

基于模糊兴趣模型与多 Agent 的个性化推荐系统^①

王晓堤¹ 文军舰¹ 张悦¹ 叶娟娟² (1.天津财经大学 管理信息系统系 天津 300222;

2.河北工业大学 管理学院 天津 300401)

摘 要: 提出了基于模糊兴趣模型与多 Agent 的个性化推荐系统框架,通过引入用户模糊兴趣模型,使以 Agent 为基础的推荐系统无法通过隐式收集用户对商品属性评价的问题得到解决,并且在客户端收集并挖掘用户的私有信息,然后从服务器中获取用户感兴趣的信息,最后生成并更新 UserProfile。

关键词: 模糊兴趣模型; 多 Agent; 个性化推荐; 隐私保护

Personalized Recommendation System Framework Based on Fuzzy Interest Model and Multi-Agent

WANG Xiao-Di¹, WEN Jun-Jian¹, ZHANG Yue¹, YE Juan-Juan²

(1.Business School, Tianjin University of Finance and Economics, Tianjin 300222, China;

2.School of Management, Hebei University of Technology, Tianjin 300401, China)

Abstract: The paper proposes a personalized recommendation system framework based on Fuzzy Interest Model and Multi-Agent to solve the recommendation system's inability to collect information from the customers' evaluation on the commodity's properties implicitly. The recommendation system creates and updates the UserProfile in accordance with customers' private information collected on the client and combined with information in which customers are interested in.

Keywords: fuzzy interest model; multi-agent; personalized recommendation; privacy protection

现有推荐系统在使用时,最重要的特点是能够实时查找、推荐商品。但随着用户和商品的增加,推荐系统算法的计算量会急剧增长,在数据量大且稀疏的情况下,导致处理能力低下、推荐准确度降低、系统性能严重下降。

近年来,为解决这类问题,学者们进行了大量研究,其中以 Agent 为基础的推荐系统得到了不少学者的重视,因为其具有智能性、自适应性、主动性等。文献[1]以多 Agent 系统为基础,建立一个基于偏好的评价模型,并在个性化的购物过程中,提出了一种基于 Agent 的决策分析模型。文献[2]针对电子商务的当前发展状况,结合多 Agent 系统的特点,提出了一种具有智能导购功能的电子商务系统。

但此类推荐系统存在的问题是:很少说明它们是何获取用户兴趣和偏好,使得以 Agent 为基础的推荐

系统存在很大局限性。本文提出以 Agent 为基础的推荐系统,在客户端隐式收集信息之后挖掘用户的兴趣和偏好,以此不断更新用户模糊兴趣模型,使得本系统在一定程度上解决了用户兴趣漂移和隐私保护的问题,并使用模糊兴趣模型提高了描述商品的准确度从而提高了推荐精度。

1 基于模糊兴趣模型的推荐系统模型

1.1 用户模糊兴趣的引入

用户兴趣模型是推荐系统产生个性化推荐的主要知识源,其捕捉用户真实偏好的能力在很大程度上决定了推荐的成功与否。

由于知道推荐系统的目标对象是用户,因而其本质都依赖于用户的兴趣。但传统的个性化推荐系统,其推荐过程是基于集合统计的原则,很少从用户兴趣

^① 收稿时间:2010-01-15;收到修改稿时间:2010-03-19

的本质考虑推荐模型, 这样的推荐难以做到真正的个性化。故本文专门建立了用户兴趣描述文件 **UserProfile**, 用于存储和更新用户的模糊兴趣模型。

1.2 用户模糊兴趣模型^[3]的建立

传统的推荐系统面临的问题是描述商品过于简单, 描述商品属性不够准确。以音乐推荐(流行, 古典, 摇滚, 校园, 大陆, 欧美)论域为例, 一首歌曲一般对应其中多项属性, 每一项的侧重都有所不同。通常的做法是用 0 或 1 进行表示, 但人们的喜好很难用喜欢或者不喜欢进行论断, 其本身是个模糊的概念, 故建立一个用户模糊兴趣模型很必要。具体过程如图 1 所示:

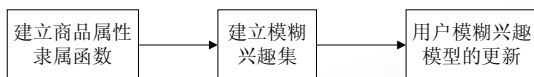


图 1 用户模糊兴趣模型的建立过程

1.2.1 商品属性隶属函数的建立

定义 2.1.^[4] 设 U 为论域, 则 U 上的一个模糊集合 A 由一个实值函数:

$$m_A = \begin{cases} U \rightarrow [0,1] \\ u \mapsto m_A(u) \end{cases} \quad (1)$$

