基于面向对象自适应粒子群算法的神经网络训练 *

徐乐华[1], 凌卫新[1], 熊丽琼[2]
(1. 华南理工大学 数学科学学院, 广州 510640; 2. 江西师范大学 计算机信息工程学院, 南昌 330022)

摘　要: 针对传统的神经网络训练算法收敛速度慢和泛化性能低的缺陷,提出一种新的基于面向对象的自适应粒子群优化算法(OAPSO)用于神经网络的训练。该算法通过改进 PSO 的编码方式和自适应搜索策略以提高网络的训练速度与泛化性能,并结合 Iris 和 Ionosphere 分类数据集进行测试。实验结果表明:基于 OAPSO 算法训练的神经网络在分类准确率上明显优于 BP 算法及标准 PSO 算法,极大地提高了网络泛化能力和优化效果,具有快速全局收敛的性能。

关键词: 神经网络; 粒子群优化算法; 面向对象方法; 拓扑结构优化
中图分类号: TP183　　**文献标志码**: A　　**文章编号**: 1001-3695(2009)01-0111-03

Neural network training based on object-oriented adaptive particle swarm optimization

XU Le-hua[1], LIN Wei-xin[1], XIONG Li-qiong[2]
(1. *School of Mathematics Science*, *South China University of Technology*, *Guangzhou* 510640, *China*; 2. *School of Computer & Information Engineering*, *Jiangxi Normal University*, *Nanchang* 330022, *China*)

Abstract: In view of the traditional neural network training algorithm defects of slow convergence speed and the low generalization, this paper proposed a novel object-oriented adaptive particle swarm optimization(OAPSO) algorithm in the neural network training. This algorithm enhanced the training speed and the generalization of network through improving the encoding method and the self-adapted search strategy of PSO. Then, used two standard data sets, Iris and Ionosphere, in the test. The experiments show that the neural network based on OAPSO algorithm is obviously superior to BP algorithm and standard PSO algorithm in the classification accuracy rate, and enhances the generalization and the optimized effect of the network. This algorithm has the performance of rapid global convergence.

Key words: neural network; particle swarm optimization algorithm; object-oriented methods; topology optimization

0　引言

神经网络训练的根本要求是构造具有足够泛化能力的网络,使构造的神经网络对训练样本集输出与输入的映射关系拥有足够的描述能力。目前,在神经网络的应用和研究中,如何有效地提高网络的泛化能力和训练速度是存在的最主要问题。为使网络具有足够的泛化能力,要求网络输出误差在相邻样本点间的变化足够小。因此,网络训练的基本要求是网络输出误差足够小和相邻样本点间网络输出误差的波动量足够小[1]。

为了达到神经网络训练的基本要求,国内外的研究者采用粒子群优化(PSO)算法训练神经网络的权值和拓扑结构[2~7]。PSO 算法通过种群间信息共享进行寻优,以获得结构优化的神经网络;但是,它存在收敛速度慢、易陷入局部最优和种群多样性易丢失问题[8~11],且缺乏有效的寻优搜索机制,算法在接近最优解时收敛缓慢甚至出现停滞现象,导致网络训练难以达到理想效果。针对 PSO 算法的不足,本文提出一种新的基于面向对象的自适应粒子群算法(object-oriented adaptive particle swarm optimization,OAPSO)应用于神经网络的训练。

1　基于面向对象的自适应粒子群算法

1.1　标准粒子群优化算法

粒子群优化算法[12]是一种基于群智能的进化计算技术,起源于对一个简化社会模型的仿真。它是受鸟群群体运动行为方式启发而提出的一种具有代表性的集群智能方法,采用速度-位移搜索模型,算法操作简单、计算复杂度低,既能以较大的概率保证最优解,又可以提高局部区域的收敛速度。PSO 算法的速度模型进化方程为

$$v_{ij}^{t+1} = w \times v_{ij}^{t} + c_1 r_1 \times (p_{ij}^{t} - x_{ij}^{t}) + c_2 r_2 \times (g_j^{t} - x_{ij}^{t}) \tag{1}$$

$$x_{ij}^{t+1} = x_{ij}^{t} + v_{ij}^{t+1} \tag{2}$$

其中:r_1 和 r_2 为[0,1]随机数;w 为惯性权重;c_1 和 c_2 为加速度权重;x_{ij}^{t} 为微粒当前位置;v_{ij}^{t+1} 为微粒新的速度;p_{ij}^{t} 为个体微粒所达到的最好位置;g_j^{t} 为所有微粒经历过的最好位置。在得到微粒新的速度后,便可获得新的位置。

