

# 电子商务推荐系统关键技术研究

曹毅<sup>1</sup>, 罗新星<sup>2</sup>

(1. 湖南工学院, 湖南 衡阳 421101; 2. 中南大学 商学院, 湖南 长沙 410083)

**摘 要:** 简要介绍了电子商务推荐系统的概念, 着重讨论了基于协同过滤的推荐技术、基于内容的推荐技术、基于效用的推荐技术、基于知识的推荐技术和基于用户统计的推荐技术, 并对这些推荐技术进行了对比分析, 最后提出了适合我国电子商务发展的组合推荐策略。

**关键词:** 电子商务; 数据挖掘; 推荐系统; 个性化推荐

**中图分类号:** TP311

**文献标识码:** A

**文章编号:** 1672-8173(2008)05-0063-04

## 0 引言

随着网络技术的发展以及机器学习、数据挖掘等知识发现技术的出现, 电子商务的发展模式对企业服务提出了许多新要求, 信息服务方式也从传统的“一对多”发展到“一对一”的个性化用户服务, 即为恰当的用户在恰当的时间方便快捷地提供恰当的信息<sup>[1]</sup>。由于供应链和物流的发展, 在电子商务的虚拟环境下, 商家能够在网上提供的商品种类和数量非常多, 但用户不可能通过一个小小的计算机屏幕很方便地发现自己感兴趣的物品, 用户既不愿意花太多时间在漫无边际的网上寻找商品, 也不可能像在物理环境下那样检查商品的质量<sup>[2]</sup>。

因此, 用户很希望电子商务系统具有一种类似采购助手的功能来帮助其选购商品, 并能根据用户的兴趣爱好自动地推荐给每个用户可能感兴趣且满意的商品。在这种环境下, 电子商务推荐系统应运而生, 它根据用户的兴趣爱好, 推荐符合用户兴趣爱好的对象, 这样就可以大大缩短用户在网络上的访问延迟, 最大限度的提高网络信息服务质量。

## 1 电子商务推荐系统

电子商务推荐系统是“利用电子商务网站向客户提供商品信息和建议, 帮助用户决定应该购买什么产品, 模拟销售人员帮助客户完成购买过程”<sup>[3]</sup>的自动化系统, 也称为个性化推荐系统。在传统商业销售中, 商场通过销售人员向顾客推荐商品; 但在电子商务环境下, 关于商品和客户的信息都是海量的, 人工处理既不现实也不经济。此时, 电子商务推荐系统扮演传统商业中销售人员的角色: 向顾客提供商品信息和建议, 帮助顾客完成购买过程。电子商务推荐系统最大的优点在于它能收集用户感兴趣的资料, 并根据用户兴趣偏好主动为用户做出个性化推荐。电子商务推荐系统从本质上要解决三个问题: 一是采用什么样的数据源; 其次是建立什么样的数据模型; 最后是推荐策略问题, 即用什么方法进行推荐<sup>[4]</sup>。

## 2 个性化推荐技术

### 2.1 基于协同过滤的推荐技术

基于协同过滤的推荐是目前研究最多的个性化推荐技术, 它根据其他用户的观点产生对目标用户的推荐

收稿日期: 2008-06-10

基金项目: 湖南省软科学研究计划项目(04ZH6005); 湖南省普通高校教学改革研究项目(2006-191); 湖南教育厅科学研究项目(07CX33)

作者简介: 曹毅(1977-), 男, 湖南长沙人, 硕士, 研究方向: 网络安全、电子商务。

列表,推荐的个性化程度高.在协同过滤推荐系统中,用户描述的典型方法是采用以商品及其评价为分量的向量来表示,向量将随着用户与系统交互时间的增加而不断增大.协同过滤推荐的核心思想是认为用户会倾向于购买具有相似意向的用户群所购买的商品,因而它在预测某个用户的商品购买倾向时是根据一个用户群的情况而决定的.现在著名的基于协同过滤的推荐系统有 Group Lens/ Net Perceptions<sup>[5]</sup>、Tapestry<sup>[6]</sup>等.协同过滤的最大优点是对推荐对象没有特殊要求,能处理非结构化的复杂对象,如音乐、电影.协同过滤推荐主要分为三类:一是最近邻协同过滤推荐;二是基于项目的协同过滤推荐;三是基于模型的协同过滤推荐.

