

面向生产管控的多源数据处理技术^①



龚 勋, 王淑莹, 崔晓宇

(西南交通大学 信息科学与技术学院, 成都 611756)

通讯作者: 龚 勋, E-mail: 1727474808@qq.com

摘 要: 为了解决流程型生产车间数据种类多、数据质量参差不齐影响生产管控的问题, 提出了面向多源数据的数据分类处理技术, 采用数据分类与滑动窗口相结合的方式进行数据处理. 首先, 根据生产数据特点建立数据模型, 然后进行数据分类. 主要分为状态数据、开关数据、逻辑数据 3 类, 不同类型的数据使用不同处理算法; 同时采用滑动窗口解决不同任务对数据实时性和完整性要求不同的难点. 最后, 该数据分类处理模型在实际生产环境投入使用, 验证了对生产数据处理的准确性和实时性. 结果表明, 将处理后的数据应用于生产管控, 管控误差率降低至不足 1%.

关键词: 流程型; 生产管控; 多源数据; 数据分类; 滑动窗口

引用格式: 龚勋, 王淑莹, 崔晓宇. 面向生产管控的多源数据处理技术. 计算机系统应用, 2020, 29(11): 237-242. <http://www.c-s-a.org.cn/1003-3254/7573.html>

Multi-Source Data Processing Technology for Production Control

GONG Xun, WANG Shu-Ying, CUI Xiao-Yu

(School of Information Science and Technology, Southwest Jiaotong University, Chengdu 611756, China)

Abstract: In order to solve the problem that there are multi-source and uneven quality data in process-oriented workshop, which affect production control, the data classification and processing technology for multi-source data is proposed. Data processing is performed using a combination of data classification and sliding windows. First, a data model is established based on the characteristics of the production data, and the data is classified. It is mainly divided into three types: state data, switch data and logical data. Different types of data use different processing algorithms. As the same time, sliding windows are used to solve the difficult that different tasks have different requirements for data real-time and integrity. Finally, the data classification processing model is put into use in the actual production environment, which verifies the accuracy and real-time performance of the production data processing. The results show that by applying the processed data to production control, the control error rate is reduced to less than 1%.

Key words: process-oriented; production control; multi-source data; data classification; sliding windows

随着智能时代的到来, 生产数据价值的重要性越来越明显. 如何根据生产工艺, 通过数据采集、处理、聚类、存储、分析和可视化等技术手段发挥数据蕴含的价值信息, 成为了热门的研究课题. 而流程型制造车间的生产环境特殊, 加上传输网络、人为因素的影响,

导致采集到的生产数据不可避免的存在一定程度的异常情况, 如数据缺失、离群和异常波动等^[1,2], 若不经过处理, 无法正常应用. 数据处理就是为了清洗掉生产现场各种因素干扰造成的异常情况, 同时也为了面向生产管控需求进行特定的数据加工^[3,4].

① 基金项目: 国家科技支撑计划 (2017YFB1401401); 山东省重大科技创新工程 (2017CXGC0608-02)

Foundation item: National Sci-Tech Supporting Plan (2017YFB1401401); Major Science and Technology Innovation Engineering of Shandong Province (2017CXGC0608-02)

