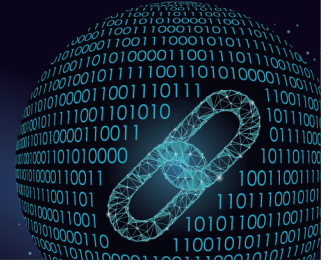




中国科技期刊卓越行动计划项目入选期刊

控制与决策

CONTROL AND DECISION



基于脱策Q学习的辊道窑温度分散 H_∞ 控制方法

陈宁, 孙嘉树, 罗彪, 李彬艳, 桂卫华

引用本文:

陈宁, 孙嘉树, 罗彪, 李彬艳, 桂卫华. 基于脱策Q学习的辊道窑温度分散 H_∞ 控制方法[J]. 控制与决策, 2024, 39(5): 1550–1556.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2022.1111>

您可能感兴趣的其他文章

Articles you may be interested in

基于2D-OTSU图像边缘检测的回转窑工况识别方法

Condition recognition method of rotary kiln based on 2D-OTSU image edge detection
控制与决策. 2021, 36(10): 2427–2434 <https://doi.org/10.13195/j.kzyjc.2020.0348>

基于反馈无源化的切换非线性系统 H_∞ 跟踪控制

Passification-based H_∞ tracking control for a class of switched nonlinear systems
控制与决策. 2021, 36(11): 2729–2734 <https://doi.org/10.13195/j.kzyjc.2020.0798>

基于T-S模糊模型的多时滞非线性网络切换控制系统非脆弱 H_∞ 控制

Non-fragile H_∞ control for multi-delay nonlinear network switching control system based on T-S model
控制与决策. 2021, 36(5): 1087–1094 <https://doi.org/10.13195/j.kzyjc.2019.1098>

一类非线性大系统分散自适应预设性能有限时间跟踪控制

Decentralized adaptive prescribed performance finite-time tracking control for a class of large-scale nonlinear systems
控制与决策. 2020, 35(12): 3045–3052 <https://doi.org/10.13195/j.kzyjc.2019.0623>

事件触发机制下分布时滞网络化控制系统 H_∞ 故障检测

Event-triggered H_∞ fault detection for networked control systems with distributed delays
控制与决策. 2020, 35(12): 3059–3065 <https://doi.org/10.13195/j.kzyjc.2019.0456>

基于脱策 Q 学习的辊道窑温度分散 H_∞ 控制方法

陈宁[†], 孙嘉树, 罗彪, 李彬艳, 桂卫华

(中南大学 自动化学院, 长沙 410083)

摘要: 辊道窑烧结过程是电池正极材料制备工艺的关键, 烧结温度的精准控制对提高材料性能、保证产品一致性至关重要。然而, 烧结过程通常面临动态信息难以获取、不同温区温度耦合严重以及存在外界干扰等问题, 给精准控制辊道窑温度带来了很大的困难。鉴于此, 提出一种新的辊道窑温度分散 H_∞ 控制方法。首先, 构造一个有界函数来描述温度关联项对当前温区控制性能的最大影响, 并根据该有界函数建立温区的极小化极大问题, 可将辊道窑温度控制问题转化为更小规模的温区温度控制问题, 通过求解所有温区的极小化极大问题的鞍点解得到辊道窑温度 H_∞ 控制策略, 实现分散控制; 然后, 采用一种脱策 Q 学习算法学习各温区极小化极大问题的鞍点解, 获得辊道窑关联系统的温度分散 H_∞ 控制器; 最后, 基于实际窑炉温度数据进行仿真实验, 实验结果表明在干扰存在的情况下, 所设计控制器仍然能够精准控制辊道窑温度稳定在设定值上。

关键词: 辊道窑; 温度控制; 强化学习; 脱策 Q 学习; H_∞ 控制; 分散控制

中图分类号: TP273

文献标志码: A

DOI: 10.13195/j.kzyjc.2022.1111

引用格式: 陈宁, 孙嘉树, 罗彪, 等. 基于脱策 Q 学习的辊道窑温度分散 H_∞ 控制方法[J]. 控制与决策, 2024, 39(5): 1550-1556.

Decentralized H_∞ control for roller kiln temperature based on off-policy Q -learning

CHEN Ning[†], SUN Jia-shu, LUO Biao, LI Bin-yan, GUI Wei-hua

(School of Automation, Central South University, Changsha 410083, China)

Abstract: Roller kiln sintering process is the key to the preparation of cathode materials, and controlling kiln temperature precisely is extremely important to reducing energy consumption, improving materials performance and heightening unity of products. However, unknown internal system dynamics, severe energy exchange between different temperature regions and frequent disturbance in the roller kiln make it very difficult to control kiln temperature accurately. A novel decentralized H_∞ control method of roller kiln temperature is proposed. First, a bounded function is constructed to describe the maximum effects caused by temperature coupling to the control performance of regions in the roller kiln, the minimax problems of different regions can be established based on bounded function above, thus, the large-scale H_∞ control problem of the whole roller kiln is turned to the small-scale H_∞ control problem of regions. The target roller kiln temperature H_∞ control policy can be obtained by solving minimax problems above, so the control method is completely decentralized. Then, the off-policy Q -learning algorithm is used to learn the roller kiln temperature decentralized H_∞ control policy. Simulation result shows that the proposed control method can not only control roller kiln temperature to reach the set point precisely, but also overcome the negative effects caused by disturbance.

Keywords: roller kiln; temperature control; reinforcement learning; off-policy Q -learning; H_∞ control; decentralized control

0 引言

现今工业上大规模制备三元正极材料主要采用高温固相法, 该方法将镍钴锰三元前驱体和含锂化合物按化学计量比均匀混合, 在高温富氧环境下烧结得

到中间产物, 再经后续破碎等工序得到最终产物, 烧结过程是其中最为核心的工序。烧结所用设备是辊道窑, 窑炉温度是系统运行过程中最关键的监控参数, 精准控制辊道窑温度对于提高材料性能、保证产

收稿日期: 2022-06-25; 录用日期: 2023-02-19.

基金项目: 国家自然科学基金重点项目(62033014).

责任编辑: 张海涛.

[†]通讯作者. E-mail: ningchen@csu.edu.cn.

