

文章编号:1000-582X(2010)04-021-05

# 智能诊断中诊断知识的实时自学习方法

孙红岩,姜雪峰

(重庆大学 机械传动国家重点实验室,重庆 400044)

**摘要:**模拟人类专家诊断知识积累过程,提出一种基于模式比较和更新的诊断知识实时自学习模型,以实现智能诊断系统知识库知识的实时积累和更新。用非正态分布假设检验算法实现待确定知识故障模式与正常模式的实时比较,用样本含量估计算法作为有效模式比较所需故障样本量确定的理论依据,用实时模式更新算法实现诊断知识模式与设备的实际运行情况一致。实际测试数据分析表明该模型有利于实现新知识的积累和诊断知识贴近实际设备的实际情况。

**关键词:**智能诊断;诊断知识;自学习;模式识别;假设检验

中图分类号:TP277

文献标志码:A

## Real-time self-learning method of diagnostic knowledge in intelligent diagnosis system

SUN Hong-yan, JIANG Xue-feng

(State Key Laboratory of Mechanical Transmission, Chongqing University,  
Chongqing 400044, P. R. China)

**Abstract:** Aiming at real-time diagnostic knowledge accumulation and updating in intelligent diagnostic system, to simulate the process experts accumulate and update fault diagnostic knowledge, a new real-time diagnostic knowledge self-learning model is proposed based on pattern comparing and updating. Abnormal distribution test of hypothesis is used to compare real-time equipment fault pattern to equipment normal pattern of contingent knowledge. The algorithm of sample size estimation algorithm is used to calculate the number of samples which is used to obtain necessary diagnostic efficiency. Real-time model updating algorithm is used to adjust the diagnostic knowledge model to actual equipment on the spot. Analysis to actual test data of equipment shows that the method could achieve new diagnostic knowledge accumulation and be adaptive to actual equipment on the spot.

**Key words:** intelligent diagnosis; diagnostic knowledge; self-learning; pattern recognition; test of hypothesis

设备故障的智能诊断过程是利用已有知识和经验的诊断<sup>[1-5]</sup>,其复杂性决定了诊断知识库的不完备性<sup>[6-7]</sup>,由于环境噪声的影响,使得设备在不同情况中表现出不同特点,与实际设备实际情况吻合的知识库更难建立。这是智能诊断理论不能实用化的重

要原因。为解决诊断知识库的完备性问题,要求智能诊断系统具有较强的自学习能力,即学习新知识和归纳、总结实时经验的能力。

在很多领域都有自适应学习方法的研究<sup>[8-12]</sup>,这些研究主要集中为 2 种方式:一种方式是用函数

收稿日期:2010-12-10

基金项目:重庆市自然科学基金资助项目(CSTC,2008BB3179)

作者简介:孙红岩(1974-),女,重庆大学博士后,从事机械工程测试与设备智能诊断的研究,(E-mail)sunhyxf@cqu.edu.cn。

或矩阵优化诊断知识。但是关于如何确定矩阵或函数,目前还没有明确的理论依据。另外一种是将知识融合在智能识别算法本身参数中,自学习过程是对这些参数进行优化。这种自学习方法受算法本身局限性限制。另外,自学习只是针对算法参数的改变,不能实现对诊断知识真正意义上的自学习,因而不能从本质上丰富诊断知识库。

用计算机模拟人类专家思维是人工智能领域研究者进行研究的指导原则和追求的目标。对于故障诊断,人类专家诊断知识库的积累过程是对多种模式的总结、对比,从而积累出用来判断设备状态的知识经验。因此,提出模拟人类专家构建基于模式比较和更新的诊断知识实时自学习模型。

从数理统计学的角度<sup>[15]</sup>,诊断知识模式中的变量是随机变量,则一个诊断知识模式可以看作是一个随机样本空间,即总体。则模式之间的比较转换为样本总体之间的比较,即故障时的模式总体与正常状态的模式总体之间的模式比较问题。因此,采用假设检验方法作为模式实时比较算法。在进行假设检验时,用于检验的样本量要满足一定的条件,才能满足检验效能,选用样本含量估计来确定样本数量。

