

1-2-2019

Health Warning and Fault Diagnosis of Pulverizing System Based on Syncretic Similarity

Songming Jiao

Department of Automation, North China Electric Power University, Baoding 071003, China;

Follow this and additional works at: <https://dc-china-simulation.researchcommons.org/journal>



Part of the Artificial Intelligence and Robotics Commons, Computer Engineering Commons, Numerical Analysis and Scientific Computing Commons, Operations Research, Systems Engineering and Industrial Engineering Commons, and the Systems Science Commons

This Paper is brought to you for free and open access by Journal of System Simulation. It has been accepted for inclusion in Journal of System Simulation by an authorized editor of Journal of System Simulation.

Health Warning and Fault Diagnosis of Pulverizing System Based on Syncretic Similarity

Abstract

Abstract: It is an arduous task to know health status of a pulverizing system in power plant by monitoring these parameters simultaneously and the fault diagnosing process is complicated. *An approach based on syncretic similarity is presented which is suitable for industrial system's health warning and fault diagnosis. The syncretic similarity couples a new type of weighted Mahalanobis distance based on principal component analysis with an improved weighted sine similarity. The approach has self-learning ability. Central parameters which are used to compute similarity can be modified along with the operation process. Simulation results show that the method is suitable for online application because of its high accuracy, fast classification, high real-time performance, reliability and simple structure.*

Keywords

fault diagnosis, similarity, pulverizing system, health warning, Mahalanobis distance

Recommended Citation

Jiao Songming. Health Warning and Fault Diagnosis of Pulverizing System Based on Syncretic Similarity[J]. Journal of System Simulation, 2018, 30(2): 595-604.

基于融合相似度的制粉系统健康预警及故障诊断

焦嵩鸣

(华北电力大学自动化系, 河北 保定 071003)

摘要: 反映火电厂制粉系统健康状况的参数众多, 同时监视这些参数难度较大且故障诊断过程过于复杂。为此提出了一种适合工业系统的基于融合相似度的健康预警实现方法和故障诊断算法。此融合相似度由一种新型的基于主成分分析的加权马氏距离和加权正弦相似度融合而成。基于该方法形成的诊断系统无需大规模样本库或记忆矩阵并具有自学习能力, 算法运行中可自行修正计算融合相似度所需要的中心参数。仿真结果表明, 该方法结构简单可靠, 诊断正确率高, 实时性强, 适合在线运用。

关键词: 故障诊断; 相似度; 制粉系统; 健康预警; 马氏距离

中图分类号: TP206+.3; TP277 文献标识码: A 文章编号: 1004-731X (2018) 02-0595-10

DOI: 10.16182/j.issn1004731x.joss.201802028

Health Warning and Fault Diagnosis of Pulverizing System Based on Syncretic Similarity

Jiao Songming

(Department of Automation, North China Electric Power University, Baoding 071003, China)

Abstract: It is an arduous task to know health status of a pulverizing system in power plant by monitoring these parameters simultaneously and the fault diagnosing process is complicated. An approach based on syncretic similarity is presented which is suitable for industrial system's health warning and fault diagnosis. The syncretic similarity couples a new type of weighted mahalanobis distance based on principal component analysis with an improved weighted sine similarity. The approach has self-learning ability. Central parameters which are used to compute similarity can be modified along with the operation process. Simulation results show that the method is suitable for online application because of its high accuracy, fast classification, high real-time performance, reliability and simple structure.

Keywords: fault diagnosis; similarity; pulverizing system; health warning; mahalanobis distance

引言

据中电联统计数据, 至 2014 年底, 火电装机总容量 92 362.61 万千瓦, 占总装机容量的 67.4%, 火力发电在中国仍然是发电主力。制粉系统是整

个火电生产过程中重要的组成部分, 其加工对象原煤中常含各种类型的杂质, 且在研磨煤粉的过程中产生大量粉尘等原因都会导致各种故障的发生。反映制粉系统生产过程状态的参数很多, 电厂的监控系统中对部分重要参数设有报警值, 当这些参数超过报警值时会有声光电形式的提示。这种方法仅能对单一参数越限时给出提示且通常报警时故障已经发生。若能将多个反映设备或系统运行状况的参数向量整合为一个反映健康程度的标量数据, 并将其以曲线的形式显示出来, 即可形成一条设备的系统



收稿日期: 2016-02-26 修回日期: 2016-05-01;
基金项目: 国家重点研发计划(2016YFB0600701), 中央高校基本科研业务费专项资金(2017 MS129);
作者简介: 焦嵩鸣(1972-), 男, 河南杞县, 博士, 副教授, 研究方向为智能算法, 先进控制策略, 复杂系统建模与仿真, 模式识别。

<http://www.china-simulation.com>

• 595 •

的健康指数曲线, 根据曲线的变化形状, 可以在故障还未形成前及时发现其存在异常, 提前解决问题, 起到健康预警的作用。计算健康指数需要事先建立系统正常工况的模型, 将生产数据向量与正常模型向量进行比对计算生成健康指数, 如果健康指数较高, 说明生产数据与正常工况模型相似程度较高, 系统处于健康状态。如果健康指数一直降低说明设备与正常工况差别逐渐增大, 设备有发生故障的趋势, 应及时加强对此设备的监测或提早检修; 健康指数低于设定阈值, 说明设备处于不健康状况, 应启动故障诊断程序发现故障原因。做到有苗头早发现, 发现问题及时处理, 这对提高设备的可靠性是极其有益的。

