

引用格式:李果,马春阳,李伟,等.基于梯次推理获益原则的故障预测方法[J].电光与控制,2019,26(9):107-110. LI G, MA C Y, LI W, et al. A failure prediction method based on echelon inference and benefit principle[J]. Electronics Optics & Control, 2019, 26(9):107-110.

## 基于梯次推理获益原则的故障预测方法

李果<sup>1</sup>, 马春阳<sup>2</sup>, 李伟<sup>3</sup>, 屈重年<sup>1</sup>, 李根全<sup>1</sup>

(1. 南阳师范学院机电工程学院, 河南 南阳 473061;

2. 河南省经济管理学校汽车工程系, 河南 南阳 473000; 3. 空装驻洛阳地区第二军事代表室, 河南 洛阳 471000)

**摘要:** 针对大数据下产生式规则推理效率低下的缺点以及异构信息融合时知识元权重分配困难的问题,以数据和音频两种异构信息为信息源,设计了基于梯次推理获益原则的产生式规则知识表示方法。根据映射特征向量与非映射特征向量知识元灵敏度的不同,采用分组分级的推理策略;采用互信息-信息增益作为获益尺度,根据知识元特征的获益值确定各知识元的权重。以某飞行数据和同步座舱音频为信源模拟了发动机故障实例并进行故障预测,验证了梯次推理和获益原则的产生式规则表示方法具有较高的执行效率、知识元结构强度和较好的推理效果,可以应用于实际工程的故障诊断和预测。

**关键词:** 故障预测; 梯次推理; 获益原则; 互信息; 知识表示

中图分类号: TP391 文献标志码: A doi:10.3969/j.issn.1671-637X.2019.09.022

## A Failure Prediction Method Based on Echelon Inference and Benefit Principle

LI Guo<sup>1</sup>, MA Chun-yang<sup>2</sup>, LI Wei<sup>3</sup>, QU Chong-nian<sup>1</sup>, LI Gen-quan<sup>1</sup>

(1. College of Mechanical & Electrical Engineering, Nanyang Normal University, Nanyang 473061, China;

2. Automobile Engineering Department, Henan Economic Management School, Nanyang 473000, China;

3. The Second Military Representative Office of Air Force Equipment Department in Luoyang District, Luoyang 471000, China)

**Abstract:** To solve the problems of poor efficiency of production rule in big data and hard weight distribution in the case of heterogeneous information fusion, a knowledge representation method of production rule based on echelon inference and benefit principle was designed taking the data and voice as the information source for fusion. A grouping & ranking inference strategy was proposed on the basis of different sensitivity between mapped knowledge elements and unmapped knowledge elements. The weight of knowledge element was specified by the benefit value obtained from mutual information and information gain. We simulated the engine fault by using the flight data and the synchronous cockpit voice as the information source for making failure prediction. The result shows that our knowledge representation method can provide higher efficiency, stronger intensity and more accurate diagnosis result, which can be used in practical fault diagnosis and fault prediction.

**Key words:** failure prediction; echelon inference; benefit principle; mutual information; knowledge representation

### 0 引言

目前,在多源信息融合的故障预测领域中,往往

需要多种不同的信息源配合使用才能获得对预测对象较为完整的信息。传统的方法只是以多个信息源数据的算术平均数作为测量信息,采用加权融合的方法处理异构信息,虽然具有一定的抗干扰能力,但难点在于权重因子的选择<sup>[1]</sup>。当前故障诊断和预测工程应用时,推理过程多采用产生式规则进行全数据顺序搜索和模式匹配,随着规则数目的增加,效率低的缺点越来越突出<sup>[2]</sup>。因此,推理方式的选择在一定程度上决定

收稿日期:2019-04-08

修回日期:2019-08-15

基金项目:国家自然科学基金(61306007);河南省科技攻关项目(172102310682,172102210095);南阳师范学院博士专项基金(2018ZX024)

作者简介:李果(1980—),男,河南南阳人,博士,高工,研究方向为信息融合技术。

了预测效率。同时,产生式表达式具有良好的自然性、灵活性、统一性和模块性,但也有推理效率低下、不直观等缺点,对于多源异构信源的特征表示,在特征知识元强度、知识表示的前件信息容量方面都有不足<sup>[3]</sup>。

本文设计了梯次推理算法,将互信息-信息增益作为权重因子进行融合,在推理效率和融合结果的精确性两方面优化知识表示方法,利用同一个预测对象的数据和音频两种异构信息进行故障预测。

## 1 梯次推理算法设计

### 1.1 推理方式分析

目前工程应用中多采用顺序推理的方式<sup>[4]</sup>,按照设定的步长(一般以时间或数据帧为单位)进行全范围扫描,最终得出相应的推理结果。这种方式最大的好处是能够充分利用信息,不会产生遗漏的问题,缺点是随着信息量的不断增加,效率逐渐降低<sup>[5]</sup>。

