

控制与决策

Control and Decision

基于局部信息增量与MPLS的质量相关故障检测方法

孔祥玉, 解建, 罗家宇, 杜柏阳, 李强

引用本文:

孔祥玉, 解建, 罗家宇, 等. 基于局部信息增量与MPLS的质量相关故障检测方法[J]. *控制与决策*, 2021, 36(7): 1647–1654.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2019.1402>

您可能感兴趣的其他文章

Articles you may be interested in

[基于自适应混合核典型变量分析的工业过程质量相关故障检测](#)

Quality-related fault detection for industrial processes based on adaptive mixed kernel canonical variable analysis

控制与决策. 2021, 36(4): 801–807 <https://doi.org/10.13195/j.kzyjc.2020.0690>

[基于多块信息提取的AUV资源勘查系统故障检测](#)

Fault detection of AUV resource exploration system based on multi-block information extraction

控制与决策. 2021, 36(4): 790–800 <https://doi.org/10.13195/j.kzyjc.2019.0732>

[基于矩阵的双论域模糊概率粗糙集增量更新算法](#)

Incremental updating of fuzzy probability rough sets over two universes based on matrix method

控制与决策. 2021, 36(3): 553–564 <https://doi.org/10.13195/j.kzyjc.2019.0692>

[基于交叉熵的改进NPE间歇过程故障检测算法](#)

Improved NPE batch process fault detection algorithm based on cross entropy

控制与决策. 2021, 36(2): 411–417 <https://doi.org/10.13195/j.kzyjc.2019.0725>

[基于不变网络模型和故障注入的分布式信息系统故障溯源方法](#)

Fault source location algorithm for distributed information system based on invariant network and fault injection

控制与决策. 2020, 35(11): 2723–2732 <https://doi.org/10.13195/j.kzyjc.2019.0214>

基于局部信息增量与MPLS的质量相关故障检测方法

孔祥玉[†], 解建, 罗家宇, 杜柏阳, 李强

(火箭军工程大学 导弹工程学院, 西安 710025)

摘要: 在工业生产中, 对系统进行故障检测具有十分重要的作用. 改进的偏最小二乘(modified partial least squares, MPLS)是在PLS基础上提出的一种扩展算法, 在质量相关故障检测中具有良好的检测效果, 但当测试数据中含有质量无关故障时, MPLS算法漏报率较高. 另外, MPLS算法的阈值为固定值会导致其误报率增加, 这些问题会对工业过程监控产生较大影响. 鉴于此, 提出一种基于局部信息增量与MPLS的质量相关故障检测方法(local information increment-MPLS, LII-MPLS). 在MPLS基础上, 通过使用局部信息增量技术对测试数据进行实时更新检测后, 质量相关故障的漏报率明显降低. 同时, 过程复杂化导致静态控制限不能满足故障检测的需求, 现有的动态控制限适用范围具有一定局限性, 因此改进静态控制限将其推广为局部动态阈值. 最后, 通过田纳西伊士曼过程(Tennessee Eastman process, TEP)仿真实验验证了所提出算法的有效性.

关键词: 故障检测; 偏最小二乘; 局部信息增量; 质量相关; 漏报率; 局部动态阈值

中图分类号: TP273

文献标志码: A

DOI: 10.13195/j.kzyjc.2019.1402

开放科学(资源服务)标识码(OSID):



引用格式: 孔祥玉, 解建, 罗家宇, 等. 基于局部信息增量与MPLS的质量相关故障检测方法[J]. 控制与决策, 2021, 36(7): 1647-1654.

Quality-related fault detection method based on local information increment and MPLS

KONG Xiang-yu[†], XIE Jian, LUO Jia-yu, DU Bo-yang, LI Qiang

(Department of Missile Engineering, Rocket Force University of Engineering, Xi'an 710025, China)

Abstract: Fault detection of the system has a very important role in industrial production. Modified partial least squares (MPLS) is an extended algorithm based on PLS, which has a good detection effect in quality-related fault detection. However, when the test data contains quality-unrelated faults, the MPLS algorithm has a high fault missed alarm rates. In addition, the fault false alarm rates of the MPLS will increase because of its static threshold, and these problems have a great influence on industrial process monitoring. To this end, this paper proposes a quality-related fault detection method based on local information increment and MPLS (LII-MPLS). On the basis of the MPLS, the fault missed alarm rates of quality-related fault is significantly reduced by using local information incremental technology to update and detect the test data in real time. Meanwhile, the complexity of the process results in static control limits that cannot meet the needs of fault detection and existing dynamic control limits have certain limitations, therefore, this paper improves the static control limit and generalizes it as a local dynamic threshold. Finally, the effectiveness of the proposed approach is verified on an industrial benchmark of Tennessee Eastman process.

Keywords: fault detection; partial least squares; local information increment; quality-related; fault missed alarm rates; local dynamic threshold

0 引言

基于数据驱动^[1-4]的多元统计过程监控^[5-8](Multivariate statistical process monitoring, MSPM)方法近年来迅速发展, 其优点在于可以处理一类具有高维特征和高度复杂特性的海量数据, 而在现代工业

过程一系列生产中的变量往往都同时具备上述特点, 这些设备和系统为满足现代工业过程生产的智能化、一体化和自动化需求, 集成了大量的各型传感器用来实时监测工业过程健康状态. 然而, 在现代工业生产中人们普遍关注产品质量这一关键性能指标, 对于影

收稿日期: 2019-10-07; 修回日期: 2020-03-05.

