

# 基于 IK-medoids 算法的飞机油耗聚类方法

陈静杰<sup>1,2,3</sup> 车 洁<sup>1,2</sup>

(中国民航大学电子信息与自动化学院 天津 300300)<sup>1</sup>

(中国民航环境与可持续发展研究中心(智库) 天津 300300)<sup>2</sup>

(综合交通大数据应用技术国家工程实验室 天津 300300)<sup>3</sup>

**摘 要** 为了分析给定外界条件下的飞机燃油消耗,提出了一种基于距离最大法的邻域搜索 K-medoids 聚类算法 (IK-medoids)。基于距离最大的样本不可能被分到同一类簇的思想,该算法首先采用距离最大法选取初始中心,并根据剩余样本与初始中心之间的标准欧氏距离计算初始中心邻域;然后利用提出的一种近邻搜索策略进行初始中心的迭代更新,直到中心点不再发生变化。在同一机型和航段、不同大小的数据集上进行对比实验,根据起飞重量、巡航高度、实飞距离以及飞行环境等特征对飞机油耗进行精准分类。实验结果表明:相对于传统的改进 K-medoids 算法,IK-medoids 算法在有效缩短分类时间的同时保证了聚类准确率,为进一步分析飞行过程中的燃油消耗提供了新视角。

**关键词** K-medoids 聚类算法,距离最大法,标准欧氏距离,近邻搜索,油耗分类,Quick Access Recorder(QAR)数据  
中图分类号 TP391.9 文献标识码 A DOI 10.11896/j.issn.1002-137X.2018.08.055

## IK-medoids Based Aircraft Fuel Consumption Clustering Algorithm

CHEN Jing-jie<sup>1,2,3</sup> CHE Jie<sup>1,2</sup>

(College of Electronic Information and Automation, Civil Aviation University of China, Tianjin 300300, China)<sup>1</sup>

(Research Center for Environment and Sustainable Development of CAAC, Tianjin 300300, China)<sup>2</sup>

(National Engineering Laboratory for Integrated Traffic Data Application Technology, Tianjin 300300, China)<sup>3</sup>

**Abstract** To analyze the aircraft fuel consumption in given external environment, this paper proposed a neighborhood search K-medoids clustering algorithm (IK-medoids) based on the maximum distance method. According to the idea that the sample points with the farthest distance cannot be divided into the same cluster, the maximum distance method is used to select the initial center. And then, the center neighborhood is determined by the standardized Euclidean distance between the initial center and rest samples. What's more, the regeneration of initial center is conducted by the proposed nearest neighbor searching strategy, efficiently reducing the iteration time. The contrast experiments were conducted on datasets with different size of the same aircraft model and flight segment, so as to classify the fuel flow data according to the gross weight, cruise altitude, flight distances and flight environment. The results demonstrate that the proposed IK-medoids algorithm outperforms common K-medoids algorithms, and provides a new angle for further analysis on the fuel consumption in flight process.

**Keywords** K-medoids clustering algorithm, Maximum distance method, Standard euclidean distance, Nearest neighbor searching strategy, Fuel consumption classification, QAR data

## 1 引言

除飞机自身能效、飞行员操纵以及飞行环境外,空中交通流量的限制也会影响飞机在飞行过程中的燃油消耗。因此,对于同一机型、航段的飞机,由于受起飞重量、巡航高度、飞行环境以及流量等的影响,其燃油消耗将不尽相同。为了分析

不同飞行特征对油耗的影响,需要根据影响油耗的主要飞行特征对飞机油耗数据进行精准分类。

目前,典型的基于划分方法的聚类算法有 K-means 和 K-medoids 算法<sup>[1]</sup>,它们能够按照给定条件将数据集聚集为若干个类簇。K-medoids 算法弥补了 K-means 算法对孤立点敏感的缺陷,但其本身对初始中心敏感,极易陷入局部最优,

到稿日期:2017-05-24 返修日期:2017-08-12 本文受国家科技支撑项目(2012BAC20B03),民航局节能减排专项计划项目(DPDSR0010)资助。

陈静杰(1967—),男,博士,教授,主要研究方向为航空系统优化与仿真、民航运输过程能效管理与碳排放控制,E-mail:jjchen@cauc.edu.cn (通信作者);车 洁(1991—),女,硕士生,主要研究方向为航空系统优化与仿真、民航运输过程能效管理与碳排放控制,E-mail:cj980324310@163.com。

