

基于 AHPSort II 的突发事件网络 舆情群体极化风险评估研究

刘逸伦 黄 微 卢国强
(吉林大学商学与管理学院, 吉林 长春 130012)

摘 要: [目的/意义] 突发事件网络舆情群体极化是危害性最大的舆情演化结果, 对突发事件网络舆情群体进行极化风险评估是风险识别与管控策略的启承环节, 也是群体极化社会风险管理的重要内容。[方法/过程] 本文首先在三元空间视域下构建了突发事件网络舆情群体极化风险结构理论框架, 进而提出了群体极化风险评估方法并构建群体极化风险评估指标体系, 再基于 AHPSort II 设计了突发事件网络舆情群体极化风险等级分类评估方法, 最后通过过往案例数据进行了实证研究。[结果/结论] 本文提出了突发事件网络舆情群体极化风险评估方法, 能够实现对群体极化风险等级的划分和计算, 为管理部门准确掌握群体极化风险的发生、发展状态与趋势, 实现风险的科学分级分类提供理论支撑。

关键词: AHPSort II; 突发事件; 网络舆情; 群体极化; 风险评估

DOI: 10.3969/j.issn.1008-0821.2024.05.009

[中图分类号] G203 [文献标识码] A [文章编号] 1008-0821 (2024) 05-0107-16

Research on Group Polarization Risk Assessment of Network Public Opinion in Emergencies

Liu Yilun Huang Wei Lu Guoqiang
(School of Business and Management, Jilin University, Changchun 130012, China)

Abstract: [Purpose/Significance] The group polarization of network public opinion in emergencies is the most harmful result of public opinion evolution. The group polarization risk assessment of network public opinion in emergencies is the starting link of risk identification and guidance strategies, and is an important part of group polarization social risk management. [Method/Process] Based on the conceptual model of group polarization risk assessment, this paper constructed a risk assessment index system, and designed a classification assessment method for the group polarization risk level of network public opinion in emergencies based on AHPSort II. Finally, it conducted an empirical study through a case. [Result/Conclusion] The group polarization risk assessment method of emergency network public opinion constructed in this paper can realize the classification and calculation of group polarization risk grade, which has certain accuracy and rationality, and provides theoretical support for the management department to accurately grasp the occurrence, development status and trend of group polarization risk, and realizes scientific classification of risk.

Key words: AHPSortII; emergencies; public opinion; group polarization; risk assessment

我国颁布的《中华人民共和国突发事件应对法》 公共卫生事件和社会安全事件。近年来, 各类突发事件
中明确指出: 突发事件包括自然灾害、事故灾难、 事件频发, 给人民群众的生命、财产安全造成了极

收稿日期: 2023-08-09

基金项目: 国家自然科学基金面上项目“重大突发事件网络舆情受众的参与行为标定、轨迹拟合与靶向导控研究”(项目编号: 72174072)。

作者简介: 刘逸伦 (1994-), 男, 博士研究生, 研究方向: 网络舆情分析、应急管理。黄微 (1967-), 男, 教授, 博士生导师, 研究方向: 网络数据分析、应急情报。卢国强 (1982-), 男, 博士研究生, 研究方向: 危机信息管理、社交媒体分析。

大危害,严重影响国家和社会稳定。据中国互联网络信息中心(CNNIC)发布的第51次《中国互联网络发展状况统计报告》显示,我国网民规模已达10.67亿,互联网用户在社交媒体平台上表达观点和态度的热情正不断增加,针对突发事件网络舆情而产生的群体极化现象已呈现上升趋势。因此,如何在突发公共事件中对网络舆情群体极化现象进行有效控制,是突发危机情境下网络信息失序治理与舆情导控的重要课题。

群体极化现象是舆情演化的重要结果之一,美国学者凯斯·桑斯坦在其著作《网络共和国》中将群体极化定义为:经由群体讨论之后,群体态度往往比讨论之前的群体成员个人态度的平均值更趋向极端化的社会现象^[1]。现有研究已经对网络舆情群体极化的影响因素、形成机理、现象识别以及风险评估进行了一定程度的探索。其中,文献[2-3]根据社会比较理论和有力论据理论对群体极化影响因素进行了探究并提出舆情预警模型。文献[4-5]则综合社会比较理论和有力论据理论,提出引发社会情绪的事件、网民群体的协同、新媒体赋能及网络社交媒体平台中的匿名性带来的责任分散等关键要素对网络舆情风险预警与识别具有重要影响。文献[6]基于社会燃烧理论提出系列解释变量,对社交媒体群体极化的生成路径进行探究。文献[7]以自我归类理论与以社会规范理论为基础且强调个体之间有限信任的Weisbuch-Deffuant(WD)模型结合,构建新的SCT模型用于群体极化的仿真研究。从主体影响因素来看,网络用户的“从众”与“集体无意识”是其不可忽视的心理基础^[8]。从客观因素来看,一方面是由于互联网的“技术风险特性”即我国经济社会转型期间社会矛盾郁积向“空间转移”^[9];另一方面,网络热点议题驱动和过滤算法推荐也对群体极化的形成具有推动性作用^[10]。此外,网络舆情群体极化的识别与测度也是目前重要的研究方向。文献[11]论证了极端观点是群体极化风险最主要的呈现形态,构建了TCMCR极端观点识别模型。文献[12-13]指出在网络舆情发展阶段,群体极端意见的数量变化是群体极化识别的重要指标,极端意见数量与群体意见总数量可以反映群体极化强度。文献[14-15]将网络用户画像、复

杂系统结构等技术思想用于群体极化的分析与识别。

然而,突发事件网络舆情群体极化风险评价是一个多层级的、复杂的系统性工作,现有研究在群体极化风险形成的多属性量化方面存在不足。本文在三元空间视域下,对散布在物理空间、社会空间、信息空间中不同类型的动力要素进行归纳,再根据风险耦合形式构建测度指标体系并对群体极化风险进行评价测度,为主管部门准确把握群体极化风险的产生、发展状态及其演化趋势,实现对相关风险的科学分类和分级提供充分的理论支持。

本文研究主要解决以下3个问题:①结合突发事件网络舆情风险在物理空间、社会空间、信息空间中的运行模式,构建突发事件网络舆情群体极化风险评估模型;②采用层次聚类法(AHPSort II)对突发事件网络舆情群体极化风险进行等级划分;③以突发事件为例开展实证研究,对所构建的突发事件网络舆情群体极化风险评估方法进行验证和评价。

1 突发事件网络舆情群体极化风险致因因素及风险评估指标体系构建

1.1 突发事件网络舆情群体极化风险的致因因素及运行模式

根据群体极化的定义,群体极化的核心要素包括非理性交互和极端观点。非理性交互是群体极化的必要条件,极端观点是群体极化的结果。突发事件网络舆情群体极化是一个复杂的系统,在物理空间、社会空间、信息空间中各种风险要素耦合及风险叠加的情况下发生并形成突发事件网络舆情群体极化风险。

为了更好地理解突发事件网络舆情群体极化风险在物理空间、社会空间、信息空间中的分布及其不同空间中风险结构的相互作用关系,本文通过整合风险要素、危机情境下的风险表现、风险传导路径形成了群体极化风险结构框架,如图1所示,用以解释风险要素耦合、风险叠加和风险传导。

在突发事件网络舆情群体极化风险结构框架中,包括物理空间的突发事件原发社会风险、社会空间的非理性交互风险和空间的信息失序风险。突发事件的原发社会风险包括外在风险和内

