

# 基于遗传小波神经网络的语音识别分类器设计

韩志艳 王 健 伦淑娴

(渤海大学信息科学与工程学院 锦州 121000)

**摘 要** 分类在语音识别中是很重要的,由于小波神经网络的学习效果对网络隐层节点数、初始权值(包括阈值)、伸缩和平移因子以及学习率和动量因子的依赖性较大,致使其全局搜索能力弱,易陷入局部极小,收敛速度减慢,甚至不收敛。而遗传算法具有的高度并行、随机、自适应搜索性能,使它在处理用传统搜索方法解决不了的复杂和非线性问题时具有明显的优势。因此,考虑把遗传算法和神经网络相结合,采用遗传算法选取初值进行训练,用小波神经网络完成给定精度的学习。仿真实验结果表明,该模型有效地提高了语音的识别率,并缩短了识别时间,实现了效率与时间的双赢,为算法的实用性奠定了基础。

**关键词** 语音识别,神经网络,遗传算法,小波分析

**中图分类号** TP391.9 **文献标识码** A

## Design of Speech Recognition Classifier Based on Genetic Wavelet Neural Network

HAN Zhi-yan WANG Jian LUN Shu-xian

(School of Information Science & Engineering, Bohai University, Jinzhou 121000, China)

**Abstract** Classification is an important problem in speech recognition, due to the fact that the learning effects of wavelet neural network strongly depend on the number of hidden nodes, the initial weights (including thresholds), the scale and displacement factors, the learning rate and momentum factor, which leads to weak global search capability, easily falling into local minimum values, low convergence rate, and even not convergent. Genetic Algorithm (GA) has height parallel performance, random and adaptive search performance, and it has obvious advantages in solving complex and nonlinear problem. Therefore, we can combine neural network and genetic algorithm by using GA to select initial value, and use wavelet neural network to finish the learning. The simulation results show that the new model effectively improves speech recognition rate, shortens the recognition time, realizes double wins in efficient and time, establishes the foundation for practicality of the algorithm.

**Keywords** Speech recognition, Neural network, Genetic algorithm, Wavelet transform

20 世纪 60 年代人类便开始了语音识别技术的研究,其目的就是使机器能够理解人类的自然语言。由于语音信号本身非常复杂,加上人类对语音学知识还很缺乏,因此目前的任何一种语音识别方法在识别率和实时反应速度上都还很难达到人类本身的听觉效果<sup>[1]</sup>。人工神经网络方法<sup>[2]</sup>自 80 年代起开始应用于语音识别,为该课题注入了新的活力。主要是因为神经网络具有以下一些特点:(1)具有自学习的功能;(2)具有泛化能力;(3)具有非线性映射功能;(4)具有高度并行处理功能。然而,目前神经网络方法受到各种条件的限制,仍无法充分发挥它的优势。为了改善这一状态,本文提出了一种新的神经网络——遗传小波神经网络,它具有响应速度快、识别精度高等优点。采用人工神经网络技术进行语音识别的过程包含两部分:一是网络学习过程,二是语音识别过程。网络学习过程是将已知语音信号作为学习样本,通过神经网络的自学习,最终得到一组连接权和偏置;语音识别过程是将待测语音信号作为网络输入,通过网络的联想得出识别结果。

## 1 小波神经网络的设计和学习算法

小波神经网络是以小波基函数为神经元激励函数的前馈神经网络模型。它作为一种前向神经网络,与基于 Sigmoid 函数的 BP 网络和径向基(RBF)网络相比,具有结构可设计性、收敛精度可控性和收敛速度快等优点<sup>[3-6]</sup>。小波神经网络一般采用单隐层结构,不影响其逼近能力。基于信号表示的小波神经网络输出层激励函数一般为线性函数,这有助于理解信号小波表示的物理意义。但在本文中采用 Sigmoid 函数代替线性函数,从而使网络具有更强的逼近能力。本文中小波神经网络采用 3 层结构,其具体的设计步骤和学习算法如下。

### 1.1 产生小波基函数

将小波母函数  $\Psi(t)$  进行伸缩和平移,得到一组连续小波函数基,即

$$\Psi_{a,b}(t) = \frac{1}{\sqrt{|a|}} \Psi\left(\frac{t-b}{a}\right) \quad a, b \in R; a \neq 0 \quad (1)$$

