

[文章编号] 1009- 6043(2013)08- 0067- 03

电子商务推荐系统中用户建模研究

——基于用户隐性行为分析的视角

赵 虎 , 余小鹏

(武汉工程大学 管理学院, 湖北 武汉 430205)

[摘 要] 用户模型是电子商务智能推荐系统的关键,它是系统智能推荐的依据,是决定个性化信息服务质量的关键因素。基于用户隐性行为分析的视角,分析用户的兴趣度,在原型系统的基础上,通过对用户行为的分析,推导出用户隐性知识,从而对用户模型进行构造和更新,构造出原型模型和个体用户模型,并探讨了二者间交互机制与个体模型的更新机制,同时,通过浏览行为估计用户兴趣度的方法无需用户主动参与,有助于提高个性化服务系统的亲和力,在理论上具有可行性,但在实际运用中还有一定难度,主要表现在时间滞后性与复杂的计算过程。

[关键词] 电子商务;推荐系统;用户模型;智能推荐;用户兴趣度预测

[中图分类号] F450

[文献标识码] A

User Modeling Research in the Recommendation System of E-Commerce from the Perspective of User Recessive Behavior Analysis

ZHAO Hu, YU Xiaopeng

Abstract: User model is the key and the basis for an intelligent recommendation system of e-commerce, and critical factor to determine the quality of personalized information service. User interest degree is analyzed on the basis of user recessive behavior. A prototype model and an individual user model are constructed and updated through user behavior analysis based on a prototype system, deducing user tacit knowledge for building and updating the user model. The interaction mechanisms between the two models and the update mechanism of the individual model are also discussed. Although the approach of predicting user interest degree by means of browsing behavior doesn't need active user participation and helps improve the affinity of the personalized service system and the method is feasible in theory, there are still some difficulties in terms of time lag and complex calculation in practice.

Key words: e-commerce, recommendation system, user model, intelligent recommendation

引言

推荐系统是利用计算机技术与人工智能技术,发现和定义用户兴趣,推荐符合用户兴趣爱好的对象,也称个性化推荐系统。

由于电子商务推荐系统所推荐的对象是商品,其推荐成功的结果是发生购买行为,为商家带来实际可见的利益,更成为智能推荐领域研究热点。然而,由于商品种类繁多、信息海量,且用户购买行为复杂,用户的兴趣领域既有交集,又各有不同,因此,对个体用户兴趣度进行预测,从而构造出精准的个体用户模型,成为电子商务推荐系统研究的重要着力点。本文在原型系统的基础上,通过对用户行为的分析,推导出用户隐性知识,从而对用户模型进行构造和更新,其研究思路如图 1 所示:

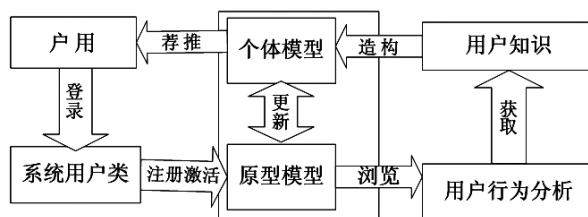


图 1 基于用户行为的用户模型构造

一、用户行为分析与用户知识识别

一个精准的个体用户模型最重要的目的和功能便是使系统准确地定义用户的兴趣取向与兴趣度,所以首先必须对用户行为进行系统分析,获取用户知识。“用户知识”是指关于用户的有组织的经验、价值观、相关信息及洞察力的动态组合,它所构成的框架可以不断地评价和

[收稿日期] 2013- 07- 10

[作者简介] 赵虎(1989-), 武汉工程大学管理学院技术经济及管理专业研究生。研究方向:管理信息系统,电子商务;

余小鹏(1974-), 武汉工程大学管理学院副教授,硕士研究生导师,工学博士,武汉大学博士后。研究方向:电子商务与信息系统数据挖掘、决策支持系统。

[基金项目] 武汉工程大学研究生创新基金资助项目。

吸收新的经验和信息。用户兴趣度表示用户对推荐商品的感兴趣程度。收集与用户兴趣关联程度比较大的行为数据,对这些数据进行处理和分析,获取隐性用户知识,是预测用户兴趣度的一种有效途径。

