

模糊逻辑在电子商务商品推荐系统中的应用^①

黄 洪, 杨卓俊, 王 奔

(浙江工业大学 计算机软件与理论, 杭州 310012)

摘 要: 在电子商务平台中提供有针对性的商品推荐机制以帮助用户快速准确地找到所需的商品非常重要。综合满意度是指根据用户关注的商品的不同属性的满意度在商品推荐中所占的权重不同, 对商品的各属性满意度进行加权运算得到的满意度度量值。根据商品综合满意度的排序进行商品的推荐可以提高推荐的准确性。商品满意度是一个模糊的概念, 提出了使用模糊逻辑来处理商品满意度的方法, 并提出了一整套基于满意度的商品推荐机制。论文介绍了一个基于商品满意度的电子商务商品推荐系统的框架, 重点介绍了其中的基于模糊逻辑的商品属性满意度处理方法和基于商品综合满意度的商品推荐规则, 并给出了一个商品推荐的模拟实例。

关键词: 电子商务商品推荐系统, 模糊逻辑, 商品满意度, 商品综合满意度

Application of Fuzzy Logic to E-Commerce Recommendation System of Commodity

HUANG Hong, YANG Zhuo-Jun, WANG Ben

(Computer Software and Theory, Zhejiang University of Technology, Hangzhou 310012, China)

Abstract: In e-commerce platform to provide targeted product recommendation mechanism to help users quickly and accurately find the merchandise is very important. Comprehensive Satisfaction is the focus of the product based on user satisfaction with different attributes of the commodity the weight recommended in the different properties of the satisfaction of all goods operations by the weighted measure of satisfaction. Comprehensive Satisfaction according to the sort of goods where goods are recommended to improve the accuracy of recommendation. Product satisfaction is a vague concept, this paper proposes the use of fuzzy logic to handle goods satisfaction methods, and proposed a set of recommendations based on product satisfaction mechanism. This paper introduces a commodity-based e-commerce product satisfaction recommendation system framework, which focuses on product attributes based on fuzzy logic approach and satisfaction with Comprehensive Satisfaction based on commodity products recommended by the rules, and gives a recommendation of goods Simulation instance.

Key words: e-commerce product recommendation system; fuzzy logic; product satisfaction; comprehensive satisfaction of goods

随着互联网的不断普及, 人们越来越习惯与在网上消费、购物, 商家也想尽办法来帮助客户发现他们想要的商品, 以便于提高自身销量。在此背景下, 电子商务商品推荐技术成为当下电子商务领域的一个研究热点。

电子商务商品推荐技术是指系统对收集到的用户基本数据和其所购买的商品信息进行分析, 得

到用户的购买习惯, 实现主动向用户智能推荐商品的一项技术^[1]。目前常见的推荐技术有: 基于内容过滤、基于协同过滤、基于人口统计、基于知识、基于效用、基于 Web 挖掘和基于领域本体的推荐技术等, 其中基于协同过滤和基于内容过滤技术在电子商务商品推荐系统中应用最为广泛^[2]。

电子商务商品推荐系统是指利用电子商务网站这

① 收稿时间:2011-06-21;收到修改稿时间:2011-07-17

个平台向客户提供其感兴趣的商品信息和购买建议,帮助用户决定该购买什么样的商品,并且模拟销售人员来帮助客户完成购买过程的应用系统。电子商务商品推荐系统的主要目的是将浏览者变成购买者,提高电子商务网站的交叉销售能力,提高客户对电子商务网站的忠诚度^[3]。

由于现有的电子商务商品推荐系统存在着多种问题,例如由于用户对商品的评价非常稀疏,这样基于用户的评价所得到的用户间的相似性可能不准确;随着用户和商品的增多,系统的性能会越来越低;随着商品不断增加,类似商品越来越多,降低了系统推荐的准确性;随着用户数量的不断增加,具有类似特点的用户越来越多,降低了系统推荐的准确性等^[4]。本文针对现有电子商务商品推荐系统存在的上述缺陷,提出了一种综合应用模糊逻辑算法、语义搜索算法和遗传算法的电子商务商品推荐系统,希望可以提高商品推荐的准确率,加快商品推荐的速度。

