

DOI: 10.16516/j.gedi.issn2095-8676.2019.04.010

基于 MI 时序处理的 GA-BP 脱硫制浆 系统能耗建模

金秀章, 李奕颖

(华北电力大学(保定) 控制与计算机工程学院, 保定 071003)

摘要: [目的] 石灰石—石膏湿法脱硫是目前燃煤电厂应用最广泛、技术最为成熟的一种烟气脱硫技术, 石灰石浆液制备作为其中一道高耗能工序, 具有生产过程复杂, 物耗与能耗间存在非线性关系等特点, 缺乏合理有效的能耗模型。为建立能够对生产参数优化提供指导的可靠的制浆系统能耗模型。[方法] 基于某 600 MW 电厂实际运行数据, 选择生产过程中的可控制量作为输入, 并基于互信息(Mutual Information)理论调整各输入变量间的时滞关系, 采用结合遗传算法(Genetic algorithm, GA)的改进 BP 神经网络建立了制浆系统的单位制浆能耗模型。[结果] 试验结果表明: 与未调整时序的 GA-BP 模型和标准 BP 算法的模型相比, 经过时序调整的 GA-BP 模型的计算结果能够更为准确地接近制浆系统生产实际数据。[结论] 所建立的模型可以应用到浆液制备过程的能耗优化研究中。

关键词: 湿法脱硫; BP 神经网络; 能耗模型; 石灰石制浆系统

中图分类号: TM611, TM73

文献标志码: A

文章编号: 2095-8676(2019)04-0064-05

Energy Consumption Modeling of GA-BP for Desulfurization Pulping System Based on MI Time Series Processing

JIN Xiuzhang, LI Yiyang

(School of Control and Computer Engineering, North China Electric Power University (Baoding), Baoding 071003, China)

Abstract: [Introduction] Limestone-gypsum wet desulfurization is a flue gas desulfurization technology with the most extensive application and the most mature technology in coal-fired power plants. Limestone slurry preparation is one of the high-energy-consuming processes, which has complex production process and difficult to correspond to material consumption and energy consumption. There is no reasonable and effective energy consumption prediction model. In order to establish a reliable pulping system energy consumption model that can guide the optimization of production parameters. [Method] Based on actual operating data of a 600 MW power plant, the controllable quantity in the production process was selected as input, and the time-delay relationship between each input variable was adjusted by mutual information theory. The improved BP neural network combined with genetic algorithm (GA) was used to establish the unit pulp energy consumption model of the pulping system. [Result] The experimental results show that compared with the unadjusted timing GA-BP model and the standard BP algorithm model, the calculation results of the GA-BP model with time series adjustment can more accurately approach the actual production data of the pulping system. [Conclusion] The established model can be applied to the energy optimization study of the slurry preparation process.

Key words: wet desulfurization; BP Neural networks; energy consumption model; limestone pulping system

SO₂ 是目前大气污染物中数量较大、影响范围

广的一种气态污染物, 其主要来源为化石燃料的燃烧过程^[1]。随着国家环保政策的推行, 火电厂普遍引入了脱硫设备^[2]。石灰石—石膏湿法脱硫由于其设备成本低, 原料来源广, 脱硫效率高等优点在火电厂中广泛使用^[3-4], 其中石灰石浆液制备过程耗

收稿日期: 2019-03-21 修回日期: 2019-05-20

基金项目: 国家科技重大专项资助“煤炭清洁高效利用和新型节能技术”(2016YFB0600701)

电量大,且缺乏有效的控制指导,极易造成物料和用电的浪费。因此,建立合理的制浆能耗模型,指导系统耗电优化是十分必要的。

石灰石浆液制备是一个复杂的物理过程,各控制参数相互影响,且物料与能耗间没有明确的对应关系,难以通过生产机理分析建立制浆系统的能耗模型。BP神经网络是常用的数据驱动模型,能够以任意精度逼近任何非线性函数,但存在收敛速度慢,易陷入局部极小值,稳定性差等缺点^[5]。本文提出了基于遗传算法的改进BP神经网络算法,对网络的权值和阈值进行优化,提高了算法收敛速度及全局寻优能力。

