粒子群优化算法与 BP 算法和遗传算法的比较研究 *

董延胜，张安年
(河南科技大学 电子信息工程学院，河南 洛阳 471003)

摘 要： 讨论了粒子群优化算法的基本原理和实现步骤，分析了该算法中各参数的设置。通过一个测试函数，对粒子群优化算法与遗传算法和 BP 算法分别进行了比较，结果表明粒子群优化算法在找寻最优解效率上好于其他两种算法。

关键词： 人工神经网络；BP 算法；遗传算法；粒子群优化算法

中图分类号： TP301.6 **文献标识码：** A

Research of particle swarm optimization algorithm comparison with back-propagation algorithm and genetic algorithm

DONG Yan Sheng, ZHANG An Nian
(School of Electronic Information Engineering, Henan University of Science and Technology, Luoyang 471003, China)

Abstract: This paper discusses particle swarm optimization algorithm principle and the step of implementation, and then analyzes the setting of each parameter. Particle swarm optimization algorithm is compared with genetic algorithm and back-propagation algorithm through the same mathematic function. The comparative result indicates that particle swarm optimization algorithm can more easily obtain the optimum solutions than genetic algorithm and back-propagation algorithm.

Key words: artificial neural networks; back-propagation algorithm; genetic algorithm; partical swarm optimization algorithm

人工神经网络 ANN(Artificial Neural Networks)是模拟大脑分析过程的简单数学模型，它是一个非线性动力学系统，其特色在于信息的分布式存储和并行协同处理。虽然单个神经元的结构极其简单，功能有限，但大量神经元构成的网络系统所能实现的行为却是极其丰富多彩的。在人工神经网络的发展进程中，学习算法的研究一直有着十分重要的地位。遗传算法 GA (Genetic Algorithm)和反向传播算法 BP Algorithm(Back-Propagation Algorithm)是该研究领域值得关注的两个研究方向。

BP 算法是最普遍的神经网络训练算法。但是，Rmulhart[1]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。

遗传算法能自适应地控制搜索过程以求得最优解，其本质是一种求解问题的高度并行性全局搜索算法。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[2]。但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，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。

粒子群优化 PSO(Partical Swarm Optimization)算法是一种基于群智能方法的演化计算技术。最早是由 Kenrredy 和 Eberhart 于 1995 年在鸟群、鱼群和人类社会行为规律的启发下提出的[3-4]。该算法模拟鸟群飞行觅食的行为，通过鸟之间的集体协作使群体达到最优。与采用传统的 BP 算法和遗传算法相比，采用 PSO 算法来训练神经网络，可以有效地克服传统算法收敛速度慢、易

* 基金项目：河南省教育厅软件科学项目(2003790350)

陷于局部极小值等缺点，并且训练出的神经网络在泛化能力上也有很大的提高。起初它是处理连续优化问题的，目前其应用已扩展到组合优化问题。由于其优点在于收敛速度快，设置参数少，简单易实现，因而发展很快，得到了众多学者的重视和研究。

1 粒子群优化算法

1.1 基本原理

设在一个 D 维的搜索空间中，由 n 个粒子组成一个群落。其中第 i 个粒子表示一个 D 维的向量，记为 $X_i=(x_{i1},x_{i2},\cdots,x_{id})$，$i=1,2,\cdots,n$ 即第 i 个粒子在 D 维的搜索空间中的位置是 X_i。每个粒子的位置就是一个潜在的解，将 X_i 带入某个为了满足问题而构造的目标函数中可计算出其适应度值，适应值 $fitness_i=f(X_i)$，然后根据适应度值的大小衡量 X_i 的优劣。第 i 个粒子的速度也是一个 D 维向量，记为 $V_i=(v_{i1},v_{i2},\cdots,v_{iD})$。每次迭代中，粒子通过跟踪两个极值更新自己：

