

基于多种群协同进化微粒群算法的径向基神经网络设计

王俊年^{1,2}, 申群太¹, 沈洪远^{1,2}, 周鲜成¹

(1. 中南大学 信息科学与工程学院, 湖南 长沙 410083; 2. 湖南科技大学 信息与电气工程学院, 湖南 湘潭 411201)

摘要: 神经网络结构和权值的联合设计一直是神经网络进化设计的一个研究方向. 本文根据基本微粒群算法的特点, 借鉴递阶编码的思想, 构造出一种多种群协同进化微粒群算法. 该算法具有种群内个体微粒自由运动特征分量与种群运动特征分量分层递阶进化的特征, 克服了标准微粒群算法在多峰函数寻优时出现的微粒“早熟”现象. 应用该算法进行径向基神经网络隐层结构和径向基函数参数联合自适应设计, 在非线性系统辨识中显示了比较好的收敛性和训练精度, 同时也使网络的泛化能力和逼近精度这一对矛盾得到了比较好的协调统一.

关键词: 微粒群算法; 多种群协同进化; 径向基神经网络; 结构优化

中图分类号: TP183 **文献标识码:** A

Evolutionary design of RBF neural network based on multi-species cooperative particle swarm optimizer

WANG Jun-nian^{1,2}, SHEN Qun-tai¹, SHEN Hong-yuan^{1,2}, ZHOU Xian-cheng¹

(1. Institute of Information Science and Engineering, Central South University, Changsha Hunan 410083, China;

2. Institute of Information and Electrical Engineering, Hunan University of Science and Technology, Xiangtan Hunan 411201, China)

Abstract: Combination design of neural network's structure and weights has been one of the research focuses in neural network's evolutionary design. In this paper, a multi-species cooperative particle swarm optimizer is proposed by combining the ideas in the standard particle swarm optimization and hierarchy method. In the new algorithm, the individual free movement of particles within the species and the species population's movement evolve in a hierarchy model. The developed algorithm overcomes the limitation of particle's "prematurity" in global optimization using the standard PSO. When this algorithm is used in the training of RBF neural network's structure and parameters, the neural network shows a satisfactory accuracy and convergence in nonlinear system identification. The resulting network is able to properly balance the relation between generation and approximation accuracy.

Key words: particle swarm optimization; multi-species cooperative; RBF neural network; structure optimization

1 引言 (Introduction)

微粒群算法 (particle swarm optimizer, PSO) 是模拟鱼、鸟等群体的社会行为的一种优化算法^[1-4], 由于算法概念简单、实现容易, 在函数优化和神经网络权值训练等方面都有很好的表现. 但由于微粒之间的共享信息是所有微粒经历过的最好位置, 亦即当前的全局最优解, 因此这种算法在求解多峰函数优化问题时出现了未成熟收敛和收敛速度慢等缺点^[5-8].

进化计算中协同进化 (co-evolution) 意指多个种群通过适应度的关联同时进化, 是在遗传算法的基础上借鉴了进化生态学的研究成果而提出的遗传

进化方法, 最早由 Rosin 等提出^[9], 主要是为了解决遗传算法应用中出现的未成熟收敛等缺陷. F. Van den Bergh 和 A. P. Engelbrecht^[10] 将这一概念引入微粒群算法, 构造出一种协同进化微粒群算法 (cooperative particle swarm optimization, CPSO), 在多峰函数优化的仿真实验中得到了比标准 PSO 满意的结果. 但 CPSO 算法中子微粒群是随机划分的, 没有充分考虑到子微粒群中个体微粒之间的共性和个体微粒与子微粒群之间的属性关系, 因此不同的子微粒群划分对算法性能影响很大. 事实上, 自然界中的鱼鸟等群体的运动, 除个体运动特征不同外, 同一种群内个体的运动有其相同的地方. 基

