

城市轨道交通道岔转辙机动作功率参考曲线推荐方法研究

朱存仁¹ 卢瑞冰² 张超凡²

(1. 卡斯柯信号有限公司, 200071, 上海; 2. 中国铁路南宁局集团有限公司电务部, 530011, 南宁 // 工程师)

摘要 介绍了一种根据道岔转辙机历史动作功率曲线, 通过无监督机器学习自动生成转辙机动作功率参考曲线的方法。该方法首先通过提取历史曲线特征, 然后形成自适应分类簇, 最后以簇为单位生成推荐的参考曲线。测试验证表明, 该方法可以降低设备维护人员工作量, 提高参考曲线设置的准确度, 提升转辙机设备的智能化维护。

关键词 城市轨道交通; 道岔转辙机; 参考曲线; 无监督机器学习

中图分类号 U213.6⁺1

DOI: 10.16037/j.1007-869x.2021.09.035

Research on Action Power Reference Curve Recommendation Method of Urban Rail Transit Switch Machine

ZHU Cunren, LU Ruibing, ZHANG Chaofan

Abstract A method to automatically generate the reference curve of switch machine action power by unsupervised machine learning according to the historical action power curve of switch machine is introduced. Firstly, the feature of history curve is extracted, and then clustering is formed by adaptive classification. Finally, the recommended reference curve is generated by clustering. Test results show that this method can reduce the workload of equipment maintenance personnel, improving the accuracy of reference curve setting and the intelligent maintenance of switch machine equipment.

Key words urban rail transit; switch machine; reference curve; unsupervised machine learning

First-author's address CASCO Signal Ltd., 200071, Shanghai, China

目前, 在城市轨道交通车站站点设置转辙机动作参考曲线的主要方式是: 在已有的道岔动作曲线中, 以个人经验选取一条能够反映设备正常工作状态的曲线作为参考曲线。国内研究也尝试过将检修维修结束前转辙机设备的最后一条动作功率曲线作为参考曲线进行自动设置^[1]。但这些方式的随

机性较大, 若设置不当则会影响运维人员对转辙机状态的判断, 从而不能有针对性地对转辙机进行预防性维修。因此, 从智能运维的角度出发, 应追踪每一组转辙机的服役性状并给予推荐的参考曲线, 以此降低运维工作量, 同时也可为转辙机设备的智能运维打好基础。

1 转辙机动作功率参考曲线推荐原理

在实际工程应用场景内, 不同地区、不同型号、不同规格、不同采样频率及不同工况下的转辙机实时监测动作曲线具有较大的差异性^[2], 因此本文拟采用无监督机器学习的方式来避免过多人工调参步骤引起的“水土不服”问题。所谓无监督机器学习, 即不通过人工干预的方式来实现自适应的机器学习过程, 相比于有监督的机器学习方法, 具有更强的适应性, 可以保证应用在使用场景方面的健壮性。转辙机参考曲线推荐原理如图1所示。

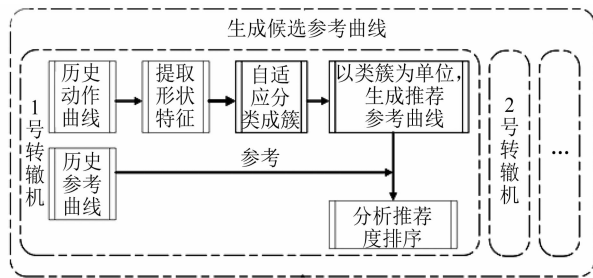


图1 转辙机参考曲线推荐原理

2 转辙机动作功率曲线形状特征提取

根据实际研究结果得知, 在不同工况条件下的转辙机的动作曲线亦可表现出一些细微差异。实际中, 由于滑床板的物理特性差异, 如水平、摩擦等, 其正向到反向与反向到正向的动作曲线在整体形状上具有微小可分的区别。鉴于上述背景, 针对转辙机的动作曲线, 拟采用具备自适应区分特性的

相似度特征来分离不同工作条件下的动作曲线,以作为推荐系统的基础输入。本文对较为常用的2种特征降维模型PCA和t-SNE算法进行了试验比较:t-SNE算法产生的结果优于PCA降维模型,PCA模型的线性方法不利于建模的曲面的流型^[3],t-SNE算法对相似度矩阵降维转化效果较好。

2.1 特征选取与相似度矩阵

令 $C\{C_1, C_2, \dots, C_m\}$ 表示一个转辙机的所有历史动作曲线,其中 m 表示 C 的样本数量。为描述方便,若无特别申明,本文中 C_i 仅表示功率电流曲线。设 P 为采用泊松采样提取 C 中元素作为参考样本 $S_p\{S_1, S_2, \dots, S_k\}$ 的概率, k 为参考样本个数,易知:

$$k = mp \quad (1)$$

综上,根据 C 与 S_p ,可以组成如下的所述的对比矩阵 $C_{m,k}$:

$$C_{m,k} = \begin{bmatrix} (C_1, S_1) & (C_1, S_2) & \cdots & (C_1, S_k) \\ (C_2, S_1) & (C_2, S_2) & \cdots & (C_2, S_k) \\ \vdots & \vdots & \ddots & \vdots \\ (C_m, S_1) & (C_m, S_2) & \cdots & (C_m, S_k) \end{bmatrix} \quad (2)$$

