

无线网络环境下数据驱动混合选别浓密过程双率控制方法

吴倩¹ 范家璐¹ 姜艺¹ 柴天佑¹

摘要 无线网络环境下赤铁矿混合选别浓密过程控制问题是以底流矿浆泵频率为内环输入, 以底流矿浆流量为内环输出外环输入, 以底流矿浆浓度为外环输出的非线性串级工业过程控制问题. 其外环反馈回路存在丢包, 且模型参数难以辨识, 故本文利用工业运行过程的在线数据, 设计不依赖模型参数的跟踪控制器. 首先, 利用浓密过程运行在工作点附近的特点进行线性化, 对流量过程设计 Q -学习控制器, 保证流量过程能够跟踪给定的流量设定值; 然后采用提升技术, 得到统一时间尺度的以底流矿浆流量设定值为输入, 以底流矿浆浓度为输出的被控对象; 最后, 考虑到在无线网络环境下浓度过程存在反馈丢包, 当前的状态可能无法获得, 故采用史密斯预估器的思想, 利用历史的数据估计系统当前的状态, 设计丢包 Q -学习设定值控制器为流量过程提供最优设定值. 通过仿真实验验证所提算法的有效性.

关键词 混合选别浓密过程, Q -学习, 丢包, 史密斯预估器

引用格式 吴倩, 范家璐, 姜艺, 柴天佑. 无线网络环境下数据驱动混合选别浓密过程双率控制方法. 自动化学报, 2019, 45(6): 1122–1135

DOI 10.16383/j.aas.c180202

Data-driven Dual-rate Control for Mixed Separation Thickening Process in a Wireless Network Environment

WU Qian¹ FAN Jia-Lu¹ JIANG Yi¹ CHAI Tian-You¹

Abstract The mixed separation thickening process (MSTP) of hematite beneficiation in a wireless network environment is a nonlinear cascade process with the frequency of underflow slurry pump as the inner loop input, the slurry flow-rate as the inner loop output and the concentration as the outer loop output. The dropout occurs in the outer feedback loop, making it difficult to identify the parameters of the model, so the tracking controller only using the data generated by operational processes and independent of the knowledge of model parameters is designed in this paper. First, linearize the thickening system near the steady states, then design a controller based on Q -learning algorithm to make the inner process trace the set-point of the slurry flow-rate. Second, use the lifting technology to obtain the uniform time scale controlled object with the set-point of the slurry flow-rate as the input and the concentration as the output. Finally, considering that the networked-induced feedback dropout exists in the feedback process, meaning the current state information may be lost, a novel Smith predictor is developed to predict the current state from historical measured data, and a dropout Q -learning method is designed to provide the optimal set-point of lower loop. A simulation experiment on MSTP is given to show the effectiveness of the proposed method.

Key words Mixed separation thickening process, Q -learning, dropout, Smith predictor

Citation Wu Qian, Fan Jia-Lu, Jiang Yi, Chai Tian-You. Data-driven dual-rate control for mixed separation thickening process in a wireless network environment. *Acta Automatica Sinica*, 2019, 45(6): 1122–1135

我国赤铁矿矿石品位低、杂质含量高、嵌布粒度细、可选性差, 只经过磨矿-磁选的选矿流程难以有效去除杂质获得较高的精矿品位. 为获得较高的精矿品位, 故一般在磁选之后采用再磨、浓密和浮选

的工艺. 浮选入选最佳浓度为 30%~35%, 而经过再磨之后浓度大概为 25%~30%, 同时由于浮选过程返回的中矿矿浆的影响使浓密机的浓度波动比较大. 浓密过程可将浓度偏低的矿浆通过重力沉降浓缩为合格浓度的底流矿浆^[1-2]. 赤铁矿混合选别浓密过程是以底流矿浆泵频率为内环输入, 以矿浆流量为内环输出、外环输入, 以矿浆浓度为外环输出的串级非线性被控过程. 由于底流矿浆流量与底流矿浆浓度具有强非线性, 其模型建立过程复杂, 因此实现矿浆浓度控制成为浓密过程的研究热点.

底流矿浆流量过程难以建模, 而对底流矿浆浓度过程的控制又要兼顾底流矿浆流量的控制, 致使

收稿日期 2018-04-10 录用日期 2018-07-02
Manuscript received April 10, 2018; accepted July 2, 2018
国家自然科学基金 (61333012, 61533015, 61304028) 和中央高校基本科研专项资金 (N160804001) 资助
Supported by Natural Science Foundations of China (61333012, 61533015, 61304028) and the Fundamental Research Funds for the Central Universities (N160804001)
本文责任编辑 侯忠生
Recommended by Associate Editor HOU Zhong-Sheng
1. 东北大学流程工业综合自动化国家重点实验室 沈阳 110819
1. State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang 110819

矿浆浓度控制难度增大. 增强学习方法对于解决无模型的最优控制存在优势, 但其计算量复杂, 对浓度过程在本地设备采用增强学习的方法难以达到高的计算性能. 与此同时, 工业无线网迅猛发展, 流量过程比较简单, 故可在本地设备实现控制; 浓度过程因计算量大, 故可通过无线网络传输到工业云计算, 从而为浓度控制提高性能. 但网络传输的过程中易受丢包等不确定性的影响^[3-5], 从而影响到浓密过程的控制性能, 甚至使其不稳定. 因此, 研究浓密等工业过程中的网络控制很有必要.

文献 [6] 针对铝土矿浓密过程, 提出一种基于规则推理的控制方法对矿浆浓度进行控制. 文献 [7] 针对混合选别的浓密过程, 提出区间智能控制, 外环采用静态模型和模糊推理切换补偿控制方法. 文献 [8] 在外环采用未建模动态补偿一步最优 PI 控制与模糊推理结合的控制方法. 文献 [9] 考虑内外环提出将未建模动态补偿驱动的一步最优 PI 控制和基于模糊推理与规则推理的切换结合的控制方法. 但现有对浓密过程控制都未考虑不同网络环境下数据通信对运行控制的影响.

针对传输不可靠的线性网络, 文献 [4, 5, 10] 利用 Lyapunov 函数得到一组线性不等式求解出稳定的反馈增益. 文献 [11-12] 针对丢包将卡尔曼滤波与最优控制相结合. 文献 [13] 针对前向通道和反馈通道都存在随机时延的情况下设计鲁棒 H_2/H_∞ 控制方法. 文献 [14] 针对时延采用网络预测的控制策略解决系统的跟踪控制问题. 文献 [15] 在网络控制系统中的反馈通道和前向通道采用时延补偿的方法. 以上这些方法需要知道系统的动力学模型. 而增强学习^[16-18] (Reinforcement learning) 是一种广泛应用于寻找未知系统动力学的最优控制策略的方法. 文献 [19] 对应用层设计了最优自适应事件触发控制器以及在数据链路层中设计了无线网络的分布式调度方案. 文献 [20] 针对具有时变系统矩阵的未知网络系统提出采用自适应估计器和 Q -学习思想的随机最优控制方法求解无限维度最优调节问题. 文献 [21] 将此方法推广到非线性情况. 但文献 [19-21] 都未考虑网络存在反馈丢包的问题.

本文的主要贡献如下: 针对未发生丢包的底层矿浆流量过程, 提出 Q -学习方法, 实现底层底流矿浆流量跟踪流量设定值. 由于底流矿浆浓度过程存在状态丢包, 当前状态可能无法获得, 从而不能采用标准的 Q -学习方法计算最优控制, 所以提出一种丢包 Q -学习方法解决线性离散浓密过程的网络控制的跟踪问题, 首先采用史密斯预估器的思想通过历史的数据估计系统当前的状态, 当丢包发生时, 这些信息可应用到在线 Q -学习方法中. 论文的组织结构如下, 第 1 节为浓密过程的问题描述; 第 2 节介绍控制

器设计; 第 3 节为整体系统的性能分析; 第 4 节为仿真实验, 其结果表明所提数据驱动控制方法只利用采集的数据实现浓密过程对设定值的跟踪并使系统稳定; 第 5 节为本文结论.

1 控制问题描述

1.1 浓密过程描述

赤铁矿混合选别的浓密过程如图 1 所示, 磁选精矿矿浆经再磨工序研磨处理后, 得到浓度相对较低的精矿矿浆, 低浓度精矿矿浆流入浓密机后, 通过浓密机耙子的搅拌作用, 矿浆颗粒在自身重力的作用下, 自然沉降, 从而在浓密机底部得到浓度较高的矿浆, 以满足浮选过程的要求.

根据参考文献 [22-23], 可建立以底流矿浆泵频率 $u(t)$ 为控制输入, 以底流矿浆流量 $y(t)$ 为内环输出外环输入且以底流矿浆浓度 $r(t)$ 为外环输出的动态模型:

$$\dot{y}(t) = -\frac{y(t)}{\tau} + \frac{1}{\tau} \sqrt{\frac{k_0 u^2(t) - \frac{\Delta \rho(t)}{g \rho(r(t))} + D}{K}} \quad (1)$$

$$\dot{r}(t) = \frac{1}{k_2 h(y, r)} \left(\frac{-r^2(t) y(t)}{r(t) + k_3 (\theta(t) + Q)} + k_1 v_p(y, r, \theta) (\theta(t) + Q) + \frac{k_1 (k_i - k_3) v_p(y, r) (\theta(t) + Q)}{r(t) + k_3 (\theta(t) + Q)} \right) \quad (2)$$

其中, $k_1 = S k_i$, $k_2 = S p$, $k_3 = k_i - \mu(p_s - p_l)/S p$, $Q = q_3 \varphi_3$, φ_3 、 q_3 是磁选精矿矿浆浓度和流量, k_i 和 S 是与浓密机结构有关的常数, $\theta(t)$ 为干扰, 且 $\theta(t) = q_1 \varphi_1 + q_2 \varphi_2$. 假设 $\theta(t)$ 恒定. 各符号及物理意义见表 1.

