

情境感知的移动阅读个性化推荐算法研究^①

谢修娟^{②*} 莫凌飞^{**} 李香菊^{*} 操凤萍^{*}

(^{*} 东南大学成贤学院计算机工程系 南京 210088)

(^{**} 东南大学仪器科学与工程学院 南京 210018)

摘 要 针对移动环境下资源的个性化推荐问题,提出一种结合情境和协同过滤的移动阅读推荐算法。该算法汇聚了两阶段的情境感知思想,首先使用朴素贝叶斯方法推算用户某情境下偏好度最高的资源类别,然后对该类别下的资源,通过情境相似度的计算,过滤出仅当前情境或者与当前情境最相似的情境的“用户-资源”二维评分模型,再运用传统的基于用户的协同过滤算法产生推荐列表。抓取新浪博文进行实验测试,结果表明相同条件下所提出算法的平均绝对偏差值比其他相关算法要低,具有更高的推荐质量。

关键词 情境感知,协同过滤,移动推荐,贝叶斯方法

0 引 言

随着移动互联网的成熟,以及移动智能终端的普及,越来越多的用户开始倾向于能够随时、随地进行移动阅读,然而“移动信息过载”问题^[1]使得用户难以在海量数据中快速找到符合自己需求的信息资源,致使移动阅读中资源的使用率和用户体验受限。个性化推荐技术依据用户的浏览历史、评价等日常信息,智能地筛选出少量典型的资源推送给用户,能有效缓解这一问题。

现有的主流推荐技术包括基于内容的推荐、基于规则的推荐和协同过滤推荐三类^[2-4]。协同过滤推荐是使用最为广泛的一种方法,该算法是根据历史数据计算目标用户/资源与全体用户/资源的相似度,建立近邻用户/资源,再由近邻集预测用户对未知资源的评分从而产生推荐集。

传统的推荐算法属于二维推荐机制,在推荐过程中仅涉及用户和资源两个维度。而移动环境下,阅读需求具有很强的情境敏感特性,同一个用户在

不同时间或不同地点,阅读需求会存在差异,因此,个性化推荐时需要考虑情境因素。基于情境感知的个性化推荐也成为一个新研究领域,受到了国内外学者们的普遍关注。Lee 等人^[5]提出一种基于地理位置的旅馆推荐系统,通过构建决策树对旅馆资源进行评分,并利用地理情境因素为旅客提供推荐服务。Dao 等人^[6]提出一种新型协同过滤推荐模型,将遗传算法和情境感知相结合,利用用户的偏好和交互的上下文信息实现位置的广告推荐服务。Kim 等人^[7]提出一种情境感知模型,将缺失的用户偏好值映射到协同过滤技术中,实现了协同过滤技术与情境感知的结合,并将其应用于泛在环境下的个性化健康服务系统中。Karatzoglou 等人^[8]提出一种基于张量分解的多元推荐算法(multiverse recommendation, MV),将情境感知与协同过滤推荐结合起来,该算法具有很好的推荐性能,但其时间复杂度太高,在实际使用中不太可行。杨君等人^[9]提出一种基于情境的多维协同过滤新方法,利用情境间的相似度计算当前用户的近邻在当前用户所在情景下对项目的评分,继而进行推荐。曾子明和李鑫^[10]运

① 国家自然科学基金(61603091)和江苏省教育信息化研究立项课题(20180054)资助项目。
② 女,1981 年,讲师,硕士;研究方向:信息提取与分析研究;联系人,E-mail: xiujuan_xie@126.com
(收稿日期:2018-10-19)

用贝叶斯方法计算用户在不同情境下的项目推荐权重,再将其与传统协同过滤中的预测评分相结合,实现基于情境的个性化推荐。田雪筠^[11]提出基于情境和基于内容相结合的推荐算法,用户对资源的兴趣度将与当前特定情境值相对应,利用与用户当前情境相似的历史情境下的资源,计算待推荐资源的相似度,从而产生推荐。洪亮等人^[12]提出一种针对移动数字图书馆用户的情境感知的个性化推荐方法,基于相似用户有相似选择的考虑,将情境因素作为参数纳入用户信任值和相似度的计算中,利用用户间隐形信任关系充实信任模型。

