

基于集团序方法的推荐系统输出

崔春生

(河南财经政法大学 计算机与信息工程学院, 郑州 450002)

摘 要 论文从推荐系统的输出形式出发, 认为系统以 top- N 形式输出时, N 值的大小影响了推荐的质量和推荐个性化. 论文探讨了输出结果差异化的可行性及一般方法, 采用集团序的方法提出以整个子集团作为推荐输出、以子集团内的产品个数作为候选 N 值的思想, 有效地避免了推荐结果中被推荐产品间差异大, 以及被推荐产品与不被推荐产品之间差异小的问题. 论文构建了推荐系统中一般产品集团序模型, 并针对 N 值大小进行了产品集团序质量评估, 进而得到了可靠的 N 值. 这一研究结果不仅丰富了推荐系统的理论成果, 也为推荐输出的个性化研究探索了新的道路.

关键词 推荐系统; 个性化推荐; 电子商务; 集团序; 推荐输出

Output of recommender systems based on aggregative rank

CUI Chun-sheng

(College of Computer and Information Engineering, Henan University of Economics and Law, Zhengzhou 450002, China)

Abstract From the output functional of recommender systems, the issue of the top- N was researched in this paper. The value of N affects the recommended quality and personalization. The feasibility and general methods of personalization of the output were studied in this paper. A new method based on aggregative rank was proposed. In the method, the entire sub-group is used as output of recommendation and the number of products in the sub-group is used as candidate value of N , in order to avoid big differences between the recommended products in the recommendation results and small differences between recommended products and un-recommended products. The general aggregative rank model of product was built in this paper. Meanwhile, aggregative rank quality was assessment based on different N , only in order to get a reliable value of N . The research results not only enrich the recommended system theoretical results, but also get a new road for personalization study of recommender systems.

Keywords recommender systems; personalization recommendation; electric commerce; aggregative rank; output functional

推荐系统是 20 世纪末随着互联网和电子商务的发展而迅速发展起来的一种技术, 其直观的作用体现在避免用户迷失于海量的产品信息中, 帮助用户选择需要的产品, 提高产品的销售量.

推荐系统的研究直到 20 世纪 90 年代才被作为一个独立的概念提出来^[1-2], 并逐渐形成为一门独立的学科. 目前被广泛引用的推荐系统的非形式化概念是 Resnick 和 Varian 在 1997 年^[3]给出的: “推荐系统是利用电子商务网站向客户提供产品信息和建议, 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”.

自第一批关于协同过滤的文章以来^[1-2,4], 推荐系统在电子商务、网络经济学和人类社会学等领域一直保持很高的研究热度. 各种推荐算法涵盖包括认知科学^[5-6]、近似性理论^[7-8]、信息检索^[9-10]、管理科学^[11-12]、市场营销建模^[13]等在内的众多研究领域^[14].

推荐系统分为输入数据 (input functional), 即为了获得推荐, 用户必须和系统之间进行通讯的信息, 包括用户信息、产品信息等; 推荐方法 (recommendation method), 即用利用输入数据形成推荐算法, 包括协

收稿日期: 2012-09-14

资助项目: 国家社会科学基金 (12BTQ011); 河南省软科学 (132400410254); 河南教育厅基础研究 (12B120001)

作者简介: 崔春生 (1974-), 男, 汉, 河南南阳人, 副教授, 博士, 研究方向: 电子商务智能推荐, 决策理论与方法, E-mail: traition@126.com.

同过滤算法、内容推荐算法、基于知识的推荐、组合推荐等^[15]; 输出结果 (output functional), 即利用算法给出合理的推荐结果, 包括输出形式等^[16].

1 问题的提出

已有的研究中推荐输入^[17-23]、推荐算法^[11,24-32]已经产生了大量的结果, 但推荐输出的研究并不很多. 推荐系统的输出形式包括以下六种形式^[33-35]: 相关产品输出、个体文本评价输出、个体评分输出、平均数值评分输出、电子邮件输出和编辑推荐输出. 生成的推荐结果一般用 Top- N 推荐集的方法表示, 即在得到用户 c_u 的最近邻集合 $C = \{c_1, c_2, \dots, c_k\}$ 后, 设 C 中各用户的评分项集合分别为 $R_1, R_2, \dots, R_k, R_u$ 为 c_u 的评分项集合, 令产品集合 $R_w = R_1 \cup R_2 \cup \dots \cup R_k - R_u$, 则对于 $\forall s_i \in R_w, R_{u,i} \neq \emptyset$, 从而可采用下式预测 c_u 对于 s_i 的评分值, 记为 $P_{u,i}$:

$$P_{u,i} = \bar{R} + \frac{\sum_{(c_u \in C) \cap (R_{c_u} \neq \emptyset)} \text{sim}(c_u, c_k) \times (R_{ki} - \bar{R}_u)}{\sum_{(c_u \in C) \cap (R_{c_u} \neq \emptyset)} (|\text{sim}(c_u, c_k)|)}.$$

