

基因算法辅助人工神经网络的结构设计

李大字

(北京化工大学自动化系, 北京 100029)

摘 要: 提出了以基因算法为基础的人工神经网络结构设计基础。首先从连接权的演化入手, 研究了基因算法的实现过程。在结构设计中, 提出了稀疏化的编码方法。仿真结果表明这种优化方法对于神经网络的选取是有效的。

关键词: 人工神经网络; 基因算法; 结构

中图分类号: TQ 015.9

引 言

人工神经网络(ANN)作为一种新型的信息描述和处理方式, 以其独特的优点引起了人们的关注。在神经网络的应用中, 多层网的研究仍有许多基本问题尚未解决^[1], 如学习速度、结构选取及泛化能力等。本文研究了其改进算法, 以便更好地应用。

基因算法(GA)是一种基于自然选择原理和自然遗传机制的搜索与寻优方法。基因算法的全局寻优能力决定了它必然在许多场合具有广泛的应用价值。作者曾研究了它在参数优化方面的应用^[2]。本文探讨了将基因算法与传统的 BP 算法相结合这一命题, 对前向网络结构的选取及泛化能力作了应用研究。

1 前向网络权连接的基因学习过程

基因演化过程的基础是由大量个体组成的种群, 每个个体可看成是搜索空间的一个状态点, 即: 问题的一个可能的解。能够较好地适应给定环境的个体, 就有较大的生存机会并繁衍后代。在神经网络的演化过程中, 个体为可能的权向量。基因算法学习神经网络权连接的过程可以看作是对适应度函数的连续优化过程, 优化过程中每个个体繁殖后代的数目取决于各自的适应度。本文研究的是一个有监督的学习过程, 所以适应度的信息是通过给定值与网络预测值之间的误差经计算得到的。基因学习过程是一个不断迭代的过程, 其过程如图 1 所示。

基因算法的使用意味着必须对个体进行基因型

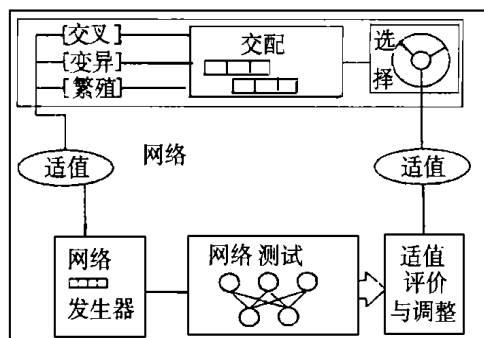


图 1 神经网络的演化过程

Fig. 1 Evolution of neural network

描述, 以便允许基因算法对其进行方便的操作, 此过程就是编码过程。

所发生的网络的基因型, 经过相应的译码过程, 构造成表现型——实际的网络权值。这些具有不同初始权值的网络, 在网络模拟器中经过相同环境(给定的训练数据)的考验, 得到各自的适值。网络对训练行为好坏的评价函数为:

$$E_{\min} = \frac{1}{PL} \sqrt{\sum_{p=1}^P \sum_{l=1}^L (d_{pl} - O_{pl})^2}$$

其中: P 是训练模式数; L 是网络输出向量的维数; d_{pl} 是期望的输出值; O_{pl} 是实际输出值。

人们学习某种知识时, 知识以什么样的形式表示出来, 对人们的记忆和学习效率影响很大。编码方案的选择十分重要, 基因算法的参数编码有两个基本原则。编码方案可以用数学方法进行严格检查, 但并不实用。有些编码方案, 从理论上说是很好的, 但在实际应用中却不一定能达到预定的效果。编码方案的选择不只是理论上的问题, 更重要的是实践问题。

作者采用了试验的方法,对常规的二进制编码和实值编码进行了比较,最终选择了二进制编码。最为直接的方法是将网络的每个权值编码为二进制子串的形式,整个二进制串是由这些子串连接在一起构成的,因此对神经网络的权值编码为多参数编码。本文选择了一个较为实用的多参数编码方案——多参数级联定点映射编码法。假设有 N 个网络权值,则利用二进制编码可描述如下:

$$\begin{array}{cccc} L_1 & L_2 & L_i & L_N \\ |b_{11} b_{12} \dots b_{1L_1}| & |b_{21} b_{22} \dots b_{2L_2}| & \dots |b_{ij} \dots| & |b_{N1} b_{N2} \dots b_{NL_N}| \end{array}$$

