

社会化推荐研究综述

王刚 蒋军 王含茹

(合肥工业大学管理学院 合肥 230009)

摘要 社会化推荐系统正随着互联社交网络的快速发展逐渐成为人们关注的热点问题。首先,介绍了社会化推荐的基础理论,阐述了社会化推荐的概念及基本框架,并在此基础上将其分类为面向个体的社会化推荐和面向群组的社会化推荐。接着,分别给出了面向个体和面向群组的社会化推荐的形式化定义,从个体和群组两个角度对社会化推荐系统的研究现状进行了综述。面向个体的社会化推荐主要包括基于评分预测的推荐方法和基于排序学习的推荐方法;面向群组的社会化推荐主要包括推荐方法的融合和推荐结果的融合。

关键词 社会化推荐,面向个体,面向群组,综述

中图分类号 TP391.3 **文献标识码** A

Review of Social Recommendation

WANG Gang JIANG Jun WANG Han-ru

(School of Management, Hefei University of Technology, Hefei 230009, China)

Abstract Social recommendation system is becoming a hot topic of concern with the rapid development of Internet social network. First of all, this paper introduced the basic theory of social recommendation, and explained the concept and basic framework of social recommendation. And on this basis, this paper classified social recommendation as individual-oriented social recommendation and group-oriented social recommendation. Then, this paper gave the formal definition for individual-oriented social recommendation and group-oriented social recommendation respectively, and summarized the current research status in the view of individual and group. Individual-oriented social recommendation includes recommended methods based on predicting and sequential learning. Group-oriented social recommendation includes recommended methods based on integration of method and the integration of results.

Keywords Social recommendation, Individual-oriented, Group-oriented, Review

1 引言

随着互联网技术的普及和飞速发展,特别是 Web2.0 理论和技术的引入,以社交网站、微博、博客、内容分享社区、维基等为代表的社交媒体受到了广泛的关注^[1]。流量排名网站 Alexa 最新的统计结果显示,世界访问排名前十的网站中,有 5 个是社交媒体网站^[2]。中国互联网中心公布的第 39 次《中国互联网发展状况统计报告》显示,到 2016 年 12 月底为止,中国网民规模达 7.31 亿,其中,网民中社交应用的使用率为 79%;微博用户规模为 2.71 亿人,使用率为 37.1%;博客和个人空间的个人用户数量达到 4.96 亿人,使用率为 67.8%^[3]。随着参与社交媒体的用户数量的急剧增加,数以亿计的用户每天在社交媒体上自助生成了大量的内容。这些用户产生的内容从个人经历(如看电影、旅行和听音乐等)到意见的表达(小到对某部电影、某个商品的评价,大到某些公共事件的舆论),无所不有^[4-5]。社交媒体具有的信息发布和接收异常简便、“核裂变”式的信息传播扩散模式、容易形成自媒体和虚拟

社区等特点,一方面丰富了信息的来源,加速了信息的扩散,另一方面也形成了所谓的信息过载问题,致使用户在面对如此海量的信息时无法选择出自己满意的或者对自己有价值的信息^[5-6]。推荐系统作为一种高效的信息过滤手段,恰是解决信息过载和实现个性化服务的有效途径之一^[7]。因此,面向社交媒体的推荐系统研究——社会化推荐,正逐渐成为当前学术界和产业界共同关注的热点问题^[8]。

社会化推荐是在传统推荐系统的基础上,把用户或者项目的社会化属性信息(如标签信息、朋友关系和信任关系等)作为重要的影响因素融合到推荐方法中,以提升推荐系统的精度和性能^[8]。从这个角度来讲,社会化推荐是一种把用户或者项目的社会化属性信息作为附加输入的推荐方法。与传统推荐系统相比,已有的社会化推荐具有了一些新的特点,首先,社会化推荐所用的数据呈现多样性的特点^[9-10]。传统推荐系统主要利用用户-项目的交互信息构建用户的偏好模型,进而利用此模型预测用户对项目的偏好程度,并主动推荐给用户,从而达到个性化服务的目的;而社会化推荐中则存在大

本文受国家自然科学基金(71471054,91646111),安徽省自然科学基金(1608085MG150),合肥工业大学应用培育计划项目(JZ2017YYPY0235)资助。

