

前馈神经网络结构动态增长-修剪方法

张米娜, 韩红桂, 乔俊飞

(北京工业大学 电子信息与控制工程学院, 北京 100124)

摘要:针对前馈神经网络隐含层神经元不能在线调整的问题,提出了一种自适应增长修剪算法(AGP),利用增长和修剪相结合对神经网络隐含层神经元进行调整,实现神经网络结构的自组织,从而提高神经网络的性能.同时,将该算法应用于污水处理生化需氧量(BOD)软测量,仿真实验结果表明,与其他自组织神经网络相比,AGP具有较好的泛化能力及较高的拟合精度,能够实现出水BOD的预测.

关键词:自适应增长修剪算法;BOD软测量;神经网络;自组织

中图分类号: TP183 **文献标识码:** A **文章编号:** 1673-4785(2011)02-0101-06

Research on dynamic feed-forward neural network structure based on growing and pruning methods

ZHANG Mi'na, HAN Honggui, QIAO Junfei

(College of Electronic and Control Engineering, Beijing University of Technology, Beijing 100124, China)

Abstract: Due to the unchangeable on-line problem of hidden neurons in feed-forward neural networks, an adaptive growing and pruning algorithm (AGP) was presented in this paper. This algorithm can insert and prune hidden neurons during the training process to adjust the structure of the network and achieve self organization of neural network structure, which can improve the performance of the neural network. Additionally, this algorithm has been applied to the biochemical oxygen demand (BOD) soft measurement of the wastewater treatment process. Experimental results show that the proposed algorithm can forecast the effluent BOD with better generalization ability and higher accuracy than other self-organizing neural networks.

Keywords: adaptive growing and pruning (AGP); BOD soft-measurement; neural network; self organization

多层前馈神经网络是目前应用最广泛的神经网络之一,然而在解决实际问题时神经网络的结构需要事先确定,且在训练过程中网络结构保持不变,这在很大程度上限制了网络的功能发挥^[1].因此,神经网络结构自适应设计已成为当前神经计算科学中人们共同关注的问题^[2].一般来说,如果神经网络的规模太小,网络对样本不能完全学习.如果网络的规模过大,则会产生过拟合现象,降低网络的泛化能力^[3].因此,在设计神经网络的规模时需要根据实际问题确定网络的结构.

近些年来,学者们提出了一些神经网络结构自组织算法,主要分为2类:修剪算法和增长算法.修

剪算法在初始时构造一个有冗余节点的网络,逐渐删除对网络输出贡献较小的节点,以达到要求的误差精度,是一种有效地去除冗余节点的方法.比较著名的修剪算法有最优脑外科算法(OBS)^[4].然而由于OBS只能修剪神经元间的连接,因此,对于初始规模较大的神经网络需要较长的调整时间. Philippe等人^[5]将扩展傅里叶振幅灵敏度算法(EFAST)运用于神经网络结构简化中,通过计算每个隐含层神经元输出值的灵敏度来确定每个隐含层神经元对输出的贡献大小,直接剔除冗余神经元来对神经网络结构进行修剪,提高了修剪效率.但是修剪算法的不可避免的问题是需要确定网络初始规模的大小^[6].与修剪算法相反,增长法在初始时构造一个小规模的网络,根据实际问题的复杂程度逐渐增加隐含层的节点,直到满足误差要求.其中较著名的是按需求增长算法(grows when required, GWR)^[7],GWR每次新增一个含有多个节点的网络模块(或子网络),

收稿日期:2010-04-22.

基金项目:国家“863”计划资助项目(2007AA04Z160);国家自然科学基金资助项目(60873043);北京市自然科学基金资助项目(4092010);高等学校博士点专项科研基金资助项目(200800050004).

通信作者:张米娜. E-mail: zhang. mi. na@163. com.

根据网络性能要求逐渐增加复杂度.但是增长法容易出现过度适应和不能有效提高网络性能等问题.

