

从故障实例数据库中挖掘振动信号特征

关惠玲¹ 张优云¹ 韩捷² 董辛旻² 郝伟²

(¹ 西安交通大学润滑理论与轴承研究所 西安,710049)

(² 郑州大学振动工程研究所 郑州,450002)

摘要 复杂、并发故障诊断的难点在于这些故障的振动信号很复杂,特征很难获取。文中阐述了从实例数据库中挖掘故障特征的整体结构,定义了信号的绝对、相对和梯度特征及相应的绝对、相对和梯度模式实例,进而阐述了应用模糊聚类分析挖掘特征模式的方法。最后以往复式压缩机实例挖掘系统为例说明了该原理的应用。

关键词:故障诊断;数据挖掘;特征提取;知识获取;实例学习

中图分类号:TH165+.3;TH17

目前,复杂、并发、非线性等疑难故障是工程诊断难题。这些故障信号很复杂,特征参数很难获得。由于机理复杂,建模分析时必然忽略掉许多因素。根据一个实例或一个实验信号也很难寻找到特征,因为一个具体对象太具有个性化,很多情况下,本机特有结构因素也混杂在信号中。如何从复杂信号中提取故障特征已被国内外学者所关注。

数据挖掘是近几年在世界范围内发展起来的从数据库中获取有价值信息的技术。应用数据挖掘技术获取疑难故障信号特征是一条值得探索的途径。现有文献,大多数在探讨一般的数据挖掘技术^[1],有一些文献开始探讨数据挖掘技术在工程中的应用^[2,3]。与工程诊断相关的文献大多是探讨一些模糊逻辑、基因算法和模糊聚类分析在已知特征前提下如何去诊断故障的方法^[4~6],探讨数据挖掘技术直

接获取特征的文献较少。文[7]对数据挖掘的快速原型处理模型进行了探讨,应用粗糙集模型对回转机械混合故障进行特征提取,研究了特征波形挖掘算法,应用二维矩和遗传编程实现了二维图象特征指标的挖掘。

本文应用模式描述信号特征,应用模糊聚类分析进行模式识别挖掘信号特征模式,进而探讨从复杂故障的实例库中挖掘特征模式的方法,解决复杂、并发故障的诊断问题。

1 从故障实例数据库中挖掘振动信号特征原理

从故障实例数据库中挖掘振动信号特征原理结构如图1所示。

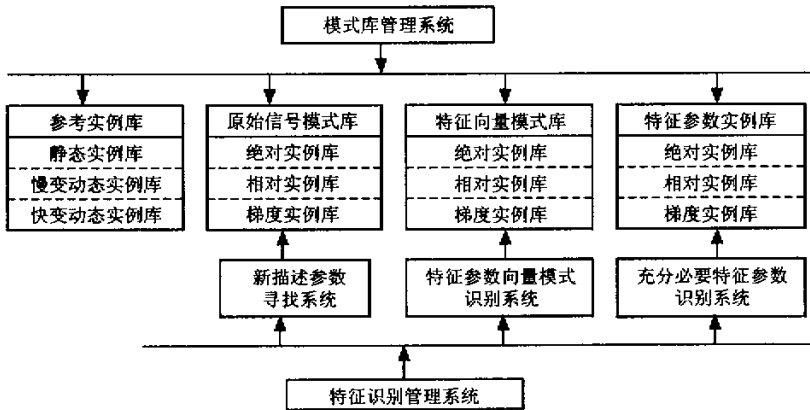


图1 从故障实例数据库中挖掘振动信号特征原理示意图

1.1 信号的绝对特征、相对特征和梯度特征

有些机组无故障时就具有很强的振动信号,如往复式压缩机,气体的非平稳压缩过程引起很强的振动(以下称其为背景振动,以 $b(t)$ 表示)。故障发生时,故障引起的振动(以 $r(t)$ 表示)和背景振动复合在一起构成总振动(以 $c(t)$ 表示)。对于线性动力学系统,振动响应满足叠加原理

$$c(t) = b(t) + r(t) \quad r(t) = c(t) - b(t) \quad (1)$$

由式(1)可知,为得到故障振动信号,需要知道背景振动 $b(t)$,因此在实例库挖掘原理中,背景振动以静态参考实例库存在。

并发故障的发生一般有一先后顺序。设有两种故障并发,引起振动分别以 $r_1(t)$ 和 $r_2(t)$ 表示,则

$$c(t) = b(t) + r_1(t) + r_2(t) \quad (2)$$

$r_1(t)$ 先于 $r_2(t)$,相对 $r_2(t)$ 而言 $b_1(t) = b(t) + r_1(t)$ 是背景振动信号,则 $b_1(t)$ 是相对 $r_2(t)$ 而言的参考实例,而且有一时间差,因此在实例库挖掘原理中, $b_1(t)$ 以慢变动态参考实例库存在。 $r(t)$ 和 $r_2(t)$ 所具有的特征称信号相对特征,以相对模式实例库表示。

一些故障信号所具有的特征表现在变化过程中,如压缩机叶轮上结垢、叶轮掉块、非线性振动的分叉等,因此信号变化的梯度(以 $g(t)$ 表示)也是很重要的特征,其是相对近期动态信号而言的,称为梯度特征,以梯度模式实例库表示。近期动态信号以快变动态参考实例库存在。

对非线性动力学系统,无法应用叠加原理,则特征只能以复合信号为对象,其所具有的特征称为绝对特征,以绝对模式实例库表示。

因此,该原理包含一个由无故障实例库(静态参考实例库)、慢变参考实例库和快变参考实例库构成的参考实例库。

一个参数描述信号特征最直接、明了,诊断推理效率也最高。然而单个参数的描述能力很有限。信号复杂时,需应用多个参数联合,这多个参数构成参数向量。参数向量维数越多描述能力越强。但维数越高,诊断推理效率越低。因此在满足故障区分能力的前提之下,维数越少越好。设原始信号的维数为 n ,描述信号的特征向量维数为 m 。当特征向量维数等于 n 时称其为原始信号模式;当 $0 < m < n$,称其为特征向量模式;当维数为1时称其为特征参数模式。本数据挖掘的目的就是获得最简单的特征向量模式库。原始信号、特征向量和特征参数实例分别由原始信号模式、特征信号模式和特征参数模式表达。