数据和音频信息是一种相互印证的异构信息类型,两类信息之间不存在一一对应的映射关系,而存在信息不对等性。因此,在保证信息可靠性的前提下,首先应考虑的是灵敏度高的特征量,从实际应用效果来看,支持某一状态模式的特征量越多,扫描结果也就越可靠<sup>[6]</sup>。在推理时,需要对特征量重构后进行扫描。

### 1.2 梯次推理流程

#### 1.2.1 特征量重组

##### 1) 分组。

两种信息类型之间存在一一映射关系的各特征量之间构成映射知识元向量组,不存在对应关系的特征量归为非映射知识元向量组。

##### 2) 分级。

映射知识元向量组内按照信息类型数量、特征量数量的顺序排序,非映射知识元向量组按照开关量、模拟量、音频特征量的顺序排序。这样的分组分级归类策略保证了信息量最大的特征优先组合,也保证了可靠性最高信源特征优先级最高。

#### 1.2.2 扫描的算法设计

##### 1) 动态矩阵时间窗口长度 $t$ 和滑动步长 $\alpha$ 的确定。

确定状态模式所需的最短时间,设定动态状态矩阵的滑动步长  $\alpha$ : 如果  $\alpha > t$ ,则将步长  $\alpha$  作  $t$  等分,分步长为  $\alpha/t$ ,窗口按照分步长滑动,各分步长顺序扫描的结果累加,得到一次扫描结果;如果  $\alpha < t$ ,则按照步长  $\alpha$  滑动,动态矩阵之间交叠  $t - \alpha$  个时间单位;如果  $\alpha = t$ ,按照顺序扫描,窗口之间无叠加。

##### 2) 推理过程。

① 设定任务,确定对应的状态模式。

② 扫描映射知识元向量组内的特征量,扫描顺序

为离散量-模拟量-音频特征量,如发现可用信息,则转向④;如未发现可用信息,则继续执行③。

③ 扫描非映射知识元向量组内的特征量,扫描顺序为开关量-模拟量-音频特征量,如发现有用信息,则继续执行④;如未发现可用信息,则返回①进行任务内容的调整。

④ 根据产生式规则推出结果组,并依据支持知识元数量赋予推论概率。

## 2 获益原则融合算法设计

为了寻求两种信源融合的最大信息量,引入互信息的概念,作为特征判别能力的量度,并将建树的方法嵌入一个外壳之中<sup>[7]</sup>。本文采用获益的原则来描述融合过程中的协同性和互补性,将每次学习中获得的互信息-信息增益作为融合机制的获益尺度。

### 2.1 获益计算

模式  $C_b (b=1, \dots, m)$  的发生概率  $P(C_b)$  组成的信源数学模型为

$$[C, P] = \begin{bmatrix} C_1, \dots, C_m \\ P(C_1), \dots, P(C_m) \end{bmatrix} \quad (1)$$

模式  $C_b$  的自信息为

$$I(C_b) = \ln \frac{1}{P(C_b)} \quad (2)$$

根据信息熵的定义及式(2)推导出

$$H(C) = \sum_{b=1}^m P(C_b) I(C_b) = \sum_{b=1}^m P(C_b) \ln \frac{1}{P(C_b)} \quad (3)$$

根据后验熵定义及式(3),信息源某一个特征参量  $U_i (1 \leq i \leq n, n$  为信息源的特征参量数) 的值为  $u_d$ , 其后验熵为

$$H_i(C | u_d) = \sum_{b=1}^m P(C_b | u_d) \ln \frac{1}{P(C_b | u_d)} \quad (4)$$

根据式(4)得到对应条件熵为

$$H_i(C | u) = \sum_{d=1}^n P(u_d) \sum_{b=1}^m P(C_b | u_d) \ln \frac{1}{P(C_b | u_d)} \quad (5)$$

根据互信息定义和式(3)、式(5)得到  $U_i$  平均互信息为

$$I_i(C | u) = H(C) - H_i(C | u) = \sum_{b=1}^m P(C_b) \ln \frac{1}{P(C_b)} - \sum_{d=1}^n P(u_d) \sum_{b=1}^m P(C_b | u_d) \ln \frac{1}{P(C_b | u_d)} \quad (6)$$

则  $I_i(C | u)$  表示融合结果从特征参量  $U_i$  处的获益。

### 2.2 权重训练算法

权重训练算法流程如图 1 所示。

1) 在当前异类信息中选取所有的特征参量  $U_i$ , 根据式(1)~式(6)求出获益集  $I = \{I_i(C | u) | 1 \leq i \leq n_s\}$ , 其中,  $n_s$  为各个信源特征参量数目之和。

