软件学报 ISSN 1000-9825, CODEN RUXUEW
Journal of Software,2015,26(6):1356−1372 [doi: 10.13328/j.cnki.jos.004831]
©中国科学院软件研究所版权所有.
E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel: +86-10-62562563

社会化推荐系统研究*

孟祥武[1,2], 刘树栋[1,2], 张玉洁[1,2], 胡 勋[1,2]

[1](智能通信软件与多媒体北京市重点实验室(北京邮电大学),北京 100876)

[2](北京邮电大学 计算机学院,北京 100876)

通讯作者: 孟祥武, E-mail: mengxw@bupt.edu.cn

摘 要: 近年来,社会化推荐系统已成为推荐系统研究领域较为活跃的研究方向之一.如何利用用户社会属性信息缓解推荐系统中数据稀疏性和冷启动问题、提高推荐系统的性能,成为社会化推荐系统的主要任务.对最近几年社会化推荐系统的研究进展进行综述,对信任推理算法、推荐关键技术及其应用进展进行前沿概括、比较和分析.最后,对社会化推荐系统中有待深入研究的难点、热点及发展趋势进行展望.

关键词: 推荐系统;协同过滤;信任推理;矩阵分解;因子分解机

中图法分类号: TP18

中文引用格式: 孟祥武,刘树栋,张玉洁,胡勋.社会化推荐系统研究.软件学报,2015,26(6):1356−1372. http://www.jos.org.cn/1000-9825/4831.htm

英文引用格式: Meng XW, Liu SD, Zhang YJ, Hu X. Research on social recommender systems. Ruan Jian Xue Bao/Journal of Software, 2015,26(6):1356−1372 (in Chinese). http://www.jos.org.cn/1000-9825/4831.htm

Research on Social Recommender Systems

MENG Xiang-Wu[1,2], LIU Shu-Dong[1,2], ZHANG Yu-Jie[1,2], HU Xun[1,2]

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

[2](School of the Computer Science, Beijing University of Posts and Telecommunication, Beijing 100876, China)

Abstract: Social recommender systems have recently become one of the hottest topics in the domain of recommender systems. The main task of social recommender system is to alleviate data sparsity and cold-start problems, and improve its performance utilizing users' social attributes. This paper presents an overview of the field of social recommender systems, including trust inference algorithms, key techniques and typical applications. The prospects for future development and suggestions for possible extensions are also discussed.

Key words: recommender system; collaborative filtering; trust inference; matrix factorization; factorization machine

推荐系统作为一种有效的信息过滤手段,是当前解决信息过载问题及实现个性化信息服务的有效方法之一.目前,主流推荐系统可以分为 4 类[1]:基于内容的推荐、协同过滤推荐、基于知识的推荐和组合推荐.

- 基于内容的推荐源于信息检索技术,不依赖于用户对项目的评价信息,侧重考察候选推荐项目与用户特征的匹配程度;
- 协同过滤推荐主要包括两类:
  - 一类是基于模型的方法.此方法利用概率统计模型或者机器学习方法,在训练集上构建用户特征模型(比如线性规划模型、统计模型、贝叶斯模型、概率相关模型、决策树模型、图模型、最大熵模型等),依此进行推荐.优点在于稳定性好;缺点是训练时间长、计算复杂性高;
  - 另一类是启发式方法,也是目前应用比较普遍的一种协同过滤推荐方法.该方法需要建立用户-

* 基金项目: 国家自然科学基金(60872051); 北京市教育委员会共建项目
收稿时间: 2014-04-25; 修改时间: 2014-12-02, 2015-02-15; 定稿时间: 2015-03-09

© 中国科学院软件研究所 http://www.jos.org.cn

项目评分矩阵,根据相似用户具有相似偏好的假设进行推荐.在用户评分信息充分的情况下,通过相似度的计算,可以快速为用户找到偏好相似的其他用户,从而实现协同推荐;但是在冷启动情况下,即用户评分信息很少,或者没有评分信息时,就显得无能为力了.因为在这种情况下,此方法找不到与该用户有相似评分模型的其他用户,也就不能基于相似用户的偏好也相似地假设进行推荐;

- 基于知识的推荐是一种基于特定领域规则或实例的推理方法,其优点在于不需要建立用户需求偏好模型,缺点是难以制定合理的推理规则;
- 组合推荐是为了克服上述各种推荐技术的弱点,对 3 种推荐方法的组合应用,其中,基于内容的推荐与协同过滤推荐组合是目前应用比较广泛的一种.

为了解决冷启动问题,学者们提出了社会化(social recommendation)推荐方法[2,3].这种方法主要根据用户之间的社会关系信息构建用户之间的社会化关系网络,对于一个新用户,只要网络中存在一个用户与此用户有直接或者间接的社会关系,就可以根据这种社会关系和已知用户兴趣模型,对新用户产生适宜推荐.这种推荐策略是合理、科学的,一方面源于社会网络分析(social network analysis,简称 SNA)的重要研究成果[4]:网络社区中,相互联系的群体,受社会因素的相互影响,往往表现出相似的兴趣爱好及行为规范;另一方面,伴随着具有 Web 2.0 特征的社会化网络的广泛应用,特别是在线社交网络的盛行,网络用户之间的活动行为表现得越来越社区化和网络化.此外,在线社会化网络的数据分析结果也证明了社会化推荐模型的正确性,而且用户的社会属性信息确实能够提高推荐系统的性能.这使得社会化推荐系统逐渐发展成为一个独立的研究方向,引起国内外学者的研究热情.

目前,社会化推荐系统逐渐成为了推荐系统领域的重要研究方向之一.有许多大学和研究机构对社会化推荐系统的理论、方法及应用展开了深入研究[5−14].同时,ACM 推荐系统年会(ACM Conf. on Recommender Systems,简称 RecSys)自 2009 年开始涉及社会化推荐系统的专题讨论会(Workshop on Recommender Systems & the Social Web[15]),在 2011 年的专题研讨会上指出了社会化推荐系统领域的几个发展和研究主题[16],讨论了该领域的研究热点和难点:(1) 案例研究及新的社会化推荐应用;(2) 以社区为基础的系统体制(economy of community-based systems):利用推荐系统鼓励用户持久地参与;(3) 社会化网络与大众分类法的进展:朋友、标签、书签、博客、音乐、社区推荐等等;(4) 推荐系统的跨界应用,Web 2.0 用户界面及多媒介的推荐系统;(5) 系统知识创作和综合人工智能;(6) 在推荐过程中直接涉及的用户或者社团推荐应用;(7) 利用大众分类法、社会化网络信息、交互信息、用户背景及社团等的推荐方法;(8) 有信任和信誉意识的社会化推荐;(9) 利用本体论或者微格局的语义网络推荐系统;(10) 社会化推荐技术的经验评估:成功和失败方法.在 2013 年专题讨论会[17]上,又进一步提出以下 4 个发展方向:(1) 异构网络上的社会化推荐;(2) 社会化推荐中的隐私保护问题;(3) 移动社会化推荐;(4) 社交网络与用户-项目关系网络的融合推荐.社会化推荐系统满足了互联网中新问题和新技术发展要求,具有较高的研究价值和应用前景.特别是移动互联网发展与应用,移动社会化推荐系统的研究也逐渐引起学术界和工业界的极大关注,并且取得了阶段性研究成果[18],例如移动业务推荐[19]、移动商务推荐[20−22]、基于位置的移动推荐服务[23,24]等等.此外,移动社交网络的位置签到与位置共享功能的出现,为获取用户日常社会活动时空数据提供了新的途径,从而推动了基于位置的社会化网络推荐系统兴起与发展[25−28].

本文主要关注当前社会化推荐领域的主流研究与应用进展,希望为未来的应用研究提供一定的参考与帮助.第 1 节概括社会化推荐系统的基本概念、形式化定义、基本框架模型及典型信任推理算法.第 2 节对社会化推荐生成技术进行归类和对比分析.第 3 节对社会化推荐系统中有待深入研究的难点和发展热点进行分析与展望.最后是结束语.

1 社会化推荐系统

推荐系统作为个性化服务的重要实现技术之一,起初主要分析用户-项目二元关系,构建用户的偏好模型,发现用户可能感兴趣的项目集,并主动推荐给用户,达到个性化服务的目的.但是随着推荐应用场景的不断扩

© 中国科学院软件研究所 http://www.jos.org.cn

展,仅仅依据用户-项目的二元关系,直接为用户生成推荐结果往往是偏颇、不合时宜的.例如:有的用户在旅游目的地想主动接收一些风味特色餐厅或者地域特产商店的信息;用户处于意志低沉状态时,不再愿意接受深沉的音乐,即使他平时比较喜欢这种音乐.考虑用户上下文信息(如时间、位置、周围人员、情绪、活动状态、网络条件等)的上下文推荐系统[29],满足了不同上下文环境下用户对信息的个性化需求,具有普适性和个性化两方面的优点.

Web 2.0 的迅速发展和广泛应用,促使传统推荐系统融合用户的社会属性信息进行推荐.在传统推荐系统中,大多数根据个体用户对项目的历史评分数据,建立用户兴趣偏好模型,并依此为用户产生推荐结果.随着Twitter,Facebook,Orkut,MySpace,YouTube 等在线社交网络的流行,用户社会属性信息及社会行为数据的收集更加容易,用户社会行为不仅仅是用户个人兴趣和爱好的体现,而且在一定程度上受到与其有社会关系的其他用户的影响.在一个虚拟社区平台中,用户间行为表现往往具有较大的相似性.社会实际生活中,社会成员之间相互推荐或者相互影响的现象也经常发生.例如:游戏玩家之间相互推荐好玩的游戏;食客之间相互推荐美食.更有甚者,把亲属或者朋友的意见作为个人决策的重要依据.因此,社会化推荐系统考虑影响用户行为偏好的社会属性因素,并将其应用到推荐中,不仅能够提高推荐的效果,而且能够使得推荐系统更符合人类生活的社会化特征.这也使得社会化推荐系统具有深远的现实意义和重要的社会价值.

