

文章编号:1671-6833(2010)05-0116-05

## 一种改进的进化神经网络优化设计方法

杨华芬<sup>1</sup>, 杨 有<sup>2</sup>, 尚 晋<sup>3</sup>

(1. 曲靖师范学院 计算机科学与工程学院, 云南 曲靖 655011; 2. 重庆师范大学 信息科学与工程学院, 重庆 401331; 3. 重庆航天职业技术学院 计算机工程系, 重庆 400021)

**摘 要:** 基于传统遗传算法优化神经网络时存在的“近亲繁殖”、基因编码冗余和难以确定隐节点数等问题, 提出改进的进化神经网络优化设计方法。通过对网络编码形式的规范, 使得基因编码与功能等价类一一对应, 从而降低编码冗余; 通过节点相关性评价, 使得低于某阈值的节点在交叉操作时被排除, 从而降低节点冗余; 通过把交叉变异概率与种群个体适应度比例相联系, 提出自适应交叉变异概率, 较好保持种群多样性。仿真实验表明, 本方法可以避免“近亲繁殖”以及由此导致的“种群早熟”, 降低编码冗余, 减少学习参数, 提高学习效率。

**关键词:** 神经网络; 遗传算法; 基因编码; 交叉概率; 变异概率; 节点相关性

**中图分类号:** TP183

**文献标识码:** A

### 0 引言

目前, 神经网络已在系统辨识、模式识别、自动控制等领域广泛应用。很多学者用遗传算法优化神经网络, 但是存在编码冗余、难以平衡选择压力和难以保持种群多样性等问题。在编码冗余方面, 文献[1-2]提出一种基于神经网络进化算法的混合系统, 但没有考虑冗余的隐节点会增加计算量, 降低学习速度。文献[3]解决种群早熟问题但存在编码冗余问题。在选择压力方面, 选择压力越大, 容易导致“早熟收敛”<sup>[4-5]</sup>。在种群多样性方面, “早熟收敛”的本质特征是群多样性较差<sup>[6]</sup>。文献[7]提出自适应交叉变异概率, 但不能清楚反映种群中当前适应度最大的那些个体之间的趋同程度。文献[8]通过多次实验确定隐节点个数, 该方法效率比较低。文献[9]将BP网络与遗传算法相结合提出了一种混合评价算法, 但对冗余隐节点影响网络进化速度考虑不够。

针对上述问题, 笔者提出改进的进化神经网络优化设计方法。首先, 通过节点的相关性评价, 删除网络内冗余的隐节点; 其次, 将每一个网络的编码转换成规范形式, 这意味着在每一个功能等价类里, 它们都有唯一的表示形式, 从而降低编码

冗余; 最后, 在遗传操作时采用基于自适应交叉概率和变异概率相结合的自适应进化方法, 以保证种群多样性和算法收敛性。改进的网络优化设计方法用于棉花耗水量预测和番茄常见病害诊断, 验证了改进方法的有效性。

### 1 神经网络的优化设计

在用遗传算法优化神经网络时, 需要对网络进行基因编码。一个具体的权重集合, 对应着多个功能相同而编码不同的网络。如果两个功能相同的个体进行交叉, 则导致“近亲繁殖”。隐节点的个数与网络的泛化能力和参数密切相关; 自适应交叉变异概率可以较好平衡选择压力和种群多样性。

#### 1.1 改进的基因编码

用遗传算法优化设计神经网络时, “近亲繁殖”是必须考虑的问题。以双曲正切函数  $\text{th}(z) = (e^z - e^{-z}) / (e^z + e^{-z})$  作为传递函数的三层网络为例, 输入向量为  $X$ , 输出向量为  $\theta$ , 则有:

$$\theta = \text{th}(H \text{th}(MX)). \quad (1)$$

式中:  $M$  是输入层到隐层的权重矩阵;  $H$  是隐层到输出层的权重矩阵;  $\text{th}$  是对称奇函数, 用作传递函数时, 可以产生结构不同, 但输入输出相同的

收稿日期: 2010-03-14; 修订日期: 2010-06-13

基金项目: 云南省自然科学基金资助项目(2009ZC128M); 重庆师范大学自然科学基金项目(10XLB006)

作者简介: 杨华芬(1981-), 女, 云南宣威人, 曲靖师范学院讲师, 硕士, 主要从事遗传算法及神经网络方面的研究, E-mail: sunnyyh@126.com.