国内外从事电站制粉系统健康预警及故障诊断的研究较少且无典型的工程应用案例。陈蔚等应用距离函数法对制粉系统故障进行分类^[1], 刘鑫沛等将非线性状态估计技术用于制粉系统故障诊断^[2], 还有不少学者把神经网络, 支持向量机等智能算法应用于制粉系统的故障诊断^[3-5]。纵观这些研究, 主要着眼于故障诊断, 从健康预警角度的研究较少。将反映设备健康程度的多维数据向量转换为一维的健康指数标量可以基于相似性原理来进行。近年来, 国内外不少学者从事相似性原理研究将并其应用于不同种类设备的故障诊断或健康预警。郭鹏应用非线性状态估计对风电设备进行设备状态预警和故障诊断^[6-7], 肖先亮将相似性原理应用于电压暂降方面的研究^[8], 韩涵将改进的马氏距离应用于模拟电路故障诊断^[9], 张胜君还将相似性理论应用于医学诊断^[10]。基于故障诊断的其它方法也有很多, 有国内外学者系统地方法对故障诊断的方法和原理进行了阐述^[11-12]。但是这些方法存在一定的局限性, 如基于智能算法的故障诊断方法因算法中存在随机运算导致一致性和确定性不足; 应用较为广泛的非线性状态估计方法需要大规模的矩阵; 而目前仅基于距离函数进行相似性度量的分类精度有待进一步提高。实际的工业应用需要保证算法的实时性和可靠性, 以上原因限制了研究成果

的应用和推广。

针对以上问题, 本文提出一种具有自学习能力的基于融合相似度的健康指数计算与故障诊断方法。此融合相似度由一种新型的加权马氏距离和新型加权正弦相似度融合而成。健康指数计算时将运行数据向量与正常工况数据中心向量进行相似程度计算, 计算量小; 故障诊断过程是将生产数据向量与各故障工况中心向量进行相似性比对。这些计算过程无需大规模样本库或记忆矩阵, 算法还能够依实时数据自动更新各相应模型的中心参数, 不仅适合于制粉系统健康度计算和故障诊断, 还适合其它实时性要求较高的系统在线健康监视与故障诊断。

1 实验数据说明

本文实验数据来自 600 MW 火电机组的中速磨正压直吹式制粉系统。其常见故障较多, 本文仅选取部分故障进行研究, 其内容如表 1 所示。

表 1 制粉系统部分工况类别

Tab.1 Part of Operation Condition of Pulverizing System

工况序号	工况类型
f_1	正常
f_2	给煤机皮带打滑(故障 1)
f_3	给煤机入口堵煤(故障 2)
f_4	磨煤机一次风管道严重堵塞(故障 3)
f_5	磨煤机润滑油冷却系统故障(故障 4)
f_6	磨煤机内部着火(故障 5)
f_7	磨煤机堵煤(故障 6)

制粉系统故障成因复杂, 故障征兆参数多且难以用数学关系表达它们之间的定量关系, 选择合适的故障参数对于制粉系统故障诊断的研究至关重要^[13-14]。通过对制粉系统各个工况下参数值变化和特性分析并应用主成分分析法剔除对故障诊断贡献率较小的数据^[15], 选择的故障征兆参数表 2。为了描述方便, 本文把反映系统或设备状态的一组数据称为状态向量, 并统一将它描述为行向量。

表 2 制粉系统故障征兆参数

序号	单位	故障征兆参数
P_1	A	磨煤机电流
P_2	°C	磨煤机齿轮箱推力瓦温度
P_3	°C	磨煤机润滑油温度
P_4	°C	磨煤机入口温度
P_5	km ³ /h	磨煤机一次风流量
P_6	kPa	磨煤机入口压力
P_7	kPa	磨煤机进出口差压
P_8	°C	磨煤机出口温度
P_9	°C	磨煤机电动机绕阻温度
P_{10}	°C	磨煤机磨辊轴承温度
P_{11}	kPa	磨煤机密封风与一次风差压
P_{12}	kPa	磨煤机分离器出口压力
P_{13}	A	给煤机电机电流信号
P_{14}	t/h	给煤量
P_{15}	°C	磨煤机送粉温度

2 相似度度量方法及改进

2.1 余弦相似度方法及改进

对于 n 维状态向量 \mathbf{x} , 距总体中心 \mathbf{u} 的余弦相似度定义为:

$$\cos(\mathbf{x}, \mathbf{u}) = \frac{\text{dot}(\mathbf{x}, \mathbf{u})}{\|\mathbf{x}\| \cdot \|\mathbf{u}\|} \quad (1)$$

式(1)中 $\text{dot}(\cdot)$ 表示内积运算, 而 $\|\cdot\|$ 为向量的 2 范数。经上式计算出来的余弦值介于 0 与 1 之间, 且越接近于 1, 从向量角度近似程度来考虑, 表示个体与总体中心越相似。但是这种计算方法没有考虑到各维度上不同量级对相似度的影响, 容易强调大量级数据对于相似度的影响。为了解决此问题, 在计算状态向量到中心的余弦相似度时, 我们可以将其改进为标准化的余弦相似度。

