

文章编号:1001-9081(2014)04-1118-04

doi:10.11772/j.issn.1001-9081.2014.04.1118

多 Agent 协同的电子商务推荐系统模型

危世民, 戴牡红^{*}

(湖南大学 信息科学与工程学院, 长沙 410082)

(*通信作者电子邮箱 komawei@126.com)

摘要:为了进一步提高电子商务推荐系统中商品推荐的准确性和高效性,通过分析传统推荐系统存在的问题和已有的优化方案,提出了多 Agent 的电子商务推荐系统模型。推荐系统通过人工智能领域中的多 Agent 技术,并应用终端自适应特性,改善了传统推荐系统在多终端情况下的电子商务系统的推荐效率,并根据用户使用的不同终端动态返回推荐结果。实验结果表明,多 Agent 协同的电子商务推荐系统在一定程度上提高了推荐效率和准确性。

关键词:多 Agent; 推荐系统; 电子商务; 个性化推荐; 聚类分析

中图分类号: TP18 **文献标志码:**A

Multi-Agent cooperative E-commerce recommender system model

WEI Shimin, DAI Muhong^{*}

(College of Information Science and Engineering, Hunan University, Changsha Hunan 410082, China)

Abstract: In order to further improve the accuracy and efficiency of product recommendation in E-commerce recommender system, by analyzing the problems of traditional recommender system and prior optimization technique, a multi-Agent recommender system model was proposed. By adopting multi-Agent technique in Artificial Intelligence (AI) and terminal self-adaption, this model improved the efficiency of traditional recommender system in the condition of using multi-terminal, and dynamically returned recommendation to users according to different terminals. The experimental results show that multi-Agent cooperative E-commerce recommender system improves the accuracy and efficiency to a certain extent.

Key words: multi-Agent; recommender system; E-commerce; personalized recommendation; cluster analysis

0 引言

如今,电子商务在社会和生活中的地位越来越显著。与此同时,电子商务规模的急剧扩大也使得用户耗费大量的时间浏览无关商品,对于销售商而言,以最合适的方式将商品推荐给用户是他们迫切希望的^[1]。于是,电子商务推荐系统逐渐发展起来并成为研究热点。基于传统推荐算法的电子商务推荐系统存在着稀疏性、冷启动、新用户等问题^[2],因此,如何解决这些问题成为研究中的重中之重。

现有的大多数关于推荐系统优化的研究主要集中于对传统算法的优化。文献[3]对基于内容的推荐算法进行改进,然而基于内容的推荐算法本身的新项目缺陷导致单独优化该算法的提升空间不大;文献[4]介绍了一种将数据集进行聚类的改进协同过滤算法,文献[5]和文献[6]分别基于聚集平滑和分布式计算改进协同过滤;文献[7]则是将基于模型与基于内存的协同过滤算法相结合,在一定程度上改善冷启动问题。这些文献的研究都只是优化了推荐响应时间,而忽略了推荐的个性化程度;或者是优化了推荐效率,却缺乏普遍性。也有一部分研究考虑应用多 Agent 技术对推荐系统模型进行改进。文献[8]介绍了应用于 Web 搜索的多 Agent 推荐系统,文献[9]详细介绍了一种多 Agent 推荐系统通信技术的具体实现。然而,很少有研究提出关于非个人电脑(Personal Computer, PC)终端访问电子商务系统对推荐系统的影响。文献[10]提出基于对等选择的分布式推荐系统架构,然而该

架构为了比较用户的偏好配置文件造成计算成本过高;文献[11]采用设备监听器来记录用户设备配置文件,并根据相似度对用户进行分组,再根据用户配置文件使用算法生成器对每个用户组生成推荐算法。在此基础上,本文提出了一种支持终端自适应的多 Agent 协同的电子商务推荐系统模型,从而提高推荐系统的个性化、自动化和持久化程度。

1 多 Agent 技术

1.1 Agent 技术的概念及其特征

Agent 技术来源于人工智能领域,中文译名为智能体。Agent 有如下特征^[12]:

自治性 Agent 具有其自身的计算资源和局部于自身的行为控制机制,可以在没有外界直接操纵的情况下,根据其内部状态和感知到的环境信息,决定和控制自身行为。

社会性 Agent 能够通过通信语言与其他 Agent 进行交互,有效地与其他 Agent 协同工作。

反映性 Agent 能够感知所处的环境并及时作出反应。

主动性 Agent 能够采取主动行为,表现出面向目标的行为。

1.2 多 Agent 协同

由于 Agent 的社会性,采用多 Agent 架构优于单个 Agent 控制的系统,通过多 Agent 协同工作,能够提高系统效率。利用 Agent 的特征,将多 Agent 技术应用到推荐系统中,可提高电子商务推荐系统中商品推荐的准确率和效率^[13~14]。

