



融合可拓学与在线评论挖掘的产品改进需求识别研究

于志刚, 成思源, 杨雪荣, 谢通

引用本文:

于志刚, 成思源, 杨雪荣, 谢通. 融合可拓学与在线评论挖掘的产品改进需求识别研究[J]. 智能系统学报, 2023, 18(5): 1047–1059.

YU Zhigang, CHENG Siyuan, YANG Xuerong, et al. Product improvement demand identification incorporating topology and online comment mining[J]. *CAAI Transactions on Intelligent Systems*, 2023, 18(5): 1047–1059.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202208052>

您可能感兴趣的其他文章

半监督自训练的方面提取

Aspects extraction based on semi-supervised self-training

智能系统学报. 2019, 14(4): 635–641 <https://dx.doi.org/10.11992/tis.201806006>

酒店在线评论数据的特征挖掘

Feature mining based on online hotel review

智能系统学报. 2018, 13(6): 1006–1014 <https://dx.doi.org/10.11992/tis.201806016>

基于ArcGIS的可拓住宅建筑设计数据库构建研究

Construction of extension residential architectural design database based on ArcGIS

智能系统学报. 2018, 13(2): 214–219 <https://dx.doi.org/10.11992/tis.201708006>

基于改进可拓学第三创造法的产品创新设计

Product innovation design using the modified third creation method

智能系统学报. 2017, 12(1): 38–46 <https://dx.doi.org/10.11992/tis.201610016>

基于因果分析的事元蕴含系及其应用

A fair-elements implication system and its application based on causal analysis

智能系统学报. 2017, 12(1): 60–66 <https://dx.doi.org/10.11992/tis.201610012>

融合情感极性和逻辑回归的虚假评论检测方法

Detection of fake reviews based on emotional orientation and logistic regression

智能系统学报. 2016, 11(3): 336–342 <https://dx.doi.org/10.11992/tis.201603027>

DOI: 10.11992/tis.202208052

网络出版地址: <https://kns.cnki.net/kcms2/detail/23.1538.TP.20230615.0950.002.html>

融合可拓学与在线评论挖掘的产品改进需求识别研究

于志刚, 成思源, 杨雪荣, 谢通

(广东工业大学机电工程学院, 广东 广州 510006)

摘要: 利用形式化、条理化的手段从在线评论当中挖掘并理清不同类别的产品改进需求是产品改进中亟待解决的问题, 考虑到产品设计信息的多层次、多特征性, 提出一种基于基元模型的在线评论产品设计信息分层表示及识别产品改进需求的可拓分析方法。利用 PYLDA-vis 交互式可视化库从在线评论中提取产品要素, 将产品要素划分为实体-功能-属性 3 个层次, 并采用物元模型进行分层表达; 结合情感倾向点互信息算法计算各个层次用户观点的评价值, 借助事元模型对评价值较低的对象特征进行用户需求表达; 利用相关网分析方法确定缺陷物元之间的关联关系, 蕴含系分析方法找出产品改进需求的实现途径。以一款养生壶的在线评论为例, 通过对在线评论的挖掘找到用户对该产品的改进需求, 并识别出实现需求的途径。融合可拓学与在线评论挖掘的产品改进需求识别方法可实现用户需求的分层表达, 形式化、条理化地理清产品改进设计信息, 为识别产品改进途径提供科学依据和有效方法。

关键词: 可拓学; 在线评论; 物元模型; 事元模型; 情感倾向点互信息算法; 相关网; 蕴含系; 需求识别

中图分类号: TP18 **文献标志码:** A **文章编号:** 1673-4785(2023)05-1047-13

中文引用格式: 于志刚, 成思源, 杨雪荣, 等. 融合可拓学与在线评论挖掘的产品改进需求识别研究 [J]. 智能系统学报, 2023, 18(5): 1047-1059.

英文引用格式: YU Zhigang, CHENG Siyuan, YANG Xuerong, et al. Product improvement demand identification incorporating topology and online comment mining[J]. CAAI transactions on intelligent systems, 2023, 18(5): 1047-1059.

Product improvement demand identification incorporating topology and online comment mining

YU Zhigang, CHENG Siyuan, YANG Xuerong, XIE Tong

(School of Electromechanics Engineering, Guangdong University of Technology, Guangzhou 510006, China)

Abstract: Formal and organized mining and clarifying of different categories of product improvement demands from online comments are necessary for the product improvement process. On considering the multilevel and multifeature nature of product design information, a topological analysis method based on a matter-element model used for hierarchical representation of online review product design information and identification of product improvement demands was proposed. PYLDAvis interactive visualization library was used to extract product elements from online comments, and the product elements were then divided into three levels, namely entity, function, and attribute. In addition, the matter-element model was used for hierarchical expression. The emotion inclination pointwise mutual information algorithm was combined to calculate the evaluation value of user opinions at all levels, and an affair-element model was used to express the user's demand for the object features with low evaluation value. The correlation network analysis method was used to determine the relationship between defective matter elements, and the implication system analysis method was used to determine the realization of the product improvement demand. The case study took the online reviews of a health pot as an example, the user's improvement demands for the product were found by mining online reviews, and means to achieve the demands were identified. A product improvement demand identification method incorporating topology and online comment mining can realize the hierarchical expression of user demands and clarify product improvement design information in a formalized and organized manner, providing a scientific basis and effective methods for identifying the improvement paths of products.

Keywords: topology; online comment; matter-element model; affair-element model; emotion inclination pointwise mutual information algorithm; correlative network; implication system; demand identification

收稿日期: 2022-08-31. 网络出版日期: 2023-06-15.

基金项目: 广东省研究生教育创新计划项目(2021JGXM043);
广东省科技计划项目(2014A040402004).

通信作者: 成思源. E-mail: imdesign@gdut.edu.cn.

面对迅速更替的新兴技术与需求不断变化的消费者, 产品和服务逐渐趋向于多样化^[1]. 对于产品设计师而言, 从不断更迭的用户反馈中准确

地获取用户需求并找到产品改进的根本途径,将有助于产品核心竞争力的提升,从而推动产品更新换代的可持续发展。然而,从用户反馈当中显在需求信息难以直接对设计者产生启发。因此,如何从用户当中挖掘出有价值的需求信息,进而为产品改进策略提供帮助,无疑成为了广大产品设计师所关注的焦点^[2]。

在产品改进设计的过程中,及时获得用户需求至关重要。传统的方法依旧采用问卷调查,用户访谈等方法,存在周期长,成本高的劣势,故而难以提高快速更迭的产品研发需求。得益于电商行业与社交网络的迅猛发展,在线评论成为了产品缺陷获取的有效途经和用户需求载体。电商平台(京东、淘宝等)实时的在线评论,描述了消费者对产品最直接的感受,在一定程度上代表了消费者对产品的观点,可以为产品设计师提供最为及时、真实、准确的客户反馈信息,为产品改进增添了新的策略与方向。

1 相关工作

在产品改进设计当中,通过挖掘庞大的在线评论来获取产品用户需求已经成为了学者们研究的热点。目前,研究者们运用各种数据挖掘方法分析在线评论当中有用的产品信息,研究方向主要聚焦于产品特征挖掘与用户情感研究。产品特征挖掘,是指从众多的在线评论当中采用数据挖掘工具或深度学习算法获取用户关注的产品特征^[3]。徐凯^[4]提出了基于 Apriori 算法及中文句子组成规律相组合的产品特征抽取方式,借助自然语言处理和文本挖掘技术对在线评论数据中的产品特征信息进行提取。李贺等^[5]利用 LDA 主题模型对多个品牌的手机在线评论进行主题和产品特征挖掘,并且有效地识别了用户需求要素。用户情感研究,指主要利用自然语言处理技术来获取潜在客户观点、文本倾向及情感状态^[6]。Yeh 等^[7]将感性工学与 SOM 聚类算法有效的结合并运用,最终获得了与消费者情感相关的产品设计因素。李江泳等^[8]运用情感倾向点互信息算法 (semantic orientation pointwise mutual information algorithm, SO-PMI) 对用户关心的产品属性进行了量化计算,通过量化结果分析出产品缺陷并应用技术创新方法对产品缺陷进行改进设计。

从不同的产品对象和特征进行描述,所表达的产品细腻度不同,无论是基于产品对象进行表达的在线评论还是基于产品特征都是具有层次性的^[9-10]。纪雪等^[11]考虑产品属性的层次性对在线

评论获取用户需求。肖寒琼等^[12]以三层次体验理论为基础,将手机评论的产品特征划分为本能层-行为层-反思层,用于分析用户对产品的偏好程度。然而,从在线评论获取用户需求的过程中,频繁使用与评论相关且无对应关系的自然语言去描述,不利于产品关键信息的提取与表达,并且缺乏描述事物本身的内部组成结构及层次关系。为此,提出结合可拓学基元模型形式化、模型化的方式表达出在线评论中的产品设计信息,使得信息表达更加全面、细腻。