并且初始中心点的随机搜索更新使得算法计算量较大,聚类精度不高。为了弥补对初始中心敏感的缺陷,减小算法的时间复杂度,提高算法的准确率,许多改进的 K-medoids 算法相继被提出。文献[1]利用近邻方差选择初始中心,解决了对初始中心敏感的问题,但是没有考虑类簇间的影响,聚类精度不高。文献[2-3]用粒计算及其改进来初始化聚类中心,降低了初始中心选取的随机性,减小了时间复杂度,但是准确率并没有得到提升。文献[4]提出了基于遗传模拟退火算法的 K-medoids 算法,增强了算法的全局搜索能力,解决了算法易陷入局部最优的问题,但是增加了算法的计算复杂度。文献[5]提出了基于密度可达思想的中心点初始化和搜索方法,减小了算法的时间复杂度,但是没有考虑类簇间的影响,准确率不高。

针对传统 K-medoids 算法的缺陷及其目前的研究现状,本文提出了一种基于距离最大法的邻域搜索 K-medoids 算法来对飞行过程中的油耗进行分类,进而对比分析飞机在不同外界条件下的燃油消耗情况。该算法基于距离最远的样本点不可能被分到同一簇的思想<sup>[6]</sup>,利用油耗数据样本间的标准欧氏距离选取初始中心,以解决传统 K-medoids 算法对初始中心敏感的问题。本文还提出了一种近邻搜索策略,将中心

点的迭代更新限制在其对应的近邻领域中,降低了算法的时间复杂度;并采用类间距离和类内距离优化适应度函数<sup>[2]</sup>来提高聚类准确率。

## 2 飞行过程中油耗分类数据的选取

参考目前已有的飞机油耗估计模型<sup>[7-10]</sup>,影响飞机油耗的飞行特征可以通过 Quick Access Recorder(QAR)数据中与油耗相关的参数统计分析得到。

为了最小化机型以及起降机场的环境影响,选择同一机型、航段下的 QAR 数据进行研究。本文选取国内常用机型 A330 飞机、北京-上海航段的某航空公司实际飞行过程中的油耗数据作为研究对象,选择起飞重量(总重)、巡航高度、风速、风向、阶梯爬升次数、阶梯下降次数以及实飞距离作为影响飞行过程中油耗的特征指标。

输入数据可以利用数据矩阵表示,矩阵的每一行描述一个航班的 8 个飞行特征。输入数据的示例如表 1 所列。其中,起飞重量、巡航高度、阶梯爬升和下降次数直接从航班 QAR 数据中得到,风速、风向是通过巡航阶段的数据加权平均得到的,实飞距离则由 QAR 数据中经纬度参数求出的每秒大圆距离累加得到。

表 1 输入数据样本  
Table 1 Input data samples

总重/kg	巡航高度/ft	风速/knot	风向/deg	实飞距离/km	阶梯爬升次数	阶梯下降次数	油耗/kg
165309.98	37065.18	40.82	298.83	1430.30	0	2	9080.38
181258.56	33083.00	37.12	302.59	1384.06	1	5	9423.64
184252.32	33085.62	38.87	306.36	1418.36	3	3	10192.08
177339.46	35082.53	48.15	271.95	1439.33	1	3	9955.71
...							

### 1) 风速、风向

$$f = \sum_{i=1}^n f_i \times w_i \quad (1)$$

其中,  $f_i$  为巡航阶段第  $i$  个数据对应的风速或风向;  $w_i = \frac{F_i}{\sum_{i=1}^n F_i}$ ,  $F_i$  为第  $i$  个数据对应的燃油流量;  $n$  为巡航阶段的数据总量,由于 QAR 数据中燃油流量每秒记录一次,因此也将  $n$  称为巡航时间。

### 2) 实飞距离

将根据 QAR 数据中经纬度参数求出的每秒大圆距离累加,即可得到实飞距离。大圆距离的计算是利用地面上任意两点的经纬度,将经纬度数据转换到球面坐标,求出两点间的直线距离,并根据地球半径求得对应的劣弧长度,即为大圆距离,主要通过 WGS-84 坐标系下的大圆距离算法实现<sup>[10]</sup>。

