

基于知识网格的电子商务推荐系统设计^{*}

陈冬林, 聂规划, 刘平峰

(武汉理工大学 经济学院 电子商务系, 湖北武汉 430070)

摘要: 提出电子商务网站、知识网格节点和虚拟知识网格服务社区三级结构模型, 建立自动实现商品知识获取、聚合和智能推荐的运作机制, 设计了基于知识网格的电子商务智能推荐系统结构。

关键词: 推荐系统; 知识网格; 语义本体论; 电子商务

中图法分类号: TP311 **文献标识码:** A **文章编号:** 1001-3695(2006)12-0307-03

E-commerce Recommendation Systems Based on Knowledge Grid

CHEN Dong lin NIE Guihua LIU Ping feng

(Dept of E-commerce, College of Economics Wuhan University of Technology Wuhan Hubei 430070 China)

Abstract An explanation is introduced term KGRS (Knowledge Grid-based Recommendation Systems), which is designed to support large scale E-commerce applications requiring highly accurate recommendations in real time

Key words Recommendation System; Knowledge Grid; Ontology; E-commerce

1 引言

电子商务推荐系统可以向客户提供商品推荐, 帮助客户找到所需的商品, 满足客户个性化的需求。目前几乎所有的大型电子商务网站均不同程度地使用了电子商务推荐技术。国内外在电子商务推荐技术和系统的研究方面已取得较多的理论和应用成果。目前国内外的电子商务推荐研究工作主要有以下几个方面:

(1) 推荐质量实时性和质量研究。大部分推荐系统在保证实时性要求的同时, 是以牺牲推荐系统的质量为前提的^[1-2]。现有推荐技术主要是基于内容的推荐技术和协同过滤技术。前者最大的缺点是必须分析资源的内容信息, 且不能对音乐、图像、视频等进行推荐^[3-4]。而协同推荐技术是当前研究的热点, 其最大优点是不需要分析对象的特征属性^[5], 如余力等人提出的客户多兴趣下的个性化推荐算法^[6]; Soe Tsy Yuan提出了一种基于语义本体论的耦合聚类算法(OBPC), 它解决了现有协同过滤算法只能应用于具有历史商品评分记录的局限性问题^[7]。协同过滤技术也存在自身的不足, 主要表现在客户对商品评分数据比较少的时候, 推荐质量比较低。为解决上述两种技术存在的不足, 人们提出基于知识的推荐技术, 它不依赖客户对商品的评分数据量, 因此可以提高推荐的质量。Bin Xiao等人采用基于案例推理的推荐技术 CBR (Case Based Reasoning), 设计出了电子商务产品智能推荐系统 PCFinder^[8]。基于知识的推荐技术研究的难点是知识的获取。

(2) 推荐系统规模性研究。到目前为止, 绝大多数电子商务推荐技术及系统研究均是基于特定网站的, 不能满足大规模网站的推荐应用。文献[9]提出了一种电子商务虚拟社区推荐系统, 它是假设在某一电子商务社区, 存在相近商品的一群电子

商务网站, 提出了基于智能代理的社区推荐系统。Cynus Shahabi等人综合基于内容的过滤和协同过程过滤两种推荐技术的优点, 设计出了离线训练模型和在线使用模型进行推荐的模式——Yoda提高了基于Web的推荐系统的质量和规模性^[10]。

(3) 网格技术下的电子商务推荐系统研究。国外刚开始基于信息网格的推荐算法和知识获取的研究工作。Lan M等人提出了电子商务推荐系统的信息网格结构^[11], 采用了基于内存的多代理节点结构来实现, 能减少客户在线的信息交互量。但该方法在推荐质量、推荐深度和大规模应用方面无显著提高。文献[12]提出了基于本体论的知识逻辑和多Agent结构模型(KS-Net), 用于解决网格条件下的用户知识获取和聚合问题。

