

# 遗传算法与神经网络的结合<sup>\*</sup>

李敏强 徐博艺 寇纪淞

(天津大学系统工程研究所, 天津 300072)

**摘要** 阐明了遗传算法和神经网络结合的必要性和可行性, 提出用多层前馈神经网络作为遗传搜索的问题表示方式的思想。用遗传算法和神经网络结合的方法求解了短期地震预报问题, 设计了用遗传算法训练神经网络权重的新方法, 实验结果显示了遗传算法快速学习网络权重的能力, 并且能够摆脱局部极点的困扰。

**关键词** 遗传算法 进化计算 神经网络

## On the Combination of Genetic Algorithms and Neural Networks

Li Minqiang Xu Boyi Kou Jisong

(Institute of Systems Engineering, Tianjin University, Tianjin 300072)

**Abstract** In this paper, we demonstrate the necessity and possibility of combining neural network (NN) with GAs. The notion of using multilayered feed forward NN as the representation method of genetic and the searching technique is introduced. We combine GA and NN for solving short term earthquake forecasting problem, design a novel method of using GAs to train connection weights of NN. The empirical test indicates the capability of the new method in fast learning of NN and escaping local optima.

**Keywords** genetic algorithms; evolutionary computation; neural networks

### 1 引言

智能可以分为三个层次: 高层次的是生物智能(BI), 其次是人工智能(AI), 处于低层次的是计算智能(CI)。计算智能是国际上新近提出的学科概念, 在计算智能中, 计算的概念是传统计算概念的拓展, 计算对象不仅局限于数和字符, 运算符号也不再局限于加减乘除等运算, 在这个范畴内的加减乘除也需赋予新的含义。但一般来说, AI偏重于逻辑推理, 而CI则偏重于数值计算。

目前, 计算智能正处于迅猛发展的阶段, 其主要技术包括模糊技术、神经网络、进化计算等<sup>[5]</sup>。这几项技术各自均有了数十年的历史, 但当时这些方法并未受到足够的重视, 一是当时这些方法还不很成熟, 二是受当时计算机软硬件的限制, 而这些方法一般需要较大的计算量, 难以取得实际应用。随着计算机技术的发展和普及, 它们在最近十年得到了突飞猛进的发展, 引起了诸多领域专家学者的关注, 成为一个跨学科的研究热点。近年来, 这些方法呈互相融合的趋势<sup>[3]</sup>, 它们之间的相互补充可增强彼此的能力, 从而获得更有力的表示和解决实际问题的能力。如对模糊神经网络、模糊遗传算法、模糊分类器系统、用遗传算法优化模糊系统的隶属度函数及神经网络的进化设计方法等的研究都体现了这种融合的优点<sup>[1,2]</sup>。

<sup>\*</sup> 本文于1997年7月16日收到

国家自然科学基金资助项目(79400013, 69574022)

天津市自然科学基金资助项目(F09523)

## 2 遗传算法与神经网络的融合

### 2.1 神经网络为什么需要遗传算法

对于一个实际问题建立神经网络通常包括下面四个阶段: 首先研究者根据自己的理论、经验和研究兴趣选择一个问题域, 如模式识别、神经控制、经济预测等。第二, 根据学习任务设计网络结构, 包括处理单元个数、各层的组织结构及处理单元之间的联结。第三, 根据已知的网络结构和学习任务用梯度下降学习算法, 如BP算法来训练联结权值。最后, 研究者以测量到的目标性能, 如解决特殊问题的能力、学习速度和泛化能力对训练过的网络进行评价。这个过程可以不断重复以获得期望的结果。

描述一个ANN模型结构的主要参数有: 网络层数、每层单元数、单元间的互连方式等。设计ANN的结构, 实际上就是根据某个性能评价准则确定适合于解决某个问题或某类问题的参数的组合。当待解决的问题比较复杂时, 用人工的方法设计ANN是比较困难的。即使小的网络的行为也难以理解, 大规模、多层、非线性网络更是十分神秘, 几乎没有什么严格的设计规则。Kosmogorov定理说明在有合理的结构和恰当的权值的条件下, 三层前馈网络可以逼近任意的连续函数, 但定理中没给出如何确定该合理结构的方法, 研究者只能凭以前的设计经验或遵循这样一句话“问题越困难, 你就需要越多的隐单元”来设计ANN的结构。而标准工程设计方法对于神经网络的设计也是无能为力, 网络处理单元间的复杂的分布交互作用使模块化设计中的分解处理技术变得不可行, 也没有直接的分析设计技术来处理这种复杂性, 更困难的是, 即使发现了一种足以完成某一特定任务的网络, 又怎能确定我们没有丢失一个性能更好的网络呢?

