

## 基于自适应小生境递阶遗传算法的 RBF 均衡器

薛富强,葛临东,王 彬

(信息工程大学 信息工程学院,郑州 450002)

(xuewh\_zh@126.com)

**摘 要:**递阶遗传算法(HGA)一次只能确定一个最优个体。采用小生境递阶遗传算法,依据进化信息自适应调整小生境区域,在均衡数据误比特率最低,隐层中心聚类有效性最佳的基础上,可以从多个进化优解中确定出最佳结构的径向基(RBF)神经网络均衡器。仿真结果验证了算法的有效性和稳定性。

**关键词:**小生境;递阶遗传算法;RBF 神经网络;均衡器

**中图分类号:** TP183 **文献标志码:** A

## RBF equalizer based on adaptive niche hierarchy genetic algorithm

XUE Fu-qiang, GE Lin-dong, WANG Bin

(Institute of Information Engineering, Information Engineering University, Zhengzhou Henan 450002, China)

**Abstract:** It is the fact that only one optimum individual can be determined by the Hierarchy Genetic Algorithm (HGA). This paper proposed a new improved hierarchy Genetic Algorithm (GA) based on the niche, in which the range of the niche could be determined by the information of the evolution. The best structure of the Radial Basis Function (RBF) neural network equalizer could be selected from several optimal solutions with the criterion of the lowest bit error rate and the best clustering degree of the hidden centers. The simulation results demonstrate that the validity and the stability of the new algorithm.

**Key words:** niche; Hierarchy Genetic Algorithm (HGA); Radial Basis Function (RBF) neural network; equalizer

### 0 引言

线性横向均衡器(Linear Transversal Equalizer, LTE)的主要目的是达到信道的逆,从而抵消信道引入的码间干扰,但其性能在恶劣信道和非最小相位信道环境下的效果较差。对此类信道,采用分类方式会收到较好的均衡效果<sup>[1]</sup>。神经网络是一种以自组织、自适应和大规模分布式并行计算为特征的非线性信号处理系统,具有强大的模式识别和泛函逼近能力,并有较好的容错性。因此,神经网络均衡器受到了较为广泛的关注<sup>[2]50-51</sup>。前馈神经网络中,从理论上分析,径向基(Radial Basis Function, RBF)神经网络隐层基函数族是闭集,可以达到与被逼近函数距离范数的下确界,因此是最佳逼近;而多层感知器网络(Multi-Layer Perceptron, MLP)的隐层作用函数是 Sigmoid 型函数,其逼近函数族是非闭集合,所以不具最佳逼近性质<sup>[3]</sup>。从结构上分析,RBF 神经网络与贝叶斯分类器结构相类似,因此 RBF 神经网络均衡器可以获得更好的均衡效果<sup>[2]54</sup>。

RBF 神经网络均衡器的设计有两个关键环节:确定隐层结构及隐层到输出层权值。神经网络的泛化能力、学习速度与其结构密切相关,如果隐含层节点数过多,则网络的泛化能力弱;如果隐含层节点数过少,则网络有可能不收敛,因此,确定网络结构是 RBF 神经网络均衡器设计的关键。固定隐层中心数目后,再确定中心位置及其宽度的方法受人为主观影响较大,对此问题,可以采用动态 K 均值算法、资源分配网络(Resource Allocating Network, RAN)算法、进化算法(Evolutionary Algorithm, EA)等方法动态确定其结构<sup>[4]</sup>。

神经网络结构设计可以归纳为结构空间的搜索问题,给定网络结构性能最优评价标准后,其性能函数是一个无界、离散、非可微和多峰的复杂曲面,每一个搜索点表示一种构造。因此,进化算法设计神经网络结构可以收到更好的效果<sup>[5-6]</sup>。遗传算法是进化算法的典型代表,是模拟生物在自然环境中的遗传和进化过程而形成的一种自适应全局优化概率搜索算法。小生境遗传算法(Niche genetic algorithm)将进化群体中的相关个体限定在一些特定范围的环境中进化,是解决多峰值函数全局最优化问题的有效方法<sup>[7]74</sup>。