王刚(1980—),男,博士,教授,主要研究方向为信息管理与信息系统,E-mail:wgedison@gmail.com;蒋军(1991—),男,硕士,主要研究方向为推荐系统;王含茹(1994—),女,硕士生,主要研究方向为推荐系统。

量的非结构化的数据,如评论、标签和微博等信息。其次,社会化推荐所面临的范围更广。传统的推荐系统大多运用于电子商务领域的商品推荐;而社会化推荐的推荐对象不再单一,除了商品推荐以外,还有好友推荐、标签推荐和旅游推荐等^[11-13]。最后,社会化推荐方法的适应性需要不断提高。传统推荐系统所选用的推荐方法主要包括基于内容的推荐、协同过滤推荐和混合推荐方法;而社会化推荐由于面临的情况更加复杂,因此需要引入一些其他领域的方法,来构建适应性更强的推荐方法^[14-16]。社会化推荐不仅面临着以上挑战,还存在推荐系统领域常见的数据稀疏性、冷启动和可解释性差等问题^[8,17]。

本研究对社会化推荐系统进行了深入研究,具有重要的理论意义和实践意义。首先,对社会化推荐的基础理论进行了系统的梳理,对社会化推荐的研究脉络和现状有了较为清晰的认识。其次,本研究深入地学习了面向个体和面向群组的社会化推荐系统的相关理论,并对国内外的研究成果进行了系统分析,有助于进一步认识和理解社会化推荐。总体来说,本研究对社会化推荐系统进行了较为完善的综述总结,丰富和完善了社会化推荐系统的理论体系,对推动社会化推荐系统的理论研究和实践应用具有重要的作用。

2 社会化推荐系统的基础理论

社会化推荐的概念最早出现在1997年。随着后期Facebook, Twitter, YouTube和新浪微博等社交媒体的兴起,社交媒体上产生了大量的用户和项目的社会化属性信息,进而使得社会化推荐成为推荐系统领域的重要研究方向之一,逐渐得到了学术界和产业界的广泛关注^[16-17]。ACM推荐系统年会从2009年开始设置与社会化推荐相关的专题讨论^[18],并在2011年的专题研讨会上讨论了社会化推荐未来的几个重要研究方向,同时提出了社会化推荐领域所面临的10个研究热点和难点^[19];在2013年的主题研讨会上,学者们又进一步提出了社会化推荐的4个研究方向^[20]。本研究主要从社会化推荐的概念及分类、社会化推荐的基本框架等方面对社会化推荐进行介绍。

2.1 社会化推荐的概念及分类

传统推荐系统主要通过挖掘用户和项目之间的二元关系,来帮助用户从海量的信息中发现自己感兴趣的内容,能够满足用户的个性化需求^[6]。而社会化推荐则在已有传统推荐系统的基础上,把用户或者项目相关的社会化属性信息(例如好友关系、信任关系和项目的标签等)作为重要影响因素引入到推荐模型中,用来提高推荐系统的整体性能。从这个角度上说,社会化推荐是把用户或者项目的社会化属性信息作为传统推荐系统附加输入的一种推荐方法^[21]。然而近些年来,随着社交媒体的发展,社会化推荐的推荐项目已经不仅仅局限在普通项目(图书、音乐和电影等),还包括社交媒体上广泛存在的标签、好友、群组等。社会化推荐方法可利用的数据也不仅仅包括用户-项目交互数据和用户之间的社会化关系信息,还包括社会化标签信息、用户的交互关系信息、用户的点击和收藏行为等。从这个角度来讲,社会化推荐是把整个社交媒体当作推荐目标的推荐方法。因此,Tang等^[17]将上述两种定义称为社会化推荐的狭义定义和广义定义。

