

基于强化学习算法的自适应 直流附加阻尼控制器

郭 力, 张 尧, 胡金磊

(华南理工大学 电力学院, 广东 广州 510640)

摘要: 提出了基于强化学习算法的直流附加阻尼控制器。控制器主体采用模糊神经网络, 利用由系统性能指标生成的强化信号在线训练控制器参数。与传统的模糊控制器相比, 由于该控制器采用自适应启发式评价算法, 将系统输出性能指标转化为强化信号反馈给控制器, 使其能够在线修改控制器参数, 因此有效地克服了传统阻尼控制器的设计对系统精确数学模型的依赖。仿真结果表明, 与传统的阻尼控制器相比, 基于强化学习算法的直流附加阻尼控制器能够有效地抑制区域间的功率振荡, 提高交直流系统的动态稳定性, 并且对多种运行方式具有一定的鲁棒性。

关键词: 强化学习; 低频振荡; 联想搜索网络; 传统附加阻尼控制器

中图分类号: TM 712; TP 18

文献标识码: A

文章编号: 1006-6047(2007)10-0087-04

0 引言

由于缺乏足够的机电振荡阻尼, 尤其是发电机所提供的电磁阻尼转矩不足, 电力系统区域间的功率会发生相应的振荡。由于其振荡频率很低, 故称为低频振荡^[1-3]。

近年来, 具有快速响应特性的灵活交流输电设备 (FACTS) 和高压直流输电 (HVDC) 成为电力系统动态调节的主要手段之一, 文献[4]指出, 在交直流互连系统中的直流调制中配置与电力系统稳定器 (PSS) 具有相类似设计结构的附加阻尼控制器 (CSDC), 以交流侧系统频率或者以与直流并联的交流联络线的传输功率为输入信号, 能够有效地抑制对通过平行交流线路互联的区域振荡模式, 且对其它振荡模式影响不大。

模糊控制、神经网络、遗传算法等^[5]人工智能技术由于计及专家经验、适应大规模非线性系统, 因此在电力系统中得到了广泛应用。文献[6]利用模糊神经网络设计 PSS, 克服了模糊控制和神经网络各自的缺点。文献[7]设计了一种自适应模糊 PSS, 其动态行为类似于传统的比例积分式控制器, 具有较好的鲁棒性。文献[8]利用神经网络在线修正模糊 PSS 的参数。文献[9]将 Prony 算法应用于模糊阻尼控制器的参数调节和整定上。文献[10]研究了互连两区域系统中由于低频振荡在系统动能和势能之间转换的关系, 以削减系统间的振荡能量为目的, 设计了附加于 HVDC 上的模糊阻尼控制器。文献[11]对控制器的隶属度函数进行了改进, 并考虑了励磁设备和附加控制器的协调设计。文献[12]提出了一种统一潮流控制器 (UPFC) 中附加模糊控制器的设计方法,

并利用序优化遗传算法对其参数进行优化。文献[13]采用类似方法设计了直流附加模糊阻尼控制器。

上述模糊阻尼控制器大多采用精确训练样本学习隐含在非线系统本身的规律, 以调整网络权系数。但在一些复杂的应用中, 精确的训练样本往往难以获得。强化学习作为一种学习算法, 只需要把系统的性能指标转换为评价指标, 对控制器的动作进行惩罚和奖励, 以最终获得对系统的优化控制动作^[14-15]。

这里将强化学习算法 (智能算法) 应用于直流附加阻尼控制器的设计中。给出了控制器的设计框图, 详细介绍了控制器的构成和学习算法, 并通过一个互联多机交直流系统在不同运行方式和振荡模式下对该附加控制器和 CSDC 的性能进行了比较。

1 基于强化学习算法的直流附加阻尼控制器

基于线性相位补偿原理的 CSDC 结构如图 1 所示, 通常采用与直流并联的交流联络线的有功功率作为控制器的输入, 输出信号作为采用定功率控制的整流侧参考功率的附加信号。其中, K_s 是放大环节增益; τ_w 是滤波环节的时间常数; $\tau_1 \sim \tau_4$ 是超前-滞后环节的时间常数; 直流整流侧采用传统的 PI 控制器。

