

文章编号: 2096-1472(2016)-09-11-03

基于遗传算法优化的RBF神经网络在MBR膜污染仿真预测中的研究

汤 佳, 李春青

(天津工业大学计算机科学与软件学院, 天津 300387)

摘 要: 膜污染问题是影响MBR推广应用的主要障碍。针对该问题, 首先运用主成分分析法确定影响膜污染的主要因素为MLSS、操作压力及温度, 然后建立基于RBF神经网络的预测模型, 再利用遗传算法对预测模型的四个参数进行优化, 使得整个网络模型达到全局最优。最后将预测结果与样本数据对比, 结果显示, 基于GA-RBF的膜污染模拟仿真器在收敛速度、预测精度等方面比单纯的RBF网络有较大提高, 达到了预期目标。

关键词: MBR; 膜通量; RBF神经网络; 遗传算法

中图分类号: TP389.1 **文献标识码:** A

Research of RBF Neural Network Based on Genetic Algorithm Optimization in MBR Membrane Pollution Simulation

TANG Jia, LI Chunqing

(School of Computer Science and Software, Tianjin University of Technology, Tianjin 300387, China)

Abstract: Membrane fouling is the main obstacle to the promotion and application of MBR. To solve this problem, the main factor is the first use of principal component analysis to determine the influence of membrane fouling is MLSS, operating pressure and temperature, and then build the prediction model based on RBF neural network, and then use genetic algorithm four parameters of the prediction model is optimized, which makes the whole network model to achieve global optimum. The predicted results compared with the sample data shows that the membrane fouling of GA-RBF simulator in terms of the convergence speed and prediction accuracy is higher than that of simple RBF neural network based on, to achieve the expected goal.

Keywords: MBR; membrane flux; RBF neural network; genetic algorithm

1 引言(Introduction)

水污染的预防和治理已成为当今世界日益严重的热点问题。膜生物反应器(MBR)作为一种将膜分离技术和生物反应技术相结合的新型高效污水处理技术, 已被广泛应用于生活污水、有机废水处理等领域^[1,2]。膜污染会影响MBR膜分离单元的性能, 缩短膜的使用寿命, 增加MBR系统的运行成本。因此研究膜污染的机理和预测控制方法, 保证膜生物反应器能够在稳定的低耗能条件下获得较大膜通量是促进MBR推广应用的关键^[3,4]。

目前在MBR膜污染领域, 常用的多种预测模型都存在一些缺陷, 如对膜污染机理分析不够透彻, 预测精度差等问题。人工神经网络在预测领域有比较成熟的应用, 而相比较传统的BP神经网络, RBF网络具有唯一最佳逼近的特性且无局部极小值的问题, 可以做更精准高效的预测分析^[5-7]。再引入遗传算法优化RBF神经网络的相关参数, 弥补单一径向基网络模型的缺陷, 就可以实现对膜污染更精准的预测。

2 RBF神经网络(RBF neural network)

人工神经网络(Artificial Neural Network, 即ANN), 是由人工建立的, 以有向图为拓扑结构的动态系统, 它从信息处理角度对人脑神经单元网络进行抽象化, 建立简单模

型, 按不同的连接方式组成不同网络。人工神经网络以其特有的非线性适应性信息处理能力, 使之在模式识别、自动控制、人工智能、预测估计等方面都有成功的应用, 表现出了良好的智能特性。

径向基函数(RBF)神经网络是一种高效的前馈式神经网络, 它由输入层、隐含层和输出层构成, 具有其他前向网络所不具有的最佳逼近性能, 并且结构简单, 训练速度快。该神经网络采用RBF作为隐单元的“基”构成隐含层空间, 将输入矢量直接映射到隐空间而不需要权值连接。当RBF的中心点确定后, 映射关系也就随之确定。隐含层空间到输出空间的映射是线性的。

径向基神经网络的网络结构如图1所示。

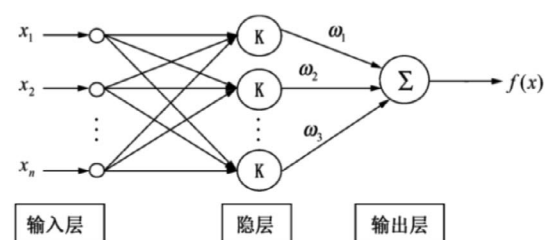


图1 RBF神经网络结构图

Fig.1 RBF neural network structure

RBF网络的激活函数采用径向基函数,以输入和权值向量之间的距离作为自变量,常见的隐含层节点的核函数为高斯函数。

$$\ell(\gamma) = \exp\left(-\frac{\gamma^2}{2\sigma^2}\right) \quad (1)$$

其中, σ 称为基函数的扩展常数或宽度, σ 越小, 径向基函数的宽度越小, 基函数就越有选择性。

3 遗传算法(Genetic algorithm)

