

采用中心聚类与 PSO 的 RBF 网络设计方法

刘俊, 商秀芹, 卢建刚, 陈金水, 孙优贤

LIU Jun, SHANG Xiu-qin, LU Jian-gang, CHEN Jin-shui, SUN You-xian

浙江大学 工业控制技术国家重点实验室, 杭州 310027

State Key Laboratory of Industrial Control Technology, Zhejiang University, Hangzhou 310027, China

E-mail: jglu@iipc.zju.edu.cn

LIU Jun, SHANG Xiu-qin, LU Jian-gang, et al. RBF design method based on center clustering and PSO. *Computer Engineering and Applications*, 2009, 45(36): 212-215.

Abstract: An algorithm for designing Radial Basis Function Neural Network based on center clustering and modified particle swarm optimization algorithm is introduced. Center clustering algorithm is proposed to process the input sample, and adaptively specify the hidden layer parameters of RBF neural network; classic PSO algorithm (Particle Swarm Optimization) with modified calculation method about globally optimal solution is used to train the parameters, and makes them tend to global optima; the output layer weights are updated on line by recursive least square algorithm with forgetting factor. At last, the hybrid algorithm is applied to design the prediction model about relationship between the sinter composition and drum strength. Then experiment validates the model with new data which are collected from sinter workshop. Experimental results show that the hybrid algorithm has high convergence speed, and model built by that has great prediction precision.

Key words: center clustering; Particle Swarm Optimization (PSO); Radial Basis Function (RBF); recursive least square; drum strength

摘要: 基于中心聚类法与微粒群 (PSO) 优化方法, 提出一种径向基函数 (RBF) 网络的设计算法。算法采用中心聚类方法对输入样本数据进行聚类处理, 自适应地确定 RBF 网络隐含层的初始参数; 利用修正全局最优解计算方法的经典 PSO 算法优化 RBF 网络隐含层参数, 进一步修正网络结构参数; 输出层权值采用带遗忘因子的递推最小二乘算法在线更新。采用该方法建立炼铁过程中烧结矿成分与转鼓强度关系的预测模型, 并用现场数据加以验证; 实验结果表明该方法收敛速度快, 所建立的模型具有较高的预测精度, 可用于复杂非线性系统建模。

关键词: 中心聚类; 微粒群优化; 径向基函数; 递推最小二乘; 转鼓强度

DOI: 10.3778/j.issn.1002-8331.2009.36.060 **文章编号:** 1002-8331(2009)36-0212-04 **文献标识码:** A **中图分类号:** TP183

1 前言

径向基函数神经网络 (RBFNN) 是一种典型的 3 层前馈神经网络, 可以用于函数逼近及分类; 与 BP 网络相比, RBF 网络结构更简洁, 学习速度更快^[1]。RBF 网络设计的主要工作是确定隐含层结构和参数, 即隐含层节点数目、中心向量及接受域宽度; 通常采用的方法是先确定 RBF 网络隐含层节点数目, 再由某种聚类算法对样本进行分类, 并由梯度训练算法或正交最小二乘算法确定网络在聚类条件下的中心向量和接受域宽度值。虽然这些方法在某些问题上能够满足使用要求, 但在样本数量大、相关程度高的条件下将显露其弊端, 如出现泛化能力差、拟合精度低等问题。

基于以上问题, 首先提出以“类中心距离”改进常规聚类算

法^[2-3]中的“个体距离”, 并利用改进算法处理样本得到初始聚类中心向量及各个子集的域半径, 然后由改进的经典粒子群优化^[4] (Particle Swarm Optimization) 算法获取当前最优隐含层参数。训练过程采用带遗忘因子的递推最小二乘算法在线求解网络的输出层权值, 从而进一步优化网络。最后文中用该方法建立基于现场数据的烧结矿成分与转鼓强度关系的预测模型, 仿真结果表明该方法可以有效减少网络的训练时间, 提高网络的预测精度, 可以在实际生产中根据预测结果调整配料系统的工作参数设定值, 优化配料。

2 RBF 网络的基本结构

RBFNN (Radius Basis Function Neural Networks) 是 1988

基金项目: 国家自然科学基金重点项目 (the Grand National Natural Science Foundation of China under Grant No.60736021); 国家高技术研究发展计划 (863) (the National High-Tech Research and Development Plan of China under Grant No.2006AA04Z184, No.2007AA041406); 浙江省科技计划项目 (the Scientific and Technical Project of Zhejiang Province of China under Grant No.2006C11066, No.2006C31051); 浙江省自然科学基金 (the Natural Science Foundation of Zhejiang Province of China under Grant No.Y4080339)。