品一致性具有极其重要的意义。

烧结过程中辊道窑相邻温区存在温差,温区间热量交换剧烈,辊道窑系统是一个多温区温度耦合的复杂关联系统.分散控制是一种解决关联系统稳定化问题的有效方法,由若干独立的控制器共同实施,这些控制器针对局部子系统设计,仅以局部信息作为反馈,具有结构简单、硬件要求低、通信带宽要求低以及控制实时性好的优点.目前,已有很多学者针对关联系统的分散控制方案设计展开研究^[1-5],其中, Saberi^[4]和 Lin等^[5]的工作验证了关联系统的分散控制问题与其局部子系统的最优控制问题存在密切的联系,只需适当设计代价函数,局部子系统的最优控制策略便可作为关联系统的分散控制策略。

电池正极材料的烧结反应对温度要求苛刻,对温度变化敏感,复杂的工业现场将给制备过程带来不可预测的扰动因素.针对这一问题,部分学者研究了分散 H_∞ 控制方法^[6-9],这些工作主要运用LMI处理 H_∞ 控制问题,以实现具有抗干扰性的控制器. Başar等^[10]建立了 H_∞ 控制理论与极小化极大问题的联系,将控制输入和干扰分别视为最小化和最大化参与者,通过找寻该极小化极大问题的鞍点解获得 H_∞ 控制策略,为解决分散 H_∞ 控制问题提供了新的思路。

目前,分散 H_∞ 控制方法往往要求系统模型完全已知^[6-9],然而,以辊道窑为代表的复杂工业系统通常具有难以确定的内部动态,导致基于模型的控制方法在应用上存在很大的困难.近年来,基于强化学习的控制方法蓬勃发展,其最大的优势在于能够克服传统控制方法对系统模型的依赖性.已有学者成功使用强化学习技术解决不确定关联系统的分散控制问题^[11-13], Al-Tamimi等^[14]针对线性离散系统的 H_∞ 控制问题,提出了一种基于Q学习的博弈代数黎卡提方程(game algebraic riccati equation, GARE)求解方法,这项奠基性的成果对后续研究产生了重要影响,但是,文献[14]中的Q学习算法属于一类依策学习算法,这意味着该算法要求只能以规定的方式调节输入干扰来产生数据,只适用于干扰可调节的情况,这在工程系统中往往是不可能实现的. Luo等^[15]意识到依策强化学习在 H_∞ 控制领域固有的局限性,提出了一种脱策强化学习算法,能够使用任何干扰信号学习 H_∞ 控制律,不要求干扰可调节.在文献[15]的启发下,后续工作越来越重视脱策强化学习方法在 H_∞ 控制器设计中的应用^[16-18]。

尽管上述研究已发展成熟,但是,很少有方法能够有效运用脱策强化学习技术来解决不确定关联系

统的分散 H_∞ 控制问题,尤其面对辊道窑这类实际的复杂工业系统,如何设计控制方法克服相邻温区温度耦合、烧结系统中存在的干扰所导致控制难的问题,仍然亟待解决。

本文提出一种基于脱策Q学习的辊道窑温度分散 H_∞ 控制方法,主要内容如下:针对相邻温区温度相互耦合、烧结系统中存在干扰所导致的控制难的问题,构造一个新的有界函数来描述温度关联项对于当前温区控制性能的最大影响,并根据该有界函数建立温区的极小化极大问题,将整体辊道窑温度 H_∞ 控制问题转化为规模更小又彼此孤立的温区温度 H_∞ 控制问题,实现分散控制.本文验证了所有温区极小化极大问题的鞍点解可共同作为辊道窑的分散 H_∞ 控制策略。

本文符号定义如下: \mathbf{R} 为实数集合; \mathbf{R}^+ 为正实数集合; \mathbf{R}^n 为 n 维向量集合; $\mathbf{R}^{n \times m}$ 为 $n \times m$ 实矩阵集合; I_n 为 n 维单位矩阵; $O_{n \times m}$ 为 $n \times m$ 零矩阵;对于实向量 v 和矩阵 M , $\|v\|_M = v^T M v$ 。

1 辊道窑温度建模及其控制问题

1.1 辊道窑多温区温度关联模型

电池正极材料烧结工艺常用的设备为电热式辊道窑,以中国某新能源企业为例,烧结镍钴锰酸锂三元正极材料所用的辊道窑结构如图1所示。

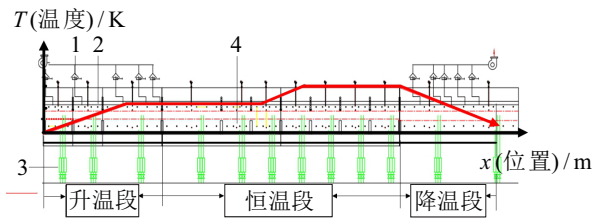


图1 辊道窑结构

工业生产中,原料装入匣钵并由辊棒运输穿过辊道窑,在此过程中完成烧结.图1中的红色实线是理想温度分散曲线,通过电加热上下两个硅碳棒维持各温区的温度稳定在设定值上。

根据不同阶段化学反应所需的理想反应温度不同,辊道窑可分为若干温区.若窑炉中通入氧气流量、各物质的比热容和总质量保持不变,则根据温度-热量计算公式 $\Delta Q = cm\Delta T$,可认为炉内温度与热量的变化呈现线性关系.假设第 i 温区的温度模型在状态空间中可写为如下形式:

$$\begin{aligned} \bar{x}_i(k+1) = & A_{i,i}\bar{x}_i(k) + A_{i,i+1}\bar{x}_{i+1}(k) + A_{i,i-1}\bar{x}_{i-1}(k) + \\ & B_{i,i}u_i(k) + D_i d(k) + C_i. \end{aligned} \quad (1)$$