若样本总体的中心为 $\mathbf{u} = (u_1, u_2, \dots, u_n)$, 状态向量 $\mathbf{x} = (x_1, x_2, \dots, x_n)$, 将所有数据都与样本中心进行相除处理, 处理后的标准化中心变为 $\mathbf{u}^* = \mathbf{I}_{1 \times n} = (1, 1, \dots, 1)$, 而标准化处理后的状态向量为 $\mathbf{x}^* = (x_1^*, x_2^*, \dots, x_n^*) = (x_1/u_1, x_2/u_2, \dots, x_n/u_n)$, 标准化的余弦相似度的公式变为:

$$\cos_{norm}(\mathbf{x}, \mathbf{u}) = \cos(\mathbf{x}^*, \mathbf{I}_{1 \times n}) = \frac{\text{dot}(\mathbf{x}^*, \mathbf{I}_{1 \times n})}{\|\mathbf{x}^*\| \cdot \|\mathbf{I}_{1 \times n}\|} \quad (2)$$

我们还可以进一步考虑不同维度上的数据对于余弦相似度计算的贡献的不同, 即计算余弦相似度应将不同维度对设备状态影响的程度考虑在内, 即需要设置权值矩阵 \mathbf{W} 。权值矩阵主对角线上的值可以根据数据样本的主成分分析结果来确定为 $w_i = i / \sqrt{\sum_{k=1}^n k^2}$, i 为根据主成分分析求得各维度数据贡献率由小至大排序序号, 主对角线外的其余元素皆为零, 加权余弦相似度公式为:

$$\cos(\mathbf{x}, \mathbf{u}, \mathbf{W}) = \frac{\text{dot}(\mathbf{x} \cdot \mathbf{W}, \mathbf{u} \cdot \mathbf{W})}{\|\mathbf{x} \cdot \mathbf{W}\| \cdot \|\mathbf{u} \cdot \mathbf{W}\|} \quad (3)$$

将 2 和 3 两因素综合考虑, 不仅顾及到数量级对余弦相似度的影响, 又考虑到对余弦相似度的贡献度, 则可将加权的余弦相似度进一步改进为公式(4), 我们把它命名为改进的加权余弦相似度。

$$\cos_{norm}(\mathbf{x}, \mathbf{u}, \mathbf{W}) = \frac{\text{dot}(\mathbf{x}^* \cdot \mathbf{W}, \mathbf{I}_{1 \times n} \cdot \mathbf{W})}{\|\mathbf{x}^* \cdot \mathbf{W}\| \cdot \|\mathbf{I}_{1 \times n} \cdot \mathbf{W}\|} \quad (4)$$

进行健康指数计算时, 较好的计算方法应能找到一条平行于横坐标的直线将不同类型数据距离正常中心区分开。定义 D_{fault_min} 为故障数据距离正常中心的最小值, D_{normal_max} 为正常数据距离正常中心的最大值, 可以绘制平行于横坐标值为 D_{normal_max} 为区分线。由以上描述, 可定义区分度

$$A_{Dif} = \frac{D_{fault_min} - D_{normal_max}}{D_{normal_max}} \quad (5)$$

显然, A_{Dif} 值越大, 越容易确定适合的阈值将工况区别为正常与故障, 当 A_{Dif} 小于 0 时, 说明无法找出一条曲线完全区分两种类型。此时区分线以上出现的正常数据点对应的相似度越多, 则算法的区分度越差(即将正常工况判断为故障工况的比率)。论文中所有相似度对比图中横坐标为反映系统健康状况数据的向量序号, 纵坐标由此向量计算出的无量纲相似度。图 1 中实线是正常数据跟正常类型中心的相似度, 点划线为磨煤机打滑故障数据与正常中心相似度。

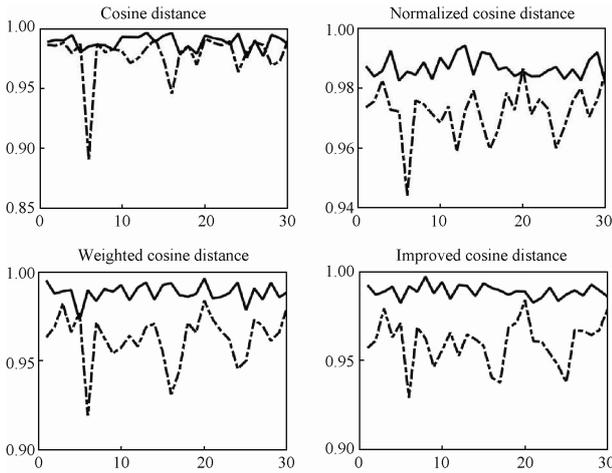


图1 四种不同的余弦相似度对比
Fig.1 Comparison of Four Different Cosine Similarity

观察图1可以发现,采用普通余弦相似度度量的正常数据及故障数据与正常中心相似度区分度最差,采用标准化的方法以及加权方法虽然有改善,但仍然有无法彻底分离的现象,采用改进的加权余弦相似度区别度最好,本文将改进后的加权余弦相似度融入健康指数计算和故障诊断。但是若 n 维向量 x 放大或缩小任意倍数,无论采用哪种形式的余弦相似度,其值均不发生变化,因此仅根据从向量的角度相似性考虑的余弦相似度计算系统健康指数是不够的。应在余弦相似度计算的基础上,进一步考虑数据从距离角度度量的相似程度。

2.2 基于距离的相似度度量方法及改进

2.2.1 几种不同的距离相似度

基于距离的相似度计算方法很多,主要有欧氏距离、标准化的欧氏距离,马氏距离等^[16]。

总体 G 的中心 u 与 n 维状态向量 x 之间的欧氏距离为:

$$D_{euclid}(\mathbf{x}, \mathbf{u}) = \sqrt{\sum_{i=1}^n (u_i - x_i)^2} \quad (6)$$