据式(2)可总结出 C 的 k 维相似度矩阵

$$D_{m,k} = \{ \{ d(C_\alpha, S_\beta) \mid \beta \in [1, k] \} \mid \alpha \in [1, m] \} \quad (3)$$

式(3)所示,使用欧氏距离算法 $d(x, y) = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}$ 计算样本间的相似度。 $D_{m,k}$ 的运算量与 m 和 p 的大小有关,其计算复杂度为 $O(m^2)$ 。考虑到转辙机历史动作曲线的记录数目,预计该步骤承担了整体算法的主要计算量。通过调整 p 可以适当控制特征总结的精度。即: p 越大, $D_{m,k}$ 的维度越高,其表征的相似特征精度越高,运算时间越长;不然则反之。

当数据量过大时,可能会产生冗余的相似度矩阵维度,造成运算时间激增。为了避免对计算机资源的过度浪费,有效控制运算时间,本系统将控制 k 变量,即通过式(1)来动态调整 p 值。

2.2 基于 t-SNE 算法的降维分析

根据上一节可知, k 维相似度矩阵 $D_{m,k}$ 可以刻画 C 的内在元素之间的相对关系,由此可以区分单一转辙机在不同工作状态下的动作曲线记录。在实际分析中,由于 k 的数值过大,使用 $D_{m,k}$ 对 C 的相似度进行分析时,必然遭受维度灾难;另一方面,当 k 数值过大时,人工无法复查,不能直观地检验采用

相似度特征进行衡量时其结果的优劣。因此,拟采用特征降维方式来转换 $D_{m,k}$ 为 $F_{m,2}$,记为 C 的2维相似度特征,即:

$$F_{m,2} = R(D_{m,k}) \quad (4)$$

式(4)中, $R(\cdot)$ 表示用过降维算法重构 $D_{m,k}$ 以得到 $F_{m,2}$ 的过程。t-SNE 算法具有最优的降维结果,如图2所示。

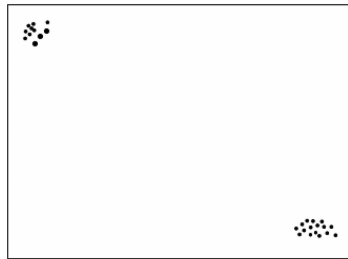


图2 样本(19#转辙机)数据降维结果

图2中每一个点都表示19#转辙机的一条历史动作曲线,通过降维以后的特征值,在二维坐标中显示出来,即可观察到其相对的相似性关系。可以看出,19#转辙机的动作曲线,在曲线形状上有两种形态,这两种形态分别为转辙机定位到反位扳动与反位到定位扳动两种工作模式。

3 自适应分类成簇

本文通过 DBSCAN 算法对相似度特征聚类,根据特征表征的动作曲线形状聚类成簇。在对大量历史数据分析中,发现转辙机的正常动作曲线能够反映出多种类簇,且每一种工况所反映出来的动作曲线都具备接近的相似性。为了实现上述功能,首先需要自动识别这些类簇,进而令系统在无监督模式下学习到这些分类方法。DBSCAN 聚类算法具备较高的聚类准确性,即使在个别转辙机存在多种工况模式的前提下,也能较合理地聚类成簇^[4]。图3为某转辙机的复杂工况自适应学习实例,图中的各个数字代表不同的工况模式。

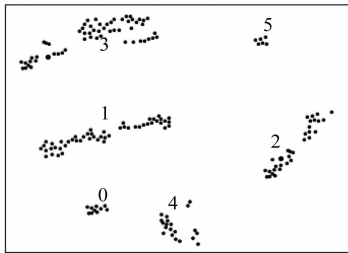


图3 转辙机各工况下数据自适应分类成簇结果

受使用频度和安装位置等因素影响,各个转辙

机之间的工况模式可以看作相互独立的事件。为了下文描述方便,以 t_i 表示任一转辙机通过无监督机器学习的方式总结得到的工况模式,以图3转辙机的动作曲线形状分布为例, $i \in [1, 6]$ 。

4 推荐转辙机动作功率参考曲线拟合方法

对于任意处于服役阶段的转辙机,据上文得知存在若干工况 $\{t_1, t_2, \dots, t_s\}$,对于每组工况 t_i ,都存在对应的动作曲线 $\{C_{1,t_i}, C_{2,t_i}, \dots, C_{\psi,t_i}\}$,即 $\{C_{\gamma,t_i} \mid \gamma \in [1, \psi]\}$,其中 ψ 表示该工况下所包含的历史动作曲线总数。本节主要阐述如何通过 $\{C_{\gamma,t_i} \mid \gamma \in [1, \psi]\}$ 生成对应工况下的候选参考曲线 C_{ref,t_i} 。

