基于模糊关联分类器的民机超限事件诊断方法

高小霞 霍纬纲 冯兴杰

(中国民航大学 计算机科学与技术学院, 天津 300300)

摘 要: 现有的民用飞机超限事件智能诊断模型大多属于“黑盒”模型,不利于分析超限事件发生的原因.为此提出了一种基于模糊关联分类器(FAC,Fuzzy Associative Classifier)的民用飞机超限事件诊断方法.该方法抽取发生超限事件时对应的 QAR(Quick Access Recorder)参数快照取值,采用模糊 C 均值(FCM,Fuzzy C-Means)聚类算法对抽取的 QAR 参数取值模糊预处理,然后基于 Apriori 算法生成模糊关联分类规则库,并由遗传算法对其进行裁剪,结合模糊分类推理方法形成 FAC.采用 B737-800 实际样本数据进行了验证.实验结果表明,所提出的 FAC 能有效诊断超限事件,FAC 识别超限事件的错误率与最小二乘支持向量机(LS-SVM,Least Squares Support Vector Machine)模型相当,但其解释性方面优于 LS-SVM.

关 键 词: 飞行品质监控; 模糊关联分类器; 超限事件; 遗传算法; 诊断模型

中图分类号: TP 18

文献标识码: A **文 章 编 号**: 1001-5965(2014)10-1366-06

Civil aircraft's exceedance event diagnosis method based on fuzzy associative classifier

Gao Xiaoxia Huo Weigang Feng Xingjie

(School of Computer Science and Technology, Civil Aviation University of China, Tianjin 300300, China)

Abstract: Most of current civil aircraft's exceedance event intelligent diagnosis models are "black box" model, which donot contribute to analyze the occurrence of civil aircraft's exceedance event. In order to overcome these shortcomings, a civil aircraft's exceedance event diagnosis method based on fuzzy associative classifier (FAC) was proposed. First, quick access recorder's (QAR's) parameters value when exceedance event occured was extracted. Fuzzy C-means (FCM) cluster algorithm was adopted to preprocess extracted QAR's parameters value. Then, the library of fuzzy associative classification rule (FACR) was generated by Apriori algorithm. Genetic algorithm was used to prune the library of FACR. Finally, fuzzy classification reasoning method was integrated to build FAC. The FAC was verified with sample data generated by B737-800. Experiment results show that the FAC can diagnose exceedance event effectively, and its classification error rate is equivalent to least squares support vector machine (LS-SVM), but its interpretability is superior to LS-SVM.

Key words: flight operations quality assurance; fuzzy associative classifier; exceedance event; genetic algorithm; diagnosis model

飞行品质监控(FOQA,Flight Operations Quality Assurance)是促进民航业持续安全的一项基础性工作,具有事先预防航空事故的重要功效.FOQA利用机载 QAR(Quick Access Recorder)记录的飞行数据,依据事先设定的监控项目和标准从中发现各种超限事件.目前我国民航 FOQA 工

收稿日期: 2013-11-18; **网络出版时间**: 2014-03-24 10:27; **DOI**: 10.13700/j.bh.1001-5965.2013.0656

网络出版地址: www.cnki.net/kcms/doi/10.13700/j.bh.1001-5965.2013.0656.html

基金项目: 国家自然科学基金资助项目(61301245,61201414,61202018); 国家自然科学基金委员会与中国民用航空局联合基金资助项目(U1233113,61179063); 中央高校基本科研业务费中国民航大学专项资助项目(ZXH2012N001)

作者简介: 高小霞(1980 -),女,河北临城人,实验师,xxgao@cauc.edu.cn.

作方式存在如下明显不足:①监控标准主要依据单一飞行参数以“[下限,上限]”形式确定,以这种方式进行超限事件检测很容易由于数据噪声而导致误判,实际上在某一超限事件发生时,可能会有若干个 QAR 参数与之相关联,如果能综合运用多个 QAR 参数进行超限事件诊断,便可以有效减少如上问题的发生;②QAR 数据利用率低,没有对其进行深入挖掘分析,查找其蕴含的安全规律.