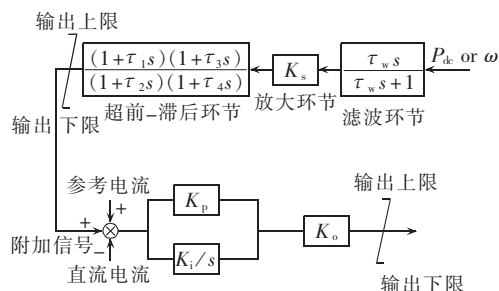


图 1 传统的附加阻尼控制器

Fig.1 Conventional supplementary damping controller

基于强化学习算法的阻尼控制器的结构可由图 2 表示,强化学习是一种试探性的评价学习过程,控制器采用启发式评价算法,由联想搜索网络(ASN)、自适应评价网络(ACN)、算法单元组成。其中 ASN 是一个 5 层的模糊神经网络,ACN 是一个前馈神经网络。

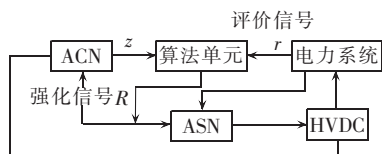


图 2 基于强化学习的自适应阻尼控制器

Fig.2 Adaptive damping controller based on RL

强化学习只需要对控制结果进行评估,并给出相应的评价(奖、罚)信号 r 以产生实际控制信号。信号 r 是由系统的运行指标转换而来,这里采用系统故障切除后的交流联络线的有功功率与额定传输的有功功率偏差作为系统的评价指标。ACN 根据当前 HVDC 整流侧控制变量来评价 HVDC 的运行情况,产生一个强化信号的预报值 z 。算法单元根据上述 2 个信号产生强化信号 R 分别给 ASN 和 ACN,从而实现了上述 2 个网络的参数在线学习。

2 控制器的设计算法

2.1 ASN 的设计结构

ASN 采用了 5 层模糊神经网络,其设计结构如图 3 所示。

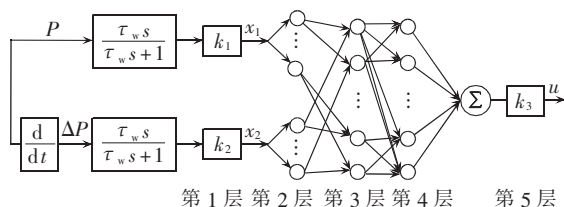


图 3 ASN 网络

Fig.3 Associative search network

第 1 层为输入层, k_1, k_2 是比例因子,其作用是将实际输入信号转化到第 2 层模糊隶属函数的定义域之内,此处采用了交流联络线路的有功功率 P 和其微分 ΔP 作为 2 个输入信号。由于测量误差或者系统状态的变化,都可能在功率偏差和加速度上产生高频分量,干扰阻尼控制器的工作,因此通常采用 2 个低通滤波器来过滤其中的高频分量^[9]。 τ_w 为滤波器时间常数。

第 2 层网络是模糊化层,每个节点代表一个模糊语言,它的作用是计算输入分量属于各个模糊语言的隶属度值。输入、输出分别定义了包含 7 个语言的模糊集合:LN、MN、SN、ZR、SP、MP、LP,分别表示负大、负中、负小、零、正小、正中、正大。输入隶属度函数采用高斯函数,计算公式为

$$A_i^j(x_i) = \exp[-(x - c_i^j)] / [2(\delta_i^j)^2] \quad (1)$$

其中, $i=1, 2, \dots, n; j=1, 2, \dots, m; A_i^j$ 代表第 i 个输入信号 x_i 隶属于第 j 个模糊语言的程度; c_i^j, δ_i^j 是高斯

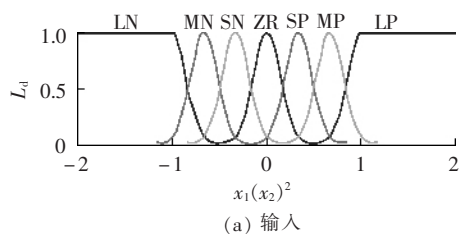
隶属度函数的参数。

输出隶属度函数采用三角函数,有 3 个参数(a, b, c),计算公式为

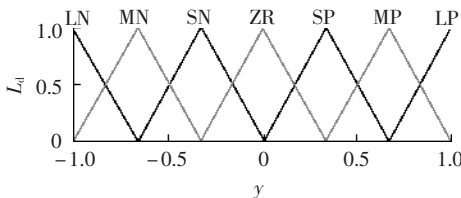
$$Y_k(y) = \begin{cases} 0 & y \leq a_k \\ \frac{y - a_k}{b - a_k} & a_k \leq y \leq b_k \\ \frac{c_k - y}{c_k - b_k} & b_k \leq y \leq c_k \\ 0 & c_k \leq y \end{cases} \quad (2)$$

