

基于个性化情景的移动商务信任推荐模型研究

杜巍¹,高长元^{1,2}

(1.哈尔滨理工大学 管理学院,黑龙江 哈尔滨 150040;

2.哈尔滨理工大学 高新技术产业发展研究中心,黑龙江 哈尔滨 150040)

摘要:【目的/意义】移动互联网时代,移动电子商务用户的个性化信息需求具有极强的情景依赖性与感知信任性,针对目前移动商务信息服务个性化和准确性较低,提出融入用户个性化情景与用户间信任关系的推荐模型。【方法/过程】首先,通过用户当前情景和历史评分数据计算出对每个用户即时信息需求影响最大的K个情景要素,以此构造用户个性化情景,然后结合不同信任环境下的用户信任度矩阵改进已有的不同信任信息环境下用户情景兴趣推荐方法,进而进行项目推荐。【结果/结论】通过Movie lens与Book-Crossing数据集对本文提出的算法和其它两种算法进行比较,实验结果表明:本模型具有较高的推荐准确率,可有效地解决移动商务环境下的个性化推荐问题。

关键词: 个性化情景;移动商务;信任关系;推荐模型

中图分类号:G252;TP391.3

文献标识码:A

文章编号:1007-7634(2017)10-23-07

DOI:10.13833/j.cnki.is.2017.10.005

Research on Mobile Commerce Trust Recommendation Model Based on Personalized Context

DU Wei¹,GAO Chang-yuan^{1,2}

(1.School of Management, Harbin University of Science and Technology, Harbin 150040,China ; 2.High-tech Industrial Development Research Center, Harbin University of Science and Technology, Harbin 150040, China)

Abstract:【Purpose/significance】In the era of mobile Internet,the personalized information needs of mobile e-commerce users have strong context-dependent and perceived trustworthiness.In view of the low personalization and accuracy of current mobile commerce information service,the paper puts forward the integration of users' personalization context and user trust collaborative recommendation model.【Method/process】Firstly, the current user context and historical data was used to calculate K context factors which has the greatest influence to each user's immediate information need, user personalization context was constructed ,and then combined with the users' trust matrix under different trust environment to improve the existing user context interest recommendation methods in different trust information environment ,then the project is recommended.【Result/conclusion】Using Movie lens and Book-Crossing data set to compare the proposed algorithm and the other two algorithms,the experimental results show that this model has a high accuracy and can effectively solve the problem of personalized recommendation in mobile commerce environment.

Keywords: personalized context;e-commerce;trust relationship;recommendation model

近年来,随着移动通信技术的快速发展及信息内容的日益增长,使得人们通过移动网络获取有价值的信息变得愈发困难,导致用户体验受到严重影响。为了有效解决“移动信息”过载问题,移动推荐系统进入了人们的视野,成为缓解此问题的有效手段,得到广泛关注。

当前,针对移动商务用户信息需求具有极强的情景依赖性与感知信任性的特点,国内外学者已开始致力于研究结合移动环境下用户情景信息与感知信任的推荐问题。Golbeck通过社交网站中的信任关系获取用户邻居,并将其应用到电影推荐中^[1]。DePessemier认为目前的信息推荐研究主要集

收稿日期:2017-03-06

基金项目:国家自然科学基金(71272191);国家自然科学基金(71672050);黑龙江省自然科学基金项目(G201301)

作者简介:杜巍(1982-),女,吉林人,讲师,博士,主要从事移动电子商务、推荐系统研究。

中在用户历史行为数据的分析,忽略了用户的情景信息(状态、环境、位置等)或社会关系,而合理地结合用户情景信息与社会网络关系,能够有效地提高推荐质量^[2]。郑孝遥等提出了一种基于信任的协同过滤推荐模型,该模型算法通过融入通信、相似、传递三种信任的综合值及通信情景信息,提高了推荐算法的预测精度^[3]。邓晓懿等人利用情景信息对用户聚类,然后运用社会网络理论分析用户间的信任关系进而改进相似度计算方法,进一步提高移动商务环境下的信息推荐质量^[4]。

以上文献可以看出,国内外学者对同时考虑用户情景兴趣与用户间信任关系进行移动商务推荐研究已取得一定成果,但研究开展大多以用户的整体情景信息为研究对象,并未考虑不同移动商务情景下对不同用户信息需求影响的差异性,忽略了对每位用户个性化情景的深入分析。为此,本文提出基于个性化情景的移动商务信任推荐模型,该模型首先通过训练集的方法为每个用户找出对其信息需求影响最大的 K 个情景要素,在此基础上融合社会网络与信任机制,分别构建基于个性化情景的移动商务富信任信息推荐模型及基于个性化情景的移动商务稀疏信任信息推荐模型,最后改进已有的不同信任环境下用户情景兴趣推荐方法,把个性化情景相似度矩阵与不同信任环境下的用户信任度矩阵分别进行合并取值,并把它作为最终的推荐权值进行推荐,此模型为提高移动商务个性化推荐服务质量,有效改善用户体验提供了新的思路。

1 模型构建

目前,融合用户情景兴趣与用户间信任关系进行推荐的研究虽然已取得一定成果,但在评分数据稀疏的情况下,同时考虑用户个性化情景与用户间信任关系对移动商务用户进行个性化信息服务的研究较少。