收稿时间: 2020-02-02; 修改时间: 2020-03-03; 采用时间: 2020-03-13; csa 在线出版时间: 2020-10-29

目前, 数据处理技术非常广泛^[5,6], 如文献 [7] 针对复杂装备的监测数据在采集过程中出现时序混乱, 影响对设备状态的分析问题进行了建模研究, 根据不同实时任务对数据的时序容错性和实时性要求不同, 采用不同大小的缓冲窗口和处理策略, 并结合数据传输延时的概率密度, 分析计算模型的参数, 最后以仿真实验的方式验证了所提出模型的有效性; 文献 [8] 针对流程型制造业个性化的环境中生产工艺参数难以快速决策影响实时管控任务的问题, 通过结合实际生产工艺进行分析, 提出了基于数据驱动的生产工艺参数匹配方法, 建立了产品工艺相似度函数, 从大量生产数据中选取与实际生产任务有关的邻近数据, 通过对比历史数据与目标产品的生产要求实现了生产工艺参数的匹配; 文献 [9] 通过分析已有时间序列数据处理的研究, 进行了大量的实验研究, 结果表明, 在特殊情况下某些观点可能太乐观, 最后重新实现了一些时间序列表示形式和相似性度量; 文献 [10] 和文献 [11] 针对缺失数据的异常情况, 基于 EM 插补算法和高斯混合模型进行改进优化, 提出一种新的数据流缺失处理算法, 有效的解决了数据缺失异常, 最后通过实验验证, 相比于传统的 EM 插补算法, 改进后的算法实时性和适应性更好, 但是该算法仅局限于对缺失数据进行修复, 其他异常情况无法处理; 文献 [12] 根据传感器数据的时序性特点, 提出了一种基于嵌套滑动窗口的数据缺失填充算法, 通过皮尔逊分析数据项的空间相关性, 然后使用固定大小的嵌套滑动窗口进行数据采样处理, 通过 MKNN 算法进行缺失数据的精准填补, 该算法的复杂度更低, 数据修复效果精准, 但是也只局限于缺失数据异常问题。

通过对现有数据处理技术的分析, 并不适合应用于流程型制造车间生产数据的处理, 因为其数据种类繁多, 格式不一, 不同数据项采集频率不一, 数据质量也是参差不齐, 加上不同任务对数据的实时性和准确性要求不同, 所以不能一概而论. 本文根据生产数据特点不同, 将其分为 3 类: 状态数据、开关数据和逻辑数据. 其中状态数据是生产过程中表示环境、设备运行状态的数据信号, 如物料重量、传送带速度、焙烧温度时间、环境温度湿度等; 开关数据是设备启动、停止, 以及阀门开关时产生的开关信号; 逻辑数据是设备运行中表示是否运行的逻辑信号, 如电流、频率, 其值在某一范围内有效. 为了满足流程型车间生产任务的实际需

求, 针对不同数据类型, 不同异常类型采用了不同的处理策略.

1 相关定义

1.1 数据模型

待处理数据是典型的时间序列^[9]数据, 即一系列按时间先后顺序、等间隔采集的数据点所组成的数据序列. 因此, 数据模型可表示为由采集点唯一标识、采集时间以及数据值组成, 记为:

$$D = \{d_1 = (p_1, t_1, v_1), \dots, d_i = (p_i, t_i, v_i), \dots, d_n = (p_n, t_n, v_n)\}$$

其中, $d_i = (p_i, t_i, v_i)$ 表示在 p_i 采集点、 t_i 时刻采集的数据值为 v_i , 并且时间是一维线性递增的, n 为时间序列的长度. 制造车间所有采集点在环境的影响下采集到的生产数据, 如图 1 所示

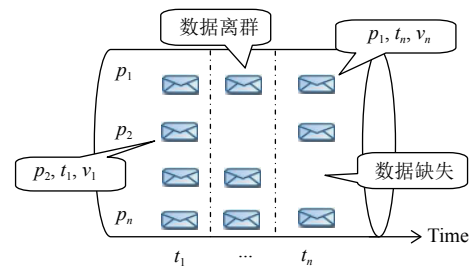


图 1 生产数据示意图

流程型制造车间生产过程中时间序列数据的特点就是数据维度高, 即数据值 v 表示的类型众多, 如温度值、速度值、重量值等; 加上时间流特点, 这决定了不能使用传统单一的数据处理方式进行处理. 因此, 降低数据维度、提高数据处理算法效率成为关键. 本文首先根据数据特征, 进行数据分类处理, 再针对每类数据采用适合的算法进行处理, 在保证数据处理质量的前提下, 最大化提高处理效率^[13].

1.2 滑动窗口

滑动窗口^[12]是一个固定长度的队列, 窗口中的数据是过去一段时间内采集数据点组成的数据流. 数据点从窗口首部流入, 尾部流出.