表 1 浓密过程符号表

Table 1 Mixed separation thickening process symbol table

符号	物理含义	符号	物理含义
S	浓密机横截面积	$\frac{\Delta \rho(t)}{g \rho(\cdot)}$	泵两端管路单位重量矿浆的势能差
μ	介质的粘度	D	阻力损失
p	平均浓度系数	k_i, \bar{K}	与浓密机结构有关的常数
p_s	矿浆内固体密度	g	重力加速度
p_l	矿浆内液体密度	$\theta(t)$	干扰
k_0	静态放大系数	$h(\cdot)$	泥层界面高度
τ	时间常数	$v_p(\cdot)$	矿浆颗粒沉降速度
φ_1	浮选中矿矿浆浓度	q_1	浮选中矿流量
φ_2	污水浓度	q_2	污水流量
φ_3	磁选精矿矿浆浓度	q_3	磁选精矿矿浆流量

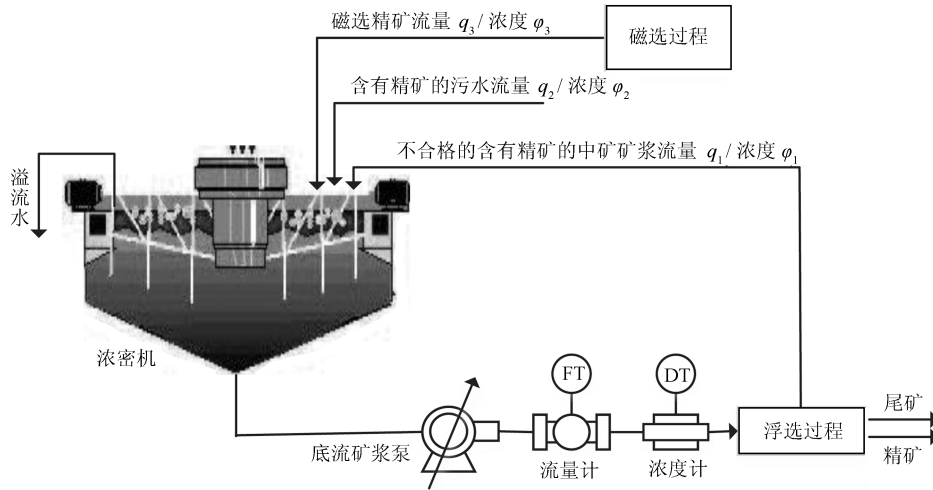


图 1 混合选别浓密过程

Fig. 1 The mixed separation thickening process

底流矿浆流量过程为快过程, 采样周期为 k ; 底流矿浆浓度过程为慢过程, 采样周期为 $T = nk$ (n 为整数). 利用工业过程通常在工作点附近稳态运行的特点, 分别对式 (1) 和式 (2) 在工作点对其线性化并离散化得到其线性模型, 则底层底流矿浆流量过程的线性模型为

$$\begin{aligned} x_1(k+1) &= A_1x_1(k) + B_1u(k) \\ y(k) &= C_1x_1(k) \end{aligned} \quad (3)$$

其中, $x_1(k)$ 维数为 1×1 , $u(k)$ 为底流矿浆泵频率且维数为 1×1 , $y(k)$ 为底流矿浆流量且维数为 1×1 . A_1 , B_1 和 C_1 的维数分别为 1×1 , 1×1 和 1×1 . 运行层底流矿浆浓度过程的线性模型为

$$\begin{aligned} x_2(T+1) &= A_2x_2(T) + B_2y(T) \\ r(T) &= C_2x_2(T) \end{aligned} \quad (4)$$

其中, $x_2(T)$ 维数为 1×1 , $y(T)$ 维数为 1×1 , $r(T)$ 为底流矿浆流量且维数为 1×1 . A_2 , B_2 和 C_2 的维数分别为 1×1 , 1×1 和 1×1 .

1.2 无线网络环境下丢包模型

系统的状态 $\eta(T)$ 通过无线网传输到控制器时可能会发生丢包, 根据参考文献 [4, 5, 24], 可知状态 $\eta(T)$ 经过网络传输后可得到的状态 $\eta_f(T)$ 为

$$\eta_f(T) = \delta(T)\eta(T) + (1 - \delta(T))\eta_f(T - 1) \quad (5)$$

其中, $\delta(T)$ 取值为 0 和 1, 当 $\delta(T) = 0$ 表示此时网络存在丢包, 反之 $\delta(T) = 1$ 表示此时通过网络传输的信号传输成功. 系统如果一直处于丢包的状态即控制系统相当于开环系统, 所以需做以下假设:

假设 1. 反馈丢包 $\delta(T)$ 的最大连续发生丢包的次数为 $\delta_{f \max}$ 有界, 即

$$\sum_{i=0}^{\delta_{f \max}} \delta(T - i) > 0 \quad (6)$$

如果当最大连续发生丢包的次数 $\delta_{f \max}$ 是无界的, 控制系统相当于开环系统, 所以此论文认为 $\delta_{f \max}$ 是有界的.

1.3 控制问题描述

本文的控制问题为对于浓密过程的线性模型 (3) 和 (4), 在网络存在丢包的情况下, 设计的控制器可完全基于采集到的数据实现系统的线性二次跟踪 (Linear quadratic tracking, LQT). 其中浓密过程的浓度的设定值为定值 r^* , 运行层流量设定值控制器为底层提供底流矿浆流量的设定值 $y^*(T)$, 为解决浓密过程底层底流矿浆流量和运行层底流矿浆浓度的跟踪问题, 故设底层的性能指标为

$$J_1(k) = \sum_{i=k}^{\infty} \gamma^{i-k} \times \left[(y^*(i) - y(i))^T Q_1 (y^*(i) - y(i)) + u^T(i) R_1 u(i) \right] \quad (7)$$

其中, γ ($0 < \gamma < 1$) 为衰减因子, 矩阵 Q_1 和 R_1 为适当维数的正定矩阵, $y^*(k)$ 维数为 1×1 .

设运行层的性能指标为

$$J_2(T) = \sum_{i=T}^{\infty} \bar{\gamma}^{i-T} \times \left[(r^* - r(i))^T Q_2 (r^* - r(i)) + y^{*T}(i) R_2 y^*(i) \right] \quad (8)$$

其中, $\bar{\gamma}$ ($0 < \bar{\gamma} < 1$) 为衰减因子, 矩阵 Q_2 和 R_2 为适当维数的正定矩阵, r^* 维数为 1×1 .

注 1. 由于时间趋于无穷时, 底层控制输入和运行层的控制输入分别与其设定值有关, 且设定值都不为 0, 故不能保证控制输入为 0, 当衰减因子为 1 时不能保证性能指标有界, 故衰减因子取小于 1.

2 控制算法

2.1 控制策略

由浓密过程的动态模型 (1) 和 (2) 可知, 底流矿浆泵的流量 $u(k)$ 首先影响底流矿浆流量 $y(k)$, 然后影响到底流矿浆浓度 $r(T)$, 故可先设计 Q -学习控制器在底层模型未知的情况下实现底流矿浆流量的跟踪得到内环底流矿浆流量闭环控制系统. 因流量过程是快过程, 浓度过程是慢过程, 此为双层结构, 考虑到在外环采样周期内, 内环流量的设定值不变, 故采用提升技术^[5, 9-10, 25] 得到一个采样周期为外环采样周期的矿浆浓度外环动态模型, 由于系统的状态通过无线网络传输时可能会发生丢包现象, 此刻系统的状态可能无法获得, 而传统的 Q -学习算法需要知道此时系统的状态值, 故设计丢包 Q -学习流量设定值控制器, 其中, 利用史密斯预估器可根据过去时刻网络传输成功时的状态量估计此刻系统的状态值, 将过去时刻的状态、过去时刻流量过程的设定值以及浓密过程的设定值进行重组为 $z(T)$, 即基于史密斯预估器的状态重组. 然后将重组后的 $z(T)$ 应用到 Q -学习流量设定值控制算法中, 从而为底流矿浆流量过程提供流量设定值. 控制策略同时考虑到底层和运行层的动态, 在网络发生丢包时不需要知道系统的模型也可以实现底流矿浆流量和底流矿浆浓度的跟踪. 图 2 为数据驱动的无线网络下浓密过程的控制结构图, 其控制策略包括 Q -学习流量控制器、提升技术和丢包 Q -学习流量设定值控制器.

2.2 Q -学习流量控制器设计

在每一个运行层的采样周期内丢包 Q -学习流量设定值控制器为底层底流矿浆流量过程提供的设定值 $y^*(k)$ 为恒定的, 所以底流矿浆流量过程的主要目标是跟踪底层的设定值 $y^*(k)$, 且

$$y^*(k+1) = F_1 y^*(k) \quad (9)$$

其中, F_1 为单位矩阵. 根据式 (3) 及式 (9) 得到底

流矿浆流量的增广系统为

$$X_d(k+1) = A_d X_d(k) + B_d u(k) \quad (10)$$

其中, $X_d(k) = \begin{bmatrix} x_1^T(k) & y^{*T}(k) \end{bmatrix}^T$, $A_d = \begin{bmatrix} A_1 & 0 \\ 0 & F_1 \end{bmatrix}$

和 $B_d = \begin{bmatrix} B_1 & 0 \end{bmatrix}^T$, 且 $X_d(k)$ 维数为 2×1 .

根据增广向量 $X_d(k)$ 的定义, 底流矿浆流量过程的性能指标 (7) 可重新写成

$$J_1(k) = \sum_{i=k}^{\infty} \gamma^{i-k} \left[X_d^T(i) Q_d X_d(i) + u^T(i) R_1 u(i) \right] \quad (11)$$

其中, $C_d = \begin{bmatrix} C_1 & -I \end{bmatrix}$, $Q_d = C_d^T Q_1 C_d$.

针对具有衰减因子的线性跟踪问题, 选取的控制器形式为

$$u(k) = K_1 x_1(k) + K_2 y^*(k) = K X_d(k) \quad (12)$$

根据参考文献 [26], 可知针对性能指标 (11) 和选取的控制策略 (12), 选取合适的衰减因子 γ 使 $(\gamma^{0.5} F_1)$ 是稳定的, 此时性能指标 (11) 可表示为线性二次型的值函数:

$$J_1(k) = V_1(k) = X_d^T(k) P_d X_d(k) \quad (13)$$

其中, $P_d = P_d^T > 0$.

