

# 一种基于进化算法的BP神经网络优化方法及应用

周国亮<sup>1</sup>, 刘希玉<sup>2</sup>

(1. 山东师范大学 信息科学与工程学院, 山东 济南 250014;

2. 山东师范大学 管理与经济学院, 山东 济南 250014)

**摘要:**针对BP网络的不足,分析了一般进化算法在神经网络结构优化过程中存在的问题,根据物种内优生优育原则和物种间相互竞争、相互学习的生物学原理,提出了一种新的基于进化算法的神经网络优化方法。该方法不但有效弥补了BP神经网络在网络结构、权值选择上的随机性缺陷,缩小了神经网络结构的解搜索空间,加快了BP网络的收敛速度,进而提高了搜索效率,而且还起到对网络的结构和权值进行同时进化的作用。实验结果表明该方法取得了良好的效果。

**关键词:**神经网络;进化算法;BP网络;适应度

中图分类号:TP183

文献标识码:A

文章编号:1673-629X(2007)08-0058-03

## A Method of BP NN Optimization Based on Evolution Algorithm and Its Application

ZHOU Guo-liang<sup>1</sup>, LIU Xi-yu<sup>2</sup>

(1. School of Information Science and Project, Shandong Normal University, Jinan 250014, China;

2. School of Management and Economy, Shandong Normal University, Jinan 250014, China)

**Abstract:** Aiming at the insufficiency of BP neural network, analyses the problems that exist in the optimization process of neural network using evolution algorithm and proposes a new NN optimization method based on evolution algorithm according to competition and adaptation principle among species. The method not only makes up BP NN's random bug when it is used in network instruction and weight choice, contracts answer's search space of NN, quickens BP network's speed in order to improve search efficiency, but also takes effect on network's structures and weights at the same time. Experimental results demonstrate that the method is effective.

**Key words:** neural network; evolution algorithm; BP network; fitness

## 0 引言

人工神经网络特有的非线性适应性信息处理能力克服了传统人工智能方法对于直觉的缺陷,使之在故障检测、智能控制、组合优化等领域得到成功应用。神经网络具有很强的自适应、自学习的能力。但在实际应用中,神经网络也暴露了一些不足:权值的初始化是随机的,易陷入局部困境,收敛时间过长,鲁棒性差等。

进化算法(Evolution Algorithm, GA)是近几年发展起来的一种崭新的全局优化算法,它借用了生物遗传学的观点,通过自然选择、遗传、变异等作用机制,实

现各个个体的适应性的提高,它的基本流程如图1所示。它具有良好的全局搜索性能,并且搜索不依赖于梯度信息,所以可以解决神经网络的优化问题。

进化算法在神经网络中的应用主要是对网络连接的权值和网络结构进行优化,而对于前者的优化已经取得了比较有效的成果<sup>[1]</sup>。但是如果只是对参数进行优化,权值只能在一个预先给定的结构下训练,而此结构有可能存在冗余,就会增加网络训练的负担,降低网络的泛化能力,因此,有必要对网络的结构进行优化。现阶段主要有两种对网络的结构训练的方法:一种是对连接权值矩阵直接编码,缺点是解搜索空间过大;另一种是只对结构相关信息进化的间接编码方法,但进化规则复杂<sup>[2]</sup>。对于两者结合中存在的问题,笔者提出了一种新的二次寻优的神经网络优化方法,以避免寻优规则复杂的问题,还可以对网络的结构和权值进行同时进化,在详细论述了该方法的执行步骤后,通过具体实验证明了其有效性和优越性。

收稿日期:2006-11-11

基金项目:山东省自然科学基金重大项目(Z2004G02);山东省中青年科学家奖励基金项目(03BS00)

作者简介:周国亮(1978-),男,山东菏泽人,硕士研究生,研究方向为进化算法、人工神经网络;刘希玉,博士,教授,博士生导师,主要研究方向为进化算法、人工神经网络、数据挖掘。

