



基于多社交数据源的协同推荐方法研究

王瑞琴¹, 潘俊², 李一啸³

(1. 湖州师范学院信息与控制技术研究所 湖州 313001;

2. 温州大学建模与数据挖掘研究所 温州 325035; 3. 浙江财经大学信息学院 杭州 310018)

摘要: 协同过滤推荐作为一种有效的推荐方法, 普遍存在数据稀疏性和冷启动问题, 利用社交网络的多项数据源对协同推荐方法进行了改进。为了克服评分矩阵的稀疏性问题, 提出结合用户评分相似度和用户信任度选择推荐邻居, 同时对用户相似度计算进行了改进; 提出了一种简单有效的信任推理方法, 能够识别出用户间隐含的间接信任关系, 进一步缓解了数据稀疏性问题; 为了解决推荐系统的冷启动问题, 提出综合利用项目的类型属性信息和领域专家信息进行联合推荐。实验结果表明, 提出的改进策略非常有效, 在精度和召回率方面都较已有方法具有明显改善。

关键词: 社交网络; 个性化推荐; 信任推理; 多数据源; 领域专家

doi: 10.11959/j.issn.1000-0801.2015113

Research on Collaborative Recommendation Method Based on Multiple Data Sources of Social Network

Wang Ruiqin¹, Pan Jun², Li Yixiao³

(1. Institute of Information and Control Technology, Huzhou University, Huzhou 313001, China;

2. Institute of Business Modeling and Data Mining, Wenzhou University, Wenzhou 325035, China;

3. School of Information, Zhejiang University of Finance and Economics, Hangzhou 310018, China)

Abstract: As an effective recommendation method, collaborative filtering typically has the data sparsity and cold-start problems. It was proposed that using multiple data sources of social network to overcome the above problems. First of all, both the rating similarity and the social trust between users were considered to resolve the data sparsity problem. Then a simple and effective trust reasoning method was proposed to identify the implicit trust relationship between users. In order to solve the cold-start problem, information of the category of items and domain experts was used for joint recommendation. Experimental results show that the proposed algorithm has significantly better precision and recall than existing methods.

Key words: social network, personalized recommendation, trust inference, multiple data source, domain expert

收稿日期: 2015-03-07; 修回日期: 2015-04-29

基金项目: 国家自然科学基金资助项目 (No.61402336, No.61403338), 国家教育部科学基金资助项目 (No.14YJCZH152), 浙江省自然科学基金资助项目 (No.LQ12F02008), 浙江省科技计划基金资助项目 (No.2013C31138, No.2012C33086)

Foundation Items: The National Science Foundation of China (No.61402336, No.61403338), Science Foundation of Ministry of Education of China (No.14YJCZH152), Zhejiang Provincial Natural Science Foundation (No.LQ12F02008), Zhejiang Provincial Technology Program (No.2013C31138, No.2012C33086)

论文引用格式: 王瑞琴, 潘俊, 李一啸. 基于多社交数据源的协同推荐方法研究. 电信科学, 2015113

Wang R Q, Pan J, Li Y X. Research on collaborative recommendation method based on multiple data sources of social network. Telecommunications Science, 2015113



1 引言

随着互联网和移动终端的飞速发展,社交网络(social network)已经成为人们生活中必不可少的组成部分。从国外的 Facebook、Twitter、Google+、YouTube 到国内的新浪微博、腾讯微博、开心网、人人网等,用户数量都在逐年飙升^[1]。社交网络在给人们带来无数便利的同时,其中海量的数据信息也会使人们感到头晕目眩,无所适从。因此,合理利用社交网络环境中的多种数据源,深入挖掘社交用户的个性化偏好及用户间的信任关系,进而实现人性化的信息推送服务,可以极大地改善用户的使用体验,从而提高社交网络的用户黏度与访问量。

