

# 一种改进的粒子群算法在 BP 网络中的应用研究

江涛 张玉芳 王银辉

(重庆大学计算机学院 重庆 400044)

**摘要** 采用 Sigmoid 激活函数的三层前向神经网络能够以任意精度模拟复杂的非线性关系,训练算法对神经网络模式分类的性能有较大影响。基于梯度下降的 BP 网络存在收敛速度慢、易陷入局部极小的缺陷。粒子群算法是一种全局优化算法。本文针对粒子群算法本身存在的不足加以改进,用改进后的粒子群算法对 BP 网络进行训练,从而克服 BP 网络的一些缺陷。采用 IRIS 分类问题验证了本文提出的方法的有效性。实验结果表明本文采用的方法比普通 PSO-BP 算法效果更好。

**关键词** 粒子群算法,全局优化, BP, 神经网络

## A Study of Application of An Improved PSO Algorithm in BP Network

JIANG Tao ZHANG Yu-Fang WANG Yin-Hui

(Computer School of Chongqing University, Chongqing 400044)

**Abstract** 3-layer forward neural network with Sigmoid activation function can simulate complex non-linear relationship with any precision. The training algorithm of BP affects the classification performance of pattern of the network. BP network with gradient descent has some defects such as low convergence speed, immersing local minimization. Particle Swarm Optimization is a global optimization algorithm. Aiming at PSO's defects this paper gives a method to improve PSO. The improved PSO is used to train BP network to overcome deficiency of BP. IRIS problem is used to validate the efficiency of the method given in this paper. Experimental results prove that this method is better than simple PSO-BP algorithm in performance.

**Keywords** Particle swarm optimization algorithm, Global optimization, BP, Neural network

## 1 引言

BP 网络是一种被广泛应用于模式识别、系统辨识、预测、控制、图像处理等的人工神经网络模型。但 BP 算法是基于梯度的方法,存在收敛速度慢、容易陷入局部极小等缺点。目前的研究针对其缺陷提出了不少改进方法。粒子群优化方法(Particle Swarm Optimization, PSO)是一种基于群体智能理论的优化算法,通过群体中粒子间的合作与竞争产生的群体智能指导优化搜索。PSO 采用速度-位移模型,操作简单<sup>[9]</sup>,易于编程,被广泛应用于求解全局优化问题。

## 2 改进的 PSO 算法(Improved PSO, IPSO)

粒子群优化算法(Particle Swarm Optimization, PSO)是由 Kennedy 和 Eberhart 于 1995 年提出的一种新的进化算法<sup>[4]</sup>,该算法来源于对鸟群体觅食行为的研究,其理论依据是演化计算理论。鸟群在随机搜索食物过程中,如果某区域里只有一块食物,那么找到食物的最简单有效的策略是搜索目前离食物最近的鸟的邻近区域。PSO 就是从这种模型中得到启发而产生的。粒子群算法通过个体之间的协作来寻找最优解。粒子群算法可用于解决大量非线性、不可微和多峰值的复杂问题的优化,已经被广泛应用于如函数优化、神经网络、模式分类等科学和工程领域,大量实际应用证明其是有效的。粒子群优化算法的收敛速度快(尤其是在进化初始阶段),运算简单,易于实现,没有遗传算法的编解码和杂交、变异等复杂运算。但粒子群优化算法也存在不足,该算法是根据全体粒子和自身的搜索经验向最优解的方向运动,在进化

后期存在速度变慢的现象以及过早收敛即早熟。

PSO 算法初始时随机产生目标函数的一组解,该解空间内的每个解被称作粒子。每个粒子通过跟随当前最优粒子的方法在求解问题的空间内搜索最优解。按照式(1)和(2)更新每个粒子的位置和速度:

$$v(i+1) = \omega \times v(i) + c_1 \times r_1 \times (pbest - x(i)) + c_2 \times r_2 \times (gbest - x(i)) \quad (1)$$

$$x(i+1) = x(i) + v(i+1) \quad (2)$$

式(1)中的  $pbest$  表示该粒子从开始到现在搜索产生的最优解,  $gbest$  表示粒子群目前找到的最优解,即运行到当前代为止所找到的全局最优。 $\omega \in [0, 1]$ , 被称作惯性因子,用于根据上一次迭代得到的速度调节本次迭代过程中粒子的运动速度。 $c_1$  和  $c_2$  是加速系数,分别调节向全局最好粒子和个体最好粒子方向飞行的最大步长,若太小,则粒子可能远离目标区域,若太大则可能会导致飞过目标区域。 $r_1$  和  $r_2$  是  $(0, 1)$  之间的随机数。根据目标函数计算的适应度值用来衡量搜索空间内哪个位置更好。适应度值驱使各粒子朝着搜索空间内个体最优和到目前为止发现的全局最优方向移动。