式中: R_{ki} 表示最近邻用户 c_k 对产品 s_i 的非空评分, \bar{R}_u 表示 c_k 在与 c_u 的共同评分产品集上的平均评分, \bar{R} 则表示 c_k 在所有产品上的平均评分. 然后按 $P_{u,i}$ 值从大到小取前 N 个项目组成 top- N 推荐集 $S_{rec} = \{s_1, s_2, \dots, s_N\}$ 并将推荐给目标用户 c_u , 从而完成整个推荐过程.

显然, N 的大小决定着推荐的质量问题, N 太小起不到推荐效果, N 太大, 推荐的覆盖度会有所提高, 但是精确度可能会降低, 由此带来的推荐成本也会成级数增加, 因为在 Top- N 推荐系统中, 每次查找邻居 (或相似产品) 都必须从整个数据库中遍历查询^[36]. 因此, 确定合适的 N 值一方面可以减少计算工作量, 另一方面可以提高推荐质量, 还能达到推荐结果个性化的目的, 从而成为推荐系统输出中的研究焦点.

2 集团序方法的引入

集团序方法是吴祈宗、侯福均^[37-38]于 2006 年提出的一种方案评估排序方法. 该方法的背景是实际管理中, 很多时候并不要求得出所有方案的严格序关系, 而只需向决策者提供方案是否隶属于某些方案集团, 此时采取集团序的方法是很合适的. 举例来说, 在网购的环节中, 很多时候用户并不注重被推荐的产品中哪些排在前哪些排在后, 仅仅注重被推荐的产品有哪些.

在推荐系统的输出结果中, 一方面需要给出推荐产品的严格排序关系, 另一方面更重要的是给出被推荐产品的数量. 这种数量不是简单的 20 或者 10, 而是针对不同的用户偏好, 不同的产品特性有选择地确定推荐数量. 例如得到的四个产品的推荐度依次为 1, 0.857, 0.593, 0.574. 显然, 同时推荐 2 个产品或者 4 个产品是合适的, 而同时推荐三个产品是不妥的, 因为第 3 个产品和第 4 个产品从推荐度来看差异不大. 直观上来看, $NR_1(s_1) = 1$ 和 $NR_1(s_2) = 0.857$ 可以看作是同一个集团, $NR_1(s_3) = 0.593$ 和 $NR_1(s_4) = 0.574$ 可以看作是另一个集团. 因此探讨产品的集团排序有助于确定 Top- N 的推荐数量 N .

需要说明的是, 采用线性变换关系求取推荐度时^[32], 产品的推荐值 $R(s_j)$ 和产品的推荐度 $NR(s_j)$ 之间存在关系:

$$NR(s_j) = \frac{R(s_j)}{\max_j R(s_j)}.$$

因此, 以产品推荐值 $R(s_i)$ 讨论 N 值等同于以推荐度 $NR(s_i)$ 讨论 N 值.

如果采用非线性关系求取推荐度, 推荐值 $R(s_j)$ 和产品的推荐度 $NR(s_j)$ 不具有线性关系, 此时需要分别讨论推荐值得到的 N 值和推荐度得到的 N 值.

3 产品集团序

3.1 产品集团序定义

所谓产品集团序是指将产品按照推荐值或者推荐度进行排序, 进而寻求不同的产品集团. 为了便于描述, 本文以推荐值定义产品集团序.

设待推荐产品集 $S = \{s_1, s_2, \dots, s_m\}$ 共有 m 个待推荐产品, $R_{ui}(u = 1, 2, \dots, n; i = 1, 2, \dots, m)$ 表示产品 s_i 相对目标用户 c_u 的推荐值.