若总体的均方差为 s ,则标准化的欧氏距离为:

$$D_{nor_euclid}(\mathbf{u}, \mathbf{x}, \mathbf{s}) = \sqrt{\sum_{i=1}^n \frac{(u_i - x_i)^2}{s_i}} \quad (7)$$

标准化的欧氏距离虽在一定程度上可消除不同维度数据量级对相似度计算的影响,但无法消除

状态向量不同维度之间的影响,但是马氏距离同时可以消除以上两种影响。若总体协方差矩阵为 Σ ,则马氏距离定义为:

$$D_{Mahal}(\mathbf{x}, G) = \sqrt{(\mathbf{x} - \boldsymbol{\mu}) \cdot \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})^T}$$

正常工况与给煤机打滑两种工况的监测数据差距较小,仿真结果表明,无论采用何种度量方法,两者的相似度量值均很接近,多数情形下无法正确区别两种类型。作为算法改进对比,以后仅对比这两种工况的相似度。

图2中采用标准化欧氏距离和马氏距离度量两种工况的距离两种方法计算的 A_{Dif} 分别为0.55和1.69。两种方法的区分度较欧氏距离度量方法的区分度-0.05有明显所改善,能够设定甄别正常工况和故障工况的阈值,但仍然存在改进空间。

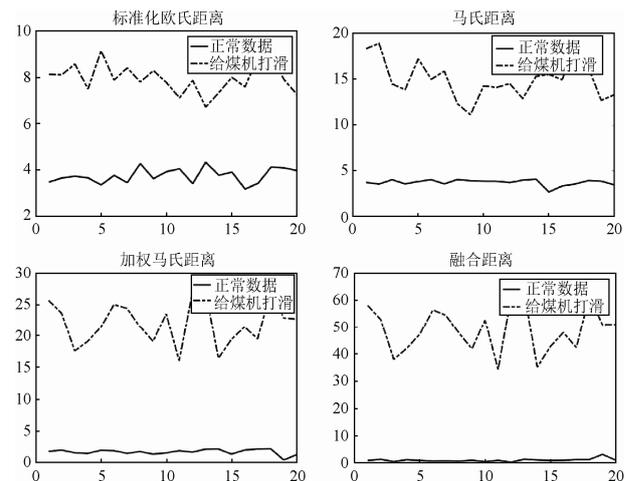


图2 正常与故障工况相似度对比
Fig.2 Comparison of Normal and Faulted Operating Condition' Similarity

2.2.2 改进的加权马氏距离

查阅有关文献可知,马氏距离过度放大了小数量数据所起作用,且对于不同维度向量起到的重要作用未能加入考虑^[17]。本文提出了一种改进的加权马氏距离,对于不同的维度,根据其值量级大小以及其主成分分析对相似度的贡献来对相似度重新进行计算。

权值矩阵的主对角线元素量值的制订规则为:如果其主成分分析的贡献率从小至大排序,式中 i 为贡献率排序序号。

$$wpca_i = i^2 / \sum_{k=1}^n k^2 \quad (8)$$

仿真表明, 采用 8 确定的权值相似度的区分度要好过直接采用 PCA 贡献率系数构成的权值矩阵。采用马氏距离计算相似度时, 数量级微小的分量在计算相似度时有被夸大因素, 因此, 在制定加权矩阵时, 应考虑这一因素。如果样本中心各分量按绝对值由小至大排列序号 1, 2, ..., 得到的权值矩阵的主对角线元素可 9 求取。

$$wval_i = i / \sqrt{\sum_{k=1}^n k^2} \quad (9)$$

综合考虑(8)因素和(9)的因素, 定义

$$\mathbf{W} = \begin{bmatrix} w_1 & 0 & \cdots & 0 \\ 0 & w_2 & \cdots & 0 \\ \vdots & \vdots & w_i & \vdots \\ 0 & 0 & \cdots & w_n \end{bmatrix} \quad (10)$$

公式(10)中 $w_i = wpca_i \cdot wval_i$, 则加权的马氏距离计算公式为:

$$D_{WMahal}(\mathbf{x}, \mathbf{u}, \mathbf{W}) = \sqrt{(\mathbf{x} - \mathbf{u}) \cdot \mathbf{W} \cdot \Sigma^{-1} (\mathbf{x} - \mathbf{u})^T} \quad (11)$$

图 2 中采用(10)式的改进加权的马氏距离算法计算的常数据与故障数据距正常工况中心对应 A_{Dif} 为 6.2, 比马氏距离及仅采用(8)式的加权马氏距离区分度都大大提高, 可将其应用于健康度指数运算。大量实验及仿真表明, 随着观测系统健康指数数据量增大, 各种度量方法的区分度都会下降, 如果仅仅根据加权的马氏距离对数据进行区分, 仍然可能会出现误分类。在对工况区分程度要求较高的故障诊断过程中应考虑将样本点距离某种工况中心的加权余弦相似度和加权马氏距离作为特征数据, 将这两个特性进行融合。

2.3 融合相似度

同时考虑向量距离以及向量夹角两个因素, 显然比仅考虑两者之一的判断结果更为准确和可靠。故障诊断方法中进行相似度计算时主要考虑了基于距离的度量, 这里将距离度量和角度度量下计算出

来的相似度量值合成为一个二维向量。对于加权的马氏距离而言, 完全重合的向量度量距离为 0, 且距离越小说明待测的个体样本与中心相似度越高。由(4)可知, 对于加权的余弦相似度, 两向量方向角越是接近重合, 即其相似程度越大, 其值越接近 1。我们衡量两个向量相似程度原则应统一为度量值越小则相似程度越高, 为此定义加权正弦相似度:

$$\sin(\mathbf{x}, \mathbf{u}, \mathbf{W}) = \sqrt{1 - \cos^2(\mathbf{x}, \mathbf{u}, \mathbf{W})} \quad (12)$$