到稿日期:2009-12-08 返修日期:2010-02-25 本文受国家自然科学基金(60974071),辽宁省教育厅重点实验室项目(2009S002)资助。

韩志艳(1982-),女,博士,讲师,CCF 会员,主要研究方向为语音信号处理,E-mail:hanzyme@126.com;王 健(1982-),男,硕士,主要研究方向为信号处理;伦淑娴(1972-),女,教授,主要研究方向为非线性系统的控制、智能控制、信号处理等。

式中,  $a, b$  分别为伸缩、平移因子。本文中采用 morlet 小波作为母小波, 即

$$\Psi(t) = \cos(rt) \exp(-t^2/2) \quad (2)$$

式中,  $r$  取为 1.75。

## 1.2 设计小波神经网络模型

该 3 层小波神经网络包括一个输入层、一个隐含层和一个输出层, 分别由  $L, M, S$  个神经元组成;  $w_{ij}$  是输入层第  $i$  个神经元与隐含层第  $j$  个神经元之间的连接权,  $v_{jt}$  是隐含层第  $j$  个神经元与输出层第  $t$  个神经元之间的连接权; 若约定  $w_{0j}$  是第  $j$  个隐层节点阈值,  $v_{0t}$  是第  $t$  个输出层节点阈值(相应输入为 -1); 第  $n$  个样本的输入为  $v, i=1, 2, \dots, L, n=1, 2, \dots, N$ , 其中  $N$  为样本总数, 网络的输出为  $y^n$ , 其中  $t=1, 2, \dots, S$ , 对应的目标输出为  $d_t^n, t=1, 2, \dots, S, n=1, 2, \dots, N$ ;  $a_j$  和  $b_j$  分别为第  $j$  个隐层节点的伸缩平移系数; 输入层激励函数为线性变换(输出=输入), 隐含层激励函数为小波函数, 输出层激励函数为 Sigmoid 函数, 则小波网络模型可表示为

$$y_t^n = f \left[ \sum_{j=0}^M v_{jt} \Psi \left[ \frac{\sum_{i=0}^L w_{ij} x_i^n - b_j}{a_j} \right] \right] \quad (3)$$

## 1.3 定义误差能量函数

实际输出与期望输出之间的误差能量函数为

$$E = \frac{1}{2} \sum_{t=1}^S (y_t^n - d_t^n)^2 \quad (4)$$

## 1.4 计算误差的偏导数

若要求出满足误差能量函数  $E$  最小的网络参数, 则要求出在第  $n$  个样本时  $E$  对每个网络参数的偏导数。

$$\frac{\partial E}{\partial v_{jt}} = (y_t^n - d_t^n) f(z) (1 - f(z)) \Psi(T) \quad (5)$$

$$\frac{\partial E}{\partial w_{ij}} = \sum_{t=1}^S (y_t^n - d_t^n) f(z) (1 - f(z)) v_{jt} \frac{\partial \Psi(T) x_i^n}{\partial T a_j} \quad (6)$$

$$\frac{\partial E}{\partial a_j} = \sum_{t=1}^S (y_t^n - d_t^n) f(z) (1 - f(z)) v_{jt} \frac{\partial \Psi(T) (-T)}{\partial T a_j} \quad (7)$$

$$\frac{\partial E}{\partial b_j} = \sum_{t=1}^S (y_t^n - d_t^n) f(z) (1 - f(z)) v_{jt} \frac{\partial \Psi(T) (-1)}{\partial T a_j} \quad (8)$$

式中,  $T$  为隐含层神经元的输入值,  $z$  为输出层神经元的输入值。

$$T = \frac{\sum_{i=0}^L w_{ij} x_i^n - b_j}{a_j} \quad (9)$$

$$z = \sum_{j=0}^M v_{jt} \Psi \left[ \frac{\sum_{i=0}^L w_{ij} x_i^n - b_j}{a_j} \right] \quad (10)$$

隐层小波基函数采用 morlet 小波函数:

$$\Psi(t) = \cos(rt) \exp(-t^2/2) \quad (11)$$

$$\frac{\partial \Psi(t)}{\partial t} = -r \sin rt \exp(-t^2/2) - t \cos rt \exp(-t^2/2) \quad (12)$$

