

DOI: 10.7652/xjtuxb201802006

效能评估可信度的客观度量方法

王军^{1,2}, 李建勋², 王兴¹, 戚宗锋^{1,3}

(1. 中国洛阳电子装备试验中心, 471000, 河南洛阳; 2. 上海交通大学自动化系, 200240, 上海;
3. 电子信息系统复杂电磁环境效应国家重点实验室, 471000, 河南洛阳)

摘要: 针对目前装备系统效能评估可信度无法客观度量的问题, 提出了一种客观度量评估可信度方法。该方法从评估数据、评估模型及单样本分类置信度 3 个方面描述评估可信度。在计算评估数据可信度方面, 通过分析评估数据的不确定性确定评估数据的可信度; 在计算评估模型可信度方面, 将模式分类准确率转化为评估模型可信度度量指标; 在计算单样本分类置信度方面, 根据不同类别样本在高维空间的分布关系, 给出了适用多数分类算法的分类置信度计算方法; 进一步通过可信度传播模型将以上 3 个方面进行综合, 得到综合评估可信度。对雷达抗干扰效能评估可信度的仿真结果表明: 该方法不使用专家经验知识, 能够客观度量效能评估的可信度, 其结果不受评估主体影响, 解决了不同专家对同一装备给出的效能评估可信度不同的问题; 与现有主观方法相比, 准确率提高了 10%。

关键词: 效能评估; 可信度; 模式分类; 客观度量

中图分类号: TP18 **文献标志码:** A **文章编号:** 0253-987X(2018)02-0037-08

An Objective Measurement Method for Credibility of Effectiveness Evaluation

WANG Jun^{1,2}, LI Jianxun², WANG Xing¹, QI Zongfeng^{1,3}

(1. Luoyang Electronic Equipment Test Center of China, Luoyang, Henan 471000, China; 2. Department of Automation, Shanghai Jiao Tong University, Shanghai 200240, China; 3. State Key Laboratory of Complex Electromagnetic Environment Effects on Electronics and Information System, Luoyang, Henan 471000, China)

Abstract: An objective measurement method for the credibility of evaluation from three aspects: evaluation data, evaluation model and single sample confidence is proposed to solve the problem that effectiveness evaluation credibility cannot be objectively measured for equipment systems. The credibility of evaluation data is determined by analyzing the uncertainty of the evaluation data; and the model classification accuracy is transformed into a credibility index of the evaluation model. A method to calculate the classification confidence is proposed based on the distribution of the different types of samples in a space with high dimension and the method is applicable to most classification algorithms. These aspects are then integrated through a credibility propagation model to get a comprehensive credibility. Finally, the concrete application of the method is discussed in detail by taking the credibility calculation of effectiveness evaluation of a radar anti-jamming efficiency as an example. It is found from the analysis of the application that the method can objectively measure the credibility of effectiveness evaluation without the needs of the

收稿日期: 2017-09-13。 作者简介: 王军(1985—), 男, 硕士生; 李建勋(通信作者), 男, 教授, 博士生导师。 基金项目: 国家自然科学基金资助项目(61673265); 国家(973 计划)资助项目(61331903); 电子信息系统复杂电磁环境效应国家重点实验室开放课题(CEMEE2013K0301A)。

网络出版时间: 2017-12-05

网络出版地址: <http://kns.cnki.net/kcms/detail/61.1069.T.20171205.2118.004.html>

knowledge of experts, the results are not affected by evaluation subjects, and the problem that different evaluation experts give different credibility for the same equipment effectiveness evaluation is solved. The calculation accuracy of the proposed method is 10% higher than that of an existing subjective method.

Keywords: effectiveness evaluation; credibility analysis; classification; objective measurement

目前效能评估在装备鉴定、网络安全评估、电力系统性能评估^[1]等领域有着广泛的应用。评估可信度是研究评估工作中各环节的可信程度,而目前关于评估可信度研究的公开文献还很少。针对可信度的研究目前大多见于仿真系统,如文献[2-5]分别对重力坝仿真系统、战争模拟系统概率仿真模型、电信网络仿真模型以及移动自组织网络仿真系统的可信度进行了研究。然而,评估系统可信度不同于仿真系统可信度,其主要区别为:①效能评估系统是一个比仿真系统更为抽象的系统,构建仿真系统模型可参考真实系统,而构建效能评估模型无真实模型可以参考;②仿真系统的仿真结果可以与真实系统进行比较,而评估系统无法获取真实值。