其中:系统状态 $\bar{x}_i \in \mathbf{R}^2$ 为第 i 温区的温度,控制输入

$u_i \in \mathbf{R}^2$ 为第 i 温区的硅碳棒加热功率, $d_i \in \mathbf{R}$ 为烧结过程干扰因素, $A_{i,i} \in \mathbf{R}^{2 \times 2}$ 、 $B_{i,i} \in \mathbf{R}^{2 \times 2}$ 、 $D_i \in \mathbf{R}^2$ 和 $C_i \in \mathbf{R}^2$ 分别为系统动态矩阵、控制系数矩阵、干扰输入矩阵和常数项, $A_{i,i-1} \in \mathbf{R}^{2 \times 2}$ 和 $A_{i,i+1} \in \mathbf{R}^{2 \times 2}$ 分别为第 i 温区与第 $i-1$ 温区和 $i+1$ 温区的温度关联矩阵。

按照上述思路建立所有温区的温度模型后, 可得到完整辊道窑的多温区温度关联模型

$$\begin{aligned} \bar{x}_i(k+1) = & \\ & A_{i,i}\bar{x}_i(k) + A_{i,i+1}\bar{x}_{i+1}(k) + A_{i,i-1}\bar{x}_{i-1}(k) + \\ & B_{i,i}u_i(k) + D_id(k) + C_i, \quad i \in \mathbf{N}. \end{aligned} \quad (2)$$

其中: $A_{1,0} \in O_{2 \times 2}$; $A_{N,N+1} \in O_{2 \times 2}$; $\mathbf{N} = \{1, 2, \dots, N\}$, N 为温区数量。

模型(2)包含了物料和匣钵的平均比热容、窑壁平均导热系数在内的大量系数, 难以从机理上分析其准确值, 只通常认为模型(2)中的系统动态信息 $\{A_{i,i}|i \in \mathbf{N}\}$ 、 $\{B_{i,i}|i \in \mathbf{N}\}$ 、 $\{D_i|i \in \mathbf{N}\}$ 、 $\{C_i|i \in \mathbf{N}\}$ 均是未知的。

由于常数项 C 的存在, 系统(2)的平衡点不位于原点, 需要使用下式对系统平衡点进行平移:

$$x(k) = \bar{x}(k) - x_e. \quad (3)$$

其中: $x_e \in \mathbf{R}^{2N}$ 为系统(2)的平衡点; \bar{x} 为辊道窑温度, 即 $\bar{x}(k) = [\bar{x}_1^T(k), \bar{x}_2^T(k), \dots, \bar{x}_N^T(k)]^T$; $x(k) = [x_1^T(k), x_2^T(k), \dots, x_N^T(k)]^T$ 为平衡点平移后的系统状态。

将式(3)代入(2), 得到

$$\begin{aligned} x_i(k+1) = & \\ & A_{i,i}x_i(k) + A_{i,i+1}x_{i+1}(k) + A_{i,i-1}x_{i-1}(k) + \\ & B_{i,i}u_i(k) + D_id(k), \quad i \in \mathbf{N}. \end{aligned} \quad (4)$$

式(4)即为最终推导出的辊道窑多温区温度关联模型, 后文所有内容均将以其为研究对象。

1.2 辊道窑温度 H_∞ 控制问题

辊道窑烧结系统(4)的 H_∞ 控制律需要满足如下假设。

假设1 若闭环系统(4)在外界干扰输入为0时 ($d=0$) 能够保持 Lyapunov 稳定, 则称闭环系统(4)是内部稳定的。

假设2 假设 $u(k) = [u_1^T(k), u_2^T(k), \dots, u_N^T(k)]^T$, $Q \in \mathbf{R}^2$ 为非负定矩阵, $R \in \mathbf{R}^2$ 为正定矩阵, $\gamma \in \mathbf{R}^+$, 若闭环系统(4)有

$$\sum_{l=0}^{\infty} (\|x(l)\|_Q + \|u(l)\|_R) \leq \gamma^2 \sum_{l=0}^{\infty} d^T(l)d(l), \quad (5)$$

对于 $d \in L_2[0, \infty)$ 均成立, 则称闭环系统(4)拥有小于等于 γ 的 L_2 增益。

本文控制的目标是寻找满足上述假设1和假设2的窑炉温度 H_∞ 控制律。

2 辊道窑温度分散 H_∞ 控制

本节将提出一种分散控制方法, 解决辊道窑温度 H_∞ 控制问题。

考虑关系统(4)的第 i 个子系统, 即第 i 温区的温度模型为

$$x_i(k+1) = A_{i,i}x_i(k) + B_{i,i}u_i(k) + D_id(k). \quad (6)$$

将不同温区的关联信息 $\{A_{i,j}|i \in \mathbf{N}, j \in \mathbf{N}, i \neq j\}$ 集中起来, 可定义

$$Z = \begin{bmatrix} O_{2 \times 2} & A_{1,2} & & & & \\ A_{2,1} & O_{2 \times 2} & A_{2,3} & & & \\ & & \ddots & \ddots & \ddots & \\ & & & A_{N-1,N-2} & O_{2 \times 2} & A_{N-1,N} \\ & & & & A_{N,N-1} & O_{2 \times 2} \end{bmatrix}. \quad (7)$$

定理1 假设系统(4)的关联项有界, 且存在 $M_Z \in \mathbf{R}^+$ 满足 $M_Z \geq \lambda_{\max}(Z^T Z)$. 若存在正定矩阵集合 $\{P_i^* \in \mathbf{R}^{2 \times 2}|i \in \mathbf{N}\}$, 非负定矩阵集合 $\{Q_i^* \in \mathbf{R}^{2 \times 2}|i \in \mathbf{N}\}$ 以及正定矩阵集合 $\{R_i^* \in \mathbf{R}^{2 \times 2}|i \in \mathbf{N}\}$, 满足下述博弈代数黎卡提方程(modified game algebraic riccati equation, MGARE)组:

$$\begin{aligned} 0 = & Q_i + 2\text{tr}(P_i^*)M_Z I_2 + 2A_{i,i}^T P_i^* A_{i,i} - P_i^* - \\ & 4[A_{i,i}^T P_i^* B_{i,i} \quad A_{i,i}^T P_i^* D_i] \times \\ & \begin{bmatrix} R_i + 2B_{i,i}^T P_i^* B_{i,i} & 2B_{i,i}^T P_i^* D_i \\ 2D_i^T P_i^* B_{i,i} & -\gamma^2 + 2D_i^T P_i^* D_i \end{bmatrix}^{-1} \times \\ & \begin{bmatrix} B_{i,i}^T P_i^* A_{i,i} \\ D_i^T P_i^* A_{i,i} \end{bmatrix}. \end{aligned} \quad (8)$$

