

基于粒子群优化的神经网络训练算法研究

高海兵, 高亮, 周驰, 喻道远

(华中科技大学工业工程系自动化所, 湖北武汉 430074)

摘要: 本文提出了基于连接结构优化的粒子群优化算法(PSO)用于神经网络训练, 该算法在训练神经网络权值的同时优化其连接结构, 删除冗余连接, 使神经网络获得与模式分类问题匹配的信息处理能力. 经 PSO 训练的神经网络应用于 Iris, Ionosphere 以及 Breast cancer 模式分类问题, 能够部分消除冗余分类参数及冗余连接结构对分类性能的影响. 与 BP 算法及遗传算法比较, 该算法在提高分类误差精度的同时可加快训练收敛的速度. 仿真结果表明, PSO 是有效的神经网络训练算法.

关键词: 粒子群优化; 神经网络; 遗传算法; 模式分类

中图分类号: TP183 **文献标识码:** A **文章编号:** 0372-2112 (2004) 09-1572-03

Particle Swarm Optimization Based Algorithm for Neural Network Learning

GAO Hai-bing, GAO Liang, ZHOU Chi, YU Dao-yuan

(Inst of Automation, Dept. of Industrial Engineering, Huazhong Univ of Sci & Tech, Wuhan, Hubei 430074, China)

Abstract: This paper proposes a structure-improving particle swarm optimization (PSO) algorithm for training artificial neural network (ANN). The algorithm is successfully applied to pattern classification problems including Iris, ionosphere and breast cancer. By tuning the structure and connection weights of ANN simultaneously, the proposed algorithm generates optimized ANN with problem-matched capacity for processing classification information. By doing this, it also eliminates some ill effects introduced by redundant input features and the corresponding structure of ANN. Compared with BP and GA based training techniques, PSO can improve the classification accuracy while speeding up the convergence process. Simulation results show that PSO is a potentially robust learning algorithm and could be extended to real world applications.

Key words: particle swarm optimization; artificial neural network; genetic algorithm; pattern classification

1 引言

工业、经济、医疗等领域的许多实际问题如质量控制、破产预测、图像识别、医疗诊断等可以转化为模式分类问题求解. 神经网络自学习、自组织、容错以及模拟非线性关系的能力使其特别适合解决上述复杂的模式分类问题. Hornik^[1]证明采用 Sigmoid 响应函数的三层前馈神经网络能够以任意精度模拟复杂的非线性关系, 而神经网络上述性能的实现依赖于对神经网络的充分训练. 因此, 训练算法对于神经网络的模式分类性能有着决定性影响.

BP 算法是最普遍的神经网络训练算法. 但是, Rumelhart^[2]研究证明, 基于梯度下降的 BP 算法依赖于初始权值的选择, 收敛速度缓慢且容易陷入局部最优. BP 的上述缺陷尤其是局部优化特性使其训练的神经网络的输出具有不一致性和不可预测性, 导致模式分类的可靠性降低.

遗传算法的并行搜索策略及全局优化特性使其成为日益普遍的神经网络训练算法. Sexton^[3]通过实验证明, 与 BP 算法比较, 遗传算法训练的神经网络在提高分类正确率的同时可以加快训练的收敛速度. 但是, 遗传算法复杂的遗传操作如选择、复制、交叉、变异使神经网络的训练时间随问题的规模及复杂程度呈指数级增长^[4]. 而且, 由于缺乏有效的局部区域搜索机制, 算法在接近最优解时收敛缓慢甚至出现收敛停滞现象^[5].

粒子群优化算法是基于群体智能理论的优化算法, 通过

种群中粒子间的合作与竞争产生的群体智能指导优化搜索^[6]. 与进化算法比较, PSO 保留了基于种群的全局搜索策略, 但是其采用的速度一位移模型操作简单, 避免了复杂的遗传操作, 在非线形函数优化^[7]、电压稳定性控制^[8]、动态目标优化^[9]等实际问题取得了成功应用. 国内对粒子群优化算法的研究刚刚起步. 文献[10]综述性地介绍了粒子群优化算法及其理论发展与实际应用. 文献[11]通过对算法基本模型的改进, 改善了 PSO 摆脱局部极值的能力. 以 5 个基准函数测试, 改进的算法优于基本 PSO 算法以及惯性权重模型的 PSO 算法. 文献[12]将粒子群优化算法用于同步发电机参数的辨识, 实验结果显示, 算法可有效地辨识同步发电机的运行参数, 简单实用, 具有较强的可行性.