作者简介: 刘俊 (1984-), 男, 硕士, 研究领域为计算机智能控制理论与应用; 商秀芹 (1983-), 女, 博士, 研究领域为复杂系统建模与优化; 卢建刚 (1968-), 男, 教授, 研究领域为复杂工程系统的建模、控制与优化; 孙优贤 (1940-), 男, 教授, 中国工程院院士, 研究领域为工业自动化。

收稿日期: 2009-09-21 **修回日期:** 2009-10-30

年 Broomhead 和 Lowe 将 Powell 于 1985 年提出的多变量插值的径向基函数应用于神经网络设计而构成的一种新的神经网络模型; 它是一种局部逼近性能很强的神经网络, 与 BP 神经网络相比 RBF 网络无论是在逼近能力方面, 还是在学习速度方面均具有优势。

RBF 网络是具有多层前向网的通用结构, 特点是它只有一个隐含层, 包含输入层共 3 层的网络结构形式。如图 1 所示是一个结构为 4-5-2 的 RBF 网络, 从输入层到输出层之间的映射关系是非线性的, 其中隐含层常用的激活函数是高斯函数, 即

$$o_i = \exp\left[-\frac{\|\mathbf{x} - \mathbf{c}_i\|^2}{2\sigma_i^2}\right] \quad (1)$$

其中, \mathbf{c}_i, σ_i 分别是网络中隐含层第 i 个节点的中心向量和接受域宽度, 也是该文设计的重点。

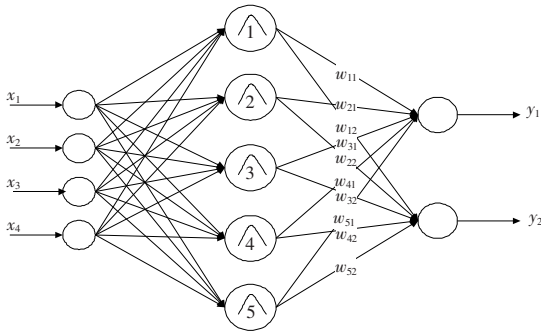


图 1 RBF 网络结构图

3 中心聚类算法

k -均值聚类算法属于聚类技术中的一种基本划分方法, 具有简单快速的优点; 其基本思路^[5]是从样本中选取 k 条记录作为初始聚类中心, 通过迭代运算将数据划分到不同的分类中, 进而使得某一分类中各个数据对象之间的欧氏距离尽可能小, 相似程度最大化。

由于 k -均值算法中聚类结果受到所选择初始聚类中心的影响, 且聚类后各个类的中心仍然是被聚类对象中的数据, 这不符合实际情况; 文献[2-3]中提出的算法在一定程度上解决了由于初始中心的选取带来的不确定性, 但是每次需要计算大量数据对象间的距离, 因此在大样本的情况下效率很低。

针对以上问题, 提出基于“类中心距离”的改进 k -均值算法, 不仅可以解决初始聚类中心选取对聚类结果的影响, 而且能够提高大样本条件下的聚类速度和精度。算法描述如下:

步骤 1 按式(2)计算数据样本 X_N^{Train} 中两两数据对象距离, 得到距离矩阵 $Dist$ 。

$$Dist = \{\|x_i - x_j\|; i, j = 1, 2, \dots, N\} \quad (2)$$

$distance_{mean} = \text{mean}(Dist)$, $distance_{hold} = Gain \cdot distance_{mean}$, $Gain$ 为该文提出的比例增益项, 反映样本密度, 取值在 0.6~0.85 之间, 实验研究中取 0.8。

步骤 2 搜索距离最短的向量组 x_i, x_j , 并作为 $Class_m$ 类的成员, 其中 m 为当前聚类对象的序号, 即已完成 $m-1$ 次聚类。

步骤 3 按式计算 $Class_m$ 的中心向量。

$$C_center_m = \frac{1}{Now} \sum_{i=1}^{Now} C_center_m(i) \quad (3)$$

步骤 4 计算 $(X_N - Class_m)$ 中 m 向量与当前 C_center_m 的距离, 并将所有距离小于 $distance_{hold}$ 的向量归为 $Class_m$, 从而增加了搜索的速度, 同时也避免了文献[2]中的逐一比较选

