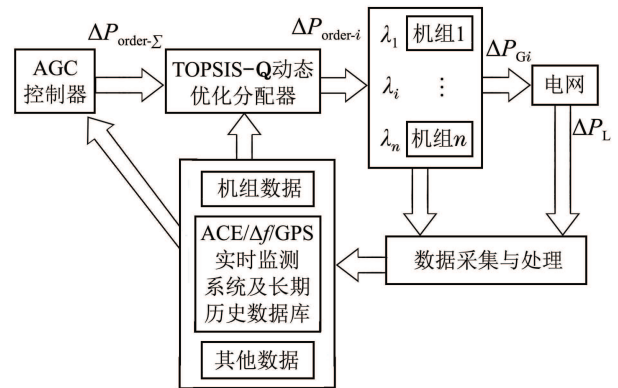


图1 AGC机组功率动态分配优化过程

Fig. 1 Dynamic optimization process for AGC units

3 改进的TOPSIS-Q算法原理(Improved TOPSIS-Q learning algorithm)

3.1 Q学习算法(Q-learning algorithm)

本文在Q学习算法设计中^[14],结合了TD(λ)算法多步回报的思想^[15],在各个目标值函数的迭代更新过程中引入资格迹^[16].其中,资格迹更新如下所示:

$$e_k(s, a) = \begin{cases} \gamma \lambda e_{k-1}(s, a) + 1, & \text{若}(s, a) = (s_k, a_k), \\ \gamma \lambda e_{k-1}(s, a), & \text{其他,} \end{cases} \quad (2)$$

式中: $e_k(s, a)$ 为第 k 次迭代对应的状态-动作对 (s, a) 下的资格迹; (s_k, a_k) 为第 k 次迭代实际的状态-动作对; γ 为折扣因子, $0 \leq \gamma \leq 1$,本文取为0.3; λ 为衰减因子, $0 \leq \lambda \leq 1$,本文取为0.5.

在引入资格迹后,对于第 m 个控制目标,其 Q 值函数迭代更新如下^[16]:

$$Q_{k+1}^m(s, a) = Q_k^m(s, a) + \alpha \delta_k^m e_k(s, a), \quad (3)$$

$$Q_{k+1}^m(s_k, a_k) = Q_{k+1}^m(s_k, a_k) + \alpha \rho_k^m, \quad (4)$$

其中:

$$\rho_k^m = R^m(s_k, s_{k+1}, a_k) + \gamma Q_k^m(s_{k+1}, a_g) - Q_k^m(s_k, a_k),$$

$$\delta_k^m = R^m(s_k, s_{k+1}, a_k) + \gamma Q_k^m(s_{k+1}, a_g) - Q_k^m(s_k, a_g),$$

式中: α 为学习因子,并且 $0 \leq \alpha \leq 1$,本文取为0.2; $R^m(s_k, s_{k+1}, a_k)$ 是第 m 个优化目标在第 k 次迭代时刻环境由状态 s_k 经动作 a_k 转移到 s_{k+1} 后的奖励函数值; a_g 为贪婪动作策略,即多目标决策评价最好的优化动作。

3.2 改进 TOPSIS 多目标决策(Improved TOPSIS for multi-objective decision)

本文提出的改进TOPSIS法主要是用来求解出多个优化目标下的贪婪动作策略 a_g 。改进TOPSIS法是利用各目标 Q 值矩阵的信息,通过构造多目标问题的理想点解,客观地赋予各目标的权重系数,并以各方案到理想点距离的加权平方和作为综合评价的数据^[12]。

1) Q 值矩阵归一化。

由于各 Q 值矩阵代表的目标量纲不同,这就需要对 Q 值矩阵进行适当的变换,化为无量纲的标准化矩阵。本文采用极差变换法对 Q 值矩阵进行归一化处理,对于第 m 个控制目标, $a \in A$, 有

$$y^m(s_k, a) = \frac{Q_k^m(s_k, a) - \min Q_k^m(s_k, a)}{\max Q_k^m(s_k, a) - \min Q_k^m(s_k, a)}. \quad (5)$$

