

基于 KPCA 与马氏距离的达林顿管故障预测

刘 强, 程进军, 谭洋波, 郭文浩, 李剑峰

(空军工程大学航空工程学院, 西安, 710038)

摘要 为了对达林顿管进行故障预测,提出了基于 KPCA 与马氏距离的达林顿管故障预测方法。通过对达林顿管进行失效机理分析,设计了加速退化试验,并获取了集电极导通电流与饱和压降性能退化数据,利用小波包分解与核主成分分析进行数据处理,滤除了原始数据中的干扰信号,得到了退化数据的主成分,结合马氏距离对处理后的数据进行特征融合,得到了可以表征达林顿管健康状态变化的健康因子。使用 2 种故障预测算法对健康因子进行预测,故障预测结果验证了文中方法的有效性,预测值与真实值的误差均在 10% 以内。

关键词 故障预测;达林顿管;核主成分分析;马氏距离;健康因子

DOI 10.3969/j.issn.1009-3516.2018.05.013

中图分类号 TP211.5 **文献标志码** A **文章编号** 1009-3516(2018)05-0071-07

A Darlington Transistor Fault Prognostics Method Based on KPCA and Mahalanobis Distance

LIU Qiang, CHENG Jinjun, TAN Yangbo, GUO Wenhao, LI Jianfeng

(Aeronautics Engineering College, Air Force Engineering University, Xi'an 710038, China)

Abstract: In order to predict the failure of Darlington transistor, a method for fault prognostics based on KPCA and Mahalanobis distance is proposed. Through the failure mechanism analysis and accelerated degradation testing of Darlington transistor, the degradation data of collector current and saturation voltage are obtained. The paper utilizes wavelet packet decomposition and KPCA to process the degradation data and filter out interference signals, obtaining the principal component of the degradation. The Mahalanobis distance is used to fuse these components into health index. And the health index could represent the healthy status of Darlington transistor in changes. Finally, two fault predict algorithms are used to predict the HI. And the availability is proved by the forecasting. The results show that RMS between the predicted value and the true value is within 10%.

Key words: fault prognostics; Darlington transistor; kernel principal component analysis; Mahalanobis distance; health index

达林顿管是一种大功率晶体管,使用达林顿管能够实现微处理器对大功率器件的驱动。达林顿管作为整个驱动电路的核心器件,其健康状态对整个电路能否正常工作起决定性的作用。因此,开展对

达林顿管故障预测研究有非常重要的意义。获得能够准确表征研究对象健康状态的特征参数是进行故障预测工作的前提^[1]。由于采集到的原始信号往往表现出非线性的特征,且掺杂噪声等不利因素,导致

收稿日期: 2018-04-03

基金项目: 国家自然科学基金(51779263)

作者简介: 刘 强(1993-),男,安徽阜阳人,硕士生,主要从事故障诊断预测与健康管理工作。E-mail: liubigapple@163.com

引用格式: 刘强,程进军,谭洋波,等.基于 KPCA 与马氏距离的达林顿管故障预测[J].空军工程大学学报(自然科学版),2018,19(5):71-77. LIU Qiang, CHENG Jinjun, TAN Yangbo, et al. A Darlington Transistor Fault Prognostics Method Based on KPCA and Mahalanobis Distance[J]. Journal of Air Force Engineering University (Natural Science Edition), 2018, 19(5): 71-77.

无法通过直接监测数据了解器件真实的健康状态。因此,在进行研究之前,要对采集到的数据进行预处理。一般的数据预处理手段分为特征识别、特征提取、特征融合等。目前,这些手段已经成功地运用于各类器件的故障诊断和预测研究当中。文献[2]用主成分分析法(Principal Component Analysis, PCA)对光伏系统的多个特征量进行特征提取,从而对系统的健康状态进行构建;文献[3]用级联神经网络融合多种数据用于完成机械转子不平衡的故障诊断;文献[4]用瞬时混合函数融合频率和时间数据,对齿轮导线的裂纹进行估计;文献[5]用基于模糊推理的特征融合方法,得到一个故障预测指示器,从而通过趋势分析对故障进行预测。