显然, 若两个向量夹角完全重合, 则对应加权正弦相似度为 0, 完全正交时加权正弦相似度为 1。现将实时向量数据的加权正弦相似度与改进的加权马氏距离定义为二维的相似度量向量, 即 $\mathbf{H}(\mathbf{x}) = [\sin(\mathbf{x}, \mathbf{u}, \mathbf{W}) \quad D_{WMahal}(\mathbf{x}, \mathbf{u}, \mathbf{W})]$, 将其简记为 $\mathbf{H} = [\sin(\mathbf{x}) \quad D_{WMahal}(\mathbf{x})]$ 。

考虑到两个维度的相似度量数据有较强的相关性且其数据的量级不同, 在进行融合相似度计算时应采用改进的加权马氏距离来计算, 即

$$H_{WMahal}(\mathbf{H}, \mathbf{W}_H) = \sqrt{(\mathbf{H} - \bar{\mathbf{H}}) \cdot \mathbf{W}_H \cdot \Sigma_H^{-1} (\mathbf{H} - \bar{\mathbf{H}})^T} \quad (13)$$

式中 \mathbf{W}_H 为复合的基于 $\mathbf{H}(\mathbf{x})$ 向量组的权值矩阵。显然, 这样构建的向量第一维参数加权正弦相似度期望值可为 0, 第二维参数期望值也可为零, 可令 $\bar{\mathbf{H}} = [0 \quad 0]$ 。融合相似度计算过程为:

步骤 1, 将采集得到运行数据向量 \mathbf{x} 与相应工况中心向量 \mathbf{u} 由(4)计算出改进的余弦相似度并进一步依(12)转换为对应的正弦相似度 $\sin(\mathbf{x})$;

步骤 2, 由 11 计算加权马氏距离 $D_{WMahal}(\mathbf{x})$;

步骤 3, 依据 13 计算出融合相似度。

表 3 列出几种距离函数在两种类型数据各 20 组, 50 组, 100 组, 200 组, 500 组时 A_{Dif} 的变化。

从表 3 可以看出, 改进的马氏距离以及融合相似度区分度最好, 各种方法随着数据量增大, 区分度逐渐变小, 改进加权马氏距离与融合相似度的区分度差距变小。

表3 随数据量增长不同度量方法 A_{Dif} 变化

度量方法	20	50	100	200	500
欧氏距离	-0.05	-0.32	-0.33	-0.39	-0.38
标准化欧氏距离	0.55	0.21	0.08	0.005	-0.002
马氏距离	1.69	0.31	0.09	-0.01	-0.01
改进加权马氏距离	6.2	1.28	1.03	0.73	0.51
融合相似度	10.15	2.81	2.62	1.62	0.55

2.4 诊断模型的中心参数自学习

在进行相似性运算时,无论是采用加权的正弦相似度运算还是进行加权的马氏距离运算,都需要与总体中心 \mathbf{u} 比对,计算改进的加权马氏距离还需要用到总体的协方差矩阵 Σ ,我们将这两个参数统称为总体模型的中心参数。假设需要诊断的系统的样本维度为 n ,在系统运行的初始时期,数据中心的确定有两种,一种是搜集一段时间的运行数据进行均值和协方差矩阵运算确定,也可以基于运行规程中规定的标准工况数据作为数据中心,根据经验由均值叠加白噪声的多于 n 组模拟向量得到初始协方差矩阵。但是随着新的运行数据进入系统,应当根据健康指数计算结果或诊断结果不断对不同工况的中心位置和协方差矩阵进行更新,若依据历史数据直接计算模型中心参数,随着历史数据规模不断增加,计算过程过于复杂,这显然不适合实时运算。为此,笔者推导了更新均值和协方差矩阵采用的迭代公式。

若参与运行的数据向量数为 k ,由均值定义 $\bar{u}(k) = \sum_{i=1}^k u_i / k$ 可知 $\sum_{i=1}^k u_i = k \cdot \bar{u}(k)$, 所以有:

$$\bar{u}(k+1) = \frac{\sum_{i=1}^k u_i + u_{k+1}}{k+1} = \frac{k}{k+1} \bar{u}(k) + \frac{u_{k+1}}{k+1} \quad (14)$$

由协方差定义,当其样本个数为 k 时,有

$$\text{cov}(\mathbf{X}, k) = \frac{\sum_{i=1}^k (\mathbf{x}_i - \bar{\mathbf{x}}(k))^T (\mathbf{x}_i - \bar{\mathbf{x}}(k))}{k-1} \quad (15)$$

而当样本个数增加至 $k+1$ 时,有

$$\text{cov}(\mathbf{X}, k+1) = \frac{\sum_{i=1}^{k+1} (\mathbf{x}_i - \bar{\mathbf{x}}(k+1))^T (\mathbf{x}_i - \bar{\mathbf{x}}(k+1))}{k}$$

上式可分解为 $(\mathbf{x}_{k+1} - \bar{\mathbf{x}}(k+1))^T (\mathbf{x}_{k+1} - \bar{\mathbf{x}}(k+1)) / k$ 和

$$\sum_{i=1}^k (\mathbf{x}_i - \bar{\mathbf{x}}(k+1))^T (\mathbf{x}_i - \bar{\mathbf{x}}(k+1)) / k。$$