由式 (11) 可以得到如下 LQT 贝尔曼 (Bellman) 方程:

$$V_1(k) = X_d^T(k) Q_d X_d(k) + u^T(k) R_1 u(k) + \gamma V_1(k+1) \quad (14)$$

将式 (13) 代入式 (14) 得到以值函数核矩阵 P_d 表示的 LQT 贝尔曼方程:

$$X_d^T(k) P_d X_d(k) = X_d^T(k) Q_d X_d(k) + u^T(k) R_1 u(k) + \gamma X_d^T(k+1) P_d X_d(k+1) \quad (15)$$

从而得到 LQT 哈密顿 (Hamiltonian) 函数:

$$H_1(k) = X_d^T(k) Q_d X_d(k) + u^T(k) R_1 u(k) + \gamma X_d^T(k+1) P_d X_d(k+1) - X_d^T(k) P_d X_d(k)$$

根据参考文献 [26], 基于最优性的必要条件, 即 $\partial H_1(k) / \partial u(k) = 0$, 可得

$$K = -(R_1 + \gamma B_d^T P_d B_d)^{-1} \gamma B_d^T P_d A_d \quad (16)$$

且矩阵 P_d 满足如下代数黎卡提方程 (Algebraic Riccati equation, ARE):

$$Q_d - P_d + \gamma A_d^T P_d A_d - \gamma^2 A_d^T P_d B_d \times (R_1 + \gamma B_d^T P_d B_d)^{-1} B_d^T P_d A_d = 0 \quad (17)$$

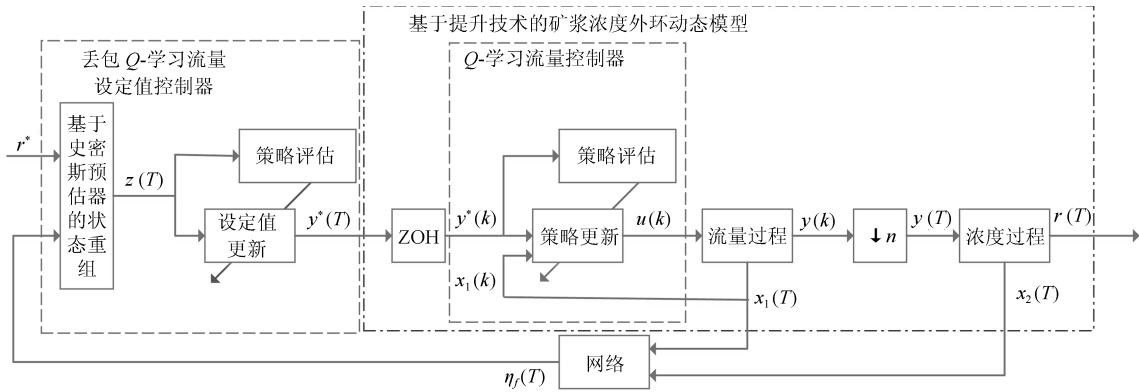


图 2 数据驱动的无线网络下浓密过程的控制结构图

Fig.2 Structure diagram of data-driven for MSTP under wireless network environment

本节设计的 Q -学习流量控制器不需要模型中 A_1 和 B_1 的值, 只利用输入输出的数据就能在线解决代数黎卡提方程 (17). 根据 LQT 贝尔曼方程 (15), 可定义离散时间的 Q -函数 (Q -function) 为

$$Q_1(k) = X_d^T(k)Q_dX_d(k) + u^T(k)R_1u(k) + \gamma X_d^T(k+1)P_dX_d(k+1) \quad (18)$$

将式 (10) 代入式 (18) 可得

$$Q_1(k) = \begin{bmatrix} X_d(k) \\ u(k) \end{bmatrix}^T H \begin{bmatrix} X_d(k) \\ u(k) \end{bmatrix} = \begin{bmatrix} X_d(k) \\ u(k) \end{bmatrix}^T \begin{bmatrix} H_{X_dX_d} & H_{X_du} \\ H_{uX_d} & H_{uu} \end{bmatrix} \begin{bmatrix} X_d(k) \\ u(k) \end{bmatrix} \quad (19)$$

其中,

$$\begin{aligned} H_{X_dX_d} &= Q_d + \gamma A_d^T P_d A_d \\ H_{X_du} &= H_{uX_d}^T = \gamma A_d^T P_d B_d \\ H_{uu} &= R_1 + \gamma B_d^T P_d B_d \end{aligned}$$

根据 $\partial Q_1(k)/\partial u(k) = 0$ 得到流量过程的控制输入:

$$u(k) = -H_{uu}^{-1}H_{uX_d}X_d(k) \quad (20)$$

故 $K = -H_{uu}^{-1}H_{uX_d}$, 显然其等价于式 (16).

根据定义的 Q -函数, 引入不依赖模型参数的 Q -学习算法就能得到底层底流矿浆流量过程的控制输入. Q -方程满足下面贝尔曼方程:

$$Q_1(k) = X_d^T(k)Q_dX_d(k) + u^T(k)R_1u(k) + \gamma Q_1(k+1) \quad (21)$$

定义 $Z_d(k) = [X_d^T(k) \quad u^T(k)]^T$, 则式 (19) 变成

$$Q_1(k) = Z_d^T(k)HZ_d(k) \quad (22)$$

其中, $Z_d(k)$ 维数为 3×1 .

将式 (22) 代入 (21) 可得到 Q -函数的贝尔曼方程:

$$Z_d^T(k)HZ_d(k) = X_d^T(k)Q_dX_d(k) + u^T(k)R_1u(k) + \gamma Z_d^T(k+1)HZ_d(k+1) \quad (23)$$

根据 Q -函数贝尔曼方程 (23) 以及流量过程的控制输入表达式 (20), 采用策略迭代的方法可以实现底流矿浆流量的跟踪, 算法如下:

算法 1. 基于策略迭代的底流矿浆流量的跟踪控制

初始化: 开始于稳定的控制策略 K^1 , 依次重复下面两个步骤直到控制输入收敛.

1) 策略评估:

$$Z_d^T(k)H^{j+1}Z_d(k) = X_d^T(k)Q_dX_d(k) + (u^j(k))^T R_1 u^j(k) + \gamma Z_d^T(k+1)H^{j+1}Z_d(k+1)$$

2) 策略提升:

$$u^{j+1}(k) = -(H_{uu}^{j+1})^{-1}H_{uX_d}^{j+1}X_d(k)$$

注 2. 算法 1 的收敛性在文献 [27–28] 中有证明. 可采用最小二乘的方法计算 H^{j+1} , 由于 H 是对称矩阵, 故执行最小二乘方法前应至少收集 6 组数据才能保证满秩的条件.

2.3 无线网络环境下浓度过程控制算法

2.3.1 基于提升技术的矿浆浓度外环动态模型

将式 (12) 代入式 (3), 得到以矿浆流量设定值 $y^*(k)$ 为输入且以矿浆流量 $y(k)$ 为输出的稳定闭环方程:

$$\begin{aligned} x_1(k+1) &= (A_1 + B_1K_1)x_1(k) + B_1K_2y^*(k) \\ y(k) &= C_1x_1(k) \end{aligned} \quad (24)$$

由于运行层设定值 Q -学习控制给出的设定值 $y^*(T)$ 是慢信号, 而底层控制的设定值 $y^*(k)$ 是快信号, 故采用提升技术, 即流量设定值 $y^*(k)$ 利用零阶保持器, 对应下采样器的参数为 n , 即

$$y^*(T) = y^*(nk) = y^*(nk+1) = \dots = y^*(nk+n-1) \quad (25)$$

将式 (24) 结合式 (25) 得到

$$\begin{aligned} x_1(T+1) &= x_1(n(k+1)) = x_1(nk+n) = \\ & (A_1 + B_1K_1)x_1(nk+n-1) + \\ & B_1K_2y^*(nk+n-1) = \dots = \\ & (A_1 + B_1K_1)^n x_1(nk) + \\ & \sum_{i=0}^{n-1} (A_1 + B_1K_1)^i B_1K_2y^*(nk) = \\ & A_o x_1(T) + B_o y^*(T) \\ y(T) &= C_1 x_1(T) \end{aligned} \quad (26)$$

其中, $A_o = (A_1 + B_1K_1)^n$ 和 $B_o = \sum_{i=0}^{n-1} (A_1 + B_1K_1)^i B_1K_2$.

将式 (26) 代入式 (4), 并与式 (4) 组成新的增广系统, 则基于提升技术的矿浆浓度外环动态模型为

$$\begin{aligned} \begin{bmatrix} x_1(T+1) \\ x_2(T+1) \end{bmatrix} &= \begin{bmatrix} A_o & 0 \\ B_2C_1 & A_2 \end{bmatrix} \begin{bmatrix} x_1(T) \\ x_2(T) \end{bmatrix} + \\ & \begin{bmatrix} B_o \\ 0 \end{bmatrix} y^*(T) \\ r(T) &= \begin{bmatrix} 0 & C_2 \end{bmatrix} \begin{bmatrix} x_1(T) \\ x_2(T) \end{bmatrix} \end{aligned}$$

令 $\tilde{A} = \begin{bmatrix} A_o & 0 \\ B_2C_1 & A_2 \end{bmatrix}$, $\tilde{B} = \begin{bmatrix} B_o \\ 0 \end{bmatrix}$, $\tilde{C} = \begin{bmatrix} 0 & C_2 \end{bmatrix}$ 和 $\eta(T) = \begin{bmatrix} x_1^T(T) & x_2^T(T) \end{bmatrix}^T$, 则

$$\begin{aligned} \eta(T+1) &= \tilde{A}\eta(T) + \tilde{B}y^*(T) \\ r(T) &= \tilde{C}\eta(T) \end{aligned} \quad (27)$$

其中, $\eta(T)$ 的维数为 2×1 .

底流矿浆浓度的设定值为 r^* , 工艺要求该浓度在一定范围内, 故设 r^* 为满足工艺要求的常数. 为解决系统的跟踪问题, 式 (27) 与底流矿浆浓度的设

定值可重组为

$$\begin{aligned} X_h(T+1) &= \begin{bmatrix} \tilde{A} & 0 \\ 0 & F_2 \end{bmatrix} \begin{bmatrix} \eta(T) \\ r^* \end{bmatrix} + \begin{bmatrix} \tilde{B} \\ 0 \end{bmatrix} y^*(T) = \\ & A_h X_h(T) + B_h y^*(T) \\ r(T) &= \begin{bmatrix} \tilde{C} & 0 \end{bmatrix} X_h(T) = C_h X_h(T) \end{aligned} \quad (28)$$

其中, F_2 为适当维数的单位矩阵, $X_h(T)$ 维数为 3×1 .