本文提出了一种基于 KPCA 与马氏距离的达林顿管故障预测方法。将小波包分解与 KPCA 相结合,利用小波包分解识别出了原始数据中所包含的信号成分,然后使用 KPCA 对识别出的数据进行特征提

取,滤除噪声干扰,提取出信号的主成分;针对多维主成分信号无法建立统一标准的情况,提出了使用马氏距离对多维信号进行特征融合的方法,构建健康因子,并利用故障预测算法对健康因子进行预测。

1 达林顿管的失效机理分析

温度、湿度应力和电应力是导致达林顿管失效的主要应力。由这 3 种应力导致的失效模式主要有晶体管正向电流传输比降低、漏电流过大或者开路以及饱和压降减小^[6]。并且每种失效模式存在多种失效机理,具体情况见表 1。

通过对达林顿管失效机理的分析,可知,达林顿管的特征参数退化主要表现为正向电流传输比降低以及饱和压降减小。因此,在综合了失效机理与实际观测点的可实现性后,本文选择对达林顿管的集电极电流 I_C 与饱和压降 I_{CE} 进行数据监测。

表 1 达林顿管主要失效机理

Tab. 1 Main failure mechanism of Darlington transistor

失效模式	与芯片相关的失效机理	与封装相关的失效机理
正向电流传输比降低	②氧化层击穿	①引线键合退化
	②发射结特性退化	②芯片和氧化铍陶瓷粘接界面的退化导致的热阻增大
	③发射极和基层多层金属化电迁移	③芯片裂纹
	④电极接触电阻退化	
漏电流过大或者开路	①氧化层、绝缘层、外延层缺陷	①管壳内含含有大量水汽
	②结击穿或缺陷	②引线键合工艺缺陷,常见的有焊料堆积或爬升至 PN 结表面等
	③介质漏电或缺陷	③芯片粘结工艺缺陷
	④金属电迁移	
饱和压降减小	①芯片与键合处电化学腐蚀	①芯片粘合处退化
	②接触电阻退化	②引线键合处退化

2 达林顿管加速退化试验

本文对驱动电路电路板进行了加速退化试验。本次试验所选取的核心试验器件为意法半导体公司(ST)制造的 TIP122 达林顿管,其封装型号为 TO-220。

文献[7]通过大量的实验和分析表明,高温高湿环境会加速晶体管的老化。通过查阅 TIP122 型达林顿管器件手册,得出其正常工作温度为 $-65^{\circ}\text{C} \sim 150^{\circ}\text{C}$,正常工作湿度为 $20\% \text{RH} \sim 95\% \text{RH}$,最适宜工作温度为 $20^{\circ}\text{C} \sim 40^{\circ}\text{C}$,适宜工作湿度为 $40\% \text{RH} \sim 60\% \text{RH}$ 。加速退化试验应当在避免破坏性试验的前提下进行,故在器件正常工作环境范围内选取试验应力高于其适宜工作环境应力,温度应力为 75°C ,高湿度环境应力为:相对湿度 $90\% \text{RH}$ 。电应力的设置形式为:从时间角度出发,在试验过程中设

置微处理器对达林顿管始终保持高电平输出,使得达林顿管时刻处于导通状态,集电极始终流通大电流。同时,选择与驱动负载电磁阀阻值相同的功率电阻作为负载,保证驱动电路的稳定性。

本试验 GWS 综合应力箱为试验板提供环境温度、湿度,整个试验过程中环境应力条件保持不变;选用 ADLINK-PCI9114 数据采集卡对数据进行采集,上位机基于 LabVIEW 平台对数据进行解析与存储。试验系统流程图见图 1。

