

基于改进 PSO 算法的 BP 神经网络的应用研究

曾万里¹, 危韧勇¹, 陈红玲²

(1. 中南大学 信息科学与工程学院, 湖南 长沙 410075;

2. 吉林大学 珠海学院, 广东 珠海 519041)

摘要: 为了克服粒子群优化算法本身存在的早熟和局部收敛的固有问題, 在描述了 BP 神经网络的基本结构的基础上, 介绍了粒子群优化算法(PSO)的基本概念, 并通过对二者优缺点的分析与比较, 结合二者的优势, 将粒子矢量位移应用到 PSO 算法中, 并在此基础上, 用改进的 PSO 算法对 BP 网络进行训练, 还利用某商场的部分消费数据进行了实验。结果表明, 基于改进的 PSO 算法的 BP 网络在收敛速度和精度上都比基于传统的 PSO 算法好。

关键词: BP 网络; 粒子群优化算法; 矢量位移; 收敛速度

中图分类号: TP301.6

文献标识码: A

文章编号: 1673-629X(2008)04-0049-03

Research and Application of BP Neural Network Based on Improved PSO Algorithm

ZENG Wan-li¹, WEI Ren-yong¹, CHEN Hong-ling²

(1. Sch. of Info. Sci. & Eng., Central South University, Changsha 410075, China;

2. School of Zhuhai, Jilin University, Zhuhai 519041, China)

Abstract: Describes the basic structure of BP neural network in order to solve the inhere problem of precocity and refraining partly. Some basic concepts of particle swarm optimizer are introduced also. Combines the advantages of them through comparing and analyzing some problems concerned to them. What's more, the concept of particle's vector shift is used in algorithm PSO. Based on this, improved algorithm PSO is used to train BP neural network. This experiment is done with a market's consumption data. As a result, it is better to use improved algorithm PSO than traditional algorithm PSO or BP alone not only in velocity but in precision of convergence.

Key words: BP neural network; particle swarm optimizer; vector shift; velocity of convergence

0 引言

近年来, BP 网络广泛应用于诸多领域并取得了很好的效果。但在实际应用中, BP 网络也暴露了一些自身固有的缺陷: 不具有全局搜索能力; 易陷入局部极小; 收敛时间过长; 结果存在一定随机性等。粒子群优化(Particle Swarm Optimizer, PSO) 算法是一类新兴的随机全局优化技术。它可以弥补 BP 网络存在的一些不足。因此, 人们考虑到了把二者结合起来, 争取做到优势互补。文献[1]研究了基于粒子群优化的 BP 网络学习算法取得了良好的效果, 但是, 由于粒子群优化算法本身存在早熟和局部收敛的问题, 因此, 基于粒子群优化的 BP 网络学习算法还有待于进一步地改进。

提出了一种基于改进粒子群优化的 BP 网络学习算法, 并分别与 BP 算法及基于粒子群优化的 BP 算法进行了对比实验, 实验取得了令人满意的结果, 说明该改进的算法可以避免 BP 网络及传统算法的固有缺陷, 充分发挥二者的优势, 达到更好的优化效果。

1 标准 PSO 算法

微粒群算法(Particle Swarm Optimization, PSO)^[2]是由美国社会心理学家 Kennedy 和电气工程师 Eberhart 于 1995 年共同提出的一种生物进化算法^[3]。PSO 算法采用速度—位置搜索模型, 将微粒比做一个没有质量和体积但具有速度和位置的点, 粒子在搜索空间以一定的速度飞行, 飞行速度根据飞行经验进行动态调整^[4]。群体中(共 m 个粒子) 第 i 个粒子在 D 维解空间的位置表示为 $x_i = (x_{i1}, x_{i2}, \dots, x_{iD})$, 其中 $i = (1, 2, \dots, m)$, $d = (1, 2, \dots, D)$; 其速度为 $v_i = (v_{i1}, v_{i2}, \dots, v_{iD})$ 。将 x_i 带入目标函数可计算出其适应值, 第 i

收稿日期: 2007-07-02

作者简介: 曾万里(1979-), 男, 湖南邵东人, 硕士研究生, 研究方向为数据仓库与数据挖掘技术; 危韧勇, 教授, 硕士生导师, 研究方向为电力谐波综合治理技术、铁道牵引电气化与自动化、计算机信息处理与控制等。

个粒子搜索到的最好位置为 $P_i = (P_{i1}, P_{i2}, \dots, P_{id})$, 称为 P_{best} , 整个粒子群搜索到的最优位置为 $P_g = (P_{g1}, P_{g2}, \dots, P_{gd})$, 整个粒子群通过速度和位置的更新来搜索整个状态空间:

$$v_{id} = wv_{id}(t) + c_1 \text{rand}() (P_{id} - x_{id}) + c_2 \text{rand}() (P_{gd} - x_{id}) \quad (1)$$

$$x_{id} = x_{id} + v_{id} \quad (2)$$

其中: w 为惯性权重 (inertia weight), 通常取值为 $0.4 \sim 1.2$; c_1 和 c_2 为加速常数 (acceleration constants), 也称学习因子; $\text{rand}()$ 为在 $[0, 1]$ 范围内变化的随机函数。