1.1 社会化推荐系统的形式化定义

一般推荐系统的描述性定义[30]是由 Resnick 和 Varian 于 1997 年给出:利用电子商务网站向客户提供商品信息和建议,帮助用户决定应该购买哪些商品,模拟销售人员协助客户完成购买过程.社会化推荐系统是在一般推荐系统的基础上,把用户社会关系信息(比如亲属关系、朋友关系等)作为重要影响因子引入推荐生成过程中,从而提高推荐系统的性能.在此角度上讲,社会化推荐是一种把社会关系信息作为附加输入的推荐方法[31].伴随着社交网络的发展,社会化推荐的目标项目不仅包括普通项目(电影、图书等),还包括标签、朋友和社团等.可利用的数据源也不仅包括用户社会关系和用户-项目评分数据,还包括社会标签、用户的互动关系和用户点击行为等.在此应用背景下,社会化推荐是以社会媒体为推荐目标的推荐方法[32].其中,文献[33]把这两者称为社会化推荐系统的狭义定义和广义定义.

从学科渊源上讲,社会化推荐系统是在社会化网络分析理论的基础上,将用户社会属性信息加权融合到传统推荐系统中,在缓解传统推荐系统中数据稀疏性及冷启动问题的同时,还提高了推荐系统的性能.社会化网络分析是社会化网络理论中一个重要工具,对人与人之间、群体之间、组织之间、计算机之间或者其他信息/知识处理体之间的关系进行描述,并对其价值进行评估.社会化网络分析将个体之间的关系视为主要考虑因素,个体本身的属性处于从属地位.虚拟网络中,大量个体之间的各种互动关系而形成拓扑网络,称为在线社会化网络(online social network).有学者[34]将其视为传统社会化网络(social network)的一部分,或者视为同一个概念.

结合社会化推荐系统基本框架模型,根据一般推荐系统的形式化定义[1],社会化推荐系统的形式化定义为:

设 U 为所有用户的集合,I 为可以推荐给用户的所有项目的集合;设 $G=(g_{i,j})_{M\times M}$ 为所有用户社会关系矩阵,其中,$M=|U|$,$N=|I|$;映射 $\mu:U\times I\to R$ 对推荐结果的评价效用函数,其中,R 是一定范围内的全序非负实数集,称为推荐的效用值,$Y=|\{u_y|u_y\in U,x\neq y,g_{x,y}\neq 0,g_{x,y}\in G\}|$,表示与用户 u_x 存在社会关系的用户总数.社会化推荐系统要研究的问题是:在具有已定社会关系的用户群体中,根据所有项目在用户群体中的评价情况,主动地为用户推荐满足其偏好需求的、效用度最大的项目集 I^*,即:

$$\forall u_x\in U, I^*=\arg\max_{i\in I}\left(\mu_i+\alpha\mu(u_x,i)+\beta\frac{1}{Y}\sum_{y}^{Y}\mu(u_y,i)\right) \tag{1}$$

其中,$\forall\alpha,\beta\in[0,1]$且$\alpha+\beta=1$,$\mu_i$ 为偏倚变量,表示项目 $i\in I$ 在用户群体中的流行声誉度.这里的用户群体可以是全体用户(例如在因子分解机模型中,偏倚变量表示的就是项目在全体用户中流行声誉度,而且所有项目的偏倚变量都相等),也可以是部分用户(例如可以只考虑与用户 u_x 存在社会关系的用户群体).

© 中国科学院软件研究所 http://www.jos.org.cn

1.2 社会化推荐系统的基本框架模型

目前,研究人员在社会化推荐系统方面做了大量的工作,取得了一些研究成果,大部分利用社会化网络结构特征、项目在社会群体中的流行度及用户间的社会关系信息对推荐策略的影响,提出不同的社会化推荐框架模型[5,21−23,35].结合社会化推荐生成技术的特点及系统流程,我们给出了社会化推荐系统的层次化基本框架(如图1所示),依次包括如下4层:

(1) 数据采集层:获取用户基本社会属性信息、用户对项目的点评记录信息及用户社交活动日志等信息;

(2) 数据预处理:采集到相关数据之后,需要对这些数据进行预处理及清洗,其处理结果作为社会化推荐系统的结构化输入.主要包括:用户重要社会属性信息的筛选和清洗;用户-项目评分矩阵的建立;用户间信任关系的确定、社会化网络模型的构建及信任值的计算等;

(3) 社会化推荐层:社会化推荐系统的核心层,不仅考虑基于用户-项目评分矩阵的基本推荐技术的实现,还要重点考虑用户社会化网络关系对推荐结果的影响,其中,影响力的大小及这种影响如何作用在推荐生成技术中,是目前社会化推荐生成技术的重要研究内容.例如:在社会化矩阵分解中,社会化网络关系往往直接作用在矩阵分解的用户及项目潜在特征向量的提取过程中;但是在基于矩阵分解的因子学习模型中,用户的社会化网络关系对推荐结果的影响体现在模型参数上,不作用于用户-项目潜在特征向量的提取过程中;

(4) 社会化推荐结果的呈现和评价:把推荐结果呈现给用户时,需要根据用户显示或者隐式反馈,利用准确度、可用性、多样性等评价指标评价社会化推荐系统的性能,并据此进行适度的扩展和改进.

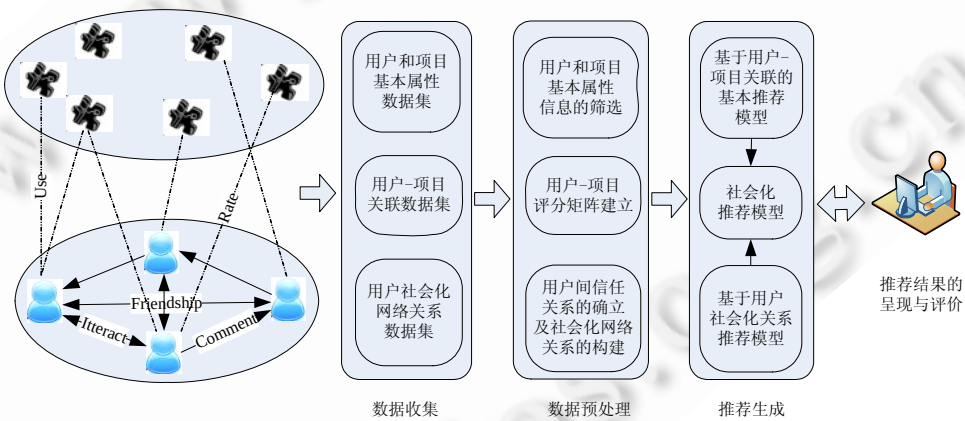


Fig.1 Basic framework of social recommender systems

图1 社会化推荐系统的基本框架

1.3 社会关系和网络模型的构建

从图1社会化推荐系统的基本框架可以看出:社会关系及网络模型的建立是紧随在数据收集模块之后,对用户社交活动数据进行挖掘,确立用户之间的社会关系,建立社会关系网络.这是社会化推荐系统有别于一般推荐系统的主要特征,也是社会化推荐系统的一个重要研究内容.它的主要任务是:考察哪些社会关系可能对用户行为和兴趣偏好有重要引导性影响,以及这种引导性影响是如何对用户的行为偏好产生影响;综合权衡这些影响因素的作用效果,改善最终的推荐结果,提高社会化推荐系统的性能.

目前,在社会化推荐系统中应用最多的社会关系网络是评分网络和信任网络.评分是在线项目(电影、图书等)推荐系统的基础,可以从用户对项目使用反馈评价中直接提取.信任关系是目前社会化推荐系统中应用最广的一种社会网络关系,用户信任关系模型的建立及信任度衡量算法的研究,是社会化网络分析的一个重要课题.在P2P(peer-to-peer network,对等网络)领域,对信任关系的研究已经有了许多优秀的模型和算法;但是社会化网

© 中国科学院软件研究所 http://www.jos.org.cn

络环境下,用户之间的信任关系的研究还处于起步阶段[36],其中:文献[37]从心理学、社会学、计算机科学等领域对信任关系的不同定义规则出发,论述了目前对信任关系的类型、特征及计算模型的研究成果;然后从社会化网络的特征角度,给出了社会化网络中个体用户的社会资本和社会信任的定义,并讨论了它们之间的差异性;最后,从社会信任关系的产生过程及信任值的计算方法,全面分析、总结了信任信息的收集模型、社会信任的评估模型和社会信任传播模型等方面的研究现状及成果.

目前,已有的信任度计算方法可以分为全局信任机制和局部信任机制[38]:局部信任机制主要考虑每一个人的个性化、主观性观点,并根据不同的用户预测不同的信任值;全局信任机制是整个信任网络中所有用户对某个用户的整体评价,是用户的全局声誉(reputation).例如,可以利用 PageRank 算法计算用户的声誉.此外,Ziegler[39]等人分别从网络视角(network perspective)、计算源点(computation locus)和链接评估(link evaluation)这 3 个方面把信任机制进一步细化(如图 2 所示),其中,全局信任机制不适合基于信任的个性化推荐系统.表 1 给出了适应社会化网络的典型信任推理算法的对比研究.