其中: $b_{ij} \in \{0, 1\}$

优化对象是前向网络,其激励函数为 $f(x) = \tan h(x)$, 对一个具有一层隐藏层的网络来说,如果指定了隐藏层大小,网络中的权值和编码个数可由下式得出:

$$N = N_i N_h + N_h N_0 + N_h + N_0$$

其中: N 为总的权值数目; N_i 为输入向量的个数; N_0 为输出向量的个数; N_h 为隐含节点数。

所有需要编码的参数均为权值,取相同的编码长度 L , 而编码长度取决于所需要的精度。假如权值的绝对值范围取为 $[0, 1]$, 取精度 优于 0.05, 可得到编码长度的估计值^[3]:

$$L = \frac{\ln \left(1 + \frac{U_{\max} - U_{\min}}{\ln 2} \right)}{\ln 2}$$

权值范围有正有负,在常规的编码之外,作者加一个符号位 S , $S = 1$ 为负, $S = 0$ 为正。

整个迭代过程如下:根据相应的适值选择的两个个体,参加下一代的繁殖;按概率 P_c 进行交叉,否则进行变异或复制。这样,新一代产生了,对其进行评价并将它们依适值插入到种群中,继续迭代直至收敛或达到预先给定的最大迭代次数。

2 前向网络的结构优化

前向网络的结构选取由于没有理论上的指导,长期以来只采用根据经验的试凑法,为保证获得需要的精度,所需的网络往往偏于复杂,即未能做到人工神经网络的“最小实现”。网络结构的不必要的复杂带来一系列不利的影响。

本文网络结构优化的基本思想是:根据双目标函数的思想,在上述权连接的基因演化方法的基础上,为提高神经网络的品质,尽可能节俭地设置网络中的节点数。在考察网络好坏的标准中增加这一指

标,不能仅仅以网络映射精度单一指标来评价。换句话说,需要实现网络的稀疏化连接,而不是全连接。除目标函数 E_{\min} 之外,引入表征网络复杂程度的函数 E_c :

$$E_c = (2 \cdot N_i + N_h) \sum_{i,j} w_{ij}^2$$

其中: N_i 为输入层的节点数(不包括被删除的节点); N_h 为隐藏层的节点数(不包括被删除的节点); w_{ij} 为从节点 i 指向节点 j 的网络权值。

双目标函数的实现保证在网络学习精度的前提下,尽量减小网络结构的复杂度,这在数学上表现为求解一个带约束的极值问题 E_{\min} , 约束条件为: $E_{\min} < \dots, > 0$ (为所允许的最大均方误差)。

常规的网络学习算法(如 BP 算法)难以直接处理非连续函数的形式,基因算法的特性使得它较适合于网络结构的优化,具体过程如下^[4]:

1) 根据所需解决问题的难易程度,给出一个相对于该问题来说比较复杂的网络,分别给出结构优化过程中要用到的初步训练精度 G_1 、最终训练精度 G_2 与网络复杂度函数 E_c 值;

2) 首先,以 G_1 为目标函数,利用改进了的基因算法对上面给定结构的网络进行学习,直到达到精度为止;

3) 以误差满足精度要求为约束条件,以 E_c 为目标函数,采用常规基因算法为训练手段,继续进行网络的学习过程,直到收敛为止;

4) 以 G_2 为目标函数,对上一步得到的网络进行学习,直到最终达到理想的精度为止。

在编码时,基于上述的二进制编码方式,将网络的连接状态直接编码,即“0”表示不连接,“1”表示连接。连接位列于符号位之前,与权值的编码共同构成基因型。

3 结果分析

本文选择了炮打飞行物的问题,使用神经网络对飞行物的位置进行预测。一个以初速度 v_0 、初始角度 抛出的飞行物在 10 s 以后的横向位移和纵向位移的表达式:

$$\begin{cases} x = v_0 \cos \theta \cdot t \\ y = v_0 \sin \theta \cdot t - \frac{1}{2} g t^2 \end{cases}$$

式中: g 为加速度; t 为飞行时间。飞行物的初速度和初始角度按下式给出:

$$\begin{cases} v_0 = 0.116(\text{m s}^{-1}) + (0.2 - 0.116)(\text{m s}^{-1}) \cdot t_1 \\ = (60 + 30 t_2) \cdot / 180 \end{cases}$$

式中: t_1, t_2 是 0~1 之间的随机数。

对于该问题, 选择隐含层节点数为 10, 则网络

结构为 3-10-2。第一步给出的精度为 $G_1 = 85\%$, 在以 E_c 为目标函数的学习达到收敛之后, 所得到的一组稀疏化连接的网络示于表 1。其所表达的网络稀疏化形式如图 2 所示。

表 1 3-10-2 网络的连接

Table 1 Connection of a 3-10-2 neural network

隐含层		$j=1$	$j=2$	$j=3$	$j=4$	$j=5$	$j=6$	$j=7$	$j=8$	$j=9$	$j=10$
输入层	$i=1$	1	1	0	1	1	1	0	0	1	0
	$i=2$	1	0	0	1	0	0	0	0	1	0
	$i=3$	1	0	1	0	0	1	0	1	1	1
输出层	$k=1$	1	0	1	1	1	0	0	0	0	0
	$k=2$	1	1	0	0	0	1	0	1	0	0

