

文章编号:1008-1542(2011)05-0455-05

改进的 BP 神经网络在故障诊断中的应用

陈金辉, 赵雷振, 杨宗宵, 郭 伟

(南京信息工程大学信息与控制学院, 江苏南京 210044)

摘 要:针对传统 BP 算法存在的收敛速度缓慢和易陷入局部极小值的固有缺陷, 提出用具有全局搜索能力的模拟退火算法优化 BP 神经网络, 避免陷入局部极小值, 提高网络的稳定性; 引入 Powell 算法优化模拟退火算法, 加快网络的收敛速度。最后, 以齿轮箱故障诊断为例进行仿真试验, 结果表明改进后的 BP 神经网络比传统 BP 神经网络的训练收敛速度快、精度高。

关键词:模拟退火算法; Powell 算法; BP 算法; PSA-BP 算法; 故障诊断

中图分类号: TP183 **文献标志码:** A

Application of improved BP neural network in fault diagnosis

CHEN Jin-hui, ZHAO Lei-zhen, YANG Zong-xiao, GUO Wei

(School of Information and Control, Nanjing University of Information Science and Technology, Nanjing Jiangsu 210044, China)

Abstract: The traditional BP algorithm has the inherent shortcomings such as low convergent speed and local minimum. So this paper proposes a simulated annealing algorithm with globalsearch property to optimize the BP neural network for avoiding the local minimum and improving its stability. In order to improve the convergent speed of neural network, the Powell algorithm is proposed to optimize the simulated annealing algorithm. Finally, the simulated results of gearbox fault diagnosis are presented to show that the improved BP neural network has higher convergent speed and precision than those of the traditional BP neural network.

Key words: simulated annealing algorithm; Powell algorithm; BP algorithm; PSA-BP algorithm; fault diagnosis

人工智能故障诊断技术因不需要建立对象的精确数学模型, 部分解决了故障诊断的模糊性和不确定性等经典常规方法不能解决各类问题, 因而在故障诊断中得到了越来越广泛的重视^[1]。人工神经网络着眼于模拟人脑, 是一个广泛连接的巨型系统, 它具有 I/O 非线性特性、信息的分布存储、并行处理和全局集体作用, 特别是其高度的自组织和自学习能力, 为故障诊断问题提供了一种崭新的途径^[2]。BP 神经网络是目前应用最广泛的人工神经网络模型之一。但是, BP 算法采用的是梯度训练法, 是一种局部搜索算法, 很容易陷入局部极小和收敛速度缓慢的问题。所以很多学者提出了不同的改进算法, 如变学习速率 BP 算法、动量 BP 算法、LM 算法、遗传算法等^[3-6], 取得了一定效果。笔者将具有全局寻优的快速模拟退火算法引入到 BP 神经网络用于故障诊断, 主要包括: 首先介绍了模拟退火算法, 阐述了模拟退火算法的不足, 用 Powell 算法优化模拟退火算法, 形成快速模拟退火算法, 加快模拟退火的速度; 然后简要说明 BP 算法的训练过程, 将快速模拟退火算法与 BP 算法结合, 形成新的 PSA-BP 算法; 最后将 PSA-BP 算法应用到齿轮箱的故障诊断中, 仿真结果表明该算法具有很好的应用价值。

收稿日期: 2011-03-06; 修回日期: 2011-09-01; 责任编辑: 陈书欣

基金项目: 国家自然科学基金资助项目(71071078)

作者简介: 陈金辉(1970-), 女, 新疆乌鲁木齐人, 副教授, 主要从事软件自动化技术、中间件技术方面的研究。

1 模拟退火算法及其改进

1.1 模拟退火算法

模拟退火(SA)算法的思想最早由 METROPOLIS 于 1953 年提出^[7],1983 年 KIRKPATRICK 等成功将其应用于组合优化问题^[8]。SA 算法是一种关于 Monte-carlo 迭代求解策略的随机寻优算法,其出发点是基于物理中固体物质的退火过程和一般组合优化问题的相似性。SA 算法是在某一初始温度下,伴随温度参数的不断下降,结合概率突跳性在解空间中随机寻找目标函数的全局最优解,在局部最优解状态下能概率性地跳出并最终趋于全局最优解^[9]。模拟退火算法过程可参考文献^[10]。