基金项目: 国家自然科学基金项目(61673387, 61833016, 61903375).

责任编辑: 吴立刚.

[†]通讯作者. E-mail: xiangyukong01@163.com.

响产品质量的一类故障被认为是质量相关故障. 实际应用中, 为了方便预知产品质量趋势, 及时发现故障并报警, 避免造成巨大的经济损失, 质量相关故障的监测具有重要意义. 相对于质量无关故障, 因为不会影响最终的产品质量, 所以对产品质量的影响可以忽略不计. 因此, 为了确保设备和系统能够正常运行并且当出现质量相关故障时可以快速监测报警, 需要设计相应的质量相关故障检测方法.

PLS是一种能够按照质量相关性将输入 X 分解为质量相关子空间和质量无关子空间的算法, 即主元子空间 \hat{X} 和残差子空间 \tilde{X} . 其存在两个缺点: 1) \hat{X} 中包含与 Y 正交的成分; 2) \tilde{X} 内部有较大的变异^[9]. 这两个缺点导致了其质量相关故障检测率较低. 针对缺点1), Yin等^[10]提出使用正交分解的MPLS算法, 消除了 \hat{X} 中对预测质量信息无用的变量. 为了解决该问题, Trygg等^[11]提出一种正交信号校正的预处理方法, 结合PLS解决了主元子空间中含有与质量变量正交成分的问题. 针对缺点2), 为了更清楚地描述输入空间, Zhou等^[12]提出了全潜结构投影(total projection to latent structures, T-PLS), 这种处理方法对PLS进一步分解并设计4个相应的统计量, 提高了对过程变量的解释能力. 由于T-PLS算法在故障幅值增大时不能降低质量无关故障的误报率, Wang等^[13]又提出OSC-MPLS算法, 结合正交信号校正的预处理算法与MPLS后处理方法解决了T-PLS算法的不足.

PLS算法在对过程数据进行分解时使用的是斜交分解, 该分解方式的优势在于可以分离出质量相关信息和质量无关信息, 但其缺陷在于分解后的主元子空间 \hat{X} 内含有与质量无关的信息, 导致在对主元子空间进行检测时, 质量无关信息的存在使得算法性能降低. 不同于PLS分解方式的是, MPLS算法对过程数据执行正交分解, 其优点在于MPLS主元子空间 \hat{X} 只含有质量相关信息, 相比PLS算法检测率明显提高. 缺点是对空间进行分解后使用传统的统计量和控制限设计方法, 该方法不同于局部信息增量技术使用固定的窗口实时更新协方差矩阵, 所以MPLS算法的统计量对数据特征提取不明显, 导致检测率较低, 静态控制限也会导致误报率较高.

针对上述问题, 本文提出一种基于局部信息增量^[14]与MPLS的质量相关故障检测方法. 首先利用正常数据建立MPLS模型, 获得模型参数; 然后将测试数据 X 分解为质量相关和质量无关空间, 在质量相关空间中求出各采样时刻的均值向量和局部协方差矩阵, 根据求得的局部协方差矩阵求出局部信息增

量均值; 最后设计基于局部信息增量均值的统计量和控制限对工业过程进行质量相关故障检测.

相比现有的多空间类PLS扩展算法, 本文通过MPLS空间分解得到质量相关子空间, 结合局部信息增量技术改进了现有算法的统计量和控制限, 解决了现有统计量对数据特征提取不明显问题. 针对传统控制限为固定值的不足, 解决了质量相关故障误报率较高的问题. 本文创新如下: 1) 用具有固定的阈值窗口更新协方差矩阵时数据有效减少, 且获取的统计量由靠近当前时刻的数据计算出来, 去除了相邻样本之间的相关性, 计算出的统计量更能反映数据实际的变化特征; 2) 不同于MPLS和OSC-MPLS, 所提出算法可以实时计算局部信息增量, 使得当含有正常数据时局部动态阈值可以根据数据的真实特性不断变化, 降低了算法的误报率. 考虑到大多数流程工业追求的是稳定的质量产出, 本文可以在实际质量相关故障检测中提供良好的故障检测率和较低的故障误报率.

1 标准PLS和MPLS算法

1.1 标准PLS算法

首先给出一对数据矩阵 X 和 Y , X 作为输入数据, Y 为输出数据, 分别具有如下形式^[10]:

$$X = \begin{bmatrix} x_1^T \\ \vdots \\ x_N^T \end{bmatrix} \in \mathbf{R}^{N \times m}, Y = \begin{bmatrix} y_1^T \\ \vdots \\ y_N^T \end{bmatrix} \in \mathbf{R}^{N \times l}, \quad (1)$$

$$x_i \in \mathbf{R}^m, y_i \in \mathbf{R}^l, i = 1, 2, \dots, N. \quad (2)$$

其中: N 为采样数, m 和 l 分别为输入变量和输出变量的个数.

将原始变量空间投影到一个不相关的潜变量子空间

$$T = [t_1 \ \dots \ t_A] \in \mathbf{R}^{N \times A}. \quad (3)$$

原始变量 X 分解为主元子空间 \hat{X} 和残差子空间 \tilde{X} , Y 分解为可预测部分 \hat{Y} 与不可预测部分 \tilde{Y} , 其外部模型^[15]如下所示:

$$\begin{cases} X = TP^T + \tilde{X} = \hat{X} + \tilde{X}, \\ Y = TQ^T + \tilde{Y} = \hat{Y} + \tilde{Y}. \end{cases} \quad (4)$$

其中: A 为主元个数, 由交叉验证得到. P 、 Q 分别为 X 、 Y 的负载矩阵. PLS建模的各个参数可以用非线性迭代偏最小二乘算法(nonlinear iterative partial least squares, NIPALS)^[16]求得.