传统的个性化推荐方法主要分为两类:基于内容(content-based)的推荐和协同过滤(collaborative filter, CF)推荐,其中协同过滤推荐技术比较成功,得到了广泛的应用。协同过滤推荐算法基于用户—项目历史评分矩阵,采用一定的相似度量算法,如 Cosine 相似度、Pearson 相关性等,计算用户间的相似程度,基于相似用户对目标项目的评分估计当前用户对目标项目的喜好程度,进而实现推荐。数据稀疏性和冷启动问题一直是协同过滤推荐方法面临的两大难题,幸运的是,目前很多电子商务站点都开设了社交网络社区,其中包含许多有价值的数据源,如用户的个人描述、用户间的社会信任关系、用户产生的信息内容、用户使用的社会标签、用户的浏览记录等。从社交网络中的多项数据源中挖掘用户的兴趣、偏好,进而提供个性化的信息推荐服务,是一个值得研究的课题。

2 国内外研究现状

2009 年, Lazer 等在《Science》杂志上提出计算社会学(computational social science)的概念,指出网络上的大量信息,如博客、论坛、聊天、消费记录、电子邮件等,都是对现实社会的人及组织行为的映射^[2],网络数据可用于分析个人和群体的行为模式,深化人们对生活、组织和社会的理解。

近年来,研究者在基于社交网络的信任计算和个性化推荐方面进行了大量研究,如 Golbeck 用信任关系定量分析社交网络中的社会关系,提出了信任网络,并开发了电影评论网 FilmTrust,利用信任关系作为电影评分的权值为用户推荐电影^[3]。Massa 等利用来自 Epinions 的信任数据建立信任预测模型 MoleTrust,以此来预测社会关系网络中

传播的信任值,并利用这种信任值进行推荐^[4]。Caverlee 等提出了在线社会网络中构建信任的 SocialTrust 框架,将链接质量和反馈评分集成起来评估信任,得到了较高的推荐准确率^[5]。

社交网络是人们在线交流的平台,也是信息传播的媒介,这个想法激发了基于信息传播的个性化推荐方法的研究。Song 等提出了基于信息扩散理论的个性化推荐方法,并从理论上证明了该信息传播模型的收敛性^[6]。Gruhl 等将信息在博客中的传播方式模拟为流行病的传播方式,从宏观上对收集来的博客内容信息进行主题聚类,从微观上通过追踪用户发帖和跟帖的时间信息对信息传播进行建模和预测,推荐用户下个时间段感兴趣的信息^[7]。Leskovec 等分析了推荐网络和社交网络中常见的传播模式,根据概率图模型给出了推理网络中传播路径的算法^[8]。Golbeck 提出了基于最短信任路径的信任推理模型 TidalTrust,模型计算复杂度低、可扩展性强^[9]。Zhang 等对 TidalTrust 模型进行扩展,引入信任评价和可信度因子,使用加权图结构计算网络实体间的信任度^[10]。Caverlee 等基于社会关系和用户反馈进行信任计算和推理,提出了基于信誉的动态信任推理模型 SocialTrust^[11]。

国内在基于社交网络的个性化推荐方面的相关研究也取得了一定的进展,如中国科学院计算所的程学旗对前 k 项排序学习(top- k learning to rank)推荐算法的研究^[12],清华大学的唐杰对学术社会网络 AmetMiner 的推荐研究^[13]和对社会关系的预测研究^[14],电子科技大学的周涛等对复杂网络、人类动力学和推荐系统的研究^[15],清华大学的林闯^[16]、武汉大学的陈晶等^[17]利用博弈原理基于用户间的交互行为对用户间信任关系的研究,北京邮电大学的李晓峰等借鉴社会心理学中信任产生的原理对信任度的计算^[18],大连海事大学的邢星利用信任传递算子对间接信任度的计算^[19]等。

综上所述,研究者们主要利用社交网络中的信任关系改善个性化推荐的性能,在一定程度上缓解了推荐的数据稀疏性问题,并取得了一定的研究成果。然而已有方法主要关注用户显式信任关系的计算及其推理,很多有价值的隐式信任关系往往被忽略,对于信任传播和信任融合方面的研究也不够深入。很明显,已有研究没有充分利用社交网络中的多种知识源,对于推荐的冷启动问题也甚少提及,这些问题都有待进一步研究。