(一)用户知识识别

显性用户知识:是经过人的整理和组织,可以以文字、公式、计算机程序等形式表现出来,并可以通过正式的、系统的方式加以传播,便于其他人学习的知识。隐性用户知识:是指与人结合在一起的经验性知识,主要包括用户使用系统的习惯、爱好、需求、个人位置等信息。隐性知识只能通过对隐式指示器收集的数据进行分析来获得。综合国内外研究现状,常见的隐式指示器有:访问时间、页面访问次数、滚动条、鼠标移动、方向键、翻页键等。

(二)用户兴趣度预测

通过对用户注册时提供的显性知识和浏览商品时的隐性行为分析,可以对用户所感兴趣的内容及兴趣度进行预测。基于用户兴趣构造用户模型,并以用户模型为支撑进行商品的智能推荐,可以大幅提高推荐系统的质量。

用户注册时提供的显性知识是反映用户特质的基本信息,体现了用户行为的规律性与一致性,描述了用户的基本特征。与消费行为相关性强的用户显性知识主要有性别、年龄、学历、职业、收入及兴趣爱好等。由于这些显性知识系统一般无法通过人工学习获得,故虽与消费行为有极高的相关性,却只能通过用户的隐性行为加以分析预测。

用户在浏览商品时会产生许多隐性行为,这些行为是用户在没有花额外时间去理解标准情况下产生的。所以隐性行为虽然不像显性行为那样具有系统性和精确性,但却是用户兴趣的最原始体现,故对隐性行为通过数据挖掘的方法进行分析,能更完善的体现用户的兴趣,是显性知识的一个很好的补充。在用户隐性行为中最易收集又能体现用户兴趣的有:访问次数、驻留时间、鼠标移动次数、滚动条滚动次数、以访问历史记录。

令 $Interest(p)$ 为用户对商品 p 的兴趣度,则 $Interest(p)$ 可以表示为:

$$Interest(p) = F(\text{Save}(p), \text{Freq}(p), \text{Duration}(p), \text{Mouse-move}(p), \text{Page-move}(p)) \quad (1)$$

公式 2 的含义如下 (1)如果用户保存了页面,则认为用户对类商品极感兴趣,可重点推荐这类商品。(2)如果用户在浏览过程中多次重复浏览每一商品,即对某浏览频率较高,则表明对该商品很感兴趣。(3)如果页面驻留时间很长,并伴随有鼠标、滚动条活动,则认为用户对该页面商品感兴趣。(4)用户浏览网页时的点击行为或许是一个习惯性动作,“click”是用户表达情绪的一种方式,尤其是用户点击查看商品大图时,表明用户对该商品感兴趣。(5)用户常常通过卷屏来浏览或细看商品内容,这说明卷屏行为同用户兴趣有一定关系。

二、用户模型表示

用户模型是用户与外部世界联系的知识模型,是人

们对外部世界的认识以及人与外部世界交互的描述。用户模型对用户进行“虚拟”表示,主要涉及用户知识表示及其动态维护问题。用户模型由描述一类用户的公共特征的原型模型与表示单个用户的特征的个体模型组成。面对各种类型的用户,系统仅有典型模型,应用是很有限的。个体模型能使系统对每个用户的需求提供一个合适的接口,并有助于兴趣情报的描述。

(一)原型系统表示

主要原型模型的构造是在系统设计时进行的,直接包含在系统的知识库中,然后在系统运行中,根据需要从它们推导新的、较特殊的原型。而所有的用户个体模型,均在系统运行中建立。从结构方面出发,一个原型具有与用户模型类似的组织结构(用户行为特征除外),它的基本结构如图 2:

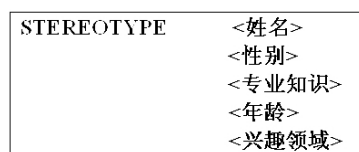


图 2 用户原型模型

每个原型都有一个激活方法,用户在首次进入系统时,系统会动态分配给用户一个固定 ID,并且原型系统中的每个槽值,如姓名、性别等也都会被赋值,此时原型模型即被激活。原名模型被激活的唯一标志是该用户拥有了一个 ID,其他信息都是可以改变的,甚至注册时该信息为默认的空值,所以此模型不能表现出用户的真正兴趣,在智能推荐中的作用不明显。一个能准确体现用户兴趣的模型称为用户个体模型,用户个体模型是通过对用户行为的分析,获取用户的兴趣领域,并把准确的信息添加到原型模型或对原型模型进行修改之后的结果。用户个体模型是对用户动态行为收集分析综合的结果。

(二)个体用户模型表示

用户模型的建立是从用户进入系统开始的,当用户进入系统后,系统首先查找个体用户模型库,如果没有该用户的模型,则被认定为新用户。系统就会从零开始建立用户模型,过程分五个步骤 (1)获取显性知识。这一过程系统会动态为该用户分配一个 ID,表示用户的主索引 i 。(2)激活原型。原型中被激活的原型称为活动原型(s_i),它们可以作为构造当前用户模型的开始点。(3)辨别原型。原型建立后用户会浏览商品,初步获取用户感兴趣的商品集合,用 $Personal_I$ 表示。(4)精炼模型。系统重复执行以下两个活动:应用合适的获取方法,收集或产生与用户有关的新信息;应用有关的实证方法检查和纠正新信息的一致性。模型精炼工作直到当前咨询结束为止。收集用户的隐性行为集合,用 $Personal_R$ 表示。(5)结束工作。当咨询结束时,当前用户的模型已初次完成,连同用户标识名一起存入用户模型库。系统经过智能分析,得出用户行为 $Personal_R$ 与用户兴趣本体 $Personal_I$ 的关系 ε ,并将用户最感兴趣前 n 件商品计入用户数据库,用户原型中的兴趣领域改变,用户建模结束。

系统经过以上步骤的交互,获取了用户信息,将获取的知识表示出来,即为用户模型。用户模型表述如下:

UserModel=(Stereo Type, Personal_I, Personal_R, ϵ) (2)

(1)StereoType:表示用户原型模型。一个系统只有一个原型类,这是在系统规划时设计好的,每个用户的进入,都会将该类实例化一次。 S_i 表示第*i*个用户的基本信息。

$S=(S_1, S_2, S_3, \dots, S_n)$ (3)

(2)Personal_I:包含用户兴趣的个性化领域本体,即某一用户浏览商品的集合, i_i 表示该用户浏览的第*i*件商品。

$I=(i_1, i_2, i_3, \dots, i_n)$ (4)

(3)Personal_R:用户行为本体,表示用户浏览商品时的行为的集合。 $(c_{s,i}, f_{s,i}, d_{s,i}, m_{s,i}, p_{s,i})$ 表示ID为*s*的用户浏览第*i*件商品时行为的集合。

$R=(C, F, D, M, P)$
 $=\{(c_{1,1}, f_{1,1}, d_{1,1}, m_{1,1}, p_{1,1}), (c_{1,2}, f_{1,2}, d_{1,2}, m_{1,2}, p_{1,2}), \dots, (c_{1,n}, f_{1,n}, d_{1,n}, m_{1,n}, p_{1,n}),$
 $(c_{2,1}, f_{2,1}, d_{2,1}, m_{2,1}, p_{2,1}), (c_{2,2}, f_{2,2}, d_{2,2}, m_{2,2}, p_{2,2}), \dots, (c_{2,n}, f_{2,n}, d_{2,n}, m_{2,n}, p_{2,n}),$
 $\dots\dots$
 $(c_{n,1}, f_{n,1}, d_{n,1}, m_{n,1}, p_{n,1}), (c_{n,2}, f_{n,2}, d_{n,2}, m_{n,2}, p_{n,2}), \dots, (c_{n,n}, f_{n,n}, d_{n,n}, m_{n,n}, p_{n,n})\}$ (5)