于这种考虑,借鉴递阶进化思想和生态学的种群划分思想,本文设计出了一种多种群协同进化微粒群算法(multi-species cooperative PSO, MCPSO),在径向神经网络结构和径向基函数进化设计中显示了良好的多峰函数寻优能力和较快的收敛速度,使网络的泛化能力和逼近精度这一对矛盾体得到了比较好的协调统一。

2 基本 PSO 算法 (Basic PSO)

基本 PSO 算法是将每个个体看作是 D 维搜索空间中的一个没有体积的微粒(点),在搜索空间中以一定的速度飞行,这个速度根据它本身的飞行经验以及同伴的飞行经验进行动态调整,直至群体中的个体飞行到对环境适应度高的区域。

假设在 n 维目标搜索空间中有 s 个微粒组成一个微粒群,其中第 i 个微粒表示为其在 n 维搜索空间中的位置向量 $x_i = \{x_{i1}, x_{i2}, \dots, x_{in}\}, i = 1, 2, \dots, s$, 将 X_i 代入一个目标函数就可以计算出其评价价值。该微粒运动的速度也是一个 n 维向量,记为 $v_i = \{v_{i1}, v_{i2}, \dots, v_{in}\}$ 。它在飞行过程中迄今为止经历过的最好位置记为 $y_i = \{y_{i1}, y_{i2}, \dots, y_{in}\}$, 整个微粒群迄今为止经历过的最好位置记为 $\hat{y} = \{\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n\}$, 则标准版本的 PSO 算法在每一代对第 i 个微粒在第 j 维 ($1 \leq j \leq n$) 上飞行速度更新公式如下^[4]:

$$v_{i,j}(t+1) = \omega v_{i,j}(t) + c_1 r_{1i} [y_{i,j}(t) - x_{i,j}(t)] + c_2 r_{2,i} [\omega_j(t) - x_{i,j}(t)]. \quad (1)$$

其中: $i = 1, 2, \dots, s, j = 1, 2, \dots, n; v_{i,j}$ 为第 i 个微粒在第 j 维搜索空间上的飞行速度,它被限定在 $[v_{\min}, v_{\max}]$ 之间,权重因子 $\omega \in (0, 1)$, 加速常数 c_1, c_2 为非负常数; $r_1(t), r_2(t)$ 为取值在 $(0, 1)$ 之间的两个随机序列。相应微粒的位置更新公式为

$$x_i(t+1) = x_i(t) + v_i(t+1). \quad (2)$$

每个微粒经历过的最好位置更新公式为

$$y_i(t+1) = \begin{cases} y_i(t), & \text{如果 } f(x_i(t+1)) \geq f(y_i(t)), \\ x_i(t+1), & \text{如果 } f(x_i(t+1)) < f(y_i(t)). \end{cases} \quad (3)$$

整个微粒群迄今为止经历过的最好位置定义为

$$\hat{y}_i(t+1) = \arg \min_{y_i} f(y_i(t+1)), 1 \leq i \leq s. \quad (4)$$

由式(1)~(4)构成的算法,一般称作基本微粒群算法或标准微粒群算法。考察式(1)可以发现,如果 $x_{i,j}(t) = y_{i,j}(t) = \hat{y}_j(t)$, 即微粒 i 当前的位置是它经历过的最好位置,也是整个微粒群迄今为止经

历过的最好位置,此时 $v_{i,j}(t+1)$ 仅仅依赖于 $\omega v_{i,j}(t)$ 这一项,当一个微粒接近全局最优位置时,其速度将接近于零,意味着标准微粒群算法最终将使所有的微粒在最优解附近都停止运动,不能保证收敛到全局最优。为克服这一缺陷, F. Van den Bergh 等对基本微粒群算法进行了如下的改进^[6], 称之为一般收敛 PSO (general convergent PSO, GCPSO)。

