

文章编号: 1000-6788(2001)10-0042-05

一种基于遗传算法的新型神经网络设计

罗 键, 刘军祥

(厦门大学自动化系, 福建 厦门 361005)

摘要: 神经网络的设计主要集中在网络权值的选取和神经网络结构的确定两个方面, 与遗传算法结合是目前研究发展的趋势。本文与一般的基于遗传算法的神经网络设计相比, 提出一个新型算子——BP算子, 并对神经网络的权值和结构同时优化。仿真结果表明该算法结果比较理想。

关键词: 遗传算法; 神经网络; BP算子

中图分类号: TP183

文献标识码: A

A New Method of Designing Artificial Neural Network Based on Genetic Algorithm

LUO Jian, LIU Jun-xiang

(Automation Department, Xiamen University, Xiamen 361005, China)

Abstract Currently research on ANN (Artificial Neural Network) is mainly focused on the initial weight and structure. It is a developing tendency to combine ANN with GA (Genetic Algorithm). Compared with general ANN design, this paper puts forward a new operator, BP operator, and optimizes the ANN's initial weights and structure at the same time. The result of the emulation is very reasonable.

Keywords: ANN; GA; BP operator

1 引言

近年来,神经网络的自动化设计成为当前的一个研究热点。按照处理方式的不同可将其分为连接主义设计法和演化设计法。连接主义设计法主要包括增补算法(Constructive Algorithm)和削减算法(Destructive Algorithm)。这两种方法采用爬山策略来设计神经网络结构,一旦某个网络结构被确定为不适,则网络将被更新。爬山策略很容易陷入局部最优,且搜索的解的范围仅仅是网络结构空间中的很小的一部分,所以很难保证设计出的网络结构是最优的。由于设计或训练神经网络需要在很大的空间中进行搜索,而且搜索空间中具有很多局部最优点,这使得传统的算法求解该问题非常困难。而以遗传算法(Genetic Algorithm,简称GA)为代表的演化算法适合大规模并行设计且能以较大的概率找到全局最优解,这就使得GA与人工神经网络(Artificial Neural Network,简称ANN)的结合成为一种趋势。Yao X在文献[1]中对此做了很好的综述。

现阶段有关应用GA来设计ANN的研究,主要集中在神经网络的结构和权值设计两个方面。与之相对应,则暴露出了一些不足。一方面,设计独立,即将ANN的结构设计和权值设计孤立开来,单独进行,而且主要集中在GA理论的改进之上,如实数编码、混合优化等。另一方面,算法独立,对于一个给定的问题,有可能在不了解如何用ANN来解决它,那么有可能首先用GA来找出解决问题的最优或次优结构,然后再用一定的算法如BP算法来找到最优或次优权值。本文的主要内容即在部分有关基于GA的ANN设计研究基础之上,提出一种新的算法,即GA-ANN优化设计算法来解决这个问题。

2 带 BP 算子的 GA-ANN 算法设计

2.1 假设条件

对于网络的结构与权值的同时进化算法, 首先给出如下假设

假设 1 网络是严格分层全连接的, 即当且仅当两相邻层的结点才有可能进行连接, 且必须有连接, 每一神经元都设成是线性阈值单元且使用 sigmoid 传递函数

假设 2 输入输出空间向量在实数空间上取值, 相连接的两个神经元如果没有影响, 则权值为 0

假设 3 给定一个训练集, 其输入模式为 $\{X^1, X^2, \dots, X^p\}$, 对应的目标输出向量为 $\{T^1, T_2, \dots, T^p\}$. 对每个输入模式, 根据网络的结构、连接权值和神经元的类型及阈值可以确定网络的实际输出. 如果输出向量是 m -维的, 则根据实际输出向量 $\{Y^1, Y^2, \dots, Y^p\}$ 和目标输出向量之间的误差可以定义一个最小二乘误差函数:

$$E(\text{net}) = \sum_{k=1}^p \sum_{i=1}^m (t_i^k - y_i^k)^2$$

t_i^k, y_i^k 分别表示向量 T^k 和 Y^k 的第 k 个分量.