1 电子商务商品推荐系统的总体结构

系统主要有四部分:基于模糊逻辑的商品信息处理模块、基于语义词典的语义搜索模块、基于遗传算法的知识学习模块、推荐引擎模块。系统总体框图如图 1 所示。

基于模糊逻辑的商品信息处理模块:以模糊数据集为基础,应用模糊逻辑算法计算相关商品属性的满意度,以此为依据构建商品满意度矩阵。

基于语义词典的语义搜索模块:以现有的语义词典 WordNet 为基础^[5],设计一个基于中文的语义词典,应用 Wu-Palmer 概念语义相似度算法,对搜索的商品进行语义分析,并结合商品综合满意度进行智能化商品推荐。

基于遗传算法的知识学习模块:以系统采集的用户行为数据为基础,通过遗传算法分析用户行为,得到相应知识,并构建商品类型推荐因素权重矩阵。

推荐引擎模块:对商品信息处理模块中取得的商品属性满意度值和知识学习模块中取得的商品商品推荐因素权重值进行加权计算,获得商品综合满意度,并给出商品推荐。

由于本文篇幅有限,在此重点讨论基于模糊逻辑的商品信息处理模块,给出具体的模糊逻辑算法,并通过算法取得相应商品的满意度值。

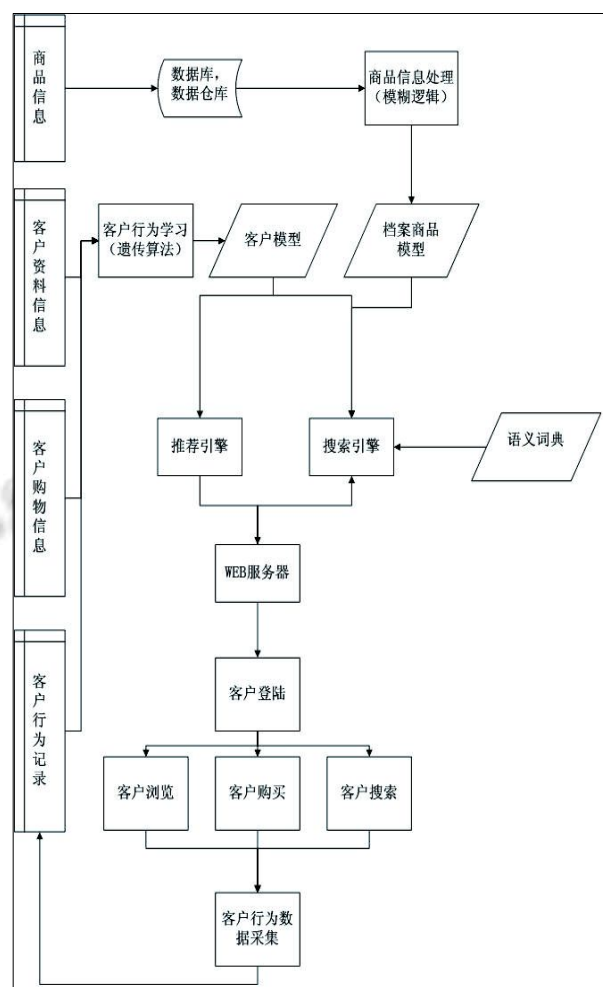


图 1 系统总体框图

2 模糊逻辑在商品推荐系统中的应用

在本节中,将首先介绍商品满意度及其相关概念,提出商品满意度矩阵和商品推荐因素权重矩阵,然后介绍模糊逻辑的概念,并给出基于模糊逻辑的商品综合满意度计算方法。

2.1 商品满意度及其相关概念

满意是一种心理状态。是客户的需求被满足后的愉悦感,是客户对产品或服务的事前期望与实际使用产品或服务后所得到实际感受的相对关系。如果用数字来衡量这种心理状态,这个数字就叫做满意度,客户满意是客户忠诚的基本条件。

商品满意度是指通过对商品各项属性的评价分值的加权计算,得到的测量用户对商品的满意程度的一种指数概念。

商品满意度矩阵是指把所有商品的属性满意度,

用一个 $M \times N$ 阶矩阵来表示, M 行代表 M 个商品, N 列代表该商品 N 个属性满意度, 第 m 行第 n 列的元素 X_{mn} 代表商品 m 在属性 n 上的满意度, 根据模糊函数的定义满意度的取值在 $[0,1]$ 区间内, 商品满意度矩阵如图 2 所示。

满意度 \ 属性	属性 1	属性 2	...	属性 n
产品 1	0.3	0.2	...	0.3
产品 2	0.1	0.2	...	0.3
...	...	...	...	...
产品 m	0.2	0.1	...	0.2