令 τ 是全局最优微粒的下标,则

$$y_\tau = \hat{y}.$$

全局最优微粒的速度更新公式做如下改进:

$$v_{\tau,j}(t+1) = -x_{\tau,j}(t) + \hat{y}_j(t) + \omega v_{\tau,j}(t) + \beta(t)(1 - 2r_{2,j}(t)). \quad (5)$$

其中: $-x_{\tau,j}(t)$ 项将微粒的位置从 $\omega_j(t)$ 进行重置, $\omega v_{i,j}(t)$ 项隐含了重新搜索的方向, $\beta(t)$ 是一个尺度因子, $\beta(t)(1 - 2r_{2,j}(t))$ 项使 PSO 产生一个围绕 $\hat{y}_j(t)$ 的随机搜索,搜索半径由 $\beta(t)$ 的大小控制。由式(2)和式(3)可以得出全局最优微粒的位置更新公式为

$$x_{\tau,j}(t+1) = \hat{y}_j(t) + \omega v_{\tau,j}(t) + \beta(t)(1 - 2r_{2,j}(t)). \quad (6)$$

3 多种群协同进化微粒群算法 (Multi-species cooperative PSO, MCPSO)

3.1 基本概念 (Basic concept)

定义 1 种群。具有一定相同飞行特征的微粒个体的集合作为一个子微粒群,称为种群。种群搜索的空间是解空间的一个子空间。

定义 2 种群特征。体现种群集体飞行有别于其它子微粒群的位置或速度分量称为种群飞行特征,简称种群特征,它唯一标识个体微粒所属的种群。

定义 3 种群划分。种群的划分是定义在种群特征之上的。设种群 S_j 中的个体微粒 X_i 表示为 $X_i = \{\rho_j, x_{i,1}, \dots, x_{i,n}\}$, 其中: ρ_j 表示特定种群的位置特征分量, $x_{i,n}$ 表示个体在子群空间中的自由位置分量。显然,对于 ρ_j 的不同定义,对应于一系列不相交的种群,所有这些种群空间的并集构成了完整的解空间。

定义 4 种群进化。种群特征分量 $\rho_j (j = 1, 2, \dots, N)$, 其中 N 是预定的种群数)的进化,称为种群进化。种群进化相比个体进化来说要漫长的多,所以其进化周期 T 在这里用个体层进化代数的正整数表示。

定义5 成熟种群. 用两次种群层进化操作间该种群中最优个体微粒适应值提高的效率表示种群的进化能力,即

$$A(t) = \frac{f_{\max}(t) - f_{\max}(t-T)}{f_{\max}(t)}. \quad (7)$$

其中 $f_{\max}(t)$ 表示进化到第 t 代时该种群中最优个体微粒的适应值. 取小正数 ε 为种群进化能力阈值, 若第 t 代时种群的进化能力 $A(t) > \varepsilon$, 则称该种群为成长种群, 反之, 称之为成熟种群.

定义6 超级个体. 进入成熟期的种群中的最优个体称为超级个体, 由一定数量超级个体组成的集合叫超级个体集合; 若成熟种群向超级个体集合提供超级个体时, 超级个体容量已达到设定值, 则用新进入的超级个体代替原来有的适应值最低的超级个体; 然后重新对超级个体的适应值进行排序.

3.2 算法设计(Design of the algorithm)

设 ρ_j 代表种群 j 的位置特征分量, $P_j X_i$ 代表种群 P_j 中的第 i 个微粒目前的自由位置矢量, $P_j Y_i$ 表示该微粒目前为止经过的最好位置, $P_j \hat{Y}$ 表示种群 P_j 中所有微粒到目前为止经历过的最好位置; P_j 中的第 i 个微粒可表示为 $b(\rho_j, P_j X_i)$, 其适应值函数表示为 $f(b(\rho_j, P_j X_i))$. 则多种群协同进化微粒群算法结构可表述如下:

初始化产生 n 个种群 $b(\rho_j, P_j X_1, P_j X_2, \dots, P_j X_m)$, $j = 1, \dots, n$; $i = 1, \dots, m$

reapt

```

for swarm  $j \in (1, \dots, n)$ 
  for particle  $i \in (1, \dots, m)$ 
    if  $f(b(\rho_j, P_j X_i)) < f(b(\rho_j, P_j Y_i))$ 
      then  $P_j Y_i = P_j X_i$ 
    if  $f(b(\rho_j, P_j X_i)) < f(b(\rho_j, P_j \hat{Y}))$ 
      then  $P_j \hat{Y} = P_j X_i$ 
  endfor

```

利用公式(1) ~ (3) 更新微粒 $b(\rho_j, P_j X_i)$

if $\text{mod}(t, T) = 0$

then 利用公式(1) ~ (3) 更新 ρ_j

if $E_j(t) < \varepsilon$

then P_j 进入超级个体集合

超级个体重新排序

endfor

until 符合算法停止条件

4 多种群协同进化 PSO 的径向基函数神经网络设计(Design of RBF neural network based on MCPSO)

4.1 RBF 神经网络分析 (Analyses of RBF neuronnetwork)

径向基函数神经网络是一种隐层激活函数采用径向基函数的单隐层的前向网络, 网络的输入输出层一般由应用问题来确定, 因此, 网络的复杂程度主要取决于隐层的节点数. 其数学模型表示为^[11]

$$\hat{q} = \sum_{n=1}^{n_c} \omega_n g(\|p - c_n\|_{\mathbb{R}^n} / \sigma_n). \quad (8)$$

其中: $p \in \mathbb{R}^n$ 为网络输入, \hat{q} 为网络输出, w_n 为网络输出层权值, $g(\cdot)$ 为径向基函数, c_n 为径向基函数的中心, σ_n 为径向基函数的敏感域, n_c 为隐层神经元数目, $\|\cdot\|_{\mathbb{R}^n}$ 为 p 与 c_n 之间的距离.

定义一个布尔变量 $L = (a_1, a_2, \dots, a_n, \dots, a_{n_{\text{hid}}})$, 其中元素 a_n 为 1 或 0, 分别对应隐层节点存在或不存在, n_{hid} 为布尔向量的长度. 作如下运算:

$$\rho = \sum_{n=1}^{n_{\text{hid}}} a_n 2^{-n_{\text{hid}}}, \quad n = 1, \dots, n_{\text{hid}}, \quad (9)$$

则必有 $0 < \rho < 1$. 这样, 当 ρ 在 $(0, 1)$ 区间内任取一随机数, 都有一个 n_{hid} 维的布尔随机向量 L 与之对应, 也就是说, RBF 网络的隐层节点可由随机数 ρ 确定的布尔向量 L 确定. 这时, 对应于训练样本 $\{p_i, q_i\}$ 的网络输出表达式可以写成

$$\hat{q}_i = \sum_{n=1}^{n_{\text{hid}}} a_n \times w_n \times g(\|p_i - c_n\|_{\mathbb{R}^n} / \sigma_n). \quad (10)$$

由式(8)(9) 容易得出, 径向基神经网络的训练即就是对训练样本集 $\{p_i, q_i\} (i = 1, 2, \dots, M)$, 寻找参数集 $\Theta = \{c_n, \sigma_n, \omega_n, \rho\}$ 使如下函数最小化^[12]:

$$\min \varepsilon_{\text{RBF}} = \frac{1}{M} \sum_{i=1}^M |q_i - \hat{q}_i|. \quad (11)$$

4.2 基于多种群协同进化微粒群算法的 RBF 神经网络学习策略 (Learning strategy of RBF neural network based on MCPSO)

1) 编码. 对于一个具有 n_c 个隐层神经元的 RBF 网络, 可以表示为微粒 $X_i = \{\rho_i, c_n^i, \sigma_n^i, w_n^i\}$, 其中: $n = 1, 2, \dots, n_{\text{hid}}$; $\rho_i \in (0, 1)$, 由 ρ_i 唯一确定的布尔变量 $L = (a_1, a_2, \dots, a_{n_{\text{hid}}})$ 确定了唯一的隐层节点数; c_n^i 表示第 n 个隐层神经元的径向基函数中心; σ_n^i 表示第 n 个隐层神经元的径向基函数宽度; w_n^i 表示第 n 个隐层节点与输出节点之间的连接权.