到目前为止, 人们花费了大量的时间和精力来解决这一难题, 而神经网络的应用也正向大规模、复杂的形式发展, 人工设计网络的方法应该抛弃, ANN需要高效的自动的设计方法, GA s则为其提供了一条很好的途径。

遗传算法用于ANN的另一个方面是用遗传算法学习神经网络的权重, 也就是用遗传算法来取代一些传统的学习算法。评价一个学习算法的标准是: 简单性、可塑性和有效性。一般的, 简单的算法并不有效, 可塑的算法又不简单, 而有效的算法则要求算法的专一性、完美性, 从而又与算法的可塑性、简单性相冲突。目前广泛研究的前馈网络中采用的是Rumelhart等人推广的误差反向传播(BP)算法, BP算法具有简单和可塑的优点, 但是BP算法是基于梯度的方法, 这种方法的收敛速度慢, 且常受局部极小点的困扰, 采用GA s则可摆脱这种困境。当然, 使用GA s可以把神经网络的结构优化和权值学习合并起来一起求解, 但这对计算机的处理能力要求很高。

### 2.2 遗传搜索技术为什么需要神经网络

现在我们所面对的系统越来越复杂, 所处理的系统问题也是越来越复杂, 用传统GA s中的串表示方式来表达复杂系统的规律性是很困难的。遗传规划中采用了树型结构的表示方式, 可以方便地表示问题中的层次性, 但对于系统中的反馈信息, 这种表示方式也是无能为力的。对许多领域来说, 网络或图为复杂问题提供了一种更自然、直观、有效的表示方式。

从广义上讲, 系统都是以网络的形式存在的, 但在很多问题的描述和建模过程中, 往往采用某种解析方法, 运筹学中的网络分析仅仅应用于一类明显具有网络形式的工程类问题求解, 而将网络作为系统的一般描述方法, 并建立相应的模型和求解策略, 则是人们正在努力研究的课题<sup>[4]</sup>。诚然, 寻找一种通用的系统的网络表示形式是不明智的, 借助目前研究比较成熟的多层前馈神经网络作为遗传搜索的问题表示方法是可行的。图1表示了遗传算法与神经网络的结合。

## 3 遗传算法在神经网络学习中的应用

### 3.1 神经网络学习的复杂性

这里所指的神经网络的学习包括拓扑结构的学习和权值的学习两个方面。ANN的学习是按照某种组合性能指标, 在可能的网络空间上, 搜索具有最优性能的网络形式。评价函数的选择一般应考虑以下几个方面的因素: 网络的收敛性、泛化能力以及网络的稳定性等。ANN的性能曲面一般是很复杂的, 可能的处

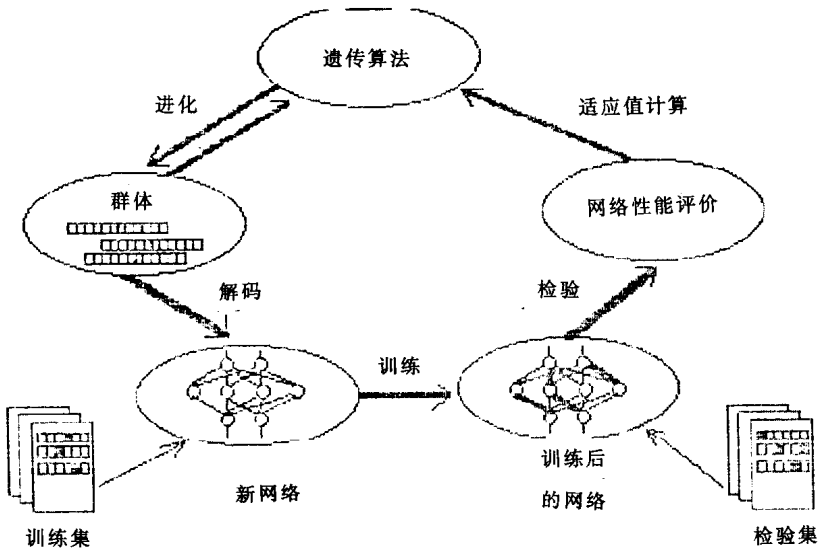


图 1 遗传算法与神经网络的结合

理单元数目及连接是无界的,因而曲面无限大,处理单元数目离散取值,导致网络的性能是非连续的,故性能曲面是不可微的。

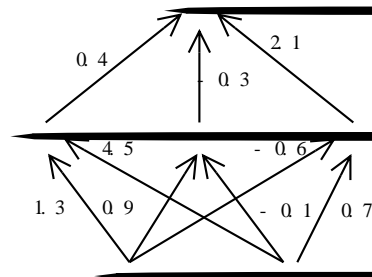
枚举搜索方法会遇到网络空间的组合爆炸问题,随机搜索方法在长的搜索过程中也不比枚举搜索方法好,难于发现有用的网络形式。而基于梯度的搜索方法需要很好定义的可微曲面,难以解决具有局部极小点的复杂的欺骗性的问题。

