

基于神经元特性的径向基函数神经网络自组织设计方法

贾丽杰, 李文静, 乔俊飞[†]

(北京工业大学 信息学部, 北京 100124; 计算智能和智能系统北京市重点实验室, 北京 100124)

摘要: 针对径向基神经网络(RBF)网络隐层结构难以确定的问题, 本文介绍了一种基于神经元特性的RBF神经网络自组织设计方法, 该方法将神经元的激活活性、显著性、相关性相结合设计RBF(ASC-RBF)神经网络. 首先利用神经元的激活活性, 实现隐层神经元的自适应增加, 结合神经元的显著性以及神经元之间的相关性, 实现神经元的自适应替换和合并, 完成网络自组织设计并提高网络的紧凑性, 然后利用二阶梯度算法对网络参数进行修正学习, 保证了RBF网络的精度; 另外, 针对网络结构自组织机制给出了稳定性分析; 最后通过两个基准非线性系统建模仿真实验以及实际污水处理过程水质参数预测实验验证, 证明该算法的有效性. 对比实验结果表明, ASC-RBF神经网络与现有的自组织网络相比, 在保证泛化性能的同时, 该网络的训练速度更快, 而且有更紧凑的网络结构.

关键词: 径向基函数神经网络; 自组织; 结构设计; 二阶算法; 非线性系统建模

引用格式: 贾丽杰, 李文静, 乔俊飞. 基于神经元特性的径向基函数神经网络自组织设计方法. 控制理论与应用, 2020, 37(12): 2618 – 2626

DOI: 10.7641/CTA.2020.00010

Self-organizing design of radial basis function neural network based on neuron characteristics

JIA Li-jie, LI Wen-jing, QIAO Jun-fei[†]

(Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China;
Beijing Key Laboratory of Computational Intelligence and Intelligent System, Beijing 100124, China)

Abstract: Aiming at the problem that the hidden layer structure of radial basis neural function (RBF) neural network is difficult to determine, this paper introduces a self-organizing design method of RBF neural network based on the characteristics of neurons. This method combines the activation activity, significance and correlation of neurons Combined design of RBF (ASC-RBF) neural network. Firstly, The network uses the activity of neurons to adaptively increase the hidden layer neurons, and combines with its significance and correlation to complete the adaptive replacement and merging of neurons. Furthermore, the self-organizing design of the neural network is completed and its compactness is improved. Then, a second-order algorithm is used to modify the network parameters to ensure the accuracy of the RBF network. In addition, a stability analysis is given for the network structure self-organization mechanism. Finally, in order to verify the effectiveness of the proposed ASC-RBF network, two benchmark nonlinear system modeling experiments and a water quality parameter prediction experiment in a wastewater treatment system are performed. The results demonstrate that compared with the existing self-organizing network, the ASC-RBF neural network has faster training speed and a more compact network structure while ensuring generalization performance.

Key words: radial basis function neural network; self-organization; structural design; second-order algorithm; nonlinear system modeling

Citation: JIA Lijie, LI Wenjing, QIAO Junfei. Self-organizing design of radial basis function neural network based on neuron characteristics. *Control Theory & Applications*, 2020, 37(12): 2618 – 2626

收稿日期: 2020-01-06; 录用日期: 2020-07-13.

[†]通信作者. E-mail: junfeiq@bjut.edu.cn.

本文责任编辑: 王凌.

国家自然科学基金项目(61890930-5, 61533002, 61603009), 国家重点研发计划项目(2018YFC1900800-5), 北京市自然科学基金项目(4182007), 北京市教委科技一般项目(KM201910005023), 北京工业大学日新人才计划项目(2017-RX(1)-04)资助.

Supported by the National Natural Science Foundation of China (61890930-5, 61533002, 61603009), the National Key Research and Development Project (2018YFC1900800-5), the Beijing Natural Science Foundation (4182007), the Beijing Municipal Education Commission Foundation (KM201910005023) and the “Rixin Scientist” Foundation of Beijing University of Technology (2017-RX(1)-04).

1 引言

径向基函数(radial basis function, RBF)神经网络因其结构简单、逼近能力强, 以及收敛速度快、不易陷入局部极小值等优点, 在模式识别^[1-2]、信号处理^[3]、故障诊断^[4-5]、非线性系统建模^[6-7]与自适应控制^[8-10]等方面得到了广泛的应用. RBF神经网络结构设计和参数优化是其成功应用的关键. 随着RBF网络应用环境复杂性的提高, 传统的通过经验法和试凑法^[11-12]来确定RBF网络结构的方法已经不能满足其实际应用. 为了寻求一个可以根据处理对象自适应调整的RBF神经网络, 大量的文章^[13-22]对网络结构的自组织设计进行了深入的研究, 主要包括: 增长法^[13]、修剪法^[14]、增长-修剪法^[15-22]、智能寻优法^[23-25]等.

Platt等人^[13]提出一种增长型资源分配网络模型(resource allocation network, RAN), 可以根据处理的对象, 通过隐含层神经元的连续生长构建网络结构, 达到处理复杂信息的目的. 它对RBF神经网络自组织的发展做出了巨大贡献, 但是对于RAN, 网络结构只增不减, 网络中必然会出现冗余的神经元, 这会影响网络的泛化性能. 为了解决冗余神经元的问题, 在RAN的基础上, Lu等人^[15]提出了一种最小资源神经网络模型(minimal resource allocation network, MRAN), 在自适应增加隐含层神经元的同时, 还通过每个神经元对整个网络输出的贡献对神经元进行修剪, 以减少RBF神经网络的复杂度, 但是MRAN增加隐节点时具有一定的盲目性, 新增加的隐节点的核函数中心随机确定, 对网络的泛化性能会有一定影响. Huang等人^[16]提出了一种广义增长和剪枝的RBF网络模型(generalized growing and pruning radial basis function, GGAP-RBF), 该模型是在GAP-RBF^[17]模型的基础上提出的一个更高级的模型. GGAP-RBF通过顺序学习的方法判断隐含层神经元的重要性, 并将其与学习精度联系起来, 从而构建神经网络结构, 得到的网络较为紧凑. 但是该网络的参数学习算法收敛速度慢, 会影响整个网络的收敛速度. 杜大军等人^[18]提出利用一种快速回归算法来确定RBF神经网络的结构和中心, 尽管该方法收敛速度快, 但是无法避免局部极小值. Han等人^[19]提出一种自组织RBF神经网络, 该网络根据RBF神经元径向作用范围和要求的稳定误差, 对隐含层神经元进行自适应增加和删减. 但是该算法涉及参数过多, 会增加网络设计的复杂度. Hao等人^[20]提出一种基于误差校正算法的RBF神经网络, 根据消除误差的思想设定新增加的RBF神经元参数. Qiao等人^[21]提出一种基于神经元活动增量的RBF网络, 网络的隐含层结构根据局部场势和平均发射率来计算神经元活动并在其基础上动态构建. 以上两个网络都利用二阶算法保证了收敛速度且避免了局部极值问题, 但是结构设计过程中没有神经元删减环节,

当网络增加到一定程度, 有可能存在冗余神经元. Xie等人^[22]提出了一种基于神经元自适应分离合并的网络结构自组织设计算法, 但是该算法要先根据历史数据对中心进行初始化, 而在实际应用有可能会因为历史数据和当前数据差异过大影响网络的性能. RBF神经网络结构确定又可以看成是网络神经元个数的寻优, 因此有很多文献利用智能寻优算法确定网络结构, 如粒子群算法^[23]、人工蜂群算法^[24]、万有引力所搜算法^[25]等, 但是这些方法都太过耗时.