针对如何开展 QAR 数据深度分析以提高飞行安全,国内外学者开展了一些研究工作.美国国家航空航天局(NASA)的 AMES 研究中心提出了一系列进行飞行品质异常检测的方法.这些方法包括基于 K-近邻的 Orca[1]、基于聚类的 IMS[2]和 SequenceMiner[3]、基于核函数的 MKAD[4].Smart 等人[5]基于支持向量机(SVM,Support Vector Machine)进行了飞行品质异常检测研究,该方法分为 2 个阶段:第 1 阶段检测某个航班在降落阶段各个高度的异常;第 2 阶段根据第 1 阶段 SVM 的输出比较航班之间的异常,从而发现最异常的航班,还应用特征选择 F-score 技术寻找与飞行异常相关的关键飞行参数.在国内,黄圣国等人[6-7]基于 SVM 构建了民用飞机的重着陆智能诊断模型,曹海鹏等人[8]基于神经网络对民用飞机重着陆事件做了研究.这些方法均通过物理机理分析获得与重着陆相关的飞行参数,然后从 QAR 数据集中截取数据构成训练样本进行建模.祁明亮等人[9]利用优选法和数理统计法,选取飞机下降率、俯仰角等飞行员综合操作指标,确定可能触发某一超限事件的高风险区域.上述与本文相关的基于 SVM 和神经网络的方法都属“黑盒”建模法,这些方法虽然能够识别超限事件,但是不能较好地解释超限事件发生的原因,即模型的解释性较差.

模糊关联分类器(FAC,Fuzzy Associative Classifier)是数据挖掘研究领域重要的分类方法之一[10-12],是在高维数据集上构建模糊分类器的一种有效方法[13],其最大特点在于完全从数据出发,发现有趣的关联,挖掘数据集中的多个变量与样本类别之间的规律性,以其较高的分类精度和较好的解释性受到了研究关注.为此,本文采用 Apriori 算法和遗传算法构建了基于模糊关联分类器的超限事件诊断模型,该模型扩展了超限事件的判定指标,实现了多因素的超限事件诊断,真实 QAR 数据集上的实验表明了本文方法的有效性.

1　模糊关联分类器的相关定义

设有数据集 $\boldsymbol{D}=[\boldsymbol{X}\ \boldsymbol{y}]_{N\times(n+1)}$,其中 $\boldsymbol{X}=[x_{k,l}]_{N\times n}$ 为在数量属性 $z_1,z_2,\cdots,z_n$ 上的 N 个取值,$\boldsymbol{y}=[y_k]_{N\times 1}$ 为在类别属性 z_{n+1} 上的取值,$\boldsymbol{D}$ 中第 $k(1\leqslant k\leqslant N)$ 个样本 x_k 在数量属性 $z_i(1\leqslant i\leqslant n)$ 上的模糊隶属度值表示为:$\boldsymbol{u}_{k,i}=[A_{i,1}(x_{k,i}),A_{i,2}(x_{k,i}),\cdots,A_{i,q_i}(x_{k,i})]$,其中 $A_{i,1},A_{i,2},\cdots,A_{i,q_i}$ 为数量属性 z_i 对应的隶属度函数,q_i 为属性 z_i 的隶属度函数的个数.数据集 $\boldsymbol{D}$ 可由各个数量属性上的隶属度函数转化为取值在[0,1]之间的模糊数据集 $D_{\mathrm{f}}=\{t_1,t_2,\cdots,t_N\}$,其中 $\boldsymbol{t}_k=[\boldsymbol{u}_{k,1},\boldsymbol{u}_{k,2},\cdots,\boldsymbol{u}_{k,n},y_k](1\leqslant k\leqslant N)$.定义 $\langle z_i:A_{i,j}\rangle$ 为模糊项,模糊项集定义为

$$\langle Z:A\rangle=\{\langle z_{i_1},A_{i_1,j}\rangle\cup\langle z_{i_2},A_{i_2,j}\rangle\cup\cdots\cup\langle z_{i_q},A_{i_q,j}\rangle\}\qquad q\leqslant n+1$$

