

基于粗糙集的一体化保障技术 属性约简问题研究

葛亚维¹, 肖明清¹, 程进军¹, 刘晓林², 杨 台¹

(1. 空军工程大学 自动测试系统实验室, 西安 710038; 2. 95007 部队, 广州 白云 510410)

摘要: 一体化保障技术是目前提高复杂武器系统测试和保障能力的研究重点, 为解决一体化保障技术中存在的指标冗余问题, 结合粗糙集相关理论, 提出了基于属性重要性和属性间影响度的约简算法; 通过在某型一体化保障设备的柴油发电机子模块中应用该算法, 在精简测试指标的同时, 保证了较高的测试准确性, 缩短了测试时间, 验证并说明该理论可以有效地解决一体化保障技术中的指标冗余问题。

关键词: 一体化保障; 粗糙集; 属性约简算法; 重要性; 影响度

Research of Attribute Reduction Problems in Integrated Support Technology Based on Rough Set

Ge Yawei¹, Xiao Mingqing¹, Cheng Jinjun¹, Liu Xiaolin², Yang Zhao¹

(1. ATS Lab, Air Force Engineering University, Xi'an 710038, China;

2. No. 95007 Unit of PLA, Guangzhou 510410, China)

Abstract: The integrated support technology is the research emphasis of improving test and support capability of complex weapon equipment recently. In order to solve the index redundancy problem existing in the integrated support technology, combining with the Rough Set Theory, the attribute reduction algorithm based on significance and influence degree among attributes is proposed. Through the application of the algorithm in the diesel generator subsystem of a certain type of integrated support equipment, when simplifying test index, high accuracy is ensured, and testing time is shortened, as the verification shows, the algorithm can effectively solve the index redundancy problem of the integrated support technology.

Keywords: integrated support; rough set; attribute reduction algorithm; significance; influence degree

0 引言

一体化保障技术是为了满足复杂武器系统的测试和保障需求, 提高保障设备的综合化和集成化水平而提出的一种新型集成方法^[1]。它将以往多种类别的保障功能集成于一体, 具有较强的数据融合能力、良好的信息共享机制以及开放的功能架构, 可以提高保障设备的机动性, 简化保障设备的操作使用过程, 从而高效、准确地完成复杂武器系统的测试和保障任务。一体化保障完成武器装备多个功能模块的测试和保障, 需要测试大量技术指标, 然而并非所有的测试指标都是必要的、重要的。如何在确保完成测试任务的条件下尽可能精简测试指标, 是一体化保障中的一个难题。

粗糙集^[2-5] (Rough Set, RS) 是由波兰科学家 Pawlak Z. 于 1982 年提出的, 它是能在保持原有知识分类不变的情况下, 约简不重要的或不相关的属性, 达到简化知识系统的目的。

为解决某型一体化保障设备“数据量大、测试指标冗余”的测试保障问题, 本文运用粗糙集理论提出了基于属性重要性和属性间影响度的约简算法, 并通过在柴油发电机子模块中具体应用, 验证并说明该算法可以有效地解决一体化保障技术中的指标冗余问题。

1 问题描述

某型一体化保障设备是模块化多功能测试保障设备, 主要用于对我军多种型号的飞机开展空调、液压、氮气、电源等一系列航空勤务保障工作, 其结构组成如图 1 所示。

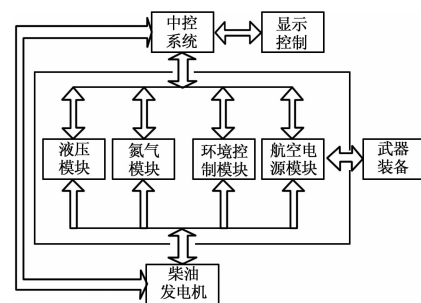


图 1 某型一体化保障设备结构组成图

某型一体化保障设备的测试指标体系如图 2 所示, 它综合了以往多个测试保障设备的功能, 在保障过程中需要测试大量的指标, 然而并非所有的指标都是必要的、重要的, 对全部指标进行测试会增加测试成本、降低测试效率, 因此需要简化冗余的测试指标。对一体化保障设备中冗余的测试指标进行简化, 本质上可以归结为粗糙集理论中的属性约简问题。