本文提出了基于粒子群优化的神经网络训练算法 PSO. 一般的训练算法仅训练全连接网络结构下的连接权值, 但是神经网络存在的冗余连接会降低神经元的信息处理效率, 大量的冗余连接甚至会影响模式分类的正确率. PSO 在训练神经网络权重的同时训练其连接结构, 删除冗余连接, 使神经网络获得与给定问题匹配的信息处理能力. 实验结果显示, PSO 是有效的神经网络训练算法, 可用于解决实际的模式分类问题

2 基于粒子群优化的神经网络训练算法

粒子群优化算法的速度一位移搜索模型操作简单, 计算复杂度低, 并通过惯性权重协调全局搜索与局部搜索, 既能以

较大的概率保证最优解, 克服 BP 算法局部最优的缺陷, 又可以提高局部区域的收敛速度, 避免 GA 在局部区域搜索过程中的收敛停滞现象。

必须指出, 狭义的神经网络训练仅训练全连接结构下的连接权值。但是神经网络需要的信息处理能力根据模式分类问题的规模和复杂度确定, 信息处理能力不足或过剩都会影响其分类性能。广义的神经网络训练应包括对连接结构的优化, 即在训练权值的同时优化其连接结构, 删除冗余连接。因为部分冗余连接由冗余的输入参数导致, 所以冗余连接的删除在一定程度上还可以消除冗余参数对神经网络分类性能的影响。

本文提出了适合优化神经网络连接结构的粒子群优化算法。首先, 给出算法的基本定义如下:

定义 1 连接结构 $\{c_{ih}\} \{c_{ho}\} \{c_{ih}\}$ 表示输入层与隐含层间的连接结构, $\{c_{ho}\}$ 表示隐含层与输出层间的连接结构, $\{c_{ih}\}$ 与 $\{c_{ho}\}$ 均为二进制变量矩阵, 对应的连接存在则该变量为 1, 否则为 0。

定义 2 连接阈值 $\{\theta_{ih}\} \{\theta_{ho}\}$: 连接阈值 $\{\theta_{ih}\}$ 与 $\{\theta_{ho}\}$ 为 $[0, 1]$ 区间的实数, 与连接变量结合用于控制神经网络的连接结构。

定义 3 连接变量 $\{\delta_{ih}\} \{\delta_{ho}\}$: 连接变量 $\{\delta_{ih}\} \{\delta_{ho}\}$ 与连接阈值结合决定神经网络的连接结构。若连接变量的值大于连接阈值, 连接阀门开启, 则连接存在。连接变量一般为 $[0, 1]$ 区间的实数。

在上述定义的基础上建立仅优化神经网络连接结构的 PSO 算法模型。其中, PSO 算法机理、基本定义及算法实现可参考文献 [9]。

连接变量速度的迭代:

$$v_{ih} \leftarrow w * v_{ih} + c_1 * \text{rand}() * (p_i - \delta_{ih}) + c_2 * \text{rand}() * (g - \delta_{ih}) \tag{1}$$
$$v_{ho} \leftarrow w * v_{ho} + c_1 * \text{rand}() * (p_i - \delta_{ho}) + c_2 * \text{rand}() * (g - \delta_{ho}) \tag{2}$$

连接变量位置的迭代:

$$\delta_{ih} \leftarrow \delta_{ih} + v_{ih} \tag{3}$$
$$\delta_{ho} \leftarrow \delta_{ho} + v_{ho} \tag{4}$$

连接结构的迭代:

$$\text{if } (\theta_{ih}/\theta_{ho} < \delta_{ih}/\delta_{ho}) \text{ then } c_{ih}/c_{ho} \leftarrow 1$$
$$\text{else } c_{ih}/c_{ho} \leftarrow 0 \tag{5}$$

仅优化连接结构的 PSO 算法模型中的连接变量速度—位置迭代公式(1)~(4)为 PSO 算法的基本搜索模型。但是与传统 PSO 算法不同, 该算法中 p_i , g 分别表示连接变量(而非连接结构本身)的个体与全局极值, 由连接变量及阈值确定的连接结构结合基本 PSO 算法训练的连接权值在给定样本下的训练误差精度决定。该算法通过间接优化可连续变化的连接变量达到训练二进制表达的连接结构的目的, 并由连接结构的迭代式(5)体现 PSO 算法对神经网络结构的更新。SPSO 算法在连接结构优化算法的基础上提出, 算法在优化连接结构的同时结合传统 PSO 算法训练神经网络的连接权值。

3 实验说明及结果分析

本文使用 Iris, Ionosphere 以及 Breast Cancer 模式分类数据

集作为 SPSO 算法的测试实例。同时, 采用 BP 及遗传算法训练全连接结构下的神经网络权值, 并使用同时训练权值与连接结构的遗传算法(SGA)与 SPSO 比较。