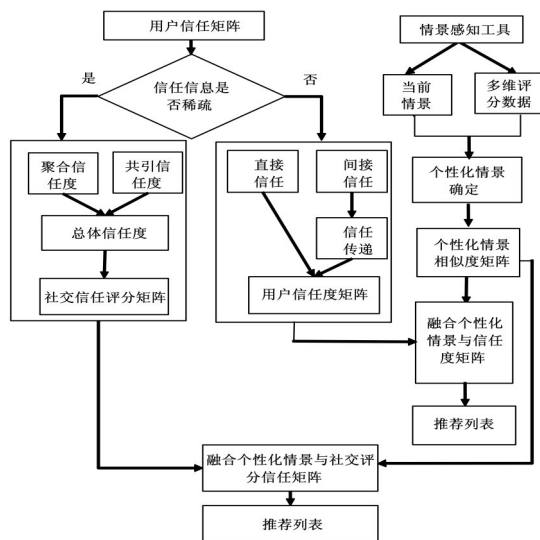


图1 基于个性化情景的移动商务信任推荐模型

鉴于此,本文提出一种基于个性化情景的移动商务信任

推荐模型,该模型首先利用用户当前情景及历史评分数据确定用户个性化情景;然后,引入社会网络分析用户间的信任关系,分别从不同信任信息环境入手,分析富信任环境下的移动商务推荐过程及稀疏信任环境下的移动商务推荐过程,模型具体运行过程如图1所示。

1.1 基于个性化情景的移动商务富信任信息推荐模型

目前,学者们对基于情景的推荐研究主要集中于对全部用户的整体情景为研究对象,对用户个性化情景进行分析略显不足,导致推荐结果在一定程度上偏离用户即时信息需求。在移动商务信息推荐过程中,情景信息因人而异,例如,旅游推荐系统中时间、地点、季节和温度是很重要的情景要素,但有的用户认为以上情景要素都重要,而有的用户认为这些情景要素对他们都没有影响,还有一部分用户认为只有时间或者地点要素对他们来说是重要的。如果能从全部的情景信息中,为每位用户找出影响其做出最终选择的情景要素加以分析,不仅可以减轻系统运行负担,而且能使预测结果更加理想^[5]。其次,已有研究表明,将用户间信任关系引入到推荐系统中,可有效提高推荐性能。据此,本文在移动商务用户间信任信息较为丰富的推荐过程中结合用户个性化情景,提出基于个性化情景的移动商务富信任信息推荐模型。该模型首先通过用户当前情景和历史评分数据计算出对每个用户即时信息需求影响最大的 K 个情景要素,以此构造用户个性化情景,进而确定用户个性化情景相似度矩阵,然后将信任的传递特征引入到信任度计算过程中,建立用户信任度矩阵,并将用户个性化情景相似度矩阵与信任度矩阵进行融合,进而产生推荐结果,此模型的建立,可有效提高个性化推荐服务的效率和准确率,具体推荐流程如图1所示。

1.1.1 个性化情景分析及确定

高旻等创造性地提出“个性化情景”的概念,周朴雄等借鉴这种思想,对个性化情景的内涵又进行了扩展,本文采用周朴雄等所给出的定义,即:个性化情景是指影响待推荐用户即时信息需求的最大的 K 个情景要素^[6]。

本文借鉴文献[6]的方法,以用户评分为衡量标准,通过在用户评分矩阵的训练集上进行训练,得到加入情景要素的推荐质量评价矩阵,每个用户 K 个最优评分下的 K 个情景要素组成该用户的个性化情景。其中, K 值为系统根据最优评分结果自动设定。个性化情景确定的具体过程如下:

Step1:将用户评分矩阵按80:20的比例划分为训练集 D 和测试集。然后将 D 按照用户的当前情景 $C = \{c_1, c_2, \dots, c_i\}$ 划分为 t 个子集,即 $D = \{d_1, d_2, \dots, d_s\}$ 。

Step2:设 $P_{X(u,D)}$ 为信息推荐质量评价标准,其中 U 为用户集, D 为训练集, X 为此评价标准所采用的推荐算法。 $P_{X(u,d_i)}$ 表示以用户 u_i 为目标用户,算法 X 在基于情景 c_i 的训练集 d_i 上训练,用户 u_i 对信息推荐质量的评分。

Step3:得到基于情景的用户推荐评价矩阵 P_c :

$$p_c = \begin{bmatrix} p_{s(u_1, d_1)} & p_{s(u_1, d_2)} & \cdots & p_{s(u_1, d_j)} \\ p_{s(u_2, d_1)} & p_{s(u_2, d_2)} & \cdots & p_{s(u_2, d_j)} \\ \vdots & \vdots & \ddots & \vdots \\ p_{s(u, d_1)} & p_{s(u, d_2)} & \cdots & p_{s(u, d_j)} \end{bmatrix} \quad (1)$$

Step4: 通过Step3为每个用户找出相应行中 K 个最优值所对应的情景要素, 组成的情景集 $C' = (c'_1, c'_2, \dots, c'_k)$, 即为用户的个性化情景。

1.1.2 个性化情景相似度计算

本文用移动用户当前个性化情景 C' 代替当前情景 C 作为过滤条件, 排除了对用户评分影响较小的情景要素所含的评分数据段, 选取所含情景与当前个性化情景 C' 足够相似的评分数据段, 并判断此数据段是否存在某一用户在不同情景下对同一项目存在不同评分的情况, 然后利用传统二维推荐算法进行评分预测, 该算法对已有基于情景的多维信息推荐算法的改进^[7], 不仅可以减轻系统运行的负担, 而且有效地提高了推荐的质量。