最近邻协同过滤推荐是当前最成功的推荐技术,它基于这样一个假设<sup>[7]</sup>:如果用户对一些项目的评分比较相似,则他们对其它项目的评分也比较相似.协同过滤推荐系统使用统计技术搜索目标用户的若干最近邻居,然后根据最近邻居对项目的评分预测目标用户对项目的评分,产生对应的推荐列表.目标用户的最近邻居查询是否准确,直接关系到整个推荐系统的推荐质量,准确查询目标用户的最近邻居是整个协同过滤推荐成功的关键<sup>[8]</sup>.

基于项目的协同过滤推荐根据用户对相似项目的评分预测该用户对目标项目的评分,它基于这样一个假设<sup>[9]</sup>:如果大部分用户对一些项目的评分比较相似,则当前用户对这些项目的评分也比较相似.基于项目的协同过滤推荐系统使用统计技术找到目标项目的若干最近邻居,由于当前用户对最近邻居的评分与对目标项目的评分比较类似,所以可以根据当前用户对最近邻居的评分预测当前用户对目标项目的评分,产生对应的推荐列表.

基于模型的协同过滤推荐,先用历史数据得到一个模型,再用此模型进行预测<sup>[10]</sup>.基于模型的推荐广泛使用的技术包括神经网络等学习技术、潜在语义检索和贝叶斯网络,训练一个样本得到模型.

## 2.2 基于内容的推荐技术

基于内容的推荐是信息过滤的继续和发展,商品通过相关特征的属性来定义,系统基于用户评价对象的特征学习用户的兴趣,依据用户资料与待预测项目的匹配程度进行推荐.基于内容的推荐需要进行匹配计算,较多地应用于可计算的文本领域,如浏览页面的推荐、新闻组中的新闻推荐等,其缺点是不能为用户发现新的感兴趣的信息,只能推荐与用户已有兴趣相似的信息.例如 NewsWeeder 新闻组过滤的文本推荐系统就是采用文本中的单词作为文本的特征<sup>[11]</sup>,系统通过学习用户已评价过的商品特征来获得对用户兴趣的描述.

用户描述类型的产生取决于系统所采用的学习算法,判定树、神经网络、基于向量的表示等技术都可应用于其中.基于内容推荐的用户描述是长效型的,它将随着系统对用户偏好的学习而不断更新,Schafer 等称之为“商品与商品的相关性推荐”<sup>[12]</sup>.

## 2.3 基于效用的推荐技术

基于效用的推荐建立在用户需要和可选集之间匹配的评估之上,通过计算商品对用户的效用来做出推荐,其核心问题是如何为每一位用户创建出合适的效用函数.基于效用的推荐其用户描述就是系统为用户创建的效用函数,采用受限制满足技术来确定最佳匹配,它的优点是能在效用函数中考虑如提供商的可靠性、产品的可获得性和可用性等非产品因素<sup>[13]</sup>.

## 2.4 基于知识的推荐技术

基于知识的推荐是通过判断用户的需要来做出推荐,它具有特定商品满足特定用户需要的知识,并由此推导出用户需要与某一推荐物品的相互关系.基于知识的推荐其用户描述可以是支持这种推导的任何知识结构,各方法因所用的知识的不同而有明显的区别<sup>[14]</sup>.例如在 Google (<http://www.google.com>) 中,用户描述仅仅是用户所构造的查询语句.系统使用的知识也可以有多种形式,如 Google 就是使用 Web 页面间的链接信息来推断流行度和权威性,而 Persona2Logic 是使用了关于功能的知识来计算效用函数的.

## 2.5 基于用户统计的推荐技术

基于用户统计的推荐是根据用户个人属性对用户分类,再基于类对类中的用户统计信息进行推荐<sup>[11]</sup>.用户个人信息可通过交互式的对话来收集,可采用机器学习来建立一个基于用户统计信息的分类器,用户的反应与人工创建的模式库相匹配.基于用户统计的推荐系统与协同过滤推荐系统实际使用的数据完全不同,其优点在于不需要用户评价历史数据.

