·先进计算与数据处理·　文章编号：1000-3428(2015)01-0065-06　文献标识码：A　中图分类号：TP312

基于用户兴趣三维建模的个性化推荐算法

王冰怡，刘　杨，聂长新，田　萱

（北京林业大学信息学院，北京 100083）

摘　要：针对推荐系统中用户的个性化需求，提出一种基于用户兴趣三维建模的个性化推荐算法。通过分析用户行为数据，从兴趣广度、兴趣深度和兴趣时效 3 个角度分析用户的兴趣构成，对用户兴趣进行三维建模，并在此基础上，逐步添加维度，设计用户之间兴趣相似度的三级计算方法。在真实推荐系统数据集上的实验结果表明，用户兴趣三维模型比一维模型、二维模型更能准确地表征用户兴趣，基于用户兴趣三维建模的个性化推荐算法能够提高个性化推荐的准确率。

关键词：个性化推荐；用户兴趣三维建模；兴趣广度；兴趣深度；兴趣时效；用户兴趣相似度

中文引用格式：王冰怡，刘　杨，聂长新，等. 基于用户兴趣三维建模的个性化推荐算法[J]. 计算机工程，2015，41(1)：65-70.

英文引用格式：Wang Bingyi，Liu Yang，Nie Changxin，et al. Personalized Recommendation Algorithm Based on Three-dimensional User Interest Modeling[J]. Computer Engineering，2015，41(1)：65-70.

Personalized Recommendation Algorithm Based on Three-dimensional User Interest Modeling

WANG Bingyi，LIU Yang，NIE Changxin，TIAN Xuan

(School of Information Science and Technology，Beijing Forestry University，Beijing 100083，China)

【**Abstract**】To realize the personalized recommendation，an algorithm based on three-dimensional user interest modeling is presented. First，by analyzing user's behavior data，three aspects of user interest are analyzed for building three-dimensional user interest model，including interest width，interest depth and user interest timeliness. Secondly，based on the three-dimensional user interest model，the dimension is gradually added and a method for calculating the interest similarity between two users is proposed. At last，a personalized recommendation algorithm is given. Experiment uses the data set derived from real recommendation system. The results show that，three-dimensional user interest model can describe user interest more accurately than two-dimensional and one-dimensional user interest model，and the proposed personalized recommendation algorithm can improve the accuracy rate of personalized recommendation.

【**Key words**】personalized recommendation；three-dimensional user interest modeling；interest width；interest depth；interest timeliness；user interest similarity

DOI：10.3969/j.issn.1000-3428.2015.01.012

1　概述

随着互联网信息的爆炸性增长，信息超载目前已成为网络用户面临的一个严重问题，用户难以从海量的在线资源中获取与用户相关或感兴趣的信息。个性化推荐系统是解决该问题的一个有力工具，个性化推荐系统[1]能够根据用户的历史行为信息建立用户兴趣模型，利用系统中的推荐算法向用户推荐其未曾产生过行为的物品。在个性化推荐系统中，用户兴趣建模是个性化推荐的基础[2]，很大程度上决定着推荐系统的质量。

当前用户兴趣模型的表示大致分为 2 类：基于关键词的兴趣模型[3]和基于语义的兴趣模型[4]。基于关键词的兴趣模型以向量空间模型表示法[5]为主，利用用户兴趣数据的特征项及其所对应的权值来表示用户兴趣模型，例如，文献[6]针对现有的用户模型不能根据用户自身兴趣实现推荐的问题，通过建立用户集、资源集，结合用户的行为特点及权

基金项目：中央高校基本科研业务费专项基金资助项目(YX2014-19)；北京林业大学国家级大学生创新基金资助项目(201310022050)。
作者简介：王冰怡(1995－)，女，本科生，主研方向：信息检索，推荐系统；刘　杨、聂长新，本科生；田　萱(通讯作者)，副教授、博士。
收稿日期：2014-04-28　**修回日期**：2014-07-03　**E-mail**：tianxuan@ruc.edu.cn