上面定义的最小二乘误差函数可用于描述对给定训练数据集当前网络的性能. 因此, 设计 ANN 及训练其权值便可看成是一个优化过程, 优化的目的便是设计出网络使得该目标函数达到最小值. 如果期望在具有相同性能的条件下网络的结构尽可能地简单, 即使网络尽量具有最少的结点数和最少的连接, 则可以在函数 $E(\text{net})$ 的右端加上一个控制项 Cn_p , 这里 C 是一个控制参数, 称为网络复杂性系数, n_p 是网络参数的个数, 通常包括网络的结点数与连接数.

2.2 编码表示

在编码时, 对网络结构采取二进制编码, 对每个权 (包括阈值) 采取实数编码来表示, 这样编码的优点在于编码长度比全部采用二进制编码时要小得多. 例如, 对于某个问题, 如果有两个隐层, 分别有 Size1 个和 Size2 个隐含单元, 输入层有 InNum 个单元, 输出层有 OutNum 个单元, 且每层的单元数用 n 位二进制码串表示, 则每层的单元数可达 $2^n - 1$ 个, 采用实数与二进制混合编码时串长为

$n \cdot (\text{InNum} + \text{OutNum} + \text{Size1} + \text{Size2}) + (\text{InNum} + 1) \cdot \text{Size1} + (\text{Size1} + 1) \cdot \text{Size2} + (\text{Size2} + 1) \cdot \text{OutNum}$
精度可达 0.0001 或更小 (根据计算机实数或浮点数的精度而定), 如果全部采用 n 位二进制编码, 则串长为

$$n \cdot [\text{InNum} + \text{OutNum} + \text{Size1} + \text{Size2} + (\text{InNum} + 1) \cdot \text{Size1} + (\text{Size1} + 1) \cdot \text{Size2} + (\text{Size2} + 1) \cdot \text{OutNum}]$$

若初始权值范围为 $[-0.5, 0.5]$, 则权值精度仅达 $\frac{1}{2^n}$. 与此同时, 实数编码不用象二进制编码表示时需要预先定义解的精度, 另外在加入一些启发式知识时不需要在译码与编码之间来回切换. 为了能同时进化网络的拓扑结构和连接权值, 允许在同一群体内的网络具有不同的拓扑结构, 即它们可以有不同的结点数和连接权数.

2.3 选择策略

采用赌轮选择策略. 使用基于适应值比例的选择策略时, 首先要确定适应值的度量方法, 设 $E(\text{net}_i, t)$ 是第 t 代进化群体中第 i 个个体所具有的目标函数 $E(\cdot)$ 的值, 其适应值如果定义为

$$f(\text{net}_i, t) = \frac{1}{1 + E(\text{net}_i, t)}$$

于是, 由赌轮选择策略可知, 网络 net_i 被选择到下一代作为遗传操作的父代的概率为

$$p(\text{net}_i, t) = \frac{f(\text{net}_i, t)}{\sum_{j=1}^n f(\text{net}_j, t)}$$

这里 n 为群体规模

2.4 遗传操作

2.4.1 个体产生

随机生成初代个体, 群体规模设置为某一整数

2.4.2 杂交算子

由于编码表示使得在同一种群内的网络可能不同构, 这样给杂交算子的设计带来了一定的难度. 在此实现两点杂交, 即在两个父代中独立地选择相同数目的连接权, 交换相应的连接权值以生成杂交后的两个后代, 如图 1 所示

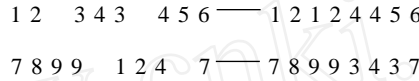


图 1 两点杂交算子

2.4.3 变异算子

对网络结构, 由于是二进制表示, 实现简单的变异方式即可, 即 0 变 1, 1 变 0. 网络结构的改变, 将导致网络权值数量增加或减少, 采取随机选择插入点或删除点的方式, 当需要增加点时, 随机生成其相应的连接权值, 这样直到权值的数量与网络结构相符为止.

2.4.4 BP 算子

现阶段, 有关结合遗传算法和神经网络的算法都是一个固定的模式, 即先用遗传算法一般的操作算子找到一个优化后的神经网络结构或者是神经网络权值, 然后再用神经网络 BP 算法进一步寻优. 这种算法把遗传算法和 BP 算法相对孤立开来 (如图 2 所示), 找到全局最优解, 其实质是一种简单迭加的混合算法.