众多网络. 这些网络有一个共同特点: 它们使用了两个变换定义的对称有限群, 通过一系列变换, 这些群中的任何一个都可以用其它有限群来构造. 用遗传算法来优化这些网络, 交叉时就容易导致“近亲繁殖”, 而且不能从父代那儿传递较多的信息到子代, 影响进化速度. 对于  $m$  个输入,  $n$  个隐节点, 1 个输出的单隐层网络, 通过如下两种变换可以得到  $2^n n!$  个功能相同但结构不同的网络: 通过组合隐节点可以得到  $2^n$  个网络; 变换隐节点和相应的权重符号可以得到  $n!$  个网络.

设有两个网络  $n_1$  和  $n_2$ , 每一个网络有 6 个相同的节点:  $a_1 \sim a_6$ , 假设:  $n_1 = \{a_1, a_2, a_3, a_4, a_5, a_6\}$ ,  $n_2 = \{a_4, a_5, a_6, a_1, a_2, a_3\}$ , 对  $n_1$  和  $n_2$  进行单点交叉, 得到后代  $c_1$  和  $c_2$ ,  $c_1 = \{a_4, a_5, a_6, a_4, a_5, a_6\}$ ,  $c_2 = \{a_1, a_2, a_3, a_1, a_2, a_3\}$ . 可见,  $c_1$  和  $c_2$  缺乏  $n_1$  和  $n_2$  的主要计算成份, 而且还破坏了父代和子代之间的联系, 其性能比父代还差.

为此, 对基因编码进行改进就势在必然. 以  $th$  作为传递函数的某网络中, 有  $n$  个隐节点,  $r$  个输入节点,  $m_{ji}$  表示第  $i$  个输入节点到第  $j$  个隐节点之间的连接权, 用  $t_j$  表示第  $j$  个隐节点的阈值. 对于输入向量  $X = (x_1, \dots, x_n) \in \mathbf{R}^n$ , 网络的第  $j$  个隐节点的输入为:  $m_j(x) = \sum_{i=1}^n m_{ji} x_i$ , 输出  $o_j(x) = th(m_j(x)) + t_j$ . 输出节点的输入和输出分别为:

$$I(x) = \sum_{j=1}^n h_{jk} o_j(x), g(x) = th(I(x)).$$

基于上面的定义, 笔者用如下两个条件来判定两个网络是否等价: ①有一个  $o_j = 0$  ( $j = 1, \dots, n$ ); ②存在  $p_1, p_2 \in \{1, \dots, n\}$  使得  $m_{p_1}, m_{p_2}$  等价. 如果条件①成立, 隐节点  $j$  对输入  $I(x)$  没任何贡献, 因此, 将网络编码对应的基因位全部置 0. 如果②成立, 则有  $p_1, p_2$  使得  $m_{p_1}$  和  $m_{p_2}$  符号等价, 即  $m_{p_1}(x) = s m_{p_2}(x)$ ,  $s \in \{-1, 1\}$ , 此时, 节点  $p_1$  和  $p_2$  对  $I(x)$  的贡献为:

$$\begin{aligned} h_{p_1} th(m_{p_1}(x)) + h_{p_2} th(m_{p_2}(x)) &= \\ h_{p_1} th(s m_{p_2}(x)) + h_{p_2} th(m_{p_2}(x)) &= \\ s h_{p_1} th(m_{p_2}(x)) + h_{p_2} th(m_{p_2}(x)) &= \\ (s h_{p_1} + h_{p_2}) th(m_{p_2}(x)). \end{aligned}$$

因此, 可以删除节点  $p_1$ , 并将  $h_{p_2}$  改为

$$h'_{p_2} = s h_{p_1} + h_{p_2}.$$

综上所述, 通过组合隐节点和符号变换, 神经网络之间可以相互转换. 笔者采用的具体规范措施为: 如果神经元的阈值为负, 则将其修改为正, 并修改相应的权值和阈值的符号. 这样就可以

将  $2^n$  个功能相同的网络简化成具有代表性的网络; 按照隐单元的阈值升序排列, 可以排除  $n!$  个功能等价的网络. 因此, 经过上述变换, 网络的基因型和表现型一一对应: 在功能等价类里面, 每一个神经网络都有唯一基因型.