重,提出了基于用户行为的兴趣模型;文献[7]通过用户关于某个主题的生成概率来反映用户对该主题的喜好程度,提出了基于主题模型的用户兴趣模型。基于语义的兴趣模型往往建立在本体技术之上,使用户兴趣模型更加富有语义信息,加强了用户模型表征力度,被认为是向量空间用户模型的一种完善。例如,文献[8]利用用户访问量,采用改进的相似度算法,实现用户分类建立用户兴趣模型;文献[9]基于用户知识本体和概念向量构建兴趣模型,实现了用户兴趣的个性化语义描述。

上述研究在用户建模的过程中,对用户兴趣的描述大都只停留在一维上,如利用用户对资源的访问量来表征用户兴趣。虽然这些研究从不同角度提高了对用户描述的精确度,但仍然缺乏对用户兴趣的深入分析和准确建模,系统需要收集更详细、有效的用户信息来描述用户兴趣[10]。例如用户兴趣范围会随着时间而变化,因此系统就要收集用户搜索的相关时间信息。传统的研究没有将时间因素加入到用户兴趣模型的构建中,导致用户特征描述的不准确性,最终影响了推荐效果。

为解决这一问题,本文通过从用户的搜索行为中获取用户数据,从用户兴趣广度、兴趣深度、兴趣时效这 3 个角度刻画用户兴趣三维模型,设计用户之间相似度的三级计算方法,提出基于用户兴趣三维建模的个性化推荐算法,并将该算法应用到个性化推荐系统中,找出目标用户的最近邻,根据最近邻的兴趣预测目标用户的兴趣,进而对目标用户产生动态推荐。

2 个性化推荐中的用户兴趣三维建模方法

2.1 用户兴趣数据的获取

要实现个性化推荐,系统必须知道用户的个性化需求,这就需要获取、分析用户信息,建立合适的用户兴趣模型。用户信息可以分为用户背景信息和用户访问信息[11],用户背景信息包括用户职业、年龄、居住地等,可以有效解决个性化推荐初期用户数据较少情况下无法做出精确推荐的“用户冷启动”问题;用户访问信息主要包括用户浏览行为和用户搜索行为,是建立用户兴趣模型的主要来源[12]。本文提出的用户兴趣三维模型主要从用户的搜索行为中获取用户数据。因为在用户的搜索行为中:

(1)用户搜索的目标、搜索的时间、浏览的时间、拖动滚动条的次数等行为数据能够体现不同年龄、性别、地区的用户兴趣;

(2)随着时间变迁,用户搜索目标的转移往往意味着用户兴趣在动态变化;

(3)用户对搜索目标的点击和浏览频度在一定程度上反映了用户的兴趣。

通过对用户的搜索行为进行分析,就能描述出用户在某一时间段内的兴趣,进而更准确地建立用户兴趣模型。

本文主要依据推荐系统中的搜索日志来分析用户兴趣。一般来说,搜索日志中包括用户的 ID,用户搜索的物品 i,用户每次搜索物品 i 时的时刻 t_i。用户 u 对物品 i 的搜索日志 $Log(u,i)$ 形式化定义为:

$$<Log(u,i)>::=<ID_u><i><t_1,t_2,\cdots,t_n>$$

2.2 用户兴趣三维模型

心理学上认为,人类的兴趣与其行为有着很大的关联[13]。兴趣可以由用户的行为判断出来。根据直觉可以得到用户搜索行为和用户兴趣的 3 个基本结论:

(1)用户对其搜索的物品感兴趣;

(2)用户搜索某一物品的次数越多,对该物品就越感兴趣;

(3)随着时间的推移,用户对某物品的兴趣会减少。

基于这 3 个结论,本文提出用户兴趣广度、兴趣深度、兴趣时效的概念,从搜索日志中获取体现用户兴趣的数据,并利用用户兴趣的广度(X)、深度(Y)、时效(Z)来构建用户兴趣三维模型,如图 1所示。

