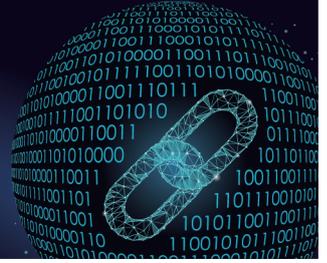




中国科技期刊卓越行动计划项目入选期刊

# 控制与决策

CONTROL AND DECISION



## 基于图犹豫不确定语言变量的商品属性分析方法

杨志辉, 俞弋峰

引用本文:

杨志辉, 俞弋峰. 基于图犹豫不确定语言变量的商品属性分析方法[J]. *控制与决策*, 2024, 39(6): 2079–2088.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2022.1590>

---

## 您可能感兴趣的其他文章

### Articles you may be interested in

#### [不确定语言信息下的大规模群体DEMATEL决策方法](#)

Large-scale group DEMATEL decision making method under uncertain linguistic information

*控制与决策*. 2021, 36(8): 2023–2033 <https://doi.org/10.13195/j.kzyjc.2019.1486>

#### [基于犹豫度和相似度的专家权重确定方法及其应用](#)

Expert weights determination method and application based on hesitancy degree and similarity measure

*控制与决策*. 2021, 36(6): 1482–1488 <https://doi.org/10.13195/j.kzyjc.2019.1382>

#### [概率区间值直觉犹豫模糊Maclaurin对称平均算子及决策方法](#)

Probabilistic interval-valued intuitionistic hesitant fuzzy Maclaurin symmetric mean operators and decision method

*控制与决策*. 2021, 36(5): 1249–1258 <https://doi.org/10.13195/j.kzyjc.2019.1370>

#### [考虑个体累积共识贡献的犹豫模糊语言自适应共识模型](#)

Adaptive consensus model with hesitant fuzzy linguistic information considering individual cumulative consensus contribution

*控制与决策*. 2021, 36(1): 187–195 <https://doi.org/10.13195/j.kzyjc.2019.0556>

#### [大群体应急决策中考虑属性关联的偏好信息融合方法](#)

Preference information fusion method of large groups emergency decision-making based on attributes association

*控制与决策*. 2021, 36(10): 2537–2546 <https://doi.org/10.13195/j.kzyjc.2020.0117>

# 基于图犹豫不确定语言变量的商品属性分析方法

杨志辉<sup>†</sup>, 俞弋峰

(东华理工大学 理学院, 南昌 330013)

**摘要:** 为研究消费者对商品属性的认可度,利用文本情感得分将顾客在线评论转化为图犹豫不确定语言数,基于图犹豫不确定语言广义 Shapely-Choquet 积分加权平均算子和商品在线评论构建商品属性分析模型. 首先,定义图犹豫不确定语言集、图犹豫不确定语言数的运算、得分函数和广义距离测度,提出图犹豫不确定语言广义 Shapely-Choquet 积分加权平均算子,并研究该算子的相关性质;然后,针对用户群体权重不完全已知的情况,建立优化模型来确定用户群体权重,并在此基础上,建立基于图犹豫不确定语言广义 Shapely-Choquet 积分加权平均算子和商品在线评论的商品属性分析模型;最后,将该模型应用于某品牌笔记本电脑的属性分析,验证所提出方法的有效性.

**关键词:** Choquet 积分; 广义 Shapely 函数; 图犹豫不确定语言集; 多属性决策; 商品属性分析

中图分类号: C934 文献标志码: A

DOI: 10.13195/j.kzyjc.2022.1590

引用格式: 杨志辉, 俞弋峰. 基于图犹豫不确定语言变量的商品属性分析方法[J]. 控制与决策, 2024, 39(6): 2079-2088.

## Product attribute analysis method based on picture hesitant uncertain linguistic variables

YANG Zhi-hui<sup>†</sup>, YU Yi-feng

(School of Science, East China University of Technology, Nanchang 330013, China)

**Abstract:** For the purpose of investigating the recognition of customer for product attributes, text sentiment score is used to convert online reviews into picture hesitant uncertain linguistic number, and the model of goods attribute analysis is constructed based on the picture hesitant uncertain linguistic generalized Shapely-Choquet weighted averaging operator and online reviews. Firstly, the picture hesitant uncertain linguistic set, the operation of picture hesitant uncertain linguistic number, scoring function and generalized distance measure are defined. The picture hesitant uncertain linguistic generalized Shapely-Choquet weighted averaging operator is proposed, and the related properties of the operator are discussed. Secondly, aiming at the situation that the weight of user group is not completely known, the optimization model is developed to determine the weight of user group. The goods attribute analysis model based on the picture hesitant uncertain linguistic generalized Shapely-Choquet weighted averaging operator and goods online review is proposed. Finally, the model is applied to the attribute analysis of a kind of laptop. The effectiveness of the proposed method is also verified.

**Keywords:** Choquet integral; generalized Shapely functions; picture hesitant uncertain linguistic set; multi-attribute decision making; goods attribute analysis

## 0 引言

在网络购物中,商品的在线评论是消费者识别商品质量好坏和商家服务水平的重要途径.有研究表明,绝大部分的消费者倾向于通过商品的在线评论来评估商品质量<sup>[1]</sup>,进而影响其购买决策,直接影响了

商品的销量<sup>[2-3]</sup>.因此,无论是消费者、厂家还是商家,对顾客在线评论进行分析显得尤为重要.

在线评论往往是基于产品体验与生活经验给出的文字性叙述,这种形式的评论难免存在一定的模糊性.对产品的在线评论进行分析,首先要应用自然语

收稿日期: 2022-09-07; 录用日期: 2023-02-10.

基金项目: 国家自然科学基金项目(71762001).

责任编委: 刘宝碇.

<sup>†</sup>通讯作者. E-mail: yang\_zhihui@aliyun.com.

\*本文附带电子附录文件,可登录本刊官网该文“资源附件”区自行下载阅览.

言处理技术对文本的情感倾向进行分析<sup>[4]</sup>,其次要将处理的结果用合理的形式表达出来,最后对处理的结果进行综合分析得到最终结果. 针对实际问题中亦此亦彼的连续过渡性现象,美国控制论专家Zadeh<sup>[5]</sup>提出了模糊集理论,用取值为区间 $[0,1]$ 上的隶属度函数来表征这种模糊性现象. Atanassov<sup>[6]</sup>在模糊集的基础上引入非隶属度,提出了直觉模糊集,用隶属度和非隶属度来表达模糊信息. Cùròng<sup>[7]</sup>在直觉模糊集的基础上引入中隶属度,提出了图模糊集理论,用积极隶属度、中隶属度和消极隶属度来刻画更加复杂的模糊信息. 由于决策问题的复杂性,决策者在评估备选方案时很难准确地用一个值来评价备选方案,往往会在多个值之间犹豫. 为此,Torra<sup>[8]</sup>提出了犹豫模糊集,更贴切地反映了决策者在决策时犹豫的态度. Wang等<sup>[9]</sup>提出了图犹豫模糊集(picture hesitant fuzzy set, PHFS). 在决策过程中,有时决策者很难确定精确的数值来描述模糊现象,可能更倾向于使用语言变量<sup>[10-11]</sup>来表达评价信息. Chen等<sup>[12]</sup>用语言术语表示隶属度和非隶属度,提出了语言直觉模糊集(linguistic intuitionistic fuzzy set, LIFS). 刘培德等<sup>[13]</sup>将直觉模糊语言集拓展成直觉不确定语言集,用一个语言区间代替直觉语言集的单一语言描述,更容易表述模糊信息.

在多属性决策中,信息聚合算子是综合考虑各个属性评价的关键工具. 刘培德等<sup>[13]</sup>提出了直觉不确定语言加权平均算子和直觉不确定语言有序加权平均算子. 在图模糊集的决策环境下,Wang等<sup>[9]</sup>提出了广义图犹豫模糊集优先级加权算子和图犹豫模糊集几何优先加权平均算子. Garg<sup>[14]</sup>提出了一系列图模糊聚合算子,即图模糊加权平均算子、图模糊有序加权平均算子和图模糊混合加权平均算子. 上述聚合算子在决策时都是在假设各个属性之间完全独立的情况下进行的. 然而,由于实际决策问题的复杂性,备选方案的属性间往往并非完全独立,可能存在着信息的冗余、交叉、互补等情况,这时属性值就不完全可加. 罗世华等<sup>[15]</sup>利用Maclaurin对称平均算子能反映多个输入参数之间的关联的特性,提出了概率区间值直觉犹豫模糊Maclaurin对称平均算子并应用于多属性决策中. Murofushi等<sup>[16]</sup>对模糊测度的非可加性对于属性间的相互影响做出了合理解释,讨论了Choquet积分的合理性并定义了基于模糊测度的Choquet积分. 陈岩等<sup>[17]</sup>提出了直觉不确定语言信息Choquet积分算子,并应用于属性间不完全独立的多属性决策中. Marichal<sup>[18]</sup>指出,离散Choquet积分是聚合不完全独立属性的一种工具. 然而,模糊测度只能

考虑两两属性的相关性,广义Shapely值能考虑属性对整体的贡献度<sup>[19]</sup>,将模糊测度与广义Shapely值相结合可以考虑非独立属性对全局的贡献.