下面通过一个示例来了解规范形式的基因编码. 图 1 为 3 输入 1 输出的单隐层神经网络, 基因编码为:  $(m_{ji}, t_j, h_{jk})$ , 其中  $i, j, k$  分别为图 1 所示网络的输入节点数、隐含节点数和输出节点数.

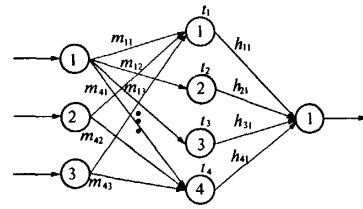


图 1 3 输入 1 输出的网络

Fig. 1 The neural network with three input-nodes and one output-node

设如图 1 所示的网络结构存在 4 个网络  $n_1, n_2, n_3, n_4$ , 其基因编码形式为:  $(m_{11}, m_{12}, m_{13}; m_{21}, m_{22}, m_{23}; m_{31}, m_{32}, m_{33}; m_{41}, m_{42}, m_{43}; t_1, t_2, t_3, t_4; h_{11}, h_{21}, h_{31}, h_{41})$ .  $n_1$  的编码为:  $(0.10, 0.20, 0.30; 0.40, 0.50, 0.60; 0.70, 0.50, 0.60; 0.34, 0.21, 0.30; 0.12, 0.23, 0.35, 0.42; 0.61, 0.62, 0.63, 0.64)$ .  $n_2$  编码为:  $(0.40, 0.50, 0.60; 0.10, 0.20, 0.30; 0.70, 0.50, 0.60; 0.34, 0.21, 0.30; 0.23, 0.12, 0.35, 0.42; 0.62, 0.61, 0.63, 0.64)$ .  $n_3$  的编码为:  $(0.40, 0.50, 0.60; 0.10, 0.20, 0.30; 0.70, 0.50, 0.60; 0.34, 0.21, 0.30; 0.23, 0.12, 0.35, 0.42; 0.62, 0.61, 0.63, 0.64)$ .  $n_4$  编码为:  $(0.40, 0.50, 0.60; 0.10, 0.20, 0.30; 0.34, 0.21, 0.30; 0.70, 0.50, 0.60; 0.23, 0.12, 0.42, 0.35; 0.62, 0.61, 0.63, 0.64)$ .

对  $n_1$  和  $n_2$  的编码进行规范以后得到完全相同的编码:  $(0.10, 0.20, 0.30; 0.40, 0.50, 0.60; 0.70, 0.50, 0.60; 0.34, 0.21, 0.30; 0.12, 0.23, 0.35, 0.42; 0.61, 0.62, 0.63, 0.64)$ . 对于  $n_3$  和  $n_4$ , 将隐层神经元的输入权、输出权和阈值看作一个整体, 根据权值进行升序排列可以得到完全相同的编码:  $(0.10, 0.20, 0.30; 0.40, 0.50, 0.60; 0.70, 0.50, 0.60; 0.34, 0.21, 0.30; 0.12, 0.23, 0.35, 0.42; 0.62, 0.61, 0.63, 0.64)$ .

由此可见, 基于改进的网络编码可以减少“近亲繁殖”, 排除网络冗余. 对于  $m$  个输入,  $n$  个

隐节点的网络,用本文的两个判定条件可以排除  $2^n n! - 2$  个功能相同而编码不同的网络,减少  $C_{(2^n n! - 2)}^2$  次“近亲繁殖”。

### 1.2 基于相关性评价的隐节点数量确定

笔者通过评价节点的相关性来确定隐节点的个数,将不相关节点的权值和阈值置0,这样并不改变网络的基因编码长度。为确定隐节点  $i$  的相关性,先计算存在该节点时网络的性能,然后删除该节点以后再计算网络的性能,对于节点  $i$  的相关性度量方法如下:

$$p_i = \frac{f_i^k - f_{-i}^k}{f_i^k} \quad (2)$$

式中:  $f_i^k$  为网络  $k$  有  $i$  节点时的性能;  $f_{-i}^k$  为网络  $k$  没有  $i$  节点时的性能。对于所有的隐节点,如果它的相关性低于某一个值  $T_r$ ,在进行交叉操作时就将其排除。这个值应该足够小,能避免删除具有微调网络作用的节点。设  $T_r = 0.001$ ,即节点的性能不能占该网络性能的  $1/1\,000$  时,就将这个节点删除。