基于系统(4)构造控制增益集合 $\{K_{i,1}^* \in \mathbf{R}^{2 \times 2}|i \in \mathbf{N}\}$ 和干扰增益集合 $\{K_{i,2}^* \in \mathbf{R}^{2 \times 1}|i \in \mathbf{N}\}$, 有

$$\begin{aligned} K_{i,1}^* = & \\ & -[R_i + 2B_{i,i}^T P_i^* B_{i,i} + \\ & 4B_{i,i}^T P_i^* D_i(\gamma^2 - 2D_i^T P_i^* D_i)^{-1} D_i^T P_i^* B_{i,i}]^{-1} \times \\ & [2B_{i,i}^T P_i^* A_{i,i} + \\ & 4B_{i,i}^T P_i^* D_i(\gamma^2 - 2D_i^T P_i^* D_i)^{-1} D_i^T P_i^* A_{i,i}], \end{aligned} \quad (9)$$

$$\begin{aligned} K_{i,2}^* = & \\ & -[2D_i^T P_i^* D_i - \gamma^2 - \\ & 4D_i^T P_i^* B_{i,i}(R_i + 2B_{i,i}^T P_i^* B_{i,i})^{-1} B_{i,i}^T P_i^* D_i]^{-1} \times \\ & [2D_i^T P_i^* A_{i,i} - \\ & 4D_i^T P_i^* B_{i,i}(R_i + 2B_{i,i}^T P_i^* B_{i,i})^{-1} B_{i,i}^T P_i^* A_{i,i}]. \end{aligned} \quad (10)$$

当关系统(4)的控制 $\{u_1, u_2, \dots, u_N\}$ 满足 $\{u_i = K_{i,1}^* x_i \triangleq u_i^*(x_i)|i \in \mathbf{N}\}$ 时, 闭环系统满足假设1和

假设2.

证明 以第*i*个子系统为例,假设存在正定矩阵 $P_i \in \mathbf{R}^{2 \times 2}, Q_i \in \mathbf{R}^{2 \times 2}, R_i \in \mathbf{R}^{2 \times 2}, \gamma_i \in \mathbf{R}^+,$ 有界函数 $\Gamma_i(x_i, u_i, d)$ 、状态 x_i 、控制策略 u_i 以及干扰信号 d 满足

$$\begin{cases} \Gamma_i(x_i, u_i, d) \triangleq \|A_{i,i}x_i + B_{i,i}u_i + D_i d\|_{P_i} + \\ \quad 2\text{tr}(P_i)M_Z x_i^T x_i, \\ 0 = \|x_i\|_{Q_i} + \|u_i\|_{R_i} - \gamma_i^2 d^T d + \\ \quad \|A_{i,i}x_i + B_{i,i}u_i + D_i d\|_{P_i} - \\ \quad \|x_i\|_{P_i} + \Gamma_i(x_i, u_i, d). \end{cases} \quad (11)$$

为了便于表述,定义 $x_{i,k} \triangleq x_i(k), u_{i,k} \triangleq u_i(k), d_k \triangleq d(k), \mathfrak{R}_i(x_i, u_i, d) \triangleq \|x_i\|_{Q_i} + \|u_i\|_{R_i} - \gamma_i^2 d^T d.$

针对系统(4)定义如下值函数:

$$V_i(x_{i,k}) = \sum_{l=k}^{\infty} \{\mathfrak{R}_i(x_{i,l}, u_{i,l}, d_l) + \Gamma_i(x_{i,l}, u_{i,l}, d_l)\}. \quad (12)$$

基于式(12)可建立一个极小化极大问题,即

$$\begin{aligned} V_i^*(x_{i,k}) = \\ \min_{u_i} \max_d \sum_{l=k}^{\infty} \{\mathfrak{R}_i(x_{i,l}, u_{i,l}, d_l) + \Gamma_i(x_{i,l}, u_{i,l}, d_l)\}. \end{aligned} \quad (13)$$

鞍点解 $\{u_i^*, d_i^*\}$ 满足

$$u_i^* = \arg \min_{u_i} V_i(x_{i,k}), \quad (14)$$

$$d_i^* = \arg \max_d V_i(x_{i,k}). \quad (15)$$

假设 $V_i^*(x_i) = \|x_i\|_{P_i^*}, V_i^*(x_{i,k})$ 满足贝尔曼最优方程

$$\begin{aligned} V_i^*(x_{i,k}) = \mathfrak{R}_i(x_{i,k}, u_{i,k}^*, d_{i,k}^*) + \Gamma_i^*(x_{i,k}, u_{i,k}^*, d_{i,k}^*) + \\ V_i^*(x_{i,k+1}). \end{aligned} \quad (16)$$

易证明 $u_i^* = K_{i,1}^* x_i, d_i^* = K_{i,2}^* x_i$ 且式(16)与MGARE是完全等价的.

由上述推导可知,关系统(4)中的每个子系统极小化极大问题的鞍点解的集合即为定理1中的 $\{u_1^*(x_1), u_2^*(x_2), \dots, u_N^*(x_N)\}.$

定义哈密顿函数为

$$\begin{aligned} H_i(x_{i,k}, u_{i,k}, d_k, \Delta V_i^*) = \\ \mathfrak{R}_i(x_{i,k}, u_{i,k}, d_k) + \|A_{i,i}x_{i,k} + B_{i,i}u_{i,k} + D_i d_k\|_{P_i^*} - \\ \|x_{i,k}\|_{P_i^*} + \Gamma_i^*(x_{i,k}, u_{i,k}, d_k). \end{aligned} \quad (17)$$

由式(16)可知, $H_i(x_{i,k}, u_{i,k}^*, d_{i,k}^*, \Delta V_i^*) = 0,$ 文献[19]已验证若极小化极大问题(13)的鞍点解存在,则必有

$$H_i(x_{i,k}, u_{i,k}^*, d_k, \Delta V_i^*) \leq 0. \quad (18)$$

由式(18)可知

$$\begin{aligned} 0 \geq H_1(x_{1,k}, u_{1,k}^*, 0, \Delta V_1^*) + H_2(x_{2,k}, u_{2,k}^*, 0, \Delta V_2^*) + \\ \dots + H_N(x_{N,k}, u_{N,k}^*, 0, \Delta V_N^*). \end{aligned} \quad (19)$$