对于神经网络来说,GA s 所处理的每个个体是单独一个网络,所涉及的参数即网络结点之间的连接权值均为实数,而标准遗传算法所处理的参数是采用二进制编码的整数。虽然可以采用二进制编码再转化为实数,但这样引入了量化误差,使参数变化为步进,如目标函数值在最优点附近变化较快的情况下,则可能错过最优点。因此,网络的权值应采用实数表示。然而,神经网络的结构参数一般是整数型的,也就是说,ANN 的参数有连续型的,也有离散型的,ANN 的学习问题是一非线性混合优化问题,从这一点讲,用一般的优化方法也是不行的,GA s 可以方便地表示逻辑值和实数值,且 GA s 并未规定一定要用数来表示,GA s 也可以应用于不好量化的对象,这使 GA s 在 ANN 的学习及人工智能的其他领域得到了广泛的应用。

### 3.2 面向神经网络权值学习的遗传算法

1) 编码方式:神经网络的权值学习是一个复杂的连续参数优化问题,如果采用二进制编码,会造成编码串过长,且需要再解码为实数,使权值变化为步进,影响网络学习精度。这里我们采用了实数编码,如图 2 所示,神经网络的各个权值按一定的顺序级联为一个长串,串上的每一个位置对应着网络的一个权值。

2) 评价函数  $f$ : 将染色体上表示的各权值分配到给定的网络结构中,网络以训练集样本为输入输出,运行后返回误差平方和的倒数作为染色体的评价函数,如式 1 所示。



编码为: (0.4, -0.3, 2.1, 1.3, 0.9, -0.6, 4.5, -0.1, 0.7)

图2 神经网络权值学习问题的编码方式

$$f = 1 / \sum_i^n e_i^2 \quad (1)$$

3) 初始化过程: 初始染色体集中, 网络的各权值是以概率分布  $e^{-|y|}$  来随机确定的, 这点有别于BP算法。在BP算法中, 初始权值一般取-1.0~1.0之间均匀分布的随机数。遗传算法的这种随机分布的取法是通过以往的大量实验得出的, 可以发现, 当网络收敛后, 权值的绝对值一般都较小, 但也有一些权值是较大的。采用以上初始化方法的原因是使遗传算法能够搜索所有可行解的范围。

4) 遗传算子: 对于不同的应用问题, 遗传算子的形式多种多样, 这里采用了权值交叉和权值变异算子, 如图3所示。

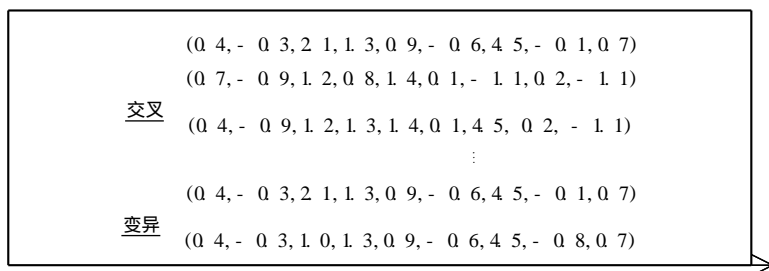


图3 交叉和变异算子的运算

权值交叉算子: 对于子代染色体中的每个权值输入位置, 交叉算子从两个亲代染色体中随机选取若干个交叉位置, 并将这一代染色体在交叉位置进行交叉运算, 这样子代染色体便含有两个亲代的遗传基因。

权值变异算子: 对于子代染色体中的每个权值输入位置, 变异算子以概率  $P_m$  在初始概率分布中随机选择一个值, 然后与该输入位置上的权值相加。

5) 选择方式: 这里不采用比例选择方式, 而使用亲代度量  $S$  决定群体中每个个体被选中的概率, 群体中的个体以式2所示的概率成为亲代染色体。

$$\begin{aligned} P_2 &= P_1 \cdot S \\ P_3 &= P_2 \cdot S \\ &\vdots \end{aligned} \quad (2)$$