根据以上的分析, 目前RBF神经网络设计仍然是一个开放性的具有挑战的问题. 为了获得结构和性能俱佳的RBF神经网络, 本文提出一种基于神经元特性的网络结构自组织算法. 首先利用神经元激活活性, 实现隐含层神经元的自适应增加, 将神经元激活活性与其显著性以及神经元之间的相关性结合, 实现神经元的自适应合并和替换, 确保可以获得紧凑的神经网络结构; 在实现网络结构自组织设计的同时, 确定网络的初始参数, 然后利用二阶梯度学习算法对网络参数进行修正, 并对网络的结构稳定性给出了证明. 最后通过非线性动态系统辨识、Mackey-Glass时间序列预测以及实际应用中污水处理过程的出水生化需氧量(biochemical oxygen demand, BOD)浓度预测实验验证, 证明了该方法的有效性.

2 RBF神经网络

RBF网络结构包括3层: 输入层、隐含层、输出层. 不失一般性, RBF神经网络采用多输入单输出的 $I-J-1$ 结构, 网络拓扑结构图如图1所示.

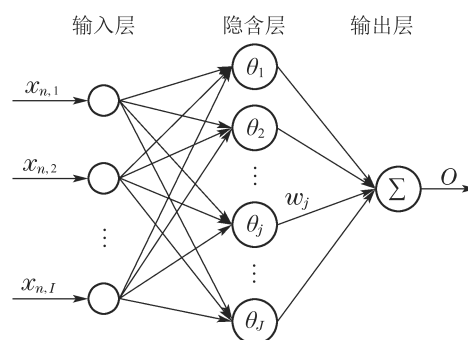


图 1 RBF网络拓扑结构图

Fig. 1 RBF network topology structure

网络的输出可描述为

$$o(\mathbf{x}) = \sum_{j=1}^J w_j \theta_j(\mathbf{x}_n), \quad (1)$$

其中: $\mathbf{x}_n = [x_{n,1} \ x_{n,2} \ \cdots \ x_{n,I}]^T$ 是网络的第 n 个 I 维输入向量, w_j 是隐含层第 j 个神经元与输出节点之间的连接权值, J 是隐含层的神经元个数, θ_j 是隐含层第 j 个神经元的输出函数. 本文激活函数选用的是高斯函数, 表达式如式(2)所示:

$$\theta_j(\mathbf{x}_n) = \exp\left(-\frac{\|\mathbf{x}_n - \mathbf{c}_j\|^2}{\sigma_j^2}\right), \quad (2)$$

其中: $\mathbf{c}_j$ 表示第 j 个神经元的中心向量; $\|\mathbf{x}_n - \mathbf{c}_j\|$ 表示样本 $\mathbf{x}_n$ 与中心 $\mathbf{c}_j$ 之间的欧几里得距离; σ_j 表示第 j 个神经元的宽度。

3 ASC-RBF网络设计

ASC-RBF神经网络设计的思想主要是基于神经元激活活性(activation activity)、显著性(significance)及神经元之间相关性特性(correlation)的判断来进行神经元的增加、替换、合并,进而确定RBF神经网络的结构,在结构确定之后利用二阶学习算法对网络的参数进行训练调整。

3.1 ASC-RBF网络结构设计

RBF神经网络最显著的特点是隐节点的基函数采用距离函数,并使用高斯函数作为激活函数,RBF关于 n 维空间的一个中心点具有径向对称性,而神经元的输入距离该中心点越远,神经元的激活程度就越低,隐节点的这个特性被称为“局部特性”。这就意味着,如果输入在任意方向远离中心点的话,输出将趋近于0,越靠近中心点,则激活程度越高,活性值越大,如图2所示。

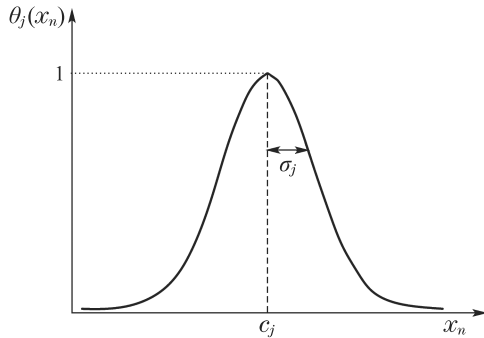


图2 神经元“局部特性”

Fig. 2 Neuron “local characteristics”

由此引入神经元激活活性评价指标,如式(3)所示:

$$\theta_j(\mathbf{x}_n) = \exp\left(-\frac{\|\mathbf{x}_n - \mathbf{c}_j\|^2}{\sigma_j^2}\right) \geq \theta_0, \quad (3)$$

这里: θ_j 为隐含层第 j 个神经元被第 n 个输入样本 $\mathbf{x}_n$ 激活后的活性值, θ_0 为活性阈值。当激活活性值大于阈值时,说明该样本可以激活神经元,并保证一定的活性;反之,不能激活神经元,说明现有的神经元不能对当前样本进行学习,需要增加新的隐含层神经元。

初始时刻,网络中没有隐含层节点,当第1个样本 $\mathbf{x}_1$ 进入网络之后,将输入的第1个向量作为隐含层第1个中心,将网络误差作为连接权值。设在第 n 时刻,假设已经存在 k 个隐含层神经元节点,当第 n 个数据样本进入网络时,计算所有神经元在样本 $\mathbf{x}_n$ 的作用下的激

活性值,如式(4)所示:

$$\theta(\mathbf{x}_n) = [\theta_1(\mathbf{x}_n) \cdots \theta_j(\mathbf{x}_n) \cdots \theta_i(\mathbf{x}_n) \cdots \theta_k(\mathbf{x}_n)]. \quad (4)$$

对所有神经元的活性值按照降序排列,找到在样本 $\mathbf{x}_n$ 作用下活性值最大的两个神经元 $j_{\max}$ 和 $i_{\max}$,如式(5)–(6)所示:

$$j_{\max} = \arg \max_{j=1,2,\dots,k} \{\theta_j(\mathbf{x}_n)\}, \quad (5)$$

$$i_{\max} = \arg \min_{i=1,2,\dots,j-1,j+1,\dots,k} \{\theta_i(\mathbf{x}_n)\}. \quad (6)$$

3.1.1 神经元增加机制

为了确保输入样本和现有神经元的相关性,当有新的数据样本进入网络时,是否添加新的神经元取决于当前所有神经元的激活活性值,如果当前所有神经元都不能被新样本激活,则添加新的神经元。