为了适应实际设备实际环境,构建实时模式更新算法,实现系统运行过程中实时更新诊断知识模式。

### 1 诊断知识模式的建立

借鉴高文等对鱼病诊断知识概念描述<sup>[13]</sup>,将在设备故障诊断中所涉及的知识分为诊断知识和诊断策略知识两个层次。前者为直接用于诊断的知识,集中体现故障原因、故障种类和故障征兆 3 个方面。诊断策略知识是诊断知识的控制策略,即诊断知识之间的关系表示。设备智能诊断主要是针对故障种类和故障征兆两方面诊断知识及其关系进行研究。因此设备诊断知识模型如图 1。

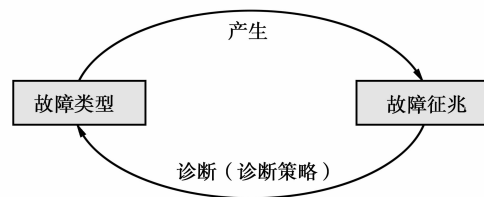


图 1 设备诊断知识

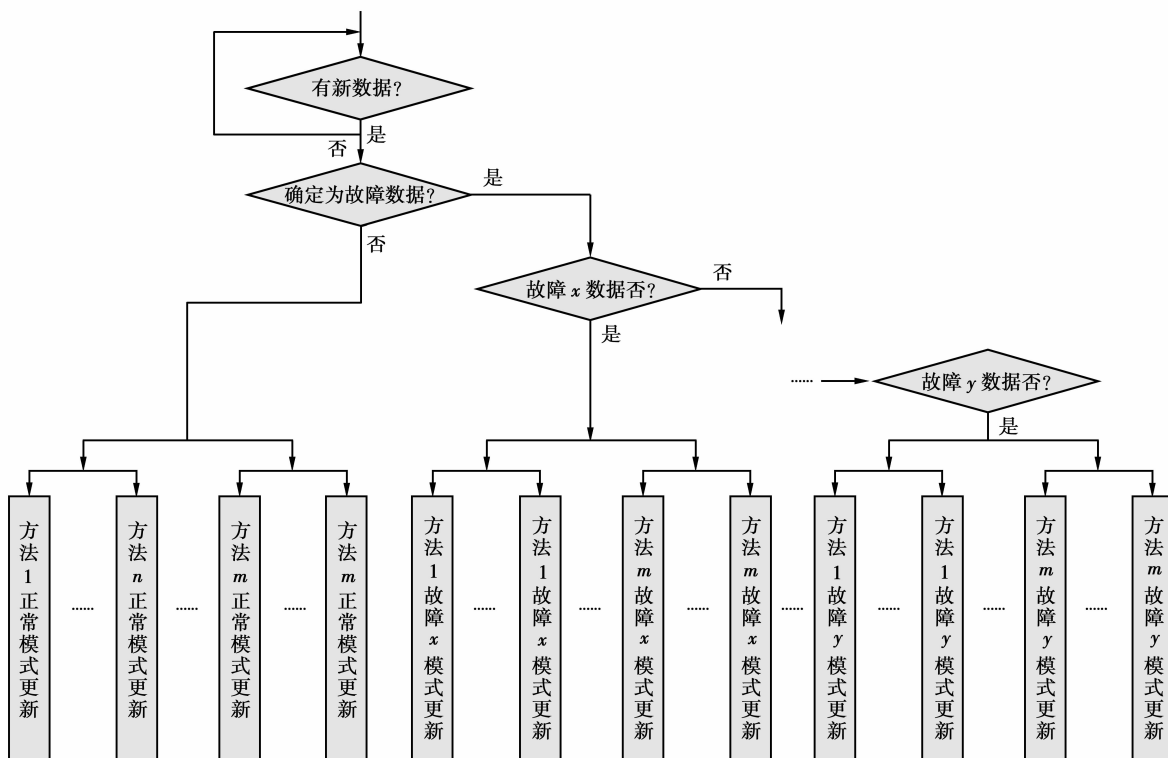


图 2 确定性知识的自学习流程

根据对诊断知识的确定情况,将诊断知识分为确定性诊断知识(即已经确认为诊断知识的知识)和待确定诊断知识(正在学习当中尚未被确认的诊断知识)。

习惯上将诊断知识表示为:故障类型,征兆。即一种故障与这种故障的一个征兆构成一条诊断知识。

根据模式的定义<sup>[14]</sup>,将诊断知识模式表示为:

故障 1 征兆 1 的模式 = {故障 1 征兆 1 数值 1, 故障 1 征兆 1 数值 2, …}。

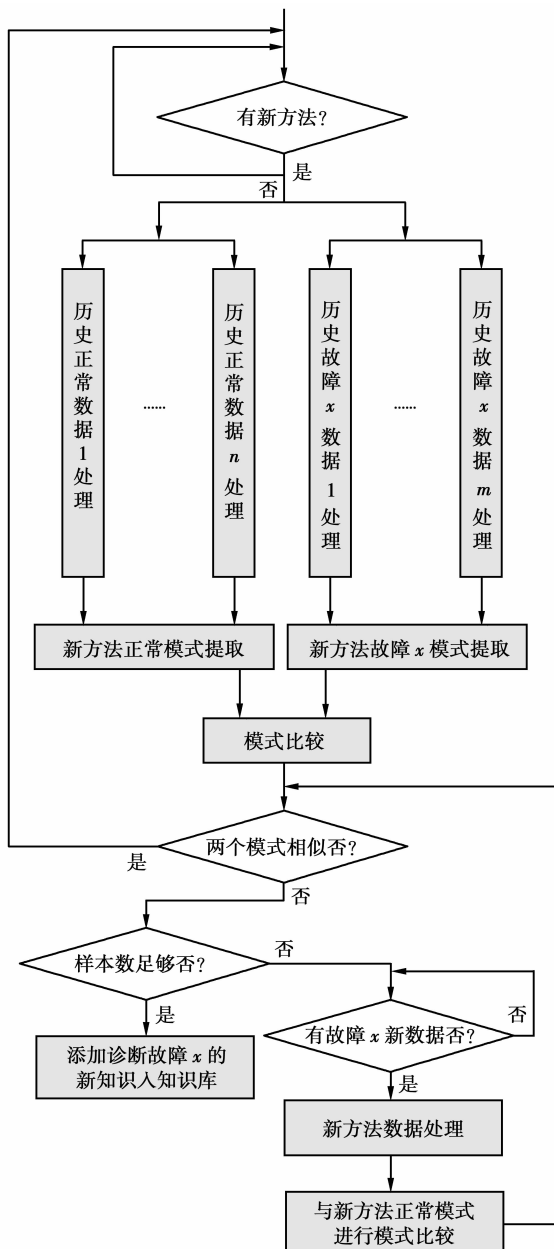


图 3 带确定诊断知识的学习流程

## 2 基于模式比较和更新的诊断知识自学习模型

### 2.1 确定性知识的自学习模型

对于确定性知识的自学习过程是通过不断对设备实际测试数据进行分析,对确定性知识进行模式更新,从而得到与实际设备实际环境相适应的模式的过程。具体过程如图 2(其中输入诊断知识自学习系统的数据由人工根据诊断结果和实际开机检验得到的结论对数据是否为故障数据或者何种故障的数据做出判断)。

### 2.2 待确定知识的自学习模型

对于待确定知识的自学习过程是当有新知识提取方法出现时,总结出新方法模式,通过对新方法故障模式与正常模式的对比、分析,在适当的情况下确定新方法诊断知识模式为确定性知识模式或者放弃新方法模式。具体流程如图 3。

## 3 诊断知识自学习算法基于模式比较和更新的诊断知识自学习模型

### 3.1 诊断知识模式中心的确定

根据模式识别理论,可以将一条诊断知识模式看成是一个聚类,选择一条诊断知识模式的均值作为聚类中心。由于在模式比较时需要样本方差和样本数,因此,将由均值、方差和样本数组成的集合作为模式中心。

### 3.2 模式比较算法

由于设备在不同状态时,其幅值谱概率密度函数不同,即诊断知识模式总体的分布符合不同规律,因此,选用非正态总体假设检验。

在对两个模式进行比较时,每个模式总体的样本数是有限的,即每个模式总体方差可知。

设  $\mu_0$  为正常状态的总体的均值,  $\xi$  为故障  $X$  的总体,均值为  $E\xi$ , 方差为  $D\xi$ , 在显著水平  $\alpha$  下,

$$H_0: E\xi = \mu_0; \quad H_1: E\xi \neq \mu_0. \quad (1)$$

$$\frac{\bar{\xi} - E\xi}{\sqrt{D\xi/n}} \xrightarrow{L} N(0, 1) \quad (n \rightarrow \infty). \quad (2)$$

以  $\frac{\bar{\xi} - E\xi}{\sqrt{D\xi/n}}$  作近似的  $u$  检验, 当  $H_0$  为真, 且  $n$  充分大时, 近似地有

$$P \left\{ \left| \frac{\bar{\xi} - \mu_0}{\sqrt{D\xi/n}} \right| \geq \mu_{1-\alpha/2} \right\} = \alpha. \quad (3)$$