由于 PSO 算法受种群初始分布的影响,尤其在高维复杂多峰函数求解问题中,可能导致种群运动陷入局部最优;同时种群引导信息较为单一,粒子将难以对整个解空间进行均匀有效的搜索,致

收稿日期: 2008-04-06; **修回日期**: 2008-06-26　　**基金项目**: 广东省自然科学基金资助项目(04300102)

作者简介:徐乐华(1983-),男,江西乐平人,硕士研究生,主要研究方向为人工神经网络、智能进化算法(xulehua616@yahoo.com.cn);凌卫新(1966-),女,广东广州人,副教授,博士,主要研究方向为模块化神经网络、图像处理和智能进化算法;熊丽琼(1983-),女,江西赣州人,硕士研究生,主要研究方向为网络安全、优化算法.

使种群的多样性易丢失,尤其是算法运算初期。为克服 PSO 算法收敛速度慢、易陷入局部最优和种群多样性易丢失问题,本文提出一种新的基于面向对象的自适应粒子群算法。

1.2　基于面向对象自适应的粒子群算法

在面向对象设计方法学中[13],现实世界中的每个个体都可看成是一个对象。对象具有属性和方法。属性部分也称为数据部分或静态部分,用于描述对象的状态;方法部分也称为函数部分或动态部分,用于改变对象的状态。

在空间中飞行的粒子,作为现实世界一个实体,就是一个对象。因此,粒子也具有自己的属性和方法。从面向对象设计方法学的角度出发,为进一步提高粒子的自适应搜索能力,本文提出 OAPSO 算法。OAPSO 作为一种全新概念的随机优化算法,为其提出了相应的粒子编码模型:

粒子 S_i =〈方法;属性〉=(飞行搜索;位置,速度,适应度);i 为粒子的编号

作为完整的粒子对象,飞行搜索为粒子的动态方法,位置、速度、适应度等为粒子的属性,粒子通过自身的动态方法与其他对象相互传递信息,改变自身的属性值。

为克服标准 PSO 算法易陷入局部最优和种群多样性易丢失问题,增强 OAPSO 种群的多样性对 OAPSO 进行多元规划。在算法初期适当考虑种群的多元最优值分布特征,引入群体中前 M 最优的对象对当前微粒的运动进行局部与全局并行搜索的多元规划。微粒的运动速度由下式决定:

a)基于个体自主的局部搜索:

$$\varphi_{id}^{t+1} = \chi_1[v_{id}^t + c_1 r_1(p_{id}^t - x_{id}^t) + c_2 r_2(x_{id}^{t-1} - x_{id}^t)] \tag{3}$$

b)基于多元最优信息规划的全局搜索:

$$\phi_{id}^{t+1} = \chi_2[v_{id}^t + c_0 r_0(p_{gd}^{t-1} - x_{id}^t) + \sum_{k=1}^{M} c_k r_k(p_{k,id}^t - x_{id}^t)] \tag{4}$$

其中:$p_{k,id}^t$ 表示在 t 时刻整个群体中排列第 k 的最优值;c_k、r_k 分别表示与其相匹配的加权系数和随机量。

OAPSO 作为一种自适应随机搜索算法,粒子对象将根据个体误差与种群误差之间的关系,采取局部与全局搜索相结合的自适应策略:

$$v_{id}^{t+1} = u\varphi_{id}^{t+1} + (1-u)\phi_{id}^{t+1} \tag{5}$$

其中:$u \in [0,1]$ 为自适应搜索因子,用于平衡粒子对象局部搜索与全局搜索的关键因素,也是粒子自适应搜索的体现,旨在充分发挥 OAPSO 算法的寻优性能,增强 OAPSO 算法的自适应性。本文自适应因子 u 采取如下搜索策略(u:局部→全局→局部):粒子对象根据个体误差与种群最优误差之间的关系,随迭代次数 $u = 1/[1+\exp(-\lambda(t-T_{\max}/2)^2)] - 0.2$ 进行自适应搜索。当 $1 \leqslant t \leqslant t_1$ 时,为增强种群对象初始化阶段的均匀分布性,保证种群一定的多样性,采取局部搜索;当 $t_1 \leqslant t \leqslant t_2$ 时,主要采取基于信息共享的全局搜索策略,对整个空间进行全局搜索;当 $t_2 \leqslant t \leqslant T_{\max}$ 时,种群各粒子都趋于进入收敛域或已进入收敛域,此时为防止各粒子在进一步搜索时跳出收敛域或在收敛域内搜索步距过大,将主要采取局部搜索。

2　基于 OAPSO 算法的神经网络训练

为提高神经网络训练算法收敛速度和泛化性能,将 OAPSO 算法用于神经网络训练学习,使其在训练神经网络权值的同时优化网络的连接结构、删除冗余连接、有效地提高网络的分类准确率。OAPSO 算法的学习步骤如下:

a)样本预处理,将样本数据值归一化为(0,1)范围内。

b)初始化网络结构,设定网络的输入层、隐层、输出层的神经元个数;初始化种群规模、位置和速度。

c)根据适应度值公式:$\text{fitness} = 1/2n\sum_{p=1}^{N}\sum_{j}^{\dim}(y_{pj} - t_{pj})^2$,计算对象的适应度值,确定个体最优值和种群最优值。其中:N 为训练样本数,dim 为输出神经元个数,y_{pj}、t_{pj} 分别为第 p 个样本的第 j 个分量的输出值和期望输出值。

d)根据式(3)~(5)更新种群粒子的位置和速度,并考虑更新后的速度和位置是否在有限的范围内:

(a)考虑速度。若 $v_{ij}(t+1) \geqslant v_{\max}$,则 $v_{ij}(t+1) = v_{\max}$;若 $v_{ij}(t+1) \leqslant -v_{\max}$,则 $v_{ij}(t+1) = -v_{\max}$;否则不变。

(b)考虑位置。若 $w_{ij}(t+1) \geqslant w_{\max}$,则 $w_{ij}(t+1) = w_{\max}$;若 $w_{ij}(t+1) \leqslant -w_{\min}$,则 $w_{ij}(t+1) = -w_{\min}$;否则不变。

其中:$w_{\min}$、$w_{\max}$ 和 $v_{\max}$ 都为常数。

当 $|w_{ij} \leqslant \beta$ 时,说明该连接权对网络误差贡献小,对网络输出影响低。设定该连接权为冗余的,应当删除其连接权,即令 $w_{ij} = 0$。

e)判断算法是否达到终止条件,如果满足终止条件,算法终止;否则返回到 c)。

3　算法实验与仿真

3.1　Iris 分类问题

Iris 数据是测试分类器性能的一个基准数据集。在数据记录中, Iris 植物有四种属性,即萼片长度、萼片宽度、花瓣长度和花瓣宽度。三种植物各有 50 组数据,总共有 150 组数据。实验将 Iris 数据集作为算法的测试实例,将 150 组数据分为训练数据和测试数据, 根据经验选取 100 个训练数据和 50 个测试数据。

在实验中首先根据经验确定一个较大的网络,选用三层的神经网络做实验;其次根据 Iris 植物属性选定输入层为四个节点,每个节点代表植物的一种属性,输出层选定为三个节点,每个输出节点代表一类植物,根据经验公式设定隐层节点数为 10 个;最后使用剪枝进行网络结构优化。

本算法采用两种终止条件分别进行算法测试取得不错效果;并与 BP、PSO 算法进行比较,实验结果表明 OAPSO 算法的有效性。

1)以最大迭代次数为终止条件的实验

实验是以最大迭代次数为终止条件,三种算法的训练误差曲线如图 1 所示。其中训练平均误差、分类正确率、网络连接数的结果如表 1 所示。

表 1　OAPSO 与 BP、PSO 训练结果比较

算法	迭代次数	平均误差	平均正确率/%	平均权数
OAPSO	500	0.024 9	96.82	46
BP	500	0.111 6	90.41	70
PSO	500	0.044 3	94.36	70