由于单纯的增长、修剪算法在实际应用中会出现上述问题,一些学者提出了增长修剪混合算法,利用修剪机制避免网络的过度拟合,提高泛化能力,又可以通过增长机制提高网络的鲁棒性.自适应合并增长算法(AMGA)就是一种增长修剪混合算法^[8].该算法在网络训练时根据隐含层节点的学习能力及训练过程自适应地合并及增加神经元,以满足性能要求.该算法使用的是一种自适应的非贪心策略,能够避免网络陷入局部极小. AMGA 在进行节点删减时,找到对网络输出贡献最小的节点,并将该节点和对应的关联性最大的节点进行合并以产生一个新的节点.但是 AMGA 在网络训练过程中,每次增加或删除节点后,均需要对网络进行完整地训练,运行时间较长. AMGA 对于分类问题能够较好地解决,但不适用于解决非线性函数的拟合问题.

因此,前馈神经网络结构优化设计方法仍是一个开放的问题^[9-10].文中提出一种自适应增长修剪算法(AGP),通过判断隐含层神经元输出对误差的影响进行增长或修剪,实现神经网络结构的自组织.最后,通过对非线性函数逼近和污水处理生化需氧量(BOD)软测量证明了 AGP 的有效性.

1 自适应增长修剪算法(AGP)

1.1 神经网络结构

自适应增长修剪算法以3层前馈神经网络为基础,神经网络结构图如图1所示.

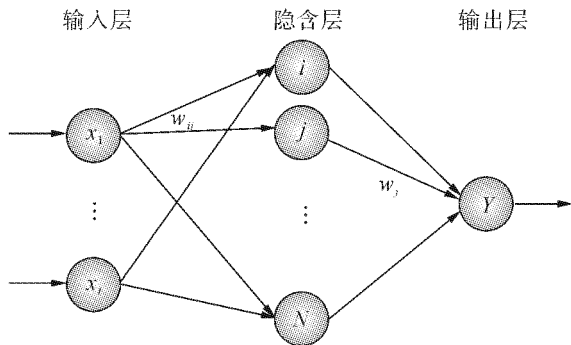


图1 前馈神经网络结构

Fig.1 The structure of feed-forward neural network

神经网络有 M 个输入神经元, N 个隐含层神经元,一个输出神经元组成.该神经网络的输出可由下式表示:

$$y = \sum_{j=1}^N w_j f\left(\sum_{i=1}^M w_{ij} x_i\right).$$

式中: $j=1, 2, \dots, N$; $i=1, 2, \dots, M$; x_1, x_2, \dots, x_M 为神经网络的输入; y 为神经网络的输出; w_{ij} 和 w_j 为隐

含层神经元 j 的输入层权值; f 为 sigmoid 函数.神经网络使用典型的梯度下降算法^[11]对神经网络连接权值进行训练.

传统3层前馈神经网络在初始结构确定后,训练过程中无法对其结构进行修改,这就限制了网络的性能发挥.

1.2 自适应增长修剪算法

为了解决前馈神经网络结构自组织问题,提出了一种自适应增长修剪算法.该算法基于3层前馈神经网络,网络的输入及输出神经元数根据实际需要给定,隐含层神经元能够在线调整.该算法通过计算隐含层各个神经元对整个网络输出的贡献值来判断该神经元对神经网络的重要性大小,并将其相关性较大的2个神经元合并,以达到精简网络的目的.当隐含层神经元合并后,经过一定步数的训练依然不能满足需要的误差精度时,对网络再进行节点增长的操作,实现神经网络隐含层结构的自动调整,以获得规模合适、泛化能力较好的神经网络.神经网络的性能评价函数为

$$E = \frac{1}{2} \sum_{v=1}^V (y_v - z_v)^2.$$

式中: y_v 和 z_v 分别为第 v 组样本的神经网络输出和期望输出, V 为总体样本数.