收稿日期:2013-10-08;修回日期:2013-12-03。 基金项目:湖南省自然科学基金资助项目(2011FJ3034)。

作者简介:危世民(1986-),男,福建三明人,硕士研究生,主要研究方向:电子商务、推荐系统; 戴牡红(1964-),男,湖南涟源人,副教授,主要研究方向:数据库、企业信息化、决策支持、知识工程。

2 终端自适应特性

在进行网页访问时,用户使用的终端不尽相同,例如PC、笔记本电脑、手机、个人数字助理(Personal Digital Assistant, PDA)等。这些终端有各自独特的构造和技术特点^[15](例如显示模式、带宽等),电子商务网站在展示推荐结果时需要区分对待;其次,在电子商务网站和各个终端之间传输数据所使用的格式存在差异;对于手机、PDA等处理能力较差、资源有限的终端,在这些终端上运行软件收集终端的配置文件并不现实。

因此,系统应用了终端自适应特性,可以解决以往用户在浏览电子商务网站获得商品推荐时因为使用不同终端造成感知差的问题。

3 系统模型介绍与多Agent应用

本文提出了一个多Agent协同的电子商务推荐系统模型(Multi-Agent Cooperative E-Commerce Recommender System, MACERS),该模型由4种Agent组成。MACERS的整体架构如图1所示。

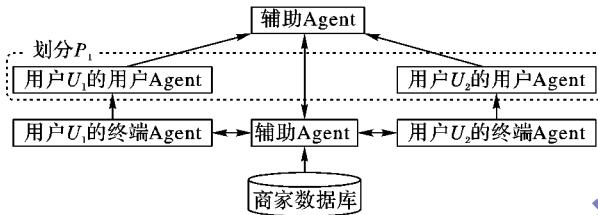


图1 MACERS整体架构

用户访问电子商务系统所使用的每个终端都与一个终端Agent关联,终端Agent监听用户行为,并建立用户与该终端关联的本地配置文件;每个用户都关联一个用户Agent,用户Agent收集终端Agent的信息,建立用户偏好配置文件,并根据偏好配置文件进行划分;辅助Agent是MACERS中最重要的组成部分,每个辅助Agent与用户Agent的一个划分关联,决定划分之间的相似度;商家Agent为辅助Agent提供的信息:网站商品的类别列表,以及与该辅助Agent关联的划分中的每个用户所访问商品的类别列表。

以下对MACERS的各个Agent的结构和职责进行详细描述。

3.1 终端Agent

在一个电子商务会话中,终端Agent保存终端信息,并根据本次会话所访问的商品类别更新用户的配置文件。

3.1.1 终端Agent结构

终端Agent包含两个数据结构,终端配置文件(Terminal Configuration File, TCF)和用户配置文件(User Configuration File, UCF)。TCF存储了一系列的配置参数,如表1所示。这些参数表明了当某个用户使用某个终端访问系统时的商品偏好。

UCF是基于整个电子商务会话建立的。当用户使用某个终端对商品点击操作,UCF会自动更新。UCF是一个元组的集合,结构为(C , IW , LU),其中 C 是电子商务网站的某个商品类别, IW 代表用户使用该终端访问电子商务网站时对于 C 的兴趣度,而 LU 表示UCF最近一次更新的日期。为了使 IW 在[0, 1]的区间内,本文使用时间 t 来定义 IW , t 代表当用户访问包含 C 的页面所花费的时间。对于一次更新, IW 的计算公式如下:

$$IW = \begin{cases} (IW + \frac{t}{T} \times A_n)/2, & t \leq T \\ (IW + A_n)/2, & t > T \end{cases}$$