文献[8]提出基于混合递阶遗传算法(Hierarchy Genetic Algorithm, HGA)的判决反馈 RBF 信道均衡器设计方法,与固定隐层中心数目的 RBF 信道均衡器比较,其结构简单,性能得到明显改善,但缺点是只能依据适应度得到一个最优解对应的结构。针对神经网络结构多峰优化的特点,本文提出基于改进小生境的递阶遗传算法,使小生境的范围随进化过程自适应改变,在误比特率最小准则下,通过对隐层中心聚类有效性的度量,可以确定最佳结构的 RBF 神经网络均衡器。

### 1 RBF 神经网络均衡器

RBF 神经网络均衡器通常包含一个输入层、一个隐层和一个输出层。输入层节点将输入信号引入网络,网络的输入为信号矢量  $\mathbf{X}(n) = [x(n), x(n-1), \dots, x(n-m+1)]^T$ ,即畸变信号  $x(n)$  的  $m$  个延迟,输入节点数与  $m$  相同;隐层节点通过径向基函数对输入激励产生一个局部响应,将数据从低维映射到高维空间内,实现非线性变换;输出层节点利用简单的线性函数,通过加权求和得到理想的均衡输出,均衡器输出

收稿日期:2008-10-21;修回日期:2008-12-15。

基金项目:国防重点预研项目(6130320);河南省基础与前沿研究基金资助项目(082300413205)。

作者简介:薛富强(1972-),男,河南南阳人,工程师,博士研究生,主要研究方向:软件无线电、计算智能;葛临东(1946-),男,安徽怀远人,教授,博士生导师,主要研究方向:软件无线电;王彬(1971-),女,河南郑州人,讲师,博士,主要研究方向:软件无线电、信道盲辨识与盲均衡。

信号如下式所示:

$$y(n) = \sum_{i=1}^n \omega_i \varphi(\|x(n) - c_i\|, \delta) \quad (1)$$

其中  $n$  是隐层中心的个数,  $\omega_i$  表示隐层与输出层的权系数,  $\varphi(\|x(n) - c_i\|, \delta)$  表示隐层径向基函数, 常采用高斯型函数,  $x(n)$  表示输入数据,  $c_i$  表示基函数的中心,  $\|\cdot\|$  表示欧式范数,  $\delta$  表示基函数的宽度, 也称扩展常数, 它决定着高斯函数的形状, 即中心点对应基函数的作用范围。图 1 为 RBF 神经网络均衡器示意图。

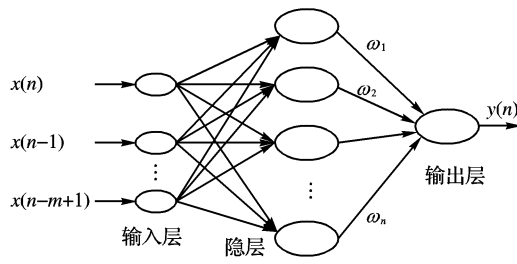


图 1 RBF 神经网络均衡器示意图

## 2 自适应小生境递阶遗传算法

由于生物倾向于与特征、形状等相类似的生物生活, 一般总是与同类繁殖后代, 加上自然地理位置的限制, 使得若干种类的生物形成了一个一个小生境。小生境的形成为新物种的形成提供了可能性, 在小生境形成初期, 小生境中物种的基因常常不同, 由于多个小生境间相对隔离缺少必要的基因交流, 使得这种基因差异得以保留, 各小生境中的生物的变异是随机发生的, 因而通常有着不同的变异方向, 这种变异的差异使物种间的基因差异不断扩大。这一差异导致了物种间基因组成产生更大的差异, 于是各物种向各自的方向发展进化, 形成“物以类聚”的局面和发展态势<sup>[9]</sup>。因此, 小生境遗传算法可以找出更多局部最优解和全局最优解<sup>[7]74</sup>。本文采用基于罚函数的小生境遗传算法, 依据群体进化信息自适应改变小生境区域, 可以减少算法计算复杂度, 提高算法的搜索能力。

### 2.1 基于罚函数的小生境遗传算法

基于罚函数的小生境遗传算法基本思想是: 首先两两比较群体中各个个体之间的距离, 若在预先指定的小生境区域  $L$  之内, 再比较两者的适应度大小, 并对其中适应度较低的个体施加一个较强的罚函数, 极大地降低其适应度。这样, 在区域  $L$  之内将只存在一个优良的个体, 从而既维护了群体的多样性, 又使得各个个体之间保持一定的距离, 能够在整个约束空间分散开来<sup>[7]75</sup>。