在对模型辅助变量的选择方面,充分考虑到其对控制过程的指导作用,选取制浆过程中可调节的控制参数作为输入训练网络。由于生产过程中各参数之间存在时间延迟,通过计算变量间的互信息对数据的时序进行了调整,进一步提高了模型的准确程度。

1 互信息时序调整

在浆液制备过程中,各变量之间存在一定的延迟,例如石灰石给料量与给水流量直接影响着成品浆液量的多少及成品浆液的密度,但给料量的变化并不会立刻反映到制浆量及制浆品质上,若不对变量间的时序关系进行调整,则会影响到模型的训练效果和精度。

目前常用的变量关联度指标有欧氏距离指标,线性相关系数, Spearman 相关系数^[6]等,但这些指标通常仅能反映变量间的线性关系,且会受归一化等无量纲处理的影响,无法对实际工程中复杂的非线性变量的因果关系进行计算。

信息是事物运动的状态和运动状态变化的反映^[7],互信息作为信息熵理论的扩展,能够衡量两个变量间的相关程度^[8]。互信息法利用计算互信息函数的最大值来确定最佳延迟时间,是一种非线性的分析方法,在混沌时间序列预测和变量筛选等领域中有广泛的应用^[9-11]。

给定两个随机变量 X 和 Y ,若它们各自的边缘概率分布和联合概率分布分别为 $p(x)$, $p(y)$ 和 $p(x, y)$,则信息熵函数可以定义为:

$$H(X) = - \sum_i p(x_i) \ln p(x_i) \quad (1)$$

式中: $p(x_i)$ 为变量 X 在状态 i 出现的概率,信息熵 $H(X)$ 能够代表 x_i 的不确定程度。 $H(Y)$ 同理,代表 y_i 的不确定度。

变量 X 和 Y 的联合熵定义为:

$$H(X, Y) = - \sum_{i,j} p(x_i, y_j) \ln p(x_i, y_j) \quad (2)$$

那么,对于给定的变量 x_i ,由于变量 y_j 的发生和二者之间的相关性,使其不确定性减少的熵值称为互信息:

$$I(x, y) = H(x) + H(y) - H(x, y) \quad (3)$$

利用实验数据计算互信息函数,其主要工作是计算 $p(x)$, $p(y)$ 和 $p(x, y)$ 。通常采用划分网格的方法^[12],将变量组成的样本空间划分为若干网格,然后通过统计各个格子中的样本数来求出其概率值。

将计算出的概率分布带入式(1)至式(3)可以求出在给定的时间序列下的互信息值,当 $I(x, y)$ 达到最大值时的时间 t 作为变量间的延迟时间。这种计算方法包含了时间序列的非线性特征,能够较为有效地调整变量间的延迟。

2 GA-BP 改进神经网络

2.1 BP 神经网络

BP神经网络是一种按误差反向传播算法训练的多层前馈神经网络,现已广泛应用于函数逼近、模式识别、自适应控制等领域^[13-15],网络主要参数有:输入层到隐层的权值 v_{ij} ,隐层到输出层的权值 w_{jk} ,隐层阈值 a 和输出层阈值 b 。当误差未能达到目标精度时,对网络权值进行更新。

由于传统BP算法权值和阈值的更新采用梯度下降算法,不可避免地存在收敛速度慢、易陷入局部极小点等缺点。而遗传算法是一种基于自然选择和模拟生物进化过程的全局优化随机搜索算法^[16]。该算法具有较强的鲁棒性,可将其应用于优化神经网络模型权值的学习和训练中。

2.2 遗传算法优化神经网络

遗传算法优化BP神经网络包含如下几个步骤:

1) 确定BP神经网络的结构,即输入变量个数 n 、输出变量个数 l 和隐含层节点数 m ,并由此确定遗传算法中个体编码的长度。由于每个个体都要包含整个网络的所有权值和阈值,故个体长度为:

$$s = n \times m + m \times l + n + l \quad (4)$$

2) 种群初始化, 确定种群规模和进化代数, 并对种群中个体进行编码。在 BP 网络模型中, 其连接权值和阈值都是连续的, 因此采用实数编码方法。

3) 确定适应度函数, BP 神经网络模型性能评价一般利用均方误差 (MSE) 函数, 因此在遗传优化算法中, 使用均方误差函数作为计算个体适应度的函数。

4) 选择操作, 采用轮盘赌法^[17]对个体进行筛选, 每个个体的选择概率为:

$$p_i = \frac{f_i}{\sum_{i=1}^N f_i} \quad (5)$$

式中: f_i 为个体 i 的适应度值的倒数; N 为种群规模。

5) 交叉操作, 由于个体为实数编码, 所以采用实数交叉法, 个体 a_k 和个体 a_j 在 i 位进行交叉操作如下:

$$\begin{aligned} a_{ki} &= a_{ki}(1-b) + a_{ji}b \\ a_{ji} &= a_{ji}(1-b) + a_{ki}b \end{aligned} \quad (6)$$

式中: b 为 $[0, 1]$ 区间的随机数。

6) 变异操作, 个体 a_i 中的第 j 个位置基因 a_{ij} 进行变异, 操作如下:

$$a_{ij} = \begin{cases} a_{ij} + (a_{ij} - a_{\max}) \times f(g), & r \geq 0.5 \\ a_{ij} + (a_{\min} - a_{ij}) \times f(g), & r < 0.5 \end{cases} \quad (7)$$

式中: a_{\max} 为基因 a_{ij} 的上界; a_{\min} 为基因 a_{ij} 的下界;

$f(g) = r_2(1 - \frac{g}{G})$; r_2 为随机数; g 为当前迭代次数; G 为最大进化代数; r 为 $[0, 1]$ 区间的随机数。

7) 对网络初始权值和阈值赋值, 在子代中选出最优的个体解码后确定最优的参数的组合, 作为神经网络训练参数。

算法实现的流程如图 1 所示。

3 建立制浆能耗模型

3.1 辅助变量的选取

为对现场控制优化的指导作用, 所建模型应正确反映生产过程中的控制量与耗电量的函数关系。通过对制浆系统的机理分析, 给料量与给水流量 (分为稀释水和研磨水) 决定了浆液产量; 再循环箱浆液密度影响成品浆液的密度; 循环泵频率及旋

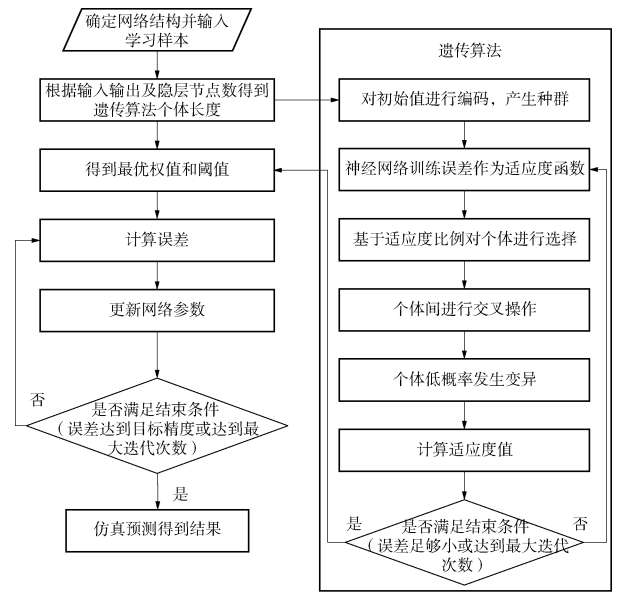


图 1 GA-BP 算法流程图

Fig. 1 GA-BP algorithm flow chart

流器压力控制浆液的分离效果; 而球磨机电流是系统中主要耗电设备。因此选取上述 7 个影响制浆过程中物耗和能耗的可控制量作为模型的输入变量。

3.2 数据预处理

数据从电厂 DCS 中导出, 采样间隔为 2 s, 为保证模型能够准确反映控制参数和制浆能耗间的关系, 样本数据覆盖了给料量 18~28 t 间的各类运行工况。此外, 从现场采集的数据含有误差及噪声, 可能导致模型性能下降或建模失败。因此首先对数据进行预处理, 具体步骤为:

1) 使用插值法将 DCS 数据中缺失的采样时间点补充。

2) 剔除各变量中超出合理工况范围的数据异常值。

3) 对波动过于剧烈的数据进行平滑降噪处理。

4) 对数据进行归一化处理, 使其转换为 $(-1, 1)$ 区间的无量纲数。

5) 采用互信息法求得辅助变量与主导变量间的时间延迟。各辅助变量与耗电量间的最大互信息值及延迟时间如表 1 所示。

3.3 构建 GA-BP 模型

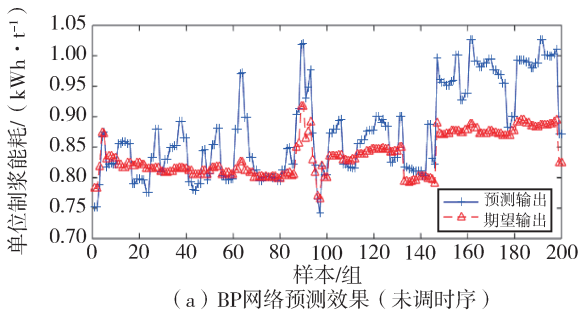
模型采用单位制浆能耗, 即制浆量/耗电量作为输出, 神经网络输入节点数为 7, 输出节点数为 1, 根据经验公式选择隐含层节点数为 4; GA 算法初始种群规模为 50, 最大迭代次数为 20 代, 交叉

概率和变异概率分别设置为 0.4 和 0.1。

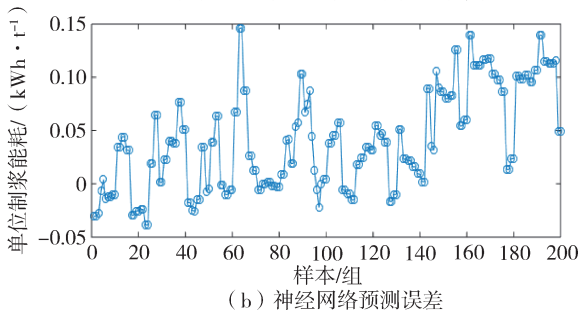
表 1 辅助变量与耗电量互信息值及时延表
Tab. 1 Auxiliary variable and power consumption mutual information value and time delay

辅助变量	MI 最大值	时延/s
给料量	0.004 2	275
稀释水流量	0.127 5	247
研磨水流量	0.113 8	283
再循环箱浆液密度	0.001 1	244
循环泵频率	0.003 6	0
旋流器出口压力	0.002 9	261
球磨机电流	0.012 8	236

模型使用经过预处理的 2 700 组数据进行仿真实验, 其中 2 500 组进行训练, 200 组用于测试, 使用均方根误差 (RMSE) 和平均相对误差 (MRE) 作为性能评价指标, 并使用标准 BP、GA-BP 与经过 MI 时序处理的 GA-BP 模型效果进行对比。各模型预测效果及预测误差如图 2~图 4 所示。



(a) BP网络预测效果 (未调时序)