来表示, 对于 $u \in U$, 函数值 $m_A(u)$ 成为 u 对于 A 的隶属度, 而函数 $m_A(u)$ 称为 A 的隶属函数。

建立商品属性隶属函数是模糊兴趣模型产生的第一步。传统上对商品属性给予是或否的简单回答将来自于 $[0, 1]$ 区间的一个模糊数所代替。同时, 这样的属性可能同时具有几个不同类型的值, 这大大提高了属性描述的准确度。

对于商品属性隶属函数的建立可以采用 **Delphi** 法。**Delphi** 法^[5]即德尔菲法是专家会议调查法的一种发展。它以匿名方式通过几轮函询, 征求专家意见。预测领导小组对每一轮意见进行汇总整理, 作为参考资料再发给每位专家, 供他们分析判断, 提出新的论证。如此多次反复, 专家意见趋于一致, 结论的可靠性越来越大。

设商品 u 的属性集论域为 $X = \{x_1, x_2, \dots, x_n\}$, 让 M 位专家分别对每个属性 x_i ($i=1, 2, \dots, n$) 给出一个隶属度的估计值, 设第 j 位专家给出的估计值为 $S_j(x_i)$ ($i=1, 2, \dots, n; j=1, 2, \dots, m$), 最后的隶属度可初步表示为:

$$S(x_i) = \frac{1}{m} \sum_{j=1}^m S_j(x_i) \quad (2)$$

Delphi 法也存在一些缺陷, 它并不能表示类如商品的新旧程度这样的随时间流逝而改变的量。对于这些属性, 我们可以使用函数表达法。

$$S_f(x_i) = F(x_i) \quad (3)$$

$S_f(x_i)$ 表示通过函数表达法得到的属性 x_i 的隶属度, x_i 为不能用(2)式来表达的属性项。最后, 商品 u 的隶属度可以用集合 $\{S'(x_i)\}$ 来表示。

$$S'(x_i) = \begin{cases} S(x_i) & x_i \text{ 可以用 Delphi 法表示} \\ S_f(x_i) & x_i \text{ 不能用 Delphi 法表示} \end{cases} \quad (4)$$

1.2.2 模糊兴趣集的建立

兴趣度就是用户对某一商品或某一类商品的兴趣的强弱, 它反应了用户对其的喜好程度。本文采用隐式方式收集用户兴趣度, 建立模糊兴趣集。用户购买行为会透露其兴趣, 用户购买记录集会反应用户的兴趣, 因此直接使用此集合构造用户兴趣集。

设 $U = \{u_1, u_2, \dots, u_l, \dots, u_p\}$ 为顾客的商品集, 其中: $u_l = (s_{l1}, s_{l2}, \dots, s_{li}, \dots, s_{ln})$ 。 s_{li} 即式(4)的结果, 它表示商品 u_l 对属性 x_i 的隶属度, 则此顾客的初级兴趣集可以表示为: $I = \{I_1, I_2, \dots, I_k, \dots, I_p\}$, 其中 I_k 表示该顾客对第 k 项属性的初级兴趣度:

$$I_k = \frac{\sum_{l=1}^p (s_{lk})^2}{p} \quad (5)$$

通过对商品属性喜好度和商品属性权重进行探讨, 建立完善的模糊兴趣模型。

1) 商品属性喜好度的获取

高阶属性喜好度可以看成一棵树, 商品喜好度的获取将从最高层向最底层逐级获取。用 IL 表示商品的属性喜好度, 但实际计算采取从底层到高层的方式。

2) 商品属性权重的获取

为获取有效的权重, 初期赋值再次采用 **Delphi** 法取得商品属性权重, 具体如下: 设商品 u 属性集为 $A = \{A_1, A_2, \dots, A_i, \dots, A_n\}$, 请 Q 位专家, 让每一个专家分别对 A_i 给出权重估计值, 记第 p 位专家给出估计值为 $W_p(A_i)$ ($i=1, 2, \dots, n; p=1, 2, \dots, Q$), 得最后的商品属性权重为:

$$I_Q(A_i) = \frac{1}{Q} \sum_{p=1}^Q W_p(A_i) \quad (6)$$

综合以上内容,利用 $\{I_L, I_Q(A_i)\}$ 表示用户模糊兴趣模型。 I_L 、 $I_Q(A_i)$ 分别表示商品属性喜好度和商品属性权重。在这个模型中,对于每个属性和量的描述都是模糊的,是一个 $[0, 1]$ 之间的数,用户的兴趣会包括多个商品属性,也可能分布在多个层次中。本文将模糊兴趣模型表示为用户兴趣描述文件(UserProfile),描述和存储用户的兴趣。