1.2 快速模拟退火算法

利用 SA 算法求解复杂问题的时候,需要大量的迭代计算,因而收敛速度缓慢。Powell 算法是局部优化直接算法中收敛性较好的一种,迭代次数远远小于 SA 算法^[11]。将 Powell 算法与模拟退火算法组合形成快速模拟退火算法(PSA 算法),可以减小迭代次数。

该算法步骤:1)从任一初始解出发,执行 Powell 算法快速得到一局部极小值点;2)调用 SA 算法进行全局搜索;3)若在指定搜索次数内找到一个比该极小值点更小的点,则初始解被替换为此次的更优解,返回 1),否则该初始解为全局最优解,退出。

2 快速模拟退火算法与 BP 神经网络的融合

2.1 BP 神经网络

BP(back propagation)网络即误差反向传播训练的神经网络,是 RUMELHART 等在 1986 年提出的。它是一种多层前馈神经网络,网络一般由三层或者三层以上的神经元组成,包括输入层、隐含层(至少包含一个隐含层)、输出层。根据 Kolmogorov 定理,一个三层 BP 网络可以逼近 n 维到 m 维的任意连续函数^[12]。所以选用三层的神经网络结构,如图 1 所示。

BP 算法的中心思想是不断调整网络权值和阈值来修正网络的输出误差,下面给出 BP 算法实现过程。

设输入层有 n 个神经元,隐含层有 q 个神经元,输出层有 m 个神经元,输入层到隐含层的权值为 w_{ij} ,隐含层的阈值为 θ_j ,隐含层到输出层的权值为 v_{jk} ,输出层的阈值为 γ_k ,

(P, T) 为含有 N 个样本的样本集合, p 为 N 个样本的任意一个,激励函数选用 $f(x) = \frac{1}{1 + \exp(-x)}$ 。其中 $i = (1, 2, \dots, n), j = (1, 2, \dots, q), k = (1, 2, \dots, m), P = (p_1, p_2, \dots, p_n), T = (t_1, t_2, \dots, t_n)$ 。

1) 正向传播过程

- ① 初始化网络权值 w_{ij}, v_{jk} ; 阈值 θ_j, γ_k ;
- ② 输入训练样本集 (P, T) ;
- ③ 计算隐含层第 j 个神经元输出

$$o_j = f\left(\sum_{i=1}^n w_{ij}x_i - \theta_j\right); \quad (1)$$

④ 计算输出层第 k 个神经元输出

$$o_k = f\left(\sum_{j=1}^q v_{jk}o_j - \theta_k\right)。 \quad (2)$$

2) 反向传播过程

① 计算样本误差

第 p 个样本的误差:

$$E^p = \sum_{k=1}^m \frac{1}{2} (t_k^p - o_k^p)^2。 \quad (3)$$

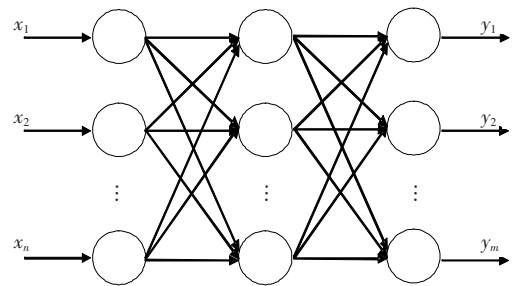


图 1 三层神经网络结构

Fig. 1 Structure of three-layer neural network

全体样本误差:

$$E_T = \sum_{p=1}^N E^p. \quad (4)$$

② 计算输出层一般化误差

$$d_k = o_k(t_k - o_k)(1 - o_k). \quad (5)$$

③ 计算隐含层一般化误差

$$e_j = o_j(1 - o_j) \left[\sum_{k=1}^m d_k v_{jk} \right]. \quad (6)$$

④ 调整权值和阈值

$$v_{jk}(N+1) = v_{jk}(N) + \alpha \cdot d_k \cdot o_j, \quad \gamma_k(N+1) = \gamma_k(N) + \alpha \cdot d_k, \quad (7)$$