可拓学中以基元模型来描述被研究对象事物间关系的特征,分析事物拓展规律及研究创新的可拓变换工具,具有统一、方便和易于操作的特点。张文旭等^[13]提出产品性能词典,对在线评论运用数据驱动手段来获取用户需求,并采用可拓学物元模型对用户需求进行形式化表达。刘晓敏等^[14]基于基元理论对产品进行多层次知识建模,构建仿生产品系统功能-原理-行为-结构的本体可拓模型。

上述研究利用基元模型实现了对在线评论的设计知识表达,但未考虑产品设计信息的多层次、多特征性;其次,针对不同类型的用户需求,实现方式也不尽相同,为了形式化、条理化地理清不同类别的用户需求实现的方法,对获取的用户需求进行相关推理及实现需求层次识别至关重要。对此,提出一种融合可拓学与在线评论层次性的设计知识建模的方法,借助产品的实体-功能-属性的层次关系,以产品对象、对象特征和用户观点为核心进行用户需求识别研究。

2 研究方法

2.1 方法流程概述

本文通过分析在线评论的基本特点,基于可拓学基元模型,提出了在线评论层次的表达方法对用户评论数据进行产品改进需求识别,针对各个层次模型计算相应的用户评价价值,为挖掘产品缺陷及用户需求提供指导依据,基本流程如图 1 所示。

1) 在线评论分析。通过爬虫工具采集产品在线评论数据,数据分析工具对在线评论文本进行预处理。

2) 在线评论层次模型。利用 PYLDA-vis 交互式可视化库提取出以产品对象和对象特征为主的产品要素。根据产品的实体-功能-属性的层次关系对在线评论产品进行划分,达到在线评论层次划分的目的,将产品要素匹配映射到相应的层次当中,完成在线评论层次的构建,并借助可拓学物元模型对在线评论层次进行表达。

3) 用户观点评价价值计算。在上述分析的基础上,利用 SO-PMI 算法计算出评论语句的评价价值,基于模型化、规范化表达的在线评论层次物元模型,再计算出各个层次用户观点的评价价值,通过分析用户观点的评价价值,对存在缺陷的产品对象

利用可拓学事元模型表达。

4) 产品改进需求识别。对存在缺陷的事元模型与在线评论层次进行分析,利用相关网分析方法确定缺陷物元之间的关联关系,利用蕴含系分析方法识别出产品改进的根本途径。

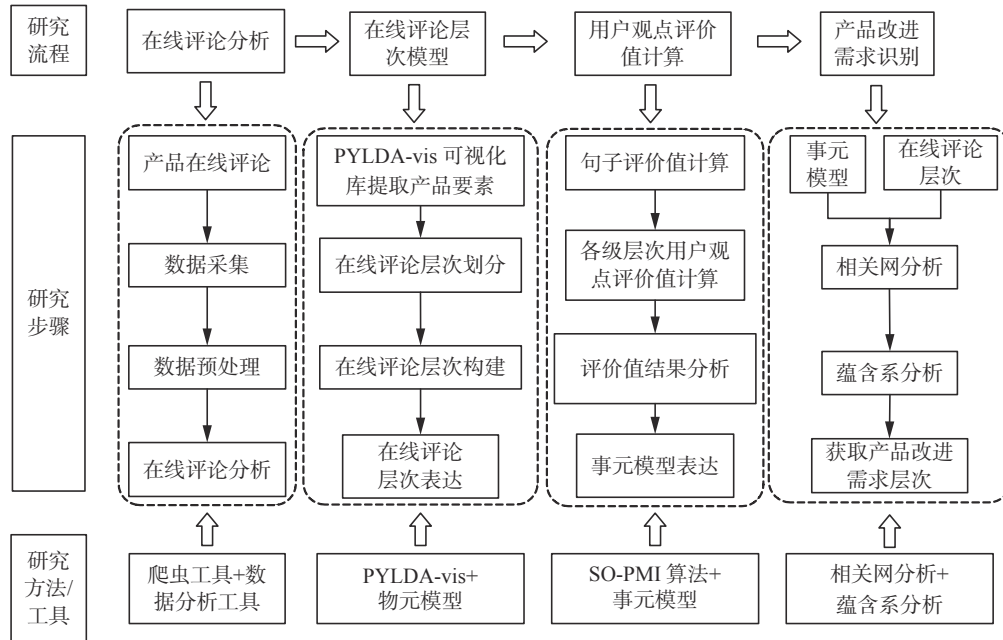


图1 产品改进需求识别基本流程

Fig. 1 Basic process of product improvement requirement identification

2.2 在线评论分析

首先借助网络爬虫工具,选择恰当的电商平台并抓取目标产品的用户评论文本,并将评论文本中对相关研究无意义的内容进行文本预处理,处理后保存最初的评论数据集用于后续相关研究。

用户通过在线评论的方式,来表达对于产品对象的评价或者对于对象特征的评价。如一款茶壶的评论“茶壶的容量有点小”,“茶壶”是产品对象,“容量”是产品对象的特征,“有点小”是用户对该对象特征的观点。从中可以看出,一条评论可以表示为<产品对象,对象特征,用户观点>的三元组模型^[13]:

$$N = \langle O_n, C_n, V_n \rangle \quad (1)$$

式中: O_n 为产品对象, C_n 为对象特征, V_n 为关于 O_n 的用户观点。对于同一款产品,其产品对象 O_n 的数量是有限的,对象特征 C_n 的数量同样是有限的。通常一个产品的产品对象具有多个特征,而一个对象特征往往被不同的用户观点所评价,因此,利用单个层次的三元组模型很难清晰地对用户的评价进行表达。

2.3 在线评论层次模型

2.3.1 基于 PYLDA-vis 可视化库提取产品要素

产品要素表达了与产品相关的设计知识,包

括产品对象和对象特征,在评论当中以词汇形式出现,学者们通常采用 LDA (latent dirichlet allocation) 主题模型来提取产品相关的设计知识。LDA 主题模型于 2003 年被学者首次提出应用于文本分类当中^[15],该主题模型引入了潜在狄利克雷分布,很好地解决了由于短文本类型造成传统主题模型中的过拟合问题^[16]。本文利用 LDA 主题模型进行主题词生成之后,并使用已有的基于 Python 开发的可视化分析工具 PYLDA-vis,该工具具有人机交互的调节面板,可以直观地显示出主题与词汇之间的关联信息,最终形成在线评论主题聚类效果来提取产品要素^[17]。假设数据集中包含 K 个主题和 W 个词汇,词汇 w 对主题 k 的关联度 r 可以表示为

$$r(w, k | \lambda) = \lambda \log(\phi_{kw}) + (1 - \lambda) \log\left(\frac{\phi_{nw}}{p_w}\right) \quad (2)$$

式中: ϕ_{kw} 被定义为词语 w 出现在主题 k 中的概率, p_w 被定义为词语 w 在语料库中的边界概率, λ 为权重参数 ($0 \leq \lambda \leq 1$)。

基于 PYLDA-vis 交互式可视化库,给出如图 2 所示的产品要素提取流程。首先,为了使得描述产品相关的信息能够准确进行分词,预先加载分

词词典,并借助分词工具对在线评论分词;分词后,加载停用词词典,将一些与产品信息描述无关的词语进行去除,加快后续产品要素提取的速度;最后,将清洗后的分词结果以二维矩阵的形式导入 PYLDA-vis 交互式可视化库中,并给定主题个数,得到最终的可视化主题聚类效果图,选定同一个主题,并调节权重参数 λ 来选择符合要求的产品要素。该流程能够快速灵活的探索主题和词汇之间的关联关系,方便地提取出每个主题下与产品设计知识相关的词汇。

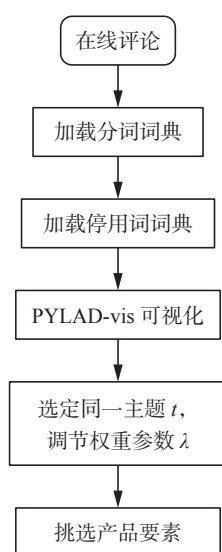


图 2 产品要素提取流程
Fig. 2 Product element extraction process

2.3.2 在线评论层次划分

为了充分挖掘用户关注的产品信息特征及与产品相关的产品设计信息,研究不同用户需求之间的关系及实现方式,根据前人的理论研究,对产品对象进行多层次知识建模,将产品对象依次划分为实体层、功能层和属性层。实体层代表了产品给用户的具体化表现形式,是指产品功能实现所需物理支撑的载体,例如:养生壶的茶壶、壶盖等;功能层是指产品内部固有的效能,由产品内部要素实体结构所决定,例如:养生壶具有保温、预约功能等;属性层是指产品本身表现出来的性质,例如:养生壶的价格、品牌等。