由于社会化推荐技术的研究还处于蓬勃发展阶段,因此到目前为止,对于社会化推荐方法的分类还没有一个被普遍认可的体系。部分学者根据社会化推荐是从传统推荐系统基础上发展而来这一特性,按照已有传统推荐方法的分类准则,将社会化推荐分为基于内容的推荐、协同过滤推荐和混合推荐等^[17]。但是伴随着社交网络分析和文本情感分析等技术在社会化推荐中的应用,近年来,几种比较典型的社会化推荐方法逐渐得到了研究者的关注,主要包括基于图结构的推荐模型、矩阵分解模型和概率模型等^[10,15-16]。因此,又有部分学者从社会推荐面向的对象的角度,将社会化推荐分为面向个体的社会化推荐和面向群组的社会化推荐,这两者分别是对社交媒体上的个体用户和群组用户进行推荐。目前,对于社会推荐方法的分类还没有统一的标准,本文主要根据面向个体的和面向群组的社会化推荐分类标准对已有的社会化推荐的基础理论进行介绍。

2.2 社会化推荐的基本框架

目前,已有很多研究者在社会化推荐的研究中做了大量的工作,并取得了一些研究成果。现有的研究成果大部分是利用用户或者项目的社会化属性信息对推荐方法的影响、社会化网络内部的结构特征和项目在社交媒体上的流行度等,提出不同的社会化推荐基本框架^[15,22-23]。本研究结合已有的社会化推荐技术的特点和系统流程,提出了如图1所示的新的社会化推荐的基本框架,该框架主要包括以下3层。

(1)社会化推荐的数据收集层:该层的主要作用是通过网络爬虫等信息技术从社交媒体上获取社会化推荐所需的数据。这些数据大致可以分为3类:用户和项目的交互信息、用户和项目的基本属性信息以及用户和项目的社会化属性信息。其中,用户和项目的交互信息主要包括用户对项目的点击、收藏和评分等信息,用户和项目的基本属性信息主要包括用户的年纪、地理位置、工作、项目的名称和类别等,用户和项目的社会化属性信息主要包括用户之间的好友信息、关注信息、信任关系,以及项目之间拥有的相同标签和相同类别信息等。

(2)社会化信息处理层:这一层的主要作用是对社会化推荐的数据收集层收集到的数据进行处理,并将处理后的结果作为社会化推荐的结构化输入。该层主要包括:利用用户和项目之间的交互信息构建用户-项目交互矩阵;综合利用用户的基本属性信息和社会化属性计算用户之间的相关性;同理,还可利用项目的基本属性和其社会化属性计算项目之间的相关性。该层的核心部分在于用户和项目的相关性计算,其是影响社会化推荐质量和精度的一个重要因素。

(3)社会化推荐结果生成层:该层是社会化推荐的核心层,其主要工作是将社会化信息处理层预处理好的数据作为设计的社会化推荐模型的输入,进而产生社会化推荐的结果。而社会化推荐模型根据面向对象的不同,又可以分为面向个体的社会化推荐模型和面向群组的社会化推荐模型。在上述两种社会化推荐模型中,不仅需要考虑到用户-项目交互信息对模型的影响,而且还要重点考虑用户和项目的社会化信息对推荐结果的影响。其中,如何将社会化信息融合到推荐模型中以及社会化信息在推荐过程中作用大小的衡量,是目前研究者普遍关注的问题。

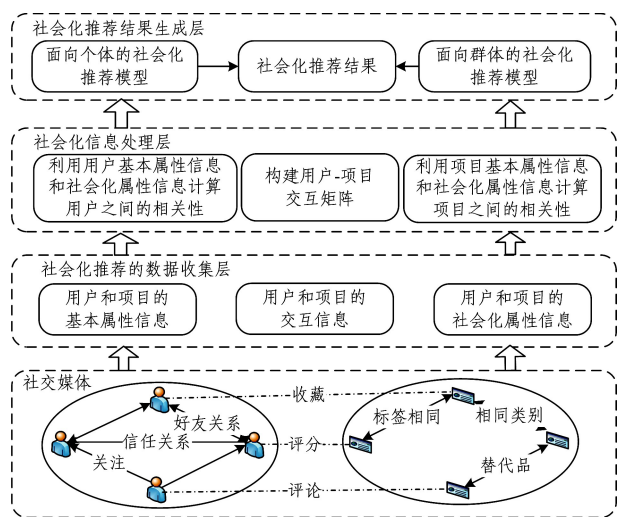


图 1 社会化推荐的基本框架

3 面向个体的社会化推荐的理论研究

面向个体的社会化推荐系统是帮助个体用户从社交媒体上用户生成的海量数据中准确地找到自己满意的信息,能为个体用户节省时间,并帮助个体用户发现自己可能感兴趣的信息,提高个体用户对信息的利用率。目前,已有众多的研究者就面向个体的社会化推荐展开研究,已对面向个体的社会化推荐形成了统一的学习框架,并且提出了许多经典的面向个体的社会化推荐方法。本研究将对面向个体的社会化推荐的形式化定义以及国内外的研究现状进行介绍。