根据窗口中存放数据方式的定义, 滑动窗口分为基于时间的滑动窗口和基于数据个数的滑动窗口; 其中基于时间的滑动窗口保存最近一段时间长度内采集到的数据; 基于数据个数的滑动窗口保存最近采集到的 K 个数据.

结合采集数据的特点分析, 数据可能存在丢失情

况,如果采用基于时间的定义方式,则在一定时间段内进入滑动窗口的数据项个数可能很少,在数据修复时准确度会下降.所以本文选择基于数据个数的滑动窗口.在数据处理过程中,结合生产工艺的具体情况,根据每类数据的特点不同,为每类数据建立不同宽度的滑动窗口缓存数据,尽可能实时高效的处理数据.

例如滑动窗口大小 N ,在 t_i 时刻,进入滑动窗口的数据流 (d_1, d_2, \dots, d_n) ,如图 2 所示^[12].

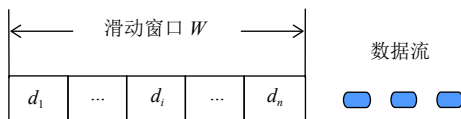


图 2 t_i 时刻滑动窗口

在 t_j 时刻,滑动窗口中前 $i-j$ 个数据移出窗口, d_{n+1} 至 d_{n+i-j} 作为最新数据进入滑动窗口,存放在最右位置,如图 3 所示.

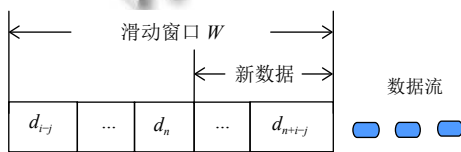


图 3 t_j 时刻滑动窗口

随着时间推移,数据流不断流入滑动窗口,窗口中的数据不断更新,并且每次只需要处理更新后的数据,即 d_{n+1} 至 d_{n+i-j} .

2 数据流处理算法

2.1 处理模型

本文针对流程型制造车间采集数据来源广、采集频率不同、格式不同等特点分析,提出基于滑动窗口的多源数据分类处理模型.如图 4 所示^[14].

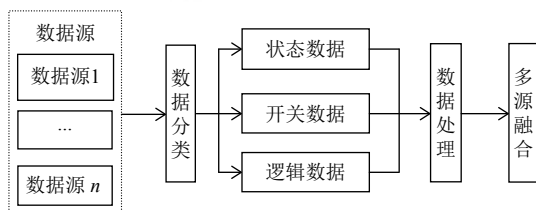


图 4 基于滑动窗口的多源数据分类处理模型

首先根据生产管控任务从数据源中选择需要的数据项,在进行处理前根据数据特点进行分类,可划分为

3 类数据,分别是状态数据、开关数据、逻辑数据;基于各类数据对完整性、准确性的需求不同,采用不同的处理算法进行修复;最后根据生产管控任务中数据项的关联关系将处理后的数据进行多源融合,方便应用.

2.2 数据分类依据

根据数据的采集频率将数据分为 3 类:逻辑数据、状态数据、开关数据.其中逻辑数据变化最频繁,采样频率最高,状态数据次之,开关数据最后;状态数据包含数据项、异常情况最多,分析处理也最为复杂.以下是对不同种类数据进行异常分析处理的过程.

2.3 状态数据

状态数据的异常情况可以分为 3 类:数据重复、数据离群和数据丢失;其中丢失数据的修复较为复杂.根据状态数据的特点,主要采用 EM 算法进行改进,结合滑动窗口技术进行处理,提高数据处理的精度和性能.

(1) 重复

根据之前对数据模型的分析,每个数据点都是由采集点唯一标识、采集时间和采集值构成,记 $d_1 = (p_1, t_1, v_1)$.所以重复数据可以根据采集时间的唯一性过滤.