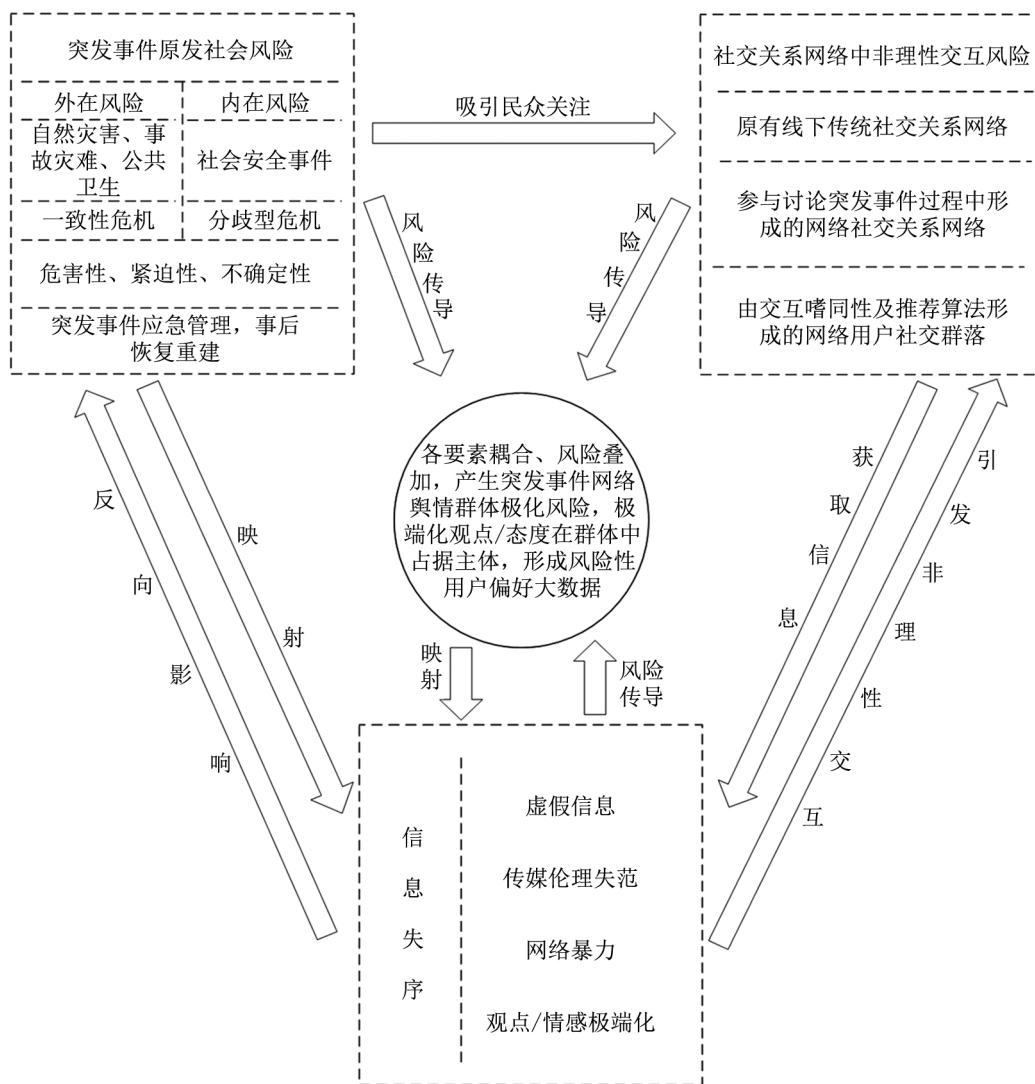


图1 突发事件网络舆情群体极化风险结构框架

Fig.1 Structural Framework of Risk for Group Polarization in Online Public Opinion on Emergencies

在风险。安东尼·吉登斯(Anthony Giddens)^[16]认为，自然灾害、事故灾难、公共卫生事件属于“外在风险(External Risk)”，社会安全事件属于“内在风险(Internal Risk)”，外在风险的处理难度通常小于内在风险。按照恩里克·克兰特利(Enricol Quarantelli)^[17]的理解，自然灾害、事故灾难、公共卫生事件都属于“一致性危机”，社会安全事件属于“分歧型危机”，常常引发群体性突发事件。社会安全事件通常将矛头直接指向核心制度、主流价值观或社会结构，挑战现存秩序的合法性和正当性，而自然灾害、事故灾难、公共卫生事件主要产生人员伤亡、财产损失，作用对象主要为普通民众，因此其社会危害更加广泛^[18]。在突发事件危机情境下，由突发事件危害性、紧迫性和不确定性引发的原发社会风险在吸引大量民众关注的同时，也通过

信息传播映射至信息空间。民众在通过社会空间中已有的社交关系网络获取突发事件信息的同时，也可从信息空间中的社交关系网络获取突发事件信息，并在社交媒体平台展开讨论，从而形成新的社交关系网络并映射至信息空间。网络用户在信息空间中失序信息的催化下开展非理性交互行为，并在交互嗜同性及推荐算法的推动下，形成了关系更为紧密的在线社交关系网络，进一步增大了用户非理性交互的风险。在此过程中，来自于物理空间和社会空间的风险，如：管理部门的应对失误、媒体信息传播的失范等相关风险将凝结成为信息空间中的信息失序风险。因此，突发事件舆情群体极化风险是在三元空间(物理空间、社会空间、信息空间)中各风险要素耦合、叠加作用下形成的。在物理和社会空间中，其表现形式为极端观点或极端情绪在

舆情受众群体中占据主体，在信息空间中则以风险性用户偏好大数据的形式存在。

1.2 突发事件网络舆情群体极化风险评估方法

突发事件网络舆情群体极化风险评价方法模型由风险评价描述、风险评价过程、风险等级划分 3 个模块构成。风险评价描述是风险评价前的准备工

作，内容包括目标定义和数据定义。风险评价过程的内容是对各个指标进行计算，也是整个评价的核心工作。风险等级划分是评估结果的输出模块，从而得到群体极化风险等级。突发事件网络舆情群体极化风险评价方法模型如图 2 所示。

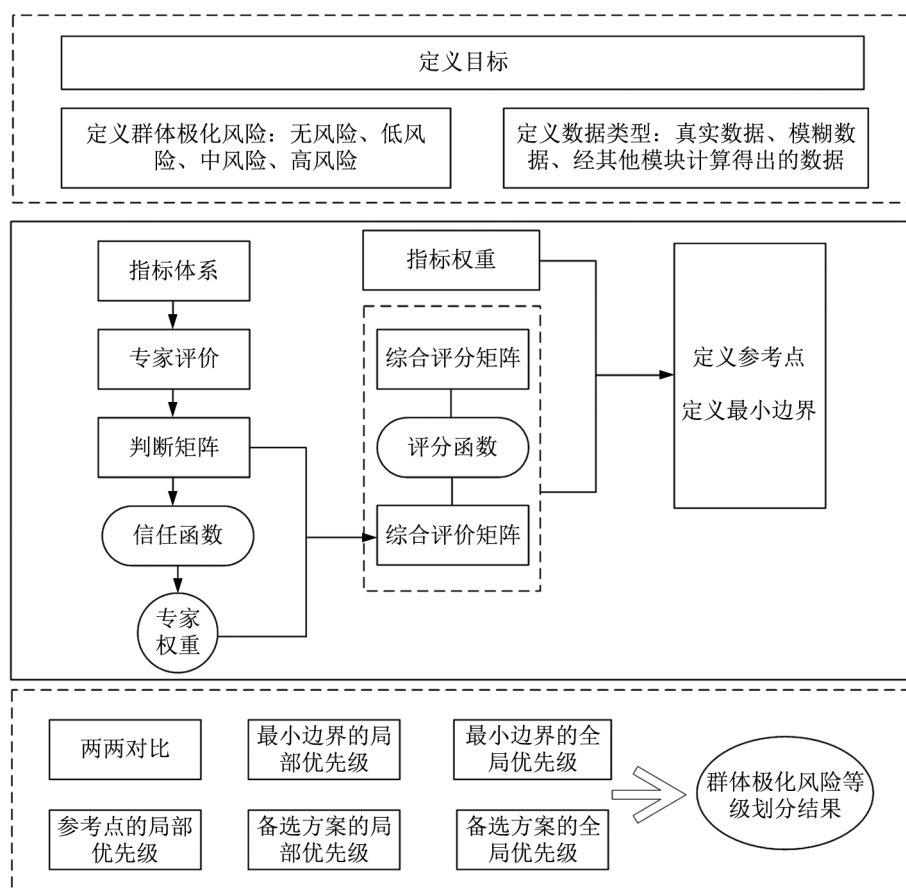


图 2 突发事件网络舆情群体极化风险评价模型

Fig. 2 Evaluation Model for Risk of Group Polarization in Online Public Opinion on Emergencies

群体极化风险评价的过程主要通过多准则分类方法，指标权重的赋值采用层次分析法（AHP）。层次分析法的主要内容包括：结合定性定量分析，将同决策的多种要素分解为 3 个层次^[19]。定义备选方案的分类是根据 AHPSort II 计算参考点和指标 C_i 的局部优先级，并依据指标权重计算边界的全局优先级和备选方案的全局优先级后，得出的群体极化风险的等级。已有研究通过多种方法，从不同角度对突发事件网络舆情风险指标的涉及进行了探索，如采用熵权法^[20]、德尔菲法和层次分析法^[21]、灰色统计法^[22]、层次分析法^[23]等，针对网络谣言^[24]、政府信息公开舆情风险^[25]等具体领域的舆情风险建立了相应的指标体系。

1.3 突发事件网络舆情群体极化风险评价指标体系设计

本文根据突发事件网络舆情群体极化特征与致因因素，参考已有研究中的突发事件网络舆情风险指标体系，构建突发事件网络舆情群体极化风险评价指标体系，并使用层次聚类法结合主观赋权方法确定指标权值。所得风险评价指标体系如表 1 所示。

突发事件网络舆情群体极化风险评价指标体系包括目标层、准则层、指标层。目标层为突发事件网络舆情群体极化风险评价，准则层为群体极化风险在三元空间中解构的突发事件原发社会风险评价、突发事件网络用户非理性交互风险评价、突发事件网络舆情信息失序风险评价。一级指标为三元空间