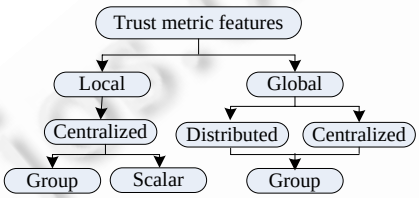


Fig.2 Trust metric classification

图 2 信任机制类型

Table 1 Comparison of classical trust reference algorithms

表 1 社会化网络中典型信任推理算法对比

算法	算法特点
Advogato[40]	(1) 用于发现在线社团中可信用户的数量,是一种基于网络最大流的群组信任机制.能够防止大规模外部攻击下的灾难性故障,接收恶意节点的数量是线性的,而且具有非常小的不变性; (2) 为了能够评价每一条边的容量(capacity),需要掌握整个网络图的结构,提高了算法的复杂度.此外,算法只给出可信用户的个数,不区分用户之间的信任值,而且这种信任关系独立于用户和项目,因此不适合基于信任的个性化推荐系统
TidalTrust[41]	(1) 基于修改的宽度优先搜索的局部信任度计算方法,用直接相连用户之间的信任度,预测计算间接相连用户之间的信任度:首先,计算源用户和目标用户之间的最短路径;然后,通过宽度优先搜索迭代计算源用户对目标用户的信任度; (2) 算法的复杂度为 $O(n)$,其中,$n=\max(V,E)$,V 和 E 是源用户到目标用户所经历的顶点数和边数
MoleTrust[42]	算法思想与 TidalTrust[37]相似,但是设置了一个搜索目标用户的最大路径长度.在计算源用户对目标用户信任值时,还参考了回溯搜索的目标用户的直接信任用户群的信任值
Appleseed[39]	(1) 基于心理学中激活扩散模型的局部群组信任传播机制,实现了信任值的标准化,输入可以是多参数化变量,输出为网络用户的可信任对象序列集.安全性上满足一般信任机制的抵御攻击特性:瓶颈特性(bottleneck property).适用于语义网、在线虚拟社区及公开的评分系统; (2) 该算法是一种内部迭代递归算法,时间复杂度为 $O(n^2)$
Sunny[43]	(1) 第 1 个包含置信度(confidence measures)计算的局部信任机制,基于贝叶斯网络概率模型的信任推理方法.利用概率抽样技术评估来自指定源用户的信任信息的置信度,在贝叶斯网络中执行抽样概率推理算法,计算信任值上、下界的概率值,并作为启示准确地预测贝叶斯网络节点之间的信任概率值; (2) 算法的难点在于:如何克服概率推理过程中,为了计算最终信任值对节点决策组合选择上的不确定性

这些信任推理算法的异同点:

- 首先,他们都是局部群组信任机制,都只能计算信任网络中存在路径相连的用户之间的信任值;
- Advagato 算法[40]和 Appleseed 算法[39]都是把用户分成入门、进阶和高阶等 3 种不同类型,每种类型间信任关系的强弱都不相同,前者采用网络流模型来计算用户的信任值,后者采用激活扩散机制来计算信任值;
- TidalTrust 算法[41]和 MoleTrust 算法[42]都是基于宽度优先搜索顺序迭代计算源用户与目标用户之间的

© 中国科学院软件研究所 http://www.jos.org.cn

信任值,二者的区别在于搜索源用户与目标用户之间路径设置上的差异;

- 不同于上述 4 种信任推理算法,Sunny 算法[39]是一种典型的基于贝叶斯网络的社会信任推荐算法,其准确率略好于其他信任推理算法.

2 社会化推荐生成技术

社会化推荐生成技术也可以分成基于内容的推荐、协同过滤推荐、基于知识的推荐和组合推荐,其中,绝大多数都是融合用户的社会化信息的协同过滤推荐方法[31,33,44,45].文献[33]从基于记忆和基于模型的协同过滤推荐方法的角度,重点讨论了社会化推荐系统中协同过滤推荐方法的研究进展,并把基于协同过滤的社会化推荐系统形式化地描述成如下形式:

一个社会化推荐的协同过滤模型(a social recommendation CF model) =

一个基本的协同过滤模型(a basic CF model) + 一个社会化信息模型(a social informantion model).

伴随着网络分析、矩阵分解等方法在社会化推荐系统中应用,最近几年,人们又提出多种比较典型的社会化推荐生成方法,主要包括基于网络图模型的推荐方法、矩阵分解方法、因子分解机模型、概率模型等.本节我们将对这些推荐生成算法进行分类总结和优缺点分析.

2.1 基于网络图模型的推荐方法

图是社会化网络关系最自然、直接的表示形式,其中,节点代表网络社会中的用户或项目,加权的边代表他们之间的连接关系.目前,把用户-项目评分网络图结构及用户之间社会关系网络图的基本特征(power-law,small word 和 scale-free)应用到推荐系统中,出现了多种社会化推荐方法[46−58],主要包括基于图结构的推荐方法[46,47]和链接预测方法[48−58].

(1) 基于图结构的推荐方法

基于图结构的推荐方法主要利用把用户-项目二分图,或者用户社会关系图的拓扑结构和基本特征信息,采用图拓扑模型分析、子图划分等网络分析方法及半监督学习方法,设计社会化推荐算法.此方法只适合于 0-1 二分类推荐系统,不能应用于评分预测的推荐系统.例如:文献[46]利用用户朋友间的社会化关系拓扑图,提出了一种三度分离的子图分割方法,并把由二度分离出来的用户视为目标用户的候选朋友;文献[47]以用户-项目对建立网络互动二分图,定义了用户-项目之间的一种联合互动关系图结构,提出一种基于图核(graph kernel)的分类推荐方法,此方法一般不需要关注图中边的权重值,只重点研究图的拓扑结构,显然,此方法对图的孤立点的推荐是无能为力的,即,存在冷启动问题.

(2) 链接预测方法

该方法把推荐问题转化为图中的边链接预测问题,这也是社会化推荐系统中常见的方法之一.首先建立用户-推荐目标网络图,常常采用复杂网络分析中的随机漫步和监督学习等算法[48,49],实现用户与目标项目的边链接预测.在评分预测推荐系统中,一般建立用户信任网络或者用户信任网络与用户-项目网络相融合的异构网络,把用户间的信任值作为用户信任网络边权重,把用户对项目的评分作为用户-项目网络边权重.借助于协同过滤或者基于用户(项目)的推荐方法的基本思想,采用随机漫步等方法,预测用户与推荐候选目标项之间的链接评分,其中,用户间的社会关系以信任值的形式作用在链接评分预测中,帮助解决用户-项目评分系统中数据稀疏性和冷启动问题.TrustWalker 模型[50]、基于随机漫步的混合协同过滤模型[51]及混合式随机漫步模型[52]是几个比较典型的实例.

在社交网络推荐系统中,用户间存在明显的社交网络关系图,这可以直接应用于边链接预测推荐系统中.大都借鉴 PageRank 算法的基本思想,把 *PageRank* 值作为社交网络关系图中边的权重,并依此预测社交网络图中任意两个节点之间的链接权重.例如:文献[53]证明了 Twitter 社交网络既是一种无标度网络,又是一种小世界网络,并把整个 Twitter 社交网络模拟成有向加权图,每个节点的 *Rank* 值为其所有入度连接点的 *Rank* 值与出度数比值之和,并采用带重启的随机漫步方法计算两个节点之间的连接度,从而实现对 Twitter 社交网络中的好友推荐.文献[54]利用 Web 应用环境下用户对各种资源的标签化命名关系构建一个异构的社会化网络结构,把信息

© 中国科学院软件研究所 http://www.jos.org.cn

推荐问题转换成排序问题,当用户搜索或者浏览信息时,为其提供信息推荐服务,提出一种基于异构社会化网络的随机漫步排序模型和异构社会化网络不同类型的边权值成对学习算法.文献[55]在社会化网络中应用带重启的随机漫步算法,把社会化网络中边连接预测问题转换成以两个节点用户的基本属性和互动属性构成的边特征向量为输入的边权重函数的监督学习问题;此外,在社会化网络边链接预测中,还可以利用监督学习算法进行预测[56−58].

Wang 等人[59]通过对两个基于位置的社会化网络 Brightkite 和 Gowalla 数据集中用户签到位置范围和存在朋友关系的两个用户的签到位置的分布情况进行分析,发现:虽然用户的签到位置大部分都是初次访问的,但是在每个用户已签到位置的 10 公里范围内,Brightkite 用户初次签到位置所占比例约为 67.57%,在 Gowalla 数据集中,这个比例达到了 81.93%;同时,在这两个数据集中,用户的首次签到位置点是其朋友或朋友的朋友已签到的比例分别为 23%和 31%.鉴于此,构建用户间的朋友社交网络和用户-位置关系网络,借用个性化的 PageRank 算法的基本思想,把朋友关系信息对位置签到的影响引入到位置推荐过程中,提出了两种位置推荐算法:基于朋友关系的标签涂色算法(friendship-based bookmark-coloring algorithm,简称 FBCA)和基于位置-朋友关系的标签涂色算法(location-friendship bookmark-coloring algorithm,简称 LFBCA).

无论是评分预测推荐系统还是社交网络推荐系统中,链接预测算法都是基于网络图模型推荐的主要方法,其中,网络结构的构建和边权重的赋值问题是此方法研究的两个核心问题.在评分预测推荐系统中,把用户的社会关系信息融合到用户-项目评分网络中,能够有效帮助缓解仅仅依靠用户-项目评分关系的推荐系统中的数据稀疏性和用户冷启动问题,但还是会存在项目冷启动问题.

Table 2 Comparison of the algorithms

表 2 算法对比

推荐方法	优点	缺点
基于图结构的推荐方法	不需要专业知识,支持已有分类算法和复杂网络分析算法	存在稀疏性和冷启动问题,需要大规模的训练数据,稳定性差,推荐面窄
链接预测方法	支持多种分类算法、复杂网络分析算法和监督学习算法,可以融合多种推荐算法,能够缓解数据稀疏性及用户冷启动问题	存在项目冷启动问题,需要大规模的训练数据,推荐效果受限于历史数据

2.2 矩阵分解方法

近些年来,矩阵分解方法(matrix factorization,简称 MF)已经广泛应用于推荐系统中,它能够把用户-项目评分矩阵分解成两个或者多个低维矩阵的乘积实现维数的规约,用低维空间数据研究高维数据的性质,主要包括奇异值分解(singular value decomposition,简称 SVD)、非负矩阵分解(nonnegative matrix factorization,简称 NMF)和概率矩阵分解(probabilistic matrix factorization,简称 PMF).其中,非负矩阵分解是由 Lee 和 Seung[60]于 1999 年提出,在许多领域有成功应用的新的矩阵分解方法.在此方法中,可以把用户对项目的评分矩阵 $R_{n\times m}$ 分解成两个实值非负矩阵 $U_{n\times k}$ 和 $V_{k\times m}$,使得 $R\approx U^TV$:

$$U=\begin{bmatrix} u_{1,1} & \cdots & u_{1,n} \\ \vdots & u_{k,x} & \vdots \\ u_{k,1} & \cdots & u_{k,n} \end{bmatrix}, V=\begin{bmatrix} v_{1,1} & \cdots & v_{1,m} \\ \vdots & v_{k,i} & \vdots \\ v_{k,1} & \cdots & v_{k,m} \end{bmatrix},$$

其中,$k<\min(m,n)$.矩阵 $U_{n\times k}$ 的第 x 列向量 U_x 和 $V_{k\times m}$ 的第 i 列向量 V_i 的乘积 $U_x^TV_i$ 代表用户 x 对项目 i 的预测评分,称 U_x 为用户 x 的潜在特征向量(latent feature vector),V_i 为项目 i 的潜在特征向量.为了得到更加准确的预测值,需要对预测值与观测值之间的损失度进行优化评估,建立如下的目标优化函数:

$$\sum_{x,i}\frac{1}{2}(r_{x,i}-U_x^TV_i)^2 \tag{2}$$

利用交替的梯度下降法,可以求出上述目标函数达到最小值时的最优解 $U_{n\times k}$ 和 $V_{k\times m}$.概率矩阵分解侧重于从条件概率最优的角度,探讨矩阵分解的概率优化分解.