输出层神经元激励函数为

$$f(t) = \frac{1}{1 + \exp(-t)} \quad (13)$$

$$\frac{\partial f(t)}{\partial t} = \frac{\exp(-t)}{(1 + \exp(-t))^2} = f(t)(1 - f(t)) \quad (14)$$

## 1.5 网络参数调整

参考常规的 BP 算法引入动量系数  $\mu$ , 网络各参数按下面公式进行调整。

$$v_{jt}(N+1) = v_{jt}(N) - \eta \frac{\partial E}{\partial v_{jt}} + \mu \Delta v_{jt}(N) \quad (15)$$

$$w_{ij}(N+1) = w_{ij}(N) - \eta \frac{\partial E}{\partial w_{ij}} + \mu \Delta w_{ij}(N) \quad (16)$$

$$a_j(N+1) = a_j(N) - \eta \frac{\partial E}{\partial a_j} + \mu \Delta a_j(N) \quad (17)$$

$$b_j(N+1) = b_j(N) - \eta \frac{\partial E}{\partial b_j} + \mu \Delta b_j(N) \quad (18)$$

式中,  $\eta$  为学习速率。只要  $\eta$  足够小, 就可以实现真正的梯度下降, 但容易陷入局部极小点; 如果  $\eta$  取得太大, BP 算法又可能产生震荡而不收敛, 所以  $\eta$  的选取很重要。  $\mu$  为动态参量,  $\mu$  使每一步迭代权值的变化不仅能反映梯度变化, 而且能够跟踪误差曲面的变化趋势。  $\Delta v_{jt}(N)$  为上次调整后的权值差值,  $\Delta v_{jt}(N+1)$  为本次调整后的权值差值, 式(16)~式(18)中同理。

## 2 使用遗传算法优化小波神经网络

遗传算法(Genetic Algorithm)是模拟生物界自然选择和自然遗传机制进化过程来求解复杂问题的全局随机搜索算法, 它以编码空间代替问题空间, 以适应度函数为评价依据, 以编码群体为进化基础, 以对群体中个体位串的遗传操作实现选择和遗传机制, 建立起一个迭代过程。在这一过程中, 通过随机重组编码位串中重要的基因, 使新一代的位串集合优于老一代的位串集合, 群体的个体不断进化, 逐渐接近最优解, 最终达到求解问题的目的<sup>[7]</sup>。具体的遗传算法描述如下。

1) 初始运行参数, 包括群体大小、交叉概率、变异概率和最大代数等。

2) 通过某种编码方法(如二进制编码、实值编码)随机产生一组初始个体, 构成初始种群, 其中每个初始个体代表了问题的一个初始解。

3) 计算群体个体的适应度。

4) 判断算法终止条件是否满足。若满足则结束, 否则执行下面步骤。

5) 根据适应度大小, 通过相应的选择算法进行选择操作。

6) 以概率  $P_c$  执行交叉操作。

7) 以概率  $P_m$  执行变异操作。

8) 保留适应度大的个体形成下一代群体。

由于小波神经网络的学习效果对网络隐层节点数、初始权值(包括阈值)、伸缩和平移因子以及学习率和动量因子的依赖性较大, 致使其易陷入局部极小, 收敛速度减慢, 甚至不收敛, 这些初始参数的取值对 ANN 的构建起到重要作用。在对这些参数无先验知识的情况下需要反复地修改各个参数值才能使 ANN 具备优良的分类性能。这种人工反复修改参数的过程其实就是一种寻优的过程。但是由于神经网络的参数多, 每个参数的取值范围分布较广, 手动调整参数的方法难以建立优良的网络分类器。而遗传算法具有高度并行、随机、自适应搜索性能, 使它在处理用传统搜索方法解决不了的复杂和非线性问题时, 具有明显的优势<sup>[8,9]</sup>。因此, 我们考虑把遗传算法和神经网络相结合, 采用遗传算法选取初值进行训练, 用小波神经网络完成给定精度的学习。

### 2.1 基本思想

首先用遗传算法对网络隐层节点数、初始权值(包括阈值)、伸缩和平移因子以及学习率和动量因子进行优化设计, 在解空间中定位出较好的搜索空间, 然后用小波神经网络算法在这些小的解空间中对网络的隐层节点数、初始权值(包括