3.1 面向个体的社会化推荐的形式化定义

面向个体的社会化推荐是在传统推荐系统的基础上发展而来的,因此参考传统推荐系统的定义,可以得到社会化推荐的形式化定义。令 U 为数据集中所有用户的集合, I 为所有待推荐项目的集合, 设 $S = \{S_{i,j}\}_{|U| \times |I|}$ 为通过社会化属性信息计算得到的用户社会化关系矩阵, 映射 $\mu: U \times I \rightarrow R$ 是社会化推荐结果的评价函数, 其中 R 用来表示推荐方法的效用值, 其取值范围是全序非负实数集, $Y = \{U_y | U_y \in U, x \neq y, S_{x,y} \neq 0, S_{x,y} \in S\}$, 表示在社交媒体上与用户 u_x 存在社会化关系的其他用户。因此, 可以将社会化推荐要研究的问题描述为: 在与用户 u_x 具有社会化关系的用户群组中, 根据待推荐项目在用户群组中的效用值, 为用户推荐与其偏好最相似的、评价效用值最大的项目集合 I^* , 即:

$$\forall U_x \in U$$

$$I^* = \arg \max (\mu_i + \alpha \mu(U_x, i) + \beta \frac{1}{Y} \sum_y \mu(U_y, i)) \quad (1)$$

其中, $\forall \alpha, \beta \in [0, 1]$, 且 $\alpha + \beta = 1$ 。

根据上述面向个体的社会化推荐的形式化定义, 现有的面向个体的社会化推荐的基础方法根据预测效用值的不同可以分为两类: 1) 预测的效用值为用户对项目的预测评分, 即基于评分预测的面向个体的推荐方法; 2) 预测效用值为用户对项目排序的偏好值, 即基于排序预测的面向个体的推荐方法^[24]。

3.2 国内外的研究现状

由于面向个体的社会化推荐所面临的推荐对象为个体用户, 与传统推荐系统极为相似, 因此有部分学者直接将传统推荐的方法运用到面向个体的社会化推荐中。而传统的面向个

体的推荐方法根据预测目标的不同可以分为两类: 基于评分预测的推荐方法和基于排序学习的推荐方法^[24]。

基于评分预测的推荐方法在推荐的过程中重点考虑用户对项目的操作信息, 如评分、点击和收藏等, 并将推荐问题最终转化为用户对项目的评分预测问题。已有的基于评分预测的推荐方法大致有 3 种, 即基于内容的推荐方法、协同过滤方法和混合推荐方法^[5, 25-26]。其中, 基于内容的推荐方法的主要原理是通过收集用户的历史行为记录, 利用其对用户的偏好进行建模, 从而形成用户的偏好特征描述^[26]。同时, 需要对待推荐项目的内容信息进行分析, 构建项目的资源特征描述。最后, 度量用户偏好特征描述和待推荐项目特征描述之间的相似性, 并将相似性较高的项目推荐给用户, 从而产生面向个体用户的推荐列表。另外, 为了提高推荐的精度, 一些研究者也将更复杂的方法引入其中, 如人工神经网络、决策树和聚类技术等^[6]。协同过滤推荐方法自提出以来, 便得到了研究者的普遍关注, 逐渐成为应用最为广泛的技术之一。它主要是基于以下假设: 具有相似兴趣的用户会喜欢相似的项目^[7, 27]。这样, 它就可以依据用户对项目的评分信息发现用户和项目之间的相关性, 然后依据这种相关性向用户推荐项目。协同过滤方法一般又分为基于内存的方法和基于模型的方法两类^[7]。其中, 基于内存的方法根据计算角度的不同又可以分为基于用户的推荐和基于项目的推荐^[25]。而基于模型的推荐方法则是通过用户的评分信息进行学习, 并据此推断用户的行为模式, 进而根据此模型预测用户对项目的评分。目前常用的模型有矩阵分解模型、聚类模型和贝叶斯网络模型等^[5, 28]。经过大量的实践证明, 基于内容的推荐和协同过滤推荐在一定程度上还存在一些弊端, 为此一些研究者通过将上述两种方法相结合, 提出了混合推荐方法^[5]。混合推荐方法通过不同的组合方式将基于内容的推荐方法和协同过滤推荐方法进行融合, 并且两种推荐方法可以在推荐的不同阶段进行融合。