# 3 推荐技术的比较研究

基于协同过滤的推荐依赖于用户间评价的重叠,故当只有较少用户对同一件商品做出评价时就会出现

“评价稀少”,特别是在新系统的开发初期,会有大量商品出现,如果用户基群不够大,用户间能共享的商品评价将会很少.故协同过滤推荐最适应于用户兴趣明显、用户兴趣密度相对较高而商品数量相对较少且恒定的情况,而不适合于“灰色绵羊”情况,即一位用户游离于不同用户群体之间<sup>[13]</sup>.

基于内容的推荐需要积累足够数量的评价才能构建出一个可靠的分类器,与基于协同过滤的推荐相似,基于内容的推荐也因与所推荐的商品紧密相关而存在不足.例如,基于内容的电影推荐系统必须依赖于为每一部电影所编写的作者名称、情节概述等描述资料,因为电影本身并不能被系统所理解.

基于效用和基于知识的推荐系统并不是基于用户数据的累积,所以都没有新用户问题(即对于做出较少评价的新用户难以对其进行分类)、新商品问题(即对于做出较少评价的新商品难以对其进行推荐)和“评价稀少”问题.

基于效用的推荐需要系统构建一个能涵盖所考虑商品全部特征的效用函数,它的优势是能综合考虑决定商品价值的各种特征因素,包括递送计划、担保条款等等,而不仅仅是产品具体的性能特点,因此,基于效用的系统会让用户表达出所有的考虑因素,以做出推荐.用户要构建一个完整的兴趣效用函数,就必须衡量每一个因素的权重,这就为用户与系统间的交互增加了额外的负担.一个完整的显式效用函数会对有特殊购买需求的专业用户有很大帮助,但对一位具备较少相关知识的临时用户就会增加麻烦,所以,基于效用的推荐系统不适合于临时浏览用户.

基于知识推荐的重点和难点是需要获取知识,它会涉及三种知识:一是目录知识,即所推荐商品及其特征的知识;二是功能知识,即系统应当能将用户的需求和可能满足该需求的商品的特征进行比较;三是用户知识,为了提供优质的推荐服务,系统必须要有一些关于用户的知识.基于知识推荐的优势在于它对用户的要求较少,因此适用于用户临时随机浏览的情况,没有新系统刚开始时常有的低质量推荐问题.

基于用户统计的推荐同样存在“灰色绵羊”情况,但它没有新用户问题,因为此系统不需要用户对商品的评价.但是,随着个人隐私问题越来越敏感,搜集必需的用户统计信息将是本推荐系统所面临的问题,用户可能拒绝绝对系统透露或虚构用来预测用户偏好的、极为关键的个人信息.

所有推荐技术都有优点和缺点,如表 1 所示.

表 1 推荐技术的比较研究

推荐技术	优 点	缺 陷
基于协同过滤的推荐	1.能跨类型推荐; 2.不需要领域知识; 3.随时间的推移,系统性能提高; 4.推荐的个性化、自动化程度高,自适应好.	1.“灰色绵羊”问题; 2.新用户问题、新商品问题; 3.推荐质量依赖于历史数据集.
基于内容的推荐	1.不需要领域知识; 2.推荐结果直观,易理解;	1.新用户问题、新商品问题; 2.复杂属性不好处理; 3.要有足够的数据构造分类器.
基于效用的推荐	3.随时间的推移,系统性能提高.	1.用户必须输入效用函数;
基于知识的推荐	1.不存在新用户问题和新商品问题; 2.对用户偏好的变化敏感; 3.能考虑非产品特性.	2.属性重叠问题; 3.静态推荐,不能学习,灵活性差.
基于用户统计的推荐	1.不存在新用户问题和新商品问题; 2.能考虑非产品特性; 3.能比较、匹配用户需求和产品特征.	1.知识的获取困难; 2.静态推荐,不能学习,灵活性差.
	1.不存在新商品问题;2.不需要领域知识.	1.“灰色绵羊”问题; 2.推荐质量依赖于历史数据集; 3.用户统计信息得到困难.