1.2 模糊聚类法从故障实例库中挖掘特征参数模式

应用模糊聚类分析从故障实例库中挖掘振动信号特征模式。由模式识别理论可知,一个向量可表示一种模式。设振动信号的分析参数有 n 个,可构成一向量 $x = \{\text{参数} 1, \text{参数} 2, \dots, \text{参数} n\}$,该向量可代表一种特征模式,实例库中有 m 个实例,可构成 m 个向量 x_1, x_2, \dots, x_m ,代表 m 种模式。由模糊聚类理论对 m 种模式进行分类。

①确定论域。将 x_1, x_2, \dots, x_m 确定为需要分类的全体,构成论域,即

$$X = \left\{ \begin{matrix} x_1 \\ x_2 \\ \dots \\ x_m \end{matrix} \right\} = \left[\begin{matrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \dots & \dots & \dots & \dots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{matrix} \right] \quad (3)$$

②建立相似模糊关系

$$\tilde{R} = \{r_{ij}\}_{m \times m}$$

应用相关系数法求 r_{ij}

$$r_{ij} = \frac{\sum_{k=1}^n |x_{ik} - \bar{x}_i| |x_{jk} - \bar{x}_j|}{\sqrt{\left(\sum_{k=1}^n (x_{ik} - \bar{x}_i)^2\right) \left(\sum_{k=1}^n (x_{jk} - \bar{x}_j)^2\right)}} \quad (4)$$

其中 $\bar{x}_i = \frac{1}{n} \sum_{k=1}^n x_{ik}, \bar{x}_j = \frac{1}{n} \sum_{k=1}^n x_{jk}$

③按直接聚类法对 m 个对象进行聚类计算

定义1 划分正确率定义为

$$\beta = \frac{\text{划分正确实例数}}{\text{总实例数}} \quad (5)$$

确定一个初始划分水平 λ_1 和步长 $\Delta\lambda$,按每一 λ_i 可计算一个结果 β_i 。在一个论域中可得出 $\beta = \max(\beta_i)$,此时的划分结果即为该论域的划分结果。

从 n 个参数中任取 $h (n \geq h \geq 1)$ 个构成一向量 $\{x_{11}, x_{12}, \dots, x_{1h}\}$, m 个实例就有 m 个向量,这 m 个向量构成一个论域,对该论域划分可得 β 和一个划分结果。

从 n 个参数中取 h 个参数可有 $s = n(n-1)\dots(n-h+1)/h!$ 种组合,可得 s 个论域,划分计算可得 s 个划分正确率 $\beta_j (j=1, 2, \dots, s)$ 。令

$$\beta_h = \max_s(\beta_j) \quad (6)$$

n 个参数可计算 n 个 $\beta_h (h=1, 2, \dots, n)$,令

$$\beta = \max_n(\beta_h) \quad (7)$$

具有 β 划分结果的参数向量即为从实例库中挖掘出的特征向量模式。

④新的描述参数的寻找

如果 $\beta=1$,说明现有参数已经能够完全描述故

障。如果 $\beta < 1$, 说明现有参数还不能完全描述故障, 还需要寻找新的参数, 需要应用其它的分析方法, 如小波分析法计算的参数。

⑤ 每个故障充分必要参数的寻找

某故障充分必要参数称为该故障的特征参数。将单独一个参数构成一个向量(一维向量), 进行划分计算, 如果划分结果有两类: 某一个故障单独为一类, 其它故障为一类, 则该参数为该故障的特征参数。不是所有故障都具有自己的特征参数。故障的特征参数可简化诊断推理。

2 应用实例

郑州大学振动工程研究所开发了往复式压缩机故障诊断系统, 其中数据挖掘系统作为其自学习系统, 以获得特征模式。其相关内容如下:

① 原始信号向量模式: {加速度 $P-P$ 值, 波形的陡峭系数, 均值, 方差, 频带宽度, 最小相关系数}

② 故障种类为 5 种: 地脚螺栓松动, 拉缸, 缸头漏气, 曲轴平衡铁装配不良, 气阀漏气。

③ 实例来源: 实例均来源于实验台信号。实验台、装置及测量设备如图 2 所示。地脚螺栓松动测点在缸体上, 气阀漏气测点在气阀座上, 缸头漏气测点在缸体顶盖上, 拉缸测点在缸体中间, 曲轴平衡铁装配不良测点在支撑座上。

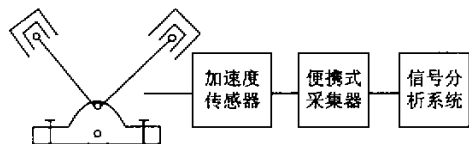


图 2 实验装置及测量设备图

④ 实例库每种故障测 4 个原始样本, 共 20 个原始样本。图 3 中显示了每种故障的一个样本信号, 其单位为 m^2/s , 样本采样长度为 4096 点。

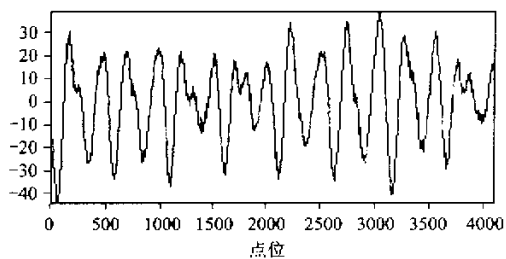
⑤ 绝对实例挖掘结果如表 1 所示, 相对实例挖掘结果如表 2 所示, 梯度实例挖掘结果如表 3 所示。