新增数据与原有数据满足同一分布,故数据与原均值不会偏离太大,再加上求取均值的数据数目庞大,由(14)式可知, $\bar{\mathbf{x}}(k+1) \approx \bar{\mathbf{x}}(k)$, 所以(15)式约等于:

$$\begin{aligned} & \frac{\sum_{i=1}^k (\mathbf{x}_i - \bar{\mathbf{x}}(k))^T (\mathbf{x}_i - \bar{\mathbf{x}}(k))}{k} + \frac{(\mathbf{x}_{k+1} - \bar{\mathbf{x}}(k+1))^T (\mathbf{x}_{k+1} - \bar{\mathbf{x}}(k+1))}{k} \\ &= \frac{k-1}{k} \cdot \text{cov}(\mathbf{X}, k) + \frac{(\mathbf{x}_{k+1} - \bar{\mathbf{x}}(k+1))^T (\mathbf{x}_{k+1} - \bar{\mathbf{x}}(k+1))}{k} \end{aligned}$$

$$\text{即有: } \text{cov}(\mathbf{X}, k+1) = \frac{k-1}{k} \text{cov}(\mathbf{X}, k) + \frac{(\mathbf{x}_{k+1} - \bar{\mathbf{x}}(k+1))^T (\mathbf{x}_{k+1} - \bar{\mathbf{x}}(k+1))}{k} \quad (16)$$

在实际的应用中,应将运行规程中规定的中心值偏离最大值设定为均值参数变化上限。

3 健康预警及故障诊断仿真

3.1 系统结构设计

本文提出的系统包括健康预警及故障诊断两大功能,其处理流程及关系如图3。

工业生产中,大多数的系统和设备处于正常工况,无需进行耗费大量计算机资源的故障诊断。仅需要进行各设备的健康度指数计算以实现健康预警功能。健康指数计算时可采用式(13)中的融合相似度,但因其计算过程较为复杂,为了保证实时性,可在进行健康指数计算时,仅采用改进的加权马氏距离进行度量。当系统监测到健康度持续降低直至低于给定阈值时,启动诊断部件进行诊断,并依据诊断结果对对应类型的模型中心参数进行修正,完成上述功能后故障诊断功能关闭。

