

# 基于改进神经网络算法的 PM<sub>2.5</sub> 污染信号分析检测

朱 洪, 蹇红梅, 刘小芳

(四川理工学院, 四川 自贡 643000)

**摘要:** 当前为了保证污染信号分析的精度, 在对 PM<sub>2.5</sub> 污染进行检测的过程中, 需处理的数据量过大, 导致经典神经网络方法遇到矛盾数据时, 需要花费大量的数据校验时间, 收敛速度下降, 检测效率大幅降低, 提出一种基于改进神经网络算法的 PM<sub>2.5</sub> 污染检测方法, 在分析标准神经网络算法的基础上, 允许信号跳变精确度范围内, 在层与层之间引入容错性变量, 同时在计算阈值的过程中融入松弛变量, 提高收敛速度; 避免神经网络陷入局部最优解; 采用改进神经网络算法, 通过不断调整网络的权值以及污染阈值, 对 PM<sub>2.5</sub> 污染信号进行高效检测; 以飞利浦公司的新一代检测系统为测试器材, 测试结果表明, 采用所提方法得到的 PM<sub>2.5</sub> 污染检测效率明显提高。

**关键词:** 改进神经网络算法; 污染检测; 网络误差

## PM<sub>2.5</sub> Pollution Detection Signal Analysis Based on Improved Neural Network Algorithm

Zhu Hong, Jian Hongmei, Liu Xiaofang

(Sichuan University of Science & Engineering, Zigong 643000, China)

**Abstract:** Currently in order to ensure the accuracy of pollution of the signal analysis, in the testing process of PM<sub>2.5</sub> pollution, need to deal with the amount of data is too large, lead to classical neural network method when you meet the contradiction between data need to spend a large amount of data checking time, convergence rate fell, the detection efficiency is greatly reduced, put forward a kind of PM<sub>2.5</sub> pollution detection method based on improved neural network algorithm, based on the analysis of the standard neural network algorithm, and allow the signal jump range, precision in the introduction of fault tolerance between layer and layer variables, at the same time in the process of calculation threshold into the slack variables, improve convergence rate; Avoid neural network into a local optimal solution. With the improved neural network algorithm, and through continuous adjust the network weights and threshold, the pollution of PM<sub>2.5</sub> pollution signal detection efficiently. In the company of a new generation of test system for test equipment, test results show that the proposed method of PM<sub>2.5</sub> pollution detection efficiency has been improved significantly.

**Keywords:** neural network algorithm; PM<sub>2.5</sub> pollution detection; network error

## 0 引言

当前, 随着社会经济的不断发展, 产生的 PM<sub>2.5</sub> 空气污染问题也逐渐增加, 严重影响了人们的日常生活<sup>[1-3]</sup>。因此, 对空气中的 PM<sub>2.5</sub> 污染有效检测是当前环保学者研究的重点方向, 受到越来越广泛的关注<sup>[4-7]</sup>。在对 PM<sub>2.5</sub> 污染进行检测的过程中, 容易出现需处理数据量大, 造成 PM<sub>2.5</sub> 污染区域相关信息数据处理量变大, 经典神经网络方法由于无法处理大量复杂数据, 不能有效准确的完成 PM<sub>2.5</sub> 污染检测<sup>[8-10]</sup>。提

出了一种基于改进神经网络算法的 PM<sub>2.5</sub> 污染检测方法, 依据 PM<sub>2.5</sub> 污染检测原理, 在分析标准神经网络算法的基础上, 在允许精确度范围内, 引入容错性理论, 在计算阈值的过程中融入松弛变量, 保证收敛速度; 通过蚁群算法对神经网络的权值进行训练, 避免神经网络陷入局部最优解, 提高收敛速度, 通过改进神经网络算法不断对网络的权值以及阈值进行调整, 使得网络的误差平方和最小, 实现 PM<sub>2.5</sub> 污染检测。实验结果表明, 采用改进神经网络算法得到的 PM<sub>2.5</sub> 污染检测结果具有很高的准确性和有效性。

## 1 PM<sub>2.5</sub> 污染检测问题的分析

当前对空气中的 PM<sub>2.5</sub> 颗粒的检测多是通过智能设备对 PM<sub>2.5</sub> 污染的信号进行采集, 根据传感器传回的信号进行分析, 其基本原理为:

- 1) 通过空气传感器的设备对某一区域的污染空气进行信号采集。
- 2) 对获取的 PM<sub>2.5</sub> 污染样本进行处理时, 必须首先确定被检测空气的污染区域, 得到待检测污染空气的空间位置。以当前已知的 PM<sub>2.5</sub> 污染相关信息完成计算。
- 3) 以待检测 PM<sub>2.5</sub> 污染区域的中点为检测初始位置, 将