(2) 离群

离群数据包括局部异常高值、低值,超出数据阈值.其中超出数据阈值异常可以通过为每个数据项配置数据阈值方式过滤;局部异常高值低值可以通过中值法修复,如下式:

$$d(p_i, t_i, v_i) = \frac{d(p_i, t_{i+1}, v_{i+1}) + d(p_i, t_{i-1}, v_{i-1})}{2} \quad (1)$$

(3) 缺失

数据缺失的异常处理较为复杂,分为连续型缺失和不连续型缺失.针对不连续缺失情况,由于缺失数据前后时间间隔较短,可以采用中值法进行修补;对于连续型缺失,由于缺失数据过多若使用中值法修复,则可能导致误差较大,所以本文采用改进的 EM (Expectation-Maximization algorithm, 最大期望算法) 插值法,处理连续性缺失^[10,11].EM 算法是一种从不完整数据集中对参数进行估值,然后通过不断迭代求参数最大似然值的经典数据修复算法之一.

文献 [10] 介绍了针对 EM 算法进行改进处理缺失数据的情况.改进的 EM 缺失插补算法有 3 个特点:(1) 设定连续缺失数据点个数的阈值为 M ,当连续缺失

数据个数小于 M 时,表明数据缺失不严重,使用中值法进行修复,大于 M 时才使用 EM 插补算法;(2) 设定 EM 算法的最高迭代次数为 K ,当算法迭代次数大于 K 时,停止迭代,使用最后一次迭代结果作为缺失数据的插补值;(3) 设定算法迭代精度 ε ,当第 k 次与 $k-1$ 次迭代结果差小于 ε 时,表明结果满足要求,停止迭代,这样可以在保证数据处理质量的前提下,提高数据处理效率.

数据修复的整个过程是基于固定大小的滑动窗口模型,窗口宽度为 N ,数据缺失阈值为 M ,最高迭代次数为 K ,迭代精度为 ε ,数据的采样频率为 T ,即时间序列数据正常到达滑动窗口的时间间隔.计算滑动窗口相邻两条数据 $d_i=(p_i,t_i,v_i)$ 和 $d_{i+1}=(p_i,t_{i+1},v_{i+1})$ 的时间间隔 $\Delta t=t_{i+1}-t_i$,比较 Δt 与 $M \times T$ 的关系,共有 3 种情况,如下:

(1) 当 $\Delta t < 2 \times T$,则采集数据无缺失情况,无须处理;

(2) 当 $2 \times T \leq \Delta t \leq M \times T$,则数据点 d_1 和 d_2 之间存在缺失点,但没有超过设定的缺失阈值 M ,使用中值法进行修复;

(3) 当 $\Delta t \geq M \times T$,则缺失数据个数超出阈值 M ,采用 EM 算法进行修复.具体如下^[12]:首先记滑动窗口中 N 个数据点构成的数据集为 S ,假设其中有 m 个连续缺失的数据点,记为 $d_i(i=1,2,\dots,m)$.扫描数据集,分别以 m 个连续缺失数据为中心进行高斯模型划分,得到 m 个高斯模型分别记为 S_1, S_2, \dots, S_m ,构成高斯混合模型^[11],表示为:

$$p(S_i) = \sum_{i=1}^m \pi_i N(S_i | \mu_i, \Sigma_i) \quad (2)$$

其中, $N(S_i | \mu_i, \Sigma_i)$ 是混合模型的第 i 个, π_i 是每个模型的权重,满足 $0 \leq \pi_i \leq 1$;在这里引入一个 m 维隐变量 z , z_i ($1 \leq i \leq m$) 只能取 0 或 1 两个值, $z_i=1$ 表示选中第 i 个分组中,即 $p(z_i=1)=\pi_i, z_i \in \{0, 1\}$.在高斯混合模型中,隐变量是缺失点属于某个模型的概率.因为 S_1, S_2, \dots, S_m 分别都是符合高斯分布的,则用条件概率表示为:

$$p(S_i | Z_k = 1) = N(S_i | \mu_i, \Sigma_i) \quad (3)$$

在贝叶斯思想下, $p(z)$ 是先验概率, $p(s|z)$ 是似然概率,得到后验概率 $p(z|s)$ 如下:

$$\gamma(Z_k) = p(Z_k = 1 | S) = \frac{\pi_i N(S | \mu_i, \Sigma_i)}{\sum_{j=1}^m \pi_j N(S | \mu_j, \Sigma_j)} \quad (4)$$