综上所述,情境是移动环境下个性化推荐不可或缺的因素。不同的移动应用所要考虑的情景因素不尽相同,如何保证基于情境的个性化推荐质量至关重要。将情境感知应用于移动阅读领域,目前国内的相关研究并不多。本文在现有研究的基础上,提出一种针对移动阅读领域中的情境感知的个性化推荐算法,该算法汇聚了两阶段的情境感知思想,首先使用贝叶斯方法推算用户在某情境下偏好度最高的资源类别,而后再对该类别下的资源进行基于情境的用户协同过滤推荐。

1 相关方法与定义

1.1 基于用户的协同过滤推荐方法

本文所使用的是协同推荐中基于用户的协同推荐算法,根据相似用户对资源的打分来预测目标用户对特定资源的评分。用户相似度记为 $\text{sim}(u, v)$, 计算公式见式(1);用户 u 对资源 i 的预测评分记为 $P_{u,i}$, 计算公式如式(2)。

$$\text{sim}(u, v) = \cos(\vec{u}, \vec{v}) = \frac{\sum_{i \in I_{u,v}} r_{u,i} r_{v,i}}{\sqrt{\sum_{i \in I_u} r_{u,i}^2 \sum_{i \in I_v} r_{v,i}^2}} \quad (1)$$

$$P_{u,i} = \bar{r}_u + \frac{\sum_{v \in \text{Neighbor}} \text{sim}(u, v) \times (r_{v,i} - \bar{r}_v)}{\sum_{v \in \text{Neighbor}} \text{sim}(u, v)} \quad (2)$$

其中, u 和 v 代表 2 个用户, $\text{sim}(u, v)$ 表示用户间的相似性, $I_{u,v}$ 表示用户 u 和 v 共同评分过的资

源集, $r_{u,i}$ 和 $r_{v,i}$ 分别表示用户 u 和用户 v 对资源 i 的评分, $\bar{r}_u$ 和 $\bar{r}_v$ 分别表示用户 u 和 v 对所有已评分资源的平均评分, Neighbor 为近邻集。

1.2 朴素贝叶斯分类方法

假设一特征集合 $A(a_1, a_2, \dots, a_n)$, 其中 a_i 表示属性特征值, 此外, 还有 m 个类, 分别用 $S_1, S_2, \dots, S_m$ 表示, 根据贝叶斯定理, 得到在特征 A 下, 某分类 S_k 的出现概率 $p(S_k | A)$ 为

$$p(S_k | A) = \frac{p(A | S_k)p(S_k)}{P(A)} = \frac{p(a_1, a_2, \dots, a_n | S_k)p(S_k)}{p(a_1, a_2, \dots, a_n)} \quad (3)$$

$P(A)$ 表示全概率, 全概率公式为

$$P(A) = \sum_{k=1}^m p(A | S_k)p(S_k) \quad (4)$$

朴素贝叶斯方法是一种应用基于独立假设的贝叶斯定理的简单分类器, 各个属性特征值间相互独立, 即联合的 $a_1, a_2, \dots, a_n$ 概率等于每个单独的属性概率的乘积, 具体公式如下:

$$p(a_1, a_2, \dots, a_n | S_k) = \prod_{i=1}^n p(a_i | S_k) \quad (5)$$

1.3 相关定义

定义 1 情境是指影响人的行为或事件的上下文信息或者场景信息^[13]。

定义 2 移动环境下的情境推荐。移动环境下, 用户的阅读内容容易受到时间、地点、网络状况等情境因素的影响, 因此, 在实现推荐时, 需要充分考虑这些因素, 数据空间应从“用户-项目”二维形式拓展到“用户-项目-情境”三维形式。