定义 1　模糊项集 $\langle Z:A\rangle$ 的支持度定义为

$$F_{\mathrm{s}}(Z:A)=\frac{\sum_{k=1}^{N}\prod_{\langle z_i:A_{i,j}\rangle\in\langle Z:A\rangle}t_k(z_i)}{N}$$

若 $F_{\mathrm{s}}(\langle Z:A\rangle)$ 大于指定的最小支持度,则称 $\langle Z:A\rangle$ 为模糊频繁项集.其中 $t_k(z_i)$ 为第 k 个数据在属性 z_i 对应的隶属度函数 $A_{i,j}$ 上的取值;$\prod$ 表示乘法运算.

定义 2　若模糊关联规则 $\langle Z:A\rangle\Rightarrow\langle\boldsymbol{y}:C\rangle$ 的后件只含类别属性,则称为模糊关联分类规则(FACR,Fuzzy Associative Classification Rule),其置信度定义为

$$F_{\mathrm{c}}(\langle Z:A\rangle\Rightarrow\langle\boldsymbol{y}:C\rangle)=\frac{F_{\mathrm{s}}(\langle Z:A\rangle\cup\langle\boldsymbol{y}:C\rangle)}{F_{\mathrm{s}}(\langle Z:A\rangle)}$$

式中 C 为样本的类别集合.

定义 3　模糊关联分类规则的模糊相关度定义为

$$F_{\mathrm{corr}}(\langle Z:A\rangle\Rightarrow\langle\boldsymbol{y}:C\rangle)=\frac{F_{\mathrm{s}}(\langle Z:A\rangle\cup\langle\boldsymbol{y}:C\rangle)-F_{\mathrm{s}}(\langle\boldsymbol{y}:C\rangle)\times F_{\mathrm{s}}(\langle Z:A\rangle)}{\sqrt{F_{\mathrm{s}}(\langle Z:A\rangle)\times(1-F_{\mathrm{s}}(\langle Z:A\rangle))\times F_{\mathrm{s}}(\langle\boldsymbol{y}:C\rangle)\times(1-F_{\mathrm{s}}(\langle\boldsymbol{y}:C\rangle))}}$$

F_{corr} 取值范围为[-1,1],规则的模糊相关度越接近 1,模糊关联分类规则的质量越高.

定义 4　模糊关联分类器的解释性:文中以分类器包含的 FACR 数目与 FACR 前件中包含的模糊项总数为衡量指标,FACR 数目和总的模糊项数目越少,模糊关联分类器的解释性越好.

2 基于模糊关联分类器的超限事件诊断建模

基于模糊关联分类器的超限事件诊断建模流程如图 1 所示. 首先从经译码、格式化后的 QAR 数据中,抽取发生超限事件时与飞行品质相关的 QAR 参数快照取值构成训练和测试数据集,采用模糊 C 均值(FCM,Fuzzy C-Means)聚类算法对数据集模糊预处理. 然后基于 Apriori 算法生成模糊频繁项集,由模糊频繁项生成模糊相关度大于 0 的 FACR 构成分类规则库. 为使模糊关联分类器具有较好的解释性,最后基于遗传算法裁剪 FACR 规则库,裁剪后的分类规则库结合模糊分类推理引擎形成超限事件诊断模型.

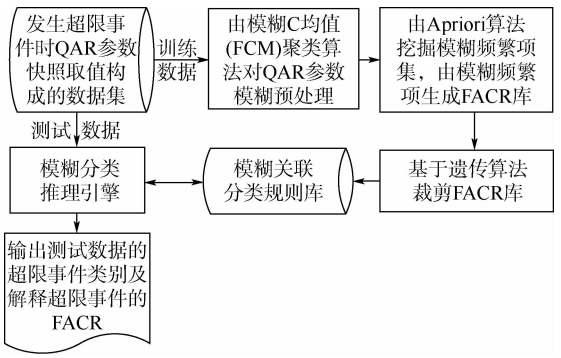

图 1 基于模糊关联分类器的超限事件诊断建模流程图

Fig. 1 Flow chart of building exceedance event diagnosis model based on fuzzy associative classifier

挖掘模糊频繁项以及生成 FACR 的具体算法描述如下.

