http://bhxb.buaa.edu.cn jbuaa@buaa.edu.cn
DOI: 10.13700/j.bh.1001-5965.2024.0627

基于BN的民机驾驶舱显示界面可用性综合评价模型

肖瑶[1]，陈登凯[1,*]，周垚[1]，肖江浩[1]，侯玉宏[1,2]

(1. 西北工业大学 工业设计与人机工效工信部重点实验室，西安 710072；2. 中国飞行试验研究院，西安 710089)

摘 要：研究民机驾驶舱显示界面可用性有助于优化设计界面，针对民机驾驶舱显示界面可用性评价的不确定性，提出一种基于贝叶斯网络(BN)的可用性综合评价模型。基于执行过程交互控制(EPIC)模型获取的可用性评价指标体系建立贝叶斯网络，定量根节点以专家眼动度量指标的加权评价为先验概率，定性根节点采用专家三角模糊数评价为先验概率；为降低专家的决策次数，使用Noisy-OR模型填充贝叶斯网络的条件概率表，并利用GeNIe软件搭建模型。以主飞行显示器界面进行实例验证，结果表明：模型因果推理结果与系统可用性量表(SUS)评价结果一致，该界面的评价等级均为“良”，验证所提模型的有效性，并对模型敏感性分析得到的敏感因子提出改进措施。

关 键 词：民机驾驶舱；显示界面；可用性评价；贝叶斯网络；眼动
中图分类号：V241.0
文献标志码：A **文章编号**：1001-5965(2026)11-4063-12

民机驾驶舱显示界面是飞行员与飞机之间的人机交互接口与核心枢纽，旨在向飞行员提供飞机的位置、速度、高度及燃油状态等基本飞行信息和各系统的状态信息。影响航空飞行事故关键致因归结为4类，即人为因素、环境因素、机械故障因素和管理因素[1]。国际民航组织(International Civil Aviation Organization, ICAO)发布的事故报告显示，人为因素造成的飞行事故占总事故数的70%，人为差错是危及世界航空安全的主要因素[2]。因此，对民机驾驶舱显示界面的可用性评价展开研究，让设计师更好地了解显示界面与飞行员认知需求的匹配度，为后续的迭代优化提供改进建议和指导有重要意义。

软件质量模型(ISO/ICE 9126)[3]规定“可用性是在特定使用环境下，用户达成预期目标的满意程度和效率”。针对民机驾驶舱显示界面的可用性研究中，Xue等[4]构建了由出错、可学习性和满意度等9项指标组成的可用性评价模型，采用多元回归法得到各指标的权重系数，对各指标的可用性评分值进行加权后得到显示界面的可用性结果。张惠琳[5]从用户行为、用户认知和设计样式维度提炼民机驾驶舱显示界面的可用性评价指标模型，利用层次分析法(analytic hierarchy process, AHP)确定不同指标的权重系数后，对所有指标的可用性评分值进行加权，以获取显示界面的可用性结果。现有民机驾驶舱显示界面的研究忽略了飞行员与显示界面交互过程的整体考量，而且可用性评价模型没有对定量指标和定性指标进行分类和差异化评价。此外，显示界面可用性评价方法可分为主、客观评价和综合评价。主观评价是通过模型、可用性原则、实际问题来构建显示界面可用性评价体系，并利用可用性量表来分析界面可用性，常用的可用性量表有系统可用性量表(system usability scale, SUS)[6]、后测系统可用性问卷(post-study system usability, PSSUQ)[7]、NASA-TLX量表[8]等。客观评价是通过脑电信号、皮肤信号、绩效数据和眼动数据等进行分析，其中，

收稿日期：2024-08-28；**录用日期**：2024-09-23；**网络出版时间**：2024-10-16 14:38
网络出版地址：link.cnki.net/urlid/11.2625.V.20241016.1259.005
基金项目：陕西省重点研发计划(2022GY-311)
*通信作者. E-mail：chendengkai@nwpu.edu.cn

引用格式：肖瑶，陈登凯，周垚，等. 基于BN的民机驾驶舱显示界面可用性综合评价模型[J]. 北京航空航天大学学报，2026，52(11)：4063-4074. Xiao Y，Chen D K，Zhou Y，et al. BN-based comprehensive evaluation model for display interface usability in civil aircraft cockpit[J]. Journal of Beijing University of Aeronautics and Astronautics，2026，52(11)：4063-4074 (in Chinese).

眼动数据可有效反映用户与图形界面交互背后的深层机制，常用于检验界面设计的可用性，如李亚鸿等[9]通过眼动追踪可用性实验对比数字孪生系统优化前后设计界面实例，验证了基于质量屋优化设计方法的可行性。综合评价整合了主观评价与客观评价，如Chui等[10]结合心率、眼动实验和NASA-TLX量表来分析老年人医疗保健系统界面的可用性。近年来，还有学者采用各种理论方法对可用性评价展开研究，如江涛等[11]运用模糊层次分析法构建工业缝纫机可用性评价模型。在现阶段的可用性评价研究中，尚未有将显示界面可用性的综合评价方法与贝叶斯网络（Bayesian network, BN）相结合的尝试。因为，贝叶斯网络依托于概率论原理，即使缺乏相关历史数据支持，也能有效地推断出民机驾驶舱显示界面的可用性评价结果。贝叶斯网络是一种概率图模型，节点表示变量，有向边表示链接节点之间的依赖关系，条件概率可直接量化该依赖关系[12]。贝叶斯网络由一个有向无环图（directed acyclic graph, DAG）和若干条件概率表（conditional probability table, CPT）组成，DAG的结构和CPT的内容共同构成了贝叶斯网络的模型。CPT包含每个节点在给定其父节点状态的条件下，该节点各个状态的概率分布。贝叶斯网络的先验概率是在无领域知识或历史数据情况下确定的，后验概率是在获取新信息后对先验概率进行修正的结果。

针对民机驾驶舱显示界面可用性评价的复杂性和多维性，利用贝叶斯网络在不确定系统中的概率推理能力，是弥补现有可用性评价研究不足的关键途径。因此，本文提出一种基于贝叶斯网络的民机驾驶舱显示界面可用性综合评价模型，还借助GeNIe软件完成评价模型的搭建，最后以实例分析验证了该模型的有效性，并针对实例的敏感因子提出改进措施。

1 民机驾驶舱显示界面可用性评价指标体系

执行过程交互控制（executive process-interactive control, EPIC）模型是人机交互领域经典的认知心理学模型，用于研究人机界面的可用性及用户的认知过程[13]。EPIC模型示意如图1所示，基于EPIC模型对用户与界面的交互过程分析，本文提出从飞行任务、飞行员感官、飞行员认知和飞行交互4个维度对民机驾驶舱显示界面可用性进行综合评价。

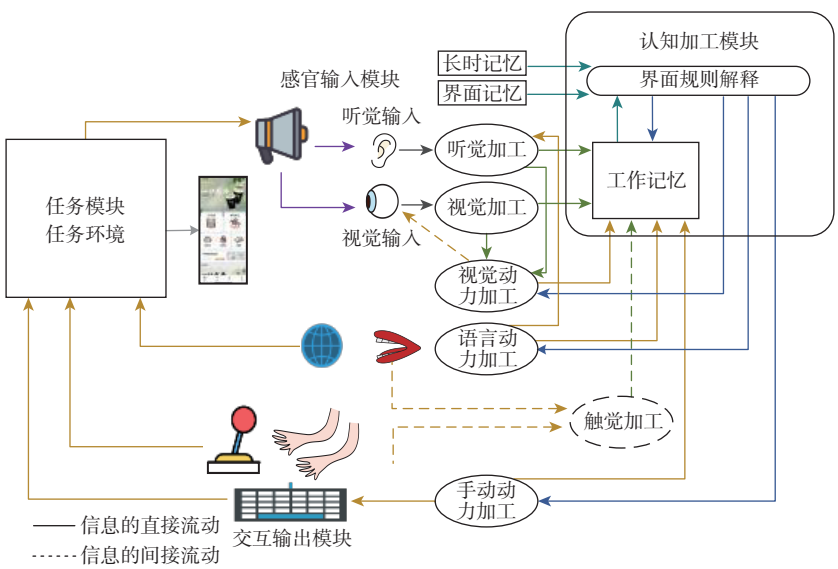


图1 EPIC认知加工模型

Fig. 1 EPIC cognitive processing model

整合国内外相关文献、行业标准与规范，完成民机驾驶舱显示界面可用性评价指标的选取工作。通过CNKI数据、Web of Science平台、万方数据等平台，以“民机驾驶舱”“显示界面”“指标体系”“可用性”等中英文关键词进行检索，获取4个维度的可用性评价初选指标，并结合Nielsen模型[14]和ISO/ICE 9126[3]对可用性评价指标进行补充和修正，最终梳理出44个民机驾驶舱显示界面可用性评价初选指标，如表1所示的指标层和删除指标。

