

文章编号: 1001-0920(2005)08-0958-03

基于复合适应度微粒群算法的神经网络训练

赵 辉^{1,2}, 刘鲁源¹, 张更新²

(1. 天津大学 电气与自动化工程学院, 天津 300072; 2 天津理工大学 自动化学院, 天津 300191)

摘 要: 为提高神经网络的泛化能力, 针对以均方误差为适应度的 PSO 算法在训练神经网络时会产生一定的过拟合问题, 提出对均方误差和误差分布均匀度进行信息融合, 构成复合适应度作为训练指标。实验结果表明, 该方法可使网络的泛化能力得到明显的改善。

关键词: 微粒群算法; 神经网络; 复合适应度; 泛化能力

中图分类号: TP18 **文献标识码:** A

Neural Network Training Based on Compound Fitness Particle Swarm Optimization

ZHAO Hui^{1,2}, LIU Lu-yuan¹, ZHANG Geng-xin²

(1. School of Electrical Engineering and Automation, Tianjin University, Tianjin 300072, China; 2 College of Automation, Tianjin University of Technology, Tianjin 300191, China Correspondent: ZHAO Hui, Email: hzhao3379@sina.com)

Abstract: The over fitting arises if the PSO (particle swarm optimization) algorithm whose fitness is mean square deviation is applied in training neural network. In order to improve generalization capacity of feedforward neural network. The compound fitness based on information merging of mean square deviation and error uniformity is proposed as the training index of PSO. The results show that the approach can improve the generalization capacity of feedforward neural networks remarkably.

Key words: Particle swarm optimization; Neural network; Compound fitness; Generalization ability

1 引 言

为使所构造的网络对训练样本集输出与输入的映射关系具有足够的描述能力, 构造具有足够泛化能力的网络是神经网络训练的根本要求。目前, 在神经网络的应用和研究中, 存在的主要问题是: 如何有效提高网络的泛化能力和训练速度问题。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]。

近年来, 随着进化计算研究的兴起, 人们在探讨采用进化计算算法训练神经网络的可能性。进化算

法具有较强的收敛能力, 不需要借助于问题的导数等梯度信息, 用它训练神经网络, 不仅能提高神经网络的泛化能力, 而且能提高神经网络的收敛速度和学习能力^[2]。微粒群算法(PSO)是一种新的进化算法, 它比遗传算法(GA)能更快地达到全局最优, 尤其是在解决多峰值问题上更具优越性。

PSO 算法的速度模型进化方程^[3]为

$$v_{ij}(t+1) = \omega v_{ij}(t) + r_1 C_1 (p_{ij} - x_{ij}(t)) + r_2 C_2 (g_j - x_{ij}(t)), \quad (1)$$

其中: r_1 和 r_2 为 0-1 随机数, ω 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} 为个体微粒所达到的最好位置,

收稿日期: 2004-09-13; 修回日期: 2004-12-14

基金项目: 天津市自然科学基金重点项目(033803311); 天津市高等学校科技发展基金项目(20041705)。

作者简介: 赵辉(1963—), 男, 天津人, 教授, 博士生, 从事智能控制理论及应用等研究; 刘鲁源(1941—), 男, 山东青岛人, 教授, 博士生导师, 从事控制理论、计算机技术等研究。

g_j 为所有微粒经历过的最好位置 在得到微粒新的速度后, 便可得到新的位置

$$x_{ij}(t+1) = x_{ij}(t) + v_{ij}(t+1). \quad (2)$$

在训练求解分类和函数逼近问题的前馈神经网络时, 采用 PSO 算法训练网络所达到的精度和泛化能力要优于基于梯度的学习算法和遗传算法^[4], 但从实验结果看, PSO 算法训练的网络也存在过度拟合问题, 即测试集上的均方误差大于训练集上的均方误差^[5]. 为有效克服这一问题, 本文在均方误差指标的基础上融合了误差均匀度, 形成复合适应度作为 PSO 算法的训练指标, 使神经网络具有更强的泛化能力

2 基于 PSO 算法的神经网络训练

2.1 PSO 算法优化神经网络流程^[6]

1) 将网络中所有神经元间的连接权编码成实数码串表示的个体, 若网络中包含 M 个优化权值, 则每个个体由 M 个权值参数组成的 M 维向量表示;

2) 初始化微粒群, 随机产生一定数目的微粒组成种群, 并初始化全局最好位置 g_{best} 和个体最好位置 l_{best} ;

3) 按照 PSO 模型进化产生新的微粒, 并对微粒适应度进行评价, 得出新的全局最好位置 g_{best} 和个体最好位置 l_{best} ;

4) 若未到达精度要求, 则返回步骤 3).