考虑到利用不确定语言变量表达不确定信息的便捷性,本文首先定义图犹豫不确定语言数及其运算、广义距离测度,将模糊测度、Choquet积分与广义Shapely值相结合,推广到决策值为图犹豫不确定语言集的决策环境中,并定义图犹豫不确定语言广义Shapely-Choquet积分平均算子(picture hesitant uncertain linguistic generalized Shapely-Choquet weighted averaging operator, PHFULGSCWA);然后给出基于图犹豫不确定语言广义Shapely-Choquet积分集成算子和商品在线评论的商品属性分析模型;最后,通过从京东商城爬取某品牌笔记本电脑的在线评论,用本文所提出的方法进行处理,得到目标商品的各个属性的综合得分值并进行排序,其排序结果有助于生产者实时捕捉关键顾客需求,以持续改进产品质量.

## 1 预备知识

### 1.1 图犹豫模糊集

定义1<sup>[9]</sup> 设 $X$ 为一个非空有限集,则定义在 $X$ 上的图犹豫模糊集(PHFS)为

$$H = \{(x, \tilde{\mu}(x), \tilde{\eta}(x), \tilde{\nu}(x)) | x \in X\}. \quad (1)$$

其中: $\tilde{\mu}(x) = \{\alpha\}$ ,  $\tilde{\eta}(x) = \{\beta\}$ ,  $\tilde{\nu}(x) = \{\gamma\}$ 为取值于区间 $[0, 1]$ 上的一系列的数组成的集合,分别表示 $x$ 对 $H$ 的积极隶属度、中隶属度和消极隶属度,且满足

$$\begin{aligned} 0 &\leq \alpha^+ + \beta^+ + \gamma^+ \leq 1, \\ \alpha^+ &= \max_{\alpha \in \tilde{\mu}(x)} \{\alpha\}, \\ \beta^+ &= \max_{\beta \in \tilde{\eta}(x)} \{\beta\}, \\ \gamma^+ &= \max_{\gamma \in \tilde{\nu}(x)} \{\gamma\}. \end{aligned}$$

定义 $\pi(x) = 1 - (\tilde{\mu}(x) + \tilde{\eta}(x) + \tilde{\nu}(x))$ 为 $x$ 对 $H$ 的拒绝隶属度.

当集合 $X$ 中只包含一个元素 $x$ 时,图犹豫模糊集退化为 $\tilde{h} = \{\tilde{\mu}(x), \tilde{\eta}(x), \tilde{\nu}(x)\}$ ,称为图犹豫模糊数(PHFN),记作 $\tilde{h} = \{\tilde{\mu}, \tilde{\eta}, \tilde{\nu}\}$ .

### 1.2 语言术语集与不确定语言变量

定义2<sup>[20]</sup> 设 $S = \{s_i | i = 0, \dots, g\}$ 是一个由奇数个语言术语项组成的集合,每一个 $s_i$ 代表一个语言术语项,若集合 $S$ 满足如下性质,则称 $S$ 为语言术语集:

- 1) 有序性: 如果 $i \leq j$ ,则 $s_i \leq s_j$ ;
- 2) 可逆性: 对于任意元素 $s_i$ 都存在逆算子 $\text{neg}(\cdot)$ ,使得 $\text{neg}(s_i) = s_j, j = g - i$ ;

3) 存在最大和最小算子: 若  $s_i > s_j$ , 则有  $\max(s_i, s_j) = s_i, \min(s_i, s_j) = s_j$ .

徐泽水<sup>[21]</sup>为了最大化保留决策信息, 将离散的语言项集  $\{s_0, s_1, \dots, s_g\}$  拓展成一个连续的语言术语集  $\tilde{S} = \{s_i | i \in [0, g]\}$ , 拓展后的语言术语集仍然满足上述3条性质.

**定义3**<sup>[22]</sup> 设  $\tilde{s} = [s_a, s_b], s_a, s_b \in \tilde{S}$  且  $a \leq b$ , 其中  $s_a, s_b$  分别表示  $\tilde{s}$  的下限和上限, 称  $\tilde{s}$  为不确定语言变量.

设  $\tilde{S}$  为所有不确定语言变量构成的集合, 对于任意的两个不确定语言变量  $\tilde{s}_1 = [s_{a_1}, s_{b_1}], \tilde{s}_2 = [s_{a_2}, s_{b_2}]$  和常数  $\lambda > 0$ , 定义它们的运算如下:

- 1)  $\tilde{s}_1 \oplus \tilde{s}_2 = [s_{a_1+a_2}, s_{b_1+b_2}];$
- 2)  $\tilde{s}_1 \otimes \tilde{s}_2 = [s_{a_1 \times a_2}, s_{b_1 \times b_2}];$
- 3)  $\lambda \tilde{s}_1 = [s_{\lambda a_1}, s_{\lambda b_1}];$
- 4)  $\tilde{s}_1^\lambda = [s_{a_1^\lambda}, s_{b_1^\lambda}].$

### 1.3 图犹豫不确定语言集

**定义4** 设  $X$  为一个非空有限集,  $[s_{L(x)}, s_{U(x)}] \in \tilde{S}$  是一个不确定语言变量, 则定义在  $X$  上的图犹豫不确定语言集为

$$A = \{ \langle x, [s_{L(x)}, s_{U(x)}], \tilde{\mu}_A(x), \tilde{\eta}_A(x), \tilde{\nu}_A(x) \rangle | x \in X \}. \tag{2}$$

其中:  $\tilde{\mu}(x) = \{\alpha\}, \tilde{\eta}(x) = \{\beta\}, \tilde{\nu}(x) = \{\gamma\}$  为取值于区间  $[0, 1]$  上的一系列的数组成的集合, 分别表示  $x$  对不确定语言变量  $[s_{L(x)}, s_{U(x)}]$  的积极隶属度、中隶属度和消极隶属度, 且满足

$$0 \leq \alpha^+ + \beta^+ + \gamma^+ \leq 1,$$

$$\alpha^+ = \max_{\alpha \in \tilde{\mu}(x)} \{\alpha\},$$

$$\beta^+ = \max_{\beta \in \tilde{\eta}(x)} \{\beta\},$$

$$\gamma^+ = \max_{\gamma \in \tilde{\nu}(x)} \{\gamma\}.$$

此外, 定义  $\pi(x) = 1 - (\tilde{\mu}(x) + \tilde{\eta}(x) + \tilde{\nu}(x))$  为  $x$  对不确定语言变量  $[s_{L(x)}, s_{U(x)}]$  的拒绝隶属度.

当集合  $X$  中只包含一个元素  $x$  时, 图犹豫不确定语言集退化为

$$\tilde{a} = \langle [s_{L(x)}, s_{U(x)}], \tilde{\mu}_A(x), \tilde{\eta}_A(x), \tilde{\nu}_A(x) \rangle,$$

则称  $\tilde{a}$  为图犹豫不确定语言数 (PHULN).