邀请某飞行研究院的5位飞行专家，通过德尔菲法完成指标筛选工作。经过筛选后，保留了16个可用性评价指标，如表1所示的指标层。本文还运用层次结构模型，构建民机驾驶舱显示界面可

用性评价指标体系。表 1 列出了指标体系的目标层、维度层、指标层、指标语义、指标类型、删除指标和指标来源。在指标类型中, 定量指标可采用实验方法收集和处理数据资料做出定量结果的价值判断, 定性指标可根据专家的知识、经验直接做出定性结论的价值判断[15]。

表 1　民机驾驶舱显示界面可用性评价指标体系

Table 1　Usability indicator system for cockpit display interfaces of civil aircraft

目标层	维度层	指标层	指标编号	指标语义	指标类型	删除指标	指标来源
民机驾驶舱显示界面可用性 C^*	飞行任务 C_1	任务导向性	c_{11}	界面是否引导飞行员进行操作	定量	时间压力、任务适应性、任务优先级、任务匹配、任务完成时间、任务熟悉度、任务导航性、任务流畅度	文献[14]
		信息吸引力	c_{12}	界面信息是否吸引飞行员执行任务			
		信息复杂度	c_{13}	界面信息是否降低飞行任务复杂度			
		任务完成效率	c_{14}	界面是否能提高飞行员的任务完成度			
	飞行员感官 C_2	一致性	c_{21}	界面色彩、图标等方面是否有一致性	定性	层次性、均衡性、语音清晰度、音频可调性、音频优化、语音告警提示、语音确认操作	文献[14]
		简洁性	c_{22}	界面是否不包括重叠、冗余信息			
		信息呈现	c_{23}	界面呈现方式是否科学合理			
		信息易读性	c_{24}	界面信息是否易于飞行员阅读和查阅			
	飞行员认知 C_3	理解	c_{31}	界面是否能让飞行员快速地理解	定性	学习时长、熟悉、学习效率、回忆、注意力、可学习性	文献[3,16-17]
		易学性	c_{32}	界面是否降低飞行员学习使用的难度			
		满意度	c_{33}	界面是否能提高飞行员的视觉满意度			
		最小记忆负荷	c_{34}	界面是否尽量减少飞行员的记忆负荷			
	飞行交互 C_4	信息实时性	c_{41}	界面信息是否实现实时更新	定性	可预测性、交互规范性、交互及时响应、防撞预警、告警提示、信息容错、更改设置	文献[18-19]
		信息精确性	c_{42}	界面信息是否精确无误			
		错误预防	c_{43}	界面是否提供错误预防与恢复机制			
		交互反馈	c_{44}	界面是否提供明确的交互信息反馈			

2　民机驾驶舱显示界面可用性综合评价模型构建

基于贝叶斯网络的民机驾驶舱显示界面可用性综合评价模型框架如图 2 所示, 步骤如下:

步骤 1　根据民机驾驶舱显示界面可用性评价指标体系建立贝叶斯网络模型。

步骤 2　贝叶斯网络定量根节点的先验概率由眼动追踪的加权客观评价获取, 定性根节点的先验概率由三角模糊主观评价获取。

步骤 3　为完善贝叶斯网络模型搭建, 利用 Noisy-OR 模型填充贝叶斯网络节点的模糊条件概率表后, 计算获取非根节点的条件概率。

步骤 4　利用贝叶斯网络进行因果推理、诊断

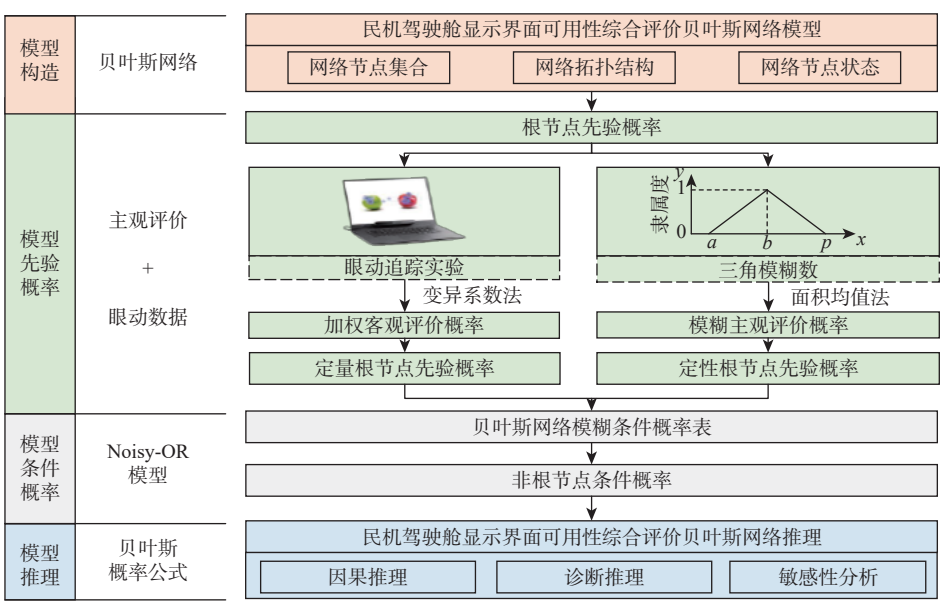


图 2　基于贝叶斯网络的民机驾驶舱显示界面可用性综合评价模型

Fig. 2　Comprehensive evaluation model for cockpit display interfaces of civil aircraft based on Bayesian network

推理和敏感性分析。

2.1 贝叶斯网络构造

根据民机驾驶舱显示界面可用性评价指标体系确定贝叶斯网络节点，节点$E=\{E_1,E_2,E_3\}$主要分为3种：①E_1为根节点，是贝叶斯网络的基本单元和底层节点，对应体系中指标层的民机驾驶舱显示界面可用性评价指标，根节点变量为$c_{im}(i=1,2,3,4;m\in N^*)$；②$E_2$为中间节点，起到连接贝叶斯网络的根节点与目标节点作用，对应维度层的4个维度，中间节点变量为$C_i(i=1,2,3,4)$；③E_3为叶节点，表示民机驾驶舱显示界面可用性评价等级，对应体系的目标层，叶节点变量为C^*。在网络节点确定的基础上，由根节点E_1指向中间节点E_2、中间节点E_2指向叶节点E_3，搭建如图3所示的民机驾驶舱显示界面可用性评价贝叶斯网络模型。假设根节点E_1及中间节点E_2由“STATE=0”和“STATE=1”2个状态组成，用$j\in\{0,\ 1\}$来表示，“STATE=0”表示对民机驾驶舱显示界面可用性的负面影响，“STATE=1”表示对民机驾驶舱显示界面可用性的积极影响，并将叶节点E_3的节点状态定义为$G_\beta=\{G_1,G_2,G_3,G_4\}=$ {优,良,中,差}。

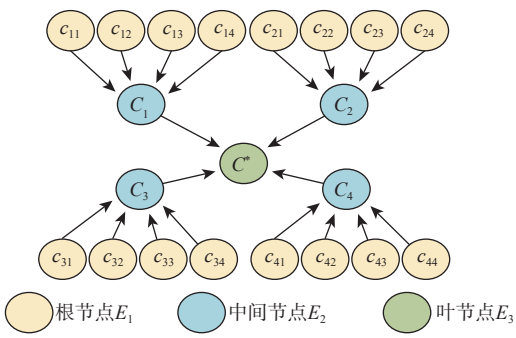


图3 民机驾驶舱显示界面可用性评价贝叶斯网络模型

Fig. 3 Bayesian network modelling for usability evaluation in cockpit display interfaces of civil aircraft

2.2 根节点先验概率获取

2.2.1 定量根节点先验概率

由飞行任务C_1中间节点对应的根节点组成定量根节点集$K=\{c_{11},c_{12},c_{13},c_{14}\}$，其中，定量根节点$c_{im}(i=1)$同样具有“STATE=0”和“STATE=1”2种状态，用$j\in\{0,\ 1\}$来表示。

眼动追踪技术通过记录被试的眼球运动轨迹注视行为，获取有关视觉注意、搜索行为等方面的数据，判断显示界面的信息复杂性、吸引性、引导性和搜索效率。设全部眼动度量指标集$Y=\{Y_1,Y_2,\cdots,Y_e\cdots,Y_{10}\}$，$e=1,2,\cdots,10$，定量根节点的度量指标分析如表2所示。