$\theta_{j_{\max}}$ 是所有神经元活性值中的最大值,如果 $\theta_{j_{\max}} < \theta_0$,就说明当前输入样本 $\mathbf{x}_n$ 不能激活任何一个神经元,即样本不能被任何现有神经元学习,所以需要添加一个新的神经元来处理该信息,新神经元的参数设置如式(7)所示:

$$\begin{cases} \mathbf{c}_{J+1} = \mathbf{x}_n, \\ w_{J+1} = |e_n|, \\ \sigma_{J+1} = \min\{\|\mathbf{c}_j - \mathbf{x}_n\|\}, j = 1, 2, \dots, k, \end{cases} \quad (7)$$

其中: $\mathbf{c}_j$ 为距离样本点 $\mathbf{x}_n$ 最近的隐含层神经元; e_n 为当 $\mathbf{x}_n$ 为网络输入时,神经网络的输出误差,计算式如式(8)所示:

$$e_n = y_n - o_n. \quad (8)$$

3.1.2 神经元替换和合并机制

当样本可以激活神经元时,检查样本所激活神经元的个数,如果只有一个神经元被激活,则计算当前样本 $\mathbf{x}_n$ 的显著性^[16]和其所激活神经元的显著性,并判断大小;如果样本可以激活多个神经元,则检查神经元活性值最大的两个神经元的相关性,如果相关性大于阈值,则合并神经元,如果相关性小于阈值,则以活性值最大的神经元为准进行下一步计算。

1) 如果当前样本 $\mathbf{x}_n$ 激活的神经元只有一个,根据式(4)–(6),即 $\theta_{j_{\max}} > \theta_0$ 和 $\theta_{i_{\max}} < \theta_0$,只有神经元 $j_{\max}$ 被激活,则计算当前样本 $\mathbf{x}_n$ 的显著性 $E_{\text{sig}}(\mathbf{x}_n)$ 和神经元 $j_{\max}$ 的显著性 $E_{\text{sig}}(\mathbf{c}_{j_{\max}})$,并比较二者大小,其中样本 $\mathbf{x}_n$ 显著性 $E_{\text{sig}}(\mathbf{c}_{j_{\max}})$ 代表了当样本 $\mathbf{x}_n$ 成为中心之后对RBF网络性能的贡献性。计算式分别如式(9)–(10)所示:

$$E_{\text{sig}}(\mathbf{x}_n) = \left| \frac{(1.8 \times \kappa \times \|\mathbf{x}_n - \mathbf{c}_{j_{\max}}\|)^I \|e_n\|}{L(\mathbf{x})} \right|, \quad (9)$$

$$E_{\text{sig}}(\mathbf{c}_{j_{\max}}) = \left| \frac{(1.8 \sigma_{j_{\max}})^I w_{j_{\max}}}{L(\mathbf{x})} \right|, \quad (10)$$

其中: κ 为径向基函数的宽度衰减因子; I 为样本输入空间的维度; $L(\mathbf{x})$ 表示样本范围的大小; e_n 为当 $\mathbf{x}_n$ 为网络输入时, 神经网络的输出误差。

如果 $E_{\text{sig}}(\mathbf{x}_n) > E_{\text{sig}}(\mathbf{c}_{j\max})$, 说明当前的输入样本 $\mathbf{x}_n$ 比神经元 $j\max$ 具有更高的显著性, 能处理更多的信息, 所以, 将神经元 $j\max$ 的中心向量替换为当前样本 $\mathbf{x}_n$, 同时设定中心点对应的宽度值和输出权值, 如式(11)所示:

$$\begin{cases} \mathbf{c}_{j\max_new} = \mathbf{x}_n \\ \sigma_{j\max_new} = \min\{\|c_i - x_n\|\}, i = 1, 2, \dots, J, \\ w_{j\max_new} = \frac{w_{j\max_old}\theta_{j\max_old}}{\theta_{j\max_new}}. \end{cases} \quad (11)$$

如果 $E_{\text{sig}}(\mathbf{x}_n) < E_{\text{sig}}(\mathbf{c}_{j\max})$, 说明当前神经元的显著性依然很高, 则只调整神经元 $j\max$ 的连接权值, 如式(12)所示:

$$w_{j\max} = e_n. \quad (12)$$

2) 如果当前样本 $\mathbf{x}_n$ 可以激活多个神经元, 根据式(4)–(6), 即有 $\theta_{j\max} > \theta_0$ 和 $\theta_{imax} > \theta_0$, 选出函数值最大的两个神经元 $j\max$ 和 $imax$ 。

计算两个神经元之间的相关性 $R_{j\max,imax}$, 利用最著名的皮尔逊矩相关系数^[26]来测量神经元间的相关关系. 计算式如式(13)所示:

$$R_{j\max,imax} = \frac{\sum_{n=1}^N (\theta_{j\max}(\mathbf{x}_n) - \bar{\theta}_{j\max})(\theta_{imax}(\mathbf{x}_n) - \bar{\theta}_{imax})}{\sigma_{\theta_{j\max}}\sigma_{\theta_{imax}}}, \quad (13)$$

其中:

$$\sigma_{\theta_{j\max}} = \sqrt{\sum_{n=1}^N (\theta_{j\max}(\mathbf{x}_n) - \bar{\theta}_{j\max})^2},$$

$$\sigma_{\theta_{imax}} = \sqrt{\sum_{n=1}^N (\theta_{imax}(\mathbf{x}_n) - \bar{\theta}_{imax})^2},$$

其中: $\theta_{j\max}(\mathbf{x}_n)$ 和 $\theta_{imax}(\mathbf{x}_n)$ 分别是当训练样本为 $\mathbf{x}_n$ 时隐含层神经元 $j\max$ 和 $imax$ 的输出; 变量 $\bar{\theta}_{j\max}$ 和 $\bar{\theta}_{imax}$ 分别表示对于目前所有训练样本神经元 $j\max$ 和 $imax$ 的输出平均值。

如果 $j\max$ 和 $imax$ 两个神经元的相关性 $R_{j\max,imax}$ 大于阈值 R_0 , 说明当前两个神经元对整个网络的学习有很强的关联性, 则可将 $j\max$ 和 $imax$ 两个神经元合并为一个神经元, 以简化神经网络的结构, 合并之后的新神经元 m 参数设置为式(14)所示:

$$\mathbf{c}_m = \frac{\mathbf{c}_{j\max} + \mathbf{c}_{imax}}{2}, \quad (14a)$$

$$\sigma_m = \max(\sigma_{j\max}, \sigma_{imax}), \quad (14b)$$

$$w_m = \frac{w_{j\max}\theta_{j\max}(\mathbf{x}_n) + w_{imax}\theta_{imax}(\mathbf{x}_n)}{\theta_m(\mathbf{x}_n)}. \quad (14c)$$