## 3 基于 IK-medoids 算法的飞机油耗分类方法

### 3.1 基于距离最大法的初始化

为了解决 K-medoids 算法对初始中心敏感的问题,基于距离大的样本的相似度小且被分到同一个类簇的可能性小的事实,将距离最远的两个样本点作为初始聚类中心<sup>[6]</sup>。由于油耗分类数据集中且各个数据指标的量纲不同,为了抑制传统欧氏距离计算中数据量纲不同带来的影响,考虑各个指标取值的差异程度,本文用标准欧氏距离代替传统的欧氏距离。

同时,为了缩短算法的计算时间,一次性计算  $N$  个数据样本间的距离而生成矩阵  $Dis$ 。选取距离矩阵中距离最大的两个样本作为初始聚类中心  $o_1$  和  $o_2$ ; 在剩余的  $N-2$  个点中选取到  $o_1$  和  $o_2$  各自距离乘积最大的样本点作为初始中心  $o_3$ ; 依次类推,找到  $k$  个初始中心。具体步骤如下:

1) 计算  $N$  个数据样本间的标准欧氏距离,生成距离矩阵  $Dis$ 。样本  $x_i$  和  $x_j$  间的标准欧氏距离可以表示为:

$$dis(i, j) = \sqrt{(x_i - x_j)^T V^{-1} (x_i - x_j)} \quad (2)$$

其中,  $x_i = (x_{i1}, x_{i2}, \dots, x_{i7})$  为输入数据矩阵的一行,包含起飞重量、巡航高度等 7 个飞行特征;  $i, j = 1, 2, \dots, N$ ;  $V^{-1}$  为输入数据的标准差矩阵的逆矩阵。

2) 比较  $N$  个数据样本两两间的距离大小,找到满足  $\{dis(1, 2) \geq dis(i, j), i, j = 1, 2, \dots, N\}$  的两个样本  $x_1, x_2$  作为初始聚类中心  $o_1, o_2$ , 并从原输入数据集中删除样本  $x_1, x_2$ 。

3) 根据已选的初始聚类中心  $o_1$  和  $o_2$ , 计算两个初始中心各自到剩余的  $N-2$  个样本的距离乘积,选取满足  $\{dis(1, 3) \times dis(2, 3) \geq dis(1, i) \times dis(2, i)\}$  的样本  $x_3$  作为初始中心  $o_3$ , 并从数据集中删除该样本。其中,  $i$  为除去初始聚类中心  $o_1$  和  $o_2$  的任意一个样本。

4) 依次类推,可以得到  $K$  个初始聚类中心。

### 3.2 近邻搜索策略

传统的 K-medoids 算法的时间开销主要来自初始中心点的搜索更新和类簇的重新分配,耗时较多。为了克服传统聚

类算法中心点搜索的随机性,减小算法的时间复杂度,根据赵湘民等提出的稠密区域中心点搜索更新策略<sup>[5]</sup>,提出了近邻搜索策略进行初始中心的迭代更新。

为定义初始中心邻域,首先根据距离矩阵选择离初始中心  $o_l$  最近的  $num$  个样本作为初始中心的近邻样本。记近邻样本集合为  $near = \{x_{l1}, \dots, x_{lnum}\}$ , 则  $o_l$  的邻域表示为:

$$neighbourhood(o_l) = \{x_i \mid dis(i, l) \leq aver(o_l)\} \quad (3)$$

其中,  $aver(o_l) = \frac{\sum_{j=1}^{num} d(j, l)}{num}$  为近邻样本的平均距离,  $i = 1, 2, \dots, N-K$ , 确定为邻域的样本将从数据集中删除, 以避免初始中心出现在同一簇中。根据式(3)依次计算得到  $K$  个初始中心对应的邻域。

将初始中心的搜索更新范围限定在  $K$  个初始相异的邻域内, 从而避免传统 K-medoids 算法中心点全局搜索更新的盲目性和随机性, 降低了算法的计算代价。