遗传算法(Genetic Algorithms, 简称GA或GAs)是由密歇根大学Joho H. Holland教授及其学生于20世纪60年代末到70年代初提出的一种经典的智能算法。遗传算法是一种基于群体寻优的方法, 具有全局搜索能力, 初值无关性以及较快的收敛速度。具体过程如下:

Step1: 采用随机的方法或者其他方法产生一个初始种群。

Step2: 根据问题的目标函数构造适值函数(Fitness Function)适值函数用来表征种群中每个个体对其生存环境的适应能力。

Step3: 根据适应值的好坏, 不断选择和繁殖, 基因通过交叉和变异得到更新。

Step4: 若干代后得到适应值最好的个体即为最优解。

4 遗传算法优化的RBF神经网络(Optimization of RBF neural network based on genetic algorithm)

RBF神经网络具有最佳逼近性和全局最优性, 可以实现对MBR膜通量较为精确的预测。但是为了更好地选取相关参数, 引入遗传算法对RBF神经网络的参数进行合理优化, 其中包含四个参数, 分别是误差、训练速度、最大神经元数目以及神经元的间隔, 经过遗传算法寻优后得到的最优个体即为进行RBF神经网络训练涉及的最优参数。GA-RBF神经网络模型完成网络训练后, 再采用仿真数据检测评估该网络的优化性能。

5 建立基于PCA的GA-RBF膜污染仿真预测模型(A simulation model of GA-RBF membrane pollution based on PCA is established.)

5.1 采用PCA进行模型输入参数选取

主成分分析(Principal Component Analysis, PCA)是一种掌握事物主要矛盾的统计分析方法, 它通过矩阵特征值分析对初始数据进行线性空间投影, 从多元事物中解析出主要影响因素, 从而达到简化复杂的目的。

膜污染的影响因素众多, 通过减少RBF网络的输入数来简化RBF网络结构, 从整体上提高RBF网络的性能。初选的六个膜污染影响因子为 $X = \{\text{总阻力, MLSS, 操作压力, COD, PH, 温度}\}$ 。

主成分分析法步骤为:

Step1: 把初选的六个膜污染影响因子组成矩阵 X 。

Step2: 对 X 数据标准化处理, 得到标准化矩阵 A 。标准化后的数据阵每个列向量的均值为0, 标准差为1, 数据无量纲。

Step3: 计算出协方差矩阵 S 。

Step4: 计算协方差矩阵 S 的特征值 λ 与相应的特征向量 u , 然后将特征值由大到小排列得到特征值矩阵 V 和特征向量矩阵 U 。

Step5: 对特征值矩阵 V 分析, 计算出贡献率, 确定主成分。通常选取贡献率为85%—95%的 m 个主成分进行综合分析。

通过计算选取主成分贡献率在90%以上的三个污染因子即: MLSS、操作压力及温度, 并以此作为RBF神经网络的输入层神经元。膜通量大小是表征膜污染程度的唯一指标, 因此以膜通量作为RBF神经网络的唯一的输出层神经元。

5.2 建立基于GA-RBF的膜污染仿真预测模型

实验建立的GA-RBF膜污染预测模型如图2所示。



图2 膜污染仿真预测模型

Fig.2 Simulation model of membrane fouling

实验步骤如下:

Step1: 对RBF神经网络的四个参数编码, 为方便操作设定误差、训练速度为实数编码, 最大神经元数目、神经元间隔为整数编码, 所有编码得到的基因位串称为一个个体。

Step2: 随机产生20个个体作为初始种群, 构建初始的交配池。

Step3: 执行RBF网络训练, 得到误差百分比作为每个个体的适应度。个体间基因进行交叉和变异, 保证基因多样性。

Step4: 采用轮盘赌选择方式, 筛选出适应度高的个体, 淘汰掉适应度低的个体。

Step5: 重复进行(3)直到结束循环, 得到的个体为最佳个体。

Step6: 训练完毕, 最佳个体即为RBF网路最优的参数值, 进行仿真测试, 将实验数据与样本数据对比分析。

6 预测结果与实验结果对比分析(Comparison of the predicted results with the experimental results)

实验数据全部来自于石家庄市某MBR污水处理厂的实验以及工业生产的历史数据, 统一采用的是孔径为0.2um的聚偏氟乙烯微滤膜处理污水方法。从实验数据中选取6组作为校验用样本, 其余24组作为神经网络的训练样本。

遗传算法操作的初始参数为: 选择种群规模:

sizepop=20；最大进化代数：maxgen=10；交叉率：pcross=0.8；变异率：pmutation=0.5.RBF训练误差范围是1e-4-1e-3；最大神经元数目范围值是80—150；显示神经元间隔范围值是10—30；速度范围值是10—25。