定义 1 给出 p 个产品的序集合 (O_1, O_2, \dots, O_p) , $p \leq m$, 并且满足条件^[37]:

① $O_1 \cup O_2 \cup \dots \cup O_p = S$;

② $\forall i, O_i \neq \emptyset$;

③ 对于任意 O_i, O_j , 如果 $i \neq j$, 则 $O_i \cap O_j \neq \emptyset$;

④ 如果 $p > 1$, 则 $\forall q (R_q^{\min} > R_{q+1}^{\max})$, 其中 $R_q^{\min} = \min\{R_{ui} | s_i \in Q_q\}$, $R_{q+1}^{\max} = \max\{R_{uj} | s_j \in Q_{q+1}\}$;
则称 (O_1, O_2, \dots, O_p) 为对应于产品集合 S 的集团序, 记为 $(O_1, O_2, \dots, O_p)^S$.

根据定义 1 可知, 如果序集合 $(O_1, O_2, \dots, O_p)^S$ 为对应于产品集合 S 的集团序, 则 $\max\{R_{ui} | s_i \in O_1\} = \max\{R_{ui} | s_i \in S\}$, 并且产品集合 S 中的产品数目 m 满足 $m \geq p$, 若 $m = p$, 则退化为一般序关系.

3.2 产品集团序优先关系

传统意义下, 产品的排序基于以下假设: 若 $R_{ui} > R_{uj}$, 则 $s_i \succ s_j$. 进而存在以下关系^[39]:

① 若 $R_{ui} - R_{uj} > \varepsilon$, 则认为对产品 s_i 的推荐在 ε 意义下优于对产品 s_j 的推荐. 在产品集合 S 中, 对产品 s_i 的推荐在 ε 意义下的优先集合为 P_i^S .

如果产品 s_i 在 ε 意义下的优先集合 P_i^S 为空集, 则产品 s_i 构成优先集团.

② 若 $|R_{ui} - R_{uj}| \leq \varepsilon$, 则认为对产品 s_i 的推荐与对产品 s_j 的推荐在 ε 意义下无差异. 在产品集合 S 中, 对产品 s_i 的推荐在 ε 意义下的无差异集合为 Q_i^S .

③ 若 $R_{ui} - R_{uj} < -\varepsilon$, 则认为对产品 s_i 的推荐在 ε 意义下劣于对产品 s_j 的推荐. 在产品集合 S 中, 对产品 s_i 的推荐在 ε 意义下的劣集合为 R_i^S .

其中, 向量 $B = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p)$ 表示等级差别, p 为推荐产品集团等级数, ε_i 为对应的级内差别.

4 产品集团序构造过程

产品集团排序的思路很容易理解, 首先从推荐值 R_{ui} 或推荐度 $NR(s_i)$ 的比较开始, 然后确定优先集, 最后得到产品集团排序, 各环节的关系如图 1 所示.

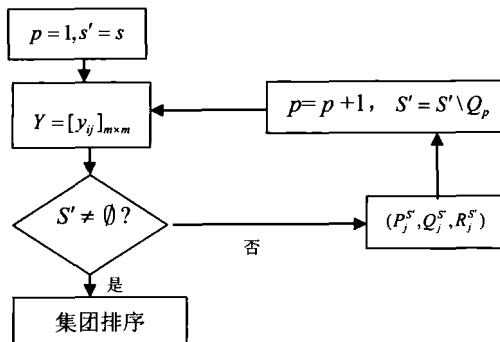


图 1 集团排序流程

基本过程描述为:

Step 1 令 $p = 1$, $S' = S$, 确定比较矩阵 $Y = [y_{ij}]_{m \times m}$, 其中 $y_{ij} = R_{ui} - R_{uj}$, $i, j = 1, 2, \dots, m$.

Step 2 如果 $S' \neq \emptyset$, 对 $\forall s_j \in S'$, 找出 $(P_j^{S'}, Q_j^{S'}, R_j^{S'})$, 进入下一步; 否则, 转入 Step 4.

Step 3 令 $p = p + 1$, $S' = S' \setminus Q_p$, 得到新的比较矩阵 Y , 转入 Step 2.

Step 4 得到集团排序, 计算结束.

5 产品集团序质量评估

产品的集团序方法实质上是一种聚类方法, 聚类的对象是产品 s_i , 而聚类计算的基础是与 s_i 一一对应的产品推荐值 R_{ui} 或推荐度 $NR(s_i)$. 这种聚类方法以 ε 值作为“聚类宽度”, 类似于一般聚类方法中的“聚类半径”. “聚类宽度”是一种粒度的概念, 粒度的大小决定了每一集团内待推荐产品的数量, 同时也决定了集团内待推荐产品的相似程度. 但是可以肯定的是, 采用这种方法一定可以保证集团间的相似程度较低. 正如一般的聚类设计一样, 这种集团序的方法也着重考虑两个因素评价聚类方法的优劣, 即: 每个簇的内部应该是紧凑的, 各个簇之间的距离应该尽可能地远.