由表 1 中参数可得: 绝对特征划分正确率 $\beta_a = \max(\beta_h) = 0.8$, 相应划分水平 $\lambda = 0.45$ 。

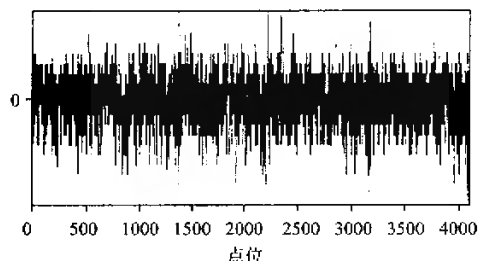
由表 2 中参数可得: 相对特征划分正确率 $\beta_r = \max(\beta_h) = 1.0$, 相应划分水平 $\lambda = 0.7$ 。

由表 3 中参数可得: 梯度特征划分正确率 $\beta_g = \max(\beta_h) = 0.6$, 相应划分水平 $\lambda = 0.3$ 。

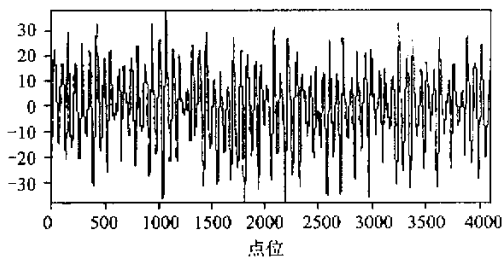
表中, 向量 $\mathbf{X} = \{ijk\}$ 代表 {参数 i , 参数 j , 参数 k } 向量, 如 $\{234\}$ 代表 {波形的陡峭系数, 均值, 方差}



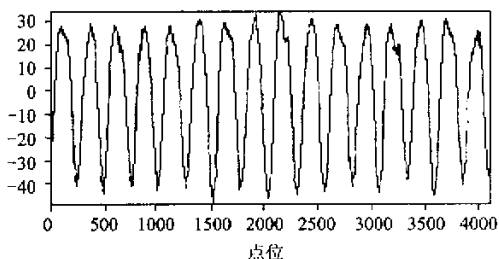
(a) 地脚螺栓松动故障样本信号



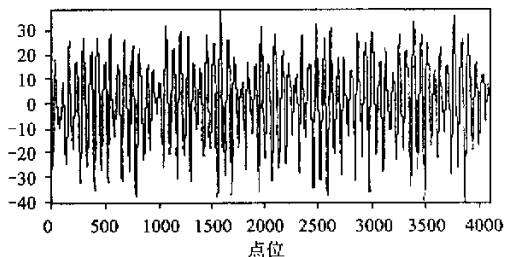
(b) 拉缸故障样本信号图



(c) 缸头漏气故障样本信号



(d) 曲轴平衡铁装配不良故障样本信号



(e) 气阀漏气故障样本信号

图 3 故障的样本信号

参数向量; β_j, β_h 的意义见式(6); λ_h 代表与 β_h 相应的划分水平。

表 1 绝对实例挖掘结果

| $h=1$ | | $h=2$ | | $h=3$ | | $h=4$ | | $h=5$ | | $h=6$ | |
|-------------|-----------|-------------|-----------|-------------|-----------|-------------|-----------|-------------|-----------|-------------|-----------|
| x | β_j | x | β_j | x | β_j | x | β_j | x | β_j | x | β_j |
| 1 | 0.4 | 12 | 0.2 | 123 | 0.2 | 1234 | 0.2 | 12345 | 0.4 | 123456 | 0.8 |
| 2 | 0.2 | 13 | 0.4 | 124 | 0.2 | 1235 | 0.4 | 12346 | 0.4 | | |
| 3 | 0.4 | 14 | 0.2 | 125 | 0.2 | 1236 | 0.4 | 23456 | 0.6 | | |
| 4 | 0.2 | 15 | 0.2 | 126 | 0.4 | 1245 | 0.2 | 13456 | 0.8 | | |
| 5 | 0.2 | 16 | 0.4 | 134 | 0.4 | 1246 | 0.2 | 12356 | 0.8 | | |
| 6 | 0.4 | 23 | 0.2 | 135 | 0.6 | 1256 | 0.6 | 12456 | 0.4 | | |
| | | 24 | 0.2 | 136 | 0.6 | 1345 | 0.4 | | | | |
| | | 25 | 0.2 | 145 | 0.4 | 1346 | 0.4 | | | | |
| | | 26 | 0.4 | 146 | 0.2 | 1356 | 0.8 | | | | |
| | | 34 | 0.4 | 156 | 0.6 | 1456 | 0.4 | | | | |
| | | 35 | 0.6 | 234 | 0.2 | 2345 | 0.4 | | | | |
| | | 36 | 0.6 | 235 | 0.6 | 2346 | 0.4 | | | | |
| | | 45 | 0.4 | 236 | 0.6 | 2356 | 0.8 | | | | |
| | | 46 | 0.4 | 245 | 0.4 | 2456 | 0.4 | | | | |
| | | 56 | 0.6 | 246 | 0.4 | 3456 | 0.4 | | | | |
| | | | | 256 | 0.6 | | | | | | |
| | | | | 345 | 0.6 | | | | | | |
| | | | | 346 | 0.6 | | | | | | |
| | | | | 356 | 0.8 | | | | | | |
| | | | | 456 | 0.6 | | | | | | |
| β_h | 0.4 | β_h | 0.6 | β_h | 0.8 | β_h | 0.8 | β_h | 0.8 | β_h | 0.8 |
| λ_h | 0.4 | λ_h | 0.4 | λ_h | 0.45 | λ_h | 0.6 | λ_h | 0.6 | λ_h | 0.6 |