2.3.2 史密斯预估器设计

由于在无线网络下系统的状态量会发生丢包, 故利用史密斯预估器的思想, 通过过去传输成功的数据估计出此刻系统的状态 $\eta(T)$. 定义在 T 步之前发生丢包的次数为 $\delta_{fn}(T)$. 由假设 1 可知, $0 \leq \delta_{fn}(T) \leq \delta_{f \max}$, 同时根据式 (5) 得到在 T 时刻状态量传输成功时 $\delta_{fn}(T) = 0$, 从而通过无线网络进行传输时在第 T 步可以获得的最近的有用数据为 $\eta_f(T) = \eta(T - \delta_{fn}(T))$.

根据丢包次数的定义, 可以将网络丢包现象认为是随机有界延迟现象, 所以可以用过去系统采集到的未发生丢包的数据和控制输入的信息预测出当前系统的状态:

$$\begin{aligned} \eta(T) &= \tilde{A}^{\delta_{fn}(T)} \eta(T - \delta_{fn}(T)) + \\ & \sum_{i=1}^{\delta_{fn}(T)} \tilde{A}^{i-1} \tilde{B} y^*(T - i) \end{aligned} \quad (29)$$

其中, $\delta_{fn}(T)$ 是已知的.

结合式 (28) 和 (29), 利用过去时刻的数据可预测出当前的增广状态 $X_h(T)$:

$$X_h(T) = Mz(T) \quad (30)$$

其中,

$$M = \begin{bmatrix} I & \tilde{A} & \dots & \tilde{A}^{\delta_{f \max}} & \tilde{B} & \tilde{A}\tilde{B} & \dots & \tilde{A}^{\delta_{f \max}-1}\tilde{B} & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & F_2 \end{bmatrix},$$

在 T 时刻, $z(T)$ 是已知的, 且 $z(T)$ 维数为 $n_z \times 1$, $n_z = (\delta_{f \max} + 1) \times 2 + \delta_{f \max} + 1$. 则当 $\delta_{fn}(T) = 0, 1, \dots, \delta_{f \max}$ 时, $z(T)$ 分别表示为

$$z(T) = \begin{bmatrix} \underbrace{\eta_f^T(T) \dots 0}_{\delta_{f \max} + 1} \underbrace{0 \dots 0}_{\delta_{f \max}} r^{*T} \\ \delta_{fn}(T) = 0 \end{bmatrix}^T,$$

$$z(T) = \begin{bmatrix} \underbrace{0 \quad \eta_f^T(T) \quad \cdots \quad 0}_{\delta_{f \max} + 1} \\ \underbrace{y^{*T}(T-1) \quad \cdots \quad 0}_{\delta_{f \max}} \quad r^{*T} \\ \vdots \\ \underbrace{0 \quad \cdots \quad \eta_f^T(T)}_{\delta_{f \max} + 1} \\ \underbrace{y^{*T}(T-1) \quad \cdots \quad y^{*T}(T - \delta_{f \max})}_{\delta_{f \max}} \quad r^{*T} \end{bmatrix}^T, \delta_{fn}(T) = 1$$

$$\delta_{fn}(T) = \delta_{f \max}$$

2.3.3 丢包 Q-学习流量设定值控制器设计

为解决运行层底流矿浆浓度的跟踪问题, 根据增广状态 $X_h(T)$ 的定义, 其性能指标 (8) 可写成

$$J_2(T) = \sum_{i=T}^{\infty} \bar{\gamma}^{i-T} \left[X_h^T(i) \tilde{Q} X_h(i) + y^{*T}(i) R_2 y^*(i) \right] \quad (31)$$

其中, $\tilde{Q} = \begin{bmatrix} \tilde{C} & -I \end{bmatrix}^T Q_2 \begin{bmatrix} \tilde{C} & -I \end{bmatrix}$. 针对发生丢包的系统, 设计底流矿浆流量的设定值形式如下

$$y^*(T) = L_1 \eta(T) + L_2 r^* = LMz(T) = \tilde{L}z(T) \quad (32)$$

根据参考文献 [24], 当选取稳定的控制策略 (32) 和合适的衰减因子 $\bar{\gamma}$ 使 $(\bar{\gamma}^{0.5} F_2)$ 是稳定的, 能够将系统的性能指标 (31) 写成二次型的形式:

$$J_2(T) = X_h^T(T) P_h X_h(T) = z^T(T) \tilde{P} z(T) \quad (33)$$

其中, $P_h = P_h^T > 0$ 和 $\tilde{P} = M^T P_h M > 0$.

由式 (31) 和 (33), 得到如下丢包形式的 LQT 贝尔曼方程:

$$z^T(T) \tilde{P} z(T) = z^T(T) M^T \tilde{Q} M z(T) + y^{*T}(T) R_2 y^*(T) + \bar{\gamma} z^T(T+1) \tilde{P} z(T+1) \quad (34)$$

从而得到如下 LQT 哈密顿函数:

$$H_2(T) = z^T(T) M^T \tilde{Q} M z(T) + y^{*T}(T) R_2 y^*(T) + \bar{\gamma} z^T(T+1) \tilde{P} z(T+1) - z^T(T) \tilde{P} z(T)$$

LQT 贝尔曼方程的稳定条件为

$$\frac{\partial H_2(T)}{\partial y^*(T)} = 2R_2 y^*(T) + \bar{\gamma} \frac{\partial z^T(T+1)}{\partial y^*(T)} \frac{\partial J_2(T+1)}{\partial z(T+1)} = 0 \quad (35)$$

结合式 (28) 和 (30), 可以得到

$$z(T+1) = M^* A_h M z(T) + M^* B_h y^*(T) \quad (36)$$

其中, $M^* = M^T (M M^T)^{-1}$ 为 M 的右逆, 将其代入式 (35) 得到

$$y^*(T) = - (R_2 + \bar{\gamma} B_h^T M^{*T} \tilde{P} M^* B_h)^{-1} \times \bar{\gamma} B_h^T M^{*T} \tilde{P} M^* A_h M z(T) - (R_2 + \bar{\gamma} B_h^T P_h B_h)^{-1} \bar{\gamma} B_h^T P_h A_h M z(T) \quad (37)$$

将式 (36) 和 (37) 代入丢包形式的 LQT 贝尔曼方程 (34) 中, 从而得到丢包形式的 LQT 黎卡提方程

$$M^T \tilde{Q} M - \tilde{P} + \bar{\gamma} M^T A_h^T M^{*T} \tilde{P} M^* A_h M - \bar{\gamma}^2 M^T A_h^T M^{*T} \tilde{P} M^* B_h (R_2 + \bar{\gamma} B_h^T M^{*T} \tilde{P} M^* B_h)^{-1} \times B_h^T M^{*T} \tilde{P} M^* A_h M = 0 \quad (38)$$

因为 M 是行满秩, 故式 (38) 可化为

$$\tilde{Q} - P_h + \bar{\gamma} A_h^T P_h A_h - \bar{\gamma}^2 A_h^T P_h B_h \times (R_2 + \bar{\gamma} B_h^T P_h B_h)^{-1} B_h^T P_h A_h = 0 \quad (39)$$

引理 1^[24]. 将式 (32) 代入具有丢包的系统 (27), 选择合适的衰减因子 $\bar{\gamma}$ 使 $(\bar{\gamma}^{0.5} F_2)$ 是稳定的, 同时参数满足

$$0 < (P_{11} - \tilde{C}^T Q_2 \tilde{C})(P_{11} + G)^{-1} < \bar{\gamma}^2 I$$

其中, $P_{11} = \sum_{i=0}^{\infty} \bar{\gamma}^i \left[(G_c^i)^T (\tilde{C}^T Q_2 \tilde{C} + L_1^T R_2 L_1) G_c^i \right]$, $G = \tilde{A}^T P_{11} \tilde{B} (R_2 + \tilde{B}^T P_{11} \tilde{B})^{-1} R_2 (R_2 + \tilde{B}^T P_{11} \tilde{B})^{-1} \times \tilde{B}^T P_{11} \tilde{A}$ 和 $G_c = \tilde{A} + \tilde{B} L_1$, 从而可以得到系统 (27) 是稳定的, 此时控制为最优的即能最小化性能指标 (31), 此部分证明放在下一节.

基于 LQT 贝尔曼方程的定义 (34), 则可将丢包 Q-函数定义为

$$Q_2(T) = z^T(T) M^T \tilde{Q} M z(T) + y^{*T}(T) R_2 y^*(T) + \bar{\gamma} z^T(T+1) \tilde{P} z(T+1) \quad (40)$$

将式 (36) 代入式 (40) 得到

$$Q_2(T) = \chi^T(T) H_2 \chi(T) \quad (41)$$

其中, $\chi(T) = \begin{bmatrix} z^T(T) & y^{*T}(T) \end{bmatrix}^T$,

$$\begin{aligned} H_2 &= \begin{bmatrix} H_{zz} & H_{zr_1} \\ H_{r_1z} & H_{r_1r_1} \end{bmatrix} \\ H_{zz} &= M^T \tilde{Q} M + \bar{\gamma} M^T A_h^T M^{*T} \tilde{P} M^* A_h M \\ H_{zr_1} &= H_{r_1z}^T = \bar{\gamma} M^T A_h^T M^{*T} \tilde{P} M^* B_h \\ H_{r_1r_1} &= R_2 + \bar{\gamma} B_h M^{*T} \tilde{P} M^* B_h \end{aligned}$$

令 $\partial Q_2(T)/\partial y^*(T) = 0$ 得到底流流量的最优设定值:

$$y^*(T) = -H_{r_1r_1}^{-1} H_{r_1z} z(T) \quad (42)$$

显然, 式 (37) 和 (42) 等价.