### 3.3 适应度函数

传统 K-medoids 算法的准则函数如式(4)所示:

$$E(w) = \sum_{i=1}^K \sum_{p \in c_i} dis^2(p, o_i) \quad (4)$$

将簇内距离平方和最小作为衡量聚类质量的标准, 忽略了聚类的核心思想: 类内距离最小, 类间距离最大。式(4)并未考虑类间距离对聚类质量的影响。本文利用潘楚等<sup>[2]</sup>提出的适应度函数衡量聚类效果。

定义类间距离  $O(w)$  为不同类簇间的距离:

$$O(w) = \sum_{i, j=1}^K dis^2(o_i, o_j) \quad (5)$$

则适应度函数为:

$$F(w) = \frac{\sqrt{E(w)}}{\sqrt{O(w)}} \quad (6)$$

根据聚类核心思想, 适应度函数值  $F(w)$  最小时, 聚类效果最优。

### 3.4 聚类算法的具体实现

聚类算法的具体过程如下:

输入: 飞机油耗分类数据矩阵, 簇的个数  $K$  以及初始中心的近邻样本个数  $num$ 。

输出: 效果最优的  $K$  个类簇和适应度函数值的总和。

1) 利用 3.1 节中描述的 1)–4) 选取  $K$  个初始聚类中心, 记为  $o_1$  到  $o_K$ 。

2) 根据 3.2 节定义的初始中心邻域计算  $K$  个初始中心的邻域, 记为  $neighbourhood(o_l)$ 。

3) 根据 3.1 节中步骤 1) 中计算的标准欧氏距离矩阵, 将剩余的  $N-K$  个样本归入距离最近的簇, 并计算每个簇的适应度函数  $F(w)$ 。

4) 在初始中心  $o_1$  到  $o_K$  对应的邻域  $neighbourhood(o_l)$  内选择一点  $o_{random}$  代替中心点对象  $o_l$ , 更新每个簇中心, 并重新

聚类, 计算适应度函数  $F(w')$ 。

5) 比较  $F(w)$  与  $F(w')$  的大小, 若  $F(w') \leq F(w)$ , 则用  $o_{random}$  替换  $o_l$ ,  $F(w') = F(w)$ 。

6) 重复步骤 4) 和步骤 5), 直到  $F(w)$  最小。

## 4 实验结果及分析

本实验基于某航空公司的 A330 飞机在北京-上海航段的 QAR 数据, 分别利用文献[1]中的基于 Num-近邻方差优化的 K-medoids(Num-Kmedoids)算法、文献[5]中基于稠密区域的 K-medoids(d-Kmedoids)算法以及本文提出的 IK-medoids 算法进行油耗数据的分类。通过类簇总的适应度值以及类簇的分布情况评价各个算法的聚类精度, 通过聚类时间来评价算法的时间复杂度。

类簇的总适应度值 ( $fit$ ) 为:

$$fit = \sum_{w=1}^K F(w) \quad (7)$$

其中,  $F(w)$  的定义如式(6)所示, 为每个类簇的适应度值。为了对比分析数据规模对不同 K-medoids 算法的影响, 分别从油耗分类数据中选取样本大小为 100, 215, 415 以及 700 的数据集进行多次实验。将数据规模为 100 的第一个数据集作为训练集来训练算法参数。根据聚类算法的核心思想, 从图 1 中可以看出, 当类簇个数增加时, 总适应度值逐渐减小, 趋于稳定。文中选择  $K=7$ , 此时适应度值随近邻样本个数的变数而变化, 如图 1 中菱形图标线所示。因此, 对于本文选取的数据集, 选择  $num=15$ 。同理, 在数据集规模大于 200 时设定  $num=25$ , 并设定 d-Kmedoids 中,  $N_d=5$ ,  $Minpts=3$ 。

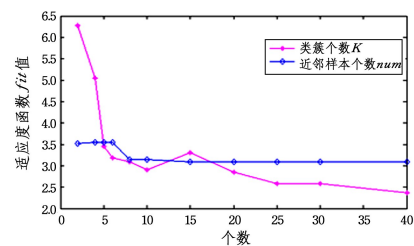


图 1 适应度值随类簇和近邻样本个数的变化

Fig. 1 Variation of fit function value with the number of clusters and nearest neighbors