表 1 突发事件网络舆情群体极化风险评估指标体系

Tab. 1 Indicator System for Risk Assessment of Group Polarization in Online Public Opinion on Emergencies						
目标层	准则层	指标层		数据类型	指标类型	获取方式
		一级指标	二级指标			
突发事件网络舆情群体极化风险评估	突发事件原发社会风险 (N1)	危机情境影响要素 (E1)	人员伤亡 (S1)	定 量	成本型	统计数据
			财产损失 (S2)	定 量	成本型	统计数据
			是否存在明确致灾人员 (S3)	逻辑值	成本型	直接判断
			发生地点是否为政治、经济中心 (S4) ^[26]	逻辑值	成本型	直接判断
			持续时长 (S5)	定 量	成本型	统计数据
			影响地域范围 (S6)	定 性	成本型	专家评价
			影响社会范围 (S7)	定 性	成本型	专家评价
			是否与其他事件存在耦合 (S8) ^[27]	逻辑值	成本型	直接判断
			是否存在事件反转 (S9) ^[28]	逻辑值	成本型	直接判断
		政府响应 (E2)	风险预警 (S10)	逻辑值	效益型	直接判断
			应对时效性 (S11)	定 性	效益型	专家评价
			应急合理性 (S12)	定 性	效益型	专家评价
	突发事件网络用户非理性交互风险评价 (N2)	网络用户特征 (E3)	信息公开时效性 (S13)	定 性	效益型	统计数据
			网络用户负向情感占比 (S14)	定 量	成本型	二次计算
			网络用户受教育水平 (S15) ^[29]	定 性	效益型	统计数据
			与突发事件相关职业占比 (S16)	定 性	效益型	统计数据
			网络用户年龄分布 (S17)	定 性	效益型	统计数据
		网络用户交互影响要素 (E4)	刻板特征影响 (S18) ^[30]	定 性	成本型	专家评价
			转评赞数量 (S19)	定 量	成本型	统计数据
			网络用户情感负向演化方向 (S20) ^[31]	逻辑值	成本型	二次计算
			观点演化速度 (S21)	定 量	成本型	二次计算
			突发事件网络舆情信息失序风险评价 (N3)	信息属性 (E5)	是否存在虚假信息 (S22)	逻辑值
	是否存在恶意信息 (S23)	逻辑值			成本型	专家评价
	信息发布者无官媒 (S24)	定 性			成本型	直接判断
	网络用户发布信息数量 (S25) ^[32]	定 量			成本型	统计数据
	信息不确定性 (S26) ^[33]	定 性			成本型	专家评价
	媒体效力 (E6)	是否存在错误信息 (S27)		逻辑值	成本型	专家评价
		是否存在“信息茧房” (S28)		逻辑值	成本型	专家评价
		信息中图片、视频占比 (S29) ^[34]		定 量	成本型	统计数据

中群体极化风险构成风险的影响要素。二级指标为群体极化风险形成的具体属性细分。

1.4 突发事件网络舆情群体极化风险评估指标值获取方式

在突发事件网络舆情群体极化风险评价指标体系中，一部分逻辑值指标可直接判断；一部分定性指标和逻辑值指标由专家进行判断，如应对时效性(S11)、应对合理性(S12)、信息不确定性(S26)等；一部分定量指标可直接获取结构化数据，如人员伤亡(S1)、财产损失(S2)、持续时长(S5)、网络用户年龄分布(S17)等；还有一部分指标，则需要调用其他方法、模型进行计算，如网络用户负向情感占比(S14)、网络用户情感负向演化方向(S20)等指标，涉及定性、定量、逻辑值3种类型。指标内容主要涉及通过对自然语言的处理，得出相应语义信息。

1.4.1 通过统计获取结构化数据

由统计网络舆情结构化数据得到指标主要包括

人员伤亡(S1)、财产损失(S2)、持续时长(S5)、信息公开时效性(S13)、网络用户受教育水平(S15)、与突发事件相关职业占比(S16)、网络用户年龄分布(S17)、转评赞数量(S19)、网络用户发布信息数量(S25)以及信息中图片、视频占比(S29)。其中人员伤亡和财产损失可由新闻报道得出。持续总时长可参考网络舆情热度指数,各时序节点的持续总时长则为具体数据。信息公开时效性为突发事件发生后至官方报道相关事件的时长,可以以小时或天为单位。其他则为获取的网络舆情数据中的具体数据,即获取的网络舆情数据不仅包括舆情内容,也包括网络用户基本信息以及网络舆情内容随时间发展的具体数值。此部分指标因涉及使用极差标准化,可以定义最大值和最小值在整体事件发生过程中的极值,实际值为风险评价时间节点上的真实值。但在这种情况下,则无法对突发事件网络舆情群体极化风险进行实时评价,可以通过专家根据过往相似突发事件给出相关极值。

1.4.2 通过专家评价获取主观评价数据

对于影响地域范围(S6)、影响社会范围(S7)、应对时效性(S11)、应急合理性(S12)、刻板特征影响(S18)、是否存在虚假信息(S22)、是否存在恶意信息(S23)、信息不确定性(S26)、是否存在错误信息(S27)、是否存在“信息茧房”(S28)等指标,由于无法直接获取结构化数据,且通过二次计算也难以实现并无法保证客观性与准确性,因此选择通过专家评价获取相关指标的主观评价数据。专家评价前首先与专家充分沟通,使专家对舆情事件、评估目标、指标体系充分了解,然后对指标体系展开讨论并形成头脑风暴,在充分理解网络舆情群体极化风险的概念、内涵、外延等内容后,遵循层次分析法(AHP)对突发事件网络舆情群体极化风险评估指标进行评价并构建判断矩阵。专家评价过程使用YAAHP软件实现,有助于保证专家评价的便捷性及评价数据的一致性。对难以使用数量级评价的指标,选择使用模糊语义方式进行评价。涉及指标包括影响地域范围(S6)、影响社会范围(S7)、应对时效性(S11)、应急合理性(S12)、刻板特征影响(S18)、信息不确定性(S26)。通过模糊语义进行评价包括极低、中低、低、中、中高、

高6个标准。模糊语义作为一种弹性语言,是指外延不确定、内涵无定指的特性语言。与精确语言相比,模糊语言具有更大的概括性和灵活性。这种概括性与灵活性集中反映在语言外延上。得到语义模糊集后,将语义模糊集转化成为直觉模糊集,然后可以参与计算。直觉模糊集是传统的模糊集的一种拓展,它同时考虑了隶属度、非隶属度和犹豫度这3个方面的信息,因而比传统的模糊集在处理模糊性和不确定性等方面更具灵活性和实用性。另外,专家评价的模糊集也是确定专家权重的重要依据。在直觉模糊多属性群决策问题中,通过直觉模糊数计算专家的权重,其本质是通过定义直觉模糊集的模糊熵计算专家判断信息的模糊程度,专家判断信息的模糊程度反映了专家对决策问题的了解程度,进而用来确定专家的权重。专家权重的合理与否对决策结果的准确性至关重要,因此,确定合理的专家权重具有重要的现实意义。

1.4.3 网络用户情感、观点相关指标计算

需要使用用户情感模型计算的指标包括:网络用户负向情感占比(S14)、网络用户情感负向演化方向(S20)、观点演化速度(S21)。

本文选择情感知识增强预训练模型(Sentiment Knowledge Enhanced Pre-training, SKEP)进行情感极性、情感值、情感演化、观点演化的识别与计算。SKEP是百度在2020年提出的情感预训练模型,在情感分析典型人物上全面超越了SOTA,相关论文被ACL2020录用^[35]。

情感是人类认知的重要组成部分,让机器具备情感分析能力是实现机器认知智能的必要环节。具体来说,情感分析旨在自动识别和提取文本中的倾向、立场、评价、观点等主观信息。它包含各式各样的任务,比如情感倾向分类、实体级情感分类、观点抽取、情绪分析等,整体上这些任务均依赖于深入的情感语义理解。近年来,基于预训练的语义理解获得了迅猛的发展,显著提升了各类自然语言处理任务的效果。相比于通用预训练主要关注事实型文本(如新闻、百科等),情感分析更侧重于分析主观型文本中蕴涵的情感和观点,因此,采用了无监督方法自动挖掘情感知识,然后利用情感知识构建预训练目标,从而让机器学会理解情感语义^[36]。

2 基于 AHPSort II 的突发事件网络舆情群体极化风险等级分类评估

多准则分类是多准则决策 (Multi-Criteria Decision Making, MCDM) 中的重要部分, 它针对的是在管理决策中将多个方案在若干准则下分门别类的问题。多准则决策又可以分为 3 类: 选择 (Choice)、排序 (Ranking) 和分类 (Sorting)^[37]。选择和排序只能解决“最优与最差”的问题, 多准则分类决策却能将每一个评价对象划分到相应的类别, 对于所有的对象都能给出合理的评价。已有研究对多准则分类进行大量研究, 从基于数据包络分析 (DEA) 方法^[38-39]、基于 ELECTRE 法^[40-43]、基于层次分析法 (AHP)^[44-45] 探索多准则分类的多种方法。

本文选择 AHPSort II 方法原因在于相比其他多准则分类方法, AHPSort II 克服了其他方法比较次数与一致性方面的缺点, 极大程度地减少了计算的复杂性, 提高了计算的效率。而相比 AHPSort 方法, AHPSort II 不需要像 AHPSort 在计算边界局部优先级时, 将各类边界内所有备选方案与特殊点两两比较。AHPSort II 是使用参考点与边界点做两两比较计算局部优先级。在本文研究的实际情况中, 并不是对多项群体极化风险进行同时分类, 而是针对某一舆情事件多个时序节点判断是否具有群体极化风险, 以及群体极化风险的等级, 即同一舆情事件在不同时序节点的状态是备选方案集。在此情况下, 各备选方案之间大多数指标值存在相同或相似的情况, 因此, AHPSort II 方法在备选方案间的比较中计算局部优先级的效果优于其他方法。

2.1 使用直觉模糊集方法的定性数据描述及专家权重赋值

在实际多准则决策过程中, 由于决策问题本身所具有的大量不确定性信息, 包括不可量化的信息、不完整的信息或不可获取的信息, 很难依靠精确数据对其进行刻画, 而模糊集为表示和操作上述数据提供了很好的方法, 因此产生了模糊多属性群决策方法^[46]。在此基础上, 随着决策问题的进一步复杂以及个人经验、知识水平的限制, 决策者会经常表现出一定的犹豫度或不确定度, 传统的模糊集理论在处理这种犹豫性时表现出很大的局限性, Atanassov K T^[47] 从隶属度、非隶属度和犹豫度 3 个方