**定义5** 设  $\tilde{a} = \langle [s_{L(x)}, s_{U(x)}], \tilde{\mu}(x), \tilde{\eta}(x), \tilde{\nu}(x) \rangle, \tilde{a}_1 = \langle [s_{L_1(x)}, s_{U_1(x)}], \tilde{\mu}_1(x), \tilde{\eta}_1(x), \tilde{\nu}_1(x) \rangle, \tilde{a}_2 = \langle [s_{L_2(x)}, s_{U_2(x)}], \tilde{\mu}_2(x), \tilde{\eta}_2(x), \tilde{\nu}_2(x) \rangle$  为 3 个 PHULN, 给定非负实数  $\lambda$ , 则定义它们之间的运算如下:

$$1) \tilde{a}_1 \oplus \tilde{a}_2 =$$

$$\langle [s_{L(a_1)+L(a_2)}, s_{U(b_1)+U(b_2)}], \bigcup_{\substack{\alpha_1 \in \tilde{\mu}_1(x), \beta_1 \in \tilde{\eta}_1(x), \gamma_1 \in \tilde{\nu}_1(x) \\ \alpha_2 \in \tilde{\mu}_2(x), \beta_2 \in \tilde{\eta}_2(x), \gamma_2 \in \tilde{\nu}_2(x)}} \{ \alpha_1 + \alpha_2 - \alpha_1 \alpha_2 \}, \{ \beta_1, \beta_2 \}, \{ \gamma_1 \gamma_2 \} \rangle,$$

$$2) \tilde{a}_1 \otimes \tilde{a}_2 =$$

$$\langle [s_{L(a_1)+L(a_2)}, s_{U(b_1)+U(b_2)}], \bigcup_{\substack{\alpha_1 \in \tilde{\mu}_1(x), \beta_1 \in \tilde{\eta}_1(x), \gamma_1 \in \tilde{\nu}_1(x) \\ \alpha_2 \in \tilde{\mu}_2(x), \beta_2 \in \tilde{\eta}_2(x), \gamma_2 \in \tilde{\nu}_2(x)}} \{ \alpha_1 \alpha_2 \}, \{ \beta_1 + \beta_2 - \beta_1 \beta_2 \}, \{ \gamma_1 + \gamma_2 - \gamma_1 \gamma_2 \} \rangle,$$

$$3) \lambda \tilde{a} =$$

$$\langle [s_{\lambda L(a)}, s_{\lambda U(b)}], \bigcup_{\alpha \in \tilde{\mu}(x), \beta \in \tilde{\eta}(x), \gamma \in \tilde{\nu}(x)} \{ 1 - (1 - \alpha)^\lambda \}, \{ \beta^\lambda \}, \{ \gamma^\lambda \} \rangle,$$

$$4) \tilde{a}^\lambda =$$

$$\langle [s_{L(a)^\lambda}, s_{U(b)^\lambda}], \bigcup_{\alpha \in \tilde{\mu}(x), \beta \in \tilde{\eta}(x), \gamma \in \tilde{\nu}(x)} \{ \alpha^\lambda \}, \{ 1 - (1 - \beta)^\lambda \}, \{ 1 - (1 - \gamma)^\lambda \} \rangle,$$

容易证明 PHULN 的运算满足下列性质:

- 1)  $\tilde{a}_1 \oplus \tilde{a}_2 = \tilde{a}_2 \oplus \tilde{a}_1;$
- 2)  $\tilde{a}_1 \otimes \tilde{a}_2 = \tilde{a}_2 \otimes \tilde{a}_1;$
- 3)  $\lambda(\tilde{a}_1 \oplus \tilde{a}_2) = \lambda \tilde{a}_1 \oplus \lambda \tilde{a}_2;$
- 4)  $\lambda(\tilde{a}_1 \otimes \tilde{a}_2) = \lambda \tilde{a}_1 \otimes \lambda \tilde{a}_2;$
- 5)  $\tilde{a}^{\lambda_1} \otimes \tilde{a}^{\lambda_2} = \tilde{a}^{\lambda_1 + \lambda_2};$
- 6)  $\tilde{a}_1^\lambda \otimes \tilde{a}_2^\lambda = (\tilde{a}_1 \otimes \tilde{a}_2)^\lambda.$

### 1.4 图犹豫不确定语言数的排序

**定义6** 设  $\tilde{a} = \langle [s_{L(x)}, s_{U(x)}], \tilde{\mu}(x), \tilde{\eta}(x), \tilde{\nu}(x) \rangle$  是一个图犹豫不确定语言数. 其中:  $[s_{L(x)}, s_{U(x)}] \in \tilde{S}, \tilde{\mu}(x) = \{\alpha_1, \dots, \alpha_m\}, \tilde{\eta}(x) = \{\beta_1, \dots, \beta_l\}, \tilde{\nu}(x) = \{\gamma_1, \dots, \gamma_q\}$ . 则定义语言图犹豫不确定语言数  $\tilde{a}$  的得分函数  $s(\tilde{a})$  和精确函数  $p(\tilde{a})$  如下:

$$s(\tilde{a}) = \frac{1}{4}(\tilde{\mu} + 1 - (\tilde{\eta} + \tilde{\nu}))(s_{L(x)} + s_{U(x)}), \tag{3}$$

$$p(\tilde{a}) = \frac{(s_{L(x)} + s_{U(x)})}{2}(\tilde{\mu} + \tilde{\eta} + \tilde{\nu}). \tag{4}$$

其中:  $\tilde{\mu} = \frac{1}{m} \left( \sum_{i=1}^m \alpha_i \right), \tilde{\eta} = \frac{1}{l} \left( \sum_{j=1}^l \beta_j \right), \tilde{\nu} = \frac{1}{q} \left( \sum_{k=1}^q \gamma_k \right)$  分别表示图犹豫不确定语言数  $\tilde{a}$  中元素  $x$  对不确定语言变量  $[s_{L(x)}, s_{U(x)}]$  的积极隶属度、中隶属度和消极隶属度的平均值.

对于任意两个图犹豫不确定语言数  $\tilde{a}_1$  和  $\tilde{a}_2$ , 若

$s(\tilde{a}_1) > s(\tilde{a}_2)$ , 则称  $\tilde{a}_1$  优于  $\tilde{a}_2$ , 记作  $\tilde{a}_1 \succ \tilde{a}_2$ ; 若  $s(\tilde{a}_1) = s(\tilde{a}_2)$  且  $p(\tilde{a}_1) > p(\tilde{a}_2)$ , 则称  $\tilde{a}_1$  优于  $\tilde{a}_2$ , 记作  $\tilde{a}_1 \succ \tilde{a}_2$ ; 若  $s(\tilde{a}_1) = s(\tilde{a}_2)$  且  $p(\tilde{a}_1) = p(\tilde{a}_2)$ , 则称  $\tilde{a}_1$  与  $\tilde{a}_2$  等价, 记作  $\tilde{a}_1 \cong \tilde{a}_2$ .

**定义7** 设  $\tilde{a}_1 = \langle [s_{L_1(x)}, s_{U_1(x)}], \tilde{\mu}_1(x), \tilde{\eta}_1(x), \tilde{\nu}_1(x) \rangle, \tilde{a}_2 = \langle [s_{L_2(x)}, s_{U_2(x)}], \tilde{\mu}_2(x), \tilde{\eta}_2(x), \tilde{\nu}_2(x) \rangle$  是论域  $X$  上的两个图犹豫不确定语言数,  $[s_{L(x)}, s_{U(x)}] \in \tilde{S}$ , 则定义它们之间的广义距离测度为

$$d(\tilde{a}_1, \tilde{a}_2) = \frac{1}{5} \left( \left| \frac{L_1(x) - L_2(x)}{g} \right|^\lambda + \left| \frac{U_1(x) - U_2(x)}{g} \right|^\lambda + |\tilde{\mu}_1 - \tilde{\mu}_2|^\lambda + |\tilde{\eta}_1 - \tilde{\eta}_2|^\lambda + |\tilde{\nu}_1 - \tilde{\nu}_2|^\lambda \right)^{\frac{1}{\lambda}} \quad (5)$$

其中:  $\tilde{\mu}_k, \tilde{\eta}_k, \tilde{\nu}_k (k = 1, 2)$  分别表示图犹豫不确定语言数  $\tilde{a}_k$  的积极隶属度、中隶属度和消极隶属度的平均值.

当  $\lambda = 1$  时,  $d(\tilde{a}_1, \tilde{a}_2)$  为  $\tilde{a}_1, \tilde{a}_2$  的 Hamming 距离; 当  $\lambda = 2$  时,  $d(\tilde{a}_1, \tilde{a}_2)$  为  $\tilde{a}_1, \tilde{a}_2$  的 Euclidean 距离.

上述定义的广义距离测度满足如下公理:

- 1)  $0 \leq d(\tilde{a}_1, \tilde{a}_2) \leq 1$ ;
- 2)  $d(\tilde{a}_1, \tilde{a}_2) = d(\tilde{a}_2, \tilde{a}_1)$ ;
- 3) 当且仅当  $\tilde{a}_1 = \tilde{a}_2$  时,  $d(\tilde{a}_1, \tilde{a}_2) = 0$ ;
- 4)  $d(\tilde{a}_1, \tilde{a}_3) \leq d(\tilde{a}_1, \tilde{a}_2) + d(\tilde{a}_2, \tilde{a}_3)$ .