在此基础上，根据眼动度量指标的实验数据计算各根节点的先验概率，具体步骤如下：

表2 定量根节点的度量指标分析

Table 2 Metrics analysis of quantitative root nodes

定量根节点	度量指标Y	正负性
任务导向性c_{11}	眼跳次数Y_1	负相关
	扫描路径时间Y_2	负相关
	扫描路径长度Y_3	负相关
信息吸引力c_{12}	AOI注视总持续时间Y_4	正相关
	AOI内注视次数Y_5	正相关
	AOI首次注视持续时间Y_6	正相关
信息复杂度c_{13}	注视点数量Y_7	负相关
	回视次数Y_8	负相关
任务完成效率c_{14}	任务准确率Y_9	正相关
	任务完成时间Y_{10}	负相关

注：兴趣区域(area of interest, AOI)。

步骤1 设参与实验的专家总数为s，收集定量根节点$c_{im}(i=1;m\in N^*)$处于状态j下s名专家的所有度量指标集Y的实验初始数据，构建眼动初始数据矩阵：

$$\boldsymbol{Y}=\begin{bmatrix} Y_{11} & Y_{12} & \cdots & Y_{1e} & \cdots & Y_{1s} \\ Y_{21} & Y_{22} & \cdots & Y_{2e} & \cdots & Y_{2s} \\ \vdots & \vdots & & \vdots & & \vdots \\ Y_{t1} & Y_{t2} & \cdots & Y_{te} & \cdots & Y_{es} \\ \vdots & \vdots & & \vdots & & \vdots \\ Y_{s1} & Y_{s2} & \cdots & Y_{se} & \cdots & Y_{s10} \end{bmatrix} \tag{1}$$

式中：$Y_{te}(e=1,2,\cdots,10;\ t=1,2,\cdots,s)$为第$t$名专家的第$e$个度量指标的原始数据。

针对正相关的度量指标原始数据，标准化公式：

$$r_{te}=\frac{Y_{te}-\min(Y_e)}{\max(Y_e)-\min(Y_e)} \tag{2}$$

针对负相关的度量指标原始数据，标准化公式：

$$r_{te}=\frac{\max(Y_e)-Y_{te}}{\max(Y_e)-\min(Y_e)} \tag{3}$$

式中：$\max(Y_e)$、$\min(Y_e)$分别为同一度量指标下s名专家原始数据的最大值、最小值；r_{te}为第t名专家的第e个度量指标原始数据归一化后的标准值。

步骤2 变异系数法是一种定量赋权法，其核心思想在于通过计算数据的标准差与平均值之比得出相对变异的度量，从而提供一种相对于平均值的标准化评估。若某个定量根节点$c_{im}(i=1;m\in N^*)$中度量指标数为ρ，利用变异系数法求某一定量根节点$c_{im}(i=1;m\in N^*)$所含度量指标权重的步骤如下：

$$\bar{r}_\rho=\sum_{t=1}^{s}r_{t\rho}/s \tag{4}$$

$$\sigma_\rho=\sqrt{\frac{\sum_{t=1}^{s}(r_{t\rho}-\bar{r}_\rho)^2}{s-1}} \tag{5}$$

$$\varphi_\rho=\bar{r}_\rho/\sigma_\rho \tag{6}$$

$$\delta_\rho = \varphi_\rho / \sum \varphi_\rho \tag{7}$$

式中: $\bar{r}_\rho$为某个定量根节点$c_{im}(i=1)$包含的第ρ个度量指标的平均值; σ_ρ为某一定量根节点$c_{im}(i=1)$包含的第ρ个度量指标的标准差; φ_ρ为第ρ个度量指标的变异系数; δ_ρ为第ρ个度量指标的权重, $0 \leqslant \delta_\rho \leqslant 1$, 且$\sum \delta_\rho = 1$。

步骤 3　已知某个定量根节点$c_{im}(i=1)$所含度量指标权重, 其加权评价概率:

$$P(c_{1mj}) = \sum \bar{r}_\rho \delta_\rho \tag{8}$$

此时, $P(c_{1mj})$为贝叶斯网络定量根节点$c_{im}(i=1)$处于状态j的先验概率。

2.2.2　定性根节点先验概率

飞行员感官C_2、飞行认知C_3和飞行交互C_4维度的根节点组成定性根节点集$X=\{c_{21},c_{22},\cdots,c_{im},\cdots,c_{43},c_{44}\}(i=2,3,4;m\in N^*)$, 其中, 定性根节点$c_{im}(i=2,3,4;m\in N^*)$有 2 种状态“STATE=0”和“STATE=1”, 用$j\in\{0,\ 1\}$表示。

三角模糊数是模糊集理论中将模糊、不确定的语言变量转化为确定数值的方法, 在评价中能解决评价对象无法准确度量、只能用自然语言进行模糊评价的问题。因此, 本文引入三角模糊数, 结合专家定性意见获取定性根节点的模糊发生概率。三角模糊数$\tilde{P}(x)=(a,b,q)$的隶属度函数表达式为

$$\tilde{P}(x)=\begin{cases}0 & x<a\\ (x-a)/(b-a) & a\leqslant x<b\\ 1 & x=b\\ (q-x)/(q-b) & b<x\leqslant q\\ 0 & x\geqslant q\end{cases} \tag{9}$$

基于专家评价与三角模糊数, 计算定性根节点的先验概率, 具体步骤如下:

步骤 1　在无法准确获取根节点发生概率的情况下, 为减少专家在定性节点状态评估时的定性差异, 可使用模糊语言变量来量化专家经验的判断结果。假设定性根节点事件发生概率的专家模糊评价集$Q(c_{im})=\{$非常低 (VL), 低 (L), 较低 (FL), 中等 (M), 较高 (FH), 高 (H), 非常高 (H)$\}(i=2,3,4;\ m\in N^*)$, 与三角模糊数$\tilde{P}(x)=(a,b,q)$的对应关系如表 3 所示。

步骤 2　采用问卷方式获取多名专家对定性根节点集X发生概率的三角模糊评价, 弥补了无法利用数据资源获得定性根节点概率的不足。已知专家总数为s, 第$t(t\in s)$名专家给出的定性根节点$c_{im}(i=2,3,4;m\in N^*)$处于状态j概率的模糊语言变量$Q_j^t(c_{im})$可根据表 3 中语言变量与三角模糊数的对应关系转换为三角模糊概率表达式:

表 3　7 个语言变量与三角模糊数的对应关系

Table 3　Corresponding relationship between 7 linguistic variables and triangular fuzzy numbers

序号	语言变量	$\tilde{P}(x)=(a,b,q)$
1	非常低(VL)	(0,0,0.1)
2	低(L)	(0,0.1,0.3)
3	较低(FL)	(0.1,0.3,0.5)
4	中等(M)	(0.3,0.5,0.7)
5	较高(FH)	(0.5,0.7,0.9)
6	高(H)	(0.7,0.9,1.0)
7	非常高(VH)	(0.9,1.0,1.0)

$$\tilde{P}_{imj}^t=(a_{imj}^t,b_{imj}^t,q_{imj}^t)\qquad t=1,2,\cdots,s \tag{10}$$

步骤 3　考虑到不同专家的经验知识对定性根节点评价的准确性, 需要对影响专家意见的教育经历和相关专业等因素进行权重判断。本文在综合s名专家意见时, 参考文献 [20] 按照不同的年龄、学历、职称和工作年限对专家意见进行加权, 专家权重如表 4 所示。

表 4　专家权重

Table 4　Weight of experts

项目	类别	权重	项目	类别	权重
年龄	⩾50岁	$w_1=1.0$	职称	机长	$w_3=1.0$
	40～50岁	$w_1=0.9$		副机长	$w_3=0.9$
	30`40岁	$w_1=0.8$		飞行学员	$w_3=0.8$
	⩽30岁	$w_1=0.7$		研究人员	$w_3=0.7$
学历	博士	$w_2=1.0$	工作年限	⩾20年	$w_4=1.0$
	硕士	$w_2=0.9$		10～20年	$w_4=0.9$
	学士	$w_2=0.8$		5～10年	$w_4=0.8$
	专科	$w_2=0.7$		⩽5年	$w_4=0.7$

此时, 定性根节点$c_{im}(i=2,3,4;m\in N^*)$处于状态j下s名专家的综合评价三角模糊概率计算式:

$$\tilde{P}_{imj}=(a_{imj},b_{imj},q_{imj})=\frac{\sum_{t=1}^{s} w_{1t}w_{2t}w_{3t}w_{4t}\tilde{P}_{imj}^t}{\sum_{t=1}^{s} w_{1t}w_{2t}w_{3t}w_{4t}}\qquad t=1,2,\cdots,s \tag{11}$$