定义: $u_k^* \triangleq [(u_{1,k}^*)^T, (u_{2,k}^*)^T, \dots, (u_{N,k}^*)^T]^T, Q = \text{diag}(Q_1, Q_2, \dots, Q_N), P^* = \text{diag}(P_1^*, P_2^*, \dots, P_N^*), R = \text{diag}(R_1, R_2, \dots, R_N), \gamma \triangleq \sqrt{\gamma_1^2 + \gamma_2^2 + \dots + \gamma_N^2}, \Gamma^*(x, u, d) \triangleq \|Ax + Bu + Dd\|_{P^*} + 2M_Z(\text{tr}(P_1^*)x_1^T x_1 + \text{tr}(P_2^*)x_2^T x_2 + \text{tr}(P_N^*)x_N^T x_N).$

基于上述定义,式(19)等同于

$$\begin{aligned} 0 \geq \|x_k\|_Q + \|u_k^*\|_R - \gamma^2 d_k^T d_k - \|x_k\|_{P^*} + \\ \|Ax_k + Bu_k^* + Dd_k\|_{P^*} + \Gamma^*(x_k, u_k^*, d_k). \end{aligned} \quad (20)$$

根据矩阵不等式的常规性质, $\Gamma^*(x, u, d)$ 满足如下不等式:

$$\begin{aligned} \Gamma^*(x, u^*, d) \triangleq \\ \|Ax + Bu + Dd\|_{P^*} + \\ 2M_Z(\text{tr}(P_1^*)x_1^T x_1 + \dots + \text{tr}(P_N^*)x_N^T x_N) \geq \\ \|Ax + Bu + Dd\|_{P^*} + \\ 2M_Z(\lambda_{\max}(P_1^*)x_1^T x_1 + \dots + \lambda_{\max}(P_N^*)x_N^T x_N) \geq \\ \|Ax + Bu + Dd\|_{P^*} + \\ 2M_Z(\|x_1\|_{P_1^*} + \dots + \|x_N\|_{P_N^*}) = \\ \|Ax + Bu + Dd\|_{P^*} + 2M_Z\|x\|_{P^*} \geq \\ \|Ax + Bu + Dd\|_{P^*} + 2\|Zx\|_{P^*}. \end{aligned} \quad (21)$$

利用式(21)可将(20)不等号右边的 $\Gamma^*(x_k, u_k^*, d_k)$ 缩小,而不等式仍然成立,有

$$\begin{aligned} -\|x_k\|_Q - \|u_k^*\|_R + \gamma^2 d_k^T d_k \geq \\ \|Ax_k + Bu_k^* + Dd_k\|_{P^*} - \|x_k\|_{P^*} + \Gamma^*(x_k, u_k^*, d_k) \geq \\ \|Ax_k + Bu_k^* + Dd_k\|_{P^*} - \|x_k\|_{P^*} + 2\|Zx_k\|_{P^*} + \\ \|Ax_k + Bu_k^* + Dd_k\|_{P^*} \geq \\ \|Ax_k + Bu_k^* + Dd_k\|_{P^*} - \|x_k\|_{P^*} + \|Zx_k\|_{P^*} + \\ 2(Ax_k + Bu_k^* + Dd_k)P^* Zx_k = \\ \|Ax_k + Bu_k^* + Dd_k + Zx_k\|_{P^*} - \|x_k\|_{P^*}. \end{aligned} \quad (22)$$

沿着状态轨迹 $x_{k+1} = Ax_k + Bu_k^* + Dd_k + Zx_k$ 对式(22)的等式两边作无穷积分,可得到

$$-\|x_k\|_{P^*} \leq \sum_{l=k}^{\infty} (\|x_l\|_Q + \|u_l^*\|_R) + \gamma^2 \sum_{l=k}^{\infty} d_l^T d_l. \quad (23)$$

由于 $\|x_k\|_{P^*} \geq 0,$ 有

$$\sum_{l=k}^{\infty} (\|x_l\|_Q + \|u_l^*\|_R) \leq \gamma^2 \sum_{l=k}^{\infty} d_l^T d_l. \quad (24)$$

至此,闭环系统满足假设2.

由式(18),有

$$0 \geq H_1(x_{1,k}, u_{1,k}^*, 0, \Delta V_1^*) + H_2(x_{2,k}, u_{2,k}^*, 0, \Delta V_2^*) + \dots + H_N(x_{N,k}, u_{N,k}^*, 0, \Delta V_N^*). \quad (25)$$

式(25)等同于

$$0 \geq \|x_k\|_Q + \|u_k^*\|_R + \|Ax_k + Bu_k^*\|_{P^*} - \|x_k\|_{P^*} + \Gamma^*(x_k, u_k^*, 0). \quad (26)$$

由式(21)和(26),得到

$$0 \geq \|x_k\|_Q + \|u_k^*\|_R + \|Ax_k + Bu_k^* + Zx_k\|_{P^*} - \|x_k\|_{P^*}. \quad (27)$$

选择 $\|x\|_{P^*}$ 作为系统(4)的 Lyapunov 函数,由式(27)可知,当干扰信号 $d=0$ 时,对于任意状态 x_k ,有

$$\|Ax_k + Bu_k^* + Zx_k\|_{P^*} - \|x_k\|_{P^*} \leq -\|x_k\|_Q - \|u_k^*\|_R \leq 0, \quad (28)$$

表明闭环系统(4)能够保持 Lyapunov 稳定,从而可证明闭环系统满足假设1. \square

定理1表明,得到 $\{u_1^*(x_1), u_2^*(x_2), \dots, u_N^*(x_N)\}$ 可解决辊道窑烧结系统的 H_∞ 控制问题,实现本文的控制目标.