如果 $j\max$ 和 $imax$ 两个神经元的相关性 $R_{j\max,imax}$ 小于阈值 R_0 , 则按照当前样本激活一个神经元 $j\max$ 来设计网络。

3) 当所有样本被学习完, 停止, 此时神经网络结构搭建完成。

3.2 网络参数学习

在RBF神经网络结构确定之后, 需要对网络的参数进行调整, 包括隐含层到输出层的连接权值以及神经元的中心和宽度. 采用二阶学习算法^[30]来训练参数, 直到达到期望精度为止. 算法的更新规则为

$$\Delta_{n+1} = \Delta_n - (Q_n + \mu_n I)^{-1} g_n, \quad (15)$$

其中: Δ 为网络的所有待训练的参数, 包括中心 $\mathbf{c}$, 宽度 σ 和权值 w ,

$$\Delta = [\omega_1 \ \dots \ \omega_J \ \dots \ c_{1,1} \ \dots \ c_{1,I} \ \dots \ c_{J,1} \ \dots \ c_{J,I} \ \dots \ \sigma_1 \ \dots \ \sigma_J]. \quad (16)$$

Q 为类海塞矩阵, g 为梯度向量. 这里的类海塞矩阵 Q 是由类海塞子矩阵 q_n 相加得到

$$Q = \sum_{n=1}^N q_n, \quad q_n = j_n^T j_n. \quad (17)$$

梯度向量由子梯度向量相加得到

$$g = \sum_{n=1}^N \eta_n, \quad \eta_n = j_n^T e_n. \quad (18)$$

误差 e_n 是期望输出 Y_n 和实际输出 o_n 的差值:

$$e_n = y_n - o_n. \quad (19)$$

j_n 为雅克比分量:

$$j_n = \begin{bmatrix} \frac{\partial e_n}{\partial \omega_1} & \dots & \frac{\partial e_n}{\partial \omega_J} & \frac{\partial e_n}{\partial c_{1,1}} & \dots & \frac{\partial e_n}{\partial c_{1,I}} & \dots \\ \dots & \frac{\partial e_n}{\partial c_{J,1}} & \dots & \frac{\partial e_n}{\partial c_{J,I}} & \frac{\partial e_n}{\partial \sigma_1} & \dots & \frac{\partial e_n}{\partial \sigma_J} \end{bmatrix}. \quad (20)$$

雅克比分量中对应的元素为

$$\frac{\partial e_n}{\partial \omega_j} = -\theta_j(\mathbf{x}_n), \quad (21)$$

$$\frac{\partial e_n}{\partial c_{j,i}} = -\frac{2\omega_j \theta_j(\mathbf{x}_n)(\mathbf{x}_{n,i} - c_{j,i})}{\sigma_j^2}, \quad (22)$$

$$\frac{\partial e_n}{\partial \sigma_j} = -\frac{2\omega_j \theta_j(\mathbf{x}_n)\|\mathbf{x}_n - \mathbf{c}_j\|^2}{\sigma_j^3}. \quad (23)$$

Δ 为学习率参数, 在每次迭代中都会随着训练误差动态变化。

选取均方根误差(root mean square error, RMSE) 为网络训练过程的评价标准, RMSE 计算公式如下所示:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{n=1}^N (y_n - o_n)^2}. \quad (24)$$

在训练过程时,将当前更新的RMSE值与之前两次的RMSE值进行比较,如果RMSE连续两次增加时,则参数训练过程结束.

4 稳定性分析

ASC-RBF神经网络的稳定性将影响网络的性能,所以下面给出网络的稳定性证明. ASC-RBF神经网络的稳定性证明主要分为3个部分:神经元增加、神经元替换和神经元合并.为了方便讨论,假设在第 n 时刻, RBF网络含有 J 个隐层神经元,网络其当前误差如式(25)所示:

$$e_J(n) = y(n) - \sum_{j=1}^J \omega_j \theta_j(\mathbf{x}(n)). \quad (25)$$

引理1 在第 n 时刻,如果新进来的样本满足神经元增加机制,则增加神经元,并设置其参数如式(7)所示,则新增的神经元不会增加网络的误差.

证 当隐层神经元由 J 个增加到 $J+1$ 个时,网络的误差为

$$\begin{aligned} e_{J+1}(n) &= \\ y(n) - \sum_{j=1}^{J+1} \omega_j \theta_j(\mathbf{x}(n)) &= \\ y(n) - \left(\sum_{j=1}^J \omega_j \theta_j(\mathbf{x}(n)) + \omega_{J+1} \theta_{J+1}(\mathbf{x}(n)) \right) &= \\ y(n) - \sum_{j=1}^J \omega_j \theta_j(\mathbf{x}(n)) - \omega_{J+1} \theta_{J+1}(\mathbf{x}(n)). \end{aligned} \quad (26)$$

结合式(7)和式(25),神经元增加后的神经网络的误差即为

$$e_{J+1}(n) = e_J(n) - \omega_{J+1} \theta_{J+1}(\mathbf{x}(n)) < e_J(n). \quad (27)$$

由此可知,神经元增加之后神经网络的误差不会变大. 证毕.

引理2 在第 n 时刻,如果新样本满足神经元替换机制,则当前神经元 $j_{\max}$ 被样本替换,替换之后新神经元参数设置如式(11)所示.神经元的个数在替换前后不会发生变化,神经元替换机制不会影响当前时刻的网络误差.

证 神经元换机制,相当于增加一个神经元的同时,删减一个神经元,则神经元替换机制发生之前, RBF网络的误差为

$$\begin{aligned} e_{J,\text{before}}(n) &= \\ y(n) - \sum_{j=1}^J \omega_j \theta_j(\mathbf{x}(n)) &= \\ y(n) - \left(\sum_{j=1}^{J-1} \omega_j \theta_j(\mathbf{x}(n)) + \omega_{j_{\max}} \theta_{j_{\max}}(\mathbf{x}(n)) \right) &= \end{aligned}$$

$$\begin{aligned} &\theta_{j_{\max,\text{old}}}(\mathbf{x}(n)) = \\ y(n) - \left(\sum_{j=1}^J \omega_j \theta_j(\mathbf{x}(n)) - \omega_{j_{\max,\text{new}}} \times \right. \\ &\left. \theta_{j_{\max,\text{new}}}(\mathbf{x}(n)) + \omega_{j_{\max,\text{old}}} \theta_{j_{\max,\text{old}}}(\mathbf{x}(n)) \right). \end{aligned} \quad (28)$$

将式(11)中的参数设置代入式(28),可知神经元替换之后的网络误差为

$$e_{J,\text{before}}(n) = y(n) - \sum_{j=1}^J \omega_j \theta_j(\mathbf{x}(n)) = e_{J,\text{after}}(n). \quad (29)$$

由此可知,当神经元发生替换,并不影响当前网络的误差. 证毕.

引理3 在第 n 时刻,如果有两个神经元满足神经元合并机制,则将两个神经元合并为一个神经元,合并之后的新神经元的参数设置如式(14)所示.神经元合并机制不会影响网络当前的误差.