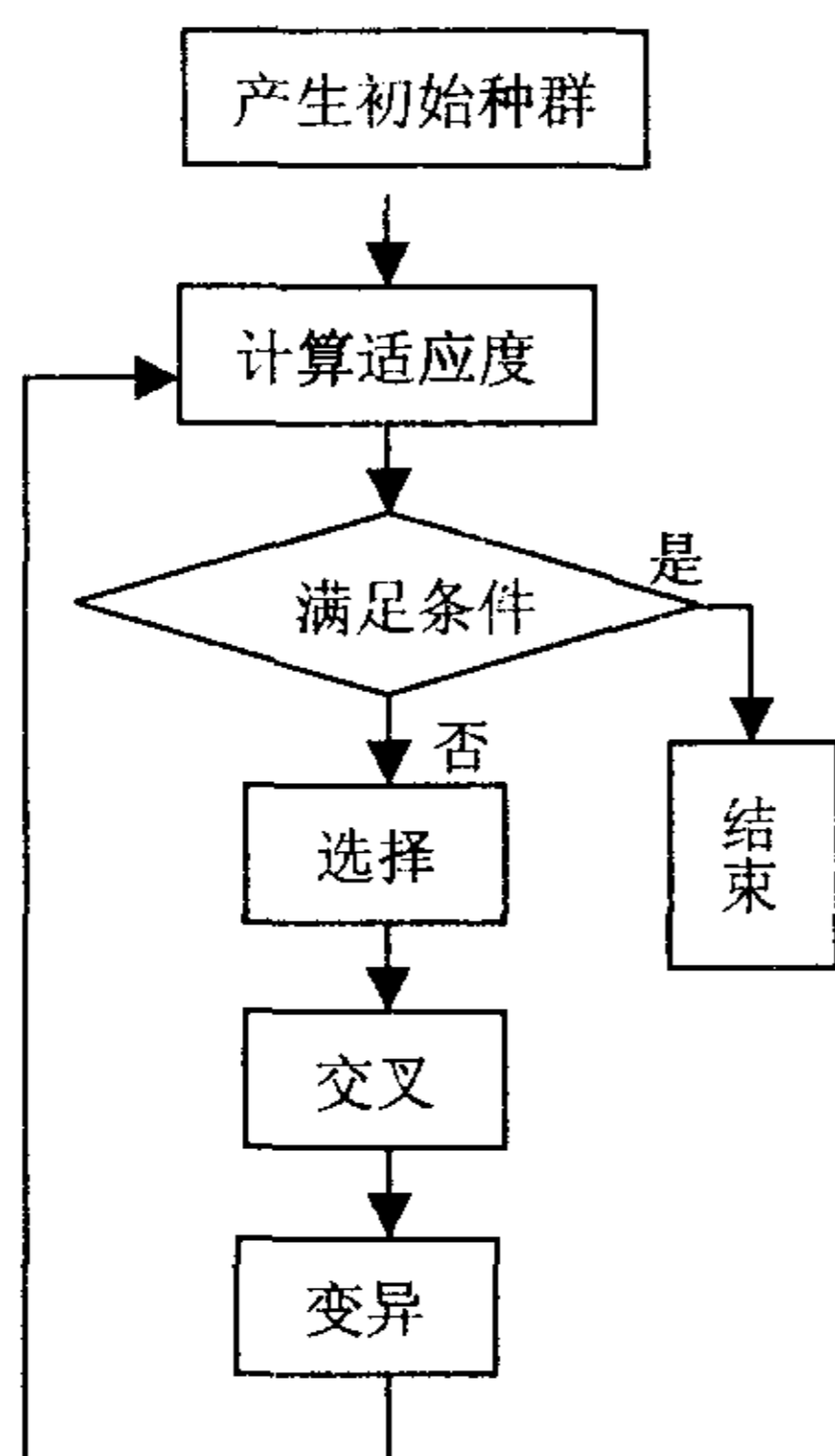


图1 GA流程

## 1 基于进化算法和多物种协同进化的二次优化方法概述

### 1.1 二次优化方法的生物学原理

根据生物进化论学说,自然界中的生物个体遵循“优胜劣汰,适者生存”的原则。但是这种适者生存的法則不仅存在于同一种群内的个体之间,在种群这一层次上仍然适用,即群落内的各种种群之间也同样存在着相互竞争和相互学习的生物界生存原则。不适应环境的种群会逐渐变得稀少,以致灭绝。物种之间的竞争和协同进化可以使它们相互进行自然选择和相互利用,从而促进自身的进化,增强自身的竞争力,逐渐淘汰适应性差的种群<sup>[3]</sup>。

### 1.2 二次优化方法的实现方案

借鉴上面的相同物种内和群落的各物种间的相互竞争和相互学习、协同进化的原理,提出了这样一种优化方法:

(1)首先利用进化算法分别对具有相同个体数目的不同结构的BP网络进行进化,通过评价适应度,找到每个种群的最佳个体并保存下来,这样就完成了同一个种群内的寻优,这里称其为第一次寻优。

(2)其次将选择出来的最佳个体按照不同的比例复制自身,形成初始种群,以便进行进化。这样就将不同物种的个体放在了同一个种群中进行进化,不同物种的个体之间可以进行基因的交叉操作,它们可以吸收彼此的优良基因,最后根据它们各自对于所求解问题适应值,选出合适的结构和权值,这里称其为第二次寻优。

## 2 具体的优化方法

由于二次优化方法的两次优化的不同之处主要在

于初始种群的设定不同,所以在这里进行统一化处理。利用进化算法解决神经网络优化问题的关键之处在于编码方法的选取、初始种群的设定、目标函数(适应度函数)的构造、遗传因子的恰当选择等方面。

### 2.1 编码方法的选取

对初始种群一般采用二进制编码,它的编码方式简单易行,容易进行交叉和变异操作,但是在处理连续性问题上存在一定误差。