其中, $k=1, 2, \dots, m_i$ 对应了输出模糊语言集合。

如图 4 所示,输入、输出隶属度函数分别被定义在 $[-1, 1]$ 之内,其中纵坐标 L_d 对应隶属度值,最大为 1。由于实际中输入信号可能很大,故将输入的 LP 和 LN 部分扩展到 2 和 -2。



(a) 输入



(b) 输出

图 4 输入、输出隶属度函数

Fig.4 Membership functions of input and output signals

第 3 层是规则层,此层输出为第 p 条规则的激活度 a_p 。

模糊规则是建立在专家经验的基础上,根据专家知识进行控制,与运行点工况联系不大,因此能够对多种情况进行识别。模糊控制规则如表 1 所示。该规则表能够较好地体现专家经验,利用专家知识进行控制,从而使得控制器具有一定的鲁棒性。

表 1 模糊控制规则表

Tab.1 Rule table of fuzzy controller

x_2	x_1						
	LP	MP	SP	ZR	SN	MN	LN
LP	LP	LP	LP	MP	MP	SP	ZR
MP	LP	LP	LP	MP	SP	ZR	SN
SP	LP	MP	MP	SP	ZR	SN	MN
ZR	MP	MP	SP	ZR	SN	MN	MN
SN	MP	SP	ZR	SN	MN	MN	LN
MN	SP	ZR	SN	MN	MN	LN	LN
LN	ZR	SN	MN	MN	LN	LN	LN

模糊规则的一般形式为

R_p^k : 如果 x_1 的隶属度值是 $A_1^p, x_2, \dots, x_n$ 隶属度值是 $A_2^p, \dots, A_n^p$, 则 y 属于第 k 个模糊语言的可信度是 w_p^k 。其中, $p=1, 2, \dots, L, L$ 是规则数,该层每个结点代表 1 条模糊规则。 w_p^k 是 $[0, 1]$ 区间的值, $w_p^k=1$ 表

示该规则绝对正确, $w_p^k=0$ 表示该规则完全不对。

第 p 条规则的激活度 a_p 定义为

$$a_p = \left(\sum_{i=1}^n A_i^p(x_i) e^{-A_i^p(x_i)} \right) / \sum_{i=1}^n e^{-A_i^p(x_i)} \quad (3)$$

当计算完每条规则的激活度后,输出变量 y 对应第 k 个模糊语言的总激活度 α'_k 可以通过第 4 层中的如下方程求解:

$$\alpha'_k = \sum_{p=1}^L (a_p w_p^k) / \sum_{p=1}^L \alpha_p \quad (4)$$

其中, $k=1,2,\dots,m_l, m_l$ 为输出变量的模糊语言集合数。

第 5 层是解模糊层,因为输出采用三角形隶属函数,则由式(2)可知 y 一般有 2 个解,取这 2 个解的平均值,即

$$y = Y_k^{-1}(\alpha'_k) = \frac{1}{2} [c_k + a_k + \alpha'_k(2b_k - c_k - a_k)] \quad (5)$$

控制输出 u 为

$$u = \sum_{k=1}^{m_l} \alpha'_k y / \sum_{k=1}^{m_l} \alpha'_k \quad (6)$$

2.2 ACN 的设计结构

ACN 单元由 3 层前馈神经网络构成。其中,输入层共有 n 个节点,分别是 $x_1, x_2, \dots, x_n$ 。 $\mathbf{x} = [x_1, x_2, \dots, x_n]$ 是被控对象的状态矢量。

隐含层共有 m 个结点,每个结点的输入和输出之间的关系是:

$$v_j^{(2)}(t+1) = g \left(\sum_{i=1}^n w_{ji}^{(12)}(t) x_i(t+1) \right) \quad (7)$$

$$g(t) = 1 / [1 + \exp(-t)] \quad (8)$$

其中, $v_j^{(2)}(t+1)$ 是 $t+1$ 时刻隐含层的输出, $w_{ji}^{(12)}(t)$ 是 t 时刻第 1 层第 i 个结点和第 2 层第 j 个结点之间的连接权系数, $j=1,2,\dots,m$ 。

