

基于 HDP-CHMM 的机械设备性能退化评估^{*}

王 恒, 季 云, 朱龙彪, 刘 肖

(南通大学机械工程学院 南通, 226019)

摘要 针对传统隐马尔可夫模型(hidden Markov model, 简称 HMM)状态数必须预先设定的不足,提出了一种基于分层狄利克雷过程-连续隐马尔可夫模型(hierarchical Dirichlet process-continuous hidden Markov model, 简称 HDP-CHMM)的机械设备性能退化评估方法。该方法利用分层狄利克雷模型的分层聚类原理,在狄利克雷过程(Dirichlet process, 简称 DP)模型的基础上进行扩展,利用多组关联数据实现了模型结构根据观测数据的自适应变化和动态调整,获得设备运行过程中的最优退化状态数,并结合连续隐马尔可夫模型(continuous hidden Markov model, 简称 CHMM)良好的分析和建模能力,获得设备退化状态转移路径,实现机械设备运行过程中的退化状态识别和性能评估。利用滚动轴承全寿命数据的多组特征值进行了应用研究,并与基于 K-S 检验算法的机械设备零部件性能退化评估方法进行了比较。结果表明,HDP-CHMM 模型可以对轴承实际运行状态转移过程进行建模,有效识别轴承运行中的不同退化状态,为基于状态的设备维修提供了理论指导。

关键词 分层狄利克雷模型;连续隐马尔可夫模型;性能退化评估;滚动轴承

中图分类号 TH 165.3

引 言

为了保证机械设备长期安全运行,工业现场需要对机械设备进行状态监测和性能评估,实现预知维护和管理。隐马尔可夫模型为一双重随机过程:一个是 Markov 链,用来描述隐藏状态间的转移概率,此为基本随机过程;另一个是描述每个隐藏状态下产生观测值的一般随机过程。HMM 这一双重随机特性,能够很好地描述机械设备运行过程中观测到的退化征兆信号(如振动、转速和位移等)与隐藏的衰退状态之间的随机关系,在机械设备性能退化评估与预测中得到广泛应用^[1-2]。

传统 HMM 的定义和估计过程存在严重不足,例如,HMM 参数训练学习中需要预先确定模型的数学结构,不能解决在模型训练过程中出现的过适应或欠适应问题。刘新民等^[3]建立了基于支持向量机(support vector machine, 简称 SVM)与 HMM 串联结构的故障诊断模型,实现了对直升机减速箱的状态识别和诊断。张继军等^[4]利用 HMM 自身的状态识别和转移回溯能力,结合多智能体遗传算法(multi-agent genetic algorithm, 简称 MAGA)实现了温控放大器的状态分类。上述文献中 HMM

的初始状态数都是依据传统经验设定的,缺乏科学性和通用性。曾庆虎^[5]提出了基于最小描述长度准则(minimum description length, 简称 MDL)学习算法,通过 MDL 优化调整状态数,但模型状态数还是需要预先确定,计算过程较为复杂。腾红智等^[6]基于 CHMM 对齿轮箱全寿命过程的退化状态识别进行了研究,提出了基于 K 均值算法和交叉验证相结合的状态数优化方法,但 K 均值聚类算法仍需要预先确定状态数,且交叉验证对不同的退化状态数都要进行训练并检验分类错误率,计算时间长,效率低。张星辉等^[7]建立了一组聚类方法评价指标,利用 K 均值聚类算法对状态特征进行聚类,通过指标评定结果从中选取模型的最优状态数。但是设备运行过程中,随着监测数据的更新,退化状态数也需要随之不断更新,关于如何有效地确定最优退化状态数还需要进一步深入研究。

针对 HMM 参数的优化问题,提出了一种基于分层狄利克雷过程(hierarchical Dirichlet process, 简称 HDP)和 HMM 相结合的机械零件性能退化评估方法,将 HDP 引入 HMM 中,结合 HMM 良好的分析和建模能力,实现运行过程中的退化状态识别。轴承退化性能评估结果表明了该方法的有效性和适用性。