3 基于脱策Q学习的辊道窑温度控制算法

引入Q函数的定义

$$Q_i(x_{i,k}, a_1, a_2) = \mathfrak{R}_i(x_{i,k}, a_1, a_2) + V_i(x_{i,k+1}). \quad (29)$$

假定当 $V_i(x_{i,k+1}) = V_i^*(x_{i,k+1})$ 时,第 i 温区最优Q函数为

$$Q_i^*(x_{i,k}, a_1, a_2) = 2Q_i^*(x_{i,k+1}, u_i^*(x_{i,k+1}), d_i^*(x_{i,k+1})) + \|x_{i,k}\|_Q + \|a_1\|_{R_i} - \gamma_i^2 a_2^T a_2 + \alpha_i^* x_{i,k}^T x_{i,k}. \quad (30)$$

定义

$$G_i^*(x_{i,k}, a_1, a_2) = \begin{bmatrix} x_{i,k} \\ a_1 \\ a_2 \end{bmatrix}^T G_i \begin{bmatrix} x_{i,k} \\ a_1 \\ a_2 \end{bmatrix}. \quad (31)$$

其中

$$G_i = \begin{bmatrix} G_{i,xx} & G_{i,xu} & G_{i,xd} \\ G_{i,xu}^T & G_{i,uu} & G_{i,ud} \\ G_{i,xd}^T & G_{i,ud}^T & G_{i,dd} \end{bmatrix} \triangleq$$

$$\begin{bmatrix} Q_i + \alpha_i^* I_2 + 2A_{i,i}^T P_i^* A_{i,i} & 2A_{i,i}^T P_i^* B_{i,i} \\ 2B_{i,i}^T P_i^* A_{i,i} & R_i + 2B_{i,i}^T P_i^* B_{i,i} \\ 2D_{i,i}^T P_i^* A_{i,i} & 2D_{i,i}^T P_i^* B_{i,i} \\ 2A_{i,i}^T P_i^* D_i \\ 2B_{i,i}^T P_i^* D_i \\ -\gamma_i^2 + 2D_{i,i}^T P_i^* D_i \end{bmatrix}. \quad (32)$$

基于上述定义可知,式(9)和(10)可改写为

$$K_{i,2}^* = -(G_{i,uu} - G_{i,ud}^T G_{i,dd}^{-1} G_{i,ud})^{-1} \times (G_{i,xu}^T - G_{i,ud}^T G_{i,dd}^{-1} G_{i,xd}^T), \quad (33)$$

$$K_{i,2}^* = -(G_{i,dd} - G_{i,ud}^T G_{i,uu}^{-1} G_{i,ud})^{-1} \times (G_{i,xd}^T - G_{i,ud}^T G_{i,uu}^{-1} G_{i,xu}^T). \quad (34)$$

注意,使用式(33)和(34)可将辊道窑温度分散 H_∞ 控制问题转化为求解最优Q函数的问题.

下面将给出一种脱策Q学习算法,该算法只需利用辊道窑烧结过程中自然产生的温度数据,便能够迭代学习最优Q函数.

算法1 基于脱策Q学习的分散 H_∞ 控制器求解算法.

step 1: 选择温区 $i = 1$.

step 2: 设定迭代次数 $j = 0$, 给定初始容许的温度控制策略增益 $K_{i,1}^{(0)}$ 和 $K_{i,2}^{(0)}$.

step 3 (策略评价): 收集数据集 S_M 并利用 S_M 求解以下方程:

$$\begin{bmatrix} x_{i,k} \\ a_1 \\ a_2 \end{bmatrix}^T G_i^{(j)} \begin{bmatrix} x_{i,k} \\ a_1 \\ a_2 \end{bmatrix} = \|x_{i,k}\|_Q + \|a_1\|_{R_i} - \gamma_i^2 a_2^T a_2 + \alpha_i^{(j)} x_{i,k}^T x_{i,k} + 2 \begin{bmatrix} x_{i,k} \\ K_{i,1}^{(j)} x_{i,k} \\ K_{i,2}^{(j)} x_{i,k} \end{bmatrix}^T G_i^{(j)} \begin{bmatrix} x_{i,k} \\ K_{i,1}^{(j)} x_{i,k} \\ K_{i,2}^{(j)} x_{i,k} \end{bmatrix}. \quad (35)$$

step 4 (策略更新): 更新第 i 温区的温度控制策略,即

$$K_{i,1}^{(j+1)} = -(G_{i,uu}^{(j)} - G_{i,ud}^{(j)} G_{i,dd}^{(j)-1} G_{i,ud}^{(j)T})^{-1} \times (G_{i,xu}^{(j)T} - G_{i,ud}^{(j)T} G_{i,dd}^{(j)-1} G_{i,xd}^{(j)T}),$$

$$K_{i,2}^{(j+1)} = -(G_{i,dd}^{(j)} - G_{i,ud}^{(j)T} G_{i,uu}^{(j)-1} G_{i,ud}^{(j)})^{-1} \times (G_{i,xd}^{(j)T} - G_{i,ud}^{(j)T} G_{i,uu}^{(j)-1} G_{i,xu}^{(j)T}). \quad (36)$$

step 5: $\|K_{i,1}^{(j+1)} + K_{i,2}^{(j+1)} - K_{i,1}^{(j)} - K_{i,2}^{(j)}\| < \gamma_i$ 时,进入 step 6; 否则,令 $j = j + 1$, 返回至 step 3.

step 6: 当 $i = N$ 时,跳出循环; 否则,令 $i = i + 1$, 返回至 step 2.

文献[18]提供了一种通过验证Q学习算法与牛顿迭代等价的思路来表明收敛性,参考这种验证方法易验证算法1是收敛的.使用数据实现算法1的具体步骤也可参见文献[18].

