

文章编号:1001-9081(2007)06-1426-04

一种优化神经网络结构的遗传禁忌算法

王淑玲,李振涛,邢 棉

(华北电力大学 数理系,河北 保定 071003)

(lixiangwangshuling@126.com)

摘 要:常用的神经网络是通过固定的网络结构得到最优权值,使网络的实用性受到影响。引入一种基于方向的交叉算子和禁忌变异算子,同时把禁忌算法(TS)引入标准遗传算法,结合标准遗传算法和禁忌算法的优点,提出一种优化神经网络结构的遗传禁忌混合算法,实现了网络结构和权值同时优化。仿真实验表明,与遗传算法和禁忌算法相比,该算法优化的神经网络收敛速度较快、预测精度较高,提高了网络的处理能力。

关键词:遗传禁忌算法;神经网络;优化;算子

中图分类号: TP183 **文献标识码:** A

Genetic and tabu search algorithm for the optimization of neural network structure

WANG Shu-ling, LI Zhen-tao, XING Mian

(Department of Mathematics and Physics, North China Electric Power University, Baoding Hebei 071003, China)

Abstract: A conventional Neural Network often optimizes the weights through invariable network structure, which has limited the extensive use of the Neural Network. The crossover operator based on direction and Tabu search mutation operator was introduced. This paper put forward Genetic and Tabu search algorithm to train the neural networks, combining the merits of genetic algorithm and that of Tabu search algorithm, which makes weights and structure of artificial neural networks be optimized together. The result shows that the neural network optimized by using the presented algorithm has the advantages of quicker convergence rate and higher precision, compared with genetic algorithm and Tabu search algorithm, and that the processing ability of networks is also raised.

Key words: genetic and tabu search algorithm; neural networks; optimization; operator

0 引言

神经网络广泛应用于故障检测、语言识别、模式识别、水文预报、信号处理等领域。神经网络具有很强的自适应、自组织和自学习的能力,具有大规模并行运算的能力。常用的神经网络系统的行为主要由两方面因素决定:拓扑结构和节点间连接权值。一般来讲,在应用神经网络过程中只是得到最优权值,结构在整个过程中是固定的,而且在确定结构时,经常根据人为经验或经验公式先拟定一个结构,在此基础上不断地训练网络、调整结构,最后得到确定的网络结构,进而应用这个固定的网络结构得到最优的权值。因此,构造一个准确定量构效关系人工神经网络模型需要耗费大量的时间,有时利用该方法构造的模型结构还可能不是最佳的^[1]。近年来,用遗传算法(Genetic Algorithms, GA)解决神经网络优化设计问题受到广泛重视。GA在寻优时搜索方向不停的发生变化,具有较强的全局搜索能力。但是,在应用遗传算法寻优的过程中,仍存在着困难^[2]:遗传算法对好的个体易造成破坏、易出现早熟。因此,许多学者采取了遗传算法和其他算法相结合^[2,3]的措施,如先用遗传算法进行全局搜索,确定最优

解所在的区域后,再用BP算法(拟牛顿LM算法)进行局部搜索,但仍然存在耗时较多、精度较低的不足。

本文把基于方向的遗传交叉算子和禁忌^[4]变异算子引入了遗传算法,提出了遗传禁忌混合算法,使在得到神经网络最优权值的同时,也自动得到最优网络结构。仿真研究表明,此算法在以负荷作为训练样本的神经网络中具有良好的实用性。

1 遗传禁忌算法优化前向神经网络

1.1 概述

1)神经网络

Kolmogorav定理证明在有合理的结构和恰当的权值条件下,三层前向网络能以任意精度逼近有界非线性函数。

对于三层神经网络来说,输入层、输出层节点的个数由目标函数结构决定,故而在优化网络结构时,是优化它的隐节点的个数。本文将对应用最为广泛的三层前向网络(如图1所示)进行优化设计,确定出最佳隐层节点个数,同时生成落在全局最优区域的权值。