基于排序学习的推荐方法一般是依据用户偏好模型来进行项目推荐的, 与用户偏好越相似的项目推荐给用户的可能性应该越大, 其主要思想则是对用户已有的收藏、评分和浏览等历史行为记录进行建模, 学习得到项目的排序函数, 最后据此为每个用户生成项目推荐列表^[29]。目前已有的基于排序学习的推荐方法根据输入模型样例的不同, 大致可以分为 3 类: 点级 (PointWise)、对级 (PairWise) 和列表级 (ListWise)^[29-30]。其中, 基于点级排序学习的推荐方法的处理对象是单个项目, 其主要思路是将基于排序学习的推荐问题转化为多分类或者回归问题, 最终模型将会依据学习到的分类或者回归函数预测用户对未评分项目的评分, 对预测评分进行排序后即可得到用户的项目推荐列表。近年来, 被众多研究者广泛采用的模型之一——隐语义模型 (Latent Factor Model) 就是采用了基于点级排序学习的思想^[31]。而基于对级排序学习的推荐方法则是将排序问题转化为二分类问题, 其侧重于对两个项目间的排序关系是否合理进行判断。其主要思路是根据用户已有的评分信息得到项目对, 将每个项目对的项目转化为特征向量后, 得到训练数据集中的一个具体训练实例, 然后选取合适的机器学习方法来学习最终的分类函数。常用的机器学习方法有很多, 如神经网络和支持向量机等^[32-33]。与基于点级和基于对级排序学习的推荐方法不

同,基于列表级排序学习的推荐方法是直接对项目的推荐列表进行优化。目前的优化方式主要有两种,一种是直接对排序学习的评价指标进行优化,例如常用的评价指标 MAP 和 NDCG 等^[30];另一种优化方式是构造损失函数,这与点级排序学习和对级排序学习的思路相同。目前,研究者已经提出了众多的构建损失函数的方法,如 ListWise 方法使用 KL 距离,RankCosine 使用预测排序与正确排序的分值向量间的夹角等^[30,34]。

上述两类传统推荐方法虽然在面向个体的社会化推荐中取得了一定的效果,但是却忽视了在社交媒体环境下,仅仅依据传统的用户-项目的交互信息直接为用户生成推荐结果往往是不合适的,很难满足用户的个性化需求。为此,一些研究者开始利用社交媒体中的朋友关系、信任关系和社会影响等因素来完善已有的社会化推荐系统。近几年,比较典型的面向个体的社会化推荐方法主要分为两类:基于社会网络分析的社会化推荐和基于文本挖掘方法的社会化推荐^[22]。其中,基于社会化网络分析的社会化推荐是对社交媒体中广泛存在的用户或者项目的社会化关系网络的拓扑结构进行分析,依据用户或者项目在网络拓扑结构中的紧密程度分别为其赋予不同的权重。目前,已有很多学者研究了基于社会网络的社会化推荐方法,如 SoNARS、GLOSSS 和社会网络协同过滤等^[8,17,35]。同时,信息之所以能够在社会化网络中实现快速的传播,主要是基于用户之间的信任关系,即形成的信任朋友关系网络。已有很多学者将信任关系引入到传统推荐方法中,提出了很多新的社会化推荐方法,主要包括 TidalTrust^[36], MoleTrust^[37], TrustWaler^[38], SoRec^[10], STE^[39], mTrust^[40] 和 SocialMF^[41] 等。而基于文本挖掘的社会化推荐方法则是利用文本挖掘技术对社交媒体上的评论、标签、博客和微博等内容信息进行更加深入的分析,从而能更深入地挖掘用户的个性化偏好,对没有项目打分或者存在恶意评价等问题都能准确得知用户的整体轮廓^[9,42-43]。现有的基于文本挖掘的社会化推荐有两个重要的分支,一个是对用户社会标签的挖掘,研究以标签的形式对用户的特征进行概述^[13,44-45];另一个则是对社交媒体中文本的整体内容进行情感分析,由于用户对商品的评价是带有个人的意见和感情的,因此如何从用户的整体文本中找到他们的真实情感和爱好对推荐系统而言至关重要^[14,43,46]。