2) 对所有特征参量的平均互信息按照从大到小进行排序,并按照信源种类进行标识,即  $I_{1,A_i}, I_{2,A_i}, \dots, I_{n_s,A_i}$ , 特征参量下标  $i(1 \leq i \leq n_s)$  表示本特征参量的排序,下标  $A_j$  表示该特征参量所示的信源类别。

3) 统计各信源的每个特征参量的获益值排序情况,并计算名次之和  $num(A_i)$ 。

4) 各信源进行权值分配,即

$$w_i = 1 - num(A_i) / \sum_{j=1}^{n_s} j \quad (7)$$

5) 根据话语音中的标注调整融合结果,如果话语音中出现了反映该状态的语音信息,则对融合的结果人为提高融合结果的概率,概率的提高幅度根据专家经验确定;如果没有出现相关话语信息,以当前融合结果作为最终融合结果进行下一步相应分析处理。

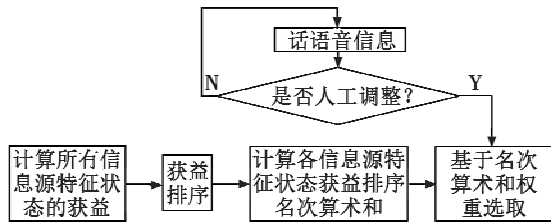


图1 获益原则的知识表示流程  
Fig.1 Knowledge representation process of benefit principle

### 2.3 知识表示方法设计

在进行知识表示时,首先对异构信息进行梯次推理扫描,得到初步分析结果,根据规则的各个特征参量和其对应的推理结论概率计算各个特征参量的融合权重,最后以各信源特征的权重和作为该信源融合知识表示时的分配权重。

## 3 实例分析

以某飞机的飞行数据和同步的座舱音频信息为例,验证知识表示方法的可行性。

预设发动机超转的状态模式,提取飞行数据中发动机高压转速和低压转速两个模拟量、发动机超转警告一个开关量,同时监控舱音信息中发动机噪声的短时平均过零率和谱流量。在一段3500 s的正常飞行数据和发动机噪声中,分别在第380 s、第2370 s和第3200 s时插入预设的状态模式,这样就模拟了飞行开始、飞行过程中和飞行快结束3个时间发生发动机超转的状态模式,分别命名为模式1、模式2和模式3。原始特征样本如图2所示。

动态状态矩阵时间窗口长度为3 s,滑动步长为7 s,其中飞行数据特征逐秒显示。舱音原始特征以帧为

单位,每帧宽度为30 ms,相邻帧之间重叠10 ms。根据时间对准策略,将飞行数据和舱音两类信息源对准,按照产生式规则进行判定。分别采用梯次推理和顺序推理方式,选取扫描时间作为判别指标,统计两种方式的扫描时间之比。

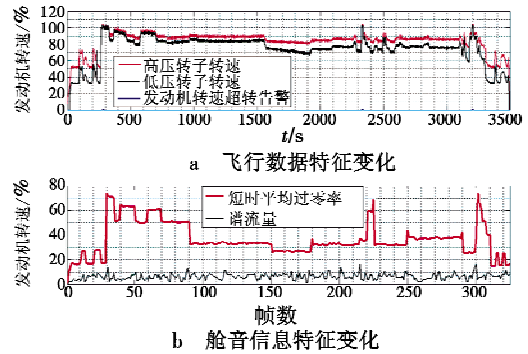


图2 原始样本图

Fig.2 The original sample

在同一台电脑上进行仿真实验,观察在同样条件下,梯次推理和顺序推理两种方式速度的快慢。扫描结果如表1所示。

表1 3种模式推理结果表

Table 1 The reasoning result of the three modes

	顺序推理			梯次推理		
	模式1	模式2	模式3	模式1	模式2	模式3
扫描时间/s	0.031	0.297	0.422	0.016	0.031	0.047

由表1可知,在同样准确地扫描到3种状态模式的情况下,梯次推理和顺序推理两种方式所用的扫描时间存在较大差异。在3种模式下,两种方式的时间之比分别为1.9375:1,9.58:1,8.97:1。

为了验证该扫描方式在大样本下的扫描速度,特征量选择与相关设置与上述实验一致。分别模拟10组、50组、100组、1000组和10000组3500 s的正常样本。在2300 s时模拟一个发动机超转的故障。分析结果如表2所示。

表2 两种推理方式结果表

Table 2 The reasoning result of the two methods

样本数量	扫描方法	扫描时间/s	时间比
10	顺序	5.083	4.94 : 1
	梯次	1.029	
50	顺序	24.894	24.08 : 1
	梯次	1.034	
100	顺序	36.092	29.15 : 1
	梯次	1.238	
1000	顺序	325.340	37.96 : 1
	梯次	8.570	
10000	顺序	4103.560	67.29 : 1
	梯次	60.980	