实际情况中, 在使用个性化情景 C' 对多维数据进行选择时, 过于明确的情景要素可能会造成数据的“稀疏性”, 所以本文采用的选择条件是 $C' \in S_c$, S_c 是个性化情景 C' 所属的上级情景, 例如 $C' = \text{“女朋友”}$, 则 $S_c = \text{“朋友”}$ 。

具体算法过程描述如下:

在上述设计思想的基础上, 对算法的输入、输出、具体过程展开描述。

输入: 多维评分数据集 MR (Multidimensional Ratings), $MR = \{ (U_j, I_k, C_l, R_{U_j, I_k, C_l}) \}$, $j = 1, \dots, m$, $k = 1, \dots, n$ (总共 m 个用户, n 个项目)。 C_l 为多维评分数据中的历史情景, $l = 1, \dots, q$ 。

用户当前个性化情景 C' , 目标用户 u_1 。

输出: 当前个性化情景 C' 下目标用户 u_1 对项目 I 的预测评分 $P_{u_1, I, C'}$ 。

算法的具体过程, 分为以下几步:

步骤 1 对 MR 进行过滤, 筛选出所含情景与用户当前个性化情景 C' 足够相似的评分数据集 MD 。

步骤 2 判断步骤 1 是否由于数据稀疏性难以构造所含情景与当前个性化情景 C' 足够相似的评分数据段 MD , 如果不存在该问题, 直接执行下一步, 否则, 采用 S_c 代替 C' 作为选择条件进行评分数据段的选取, S_c 为个性化情景 C' 所属的上级情景。

步骤 3 判断 MD 中是否存在某一用户在不同情景下对同一项目存在不同评分的情况, 如果不存在该问题, 直接执行步骤 4, 否则, 对该用户在不同情景下对同一项目的不同评分进行聚合计算, 得到该用户对某一项目的最终评分, 即: $R_{U_j, I_k} = \text{AVG}(U_j, I_k, C_l)$, 这样保证了该用户对同一个项目有唯一一评分值, 从而得到进行了个性化情景过滤后的最终的评分数据段 MD 。

步骤 4 对 MD 采用传统二维推荐算法进行相似度计算。

即:

$$\text{Sim}(u_i, U_j) = \frac{\sum_{s \in I(u_i, U_j)} (R_{u_i, s} - \bar{R}_{u_i}) (R_{U_j, s} - \bar{R}_{U_j})}{\sqrt{\sum_{s \in I(u_i, U_j)} (R_{u_i, s} - \bar{R}_{u_i})^2} \sqrt{\sum_{s \in I(u_i, U_j)} (R_{U_j, s} - \bar{R}_{U_j})^2}} \quad (2)$$

其中, MD 中目标用户 u_i 对所有项目评分的平均值用 \bar{R}_{u_i} 表示, 用户 U_j 对所有项目评分的平均值记为 \bar{R}_{U_j} 。

1.1.3 基于社会网络的用户信任矩阵

引入用户间信任关系的推荐模型需构建用户信任矩阵, 用于描述社会网络中用户间的信任关系。其中, 节点表示用户, 每一条有向的边表示用户间的信任关系, 有向边的权重表示信任程度。由此, 移动商务用户信任网络中用户间信任关系为用户-用户信任矩阵来表示, 记为 $TU_{E \times I}$, 如公式(2)所示。

$$TU_{E \times I} = \begin{bmatrix} tu_{u_1 u_1} & tu_{u_1 u_2} & \cdots & tu_{u_1 u_j} \\ tu_{u_2 u_1} & tu_{u_2 u_2} & \cdots & tu_{u_2 u_j} \\ \vdots & \vdots & \ddots & \vdots \\ tu_{u_k u_1} & tu_{u_k u_2} & \cdots & tu_{u_k u_j} \end{bmatrix} \quad (3)$$

其中, 用户 u_i 对用户 u_j 的信任度用 tu_{u_i, u_j} 来表示, 其取值范围为 $[0, 1]$ 区间的实数。

1.1.4 基于个性化情景的富信任信息推荐算法

(1) 算法设计思想。在信任信息较为丰富的移动商务推荐系统中, 部分用户间不存在直接的信任关系, 致使初始化信任矩阵较为稀疏, 由此导致推荐质量下降。鉴于此, 本文在用户间接信任度计算过程中引入信任的传递特征, 并将用户个性化情景相似度矩阵与信任度矩阵进行合并, 计算两者合并后的综合值, 最后将它作为最终的推荐权值进行推荐。

(2) 算法过程描述 算法的具体过程, 分为以下几步。

Step 1: 计算直接信任度 $T(u, v)$ 。在移动商务用户信任关系网络中, 若有一条从 u 直接指向 v 的边, 存在则直接信任值 $T(u, v) = \text{权值 } w(u, v)$; 不存在则直接信任值 $T(u, v) = 0$ 。

Step 2: 计算间接信任度 $T'(u, v)$ 。在移动商务用户信任关系网络中, 若两用户间建立信任关系需要通过其他用户的间接推荐, 则需要引入信任的传递性, 而信任度的传递过程需要遵循两个规则, 即: 单调递减和有界^[8]。因此, 对于用户 u 与 v 属于 $k_{\text{路径}} (k \geq 2)$ 度好友, 根据小世界理论设置 $\max k \leq 6$, 交互路径 $u, u_1, u_2, \dots, u_{k-1}, v$ 上的间接信任度计算方法如公式所示:

$$\begin{aligned} T'(u, v) &= T(u, u_1) \times T(u_1, u_2) \times \dots \times T(u_{k-1}, v) \\ &= T(u, u_1) \times \left(\prod_{i=1}^{k-2} T(u_i, u_{i+1}) \right) \times T(u_{k-1}, v) \end{aligned} \quad (4)$$