由以上可知,面向个体的社会化推荐的研究现状及发展动态可总结为:传统推荐系统主要是基于用户评分偏好矩阵进行推荐的,对社交媒体中蕴含的大量用户自助生成的内容却没有加以分析和利用;目前已有学者认识到这个问题,并从社会网络分析和文本情感分析等多个角度出发,将社交媒体上的社会化信息融入到传统推荐系统中,构建了社会化推荐系统。

4 面向群组的社会化推荐的理论研究

近些年来,随着社交媒体的不断发展,其中的许多用户逐渐因为兴趣爱好相同组成了大量的群组,它将推荐对象由单个用户扩展为了一个群组^[47-49]。因此,针对这些群组用户的社会化推荐技术受到研究者的普遍关注。下面主要对面向群组的社会化推荐的形式化定义以及国内外研究现状进行介绍。

4.1 面向群组的社会化推荐的形式化定义

目前,对于面向群组的社会化推荐还没有一个统一的形式化定义,普遍认可的定义是基于共识分数的。共识分数使用整个群组的偏好程度和群组内成员之间的差异性来衡量待推荐项目对群组的效用^[47]。因此,可以根据以上两个因素定义群组差异性和群组预测评分,并定义了共识函数,最后通过群组分歧度和群组预测评分来计算共识分数,具体的形式化定义如下。

定义 1(群组预测评分) 群组 G_g 对项目 I_i 的预测评分 $RG_{g,i}$ 可以由群组内所有用户的预测评分 $\overline{R}_{x,i}$ 通过融合策略计算得到。下面将以群推荐中最常用的融合策略均值策略为例,给出群组预测评分 $RG_{g,i}$ 的计算方法:

$$RG_{g,i} = \frac{1}{|G_g|} \sum_{U_x \in G_g} \overline{R}_{x,i} \quad (2)$$

定义 2(群组差异性) 群组差异性 $dis(G_g, I_i)$ 用来表示群组 G_g 中所有用户对项目 I_i 预测评分之间的差异程度。目前已有多种方法可以计算群组的差异性,最常用的是利用分歧方差计算,其中 $\overline{RG}_{g,i}$ 表示 G_g 中所有用户对项目 I_i 预测评分的均值:

$$dis(G_g, I_i) = \frac{1}{|G_g|} \sum_{U_x \in G_g} (\overline{R}_{x,i} - \overline{RG}_{g,i}) \quad (3)$$

定义 3(共识函数) 共识函数 $F(G_g, I_i)$ 的主要作用是通过群组分歧度和预测评分来计算共识分数,其中 ω_1 表示群组预测评分 $RG_{g,i}$ 在最后的共识分数中所占的权重:

$$F(G_g, I_i) = \omega_1 \times RG_{g,i} + (1 - \omega_1) \times (1 - dis(G_g, I_i)) \quad (4)$$

定义 4(TopN 推荐) 群推荐系统通常选择共识分数最高的前 N 个项目组成推荐列表 I_{G_g} 展示给群组,即 TopN 推荐。而推荐列表 I_{G_g} 需要满足下列条件:

$$\forall j \in I, \text{ s. t. } F(G_g, I_j) < F(G_g, I_i), I_j \notin I_{G_g}, I_i \in I_{G_g} \quad (5)$$

从上述群推荐的形式化定义可看出,面向群组的社会化推荐是建立在个体推荐的基础上的,首先需要确定在群推荐中使用什么样的个体推荐方法,并选择合适的融合方法,其次需要针对不同的应用场景选择合适的合成策略,最终才能产生满足群组偏好的推荐结果。因此,群推荐中融合方法和合成策略的选择对于面向群组的社会化推荐至关重要。