(1) 社会化矩阵分解

© 中国科学院软件研究所 http://www.jos.org.cn

社会化推荐系统中,矩阵分解方法关注用户的社会化网络信息对用户潜在特征向量的影响,把用户各种社会化网络关系信息融合到矩阵的优化分解过程中,以提取更优的潜在特征向量,称为社会化矩阵分解[5,61].文献[62]根据社会化网络信息在对潜在特征向量优化求解时发挥的不同作用,将社会化矩阵分解方法分成两类:

- 第 1 类方法称为社会正则化(social regularization),其中,两个用户 x 和 y 之间社会关系信息 $S_{x,y}$ 对用户潜在特征向量的约束求解影响方式有两种:
  - 第 1 种称为社会正则化[63,64],具体实现方法如下:

$$\sum_x \sum_{y\in friends_x} \frac{1}{2}(S_{x,y} - \langle U_x, U_y\rangle)^2 \tag{3}$$

  其中,$\langle\cdot,\cdot\rangle$表示向量之间的内积;
  - 第 2 种称为社会谱的正则化(social spectral regularization)[6,13,65]:

$$\sum_x \sum_{y\in friends_x} \frac{1}{2} S^+_{x,y} \| U_x - U_y \|_2^2 \tag{4}$$

  此外,在 SoRec 系统[66]中,提出一种社会谱正则化的变形方法:

$$\sum_x \sum_{y\in friends_x} \frac{1}{2}(\bar{S}_{x,y} - \sigma(\langle U_x, Z_y\rangle))^2 \tag{5}$$

  其中,$\sigma = 1/(1+\mathrm{e}^{-z})$,$\bar{S}_{x,y} = \sigma(S_{x,y})$.$Z_{N\times N}$ 为用户之间关系矩阵,$U_x^T Z_y$ 能够预测用户之间关系的强弱.
- 第 2 类方法称为加权朋友平均(weighted friend average)方法,典型的例子是文献[12]提出的社会信任模型(social trust ensemble,简称 STE).将用户 x 对项目 i 的预测评分与用户 x 的所有朋友对项目 i 的评测评分加权平均在一起,目标优化函数形式如下:

$$\sum_x \frac{1}{2}(r_{x,i} - \sigma(\alpha U_x^T V_i + (1-\alpha)\sum_{y\in friends_x} T_{x,y} U_y^T V_i))^2 \tag{6}$$

文献[62]结合实现情况,针对上述矩阵分解技术中存在的缺陷提出了 3 种新的社会化矩阵优化分解方法:

第 1 种方法:特征社会正则化(feature social regularization).

添加用户的性别特征对社会正则化方法进行改进,实现了性别特征对用户潜在特征向量提取的影响.改进后的目标优化函数分别为

$$\sum_x \sum_{y\in friends_x} \frac{1}{2}(S_{x,y} - x^T U_x^T U_y)^2 \tag{7}$$

$$\sum_x \sum_{y\in friends_x} \frac{1}{2} S^+_{x,y}(x-y)^T U_x^T U_y (x-y) \tag{8}$$

第 2 种方法:混合信息融合的社会化矩阵优化分解方法.

利用线性支持向量机提取包含用户 x 和项目 i 社会网络特征的信息融合特征向量 $f_{x,i}\in R^F$,并学习得到权重向量 $w\in R^F$,将线性回归 $w^T f_{x,i}$ 视为信息融合的观测结果,并把这个结果与 Matchbox 的预测结果叠加在一起,得到最终的预测结果,其中,目标优化函数形式如下:

$$\sum_x \frac{1}{2}(r_{x,i} - [\sigma]w^T f_{x,i} - [\sigma]x^T U_x^T V_i)^2 \tag{9}$$

第 3 种方法:共同偏好正则化(co-preference regularization).

提取用户之间的共同选择性偏好,实现了用户 x 和用户 z 对同一个项目具有相同或者相反偏好的预测,其目标优化函数形式如下:

$$\sum_{x,y,i} \frac{1}{2}(P_{x,y,i} - x^T U_x^T diag(V_i) U_y)^2 \tag{10}$$

(2) 基于矩阵分解的因子学习模型

不同于社会化矩阵分解方法,基于矩阵分解的因子学习模型是以矩阵分解得到的用户潜在特征向量与项

© 中国科学院软件研究所 http://www.jos.org.cn

目潜在特征向量的内积$U_x^T V_i$代表用户x对项目i的预测评分,把用户和项目的社会化信息作为因子参数,以线性组合或数乘形式与用户和项目的潜在特征向量相融合,体现这些信息对项目评分预测的影响,形成对项目评分预测的优化学习模型.大量研究[67−71]表明,此类模型对解决数据稀疏性有显著的效果.此类模型的最简单形式如下[68]:

$$\hat{r}_{x,i} = \mu + b_x + b_i + U_x^T V_i \tag{11}$$

其中,μ代表全局偏倚参数,b_x代表用户x的偏倚参数,b_i代表项目i的偏倚参数.优化学习目标函数为

$$\sum_{(x.i)} (r_{x,i} - \hat{r}_{x,i})^2 + \lambda(\| U_x \|^2 + \| V_i \|^2 + b_x^2 + b_i^2) \tag{12}$$

然后,以上述模型为基础,学者们提出了多种扩展模型[69−72].例如,Chen 等人[69,70]提出的基于特征的矩阵分解模型:

$$\hat{r}_{x,i} = \left(\sum_{c \in C(x)} \alpha_c^{(x)} U_c \right)^T \left(\sum_{c \in C(i)} \beta_c^{(i)} V_c \right) + \sum_{c \in C(x,i)} \gamma_c^{(u,i)} g_c \tag{13}$$

其中,α,β,γ分别表示用户、项目及二元特征系数,g 为二元偏倚项,$C(x)$表示用户 x 所属的类别,$\alpha_c^{(x)}$ 表示用户所有类别的权重系数.针对不同的实际应用问题,提出了多种改进模型.例如:

文献[71]针对 Twitter 数据的特点,重点考虑 tweet 主题水平因素(tweet topic level factors)T_i和用户社会关系因素$d_{U(i)}$及其他显著特征,提出了一种基于特征的矩阵分解模型,实现对用户的 tweet 个性化协同推荐.模型具体如下:

$$\hat{r}_{x,i} = \sum_j b_j \gamma_j + U_x^T \left(\frac{1}{Z} \sum_{\omega_j \in T_i} V_{\omega_j} + \alpha d_{U(i)} \right) \tag{14}$$

文献[72]建立递增树林模型,获取用户的活动及其连续模式,将用户社会特征、参与的互动、社会网络关系及项目的分类信息融合在一起,构造了一个因子分解模型,实现对用户的社交网络项目(包括人、机构、群等)推荐:

$$\hat{r}_{x,i} = \left(\sum_{c \in C(x)} \alpha_c^{(x)} U_c \right)^T \left(\sum_{c \in C(i)} \beta_c^{(i)} V_c \right) + \sum_{c \in C(x,i)} \gamma_c^{(u,i)} g_c + \sum_{s=1}^{S} f_{s,root(s,i)}(x_{x,i}) \tag{15}$$

此外,基于矩阵分解的因子模型,还有 Ma 等人[73]提出的带网络图的非对称因子模型:

$$\hat{r}_{x,i} = V_i^T \cdot (|N(x)|^{-1/2} \sum_{j \in N(x)} V_j) + b_x + b_i + \alpha_x \cdot c_{xi} + b_{xt}(t_{xi}) \cdot q_t(t_{xi}) \tag{16}$$

其中,$N(x)$表示用户 x 在社交网络中关注的项目集合;$|N(x)|^{-1/2} \sum_{j \in N(x)} V_j$ 表示用户虚拟特征;$\alpha_x c_{xi}$ 表示点击率偏倚参数,此点击率被定义为给一组特定用户推荐不同项目所被接受的比例;$b_{xt}(t_{xi}) \cdot q_t(t_{xi})$表示不同时间段对推荐结果影响因子.

基于矩阵分解因子学习模型,还有许多学者结合社会化网络中其他信息,提出多种满足特殊应用要求的社会化推荐方法.例如,在社会化网络中利用最近邻信任信息的 Top-N 推荐[74]、利用特定类型用户社会信任环的项目推荐[75]、融合用户社会活动时间和空间限制的连续位置推荐[76]等.

Table 3 Comparison of the algorithms

表 3 算法对比

推荐方法	优点	缺点
社会化矩阵分解	不需要专业知识,能够缓解数据稀疏性和用户冷启动问题,以用户为中心自动快速生成推荐结果	项目冷启动问题,推荐结果的解释性差
基于矩阵分解的因子学习模型	融合用户和项目的多种社会学属性,能够缓解数据稀疏性和冷启动问题	需要学习的参数多,推荐效果受限于历史数据

作为协同过滤推荐系统中一种新型推荐生成算法,矩阵分解技术对高维数据降维及解决数据稀疏性等方面有非常好的效果.把用户的社会关系信息加权应用到矩阵优化分解过程中,实现用户的社会关系信息对用户

© 中国科学院软件研究所 http://www.jos.org.cn

推荐结果的影响,在一定程度上缓解了冷启动用户问题.但是在实际网络环境中,用户对项目的评分信息与用户之间的社会关系网络信息往往是不属于同一个数据源,同时,用户之间社会关系强弱划分也没有统一的标准,这是此方法推广应用过程中需要克服的困难.