收稿日期: 2014-01-15; 修回日期: 2014-03-16。

作者简介: 葛亚维(1990-), 男, 山东枣庄人, 硕士研究生, 主要从事自动测试系统集成等方向的研究。

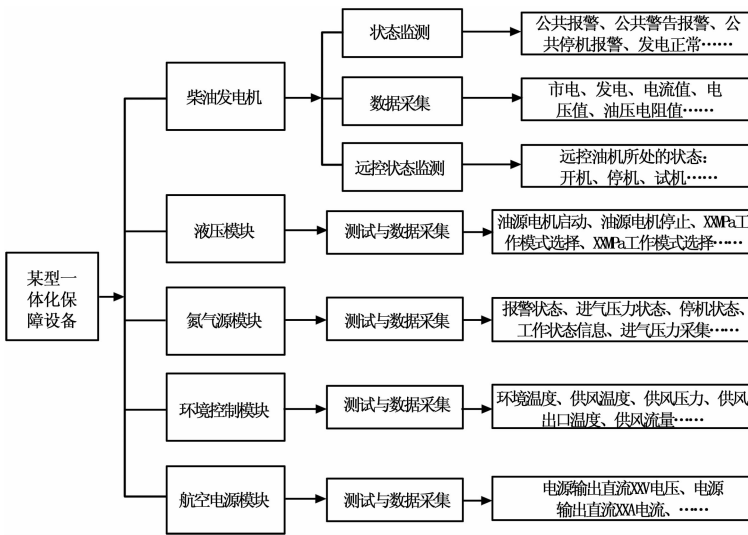


图 2 某型一体化保障设备各功能模块测试指标图

影响度函数为：

$$IMP(a) = \frac{count(RED \cap a)}{count(a)} \quad (2)$$

式中， RED 代表一个约简， $count(a)$ 为属性 a 在可辨识矩阵中出现的次数， $count(RED \cap a)$ 为属性 a 与约简集中相同属性同时出现的次数。

2.2 约简算法实现过程

下面给出给定信息决策系统 $L = (U, C, D, V, F)$ 基于属性重要性和属性间影响度求取约简的一般算法。

(1) 系统初始化：

$$core = \phi, B = C$$

(2) 求取可辨识矩阵^[8-9]：

即可辨识矩阵 G_{ij} 为：

$$G_{ij} = \begin{cases} a \in A: a(x_i) \neq a(x_j) & D(x_i) \neq D(x_j) \\ 0 & D(x_i) = D(x_j) \\ -1 & a(x_i) = a(x_j), D(x_i) \neq D(x_j) \end{cases} \quad (3)$$

式中， $a(x)$ 是 x 在属性 a 上的值。定义系统 $S = (U, A)$ ，可辨识矩阵计算过程如下：

$$G = DisMat(S)$$

$$for(i = 0; i < n; i++)$$

$$for(j = i + 1; j < n; j++)$$

$$for(k = 1; k \leq |c|; k++)$$

$$if((C_k(X_i) \neq C_k(X_j)) \&\& (D(X_i) \neq D(X_j)))$$

$$G_{ij} = G_{ij} \cup \{c_k\}$$

(3) 求核：

令 R 为一族等价关系， $r \in R$ ，如果 $ind(R) = ind(R - \{r\})$ ，则称 r 为 R 中必要的。设 $Q \subseteq T$ ，如果 Q 是独立的，且 $ind(Q) = ind(T)$ ，则称 Q 为 T 的一个约简。 T 中所有必要关系组成的集合称为核，记为 $core(T)$ 。

具体计算步骤如下：

① 计算决策属性 D 的条件属性集 C 的正域 $POS_C(D)$ ；

② 如果 $B = \phi$ 则转 ⑤，否则，任取 $a_i \in B$ ，计算 $POS_{C-\{a_i\}}(D)$ ；

③ $B = B - \{a_i\}$ ，并判断 $POS_C(D)$ 和 $POS_{C-\{a_i\}}(D)$ ，如果 $POS_C(D) \neq POS_{C-\{a_i\}}(D)$ ，根据上述定义则说明属性 a_i 是属性集 C 中不可省略的，否则转 ②；

④ $core = core \cup \{a_i\}$

⑤ 输出核。

(4) 求约简集：

① 数据初始化 $R = core$ ；

② $T = R, \forall x_i \in (C - R)$ ，如果 $f(\gamma_{R \cup \{x_i\}}(D) > \gamma_R(D))$ ，则令 $T = R \cup \{x_i\}$ ， $R = T$ 并转到 ③，否则继续循环 ②；若没有剩余属性可选，也转到 ③；其中 $i = 1, 2, 3, \dots, n (n \leq |C - R|)$ ；