4.2 国内外的研究现状

目前,已有一些研究者针对社交媒体上面向群组的推荐问题展开了研究,根据生成群推荐的方法不同,已有的面向群组的社会化推荐方法大致可以分为两类:推荐方法融合和推荐结果融合^[50]。

推荐方法融合是先根据群组内成员的偏好或者最近邻居融合生成群组的偏好或者最近邻居,然后利用面向个体的推荐模型生成面向群组的推荐结果^[47]。根据使用的面向个体的推荐模型的不同,推荐方法融合又可以分为基于协同过滤的推荐方法融合和基于内容的推荐方法融合。因为协同过滤推荐方法主要分为生成用户偏好、产生用户最近邻居、计算用户对项目的预测评分和产生用户推荐列表 4 个阶段^[28,51],所以基于协同过滤的推荐方法融合主要发生在生成用户偏好和产生用户最近邻居阶段,而后两个阶段则属于推荐结果融合,在群推荐的推荐结果融合方法中再进行介绍^[51]。其中,在生成用户偏好阶段,基于协同过滤的推荐方法融合主要是利用融合策略对用户项目的评分、点击和收藏等用户偏好进行

融合,生成群组对项目的评分、点击和收藏等偏好信息,再依据群组偏好信息进行推荐^[52-53]。而在产生用户最近邻居阶段,基于协同过滤的推荐方法融合主要是根据群组中用户的偏好信息,先建立各自的最近邻居用户集合,接着再将各自的最近邻居集合进行融合得到群组的最近邻居集合,然后再根据群组的最近邻居集合进行推荐^[51-52]。而基于内容的推荐方法主要分为生成用户偏好、计算用户与项目的相似度和产生用户的项目推荐列表 3 个阶段^[5,54]。因此,基于内容的推荐方法融合主要发生在生成用户偏好阶段,而基于内容的推荐的后两个阶段属于推荐结果融合,在群推荐的推荐结果融合方法中再进行详细讨论。其中,针对生成用户偏好阶段,基于内容的推荐方法融合首先是将每个用户的兴趣偏好表示成一组关键词的向量,然后利用融合策略对个体用户偏好进行线性组合,产生一组关键词向量表示的群组偏好模型,再基于此进行群推荐^[54]。

与推荐方法融合不同,推荐结果融合则是先利用面向个体的推荐模型对每个群组成员生成推荐结果,然后利用合成策略将群组内所有用户的推荐结果融合,从而得到群组的推荐结果^[50-51]。因为推荐结果融合是融合群组内用户的预测评分或推荐项目列表,所以最终得到的是群组的预测评分或项目推荐列表。因此,对于评分预测和获得项目推荐列表两类问题,推荐结果融合又可分为两类:预测评分融合和推荐列表融合^[55-56]。其中,预测评分融合是先根据个体用户偏好模型生成个体用户对项目的预测评分,然后利用合成策略融合群组内成员的预测评分,从而得到群组对项目的预测评分^[49]。由于涉及到个体用户的预测评分问题,采用预测评分融合的面向群组的推荐系统可以与协同过滤方法结合,作用在协同过滤推荐方法的计算用户对项目的预测评分阶段;也可以与基于内容的推荐方法关联起来,作用在其计算用户与项目的相似度阶段。而推荐列表融合的思想是先单独对个体用户进行推荐,为每个群组成员生成推荐列表,然后选取合适的合成策略将群组内所有成员的项目推荐列表融合,从而得到群组推荐列表^[51-55]。但是,推荐列表融合通常会面临两个方面的问题:1)该方法容易推荐出与某个用户偏好很相似但群组内其他用户不感兴趣的项目集合;2)利用群组内用户的推荐列表进行融合容易导致推荐后的群推荐项目列表过大,不利于产生群组决策,因此为了缩减群推荐列表的长度,需要更加深入地提取群组偏好^[57-58]。