为了方便计算模型的参数,在式 (4) 中用 $\gamma(z_k)$ 表示第 k 个分组的后验概率.这里有 3 个参数 π, μ, Σ 每次计算需要更新,需要分别求出这 3 个参数的最大似然函数,求似然函数的方法有多种,这里直接给出最大似然函数:

$$\mu_i = \frac{1}{C_i} \sum_{n=1}^C \gamma(Z_{ni}) x_n \quad (5)$$

$$\Sigma_i = \frac{1}{C_i} \sum_{n=1}^C \gamma(Z_{ni}) (x_n - \mu_i)(x_n - \mu_i)^T \quad (6)$$

$$\pi_i = \frac{C_i}{C} \quad (7)$$

其中, C 表示每个分组中包含数据点的个数, x_n 表示当前分组中的第 n 个点, $\gamma(z_{ni})$ 表示第 n 点属于分组 S_i 的后验概率.

然后设定最大迭代次数 K ,初始迭代次数 $K_0=1$,迭代精度 ε ,数据缺失阈值 M ,以及每个分组 π_i, μ_i, Σ_i 的初始值进行结算.

求期望:根据当前 π_i, μ_i, Σ_i 计算后验概率 $\gamma(z_n i)$,由上述 $\gamma(z_n i)$ 表达式如下:

$$\gamma(z_n i) = \frac{\pi_i N(x_n | \mu_i, \Sigma_i)}{\sum_{j=1}^m \pi_j N(x_n | \mu_j, \Sigma_j)} \quad (8)$$

求极大:根据求期望步骤可得的 $\gamma(z_n i)$,再代入式 (5)~式 (7) 计算 π_i, μ_i, Σ_i 的值,在每次迭代之前都需要判断迭代次数 $K_0 > K$ 和迭代精度是否小于 ε ,若不成立则迭代执行求期望步骤;若成立则停止迭代,输出最终参数 μ_i 作为缺失点的插补值.

2.4 开关数据

开关数据主要有两种状态开和关,有两种表现形,第一种分别用 True 和 False 表示开和关,另一种用 1 和 0 表示用开和关,两种表示形式的数据处理都不复杂.首先需要将非数字量转换为数字量的 0 或 1;然后结合生产工艺,联合相关数据项进行处理,这里不做详细介绍.对于同一采集点的开关数据处理结果一定是 $\{0,1,0,1,\dots\}$.

2.5 逻辑数据

逻辑数据是另一种形式表示设备开关状态的数据,其采样频率高,但是只有在一定范围内才有效,如机器运转的频率、电流、电压等,只有频率达到某个范围才表示机器在运转中,而在这个范围内,数据波动频繁,

有时出现异常情况可能跳出范围内. 所以对该类数据的容错性较高, 即使出现部分数据缺失、离群对整体数据的影响不大, 处理的重点在于找到数据趋于稳定的点, 将稳定在某范围内的数据按逻辑 1 处理, 其余按逻辑 0 处理. 这样便于对逻辑数据的分析时找到从 0 跳变到 1 或者从 1 跳变到 0 的时间点.

本文采用二次稳定度算法判断逻辑数据的稳定性, 具体如算法 1^[15].