在产品的开发过程中,设计者往往会对待开发产品实体层的设计结构进行详细的展示。有学者将专利产品的实体进行了划分,依次划分为系统、子系统和特征结构^[18],通过对产品实体进行分层划分,可以更加清晰地表达专利产品设计和开发者的设计意图。结合以上研究者的分析方式,本文将实体层进一步划分为子系统和特征结构,例如,养生壶是一个系统,茶壶、杯子就可

以分别作为一个子系统,壶盖和壶体就可以看作是茶壶的特征结构。

有关研究表明,在获取用户需求的过程中,由产品开发者构建的、与由用户构建隐含的产品层次是不同的^[19]。整个产品层次的划分主要依据产品设计知识库,典型的方式是通过产品规格说明书进行划分^[20]。产品规格说明书在产品开发者和消费者之间起到纽带作用。因此,可以通过产品规格说明书来调整和优化产品层次的表述及层次区域的大小。接下来,先将在线评论当中提取的产品对象与各个产品层次区域进行匹配映射,然后将对象特征匹配映射到相应的产品对象中,从而达到在线评论层次划分的目的。在线评论层次映射关系如图 3 所示。将在线评论当中提取到的产品要素按照映射关系划分到不同层次的子系统或特征结构当中,使得在线评论被逐级分解到不同产品模块下的不同区域内,达到了分类组合的效果,从而更好地挖掘用户需求。

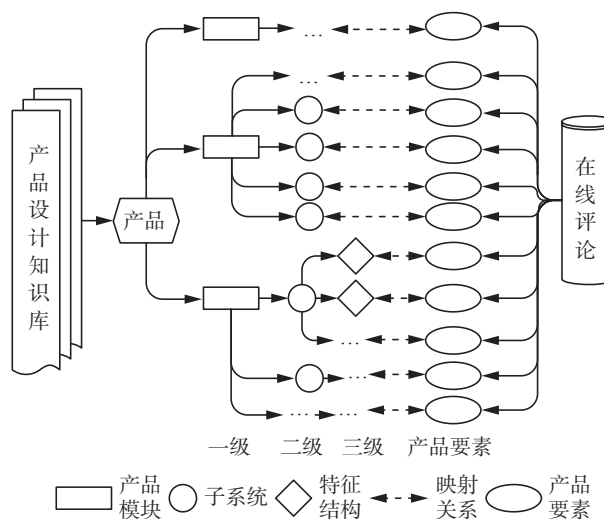


图 3 在线评论层次映射关系

Fig. 3 Online comment hierarchy mapping relationship

2.3.3 在线评论层次表达

在可拓学中采用以物元、事元和关系元为核心的基元表达形式来形式化、模型化地表达事物。其中基元的组成结构包括对象 O 、特征 C 和量值 V ,一般采用一维基元或多维基元的表达方式来描述对象的各个特征和参数,其中多维基元的一般表达式为^[21]

$$B = (O, C, V) = \begin{bmatrix} O_b, & c_1, & v_1 \\ & c_2, & v_2 \\ & \vdots & \vdots \\ & c_n, & v_n \end{bmatrix} \quad (3)$$

物元为描述事物提供了量化表达工具,本文将根据在线评论划分的实体层-功能层-属性层

采用 n 维物元模型 $M_{H_{Si}}$ 表达有:

$$M_{H_{Si}} = \begin{bmatrix} H_{Si}, & O_{S1}, & \{S_{S1}\} \\ & O_{S2}, & \{S_{S2}\} \\ & \vdots & \vdots \\ & O_{Sn}, & \{S_{Sn}\} \end{bmatrix} \quad (4)$$

式中: $H_{Si} (i=1,2,\dots,n)$ 表示划分的实体层-功能层-属性层, $O_{Si} (i=1,2,\dots,n)$ 表示该层下的产品对象, $\{S_{Si}\}$ 表示该产品对象的对象特征集合 $\{S_{Si1}, S_{Si2}, \dots, S_{Sin}\}$ 。每个产品对象的物元模型 $M_{O_{Si}}$ 可以表示为

$$M_{O_{Si}} = \begin{bmatrix} O_{Si}, & S_{Si1}, & v_{S1} \\ & S_{Si2}, & v_{S2} \\ & \vdots & \vdots \\ & S_{Sin}, & v_{Sn} \end{bmatrix} \quad (5)$$

特别的,对于实体层的产品对象可以继续划分,进一步划分该产品对象下的特征结构,每个 O_{Si} 可按特征结构组成,采用 n 维物元 $M_{H_{Ti}}$ 表示为

$$M_{H_{Ti}} = \begin{bmatrix} H_{Ti}, & O_{T1}, & \{T_{T1}\} \\ & O_{T2}, & \{T_{T2}\} \\ & \vdots & \vdots \\ & O_{Tn}, & \{T_{Tn}\} \end{bmatrix} \quad (6)$$

式中: $H_{Ti} (i=1,2,\dots,n)$ 表示子系统的特征结构, $O_{Ti} (i=1,2,\dots,n)$ 表示特征结构, $\{T_{Ti}\}$ 表示该特征结构的对象特征集合 $\{T_{Ti1}, T_{Ti2}, \dots, T_{Tin}\}$ 。每个特征结构的物元模型 $M_{O_{Ti}}$ 可以表示为

$$M_{O_{Ti}} = \begin{bmatrix} O_{Ti}, & T_{Ti1}, & v_{T1} \\ & T_{Ti2}, & v_{T2} \\ & \vdots & \vdots \\ & T_{Tin}, & v_{Tn} \end{bmatrix} \quad (7)$$

结合物元模型,式(4)和式(6)分别给出了在线评论二级和三级层次的形式化表达方式,式(5)和式(7)给出了在线评论产品对象-对象特征对应关系的表达方式。基于物元模型的表示方法,不仅再现了产品对象相关的信息和在线评论者的用户反馈,而且对在线评论的层次进行了表达,使得描述同一层次的在线评论聚合到同一层次区域,相当于对在线评论完成了层次聚类,最大限度保证不同层次之间的用户观点的独立性。

2.4 基于 SO-PMI 算法的用户观点评价价值计算

2.4.1 句子评价价值计算

用户观点表达了对产品对象及其对象特征的评价。在一个完整的句子当中最能反映人类情感的词包括情感极性词、程度副词和否定词,情感倾向点互信息 (SO-PMI) 算法可以综合分析这三类词在句意表达上的情感倾向^[22],对每一类词都

可以采用相应的计算公式来量化句子的极性程度,在情感词典的基础上,计算每个在线评论层次中用户观点评价价值。

SO-PMI 算法是以点互信息 (point mutual Information, PMI) 为计算基础对用户观点进行量化分析的^[23],PMI 在用户观点量化分析中,主要用来计算两个词汇之间的关联性,其计算公式为

$$I_{PM}(w_1, w_2) = \log_2 \left[\frac{P(w_1, w_2)}{P(w_1) \cdot P(w_2)} \right] \quad (8)$$

其中: w_1, w_2 分别表示不同的两个词语, $P(w_1), P(w_2)$ 表示这两个词语在整个语料库中 w_1, w_2 出现的概率, $P(w_1, w_2)$ 表示 w_1, w_2 在语料库中共现的概率。

2.4.2 极性词强度的量化计算

根据已知的情感词语的极性强度,将其与已知的参照词语进行比较,通过公式量化计算,得出其未知词汇的极性强度。未知词 W 极性强度的量化计算公式为

$$s(W) = \frac{\sum_{w_+ \in W_+} I_{PM}(W, w_+)}{N_+} - \frac{\sum_{w_- \in W_-} I_{PM}(W, w_-)}{N_-} \quad (9)$$

式中: $s(W)$ 代表未知词 W 的极性强度值, N_+ 为所有褒义词的数量, N_- 为所有贬义词的数量, $I_{PM}(W, w_+)$ 为词 W 与褒义词列表中每一个词 w_+ 的 PMI 值, $I_{PM}(W, w_-)$ 为词 W 与贬义词列表中每一个词 w_- 的 PMI 值,其中褒义词列表和贬义词列表来自知网 HowNet 情感词典。

2.4.3 程度副词和否定词的量化计算

首先建立程度副词词典,将其划分成 4 种类别,各个类别的副词强度值 d 如表 1 所示。

表 1 各个层级的副词程度值
Table 1 Each level of adverb degree values

强度值	数量	举例
1.4 (极量词)	11	太、非常、特别……
1.2 (高量词)	10	很、更加、十分……
0.8 (中量词)	9	较为、挺、确实……
0.6 (低量词)	14	小、有些、有点……