^{*} 国家自然科学基金资助项目(51405246);江苏省自然科学基金面上资助项目(BK20151271);南通市应用基础研究-工业创新资助项目(GY12016010);江苏省“六大人才高峰”高层次人才资助项目(GDZB-048)

收稿日期:2016-11-02;修回日期:2017-03-28

1 HDP 模型

1.1 DP 定义

狄利克雷过程定义为关于一组分布或者随机测度的分布,假设参数服从一类样本空间上的宽先验分布,参数的后验分布通过采样推断,该狄利克雷模型及其扩展模型则具有良好的聚类特性。近几年,狄利克雷模型已经在机器学习、生物信息学、文本聚类和图像分割等方面有较好的应用。Ferguson 首次提出狄利克雷过程的定义, G_0 为测度空间 Θ 上的随机概率测度,参数 α 为正实数。对于测度空间 Θ 的任意有限划分 A_1, \dots, A_r , 如果存在如下关系 $(G(A_1), \dots, G(A_r)) \sim \text{Dir}(\alpha G_0(A_1), \dots, \alpha G_0(A_r))$

(1)

则 G 服从由基分布 G_0 和参数 α 组成的 DP, 即

$$G \sim \text{DP}(\alpha, G_0) \quad (2)$$

1.2 HDP 过程

狄利克雷过程可以实现一组数据的聚类,但在研究多组数据的聚类问题时,单纯利用狄利克雷混合模型无法进行建模分析,这大大限制了其应用。针对这个问题,引入分层狄利克雷过程^[8-9]。分层狄利克雷过程将基础分布 G_0 扩展为一个服从狄利克雷过程的随机概率测度 G , 即 $G \sim \text{DP}(\gamma, H)$, 多个数据源 $G_j \sim \text{DP}(\alpha, G)$ 将共享基础分布 G 的离散原子,实现多组关联数据的聚类。

HDP 无法实现对其过程的采样,在实际应用中往往采用不同形式的构造实现 HDP 过程的应用。HDP 的 Stick-breaking 构造可以分为两层进行^[10]。第 1 层 Dirichlet 过程为

$$\begin{cases} \varphi_k \sim H(\lambda) & (k=1, 2, \dots) \\ \beta \sim \text{GME}(\gamma) \\ G(\varphi) = \sum_{k=1}^{\infty} \beta_k \delta(\varphi, \varphi_k) \end{cases} \quad (3)$$

其中: H 为基础分布; β 为第 1 层分布的权重; k 为任意初始聚类数; φ_k 为抽样的原子。

第 2 层每组 G_j 的 Stick-breaking 构造方式为

$$\begin{cases} G_j(\varphi) = \sum_{k=1}^{\infty} a_{jk} \delta(\varphi, \varphi_k) \\ a_j = (a_{jk})_{k=1}^{\infty} \end{cases} \quad (4)$$

其中: G_j 为扩展分布; a_j 为第 2 层分布中每组的权重。

在 HDP 模型中,狄利克雷过程作为参数的先验分布存在,假设模型中的观测数据 x_{ji} 表示第 j 组第 i 个观测数据,其分布形式定义如式(5),两层狄利克雷过程分别用上述的 Stick-breaking 构造^[10]。

$$\begin{cases} x_{ji} | \theta_{ji} \sim F(\theta_{ji}), & \theta_{ji} | G_j \sim G_j \\ G_j | \alpha, G \sim \text{DP}(\alpha, G), & a_j | \alpha, \beta \sim \text{Dir}(\alpha, \beta) \\ G | \gamma, H \sim \text{DP}(\gamma, H), & \beta | \gamma \sim \text{Dir}(\gamma) \end{cases} \quad (5)$$

HDP 模型参数的含义如下: x_{ji} 为观测数据; k 为类别标签的初始聚类数; γ, α 为 HDP 模型的超参数; β, a_j 为 HDP 模型每层的权重系数; λ 为基础分布的采样参数; φ_k 为随机抽样的离散原子。