由于 Q 算法是在学习过程中使得智能体的长期积累奖励值最大,结合式(5),本文提出的TOPSIS理想点解为(1, 1, 1)。

2) 动态最优权重系数求解。

TOPSIS法用各动作方案到理想点解的距离平方作为综合评价准则,如下所示:

$$d_a = \sum_{m=1}^3 [y^m(s_k, a) - 1]^2 \omega_m^2. \quad (6)$$

为确定最优权重系数 ω_m^2 ,可构造最优化模型如下:

$$\begin{cases} \min Z = \sum_{a \in A} d_a = \sum_{a \in A} \sum_{m=1}^3 [y^m(s_k, a) - 1]^2 \omega_m^2, \\ \text{s.t. } \sum_{m=1}^3 \omega_m^2 = 1. \end{cases} \quad (7)$$

通过构造拉格朗日函数,求解上述模型,可得最优

权重系数为^[12]

$$\omega_m = \frac{1}{\left(\sum_{m=1}^3 \frac{1}{d^m}\right)(d^m)}, \quad (8)$$

其中 $d^m = \sum_{a \in A} [y^m(s_k, a) - 1]^2$ 。

3) 动作策略选择。

在得到最优权重后,由式(6)即可得到贪婪动作策略,如下:

$$a_g = \arg \min_{a \in A} \left(\sum_{m=1}^3 [y^m(s_k, a) - 1]^2 \omega_m^2 \right). \quad (9)$$

如果在改进TOPSIS-Q算法学习过程中每次迭代都执行贪婪策略,算法就不能充分搜索动作空间 A 的其他动作,就有可能得到局部最优解。本文采用一种基于概率分布选择动作的追踪算法^[17]来构造动作选择策略。初始化时,每个动作的选择概率是相等的,在算法学习过程中,各状态下的动作概率分布更新如下^[18]:

$$\begin{cases} P_s^{k+1}(a_g) = P_s^k(a_g) + \beta(1 - P_s^k(a_g)), \\ P_s^{k+1}(a) = P_s^k(a)(1 - \beta), \forall a \in A, a \neq a_g, \\ P_{s'}^{k+1}(a) = P_{s'}^k(a), \forall a \in A, \forall s' \in S, s' \neq s, \end{cases} \quad (10)$$

式中: β 代表动作搜索速度, β 值越小,则搜索范围越大,本文取 β 为0.1; $P_s^k(a)$ 代表第 k 次迭代时状态 s 下动作 a 被选择的概率。

4 基于改进TOPSIS-Q的AGC机组功率分配算法 (Improved TOPSIS-Q learning algorithm for AGC order dispatch)

4.1 算法状态-动作设计(Selection of state-action)

如图1所示,优化分配器在每个控制周期都将输入一个功率总指令 $\Delta P_{\text{order}-\Sigma}$,本文将 $\Delta P_{\text{order}-\Sigma}$ 分成若干个离散的范围区间,每个区间对应一个状态。

本文将AGC机组的分配因子组合与算法的动作 a 相对应。当存在 n 个AGC机组, r 个动作组合时,其动作空间可表示如下:

$$A = [a_1, a_2, \dots, a_r] = [(\lambda_{11}, \lambda_{12}, \dots, \lambda_{1n}), (\lambda_{21}, \lambda_{22}, \dots, \lambda_{2n}), \dots, (\lambda_{r1}, \lambda_{r2}, \dots, \lambda_{rn})], \quad (11)$$

式中 λ 代表AGC机组的功率分配因子。

4.2 奖励函数设计(Design of reward function)

笔者在之前的 $Q(\lambda)$ 算法多目标研究工作中,只是简单地在奖励函数中采用线性加权进行多目标处理。在本文提出的3个目标函数下,其奖励函数如下所示:

$$R(s_k, s_{k+1}, a_k) =$$

$$-\sum_{i=1}^n \Delta P_{\text{error}-i}^2(k) - \mu_1 \sum_{i=1}^n C_i [\Delta P_{G_i}(k)] - \mu_2 \sum_{i=1}^n D_i [\Delta P_{G_i}(k)], \quad (12)$$