现在将神经网络的 BP 算法应用到遗传算法中, 如图 3 中虚线所示. 对比图 2, 可以看出, 图 3 新增加了一种遗传操作, 这种操作对遗传算法中每代适应值满足一定条件的个体进行 BP 训练, 得到新的个体, 称之为 BP 算子.

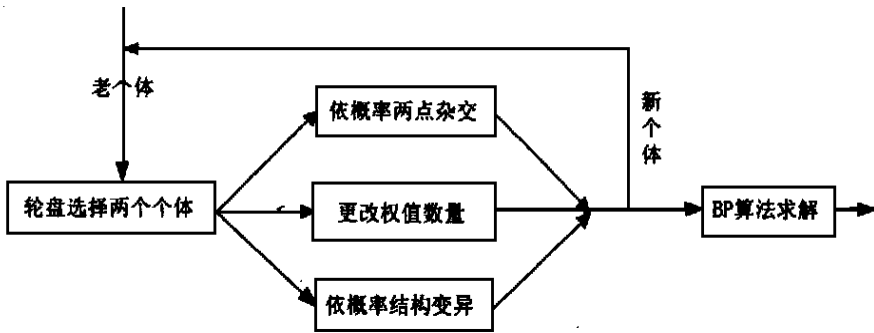


图 2 传统的 GA 操作

如图 3 所示, 对每代个体, 除了执行传统的 GA 操作外, 若其适应值或某一随机概率值达到一定大小, 就进行 BP 操作. 为简单操作, 仅取那些第一隐层单元数不为 0 的个体为正确变异个体, 而对于那些不符合条件的个体, 如第一隐层单元数为 0, 给这些小个体一个固定的染色体串和适应值. 可以看出, BP 算子实质是对群体中某些个体的变异操作, 染色体串记录经过了 BP 算子的操作结果.

2.5 进化规则

对每相连两代中得到的最优个体进行比较, 若后代的最优个体不及前代的最优个体, 则以前代的最优个体来代替后代的最差个体, 这样达到最优繁殖.

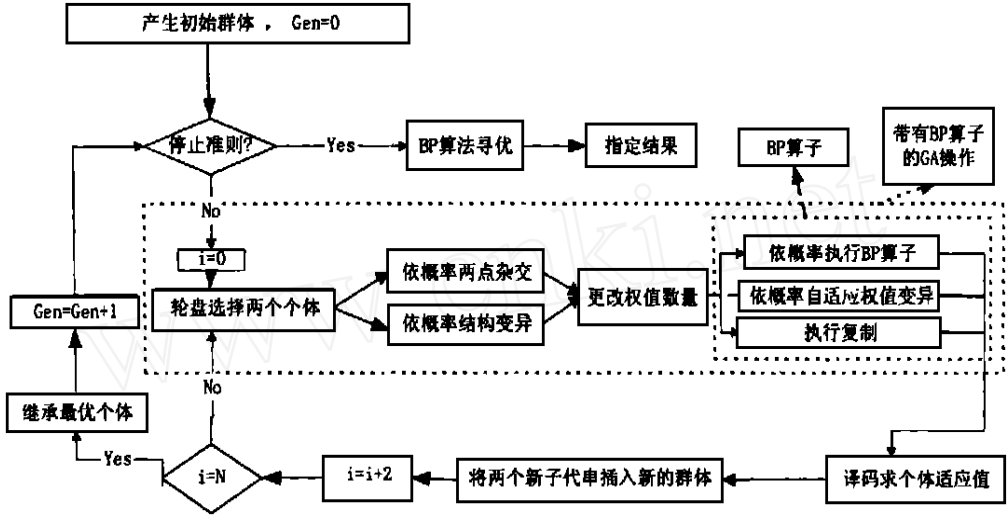


图 3 带有BP算子的 GA-ANN 设计流程图

3 算法仿真

以文献[2]中的模式识别学习任务(见表 1)为例,用此算法进行仿真。给出的模式识别学习任务中,在输入位中,开头两位是噪声位,与输出模式没有关系;最右边两位与输出之间的关系是二进制整数到它们对应的 Gray 码。这个任务的复杂性在于增加了不相关输入的单元和要求再编码有意义的输入。