为了实现这一目标只需考察集团 O 内产品推荐值的差异 $w(O)$ 和集团间差异 $b(O)$. 这里, 探讨两种方法, 一种是以集团的聚类中心来考虑问题, 另一种是以集团的边界考虑问题.

5.1 以集团聚类中心评估

以集团的聚类中心评估相对简单, 适用于一般的聚类方法, 类内产品差异可以采用多种距离函数来定义, 最简单的是计算集团内每一个产品推荐值到它所属类中心的距离的平方和, 即:

$$w(O) = \sum_{I=1}^K w(O_I) = \sum_{I=1}^K \sum_{s_i \in O_I} d(R_{ui}, O_I)^2 \quad (1)$$

其中, $O_I (I = 1, 2, \dots, K)$ 表示第 I 个集团; $d(R_{ui}, O_I)$ 表示 O_I 内产品 s_i 到该集团中心的距离.

集团间差异 $b(O)$ 定义为不同聚类中心间的距离, 即

$$b(O) = \sum_{1 \leq J \leq I \leq K} d(O_I, O_J)^2 \quad (2)$$

其中, $d(O_I, O_J)$ 表示 O_I 集团和 O_J 集团聚类中心的距离.

集团内差异衡量集团的紧凑性, 集团间差异衡量不同集团之间的相离性, 集团的总体质量可以被定义为类内差异和类间差异的组合. 基于这一思想, 将质量评价函数取为集团内差异和集团间差异平方的二次方根^[40]:

$$J(O, q) = \sqrt{(w(O))^2 + (b(O))^2} \quad (3)$$

当质量评价函数达到最小时, 将得到最优的集团结果, 其中 q 的选择如下:

$$\min\{J(O, q)\}, q = 1, 2, \dots, m \quad (4)$$

5.2 以集团边界评估

以集团边界评估的方法更容易描绘集团序的思想, 可以借助 TOPSIS 方法实现质量评估.

定义 2 相邻集团 集团排序得到的任意三个相邻集团表示为 O_{I-1}, O_I, O_{I+1} , 并且存在 $O_{I-1} \succ O_I \succ O_{I+1}$.

为了方便描述, 对符号做以下简单约定:

集团 O_{I-1} 内共有 p_{I-1} 个产品, 依次为 $s_i^{I-1} (i = 1, 2, \dots, p_{I-1})$, 对应的推荐值为 R_{ui}^{I-1} ; 集团 O_I 内共有 p_I 个产品, 依次为 $s_i^I (i = 1, 2, \dots, p_I)$, 对应的推荐值为 R_{ui}^I ; 集团 O_{I+1} 内共有 p_{I+1} 个产品, 依次为 $s_i^{I+1} (i = 1, 2, \dots, p_{I+1})$, 对应的推荐值为 R_{ui}^{I+1} . 各推荐值间存在关系 $R_{ui}^{I-1} > R_{ui}^I > R_{ui}^{I+1}$.

定义 3 理想解与负理想解 集团 O_I 的理想解定义为 O_I 内的最优值, 即: $O_I^{\max} = \max_{\substack{s_i \in O_I \\ i=1,2,\dots,p}} R_{ui}^I, (I = 1, 2, \dots, K)$; 相应定义集团 O_I 的负理想解为 O_I 内的最劣值, 即: $O_I^{\min} = \min_{\substack{s_i \in O_I \\ i=1,2,\dots,p}} R_{ui}^I, (I = 1, 2, \dots, K)$.

对于任意三个相邻集团, 存在关系:

$$O_{I-1}^{\max} \succ O_{I-1}^{\min} \succ O_I^{\max} \succ O_I^{\min} \succ O_{I+1}^{\max} \succ O_{I+1}^{\min}.$$

推理

① 若 $I < J$, 则 $O_I^{\max} \succ O_J^{\max}, O_I^{\min} \succ O_J^{\min}$;

② 若 $I = J$, 则 $O_I^{\max} \sim O_J^{\max}, O_I^{\min} \sim O_J^{\min}$;

③ 若 $I > J$, 则 $O_I^{\max} \prec O_J^{\max}, O_I^{\min} \prec O_J^{\min}$.