式中 μ_1, μ_2 分别代表调节成本和碳排放的目标权重系数。

Q算法是不断最大化期望折扣奖励R的学习过程。因此,由式(1)可将改进TOPSIS-Q算法下各个目标的奖励函数设计如下:

$$\begin{cases} R^1(s_k, s_{k+1}, a_k) = -\sum_{i=1}^n \Delta P_{\text{error}-i}^2(k), \\ R^2(s_k, s_{k+1}, a_k) = -\sum_{i=1}^n C_i [\Delta P_{G_i}(k)], \\ R^3(s_k, s_{k+1}, a_k) = -\sum_{i=1}^n D_i [\Delta P_{G_i}(k)]. \end{cases} \quad (13)$$

通过比较式(12)–(13)可以发现, $Q(\lambda)$ 算法下各个目标的权重系数难以确定,需要通过大量仿真试验才能找到较为合理的权重系数。而在TOPSIS-Q算法下,不同量纲的目标函数先按式(5)进行归一化处理,然后通过构造最优化模型实时求解出各个目标的最优权重系数,更加合理简便。

4.3 算法流程(Algorithm flow)

综上所述,本文提出的最优分配算法主要包括以下3个过程:①首先,获得当前状态和奖励函数值,更新各个优化目标的Q值矩阵;②利用改进TOPSIS多目标决策法求解出贪婪动作策略;③根据贪婪动作策略更新状态动作概率分布,执行动作。具体可见图2。

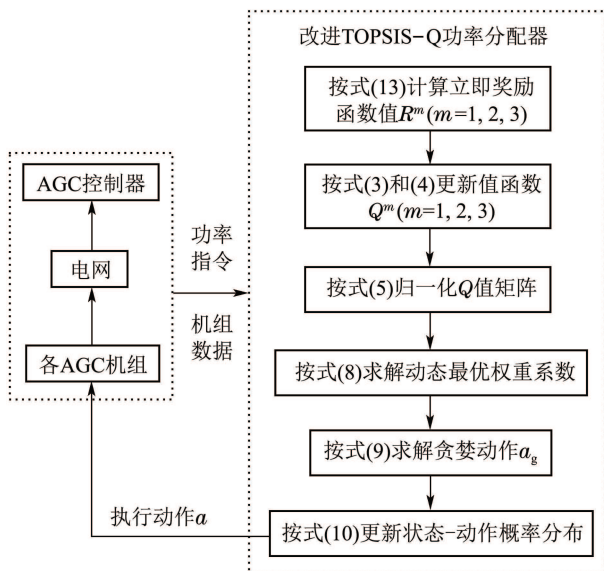


图2 改进TOPSIS-Q算法流程图

Fig. 2 Flow diagram of improved TOPSIS-Q algorithm

5 仿真研究(Simulation studies)

5.1 仿真模型(Simulation model)

本文以标准IEEE两区域互联系统的负荷频率控制

模型^[19]作为研究对象,在A区域使用4种机组模型替代原来的1台机组,分别为燃煤、液化天然气(liquefied natural gas, LNG)、水电机组和风电机组,其中4种机组模型参考文献[5, 20]建立, B区域仍使用原来的1台机组模型。其中, AGC控制周期设置为8 s, 系统模型具体相关参数如表1–2所示。

表1 两区域互联系统模型参数

Table 1 System parameters of the two-area LFC model

| T_g/s | T_t/s | T_p/s | $R/(\text{Hz}\cdot\text{pu}^{-1})$ | $K_p/(\text{Hz}\cdot\text{pu}^{-1})$ | $T_{12}\cdot\text{s}^{-1}$ |
|---------|---------|---------|------------------------------------|--------------------------------------|----------------------------|
| 0.08 | 0.3 | 20 | 2.4 | 120 | 0.545 |

表2 各类型机组的相关参数

Table 2 Parameters of various types of models

| 类型 | $\Delta P_{G_i}^{\max}$ | $\Delta P_{G_i}^{\min}$ | $P_{\text{rate}i}^+$ | $P_{\text{rate}i}^-$ | C_i | D_i |
|-----|-------------------------|-------------------------|----------------------|----------------------|--------|-------|
| 煤电 | 2600 | -2600 | 100.44 | -100.44 | 127.04 | 1.06 |
| LNG | 750 | -750 | 80.1 | -80.1 | 253.40 | 0.5 |
| 水电 | 500 | 0 | 500 | -500 | 93.65 | 0 |
| 风电 | 300 | 0 | 300 | -300 | 52.30 | 0 |