在句子中出现否定词,则 $f=-0.5$, 否则 $f=1$ 。计算公式为

$$S = \frac{\sum_{i=1}^m s(W_i) \cdot d \cdot f}{m} \quad (10)$$

式中: m 为情感词总数, S 为一个句子中的评价价值。

2.4.4 用户观点的正负评价价值计算

在计算完所有在线评论层次对应的全部句子的情感倾向值后,然后加权求和得出同一层次下的综合评价价值,计算公式为

$$Q_+ = \frac{\sum_{i=1}^p S_+}{p}; Q_- = \frac{\sum_{i=1}^q S_-}{q} \quad (11)$$

式中: Q_+ 和 Q_- 分别为同一层次中的正面评价价值和负面评价价值, S_+ 和 S_- 分别表示为一个层次中大于零的评价值和小于零的评价值, p 和 q 分别为 S_+ 和 S_- 的总数。

2.5 基于相关网与蕴含系的用户需求识别

在线评论用户的需求是事件,因此可以采用事元模型来描述。事元可以形式化表达出用户做什么、谁做、为谁做等信息。本文对用户观点评级值较低的用户需求建立用户需求事元模型,采用事元模型 A_{ij} 表示,根据式 (3) 有^[21]:

$$A_{ij} = \begin{bmatrix} O_a, & c_{a1}, & v_{a1} \\ & c_{a2}, & v_{a2} \\ & \vdots & \vdots \\ & c_{an}, & v_{an} \end{bmatrix} \quad (12)$$

式中: O_a 为用户需求动作; $c_{a1}, c_{a2}, \dots, c_{an}$ 为动作的基本特征(支配对象、施动对象、接受对象等); $v_{a1}, v_{a2}, \dots, v_{an}$ 为量值。

产品设计信息具有多层次、多特征的特点,部分用户评论并未清晰地指明产品缺陷存在的部位,评论观点只停留在表面,而且对于产品各层次缺陷信息之间的关联性缺乏分析,无法找到对应的改进策略。在产品对象与产品特征当中,影响某个产品对象用户评价价值高低的因素有很多,既有同一层次产品对象之间的影响,也有不同层次产品对象之间的影响,这些因素并不是孤立存在的。基于此,引入可拓学拓展分析方法,相关网和蕴含系辅助分析产品用户需求识别问题。

2.5.1 相关网方法

产品对象之间、产品特征之间往往存在着千丝万缕的联系。在可拓学中,物元以形式化的方式表达世间万物,当难以直接或有效地对某个物元进行研究分析时,可以考虑利用物元对象之间、物元特征之间的相关关系进行分析。根据可拓学中已经定义的相关网规则(包括同对象异特征相关规则、异对象同特征相关规则和异对象异特征相关规则),这 3 种相关准则的判断依据主要是基于常识或领域知识。每个物元可能与其他物

元之间存在相关关系,用模型表示,会形成一个网状结构,这种方法称之为相关网方法^[24]。

根据相关准则及领域知识对给定以下存在缺陷的物元确定相关关系,对于物元 $M_1 = (O_1, c_1, v_1)$, 若存在同对象异特征物元 $M_2 = (O_1, c'_1, v'_1)$ 或异对象同特征物元 $M_3 = (O'_1, c_1, v'_1)$ 或异对象异特征物元 $M_4 = (O'_1, c'_1, v'_1)$, 使 $(c_1)M_1 \sim (c'_1)M_2$, $(c_1)M_1 \sim (c_1)M_3$, $(c_1)M_1 \sim (c'_1)M_4$ 。其中符号“ \sim ”表示两个缺陷物元之间相关,最终建立缺陷物元 M_1 的相关网如图 4 所示。其中,物元 M_1 的特征 c_1 与物元 M_2 的特征 c'_1 相关,物元 M_1 的对象 O_1 与物元 M_3 的对象 O'_1 相关,物元 M_1 的对象 O_1 和特征 c_1 与物元 M_4 的对象 O'_1 和特征 c'_1 相关。

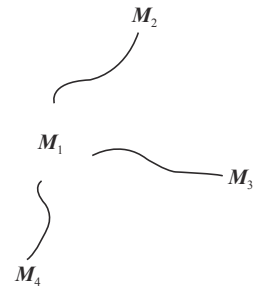


图 4 缺陷物元之间的相关网

Fig. 4 Correlation network between defective matter-elements

基于前述在线评论层次的划分,对每个层次不同层级中存在缺陷的物元采用相关网方法进行分析。为了达到分析效果,只对功能层与实体层、属性层与实体层和实体层之间进行相关网分析,并对一些较复杂的关系进行简化,其基本分析步骤如下:

1) 列出存在缺陷的物元。

2) 利用相关准则列出物元的相关网,根据层次关系,分为 3 类情况:功能层和实体层之间,属性层和实体层之间的产品对象及特征必然互不相同,故采用异对象异特征相关准则进行分析;实体层不同层级之间具有不同对象相同特征的物元,故采用异对象同特征相关准则进行分析;为了探索不同产品特征对用户需求的影响,避免引入相同对象之间的相关性对后续分析结果的影响,对实体层相同层级之间不再进行分析。

通过相关网方法来分析缺陷物元之间的关联关系,为识别用户需求的改进途径提供指引。

2.5.2 蕴含系方法

蕴含系方法是以基元模型对物、事和关系进行形式化的分析方法。设有事元 A 、 A_1 和 A_2 :

若 A_2 实现必有 A_1 实现,称事元 A_2 蕴含 A_1 ,记作 $A_2 \Rightarrow A_1$;若 A_1 与 A_2 同时实现必有 A 实现,则 A_1 、 A_2 与蕴含 A ,记作 $A_2 \wedge A_1 \Rightarrow A$,推广到更一般的情况:

$$\bigwedge_{i=1}^n A_i \Rightarrow A \quad (13)$$

根据蕴含系方法,利用产品设计专业知识将难于实现的上位事元转化为它的易于实现的下位事元,从而保证最终目标事元的实现。在本文中对用户观点评价价值较低的产品对象建立用户需求事元模型并提取出相应的缺陷物元,用事元来表示用户需求。通过利用相关网方法分析出哪些缺陷物元之间存在着相关关系,并根据蕴含系方法,利用确定出相关关系的缺陷物元来指导用户需求事元蕴含系的构建,其蕴含关系建立的基本步骤如下:

1) 列出用户需求事元 A_{ij} 。

2) 结合缺陷物元之间的相关网,确定用户需求事元之间的蕴含关系。若 $(c_1)M_1 \sim (c'_1)M_2$, A_{11} 和 A_{21} 分别为相应的用户需求事元,且 M_1 所在层级低于 M_2 所在层级,则有 A_{21} 蕴含 A_{11} 。

3) 将功能层和属性层较难实现的事元划分为上位基元,将实体层中的子系统所表示的事元划分为中位基元,将实体层中的特征结构所表示的事元划分为下位基元,完成蕴含关系的构建。

通过上述步骤,用户需求事元蕴含系统一般形式如图5所示。

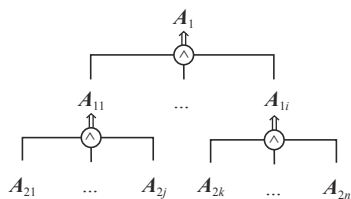


图5 用户需求事元蕴含系统一般形式

Fig. 5 User requirement affair-elements implication system in general form

事元之间的蕴含关系是比较复杂的,往往是多层级的^[25]。结合物元的相关网,利用蕴含系对功能层、属性层的事元模型,蕴含出实体层的事元模型,形式化地分析问题功能、问题属性与实体层之间的关系,形成自下而上的用户需要实现的步骤和方法,找到满足用户需求的有效途径。

3 实例研究

近年来,养生壶深受广大用户青睐,逐渐成为各大电商平台及家电行业的热销产品,并且用户的在线评论信息也较为丰富。因此,通过以该产

品作为案例,对于挖掘在线评论中的改进需求信息,提升用户对产品体验度研究具有代表性。本文选取京东购物平台中的一款养生壶作为分析对象,养生壶实物结构如图6所示(图片来源于京东小熊厨房电器官方旗舰店)。