2)遗传算法^[5,6]根据生物进化理论和遗传变异理论,通

收稿日期:2006-12-20;修订日期:2007-02-12 基金项目:国家自然科学基金资助项目(50077007)

作者简介:王淑玲(1972-),女,河北保定人,硕士研究生,主要研究方向:信息处理技术;李振涛(1975-),男,河北保定人,硕士研究生,主要研究方向:智能优化理论及其应用;邢棉(1962-),女,河北保定人,教授,主要研究方向:电力市场与电力负荷预测研究。

过选择复制和遗传算子的作用使种群不断进化,最终收敛到优化解。遗传算法是一种并行搜索过程,它具有较强的全局寻优能力。

3) 禁忌算法(TS)^[4]是一种亚启发式随机搜索算法。TS的特点是通过禁忌表和禁忌准则来避免迂回搜索。为了回避局部邻域搜索陷入局部最优的不足,禁忌搜索算法用禁忌表记录下已经到达过的局部最优点,在下一次的搜索中,利用禁忌表中的信息避免或有选择的搜索这些点,以此来跳出局部最优点。因此TS具有较强的局部搜索能力。TS的缺陷是在搜索过程中初始解只能有1个。

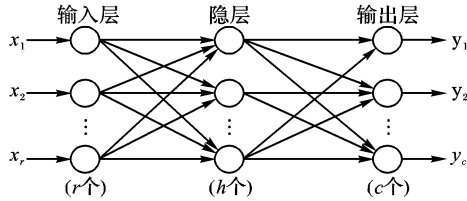
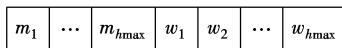


图1 3层BP神经网络

1.2 优化设计^[3]

1.2.1 编码方式

为了计算简便,本文省略神经网络阈值。个体的长度由最大隐节点个数决定,即 $l = (r + c) \cdot h_{\max} + h_{\max}$,其中: r 为输入节点数、 h_{\max} 为最大隐层节点数、 c 为输出节点数。将同一个隐层节点所连接的权值排在一起,当删除或增加隐层节点时,方便操作。将隐节点个数及网络权值进行混合实数编码,另外对隐含层节点增加了一个0-1组成的数串用作控制码。控制码主要是控制隐节点的个数,0-1数串中的一个数控制一个隐层节点。具体编码方法^[5]如下所示。



其中, m_1 是表示第一个隐层节点有效性的0-1控制码, $m_{h_{\max}}$ 是表示第 h_{\max} 个隐层节点有效性的0-1控制码, w_1 是与第一个隐层节点相关的权值, w_2 是与第二个隐层节点相关的权值, $w_{h_{\max}}$ 是与第 h_{\max} 个隐层节点相关的权值。

初始群体中不同隐层节点个数的个体按照一定的比例随机产生。当0-1数串中的数为0时,其所对应的隐含层节点对输出层没有作用(即相当于不存在此神经元),当数串中的数为1时,就相当于此隐含层节点对输出层有作用。

采用上述方法编码,每个个体译码后的网络结构都满足前面所述的研究前提条件,因此不存在所对应的网络结构无效的个体。

1.2.2 适应度函数的确定及选择运算

本文采用如(1)式所示适应度函数:

$$F = \frac{1}{E} \cdot [1 + a \cdot (1 - \frac{N_y}{N_{y_{\max}}})] \quad (1)$$

式中: F 为神经网络的适应度, E 为神经网络的误差平方和, a 为隐含层神经元数量的影响系数, $N_{y_{\max}}$ 为隐含层神经元的最大数量, N_y 为对输出层节点有作用的隐含层节点的数量。对于选择操作采用轮盘赌方法。