表 2 相对实例挖掘结果

| $h=1$ | | $h=2$ | | $h=3$ | | $h=4$ | | $h=5$ | | $h=6$ | |
|-------------|-----------|-------------|-----------|-------------|-----------|-------------|-----------|-------------|-----------|-------------|-----------|
| x | β_j | x | β_j | x | β_j | x | β_j | x | β_j | x | β_j |
| 1 | 0.4 | 12 | 0.6 | 123 | 0.4 | 1234 | 0.4 | 12345 | 0.8 | 123456 | 1.0 |
| 2 | 0.4 | 13 | 0.4 | 124 | 0.4 | 1235 | 0.8 | 12346 | 0.8 | | |
| 3 | 0.2 | 14 | 0.4 | 125 | 0.6 | 1236 | 0.4 | 23456 | 0.6 | | |
| 4 | 0.2 | 15 | 0.4 | 126 | 0.6 | 2345 | 0.4 | 13456 | 0.6 | | |
| 5 | 0.4 | 16 | 0.6 | 134 | 0.2 | 2346 | 0.4 | 12356 | 0.8 | | |
| 6 | 0.4 | 23 | 0.2 | 135 | 0.4 | 3456 | 0.2 | 12456 | 1.0 | | |
| | | 24 | 0.2 | 136 | 0.2 | 1345 | 0.4 | | | | |
| | | 25 | 0.4 | 145 | 0.4 | 1346 | 0.4 | | | | |
| | | 26 | 0.4 | 146 | 0.4 | 1456 | 0.6 | | | | |
| | | 34 | 0.2 | 156 | 0.3 | 2456 | 0.6 | | | | |
| | | 35 | 0.2 | 234 | 0.2 | 1245 | 0.8 | | | | |
| | | 36 | 0.2 | 235 | 0.4 | 1246 | 0.8 | | | | |
| | | 45 | 0.4 | 236 | 0.4 | 1256 | 1.0 | | | | |
| | | 46 | 0.6 | 245 | 0.4 | 1356 | 0.6 | | | | |
| | | 56 | 0.6 | 246 | 0.4 | 2356 | 0.8 | | | | |
| | | | | 256 | 0.4 | | | | | | |
| | | | | 345 | 0.2 | | | | | | |
| | | | | 346 | 0.2 | | | | | | |
| | | | | 356 | 0.4 | | | | | | |
| | | | | 456 | 0.4 | | | | | | |
| β_h | 0.4 | β_h | 0.6 | β_h | 0.6 | β_h | 1.0 | β_h | 1.0 | β_h | 1.0 |
| λ_h | 0.4 | λ_h | 0.5 | λ_h | 0.5 | λ_h | 0.7 | λ_h | 0.7 | λ_h | 0.7 |

表3 梯度实例挖掘结果

| $h=1$ | | $h=2$ | | $h=3$ | | $h=4$ | | $h=5$ | | $h=6$ | |
|-------------|-----------|-------------|-----------|-------------|-----------|-------------|-----------|-------------|-----------|-------------|-----------|
| x | β_j | x | β_j | x | β_j | x | β_j | x | β_j | x | β_j |
| 1 | 0.2 | 12 | 0.2 | 123 | 0.2 | 1234 | 0.2 | 12345 | 0.2 | 123456 | 0.6 |
| 2 | 0.0 | 13 | 0.2 | 124 | 0.2 | 1235 | 0.4 | 12346 | 0.4 | | |
| 3 | 0.2 | 14 | 0.0 | 125 | 0.2 | 1236 | 0.4 | 23456 | 0.4 | | |
| 4 | 0.0 | 15 | 0.0 | 126 | 0.2 | 2345 | 0.2 | 13456 | 0.6 | | |
| 5 | 0.2 | 16 | 0.0 | 234 | 0.2 | 2346 | 0.2 | 12356 | 0.6 | | |
| 6 | 0.2 | 23 | 0.0 | 235 | 0.4 | 3456 | 0.2 | 12456 | 0.4 | | |
| | | 24 | 0.0 | 236 | 0.4 | 1345 | 0.4 | | | | |
| | | 25 | 0.2 | 345 | 0.4 | 1346 | 0.4 | | | | |
| | | 26 | 0.2 | 346 | 0.4 | 1456 | 0.4 | | | | |
| | | 34 | 0.2 | 456 | 0.4 | 2456 | 0.2 | | | | |
| | | 35 | 0.4 | 134 | 0.4 | 1245 | 0.2 | | | | |
| | | 36 | 0.2 | 135 | 0.4 | 1246 | 0.4 | | | | |
| | | 45 | 0.2 | 136 | 0.4 | 1256 | 0.4 | | | | |
| | | 46 | 0.2 | 145 | 0.4 | 1356 | 0.6 | | | | |
| | | 56 | 0.4 | 146 | 0.4 | 2356 | 0.4 | | | | |
| | | | | 245 | 0.2 | | | | | | |
| | | | | 246 | 0.2 | | | | | | |
| | | | | 256 | 0.4 | | | | | | |
| | | | | 356 | 0.4 | | | | | | |
| β_h | 0.2 | β_h | 0.4 | β_h | 0.4 | β_h | 0.6 | β_h | 0.6 | β_h | 0.6 |
| λ_h | 0.3 | λ_h | 0.3 | λ_h | 0.35 | λ_h | 0.3 | λ_h | 0.4 | λ_h | 0.45 |

从表1、表2和表3可看出,对以上五种故障而言,往复式压缩机的相对特征最明显,梯度特征最不明显。说明该系统是拟线性系统,振动平稳。

3 结论

以上分析可知:

①对于复杂故障,应用模式可增强信号特征的描述能力;

②应用数据挖掘技术从大量实例库中获取信号特征模式是一条可行的途径;