算法 1. 二次稳定度算法

- 1) 结合生产工艺和数据特点, 选择合适的波动率 (R_1, R_2), 其中 R_1 、 R_2 分别是第一次和第二次稳定性判断的波动率, 以及稳定指标 (S_1, S_2); 一般情况下波动率 R_1 大于 R_2 , 稳定指标 S_1 小于 S_2 ;
- 2) 从计算滑动窗口中所有数据的总值 V_{sum} 和平均值 V_{avg} ; 当每次滑动窗口向后移动时只需要用 V_{sum} 减去移出数据值, 再加上移入数据值即可得到新的总值;
- 3) 根据波动率 R_1 确定均值上下的一次稳定区间 [$V_{avg}-V_{avg}\times R_1$, $V_{avg}+V_{avg}\times R_1$], 然后计算滑动窗口中数据在一次稳定区间中的概率, 即数值在稳定区间的个数占总数的比值 P_1 ; 如果 P_1 小于 S_1 , 则判定数据不稳定, 将滑动窗口首位置为 0, 否则执行 4);
- 4) 根据波动率 R_2 计算均值上下二次稳定区间 [$V_{avg}-V_{avg}\times R_2$, $V_{avg}+V_{avg}\times R_2$], 然后计算滑动窗口中数据在二次稳定区间中的概率 P_2 ; 如果 P_2 大于等于 S_2 , 则判定数据稳定, 将滑动窗口首位置为 1, 否则置为 0.

3 实验结果

3.1 实验环境说明

本文采用的实验环境是 Windows 7, CPU 3.6 GHz, 内存 16 GB 的计算机; 实验数据是某锂电池材料生产车间产生的数据, 该生产车间是典型的流程型生产车间, 数据种类多, 质量参差不齐, 采样间隔 1 s, 如表 1 所示.

表 1 部分生产数据

时间(s)	Bunker weigh(kg)	传送速度(m/s)	温度(°C)
1	27.0	0.2	180
2	28.1	0.3	182
3	29.3	0.4	185
4	null	null	189
5	31.5	0.6	null
6	null	0.6	192
7	null	0.6	null
16	null	null	193
17	null	null	null
18	35.8	null	196
19	36.9	null	198
...

3.2 实验结果

本文分别选取滑动窗口的宽度 N 为 150, 160, 170,

180, 190, 200; EM 算法循环迭代精度 $\varepsilon=0.1$; 数据缺失阈值 $W=4,5,6,7,8$; 分别进行实验验证, 最后选定滑动窗口宽度 $N=180$, 数据缺失阈值 $W=7$ 时数据插补效果最好.

同时对于逻辑数据由于生产环境因素影响造成其值在一定范围内波动的情况, 如图 5 所示, 为原料高混机运行时产生的电流数据.

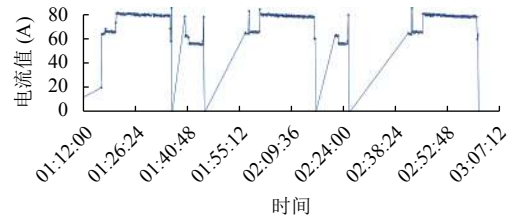


图 5 高混机电流数据

为了便于找到数据稳定趋势, 分析从逻辑 0 跳变到逻辑 1, 或者相反跳变的时间点, 采用二次稳定度算法, 结合生产工艺设置波动率 R_1 、 R_2 分别是 10%、5%, 稳定指标 S_1 、 S_2 分别是 90%、80%, 将稳定在某范围内的数据值按逻辑 1 处理, 其余按照逻辑 0 处理, 结果如图 6 所示.

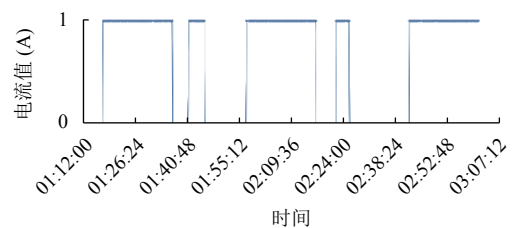


图 6 处理后的电流数据

最后为了验证数据处理效果, 将经过处理后的数据应用于流程型制造车间锂电池材料产品生产批次判断, 与实际人工统计结果比较, 误差率不足 1%.

4 结论与展望

本文针对流程型车间生产数据的多样性、数据质量参差不齐等问题, 采用数据分类处理技术, 相比与传统的单一数据处理技术, 提高了数据处理效率, 为数据应用于过程控制提供了保障. 最后以某锂电池材料生产车间生产数据为例, 验证了所设计的数据分类处理模型. 结果表明: 该方法在保证数据处理质量的前提下, 较传统数据处理方法更加简洁、高效, 具有实际应用价值.

