

基于多步回溯 $Q(\lambda)$ 的 PSS 最优控制方法的研究

余涛, 甄卫国

(华南理工大学电力学院, 广东 广州 510640)

摘要: 电力系统稳定器 (PSS) 是用来产生能抑制低频电力系统振荡的励磁系统辅助控制信号, 具备自学习和参数在线整定能力是未来智能电网 PSS 控制器的一个发展趋势。提出一种基于多步回溯 $Q(\lambda)$ 学习的新颖电力系统稳定器设计方法。利用多步回溯 $Q(\lambda)$ 控制器代替整个传统 PSS 作为励磁附加控制, 与传统 PSS 和 Q 学习控制器进行比较。仿真研究显示, 引入基于多步回溯 $Q(\lambda)$ 学习的 PSS 控制后显著增强了整个系统的鲁棒性, 有效提高了系统抑制低频电力系统振荡的能力, 较好地解决了 Q 学习控制器收敛速度慢的问题。

关键词: 电力系统稳定器 (PSS); 马尔可夫策略 (MDP); 强化学习; Q 学习; 多步回溯 $Q(\lambda)$ 学习

Optimal control method of PSS based on multi-step backtrack $Q(\lambda)$ learning

YU Tao, ZHEN Wei-guo

(College of Electric Power, South China University of Technology, Guangzhou 510640, China)

Abstract: Power system stabilizers (PSS) are used to generate supplementary control signals for the excitation system in order to damp the low frequency power system oscillations. With the development of smart grids, the multiply PSS controllers with the abilities of self-learning and self-tuning become the attractive trend. A novel control method of power system stabilizer (PSS) based on multi-step backtrack $Q(\lambda)$ learning is proposed in this paper. The multi-step backtrack $Q(\lambda)$ controller is used to replace the conventional PSS to generate supplementary control signals for the excitation system, and is compared with the conventional PSS and Q -learning controller. Results show that the $Q(\lambda)$ controller strengthens the robustness of the power system and enhances the ability of damping the low frequency power system oscillations. Besides, it can solve the problem of the slow convergence rate of Q -learning controller.

This work is supported by National Natural Science Foundation of China (No. 50807016) and Natural Science Foundation of Guangdong Province (No. 9151064101000049).

Key words: power system stabilizer (PSS); Markov decision process (MDP); reinforcement learning; Q -learning; multi-step backtrack $Q(\lambda)$ learning

中图分类号: TM712 文献标识码: A 文章编号: 1674-3415(2011)03-0018-06

0 引言

强化学习^[1] (Reinforcement Learning, RL), 又称再励学习、评价学习, 既可看作是人工智能领域中一种重要的机器学习方法, 也被认为是属于马尔可夫决策过程^[2] (Markov Decision Process, MDP) 和动态优化方法的一个独立分支。

RL 方法基本思想是一个学习系统与其环境的反复交互作用, 仅从所在环境中自身经历产生反馈的信息来学会执行一个任务并不断地对系统性能进

行自我改进。RL 具有强大的在线自学习能力, 易于进行动态并行计算, 且算法对研究对象的数学模型不敏感。电力系统是一个复杂的动态大系统, 由于电力系统具有复杂的非线性和不确定性, 传统方法是采用系统辨识或工程经验的途径, 其应用效果往往受到实际环境的限制, 而 RL 可以把控制系统的性能指标要求转化为一种评价指标, 智能体 (Agent) 通过奖励和惩罚进行学习。笔者在文献[3]中综合介绍了 RL 理论在电力系统中的应用研究现状, 并探讨了该课题在电力系统运行控制中的巨大潜力。

目前广泛使用的传统电力系统稳定器 (CPSS) 的设计是建立在频域的相位补偿理论^[4-5], 作为一种

基金项目: 国家自然科学基金项目 (50807016); 广东省自然科学基金项目 (9151064101000049); 中央高校基本科研业务费专项资金资助

转速或功率的相位滞后补偿。CPSS的参数是基于线性模型的电力系统。为了使到CPSS在很宽的范围内能提供良好的阻尼作用,它的参数需要被微调来适应两种类型的振荡。因为电力系统是高度非线性系统,它的结构和参数都是随着时间的变化而变化的。基于线性化模型的电力系统的CPSS是不能保证它在实际运行环境中的控制性能。因此,研究适应电力系统复杂运行状态变化的自适应PSS一直是此领域研究的重点。

为了改善CPSS的性能,许多智能化优化技术已经提出,例如:智能优化方法^[6-8](如:模拟退火算法、遗传算法),模糊逻辑^[9-10]、神经网络^[11]和许多其他的非线性控制技术。智能优化算法是在离线的环境下通过优化一个特征值来优化CPSS的参数。因为这些方法是建立在线性模型上,而且它们的参数是不能在线改变的。因此,它们在实际运行中缺乏令人满意的性能。众所周知,基于规则的模糊逻辑控制是很难取得和调整它的参数的,特别是在线的情况下。