效能评估可信度分析可以概括为3个方面:①效能评估过程的可信度分析;②效能评估稳健性实现方法;③效能评估结果可信度分析^[6]。在效能评估稳健性实现方法方面,随着机器学习方法的不断发展,大量的研究人员将机器学习算法应用到效能评估中,这在一定程度上解决了效能评估不稳健的问题。在效能评估过程及结果可信度分析方面,文献[7]给出了一种基于元评估的可信度分析方法,该方法通过描述评估过程是否科学严谨,分析评估过程可信度并给出评估结果的可信度结果,本质是一种主观评价方法。评估可信度直接关系到评估结论的使用,只有建立一种效能评估可信度客观度量方法,才能确保效能评估的真正价值。因此,本文研究评估可信度的客观度量方法,解决目前实际工程应用中的问题,以弥补效能评估中的不足。

1 效能评估与模式分类的关系

效能评估流程主要包括评估指标体系构建、指标数据采集、评估模型构建及评估结果分析等^[8],如图1所示。在指标体系构建时,要求所选取的每个指标值能描述同类系统不同对象间的区别;在指标数据采集时,通过试验采集相应的指标数据;在评估模型构建时,根据评估工作的需要,建立科学合理的评估模型;在评估结果分析时,通过分析评估结果指导评估工作前面环节的完善。

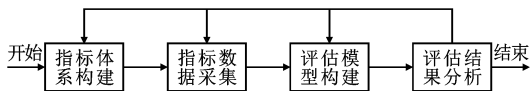


图1 效能评估主要流程

模式分类是指对表征事物或现象的各种形式信息进行处理,以对事物或现象进行描述、辨认、分类和解释的过程^[9]。模式是模式分类系统中的基本元素,根据研究对象的不同,将具有不同特征的对象看作不同的模式,很多问题可以采用模式分类的方法解决,例如故障诊断^[10-12]、图像检测^[13]等。典型模式分类流程包括数据采集、特征选择或抽取、分类器设计以及分类结果分析^[9],如图2所示。

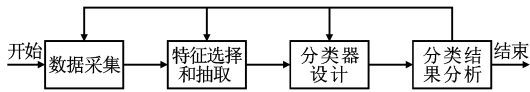


图2 典型模式分类流程

效能评估中的指标选择其实是选择不同的特征,这与模式分类中的特征选择或提取的本质是一样的。如果将不同的效能等级看作不同的模式,那么对被评估对象的效能评估过程也是一种模式分类的过程。因此,为了研究效能评估可信度,本文从模式分类的角度研究评估可信度问题。

2 评估模型可信度

2.1 模型性能评价方法

通过上面的分析可以看出,评估模型实际上可以看作是一个分类模型。对分类模型的性能评价常用的指标是分类准确率、运算速度、成本等,在对效能评估模型性能分析时由于运算速度、成本等本身并不影响分类结果的正确性,因此这里仅考虑模型分类准确率。

一般在二分类问题中,评价模型分类性能的直观方法是根据正负样本分类结果绘制 ROC 曲线和 AUC 曲线^[14-15]。在多分类模型中,无正负样本之分,但是仍然可以通过计算分类准确率描述分类模型的性能。多分类模型中,模型对不同类别样本的分类能力是不同的,因此需要分别计算模型对每类样本的分类准确率。第 i 类样本的分类准确率 P_{Ti}

的计算公式为

$$P_{Ti} = \frac{N_{Ti}}{N_i} \quad (1)$$

式中: N_{Ti} 表示 i 类样本被正确划分的个数; N_i 表示 i 类样本的总个数。

2.2 评估模型可信度

上面从分类的角度给出了评估模型性能评价方法。在评估系统中,模型的可信度代表着评估者对评估模型的认可程度,度量的是一种人主观想法的表现,不同的人对同一模型的认可程度不同,但是人的主观判断是基于客观数据的推断,因此只有基于客观数据通过科学的数学方法得出的可信度才更准确,同时也能够避免不同人对同一模型认可程度不同的问题。

在不确定性理论中,可信度是建立在可能性及必要性测度上的。假设某一事件 A 发生的可能性测度为 $P_{\text{pos}}(A)$,事件 A 发生的必要性测度为 $P_{\text{nec}}(A)$,那么事件 A 发生的可信度为

$$C(A) = \lambda P_{\text{pos}}(A) + (1 - \lambda) P_{\text{nec}}(A) \quad (2)$$

$$0 \leq \lambda \leq 1$$

式中: λ 为可能性和必要性平衡系数,通常可以认为可能性与必要性同等重要,此时 $\lambda = 1/2$ 。

本文借鉴不确定性理论中可信度的概念,讨论评估模型的可能性及必要性测度进而确定评估模型的可信度。事件 A 定义为通过评估模型将应属于 A 类的样本划分为 A 类,则事件 A 发生的可能性测度 $P_{\text{pos}}(A)$ 定义为事件 A 发生的概率,即