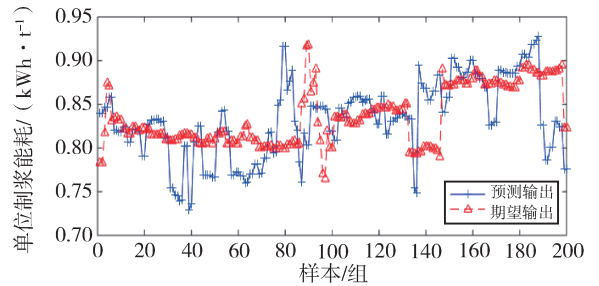


(b) 神经网络预测误差

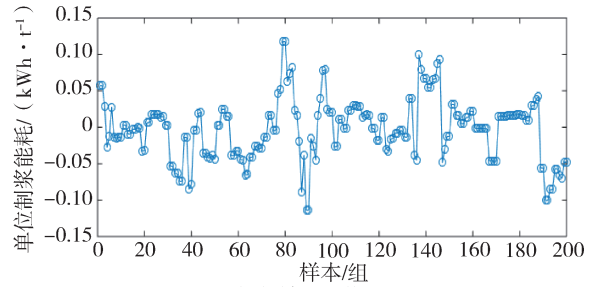
图 2 BP 模型 (未调时序) 预测结果

Fig. 2 BP model (unregulated timing) prediction results

由图 3 可见, 未经时序调整的 GA-BP 模型预测值与实际值存在约 30 s 的时间延迟, 导致预测误差偏大, 误差最大值达到 0.12 kw · h/t, 而根据表 1 所列延迟时间对模型输入数据进行调整后得到如图 4 所示的预测结果, 预测值和实际值的偏差明显减小, 最大误差值仅为 0.047 kw · h/t, 最大相对



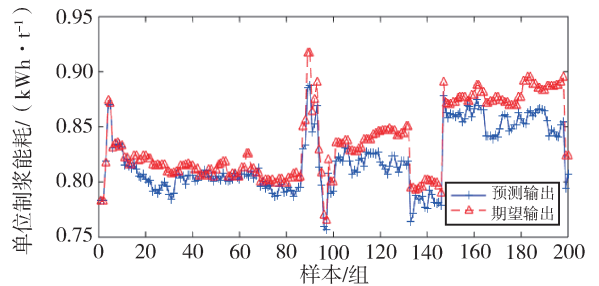
(a) GA-BP网络预测效果 (未调时序)



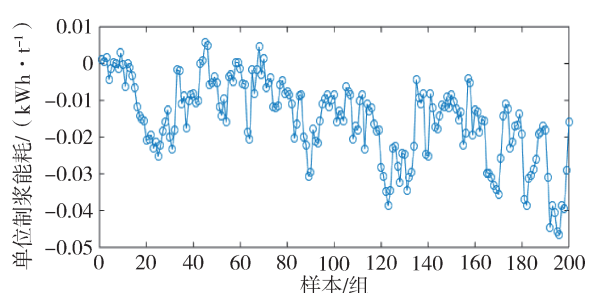
(b) 神经网络预测误差

图 3 GA-BP 模型 (未调时序) 预测结果

Fig. 3 GA-BP model (unregulated timing) prediction results



(a) GA-BP网络预测效果 (调整时序)



(b) 神经网络预测误差

图 4 GA-BP 模型 (调整时序) 预测结果

Fig. 4 GA-BP model (adjusted timing) prediction results

误差由 13% 降至 5%, 平均相对误差由 5% 降至 1.5%, 说明经过互信息处理后的输入变量能够更好的反映实际生产情况, 减少变量间时延造成的预测误差。

模型性能指标对比如表 2 所示:

表2 预测性能对比

Tab. 2 Comparison of predicted performance

预测样本数	RMSE/(kWh · t ⁻¹)	MRE/%	迭代步数
BP(未调时序)	0.062 4	5.19	278
GA-BP(未调时序)	0.042 0	3.92	126
GA-BP(调整时序)	0.018 8	1.55	133

从表2及仿真结果可以看出,经过MI时序处理后的GA-BP模型具有较高的预测精度和泛化能力,预测值与实际值的均方根误差及平均相对误差在2%以内,本文提出的方法具有一定的实用性。

4 结论

本文设计了基于MI时序处理的GA-BP神经网络脱硫脱硝浆系统能耗模型,该模型利用遗传算法出色的全局搜索能力对BP网络的结构进行优化,改善了传统BP神经网络收敛速度慢,易陷入局部极小的缺点,提高了预测精度。在模型建立过程中,本文使用互信息计算了变量间相关性,对辅助变量的时序进行了调整,解决了生产过程中的延迟问题,进一步提高了模型的准确度。通过现场数据建模实验得到以下结论:

1)变量间的时序关系对预测模型的准确度影响较大,通过互信息对时序进行调整能够有效减小误差。

2)GA-BP算法具有较快的迭代速度和全局搜索能力,能够有效克服BP算法收敛速度慢和训练时间长的缺点。

3)基于MI时序处理的GA-BP神经网络模型在各项性能指标上均有较好表现,计算速度快,预测精度高,所设计的能耗模型能够为电厂制浆系统的运行优化提供指导。

参考文献:

- [1] 吕新锋. 石灰石-石膏湿法烟气脱硫设施常见故障及影响脱硫效率因素分析[J]. 电力科技与环保, 2018, 34(2): 27-29.
- [2] HAO R, ZHANG Y, WANG Z, et al. An advanced wet method for simultaneous removal of SO₂ and NO from coal-fired flue gas by utilizing a complex absorbent [J]. Chemical Engineering Journal, 2017, 307: 562-571.
- [3] 李伟科. 火电厂排放烟气白雾治理新方案及节能分析[J]. 南方能源建设, 2018, (4): 50-54.
- [4] 杨伟, 倪黔东, 吴军基. BP神经网络权值初始值与收敛性问题研究[J]. 电力系统及其自动化学报, 2002(1): 20-22.

- [5] 史坤鹏, 乔颖, 赵伟, 等. 计及历史数据熵关联信息挖掘的短期风电功率预测[J]. 电力系统自动化, 2017, 41(3): 13-18.
- [6] SHANNON C. A mathematical theory of communication [J]. Bell System Tech. J., 1948, 27(4): 379-423.
- [7] WANG Z C, LI M Q, LI J Z. A multi-objective evolutionary algorithm for feature selection based on mutual information with a new redundancy measure [J]. Information Sciences, 2015, 307(1): 73-78.
- [8] 赵彦涛, 单泽宇, 常跃进, 等. 基于MI-LSSVM的水泥生料细度软测量建模[J]. 仪器仪表学报, 2017, 38(2): 487-496.
- [9] 刘吉臻, 秦天牧, 杨婷婷, 等. 基于偏互信息的变量选择方法及其在火电厂SCR系统建模中的应用[J]. 中国电机工程学报, 2016, 36(9): 2438-2443.
- [10] 马平, 李珍, 梁薇. 基于互信息的辅助变量筛选及在火电厂NO_x软测量模型中的应用[J]. 科学技术与工程, 2017, 17(22): 249-254.
- [11] CELLUCCI C J, ALBANO A M, RAPP P E. Statistical validation of mutual information calculations: comparison of alternative numerical algorithms [J]. Physical Review E, 2005, 71(6): 066208.
- [12] 冯长敏, 张炳江. 基于BP神经网络的分段函数连续优化处理[J]. 北京信息科技大学学报(自然科学版), 2019, 34(1): 18-22.
- [13] SHAN Y, YI J L, F J G. The Application of BP neural network algorithm in optical fiber fault diagnosis [C]// IEEE International Symposium on Distributed Computing & Applications for Business Engineering & Science, Guiyang, 2015. Guiyang: IEEE, 2015.
- [14] 陈浩广, 王银河. 基于扩展BP神经网络的一类非线性系统自适应控制设计[J]. 计算机应用, 2017, 37(6): 1670-1673.
- [15] 张铃, 张钊. 遗传算法机理的研究[J]. 软件学报, 2000, 11(7): 945-952.
- [16] 郭小龙, 郭兵, 沈艳. 改进BP神经网络在软件能耗分析中的应用[J]. 计算机工程, 2012, 38(10): 185-187.

作者简介:



JIN X Z

金秀章

1969-, 男, 河北衡水人, 博士, 华北电力大学控制与计算机工程学院副教授, 硕士生导师。近期主要研究兴趣: 先进控制策略在大型电力机组的控制研究、信息融合技术 (e-mail) jinxsyzs@163.com。

李奕颖(通信作者)

1995-, 女, 吉林吉林人, 华北电力大学硕士生, 主要从事机器学习算法研究及火电厂控制策略研究 (e-mail) 15733220440@163.com。

(责任编辑 郑文棠)