Q学习算法是无需环境模型的一种目前国际上最为广泛应用的RL算法,它可以被认为是异步动态规划的一种方法。它提供Agent在马尔科夫环境中,利用经历的动作序列执行最优动作的一种学习能力。笔者在文献[12]中详细分析了CPSS的局限性和现代控制算法的非在线性,并在文中提出了基于Q学习算法的新型电力系统稳定器设计方法,提高了系统抑制低频振荡的能力。但基于单步Q学习算法的PSS需要较长的时间来搜索到最优动作。而本文引入具备多步回溯能力的Q(λ)学习,该在线回溯算法显式地利用资格迹对将来多步决策的在线强化信息进行高效地回溯操作,能够有效地解决上述问题,提高算法收敛速度。

1 多步回溯Q(λ)学习算法

多步Q(λ)学习(Multi-step Q(λ)learning)^[13]是基于离散马尔可夫决策过程的经典Q学习^[14]结合了TD(λ)算法^[15]多步回报的思想。Q(λ)值函数的回溯更新规则利用资格迹来获取算法行为的频度和渐新度两种启发信息,从而考虑了未来控制决策的影响。资格迹^[7]用于解决延时强化学习的时间信度分配问题,第k步迭代时刻的矩阵形式即 $e_k(s, a)$,是对过去所访问状态与动作信息的一种临时记录。对任何状态-动作对而言,资格迹都将以时效性按指数 λ^k 衰减,一旦执行非贪婪探索动作时,迹则可以复位设置为零。资格迹更新公式定义为:

$$I_{xy} = \begin{cases} 1 & x = y \\ 0 & \text{其他} \end{cases} \quad (1)$$

$$e_k(s, a) = I_{ss_k} \cdot I_{aa_k} + \begin{cases} \gamma \lambda e_{k-1}(s, a) & Q_{k-1}(s_k, a_k) = \max_a Q_{k-1}(s_k, a) \\ 0 & \text{其他} \end{cases} \quad (2)$$

式中: I_{xy} 是迹特征函数; $0 < \gamma < 1$, 为折扣因子; λ 为迹衰退系数。

Q(λ)充分考虑了随机过程的时间回溯因素, λ-回报算法将值函数与资格迹相互融合,如图1,其“后向估计”将时间信度和值函数误差对所经历的“状态流”进行合理回溯分配,对于非马氏环境具有很好的适应性^[16]。

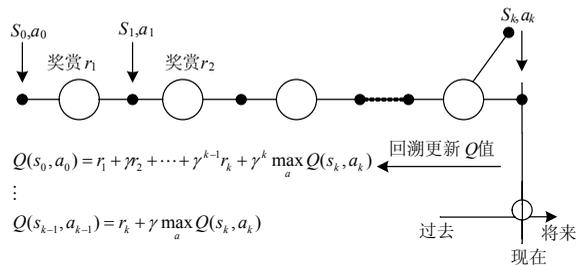


图1 Q(λ)方法“后向估计”回溯学习更新图

Fig.1 Q(λ)backtrack learning of backward estimation

资格迹λ-回报算法的“后向估计”机理提供了一个逼近最优值函数Q*的渐进机制,而这类对所有状态-动作对Q值的高效持续更新是以提高算法复杂度和增加计算量为代价的。设Q_k代表Q*估计值的第k次迭代值, Q(λ)学习迭代更新公式为:

$$\delta_k = R(s_k, s_{k+1}, a_k) + \gamma \max_{a'} Q_k(s_{k+1}, a') - Q_k(s_k, a_k) \quad (3)$$

$$Q_{k+1}(s, a) = Q_k(s, a) + \alpha \delta_k e_k(s, a) \quad (4)$$

式中: $0 < \alpha < 1$, 称为学习因子; $R(s_k, s_{k+1}, a_k)$ 是第k步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; $Q(s, a)$ 代表s状态下执行动作a的Q值函数,其实现方式均采用lookup查表法。文中多步Q(λ)学习中动作选择策略则通常采用由概率矢量法派生的追踪算法或boltzmann分布法^[14]。

2 基于Q(λ)学习的PSS最优控制原理

2.1 励磁附加控制结构

一旦系统出现低频振荡,可通过减少线路输送容量,改用常规励磁调节器等来处理,但是这两者都不利于经济性和大扰动下的暂态稳定。而在励磁系统中引入附加控制信号解决了以上的问题。传统PSS以 $\Delta \omega$ 、 ΔP_e 或两者的组合作为输入信号,通过放大环节、复位环节、相位补偿环节、限幅环节后作为励磁附加信号。传统PSS各参数的整定是根据