成的时间消耗, 同时计算算术中心与其他向量之间的距离要比计算 $Class_m$ 中每个向量与 $(X_N - Class_m)$ 中向量距离合理。

步骤 5 判断 $Sample = \{X_N^{Train} - Class_m\}$ 是否为空。若是则完成聚类, 进入步骤 6; 否则 $m = m + 1$, $X_N^{Train} = Sample$, 并返回步骤 1 进行下一轮聚类。

步骤 6 当前聚类中心向量构成的矩阵为:

$$Center = [Center_1, Center_2, \dots, Center_m] \quad (4)$$

每个聚类子样本的接受域半径构成向量为:

$$Distance_{hold} = [distance_{hold}^1, distance_{hold}^2, \dots, distance_{hold}^m] \quad (5)$$

每个聚类子样本数据对象为:

$$Class_i = [x_1^i, x_2^i, \dots, Center_i, \dots] \quad (6)$$

其中, $1 \leq i \leq m$, m 为聚类结果中类的数量。完成计算。

4 改进 PSO 算法

作为一种高效并行优化方法, PSO 算法可以用于求解非线性、不可微和多峰值的优化问题^[6-7]; 与遗传算法、蚁群算法等优化算法相似, PSO 算法也是一种基于群体的算法, 但是相比之下 PSO 算法的特点是: (1)任何个体都具有位置和飞行速度两个属性, 并随着优化过程的进行不断更新; (2)个体具有记忆能力, 能记住个体的最佳位置; (3)最优解是通过个体自身的学习和相互合作两者共同作用下达到的。由于 PSO 结构简洁、易于实现, 该文对经典 PSO 算法^[8]中当前全局最优解的选取方法进行了改进, 使得收敛的速度得到提升。算法的步骤如下:

步骤 1 初始化粒子。采用上述改进聚类算法对输入样本 X_N^{Train} 进行处理后, 得到 m 个聚类中心向量, 并将初始 RBF 网络的隐含层节点数目定为 m 。获得初始中心向量集合 $Center = [Center_1, Center_2, \dots, Center_m]$ 。在每个已分类的子样本 $Class_i = [x_1^i, x_2^i, \dots, Center_i, \dots]$ 中挑出一个向量, 共得到 m 个向量集 $[x_1^{index1}, x_2^{index2}, \dots, x_m^{indexm}]$, 其中向量的上标表示该向量在所属样本中的序号, 下标表示在挑出向量组中的序号。计算向量集中任意两两向量之间的平均距离得到一个距离向量 $Mean_Dist_vector = [dist_1, dist_2, \dots, dist_m]$, 将其乘以比例增益 $Gain$, 运算后的变换距离向量为 $Dist_vector = Mean_Dist_vector \cdot Gain$; 标准粒子群算法中粒子个数为 30, 因此同样类型的距离向量矩阵 $Dist_vector$ 共 30 个^[9], 将此设为各个粒子的初始位置, 实时全局最优位置为 $Global_Best = Distance_{hold} * \lambda$, 其中 λ 为该文提出的叠合系数, 反映神经元节点之间接受域宽度的叠合区比例, $Distance_{hold}$ 由式(5)确定, 它决定了聚类样本之间重叠的系数, λ 越大重叠越多, 取值范围为 1~2, 实验研究取 $\lambda = 1.4$ 。

步骤 2 求解每个粒子的适应度函数值, 文中适应度函数选用输出误差平方的负指数函数, 计算出当前最优接受域宽度向量, 同时与上一周期中的全局最优解相比较, 如果比上一周期最优解的适应度大则更新最优解。

$$Fitness_{particle} = \text{Fit}(Dist_Vector^{particle}) = \exp(-\|T_i^{particle} - Y_i^{particle}\|^2) \quad (7)$$

其中, $1 \leq particle \leq 30$ 。

$$Y_i = W \cdot O_i = W \cdot [o_1^i, o_2^i, \dots, o_{m-1}^i, o_m^i]$$

其中, W 为输出层权值矩阵, 有带遗忘因子 ρ 的递推最小二乘

法在训练过程中不断更新,并选择 $\rho=0.95$ 。隐含层的输出 o_{index}^i 由式(8)计算得出,即:

$$o_j^i = \Psi[(x_i - c_j), \sigma_{particle}^j] \quad (8)$$

其中, $1 \leq j \leq m$ 。

按式(9)计算当前各个节点接受域宽度的当前最优解:

$$Candidate = \frac{Fitness}{\sum_{particle=1}^{30} Fitness_{particle}} \begin{bmatrix} Dist_Vector^1 \\ Dist_Vector^2 \\ \vdots \\ Dist_Vector^{30} \end{bmatrix} \quad (9)$$

