

改进ELMAN网络的Q学习温度切换控制

乔景慧^{1†}, 柴天佑²

(1. 沈阳工业大学 机械工程学院, 辽宁 沈阳 110870; 2. 东北大学 流程工业综合自动化国家重点实验室, 辽宁 沈阳 110819)

摘要: 在水泥生料分解过程中, 由于生料边界条件频繁变化且生料流量波动大, 使过程处于正常工况或异常工况, 致使分解炉温度和预热器C5出口温度很难控制在工艺要求的范围内. 传统的控制方法经常导致预热器C5下料管堵塞. 为了解决上述问题, 本文提出了一个带有前馈补偿的温度智能切换控制策略, 由基于T-S的模糊控制器、基于改进的ELMAN网络的Q学习异常工况控制器和切换机制组成. 实际应用结果表明, 所提出的控制策略能够根据当前工况的变化选择正确的控制器, 并且使生产远离故障工况.

关键词: 生料分解过程; 分解炉温度; 改进ELMAN网络; Q学习控制; 切换控制

中图分类号: TP18 文献标识码: A

Temperature switching control integrated with improved ELMAN network and Q learning

QIAO Jing-hui^{1†}, CHAI Tian-you²

(1. School of Mechanical Engineering, Shenyang University of Technology, Shenyang Liaoning 110870, China;

2. State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang Liaoning 110819, China)

Abstract: Because of the frequent change of the boundary conditions as well as the large variation in the raw meal flow in calcination process, the process is often switching between the normal and abnormal working conditions. This causes difficulties to maintain the calciner temperature and the outlet temperature of preheater C5 (i.e., the No. 5 preheater) within their desired range of operations. Conventional control may even lead to operational failure, such as the clogging of feeding tube C5. To deal with this problem, we develop an intelligent temperature control scheme with feedforward compensation, which consists of a T-S-based fuzzy controller, an abnormal condition controller, an improved ELMAN network with Q learning, and a switching mechanism. The application results show that the proposed approach can select a right controller based on the current conditions to maintain the operation far from the faulty operation conditions.

Key words: raw meal calcination process; calciner temperature; improved ELMAN network; Q learning control; switching control

1 引言(Introduction)

生料分解在整个水泥熟料生产过程中起着至关重要的作用, 它是将生料经过分解转变成合格煅烧熟料的过程, 其中分解炉温度决定产品的质量指标生料分解率, 预热器C5出口温度是过程的安全指标. 由于生料边界条件频繁变化及生料流量波动大, 使过程经常处于正常工况或异常工况, 使得分解炉温度和预热器C5出口温度很难控制在工艺要求的范围内.

对生料分解过程的研究, 大多数文献均集中研究反应机理及仿真方面, 对分解炉温度控制研究很少. 文[1]为了优化分解炉的设计, 建立了一个碳粒燃烧和石灰石分解非等温的扩散动态数学模型. 文[2]建立了

一个分解率的仿真实验模型, 通过调整分解炉给煤量控制生料分解率. 文[3]使用商业化有限元软件, 分析了水泥分解反应的机理并验证. 文[4]对分解炉内气固两相流的流动特性进行了模拟. 文[5]对一个1.7 MW的电厂分解炉进行研究, 建立了两个模型: 一个是3-D分解炉模型, 另一个是1-D过程模型.

上述生料分解过程生料流量、三次风温度和窑尾温度变化小, 煤粉质量(细度)好, 因此采用上述控制策略能够使分解炉温度跟踪设定值, 预热器C5出口温度低于最大值. 从而保证产品质量指标即生料分解率在工艺要求的范围内. 但是, 实际生产中生料边界条件变化频繁, 生料流量、三次风温度和窑尾温度经常发

收稿日期: 2014-07-10; 录用日期: 2015-01-27.

[†]通信作者. E-mail: qiaojh2002@163.com; Tel.: +86 24-25496271.

中国博士后科学基金项目(2014M561249, 2015T80268), 国家自然科学基金面上项目(61273177)资助.

Supported by China Postdoctoral Science Foundation (2014M561249, 2015T80268) and National Nature Science Foundation of China (61273177).

生变化,致使生料分解过程经常处于异常工况,导致预热器C5下料口堵塞. 为了克服以上缺点,本文提出了一个带有前馈补偿的温度智能切换控制方法,由基于T-S的模糊控制器、基于改进的ELMAN网络的Q学习异常工况控制器和切换机制组成. 所提方法能根据当前工况的变化选择正确的控制器并使分解炉温度跟踪设定值,同时预热器C5出口温度低于最大值.

2 生料分解过程问题描述及动态特性分析 (Problem description of raw meal calcination process and dynamic characteristics analysis)

2.1 生料分解过程问题描述(Problem description of raw meal calcination process)

生料分解过程工艺流程如图1所示. 系统具有5级预热器、1台分解炉和1台回转窑. 在高温风机和排风机的作用下,来自回转窑和分解炉的废气经过旋风筒

C5到达增湿塔. 来自均化库的生料经过C1旋风筒与分解炉和窑尾废气进行热交换. 然后,生料继续下行进入下一级旋风筒,这样热交换重复进行. 生料在下降过程中被逆向前进的热气流加热,煤粉和辅助风混合成气固两相流与三次风一起喷入分解炉,煤粉燃烧的热量使生料在分解炉内的分解率达到0.85%~0.94%. 同时窑头煤粉与来自冷却风机的二次风一起燃烧后的烟气进入分解炉内.