经典控制理论在某一特定的振荡模式下得出，无法保证在各种扰动下都能较好地抑制低频振荡。如图 2 所示，本文基于多步 $Q(\lambda)$ 学习算法，以 $Q(\lambda)$ 控制器代替整个的传统 PSS，其输出作为励磁附加控制信号。在不同扰动下， $Q(\lambda)$ 控制器通过试错学习在系统不同状态下，搜索出最佳的附加控制信号，而不需类似传统 PSS 那样设定各参数值。 $Q(\lambda)$ 控制器具有更强的适应性。

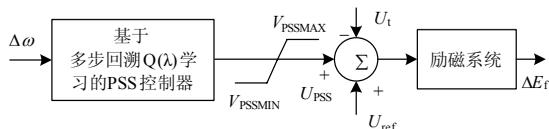


图 2 基于多步回溯 $Q(\lambda)$ 学习的 PSS 传递函数框图

Fig.2 Structure of PSS transfer function based on multi-step $Q(\lambda)$ learning

2.2 多步回溯 $Q(\lambda)$ 学习的优化控制算法设计

励磁系统的附加信号主要用来抑制低频振荡，为了防止它在某些特殊情况下，造成发电机电压过大的变化，需要限制该附加信号，使其对定子电压幅值的改变在某个限定范围内。本文取附加控制信号的范围为 $-0.15 \sim 0.15$ 之间（归一化后的标么值），通过对该范围进行离散化，得出该优化控制算法的动作集为 $[-0.15:0.01:0.15]$ ，则动作集被量化为 31 个动作。

为了能反映系统受低频振荡的影响程度，本文选取 Δw 作为状态值。对 Δw 状态空间进行分区，仿真比较显示，将 Δw 划分为 $(+\infty, 0.002)$, $(0.002, 0.001)$, $[0.001, 0]$, $(0, -0.002)$, $(-0.002, -0.004)$, $[-0.004, -\infty)$ 较为合适，则一维输入空间被量化为 6 个不同状态。

电力系统稳定器的目的就是抑制系统中有功功率、频率等量的振荡，使其保持正常值。通常，对于电力系统稳定器的控制目标可描述为使系统输出误差最小，误差为有功功率、转子角频率以及发电机端电压与它们各自正常值的相对误差绝对值之和。基于多步 $Q(\lambda)$ 学习的控制系统通过试错与环境进行交互式学习，从长期的观点构造控制策略，以期从环境获得的长期积累奖励值最大，而将系统输出误差取负值作为多步 $Q(\lambda)$ 学习的奖励值，则长期以来，系统输出误差越小，它的长期积累奖励值最大，这与电力系统稳定器的控制目标相符，因此，将系统输出误差取负值作为多步 $Q(\lambda)$ 学习的奖励函数很恰当合理。多步 $Q(\lambda)$ 学习的奖励函数 Reward 表示为：

$$R_i = -\int_t^{t+0.1} (K_p |\Delta P| + K_w |\Delta w| + K_v |\Delta V|) dt \quad (5)$$

式中： R_i 为每次迭代步所得到的奖励值； ΔP 为机械功率和电磁功率的差值； Δw 为发电机转速偏差； ΔV 为发电机机端电压偏差； K_p 、 K_w 、 K_v 为权值系数。

多步回溯 $Q(\lambda)$ 学习算法中的 γ 、 λ 、 α 分别为折扣因子、迹衰退系数和学习因子。 α 指明了要给改善的更新部分多少信任度，较大的 α 值会加快学习算法的收敛速度，而较小的 α 值能保证控制器的搜索空间，从而提高学习收敛的稳定性。它们的取值影响着算法的收敛性。通过仿真比较，本文取值为 0.9 、 $\lambda=0.9$ 、 $\alpha=0.5$ 。

在确定了控制动作集、状态集和奖励函数后，即可进行多步 $Q(\lambda)$ 学习控制器在线自学习和动态优化，其步骤如下：

对所有 (s, a) ，初始化各参数及当前状态 s_0 ，令 $k=0$ ；

Repeat forever

1) 由动作概率分布法选择并执行调度动作 a_k ；

2) 观察下一时刻的状态 s_{k+1} ，即 CPS1/CPS2 滚动指标；

3) 由式 (5) 得到一个奖励信号 $R(k)$ ；

4) 根据式 (3) 计算值函数误差 δ ；

5) 更新资格迹，令 $e(s_k, a_k) \leftarrow e(s_k, a_k) + 1$ ；

6) 式 (4) 更新 Q 值函数表格；

7) If a_{k+1} 为贪婪动作 Then $e(s, a) \leftarrow \gamma \lambda e(s, a)$

Else $e(s, a) \leftarrow 0$ ；

8) $k=k+1$ ，返回步骤 1)；

End

3 仿真算例研究

为评价多步回溯 $Q(\lambda)$ 控制器的鲁棒性和收敛性，本文以典型的四机两区域互联系统电力系统稳定器控制模型作为研究对象（如图 3 所示），对基于多步回溯 $Q(\lambda)$ 学习的控制器、基于 Q 学习的控制器^[12]、传统 PSS 的性能进行仿真比较。这些比较是