2 基于 HDP-CHMM 的设备性能退化评估算法

2.1 基于 HDP 的退化隐状态数确定

隐状态数 K 的确定是 HMM 模型训练和测试的关键,但目前 HMM 隐状态数大多根据经验人为设定,或者通过训练一个最可能的隐状态数来达到近似的目的,这些方法很难将各种类别都考虑到,且新的数据中也可能有未知类型出现,依靠训练样本得到的固定模型结构对新的观测数据的适用性和涵盖性不强。HDP 算法不依赖于训练样本,而且随着数据的变化,模型结构能够实现自适应调整,实现动态聚类。基于 HDP 的隐状态数确定算法流程图如图 1 所示,主要步骤如下。

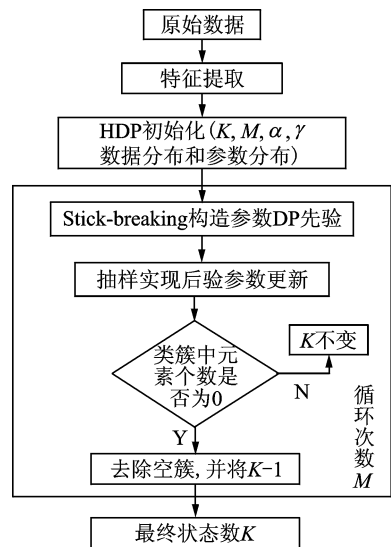


图1 基于 HDP 模型的隐状态数确定

Fig. 1 The number of hidden states of HDP

1) 初始化 HDP 模型的参数,任意给定隐状态数 K , 迭代次数 M , 超参数 α, γ 。

2) 假设输入的多组观测数据 x_{ji} 服从多项式分布,分布密度函数为 $f(\cdot | \theta)$; 观测数据分布的参数 θ_{ji} 服从其共轭分布 DP 分布,分布密度函数为 $h(\cdot)$; HDP 的权重系数 β 和 a_j 分别通过式(3)和式(4)中的 Stick-breaking 构造先验分布,观测数据 x_{ji}

的条件分布为

$$f_k^{\setminus x_{j_i}}(x_{j_i}) = \frac{\int f(x_{j_i} | \theta) \prod_{j_i' \neq j_i} f(x_{j_i'} | \theta) h(\theta) d\theta}{\int \prod_{j_i' \neq j_i} f(x_{j_i'} | \theta) h(\theta) d\theta} \quad (6)$$

其中: $f_k^{\setminus x_{j_i}}(x_{j_i})$ 为观测数据 x 中除了 x_{j_i} 外的条件概率分布。

3) 基于马尔科夫链蒙特卡罗 (Markov chain Monte Carlo, 简称 MCMC) 统计模拟方法简化条件概率的计算, 并实现模型后验参数的更新, 利用 Augmented representation 后验采样算法^[11] 实现对参数 β 的更新 (a_{jk} 类似), 根据超参数 γ 和 α 对 β 进行采样

$$(\beta_1, \dots, \beta_K, \beta_{\bar{K}}) \sim \text{Dir}(m_{1,1}, \dots, m_{1,k}, \gamma) \quad (7)$$

其中: $m_{1,1}, \dots, m_{1,k}$ 利用直接分配后验采样算法^[12] 进行采样。

$$p(m_{jk} = m | Z, M^{jk}, \alpha, \beta) = \frac{\Gamma(\alpha\beta_k)}{\Gamma(\alpha\beta_k + n_{j,k})} s(n, m) (\alpha\beta_k)^m \quad (8)$$

其中: m_{jk} 为第 j 组的聚类数目; M^{jk} 为除了第 j 组的第 k 类之外的聚类数; $n_{j,k}$ 为第 j 组观测数据中属于 k 类的数据总数; $s(n, m)$ 为 Stirling 数; γ 为权重的第 $K+1$ 个值。

通过对 β 和 m 的交叉采样实现参数更新, 并在 Stick-breaking 构造中运用直接分配后验采样算法对观测数据的指示因子 Z_{j_i} 进行采样, 每次只更新一个数据的聚类属性, 当某个类簇中元素个数为 0 时, K 减 1, 继续迭代, 待聚类数目稳定时停止迭代, 获得最终的状态数目 K 。