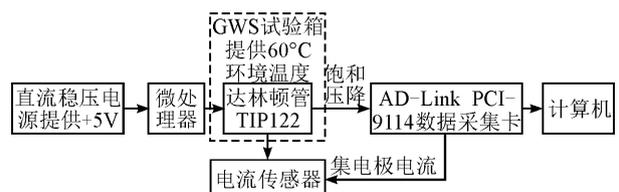


图 1 试验系统流程图

Fig. 1 The flow diagram of ADT

经过试验得到了 12 组样本的集电极电流及饱和压降数据,经对比发现各样本数据均呈现了相似的变化趋势,即集电极导通电流与饱和压降均减小。试验结果与达林顿管失效机理的结论相同。下面以 1 号样本的导通电流 I_C 与饱和压降 I_{CE} 的数据(0~39 000 min)为例,对本文所提出的方法进行阐述和验证。1 号样本的数据情况见图 2、图 3。

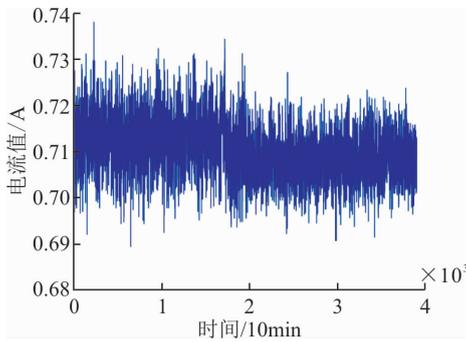


图 2 达林顿管集电极导通电流数据图

Fig. 2 The data image of breakover current

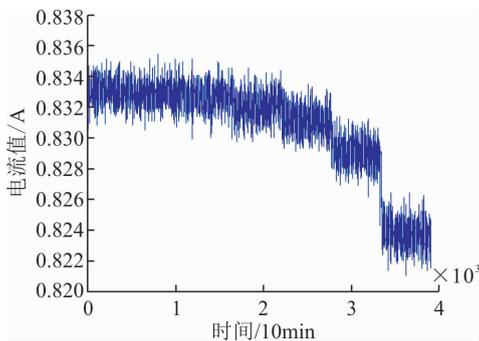


图 3 达林顿管饱和压降数据图

Fig. 3 The data image of saturation voltage drop

从原始数据图中可以看出,在监测的时间范围内,达林顿管集电极导通电流 I_C 与饱和压降 I_{CE} 均发生了一定趋势的变化,集电极导通电流饱和压降均减小,表现出了达林顿管呈现出逐渐老化的现象。若直接使用原始数据作为评估器件健康状态的特征参数,数据所含信息片面,无法反映出器件真实的健康状态;且原始数据含有大量噪声,使用故障预测算法对其进行处理,会增加处理算法的复杂性,降低各类算法的性能,并且会引入巨大的误差。因此,选择对原始数据进行预处理,构建健康因子。

3 故障预测的模型原理

3.1 小波包分解

小波包分解是建立在小波分析基础上的一种对信号更加细致分析和重构的方法,它对信号进行宽

频带划分,对小波分解未涉及的高频部分也进一步分解,提高了信号的时频分辨率^[8]。

给定正交尺度函数 $u_n(t)$ 和小波函数 $u_{2n}(t)$,其满足如下关系^[9]:

$$u_{2n}(t) = \sqrt{2} \sum_{k \in Z} h(k) u_n(2t - k) \quad (1)$$

$$u_{2n+1}(t) = \sqrt{2} \sum_{k \in Z} g(k) u_n(2t - k) \quad (2)$$

式中: $g(k)$ 和 $h(k)$ 分别为低通和高通滤波器系数。以上定义的集合 $\{u_n(t)\}$ 即为正交小波包。

3.2 核主成分分析法

KPCA 是在 PCA 基础上,通过引入核函数,将高维数据的计算转化为原始数据在低维数据的计算,从而大大降低了 PCA 计算高维数据的难度,使得 PCA 能够适用于高维数据空间^[10]。

设原始数据有 m 组输入样本: $X = [x_1, x_2, \dots, x_m]$,每组样本有 N 个变量,定义其为输入空间 R^N 的数据集。定义非线性映射 $\Phi: x_i \rightarrow \Phi(x_i), i = 1, 2, \dots, m$ 。将原始数据从低维输入空间 R^N 通过非线性变换 Φ 映射到高维空间 F 。