知识网格是一个智能互连环境, 它能使客户有效地获取、发布、共享和管理知识资源, 并为客户提供所需要的知识服务, 辅助实现知识创新、协同工作、问题解决和决策支持。2001年5月, Web之父Tim Berners-Lee正式提出了语义网的概念。为了实现知识的共享和重用, 在语义网研究中引入本体技术, 且正在被不断地实践。如果广泛采用语义本体论技术, 那么精细、准确和自动化的搜索及智能推荐就能够实现^[13-15]。

国内外在基于语义本体论的知识网格应用研究方面有一定的进步。中国科学院计算技术研究所诸葛海研究员领导的知识网格研究小组已经初步研究出了知识网格平台、语义链建造工具和智能浏览器的原型; 中国E-science知识网格环境IMAGINE-1将知识网格的理念付诸实施, 它为科学研究活动提供了知识共享的平台^[16]; 吴朝晖等人结合语义本体论和知识网格技术, 开发了基于知识网格的中医药信息共享与服务平台TCM, 该平台建立TCM本体库、TCM KB Grid知识网格模型和知识网格服务结构, 提供智能化中医药信息与知识服务^[17, 18]。

综上所述, 现有电子商务推荐技术和系统实质上只适用于单个电子商务网站, 远远不能满足网络条件下大规模电子商务推荐的需求。目前我国电子商务推荐技术和系统的研究相对

收稿日期: 2005-09-29 修返日期: 2005-11-27

基金项目: 国家自然科学基金资助项目(70572079)

国外差距较大^[18]。结合知识网格技术、语义本体理论和电子商务推荐技术,设计了基于知识网格的电子商务智能推荐系统(Knowledge Grid based Recommendation System, KGRS),可以满足网格条件下商品知识有效获取、聚合和智能推荐的要求,并提高电子商务推荐系统的规模性、质量和实时性。

2 系统原理与结构

2.1 基本原理

采用电子商务网站、知识网格节点和虚拟知识网格服务社区三级结构模型,建立自动实现商品知识获取、聚合和智能推荐的运作机制。

其实现原理和动作机制为:①以电子商务网站为商品信息源,由最相邻的知识网格负责进行从这些网站获取和聚合知识;②对于客户提出的需求,结合客户知识库,首先在电子商务推荐系统的知识网格服务目录库中寻找最相关的虚拟知识网格服务社区,如果社区中商务知识不足或社区过小,可启动知识网格智能搜索子系统,采用最相邻知识网格搜索和过滤技术,重新动态形成社区,并将结果存储于知识网格目录服务器和知识库中;③智能推荐系统在虚拟知识网格服务社区中调用智能推荐子系统,进行商品知识与客户知识的匹配。

2.2 智能推荐系统结构设计

该系统由以下四个层次组成(图1):

(1)外部接口层。该层提供与客户的交互接口;提供数据库、本体库管理界面;提供与电子商务计算网格的接口;实现与商品知识网格的互连。

(2)智能推荐与质量评估层。该层主要完成任务监视与负载均衡、智能推荐、推荐质量评估和结果解释。任务监视与负载均衡器监视人机接口提交的活动,进行推荐活动任务的智能负载均衡,负责与计算网格接口交互,请求电子商务计算网格的协同计算;推荐质量评估和结果解释器,进行推荐质量的评估、负责向客户解释推荐产生的原因。智能推荐引擎根据客户需求或个性化电子商务服务的要求,从虚拟知识网格服务社区的知识库选择合适的知识进行智能推荐,如果知识不足或社区过小,则调用下层的知识网格智能搜索子系统进行相关知识的获取和聚合。

(3)知识集成与功能库管理子系统。该系统负责在网上搜索最相邻知识网格,完成知识网格的商品知识集成,管理知识网格目录,管理数据库和本体库。知识网格智能搜索子系统采用最相邻知识网格搜索和过滤技术,通过规划商品知识供应链和恰当地实施知识供应,主动发现商品知识需求和商品知识资源,并确保合适的知识闭包(即最小完备知识集)。该子系统具有自学习和自更新的能力,实现基于语义的资源集成服务(包括客户知识集成、商品知识集成、推荐模型集成),收集相关资源,并形成匹配多种客户需求的完整内容。统一的资源管理模型有助于实现知识服务的动态聚合,并消除不同知识网格之间存在的商品知识冗余;知识网格目录服务子系统统一管理KGRS知识网目录资源。