即拒绝域为

$$\Gamma = \left\{ (x_1, x_2, \dots, x_n) : \frac{|\bar{\xi} - \mu_0|}{\sqrt{D\xi n}} \geq \mu_{1-\alpha/2} \right\}. \quad (4)$$

于是,模式比较的法则为:

若  $\frac{|\bar{\xi} - \mu_0|}{\sqrt{D\xi n}} \geq \mu_{1-\alpha/2}$ , 则拒绝  $H_0$ , 认为故障  $X$  的

模式与正常模式有显著差异,即不相似;

若  $\frac{|\bar{\xi} - \mu_0|}{\sqrt{D\xi n}} < \mu_{1-\alpha/2}$ , 则接受  $H_0$ , 认为故障  $X$  的

模式与正常模式有没有显著差异,即相似。

### 3.3 样本量的确定

在故障模式和正常模式进行比较的时候,故障模式中样本数量(即故障数据)越多,则比较的结论越准确。样本数量的确定可以采用样本含量估计的方法。

$$Z = \frac{\bar{\xi} - \mu_0}{\frac{D\xi}{\sqrt{n}}}, \quad (5)$$

$$n = \left( \frac{(z_\alpha + z_\beta) D\xi}{\mu - \mu_0} \right)^2 = \left( \frac{(z_\alpha + z_\beta) D\xi}{\delta} \right)^2. \quad (6)$$

$\bar{\xi}$  为样本均值,  $D\xi$  为样本方差,  $n$  为样本数,  $\delta$  为要求的区分度(即规定有专业意义的差值),  $\beta$  为允许犯 II 类错误的概率(通常所期望的检验效能 power, 用  $1-\beta$  表示。在科研设计中常把  $1-\beta$  定为 0.90 或 0.80。本文取  $1-\beta=0.80$ ),  $Z_\alpha, Z_\beta$  分别是对应于  $\alpha$  和  $\beta$  的  $u$  值(可由  $t$  界值表, 自由度  $\nu=\infty$  一行查出来, 在这里  $\alpha$  为双侧,  $\beta$  只取单侧值)。

### 3.4 模式更新

样本均值的更新

$$\bar{\xi} = \frac{\bar{\xi}n + x_{n+1}}{n+1}. \quad (7)$$

方差的更新

$$D\xi = \sqrt{\frac{D\xi^2 n + x_{n+1}^2}{n+1}}. \quad (8)$$

样本数的更新

$$D\xi = \sqrt{\frac{D\xi^2 n + (x_{n+1} - \bar{\xi})^2}{n+1}}. \quad (9)$$

其中:  $\bar{\xi}$  为样本均值,  $D\xi$  为样本方差,  $n$  为样本数,  $x_{n+1}$  为新增样本值

## 4 实例

用某大型钢铁企业棒材厂 16 V 轧机一轴轴向测点(加速度传感器)在线监测的正常和故障数据的峭度指标进行分析。(事先从知识库中去掉滚动轴承裂纹类故障的峭度指标知识)

用连续运行一个月的实时数据正常状态下的模式均值与方差对比, 如表 1。

表 1 某钢铁公司 2 号轧机减速机  
输入轴垂直方向的部分振动测试峭度指标

时间	第一天	第二天	第三天	...	第三十天
均值	3.35	3.40	3.39	...	3.39
方差	0.09	0.10	0.12	...	0.15
...	...	...	...	...	...

理论上,滚动轴承的正常峭度指标是 3。从此设备一个月运行得到的正常状态模式均值和方差的对比可知,这一设备正常状态的峭度指标约为 3.37, 而且这一值根据运行时间的增加而变化。

对于某次历史外圈裂纹故障数据的分析,对于不同的样本数的均值和方差对比如表 2。

表 2 某钢铁公司 2 号轧机减速机  
输入轴垂直方向的部分振动峭度指标

样本数	40	60	80	100
均值	150.3	155.7	152.4	153.1
方差	3.09	3.10	3.52	4.10
...	...	...	...	...