浮点数编码可以提高解的精度和搜索的效率,因此文中采用实数编码。在第一次优化时:网络结构第 $K$ 隐节点及其连结权值的编码串为: $w_{1k}w_{2k}\cdots w_{ik}\cdots w_{mk}v_{k2}\cdots v_{kj}\cdots v_{kn}$ ,把所有隐层节点的编码串按顺序连接起来即为整个个体的编码;第二次优化时:由于是几种不同的结构在同一种群中进行优化,所以将第一次优化时产生的最佳个体先进行解码然后重新编码,并以隐节点数最多的结构为编码标准,其他结构的网络在编码时只需改变表示节点的位上的数值即可。用 $\theta_i \in \{0,1\}$ 表示第 $i$ 节点的有无。当 $\theta_i = 1$ 时表示该节点存在,当 $\theta_i = 0$ 时表示该节点不存在。网络结构第 $K$ 隐节点及其连结权值的编码串为: $\theta_k w_{1k}w_{2k}\cdots w_{ik}\cdots w_{mk}v_{k2}\cdots v_{kj}\cdots v_{kn}$ ,把所有隐层节点的编码串按顺序连接起来即为整个个体的编码。

### 2.2 初始种群的设定

进化算法的应用主要是利用其并行性和全局搜索的特点来进行优化,整个算法在解的群体上进行,因此初始种群的设定对GA的运行性能起到重要的作用。针对三层BP神经网络,输入节点和输出节点是由实际问题决定的,所以确定网络的结构实际上就是确定网络结构的隐层节点数目。这里暂时选择隐层节点数目为3,4,5,6,7的网络结构进行训练,在第一次优化时,对每种结构随机选 $N$ 个个体分别进行进化;在第二次优化时,为避免某些优秀个体在充分进化前就被淘汰,按照一定的比例复制第一次优化选出的各结构最佳个体本身,以形成第二次优化时的初始化种群。