定义 3 情境因素。影响用户需求的情境因素可以概括为 4 点: 第一, 用户偏好, 不同用户兴趣偏好亦不同, 应根据用户的偏好进行个性化推荐; 第二, 阅读目的, 用户阅读是一种有目的的活动, 及时感知用户的阅读目的可以提高推荐的质量; 第三, 网络状况, 网络条件的好坏决定用户选择所要阅读资源的类型, 如文本、图片、视频等; 第四, 移动终端特点, 应充分考虑移动终端的电耗、屏幕尺寸等, 做到人性化推荐。

2 融合情境的移动阅读个性化推荐算法

融合情境的移动阅读个性化推荐算法, 充分考

虑到情境因素对于用户移动阅读偏好的影响,提出两阶段的情境感知推荐思想。第一阶段是使用朴素贝叶斯方法,计算得出目标用户在当前情境下对每个阅读类别的偏好值,选取偏好值最大的类别作为推荐对象;第二阶段,是对第一阶段中偏好值最大的阅读类别下的资源,通过情境相似度的计算,筛选出当前情境或者与当前情境最为相似的二维评分矩阵模型,并运用传统的基于用户的协同过滤算法产生最终推荐集。本节将分别对本文算法所涉及的数据参数、基于贝叶斯推理的情景化用户阅读类别偏好、情境相似度计算、本文所提出算法的过程进行叙述。

2.1 本文算法所涉及的数据参数

本文算法的用户数据模型包括用户、项目资源和情境 3 个维度,记作 $\{U-I-C\}$, 每个维度由各自属性值组成的向量表示,下面列出本文的主要数据参数:

资源类别 $Sort = \{S_{kj}, S_{yl}, S_{qj}, S_{sh}, S_{ty}, S_{jy}\}$, 包括科技、娱乐、财经、社会、体育和教育 6 种类别;

情境 $Context = \{Time, Week, Location\}$, 包括阅读时间、是否周末和地点;

阅读时间 $Time = \{T1, T2, T3, T4\}$, 时间的属性取值分别为早晨、中午、下午和晚上;

周末 $Week = \{W1, W2\}$, 属性的 2 个取值分别表示周末和工作日;

地点 $Location = \{L1, L2, L3, L4\}$, 地点的属性取值分别表示工作场所、家中、途中和其他。

2.2 基于贝叶斯推理的情景化用户阅读类别偏好

用户的阅读偏好与其当前的情境有一定联系,例如,一些用户习惯早晨上班的途中看社会新闻,了解时事,而在中午休息时间则喜欢看娱乐方面的资讯。此外,不同用户对同样情境的敏感度又不相同,如,同样是“早晨上班途中”这个情境,用户甲喜欢看社会新闻类,而用户乙更喜欢阅读体育类。因此,分析单个用户的情境化阅读偏好,有利于更精确地向目标用户推荐资源。

鉴于朴素贝叶斯方法在处理大量数据时具有较高的分类准确性和运算性能,本文利用朴素贝叶斯分类方法计算用户的阅读偏好值。考虑到阅读资源数量很大,用户评价相对较少,而资源间内容或者题

材具有一定的相似性,故先将每个阅读资源划分到 $Sort$ 对应的 6 个类别中,再计算不同情境下用户对每个资源类别的偏好情况。

欲求用户 u 在给定情境 ($Time, Week, Location$) 下对资源类别 S_k 的阅读偏好值 $p(S_k | C_t, C_w, C_l)$, 这里的 C_t, C_w, C_l 分别表示时间、周末和地点三个情境的具体属性值。首先,根据三维模型 $\{U-I-C\}$ 中的评分数据,统计用户 u 在每个情境属性值下关于资源类别 S_k 的条件概率,以及资源类别 S_k 在所有类别中的出现概率,分别记作 $p(C_i | S_k)$ 和 $p(S_k)$, 然后运用朴素贝叶斯相关定理,推导得到 $p(S_k | C_t, C_w, C_l)$, 推导过程如下。

由贝叶斯定理,即公式(3),将 $p(S_k | C_t, C_w, C_l)$ 展开,得到式(6)。

$$p(S_k | C_t, C_w, C_l) = \frac{p(C_t, C_w, C_l | S_k)p(S_k)}{p(C_t, C_w, C_l)} \tag{6}$$