在学习过程中,为了对每个由遗传算法产生的个体进行评价,首先从 8 个样本模式中随机选取一个放在一边;利用剩下的七个样本(训练集)对染色体个体所对应的网络结构进行训练;当误差平方和达到预定的

目标值或预先确定的学习迭代时,就停止训练;一旦一个网络已经学习到了判别准则,就用先前挑选出来的训练实例对它进行评价。每个个体的适应值就是其对这个实例的均方误差,它表示在学习阶段中对所获得的知识的一般性的估计。针对该问题,设置如下网络参数进行实验:

- 群体规模: 30;
- 最大选带代数: 15;
- 杂交率: 0.95;
- 结构变异率: 0.005;
- 权值变异率: 0.8
- 最优 ANN 个体训练代数: 10000;
- BP 算子的训练代数: 50;
- BP 算子的下临界适应值: 140;
- 适应值函数: 如果译码值大于 70, 则 $f = 5$; 否则

$$f = \exp((70 - x - perf) / 13 - 0.002 * lchrom)$$

其中 $lchrom$ 为网络权数, 0.002 为惩罚因子, 依具体情况不同而有所改变, 两者之积作为一个惩罚项, 用以使得最优个体不仅网络权值接近最优, 而且结构简单。另 x 为译码值, $perf$ 为 BP 训练误差。

自适应并行遗传算法仿真得到的最优个体为:

表 1 模式识别任务

输入	输出	输入	输出
0000	00	0100	00
1100	00	1000	00
1001	01	0001	01
1101	01	0101	01
0010	11	1010	11
0110	11	1110	11
0011	10	0111	10
1011	10	1111	10

```

0      0      1 0000  1 0000  0      0      0 9536  - 0 1557
- 1.5574  0 8943  - 1 6028  1 0334  0 5499  2 0176  - 0 6425
- 1.7765  - 2 4811  - 2 6134  0 1162  0 8586  - 1 3909  - 1 0949
- 0 4145  1 7207  3 8867  2 1118  1 9203  4 0778  4 3323

```

可见, 权值数为 23 (包括阈值), 适应值为 178 6976, 其中前 6 位就是网络结构的二进制编码, 后 23 位是得到的最优个体没经过 BP 进一步训练的网络权值 更进一步的结果如图 4 所示, 最优 ANN 个体训练如图 5 所示