根据丢包 Q -函数的定义 (40) 结合式 (34), 则 Q -函数满足丢包 LQT 贝尔曼方程:

$$\begin{aligned} Q_2(T) &= z^T(T) M^T \tilde{Q} M z(T) + \\ & y^{*T}(T) R_2 y^*(T) + \bar{\gamma} Q_2(T+1) \end{aligned} \quad (43)$$

将式 (41) 代入式 (43) 得到丢包 Q -函数贝尔曼方程:

$$\begin{aligned} \chi^T(T) H_2 \chi(T) &= z^T(T) M^T \tilde{Q} M z(T) + \\ & y^{*T}(T) R_2 y^*(T) + \bar{\gamma} \chi^T(T+1) H_2 \chi(T+1) \end{aligned} \quad (44)$$

定义

$$\begin{aligned} z(T) &= \begin{bmatrix} z_1(T) \\ z_2(T) \\ r^* \end{bmatrix}, M = \begin{bmatrix} I & \bar{M} & 0 \\ 0 & 0 & I \end{bmatrix} \\ \tilde{Q} &= \begin{bmatrix} \tilde{C}^T Q_2 \tilde{C} & -\tilde{C}^T Q_2 \\ -Q_2 \tilde{C} & Q_2 \end{bmatrix} \end{aligned}$$

其中, $z_1(T)$ 是 $z(T)$ 从第一列第 1 个元素到第 2 个元素, 所以当 $\delta(T) = 0$ 时 $z_1(T) = 0$ 和当 $\delta(T) = 1$ 时 $z_1(T) = \eta(T)$, 且 $z_2(T)$ 是 $z(T)$ 去掉 $z_1(T)$ 和 r^* 元素之后剩下的元素, 从而

$$\begin{aligned} z^T(T) M^T \tilde{Q} M z(T) &= z_1^T(T) \tilde{C}^T Q_2 \tilde{C} z_1(T) + \\ & r^{*T} Q_2 r^* + z_2^T(T) \bar{M}^T \tilde{C}^T Q_2 \tilde{C} \bar{M} z_2(T) - \\ & 2r^{*T} Q_2 \tilde{C} z_1(T) - 2r^{*T} Q_2 \tilde{C} \bar{M} z_2(T) \end{aligned}$$

利用克罗内克积展开, 即 $a^T W b = (b^T \otimes a^T) \text{vec}(W)$, 定义 $U(T) = \chi^T(T) \otimes \chi^T(T)$, $V(T) = z_2^T(T) \otimes z_2^T(T)$, $W(T) = z_2^T(T) \otimes r^{*T}$ 和 $\varepsilon(T) = z_1^T(T) \tilde{C}^T Q_2 \tilde{C} z_1(T) + r^{*T} Q_2 r^* - 2r^{*T} Q_2 \tilde{C} z_1(T) + y^{*T}(T) R_2 y^*(T)$, 从而丢包 Q -函数贝尔曼方程 (44) 可表示为

$$\begin{aligned} U(T) \text{vec}(H_2) &= \varepsilon(T) + V(T) \text{vec}(\bar{M}^T \tilde{C}^T Q_2 \tilde{C} \bar{M}) - \\ & 2W(T) \text{vec}(Q_2 \tilde{C} \bar{M}) + \bar{\gamma} U(T+1) \text{vec}(H_2) \end{aligned}$$

或者等价于

$$\begin{bmatrix} U(T) - \bar{\gamma} U(T+1) & -V(T) & 2W(T) \end{bmatrix} \times \begin{bmatrix} \text{vec}(H_2) \\ \text{vec}(\bar{M}^T \tilde{C}^T Q_2 \tilde{C} \bar{M}) \\ \text{vec}(Q_2 \tilde{C} \bar{M}) \end{bmatrix} = \varepsilon(T)$$

定义

$$\begin{aligned} \sigma(T) &= \begin{bmatrix} U(T) - \bar{\gamma} U(T+1) & -V(T) & 2W(T) \\ \vdots & \vdots & \vdots \\ U(T+s) - \bar{\gamma} U(T+s+1) & -V(T+s) & 2W(T+s) \end{bmatrix} \cong \\ & \begin{bmatrix} \bar{\sigma}(T) & 0 \end{bmatrix} = \sigma(T) N \end{aligned}$$

$$\xi(T) = \begin{bmatrix} \varepsilon(T) \\ \vdots \\ \varepsilon(T+s) \end{bmatrix}$$

其中, N 为列初等变换矩阵, s 是依赖于反馈丢包的连续最大丢包次数 $\delta_{f \max}$ 的整数. 使用最小二乘的方法, 需要满足秩条件

$$\text{rank} \{ \sigma^T(T) \sigma(T) \} = S_r$$

其中, $S_r = \sum_{i=0}^{\delta_{f \max}} (4+i) \times (5+i)/2 - 3\delta_{f \max} + (2 + \delta_{f \max}) \times (3 + \delta_{f \max})/2 + (2 + \delta_{f \max})$.

丢包 Q -函数的贝尔曼方程变为

$$\sigma(T) \begin{bmatrix} \text{vec}(H_2) \\ \text{vec}(\bar{M}^T \tilde{C}^T Q_2 \tilde{C} \bar{M}) \\ \text{vec}(Q_2 \tilde{C} \bar{M}) \end{bmatrix} = \xi(T)$$

或者等价于

$$\bar{\sigma}(T) \text{vec}(\bar{H}_2) = \xi(T)$$

其中,

$$\begin{bmatrix} \text{vec}(\bar{H}_2) \\ \text{vec}(\hat{H}_2) \end{bmatrix} = N^{-1} \begin{bmatrix} \text{vec}(H_2) \\ \text{vec}(\bar{M}^T \tilde{C}^T Q_2 \tilde{C} \bar{M}) \\ \text{vec}(Q_2 \tilde{C} \bar{M}) \end{bmatrix}$$

最终可得

$$\text{vec}(\bar{H}_2) = (\bar{\sigma}^T(T) \bar{\sigma}(T))^{-1} \bar{\sigma}^T(T) \xi(T)$$

为实现底流矿浆浓度的跟踪, 为底流流量过程提供最优设定值 $y^*(T)$. 使用策略迭代的方法利用在网络环境下采集到的系统数据 $\eta_f(T)$ 在线解决 Q -函数, 其算法如下所示.

算法 2. 底流流量过程最优设定值 $y^*(T)$ 的丢包 Q -学习算法

初始化: 给定初始稳定的控制策略 \tilde{L}^1 , 依次重复下面两个步骤直到控制输入收敛.

1) 策略评估: 利用最小二乘的方法计算出 \tilde{H}_2^{j+1}

$$\begin{bmatrix} U(T) - \bar{\gamma}U(T+1) & -V(T) & 2W(T) \end{bmatrix} \\ NN^{-1} \begin{bmatrix} \text{vec}(H_2) \\ \text{vec}(\tilde{M}^T \tilde{C}^T Q_2 \tilde{C} \tilde{M}) \\ \text{vec}(Q_2 \tilde{C} \tilde{M}) \end{bmatrix}^{j+1} = \\ z_1^T(T) \tilde{C}^T Q_2 \tilde{C} z_1(T) + r^{*T} Q_2 r^* - \\ 2r^{*T} Q_2 \tilde{C} z_1(T) + (y^{*j}(T))^T R_2 y^{*j}(T)$$

2) 策略提升:

$$y^{*j+1}(T) = -(H_{r_1 z}^{j+1})^{-1} H_{r_1 z}^{j+1} z(T)$$

注 3. 根据参考文献 [26], 选取比较大的半正定矩阵和合适的衰减因子能够得到比较小的跟踪误差.

注 4. 本文为双层架构的控制算法, 首先以底层和运行层稳定的控制策略运行, 运行层为底层提供设定值, 底层通过算法 1 不依赖于系统的模型参数计算得到最优的控制策略; 然后在底层稳定的情况下, 再通过算法 2 为底层提供最优的控制设定值. 此算法不需要知道系统的模型.

注 5. 算法 1 和算法 2 都需要持续激励的条件, 从而对状态空间进行充分的探索得到足够充足的数据. 如果状态收敛到期望位置, 持续激励的条件就不再需要. 可以在控制输入中加入探测噪声从而确保持续激励的条件, 此处探测噪声选择为白噪声.

3 控制性能分析

将控制策略 (32) 代入系统 (27) 得到闭环系统:

$$\eta(T+1) = (\tilde{A} + \tilde{B}L_1)\eta(T) + \tilde{B}L_2 r^* = \\ A_p \eta(T) + B_p r^*$$

如果闭环系统 A_p 的特征值在单位圆内, 则闭环系统是稳定的.

假设 λ 是闭环系统 A_p 的一个特征值, 可知满足 $A_p x_\lambda = \lambda x_\lambda$ (x_λ 是矩阵 A_p 对应 λ 的一个特征向量). 根据参考文献 [26] 可得 $P_h = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}$ 的具体形式, 从而 LQT 黎卡提方程可以化为

$$\tilde{C}^T Q_2 \tilde{C} - P_{11} + \bar{\gamma} A_p^T P_{11} A_p + L_1^T R_2 L_1 = 0$$

其中, $L_1 = -(R_2 + \bar{\gamma} \tilde{B}^T P_{11} \tilde{B})^{-1} \bar{\gamma} \tilde{B}^T P_{11} \tilde{A}$, 对其左乘 x_λ^T 右乘 x_λ 可得

$$x_\lambda^T \tilde{C}^T Q_2 \tilde{C} x_\lambda - x_\lambda^T P_{11} x_\lambda + \bar{\gamma} |\lambda|^2 x_\lambda^T P_{11} x_\lambda + \\ \bar{\gamma}^2 x_\lambda^T \tilde{A}^T P_{11} \tilde{B} (R_2 + \bar{\gamma} \tilde{B}^T P_{11} \tilde{B})^{-1} \times \\ R_2 (R_2 + \bar{\gamma} \tilde{B}^T P_{11} \tilde{B})^{-1} \tilde{B}^T P_{11} \tilde{A} x_\lambda = 0$$

因为 R_2 和 Q_2 是正定的且衰减因子 $0 < \bar{\gamma} < 1$, 故 $(R_2 + \bar{\gamma} \tilde{B}^T P_{11} \tilde{B})^{-1} > (R_2 + \tilde{B}^T P_{11} \tilde{B})^{-1}$, 所以可得

$$(1 - \bar{\gamma} |\lambda|^2) P_{11} - \tilde{C}^T Q_2 \tilde{C} \geq \bar{\gamma}^2 |\lambda|^2 \tilde{A}^T P_{11} \tilde{B} \times \\ (R_2 + \tilde{B}^T P_{11} \tilde{B})^{-1} R_2 (R_2 + \\ \tilde{B}^T P_{11} \tilde{B})^{-1} \tilde{B}^T P_{11} \tilde{A}$$

定义

$$G = \tilde{A}^T P_{11} \tilde{B} (R_2 + \tilde{B}^T P_{11} \tilde{B})^{-1} \times \\ R_2 (R_2 + \tilde{B}^T P_{11} \tilde{B})^{-1} \tilde{B}^T P_{11} \tilde{A},$$

所以

$$P_{11} - \tilde{C}^T Q_2 \tilde{C} > \bar{\gamma}^2 |\lambda|^2 (P_{11} + G)$$

或等价于

$$\frac{1}{\bar{\gamma}^2} (P_{11} - \tilde{C}^T Q_2 \tilde{C}) (P_{11} + G)^{-1} > |\lambda|^2 I$$

如果闭环系统 A_p 的特征值在单位圆内, 即 $|\lambda| \leq 1$, 闭环系统是稳定的. 所以当满足条件 $0 < (P_{11} - \tilde{C}^T Q_2 \tilde{C}) (P_{11} + G)^{-1} < \bar{\gamma}^2 I$, 闭环系统的稳定性成立. 采用 Q -学习流量控制器能保证内环稳定, 又因为外环给内环的设定值是有界的, 故双率控制结构下整体稳定.