而对于 $k_{\text{路径}} (2 \leq k \leq 6)$ 度好友, 若存在多条交互路径, 则对其间接信任度的融合采用算术平均值表示综合信任度。计算方法如式(5)所示:

$$T'(u, v) = \frac{\sum_{i=1}^{m_{\text{路径}}} T'_i(u, v)}{m_{\text{路径}}} \quad (5)$$

Step 3: 计算用户信任度 T_w 。本文把移动商务用户间的

信任度分为直接信任和间接信任,故用户的信任度是它们的综合值。设 T_{uv} 表示用户 u 对用户 v 的信任度, A 表示 $T(u,v)$, B 表示 $T'(u,v)$, 则用户信任度计算公式如下:

$$T_{uv} = \begin{cases} A & A > 0 \wedge B = 0 \\ B & A = 0 \wedge B > 0 \\ \frac{2AB}{A+B} & A > 0 \wedge B > 0 \end{cases} \quad (6)$$

Step 4: 个性化情景与信任度矩阵合并具体方法如下:

$$Weight = \begin{cases} 0 & \text{if } SimC'(u,v) > 0 \text{ and } Trust(u,v) = 0 \\ Trust(u,v) & \text{if } SimC'(u,v) = 0 \text{ and } Trust(u,v) > 0 \\ \frac{2 \times SimC'(u,v) \times Trust(u,v)}{SimC'(u,v) + Trust(u,v)} & \text{if } SimC'(u,v) > 0 \text{ and } Trust(u,v) > 0 \end{cases} \quad (7)$$

其中, 用户 u 与用户 v 间的个性化情景相似度采用 $SimC'(u,v)$ 表示, 用户 u 与用户 v 的信任度记为 $Trust(u,v)$, $Weight$ 为合并后的综合值。

Step 5: 采用加权平均值方法预测待推荐项目的评分。

$$P_{u,i} = \bar{R}_u + \frac{\sum_{i=1}^m (R_v(i) - \bar{R}_v) \times Weight(u,v)}{\sum_{j=1}^m Weight(u,v)} \quad (8)$$

其中, 邻居个数为 m , $R_u(i)$ 表示用户 u 对项目 i 的评分值, \bar{R}_u 表示用户 u 的平均评分值, \bar{R}_v 为用户 v 的平均评分值, 用户 u 对用户 v 的推荐权值定义为 $Weight(u,v)$ 。

Step 6: 根据评分值的大小进行降序排列, 并取出排名最靠前的 TOP-N 作为推荐集推荐给目标用户。

1.2 基于个性化情景的移动商务稀疏信任信息推荐模型

本文上一节构建了基于个性化情景的移动商务富信任信息推荐模型。但现实情况下大多被推荐用户可用信任信息较少, 难以获得理想的推荐结果。鉴于此, 本节在用户存在少量初始信任关系的环境下, 引入文献[9]对“聚合”与“同引”的潜在信任发现方法, 结合用户个性化情景, 提出基于个性化情景的移动商务稀疏信任信息推荐模型, 首先从“聚合”与“同引”信任两方面入手, 挖掘稀疏信任信息环境下用户间

潜在的信任关系, 在此基础上将用户信任度矩阵与个性化情景相似度矩阵合并为一个复合矩阵进行推荐, 具体推荐流程如图1所示。

1.2.1 显性信任描述

在基于个性化情景的移动商务稀疏信任信息推荐模型中, 根据目标用户的评分数据, 可以建立目标用户与其好友间的信任关系及信任程度, 因此可将社会信任网络定义为一个有向加权图, 即: $G(U, F, W)$ 表示。其中, U 为 G 中用户节点集合; 边的集合 $F = \{(u_1, u_2), (u_1, u_3), \dots, (u_{m-1}, u_m)\}$ 主要用于表示用户与邻居间的相互信任关系; $W = \{t_{u_1, u_2}, t_{u_1, u_3}, \dots, t_{u_{m-1}, u_m}\}$ 为权值集, t_{u_1, u_2} 的信任度取值范围为 $[0, 1]$ 。

通过上述分析可以看出, 社会信任网络有以下三个特点:

(1) 有向性: 指具有一定指向行为的信任关系, 如用户 u_1 信任 u_2 , 但用户 u_2 有可能完全不信任 u_1 ;

(2) 非对称性: 是指用户间的信任程度是不具有对称性的, 如用户 u_1 信任 u_2 , 并不等于 u_2 也同样程度的信任 u_1 ;

(3) 有条件的传递性: 是指信任在一定约束条件下具有传递性, 如用户 u_1 信任 u_2 , u_2 信任 u_3 , 根据信任传递性可推知 u_1 信任 u_3 。

根据上述三种特性, 计算目标用户 u_a 和邻居 u_b 间信任度:

$$T(u_a, u_b) = \begin{cases} \frac{\sum_{u_i \in H_{\max}(u_a)} t_{u_i, u_b} \left(\prod_{u_j \in A_{H_{\max}}(u_i)} t_{u_j, u_i} \right) t_{u_i, u_b}}{\sum_{u_i \in H_{\max}(u_a)} t_{u_i, u_b} \left(\prod_{u_j \in A_{H_{\max}}(u_i)} t_{u_j, u_i} \right)} & t_{u_a, u_b} \neq \phi, A_{H_{\max}}(u_a) \neq \phi \\ 0, & \text{otherwise} \end{cases} \quad (9)$$