4 组合推荐策略

由于各种推荐方法都有优缺点,所以在实际中常采用组合推荐.研究和应用最多的是内容推荐和协同推荐的组合<sup>[15]</sup>.最简单的做法是分别用基于内容的方法和协同推荐方法,产生一个推荐预测结果,然后用某方法组合其结果.例如,利用用户评价数据得到的文档/词矩阵产生一个基于内容用户资料的矩阵,通过潜在语义索引计算一个基于内容用户资料的排序形式,然后加权用户档案中的词义向量产生推荐.

尽管理论上有很多推荐组合方式,但在某一具体问题中并不见得都有效,组合推荐一个最重要的原则,就是通过组合后要能避免或弥补各自推荐技术的弱点,如对于基于内容和协同过滤的技术,不论如何组合,总是存在新用户、新商品问题,因为它们都要求有一个历史的评价数据;但若把基于协同过滤与基于知识或基于效用的技术相结合,则能有效解决这些问题,因为后两种技术不需要许多用户历史数据.所以,在组合方式上有以下思路<sup>[16]</sup>:

- (1)加权式:对多个推荐方法的计算结果加权组合而成;
- (2)混合式:同时为用户提供多种推荐方法的结果;
- (3)变换式:根据问题背景和实际情形采用不同的推荐方法;
- (4)特征组合式:组合来自不同推荐数据源的特征被另一种推荐方法所采用;
- (5)层叠式:先用一种推荐方法产生一种粗糙的推荐结果,第二种推荐方法在此推荐结果的基础上进一步作出更精确的推荐;
- (6)特征递增式:第一种推荐方法产生附加的特征信息嵌入到其他推荐方法的特征输入中;
- (7)元层次式:用前一种推荐方法所产生的模型作为后一种推荐方法的输入.

## 5 结束语

通过推荐系统实现个性化服务已成为电子商务应用的一项新兴技术,尽管国外目前已取得了很好的研究成果,但还存在很多问题,如数据获取主要依赖用户的显式评价,在自动获得用户的隐式信息方面做得也不够等等.目前,我国电子商务的推荐功能相对国外存在较大的差距,在个性化推荐和自动推荐方面的理论研究刚刚起步,对协同推荐研究则更少,严重影响我国电子商务的发展.推荐系统在电子商务系统中具有良好的发展和应用前景,逐渐成为电子商务 IT 技术的一个重要研究领域,得到了越来越多研究者的关注.

## 参考文献:

- [1] 曹毅.基于内容和协同过滤的混合模式推荐技术研究:[硕士论文].长沙:中南大学,2007.
- [2] 余力,刘鲁.电子商务个性化推荐研究[J].计算机集成制造系统,2004,10(10):1306-1313.
- [3] Resnick and Varian. Recommender systems[J]. Communications of the ACM, 1997, 40(3):56-58.
- [4] 宁彬,袁磊.电子商务推荐系统的应用[J].商场现代化,2007,(9):95.
- [5] Resnick P, Iacovou N, Suchak M, et al. Group lens: an open architecture for collaborative filtering of netnews[A]. Proceedings of the Conference on Computer Supported Cooperative Work[C]. Chapel Hill, NC, 1994, 175-186.
- [6] Goldberg, Nichols D, Okib M, et al. Using collaborative filtering to weave an information apestry[J]. Communications of the ACM, 1992, 35(12):61-70.
- [7] Breese J, Hecherman D, Kadie C. Empirical Analysis of Predictive Algorithms for Collaborative Filtering[A]. Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence (UAI-98)[C]. 1998, 43-52.
- [8] 温会平,陈俊杰.基于用户模糊聚类的个性化推荐算法[J].计算机与数字工程,2008,36(2):13-16.
- [9] Sarwar B, Karypis G, Konstan J and Riedl J. Item-based collaborative filtering recommendation algorithms[C]. In: Proceedings of the Tenth International World Wide Web Conference, 2001, 285-295.
- [10] Ben J, Konstan J A, John R. E-commerce recommendation applications[R]. University of Minnesota, 2001.
- [11] Lang K. Newsweeder: learning to filter news[A]. Proceedings of the 12th International Conference on Machine Learning[C]. 1995, 331-339.
- [12] Schafer J B, Konstan J, Riedl J. Recommender Systems in E2Commerce[A]. Proc of ACM Conf on Electronic Commerce[C]. 1999, 158-166.
- [13] 黎星星,黄小琴,朱庆生.电子商务推荐系统研究[J].计算机工程与科学,2004,26(5):7-10.
- [14] 刘平峰,聂规划,陈冬林.基于知识的电子商务智能推荐系统平台设计[J].计算机工程与应用,2007,43(19):199-201.
- [15] Pazzani M J. A framework for collaborative, contentbased and demographic filtering[J]. Artificial Intelligence Review, 1999, 13(5/6):393-408.
- [16] Robin B. Hybrid recommender systems: survey and experiments[R]. Department of Information Systems and Decision Sciences, California State University, Fullerton.