5.2 仿真设计(Simulation design)

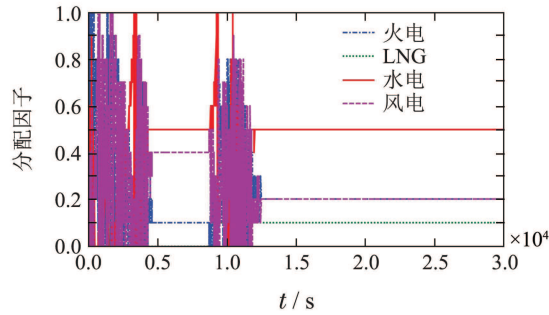
算法的状态空间以功率总指令 $\Delta P_{\text{order}-\Sigma}$ 为输入状态,并将其离散化为 $(-\infty, -1000)$, $[-1000, -500)$, $[-500, 0)$, $[0, 500)$, $[500, 1000)$, $[1000, +\infty)$, 单位MW。动作空间 $A = [(0, 0, 0, 1), (0, 0, 0.9, 0.1), (0, 0, 0.8, 0.2), \dots, (0.1, 0.9, 0, 0), (1, 0, 0, 0)]$, 共286个离散动作。

在仿真研究中,本文还引入按相同可调容量比例分配方法PROP^[3], $Q(\lambda)$ 算法^[6]与改进TOPSIS-Q算法进行比较分析。其中, $Q(\lambda)$ 算法利用奖励函数线性加权进行多目标优化,如式(12)所示,在仿真算例中, μ_1, μ_2 分别取为0.5。

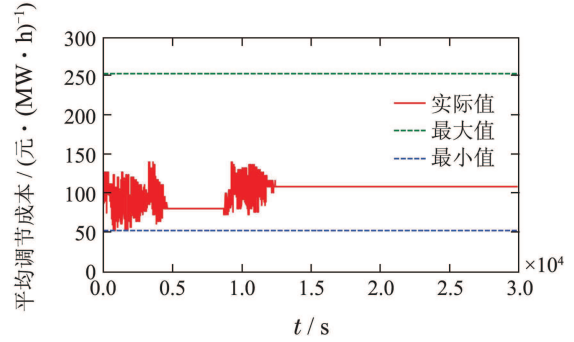
5.3 预学习仿真(Study on pre-learning process)

强化学习算法的应用需要经历探索与试错的“预学习”过程。其中,学习模式有两种:在线学习和离线学习。假如预学习过程采用在线模式,算法在随机探索寻优的同时可能危害系统稳定性。因此,本文采用离线模式进行算法预学习,直到所有的Q值矩阵都收敛。在完成预学习后,再把收敛后的Q值矩阵及其算法中应用到实际AGC系统中在线优化运行。

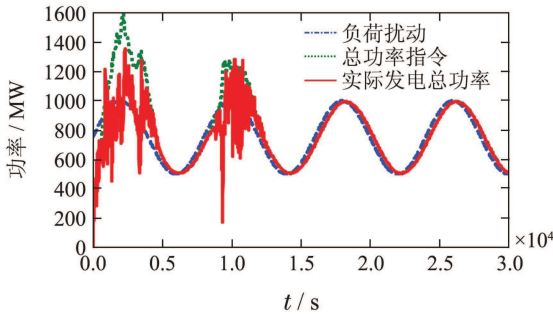
图3给出了负荷扰动为[500, 1000) MW范围内(即状态5)的算法预学习收敛过程,从图3中可以看出在15000 s左右时算法开始收敛。其中:如图3(a)所示,调节速度较快、经济性高、无碳排放的水电机组的分配因子最大,承担了大部分负荷扰动;从图3(c)可以看到,当算法迭代到2000次时,各个动作点与理想点的分布情况,评价最好的贪婪动作点则最接近理想点。