首先给出算法的参数设置。基于粒子群优化的训练算法(PSO 与 SPSO)种群规模 $n=40$, 初始惯性权重 $w(0)=0.9$ 并随迭代次数线性递减至 0.4, $c_1=c_2=2$, 连接权值为 $[-1, 1]$ 区间变量, Iris 问题的连接阈值 $\theta_{ih}=\theta_{ho}=0.5$ 。由经验公式确定神经网络隐含层节点数 $n_h=8$ 。所有模式分类问题均以最大迭代次数 Itemax 为算法停止条件。Iris 数据集中各种算法训练迭代次数为 400(BP 算法为 1000), Ionosphere 数据集中 $\theta_{ih}=\theta_{ho}=0.4$, $n_h=15$, 各种算法训练迭代次数为 800(BP 算法为 2000), Breast Cancer 数据集 $n_h=12$, 其余参数设置同 Iris。

SPSO 及 SGA 算法随迭代次数的训练误差曲线如图 1~3 所示。SPSO、PSO、SGA、GA 及 BP 算法的测试误差、分类正确率、神经网络连接数以及训练阶段的 CPU 运行时间见表 1。

表 1 神经网络训练性能比较				
算法	性能指标	Iris	Ionosphere	Breast cancer
SPSO	误差指标	1.117	1.900	0.627
	分类正确率	97.959%	97.351%	98.995%
	连接数	29	265	61
	CPU 时间(s)	16.304	198.973	92.490
PSO	误差指标	1.613	2.734	0.797
	分类正确率	93.877%	94.702%	98.492%
	连接数	56	525	120
	CPU 时间(s)	13.439	194.730	91.701
SGA	误差指标	1.772	4.809	1.168
	分类正确率	95.918%	93.377%	98.995%
	连接数	32	272	71
	CPU 时间(s)	49.711	899.264	308.424
GA	误差指标	2.928	4.369	1.499
	分类正确率	93.877%	93.377%	98.492%
	连接数	56	525	120
	CPU 时间(s)	48.369	817.596	286.782
BP	误差指标	2.198	4.037	1.556
	分类正确率	95.918%	93.377%	98.492%
	连接数	56	525	120
	CPU 时间(s)	55.340	811.477	382.580

由图 1~3 可知, SPSO 在训练阶段收敛速度快且误差精度高。此外, 与 SGA 比较, 其最优及平均误差曲线具有收敛一

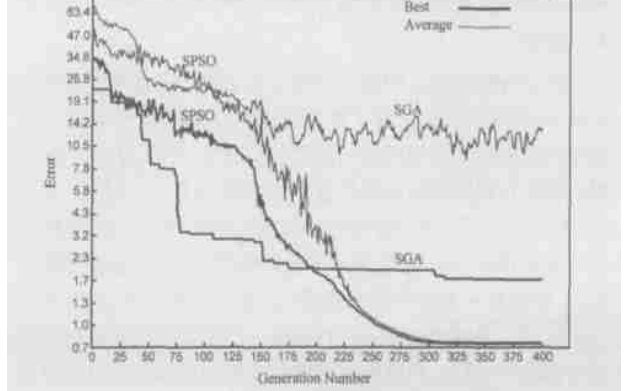


图 1 SPSO 与 SGA 训练误差曲线图(Iris)

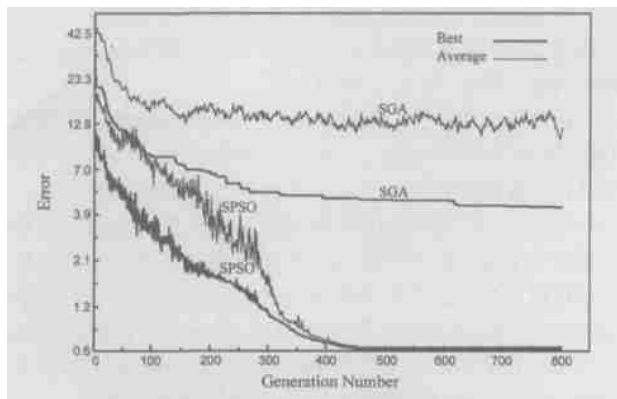


图2 SPSO与SGA训练误差曲线图(Ionosphere)

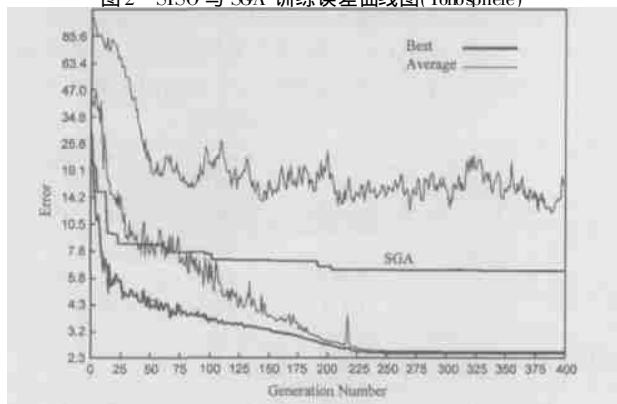


图3 SPSO与SGA训练误差曲线图(Breast cancer)