再运用全概率公式,以及朴素贝叶斯方法中“联合的概率等于每个单独属性概率的乘积”的特性,对式(6)进一步展开,得到:

$$p(S_k | C_t, C_w, C_l) = \frac{\prod_{i=1}^3 p(C_i | S_k)p(S_k)}{\sum_{k=1}^6 p(C_t, C_w, C_l | S_k)p(S_k)} \tag{7}$$

对式(7)中的 $p(C_t, C_w, C_l | S_k)$, 再次运用全概率公式,得到式(8),至此,推导过程结束。

$$p(S_k | C_t, C_w, C_l) = \frac{\prod_{i=1}^3 p(C_i | S_k)p(S_k)}{\sum_{k=1}^6 (\prod_{i=1}^3 p(C_i | S_k))p(S_k)} \tag{8}$$

2.3 情境相似度

根据用户 u 在不同情境 C_x 和 C_y 下对相同资源的评分,利用余弦相似度算法,计算基于用户 u 的不同情境间的相似度,计算公式如下:

$$Sim(u, C_x, C_y) = \frac{\sum_{i \in I} r_{u,i,C_x} r_{u,i,C_y}}{\sqrt{\sum_{i \in I} r_{u,i,C_x}^2 \sum_{i \in I} r_{u,i,C_y}^2}} \tag{9}$$

其中, I 表示用户 u 在情境 C_x 和 C_y 下共同评分的资源集合, r_{u,i,C_x} 和 r_{u,i,C_y} 分别表示用户 u 在情境 C_x 和情境 C_y 下对资源 i 的评分值。

2.4 算法过程描述

输入:“用户 U -资源 I -情境 C ”三维数据模型;
目标用户 u ;当前情境 C_i ;

输出:用户 u 在当前情境 C_i 下偏好最大的 N 个资源推荐集。

步骤 1 收集属于当前用户 u 的情境化资源评分数据,以情境为单位,将资源划分到 $Sort$ 对应的 6 个类别中,并统计情境信息及资源类别所涉及的条件概率;

步骤 2 利用式(8),计算当前用户 u 在当前情境 C_i 下对每个资源类别的阅读偏好值,选取偏好值最大的类别作为推荐对象,记为 RC ;

步骤 3 从“ U - I - C ”三维模型中提取所有包含 RC 中资源的用户评分记录,如果某用户 v 对某资源 i 在当前情境下有评分,则直接予以保留;如果用户 v 在除当前情境以外的其他情境下对资源 i 有评分,则先根据式(9)计算基于用户 v 的每个情境与当前情境的相似度,保留相似度最大的情境所对应的评分记录,至此,可以过滤出与当前情境相关的“ U - I ”二维评分矩阵模型;

步骤 4 将步骤 3 中的“ U - I ”模型,首先通过余弦相似度公式计算用户间的相似度,找到目标用户 u 的近邻集,然后根据近邻对资源的评价,预测目标用户 u 对未知资源的评分,筛选出评分最高的前 N 项,即产生基于情境的 Top- N 推荐列表。

3 实验与结果分析

3.1 实验数据集

本文的测试资源为网络博文,测试用户是在校的 250 位大学生。用网络爬虫工具按照 $Sort$ 的 6 个类别分别抓取来自新浪上的博文,每个类别 20 篇文章,共计 120 篇文章。采用人工问卷的方式收集数据,每张问卷上给出所有文章的链接地址,要求用户在规定的情境下选择感兴趣的文章进行 5 级制评分,每个用户至少测试 3 个情境,每个情境不少于 20 篇文章。评分范围 1~5,1 表示非常不感兴趣,2 表示不感兴趣,3 表示中等,4 表示感兴趣,5 表示非常感兴趣。阅读情境包括阅读时间、周末、地点 3 个属性,时间属性有 4 个取值,周末属性有 2 个取值,地点有 4 个取值,共有 $4 \times 2 \times 4 = 32$ 个情境。