4) 超参数 γ 和 α 决定 DP 过程的离散程度, 初始值的选取对聚类结果影响很大, 大大影响了 HDP 模型的通用性^[13]。因此, 在进行后验参数的更新时, 同时对超参数进行更新。设 $\alpha \sim \Gamma(a, b)$, 对任意 $k = 1, 2, \dots, n, P(k | \alpha)$ 服从 Beta 分布, 则

$$p(\alpha | k) \propto p(\alpha) P(k | \alpha) \propto p(\alpha) \alpha^{k-1} (\alpha + n) \beta'(\alpha + 1, n) \quad (9)$$

其中: β' 为 Beta 分布。

从 $\beta'(\alpha + 1, n)$ 中抽样参数 η , 则参数 α 和 η 的联合分布为

$$p(\alpha, \eta | k) \propto p(\alpha) \alpha^{k-1} (\alpha + n) \eta^{\alpha} (1 - \eta)^{n-1} \quad (10)$$

参数 α 的后验分布为

$$p(\alpha | \eta, k) \propto \alpha^{a+k-2} (\alpha + n) e^{-\alpha(b - \log(\eta))} \quad (11)$$

则

$$(\alpha | \eta, k) \sim a_{\eta} G(a + k, b - \log(\eta)) + (1 - a_{\eta}) G(a + k - 1, b - \log(\eta)) \quad (12)$$

其中: a_{η} 为权重。

$(\alpha | \eta, k)$ 服从两个 Gamma 分布之和, 且分布系数分别为 a_{η} 和 $1 - a_{\eta}$, 超参数 α 和 γ 分别从后验分布 $a_{\eta}/(1 - a_{\eta})$ 中不断迭代采样实现更新。

2.2 基于 CHMM 的机械设备退化状态评估

基于 CHMM 的机械设备退化状态评估算法的流程图如图 2 所示, 主要步骤如下。

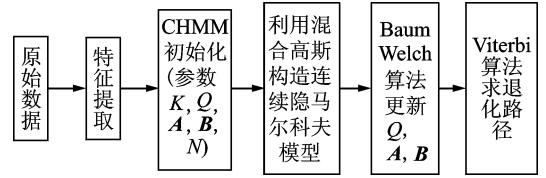


图 2 基于 CHMM 的设备退化状态识别

Fig. 2 Identification of equipment degradation status based on CHMM

1) 根据步骤 1~4 确定的状态数 K 作为 CHMM 模型的输入参数, 在一定约束条件下, 随机初始化模型的其他参数 (初始状态 Q , 转移矩阵 A , 观测值矩阵 B , 混合高斯数 N)。

2) 运用 CHMM 模型进行建模, 将观测数据作为模型的输入, 通过 Baum-Welch 算法对观察数据进行训练, 利用 EM 算法求概率参数模型的最大似然估计, 重估权值 w 、均值 μ 和方差 ξ , 通过不断迭代估计模型参数 $\lambda = (Q, A, B)$ 。

3) 根据训练得到的 CHMM 模型, 利用 Viterbi 算法计算已知模型参数 λ 最可能的隐藏状态序列, 即 $P(O | \lambda)$ 的最大值, 通过路径回溯求得每个观察序列最可能的状态, 即得到退化状态转移曲线。

3 应用

应用研究采用美国 USFI/UCR 的智能维护中心提供的轴承全寿命数据^[14], 图 3 为实验装置。4 个 ZA-2115 双列轴承并列安装在同一轴上, 由恒定转速 2 kr/min 的直流电机驱动, 在轴承和轴上加载约 26 671 N 的弹性径向载荷, 采用加速度传感器采集振动信号, 采样频率为 20 kHz。轴承 1 在连续运