面描述不确定信息, 提出了直觉模糊集 (Intuitionistic Fuzzy, IF), 在处理不确定信息时具有更强的表现能力。

定义 1: 设 X 是一个非空集合, $A = \{ \langle x, \mu_A(x), v_A(x) \rangle | x \in X \}$ 为直觉模糊集, 其中, $\mu_A(x)$ 、 $v_A(x)$ 分别为元素 x 属于集合 X 的隶属度和非隶属度, 且满足条件 $0 \leq \mu_A(x) + v_A(x) \leq 1, \forall x \in X, \pi_A(x) = 1 - \mu_A(x) - v_A(x)$ 为 X 中元素 x 属于 X 的犹豫度, $0 \leq \pi_A(x) \leq 1$ 。

由于客观事物的复杂性和不确定, $\mu_A(x)$ 和 $v_A(x)$ 难以用精确数表示, 故对直觉模糊集进行扩展, 使区间数 $\mu_A(x)$ 和 $v_A(x)$ 满足 $\mu_A(x) \subset [0, 1], v_A(x) \subset [0, 1]$, 且 $\mu_A^{\max}(x) + v_A^{\max}(x) \leq 1, \forall x \in X$ 。

本文选择使用直觉模糊数集表示无法获取结构化数据的指标。将 $\alpha = (\mu_\alpha, \gamma_\alpha, \pi_\alpha)$ 作为本文研究中的直觉模糊数。 $\mu_\alpha \in [0, 1], \gamma_\alpha \in [0, 1], \mu_\alpha + \gamma_\alpha \leq 1, \pi_\alpha = 1 - \mu_\alpha - \gamma_\alpha, \alpha^+ = (1, 0, 0)$ 和 $\alpha^- = (0, 1, 0)$ 分别是本文研究中直觉模糊数的最大数和最小数。例如 $(0.5, 0.3, 0.2)$ 表示专家认为此指标的评价结果是 50% 支持、30% 反对、20% 中立。

为了便于计算, 还需将直觉模糊数集利用评分函数转成具体代数值^[48], 评分函数如式 (1) 所示。

$$J(x) = \mu_A(x) + \frac{1 + \mu_A(x) - \gamma_A(x)}{2} \cdot \pi_A(x) \quad (1)$$

$J(x) \in [0, 1]$ 。评分函数中由 $\mu_A(x)$ 和犹豫度偏向 $\mu_A(x)$ 方向的值组成。这意味着, 可以投票支持的中立者的比例一开始被认为是 0.5, 然后分配的比例被修正为赞成票和反对票之差的一半, 以反映投票支持的人越多, 倾向于投票支持的中立者的比例越大, 反之亦然。

在指标体系 $C_i (i = 1, 2, \dots, m)$ 中, 能够用定量数据表达的指标为 C_r , 通过专家判断得到结果的指标为 $C_f, C_i = C_r + C_f$ 。设有 t 位专家, 由 t 位专家针对方案使用模糊数集对 C_f 做出评价, 得到评价矩阵 $R_f = a_{jf}^k$:

$$R_f = \begin{bmatrix} a_1^1 & a_1^2 & \cdots & a_1^t \\ a_2^1 & a_2^2 & \cdots & a_2^t \\ \cdots & \cdots & \cdots & \cdots \\ a_{m_f}^1 & a_{m_f}^2 & \cdots & a_{m_f}^t \end{bmatrix}$$

其中, 矩阵中每一个 a_{ij}^k 均为模糊数集, 每一列为 1 位专家针对某一方案对所有定性指标评价的模糊数集, 将 $a_{i1}^1, a_{i2}^1, \dots, a_{im}^1$ 中对应的 $\mu_\alpha, \gamma_\alpha, \pi_\alpha$ 求算术平均, 得出每位专家的模糊数集, 利用式 (2)^[49] 计算出专家的权重 λ_k 。

$$\lambda_k = \frac{\left(\mu_k + \pi_k \left(\frac{\mu_k}{\mu_k + \gamma_k} \right) \right)}{\sum_{k=1}^t \left(\mu_k + \pi_k \left(\frac{\mu_k}{\mu_k + \gamma_k} \right) \right)} \quad (2)$$

2.2 参考点和边界的局部优先级计算

将定量表达的指标值使用极差标准化方法进行标准化、去量纲后, 极差标准化如式 (3) 所示。

$$\hat{a} = \frac{a - a_{\min}}{a_{\max} - a_{\min}} \quad (3)$$

将专家定性评价矩阵利用式 (1) 的评分函数计算出具体代数值, 与定量指标值结合, 得到完整指标体系的评价矩阵 R 。

$$R = \begin{bmatrix} \hat{a}_1^1 & \hat{a}_1^2 & \cdots & \hat{a}_1^t \\ \hat{a}_2^1 & \hat{a}_2^2 & \cdots & \hat{a}_2^t \\ \cdots & \cdots & \cdots & \cdots \\ \hat{a}_m^1 & \hat{a}_m^2 & \cdots & \hat{a}_m^t \end{bmatrix}$$

局部边界是指针对一个指标, 由专家判断的将备选方案划分为不同类的值。如本文将群体极化风险划分为 4 类, 则每个指标上均有 3 个边界点。可以用边界表示, 也可以用中界表示。

多准则分类方法的关键点是确定每个分组的界, 从而对类进行定义, 分组的界包括边界和中界。定义类 $S_y (y=1, 2, \dots, k)$, 其中 k 为类的数量。类被定义为有序类, 并且每一个类都有一个对应的名称 (如: 好, 中, 差), 且类 S_1 为最偏好的类, 类 S_k 为最不偏好的类。

定义 2^[39]: 定义边界 $lp_y (y=1, 2, \dots, k-1)$, 表示某个指标下能将方案划分到类 S_y 的最低表现, 可以被用来定义类的边界。以边界定义类如图 3 所示。

定义 3^[39]: 定义中界 $cp_y (y=1, 2, \dots, k)$, 表示在某个指标下方方案属于类 S_y 的一般 (典型) 表现, 可以被用来定义类的界限。以中界定义类如图 4 所示。

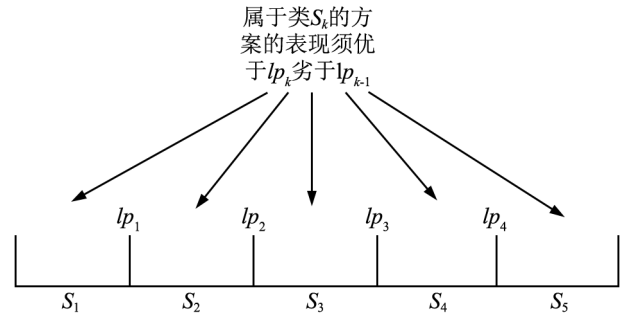


图 3 以边界定义的类

Fig. 3 The Class Defined By Boundaries

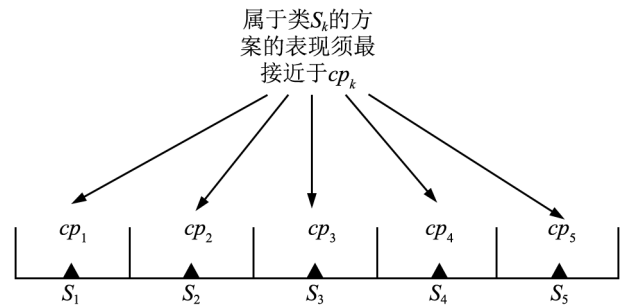


图 4 以中界定义的类

Fig. 4 The Class Defined By Median Boundaries

可以通过式 (4) 将中界推导出边界。

$$lp_y = \frac{cp_y + cp_{y+1}}{2} \quad (4)$$

在本文研究中, 使用边界进行突发事件网络舆情群体极化风险的分类。

设置参考点是 AHPSortII 方法的重要步骤, 参考点的不同选取方法, 将得到不同的指标 C_i 的局部最优。由于本文得到最后的评价矩阵 R 中的值均为标准化后的值, 且位于区间 (0, 1), 因此将 (0, 1) 区间平均划分 10 等分, 划分点 (0.1, 0.2, ..., 1) 作为参考点。定义每个风险等级的局部边界 $lp_y^q, q=1, 2, \dots, l-1$ 。在指标数为 m 下, 一共需要定义 $m \cdot (l-1)$ 个边界。参考点和边界定义如图 5 所示。

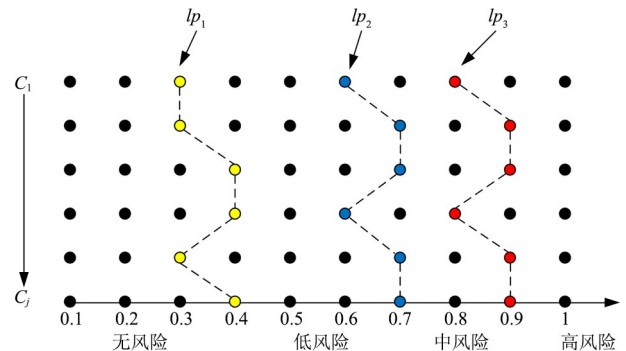


图 5 参考点和边界定义

Fig. 5 Reference Point and Boundary Definition

两两比较参考点和边界, 评价矩阵由专家作出判断, 用特征值方法推导出参考点的局部优先级 p_q^i 和边界的局部优先级 p_y^i 。

2.3 边界与方案的全局优先级计算

对于定量指标, 其值经标准化后可直接参与计算, 而对于定性指标, 由不同专家对方案的评价需经数据聚合、评分函数后, 得出方案在指标 C_j 上的值。数据聚合如式 (5) 所示。

$$g_j = \lambda_1 r_{ij}^1 \oplus \lambda_2 r_{ij}^2 \oplus \cdots \oplus \lambda_i r_{ij}^i \quad (5)$$