参考文献

- 1 Ponce LM, dos Santos W, Meira Jr W, *et al.* Upgrading a high performance computing environment for massive data processing. *Journal of Internet Services and Applications*, 2019, 10(1): 19. [doi: [10.1186/s13174-019-0118-7](https://doi.org/10.1186/s13174-019-0118-7)]
- 2 Van Nguyen T, Zhou L, Chong AYL, *et al.* Predicting customer demand for remanufactured products: A data-mining approach. *European Journal of Operational Research*, 2019, 281(3): 543–558.
- 3 Niu G, Yang BS, Pecht M. Development of an optimized condition-based maintenance system by data fusion and reliability-centered maintenance. *Reliability Engineering & System Safety*, 2010, 95(7): 786–796.
- 4 赵颖, 侯俊杰, 于成龙, 等. 面向生产管控的工业大数据研究及应用. *计算机科学*, 2019, 46(S1): 45–51.
- 5 Nardelli M, Cardellini V, Grassi V, *et al.* Efficient operator placement for distributed data stream processing applications. *IEEE Transactions on Parallel and Distributed Systems*, 2019, 30(8): 1753–1767. [doi: [10.1109/TPDS.2019.2896115](https://doi.org/10.1109/TPDS.2019.2896115)]
- 6 Liu TH, Wei HK, Zhang KJ. Wind power prediction with missing data using Gaussian process regression and multiple imputation. *Applied Soft Computing*, 2018, 71: 905–916. [doi: [10.1016/j.asoc.2018.07.027](https://doi.org/10.1016/j.asoc.2018.07.027)]
- 7 曹原, 刘英博, 肖利, 等. 状态监测数据流时间乱序问题建模与研究. *计算机集成制造系统*, 2013, 19(12): 2960–2967.
- 8 程进, 王坚. 数据驱动的流程制造工艺参数匹配方法. *计算机集成制造系统*, 2017, 23(11): 2361–2370.
- 9 Wang XY, Mueen A, Ding H, *et al.* Experimental comparison of representation methods and distance measures for time series data. *Data Mining and Knowledge Discovery*, 2010, 26(2): 275–309.
- 10 张宙锋, 张磊, 王志超. 改进的数据流缺失数据处理算法. *微电子学与计算机*, 2012, 29(3): 55–59.
- 11 王凯南, 金立左. 基于高斯混合模型的EM算法改进与优化. *工业控制计算机*, 2017, 30(5): 115–116, 118. [doi: [10.3969/j.issn.1001-182X.2017.05.049](https://doi.org/10.3969/j.issn.1001-182X.2017.05.049)]
- 12 许江, 陈志奎, 张清辰. 基于嵌套滑动窗口的数据流缺失数据填充算法. *西南师范大学学报(自然科学版)*, 2015, 40(11): 130–136.
- 13 Liu GY, Yang J, Hao Y, *et al.* Big data-informed energy efficiency assessment of China industry sectors based on K-means clustering. *Journal of Cleaner Production*, 2018, 183: 304–314. [doi: [10.1016/j.jclepro.2018.02.129](https://doi.org/10.1016/j.jclepro.2018.02.129)]
- 14 徐卫, 朱翔宇, 刘志. 城市多源交通数据分析处理关键技术研究. *浙江工业大学学报*, 2018, 46(3): 305–309, 315. [doi: [10.3969/j.issn.1006-4303.2018.03.012](https://doi.org/10.3969/j.issn.1006-4303.2018.03.012)]
- 15 李娜, 黄孝彬, 田志强, 等. 生产过程数据稳定性判断的一种方法. *华电技术*, 2010, 32(1): 32–34, 38. [doi: [10.3969/j.issn.1674-1951.2010.01.011](https://doi.org/10.3969/j.issn.1674-1951.2010.01.011)]