证 两个神经元合并,相当于删减了两个神经元的同时增加了一个新的神经元,此时网络隐层神经元个数变为 $J-1$ 个,当前误差为

$$\begin{aligned} e_{J-1}(n) &= \\ y(n) - \sum_{j=1}^{J-1} \omega_j \theta_j(\mathbf{x}(n)) &= \\ y(n) - \left(\sum_{j=1}^J \omega_j \theta_j(\mathbf{x}(n)) - \omega_{j_{\max}} \theta_{j_{\max}}(\mathbf{x}(n)) - \right. \\ &\left. \omega_{j_{\max}} \theta_{j_{\max}}(\mathbf{x}(n)) + \omega_m \theta_m(\mathbf{x}(n)) \right). \end{aligned} \quad (30)$$

结合式(14)和式(30),神经元合并之后网络的误差为

$$e_{J-1}(n) = y(n) - \sum_{j=1}^J \omega_j \theta_j(\mathbf{x}(n)) = e_J(n). \quad (31)$$

由此可知,当神经元发生合并,并不影响当前网络的误差. 证毕.

由以上分析可知,ASC-RBF神经网络的自组织机制并不影响网络的稳定性.同时保留显著性最高的神经元,可以避免有冗余的神经元,获得更为简洁的网络结构.

5 仿真实验

本文选取非线性动态系统辨识、Mackey-Glass时间序列预测以及实际应用中污水处理过程的出水BOD浓度预测实验对ASC-RBF神经网络的性能进行测试,并与现有的网络自组织算法进行了实验对比.在所有实验中,相关阈值 R_0 设置为0.8,以保证两个神经元只有在强相关时才合并.在LM(Levenberg-Marquardt)算法,学习系数初始值参考对比算法设置为 $\Delta = 0.01$.

所有仿真实验均在理想的软硬件环境中进行:联想台式机(i7-7700, CPU为3.60 GHz, 8.0 GB RAM),

操作系统Windows 7, MATLAB软件版本2014a.

5.1 非线性系统辨识

非线性动态系统辨识常被用来验证所设计的神经网络的性能, 本实验中所用到的非线性系统如式(32)所示:

$$y(t+1) = \frac{y(t)y(t-1)[y(t)+2.5]}{1+y^2(t)+y^2(t-1)} + u(t), \quad (32)$$

其中: 非线性动态系统的初始值为 $u(t) = \sin \frac{2\pi t}{25}$, $y(0) = 0, y(1) = 0$.

该模型描述如式(33)所示:

$$\hat{y}(t+1) = f(y(t), y(t-1), u(t)). \quad (33)$$

输入向量为 $[y(t) \ y(t-1) \ u(t)]$, 预测 $[y(t+1)]$ 的值. 本实验通过式(32)产生1000个实验数据用来验证网络模型的性能, 其中用800个样本进行训练($t = 1, 2, \dots, 800$), 用200个样本进行测试($t = 801, \dots, 1000$). 均方误差(mean square error, MSE)期望的值参考对比算法, 设置为0.001. 激活活性值阈值 $\theta_0 = 0.7$.

图3-5分别表示了ASC-RBF网络的非线性系统辨识结果、训练RMSE、隐含层神经元个数变化曲线图. 实验结果分别与现有的自组织算法做了对比. 所有算法使用相同的训练数据和测试数据. 表1为各个算法的测试RMSE、隐含层神经元个数、训练时间的比较.

由下图的结果可知, 所提出的ASC-RBF算法性能良好, 训练误差不断下降. 在此过程中, 网络结构不断的根据迭代调整, 当网络结构为4时性能最佳, 且由图4可知, 在第7步是基本完全收敛, 表明了收敛速度快. 除此之外, 与NA-RBF^[22], FS-RBF^[22], AANN^[31], FDC-RBF^[28]进行对比, 由表1结果可知, ASC-RBF神经网络具有最紧凑的网络结构(4个隐含层神经元), 除了NARBF, 该算法在训练时间上要远远优于其他几种算法, 而且ASC-RBF神经网络跟其他算法相比预测精度相对最高.

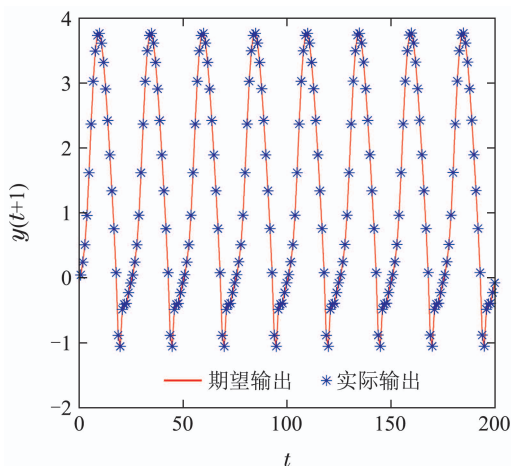


图3 非线性系统辨识结果

Fig. 3 The testing results of the nonlinear system identification

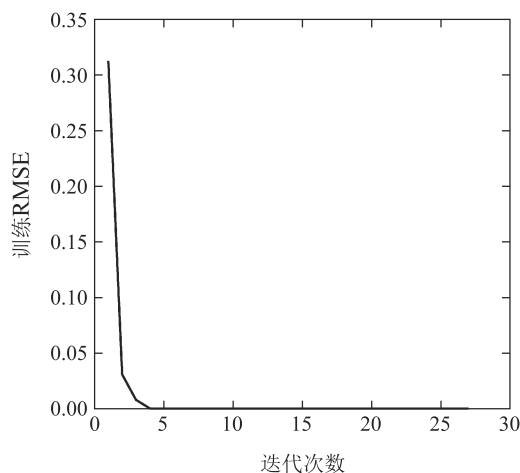


图4 非线性系统辨识训练RMSE

Fig. 4 The training RMSE of the nonlinear system identification

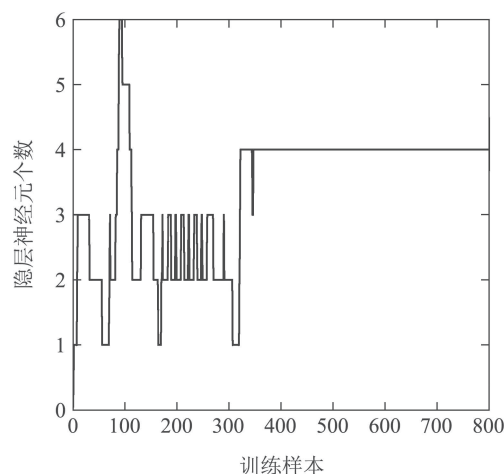


图5 隐含层神经元个数变化曲线

Fig. 5 Variation curve of the number of hidden layer neurons

表1 非线性系统辨识不同算法结果对比

Table 1 Comparison of the performance of different algorithms on nonlinear system identification

算法	训练时间/s	测试RMSE	隐节点个数
ASC-RBF	0.0621	0.0192	4
NARBF ^[22]	0.5702	0.029	6
FDC-RBF ^[28]	15.26	0.0979	10
FS-RBF ^[22]	13.325	0.031	12
AANN ^[30]	13.31	0.0424	15

由此可以得出, 在用ASC-RBF神经网络进行非线性动态系统辨识时, 在保证相当测试精度和收敛速度的前提下, 结构更为紧凑.