为证明最优性, 设

$$U_2(i) = (r^* - r(i))^T Q_2 (r^* - r(i)) + y^{*T}(i) R_2 y^*(i)$$

则性能指标 (8) 可写成

$$J_2(T) = U_2(T) + \bar{\gamma} J_2(T+1)$$

对其左乘 $\bar{\gamma}^T$ 得到

$$\bar{\gamma}^T J_2(T) = \bar{\gamma}^T U_2(T) + \bar{\gamma}^{T+1} J_2(T+1)$$

移项可得

$$\bar{\gamma}^{T+1} J_2(T+1) - \bar{\gamma}^T J_2(T) = -\bar{\gamma}^T U_2(T) \quad (45)$$

对其两边从 T 到 ∞ 进行累加和, 得

$$\bar{\gamma}^\infty J_2(\infty) - \bar{\gamma}^T J_2(T) = -\sum_{i=T}^{\infty} \bar{\gamma}^i U_2(i) \quad (46)$$

因为 $\bar{\gamma}^\infty J_2(\infty) = 0$, 得到

$$\bar{\gamma}^T J_2(T) = \sum_{i=T}^{\infty} \bar{\gamma}^i U_2(i) \quad (47)$$

对式 (45) 的等号左边从 T 到 ∞ 累加得到

$$\bar{\gamma}^T J_2(T) = \sum_{i=T}^{\infty} [\bar{\gamma}^i J_2(i) - \bar{\gamma}^{i+1} J_2(i+1)] = \\ \sum_{i=T}^{\infty} [\bar{\gamma}^i z^T(i) \tilde{P} z(i) - \bar{\gamma}^{i+1} z^T(i+1) \tilde{P} z(i+1)]$$

将丢包形式的黎卡提方程 (38) 代入上式, 结合式 (47), 从而得到

$$\begin{aligned} \bar{\gamma}^T J_2(T) = & \bar{\gamma}^T z^T(k) \tilde{P} z(k) + \\ & \sum_{i=T}^{\infty} \left[y^*(i) + (R_2 + \bar{\gamma} B_h^T M^{*T} \tilde{P} M^* B_h)^{-1} \times \right. \\ & \left. \bar{\gamma} B_h^T M^{*T} \tilde{P} M^* A_h M z(i) \right]^T \times \\ & (R_2 + \bar{\gamma} B_h^T M^{*T} \tilde{P} M^* B_h) \times \\ & \left[y^*(i) + (R_2 + \bar{\gamma} B_h^T M^{*T} \tilde{P} M^* B_h)^{-1} \times \right. \\ & \left. \bar{\gamma} B_h^T M^{*T} \tilde{P} M^* A_h M z(i) \right]^T \end{aligned}$$

因为 $(R_2 + \bar{\gamma} B_h^T M^{*T} \tilde{P} M^* B_h)$ 是正定的, 故为了最小化性能指标, 则最优控制输入应该满足控制策略 (37).

4 仿真实验

为了验证在无线网络环境下针对浓密过程本文提出数据驱动算法的有效性, 设计的对比实验为浓度过程发生丢包时将最近的有用数据作为采集的数据采用 Q -学习的方法, 浓度过程未发生丢包时采用 Q -学习的方法和对本文方法增大权重的方法.

4.1 仿真实验参数选择

针对赤铁矿混合选别的浓密过程 (1) 和 (2), 进行本文提出的无线网络环境下增强学习控制方法的仿真实验, 根据实际混合选别过程可确定其参数如下^[9], $k_i = 0.001$, $k_1 = 1.9625$, $k_2 = 19.625$, $k_3 = 0.0049$, $k_0 = 47.97$, $h = 6$, $S = 1962.5 \text{ m}^2$, $\tau = 3.25$, $v_p = 1.825$, $D = 100000$, $\Delta\rho(t)/[g\rho(\cdot)] = 151.0748$, $\bar{K} = 1.12$.

底流矿浆流量为快过程, 其采样周期为 $k = 1 \text{ s}$, 底流矿浆浓度为慢过程, 其采样周期为 $T = 15 \text{ s}$. 在工作点处对其进行线性离散化, 则底流矿浆流量过程可以表示为

$$\begin{cases} A_1(z^{-1})y(k+1) = B_1(z^{-1})u(k) \\ \begin{cases} A_1(z^{-1}) = 1 + 0.1905z^{-1} \\ B_1(z^{-1}) = 12.6027 \end{cases} \end{cases}$$

底流矿浆浓度过程可以表示为

$$\begin{cases} A_2(z^{-1})r(T+1) = B_2(z^{-1})y(T) \\ \begin{cases} A_2(z^{-1}) = 1 - 0.3701z^{-1} \\ B_2(z^{-1}) = -0.1 \end{cases} \end{cases}$$

设底流矿浆浓度的设定值 r^* 为 31%, 运行层的最大丢包间隔为 $\delta_{f_{\max}} = 1$. 选取底层系统的初始策

略为 $K^1 = \begin{bmatrix} 0.002 & 0.06 \end{bmatrix}$ 和运行层的初始策略为 $\tilde{L}^1 = \begin{bmatrix} -0.31 & 1.3 & -0.12 & 0.3 & -0.35 & -9.9 \end{bmatrix}$, 选取底层系统的权重为 $Q_1 = 10$, $R_1 = 1$. 选取运行层的权重为 $Q_2 = 10000$, $R_2 = 1$, 且 $\gamma = 0.95$ 和 $\bar{\gamma} = 0.95$, 计算得 $H_{uX_d}^* = \begin{bmatrix} -22.8083 & -119.7113 \end{bmatrix}$, $H_{uu}^* = 1509.9$, 从而得到底层系统的最优策略为 $K^* = \begin{bmatrix} 0.0151 & 0.0793 \end{bmatrix}$, 计算得到丢包 Q -学习流量设定值控制器的 $H_{r_1z} = \begin{bmatrix} 33.4223 & -123.6960 & 12.3696 & -45.7799 & 33.3456 & 905.2622 \end{bmatrix}$ 和 $H_{r_1r_1}^* = 91.2335$, 则丢包 Q -学习流量设定值控制器的最优策略为 $\tilde{L}^* = \begin{bmatrix} -0.3663 & 1.3561 & -0.1354 & 0.5029 & -0.3661 & -9.9236 \end{bmatrix}$.

4.2 仿真结果

算法 1 经过迭代 3 次收敛, 得到 Q -学习流量控制器的 $H_{uX_d} = \begin{bmatrix} -22.8083 & -119.7113 \end{bmatrix}$, $H_{uu} = 1509.9$, 则增益 $K = \begin{bmatrix} 0.0151 & 0.0793 \end{bmatrix}$. 待 Q -学习流量控制器收敛时, 之后算法 2 迭代 6 次就能收敛得到丢包 Q -学习流量设定值控制器的 $H_{r_1z} = \begin{bmatrix} 33.4122 & -123.7178 & 12.3686 & -45.7484 & 33.3905 & 905.1541 \end{bmatrix}$ 和 $H_{r_1r_1} = 91.2114$, 从而得到 $\tilde{L} = \begin{bmatrix} -0.3663 & 1.3564 & -0.1356 & 0.5016 & -0.3661 & -9.9237 \end{bmatrix}$.

由图 3 可知, Q -学习流量控制器能使流量 $y(k)$ 跟踪丢包 Q -学习控制器提供的设定值 $y^*(T)$, 并且使浓密过程的浓度输出 $r(T)$ 跟踪浓度的设定值 r^* , 同时在系统稳定时, 矿浆泵频率 u 的输入也趋于稳定. 图 4 表明流量过程控制增益 K 在学习的过程中与最优的控制增益 K^* 差值的 2 范数逐渐变小, 且趋于 0. 图 5 为底流矿浆流量过程在学习的过程中 H 收敛到最优值 H^* . 图 6 表明浓度过程控制增益 \tilde{L} 在学习的过程中与最优的控制增益 \tilde{L}^* 差值的 2 范数逐渐变小, 且趋于 0. 图 7 为浓度过程在学习的过程中 \bar{H}_2 收敛到最优值 \bar{H}_2^* . 仿真结果表明, 本文提出的算法在不知道浓密过程的模型时, 在无线网络环境下, 只利用在线采集到的输入输出的数据就能实现最优控制.