### 2.2 递阶遗传算法<sup>[10]</sup>

在生物学领域, 染色体的结构是一系列基因按层次排列而成的, 一些基因控制着另一些基因, 染色体可表示为包括控制基因和参数基因的递阶结构, 参数基因处于下级, 控制基因处于上级, 下级基因受上级基因的控制。控制基因主要用于控制整个网络的拓扑结构, 参数基因则用于表示网络的中心和扩展常数。在基因编码时, 控制基因常采用二进制编码: “1”表示对应的基因处于激活状态, 与该基因相联系的下级基因串处于有效状态; “0”表示对应的基因处于非激活状态, 与该基因相联系的下级基因串处于无效状态。这种定义染色体结构的遗传算法称为递阶遗传算法。递阶结构的染色体中不仅包含了参数信息而且还包含了结构信息, 而基本遗传算法中, 染色体的参数结构是预先设定的, 遗传过程中只能对其参数进行优化, 因此前者比后者包含更多的信息, 能处理更复

杂的问题。二级递阶结构的染色体结构如图 2 所示。

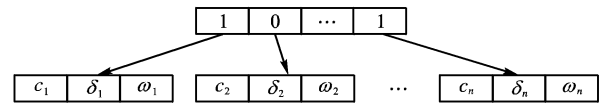


图 2 递阶染色体编码结构

图 2 中,  $c_n$  表示第  $n$  个中心的位置,  $\delta_n$  表示第  $n$  个中心的宽度,  $\omega_n$  表示第  $n$  个中心与输出层连接权。

### 2.3 自适应小生境区域

遗传算法进化初期, 个体相异度较大, 个体间的汉明距离也因此较大; 随着群体不断进化, 个体相异度相对减小, 个体间的汉明距离逐渐降低。固定小生境区域存在一些问题, 如果区域设定较大, 会增大小生境范围, 有可能抑制一些优解的进化; 如果区域设定较小, 又会增加算法的计算复杂度, 因此固定小生境区域限制了小生境群体进化的性能和效率。对于相同的输入, RBF 神经网络的隐层中心, 会因为不同的中心数目和位置产生不同的激励输出, 因此, 可以利用控制基因的平均汉明距离作为小生境区域, 使其能够自适应地随着进化呈单调下降的趋势。

### 2.4 改进小生境递阶遗传算法的设计

#### 2.4.1 编码

编码是遗传算法的一个关键部分, 它不仅决定并影响着个体基因型和表现型之间的转换, 而且也影响到交叉算子、变异算子的操作效率。常见的编码方式主要有三大类<sup>[7]33</sup>: 二进制编码、浮点数编码和符号编码。其中二进制编码的优点是编码、解码简单易行, 遗传算子操作便于实现。但二进制编码存在连续函数离散化时的映射误差, 虽然加长个体编码串可以提高求解精度, 但同时也急剧扩大了搜索空间。而浮点数编码省去了编解码环节, 改善了遗传算法的计算复杂度, 提高了算法的运算效率, 适于处理多维、高精度连续空间。

本文针对递阶遗传算法编码特点, 结合二进制编码和浮点数编码的优点, 采用各自独立的编码方式, 将隐节点中心和宽度参数分开编码, 使其在不同的搜索空间进化。其中控制基因采用二进制编码, 参数基因采用多参数级联浮点数编码, 两种形式的编码长度相同, 只与网络的初始中心数有关, 与训练数据的长度无关。

#### 2.4.2 适应度函数

RBF 神经网络的复杂度由隐层节点数决定, 精度由网络输出与期望输出之间的误差决定。进化 RBF 神经网络的目标是使网络的精确度函数和网络的复杂度函数都达到最小, 这是一个双目标优化问题, 可以在两个目标函数与染色体适应度之间建立起合适的映射关系, 将双目标优化转换为单目标优化<sup>[8]228</sup>, 适应度函数如下式所示:

$$f = 2N / \left( (a + be^{n_c/km}) \sum_{i=1}^N (d_i - y_i)^2 \right) \quad (2)$$