1.2 MPLS算法

MPLS算法给出期望输出变量 Y 的分解^[10]如下:

$$Y = XM + E_y = \hat{Y} + E_y. \quad (5)$$

由于MPLS对Y执行的是正交分解,式(5)第1部分 \hat{Y} 与X相关,第2部分 E_y 与X正交. 则有

$$\text{cov}(e_y, x) = \varepsilon\{e_y x^T\} = 0. \quad (6)$$

基于式(5)和(6),可以推导出如下关系^[9]:

$$\frac{1}{N}Y^T X = \frac{1}{N}M^T X^T X + \frac{1}{N}E_y^T X \approx M^T \frac{X^T X}{N}, \quad (7)$$

则X与Y之间的相关矩阵为

$$M = (X^T X)^\dagger X^T Y. \quad (8)$$

其中, $(X^T X)^\dagger$ 是由于 X^T 与X的乘积非满秩而对 $X^T X$ 进行奇异值分解(singular value decomposition, SVD)^[17]求得的. 计算出总体参数的最小二乘估计量M后,对 MM^T 再进行一次SVD分解得到正交投影算子. 将X沿着正交投影算子分别向 $\text{span}\{M\}$ 和 $\text{span}\{M\}^\perp$ 上投影,得到最终的MPLS外部模型为

$$\begin{cases} X = \hat{X} + \tilde{X}, \\ Y = XM + E_y. \end{cases} \quad (9)$$

2 移动窗口协方差矩阵的局部信息增量技术

局部数据移动窗口协方差矩阵的信息增量技术是在全局协方差矩阵信息增量故障检测算法^[18]的基础上改进而来. 该方法主要通过移动窗口将全体采样数据的全局协方差矩阵改为局部协方差矩阵,从而求出局部信息增量矩阵和局部信息增量均值. 其具体思想如下^[14].

给定一组观测数据

$$X_n = \begin{bmatrix} x_1(1) & x_1(2) & \dots & x_1(N) \\ x_2(1) & x_2(2) & \dots & x_2(N) \\ \vdots & \vdots & \ddots & \vdots \\ x_p(1) & x_p(2) & \dots & x_p(N) \end{bmatrix} \in \mathbf{R}^{p \times N}. \quad (10)$$

其中:p为采样变量个数,N为每个变量的采样数. 对采样数据进行预处理,求得 X_n 的均值向量

$$b_n = \frac{1}{n}X_n l_n, \quad (11)$$

其中 $l_n = [1, 1, \dots, 1]^T \in \mathbf{R}^{N \times 1}$. 将原始数据进行预处理得到

$$X_n^l = X_n - b_n l_n^T. \quad (12)$$

之后,从正常数据集中选取固定窗口长度为L的采样数据作为局部数据矩阵,有

$$X_n^L = [X(i_{n'-L+1}), X(i_{n'-L+2}), \dots, X(i_{n'})], \quad (13)$$

其中 $i_{n'}$ 为正常采样数据中的某一时刻,在第n+1个

时刻的样本到来之后,形成的局部数据矩阵变为

$$X_{n+1}^L = [X(i_{n'-L+2}), \dots, X(i_{n'}), X(n+1)]. \quad (14)$$

通过对比式(13)与(14)发现,二者公共部分所形成的局部数据矩阵为

$$Y_{n,n+1}^L = [X(i_{n'-L+2}), \dots, X(i_{n'})]. \quad (15)$$

由式(15)可以得到各采样时刻所定义的均值向量为

$$y_{n,n+1}^L = \frac{1}{L-1}Y_{n,n+1}^L i_n, \quad (16)$$

其中 $i_n = [1, 1, \dots, 1]^T \in \mathbf{R}^{(L-1) \times 1}$. 为了计算方便,定义

$$K_n^L = X_n^L (X_n^L)^T. \quad (17)$$

由此得到在n时刻的局部协方差矩阵为

$$R_n^L = \frac{K_n^L}{L-1} - \frac{L b_n^L (b_n^L)^T}{L-1}, \quad (18)$$

其中 b_n^L 为n时刻的均值向量,定义为

$$b_n^L = \frac{(L-1)y_{n,n+1}^L + X(i_{n'-L+1})}{L}. \quad (19)$$

类似于式(17),可以定义在n+1时刻有

$$K_{n+1}^L = X_{n+1}^L (X_{n+1}^L)^T, \quad (20)$$

即 $K_{n+1}^L = Y_{n,n+1}^L (Y_{n,n+1}^L)^T + X(n+1)X(n+1)^T$,那么在n+1时刻的局部协方差矩阵为

$$R_{n+1}^L = \frac{K_{n+1}^L}{L-1} - \frac{L b_{n+1}^L (b_{n+1}^L)^T}{L-1}. \quad (21)$$

由式(18)和(21)可以得到局部信息增量矩阵为

$$D_{n+1}^L = R_{n+1}^L - R_n^L. \quad (22)$$

至此可得到局部数据矩阵的局部信息增量均值如下所示:

$$\lambda_{n+1}^L = \frac{\sum_{i=1}^p \sum_{j=1}^p |D_{n+1}^L[i, j]|}{p^2}. \quad (23)$$