试验中得到遗传算法代数与误差曲线如图3所示。

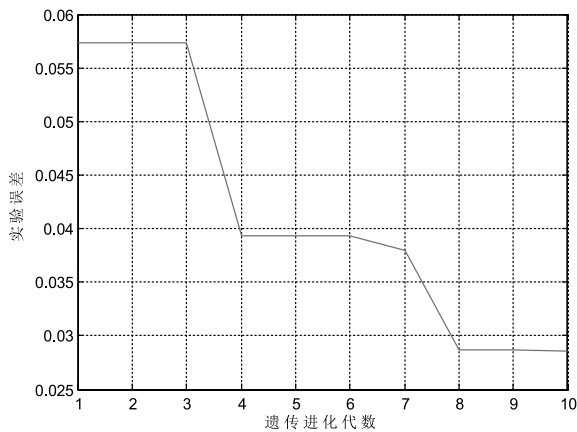


图3 实验误差分析图

Fig.3 Experimental error analysis

图3显示了网络训练的误差随着遗传代数增加表现出来的变化趋势，随着进化代数增多误差降低，进化至第8代误差仅为0.02871，到第8代以后图像趋于缓和，基本已达到优化极限,实验收敛速度很快，优化效果明显。

为更好地体现优化效果,实验使用相同的样本数据进行测试，设置遗传算法进化到10代，分别对RBF网络和GA-RBF网络进行训练。首先得到基于RBF神经网络的预测结果，如图4所示，RBF网络的训练结果平均误差值为0.1130，已经能够完成初步的膜通量预测。然后再进行GA-RBF网络训练，预测结果对比如图5所示，经过GA优化参数之后的预测模型平均误差值仅为0.0275.准确度有非常明显的提高。实验结果表明，经过遗传优化过的RBF神经网络误差明显减小，得到的训练结果明显优于优化前的结果，预测精度较为理想。预测误差数据分析见表1。

表1 试验预测误差分析

Tab.1 Test prediction error analysis

样本	期望输出	RBF预测值	RBF相对误差	GA-RBF预测值	GA-RBF相对误差
1	45.1000	44.2654	-0.0185	42.5184	-0.0572
2	28.9000	36.4103	0.2599	28.7122	-0.0065
3	9.4000	11.1028	0.1811	9.4673	0.0071
4	43.4000	43.3223	-0.0018	42.9908	-0.0094
5	28.9000	35.3481	0.2231	28.6398	-0.0090
6	8.4000	10.2058	0.2150	7.3956	-0.1196

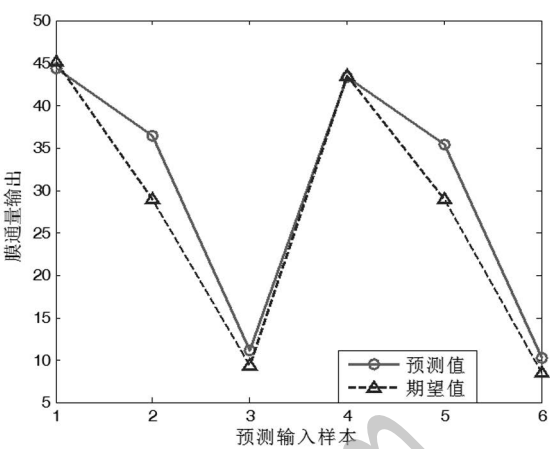


图4 RBF神经网络预测图

Fig.4 RBF neural network prediction map

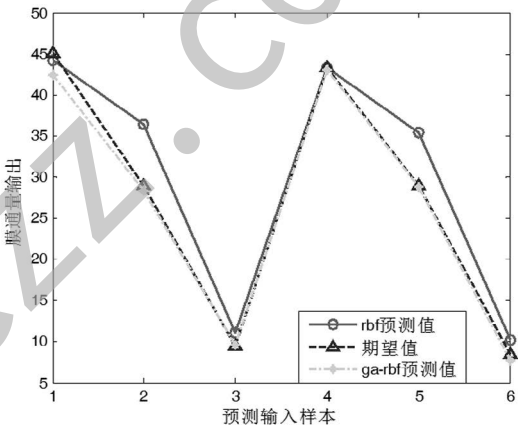


图5 试验预测结果对比图

Fig.5 Comparison of test results

7 结论(Conclusion)

由于膜污染过程具有非线性、参数时变性、随机干扰等复杂的机理，造成预测研究过程十分困难，本文在用主成分分析法简化输入参数的基础上，建立了RBF神经网络仿真模型，实验取得了良好的预测效果，误差符合实验标准。由于预测精度是否准确很大程度上取决于网络模型参数的选取，因此引入遗传算法，优化RBF网络相关参数，建立基于遗传算法优化的径向基函数神经网络预测模型并成功应用于MBR膜污染仿真预测中，研究结果表明，遗传算法优化后的RBF神经网络不仅提高了预测速度,还明显提高了预测精确度,从而能更好地实验对膜通量的检测和控制。整个实验过程具有一定的理论价值和实践意义，对MBR实际工程应该可以起到积极的指导作用。

参考文献(References)

[1] Yu Zhang,et al.Risk Assessment of Giardia from a Full Scale MBR Sewage Treatment Plant Caused by Membrane Integrity Failure[J].Journal of Environmental Sciences,2015(04):252-258.
[2] Kun Li,et al.Advanced Treatment of Municipal Wastewater