

doi: 10.7690/bgzdh.2020.02.001

决策树在装备维修器材筹措中的应用

代冬升¹, 李雅峰¹, 钟 华¹, 李铮铮²

(1. 中国人民解放军 32181 部队一室, 西安 710032; 2. 陆军工程大学石家庄校区教练勤务连, 石家庄 050003)

摘要: 针对装备维修器材筹措工作中品种选择的问题, 应用决策树分类方法, 对典型装备正常周转器材库存数据进行挖掘分析。对器材筹措基本业务、筹措品种选择需求以及筹措决策树的作用进行描述, 应用 C4.5 算法, 通过计算信息增益比例来选择分裂属性, 对装备维修器材进行相关属性约简、属性值域简化和决策树构造, 得到器材筹措品种选择需要重点考虑的因素及各因素的重要性。应用结果表明, 该分析能够为装备保障方案优化提供重要参考依据。

关键词: 决策树; 装备维修器材; 器材筹措

中图分类号: TJ02 **文献标志码:** A

Application of Decision Tree in Equipment Maintenance Material Raising

Dai Dongsheng¹, Li Yafeng¹, Zhong Hua¹, Li Zhengzheng²

(1. No. 1 Department, No. 32181 Unit of PLA, Xi'an 710032, China;

2. Training Service Company, Shijiazhuang Campus of PLA University of Army Engineering, Shijiazhuang 050003, China)

Abstract: Aiming at the problem of varieties selection in equipment maintenance equipment raising, it mines and analyzes the stock data of typical equipment normal working material by the decision tree classification method. It describes the basic business of equipment raising, the selection requirements of raising varieties, and the function of raising decision tree. Applying the C4.5 algorithm, it selects split attributes by calculating the information gain ratio, reduces the relative attribute, simplifying the attribute range, and constructing the decision tree for equipment maintenance material. And then, it obtains the key factors and the importance of each factor. The application results show that this analysis can provide an important reference for equipment support scheme optimization.

Keywords: decision tree; equipment maintenance material; material raising

0 引言

决策树^[1-2]是指从一组无次序、无规则的训练样本集中进行归纳学习, 形成以决策树为表现形式的分类规则, 并应用分类规则对目标数据集中记录进行分类的方法。决策树分类的关键是分类规则的确定。决策树是一个树形结构, 树的一个节点代表某条件下的一个记录集, 根据节点属性的不同取值建立树的分支, 在每个分支子集中重复上述工作, 逐步建立下层的节点和分支。构建决策树的常用算法包括 ID3、C4.5、CART、SPRINT 等。决策树容易被理解且效率较高, 易于转换成 IF-THEN 分类规则来辅助实际决策, 使得其在数据挖掘领域中有重要的地位; 因此, 笔者对决策树在装备维修器材筹措中的应用进行探讨。

1 装备维修器材筹措问题分析

1.1 筹措业务基本描述

装备维修器材是装备故障维修、维护保养等工

作中所需各种备件、设备和原材料的统称, 是部队进行装备维修、维护保养的重要物质基础。装备维修器材筹措是装备维修器材保障的首要环节, 是装备维修器材保障部门通过各种形式和渠道, 有组织、有计划、有选择地进行申请、采购、订货、生产等一系列筹集活动的统称^[3]。为满足联合作战装备保障的需要, 装备维修器材筹措需要做到适用、及时、齐备。其中, 适用主要是指筹措的器材在品种、数量上要与实际需求相匹配; 及时主要是指器材能够按要求储备到指定地点, 并满足部队作战需要; 齐备主要是指器材品种要齐全, 不能因为有的品种数量庞大, 导致积压浪费, 有的品种数量过少或没有储备导致缺货, 影响部队使用。

装备维修器材筹措主要是完成将装备维修器材从生产厂家购置到军队器材仓库的过程, 简要过程可描述为: 首先, 预计部队装备维修、维护保养等的工作量和需求量, 在此基础上确定需要购置的器材品种和数量; 然后, 按照有关要求制定器材订货