聚类精度通过式(7)评价, 各算法的聚类精度和算法耗时如表 2 所列。从表 2 可以看出, 与 d-Kmedoids 和 Num-Kmedoids 算法相比, IK-medoids 算法的  $fit$  值最小, 聚类精度相对较高, 算法耗时最少, 并且数据集规模大小对 IK-medoids 算法的聚类精度和算法耗时的影响较小。表 2 的结果表明, 本文提出的算法有效改善了现有 K-medoids 算法的聚类性能, 提升了算法的伸缩性。

表 2 各个聚类算法在不同规模数据集上的聚类效果

Table 2 Clustering effect of investigated clustering algorithms on different scale datasets

样本大小	适应度值 $fit$					算法耗时/s				
	100(训练集)	100	215	415	700	100(训练集)	100	215	415	700
Num-Kmedoids	3.514	4.091	8.842	11.435	29.485	13.564	12.333	63.483	180.755	443.008
d-Kmedoids	3.716	3.674	5.046	12.779	14.162	14.941	14.870	45.954	89.168	86.838
IK-medoids	3.299	3.310	4.687	6.439	8.687	4.480	4.678	12.016	21.770	27.051

为了更直观地分析各个算法的聚类效果,绘制数据规模为 100 时的第二个油耗分类数据在不同总重和实飞距离下的分类效果图,不同算法的类簇分布情况如图 2—图 4 所示。对比图 2—图 4 发现,d-Kmedoids 算法与 Num-Kmedoids 算法类簇间的重叠较多;Num-Kmedoids 算法类簇的类内距离较大,聚类效果不理想。从图 4 可以看出,本文所提算法的类内距离较小,类簇间几乎没有重叠,因此对于油耗分类数据更为适用,能够对油耗数据进行精确的分类。此时,同一类簇内的燃油消耗较为相似,可以分析此时同一类簇的飞行特征,确定给定条件下的燃油消耗量,从而为飞行过程中的燃油消耗预测提供可靠的依据。

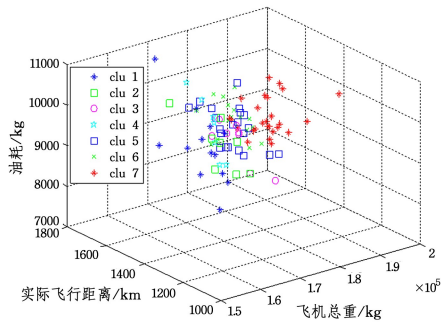


图 2 Num-Kmedoids 算法的聚类效果

Fig. 2 Clustering effect of Num-Kmedoids algorithm

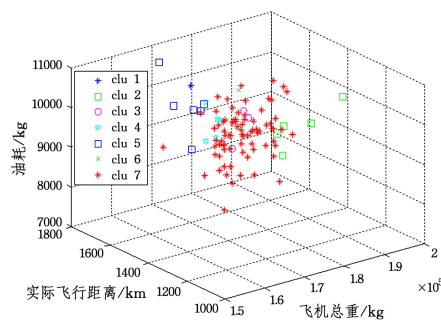


图 3 IK-medoids 算法的聚类效果

Fig. 3 Clustering effect of IK-medoids algorithm

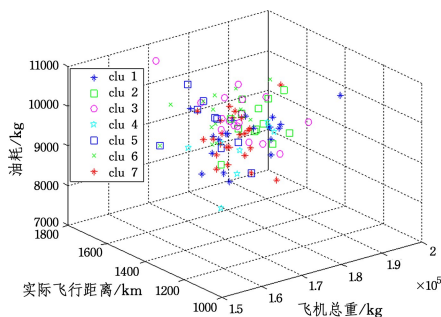


图 4 d-Kmedoids 算法的聚类效果

Fig. 4 Clustering effect of IK-medoids algorithm

**结束语** 本文在分析现有 K-medoids 聚类算法的适用性及不足的基础上,参考样本相似性度量方法,提出了基于距离最大法的邻域搜索 K-medoids 算法,以期对飞行过程中的油耗进行精准分类。通过引入标准欧氏距离来度量样本间的相

似性,减小了样本各参量纲不同带来的误差,使得距离最大法的使用更为准确,弥补了传统 K-medoids 算法对初始中心敏感的缺陷。中心点的邻域搜索更新在很大程度上减少了算法的迭代更新时间。实验证明,该算法能够对飞行过程中的油耗进行精确分类,为航空公司分析飞机在飞行过程中的油耗提供了新视角。该方法具有聚类精度高、时间复杂度低以及伸缩性好等特点,并且能够适用于大数据集的聚类。