输入:经模糊预处理后的数据集 D_{f},最小模糊支持度计数 m_{sup};

输出:D_{f} 中的所有模糊频繁项集 L;

扫描 D_{f} 生成 1-模糊频繁项集,记为 L_1;

For ($k=2; L_k \neq \varnothing; k++$) do

C_k = apriori-Gen(L_{k-1});/ * 由($k-1$)-模糊频繁项集生成候选模糊项集 C_k * /

For each $t \in D_{\mathrm{f}}$ do/ * 扫描 D_{f},求 C_k 中模糊项集的支持度计数 * /

C_t = subset(C_k, t);/ * 求事务 t 所能包含 C_k 中的模糊项集 * /

For each 模糊项 $c \in C_t$ do

$c_{\mathrm{count}} = c_{\mathrm{count}} + T(u_{1t}, u_{2t}, \cdots, u_{kt})$;

/ * 计算模糊项 c 的支持度计数 c_{count},T 为 t-模算子,文中取乘法运算,u_{it} 表示 c 中的第 i ($1 \leqslant i \leqslant k$)项在事务 t 上的取值; * /

End for

End for

$L_k = \{c \in C_k \mid c_{\mathrm{count}} \geqslant m_{\mathrm{sup}}\}$

End for

所有模糊频繁项集 $L = L_1 \cup L_2 \cup \cdots \cup L_k$;

上述算法中由($k-1$)-模糊频繁项集生成候选模糊项集 C_k 的 apriori-Gen(L_{k-1})具体步骤如下:

For each 模糊项 $l_1 \in L_{k-1}$

For each 模糊项 $l_2 \in L_{k-1}$

If $(l_1[1] = l_2[1]) \wedge (l_1[2] = l_2[2]) \wedge \cdots \wedge (l_1[k-2] = l_2[k-2]) \wedge l_1[k-1]$与 $l_2[k-1]$ 为两个不同数量性对应的模糊项 Then

由 l_1, l_2 生成新的模糊项 $c = l_1[1] l_2[2] \cdots l_1[k-2] l_1[k-1] l_2[k-1]$;

If c 的每个包含 $k-1$ 模糊项的子集 s 都属于 L_{k-1} Then

新生成的模糊项 c 加入 C_k;

End if

End if

End for

End for

ProcMineFuzzyClassifyRule (L)

/ * 算法的输入为除 1-模糊频繁项的模糊频繁项集 L * /

For each $l \in L$

For each $l' \subset l$

If $l - l'$中只含类别属性并且 $F_{\mathrm{corr}}(l' \Rightarrow l - l') > 0$ then

$R = R \cup (l' \Rightarrow l - l')$;/ * 集合 R 为初始模糊关联分类规则集 * /

End if

End for

End for

2.1 基于遗传算法的模糊关联分类规则库裁剪方法

遗传进化种群的个体采用二进制编码方式,设初始分类规则库有 M 个 FACR,则群体中的每个个体可表示为长度 M 的二进制串:$s_1 s_2 \cdots s_M$,每个 s_i($1 \leqslant i \leqslant M$)取值 0 或 1.0 表示该基因位对应的 FACR 不在构成 FAC 的规则库中,1 则表示对应的 FACR 包含在 FAC 的规则库中. 种群个体间的交叉方式采用 HUX(Half Uniform Crossover)交叉算子,为尽量减少 FAC 中分类规则的数量,采用有偏向的变异算子,即个体中的每个基因位由 0 至 1 的变异概率高于 1 至 0 的变异概率,交叉和变异算子的具体描述见文献[14]. 模糊关联分类器的准确率和解释性是两个相互制约的优化目