### 2.3 适应度函数的构造

BP网络的评估标准就是输出值和期望值之间的误差平方和越小越好。进化算法利用适应度来评估染色体的生存能力。将均方误差函数作为适应度函数可以充分地把两者的评估标准融为一体,提高网络的优化能力,文中将 $F(x) = c_{\max} - f(x)$ 作为GA的适应度函数,其中: $f(x)$ 是BP网络均方误差函数, $c_{\max}$ 是一个较大的常数值。

### 2.4 遗传算子的选择

遗传算子对群体的进化进程起到调节和控制的作用,主要包括:选择、交叉和变异。遗传算子的恰当选

择对进化的趋势起到重要的导向作用。

(1)选择算子来实现对群体中的个体进行优胜劣汰操作,其任务就是按某种方法从父代群体中选取一些个体,以便进行交叉和变异操作。文中选择算子采用轮盘赌选择方法:多次选择,而概率低的就可能被淘汰,如图 2 所示。

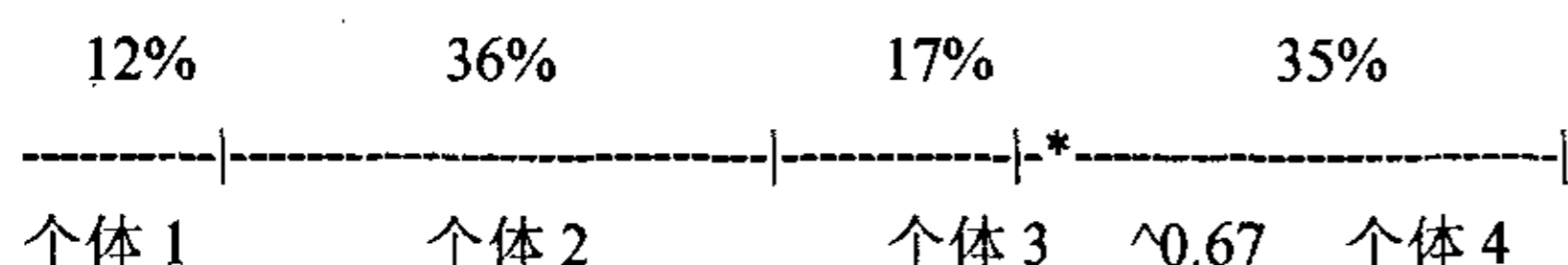


图 2 轮转法示意图

随机数为 0.68 落在了个体 4 的端内,本次选择了个体 4。

(2)交叉算子的选择。交叉的目的是为了将优秀的基因遗传给下一代个体,文中采用单点交叉,将节点以及连接权值的位串当成一个点,与其他的点进行交换,其中交叉个体和交叉位置都是随机选取的。假设两输入两输出的 BP 网络的两个个体进行交叉:

... $w_{12}w_{22}v_{21}v_{22}$  和  $w_{15}w_{25}v_{51}v_{52}$ ... 交换(第一次进化中)

或者

... $\theta_2w_{12}w_{22}v_{21}v_{22}$  和  $\theta_5w_{15}w_{25}v_{51}v_{52}$ ... 交换(第二次进化中)

### 2.5 二次优化方法实现的具体步骤

二次优化方法实现的具体步骤如下:

(1)将 BP 网络按照上面的编码要求将每一个网络的结构按隐层节点的顺序级联起来,随机产生  $N$  个这样的染色体;

(2)将  $F(x) = c_{max} - f(x)$  作为适应度函数,其中  $f(x)$  为均方函数,计算每个染色体的适应度,并判断是否符合优化标准或循环次数,如果符合,则转向(4);

(3)根据适应度选择个体,按一定的交叉率和变异率产生新的个体,返回(2);

(4)保存最佳个体,如果循环次数小于网络个数(隐含层节点不同对应的网络不同),则改变隐层节点的数目返回(1),否则将选出的各结构最佳个体按比例复制自身,形成第二次优化的初始种群;

(5)根据每个个体的编码得到相应的神经网络结构及其权值;

(6)进行 BP 网络的前向传播,根据适应度函数计算每一个个体的适应度,评价相对应的神经网络性能;