3 基于LII-MPLS的质量相关故障检测方法

基于LII-MPLS的质量相关故障检测方法详细步骤如下.

step 1: 对过程数据X和质量数据Y进行标准化处理后对 $X^T X$ 进行SVD分解,进而求出 $(X^T X)^\dagger$,有

$$X^T X = \begin{bmatrix} P_{pc} & P_{res} \end{bmatrix} \begin{bmatrix} A_{pc} & \mathbf{0} \\ \mathbf{0} & A_{res} \end{bmatrix} \begin{bmatrix} P_{pc}^T \\ P_{res}^T \end{bmatrix}, \quad (24)$$

$$(X^T X)^\dagger = P_{pc} A_{pc}^{-1} P_{pc}^T. \quad (25)$$

step 2: 由式(26)求解 M_b ,即

$$M_b = (X^T X)^\dagger X^T Y. \quad (26)$$

注意到, M_b 由下式推导而得:

$$\begin{aligned} \frac{1}{N} Y^T X &= \frac{1}{N} (\hat{Y} + E_y)^T X = \\ &= \frac{1}{N} M_b^T X^T X + \frac{1}{N} E_y^T X \approx \\ &= M_b^T \frac{X^T X}{N}. \end{aligned} \quad (27)$$

step 3: 利用下式对 $M_b M_b^T$ 进行一次SVD分解得到 $\hat{\Gamma}_\psi \in \mathbf{R}^{m \times l}$, $\tilde{\Gamma}_\psi \in \mathbf{R}^{m \times (m-l)}$ 和 $\Lambda_\psi \in \mathbf{R}^{l \times l}$, 即

$$M_b M_b^T = [\hat{\Gamma}_\psi \quad \tilde{\Gamma}_\psi] \begin{bmatrix} \Lambda_\psi & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \hat{\Gamma}_\psi^T \\ \tilde{\Gamma}_\psi^T \end{bmatrix}. \quad (28)$$

step 4: 根据上述SVD分解得到的结果求得正交投影算子

$$\Xi_\psi = \hat{\Gamma}_\psi \hat{\Gamma}_\psi^T \in \mathbf{R}^{m \times m}, \quad (29)$$

$$\Xi_\psi^\perp = \tilde{\Gamma}_\psi \tilde{\Gamma}_\psi^T \in \mathbf{R}^{m \times m}. \quad (30)$$

step 5: 将测试数据 X_n 投影到 Ξ_ψ 和 Ξ_ψ^\perp 上, 计算出两个互相正交的空间 \hat{X}_n 和 \tilde{X}_n 分别为

$$\hat{X}_n = X_n \Xi_\psi = T_{\hat{X}_n} \hat{\Gamma}_\psi^T \in S_{\hat{X}_n} \equiv \text{span}\{M_b\}, \quad (31)$$

$$\tilde{X}_n = X_n \Xi_\psi^\perp = T_{\tilde{X}_n} \tilde{\Gamma}_\psi^T \in S_{\tilde{X}_n} \equiv \text{span}\{M_b\}^\perp, \quad (32)$$

其中 $T_{\hat{X}_n} = X_n \hat{\Gamma}_\psi$, $T_{\tilde{X}_n} = X_n \tilde{\Gamma}_\psi$ 分别为 \hat{X}_n 和 \tilde{X}_n 的得分矩阵.

step 6: 求得 \hat{X}_n 的均值向量 c_n , 原始数据预处理得到

$$c_n = \frac{1}{n} \hat{X}_n l_n, \quad (33)$$

$$\hat{X}_n^l = \hat{X}_n - c_n l_n^T, \quad (34)$$

其中 $l_n = [1, 1, \dots, 1]^T \in \mathbf{R}^{N \times 1}$.

step 7: 假定当前时刻为 n , 从正常数据集中选取固定窗口长度为 L 的采样数据作为 n 时刻的局部数据矩阵, 有

$$\hat{X}_n^L = [\hat{X}(n' - L + 1), \hat{X}(n' - L + 2), \dots, \hat{X}(n')], \quad (35)$$

其中 n' 为正常采样数据中的某一时刻. 则 $n + 1$ 时刻的局部数据矩阵为

$$\hat{X}_{n+1}^L = [\hat{X}(n' - L + 2), \dots, \hat{X}(n'), \hat{X}(n + 1)]. \quad (36)$$

step 8: 求式 (35) 和 (36) 公共部分所形成的局部数据矩阵为

$$Y_{n,n+1}^L = [\hat{X}(n' - L + 2), \dots, \hat{X}(n')], \quad (37)$$

$$i_n = [1, 1, \dots, 1]^T \in \mathbf{R}^{(L-1) \times 1}. \quad (38)$$

得到各采样数据所定义的均值向量为

$$y_{n,n+1}^L = \frac{1}{L-1} Y_{n,n+1}^L i_n. \quad (39)$$

step 9: 求得 n 时刻的局部协方差矩阵为

$$R_n^L = \frac{K_n^L}{L-1} - \frac{L b_n^L (b_n^L)^T}{L-1}. \quad (40)$$

其中

$$K_n^L = \hat{X}_n^L (\hat{X}_n^L)^T,$$

$$b_n^L = \frac{(L-1)y_{n,n+1}^L + \hat{X}(n' - L + 1)}{L}.$$

step 10: 类似地, $K_{n+1}^L = \hat{X}_{n+1}^L (\hat{X}_{n+1}^L)^T$, 可得 $n + 1$ 时刻局部协方差矩阵为