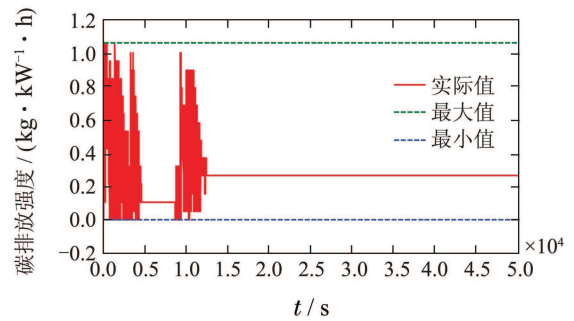
(a) 分配因子实时曲线



(b) 平均调节成本实时曲线



(b) 功率输出曲线



(c) 碳排放强度实时曲线

图4 各目标的收敛曲线

Fig. 4 Convergence curve of each objective

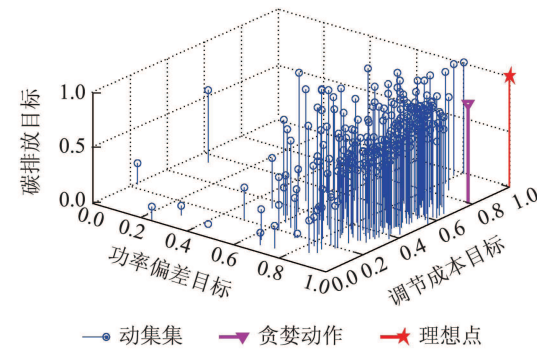
算法在其他状态下也进行相应的预学习, 最后把学习完毕的Q值矩阵投入在线学习即可。

5.4 随机负荷扰动仿真(Study on statistical experiments)

为验证改进TOPSIS算法在随机变化的复杂电网的适应性, 本节将在A区域施加周期为2000 s、幅值不超过1400 MW的随机负荷扰动进行24 h的统计性实验。其中, 在仿真算例中还考虑不同季节、不同时段下水电及风电机组可调容量的变化。在丰水期水电机组可调容量为500 MW, 枯水期降至300 MW; 风电机组以海上风电为例^[21], 晚上(0 ~ 12 h)风速较大, 可调容量为300 MW, 白天(12 h~24 h)风速较小, 可调容量为200 MW。

1) 丰水期仿真分析。

丰水期的仿真试验结果如表3所示。其中: $|\Delta f|$, $|E_{ACE}|$, CPS1均为仿真时间内平均值, 费用与碳排放量均为仿真时间内的累积值。从表3中可以看出: 3种算法下的CPS1平均值相差不大, 然而改进TOPSIS-Q与 $Q(\lambda)$ 这两种强化学习算法相比PROP方法, 调节费用和碳排放量都少得多。这说明了在PROP方法下, 水电和风电机组容量承担负荷扰动的比重较小。其中, 相比 $Q(\lambda)$ 算法, 改进TOPSIS-Q算法下机组功率分配的经济性更好、更加低碳。此外, 相比白天时段, 晚上时段各CPS1平均值更高, 调节费用和碳排放量也更少, 这也说明了风电机组的快速调节能力、经济性好

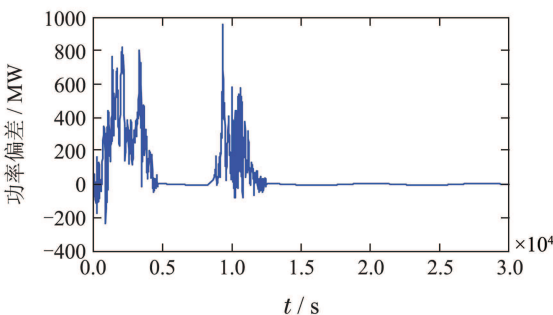


(c) 改进TOPSIS多目标决策(迭代步数= 2000时)

图3 状态5的预学习收敛过程

Fig. 3 Pre-learning convergence process of the fifth state

图4给出了相应的各个目标的收敛曲线。从图4(a)可以看到, 功率偏差在收敛后基本趋近于零。图4(b)与4(c)的平均调节成本曲线和碳排放强度曲线都没有收敛到最小值, 这是由于水电与风电机组的调节容量所决定的; 另一方面, 这两个目标值都远小于其最大值, 说明了经济性更高、无碳排放的水电和风电机组承担了大部分负荷扰动。



(a) 功率偏差实时曲线

与绿色环保等特点.