其中  $P_1, P_2, P_3, \dots$  分别表示评价函数最优、次优、第三...的个体的概率。

使用实数编码和使用二进制编码之间的主要区别:

- 1) 每个权重都由单一的实数值表示, 因而, 重组只能发生在各权值之间。
- 2) 变异操作是将一个随机值加到现存的权值上, 而不是取代它。
- 3) 通常使用较小的群体。
- 4) 变异操作的使用率较高, 因为交叉不能改变每个权值。
- 5) 不需要二进制和实数之间的转换。

### 3.3 一种改进的算法(MGA)

在改进的算法(MGA)中, 我们采用了两种技术来改善3.2中遗传算法的运行性能。

1) 采用自适应交叉率和变异率

交叉率  $P_c$  和变异率  $P_m$  的大小对遗传算法的运行性能有较大的影响, 较理想的情况下,  $P_c$  与  $P_m$  的取值应在算法运行过程中随着适应值的变化而自适应改变。我们用适应值来衡量算法的收敛状况, 对于适应值高的解, 取较低的  $P_c$  和  $P_m$ , 使该解进入下一代的机会增大; 而对于适应值低的解, 则应取较高的  $P_c$  和  $P_m$ , 使该解被淘汰掉; 当成熟前收敛发生时, 应加大  $P_c$  和  $P_m$ , 加快新个体的产生。式(3)是  $P_c$  和  $P_m$  随着适应值的变化而自适应改变的计算公式。

$$\begin{aligned} P_c &= \begin{cases} k_c(f_{\max} - f_c)/(f_{\max} - \bar{f}), & f_c > \bar{f} \\ k_c, & f_c < \bar{f} \end{cases} \\ P_m &= \begin{cases} k_m(f_{\max} - f_m)/(f_{\max} - \bar{f}), & f_m > \bar{f} \\ k_m, & f_m < \bar{f} \end{cases} \end{aligned} \quad (3)$$

其中  $k_c$ ,  $k_m$  是小于 1 的常数,  $f_c$  是要交叉的两个个体中适应值大的一个,  $f_m$  是要变异的个体的适应值,  $f_{\max}$ ,  $\bar{f}$  分别是群体的最大适应值和平均适应值,  $f_{\max} - \bar{f}$  体现了群体的收敛程度, 若此值小, 说明群体已趋向收敛, 应加大  $P_c$  和  $P_m$ 。

## 2) 采用 GA 与 BP 的混合技术

前述的不采用加速收敛技术的 GA 对于全局搜索有较强的鲁棒性和较高的效率, 但不适合候选解的精确, 难以确定它们的确切位置。组合 GA 与 BP 的目的是避免它们各自存在的问题, 综合它们各自的优点, 即 GA 的全局收敛性和 BP 局部搜索的快速性, 使它们更加有效地应用于神经网络的学习中。GA 与 BP 可以有不同的组合方法。

这里, 我们先使用 GA 反复优化神经网络的权值, 直到平均值不再有意义地增加为止, 此时解码得到的参数组合已经充分接近最佳参数组合, 在此基础上再用 BP 算法对它们进行细调, 这种方法的通用性较好。

## 4 三种算法在短期地震预报中的应用

我们将误差反传算法 (BP) 和 3.2 中所研究的实数编码遗传算法 (GA), 3.3 中的改进型 (MGA) 三种算法应用于神经网络短期地震预报中, 并给出了三种方法的结果比较。

地震预测问题的输入空间是一多维空间, 选取网络的输入变量时如果输入变量过少, 可能引起建模不充分, 过多的输入变量则会增加模型输入输出关系的复杂度, 并延长网络的收敛时间。本方案中, 我们选用前一年的最大震级、最大震级之差、累计能量及累计能量之差为网络的输入, 网络的输出是下一年的震级。我们将某地区过去八十年的地震数据处理后, 从中取出 50 个样本组成建模集<sup>[6]</sup>。

网络选用三层 BP 网络, 隐层节点数为 30, 这个 4-30-1 网络共有 150 个可调权值。各层的激发函数均为 Sigmoid 型函数, 这种函数的输出在 0~1 之间变化, 且只有当输出为 - 和 + , 输出才达到 0 和 1, 比较合适的范围是 0.1~0.9, 网络的输出也应尽量使 Sigmoid 型函数工作于非饱和区, 因此, 首先要对样本进行规整, 以符合上述要求。