③绝对、相对和梯度特征是从不同角度描述故障信号特征的有效方法。

为提高该方法的实用性,对每一个系统需要大量的实例。因此,进一步需要研究能够将现有文献上的实例转化为系统可接受的、统一表达形式的实例的方法。

参 考 文 献

1 Milne R, Nelson C. Knowledge guided data mining.

IEE Colloquium on Case Based Reasoning: Prospects for Applications, London, UK, IEE, London, UK, 10/11—13(355)

- 2 吉根林,帅克,孙志挥. 数据挖掘技术及其应用. 南京师范大学学报, 2000;(2):25—27
- 3 窦万春,李东波,王玉等. 基于知识应用的数据挖掘技术理论分析与应用研究. 计算机集成制造系统-CIMS, 2000;(4):65—69
- 4 Carse, Davies T S, Pipe A G, et al. A fuzzy logic based approach to power system fault location. IEE Colloquium on Knowledge Discovery and Data Mining, London, UK, IEE, UK, 1996;335(8):1—4
- 5 Rao M, Wang Q. Generic integrated intelligence system architecture for diagnosis and maintenance. Canadian Artificial Intelligence, 1996;330:35—51
- 6 Wang Z, Lawrenz W, Rao R B K N. Feature-filtered fuzzy clustering for condition monitoring of tool wear. Intelligent Manufacturing, Chapman&Hall, 1996; 22: 13—22
- 7 高毅龙. 数据挖掘技术及其在工程诊断中的应用:[博士学位论文]. 西安:西安交通大学,2000

Data-Mining Diagnostic Features of Vibration Signal from Fault-Case Database

Guan Huiling¹ Zhang Youyun¹ Han Jie² Dong Xinmin² Hao Wei²

(¹Institute of Lubrication Theory and Bearing ,Xi'an Jiaotong University Xi'an,710049)

(²Institute of Vibration Engineering ,Zhengzhou University Zhengzhou,450002)

Abstract The difficult problem of complex and syndrome fault diagnosis is that the diagnostic features might not be obtained because the signal of the fault is very puzzling. This paper proposes a diagram of the system of data-mining diagnostic patterns, defines an absolute, relative and gradient diagnostic parameters and correspondingly patterns of absolute, relative and gradient fault-cases. The method of data-mining diagnostic patterns using fuzzy clustering is discussed. In the last, the diagnostic pattern data-mining system from fault-case database of reciprocating compressor is demonstrated.

Key words: fault diagnosis; data mining; characteristic obtaining; knowledge discovery; machine learning

第一作者 关惠玲 女,教授,博士研究生,1962年4月生。电话:(0371)3887383; E-mail: hlguan0@yahoo.com

从故障实例数据库中挖掘振动信号特征

作者: 关惠玲, 张优云, 韩捷, 董辛旻, 郝伟

作者单位: 关惠玲, 张优云(西安交通大学润滑理论与轴承研究所, 西安, 710049), 韩捷, 董辛旻, 郝伟(郑州大学振动工程研究所, 郑州, 450002)

刊名: 振动工程学报 

英文刊名: JOURNAL OF VIBRATION ENGINEERING

年, 卷(期): 2002, 15(3)

引用次数: 3次

参考文献(7条)

1. Milne R, Nelson C Knowledge guided data mining
2. 吉根林, 帅克, 孙志挥 数据挖掘技术及其应用[期刊论文]-南京师大学报(自然科学版) 2000(2)
3. 窦万春, 李东波, 王玉, 张世琪 基于知识应用的数据挖掘技术理论与应用研究[期刊论文]-计算机集成制造系统 2000(4)
4. Carse, Davies T S, Pipe A G A fuzzy logic based approach to power system fault location 1996
5. Rao M, Wang Q Generic integrated intelligence system architecture for diagnosis and maintenance 1996
6. Wang Z, Lawrenz W, Rao R B K N Feature-filtered fuzzy clustering for condition monitoring of tool wear 1996
7. 高毅龙 数据挖掘技术及其在工程诊断中的应用 2000

相似文献(10条)

1. 学位论文 万鹏 机械故障诊断中基于神经网络的数据挖掘方法研究 2006

随着生产的发展, 机械故障诊断的重要性越来越明显。传统的诊断技术和理论方法对于具有多故障、多过程、突发性故障的现代化机械设备, 往往显示出较大的局限性, 难以从大量的故障信息中发现原因和准确地诊断机械故障。而数据挖掘是从大型数据库或数据仓库中发现并提取隐藏在其中的信息的一种新技术, 它能从数据仓库中自动分析数据, 进行归纳性推理, 从中挖掘出潜在的模式; 或者产生联想, 建立新的业务模型, 帮助决策者做出正确的决策。数据挖掘研究和实践表明: 知识就隐藏在日常积累下来的大量数据之中, 然而仅靠复杂的算法和推理并不能发现知识。人工智能技术, 特别是神经网络技术与数据挖掘的结合为机械故障诊断的研究指出了一条新的道路。首先, 本文对机械故障诊断、数据挖掘的研究现状、发展趋势及应用领域进行了综述, 阐述了数据挖掘、神经网络的基本概念、基本模型和实现方法。其次, 针对数据挖掘具体的研究内容, 提出并实现了神经网络数据挖掘中的数据准备的一般处理方法和基于自组织神经网络的聚类分析的模型和算法, 并对算法进行了改进。最后, 根据基于神经网络数据挖掘的研究成果, 进行了轴承和齿轮的故障诊断研究。研究内容包括: 问题定义、数据准备和预处理、数据挖掘, 以及结果的解释和评估。首先通过传感器测得轴承及齿轮在正常和非正常状态下的信号; 然后对所得数据进行归一化; 考虑到数据比较庞大, 利用曲线分析进行降维; 再利用自组织神经网络进行训练, 网络对不同状态下的输入具有明显不同的输出。文中利用MATLAB神经网络工具箱来实现上述算法。实例仿真表明, 这个算法可以快速准确地提取出轴承与齿轮故障特征值, 并通过聚类算法完成轴承与齿轮的故障诊断。基于神经网络数据挖掘方法为机械故障诊断提供了新的思路, 它与一般数据挖掘方法相比较, 在信息处理能力和表示能力方面有较大提高。