图6 养生壶实物结构图

Fig. 6 Physical structure diagram of health pot

3.1 抓取在线评论数据集及文本预处理

首先,采用Python爬虫工具抓取养生壶的评论数据集,总共统计原始的在线评论10160条。之后,对抓取的养生壶评论文本进行文本预处理,处理的内容或步骤主要包括以下内容:1)删除系统默认好评,如“系统默认好评”;2)删除评论字数较短或无实际研究价值的文本内容,如“不错”“喜欢”“赞”等;3)重复出现的评论;4)删除非汉字的内容,如特殊字符、表情、数字和字母等。文本处理后剩余的评论合计5319条。将处理之后的原始评论采用Python汉语词汇分割工具——jieba库进行分词,并导入停用词表删除限用词。

3.2 产品在线评论层次模型构建

3.2.1 产品要素提取

对在线评论进行数据集预处理后,应用PYLDAvis交互式可视化库提取产品要素,调节权重参数 λ ,并设置主题个数为10,最终效果图如图7所示。可视化效果图有左右两个面板,其中左边面板与右边面板是有关联的。左边面板展示了整个评论的主题全局视图,每个气泡图的大小与话题在语料库中的相对关联度成正比,在左边面板中选择一个用气泡图形表示的一个主题,就会在右边同步显示这个主题所对应的前30个特征词。右边面板中,蓝色条形图的宽度表示每个词在整个语料库中出现的频率,红色条形图的宽度代表每个词在该主题下所占的权重。参数 $\lambda(0 \leq \lambda \leq 1)$ 反应了词汇和主题之间的关联关系,其中,当 λ 越接近于1,表示该主题下更频繁出现的词与主题

更相关,当 λ 越接近于 0,表示该主题下较少出现、独有的词与主题更相关^[15]。通过调节 λ 的值,最

终提取出产品要素,并总结获得如图 8 所示的产品要素与各个产品层次区域的映射关系。

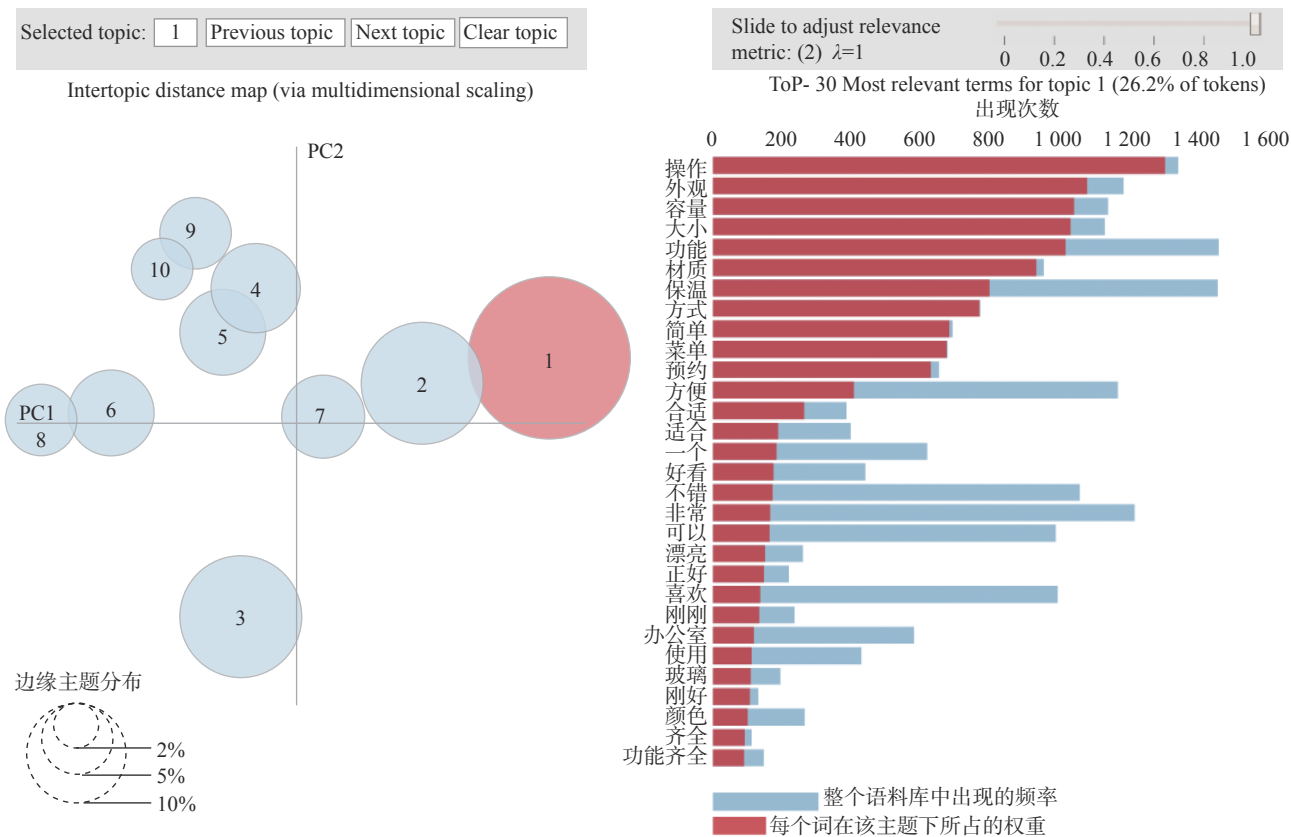


图 7 养生壶评论主题分布交互式可视化效果

Fig. 7 Interactive visualization of health pot comment topic distribution

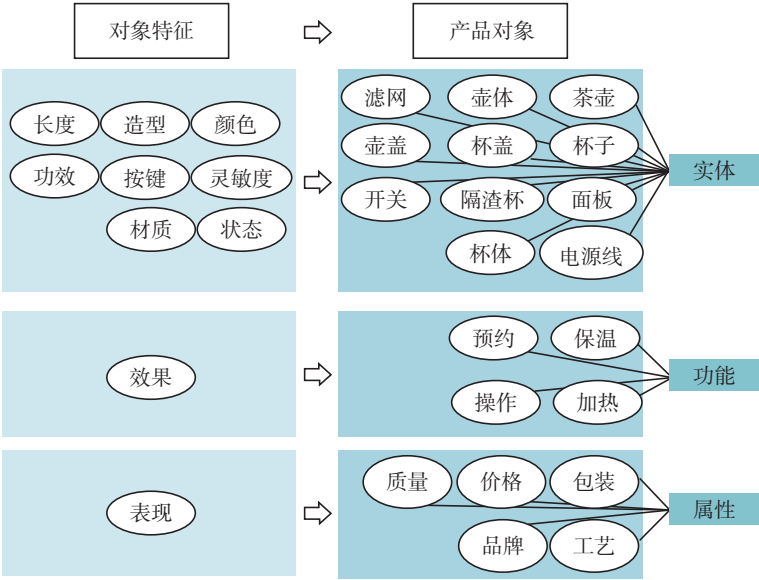


图 8 产品要素与产品层次映射关系

Fig. 8 Mapping relationship between product elements and product hierarchy

3.2.2 在线产品层次模型表达

在线评论二级层次根据式 (4) 表达如下,以实体层和功能层为例:

$$M_{H_{S1}} = \begin{bmatrix} H_{S1}, & \text{茶壶 } O_{S1}, & \{\text{颜色, 材质, 造型}\} \\ & \text{杯子 } O_{S2}, & \{\text{颜色, 材质, 造型}\} \\ & \text{面板 } O_{S3}, & \{\text{造型, 材质, 灵敏度}\} \\ & \text{电源线 } O_{S4}, & \{\text{长度, 材质}\} \end{bmatrix}$$

$$M_{H_{S2}} = \begin{bmatrix} H_{S2}, & \text{保温}O_{S5}, & \left\{ \begin{matrix} \text{效果} \end{matrix} \right\} \\ & \text{预热}O_{S6}, & \left\{ \begin{matrix} \text{效果} \end{matrix} \right\} \\ & \text{加热}O_{S7}, & \left\{ \begin{matrix} \text{效果} \end{matrix} \right\} \\ & \text{操作}O_{S8}, & \left\{ \begin{matrix} \text{效果} \end{matrix} \right\} \end{bmatrix}$$

二级层次中的产品对象根据式(5)表达如下,以茶壶 O_{S1} 为例:

$$M_{O_{S1}} = \begin{bmatrix} \text{茶壶}O_{S1}, & \text{颜色}S_{S11}, & v_{S1} \\ & \text{材质}S_{S12}, & v_{S2} \\ & \text{造型}S_{S13}, & v_{S3} \end{bmatrix}$$