1.2.3 用户模糊兴趣模型的更新

本系统用 Web 数据挖掘的方法得到用户的浏览行为和访问模式,了解用户浏览了哪些商品、浏览时间、是否购买、购买的数量,这些数据来自 Web 日志文件,通过分析 Web 日志可以获得相关页面、用户访问模式等信息,对日志信息进行分析,将用户浏览或购买过某种商品所表现出的兴趣度的隐性信息转化为用户对该商品的显性评分数值。本个性化系统利用这些信息创建或更新 UserProfile。这样不断地进行 Web 挖掘并更新 UserProfile,在一定程度上减轻了用户的负担,同时在一定程度上解决了用户兴趣变化和漂移的问题[6,7]。

2 增加隐私保护的个性化服务

W3C 已经提出了隐私参数平台(Platform for privacy Preferences, P3P)[8]。用户在 P3P 提供的个人隐私保护策略下,能够清晰地明白网站对自己隐私信息做何种处理,并且 P3P 向用户提供了个人隐私信息在保护性上的可操作性。它允许服务器把它的隐私条款用机器可读的格式编码,比如信息将被如何收集,允许哪些人以什么目的访问这些数据等。

P3P 是一个扩展语法和数据元集,它受 RDF(Resource Description Framework, 资源描述框架)语言的规范。它允许用户通过浏览器与网站,通过临时或是成对的 ID(TUID/PUID)进行协商,以明确可以透露哪些个人信息并且作何用途。双方最终通过一个双向的选择达成用户个人隐私策略。

用户兴趣模型是个性化推荐系统的基础和核心,用户兴趣模型在个性化推荐系统中表现为 UserProfile。本文基于隐私保护的考虑,将 UserProfile 放置在客户端。

3 结合模糊兴趣模型和多Agent的个性化推荐系统模型实现模型

3.1 各 Agent 介绍

在研究提出的推荐模型中,用多个 Agent 通过协同合作完成系统推荐任务,如图 2 所示。整个推荐系

统内部模块结构可划分为用户交互 Agent、兴趣获取 Agent、推荐客户端 Agent 和推荐服务器四个部分,由这四部分共同分工协作,完成系统推荐任务。

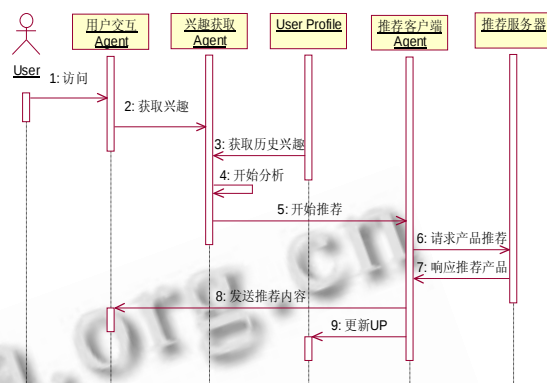


图 2 推荐系统顺序图

1) 用户交互 Agent 从用户使用系统过程中收集信息,进行隐式挖掘用户的兴趣。用户交互 Agent 负责用户与其他 Agent 交互。它一直处于活动状态,便可以实时处理用户的购物信息。当用户登录系统后,用户交互 Agent 将会把用户信息传给兴趣获取 Agent,而推荐客户端 Agent 则把商品的推荐请求发到推荐服务器。

2) 兴趣获取 Agent 提取并分析用户兴趣和历史数据,在客户端创建一个 UserProfile 用于表示用户模糊兴趣模型,并在以后用户访问过程中,进行隐式挖掘收集用户的信息,同时更新 UserProfile。首先,兴趣获取 Agent 从用户交互 Agent 获得用户信息和购物信息,然后从 UserProfile 调取相应数据传递到推荐客户端 Agent,并实时处理;当用户浏览或者购买完毕后,通过用户交互 Agent 将浏览或购买记录传递到兴趣获取 Agent,并进行隐式挖掘用户兴趣。

3) 推荐客户端 Agent 使用 UserProfile 选择并过滤推荐服务器发送的内容。客户端能够被部署在电脑或者其他客户端设备上,给用户有效的分类和过滤服务,收集用户的反馈给兴趣获取 Agent,使得 UserProfile 能动态的改进和更新。整个工作过程在客户端独立完成并且不需要服务器的干涉,因此本系统较好地处理了隐私保护问题。推荐客户端 Agent 是模型最重要的部分,它负责对商品候选集进行多属性决策评价和商品推荐。首先推荐客户端 Agent 从兴趣获取 Agent 和推荐服务器获取相应商品属性模糊兴趣集和候选商品集,然后利用 PROMETHEE II 法[9,10]进行多属性评价。最后对候选商品两两进行比较后进行降