其中 A_n 为表1中的系数($n = 1, 2, 3$)。计算 IW 的同时,根据两次 LU 之间的时间间隔,使用参数 p 降低未访问的商品类别的 IW 值。

表1 TCF配置参数

参数名	参数描述
CAS	与用户归属的划分相关联的辅助Agent集合
$Text_{max}$	一次电子商务会话中处理字符数据的最大数值(单位:kB)
$Media_{max}$	一次电子商务会话中处理非字符数据的最大数值(单位:kB)
A_1	用户访问某个商品的页面所产生的偏好权重
A_2	用户购买某个商品所产生的偏好权重
A_3	用户收藏某个商品所产生的偏好权重
T	评估用户对于某商品的兴趣度的系数
ω	用户对某个商品进行操作与上一次对该商品进行操作的天数间隔
ψ	区间为[0,1]的参数,用于每过 ω 天用户对某商品不再进行操作所降低的兴趣度
r	每个相似的Agent给出的推荐个数
k	辅助Agent生成基于内容推荐时获取的兴趣商品类别的最大个数
z	辅助Agent生成协同过滤推荐时获取的相似Agent的个数

3.1.2 终端Agent职责

终端Agent根据用户在本次电子商务会话中所访问的商品类别创建用户配置文件UCF。终端Agent周期性地将UCF发送给用户Agent。当用户访问电子商务网站时,终端Agent将与用户使用的设备有关的参数发送给商家Agent,为用户生成个性化的商品推荐。最后,根据本次的电子商务会话时间,终端Agent周期性地更新 IW 的值。

3.2 用户Agent

在MACERS中,每个用户都与一个用户Agent相关联。用户Agent监听用户的电子商务会话,采用离线模式从终端Agent获得该用户访问的商品类别信息,之后,将这些信息发送给辅助Agent。

3.2.1 用户Agent结构

用户Agent的数据结构由连接参数(Connection Parameters, CP)和全局用户配置文件(Global User Configuration File, GUCF)两部分组成。CP包含以下内容: ND 表示与用户关联的终端Agent的个数; C 是包含 ND 元素的一个数组,其中每个元素 c_i 表示使用第*i*个终端的Internet连接成本; Σ 表示在[0, 1]的一个阈值,用于计算用户划分的相似度; α 表示划分相似度的计算周期。

与用户的商品访问记录相关的数据保存在GUCF中。GUCF以(C , GIW)的元组形式存储,其中: C 表示用户访问过的商品类别, GIW 表示使用不同终端访问电子商务网站时对 C 的兴趣度。 GIW 的计算公式如下:

$$GIW = \frac{\sum_{i=1}^{ND} c_i \times IW_i}{\sum_{i=1}^{ND} IW_i}$$

其中: IW_i 为使用第*i*个终端访问网站时对 C 的兴趣度, c_i 为使

用第 i 个终端访问网站的 Internet 连接成本。

3.2.2 用户 Agent 职责

用户 Agent 有两个职责,一是利用终端 Agent 发送给用户 Agent 的数据,对全局用户配置文件 GUCF 进行更新;二是确定用户归属的划分。用户划分的确定步骤如下:

假设存在 n 个划分 P_1, P_2, \dots, P_n , 每个划分 P_i 分别与一个辅助 Agent 关联。用户 Agent 周期性执行以下操作:

1) 用户 Agent 给每个辅助 Agent 发送请求,并得到辅助 Agent 的响应,响应内容为 $GUCF_{ik}$ 的列表 L_i , $GUCF_{ik}$ 为归属于划分 P_i 的用户 k 的全局用户配置文件。

2) 用户 Agent 对存储在用户 Agent 中的 GUCF 和存储在 L_i 中的 $GUCF_{ik}$ 进行相似度 S_i 的计算。本文采用 Jaccard 相似度度量用户 a 与用户 b 的全局配置文件的相似度。假设用户 a 与 b 的全局配置文件分别为 $GUCF_a, GUCF_b$, 则 Jaccard 相似度度量为

$$\sigma_{ab} = \frac{|GUCF_a \cap GUCF_b|}{|GUCF_a \cup GUCF_b|}$$

划分 P_i 的平均相似度 S_i 的计算公式为

$$S_i = \frac{1}{L_i} \sum_{k \in P_i} \sigma_{ab}$$

3) 对于平均相似度 S_i 大于阈值 Σ 的每个划分 P_k , 用户 Agent 发送全局用户配置文件 GUCF。

4) 根据存储在 GUCF 中的 χ 参数,划分的相似度计算操作按照周期反复进行。

总之,用户 Agent 计算出的相似度决定了用户的划分。该相似度会动态地进行重新计算,并且部分重叠(一个用户可以归属不同的划分)。这表明聚类是在离线模式下进行的,完全独立于用户行为。因此,聚类不会影响到用户行为的效率,尤其是在推荐生成的过程中。然而,这种聚类架构并非最优选择,因为聚类是周期性执行的,而不是数据发生变化时就执行。