1.2.3 交叉率和变异率的确定

本文采用自适应的交叉率和变异率,其公式如下:

$$P_c = \begin{cases} k_1 \cdot (f_{\max} - f^1) / (f_{\max} - f_{avg}), & f^1 \geq f_{avg} \\ k_2, & f^1 < f_{avg} \end{cases}$$

$$P_m = \begin{cases} k_3 \cdot (f_{\max} - f_i) / (f_{\max} - f_{avg}), & f_i \geq f_{avg} \\ k_4, & f_i < f_{avg} \end{cases}$$

式中 k_1, k_2, k_3, k_4 是取值范围为 $[0, 1]$ 的常数, f^1 是要交叉的两个个体适应度中较大的一个, f_i 是要变异的个体的适应度, f_{\max} 是种群中最大的适应度, f_{avg} 是种群的平均适应度。

1.2.4 交叉算子

对控制基因和实数编码的基因分别采用不同的交叉算子。对于控制基因,采用两点交叉;对于用实数编码的代表各个节点权值的基因,采用下面所述交叉方法:

假设要参与交叉操作的两个个体分别为 X_i, X_j ,两个个体上相对应位的权值分别为 x_i, x_j ,并且 X_i 的适应度高于 X_j 的适应度,定义式(2)、(3)所示的 Δ_i^1, Δ_j^1 两个中间变量:

$$\Delta_i^1 = \begin{cases} \min(x_i + \frac{1 + P_c}{2}(x_i - x_j), x^{\max}), & x_i \geq x_j \\ \max(x_i + \frac{1 + P_c}{2}(x_i - x_j), x^{\min}), & x_i < x_j \end{cases} \quad (2)$$

$$\Delta_j^1 = \begin{cases} \max(x_j + \frac{1 - P_c}{2}(x_j - x_i), x^{\min}), & x_i \geq x_j \\ \min(x_j + \frac{1 - P_c}{2}(x_j - x_i), x^{\max}), & x_i < x_j \end{cases} \quad (3)$$

式中 x^{\max}, x^{\min} 分别为 $x_i(x_j)$ 的取值上限和下限,这样,交叉后所产生的两个新个体 y_i, y_j 可以由式(4)、(5)确定:

$$y_i = \frac{1 + P_c}{2} \cdot \Delta_i^1 + \frac{1 - P_c}{2} \Delta_j^1 \quad (4)$$

$$y_j = \alpha \cdot x_i + (1 - \alpha) \cdot x_j \quad (5)$$

式中 α 为一个在 $[0, 1]$ 之间的随机数。从(2)~(5)式可以看出,这种交叉算子可以保证产生的两个个体,一个在两个父代所在的区域之间搜索,另一个向适应度高的父代个体的方向搜索新的领域。

交叉后,根据有作用的隐层节点数确定相应长度的权值域,使得有作用的网络权值的数量与网络结构相符。

1.2.5 禁忌变异算子

记禁忌变异算子为TSMO,在禁忌表中记录着染色体的适应值。渴望水平取父代群体适应值的平均值。进行TSMO操作时,先随机取出需要变异的一个个体,再从控制码中随机选出一个数,这个数的变异采用基本变异算子。若这个数变异后的控制基因是1,则从控制基因1对应的权值串中选一随机位,设这一随机位上的权值为 x_c, x_c 的变异为:在区间 $[x_1, x_2]$ 上随机取一数 x_c^1 代替 x_c 。区间 $[x_1, x_2]$ 如式(6)、(7)确定。计算变异后个体的适应值 g_2 。选NS次随机位,得到NS个值 g_2 。把 g_2 的最大值与渴望水平比较,如果比渴望水平好,让 g_2 的最大值对应的个体进入下一代(即使在禁忌表中也被禁忌);如果子代比渴望水平差,但不属于禁忌,也接受这个子代,同时把 g_2 的最大值存入禁忌表;若是属于禁忌,仍让变异前个体进入下一代;若这个数变异后的控制基因是0,则从是1的控制基因中选出一个随机数,重复上面的过程。对需要变异的每个个体,整个过程重复NH次。

$$x_1 = x^{\min} + \left| \frac{x^{\min} \cdot p_m \cdot f_i}{f_{\max}} \right| \quad (6)$$

$$x_2 = x^{\max} + \left| \frac{x^{\max} \cdot p_m \cdot f_i}{f_{\max}} \right| \quad (7)$$