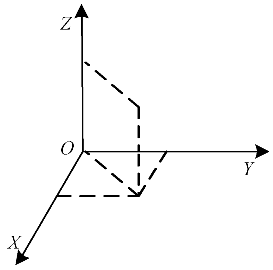

图 1 用户兴趣三维模型

2.2.1 用户兴趣广度

在用户兴趣三维模型中,本文用用户兴趣广度来描述用户的兴趣范围。在搜索日志 $Log(u, i)$ 中可以统计出用户 u 搜索的物品集合,这个集合在一定程度上描述了用户感兴趣的物品范围,用它来标记用户 u 的兴趣广度,记为 $Width(u)=\{i_1,i_2,\cdots\}$。

直觉上,用户 u,v 兴趣广度交集元素个数越多,即共同搜索的物品越多,用户 u,v 的兴趣相似度越高,在对目标用户 u 进行个性化推荐时,更倾向利用用户 v 的兴趣来预测目标用户 u 的兴趣。

根据这一直觉,定义第 1 个启发式规则:

启发式规则 1 设用户 u,v,t 的兴趣广度分别为 $Width(u)$,$Width(v)$,$Width(t)$,若 $|Width(u)\cap Width(v)|>|Width(u)\cap Width(t)|$,则用户兴趣相似度 $W_{uv}>W_{ut}$。

2.2.2　用户兴趣深度

在用户兴趣三维模型中,本文用用户兴趣深度来表示用户 u 对物品 i 的感兴趣程度。在搜索日志 $Log(u,i)$ 中,可以统计出用户 u 对物品 i 的搜索次数,这个搜索次数在一定程度上反映了用户 u 对物品 i 的感兴趣程度。用这个搜索次数作为用户 u 对物品 i 的兴趣深度,记为 $Depth_{ui}$。

直觉上,$Depth_{ui}$ 和 $Depth_{vi}$ 差距越小,即用户 u,v 对物品 i 的搜索次数越接近,用户 u,v 的兴趣相似度越高,在对目标用户 u 进行个性化推荐时,更倾向利用用户 v 的兴趣来预测目标用户 u 的兴趣。根据这一直觉,定义第 2 个启发式规则:

启发式规则 2　设用户 u,v,t 对物品 i 的兴趣深度分别为 $Depth_{ui}$,$Depth_{vi}$,$Depth_{ti}$,若 $|Depth_{ui}-Depth_{vi}|>|Depth_{ui}-Depth_{ti}|$,则用户兴趣相似度 $W_{uv}>W_{ut}$。

2.2.3　用户兴趣时效

在一般情况下,用户的兴趣会随着时间而变化。事实上,随着用户 u 搜索物品 i 的时间推移,用户 u 对物品 i 的兴趣是衰减的[14]。在用户兴趣三维模型中,本文用用户兴趣时效来表示和时间有关的兴趣衰减。在搜索日志 $Log(u,i)$ 中,可以统计出用户 u 搜索物品 i 的最后时间,记为 T_{ui}。直觉上,在当前时刻 T,若用户 u 刚刚不久前才搜索过物品 i,即 T_{ui} 距离当前时间 T 越近,用户 u 对物品 i 的兴趣时效就越大。将用户 u 对物品 i 的兴趣时效记为 $Time_{ui}$,$Time_{ui}$ 可表示为:

$$Time_{ui}=\frac{1}{1+\alpha|T-T_{ui}|} \tag{1}$$

其中,α 为平衡参数。考虑到,如果用户 u 经常搜索物品,对物品 i 来说,$|T-T_{ui}|$ 偏小,但这只能说明用户兴趣变化得快,并不意味着用户 u 对用户 i 的兴趣时效大,因此,本文定义了平衡参数 α 来消去这一差异。规定如果某用户产生搜索行为的频率很高、用户较活跃,则选取较大的 α,反之则选取较小的 α。