映射后的协方差矩阵:

$$C = \frac{1}{m} \sum_{i=1}^m \Phi(x_i) \Phi^T(x_i) \quad (3)$$

C 的特征向量 V 和特征值 λ 满足关系式:

$$CV = \lambda V \quad (4)$$

定义核对称矩阵 $K_{m \times m}$:

$$K(x_i, x_j) = \Phi(x_i) \Phi^T(x_j) = \langle \Phi(x_i), \Phi(x_j) \rangle \quad (5)$$

则可得样本 X 在特征空间中的第 $k(k = 1, 2, \dots, N)$ 个主元 t_k ^[11]:

$$t_k = \langle v^k, \Phi(x) \rangle = \sum_{i=1}^m \alpha_{ik} \langle \Phi(x_i), \Phi(x) \rangle \quad (6)$$

3.3 马氏距离

马氏距离(Mahalanobis Distance, MD)是由印度统计学家 Mahalanobi 提出的一种统计量指标,用来计算数据样本之间的协方差距离^[12]。马氏距离越大,说明样本之间差异性越明显。与另一个能够表征样本相似性的统计量指标——欧式距离相比,马氏距离不受量纲影响,可以体现各个特征参数之间的联系,并排除特征参数之间的相关性干扰^[13]。

样本 Y 到样本集 $X_{m \times n}$ 之间马氏距离的计算公式如下:

$$d_{MD} = \sqrt{(y - \bar{x}) \sum^{-1} (y - \bar{x})^T} \quad (7)$$

式中: $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ 为样本集 X 的重心; $\sum =$

$\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(x_i - \bar{x})^T$ 为 \mathbf{X} 的协方差矩阵。

4 结果分析

4.1 基于小波包分解的特征识别

小波包分解通过一系列低通滤波器和高通滤波器对信号进行滤波,将信号分解成为二叉树的结构形式。本文采用3层分解,分别提取第3层从低频到高频的8个特征信号。具体分解见图4。

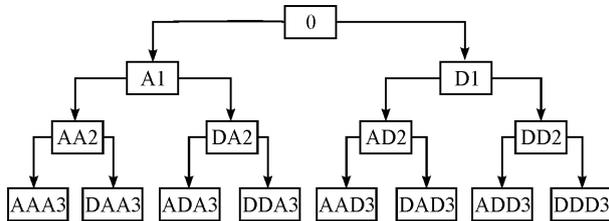


图4 3层小波包分解树

Fig. 4 Three layers structure of wavelet packet

图4中,A表示低频,D表示高频,末尾的数字表示当前小波包分解的层数,具体分解关系为:

$$S = AAA3 + DAA3 + ADA3 + DDA3 + AAD3 + DAD3 + ADD3 + DDD3 \quad (8)$$

本文选择 db10 小波基函数进行3层小波包分解,设定每100 min为一个周期,对周期内的数据进行小波包分解。经过小波包分解得到的达林顿管集电极导通电流与饱和压降数据图见图5、图6。

