

加权 k 最近邻重构分析的工业过程故障诊断

王国柱^{1†}, 刘建昌¹, 李 元², 商亮亮¹

(1. 东北大学 信息科学与工程学院, 辽宁 沈阳 110819; 2. 沈阳化工大学 信息工程学院, 辽宁 沈阳 110142)

摘要: k -最近邻(k -nearest neighbor, k -NN)是一种有效的基于数据驱动的故障检测方法,该方法在工业过程监视方面已经得到了广泛的应用.但在过程中存在故障时,精确地寻找故障根源和识别故障变量是故障诊断的重要目标,也是保证工业过程安全生产的重要任务.本文在 k -NN故障检测技术的基础上,提出了一种加权的 k -NN重构方法,对使控制指标减小最大(maximize reduce index, MRI)的过程变量依次进行重构,进而确定发生故障的传感器.根据理论分析并结合数值仿真对提出的方法进行了验证,数值仿真先从精度方面验证了该方法能够有效地对故障传感器数值进行重构,然后验证了该方法不仅适用于单一传感器故障诊断,对于同时发生或者因变量相关性而传播的传感器故障也具有很好的效果.最后,该方法被成功应用于TE(Tennessee Eastman)化工过程.

关键词: 故障检测; 故障诊断; k -最近邻; 数据重构; 指标减小最大

中图分类号: TP273 文献标识码: A

Fault diagnosis of industrial processes based on weighted k -nearest neighbor reconstruction analysis

WANG Guo-zhu^{1†}, LIU Jian-chang¹, LI Yuan², SHANG Liang-liang¹

(1. College of Information Science and Engineering, Northeastern University, Shenyang Liaoning 110819, China;

2. College of Information Engineering, Shenyang University of Chemical Technology, Shenyang Liaoning 110142, China)

Abstract: The k -nearest-neighbor (k -NN) is an effective fault detection method based on data driven, which has been widely used in industrial process monitoring. However, investigating the root causes of abnormal events is a crucial task when the process faults have been detected, and isolating the faulty variables provides additional information for investigating the root causes of the faults. In this paper, a novel fault diagnosis method is derived using weighted k -NN reconstruction on maximize reduce index (MRI) variables, it can effectively identify the faulty variables and locate the faulty sensors. A numerical simulation is provided to validate the performance of weighted k -NN in the aspect of data reconstruction; it also show that this method is suitable not only for a single sensor fault, but also with good results for multiple sensor faults which are existing simultaneously or in propagation through variable correlation. Finally, this method is applied to TE chemical process successfully.

Key words: fault detection; fault diagnosis; k -NN; data reconstruction; maximize reduce index (MRI)

1 引言(Introduction)

工业过程中故障的存在不仅会影响系统的安全运行,甚至会造成产品质量下降以及人员伤亡.因此,工厂和企业对控制过程的安全性越来越重视,工业过程的故障监视与诊断技术也成为国际自控领域的热点研究方向之一^[1-3].随着自动化控制系统的日益复杂化,计算机技术和传感器技术的飞速发展,越来越多的过程数据能够被存储和应用,这也成为基于数据驱动的故障诊断技术得以发展的前提.近些年,多变量统计过程控制(multivariate statistical process control,

MSPC)方法已经被深入地研究且取得了可喜的成果,例如:主元分析(principal components analysis, PCA)、部分最小二乘(part least square, PLS)、相似性分析和 k 最近邻(k -nearest neighbor, k -NN)等检测方法及其扩展^[4-20]已经被广泛地应用于实际工业过程.

当工业过程中存在故障时,为了及时使其恢复到正常工况,故障的诊断与识别是重要的.及时定位故障变量,发现故障根源是故障检测后的另一个目的,也是从根本上消除故障的关键.其中,贡献图^[21-22]作为传统的故障诊断方法,已经得到了广泛的应用.

收稿日期: 2015-01-11; 录用日期: 2015-04-10.

[†]通信作者. E-mail: wang.guo.zhu@163.com; Tel.: +86 15998347975.

国家自然科学基金项目(61374137, 61174119, 61034006), 流程工业综合自动化国家重点实验室基础科研业务项目(2013ZCX02-03)资助.

Supported by the National Natural Science Foundation of China (61374137, 61174119, 61034006) and the State Key Laboratory of Integrated Automation of Process Industry Technology and Research Center of National Metallurgical Automation Fundamental Research Funds (2013ZCX02-03).

Kourti和Mac-Gregor^[23]将质量变量和过程变量的贡献图用于寻找高压低密度聚乙烯反应器的故障变量,经过试验,他们认为这种方法并不能一直准确地揭示故障根源;Qin和Dunia^[24]提出一种基于重构的故障诊断方法在故障子空间中隔离故障变量,该方法已经被应用于重构故障变量;Yue和Qin^[25]运用联合统计指标的故障检测方法,并在其基础上深入研究了基于贡献重构(RBC)的故障变量识别方法,同时,也说明了该方法在一定程度上和传统贡献图方法一样具有故障的拖尾效应。

近年来,基于 k -NN的故障检测方法被提出并成功应用于连续过程和间歇过程^[17,26-27]。然而,由于该方法的计算量大,需要存储的中间值多, k -NN方法对计算机的计算速度与存储空间有较高的要求。为了解决这些问题,He等^[17,26]提出了PC- k -NN方法,该方法采用原始样本的主成分作为建模样本,既降低了计算距离的运算量,又节省了存储空间,但该方法忽略了发生在残差空间的异常信息,影响了故障检测的精确性。随后,郭小萍等^[27]等人提出了一种基于特征空间 k 最近邻(FS- k -NN)的批次过程监视方法,该方法结合特征空间的主元和残差全面表示了原始数据的有用信息。尽管 k -NN方法在故障监视方面已经取得了很好的成果,但在故障诊断和故障变量识别方面少有研究。