(4)电子商务智能推荐系统底层本体库及知识库。它是由商品本体库(包括客户数据、客户本体、商品本体)、知识网格服务目录库组成的电子商务智能推荐知识库。电子商务智

能推荐商品知识库存储从知识网格中获取和聚合的商品知识、客户知识以及推荐技术模型等。

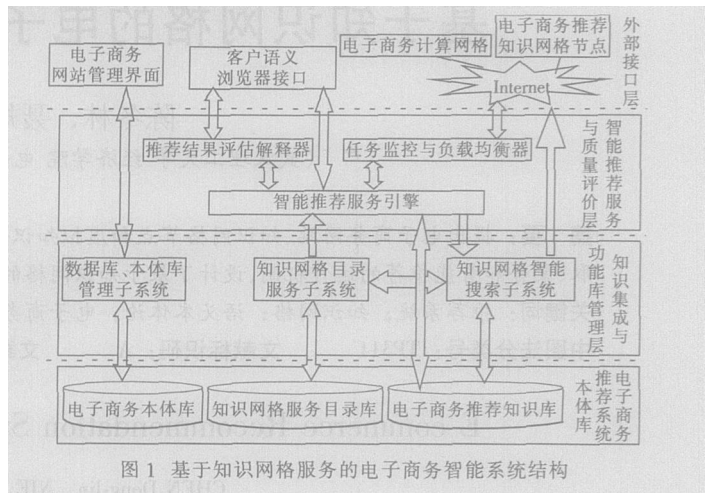


图1 基于知识网格服务的电子商务智能系统结构

3 关键技术

3.1 商品本体库设计

商品本体库是电子商务实用分类系统,其主要成分是一整套对商品知识进行表述的词和术语。根据该知识领域的结构,将这些商品的词和术语组成等级类目,同时规定类目的特性及其之间的关系。这些词和术语被称为商品元数据(Metadata),采用XML语言、W3C的本体论语言(OWL)或者资源描述框架(RDF)表达。商品本体库与庞大通用的本体库建设正好相反,KGRS知识网格本体库的规模可以小很多,由此设计了电子商务智能推荐系统的本体库。在商品本体库的基础上,结合我国统计分类中有关商品分类目录,构造了中文环境下的商品本体原型库。

3.2 商品知识获取、聚合和智能推荐技术

具体包括以下推荐技术和算法的研究:①基于商品评分的灰色关联度协同过滤算法。灰色关联度分析不需要大量的样本及数据的典型分布,就可以解决客户评分数据的极端稀疏性问题。②基于商品效用的推荐理论研究。通过计算商品对客户的作用来作出推荐。③基于知识服务的推荐算法,通过推断客户的需要进行推荐,搜索推荐系统知识库满足客户需要的商品知识。④基于商品模糊聚类的协同过滤算法。通过客户对商品评分的相似性对商品模糊聚类,选择与目标商品相似性最高的聚类空间,搜索目标商品,从而能够在尽量少的商品空间上搜索到目标。对商品的模糊聚类计算比较耗时,因此采用离线进行处理并转换为知识网格的知识,减少类似推荐时的运算时间。⑤基于客户评分的商品知识获取和聚合算法。⑥通过电子商务网站数据资源进行商品知识的抽取、挖掘、归纳、演绎和聚合的方法。

3.3 KGRS知识网格模型

建立KGRS知识网格模型,为电子商务智能推荐系统提供一个智能互连的环境,使电子商务产品或服务的提供者及消费者能有效地获取、发布、共享和管理商品知识,提供电子商务智能推荐应用,并实现商品知识的更新、知识网格间的协同工作等。KGRS知识网格模型由商品知识智能搜索聚合组件、商品知识目录服务组件、商品本体库组成。商品知识智能搜索聚合组件主要完成对KGRS知识网格节点所汇聚的电子商务网站