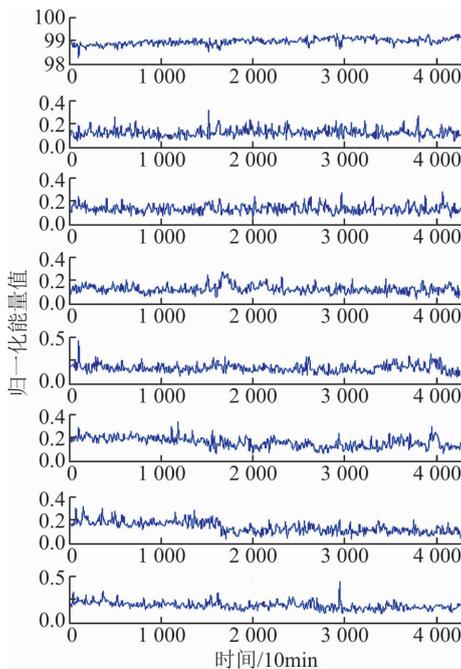


图5 达林顿管集电极导通电流小波包分解

Fig. 5 Wavelet packet decomposition of collector current

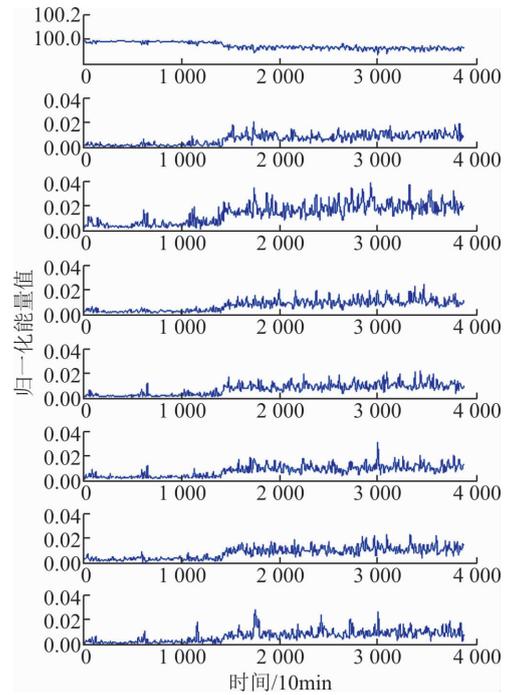


图6 达林顿管饱和压降小波包分解

Fig. 6 Wavelet packet decomposition of saturation voltage drop

图像从上至下依次为 AAA3, DAA3, ..., DDD3 频段内的信号能量值图像。

通过对图像的分析,可以得知,AAA3 频段的信号所含有的能量值远远大于其他各个频段上信号的能量值,说明该频段内所含原始数据的成分较多。从能量趋势上来看,各个频段上的信号能量均随着时间的增加发生了变化,从侧面说明了监测参数性能的退化。

通过小波包分解后数据维度为16。为了提取出这些数据中的主要成分,本文使用 KPCA 对经过特征识别出的数据进行了特征提取,对数据维度进行约减。

4.2 基于核主成分分析的特征提取

基于核主成分分析方法中的核心环节在于对核函数的选取,常用的核函数有多项式核函数、Sigmoid 核函数以及 RBF 核函数^[14]。

本文使用了 PCA 方法、基于多项式核函数的 KPCA(P-KPCA)、基于径向基(RBF)核函数的 KP-CA(RBF-KPCA)对数据进行了特征提取。具体结果见表2。

分析结果可知,KPCA 方法对数据的降维效果明显好于 PCA 方法。PCA 方法前9个主成分累计贡献率小于99%;多项式 KPCA,当 $\sigma=2$ 时前5个主成分累计贡献率小于99%; $\sigma=3,4,5$ 时,前4个主成分累计贡献率小于99%;径向基

KPCA,当 $\sigma=2$ 时前 3 个主成分累计贡献率小于 99%; $\sigma=3$ 时前 4 个主成分累计贡献率小于 99%。通过对比分析,径向基 KPCA($\sigma=2$) 提取出的特征信息最为集中,降维效果最为明显。因

此,本文选取 $\sigma=2$ 的径向基 KPCA 作为特征提取的核函数。选取通过该函数处理使得主成分累计贡献率大于 99% 的前 4 个主成分作为特征提取后的特征参数。

表 2 前 10 个主成分累计贡献率

Tab.2 Accumulative contribution rate of the top ten main components

Method	σ	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10	%
PCA	—	61.08	68.88	75.73	81.24	86.29	90.99	94.96	98.52	98.93	99.24	
	2	78.12	90.14	93.45	97.58	98.76	99.21	99.58	99.79	99.90	99.97	
	3	78.64	90.35	93.80	98.01	99.14	99.35	99.62	99.81	99.94	100.00	
P-KPCA	4	77.27	90.44	94.57	98.55	99.21	99.47	99.70	99.84	99.97	100.00	
	5	76.86	90.45	94.77	95.68	99.30	99.52	99.74	99.88	99.95	100.00	
RBF-KPCA	2	83.34	91.47	95.83	99.75	99.80	99.87	99.92	99.98	100.00	100.00	
	3	81.14	91.52	95.36	98.79	99.45	99.63	99.81	99.01	99.99	100.00	