标,文中进化群体中个体 S 的适应度函数设计如式(1):

$$f(S)=\frac{N_{\text{right}}}{N_{\text{rule}}+N_{\text{fitem}}} \tag{1}$$

式中,N_{right}为被个体 S 表示的模糊关联分类规则库正确分类的训练样本数目;N_{rule}和 N_{fitem}分别为个体 S 表示的模糊关联分类规则库中分类规则数目和所有规则前件中包含的模糊项数目总和. 裁剪模糊关联分类规则库的具体算法步骤为:对初始模糊关联分类规则库进行二进制编码,初始化进化种群;由式(1)计算进化种群中每个个体的适应度值,采用比例选择算子生成下一代进化种群,对新种群实施 HUX 交叉算子和有偏向的变异算子;若达到预先设定的最大迭代次数,则输出进化种群中最优个体表示的模糊关联分类规则库,否则继续上述的遗传进化过程.

2.2　模糊分类推理方法

图 1 中的模糊分类推理引擎使用的分类推理方法描述如下[15]:设由遗传算法裁剪后的模糊关联分类规则库由 R 条规则组成,待分类的样本 x_k 对分类模型中的第 j 条规则的激励强度按式(2)计算:

$$\beta_j(x_k)=\prod_i \boldsymbol{t}_k(z_i) \qquad \langle z_i:A_{i,j}\rangle \in \langle Z:A\rangle \tag{2}$$

设训练样本集中的样本有 s 个类别,表示为集合 $C=\{c_1,c_2,\cdots,c_s\}$,对 s 个样本类别设置积分向量 $\boldsymbol{V}=[v_1,v_2,\cdots,v_s]$,初始时每个 $v_i(1\leqslant i\leqslant s)$ 的值为 0. 若分类规则库中的第 j 条规则记为

$$r_j:\langle Z:A\rangle\Rightarrow\langle \boldsymbol{y}:C(r_j)\rangle \tag{3}$$

式中 $C(r_j)$ 为样本类别集合 C 中的第 p 类 c_p,则由规则 r_j 积分向量 $\boldsymbol{V}$ 可更新为

$$\boldsymbol{V}=[v_1,\cdots,v_{p-1},v_p+(\beta_j(x_k)\times F_{c_j}),v_{p+1},\cdots,v_s] \tag{4}$$

由式(4)可知积分向量 $\boldsymbol{V}$ 中第 p 个分量值增加了样本 x_k 对第 j 条规则的激励度 $\beta_j(x_k)$ 与规则置信度 F_{c_j}的乘积,按此方法,对于分类规则库中的所有规则,积分向量 $\boldsymbol{V}$ 中的每个分量值可按式(5)计算:

$$v_{C(r_j)}=\sum_j \beta_j(x_k)\times F_{c_j} \qquad j=1,2,\cdots,R \tag{5}$$

待分类样本 x_k 的类别为积分向量 $\boldsymbol{V}$ 中积分值最大的分量对应的类别,即 $y_k=\arg\max(\boldsymbol{V})$.

3　实际应用及结果分析

实验数据来自某航空公司 B737-800 的机载 QAR 数据集,从中抽取发生“50 ft 至接地距离长”和“拉平着陆坡度大”两类三级超限事件时各 200 个航班对应的 QAR 参数快照取值. 根据领域工程师经验知识,取与所选超限事件相关的 15 个 QAR 参数,对每个参数取值采用 Matlab 中的 FCM 聚类算法模糊预处理,聚类数目均为 3,隶属度矩阵的加权指数 m 取值为 2,对应为低、中、高 3 个模糊项. 选取的 15 个参数及其取值范围、聚类中心描述如表 1 所示.