2.2 微粒的评价

由于每个微粒代表一组网络权值, 对微粒的评价就是对该组权值在训练集上产生的均方误差的评价. 可将均方误差作为 PSO 算法的适应度^[7]

$$E(X) = \frac{1}{2n} \sum_{p=1}^n \sum_{k=0}^c (Y_{k,p}(X) - t_{k,p})^2. \quad (3)$$

其中: X 为微粒向量, $t_{k,p}$ 为训练样本 p 在 k 输出端的给定输出, n 为训练集样本个数, c 为输出端个数, $Y_{k,p}$ 为微粒对应的网络输出

对于网络的泛化能力, 可用测试集上的均方误差来评价

$$F(X) = \frac{1}{2m} \sum_{p=1}^m \sum_{k=0}^c (Y_{k,p}(X) - t_{k,p})^2, \quad (4)$$

其中 m 为测试集样本个数

2.3 基于信息融合的 PSO 算法的复合适应度

在神经网络的学习过程中, 通常都是以训练集上的均方误差作为指标^[8], 但对同样的均方误差, 各样本点误差的平方与均方误差的偏差是不同的. 如果每个样本点的误差平方与均方误差都很接近, 则说明该网络在误差分布上具有良好的均匀度^[9]. 在同样的均方误差下, 误差分布均匀度好的网络在泛化能力上优于均匀度差的网络. 本文综合考虑均方

误差和误差均匀度 (EU), 形成复合适应度 (CF), 并将其作为 PSO 算法训练神经网络的指标. 误差均匀度为

$$EU = \eta \left| \frac{1}{2n} \sum_{k=0}^c (Y_{k,p}(X) - t_{k,p})^2 - E(X) \right|, \quad (5)$$

复合适应度为

$$CF = EU + E(X). \quad (6)$$

其中 η 为误差均匀度系数, 取值范围一般为 (0 ~ 1.5). 为提高 PSO 算法的效率, 可使 η 随进化代数线性增大. 这里对误差均匀度采用的是绝对值而不是均方的形式. 这是因为每个样本点的误差在 EU 中已有过一次平方, 若再使用均方的形式将会形成四次方, 使大于 1 的误差过大, 小于 1 的误差过小, 从而使均匀度失去意义.

2.4 算法实验

为检验 PSO 算法在提高神经网络泛化能力上的效果, 以 $f(z) = z^3 - 0.05z$ 的函数逼近问题作为检验实例. 其中训练集取 0, 0.2, 0.4, ..., 共 20 个点; 测试集取 0.1, 0.3, 0.5, ..., 共 20 个点. 微粒数取 200, 惯性权重按进化代数由 0.9 线性减少至 0.4^[5], c_1 和 c_2 取 1.5. 前馈神经网络结构取 1-4-1 结构, 隐层激励函数为 Sigmoid 函数^[6]

$$\mathcal{Q}(v) = \frac{1}{1 + \exp(-0.2v)}. \quad (7)$$

分别以均方误差 E 和复合适应度 CF 为适应度进行计算, 实验结果如表 1 (分裂因子为 2) 和表 2 (分裂因子为 3) 所示. 其中, 梯度训练算法、遗传训练算法和微粒群算法以训练集和测试集上的两种均方误差作为性能指标, 复合适应度微粒群算法以本文提出的复合适应度作为训练指标, 对比了不同算

表 1 实验结果 1

| 训练算法 | 训练集均方误差 | 测试集均方误差 | 测试集均方差 - 训练集均方差 |
|------------|-----------------------|-----------------------|-----------------------|
| 梯度训练算法 | 1.00×10^{-4} | 2.19×10^{-4} | 1.19×10^{-4} |
| 遗传训练算法 | 1.17×10^{-5} | 2.38×10^{-5} | 1.21×10^{-5} |
| 微粒群算法 | 9.45×10^{-6} | 2.39×10^{-5} | 1.45×10^{-5} |
| 复合适应度微粒群算法 | 2.18×10^{-6} | 3.22×10^{-6} | 1.04×10^{-6} |

表 2 实验结果 2

| 训练算法 | 训练集均方误差 | 测试集均方误差 | 测试集均方差 - 训练集均方差 |
|------------|-----------------------|-----------------------|-----------------------|
| 梯度训练算法 | 2.74×10^{-1} | 5.38×10^{-1} | 2.64×10^{-1} |
| 遗传训练算法 | 5.99×10^{-4} | 9.84×10^{-4} | 3.85×10^{-4} |
| 微粒群算法 | 3.35×10^{-5} | 8.12×10^{-5} | 4.77×10^{-5} |
| 复合适应度微粒群算法 | 1.23×10^{-5} | 2.01×10^{-5} | 7.80×10^{-6} |

法所训练的神经网络的优劣

3 结 论

研究结果表明,复合适应度微粒群算法所训练的网络,无论是对训练集还是对测试集,所得到的均方误差都最小,说明复合适应度微粒群算法训练网络的效果最好,网络在测试集上的均方误差与在训练集上的均方误差最接近,在一定程度上解决了所训练网络存在的过度拟合问题.在训练求解分类和函数逼近问题的神经网络时,采用复合适应度微粒群算法所训练的网络所达到的精度和泛化能力,要优于梯度算法、遗传算法和常规微粒群算法.