4.3 基于马氏距离的特征融合

经过 KPCA 特征提取后的数据维数为 4。使用马氏距离完成特征融合,首先要进行健康样本的选取。从原始数据图上并不能直观得到数据的健康样本以及参数变化的确切时间点。因此,对原始数据分组求平均值,以 100 min 为一组,分析参数变化状况。具体情况见图 7、图 8。

为该段时间内器件状态良好,未发生性能退化。故选择特征处理后的前 150 个样本作为健康样本。按照时间顺序,计算当前样本与健康样本的马氏距离作为健康因子。图 9 为通过特征处理后的归一化马氏距离计算后的数据信息。

为了评价本文所提出方法的优劣性,在此列举了对特征提取数据使用欧氏距离方法进行特征融合的结果。具体数据情况见图 10。

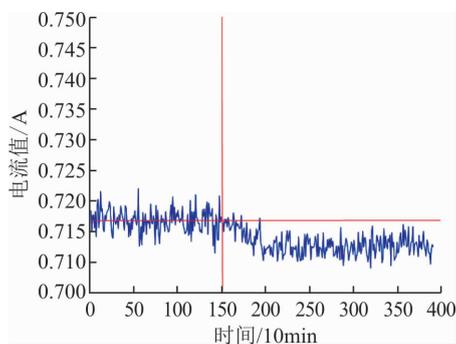


图 7 集电极电流均值

Fig.7 The mean value of breakover current

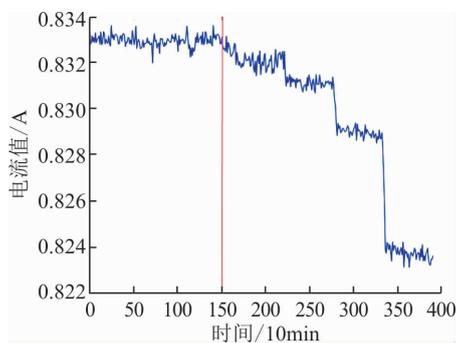


图 8 饱和压降均值

Fig.8 The mean value of saturation voltage drop

结果表明,集电极电流与饱和压降的前 150 个均值基本保持不变,说明达林顿管正向电流 I_C 与饱和压降 V_{CE} 在 0~15 000 min 内基本保持不变,可认

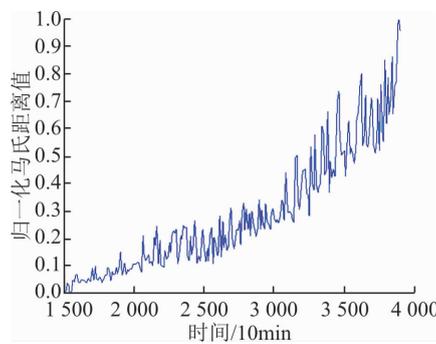


图 9 特征处理得到的马氏距离

Fig.9 Feature processing Mahalanobis distance

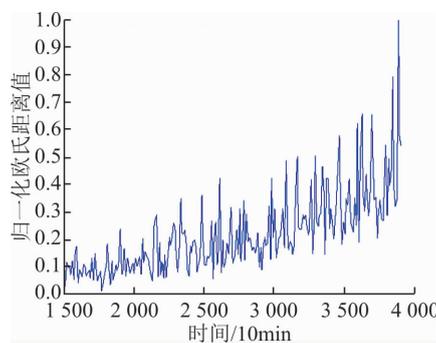


图 10 特征处理得到的欧氏距离

Fig.10 Feature processing Euclidean distance

通过直观对比,可以看出,2 组数据均表现出了一定的趋势性,数值总体随着时间不断增大,当前状

态样本与健康状态样本的差异性随时间不断变大,达林顿管呈现出不断劣化的状态,符合器件真实的寿命变化情况。对比马氏距离与欧氏距离,欧氏距离数据表现出的趋势性较为平缓且波动较大,这是由于KPCA提取出的数据存在数量级的差异,对于不同数量级数据计算欧氏距离时会自动生成相应的权重,引入尺度造成的误差;而马氏距离不受量纲影响,且能够反映数据的微小变化,因此,马氏距离数据的趋势性更加明显且反映样本之间的差异性更加真实。