进入步骤3。

步骤3 更新粒子个体最优位置和该文改进的全局最优解计算。对每个粒子,比较这一代与上一代适应值,如果比上一代好则更新该粒子的最优解,即:

$$\text{If } Fitness_{particle}^{New} > Fitness_{particle}^{Old} \text{ then}$$

$$P_l^{particle} = Dist_Vector_{particle}$$

Else

$$P_l^{particle} = P_l^{particle}$$

End if

分别计算 *Candidate* 和 *Global_Best* 的适应度,如果前者比后者的适应度大则更新当前全局最优解 *Global_Best=Candidate*,并进入步骤4;否则保留当前全局最优解不变进入步骤4。

步骤4 判断是否已经达到最大训练次数 *maxepoch* 或 *Error<GivenValue*, *Error* 由式(10)计算:

$$Error = \sum_{j=1}^{Batch} \| Y_j - T_j \|^2 = \sum_{j=1}^{Batch} [W \cdot \Psi(x_j - \sigma_{Best}) - T_j]^2 \quad (10)$$

其中, *Batch* 为批量值,即累计 *Batch* 次的误差再与给定精度比较;如果上述两个条件满足其中之一则给出全局最优解 *Global_Best* 并退出,都不满足时进入步骤5。

步骤5 由改进经典 PSO 算法对粒子飞行速度和位置进行更新。

速度更新:

如果 $\min(Dist) \leq Dist_Vector^{index} \leq \max(Dist)$ 成立,则

$$V_{index} = K(V_{index} + c_1 r_1 (P_{index,d} - Dist_Vector^{index}) + c_2 r_2 (Candidate - Dist_Vector^{index}))$$

否则 $V_{index} = 0$ 。

位置更新:

如果 $\min(Dist) \leq Dist_Vector^{index} \leq \max(Dist)$ 成立,则

$$Dist_Vector^{index} = Dist_Vector^{index} + V_{index}$$

否则判断 $(Dist_Vector^{index} + V_{index}) \geq \max(Dist)$ 是否成立,若成立则有 $Dist_Vector^{index} = \max(Dist)$,若不成立则有 $Dist_Vector^{index} = \min(Dist)$ 。其中涉及参数^[9]如下:

$$K = \frac{2}{|2 - C - \sqrt{C^2 - 4C}|}, c_1 = 2.8, c_2 = 1.3, C = c_1 + c_2。$$

返回步骤2,进行下一个迭代周期。

5 仿真实验

利用上述算法设计 RBF 网络,建立炼铁过程中烧结矿成分与转鼓强度之间的关系模型,并利用现场数据验证算法的有

效性和模型精度。

在 Al_2O_3 含量比较稳定的情况下,通过主元分析^[10-11]和工艺分析,烧结矿成分参数中与转鼓强度有关的主要有 CaO、 SiO_2 、FeO、MgO 的含量^[12-14]。烧结矿在线取样并送至化验,每天有 12 组数据,取 360 组数据作为样本;从 360 组数据中任意选取 200 组作为 Part1,任意选取 200 组作为 Part2,Part1 用做聚类的对象和网络训练时使用的样本,Part2 用于测试网络的性能。部分数据如表 1。

表 1 成分参数与转鼓强度数据表

FeO	CaO	MgO	SiO_2	转鼓强度	FeO	CaO	MgO	SiO_2	转鼓强度
9.82	8.87	2.60	5.17	75.52	9.54	9.95	2.78	5.79	75.68
9.40	8.79	2.66	4.91	75.13	9.25	10.50	2.82	6.03	75.72
9.25	9.29	2.76	5.23	74.87	8.26	10.08	2.83	6.01	76.00
9.11	9.61	2.72	5.29	75.13	8.54	10.56	2.77	5.81	76.53
9.25	9.33	2.71	4.99	75.73	9.25	10.45	2.71	5.66	76.67
8.54	9.37	2.76	5.14	76.10	8.12	10.37	2.91	5.84	77.12
8.83	9.42	2.66	5.06	76.93	9.40	10.16	2.65	5.46	76.85
8.68	9.34	2.64	5.21	77.06