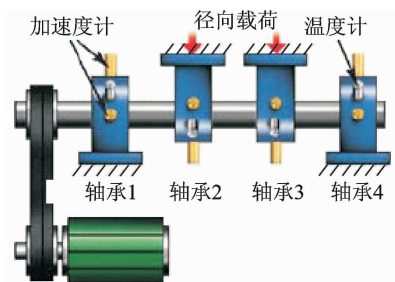


图 3 轴承加速寿命实验装置示意图

Fig. 3 Schematic diagram of the bearing life device

转约 163 h 外圈出现严重损伤,共采集 982 组数据,采用轴承 1 的全寿命数据进行建模与分析。

DP 模型的扩展模型 HDP 利用分层构造 DP 原理,能够实现多组特征值的聚类。笔者基于 HDP 模型对轴承偏度指标(skewness)和峭度指标(kurtosis)进行训练,以获取 CHMM 模型中的最佳退化状态数。以轴承峭度指标为例,分析 HDP 模型中超参数 α 的不同初始值对聚类结果的影响,由图 4 可知,当 α 分别取 2,20,2 000 时,该模型聚类结果均能收敛到相同的值,可见,HDP 模型的超参数初始值选取对聚类结果不敏感。

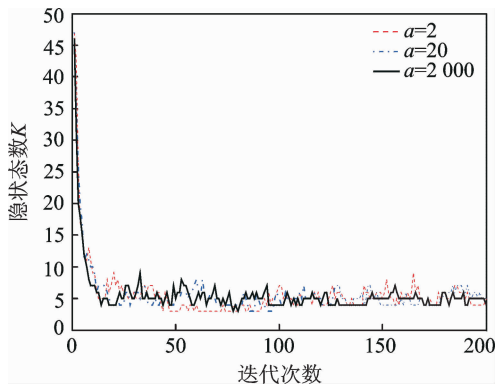


图 4 不同 α 值影响分析

Fig. 4 Effect analysis of different values α

选取峭度指标和偏度指标作为 HDP 模型的观测数据。假设观测数据服从 Multinomial 分布,观测数据分布的参数服从 DP 分布。设定初始聚类数目 $K=50$,聚集参数 $\alpha = 20, \gamma = 1$,迭代次数 $M=300$,通过 Stick-breaking 构造 HDP,利用不断采样更新参数,实现基于 HDP 的自动聚类,结果如图 5 所示,聚类结果趋近于 5,且聚类结果比图 4 中单一特征量输入更稳定。因此,将 $K=5$ 作为 CHMM 模型隐状态数进行模型的训练和测试。

利用混合高斯模型来拟合各状态下的观测值概率密度函数,在构造连续隐马尔科夫模型时,高斯分量数目 N 取 3,基于 Baum-Welch 算法重估参数 ($Q,$

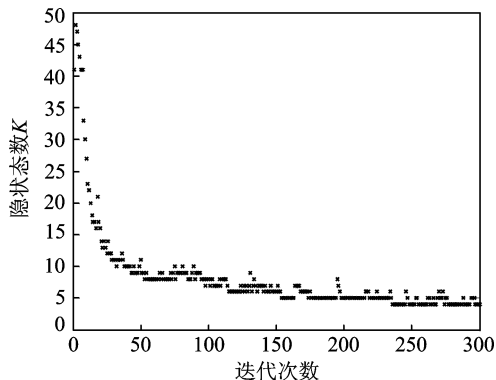


图 5 基于 HDP 模型的多组观测数据混合聚类

Fig. 5 Multi group observation data mixed clustering of HDP

A, B), 获得经过多次迭代的重估模型。利用 Viterbi 算法计算 $P(O|\lambda)$ 的最大值,即得到退化状态转移曲线。

基于 HDP-CHMM 获得的轴承退化状态转移曲线如图 6 所示。轴承从正常状态到失效状态的全寿命历程中一共出现了 5 次不同状态,分别为正常状态、早期退化状态 1、中度退化状态 2、严重退化状态 3 和失效状态。通过该模型可以找出轴承运行时的早期故障点在 576 号文件处,严重故障点在 930 号文件处,并能识别轴承在运行过程中的一系列退化状态,为轴承的早期维护和维修提供理论指导。