(下转第 74 页)

- [7] H Kasahara, S Narita. Practical multiprocessing scheduling algorithms for efficient parallel processing[J]. IEEE Transactions on Computers, 1984,33(11):1023 - 1029.
- [8] J H Holl and Adaptation in Natural and Artificial Systems[M]. MIT Press, Cambridge, MA, USA, 1992.
- [9] J Levine and F Ducatelle. Ant colony optimization and local search for bin packing and cutting stock problems[M]. Journal of the Operational Research Society. 2003.
- [10] G Ritchie. Static multi-processor scheduling with ant colony optimisation and local search[M]. Master's thesis, University of Edinburgh. 2003.
- [11] 吴 斌,史忠植.一种基于蚁群算法的 TSP 问题分段求解算法[J].计算机学报,2001,24(12):1328 - 1333.
- [12] 吴庆洪,张纪会,徐心和.具有变异特征的蚁群算法[J].计算机研究与发展,1999,36(10):1240 - 1245.

## A Hybrid Approach Based on Ants Colony Optimization (ACO) and Genetic Operators

*LI Sheng-xin<sup>1,2</sup>*

- (1. Department of Computer Science, Xiangnan University, Chenzhou 423000, China;  
2. School of Computer and Communication, Hunan University, Changsha 410082, China)

**Abstract:** Despite the numerous applications of ACO(ant colony optimization) algorithm in optimization computation, it remains a computational bottleneck that the ACO algorithm costs too much time in order to find an optimal solution for large-scaled optimization problems. The objective is to find an optimal schedule which will minimize the makespan so as to ensure the proper utilization of the resources and minimize the delay in achieving user specified deadlines. A hybrid approach based on ACO and Genetic operators is proposed for resource allocation and scheduling in computational grids.

**Key words:** grid scheduling; computational grids; genetic operators; ants colony optimization

---

(上接第 66 页)

## Research on Recommending System of Primary Technologies in E-commerce

*CAO Yi<sup>1</sup>, LUO Xin-xing<sup>2</sup>*

- (1. Hunan Institute of Technology, Hengyang 421101, China;  
2. School of Business Central South University, Changsha 410083, China)

**Abstract:** Concept of E-commerce recommending system is briefly introduced at the beginning of the paper. And then five main recommendation technologies such as collaborative filtering recommendation, recommendation based on contents, utility, knowledge and demographic are mentioned. Analysis and comparison among these above-mentioned technical recommendations are also provided. The combined recommendation strategy suitable for the development of E-commerce in China is presented at last.

**Key words:** E-commerce; data mining; recommending system; personalized recommendation

作者: [曹毅](#), [罗新星](#), [CAO Yi](#), [LUO Xin-xing](#)  
作者单位: [曹毅, CAO Yi \(湖南工学院, 湖南, 衡阳, 421101\)](#), [罗新星, LUO Xin-xing \(中南大学, 商学院, 湖南, 长沙, 410083\)](#)  
刊名: [湘南学院学报](#)  
英文刊名: [JOURNAL OF XIANGNAN UNIVERSITY](#)  
年, 卷(期): 2008, 29 (5)  
被引用次数: 1次