4 仿真验证

本节将通过一个数值仿真案例表明所提出控制方法的有效性.选取辊道窑第9温区~第11温区的温段作为控制对象,系统模型如下式所示:

$$\begin{aligned} \bar{x}_i(k+1) = & A_{i,i}\bar{x}_i(k) + A_{i,i+1}\bar{x}_{i+1}(k) + A_{i,i-1}\bar{x}_{i-1}(k) + \\ & B_{i,i}u_i(k) + D_i d(k) + C_i. \end{aligned} \quad (37)$$

其中

$$\begin{aligned} A_{9,9} &= \begin{bmatrix} 0.300 & 0.298 \\ 0.156 & 0.285 \end{bmatrix}, & A_{9,10} &= \begin{bmatrix} 0.122 & 0 \\ 0 & 0.269 \end{bmatrix}, \\ A_{10,10} &= \begin{bmatrix} 0.295 & 0.165 \\ 0.220 & 0.298 \end{bmatrix}, & A_{10,9} &= \begin{bmatrix} 0.251 & 0 \\ 0 & 0.273 \end{bmatrix}, \\ A_{10,11} &= \begin{bmatrix} 0.299 & 0 \\ 0 & 0.220 \end{bmatrix}, & A_{11,10} &= \begin{bmatrix} 0.105 & 0 \\ 0 & 0.270 \end{bmatrix}, \\ A_{11,11} &= \begin{bmatrix} 0.296 & 0.299 \\ 0.300 & 0.129 \end{bmatrix}, & B_{9,9} &= \begin{bmatrix} 0.300 & 0 \\ 0 & 0.217 \end{bmatrix}, \\ B_{10,10} &= \begin{bmatrix} 0.002 & 0 \\ 0 & 0.135 \end{bmatrix}, & B_{11,11} &= \begin{bmatrix} 0.078 & 0 \\ 0 & 0.300 \end{bmatrix}, \\ C_9 &= \begin{bmatrix} 0.295 \\ 0.045 \end{bmatrix}, & C_{10} &= \begin{bmatrix} 0.290 \\ 0.020 \end{bmatrix}, & C_{11} &= \begin{bmatrix} 0.006 \\ 0.213 \end{bmatrix}, \\ D_9 &= \begin{bmatrix} -0.177 \\ 0.0006 \end{bmatrix}, & D_{10} &= \begin{bmatrix} 0.724 \\ -0.586 \end{bmatrix}, \\ D_{11} &= \begin{bmatrix} -0.323 \\ 0.365 \end{bmatrix}. \end{aligned}$$

选择 $\{Q_i|i = 9, 10, 11\}$ 和 $\{R_i|i = 9, 10, 11\}$ 均为 I_2 ,选择 $M = 25$,选择 $\{\gamma_i|i = 9, 10, 11\}$ 为0.2且 $\{G_i^{(0)}|i = 9, 10, 11\}$ 中所有数值均取值于 $[0, 0.01]$,并以第10温区为例来表明所提出控制方法的有效性.

图2~图4表明算法可收敛.注意:图2为矩阵 G_{10} 的模的变化过程;图3和图4中每条曲线均为矩阵中某一元素的收敛过程.

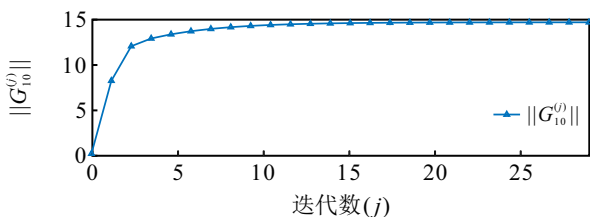


图2 $\|G_{10}^{(j)}\|$ 随迭代的收敛过程

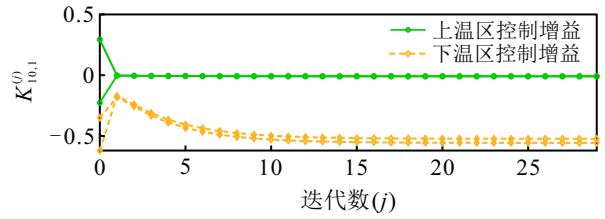


图3 $K_{10,1}^{(j)}$ 随迭代的收敛过程

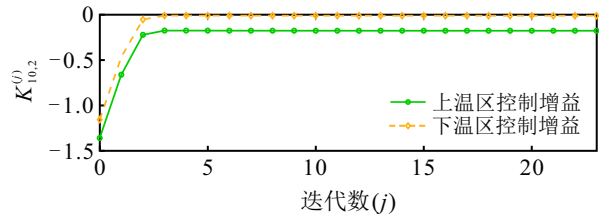


图4 $K_{10,2}^{(j)}$ 随迭代的收敛过程

通过学习获得的Q函数、分散 H_∞ 温度控制增益以及干扰增益为

$$\begin{aligned} G_{10}^* &= \begin{bmatrix} 7.823 & 2.082 & 0.014 & 0.566 & 1.243 \\ 2.082 & 7.404 & 0.009 & 0.652 & -0.305 \\ 0.014 & 0.009 & 1.000 & 0.001 & 0.023 \\ 0.566 & 0.652 & 0.001 & 1.261 & -0.778 \\ 1.243 & -0.305 & 0.023 & -0.778 & 9.502 \end{bmatrix}, \\ K_{10,1}^* &= \begin{bmatrix} -0.009 & -0.008 \\ -0.557 & -0.524 \end{bmatrix}, \\ K_{10,2}^* &= [-0.176 \quad -0.011]. \end{aligned}$$

接下来将学习获得的分散 H_∞ 温度控制策略应用于辊道窑中,随即给定一有限能量的干扰信号,记录第10温区温度和电加热功率变化曲线,上下温区初始温度分别为 44°C 和 34°C ,设定温度为 1057°C 和 1039°C ,如图5和图6所示.

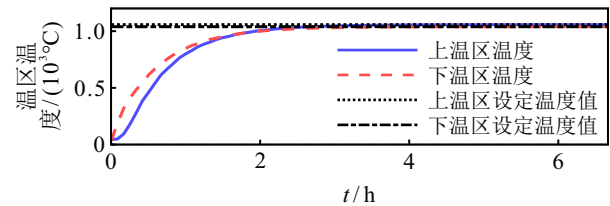


图5 第10温区温度变化曲线

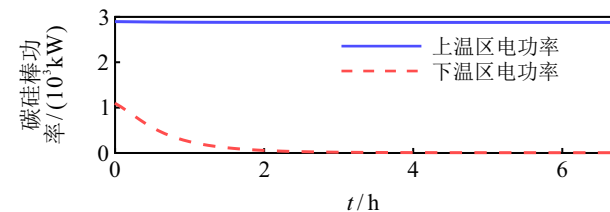


图6 第10温区电加热功率变化曲线

图5和图6的仿真结果表明,即使辊道窑内不同温区的温度存在强烈的关联耦合、辊道窑烧结系统内存在干扰,分散 H_∞ 温度控制策略仍然能够令各温

区的温度稳定于理想温度设定值,表明了所提出控制方法的有效性.