如果 g_j 落在两个连续的参考点内, 则方案在指标 C_j 上的局部优先级由式 (6) 得出。

$$p_j = p_j^o + \frac{p_j^{o+1} - p_j^o}{s_j^{o+1} - s_j^o} \cdot (g_j - s_j^o) \quad (6)$$

其中, s_j^o 和 s_j^{o+1} 是指标 C_j 上两个连续的参考点; p_j^o 和 p_j^{o+1} 是两个连续参考点的局部优先级; g_j 是方案在指标 C_j 上得分值; p_j 是方案在指标 C_j 上的局部优先级。对于成本型指标, g_j 直接参与计算, 对于效益型指标, $\hat{g}_j = 1 - g_j$ 。得到所有指标的局部优先级后, 结合指标权重即可计算方案的全局优先级以及边界的全局优先级, 如式 (7) 所示。

$$p_x = \sum_{j=1}^m p_j \omega_j \quad (7)$$

$$lp_q = \sum_{j=1}^m p_j^q \cdot \omega_j$$

根据方案的全局优先级以及边界的全局优先级, 即可得出方案的风险等级。

$$p_x < lp_1 \Rightarrow A_x \in S_1$$

$$lp_1 \leq p_x < lp_2 \Rightarrow A_x \in S_2$$

...

$$p_x \geq lp_q \Rightarrow A_x \in S_l$$

3 实证研究

2021 年 7 月 17 日起, 我国河南省部分地区遭遇极端强降雨袭击, 河南省气象台 2021 年 7 月 20 日 9 时 00 分发布了暴雨红色预警, 截至 2021 年 7 月 21 日, 郑州市累积平均降水量达 449 毫米。期间, “河南暴雨” “郑州地铁” “京广隧道” 等多个话题登上微博平台热搜, 吸引大量用户关注。据“知微事见” 全网热点舆情事件平台分析可见, “河南 7·20 特大暴雨” 突发舆情事件相比同期其他热点

事件具有更高的网络关注度和传播广泛度, 居于 2021 年度网络事件影响力和热度榜榜首。在暴雨期间, 由于灾难本身的突发性和不确定性, 以及主管部门的救灾响应不及时、针对网上舆情应对失措等原因, 导致相关舆情用户出现非理性交互现象, 部分谣言的迅速传播和网络舆论呈现对当地应急管理部门“一边倒”式的指责, 其具备群体极化的明显特征。“河南 7·20 暴雨”事件的热度趋势如图 6 所示。从事件热度趋势图可以看出, 事件的参与热度波峰与事件的节点高度匹配。事件参与热度的第一、二个峰值为 2021 年 7 月 20—21 日, 即郑州地区遭遇千年一遇暴雨, 3 日内降雨量超过该地一年内降雨量总和, 城内的交通、电力、通讯等基础设施受到极大影响; 其中 7 月 20 日晚, 郑州地铁五号线出现严重积水, 多名乘客被困车厢内, 有乘客在微博平台发出求救信息, 引发大量网民关注。7 月 21 日上午, 郑州官方媒体公布了地铁五号线人员伤亡情况, 其中 12 人死亡, 5 人受伤。本次舆情事件的最后一波热度峰值出现在 2021 年 8 月 2 日, 当日河南省政府通报, 此次特大洪涝灾害致 302 人遇难, 50 人失踪, 国务院同时宣布成立调查组, 对“7·20”特大暴雨灾害展开调查。

3.1 “河南 7·20 暴雨”事件群体极化风险评估指标体系赋权

通过 AHPSort II 对“河南 7·20 暴雨”舆情事件群体极化风险进行评估, 结合层次分析法对突发事件网络舆情群体计划风险评估指标进行赋权。目前, 我国对网络舆情的风险等级划分还没有统一的标准。网络舆情风险等级划分是风险评估的关键。本文借鉴已有研究, 将突发事件网络舆情群体极化风险划分为无风险、低风险、中风险、高风险。选择 4 位权威专家对本文构建的突发事件网络舆情群体极化风险评估指标使用层次分析法进行赋权。4 位专家包括一位致力于网络舆情领域科学研究及从事网络舆情分析人员, 一位长期致力于网络舆情领域科研人员, 一位互联网媒体从业人员, 一位政府应急事件管理人员, 以上专家均长期使用微博平台。突发事件网络舆情群体极化风险评估指标权重如表 2 所示。

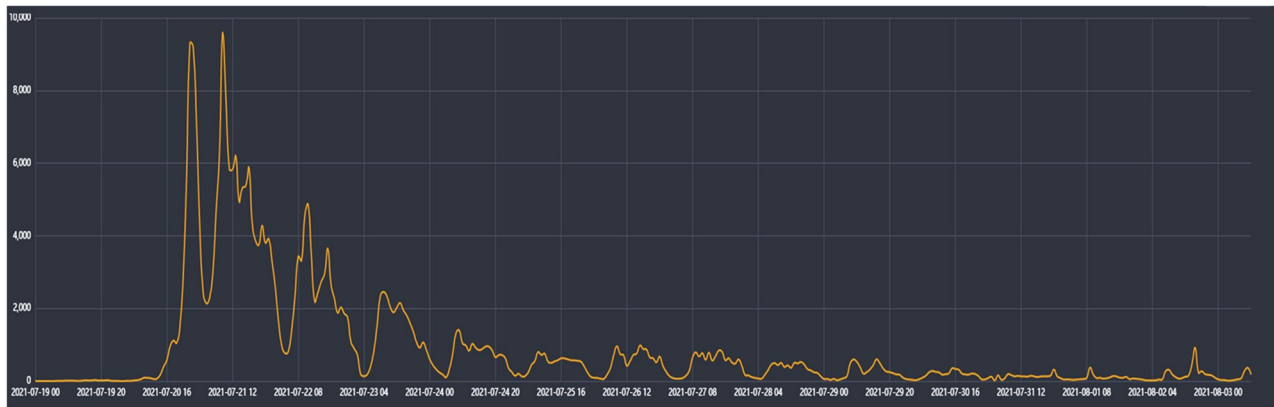


图 6 “河南 7·20 暴雨”事件热度趋势图(来源:知微事见)

Fig. 6 The Heat Trend Chart of the “July 20 Heavy Rainfall in Henan” Event (Source: Zhiwei Insight)

表 2 突发事件网络舆情群体极化风险评估指标权重

Tab. 2 The Weights of Risk Assessment Indicators for Group Polarization in Online Public Opinion on Emergencies

指 标	权 重
人员伤亡(S1)	0.0951
财产损失(S2)	0.0113
是否存在明确致灾人员(S3)	0.0487
发生地点是否为政治、经济中心(S4)	0.0051
持续时长(S5)	0.0071
影响地域范围(S6)	0.002
影响社会范围(S7)	0.0086
是否与其他事件存在耦合(S8)	0.0181
是否存在事件反转(S9)	0.0029
风险预警(S10)	0.0517
应对时效性(S11)	0.0762
应急合理性(S12)	0.0762
信息公开时效性(S13)	0.0762
网络用户负向情感占比(S14)	0.061
网络用户受教育水平(S15)	0.0538
与突发事件相关职业占比(S16)	0.0093
网络用户年龄分布(S17)	0.0056
刻板特征影响(S18)	0.0756
转评赞数量(S19)	0.0071
网络用户情感负向演化方向(S20)	0.0203
观点演化速度(S21)	0.011
是否存在虚假信息(S22)	0.009
是否存在恶意信息(S23)	0.0094
信息发布者无官媒(S24)	0.0061
网络用户发布信息数量(S25)	0.0211
信息不确定性(S26)	0.0975
是否存在错误信息(S27)	0.0985
是否存在“信息茧房”(S28)	0.0329
信息中图片、视频占比(S29)	0.0025

在突发事件网络舆情群体极化风险评估指标体系中,有 10 项指标需要专家评价,记为 C_f ,其中 6 项指标需要使用模糊集方法评价,数据类型为逻辑值的指标可以表达为犹豫值为 0 的模糊值。语义模糊数转换为直觉模糊数如表 3 所示。

表 3 语义模糊数转换为直觉模糊数

Tab. 3 Converting Semantic Fuzzy Numbers Into Intuitionistic Fuzzy Numbers

语义模糊表达	省略标记	直觉模糊数
极低	EL	(0.05,0.95,0.00)
很低	VL	(0.15,0.80,0.05)
低	L	(0.25,0.65,0.10)
中低	ML	(0.35,0.55,0.10)
中	M	(0.50,0.40,0.10)
中高	MH	(0.65,0.25,0.10)
高	H	(0.75,0.15,0.10)
很高	VH	(0.85,0.10,0.05)
极高	EH	(0.95,0.05,0.00)

对于数据类型为逻辑值的指标,专家的评价没有模糊熵的意义,因此,在使用模糊集计算专家权重时不考虑此部分指标。专家根据“河南 7·20 暴雨”事件网络舆情实际情况,结合自身对相关情况的理解与认知,对其他 6 项指标应用语义模糊方法进行评价结果并转化为直觉模糊集,并由式(2)可得 4 位专家权重分别为 0.2073、0.2565、0.2684、0.2677。

3.2 “河南 7·20 暴雨”事件群体极化风险评估指标值获得

根据“河南 7·20 暴雨”事件发展实际情况并结

合图 6, 本次突发事件舆情共持续约 16 天。因此, 以每两天作为一个时间节点进行群体极化风险评估。新浪微博在国内使用广泛、数据量大且具有代表性。因此, 本文选择以新浪微博网络舆情数据为代表, 对“河南 7·20 暴雨”事件网络舆情群体极化风险进行评估。获取数据的时间区间为 7 月 19 日—8 月 3 日, 获取数据属性包括原发博文内容、发表时间、转发、评论内容、用户名等, 在筛除无效数据后, 共计获取有效数据 177 895 条。

