

# 基于遗传算法与 BP 算法的水质评价模型

蚩志锋<sup>1,2</sup> 闫珍珠<sup>1</sup> 黄彪<sup>2</sup>

(1.信阳师范学院, 信阳 464000; 2.华中师范大学, 武汉 430079)

**摘要:**BP 神经网络广泛应用于函数逼近、模式分类、水质评价等方面,但标准 BP 算法收敛速度较慢,容易陷入局部最优;而遗传算法是一种全局寻优搜索算法,能够有效克服上述缺陷。提出了基于遗传算法与 BP 算法的混合算法,既保留了神经网络原有的优点,又克服了上述缺点,并建立了水质评价模型。以信阳南湾水库为例进行评价,实验结果表明该混合算法模型评价精度较高,完全可以应用于水质评价工作。

**关键词:**遗传算法;BP 算法;水质评价;局部最优

**中图分类号:**TP301

**文献标识码:**A

**文章编号:**1673-1980(2009)01-0122-03

人工神经网络是近年迅速发展起来的人工智能科学的一个分支,它具有自学习、自组织、容错性好等特性,受到众多领域学者的关注。人工神经网络除了应用在信息处理、自动控制等领域外,还逐渐拓展到工程应用领域,如水质评价等方面,其准确性优于灰色聚类法、模糊综合评判法<sup>[1]</sup>。

目前,人工神经网络在水质分析和评价中已逐步得到应用,但现有评价模型大多采用标准 BP 算法,由于其算法本身的缺陷,如迭代次数多,容易陷入局部最优等,因此评价效果有待改进。本文在分析遗传算法<sup>[2]</sup>和 BP 算法的基础上,提出了一种混合算法,建立了新的水质评价模型,在提高算法收敛速度的同时,保持算法结果的全局最优。系统仿真实验以信阳南湾水库的水质监测结果为例进行评价,实验结果表明该混合算法模型收敛速度快,评价结果与实际监测结果基本一致,效果较好。

## 1 BP 网络算法和遗传算法原理

标准神经网络模型是由一个输入层、一个输出层和至少一个隐含层组成,一个标准的 BP 学习算法是:

(1)计算网络各层节点的输出。

$$y_i = f\left(\sum_{j=1}^{L-1} w_{ij} x_j - \theta_j\right) \quad i=1, 2, \dots, L-1 \quad (1)$$

式中:非线性函数采用  $f(a) = 1/(1+e^{-a})$ ;  $w_{ij}$ —连接权值;  $\theta_j$ —节点阈值;  $L$  为输入层、输出层单元节点个数。

(2)计算网络误差。

$$\delta_i = (t_i - y_i) \times y_i \times (1 - y_i) \quad (2)$$

式中:  $t_i$ —期望误差;  $\delta_i$ —网络误差。

(3)将网络误差与最大允许误差进行比较,若满足,则结束,否则按下面公式修正权值后,执行步骤(1)。

$$w_{ij}(k+1) = w_{ij}(k) + \eta \delta_i x_j \quad (3)$$

式中:  $w_{ij}(k+1)$ —下次循环的连接权值;  $w_{ij}(k)$ —当前的连接权值;  $\eta$ —学习率。

BP 神经网络<sup>[5-7]</sup>具有许多优秀的品质,如自适应、自组织性、容错性等,并且善于从近似的、不确定的、甚至相互矛盾的知识环境中做出评价。但仍然存在不少问题,如学习算法收敛速度较慢、容易陷入局部最优、网络运行时只是单向传播等,从而使学习结果不令人满意。

遗传算法(Genetic Algorithms,简称 GA)是一种基于自然选择<sup>[8]</sup>和群体遗传机理的搜索算法,它模拟了自然选择和自然遗传过程中发生的繁殖、杂交和突变现象。在利用遗传算法求解问题时,问题的每个可能的解都被编码成一个“染色体”,即个体,若干个个体构成了群体(所有可能解)。遗传算法是一个由可行解组成的群体逐代进化的过程。遗传算法开始时随机地产生一些个体,根据预定的目标函数对每个个体进行评价,给出了一个适应度值。基于此适应度值,选择个体用来复制下一代。选择操作体现了“适者生存”原理,“好”的个体被选择用来复制,而“坏”的个体则被淘汰。然后选择出来的个体经过交叉和变异算子进行再组合生成新一代。这群新个体继承

收稿日期:2008-11-18

基金项目:河南省软科学研究计划项目(072400421130)

作者简介:蚩志锋(1981-),男,在读硕士研究生,研究方向:智能优化。

了上一代的一些优良性状,在性能上要优于上一代。

GA 对于全局搜索具有较强的鲁棒性和较高的效率,但不适合候选解的精调,难以确定它们的确切位置。因此,文章将遗传算法和 BP 算法相结合提出了一种混合算法,实现了优势互补,能更好地获取训练样本的经验,做出更好的水质评价。