1) 统计拟合方法。本文尝试了统计拟合方法和正态拟合方法进行参考曲线的拟合工作。试验发现:统计拟合方法可以快速地反映出 St_i 下曲线形状的主体趋势,可以通过均值和中位数方式快速拟合出具有普遍规律的候选参考曲线 C_{ref,St_i} 。以均值拟合为例,其计算方法如下:

$$C_{\text{ref},t_i} = \left\{ \sum_{\gamma=1}^{\psi} C_{\gamma,t_i}(\alpha) / \psi \mid \alpha \in [1, l] \right\} \quad (5)$$

式(5)中, $C_{\gamma,t_i}(\alpha)$ 表示 t_i 下第 γ 组动作曲线的第 α 个采样值; l 表示 C_{ref,St_i} 的采样点数,该数值与对应的转辙机物理特性有关。

2) 正态拟合方法。统计方法存在一定潜在缺陷,当 St_i 对应的动作曲线集合中存在游离的动作曲线,即因极个别漏报警所导致的故障数据混入 $\{C_{\gamma,t_i} \mid \gamma \in [1, \psi]\}$ 中时,根据式(7)生成的候选曲线会产生一定畸变,需要人工进行调整。

为了避免这种潜在缺陷,可以借助正态分布 $X \sim N(\mu, \sigma^2)$ 来有效削弱游离动作曲线所带来的畸变。其中,参数 μ 为 $\{C_{\gamma,t_i} \mid \gamma \in [1, \psi]\}$ 的平均数,参数 σ 为 $\{C_{\gamma,t_i} \mid \gamma \in [1, \psi]\}$ 的标准差。根据常识可知,可游离动作曲线中的数值偏离 $\{C_{\gamma,t_i} \mid \gamma \in [1, \psi]\}$ 内的主流数值,因此引入 $N(\mu, \sigma^2)$ 的概率密度函数 P :

$$P(x, \mu, \sigma) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (6)$$

式(6)作为式(7)中的加权重,即可使得生成的候选参考曲线尽可能地接近 t_i 下的动作曲线的样本核心。

鉴于此,可以通过如下步骤来生成候选参考曲线:

$$\mu_{\alpha} = \sum_{\alpha=1}^{\psi} C_{\gamma,t_i}(\alpha) / \psi, \alpha \in [1, l] \quad (7)$$

$$\sigma_{\alpha} = \sum_{\alpha=1}^{\psi} [C_{\gamma,t_i}(\alpha) - \mu_{\alpha}]^2 / \psi, \alpha \in [1, l] \quad (8)$$

$$W_{\alpha,i} = P(C_{\gamma,t_i}(\alpha), \mu_{\alpha}, \sigma_{\alpha}) \quad (9)$$

$$\omega_{\alpha,i} = W_{\alpha,i} / \sum_{\alpha=1}^{\psi} W_{\alpha,i} \quad (10)$$

结合式(7)-式(10)及式(5),得到正态拟合方法如下:

$$C_{\text{ref},t_i} = \left\{ \sum_{\gamma=1}^{\psi} \omega_{\alpha,i} C_{\gamma,t_i}(\alpha) \mid \alpha \in [1, l] \right\} \quad (11)$$

与式(5)相比,式(11)具备自适应去噪的特性。图4为使用正态拟合方法得出的转辙机动作功率参考曲线。

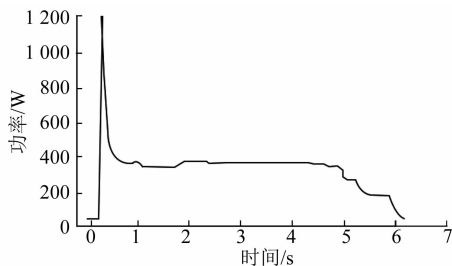


图4 正态拟合方法生成的转辙机动作功率参考曲线

5 结语

针对城市轨道交通中道岔转辙机设备类型多样化、人工设置参考曲线受制于经验知识、数据分段分析方法不具普适性等问题,本文提出了一种基于转辙机动作曲线相似度特征的转辙机工况自适应识别方法和不同工况下转辙机参考曲线推荐方法。该方法通过分析转辙机动作曲线相似度,实现了自适应分类成簇,完成了不同工况下的参考曲线推荐,以最小的人力成本为海量繁杂的转辙机历史动作曲线数据附加合理的数据标签,并能根据每个转辙机的服役性状给予推荐的参考曲线,解决了目前设置参考曲线受制于人工经验知识的问题。该识别方法在21万余条转辙机动作功率曲线样本构成的数据集上测试验证,取得了良好的效果。

参考文献

- [1] 张建军. 浅谈利用微机监测道岔电流曲线指导道岔设备维修[J]. 中国科技纵横, 2015(8): 103.
- [2] 张天赋. S700K 转辙机动作电流曲线智能分析研究[J]. 制造业自动化, 2014(20): 71.
- [3] 董迎朝, 王彬, 马洒洒, 等. 基于 t-SNE 的脑网络状态观测矩阵降维方法研究[J]. 计算机工程与应用, 2018(1): 42.
- [4] 荣秋生, 颜君彪, 郭国强. 基于 DBSCAN 聚类算法的研究与实现[J]. 计算机应用, 2004(4): 45.

(收稿日期: 2020-09-03)