收稿日期: 2019-11-02; 修回日期: 2019-12-24

作者简介: 代冬升(1979—), 男, 山东人, 博士, 助理研究员, 从事装备保障信息化、装备保障数据分析研究。E-mail: dds1979@sina.com。

计划,明确从哪些生产厂家、购置哪些器材、共计多少金额等;接着,订货计划经过审批后,军方与生产厂家签订订货合同,生产厂家按照合同进行生产,并将相应器材交付指定仓库或部队。

1.2 筹措品种选择需求

筹措是装备维修器材供应保障的关键环节,其核心内容是确定“买什么、买多少”的问题,即确定装备维修器材购置的品种和数量。单独从装备维修器材品种的角度看,其在装备维修器材筹措中主要需求表现为:

1) 需要筹措的器材品种多。当前,我军武器装备技术密集、功能强大、结构复杂,在装备使用维护、故障维修过程中,需要使用的器材品种和数量非常大。这些器材大部分都需要进行筹措,才能满足部队装备维修保障、维护保养的需要。以某型自行火炮为例,在使用维修过程中涉及的维修器材品种达上万个,百门份装备年筹供维修器材数量达上百万个。

2) 筹措品种选择难度大。受购置经费、部队需求和库存条件等多种因素的影响和制约,从品种众多的维修器材中选择恰当的品种进行购置和储备成为装备维修器材保障中的难点问题。有的器材单装用数少、消耗量小,但对于装备却非常重要,需要少量购置;有的器材单装用数多、消耗量大,但因库存数量较大却不一定需要购置;有的器材间存在组合配对关系,必须同时购买。因此,需要结合实际情况,制定器材筹措品种选择的原则和依据,确保购置的器材能够满足部队实际需求。

3) 对器材保障效益发挥影响重大。从军事效益上看,器材品种配套齐全能够为部队顺利完成作战任务提供保证,如果器材品种缺失,势必影响部队装备保障工作,进而影响部队战备训练和遂行多样化作战任务。从经济效益上看,将有限的资金用于购买恰当的器材品种,能够节约经费,避免浪费。

1.3 筹措决策树作用分析

通过上述分析可以发现:正确选择合适的器材品种进行器材筹措与储备难度很大,在日常工作中对专家经验的依赖较大。同时,装备维修器材的单装用数、筹供标准、关重性、易损性等属性差异较大,器材筹措受生产周期、储存条件、物流运输能力、维修任务量及急迫程度、部队修理能力等多种因素影响,难以一次性将所有器材全部保障到位。

构建装备维修器材筹措决策树,能够以大量装备维修器材筹措历史数据为样本,根据样本数据的属性集进行分类,从中提取不同条件下器材的分类规则,从而形成针对不同情况下器材品种选择的策略,为装备维修器材品种的选择分析提供指导。

2 装备维修器材筹措决策树构建方法

2.1 基本思路

从正常周转器材库存事实中抽取部分样本,直接应用 C4.5 算法构造决策树存在 2 点突出问题: 1) 有的属性字段的属性值较为分散,例如单装用数,难以直接作为条件属性; 2) 事实表属性字段较多,如果全部用来构造决策树,决策树的规模较大,而且效率较低。针对上述问题,在进行器材筹措决策树构建过程中,首先针对样本数据集属性进行简化处理,包括属性数量的简化和属性值的简化 2 部分。对属性经过简化处理以后,再应用 C4.5 算法进行决策树的构建。其基本过程如图 1 所示。