图1中相关变量及符号解释如下: M 为电机, C_i 为第*i*个预热器, $F(t)$ 为生料流量, TT 为温度传感器, C 为控制器, T_{csp} 为分解炉温度设定值, T_{c5max} 为预热器C5温度最大值, F_{ref} 为生料流量参考值, $T_c(t)$ 和 $T_{c5}(t)$ 分别为分解炉温度和预热器C5温度的反馈值, $\Delta u_{c1}(t)$, $\Delta u_{c2}(t)$ 和 $\Delta u_f(t)$ 分别为控制器 TC_1 , TC_2 和 FC 的输出值, $\Delta u_c(t)$ 为 t 时刻 $\Delta u_{c1}(t)$ 或 $\Delta u_{c2}(t)$ 的值, 而 $u(t)$ 为 $\Delta u_c(t)$ 与 $\Delta u_f(t)$ 的和.

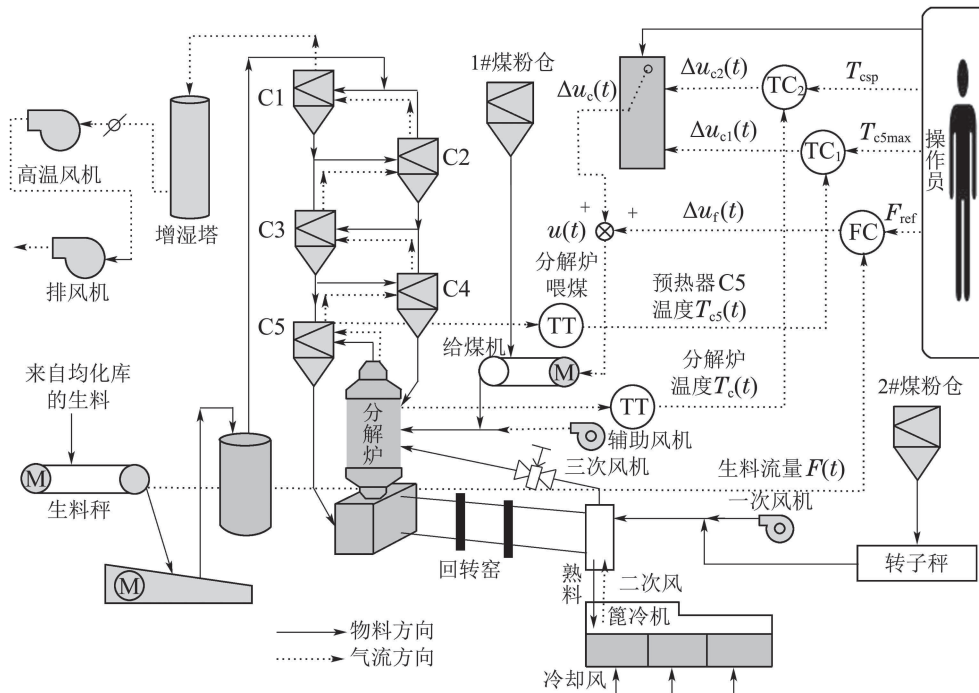


图1 生料分解过程工艺流程及控制现状

Fig. 1 Process flow diagram and current control of the raw meal calcination process

长期以来,生料分解过程温度主要由操作员通过经验调整给煤量,使生产逐渐远离故障工况,由于人工诊断和判断不及时,常常导致故障工况的发生,严重影响生料分解率,甚至导致预热器C5温度高于上限而停产.

生料分解过程的控制目标是通过调整分解炉给煤量 $u(t)$, 使分解炉温度 $T_c(t)$ 跟踪设定值 $T_{csp}(t)$, 同时保证预热器C5出口温度 $T_{c5}(t)$ 低于上限 T_{c5max} .

2.2 生料分解过程动态特性分析(Dynamic characteristics analysis of raw meal calcination process)

根据文[6-8]可知,分解炉给煤量 $u(t)$ 与分解炉温度 $T_c(t)$ 和预热器C5出口温度 $T_{c5}(t)$ 之间的动态模型:

$$\dot{T}_c(t) = (1/C_{p,g}m_g)\{r_{combg}H_{combg} - \sigma C_{p,c}m_c(t)T_c(t) +$$

$$\begin{aligned}
 & 8.1 \times 10^9 \pi r_c^2 \int_0^\theta \exp\left(-\frac{E_c}{RT_c(t)}\right) d\theta - \\
 & C_{p,r} m_r(t) T_c(t) + C_{p,g} m_{g,in} T_{g,in} - \\
 & 6.3 \times 10^8 k_{s,r} \pi r_r^2 \int_0^\psi \exp\left(-\frac{E_r}{RT_c(t)}\right) d\psi - \\
 & C_{p,g}(v_T(t)) \sqrt{\frac{43}{f_T T_{g,T}}} + m_{g,s} + 0.28 F(t) + \\
 & 0.64 m_{c,k} + m_{g,p} + 0.18 u(t) T_c(t)\}, \quad (1) \\
 & T_{c5}(t) = k_0 \cdot (1 - \sigma) \cdot m_c(t) \cdot f_1(T_c(t), r_c), \quad (2) \\
 & m_c(t) = m_c(t_0) + \int_{t_0}^t \left(\frac{u(t)}{\alpha_1} - \frac{C_{p,r} m_r(t)}{C_{p,c}}\right) dt, \quad (3) \\
 & m_r(t) = m_r(t_0) + \int_{t_0}^t \left(\frac{F(t) - G_m(t)}{\alpha_2}\right) dt, \quad (4) \\
 & G_m(t) = \frac{G_{max} m_c(t) (2 - m_c(t)) f_2(T_c(t), r_c)}{m_{c,max}^2}, \quad (5)
 \end{aligned}$$

其中: k_0 是比例系数, 3次风阀 $v_T(t)$ 保持最大开度(即100%), $f_1(T_c(t), r_c)$ 和 $f_2(T_c(t), r_c)$ 是两个非线性函数. 式(1)–(5)中各符号的含义如表1所示.