因此, 集团内产品的差异 $w(O)$ 可以定义为集团内任意评分值到理想解的距离, 即:

$$w(O) = \sum_{I=1}^K w(O_I) = \sum_{I=1}^K \sum_{\substack{i=1 \\ s_i \in O_I}}^p (O_I^{\max} - R_{ui}) \quad (5)$$

集团间差异 $b(O)$ 定义为相邻集团间负理想解与理想解之间的距离, 即

$$b(O) = \sum_{I=2}^K b(O_I) = \sum_{I=2}^K (O_{I-1}^{\min} - O_I^{\max}) \quad (6)$$

同样根据“集团内耦合度高, 集团间组合度低”的思想, 得到质量评价函数, 即

$$J(O, q) = \frac{w(O)}{w(O) + b(O)} \quad (7)$$

质量评价函数依赖于集团序等级向量 $B = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p)$, 当质量评价函数达到最小时, 将得到最优的集团结果, 其中 q 的选择如下:

$$\min\{J(O,q)\}, q=1,2,\cdots,m$$

(8)

相对以集团聚类中心的评估方法, 该方法通过集团排序结果, 以子集团最优值作为理想解, 以子集团最劣值作为负理想解, 很好地诠释了集团序的思想, 进而可以得到更优的集团排序结果.

6 *N* 值的确定

通过集团序方法得到待推荐产品的集团排序, 进而可以确定 Top-*N* 推荐系统中 *N* 的选取策略:

Step 1 确定初始集团序

在获得待推荐产品推荐值 R_{ui} (或推荐度 $NR(s_i)$) 的基础上, 选择初始向量 $B = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_p)$, 确定推荐值 (或推荐度) 的初始集团排序.

Step 2 计算评价函数

采用公式 (1)–(3) 或 (5)–(7) 计算初始集团排序质量评价函数.

Step 3 选择最优集团排序

调整 *B* 值, 重复 Step 1 和 Step 2, 获得多个集团排序, 采用公式 (4) 或公式 (8) 检验各个排序的质量, 从中选择最优集团排序结果, 记为: (O_1, O_2, \cdots, O_p) .

Step 4 确定推荐系统的*N*值

取 $N = \sum_{i=1}^q |O_i|$, ($q = 1, 2, \cdots, p$), 其中 $|O_i|$ 表示子集团 O_i 中元素的个数. 即: 推荐系统中的 *N* 值等于整个子集团产品的个数.

显然采用基于集团序的 Top-*N* 推荐方法不再是简单的确定 10 或者 20 这样的数值作为 *N* 值, 而是根据用户 c_u 的不同推荐结果, 有针对地选择整个子集团内产品作为推荐输出结果, 以子集团内的产品个数作为候选 *N* 值, 这种方法的提出不仅提高了推荐系统的个性化程度, 也提高了系统的推荐精度.

7 案例计算

随机抽取 Movielens 中的用户 804 和用户 901, 以两个用户推荐值最高的前 30 个产品进行计算, 如表 1 所示.

Step 1 由于待推荐产品的数量不多 (案例中仅选取了 30 个), 且评分值离差度不大, 因此选择 $B = (0.4, 0.4, 0.3, 0.3)$, 按推荐值将待推荐产品分成不同集团.

根据第 4 节提出的方法, 可以得到用户 804 对应产品推荐值的集团排序为: $(O_1, O_2, O_3, O_4)^S$, 其中, $O_1 = \{71, 756, 571, 3, 1457\}$, $O_2 = \{124, 1507, 313, 815, 523, 23, 615, 1401, 328\}$, $O_3 = \{1056, 210, 563, 208, 568, 866\}$, $O_4 = \{445, 720, 840, 234, 488, 38, 257, 970, 797, 588\}$.

用户 901 对应产品推荐值的集团排序为: $(O_1, O_2, O_3, O_4, O_5)^S$, 其中, $O_1 = \{562\}$, $O_2 = \{586, 716, 192, 828, 978, 11, 1, 721, 211\}$, $O_3 = \{475, 15\}$, $O_4 = \{614, 571, 99, 641, 770, 132, 25, 195, 1674, 32, 463, 169, 657\}$, $O_5 = \{924, 810, 1007, 302, 708\}$.