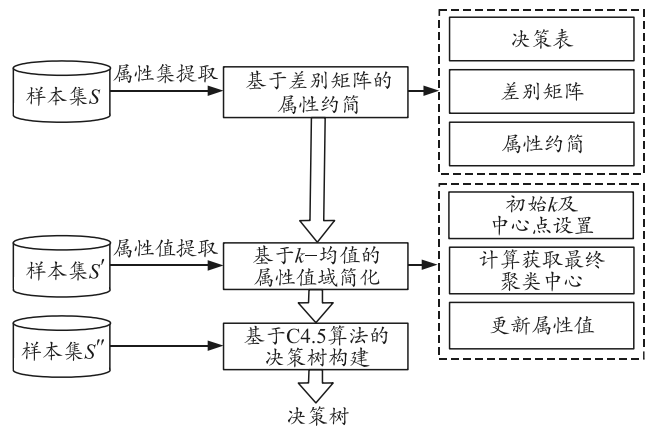


图 1 器材筹措决策树构建基本过程

在应用差别矩阵进行属性约简时,首先需要建立决策表,依据决策表生成差别矩阵,然后对差别矩阵进行分析得到决策属性的核,并根据启发式规则,最终得到满足条件的属性约简。在应用 k -均值方法进行属性值域简化时,设置初始 k 值和中心点,通过计算误差平方和求取最终聚类中心,并将不同类中的属性值设定为一个统一值,以便分类。在应用 C4.5 算法进行决策树构建时,通过计算信息增益比例来选择分裂属性,最终生成决策树。

2.2 基于差别矩阵的属性约简^[1,4]

装备维修器材的属性字段较多,但重要程度存在差异,全部用于决策树的构造既耗时耗力,又掩盖了很多重要字段的决定性作用,最终的决策树也因过于庞大不能起到很好的辅助决策作用。差别矩

阵是关于主对角线的对称矩阵，可以通过构建决策表 S 及其差别矩阵 $E_{n \times n}$ ，判断装备维修器材属性的独立性，将具有依赖性的冗余属性字段剔除，保留具有独立性、对器材筹措决策具有直接作用的条件属性字段，即可实现装备维修器材属性的约简。

设 C 是条件属性集， D 是决策属性集，差别矩阵 $E_{n \times n}=(c_{ij})_{n \times n}$ ， B 为约简结果。基于差别矩阵的属性约简算法的基本过程如下：

- 1) 计算差别矩阵 $E_{n \times n}$ ；
- 2) 获取决策表的相对核 $core_D(C)$ ，令 $B=core_D(C)$ ；
- 3) 对任意 C_{ij} ，如果 $c_{ij} \cap B = \emptyset$ ，则令 $c_{ij} = \emptyset$ ；
- 4) 对任意 C_{ij} ，如果都有 $c_{ij} = \emptyset$ ，则转 6)，否则继续；
- 5) 计算当前 $E_{n \times n}$ 中出现次数最多的元素 a_m ，令 $B=B \cup \{a_m\}$ ，转至 3)；
- 6) $B \in red_D(C)$ ，结束。

通过上述计算，即可得到属性集 C 的所有约简 $red_D(C)$ ，用于装备维修器材决策树构造。

2.3 基于 k -均值聚类的值域简化

装备维修器材很多属性字段取值较分散，不利于直接用作条件属性，笔者采用 k -均值聚类的方法将值域划分成若干子类，并重新对子类进行定义以实现值域的简化处理。

首先，应用 k -均值聚类的方法，将某一字段的值域划分为 k 个子类。在设定参数 k 的时候，需要紧密结合器材筹措决策树的分类需求， k 值过大会导致决策树分支过多。具体过程^[2]为：

- 1) 设定 k 个初始聚类中心 $Z_j(I)$ ；
- 2) 通过计算每个数据对象与聚类中心的距离 $D(x_i, Z_j(I))$ 进行对象归类；
- 3) 重新计算 k 个新的聚类中心 $Z_j(I+1) = \frac{1}{n} \sum_{i=1}^{n_j} x_i^{(j)}$ ；
- 4) 根据误差平方和准则函数 $J_c(I) =$

$\sum_{j=1}^k \sum_{k=1}^{n_j} \|x_k^{(j)} - Z_j(I)\|^2$ ，计算 $J_c(I) - J_c(I-1)$ 直至小于设定阈值，否则转 2) 继续分类。

其次，观察字段取值的大致情况，对聚类形成的 k 个类别的值进行简化处理。例如，某种装备维修器材的年消耗量小于 10 描述为“一般”，年消耗量介于 10~20 描述为“较多”，年消耗量大于 20 描述为“很多”，这是一个从定量到定性的过程。