2. 学位论文 辛旭明 基于过程监控的数据挖掘系统及其应用 2003

该课题对工业过程故障诊断进行研究和探索。随着计算机与网络信息技术的发展, 故障诊断中采集的数据急剧增加, 因而将数据挖掘技术应用于故障诊断中。为此, 该文提出了基于故障诊断的数据挖掘系统。基于故障诊断的数据挖掘系统, 既是一个数据挖掘系统, 又必须满足故障诊断的特殊性。基于故障诊断的数据挖掘系统针对故障诊断提供解决方案。该系统以粗集为关键技术。从实际中获得的数据常常包含噪声, 存在不确定因素和不完整信息。运用纯数学上的假设来消除不确定性, 往往效果不理想。粗集作为数据挖掘方法的一种, 是一种处理不精确或不完整信息的强有力数学工具, 在故障诊断中的应用是一个新的研究方向。粗集能从大量数据中挖掘出有用知识, 所以能有效地应用于故障诊断, 并且可以克服以往存在的知识获取瓶颈。在设计算法的时候, 充分考虑到故障诊断的数据和需求的特殊性, 并作了优化。引入规则的置信度和支持度, 提出了在故障诊断中的基于粗集的数据挖掘算法, 提出了在故障诊断中的时态信息系统转换为信息系统的方法。使用数据库连接池技术, 连接池技术尽可能多的重用了消耗内存的资源, 大大地省了内存, 提高了程序效率, 同时, 通过其自身的管理机制来监视数据库的连接数量、使用情况等。该系统以粗集为关键技术, 采用适合故障诊断的算法, 从现有的故障诊断的数据出发, 能有效地处理不完备数据、不一致数据, 实现提取诊断规则的目的, 高效快速地导出故障诊断规则。该文论述了粗集基本理论、扩展理论及时间序列数据的处理方法, 提出了基于故障诊断的具体算法, 并介绍了基于故障诊断的数据挖掘系统结构及功能, 最后将系统应用于机械故障诊断。

3. 学位论文 孙卫祥 基于数据挖掘与信息融合的故障诊断方法研究 2007

随着科学技术的飞速发展和市场竞争的日益激烈, 工业生产表现出大型化、分布化、高速化、自动化和复杂化等特点。由于这些大型系统一般都是作为能源、石化、冶金以及其他国民经济支柱产业中的关键设备, 一旦发生故障, 轻则降低生产效率, 重则设备停机、生产停顿, 有时甚至发生机毁人亡的恶性事故, 造成灾难性的后果。所以, 有计划、有组织、有针对地对关键设备进行实时监测与诊断, 做到尽早地发现设备在运行过程中的各种隐患, 从而防止灾难性事故的发生, 成为机械设备故障诊断系统面临和解决的首要问题。数据挖掘、信息融合理论以及现代计算机、测试、信号分析处理等技术的发展, 为机械状态监测与故障诊断提供了一种崭新的解决方案。本文在分析了机械故障诊断的现状、发展趋势以及数据挖掘、信息融合技术对于机械故障诊断的重要性的基础上, 主要从事了以下几点研究: (1) Bently转子台模拟故障试验: Bently转子台是一个通用的转子故障模拟试验平台, 本文使用的故障诊断数据来自该试验台。本文共模拟转子的5种运行状态, 一种为正常状态, 另四种为故障状态包括不平衡、摩擦、涡动以及不平衡+摩擦。

(2) 信息融合包括数据层、特征层和决策层三个层次的融合, 数据层融合数据量巨大, 计算效率较低。因此, 本文首先从特征融合角度, 研究利用多特征来提高故障诊断的精度。本文使用的诊断特征有三类, 包括时域统计指标(6维)、频域幅值谱(13维)以及时频域小波能量谱(25维), 采用PCA进行特征融合。基于Bently转子台的试验数据诊断分析表明, 基于PCA的特征融合诊断方法有效地提高了故障诊断的精度。

(3) 为了有效地利用数据挖掘的先进技术, 对数据挖掘中两种分类方法C4.5决策树以及SVM进行了改进。根据Fayyad边界点判定定理, 改进了C4.5算法连续属性离散化过程中最优阈值的选择方法, 提高了连续属性离散化效率。根据传统SVM的不足, 提出一种支持向量机决策树(SVM Decision Tree, SVMDT)分类模型。它解决了“1对1”与“1对多”策

略下SVM分类方法的拒识问题,同时它对输出结果进行了的概率赋值,方便了后续的决策融合。在此基础上,以C4.5决策树、SVMDT以及后向传播神经网络(BPNN)作为基本分类方法,分别针对三种特征进行诊断决策,最后采用改进D-S证据合成理论进行决策融合。试验数据的诊断分析表明,基于多特征、多分类器的信息融合方法能够大大提高故障诊断精度。振动、冲击、噪声国家重点实验室。(4)为了确定决策融合时多分类器的选择以及说明多分类器决策融合对于提高诊断精度的有效性,本文从分类器间的差异性角度对多分类器融合进行了研究,同时给出一种有效差异性(Efficient Diversity, ED)度量指标作为分类器选择与融合的依据,有效差异度ED越大说明分类器间的差异性越大,分类器决策融合后对诊断精度的提高越显著。

