

一种基于信息老化的协同过滤推荐算法

王玉斌* 孟祥武 胡 勋

(北京邮电大学智能通信软件与多媒体北京市重点实验室 北京 100876)

(北京邮电大学计算机学院 北京 100876)

摘 要 协同过滤推荐算法中存在推荐信息低时效性问题, 该文针对此问题, 结合信息老化理论, 提出一种基于信息老化的协同过滤推荐算法。该算法利用用户的点击记录, 构建项目的时效性评价模型来预测项目当前时刻被点击的概率; 将模型与基于项目协同过滤推荐算法结合, 综合考虑用户的兴趣和项目的时效性来发现项目的最近邻居, 从而进行高时效性的推荐。实验结果表明, 与传统基于项目的协同过滤推荐算法相比, 该算法提高了推荐结果的时效性。

关键词: 信息处理; 协同过滤; 推荐系统

中图分类号: TP391

文献标识码: A

文章编号: 1009-5896(2013)10-2391-06

DOI: 10.3724/SP.J.1146.2012.01743

Information Aging-based Collaborative Filtering Recommendation Algorithm

Wang Yu-bin Meng Xiang-wu Hu Xun

(Beijing Key Laboratory of Intelligent Telecommunications Software and Multimedia, Beijing University of Posts and Telecommunications, Beijing 100876, China)

(School of Computer Science, Beijing University of Posts and Telecommunications, Beijing 100876, China)

Abstract: Recommendations from Collaborative Filtering (CF) recommender algorithms have low timeliness. To solve the problem, an information aging-based collaborative filtering algorithm is proposed by combining information age method. The algorithm builds a model to evaluate the timeliness of items based on users' hit records to predict the probabilities of the items being clicked at the present time. To consider comprehensively users' interests and the timeliness of items, the model and item-based collaborative filtering recommender algorithm are combined to find the nearest neighbor collection. Experimental results show that comparing with traditional collaborative filtering recommender algorithm the proposed algorithm can improve the timeliness of recommendations.

Key words: Information processing; Collaborative filtering; Recommendation systems

1 引言

随着网络技术快速发展, 用户无法有效地在海量的信息中获取对自己有用的信息, 这样造成了信息迷航与信息过载的问题。为了让用户更加方便地获取有用信息, 个性化推荐作为信息过滤的有效手段, 在许多领域都得到应用^[1,2]。推荐系统主要依据用户的兴趣进行推荐, 为用户推荐符合其兴趣偏好的信息。其中协同过滤算法^[3,4]是推荐系统中应用最广泛的推荐算法, 协同过滤算法基于相似用户偏好也相似的假设进行预测推荐, 主要根据相似用户的

历史行为来预测, 如用户 A 与 B 非常相似, 那么 A 喜欢项目 i , 那么 B 也可能喜欢项目 i , 但是这其中并没有考虑项目的时效性, 如用户 A 阅读了半年前的一条新闻, 根据协同过滤的思想用户 B 很可能被推荐该新闻, 但是新闻的时效性很高, 过期的新闻对用户的吸引力会剧减, 过期新闻的推荐就没有意义。即现有的推荐系统主要依据用户的兴趣进行推荐, 为用户推荐符合其兴趣偏好的信息, 但是系统中有些信息可能是过期无用的, 即时效性低的信息, 如果依据传统算法将其推荐给目标用户, 会降低推荐的成功率。