输出层只有 1 个结点,它接受来自第 2 层的输出信号,输出对评价信号 r 的预报值 z ,即

$$z(t+1) = \sum_{j=1}^m w_j^{23}(t) v_j(t+1) \quad (9)$$

其中, $w_j^{23}(t)$ 是 t 时刻第 2 层第 j 个结点和输出层之间的连接权系数。

算法单元根据 ACN 输出的预报评价信号 z 和评价信号 r 来产生内部强化信号 R 。其算法如下:

$$R(t+1) = r(t+1) - z(t+1) \quad (10)$$

$$r(t+1) = \begin{cases} 0 & |e(t)| < \varepsilon \\ 1 & |e(t)| > \varepsilon \end{cases} \quad (11)$$

其中, $e(t) = P(t) - P_0(t)$ 。

$r(t+1)$ 是 $t+1$ 时刻的评价信号,它在输出满足系统要求的性能指标时为 0,输出不满足系统要求的性能指标时为 -1。

ACN 单元学习的目的是通过调整网络权值使得内部强化信号 r 最大。因此, $w_j^{(23)}(j=1,2,\dots,m)$ 采用以下学习算法:

$$w_j^{(23)}(t+1) = w_j^{(23)}(t) + \beta R(t+1) v_j^{(2)}(t, t+1) \quad (12)$$

其中, $\beta > 0$ 是学习率。

$w_{ji}^{(12)}$ 采用改进的 BP 算法进行学习:

$$w_{ji}^{(12)}(t+1) = w_{ji}^{(12)}(t) + a R(t+1) v_j^{(2)}(t+1) \times [1 - v_j^{(2)}(t+1) \operatorname{sgn}(w_{ji}^{(12)}(t) x_i(t+1))] \quad (13)$$

其中, $a > 0$ 是学习率。采用 $\operatorname{sgn}(w_j^{(23)}(t))$ 是为了使算法更具鲁棒性。

2.3 参数学习算法

ASN 网络的初始权系数均为 1,其可调参数 p 有:第 4 层的参数 w 、第 5 层的输出隶属度函数参数 (a_k, b_k, c_k) 和输出因子 k_3 。它学习的目的是使 ACN 网络的输出 z 最大,采用梯度法进行学习。

$$p_i(t+1) = p_i + \Delta p_i \quad (14)$$

$$\Delta p_i = \gamma R(t) \frac{\partial z}{\partial p_i} = \gamma R(t) \frac{\partial z}{\partial u} \frac{\partial u}{\partial p_i} \quad (15)$$

其中, $\gamma > 0$ 为学习率。由于 z 和 u 之间没有明确的关系,所以采用近似的方法求出 $\partial z / \partial u$,即

$$\frac{\partial z}{\partial u} = \operatorname{sgn} \frac{z(t+1) - z(t)}{u(t+1) - u(t)} \quad (16)$$

$\partial u / \partial a_k, \partial u / \partial b_k, \partial u / \partial c_k, \partial u / \partial k_3, \partial u / \partial w_p^k$ 可由式(17)求出,即

$$\begin{aligned} \frac{\partial u}{\partial a_k} &= \frac{k_3 \alpha'_k}{2 \sum_{k=1}^{m_l} \alpha'_k} (1 - \alpha'_k), & \frac{\partial u}{\partial b_k} &= \frac{k_3 \alpha'_k}{\sum_{k=1}^{m_l} \alpha'_k} \\ \frac{\partial u}{\partial c_k} &= \frac{k_3 \alpha'_k}{2 \sum_{k=1}^{m_l} \alpha'_k} (1 - \alpha'_k), & \frac{\partial u}{\partial k_3} &= \frac{\sum_{k=1}^{m_l} \alpha'_k Y_k^{-1}(\alpha'_k)}{\sum_{k=1}^{m_l} \alpha'_k} \\ \frac{\partial u}{\partial w_p^k} &= \frac{\sum_{k=1}^{m_l} \alpha'_k [Y_k^{-1}(\alpha'_k) + \alpha'_k(2b - c - a)\alpha'_k] - p}{\left(\sum_{k=1}^{m_l} \alpha'_k \right)^2} \frac{\alpha_p}{\sum_{j=1}^L \alpha_j} \end{aligned} \quad (17)$$