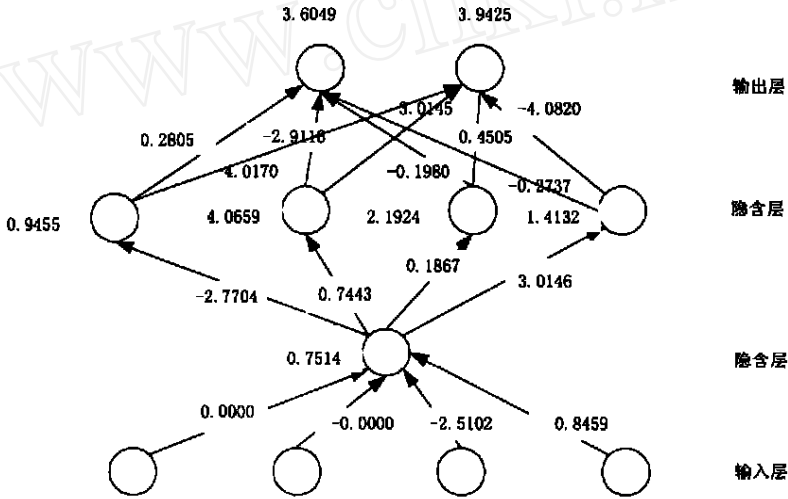


图 4 带BP 算子的 GA-ANN 设计结果图

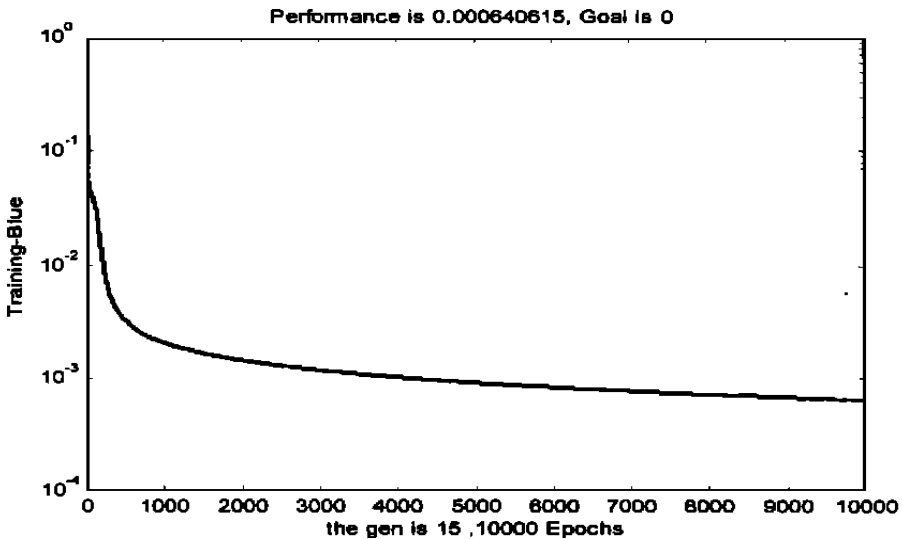


图 5 带BP 算子的 GA-ANN 训练图

4 结论

与文献[2]中的结果进行比较, 从图 5 可以看出, 基于带有 BP 算子的遗传算法的神经网络设计在利用 (下转第 70 页)

续表 3 1998 年上海股票市场涨幅前 20 名股票

代码	股票名称	97 年收盘价	98 年收盘价	调整后涨幅	97 年总股本	98 年总股本	股本扩张速度	利润增幅
600601	延中实业	9.28	12.79	106.73	10368	15552	50	10.32
600733	前锋股份	10.49	20.9	99.24	10977	10977	0	16.48
600711	龙舟股份	7.25	14.28	96.97	6036	6036	0	0
600747	大显股份	9.79	17.1	91.8	12837	16535	29	83.03
600651	飞乐音响	7.5	11.06	91.71	8112	10546	30	15.03
600862	ST 通机	5.35	10.24	91.4	7435	7435	0	811.55
600778	友好集团	6.66	10.29	90.04	10700	17109	60	49.16
600782	新华股份	10	18.2	85.34	10795	10795	0	-15.87

注 以 1997 年股本为基准数据, 总股本的平均值(mean)为 25959 万股, 总股本的中位数(median)为 13680 万股

参考文献:

- [1] Edw in J Elton, Martin J Gruber. Modern portfolio theory and investment analysis[M]. New York: John Wiley, 1991.
- [2] 霍学文, 赵军. 关于股票定价理论的发展脉络[J]. 南开经济研究, 1998, (1): 48- 53
- [3] Fama, Eugene F, French Kenneth R. Size and book to market factors in earnings returns[J]. Journal of Finance, 1995, 50: 131- 155

(上接第 46 页)

BP 算法搜索全局最优解时, 在第 0 代就已逼近全局最优. 利用 BP 算子, GA 算法个体能记忆其运作结果, 使各个个体能迅速地逼近最优结构或最优权值, 仿真结果表明, 算法理想

参考文献:

- [1] Yao X. A review of evolutionary artificial neural networks[J]. International Journal of Intelligent Systems, 1993, 8(4): 539- 567.
- [2] 刘勇, 康立山, 等. 非数值并行算法——遗传算法[M]. 北京: 科学出版社, 1995
- [3] 潘正君, 康立山, 陈毓屏. 演化计算[M]. 北京: 清华大学出版社; 南宁: 广西科学技术出版社, 1998
- [4] Chen D, Giles C, Sun G, Chen H, Less Y, Goudreau M. Constructive learning of recurrent neural networks[J]. IEEE International Conference on Neural Networks'93, 1993, 3: 1196- 1201
- [5] 陈荣, 徐用懋, 等. 多层前向网络的研究——遗传 BP 算法和结构优化策略[J]. 自动化学报, 1997, 23(1): 43- 49
- [6] 王保中, 康立山, 何巍. 基于实数编码遗传算法的多层神经网络 BP 算法[J]. 武汉大学学报(自然科学版), 1998, 44(3): 289- 291
- [7] 孙亚军, 等. 一种并行的遗传/神经网络混合学习算法[J]. 电子科技大学学报, 1996, 25(4): 373- 376