其中: $k = 1, 2, \dots, m; j = 1, 2, \dots, q; 0 < \alpha < 1$ 。

$$w_{ij}(N+1) = w_{ij}(N) + \alpha \cdot e_j \cdot p_i, \quad \theta_j(N+1) = \theta_j(N) + \beta \cdot e_j, \quad (8)$$

其中: $i = 1, 2, \dots, n; j = 1, 2, \dots, q; 0 < \beta < 1$ 。

⑤ 返回1)③,重新计算,直至 E_T 小于设定的值或者达到训练次数,训练结束。

2.2 快速模拟退火算法与BP算法的融合

传统BP算法是通过梯度最速下降法来修正权值、阈值向量,很容易陷入局部极小值,且收敛速度慢。为了克服BP算法的缺陷,将具有全局搜索能力的PSA算法和局部搜索的BP算法完美地结合在一起,形成PSA-BP算法,既避免了网络陷入极小值的发生,又加快了训练的收敛速度。

1) 构造目标函数,即所有训练样本的平方误差的均值:

$$E = \frac{1}{2p} \sum_p \sum_k (t_k^p - o_k^p)^2. \quad (9)$$

2) 采用 $T = T_0(1 + \ln(i))$ 的降温策略。

3) 初始化网络的权值 W , 阈值 B , 温度 T_0 , 迭代次数 N , 网络最大允许误差 E_GOAL 。

4) 输入训练样本。

5) 计算当前的误差 E , 若 E 小于 E_GOAL , 退出程序。

6) 将网络的权值 W , 阈值 B 作为初始解, 调用 Powell 算法快速搜索到某个局部极小值, 得到一组新的网络权值 W' 、阈值 B' , 计算当前的误差记为 E' 。令 $W=W', B=B', E=E'$, 转7)。

7) 调用模拟退火算法:

① 加入随机扰动, 令 $W^* = W + \Delta W, B^* = B + \Delta B$, 其中 ΔW 和 ΔB 为很小的均匀分布的随机扰动。

② 重新计算误差 E , 记为 E^* 。若 $E^* < E$, 则 $W = W^*, B = B^*$; 否则计算接受概率 $r = \exp[-(E^* - E)/(kT_i)]$, T_i 为当前温度, 产生一个 $[0, 1]$ 区间的随机数 $p = \text{random}(0, 1)$, 若 $r > p$ 则 $W = W^*, B = B^*$, 否则 $W = W, B = B$ 。

③ 若 E^* 小于 E_GOAL 或者超过迭代次数 N , 退出。否则 $i = i + 1$ 。

8) 若 $E^* < E$, 转6); 若 $E^* > E$, 转7)。

3 仿真试验

利用新的PSA-BP算法,对齿轮箱故障进行仿真。样本集^[13]选择15个特征信号作为齿轮箱故障信号,故障类型有3种,所以输入层神经元数目为15个,输出层神经元数目为3个。根据经验,试验后得出隐含层神经元数目为6个,得到15-6-3的网络结构。隐含层采用S型正切函数 tansig , 输出层采用S型对数函数 logsig 。参数设置:初始温度 $T = 30^\circ\text{C}$, SA最大迭代次数 $N = 50$, 系统最大误差 0.001, 网络总的迭代次数 10 000次。归一化后的样本数据如表1所示。

图2是采用传统BP神经网络训练误差曲线图,图3是采用PSA-BP算法的神经网络训练误差曲线图,从两图的比对结果来看,传统BP神经网络经过10 000次的训练,精度远远大于给定值,收敛速度慢,训练时间长;采用PSA-BP算法的神经网络只需要经过493次的训练,精度达到并小于给定值,收敛速度快,训练时间短。同时,重复多次试验的统计结果表明,采用PSA-BP算法的神经网络更易获取全局最优解。

