

# 基于 Web 的多渠道用户需求知识 获取框架研究

何忠秀<sup>1</sup>, 王 霜<sup>2</sup>, 杜亚军<sup>1</sup>

(1. 西华大学 数学与计算机学院, 四川 成都 610039;

2. 西华大学 机械工程与自动化学院, 四川 成都 610039)

**摘 要:**用户需求知识是企业进行设计产品的依据、源头和关键。用户在产品网站内浏览、查询及提出意见时,会有意无意地留下自己对某个产品的倾向和偏好。鉴于此,文中提出了基于 Web 的多渠道用户需求知识获取框架,从用户浏览方式、用户对产品的查询、站内 BBS 和用户投诉 4 种渠道中获取需求信息构建用户关注度矩阵和用户满意度矩阵,给出需求指标综合计算公式;并将从多渠道获取的用户关注度和用户满意度进行融合得到用户需求矩阵,该矩阵对新产品开发具有指导意义。

**关键词:**用户需求; Web; 多渠道; 知识获取

中图分类号: TP391

文献标识码: A

文章编号: 1673-629X(2010)04-0124-04

## Research on Multi-channel's Knowledge Acquisition Frame for Customer Requirements Based on Web

HE Zhong-xiu<sup>1</sup>, WANG Shuang<sup>2</sup>, DU Ya-jun<sup>1</sup>

(1. Sch. of Mathematics & Computer Eng., Xihua Univ., Chengdu 610039, China;

2. Sch. of Mechanical Eng. & Automation, Xihua Univ., Chengdu 610039, China)

**Abstract:** Customer requirements knowledge is most important for enterprise to product design. When a customer browses some Web pages, inquiries and puts forward his opinion, his tendency and fancy of a certain product may leave. In view of this, carries out a multi-channel's knowledge acquisition frame for customer requirements based on Web, and lists four channels to obtain customer requirements in order to set up customer attention degree matrix(CA) and customer satisfaction degree matrix(CS). The four kinds of channels are customer browsing way on webs, product inquiries of customer, BBS in website and customer's complaints. An algorithm is followed to synthesize the degrees of customer attention and customer satisfaction in order to get customer requirement degree matrix(CR), which can guide significantly in new product design process.

**Key words:** customer requirement; Web; multi-channel; knowledge acquisition

## 0 引言

用户需求的提取可能是产品设计过程中最为关键也是最为困难的一步,它直接关系到产品的成功与否。要设计出最受用户欢迎的产品,需要全面、及时地获取用户需求。

Chen C H 等人<sup>[1]</sup>研究了用自适应谐振神经网络(Adaptive Resonance Theory2, ART2)获取用户需求的方法。Hsin-Hung Wu 等人<sup>[2]</sup>则是在质量功能展开(Quality Function Deployment, QFD)中用 Markov 链模型分析用户需求。文献[3]在研究产品质量指标体系的基础上,分析用户对产品的实际需求,将需求分为时间、质量、费用、服务与社会性五个方面,共 20 个小项的内容。王美清等人<sup>[4]</sup>研究了产品设计中的用户需求与产品质量特征映射方法,将获取的用户需求用网络分析法(The Analytic Network Process, ANP)确定用户需求间的重要度,进而建立产品质量特征集。这些研究集中在对用户需求的分析与处理上,忽略了用户需求来源的多样性。Bailetti A J 等人<sup>[5]</sup>从管理者的角度

收稿日期: 2009-06-29; 修回日期: 2009-09-11

基金项目: 国家自然科学基金(60872089); 西华大学人才培养课题(R0820208)

作者简介: 何忠秀(1974-),女,硕士,讲师,研究方向为人工智能与数据挖掘;王 霜,博士,副教授,研究方向为智能设计;杜亚军,博士,教授,研究方向为 Web 数据挖掘与知识发现,智能搜索引擎与智能网络。

强调了不同来源的用户需求与设计师经验相结合的重要性。邓家祺等<sup>[6-8]</sup>获取需求是通过市场调查以及统计分析调查信息而得到的,包含了目标市场信息和用户需求信息。随着互联网的广泛应用,大量用户的 Web 访问行为中包含了对产品的需求信息,可以被提取成用户需求。