式中: w_{1t}为第t名专家的年龄权重; w_{2t}为第t名专家的学历权重; w_{3t}为第t名专家的职称权重; w_{4t}为第t名专家的工作年限权重; $\tilde{P}_{imj}^t$为第t名专家对定性根节点c_{im}处于状态j下发生概率的模糊语言所转化的三角模糊概率。

步骤 4　由于三角模糊概率不利于贝叶斯网络的推理计算, 应将综合专家评价三角模糊概率进行解模糊处理。本文采取面积均值法将综合专家评价三角模糊概率转化为精确概率, 定性根节点$c_{im}(i=2,3,4;m\in N^*)$处于状态j的精确概率:

$$P_{imj}=(a_{imj}+2b_{imj}+q_{imj})/4 \tag{12}$$

经过去模糊处理后，为满足同一定性根节点在各状态下的概率之和为1，需对各定性根节点所有状态的精确概率进行归一化处理。经过归一化后，定性根节点$c_{im}(i=2,3,4;m\in N^*)$处于状态j的先验概率：

$$P(c_{imj})=P_{imj}\Big/\sum_{j=0}^{1}P_{imj} \tag{13}$$

2.3 非根节点条件概率获取

已知贝叶斯网络存在节点A、B且A为B的父节点，节点B的条件概率为该节点在父节点A所有状态组合下处于不同状态的概率，即在节点A发生的条件下节点B发生的条件概率：

$$P(B|A)=\frac{P(B)}{P(A)}\qquad P(A>0) \tag{14}$$

非根节点(包括中间节点和叶节点)的条件概率通过三角模糊隶属度函数和7级语言变量获取。贝叶斯网络的条件概率存在指数爆炸问题。对于有n个父节点的子节点，如果每个父节点有2种状态，则需要2^n个条件概率参数，增加了专家的工作负荷，还会累积评价误差。为解决上述问题，本文运用Noisy-OR模型来填充CPT，简化条件概率获取步骤并减少专家决策次数。通过获取专家在单一父节点影响下的条件概率，便可计算多个父节点对同一子节点的综合影响。

假定二元变量Y具有n个父节点X，每个节点X存在0和1(事件不发生为“0”，事件发生为“1”)2种状态，Noisy-OR模型如图4所示。

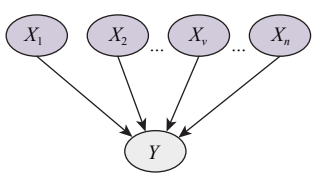


图4 Noisy-OR模型

Fig. 4 Noisy-OR model

根据贝叶斯网络条件概率公式，由事件$X_v(v\in 1,2,\cdots,n)$发生独立引起事件Y发生的概率：

$$P(Y\leftarrow X_v)=P(Y|\overline{X}_1,\ \overline{X}_2,\cdots,X_v,\cdots,\overline{X}_n) \tag{15}$$

式中：X_v表示事件v发生，$\overline{X}_1$、$\overline{X}_2$、$\overline{X}_n$表示事件1、2、n不发生；$P(Y\leftarrow X_v)$表示在仅考虑父节点X_v发生、其余父节点均不发生条件下，子节点Y的发生概率。

令父节点$X_v\in(X_1,X_2,\cdots,X_n)$构成集合$X_a$，在获取$n$个父节点独立影响子节点$Y$的条件概率后，子节点$Y$受$n$个父节点$X$共同影响下的条件概率：

$$P(Y|X_p)=1-P(Y\leftarrow X_p)=1-\prod_{v:\ X_v\in X_p}(1-P(Y\leftarrow X_v)) \tag{16}$$

2.4 贝叶斯网络推理

2.4.1 因果推理

基于贝叶斯网络的条件概率表，可依次求得中间节点与根节点处于不同状态的先验概率。对贝叶斯网络进行正向因果推理，可得到民机驾驶舱显示界面可用性评价等级状况。

在贝叶斯网络中，假设存在节点A、B_1、B_2、$\cdots$、$B_n(B_i\in B_n)$，$P(B_i)>0$，$i=1,2,\cdots,n$，则节点A的全概率公式表示为

$$P(A)=\sum_{i=1}^{n}P(B_i)P(A|B_i) \tag{17}$$

叶节点C^*民机驾驶舱显示界面可用性评价等级为$G_\beta=\{G_1,G_2,G_3,G_4\}=\{$优,良,中,差$\}$，利用贝叶斯网络的全概率公式计算在中间节点$C_i(i=1,2,3,4)$影响下的叶节点$C^*$在$G_\beta(\beta$=1,2,3,4)状态的全概率：

$$P(C^*=G_\beta)=\sum_{i=1}^{4}P(C_i)P(C^*=G_\beta|C_i) \tag{18}$$

将叶节点C^*的全概率$P(C^*=G_\beta)(\alpha$=1,2,3,4)数值进行对比，数值最大的节点状态即为民机驾驶舱显示界面可用性评价等级。本文采用金字塔等级来划分民机驾驶舱显示界面可用性评价等级$G_\beta(\beta=1,2,3,4)$的[0,1]区间，等级区间如表5所示。

表5 民机驾驶舱显示界面可用性评价等级区间

Table 5 Grade scale for cockpit display interface of civil aircraft usability evaluation

目标	评价等级G_β	等级区间
民机驾驶舱显示界面可用性评价等级	G_1=优	[0.95,1.00]
	G_2=良	[0.85,0.95)
	G_3=中	[0.70,0.85)
	G_4=差	[0.00,0.70)

2.4.2 诊断推理

诊断推理是从结果到原因的逆向推理，即叶节点C^*在可用性评价等级为“G_4=差”状态的发生概率为1时，利用贝叶斯网络的结构和概率分布逆向推理根节点不同状态的后验概率，获取可能的致因链、存在大影响力的父节点。

在贝叶斯网络中，若节点A为B的父节点，贝叶斯公式描述了节点A、B先验概率和后验概率的关系，其表达式为

$$P(A|B)=\frac{P(B|A)P(A)}{P(B)} \tag{19}$$

叶节点C^*在可用性评价等级G_4下，根节点

$c_{im}(i=1,2,3,4;m\in N^*)$在状态 $j=1$ 时的后验概率:

$$P\left(c_{im}=c_{im}^{j=1}|C^*=G_4\right)=\frac{P\left(C^*=G_4|c_{im}=c_{im}^{j=1}\right)P\left(c_{im}=c_{im}^{j=1}\right)}{P(C^*=G_4)} \tag{20}$$

此时, 根节点 $c_{im}(i=1,2,3,4;m\in N^*)$在状态 $j=0$ 时的后验概率:

$$P\left(c_{im}=c_{im}^{j=0}|C^*=G_4\right)=1-P\left(c_{im}=c_{im}^{j=1}|C^*=G_4\right) \tag{21}$$

2.4.3　敏感性分析

敏感性分析是指从众多不确定性父节点因素中找出对目标节点有重要影响的敏感因子, 即研究参数变化对目标节点条件概率所造成的影响。在贝叶斯网络诊断推理基础上, 采用风险促进值重要度 W、风险降低值重要度 R 和变化横向重要度 M 来衡量根节点 $c_{im}(i=1,2,3,4;m\in N^*)$对可用性评价等级的影响程度计算公式:

$$W(c_{im})=\frac{\max\left\{P\left(C^*=G_4|c_{im}=c_{im}^{j}\right)\right\}-P(C^*=G_4)}{P(C^*=G_4)} \tag{22}$$

$$R(c_{im})=\frac{P(C_*=G_4)-\min\left\{P\left(C^*=G_4|c_{im}=c_{im}^{j}\right)\right\}}{P(C^*=G_4)} \tag{23}$$

$$M(c_{im})=(A(c_{im})+R(c_{im}))/2 \tag{24}$$

3　实例分析

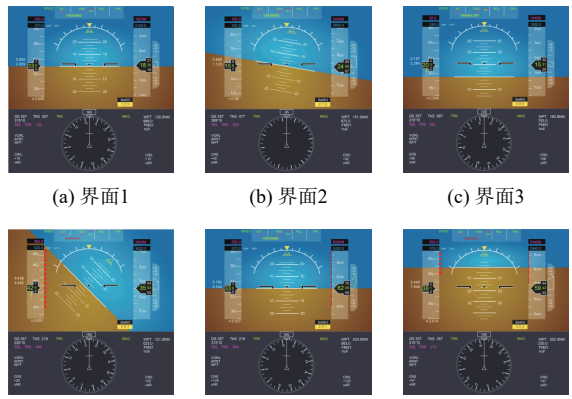
3.1　实验被试

本实验在西北工业大学机电学院进行, 共招募 15 名被试, 年龄范围在 24～30 岁, 平均年龄为 25.6 岁(S_D=2.616 岁), 大部分为与飞行或界面设计相关专业背景的博士、硕士研究生。在接受本实验任务前, 所有被试均受过实验飞行任务的培训, 熟练掌握实验操作方法。