本文对基于多社交数据源的个性化推荐进行了深入

的探索,主要有:采用协同过滤的思想,结合用户历史评分信息和用户间的信任关系克服推荐中的数据稀疏性问题;基于信任的传播性特征,设计相应的信任推理方法,发现用户间潜在的间接信任关系,扩大用户的社交范围,进而提高推荐的召回率;为了解决系统的冷启动问题,进一步探索用户间的隐式信任关系,利用项目的类型属性,通过用户对同类型项目的评分来估计用户对目标项目的喜好程度,同时利用领域专家对新用户进行推荐。

3 基于多社交数据源的协同推荐

3.1 基于用户历史评分的用户相似度计算

用户通常会对已购买或使用过的项目给出相应的评价,这些历史评价信息是社交网络中最常见的一类数据源。简单起见,大多数社交站点要求将评分量化到一定的数量范围内,如参考文献[1,5]。通过数据抽取可以得到一个表示用户—项目历史评分的矩阵 R ,其中,行表示用户 U ,列表示项目 T ,矩阵元素 r_{ij} 表示用户 u_i 对项目 t_j 的评分。

传统的协同过滤推荐中,通常以 R 的行作为用户向量,采用 Pearson 相关系数度量用户间的相似度,表示为:

$$\text{sim}(i, j) = \frac{\sum_{k \in T_{ij}} (r_{ik} - \bar{r}_i) \times (r_{jk} - \bar{r}_j)}{\sqrt{\sum_{k \in T_{ij}} (r_{ik} - \bar{r}_i)^2} \times \sqrt{\sum_{k \in T_{ij}} (r_{jk} - \bar{r}_j)^2}} \quad (1)$$

其中, $\bar{r}_i$ 和 $\bar{r}_j$ 表示用户 u_i 和 u_j 的评价均值, T_{ij} 表示用户 u_i 和 u_j 共同评价过的项目集, r_{ik} 和 r_{jk} 表示用户 u_i 和 u_j 对项目 t_k 的评分值。

在此基础上,基于用户间的共评项目数 $|T|$ 对用户相似度的计算进行了改良。假设 $\text{sim}(i, j) = \text{sim}(i, k)$, 但是 $|T_{ij}| > |T_{ik}|$, 即用户 u_i 和 u_j 之间的共评项目数大于用户 u_i 和 u_k 之间的共评项目数,显然,用户 u_i 和 u_j 之间的相似度 $\text{sim}(i, j)$ 应该比用户 u_i 和 u_k 之间的相似度 $\text{sim}(i, k)$ 大。改进的用户相似度计算如下:

$$\text{sim}(i, j)' = \text{sim}(i, j) \times \frac{1}{1 + e^{-\frac{|T_{ij}|}{2}}} \quad (2)$$

式(2)使用指数函数来避免用户间的共评项目数过大对相似度造成的影响,从而使用户相似度落在 $[0, 1]$ 区间。当用户共评项目数足够大时,式(2)的右边项趋向于 1; 对于很小的用户共评项目数,该项值约为 0.6; 当用户共评项目数大于 5 时,该项值大于 0.9。

3.2 结合用户信任关系的协同过滤推荐

传统的协同过滤推荐方法基于相似用户对目标项目

的评分来估计当前用户对目标项目的喜好程度,该方法中的 r_{ik} 表示为:

$$r_{ik} = \bar{r}_i + \frac{\sum_{j \in KNN(i)} \text{sim}(i, j)' \times (r_{jk} - \bar{r}_j)}{\sum_{j \in KNN(i)} \text{sim}(i, j)'} \quad (3)$$

其中, $KNN(i)$ 表示用户 u_i 的 K -近邻集。通常,随着用户数和项目规模的不断增加,历史评分矩阵会变得越来越稀疏,导致 $\text{sim}(i, j)'$ 很不准确,从而影响评分预测和个性化推荐的性能。