从(6)、(7)两式可以看出,适应度大的个体变异区间较小,适应度小的个体的变异区间较大。这样,既能减少变异操作对好的个体的破坏,又能够保证遗传禁忌算法的搜索能力。变异后,根据有作用的隐层节点数确定相应长度的权值域,使得有作用的网络权值的数量与网络结构相符。

2 训练步骤

步骤1 设定三层前向神经网络的输入节点 r 、隐含层节点 h 、输出节点 c ;

步骤2 在 $0 \sim 1$ 范围内随机产生如 1.2.1 编码所示的 pp 条染色体,得到初始种群,并设初始进化代数 $L = 0$;

步骤3 由设定的选择概率按选择机制对个体选择;

步骤4 由自适应交叉概率按交叉算子进行交叉操作;

步骤5 由自适应变异概率按算子 TSMO 进行遗传禁忌变异操作;

步骤6 把由步骤 3 ~ 5 得到的个体一起作为新一代种群;

步骤7 如果进化代数小于预先设定的最大遗传代数,令 $L = L + 1$,转步骤 3,否则,终止网络的训练,选出最优个体,转步骤 8;

步骤8 将最优个体按顺序拆分,作为 BP 网络的最优权值;

步骤9 输入与训练样本集模式相一致的预测样本,得到预测值,并计算出相对误差。

3 仿真试验及结果

本文在大连地区 1999 年 7 月 1 日至 2003 年 12 月 31 日的历史负荷中筛选了多组样本数据,训练神经网络。在训练神经网络时所用到的主要参数有:输入节点数:6;隐层节点数:4-9;输出节点数:1;种群规模:30;权值范围: $[-1, 1]$ 。

表 1 算法比较

	本文算法	禁忌算法
NS 次数	10	50
NH 次数	20	50

训练神经网络时所用的激活函数是 Sigmoid 函数。在交叉和变异的过程中,可能出现隐层节点为 1、2、3 的个体,但由于它们的适应度较小,不会影响最优结构及最优权值。