(1)第 i 个粒子迄今为止搜索到的最优位置 $P_i=(P_{i1},P_{i2},\cdots,P_{iD})$，也称为 p_{best}。

(2)整个粒子群迄今为止搜索到的最优位置 $P_g=(p_{g1},p_{g2},\cdots,p_{gD})$，也称为 g_{best}。

PSO 算法的计算公式如下：

$$X_i(m+1)=X_i(m)+V_i(m+1)\Delta t \tag{1}$$

$$V_i(m+1)=wV_i(m)+c_1r_1(P_i-X_i(m))/\Delta t+c_2r_2(P_g-X_i(m))/\Delta t \tag{2}$$

其中 $i=1,2,3\cdots,n$；学习因子 c_1 和 c_2 是非负常数，通常令 $c_1=c_2=2$，调节学习最大步长；r_1 和 r_2 是介于 0 和 1 之间的随机数，增加搜索的随机性；w 为惯性权重，是一个位于区间[0,1]中的常数，调节对解空间的搜索范围；m 为迭代次数；X_i 为第 i 个粒子的位置向量；V_i 为第 i 个粒子速度向量；Δt 为时间间隔，常取单位时间。另外，粒子在每一维的速度 V_{ik}，$k\in\{1,2,\cdots,D\}$ 都被一个最大速度 V_{max} 所限制。如果当前粒子的加速度导致它在某一维的速度超过该维上的最大速度 V_{max}，则该维的速度被限制为最大速度。

迭代终止条件一般选为最大迭代次数和粒子群迄今为止搜索到的全局最优位置满足适应阈值。

1.2 实现步骤

粒子群优化算法的基本实现步骤描述如图 1 所示。

(1)初始化粒子群，设群体规模为 n，在允许的范围内随机设置粒子的初始位置和初始速度；

(2)评价每个粒子的适应度，即计算每个粒子的目标函数 $fitness$；

(3)对每个粒子，比较它的适应值和它经历过的最好位置的适应值，若更好，更新当前最好位置 $P_i=X_i$；

(4)对每个粒子，比较它经历过的最好位置的适应值和群体所经历的最好位置的适应值，若更好，更新全局最好位置 $P_g=P_i$；

(5)根据式(1)和式(2)进化粒子速度和位置；

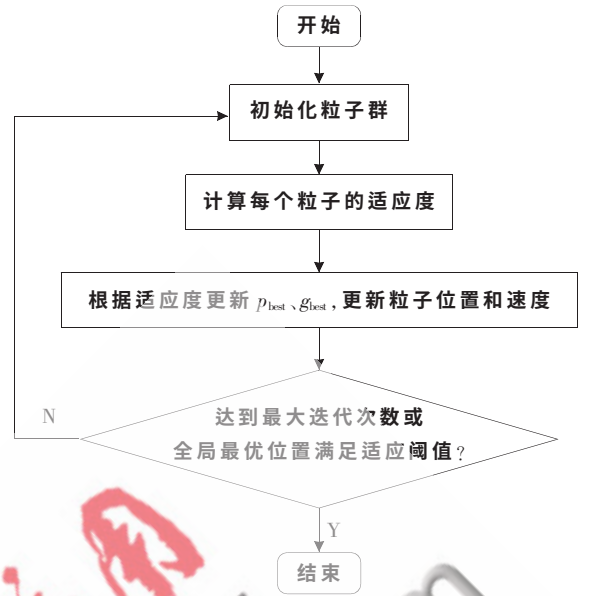

图 1 粒子群优化算法流程图

(6)如果达到结束条件(足够好的位置或最大迭代次数)，则结束，否则转(2)。

1.3 参数设置

粒子群优化算法的参数主要包括：群体规模 n、惯性权重 w、学习因子 c_1 和 c_2、粒子最大速度 V_{max} 以及终止条件。

(1)群体规模 n。即粒子的个数，一般情况下取 20~40。其实对于大部分的问题 10 个粒子已经足够可以取得好的结果，不过对于比较难的问题或者特定类别的问题，粒子数可以取到 100 或 200。