③ 判断 $\gamma_R(D) = \gamma_C(D)$ ，如果是则结束，转 ④，否则转到 ②；

④ 返回约简集 R 。

(5) 影响度^[10]分析过程：

计算各属性的重要性：

$$SIG(a_i) = \sum_{i=1}^n count(a_i) / i \quad (4)$$

2 基于属性重要性和属性间影响度的约简算法

2.1 基于粗糙集的属性约简方法

在数据分析^[6]中，设信息系统 $L = (U, C, D, V, F)$ 是一个决策表，其中 $U = \{x_1, x_2, \dots, x_n\}$ 表示对象的非空有限集合，称为论域； C, D 分别表示条件属性和决策属性的非空有限集合， $C \cap D = \phi$ ，而 $C \cup D = A$ ，即 A 为属性集合； V 为属性集 A 的值域； $F: U \times A \rightarrow V$ 为一个信息函数，表示任一对象的属性在 V 上的取值， $F(x, r) \in V_r$ ，表示 U 中每个 x 的属性值。但决策表中的属性并不是同等重要的，冗余属性的存在不仅浪费资源，还会干扰人们做出正确的决策。

一个决策表^[7]的条件属性对于决策属性的相对约简定义为：

令 $M \subseteq C, N \subseteq D, U/N \neq \{U\}$ 。可以找到一个极小子集 $M_0 \subseteq M$ ，满足 $S_{M_0}(N) = S_M(N)$ ，且如果 $M' \subset M_0$ ，则 $S_{M'}(N) \neq S_M(N)$ ，则称 M_0 是 M 的一个约简。同一个决策表可能存在多个约简，而约简中属性的多少直接影响着决策规则的繁简和性能。

解决此问题的方法是将属性重要性作为启发信息引入到属性约简中，并综合考虑属性间的相互影响度因素，减小样本的搜索空间，从而提高约简效率。

其中属性重要性的定义为：

令 $M \subseteq C, N \subseteq D, U/N \neq \{U\}$ ，给定 $m \in M$ ，如果 $S_m(N) \neq S_{M-\{m\}}(N)$ ，则称 m 在 M 中是重要的。

考虑属性间相互影响进行属性约简的方法是：以属性核作为初始约简集的基础，计算属性的重要性；选择重要性高的属性，计算其对于属性核的影响度，依据影响度标准综合考虑是否加入属性约简集。依次按照属性的重要性高低重复上述过程，直到该属性集合对决策属性的依赖程度与初始约简集的一致为止，最终得到最优约简集。

其中属性间影响度的定义为：

如果 $S_m(N) = S_{M-\{m\}}(N)$ ，则称 x 在 X 中是不重要的（对于 N 而言），则属性子集 $C' \subseteq C$ 关于 D 的重要性定义为：

$$\sigma_{CD}(C') = \gamma_C(D) - \gamma_{C-C'}(D) \quad (1)$$

对于任意属性 $\{a | a \in B, B \subseteq C\}$ 相对于约简集中属性的

制式摄像头图像采集频率为 30 帧每秒, 图像的分辨率为 720×576 , 实验中参数设置为, $\epsilon=1$, $th1=10.89th$, 使用本文的方法在平台上进行验证。效果图如图 5 所示。

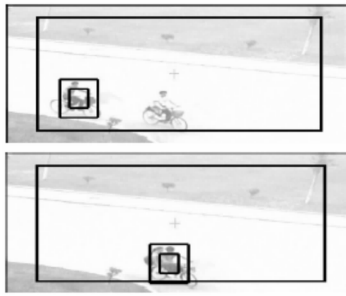


图 5 实验目标跟踪效果图

本文目标跟踪算法将其运用到 DM642 中, 成功实现了遮挡和无遮挡情况下运动目标的实时跟踪。实验结果表明该方法具有较好的实际应用价值。

5 结论

本文首先对文中使用的 DSP 实时图像处理硬件平台和使用的软件开发环境 CCS 进行介绍, 接着介绍将文中算法移植

到 DSP 开发平台所用到的算法移植和优化的方法。实验结果表明, 本章所使用的算法移植和优化方法效果明显, 可以在 DSP 开发平台上实现运动目标的实时检测与跟踪。

参考文献:

- [1] 吴 薇, 朱光喜, 陈云锋. 基于视频技术的嵌入式交通信息采集系统 [J]. 计算机应用研究, 2008, 25 (7): 2104-2106.
- [2] 李功燕, 陈晓鹏, 李 斌, 等. 基于 Xscale 和多 DSP 的智能视频监控系统的设计与实现 [J]. 计算机工程与设计, 2009, (2): 003.
- [3] 郑伟亮, 蔡 铁. 基于 DM642 的视频稳像系统设计 [J]. 科学技术与工程, 2009, 9 (5): 1143-1145.
- [4] 胡 波, 陈 息, 徐建瑜, 等. 基于 Kalman 预测和 Mean-Shift 算法的视频目标跟踪 [J]. 光电子·激光, 2009, 20 (11): 1517-1522.
- [5] 周尚波, 胡 鹏, 柳玉炯. 基于改进 Mean-Shift 与自适应 Kalman 滤波的视频目标跟踪 [J]. 计算机应用, 2010, (6): 1573-1576.
- [6] 左军毅, 程咏梅, 王正平. 一种对背景干扰及遮挡鲁棒的视频目标跟踪算法 [J]. 计算机测量与控制, 2009, 17 (11): 2292-2294.
- [7] 张 娟, 毛晓波, 陈铁军. 运动目标跟踪算法研究综述 [J]. 计算机应用研究, 2009, 26 (12): 4407-4410.
- [8] 宦若虹, 王浙沪, 唐晓梅, 等. 动态背景下基于粒子滤波的运动目标跟踪方法 [J]. 计算机应用与软件, 2011, 28 (5): 109-111

*** **

(上接第 2191 页)

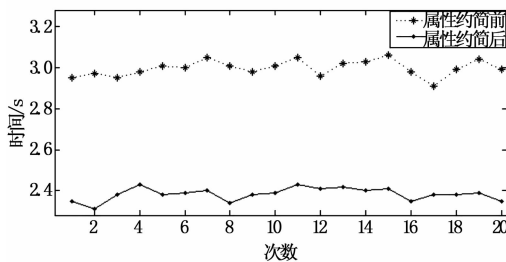


图 3 测试任务完成时间对比

在某型一体化保障设备上上进行多次实验, 记录每次属性约简前后的测试时间, 剔除数据中的坏值, 随机选取 20 组数据, 绘制折线图并求取平均值, 可以得到属性约简前所需时间为 2.997 s, 而经过属性约简后的完成时间为 2.383 5 s, 测试时间缩短了 20.47%。

通过上述分析及验证表明, 某型一体化保障设备在经过属性约简后, 可以在确保完成测试任务的条件下, 精简测试指标, 并缩短测试时间。在该设备的其他功能模块中采用该算法进行属性约简, 可以得到优化的测试指标体系, 使测试设备更加高效地完成测试保障任务。

5 结论

本文针对一体化保障集成技术中的测试指标冗余问题, 运用粗糙集理论, 提出了基于属性重要性和属性间影响度的约简算法, 通过在某型一体化保障设备的柴油发电机子模块中应用

该算法, 验证并说明了该算法可以有效地精简一体化保障中的测试指标体系, 并提高保障效率。

参考文献:

- [1] 文 莹, 肖明清, 胡雷刚, 等. 基于粗糙神经网络的航空电子设备故障预测研究 [J]. 计算机测量与控制, 2010, 18 (4): 807-809.
- [2] 徐 权, 倪世宏, 张 鹏. 基于粗糙集的专家系统知识库约简研究 [J]. 计算机测量与控制, 2012, 20 (5): 1333-1335.
- [3] 马彦涛, 肖明清. 基于粗糙集的复杂航空电子设备故障诊断 [J]. 计算机测量与控制, 2012, 20 (7): 1757-1759.
- [4] Pawlak Z. Rough sets [J]. International Journal of Computer and Information Sciences, 1982, 11 (5): 341-356.
- [5] Parthala N M, Shen Q, Jensen R. A Distance Measure Approach to Exploring the Rough Set Boundary Region for Attribute Reduction [J]. IEEE Transactions on Knowledge and Data Engineering, 2010, 22 (3): 305-317.
- [6] 张文东, 李明壮, 石小艳. 基于粗糙集理论的属性约简算法 [J]. 计算机工程与设计, 2008, 29 (22): 5795-5797.
- [7] 索中英, 程嗣怡, 苟新禹. 基于极小析取范式的属性约简算法 [J]. 空军工程大学学报: 自然科学版, 2012, 13 (6): 66-70.
- [8] 杨 尧. 基于粗糙集理论的决策信息系统属性约简算法研究 [D]. 青岛: 青岛大学, 2011.
- [9] 安芹力, 李安平. 不协调决策表的属性约简模型及规则提取 [J]. 空军工程大学学报: 自然科学版, 2005, 6 (3): 88-91.
- [10] 夏春艳, 李树平, 宁志超. 基于粗糙集理论属性约简的改进算法 [J]. 微计算机信息, 2010, 26 (12): 282-283.