4.4 健康因子的故障预测

本文采用了自回归滑动平均模型(Autoregressive Moving Average, ARMA)模型和RBF神经网络对以上2类数据进行故障预测。预测后的结果见表3、图11、图12。

表3 预测情况表

Tab. 3 The table of prediction %

健康因子	ARMA(RMS)	RBF-ANN(RMS)
马氏距离	6.58	4.20
欧式距离	10.20	7.64

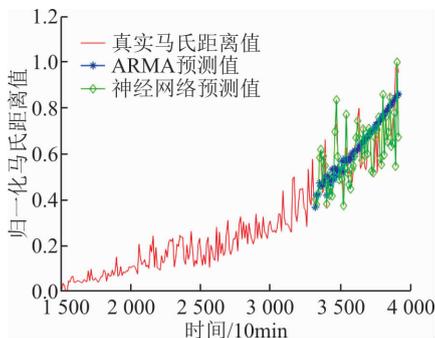


图11 特征处理马氏距离预测情况

Fig. 11 Forecasting results of feature processing Mahalanobis distance

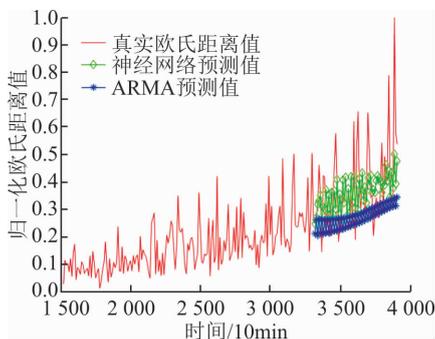


图12 特征处理欧氏距离预测情况

Fig. 12 Forecasting results of feature processing Euclidean distance

ARMA模型的预测方式为:对数据选取前180样本点作为模型的训练样本,对后60个样本点进行

预测。(由于针对不同数据,ARMA模型的参数 p, q 选取不同,为了保证控制变量,本文针对每组数据在AIC准则下均选取了最优参数 p, q 进行预测)。

RBF神经网络预测方式则选用文献[15]中提出的预测模型,即:选取前160个数据作为训练样本;其中,将每10个样本分为1组作为输入数据,组后排序第1位的数据作为输出。

分析图、表可知,对于以上2种健康因子而言,RBF神经网络较之ARMA预测模型拥有更好的预测精度;两种预测算法对本文所提出的健康因子的预测精度明显高于以欧氏距离作为健康因子的方法;从预测具体情况来看,RBF神经网络对数据的动态变化更为敏感,能够捕捉数据的突变情况;而ARMA预测模型则对整体的变化趋势更为敏感。2种方法平均相对误差均小于10%,取得了较好的预测效果。

5 结论

1) 通过对达林顿管的失效机理分析和加速退化试验,得出达林顿管健康状态的劣化可表现为集电极导电电流与饱和压降的退化。

2) 使用KPCA特征提取方法提取出了原始信号中的主要成分,证明了KPCA具有比PCA方法更优异的降维性能。

3) 使用ARMA预测模型及RBF神经网络对本文所提出的健康因子进行了故障预测,两种预测算法的平均相对误差分别为6.8%和4.2%,取得了良好的预测效果。

参考文献(References):

- [1] 曾声奎. 故障预测与健康(PHM)技术的现状[J]. 航空学报, 2005, 26(5): 626-631.
ZENG S K. Status and Perspectives of Prognostics and Health Management Technologies[J]. Acta Aeronautica et Astronautica Sinica, 2005, 26(5): 626-631. (in Chinese)
- [2] 丁坤, 刘振飞, 高列, 等. 基于主成分分析和马氏距离的光伏系统健康状态研究[J]. 可再生能源, 2017(1): 1-7.
DING K, LIU Z F, GAO L, et al. Research on Photovoltaic System Health State Based on PCA-MD method [J]. Renewable Energy Resources, 2017(1): 1-7. (in Chinese)
- [3] LIU Q, WANG H P. A Case Study on Multisensor Data Fusion for Imbalance Diagnosis of Rotating Machinery [M]. Cambridge: Cambridge University