## 参考文献(16条)

1. [曾毅](#) [基于内容和协同过滤的混合模式推荐技术研究](#) 2007
2. [余力](#); [刘鲁](#) [电子商务个性化推荐研究](#)[期刊论文]-[计算机集成制造系统-CIMS](#) 2004 (10)
3. [Resnick](#); [Varian](#) [Recommender systems](#) 1997 (03)
4. [宁彬](#); [袁磊](#) [电子商务推荐系统的应用](#)[期刊论文]-[商场现代化](#) 2007 (09)
5. [Resnick P](#); [Lacovou N](#); [Suchak M](#) [Group lens: an open architecture for collaborative filtering of netnews](#) 1994
6. [Coldbergd](#); [Nichols D](#); [Okib M](#) [Using collaborative filtering to weave information apestry](#) 1992 (12)
7. [Breese J](#); [Hecherman D](#); [Kadie C](#) [Empirical Analysis of Predictive Algorithms for Collaborative Filtering](#) 1998
8. [温会平](#); [陈俊杰](#) [基于用户模糊聚类的个性化推荐算法](#)[期刊论文]-[计算机与数字工程](#) 2008 (02)
9. [Sarwar B](#); [Karypis G](#); [Konstan J](#); [Riedl J](#) [Item-based collaborative fitering recommendation algorithms](#) 2001
10. [Ben J](#); [Komtan J A](#); [John R](#) [E-commerce recommendation applications](#) 2001
11. [Lang K](#) [Newsweeder: learning to filter](#) 1995
12. [Schafer J B](#); [Konstan J](#); [Riedl J](#) [Recommender Systems in E2Commerce](#) 1999
13. [黎星星](#); [黄小琴](#); [朱庆生](#) [电子商务推荐系统研究](#)[期刊论文]-[计算机工程与科学](#) 2004 (05)
14. [刘平峰](#); [聂规划](#); [陈冬林](#) [基于知识的电子商务智能推荐系统平台设计](#)[期刊论文]-[计算机工程与应用](#) 2007 (19)
15. [Pazzani M J](#) [A framework for collaborative, contentbased and demographic filtering](#) 1999 (5/6)
16. [Robin B](#) [Hybrid recommender systems: survey and experiments](#)

## 本文读者也读过(10条)

1. [胡炜](#). [HU Wei](#) [协同过滤算法在推荐系统中的应用](#)[期刊论文]-[计算机时代](#) 2009 (11)
2. [顾桢](#) [浅析电子商务中的WEB数据的挖掘](#)[期刊论文]-[商场现代化](#) 2008 (16)
3. [徐春选](#). [姜伟](#). [XU Chun-xuan](#). [JIANG Wei](#) [个性化推荐技术在电子商务中的实现](#)[期刊论文]-[电脑知识与技术](#) 2008, 2 (16)
4. [张海鹏](#). [李烈彪](#). [李仙](#). [周亚峰](#). [Zhang Haipeng](#). [Li Liebiao](#). [Li Xian](#). [Zhou Yafeng](#) [基于项目分类预测的协同过滤推荐算法](#)[期刊论文]-[情报学报](#) 2008, 27 (2)
5. [杨帆](#). [姜建国](#). [YANG Fan](#). [JIANG Jian-guo](#) [基于客户偏好的页面隐式电子商务推荐技术探究](#)[期刊论文]-[电脑知识与技术](#) 2008, 2 (11)
6. [龚松杰](#) [个性化推荐中一种新的相似性计算方法](#)[期刊论文]-[计算机系统应用](#) 2008, 17 (7)
7. [龚松杰](#) [电子商务中协同过滤推荐技术研究](#)[期刊论文]-[商场现代化](#) 2008 (3)
8. [查文琴](#). [梁昌勇](#). [曹锺](#). [ZHA Wen-qin](#). [LIANG Chang-yong](#). [CAO Lei](#) [基于用户聚类的协同过滤推荐方法](#)[期刊论文]-

[计算机技术与发展](#)2009, 19(6)

9. [曾子明](#) [基于Agent和CBR的电子商务推荐系统模型研究](#)[期刊论文]-[现代情报](#)2008, 28(3)

10. [刘旭东](#) [个性化网页推荐系统在电子商务中的设计与应用](#)[期刊论文]-[烟台职业学院学报](#)2008, 14(4)

#### 引证文献(1条)

1. [蒋宗礼](#), [陆晨](#) [基于级联二部图的动态推荐算法](#)[期刊论文]-[计算机工程与设计](#) 2013(12)

引用本文格式: [曹毅](#), [罗新星](#), [CAO Yi](#), [LUO Xin-xing](#) [电子商务推荐系统关键技术研究](#)[期刊论文]-[湘南学院学报](#) 2008(5)