### 1.3 改进的交叉和变异概率

交叉概率  $p_c$  和变异概率  $p_m$  等控制参数对系统性能有重要的影响。交叉概率  $p_c$  的高低决定解群体的更新和搜索速度的快慢,  $p_c$  越大越容易探测到新的优良个体,增加算法的收敛速度;变异对于保持解群体结构多样性,防止算法“早熟”是一种重要手段。  $p_m$  太大会使遗传算法变成随机搜索,从而失去其优良特性;  $p_m$  太小又难以产生新的基因块。

为此,笔者提出一个自适应交叉和变异概率。设第  $t$  代种群由  $M$  个个体  $X_1^t, X_2^t, \dots, X_M^t$  构成,适应度分别为  $F_1^t, F_2^t, \dots, F_M^t$ ,则个体的平均适应度为

$$\bar{F}_t = \frac{1}{M} \sum_{i=1}^M F_i^t \quad (3)$$

则  $p_m$  和  $p_c$  为

$$p_c = \frac{1}{1 + \exp\left(\frac{x}{M} \cdot \frac{x}{y}\right)} \quad (4)$$

$$p_m = \frac{1}{1 + \exp\left(-\frac{x}{M} \cdot \frac{x}{y}\right)} \quad (5)$$

式中:  $M$  为种群规模;  $x$  为适应度大于  $\bar{F}_t$  的个体数目;  $y$  为适应度小于  $\bar{F}_t$  的个体数目;进化过程中  $p_m$  和  $p_c$  会跟据  $\frac{x}{M} \cdot \frac{x}{y}$  取值动态进行调整。当种群个体趋于离散时,  $\frac{x}{M} \cdot \frac{x}{y}$  相应变小,  $p_c$  增大,

$p_m$  减小,种群多样性得到保持。

## 2 优化设计仿真实验

### 2.1 设计流程

利用改进的编码方案、节点相关性评价和自适应交叉变异概率,笔者提出的神经网络优化设计算法描述如下,并得到图2所示的算法流程图。

(1) 随机生成一定规模的初始种群;

(2) 基于规范形式,对网络节点进行基因编码,并排除功能相同的网络;

(3) 根据式(2)进行节点相关性评价,确定隐节点的数量,并删除冗余节点;

(4) 根据式(4)和式(5)得到自适应交叉变异概率,对种群进行交叉和变异操作;

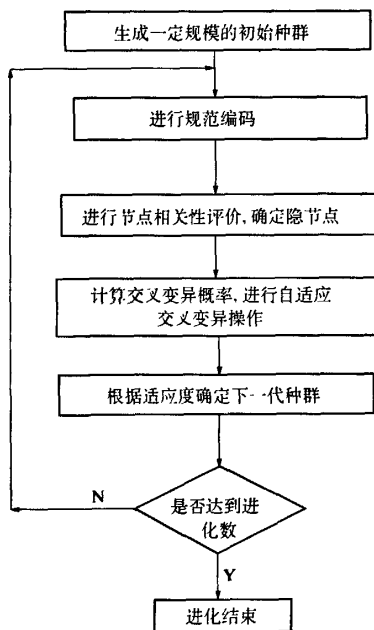


图2 算法流程图

Fig.2 The flow chart of algorithm proposed

(5) 根据个体的适应度确定下一代种群;

(6) 判断进化是否达到最大迭代次数,如果达到则停止,否则转(2)。

在该算法中,  $\frac{x}{M} \cdot \frac{x}{y}$  和文献[10]  $w'$  的计算量

相当,但是  $\frac{x}{M} \cdot \frac{x}{y}$  更能反应种群早熟程度。节点相关性评价和改进编码这两个步骤对存储空间影响不大,虽然增加了遗传操作的计算量,但是加快进化速度,而且维持了种群的多样性。在遗传操作中,保持种群多样性才能搜索到近似最优解以及