$$P_{\text{pos}}(A) = P_{TA} = \frac{N_{TA}}{N_A} \quad (3)$$

式中: P_{TA} 表示事件 A 发生的概率; N_{TA} 表示 A 类样本被正确分类个数; N_A 表示 A 类样本总数。

事件 A 发生的必要性定义为不属于 A 类的样本被划分到 A 类的不可能性,于是事件 A 发生的必要性测度 $P_{\text{nec}}(A)$ 定义为

$$P_{\text{nec}}(A) = 1 - \frac{N_{FA}}{N - N_A} \quad (4)$$

式中: N_{FA} 表示不属于 A 类的样本被错误分为 A 类的个数; N 表示样本总数。

由式(2)~式(4)可以得到事件 A 发生的可信度为

$$C(A) = \lambda P_{\text{pos}}(A) + (1 - \lambda) P_{\text{nec}}(A) = \lambda \frac{N_{TA}}{N_A} + (1 - \lambda) \left(1 - \frac{N_{FA}}{N - N_A} \right) \quad (5)$$

3 单样本分类置信度

对一待测样本,通过分类模型进行分类后,在不考虑分类结果是否正确,的情况下,该样本被划分为某类的可信性程度称为单样本分类置信度。单样本分类置信度仅代表模型将样本分类的结果和其相似样本分类结果一致性的度量,与分类结果是否正确无关。例如,样本 x_1 通过某评估模型得到的评估分类结果为 y_1 ,模型在对样本 x_1 评估时,经过计算认为 x_1 属于 y_1 的置信度为 95%,此时不管分类结果是否正确,样本的评估分类结果置信度就是 95%。

关于分类置信度的计算,文献[16-17]给出了一种最近邻分类器的分类置信度计算方法。文献[16]同时给出了 SVM 分类器分类置信度计算方法。综合各种分类算法的主要思想可知,分类器是通过样本间的距离将进行分类的,这种距离包括同类样本距离与异类样本距离,因此可以通过距离度量样本分类置信度。另外,从高维空间的角度分析,样本在高维空间的分布影响着分类的结果,对单个待测样本来讲,单样本与邻域样本中的同类样本分布关系和异类样本分布关系都影响其分类结果,这些关系包括样本数量和样本空间距离。影响样本分类置信度的因素如图 3 所示。

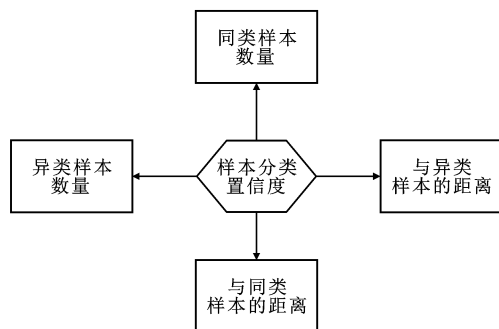


图3 影响样本分类置信度的因素

基于此,本文提出一种新的满足多数分类算法的单样本分类置信度计算方法。

定义1 设 $m_s(x)$ 为测试样本中与待测样本 x 同类的邻域样本个数, $m(x)$ 为测试样本中待测样本 x 的邻域样本总数, $D_s(x)$ 为待测样本 x 与邻域样本中所有同类样本的欧氏距离和, $D_d(x)$ 为待测样本 x 与邻域样本中异类样本的欧氏距离和。定义待测样本 x 的分类置信度为

$$f(x) = \frac{m_s(x)}{m(x)} \left[\exp\left(\frac{D_d(x)}{D_s(x)}\right) \right]^{-1} \quad (6)$$

特别规定:当待测样本邻域内无异类样本时

$D_d(x)=0$ 。显然,该置信度取值范围是 $0\sim 1$ 。

4 效能评估综合可信度度量方法

以上分析了评估模型可信度及待测样本置信度,对于待评估的对象,其评估可信度应包括被评估对象样本数据的可信度、模型可信度及样本分类置信度等 3 个方面。试验数据的可信度主要受试验活动中的不确定性影响,下面讨论试验数据的不确定性对样本数据可信度的影响及处理方法,以及如何将试验数据可信度、评估模型可信度及单样本分类置信度 3 个方面进行综合,得到最终评估可信度。

4.1 不确定性数据的处理

通常采用随机变量、模糊变量以及区间变量描述试验过程中的不确定性。对于随机变量和模糊变量可以直接计算不确定性测度,进而确定试验数据的可信度,或者将随机变量及模糊变量转化为区间数据再进行处理。文献[18]给出了模糊变量及随机变量的不确定性测度计算方法,本文不再讨论。本文重点研究包含区间变量的评估可信度度量问题。设 $x_i=(x_{i1},x_{i2},\cdots,x_{im})$ 为某装备的试验样本,假设样本分量 x_{i1} 的值是区间值,即 x_{i1} 取值范围是 $a\sim b$ 。在对样本 x_i 进行效能评估时可以分别采用区间上限及区间下限进行计算,此时计算结果会出现 2 种情况。一是采用区间上限和区间下限得到的评估分类结果一致,此时区间数据不影响被评估对象的效能等级划分,但是影响效能评估的可信度;二是采用区间上限及区间下限时得到的评估结果不同。进一步假设当取区间上限 b 时,评估结果为 y_1 ,当取区间的下限 a 时,评估结果为 y_2 ,此时需要分别计算该样本属于效能等级 y_1 与 y_2 的可信度。在已知 x_{i1} 在 $a\sim b$ 内的概率密度函数的情况下,假设通过判断 x_{i1} 取区间中某点 c 时,样本 $x_i=(x_{i1},x_{i2},\cdots,x_{im})$ 处在分界面上,此时将区间 $a\sim b$ 分为 $a\sim c$ 和 $c\sim b$,分别计算 $a\sim c$ 及 $c\sim b$ 的概率分布值,此时概率分布描述的是仅考虑试验数据时的评估可信度。