在推荐系统领域并没有关于如何过滤过期信息的相关研究, 而通常的信息过滤技术是基于信息处理、语义理解的相关技术对信息进行过滤^[5], 但对信

2012-12-31 收到, 2013-05-03 改回

国家自然科学基金(60872051)和北京市教育委员会共建项目资助课题

*通信作者: 王玉斌 wangyubin1988999@163.com

息的内容过期的情况无法有效地过滤。为了解决系统中过期信息的问题,则需要给出信息过期的评价标准。文献[6-8]提出了应用信息计量学中的文献老化理论来描述网络信息的价值,即信息时效性随时间的变化规律。国内外在研究时间属性在推荐系统中作用已经有先行者,例如,文献[9-11]中提出:同一用户在不同时期产生的偏好行为不能同等对待,发生越晚的偏好行为能更好地体现用户当前的兴趣,即用户的偏好行为是有时效性的,同时文献[9]提出了以遗忘函数、线性或者非线性的退化函数来构建用户行为的时效性量化模型,通过与推荐算法的结合来提高推荐效果。此外,在文献[12]中作者介绍了利用时间属性的混合推荐算法;文献[13]提出了用本体论模型与项目的时效性的曲线来校正推荐结果的推荐算法。以上这些结合时间属性的推荐算法主要是考虑用户偏好行为的时效性,但是并没有考虑到信息本身的时效性问题,对传统推荐系统中过期信息的问题还是无法解决。

本文为了解决传统推荐系统中过期信息的问题,结合信息老化分析方法,提出一种基于信息老化的协同过滤推荐算法。利用系统中过往的点击记录,建立项目的时效性评价模型,预测项目当前时刻的价值大小;并结合基于项目协同过滤^[14-18]得到的相似度等数据,综合考虑用户的兴趣和项目的价值来发现项目的最近邻居,找到时效性较高的项目集合,最后进行评分预测与结果优化,从而进行高时效性的推荐。实验结果表明,该算法在充分考虑用户兴趣的前提下提高了推荐结果的时效性。本文第2节提出一种分析挖掘系统点击记录并建立信息时效性评价模型的方法;第3节提出一种基于信息老化的协同过滤推荐算法;第4节利用公开数据集进行了实验并对实验结果进行分析;最后给出结论。

2 信息时效性模型

信息都有自己的规律^[11],从信息的诞生,经历高潮,衰退以及最后消失,都是遵守着这个规律。为了描述这个规律,本文在文献[12]的基础上引入生命周期、半衰期两个概念,定义如下:

定义 1(生命周期 T_s) 信息的生命周期是指信息发布时刻 t 到信息不再被点击或评论为止的时刻 $t + T_s$ 之间的时间段 T_s ,即信息有影响力的这段时间。

定义 2(半衰期 T_h) 信息的半衰期是指信息自发布的时刻 t 开始到信息的影响力下降到一半的时刻 $t + T_h$ 之间的时间段 T_h ,即信息的影响力减半的时间。

由于网络信息是信息的一种,所以网络信息老化继承了信息老化的经典模型——负指数的模型^[13],如式(1)所示。

$$C(t_i, t_f) = e^{-a(t_i - t_f)} \quad (1)$$

其中 t_f 表示信息发布时间, t_i 表示当前的时间, $C(t_i, t_f)$ 表示信息在 t_i 时刻的影响力大小, a 代表的是信息的老化率系数。

为了说明式(1)能够描述出信息在生命周期 T_s 中各个时刻的影响力的大小,本文将所有用户对信息的所有偏好行为作为考察集 S ,其中偏好行为可以是用户的点击、评论或者评分等操作,本文把集合 S 以信息为单位进行子集划分,即所有用户对某条信息的偏好行为将被划分为一个子集,从而将集合 S 划分成一系列的子集 s 。另外,式(1)主要是为了定量描述信息在生命周期 T_s 中的每个时刻的影响力的大小。为了更好地说明,本文对信息的影响力做如下的定义。

定义 3(信息的影响力 $C(t_i, t_f)$) 信息的发布时间 t_f , 信息的生命周期 T_s , 当信息在 t_i 时刻,信息在时间段 $(t_i, t_f + T_s)$ 被点击或者评论的数量 $Ct(t_i, t_f + T_s)$ 与在整个生命周期 T_s 内被点击的数量 $Ct(T_s)$ 之比,具体如式(2)所示。

$$C(t_i, t_f) = (Ct(t_i, t_f + T_s) / Ct(T_s)) \times 100\% \quad (2)$$

根据上面的描述,只要知道当前时刻 t_i (假设本信息发布时间 t_f , 生命周期 T_s), 根据式(1), 可以快速地估计出信息在当前时刻的影响力大小,但是这必须是在知道信息的老化率系数 a 的基础之上。为了得到信息老化率 a , 通过化简式(1)可得式(3),具体如下:

$$a = -\ln(C(t_i, t_f)) / (t_i - t_f) \quad (3)$$

$$a = -\ln(C(T)) / T \quad (4)$$

根据式(3),如果能够得到一个信息在经过时间段 $t_i - t_f$ 后还拥有的影响力 $C(t_i, t_f)$ 的值就可以计算出信息的老化率系数 a 。为了方便说明,本文把式(3)中的 $t_i - t_f$ 替换为 T 得到式(4),其表示每条信息在发布后经历过时间段 T 后还拥有的影响力的大小。式(2)可计算出信息在时刻 t_i 拥有的影响力大小 $C(t_i, t_f)$,但式(2)考察的目标是单条信息 s ,存在很大的片面性。为解决上述问题,本文通过式(5),可计算 S 中的子集 s 对应 $T = t_i - t_f$ 的平均值,记为 \bar{T} 。

$$\bar{T} = \sum_{s \in S} T / |S| \quad (5)$$

其中 $|S|$ 为考察集合 S 中子集的个数,把式(5)与式(4)结合可以得到式(6):

$$a = -\ln(C(T)) / \bar{T} \quad (6)$$

在式(6)中, \bar{T} 由式(5)计算出, 并且此时 $C(T)$ 为一个定值, 这样就可以求出信息老化系数 a 。所以当本文选择 $C(T) = 1/2$, 即 $T = t_t - t_f$ 实际是 s 的半衰期 T_h 。把以上这些信息代入式(6), 可以得到式(7)。

$$a = -\ln(0.5) / \bar{T}_h \quad (7)$$

通过式(7), 发现只需要得到每个子集 s 的半衰期 T_h 就可以很方便计算得到老化系数 a 。

在信息老化中得到半衰期^[11,12], 主要有两种方式: 其一, 根据半衰期的定义通过统计的方法得到; 其二, 在信息老化中由巴尔顿-凯普勒老化方程、布鲁克斯老化方程可以求出, 但是这两个方程需要的参数也是通过统计方式得到的, 并且这些参数的设定都有着文献老化的特征, 对于信息老化并不合适, 因此本文选用方式一。

根据半衰期的定义, 半衰期等于信息最中间点击时刻减去信息的发布时刻, 而本文有单个信息点击量的集合 s , 这很容易在其中找出中间点击的时刻 t_0 , 信息的发布的时刻 t_f , 所以信息的半衰期为 $T_h = t_0 - t_f$ 。通过这种方式就可以找到对于每个信息的点击的半衰期的值, 最后根据式(7)计算出信息的老化系数 a 。

3 高时效性推荐方法

3.1 方法设计思想

为了解决传统推荐算法无法过滤过期信息的情况, 本文结合上文提出的时效性评价模型, 综合考虑用户兴趣与信息价值, 通过改变选取项目最近邻居选取方法以及过滤推荐结果, 来提高推荐的时效性。

3.2 建立数据模型

(1) 用户集合, 用集合 $U = \{u_1, u_2, \dots, u_n\}$ 表示 n 个用户的集合。

(2) 网络信息的集合, 即需要处理的所有网络信息或服务的集合, 用 $I = \{i_1, i_2, \dots, i_m\}$ 表示 m 个信息的集合。

(3) 用户-项目偏好矩阵, 即 U 中的用户对 I 中的项目的评分矩阵, 这些数据来源于用户对信息的反馈或评价的记录。这些记录可能是用户的评分, 或者是根据用户的历史行为计算得出, 能够反映出用户对项目的兴趣程度, 这里用矩阵 $R_{n \times m}$ 来表示。

3.3 近邻选取改进方法

(1) 基于项目相似度计算^[3,4] 基于项目 i 与 j 的评分相似度计算, 通常有 3 种计算方式, 余弦相似度计算方法, 修正的余弦相似度计算方法和相关相似性的计算方法。余弦相似度是协同过滤算法中计