文中提出了一种从互联网上多渠道获取用户需求的方法,试图综合用户浏览行为、方式,用户对产品的查询,网站提供的 BBS 以及用户投诉处理等多渠道表现的需求信息。

### 1 用户需求指标体系

分析用户对产品的实际需求,可以将需求分为时间、质量、费用、服务与社会性五个方面,共 20 个小组的内容。参照质量功能配置(Quality Function Deployment, QFD)建立质量屋的方法,分析出产品指标术语集合,建立用户需求指标体系。

在获取用户需求时,将模糊的需求表述转化为对该需求项的量化值。文中采用评分法进行需求量化,设有待研究产品的用户需求体系有  $m$  维,每维需求数最大有  $n$  项,则有  $R_{ij} \in [0, 1], i=1, 2, \dots, m; j=1, 2, \dots, n$  为第  $i$  维  $j$  项指标的量化值<sup>[9]</sup>。“完全不需要”为“0”,“非常需求”为“1”,中间状态在 0~1 区间内取值。

例如铅笔,可以根据情况分析出如下的用户需求指标体系的术语集合:笔芯(粗细、横截面形状、颜色、硬度)、外形(粗细、长度、横截面形状、图案、文字)、附件(带橡皮、带闪烁灯)。

不同的用户有不同的需求,记录不同用户的需求可以分析不同用户群体的不同需求,从而将二维的用户需求指标按照顺序排列为二维的,例如上面铅笔的转化为 11 列,外形中的粗细对应为第 5 列。相应的,将以用户需求指标项作为列,用户作为行,得到一个用户需求矩阵(Customer Requirement Degree Matrix)  $CR_{mn}$  ( $m$  为用户数目,  $n$  为用户需求指标项数),  $CR[i, j] \in [0, 1]$ , 数值表示了用户对该指标的需求度量值。

为了分析出不同类型的用户对产品的不同的关注和满意情况,处理不同用户群体的需求信息,需要建立一个用户信息库,记录每个用户的用户名、ID 号、常驻城市、地区、生日、职业、收入、爱好等。

识别不同的用户有三种基本方法:让用户注册获得用户唯一的 ID,对未注册用户使用日志中的 IP 地址和浏览器版本号来识别,使用 Cookie 技术让客户端

驻留一个唯一标识。

### 2 基于 Web 的用户需求知识获取方法

随着网络的发展,现在用户在购买产品前后大多会对相关网站进行浏览。在用户浏览、关注自己喜欢的网页时,就会有意无意地留下自己对某个产品的倾向和偏好。这些信息可能是用户自己也无法阐述清楚的,而网络服务器会记录下用户在网站的行为,这就提供了从网络中获取用户需求的可能。从 Web 上获取用户的需求可以从用户浏览方式,用户对产品的查询,网站提供的 BBS,用户投诉,以及网上用户问卷调查等方面来获取。图 1 给出了从这些渠道获取用户需求的框架。用户在浏览和查询时往往不会直接给出关于对产品性能指标的满意度,只能从中得到用户对一些需求指标的关注度,因此建立用户关注度矩阵  $CA$ ,而在 BBS 和用户投诉中既有对需求指标的关注信息又有对具体指标的满意度信息,据此可以建立用户满意度矩阵  $CS$ 。再由  $CA$  和  $CS$  矩阵进行融合,获得用户需求矩阵  $CR$ 。

