

基于免疫进化规划的多层前馈网络设计*

曹先彬 刘克胜 王照法

(中国科学技术大学计算机科学技术系 合肥 230026)

摘要 文章用一种免疫进化规划来设计多层前馈神经网络,该免疫进化规划在保留传统进化规划的随机全局搜索能力的基础上,引进生物免疫中抗体通过浓度相互作用的机制和多样性保持机制.免疫进化规划的全局收敛性更优,并且具有很强的自适应环境的能力.实验结果验证了免疫进化规划在设计神经网络时的高效能.

关键词 进化规划,未成熟收敛,免疫,浓度,多样性保持.

中图法分类号 TP183

多层前馈网络在神经网络的研究中占有十分重要的地位,也是目前应用最为广泛的神经网络模型之一.近 10 年来,多层前馈网络一直受到普遍的重视,在模式识别和分类方面获得了很大成功^[1,2].

用多层前馈网络解决问题的过程主要是设计适当的网络结构和构造相应的网络权值的过程.目前的设计方法主要有两种.前几年一般采用探测训练法^[3,4],权值是通过不断前向计算和反向传播得到的,而结构的获取采用逐步增减式的探测方法,这既不能避免反向传播算法所固有的不足(特定问题的网络结构难以确定,网络训练慢且经常陷入局部收敛),也必将使训练出的结果网络的优劣依赖于具体的探测过程;并且一旦当前探测过程失败,目前得到的中间结果也不能继续沿用.近几年来,人们提出用进化算法 EA(evolutionary algorithms)来设计神经网络^[5].进化算法包括遗传算法 GA(genetic algorithm)、进化策略 ES(evolution strategy)和进化规划 EP(evolutionary programming).作为一类随机优化处理方法,进化算法结合了达尔文的适者生存规律和随机信息交换思想,可使优化过程趋于全局极优,在这一点上它优于传统优化方法.但进化算法仍然存在许多不足.例如,由于初始群体是随机生成的,算法在解群分布不均匀时易出现未成熟收敛,陷入局部极优;由于在整个优化过程中多采用固定不变的进化策略,基本未考虑遗传群体的环境变化,使个体自适应于环境的能力较弱.用进化算法来设计神经网络也同样存在这一问题.

本文用一种改进的进化规划来设计多层前馈网络.这种算法将进化规划同生物免疫思想中的浓度机制及个体多样性保持策略相结合,既保留了进化规划的全局搜索特性,又能在很大程度上增强个体适应环境的能力,避免未成熟收敛.实验结果显示出该算法在设计多层前馈神经网络时的优越性能,并表明该算法同样可用于设计其他神经网络.

1 免疫进化规划

进化规划是进化算法的一种.它与遗传算法的不同主要在于,遗传算法关注的是个体行为的遗传,通过个体之间的信息交换来产生下一代个体,而进化规划侧重于群体的行为变化,只采用变异一个操作算子.本文选择进化规划来设计神经网络是基于以下两点考虑.一是由于进化规划采用实数编码方案,这就省略了二进制数编码方案复杂的网络编码表示和解码计算,更适合表示和操作神经网络;二是进化规划只需用变异算子,避免了在遗

* 本文研究得到国家自然科学基金和安徽省“九五”重点攻关项目资金资助.作者曹先彬,1969年生,博士,讲师,主要研究领域为进化计算,神经网络,专家系统.刘克胜,1968年生,讲师,主要研究领域为进化计算.王照法,1948年生,教授,博士生导师,主要研究领域为计算智能,人工生命.

本文通讯联系人:曹先彬,合肥 230027,中国科学技术大学计算机科学技术系

本文 1998-07-21 收到原稿,1998-12-01 收到修改稿

仿算法中交叉算子在操作时很容易破坏已进化出的网络结构这一问题。

在设计神经网络时,进化规划是比较好的一种. 1994年,在 IEEE Transactions on Neural Network 发表进化算法与神经网络结合的专集前后,已有少数工作涉及用进化规划自进化设计神经网络^[6,7]. 但由于进化算法同样无法很好保持个体的多样性以及个体的进化不能自适应环境,导致基于进化算法的神经网络设计还远远不能满足要求。

我们从生物免疫机制得到启示. 生物免疫原理认为^[8],生物的免疫系统对外来侵犯的抗原可自动产生相应的抗体来抵御,抗体与抗原结合后,会通过一系列的反应而破坏抗原,并且抗体之间也有相互刺激与抑制反应. 这种反应是根据抗体浓度进行的. 生物的免疫过程有两个明显的特点。

• 抗体的多样性. 生物免疫机制保证各种抗体在任何时刻都存在于生物机体中;