阈值)、伸缩和平移因子以及学习率和动量因子再次寻优,搜索出最优解。

## 2.2 算法实现的关键技术

### (1) 编码与解码方案

在传统的遗传算法中采用的是二进制编码。二进制编码在处理自变量较多的问题时,会导致染色体长度偏长、搜索空间大、搜索效率低等问题。另外在二进制数转化成十进制数时,存在量化误差,影响了搜索精度。针对本次研究问题的自变量数目较多,且要与神经网络算法相结合的特点,采用实值编码。这样可以有效地提高运算速度和精度,且避免了编码中带来的附加问题,更便于和神经网络算法相结合。

将所要构建的神经网络的所有隐层节点数、初始权值(包括阈值)、伸缩和平移因子以及学习率和动量因子作为一组染色体,采用实值编码方法分别对它们进行编码。每一个个体编码是由7部分组成的:第一部分是隐层节点数  $M$  对应的实值编码;第二部分是输入层节点到隐层节点的初始权值(包括阈值)  $w_{ij}$  对应的实值编码;第三部分是隐含层节点到输出层节点的初始权值(包括阈值)  $v_k$  对应的实值编码;第四部分是伸缩因子  $a_j$  对应的实值编码;第五部分是平移因子  $b_j$  对应的实值编码;第六部分是学习率  $\eta$  对应的实值编码;第七部分是动量因子  $\mu$  对应的实值编码,如图1所示。

$M$	$w_{ij}$	$v_k$	$a_j$	$b_j$	$\eta$	$\mu$
-----	----------	-------	-------	-------	--------	-------

图1 个体编码的组成

### (2) 适应度函数的选择

衡量网络性能的主要指标是网络的实际输出值与期望输出值之间的误差平方和。该误差平方和小,则表示该网络性能好。

定义适应度函数<sup>[10]</sup>为

$$f = \frac{1}{R+1} \quad (19)$$

式中,  $R = \frac{1}{N} \sum_{n=1}^N \sum_{t=1}^S (y_n^t - d_n^t)^2$ ,  $N$  为训练集样本总数,  $S$  为输出层神经元总数,  $y_n^t$  是第  $n$  个样本输出层单元  $t$  的实际输出,  $d_n^t$  是第  $n$  个样本输出层单元  $t$  的理想输出。

### (3) 遗传操作

选择算子:选择就是从当前种群中选择优胜(即适应度高)个体而淘汰劣质个体的操作。目前常用的选择算子有以下几种:适应度比例选择法、最佳个体保存法、期望值法、排序选择法及联赛选择法等。本文采用适应度比例选择法来进行选择操作。适应度比例选择法是目前遗传算法中最基础也是最常用的选择方法,也叫轮盘赌选择法。在该方法中,每个个体的选择概率和它的适应度值成正比,其中个体被选择的概率  $P_s$  可以用下式来确定

$$P_s = \frac{f[i]}{\sum_{i=1}^{pop\_size} f[i]} = \frac{f[i]}{sumf} \quad (20)$$

式中,  $pop\_size$  为种群规模,  $f[i]$  为个体  $i$  的适应度值,  $sumf$  为当前代种群的适应度和。可通过下述步骤实现轮盘赌选择:首先,计算当代种群的适应度总和,然后产生一个(0,1)之间的随机数  $r_0$ ,取  $r = r_0 * sumf$ ,并把满足下式

$$\sum_{i=0}^{j-1} f[i] < r \text{ 且 } \sum_{i=0}^j f[i] \geq r \quad (21)$$

的第  $j$  个个体选中,进入交叉操作。重复  $pop\_size$  次这样的

操作,最终得到  $pop\_size$  个个体,一起进入交叉操作。

交叉算子:在遗传算法中,交叉操作是最主要的遗传操作。根据指定的参数  $P_c$  作为交叉的概率,这个概率说明种群中有期望值为  $P_c \times pop\_size$  个染色体进行交叉操作。为确定交叉操作的父代,从  $i=1$  到  $pop\_size$  重复以下过程:从区间  $[0,1]$  中产生一个一致分布的随机数  $r$ ,如果  $r < P_c$ ,则选择  $v_i$  作为一个父代。如果用  $v_1', v_2', v_3', \dots$  表示上面选择的父代,并把它们随机分成下面的对  $(v_1', v_2'), (v_3', v_4'), \dots$  进行交叉,以  $(v_1', v_2')$  为例进行的交叉操作解释如下。

$$\begin{cases} x = c \times v_1' + (1-c) \times v_2' \\ y = (1-c) \times v_1' + c \times v_2' \end{cases} \quad (22)$$