Step 2 根据第 5 节提出的方法, 采用集团边界评估, 计算质量评估函数值, 如表 2 所示.

表 2 推荐结果质量评价

用户	$w(O)$	$b(O)$	$J(O, q)$
804	3.8015	0.3065	0.925
901	4.1295	1.097	0.790

Step 3 调整 *B* 值可以获得不同的集团排序结果和质量评价函数. 不失一般性, 假设以上得到的结果均为最优结果.

Step 4 得到集团排序的结果如下:

表 1 用户 804 和用户 901 的预测值

用户 804		用户 901	
MovieID	Ratings	MovieID	Ratings
71	4.9127	562	4.8851
756	4.7001	586	4.1519
571	4.6196	716	4.116
3	4.5153	192	4.05
1457	4.5099	828	4.0437
124	4.4431	978	3.9309
1507	4.4123	11	3.8763
313	4.3912	1	3.8136
815	4.2875	721	3.7587
523	4.1923	211	3.755
23	4.1802	475	3.7296
615	4.1231	15	3.5877
1401	4.1212	614	3.2733
328	4.0445	571	3.1495
1056	4.0262	99	3.1206
210	4.0119	641	3.119
568	3.9023	770	3.1187
563	3.9517	132	3.1121
208	3.9102	25	3.1121
866	3.827	195	3.1121
445	3.6065	1674	3.1121
720	3.6048	32	3.1121
840	3.6004	463	3.1121
234	3.5982	169	3.1121
488	3.5982	657	2.9801
38	3.5823	924	2.9561
257	3.5821	810	2.9527
970	3.582	1007	2.9367
797	3.5712	302	2.911
588	3.5641	708	2.9144

根据用户 804 的集团排序结果 $(O_1, O_2, O_3, O_4)^S$ 和用户 901 的集团排序结果 $(O_1, O_2, O_3, O_4, O_5)^S$, 向用户 804 推荐的产品数量 N 为 5, 14, 20, 30; 而向用户 901 推荐的产品数量 N 为 1, 10, 12, 25, 30.

从以上结果易得: 在同一层级内用户 804 获得的推荐数量略高于用户 901 获得的推荐数量. 原因在于用户 804 推荐值的离差度更小, 即用户 804 各个推荐值的相似度更高, 这与原始数据的表现是一致的.

8 结论

论文研究了推荐系统的输出模块, 立足于推荐输出 Top- N 推荐策略中 N 的取值问题, 取得了两个方面的有效结果: 第一、对应于推荐系统分析了产品集团排序方法. 通过分析推荐输出中推荐值排序和推荐度排序的异同, 探讨了集团排序质量评估的一般思路. 并基于以上分析, 构建了依据推荐值进行产品集团排序的一般方法. 第二、从集团排序出发提出了推荐输出中 N 的取值策略. 通过分析, 找到了借助集团排序获得推荐输出的一般过程, 提出以整个子集团作为推荐输出、以子集团内的产品个数作为候选 N 值的思想, 有效地避免了推荐结果中被推荐产品差异大, 以及被推荐产品与不被推荐产品之间差异小的问题. 同时, 这种方法的提出也有效地满足了推荐结果差异化、个性化、高质量的要求.

但是论文的研究尚存在不足之处: 等级差别向量 $B = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_p)$ 是确定集团边界的关键因素, B 值选择一方面与备选方案的多少有关, 另一方面与备选方案的离差度有关, 而论文中没有对 B 值的确定给出具体方案. 这是论文研究的一个遗憾, 也是本文作者后期研究的一个方向.