直觉上,$Time_{ui}$ 和 $Time_{vi}$ 差距越小,即用户 u,v 搜索同一物品 i 的时间越接近,用户 u,v 的兴趣相似度越高,在对目标用户 u 进行个性化推荐时,更倾向利用用户 v 的兴趣来预测目标用户 u 的兴趣。根据这一直觉,定义第 3 个启发式规则:

启发式规则 3　设用户 u,v,t 对物品 i 的兴趣时效分别为 $Time_{ui}$,$Time_{vi}$,$Time_{ti}$,若 $|Time_{ui}-Time_{vi}|>|Time_{ui}-Time_{ti}|$,则用户兴趣相似度 $W_{uv}>W_{ut}$。

3　用户兴趣相似度的三级计算方法

用户之间相似性可以通过 Jaccard 公式度量,本文利用并改进 Jaccard 公式来计算用户之间的相似度。Jaccard 公式表示为:

$$W_{uv}=\frac{|N(u)\cap N(v)|}{\sqrt{|N(u)|\cdot|N(v)|}} \tag{2}$$

其中,W_{uv} 表示用户 u 与用户 v 之间的兴趣相似程度;$N(u)$ 表示用户 u 搜索过的物品集合;$N(v)$ 表示用户 v 搜索过的物品集合。

上文从用户兴趣广度、兴趣深度、兴趣时效这 3 个方面来构建用户兴趣三维模型,下文将根据用户兴趣三维模型的每一维来逐步改进计算用户兴趣相似度的 Jaccard 方法。

3.1　用户兴趣相似度的第 1 级计算方法

在式(2)的基础上,本文引入用户兴趣三维模型中的第 1 维——兴趣广度来刻画用户兴趣相似度,设用户 u,v 的兴趣广度分别为 $Width(u)$,$Width(v)$,根据中启发式规则 1,用户 u,v 之间的兴趣相似度 W'_{uv} 可定义如下:

$$W'_{uv}=\frac{|Width(u)\cap Width(v)|}{\sqrt{|Width(u)|\cdot|Width(v)|}} \tag{3}$$

从式(3)可以看出,如果用户 u 和用户 v 共同产生过行为的物品数目越多,用户 u 和用户 v 之间的兴趣相似度就越高。这个公式完全沿用了 Jaccard 方法。

3.2　用户兴趣相似度的第 2 级计算方法

在式(3)的基础上,本文引入用户兴趣三维模型中的第 2 维——兴趣深度来进一步分析用户兴趣相似度。设用户 u,v 对物品 i 的兴趣深度为 $Depth_{ui}$,$Depth_{vi}$,根据上文中启发式规则 2,用户 u 和用户 v 的兴趣相似度 W'_{uv} 可改进为:

$$W''_{uv}=W'_{uv}\cdot\sum_{i\in Width(u)\cap Width(v)}\left(1+\frac{1}{1+|Depth_{ui}-Depth_{vi}|}\right) \tag{4}$$

其中,i 为用户 u 和用户 v 都搜索过的物品。从式(4)可以看出,用户 u 和用户 v 对物品 i 的搜索次数差别越小,用户 u 和用户 v 的兴趣相似度就越大。

3.3　用户兴趣相似度的第 3 级计算方法

在式(4)的基础上,本文引入用户兴趣三维模型中的第 3 维——用户兴趣时效来更进一步分析用户兴趣相似度。设用户 u,v 对物品 i 的兴趣时效分别为 $Time_{ui}$,$Time_{vi}$,根据上文中启发式规则 3,用户 u 和用户 v 的兴趣相似度 W''_{uv} 可改进为:

$$W'''_{uv} = W''_{uv} \cdot \sum_{i \in Width(u) \cap Width(v)} \left(1 + \frac{1}{1 + |Time_{ui} - Time_{vi}|}\right) \quad (5)$$