$$R_{n+1}^L = \frac{K_{n+1}^L}{L-1} - \frac{L b_{n+1}^L (b_{n+1}^L)^T}{L-1}. \quad (41)$$

step 11: 由式 (40) 和 (41) 得到局部信息增量矩阵为

$$D_{n+1}^L = R_{n+1}^L - R_n^L. \quad (42)$$

step 12: 由此得到局部数据矩阵的局部信息增量均值如下:

$$\lambda_{n+1}^L = \frac{\sum_{i=1}^m \sum_{j=1}^m |D_{n+1}^L[i, j]|}{m^2}. \quad (43)$$

step 13: 根据上述公式, 计算局部正常采样数据的动态阈值为

$$\rho_{n+1}^L = \frac{1.5}{L} \sum_{s=n'-L+1}^{n'} \lambda_s^L. \quad (44)$$

其中: n' 为 n 时刻非异常数据标号, 采样数 $n > L + 2$, λ_s^L 为未发生故障的局部信息增量均值.

step 14: 设计故障判据如下:

- 1) 当 $\lambda_{n+1}^L \geq \rho_{n+1}^L$ 时, 有质量相关故障发生;
- 2) 当 $\lambda_{n+1}^L < \rho_{n+1}^L$ 时, 无质量相关故障发生.

LII-MPLS 算法步骤总结如下.

离线建模:

- step 1: 对正常数据进行标准化处理;
- step 2: 使用 MPLS 对正常数据进行离线建模;
- step 3: 得到 MPLS 的模型参数.

在线检测:

- step 4: 对测试数据进行标准化处理;
- step 5: 对标准化后的测试数据使用 LII-MPLS 算法;
- step 6: 计算基于局部信息增量的统计量和阈值.
- 故障判断:
- step 7: 对比统计量与阈值大小;
- step 8: 根据故障判据得出质量相关故障检测结果.

注 1 式 (44) 所定义的局部正常采样数据动态阈值通过求取各采样时刻的局部协方差矩阵, 然后计

算相邻时刻的局部协方差矩阵的差得到局部信息增量矩阵,并利用局部信息增量矩阵计算局部信息增量均值而得到.记发生故障时刻的局部信息增量均值全体为 $\{\lambda_1^L, \dots, \lambda_s^L, \dots, \lambda_n^L\}$,未发生故障的局部信息增量均值为 $\{\lambda_{n'+1}^L, \dots, \lambda_{n+1}^L\}$,则 s 时刻的阈值可以通过对 s 时刻附近固定窗口长度为 L 的所有局部信息增量均值求和,再求取平均值而得到;式(44)中的“1.5”为局部动态阈值的优化系数,该系数是以故障检测的误报率和漏报率达到最佳为目标函数,在离线建模时通过对大量已知样本数据使用粒子群(particle swarm optimization, PSO)智能优化算法训练得到的一个较好的参数.因为本文研究重点是如何提高算法的检测率,对于该算法本文不具体展开说明,阈值计算方法见文献[14, 18].

注2 MPLS和OSC-MPLS算法使用的统计量计算方法没有考虑样本之间的相关性,忽略了相邻样本之间的相互作用会导致提取的数据特征不明显;而传统的控制限又为固定值,不能适应数据实际情况,因此造成了较高的故障误报率.本文所提出的基于LII-MPLS的质量相关故障检测算法在对过程数据执行正交分解后,在质量相关子空间内结合局部信息增量技术设计新的统计量和控制限计算方法.将MPLS算法扩展为LII-MPLS算法的优势在于,随着采样数据的增多,本文算法可以在不断更新局部数据协方差的同时使所用数据的长度固定,通过计算局部信息增量消除前一时对下一时刻的影响,构造的统计量会更准确,更能反映数据的实际变化情况.同时,基于局部信息增量的局部动态阈值可以根据数据的

特征而变化,减小了算法的误报率.

注3 本文所提出的LII-MPLS算法step 1~step 5使用了MPLS算法,通过MPLS算法将过程变量正交分解为两个空间,即质量相关空间和质量无关空间.从step 6开始将局部信息增量技术引入到MPLS分解后的质量相关空间,重新设计了基于局部信息增量的统计量和控制限.LII-MPLS算法改进传统的统计量和控制限,可以解决如下问题:1)现有统计量设计方法不能去除连续样本之间的相关性造成提取数据特征不明显的问题;2)现有静态控制限不能适应数据实际变化情况造成误报率较高的问题.

4 TEP仿真实验

田纳西伊士曼过程(Tennessee Eastman process)是美国Tennessee Eastman化学公司过程控制部门在1993年提出的一个基于实际工业过程的仿真实例^[19].图1为TEP流程图^[20],该过程现已广泛应用于评估过程监控的性能,并取得了很好的应用效果.TEP的整个过程系统共有12个操纵变量和41个测量变量(含有22个连续变量和19个成分变量),表1和表2分别给出了12个操纵变量和41个测量变量在系统中的实际意义.