在线评论三级层次根据式(6)表达如下,以茶壶的特征结构为例:

$$M_{H_{T1}} = \begin{bmatrix} H_{T1}, & \text{壶盖}O_{T1}, & \left\{ \begin{matrix} \text{颜色, 造型, 材质} \end{matrix} \right\} \\ & \text{隔渣杯}O_{T2}, & \left\{ \begin{matrix} \text{材质, 功效} \end{matrix} \right\} \\ & \text{滤网}O_{T3}, & \left\{ \begin{matrix} \text{材质, 功效} \end{matrix} \right\} \\ & \text{壶体}O_{T4}, & \left\{ \begin{matrix} \text{颜色, 造型, 材质} \end{matrix} \right\} \end{bmatrix}$$

三级层次中的产品对象根据式(7)表达如下,以壶盖 O_{T1} 为例:

$$M_{O_{T1}} = \begin{bmatrix} \text{壶盖}O_{T1}, & \text{颜色}T_{T11}, & v_{T1} \\ & \text{材质}T_{T12}, & v_{T2} \\ & \text{造型}T_{T13}, & v_{T3} \end{bmatrix}$$

3.3 用户观点评价值计算

3.3.1 产品对象词典和对象特征词典构建

在线评论对产品对象和对象特征的描述往往较口语化,因此在计算用户观点评价值之前,需要对产品对象和对象特征整理、归纳,最终构建产品对象词典和属性词典,见表2和表3。

表2 产品对象词典(部分展示)

Table 2 Product object dictionary (partial display)

产品对象	举例	数量
茶壶	水壶、养生壶、壶、玻璃壶、热水壶、茶壶	6
杯子	水杯、小杯、杯、玻璃杯、茶杯、杯子	6
面板	触摸屏、菜单、按键、触摸板、面板、桌面	7
电源线	电源、电线、充电线、电源线、数据线	5
...

表3 对象特征词典(部分展示)

Table 3 Object feature dictionary (partial display)

对象特征	举例	数量
颜色	色彩、粉色、颜值、色泽、色差...	7
材质	材料、玻璃、铝合金、塑料、选料...	8
造型	外观、外形、样貌、表面、形状...	12
灵敏度	灵活、灵敏、迟钝、反应、反应度...	7
...

3.3.2 各个层次用户观点评价值计算

通过上述分析,将产品对象及对应的对象特征采用物元模型进行了分层次表示。用户观点评价值的计算准则是以句子为单位进行的,因此在计算各个层次的评价值需要将所有有效评论按照“。?! ”三个符号进行断句。

在上述准则下,利用SO-PMI算法,计算二级层次中实体层的用户观点评价值及二级层次中功能层和属性层的用户观点评价值,计算结果分布如表4、表5所示。

表4 二级层次中实体层的用户观点评价值

Table 4 User opinion evaluation value of the entity layer in the second level

实体对象	特征	评价值	
		正面(Q_+)	负面(Q_-)
茶壶	颜色	0.086	-0.002
	材质	0.034	-0.045
	造型	0.015	-0.032
杯子	颜色	0.035	0
	材质	0.012	-0.057
	造型	0.054	-0.003
面板	材质	0.020	0
	造型	0.032	0
	灵敏度	0.002	-0.032
电源线	材质	0	0
	长度	0.003	-0.026

表5 二级层次中功能层和属性层的用户观点评价值

Table 5 User opinion evaluation value of function layer and attribute layer in the second level

功能/属性对象	效果/表现	评价值	
		正面(Q_+)	负面(Q_-)
保温	效果	0.051	-0.004
预约	效果	0.052	-0.006
加热	效果	0.086	-0.001
操作	效果	0.005	-0.053
包装	表现	0.089	-0.005
价格	表现	0.021	-0.010
质量	表现	0.012	-0.056
工艺	表现	0.008	-0.009
品牌	表现	0.023	-0.003

根据表4计算结果,可以看出用户关注的实体层中二级层次的部分实体对象的正面特征评价值较高,有少部分实体对象的负面特征评价值较低(负面特征评价值越低则代表用户对该对象的特征越不满意)。根据计算结果,将产品对象的正面特征评价值和负面特征评价值分别划分为3个区间: $Q_+ > 0.02$ 且 $0 \geq Q_- \geq -0.01$ (表示很满意)、 $0.02 \geq Q_+ \geq 0$ 且 $0 \geq Q_- \geq -0.01$ (表示满意)、 $0.02 \geq$

$Q_+ \geq 0$ 且 $-0.01 \geq Q_-$ (表示不满意)。按照上述划分规则,可以看出对于茶壶的颜色和材质、杯子的颜色和造型、面板的材质和造型是用户很满意的特征;而茶壶的造型、杯子的材质、面板的灵敏度和电源线的长度是用户不满意的特征;特别地对于电源线的材质的评价值均为 0,是由于用户并未对该对象的特征做出评价。根据表 5,从功能对象和属性对象的评价结果显示,可以看出除了操作功能和质量属性是用户不满意的,其他的都是用户很满意的对象。

三级层次各个实体对象的用户观点评价值计算结果如表 6 所示。

表 6 三级层次的用户观点评价值
Table 6 User opinion evaluation value in the three level

实体对象	特征	评价值	
		正面(Q_+)	负面(Q_-)
壶盖	颜色	0	0
	材质	0.022	-0.001
	造型	0.012	-0.027
隔渣杯	材质	0.010	-0.002
	功效	0.009	-0.019
滤网	材质	0.005	-0.012
	功效	0.002	-0.010
壶体	颜色	0.049	-0.001
	材质	0.023	-0.036
	造型	0.043	-0.004
杯盖	颜色	0	0
	材质	0.023	-0.010
	造型	0	0
杯体	颜色	0.034	-0.004
	材质	0.007	-0.036
	造型	0.043	-0.006
开关	状态	—	—

根据表 6 计算结果,壶盖的造型、隔渣杯的功效、杯体的材质是用户不满意的特征。其中,出现部分特征正面评价值和负面评价值均为 0 的情况,同样是因为用户并没有对该对象的特征做出评价。查看实际评论数据发现,该产品本身没有开关这个实体对象,所以将不存在特征的评价值定义为不存在。

3.4 结果分析及改进需求识别

通过计算各个层次中产品对象对应的对象特征的用户观点评价值,确定出需要改进的产品对象及对象特征。针对用户观点评价值较低的产品需求,根据式 (12) 建立用户需求事元模型:

$$\begin{aligned}
 A_{11} &= \left[\begin{array}{l} \text{改善,} \\ \text{支配对象,} \\ \text{接受对象,} \end{array} \begin{array}{l} \text{造型 } S_{S13} \\ \text{茶壶 } O_{S1} \end{array} \right] \\
 A_{12} &= \left[\begin{array}{l} \text{改善,} \\ \text{支配对象,} \\ \text{接受对象,} \end{array} \begin{array}{l} \text{材质 } S_{S22} \\ \text{杯子 } O_{S2} \end{array} \right] \\
 A_{13} &= \left[\begin{array}{l} \text{提高,} \\ \text{支配对象,} \\ \text{接受对象,} \end{array} \begin{array}{l} \text{灵敏度 } S_{S33} \\ \text{面板 } O_{S3} \end{array} \right] \\
 A_{14} &= \left[\begin{array}{l} \text{延长,} \\ \text{支配对象,} \\ \text{接受对象,} \end{array} \begin{array}{l} \text{长度 } S_{S41} \\ \text{电源线 } O_{S4} \end{array} \right] \\
 A_{15} &= \left[\begin{array}{l} \text{改善,} \\ \text{支配对象,} \\ \text{接受对象,} \end{array} \begin{array}{l} \text{材质 } S_{S12} \\ \text{茶壶 } O_{S1} \end{array} \right] \\
 A_{21} &= \left[\begin{array}{l} \text{改善,} \\ \text{支配对象,} \\ \text{接受对象,} \end{array} \begin{array}{l} \text{造型 } T_{T13} \\ \text{壶盖 } O_{T1} \end{array} \right] \\
 A_{22} &= \left[\begin{array}{l} \text{改善,} \\ \text{支配对象,} \\ \text{接受对象,} \end{array} \begin{array}{l} \text{材质 } T_{T42} \\ \text{壶体 } O_{T4} \end{array} \right] \\
 A_{23} &= \left[\begin{array}{l} \text{改善,} \\ \text{支配对象,} \\ \text{接受对象,} \end{array} \begin{array}{l} \text{材质 } T_{T62} \\ \text{杯体 } O_{T6} \end{array} \right] \\
 A_{24} &= \left[\begin{array}{l} \text{增加,} \\ \text{支配对象,} \\ \text{接受对象,} \end{array} \begin{array}{l} \text{开关 } O_{T7} \\ \text{面板 } O_{S3} \end{array} \right] \\
 A_{16} &= \left[\begin{array}{l} \text{提高,} \\ \text{支配对象,} \\ \text{接受对象,} \end{array} \begin{array}{l} \text{操作 } O_{S5} \end{array} \right] \\
 A_{17} &= \left[\begin{array}{l} \text{提高,} \\ \text{支配对象,} \\ \text{接受对象,} \end{array} \begin{array}{l} \text{质量 } O_{S6} \end{array} \right]
 \end{aligned}$$