其中, $T(u_a, u_b)$ 为用户 u_a 和用户 u_b 间的信任度, 取值范围为 $[0, 1]$ 。 $A_{H_{\max}}(u_a)$ 是基于信任的有条件传递规则、在最大信任传递距离 H_{\max} 范围内获得用户 u_a 的信任邻居集合。

基于以上对移动商务用户间社交网络显性信任的描述, 为更好地解决稀疏信任信息环境下的移动商务个性化服务问题, 本文引入文献[9]“聚合”与“同引”的潜在信任发现方法, 挖掘稀疏信任信息环境下潜在的用户信任关系, 并结合本文提出的个性化情景, 在此基础上计算个性化情景相似度

与信任度的复合值进行推荐。

1.2.2 基于个性化情景的稀疏信任信息推荐算法

具体步骤如下:

Step1: 计算聚合信任度。

聚合信任是指在移动商务社会信任网络中, 若用户 u_a 与 u_b 存在共同信任的用户, 则 u_a 与 u_b 存在聚合信任关系, 且 u_a 与 u_b 共同信任邻居的数目影响二者信任度的大小, 目标用户 u_a 对 u_b 的聚合信任经验 $T_{\text{聚合}}(u_a, u_b)$ 计算方法为:

$$T_{\text{聚合}}(u_a, u_b) = \frac{\sum_{u_c \in A_{H_{\max}}^{\text{聚合}}(u_a, u_b)} E(T_{\text{聚合}}(u_a, u_c), T_{\text{聚合}}(u_b, u_c))}{\sum_{u_g \in A_{H_{\max}}^{\text{聚合}}(u_a)} E(T_{\text{聚合}}(u_a, u_g))} \times \log \frac{U_{\text{社网}}}{|A_{H_{\max}}^{\text{聚合}}(u_a, u_b)|} \quad (10)$$

其中,目标用户 u_a 对 u_b 的信任经验采用 $T_{\text{聚合}}(u_a, u_b)$ 表示,整个社会信任网络中的用户数目记为 $U_{\text{社网}}$ 。 $A_{H_{\max}}^{\text{聚合}}(u_a)$ 是根据信任的有条件传递规则、在最大信任传递距离 H_{\max} 范围内获得的用户 u_a 的信任邻居集合; $A_{H_{\max}}^{\text{聚合}}(u_a, u_b)$ 表示用户 u_a 和用户 u_b 共同信任邻居的集合。设目标用户 u_a 对其邻居 u_g 完全信任,则 $T_{\text{聚合}}(u_a, u_g) = 1$, 否则为 0; 用户 u_a 和用户 u_b 的共同信任经验记为 $E(T_{\text{聚合}}(u_a, u_c), T_{\text{聚合}}(u_b, u_c))$; 对数用于衡量邻居的重要性,其中 $|A_{H_{\max}}^{\text{聚合}}(u_a, u_b)|$ 表示 u_a 和用户 u_b 共同信任

邻居的数量^[10]。

根据 u_a 和 u_b 对共同信任邻居的信任度大小,可计算用户 u_a 对 u_b 的信任水平:

$$T_{\text{聚合}}(u_a, u_b) = \frac{\sum_{u_c \in A_{H_{\max}}^{\text{聚合}}(u_a, u_b)} L(T_{\text{聚合}}(u_a, u_c)) \times L(T_{\text{聚合}}(u_b, u_c))}{\sum_{u_g \in A_{H_{\max}}^{\text{聚合}}(u_a)} L(T_{\text{聚合}}(u_a, u_g))} \quad (11)$$

其中, $T(\cdot)$ 的计算方法为:

结合公式(10)和(11), u_a 对 u_b 的聚合信任度 $T_{\text{聚合}}(u_a, u_b)$

$$T(u_a, u_b) = \begin{cases} \frac{\sum_{u_i \in A_{H_{\max}}^{\text{聚合}}(u_a)} t_{u_a, u_i} \left(\prod_{u_j \in A_{H_{\max}}^{\text{聚合}}(u_a)} t_{u_i, u_j} \right) t_{u_i, u_b}}{\sum_{u_i \in A_{H_{\max}}^{\text{聚合}}(u_a)} t_{u_a, u_i} \left(\prod_{u_j \in A_{H_{\max}}^{\text{聚合}}(u_a)} t_{u_i, u_j} \right)} & t_{u_a, u_b} \neq \phi, A_{H_{\max}}(u_a) \neq \phi \\ 0, & \text{otherwise} \end{cases} \quad (12)$$

的计算方法为:

$$T_{\text{聚合}}(u_a, u_b) = \frac{T_{\text{聚合}}(u_a, u_b) \times T_{\text{聚合}}(u_a, u_b)}{\sum_{u_g \in A_{H_{\max}}^{\text{聚合}}(u_a)} T_{\text{聚合}}(u_a, u_g) \times T_{\text{聚合}}(u_a, u_g)} \quad (13)$$