(7)评价神经网络的性能是否收敛到期望的精度范围内,如果能就输出优化好的神经网络,如果没有就执行第(5)步;

(8)对群体进行遗传操作,产生新一代群体,跳转到第(5)步继续循环。

### 3 试验结果与分析

结合动态称重系统来进行试验,该系统由埋在路面下的传感器在车辆通过时自动获取车辆的轴重、车速、加速度、车型、总轴距等信息,把这些信息输入到用神经网络设计的称重系统中,计算出该车的重量。文中从实际数据中选取 300 对样本,其中 200 对样本作为训练样本,100 对样本作为测试样本,每对样本包括 5 个数据:速度、加速度、总轴距、轴重、期望轴重。前 4 个数据作为输入向量中的数据,最后的期望重量作为期望输出值,并且将这些数据归一化为  $[0 \sim 1]$  之间的数。结果如表 1 所示。

表 1 部分归一化数据

样本	速度	加速度	总轴重	轴重	期望重量
1	0.822	0.4	0.1258	0.1271	0.1824
2	0.543	-0.1	0.1257	0.5721	0.8602
3	0.521	0.1	0.0666	0.0288	0.0263
4	0.651	0.4	0.1434	0.1942	0.2212
5	0.366	-0.5	0.1856	0.2641	0.2925

当采用二次优化算法训练网络时,网络结构的个数为 5,初始种群为 30,其它初始设置按照上面算法的要求去做,得到的数据如表 2 所示。

表 2 进化后神经网络的实验数据

个体	最大适应度	训练样本的最小均方差	测试样本的最小均方差	
结构	4-3-1	0.018 264	0.000 401 48	0.000 387 23
	4-4-1	0.018 393	0.000 338 67	0.000 346 36
	4-5-1	0.019 341	0.000 325 05	0.000 321 59
	4-6-1	0.019 538	0.000 232 74	0.000 225 08
	4-7-1	0.019 637	0.000 161 83	0.000 132 58

从训练结果可以看出隐含层节点数为 7 的网络最佳个体对训练样本和测试样本的最小均方差都比其余结构小,可见其具有较好的逼近能力和适应能力。

### 4 结束语

在分析了一般进化算法对神经网络优化过程中存在的问题之后,根据物种内优生优育的原则和多物种之间的相互竞争、相互适应原理<sup>[4]</sup>,提出了一种改进的基于进化算法的二次优化方法。该方法有效地弥补了 BP 网络在网络结构、权值选择上的随机性缺陷<sup>[5]</sup>,缩小了神经网络结构的解搜索空间,提高了进化算法在神经网络优化应用中的收敛速度。试验证明该方法可行并且是有效的。

#### 参考文献:

[1] 李敏强. 进化算法与神经网络结合[J]. 系统工程理论与实践, 1999, 2 (2): 66 - 69.

(下转第 63 页)

结构启发式信息的 G 算法这里分别称为 G-CMI 算法、G-MI 算法。

G 算法避免了 K2 算法由于缺乏先验知识无法确定,或只能盲目确定结点次序的不足。在 TAN 结构的基础上学习得到的贝叶斯网络结构更合理。

G 算法的时间开销包括两部分:次优结构的确立和 K2 算法的执行时间。由前文可知,K2 算法的时间复杂度是  $O(n^2mr)$ ,学习 TAN 结构的时间复杂度是  $O(n^2m)$ , $m$  是实例个数, $n$  是属性个数, $r$  是属性最大状态数<sup>[2]</sup>。由此可知 G 算法的时间复杂度是  $O(n^2mr) + O(n^2m)$ ,属于多项式时间复杂度算法。G 算法的时间开销仅比 K2 算法多出建构 TAN 结构的时间。在作者开发的贝叶斯分类器实验平台 MBNC (Bayesian Networks Classifier using Matlab)<sup>[3]</sup> 实验中,用从 UCI (University of California in Irvine)<sup>[5]</sup> 下载的标准数据集建构 TAN 结构,时间一般都不超过 5 秒。