2.3 因子分解机模型(factorization machine)

类似于基于矩阵分解的因子学习模型,因子分解机模型(factorization machines)[77]也是一种面向任务的模型学习方法,是由 Rendle 于 2010 年提出的一种新型分类预测方法.在数据稀疏条件下,此模型具有良好的预测效果,其时间复杂是线性的.模型(二阶 d=2)具体如下:

$$\hat{r}(\bar{x})=\omega_0+\sum_{i=1}^{n}\bar{\omega}_i x_i+\sum_{i=1}^{n}\sum_{j=i+1}^{n}\langle\bar{v}_i,\bar{v}_j\rangle x_i x_j \tag{17}$$

其中,$\langle\cdot,\cdot\rangle$表示两个 k 阶向量之间的内积 $\omega_{i,j}:\langle\bar{v}_i,\bar{v}_j\rangle=\sum_{f}^{k}v_{i,f}\cdot v_{j,f}$;$\bar{v}_i$ 表示矩阵 V 中包含 k 个元素的第 i 行;$k\in N^*$ 是提前设置的参数,表示分解因子的维数.需要学习的模型参数Θ包括:

$$\omega_0\in R,\bar{\omega}\in R^n,V\in R^{n\times k},$$

其中,ω_0表示全局偏倚,ω_i模拟第 i 个变量的权重,$\hat{\omega}_{i,j}$ 模拟第 i 个变量与第 j 个变量之间的互动作用.此模型简化计算方法及正则化最小二乘优化目标函数分别为

$$\hat{r}(\bar{x})=\omega_0+\sum_{i=1}^{n}\bar{\omega}_i x_i+\frac{1}{2}\left(\left(\sum_{i=1}^{n}v_{i,f}x_i\right)^2-\sum_{i=1}^{n}v_{i,f}^2x_i^2\right) \tag{18}$$

$$\sum_{(\bar{x},y)}(\hat{r}(\bar{x})-r)^2+\sum_{\theta\in\Theta}\lambda_\theta\theta^2 \tag{19}$$

利用此模型的关键在于模型输入特征向量的构造,例如:文献[78]将用户的情绪信息、陪同观看电影的用户信息作为类型变量与用户信息、指示变量(电影信息)编码在一起,构成模型输入特征向量,实现对用户情绪和陪同人员感知的电影评分预测;文献[79]将用户年龄、性别、微博数目、在社交网络中关注的用户数量及一些标签和关键词作为类型变量融合在输入特征向量中,结合贝叶斯推理及马尔科夫链的蒙特卡洛方法,实现了基于时间的社交网络项目(包括人、机构、群等)推荐;文献[80]利用因子分解机模型,模拟用户兴趣潜在特征并实现对 Twitter 用户的个体决策分析与预测.

因子分解机模型优点在于:能够把多种上下文信息与用户-项目关系矩阵融合在一起,形成模型的多维输入特征向量,使上下文信息对用户潜在特征向量和项目潜在特征向量产生从始至终的影响,在基于矩阵分解的因子模型中,这种影响力仅仅发生在模型参数的确定过程中,对用户和项目潜在特征向量的提取没有任何影响.缺点在于:模型需要学习的参数多,复杂性高,推荐结果的可解释性差等.

2.4 概率模型

概率模型是社会化推荐系统中应用比较广泛的一种模型学习方法,是以朴素贝叶斯定理为理论基础,把用户-项目的联合概率以及用户和项目概率作为模型学习参数,在此条件下,推断已知用户条件下推荐项目的条件概率.应用于推荐系统中的简单概率模型[81]形式如下:

已知系统中用户集合 $U=\{u_1,u_2,u_3,\ldots,u_N\}$和项目集合 $I=\{i_1,i_2,i_3,\ldots,i_M\}$,能够获取用户潜在兴趣和项目特征,用一个潜在主题集合 $Z=\{z_1,z_2,z_3,\ldots,z_K\}$表示.用户 $u\in U$ 使用的每一个项目 $i\in I$ 都关联着一个潜在主题变量 $z\in Z$,并且用户依据其兴趣偏好分布选择的主题 $z\in Z$,能够按照与其关联的项目的概率分布生成项目 $i\in I$.其中,假设用户与已知所属主题的项目是相互独立的,则用户 u、项目 i 和主题 z 的联合概率为

$$\Pr(u,z,i)=\Pr(u)\Pr(z|u)\Pr(i|z)=\Pr(z)\Pr(u|z)\Pr(i|z) \tag{20}$$

那么,用户 u 和项目 i 之间的联合概率为

$$\Pr(u,i)=\sum_{z\in Z}\Pr(u,z,i)=\sum_{z\in Z}\Pr(z)\Pr(u\mid z)\Pr(i\mid z) \tag{21}$$

其中,需要学习的模型参数Θ包括 $\Pr(z)$,$\Pr(u|z)$和 $\Pr(i|z)$,学习的优化目标函数如下:

© 中国科学院软件研究所 http://www.jos.org.cn

$$\max \sum_{\forall(u,i),\theta\in\Theta} \log(\Pr(u,i\mid\theta)) \tag{22}$$

那么,用户 u 选择项目 i 的概率为

$$\Pr(i\mid u)=\frac{\Pr(u,i)}{\Pr(u)}\propto \Pr(u,i) \tag{23}$$

在此基础上,Ye 等人[82]将朋友关系信息对用户活动的影响引入到上述概率模型中,此时,用户 u 对项目 i 选择的概率正比于用户 u、朋友 $f\in f(u)$($f(u)$表示用户 u 的朋友集合)、主题 f 和项目 i 之间的联合概率分布:

$$\Pr(i|u)\propto \Pr(u,f,z,i)=\Pr(z)\Pr(u|f)\Pr(f|z)\Pr(i|z) \tag{24}$$

此时,需要学习的模型参数为 $\Theta=\{\Pr(z),\Pr(u|f),\Pr(f|z),\Pr(i|z)\}$,用户和项目之间的联合概率为

$$\Pr(u,i)=\sum_{z\in Z}\sum_{f\in F(u)}\Pr(z)\Pr(u\mid f)\Pr(f\mid z)\Pr(i\mid z) \tag{25}$$

文献[71]还将用户的社会影响、协同过滤与基于内容项目推荐方法相结合,提出一种统一概率生成模型(W 表示项目内容空间):

$$\Pr(i|u)\propto \Pr(u,f,z,i,w)=\Pr(z)\Pr(u|f)\Pr(f|z)\Pr(i|z)\Pr(w|z) \tag{26}$$

文献[83]在社会网络图中考虑用户偏好、项目普遍接受率及用户直接朋友关系等影响因子,建立用户对项目评分预测概率模型,通过计算用户对项目评分期望的形式,预测用户对项目的评分:

$$\hat{r}_{ui}=\sum_k k\cdot \Pr(r_{u,i}=k\mid A=a_h,A'=a'_i,\{r_{vi}=k':\forall V\in U(i)\cap N(u)\}) \tag{27}$$

在此模型中,用 $G=\{U,E\}$表示所有用户的社会网络图,图中所有节点代表用户集 U,所有边代表用户之间的直接朋友关系,用 $A_u=\{a_h|h=1,2,\ldots,H\}$表示每一个用户的属性集,$N(u)=\{v|(u,v\in E)\}$表示用户直接朋友集,$A'_l=\{a'_l\mid l=1,2,\ldots,L\}$表示项目集 I 每一个项目 i 的属性集,$I(u)$表示用户 u 已经评分的项目集合,$U(i)$表示已对项目 i 评分的用户集.条件概率的具体形式如下(Z 是一个标准化常量):

$$\begin{aligned}&\Pr(r_{u,i}=k\mid A'=a'_l,A=a_h,\{r_{vi}=k':\forall V\in U(i)\cap N(u)\})=\\&\frac{1}{Z}\Pr(r_{ui}=k\mid A=a_h)\cdot\Pr(r_{ui}=k\mid A'=a'_l)\cdot\Pr(r_{ui}=k\mid r_{vi}=k':\forall v\in U(i)\cap N(u))\end{aligned} \tag{28}$$

此外,在基于位置服务的社会化网络推荐系统中,概率模型也得到了广泛的应用.例如:Yin 等人[27,28]提出两种潜在类统计混合模型——时间上下文感知混合模型和位置内容感知概率混合生成模型,分别用于模拟时间、空间两种上下文因素对用户行为的影响,然后提出两种 Top-N 社会化网络推荐方法.文献[84]分析同一用户签到的不同兴趣点(point of interest,简称 POI)在地理空间中分布情况,采用幂律分布模拟用户在一定距离空间内的两个兴趣点的签到概率,把社会朋友关系信息对用户签到位置的影响加入到推荐算法的设计过程中,提出一种基于朴素贝叶斯网络位置推荐方法.

还有一些面向社交网络不同应用类型的概率推荐方法,例如:基于贝叶斯推理的朋友圈信息咨询推荐[85]、基于极大似然估计的用户位置预测[86]和 Digg 博文信息推荐[87];基于狄利克雷分布(latent Dirichlet allocation)的 followee 和 tweet 推荐[88].总之,概率模型充分发挥了贝叶斯理论、极大似然估计等在知识推理方面的优势,使得推荐系统具备较好理论依据,具有较高推荐准确率.此方法的缺点在于学习参数多、时间长.

作为解决社会化推荐系统中数据稀疏性及冷启动问题的有效方法之一,矩阵分解方法和概率模型是目前推荐系统中研究热点,其优点在于:能够把用户、项目的特征信息及用户之间社交网络关系信息应用到推荐预测的过程中,实现多种信息对推荐结果的综合影响.缺点在于:模型学习时间较长,复杂性较高,这也是该方法在实际应用过程中需要克服的主要问题之一.在社会化矩阵分解中,用户的社会属性信息对推荐预测结果的影响主要体现在用户/项目的潜在特征向量的分解过程中;基于矩阵分解的因子学习模型中,虽然大部分也是基于矩阵分解的方法,但是用户社会属性信息或者项目属性信息对推荐预测结果的影响主要表现在融合参数的学习与设置上,这也是在此将它们视为两种不同类型的一个重要原因.表 4 中列出了各种社会化网络推荐生成技术涉及的核心理论和算法、适用的推荐类型.