2) 初始化种群. 为保证初始化的种群均匀分布在解空间, 采用 Faure 序列产生均匀分布在 $(0, 1)$ 区间的 M 个随机数 $\rho_i (i = 1, 2, \dots, M)$, 用以确定 M 个初始化种群; 每个种群中随机产生 3 ~ 5 个个体微粒的初始位置矢量和相应微粒的初始飞行

速度矢量. 矢量的长度为 $(1 + m + k)n_{hid}$, 其中: m 为输入维数, k 为输出维数.

3) 适应值选择. 考虑到网络的复杂度, 个体微粒 X_i 的适应值按下式计算^[13]:

$$J(a_1, a_2, \dots, a_{n_{hid}}, c, \sigma, w) = \ln \varepsilon_{RBF} + \lambda \frac{n_{hid}}{N} = \ln \varepsilon_{RBF} + \frac{\lambda}{N} \sum_{n=1}^{n_{hid}} a_n. \quad (12)$$

4) 种群内微粒进化. 这里用 GPCSO 算法代替标准 PSO 算法, 种群内个体微粒 X_i 运动公式为

当 X_i 不是该种群内当前最优微粒时, 其运动迭代按式(1)(2)进行;

当 X_i 为该种群内当前最优微粒时, 其运动迭代按式(3)(4)进行.

5) 成熟种群层进化. 亦按 GPCSO 算法进行迭代.

6) 算法终止条件与最优解. 以超级个体集合中适应值最好的个体微粒作为当前全局最优解. 当该最优解的适应值达到预定要求或进化代数达到预定代数时, 算法终止并输出该最优解. 将该最优解按式(10)(11)解码为 RBF 神经网络模型.

4.3 仿真研究 (Simulation studying)

二维非线性系统 $Y_p = \cos(2\pi k_1 m + 2\pi k_2 n) \sin(2\pi k_2 n)$ 辨识, 其中 $0 < m, n < 36$; 取 $k_1 = 0.05, k_2 = 0.02$. 在 $m \in (0, 36), n \in (0, 36)$ 的区间内产生 200 组数据 $\{q, m, n\}$, 以其中的 100 组数据为训练样本, 另外 100 组为检验样本, 分别用 BP 算法、标准 PSO 算法和本文提出的 MCPSO 算法对 RBF 神经网络训练 500 步, 训练结果如表 1 所示, 其中 BP 算法的学习效率 $\eta = 0.2$, 两种 PSO 算法中 $\omega = 0.8, c_1 = c_2 = 0.5$.

表 1 3 种算法训练结果的比较

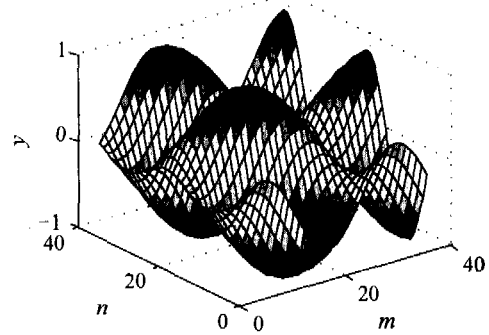
Table 1 Comparison of three algorithms

算法	网络结构	训练时间 /s	均方无误差	
			训练误差 E	检验误差 e
BP	2-10-1	4.32	3.52×10^{-3}	7.58×10^{-2}
CPSO	2-10-1	6.78	1.73×10^{-3}	4.62×10^{-2}
MCPSO	2-6-1	6.23	1.7569×10^{-4}	6.24×10^{-4}