### 3 实验设计与结果

基于 TAN 结构的启发式贝叶斯结构学习算法的测试在 MBNC 实验平台上完成。选择了从 UCI 下载的 Corral 数据集进行实验验证。先进行数据预处理,如识别各种格式的数据集,转换为统一的格式,以及打乱数据集记录的次序,缺失数据的例子被删除或补齐,连续属性的值进行离散化处理,明显对分类的作用微小的属性被忽略等等。

Corral 数据集是一个 7 个属性 (A, B, C, D, E, F, Class) 的人工数据集,其中,Class 是类结点,A 跟 B 相关,C 跟 D 相关,E 与 Class 不相关,F 与 Class 相关,干扰较大。表 2 列出了不同算法学习 Corral 数据集生成结构 BD 测度的分值。

表 2 G 算法和原算法生成结构的分值比较

分类算法	TANC-MI	TANC-CMI	G-MI	G-CMI
结构分值	-520.603728	-487.637297	-473.868599	-473.293974

用 G 算法得到结构的分值较大,表明所生成的结构跟数据拟合的更好。用 G-MI 算法和 G-CMI 算法都能较准确地学习得到“关键”特征,所添加的弧较少,网络结构比较合理,避免了大的存储空间和计算条件概率的复杂度。

G 算法取得了较好的实验效果,有如下特点:

(1) 结合树形结构和网形结构,按“需”添加扩展弧,所添加的弧数目比较少,比 TAN 结构更加简洁,更加合理。

(2) G 算法不需要事先确定结点次序,无须先验知识。

与一般的分类相比,文本分类面临着文本数据的属性个数很多和属性之间存在某些依赖关系的情况。用 TAN 建构文本分类模型,添加的扩展弧较多,结构模型比较复杂,计算复杂度也较高。与此相比,文中提出的 G 启发式算法建构分类模型有很好的适应性。

总之,采用 G 算法所生成的结构较之 TAN 结构更加合理,在某些领域(如文本分类)内有更大的适应性。

### 4 结论与展望

提出了启发式 BNC 结构学习 G 算法,通过在 MBNC 实验平台上验证表明算法的有效性,进一步的研究如下:

(1) 由 TAN 结构确定结点次序,还不是最优的结果,采用更好的启发式信息,减少节点次序的搜索空间是今后的研究方向。

(2) 网络结构中最大父结点数限制为 2,如最大父结点数取为 3 或更高,虽然计算复杂度会大幅提高,但能学习得到更加合理的网络结构。

#### 参考文献:

- [1] Duda R, Hart P. Pattern Classification and Scene Analysis [M]. New York: John Wiley and Sons, 1973.
- [2] Friedman N. Bayesian network classifiers [J]. Machine Learning, 1997(29): 131-163.
- [3] Cooper G, Herskovits E. A Bayesian method for the induction of probabilistic networks from data [J]. Machine Learning, 1992, 9: 309-347.
- [4] 程泽凯, 林士敏, 陆玉昌. 基于 Matlab 的贝叶斯分类器平台 MBNC [J]. 复旦学报, 2004, 43(5): 729-732.
- [5] Blake C, Keogh E, Merz C. UCI repository of machine learning database [EB/OL]. 1998. <http://www.ics.uci.edu/mllearn/MLRepository.html>.

(上接第 60 页)

- [2] 罗 菲, 何明一. 基于免疫进化算法的多层前向神经网络设计 [J]. 计算机应用, 2005, 25(7): 1661-1663.
- [3] 吴清佳. 遗传神经网络的智能天气预报系统 [J]. 计算机工程, 2005, 31(14): 176-189.

- [4] 林 雄. 自适应模糊神经网络研究 [J]. 微计算机信息, 2006, 19(3): 16-17.
- [5] 史永胜, 宋云雪. 基于进化算法和 BP 神经网络的故障诊断模型 [J]. 计算机工程, 2004, 30(14): 125-127.