算相似度的最基本的方法, 而修正的余弦相似度和相关相似性(主要是皮尔森相关系数)考虑了用户自身评价标准, 由于本文重点关注信息时效性对推荐的影响, 并不需要分析 3 种不同相似度的计算方法的影响, 所以本文采用基础的余弦相似度来计算项目 i 与 j 的相似度。项目 i 与 j 的相似度记为 $\text{sim}(i, j)$, 用户 u 对项目 i 的评分记为 r_{ui} :

$$\text{sim}(i, j) = \sum_{u=1}^m r_{ui} \cdot r_{uj} / \sqrt{\sum_{u=1}^m (r_{ui})^2 \cdot \sum_{u=1}^m (r_{uj})^2} \quad (8)$$

(2) 近邻选取改进 本算法通过在选取项目的最近邻居时, 融入信息的时效性, 全局考虑用户的兴趣与信息的时效性在推荐中的作用, 能解决传统推荐算法在推荐过程中只考虑用户兴趣, 而有可能推荐出过期无用的信息的问题。

在推荐过程中通过式(8), 可以得到项目 i 与其它项目的相似度。传统的基于项目的协同过滤算法主要是挑选与项目 i 相似度排序靠前的 N 个项目作为项目 i 的一个近邻集合。但是由于传统的算法, 只是考虑用户的兴趣, 没有对信息进行过滤, 推荐出来的信息不能保证质量。基于以上的考虑, 本算法在选取项目最近邻的时候, 根据用户兴趣与信息的时效性的综合评价, 对项目 i 进行最近邻的选取。对项目 i 与项目 j 用加权的方式重新计算其有效权值, 公式如式(9):

$$st(i, j) = \lambda \cdot \text{sim}(i, j) + (1 - \lambda) \cdot C(t, t_j) \quad (9)$$

其中 $\text{sim}(i, j)$ 是项目 i 与项目 j 的相似度; $C(t, t_j)$ 是项目 j 在当前时间的时效性, 其值范围由上文可知是在 $(0, 1)$ 之间的; λ 是融合参数, 值在 $(0, 1)$ 之间, 当 λ 越大推荐系统对用户的兴趣的关注就会更大, 相反就会对信息的时效性要求更高; λ 在不同的应用环境, 设定是不同的, 主要根据经验与交叉实验的方式进行确定, 本文在这里取为 0.5, 即认为时效性与用户兴趣是等同的效果。通过以上的计算, 可以得到项目 i 与其它项目的加权重, 最后在其中选择权重值最大的 N 个项目作为项目 i 的近邻集合 I_{Ni} 。

(3) 评分预测 根据项目 i 的时效性加权的最近邻集合 I_{Ni} , 可以得出用户 u 对其未评分的项目的评分 $p_{u,i}$, 其加权平均值预测公式如式(10), 其中 $r_{u,j}$ 表示用户 u 对项目 j 的真实评分值。

$$p_{u,i} = \frac{\sum_{j \in I_{Ni}} \text{sim}(i, j) \cdot r_{u,j}}{\sum_{j \in I_{Ni}} \text{sim}(i, j)} \quad (10)$$

3.4 推荐结果过滤方法

通过以上的计算, 得到对用户评分预测最高的

前 $2N$ 个项目集合 I_{rec} , 为了实现对推荐结果的优化, 选取时效性更高的信息推荐给用户, 在这里对集合 I_{rec} 中的元素再一次进行筛选, 对集合 I_{rec} 中的项目进行相似度与时效性的结合, 这样能够兼顾信息的时效性与用户的兴趣, 结合公式如下:

$$Cp(x, i) = C(t_i, t_f) \cdot p_{x,i} \quad (11)$$

其中 $Cp(x, i)$ 为用户 x 对项目 i 的最后预测评分, 从中挑选出评分排名前 N 个项目形成最后推荐集合 $I_{\text{real}} = \{i_1, i_2, \dots, i_n\}$ 。