从式(5)可以看出,若用户 u 和用户 v 对物品 i 产生搜索行为的时间点越接近,用户 u,v 之间的兴趣相似度就越大。

在以上给出的三级算法中,第 3 级用户兴趣相似度在第 2 级用户兴趣相似度的基础上增加了用户兴趣时效因素,能够更准确地从不同角度分析用户之间的相似性,帮助提高用户个性化推荐效果。这将在本文的实验部分进行验证。

4 基于用户兴趣三维建模的个性化推荐

一般推荐系统预测目标用户对某一物品感兴趣程度的做法是:选取与目标用户兴趣最接近的用户群体 U,通过用户群体 U 对该物品的感兴趣程度来预测目标用户对其感兴趣程度。从物品中挑选出用户感兴趣程度 $Top(K)$ 的 K 个物品推荐给目标用户。

4.1 用户对物品的感兴趣程度计算

本文已经通过构建用户兴趣三维模型得出了用户 u 和用户 v 之间兴趣相似度 W'''_{uv},那么,目标用户 u 对未产生过行为的物品 j 的感兴趣程度可表示为:

$$P(u,j) = \sum_{v \in (S(u,K) \cap M(j))} W'''_{uv} \cdot Depth_{vj} \quad (6)$$

其中,$S(u,K)$ 表示与目标用户 u 兴趣最接近的 K 个用户,可通过对 W'''_{uv} 进行 top-N 运算求得;$M(j)$ 表示对物品 j 产生过搜索行为的用户集合;T 为当前时间。由式(6)可知,在当前时间 T,与用户 u 兴趣相似的用户 v 对物品 j 的感兴趣程度越高,用户 u 对物品 j 的感兴趣程度就越高。

4.2 算法描述

综上,可以看到,基于用户兴趣三维建模的个性化推荐算法的过程可以分为 2 步:

(1)通过 3.3 节中第 3 级计算方法得到用户 u 的最近邻集合 $S(u,K)$;

(2)通过 4.1 节方法计算用户 u 对物品 j 的感兴趣程度 $P(u,j)$,最后向用户 u 推荐 $P(u,j)$ 排名 $Top(K)$ 的物品 j 即可。

将本文算法分为 2 个阶段来描述:

阶段 1 求用户 u 的最近邻集合 $S(u,K)$

输入 用户 u,用户整体集合 $USERS$

输出 用户 u 的最近邻集合 $S(u,K)$

```
Begin:
For (v ∈ USERS) /*计算用户 u 和用户 v 的兴趣相似度*/
  { if (Width(u) ∩ Width(v) ≠ NULL)
    { W_uv' = |Width(u) ∩ Width(v)|/
  √(|Width(u)| · |Width(v)|);
     tempW1 = 0;
     tempW2 = 0;
     for (∀ i ∈ Width(u) ∩ Width(v))
       { tempW1 = tempW1 + {1 + 1/(1 + |Depth_ui -
       Depth_vi|);
        tempW2 = tempW2 + {1 + 1/(1 + |Time_ui -
        Time_vi|);}
    W_uv'' = W_uv' * tempW1;
    W_uv''' = W_uv'' * tempW2;} }
For (v ∈ USERS)//Top(K)统计 K 个 W_uv'''
  { 排序 W_uv''',统计 K 个 W_uv''',得到用户 u 的最近邻集合 S(u,K);}
End
```

该阶段算法的时间复杂度为 $O(mn)+O(m\text{lb}m)$,其中,$m=|USERS|$,是全部用户的个数;n 为用户 u 与其他用户共同搜索的物品的最大个数,即 $n=\text{Max}(|Width(u) \cap Width(v)|)$。时间复杂度的前一项 $O(mn)$ 为计算用户 u 和其他用户兴趣相似度的时间复杂度,事实上,由于用户搜索的物品个数远远小于用户个数,即 $n \ll m$,因此该阶段算法的时间复杂度可以简化为 $O(m)$。时间复杂度的后一项 $O(m\text{lb}m)$ 为堆排序 W_{uv}''' 的时间复杂度。