图 2 商品满意度矩阵

商品推荐因素权重矩阵是指通过建立一个 $M \times N$ 阶矩阵, 把影响商品推荐的不同因素的权重表示出来, M 行代表 M 类商品, N 列代表影响商品推荐的 N 个因素, 第 m 行第 n 列的元素 X_{mn} 代表第 m 类商品的第 n 个属性的推荐权重值。商品类型推荐因素权重矩阵如图 3 所示。

权重 \ 属性	属性 1	属性 2	...	属性 n
类别 1	0.2	0.4	...	0.1
...	...	...	...	...
类别 m	0.2	0.1	...	0.3

图 3 商品推荐因素权重矩阵

同一商品不同属性的满意度在商品推荐中所占的比重是不同的, 对商品的属性满意度进行加权运算得到的商品满意度, 称为商品的综合满意度。

2.2 满意度的模糊逻辑表示

2.2.1 商品属性满意度的模糊逻辑表示

“模糊逻辑”是 1965 年由美国工程师 LA Zadeh 在其改进计算机程序的“模糊集合理论”中提出的一个概念。传统的计算机通常只能按照“是与否”、“对与错”、“0 与 1”这样的二元逻辑进行识别, 而对冷、热、大、小这样的模糊概念无能为力。有了模糊逻辑, 计算机开始能够在中间地带发挥作用, 可以在信息有限的条件下, 提供精确的答案^[6]。

在电子商务网站中, 存在着一系列可供选择的商品, 每件商品都具有多个属性, 这些属性存在着相互

间的冲突并且难以被精确定义, 用户通常只能模糊地描述这些属性^[7]。例如, 在购买笔记本电脑的问题中, 用户可能对商品有以下要求: 价格在 1000\$ 左右; 硬盘空间大于 40G; 内存大小大于 2G; 有较快的运行速度; 有较低的返修率。以上这些说法在生活中是较为常见的, 但在计算的意义上却是模糊的。因此, 我们使用模糊逻辑对电子商务商品的属性进行量化处理, 使用模糊函数把日常术语表示出来。

在电子商务网站的商品购买中, 用户的需求一般是具有明确的方向性的, 例如随着商品价格的提高用户对商品的满意度是逐步减少的, 而随着产品性能的提高用户对商品满意度是逐步增加的, 因此我们使用以下两个模糊函数来描述用户对商品的满意程度。本文使用正态隶属函数来建立模糊函数^[8], 假如商品属性 X 的取值为 x , 则对该商品在属性 X 上的满意度可以用公式(1.1)或(1.2)表示, 当满意度随着商品属性值 x 的增加而减少时, 我们选择公式(1.1), 当满意度随着商品属性值 x 的增加而增加时, 我们选择公式(1.2), 具体公式如下:

$$\mu(x) = \begin{cases} 1, & x \leq a \\ e^{-k(x-a)^2}, & x > a \end{cases} \quad (1.1)$$

$$\mu(x) = \begin{cases} 1, & x \geq a \\ e^{-k(x-a)^2}, & x < a \end{cases} \quad (1.2)$$

隶属函数中的 K 值一般使用模糊统计法、例证法, 专家经验法和二元对比排序法等方法来确定。所谓专家经验法是指根据专家的实际经验给出模糊信息的处理算式或相应权系数值来确定隶属函数的一种方法。在许多情况下, 经常是初步确定粗略的隶属函数, 然后再通过“学习”和实践检验逐步修改和完善, 而实际效果正是检验和调整隶属函数的依据^[9]。在此我们使用专家经验法选择 $K=0.001$, a 为商品对应属性 X 在满意度最高时的属性取值, x 为商品在属性 X 的取值。如价格在 1000\$ 左右的模糊函数如图 4 所示, 此图对应公式(1.1), 当价格小于或等于为 1000\$ 时, 用户对价格属性的满意度为 1, 当价格增加时, 用户的满意度逐步减少。