在现实社会中,人们在选购商品或做其他决定时,往往更加倾向于采纳信任朋友给出的意见。所以一些电子商务站点中的 SNS 社区让用户同自己的朋友保持联系,及时了解对方的最新状况和动态信息,以获得更可靠的购物经验与建议。本文拟引入社交网络中用户间的信任关系来克服推荐系统的数据稀疏性问题,提出改进的预测评分计算方法,表示如下:

$$r_{ik} = \bar{r}_i + \frac{\sum_{j \in KNN(i)} \text{corr}(i, j) \times (r_{jk} - \bar{r}_j)}{\sum_{j \in KNN(i)} \text{corr}(i, j)} \quad (4)$$

其中, $\text{corr}(i, j)$ 为用户 u_i 和 u_j 之间的相关性,表示为:

$$\text{corr}(i, j) = \frac{\text{sim}(i, j)' + \text{tr}(i, j)}{2} \quad (5)$$

其中, $\text{tr}(i, j)$ 表示用户 u_i 和 u_j 之间的信任度(具体计算方法见第 3.3 节)。这样一来,用户推荐邻居的选择不仅与用户的历史评分有关,还与用户间的社会关系有关,更符合现实中用户间的信任规律,包含熟悉信任度和相似信任度两个方面。

3.3 基于最短路径的信任推理方法

尽管大多数社交网络都允许用户建立自己的信任圈,但现实中却发现,用户很少显式地声明自己信任的人,而是简单地维系几个最为重要的朋友。所以,只利用已有的信任关系对推荐算法进行改善,效果并不理想。本文拟基于社交网络中有限的显式信任关系,采用相应的信任推理策略,发现用户间的间接信任关系,进而提高推荐的性能。

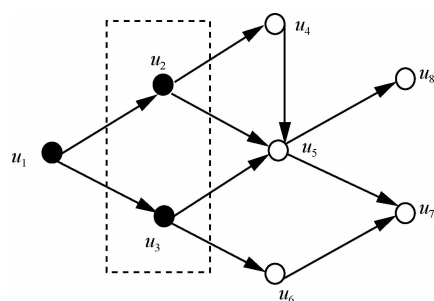
信任是一个社会学范畴,具有传递性,即如果 A 信任 B、B 信任 C,那么 A 对 C 也具有一定程度的信任,正如在现实社会中常常把“朋友的朋友”也当成自己的朋友一样。但是随着信任传递路径的增长,信任程度将会逐渐减弱,过于弱化的信任关系使得推荐的效果不够准确,因此,本



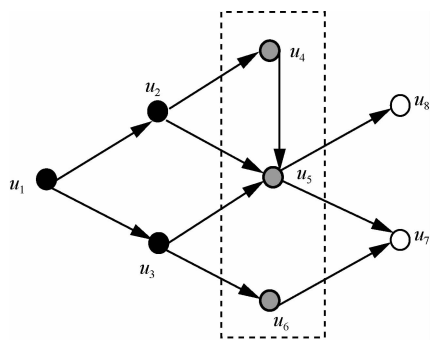
文只进行了 2-step 信任推理。推理过程非常简单,如图 1 所示,首先,利用用户声明的直接信任关系构造有向信任图,然后采用广度优先遍历(breadth first traversal,BFS)方法寻找当前用户的间接信任用户,最后采用一定的信任融合机制计算用户间的间接信任度。

直接信任用户展示如图 1(a)所示,假设用户 u_1 为当前用户, u_2, u_3 为 u_1 的直接信任用户,此时要获得 u_1 的间接信任用户,只需在信任图中进行 BFS 遍历即可。遍历一步之后的情形如图 1(b)所示,此时,用户 u_4, u_5, u_6 成为 u_1 的 1-step 间接信任用户;遍历 2 步之后的情形如图 1(c)所示,此时,用户 u_7, u_8 成为 u_1 的 2-step 间接信任用户。