从式(1)可以看出:当一个粒子的当前位置接近但并未到达全局最优位置时,如果该粒子的上一次迭代中的运动速度和惯性权值非零,则该粒子将渐渐运动远离该位置。另外,一旦所有粒子追赶上当前全局最优粒子则这些粒子将停止移动。该现象被称为早熟,即算法过早收敛。

为避免早熟,考虑在算法搜索过程中引入一些干扰因素。算法过早收敛的发生有两个必要条件:(1)各粒子追赶上全局最优粒子;(2)粒子的速度迅速降低至接近零。惯性权值接近

零时对速度也有影响。如果在迭代搜索的过程中惯性权值被设为随机的接近零,该影响可忽略。对一个随机搜索过程来说,条件(1)是不能避免的,原因有两点:a)循环过程中的每一步是不能控制的;b)当找到全局最小时该情况也会发生。因此避免 PSO 算法的早熟就考虑破坏条件 2 出现的可能。

当早熟发生时,为粒子的速度加上干扰条件。描述如下:

$$v(i) = v(i) + \text{random} \quad (3)$$

$$x(i) = x(i) - 2v(i) \quad (4)$$

当早熟发生时,式(3)为速度增加一个随机数,从而改变了该粒子的运动速度,式(4)将粒子移动到一个新的位置,改变搜索方向。式(3)适合单模函数,式(4)适合多模函数。式(3)和(4)描述的干扰使得粒子在新的位置以新的速度开始搜索,跳出了原早熟区域,避免了局部最优解,能够更好地找到全局最小。

当过早收敛发生时,干扰被引入。在搜索过程中引入搜索因子来发现早熟发生的时机,该因子被定义为  $f_m$ 。迭代搜索过程中记录下每一个有效搜索步骤,每一有效搜索步骤的输出记为  $f_g$ 。当迭代过程中某一步的输出大于早熟因子(即  $f_g > f_m$ )时,式(3)和(4)描述的干扰产生作用,破坏过早收敛的条件。早熟因子  $f_m$  根据目标函数的不同取不同值。

### 3 训练 BP 网络

BP 算法的训练是一种有监督的方法,采用误差反向传播的方法使网络的权值随机的沿梯度下降。每个权值都会影响输出层产生的误差,因此为了减少该误差,要根据输出层的误差调整权值矩阵,形成将输出端表现出的误差沿着与输入信号相反的方向逐渐向输入端传递的过程。

将改进后的 PSO(Improved PSO, IPSO)用于训练 BP 网络的方法是:粒子群中每个粒子的位置表示 BP 网络中当前迭代中的权值集合,每个粒子的维数表示网络中起连接作用的权的数量,以给定训练样本集的神经网络输出误差作为神经网络训练问题的适应函数,适应度值表示神经网络的误差阈值,误差越小则表明粒子在搜索中具有更好的性能。粒子在权值空间内移动搜索使得网络输出层的误差最小,改变粒子的速度即更新网络的权值,以减少 MSE。通过这种方式,PSO 优化搜索训练神经网络的权值来获得更小的输出误差。在每一代过程中,所有粒子根据计算出的新的速率更新位置,向新的方位移动,新的位置即新的权值集合,根据该权值集合得到新的 MSE。如果某粒子运动到的新方向使得新的 MSE 不能减小,则粒子不朝该新方向移动,即丢弃新的权值矩阵。所有粒子重复该操作。每次迭代过程中产生 MSE 最小的粒子为目前全局最优粒子,即 PSO 算法中的 gbest。该训练过程不断重复直到产生满足要求的误差或者超过预先设定的计算次数为止。算法结束时得到的权值集合即最终结果。

从上面的描述可以看出:采用了改进后的 PSO 算法对 BP 网络进行训练,避免了原 BP 算法中的向后传播过程,加快了搜索速度并且能够防止整个算法的过早收敛。

算法步骤:

1)初始化参数,包括种群大小、加速系数、学习率、迭代次数等

2)如果  $mse \leq \epsilon$  或者迭代次数超过上限,算法结束

3)对每个粒子  $i$ ,

在网络中正向传播,计算其在输出层的输出,得到误差  $E_i$ ,如果  $E_i < pb_i$ ,则  $pb_i = E_i$ ,  $p_i = x_i$ ,如果  $E_i < gb_i$ ,则  $gb_i = E_i$ ,  $p_g = x_i$