4.2 可信度传播模型

设评估系统数学模型为

$$D=f(X) \tag{7}$$

式中: D 为评估结果,即效能值; $X=\{x_1,x_2,\cdots,x_n\}$ 为评估系统的输入。

可信度传播分为模块间的串联和并联 2 种情况,可信度传播模型如图 4 所示。图 4a 代表的是一种串联情况,其可信度传播模型为

$$C(D)=C(Y)C[f(X)] \tag{8}$$

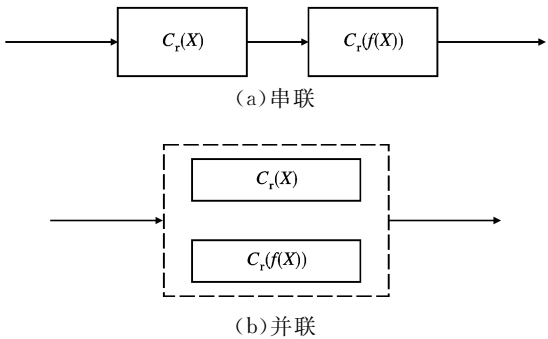


图 4 可信度传播模型

图 4b 代表的是一种并联情况,其可信度传播模型为

$$C(D)=\min\{C(X),C[f(X)]\} \tag{9}$$

效能评估可信度是评估数据、评估模型及样本置信度 3 个因素的累积影响,是一种串联情形。假设试验数据的置信度用 $C_d(x)$ 表示,模型的可信度用 $C_m(f)$ 表示,单样本的评估分类的置信度用 $f(x)$ 表示,那么待评估样本 x 的综合评估可信度为

$$C(x)=C_d(x)C_m(f)f(x) \tag{10}$$

上面分别讨论了试验数据的不确定性处理方法、评估模型可信度分析方法以及单样本分类置信度计算方法,并给出了一种简单的可信度传播模型。

4.3 综合可信度分析方法原理

从上面的分析可以给出综合效能评估可信度分析方法原理,如图 5 所示。

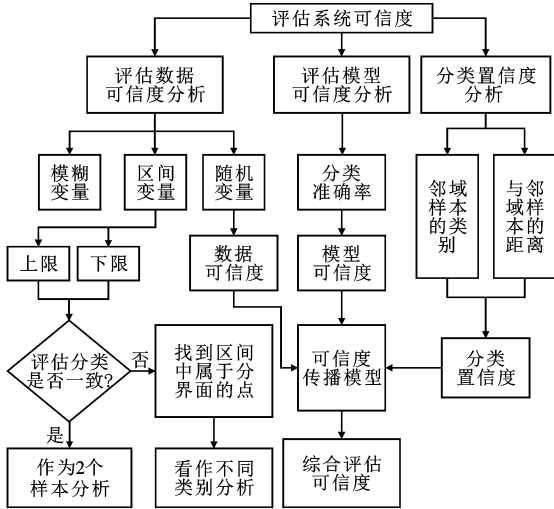


图 5 评估可信度分析方法原理

从图 5 可以看出,本文将评估可信度分为评估数据可信度、评估模型可信度以及单样本置信度 3 个方面。在计算评估数据可信度时,根据所包含的不确定性数据类型,直接计算可信度大小或者转化为区间数据进行处理;在计算评估模型可信度时,将

评估模型与模式分类模型建立联系,结合不确定性理论中的可信度概念,将分类性能评价指标转化为可信度指标;在计算单样本分类置信度时,结合现有分类置信度计算的主要思想,综合考虑邻近样本的类别与距离信息,确定待测样本分类置信度。最后,将影响评估可信度的 3 个方面通过简单的可信度传播模型进行综合,得到最终的评估可信度。

5 实例分析

以雷达抗干扰效能评估为例,选取抗干扰拖引

成功率(P_1)、抗假目标欺骗成功率(P_2)、发现概率(P_3)、跟踪误差(E)、抗干扰扇面有效度(θ)以及平局虚假航迹改善因子(I)等 6 个评估指标。通过仿真试验获取 140 组数据,将效能等级划分为 4 个等级(1、2、3、4)得到 140 个样本的仿真数据,其中部分组数据如表 1 所示。