其中:  $a, b, k$  为待定参数,  $N$  为输入样本数,  $n_c$  为隐层节点数,  $m$  为输入维数,  $d_i$  为期望输出值,  $y_i$  为网络输出。

#### 2.4.3 遗传算子

遗传算子是遗传算法的一个重要组成部分, 不同的算子将会对算法性能产生不同的影响。控制基因和参数基因都采用轮盘赌选择方式及精英保留策略, 由于两类基因的编码方式不同, 基于二进制编码的控制基因适于采用两点交叉, 基本位变异; 而基于实数编码的参数基因则采用算术交叉, 均匀变异。

#### 2.4.4 算法具体步骤

1) 计数器置初值:  $t \leftarrow 1$ ; 随机产生  $M$  个个体组成的初始

群体  $P(t)$ , 并求出各个个体的适应度  $F_i (i = 1, 2, \dots, M)$ 。

2) 依据各个个体的适应度对其进行降序排列, 记忆前  $N$  个个体 ( $N < M$ )。

3) 控制基因和参数基因遗传操作。

4) 小生境运算。根据群体平均汉明距离确定小生境范围  $L$ , 将 3) 中得到的  $M$  个个体和 2) 中记忆的  $N$  个个体合并在一起, 对这  $M + N$  个个体组成的新群体按照下式求出每两个个体  $X_i$  和  $X_j$  之间的汉明距离  $D_{ham}$ ; 当  $D_{ham} < L$  时, 比较个体  $X_i$  和  $X_j$  的适应度大小, 并对其中适应度较低的个体处以惩罚函数, 令其适应度为 0, 如下式:

$$\min (F(x_i), F(x_j)) = 0 \quad (3)$$

5) 依据这  $M + N$  个个体的新适应度对各个个体降序排列, 记忆前  $N$  个个体。

6) 终止条件判断。若不满足终止条件, 则  $t \leftarrow t + 1$ , 并将 5) 中的前  $M$  个个体作为新的下一代群体  $P(t)$ , 然后转至 3); 否则输出计算结果, 算法结束。

## 2.5 聚类有效性分析

RBF 神经网络隐层的作用类似于将输入数据以隐层中心点为响应中心进行聚类<sup>[11]</sup>, 因此, RBF 神经网络均衡器结构进化完成后, 对隐层中心进行聚类有效性评价, 可以判断其进化优劣。

聚类有效性用来度量各聚类之间的独立性和聚类中模式的紧密性。DB (Davies-Bouldin) 指标计算快捷, 受噪声向量影响较小<sup>[12]</sup>。设  $s_i$  为聚类  $c_i$  的分散程度的测度,  $d(c_i, c_j) = d_{ij}$  表示两个聚类间的不相似性,  $R_{ij}$  表示两个聚类间的相似性, 如果  $c_i$  和  $c_j$  重叠于一点, 则  $R_{ij} = 0$ 。在  $d_{ij}$  对称的前提下, 聚类有效性  $DB_m$  由下式确定:

$$DB_m = \frac{1}{m} \sum_{i=1}^m R_i \quad (4)$$

其中:  $R_i = \max_{j=1, \dots, m, j \neq i} R_{ij}$ ,  $R_{ij} = \frac{s_i + s_j}{d_{ij}}$ ,  $m$  表示聚类中心个数。  $DB_m$  是聚类  $c_i, i = 1, \dots, m$  与其最相似性聚类之间的平均相似性,  $DB_m$  越小, 聚类的致密性和分离性就越好。

## 3 仿真与分析

本文仿真信道同文献[8], 仿真信号采用 2PAM, 训练数据 1000, 测试数据 1000, 信噪比为 15 dB。采用三种算法: 递阶遗传算法、小生境递阶遗传算法和改进小生境递阶遗传算法。递阶遗传算法参数: 种群个体 40; 最大进化代数 50; 交叉概率 0.9; 变异概率 0.01; 小生境参数  $L = 2$ ; 式(2)中参数  $a$  为 0.5,  $b$  为 0.5,  $k$  为 3。算法分别独立运行 5 次, 对应误比特率 (Bit Error Rate, BER) 及聚类有效性分别如表 1、2 所示。