在进行贝叶斯推理时,需要将评分数据转换为概率数据,大致过程是:先统计目标用户在每个情境属性取值下对每个类别的文章评分数和(表 1 情境行中竖线的左侧数据),以及每个类别的评分数和,再由这两个量相除得到情境信息的阅读类别条件概率 $p(C_i|S_k)$ (表 1 情境行中竖线的右侧数据),而后用每个类别的评分和除以所有类别的评分总和得到每个资源类别的出现概率 $p(S_k)$ (表 1 最后一行竖线右侧数据),如表 1 所示。

表 1 情境信息的条件概率

类别		科技	娱乐	财经	社会	体育	教育
情境							
时间	上午	28 0.431	12 0.293	15 0.375	23 0.390	6 0.231	3 0.25
	中午	19 0.292	8 0.195	10 0.25	11 0.186	3 0.115	4 0.333
	下午	7 0.108	5 0.122	3 0.075	6 0.102	2 0.077	0 0
	晚上	11 0.169	16 0.390	12 0.3	19 0.322	15 0.577	5 0.417
周末	是	45 0.692	28 0.683	29 0.725	32 0.542	15 0.577	7 0.583
	否	20 0.308	13 0.317	11 0.275	27 0.458	11 0.423	5 0.417
地点	工作场所	7 0.108	2 0.049	18 0.45	6 0.102	6 0.231	4 0.333
	家中	28 0.431	14 0.341	11 0.275	15 0.254	7 0.269	0 0
	途中	22 0.338	19 0.463	7 0.175	27 0.458	10 0.385	5 0.417
	其他	8 0.123	6 0.146	4 0.1	11 0.186	3 0.115	3 0.25
总计	243	65 0.267	41 0.169	40 0.165	59 0.243	26 0.107	12 0.049

为了分析不同数据集对本文算法的影响,分别随机抽取了 150、200、250 位用户的评分数据组成 3 个数据集,记为 D_1 、 D_2 、 D_3 ,在每个数据集中,实验训练集占 80%,测试集占 20%。

3.2 度量标准

本次实验采用平均绝对偏差^[14] (mean absolute error, MAE) 作为度量标准,验证算法的有效性。MAE 通过计算预测评分与实际评分之间的偏差来度量预测的准确性,MAE 值越小,推荐质量越高。MAE 的计算公式如下:

$$MAE = \frac{\sum_{i \in N} |p_i - r_i|}{N} \tag{10}$$

其中, p_i 表示目标用户对资源 i 的预测评分, r_i 表示目标用户对资源 i 的实际评分, N 表示测试集的大小。

3.3 对比方法介绍

在实验结果分析中,将本文所提出的情境感知的移动阅读个性化推荐算法 (context-aware mobile reading personalized recommendation algorithm, CPRG) 与下列方法进行比较。

- (1) 基于用户的协同过滤算法 (user _ based collaborative filtering algorithm, UCF): 该方法仅依靠“用户-资源”的二维评分数据,计算用户间的相似度,而后根据相近用户的行为对目标用户作出推荐。
- (2) 一阶段情境感知的移动阅读个性化推荐算

法 (CPRG₁): 该方法是 CPRG 方法中一阶段情境感知的独立版本,运用朴素贝叶斯方法得到用户对资源类别的情境化偏好值,再将其与 UCF 的预测评分进行乘法结合,得到基于一阶段情境感知的推荐结果。