(4) ϵ :表示基于某一具体原型的用户从用户行为到用户个性化的映射。 $\epsilon(s, i)$ 表示ID为*s*的用户对第*i*件商品的兴趣度。

$\epsilon(s, i)=F(R(s, i))$
 $=F(c_{ij}, f_{ij}, d_{ij}, m_{ij}, p_{ij})$ (6)

三、用户模型的更新

如果用户是登陆系统,即系统已知当前用户,那么该用户模型已包含在用户模型库中,所以用户第二次操作便是对模型的更新,而不需要重新建立模型了。更新用户模型的操作有所不同,与建模相比,步骤(1)~(3)由另两个新的活动代替。更新模型的操作包含四个步骤:

(1)检索模型。识别当前用户为已知用户,即从用户个

体模型库检索该用户的模型。在 $(S_1, S_2, S_3, \dots, S_n)$ 依次遍历查询,若有进行智能推荐。

(2)智能推荐。根据已有的用户模型,查找出用户最感兴趣的商品或商品类目,优先呈现给用户。即根据 ϵ 的结果将用户最感兴趣前 λ 件商品和用户原型中的兴趣领域在用户进入系统时进行推荐。并在此基础上继续收集用户行为,并将在当前咨询中进一步精炼。

(3)精炼模型。

(4)结束工作。

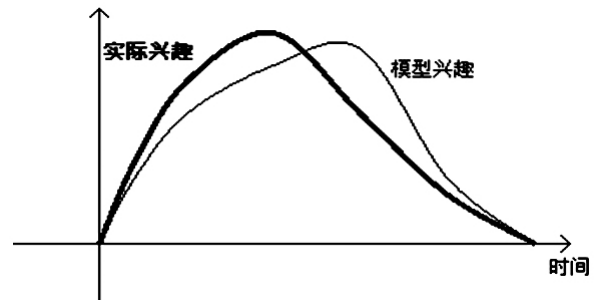


图3 用户模型更新示意图

四、结论与展望

通过浏览行为估计用户兴趣度的方法无需用户主动参与,有助于提高个性化服务系统的亲和力,在理论上具有可行性,但在实际运用中还有一定难度,主要表现在时间滞后性与复杂的计算过程,在下一阶段的研究工作中将主要研究如何解决这一问题。

[参 考 文 献]

- [1]李勇.智能信息检索中基于本体的个性化用户建模技术及应用[C].国防科技大学硕士论文,2002
- [2]单蓉.个性化信息系统中一种新的用户兴趣模型建立[J].科学技术与工程,2009,7(8):47-49
- [3]蓝惠群.数字时代图书馆个性化信息服务研究[J].内蒙古科技与经济,2009,7(14):154-155
- [4]余小鹏.基于信息构建的电子商务网站搜索系统研究[J].情报科学,2011,5,29(5):778-781

[责任编辑 潘洪志]

(上接第48页)

现代连锁企业对组织创新不够重视,面对新的市场和经营环境,现代连锁企业必须调整现有的组织结构,充分利用现代信息技术,将企业信息化建设与新型组织结构结合起来,才能符合物联网环境对企业的要求。

综上所述,对物联网环境下连锁企业的组织模式进行了重新思考与定位,物联网时代企业的组织结构应该是扁平化、动态虚拟的,现代连锁企业的组织结构、管理模式、经营理念的创新是应用物联网技术的基础。应用物联网旨在为顾客带来更大的价值,降低企业的成本,流程型组织结构的目标与物联网时代的企业宗旨相吻合,对现代连锁企业应用物联网将起到推动作用。

[参 考 文 献]

- [1]岳彭,任浩,张军果.现代企业组织的发展趋势——流程型组织浅析[J].工业技术经济,2006
- [2]岳彭.流程型组织的构建研究[D].同济大学博士论文,2006
- [3]周萍.职能导向与流程导向组织绩效影响因素研究[J].山东大学学报,2008
- [4]谢守祥,周菲菲.信息时代下的无边界组织研究[J].江苏商论,2010
- [5]尹育新,戎宏娜.浅谈一种全新的网络技术——物联网[J].信息系统工程,2010
- [6]赵明地.信息化时代的企业组织创新[D].哈尔滨工程大学硕士论文,2003

[责任编辑 王凤娟]