**1.5 模糊测度、广义 Shapely 值和 Choquet 积分**

**定义8**<sup>[23]</sup> 设  $P(X)$  为  $X = \{x_1, x_2, \dots, x_n\}$  的幂集,  $c: P(X) \rightarrow [0, 1]$ , 给定  $\lambda \geq -1$ , 若  $c$  满足如下条件, 则称  $c$  为  $X$  上的  $c_\lambda$  模糊测度:

- 1)  $c(\phi) = 0, c(X) = 1$ ;
- 2)  $\forall E, F \in P(x)$ , 若  $E \subset F$ , 则  $c(E) \leq c(F)$ ;
- 3)  $\forall E, F \in P(x)$ , 若  $E \cap F = \phi$ , 则有  $c(E \cup F) = c(E) + c(F) + \lambda c(E)c(F)$ .

如果  $X$  代表决策时的属性集, 则  $\forall E, F \in P(X)$ ,  $c(E), c(F)$  分别表示属性  $E$  和属性  $F$  的权重. 当  $-1 \leq \lambda \leq 0$  时, 有  $c(E \cup F) > c(E) + c(F)$ , 表示信息存在冗余关系; 当  $\lambda = 0$  时,  $c(E \cup F) = c(E) + c(F)$ , 表示属性间相互独立; 当  $\lambda > 0$  时,  $c(E \cup F) < c(E) + c(F)$ , 表示属性之间存在信息互补的关系.

**定义9**<sup>[24]</sup> 设  $P(X)$  为  $X = \{x_1, x_2, \dots, x_n\}$  的幂集,  $\forall S \in P(X)$ , 其广义 Shapely 值定义为

$$\varphi(S) = \sum_{T \subseteq X \setminus S} \frac{(n-s-t)!t!}{(n-s+1)!} [c(S \cup T) - c(T)]. \quad (6)$$

其中:  $c$  为定义在  $X$  上的  $c_\lambda$  模糊测度,  $S$  为  $X$  上的任意子集,  $X \setminus S$  为  $X$  与  $S$  的差集,  $T$  为  $X \setminus S$  的子集,  $n, s, t$  分别表示集合  $X, S, T$  的基数. 当  $X$  中只包含一个

元素  $x$  时, 退化为 Shapely 函数, 即

$$\varphi(\{x_i\}) = \sum_{T \subseteq X \setminus x_i} \frac{(n-1-t)!t!}{n!} [c(x_i) - c(T)]. \quad (7)$$

**定义10**<sup>[16]</sup> 设  $X$  为一非空有限集, 若  $f$  为定义在  $X$  上的非负函数,  $c$  为定义在  $X$  上的模糊测度, 则  $f$  关于模糊测度  $c$  的离散 Choquet 积分为

$$\int f dc = f(x_{(i)}) [c(A_{(i)}) - c(A_{(i+1)})]. \quad (8)$$

其中:  $f(x_{(1)}) \leq f(x_{(2)}) \leq \dots \leq f(x_{(n)})$ ,  $A_{(i)} = \{x_i, x_{i+1}, \dots, x_n\}$  且  $A_{(n+1)} = \phi$ .

**2 图犹豫不确定语言广义 Shapely-Choquet 积分集成算子**

**定义11** 设  $\tilde{A} = (\tilde{a}_1, \tilde{a}_2, \dots, \tilde{a}_n)$  是一组定义在  $X$  上的图犹豫不确定语言数,  $\varphi(x)$  为定义在  $X$  的广义 Shapely 函数, 则定义  $\Omega^n \rightarrow \Omega$  的映射 PHULGSCWA 为

$$\int \tilde{A} d\varphi = \bigoplus_{i=1}^n \tilde{a}_{(i)} [\varphi(X_{(i)}) - \varphi(X_{(i+1)})], \quad (9)$$

称映射 PHULGSCWA 为图犹豫不确定语言广义 Shapely-Choquet 积分平均算子. 其中:  $\tilde{a}(x_1) \leq \tilde{a}(x_2) \leq \dots \leq \tilde{a}(x_n)$ ,  $X_{(i)} = \{x_i, x_{i+1}, \dots, x_n\}$  且  $X_{(n+1)} = \phi$ .

**定理1** 设  $\tilde{A} = (\tilde{a}_1, \tilde{a}_2, \dots, \tilde{a}_n)$  是一组定义在  $X$  上的图犹豫不确定语言数, 则  $\tilde{A}$  经图犹豫不确定语言广义 Shapely-Choquet 积分平均算子集成后的结果为

$$\int \tilde{A} d\varphi = \bigoplus_{i=1}^n \tilde{a}_{(i)} [\varphi(X_{(i)}) - \varphi(X_{(i+1)})] = \left\langle \left[ s \sum_{i=1}^n L(\tilde{a}_{(i)}), s \sum_{i=1}^n U(\tilde{a}_{(i)}) \right], \left\{ 1 - \prod_{i=1}^n (1 - \alpha_{(i)})^{[\varphi(x_{(i)}) - \varphi(x_{(i+1)})]} \right\}, \left\{ \prod_{i=1}^n (\beta_{(i)})^{[\varphi(x_{(i)}) - \varphi(x_{(i+1)})]} \right\}, \left\{ \prod_{i=1}^n (\gamma_{(i)})^{[\varphi(x_{(i)}) - \varphi(x_{(i+1)})]} \right\} \right\rangle. \quad (10)$$

其中:  $\tilde{a}(x_1) \leq \tilde{a}(x_2) \leq \dots \leq \tilde{a}(x_n)$ ,  $X_{(i)} = \{x_i, x_{i+1}, \dots, x_n\}$  且  $X_{(n+1)} = \phi$ ,  $\varphi(x_{(i)})$  为  $X_{(i)}$  的广义 Shapely 值.

**定理2** 设  $X = \{x_1, x_2, \dots, x_n\}$  为一非空有限集,  $c$  为定义在  $X$  上的  $c_\lambda$  模糊测度, 若  $X$  中的元素相互独立, 即  $\lambda = 0$ , 此时  $\varphi(X) = \varphi(x_1) + \varphi(x_2) + \dots + \varphi(x_n) = 1$ , 则 PHULGSCWA 算子退化成普通的图犹豫不确定语言加权平均算子.

**定理3** 设  $\tilde{A} = (\tilde{a}_1, \tilde{a}_2, \dots, \tilde{a}_n)$  是一组定义在  $X$  上的图犹豫不确定语言数,  $\varphi(S)$  为集合  $S$  的广义

Shapely 值,  $\tilde{a}_{\sigma_1}, \tilde{a}_{\sigma_2}, \dots, \tilde{a}_{\sigma_n}$  为  $\tilde{a}_1, \tilde{a}_2, \dots, \tilde{a}_n$  的任意一置换序列, 则有

$$\begin{aligned} & \text{PHFULGSCWA}(\tilde{a}_{\sigma_1}, \tilde{a}_{\sigma_2}, \dots, \tilde{a}_{\sigma_n}) = \\ & \text{PHFULGSCWA}(\tilde{a}_1, \tilde{a}_2, \dots, \tilde{a}_n). \end{aligned}$$

**定理4** 设  $\tilde{A} = (\tilde{a}_1, \tilde{a}_2, \dots, \tilde{a}_n)$  是一组定义在  $X$  上的图犹豫不确定语言数,  $\varphi(S)$  为集合  $S$  的广义 Shapely 值, 若定义  $\tilde{a}^+ = \langle [s_{L(x)}, s_{U(x)}]^+, \alpha^+, \beta^-, \gamma^- \rangle, \tilde{a}^- = \langle [s_{L(x)}, s_{U(x)}]^-, \alpha^-, \beta^+, \gamma^+ \rangle$ , 则有

$$\begin{aligned} & \text{PHFULGSCWA}(\tilde{a}^-, \tilde{a}^-, \dots, \tilde{a}^-) \leq \\ & \text{PHFULGSCWA}(\tilde{a}_1, \tilde{a}_2, \dots, \tilde{a}_n) \leq \\ & \text{PHFULGSCWA}(\tilde{a}^+, \tilde{a}^+, \dots, \tilde{a}^+). \end{aligned}$$

其中

$$\begin{aligned} & [s_{L(x)}, s_{U(x)}]^+ = [\max(s_{L(x)}, \min(s_{U(x)})), \max(s_{U(x)})], \\ & [s_{L(x)}, s_{U(x)}]^- = [\min(s_{L(x)}, \max(s_{U(x)})), \min(s_{U(x)})], \\ & \alpha^+ = \bigcup_{\alpha \in \tilde{\mu}(x)} \max\{\alpha\}, \alpha^- = \bigcup_{\alpha \in \tilde{\mu}(x)} \min\{\alpha\}, \\ & \beta^+ = \bigcup_{\beta \in \tilde{\eta}(x)} \max\{\beta\}, \beta^- = \bigcup_{\beta \in \tilde{\eta}(x)} \min\{\beta\}, \\ & \gamma^+ = \bigcup_{\gamma \in \tilde{\nu}(x)} \max\{\gamma\}, \gamma^- = \bigcup_{\gamma \in \tilde{\nu}(x)} \min\{\gamma\}. \end{aligned}$$

由定理4易得如下定理.