对于功能层和属性层的产品对象缺陷的用户需求目标并不容易实现,但对于实体层中的产品对象缺陷可以直接对产品对象或其对应的对象特征进行改进,所以,可以通过找到实体层对功能层、属性层的影响关系,来实现功能层和属性层的用户需求目标。基于物元模型表达的在线评论层次及用户观点评价值的计算结果,筛选出存在缺陷的产品对象及对应对象特征的物元模型如下:

$$\begin{aligned}
 M_{O_{S1}} &= \left[\begin{array}{l} \text{茶壶 } O_{S1}, \text{ 材质 } S_{S12}, \text{ } v_{S2} \end{array} \right] \\
 M_{O_{S1}} &= \left[\begin{array}{l} \text{茶壶 } O_{S1}, \text{ 造型 } S_{S13}, \text{ } v_{S3} \end{array} \right] \\
 M_{O_{S2}} &= \left[\begin{array}{l} \text{杯子 } O_{S2}, \text{ 材质 } S_{S22}, \text{ } v_{S5} \end{array} \right] \\
 M_{O_{S3}} &= \left[\begin{array}{l} \text{面板 } O_{S3}, \text{ 灵敏度 } S_{S33}, \text{ } v_{S9} \end{array} \right] \\
 M_{O_{S4}} &= \left[\begin{array}{l} \text{电源线 } O_{S4}, \text{ 长度 } S_{S41}, \text{ } v_{S10} \end{array} \right] \\
 M_{O_{T1}} &= \left[\begin{array}{l} \text{壶盖 } O_{T1}, \text{ 造型 } T_{T13}, \text{ } v_{T3} \end{array} \right] \\
 M_{O_{T4}} &= \left[\begin{array}{l} \text{壶体 } O_{T4}, \text{ 材质 } T_{T42}, \text{ } v_{T9} \end{array} \right] \\
 M_{O_{T6}} &= \left[\begin{array}{l} \text{杯体 } O_{T6}, \text{ 材质 } T_{T62}, \text{ } v_{T12} \end{array} \right] \\
 M_{O_{T7}} &= \left[\begin{array}{l} \text{开关 } O_{T7}, \text{ 状态 } T_{T71}, \text{ } v_{T13} \end{array} \right] \\
 M_{O_{S5}} &= \left[\begin{array}{l} \text{操作 } O_{S5}, \text{ 效果 } S_{S51}, \text{ 差} \end{array} \right] \\
 M_{O_{S6}} &= \left[\begin{array}{l} \text{质量 } O_{S6}, \text{ 表现 } S_{S61}, \text{ 差} \end{array} \right]
 \end{aligned}$$

根据该产品相关设计信息与实际在线评论数据,借助 3 种相关网准则分析得到各个存在缺陷对象的物元之间的相关关系表示如下。以功能层和实体层之间为例,根据实际评论数据以及异对

象异特征相关准则,分析出茶壶的造型与产品的操作功能相关,即 $(S_{S13})M_{O_{S1}} \sim (S_{S51})M_{O_{S5}}$,面板的灵敏度与产品的操作功能相关,即 $(S_{S33})M_{O_{S3}} \sim (S_{S51})M_{O_{S5}}$,电源线的长度与产品操作功能相关,即 $(S_{S41})M_{O_{S4}} \sim (S_{S51})M_{O_{S5}}$,面板的开关状态与产品操作功能相关,即 $(T_{T71})M_{O_{T7}} \sim (S_{S51})M_{O_{S5}}$ 。对实体层不同层次采用异对象同特征相关准则,分析出壶盖的造型与茶壶的造型相关,即 $(T_{T13})M_{O_{T1}} \sim (S_{S13})M_{O_{S1}}$ 。而对于实体层同一层次的缺陷物元,如 $(S_{S12})M_{O_{S1}}$ 和 $(S_{S13})M_{O_{S1}}$ 之间建立相关网,就会导致在探索实体层分别与功能、属性之间的关系时引入产品对象这一无关因素,因而针对这一类不再进行相关网分析。

根据上述分析,同理可进行实体层和属性层的分析,最终可建立如图9所示的缺陷物元模型相关网。

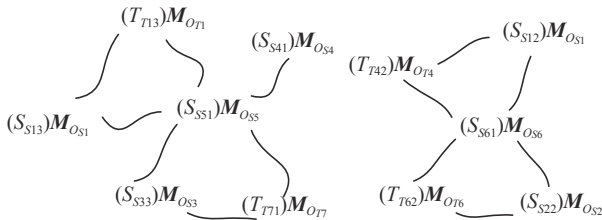


图9 缺陷物元模型相关网

Fig. 9 Correlation network of defective matter-elements model

在图9中,根据相关网确定物元模型之间的关联关系,虽然很难对产品操作功能 $M_{O_{S5}}$ 和产品质量属性 $M_{O_{S6}}$ 直接进行改进,但这两个物元与实体层的其他物元均相关。结合缺陷物元模型之间的相关网分析,如 $(T_{T13})M_{O_{T1}} \sim (S_{S13})M_{O_{S1}}$,可以从 A_{21} 蕴含出 A_{11} ,如 $(S_{S13})M_{O_{S1}} \sim (S_{S51})M_{O_{S5}}$,可以从 A_{11} 蕴含出 A_{16} ,其余同理可得。将功能层和属性层的用户需求事元 A_{16} 和 A_{17} 作为较难实现的上位基元,实体层中子系统的用户需求事元 $A_{11}, A_{12}, \dots, A_{15}$ 作为中位基元,实体层中特征结构的用户需求事元 $A_{21}, A_{22}, \dots, A_{24}$ 为下位基元,通过实现用户对特征结构的需求从而实现用户对子系统的需求,通过实现用户对所有子系统的需求从而实现用户对功能层和属性层的需求。最终得到该产品的两个用户需求事元实现层次蕴含系统如图10。

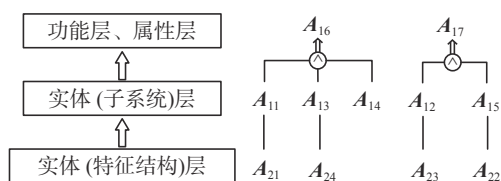


图10 用户需求事元实现层次蕴含系统

Fig. 10 Implementation hierarchical implicit system of user requirement affair-elements

在图10中, A_{16} 和 A_{17} 表示用户对操作功能和属性层的需求; $A_{11}, A_{12}, \dots, A_{15}$ 表示用户对实体层中子系统的用户需求; $A_{21}, A_{22}, \dots, A_{24}$ 表示用户对实体层中特征结构的用户需求。通过蕴含系分析可以得出:当功能层和属性层的用户需求目标很难实现时,可以寻找与之关联的实体层中的子系统。例如:改进操作功能(A_{16})可以通过改善茶壶的造型(A_{11})、提高面板的灵敏度(A_{13})和延长电源线的长度(A_{14})来实现,提高产品(A_{17})质量可以通过改善茶壶的材质(A_{12})和提高杯子的材质(A_{15})来实现。当子系统内的用户需求目标很难实现时,可以寻找与之关联的实体层中的特征结构。例如:改善茶壶造型(A_{11})可以通过改善壶盖的造型(A_{21})来实现,提高面板的灵敏度(A_{13})可以通过增加开关按键(A_{24})来实现。上述分析结果与用户评论中反映出来的主观信息也是一致的,从而说明了本文方法的可行性。

以相关网和蕴含系相结合识别产品改进的方式,深度剖析了存在缺陷的产品对象之间的层次需求关系,指明了产品实体存在缺陷的部位,进而找到产品功能和属性的改进途径。这也有利于设计人员形式化、条理化地处理多层次、多特征的产品改进设计信息,从而更加具有针对性地满足产品改进需求。

4 结束语

1)本文基于在线评论的基本构成信息,考虑多层次、多特征的产品,根据产品的特点,将产品划分为实体层、功能层和属性层,借助数据挖掘方法,从在线评论当中获取产品要素,再应用可拓学物元模型,建立在线评论层次模型并对其进行表达。

2)根据在线评论层次模型,归纳、总结出各个层次中的产品对象词典和产品特征词典,应用SO-PMI算法计算出各个产品对象不同的对象特征评价价值,找到评价价值较低的产品对象所对应的特征。