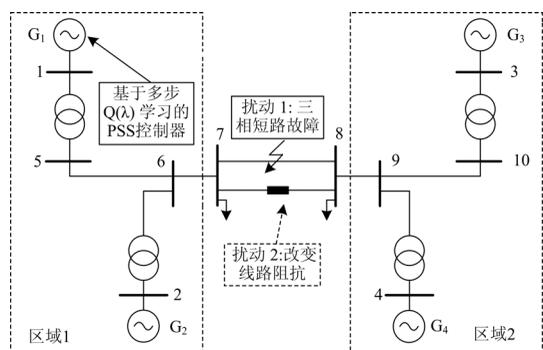


图 3 Kundur 四机两区互联系统框图

Fig.3 Kundur's four-machine two-area test system

在各种不同类型的操作条件和干扰下进行的。这些扰动是：在区域联络线路上发生三相短路，改变输电线路的阻抗。故障位置和类型如图 3 所示。

图 4 给出了传统 PSS (CPSS) [17] 的传递函数框图，该 CPSS 由低通/高通滤波器、放大环节、超前-滞后环节和限幅环节构成，CPSS 中各参数设置值见表 1。

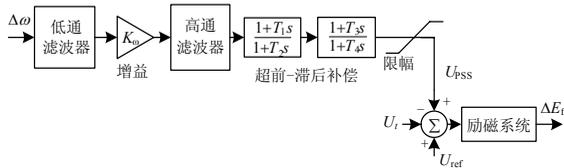


图 4 传统 PSS 传递函数框图

Fig.4 Transfer function diagram of conventional PSS

表1 传统PSS参数值

Tab.1 Parameters of conventional PSS

K_w	T_1	T_2	T_3	T_4
5	50e-3 s	20e-3 s	3 s	5.4 s

3.1 $Q(\lambda)$ 与 Q 学习算法在两种扰动下的收敛性比较

$Q(\lambda)$ 学习控制器与 Q 学习控制器在最优策略学习的初始阶段动作的选取会有很强的随机性，这样会造成系统的不稳定，甚至使系统中各发电机的功角差超过 180° 而造成失步。因此，算法在投入实际运行之前要经历离线的预学习过程。经过仿真测试，设置式 (5) 中各权值系数后，表 2 给出了 $Q(\lambda)$ 学习控制器和 Q 学习控制器的平均收敛时间。图 5 和图 6 给出了 $Q(\lambda)$ 学习控制器和 Q 学习控制器的附加控制信号输出值在三相短路扰动下的预学习收敛典型曲线图。

由表 2、图 5 和图 6 可见，在两种扰动下， $Q(\lambda)$ 控制器的收敛时间都优于 Q 学习控制器。 $Q(\lambda)$ 控制器在鲁棒性和收敛性上都优于 Q 学习控制器。

表2 不同故障情况下算法收敛性比较

Tab.2 Convergence time of $Q(\lambda)$ leaning algorithm compared with Q-learning

算法	扰动	平均收敛时间/s
$Q(\lambda)$ 学习	区域联络线路发生三相短路故障	540
Q 学习		2 020
$Q(\lambda)$ 学习	改变线路阻抗值	1 160
Q 学习		1 640