4.3 对比实验

对比实验 1 为流量过程采取相同的控制策略, 在网络发生丢包时, 因为没有数据传输过来, 故将最

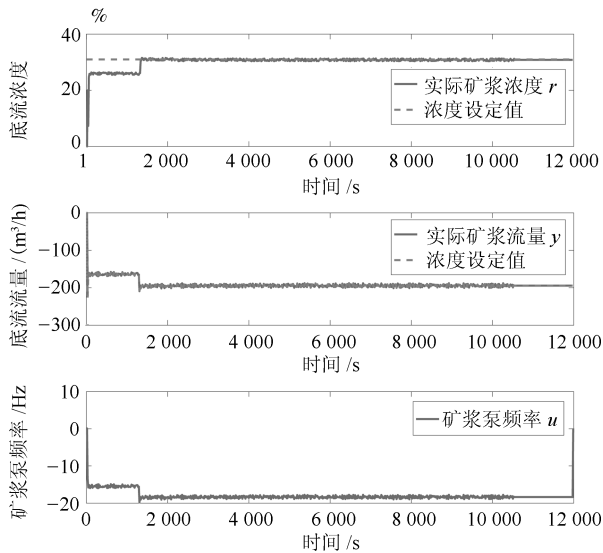


图 3 浓密过程中浓度、流量的跟踪曲线以及底流泵转速的输入的曲线

Fig. 3 The tracing result of the slurry concentration and the slurry flow-rate, and the input of the frequency of slurry pump

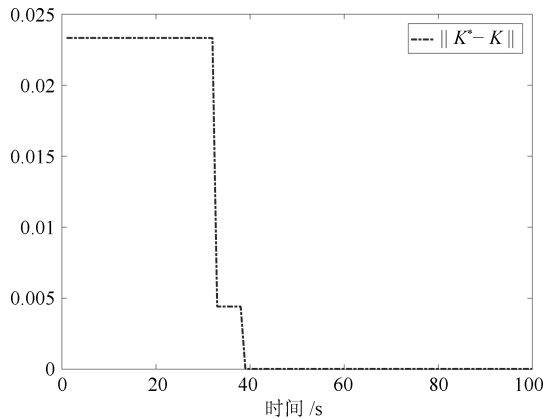


图 4 流量过程控制增益 K 的收敛过程

Fig. 4 Convergence of K to its optimal value K^*

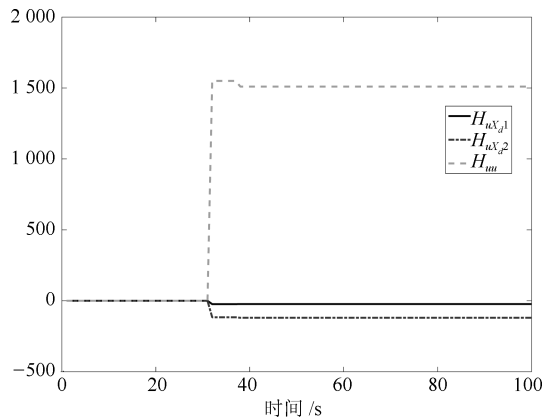


图 5 流量过程 Q -学习的结果

Fig. 5 The result of the slurry flow-rate process during the Q -learning process

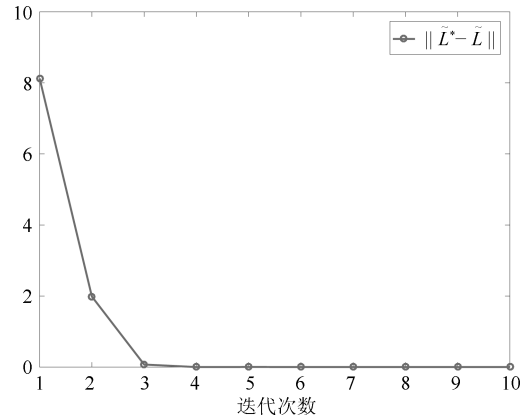


图 6 浓度过程控制增益 \tilde{L} 的收敛过程

Fig. 6 Convergence of \tilde{L} to its optimal value \tilde{L}^*

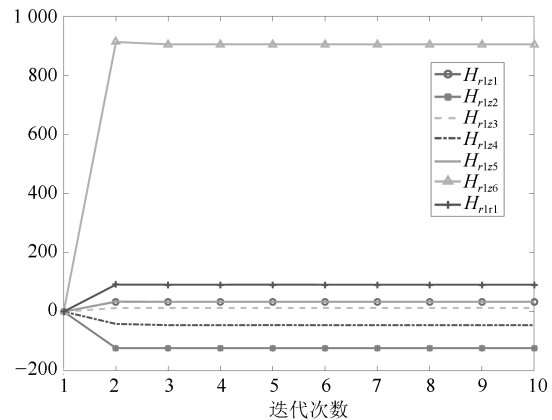


图 7 浓度过程丢包 Q -学习的结果

Fig. 7 The result of the slurry concentration process during the dropout Q -learning process

近的有用数据作为这次采集的数据,此时采用 Q -学习的方法计算得到流量设定值增益. 选取 $Q_2 = 10000$ 和 $\bar{\gamma} = 0.95$, 第一次迭代得到的增益为 $[0.0476 \quad -0.7379 \quad -4.9454]$, 第二次迭代得到的增益为 $[1.2071 \quad 4.8951 \quad -3.5455]$, 第三次迭代得到的增益为 $[0.9464 \quad 8.6217 \quad -8.5626]$, 三次迭代得到的增益不能收敛且变化大, 其作用到浓密过程得到仿真结果为图 8.

从图 8 可知将最近的有用数据作为这次采集的数据, 采用 Q -学习的方法迭代计算的三次增益分别作用到系统中不能实现对设定值的跟踪, 且随着迭代次数的增多使浓密过程越来越不稳定.

对比实验 2 为不考虑网络存在丢包的情况下, 流量过程采取相同的控制策略, 对浓度过程采取 Q -学习控制算法, 选取 $Q_2 = 10000$ 和 $\bar{\gamma} = 0.95$, 经过迭代得到 Q -学习流量设定值增益为 $[-0.3664 \quad 1.3560 \quad -9.9237]$, 其仿真结果为图 9.

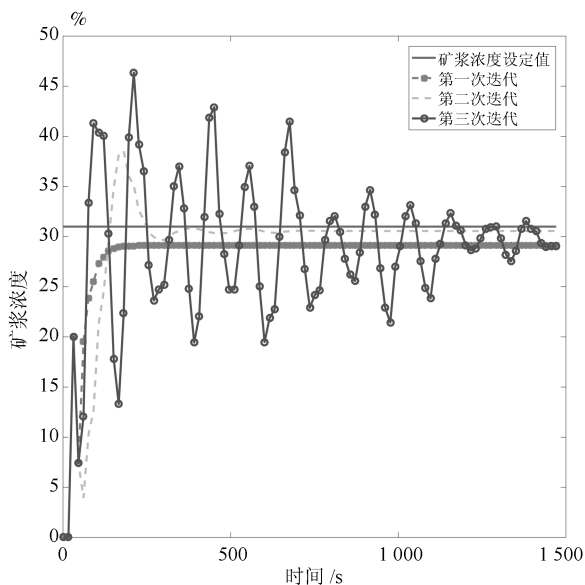


图8 浓度过程 Q-学习的结果

Fig.8 The result of the slurry concentration process during the Q-learning process

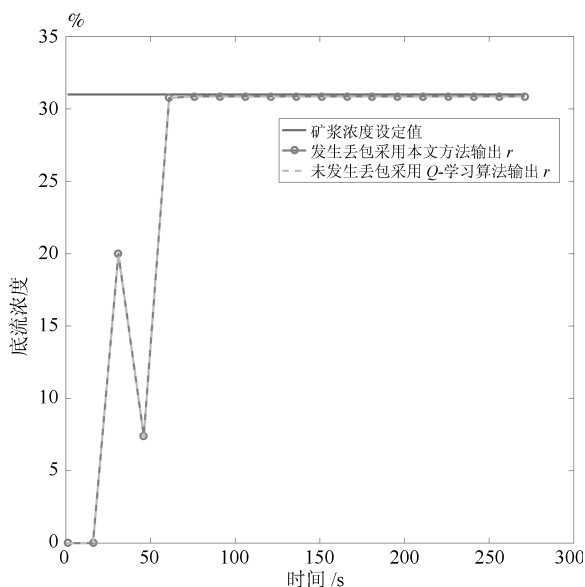


图9 对比实验 2 的仿真结果图

Fig.9 The result of experiment 2

对比实验 3 为流量过程采取相同的控制律, 对浓度过程的性能指标增大权重 $Q_2 = 500\ 000$, 得到的仿真结果如图 10 所示.

为了评价本文的控制效果, 采用绝对误差积分 (Integral absolute error, IAE) 与误差均方差 (Mean square error, MSE)^[16, 29], 其公式为

$$IAE = \sum_{T=1}^{T^*} |r^* - r(T)|$$

$$MSE = \sqrt{\frac{1}{T^*} \sum_{T=1}^{T^*} |r^* - r(T)|^2}$$

对比实验 2 和 3 的评价指标结果如表 2 所示.

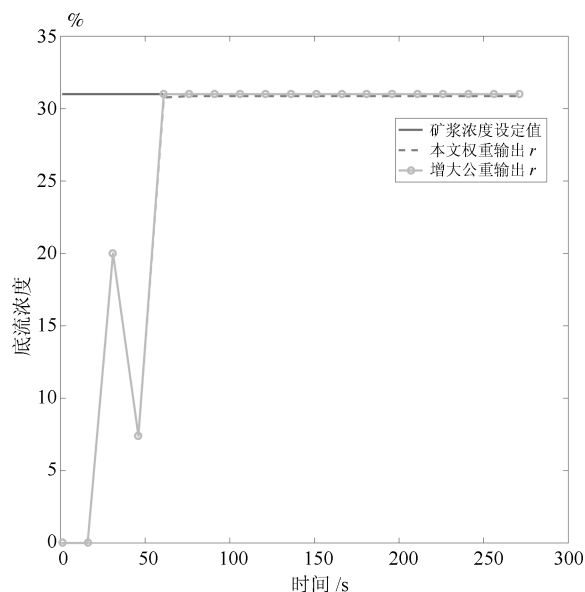


图 10 增大 Q_2 仿真结果图

Fig.10 The result of increasing Q_2

表 2 对比实验 2 和 3 评价指标

Table 2 Performance index of comparison experiment

	IAE	MSE
本文 Q_2	8.4224	0.0191
未丢包	8.4093	0.0190
增大 Q_2	0.0418	6.63×10^{-7}

从表 2 中可知, 选取相同的权重和衰减因子, 当系统发生丢包采取本文的方法得到的控制效果与未发生丢包采取 Q-学习的控制效果基本相同, 表明本文的方法对网络环境下浓密过程存在丢包的情况有效. 对本文的方法增大 Q_2 时, 浓度过程的输出能很好的跟踪浓度设定值, 余差变小且性能评价指标变优, 且本文不需要系统的模型, 在丢包时利用采集到的数据也能实现跟踪.

