

Optimization on FNN based on genetic algorithm and its application on CCR soft sensor

Shihuai Chen, Ziqiang Sun, Xingsheng Gu

Research Institute of Automation, East China University of Science and Technology, Shanghai 200237, China

Abstract—The Continuous Catalyst Reforming Unit plays an important role in the refinery. In the regeneration tower, oxygen content which is very important to catalyst regeneration process is hard to measure. So T-S fuzzy model is used on estimation of oxygen content. Based on NARMAX model identification of T-S fuzzy-neural-network (FNN), the genetic algorithm is applied to optimize the membership functions and network parameters, rule sets which can only be acquired by experiences. So that it can make better performance with the optimized parameters. The improved GA combines the advantages of GA's strong search capacity and conventional optimization technologies's fast convergence and accuracy merits. Therefore, the algorithm achieves a trade-off between accuracy, reliability and computing time in global optimization. Finally validity and accuracy of the present model is verified by the On-field performance on oxygen content soft sensor modeling that was put into operation recently.

Keywords—fuzzy-neural-network, genetic algorithm, soft sensor

基于遗传算法的 T-S 模型优化及其在再生器氧含量软测量中的应用

陈诗淮 孙自强 顾幸生

华东理工大学自动化研究所, 上海 200237

摘要 连续重整装置在整个炼油装置中有举足轻重的作用, 而再生塔的氧含量对催化重整过程很重要但很难测得。因此 T-S 模型被用来预测氧含量, 这里从基于 NARMAX 模型辨识的 T-S 模糊神经网络入手, 利用遗传算法对 T-S 模糊神经网络模糊控制规则进行寻优, 并有效地确定出较优的控制规则和隶属度函数, 从而达到了预定的要求。算法综合了 GA 强大空间搜索能力和传统优化方法的快速收敛和高精度的优点, 在保证全局优化的前提下, 综合考虑了模糊控制器的复杂程度、训练速度和控制精度。将该方法用于某炼油厂连续催化再生器氧含量模型的软测量建模, 解决了再生器氧含量实时测量难题。现场实际运行结果表明所建氧含量软测量模型精度较高, 优于采用一般的 T-S 模糊神经网络所建模型。

关键词 模糊神经网络, 遗传算法, 软测量

1. 引言

连续催化重整是石油加工过程中重要的二次加工方法, 能生产高辛烷值汽油和芳烃、氢气, 在炼油厂占有重要的地位。在催化剂再生过程中, 再生器燃烧区入口氧含量是一个重要的被控变量。较高的氧含量会导致较高的燃烧温度, 造成催化剂损失; 较低的氧含量会造成焦炭燃烧较慢, 使催化剂不能在燃烧区完全燃烧。现有的再生器氧

含量测量办法是利用一种特殊的进口氧气分析仪, 它基于氧化锆测量原理工作。由于测量环境和采样点条件恶劣, 氧分仪工作不很稳定, 且极易腐蚀损坏; 同时, 由于该分析仪是进口产品, 除了仪表价格高外, 采购周期也过长。一旦停止对再生器氧含量的测量, 就会直接影响再生器燃烧区温度的控制, 从而影响整个生产过程。

本文提出一种基于 T-S 模糊神经网络的软测量方法,

实现对再生器氧含量的软测量。T-S 模型是把输入信号引入解模糊环节,即输出以输入变量的显函数形式表示,形成一种从输入、输出数据集中系统地产生模糊规则的方法。在设计 T-S 模型时,需要根据有经验的操作者或专家的经验制定出相应的控制规则和隶属度函数。然而,对于控制规则和隶属度函数的获取,专家的经验只能起到一个指导的作用,很难确定出各项参数,只能采用试探的方法不断调整,这实际上是一个寻优的过程。

遗传算法(GA)是以达尔文生物进化论为起源创建的,基于进化中优胜劣汰、自然选择、适者生存和物种遗传思想的搜索算法。由于其具有对问题依赖性小,能求得全局最优解等优点,可以用于模糊控制规则的寻优过程,并能有效地确定出较优的控制规则和隶属度函数,为模糊控制系统的建立提供了极为有利的条件。