计算样本值为 69。因此当样本增加到 69 时,查询故障知识库,发现关于滚动轴承裂纹类故障的新知识峭度指标,模式均值为 153.1, 方差为 3.10, 样本量为 69。

从对历史外圈裂纹故障数据的分析可知,该方法可以实现新知识的积累。

## 5 结论

1) 基于模式比较和更新的自学习模型有利于实现智能诊断系统在运行过程中诊断新知识的积累与诊断知识更加贴近实际设备实际情况;

2) 基于非正态分布假设检验的模式比较算法, 能实现待检验方法的故障模式与正常模式的有效比较;

3) 基于样本含量估计的故障模式样本量的确定, 为能有效比较故障模式与正常模式的故障数据量提供了理论依据;

4) 模式更新算法使设备诊断知识与实际情况一致。

## 参考文献:

- [1] LEI Y G, HE Z J, ZI Y Y. Application of an intelligent classification method to mechanical fault diagnosis[J]. Expert Systems with Applications, 2009, 36(6):9941-9948.
- [2] YANG B S, HAN T, AN J L. ART-KOHONEN neural network for fault diagnosis of rotating machinery[J]. Mechanical Systems and Signal Processing, 2004, 18(3):645-657.
- [3] DE MIGUEL L J, FELIPE BLAZQUEZ L. Fuzzy logic-based decision-making for fault diagnosis in a DC motor[J]. Engineering Applications of Artificial Intelligence, 2005, 18(4):423-450.
- [4] OLSSON E, FUNK P, XIONG N. Fault diagnosis in industry using sensor readings and case-based reasoning[J]. Journal of Intelligent and Fuzzy Systems, 2004, 15(1):41-46.
- [5] DIALLO D, BENBOUZID M E H, HAMAD D, et al. Fault detection and diagnosis in an induction machine drive: a pattern recognition approach based on concordia stator mean current vector [J]. IEEE Transactions on Energy Conversion, 2005, 20(3):512-519.
- [6] 李伟. 复杂系统的智能故障诊断技术现状及其发展趋势[J]. 计算机仿真, 2004, 21(10):4-7.  
LI WEI. Advance of intelligent fault diagnosis for complex system and its present system[J]. Computer Simulation, 2004, 21(10):4-7.
- [7] 曾志强, 王俊元. 复杂系统故障诊断的现状和发展趋势[J]. 机械管理开发, 2007(5):78-79.  
ZENG ZHI-QIANG. The Status Quo of Complex System Fault Diagnosis and Developing Trend [J]. Mechanical Management and Development, 2007(5):78-79.
- [8] 杨炜, 高夫燕, 徐晔. 基于粗糙集的注塑成型缺陷智能诊断的专家系统研究[J]. 中国制造业信息化, 2008, 37(3):65-68.  
YANG WEI, GAO FU-YAN, XU YE. Study of expert system in intelligent diagnosis of injection molding limitations based on RS theory[J]. Manufacture Information Engineering of China, 2008, 37(3):65-68.
- [9] 刘海亮, 熊静琪. 基于BP神经网络液压制动故障诊断研究[J]. 微计算机信息, 2007, 23(2):186-219.  
LIU HAI-LIANG, XIONG JING-QI. The research on fault intelligent diagnosis of some hydraulic braking equipment based on BP NN [J]. Control and Automation, 2007, 23(2):186-219.
- [10] MALIKOPOULOS A A, ASSANIS D N, PAPALAMBROS P Y. Real-time self-learning optimization of diesel engine calibration[J]. Journal of Engineering for Gas Turbines and Power, 2009, 131(2):022803-1-022803-7.
- [11] IBRAHIM W R A, MORCOS M M. An adaptive fuzzy self-learning technique for prediction of abnormal operation of electrical systems[J]. IEEE Transactions on Power Delivery, 2006, 21(4):1770-1777.
- [12] LI W, MAIBER P, ENGE H. Self-learning control applied to vibration control of a rotating spindle by piezopusher bearings[J]. Proceedings of the Institution of Mechanical Engineers. Part I: Journal of Systems and Control Engineering, 2004, 218(3):185-196.
- [13] 高文, 张小栓, 傅泽田, 等. 基于OWL的鱼病诊断本体模型[J]. 计算机工程与设计, 2007, 28(19):4770-4773.  
GAO WEN, ZHANG XIAO-SHUAN, FU ZE-TIAN, et al. Fish disease diagnosis ontology model based on OWL [J]. Computer Engineering and Design, 2007, 28(19):4770-4773.
- [14] 马奎斯德萨. 模式识别-原理、方法及应用[M]. 北京: 清华大学出版社, 2002.
- [15] 庄楚强, 何春雄. 应用数理统计基础[M]. 广州: 华南理工大学出版社, 2006.

(编辑 张小强)