3.2.1 通过专家评价及直接判断获得对应指标值

对于是否存在虚假信息(S22)、是否存在恶意信息(S23)、是否存在错误信息(S27)、是否存在“信息茧房”(S28), 4 位专家的评价结果均为否, 即 4 个指标的逻辑值均为 0。将 4 位专家对指标 C_j ($C_j \in C_f$) 的模糊评价结合专家权重做算术平均, 得出 C_j 的模糊数, 通过式(1)得出指标 C_j 的代数值得。由专家评价得出指标值如表 4 所示。

表 4 专家评价指标值
Tab. 4 Expert Evaluation Index Value

指 标	指标 代数值	数据 类型	转成本型 的代数值
影响地域范围(S6)	0.1969	成本型	0.1969
影响社会范围(S7)	0.6663	成本型	0.6663
应对时效性(S11)	0.4015	效益型	0.5985
应急合理性(S12)	0.4035	效益型	0.5965
刻板特征影响(S18)	0.6246	成本型	0.6246
信息不确定性(S26)	0.7924	成本型	0.7924

由直接判断获得指标逻辑值如表 5 所示。

表 5 由直接判断获得指标逻辑值
Tab. 5 Obtaining Indicator Logical Values Through
Direct Judgment

评估二级指标	指标 类型	指标值
是否存在明确致灾人员(S3)	成本型	1
发生地点是否为政治、经济中心(S4)	成本型	1
是否与其他事件存在耦合(S8)	成本型	1
是否存在事件反转(S9)	成本型	0
风险预警(S10)	效益型	0
信息发布者无官媒(S24)	成本型	0

3.2.2 通过统计结构化数据获得对应指标值

由统计网络舆情结构化数据得到指标主要包括人员伤亡(S1)、财产损失(S2)、持续时长(S5)、信息公开时效性(S13)、网络用户受教育水平(S15)、与突发事件相关职业占比(S16)、网络用户年龄分布(S17)、转评赞数量(S19)、网络用户发布信息数量(S25)以及信息中图片、视频占比(S29)。对于“河南 7·20 暴雨”事件, 存在一定人员伤亡(S1)和财产损失(S2)。持续时长参考图 6 及事件实际发展, 可以认定为从 7 月 19 日起至 8 月初讨论热度衰退。鉴于前文选择的 8 个时间节点, 持续时长可以被认定为从 7 月 19 日至 8 月 3 日, 共计 16 天。各时序节点的持续时长则以 16 天为参考计算比值。“河南 7·20 暴雨”事件中不存在物理空间中突发事件相关事件信息公开时效性, 或也可以认为相关事件发生后, 官方媒体报道没有时滞。因此, 信息公开时效性最大值为 1, 又由于信息公开时效性效益型指标, 因此转为成本型即为 0。网络用户受教育水平(S15)以 8 个群体极化风险评估时间节点内(即各时间段内)本科以上学历用户占此时间段内参与讨论网络用户总量的比值。与突发事件相关职业占比的计算方式与网络用户受教育水平计算方式相同。两者均为效益型指标, 因此取与 1 之差的绝对值参与最终计算。网络用户年龄分布(S17)采用网络用户年龄平均值与 100 的比值代表各时段内参与讨论的网络用户年龄在 0~100 之间的值分布。转发评论数量(S19)(本文选择交互行为中代表性较强的转发、评论数量)、网络用户发布信息数量(S25)通过各时间段内总量与生命周期内总量比值计算。信息中图片、视频占比(S29)通过获取网络舆情数据的文本特征判断是否包含图片或视频, 计算方法及标准化与转发评论数量(S19)、网络用户发布信息数量(S25)的计算相同。第一个时间节点通过统计结构化数据获得对应指标值如表 6 所示。

3.2.3 通过二次计算获得对应指标值

通过二次计算获得值的指标包括网络用户负向情感占比(S14)、网络用户情感负向演化方向(S20)、观点演化速度(S21)。网络用户负向情感占比(S14)是指在某一时间段内具有负向情感网络用户占此时

间段内用户总量的比例。网络用户情感负向演化方向为逻辑型数值,通过此时段内平均情感值与前一时段平均情感值对比,如负向情感平均值增加,则逻辑值为 1,否则为 0。根据群体极化的定义,网络用户经非理性交互后,观点更为极端,同时也意

味着观点数量相对变得更少。因此,观点演化速度的度量可以以观点数量的变化与时长的比值来表现,观点数量越来越少,则群体极化风险越大,值为正值,否则值为负值。第一个时间节点通过二次计算获得对应指标值如表 7 所示。

表 6 第一个时间节点通过统计结构化数据获得对应指标值

Tab. 6 The Corresponding Indicator Values are Obtained Through Statistical Structured Data at the First Time Node

时间节点	需要统计结构化数据的指标	指标值	转成本型代数值
7 月 20 日	人员伤亡(S1)	0.1811	0.7189
7 月 20 日	财产损失(S2)	0.3798	0.6202
7 月 20 日	持续时长(S5)	0.125	0.125
7 月 20 日	信息公开时效性(S13)	1	0
7 月 20 日	网络用户受教育水平(S15)	0.1816	0.8184
7 月 20 日	与突发事件相关职业占比(S16)	0.0207	0.9793
7 月 20 日	网络用户年龄分布(S17)	0.3745	0.6265
7 月 20 日	转发评论数量(S19)	0.1083	0.1083
7 月 20 日	网络用户发布信息数量(S25)	0.1206	0.1206
7 月 20 日	信息中图片、视频占比(S29)	0.02736	0.02736

表 7 第一个时间节点通过二次计算获得对应指标值

Tab. 7 The Corresponding Indicator Values are Obtained Through Secondary Calculation at the First Time Node

时间 区间	负向情感 占比(%)	信息数量 (万)	观点数 (个)	负向情感 演化方向	观点演化速度 (个/2 天)	转发评论 数量(万)	图片、视频 占比(%)
1	36.65	1.8522	3	1	1.5	1.1245	18.05
2	51.98	4.3221	8	2	2.5	2.2565	27.65
3	39.08	2.3315	10	1	1	2.0605	11.14
4	33.43	1.2141	8	1	-1	1.9525	6.63
5	26.16	0.5631	6	1	-1	1.1037	7.35
6	14.53	0.4081	5	0	-0.5	0.6805	6.54
7	16	0.4246	3	0	-1	0.5784	4.15
8	37.15	1.3732	1	1	-1	1.3315	13.79

3.3 群体极化风险评估的参考点和边界的局部优先级计算

参考点和边界的局部优先级计算是 AHPSort II 的较为复杂的核心步骤。本文指标值均为[0,1]。因此,参考点在[0,1]之间以 10 等分平均分布。以持续时长(S5)为例,对所有指标用于群体极化风险评估参考点的设置和局部优先级计算进行说明。对于“河南 7·20 暴雨”事件而言,持续时长最长为 16 天,即 7 月 19 日—8 月 3 日,最小为两天,即 7 月 19—20 日。根据专家知识确定持续时长

的边界为 $lp_1=10$ 、 $lp_2=12$ 、 $lp_3=14$ 。各时序节点即各备选方案的持续时长原始决策矩阵如表 8 所示。

参考点与边界比较矩阵如表 9 所示。其中局部优先级采用特征根方法计算。以 rp_{10} 为基准进行归一化。AHPSort II 分类方法的特点是可以使用 1~7 或 1~9 的自然数对参考点与边界进行线性近似评估。但指标持续时长是结构化数据,且在各时序节点具有明确分布,因此可使用具体数据做客观准确评价代替主观线性近似评价。

表 8 各备选方案的持续时长决策矩阵

Tab. 8 Decision Matrix for the Duration of Each Alternative Plan

备选方案	7 月 20 日	7 月 22 日	7 月 24 日	7 月 26 日	7 月 28 日	7 月 30 日	8 月 1 日	8 月 3 日
持续时长	2	4	6	8	10	12	14	16
标准化	0.125	0.25	0.375	0.5	0.625	0.75	0.875	1