收稿日期: 2014-09-12; 修回日期: 2014-10-23。

**基金项目:** 企业信息化与物联网测控技术四川省高校重点实验室(2014WZJ01); 四川理工学院人才引进项目(2012RC21); 四川理工学院学科建设工程项目(2014JC01)。

**作者简介:** 朱 洪(1975-), 女, 江苏南京人, 硕士, 副教授, 主要从事计算机应用技术方向的研究。

蹇红梅(1980-), 女, 四川自贡人, 硕士, 副教授, 主要从事多媒体技术方向的研究。

刘小芳(1969-), 女, 重庆人, 博士, 教授, 硕士生导师, 主要从事智能信息处理、数据挖掘、模式识别与智能系统方向的研究。

该点作为中心点，划分出一个矩形区域，对角的坐标为  $(x, y, z)$ ,  $(x', y', z')$ ，区域高度为  $h$ 。通过式 (1) 与式 (2) 对 PM2.5 污染区域的空气流动强度  $F$ ,  $D$  为单个区域的空气流动强度，对污染强度  $J_{are}$  进行计算：

$$F = \sum_{\substack{y' > y > y' - x' \\ z' > z > z' - h}} (x, y, z) * D \quad (1)$$

$$J_{are} = \frac{\sum K}{F} \quad (2)$$

4) 上式  $K$  通过传感器采集得到，采用求和的方式可以计算出，为污染系数。通过式 (3) 可求出需完成信号的衰减度：

$$T = J_{are} * F \quad (3)$$

5) 通过数据传递网络把获取的参数以及空间位置发送至以神经网络为基础的检测设备终端。数据传递的距离通常很长，因此必须考虑信号衰减。

## 2 标准神经网络算法

神经网络是一种以模仿动物神经网络行为特征为基础，进行分布式并行信息处理的算法。这种网络依靠系统的复杂程度，通过调整内部大量节点之间相互连接的关系，达到处理信息的目的。神经网络模型的种类很多，通常采用的是 BP 神经网络。其是一种按误差逆传播算法训练的多层前馈网络，可学习和存贮大量的输入—输出模式映射关系，无需事前知晓描述这种映射关系的数学方程。其学习规则是使用最速下降法，通过反向传播不断调整网络的权值和阈值，使网络的误差平方和最小。

如图 1 所示，BP 神经网络主要由输入层、隐含层以及输出层组成，其中可包含多个隐含层，一般情况下采用三层结构。神经网络每一层中的神经元只和与其相邻层神经元之间存在联系，和同一层神经元无任何联系。

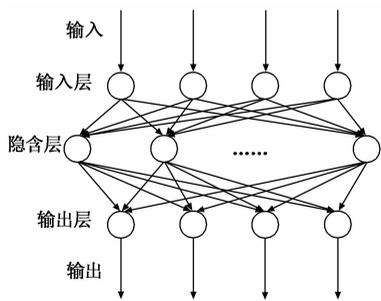


图 1 神经网络结构

BP 神经网络的基本思想为：输入学习样本，利用正向传播的学习信号获取期望输出和实际输出的误差，依据得到的误差值，通过反向传播算法完成网络权值与阈值的修正，通过不断修正与训练使得实际输出逐渐与期望值吻合，直至误差达到可接受范围内，结束训练，同时存储当前的网络权值以及阈值。

网络各层神经元的输出可描述成：

$$y = f\left(\sum_{i=1}^n w_i x_i - \theta\right) \quad (5)$$

式中， $\theta$  表示阈值， $f(X)$  表示激活函数，其可为线性函数，也可为非线性函数。

通常以 Sigmoid 函数为激活函数，将其看作是隐含层的传

递函数，也就是：

$$f(x) = \frac{1}{1 + e^{-x}} \quad (6)$$

全局误差函数可描述成：

$$E = \sum_{p=1}^p E_p = \frac{1}{2} \sum_{p=1}^p \sum_{j=1}^m (t_{pj} - y_{pj})^2 \quad (7)$$

将输出层权值变成：

$$\Delta \omega_{jk} = \eta \sum_{p=1}^p \left( -\frac{\partial E_p}{\partial net_{pj}} \cdot \frac{\partial net_{pj}}{\partial \omega_{jk}} \right) = \eta \sum_{p=1}^p (t_{pj} - y_{pj}) \cdot y_{pj} (1 - y_{pj}) Z_{pk} \quad (8)$$