选取 140 组样本中的 75 组为训练样本、65 组为测试样本。将测试样本中效能等级为 1 的样本集记为 S_1 ,共计 15 个;将效能等级为 2 的样本集记为 S_2 ,共计 12 个;将效能等级为 3 的样本集记为 S_3 ,

表 1 雷达抗干扰效能评估仿真数据

样本组号	抗距离拖引成功率/%	抗假目标欺骗成功率/%	发现概率/%	跟踪误差/m	抗干扰扇面有效度/(°)	平均虚假航迹改善因子/dB	效能等级
1	54.86	64.22	73.32	61.94	59.48	6.91	1
2	54.57	64.98	71.53	62.10	60.08	6.88	1
3	59.65	70.52	78.22	61.28	86.74	7.99	2
4	60.23	69.68	77.91	60.63	88.44	8.14	2
5	73.73	79.02	91.28	68.29	88.55	9.18	3
6	75.50	80.50	86.99	73.42	85.88	8.63	3
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
140	61.72	74.09	78.10	61.07	85.09	8.15	2

共计 18 个;将效能等级为 4 的样本集记为 S_4 ,共计 20 个。采用 SVM 评估模型进行计算,评估程序运行结果界面如图 6 所示,图中左侧为样本实际分类结果(类标签),右侧是运算错误结果。

Original Label is:1 Wrongly Predicted Label is 2
Original Label is:1 Wrongly Predicted Label is 2
Original Label is:2 Wrongly Predicted Label is 1
Original Label is:2 Wrongly Predicted Label is 4
Original Label is:3 Wrongly Predicted Label is 2
Original Label is:3 Wrongly Predicted Label is 4
Original Label is:3 Wrongly Predicted Label is 1
Original Label is:4 Wrongly Predicted Label is 2
Original Label is:4 Wrongly Predicted Label is 3
Original Label is:4 Wrongly Predicted Label is 3

图 6 评估程序运行结果界面

根据图 6 的计算结果及式(3)、(4),可得每一类样本分类结果的可能性及必要性分别为

$$\left. \begin{aligned} P_{\text{pos}}(S_1) &= \frac{N_{\text{TS1}}}{N_{S1}} = \frac{13}{15} \\ P_{\text{pos}}(S_2) &= \frac{N_{\text{TS2}}}{|S_2|} = \frac{5}{6} \\ P_{\text{pos}}(S_3) &= \frac{N_{\text{TS3}}}{N_{S3}} = \frac{5}{6} \\ P_{\text{pos}}(S_4) &= \frac{N_{\text{TS4}}}{N_{S4}} = \frac{17}{20} \end{aligned} \right\} \quad (11)$$

$$\left. \begin{aligned} P_{\text{nes}}(S_1) &= 1 - \frac{N_{\text{FS1}}}{N - N_{S1}} = 1 - \frac{2}{65 - 15} = \frac{24}{25} \\ P_{\text{nes}}(S_2) &= 1 - \frac{N_{\text{FS2}}}{N - N_{S2}} = 1 - \frac{4}{65 - 12} = \frac{49}{53} \\ P_{\text{nes}}(S_3) &= 1 - \frac{N_{\text{FS3}}}{N - N_{S3}} = 1 - \frac{2}{65 - 18} = \frac{45}{47} \\ P_{\text{nes}}(S_4) &= 1 - \frac{N_{\text{FS3}}}{N - N_{S3}} = 1 - \frac{2}{65 - 20} = \frac{43}{45} \end{aligned} \right\} \quad (12)$$

根据式(5),取 $\lambda=0.5$,得到模型对每类样本分类的可信度为

$$\left. \begin{aligned} C(S_1) &= \frac{1}{2}[P_{\text{pos}}(S_1) + P_{\text{nes}}(S_1)] = 0.91 \\ C(S_2) &= \frac{1}{2}[P_{\text{pos}}(S_2) + P_{\text{nes}}(S_2)] = 0.88 \\ C(S_3) &= \frac{1}{2}[P_{\text{pos}}(S_3) + P_{\text{nes}}(S_3)] = 0.90 \\ C(S_4) &= \frac{1}{2}[P_{\text{pos}}(S_4) + P_{\text{nes}}(S_4)] = 0.90 \end{aligned} \right\} \quad (13)$$

表 2 为试验采集的 3 组待评估装备样本的指标数据,其中部分指标无具体的数据,只能得到其取值的范围。 S_{T2} 中的 θ 以及 S_{T3} 中的 I 的试验数据是区间值,需要分别考虑区间上限及下限情况,判断其效