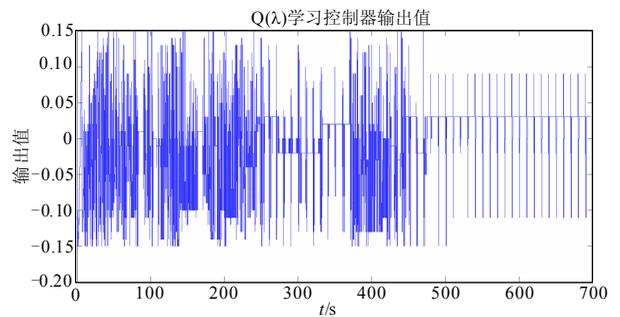


图 5 $Q(\lambda)$ 学习控制器输出值

Fig.5 Output of $Q(\lambda)$ learning controller

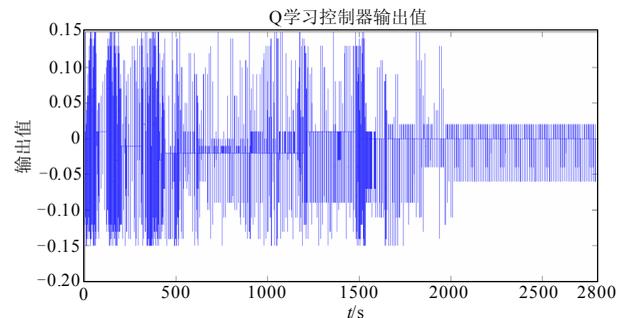


图 6 Q 学习控制器输出值

Fig.6 Output of Q-learning controller

3.2 权值系数 K_p 、 K_w 、 K_v 的影响

式 (5) 中奖励函数的选取引入了线性最优二次型调节器 (LQR) 的思想，奖励函数为对象状态和控制输入的二次型函数。权值系数 K_p 、 K_w 、 K_v 的选值对收敛结果影响较大，表 3 给出了 $Q(\lambda)$ 学习过程中所选取的权值，并比较了各权值对收敛结果的影响。

表3 权值系数 K_p 、 K_w 、 K_v

Tab.3 Value of K_p , and K_w , and K_v

序号	K_p	K_w	K_v
情况 1	1	50	10
情况 2	1	50	20
情况 3	1	50	30

表 3 中 K_v 选取了不同的值，这表示了对电压偏差值进行了不同程度的考虑， K_v 值越大，表示越对电压偏差的关注程度越大，反之越小就是关注程度越小。图 7 和图 8 分别给出了在线路发生短路故障的情况下，不同的权值取值下 $Q(\lambda)$ 学习控制器的控制效果曲线图。由图中可知，取不同的权值会影响 $Q(\lambda)$ 学习控制器的控制效果。在预学习中，需对各权值进行优化，以得到最优的控制效果。比较图 7 和图 8 中各曲线，情况 1 所得到的控制效果优于另两种。本文选取情况 1 的权值系数所得的优化结

果与 Q 学习控制器、传统 PSS 进行比较。

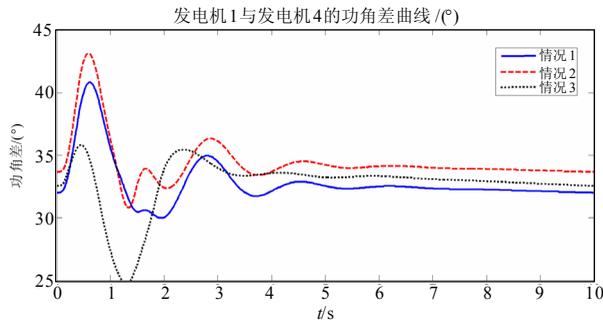


图 7 发电机 1 与发电机 4 的功角差曲线
Fig.7 Power angle difference of G1 with G4

性能最优。

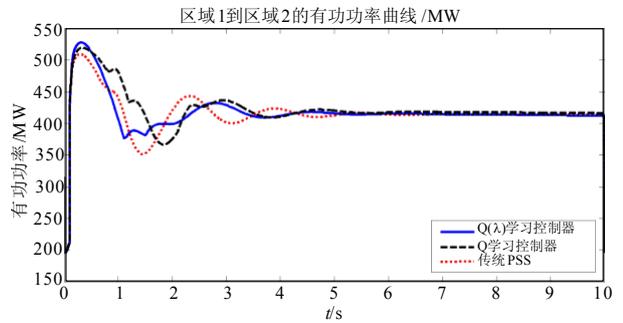


图 10 区域 1 到区域 2 的有功功率值
Fig.10 Active power from area 1 to area 2

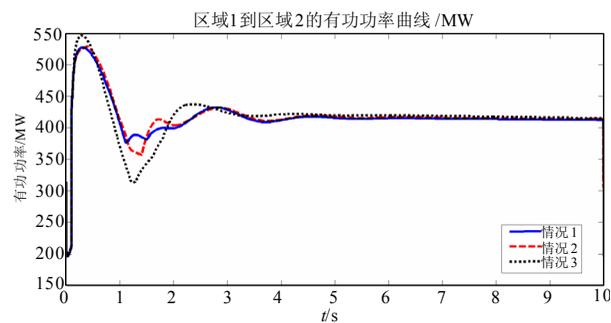


图 8 区域 1 到区域 2 的有功功率值
Fig.8 Active power from area 1 to area 2

3.4 改变线路阻抗值

为了测试所提出的 $Q(\lambda)$ 控制器对电网参数变化的鲁棒性，为本文通过在一回区域联络线路上设置断路器来模拟输电线路阻抗的突变，如图 3 所示。通过断路器的开断，输电线路阻抗在开断前后相差 2 倍。