(2)惯性权重 w。它是与前一次速度有关的一个比例因子，而速度更新方程为式(2)，用惯性权重来控制前面的速度对当前速度的影响，较大的 w 可以加强粒子群优化算法的全局搜索能力，而较小的 w 能加强局部搜索能力。基本的粒子群优化算法可以看作 $w=1$。

(3)学习因子 c_1 和 c_2。c_1 和 c_2 通常等于 2，不过在文献中也有其他的取值[6]。但是一般 c_1 等于 c_2 并且范围在 0 和 4 之间。

(4)粒子最大速度 V_{max}。它决定粒子在一个循环中最大的移动距离，通常设定为粒子的范围宽度，例如粒子 (V_1,V_2,V_3)，V_1 属于[-10,10]，那么 V_{max} 的大小就是 20。

(5)终止条件。达到最大迭代次数或全局最优位置满足适应阈值，这个终止条件由具体的问题确定。

2 算法比较与举例分析

2.1 遗传算法与粒子群优化算法的比较

遗传算法和粒子群算法都是作用在由若干个个体(individuals)所构成的群体(population)上的，每个个体代表一个潜在解(solution)，或是潜在解的一部分。每个个体又是由一个特征(characteristics)集合所构成的，正是这些特征将个体映射到某个特定解上。这个特征集合定义

了群体中所有个体所处的一个状态空间。算法的目标是通过个体的适应度，找到状态空间的一个最优解。

遗传算法与粒子群优化算法的不同点在于：(1)遗传算法的个体为染色体(chromosomes)，特征为基因(gene)；而粒子群算法的个体为鸟(bird)或粒子(particle)，特征为位置(position)、速度(velocity)和邻居(neighbourhood-group)。(2)粒子群算法是采用实数编码，遗传算法采用的是二进制编码(或者采用针对实数的遗传操作)。例如对于问题 $f(k)=0.5k_1+k_2^2+k_3^2$ 求解，粒子可以直接编码为 (k_1,k_2,k_3)，而适应度函数就是 $f(k)$。接着就可以利用前面的过程去寻优；而遗传算法则需要将问题的解 k_1,k_2,k_3 映射成0、1组成的位串，然后在位串上进行复制、交换、变异等遗传操作。(3)粒子还有一个重要的特点，就是有记忆。它通过"自我"学习和向"他人"学习，使其下一代能从先辈那里有针对性地继承更多的信息，从而能在更短的时间内找到最优解。

2.2 BP算法与粒子群优化算法的比较

BP算法是一种有监督式的学习算法。其主要思想是：输入学习样本，使用反向传播算法对网络的权值和偏差进行反复地调整训练，使输出的向量与期望向量尽可能地接近，当网络输出层的误差平方和小于指定的误差时训练完成，保存网络的权值和偏差。实质上它是把一组样本输入/输出问题转化为一个非线性优化问题，并通过负梯度下降算法，利用迭代运算求解权值问题的一种学习方法。

BP算法与粒子群优化算法的不同在于：(1)BP算法是严格遵从误差梯度下降的原则调整权值的，当训练进入平坦区后，尽管输出向量与输入向量相差很大，但由于误差梯度小而使权值调整力度减小，训练只能以增加迭代次数为代价缓慢进行；而粒子群算法并不遵从误差梯度降的原则，它通过粒子的自我更新，使其"后代"从先辈那里有针对性地继承更多的信息，从而能在更短的时间内找到最优解。(2)二维权空间的误差曲面像一片连绵起伏的山脉，其低凹部分就是误差函数的极小值点。可以想象，高维权空间的误差曲面"山势"会更加复杂，因而会有更多的极小点。多数极小点都是局部极小，但其特点都是误差梯度为零。误差梯度的这一特点是BP算法无法辨别极小点的性质，因而训练经常陷入某个局部极小点而不能自拔。而粒子群算法通过更新当前每个粒子的最好位置和整个群体的最好位置，较之BP算法最终能够达到一个对全局优化的效果。