- (3) 基于张量分解的多元推荐算法 (MV)^[8], 该方法将情境感知与协同过滤推荐结合起来,将数据模拟为“用户-资源-情境”的 N 维张量。

3.4 实验结果分析

(1) 近邻数对 MAE 的影响

第一组实验采用用户数及评分数最多的 D_3 数据集,近邻数从 5 递增到 35,步长为 5,将 4 种方法进行对比实验,结果如图 1 所示。由于每个用户要完成 3 个情境的测试,因此 MAE 是取 3 种情境的均值,从图 1 中不难发现,对于给定的不同邻居数, CPRG 总体上比其他方法有更小的 MAE。此外,每个方法的 MAE 值跟近邻数有着密切的关系,初期,随着近邻数的增加,MAE 值呈递减的趋势,说明推荐过程中,参考信息越多,产生的推荐结果越理想。但当邻居数递增到一定值后,MAE 值开始逐渐变大,其原因可能是随着近邻集合的扩大,集合中包含相似度不大的近邻用户,导致推荐质量下降。图 1 中, CPRG、MV 和 CPRG₁ 3 种方法均在近邻数为 25 时,取得最小的 MAE 值,分别为 0.547、0.571 和 0.686,而 UCF 是在近邻数为 20 时,MAE 突然骤降,

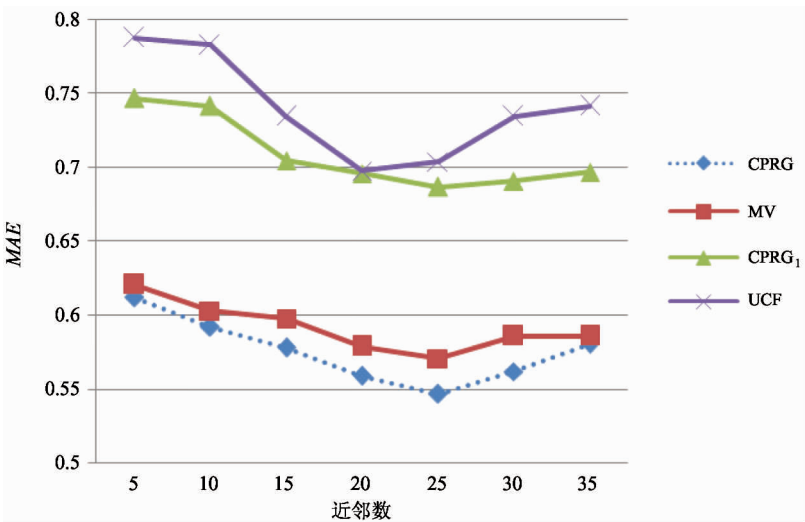


图 1 不同算法中近邻数对 MAE 的影响对比

取到最小值 0.698,该值与 CPRG_1 方法相同近邻数下的 MAE 值接近,随后开始递增,这说明情境因素在近邻数为 20 时,对于推荐质量影响程度相对较小。综上,每个算法存在一个最佳的近邻值,此时推荐质量达到最佳,为客观起见,在后面的样本量的实验中,4 种算法分别取最佳的近邻值。

(2) 样本量对 MAE 的影响

第二组实验在数据集 D_1 、 D_2 、 D_3 上进行,分别计算 4 种方法在不同样本量下的 MAE , CPRG 、 MV 和 CPRG_1 3 种方法的近邻数取值为 25,UCF 近邻数取值为 20,结果如图 2 所示。4 种算法均表现出训练集样本量越多,推荐结果越好。因为训练集中用户数越多,获得的目标用户的近邻就越准确,评分的预测就越可靠,从而降低 MAE 。此外,对于同一个样本,在近邻数一定时,本文提出的方法的 MAE 小于其他方法的 MAE ,当使用样本量为 150 的 D_1 数据集时, CPRG 方法比 MV 方法在 MAE 值上低 3.97%,比 CPRG_1 方法低 12.03%,比 UCF 方法低 14.7%;当使用样本量为 200 的 D_2 数据集时, CPRG 方法比 MV 方法在 MAE 值上低 6.33%,比 CPRG_1 方法低 15.89%,比 UCF 方法低 18.73%;当使用样本量为 250 的 D_3 数据集时, CPRG 方法比 MV 方法在 MAE 值上低 4.2%,比 CPRG_1 方法低 20.38%,比 UCF 方法低 21.63%;这些数据体现出 CPRG 方法在推荐质量上优于其他 3 种方法,进一步验证了 CPRG 方法的有效性。