## 参 考 文 献

- [1] XIE J Y,GAO R. Optimized K-medoids clustering algorithm by variance of Num-near neighbour[J]. Application Research of Computers,2015,32(1):30-34. (in Chinese)  
谢娟英,高瑞. Num-近邻方差优化的 K-medoids 聚类算法[J]. 计算机应用研究,2015,32(1):30-34.
- [2] PAN C,LUO K. Improved K-medoids clustering algorithm based on improved granular computing[J]. Journal of Computer Applications,2014,34(7):1997-2000. (in Chinese)  
潘楚,罗可. 基于改进粒计算的 K-medoids 聚类算法[J]. 计算机应用,2014,34(7):1997-2000.
- [3] SONG H H,YAN H W. Novel K-medoids clustering algorithm based on dynamic search of microparticles under optimized granular computing[J]. Intelligent Computer and Applications,2016,6(2):9-13. (in Chinese)  
宋红海,颜宏文. 基于优化粒计算下微粒子动态搜索的 K-medoids 聚类算法[J]. 智能计算机与应用,2016,6(2):9-13.
- [4] HAN X,LIU S F,XU T Q. Improved K-medoids algorithm based on genetic simulated annealing algorithm[J]. Journal of Jilin University (Engineering and Technology Edition),2015,45(2):619-623. (in Chinese)  
韩啸,刘淑芬,徐天琦. 基于遗传模拟退火算法的改进 K-medoids 算法[J]. 吉林大学学报(工学版),2015,45(2):619-623.
- [5] ZHAO X M,CHEN X,PAN C. Novel K-medoids clustering algorithm based on dense regional block[J]. Computer Engineering and Applications,2016,52(16):85-89. (in Chinese)  
赵湘民,陈曦,潘楚. 基于稠密区域的 K-medoids 聚类算法[J]. 计算机工程与应用,2016,52(16):85-89.
- [6] ZHAI D H,YU J,GAO F,et al. K-means text clustering algorithm based on initial cluster centers selection according to maximum distance[J]. Application Research of Computers,2014,31(3):713-715,719. (in Chinese)  
翟东海,鱼江,高飞,等. 最大距离法选取初始簇中心的 K-means 文本聚类算法的研究[J]. 计算机应用研究,2014,31(3):713-715,719.
- [7] CHEN J J,YAN Y,LIU J X. Fuel Consumption Estimation for Cruise Phase Based on BP Neural Network and Its Analysis of Structural Factors [J]. Computer Measurement & Control,2015,23(6):2135-2138. (in Chinese)  
陈静杰,颜艳,刘家学. 基于 BP 网络巡航段油耗估计及影响结构分析[J]. 计算机测量与控制,2015,23(6):2135-2138.

另外,已有的方法直接删除了条件非频繁行为,对频繁行为进行分析处理。而本文并没有全部去掉条件非频繁行为,而是适当地保留了部分含有重要信息的行为。相比于已有算法,本文方法在时间和空间效率上的代价会更大,基于此,本文没有在时间和空间上进行比较分析。

**结束语** 本文在已有研究的基础上,给出了基于通讯行为轮廓挖掘条件非频繁行为的优化方法。该方法以模块网之间的通讯行为轮廓为基础,首先给定业务过程源模型,查询其可执行事件日志,找出频数较低的事件轨迹并赋予其属性值。依据属性化的条件非频繁轨迹,计算不同模块网之间活动对的条件依赖度量值并进行比较,从而确定轨迹是否被保留或者删除,得到优化事件日志。依据优化事件轨迹集挖掘出业务流程优化通讯模型。