表 1 分解炉温度模型各变量的含义

Table 1 The meaning of variables in calciner temperature model

变量	含义	变量	含义
$C_{g,p}$	气体比热	θ	煤粉停留时间
$C_{p,c}$	煤粉比热	ψ	生料停留时间
$C_{p,r}$	生料比热	$k_{s,r}$	分解常数
m_g	气体质量	R	气体常数
$m_c(t)$	煤粉质量	E_c	活化能
$m_r(t)$	生料质量	E_r	分解活化能
$m_{g,s}$	二次风质量	r_c	碳粒子半径
$m_{c,k}$	入窑煤粉质量	r_r	生料粒子半径
$m_{g,p}$	一次风质量	r_{combg}	挥发燃烧速率
$m_{g,in}$	进入气体质量	H_{combg}	气相燃烧热
$F(t)$	生料质量	$G_m(t)$	台时产能
$v_T(t)$	三次风阀开度	G_{max}	产能最大值
f_T	粘滞系数	$T_{g,in}$	气体初始温度
σ	煤粉燃烧速率	$T_{g,T}$	三次风温度

式(1)–(5)表明, 分解炉温度 $T_c(t)$ 与分解炉给煤量 $u(t)$ 之间具有非线性关系. 本文采用gap metric方法^[9-10]分析系统非线性的强弱, 将运行空间 Ω 分成 N_g 单元, 定义式(6)的非线性度量:

$$v = \max_{i,j=1,\dots,N_g} \{\delta(L_i, L_j)\}, \quad (6)$$

其中 L_i 和 L_j 分别代表运行空间 Ω 的第 i 个和第 j 个非线性运行点. 定义运行点

$$N_g = 150, T_c \in [850, 890], u \in [2.1, 6.5],$$

这样分解炉温度 $T_c(t)$ 和给煤量 $u(t)$ 之间的非线性度

量值等于1.

图2表明对于任意的 $\gamma < 1$, 都有 $v > \gamma$. 这样分解炉温度 $T_c(t)$ 与给煤量 $u(t)$ 之间具有强非线性. 文献[11]表明一个单一的线性控制器不能使系统稳定. 因此, 需要采用切换控制器.

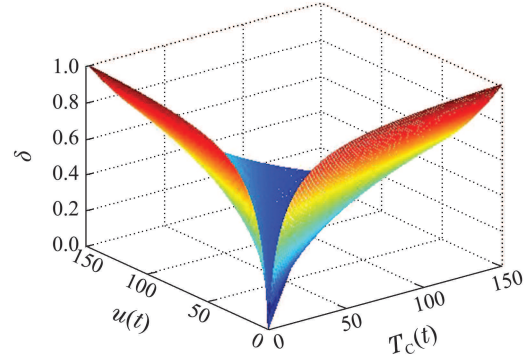


图 2 分解炉温度 $T_c(t)$ 和给煤量 $u(t)$ 之间的gap metric
Fig. 2 Gap metric between calciner temperature $T_c(t)$ and coal feeder $u(t)$

因此, 采用已有的控制方法难以通过调整分解炉给煤量 $u(t)$ 将分解炉温度 $T_c(t)$ 控制在设定值范围内, 同时预热器C5出口温度低于上限 $T_{c5,max}$, 满足

$$\begin{cases} T_{c \min} \leq T_c(t) \leq T_{c \max}, \\ T_{c5}(t) \leq T_{c5 \max}. \end{cases} \quad (7)$$

3 生料分解过程温度智能切换控制方法(Intelligence-based temperature switching control method for raw meal calcination process)

3.1 生料分解过程控制策略(Control strategy for raw meal calcination process)

为了实现式(7)的控制问题且降低生料分解过程故障工况发生的概率, 本文提出了一个带有前馈补偿的温度智能切换控制方法, 如图3所示. 图3中, $e_c(t)$ 和 $e_{c5}(t)$ 分别为分解炉温度设定和预热器C5出口温度设定与各自的反馈值之差, 图3中其他符号的含义见表2.

表 2 图3中各变量的含义

Table 2 The meaning of variables in Fig. 3

变量	含义	变量	含义
$T_c(t)$	分解炉温度	$\Delta u_{c1}(t)$	异常工况控制器输出
$T_{c5}(t)$	C5出口温度	$\Delta u_{c2}(t)$	T-S模糊控制器输出
T_{csp}	分解炉温度设定	$\Delta u_f(t)$	前馈控制器输出增量
$T_{c5 \max}$	C5温度最大值	$\Delta u_c(t)$	切换控制器输出增量
$F(t)$	生料流量	$u(t)$	分解炉给煤量
F_{ref}	生料流量参考值	$\Delta F(t)$	生料流量与参考值差

如图3所示, 生料分解过程温度智能切换控制器由切换机制、T-S模糊控制器、改进ELMAN网络

的Q学习异常工况控制器和前馈控制器组成. 其中切换机制、T-S模糊控制器和前馈控制器的功能和

控制算法采用文献[8]所述的方法, 本文详细介绍改进ELMAN网络的Q学习异常工况控制器.

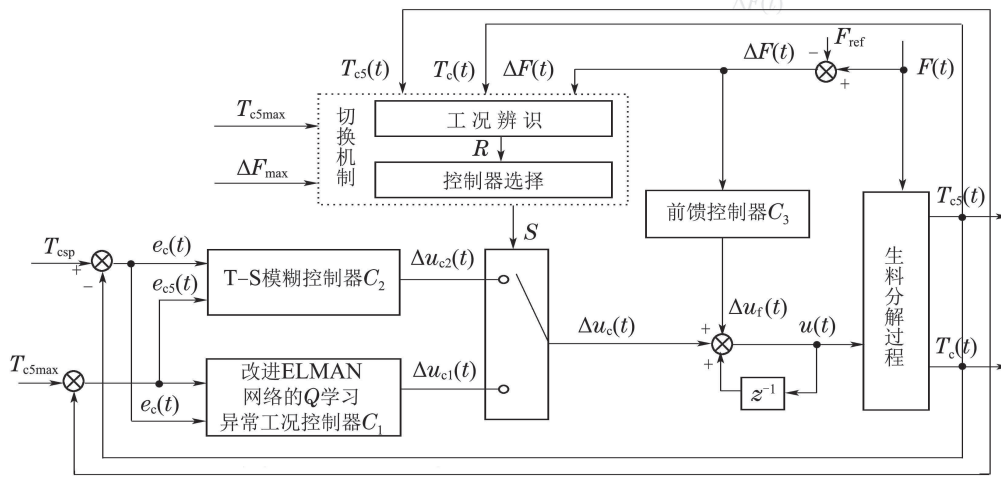


图3 生料分解过程温度智能切换控制结构

Fig. 3 The proposed intelligence-based temperature switching control for raw meal calcination process