**定理5** 对于任意两组定义在  $X$  上的图犹豫不确定语言数  $\tilde{A} = (\tilde{a}_1, \tilde{a}_2, \dots, \tilde{a}_n), \tilde{B} = (\tilde{b}_1, \tilde{b}_2, \dots, \tilde{b}_n)$ , 若  $\forall i$ , 都有  $[s_{L(\tilde{a}_i)}, s_{U(\tilde{a}_i)}] \leq [s_{L(\tilde{b}_i)}, s_{U(\tilde{b}_i)}], \alpha_{\tilde{a}_i} \leq \alpha_{\tilde{b}_i}, \beta_{\tilde{a}_i} \geq \beta_{\tilde{b}_i}, \gamma_{\tilde{a}_i} \geq \gamma_{\tilde{b}_i}$ , 则有  $\text{PHFULGSCWA}(\tilde{a}_1, \tilde{a}_2, \dots, \tilde{a}_n) \leq \text{PHFULGSCWA}(\tilde{b}_1, \tilde{b}_2, \dots, \tilde{b}_n)$ .

### 3 基于图犹豫不确定语言集与用户评价的商品口碑分析模型

在网络购物中, 要了解一个商品各个属性的表现, 除了标定参数外, 还有一个重要途径就是商品的在线评论, 这能直接反映商品的用户体验, 为消费者在选择商品时提供参考. 与此同时, 厂商也能从商品评论中获取用户真正在意的商品属性. 假设  $C = \{C_1, C_2, \dots, C_m\}$  为商品  $D$  的属性集, 有相关消费群体  $A = \{A_1, A_2, \dots, A_n\}$  对商品的用户体验进行评价. 由于每个用户掌握的知识和信息不完全相同, 给出的评价不完全独立, 会存在着信息冗余、互补等情况. 利用爬虫技术, 收集不同用户的评价信息, 然后利用 python 的情感倾向分析, 将不同用户群体的评价信息转化成图犹豫不确定语言数, 形成以图犹豫不确定语言数构成的决策矩阵, 记为  $R = [r_{ij}]_{m \times n}$ . 最后, 用 PHFULGSCWA 算子对各个用户群体给出的评价意见

进行聚合, 得出各个属性的综合得分并进行排序, 为厂商在研发产品时提升产品竞争力提供参考.

#### 3.1 用户群体测度的求解模型

在实际决策过程中, 存在专家或属性权重难以确定的情形, 只能凭经验确定大致的权重范围. 为了求解出精确的权重, 引入  $c_\lambda$  测度和广义 Shapely 值, 使每个专家给出的评价对全局最优的效用值最大. 专家对属性的评价和最理想评价的距离越小, 专家评论对属性的最优评价贡献越大; 反之, 专家对属性的评价值和负理想评价值的距离越大, 专家评论对属性的最优评价贡献越大. 基于专家的评价与不完全确定的专家权重信息, 建立关于专家群体的模糊测度优化模型的步骤如下:

1) 确定属性  $c_i$  的正理想评价  $r_i^+$  和负理想评价  $r_i^-$ .

2) 根据下式求专家对各属性评论与负理想评价的贴进度系数:

$$D_{ij}^- = \frac{d_{ij}(r_{ij}, r_i^-)}{d_{ij}(r_{ij}, r_i^-) + d_{ij}(r_{ij}, r_i^+)}. \quad (11)$$

其中:  $i = 1, 2, \dots, m; j = 1, 2, \dots, n; d_{ij}(r_{ij}, r_i^-)$  为专家  $j$  对属性  $i$  的评价与属性  $i$  的负理想评价之间的距离.

3) 为了使各个专家的评价与负理想评价之间的距离达到最大, 构建以下最优化模型确定各个用户群体的模糊测度值:

$$\begin{aligned} & \max \sum_{i=1}^m \sum_{j=1}^n D_{ij}(r_{ij}, r_i^-) \varphi(A_j), A_j \subseteq A. \\ & \text{s.t.} \begin{cases} c(A) = 1; \\ c(S) \leq c(T) \leq 1, \forall S, T \in A, S \subseteq T; \\ c(A_j) \in U_j, c(A_j) \geq 0, j = 1, 2, \dots, n. \end{cases} \end{aligned} \quad (12)$$

其中:  $A = \{A_1, A_2, \dots, A_n\}, \varphi(A_j)$  为用户群体  $A_j$  的广义 Shapely 值,  $U_j$  为先验权重信息.

4) 通过 Matlab 求取不同用户群体的模糊测度值  $c(A_j)$ ,  $A_j$  为用户群体  $A$  的子集.

#### 3.2 数据的采集与预处理

通过对网站结构、URL 参数(统一资源定位器)、数据请求和响应规则的分析, 从京东商城中爬取某品牌笔记本电脑(产品定位为游戏本)的所有顾客在线评论. 对获得的数据进行如下操作:

1) 剔除空白评论, 商品评论中可能存在一些缺失值, 需要将其剔除.

2) 短句过滤和数据去重, 收集来的数据中存在重

复评论以及无意义的“默认好评”,要将其剔除。

3) 将去重后的数据划分为两类,分别为“自由评论”和“模板评论”,由于京东对电脑评论内置了模板,分别是“外形外观”“画面品质”“跑分评测”“运行速度”“游戏效果”。将自由评论剔除,保留模板评论。

4) 用python的字符串分割法对模板评论进行分割,将5个商品属性的评论分别保存。

5) 删去停用词,删除评论文本中无情感倾向的文本及标点,具体内容参考停用词表。

6) 调用SnowNLP库对评论进行情感分析,对每条评论进行打分。

### 3.3 利用图犹豫不确定语言数和PHFULGSCWA算子对商品属性排序

step 1: 利用原始数据经过预处理得到的商品属性情感得分的结果,给定语言术语集 $S$ 将商品评论转化为图犹豫不确定语言数组成的决策矩阵,由决策团队给出各个用户群体的权重范围。

step 2: 利用3.1节所述模型求取 $A_j$ 的模糊测度,使得各个专家评价与负理想评价的距离之和最大。

step 3: 将step2求得的结果代入式(6),求 $A_j$ 的广义Shapely值。

step 4: 利用PHFULGSCWA算子对不同用户群体 $A_j$ 对属性 $C_i$ 的评价值进行聚合,得到各属性 $C_i$ 的综合评价。

step 5: 根据定义6,求解各属性的得分函数值。

step 6: 比较各属性 $C_i$ 得分值的大小,对属性进行排序。

## 4 基于图犹豫不确定语言数与用户评价的商品属性分析的应用实例

本文选择在京东商城上爬取某品牌笔记本电脑的用户评论。其中,京东内置了评论模板,将评论模板中的“外形外观”“画面品质”“跑分评测”“运行速度”“游戏效果”作为待排序对象,为保证各个属性能得到客观评价,应分类计算plus会员和非plus会员( $A_1, A_2$ )对每一属性的评论。同时由于京东上的用户对电脑的专业知识掌握的水平参差不齐,本文还收集了网络上关于该笔记本的专业评测文章,作为用户群体( $A_3$ ),同样提取其中对“外形外观”“画面品质”“跑分评测”“运行速度”“游戏效果”的评价内容。对收集来的数据进行文本情感倾向分析,确定以图犹豫不确定语言数表示的决策矩阵,给定 $A_1, A_2, A_3$ 的权重范围,用优化模型来确定相应用户群体的模糊测度和广义Shapely值。最后,用PHFULGSCWA算子对各个用户群体给出的评价意见进行聚合,得出各个属性

的综合得分,分析用户群体对商品各种属性的重视程度并进行排序,为厂商在研发产品时提升产品竞争力提供参考。

### 4.1 数据收集与数据预处理

本文选择爬取京东在线购物平台和网络上大量数码博主对某品牌笔记本电脑的评论文章作为原始数据。获取的在线评论分为两种,分别为自主评论与模板评论,如表1所示,评测文章对每个属性的表现也有详细评价。