其中, $k=1,2,\dots,m_l; p=1,2,\dots,L$ 。

3 算例分析

为了验证所提方法的有效性,采用如图 5 所示的两区域交直流互联系统进行测试。算例的具体参数见参考文献[1]。

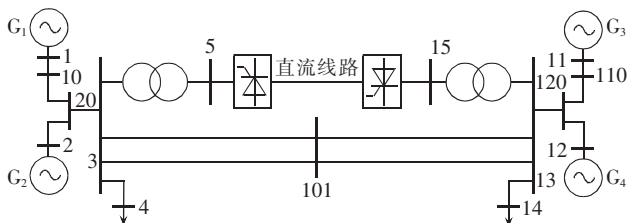


图 5 四机两区域交直流互联测试系统

Fig.5 AC/DC test system with four generators and two areas

发电机采用次暂态模型,励磁系统采用 3 阶模型,装有 PSS,不考虑调速器,负荷采用 ZIP 模型。采用 2 种工况进行仿真。在第 1 种工况下,在 0.1 s 时母线 3 发生三相接地短路故障,0.2 s 故障切除。第 2 种工况下,3-13 之间的传输功率提高 20%,故障及其切除时间不变。

图 6 给出了直流整流侧未加装附加控制器时发电机 1 和 3 之间的功角差 δ_{G1-G3} 随时间变化的曲线,分别对 HVDC 加装 CSDC 和本文所提出的阻尼控制器进行仿真,仿真结果如图 7 所示。

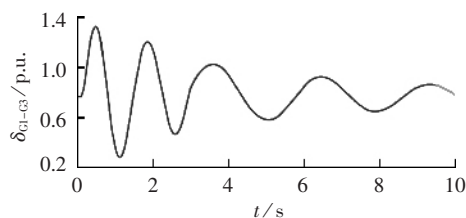


图 6 无附加阻尼控制器的响应曲线

Fig.6 Response of HVDC without supplementary controller

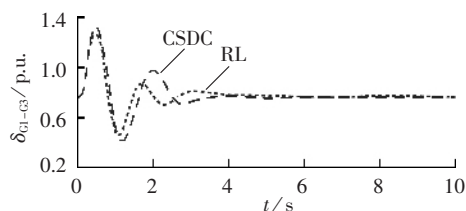


图 7 HVDC 附加阻尼控制器的响应曲线

Fig.7 Responses of HVDC with CSDC and HVDC with adaptive supplementary controller(case 1)

图 8 为在第 2 种工况下,对 2 种不同控制器进行时域仿真所得到的响应曲线。

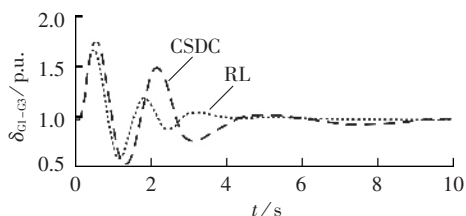


图 8 HVDC 附加阻尼控制器的响应曲线

Fig.8 Responses of HVDC with CSDC and HVDC with adaptive supplementary controller(case 2)

仿真结果表明,基于 RL 的模糊神经阻尼控制器在不同工况下比传统的阻尼控制器具有更好的响应,且体现出了较好的鲁棒性,更好地适应了各种运行状况的变化。

4 结论

利用强化学习算法设计了一种自适应直流附加阻尼控制器,用于抑制交直流互联系统区域之间的功率振荡。算例表明,与传统的阻尼控制器相比,该方法具有有效的自我学习效率,以使控制系统达到预期的控制效果,因此适合于电力系统这种存在很多不确定性因素及大扰动的系统。

参考文献:

- [1] KUNDUR P. Power system stability and control[M]. New York: McGraw- Hill, 1994.
- [2] GRAHAM R. Power system oscillations [M]. Boston: Kluwer Academic, 2000.
- [3] LARSEN E, SWANN D. Applying power system stabilizers[J]. IEEE Trans PAS, 1981, 100(6): 3017- 3046.
- [4] 黄莹, 徐政. 基于同步相量测量单元的直流附加控制器研究[J]. 中国电机工程学报, 2004, 24(9): 7- 12.

the CSEE, 2004, 24(9): 7- 12.