式中,  $x$  和  $y$  为后代,  $x$  和  $y$  的替换对象分别为两个父代  $v_1'$  和  $v_2'$ ,  $c$  为(0,1)之间的随机数。

变异算子:设定变异概率  $P_m$ ,这个概率表明总体中有  $P_m \times pop\_size$  个染色体(期望值)用来进行变异操作。类似于交叉过程中父代的选择,从  $i=1$  到  $pop\_size$  重复以下过程:从区间  $[0,1]$  中产生一个随机数  $r$ ,如果  $r < P_m$ ,则选择  $v_i$  作为一个父代。如果用  $v_1', v_2', v_3', \dots$  表示上面选择的父代,以  $v_1'$  为例的变异操作如下。

$$z = v_1' + c \quad (23)$$

式中,  $v_1'$  为父代个体的变异基因,  $z$  是变异后相应的子代个体的基因。

为了防止变异操作产生空类,先为被选择变异的个体作备份,变异操作完成后,如果检查出不合格的个体,则用副本还原原来的个体,否则选用变异后的新个体代替旧个体。

### (4) 终止条件

采用给定的迭代次数  $g\_max$  和适应度来控制遗传算法的终止。如果迭代次数已经用完或某代的最优个体的适应度值大于等于预先给定的值,算法就结束。

## 2.3 算法步骤与流程

基于遗传算法优化的神经网络学习算法,具体步骤如下:

1) 随机产生一组实值串种群,每一个个体表示网络隐层节点数、初始权值(包括阈值)、伸缩和平移因子以及学习率和动量因子的一个集合。

2) 对实值串中的隐层节点数部分进行解码,生成相应的网络结构。

3) 对实值串中的其余6部分分别进行解码,生成网络的初始权值(包括阈值)、伸缩和平移因子以及学习率和动量因子。

4) 正向运行网络,根据式(19)计算适应度值,评价网络性能。

5) 通过选择、交叉和变异等遗传操作,产生下一代种群,形成下一代网络。

6) 重复步骤2)一步骤5),直到  $R \leq R_{max}$  或达到进化代数  $g\_max$ 。此时,将最终种群中的个体进行解码,从而得到通过GA优化后的值。其中  $R_{max}$  为遗传算法所要达到的性能指标。

7) 将GA优化后的网络隐层节点数、初始权值(包括阈值)、伸缩和平移因子以及学习率和动量因子作为小波神经网络使用的参数。

8) 双向运行网络,按照式(15)一式(18)调节网络权值(包

括阈值)和伸缩平移因子,评价网络性能。

9)重复步骤8),直到  $R \leq R_{\min}$ 。其中  $R_{\min}$  为小波神经网络算法所要达到的网络性能指标,即最终要求的性能指标。保存网络隐层节点数、初始权值(包括阈值)、伸缩和平移因子以及学习率和动量因子,学习过程结束。

### 3 网络性能分析与评估

为了验证遗传小波神经网络算法的优越性,我们对算法的性能进行了测试,并与 BP 算法和小波神经网络算法进行了比较。

语音数据由五男五女的发音(0到9共10个数字的普通话发音)数据组成,每个音每人用普通话发10遍,共计1000次发音。其中以每人每个发音的前5次作为训练样本,后5次作为测试样本。录音环境为安静的实验室,采样频率为11.025kHz。我们对每帧语音信号分别提取12维的Mel倒谱系数(MFCC)作为特征向量,而且每个语音信号只选择中间的6帧来进行实验,这样就大大减少了运算量。

神经网络的输入层有72个神经元,输出层有10个神经元,经遗传算法优化的隐含层节点数为17个。在相同条件下,分别用BP算法、小波神经网络算法及遗传小波神经网络算法对网络进行训练。网络训练时,各参数取值为  $P_c = 0.6$ ,  $P_m = 0.1$ ,  $pop\_size = 70$ ,  $g\_max = 500$ ,  $R_{max} = 0.01$ ,  $R_{min} = 0.001$ , 经遗传算法优化的学习率  $\eta = 0.955$ , 动量因子  $\mu = 0.300$ 。各算法的迭代次数和均方误差的关系曲线如图2—图4所示,其中图2为BP算法误差学习曲线,图3为小波神经网络算法误差学习曲线,图4为遗传小波神经网络算法的误差学习曲线。