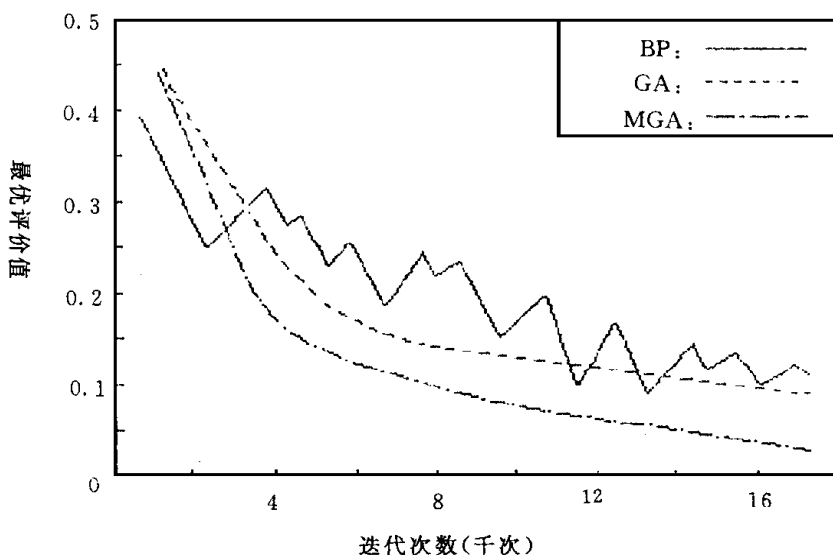


图 4 三种方法的运算结果比较

在实验过程中, 各个参数的取值如下:

BP 算法中, 学习率  $\eta = 0.7$ ; 惯性系数  $\alpha = 0.2$ 。

(下转第 112 页)

$$\begin{aligned} K_1(x_1) &= 0, K_1(x_2) = 1.57, K_1(x_3) = -1, K_1(x_4) = -0.65 \\ K_2(x_1) &= -1.6, K_2(x_2) = 2.8, K_2(x_3) = 0, K_2(x_4) = -0.86 \\ K_3(x_1) &= -0.31, K_3(x_2) = 0.083, K_3(x_3) = -0.054, K_3(x_4) = 0.158 \end{aligned}$$

有

$$K_1(P_0) = 0.2464, K_2(P_0) = 0.52, K_3(P_0) = -0.00356$$

从而

$$K_2(P_0) = \max_{j=1,2,3} K_j(P_0)$$

知  $P_0$  即  $N(1)$  区属于“中效益”区

同理可得: 高效益区:  $\{N(2)\}$ , 中效益区:  $\{N(1), N(3), N(4), N(5), N(6), N(7)\}$ .

### 3 结语

本文提出的物元评判方法简便易行, 对于经济决策效果好, 与其它方法相比具有优越性, 它可以推广应用於其它相关问题

### 参考文献

- 1 蔡文. 物元模型及其应用. 北京: 科学技术文献出版社, 1994
- 2 山西省统计局. 山西统计年鉴, 山西人民出版社, 1987
- 3 Cai Wen. Method of Matter Elements Analysis Busefal, 1984(20)

(上接第 69 页)

实数 GA 中, 群体容量  $n = 40$ , 交叉率  $P_c = 0.12$ , 变异率  $P_m = 0.1$ , 亲代度量  $S = 0.9$ , 各参数在算法运行过程中保持不变。

MGA 中, 常量  $k_c = 0.1$ ,  $k_m = 0.11$ 。

三种算法的运行结果如图 4 所示。

实验结果显示了遗传算法快速学习网络权重的能力, 并且能够摆脱局部极小点的困扰。组合 GA 与 BP 可以改善单纯使用 GA 确定的参数组合比较粗糙以及单纯使用 BP 收敛性差、易陷入局部极小等问题。

### 参考文献

- 1 Bornholdt A. General A symmetric Neural Networks and Structure Design by Genetic Algorithms Neural Networks, 1992, 5(2): 327~ 334
- 2 David J J, Frenzel J F. Training Product Unit Neural Networks with Genetic Algorithms IEEE Expert, 1993, 8(5): 26~ 33
- 3 Schaffer J D, Whitley D (ed). Proceedings of the Workshop on Combinations of Genetic Algorithms and Neural Networks 1992, Los Alamitos, CA: The IEEE Computer Society Press, 1992
- 4 李敏强, 寇纪淞, 纪仕光. 基于网络描述的系统模型及其管理系统. 复杂巨系统理论·方法·应用. 北京: 科学技术文献出版社, 1994
- 5 刘豹. 模糊工程. 决策与决策支持系统, 1995, 5(3): 1~ 5