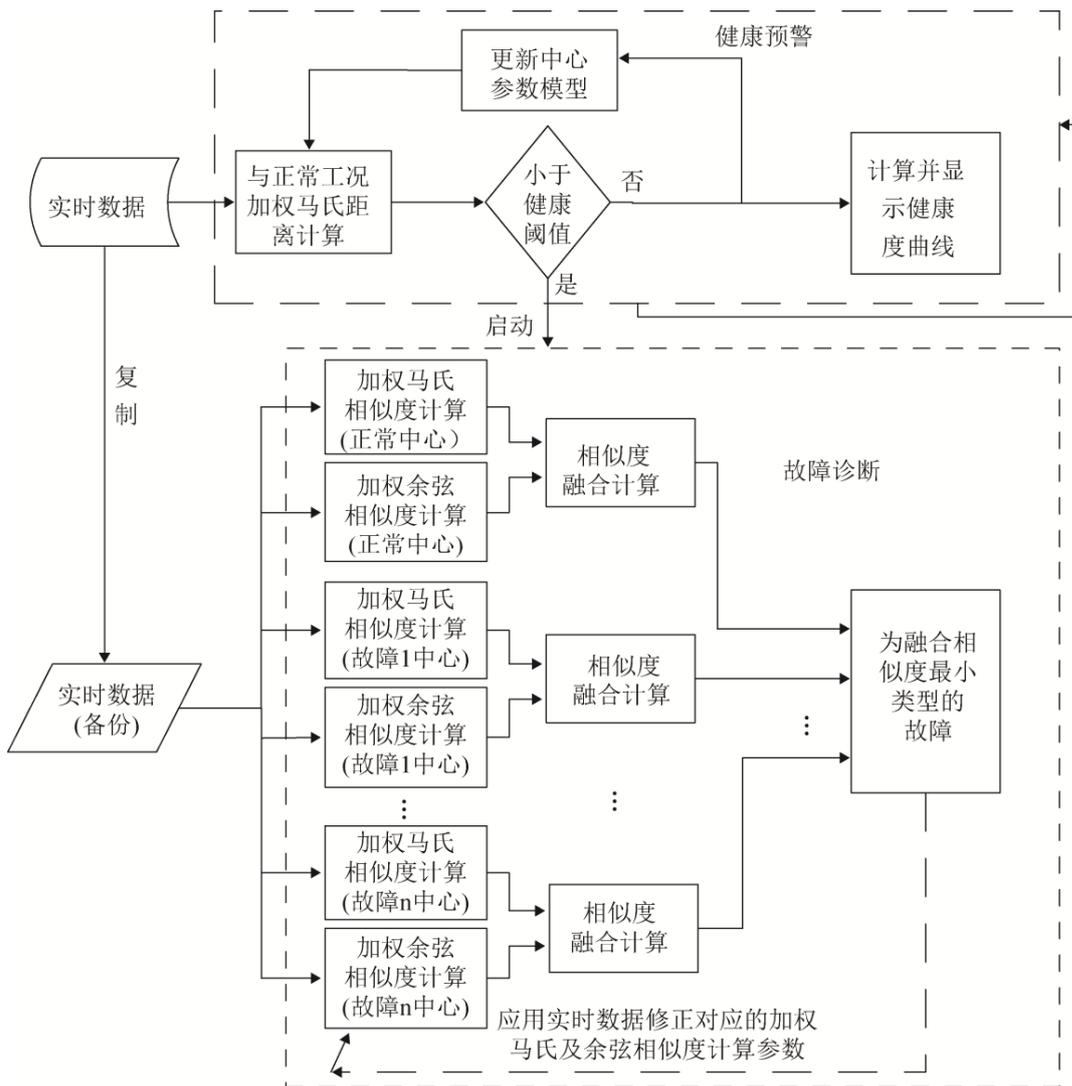


图 3 健康指数计算及故障诊断系统示意图

Fig.3 Schematic Diagram of Health Index Computation and Fault Diagnosis System

3.2 制粉系统健康指数计算

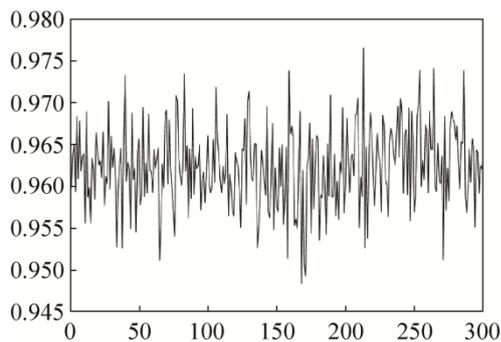
健康预警功能依据健康指数计算来进行, 在计算系统或设备的健康指数时, 应将反映与正常工况中心接近程度的改进加权马氏距离映射到(0, 100%)之间的值。根据监测习惯, 若设备的状态向量与正常工况模型中心重合, 健康程度 100%, 距离变远则说明健康程度下降, 根据这种需要, 为此, 可定义设备的健康指数为:

$$E_{health} = 0.9 \frac{k \cdot D_{WMahala}(x, \Sigma, u)}{D_b} \times 100\% \quad (17)$$

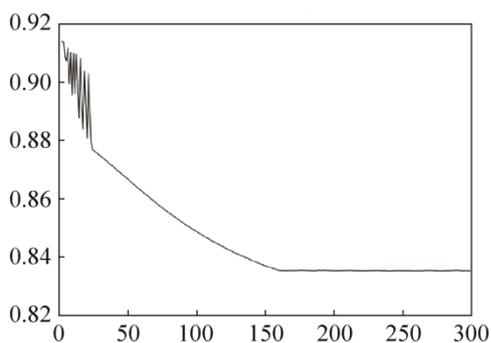
式中: D_b 为系统健康度阈值, 可根据系统的运行规程允许的 normal 值偏离上限计算确定, 系数 k 可调节

边界值对应的期望分值。当设备低于事先设置的阈值且健康指数持续下滑时, 说明设备有出现故障的可能或已出现故障, 可启动故障诊断程序进行诊断。

图 4 为从现场采集两种工况进行健康指数计算过程仿真。两种工况分别为正常工况以及从正常工况发展到磨煤机堵煤工况。图 4(a)为正常工况, 其健康指数一直位于阈值对应的健康指数 90% 以上。图 4(b)中的磨煤机在开始即有健康指数持续下降的趋势, 说明故障在逐渐形成, 在 20 s 左右超过健康指数阈值并迅速下降, 故障程度在扩大, 在 150 s 健康指数变化趋缓。说明设备故障状况趋于稳定。系统应于 10 s 左右启动故障诊断程序进行故障诊断。



(a) 正常运行的系统健康指数曲线



(b) 系统出现故障时的健康指数曲线

图4 两种工况健康度曲线

Fig. 4 Health Index Curves of Two Operation Conditions

3.3 故障诊断过程仿真

故障诊断仅在健康指数持续下降并越过事先设定的健康指数阈值时进行。系统会将生产数据形成的状态向量与故障模型库中的各种故障模型进行基于融合相似度的比对运算，与哪种故障模型生成的融合相似度越小，则说明此时系统为其所对应的故障类型。为了检验本文提出的故障诊断的有效性，作者选取在100%负荷时对应7种不同工况的各100组生产数据的状态向量来进行故障类型分类。为

了对比与其它故障诊断方法的效果，作者还将这些生产数据用其它的诊断方法进行分类测试。

方案一：应用BP神经网络进行分类。神经网络结构为15-20-7结构，输入层为数目对应状态向量各分量，输出对应七种工况，隐层作用函数为双曲正切函数，输出层作用函数为线性函数，训练方法采用L-M法；事先取各工况500组数据进行训练，完成训练后将各工况的100组数据进行分类测试，某输出层神经元输出值最高则故障为其所对应工况。

方案二：采用最小二乘支持向量机(LSSVM)进行分类。首先对训练使用的数据进行预处理且作归一化处理，处理后的数据应用支持向量机进行训练建模，仿真表明，对支持向量机核参数和惩罚因子先采用PSO算法进行优化能更有效的分类，因此这里采用的方法为PSO-LSSVM方法。

方案三：采用文献[6]提出的非线性状态估计方法，其具体实现过程请参阅相关文献，在此不在赘述。

方案四：基于融合相似度最小值原则进行分类。首先应用各工况的数据计算出其对应的模型中心参数，即各自对应的总体中心 \mathbf{u} 和协方差矩阵 $\mathbf{\Sigma}$ ，记录产生模型中心参数使用的数据个数，以便以后模型中心参数自修正计算。计算验证数据与各种工况中心的融合相似度，数据与第几种工况中心融合相似度最小，则判定当前设备状况为其对应工况。