当神经网络的误差过大时,认为神经网络现有结构不能满足需要,在调整网络连接权值的同时调整神经网络隐含层的结构,网络调整的依据是^[12]:

$$\eta_j = \frac{\sqrt{\mu_j/V}}{\sigma_j}. \quad (1)$$

式中: η_j 为第 j 个隐含层神经元的显著性即第 j 个神经元对整个网络输出的影响程度; μ_j 为第 j 个隐含层神经元的强度; σ_j 为强度 μ_j 的方差; V 为总体样本数.强度 μ_j 为

$$\mu_j = \sum_{i=T_1}^{T_2} M_i.$$

式中: M_i 为神经网络每次循环后隐含神经元的输出对整个神经网络的输出贡献值; T_1 为前次调整时训练步数; T_2 为当前调整时训练步数.

由式(1)可以看出,当隐含层神经元对整个网络的显著性阈值为 η_{th} 时,将该节点与其对应的相关性大且显著性高的节点进行合并.通过分别计算需要合并的2个节点所占显著性的比率,并将该比率与权值相乘,得到新增节点的权值.如果神经元 a 和 b 需要合并,经过合并后,生成新的神经元 m 的权值由式(2)、(3)所示.

$$w_{im} = P_a w_{ia} + P_b w_{ib}, \quad (2)$$

$$w_m = w_a + w_b. \tag{3}$$

式中: w_m 是新神经元 m 与第 i 个输入神经元间的连接权值; w_m 是新神经元 m 与输出神经元间的连接权值;权值 w_{ia} 与 w_{ib} 是输入层第 i 个神经元与隐含层第 a 个及第 b 个神经元之间的连接权值; w_a 与 w_b 是输出层与第 a 个及第 b 个神经元之间的连接权值.

当经过节点“合并”后,神经网络依然不能满足所要求的误差精度,该算法则假定隐含层的神经元不足以完全对样本进行学习,此时需要对网络增加节点.该算法在进行节点增长时采用了“细胞分裂”的思想,该种增长机制可以看作是删除了一个节点,同时增加了 2 个新的节点. AGP 在现有的隐含层神经元中挑选出显著性最高的神经元,以此作为“父节点”,将其分裂成若干个“子节点”,如隐含层神经元 a 分裂成 Q 个新神经元,分裂后神经元的权值为

$$w_{inew} = w_{ia}, \tag{4}$$

$$w_{new} = a_{new} w_a. \tag{5}$$

$$\sum_{new=1}^Q a_{new} = 1, new = 1, 2, \cdots, Q.$$

式中: w_{inew} 是新神经元 new 与第 i 个输入神经元间的连接权值; w_{new} 是新神经元 new 与输出神经元间的连接权值;权值 w_{ia} 是输入层第 i 个神经元与隐含层第 a 个之间的连接权值; w_a 是输出层与第 a 个神经元之间的连接权值. AGP 通过“细胞分裂”的方式增加新的节点,能够保留“父节点”及“子节点”之间的关联性,减小结构调整引起的误差.

在设计神经网络的结构时,AGP 不同于普通的神经网络结构调整算法的关键一点在于,隐含层神经元的修剪不是简单的“删减”,而是通过合并隐含层的节点以减少冗余神经元.该合并机制相当于删除了 2 个节点,同时又增加了 1 个新的节点.同时,神经网络结构的生长也不是每一次调整只增加 1 个神经元,而是根据神经元的显著性大小给定新增神经元的多少.

1.3 神经网络结构优化算法流程

本文提出的神经网络结构优化算法在结构设计上采用的是一种根据隐含层神经元的显著性调整神经网络达到目标误差的策略,合并及增加神经元,可以避免初始神经网络设计过大或者过小的问题.该算法初始时构造一个 3 层前馈神经网络,隐含层神经元数为较小自然数,采用 BP 学习算法及对网络进行训练.