3.2　实验设计

以根节点和中间节点的“STATE=1”状态展开实验设计。鉴于主飞行显示器界面是民机驾驶舱显示界面的核心, 本实验以主飞行显示器为研究对象。实验材料使用作者自绘的主飞行器界面, 如图 5 所示。该系列界面涵盖了巡航、起飞、告警 3 种不同驾驶状态, 对空速、高度、俯仰角、横滚角和航向角的数值和状态进行调整, 异常信息设定范围如表 6 所示。

实验任务为任务 1、任务 2 和任务 3, 并增加无任务设置。无任务要求被试在界面无异常信息时, 自由观察主飞行显示界面 30 s; 任务 1 要求被试自由观察主飞行显示界面 20 s, 记忆关键信息数值后, 点击关键信息的答案; 任务 2 要求被试在已知界面有几处异常值时, 观察界面 20 s 并记忆, 回复异常部位及其数值; 任务 3 要求被试在已知界面存在异常时, 观察界面 20 s 并记忆, 回复异常部位及其数值。



(a) 界面1　(b) 界面2　(c) 界面3

(d) 界面4　(e) 界面5　(f) 界面6

图 5　主飞行显示器界面实验材料

Fig. 5　Experimental materials of primary flight display interface

表 6　主飞行显示器的异常信息范围设置

Table 6　Abnormal message range settings for primary flight display

仪表类型	仪表信息	异常值
姿态仪	横滚角	>20°
	俯仰角	仰角>30°; 俯角>15°
罗盘标尺	航向角	>180°
空速表	空速	>400 kn
高度和垂直速度表	高度	>1 828.80 m

3.3　实验环境

为模拟真实民机驾驶舱情境, 实验在全封闭的室内环境中完成。实验设备包括联想 GeekPro 工作站一台, 被试距离屏幕 500～600 mm, 切视线与屏幕的夹角成 90°。实验选用 Tobii Pro Fusion 遥测式眼动仪, 分辨率设定为 120 Hz, 以记录定量根节点的度量指标参数。

3.4　实验流程

实验开始后让被试进行眼动校准, 接着界面会引导被试进入实验任务。被试完成任务 1、任务 2、任务 3 期间, 记录被试点击任务 1 的答案和口头回复任务 2、任务 3 的答案。眼动实验完成后每名被试需填写主观模糊评价量表、SUS 系统可用性量表、模糊条件评价量表, 其中, 模糊条件量表包括父节点对中间节点 $C_i(i=1,2,3,4)$和叶节点 C^*的影响。

4　实验结果分析

4.1　节点先验概率和条件概率计算

4.1.1　根节点先验概率

1) 定量根节点先验概率。将 15 名被试完成任务 3 的眼动数据作为定量根节点先验概率的来源,

由式(2)和式(3)得，定量根节点$c_{im}(i=1;m\in N^*)$在“STATE=1”状态下眼动度量指标的正相关和负相关标准化结果如表7所示。

利用变异系数法求度量指标权重时，根据式(1)～式(7)计算定量根节点$c_{im}(i=1;m\in N^*)$包含所有度量指标Y_e的平均值$\bar{r}_\rho$、标准差σ_ρ、变异系数φ_ρ和权重δ_ρ。根据式(8)可得，定量根节点$c_{im}(i=1;m\in N^*)$先验概率为所包含度量指标的平均值$\bar{r}_\rho$与权重δ_ρ的乘积之和，最终变异系数法求定量根节点的先验概率$P(c_{1mj})$结果如表8所示。

表7 “STATE=1”状态下眼动度量指标部分标准化结果

Table 7 Results of partial standardisation of eye movement metrics in “STATE=1” state

被试t	Y_1/次	Y_2/ms	Y_3/像素	Y_4/ms	Y_5/ms	Y_6/ms	Y_7/个	Y_8/次	Y_9/%	Y_{10}/ms
t_1	0.772 3	0.545 0	0.533 3	0.547 3	0.941 7	0.474 9	0.533 3	0.121 2	0.25	0.062 0
t_2	0.960 4	0.472 7	0.644 8	0.814 6	0.210 8	0.698 6	0.644 8	0.151 5	0.25	0.076 3
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
t_{14}	0.643 6	0.728 3	0	0.520 5	1	0.783 1	0	0.939 4	0.25	0.065 2
t_{15}	0.703 0	1	0.433 2	0.246 8	0.403 6	0.421 2	0.433 2	0.484 8	0.5	0.097 8

表8 “STATE=1”状态下定量根节点的先验概率

Table 8 Prior probabilitiy of eye movement metrics for quantitative root nodes in “STATE=1” state

c_{1m}	Y	$\bar{r}_\rho$	σ_ρ	φ_ρ	δ_ρ	$P(c_{1mj})$
c_{11}	Y_1	0.674 0	0.263 0	0.390 0	0.242 3	0.528 8
	Y_2	0.474 0	0.310 0	0.654 0	0.406 1	
	Y_3	0.492 0	0.279 0	0.566 0	0.351 7	
c_{12}	Y_4	0.579 0	0.317 0	0.548 0	0.339 2	0.536 3
	Y_5	0.469 0	0.303 0	0.646 0	0.399 9	
	Y_6	0.584 0	0.246 0	0.421 0	0.260 8	
c_{13}	Y_7	0.612 0	0.319 0	0.521 0	0.438 8	0.523 3
	Y_8	0.454 0	0.302 0	0.666 0	0.561 2	
c_{14}	Y_9	0.555 0	0.358 0	0.646 0	0.488 5	0.547 8
	Y_{10}	0.541 0	0.366 0	0.677 0	0.511 5	

2) 定性根节点先验概率。根据表4可得被试年龄、学历、职称、工作年限所对应的权重w_1、w_2、w_3、w_4。由于本实验只考虑“STATE=1”情况，不进行标准化处理。由式(11)～式(13)计算定性根节点$c_{im}(i=2,3,4;m\in N^*)$在“STATE=1”状态下被试的三角模糊概率$\tilde{P}^t_{imj}$、精确概率P_{imj}、先验概率$P(c_{imj})$结果如表9所示。

4.1.2 非根节点条件概率

非根节点的模糊条件概率获取同样采用7级自然语言变量来获取。综合模糊条件概率计算参考定性根节点的综合模糊评价概率计算过程，通过式(11)～式(13)得到父节点影响下的面积均值条件概率表。

重复条件概率计算过程，得到父节点影响下中间节点$C_i(i=1,2,3,4)$处于“STATE=1”状态和叶节点C^*处于$G_\beta=\{G_1,G_2,G_3,G_4\}=\{$优,良,中,差$\}$状态的条件概率表，其中叶节点$C^*$的条件概率如表10所示。

4.2 可用性综合评价模型贝叶斯网络推理

1) 因果推理。在GeNIe软件中填充所有节点的条件概率表，利用软件正向计算界面的可用性评价等级，因果推理结果如图6所示。叶节点C^*全概率$P(C^*=G_2)>P(C^*=G_1)>P(C^*=G_3)>P(C^*=G_4)$，结合表5可知，该主飞行显示器界面的可用性评价等级为“$G_2=$良”。

2) 模型有效性验证。系统可用性量表(system usability scale, SUS)由10个问题组成，其中，奇数项问题着重于积极描述，偶数项问题则偏向消极描述，通常采用5级李克特量表(Likert scale)评分。通过对15名被试的SUS系统可用性量表进行计算，能得到该主飞行显示器界面处于[0,1]区间的归一化SUS系统可用性评分：

$$S_{\mathrm{US}}=[2.5\times(\overline{S}_{\mathrm{q}}-\overline{S}_{\mathrm{p}}+20)]/100=0.87$$

式中：$\overline{S}_{\mathrm{q}}$和$\overline{S}_{\mathrm{p}}$分别表示15名被试对SUS量表中奇数项问题和偶数项问题评分的平均值之和。将主飞行显示器界面处于[0,1]区间的SUS系统可用性分数与表5比较，得到该主飞行显示器界面的可用性评价等级同样为“$G_2=$良”，验证了本文民机驾驶舱显示界面可用性综合评价模型的有效性。