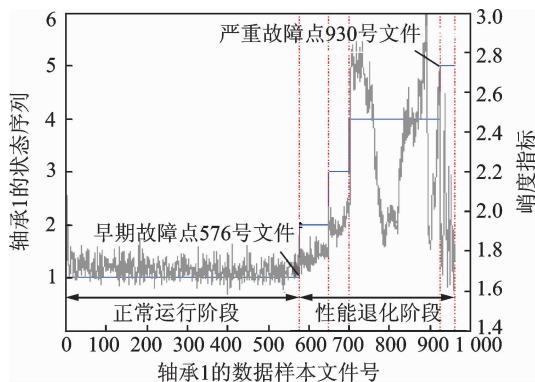


图 6 基于 HDP-CHMM 的轴承退化状态识别

Fig. 6 Identification of bearing degradation status based on HDP-CHMM

为了验证 HDP-CHMM 算法的有效性,与文献 [15]中基于 K-S 检验的轴承退化状态识别方法进行了对比。由图 6,7 可知,轴承 1 全寿命数据在基于 K-S 检验的退化评估中,早期故障点在第 533 序列处,严重故障点在 962 处,与基于 HDP-CHMM 算法识别的早期故障和严重故障点大体一致。在轴承性能退化阶段,由于振动波动异常剧烈,基于距离的 K-S 检验法将轴承振幅跳变剧烈处的状态划分过多,在工作过程中不利于轴承实际状态的判定。

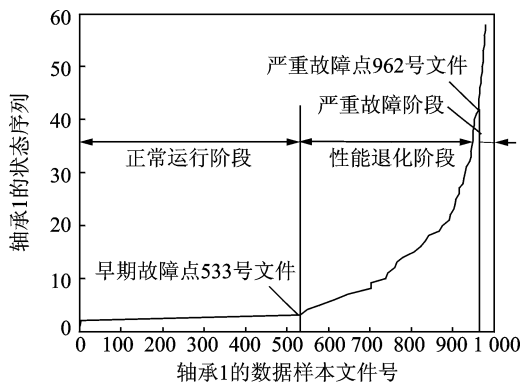


图 7 基于 K-S 检验的轴承退化状态识别

Fig. 7 Identification of bearing degradation status based on K-S test

4 结 论

1) 提出了一种基于 HDP-CHMM 的机械设备性能退化评估方法,利用 HDP 的分层聚类特性,将轴承多个特征指标作为模型的输入,使聚类结果更加准确,解决了传统的 CHMM 状态数必须预先设定的不足,实现了模型结构的自适应调整。

2) 滚动轴承全寿命数据建模结果表明,HDP-CHMM 可以有效地找出轴承运行时的早期故障点和严重故障点,便于设置预警机制,并能够识别轴承在运行过程中的一系列退化状态,为基于状态的设备退化评估提供了一种新的方法。

3) HDP-CHMM 算法和 K-S 检验算法对比可知,HDP-CHMM 算法可以对轴承实际运行过程中的状态转移进行建模,有效识别出轴承运行中的不同退化状态,比 K-S 检验算法基于距离的诊断方法更符合设备实际退化过程。

参 考 文 献

- [1] Peng Ying, Dong Ming. A prognosis method using age-dependent hidden semi-Markov model for equipment health prediction [J]. Mechanical System and Signal Processing, 2011, 25 (1): 237-252.
- [2] Liu Qinming, Dong Ming, Peng Ying. A novel method for online health prognosis of equipment based on hidden semi-Markov model using sequential Monte Carlo methods [J]. Mechanical Systems and Signal Processing, 2012, 32: 331-348.
- [3] 刘新民,刘冠军,邱静. 基于 HMM-SVM 的故障诊断模型及应用[J]. 仪器仪表学报, 2006, 27(1): 20-26. Liu Xinmin, Liu Guanjun, Qiu Jin. The fault diagnosis model and application based on HMM-SVM [J]. Chinese Journal of Scientific Instrument, 2006, 27 (1): 20-26. (in Chinese)
- [4] 张继军,马登武,张金春. 基于 HMM 的电子设备状态监测与健康评估[J]. 系统工程与电子技术, 2013, 35(8): 1692-1695. Zhang Jijun, Ma Dengwu, Zhang Jinchun. State of electronic equipment inspection and health assessment based on HMM [J]. Journal of Systems Engineering and Electronics, 2013, 35 (8): 1692-1695. (in Chinese)
- [5] 曾庆虎. 机械动力传动系统关键部件故障预测技术研究[D]. 长沙:国防科学技术大学, 2010: 65-72.
- [6] 滕红智,赵建民,贾希胜,等. 基于 CHMM 的齿轮箱状态识别研究[J]. 振动与冲击, 2012, 31 (5): 92-96. Teng Hongzhi, Zhao Jianmin, Jia Xisheng, et al. Research on the state identification of gear box based on CHMM [J]. Journal of Vibration and Shock, 2012,