2.4 决策树构造

装备维修器材筹措决策树采用 C4.5 算法进行构建。C4.5 算法是对 ID3 算法的改进，通过增加信息增益比例来选择分类属性，克服了 ID3 算法在属性多值偏向性方面的不足，能够对连续值属性进行离散化处理，并可以通过属性值预测进行空缺属性值补充，表现出了较好的特性^[5-6]。

在装备维修器材筹措决策树构造过程中，首先从经过约简处理的属性集中将计算信息增益最大的属性选为分裂属性，根据分裂属性值进行叶节点设置，然后重复这一过程。当样本集为空，或者样本集中所有记录都属于同一个类别，或者候选属性集为空时，决策树构造结束。

3 装备维修器材筹措决策树构建实例

装备维修器材信息管理过程中，正常周转器材库存事实表包括单位代码、库存类型、装备代码、器材代码、规格件号、计量单位等信息，与器材目录标准、单位维、装备维、成套类型维等维表关联，形成星形组织结构。笔者以某型火炮为例，从该装备某年正常周转器材库存数据中抽取 10 027 条记录作为训练样本集，在此基础上进行相关属性约简、属性值域简化和决策树构造。

3.1 计算过程

- 1) 属性约简。

为简化计算，选择库存量、单装用数、筹供标准、单装配套、关键性、易损性和设计寿命等 7 个属性字段，建立如表 1 所示的决策表。

表 1 决策表

U	库存量(a_1)	单装用数(a_2)	筹供标准(a_3)	单装配套(a_4)	关键性(a_5)	易损性(a_6)	设计寿命(a_7)	是否储备(D)
1	1	1	1	1	2	1	1	1
2	2	2	2	1	2	1	1	0
3	2	1	1	1	2	1	1	0
4	3	2	1	3	1	3	1	0
5	2	3	2	1	2	1	1	1
6	1	2	1	2	1	3	2	1
7	3	1	1	2	1	2	2	0
8	2	3	2	1	1	1	1	0
9	3	1	1	2	1	3	2	1
10	2	1	3	1	2	1	1	1

经过计算，得到差别矩阵核属性为 $\{a_1, a_2, a_3, a_4, a_5, a_6\}$ 。由于差别矩阵中包含核属性的属性组合，可确定决策表的属性约简即为核属性 $\{a_1, a_2, a_3, a_4, a_5, a_6\}$ ，而属性 a_7 可以从决策过程中删除。

2) 值域简化。

表 2 属性值域聚类处理结果

聚类属性	k	ζ	I	Z_1	Z_2	Z_3	C_1	C_2	C_3
库存量	3	10	22	218.2	3 100.8	7 473.9	8 226	1 230	571
筹供标准	3	10	12	9.1	348.6	4 750	9 921	102	4
单装用数	3	10	11	3.2	190.5	1 755	9 980	45	2

表中： I 表示迭代次数； Z_1 、 Z_2 、 Z_3 分别代表最终计算的 3 个中心的值； C_1 、 C_2 、 C_3 分别代表对应以 Z_1 、 Z_2 、 Z_3 为中心的类别中对象的个数。

3) 决策树构建。

在本例中，样本数 $S=10\ 027$ 。决策属性值只有购买和不购买。对应 C4.5 算法中， $m=2$ ， $s_1=8\ 947$ ， $s_2=1\ 080$ ， $P_1=0.892$ ， $P_2=0.108$ 。条件属性有 5 个，即库存量、单装用数、筹供标准、关键性和易损性。

通过比较可以看出：库存量的信息增益比例最

通过对正常周转器材库存事实相关属性取值情况的仔细观察，发现单装用数、筹供标准、库存量等属性字段的取值较为分散，需要通过属性值的聚类处理，将属性值划分为不同的类别，并进行简化处理。属性值域聚类处理结果如表 2 所示。