AGP 的具体步骤如下:

1) 创建一个初始的前馈神经网络,用 BP 算法进行训练;

2) 判断网络误差是否满足终止条件,若满足则转 8),否则转向 3);

3) 计算隐含层神经元的显著性,将显著性小于阈值 η_{th} 的神经元放入集合 S 中,否则转向 7);

4) 计算 S 集合中的神经元与非 S 集合神经元之间的相关性;

5) 合并相关性较大的 2 个神经元,以产生 1 个新的神经元,新神经元的初始连接权值根据式(2)和(3)设定;

6) 重新训练网络,若网络误差能达到第 1 次训练的水平,转 2),否则转向 7);

7) 若经过神经元合并后网络依然不能满足性能要求,则增加 1 个新的隐节点,新增神经元的连接权值根据式(4)和(5)设定,转向 2);

8) 神经网络训练结束.

研究表明,AGP 能够较好地调整神经网络的结构,通过对神经网络结构和参数的调整,最终获得的神经网络结构比较紧凑.

2 仿真实验

为了验证 AGP 的有效性和高效性,将 AGP 用于非线性函数逼近及污水处理过程 BOD 的预测中,并与 GWR^[7]和 AMGA^[8]算法相比较,证明了该算法具有较高的函数拟合精度及较好的泛化能力.

2.1 神经网络结构优化算法的非线性函数逼近

利用 AGP 对非线性函数进行逼近,所采用的非线性函数为

$$f(x) = 2.5 \times \frac{\tan x}{x^2} + x \sin x + \cos x.$$

式中: $x \in [1, 2\pi]$. 实验采用 64 组数据作为训练样本,36 组为测试样本,初始隐含层神经元个数为 12 个,经过神经网络优化算法进行网络训练,剩余隐节点个数为 8 个.图 2 显示神经网络对非线性函数的逼近效果图,图 3 显示神经网络对非线性函数逼近的误差曲线,训练过程中误差变化曲线如图 4 所示,与 GWR 算法及 AMGA 算法的详细比较结果如表 1 所示.

表 1 3 种优化算法性能比较

Table 1 Performance of the three algorithms				
优化算法	优化后 神经元	测试 误差	训练 步长	训练 时间/s
AGP 算法	8	0.03	3 729	11.8
AMGA 算法	10	0.09	6 508	16.3
GWR 算法	10	0.16	5 328	13.4

由图2和图3可以看出,AGP能够较好地对非线性函数进行拟合,具有较好的拟合精度,通过与其他结构调整算法的比较不难发现,达到相同设定误差时,AGP具有更快的训练速度和简单的网络结构。