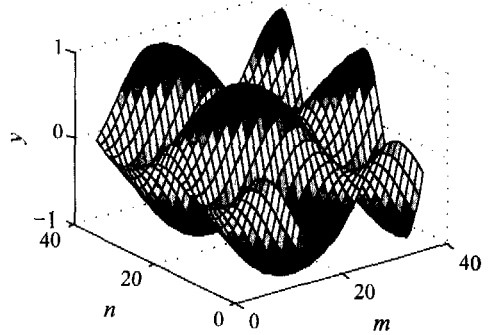
从表 1 看到, 用本文提出的 MCPSO 算法训练 RBF 神经网络时, 在训练时间上虽然和另外两种算法相比没有什么优势, 但在几乎相同的训练时间内,

得到的神经网络在隐层神经元减少的同时, 训练误差降低了一个数量级. 另外, 从检验误差来看, 采用 MCPSO 算法训练的 RBF 神经网络, 检验误差有明显降低, 说明网络的泛化能力得到了比较大的提高.

由 MCPSO 算法训练得到的 2-6-1 RBF 神经网络模型, 对系统的辨识的样本图形图 1(a) 所示, 图 1(b) 为辨识结果, 样本图和辨识图基本没有差别. 图 2 为训练误差, 由图 2 看出大部分区域的辨识误差都在 10^{-4} 以下, 只有个别的边缘地方辨识误差在 10^{-3} 水平. 图 3 为算法的收敛性比较, 可以看出 MCPSO 算法在训练 RBF 网络时, 在较少的迭代次数 t 内就能达到很高的训练精度.



(a) 样本图



(b) 辨识结果图

图 1 非线性系统辨识

Fig. 1 Nonlinear system identification

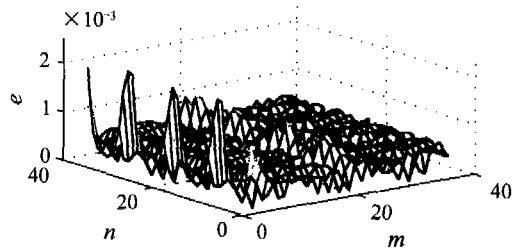


图 2 非线性系统辨识误差

Fig. 2 Error of nonlinear system identification

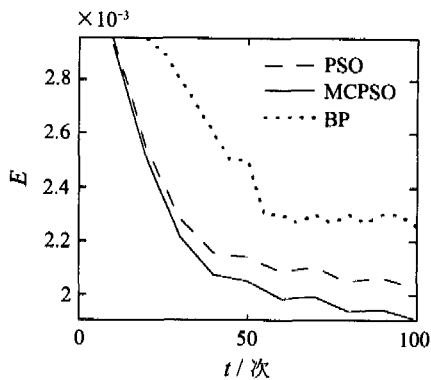


图3 训练误差比较

Fig. 3 Comparison of training error

5 结论 (Conclusion)

神经网络的进化设计一直是神经网络和进化计算中的一个研究热点. 本文根据基本微粒群算法的特点, 借鉴分层递阶编码的思想, 构造出一种基于种群特征分量的多种群协同进化微粒群优化算法, 用于径向神经网络的进化设计, 具有以下特点: 1) 采用种群内个体微粒自由运动特征的微粒群进化算法与种群运动特征分量的微粒群进化分层递阶进行, 以及各种群并行进化的方法, 有利于在并行计算机上运行; 2) RBF神经网络结构和连接权值分层递阶进化思想模仿了人工设计神经网络时连接权值调节-网络结构调节-连接权值调节这样的循环过程, 容易理解; 3) 利用MCPSO算法训练RBF神经网络, 使网络的逼近精度和泛化能力这一对矛盾体得到了比较好的协调统一.