下面列出了本文算法独立训练神经网络 7 次对应的最优神经网络结构及权值(如表 2 所示),以及由三种算法得到的负荷预测结果(如表 3 所示)。

表 2 7 次得到的最优神经网络结构及权值

序号	1	2	3	4	5	6	7
隐层节点	4	4	6	5	5	5	4
相应权值	-0.6876	0.0784	0.8818	-0.2186	0.5454	0.4313	0.7616
	0.3376	-0.9186	-0.0254	0.3150	-0.3391	-0.1455	0.2574
	0.4777	-0.7181	-0.6335	0.2324	-0.2220	0.1675	-0.7931
	0.4805	0.2743	-0.0912	-0.4753	0.9116	-0.7175	-0.5662
	-0.3555	-0.5588	-0.4615	-0.9025	0.3944	-0.5940	0.8441
	-0.2362	0.1904	-0.3014	-0.0518	-0.9616	-0.4244	0.0897
	0.2061	0.6409	-0.6454	0.7873	-0.4476	0.6532	0.2231
	-0.2339	-0.2592	0.7782	0.6341	-0.1023	0.5208	0.3367
	-0.5616	0.2632	0.2268	0.8270	0.3066	-0.2007	-0.9047
	-0.7504	0.4961	-0.3757	0.3252	0.0246	0.4547	-0.0022
	0.7760	0.5483	0.3471	0.4565	-0.7041	0.9899	0.1078
	0.7964	-0.1771	-0.8885	-0.2097	0.4080	-0.0938	0.4389
	-0.5226	0.1390	-0.5165	-0.9957	0.9307	0.6788	-0.9857
	-0.1261	-0.6605	-0.0202	-0.5955	0.5335	0.5524	0.3189
	-0.6758	-0.4159	0.0862	0.0720	0.5766	-0.5644	-0.6874
	-0.4661	0.1845	0.9521	-0.1784	-0.7119	0.6089	-0.1454
	-0.8896	0.4684	0.3971	0.1777	-0.1603	0.1204	-0.2746
	-0.7039	0.1674	0.9982	0.3317	0.4178	-0.6928	-0.9087
	0.4078	-0.1860	0.7181	0.3491	-0.7886	0.7172	0.8042
	-0.0475	0.8245	0.9642	-0.0129	-0.1376	0.4218	-0.7207
	0.6903	0.7282	-0.0877	-0.4675	-0.6470	-0.1723	-0.8818
	-0.6159	0.7787	-0.6558	0.0118	-0.2654	-0.6171	0.4832
	-0.8506	-0.8194	0.1955	0.3778	0.0633	0.3810	-0.6022
	0.8474	-0.2342	-0.1696	-0.2992	-0.0582	-0.0823	-0.4683
	0.3650	-0.9103	0.9547	0.2849	0.1631	0.6732	0.1873
	-0.6132	0.2241	-0.7592	-0.8453	-0.0650	0.9170	0.2205
	-0.5482	0.8329	-0.0576	-0.7723	-0.3753	-0.8571	0.4389
	-0.6342	-0.8468	-0.9937	0.5162	-0.1310	-0.8920	0.9609
			-0.8579	0.8410	-0.6102	0.6130	
			-0.1087	-0.6911	-0.3324	-0.5315	
			0.2125	-0.3395	0.6209	-0.2537	
			-0.6375	-0.9461	-0.9796	-0.2844	
			0.5563	-0.7887	0.5538	0.8352	
			0.5610	0.6793	-0.4500	0.7350	
			0.7529	0.8987	0.3624	0.3927	
			0.2082				
			-0.2399				
			0.6128				
			-0.2832				
			0.2585				
			0.0653				
			0.8709				

表 2 说明本文算法确实有效地使神经网络的结构和权值

同时得到了优化。表 3 说明本文算法优化神经网络应用于预

测,确实提高了收敛速度,得到了较高的预测精度。

4 结语

本文提出的遗传禁忌算法是一种同时优化神经网络结构和权值的算法。它是在遗传算法中对控制码和权值域分别引进算子。在交叉过程中,对权值应用自适应交叉和变异率,引

入向适应度高方向进化的交叉算子,对控制基因用两点交叉。在变异时,对控制基因采用基本变异算子,对权值引入禁忌算子,通过禁忌算子作用于变异算子,这样充分发挥了遗传算法和禁忌算法的优点,克服了遗传算法在神经网络优化过程中存在的不足。

表 3 负荷预测结果

日期	实际值	本文算法		标准 GA 算法		禁忌算法	
		预测值	误差	预测值	误差	预测值	误差
6月16日	801.0700	804.2255	0.3924	808.8309	0.9595	809.1028	0.9928
6月23日	805.0100	796.0387	-1.1270	795.5425	-1.1901	814.6648	1.1851
6月30日	815.0800	809.1898	-0.7403	828.5415	1.6126	804.0019	-1.3903
7月6日	832.2500	828.8769	-0.4069	819.6557	-1.5365	821.4860	-1.3103
7月13日	818.2100	811.0454	-0.8834	805.7339	-1.5484	807.9459	-1.2704
7月20日	836.9800	833.5836	-0.4074	851.5124	1.7067	828.1788	-1.0627
7月27日	843.9500	851.3557	0.8699	826.9286	-2.0584	821.1931	-2.7712
绝对平均误差		0.6896		1.5160		1.4261	

参考文献:

- [1] 高大文,王鹏. 人工神经网络中隐含层节点与训练次数的优化[J]. 哈尔滨工业大学学报,2003,35(2):207-209.
- [2] 金朝红,吴汉松,李腊梅,等. 一种基于自适应遗传算法的神经网络学习算法[J]. 微计算机信息,2005,21(10S):49-51.
- [3] 田旭光,宋彤,刘宇新. 结合遗传算法优化 BP 神经网络的结构和参数[J]. 计算机应用与软件,2004,21(6):69-71.
- [4] 贺一,刘光远,雷开友,等. 多层前向神经网络的自适应禁忌搜索训练[J]. 计算机科学,2005,32(6):118-120.
- [5] 叶德谦,康建红,杨樱. 实数编码遗传算法的前向神经网络优化设计[J]. 计算机工程,2005,31(16):163-164,175.
- [6] 徐红. 基于遗传算法的人工神经网络优化设计[J]. 燕山大学学报,2004,28(4):337-340.

(上接第 1419 页)

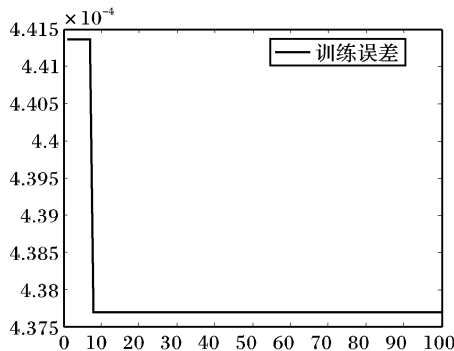


图 1 验证误差函数曲线

表 1 LS-SVM 训练及预测结果

年份	实际值(万人)	预测值(万人)	相对误差
1995	542.12	541.4277	-0.0013
1996	543.45	544.5392	0.0020
1997	549.20	548.8854	-0.0006
1998	553.54	553.1116	-0.0008
1999	557.63	557.3813	-0.0004
2000	562.65	562.7079	0.0001
2001	569.00	568.9100	-0.0002
2002	575.01	575.1641	0.0003
2003	582.56	582.2788	-0.0005
2004	590.08	590.4413	0.0006

3.2 比较

用三层 BP 进行预测,根据 Kolmogorov 定理隐节点数取 $m = 2n + 1$,其中 n 为输入节点数,这里取 11 个隐节点,设定训练均方误差为 0.001,最大进化代数 2500 次,训练及预测结果如表 2 所示。

表 2 LS-SVM 与 BP 网络预测结果比较

	学习误差	预测值(万人)	相对误差	迭代次数
LS-SVM	0.00010262	590.4413	0.0006	100
BP 网络	0.00130857	572.5066	-0.0298	2500

应用遗传算法优选参数后的最小二乘支持向量机进行预测,并与传统 BP 网络预测结果进行比较,说明遗传算法优选参数后的最小二乘支持向量机的结构是比较合理的,训练的速度是比较快的,预测的精度是比较高的。

4 结语

本文尝试了用最小二乘支持向量机回归的方法在非线性经济中的预测,首先用遗传算法确定支持向量机的最佳参数,进而建立预测模型,从预测结果可以看出,该方法用于人口预测具有更高的精度,相对误差比用 BP 网络预测小的多。该方法在训练过程中,所需时间短,用遗传算法优化有关参数后能有效地避免过拟合和欠拟合的现象,具有较强的泛化能力。

参考文献:

- [1] VAPNIK V, LEVIN E, LE CUN Y. Measuring the VC-dimension of a learning machine[J]. Neural Computation, 1994, (6): 851-876.
- [2] VAPNIK VN. The nature of statistical learning theory [M]. New York: Springer, 1995.
- [3] SUYKENS JAK, VANDEWALL J. Least squares support vector machine classifiers[J]. Neural Processing Letters. 1999, 9(3):293-300.
- [4] PELCKMANS K, SUYKENS JAK, DE MOOR B. Building sparse representations and structure determination on LS-SVM substrates [J]. Neurocomputing, Special Issue, 2005, 64: 137-159.
- [5] 济南市统计局. 济南统计年鉴—2005[M]. 北京: 中国统计出版社,2005.