2 改进的 PSO 算法

通过分析标准 PSO 算法可知, 惯性权重 w 具有维护全局和局部搜索能力的平衡的作用, 当 w 较大时, PSO 算法具有全局收敛性, 但运算量很大; 当 w 较小时, 算法容易收敛, 但容易陷入局部最优^[5]。加速常数 c_1 和 c_2 分别用来调节向个体最好粒子和全局最好粒子方向运动的最小单位, 如果太小, 则粒子可能远离目标, 如果太大, 则粒子可能跳过目标。于是有的学者提出利用模糊系统来调整 w , 可是这样会增加系统的随机性和不可控性, 并将会使算法变得更加复杂。

因此, 文中提出基于矢量位移的 PSO 算法来调整 w 的策略。该算法就是让惯性权重 w 随粒子朝目标不断地靠近而线性地变化, 这将使得 PSO 算法在初期具有较强的全局收敛能力, 而晚期具有较强的局部收敛能力^[6]。设 w_{\max} 为最大惯性权重, w_{\min} 为最小惯性权重; x_d^{\min} 、 x_d^{\max} 分别为粒子在第 d 维位置的上、下界。 w 按下式进行动态调节:

$$w = w_{\max} - \frac{x_{id} - x_d^{\min}}{x_d^{\max} - x_d^{\min}} (w_{\max} - w_{\min}) \quad (3)$$

由于粒子的不断运动, 它与目标的矢量距离逐渐缩小, 而 $x_{id} - x_d^{\min}$ 逐渐增大, 由式(3)可看出, w 也会跟着线性下降。因此, 算法在开始时搜索较大的区域, 较快地确定最优解的大致位置。随着 w 的减小, 粒子的速度减慢, 开始精细的局部搜索。这种方法加快了收敛速度, 提高了 PSO 算法的性能。

3 基于改进 PSO 算法的 BP 网络学习

3.1 改进 PSO 算法与 BP 网络的结合

①BP 网络结构。BP 网络是通过迭代地处理一组训练样本, 将每个样本的网络预测值与实际值比较, 进行学习, 如果未满足条件, 则不断地修改权值, 继续进行学习, 直至最终收敛或超出允许的范围。这与 PSO

算法中粒子通过修改惯性权重和加速度权重不断地朝目标运动非常相似。

②编码。用粒子群中每个粒子的位置表示 BP 网络中当前迭代的权值集合, 用每个粒子的维数表示网络中权的数量, 以给定训练样本集的输出误差作为适应函数, 以适应度值表示误差阈值, 误差越小则表明粒子在搜索中具有更好的性能。

③动态性。动态性不仅表现在二者的学习过程和运动过程都要不断地修改权值, 还表现在二者最后的结果收敛性具有不确定性。

3.2 基本思想

BP 网络是一种多层前馈网络, 通常由输入层、输出层和若干个隐层构成。对 BP 网络进行训练, 也即对网络的连接权系数进行学习和调整, 以使网络实现给定的输入输出映射关系。

基于改进的 PSO 算法的 BP 神经网络的具体训练方法如下:

在每一次迭代过程中, 所有粒子根据计算出的新的速率更新位置, 向新的方向移动, 新的位置即新的权值集合, 根据该权值集体得到新的网络误差。为了使网络输出层的误差最小, 粒子在权值范围内不断地运动, 也就是通过移动搜索不断地更新网络权值^[1]。

3.3 算法步骤

①初始化种群;

②设立终止准则, 符合条件即终止, 否则执行下一步;

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, 并比较得到误差;

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 , 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。

基于改进 PSO 的 BP 网络流程图如图 1 所示。

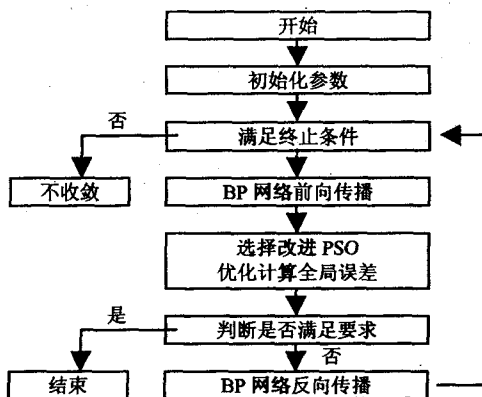


图 1 基于改进 PSO 的 BP 网络流程图

4 实验

随着经济的不断增长,商业在整个国民经济中所占的比重也越来越高,在以消费带动经济增长的政策理念下,客户的消费行为显得尤为重要。BP 网络已广泛应用于模式识别、质量检测、风险预测、图像处理等领域。但将 BP 网络用于客户消费行为的分析和预测尚不多见。