未来将以不同模块间的通讯行为轮廓理论为基础,对于事件数量较多的频数较低事件轨迹集,挖掘带有配置信息的非频繁行为,以提高业务过程系统的运作效率。

### 参考文献

- [1] VAN DER AALST W M P. Process mining data science in action(Second Edition)[M]. Springer,2016.
- [2] WEERDT J D,BACKER M D,VANTHIENEN J,et al. A multi-dimensional quality assessment of state-of-the-art process discovery algorithms using real-life event logs[J]. Information Systems,2012,37(7):654-676.
- [3] SURIADI S,ANDREWS R,TER HOFSTED E A H M,et al. Event log imperfection patterns for process mining: Towards a systematic approach to cleaning event logs[J]. Information Systems,2017,64(C):132-150.
- [4] LIESAPUTRA V,YONGCHAREON S,CH AISIRI S. Efficient process model discovery using maximal pattern mining[C]// International Conference on Business Process Management. Springer,Cham,2015:441-456.
- [5] PONCE-DE-LEÓN H,CARMONA J,VANDEN BROUCKE S K L M. Incorporating negative information in process discovery [C]// International Conference on Business Process Management. Springer,2015:126-143.
- [6] BELLODI E,RIGUZZI F,LAMMA E. Statistical relational learning for workflow mining[J]. Intelligent Data Analysis,2016,20(3):515-541.
- [7] CONFORTI R,ROSA M L,HOFSTED E T,et al. Filtering out infrequent behavior from business process event logs[J]. IEEE Transactions on Knowledge & Data Engineering,2017,29(2),300-314.
- [8] ZELST S J V,DONGEN B F V,AALST W M P V D,et al. Discovering relaxed sound workflow nets using integer linear programming[J]. Computing,2017(4).
- [9] LEEMANS S J J,FAHLAND D,VAN DER AALST W M P. Discovering block-structured process models from event logs containing infrequent behaviour[C]// International Conference on Business Process Management. Springer,Cham,2013:66-78.
- [10] VAN DER WERF J M,KAATS E. Discovery of functional architectures from event logs[C]// PNSE@ Petri Nets. 2015:227-243.
- [11] MANNHARDT F,DE LEONI M,REIJERS H A,et al. Data-driven process discovery-revealing conditional infrequent behavior from event logs[C]// International Conference on Advanced Information Systems Engineering. Springer, Cham, 2017: 545-560.
- [12] VAN DER AALST W M P,WEIJTERS T,MARUSTER L. Workflow mining: discovering process models from event logs [J]. IEEE Transactions on Knowledge Data Engineering,2004,16(9):1128-1142.
- [11] WANG X H,CHEN X F. A Support Vector Method for Modeling Civil Aircraft Fuel Consumption with ROC Optimization [C]//2014 Second International Conference on Enterprise Systems. Shanghai,China,2014:112-116.
- [12] AROCRA P,DEEPALI D,VARSHNEY S. Analysis of K-Means and K-Medodis Algorithm For Big Data[C]// International Conference on Information Security & Privacy(ICISP2015). Nagpur,INDIA,2015:507-512.
- [13] HAN L S, XIANG L S, LIU X Y, et al. The K-medodis Algorithm with Initial Centers Optimized Based on a P System[J]. Journal of Information & Computational Science,2014,11(6):1765-1773.
- [14] HE Y,PI D C. Iterative Imputation Algorithm Based on Reduced Relational Grade for Gene Expression Data[J]. Computer Science,2015,42(11):251-255,283. (in Chinese)  
何云,皮德常. 基于精简关联度的基因表达数据迭代填补算法[J]. 计算机科学,2015,42(11):251-255,283.

(上接第309页)

- [8] LIU J X,MA T,CHEN J J. RELAX-Based Method for Aircraft Fuel Consumption Performance Evaluation[J]. Electronics Optics & Control,2015,22(9):101-105. (in Chinese)  
刘家学,马涛,陈静杰. 基于 RELAX 算法的飞机油耗性能估计方法[J]. 电光与控制,2015,22(9):101-105.
- [9] CHEN J J,LI L Q. Fuel Consumption Classification of Aircraft in Descent Based on K-means Algorithm[J]. Measurement & Control Technology,2015,34(11):16-19. (in Chinese)  
陈静杰,李吕琪. 基于 K-means 算法的飞机下降过程油耗分类[J]. 测控技术,2015,34(11):16-19.
- [10] CHEN J J,XIAO G P. Analysis tool design of aircraft fuel consumption[J]. Computer Engineering and Design,2014,35(11):4012-4016. (in Chinese)  
陈静杰,肖冠平. 飞机油耗分析工具设计[J]. 计算机工程与设计,2014,35(11):4012-4016.