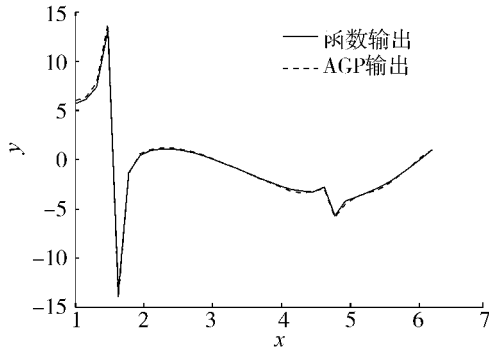


图2 AGP算法对函数逼近效果

Fig. 2 The fitting chart of nonlinear function based on AGP algorithm

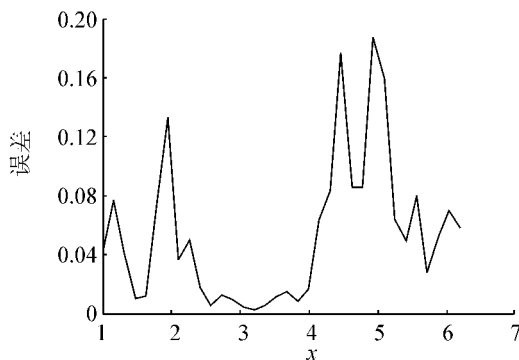


图3 AGP算法对非线性函数逼近的绝对误差曲线

Fig. 3 Absolute error of the nonlinear function based on AGP algorithm

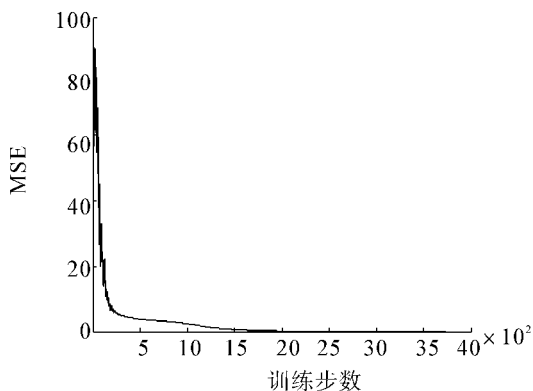


图4 AGP算法对网络的训练误差曲线

Fig. 4 Error value during the training process of AGP algorithm

2.2 神经网络结构优化算法的BOD软测量应用

由于污水处理过程机理复杂,且具有非线性、大时滞等特性.为了使污水处理系统运行良好,获得更好的出水水质,并及时取得污水处理过程中一些重要的过程参数和水质参数,以此对污水处理系统进行实时控制.其中,生化需氧量BOD是一种用微生物代谢作用所消耗的溶解氧的量来间接表示水体被有机物污染程度的一个重要指标.目前,BOD的测量很难由仪器直接完成.传统的方式为五日生化需氧量法,但获得测量结果严重滞后.因此,可采用软测量的方式实现.通过一些容易测得的数据对不可测变量进行预测,并利用神经网络建立软测量模型。

在仿真实验中,利用AGP及其他结构调整算法对污水处理过程中的生化需氧量BOD进行预测,并将实际数据与神经网络的预测数据进行对比。

本实验数据来源于2005年昆明市某小型污水处理厂数据日报表,得到的实际数据经过误差处理及主元分析后,神经网络的输入参量为:进水量 Q_{in} 、pH值、进水COD、悬浮物SS及总氮TN,神经网络的输出为出水BOD.实验中神经网络结构为5-24-1,模型结构图如图5所示。

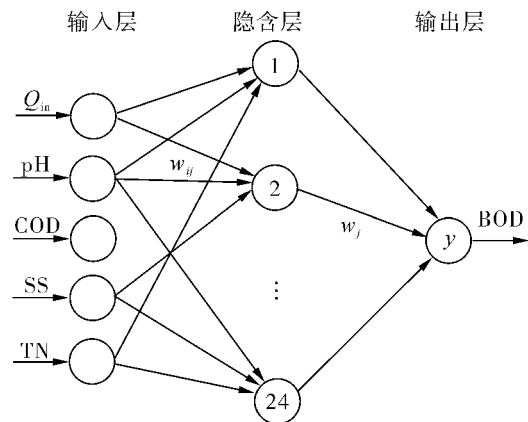


图5 前馈神经网络结构图

Fig. 5 The structure of feed-forward neural network

实验中训练样本为56组,测试样本为24组. AGP算法与AMGA算法都选择相同的初始隐含层神经元个数18个,GWR算法的初始隐含层神经元个数为3个,网络训练的目标误差均为0.01.图6显示神经网络对出水BOD的预测效果图,图7显示神经网络对出水BOD的预测结果与实际出水BOD之间的误差,训练过程中误差变化曲线如图8所示,图9所示为AGP算法训练过程中隐含层神经元调整过程.3种算法的实验结果相关数据的详细比较如表2所示。

表2 3种优化算法性能比较
Table 2 Performance of the three algorithms

优化算法	优化后 神经元	测试 误差	训练 步长	训练 时间/s
AGP 算法	12	0.046	6 358	23.8
AMGA 算法	13	0.125	7 815	26.6
GWR 算法	14	0.192	7 627	28.7

仿真实验表明,当训练结束时,AGP 的剩余神经元个数比 AMGA 算法和 GWR 算法略少,且需要的训练步长少.利用训练后的神经网络对样本进行测试时,检测误差也比 AMGA 算法和 GWR 算法的检测误差小.图 6 和图 7 表明 AGP 能够很好地预测出水 BOD 的值(实线表示实际出水 BOD,虚线表示网络预测输出 BOD),证明了基于 AGP 算法的模型的有效性.表 2 中所示 3 种优化算法性能比较,AGP 算法与 AMGA 算法及 GWR 算法相比,测试误差的精度有所提高.在训练时间上虽没有较明显的优势,但由于污水处理过程时间较长,对出水参数预测时神经网络的训练时间较长对其影响不大.