3.2 改进ELMAN网络的Q学习异常工况控制器 (An abnormal condition controller integrated with improved ELMAN network and Q learning)

生料分解过程中, 当预热器C5出口温度大于高限时, 表明生料分解过程处于异常工况, 即切换机制中C5出口温度 $T_{c5}(t) \geq T_{c5max}$. 当系统处于异常工况时, 操作员根据C5出口温度及分解炉温度调整分解炉给煤量, 存在状态和动作的一一对应关系. 在实际生产中, 由于煤粉细度过粗或生料流量波动大, 经常出现预热器C5出口温度高于上限, 致使C5下料管堵塞现象发生. 针对所出现的异常工况, 操作员根据系统当前状态, 采取一系列动作, 观察其效果, 再学习的控制策略. ELMAN网络是一种带有局部

反馈的神经网络, 它可以很好的实现非线性系统的逼近^[12], 它除了具有输入层、隐含层、输出层外, 还有一个特殊的单元-关联层. 其结构单元与通常的多层前馈网络相同, 输入层单元仅起信号传输的作用, 输出层单元起线性加权作用, 隐含层单元可由线性或非线性激发函数进行计算, 关联层则用来记忆隐层单元前一时刻的输出值, 可认为是一个一步时延算子, 它使该网络具有动态的记忆功能. 这里的前馈连接部分可进行连接权修正, 而递归部分是固定的, 即不能进行学习修正, 从而ELMAN网络仅是部分递归的. 为了更好地提高递归结构的作用, 本文首先提出将输出单元与隐层进行全反馈连接.

因此, 本文提出一种改进的ELMAN网络来进行Q值函数学习控制, 如图4所示.

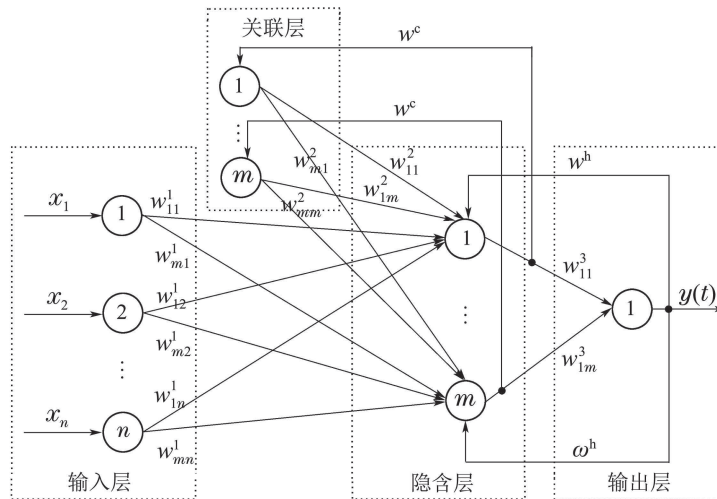


图4 改进ELMAN网络与Q学习网络结构

Fig. 4 Network structure integrated with improved ELMAN network and Q learning

图4中,网络的输入为

$$x = [x_1 \ x_2 \ \cdots \ x_n]^T,$$

$x_c(t)$ 和 $x_h(t)$ 分别表示关联层和隐含层的输出向量,且均为 m 维. $y(t)$ 表示输出层的输出, $w_{ji}^1 (i=1, \dots, n, j=1, \dots, m)$ 为输入层与隐含层的连接权, $w_{jl}^2 (j=1, \dots, m, l=1, \dots, m)$ 为关联层与隐含层的连接权, $w_{1j}^3 (j=1, \dots, m)$ 为隐含层与输出层的连接权, w^c 为隐含层输出与关联层的权值($w^c = 1$), w^h 为输出层输出与隐含层的权值($w^h = 1$).隐含层的激活函数 $f(x)$ 为tansig函数,如式(8)所示.输出层 $g(x)$ 为purelin线性函数.

$$f(x) = \frac{2}{1 + e^{-2x}} - 1. \quad (8)$$

因此,各层之间的表达式如下:

$$Q(s_t, a_t) = g(w^3 x_h(t)), \quad (9)$$

$$x_h(t) = f(w^1 x(t-1) + w^2 x_c(t) + I_{m \times 1} y(t-1)), \quad (10)$$

$$x_c(t) = x_h(t-1), \quad (11)$$

其中: $I_{m \times 1}$ 为单位矩阵, w^1, w^2 和 w^3 分别为输入层与隐含层的连接权矩阵,关联层与隐含层的连接权矩阵,隐含层与输出层的连接权矩阵分别如式(12)–(14)所示:

$$w^1 = \begin{bmatrix} w_{11}^1 & w_{12}^1 & \cdots & w_{1n}^1 \\ \vdots & \vdots & \dots & \vdots \\ w_{m1}^1 & w_{m2}^1 & \cdots & w_{mn}^1 \end{bmatrix}, \quad (12)$$

$$w^2 = \begin{bmatrix} w_{11}^2 & w_{12}^2 & \cdots & w_{1m}^2 \\ \vdots & \vdots & \dots & \vdots \\ w_{m1}^2 & w_{m2}^2 & \cdots & w_{mm}^2 \end{bmatrix}, \quad (13)$$