5.2 Mackey-Glass时间序列预测

为检验该网络模型对混沌系统预测的能力, 选取典型混沌时间序列Mackey-Glass作为仿真实例进行预测研究. Mackey-Glass微分方程被认为是基准时间

序列问题之一,它是由以下时滞微分方程产生的:

$$\frac{dy(t)}{dt} = \frac{ay(t-\tau)}{1+y^{10}(t-\tau)} - by(t).$$

(34)

实验时,方程中的参数设置: $a = 0.2, b = 0.1, \tau = 17$,采用四阶-龙库塔方法产生4500个实验数据,用4000个样本对神经网络进行训练($t = 1, 2, \cdots, 4000$),用500个样本对神经网络进行测试($t = 4001, \cdots, 4500$).输入向量为 $[y(t) \ y(t-6) \ y(t-12) \ y(t-18)]$,输出参考对比算法预测50步 $[y(t+50)]$ 的值.激活活性值阈值 $\theta_0 = 0.5$.

图6-8分别表示了ASC-RBF网络的Mackey-Glass时间序列预测结果、训练RMSE以及隐含层神经元个数变化曲线.此外将实验结果分别与SAS-RBF^[22], NARBF^[21], AI-RBF^[28], GGAP-RBF^[16], RAN^[16]等算法做了对比.所有算法都使用相同的训练数据和测试数据.对比结果如表2所示.

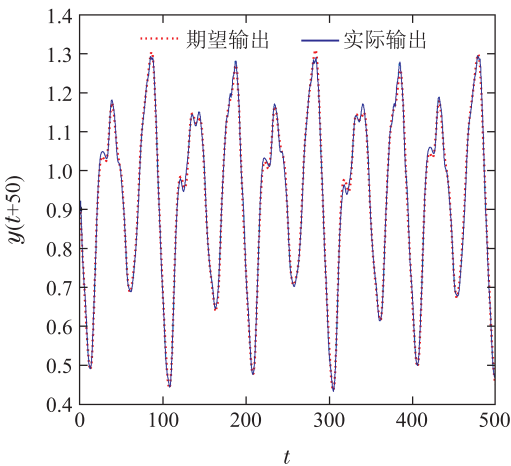


图 6 Mackey-Glass时间序列预测结果
Fig. 6 The testing results of Mackey-Glass time series

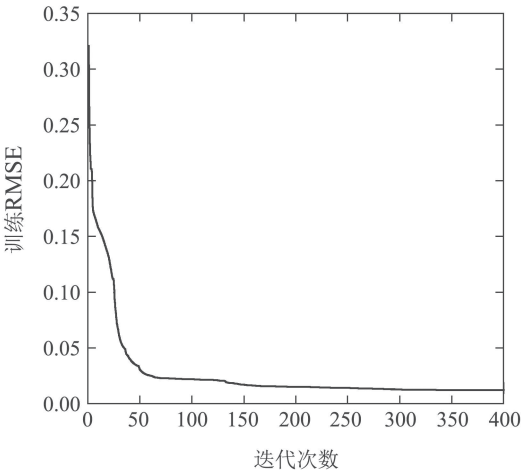


图 7 Mackey-Glass时间序列预测网络训练RMSE
Fig. 7 The training RMSE of Mackey-Glass time series prediction

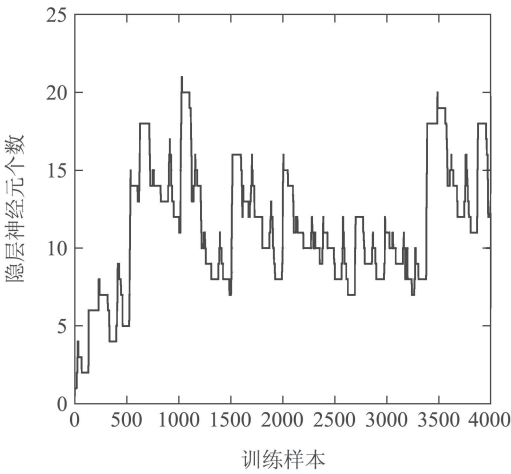


图 8 隐层神经元个数变化曲线
Fig. 8 Variation curve of the number of hidden layer neurons

表 2 Mackey-Glass时间序列预测不同算法结果对比
Table 2 Comparison of the performance of different algorithms on Mackey-Glass time series prediction

算法	训练时间/s	测试RMSE	隐节点个数
ASCRBF	9.319	0.0127	13
SAS-RBF ^[22]	10.15	0.0117	129
AI-RBF ^[28]	892.9	0.151	11
NARBF ^[21]	14.705	0.0131	16
GGAP-RBF ^[16]	28.868	0.0312	19
RAN ^[16]	58.127	0.0466	39

由表2可知,ASC-RBF和NARBF两种算法得到的网络性能具有可比性,但是ASC-RBF算法比NARBF算法的在训练时间和网络规模更有优势.与另外几种算法相比,虽然SAS-RBF算法的网络性能和训练时间和本文算法相当,但是网络规模结构过大,而AI-RBF算法训练网络太过耗时.此外,与GGAP-RBF相比,ASC-RBF算法实现了用最短训练时间,最简结构获得最佳的网络性能.

5.3 污水处理过程中出水BOD浓度预测

为了控制和优化污水处理工艺,需要对污水处理过程中的许多质量参数进行测量和控制.虽然这些参数可以通过实验室分析来测量,但是需要的测量时间长(从几分钟到几天),影响出水水质运行的有效性.因此,设计一个有效的污水处理过程水质预测模型是十分必要的.

本节利用所提出的ASC-RBF神经网络建立模型,对污水处理过程中的出水生化需氧量(biochemical oxygen demand, BOD)浓度进行准确、快速的预测.实验数据来自北京市某污水处理厂,一共有365组数据,其中265组数据用来训练网络,其余的100组数据用来测试网络性能.分别选取进水pH、出水pH、进水固体

悬浮物浓度(suspend solids, SS)、出水SS、进水BOD浓度、进水化学需氧量(chemical oxygen demand, COD)浓度、出水COD浓度、生化池污泥沉降比(settling velocity, SV)、生化池混合液悬浮固体浓度(mixed liquid suspended solids, MLSS)、生化池溶解氧浓度(dissolved oxygen, DO)等10个变量作为输入。出水BOD浓度作为输出。

因为实际的污水厂受天气、环境等因素影响, 各变量之间存在很大的差异, 所以需要先对输入的10个变量进行归一化处理。将输入变量归一化至 $[-1, 1]$, 输出变量归一化至 $[0, 1]$, 训练结束后, 网络输出进行反归一化得到出水BOD实际预测浓度。