阶段 2 求用户 u 对物品感兴趣程度 $Top(K)$ 的物品集合 J'

输入 用户 u 的最近邻集合 $S(u,K)$,物品集合 J

输出 用户 u 对物品感兴趣程度 $Top(K)$ 的物品集合 J'

```
Begin:
For (∀ j ∈ J∩j ∉ Width(u)) /*计算用户 u 对物品 j 的感兴趣程度 P(u,j)*/
  {P(u,j) = 0;
   For (v ∈ S(u,K))
     If (j ∈ Width(v))
     P(u,j) = P(u,j) + W_uv''' * Depth_vj;}
For (∀ j ∈ J) /*Top(K)统计 K 个 P(u,j)*/
  { 排序 P(u,j),统计 K 个 P(u,j),得到向用户 u 推
  荐的物品集合 J' }
End
```

该阶段算法的时间复杂度为 $O(x|K|)+O(x\text{lb}x)$,其中,x 为用户 u 未搜索过的物品个数,最大为整个物品集合的数目 $|J|$;$|K|$ 是系统选取的最近邻的数目。时间复杂度的前一项 $O(x \times |K|)$ 为计算用户 u 对物品 j 的感兴趣程度的时间复杂度,一般

来说,系统选取的最近邻数目远远小于用户未搜索过的物品个数,即$|K|\ll x$,所以该阶段算法的时间复杂度可以简化为 $O(x)$。时间复杂度的后一项 $O(x\mathrm{lb}x)$ 为堆排序 $P(u,j)$ 的时间复杂度。

5　实验结果与分析

5.1　实验环境与数据集

个性化推荐中常用的公开数据集包括 MovieLens, BookCrossing, Jester 等,由于本文算法所需的数据来源于搜索日志,因此这些数据集并不适用于本文算法。目前没有公开的推荐系统的搜索日志数据集,已有的公开搜索日志数据集如 Sogou 日志难以用在个性化推荐系统中,因此,本文实验中采用北京林业大学数字标本馆 Web 系统的用户搜索日志,用以考察系统向用户推荐标本的准确性。

经过对数据的处理和筛选后,选择了包含467 个用户关于788个标本的 4 783 条搜索记录,每个用户至少有 10 次的搜索行为。这些数据可以比较充分地对本文提出的算法进行测试。实验中按照本文第 2 节的部分来构建用户兴趣三维模型。在实验过程中,将整个搜索日志数据集划分为 2 个部分,80% 用作训练集,主要用于构建用户兴趣模型,剩下的 20% 用作测试集,用于验证模型的实际效果。

5.2　评价标准

实验比较了基于一维、二维、三维兴趣模型的个性化推荐算法对用户兴趣描述的准确性。一维兴趣模型只用用户兴趣广度表征用户兴趣;二维兴趣模型利用用户兴趣广度和深度来表征用户兴趣,也就是常见的利用搜索次数和搜索范围来描述用户兴趣,实验以基于用户兴趣二维模型的推荐算法为参照,来验证本文提出的基于用户兴趣三维模型的推荐算法的质量;三维兴趣模型利用用户兴趣广度、深度、时效来表征用户兴趣。分别根据用户兴趣的一维、二维、三维模型来计算用户兴趣相似度,产生最近邻,预测用户的兴趣,产生推荐物品,比较推荐的准确性。

本文采用 2 种指标来评价实验结果:平均绝对值误差(Mean Absolute Error, MAE)[15]和推荐精度(*Precision@N*)[16]。

平均绝对值误差方法通过计算预测的用户感兴趣物品与实际的用户感兴趣物品之间的误差,来表示预测的准确性,如式(7)所示。MAE 值越小,说明推荐越精确。MAE 值计算公式如下:

$$M_{\mathrm{MAE}} = \frac{\sum_{i=1}^{N} |P_i - Q_i|}{N} \tag{7}$$

其中,N 为测试集中物品的个数;P_i 是对物品 i 的预测感兴趣程度;Q_i 是实际感兴趣程度。

Precision@N 通过计算预测推荐物品与实际推荐物品的相关程度来反映推荐的精度,该值越大,精度越高,推荐的质量就越好,如式(8)所示。

$$Precision@N = \frac{|Prediction(N) \cap Data(test)|}{N} \tag{8}$$

其中,N 为推荐给目标用户的物品个数;$Prediction(N)$ 表示通过训练集对目标用户进行兴趣预测后,推荐给目标用户的物品集合;$Data(test)$ 表示测试集中实际应该推荐给目标用户的物品集合。