CPRG 方法之所以相较其他 3 种方法具有更精准的推荐值,分析其原因,主要有:第一,UCF 方法只考虑用户间的相似度,忽视了情境因素对推荐结果的影响,因此,推荐效果相对比较差;第二, CPRG_1 方法只进行了第一阶段的情境感知,获得的是用户对资源类别的情境偏好值,资源类别与具体资源的情境偏好值具有一定偏差,所以某种程度上对于推荐结果有一定影响,比如某用户爱好阅读篮球、足球之类的体育赛事文章,并不代表他对其他体育类文章也一样感兴趣,因此仅用资源类别的偏好作推荐会存在以点概面的问题;第三, MV 方法在 3 个数据集下的 MAE 值均与 CPRG 方法最接近,但是由于受评分数据稀疏的影响,缺少 CPRG 方法中使用最相似情境评分代替缺失情境评分的处理,致使推荐质量受到影响;第四, CPRG 方法汇聚两阶段的情境感知思想,不仅考虑资源类别的情境偏好,还注意情境因素对具体资源的影响,因此得出的推荐列表更为准确。

(3) 时间复杂度分析

CPRG 方法中,第一阶段使用朴素贝叶斯方法计算得出目标用户在当前情境下对每个阅读类别的偏好值,时间复杂度为 $O(n^2)$,第二阶段通过情境相似度计算筛选出当前情境或者与当前情境最为相似的二维评分矩阵模型,时间复杂度为 $O(n^2)$,以及根据近邻集预测评分的时间复杂度也为 $O(n^2)$,得到

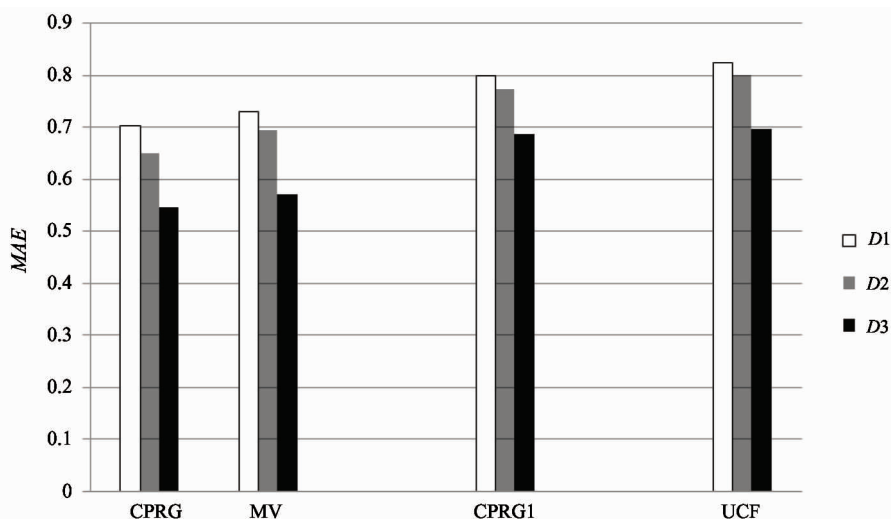


图2 不同算法中样本量对 MAE 的影响对比

CPRG方法总的时间复杂度为 $O(n^2) + O(n^2) + O(n^2)$ 。传统UCF方法的时间复杂度为 $O(n^2)$,而MV方法的时间复杂度随因子维度和分解维度成指数级增长 $O(k^m)$,在基于情境的协同推荐过程中,时间复杂度至少是立方级的。因此本文的CPRG方法时间复杂度与传统UCF方法属于同一个数量级,优于MV方法,可见CPRG方法并没有因为增加情境维度而过多地影响算法的时间复杂度。

4 结 论

本文提出一种结合情境和用户协同过滤的移动阅读推荐算法。首先利用朴素贝叶斯方法,推理得出目标用户在当前情境下的阅读类别偏好;然后对偏好值最大的类别下的资源,通过情境相似度的计算,筛选出当前情境或者与当前情境最为相似的评分记录,从而将“*U-I-C*”三维模型降至“*U-I*”二维模型;最后运用传统的基于用户的协同过滤算法产生推荐列表。为了对本文算法进行评价,与其他相关算法进行对比实验,结果表明,本文算法总体上推荐质量更高,具有一定的可行性和有效性。