表 1　本文采用的 QAR 参数及 FCM 聚类中心描述

Table 1　Parameters of QAR and description of FCM cluster center

参数	取值范围	聚类中心		
		低	中	高
磁航向/(°)	[0,363]	34.47	194.19	333.15
仰角/(°)	[-2.1,5.3]	0.59	2.37	4.01
坡度/(°)	[-9.6,7.7]	-5.09	-0.26	5.12
大气总温/℃	[-11,34]	2.33	13.12	23.59
N1/(% r/min)	[26,68.9]	31.30	43.63	57.29
全重/lb	[55 280,143 920]	60 345	112 000	133 000
方向舵位置/(°)	[-10,10.6]	-5.42	-1.01	1.88
地速/kn	[101,169]	119.83	135.52	148.72
左副翼位置/(°)	[-10.9,13.8]	-4.40	2.18	7.79
右副翼位置/(°)	[-12.5,11.1]	-5.93	0.77	6.58
左升降舵位置/(°)	[-9.5,6.2]	-5.40	-2.40	1.97
右升降舵位置/(°)	[-9.5,6]	-5.31	-2.37	1.91
风速/kn	[0,41]	4.62	15.40	28.94
风向/(°)	[-176.8,345.8]	-134.08	0.98	137.71
压力高度/ft	[-576,6 844]	63.28	2 045.70	6 509.20

注:1 lb = 0.453 592 37 kg; 1 ft = 0.304 8 m.

将 400 组数据采用分层抽样的方法随机划分为 3 个部分,其中 2 部分作为训练集,剩余部分作为测试集. 裁剪模糊关联分类规则库的遗传算法参数设置如表 2 所示. 实验环境为 WinXP, CPU 2.93 GHz,2.0 GB 内存,Visual C + +6.0.

表 2　遗传算法参数设置

Table 2　Parameter value of genetic algorithm

参数	数值
进化种群大小	200
进化迭代次数	100
交叉概率	0.9
变异概率(1→0)	0.1
变异概率(0→1)	0.001

表 3 列出了本文的 FAC 在不同最小支持度取值下生成的初始分类规则数目、经遗传算法裁剪后分类规则库中规则数目及所有规则前件中包含的模糊项总数及其对应的分类模型在训练集和测试集上的分类错误率. 从表 3 中可以看出随着最小支持度的增大,FAC 在训练集上的分类错误

率整体上呈下降趋势. 这是因为最小支持度的增大使得初始分类规则集中的规则质量提高,而且分类规则数目变少,缩小了遗传算法的搜索空间,有利于筛选较好的分类规则. 但从表 3 中也发现最小支持度为 0.10 和 0.15 时,生成的 FAC 在测试集上的分类错误率还低于在训练集上的分类错误率,这可能是因为最小支持度为 0.10 和 0.15 时,经裁剪后分类规则库中包含的分类规则数目相对较多,分类推理能力较强,而测试集中的样本数目相对训练集较少,从而造成测试集上的分类错误率低于训练集.

表 3 本文模糊关联分类器的分类性能

Table 3 Classification performance of this paper's fuzzy associative classifier

最小支持度	初始规则数	裁剪后的规则数及模糊项总数	分类错误率	
			训练集	测试集
0.10	3 077	674,2 677	0.08	0.07
0.15	943	96,312	0.03	0.02
0.20	319	11,22	0.04	0.05
0.25	137	4,5	0.04	0.07
0.30	74	3,3	0.01	0.04
0.35	39	2,2	0.01	0.01

限于篇幅,表 4 只列出了最小支持度 0.25 与 0.30 对应的 FAC 分类规则库中的规则及每条规则的置信度. 可以看出,模糊项坡度高、风速中与超限事件"拉平着陆坡度大"关联性较大,而坡度中、风速低与超限事件"50 ft 至接地距离长"关联性较大.

表 4 模糊关联分类器的规则

Table 4 Rules of fuzzy associative classifier

最小支持度	模糊关联分类规则	置信度
0.25	坡度高→拉平着陆坡度大	0.98
	方向舵位置中→拉平着陆坡度大	0.53
	风速中→拉平着陆坡度大	0.88
	坡度中,风速低→50 ft 至接地距离长	0.99
0.30	坡度中→50 ft 至接地距离长	0.97
	风速中→拉平着陆坡度大	0.88
	压力高度低→拉平着陆坡度大	0.53

图 2 为最小支持度取值为 0.20,0.25,0.30,0.35时,基于遗传算法裁剪模糊关联分类规则库过程中适应度函数取值的变化趋势图. 从图 2 中不难发现:随着迭代次数的增加,上述 4 个最小支持度取值对应的适应度取值逐步增大,最后均能收敛于某个值. 而且最小支持度值越大,遗传适应度函数的收敛值越大. 根据文中 2.1 节设计的适应度函数分析可知,最小支持度取值较大时,FAC 在训练集中正确分类的样本数目多,分类规则库中包含规则数目和模糊项数目总和较少,这与表 3 中的实验结果一致.