令系统处于稳态后 0 s 时阻抗变为初始值的 2 倍，10 s 时阻抗恢复到初始值。图 11 和图 12 分别给出了功角差曲线和区域联络线路传输的有功功率曲线。由图 11 可知，Q 学习控制器的控制效果要优于传统 PSS。而 Q 学习控制器与 $Q(\lambda)$ 控制器的控制效果相近。由图 12 可知， $Q(\lambda)$ 控制器对于区域交换功率的抑制优于其他两个控制器。

3.3 区域联络线路上发生三相短路

如图 3 所示，在区域联络线路中点处发生瞬时三相短路故障，短路持续时间为 0.1 s，0.1 s 后短路故障消失。以 #4 发电机作为参考发电机，图 9 给出的是发电机 1 与发电机 4 的功角差曲线。分析图中曲线可知，三种控制器均能有效抑制系统的低频振荡，其中 $Q(\lambda)$ 学习控制器具有最优的控制性能，其振荡调节时间和超调量均优于 Q 学习控制器和传统 PSS。

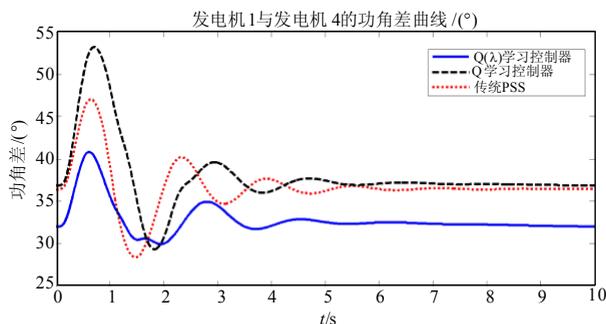


图 9 发电机 1 与发电机 4 的功角差曲线
Fig.9 Power angle difference of G1 with G4

图 10 给出的是区域 1 向区域 2 传输的有功功率曲线。比较图中曲线可知， $Q(\lambda)$ 学习控制器控制下的区域交换有功功率值振荡幅度最小，传统 PSS 与 Q 学习控制器的性能相近， $Q(\lambda)$ 学习控制器的控制

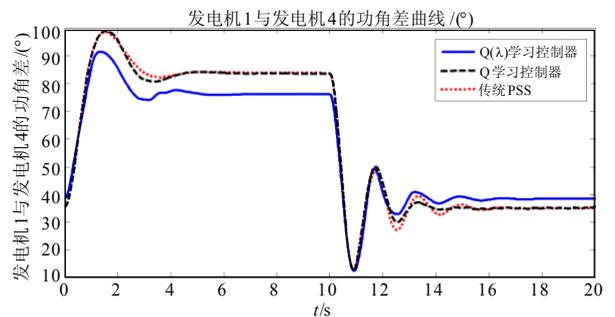


图 11 发电机 1 与发电机 4 的功角差曲线
Fig.11 Power angle difference of G1 with G4

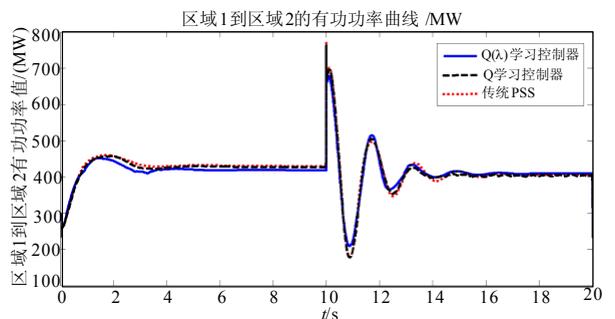


图 12 区域 1 到区域 2 的有功功率值
Fig.12 Active power from area 1 to area 2

综合以上仿真结果,三种控制器均能有效抑制系统的低频振荡,而 $Q(\lambda)$ 学习控制器在各种故障下具有更好的鲁棒性。

同时,该扰动也表明了在线扰动中 $Q(\lambda)$ 和 Q 学习算法具有在线学习、更新动作策略的能力,而 $Q(\lambda)$ 学习可以利用资格迹进行多步回溯,更好地利用经验,因此表现出更突出的抑制效果。

4 结语

尽管传统 PSS 在整定良好的情况下也具有优良的控制性能,但其最大的问题是缺乏在线自学习和自整定能力,在电网渐变的过程中,缺乏参数自整定的一类 PSS 是无法满足未来智能电网的高度智能化和适应性要求。本文介绍了一种基于多步回溯 $Q(\lambda)$ 学习属于一类半马氏决策过程的新型方法,据作者知识范围,多步回溯 $Q(\lambda)$ 学习尚未在国内外电力控制领域有文献正式发表。通过对仿真结果进行分析,可以得出以下结论:

(1) $Q(\lambda)$ 算法是一种模型无关的算法,它不像监督学习方法,它不依赖于某一种特定的模型进行学习,不像传统 PSS 那样需要整定参数。仿真结果表明, $Q(\lambda)$ 算法在系统受扰动后减少振荡的超调量和调节时间,具有更优的抑制效果。

(2) $Q(\lambda)$ 算法具有在线回溯算法显式地利用资格迹对将来多步决策的在线强化信息进行高效地回溯操作,仿真结果表明 $Q(\lambda)$ 学习算法收敛速度要优于 Q 学习算法。

(3) $Q(\lambda)$ 算法具有在线学习、更新动作策略的能力,能适用连续扰动后环境的变化。通过预学习后,然后在线继续进行交互式学习,以此在线适应工况的变化。

在后续的工作中,对于多机复杂系统,分布式强化学习的引入可有助于多机 PSS 的优化,将各 PSS 视为学习的学习单元,每个单元独立地执行部分或者全部的任务,达到整个系统意义上的学习目标。另外,可以结合其他离线优化方法,利用经验回放、输入空间量化等方法提高学习速度。

参考文献

[1] Sutton R S, Barto A G. Reinforcement learning: an introduction[J]. Cambridge: MIT Press, 1998.

[2] Mine H, Osaki S. Markov decision processes[M]. New York: Eisevier, 1970.

[3] 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[J]. 电力系统保护与控制, 2009, 37(14): 122-128.

YU Tao, ZHOU Bin, ZHEN Wei-guo. Application and

development of reinforcement learning theory in power systems[J]. Power System Protection and Control, 2009, 37(14): 122-128.

- [4] 刘取. 电力系统稳定性及发电机励磁控制[M]. 北京: 中国电力出版社, 2007.
- [5] 倪以信,陈寿孙,张宝霖. 动态电力系统的理论和分析[M]. 北京: 清华大学出版社, 2008: 264-291.
- [6] Abido M A. Robust design of multi-machine power system stabilizers using simulated annealing[J]. IEEE Transaction on Energy Conversion, 2000: 297-304.
- [7] 文劲宇,王国兴,程时杰,等. 用遗传算法设计模糊式电力系统稳定器[J]. 电力系统自动化, 1999, 23(4): 22-24.
- WEN Jin-yu, WANG Guo-xing, CHENG Shi-jie, et al. A design of fuzzy power system stabilizer using genetic algorithm[J]. Automation of Electric Power Systems, 1999, 23(4): 22-24.
- [8] 刘杨名,严正,胥国毅. 免疫遗传算法在多机电力系统 PSS 系统优化中的运用[J]. 电力系统保护与控制, 2007, 35(7): 19-23, 45.
- LIU Yang-ming, YAN Zheng, XU Guo-yi. Parameter optimization of power system stabilizers using immune genetic algorithm in multi-machine system[J]. Power System Protection and Control, 2007, 35(7): 19-23, 45.
- [9] 赵书强,丁峰,侯子利,等. 自寻优模糊电力系统稳定器的设计[J]. 电工技术学报, 2004, 19(3): 94-98.
- ZHAO Shu-qiang, DING Feng, HOU Zi-li, et al. Design of self-optimizing fuzzy power system stabilizer[J]. Transactions of China Electrotechnical Society, 2004, 19(3): 94-98.
- [10] 王明东,刘宪林,于继来. 基于改进遗传算法的发电机调速器侧模糊 PSS 优化设计[J]. 继电器, 2006, 34(13): 25-28.
- WANG Ming-dong, LIU Xian-lin, YU Ji-lai. Optimization design of fuzzy PSS for governor of generator based on an improved GA[J]. Relay, 2006, 34(13): 25-28.
- [11] 梁有伟,胡志坚,陈允平. 基于神经网络逆系统的电力系统稳定器的研究[J]. 电工技术学报, 2004, 19(5): 61-65, 14.
- LIANG You-wei, HU Zhi-jian, CHEN Yun-ping. A power system stabilizer based on neural network inverse system[J]. Transactions of China Electrotechnical Society, 2004, 19(5): 61-65, 14.
- [12] YU Tao, ZHEN Wei-guo. A reinforcement learning approach to power system stabilizer[C]. //Power & Energy Society General Meeting, PES '09. IEEE. 2009: 1-5.
- [13] Jing Peng, Williams R J. Incremental multi-step Q-learning [J]. Machine Learning, 1996, 22: 283-290.

(下转第 28 页 continued on page 28)

障或出口处反方向故障时, 该线路纵联零序方向保护不受零序互感的影响, 能够正确动作; 对于其相邻线路来说电气联系越紧密, 零序方向元件误动的可能性越小; 反之, 误动可能性越大。因此, 应综合考虑相邻互感线路纵联零序方向保护的配置^[6-10], 增加必要的防范措施, 确保电网稳定运行。

参考文献

[1] 李一泉, 焦邵麟, 张弛, 等. 平行线路纵联零序方向保护安全性分析[J]. 电力系统自动化, 2008, 32(6): 104-107.
LI Yi-quan, JIAO Shao-lin, ZHANG Chi, et al. Safety analysis on pilot zero sequence direction relay applied in parallel lines[J]. Automation of Electric Power Systems, 2008, 32(6): 104-107.