用户问卷方法在网络上应用较多,一个设计优秀的问卷和有效的获取问卷答案的方式,也可获得用户关注度或用户满意度。由于方法较为成熟,在此不再赘述。

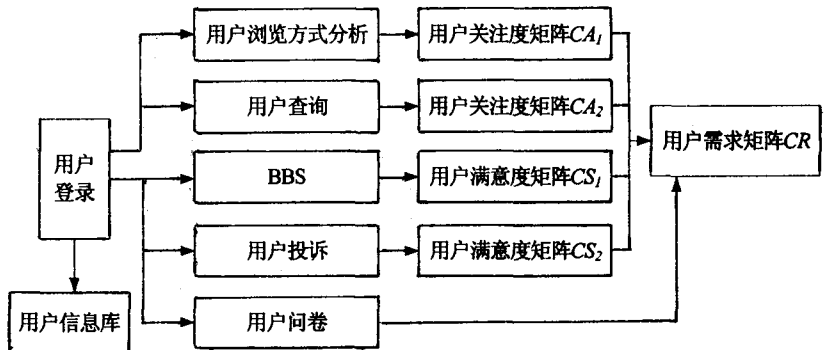


图 1 基于 Web 的多渠道用户需求知识获取框架

#### 2.1 用户浏览行为

在用户对专业网站进行浏览时,用户就会对感兴趣的网站、网页进行浏览,并且用户点击浏览网页的信息很多、用户数目很大,能够反映用户需求的统计规律,有些是用户本身也无法清楚表达的需求。用户浏览网页时在网站服务器上会留下该浏览者的浏览信息。在服务器端记录用户的访问过程,跟踪用户的 Web 访问习惯,记录在网络日志中。利用网络日志对人们在访问某领域网页时点击超级链接的情况和使用搜索引擎了解产品知识时所关注的方面进行跟踪记录,在此基础上进行需求知识发现,可以归纳出用户对于企业所关注的产品兴趣所在,并在一定程度上反映

出用户对该产品的潜在需求。图 2 显示了基于 Web 访问分析的用户需求获取过程。

在分析网络日志的基础上,获取用户点击浏览各网页数据,建立用户关注度矩阵(Customer Attention Degree Matrix)  $CA[]$ ,从网络日志中分析得到的用户关注矩阵记为  $CA_1[]$ ,  $CA_1[i, j] = a$ ,  $a$  为第  $i$  个用户在浏览网页时对第  $j$  项需求指标的关注度,这里以点击次数表示关注度。

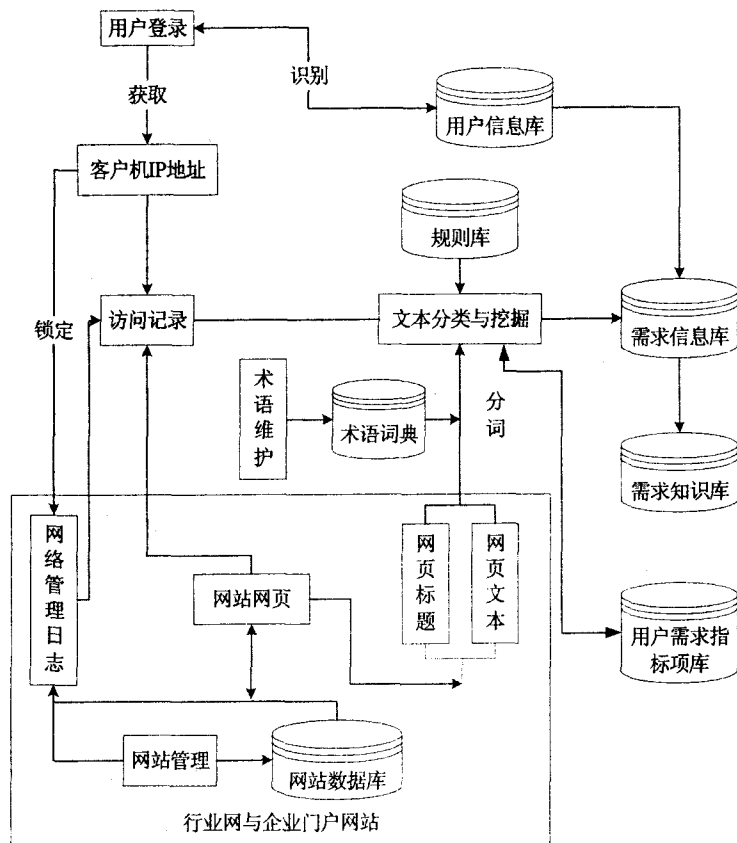


图 2 基于 Web 访问分析的用户需求获取方法

用户点击一次介绍产品的第  $j$  维指标的网页,则表明该用户对该项指标关注,在  $CA_1$  矩阵相应的  $i$  行和  $j$  列应增加 1。点击次数越多,反映该用户对该产品的这项指标越关注,该指标对该用户越重要。需要注意的是,如果一个网页包含了涉及产品的几个不同的指标,则应该在对应指标所在列上增加关注次数。

要分析网页涉及哪些指标,需要将网页中的文本信息提取出来,分析文本对应了产品指标的术语集中的词语,如图 2 所示。这时就需要将网页的标题、内容文本提取出来进行文本分析和挖掘,分析该网页关注了哪些用户需求项中的词语,才能在矩阵  $CA_1$  相应行列中增加关注次数。