- [5] BUCKLEY J J. Fuzzy and neural: interactions and applications [M]. New York: Heidelberg, 1999.
- [6] HARIRI A, MALIK O P. Adaptive - network - based fuzzy logic power system stabilizer[J]. IEEE Transactions on Power System, 1998, 13(4): 1323- 1329.
- [7] MRAD F, KARAKI S, COPTI B. An adaptive fuzzy - synchronous machines stabilizer[J]. IEEE Transactions on System, 2000, 30(1): 131- 137.
- [8] YOU Ru - hua, EGHBALI H J, NEHRIR H. An online adaptive neuro - fuzzy power system stabilizer for multi - machine systems [J]. IEEE Transactions on Power System, 2003, 18(1): 128- 135.
- [9] 魏伟, 赵书强, 马燕峰. 基于 Prony 算法的模糊电力系统稳定器设计[J]. 电力自动化设备, 2005, 25(3): 54- 59.
- WEI Wei, ZHAO Shu - qiang, MA Yan - feng. Design of fuzzy power system stabilizer based on prony algorithm[J]. Electric Power Automation Equipment, 2005, 25(3): 54- 59.
- [10] 房大中, 杨晓东, 宋文南. 提高交直流电力系统稳定性的 HVDC 模糊逻辑控制器[J]. 电力系统自动化, 2002, 26(6): 23- 27.
- FANG Da - zhong, YANG Xiao - dong, SONG Wen - nan. A fuzzy logic controller for the HVDC system to enhance stability of AC/DC power system[J]. Automation of Electric Power Systems, 2002, 26(6): 23- 27.
- [11] 徐光虎, 孙衡, 陈陈. HVDC 模糊协调阻尼控制器的设计[J]. 电力系统自动化, 2004, 28(12): 18- 23.
- XU Guang - hu, SUN Qu, CHEN Chen. Design on HVDC fuzzy coordinated damping controller[J]. Automation of Electric Power Systems, 2004, 28(12): 18- 23.
- [12] 刘皓明, MOK T K, 李扬, 等. 基于遗传算法的 UPFC 模糊阻尼控制器参数优化及与梯度下降法的比较[J]. 电力自动化设备, 2005, 25(11): 5- 10.
- LIU Hao - ming, MOK T K, LI Yang, et al. Comparison between GA and gradient descent algorithm in parameter optimization of UPFC fuzzy damping controller[J]. Electric Power Automation Equipment, 2005, 25(11): 5- 10.
- [13] 朱浩骏, 蔡泽祥, 刘皓明, 等. 多馈入交直流输电系统的模糊控制器协调优化算法[J]. 中国电机工程学报, 2006, 26(13): 7- 13.
- ZHU Hao - jun, CAI Ze - xiang, LIU Hao - ming, et al. Coordinate optimization algorithm of fuzzy controller in multi - infeed AC/DC power systems[J]. Proceedings of the CSEE, 2006, 26(13): 7- 13.
- [14] 郭红霞, 吴捷. 基于强化学习算法的静止同步补偿电压控制器[J]. 电网技术, 2004, 28(19): 9- 13.
- GUO Hong - xia, WU Jie. Application of reinforcement learning to STATCOM controller [J]. Power System Technology, 2004, 28(19): 9- 13.
- [15] 郭红霞, 吴捷. 基于强化学习的模型参考自适应控制[J]. 控制理论与应用, 2005, 22(2): 291- 294.
- GUO Hong - xia, WU Jie. Model reference adaptive control based on reinforcement learning[J]. Control Theory & Application, 2005, 22(2): 291- 294.

(责任编辑: 柏英武)

作者简介:

郭力(1981-), 男, 山东济宁人, 博士研究生, 研究方向为电力系统运行与稳定(E-mail: guolidir@sohu.com);

张尧(1948-), 男, 广东廉江人, 教授, 博士研究生导师, 长期从事电力系统运行分析与控制、电压稳定性研究等;

胡金磊(1981-), 男, 河南商丘人, 博士研究生, 主要从事电力系统运行与控制方面的研究。

Adaptive HVDC supplementary damping controller based on reinforcement learning

GUO Li,ZHANG Yao,HU Jin-lei

(South China University of Technology,Guangzhou 510640,China)

Abstract: An adaptive supplementary damping controller of HVDC is presented based on reinforcement learning algorithm,which is mainly composed of neuro - fuzzy network using reinforcement signal to train the parameters of controller. Contrary to conventional fuzzy controller,the reinforcement signal,transformed from the output performance index of power system by the adaptive heuristic assessment algorithm,is fed back to the controller to update the key parameters, which effectively reduces the dependence of damping controller on accurate mathematical model. Simulation results show that this supplementary controller efficiently damps the power oscillation among areas,improves system stability,and has better robustness in various operation modes than the conventional supplementary damping controller.

This project is supported by the Key Program of National Natural Science Foundation of China(50337010).

Key words: reinforcement learning; low frequency oscillation; associative search network; conventional supplementary damping controller