加快进化速度,否则陷入局部最优。

2.2 仿真实验

2.2.1 实验1

为验证笔者所提出的优化设计方法的有效性,将改进的神经网络用于棉花耗水量预测。笔者选取文献[11]中2005年和2006年7月上旬数据作为训练样本,随机初始100个3输入1输出的单隐层网络,进化100代。网络输入分别为平均气温、日照时数、平均风速,输出为棉花耗水量,并将训练后的网络用于文献[11]中预测2007年和2008年7月上旬的棉花耗水量,预测结果如表1所示。从表1可以看出,用本文方法预测2007年7月上旬的棉花耗水量的平均误差为0.794%。从表2可以看出,用本文方法预测2008年7月上旬的棉花耗水量的平均误差为0.747%,文献[11]预测的棉花耗水量平均误差为0.970%。由此可以看出,采用笔者方法预测棉花耗水量更接近于实际耗水量,这说明笔者所采用的优化方法较为有效。

表1 本文预测2007年的结果  
Tab.1 the predicting results got with the methods proposed in this paper for 2007

日期	棉花耗水量/(mm·d <sup>-1</sup> )		本文算法 误差/%
	实际	预测	
20070701	5.025 4	5.020 0	0.11
20070702	4.638 0	4.625 0	0.28
20070703	3.715 3	3.732 0	0.45
20070704	5.032 1	5.292 4	5.17
20070705	4.605 7	4.640 0	0.74
20070706	6.351 1	6.350 0	0.02
20070707	7.265 7	7.230 0	0.49
20070708	0.189 2	0.189 2	0.00
20070709	4.824 7	4.820 0	0.10
20070710	6.537 8	6.500 0	0.58

表2 本文算法与文献[11]算法预测2008年的结果

Tab.2 The predicting result got with the proposed method in this paper and literature [11] for 2008

日期	棉花耗水量/(mm·d <sup>-1</sup> )			本文误 差/%	文献[11] 误差/%
	实际	本文	文献[11]		
20080701	6.181 7	6.178 9	6.196 2	0.05	0.23
20080702	5.117 8	5.231 0	5.119 6	2.21	0.04
20080703	6.919 6	6.934 0	6.840 0	0.21	1.15
20080704	3.157 3	3.247 0	3.117 0	2.84	1.28
20080705	4.938 7	4.950 0	4.922 7	0.23	0.32
20080706	4.478 5	4.465 0	4.434 6	0.30	0.98
20080707	3.812 6	3.860 0	3.881 2	1.24	1.80
20080708	6.123 6	6.121 4	6.189 0	0.04	1.07
20080709	6.636 5	6.651 2	6.732 7	0.22	1.45
20080710	6.840 4	6.831 7	6.933 7	0.13	1.36

2.2.2 实验2

用笔者设计方法初始化30个7输入4输出的网络,进化100代。以文献[12]中的前8组数据作为训练样本,后8组数据作为测试样本,用本文改进的神经网络进行预测。7个输入参数为: $x_1$ 为发病部位; $x_2$ 为病斑颜色; $x_3$ 为病斑形状; $x_4$ 为霉层颜色; $x_5$ 为霉层形状; $x_6$ 为生长特征; $x_7$ 为其他特征。4个网络输出为 $y_1, y_2, y_3, y_4$ ,其输出的16种状态分别表示16种常见的病害(番茄茎基腐病、番茄白绢病、番茄斑枯病等),预测结果如表3所示。本文输出正确率为75%,文献[12]的正确率为68.75%,这说明本文的优化设计方法具有一定的通用性,虽然25%输出有误,但是对诊断病害无影响。

3 结束语

自适应交叉变异概率更好反应当前种群的多样性,用于遗传操作,一定程度上缓解“近亲繁殖”问题。对于 $m$ 个输入, $n$ 个隐节点的网络,用本文的编码方法可以排除 $2^n n! - 2$ 个功能相同而

表3 实际输出和期望输出  
Tab.3 Actual output and expected output

样本 编号	期望输出				本文算法输出				文献[12]算法输出			
	$y_1$	$y_2$	$y_3$	$y_4$	$y_1$	$y_2$	$y_3$	$y_4$	$y_1$	$y_2$	$y_3$	$y_4$
1	1	0	0	0	1.00	0.00	0.00	0.00	0.99	0.00	0.00	0.99
2	1	0	0	1	1.00	0.00	0.00	1.00	1.00	0.00	0.00	0.00
3	1	0	1	0	1.00	0.00	0.98	0.00	0.99	0.00	0.99	0.00
4	1	0	1	1	1.00	0.00	1.00	1.00	1.00	0.00	0.99	0.99
5	1	1	0	0	1.00	1.00	0.00	0.00	1.00	0.99	0.00	0.00
6	1	1	0	1	1.00	0.99	0.00	1.00	0.99	0.99	0.00	0.99
7	1	1	1	0	1.00	1.00	1.00	0.00	1.00	0.99	0.98	0.00
8	1	1	1	1	1.00	1.00	1.00	1.00	0.99	1.00	0.99	0.99