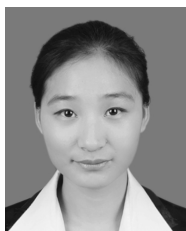
5 结论

本文针对在网络环境下的浓密过程设计 Q-学习流量控制器和丢包 Q-学习流量设定值控制器, 保证浓密过程存在网络丢包时, 不需要知道浓密过程的模型仅利用在线采集到的输入输出的数据实现能够很好地跟踪浓度设定值, 且仅依赖采集的数据为流量过程提供最优的设定值. 其仿真结果表明该方法的有效性, 保证丢包时不依赖模型参数, 仅利用采集的数据实现对底流矿浆浓度的跟踪.

References

- 1 Diehl S. A regulator for continuous sedimentation in ideal clarifier-thickener units. *Journal of Engineering Mathematics*, 2008, **60**(3–4): 265–291
- 2 Betancourt F, Bürger R, Diehl S, Farås S. Modeling and controlling clarifier-thickeners fed by suspensions with time-dependent properties. *Minerals Engineering*, 2014, **62**: 91–101
- 3 Cao X H, Cheng P, Chen J M, Sun Y X. An online optimization approach for control and communication codesign in networked cyber-physical systems. *IEEE Transactions on Industrial Informatics*, 2013, **9**(1): 439–450
- 4 Fan J L, Jiang Y, Chai T Y. MPC-based setpoint compensation with unreliable wireless communications and constrained operational conditions. *Neurocomputing*, 2017, **270**: 110–121
- 5 Fan Jia-Lu, Jiang Yi, Chai Tian-You. Operational feedback control of industrial processes in a wireless network environment. *Acta Automatica Sinica*, 2016, **42**(8): 1166–1174
(范家璐, 姜艺, 柴天佑. 无线网络环境下工业过程运行反馈控制方法. *自动化学报*, 2016, **42**(8): 1166–1174)
- 6 Sidrak Y L. Control of the thickener operation in alumina production. *Control Engineering Practice*, 1997, **5**(10): 1417–1426
- 7 Li Hai-Bo, Chai Tian-You, Zhao Da-Yong. Intelligent switching control of underflow slurry concentration and flowrate intervals in mixed separation thickener. *Acta Automatica Sinica*, 2014, **40**(9): 1967–1975
(李海波, 柴天佑, 赵大勇. 混合选别浓密机底流矿浆浓度和流量区间智能切换控制方法. *自动化学报*, 2014, **40**(9): 1967–1975)
- 8 Chai T Y, Jia Y, Li H B, Wang H. An intelligent switching control for a mixed separation thickener process. *Control Engineering Practice*, 2016, **57**: 61–71
- 9 Wang Lin-Yan, Li Jian, Jia Yao, Chai Tian-You. Dual-rate intelligent switching control for mixed separation thickening process. *Acta Automatica Sinica*, 2018, **44**(2): 330–343
(王琳岩, 李健, 贾瑶, 柴天佑. 混合选别浓密过程双速率智能切换控制. *自动化学报*, 2018, **44**(2): 330–343)
- 10 Jiang Y, Fan J L, Chai T Y, Chen T W. Setpoint dynamic compensation via output feedback control with network induced time delays. In: Proceedings of the 2015 American Control Conference (ACC). Chicago, IL, USA: IEEE, 2015. 5384–5389
- 11 Schenato L, Sinopoli B, Franceschetti M, Poolla K, Sastry S S. Foundations of control and estimation over lossy networks. *Proceedings of the IEEE*, 2007, **95**(1): 163–187
- 12 Sinopoli B, Schenato L, Franceschetti M, Poolla K, Jordan M I, Sastry S S. Kalman filtering with intermittent observations. *IEEE Transactions on Automatic Control*, 2004, **49**(9): 1453–1464
- 13 Shi Y, Yu B. Robust mixed H_2/H_∞ control of networked control systems with random time delays in both forward and backward communication links. *Automatica*, 2011, **47**(4): 754–760
- 14 Zhang H, Shi Y, Wang J M. Observer-based tracking controller design for networked predictive control systems with uncertain Markov delays. *International Journal of Control*, 2013, **86**(10): 1824–1836
- 15 Zhang J H, Lin Y J, Shi P. Output tracking control of networked control systems via delay compensation controllers. *Automatica*, 2015, **57**: 85–92
- 16 Jiang Y, Fan J L, Chai T Y, Li J N, Lewis F L. Data-driven flotation industrial process operational optimal control based on reinforcement learning. *IEEE Transactions on Industrial Informatics*, 2018, **14**(5): 1974–1989
- 17 Gao W N, Jiang Z P, Lewis F L, Wang Y B. Leader-to-formation stability of multi-agent systems: an adaptive optimal control approach. *IEEE Transactions on Automatic Control*, 2018, **63**(10): 3581–3587
- 18 Gao W N, Jiang Z P. Learning-based adaptive optimal tracking control of strict-feedback nonlinear systems. *IEEE Transactions on Neural Networks and Learning Systems*, 2018, **29**(6): 2614–2624
- 19 Xu H, Sahoo A, Jagannathan S. Stochastic adaptive event-triggered control and network scheduling protocol co-design for distributed networked systems. *IET Control Theory & Applications*, 2014, **8**(18): 2253–2265
- 20 Xu H, Jagannathan S, Lewis F L. Stochastic optimal control of unknown linear networked control system in the presence of random delays and packet losses. *Automatica*, 2012, **48**(6): 1017–1030
- 21 Xu H, Jagannathan S. Stochastic optimal controller design for uncertain nonlinear networked control system via neuro dynamic programming. *IEEE Transactions on Neural Networks and Learning Systems*, 2013, **24**(3): 471–484
- 22 Kim B H, Klima M S. Development and application of a dynamic model for hindered-settling column separations. *Minerals Engineering*, 2004, **17**(3): 403–410
- 23 Zheng Y Y. Mathematical Mode of Anaerobic Processes Applied to the Anaerobic Sequencing Batch Reactor [Ph.D. dissertation], University of Toronto, Canada, 2003
- 24 Jiang Y, Fan J L, Chai T Y, Lewis F L, Li J N. Tracking control for linear discrete-time networked control systems with unknown dynamics and dropout. *IEEE Transactions on Neural Networks and Learning Systems*, 2018, **29**(10): 4607–4620
- 25 Jiang Yi. Operational Feedback Multi-rate Interval Switch Control of Flotation Processes [Master thesis], Northeastern University, China, 2016
(姜艺. 浮选过程运行反馈双率区间切换控制方法 [硕士学位论文], 东北大学, 中国, 2016)
- 26 Kiumarsi B, Lewis F L, Modares H, Karimpour A, Naghibi-Sistani M B. Reinforcement Q -learning for optimal tracking control of linear discrete-time systems with unknown dynamics. *Automatica*, 2014, **50**(4): 1167–1175
- 27 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 -infinity control. *Automatica*, 2007, **43**(3): 473–481

- 28 Gao W N, Huang M Z, Jiang Z P, Chai T Y. Sampled-data-based adaptive optimal output-feedback control of a 2-degree-of-freedom helicopter. *IET Control Theory & Applications*, 2016, **10**(12): 1440–1447
- 29 Jiang Yi, Fan Jia-Lu, Jia Yao, Chai Tian-You. Data-driven flotation process operational feedback decoupling control. *Acta Automatica Sinica*, 2019, **45**(4): 759–770
(姜艺, 范家璐, 贾瑶, 柴天佑. 数据驱动的浮选过程运行反馈解耦控制方法. *自动化学报*, 2019, **45**(4): 759–770)



吴倩 东北大学流程工业综合自动化国家重点实验室硕士研究生. 主要研究方向为工业过程运行控制, 网络控制, 强化学习. E-mail: wuqian_neu@163.com
(**WU Qian** Master student at the State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University. Her research

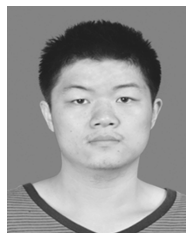
interest covers industrial process operational control, networked control and reinforcement learning.)



范家璐 东北大学流程工业综合自动化国家重点实验室副教授. 2011 年获得浙江大学博士学位 (与美国宾夕法尼亚州立大学联合培养). 主要研究方向为工业过程运行控制, 工业无线传感器网络与移动社会网络. 本文通信作者.
E-mail: jlfan@mail.neu.edu.cn

(**FAN Jia-Lu** Associate professor at the State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University. She received her Ph.D. degree from Zhejiang University in 2011. She was a Visiting Scholar with the Pennsylvania State Univer-

sity during 2009~2010. Her research interest covers networked operational control, industrial wireless sensor networks and mobile social networks. Corresponding author of this paper.)



姜艺 东北大学流程工业综合自动化国家重点实验室博士研究生. 2016 年获得东北大学控制理论与控制工程硕士学位. 主要研究方向为工业过程运行控制, 网络控制, 自适应动态规划, 强化学习.

E-mail: JY369356904@163.com

(**JIANG Yi** Ph. D. candidate at the State Key Laboratory of Synthetical

Automation for Process Industries, Northeastern University. He received his master degree in control theory and engineering from Northeastern University in 2016. His research interest covers industrial process operational control, networked control, adaptive dynamic programming and reinforcement learning.)



柴天佑 中国工程院院士, 东北大学教授, IEEE Fellow, IFAC Fellow. 1985 年获得东北大学博士学位. 主要研究方向为自适应控制, 智能解耦控制, 流程工业综合自动化理论、方法与技术.

E-mail: tychai@mail.neu.edu.cn

(**CHAI Tian-You** Academician of Chinese Academy of Engineering, professor at Northeastern University, IEEE Fellow, IFAC Fellow. He received his Ph. D. degree from Northeastern University in 1985. His research interest covers adaptive control, intelligent decoupling control, and integrated automation theory, method and technology of industrial process.)