其中,聚合社交网络信任度 $T_{\text{聚合}}(u_a, u_b)$ 的取值范围为 $[0, 1]$ 。

Step2: 计算同引信任度。

同引信任是指在移动商务社会信任网络中,若用户 u_a 与 u_b 同时被其他共同的邻居所信任,则 u_a 与 u_b 存在同引信任关系,且信任度的大小受 u_a 和 u_b 被共同信任邻居数目的影响。目标用户 u_a 对 u_b 的同引信任经验 $T_{\text{同引}}(u_a, u_b)$ 计算方法为:

$$T_{\text{同引}}(u_a, u_b) = \frac{\sum_{u_d \in A_{H_{\max}}^{\text{同引}}(u_a, u_b)} E(T_{\text{同引}}(u_d, u_a), T_{\text{同引}}(u_d, u_b))}{\sum_{u_k \in A_{H_{\max}}^{\text{同引}}(u_a)} E(T_{\text{同引}}(u_k, u_a))} \times \log \frac{U_{\text{社网}}}{|A_{H_{\max}}^{\text{同引}}(u_a, u_b)|} \quad (14)$$

其中,目标用户 u_a 对 u_b 的同引信任经验为 $T_{\text{同引}}(u_a, u_b)$, 整个社会信任网络中的用户数目采用 $U_{\text{社网}}$ 表示。用户 u_a 和用户 u_b 被共同邻居信任的集合定义为 $A_{H_{\max}}^{\text{同引}}(u_a, u_b)$; 设目标 u_k 对其邻居 u_a 完全信任,则 $E(T_{\text{同引}}(u_k, u_a)) = 1$, 否则为 0; 同时信任用户 u_a 和用户 u_b 的共同邻居的数量用 $|A_{H_{\max}}^{\text{同引}}(u_a, u_b)|$

表示,用户 u_a 和用户 u_b 同时被信任的经验值记为 $E(T_{\text{同引}}(u_d, u_a), T_{\text{同引}}(u_d, u_b))$ ^[10]。

根据 u_a 和 u_b 的被共同信任邻居的信任度大小计算 u_a 对 u_b 的同引信任水平,如式(15)所示:

$$T_{\text{同引}}(u_a, u_b) = \frac{\sum_{u_d \in A_{H_{\max}}^{\text{同引}}(u_a, u_b)} L(T_{\text{同引}}(u_d, u_a)) \times T_{\text{同引}}(u_d, u_b)}{\sum_{u_k \in A_{H_{\max}}^{\text{同引}}(u_a)} L(T_{\text{同引}}(u_k, u_a))} \quad (15)$$

结合公式(14)和(15), u_a 对 u_b 的同引信任度 $T_{\text{同引}}(u_a, u_b)$ 的计算方法为:

$$T_{\text{同引}}(u_a, u_b) = \frac{T_{\text{同引}}(u_a, u_b) \times T_{\text{同引}}(u_a, u_b)}{\sum_{u_k \in A_{H_{\max}}^{\text{同引}}(u_a)} T_{\text{同引}}(u_a, u_k) \times T_{\text{同引}}(u_a, u_k)} \quad (16)$$

Step3: 计算综合信任度。

综合聚合信任与同引信任更宜实现优势互补,提高推荐质量,鉴于此,本节将“聚合”信任与“同引”信任进行融合,计算二者综合信任度,具体如公式(17)所示:

$$T_{\text{综合}}(u_a, u_b) = \alpha T_{\text{聚合}}(u_a, u_b) + (1 - \alpha) T_{\text{同引}}(u_a, u_b) \quad (17)$$

其中, α 为 $[0, 1]$ 的可调参数,该参数可由人工设定,也可根据实验验证的方式确定最优的取值^[11]。

Step4: 矩阵合并。个性化情景与信任度矩阵合并方法与本文公式(7)与公式(8)相同,具体如式(18)所示:

$$\text{Weight} = \begin{cases} 0 & \text{if } \text{Sim}C'(u, v) > 0 \text{ and } \text{Trust}(u, v) = 0 \\ \text{Trust}(u, v) & \text{if } \text{Sim}C'(u, v) = 0 \text{ and } \text{Trust}(u, v) > 0 \\ \frac{2 \times \text{Sim}C'(u, v) \times \text{Trust}(u, v)}{\text{Sim}C'(u, v) + \text{Trust}(u, v)} & \text{if } \text{Sim}C'(u, v) > 0 \text{ and } \text{Trust}(u, v) > 0 \end{cases} \quad (18)$$

Step5:产生推荐。采用加权平均值方法计算推荐值,将权值 $Weight(u,v)$ 改为用户个性化情景相似度与信任度的复合值进行推荐:

$$P_{u,i} = \bar{R}_u + \frac{\sum_{i=1}^m (R_i(i) - \bar{R}_i) \times Weight(u,v)}{\sum_{j=1}^m Weight(u,v)} \quad (19)$$

2 实验结果及分析

2.1 实验数据集与评价标准

为验证基于个性化情景的富信任信息推荐算法及基于个性化情景的稀疏信任信息推荐算法的预测效果,本文采用美国 Minnesota 大学公开提供的 Movie lens 电影评分数据集^[12]以及 Book-Crossing 提供的图书评分数据集进行数值实验^[13],按 80:20 的比例将评分数据随机地分为训练集和测试集^[14],并选取富信任信息环境下及稀疏信任信息环境下用户情景兴趣推荐算法进行对比分析^[10]。