参考文献 (References):

- [1] KENNEDY J, EBERHART R. Particle swarm optimization [C]// *Proc of IEEE Int Conf on Neural Networks*. Piscataway, NJ: IEEE Press, 1995: 1942-1948.
- [2] EBERHART R, KENNEDY J. A new optimizer using particle swarm theory [C]// *Proc of the 6th Int Symposium on Micromachine and Human Science*. Piscataway, NJ: IEEE Press, 1995: 39-43.
- [3] KENNEDY J. The particle swarm: Social adaptation of knowledge [C]// *Proc of IEEE Int Conf on Evolutionary Computation*. Piscataway, NJ: IEEE Press, 1997: 303-308.
- [4] SHI Yuhui, EBERHART R. A modified particle swarm optimizer [C]// *Proc of IEEE Int Conf on Evolutionary Computation*. Piscataway, NJ: IEEE Press, 1998: 67-73.
- [5] TRELEA I C. The particle swarm optimization algorithm: con-

vergence analysis and parameter selection [J]. *Information Processing Letters*, 2003, 85(9): 317-325.

- [6] BERGH F, ENGELBRECHT A P. A new locally convergent particle swarm optimization [C]// *Proc of IEEE Int Conf on System, Man and Cybernetics*. Piscataway, NJ: IEEE Press, 2002: 625-631.
- [7] 谢晓峰, 杨文俊, 杨之廉. 微粒群算法综述[J]. *控制与决策*, 2003, 18(2): 129-134.
(XIE Xiaofeng, YANG Wenjun, YANG Zhilian. Overview of particle swarm optimization [J]. *Control and Decision*, 2003, 18(2): 129-134.)
- [8] 彭宇, 彭喜元, 刘兆庆. 微粒群算法参数效能的统计分析[J]. *电子学报*, 2004, 32(2): 209-213.
(PENG Yu, PENG Xiyuan, LIU Zhaoqing. Statistic analysis on parameter efficiency of particle swarm optimization [J]. *Acta Electronica Sinica*, 2004, 32(2): 209-213.)
- [9] ROSIN C, BELEW R, MORRIS G, et al. New methods for competitive coevolution [J]. *Evolutionary Computation*. 1997, 5(1): 1-29.
- [10] BERGH F, ENGELBRECHT A P. A cooperative approach to particle swarm optimization [J]. *IEEE Trans on Evolutionary Computation*, 2004, 8(3): 1-15.
- [11] HARM F M, KOSTANIC I. *Principles of Neurocomputing for Science and Engineering* [M]. Columbus: McGraw Hill, 2001.
- [12] 石红瑞, 刘勇, 刘宝坤, 等. 基于混合递阶遗传算法的径向神经网络学习算法及其应用[J]. *控制理论与应用*, 2002, 19(4): 627-630.
(SHI Hongrui, LIU Yong, LIU Baokun, et al. RBFNN algorithm based on hybrid hierarchy genetic algorithm and its application [J]. *Control Theory & Applications*, 2002, 19(4): 627-630.)
- [13] 李祥飞, 敏恩, 张泰山. 基于混沌变量的前向神经网络结构优化设计[J]. *控制与决策*, 2003, 18(6): 705-707.
(LI Xiangfei, ZOU En, ZHANG Taishan. Optimization design of feed-forward neural network structure based on chaos variables [J]. *Control and Decision*, 2003, 18(6): 705-707.)

作者简介:

王俊年 (1968—), 男, 副教授, 博士研究生, 主要研究方向为智能控制及信号处理, E-mail: jnwang@hust.edu.cn;

申群太 (1945—), 男, 教授, 博士生导师, 主要研究方向为智能控制, E-mail: zhouyi_sq@ sina.com;

沈洪远 (1958—), 男, 副教授, 博士研究生, 主要研究方向为智能控制及数据挖掘, E-mail: hyshen@hust.edu.cn;

周鲜成 (1962—), 男, 副教授, 博士研究生, 主要研究方向为智能控制, E-mail: zxc650@ sina.com.