• 抗体间的相互反应是基于浓度进行的. 抗体的浓度越高,则越受到抑制;浓度越低,则越受到促进. 这使抗体具有自适应环境的能力。

根据从生物免疫机制中抽象出的计算模型,我们将待求解的问题对应为抗原,而问题的解对应为抗体,在进化规划的基础上引入浓度机制,提出免疫进化规划 IEP(immune EP)算法,并用来说设计神经网络。

免疫进化规划的算法框图如图 1 所示。

IEP 首先接收一个抗原(对应特定问题),然后随机产生一组初始抗体(对应初始候选解,也即初始进化个体);接着计算每一抗体的适应度;通过抗体多样性保持策略来维持抗体的多样性;对抗体进行高斯变异;再通过基于浓度的群体更新策略生成下一代抗体群,直至满足终止条件,算法结束. 其中,抗体的浓度 C 定义为群体中具有最大适应度或近似最大适应度的抗体个数 t 与群体中抗体总数 N 的比率,即 $C=t/N$ 。下面结合多层前馈网络的设计进一步说明 IEP 算法。

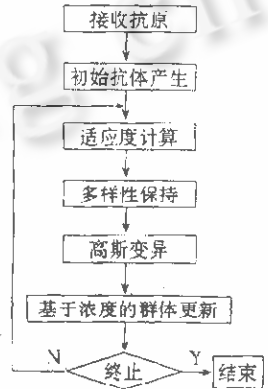


图1 免疫进化规划算法框图

2 基于 IEP 的多层前馈网络设计

在设计 NN 时,假设其输入(包括输入节点数及输入值)和输出(包括输出节点数及输出值)是已知的,节点转换函数也预先确定,如取 S 型函数,隐节点和输出节点的阈值为 θ ;另外,下列条件必须满足:

- Rule1. 输入节点没有输入联接;
 - Rule2. 输出节点没有输出联接;
 - Rule3. 节点 x, y 间至多只有一个联接存在.
- 对本文考虑的多层前馈网络,则还必须满足
- Rule4. 前层节点到后层节点没有联接存在.

2.1 抗体编码

每个抗体对应一种网络结构,我们将隐节点数及网络权值进行混合实数编码,每一抗体码串如图 2 所示。

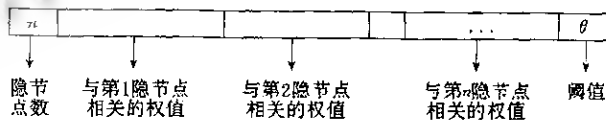


图2 抗体编码形式

抗体码长随隐节点数目的不同而不同。

2.2 适应度函数的设计

设抗体 P_i 对应的网络的误差为 e_i , 能量函数为 E_i , 则适应度函数 $F(i)$ 可直接定义为 e_i 或 E_i 的函数. 本文提供了下面 3 种可选择的方案:

$$F(i) = 1/e_i. \tag{1}$$

$$F(i) = e^{-Y}, \quad (2)$$

$$F(i) = 1/(E_i + const). \quad (3)$$

其中 $Y, const$ 为大于零的常量, 引进 $const$ 的目的是避免算法因为分母为 0 而产生的溢出. 对于本文所考虑的单隐层前馈网络, 有

$$e_i = \sum_p \sum_{out} (T_{p,out} - Y_{p,out}), \quad (4)$$

$$E_i = \sum_p \sum_{out} (T_{p,out} - Y_{p,out})^2, \quad (5)$$

$T_{p,out}$ 和 $Y_{p,out}$ 分别为第 p 个训练样本的第 out 个输出节点的期望输出和实际输出.

2.3 新抗体生成

在 IEP 中, 与抗体 P_i 对应的新抗体 P'_i 的生成需要经过 3 个阶段.

(1) 多样性保持

IEP 的多样性保持策略只针对隐节点数. 为保持抗体中隐节点分量的多样性, IEP 规定

$$\text{Minnum} \leq \text{Num}(k) \leq \text{Maxnum}, \quad (6)$$

其中 $\text{Num}(k)$ 表示隐节点数为 k 的抗体数目, $\text{Maxnum}, \text{Minnum}$ 分别是其最大值和最小值, 由具体问题决定. 在下面的实验中, 我们设 $\text{Maxnum} = 2 \times (\text{IN} + \text{ON}), \text{Minnum} = 1, \text{IN}, \text{ON}$ 分别为特定网络的输入和输出数.

(2) 高斯变异

变异只针对每个抗体中的权值分量进行. 操作过程如下:

(a) 将各抗体 P_i 解码为相应的网络结构;

(b) 依次按下式变异网络的所有权值

$$W'_{ij} = W_{ij} + a \cdot \sqrt{F(i)} \cdot \mu(0, 1), \quad (7)$$

其中 $F(i)$ 为抗体 P_i 的适应度, a 为 $-1 \sim 1$ 之间的一个参数, $\mu(0, 1)$ 为高斯算子;

(3) 重新由隐节点分量和权值分量组成一个新抗体.