由于样本的量纲和取值范围不同,需对样本进行归一化处理再按照文中提出的算法进行聚类和优化,聚类的数据对象是 4 个数据项 FeO、CaO、MgO、 SiO_2 构成的记录,实验中相关参数的选择如表 2。如此选择的参数所建立的 RBF 网络最终结构为 4-7-1,即隐含层具有 7 个节点。

表 2 实验参数表

参数名称	取值范围	实验取值
比例增益 <i>Gain</i>	0.6~0.85	0.8
叠合系数 λ	1~2	1.4
聚类子集数 <i>m</i>	>1	7
粒子数 <i>particle</i>	>1	30
遗忘因子 ρ	0.9~1	0.95

如图 2 所示,只包含 FeO、CaO、MgO 3 个数据项的聚类结果;在同一台计算机上,文中提出的改进聚类算法耗时 1.084 6 秒,文献[2-3]聚类算法均需要 1.7 秒以上。

从图 2(a)和图 2(b)的聚类结果可以分析出:

(1)基于“个体距离”所得到的类中心是样本点中的某些值,而样本点直接作为中心不能真实反映实际的中心,因此聚类结果的可信度比较低;基于“类中心距离”聚类后得到的聚类中心值不会局限于现有样本点的值,而是通过欧式距离计算比较后得到的某些样本点加权计算得到的中心向量,是样本点的几何中心,能够反映中心的实际意义,因此在实验中聚类结果也比较理想。

(2)经该文改进的动态聚类算法对样本聚类处理后,所得的样本中心分布及其位置的数量可以直接反应中心区域范围内样本出现的概率,而基于“个体距离”聚类仅能体现一个局部特性,在变量维数较小(如二维)的情况下能够得到较好的聚类结果。如图 2(a)和图 2(b)左侧是样本出现概率较大的部分,因此需要一定数量且分配均匀的样本中心点,图 2(a)仅有 2 个分配在纵向上样本水平中心线上,不能达到要求,而图 2(b)则有 3 个分配在 3 个方向的样本中心线上,因此聚类效果较好。

经过该文提出的动态聚类方法的结果初始化 RBF 网络的隐含层参数,再按照文中描述的改进的经典 PSO 算法优化网络的参数,Part2 数据样本分别测试优化前后的 RBF 网络如图 3 所示。