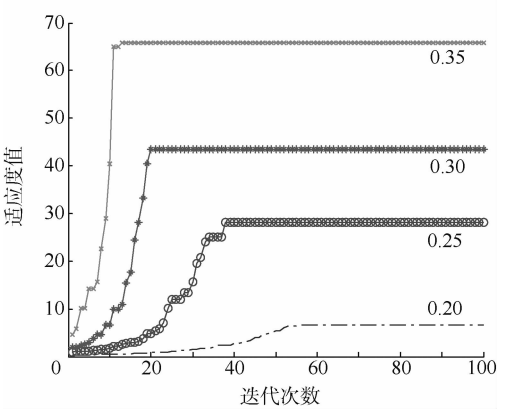

图 2 不同的最小支持度对应的适应度值变化图

Fig. 2 Changing trend of fitness value corresponding to different minimum support threshold

本文比较了 LS-SVM 与文中 FAC 在所选取的 QAR 数据集上的分类性能. 实验采用美国北卡罗莱纳那州立大学开发的基于 LS-SVMlab1.8 软件的 Matlab 工具箱,选取常用的径向基(RBF)函数作为 LS-SVM 的核函数对模型进行训练. 实验过程中首先在选取的 400 组数据上,基于 3-折交叉验证,分别采用单纯形法、网格搜索法对 LS-SVM 的参数 gam 和 sig2 进行优化,然后采用与上文训练模糊关联分类器相同的样本划分方法训练 LS-SVM 分类模型. 表 5 列出了两种参数优化方法对应的 gam 和 sig2 取值及由其所训练的 LS-SVM 分类模型在测试集上的分类错误率. 结合表 3 的结果对比分析,最小支持度为 0.30 时,FAC 在测试集上的分类错误率为 0.04,与表 5 中单纯形法对应的 LS-SVM 分类错误率相同;最小支持度0.35 时,FAC 在测试集上的分类错误率为 0.01,低于表 5 中网格搜索法对应的 LS-SVM 分类错误率 0.02. 而且本文 FAC 在解释性方面也优于 LS-SVM.

表 5 LS-SVM 的分类性能

Table 5 Performance of LS-SVM

参数优化法	gam	sig2	分类错误率
单纯形法	52 971	8.12	0.04
网格搜索法	201.97	5.06	0.02

4 结 论

1) 建立的基于模糊关联分类器的民用飞机超限事件诊断模型能较好地识别超限事件,模型中包含的模糊关联分类规则有助于实现多因素的超限事件诊断.

2) 与目前用于民机超限事件诊断的 LS-SVM 模型相比,本文模型在所采用 B737-800 的数据集上分类准确率相当,对领域工程师分析超限事件

发生的原因具有辅助作用.

3）本文的民用飞机超限事件建模方法完全基于数据驱动,与实际基于监控标准阈值判定超限事件的方法相比,提高了 QAR 数据的利用率.下一步可把领域工程师的经验知识融入模型,进一步提高该诊断模型的可用性.