2.4 基于浓度的群体更新

在 \mathcal{N} 个父代个体每个经过一次变异产生 λ 个子代个体之后, 传统 EP 利用一种随机竞争 q 选择方法从父代和子代的集合中选出若干个体, 组成下一群体. 具体方法是: 从 $(\mathcal{N} + \lambda)$ 个个体中随机选出 q 个个体, 把它们按适应度值与每个个体比较, 计算出其中比 a_i 差的个体数目 m_i , 并把 m_i 作为 a_i 的得分; 再按得分从大到小对每个个体排序, 选出前若干个个体组成下一群体.

IEP 的群体更新策略结合免疫机制中抗体间基于浓度的相互抑制作用, 引入浓度因子调整个体的随机 q 得分. 抗体的浓度过高, 则在进化过程中容易陷入未成熟收敛. 因此, IEP 需要抑制浓度过高的抗体, 同时保证适应度高的个体被选中的概率大. 具体方法是, 在抗体 i 原随机 q 得分 $P(i).scores$ 的基础上加上基于浓度和适应度的修正值,

$$P(i).scores = P(i).scores + \alpha \cdot C \cdot \left(1 - \frac{\text{Fitness}(i)}{\text{MaxFitness}}\right) \cdot P(i).scores + \beta \cdot \frac{\text{Fitness}(i)}{\text{MaxFitness}} \cdot P(i).scores, \quad (8)$$

其中 α, β 为 $0 \sim 1$ 间的可调参数, $\text{Fitness}(i)$ 为第 i 个抗体的适应度, MaxFitness 为抗体的最大适应度, C 为抗体的浓度.

由上式可以看出, 对于高浓度的 r 个抗体, 其中适应度较高的抗体获得的得分修正反而较少 (上式中的第 2 项); 若抗体的浓度不高, 则上式中的第 3 项保证高适应度的抗体得到的得分修正值相对也高.

3 实验

本节通过两个实验来验证 IEP 算法的收敛效能.

3.1 实验 I——用 BP 网络解决 XOR 问题

XOR 问题被广泛应用于评价一个前馈网络的优劣. 本实验规定: 阈值 $\theta = 0, \text{Maxnum} = 6, \text{Minnum} = 1$; 适应

度函数按式(3)来计算,式中 $const=0.01$;群体规模为 50;式(7)中参数 $a=0.1$;式(8)中 $a=0.5, \beta=0.5$;浓度 C 定义为

$$C = \frac{0.8 \cdot \text{MaxFitness 至 MaxFitness 间的抗体数}}{\text{群体中抗体总数}}$$

在选择时,每次按得分大小从高到低选前半部分个体组成新群体。在规定当网络误差 $e \leq 0.001$ 时算法收敛的前提下,表 1 显示出算法进行 10 次的结果。图 3 显示了几种方法的学习过程。可以看出,在兼顾全局收敛性和收敛速度时,IEP 的效能最为理想。

表 1 IEP 运行 10 次的结果

运行次数	迭代次数	误差 e	是否收敛
1	200	0.000 607	收敛
2	200	0.000 643	收敛
3	200	0.000 687	收敛
4	200	0.000 438	收敛
5	200	0.000 141	收敛
6	200	0.000 430	收敛
7	200	0.000 167	收敛
8	200	0.000 284	收敛
9	200	0.000 002	收敛
10	200	0.000 162	收敛

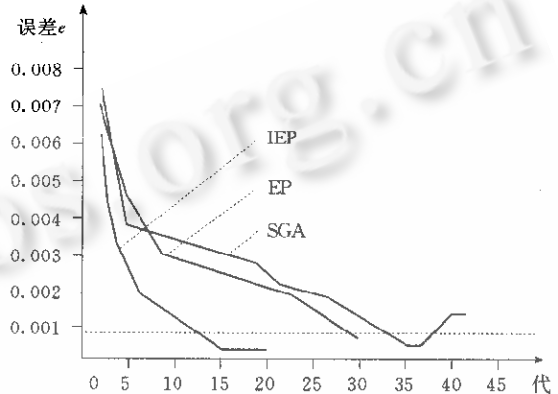


图 3 几种算法的学习过程

3.2 实验 2——农业小麦产量预报

农业作为基础产业,在我国一直受到特别重视。本实验用 IEP 设计预报安徽省农业冬小麦产量的神经网络系统,这一系统的输入有 10 个气象因子,输出为预报的产量,以安徽省 '93~'95 这 3 年为例,相关气象因子的数据见表 2。