商品知识的搜索、模糊聚合和排序工作。

3.4 基于语义本体论的最相邻知识网格搜索和过滤技术

这是动态构建虚拟知识网格服务社区关键技术的基础。知识网格中的资源极其丰富,但要按照客户需求寻找合适的虚拟知识网格服务社区仍面临不少问题。在现有的相关或类似技术中,较有影响的有 Web Service 中的 UDDI G bbus Toolkit 中基于 LDAP 查询的 MDS2 机制以及 G bbus 中基于服务数据匹配的 Index Service。它们的共同缺点是:在实际应用中,电子商务网站与客户往往无法做到事先沟通,因而简单的关键字匹配的方式很难提供足够的灵活性和高质量,推荐结果难以令人满意^[15-16]。本文使用 OWL 语言进行描述、匹配,提出在最相邻知识网格服务中使用本体论进行描述和搜索过滤算法。该算法完成以下功能:①智能搜索与客户需求知识相关的最相邻知识网格;②从 P2P 网络系统实际结构出发,分析虚拟知识网格社区拓扑及特征节点;③从一个客户实际需求出发,分析语用搜索方法的具体交互过程;④能自动演化出新的商品知识理论体系,即自动知识进化。

4 结论

电子商务推荐系统是提高电子商务网站竞争力、满足客户个性化需求的必然趋势。现有的推荐技术和系统不能满足大规模、高质量和实时的网络化电子商务智能推荐需要,并缺乏有效的推荐结果来解释模型和评估体系。基于知识的推荐技术是解决该问题最有效的方法,但知识的获取、聚合和智能推荐一直是其研究的难点。知识网格和语义本体论的发展及相关领域的应用成果为该电子商务智能推荐系统提供了一种全新的研究方法。基于知识网格的电子商务智能推荐系统,可以满足商品知识有效获取、聚合和智能推荐的要求,并提高电子商务智能推荐系统的规模性、质量和实时性。

参考文献:

[1] Qianbang Li, Rajiv Khosla. An Adaptive Algorithm for Improving Recommendation Quality of E-recommendation Systems[C]. Lugano Switzerland International Symposium on Computational Intelligence for Measurement Systems and Applications. 2003. 29-31.

[2] Kouichi. An Idea of the Agent-based Information Recommending System Using the Statistical Information[C]. IEEE. 2000. 143-146.

[3] 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.

[4] Chih-Ming Chen, Ling-Juin Duh, Chao-Yu Liu. A Personalized Courseware Recommendation System Based on Fuzzy Item Response

Theory[C]. Proceedings of the 2004 IEEE International Conference on E-technology, E-commerce and E-service (EEE' 04), 2004. 1-4.

[5] 邓爱林, 左子叶, 朱扬勇. 基于项目聚类的协同过滤推荐算法[J]. 小型微型计算机系统, 2004. 25(9): 1665-1670.

[6] 余力, 刘鲁, 李雪峰. 用户多兴趣下的个性化推荐算法研究[J]. 计算机集成制造系统, 2004. 10(12): 1610-1615.

[7] Soe-Tsyi Yuan, Chiah-shin Cheng. Ontology-based Personalized Clustering for Heterogeneous Product Recommendation in Mobile Marketing[J]. Expert Systems with Applications. 2004. 26: 461-476.

[8] Bin Xiao, Esma Aïmeur, José Manuel Fernández PCF. Indexing an Intelligent Product Recommendation Agent for E-commerce[C]. IEEE International Conference on E-Commerce. 2003. 24-27.

[9] Chuan-Feng Chiu, Timothy K. Shih, Ying-Hong Wang. An Integrated Analysis Strategy and Mobile Agent Framework for Recommendation System in E-commerce over Internet[J]. Tamkang Journal of Science and Engineering. 2002. 5(3): 159-174.

[10] Cyrus Shahabi, Farnoush Banaei-Kashani, Yoda. An Accurate and Scalable Web-based Recommendation System[C]. Grid and Cooperative Computing. The 2nd International Workshop. 2003. 1-14.