本文正是运用遗传算法(GA)的上述特点,对 T-S 模型的模糊规则和隶属度函数进行优化,最终建立基于遗传算法的 T-S 模糊神经网络模型。

2. T-S 模糊神经网络

T-S 模型有如下一组模糊规则表示:

$$R^i: \text{if } x_1 \text{ is } A_1^i \text{ and } x_2 \text{ is } A_2^i, \dots, x_m \text{ is } A_m^i \text{ then } y^i = p_0^i + p_1^i x_1 + p_2^i x_2 + \dots + p_m^i x_m \quad (1)$$

其中, R^i 表示第 i 条规则, x_i 为模糊模型的第 i 个输入变量, A_j^i 为变量 x_j 的第 j 个模糊子集, y^i 为第 i 条规则

的输出, p_j^i 为变量 x_j 与 y^i 的关系值, p_0^i 为一实数。

这里采用基于 T-S 模糊推理系统的 NARMAX 模型便是来建立模糊神经网络:

NARMAX 模型的辨识,实际上就是辨识非线性函数 $f(\cdot)$ 。令

$$y(k) = \begin{bmatrix} y_1(k) \\ y_2(k) \\ \vdots \\ y_m(k) \end{bmatrix}, \quad x(k) = \begin{bmatrix} x_1(k) \\ x_2(k) \\ \vdots \\ x_r(k) \end{bmatrix} \quad (2)$$

分别为系统的输出、输入, n_y 、 n_u 分别为系统的输出、输入的最大滞后, 把它们的延时量 $y(k), \dots, y(k-n_y); x(k-1), \dots, x(k-n_u)$ 作为网络的增广输入。这样可以得到:

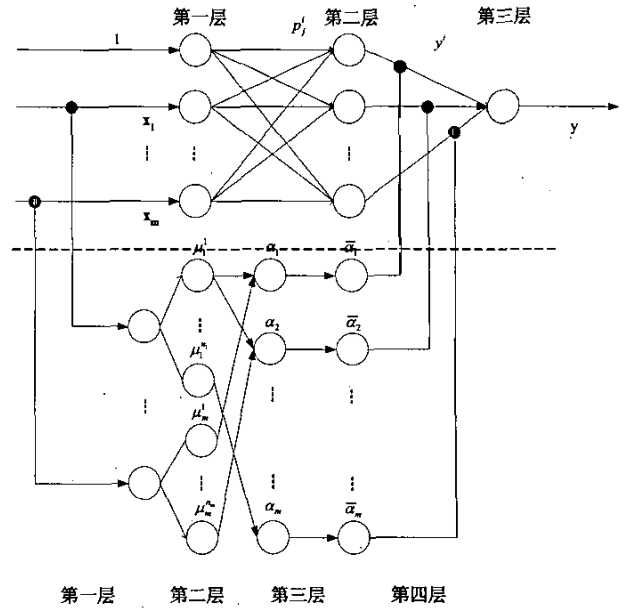


图1 T-S 模糊神经网络结构图

$$y(k) = f(y(k-1), \dots, y(k-n_y), x(k-1), \dots, x(k-n_u)) \quad (3)$$

因此, 一个简单的网络结构见图1。

即由输入变量和前一时刻的输出变量作为网络的输入, 则第 i 条模糊规则为:

$$y^i(k) = p_0^i + p_1^i x_1(k-1) + \dots + p_{m-1}^i x_{m-1}(k-1) + p_m^i y^i(k-1) \quad (4)$$

(图一中虚线以上部分为后件网络, 虚线以下部分为前件网络)

(1) 前件网络:

第一层为输入层 $x = [x_1, x_2, \dots, x_n]^T$ 。

第二层计算各输入分量模糊集合的隶属度 μ_j^i 表示第 i 个输入变量的第 j 个模糊隶属度。

假定输入变量的模糊分割数分别为 m_1, m_2, \dots, m_n , 隶属函数采用广义钟形隶属函数, 其形式如下:

$$f(x, a, b, c) = \frac{1}{1 + \left| \frac{x-c}{a} \right|^{2b}} \quad (5)$$

它的三个参数 a, b, c 都是需要学习的参数。

第三层计算出每条规则的适用度 α_j , 即

$$\alpha_j = \min\{\mu_1^i, \mu_2^i, \dots, \mu_n^i\} \text{ 或 } \alpha_j = \mu_1^i \mu_2^i \dots \mu_n^i \quad \text{其中}$$

$$i_1 \in \{1, 2, \dots, m_1\}, i_2 \in \{1, 2, \dots, m_2\}, \dots, i_n \in \{1, 2, \dots, m_n\}, j = 1, 2, \dots, m,$$

$$m = \prod_{i=1}^n m_i$$

该层节点数为 m 。

第四层实现的是归一化运算, 即

$$\bar{\alpha}_j = \alpha_j / \sum_{j=1}^m \alpha_j; \quad j = 1, 2, \dots, m \quad (6)$$