5 结论

为了解决相邻温区温度相互耦合、烧结系统中存在干扰所导致的控制难的问题,本文提出了一种基于脱策 Q 学习的辊道窑温度分散 H_∞ 控制方法.首先,在建立了窑炉多温区温度关联模型的基础上,描述了辊道窑温度 H_∞ 控制问题;然后,根据不同温区间的温度关联项设计合适的有界函数,并根据设计的函数,建立各温区的极小极大化设计问题.本文验证了通过求解上述问题的鞍点解能够获得辊道窑温度 H_∞ 控制策略,从而实现控制策略的分散化.关键的分散 H_∞ 控制策略使用一种脱策 Q 学习进行迭代学习,学习过程不需要人为调节扰动,只需收集辊道窑自然运行过程中累积的离线温度数据即可完成.基于真实数据的仿真实验表明,所提出辊道窑温度分散 H_∞ 控制方法是有效的.

参考文献(References)

- [1] Kalsi K, Lian J, Zak S H. Decentralized dynamic output feedback control of nonlinear interconnected systems[J]. IEEE Transactions on Automatic Control, 2010, 55(8): 1964-1970.
- [2] Spooner J T, Passino K M. Decentralized adaptive control of nonlinear systems using radial basis neural networks[J]. IEEE Transactions on Automatic Control, 1999, 44(11): 2050-2057.
- [3] Ioannou P. Decentralized adaptive control of interconnected systems[J]. IEEE Transactions on Automatic Control, 1986, 31(4): 291-298.
- [4] Saberi A. On optimality of decentralized control for a class of nonlinear interconnected systems[J]. Automatica, 1988, 24(1): 101-104.
- [5] Lin F, Brandt R D, Sun J. Robust control of nonlinear systems: Compensating for uncertainty[J]. International Journal of Control, 1992, 56(6): 1453-1459.
- [6] 刘碧玉, 桂卫华, 陈宁. 一类关联电力系统的时滞相关分散 H_∞ 控制[J]. 控制与决策, 2007, 22(6): 702-706. (Liu B Y, Gui W H, Chen N. Delay-dependent decentralized H_∞ control for a class of interconnected power systems[J]. Control and Decision, 2007, 22(6): 702-706.)
- [7] Zhai G S, Ikeda M, Fujisaki M. Decentralized H_∞ controller design: A matrix inequality approach using a homotopy method[J]. Automatica, 2001, 37(4): 565-572.
- [8] Chen N, Ikeda M, Gui B W. Design of H_∞ control of interconnected systems: A homotopy method[J]. International Journal of Control Automation Systems, 2005, 3(2): 143-151.
- [9] Chen N, Gui W H, Zhai G S. Robust decentralized H_∞ control for interconnected descriptor systems with norm-bounded uncertainties[J]. Asian Journal of Control, 2009, 11(1): 78-88.
- [10] Başar T, Bernhard P. H_∞ optimal control and related minimax design problems: A dynamic game approach[M]. Boston: Birkhäuser, 2008: 1-411.
- [11] Liu D R, Wang D, Li H L. Decentralized stabilization for a class of continuous-time nonlinear interconnected systems using online learning optimal control approach[J]. IEEE Transactions on Neural Networks and Learning Systems, 2014, 25(2): 418-428.
- [12] Zhao B, Wang D, Shi G, et al. Decentralized control for large-scale nonlinear systems with unknown mismatched interconnections via policy iteration[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2018, 48(10): 1725-1735.
- [13] Wang D, Liu D R, Mu C X, et al. Decentralized guaranteed cost control of interconnected systems with uncertainties: A learning-based optimal control strategy[J]. Neurocomputing, 2016, 214: 297-306.
- [14] Al-Tamimi A, Lewis F L, Abu-Khalaf M. Model-free Q -learning designs for linear discrete-time zero-sum games with application to H_∞ control[J]. Automatica, 2007, 43(3): 473-481.
- [15] Luo B, Wu H N, Huang T W. Off-policy reinforcement learning for H_∞ control design[J]. IEEE Transactions on Cybernetics, 2015, 45(1): 65-76.
- [16] Modares H, Lewis F L, Jiang Z P. H_∞ tracking control of completely unknown continuous-time systems via off-policy reinforcement learning[J]. IEEE Transactions on Neural Networks and Learning Systems, 2015, 26(10): 2550-2562.
- [17] Kiumarsi B, Lewis F L, Jiang Z P. H_∞ control of linear discrete-time systems: Off-policy reinforcement learning[J]. Automatica, 2017, 78: 144-152.
- [18] Luo B, Yang Y, Liu D R. Policy iteration Q -learning for data-based two-player zero-sum game of linear discrete-time systems[J]. IEEE Transactions on Cybernetics, 2021, 51(7): 3630-3640.
- [19] Lin W, Byrnes C I. H_∞ -control of discrete-time nonlinear systems[J]. IEEE Transactions on Automatic Control, 1996, 41(4): 494-510.

作者简介

陈宁(1970—),女,教授,博士生导师,从事复杂工业过程建模与优化控制、电池材料智能制造、自适应动态规划等研究, E-mail: ningchen@csu.edu.cn;

孙嘉树(1996—),男,硕士生,从事辊道窑温度场的建模与优化控制方法等研究, E-mail: 374110675@qq.com;

罗彪(1984—),男,教授,博士生导师,从事分布式参数系统、智能控制、强化学习、深度学习等研究, E-mail: biao.luo@hotmail.com;

李彬艳(1992—),女,博士生,从事辊道窑温度场的建模与优化控制方法等研究, E-mail: binyan.li@hotmail.com;

桂卫华(1950—),男,教授,博士生导师,中国工程院院士,从事复杂工业过程建模、优化控制与应用、故障诊断与分布式鲁棒控制等研究, E-mail: gwh@csu.edu.cn.