因此,本文提出了一种新的基于加权 k -NN重构分析的故障诊断方法,该方法的发展建立在 k -NN监视算法的基础上,是 k -NN方法在故障诊断以及故障变量识别方面的进一步研究。首先,通过 k -NN故障检测方法对过程进行监视,判定过程中是否存在故障;其次,当系统中有故障存在时,采用加权 k -NN重构算法重构每一个变量并计算重构后的控制指标,依次对使控制指标减小最大的变量进行替换,迭代至控制指标在控制范围之内。该方法的目的是寻找引起过程发生故障的所有变量。该方法具有以下两个优势:

1) 对数据特性没有限制,可以有效地应用于非高斯、非线性、多工况数据;

2) 不仅仅适用于单一传感器故障,对多个传感器同时故障同样有效,能够识别到所有发生故障的传感器;对于主要故障变量和因变量相关性引起的其它故障变量,该方法也能根据故障发生时刻和控制指标减小排序做出初步的解释。本文最后采用数值仿真验证了该算法能够有效地对故障传感器数值进行重构,并且通过TE过程对所提方法进行了实验,验证了方法的有效性和可靠性。

2 基本理论(Basic theory)

2.1 k -NN基本规则(Basic rule of k -NN)

k -NN近邻法,又称 k -NN分类法,它是一种可以预测未知类或标签的非参数监督分类方法,对于未知

分布和非正态分布的数据集可以获得较高的分类准确率,具有健壮性强、概念清晰和适用非高斯、非线性数据等诸多优点^[27]。假若一个样本在特征空间中的 k 个最相似(距离最近)样本大多数都属于某一个类别,则此样本也属于这个类别;只有当新的样本与其在训练样本的前 k 个最近邻的距离超出类别属性时,称该样本与训练样本具有较大差异。目前,该思想已经被应用于工业过程故障监视。

2.2 基于 k -NN规则的故障检测(Fault detection based k -nearest neighbors rule)

基于最近邻法的过程异常检测是直接利用正常条件下的过程数据分布临近的特征进行选取临近个数据样本建立过程统计模型,提取正常样本之间的最近邻距离特征量,建立统计控制限。其指导思想是异常样本与正常样本所属类不同,异常样本与建模数据中正常样本的最近邻距离会大于正常样本之间的最近邻距离,可以表示为

$$\min d_f > \min d_n, \quad (1)$$

其中: $\min d_f$ 是异常样本与训练样本之间的最近邻距离; $\min d_n$ 是正常样本与训练样本之间的最近邻距离^[27]; k -NN故障建模与检测步骤见图1。

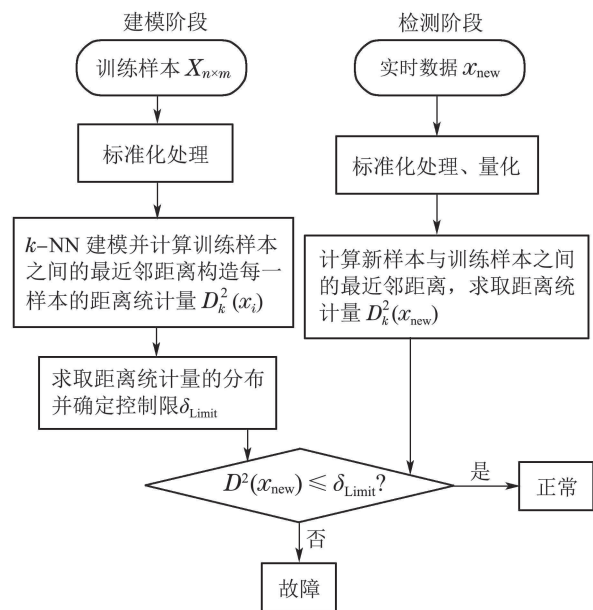


图1 k -NN建模与检测过程流程图

Fig. 1 Flow chart of k -NN modeling and fault detection

详细的建模与检测详细步骤如下:

1) 采集正常过程数据并标准化处理,将其转化为零均值,单位标准差的矩阵。

2) 根据式(2)计算样本间的距离统计量,并采用欧氏距离计算每一个训练样本与其他训练样本之间的 k 个最近距离平方和 $D_k^2(x_i)$,

$$D_k^2(x_i) = \sum_{j=1}^k \|x_i - N_j(x_i)\|^2, \quad (2)$$

其中: $k = 3$, $N_j(x_i)$ 代表 x_i 的第 j 个最近邻样本.

3) 通过非中心卡方分布确定统计量正常分布置信度为 99% 的控制限 δ_{Limit} [16, 27-28].

4) 对于新采集到的实时数据, 采用建模时的均值及标准差进行标准化处理.

5) 求取其与训练样本的前 k 个最近邻的距离并计算统计指标 $D_k^2(x_{new})$ (k 个最近邻距离平方和).

6) $D_k^2(x_{new})$ 与正常统计量控制限 δ_{Limit} 比较, 如果小于控制限, 则表示该样本正常; 反之, 过程异常.

3 加权的 k -NN 重构分析方法 (Weighted k -NN reconstruction method)

当统计指标大于控制限时, 系统有故障存在, 为了确保设备能够恢复到正常状态, 查找引起故障的根源是必要的. 因此, 对故障进行定位, 辨识发生故障的变量或传感器是故障诊断的目的, 也是本文的研究重点.

首先对几个符号和权值进行如下定义:

1) x_{test} 表示检测到的故障样本. x'_{test} 表示样本 x_{test} 去掉需要重构的变量后剩下变量组成的样本, 参见图 2.

2) $n_j(x'_{test})$ 表示 x'_{test} 在 $x'_{n \times (m-1)}$ 中的第 j 个临近样本的标签, 其中 $j = 1, 2, \dots, k$.

3) w 为权值, 表示测试样本与建模样本的相似程度, 权值越大相似度越大; 反之则越小.

故障变量识别步骤(结合图 2):

1) 重构第 1 个变量, 根据式(2)找到 $x'_{n \times (m-1)}$ 中 x'_{test} 的 k 个最近邻样本, 标签为 $n_j(x'_{test})$ ($j = 1, 2, \dots, k$), 并标记距离 d_1, d_2, \dots, d_k .