(2) 后件网络:

第一层为输入层, 数字 1 的作用是提供模糊规则后件

中的常数项 $p_0^i (i = 1, 2, \dots, m)$ 。

第二层计算每一条规则的后件, 即

$$y^i = p_0^i + p_1^i x_1 + \dots + p_n^i x_n, i = 1, \dots, m \quad (7)$$

第三层计算系统输出, 即

$$y = \sum_{i=1}^m \bar{\alpha}_i y^i \quad (8)$$

这里 $P_0^i, P_1^i, \dots, P_n^i$, 是需要学习的参数。

3. 遗传算法

用遗传算法优化 T-S 网络分为以下几个步骤: (1) 染色体编码 (2) 指定适应度函数 (3) 遗传算子的制定 (4) 染色体的初始化 (5) 参数设定: 设定诸如染色体数量、染色体长度、交叉概率、突变概率、终止条件等参数 (6) 由遗传算法按复制, 交叉, 变异三种操作寻找最佳串 (7) 根据最佳串 S_{np} 给出实际问题的最优解。

针对上面的 T-S 网络结构, 利用遗传算法对其前件网络第二层的隶属度函数参数 a, b, c (式-5) 和后件网络第二层的 $P_0^i, P_1^i, \dots, P_m^i$ (式-7) 进行优化。假定输入为 x_1, x_2, x_3 , 其中 x_1, x_2 为输入变量 x_3 为前一刻的输出变量作为当前的输入变量, 输出为 y , 设每个变量的模糊分割数为 2, 这样前件网络有 $6 * 3 = 18$ 个参数, 后件网络有 2^5 条规则即 $2^5 * 7 = 56$ 个参数, 前后共 74 个参数, 把这 74 个参数作为一个基因编码, 为了提高运算准确性和减小耗时, 选用浮点数编码, 从前到后依次 1, 2, 3, ..., 74, 组成了一条染色体进行寻优。

3.1 交叉运算和变异运算

首先, 为了提高运算的准确性, 选择单点交叉算子和

高斯变异算子。同时这里我们采用一种自适应 P_c 和 P_m 方法, 用适合度函数来衡量算法的收敛状况, 其表达式为:

$$P_c = K_1 / (f_{MAX} - f_{AVG}) \quad (9)$$

随着遗传算法在线性性能的提高, 可以增大交叉概率 P_c 的取值。

$$P_m = K_2 / (f_{MAX} - f_{AVG}) \quad (10)$$

随着遗传算法性能的下降, 可以减小变异概率 P_m 的取值。其中, f_{MAX} , f_{AVG} 分别为群体中的最大适合度和平均适合度, $f_{MAX} - f_{AVG}$ 体现了群体的收敛程度, $K_1, K_2 \leq 1.0$ (为常数)。

3.2 选择适当的适应值函数

在遗传算法中, 选择过程是根据解的适应值进行的: 调整解的适应值函数和其目标函数间的关系可以控制遗传算法的迭代收敛速率: 本文选择的适应度函数为:

$F = 1 / (1 + E)$, 式中 E 为与控制目标的误差:

$$E = \sqrt{\frac{\sum_{i=1}^n (Y - Y^i)^2}{n}} \quad (11)$$

Y^i 为模型输出, Y 为系统分析值。

3.3 计算最佳个体

这里采用最优保存策略, 将每一代种群中的最佳个体 (适应值最大的个体) 保留下来, 不参加杂交和变异过程, 使之直接进入下一代, 这样可以防止优秀的个体由于选择、交叉或变异中的偶然因素而被破坏。具体步骤是:

- (1) 找出当前群体中适应度最高的个体和适应度最低的个体。
- (2) 若当前群体中最佳个体的适应度总比迄今为止的最好的个体的适应度还要高, 则以当前群体中的最佳个体作为迄今为止最好的个体。
- (3) 用迄今为止最好个体替换掉当前群体中的最差个体。

算法中各有关参数确定如下: 种群数目 $n = 50$, 迭代代数 $N = 30$, 交叉概率 $P_c = 0.8$, 变异概率 $P_m = 0.15$ 。这些参数不仅影响到遗传算法的内存量与计算时间, 而且会影响到寻优结果。

4. 实际应用

将上述方法用于某炼油厂连续催化重整再生器氧含量软测量建模。根据工艺技术分析, 确定可用于建模的辅助变量有 21 个, 包括再生烧焦温度、再生气体温度、再生气体流量、再生气体压力、催化剂循环量、待生催化剂结焦

含量等。从现场生产记录中,选取了300组数据,其中前200组用于建模训练,后100组用于模型检验。由于辅助变量较多,因此先对辅助变量进行主元分析,根据总的方差贡献率大小,选择5个主元变量作为模型的输入,另外再将上一时刻的输出变量作为当前时刻的一个输入变量。这样模型共有6个输入变量和一个输出变量。利用遗传算法的T-S模型和一般的T-S模型的检测输出结果分别如图2和图3,图中虚线为氧含量的实际分析值,实线为模型输出值。前者的均方根误差为 1.5×10^{-4} ,后者的为 3.5×10^{-4} 。结果显示:利用遗传算法的T-S模型得到了满足设计要求的参数值,改善了T-S模糊控制品质,从而表明了这种方法的可行性。

5. 结论

本文利用遗传算法对问题依赖性小,无需先验知识,能求得全局最优解等优点,用于T-S模糊神经网络模糊控制规则的寻优过程,并有效地确定出较优的控制规则和隶属度函数,将这个用于连续催化再生器氧含量软测量的建模,投入实际运行取得了良好的效果。

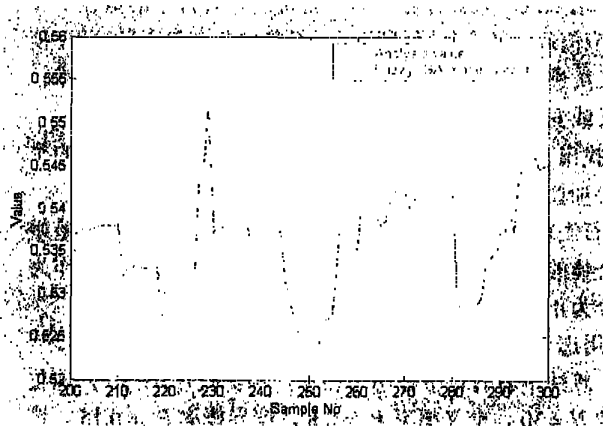


图2 Fuzzy+GA 模型检验

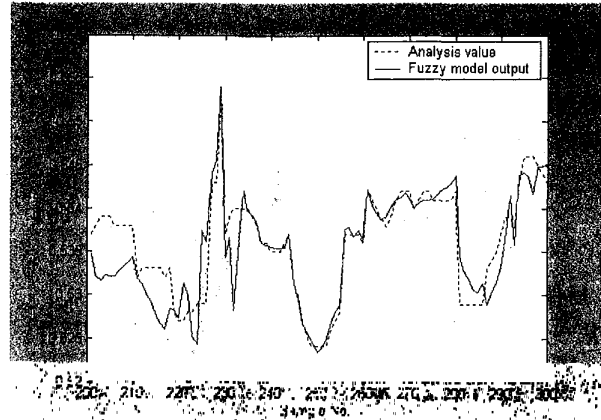


图3 Fuzzy 模型检验

参考文献

- [1] 刘增良, 模糊技术与应用选编(3), 北京航空航天大学出版社, 1998.
- [2] J. J. , , 1 eN: . 9kS 6 F 5 E9L , 1 9 9 7 .
- [3] Johansen T. A. , Shor and identification o models . IEEE Transact 2 9 7 - 3 1 3 .
- [4] MP. o, tPtemaarnsno n R. K. . B with output multipli 1 3 1 - 1 4 0 .
- [5] Goldberg D. E. , Gen Optimization and Mac 1 9 8 9 .