表1 评论内容分类表

用户名	级别	评论内容
北***枫	非plus会员	想换一台游戏本,从多方面了解后觉得今年拯救者系列性价比还是最高的,首发幸运入手.京东送货很快,第2天就邮寄到了.体验了两天感觉体验很好,sn730的固态,内存是三星双面颗粒双通道,显卡对我来说也完全够用了,3A大作只要不是最高画质都很流畅,此外,第1次用165Hz的屏幕感觉打开了新世界的大门,纵享丝滑
V***你	plus会员	外形外观:外观高大尚,有些灯很漂亮;画面品质:清晰度够用;跑分评测:反应很快,大游戏也很流畅;运行速度:开机几秒真的快,一直这样就可以了,希望以后售后更好

现对3种评论进行筛选,具体步骤如下:

1) 商品评论中可能存在一些缺失值,需要将其剔除。

2) 短句过滤和数据去重,收集来的数据中有许多重复的,以及无意义的“默认好评”“此用户暂未填写评论”,要将其剔除。

3) 设置特定识别对象,利用python中shell awk函数筛选出评论模板的位置下标。

4) 分别提取不同用户群体评价中属性信息,索引文本内容中“外形外观”“画面品质”“运行速度”以及“跑分评测”的字符串下标,做差截取每个属性对应的评论并分类保存。例如,评论为“运行速度:开机几秒真的快,一直这样就可以了,希望以后售后更好。外形外观:外观高大尚,有些灯很漂亮。”python自动识别出“运行速度”,并标记字符串“运行速度”的“运”位置为0号,“外形外观”的“外”位置为32号,利用字符串下标做差“32-1=31”,因此,截取“0~31”号下标的字符串序列“运行速度:开机几秒真的快,一直这样就可以了,希望以后售后更好”即为“运行速

度”对应的评价.

5) 删去停用词,即无关紧要的词,对得分结果不产生影响.

6) 利用 snowNLP 包中自带的计算方式,计算各个属性中每条评论的情感得分.

#### 4.2 利用图犹豫不确定语言数和PHFULGSCWA算子对商品属性排序

step 1: 将情感得分转化为图犹豫不确定语言数,并给定权重范围.

按属性分类提取的商品评论内容经 SnowNLP(txt). Sentiments 的计算后,返回一个该文本趋于积极的倾向的得分值记为  $a$  (该值在  $0 \sim 1$  之间). 在大数据背景下,情感得分遍布于  $0.1 \sim 0.99$  之间所有数,因此,存在多个积极隶属度、中隶属度与非

隶属度的情况. 给定语言术语集  $S = \{s_i | i = 0, 1, 2, \dots, 8\}$ . 其中:  $s_0$  表示极差,  $s_1$  表示非常差,  $s_2$  表示差,  $s_3$  表示较差,  $s_4$  表示不好不坏,  $s_5$  表示较好,  $s_6$  表示好,  $s_7$  表示非常好,  $s_8$  表示极好. 结合情感倾向得分值  $a$  将商品评论转化为图犹豫不确定语言数,具体方法如下: 不确定语言变量的下限  $s_{L(x)}$  取  $s_{L(x)=\lfloor g \times \bar{a} \rfloor}$ , 因为在线评论的好评率在 90% 以上, 所以不确定语言变量的上限为  $s_{U(x)=g}$ ; 为了便于利用 PHFULGSCWA 算子计算, 分别将积极隶属度、中隶属度和消极隶属度保留 1 到 2 个元素, 积极隶属度由评论正向情感得分确定, 取  $\tilde{\mu}(x) = \left\{ \frac{\lfloor a \times 10 \rfloor}{10}, \bar{a} \right\}$ ; 中隶属度和消极隶属度用 python 中 randint( $a, b$ ) 产生  $[0, 1 - \bar{a}]$  的随机数. 得到图犹豫不确定语言数表示的属性决策矩阵, 如表 2 所示.

表 2 以图犹豫不确定语言变量组成的商品属性决策矩阵

	非 plus 会员	plus 会员	媒体评测
外形外观	$\langle [s_6, s_8], \{0.855\ 3, 0.94\}, \{0.03\}, \{0.01\} \rangle$	$\langle [s_7, s_8], \{0.893\ 1, 0.980\ 6\}, \{0.02\}, \{0.03, 0.06\} \rangle$	$\langle [s_7, s_8], \{0.9, 0.98\}, \{0.01\}, \{0.01\} \rangle$
屏幕品质	$\langle [s_5, s_8], \{0.634\ 6, 0.692\ 4\}, \{0.05\}, \{0.05\} \rangle$	$\langle [s_6, s_8], \{0.755, 0.798\ 0\}, \{0.01\}, \{0.03, 0.02\} \rangle$	$\langle [s_6, s_8], \{0.882\ 7, 0.7\}, \{0.05\}, \{0.01\} \rangle$
跑分评测	$\langle [s_7, s_8], \{0.884\ 6, 0.964\ 6\}, \{0.01\}, \{0.01\} \rangle$	$\langle [s_6, s_8], \{0.846\ 9, 0.915\ 1\}, \{0.02\}, \{0.08\} \rangle$	$\langle [s_7, s_8], \{0.953\ 0\}, \{0.02\}, \{0.01, 0.02\} \rangle$
运行速度	$\langle [s_5, s_8], \{0.781\ 7, 0.85\}, \{0.05\}, \{0.07\} \rangle$	$\langle [s_6, s_8], \{0.834\ 1, 0.904\ 2\}, \{0.01\}, \{0.04\} \rangle$	$\langle [s_6, s_8], \{0.850\ 7, 0.9\}, \{0.02\}, \{0.03\} \rangle$
游戏体验	$\langle [s_5, s_8], \{0.736\ 2, 0.8\}, \{0.1\}, \{0.08\} \rangle$	$\langle [s_6, s_8], \{0.818\ 6, 0.882\ 5\}, \{0.03\}, \{0.01\} \rangle$	$\langle [s_7, s_8], \{0.934\ 7, 0.9\}, \{0.01\}, \{0.01\} \rangle$

权重范围的确定: 首先, 京东商城的会员等级与用户信誉、购物次数、购买力等因素相关, 即 plus 会员的评论的可信度要稍优于非 plus 会员; 其次, 媒体评测的文章是由专业知识较强的媒体人用专业软件及仪器测试之后给出的评价, 专业性较强, 可信度较高. 基于这两点原因, 给定用户群体  $\{A_1, A_2, A_3\}$  的权重

范围为  $\omega = \{0.25 \leq \omega_1 \leq 0.35, 0.25 \leq \omega_2 \leq 0.45, 0.35 \leq \omega_3 \leq 0.5\}$ .

step 2: 建立优化模型确定各个用户群体的模糊测度.

1) 确定属性  $c_i$  的正理想评价  $r_i^+$  和负理想评价  $r_i^-$ , 如表 3 所示.

表 3 各属性的正理想评价  $r_i^+$  与负理想评价  $r_i^-$

	正理想评价	负理想评价
外形外观	$\langle [s_7, s_8], \{0.980\ 0\}, \{0.01\}, \{0.01\} \rangle$	$\langle [s_6, s_8], \{0.855\ 3\}, \{0.03\}, \{0.06\} \rangle$
屏幕品质	$\langle [s_6, s_8], \{0.882\ 7\}, \{0.01\}, \{0.01\} \rangle$	$\langle [s_5, s_8], \{0.634\ 6\}, \{0.05\}, \{0.05\} \rangle$
跑分评测	$\langle [s_7, s_8], \{0.964\ 6\}, \{0.01\}, \{0.01\} \rangle$	$\langle [s_6, s_8], \{0.846\ 9\}, \{0.02\}, \{0.08\} \rangle$
运行速度	$\langle [s_6, s_8], \{0.904\ 2\}, \{0.01\}, \{0.03\} \rangle$	$\langle [s_5, s_8], \{0.781\ 7\}, \{0.05\}, \{0.07\} \rangle$
游戏体验	$\langle [s_7, s_8], \{0.934\ 7\}, \{0.01\}, \{0.01\} \rangle$	$\langle [s_5, s_8], \{0.736\ 2\}, \{0.03\}, \{0.08\} \rangle$

2) 利用式(5) ( $\lambda = 1$ ) 和(11), 求各个评价值与负理想评价的贴进度系数  $D_{ij}^-$ , 结果如表 4 所示.