大，首先选择库存量作为分裂节点。在表 2 中给出了库存量为“较少”“一般”“较多”3 种情况下的器材种类，按此将库存量节点分为 3 个分支。由于库存量为“较少”的分支所有记录决策属性都为“购买”，该分支节点即为叶节点。然后计算剩余条件属性对于库存量为“一般”的记录的信息增益比例，选择信息增益比例最大的属性为分裂节点，并重复上面的工作，最终构建如图 2 所示的决策树。其中，每个分类属性后的括号中的数值表示该节点下的记录数。

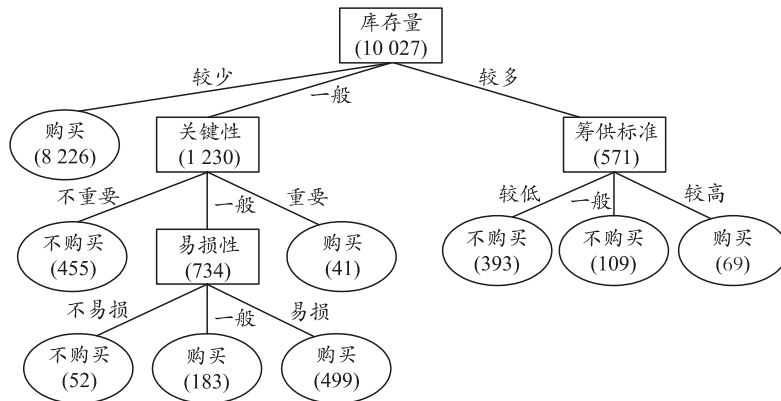


图 2 装备维修器材筹措决策树

3.2 结果分析

从图 2 的决策树，可以得到分类规则如下：

IF 库存量=较少 THEN 购买，

IF 库存量=一般 AND 关键性=不重要 THEN 不购买，

IF 库存量=一般 AND 关键性=一般 AND 易损性=不易损 THEN 不购买，

IF 库存量=一般 AND 关键性=一般 AND 易损性=一般 THEN 购买，

IF 库存量=一般 AND 关键性=一般 AND 易损性=易损 THEN 购买，

IF 库存量=一般 AND 关键性=重要 THEN 购买，

IF 库存量=较多 AND 筹供标准=较低 THEN 不

购买，

IF 库存量=较多 AND 筹供标准=一般 THEN 不购买，

IF 库存量=较多 AND 筹供标准=较高 THEN 不购买。

根据决策树结果，从装备维修器材筹措的整体上来看，库存量对器材购买决策的影响最大，其次是关键性、易损性和筹供标准，单装用数影响较小^[7-8]；因此，装备维修器材筹措过程需要优先从库存量、关键性、易损性和筹供标准等方面进行考虑，确定相应的器材品种，确保能够及时储备符合需求的装备维修器材，有效地完成相应的保障任务。

5 结束语

笔者运用网络信息技术搭建军用车辆装备管理平台,实现运输保障的互联互通互操作,提升军车保障能力和服务质量,以适应信息化战争的需求。虽然因为军事装备高保密性要求,形成了军用信息平台建设与互联网运用之间的天然沟壑,但随着“北斗”系统的不断完善和军队网络的全覆盖,拥有完全自主的网络体系,“网约”军用车辆装备也能在不久的将来得以实现。

参考文献:

[1] 王静. 中国网约车的监管困境及解决[J]. 行政法究, 2016(2): 49-59.
 [2] 李胤. 网约车平台法律责任边界研究[J]. 交通世界, 2018(34): 10-11.
 [3] 侯登华. “四方协议”下网约车的运营模式及其监管路径[J]. 法学杂志, 2016, 37(12): 68-77.
 [4] 白莹莹, 刘均禄, 卢滢伊. 新规视野下网约车运营模式及其主体责任分析[J]. 经贸实践, 2017(7): 9-11.