$$w^3 = [w_{11}^3 \ w_{12}^3 \ \cdots \ w_{1m}^3]. \quad (14)$$

由于输出层 $g(x)$ 为purelin线性函数,所以式(9)可以写为

$$Q(s_t, a_t) = w^3 x_h(t). \quad (15)$$

异常工况控制器的输入为 $e_c(t)$ 和 $e_{c5}(t)$,输出为分解炉给煤量增量 $\Delta u_{c1}(t)$.由于学习控制是状态和动作的一一对应,因此,图4中的 x 写为 $x = [e_c(t) \ e_{c5}(t) \ \Delta u_{c1}(t)]^T$,其中: $e_c(t)$ 和 $e_{c5}(t)$ 为学习的状态量, $\Delta u_{c1}(t)$ 为学习的一组动作.为了描述方便,向量 x 表述为 $x = [s_t \ a_t]^T \in \mathbb{R}^3$,表示在时刻 t 时的状态-动作对,其中 $s_t = [s_1 \ s_2]^T \in \mathbb{R}^2$ 表示二维系统状态变量, $a_t = a_1 \in \mathbb{R}$ 表示一维动作向量,网络输出由1个节点组成,且描述为 $Q(s_t, a_t) \in \mathbb{R}$.

改进ELMAN网络的Q学习网络的输出为

$$Q(s_t, a_t) = \sum_{j=1}^m w_{1j}^3 x_{hj}(t) = \sum_{j=1}^m w_{1j}^3 f\left(\sum_{i=1}^n w_{ji}^1 x_i(t-1) + \sum_{l=1}^m w_{jl}^2 x_{hl}(t-1) + I_{j \times 1} w_{1j}^3 x_{hj}(t-1)\right), \quad (16)$$

其中: w_{ji}^1, w_{jl}^2 和 w_{1j}^3 分别表示输入层与隐含层、隐含层与连接层以及隐含层与输出层之间的连接权值; $x_{hl}(t-1)$ 和 $x_{hj}(t-1)$ 分别为隐含层第 l 个和第 j 个神经元 $t-1$ 时刻的输出,隐含层节点数为 m 个, $n=3$.

针对式(16)的值,当系统处于状态 s_t 时,如何选择动作 a_k 是能否获得最优策略的关键.因此,采用近似贪心且连续可微的Boltzmann-Gibbs分布作为动作选择策略,则动作 a_k 被选择的概率为

$$P(a_t = a_k | s_t) = \frac{\exp(Q(s_t, a_k)/T_t)}{\sum_a \exp(Q(s_t, a)/T_t)}, \quad (17)$$

其中 $T_t > 0$ 为温度参数.为了提高学习的速度,采用式(18)进行动态调整温度值^[13],即在学习的初期选择较大的温度,以保证动作选择的随机性较大,增加搜索能力,在学习的过程中逐渐降低温度,保证以前的学习效果不被破坏.

$$\begin{cases} T_0 = T_{\max}, \\ T_{t+1} = T_{\min} + \xi(T_t - T_{\min}), \end{cases} \quad (18)$$

其中: $0 \leq \xi < 1$ 为退火因子, T_{\min} 和 T_{\max} 分别为温度参数的最小值和最大值.

采用动作集中概率大的动作 a_k 作用于状态 s_t 上,必须知道ELMAN网络的权值才能根据式(16)计算 $Q(s_t, a_t)$ 值.因此本文采用最速下降梯度法^[14]对ELMAN网络的权值(包括输入层与隐含层之间的连接权值 $w_{ji}^1 (i=1, \dots, n; j=1, \dots, m)$ 、关联层与隐含层之间的连接权值 $w_{jl}^2 (j=1, \dots, m; l=1, \dots, m)$ 以及隐含层与输出层之间的连接权值 $w_{1j}^3 (j=1, \dots, m)$)进行学习.学习算法的学习过程是循环地减小对相邻状态-动作的值估计之间的差异,在这个意义上,学习是更广泛的时间差分(temporal difference, TD)算法中的特例^[15],定义TD误差如式(19)所示^[16-17]:

$$\delta_{TD}(t) = r_t + \gamma Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t), \quad (19)$$

其中: $\delta_{TD}(t)$ 表示TD误差, r_t 为立即回报, $0 \leq \gamma < 1$ 为折扣因子,它确定了延迟回报与立即回报的相对比例.

采用如式(20)所示的TD误差的平方和作为网络权值更新的规则。

$$E(t) = \frac{1}{2} \delta_{TD}^2(t). \quad (20)$$

下面以输入层与隐含层之间连接权值 w_{ji}^1 ($i = 1, \dots, n; j = 1, \dots, m$)更新为例, 则输入层与隐含层之间的连接权值 w_{ji}^1 ($i = 1, \dots, n; j = 1, \dots, m$)表示为

$$w_{ji}^1(t+1) = w_{ji}^1(t) - \Delta w_{ji}^1(t), \quad (21)$$

其中 $\Delta w_{ji}^1(t)$ 为第 i 个输入对隐含层第 j 个神经网络权值的更新, 如式(22)所示:

$$\Delta w_{ji}^1(t) = -\alpha g_{ji}^1(t), \quad (22)$$

其中: α 为网络权值的学习速率, $g_{ji}^1(t)$ 为梯度向量, 如式(23):

$$g_{ji}^1(t) = \frac{\partial E(t)}{\partial w_{ji}^1(t)}. \quad (23)$$

根据式(19)–(20)和(23), 式(22)可以表达为

$$\begin{aligned} \Delta w_{ji}^1(t) &= -\alpha \frac{\partial E(t)}{\partial w_{ji}^1(t)} = \\ &\alpha [r_t + \gamma Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)] \frac{\partial Q(s_t, a_t)}{\partial w_{ji}^1(t)}, \end{aligned} \quad (24)$$