[2] 郭润生, 何彩红, 鄧建杰. 相邻线路零序互感对线路零序纵联方向保护的影响[J]. 继电器, 2004, 32(9): 71-73.
GUO Run-sheng, HE Cai-hong, ZHI Jian-jie. Influence of zero-sequence mutual inductance to pilot protection in parallel lines[J]. Relay, 2004, 32(9): 71-73.

[3] 康小宁, 梁振锋, 索南加乐. 相邻线路零序互感对平行双回线电流平衡保护的影响及改进措施[J]. 继电器, 2005, 33(20): 6-9.
KANG Xiao-ning, LIANG Zhen-feng, SUONAN Jia-le. Influence of zero-sequence mutual inductance to transverse differential current protection and approaches to the improved measures[J]. Relay, 2005, 33(20): 6-9.

[4] 曾耿晖, 黄明辉, 刘之尧, 等. 同杆线路纵联零序保护误动分析及措施[J]. 电力系统自动化, 2006, 30(20): 103-107.
ZENG Geng-hui, HUANG Ming-hui, LIU Zhi-yao, et al. Analysis and countermeasures of misoperation of zero sequence pilot protection on circuit lines of same pole[J]. Automation of Electric Power Systems, 2006, 30(20): 103-107.

[5] 胡宁, 郑罡, 胡志坚. 基于积分方程的互感线路参数带电测量研究[J]. 继电器, 2005, 33(16): 22-25.
HU Ning, ZHENG Gang, HU Zhi-jian. Study of live

line measurement of parameters of transmission lines with mutual inductance based on integral equations[J]. Relay, 2005, 33(16): 22-25.

[6] 李钢, 冯辰虎, 孙集伟, 等. 平行运行线路互感对纵联零序方向保护的影响[J]. 华北电力技术, 2007, 12: 1-4.
LI Gang, FENG Chen-hu, SUN Ji-wei, et al. Effects of mutual inductance of parallel transmission lines on longitudinal zero-sequence protection[J]. North China Electric Power, 2007, 12: 1-4.

[7] 朱景富. 零序互感对线路接地距离保护的影响分析[J]. 电力系统保护与控制, 2009, 37(9): 113-115.
ZHU Jing-fu. Zero-sequence mutual inductance on the ground distance elements impact analysis[J]. Power System Protection and Control, 2009, 37(9): 113-115.

[8] 高中德, 舒治淮, 王德林. 国家电网公司继电保护培训教材[J]. 北京: 中国电力出版社, 2009.

[9] 樊占峰, 叶东印, 李瑞生, 等. 平行线弱电强磁模型下零序方向元件改进[J]. 电力系统自动化, 2008, 32(17): 100-103.
FAN Zhan-feng, YE Dong-yin, LI Rui-sheng, et al. Improvement of zero-sequence directional relay for the parallel line with magnetically strong and electrically weak connection[J]. Automation of Electric Power Systems, 2008, 32(17): 100-103.

[10] 丁晓兵, 赵曼勇, 徐振宇. 接地故障零序方向元件拒动保护改进方案[J]. 电力系统自动化, 2006, 30(9): 88-90.
DING Xiao-bing, ZHAO Man-yong, XU Zhen-yu. Improvement on zero-sequence current protection when directional element fails to operate during earth faults[J]. Automation of Electric Power Systems, 2006, 30(9): 88-90.

收稿日期: 2010-03-01; 修回日期: 2010-05-06

作者简介:

吴麟琳(1982-), 女, 博士研究生, 研究方向为电力系统继电保护; E-mail: aileenwu500@163.com

黄少锋(1958-), 男, 教授, 博士生导师, 研究方向为电力系统继电保护。

(上接第 23 页 continued from page 23)

[14] Watkins J C H, Dayan Peter. Q-learning [J]. Machine Learning, 1992, 8: 279-292.

[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: 69-100.

[16] Sutton R S, Barto A G. Reinforcement learning: an introduction [M]. Cambridge: MIT Press, 1998: 87-160.

[17] Kundur P. Power system stability and control[M]. McGraw-Hill, 1994.

收稿日期: 2010-02-25; 修回日期: 2010-06-10

作者简介:

余涛(1974-), 男, 副教授, 博士, 主要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

甄卫国(1985-), 男, 硕士研究生, 主要研究方向为电力系统稳定运行与控制。E-mail: zwg1116@126.com