另外,该过程还包含20种扰动,其中15种扰动为已知故障,分别为IDV(1)~IDV(15).这15种已知故障中分为两种类型故障:一种是质量相关故障,如IDV(1)、IDV(2)、IDV(5)~IDV(8)、IDV(10)、IDV(12)、IDV(9)、IDV(11)、IDV(15).这些样本中的每个正常数据集包含480个样本,每个故障数据集包含960个样

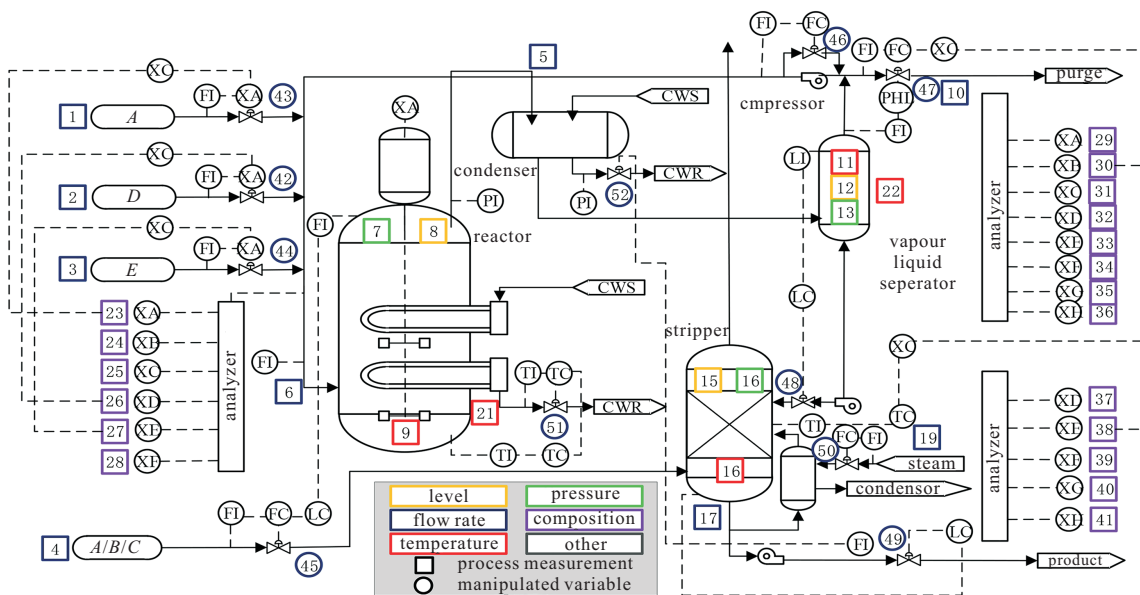


图1 TEP流程图

表1 TEP操纵变量

编号	变量名称	编号	变量名称
1	物料D的流量	7	分离器液体流量
2	物料E的流量	8	解吸塔液体流量
3	物料A的流量	9	汽提器水流阀
4	物料A和C的流量	10	反应器冷水流量
5	压缩机循环阀	11	冷凝器冷水流量
6	放空阀	12	搅拌器速度

表2 TEP测量变量

编号	变量名称	编号	变量名称
1	物料A流量	22	分离器冷却水出口温度
2	物料D流量	23	物料6中A的摩尔含量
3	物料E流量	24	物料6中B的摩尔含量
4	物料A和C流量	25	物料6中C的摩尔含量
5	循环流量	26	物料6中D的摩尔含量
6	反应器进料流量	27	物料6中E的摩尔含量
7	反应器压力	28	物料6中F的摩尔含量
8	反应器液位	29	物料9中A的摩尔含量
9	反应器温度	30	物料9中B的摩尔含量
10	放空速率	31	物料9中C的摩尔含量
11	分离器温度	32	物料9中D的摩尔含量
12	分离器液位	33	物料9中E的摩尔含量
13	分离器压力	34	物料9中F的摩尔含量
14	分离器底部流量	35	物料9中G的摩尔含量
15	汽提器液位	36	物料9中H的摩尔含量
16	汽提器压力	37	物料11中D的摩尔含量
17	汽提器底部流量	38	物料11中E的摩尔含量
18	汽提器温度	39	物料11中F的摩尔含量
19	汽提器流量	40	物料11中G的摩尔含量
20	压缩机功率	41	物料11中H的摩尔含量
21	反应器冷却水出口温度		

本,960个样本中前160个样本是正常数据.所有数据在建模之前都应该进行标准化处理,之后使用正常数据建立回归模型,故障数据集则被用来进行检测.本实验从故障数据集中选取100个正常数据样本和800个故障数据样本组成测试数据进行故障检测.

本节将LII-MPLS算法应用到TEP仿真实验上进行故障检测,并与MPLS和OSC-MPLS算法进行比较.首先对实验参数初始化:选取数据集中22个过程变量XMEAS(1~22)和11个操纵变量XMV(1~11)作为输入X,选取过程变量XMEAS(35)为质量变量Y,LII-MPLS算法中主元个数取A=9由交叉验证得到,采用PSO智能优化算法寻求最优窗长L=3.

图2和图3中 T^2 代表MPLS、OSC-MPLS算法的统计量,图4中 $T_{LII-MPLS}$ 代表LII-MPLS算法的统计

量, σ 代表阈值.图2为MPLS算法在IDV(1)下的故障检测图,图中900~1000采样区段表示当有新数据到来时算法继续进行检测的实时过程.可以看出在前100个正常采样数据中,统计量始终在阈值下方,表明无质量相关故障发生.在故障采样数据到来后,MPLS有多处统计量在阈值下方,表明MPLS在对IDV(1)检测时具有较高的漏报率.