文中从某商场的消费数据中选取一部分作为样本对网络进行训练,同时设定运算次数为 2000,允许的精度为 0.001。其中对顾客消费行为影响比较大的几个因子如表 1 所示。BP 网络训练结果如表 2 所示,由表可以看出,训练结果基本上能达到所需的目的。但比较三种不同方法的预测值不难发现,也就是采用改进的 PSO 算法来训练 BP 网络可以达到更高的精度。

表 1 影响顾客消费行为的 4 个影响因子

RID	age	Income	Student	Credit-rating
1	<= 30	High	No	Fair
2	<= 30	High	No	Excellent
3	31~40	High	No	Fair
4	>40	Medium	No	Fair
5	>40	Low	Yes	Fair
6	>40	Low	Yes	Excellent
7	31~40	Low	Yes	Excellent

BP 网络的收敛情况如图 2 所示,很快可以发现,图中光滑曲线实线收敛最迅速,也就是采用改进 PSO

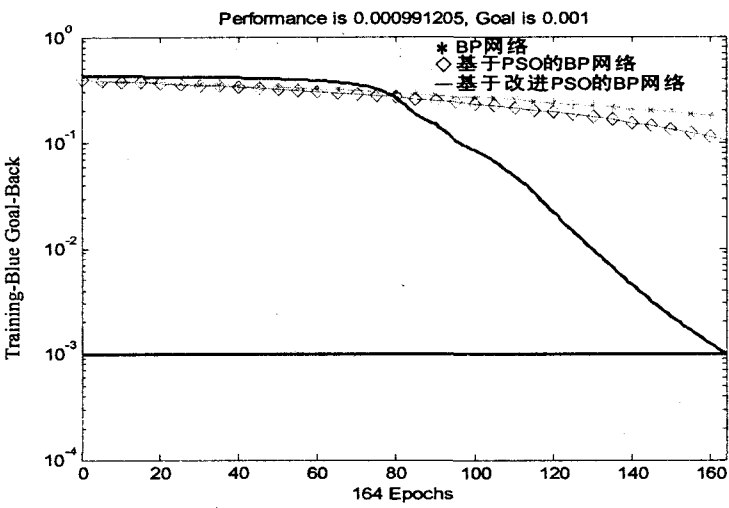


图 2 BP 网络的收敛情况

算法来训练 BP 网络可使收敛速度更快。

通过实验可以看出,采用改进后的 PSO 算法对 BP 网络进行训练能有效地避免 PSO 算法和纯 BP 网络本身固有的缺陷,充分发挥二者的优势,从而做到快速收敛,减少训练时间,提高精度,减少误差。

5 结 语

标准 PSO 算法往往不能兼顾收敛速度、全局探索能力和局部精细搜索能力。文中将粒子位移的矢量距离引入 PSO,并以此为基础,对粒子群优化算法加以改进,并用某商场的部分消费数据为样本,结合 BP 网络进行训练。结果表明,基于改进的 PSO 算法的 BP 网络在收敛速度和精度上都要比纯 BP 网络或基于传统 PSO 算法的 BP 网络要好。

表 2 采用三种不同方法的预测值比较

实际值	0	1	1	1	1	1	0
BP 网预测值	0.15376	0.70204	0.81763	0.67492	0.71082	0.80731	0.20491
PSO 算法 BP 网预测值	0.081241	0.9265	0.87212	0.89527	0.78826	0.92375	0.14387
改进 PSO 算法 BP 网预测值	0.031225	0.99985	0.97275	0.99828	0.94485	0.98999	0.043534

参考文献:

[1] 江 涛,张玉芳,王银辉.一种改进的粒子群算法在 BP 网络中的应用研究[J]. 计算机科学,2006,33(9):164-190.

[2] 熊伟丽,徐保国,周其明.基于改进粒子群算法的 PID 参数优化方法研究[J]. 计算机工程,2005,31(24):41-43.

[3] Kennedy J, Eberhart R. Particle Swarm Optimization[C]// In:Proc. of IEEE Int. Conf. on Neural Networks. Perth:[s. n.],1995:1942-1948.

[4] 延丽平,曾建潮.具有自适应随机惯性权重的 PSO 算法[J]. 计算机工程与设计,2006,27(24):4677-4706.

[5] 孙红光,潘毓学.基于运动目标路径的粒子群优化算法研究[J]. 仪器仪表学报,2004,25(4):946-947.

[6] Parsopoulos K E, Vrahatis M N. Recent approach to global optimization problems through particle swarm optimization[J]. Natural Computing,2002,1(2):235-306.