2) 枯水期仿真分析.

枯水期的仿真试验结果如表4所示. 从表中可以看出: 相比丰水期, 枯水期下各算法的仿真效果更差, 这也表明了水电机组的快速、经济、低碳调节对AGC控

制的贡献. 与丰水期一样, PROP方法下的调节费用与碳排放量最多, 白天比晚上时段的总体效果更差. 此外, 在枯水期内, $Q(\lambda)$ 算法下的经济性更好, 改进TOPSIS-Q算法下AGC系统的CPS1平均值更高, 更加低碳.

表3 丰水期仿真试验比较表

Table 3 Simulation experiment comparison in July

| 时段 | 算法 | $ \Delta f /\text{Hz}$ | $ E_{ACE} /(\text{MW})$ | CPS1/% | 费用/万元 | 碳排放量/t |
|----|--------------|------------------------|-------------------------|--------|-------|--------|
| 晚上 | 改进TOPSIS-Q | 0.0194 | 60.30 | 193.41 | 41.76 | 1185 |
| | $Q(\lambda)$ | 0.0194 | 60.39 | 193.35 | 44.64 | 1107 |
| | PROP | 0.0197 | 61.43 | 193.01 | 55.02 | 2956 |
| 白天 | 改进TOPSIS-Q | 0.0234 | 72.23 | 190.51 | 53.53 | 1089 |
| | $Q(\lambda)$ | 0.0230 | 71.44 | 190.81 | 58.89 | 1643 |
| | PROP | 0.0230 | 72.19 | 190.60 | 63.07 | 3390 |

表4 枯水期仿真试验比较表

Table 4 Simulation experiment comparison in December

| 时段 | 算法 | $ \Delta f /\text{Hz}$ | $ E_{ACE} /\text{MW}$ | CPS1/% | 费用/万元 | 碳排放量/t |
|----|--------------|------------------------|-----------------------|--------|-------|--------|
| 晚上 | 改进TOPSIS-Q | 0.0198 | 61.37 | 193.17 | 52.87 | 1489 |
| | $Q(\lambda)$ | 0.0202 | 61.99 | 192.94 | 50.27 | 1533 |
| | PROP | 0.0197 | 61.62 | 192.92 | 55.97 | 3099 |
| 白天 | 改进TOPSIS-Q | 0.0237 | 73.82 | 190.31 | 60.43 | 1431 |
| | $Q(\lambda)$ | 0.0242 | 75.03 | 189.91 | 51.45 | 2106 |
| | PROP | 0.0232 | 72.56 | 190.44 | 63.67 | 3572 |

5.5 讨论(Discussions)

与其他多目标优化算法一样, 当优化目标个数增加时, TOPSIS-Q算法可能收敛到不同的最优折中解, 其求解速度也会大大下降, 就有可能难以满足AGC的实时控制要求. 然而, 本文采用的多目标强化学习方法, 每个 Q 值矩阵分别代表一个目标函数, 可以单独进行迭代更新, 易于利用多台计算机开展并行计算, 然后再集中利用所有的矩阵信息进行多目标决策, 如图5所示.

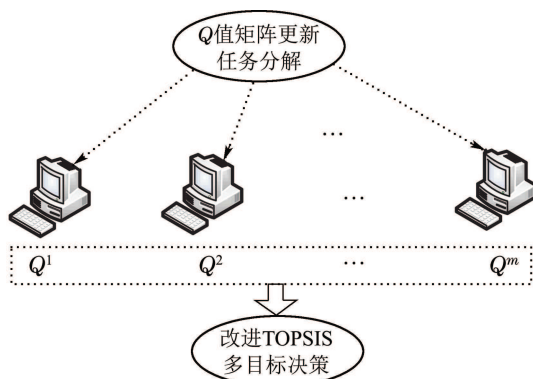


图5 Q 值矩阵并行计算示意图

Fig. 5 Parallel computing of Q -value matrixes

当采用并行计算方式对不同的 Q 值进行迭代更新时, 算法收敛计算时间也将基本不变, 可以满足AGC的动态实时控制.