3) 由模型(12) 建立以下优化模型求取各个用户群体的模糊测度:

$$\begin{aligned} \max & -0.720\ 4c(A_1) + 0.011\ 4c(A_2) + 0.708\ 9c(A_3) - \\ & 0.708\ 9c(A_1, A_2) - 0.011\ 4c(A_1, A_3) + \\ & 0.720\ 4c(A_2, A_3) + 2.876\ 2c(A_1, A_2, A_3). \end{aligned}$$

$$\text{s.t.} \begin{cases} c(A_1, A_2, A_3) = 1; \\ c(S) \leq c(T) \leq 1, \forall S, T \in \{A_1, A_2, A_3\}, S \subseteq T; \\ 0.25 \leq c(A_1) \leq 0.35, 0.25 \leq c(A_2) \leq 0.45, \\ 0.35 \leq c(A_3) \leq 0.5. \end{cases}$$

经 Matlab 求解得  $c(A_1) = 0.25, c(A_2) = 0.25, c(A_3) = 0.5, c(A_1, A_2) = 0.25, c(A_1, A_3) = 0.5, c(A_2, A_3) = 1, c(A_1, A_2, A_3) = 1$ .

step 3: 将用户集中各子集的模糊测度代入式(7), 得  $\varphi(A_1) = 0.0833, \varphi(A_2) = 0.3333, \varphi(A_3) = 0.5833, \varphi(A_1, A_2) = 0.375, \varphi(A_1, A_3) = 0.625,$

表4 用户群体评价与负理想评价的贴近度系数

	非 plus 会员	plus 会员	媒体评测
外形外观	0.288 87	0.640 21	0.874 89
屏幕品质	0.069 96	0.716 69	0.652 50
跑分评测	0.876 05	0.105 67	0.917 57
运行速度	0.104 27	0.862 44	0.881 37
游戏体验	0.096 38	0.574 47	0.936 46

$\varphi(A_2, A_3) = 0.875, \varphi(A_1, A_2, A_3) = 1$ .

step 4: 利用本文定义的 PHFULGSCWA 算子对属性的评价进行聚合, 得到各属性  $C_i (i = 1, \dots, 5)$  的综合评价值和得分函数, 如表5所示.

表5 PHFULGSCWA 聚合结果

	PHFULGSCWA 聚合结果	得分值 $s(C_i)$
外形外观	$\langle [s_{6.875}, s_8], \{0.8929, 0.9041, 0.9552, 0.9599, 0.9394, 0.9457, 0.9746, 0.9733\}, \{0.0145\}, \{0.0144, 0.0182\} \rangle$	$s(C_1) = s_{7.1}$
屏幕品质	$\langle [s_{5.875}, s_8], \{0.832, 0.8356, 0.7206, 0.7266, 0.8379, 0.8414, 0.7305, 0.7362\}, \{0.0292\}, \{0.0154, 0.0176\} \rangle$	$s(C_2) = s_{6.0}$
跑分评测	$\langle [s_{6.625}, s_8], \{0.924, 0.9391, 0.9277, 0.942\}, \{0.0194\}, \{0.0218, 0.0327\} \rangle$	$s(C_3) = s_{6.9}$
运行速度	$\langle [s_{5.875}, s_8], \{0.8385, 0.8459, 0.8624, 0.8687, 0.8722, 0.8781, 0.8911, 0.8961\}, \{0.018\}, \{0.0363\} \rangle$	$s(C_4) = s_{6.3}$
游戏体验	$\langle [s_{6.4583}, s_8], \{0.8967, 0.8988, 0.909, 0.9109, 0.8675, 0.8703, 0.8833, 0.8857\}, \{0.0138\}, \{0.013\} \rangle$	$s(C_5) = s_{6.7}$

step 5: 根据得分值越大越优的原则, 得到属性的排序结果为  $C_1 \succ C_3 \succ C_5 \succ C_4 \succ C_2$ , 故  $C_1$  (外形外观) 为顾客最在意的属性值, 其次是顾客得以直观了解电脑性能的跑分评测结果  $C_3$ , 再次就是游戏体验  $C_5$  与运行速度  $C_4$ , 最后是屏幕品质  $C_2$ . 以上得到的结果与本产品的游戏笔记本的定位相符, 符合购买游戏本的消费者心理预期.

针对高评分属性, 厂家的营销部门可以对此加大宣传力度, 以便得到更多消费者青睐; 针对低评分属性, 产品设计部门应采取适当措施, 调整屏幕设计, 例如从材料、分辨率、清晰度等方面入手加以改善.

### 4.3 对比分析

为了说明本文决策方法的有效性, 将本文数据用图犹豫不确定语言加权平均算子 (PHULWA) 进行聚合, 由定义8可知, 用户群体的模糊测度可以直接作为其权重, 故取各用户群体的权重为  $\omega = (0.25, 0.25, 0.5)$ . 经计算: PHULWA 算子聚合结果为  $s(C_1) = s(C_3) \succ s(C_5) \succ s(C_4) \succ s(C_2)$ , 无法区分  $C_1$  和  $C_2$  的优劣; PHFULGSCWA 算子的聚合结果为  $s(C_1) \succ s(C_3) \succ s(C_5) \succ s(C_4) \succ s(C_2)$ , 区分度

较好. PHULWA 算子聚合时默认用户群体之间是相互独立的, 而 PHFULSCWA 算子引入模糊测度、广义 Shapely 值和 Choquet 积分, 考虑了决策群体对全局的贡献, 以及决策群体之间的相互影响, 因此更符合实际决策情况.

### 4.4 参数敏感性分析

在利用式(5)求解用户评论与正、负理想评价之间的距离的过程中, 涉及参数  $\lambda$  的选取, 为了理解参数  $\lambda$  的变化趋势对用户群体模糊测度的影响以及对最终决策结果的影响, 分别选取参数  $\lambda = 1, 2, \dots, 6$  对本文数据进行分析. 求得各个用户群体的模糊测度如图1所示. 可以发现, 当  $\lambda > 2$  时,  $c(A_1, A_3)$  发生了改变, 经计算得到最终的排序结果为  $C_3 \succ C_1 \succ C_5 \succ C_4 \succ C_2$ . 通过对比分析可知, 当  $\lambda = 3, 4, 5, 6$  与  $\lambda = 1, 2$  时的结果不同体现在“外形外观”和“跑分评测”的顺序上. 而“外形外观”和“跑分评测”的在线评论经 PHFULWA 算子聚合后的结果得分值相等, 所以这两种排序结果都是合理的. 因此, 决策者可以根据其风险偏好选择合适的参数进行决策.