可将隐含层权值变成：

$$\Delta v_{ki} = \eta \sum_{p=1}^p \left( -\frac{\partial E_p}{\partial net_{pk}} \cdot \frac{\partial net_{pk}}{\partial v_{ki}} \right) = \eta \sum_{p=1}^p \left( \sum_{j=1}^m \delta_{pj} \omega_{jk} \right) Z_{pk} (1 - Z_{pk}) X_{pi} \quad (9)$$

综上所述，BP 神经网络首先利用输入层对样本数据进行正向传播，然后通过目标值求出误差值，将误差按层反向传播，最后通过最速下降法对权值以及阈值进行不断修正。因为其具有多层隐含层，所以 BP 神经网络具备很好的非线性能力，能够对所有函数进行逼近，同时可依据误差完成自适应的调整，使得 BP 神经网络能够传播开来。

## 3 改进神经网络算法

在允许精确度范围内，为了获取更大的几何间隔，引入容错性理论，在上一章计算阈值的过程中融入松弛变量。因为计算样本往往具有随机性，进行归一化处理后，在 BP 神经网络进行学习的初始阶段，样本的训练误差相对较大，收敛速度快，分析式 (8)、式 (9) 可知，一段时间之后，特别是当样本的训练结果逐渐接近 1 或 0 时，训练将进入平台期，造成训练中止。

式 (8)、式 (9) 中分别含有  $y_{pj}(1 - y_{pj})$  与  $Z_{pk}(1 - Z_{pk})$ ，如果隐含层与输出层的输出  $Z_{pk}$  以及  $y_{pj}$  近似于 1 或 0 时，则权值几乎不发生变化，导致 BP 神经网络的训练中止。为了解决上述问题，必须用时间交换空间，增加训练次数，使得训练从平台区域中退出，但上述操作将造成收敛速度降低。

为了在时间与空间交换，增加训练次数的同时，保证收敛速度，引入容错因子，在收敛精度允许范围内，在隐含层与输出层的输出  $Z_{pk}$  以及  $y_{pj}$  近似于 1 或 0 时，将隐含层的权值变化值以及输出的权值变化值与一个放大因子相乘，改变权值的调节量，假设其用  $g(x)$  描述， $g(x)$  应符合下述条件：当  $x \rightarrow 0$  或 1 时，有：

$$\lim_{x \rightarrow 0.1} \frac{g(x)}{x(1-x)} > 0 \quad (10)$$

即当  $x \rightarrow 0$  或 1 时， $g(x) \rightarrow \infty$ ，假设：

$$g(x) = -\ln x(1-x) \quad (11)$$

$g(x)$  的函数图像用图 2 进行描述。

分析图 2 可以看出，当  $x$  趋近 0 或 1 时， $g(x)$  趋近无穷大，也就是当神经元的输出是 0 或 1 时，误差信号也逐渐变大，使得权值调整值达到允许范围内。

经大量研究表明，融入放大因子对隐含层和输出层之间的权值进行修正的效果明显由于隐含层和输入层，因此本文将放大因子应用在隐含层和输出层间的权值修正上。

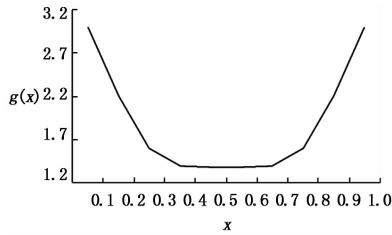


图 2 函数  $g(x)$  的图像

### 4 实验结果分析

为了验证本文提出的改进神经网络算法在检测 PM2.5 污染中的有效性, 需要进行一次实验。以飞利浦的 PLC 小型检测设备为基础, 对算法进行改进。假设待测空间的面积参数是  $S$ , 空气流动性强度系数是  $\eta$ , 空气采集装置数目是  $N$ , 采集的空气样本构成的数据集是  $\{f_1, f_2, \dots, f_n\}$ 。通过下式可求出 PM2.5 污染检测第  $i$  次漏检系数:

$$\epsilon = \frac{\eta(f_i - s^2)}{N + 1} \quad (12)$$

将 PM2.5 污染检测漏检系数作为 PM2.5 污染检测的衡量标准, 污染检测的漏检系数越大, 则检测方法的性能越差。

塑造 PM2.5 污染模拟环境, 在上述环境中完成 10 次 PM2.5 污染检测实验, 实验相关数据如表 1 所示。

表 1 污染检测数据表

序号	空气中有害气体浓度 / (ml/m <sup>3</sup> )	污染程度
1	0.12	中度污染
2	0.31	重度污染
3	0.07	轻度污染
4	0.15	中度污染
5	0.24	中度污染
6	0.09	轻度污染
7	0.34	重度污染
8	0.05	轻度污染
9	0.45	重度污染
10	0.31	重度污染