2) 根据式(3)计算权值 w , 其中 $\sum_{j=1}^k w_j = 1$, 进一步根据式(4)重构第 1 个变量:

$$w_j = \frac{1}{d_j} / \sum_{l=1}^k \frac{1}{d_l}, \quad (3)$$

$$v'_1 = \sum_{i=1}^k w_i (x_{n_i(x'_{test})})_1, \quad (4)$$

其中: $(x_{n_i(x'_{test})})_1$ 表示 x'_{test} 在 $x'_{n \times (m-1)}$ 中第 i 个近邻标签下的 x 的第 1 个变量值, v'_1 表示重构后的 x_{test} 中第 1 个变量值.

3) 计算重构第 1 个变量后的统计指标 d'_1 .

4) 根据上述方法, 依次重构 x_{test} 中的所有变量, 并计算重构后的统计指标 d'_1, d'_2, \dots, d'_m .

5) 将重构每个变量后的统计指标按从小到大排序, 假设排序后如图 3 所示, 寻找指标减小最大(maximize reduce index, MRI)的变量, 指标越小, 说明相比故障时的指标减小的越多.

6) 在图 3 中首先将指标减小最大的变量(x_{test} 中的变量 1) 替换为重构后的 v'_1 , 并计算控制指标 D_1 , 判断

是否小与控制限, 若大于控制限, 将重构后的 v'_2 替换 x_{test} 中的变量 2, 计算控制指标 D_{12} , 从小到大顺序依次迭代, 直到控制指标小于控制限; 假设

$$x_{t1} = [v'_1 \ v_2 \ \dots \ v_m], \quad x_{t12} = [v'_1 \ v'_2 \ \dots \ v_m],$$

D_1 和 D_{12} 的计算方法如式(5)-(6):

$$D_1 = \sum_{j=1}^k \|x_{t1} - N_j(x_{t1})\|^2, \quad (5)$$

$$D_{12} = \sum_{j=1}^k \|x_{t12} - N_j(x_{t12})\|^2. \quad (6)$$

7) 假设系统中只有两个变量(变量 1 和 2) 发生异常, 得到图 3 结论, 故障变量识别结束.

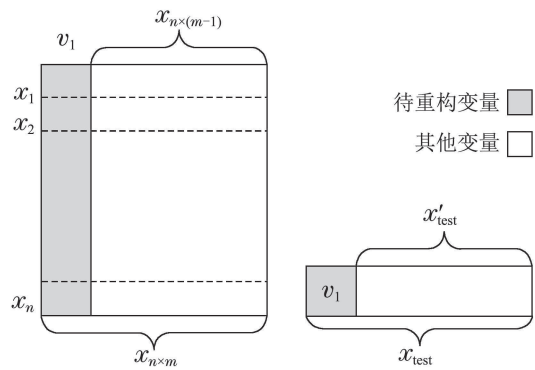


图 2 变量重构细节图

Fig. 2 The detail of variable reconstruction

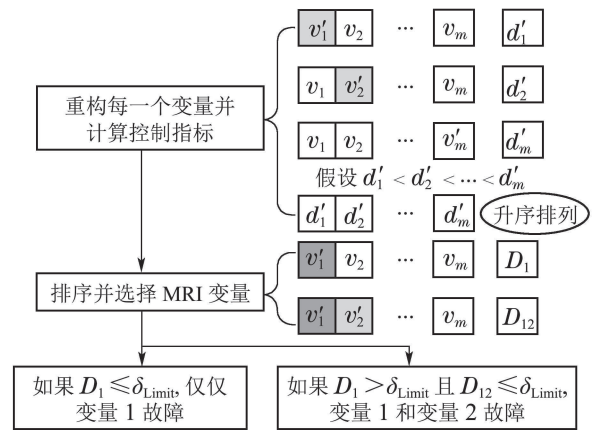


图 3 故障变量识别过程流程图

Fig. 3 Flow chart of fault variable identification

4 变量重构与故障诊断实例研究 (Case study of variable reconstruction and fault diagnosis)

4.1 数值仿真 (Numerical simulation)

数值仿真实例首先对变量重构的精度进行了验证, 然后从故障变量识别的角度进一步做出了分析与讨论: 最终验证了该方法在故障变量识别方面的有效性. 数值实例由两个潜隐变量 s_a, s_b 构造的 7 个变量组成, 可以根据式(7)得到:

$$\begin{cases} x_1 = 0.3217s_a + 0.4821s_b + e_1, \\ x_2 = 0.2468s_a + 0.1766s_b + e_2, \\ x_3 = 0.8291s_a + 0.4009s_b^2 + e_3, \\ x_4 = 0.7382s_a^2 + 0.0566s_b + e_4, \\ x_5 = 0.3972s_a^2 + 0.8045s_b^2 + e_5, \\ x_6 = 0.6519s_a s_b + 0.2071s_b + e_6, \\ x_7 = 0.4817s_a + 0.4508s_a s_b + e_7, \end{cases} \quad (7)$$

其中: e_1-e_7 是均值为零,标准差为0.01的噪声, s_a 是从-10到-7之间的随机数, s_b 服从 $N(-15, 1)$.

根据上式生成500个训练样本,同时产生500个待验证样本(假设变量 x_1 从301到350时刻数据缺失).为了验证加权 k -NN重构方法的有效性,使用所提方法对缺失数据进行重构补充,图4显示了重构结果与实际值的比较,经计算,50个数值的平均误差为0.04%.