(5)基于数据挖掘的过程模型,结合信息融合思想与方法,设计了一种新的基于数据挖掘与信息融合的知识获取与故障诊断系统框架结构。该系统包括数据采集、数据变换、特征选择、特征融合、单一分类器故障诊断、规则或模型知识表示、决策融合等模块。实验室测试以及某风机监测诊断的现场应用验证了系统设计的有效性。

4. 学位论文 [侯淑旻](#) [时序数据挖掘及其在故障诊断中的应用研究](#) 2006

数据挖掘是一个复杂的过程,而针对时间序列具有维数高和动态性等特点,如何高效地对时间序列数据进行挖掘是一个十分有意义的研究课题。本文对时间序列数据非线性检验、降噪、分割等数据挖掘中的预处理工作进行了研究,并将时间序列数据挖掘技术引入机械故障诊断领域;通过应用实例,验证了这些方法和理论。论文主要完成工作如下: 1)时间序列数据非线性检验研究提出了一种随机迭代修正幅值的傅立叶变换(SIAFFT)算法和KS检验法相结合的时间序列数据非线性检验方法,通过对弱非线性信号、强非线性信号以及含噪信号的检验,该方法均能得到正确的判断结果。与传统方法相比,该方法收敛速度快,具有较强的抗噪声能力,对非线性信号具有较高的敏感性。 2)全局投影算法用于降噪和故障特征提取全局投影降噪算法具有良好的适用性,运算效率和降噪效果明显优于局部投影降噪算法。将全局投影算法应用于转子轴心轨迹的提纯,比基于谐波小波等的轴心轨迹提纯方法运算效率高得多。 3)将全局投影降噪算法和共振解调技术相结合,提出了一种适用于低速重载轴承故障诊断的方法。利用此方法对某炼钢厂炉轴机构悬挂齿轮箱耳轴轴承进行故障诊断,诊断结果与实际相符。 4)基于GG聚类的时间序列数据分割方法针对传统时间序列分割算法的不足,提出了一种基于GG聚类的在线数据分割算法。该算法具有自合并功能,能不借助相关领域专家的支持自动寻找最优的分割子集数,是一种有工程应用价值的在线分割方法。 5)基于KS检验的时序数据分类挖掘系统提出了一种基于KS检验的时序数据分类挖掘方法,在此方法基础上,建立了故障分类系统。通过仿真试验和齿轮及轴承的故障诊断,说明该方法在数据样本含有一定噪声时也能正确判断故障类型。该方法算法简单、计算效率高、实用性强,在机械故障智能诊断领域具有广泛的应用前景和推广价值。 6)改进型支持向量机(ES-SVM)用于时间序列数据趋势预测将传统的基于支持向量机的时间序列预测方法进行改进,采用进化策略搜索惩罚因子C, ϵ 不敏感损失函数和高斯核参数 σ 。Lorenz信号的数值仿真实验预测结果证明改进后的ES-SVM算法比原SVM算法的预测精度更高。另外,由于在工程实际中所抽取的信号不可避免会含有噪声,因此,将全局投影算法和ES-SVM相结合的预测方法具有更大的工程应用价值。 7)构建基于时序数据挖掘的远程监测与诊断系统作为时序数据挖掘应用的重要组成部分,本文以机械在线监测系统为对象,对数据库结构、数据类型进行了探讨;并针对机械信号的特点,建立了时间序列数据挖掘快速处理模型。在此基础上,结合企业对具体远程监测与诊断系统的要求,构建了基于时间序列数据挖掘的远程监测与诊断系统,并成功应用于某炼钢厂抽风机。

5. 学位论文 [徐宗龙](#) [数据挖掘在齿轮箱故障诊断中的应用研究](#) 2006

随着数据库和数据仓库技术在故障诊断领域的广泛应用,以及当前大部分企业对重要设备安装了在线监测系统,并且已经形成存储设备运行状态数据的大型数据仓库。如何从这些大量的、杂乱无章的数据中挖掘出对故障诊断有用的决策和知识,成为本文研究的重点。 本文研究了数据挖掘技术、OLAP技术和联机分析挖掘理论,将数据挖掘中关联规则方法与数据立方体相结合,构建基于OLAP的关联规则挖掘模型,并将产生的规则应用于故障诊断领域,为故障诊断提供决策依据。 本文讨论了齿轮箱故障诊断的主要内容,介绍了常见的齿轮箱故障的振动特征的提取和分析方法,提出了基于数据挖掘的故障诊断思想,介绍了基于OLAP的关联规则挖掘模型结构及工作原理和过程,并结合齿轮箱的故障诊断对模型组成进行了研究。 在WindowsXP操作平台下,使用SQL Server 2005,开发了关联规则挖掘模型。利用振动试验台的模拟故障数据对提出的模型进行了训练和考核。结果表明,关联规则挖掘模型产生的规则,能够为故障诊断提供决策。

6. 学位论文 [许苗](#) [数据挖掘在轴承故障诊断中的应用](#) 2005

本文以大连海事大学自动化研究所与瓦轴集团轴承试验测试中心合作的轴承疲劳寿命二期技术改造与增容项目为选题背景,结合当前国内外关于轴承寿命研究的发展现状,针对数据库和数据仓库技术在故障诊断领域的广泛应用以及现在各工厂基本上都对重要设备实施了实时监控,由传感器不间断的传回试验机组运行状态的各种数据及其参数,已经形成大型数据库或数据仓库这一事实,提出将数据挖掘技术应用于轴承的故障诊断领域,找出典型故障数据中的潜在知识,完成典型故障的识别问题,以推动故障诊断技术的不断向前发展,同时,对轴承寿命的预测做了探讨性的尝试。 决策树、关联规则等数据挖掘方法能产生显示的规则,并能有效解决海量数据中知识的发现问题。决策树是一个类似于流程图的树结构,主要用途是提取分类规则,进行分类预测;关联规则挖掘寻找给定数据集中项之间的有趣联系,并产生规则。 本文选用C++Builder自带数据库作为存储设备状态信号的数据库管理系统,建造用于存储设备数据状态的简易数据库。在WINDOWS98开发平台上,使用C++Builder开发工具采用面向对象的程序设计思想和模块化程序设计方法,对给予数据挖掘的轴承故障诊断系统进行软件实现。 鉴于将数据挖掘方法应用于轴承的故障诊断上,毕竟是一种新方法的尝试,若采用单一的分类方法未必会取得良好的效果,本文就是采用上述两种数据挖掘的方法来处理实际问题的。实践证明,根据具体问题,有针对性的采用几种方法相结合使用,可提取出很有价值的轴承故障规律。 另外,文章最后调用SPSS软件包,应用回归分析原理中的曲线估计对轴承的峰值指标进行简单的曲线拟合,预测轴承疲劳寿命。实验结果表明,建立了较准确的预测模型,可以较准确地预测轴承寿命。