4)判断是否出现早熟

如果是,则按式(1)、(3)更新每个粒子的速度  $v_i$ ,按式(2)、(4)更新粒子的位置  $x_i$

否则按式(1)和(2)分别计算并更新每个粒子的速度和位置

5)转 2)

6)结束

针对粒子群存在的过早收敛现象,采用改进后的带干扰条件的粒子群方法训练的 BP 网络在产生下一代解集上具有非常大的随机性,所以不容易陷入局部极小,而且由于每代所有解信息的共享性和各个解的质量的提高,使得每代种群中的解具有“自我”学习提高和向“他人”学习的双重优点,所以具有较快的收敛速度,在算法的后期对新个体还进行训练优化,更提高了收敛速度。由于不存在随机寻找最优的退化现象,因此与采用遗传算法训练的 BP 网络相比,后期收敛相对更平稳,波动现象大大减少。与使用普通 PSO 训练的 BP 网络比较,也具有更好的寻优能力和更好的泛化性能。

### 4 实验

IRIS 数据库是广泛使用的模式分类实例系统。本文采用 IRIS 来验证改进后的算法效率。IRIS 含有 150 个例子,分为 Setosa, Versicolor 与 Virginica 三类,每类包含 50 个例子,由四个实数特征值描述,分别表示花萼(Sepal)长度、花萼宽度、花瓣(Petal)长度、花瓣宽度。模式分类问题根据特性值区分三种 IRIS 植物<sup>[8]</sup>。取 101 组数据用于训练,49 组用于测试训练后的网络的分类性能。

网络结构为 4-3-3,包括偏差权在内总共 27 个权值,粒子群大小为 10,  $c_1$  和  $c_2$  均设为 2,惯性权值从 0.7 到 0.3 线性下降,网络最大迭代次数为 4000,因此向前传播的总数量为:种群数 \* 最大迭代数 = 40000。

参照验证 BP 算法性能常用的指标。实验考察 4 个指标:  $MSE_G$ 、 $MSE_T$ 、 $\epsilon_G$  和  $\epsilon_T$ 。  $MSE_G$  是在测试样本上得出的 MSE,  $MSE_T$  是在训练样本得出的 MSE,  $\epsilon_G$  和  $\epsilon_T$  分别是由训练样本集和测试样本得出的误差,其表达式为:

$$\epsilon_G = \frac{\text{训练集中分类错误的数量}}{\text{训练样本个数}} \times 100$$

$$\epsilon_T = \frac{\text{测试集中分类错误的数量}}{\text{测试样本个数}} \times 100$$

	BP	PSO-BP	IPSO-BP
$MSE_T$	0.0508	0.0106	0.0130
$MSE_G$	0.0586	0.0269	0.0184
$\epsilon_T$	3.4	2.35	2.45
$\epsilon_G$	4.5	5.1	3.0

图 1 误差结果

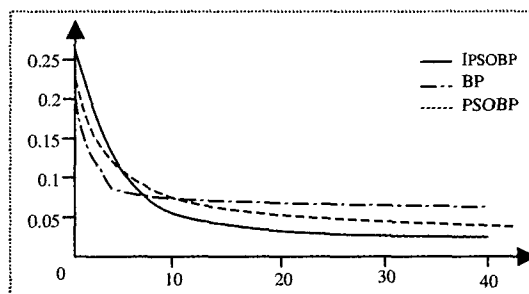


图 2 误差随时间变化情况

JXTA的主要技术是:通过 Peer 元素组建 Peer Group;提供一套通用服务、协同工作的 Peer 集合;通过 Pipe 通道传输 Message;由一个 XML 结构文档 Advertisement 来命名、描述或发布一个资源。JXTA 技术实际上就是一组协议,每一种协议定义参加协议的各个成员间如何交换信息。其中 PDP(Peer Discovery Protocol)是核心协议之一,具有网络发现、被动发现、瀑布式发现、通过集合点发现;使用 PRP(Peer Resolver Protocol)协议来向其他 Peer 发出通用查询消息,并获得查询结果;通过 PIP(Peer Information Protocol)获知其他 Peer 的状态、运行时间、负载、能力等;通过 PMP(Peer membership protocol)协议使 Peer 能自己组成一个或加入一个 Peer 组;通过 PBP(Peer Biding Protocol)协议使 Peer 能建立一个或多个连接管道。

本系统远程教育软件平台在内容方面主要实现以下几个功能模块。

(1)教学资料模块:主要包括教学大纲模块、课件下载模块、教学理论模块,该模块主要是针对教师的应用。教师可以根据需要上传教学大纲、课件等相关资料,并可以对上传的资料进行相应的编辑、修改等操作。同时,教师可以相互交流教学理论,探讨教学方法。