由表 1 和图 1 可知,在最大迭代次数条件下,OAPSO 算法不仅具有较快的收敛速度,而且还具有较强的泛化性能。

2)以目标误差为终止条件的实验

实验是以目标误差为终止条件,三种算法的训练误差曲线如图 2 所示。其中训练平均误差、分类正确率、网络连接数的结果如表 2 所示。

由表 2 和图 2 可知,在以相同目标误差为终止条件时,OAPSO 算法达到目标误差的训练次数少于 BP 和 PSO 算法,具

有快速收敛速度,表现出较强的泛化性能。

表 2　OAPSO 与 BP、PSO 训练结果比较

算法	目标误差	平均次数	平均正确率/%	平均权数
OAPSO	0.04	385	96.82	44
BP	0.04	928	92.43	70
PSO	0.04	569	94.43	70

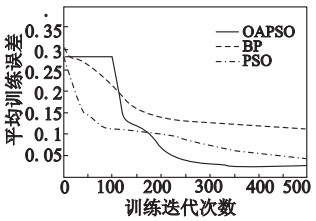

图1　OAPSO与BP、PSO 误差曲线图

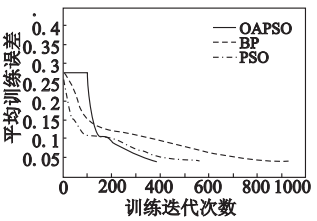

图2　OAPSO与BP、PSO 误差曲线图

3.2　Ionosphere 分类问题

Ionosphere 为雷达信号分类数据,一类是 good 数据,该数据反映出电离层结构特征;另一类属于 bad 数据,不能反映电离层特征。该数据集共有 351 个数据,每组数据有 34 维的特征属性。根据经验选用 200 个作为训练数据,151 个作为测试数据。

实验中,首先根据经验确定一个较大的网络,选用三层的神经网络做实验;其次根据 Ionosphere 数据特性选定输入层为 34 个节点,每个节点表征数据一种属性,输出层选定为一个节点,输出节点代表是有用信号或无用信号,根据经验公式确定隐层节点数为 15 个;最后使用剪枝进行网络结构优化。

本算法采用两种终止条件来进行算法测试,取得不错结果;并与 BP、PSO 算法进行比较,实验结果表明 OAPSO 算法的有效性。

1)以最大迭代次数为终止条件实验

实验以最大迭代次数为终止条件,三种算法的训练误差曲线如图 3 所示。其中训练平均误差、分类正确率、网络连接数的结果如表 3 所示。

表 3　OAPSO 与 BP、PSO 训练结果比较

算法	迭代次数	平均误差	平均正确率/%	平均权数
OAPSO	500	0.023 1	95.76	432
BP	500	0.080 8	90.48	525
PSO	500	0.034 6	94.89	525

由表 3 和图 3 可知,在最大迭代次数条件下,OAPSO 算法不仅具有较快的收敛速度,而且还具有较强的泛化性能。

2)以目标误差为终止条件实验

实验是以最大迭代次数为终止条件,三种算法的训练误差曲线如图 4 所示。其中训练平均误差、分类正确率、网络连接数的结果如表 4 所示。

表 4　OAPSO 与 BP、PSO 训练结果比较

算法	目标误差	平均次数	平均正确率/%	平均权数
OAPSO	0.04	306	95.76	463
BP	0.04	1 392	94.89	525
PSO	0.04	391	94.69	525

由表 4 和图 4 可知,在以相同目标误差为终止条件时,OAPSO 算法达到目标误差的训练次数明显少于 BP 和 PSO 算法,具有快速收敛速度,表现出较强的泛化性能。