图 3 描述的是采用改进神经网络算法方法和经典神经网络算法对 PM2.5 污染进行检测的结果比较。分析图 3 可以看出, 与经典神经网络算法相比, 采用改进神经网络算法得到的漏检系数明显降低, 验证了改进神经网络算法的有效性。

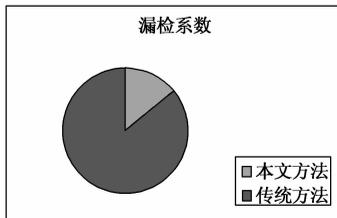


图 3 空气流动性较差时检测结果比较

为了进一步验证改进神经网络算法的有效性, 分别采用改进神经网络算法和经典神经网络算法对上述塑造的 PM2.5 污染模拟环境进行 PM2.5 污染检测, 得到的结果用表 2 进行描述。

表 2 改进神经网络算法和经典神经网络算法 PM2.5 污染检测结果

序号	改进神经网络算法检测有害气体浓度 / (ml/m <sup>3</sup> )	经典神经网络算法检测有害气体浓度 / (ml/m <sup>3</sup> )
1	0.11	0.15
2	0.32	0.41
3	0.07	0.16
4	0.17	0.23
5	0.2	0.24
6	0.11	0.06
7	0.34	0.3
8	0.07	0.11
9	0.46	0.38
10	0.29	0.34

将表 2 得到的数据与表 1 中的真实数据进行对比可以发现, 与经典神经网络算法相比, 采用改进神经网络算法得到的 PM2.5 污染检测结果与真实数据更加吻合, 验证了改进神经网络算法的准确性。

### 5 结论

本文提出一种基于改进神经网络算法的 PM2.5 污染检测方法, 依据 PM2.5 污染检测原理, 在分析标准神经网络算法的基础上, 在允许精确度范围内, 引入容错性理论, 在计算阈值的过程中融入松弛变量, 保证收敛速度; 通过蚁群算法对神经网络的权值进行训练, 避免神经网络陷入局部最优解, 提高收敛速度, 通过改进神经网络算法不断对网络的权值以及阈值进行调整, 使得网络的误差平方和最小, 实现 PM2.5 污染检测。实验结果表明, 采用改进神经网络算法得到的 PM2.5 污染检测结果具有很高的准确性和有效性。

#### 参考文献:

- [1] 陈斌, 万江文, 吴银锋, 等. 神经网络和证据理论融合的管道泄漏诊断方法 [J]. 北京邮电大学学报, 2000, 32 (1): 5-9.
- [2] 王晓敏, 刘希玉, 戴芬. BP 神经网络预测算法的改进与应用 [J]. 计算机技术与发展, 2009, 19 (11): 64-67.
- [3] 卢继霞, 夏连海, 丁思变, 等. 污染度检测传感器分类及检测特点分析 [J]. 润滑与密封, 2011 (7): 99-102.
- [4] 周德新, 樊智勇. 环境泄漏监测与控制技术的研究 [J]. 计算机测量与控制, 2005, 13 (4): 237-240.
- [5] 王俊, 张景义, 陈双基. 室内空气中总挥发性有机物 (TVOCs) 的污染 [J]. 北京联合大学学报, 2002, 16 (3): 52-56.
- [6] 袁华丽, 高松亭, 韩朔睽. 室内空气中挥发性有机物采样方法进展 [J]. 环境污染与防治, 2002, 24 (5): 297-299.
- [7] 郑小萍. 环境空气中 VOCs 的监测技术新进展 [J]. 环境监测管理与技术, 2001, 13 (3): 15-17.
- [8] 游静, 尤进茂, 王国俊. 离线超临界流体萃取和气相色谱质谱联用对实验室内空气中气相有机污染物的测试 [J]. 分析化学, 1998, 26 (7): 86-89.
- [9] 熊邦宏, 李扬, 朱洪雷, 等. 无线传感器网络在环境污染监测中的应用设计 [J]. 计算机测量与控制, 2009 (2): 258-260.
- [10] 梅红. 基于蚁群优化的前向神经网络 [J]. 武汉理工大学学报, 2009, 33 (3): 531-533.