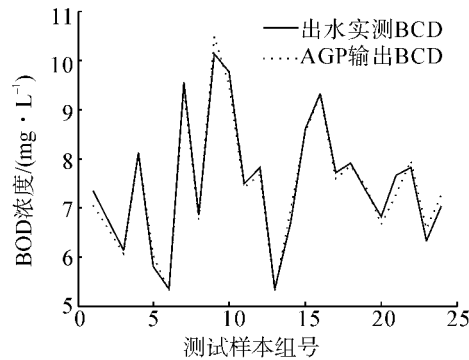


图6 AGP 算法预测 BOD 出水值结果
Fig.6 The predictive effluent BOD based on AGP algorithm

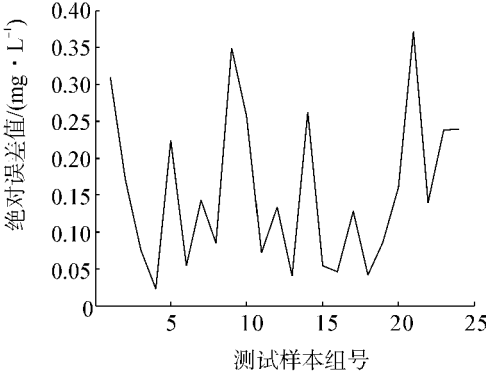


图7 AGP 算法的 BOD 预测值与实际值绝对误差曲线
Fig.7 Absolute error of the real effluent BOD and predictive value based on AGP algorithm

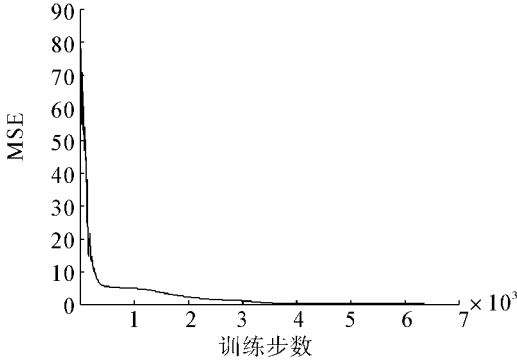


图8 AGP 算法对网络的训练误差曲线
Fig.8 Error value during the training process of AGP algorithm

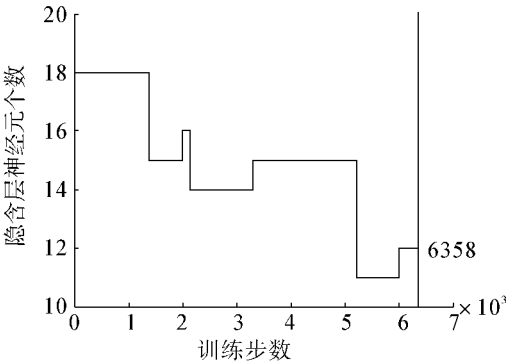


图9 AGP 算法隐含层神经元调整过程图
Fig.9 The adjusting process of hidden layer neurons

3 结 论

提出了一种神经网络结构自适应增长修剪算法 (AGP),该算法能够在线调整神经网络结构,从而提高神经网络的性能.将 AGP 应用于非线性函数进行逼近和污水处理过程 BOD 软测量中,取得较好的效果,通过与 AMGA 算法和 GWR 算法进行比较,AGP 具有如下的优点:

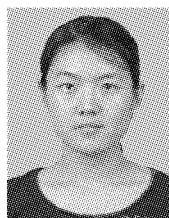
- 1)能够根据实际问题的复杂程度,对网络结构自动调整,解决了一般前馈神经网络结构不能在线调整的问题.
- 2)AMGA 算法和 GWR 算法相比,AGP 具有更少的训练步长、较短的训练时间以及更简单的网络结构.
- 3)具有较好的泛化能力及更高的拟合精度,能够对生化需氧量 BOD 实现有效的预测,在解决污水处理过程参数的预测问题上更具有一定的实用性.