表 2 3 组待评估装备试验数据

样本	$P_1/\%$	$P_2/\%$	$P_3/\%$	E/m	$\theta/(^{\circ})$	I	效能等级
S_{T1}	85.40	92.11	92.80	50.87	84.07	9.81	2
S_{T2}	86.38	84.34	76.25	65.11	76.3~80.27	8.84	3
S_{T3}	83.86	85.29	80.85	74.17	60.87	8.79~8.85	3 或 4

能等级是否一致。经过分析, S_{T2} 中 θ 取区间上限及下限时,效能等级相同。 S_{T3} 中的指标 I 取区间下限时,效能等级为 3,取区间上限时,效能等级为 4,进一步通过区间搜索发现:当 S_{T3} 中的指标 I 取值为 8.81 时,效能等级为 3,取值为 8.82 时,效能等级为 4,于是近似认为 8.815 为 S_{T3} 样本的效能分界点。假设该参数在区间内取每一点值的概率相同,于是对于样本 S_{T3} 有:当指标 I 取值为 8.79~8.815 时,效能等级为 3,其概率大小为 0.42;当指标 I 取值为 8.815~8.85 时,效能等级为 4,其概率大小为 0.58。如果想要获取准确的效能等级,需要通过试验进一步确定指标 I 的精确值。

式(13)给出的模型对每类样本分类可信度在一定程度上描述了评估的可信度。下面进一步分析单样本的置信度。表 3 给出了测试样本中 S_{T1} 样本的邻域样本及 S_{T1} 样本与邻域样本的距离 d_{T1} 。

由式(6)得到 S_{T1} 样本的分类置信度为

$$f(S_{T1})=\frac{m_s(S_{T1})/m(S_{T1})}{\exp[D_d(S_{T1})/D_d(S_{T1})]}=0.68 \quad (14)$$

记 S_{T2} 中指标 I 取区间下限时样本为 S_{T21} ,记取区间上限时的样本为 S_{T22} ,分别计算 S_{T21} 和 S_{T22} 的在测试样本中的邻域样本及其与邻域样本的距离 d_{T21} 、 d_{T22} 样本,经计算测试样本中 S_{T21} 和 S_{T22} 的邻域样本相同,距离信息如表 4 所示。

由表 4 及式(6)可以得到 T_{21} 及 T_{22} 样本的分类

置信度分别为

$$f(S_{T21})=\frac{m_s(S_{T21})/m(S_{T21})}{\exp[D_d(S_{T21})/D_d(S_{T21})]}=0.837 \quad 8 \quad (15)$$

$$f(S_{T22})=\frac{m_s(S_{T22})/m(S_{T22})}{\exp[D_d(S_{T22})/D_d(S_{T22})]}=0.838 \quad 0 \quad (16)$$

由式(15)及式(16)的结果可以看出, S_{T2} 样本的区间上限和下限结果基本相同,而且效能评估分类结果相同,因此可以忽略指标 I 的区间数据对 S_{T2} 样本评估及评估可信度的影响。

对于 S_{T3} 样本可采用同样的方法进行分析,求得其分类置信度,这里不再赘述。

上面分析了区间数据对评估结果及评估可信度的影响,并得到了模型对每一类样本分类的可信度,同时得到了单样本的置信度。于是,根据串联模块可信度传播模型可以得到样本 S_{T1} 和 S_{T2} 的评估可信度为

$$C(S_{T1})=C(S_2)f(S_{T1})=0.88\times0.68=0.60 \quad (17)$$

$$C(S_{T2})=C(S_2)f(S_{T2})=0.90\times0.84=0.76 \quad (18)$$

式(17)和式(18)表示样本 S_{T1} 和 S_{T2} 所代表的装备效能评估结果为等级 2 和等级 3 的可信度分别为 60%和 76%,而 3 个装备通过蒙特卡洛仿真试验

表 3 S_{T1} 样本的邻域样本及相关数据信息

邻域样本序号	$P_1/\%$	$P_2/\%$	$P_3/\%$	E/m	$\theta/(^{\circ})$	I	效能等级	d_{T1}/m
1	86.41	87.36	84.40	55.80	87.34	85.22	2	76.2
2	80.27	78.42	75.26	66.29	75.50	80.80	4	76.6
3	91.26	87.53	87.25	57.40	82.15	85.79	2	76.9
4	87.98	86.36	90.17	56.09	84.03	86.59	2	77.2
5	81.01	84.94	77.53	72.48	78.23	81.89	3	77.4
6	89.96	85.94	88.34	57.05	83.99	86.94	2	77.8
7	82.62	86.89	76.96	70.05	76.97	83.16	2	78.0
8	81.64	86.05	81.85	70.46	82.79	84.33	2	78.1
9	81.95	85.69	78.63	68.28	77.60	84.13	2	78.2
10	89.67	85.86	90.90	57.23	80.93	87.39	2	78.3
11	89.29	84.61	83.78	57.29	82.31	87.67	2	78.3
12	85.47	83.53	84.31	61.98	82.66	87.15	2	78.3