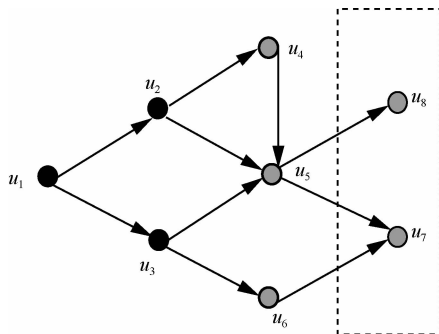
通常假定直接信任用户之间的信任度为 1,那么定义



(a) 直接信任用户展示



(b) 1-step信任用户展示



(c) 2-step信任用户展示

图1 信任推理过程

间接信任度的计算方法如下:

$$tr(i, k) = \frac{1}{2^d} \times \frac{1}{1 + e^{-\frac{n}{2}}} \quad (6)$$

其中, d 表示信任推理的步数, n 表示从用户 u_i 到 u_k 的 d -step 信任推理路径的个数。式(6)中的指数函数的使用有两个作用:保证了信任推理路径越多的用户之间的信任度越高;避免了过多信任推理路径导致的信任度过大的问题。

图 2 展示了间接信任度的计算过程。 u_1 和 u_2, u_1 和 u_3 具有直接信任关系,因此, $tr(1,2)=tr(1,3)=1$; u_1 和 u_4 之间没有直接信任关系,但是有 1 条 1-step 信任推理路径 ($u_1-u_2-u_4$),所以 $d=1, n=1$,通过式(6)计算得到 u_1 对 u_4 的间接信任度 $tr(1,4)=0.31$ 。同理,从 u_1 到 u_5 有 2 条 1-step 信任推理路径,所以 $d=1, n=2, tr(1,5)=0.37$;从 u_1 到 u_6 有 1 条 1-step 信任推理路径,所以 $d=1, n=1, tr(1,6)=0.31$;从 u_1 到 u_7 有 3 条 2-step 信任推理路径,所以 $d=2, n=3, tr(1,7)=0.2$ 。至此,得到了 u_1 的所有间接信任用户及其信任度估计值。

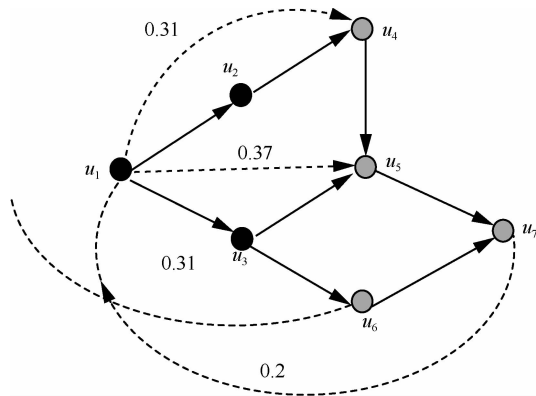


图2 间接信任度的计算

3.4 基于项目属性对新项目的评分预测

基于以上步骤,数据稀疏性问题基本解决了,但是冷启动问题仍然存在。对于新加入的项目,从没有被用户评价过,要想获得某用户对该类项目的喜好程度,需要从项目本身的属性着手。一般的电子商务网站中发布的项目都具有一定的属性,如名称、类别、价值、产地等。可基于项目属性度量项目之间的相似程度,利用用户对相似项目的历史评分来估计用户对目标项目的评分。简单起见,本文基于项目的类别属性,使用用户对一类项目的评分均值来预测用户对目标项目的喜好程度,仅以此作为基于项目属性解决冷启动问题有效性的判据。用户 i 对目标 k 的喜好程度表示为:

$$pr(i, k) = \frac{1}{N} \sum_{t_j \in C_k} r(i, j) \quad (7)$$

其中, C_k 表示目标项目 t_k 所属的类别, N 表示用户 u_i 评价过的类别为 C_k 的项目总数。

3.5 基于领域专家对新用户的评分预测

对于一个刚进入网站的新用户, 或者一个“潜水”型用户, 他们还没有对网站中任何项目进行过评分, 要预测其对某项目的喜好程度, 传统的推荐方法显得力不从心。此时, 可以将现实社会中人们对权威专家的隐式信任引入推荐过程中, 基于领域专家的评分信息对冷启动用户的项目评分进行预测, 基于预测结果进行推荐。一些社交网络中的领域专家已在系统数据集中显式给出, 例如, 本文采用的测试数据集 Rich Epinions 中就给出了部分类别的领域专家, 并按专业化程度将专家进一步划分为 Category Lead、Top Reviewer、Advisor 3 类。基于领域专家的评分预测方法如下:

$$pr'(i, k) = \frac{1}{N} \sum_{j \in E(C_k)} w_j \times r(j, k) \quad (8)$$

其中, $E(C_k)$ 表示 C_k 类型中的领域专家集, w_j 表示专家 j 的专业化程度权重。

如果社交网络中没有显式指明领域专家, 可以通过用户的个体特征、结构重要性、社会影响力等因素综合评价用户在社交网络中的重要性, 以此作为领域专家的判据: 用户的个体特征主要指用户的专业知识水平、信息含量、所含信息的质量等; 用户的结构重要性可以在用户交互图中利用图论方法来衡量, 如点度中心度 (degree)、中介中心度 (betweenness) 等度量指标; 用户的社会影响力可通过用户在社交网络中的活跃程度、信息传播能力以及社会信誉度等进行度量。

4 实验结果与分析

本文采用的测试数据集为 Rich Epinions, 该数据集是由 Simon 等人于 2011 年 9 月从 Epinions 网站收集到的, 是

原有 Epinions 数据集的一个扩充版本, 其中包含了用户对项目的评分信息、用户间的信任关系、项目的类型属性、类型的层次结构信息和领域专家等信息。数据集中包含用户 131 228 个、项目 317 755 个、评分 1 127 673 个、信任关系 538 817 个, 数据稀疏度为 0.003%, 有 4 287 个用户既没有评分信息也没有信任关系, 属于新用户或潜水用户。

预测精度 (precision) 和召回率 (recall) 是衡量推荐算法的重要参数。预测精度通过比较预测的推荐评价值与测试集中的真实评价值来衡量预测的准确程度, 其中绝对平均误差 (mean absolute error, MAE) 是一个广泛用于测量评分差异的参数, MAE 值越低, 表示推荐系统预测的精度越高。MAE 表示为:

$$MAE = \frac{\sum_{i=1}^n |PR_i - P_i|}{n} \quad (9)$$

其中, PR_i 为预测得到的评分值, P_i 为真实评分值, n 为预测的次数。

召回率也叫查全率, 指的是通过某算法预测出来的评分数目与所有待测评分数目的比值。假设测试集中的待测项目为 n 个, 通过某算法得到其中 m 个项目的预测评分, 那么该算法的召回率为 $RC=m/n$ 。F1 则是一个综合考虑预测精度 (PC) 和召回率 (RC) 的参数, 表示为:

$$F1 = \frac{2 \times PC \times RC}{PC + RC} \quad (10)$$

当 F1 较高时, 说明该试验方法比较有效。

为了验证信任推理对推荐性能的影响, 将信任推理的步数逐渐增加, 同时观察推荐的性能。六度空间理论认为, 一个人和世界上任何一个陌生人之间所间隔的人数不会超过 6 个, 也就是说, 最多通过 6 个中间人就能够认识任何一个陌生人。因此, 本文设定信任推理的最大步数为 6, 实验结果见表 1。

从表 1 中可以看出, 2-step 信任推理的整体推荐最好, 多步信任推理虽然可以进一步提高推荐的召回率, 但同时也降低了推荐的精度。原因在于, 多步信任推理得到了许

表 1 基于不同步数信任推理的推荐性能比较

	1-step	2-step	3-step	4-step	5-step	6-step
绝对平均误差	0.69	0.62	0.74	0.88	1.13	1.98
召回率	0.28	0.44	0.45	0.45	0.46	0.46
F1	0.42	0.59	0.58	0.58	0.58	0.52



多信任关系较弱的用户邻居,弱信任邻居的推荐很可能不准确,从而影响了推荐的精度。进一步观察实验结果发现,多步信任推理并没有如预想中那样稳步提高推荐的召回率,通过分析原始信任数据发现,信任用户之间形成了高内聚低耦合的信任簇,所以,继续扩展推荐的深度也鲜少有新用户的增加,这一现象和现实生活中固定的社团、社区比较类似,容易理解。