6 结论(Conclusions)

本文提出了一种基于改进TOPSIS-Q的AGC机组功率最优分配算法, 主要有以下优势:

1) 结合了改进TOPSIS多目标决策方法, 满足了实时性要求很高的AGC闭环控制系统的动态优化需求, 在求解动态最优权重系数后, 贪婪动作策略的确定更加合理.

2) 与经典强化学习算法一样, 具有很好的在线学习能力, 有效解决了风电接入后随机电力系统的AGC功率指令动态分配随机优化问题.

3) 改进TOPSIS-Q算法能在保证CPS性能的同时, 有效降低AGC系统的发电调节成本, 同时使得AGC的调节更加低碳环保.

参考文献(References):

- [1] JALEELI N, VANSLYCK L S, EWART N, et al. Understanding automatic generation control [J]. *IEEE Transactions on Power Systems*,

- 1992, 7(3): 1106 – 1122.
- [2] 余涛, 张水平. 基于5要素试错更新算法SARSA(λ)的自动发电控制[J]. 控制理论与应用, 2013, 30(10): 1246 – 1251.
(YU Tao, ZHANG Shuiping. Automatic control of electricity generation based 5-componnet update learning algorithm SARSA(λ) [J]. *Control Theory & Applications*, 2013, 30(10): 1246 – 1251.)
- [3] 高宗和. 自动发电控制算法的几点改进[J]. 电力系统自动化, 2001, 25(22): 49 – 51.
(GAO Zonghe. Some algorithm improvement on AGC software [J]. *Automation of Electric Power Systems*, 2001, 25(22): 49 – 51.)
- [4] 刘斌, 王克英, 余涛, 等. PSO算法在互联网CPS功率调节中的应用研究[J]. 电力系统保护与控制, 2009, 37(6): 36 – 39.
(LIU Bin, WANG Keying, YU Tao, et al. Study on the application of particle swarm optimization algorithm to power regulation of CPS in interconnected power grids [J]. *Power System Protection and Control*, 2009, 37(6): 36 – 39.)
- [5] 余涛, 王宇名, 刘前进. 互联网CPS调节指令动态最优分配Q-学习算法[J]. 中国电机工程学报, 2010, 30(7): 62 – 69.
(YU Tao, WANG Yuming, LIU Qianjin. Q-learning based dynamic optimal allocation algorithm for CPS order of interconnected power grids [J]. *Proceeding of CSEE*, 2010, 30(7): 62 – 69.)
- [6] 余涛, 王宇名, 甄卫国, 等. 基于多步回溯Q学习的自动发电控制指令动态优化分配算法[J]. 控制理论与应用, 2011, 28(1): 58 – 69.
(YU Tao, WANG Yuming, ZHEN Weigu, et al. Multi-step backtrack Q-learning based dynamic optimal algorithm for auto generation control order dispatch [J]. *Control Theory & Applications*, 2011, 28(1): 58 – 69.)
- [7] 余涛, 王宇名, 叶文加, 等. 基于改进分层强化学习的CPS指令动态优化分配算法[J]. 中国电机工程学报, 2011, 31(19): 90 – 96.
(YU Tao, WANG Yuming, YE Wenjia, et al. Multi-objective dynamic optimal dispatch method for CPS order of interconnected power grids using improved hierarchical reinforcement learning [J]. *Proceeding of CSEE*, 2011, 31(19): 90 – 96.)
- [8] YU T, WANG Y M, YE W J, et al. Stochastic optimal generation command dispatch based on improved hierarchical reinforcement learning approach [J]. *IET Generation, Transmission & Distribution*, 2011, 5(8): 789 – 797.
- [9] 许叶军, 达庆利. 区间混合判断矩阵排序的线性目标规划法[J]. 系统工程与电子技术, 2008, 30(6): 1079 – 1085.
(XU Yejun, DA Qingli. Linear goal programming method for priorities of interval hybrid [J]. *Systems Engineering and Electronics*, 2008, 30(6): 1079 – 1085.)
- [10] 韩亮, 王守相, 赵歌. 基于区间TOPSIS与遗传算法混合的分布式电源优化配置[J]. 电力系统自动化, 2013, 37(24): 37 – 42.