表1 齿轮箱状态训练样本数据
Tab.1 Training sample data of gearbox state

序号	特征样本					齿轮状态
1	0.228 6	0.129 2	0.072 0	0.159 2	0.133 5	无故障 (1,0,0)
	0.073 3	0.115 9	0.094 0	0.052 2	0.134 5	
	0.009 0	0.126 0	0.361 9	0.069 0	0.182 8	
2	0.209 0	0.094 7	0.139 3	0.138 7	0.255 8	无故障 (1,0,0)
	0.090 0	0.077 1	0.088 2	0.039 3	0.143 0	
	0.012 6	0.167 0	0.245 0	0.050 8	0.132 8	
3	0.044 2	0.088 0	0.114 7	0.056 3	0.334 7	无故障 (1,0,0)
	0.115 0	0.145 3	0.042 9	0.181 8	0.037 8	
	0.009 2	0.225 1	0.151 6	0.085 8	0.067 0	
4	0.260 3	0.171 5	0.070 2	0.271 1	0.149 1	齿根裂痕 (0,1,0)
	0.133 0	0.096 8	0.191 1	0.254 5	0.087 1	
	0.006 0	0.179 3	0.100 2	0.078 9	0.090 9	
5	0.369 0	0.222 2	0.056 2	0.515 7	0.187 2	齿根裂痕 (0,1,0)
	0.161 4	0.142 5	0.150 6	0.131 0	0.050 0	
	0.007 8	0.034 8	0.045 1	0.070 7	0.088 0	
6	0.035 9	0.114 9	0.123 0	0.546 0	0.197 7	齿根裂痕 (0,1,0)
	0.124 8	0.062 4	0.083 2	0.164 0	0.100 2	
	0.005 9	0.150 3	0.183 7	0.129 5	0.070 0	
7	0.175 9	0.234 7	0.182 9	0.181 1	0.292 2	断齿 (0,0,1)
	0.065 5	0.077 4	0.227 3	0.205 6	0.092 5	
	0.007 8	0.185 2	0.350 1	0.168 0	0.266 8	
8	0.072 4	0.190 9	0.134 0	0.240 9	0.284 2	断齿 (0,0,1)
	0.045 0	0.082 4	0.106 4	0.190 9	0.158 6	
	0.011 6	0.169 8	0.364 4	0.271 8	0.249 4	
9	0.263 4	0.225 8	0.116 5	0.115 4	0.107 4	断齿 (0,0,1)
	0.065 7	0.061 0	0.262 3	0.258 8	0.115 5	
	0.005 0	0.097 8	0.151 1	0.227 3	0.322 0	