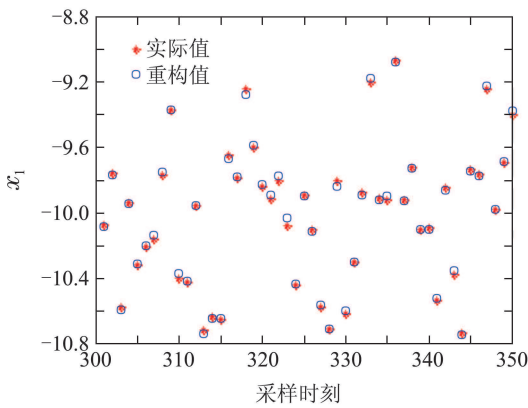


图4 重构后数据与原始数据的比较

Fig. 4 The comparison of reconstructed data and original data

另外,为了验证所提方法在故障变量识别方面的有效性,500个测试样本也被生成,其中故障如下:从101到150时刻对 x_1 加入幅值为8%的阶跃故障;从401到450时刻分别对 x_2 和 x_7 加入10%和5%的阶跃故障.图5给出了对测试样本的检测结果,可以明显地看到分别在101~150和401~450时刻系统都有故障发生.由于在设置故障时,101~150时刻仅仅变量1添加了故障,当变量1被重构后,如图6所示,控制指标已经小于控制限,因此在101~150时刻,识别出系统中只有一个故障变量为 x_1 .

同理,对于401~450时刻的故障,笔者首先给出如图7(a)-7(g)所示的重构每一个变量后的检测结果,可以看出,图7(a),7(c)-7(f)检测结果基本没有变化,和故障时的控制指标一样,而7(b)和7(g)存在较大的差异,通过比较有 $d'_2 < d'_7 < d'_1 < d'_6 < d'_3 < d'_5 < d'_4$.此时笔者得出故障可能为变量2和7,即 x_2 和 x_7 .当仅仅重构 x_2 和 x_7 后,得到的检测结果为图7(h),此时控制指标已经处于可控范围之内.因此,可以确定401~450时刻的故障变量为 x_2 和 x_7 .

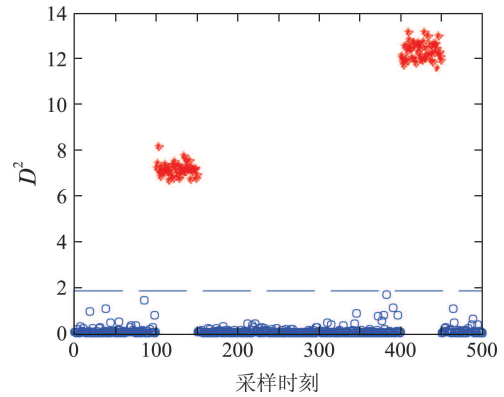


图5 测试数据的故障检测结果

Fig. 5 Fault detection results of test data

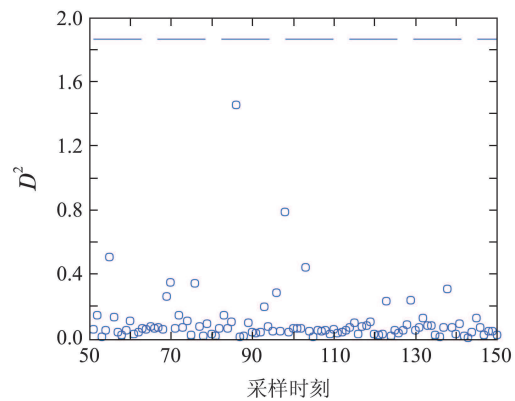
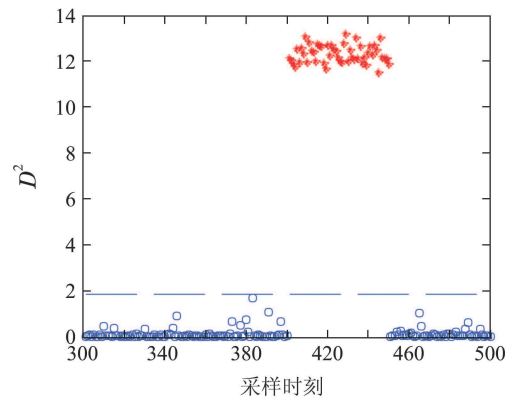
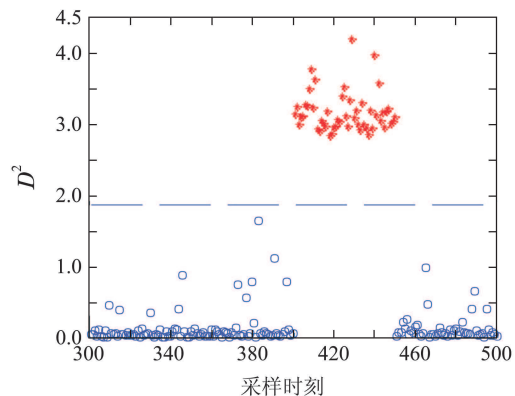


图6 重构变量1后的检测结果

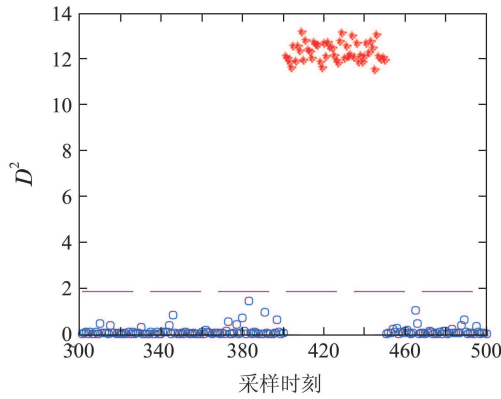
Fig. 6 The detection results after reconstructing variable 1