(HAN Liang, WANG Shouxiang, ZHAO Ge. Optimal planning of distributed generators based on combination of interval TOPSIS method and genetic algorithm [J]. *Automation of Electric Power Systems*, 2013, 37(24): 37 – 42.)
- [11] 胡国强, 贺仁睦. 基于交互式多目标决策方法的水火电力系统日有功负荷优化分配[J]. 电网技术, 2007, 31(18): 37 – 42.
(HU Guoqiang, HE Renmu. Optimal daily active power load dispatching of hydrothermal power system based on interactive multi-objective decision-making method [J]. *Power System Technology*, 2007, 31(18): 37 – 42.)
- [12] 赵新泉, 彭勇行. 管理决策分析[M]. 北京: 科学出版社, 2008.
(ZHAO Xinquan, PENG Yongxing. *Management and Decision Analysis* [M]. Beijing: Science Press, 2008.)
- [13] 高宗和, 滕贤亮, 涂力群. 互联网AGC分层控制与CPS控制策略[J]. 电力系统自动化, 2004, 28(1): 78 – 81.
(GAO Zonghe, TENG Xianliang, TU Liqun. Hierarchical AGC mode and CPS control strategy for interconnected power systems [J]. *Automation of Electric Power Systems*, 2004, 28(1): 78 – 81.)
- [14] 张汝波, 顾国昌, 刘德照, 等. 强化学习理论、算法及应用[J]. 控制理论与应用, 2000, 17(5): 637 – 642.
(ZHANG Rubo, GU Guochang, LIU Dezhao, et al. Reinforcement learning theory, algorithms and its application [J]. *Control Theory & Applications*, 2000, 17(5): 637 – 642.)
- [15] 张汝波. 强化学习理论及应用[M]. 哈尔滨: 哈尔滨工程大学出版社, 2001.
(ZHANG Rubo. *Reinforcement Learning Theory and Its Application* [M]. Harbin: Harbin Engineering University Press, 2001.)
- [16] YU T, ZHOU B, CHAN K W, et al. Stochastic optimal relaxed automatic generation control in non-Markov environment based on multi-step $Q(\lambda)$ learning [J]. *IEEE Transactions on Power Systems*, 2011, 26(3): 1272 – 1282.
- [17] SUTTON R S, BARTO A G. *An Introduction to Reinforcement Learning* [M]. Cambridge: The MIT Press, 1998.
- [18] YU T, LIU J, CHAN K W, et al. Distributed multi-step $Q(\lambda)$ learning for optimal power flow of large-scale power grids [J]. *International Journal of Electrical Power and Energy Systems*, 2012, 42(1): 614 – 620.
- [19] WEISSGERBER J. Dynamic models for steam and hydro turbines in power system studies [J]. *IEEE Transactions on Power Apparatus and Systems*, 1973, 92(6): 1904 – 1951.
- [20] EKANAYAKE J B, JENKINS N, STRBAC G. Frequency response from wind turbines [J]. *Wind Engineering*, 2008, 32(6): 573 – 586.
- [21] 梁双, 胡学浩, 张东霞, 等. 考虑风速变化特性的风电容量可信度评估方法[J]. 中国电机工程学报, 2013, 33(10): 18 – 26.
(LIANG Shuang, HU Xuehao, ZHANG Dongxia, et al. Capacity credit evaluation of wind generation considering wind speed variation characteristics [J]. *Proceeding of CSEE*, 2013, 33(10): 18 – 26.)

作者简介:

张孝顺 (1990–), 男, 博士研究生, 目前研究方向为电力系统优化运行与控制, E-mail: xszhang1990@sina.cn;

余涛 (1974–), 男, 教授, 博士生导师, 目前研究方向为复杂电力系统的非线性控制理论和仿真研究, E-mail: taoyu1@scut.edu.cn.