3) 诊断推理。当该主飞行显示器界面的可用性评价等级为“$G_4=$差”时(即$P(C^*=G_4)=1$)，为保证根节点后验概率的负增长，本文还设定中间节

表 9　定性根节点的三角模糊概率和先验概率

Table 9　Triangular fuzzy and prior probabilities for qualitative root nodes

$c_{im}(i=2,3,4;m\in N^*)$	t_1	t_2	t_3	…	t_{15}	三角模糊概率	先验概率
一致性c_{21}	高(H) (0.7,0.9,1.0)	较高(FH) (0.5,0.7,0.9)	高(H) (0.7,0.9,1.0)	…	较高(FH) (0.5,0.7,0.9)	(0.588 6,0.756 6,0.883 8)	0.746 4
简洁性c_{22}	较高(FH) (0.5,0.7,0.9)	高(H) (0.7,0.9,1.0)	中等(M) (0.3,0.5,0.7)	…	较高(FH) (0.9,1.0,1.0)	(0.538 4,0.702 6,0.845 5)	0.696 4
信息呈现c_{23}	中等(M) (0.3,0.5,0.7)	高(H) (0.7,0.9,1.0)	较高(FH) (0.7,0.9,1.0)	…	中等(M) (0.3,0.5,0.7)	(0.581 7,0.756 1,0.894 3)	0.747 0
信息易读性c_{24}	较高(FH) (0.5,0.7,0.9)	高(H) (0.7,0.9,1.0)	高(H) (0.7,0.9,1.0)	…	较高(FH) (0.9,1.0,1.0)	(0.606 2,0.768 2,0.894 3)	0.759 2
理解c_{31}	中等(M) (0.3,0.5,0.7)	高(H) (0.7,0.9,1.0)	中等(M) (0.3,0.5,0.7)	…	较高(FH) (0.7,0.9,1.0)	(0.610 8,0.779 0,0.896 1)	0.766 2
易学性c_{32}	高(H) (0.7,0.9,1.0)	非常高(VH) (0.9,1.0,1.0)	高(H) (0.7,0.9,1.0)	…	中等(M) (0.3,0.5,0.7)	(0.567 8,0.729 8,0.868 7)	0.724 0
满意度c_{33}	中等(M) (0.3,0.5,0.7)	高(H) (0.7,0.9,1.0)	较高(FH) (0.7,0.9,1.0)	…	中等(M) (0.3,0.5,0.7)	(0.591 3,0.758 4,0.887 4)	0.748 9
最小记忆负荷c_{34}	中等(M) (0.3,0.5,0.7)	高(H) (0.7,0.9,1.0)	中等(M) (0.3,0.5,0.7)	…	中等(M) (0.3,0.5,0.7)	(0.490 9,0.665 1,0.807 1)	0.657 0
信息实时性c_{41}	较低(FL) (0.1,0.3,0.5)	高(H) (0.7,0.9,1.0)	较低(FL) (0.1,0.3,0.5)	…	高(H) (0.7,0.9,1.0)	(0.454 5,0.628 7,0.785 4)	0.624 4
信息精确性c_{42}	较高(FH) (0.7,0.9,1.0)	高(H) (0.7,0.9,1.0)	较高(FH) (0.7,0.9,1.0)	…	高(H) (0.7,0.9,1.0)	(0.478 1,0.638 6,0.767 8)	0.630 7
错误预防c_{43}	较低(FL) (0.1,0.3,0.5)	较高(FH) (0.7,0.9,1.0)	较高(FH) (0.7,0.9,1.0)	…	高(H) (0.7,0.9,1.0)	(0.351 9,0.519 9,0.685 6)	0.519 3
交互反馈c_{44}	低(L) (0,0.1,0.3)	较高(FH) (0.7,0.9,1.0)	中等(M) (0.3,0.5,0.7)	…	较高(FH) (0.7,0.9,1.0)	(0.430 4,0.584 3,0.738 2)	0.584 3

表 10　“STATE=1”状态下父节点对叶节点C^*影响下的条件概率表

Table 10　Conditional probability table under the influence of parent nodes on leaf node C^* in “STATE=1” state

C^*处于G_1状态时	条件概率值	C^*处于G_2状态时	条件概率值	C^*处于G_3状态时	条件概率值	C^*处于G_4状态时	条件概率值
$P(C^*=G_2\leftarrow C_1)$	0.315 4	$P(C^*=G_2\leftarrow C_1)$	0.425 6	$P(C^*=G_3\leftarrow C_1)$	0.213 3	$P(C^*=G_4\leftarrow C_1)$	0.047 5
$P(C^*=G_1\leftarrow C_2)$	0.421 7	$P(C^*=G_2\leftarrow C_2)$	0.512 5	$P(C^*=G_3\leftarrow C_2)$	0.065 5	$P(C^*=G_4\leftarrow C_2)$	0.003 0
$P(C^*=G_1\leftarrow C_3)$	0.414 6	$P(C^*=G_2\leftarrow C_3)$	0.425 9	$P(C^*=G_3\leftarrow C_3)$	0.147 6	$P(C^*=G_4\leftarrow C_3)$	0.019 9
$P(C^*=G_1\leftarrow C_4)$	0.145 2	$P(C^*=G_2\leftarrow C_4)$	0.264 3	$P(C^*=G_3\leftarrow C_4)$	0.387 2	$P(C^*=G_4\leftarrow C_4)$	0.203 3
$P(C^*=G_1\leftarrow C_1,C_2)$	0.369 7	$P(C^*=G_2\leftarrow C_1,C_2)$	0.421 8	$P(C^*=G_3\leftarrow C_1,C_2)$	0.129 7	$P(C^*=G_4\leftarrow C_1,C_2)$	0.078 8
$P(C^*=G_1\leftarrow C_1,C_3)$	0.258 7	$P(C^*=G_2\leftarrow C_1,C_3)$	0.346 4	$P(C^*=G_3\leftarrow C_1,C_3)$	0.358 8	$P(C^*=G_4\leftarrow C_1,C_3)$	0.036 1
$P(C^*=G_1\leftarrow C_1,C_4)$	0.123 1	$P(C^*=G_2\leftarrow C_1,C_4)$	0.310 5	$P(C^*=G_3\leftarrow C_1,C_4)$	0.456 3	$P(C^*=G_4\leftarrow C_1,C_4)$	0.110 1
$P(C^*=G_1\leftarrow C_2,C_3)$	0.454 3	$P(C^*=G_2\leftarrow C_2,C_3)$	0.425 1	$P(C^*=G_3\leftarrow C_2,C_3)$	0.116 0	$P(C^*=G_4\leftarrow C_2,C_3)$	0.004 6
$P(C^*=G_1\leftarrow C_2,C_4)$	0.297 6	$P(C^*=G_2\leftarrow C_2,C_4)$	0.402 0	$P(C^*=G_3\leftarrow C_2,C_4)$	0.214 1	$P(C^*=G_4\leftarrow C_2,C_4)$	0.086 3
$P(C^*=G_1\leftarrow C_3,C_4)$	0.265 2	$P(C^*=G_2\leftarrow C_3,C_4)$	0.396 8	$P(C^*=G_3\leftarrow C_3,C_4)$	0.278 1	$P(C^*=G_4\leftarrow C_3,C_4)$	0.059 9
$P(C^*=G_1\leftarrow C_1,C_2,C_3)$	0.394 4	$P(C^*=G_2\leftarrow C_1,C_2,C_3)$	0.362 3	$P(C^*=G_3\leftarrow C_1,C_2,C_3)$	0.156 3	$P(C^*=G_4\leftarrow C_1,C_2,C_3)$	0.087 0
$P(C^*=G_1\leftarrow C_1,C_2,C_4)$	0.375 2	$P(C^*=G_2\leftarrow C_1,C_2,C_4)$	0.412 0	$P(C^*=G_3\leftarrow C_1,C_2,C_4)$	0.118 9	$P(C^*=G_4\leftarrow C_1,C_2,C_4)$	0.093 9
$P(C^*=G_1\leftarrow C_2,C_3,C_4)$	0.218 9	$P(C^*=G_2\leftarrow C_1,C_3,C_4)$	0.371 3	$P(C^*=G_3\leftarrow C_1,C_3,C_4)$	0.301 1	$P(C^*=G_4\leftarrow C_1,C_3,C_4)$	0.106 9
$P(C^*=G_1\leftarrow C_2,C_3,C_4)$	0.192 3	$P(C^*=G_2\leftarrow C_2,C_3,C_4)$	0.421 0	$P(C^*=G_3\leftarrow C_2,C_3,C_4)$	0.376 4	$P(C^*=G_4\leftarrow C_2,C_3,C_4)$	0.010 3
$P(C^*=G_2\leftarrow C_1,C_2,C_3,C_4)$	0.365 4	$P(C^*=G_2\leftarrow C_1,C_2,C_3,C_4)$	0.552 1	$P(C^*=G_3\leftarrow C_1,C_2,C_3,C_4)$	0.075 1	$P(C^*=G_4\leftarrow C_1,C_2,C_3,C_4)$	0.007 4
$P(C^*=G_2\leftarrow \overline{C_1},\overline{C_2},\overline{C_3},\overline{C_4})$	0.075 0	$P(C^*=G_2\leftarrow \overline{C_1},\overline{C_2},\overline{C_3},\overline{C_4})$	0.008 6	$P(C^*=G_3\leftarrow C_1,C_2,C_3,C_4)$	0.258 2	$P(C^*=G_4\leftarrow C_1,C_2,C_3,C_4)$	0.658 2