3.5 算法复杂度分析

(1) 时间复杂度 本算法相对于传统的基于项目的协同过滤算法增加式(9)的计算过程以及集合 I_{rec} 的过滤过程, 其中式(9)的计算的时间复杂度为 $O(m)$, m 为项目的个数; 而在集合 I_{rec} 的过滤过程中, 本算法对集合 I_{rec} 进行了重新权值计算与排序, 其中权值计算的时间复杂度为 $O(2N)$, 排序部分的时间复杂度为 $O(2N \times \lg(2N))$, 其中 N 为定值且较小。已知传统的基于项目的协同过滤算法的时间复杂度由两部分组成: (1) 计算项目之间的相似度的时间复杂度为 $O(m^2)$, (2) 评分预测阶段的时间复杂度是 $O(k \times n)$, n 表示用户数量, k 表示所有用户尚未评价的项目数量的均值, 由于在绝大多数系统中用户对项目的评分是相当少的, 所以可以假设 k 等于 m , 此时若 m 大于 n , 则时间复杂度为 $O(m^2)$, 相反时间复杂度即为 $O(m \times n)$ 。由以上可知, 本算法新增的计算时间并不影响基于项目的协同过滤算法的时间复杂度量级。

(2) 空间复杂度 本文提出的算法需要增加存储项目时间属性的空间, 以及式(9)组合部分的计算空间, 其中存储时间属性需要耗费 m 个地址空间, 而式(9)的空间复杂度为 $O(m)$, 传统推荐算法中的项目之间的相似度的空间复杂度为 $O(m^2)$, 因此本算法的空间复杂度有所增长, 但并不影响传统基于项目的协同过滤的空间复杂度的量级。

4 实验

本节通过实验验证该文提出的方法可行性, 包括数据集介绍、评价指标以及实验结果和分析。其实验环境为: 8 G 内存, 2.60 GHz 双核 CPU, Windows 7 操作系统, Java1.6 开发语言, My Eclipse8.5 集成环境, MySQL5.5 数据库。

4.1 数据集介绍

本文使用的是 2006 年 Netflix^[19] 对外发布了一个电影评分数据集。Netflix 中有 480189 个用户在 1988 年 10 月到 2005 年的 12 月期间对 17770 部电影的 103297638 个评分, 所有的评分值都是在 1 到

5 之间的整数值, 其中越高的分数代表用户对相应的电影的喜欢度更高。Netflix 中有具体每个电影的发布时间, 以及被用户进行评论的记录, 包括被评论的时间, 用户 ID, 电影 ID, 以及用户评分。

4.2 评价方式

本文主要通过数学统计及曲线拟合的方式对时效性进行验证与说明, 并通过 SSE(The Sum of Squares due to Error), R-square(Coefficient of determination), Adjusted R-square(Degree of Freedom Adjusted R-square) 与 RMSE(Root Mean Squared Error) 等指标进行度量。其中 SSE 是误差平方和, 越接近 0, 曲线的拟合效果越好, R-square 越接近 1, 曲线的拟合效果越好, Adjusted R-square 是调整自由度以后的残差的平方, 数值越接近 1, 曲线的拟合效果越好, RMSE 表示根的均方误差, 越接近 0, 效果越好。

本文对推荐结果的时效性高低的判断, 主要是通过推荐结果求其时效性平均值, 从而进行衡量, 公式如下:

$$Tm = \sum_{i \in I} e^{-a(t_i - t_{i,f})} / |I| \quad (12)$$

其中 I 为符合某项条件的电影集合, t_i 为当前年份, $t_{i,f}$ 为电影 i 发布的年份, a 为老化系数, Tm 是每部电影时效性的平均值。通过对推荐结果进行时效性均值的求取, 可以看出结果的时效性的高低, 当 Tm 的值越高, 证明推荐结果越来越好。

4.3 实验结果与分析

(1) 时效性验证实验: 通过对 Netflix 数据集近十年来的数据进行处理, 对其中的电影按照发布时间, 分成了 6 个时间段。并且对这 6 个时间段发布的电影进行信息统计并且求出在发布后每年被评论数量的平均值, 可以得到这 6 个时间段发布的电影在以后的时间段被评论的情况如表 1 所示, 其中列表示电影发布的年份, 行表示某一年中电影被评论的均值:

表 1 电影被评论数量的平均值

年份	时间段					
	2000	2001	2002	2003	2004	2005
2000	23.17	222.01	493.37	677.99	1468.82	2433.40
2001		46.97	593.78	918.01	1763.66	2793.99
2002			145.31	1498.76	2128.37	2823.62
2003				573.06	3503.34	3458.08
2004					1450.17	5831.14
2005						3874.39

从表 1 中,可以看出每个分类的电影在被发布后,并没有随着时间的推移而出现评论量下降的情况。通过分析可知,这主要是由于数据集本身的用户量在这几年不断地增加,所以造成了统计结果在环比上的变化不符合预期。但是对 2005 年的整个年份的电影的评论行为进行统计,发现不同年份的电影的评论量是随着电影的发布年限增长而降低的,与预期相符合。从这可以初步认为信息的价值是随着信息发布时间增长而下降的,同时由于实验所用的数据集是一个电影的数据集,其时效性并不强,如果研究的主体是新闻等高时效性的信息种类,这种趋势会更加明显。

为了验证负指数模型的正确性,本文取 2005 年的全年的统计数据,即表 1 的最后一列数据进行曲线拟合,其中与模型 $y = a \cdot \exp(b \cdot x)$ 的拟合结果如图 1 曲线 1 所示,与模型 $y = a \cdot \exp(b \cdot x) + c \cdot \exp(d \cdot x)$ 拟合的结果如图 1 曲线 2 所示,两次拟合的结果参数如表 2 所示。

表 2 曲线拟合结果参数

拟合曲线	SSE	R-square	Adjusted	RMSE
图 1 曲线 1	4.592e-2	9.178e-1	8.972e-1	1.072e-1
图 1 曲线 2	8.992e-5	9.998e-1	9.996e-1	6.705e-3

依据图 1 与表 2 可以看出统计出的数据与负指数模型的拟合度很高,其中 R-square, Adjusted R-square 都接近 1, RMSE 接近 0, 这些说明信息的时效性是随着发布的时间的推移而下降,并且符合负指数的模型。

(2)为了验证改进过后方法的效果,本文通过对原始数据集抽取 600 部电影,与 500 个用户的信息形成新的数据集。同时,为了验证实验中的二次改进,本实验分别对二次改进进行推荐列表的时效性均值求取。

对于过滤推荐结果,首先本文使用传统的基于

项目协同过滤算法可以得到一个推荐列表 RL_1 ,再通过时效性过滤得到推荐结果 RL_2 ,这时老化系数 a 取 0.05,最后对 RL_1 与 RL_2 挑选排名靠前的 N 个通过式(12)计算时效性均值,结果如图 2 所示,其横轴是结果的推荐个数,纵轴为时效性均值。

对于改进近邻选取,本文通过基于项目的协同过滤可以得到推荐结果集 RL_3 ,再通过改进后的推荐算法可以得到推荐结果 RL_4 ,其中在老化系数设为 0.1,最后通过式(12)计算时效性均值如图 3 所示,其中横坐标为选择推荐结果的个数,纵坐标为时效性均值。

从图 2,图 3 可以看出,在推荐时用改进后的算法,推荐结果的时效性明显高于未改进之前的效果。再综合考虑前面实验结论可以得出,改进后的方法综合考虑用户兴趣与信息价值因素,从而推荐的结果更加有价值,能够屏蔽一些过期无用的信息。此外,图 3 中的对时效性的改进效果没有图 2 中的效果明显。通过分析可知,通过优化推荐结果的方式在对时效性的提升方面更加显著,但是通过改进推荐算法近邻选择方式能够更加全面考虑用户兴趣与时效性。

从上面分析可知本文提出算法能提高推荐结果的时效性,理论上适用于所有信息,对不同时时效性的信息,可以通过修改算法的参数来进行调整。但是本文是针对用户更加喜欢时效性高的信息的假设而设定的,对于那些时效性不太敏感的信息,即使通过参数的调整,最终的推荐效果并不会会有明显的提高,因此本文的方法对那些时效性敏感的信息的推荐效果更明显。

5 总结

本文根据系统中用户对信息的评价记录,建立信息时效性模型,能够定量描述信息在当前时刻的价值大小,并且在实验中,通过对公开数据集的统计与曲线拟合的方式验证了时效性模型的准确性。同时本文提出一种基于信息老化的协同过滤推荐算法,本算法综合考虑信息的时效性与用户的兴趣,