参考文献

- [1] Resnick P, Iacovou N, Sushak M, et al. GroupLens: An open architecture for collaborative filtering of net-news[C]// Proceedings of ACM 1994 Conference on Computer Supported Cooperative Work, Chapel Hill: Association for Computing Machinery, 1994: 175-186.
- [2] Hill W, Stead L, Rosenstein M, et al. Recommending and evaluating choices in a virtual community of use[C]// Proceeding of the Conference on Human Factors in Computing Systems, New York: ACM Press, 1995: 194-201.
- [3] Resnick P, Varian H R. Recommender systems[J]. Communications of the ACM, 1997, 40(3): 56-58.
- [4] Shardanand U, Maes P. Social information filtering: Algorithms for automating "Word of Mouth"[C]// Proceeding of the Conference on Human Factors in Computing Systems, New York: ACM Press, 1995: 210-217.
- [5] 白云. P2P 环境中基于语义的资源自组织、发现及推荐研究 [D]. 成都: 西南大学, 2008.
Bai Y. Implements the self-organizing, retrieving and recommending of the resources in P2P networks to support the semantic understanding[D]. Chengdu: Southwest University, 2008.
- [6] Basu C, Hirsh H, Cohen W. Recommendation as classification: Using social and content based information in recommendation[C]// AAAI-98 Proceedings, Menlo Park: AAAI Press, 1998: 714-720.
- [7] Yang M H, Gu Z M. Personalized recommendation based on partial similarity of interests[J]. Advanced Data Mining and Applications, 2006, 4093: 509-516.
- [8] 王太雷. 个性化推荐系统中相似模式聚类研究 [J]. 计算机工程, 2005, 31(10): 156-158.
Wang T L. Research on similar pattern clustering in personality recommendation system[J]. Computer Engineering, 2005, 31(10): 156-158.
- [9] Belkin N, Croft W B. Information filtering and information retrieval: Two sides of the same coin?[J]. Communications of the ACM, 1992, 35(12): 29-37.
- [10] 朱鲲鹏, 刘文涵, 王晓龙, 等. 基于日志挖掘的检索推荐系统 [J]. 沈阳建筑大学学报: 自然科学版, 2009, 25(2): 366-370.
Zhu K P, Liu W H, Wang X L, et al. A recommendation system for information retrieval based on web logs mining[J]. Journal of Shenyang Jianzhu University: Natural Science, 2009, 25(2): 366-370.
- [11] Murthi B P S, Sarkar S. The role of management sciences in research on personalization[J]. Management Science, 2003, 49(10): 1344-1362.
- [12] 李欣璐, 刘鲁. 基于协同过滤的银行产品推荐系统建模 [J]. 计算机与数字工程, 2006, 35(9): 6-10.
Li X L, Liu L. Modeling of recommendation system for financial products based on collaborative filtering[J]. Computer & Digital Engineering, 2006, 35(9): 6-10.
- [13] 叶群来. 营销与网络推荐系统 [J]. 电子商务, 2007(10): 64-66.
- [14] Adomavicius G, Tuzhilin A. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions[J]. IEEE Transactions on Knowledge and Data Engineering, 2005, 17(6): 734-749.
- [15] Blanco-Fernandez Y, Lopez-Nores M, Pazos-Arias J, et al. An improvement for semantics-based recommender systems grounded on attaching temporal information to ontologies and user profiles[J]. Engineering Applications of Artificial Intelligence, 2011, 24: 1385-1397.
- [16] 马辉民, 周凤林. 电子商务下的柔性推荐系统 [J]. 武汉理工大学学报: 信息与管理工程版, 2007, 29(2): 119-123.