点$C_i(i=1,2,3,4)$处于“STATE=0”消极状态的发生概率为 1。根据式(20)和式(21)计算各根节点的后验概率, 通过 GeNIe 软件建立的逆向诊断推理结果如图 7 所示。

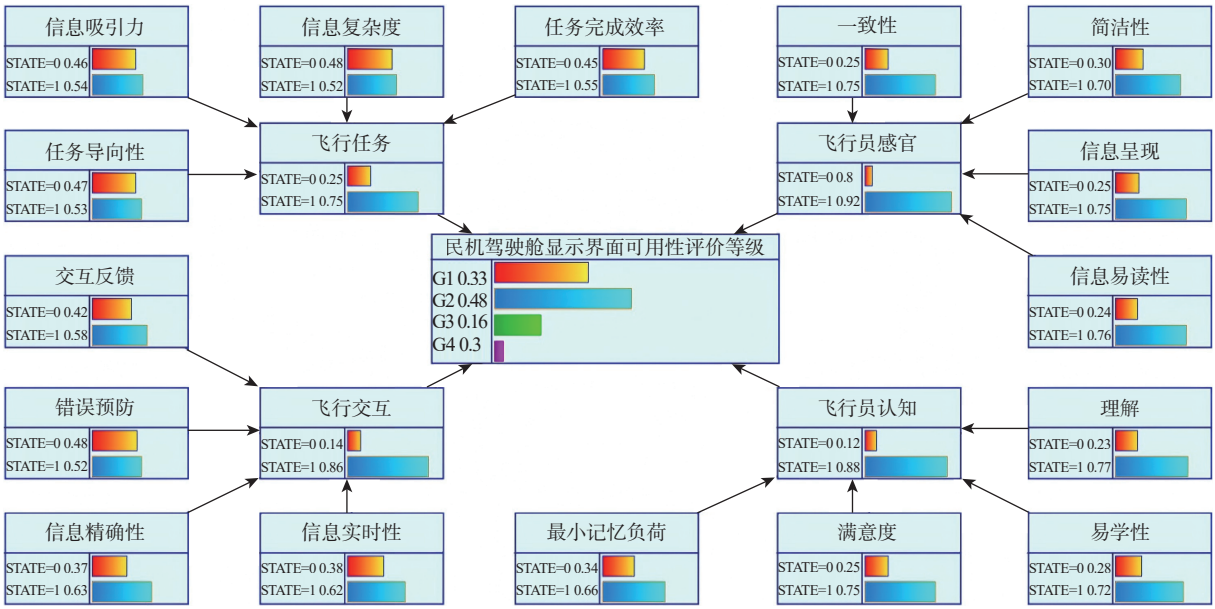


图6 可用性评价模型因果推理结果

Fig. 6 Causal reasoning result of usability evaluation model

根节点后验概率与先验概率的差值为变化百分比，若变化百分比越大，说明该根节点最能影响叶节点和中间节点。图8给出了根节点处于“STATE=1”状态下的先验概率、后验概率和变化百分比。由图8可知，该主飞行显示器界面对应的可用性评价指标均明显降低。当界面可用性状态最差时，根节点变化百分比＞10%的可用性评价指标有任务导向性c_{11}、一致性c_{21}、简洁性c_{22}、信息呈现c_{23}、信息易读性c_{24}、信息实时性c_{41}和信息精确性c_{42}。结果表明：在缺乏其他证据的情况下，上述有明显变化百分比的可用性评价指标可能是导致该界面可用性下降的最主要原因。

4）敏感性分析。根据式(22)～式(24)计算该界面中各根节点$W(c_{im})$、$R(c_{im})$和$M(c_{im})$的值，如图9所示。由图9可知，3项重要程度指标的排序结果相似，一致性c_{21}、简洁性c_{22}、信息呈现c_{23}、信息易读性c_{24}、信息实时性c_{41}和信息精确性c_{42}这6个敏感因子的排序最为靠前，应重点优化改进，具体措施如下。

在飞行员感官C_2维度，界面需弱化横滚针刻度线和航向角标识，选择更直观的图形或图标；替换背景颜色代表“天空”“大地”的色彩方案与对比度，调整字符排列和密度、各类仪表信息量大小。

在飞行交互C_4维度，界面需包含明确的反馈机

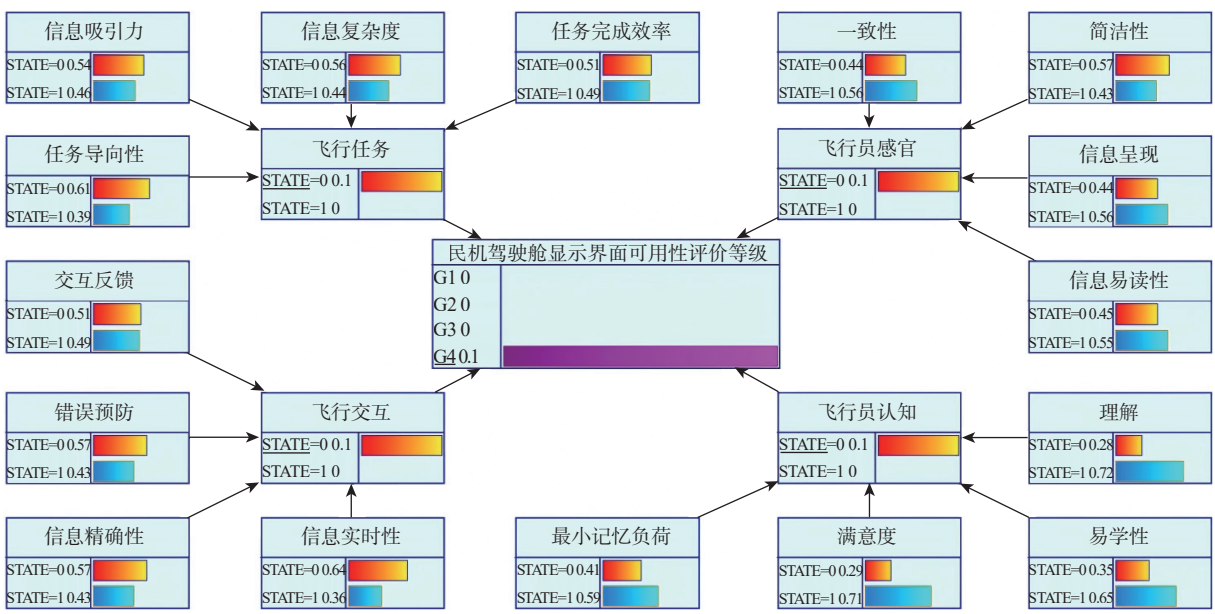


图7 可用性评价模型诊断推理结果

Fig. 7 Diagnostic reasoning result of usability evaluation model

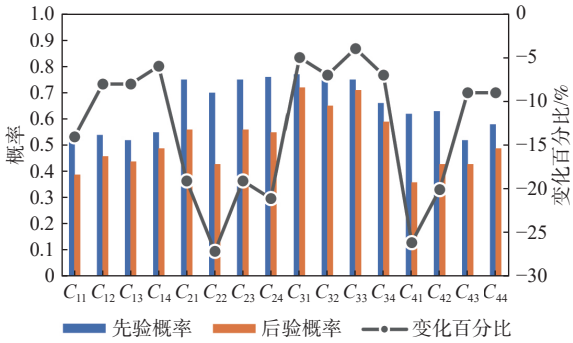


图 8　根节点处于“STATE=1”状态下的先验概率、后验概率和变化百分比

Fig. 8　Priori probability, posterior probability and percentage change of each root node in "STATE=1" state