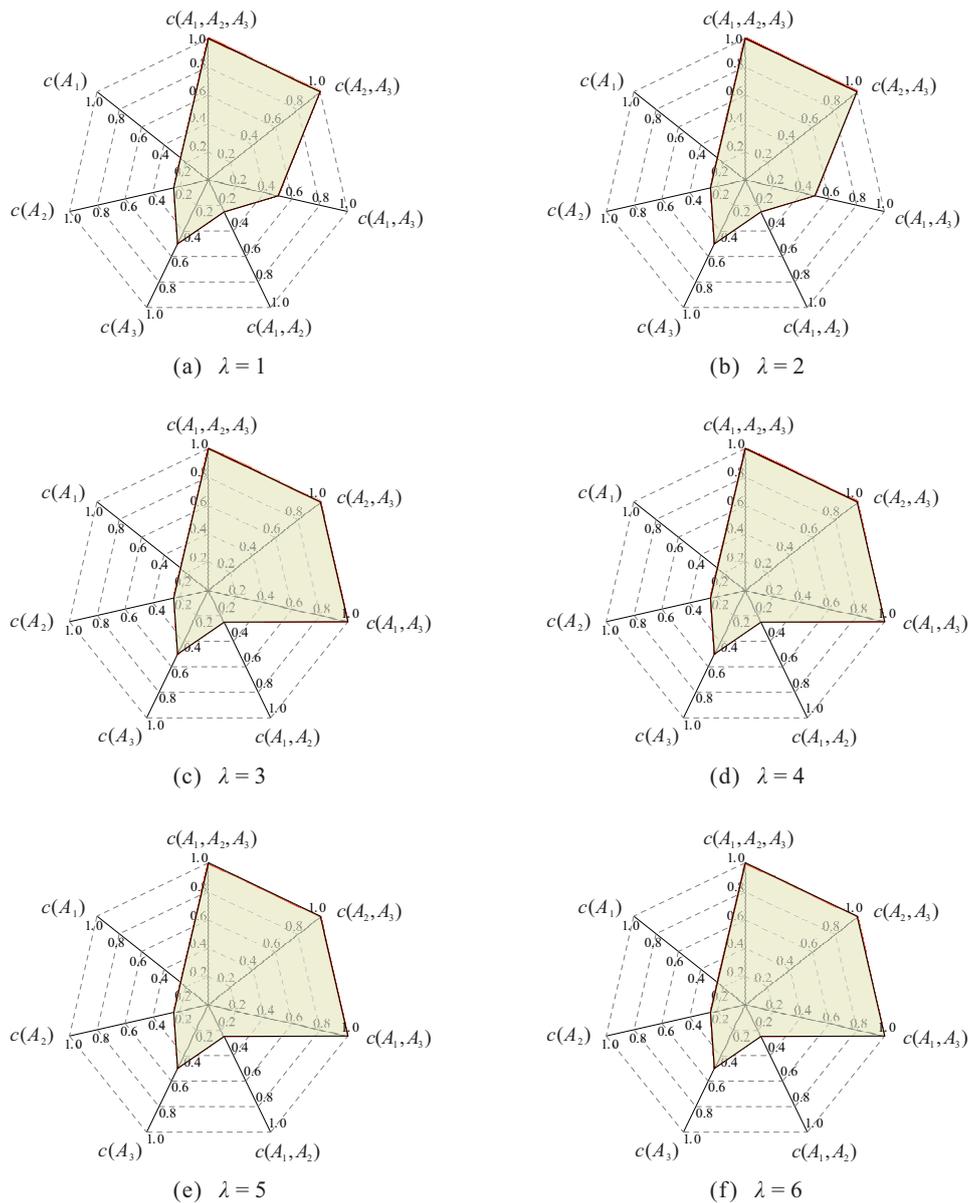


图1 参数λ对用户群体模糊测度的敏感性分析

### 5 结论

本文将图犹豫模糊集与不确定语言变量相结合, 将图犹豫模糊集拓展为图犹豫不确定语言集, 提出了一种利用图犹豫不确定语言数和用户评论的商品属性分析方法, 并应用于笔记本电脑的商品属性分析. 首先, 本文基于不确定语言变量和图犹豫模糊集, 定义了图犹豫模糊不确定语言数及其运算, 构建了图犹豫不确定语言数的得分函数、精确函数与广义距离测度. 其次, 定义了图犹豫不确定语言集环境下的广义 Shapely-Choquet 加权平均算子, 并研究了该算子的有界性、单调性和置换不变性等性质. 该算子可以考虑不同聚合参数间的交互关系, 且可以考虑不同聚合参数对全局的贡献率. 然后, 针对专家权重不完全已知的情况, 构建优化模型求解其模糊测度和广义

Shapely 值. 最后, 使用 python 提取京东商城某品牌笔记本电脑的在线评论, 并利用 SnowNLP 库计算每个属性评论的情感得分, 将其转化为图犹豫不确定语言集, 用所提出的 PHULGSCWA 算子对笔记本电脑的 5 种属性排序, 并针对排序结果对厂家及商家提出了合理化建议.

### 参考文献(References)

[1] Lin X L, Featherman M, Brooks S L, et al. Exploring gender differences in online consumer purchase decision making: An online product presentation perspective[J]. Information Systems Frontiers, 2019, 21(5): 1187-1201.

[2] 杜学美, 丁璟好, 谢志鸿, 等. 在线评论对消费者购买意愿的影响研究[J]. 管理评论, 2016, 28(3): 173-183. (Du X M, Ding J Y, Xie Z H, et al. An empirical study on the impact of online reviews on consumers'

- purchasing intention[J]. *Management Review*, 2016, 28(3): 173-183.)
- [3] Li X L, Wu C J, Mai F. The effect of online reviews on product sales: A joint sentiment-topic analysis[J]. *Information & Management*, 2019, 56(2): 172-184.
- [4] 王颖洁, 朱久祺, 汪祖民, 等. 自然语言处理在文本情感分析领域应用综述[J]. *计算机应用*, 2022, 42(4): 1011-1020.  
(Wang Y J, Zhu J Q, Wang Z M, et al. Review of applications of natural language processing in text sentiment analysis[J]. *Journal of Computer Applications*, 2022, 42(4): 1011-1020.)
- [5] Zadeh L A. Fuzzy sets[J]. *Information and Control*, 1965, 8(3): 338-353.
- [6] Atanassov K T. Intuitionistic fuzzy sets[J]. *Fuzzy Sets and Systems*, 1986, 20(1): 87-96.
- [7] Cường B C. Picture fuzzy sets[J]. *Journal of Computer Science and Cybernetics*, 2015, 30(4): 409-420.
- [8] Torra V. Hesitant fuzzy sets[J]. *International Journal of Intelligent Systems*, 2010, 25(6): 529-539.
- [9] Wang R, Li Y L. Picture hesitant fuzzy set and its application to multiple criteria decision-making[J]. *Symmetry*, 2018, 10(7): 295.
- [10] Zadeh L A. The concept of a linguistic variable and its application to approximate reasoning— I[J]. *Information Sciences*, 1975, 8(3): 199-249.
- [11] Zadeh L A. The concept of a linguistic variable and its application to approximate reasoning— II[J]. *Information Sciences*, 1975, 8(4): 301-357.
- [12] Chen Z C, Liu P H, Pei Z. An approach to multiple attribute group decision making based on linguistic intuitionistic fuzzy numbers[J]. *International Journal of Computational Intelligence Systems*, 2015, 8(4): 747.
- [13] 刘培德, 张新. 直觉不确定语言集成算子及在群决策中的应用[J]. *系统工程理论与实践*, 2012, 32(12): 2704-2711.  
(Liu P D, Zhang X. Intuitionistic uncertain linguistic aggregation operators and their application to group decision making[J]. *Systems Engineering—Theory & Practice*, 2012, 32(12): 2704-2711.)
- [14] Garg H. Some picture fuzzy aggregation operators and their applications to multicriteria decision-making[J]. *Arabian Journal for Science and Engineering*, 2017, 42(12): 5275-5290.
- [15] 罗世华, 方童, 刘俊. 概率区间值直觉犹豫模糊 Maclaurin 对称平均算子及决策方法[J]. *控制与决策*, 2021, 36(5): 1249-1258.  
(Luo S H, Fang T, Liu J. Probabilistic interval-valued intuitionistic hesitant fuzzy Maclaurin symmetric mean operators and decision method[J]. *Control and Decision*, 2021, 36(5): 1249-1258.)
- [16] Murofushi T, Sugeno M. An interpretation of fuzzy measures and the Choquet integral as an integral with respect to a fuzzy measure[J]. *Fuzzy Sets and Systems*, 1989, 29(2): 201-227.
- [17] 陈岩, 李庭. 基于 Choquet 积分的直觉不确定语言信息集结算子及其应用[J]. *控制与决策*, 2016, 31(5): 842-852.  
(Chen Y, Li T. Intuitionistic uncertain linguistic information aggregation operators based on Choquet integral and their application[J]. *Control and Decision*, 2016, 31(5): 842-852.)
- [18] Marichal J L. An axiomatic approach of the discrete Choquet integral as a tool to aggregate interacting criteria[J]. *IEEE Transactions on Fuzzy Systems*, 2000, 8(6): 800-807.
- [19] Shapley L S. A value for  $n$ -persons games[J]. *Annals of Mathematics Studies*, 1953, 28(7): 307-318.
- [20] Zadeh L A. The concept of a linguistic variable and its application to approximate reasoning— III[J]. *Information Sciences*, 1975, 9(1): 43-80.
- [21] 徐泽水. 基于语言标度中术语指标的多属性群决策法[J]. *系统工程学报*, 2005, 20(1): 84-88.  
(Xu Z S. A multi-attribute group decision making method based on term indices in linguistic evaluation scales[J]. *Journal of Systems Engineering*, 2005, 20(1): 84-88.)
- [22] Xu Z S. Uncertain linguistic aggregation operators based approach to multiple attribute group decision making under uncertain linguistic environment[J]. *Information Sciences*, 2004, 168(1/2/3/4): 171-184.
- [23] Xu Z S. Choquet integrals of weighted intuitionistic fuzzy information[J]. *Information Sciences*, 2010, 180(5): 726-736.
- [24] Tan C Q. A multi-criteria interval-valued intuitionistic fuzzy group decision making with Choquet integral-based TOPSIS[J]. *Expert Systems with Applications*, 2011, 38(4): 3023-3033.

## 作者简介

杨志辉(1975—), 男, 教授, 博士, 硕士生导师, 从事模糊决策、质量管理等研究, E-mail: yang\_zhihui@aliyun.com;  
俞弋峰(1996—), 男, 硕士生, 从事模糊决策的研究, E-mail: 1281066589@qq.com.