在聚类行为中,如果时间周期 χ 增加,用户 Agent 的任务将不再那么繁重;然而,如果 χ 太大,系统的推荐则不够准确。因此, χ 的最优值必须要由经验丰富的人确定。

3.3 商家 Agent 与辅助 Agent

3.3.1 商家 Agent 与辅助 Agent 结构

由于商家 Agent 结构仅包含电子商务网站的商品列表,因此本节仅详细描述辅助 Agent 的结构。

辅助 Agent 数据结构包含三部分:商家商品目录(Seller Product Catalog, SPC)、全局配置文件集合(Global Configuration File Set, GCFS)和配置文件收集器(Configuration File Collector, CFC)。SPC 中包含电子商务网站的所有商品;GCFS 包含与辅助 Agent 相关联用户的所有配置文件;CFC 包含一个数据段(Data Section, DS),与电子商务网站相关联。在 DS 中包含与电子商务网站访问者相关联的用户配置文件列表。将每个配置文件标记为 $DS[u, t]$, u 代表某个用户, t 代表该用户使用的终端。每当一个访问者的电子商务会话结束时,商家 Agent 将与访问者相关联的 DS 发送给辅助 Agent。

3.3.2 商家 Agent 与辅助 Agent 职责

商家 Agent 与辅助 Agent 共同生成商品推荐。

假设某个用户 u 使用某个终端 t 访问电子商务网站,首先,用户 u 的终端 Agent 将终端配置文件 TCF 发送给商家 Agent。根据计算,用户属于某些划分,每个划分与一个辅助 Agent 相关联。之后,商家 Agent 与每个辅助 Agent 通信,并将 TCF 发送给辅助 Agent。为了支持基于内容的推荐,辅助 Agent 创建一个商品列表 CB,商品归属于用户 u 的 GCF 中的

商品类别。接下来,辅助 Agent 根据 IW 系数将 CB 按照 IW 系数降序排列,保留前 k 项并删除列表中的其他元素。同时,为了支持协同过滤的推荐,辅助 Agent 将用户 u 的 $DS[u, t]$ 与所有使用相同终端访问电子商务网站的用户 v_i ($i = 1, 2, \dots, n$) 的 $DS[v_i, t]$ 进行比较,最终得到另一个商品列表 CF,商品归属于与用户 u 差别度最小的 z 个用户所访问的商品类别。 k 与 z 均为终端 t 的 TCF 参数。

使用相同设备 t 的用户 u 与用户 v 的差别度计算如下:设商品类别 c 同属于数据段 $DS[u, t]$ 和 $DS[v, t]$, $IW_u(c)$ 为用户 u 对商品类别 c 的兴趣度,它保存在用户 u 的配置文件 UCF 中;相应地, $IW_v(c)$ 为保存在用户 v 的 UCF 中的用户 v 的兴趣度。因此对于商品类别 c ,用户 u 与用户 v 的差别度公式如下:

$$D(u, v, t) = |IW_u(c) - IW_v(c)|$$

一般地,对于数据段 $DS[u, t]$ 和 $DS[v, t]$ 共同包含的所有商品类别,用户 u 与用户 v 的全局差别度公式如下:

$$D(u, v, t) = \sum_{c \in DS[u, t] \cap DS[v, t]} |IW_u(c) - IW_v(c)|$$

4 多 Agent 协同工作流程

图 2 阐述了 MACERS 的工作流程。当某个用户 u 访问电子商务网站时, u 的终端 Agent 与商家 Agent 进行交互,并将 u 的偏好信息发给商家 Agent。这些偏好信息存储在终端配置文件 TCF 中。用户 Agent 根据 u 的偏好信息,决定 u 归属于哪些划分。假设用户 u 归属于 n 个划分,那么商家 Agent 与 u 归属的 n 个划分所关联的辅助 Agent 进行通信,并将 u 的终端配置文件发送给辅助 Agent。为了同时支持基于内容的推荐和协同过滤的推荐,辅助 Agent 预先计算了与 u 的终端配置文件匹配度最高的那些商品,以及与用户 u 相似并且使用相同终端的用户所访问的商品。之后,这些预计算出的推荐商品被保存到一个列表中,并返回给商家 Agent。最后,商家 Agent 将做出的个性化推荐返回给终端 Agent,并展示给用户。