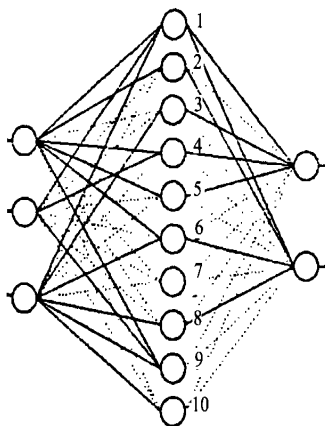


图 2 3-10-2 网络的稀疏化连接 (虚线部分表示不连接)

Fig. 2 Sparse connection of a 3-10-2 neural network

从图 2 可以看出, 网络的输入节点并没有被稀疏掉, 说明在演化过程中, 只要给出适当的搜索指导, 即适值函数, GA 就能正确地完成任务。在上述网络结构的基础上, 继续对它进行训练, 直到收敛到 G_2 。

隐节点 7、9、10 都是被稀疏掉的点, 即没有输出到输出层的节点。去掉这些节点后, 网络的实际结构如图 3 所示。

用一组训练网络时未使用过的数据来检验其泛化能力, 得表 2。

训练两种网络的收敛趋势示于图 4, 命中率为 x , 迭代次数为 n 。

由上可见, 利用稀疏化编码演化出的稀疏化网络结构可以提高网络的泛化能力及训练速度。

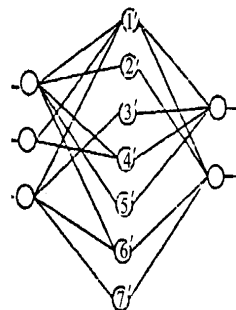


图 3 实际的 ANN 结构

Fig. 3 The actual ANN configuration

表 2 网络泛化能力的比较

Table 2 Generalization ability comparison of the two neural networks

t_1	t_2	E_{\min} (稀疏连接)	E_{\min} (满连接)
0.480 4	0.050 8	0.002 67	0.004 04
0.086 3	0.047 6	0.002 38	0.003 20
0.989 1	0.710 7	0.006 54	0.010 45

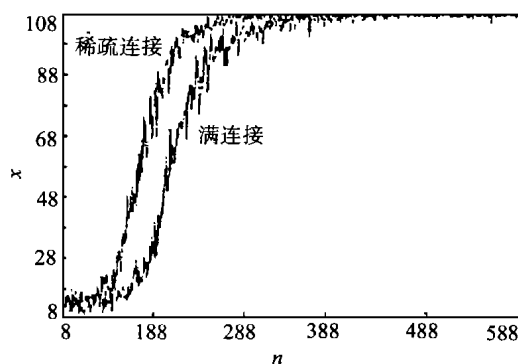


图 4 收敛趋势比较

Fig. 4 Comparison of convergence tendency

(下转第 55 页)

蚀金属材料及防护技术分册. 北京:化学工业出版社,
1990

[9] 黄少斌,肖锦. 酸缓蚀剂在 A3 钢表面吸附成膜的研究. 材料保护,1998,31(2):3~4

Study of occluded cell corrosion of A3 steel in alkali NaCl solution

XU Chun-chun ZHANG Xiao-bo LIU Your-ping

(Department of Applied Chemistry, Beijing University of Chemical Technology, Beijing 100029, China)

Abstract: A simulated occluded corrosion cell (OCC) was used to study chemical changes within corrosion cracks or crevices of mild steel (A3) in 0.01 (mol L⁻¹) NaCl solution at 50 °C. It has been found that as the anodic current passes through the cell, the pH value falls quickly at the beginning, then slowly down to about 4.55. Meanwhile Cl⁻ migrates into the occluded cell, and the Cl⁻ concentration is proportional to polarization time. Polarization curves show that in solutions with a pH value above 11, obviously visible localized corrosion occurs to A3 steel.

Key words: partial corrosion; occluded cell; pH

(上接第 48 页)

参 考 文 献

- [1] 焦李成. 神经网络系统理论. 西安:西安电子科技大学出版社,1992
- [2] 高彦臣,李大字. 利用基因算法实现参数优化的研究.

北京化工大学学报,1995,22(1):53~58

- [3] 田盛丰,陈峰,任建宏. 人工智能原理与应用. 北京:北京理工大学出版社,1993
- [4] 李大字. 基因算法辅助人工神经元的结构设计. [学位论文]. 北京:北京化工大学,1995

Genetic algorithm aided architecture design of artificial neural network

LI Da-zi

(Department of automation, Beijing University of Chemical Technology, Beijing 100029, China)

Abstract: A system for the architecture design of artificial neural network based on the genetic algorithm (GA) is proposed. With the evolution of connection weights first, the author researches into the materialization of GA, and a combined genetic/back-propagation learning algorithm is hence proposed. In the development of neural network's architectures, a sparse coding is put forward. Simulation results show that this optimization method is efficient for selecting ANN's structures.

Key words: artificial neural network; genetic algorithm; architecture