- Ma H M, Zhou F L. A flexible recommendation system in e-commerce[J]. Journal of Wuhan University of Technology: Information & Management Engineering, 2007, 29(2): 119-123.
- [17] Kim J, Oard D W, Romanik K. Using implicit feedback for user modeling in internet and intranet searching[D]. College of Library and Information Services, University of Maryland at College Park, 2000.
- [18] 刘建国, 周涛, 汪秉宏. 个性化推荐系统的研究进展 [J]. 自然科学进展, 2009, 19(1): 1-15.
- [19] 崔春生. 推荐系统中显式评分输入的用户聚类方法研究 [J]. 计算机应用研究, 2011, 28(8): 2856-2868.
- Cui C S. Aggregation method of recommender systems with way of sparse matrix reduction by sociality structure[J]. Application Research of Computers, 2011, 28(8): 2856-2868.
- [20] 崔春生. 基于隐式浏览输入的用户聚类分析 [J]. 计算机应用研究, 2011, 28(8): 2862-2864.
- Cui C S. User clustering analysis based on implicit navigation[J]. Application Research of Computers, 2011, 28(8): 2862-2864.
- [21] 王霞. 电子商务推荐系统评述 [J]. 福建电脑, 2006(8): 60-61.
- [22] 熊馨, 王卫平, 叶跃祥. 电子商务个性化产品推荐策略研究 [J]. 科技进步与对策, 2005, 22(7): 163-165.
- [23] Nichols D M. Implicit rating and filtering[C]// Proceedings of the 5th DELOS Workshop on Filtering and Collaborative Filtering, Budapest: ERCIM, 1997: 31-36.
- [24] 许海玲, 吴潇, 李晓东, 等. 互联网推荐系统比较研究 [J]. 软件学报, 2009, 20(2): 350-362.
- Xu H L, Wu X, Li X D, et al. Comparison study of internet recommendation system[J]. Journal of Software, 2009, 20(2): 350-362.
- [25] Balabanović M, Shoham Y. FAB: Content-based collaborative recommendation[J]. Communications of the ACM, 1997, 40(3): 66-72.
- [26] Schafer J B, Konstan J, Riedl J. Recommender systems in e-commerce[C]// Proceeding of the 1st ACM Conference on Electronic Commerce, New York: ACM Press, 1999: 158-166.
- [27] Mooney R J, Bennett P N, Roy L. Book recommending using text categorization with extracted information[C]// Proceeding of the AAAI'98/ICML'98 Workshop on Learning for Text Categorization, Madison: AAAI Press, 1998: 49-54.
- [28] Terveen L, Hill W, Amento B, et al. PHOAKS: A system for sharing recommendations[J]. Communications of the ACM, 1997, 40(3): 59-62.
- [29] Delgado J, Ishii N. Memory-based weighted-majority prediction for recommender systems[C]// Proceeding of the ACM SIGIR'99 Workshop Recommender Systems: Algorithms and Evaluation, New York: ACM Press, 1999: 99-107.
- [30] Konstan J A, Miller B N, Maltz D, et al. GroupLens: Applying collaborative filtering to usenet news[J]. Communications of the ACM, 1997, 40(3): 77-87.
- [31] Goldberg K, Roeder T, Gupta D. Eigentaste: A constant time collaborative filtering algorithm[J]. Information Retrieval, 2001, 4(2): 133-151.
- [32] 崔春生, 吴祈宗. 基于 Vague 集的内容推荐算法研究 [J]. 计算机应用研究, 2010, 27(6): 2108-2110.
- Cui C S, Wu Q Z. Research on content-based recommendation based on vague sets[J]. Application Research of Computers, 2010, 27(6): 2108-2110.
- [33] 邓爱林. 电子商务推荐系统关键技术研究 [D]. 上海: 复旦大学, 2003.
- Deng A L. The research on key technologies of recommendation system in e-commerce[D]. Shanghai: Fudan university, 2003.
- [34] Schafer J B, Konstan J A, Riedl J. E-commerce recommendation applications[J]. Data Mining and Knowledge Discovery, 2001, 5(1-2): 115-153.
- [35] 张富国. 基于信任的电子商务个性化推荐关键问题研究 [D]. 南昌: 江西财经大学, 2009.
- Zhang F G. Research on the key issues of trust-based personalized recommender systems in e-commerce[D]. Nanchang: Jiangxi University of Finance and Economics, 2009.
- [36] 曹波, 苏一丹. 基于蚁群聚类的 top-N 推荐系统 [J]. 微计算机信息, 2009, 25(9): 225-227.
- Cao B, Su Y D. Web log mining based on ant clustering[J]. Microcomputer Information, 2009, 25(9): 225-227.
- [37] 吴祈宗, 侯福均. 方案集团序及其应用 [J]. 北京理工大学学报, 2006, 26(6): 521-524.
- Wu Q Z, Hou F J. Aggregative rank of alternatives and its application[J]. Transactions of Beijing Institute of Technology, 2006, 26(6): 521-524.
- [38] 侯福均, 吴祈宗, 管欣. 集团序及其应用 [J]. 数学的实践与认识, 2006, 36(5): 73-76.
- Hou F J, Wu Q Z, Zan X. Aggregative rank and its application squares estimator is BLUE[J]. Mathematics in Practice and Theory, 2006, 36(5): 73-76.
- [39] 苏白云, 崔春生. 基于集团序的软件和信息服务业收入评估方法研究 [J]. 数学的实践与认识, 2012, 42(9): 85-90.
- Su B Y, Cui C S. Research on the evaluate method of software and information services industry income based on aggregative rank[J]. Mathematics in Practice and Theory, 2012, 42(9): 85-90.
- [40] 汪中, 刘贵全, 陈恩红. 一种优化初始中心点的 K-means 算法 [J]. 模式识别与人工智能, 2009, 22(2): 299-304.
- Wang Z, Liu G Q, Chen E H. A K-means algorithm based on optimized initial center points[J]. Pattern Recognition and Artificial Intelligence, 2009, 22(2): 299-304.