参考文献

- [1] 王立才,孟祥武,张玉洁. 移动网络服务中基于认知心理学的用户偏好提取方法[J]. 电子学报,2011,39(11): 2547-2553
- [2] Greg L, Brent S, Jeremy Y. Amazon. Com Recommendations: item-to-item collaborative filtering[J]. *IEEE Internet Computing*, 2003,7(1): 76-80
- [3] Deng A L, Zhu Y Y, Shi B L. A collaborative filtering recommendation algorithm based on item rating prediction [J]. *Journal of Software*, 2003,14(9): 1621-1624
- [4] Karypis G. Evaluation of item-based top-*N* recommendation algorithms[C]. In: Proceedings of CIKM 2001, Atlanta, USA, 2001. 247-254
- [5] Lee B H, Kim H N, Jung J G, et al. Location- based service with context data for a restaurant recommendation [C]. In: Proceedings of the 17th International Conference DEXA, Krak6w, Poland, 2006. 430-438
- [6] Dao T H, Jeong S R, Ahn H. A novel recommendation model of location-based advertising: context-aware collaborative filtering using GA approach[J]. *Expert Systems with Applications*, 2012,39(3): 3731-3739
- [7] Kim J, Lee D, Chung K Y. Item recommendation based on context-aware model for personalized u-healthcare service[J]. *Multimedia Tools and Applications*, 2014,71(2): 855-872
- [8] Karatzoglou A, Amatriain X, Baltrunas L, et al. Multi-verse recommendation: n-dimensional tensor factorization for context-aware collaborative filtering [C]. In: Proceedings of ACM Conference on Recommender Systems, Barcelona, Spain. 2010. 79-86
- [9] 杨君,汪会玲,艾丹祥. 一种基于情景的多维协同过滤新方法研究[J]. 图书情报工作, 2011, 55(21): 110-114
- [10] 曾子明,李鑫. 移动环境下基于情境感知的个性化信息推荐[J]. 情报杂志, 2012, 31(8): 166-170
- [11] 田雪筠. 基于情境感知的移动电子资源推荐技术研究[J]. 情报理论与实践, 2015, 38(5): 86-89
- [12] 洪亮,钱晨,樊星. 移动数字图书馆资源的情境感知个性化推荐方法研究[J]. 现代图书情报技术, 2016, 32(7): 110-119
- [13] Koren Y. Factorization meets the neighborhood: a multifaceted collaborative filtering model[C]. In: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Las Vegas, USA, 2008. 426-434
- [14] Fazeli S, Loni B, et al. Implicit vs explicit trust in social matrix factorization[C]. In: Proceedings of the 8th ACM Conference on Recommender System, Silicon Valley, USA, 2014. 317-320

Research on context-aware mobile reading personalized recommendation algorithm

Xie Xiujuan^{*}, Mo Lingfei^{**}, Li Xiangju^{*}, Cao Fengping^{*}

(^{*} Department of Computer Engineering, Chengxian College, Southeast University, Nanjing 210088)

(^{**} College of Instrument Science and Engineering, Southeast University, Nanjing 210018)

Abstract

Aiming at the problem of personalized recommendation of resources in mobile environment, a mobile reading recommendation algorithm combining context and collaborative filtering is proposed. Two stages of context awareness are intergrated in the algorithm. Firstly, the naive Bayesian method is used to estimate the highest preference categories in a given context. Then, through the calculation of context similarity, the ‘user resource’ two-dimensional scoring model is filtered out which only belongs to the current context or the most similar context to the current context. Finally, the traditional user-based collaborative filtering algorithm is used to generate the recommendation list. Experiments on Sina Bowen are carried out, and the results show that the mean absolute error (MAE) of the proposed algorithm is lower than the related algorithm under the same conditions, and has higher recommendation quality.

Key words: context-aware, collaborative filtering, mobile recommendation, Bayesian method