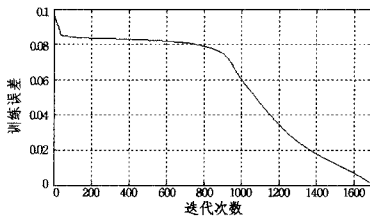


图2 BP算法误差学习曲线

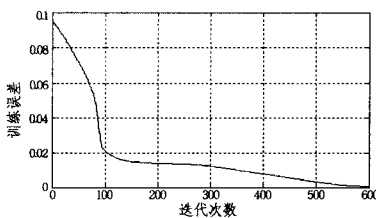


图3 小波神经网络算法误差学习曲线

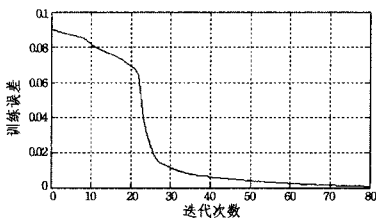


图4 遗传小波神经网络算法误差学习曲线

从图2—图4可以看到,基于遗传小波神经网络的算法只迭代了76次便使误差值为0.001,而小波神经网络算法和BP算法要分别迭代573次和1695次。由此可知,无论在收敛速度上,还是在迭代相同次数所能达到的精度上,基于遗传小波神经网络的算法都比现有的基本BP算法和小波神经网络算法取得了更好的效果。基于遗传小波神经网络的算法避免了局部极小,从而快速地找到最优解,降低了学习时间。

**结束语** 利用遗传算法强大的全局搜索能力,对小波神经网络的隐层节点数、初始权值(包括阈值)、伸缩和平移因子以及学习率和动量因子进行全面优化,在解空间中定位出较好的搜索空间,然后用小波神经网络算法在这些小的解空间中搜索出最优解,形成了一种遗传小波神经网络学习算法。实例表明了该算法的收敛速度和精度均优于BP算法和小波神经网络算法,它是可行的、有效的。这是由于遗传算法能以较快的速度减少搜索空间范围,而且不易陷入局部极小点,有很强的全局搜索能力,而小波神经网络则具有局部搜索效率高的特点,将两者结合起来便可提高算法的整体性能。

### 参考文献

- [1] Juang B H, Tsuhan C. The Past, Present, and Future of Speech Processing[J]. IEEE Singal Processing Magazine, 1998, 15(3): 24-48
- [2] Ryeu J K, Chung H S. Chaotic Recurrent Neural Networks and Their Application to Speech Recognition[J]. 1996, 13(2-4): 281-294
- [3] 林遂芳,潘永湘,孙旭霞. 基于HMM和小波网络模型的抗噪声语音识别方法[J]. 系统仿真学报, 2005, 17(7): 1720-1723
- [4] Zhang Q H, Benveniste A. Wavelet Networks [J]. IEEE Transactions on Speech and Neural Networks, 1992, 3(2): 889-898
- [5] Yoshihiro Y, Nikiforuk P N. A New Supervised Learning Algorithm for Multilayered and Inter-connected Neural Networks [J]. IEEE Transactions on Speech and Neural Networks, 2000, 11(1): 36-46
- [6] Zhang J, Walter G. Wavelet Neural Networks for Function Learning[J]. IEEE Transactions on Signal Processing, 1995, 43(6): 1485-1497
- [7] Pan S T, Wu C H, Lai C C. The application of Improved Genetic Algorithm on the Training of Neural Network for Speech Recognition[C]// The Second International Conference on Innovative Computing, Information and Control (ICIC '2007). 2007: 168-168
- [8] Slowik A, Bialko M. Training of Artificial Neural Networks Using Differential Evolution Algorithm[C]// The International Conference on Human System Interactions, 2008: 60-65
- [9] 李影,徐涛,邢伟. 基于进化遗传算法的神经网络优化[J]. 长春理工大学学报, 2006, 29(3): 48-50
- [10] Chen M, Yao Z M. Classification Techniques of Neural Networks Using Improved Genetic Algorithms[C]// The Second International Conference on Genetic and Evolutionary Computing, 2008: 115-119