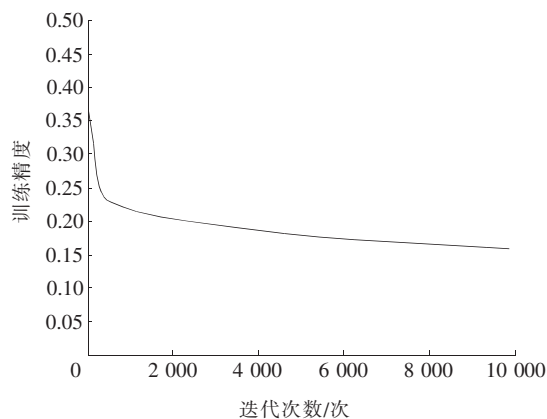


图2 传统BP神经网络训练误差曲线图

Fig.2 Curve of traditional BP neural network training error

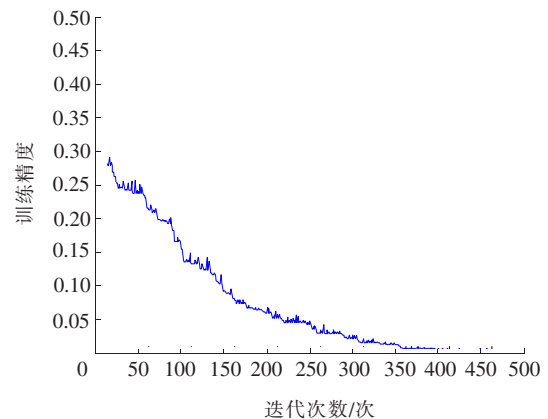


图3 PSA-BP算法的神经网络训练误差曲线图

Fig.3 PSA-BP algorithm based training error curve of neural network

4 预测结果分析

为了验证 PSA-BP 算法的有效性,选取 6 组测试样本进行预测试验,归一化后的样本数据如表 2 所示,测试结果如表 3 所示。

表 2 齿轮箱状态测试样本数据
Tab. 2 Test sample data of gearbox state

序号	特征样本					齿轮状态
1	0.209 7	0.093 7	0.140 5	0.138 7	0.256 6	无故障 (1,0,0)
	0.088 8	0.076 6	0.088 2	0.039 4	0.144 1	
	0.012 1	0.168 1	0.246 1	0.050 7	0.133 8	
2	0.043 4	0.087 2	0.114 1	0.057 2	0.334 3	无故障 (1,0,0)
	0.115 0	0.144 1	0.042 5	0.181 0	0.037 9	
	0.009 4	0.224 3	0.152 6	0.086 7	0.067 8	
3	0.369 3	0.223 0	0.056 2	0.516 1	0.186 1	齿根裂痕 (0,1,0)
	0.160 3	0.143 6	0.150 0	0.131 6	0.050 2	
	0.007 7	0.033 7	0.045 3	0.069 5	0.088 5	
4	0.035 4	0.114 5	0.122 9	0.545 5	0.197 3	齿根裂痕 (0,1,0)
	0.125 6	0.063 3	0.082 0	0.164 8	0.101 4	
	0.005 7	0.149 3	0.183 3	0.130 3	0.070 3	
5	0.072 9	0.190 6	0.134 6	0.240 1	0.283 0	断齿 (0,0,1)
	0.044 4	0.082 8	0.106 2	0.191 9	0.157 6	
	0.010 7	0.169 0	0.365 0	0.271 6	0.250 6	
6	0.263 6	0.226 3	0.117 3	0.116 2	0.107 6	断齿 (0,0,1)
	0.066 7	0.060 8	0.262 4	0.258 3	0.115 0	
	0.004 7	0.099 0	0.152 2	0.227 8	0.321 9	

从表 3 中可以看出,采用 PSA-BP 算法的神经网络诊断正确率明显高于传统的 BP 神经网络,特别是阈值越大,优势越明显。

5 结 语

BP 网络具有良好的非线性映射能力,故可以应用到故障诊断中。但是传统 BP 神经网络存在的收敛速度缓慢和易陷入局部极小值的固有缺陷,阻碍了其应用范围。为了克服 BP 网络的缺陷,将具有全局搜索能力的快速模拟退火算法加入到 BP 神经

网络权值和阈值的优化中,避免了传统 BP 神经网络容易陷入极小值的问题,加快了网络训练速度。最后对齿轮箱故障的仿真结果表明,应用 PSA-BP 算法的神经网络更稳定、更可靠。

参考文献:

- [1] 李春华,张生,杨智慧.基于 BP 神经网络电路故障诊断系统研究[J].煤矿机械(Coal Mine Machinery),2009,30(3):196-198.
- [2] 周世官,张云.基于进化神经网络的模拟电路故障诊断[J].计算机测量与控制(Computer Measurement and Control),2007,15(8):991-993.
- [3] 周开利,康耀红.神经网络模型及其 MATLAB 仿真程序设计[M].北京:清华大学出版社,2005.
- [4] 陈玉芳.BP 神经网络的算法改进及应用研究[D].成都:电子科技大学,2004.

表 3 PSA-BP 算法与传统 BP 算法预测结果

Tab. 3 The prediction results of PSA-BP algorithm and the traditional BP algorithm

	阈值	测试结果正确数 / 测试样本总数
传统 BP	0.5	5 / 6
	0.7	3 / 6
	0.9	0 / 6
PSA-BP	0.5	6 / 6
	0.7	6 / 6
	0.9	6 / 6

(下转第 495 页)

为 87.87%)。HPLC 检测值为 99.75%，文献[8]中的收率为 71.0%。

2 结果与讨论

1)在合成化合物 3 时,用沸点较高的异丙醇为溶剂,适当提高反应温度,不但可以缩短反应时间,还可以提高反应收率。此外,反应没有采取文献中在高温下加入浓硝酸使产品以硝酸盐的形式析出。而是在反应后处理时调节反应液的 pH 值,使产物以游离碱的形式直接析出。这不但可以提高产物的纯度,同时可以避免下一步环合反应中加入氢氧化钠中和其中的硝酸,使环合反应在近中性条件下进行,可以避免更多副反应的发生。

2)在合成化合物 5 时,用 $\text{FeCl}_3 \cdot 6\text{H}_2\text{O}$ /活性炭体系还原硝基替代昂贵的金属钯/碳,极大降低了生产成本,同时减少了生产过程中的危险性。