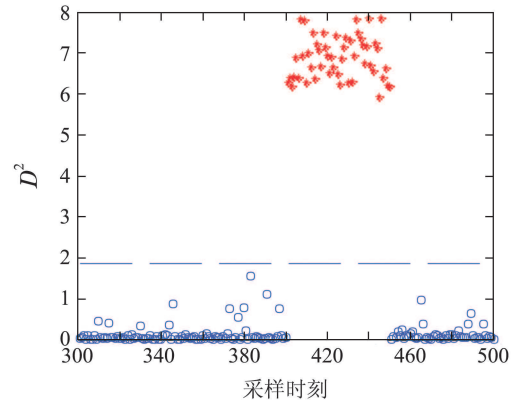
(a) 重构 x_1 后的结果



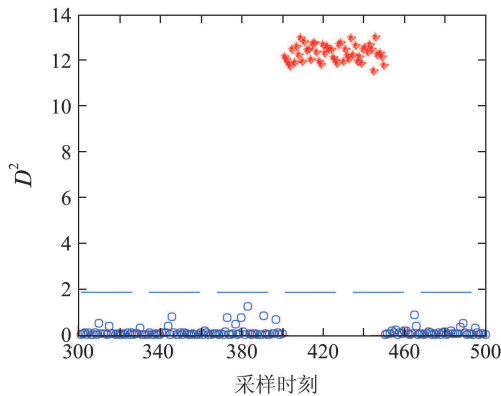
(b) 重构 x_2 后的结果



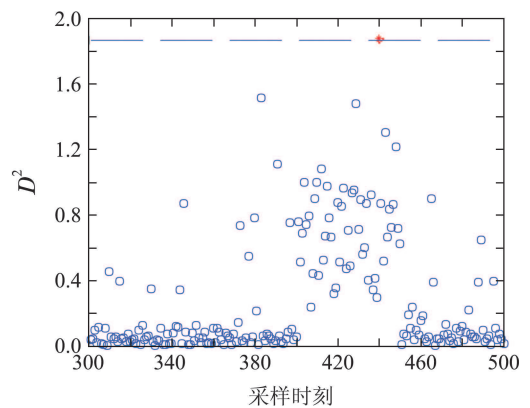
(c) 重构 x_3 后的结果



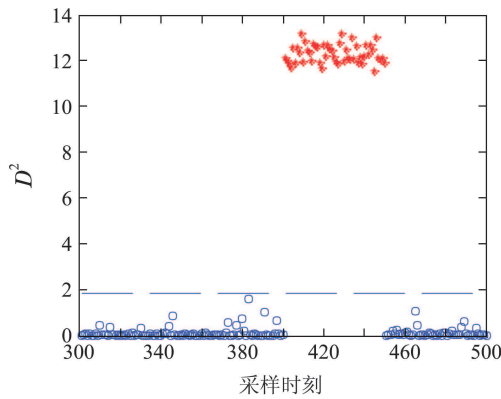
(g) 重构 x_7 后的结果



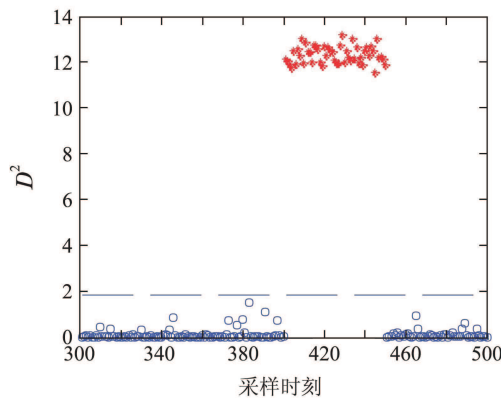
(d) 重构 x_4 后的结果



(h) 重构 x_2 和 x_7 后的结果



(e) 重构 x_5 后的结果



(f) 重构 x_6 后的结果

图 7 重构变量后的监视结果(从401–450)

Fig. 7 Monitoring results of reconstructed data from 401 to 450 samples

4.2 TE过程应用研究(Application research on TE process)

TE(Tennessee Eastman)过程是基于实际工业过程的仿真案例,由美国TE化学公司过程控制小组的J.J.Downs和E.F.Vogel于1993年提出^[29]. TE过程适合于研究过程控制技术、过程监控、故障检测与诊断等方向. 该过程由压缩机、分离器、反应器、冷凝器和汽提塔5个主要的操作单元组成,包括A-H8种成分,过程中有41个测量变量和12个控制变量,详细的工艺流程图如图8所示.

本文选取其中的22个测量变量(见表1)作为监测变量,数据通过仿真软件获得^[30]. 该过程共有21种故障可以引入,本文仅以故障1(故障类型为阶跃,扰动为A/C进料流量比发生变化,组分B含量保持不变)为例来进行仿真试验. 正常工况的960个样本来自过程平稳运行状态,故障工况前160个样本为无故障样本,后800个样本为故障样本.