表 1 显示递阶遗传算法一次进化只能得到一个最优解, 而基于小生境的递阶遗传算法不仅可以同时得到多个满意解, 而且较优个体确定的 RBF 神经网络均衡器, 能够得到更低的误比特率。采用自适应的小生境范围后, 算法的搜索性能明显有所提升, 三个算法中, 本文算法获取的 BER 最低。表 2 中, 改进小生境递阶遗传算法单次运行得到 6 个 BER 相同的个体, 其中有三个个体的隐层中心数同为 9, 但其  $DB_m$  有差异, 表明中心位置的变化对网络性能有较大的影响。在 BER 相同的情况下, 聚类有效性指标  $DB_m$  可以作为结构选取的重要参考, 按照  $DB_m$  指标越小越好原则, 选取聚类中心为 8 的个体作为最优网络结构, 既简化了 RBF 神经网络均衡器结构,

又实现了最佳均衡。

表 1 仿真结果

算法类型	最低 BER	平均最优个体数 (对应 BER ≤ 0.01 的个体)
递阶遗传算法	0.008	1
小生境递阶遗传算法	0.003	11
自适应小生境递阶遗传算法	0.001	13

表 2 改进小生境递阶遗传算法单次运行结果

BER	隐层中心数	训练数据 $DB_m$	测试数据 $DB_m$
0.001	9	1 193	1 147
0.001	9	1 037	1 022
0.001	9	1 036	1 021
0.001	8	633	619
0.001	10	997	980
0.001	7	1 423	1 559

## 4 结语

神经网络结构最优设计属于多峰函数优化问题。递阶遗传算法虽然能够从进化角度动态确定 RBF 神经网络隐层中心的数目和位置, 但算法基于简单遗传算法, 最终只能搜寻到解空间中的一个最优解。小生境递阶遗传算法适合求解此类问题, 可以同时搜索到多个较优解, 但需要先确定合适的小生境范围。本文算法依据群体进化信息, 将平均汉明距离作为小生境范围, 使其随进化过程自适应改变, 有效地提高了算法的搜索性能。在误比特率同样最小的情况下, 聚类有效性指标  $DB_m$  是度量隐层中心独立性和聚类紧密性的重要参考,  $DB_m$  值最小的个体对应 RBF 神经网络均衡器的最优结构。仿真结果验证了本文算法的有效性和稳定性。

## 参考文献:

- [1] GIBSON G J, SIU S, COWAN C F N. The application of nonlinear structures to the reconstruction of binary signals[J]. IEEE Transactions on Signal Processing, 1991, 39(8): 1877-1884.
- [2] MULGREW B. Applying radial basis functions[J]. IEEE Signal Processing Magazine, 1996, 13(2): 50-65.
- [3] 柴杰, 江青茵, 曹志凯. RBF 神经网络的函数逼近能力及其算法[J]. 模式识别与人工智能, 2002, 15(3): 310-316.
- [4] 魏海坤. 神经网络结构设计的理论与方法[M]. 北京: 国防工业出版社, 2005.
- [5] 李智勇, 童调生. 基于多种群进化小生境遗传算法的神经网络进化设计方法研究[J]. 控制与决策, 2003, 18(5): 607-610.
- [6] XIN YAO. Evolving artificial neural networks[J]. Proceedings of the IEEE, 1999, 87(9): 1423-1447.
- [7] 周明, 孙树栋. 遗传算法原理及应用[M]. 北京: 国防工业出版社, 1999.
- [8] 王俊年, 申群太, 周少武. 基于混合递阶遗传算法的判决反馈 RBF 网络信道均衡器设计[J]. 信号处理, 2006, 22(2): 226-229.
- [9] 郑宣耀, 王芳. 一种改进的小生境遗传算法[J]. 重庆邮电学院学报: 自然科学版, 2005, 17(6): 721-723.
- [10] 李方方, 赵英凯, 许必熙. 基于递阶遗传算法优化神经网络的研究[J]. 机械与电子, 2006(2): 41-44.
- [11] PANDYA A S. 神经网络模式识别及其实现[M]. 徐勇, 荆涛, 译. 北京: 电子工业出版社, 1999.
- [12] THEODORIDIS S, KOUTROUMBAS K. 模式识别[M]. 李晶皎, 王爱侠, 张广渊, 等译. 3 版. 北京: 电子工业出版社, 2006.