(上接第 4 页)

4 结论

笔者针对装备维修器材筹措工作中的品种选择问题,应用决策树分类方法进行典型装备正常周转器材库存数据挖掘分析,得到器材筹措品种选择需要重点考虑的因素及各因素的重要性。该研究能够为装备保障方案优化提供依据,为实际工作中的装备维修器材筹措提供指导。

参考文献:

[1] 朱明. 数据挖掘[M]. 合肥: 中国科学技术大学出版社, 2008: 63-64.
 [2] 范明, 牛常勇. 数据挖掘基础教程[M]. 北京: 机械工

(上接第 32 页)

[8] 刘钢, 老松杨, 谭东风. 基于功能区域的反舰导弹逆向航路规划[J]. 系统工程与电子技术, 2011, 33(4): 799-805.
 [9] 赵永涛, 胡云安, 熊厚情. 基于 HLA 的舰空巡飞弹协同指导作战仿真系统方案[J]. 海军航空工程学院学报, 2012, 27(2): 133-137.
 [10] WANG G, SUN X, ZHANG L, et al. Saturation attack based route planning and threat avoidance algorithm for cruise missiles[J]. Journal of Systems Engineering and

Transportation Administration Office. Notice on Issuing the Management Measures for Network Booking Taxi Supervision Information Interaction Platform Operation Management [EB/OL]. Transport (2018)24 No. Baidu Encyclopedia [2018-02-26].
 [6] 杨华, 朱闽, 邵文韬. 车辆装备保障数据服务平台业务需求与设计建设研究[J]. 兵器装备工程学报, 2018(3): 143-146.
 [7] 交通部. 全国网约车监管信息交互平台已收到 49 家平台公司运营数据 [EB/OL]. 中国交通新闻网 [2018-08-12].
 [8] 深圳市都市交通规划设计研究院. 互联网+,网约车监管利器: 深圳市网约车监管平台技术方案[EB/OL]. 搜狐网[2017-05-11].
 [9] 周鹏. 数据挖掘技术下的智能化车辆管理系统实现[J]. 现代电子技术, 2016, 39(16): 52-54.
 [10] 王冬梅. 大数据背景下人工智能如何服务于政府经济决策[J]. 中国统计, 2017(4): 8-10.
 [11] 曹宏炳, 贾严冬. 战斗车辆乘员站综合集成设计[J]. 兵工自动化, 2018, 37(10): 13-16.
 [12] 胡迪·利普森, 梅尔巴·库曼. 无人驾驶[M]. 上海: 文汇出版社, 2017: 283.

业出版社, 2012: 40-41.
 [3] 曹小平, 路广安. 装备维修器材保障[M]. 北京: 国防大学出版社, 2005: 105.
 [4] 刘帅. 基于决策树和信息熵的属性约简算法研究[D]. 沈阳: 东北大学, 2010.
 [5] 邓自洋. 改进决策树算法在高效就业管理中的应用研究[D]. 上海: 华东理工大学, 2012.
 [6] 杨琦, 杨学强, 王张杰, 等. 基于决策树的装甲器材库存策略选择研究[J]. 军事物流, 2010(224): 141-143.
 [7] 黄泽波, 李占峰, 熊亮, 等. 基于故障树分析法与寻址技术的航空电源系统故障诊断系统[J]. 兵工自动化, 2018, 37(10): 51-56.
 [8] 石平. 基于决策树分类的成绩分析系统研究[J]. 长春师范大学学报, 2018(2): 56-59.

Electronics, 2012, 22(6): 948-953.
 [11] 汲万峰, 姜礼平, 朱建冲. 基于遗传算法的航路规划模型研究[J]. 军事运筹与系统工程, 2010, 24(2): 52-55.
 [12] 李子杰, 刘湘伟. 基于进化算法的多无人机协同航路规划[J]. 火力与指挥控制, 2015, 40(2): 85-89.
 [13] 田鹤, 李启华, 孟一鸣. 基于改进型遗传算法的舰艇航路规划研究[J]. 舰船电子工程, 2011, 31(10): 46-48.
 [14] 乔莎莎, 吴勇, 张建东, 等. 基于遗传算法和人工势场法的路径规划[J]. 现代电子技术, 2012, 35(12): 75-78.