© 中国科学院软件研究所 http://www.jos.org.cn

Table 4 Applications of social recommendation generating approaches

表 4 社会化推荐生成方法应用

推荐方法	核心理论、算法	适用的推荐类型	代表文献
基于网络图模型的推荐方法	子图划分、图的遍历、社团发现算法等复杂网络分析方法、分类算法、监督学习算法、随机漫步算法、信任推理算法	朋友、社团、微博推荐等社交网络推荐和电影、音乐、图书等推荐	文献[20,23,24,46,47,54–58]
矩阵分解方法	矩阵分解、最小二乘、信任推荐算法	基于项目评分预测的推荐应用(如电影、音乐、电子商务等)	文献[6,12,13,9,62–76]
因子分解机模型	因子分解、最小二乘、贝叶斯推理等	基于项目评分预测的推荐应用(如电影、音乐、电子商务等)及社交网络推荐应用	文献[77–80]
概率模型	贝叶斯理论、最小二乘、极大似然估计	基于项目评分预测的推荐应用、社交网络推荐应用(个体、结构、社团推荐等)	文献[81–88]

3 社会化推荐系统研究的热点和难点

社交网络的不断发展与应用,为推荐系统的研究和应用提供了新的发展方向与前进动力.普适计算、人工智能、信息检索及人类社会学等领域知识的引进,推动了社会化推荐系统研究与发展,并取得了一定的研究成果.但是作为一个新兴研究领域,仍然存在许多问题和挑战亟待解决和研究.这主要包括以下几个方面.

(1) 社会属性信息的确定和量化

社会属性信息主要是指在社会网络推荐系统中涉及对用户行为偏好有影响的个体属性信息(例如性别、民族及位置信息等)和社会关系信息,其中,个体属性信息一般是比较明确的.在现实生活中,人与人之间的社会关系可以分为很多种,例如亲属关系、朋友关系、同事关系、同学关系等.但是在虚拟的网络环境中,用户之间各种社会关系常常是模糊的,而且具有统一的表现形式:网络通信互动关系,一些社会关系能够从网络互动内容的语义上确定.但是到目前为止,在社会化推荐系统中,用户之间社会关系及不同社会关系之间对用户偏好影响的量化还没有统一的标准.在基于矩阵分解的因子学习模型[72,73]中,这种影响往往体现在对全局参数的设定上或者模型的优化学习过程中,在基于网络图模型的推荐方法[49]中,大都将用户之间的各种社会关系转化成信任关系,并基于用户之间的网络互动频率及次数,赋予不同的信任值,实现用户社会关系的量化.

(2) 社会化网络模型的建立

社会化网络模型的建立是社会化推荐系统的前提条件,社会化推荐系统中,一般涉及两种网络拓扑图模型:

- 一种是所有用户之间的社会关系网络图,图中的节点代表用户,节点之间的连线表示用户间社会关系,同时制定评价标准及推理规则,衡量用户之间社会关系的强弱.信任网络是目前该类模型中应用最普遍的方法;
- 另一类是用户-项目评分网络图模型.不同的推荐技术,其构建的社会化网络模型也不一样.目前已有的研究成果中,构建的网络拓扑图基本上都是静态的、单一化的,很少涉及网络动态变化问题和多层次网络.文献[89]分别对用户的通讯录、使用标签、所在社区、兴趣模型和观点倾向进行分析,建立用户多维社会化网络模型,为多媒体共享用户推荐新的网络连接.

此外,社会化网络图模型的大小及计算复杂性问题,同样也是一个值得考虑和研究的问题,特别是应用在存储和计算能力都有限的移动推荐服务系统中.

(3) 社会化推荐系统中的数据稀疏性和冷启动问题

传统的推荐系统中,用户-项目评分矩阵中可能存在数据稀疏性问题,这使得很难为较少项目评分的用户找到兴趣相似用户.此时引入用户社会关系信息,可以把与此用户存在社会关系的用户作为兴趣相似用户,从而解决由于数据稀疏性导致的推荐算法失效问题.

此外,对于新加入用户或者项目,系统中没有或者很少有关他们的记录信息,推荐算法很难对他们进行处理.用户社会网络关系的引入,能够帮助建立新加入用户特征模型,在一定程度上缓解了推荐系统中新用户的冷

© 中国科学院软件研究所 http://www.jos.org.cn

启动问题[3],这也使得社会化推荐系统在社会关系推荐方面有比较明显的优势.但是对于新加入项目,在传统推荐系统中,只有该项目被用户使用后,依靠协同过滤推荐算法向其他用户推送;在社会化推荐系统中,除了可以采取上述方法外,社会网络中用户之间还能够对新加入项目进行自动推荐与传播.

(4) 社会化推荐系统的效用评价

推荐系统的大部分性能评价指标源于信息检索领域,推荐的准确度和效率是推荐系统的一般性能指标.在评分系统中,主要采用平均绝对误差(mean absolute error,简称 MAE)、均方根误差(root mean squared error,简称 RMSE)和关联度(correlation)等,这些评价指标同样也适用社会化推荐系统.在 KDD Cup 2012 Track1[79]中,采用了 MAP@3 的评价指标.目前,不同的研究工作针对不同的实际问题,使用不同的数据集,在推荐生成技术上存在较大的差异,一方面不能形成统一的效用评价标准;另一方面,对用户社会网络关系在推荐系统的中作用效果的评价研究还比较少,大多数的评价机制还是从整体上对推荐系统进行性能评估.

(5) 社会化推荐系统的隐私保护和安全性问题

隐私保护和安全性问题是所有推荐系统都不可避免的问题,2010 年,Netflix prize 比赛由于涉及用户隐私问题而停止举办[29].社会化推荐系统中,隐私保护和安全性问题也尤为重要,因为社会化推荐系统中不仅涉及用户本身的属性信息,还涉及用户的社会关系信息,这些信息可能揭示用户的个人隐私.目前,对这方面的研究还比较少.文献[8]从保护用户间的连接关系入手,提出了一种具有隐私保护机制的社会化推荐系统;文献[90]把推荐准确率和保护用户隐私进行折衷处理,在不揭示用户间敏感连接的前提下,提出两种个性化网络推荐方法.

(6) 社会化推荐系统大规模数据处理问题

社会化网络中,特别是在移动社会化网络中,移动终端具有较高的移动性和灵活性,这就要求系统能够实时地处理(采集、存储等)由用户随时随地频繁交互而产出的大量数据.在此以后,推荐系统才能从这些数据中挖掘用户兴趣偏好模型及其社会网络关系,并选择合理的推荐生成算法,为用户快速的产生推荐结果.这务必给整个数据处理系统带来额外的负担,而且整个推荐过程对数据处理的复杂度远远高于对这些数据的简单处理(采集、存储).在实际的工业应用中,这种代价是相当昂贵的.目前,对这方面的研究还非常少,而且已有的研究成果中涉及的数据集大多数都是小规模的或者模拟的数据集.因此,面对实际应用需求,利用社会网络信息,设计能够处理对实时性要求较高的快速推荐算法,将逐渐成为社会化推荐系统领域新的挑战.

(7) 社会化推荐系统的可扩展性与主动性

可扩展性一直是困扰推荐系统应用过程中一个难点问题.在基于用户的协同过滤推荐系统中,用户相似度计算的时间复杂度为 $O(n^2\cdot m)$,其中,n 为用户数,m 为用户所关注的平均项目数.需要搜索所有项目,为用户寻找相似用户.随着用户和项目数量的增加,相似度的计算量也会增大[91].目前,主要从聚类方法、数据集的缩减、维数降解及建立线性模型等方面解决协同过滤可扩展性的问题.在社会化网络中,用户之间随时随地的信息互动会产生大规模的数据集,将这些数据信息引入到推荐系统中,势必进一步增加推荐算法的计算复杂度,并使问题变的更加严重.目前,对于此问题的研究还比较少.特别是在移动社交网络中,用户对推荐结果的要求不再仅仅局限于内容的准确性,而且还要在时间、地点等方面具有普遍的适应性.这就要求推荐算法不仅要照顾用户的长期兴趣,还要照顾用户随时间、位置变化而产生的短期兴趣.

4 总 结

在一个信息资源爆炸式增长的时代,推荐技术是作为解决信息过载问题的主要方法之一.在过去的数十年中,无论在学术研究领域还是在工业应用领域,都取得了长足的进步与发展.在社会化网络日益盛行的当下,把用户的社会属性信息引入到推荐系统中,不仅满足社会化网络发展对推荐技术的要求,而且符合人类社会学有关推荐决策的基本原理,同时也促进了推荐系统的快速发展.本文对社会化推荐系统的基本原理、关键技术、研究进展和发展趋势进行总结与分析,希望能够有助于促进该领域的发展和进步.

© 中国科学院软件研究所 http://www.jos.org.cn

References:

[1] Xu HL, Wu XD, Li X, Yan BP. Comparison study of Internet recommendation system. Ruan Jian Xue Bao/Journal of Software, 2009,20(2):350−362 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/3388.htm [doi: 10.3724/SP.J.1001.2009.00350]

[2] Wang Z, Sun LF, Zhu WW, Yang SQ, Li HZ, Wu DP. Joint social and content recommendation for user-generated videos in online social network. IEEE Trans. on Multimedia, 2013,15(3):698−710. [doi: 10.1109/TMM.2012.2237022]

[3] Quijano-Sanchez L, Recio-Garcia J, Diaz-Agudo B. Social factors in group recommender systems. ACM Trans. on Intelligent Systems and Technology, 2013,4(1):Article No.8. [doi: 10.1145/2414425/2414433]

[4] Social Network Analysis. A brief introduction. 2007. http://orgnet.Com/sna.html

[5] Jamali M, Ester M. A transitivity aware matrix factorization model for recommendation in social networks. In: Proc. of the IJCAI. AAAI Press, 2011. 2644−2649. [doi: 10.5591/978-1-57735-516-8/IJCAI11-440]

[6] Jamali M, Ester M. A matrix factorization technique with trust propagation for recommendation in social networks. In: Proc. of the RecSys 2010. New York: ACM Press, 2010. 135−142. [doi: 10.1145/1864708.1864736]

[7] Moghaddam S, Jamali M, Ester M. ETF: Extended tensor factorization model for personalizing prediction of review helpfulness. In: Proc. of the ACM WSDM 2012. New York: ACM Press, 2012. 163−172. [doi: 10.1145/2124295.2124316]

[8] Hoens TR, Blanton M. A private and reliable recommendation system for social networks. In: Proc. of the IEEE SocialCom. Washington: IEEE Computer Society, 2011. 816−825. [doi: 10.1109/SocialCom.2010.124]

[9] Berjani B, Strufe T. A recommendation system for spots in location-based online social networks. In: Proc. of the SNS 2011. New York: ACM Press, 2011. 1−4. [doi: 10.1145/1989656.1989660]

[10] Kim HK, Ryu YU. Cho Y. Customer-Driven content recommendation over a network of customers. IEEE Trans. on Systems, Man and Cybernetics, Part A, 2012,42(1):48−56. [doi: 10.1109/TSMCA.2011.2147306]

[11] Purushotham S, Liu Y, Kuo CC. Collaborative topic regression with matrix factorization for recommendation systems. In: Proc. of the ICML. 2012. 759−766.