表4为测试的综合结果，表中数据为各工况对应的100组检验向量中的正确分类数目。

表4 几种不同诊断方法对比结果

Tab.4 Results Comparison of Different Diagnosis Methods

分类法	正常	故障1	故障2	故障3	故障4	故障5	故障6	总正确率(%)
BP	97	52	94	100	100	100	100	91.86
LSSVM-PSO	96	96	100	100	100	100	100	98.86
NSET	95	86	100	100	100	100	99	97.14
融合相似	95	98	100	100	100	100	100	99.00
融合相似(15%噪声)	99	86	100	100	100	100	100	97.86
融合相似(30%噪声)	88	91	96	100	100	100	95	95.71

四种方案中前三种方案的数据需要先进行归一化处理, 方案四无需处理。对比四种方案, 方案一效果最差且训练过程较为复杂; 方案二与方案三效果相当, 但方案二训练过程复杂; 方案三分类效果不如方案二和方案四, 且其记忆矩阵规模较大, 若减小记忆矩阵规模分类效果也会随之降低; 方案四初始的模型中心参数选择仅需要少量数据(通常可选取比向量维度多 10), 且本方案具有中心参数自学习功能, 其计算量远远小于前三种方案。为了测试融合相似度最小值原则诊断方法的推广能力, 除了应用采集的原始数据测试, 还在原始数据上叠加 15% 以及 30% 具有白噪声特性的随机干扰值重复以上实验, 实验结果表明该方法具有很强的适应性。

当完成诊断过程后, 还应按式(14)和式(16)自动更新对应故障类型的模型中心参数, 提高后续诊断正确率。

4 结论

本文提出了一种新型的加权马氏距离计算方法和一种加权的正弦相似度计算方法, 并进一步将两个相似度合并为一个融合相似度。为保证实时性, 所设计的系统可仅依据改进的加权马氏距离对系统设备的健康指数进行在线计算并实时更新正常模型中心参数, 当系统健康指数超过阈值后, 启动基于融合相似度的故障诊断程序, 并依据诊断结果对相应的故障模型参数进行更新。基于电站制粉系统生成数据的仿真结果表明, 此系统算法过程简单, 实时性强, 对于故障诊断准确率高, 系统适合于工业生产的设备及系统进行健康状况监测及故障诊断。

参考文献:

- [1] 陈蔚, 贾民平. 距离函数分类法在制粉系统故障诊断中的应用[J]. 振动、测试与诊断, 2009, 29(3): 282-286.
CHEN Wei, JIA Minping. Application of Distance Function to Fault Diagnosis of Coal Pulverizing System [J]. Journal of Vibration, Measurement & Diagnosis, 2009, 29(3): 282-286.
- [2] 刘鑫沛, 翟永杰, 张君颖. 基于聚类分析和状态估计

- 的制粉系统故障预警[J]. 计算机仿真, 2013, 30(8):151-154.
- LIU Xinpei, ZHAI Yongjie, ZHANG Junying. Fault Warning of Pulverizing System Based on Cluster Analysis and State Estimation[J]. Computer Simulation, 2013, 30(8):151-154
- [3] 赵刚, 于向军, 吕震中. 运用 RBF 神经网络的制粉系统故障诊断[J]. 发电设备, 2006, 20(6): 449-453.
ZHAO Gang, YU Xiangjun, Lv Zhenzhong. Diagnosis with RBF Neural Networks of Faults Occurring in Pulverizing Systems[J]. Power Equipment, 2006, 20(6): 449-453.
- [4] 江若玫, 龚春琼. 基于极化因子神经网络的火电厂制粉系统故障诊断技术[J]. 计算机测量与控制, 2015, 23(5):1476-1478.
Jiang Ruomei, Gong Chunqiong. Neural Network with Polarization Factor for Pulverizing System Fault Diagnosis[J]. Computer Measurement & Control. 2015, 23(5): 1476-1478.
- [5] 费树岷, 李延红, 柴琳. 基于 RSPNN 的制粉系统故障诊断[J]. 控制工程, 2012, 19(3): 412-415.
FEI Shumin, LI Yanhong, CHAI Lin. Fault Diagnosis Based on RSPNN for Pulverizing Systems[J]. Control Engineering of China. 2012, 19(3):412-415
- [6] Guo P, Infield D. Wind turbine tower vibration modeling and monitoring by the nonlinear State Estimation Technique (NSET)[J]. Energies (S1996-1073), 2012, 5(12): 5279-5293.
- [7] 郭鹏, Infield David, 杨锡运. 风电机组齿轮箱温度趋势状态监测及分析方法[J]. 中国电机工程学报, 2011, 31(32): 129-136.
GUO Peng, DAVID Infield, YANG Xiyun. Wind Turbine Gearbox Condition Monitoring Using Temperature Trend Analysis[J]. Proceedings of the CSEE, 2011, 31(32): 129-136.
- [8] 肖先勇, 崔灿, 汪洋, 等. 电压暂降分类特征可比性, 相关性及马氏距离分类法[J]. 中国电机工程学报, 2015, 35(6):1299-1305.
XIAO Xianyong, CUI Can, WANG Yang, et al. Comparability and Relevance of Voltage Sag Classifying Features and Classification Methods Based on Mahalanobis Distance[J]. Proceedings of the CSEE, 2015, 35(6):1299-1305.
- [9] 韩涵, 王厚军, 龙兵, 等. 基于改进马氏距离的模拟电路故障诊断方法[J]. 控制与决策, 2013, 28(11): 1713-1722.
HAN Han, WANG Houjun, LONG Bing, et al. Method for Analog Circuit Fault Diagnosis Based on Improved