致性,从而提高了训练收敛的可靠性。如表1所示,SPSO对结构的优化提高了神经网络的模式分类性能,表明SPSO删除的神经网络连接为冗余连接。同时也验证了冗余连接对神经网络性能的影响以及信息处理能力过剩导致的过度训练问题。由实验结果还可以推出,粒子群优化算法与遗传算法相比,在训练时间接近时,分类性能显著提高;若达到同样误差目标,采用粒子群优化算法收敛所需的训练迭代次数明显降低。由此可见,粒子群优化算法不仅使训练的收敛速度大大提高,且其训练的神经网络的性能也显著增强。

需要指出,尽管对结构的优化增加了SPSO训练阶段的复杂度,但是经过连接结构优化的神经网络分类模型在实际应用中信息处理效率提高。经过连接结构优化的神经网络尤其适合大规模数据的实时处理。

4 总结

模式分类问题可能存在冗余分类参数,从而导致了神经网络中的冗余连接,降低了神经网络的处理速度。大量的冗余连接甚至会影响神经网络的模式分类精度。本文提出的SPSO训练算法在训练神经网络权值的同时删除其冗余连接,实现连接结构的优化。与训练全连接结构权值的PSO比较,算法在保证分类正确率的同时减少了连接数目。与BP及遗传算法比较,在提高分类误差精度的同时加快了训练收敛速度。实验结果验证了SPSO算法的有效性。进一步的工作可以从以下两个方面展开:首先,根据神经网络的实时训练性能自适应地改变连接阈值的设置。其次,根据完成训练的神经网络评价输入

参数的影响,并根据重要参数由PSO算法提取分类规则,这些都是基于PSO的神经网络模式分类的改进方向。

参考文献:

- [1] Hornik K, Stinchcombe M, White H. Multilayer feed-forward networks are universal approximators[J]. Neural Networks, 1989, 2(5): 359-366.
- [2] Rumelhart D E, Hinton G E, Williams R J. Learning representations by back propagating errors[J]. Nature, 1986, 323(11): 533-536.
- [3] Sexton R S, Dorsey R E. Reliable classification using neural networks: a genetic algorithm and backpropagation comparison[J]. Decision Support Systems, 2000, 30(1): 11-22.
- [4] Yang J M, Kao C Y. A robust evolutionary algorithm for training neural networks[J]. Neural Computing and Application, 2001, 10(3): 214-230.
- [5] Franchini M. Use of a genetic algorithm combined with a local search method for the automatic calibration of conceptual rainfall-runoff models[J]. Hydrological Science Journal, 1996, 41(1): 21-39.
- [6] Kennedy J, Eberhart R C. Particle swarm optimization[A]. Proceedings of IEEE International Conference on Neural Networks[C]. Australia: IEEE, 1995. 1942-1948.
- [7] Shi Y H, Eberhart R C. Empirical study of particle swarm optimization[A]. Proceedings of IEEE International Congress on Evolutionary Computation[C]. USA: IEEE, 1999. 6-9.
- [8] Yoshida H, Kawata K, Yoshikazu F. A Particle swarm optimization for reactive power and voltage control considering voltage security assessment[J]. IEEE Transaction on Power System, 2000, 15(4): 1232-1239.
- [9] Carlisle A, Dozier G. Adapting particle swarm optimization to dynamic environments[A]. Proceedings of International Conference on Artificial Intelligence[C]. USA: IEEE, 2000. 11-15.
- [10] 李爱国, 覃征, 鲍复民, 等. 粒子群优化算法[J]. 计算机工程与应用, 2002, (21): 1-3.
- [11] 王岁花, 冯乃勤, 等. 一类新颖的粒子群优化算法[J]. 计算机工程与应用, 2003, (13): 109-110.
- [12] 龙云, 王健全. 基于粒子群算法的同步发电机参数辨识[J]. 大电机技术, 2003, (1): 8-11.

作者简介:



高海兵 男, 1978年9月出生于江苏省南通市, 华中科技大学工业工程系硕士研究生, 研究方向为计



高亮 男, 1974年8月出生于山东省临清市, 现为华中科技大学工业工程系副教授、副主任, 研究方向运筹学与优化, 在国内外发表学术论文20余篇, 算智能、运筹学与优化。