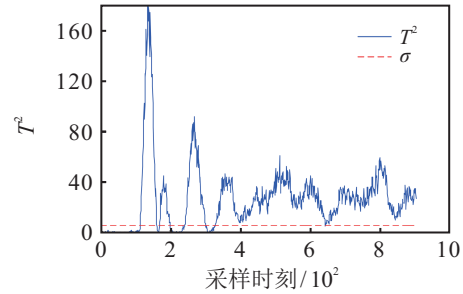


图2 MPLS的IDV(1)故障检测结果

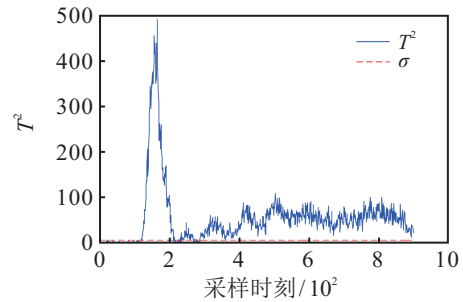


图3 OSC-MPLS的IDV(1)故障检测结果

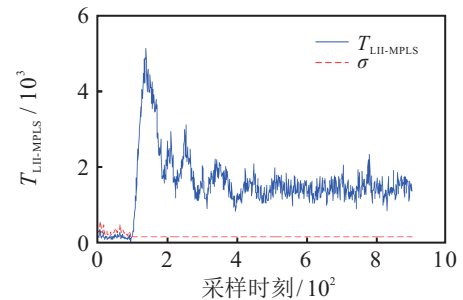


图4 LII-MPLS的IDV(1)故障检测

图3为OSC-MPLS在IDV(1)下的故障检测.由图3可见,前100个正常采样数据中OSC-MPLS存在一定的误报率,在故障采样数据到来后OSC-MPLS也存较高的漏报率.

图4为LII-MPLS在IDV(1)下的故障检测图,不同于图2、图3的MPLS和OSC-MPLS的静态控制限,图4虚线部分代表可以根据数据实际特性变化的LII-MPLS算法的局部动态阈值.可以看出,其在前100个正常采样数据中误报率为0%;在故障数据到来后准确检测到了故障并报警,后800个故障数据中漏报率很低.

表3为3种算法所有质量相关故障检测率,粗体

部分为3种算法中质量相关检测率最好的组. 由表3数据可见, 所提出的LII-MPLS算法除IDV(5)略低于MPLS和OSC-MPLS 6.62%外, 其他故障的检测率均高于MPLS和OSC-MPLS. 其中IDV(1)、IDV(2)、IDV(6)、IDV(7)、IDV(8)、IDV(12)和IDV(13)的故障检测率均达到95%以上, 所提出算法可以对所有质量相关故障准确报警.

表4为3种算法所有质量相关故障的误报率, 粗体部分为3种算法中质量相关误报率最高的组. 由

表4数据可见, MPLS的误报率较低, 仅有IDV(12)达到了6%, 而OSC-MPLS质量相关故障的误报率则相对较高, 误报率最高的一组IDV(10)达到了7%. LII-MPLS算法除IDV(7)和IDV(12)有1%的误报率外, 其他质量相关故障的误报率均为0%. 通过对3种算法多次实验得到的检测率数据进行分析可得, 所提出算法可以准确反映故障信息, 并在质量相关故障检测中伴随有正态噪声影响和较小动态特性波动条件下仍然保持良好的检测效果.

表3 3种算法的质量相关故障检测率

fault ID	故障描述 已知的过程故障	故障检测率/%		
		MPLS	OSC-MPLS	LII-MPLS
IDV(1)	A/C进料比, B组成常数(流4)	88.63	85.64	99.75
IDV(2)	B组成, A/C比率常量(流4)	91.76	94.88	100.00
IDV(5)	D进料温度(流2)	99.87	99.87	93.25
IDV(6)	反应堆冷却水入口温度	99.00	98.87	100.00
IDV(7)	冷凝器冷却水入口温度	18.85	41.19	99.63
IDV(8)	进料损失(流1)	63.42	82.77	99.50
IDV(10)	C集管压力损失-可用性降低(流4)	20.97	64.41	89.00
IDV(12)	A, B, C进料成分(流4)	80.39	98.38	99.75
IDV(13)	D进料温度(流2)	94.25	79.52	99.38

表4 3种算法的质量相关故障误报率

fault ID	故障描述 已知的过程故障	故障误报率/%		
		MPLS	OSC-MPLS	LII-MPLS
IDV(1)	A/C进料比, B组成常数(流4)	0	2.00	0
IDV(2)	B组成, A/C比率常量(流4)	0	3.00	0
IDV(5)	D进料温度(流2)	0	4.00	0
IDV(6)	反应堆冷却水入口温度	0	2.00	0
IDV(7)	冷凝器冷却水入口温度	0	6.00	1.00
IDV(8)	进料损失(流1)	0	2.00	0
IDV(10)	C集管压力损失-可用性降低(流4)	0	7.00	0
IDV(12)	A, B, C进料成分(流4)	6.00	3.00	1.00
IDV(13)	D进料温度(流2)	0	1.00	0

5 结论

本文通过将局部信息增量技术应用到MPLS上, 提出了一种基于LII-MPLS的质量相关故障检测方法. 该方法首先利用正常数据建立算法模型, 然后将测试数据分解为质量相关空间和质量无关空间, 结合局部信息增量技术进行故障检测. 通过与MPLS和OSC-MPLS的实验结果对比分析, 所提出算法不仅提高了质量相关故障的检测率, 而且其误报率也较低. 值得一提的是, LII-MPLS算法的质量无关空间虽然对于产品质量的影响可以忽略不计, 但是对质量无关空间的监控也是所提算法今后研究的一个方向.