## 2 基于遗传算法与 BP 算法的混合算法

混合算法主要是在 BP 神经网络训练过程中,利用遗传算法善于发现最优解区域的特点来优化网络权重值和阈值,避免了标准 BP 算法容易陷入局部最优的缺陷,同时也减少了迭代次数。该算法具体步骤如下:

Step1 确定遗传算法及 BP 网络的有关参数,包括遗传算法的群体容量、交叉率、变异率、BP 网络的结构参数及精度等。

Step2 随机产生一组权重值和阈值分布 (染色体),用遗传算法优化权重值和阈值。

按神经网络生成初始权重值和阈值的常规办法来生成网络权重值和阈值,任一组完整的权重值和阈值相当于一个染色体,即为父代个体。染色体采用实数编码。计算个体适应度,如果个体适应度满足条件,直接跳到 Step4,否则转入 Step3。

Step3 遗传操作产生新一代个体,淘汰父代个体,转到 Step2。遗传算子选择继承、交叉、选择 3 个常用的算子。这一步就是用遗传算法优化网络权重值和阈值。

Step4 用 BP 算法进行微调,直到满足条件结束。混合算法流程图如图 1 所示。

## 3 水质评价模型

根据国家地面水环境质量标准 (国标 GB3838-2002 水质标准<sup>[3]</sup>,见表 1),算法在构建输入模式时

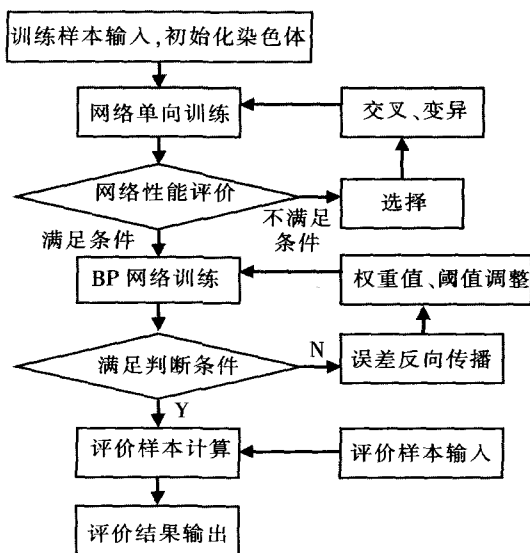


图 1 混合算法流程图

选用其中对水质影响比较大的十个评价指标来构建模型的输入矢量。在水环境质量评价模型中引入隶属度的概念,建立了隶属度网络模型,核心思想是为了解决水质评价的客观性和不确定性问题。结合水质评价来分析,就是将 I—V 类水质级别看作是论域,根据混合算法模型对各水质样本的中间结果输出值,计算出实测样本属于各级水质的隶属度值,最终输出为测试样本的水质类别。输出层确定为一个神经元,其输出为网络的隶属度值。

由于三层的 BP 网络可以模拟任何复杂的空间曲线,这里所建立的评价模型为三层结构,即输入层、单一隐含层、输出层。隐含层单元个数按柯尔莫哥洛夫定理,取值为  $2m+1$  ( $m$  为输入数据个数),所以隐含层单元个数为 21,因此网络评价模型结构为 10-21-1。

### 3.1 数据预处理

由于各种神经网络的激活函数和学习规则等不同,因此在输入之前需对输入样本作规范化处理。由

表 1 国家地面水环境质量标准

水质	溶解氧	BOD <sub>5</sub>	锰酸盐	总氮	总磷	挥发酚	氰化物	汞	砷	六价铬
I	7.5	3	2	0.2	0.02	0.002	0.005	$5 \times 10^{-5}$	0.05	0.01
II	6	3	4	0.5	0.1	0.002	0.05	$5 \times 10^{-5}$	0.05	0.05
III	5	4	6	1.0	0.2	0.005	0.02	$10 \times 10^{-5}$	0.05	0.05
IV	3	6	8	1.5	0.3	0.01	0.2	$100 \times 10^{-5}$	0.1	0.05
V	2	10	10	2.0	0.4	0.1	0.2	$100 \times 10^{-5}$	0.1	0.1

于 Sigmoid 函数在  $[0,0.1]$  和  $[0.9,1]$  区间内变化极为缓慢,容易陷入神经网络饱和区,因而将各输入量均规范化至  $[0.1,0.9]$  区间。为此,实行如下规范化转化:

$$y = 0.1 + \frac{0.8(x - x_{\min})}{x_{\max} - x_{\min}} \quad (4)$$

式中:  $x$  是所收集的一组时间序列的数据  $x_{\min}$ 、 $x_{\max}$  分别是这一组数据中的最小值和最大值;  $y$  是映射后的数据。