美国 Minnesota 大学公开提供的 Movie lens 数据集提供了 100000 条电影评分记录,广泛用于推荐技术实证检验中。该数据集包含了用户性别、年龄、职业等用户情景信息,时间情景(如时间及节假日)则可通过数据集中的时间戳数据间接获得,并且经过清理与量化的 Movie lens 数据集,可间接获得用户偏好,能更好地被推荐算法直接使用。Book-Crossing 数据集是由 Cai-Nicolas Ziegler 使用爬虫程序从 Book-Crossing 图书社区上采集的 278858 个用户对 271379 本书进行的行为信息。包含了用户 ID、置、年龄的基本信息及时间情景信息(早上中午、下午、晚上)、位置情景信息(办公场所、家、学校、餐厅)、环境情景信息(声音环境、空间环境)、状态信息(娱乐、工作、学习)等情景信息。

本文采用推荐质量测试方法中广泛使用的平均绝对误差 MAE^[15]及覆盖率 Coverage^[16]作为度量标准。

MAE 通过计算预测的评分与实际评分之间的偏差来度量推荐结果的准确性,偏差越小,预测精度越高,其计算公式为:

$$MAE = \frac{\sum_{i=1}^N |p_i - q_i|}{N} \quad (20)$$

而覆盖率 Coverage 是一项被广泛采用、用于评价推荐系统推荐覆盖度的评判度量指标,它描述了推荐系统向用户推荐的项目集合对用户兴趣的覆盖范围, Coverage 越高,覆盖能力越强,具体计算公式为:

$$Coverage = \frac{\sum_{u \in U} |IP(u) \cap IR(u)|}{\sum_{u \in U} |IR(u)|} \quad (21)$$

2.2 实验结果及分析

(1)个性化情景确定。依据 2.1.1 中(1)的具体步骤对训练集上的用户评分进行训练,从而得到基于情景的用户推荐

评价矩阵:

$$P_c = \begin{bmatrix} \text{性别} & \text{时间} & \dots & \text{偏好} \\ 4.869, & 4.881, & \dots, & 3.856 \\ 3.701, & 2.685, & \dots, & 4.189 \\ \vdots & \vdots & & \vdots \\ 4.793, & 0, & \dots, & 4.236 \end{bmatrix}$$

Movie lens 数据集基于情景的用户推荐评价矩阵

$$p_c = \begin{bmatrix} \text{时间} & \text{位置} & \dots & \text{状态} \\ 4.858, & 4.782, & \dots, & 2.957 \\ 3.775, & 1.856, & \dots, & 4.891 \\ \vdots & \vdots & & \vdots \\ 2.973, & 4.579, & \dots, & 3.632 \end{bmatrix}$$

Book-Crossing 数据集基于情景的用户推荐评价矩阵

矩阵中的行代表各个用户,列代表对应的情景,每个数值代表用户 u_i 在 c_i 下用户对信息推荐质量的评分,若系统设定 $K=5$,则取各行中评分最高的 5 个情景组成用户的个性化情景 C' 。

(2)预测结果分析。使用本文提出的基于个性化情景的富信任信息推荐算法及基于个性化情景的稀疏信任信息推荐算法与不同信任信息环境下用户情景兴趣推荐算法进行比较,具体的比较结果如图 2 所示。

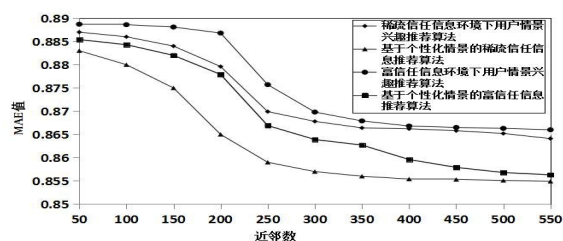


图2 四种算法在 Movie lens 数据集上的 MAE 对比结果

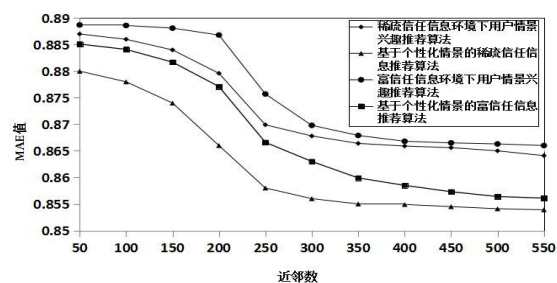


图3 四种算法在 Book-Crossing 数据集上的 MAE 对比结果

如图 2 所示,由于 Movie lens 数据集缺乏部分外部情景信息,致使选取所含情景与当前个性化情景 C' 相似的评分数据段具有一定的稀疏性,导致近邻数较少在 0~150 个邻居用户时,4 种算法的 MAE 值均较大,但本文两种的算法随着近邻数的增加,MAE 值呈现较快的下降趋势并趋于稳定,且 MAE 值小于其他两种算法。

在图 3 中,四种算法 MAE 值随邻居用户数目增加均呈现较快的下降趋势,且本文算法的 MAE 值最小,预测精度最高,这说明本文提出的两种算法优于不考虑或只考虑用户整体情景信息的信任推荐算法,因为不同移动商务情景对不同