### 2.2 用户查询

用户在 Web 页面进行资料查询时,其查询的内容是用户更加感兴趣、更为关注看重的,并往往以关键词

的形式出现,便于处理。记录用户的查询,可以直接获得用户所关注的产品兴趣,更能反映出用户对产品某些指标的潜在需求。用户查询在数量上虽远远不如用户点击,但是用户的查询通常是其最关注的主题,所以其比点击更有目的性,因此应该给予比浏览方式更大的权重值。同样的将这些查询信息记录在  $CA$  矩阵中,这里以  $CA_2[]$  表示,  $CA_2[i, j] = b$ ,  $b$  为第  $i$  个用户对第  $j$  项需求指标项进行的查询次数。

### 2.3 从 BBS、用户投诉中提取

BBS 和用户投诉的内容是用户自行输入,较真实地反映了用户真正的需要和满意情况,或者包含了对现有产品缺陷的改进意见。将 BBS 和用户投诉中的文本内容提取出来,对应于产品指标术语集中的术语进行收集和整理,建立一个用户满意度矩阵(Customer Satisfaction Degree Matrix)  $CS[]$ ,从 BBS 中获取的以  $CS_1[]$  表示,从用户投诉中获取的以  $CS_2[]$  表示。

文本内容的分析需要认知词集和意愿词集的支持。在实际中,由于用户表达方式的不同、语言的多义性和用户的不专业,所以需要尽可能地将各种和用户需求项相关的词语收集成用户词汇并逐一分类,生成用户认知词集。认知词集是用户对产品属性的认知表述,是用户需求指标集合的超集,和用户需求指标集合有联系但有所区别。建立一个从用户认知词集到用户需求指标集的映射  $\phi: C \rightarrow R$ ,  $C$  是用户认知词集,  $R$  是用户需求指标集,并且  $\forall c \in C, \phi(c) = r, r \in R$ 。

意愿词集是描述用户对某项主题的意愿程度的词汇汇集。意愿词集分为评价、好恶、程度等子集,子集里的每一元素都对应了意愿分  $\omega$ ,可以从用户意愿词与意愿分对应表中查表得到。意愿分  $\omega \in [-1, 1]$ ,  $\omega = -1$ ,表示非常不满意;  $\omega = 1$ ,表示非常满意。意愿分  $\omega$  在确定用户需求指标值时是一项重要的参数。

在 BBS 和用户投诉中,使用分词技术,获得用户对需求指标项的认知词和相应的意愿词(见图 3),在用户认知词集和意愿词集查找,如果都查找到,则取得对应的需求指标项和意愿分  $\omega$  值,计入  $CS_1$  或  $CS_2$  矩阵;如果不能获得用户认知词集或意愿词集中的词,则需要人工确定、手动添加进入用户认知词集或意愿词集。  $CS_m[i, j] = c$ ,  $c$  为第  $i$  个用户在 BBS ( $m = 1$ ) 或用户投诉中 ( $m = 2$ ) 对第  $j$  项用户需求指标的意愿分,意愿分越高表示用户对该项指标越满意,越低则表示

越不满意。

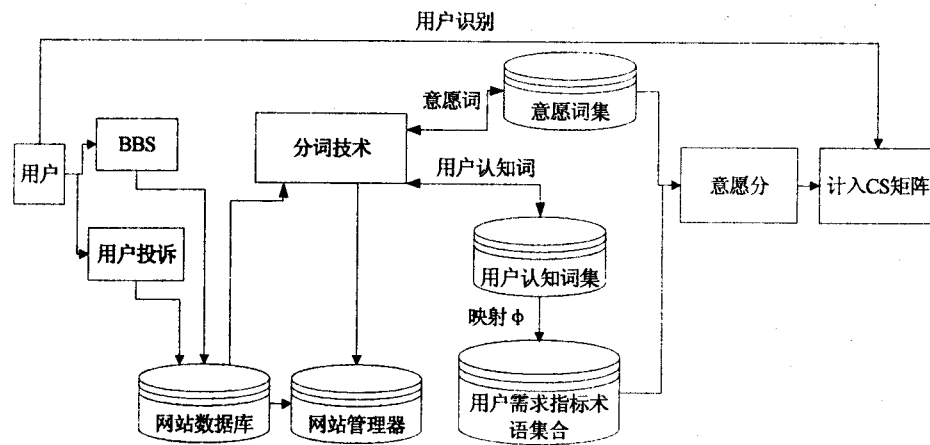


图 3 从 BBS 和用户投诉中的用户需求获取方法

### 3 需求指标的综合计算

从用户浏览网页行为和用户查询的分析中获得的用户关注度应按该渠道的可信度分别给予一定的权重进行叠加,并进行归一化处理以获得综合的用户关注度;从 BBS 和用户投诉中获得的用户满意度也按该渠道的可信度给予一定的权重叠加获得用户满意度,该两渠道的权重之和为 1。则满意度值空间为  $[-1, 1]$ , 其含义与意愿分  $\omega$  类似,绝对值大小表明意愿强烈程度,正值为满意,负值为不满意。