表 9 参考点与边界比较矩阵

Tab. 9 Comparison Matrix of Reference Points and Boundaries

	rp_1	rp_2	rp_3	rp_4	rp_5	rp_6	lp_1	rp_7	lp_2	rp_8	lp_3	rp_9	rp_{10}	局部 优先级	归一化局 部优先级
lp_1	1	4	3	3	2	2	$\frac{2}{1/2}$	$\frac{1}{3/4}$	$\frac{1}{2/5}$	3	$\frac{3}{29/57}$	$\frac{3}{1/2}$	9	0.1587	0.8512
rp_2	$\frac{1}{4}$	1	5	3	8	$\frac{7}{1/2}$	$\frac{4}{1/3}$	$\frac{3}{1/2}$	$\frac{1}{2/5}$	3	4	7	8	0.1865	1.0000
rp_3	$\frac{1}{3}$	$\frac{1}{5}$	1	7	6	$\frac{3}{1/2}$	5	$\frac{2}{1/2}$	4	3	3	$\frac{1}{1/2}$	7	0.1449	0.7767
rp_4	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{7}$	1	5	$\frac{2}{1/2}$	4	3	$\frac{1}{1/2}$	$\frac{1}{1/2}$	2	2	6	0.0879	0.4711
rp_5	$\frac{1}{2}$	$\frac{1}{8}$	$\frac{1}{6}$	$\frac{1}{5}$	1	4	3	3	3	2	2	5	5	0.0862	0.4622
rp_6	$\frac{1}{2}$	$\frac{2}{15}$	$\frac{2}{7}$	$\frac{2}{5}$	$\frac{1}{4}$	1	$\frac{2}{1/2}$	4	$\frac{1}{1/2}$	3	$\frac{1}{1/2}$	2	4	0.0662	0.3548
lp_1	$\frac{2}{5}$	$\frac{3}{13}$	$\frac{1}{5}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{2}{5}$	1	3	3	$\frac{1}{1/2}$	2	3	3	0.0662	0.3548
rp_7	$\frac{4}{7}$	$\frac{2}{7}$	$\frac{2}{5}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{3}$	1	$\frac{14}{15}$	$\frac{7}{8}$	$\frac{4}{5}$	4	$\frac{2}{1/2}$	0.0412	0.2210
lp_2	$\frac{5}{7}$	$\frac{4}{7}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{2}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{1/14}$	1	$\frac{2}{3}$	$\frac{6}{7}$	3	2	0.0436	0.2337
rp_8	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{2}{3}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{2}{3}$	$\frac{1}{1/7}$	$\frac{1}{1/2}$	1	$\frac{6}{7}$	3	2	0.0418	0.2239
lp_3	$\frac{2}{7}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{2}{3}$	$\frac{1}{2}$	$\frac{1}{1/4}$	$\frac{1}{1/6}$	$\frac{1}{1/6}$	1	3	4	0.0429	0.2302
rp_9	$\frac{2}{7}$	$\frac{1}{7}$	$\frac{2}{3}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{1}{3}$	1	5	0.0282	0.1512
rp_{10}	$\frac{1}{9}$	$\frac{1}{8}$	$\frac{1}{7}$	$\frac{1}{6}$	$\frac{1}{5}$	$\frac{1}{4}$	$\frac{1}{3}$	$\frac{2}{5}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{5}$	1	0.0268	0.1439

3.4 群体极化风险评估的风险等级分类

得到参考点的局部优先级后,可以计算各时序节点即各评估方案在指标“持续时长”下的局部优先级。以 7 月 20 日为例,持续时长为两天,标准化后为 0.125,落在 0.1 与 0.2 两个参考点范围内。因此, $s_j^o=0.1$, $s_j^{o+1}=0.2$, p_j^o 和 p_j^{o+1} 是参考点 s_j^o 和 s_j^{o+1} 对应的归一化局部优先级, $p_j^o=0.0969$, $p_j^{o+1}=0.1939$ 。 g_j 是方案在指标 C_j 上标准化后得分值, $g_j=0.125$, 可得方案在指标持续时长上的局

部优先级 $p_j=0.8512+(1-0.8512)/(0.2-0.1)\times(0.125-0.1)=0.8884$ 。

同理可得,各时序节点在指标持续时长上的局部优先级如表 10 所示。

得到所有指标的局部优先级后,结合指标权重即可计算方案的全局优先级以及边界的全局优先级,计算方法如 2.3 小节的式(6)所示。“河南 7·20 暴雨”事件网络舆情各时序节点群体极化风险如图 7 所示。

表 10 各时序节点在指标持续时长上的局部优先级
Tab. 10 Local Priority of Each Timing Node in Terms of Indicator Duration

备选方案	7 月 20 日	7 月 22 日	7 月 24 日	7 月 26 日	7 月 28 日	7 月 30 日	8 月 1 日	8 月 3 日
持续时长	2	4	6	8	10	12	14	16
局部优先级	0. 8884	1	0. 8624	0. 5807	0. 4029	0. 3115	0. 2479	0. 1239

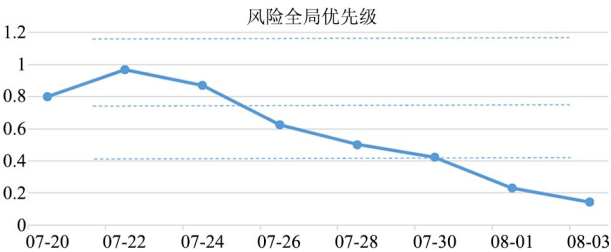


图 7 事件网络舆情各时序节点群体极化风险
Fig. 7 The Risk of Group Polarization at Each Timing Node of Online Public Opinion on Events

由图 7 可知，随着“河南 7·20 暴雨”事件网络舆情逐渐发酵，群体极化风险逐渐升高；随着舆情热度衰退，群体极化风险降低。其中，7 月 20—25 日为高风险，7 月 25—30 日为中风险，7 月 30 日之后为低风险。

4 群体极化风险评估结果讨论

4.1 结果讨论

通过 AHPSort II 对突发事件网络舆情群体极化风险进行评估，当对各时序节点进行风险分类时，采用了结合影响群体极化风险评估指标与对应数值做线性加和的方式进行计算。因此，可计算各指标的局部优先级占群体极化风险的全局优先级的比例，并以此比例表现出各评估指标对群体极化风险的贡献率。

其中，在 7 月 20 日时序节点的“河南 7·20 暴雨”事件网络舆情群体极化风险中，对风险形成占较大的评估指标主要包括：网络用户负向情感占比、与其他事件耦合、转发评论数量、网络用户发布信息数量。网络用户负向情感占比对群体极化风险形成贡献最高且符合群体极化的特征，即网络用户经非理性讨论后形成极端观点。在我国应急管理部门发布河南暴雨受灾害情况的同时，有媒体发布了郑州地铁五号线进水、快速路隧道被淹等紧急情况，网络用户展开了热烈讨论，在突发事件网络舆情话题关系网络中，各节点逐渐趋于同步，观点逐渐演化向极端态度，即“一边倒”地质疑管理部

门的一切应急处置措施、极端不信任政府做出的救灾举措。由于网络舆情发展过程中，网络用户具有沉默螺旋效应，因此鲜有相关领域专业人员发表权威性言论进行说明。“河南 7·20 暴雨”事件网络舆情中存在的负面情感与应急管理部门应对措施失当，导致网络用户在非理性状态下对相关部门的指责与不信任存在明显耦合关系。转评赞数量以及网络用户发布信息数量则体现为“河南 7·20 暴雨”事件吸引大量网络用户参与讨论以及交互，是群体极化的外在条件，也是非理性交互的载体，助推了极端观点的产生以及群体极化风险的形成。信息不确定性主要是由于在此事件中，缺少具有公信力的专家和机构对当地受灾情况和突发灾难事件进行权威解读，同时，灾难的突发性和破坏力天然具备不确定性，从而触发了网络用户对政府应急响应和救灾举措持否定态度的现象。

在 7 月 22 日时序节点的“河南 7·20 暴雨”事件网络舆情群体极化风险中，对风险形成占比较大的评估指标主要包括：网络用户负向情感占比、与其他事件耦合、转发评论数量、信息不确定性、网络用户发布信息数量。7 月 22 日，随着人员伤亡情况以及郑州地铁五号线进水、郑州周边地区高铁停运、高速封闭等受灾信息的快速传播、发酵，网络用户的诸多负面情绪持续增加。转评赞数量、持续时长、网络用户发布信息数量持续增加，出现了网络用户情感进一步向负向演化的趋势。由于突发事件事发地为省会，因此事件发生地点是否为经济中心也对群体极化风险形成产生了作用。随着中央媒体加强了对灾难成因、灾难趋势及中央政府救灾措施的权威解读报道，同时相关灾害逐步减轻，使人们对自然灾害破坏力的认知进一步明确。8 月 3 日，国务院宣布成立调查组并通报了整个灾害事件的人员伤亡与财产损失情况，正式宣布灾情已经得到解决，当地政府积极开展灾后重建。网络用户的

非理性观点的增长趋势已经有所缓解,讨论话题逐步向灾后重建、经验总结和灾区援助等方面演进。讨论话题逐步向灾后重建、经验总结和向灾区援助等方向演化。随着“河南7·20暴雨”事件的舆情信息传播达到衰退期,至5月初至8月初群体极化风险已经可以基本忽略群体极化风险已经可以基本忽略。

4.2 研究方法评价

通过实证研究结果可知,本文构建的突发事件网络舆情群体极化风险评估方法可以有效对群体极化风险进行较为准确的评估。方法的有效性主要源自于群体极化风险评价指标体系的全面性以及AHP-SortII方法的使用。首先,本文构建的风险评价指标体系综合考虑了物理空间中的突发危机情境、社会空间中用户非理性交互以及信息空间中的信息失序的具象化表现。评价指标的全面性和综合性是风险评价的重要保障;其次,AHP-Sort 分类方法可通过将方案与组别的边界进行比较从而将备选方案分类到预定义的有序类别,且与传统多准则决策方法AHP相比,大大减少了比较次数。AHP-Sort II在AHP-Sort 基础之上通过比较选定的参考点得到其权重,并通过线性近似得到方案的权重,从而不需要将方案进行直接比较,需要更少比较次数。AHP-Sort II方法的应用也是实现准确评估群体极化风险的重要辅助工具。但AHP-Sort II在应用过程中存在部分需要定性评价的工作,对群体极化风险评估的便捷性、自动性带来一定影响。在实际应用中,本文所构建的群体极化风险评估指标体系与评价方法可以直接应用于突发舆情的监测与导控中。主管部门可结合实际情况组织多位相关领域专家,并根据上述评价方法和指标体系对舆情事件进行实时监测,以便及时调整导控策略。