2.2.2 商品综合满意度

在电子商务商品交易中, 用户在选择类似商品的时候通常会产生一定的模糊性。例如, 为什么选择这件商品或为什么不选择同类的那件商品。在解释这些

问题的时候,用户往往会提及一系列模糊概念,而电子商务商品推荐系统则要帮助用户把这些概念量化,以数字的形式加以描述,并把得到的结果作为推荐的依据。

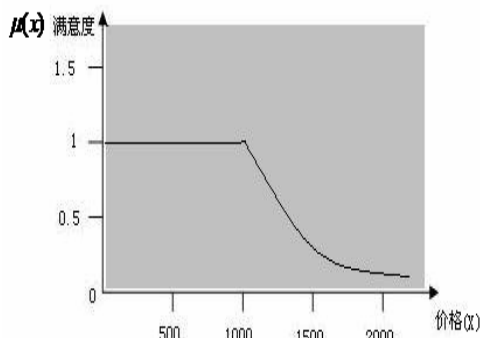


图 4 笔记本价格的模糊函数

文献[10]提出的商品分类算法在商品分类领域有很好的实用性,但其选择具有主观性,极大地降低了分类的准确性。

本文提出的建立商品满意度的方法很好地解决了上述问题。本文先对商品属性满意度进行取均值,并定义商品某一属性的平均满意度为用户所能容忍的最低满意度,用户对该商品在该属性上的满意度与用户对该商品属性的平均满意度之差,称为属性的满意偏离量。如图 4 所示,假设两款 SONY 笔记本的价格满意度为 0.7 和 0.9,如果商品的价格属性的平均满意度为 0.8,则 A 商品的属性满意偏离量为 -0.1,而 B 商品的属性满意松弛度为 0.1,由此可知 B 商品在该属性上符合用户的要求,而 A 商品在该属性上不符合用户的要求。当某商品属性的满意偏离量为负值,而综合满意度满足用户需求,则用户对此商品属性需要给出让步,称为让步满意。当两件商品具有相同的综合满意度时,优先推荐用户对所有属性均满意的商品,然后再推荐用户对商品属性让步满意的商品。通过这种方式获得的商品属性满意度评分,极大地提高了推荐的准确性。

2.2.3 基于模糊逻辑的商品满意度算法

设电子商务网站一共有 m 件商品,商品 P_i ($i=1, 2, \dots, m$) 有 n 个属性, Q_{ij} ($i=1, 2, \dots, m$ 且 $j=1, 2, \dots, n$) 代表商品满意度矩阵中的第 i 件商品的第 j 个属性的满意度。模糊集 $\{x_{i1}, x_{i2}, \dots, x_{in}\}$ 是用户对第 i 种商品各属性的满意度值。第 i 件商品的第 j 个属性的

满意度值可以表示为 x_{ij} ($i=1, 2, \dots, m$ 且 $j=1, 2, \dots, n$), 则商品属性的平均满意度值为:

$$AVG_j = \frac{1}{m} \sum_{i=1}^m x_{ij} \quad (j=1, 2, \dots, n) \quad (2.1)$$

由公式(2.2)可得,商品属性的满意偏离量为:

$$H_i = x_{ij} - AVG_j \quad (j=1, 2, \dots, n) \quad (2.2)$$

我们可以通过下列公式得到第 i 件 ($i=1, 2, \dots, m$) 商品 k 个属性的综合满意度值 I_i , 其中 m_k 为商品类型推荐因素权重矩阵中的相对应属性权值:

$$I_i = \sum_{j=1}^k x_{ij} \times m_k \quad (2.3)$$

通过商品属性的满意偏移量和商品的综合满意度,我们可以得到用户满意与让步满意的商品,当两件商品综合满意度相同时,我们先选择用户满意的商品进行推荐,然后才选择用户让步满意的商品进行推荐。

2.2.4 基于满意度的商品推荐机制

已知商品集合 $P=\{p_1, p_2, \dots, p_k\}$, 若选择 $1, 2, \dots, k$ 个属性, 则商品综合满意度求值过程如下:

- (1) 获得商品参与运算的属性, 建立模糊集 $\{x_{i1}, x_{i2}, \dots, x_{in}\}$;
- (2) 根据公式(2.1)计算商品的平均满意度值;
- (3) 根据公式(2.2)得到商品属性的满意偏离量;
- (4) 根据公式(2.3)结合商品类型推荐因素权重矩阵, 获得商品的综合满意度值。
- (5) 根据获得的商品属性满意偏离量和商品的综合满意度, 选择用户满意与让步满意的商品进行推荐。假设商品数量为 $10n(n=1, 2, \dots, x)$ 件, 则选择综合满意度值排名前 $10m(m=1, 2, \dots, x)\%$ 的商品进行推荐, m 可以根据具体的情况进行动态地设置。