5.3　结果分析

基于用户的个性化协同过滤推荐算法中,最近邻数目的选取通常会影响计算结果,最近邻数目越多,精度越高。在最近邻用户数 K 分别取 4,8,12,16,20 时,对基于 3 种用户兴趣模型的个性化推荐算法进行 MAE 值进行比较,实验结果如图 2 所示。

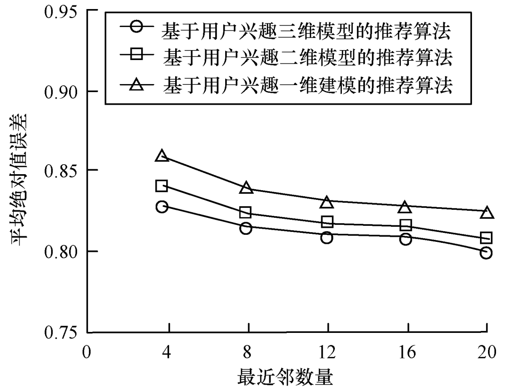

图 2　3 种个性化推荐算法的 MAE 值对比

由图 2 可见,首先,随着最近邻用户数 K 的增加,基于 3 种不同兴趣模型的推荐算法 MAE 值都降低,即推荐结果都更精确。这是因为随着最近邻数目的增加,能够体现用户兴趣的数据增多,构建的用户兴趣模型就更精确。其次,在任意 K 值下,随着建模兴趣维度的增加,MAE 值随之降低,说明建模时每增加用户兴趣的一个维数,对用户兴趣的描述就更准确些,证明本文提出的用户相似度三级计算方法是科学的,能够逐步提高推荐精确度。并且在取同样 K 值条件下,基于用户兴趣三维建模的推荐算法都取得了最低的 MAE 值,这是因为基于用户兴趣三维建模的个性化推荐算法从用户的兴趣广度、深度、时效 3 个方面来构建用户兴趣模型,较另外两种兴趣模型而言,对用户兴趣的表征更精确。同时,也说明本文提出的算法推荐质量高于常见的基于用户兴趣二维模型算法的推荐质量。

基于3种不同兴趣模型的个性化推荐推荐算法的 *Precision@ N* 值的比较结果如图3所示。可以看出,随着推荐资源个数的增加,3种基于不同兴趣模型推荐算法的推荐精度都在逐步增加,基于用户兴趣三维建模的算法始终比基于兴趣一维和二维建模算法取得更大的推荐精度,在推荐精度为15时,与另外2种推荐算法差距达到最大。与常见的基于用户兴趣二维模型的推荐算法相比,用户兴趣三维模型更能够准确地表征用户兴趣,在实际应用中,基于用户兴趣三维建模的个性化推荐算法具有更高的使用价值。