[11] Lan M., Zhou W. An Adaptive Information Grid Architecture for Recommendation System[C]. Germany. Grid and Cooperative Computing. The 2nd International Workshop. GCC Revised Papers. Part I. Springer-Verlag. 2003. 584-591.

[12] Alexander Smimov, Mikhail Pashkin, Nikolai Chikov. Knowledge Logistics in Information Grid Environment[J]. Future Generation Computer Systems. 2004. 20: 61-79.

[13] 金一娇. 语义网和语义网格中的本体研究综述[R]. Technical Report. 2004. 1-9.

[14] 潘红艳, 林鸿飞, 赵晶. 基于 Ontology 的个性化推送系统[J]. 计算机工程与应用, 2005. 41(20): 176-180.

[15] 顾慧翔, 俞勇. 基于领域本体和知识推理的语义互联网应用[J]. 上海交通大学学报, 2004. (4): 583-585.

[16] H. Zhuge. China's E-science Knowledge Grid Environment[J]. IEEE Intelligent Systems. 2004. 19(1): 13-17.

[17] Wu Zhaohui, Chen Huajun, Xu Jiefeng. Knowledge Base Grid: A Generic Grid Architecture for Semantic Web[J]. Journal of Computer Science & Technology. 2003. 18(3): 462-473.

[18] Huajun Chen, Zhouhui Wu, Chang Huang. Open Grid Services of Traditional Chinese Medicine[C]. IEEE International Conference on Systems, Man and Cybernetics. 2003. 4546-4551.

[19] 余力, 刘鲁, 罗掌华. 我国电子商务推荐策略的比较分析[J]. 系统工程理论与实践, 2004. (8): 96-101.

作者简介:

陈冬林(1970),男,湖北安陆人,副教授,博士,主要研究方向为电子商务、知识管理、网格技术;聂规划(1957),男,河南人,教授,博导,主要研究方向为知识管理、人工智能、电子商务。

(上接第 306 页)

[33] Yongguang Zhang, Wenke Lee. Intrusion Detection in Wireless Ad hoc Networks[C]. MOBIKOM' 00. 2000. 275-283.

[34] Y. Zhang *et al*. Intrusion Detection Techniques for Mobile Wireless Networks[J]. ACM Wireless Networks. 2003. 9(5): 545-556.

[35] Sergio Martí *et al*. Mitigating Routing Misbehavior in Mobile Ad hoc Networks[C]. IEEE MOBIKOM' 00. 2000. 255-265.

[36] Sonja Buchegger, Jean-Yves Le Boudec. The Selfish Node: Increasing Routing Security in Mobile Ad hoc Networks[R]. IBM Research Report RZ 3354. 2001.

[37] Sonja Buchegger, Jean-Yves Le Boudec. Nodes Bearing Girdges: Towards Routing Security, Fairness, and Robustness in Mobile Ad hoc Networks[C]. Spain. The 10th European Workshop on Parallel Dis-

tributed and Network-based Processing. 2002. 403-410.

[38] Sonja Buchegger, Jean-Yves Le Boudec. Performance Analysis of the CONFIDANT Protocol: Cooperation of Nodes, Fairness in Dynamic Ad hoc Networks[C]. IEEE ACM MOBHOC' 02. 2002. 226-236.

[39] Di-Ho Chen, He-Ling-Ge Jiang *et al*. A Chaotic Map with Infinite Collapses[C]. Kuala Lumpur, Malaysia. IEEE 2000 TENCON Proceedings. 2000. 95-99.

作者简介:

何敏(1975),女,博士研究生,主要研究方向为计算机网络、宽带网络与通信;刘心松(1940),男,教授,博导,主要研究方向为宽带网络与通信、分布式并行操作系统、分布式并行数据库、分布式并行多媒体服务器、分布式并行防火墙等;赵东风(1957),男,教授,博导,主要研究方向为随机多址通信、轮询多址通信等。