3)在合成甲磺酸伊马替尼时,采取“一锅煮”的方法,简化了许多繁琐的操作与分离过程,减少了损失,从而提高了收率,节约了成本。

4)采用经过改进与优化的合成工艺制备甲磺酸伊马替尼,反应中所使用的溶剂原料价廉易得,反应过程及后处理都十分简单,反应的总收率较文献[7]报道的 33.54%要高,为 41.9%以上,且纯度可达 99.75% (HPLC)。

参考文献:

- [1] HUANG A, LIU X, MAZKERET-BATIA L Z, et al. Process for Preparing Imatinib[P]. US: 7507821, 2009-03-24.
- [2] WOJCIECH S, WOJCIECH L, LUKASZ K, et al. A Process for Preparation of Imatinib Base[P]. WO: 2006071130, 2005-03-12.
- [3] ZIMMERMANN J D, GIVEN N N. Pyrimidine Derivatives and Process for Their Preparation[P]. EP: 0564409, 1993-10-07.
- [4] RAJENDRA N K, DHARMARAJ R R. A Process of Preparation Imatinib and Imatinib Prepared Thereby [P]. WO: 2004074502, 2004-02-09.
- [5] GENTILINO P M, BALERNA P R. Process for the Preparation of Imatinib[P]. US: 20080103305, 2008-05-01.
- [6] AMALA K, ADIBHATLA B R, NANNAPANENI V C. Process for the Preparation of the Anti-cancer Drug Imatinib and Its Analogues [P]. WO: 2004108699, 2004-12-16.
- [7] 李铭东,李 东,吉 民. 甲磺酸伊马替尼的合成[J]. 中国药学杂志(Chinese Pharmaceutical Journal), 2008, 43(3): 228-229.
- [8] AMALA K K, BHUJANGA R A, KHADGAPATHI P, et al. Polymorphic Forms of Imatinib Mesylate[P]. WO: 200605431, 2006-05-26.
- [9] 安 静,李雪艳. 头孢美唑合成方法研究[J]. 河北科技大学学报(Journal of Hebei University of Science and Technology), 2006, 27(4): 288-293.
- [10] 姚 军,张 显,王颖明,等. 手性药物富马酸比索洛尔的合成研究[J]. 河北科技大学学报(Journal of Hebei University of Science and Technology), 2010, 31(3): 317-320.

(上接第 459 页)

- [5] CHIEN C Y, BIN D L. A backpropagation algorithm with adaptive learning rate and momentum coefficient[J]. Neural Networks, 2002(2): 1 218-1 223.
- [6] VAN ROOIJ A J F, JAIN L C, JOHNSON R P. Neural Network Training Using Genetic Algorithm[M]. [S. l.]: World Scientific, 1996.
- [7] METROPOLIS N. Equations of state calculations by fast computing machines[J]. J Chem Phys, 1953, 21: 1 087-1 091.
- [8] KIRKPATRICK S, GELATT C D Jr, VECCHI M P. Optimization by simulated annealing[J]. Science, 1983, 220: 671-680.
- [9] 王 凌. 智能优化算法及其应用[M]. 北京:清华大学出版社, 2003.
- [10] 田景文,高美娟. 基于改进的模拟退火人工神经网络的薄互储层参数预测[J]. 信息与控制(Information and Control), 2002, 31(2): 180-184.
- [11] 肖玉刚. 一种新型模拟退火神经网络及其应用[J]. 控制工程(Control Engineering of China), 2006, 13(6): 550-552.
- [12] KOLMOGOROV A N. On the representation of continuous functions of many variables by superposition of continuous functions of one variable and addition[J]. Dokl Akad Nauk SSSR, 1957, 114(5): 953-956.
- [13] 张 捷. 基于神经网络的齿轮箱智能故障诊断技术的研究[D]. 苏州:江苏大学, 2003.
- [14] 陈晓明,屈宝存,王小菽,等. 混沌神经网络在球磨机故障诊断中的应用[J]. 河北工业科技(Hebei Journal of Industrial Science and Technology), 2008, 25(5): 301-303.
- [15] 李森林,邓小武. 基于二参数的 BP 神经网络算法改进与应用[J]. 河北科技大学学报(Journal of Hebei University of Science and Technology), 2010, 31(5): 447-450.