用户信息需求影响存在差异性,并且由结果可知,同时考虑个性化情景及用户间信任关系对用户的影响能够有效地提高推荐质量。

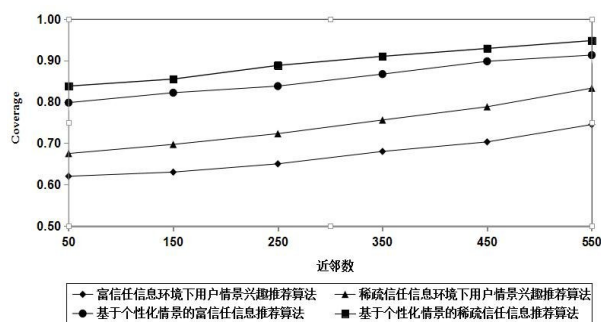


图4 四种算法在Movie lens数据集上的Coverage对比结果

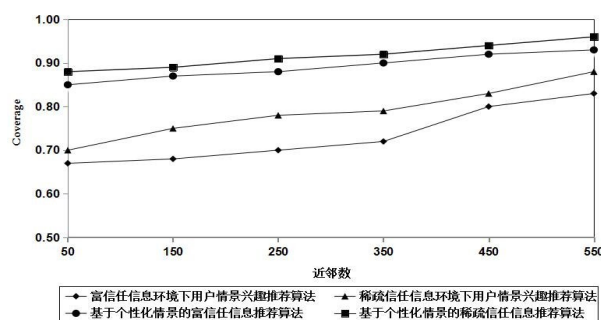


图5 四种算法在Book-Crossing数据集上的Coverage对比结果

通过图4、图5可以看出,在Coverage性能比较中,同时考虑用户个性化情景与用户间信任关系的推荐算法覆盖率指标明显高于已有的不同信任信息环境下用户情景兴趣推荐方法。首先,“聚合”与“同引”信任推荐算法在选取最近邻居段时,运用社会网络分析法挖掘出用户间潜在的信任关系,将二者相结合形成优势互补,覆盖能力更强,推荐性能更优;其次,分别加权用户个性化情景相似性与潜在信任值,进一步加强了目标用户个性化需求意愿的相似性,从而提高了推荐结果准确性。综上可知,本文提出的两种算法在合理考虑个性化情景要素的情况下,结合不同信任环境下用户间的信任关系具有更高的推荐质量。

3 结 语

针对目前移动商务信息服务个性化和准确性较低的问题,提出了基于个性化情景的不同信任信息环境下的移动商务信任推荐模型。其中,基于个性化情景的移动商务富信任信息推荐模型,通过融合用户个性化情景信息与移动商务用户间信任关系,解决数据稀疏性及个性化服务程度较低问题。其次,针对现实情况下用户间可用信任信息较少,难以获得理想的推荐结果,提出基于个性化情景的移动商务稀疏信任信息推荐模型,从“聚合”与“同引”信任两个维度入手,挖掘稀疏信任信息环境下潜在的用户信任关系,并结合用户个性化情景进行推荐。实验结果表明,本文提出的两种算法

与已有的推荐算法相比,均取得了更优的推荐效果(MAE值更小,Coverage更大),尤其是基于个性化情景的稀疏信任信息推荐算法预测精度更高(MAE值最小,Coverage最大),更为准确地预测了用户个性化信息需求。本文提出算法是在单机上实现的,面向海量数据时运算速度缓慢,未来的研究工作将考虑在并行环境中进行,以期进一步提高运行效率。

参考文献

- Golbeck J. Computer science—Weaving a Web of trust[J]. Science, 2008, 321(5896): 1640–1641.
- DePessemer T. Context aware recommendations for user-generated content on a social network site[C]// Euroitv'09: Proceedings of the Seventh European Interactive Television Conference, 2009: 133–136.
- 郑孝遥, 鲍煜, 孙忠宝, 等. 一种基于信任的协同过滤推荐模型[J]. 计算机工程与应用, 2016, 52(5): 50–54.
- 邓晓懿, 金淳, 韩庆平, 等. 基于情境聚类和用户评级的协同过滤推荐模型[J]. 系统工程理论与实践, 2013, 33(11): 2945–2946.
- 高旻, 吴中福. 基于个性化情境和项目的协同推荐研究[J]. 东南大学学报: 自然科学版, 2009, (1): 29–30.
- 周朴雄, 薛玮炜, 赵龙文. 基于个性化情境的 Multi-Agent 信息推荐研究[J]. 情报杂志, 2015, (5): 182–183.
- 杨君, 莫赞, 艾丹祥, 等. 基于降维的输入情景化多维信息推荐研究[J]. 情报理论与实践, 2013, (11): 64–65.
- 张朝旭. 移动社交网络中上下文感知推荐机制的研究与设计[D]. 北京: 北京交通大学, 2013.
- 刘海鸥. 面向大数据知识服务推荐的移动 SNS 信任模型[J]. 图书馆论坛, 2014, (10): 68–75.
- 刘海鸥. 云环境用户情景兴趣的移动商务推荐模型及应用研究[D]. 秦皇岛: 燕山大学, 2013.
- 葛冰玉. 基于信任和不信任网络的推荐算法研究[D]. 长春: 吉林大学, 2016.
- 高全力, 高岭, 杨建锋, 等. 上下文感知推荐系统中基于用户认知行为的偏好获取方法[J]. 计算机学报, 2015, 28(9): 1772–1774.
- 胡勋, 孟祥武, 张玉洁, 等. 一种融合项目特征和移动用户信任关系的推荐算法[J]. 软件学报, 2014, 25(8): 1824–1825.
- KIM J, LEE D, CHUNG K. Item Recommendation Based On Context-Aware Model for Service[J]. Multimedia Tools and Applications, 2014, 71(2): 855–872.
- 易明, 邓卫华, 徐佳. 社会化标签系统中基于组合策略的个性化知识推荐研究[J]. 情报科学, 2011, 29(7): 1093–1097.
- 谭龙江. 基于用户活动事件的社会网络推荐模型[J]. 情报科学, 2015, 32(11): 140–144.

(责任编辑: 张连峰)