图9为出水BOD浓度预测结果。图10为出水BOD浓度预测误差, 由误差图可以看出ASC-RBF网络对出水BOD浓度的预测误差波动范围在 $[-1, 1]$ 之间, 说明网络对测试样本进行了比较精确的预测。

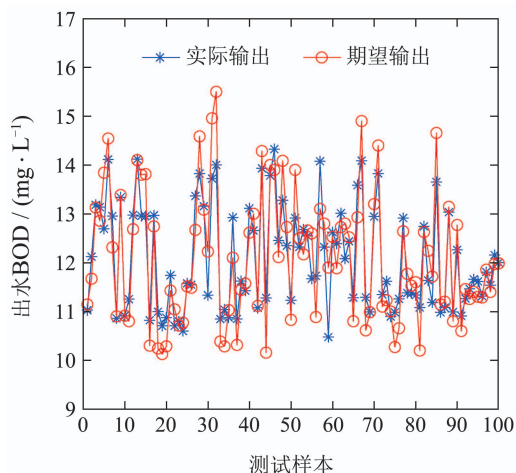


图9 出水BOD浓度预测结果

Fig. 9 The testing results of the effluent BOD concentration

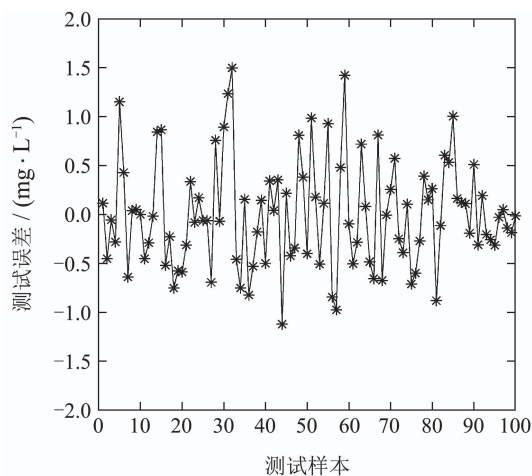


图10 出水BOD浓度预测误差

Fig. 10 The testing error of the effluent BOD concentration prediction

为了进一步评价本文所提的ASC-RBF网络的性能, 将实验结果分别与NARBF^[21], GGAP-RBF^[31],

FS-RBF^[31]和ES-RBF^[32]等自组织算法做了对比。所对比算法的文献中的输入量是基于人工经验选取的5个特征参量, 为了便于对比, 利用本文提出的算法, 同样以这5个特征参量作为输入, 来预测出水BOD浓度的值, 得到的对比结果如表3所示。

表3 出水BOD浓度预测不同算法结果对比

Table 3 Comparison of the performance of different algorithms on the effluent BOD concentration prediction

算法	训练时间/s	测试MSE	隐节点个数
ASC-RBF	0.1486	0.0962	4
NA-RBF ^[21]	0.3275	0.0977	6
ES-RBF ^[32]	4.82	0.2886	5
GGAP-RBF ^[31]	73.59	0.3325	11
FS-RBF ^[31]	63.24	0.3015	11

由对比结果可知, 虽然ASC-RBF和NARBF的测试误差相当, 但是前者比后者的网络结构更为紧凑, 而对比ES-RBF, GGAP-RBF, FS-RBF, 本文提出的ASC-RBF网络训练时间更短, 同时具有更紧凑的结构和较高的预测精度, 可以较为精准的预测出污水处理过程中水质参数。

5.4 讨论

本文的自组织机制是基于神经元特性设计的, 利用神经元激活活性的设计神经元增加机制, 随着神经元的增加在一定范围内, 会提高网络的预测精度; 通过神经元和样本的显著性判断, 执行神经元替换机制, 确保神经元的高显著性特性, 在保证规模不变的情况下提升网络性能; 通过计算神经元之间的相关性, 执行合并机制, 在不影响网络性能的前提下, 精简网络结构。由以上实验结果可知, 本文所提出的网络具有最佳的网络结构, 且具有较好的网络性能。

6 结论

为了提高RBF神经网络的精确性和经济性, 本文基于神经元相关特性提出了一种新的结构自组织的ASC-RBF神经网络, 实验仿真结果表明了该网络的有效性, 该网络的优点主要有以下方面: 1) ASC-RBF神经网络结构紧凑。首先根据能否激活神经元来判断是否增加新的神经元, 这样确保每个样本都可以被学习, 其次不仅仅检查样本能否激活神经元, 而且检查激活几个神经元, 然后判断显著性和相关性进行神经元的合并和替换, 只有显著性高的神经元才会被选为隐含层神经元, 使得整个网络的结构紧凑; 2) 神经网络结构的稳定性是保证网络实际应用的关键因素, 本文讨论了网络的自组织增加、替换、合并机制, 并给出了稳定性证明; 3) 根据两个基准仿真实验和一个实际应用预测实验的实验结果以及和现有的一些自组织网络模型的对比结果可知, 本文所提出的ASC-

RBF神经网络在保证网络紧凑结构的前提下可以获得较好的泛化性能与较高的收敛速度。

但是所提的算法也有一些不足需要改进,比如算法关键参数的确定,如何可以根据样本和期望的误差自适应的确定关键参数的值,这是以后要考虑的问题。

参考文献:

- [1] XI Y H, PENG H, MO H. Parameter estimation of RBF-AR model based on the EM-EKF algorithm. *Acta Automatica Sinica*, 2017, 43(9): 1636 – 1643.
- [2] CARLOS B P, WEI H L, ADRIAN R S. Generalized multiscale RBF networks and the DCT for breast cancer detection. *International Journal of Automation and Computing*, 2020, 17(1): 55 – 70.
- [3] LI Junzhi, WU Haifeng, ZENG Yu, et al. Recovery of collided RFID tags with frequency drift on physical layer. *Acta Electronica Sinica*, 2018, 46(8): 2004 – 2010.
(李俊志, 吴海峰, 曾玉, 等. 频率漂移的RFID冲突标签的物理层恢复. 电子学报, 2018, 46(8): 2004 – 2010.)
- [4] ZHOU W, LI X L, YI J, et al. A novel UKF-RBF method based on adaptive noise factor for fault diagnosis in pumping unit. *IEEE Transactions on Industrial Informatics*, 2019, 34(6): 1415 – 1424.
- [5] WU Yuxiang, ZHANG Jing, WANG Cong. Fault diagnosis of cracked rotor systems based on radial basis function neural networks. *Control Theory & Applications*, 2014, 31(8): 1061 – 1068.
(吴玉香, 张景, 王聪. 基于径向基函数神经网络的转子系统裂纹故障诊断. 控制理论与应用, 2014, 31(8): 1061 – 1068.)
- [6] FAN Jiahua, MA Lei, ZHOU Pan, et al. Modeling and control of piezoelectric actuator based on radial basis function neural network. *Control Theory & Applications*, 2016, 33(7): 856 – 862.
(范家华, 马磊, 周攀, 等. 基于径向基神经网络的压电作动器建模与控制. 控制理论与应用, 2016, 33(7): 856 – 862.)
- [7] DUAN Xiaochen, XU Jia. Research on estimating models of whole life cost for road engineering based on SOM-RBF neural network. *Journal of the China Railway Society*, 2020, 42(1): 9 – 14.
(段晓晨, 徐佳. 道路工程SOM-RBF神经网络估价模型研究. 铁道学报, 2020, 42(1): 9 – 14.)
- [8] LIAO Liefu, YANG Ligu. Adaptive radial basis function neural network bi-quadratic functional optimal control for manipulators. *Control Theory & Applications*, 2020, 37(1): 47 – 58.
(廖列法, 杨翌隼. 机械臂的自适应径向基函数神经网络双二次泛函最优控制. 控制理论与应用, 2020, 37(1): 47 – 58.)
- [9] ZHAO Haibo, WANG Chengguang. Radial-basis-function neural network backstepping adaptive control of dual-motor driving servo system. *Control Theory & Applications*, 2018, 35(9): 1272 – 1284.
(赵海波, 王承光. 双电机驱动伺服系统径向基函数神经网络反推自适应控制. 控制理论与应用, 2018, 35(9): 1272 – 1284.)
- [10] KOSIC, DINO. Fast clustered radial basis function network as an adaptive predictive controller. *Neural Networks*, 2015, 63(2015): 79 – 86.
- [11] BORS A G, PITAS I. Median radial basis function neural network. *IEEE Transactions on Neural Networks*, 1996, 7(6): 1351 – 1364.
- [12] YIN H J, ALLINSON N M. Self-organizing mixture networks for probability density estimation. *IEEE Transactions on Neural Networks*, 2001, 12(2): 405 – 411.
- [13] PLATT J. A resource-allocating network for function interpolation. *Neural Computation*, 1991, 3(2): 213 – 225.
- [14] HONG X, BILLINGS S A. Givens rotation based fast backward elimination algorithm for RBF neural network pruning. *IEEE Proceedings—Control Theory and Applications*, 1997, 144(5): 381 – 384.
- [15] LU Y W, SUNDARARAJAN N, SARATCHANDRAN P. A sequential learning scheme for function approximation using minimal radial basis function neural networks. *Neural Computation*, 1997, 9(2): 461 – 478.
- [16] HUANG G B, SARATCHANDRAN P, SUNDARARAJAN N. A generalized growing and pruning RBF (GGAP-RBF) neural network for function approximation. *IEEE Transactions on Neural Networks*, 2005, 16(1): 57 – 67.
- [17] HUANG G B, SARATCHANDRAN P, SUNDARARAJAN N. An efficient sequential learning algorithm for growing and pruning RBF (GAP-RBF) networks. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 2004, 34(6): 2284 – 2292.
- [18] DU Dajun, FEI Minrui, LI Lixiong. Radial-basis-function neural network based on fast recursive algorithm and its application. *Control Theory & Applications*, 2008, 25(5): 827 – 830.
(杜大军, 费敏锐, 李力雄. 基于快速回归算法的RBF神经网络及其应用. 控制理论与应用, 2008, 25(5): 827 – 830.)
- [19] HAN H G, CHEN Q L, QIAO J F. Research on an online self-organizing radial basis function neural network. *Neural Computing & Applications*, 2010, 19(5): 667 – 676.
- [20] HAO Y, REINER P D, XIE T T. An incremental design of radial basis function networks. *IEEE Transactions on Neural Networks and Learning Systems*, 2014, 25(10): 1793 – 1803.
- [21] QIAO J F, MENG X, LI W J. An incremental neuronal-activity-based RBF neural network for nonlinear system modeling. *Neurocomputing*, 2018, 302(2018): 1 – 11.
- [22] XIE Y F, JIN J, XIE S W. On-line prediction of ferrous ion concentration in goethite process based on self-adjusting structure RBF neural network. *Neural Networks*, 2019, 116(2019): 1 – 10.
- [23] HAN H G, LU W, HOU Y, et al. An adaptive-PSO-based self-organizing RBF neural network. *IEEE Transactions on Neural Networks and Learning Systems*, 2018, 29(1): 104 – 117.
- [24] SHAO Guangcheng, ZHANG Kun, WANG Zhiyu, et al. Groundwater depth prediction model based on IABC-RBF neural network. *Journal of Zhejiang University (Engineering Science)*, 2019, 53(7): 1323 – 1330.
(邵光成, 章坤, 王志宇, 等. 基于IABC-RBF神经网络的地下水埋深预测模型. 浙江大学学报(工学版), 2019, 53(7): 1323 – 1330.)
- [25] ALJARAHI I, FARIS H, MIRJALILI S, et al. Training radial basis function networks using biogeography-based optimizer. *Neural Computer & Applications*, 2018, 29(2018): 529 – 553.
- [26] ISLAM M, SATTAR A, AMIN F, et al. A new adaptive merging and growing algorithm for designing artificial neural networks. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 2009, 39(3): 705 – 722.
- [27] MENG Xi, QIAO Junfei, LI Wenjing. Construction of RBF neural networks via fast density clustering. *Caai Transactions on Intelligent Systems*, 2018, 13(3): 331 – 338.
(蒙西, 乔俊飞, 李文静. 基于快速密度聚类的RBF神经网络设计. 智能系统学报, 2018, 13(3): 331 – 338.)
- [28] LI Fei, YANG Cuili, QIAO Junfei. A novel RBF neural network design based on immune algorithm system. *The 36th Chinese Control Conference (CCC)*. Dalian, Liaoning: IEEE, 2017: 4598 – 4603.
- [29] XIE T T, HAO Y, HEALETT J, et al. Fast and efficient second-order method for training radial basis function networks. *IEEE Transactions on Neural Networks and Learning Systems*, 2012, 23(4): 60 – 619.
- [30] HAN H G, WANG L D, QIAO J F. Efficient self-organizing multilayer neural network for nonlinear system modeling. *Neural Networks*, 2013, 43(2013): 22 – 32.
- [31] HAN H G, CHEN Q L, QIAO J F. An efficient self-organizing RBF neural network for water quality prediction. *Neural Networks*, 2011, 24(7): 717 – 725.
- [32] LI Wenjing, LI Meng, QIAO Junfei. Effluent BOD soft measurement based on mutual information and self-organizing RBF neural network. *CIESC Journal*, 2019, 70(2): 687 – 695.
(李文静, 李萌, 乔俊飞. 基于互信息和自组织RBF神经网络的出水BOD软测量方法. 化工学报, 2019, 70(2): 687 – 695.)

作者简介:

贾丽杰 博士研究生, 目前研究方向为神经网络分析建模, E-mail: jialijie@126.com;

李文静 博士, 副教授, 目前研究方向为人工神经网络、模式识别, E-mail: wenjing.li@bjut.edu.cn;

乔俊飞 教授, 博士生导师, 目前研究方向为污水处理过程知识自动化, E-mail: junfeiq@bjut.edu.cn.