其中 $\frac{\partial Q(s_t, a_t)}{\partial w_{ji}^1(t)}$ 可由式(16)递推求得。

同理, 可以求解出关联层与隐含层之间的连接权值的更新 Δw_{jl}^2 ($j = 1, \dots, m; l = 1, \dots, m$)以及隐含层与输出层之间的连接权值的更新 Δw_{1j}^3 ($j = 1, \dots, m$), 分别如式(25)–(26)所示:

$$\begin{aligned} \Delta w_{jl}^2(t) &= -\alpha \frac{\partial E(t)}{\partial w_{jl}^2(t)} = \\ &\alpha \delta_{TD}(t) \frac{\partial Q(s_t, a_t)}{\partial w_{jl}^2(t)} = \\ &\alpha [r_t + \gamma Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)] \frac{\partial Q(s_t, a_t)}{\partial w_{jl}^2(t)}, \end{aligned} \quad (25)$$

$$\begin{aligned} \Delta w_{1j}^3 &= -\alpha \frac{\partial E(t)}{\partial w_{1j}^3} = \\ &\alpha \delta_{TD}(t) \frac{\partial Q(s_t, a_t)}{\partial w_{1j}^3} = \\ &\alpha [r_t + \gamma Q(s_{t+1}, a_{t+1}) - Q(s_t, a_t)] \frac{\partial Q(s_t, a_t)}{\partial w_{1j}^3}, \end{aligned} \quad (26)$$

其中: $\frac{\partial Q(s_t, a_t)}{\partial w_{jl}^2(t)}$ 和 $\frac{\partial Q(s_t, a_t)}{\partial w_{1j}^3}$ 可由式(16)递推求得。

则关联层与隐含层之间的连接权值 $w_{jl}^2(t)$ ($j = 1, \dots, m; l = 1, \dots, m$)和隐含层与输出层之间的连接权值 w_{1j}^3 ($j = 1, \dots, m$)如式(27)–(28)所示:

$$w_{jl}^2(t+1) = w_{jl}^2(t) - \Delta w_{jl}^2(t), \quad (27)$$

$$w_{1j}^3(t+1) = w_{1j}^3(t) - \Delta w_{1j}^3(t), \quad (28)$$

其中 $\Delta w_{jl}^2(t)$ 和 Δw_{1j}^3 如式(25)–(26)所示。

下面详细介绍改进Elam网络的Q学习异常工况控制器算法步骤:

- a) 初始化系统的状态和学习控制器参数, 包括网络权值 $w^1(0)$, $w^2(0)$ 和 $w^3(0)$ 、学习率 α 、折扣因子 γ 、退火因子 ξ 、温度参数 T_{\min} 和 T_{\max} ;
- b) 检测当前状态 s_t ;
- c) 将状态–动作对 (s_t, a) 作为ELMAN网络的输入, 由式(16)计算网络的输出 $Q(s_t, a)$, $\forall a \in A$;
- d) 根据式(17)计算概率分布, 选择一个动作 a_t 作用于系统;

- e) 对被选择的动作 a_t 根据式(16)计算

$$\frac{\partial Q(s_t, a_t)}{\partial w_{ji}^1(t)}, \frac{\partial Q(s_t, a_t)}{\partial w_{jl}^2(t)}, \frac{\partial Q(s_t, a_t)}{\partial w_{1j}^3},$$

观察新状态 s_{t+1} 和立即回报 r_t ;

- f) 应用 (s_{t+1}, a) 作为ELMAN网络的输入, 由式(16)计算网络的输出 $Q(s_{t+1}, a)$, $\forall a \in A$;

- g) 根据式(17)计算概率分布, 选择下一动作 a_{t+1} 作用于系统;

- h) 根据式(24)–(26)更新网络的权值;

- i) 若满足学习结束条件, 结束; 否则, $t \leftarrow t + 1$, 转到b)。

4 工业应用(Industrial applications)

本文所提出的改进ELMAN网络的Q学习温度切换控制方法已经成功应用于中国酒钢宏达水泥厂生料分解过程中。此算法的硬件平台是西门子S7-400, 软件平台是Step7 V5.4 与WinCC V6.2。

4.1 算法实现(Algorithm realization)

项目设有1台工程师站和3台操作员站, 其中工程师站除了装有Step 7和WinCC软件外, 还安装Visual Basic 6.0软件, 算法在Visual Basic 6.0中编程, 通过OPC通讯实现的, 算法实现控制流程图如图5所示。

4.2 参数选择(Parameters selection)

根据第3节的控制策略, 控制器参数按照如下进行选择: 1) 切换机制、前馈控制器和T-S模糊控制器的参数按照文献[8]选择; 2) 改进ELMAN网络的Q学习异常工况控制器网络权值 $w^1(0)$, $w^2(0)$ 以及

$w^3(0)$ 分别为

$$w^1(0) = \begin{bmatrix} w_{11}^1(0) & w_{12}^1(0) & w_{13}^1(0) \\ \vdots & \vdots & \vdots \\ w_{91}^1(0) & w_{92}^1(0) & w_{93}^1(0) \end{bmatrix},$$

$$w_{ij}^1(0) \in [0, 1], i = 1, \dots, 9; j = 1, 2, 3,$$

$$w^2(0) = \begin{bmatrix} w_{11}^2(0) & w_{12}^2(0) & \dots & w_{19}^2(0) \\ \vdots & \vdots & \dots & \vdots \\ w_{91}^2(0) & w_{92}^2(0) & \dots & w_{99}^2(0) \end{bmatrix},$$

$$w_{ij}^2(0) \in [0, 1], i = 1, \dots, 9; j = 1, \dots, 9,$$

$$w^3(0) = [w_{11}^3(0) \ w_{12}^3(0) \ \dots \ w_{19}^3(0)],$$

$$w_{1j}^3(0) \in [0, 1], j = 1, \dots, 9.$$

学习率 $\alpha = 1.2$, 折扣因子 $\gamma = 0.5$, 退火因子 $\xi = 0.3$, 温度参数 $T_{\min} = 100$, $T_{\max} = 150$.