7. 期刊论文 [何春芳](#)、[傅成华](#)、[郭辉](#) [基于数据挖掘与仿真的数控机床旋转部件故障诊断方法研究](#) -科技信息(学术版) 2007(24)

针对数控机床旋转部件故障信息的特点,提出了系统仿真与数据挖掘相结合的综合故障诊断方法。文中将旋转部件故障归为4类,提出了一种改进的Apriori算法,结合粗糙集理论的数据挖掘方法,通过实例证明了两者在数控机床系统旋转部件故障诊断中,可以提高数据挖掘的速度。分析了基于仿真的故障诊断方法,在此基础上,结合仿真与数据挖掘各自在故障诊断方面的优势,进一步提出了基于仿真与数据挖掘的综合诊断方法,给出了诊断的方法流程图。

8. 期刊论文 [宋长新](#)、[吴晓明](#)、[SONG Chang-xin](#)、[WU Xiao-ming](#) [工业锅炉故障诊断的改进数据挖掘方法](#) -微计算机信息 2009, 25(19)

针对工业锅炉的常见故障,提出了一种基于数据挖掘方法的锅炉故障诊断技术。通过建立一个智能化的数据挖掘工具,直接从大量实时数据中获取故障诊断知识进行故障诊断。数据挖掘工具的核心是采用信息熵技术来辅助遗传算法初始种群的生成。遗传算法和信息熵的有效集成,极大地提高了该数据挖掘方法的工作绩效。将本方法应用于火电厂锅炉的一个复杂故障事例,结果表明其诊断的精度可以满足现场应用的要求。

9. 学位论文 [石金彦](#) [基于规则的数据挖掘方法在故障诊断中的应用](#) 2003

决策树是一个类似于流程图的树结构,主要用途是提取分类规则,进行分类预测。粗糙集理论无需提供除问题相关的数据集外的任何先验信息,适合于发现数据中潜在的规律、不准确数据或噪声数据内在的结构联系,可以解决重要的分类问题。关联规则反映一个事件和其他事件之间的依赖或关联的知识。如果两项或多项属性之间存在关联,那么其中一项的属性值就可以依据其他属性值进行预测。应用于故障诊断领域,对大量的机组状态数据进行挖掘,发现故障数据中存在的规律,以规则的形状体现出来,可以为故障诊断提供决策依据。利用转子试验台的模拟故障数据对提出的方法进行考核。结果表明,系统采用的几种方法所得出的规则是正确的,并能正确用于故障的分类工作,也可为专家提供有价值的信息。在WINDOWS98开发平台上,结合开发工具Microsoft Visual C++6.0,SQL Server数据库管理系统,采用面向对象的程序设计思想和模块化程序设计方法,对基于显式规则的几种数据挖掘方法进行了软件实现,开发了基于故障诊断的挖掘系统,该系统具有一定的实用价值。该文的创新点在于将数据挖掘产生规则的方法引入了故障诊断,结合先进的数据库管理系统,将数据挖掘中关联规则、粗糙集以及决策树方法用于故障的分类工作,并最终软件实现了以上的构想。

10. 期刊论文 [张素兰](#)、[胡骏](#)、[ZHANG Su-lan](#)、[HU Jun](#) [基于仿真与数据挖掘的故障诊断方法研究](#) -计算机仿真 2007, 24(1)

针对导弹武器系统故障信息的特点,提出了系统仿真与数据挖掘相结合的综合故障诊断方法。文中将导弹武器系统故障检测信息分为3类,即:离散交

互特征信息、连续动态特征信息和离散事件特征信息,通过实例阐述了数据挖掘在导弹武器系统故障诊断中的应用,分析了基于仿真的故障诊断方法,在此基础上,结合仿真与数据挖掘各自在故障诊断方面的优势,进一步提出了基于仿真与数据挖掘的综合诊断方法,给出了方法的步骤和诊断流程。

引证文献(7条)

1. 侯明, YANG Jun-dong 机电设备传感器嵌入式WEB监控系统研究[期刊论文]-微计算机信息 2008(20)
2. 刘燕, 李世其, 宋义超 舰船动力装置智能诊断方法研究[期刊论文]-中国修船 2007(02)
3. 万绪财 基于数据挖掘技术的锅炉高温受热面寿命管理的研究[学位论文]硕士 2006
4. 刘兵 数据挖掘技术在某型测量船动力装置监测及维修管理中的应用[学位论文]硕士 2006
5. 伍星 基于数据挖掘的设备远程监测和故障诊断系统的研究[学位论文]博士 2005
6. 袁艳霞 基于组件的设备状态监测与故障诊断系统的研究[学位论文]硕士 2005
7. 刘钊, 杨林权, 吕维先 基于粒子群优化算法的自适应滤波[期刊论文]-计算机与现代化 2004(04)

本文链接: http://d.g.wanfangdata.com.cn/Periodical_zdgcxb200203018.aspx

下载时间: 2009年12月31日