为了进一步验证本文方法的有效性,从实验数据集中随机抽取 1 000 组评分数据,从传统的 K-近邻协同过滤算法开始,逐个增加推荐改进策略,进行评分预测,取 10 次测试的平均值进行对比分析,实验结果见表 2。其中,BASE 代表基于项目评分均值的推荐,CF 代表传统的协同过滤推荐方法,TCF 代表基于用户信任度的 CF,STCF 代表融合用户相似度与用户信任度的 CF,STCF1 代表基于改进用户相似度的 STCF,STCF2 代表基于 2-step 信任推理的 STCF1,STCF3 代表结合项目类型属性的 STCF2,STCF4 代表结合领域专家推荐的 STCF3。

表 2 不同推荐算法的性能比较

方法	绝对平均误差	召回率	F1
BASE	2.71	0.94	0.62
CF	0.80	0.16	0.27
TCF	0.54	0.12	0.21
STCF	0.71	0.20	0.32
STCF1	0.69	0.20	0.32
STCF2	0.62	0.44	0.59
STCF3	0.65	0.70	0.78
STCF4	0.66	0.72	0.79

从表 2 中可以看出,基于项目流行度的 BASE 可以达到的召回率很高,但精度却非常低,达不到推荐的目的。TCF 的预测精度最高,说明基于信任朋友的推荐非常可靠,遗憾的是,召回率很低,这是由于数据集中信任关系的稀疏性所致,因此,需要进一步结合用户历史评分数据进行推荐。STCF 较 TCF 和 SCF 的召回率都有所提高,说明评分数据和信任数据所涵盖的范围有所不同,在预测推荐时可以互为补充。从 STCF1 到 STCF4 的精度和召回率在逐步上升,说明本文提出的改进策略起到了一定的作用,其中信任推理和类型属性的应用对于召回率的提高发挥了较大的作用。从 STCF2 到 STCF3 的精度有所下降,说明为了解决新项目冷启动问题,基于类型

属性的评分预测引入了一定的噪声数据,究其原因,在于基于类型决定项目相似度的粒度过大,在有条件的情况下,应进一步结合项目的其他属性信息进行项目相似度的计算。

5 结束语

日益流行的社交网络包含了丰富的数据信息,本文尝试利用其中的历史评分数据、用户信任关系、项目属性信息、领域专家信息等多种数据源,并结合相应的用户相似度改进策略与信任推理方法,对传统的协同过滤推荐算法进行了一系列改进。对比实验结果表明,本文提出的方法在改善推荐性能方面非常有效,这为基于社交网络的个性化推荐技术的研究指明了方向。

下一步将继续挖掘社交网络中的其他有效信息源,如用户标签、交互信息等,同时引入更加有效的推荐算法,并考虑信任的动态性和上下文相关性等特征,深入研究社交网络环境中的个性化推荐技术。

参考文献

- 1 徐雅斌,刘超,武装.基于用户兴趣和推荐信任域的微博推荐.电信科学,2015,31(1):9~16
Xu Y B, Liu C, Wu Z. Micro-blog recommendation based on user interests and recommendation trust domain. Telecommunications Science, 2015, 31(1): 9~16
- 2 Lazer D, Pentland A S, Adamic L, *et al.* Computational social science. Science, 2009, 323(1): 721~723
- 3 Golbeck J. Weaving a web of trust. Computer Science, 2008, 321 (5896):1640~1641
- 4 Massa P, Bhattacharjee B. Using trust in recommender systems: an experimental analysis trust management. Lecture Notes in Computer Science, 2004(2995): 221~235
- 5 Caverlee J, Liu L, Webb S. The socialtrust framework for trusted social information management: architecture and algorithms. Information Sciences, 2010, 18(1): 95~112
- 6 Song X D, Tseng B L, Lin C Y, *et al.* Personalized recommendation driven by information flow. Proceedings of the ACM SIGIR Conference on Research and Development in Information Retrieval, Seattle, USA, 2006: 509~516
- 7 Gruhl D, Guha R, Liben-Nowell D, *et al.* Information diffusion through blogspace. Proceedings of the 13th International World Wide Web Conference, New York, USA, 2004: 43~52
- 8 Rodriguez M G, Leskovec J, Krause A. Inferring networks of diffusion and influence. Proceedings of the 16th ACM SIGKDD Conference on Knowledge Discovery and Data Mining, Washington, USA, 2010: 1~38