序排列,之后进行 Top-N 推荐,并将结果传递到用户交互 Agent。

① 整体推荐算法如下:

```

Array Calculate_Top_N(set1 &userIn-
terest, set2 &productSet){
    for each product in ProductSet{
        if(userInterest.amount < lowerLimit &&
implicitMining() == NULL)
            //假如用户兴趣信息低于某个下限并且隐式挖掘也没有结果,则给出一般的 Top-N 推荐
            return orig_Top_N;
        else{
            if(implicitMining(log))userInterest.up
date(implicitMining());
            recommend(PROMETHEE_II(UserProfile));
            //根据 UserProfile 进行推荐
            returnDescendingSort(Top_N);
        }
    }
}

```

4) 推荐服务器管理商品内容并且根据客户端的请求给它们发送商品内容,并把要推荐的商品内容组织成分类树。

4 实验与分析

本文以电影推荐为例,收集 100 位用户的 3000 条电影观看记录,用 MAE(Mean Absolute Error)比较本文算法和协同过滤算法推荐性能。实验结果如图 3 所示。

由于本文是用模糊兴趣模型来描述用户兴趣,所以描述用户兴趣更准确,从图 3 可以看出本文算法无论在数据稀疏时还是在数据丰富时,准确度都要高于协同过滤算法,故本文提出的算法是有效的。

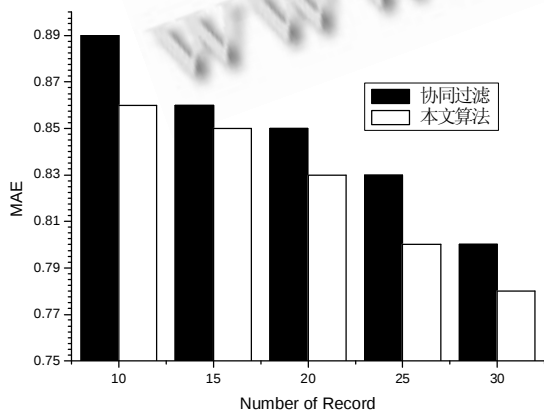


图 3 推荐准确度示意图

5 小结

本文提出了基于模糊兴趣模型与多 Agent 的个性化推荐系统框架,通过引入用户模糊兴趣模型,使以 Agent 为基础的推荐系统无法通过隐式收集用户对商品属性评价的问题得到解决,并且在客户端收集并挖掘用户的私有信息,生成、存储并更新 UserProfile,然后再从服务器中获取用户感兴趣的信息。由于对用户私有信息相关的操作都在客户端执行,用户隐私会得到一定的保护,并可将其应用于电影、音乐、书籍以及其它频繁购买商品推荐上。

参考文献

- 1 曾子明.基于 Agent 和 CBR 的电子商务推荐模型的研究.现代情报,2008,(3):209-213.
- 2 陈德军,李婷.基于多 Agent 的智能电子商务系统研究.计算机与数字工程,2007,35(11):41-43.
- 3 唐灿,朱征宇.基于模糊兴趣模型的个性化推荐算法.计算机工程与应用,2006,(9):166-168.
- 4 杨纶标,高英仪.模糊数学原理及应用.4th ed.,广州:华南理工大学出版社,2006:69-82.
- 5 谭跃进.定量分析方法.北京:中国人民大学出版社,2006:70-72.
- 6 Widmer G, Kubat M. Learning in the Presence of Concept Drift and Hidden Contexts. Machine Learning, 1996,23(1):69.
- 7 Dan ZG, Cai LN, Zheng L. Improved Multi-Agent System for the Vehicle Routing Problem with Time Windows. TSINGHUA SCIENCE AND TECHNOLOGY, 2009, 14(3):407-412.
- 8 Lorrie Cranor, Rigo Wenning.Platform for Privacy Preferences, 2007. <http://www.w3.org/P3P/>.
- 9 Brans JP, Vincke P. A preference ranking organization method: the PROMETHEE Method. Management Science,1985, 31:647-656.
- 10 Brans JP, Vincke P, Mareschal B. How to select and how to rank projects: The PROMETHEE method. Euro.J.Opl.Res.1986,24:228-238.