编码不同的网络,减少  $C_{2n-2}^2$  次“近亲繁殖”。节点的相关性评价可以删除网络内部的冗余节点,减少学习参数,提高学习效率。棉花耗水量和番茄病害诊断两个实验验证了笔者提出的优化设计方法的有效性和实用价值,但是初始权值和阈值影响算法的进化速度。

### 参考文献:

- [1] HU Z H. A Hybrid system based on neural network and immune co - evolutionary algorithm for garment pattern design optimization[J]. Journal of Computers, 2009,4(11):1151 - 1158.
- [2] ANAM S, SHOHIDUL I M. Face Recognition Using Genetic Algorithm and Back Propagation Neural Network[C]//The International Multiconference of Engineeris and computer Scientists IMECS 2009 Hong Kong; The International Assouation of Engineers. 2009:18 - 20.
- [3] NICOLÁS G P, DOMINGO O B, CÁsar H M. An alternative approach for neural network evolution with a genetic algorithm: crossover by combinatorial optimization[J]. Neural Networks, 2006,19(4):514 - 528.
- [4] KHARBAT F, BULL L, ODEH M. Revisiting genetic Selection in the XCS Learning Classifier System[C]. England; Sch of Comput Sci, 2005:2061 - 2068.
- [5] SERGEI L, POND K, FROST S D. A genetic algorithm approach to detecting lineage - specific variation in selection pressure[J]. Molecular Biology and Evolution, 2005,22(3):478 - 485.
- [6] KHAN A U, BANDOPADHYAYA T K. Genetic algorithm based backpropagation neural network performs better than backpropagation neural network in stock rates prediction[J]. International Journal of Computer Science and Network Security, 2008,8(7):162 - 166.
- [7] LAW N L, SZETO K Y. Adaptive Genetic Algorithm with Mutation and Crossover Matrices[C]. USA: Morgan Kaufmann Publishers Inc, 2007:2330 - 2333.
- [8] 王利平,刘志强,宗永臣. 黄河流域年降水量反算与分析[J]. 人民黄河, 2007,29(11):36 - 38.
- [9] 冯冬青,郭艳. 遗传算法改进 BP 神经网络在地下水水质评价中的应用[J]. 郑州大学学报:工学版, 2009,30(3):126 - 129.
- [10] 张晓馈,戴冠中,徐乃平. 遗传算法种群多样性的分析研究[J]. 控制理论与应用, 1998,15(1):96 - 101.
- [11] 刘婧然,马英杰. 径向基函数人工神经网络在棉花耗水量预测中的应用[J]. 新疆农业大学学报, 2009,32(1):85 - 88.
- [12] 柴毅,尹宏鹏. 基于改进遗传算法的 BP 神经网络自适应优化设计[J]. 重庆大学学报:自然科学版, 2007,30(4):91 - 96.

## An Improved Evolution Optimization of Neural Network

YANG Hua - fen<sup>1</sup>, YANG You<sup>2</sup>, SHANG Jin<sup>3</sup>

(1. Department of Computer Science and Engineering, Qujing Normal College, Qujing 655011, China; 2. School of Information Science and Engineering, Chongqing Normal University, Chongqing 401331, China; 3. Computer Engineering Department, Chongqing Aerospace Polytechnic, Chongqing 400021, China)

**Abstract:** Improved evolutionary optimization of neural network design is proposed, because there are inbreeding and coding redundancy and it is difficult to determine the number of hidden node. There is a one to one correspondence between gene coding and functional equivalence class, through normalizing of network coding, which decreases the coding redundancy. Through the evaluation of node correlations, the nodes whose correlation value is less than a certain threshold would be deleted during crossover operation so that decreasing the node redundancy. Furthermore, through the combination of crossover probability and mutation probability, the diversity of the network could be held. The experiments show that proposed approach of neural network avoids inbreeding and premature convergence aroused by inbreeding, while increasing the learning speed, and reducing the code redundancy and learning parameters.

**Key words:** neural network; genetic algorithm; gene coding; crossover probability; mutation probability; correlation between nodes