- Press, 2001.
- [4] JARDINE A K S, LIN D, BANJEVIC D. A Review on Machinery Diagnostics and Prognostics Implementing Condition-Based Maintenance[J]. *Mechanical Systems & Signal Processing*, 2006, 20(7): 1483-1510.
- [5] GOEBEL K, BONISSONE P. Prognostic Information Fusion for Constant Load Systems[C]// *International Conference on Information Fusion*. Philadelphia, PA: IEEE, 2006: 1-9.
- [6] 杜磊. 基于大功率晶体管 BUX10 性能退化的可靠性分析研究[D]. 鞍山: 辽宁科技大学, 2012.
- DU L. Research on Reliability Evaluation for BUX10 Based on Performace Degradation Data[D]. Anshan: University of Science and Technology Liaoning, 2012. (in Chinese)
- [7] 杨家铿, 翁寿松. 晶体管非工作期失效率预计[J]. *半导体学报*, 1994, 15(8): 558-564.
- YANG J K, WENG S S. Nonoperating Failure Rate Prediction for Transistors[J]. *Chinese Journal of Semiconductors*, 1994, 15(8): 558-564. (in Chinese)
- [8] 孙健, 王成华, 杜庆波. 基于小波包能量谱和 NPE 的模拟电路故障诊断[J]. *仪器仪表学报*, 2013, 34(9): 2021-2027.
- SUN J, WANG C H, DU Q B. Analog Circuit Fault Diagnosis Based on Wavelet Packet Energy Spectrum and NPE[J]. *Chinese Journal of Scientific Instrument*, 2013, 34(9): 2021-2027. (in Chinese)
- [9] WANG Y, XU G H, LIANG L, et al. Detection of Weak Transient Signals Based on Wavelet Packet Transform and Manifold Learning for Rolling Element Bearing Fault Diagnosis[J]. *Mechanical Systems and Signal Processing*, 2015(54-55): 259-276.
- [10] 周志华. 机器学习[M]. 北京: 清华大学出版社, 2016.
- ZHOU Z H. Machine Learning[M]. Beijing: Tsinghua University Press, 2016. (in Chinese)
- [11] ZHANG Y W. Enhanced Statistical Analysis of Non-linear Processes Using KPCA, KICA and SVM[J]. *Journal of Chemical Engineering Science*, 2009, 64: 801-811.
- [12] YOUNIS K, KARIM M, HARDIE R, et al. Cluster Merging Based on Weighted Mahalanobis Distance with Application in Digital Mammo-Graph[C]// *Aerospace and Electronics Conference*. Dayton, OH: IEEE, 1998: 525-530.
- [13] NIU G, SINGH S, HOLLAND S W, et al. Health Monitoring of Electronic Products Based on Mahalanobis Distance and Weibull Decision Metrics[J]. *Microelectronics Reliability*, 2011, 51: 279-284.
- [14] 孙靖杰, 赵建军, 王汉昌, 等. 基于 FRFT-KPCA 的模拟电路非线性故障特征提取[J]. *电机与控制学报*, 2013, 17(8): 100-106.
- SUN J J, ZHAO J J, WANG H C, et al. Nonlinear Fault Features Extraction for Analog Circuit Based on FRFT-KPCA [J]. *Electric Machines and Control*, 2013, 17(8), 100-106. (in Chinese)
- [15] 李瑞莹, 康锐. 基于神经网络的故障率预测方法[J]. *航空学报*, 2008, 29(2): 357-362.
- LI R Y, KANG R. Failure Rate Forecasting Method Based on Neural Networks[J]. *Acta Aeronautica & Astronautica Sonica*, 2008, 29(2): 357-362. (in Chinese)

(编辑: 徐楠楠)