3.3　OAPSO 算法性能分析

通过 Iris 和 Ionosphere 两组数据实验可以看出,OAPSO 算法具有很强的自适应搜索能力,它根据种群对象的适应度值,采取局部与全局相结合的搜索策略。a)与 BP 算法相比,OAPSO 算法收敛速度快,不易陷入局部最优,且无须使用函数的梯度值。b)与 PSO 算法相比,虽 PSO 算法初期收敛速度较快,但 PSO 算法种群多样性易丢失,以至算法后期收敛速度较慢;OAPSO 算法运算初期速度虽然较慢,但初期实行局部为主搜索有利于粒子均匀分布整个搜索空间,有利于保持种群的多样性,同时 OAPSO 算法后期采取局部为主搜索,有利于粒子在收敛域内搜索,防止粒子跳出收敛域或在收敛域搜索步距过大,这样更有利快速找到最优值。

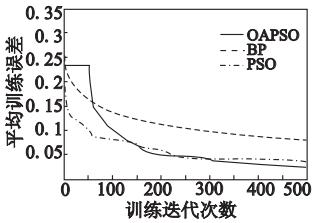

图3　OAPSO与BP、PSO 误差曲线图

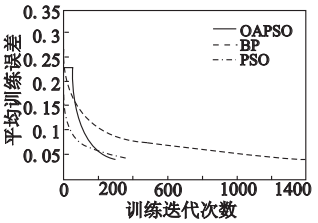

图4　OAPSO与BP、PSO 误差曲线图

4　结束语

本文提出一种基于面向对象的自适应粒子群优化算法,并将该算法用于神经网络的训练中,取得了较好的结果。实验证明:该算法与 BP 算法、PSO 算法相比,取得了明显的网络优化效果,提高了网络的泛化性能,具有快速全局收敛速度。进一步的研究工作将着手引入合理的种群多样性评判规则以及神经网络分类规则。

参考文献:

[1] 张德贤. 基于输出误差与偏导数误差信息融合的神经网络训练[J]. 计算机工程与应用,2002, 38(24):55-57.

[2] 高海兵,高亮,周驰,等. 基于粒子群优化的神经网络训练算法研究[J]. 电子学报,2004,32(9):1572-1574.

[3] 刘洪波,王秀坤,孟军,等. 神经网络基于粒子群优化的学习算法研究[J]. 小型微型计算机系统,2005,26(4):638-640.

[4] GRIMALDI E A, GRIMACCIA F, MUSSETTA M, *et al*. PSO as an effective learning algorithm for neural network applications[C]//Proc of the 3rd International Conference on Computational Electromagnetics and Its Applications. 2004:557-560.

[5] GUDISE V G, VENAYAGAMOORTHY G K. Comparison of particle swarm optimization and backpropagation as training algorithms for neural networks[C]//Proc of IEEE Swarm Intelligence Symposium. 2003: 110-117.

[6] ZHANG Jing-ru, ZHANG Jun, LOK T M, *et al*. A hybrid particle swarm optimization-back-propagation algorithm for feedforward neural network training[J]. Applied Mathematics and Computation, 2006,185(2):1026-1037.

[7] DENG Chang-hui, WEI Xin-jiang, GUO Lian-xi. Application of neural network based on PSO algorithm in prediction model for dissolved oxygen in fishpond[C]//Proc of the 6th World Congress on Intelligent Control and Automation. 2006:9401-9405.

[8] 毛恒, 王永初. 一种基于差异演化变异的粒子群优化算法[J]. 计算机工程与应用,2007,43(30):56-59.

[9] 陆克中, 王汝传, 帅小应. 保持粒子活性的改进粒子群优化算法[J]. 计算机工程与应用, 2007,43(11):35-38.

[10] 何庆元, 韩传久. 带有扰动项的改进粒子群算法[J]. 计算机工程与应用,2007,43(7):84-86.

[11] 李绍军, 王惠, 钱锋. 基于模式优化思想改进的粒子群优化算法[J]. 控制与决策,2006,21(10):1193-1196.

[12] 杨维, 李歧强. 粒子群优化算法综述[J]. 中国工程科学, 2004, 6(5):87-94.

[13] 文登敏, 张丽梅. 面向对象理论与实践[M]. 成都: 西南交通大学出版社, 2005.