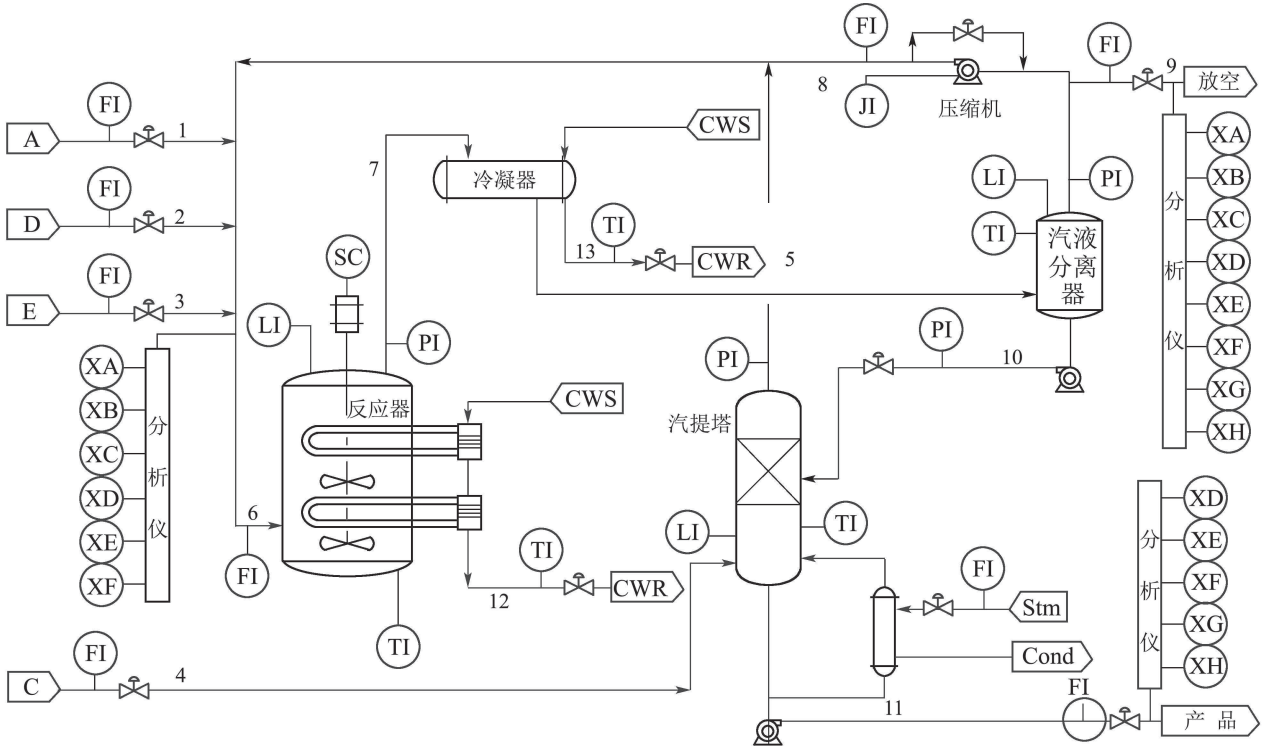


图8 TE化工过程流程图

Fig. 8 Flow diagram of the Tennessee Eastman chemical process

表1 TE过程监测变量

Table 1 Monitoring variables in the TE process

No.	变量名称	No.	变量名称
1	物料A流量	12	分离器液位
2	物料D流量	13	分离器压力
3	物料E流量	14	分离器底部流量
4	物料A和C流量	15	解析塔液位
5	循环流量	16	解析塔压力
6	反应器进料流量	17	解析塔底部流量
7	反应器压力	18	解析塔温度
8	反应器液位	19	解析塔蒸汽流量
9	反应器温度	20	压缩机功率
10	放空速率	21	反应器冷却水出口温度
11	分离器温度	22	分离器冷却水出口温度

首先, k -NN故障检测方法被应用于故障1, 从图9给出的监视结果可以得到故障引入时间为161时刻, 与设定故障时刻相同. 另外, 根据控制指标减小的大小可以得到, 变量1和变量4发生故障较早, 并且故障幅值较大, 这两个变量可能为故障根源(与引入故障相符). 在故障引入初期的180到300时刻, 由于变量之间的相关性, 控制指标出现了很大的波动, 与控制限差异很大; 后期由于系统的自调节功能, 许多可以自行恢复的过程故障变量已经回归正常, 导致控制指标有了一定程度的回落; 而在过程末期仍然没有恢复正常的变量可以认为是过程故障的根源.

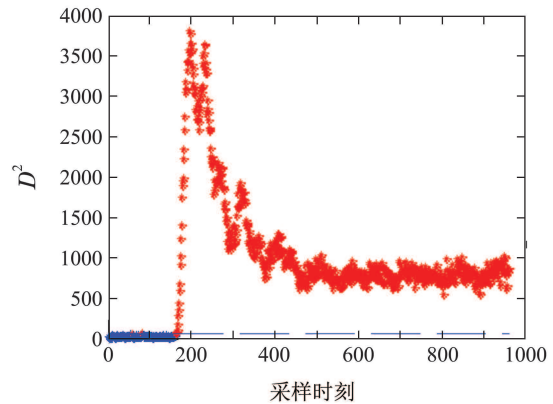


图9 故障1的故障检测结果

Fig. 9 Fault detection results of fault 1

为了识别产生故障的根源并找到发生故障的变量, 利用本文提出的加权 k -NN重构方法对每一个变量进行重构, 随后按照控制指标减小从大到小(MRI)的顺序排列变量, 如图10所示(颜色越深, 指标的减小幅度越大), 从图10中可以看到, 变量1(物料A流量)从故障引入时刻开始就已经发生异常(与引入故障相符), 直到过程结束. 重构变量1后, 图11显示出控制指标有明显下降, 但仍然大于控制限. 依次重构其他减小指标较大的变量, 重构变量1和4, 重构变量1, 4, 18和19后的检测结果分别如图12-13所示, 控制指标有明显的减小. 实验以951到960时刻为例, 在这段时刻重构后系统已经完全处于正常

状态. 因此, 此时段故障变量为变量 1, 4, 18 和 19, 所对应的变量分别为物料 A 流量, 物料 A 和 C 的流量, 解析塔温度和解析塔蒸汽流量, 与工艺流程和引入故障分析一致. 当过程中所有的故障变量都被重构之后, 故障检测结果如图 14 所示, 每个时刻的控制指标均在控制范围之内, 系统中所有的故障均被找到.