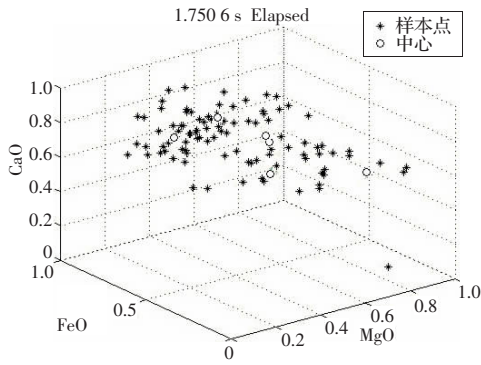


图 2(a) 文献[2]算法聚类结果

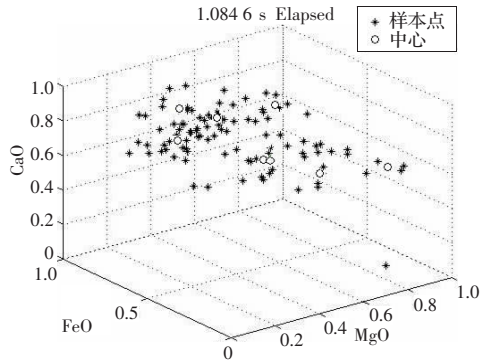


图 2(b) 中心聚类算法聚类结果

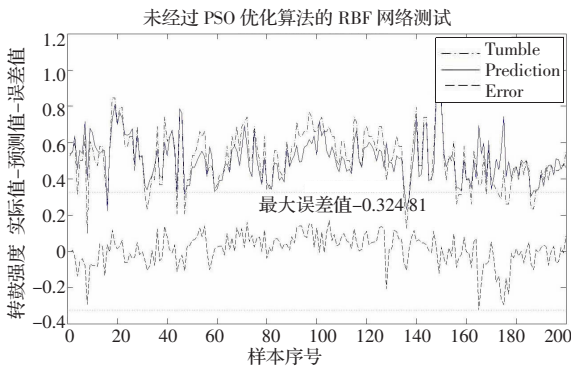


图 3(a) 优化前测试结果

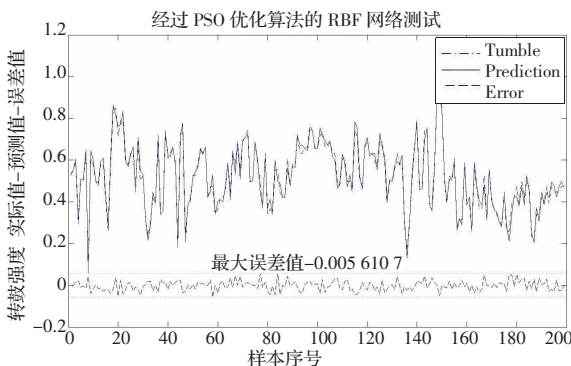


图 3(b) 改进 PSO 算法优化后的测试结果

从仿真结果可知优化前的最大误差绝对值为 0.324 81, 优化后的最大误差绝对值为 0.005 610 7, 经优化的网络预测精度

较高, 误差控制在 6% 以内; 而未经任何优化的网络输出误差相对较大, 说明了算法的有效性, 由于生产过程是连续进行的, 设计定期自动更新模式的优化将进一步提高预测的精度。实验测试数据直接来自现场, 未经过任何筛选, 故算法处理数据的鲁棒性强, 能够适应数据种类多、数据量大的工业过程, 因此基于该文提出的 RBF 网络设计方法在建立大时滞对象的预测模型上有很好的应用前景。

6 结束语

针对 RBF 网络隐含层参数初始化与在线更新问题进行了分析, 并将中心聚类算法和修改后的经典微粒群算法分别用于 RBF 网络隐含层参数的初始化阶段和参数的优化训练阶段。该聚类算法能有效减少聚类的时间消耗, 提高聚类的效率; 运用加权计算获得最优解的方法代替经典 PSO 算法中的全局最优解的选取方法, 增大了全局最优解的搜索范围, 同时也提高了网络的预测精度。实验结果表明文中提出的 RBF 网络设计方法能用于建立复杂预测模型, 并且具有很好的预测精度, 在大滞后系统中有很好的推广价值和应用前景。

参考文献:

- [1] 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.
- [2] 刘艳丽, 刘希云. 一种基于密度的 K -均值算法[J]. 计算机工程与应用, 2007, 43(32): 153-155.
- [3] 徐义峰, 陈春明, 徐长青. 一种改进的 k -均值聚类算法[J]. 计算机应用与软件, 2008(3): 76-78.
- [4] Kennedy J, Eberhart R C. Particle swarm optimization[C]//Proc IEEE Int Conf Neural Networks. Piscataway, NJ: IEEE Press, 1995: 1942-1948.
- [5] 李国勇. 智能控制及其 MATLAB 实现[M]. 北京: 电子工业出版社, 2007: 31-36.
- [6] Kanungo T, Mount D M, Netanyahu N S, et al. An efficient k -means clustering algorithm: Analysis and implementation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 27(7): 881-892.
- [7] 谢晓锋, 张文俊, 杨之廉. 微粒群算法综述[J]. 控制与决策, 2003(2): 129-134.
- [8] Parsopoulos K E, Vrahatis M N. Particle swarm optimization method for constrained optimization problems[M]//Intelligent Technologies: From Theory to Applications. Amsterdam: IOS Press, 2002: 214-220.
- [9] 曾建潮, 介婧, 崔志华. 微粒群算法[M]. 北京: 科学出版社, 2004.
- [10] 张顶学, 关治洪, 刘新芝. 基于 PSO 的 RBF 神经网络学习算法及其应用[J]. 计算机工程与应用, 2006, 42(20): 13-15.
- [11] 翟军勇, 费树岷. 基于主元分析的 RBF 神经网络多模型切换控制[J]. 系统工程与电子技术, 2006(7): 1051-1054.
- [12] 于素荣, 于淑娟, 刘艳辉, 等. MgO 含量对烧结矿烧结指标及冶金性能的影响[J]. 鞍钢技术, 2008(5): 23-26.
- [13] 张金柱, 邓海亮, 敖万忠, 等. 烧结混合料中 SiO_2 和 FeO 含量对烧结矿强度的影响[J]. 钢铁, 2008(5): 18-21.
- [14] 张斌, 雷仕江. 水钢烧结矿转鼓强度下降的原因分析及对策[J]. 烧结球团, 2008(1): 56-58.