表 4 S_{T2} 样本的邻域样本及相关数据信息

邻域样本序号	$P_1/\%$	$P_2/\%$	$P_3/\%$	E/m	$\theta/(\text{^\circ})$	I	效能等级	d_{T21}/m	d_{T22}/m
1	80.27	78.42	75.26	66.29	75.50	80.80	4	72.5	72.6
2	79.24	80.51	75.17	73.04	76.71	81.27	4	73.3	73.4
3	81.01	84.94	77.53	72.48	78.23	81.89	3	73.7	73.7
4	79.25	77.62	75.78	71.31	74.97	81.63	4	73.7	73.9
5	77.15	79.20	75.80	72.19	73.31	81.71	4	74.0	74.3
6	83.37	83.35	73.27	74.50	72.52	82.23	4	742	74.5
7	78.51	81.24	74.96	72.82	73.42	82.35	4	745	74.7
8	82.47	83.06	75.74	75.21	74.97	82.61	4	746	74.8
9	81.86	79.63	70.08	69.52	72.79	83.19	4	75.1	75.4
10	75.41	84.07	74.82	70.11	73.73	83.23	4	75.4	75.7
11	78.00	80.65	73.18	76.93	76.63	83.22	4	761	76.1
12	80.18	77.82	73.94	70.75	74.61	84.14	4	76.1	76.2

验证的效能等级分别为 2、3、4,其可信度分别为 0.6~0.68、0.7~0.79 以及 0.5~0.56。同时,邀请雷达领域 4 名专家给出的 3 个装备效能等级分别为 2、3、4,其综合可信性分别为 0.73、0.86、0.68。

从上面具体实例可以看出,本文提出的效能评估可信度度量方法并没有利用专家经验知识,避免了不同专家对同一被评估对象评估结果分析的不同,便于编程实现评估结果的自动分析,能够提高计算效率,且计算结果与蒙特卡洛仿真试验结果相符,取蒙特卡洛仿真结果中间值作为可信度基准值,经计算本文方法与专家主观判断结果相比,准确率提高了约 10%,证明了本文可信度客观度量方的有效性。

6 结 论

本文为解决评估可信度缺少客观度量方法的难题,在分析现有效能评估与模式分类相关性的基础上,创新性地将模式分类相关理论应用到效能评估中,提出了从试验数据可信度到评估模型可信度,再到单样本分类置信度 3 个方面客观度量评估可信度的方法。在评估模型可信度分析方面,结合不确定性理论给出了计算评估模型可信度的方法;在单样本分类置信度计算方面,给出了一种适用多数分类算法的相对置信度计算方法;同时本文还讨论了不确定性试验数据对评估可信度的影响及处理方法,并重点研究了区间值数据的处理方法,并给出了一种简单的可信度传播模型,基于可信度传播模型,可得到效能评估综合可信度。最后,以雷达抗干扰效能评估可信性度量为例,验证了本文提出的方法能够在不使用专家经验知识的情况下,客观地度量效能评估可信度,计算结果更加客观、准确,有效地解

决了评估可信度难以客观度量的问题,提高了装备试验鉴定效率。

参考文献:

[1] 王秀丽,张择策,侯雨伸. 主动配电网多维度静态安全评估 [J]. 西安交通大学学报, 2016, 50(8): 110-116.
WANG Xiuli, ZHANG Zece, HOU Yushen. Multidimensional static security assessment for active distribution network [J]. Journal of Xi'an Jiaotong University, 2016, 50(8): 110-116.

[2] CAO X, GU C. Risk analysis of gravity dam instability using credibility theory Monte Carlo simulation model [J]. Springer Plus, 2016, 5(1): 1-14.

[3] 李震,董鸿鹏,姜本清. 作战仿真概率模型可信性测试 [J]. 计算机与现代化, 2016(1): 82-86.
LI Zhen, DONG H P, JIANG B Q. Credibility test of probability model in warfare simulation system [J]. Computer & Modernization, 2016(1): 82-86.

[4] PAWLIKOWSKI K, JEONG H D J, LEE J S R. On credibility of simulation studies of telecommunication networks [J]. IEEE Communications Magazine, 2002, 40(1): 132-139.

[5] ANDEL T R, YASINSAC A. On the credibility of manet simulations [J]. Computer, 2006, 39(7): 48-54.

[6] 汪连栋,曾勇虎,申绪润. 电子信息系统复杂电磁环境效应研究路线图 [M]. 北京: 国防工业出版社, 2013: 8.