[12] Ma H, King I, Lyu MR. Learning to recommend with social trust ensemble. In: Proc. of the SIGIR 2009. New York: ACM Press, 2009. 203−210. [doi: 10.1145/1571941.1571978]

[13] Ma H, Zhou D, Li C. Recommender systems with social regularization. In: Proc. of the WSDM 2011. New York: ACM Press, 2011. 287−296. [doi: 10.1145/1935826.1935877]

[14] Huang WH, Meng XW, Wang LC. A collaborative filtering algorithm based on users' social relationship mining in mobile communication network. Journal of Electronics & Information Technology, 2011,33(12):3002−3007 (in Chinese with English abstract).

[15] Jannach D, Geyer W, Dugan C, Freyne J. RecSys 09: Workshop on recommender systems and the social Web. In: Proc. of the ACM RecSys 2009. New York: ACM Press, 2009. 421−422.

[16] Freyne J, Anand SS. Recommender systems and the social Web. In: Proc. of the ACM RecSys 2011. New York: ACM Press, 2011. 383−384. [doi: 10.1145/2043932.2044014]

[17] Mobasher B, Jannach D, Geyer W, Freyne J. Recommender systems and the social Web. In: Proc. of the ACM RecSys 2013, ACM Press, 2013. 477−478. [doi: 10.1145/2507157.2508001]

[18] Meng XW, Hu X, Wang LC, Zhang YJ. Mobile recommender systems and their applications. Ruan Jian Xue Bao/Journal of Software, 2013,24(1):91−108 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/4292.htm [doi: 10.3724/SP.J.1001.2013.04292]

[19] Min JK, Cho SB. Mobile human network management and recommendation by probabilistic social mining. IEEE Trans. on Systems, Man and Cybernetics, Part B, 2011,41(3):761−771. [doi: 10.1109/TSMCB.2010.2092424]

[20] Jung JJ. Contextualized mobile recommendation service based on interactive social network discovered from mobile users. Expert Systems with Applications, 2009,36(1):11950−11956. [doi: 10.1016/j.eswa.2009.03.067]

[21] Saravanan M, Ericsson RD. Bayesian filters for mobile recommender systems. In: Proc. of the Advances in Social Networks Analysis and Mining (ASONAM). Washington: IEEE Computer Society, 2011. 25−27. [doi: 10.1109/ASONAM.2011.51]

[22] Zanda A, Eibe S, Menasalvas E. SOMAR: A social mobile activity recommender. Expert Systems with Applications, 2012,39: 8423−8429. [doi: 10.1016/j.eswa.2012.01.178]

[23] Symeonidis P, Papadimitriou A, Manolopoulos Y. Geo-Social recommendations based on incremental tensor reduction and local path traversal. In: Proc. of the 3rd ACM SIGSPATIAL Int'l Workshop on Location-Based Social Networks. 2011. 89−96. [doi: 10.1145/2063212.2063228]

[24] Kazienko P, Musial K, Kajdanowicz T. Multidimensional social network in the social recommender system. IEEE Trans. on Systems, Man, and Cybernetics—Part A: Systems and Humans, 2011,41(4):746−759. [doi: 10.1109/TSMCA.2011.2132707]

[25] Yuan Q, Cong G, Ma Z, Sun A. Time-Aware point-of-interest recommendation. In: Proc. of the SIGIR. New York: ACM Press, 2013. 363−372. [doi: 10.1145/2484028.2484030]

© 中国科学院软件研究所 http://www.jos.org.cn

[26] Levandoski JJ, Sarwat M, Eldawy A, Mokbel M. LARS: A location-aware recommender system. In: Proc. of the ICDE. Washington: IEEE Computer Society, 2012. 450−461. [doi: 10.1109/ICDE.2012.54]

[27] Yin H, Cui B, Chen L, Hu Z, Huang Z. A temporal context-aware model for user behavior modeling in social media systems. In: Proc. of the SIGMOD/PODS. New York: ACM Press, 2014. 1543−1554. [doi: 10.1145/2588555.2593685]

[28] Yin H, Sun Y, Cui B, Hu Z, Chen L. LCARS: A location-content-aware recommender system. In: Proc. of the KDD. New York: ACM Press, 2013. 221−229. [doi: 10.1145/2487575.2487608]

[29] Wang LC, Meng XW, Zhang YJ. Context-Aware recommender systems: A survey of the state-of-the-art and possible extension. Ruan Jian Xue Bao/Journal of Software, 2012,23(1):1−20 (in Chinese with English abstract). http://www.jos.org.cn/1000-9825/4100.htm [doi: 10.3724/SP.J.1001.2012.04100]

[30] Resnick P, Varian HR. Recommender systems. Communications of the ACM, 1997,40(3):56−58. [doi: 10.1145/245108. 245121]

[31] Yang X, Guo Y, Liu Y, Steck H. A survey of collaborative filtering based social recommender systems. Computer Communications, 2014,41:1−10. [doi: 10.1016/j.comcom.2013.06.009]

[32] Guy I, Carmel D. Social recommender systems. In: Proc. of the 20th Int'l Conf. on Companion on World Wide Web. New York: ACM Press, 2011. 283−284.

[33] Tang J, Hu X, Liu H. Social recommendation: A review. Social Networks Analysis and Mining, 2013,3(4):1113−1133. [doi: 10.1007/s13278-013-0141-9]

[34] Wellman B, Salaff J. Computer networks as social networks: Collaborative work, telework and virtual community. Annual Review Social, 1996,22:213−238. [doi: 10.1146/annurev.soc.22.1.213]

[35] Li YM, Hsiao HW. Recommender service for social network based applications. In: Proc. of the 11th Int'l Conf. on Electronic Commerce (ICEC). New York: ACM Press, 2009. 378−381. [doi: 10.1145/1593254.1594683]

[36] Jiang W, Wang G, Wu J. Generating trusted graphs for trust evaluation in online social networks. Future Generation Computer Systems, 2014,31:48−58. [doi: 10.1016/j.future.2012.06.010]

[37] Sherchan W, Nepal S, Paris C. A survey of trust in social networks. ACM Computing Surveys, 2013,45(4):Article No.47. [doi: 10.1145/ 2501654.2501661]

[38] Massa P, Avesani P. Trust metrics on controversial users: Balancing between tyranny of the majority and echo chambers. Int'l Journal on Semantic Web and Information Systems, 2007,3(1):1−21. [doi: 10.4018/jswis.2007010103]

[39] Ziegler CN, Lausen G. Spreading activation models for trust propagation. In: Proc. of the IEEE Int'l Conf. on e-Technology, e-Commerce, and e-Service. Washington: IEEE Computer Society, 2004. 83−97. [doi: 10.1109/EEE.2004.1287293]

[40] Levien, Aiken. Advogato's trust metric. 002. http: //advogato.org/trust-metric.html

[41] Golbeck J. Computing and applying trust in Web-based social networks [Ph.D. Thesis]. University of Maryland College Park, 2005.

[42] Massa P, Avesani P. Trust-Aware recommender systems. In: Proc. of the RecSys 2007. New York: ACM Press. 2007. 17−24. [doi: 10.1145/1297231.1297235]

[43] Kuter U, Golbeck J. Using probabilistic confidence models for trust inference in Web-based social networks. ACM Trans. on Internet Technology, 2010,10(2):1−23. [doi: 10.1145/1754393.1754397]

[44] Liu F, Lee HJ. Use of social network information to enhance collaborative filtering performance. Expert System with Applications, 2010,37(7):4772−4778. [doi: 10.1016/j.eswa.2009.12.061]

[45] Yu L, Pan R, Li Z. Adaptive social similarities for recommender systems. In: Proc. of the ACM RecSys 2011. ACM Press, 2011. 257−260. [doi: 10.1145/2043932.2043978]

[46] Silva NB, Tsang IR, Cavacanti GDC. A graph-based friend recommendation system using genetic algorithm. In: Proc. of the 2010 IEEE Congress on Evolutionary Computation (CEC). Washington: IEEE Computer Society, 2010. 1−7. [doi: 10.1109/CEC.2010.5586144]

[47] Li X, Chen H. Recommendation as link prediction in bipartite graphs: A graph kernel-based machine learning approach. Decision Support Systems, 2013,54(2):880−890. [doi: 10.1016/j.dss.2012.09.019]

[48] Liben-Nowell D, Kleinberg J. The link prediction problem for social networks. Journal of the American Society for Information Science and Technology, 2007,58(7):1019−1031. [doi: 1002/asi.20591]

[49] Lü L, Zhou T. Link prediction in complex networks: A survey. Physica A: Staticstical Methanics and its Application, 2011,390(6): 1150−1170. [doi: 10.1016/j.physa.2010.11.027]

[50] Mohsen J, Martin E. TrustWalker: A random walk model for combing trust-based and item-based recommendation. In: Proc. of the KDD. New York: ACM Press, 2009. 397−406. [doi: 10.1145/1557019.1557067]

[51] Shang S, Kuikarni SR, Cuff PW. A random walk based model incorporating social information for recommendations. In: Proc. of the 2012 IEEE Int'l Workshop on Machine Learning for Signal Processing (MLSP). 2012. 1−6. [doi: 10.1109/MLSP.2012.6349732]

[52] Jiang M, Cui P, Wang F. Social recommendation across multiple relational domains. In: Proc. of the CIKM. New York: ACM Press, 2012. 1422−1431. [doi: 10.1145/2396761.2398448]