2.3 举例与分析

Iris(一种鸢尾属植物)分类问题是用于测试分类器性能的一个著名的基准函数数据集。在数据记录中，每组数据包含Iris花的4种属性：萼片长度、萼片宽度、花瓣长度和花瓣宽度，3种不同的花各有50组数据，这样总共有150组数据。将150组数据分为两个集合：训练集合90组数据与测试集合60组数据。

实验用3层网络对Iris问题来做分类。将网络的输入节点定为4个，输出节点定为3个；每个输入节点代表Iris花的一个属性，而每个输出节点代表一种Iris花。隐层节点和输出节点的激活函数为单极性的Sigmoid函数曲线[6]。即 $f(x)=\frac{1}{1+e^{-x}}$，其中 x 是神经元的加权输入函数。

评价网络分类效果的指标如下：

$$E=1-\frac{en}{N}\times 100\% \tag{3}$$

其中，en 表示错误分类的样本数目，N 表示总的样本数目。

在实验中，PSO算法的参数固定取粒子数为20，$c_1=c_2=2$，$V_{max}=1.2$，w 随迭代次数由0.7线性地减小为0.3，GA的种群数为20，迭代次数取400次。

3种学习算法的实验结果如表1所示。表中GA和PSO算法的分类正确率(即式(3)中的 E)为程序运行10次的平均值。

表1 3种学习算法试验结果比较

算法	进化代数	分类正确率 E(%)
BP算法	600	94.241
GA算法	400	93.182
PSO算法	300	96.254

由表1中的数据可看出，PSO算法的代数小于BP和GA算法的代数，并且PSO算法的分类性能优于BP和GA算法的分类性能。作者还试验当PSO算法的代数与BP算法的代数相同时，PSO算法的分类正确率会提高。PSO算法的运行代数比BP和GA算法的运行代数少，且PSO算法的分类正确率比BP和GA算法的高，可看出PSO算法可克服BP和GA算法易陷入局部极小、收敛速度慢的缺点。由此可得出结论：PSO算法是有效的神经网络学习算法。

文中选择了Iris分类问题对BP算法、标准遗传算法和粒子群算法分别进行测试，并比较其结果，发现粒子群算法通过粒子间的相互作用在复杂搜索空间中发现最优区域。与传统的BP算法和遗传算法相比，采用粒子群算法来训练神经网络，可以有效地克服传统算法收敛速度慢、易陷于局部极小值等缺点，并且训练出的神经网络在泛化能力上也有很大的提高。尽管对PSO的研究已取得了一些进展，但算法本身的工作原理、算法内部机理及其数学基础仍未真正建立，算法中位置和速度的构造及参数的设计理论也不成熟，对参数分析尚处于试验水平，缺乏实质性的认识与理论支持，算法实用性还需进一步提高。

参考文献

[1] RUMELHART D E, HINTON G E, WILLIAMS R J.Learning representations by back propagating errors[J]. Nature, 1986,323(11):533-536.

[2] SEXTON R S, DORSEY R E. Reliable classification using neural networks: a genetic algorithm and backpropagation comparison[J]. Decision Sup2 port Systems,2000,30(1):11-22.

[3] KENNEDY J, EBERHART R. Particle swarm optimization[C]. IEEE International Conference on Neural Networks,Perth, Australia,1995:1942-1948.

[4] EBERHART R, KENNEDY J. A new optimizer using particle swarm theory[C]. Pro of the Sixth International Symposium on Micro Machine and Human Science, Nagoya, Japan,1995:39-43.

[5] 张丽平,陈德钊.粒子群优化算法的理论与实践[D].杭州:浙江大学,2005.

[6] 韩力群.人工神经网络教程[M].第1版.北京:北京邮电大学出版社,2006.

(收稿日期:2009-03-06)

APPLICATION OF ELECTRONIC TECHNIQUE www.chinaAET.com