参考文献:

[1] 乔俊飞, 张颖. 一种多层前馈神经网络的快速修剪算法 [J]. 智能系统学报, 2008, 3(2): 173-176.
QIAO Junfei, ZHANG Ying. Fast unit pruning algorithm for

- multilayer feedforward network design[J]. CAAI Transactions on Intelligent Systems, 2008, 3(2): 173-176.
- [2] 杨慧中, 王伟娜. 神经网络的两种结构优化算法研究[J]. 信息与控制, 2006, 35(6): 700-704.
- YANG Huizhong, WANG Weina. Two structure optimization algorithms for neural networks[J]. Information and Control, 2006, 35(6): 700-704.
- [3] BORTSMAN M, ALADIEM M. A growing and pruning method for radial basis function networks[J]. IEEE Transactions on Neural Networks, 2009, 20(6): 1039-1045.
- [4] HASSIBI B, STORK D G. Second order derivatives for network pruning: optimal brain surgeon[C]//Advances in Neural Information Processing Systems. San Mateo, USA: Morgan Kaufman, 1993: 164-171.
- [5] LAURET P, FOCK E, MARA T A. A node pruning algorithm based on a Fourier amplitude sensitivity test method[J]. IEEE Transactions on Neural Networks, 2006, 17(2): 273-293.
- [6] XU Jinhua, DANIEL W. A new training and pruning algorithm based on node dependence and Jacobian rank deficiency[J]. Neurocomputing, 2006, 70(1): 544-558.
- [7] MARSLAND S, SHAPIRO J. A self-organizing network that grows when required[J]. Neural Networks, 2002, 15(8): 1041-1058.
- [8] ISLAM M, SATTAR A, AMIN F, YAO Xin. A new adaptive merging and growing algorithm for designing artificial neural networks[J]. IEEE Trans Systems, Man, Cybernetics-Part B: Cybernetics, 2009, 39(3): 705-722.
- [9] QIAO Junfei, HAN Honggui. A repair algorithm for RBF neural network and its application to chemical oxygen demand modeling[J]. International Journal of Neural Systems, 2010, 20(1): 63-74.
- [10] JOHNSON C, VENAYAGAMOORTHY G K, MITRA P. Comparison of a spiking neural network and an MLP for robust identification of generator dynamics in a multimachine power system[J]. Neural Networks, 2009, 22(5/6): 833-841.
- [11] RUMELHART D E, HINTON G E, WILLIAMS R J. Learning internal representations by error propagation[M]. Parallel Distributed Processing. Cambridge, USA: MIT Press, 1986: 318-362.
- [12] 韩红桂, 甄博然, 乔俊飞. 动态结构优化神经网络及其在溶解氧控制中的应用[J]. 信息与控制, 2010, 39(3): 354-360.
- HAN Honggui, ZHEN Boran, QIAO Junfei. Dynamic structure optimization neural network and its applications to dissolved oxygenic (DO) control[J]. Information and Control, 2010, 39(3): 345-360.
- [13] PENG Jianxun, LI Kang, HUANG Deshuang. A hybrid forward algorithm for RBF neural network construction[J]. IEEE Transactions on Neural Networks, 2006, 17(6): 1439-1451.

作者简介:



张米娜, 女, 1986 年生, 硕士研究生, 主要研究方向为神经网络结构优化设计、智能控制理论与应用。



韩红桂, 男, 1983 年生, 博士研究生, 主要研究方向为智能信息处理、智能控制理论与应用。



乔俊飞, 男, 1968 年生, 教授, 博士, 主要研究方向为神经网络结构分析与设计、计算智能与智能优化控制。