- 9 Golbeck J A, Hendler J. Computing and applying trust in web-based social networks (doctor dissertation). University of Maryland at College Park, 2005
- 10 Zhang Y, Chen H J, Wu Z H. A social network-based trust model for the semantic Web. Proceedings of the Third International Conference on Autonomic and Trusted Computing, Wuhan, China, 2006: 183~192
- 11 Caverlee J, Liu L, Webb S. Towards robust trust establishment in web-based social networks with socialtrust. Proceedings of the International Conference on World Wide Web, Beijing, China, 2008: 1163~1164
- 12 刘建国,周涛,郭强等. 个性化推荐系统评价方法综述. 复杂系统与复杂性科学, 2009, 6(3): 10~15
Liu J G, Zhou T, Guo Q, *et al.* Overview of the evaluated algorithms for the personal recommendation systems. Complex Systems and Complexity Science, 2009, 6(3): 10~15
- 13 田立勤,林闯. 可信网络中一种基于行为信任预测的博弈控制机制. 计算机学报, 2007, 30(11): 1930~1939
Tian L Q, Lin C. A kind of game-theoretic control mechanism of user behavior trust based on prediction in trustworthy network. Chinese Journal of Computer, 2007, 30(11): 1930~1939
- 14 陈晶,杜瑞颖,王丽娜等. 网络环境下一种基于概率密度的信任博弈模型. 电子学报, 2010, 38(2): 427~433
Chen J, Du R Y, Wang L N, *et al.* A trust game method basing on probability model in networks. Chinese Journal of Electronics, 2010, 38(2): 427~433
- 15 乔秀全,杨春,李晓峰等. 社交网络服务中一种基于用户上下文的信任度计算方法. 计算机学报, 2011, 34(12): 2403~2414
Qiao X Q, Yang C, Li X F, *et al.* A trust calculating algorithm based on social networking service users' context. Chinese Journal of Computer, 2011, 34(12): 2403~2414
- 16 邢星. 社交网络个性化推荐方法研究 (硕士学位论文). 大连海事大学, 2013
Xing X. Research on recommendation methods in social networks (master dissertation). Dalian Maritime University, 2013
- 17 Simon M, Emmanuel G, Lionel M, *et al.* RED: a rich epinions dataset for recommender systems. Tech Report at LIRIS UMR CNRS, 2012

[作者简介]



王瑞琴,女,博士,湖州师范学院讲师,主要研究方向为数据挖掘、自然语言处理。



潘俊,男,博士,温州大学讲师,主要研究方向为机器学习、数据挖掘。



李一啸,男,博士,浙江财经大学副教授,主要研究方向为复杂系统、复杂网络。

2015113-7

2015年4月电话用户分省情况

2015年4月份统计数据显示,全国固定电话用户数为24 388.6万户,其中城市电话用户数为17 705.7万户,农村电话用户数为6 682.9万户;全国移动电话用户数为129 295.5万户。

全国统计的电话用户主要分布在东部、中部和西部地区。东部固定电话用户数为13 129.2万户,其中城市电话用户数为9 452.7万户,农村电话用户数为3 676.5万户;

东部移动电话用户数为64 025.9万户。中部固定电话用户数为5 956.8万户,其中城市电话用户数为4 209.1万户,农村电话用户数为1 747.7万户;中部移动电话用户数为33 694.3万户。西部固定电话用户数为5 302.7万户,其中城市电话用户数为4 044.0万户,农村电话用户数为1 258.7万户;西部移动电话用户数为31 574.8万户。

(来源:工业和信息化部运行监测协调局)