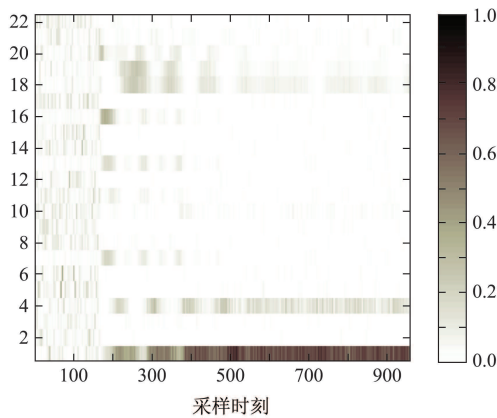


图 10 控制指标减小排序图

Fig. 10 The ordering of control index

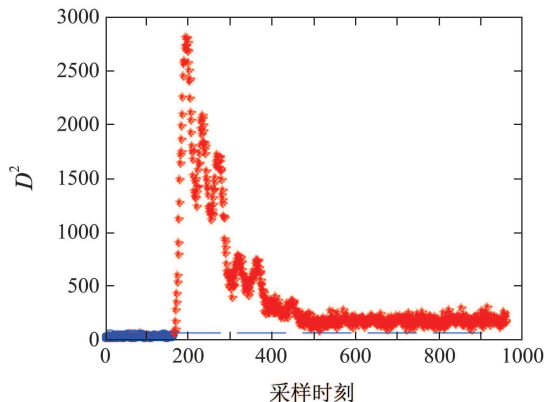


图 11 重构变量 1 后的检测结果

Fig. 11 The detection results after reconstructing variable 1

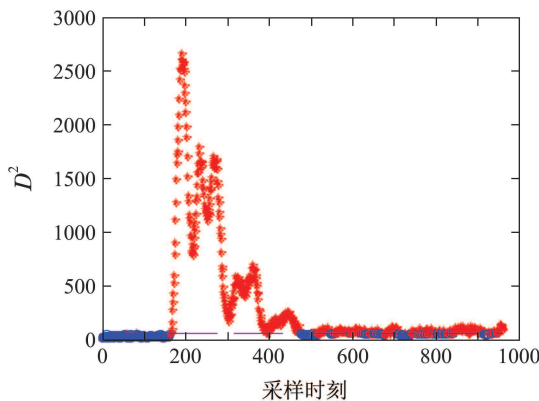


图 12 重构变量 1 和 4 后的检测结果

Fig. 12 The detection results after reconstructing variables 1 and 4

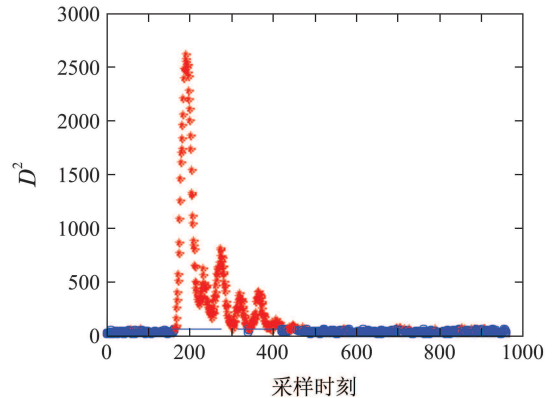


图 13 重构变量 1, 4, 18 和 19 后的检测结果

Fig. 13 The detection results after reconstructing variables 1, 4, 18 and 19

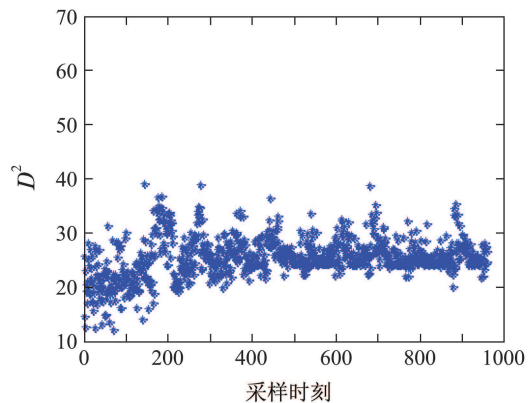


图 14 重构所有变量后的检测结果

Fig. 14 The detection results after reconstructing all variables

5 结论和进一步的工作(Conclusion and future works)

在 k -NN 故障检测的基础上, 本文首次将该算法应用于缺失数据补充和变量重构, 提出了一种新的基于加权 k -NN 重构分析的故障诊断方法, 该方法的目标是精确地识别过程中所有的异常变量. 文章分别使用数值仿真和 TE 实例验证了加权 k -NN 数据重构的精确性和多变量故障识别的有效性. 下一步计划在本文基础上进一步探讨参数 k 对重构精度的影响, 并用马氏距离代替欧式距离, 通过理论分析结合实验仿真说明两者的诊断效果以及本质区别.

参考文献(References):

- [1] 周东华, 李钢, 李元. 数据驱动的工业过程故障检测与诊断技术 [M]. 北京: 科学出版社, 2011. (ZHOU Donghua, LI Gang, LI Yuan. *Fault Detection and Diagnosis Industrial Process Based on Data-Driven* [M]. Beijing: Science Press, 2011.)
- [2] 周东华, 胡艳艳. 动态系统的故障诊断技术 [J]. 自动化学报, 2009, 35(6): 748 - 758. (ZHOU Donghua, HU Yanyan. Fault diagnosis techniques for dynamic systems [J]. *Acta Automatica Sinica*, 2009, 35(6): 748 - 758.)