(2)网络课堂模块:主要包括双向视频课堂(包括视频点播)模块、网上课程模块和电子教材模块 3 个部分,该模块为客户提供网络教学环境,为用户提供点播源和广播源的制作,可以根据需要更新、调整教学视频流。

(3)在线考试模块:主要包括试题集锦、试题答案、在线分析 3 个部分,采用 ASP 将脚本、HTML 和数据库访问功能相结合,为客户提供在线考试环境,并提供试题答案和在线分析

等相关内容。

(4)疑难解答模块:主要包括视频交流、专题讨论、疑难分析 3 个子模块,学生可以提出相关的问题,教师针对学生提出的问题进行分析探讨,同时教师可以针对相关专题进行讨论研究,充分体现网络教学的互动性。

**结论** 本系统设计上采用基于 P2P 的分布式视频教学系统结构。通过合理使用用户计算机空闲的资源提供部分服务,使得整个系统的能力得到极大的提高。网络中的资源和服务分散在所有节点上,体现了 P2P 的非中心化;整个体系是全分布的,随着用户的加入,不仅服务的需求增加了,系统整体的资源和服务能力也在同步地扩充,体现了可扩展性;由于服务是分散在各个节点之间进行的,部分节点或网络遭到破坏对其它部分的影响很小,还能够根据网络带宽、节点数、负载等变化不断地做自适应式的调整,体现了健壮性;由于每个节点既是服务器又是客户机,更好地实现了整个网络的负载均衡。随着用户的增加,系统容量扩大,带宽资源增加,系统能力也不断地增强。

## 参考文献

- 1 魏苏林,等.基于 P2P 模式的新型视频点播系统.微计算机信息,2005(12)
- 2 葛苏慧,等.IP 多播技术在网络教学系统中的应用探究.教育信息化,2005(10)
- 3 方炜,等.基于 P2P 的流媒体应用及其关键算法研究.计算机应用与软件,2005(5)
- 4 易光华,等.基于对等网络的 VOD 模型研究.网络安全技术与应用,2005(2)

(上接第 165 页)

从实验结果看,用改进后的 PSO 对 BP 进行训练能够避免由于 PSO 本身的缺陷导致在 BP 训练过程中不能更好地找到最小解,减少了训练时间,能够有效地进行全局寻优,使得出的误差更小。

**结论** 本文分析了 PSO 算法的早熟现象,并进行了改进。改进后的 PSO 算法被用于 BP 网络的训练,在克服局部极小,加快收敛速度方面有显著的效果。通过 IRIS 问题验证了使用改进后的 PSO 作为 BP 网络的训练算法的有效性。

## 参考文献

- 1 van den F B, Engelbrecht AP. Cooperative Learning in Neural Networks using Particle Swarm Optimizer. In: Proc. SAICSIT, 2000
- 2 Zhang Chunkai, Shao Huihe, Li Yu. Particle Swarm Optimization for Evolving Artificial Neural Network. IEEE
- 3 Clerc M, Kennedy J. The Particle Swarm- Explosion, Stability, and Convergence in A Multidimensional Complex Space [J]. IEEE Trans. on Evolutionary Computation, 2002, 6 (1):58~73
- 4 Kennedy J, Eberhart R C. Particle Swarm Optimization. Proc. IEEE International conference on Neural Networks, 1995. 1942~

1948

- 5 Shi Y H, Eberhart RC. Empirical study of particle swarm optimization [A]. In: Proceedings of IEEE International Congress on Evolutionary Computation [C]. USA: IEEE, 1999. 6~9
- 6 Yang J M, Kao C Y. A robust evolutionary algorithm for training neural networks [J]. Neural Computing and Application, 2001, 10(3):214~230
- 7 Van den B F, Engelbrecht A P, Training Product Unit Networks Using Cooperative Particle Swarm Optimizers [C]. In: Proc. of the third Genetic and Evolutionary Computation Conference, 2001
- 8 Maniezzo V. Genetic evolution of the topology and weight distribution of neural networks. IEEE Trans. Neural Networks, 1994, 5(1):39~53
- 9 周弛,高海兵,高亮,等.粒子群优化算法[J].计算机应用研究,2003,(12):7~12
- 10 李孝安,张晓缙.神经网络与神经计算机导论[M].西北工业大学出版社,1995
- 11 侯志荣,吕振肃.基于 MATLAB 的粒子群优化算法及其应用[J].计算机仿真,2003,20(10):68~70
- 12 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