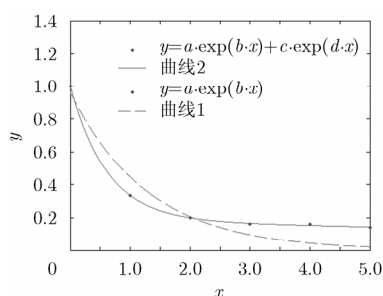


图 1 曲线拟合结果

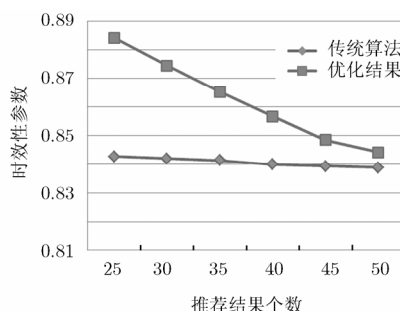


图 2 推荐结果优化

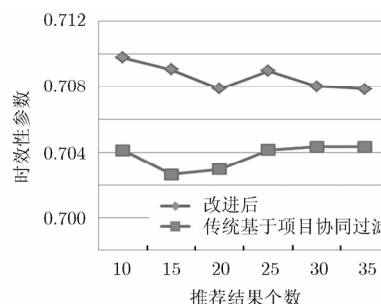


图 3 推荐算法改进

通过改变项目的最近邻居的选取规则,寻找时效性高的邻居集合,最后根据近邻集合的偏好进行推荐;本算法还结合时效性模型对推荐结果进行过滤。通过在公开数据集上的实验,验证了本算法能够有效改进推荐结果的时效性,能在一定程度上解决传统推荐算法无法屏蔽过期信息的问题。同时根据实验可以看出,对结果进行过滤在整个提高推荐的时效性过程中起着更大的作用。