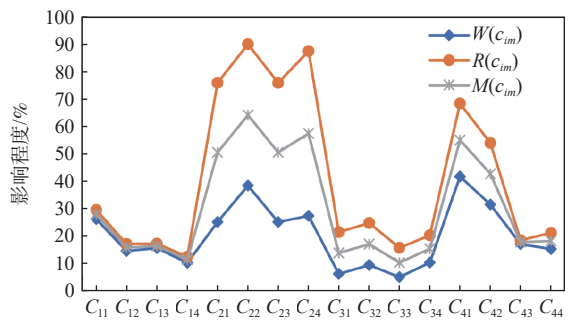


图 9　根节点重要程度

Fig. 9　Importance of root nodes

制, 如突出飞行空速、高度等关键参数的变化趋势和异常数值提示; 采用线框标明不同飞行状态, 增加对告警状态的预警响应。

5　结　论

1) 本文基于 EPIC 模型提取民机驾驶舱显示界面可用性评价指标的飞行任务、飞行员感官、飞行员认知和飞行交互维度, 在 4 个维度基础上, 提取并筛选出 4 个定性指标和 12 个定量指标。

2) 本文提出一种基于贝叶斯网络的民机驾驶舱显示界面可用性综合评价模型, 通过主飞行显示器界面的可用性评价模型因果推理结果与 SUS 量表评价结果相同, 验证了本文模型的有效性。

3) 可用性综合评价模型的定量根节点先验概率由眼动度量指标加权评价获得, 通过客观的行为数据减少评价偏差; 而定性根节点先验概率是基于专家和三角模糊数综合评价, 能有效处理评价的不确定性和主观性。

为提高模型贝叶斯网络中定量根节点先验概率的准确性, 可融合 Kinect、心率、呼吸等多源评价数据。

参考文献（References）

[1] 郑秀梅, 田晓康, 柳青, 等. 通用航空事故致因文本挖掘和社会网络分析[J]. 安全与环境学报, 2024, 24(2): 602-609.
Zheng X M, Tian X K, Liu Q, et al. Text mining and social network analysis of general aviation accident causes[J]. Journal of Safety and Environment, 2024, 24(2): 602-609(in Chinese).

[2] 王君, 孙宏. 航空飞行事故原因分析及预防[J]. 物流科技, 2021, 44(1): 81-84.
Wang J, Sun H. Analysis and prevention of aviation accidents[J]. Logistics Sci Tech, 2021, 44(1): 81-84(in Chinese).

[3] Al-Kilidar H, Cox K, Kitchenham B. The use and usefulness of the ISO/IEC 9126 quality standard[C]//Proceedings of the 2005 International Symposium on Empirical Software Engineering. Piscataway: IEEE Press, 2005: 7.

[4] Xue H J, Li T, Zhang X Y. An approach for assessing the usability of cockpit display system[C]//Proceedings of the Engineering psychology and cognitive ergonomics: Cognition and design. Berlin: Springer, 2017: 317-325.

[5] 张惠琳. 飞机驾驶舱数字界面的可用性评估研究[D]. 南京: 南京航空航天大学, 2020.
Zhang H L. Research on usability evaluationof aircraft cockpit digital interface[D]. Nanjing: Nanjing University of Aeronautics and Astronautics, 2020(in Chinese).

[6] Lewis J R. The system usability scale: past, present, and future[J]. International Journal of Human-Computer Interaction, 2018, 34(7): 577-590.

[7] Zhong X, Cheng Y N, Yang J H, et al. Evaluation and optimization of in-vehicle HUD design by applying an entropy weight-VIKOR hybrid method[J]. Applied Sciences, 2023, 13(6): 3789.

[8] Choi H, Kim S, Jang W. User experience evaluation of a spinal surgery robot: workload, usability, and satisfaction study[J]. JMIR Human Factors, 2024, 11: e54425.

[9] 李亚鸿, 董占勋, 储程. 基于质量屋的数字孪生系统用户界面优化和可用性评价[J]. 计算机集成制造系统, 2023, 29(6): 1941-1949.
Li Y H, Dong Z X, Chu C. User interface optimization and usability evaluation for digital twin system based on house of quality[J]. Computer Integrated Manufacturing Systems, 2023, 29(6): 1941-1949(in Chinese).

[10] Chiu M C, Huang P H, Tsao S H. Redesigning the user interface of a healthcare management system for the elderly with a systematic usability testing method[J]. Journal of Industrial and Production Engineering, 2019, 36(5): 324-334.

[11] 江涛, 葛正浩, 雷静, 等. 基于模糊层次分析法的工业缝纫机可用性评价[J]. 机械设计, 2022, 39(12): 119-124.
Jiang T, Ge Z H, Lei J, et al. Usability evaluation of industrial sewing machine based on fuzzy analytic hierarchy process[J]. Journal of Machine Design, 2022, 39(12): 119-124(in Chinese).

[12] Kabir S, Papadopoulos Y. Applications of Bayesian networks and Petri nets in safety, reliability, and risk assessments: a review[J]. Safety Science, 2019, 115: 154-175.

[13] 杨海波, 汪洋, 张磊. 人机交互的 GOMS 模型与 EPIC 模型的比较[J]. 包装工程, 2015, 36(8): 96-99.
Yang H B, Wang Y, Zhang L. Comparison of human-computer interaction model of GOMS and EPIC[J]. Packaging Engineering, 2015, 36(8): 96-99(in Chinese).

[14] Nielsen J. The usability engineering life cycle[J]. Computer, 1992, 25(3): 12-22.

[15] 徐美姣, 薛善良, 张惠, 等. 基于 BP-AdaBoost 算法的复杂产品装配制造成熟度等级评估方法[J]. 中国机械工程, 2023, 34(20): 2513-2519.
Xu M J, Xue S L, Zhang H, et al. Manufacturing readiness level assessment method of complex product assembly based on BP-AdaBoost algorithm[J]. China Mechanical Engineering, 2023, 34(20): 2513-2519(in Chinese).
[16] Weichbroth P. Usability of mobile applications: a systematic literature study[J]. IEEE Access, 2020, 8: 55563-55577.
[17] Chen D K, Zhu M Y, Qiao Y D, et al. An ergonomic design method of manned cabin driven by human operation performance[J]. Advanced Design Research, 2023, 1(1): 12-20.
[18] 赵欣, 丁怡, 侯文军, 等. 复杂信息系统界面可用性指标体系研究[J]. 图学学报, 2018, 39(4): 716-722.
Zhao X, Ding Y, Hou W J, et al. On the usability evaluation index system of complex information system interface[J]. Journal of Graphics, 2018, 39(4): 716-722(in Chinese).
[19] Mohanavelu K, Poonguzhali S, Ravi D, et al. Cognitive workload analysis of fighter aircraft pilots in flight simulator environment[J]. Defence Science Journal, 2020, 70(2): 131-139.
[20] 李敏, 林志军, 鲁义, 等. 基于模糊贝叶斯网络的煤矿瓦斯爆炸风险评估[J]. 煤炭学报, 2023, 48(增刊 2): 626-637.
Li M, Lin Z J, Lu Y, et al. Risk assessment of gas explosionin coal mines based on fuzzy Bayesian network[J]. Journal of China Coal Society, 2023, 48(Sup 2): 626-637(in Chinese).

BN-based comprehensive evaluation model for display interface usability in civil aircraft cockpit

XIAO Yao[1], CHEN Dengkai[1,*], ZHOU Yao[1], XIAO Jianghao[1], HOU Yuhong[1,2]

(1. Key Laboratory of Industrial Design and Ergonomics, Ministry of Industry and Information Technology, Northwestern Polytechnical University, Xi'an 710072, China;
2. Chinese Flight Test Establishment, Xi'an 710089, China)

Abstract: Optimizing the design interface is aided by research on the usability of display interfaces in civil aircraft cockpits. Aiming at the uncertainty of usability evaluation of civil aircraft cockpit display interface, this paper proposes a comprehensive usability evaluation model based on the Bayesian network (BN). Firstly, a Bayesian network is established based on the usability evaluation indicator system obtained from the executive process-interactive control model (EPIC). Both the experts' triangle fuzzy number evaluation and the quantitative root nodes' weighted evaluation of expert eye movement measures served as prior probabilities. Secondly, in order to reduce the number of experts' decision-making, the Noisy-OR model was used to fill the conditional probability table of the Bayesian network, and the model was built using GeNIe software. A primary flight display interface is taken as an example for instance verification. The results show that the causal reasoning results of the model are consistent with evaluation results of the system usability scale (SUS) scale. The evaluation grade of the interface is "good", which verifies the effectiveness of the model, and proposes improvement measures for sensitive factors obtained by the model sensitivity analysis.

Keywords: civil aircraft cockpit; display interface; usability evaluation; Bayesian network; eye movement

Received: 2024-08-28; **Accepted**: 2024-09-23; **Published Online**: 2024-10-16 14:38
URL: link.cnki.net/urlid/11.2625.V.20241016.1259.005
Foundation item: Key Research and Development Program of Shaanxi (2022GY-311)
* **Corresponding author**. E-mail: chendengkai@nwpu.edu.cn