由表 2 可知,梯次推理方式随着样本的增大,扫描时间的优势较之顺序扫描方式十分明显。

实验表明,梯次推理的扫描策略实质上实现了特征的分阶段维数约减,优先利用了误差项最小的特征量,间接优化了特征量的组合效率,直接导致了推理时间效率的提高。

设计故障算例如表 3 所示。

表 3 故障征兆描述

Table 3 The failure symptom

故障类型	故障征兆描述
A	稳态转速低于正常范围,发动机噪声逐渐增大、发动机振动较大并逐渐接近阈值,滑油压力出现明显波动
B	稳态转速波动较大,振动值呈周期性的突变峰值,调变频率为频率的两倍,滑油压力出现明显波动,发动机噪声呈现周期性跳变
C	发动机噪声出现类似“金属撞击”声,稳态转速出现明显波动

提取 8 次故障 A、10 次故障 B 和 12 次故障 C 以及故障发生前 5 个架次的特征数据,分别利用飞行数据、舱音信息和两种信息的融合信息基于产生式规则按照 5 s 的步长进行故障预测,预测结果如表 4 所示。

表 4 故障预测结果表

Table 4 The predicting outcomes

信息类型	故障类型	实际发生次数	预测正确次数	误判次数	总准确率/%	总误报率/%
数据	A	8	6	2	70	26.6
	B	10	8	2		
	C	12	7	4		
舱音信息	A	8	2	2	40	23.3
	B	10	8	4		
	C	12	2	1		
飞行数据舱音信息	A	8	7	1	73.3	13.3
	B	10	9	1		
	C	12	6	2		

由表 4 可知,按获益原则融合的预测结果与飞行数据预测结果相比,在准确率相差不大的情况下,大幅度降低了误报率,这是由于飞行数据和舱音信息特征相互补充,减小了不确定性。由于舱音信息对于故障 A、C 的敏感性较低,而对故障 B 较为敏感,造成了较低的预测准确率和误报率。获益原则融合知识的预测误报率略低于舱音信息,但准确率远高于舱音信息,这是因为飞行数据与舱音特征信息相互印证,提高了预测准确率。

图 3 所示为获益原则知识表示融合的预测准确率、误报率和步数之间的变化关系。

由图 3 可知,随着预测步数的增加,误报率有所上升,准确率有所下降,因此在利用获益原则融合的知识表示方法进行预测时,应该根据样本结果,选择适宜的步数以保证结果的准确率。

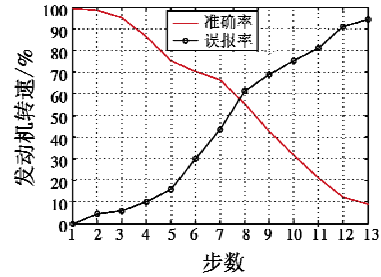


图 3 预测准确率和误报率变化曲线

Fig. 3 The curves of accuracy rate & false alarm rate

### 4 结束语

本文针对异构信息设计了梯次推理算法,利用数据和音频异构信息融合进行知识表示,利用互信息-信息增益尺度设计了基于获益原则的权重确定方法。

基于飞行数据和座舱音频信息样本进行了梯次推理和获益原则的权重提取,实验结果证明,梯次推理较之顺序推理有着明显的时间优势和较好的知识元结构强度,提高了故障预测准确率,降低了误报率。

### 参考文献

- [1] 杨菁华,熊智,刘建业. 基于异类传感器观测信息融合的 UKF 算法[J]. 传感器与微系统,2019,38(3):131-133.
- [2] 文天柱,许爱强,汪定国. 故障诊断专家系统的可拓知识表示和匹配研究[J]. 计算机测量与控制,2014,22(6):213-217.
- [3] 董泽,闫来清. 基于互信息和多尺度小波核偏最小二乘的 SCR 脱硝系统预测模型[J]. 动力工程学报,2019,39(1):50-58.
- [4] 杨凯,杨晓元,马建峰. 无线传感器网络中一种基于 D-S 证据理论的监测机制[J]. 四川大学学报:工程科学版,2016,48(2):118-124.
- [5] 谭青,向阳辉. 加权证据理论信息融合方法在故障诊断中的应用[J]. 振动与冲击,2008,27(4):112-116.
- [6] 刘培奇,李增智,赵银亮. 扩展产生式规则知识表示方法[J]. 西安交通大学学报,2004,38(6):234-238.
- [7] 韩敏,刘晓欣. 基于 Copula 熵的互信息估计方法[J]. 控制理论与应用,2013,30(7):542-548.