表 2 10 个气象因子的数值

年份	$RR_{3中}$	$RR_{2下}$	$RR_{2中下}$	$S_{3中}$	$RR_{4上}$	$Rd_{3下}$	$RRd_{3下-4中}$	$RRd_{4上中}$	$S_{3下-4中}$	$S_{3中-4上}$
93	62.7	0	62.7	52.3	6.3	6.7	296.3	14	181	161.3
94	6.7	13.3	20.0	53	2.7	1.3	71.7	282.7	166.3	174.3
95	14.7	0	14.7	62.7	0	3.7	80.3	43.3	168.3	186.3

其中 RR 表示降水量, Rd 为雨日, S 为日照时间;右下标表示时间间隔,如 $S_{3中-4上}$ 为 3 月中旬至 4 月上旬的日照时间。

在采用 IEP 免疫进化时,规定 $Maxnum=22, Minnum=1$,其他参数与实验 1 一致。表 3 列出了本文的预报结果,同时给出了采用传统专家系统(方法 1)、神经网络专家系统(方法 2)、传统数值拟合(方法 3)以及进化神经网络系统(方法 4)的结果^[9],可以看出,采用 IEP 设计的神经网络系统取得了相对较好的结果。

表 3 几种方法的预报结果

年份	实际产量	方法 1 结果	方法 2 结果	方法 3 结果	方法 4 结果	本文的结果
'93	456	410 -10.0%	490 7.5%	409 -10.5%	420 -7.6%	426 -6.6%
'94	473	432 -8.6%	401 15.2%	414 -12.5%	431 -8.8%	428 -9.7%
'95	506	466 -8.0%	469 7.4%	465 -8.1%	473 -6.5%	484 -4.3%

4 结束语

在用神经网络解决具体问题时,最主要的工作就是设计出适当的网络结构和连接权值。现有的一些方法都

存在不足,进化规划虽然相对较好,但也存在易陷入局部极优和不能自适应环境两大缺陷.本文从生物免疫机制中抽象出抗体之间基于浓度的关系模型,提出一种免疫进化规划算法.此方法在设计神经网络时,既保留了进化规划的强全局搜索能力,又能在很大程度上克服它的不足,在全局收敛性和收敛速度两方面都达到了较好的效能.

致谢 本文研究得到国家自然科学基金资助.此项目编号为 69671029,在此特表示感谢!

参考文献

- 1 Weymaere N, Martens J P. On the initialization and optimization of multilayer perceptions. *IEEE Transactions on Neural Networks*, 1994,5(1):738~751
- 2 Park Y. A comparison of neural net classifiers and linear tree classifiers; there similarities and differences. *Pattern Recognition*, 1994,27(11):1493~1503
- 3 Ash T. Dynamic node creation in backpropagation networks. *Connection Science*, 1989,1(4):365~375
- 4 Azimi-Sadjadi M R *et al.* Recursive dynamic node creation in multilayer neural networks. *IEEE Transactions on Neural Networks*, 1993,4(2):242~256
- 5 Vittorio Maniezzo. Genetic evolution of the topology and weight distribution of neural networks. *IEEE Transactions on Neural Networks*, 1993,5(1):39~53
- 6 Fogel D B *et al.* Evolving neural networks. *Biological Cybernetics*, 1990,63(1):487~493
- 7 Angeline P J *et al.* An evolutionary algorithm that constructs recurrent neural networks. *IEEE Transactions on Neural Networks*, 1994,5(1):54~65
- 8 吴敏毓,刘恭植.医学免疫学.合肥,中国科学技术大学出版社,1995
(Wu Min-yu, Liu Gong-zhi. *Medicine Immune Science*. Hefei: Press of University of Science and Technology of China, 1995)
- 9 曹先彬.基于 GA 和 NN 驱动的智能处理系统研究[博士学位论文].合肥,中国科学技术大学,1996
(Cao Xian-bin. *Research on intelligent processing system based on GA and NN* [Ph.D. Thesis]. Hefei: University of Science and Technology of China, 1996)

Design Multilayer Feed-forward Networks Based on Immune Evolutionary Programming

CAO Xian-bin LIU Ke sheng WANG Xu-fa

(Department of Computer Science and Technology University of Science and Technology of China Hefei 230026)

Abstract The authors use an immune evolutionary programming to design multilayer feed-forward networks in this paper. The immune evolutionary programming retains the ability of stochastic global searching of traditional evolutionary programming, and draws into the interaction mechanism based on density and the diversity maintaining mechanism which exists in living beings' immune procedure. The immune evolutionary programming has better global convergence and very strong self-adaptive ability with environment. The experimental results prove the high efficiency of the immune evolutionary programming in designing neural networks.

Key words Evolutionary programming, premature convergence, immune, density, diversity maintain.