2.3 一个商品推荐的模拟示例

假设有以下 6 种笔记本电脑, 相应电脑的属性如图 5 所示:

使用公式 2.1 有得到价格的平均满意度为 0.63, 保修年限的平均满意度为 0.75, 重量的平均满意度为 0.7。通过公式 2.2, 得到价格属性的让步满意集为 $\{1, 2, 5\}$, 保修年限属性的让步满意集为 $\{4, 6\}$, 重量属性的让步满意集为 $\{1, 4, 6\}$ 。

我们在这里假设价格、保修年限、重量这三个属性的权重值为 0.5、0.2、0.3, 则六件待推荐商品的综合满意度为 $\{0.68, 0.64, 0.79, 0.72, 0.60, 0.73\}$, 由

获得的结果可知笔记本{3}拥有最高综合满意度,而笔记本{1, 2, 3, 5, 6}在属性上是让步满意的,因此推荐顺序应为{3, 6, 4, 1, 2, 5},第 3 件商品最符合用户的需求。

值 \ 属性 \ 笔记本	价格(元)	保修年限 (年)	重 量 (kg)
1	5000	2	1.6
2	6000	2	1.4
3	4500	2	1.4
4	3000	1	1.8
5	7000	2	1.2
6	3500	1	1.6

图 5 笔记本属性

笔记本的价格,保修年限,重量三个属性的满意度如图 6 所示:

满意度 \ 属性 \ 笔记本	价格	保修年限	重量
1	0.6	1	0.6
2	0.4	1	0.8
3	0.7	1	0.8
4	1	0.5	0.4
5	0.2	1	1
6	0.9	0.5	0.6

图 6 笔记本属性满意度

3 结语

随着电子商务网站提供的商品不断增多,商品推荐难度越来越大,本文给出了一种商品推荐方法,以模糊逻辑算法为基础,建立商品满意度矩阵,同时结合商品类型推荐因素矩阵,得到综合满意度值,旨在提高推荐的效率和准确性。

本文提到的基于模糊逻辑的商品信息处理方法可以作为用户在电子商务网站上的购买支持,为了提高推荐的准确性,参与运算的属性权重值较为重要,商品类型推荐因素权重表的建立是之后工作的重点,将在本系统的基于遗传算法的知识学习模块中进行处理。

参考文献

1 周惠宏等.推荐技术在电子商务中的运用综述.计算机应用研究,2004,(1):8-9.

2 刘玮.电子商务系统中的信息推荐方法研究.情报科学,2006,(2):10-11.

3 黄晓斌.网络信息挖掘.北京:电子工业出版社,2005.5-6.

4 黎星星.电子商务推荐系统研究.计算机工程与科学,2004,26(5):7-10.

5 Budan IA, Graeme H. Evaluating WordNet-Based Measures of Semantic Distance. Computational Linguistics, 2006,32(1): 13-47.

6 扎德,陈国权.模糊集合、语言变量及模糊逻辑.北京:科学出版社,1982.15-16.

7 Shipley MF, de Korvin A, Omer K. A fuzzy logic-based decision model to satisfy goals for successful product/service introduction. European Journal Operational Research, 2001(133): 209-219.

8 Dubois D, Rade HP. Addition of interactive fuzzy numbers. IEEE Trans, Automat. Contr. AC, 1981(26):926-936.

9 Ryu YU. A hierarchical constraint satisfaction approach to product selection for the electronic shopping support. IEEE Trans, Syst Man Cybern, Part A, Syst Humans, 1999(29):525-532.

10 李擎,郑德玲,唐勇,陈占英.一种新的模糊遗传算法.北京科技大学学报,2001.

(上接第 54 页)

Detection. 2008:116-134.

3 Peng JF, Chen H. CUGrep: A GPU-based high performance multi-string matching system. International Conference on Future Computer and Communication, 2010:77-81.

4 Wu S, Manber U. A fast algorithm for multi-pattern searching. University of Arizona, Technical Report TR- 94-17, 1994.

5 NVIDIA Corporation.NVIDIA CUDA Programming Guide 3.0..http://www.nvidia.com/.

6 张舒,褚艳利,赵开勇,张钰勃.GPU 高性能运算之 CUDA.北京:中国水利水电出版社,2009.

7 Boyer RS, Moore JS. A fast string searching algorithm. Communications of the ACM, 1977,20(10):762-772.