4.3 应用结果(Application results)

将温度智能切换控制系统应用于生料分解过程, 在2014年5月2日从10:00 AM至11:30 AM和2014年5月1日从12:30 PM至2:00 PM时运行中的分解炉温度设定 $T_{csp}(t)$ 、分解炉温度 $T_c(t)$ 、预热器C5出口温度最大值 T_{c5max} 、预热器C5出口温度 $T_{c5}(t)$ 、生料流量 $F(t)$ 、分解炉给煤 $u(t)$ 的运行曲线, 分别如图6-7所示. 曲线的横坐标为时间轴, 纵坐标为各变量的数值轴.

图6为生料分解过程在2014年5月2日从10:00 AM至11:30 AM时的运行曲线. 在10:00 AM至10:10 AM及10:20 AM至10:30 AM时, 由于预热器C5出口温度 $T_{c5}(t) \geq T_{c5max}$, 所以生料分解处于异常工况, 切换机制选择基于改进ELMAN网络的Q学

习控制器.

图7为生料分解过程在2014年5月1日从12:30 PM至2:00 PM时的运行曲线. 在12:30 PM至1:12 PM时, 由于预热器C5出口温度 $T_{c5}(t) < T_{c5max}$, 所以生料处于正常工况, 切换机制选择基于T-S的模糊控制器.

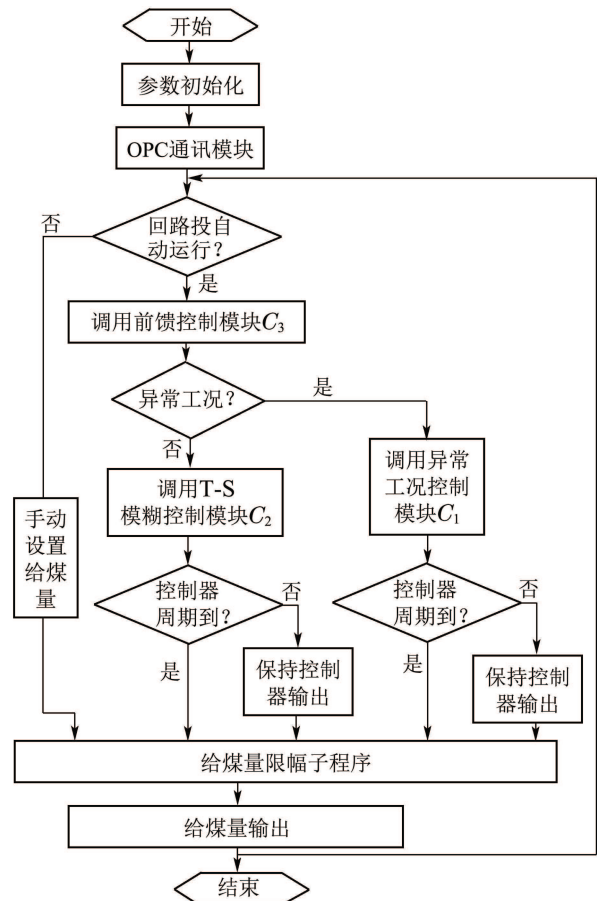


图 5 算法实现控制流程图

Fig. 5 The flow chart of algorithm realization

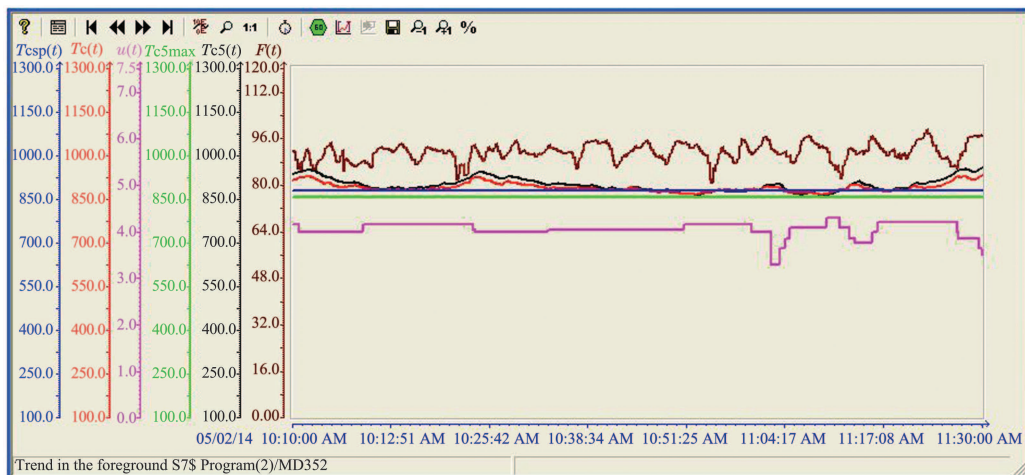


图 6 采用温度切换控制从10:00 AM至11:30 AM时的控制效果

Fig. 6 The control result of raw meal calcination process from 10:00 AM to 11:30 AM using the switching control