参 考 文 献

- [1] 孟祥武, 胡勋, 王立才, 等. 移动推荐系统及其应用[J]. 软件学报, 2013, 24(1): 91-108.
Meng Xiang-wu, Hu Xun, Wang Li-cai, et al. Mobile recommender systems and their applications[J]. *Journal of Software*, 2013, 24(1): 91-108.
- [2] 谢海涛, 孟祥武. 适应用户需求进化的个性化信息服务模型[J]. 电子学报, 2011, 39(3): 643-648.
Xie Hai-tao and Meng Xiang-wu. A personalized information service model adapting to user requirement evolution[J]. *Acta Electronica Sinica*, 2011, 39(3): 643-648.
- [3] 黄武汉, 孟祥武, 王立才. 移动通信网中基于用户社会化关系挖掘的协同过滤算法[J]. 电子与信息学报, 2011, 33(12): 3002-3007.
Huang Wu-han, Meng Xiang-wu, and Wang Li-cai. A collaborative filtering algorithm based on users' social relationship mining in mobile communication network[J]. *Journal of Electronics & Information Technology*, 2011, 33(12): 3002-3007.
- [4] 徐风苓, 孟祥武, 王立才. 基于移动用户上下文相似度的协同过滤推荐算法[J]. 电子与信息学报, 2011, 33(11): 2785-2789.
Xu Feng-ling, Meng Xiang-wu, and Wang Li-cai. A collaborative filtering recommendation algorithm based on context similarity for mobile users[J]. *Journal of Electronics & Information Technology*, 2011, 33(11): 2785-2789.
- [5] 张波, 向阳, 王坚, 等. 一种基于语义可理解的信息过滤算法[J]. 电子与信息学报, 2010, 32(10): 2324-2330.
Zhang Bo, Xiang Yang, Wang Jian, et al. Information filtering algorithm based on semantic understanding[J]. *Journal of Electronics & Information Technology*, 2010, 32(10): 2324-2330.
- [6] 杨思洛, 邱均平. 网络信息资源的老化规律研究——以网络引文的可追溯性为视角[J]. 情报学报, 2012, 31(6): 619-629.
Yang Si-luo and Qiu Jun-ping. Study on obsolescence of Web information resources——a perspective from the traceability of Web citations[J]. *Journal of the China Scientific and Technical Information*, 2012, 31(6): 619-629.
- [7] 马费成, 望俊成. 信息生命周期研究述评(I)——价值视角[J]. 情报学报, 2010, 29(5): 939-947.
Ma Fei-cheng and Wang Jun-cheng. A literature review of studies on information lifecycle(I)——the perspective of value[J]. *Journal of the China Scientific and Technical Information*, 2010, 29(5): 939-947.
- [8] 黄利平. 基于引文分析的文献老化研究——以图书情报和管理学为例[J]. 情报杂志, 2011, 30(10): 30-35.
Hang Li-ping. Literature aging research based on citation analysis——an analysis in library and information science and management science[J]. *Journal of Intelligence*, 2011, 30(10): 30-35.
- [9] 印桂生, 崔晓晖, 马志强, 等. 遗忘曲线的协同过滤推荐模型[J]. 哈尔滨工程大学学报, 2012, 33(1): 85-90.
Yin Gui-sheng, Cui Xiao-hui, Ma Zhi-qiang, et al. Forgetting curve-based collaborative filtering recommendation model[J]. *Journal of Harbin Engineering University*, 2012, 33(1): 85-90.
- [10] 于洪, 李转运. 基于遗忘曲线的协同过滤推荐算法[J]. 南京大学学报(自然科学版), 2010, 46(5): 520-527.
Yu Hong and Li Zhuan-yun. A collaborative filtering recommendation algorithm based on forgetting curve[J]. *Journal of Nanjing University (Natural Sciences)*, 2010, 46(5): 520-527.
- [11] 印桂生, 崔晓晖, 马志强, 等. 基于差异时效的服务评价模型[J]. 西南交通大学学报, 2012, 47(4): 652-661.
Yin Gui-sheng, Cui Xiao-hui, Ma Zhi-qiang, et al. Web services evaluation model based on variant time utility[J]. *Journal of Southwest Jiaotong University*, 2012, 47(4): 652-661.
- [12] Farman U, Ghulam S, Lee Sung chang, et al. Hybrid recommender system with temporal information[C]. Proceedings of International Conference on Information Networking, Bali, 2012: 421-425.
- [13] Blanco-Fernandez Y, Lopez-Nores M, Pazos-Arias J 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(8): 1385-1397.
- [14] Li D and Lv Q. Interest-based real-time content recommendation in online social communities[J]. *Knowledge Based System*, 2012, 28(1): 1-12.
- [15] 王玉祥, 乔秀全, 李晓峰, 等. 上下文感知的移动社交网络服务选择机制研究[J]. 计算机学报, 2010, 33(11): 2126-2135.
Wang Yu-xiang, Qiao Xiu-quan, Li Xiao-feng, et al. Research on context-awareness mobile SNS service selection mechanism[J]. *Chinese Journal of Computers*, 2010, 33(11): 2126-2135.
- [16] Rawat R, Nayak R, and Li Y F. Effective hybrid recommendation combining users-searches correlations using tensors[C]. Proceedings of the APWeb 2011, Berlin, Springer-Verlag, 2011: 131-142.
- [17] Deng Y, Wu Z, and Tang C. A hybrid movie recommender based on ontology and neural networks[C]. Proceedings of the GREENCOM-CP SCOM 2010, IEEE Computer Society, Hangzhou, 2010: 846-851.
- [18] Kim H K, Hee Y, and Gu J C. Commenders: a recommendation procedure for online book communities[J]. *Electronic Commerce Research and Applications*, 2011, 10(5): 501-509.
- [19] Netflix. Netflix prize. URL <http://www.netflixprize.com>, 2006, 10.

王玉斌: 男, 1988年生, 硕士生, 研究方向为信息计量与推荐系统。

孟祥武: 男, 1966年生, 博士, 教授, 博士生导师, 主要研究领域为网络服务、智能信息处理、通信软件。

胡勋: 男, 1985年生, 博士生, 研究方向为推荐系统、个性化服务。