由此得到需求指标值的计算公式为:

$$CR[i, j] = \tau(v_1 CA_1[i, j] + v_2 CA_2[i, j]) * (v_3 CS_1[i, j] + v_4 CS_2[i, j])$$

其中:

$CR[i, j]$  - 体系中第  $i$  个用户第  $j$  用户需求指标的指标值;

$CA_1[i, j]$  - 从用户浏览行为分析中获得的第  $i$  个用户对第  $j$  项需求指标关注度的量化值;

$CA_2[i, j]$  - 从用户查询中获得的第  $i$  个用户对第  $j$  项需求指标关注度的量化值;

$CS_1[i, j]$  - 从分析 BBS 中获得的第  $i$  个用户对第  $j$  项需求指标满意度的量化值;

$CS_2[i, j]$  - 从分析用户投诉中获得的第  $i$  个用户对第  $j$  项需求指标满意度的量化值;

$\tau$  - 用户关注度的归一化算子;

$v_1, v_2, v_3, v_4$  - 各指标值的权值;

\* 为矩阵的按位相乘。

从公式中可以看到,用户对某一需求指标的关注度越高,相应的满意度越低(为负数),说明对该项指标需求越关心但是不满意,企业越应在这个指标上进行改进;用户对某一需求指标的关注度越高,相应的满意

度越高,说明对该项指标需求越满意。因此需求指标值矩阵  $CR$  能给出用户对产品各需求指标的反馈,从而指导企业改进产品。

需要指出的是,  $CR$ 、 $CA$  及  $CS$  矩阵是稀疏矩阵,为了压缩数据应使用三元组来表示。

### 4 结束语

从网络不断产生的海量信息中挖掘用户需求是一个巨大的挑战,笔者试图从企业和行业网站的现存信息中找到进行用户需求知识发现的方法,分别讨论了从 4 种渠道中获取需求信息的方法,并根据所获取信息的特点分别构建了用户关注度矩阵和用户满意度矩阵,给出计算公式将两类矩阵融合为代表用户需求知识的需求矩阵,对于企业的产品开发计划具有一定参考价值。在此基础上根据企业需要进行针对性的知识发现则具有更为广阔的发展前景。

#### 参考文献:

- [1] Chen C H, Khoo L P, Yan W. A strategy for acquiring customer requirement patterns using laddering technique and ART2 neural network[J]. Advanced Engineering Informatics, 2002, 16(3): 229-240.
- [2] Wu Hsin - Hung, Shieh Jiunn - I. Using a Markov chain model in quality function deployment to analyse customer requirements[J]. The International Journal of Advanced Manufacturing Technology, 2005, 30(1-2): 141-146.
- [3] 吴昭同, 余忠华, 陈文华, 等. 机械设计丛书——保质设计[M]. 北京: 机械工业出版社, 2004.
- [4] 王美清, 唐晓青. 产品设计中的用户需求与产品质量特征映射方法研究[J]. 机械工程学报, 2004, 40(5): 136-140.
- [5] Bailetti A J, Litva P F. Integrating Customer Requirements into Product Designs[J]. Journal of Product Innovation Management, 1995, 12(1): 3-15.
- [6] 乌兰木其, 吴 斌, 邓家祺. 面向创新的产品需求获取与处理研究[J]. 北京航空航天大学学报, 2001(4): 202-205.
- [7] 吴 斌, 乌兰木其, 邓家祺. 创新产品需求敏捷获取与处理系统的实现研究[J]. 机械设计, 2001(9): 9-11.
- [8] 王丽英, 吴 斌, 乌兰木其, 等. 面向创新的用户需求获取与处理系统研究[J]. 机械设计, 2001(5): 1-2.
- [9] 王 霜, 殷国富, 何忠秀. 基于 Kano 模型的用户需求指标体系研究[J]. 包装工程, 2006, 27(4): 209-210.