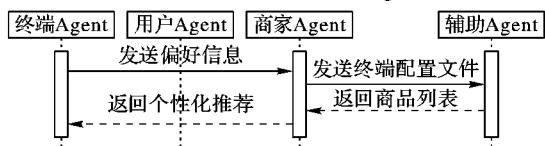


图 2 MACERS 工作流程

5 实验结果与分析

为了验证 MACERS 的推荐性能,与 MASHA^[11]、Ringo^[16]、EC-XAMAS^[17] 和 GroupLens^[18] 进行对比实验。MASHA 与 EC-XAMAS 是多 Agent 推荐系统,在个性化推荐过程中均考虑到设备影响;Ringo 与 GroupLens 是基于协同过滤的推荐系统,然而在本次实验中本文将 Ringo 与 GroupLens 以多 Agent 方式实现,以便对照。实验中,对电子商务会话中的用户进行监听。元组 (P_a, P_b, t, r) 表示用户从商品 P_a 通过链接查看商品 P_b ,并在 P_b 页面上的停留时间 t ,而 r 表示对商品的评分,该数值由用户提供。为了创建用户初始的配置文件 UCF,对于每个用户,采用 20 个元组作为训练集。收集了 120 个用户的数据,并将它们分为 3 组,分别有 20、40 和 60 个用户。

对于每个用户,对应所有属于测试集的元组 (P_a, P_b, t, r) ,生成一个商品推荐列表 $R(a)$,并验证 P_b 是否在该列表中,并将结果保存在值 δ_a 中。 δ_a 的计算公式如下:

$$\delta_a = \begin{cases} 1, & b \in R(a) \\ 0, & b \notin R(a) \end{cases}$$

推荐系统的平均预测精确度(P) 定义为 δ_a 的平均值。

表2展示了五种方法对于用户Agent数量不同的三种情况 S_1, S_2, S_3 的实验结果。从表2可以很清楚地看到,MACERS的性能比其他参与实验的推荐系统更好。

表2 推荐系统平均精确度比较

推荐系统	S_1	S_2	S_3
MACERS	0.68	0.70	0.71
MASHA	0.47	0.54	0.67
Ringo	0.52	0.55	0.54
EC-XAMAS	0.40	0.45	0.49
GroupLens	0.51	0.53	0.54

最后,本文比较了推荐算法对电子商务网站性能的影响。图3说明了当用户访问使用实验中的推荐系统后的电子商务网站的平均等待时间,该平均值基于不同的访问用户数和访问终端数来计算。实验结果表明,MACERS的等待时间明显低于其他测试系统,当访问用户数和访问终端数增加时,在时间成本上获得的增益也同时增加。例如,当有60个用户访问使用MACERS的电子商务网站时,平均等待时间是1.3 s,而MASHA的平均等待时间是4.1 s,EC-XAMAS是6.7 s,Ringo和GroupLens分别是4.3 s和3.7 s。MACERS取得良好性能的关键在于辅助Agent的预算算。事实上,商家Agent的计算复杂度是 $O(P * N_p)$,其中:P是商品数,N_p是相似用户的划分数;而MASHA、GroupLens和Ringo的计算复杂度取决于用户数的平方;EC-XAMAS的时间成本最高,因为该系统的推荐计算均在客户端完成。从实验结果可以得出,MACERS为终端Agent提供的推荐比其他系统更加有效。

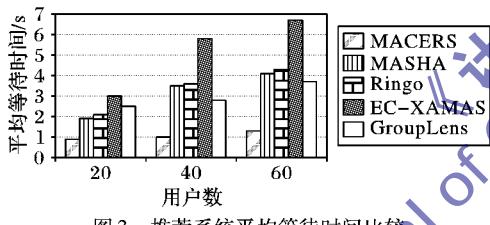


图3 推荐系统平均等待时间比较

6 结语

本文提出了一个支持终端自适应的多Agent协同的电子商务推荐系统MACERS,当用户使用不同的设备访问电子商务网站,推荐系统则根据用户的设备配置文件和用户配置文件产生合适的推荐。未来还需要对以下问题进行研究:首先,确定 χ 的最优值得到最优的用户聚类;第二,在数据稀少的情况下,做到准确推荐;第三,将系统扩展为推荐平台,为不同电子商务网站实现交叉推荐。

参考文献:

- [1] YU W, ZHANG R, WANG L. Recommendation in e-commerce [J]. Journal of East China Normal University: Natural Science Edition, 2013, 2013(3):47–48. (于文皓,张蓉,王立.电子商务中的商品推荐系统[J].华东师范大学学报:自然科学版,2013,2013(3):47–48.)
- [2] YANG J. Data filtrating for recommendation system in electronic commerce[J]. Science Technology and Industry, 2012, 12(6): 137–138. (杨静.电子商务推荐系统的数据过滤[J].科技和产业,2012,12(6):137–138.)
- [3] CAO Y, HE W. Research on E-commerce recommender system based on content-based filtering [J]. Computer Technology and Development, 2009, 19(6): 182–185. (曹毅,贺卫红.基于内容过滤的电子商务推荐系统研究[J].计算机技术与发展,2009,19(6):182–185.)
- [4] CHI N. Research on the improved collaborative filtering algorithm in recommendation system [D]. Fuxin: Liaoning Technical University, 2010. (迟娜.改进的协同过滤算法在推荐系统中的研究[D].阜新:辽宁工程技术大学,2010.)
- [5] XUE G-R, LIN C, YANG Q, et al. Scalable collaborative filtering using cluster-based smoothing [C]// Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval. New York: ACM, 2005: 114–121.
- [6] WENG L-T, XU Y, LI Y, et al. A fair peer selection algorithm for an ecommerce-oriented distributed recommender system [C]// Proceedings of the 2006 Conference on Advances in Intelligent IT: Active Media Technology 2006. Amsterdam: IOS Press, 2006: 31–37.
- [7] JEONG J H, KIM J W. Personalized recommendation based on collaborative filtering with social network analysis [C]// ICICA 2012: Proceedings of the 2012 International Conference on Information and Computer Applications. Singapore: IACSIT Press, 2012: 67–71.
- [8] ALIAKSANDR B, ENRICO B, PAOLO G. Implicit: a multi-Agent recommendation system for Web search [J]. Autonomous Agent and Multi-Agent Systems. Springer, 2012, 24(1): 141–174.
- [9] MORAIS A J, OLIVEIRA E, JORGE A M. A multi-Agent recommender system [C]// Proceedings of the 9th International Conference on Distributed Computing and Artificial Intelligence, Advances in Intelligent and Soft Computing Volume 151. Berlin: Springer-Verlag, 2012: 281–288.
- [10] HE M-H, CHI N. Research on improvement of multi-Agent-based E-commerce recommendation system [C]// JCSIT 2011: Proceedings of the 13th IEEE Joint International Computer Science and Information Technology Conference. Piscataway: IEEE, 2011: 356–359.
- [11] ROSACI D, SARNÉ G M. MASHA: a multi-Agent system handling user and device adaptivity of Web sites [J]. User Modeling and User-Adapted Interaction, 2006, 16(5): 435–462.
- [12] DENG H, ZHANG X, LIU J. Cooperation mechanism of multi-Agent system driven by utility [J]. Journal on Communications, 2013, 34(7): 124–125. (邓寒冰,张霞,刘积仁.效用驱动的多Agent合作机制[J].通信学报,2013,34(7):124–125.)
- [13] YAN Q. Personalized recommendation system based on multi-Agent [D]. Tianjin: Tianjin University, 2009. (闫倩.基于多Agent的个性化推荐系统研究[D].天津:天津大学,2009.)
- [14] KANG X. Technology of multi-Agent used in E-commerce [J]. Computer and Digital Engineering, 2010, 38(2): 170–173. (康晓凤.多Agent技术在电子商务中的应用[J].计算机与数字工程,2010,38(2):170–173.)
- [15] MISO S, HORNOS M J, RODRIGUEZ M L. Highly adaptive and context-aware intelligent system of urban events for mobile devices [C]// Proceedings of the 2012 Eighth International Conference on Intelligent Environment. Piscataway: IEEE, 2012: 122–129.
- [16] SHARDANAND U, MAES P. Social information filtering: algorithms for automating "word of mouth" [C]// Proceedings of the 1995 SIGCHI Conference on Human Factors in Computing Systems. New York: ACM, 1995: 210–217.
- [17] de MEO P, ROSACI D, SARNÉ G M L, et al. EC-XAMAS: supporting E-commerce activities by an XML-based adaptive multi-Agent system [J]. Applied Artificial Intelligence, 2007, 21(6): 529–562.
- [18] RESNICK P, IACOVOU N, SUCHAK M, et al. GroupLens: an open architecture for collaborative filtering of netnews [C]// Proceedings of the 1994 ACM Conference on Computer Supported Cooperative Work. New York: ACM, 1994: 175–186.