参考文献(References)

- [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, 24(8): 55-57.
(Zhang D X. A New Approach for the Training of Feedforward Neural Network Based on the Information Merging of Output Errors and Derivative Errors[J]. *Computer Engineering and Applications*, 2002, 24(8): 55-57.)
- [2] Coelho C A, Pulido G T, Lechuga M S. Handling Multiple Objectives with Particle Swarm Optimization[J]. *IEEE Trans on Evolutionary Computation*, 2004, 8(3): 256-279.
- [3] Naka S, Genji T, Yura T, et al. A Hybrid Particle

Swarm Optimization for Distribution State Estimation[J]. *IEEE Trans on Power Systems*, 2003, 18(1): 60-68

- [4] Kassabalidis I N, Sharkawi M A, Marks R J, et al. Dynamic Security Border Identification Using Enhanced Particle Swarm Optimization[J]. *IEEE Trans on Power Systems*, 2002, 17(3): 723-729
- [5] Parsopoulos K E, Vrahatis M N. On the Computation of All Global Minimizers through Particle Swarm Optimization[J]. *IEEE Trans on Evolutionary Computation*, 2004, 8(3): 211-224
- [6] Wachowiak M P, Smolikova R, Zheng Y, et al. An Approach to Multimodal Biomedical Image Registration Utilizing Particle Swarm Optimization[J]. *IEEE Trans on Evolutionary Computation*, 2004, 8(3): 289-301
- [7] Boeringer D W, Werner D H. Particle Swarm Optimization Versus Genetic Algorithms for Phased Array Synthesis[J]. *IEEE Trans on Antennas and Propagation*, 2004, 52(3): 771-779
- [8] Vanden Bergh F, Engelbrecht A P. A Cooperative Approach to Particle Swarm Optimization[J]. *IEEE Trans on Evolutionary Computation*, 2004, 8(3): 225-239
- [9] Robinson J, Rahmat Samii Y. Particle Swarm Optimization in Electromagnetics[J]. *IEEE Trans on Antennas and Propagation*, 2004, 52(2): 397-407.

(上接第 957 页)

参考文献(References)

- [1] Or I, Demirkol E. Optimization Issues in Automated Production of Printed Circuit Boards: Operations Sequencing and Feeder Configuration Problems[A]. *ETFA '95 Proc[C]*. Paris, 1995: 479-487.
- [2] Lee S H, Lee B H, Park T H. A Hierarchical Method to Improve the Productivity of a Multi-head Surface Mounting Machine[A]. *Proc of the 1999 IEEE Int Conf on Robotics and Automation[C]*. Detroit, 1999: 2110-2115
- [3] Hong J M, Lee S H, Kim D W, et al. An Algorithmic Approach to the Improvement of Efficiency for Surface Mounting Machines[A]. *IEEE Int Conf on Robotics and Automation[C]*. Nagoya, 1995: 806-811
- [4] Jinhyeon S, Sungsoo P. Efficient Operation of a Surface Mounting Machine with a Multihead Turret[J]. *Int J Production Research*, 1996, 34(4): 1131-1143
- [5] 曾又姣, 严隽琪. 贴片机贴装印刷电路板中供料器的分配问题[J]. *上海交通大学学报*, 2003, 37(11): 1719-

1722

(Zeng Y J, Yan J Q. Feeder Assignment Problem of Placement Machine Assembling Printed Circuit Board[J]. *J of Shanghai Jiaotong University*, 2003, 37(11): 1719-1722)

- [6] Aristides D, Peter C N, Thomas M T. Optimization of High-mix Printed Circuit Card Assembly Using Genetic Algorithms[J]. *Annals of Operations Research*, 1997, 75(1): 303-324
- [7] Peter C, Thomas M T, Peter C N. Optimization of a High-speed Placement Machine Using Tabu Search Algorithms[J]. *Annals of Operations Research*, 2000, 96(1): 125-147
- [8] 邢文训, 谢金星. *现代优化计算方法*[M]. 北京: 清华大学出版社, 1999.
(Xing W X, Xie J X. *Advanced Optimization Algorithms*[M]. Beijing: Tsinghua University Press, 1999.)