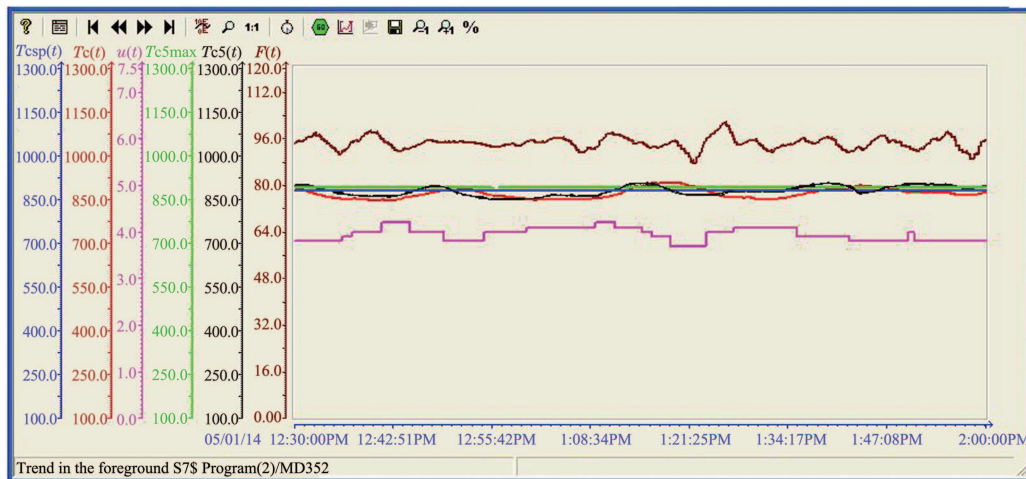


图7 采用温度切换控制从12:30 PM至2:00 PM时的控制效果

Fig. 7 Control result of raw meal calcination process from 12:30 PM to 2:00 PM using the switching control

5 结论(Conclusions)

本文提出的改进ELMAN网络的Q学习温度切换控制方法能够将分解炉温度控制在设定值上,同时使预热器C5出口温度低于最大值,并且能够根据当前工况的变化选择正确的控制器.所提出的方法已经成功应用于酒钢宏达水泥厂水泥生料分解过程.

参考文献(References):

- [1] ILIUTA I, JOHANSEN K D, JENSEN L S. Mathematical modeling of an in-line low-NO_x calciner [J]. *Chemical Engineering Science*, 2002, 57(5): 805 – 820.
- [2] KOUMBOULIS F N, KOUVAKAS N D. Indirect adaptive neural control for precalcination in cement plants [J]. *Mathematics and Computers in Simulation*, 2002, 60(5): 325 – 334.
- [3] MIKULCIC H, BERG E V, VUJANOVIC M, et al. Numerical modelling of calcination reaction mechanism for cement production [J]. *Chemical Engineering Science*, 2012, 69(1): 607 – 615.
- [4] ZHENG J X, YAN T Z, YANG J. Numerical simulation of gas and solid flow behavior in the pre-calciner with large eddy simulation approach [J]. *Energy Procedia*, 2012, 17(B): 1535 – 1541.
- [5] YLATALO J, PARKKINEN J, RITVANEN J, et al. Modeling of the oxy-combustion calciner in the post-combustion calcium looping process [J]. *Fuel*, 2013, 113(1): 770 – 779.
- [6] MUJUMDAR K S, GANESH K V, KULKARNI S B, et al. Rotary cement kiln simulator (RoCKS): integrated modelling of pre-heater, calciner, kiln and clinker cooler [J]. *Chemical Engineering Science*, 2007, 62(9): 2590 – 2607.
- [7] QIAO J H. *Study on intelligent control system for the cement raw meal calcination process* [D]. Shenyang: Northeastern University, 2012.
- [8] QIAO J H, CHAI T Y. Intelligent-based temperature switching control for cement raw meal calcination process [J]. *IEEE Transactions on Control Systems Technology*, 2014, doi:10.1109/TCST.2014.2325896.
- [9] DU J J, SONG C Y, LI P. Application of gap metric to model bank determination in multilinear model approach [J]. *Journal of Process Control*, 2009, 19(2): 231 – 240.
- [10] DU J J, SONG C Y, LI P. A gap metric based nonlinearity measure for chemical processes [C] // *2009 American Control Conference (ACC2009)*. Louis, USA: IEEE, 2009, 6: 4440 – 4445.
- [11] LIBERZON D. *Switching in Systems and Control* [M]. Boston: Birkhauser Boston Press, 2003.
- [12] 余涛, 于文俊, 李章文. 基于Q学习算法的变论域模糊控制新算法 [J]. *控制理论与应用*, 2011, 28(11): 1645 – 1650. (YU Tao, YU Wenjun, LI Zhangwen. A novel optimal method of variable-universe fuzzy control based on Q-learning algorithms [J]. *Control Theory & Applications*, 2011, 28(11): 1645 – 1650.)
- [13] 王雪松, 程玉虎, 易建强, 等. 基于ELMAN网络的非线性系统增强式学习控制 [J]. *中国矿业大学学报*, 2006, 35(5): 653 – 657. (WANG Xuesong, CHENG Yuhu, YI Jianqiang, et al. Reinforcement learning control of nonlinear system based on ELMAN network [J]. *Journal of China University of Mining and Technology*, 2006, 35(5): 653 – 657.)
- [14] RICHARDSON W B. Steepest descent using smoothed gradients [J]. *Applied Mathematics and Computation*, 2000, 112(2): 241 – 254.
- [15] MITCHELL T M. *机器学习* [M]. 北京: 机械工业出版社, 2002. (MITCHELL T M. *Machine Learning* [M]. Beijing: China Machine Press.)
- [16] SUTTON R S, BARTO A G. *Reinforcement Learning* [M]. Cambridge, MA: MIT Press, 1998.
- [17] SIMON B, BEN W. An automated signalized junction controller that learns strategies by temporal difference reinforcement learning [J]. *Engineering Applications of Artificial Intelligence*, 2013, 26(1): 652 – 659.

作者简介:

乔景慧 (1976–), 男, 博士, 讲师, 硕士研究生导师, 目前研究方向为复杂工业过程及机电系统的建模与控制、模糊控制、软测量、机器学习、设备振动及故障诊断, E-mail: qiaojh2002@163.com;

柴天佑 (1947–), 男, 中国工程院院士, 教授, 博士生导师, 目前研究方向为自适应控制、多变量智能解耦控制、流程工业过程综合自动化等, E-mail: tychai@mail.neu.edu.cn.