- 31(5):92-96. (in Chinese)
- [7] 张星辉,康建设,高存明,等. 基于 MoG-HMM 的齿轮箱状态识别与剩余使用寿命预测研究[J]. 振动与冲击, 2013, 32(15): 20-25. Zhang Xinghui, Kang Jianshe, Gao Cunming, et al. Research of gear box state identification and residual service life prediction based on MoG-HMM [J]. Journal of Vibration and Shock, 2013, 32(15): 20-25. (in Chinese)
- [8] 周建英,王飞跃,曾大军. 分层 Dirichlet 过程及其应用综述[J]. 自动化学报, 2011, 37(4): 389-407. Zhou Jianying, Wang Feiyue, Zeng Dajun. Review on hierarchical Dirichlet processes and application [J]. Acta Automatica Sinica, 2011, 37(4): 389-407. (in Chinese)
- [9] Xu Tianbing, Zhang Zhongfei, Philip S Y, et al. Dirichlet process based evolutionary clustering [C] // Proceedings of the 8th IEEE International Conference on Data Mining. Pisa, Italy: IEEE, 2008: 648-657.
- [10] 梅素玉,王飞,周水庚. 狄利克雷过程混合模型、扩展模型及应用[J]. 科学通报, 2012, 57(34): 3243-3257. Mei Suyu, Wang Fei, Zhou Shuigeng. Dirichlet processes mixture model, extended model and application [J]. Chinese Science Bulletin, 2012, 57 (34): 3243-3257. (in Chinese)
- [11] Teh Y W, Jordan M I, Beal M J, et al. Hierarchical dirichlet processes[J]. Journal of the American Statistical Association, 2006, 101(476): 1566-1581.
- [12] Antoniak C E. Mixtures of dirichlet processes with applications to bayesian nonparametric problems[J]. The Annals of Statistics, 1974, 2(6): 1152-1174.
- [13] Neal R M. Markov chain sampling methods for Dirichlet process mixture models[J]. Journal of Computational and Graphical Statistics, 2000, 9(2): 249-265.
- [14] Qiu Hai, Lee J, Lin Jing. Wavelet filter-based weak signature detection method and its application on roller bearing prognostics[J]. Journal of Sound and Vibration, 2006, 289(4):1066-1090.
- [15] 王恒,马海波,黄希,等. 基于 Kolmogrov-Smirnov 检验和 LS-SVM 的机械设备故障预测研究[J]. 中南大学学报:自然科学版, 2016, 47(6): 1925-1929. Wang Heng, Ma Haibo, Huang Xi, et al. Research on fault prediction of mechanical equipment based on Kolmogrov-Smirnov and LS-SVM[J]. Journal of Central South University: Nature Science Edition, 2016, 47 (6):1925-1929. (in Chinese)



第一作者简介:王恒,男,1981年11月生,博士、副教授、硕士生导师。主要研究方向为机械故障诊断学及机电测控技术。曾发表《基于 Kolmogrov-Smirnov 检验和 LS-SVM 的机械设备故障预测研究》《中南大学学报:自然科学版》2016年第47卷第6期)等论文。
E-mail: wangheng@ntu.edu.cn