5 结语

突发事件网络舆情群体极化风险评估是风险识别与导控策略的中间环节,是群体极化社会风险管理的重要内容。本文解析了突发事件网络舆情群体极化风险评价方法模型,构建了突发事件网络舆情群体极化风险评价指标体系,阐述了突发事件网络舆情群体极化风险评价使用的AHP-Sort II计算方法,包括指标赋权、专家赋权以及详细计算过程,

并通过实证研究验证了其合理性与可行性,以期为主管部门对群体极化风险的评估与测度、突发网络舆情的导控提供一套行之有效的方法。

参 考 文 献

- [1] 张志剑,刘政昊,马费成.面向互联网舆情事件的企业风险识别[J].工程管理科技前沿,2022,41(1):65-73.
- [2] 江志英,李宇洋,李佳桐,等.基于层次分析的长短记忆网络(AHP-LSTM)的食品安全风险预警模型[J].北京化工大学学报(自然科学版),2021,48(6):98-107.
- [3] 林玲,陈福集,谢加良,等.考虑风险偏好的网络舆情预警模型——基于直觉模糊和Choquet积分[J].情报杂志,2021,40(10):52-58.
- [4] 黄微,刘熠,孙悦.多媒体网络舆情语义识别的关键技术分析[J].情报理论与实践,2019,42(1):134-140.
- [5] 秦琴.突发自然灾害网络舆情风险评价研究[D].合肥:中国科学技术大学,2019.
- [6] 彭国超,程晓.热点事件中社交媒体群体极化形成机理研究[J].信息资源管理学报,2023,13(2):42-52.
- [7] 张鑫,田雪灿,刘鑫雅.反复性视角下网络舆情风险评估指标体系研究[J].图书与情报,2020,(6):123-135.
- [8] 梁冠华,鞠玉梅.基于舆情演化生命周期的突发事件网络舆情风险评估分析[J].情报科学,2018,36(10):48-53.
- [9] 秦程节.网络群体极化:风险,成因及其治理[J].电子政务,2017,(4):49-57.
- [10] 蒋宇,兰月新,刘冰月,等.面向舆情预测的突发事件首发信息风险评估研究[J].图书与情报,2016,(3):19-27.
- [11] 卢国强,黄微,刘毅洲.群体极化视域下突发事件网络舆情极端观点识别研究[J].情报资料工作,2023,44(1):42-51.
- [12] 杨洋洋,谢雪梅.网络谣言风险测度与治理路径研究[J].情报科学,2021,39(9):170-177.
- [13] 薛新波,周中林.政府信息公开的舆情风险评估指标体系构建及其策略研究[J].西南科技大学学报(哲学社会科学版),2018,35(3):69-75,79.
- [14] 张岩,魏玖长.风险态度、风险认知和政府信赖——基于前景理论的突发状态下政府信息供给机制分析框架[J].华中科技大学学报(社会科学版),2011,25(1):53-59,90.
- [15] Liu J, Liu L Y, Tu Y, et al. Multi-stage Internet Public Opinion Risk Grading Analysis of Public Health Emergencies: An Empirical Study on Microblog in COVID-19 [J]. Information Processing and Management, 2022, 59(1): 102796.
- [16] Giddens A. The Consequences of Modernity [M]. California: Stanford University Press, 1990.
- [17] Quarantelli E L. Conceptualizing Disaster From a Sociological Perspective [J]. International Journal of Mass Emergencies and Disasters, 1989, 7(3): 243-251.

- [18] 童星, 张海波. 群体性突发公共卫生事件及其治理——社会风险与公共危机综合分析框架下的再考量 [J]. 学术界, 2008, (2) 35-45.
- [19] 李丽娜. 网络信息资源质量评价研究综述 [J]. 图书情报工作, 2011, 55 (15): 62-66, 61.
- [20] 秦琴. 突发自然灾害网络舆情风险评估研究 [D]. 合肥: 中国科学技术大学, 2019.
- [21] 张鑫, 田雪灿, 刘鑫雅. 反复性视角下网络舆情风险评估指标体系研究 [J]. 图书与情报, 2020, (6): 123-135.
- [22] 梁冠华, 鞠玉梅. 基于舆情演化生命周期的突发事件网络舆情风险评估分析 [J]. 情报科学, 2018, 36 (10): 48-53.
- [23] 蒋宇, 兰月新, 刘冰月, 等. 面向舆情预测的突发事件首发信息风险评估研究 [J]. 图书与情报, 2016, (3): 19-27.
- [24] 杨洋洋, 谢雪梅. 网络谣言风险测度与治理路径研究 [J]. 情报科学, 2021, 39 (9): 170-177.
- [25] 薛新波, 周中林. 政府信息公开的舆情风险评估指标体系构建及其策略研究 [J]. 西南科技大学学报 (哲学社会科学版), 2018, 35 (3): 69-75, 79.
- [26] 贾若男, 王晰巍, 于雪, 等. 突发公共事件网络舆情时空演化分析模型及算法研究 [J]. 现代情报, 2023, 43 (2): 137-145.
- [27] 魏永. 重大突发事件网络舆情群体极化形成机制、传播过程及引导处置研究 [D]. 长春: 吉林大学, 2021.
- [28] 王晰巍, 贾若男, 孙玉姣. 数据驱动的社交网络舆情极化群体画像构建研究 [J]. 情报资料工作, 2021, 42 (6): 21-30.
- [29] Wang X W, Li Y Q, Li J X, et al. A Rumor Reversal Model of Online Health Information During the COVID-19 Epidemic [J]. Information Processing & Management, 2021, 58 (6): 102731.
- [30] Campbell D T. Common Fate, Similarity and Other Indices of the Status of Aggregates of Persons as Social Entities [J]. Behavioural Science, 1958, 3: 14-25.
- [31] 卢国强, 黄微, 杨佩霖. 基于风险耦合的突发事件网络舆情群体极化演化分析 [J]. 现代情报, 2023, 43 (3): 96-109.
- [32] 卢国强, 黄微, 杨佩霖, 等. 基于主题互感耦合的微博主题影响力研究 [J]. 现代情报, 2022, 42 (8): 136-148.
- [33] 黄微, 孙悦. 社交媒体平台用户参与行为强度测定与预判 [J]. 现代情报, 2022, 42 (11): 97-110, 177.
- [34] 周昕, 李瑞, 黄微. 多媒体网络舆情危机响应机理及风险分型研究 [J]. 图书情报工作, 2019, 63 (20): 6-16.
- [35] 百度 NLP. 百度发布情感预训练模型 SKEP [EB/OL]. <https://baijiahao.baidu.com/s?id=1667-378756041935660&wfr=spider&for=pc>, 2020-05-22.
- [36] Tian H, Gao C, Xiao X Y, et al. SKEP: Sentiment Knowledge Enhanced Pre-training for Sentiment Analysis [C] //58th Annual Meeting of the Association for Computational Linguistics (ACL 2020), Assoc Computat Linguist, 2020: 4067-4076.
- [37] Ehrgott M, Greco S, Figueira J R. Multiple Criteria Decision Analysis: State of the Art Surveys [M]. New York: Springer New York, 2016.
- [38] Karasakal E, Aker P. A Multicriteria Sorting Approach Based on Data Envelopment Analysis for R&D Project Selection Problem [J]. Omega, 2017, 73: 79-92.
- [39] Ishizaka A, Lolli F, Balugani E, et al. DEASort: Assigning Items with Data Envelopment Analysis in ABC Classes [J]. International Journal of Production Economics, 2018, 199: 7-15.
- [40] Dias L C, Antunes C H, Dantas G, et al. A Multi-criteria Approach to Sort and Rank Policies Based on Delphi Qualitative Assessments and ELECTRE TRI: The Case of Smart Grids in Brazil [J]. Omega, 2018, 76: 100-111.
- [41] Ishizaka A, Nemery P. Assigning Machines to Incomparable Maintenance Strategies with ELECTRE-SORT [J]. Omega, 2014, 47: 45-59.
- [42] Almeida-Dias J, Figueira J R, Roy B. ELECTRE TRI-C: A Multiple Criteria Sorting Method Based on Characteristic Reference Actions [J]. European Journal of Operational Research, 2010, 204 (3): 565-580.
- [43] Almeida-Dias J, Figueira J R, Roy B. A Multiple Criteria Sorting Method Where Each Category is Characterized By Several Reference Actions: The Electre Tri-nC Method [J]. European Journal of Operational Research, 2012, 217 (3): 567-579.
- [44] Ishizaka A, Pearman C, Nemery P. AHPSort: An AHP-based Method for Sorting Problems [J]. International Journal of Production Research, 2012, 50 (17): 4767-4784.
- [45] Miccoli F, Ishizaka A. Sorting Municipalities in Umbria According to the Risk of Wolf Attacks with AHPSort II [J]. Ecological Indicators, 2017, 73: 741-755.
- [46] 赵萌, 秦松松, 谢佳恒, 等. 考虑决策者风险偏好的区间直觉模糊多属性群决策方法 [J]. 运筹与管理, 2018, 27 (1): 7-16.
- [47] Atanassov K T. Intuition Fuzzy Sets [J]. Fuzzy Sets and System, 1986, (20): 87-96.
- [48] Wang W Q, Xin X L. Distance Measure Between Intuitionistic Fuzzy Sets [J]. Pattern Recognition Letters, 2005, 26 (13): 2063-2069.
- [49] Zhang S F, Liu S Y. A GRA-based Intuitionistic Fuzzy Multi-criteria Group Decision Making Method for Personnel Selection [J]. Expert Systems with Applications, 2011, 38 (9): 11401-11405.

(责任编辑: 郭沫含)