- [3] 李娟, 周东华, 司小胜, 等. 微小故障诊断综述 [J]. 控制理论与应用, 2012, 29(12): 1515 – 1529.
(LI Juan, ZHOU Donghua, SI Xiaosheng, et al. Review of incipient fault diagnosis methods [J]. *Control Theory & Applications*, 2012, 29(12): 1515 – 1529.)
- [4] NOMIKOS P, MACGREGOR J F. Monitoring batch processes using multiway principal component analysis [J]. *American Institute of Chemical Engineers Journal*, 1994, 40(8): 1361 – 1375.
- [5] KRESTA J, MACGREGOR J F, MARLIN T E. Multivariate statistical monitoring of process operating performance [J]. *Canadian Journal of Chemical Engineering*, 1991, 69(1): 35 – 47.
- [6] WISE B M, RICKER N L. Recent advances in multivariate statistical process control: improving robustness and sensitivity [C] // *Proceedings of the IFAC International Symposium, ADCHEM' 91*. Toulouse, France: Pergamon Press, 1991: 125 – 130.
- [7] MACGREGOR J F, JAECKLE C, KIPARISSIDES C, et al. Process monitoring and diagnosis by multiblock PLS methods [J]. *AIChE Journal*, 1994, 40(5): 826 – 838.
- [8] NOMIKOS P, MACGREGOR J F. Monitoring batch process using multiway principal component analysis [J]. *AIChE Journal*, 1994, 40(8): 1361 – 1375.
- [9] 邱天, 白晓静, 郑茜予, 等. 多元指数加权移动平均主元分析的微小故障检测 [J]. 控制理论与应用, 2014, 31(1): 19 – 26.
(QIU Tian, BAI Xiaojing, ZHENG Xiyu, et al. Incipient fault detection of multivariate exponentially weighted moving average principal component analysis [J]. *Control Theory & Applications*, 2014, 31(1): 19 – 26.)
- [10] 张家良, 曹建福, 高峰, 等. 结合非线性频谱与核主元分析的复杂系统故障诊断方法 [J]. 控制理论与应用, 2012, 29(12): 1558 – 1564.
(ZHANG Jialiang, CAO Jianfu, GAO Feng, et al. Fault diagnosis of complex system based on nonlinear spectrum and kernel principal component analysis [J]. *Control Theory & Applications*, 2012, 29(12): 1558 – 1564.)
- [11] KRUGER U, DIMITRIADIS G. Diagnosis of process faults in chemical systems using a local partial least squares approach [J]. *AIChE Journal*, 2008, 54(10): 2581 – 2596.
- [12] QIN S J. Statistical process monitoring: basics and beyond [J]. *Journal of Chemometrics*, 2003, 17(8/9): 480 – 502.
- [13] ZHANG Y W, QIN S J. Improved nonlinear fault detection technique and statistical analysis [J]. *AIChE Journal*, 2008, 54(12): 3207 – 3220.
- [14] ZHANG Y W, ZHOU H, QIN S J, et al. Decentralized fault diagnosis of large-scale processes using multiblock kernel partial least squares [J]. *IEEE Transactions on Industrial Informatics*, 2010, 6(1): 3 – 10.
- [15] WANG G Z, LIU J C, LI Y. A progressive fault detection and diagnosis method based on dissimilarity of process data [C] // *Proceedings of the 2014 IEEE International Conference on Information and Automation (ICIA)*. Piscataway, NJ, USA: IEEE, 2014: 1210 – 1216.
- [16] HE Q P, WANG J. Fault detection using the k -nearest neighbor rule for semiconductor manufacturing processes [J]. *IEEE Transactions on Semiconductor Manufacturing*, 2007, 20(4): 345 – 354.
- [17] HE Q P, WANG J. Principal component based k -nearest neighbor rule for semiconductor process fault detection [C] // *Proceedings of the 2008 American Control Conference, ACC*. Piscataway, NJ, USA: IEEE, 2008: 1606 – 1611.
- [18] HE Q P, WANG J. Large-scale semiconductor process fault detection using a fast pattern recognition-based method [J]. *IEEE Transactions on Semiconductor Manufacturing*, 2010, 23(2): 194 – 200.
- [19] VERDIER G, FERREIRA A. Fault detection with an adaptive distance for the k -nearest neighbors rule [C] // *Proceedings of the 2009 International Conference on Computers and Industrial Engineering*. Piscataway, NJ, USA: IEEE, 2009: 1273 – 1278.
- [20] WANG G Z, LIU J C, ZHANG Y W, et al. A novel multi-mode data processing method and its application in industrial process monitoring [J]. *Journal of Chemometrics*, 2015, 29(2): 126 – 138.
- [21] WESTERHUIS J A, GURDEN S P, SMILDE A K. Generalized contribution plots in multivariate statistical process monitoring [J]. *Chemometrics and Intelligent Laboratory Systems*, 2000, 51(1): 95 – 114.
- [22] CONLIN A K, MARTIN E B, MORRIS A J. Confidence limits for contribution plots [J]. *Journal of Chemometrics*, 2000, 14(5/6): 725 – 736.
- [23] KOURTI T, MACGREGOR J F. Multivariate SPC methods for process and product monitoring [J]. *Journal of Quality Technology*, 1996, 28(4): 409 – 428.
- [24] QIN S J, YUE H H, DUNIA R. Self-validating inferential sensors with application to air emission monitoring [J]. *Industrial and Engineering Chemistry Research*, 1997, 36(5): 1675 – 1685.
- [25] YUE H, QIN S J. Reconstruction based fault identification using a combined index [J]. *Industrial and Engineering Chemistry Research*, 2001, 40(20): 4403 – 4414.
- [26] HE Q P, WANG J. Large-scale semiconductor process fault detection using a fast pattern recognition-based method [J]. *IEEE Transactions on Semiconductor Manufacturing*, 2010, 23(2): 194 – 200.
- [27] 郭小萍, 袁杰, 李元. 基于特征空间 k 最近邻的批次过程监视 [J]. 自动化学报, 2014, 40(1): 135 – 142.
(GUO Xiaoping, YUAN Jie, LI Yuan. Feature space k -nearest neighbor based batch process monitoring [J]. *Acta Automatica Sinica*, 2014, 40(1): 135 – 142.)
- [28] LI Y, ZHANG X M. Diffusion maps based k -nearest-neighbor rule technique for semiconductor manufacturing process fault detection [J]. *Chemometrics and Intelligent Laboratory Systems*, 2014, 136: 47 – 57.
- [29] DOWN J J, VOGEL E F. A plant-wide industrial process control problem [J]. *Computers and Chemical Engineering*, 1993, 17(3): 245 – 255.
- [30] RICKER N L. Decentralized control of the Tennessee Eastman challenge process [J]. *Journal of Process Control*, 1996, 6(4): 205 – 221.

作者简介:

王国柱 (1984–), 男, 博士研究生, 目前研究方向为复杂工业过程的故障检测与故障诊断研究, E-mail: wang.guo.zhu@163.com;

刘建昌 (1960–), 男, 教授, 博士生导师, 目前研究方向为智能控制理论与应用、复杂系统建模、控制与优化、故障诊断等, E-mail: liujianchang@ise.neu.edu.cn;

李元 (1964–), 女, 教授, 博士生导师, 目前研究方向为过程控制与故障诊断, E-mail: li-yuan@mail.tsinghua.edu.cn;

商亮亮 (1981–), 男, 博士研究生, 目前研究方向为工业过程的子空间辨识建模方法, E-mail: sl-15@163.com.