© 中国科学院软件研究所 http://www.jos.org.cn

[53] Li J, Ma S, Hong S. Recommendation on social network based on graph model. In: Proc. of the 31th Chinese Control Conf. Washington: IEEE Computer Society, 2012. 7548−7551.
[54] Dong Y, Tang J, Wu S, Tian JL, Chawla NV, Rao JH, Cao HH. Link prediction and recommendation across heterogeneous social networks. In: Proc. of the ICMD. Washington: IEEE Computer Society, 2012. 181−190. [doi: 10.1109/ICDM.2012.140]
[55] Bachstrom L, Leskovec J. Supervised random walks: Predicting and recommending links in social networks. In: Proc. of the WSDM. New York: ACM Press, 2011. 635−644. [doi: 10.1145/1935826.1935914]
[56] Lichtenwalter RN, Lussier JT, Chawla N. New perspectives and methods in link prediction. In: Proc. of the KDD. New York: ACM Press, 2010. 243−252. [doi: 10.1145/1835804.1835837]
[57] Kuo TT, Yan R, Huang YY, Kung PH. Unsupervised link prediction using aggregative statistics on heterogeneous social networks. In: Proc. of the KDD. New York: ACM Press, 2013. 755−783. [doi: 10.1145/2487575.2487614]
[58] Liu J, Zhang F, Song X. What's in a name? An unsupervised approach to link users across communities. In: Proc. of the WSDM. New York: ACM Press, 2013. 495−504. [doi: 10.1145/2433396.2433457]
[59] Wang H, Manolis T, Nikos M. Location recommendation in location-based social networks using user check-in data. In: Proc. of the GIS. 2013. 364−373. [doi: 10.1145/1869790.1869861]
[60] Lee DD, Seung HS. Learning the parts of objects by non-negative matrix factorization. Journal of Nature, 1999,401(6755):788−791. [doi: 10.1038/44565]
[61] Jiang M, Cui P, Liu R, Yang Q. Social contextual recommendation. In: Proc. of the CIKM. New York: ACM Press, 2012. 45−55. [doi: 10.1145/2396761.2396771]
[62] Noel J, Sanner S, Tran KN. New objective functions for social collaborative filtering. In: Proc. of the 20th Int'l World Wide Web Conf. (WWW 2012). New York: ACM Press, 2012. 859−869. [doi: 10.1145/2187836.2187952]
[63] Cui P, Wang F, Liu S, Ou M. Who should share what? Item-Level social influence prediction for users and posts ranking. In: Proc. of the SIGIR. New York: ACM Press, 2011. 185−194. [doi: 10.1145/2009916.2009945]
[64] Yang SH, Long B, Smola A, Sadagopan N. Like like alike: Joint friendship and interest propagation in social networks. In: Proc. of the 19th Int'l World Wide Web Conf. (WWW 2011). New York: ACM Press, 2011. 537−546. [doi: 10.1145/1963405.1963481]
[65] Li WJ, Yeung DY. Relation regularized matrix factorization. In: Proc. of the IJCAI 2009. 2009. 1126−1131.
[66] Ma H, Yang H. Lyu MR. Sorec: Social recommendation using probabilistic matrix factorization. In: Proc. of the CIKM 2008. New York: ACM Press, 2008. 931−940. [doi: 10.1145/1458082.1458205]
[67] Lian D, Zhao C, Xie X, Sun G, Chen E, Rui Y. GeoMF: Joiny geographical modeling and matrix factorization for point-of-interest recommendation. In: Proc. of the KDD. New York: ACM Press, 2014. 831−840. [doi: 10.1145/2623330.2623638]
[68] Koren Y, Bell RM, Volinsky C. Matrix factorization techniques for recommender systems. IEEE Computer, 2009,42(8):30−37. [doi: 10.1109/MC.2009.263]
[69] Chen T, Zheng Z, Lu Q. Informative ensemble of multi-resolution dynamic factorization models. In: Proc. of the KDD Cup Workshop. New York: ACM Press, 2011.
[70] Chen T, Zheng Z, Lu Q. Feature-Based matrix factorization. CoRR abs/1109.2271, 2011.
[71] Chen K, Chen T, Zheng Q. Collaborative personalized tweet recommendation. In: Proc. of the SIGIR. New York: ACM Press, 2012. 661−670. [doi: 10.1145/2348283.2348372]
[72] Chen T, Tang L, Liu Q. Combining factorization model and additive forest for collaborative followee recommendation. In: Proc. of the KDD Cup Workshop. New York: ACM Press, 2012.
[73] Ma T, Yang Y, Wang Y. Recommending people to follow using asymmetric factor models with social graphs. In: Proc.of the 17th Online Word Conf. on Soft Computing in Industrial Applications, Vol.223. 2014. 265−276. [doi: 10.1007/978-3-319-00930-8_24]
[74] Yang X, Steck H, Guo Y, Liu Y. On top-*k* recommendation using social networks. In: Proc. of the RecSys 2012. New York: ACM Press, 2012. 67−74. [doi: 10.1145/2365952.2365969]
[75] Yang X, Steck H, Liu Y. Circle-Based recommendation in online social networks. In: Proc. of the KDD. New York: ACM Press, 2012. 1267−1275. [doi: 10. 1145/2339530.2339728]
[76] Cheng C, Yang HQ, Lyu MR, Irwin K. Where you like to go next: Successive point-of-interest recommendation. In: Proc. of the IJCAI. 2013. 2605−2611.
[77] Rendle S. Factorization machines. In: Proc. of the 10th IEEE Int'l Conf. on Data Mining. Washington: IEEE Computer Society, 2010. 995−1000. [doi: 10.1109/ICDM.2010.127]
[78] Rendle S, Zeno G, Christoph F. Fast context-aware recommendations with factorization machines. In: Proc. of the SIGIR. New York: ACM Press, 2011. 635−644. [doi: 10.1145/2009916.2010002]
[79] Rendle S. Social network and click-through prediction with factorization machines. In: Proc. of the KDD Cup Workshop. New York: ACM Press, 2012.
[80] Hong L, Doumith AS, Davison BD. Co-Factorization machines: Modeling user interests and predicting individual decisions in twitter. In: Proc. of the WSDM. New York: ACM Press, 2013. 557−566. [doi: 10.1145/2433396.2433467]
[81] Hofmann T, Puzicha J. Latent class models for collaborative filtering. In: Proc. of the IJCAI. 1999. 688−693.

© 中国科学院软件研究所 http://www.jos.org.cn

[82] Ye M, Liu X, Lee WH. Exploring social influences for recommendation-a probabilistic generative model approach. In: Proc. of the SIGIR. New York: ACM Press, 2012. 671−680. [doi: 10.1145/2348283.2348373]

[83] He J, Chu W. Design considerations for a social network-based recommendation system (SNRS). In: Proc. of the Community-Built Databases. 2011. 73−106. [doi: 10.1007/978-3-642-19047-6_4]

[84] Yang X, Guo Y, Liu Y. Bayesian-Inference based recommendation in online social networks. IEEE Trans. on Parallel and Distributed System, 2013,24(4):642−651. [doi: 10.1109/TPDS.2012.192]

[85] Ye M, Yin P, Lee W. Exploiting geographical influence for collaborative point-of-interest recommendation. In: Proc. of the SIGIR. New York: ACM Press, 2011. 325 −334. [doi: 10.1145/2009916.2009962]

[86] Li R, Wang S, Deng H. Towards social user profiling: Unified and discriminative influence model for inferring home locations. In: Proc. of the KDD. New York: ACM Press, 2012. 1023−1031. [doi: 10.1145/2339530.2339692]

[87] Kim Y, Park Y, Shim K. DIGTOBI: A recommendation system for dig articles using probabilistic modeling. In: Proc. of the WWW. New York: ACM Press, 2013. 691−701.

[88] Kim Y, Shim K. TWILITE: A recommendation system for twitter using a probilistic model based on latent dirichlet allocation. Information Systems, 2014,42:59−77. [doi: 10.1016/j.is.2013.11.003]

[89] Yu X, Pan A, Tang LA, Han JW. Geo-Friends recommendation in GPS-based cyber-physical social network. In: Proc. of the Int'l Conf. on Advances in Social Networks Analysis and Mining. Washington: IEEE Computer Society, 2011. 361−368. [doi: 10.1109/ASONAM.2011.118]

[90] Machanavajjhala A, Korolova, Sarma AS. Personalized social recommendations-accurate or private? In: Proc. of the VLDB. New York: ACM Press, 2011. 440−450. [doi: 10.14778/1988776.1988780]

[91] Phuong ND, Thang LQ, Phuong TM. A graph-based method for combining collaborative and content based filtering. In: Proc. of the 10th Pacific Rim Int'l Conf. on Artificial Intelligence. Berlin, Heidelberg: Springer-Verlag, 2008. 859−869. [doi: 10.1007/978-3-540-89197-080]

附中文参考文献:

[1] 许海玲,吴潇,李晓东,阎保平.互联网推荐系统比较研究.软件学报,2009,20(2):350−362. http://www.jos.org.cn/1000-9825/ 3388.htm [doi: 10.3724/SP.J.1001.2009.00350]

[14] 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.电子与信息学报,2011,33(12):3002−3007.

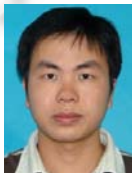
[18] 孟祥武,胡勋,王立才,张玉洁.移动推荐系统及其应用.软件学报,2013,24(1):91−108. http://www.jos.org.cn/1000-9825/4292.htm [doi: 10.3724/SP.J.1001.2013.04292]

[29] 王立才,孟祥武,张玉洁.上下文感知推荐系统研究进展.软件学报,2012,23(1):1−20. http://www.jos.org.cn/1000-9825/4100.htm [doi: 10.3724/SP.J.1001.2012.04100]

孟祥武(1966－),男,山东招远人,博士,教授,博士生导师,CCF 高级会员,主要研究领域为网络服务,通信软件,人工智能.

张玉洁(1969－),女,副教授,主要研究领域为智能信息处理,通信软件,网络服务.

刘树栋(1984－),男,博士生,主要研究领域为推荐系统.



胡勋(1985－),男,博士,主要研究领域为社会网络分析,推荐系统.

© 中国科学院软件研究所 http://www.jos.org.cn