### 3.2 确定遗传算法与 BP 网络参数

设遗传算法的群体规模为 300,选择概率为 0.8,

表 2 2003~2005 年南湾水库水环境质量监测结果

指标	溶解氧	BOD <sub>5</sub>	锰酸盐	总氮	总磷	挥发酚	氰化物	汞	砷	六价铬
2003 年	8.34	0.98	2.35	0.521	0.015	0.001	0.002	2	0.004	0.002
2004 年	8.02	0.90	2.13	0.572	0.026	0.001	0.002	2	0.005	0.002
2005 年	8.46	2.2	2.38	0.699	0.02	0.001	0.003	2	0.005	0.002

杂交概率取为 0.8, 变异率取为 0.2。BP 网络的学习速率为 0.05, 最大训练次数为 200 次, 精度为  $10^{-4}$ 。表 2 列出了 2003~2005 年南湾水库检测结果, 实现采用 MATLAB 编程, 用表 1 和表 2 中的数据分别作为训练样本、检验样本。

仿真实验采用了信阳市南湾水库 2003~2005 年内的水质数据<sup>[4]</sup>, 选用了其中最重要的指标, 如表 2 所示。

利用训练样本对基于混合算法的 BP 神经网络进行训练, 训练结果如图 2 所示。

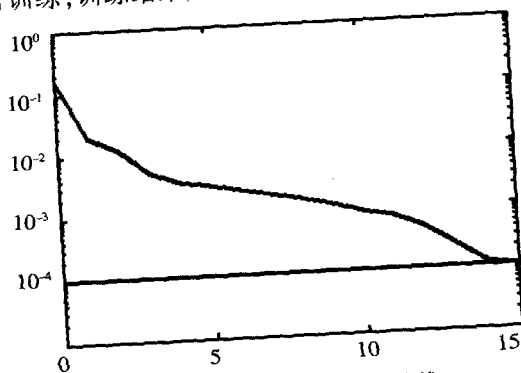


图 2 混合算法模型训练误差曲线

用表 2 中数据作为输入, 用训练后的网络对南湾水库水质状况进行评价, 其评价结果见表 3。

实验结果表明, 基于混合算法的水质评价模型所得到的评价结果与南湾水库水质情况基本一致。

#### 4 结语

介绍了 BP 网络算法和遗传算法, 然后根据遗

表 3 评价结果

年份/年	网络输出结果	目标结果	识别类别	实测类别
2003	0.446 6	0.5	Ⅲ	Ⅲ
2004	0.438 2	0.5	Ⅲ	Ⅲ
2005	0.535 8	0.5	Ⅲ	Ⅲ

传算法优化 BP 神经网络的策略, 提出了一种实数编码的混合算法, 克服了传统 BP 网络算法中的局部极小缺陷, 且训练速度有了很大提高, 并建立了水质评价模型, 通过仿真实验说明了该算法在水质评价中的实用性, 完全可以应用于实际的水库水质评价工作。

#### 参考文献

- [1] 黄胜伟. 自适应变步长 BP 神经网络在水质评价中的应用[J]. 水利学报, 2002(10): 119-123.
- [2] 肖青. 一种基于遗传算法的网络任务调度算法[J]. 计算机技术与发展, 2008, 18(8): 32-38.
- [3] GB 3838—2002. 中华人民共和国地表水环境质量标准[S].
- [4] 李自荣. 南湾水库水质状况调查及治理对策[J]. 水利渔业, 2007, 27(5): 65-66.
- [5] 徐昕, 贺汉根. 神经网络增强学习的梯度算法研究[J]. 计算机学报, 2003, 26(2): 227-233.
- [6] Anurag More, M. C. Deo. Forecasting Wind with Neural Networks[J]. Marine Structure, 2003(16): 35-49.
- [7] Eberhart R C. Neural Network PC Tools[M]. [S.L.] Academic Press, 1990.
- [8] Baker J E. Reducing Bias and Inefficiency In The Selection Algorithm [C]. Proc. ICGA2, 1987: 14-21.

## Water Quality Evaluation Model Based on Genetic Algorithm and BP Algorithm

CHI Zhi-feng<sup>1,2</sup> YAN Zhen-zhu<sup>1</sup> HUANG Biao<sup>2</sup>

(1. Xinyang Normal University, Xinyang 464000; 2. Huazhong Normal University, Wuhan 430079)

**Abstract:** BP neural network has been widely used function approximation, pattern classification, water quality evaluation. But the convergence criteria BP algorithm is slow, and vulnerable to local optimum. The genetic algorithm is a global optimization search algorithm, and can effectively overcome these shortcomings. So the hybrid algorithm based on genetic algorithm and BP algorithm, not only retains the original merits of the neural network, but also overcomes these shortcomings, and establishes water quality evaluation model. Xinyang Nanwan reservoir as an example to evaluate, the experimental results show that the hybrid algorithm model has evaluation of high precision, and can be applied to water quality evaluation.

**Key words:** genetic algorithm; BP algorithm; water quality evaluation; local optimal