[7] 宋彦学,张志峰,齐立辉. 基于元评估的武器装备作战效能评估可信性研究 [J]. 火力与指挥控制, 2009 (S1): 128-131.
SONG Yanxue, ZHANG Zhifeng, QI Lihui. Research

- on the creditability of operational effectiveness evaluation with weapon equipment based on meta-evaluation [J]. Fire Control and Command Control, 2009(S1): 128-131.
- [8] 黄炎焱. 武器装备作战效能稳健评估方法及其支撑技术研究 [D]. 长沙: 国防科学技术大学, 2006: 22-25.
- [9] 姜斌, 黎湘, 王宏强, 等. 模式分类方法研究 [J]. 系统工程与电子技术, 2007, 29(1): 99-102.
- JIANG Bin, LI Xiang, WANG Hongqiang, et al. Methods for pattern classification [J]. Systems Engineering and Electronics, 2007, 29(1): 99-102.
- [10] 张西宁, 雷威, 李兵. 主分量分析和隐马尔科夫模型结合的轴承监测诊断方法 [J]. 西安交通大学学报, 2017, 51(6): 1-7.
- ZHANG Xining, LEI Wei, LI Bing. Bearing fault detection and diagnosis method based on principal component analysis and hidden Markov model [J]. Journal of Xi'an Jiaotong University, 2017, 51(6): 1-7.
- [11] WANG T, QI J, XU H, et al. Fault diagnosis method based on FFT-RPCA-SVM for cascaded-multilevel inverter [J]. ISA Transactions, 2016, 60: 156-163.
- [12] YIN Z, LIU J, KRUEGER M, et al. Introduction of SVM algorithms and recent applications about fault diagnosis and other aspects [C] // Proceedings of 2015 IEEE International Conference on Industrial Informatics. Piscataway, NJ, USA: IEEE, 2015: 550-555.
- [13] 张鹏, 陈湘军, 阮雅端, 等. 采用稀疏 SIFT 特征的车牌识别方法 [J]. 西安交通大学学报, 2015, 49(12): 137-143.
- ZHANG Peng, CHEN Xiangjun, RUAN Yaduan, et al. A vehicle classification technique based on sparse coding [J]. Journal of Xi'an Jiaotong University, 2015, 49(12): 137-143.
- [14] 张小利. 图像融合及其性能评估若干问题研究 [D]. 长春: 吉林大学, 2016: 115-119.
- [15] 王颖, 李金, 王磊, 等. 基于机器学习的 microRNA 预测方法研究进展 [J]. 计算机科学, 2015, 42(2): 7-13.
- WANG Ying, LI Jin, WANG Lei, et al. Research and progress of microRNA prediction methods on machine learning [J]. Computer Science, 2015, 42(2): 7-13.
- [16] 赵行. SVM 分类器置信度的研究 [D]. 北京: 北京邮电大学, 2010: 8-20.
- [17] 薛磊, 杨晓敏, 吴炜, 等. 一种基于 KNN 与改进 SVM 的车牌字符识别算法 [J]. 四川大学学报(自然科学版), 2006, 43(5): 1031-1036.
- XUUE Lei, YANG Xiaomin, WU Wei, et al. An algorithm based on KNN and improved SVM for license plate recognition [J]. Journal of Sichuan University (Natural Science Edition), 2006, 43(5): 1031-1036.
- [18] 柯宏发, 陈永光, 刘思峰. 电子装备试验数据的不确定性分析方法 [J]. 应用基础与工程科学学报, 2011, 19(4): 653-663.
- KE Hongfa, CHEN Yongguang, LIU Sifeng. New method of uncertainties analysis for electronic equipment test data [J]. Journal of Basic Science and Engineering, 2011, 19(4): 653-663.

(编辑 刘杨)

(上接第 29 页)

- [12] LIN Z, CHEN M, MA Y. The augmented Lagrange multiplier method for exact recovery of corrupted low-rank matrices [EB/OL]. (2013-10-18) [2015-12-10]. <https://arxiv.org/pdf/1009.5055v3.pdf>.
- [13] MAIRAL J. SPAMS: a sparse modeling software, v2.5 [EB/OL]. (2014-05-25) [2015-12-26]. <http://spams-devel.gforge.inria.fr>.
- [14] MARSZALEK M, LAPTEV I, SCHMID C. Actions in context [C] // Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2009: 2929-2936.
- [15] LIU J, LUO J, SHAH M. Recognizing realistic actions from videos in the wild [C] // Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2009: 1996-2003.
- [16] WANG Xiaofang, QI Chun, LIN Fei. Combined trajectories for action recognition based on saliency detection and motion boundary [J]. Signal Processing Image Communication, 2017, 57: 91-102.
- [17] CHO J, LEE M, CHANG H J, et al. Robust action recognition using local motion and group sparsity [J]. Pattern Recognition, 2014, 47(5): 1813-1825.
- [18] PENG X, QIAO Y, PENG Q. Motion boundary based sampling and 3D co-occurrence descriptors for action recognition [J]. Image and Vision Computing, 2014, 32(9): 616-628.
- [19] WU J, ZHANG Y, LIN W. Towards good practices for action video encoding [C] // Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2014: 2577-2584.

(编辑 刘杨)