3)以用事元模型为工具对评价价值较低的产品对象所对应的特征进行用户需求表达,通过相关网方法判断其缺陷物元模型之间的关联关系、蕴含系方法构建用户需求实现的层次蕴含系统,为利用产品的在线评论进行改进提供指导意见。将该流程应用于一款养生壶的用户需求识别中,形式化、条理化地理清该产品改进设计信息,找到了用户对不同层次需求实现的步骤。

参考文献:

- [1] LAI Xinjun, ZHANG Qixiang, CHEN Qingxin, et al. The analytics of product-design requirements using dynamic Internet data: application to Chinese smartphone market[J]. *International journal of production research*, 2019, 57(18): 5660–5684.
- [2] 冯立杰, 关柯楠, 王金凤. 基于在线评论考虑潜在用户需求的产品创新方案识别研究 [J]. *情报理论与实践*, 2022, 45(6): 129–137.
FENG Lijie, GUAN Kenan, WANG Jinfeng. Research on product innovation scheme identification considering potential user needs based on online comments[J]. *Information studies: theory & application*, 2022, 45(6): 129–137.
- [3] 王慧. 基于特征提取和情感分析的用户需求挖掘研究 [D]. 杭州: 浙江理工大学, 2019: 5–17.
WANG Hui. Research on user demand mining based on feature extraction and emotion analysis[D]. Hangzhou: Zhejiang Sci-Tech University, 2019: 5–17.
- [4] 徐凯. 基于产品特征的用户评论情感倾向分析研究 [D]. 合肥: 合肥工业大学, 2015: 8–11.
XU Kai. Research on sentiment orientation analysis of user reviews based on product features[D]. Hefei: Hefei University of Technology, 2015: 8–11.
- [5] 李贺, 曹阳, 沈旺, 等. 基于 LDA 主题识别与 Kano 模型分析的用户需求研究 [J]. *情报科学*, 2021, 39(8): 3–11, 36.
LI He, CAO Yang, SHEN Wang, et al. User demand based on LDA subject identification and Kano model analysis[J]. *Information science*, 2021, 39(8): 3–11, 36.
- [6] DAY M Y, TENG H C. A study of deep learning to sentiment analysis on word of mouth of smart bracelet[C]// *Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*. New York: ACM, 2017: 763–770.
- [7] YE H C T, CHEN Muchen. Applying Kansei Engineering and data mining to design door-to-door delivery service[J]. *Computers & industrial engineering*, 2018, 120: 401–417.
- [8] 李江泳, 张伟. 在线评论数据挖掘视角下的产品改良设计研究 [J]. *包装工程*, 2021, 42(6): 135–141.
LI Jiangyong, ZHANG Wei. Product improvement design from the perspective of online comment data mining[J]. *Packaging engineering*, 2021, 42(6): 135–141.
- [9] 萧箏. 客户需求信息处理理论和方法研究 [D]. 武汉: 武汉理工大学, 2013: 5–20.
XIAO Zheng. Research on theories and methods of customer requirement information processing[D]. Wuhan: Wuhan University of Technology, 2013: 5–20.
- [10] 赵云彦, 韩冬楠, 尤立思, 等. 基于用户需求层次分析的女性安全防护产品设计研究 [J]. *包装工程*, 2021, 42(16): 120–126.
ZHAO Yunyan, HAN Dongnan, YOU Lisi, et al. The design of women's safety protection products based on the level of users' demand[J]. *Packaging engineering*, 2021, 42(16): 120–126.
- [11] 纪雪, 高琦, 李先飞, 等. 考虑产品属性层次性的评论挖掘及需求获取方法 [J]. *计算机集成制造系统*, 2020, 26(3): 747–759.
JI Xue, GAO Qi, LI Xianfei, et al. Reviews mining and requirements elicitation methodology considering product attributes' hierarchy[J]. *Computer integrated manufacturing systems*, 2020, 26(3): 747–759.
- [12] 肖寒琼, 张馨遇, 肖宇晗, 等. 基于方面词的用户消费心理画像方法 [J]. *数据分析与知识发现*, 2022, 6(6): 22–31.
XIAO Hanqiong, ZHANG Xinyu, XIAO Yuhan, et al. Creating consumer psychology portrait with aspect words[J]. *Data analysis and knowledge discovery*, 2022, 6(6): 22–31.
- [13] 张文旭, 肖人彬, 林文广. 基于产品性能词典的评论数据驱动客户需求模型研究 [J]. *中国机械工程*, 2020, 31(15): 1866–1876.
ZHANG Wenxu, XIAO Renbin, LIN Wenguang. Review data driven customer need model research based on product performance lexicon[J]. *China mechanical engineering*, 2020, 31(15): 1866–1876.
- [14] 刘晓敏, 李娇蓉, 陈亮, 等. 融合可拓学与多层次知识建模的仿生原型获取 [J]. *机械工程学报*, 2019, 55(21): 150–160.
LIU Xiaomin, LI Jiaorong, CHEN Liang, et al. Bionic prototype acquisition incorporating extension and multi-level knowledge modeling[J]. *Journal of mechanical engineering*, 2019, 55(21): 150–160.
- [15] BLEI D, NG A, JORDAN M. Latent dirichlet allocation[J]. *Journal of machine learning research*, 2003(3): 933–1022.
- [16] 杨凡. 基于 LDA 主题模型的在线评论聚类分析与推荐 [D]. 大连: 大连理工大学, 2018.
YANG Fan. Online reviews clustering analysis and recommendation based on LDA topic model[D]. Dalian: Dalian University of Technology, 2018.
- [17] SIEVERT C, SHIRLEY K. LDAvis: a method for visualizing and interpreting topics[C]// *Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces*. Stroudsburg: Association for Computational Linguistics, 2014: 63–70.
- [18] 陈晓菁, 成思源, 杨雪荣. 面向专利产品创新设计的可拓变换方法研究 [J]. *包装工程*, 2020, 41(24): 137–142.
CHEN Xiaojing, CHENG Siyuan, YANG Xuerong. Ex-

- tension transform method for patent product innovation design[J]. *Packaging engineering*, 2020, 41(24): 137–142.
- [19] GRIFFIN A, HAUSER J R. The voice of the customer[J]. *Marketing science*, 1993, 12(1): 1–27.
- [20] 高新勤, 金雨昊, 王雪萍, 等. 基于在线评论挖掘的产品感性评价方法研究 [J]. *现代制造工程*, 2021(12): 13–20. GAO Xinqin, JIN Yuhao, WANG Xueping, et al. Research on product perceptual evaluation method based on online review mining[J]. *Modern manufacturing engineering*, 2021(12): 13–20.
- [21] 杨春燕. 可拓创新方法 [M]. 北京: 科学出版社, 2017.
- [22] TURNEY P D, LITTMAN M L. Measuring praise and criticism[J]. *ACM transactions on information systems*, 2003, 21(4): 315–346.
- [23] 彭德中, 章毅, 吕建成, 等. 一种基于 SO-PMI 商品评价信息的情感分析方法: CN105022805B[P]. 2018-05-04. PENG Dezhong, ZHANG Yi, LYU Jiancheng, et al. Emotional analysis method based on SO-PMI (semantic orientation-pointwise mutual information) commodity evaluation information: CN105022805B[P]. 2018-05-04.
- [24] 李文军, 杨春燕, 汤龙, 等. 可拓学中相关关系的变换方法研究 [J]. *智能系统学报*, 2019, 14(4): 619–626. LI Wenjun, YANG Chunyan, TANG Long, et al. Research on the transformation method for the correlation relation in extenics[J]. *CAAI transactions on intelligent systems*, 2019, 14(4): 619–626.
- [25] 花黄伟, 杨春燕. 基于因果分析的事元蕴含系及其应用 [J]. *智能系统学报*, 2017, 12(1): 60–66.

HUA Huangwei, YANG Chunyan. A ffair-elements implication system andits application based on causal analysis[J]. *CAAI transactions on intelligent systems*, 2017, 12(1): 60–66.

作者简介:



于志刚, 硕士研究生, 主要研究方向为可拓创新方法、数据挖掘。



成思源, 教授, 博士, 主要研究方向为技术创新方法、专利挖掘、逆向工程技术。主持并完成国家自然科学基金项目 1 项、广东省自然科学基金项目 2 项、广东省科技计划项目 4 项。获国家级教学成果二等奖 1 项、广东省教学成果一等奖 3 项, 获得发明专利授权 20 余项。出版著作 8 部, 发表学术论文 70 余篇。



杨雪荣, 副教授, 博士, 主要研究方向为机器视觉、逆向工程技术、TRIZ 理论及应用。主持并完成国家自然科学基金项目 1 项、广东省自然科学基金项目 1 项、广东省科技计划项目 2 项。获广东省教学成果一等奖 1 项, 获发明专利授权 10 余项。出版著作 5 部, 发表学术论文 40 余篇。