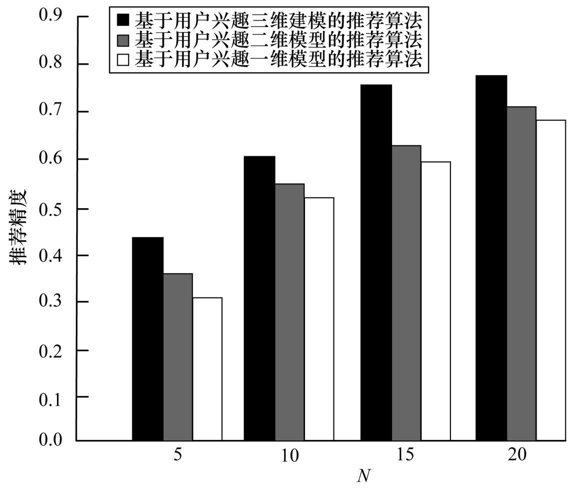

图3 3种算法的推荐精度对比

6 结束语

个性化推荐系统的目的是在应用领域预测用户需求兴趣并向用户推荐可用的信息。本文针对大多数推荐算法对用户兴趣描述不准确的问题,提出了一种基于用户兴趣三维建模的个性化推荐算法。通过分析用户行为数据,从用户兴趣广度、深度、时效3个维度来构建用户兴趣三维模型,使用户兴趣模型更能准确地表征用户的兴趣,基于此提出了用户兴趣相似度的三级计算方法。实验结果证明,用户兴趣三维模型具有一定的科学性,能够更准确地表征用户兴趣,与传统基于用户兴趣二维模型的推荐算法相比,基于用户兴趣三维建模的个性化推荐算法能够提高推荐的准确性,将其应用到推荐系统中,能够有效地提高系统的推荐质量。

参考文献

[1] Wei Chen, Hsu W, Lee M L. Modeling User's Receptiveness over Time for Recommendation[C]//Proceedings of the 36th ACM SIGIR Conference. Dublin, Ireland:[s. n.],2013:373-382.

[2] 王巧容,赵海燕,曹 健. 个性化服务中的用户建模技术[J]. 小型微型计算机系统,2011,32(1):39-46.

[3] 陈文涛,张小明,李舟军. 构建微博用户兴趣模型的主题模型的分析[J]. 计算机科学,2013,40(4):127-135.

[4] 蒲国林,杨清平,王 刚,等. 基于语义的个性化用户兴趣模型[J]. 计算机科学,2008,35(7):181-184.

[5] 石 林,徐 飞,徐守坤. 基于用户兴趣建模的个性化推荐[J]. 计算机应用与软件,2013,30(12):211-214.

[6] 王微微,夏秀峰,李晓明. 一种基于用户行为的兴趣度模型[J]. 计算机工程与应用,2012,48(8):148-151.

[7] 邱云飞,王琳颖,邵良杉,等. 基于微博短文本的用户兴趣建模方法[J]. 计算机工程,2014,40(2):275-279.

[8] 蒋秀林,谢 强,丁秋林. 基于领域本体的用户模型的研究[J]. 计算机应用研究,2012,29(2):606-608.

[9] 吴 蓉,丁二玉,骆 斌. 基于加权本体的个性化语义搜索[J]. 计算机工程与设计,2008,29(19):5051-5053.

[10] 陈 婷,韩伟力,杨 珉. 基于隐私保护的个性化推荐系统的研究与实现[J]. 计算机工程,2009,35(8):283-285.

[11] 刘滨强. 移动环境下的个性化推荐用户兴趣建模研究[D]. 北京:北京邮电大学,2009.

[12] 杨继萍,王 跃,高雪松. 个性化流媒体服务中基于行为分析的用户兴趣建模[J]. 计算机应用与软件,2011,28(8):247-250.

[13] Li Jun, Zhang Peng. Mining Explainable User Interest from Scalable User Behavior Data[J]. Procedia Computer Science,2013,17:789-796.

[14] 康雨洁. 基于协同过滤的个性化社区推荐方法研究[D]. 合肥:中国科学技术大学,2001.

[15] 邸佳奇,王霓虹. 一种基于GridGIS的增量式协同过滤算法[J]. 计算机科学,2013,40(12):219-222.

[16] Xu Guandong, Gu Yanhui, Zhang Yanchun. Toast:A Topic-oriented Tag-base Recommender System[C]//Proceedings of the 12th International Conference on Web Information System Engineering. Sydney, Australia:[s. n.],2011:158-171.

编辑 金胡考