参考文献(References)

- [1] Yin S, Gao H J, Qiu J B, et al. Fault detection for nonlinear process with deterministic disturbances: A just-in-time learning based data driven method[J]. IEEE Transactions on Cybernetics, 2017, 47(11): 3649-3657.
- [2] Yin S, Li X W, Gao H J, et al. Data-based techniques focused on modern industry: An overview[J]. IEEE Transactions on Industrial Electronics, 2015, 62(1): 657-667.
- [3] Wen L, Li X Y, Gao L, et al. A new convolutional neural network-based data-driven fault diagnosis method[J]. IEEE Transactions on Industrial Electronics, 2018, 65(7):

- 5990-5998.
- [4] Yin S, Ding S X, Haghani A, et al. A comparison study of basic data-driven fault diagnosis and process monitoring methods on the benchmark Tennessee Eastman process[J]. *Journal of Process Control*, 2012, 22(9): 1567-1581.
- [5] Ji H Q, Xiao X, Shang J, et al. Incipient fault detection with smoothing techniques in statistical process monitoring[J]. *Control Engineering Practice*, 2017, 62: 11-21.
- [6] Chen A, Zhou H, An Y, et al. PCA and PLS monitoring approaches for fault detection of wastewater treatment process[C]. *The 25th International Symposium on Industrial Electronics*. Piscataway: IEEE, 2016: 1022-1027.
- [7] Haghani A, Jeansch T, Ding S X. Quality-related fault detection in industrial multimode dynamic processes[J]. *IEEE Transactions on Industrial Electronics*, 2014, 61(11): 6446-6453.
- [8] Jiang Q, Gao F, Hui Y, et al. Multivariate statistical monitoring of key operation units of batch processes based on time-slice CCA[J]. *IEEE Transactions on Control Systems Technology*, 2018, 99: 1-8.
- [9] 孔祥玉, 曹泽豪, 安秋生, 等. 偏最小二乘线性模型及其非线性动态扩展模型综述[J]. *控制与决策*, 2018, 33(9): 1537-1548.
(Kong X Y, Cao Z H, An Q S, et al. Review of partial least squares linear models and their nonlinear dynamic expansion models[J]. *Control and Decision*, 2014, 33(9): 1537-1548.)
- [10] Yin S, Ding S X, Zhang P, et al. Study on modifications of PLS approach for process monitoring[J]. *IFAC Proceedings Volumes*, 2011, 44(1): 12389-12394.
- [11] Trygg J, Wold S. Orthogonal projections to latent structures (O-PLS)[J]. *Journal of Chemometrics*, 2002, 16(3): 119-128.
- [12] Zhou D, Li G, Qin S J. Total projection to latent structures for process monitoring[J]. *Aiche Journal*, 2010, 56(1): 168-178.
- [13] Wang G, Yin S. Quality-related fault detection approach based on orthogonal signal correction and modified PLS[J]. *IEEE Transactions on Industrial Informatics*, 2017, 11(2): 398-405.
- [14] 文成林, 胡玉成. 基于信息增量矩阵的故障诊断方法[J]. *自动化学报*, 2012, 38(5): 1-8.
(Wen C L, Hu Y C. Fault diagnosis based on information incremental matrix[J]. *Acta Automatica Sinica*, 2012, 38(5): 1-8.)
- [15] Wold H. Nonlinear iterative partial least squares (NIPALS) modelling: Some current developments[M]. New York: Academic Press, 1973: 383-407.
- [16] Björck Å, Indahl U G. Fast and stable partial least squares modelling: A benchmark study with theoretical comments[J]. *Journal of Chemometrics*, 2017, 31(8): e2898.
- [17] Xu G U, Hu G R, Song Y. Singular value decomposition (SVD)-based nonlinear filtering algorithm[J]. *Journal of Shanghai Jiaotong University*, 2005, 31(1): 1-9.
- [18] Yang H Y. Advanced prognosis and health management of aircraft and spacecraft subsystems[D]. Cambridge: Massachusetts Institute of Technology, Department of Electrical Engineering and Computer Science, 2000: 6-35.
- [19] Chiang L H, Russell E L, Braatz R D. Tennessee eastman process[C]. *Fault Detection and Diagnosis in Industrial Systems*. London: Springer, 2001: 103-112.
- [20] Jiao J F, Yu H, Wang G. A quality-related fault detection approach based on dynamic least squares for process monitoring[J]. *IEEE Transactions on Industrial Electronics*, 2015, 63(4): 2625-2632.

作者简介

孔祥玉(1967—), 男, 教授, 博士生导师, 从事随机系统分析、非线性系统建模及其应用、复杂系统故障诊断等研究, E-mail: xiangyukong01@163.com;

解建(1995—), 男, 硕士生, 从事系统特征提取、故障检测与诊断的研究, E-mail: hdxiejian@163.com;

罗家宇(1994—), 男, 硕士生, 从事数据驱动的复杂系统下故障监测与诊断的研究, E-mail: 540629964@qq.com;

杜柏阳(1993—), 男, 博士生, 从事数据驱动的复杂系统下特征提取、故障监测与诊断的研究, E-mail: duboyangpegc@163.com;

李强(1994—), 男, 硕士生, 从事复杂系统故障监测与诊断的研究, E-mail: 1058097905@qq.com.

(责任编辑: 郑晓蕾)