参考文献（References）

[1] Bay S D, Schwabacher M. Mining distance-based outliers in near linear time with randomization and a simple pruning rule[C]// Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. New York: ACM, 2003: 29 – 38

[2] Iverson D L. Inductive system health monitoring[C]//Proceedings of the International Conference on Artificial Intelligence, IC-AI'04. Las Vegas: CSREA Press, 2004: 605 – 611

[3] Budalakoti S, Srivastava A N, Otey M E. Anomaly detection and diagnosis algorithms for discrete symbol sequences with applications to airline safety[J]. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, 2009, 39(1): 101 – 113

[4] Das S, Matthews B L, Srivastava A N, et al. Multiple kernel learning for heterogeneous anomaly detection: algorithm and aviation safety case study[C]//16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD-2010. New York: ACM, 2010: 47 – 56

[5] Smart E, Brown D, Denman J. Combining multiple classifiers to quantitatively rank the impact of abnormalities in flight data[J]. Applied Soft Computing, 2012, 12(8): 2583 – 2592

[6] 聂磊,黄圣国,舒平,等. 基于支持向量机的民用飞机重着陆智能诊断研究[J]. 中国安全科学学报,2009,19(7):149 – 153
Nie Lei, Huang Shengguo, Shu Ping, et al. Intelligent diagnosis for hard landing of aircraft based on SVM[J]. China Safety Science Journal, 2009, 19(7): 149 – 153(in Chinese)

[7] 许桂梅,黄圣国. 基于优化支持向量机的飞机重着陆智能诊断[J]. 计算机测量与控制,2011,19(2):256 – 259
Xu Guimei, Huang Shengguo. Airplane's hard landing diagnosis based on optimized support vector machine[J]. Computer Measurement & Control, 2011, 19(2): 256 – 259(in Chinese)

[8] 曹海鹏,舒平,黄圣国. 基于神经网络的民用飞机重着陆诊断技术研究[J]. 计算机测量与控制,2008,16(7):906 – 908
Cao Haipeng, Shu Ping, Huang Shengguo. Study of aircraft hard landing diagnosis based on neural network[J]. Computer Measurement & Control, 2008, 16(7): 906 – 908(in Chinese)

[9] 祁明亮,邵雪焱,池宏. QAR 超限事件飞行操作风险诊断方法[J]. 北京航空航天大学学报,2011,37(10):1207 – 1210
Qi Mingliang, Shao Xueyan, Chi Hong. Flight operations risk diagnosis method on quick access record exceedance[J]. Journal of Beijing University of Aeronautics and Astronautics, 2011, 37(10): 1207 – 1210(in Chinese)

[10] 董杰,韩敏. 基于自适应区间划分的模糊关联分类[J]. 系统仿真学报,2009,21(9):2675 – 2678
Dong Jie, Han Min. Fuzzy associative classification based on adaptive interval partition[J]. Journal of System Simulation, 2009, 21(9): 2675 – 2678(in Chinese)

[11] 董杰,沈国杰. 一种基于模糊关联分类的遥感图像分类方法[J]. 计算机研究与发展,2012,49(7):1500 – 1506
Dong Jie, Shen Guojie. Remote sensing image classification based on fuzzy associative classification[J]. Journal of Computer Research and Development, 2012, 49(7): 1500 – 1506(in Chinese)

[12] 霍纬纲,邵秀丽. 一种基于多目标进化算法的模糊关联分类方法[J]. 计算机研究与发展,2011,48(4):567 – 575
Huo Weigang, Shao Xiuli. A fuzzy associative classification method based on multi-objective evolutionary algorithm[J]. Journal of Computer Research and Development, 2011, 48(4): 567 – 575(in Chinese)

[13] Alcalá-Fdez J, Alcalá R, Herrera F. A fuzzy association rule-based classification model for high-dimensional problems with genetic rule selection and lateral tuning[J]. IEEE Transactions on Fuzzy Systems, 2011, 19(5): 857 – 872

[14] 霍纬纲,高小霞. 一种适用于多类不平衡数据集的模糊关联分类方法[J]. 控制与决策,2012,27(12):1833 – 1838
Huo Weigang, Gao Xiaoxia. A fuzzy associative classification method for multi-class imbalanced dataset[J]. Control & Decision, 2012, 27(12): 1833 – 1838(in Chinese)

[15] Pach F P, Gyenesei A, Abonyi J. Compact fuzzy association rule-based classifier[J]. Expert Systems with Application, 2008, 34(4): 2406 – 2416