从上述两类面向群组的推荐方法可以看出,推荐方法融合和推荐结果融合两类群推荐方法中都需要使用融合策略,为此如何选择和设计有效的融合策略是影响群推荐质量的重要因素之一^[52,59]。目前,常用的融合策略大致可以分为 3 类:基本融合策略、加权策略和多种策略融合^[50]。其中,基本融合策略是由社会选择理论发展而来,主要包括均值策略、最大幸福策略、最小痛苦策略、公平策略、痛苦避免均值策略、多数票策略、连乘策略、Borda 计数策略、Copeland 策略和许可投票策略等^[47,52,60]。加权策略则是根据群组内成员的角色、特征和影响力等因素,在群推荐的过程中为群组内成员分配不同的权重^[59,61]。加权策略根据计算权重方式的不同又可分为静态加权策略和交互加权策略^[48]。其中,静态加权策略通常是利用一些启发式方法计算群组内成员的权重,而交互加权策略则是根据群组内用户与项目的交互行为来计算权重,例如若群组内某个用户的评分数量比其他用户多,则表明

其在群组内的表现较为活跃,从而应该为其赋予更大的权重^[61-62]。多种策略融合则是通过组合使用多种融合策略来避免单一融合策略的缺陷。例如,有的研究者通过启发式学习方法将最大幸福策略、最小痛苦策略和一般加权模型进行组合^[58]。有的研究者通过概率方法来构建群推荐模型,从而使该群推荐方法能够与多种融合策略相结合^[63]。

综上所述,面向群组的社会化推荐是一个具有重要价值的研究方面,对面向群组的社会化推荐方法的改进和应用都是十分重要的研究问题。

结束语 本文分别通过面向个体的社会化推荐与面向群组的社会化推荐两方面对社会化推荐的研究进行综述,目的是使人们更加系统地认识和理解社会化推荐及其发展现状。自 Facebook 等各大社交媒体兴起以来,社交媒体上产生了大量的用户和项目的社会化属性信息,使社会化推荐成为推荐系统领域的重要研究方向之一,社会化推荐正逐渐成为当前学术界和产业界共同关注的热点问题。目前,面向个体与面向群组的推荐从多个角度出发,将社会化信息融入到传统的推荐系统中,构建社会化推荐系统。其中,面向群组的社会化推荐研究尚处于起步阶段,还有大量可以改进的方法,充分挖掘群组内成员间的相关性,能够让面向群组推荐更加优化,同时也能使社会化推荐的研究更进一步。

参 考 文 献

- [1] CHEN H, CHIANG R H L, STOREY V C. Business intelligence and analytics: from big data to big impact[J]. *Mis Quarterly*, 2012, 36(4): 1165-1188.
- [2] 包红云, 李秋丹, 高珩, 等. 集成组内标签与用户链接关系的 Flickr 组推荐模型[J]. *模式识别与人工智能*, 2013(2): 176-181.
- [3] 中国互联网络信息中心. 第 34 次中国互联网络发展状况统计报告[J]. *互联网天地*, 2014(7): 71-89.
- [4] 项亮. 推荐系统实践[M]. 北京: 人民邮电出版社, 2012.
- [5] ADOMAVICIUS G, TUZHILIN A. Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions[J]. *IEEE Transactions on Knowledge & Data Engineering*, 2005, 17(6): 734-749.
- [6] RESNICK P, VARIAN H R. Recommender systems[J]. *Communications of the Acm*, 1997, 40(3): 56-58.
- [7] SCHAFFER J B, DAN F, HERLOCKER, et al. Collaborative Filtering Recommender Systems[M]// *The adaptive Web*. Berlin: Springer, 2007.
- [8] 孟祥武, 刘树栋, 张玉洁, 等. 社会化推荐系统研究[J]. *软件学报*, 2015, 26(6): 1356-1372.
- [9] JIANG M, CUI P, WANG F, et al. Scalable et al. Recommendation with Social Contextual Information[J]. *IEEE Transactions on Knowledge & Data Engineering*, 2014, 26(11): 2789-2802.
- [10] MA H, YANG H, LYU M R, et al. SoRec: Social Recommendation Using Probabilistic Matrix Factorization[C]// *ACM Conference on Information and Knowledge Management (CIKM 2008)*. Napa Valley, California, USA, 2008.
- [11] FIGUEIREDO F, PINTO H, BELÉM F, et al. Assessing the quality of textual features in social media[J]. *Information Processing & Management*, 2013, 49(1): 222-247.
- [12] MA H, ZHOU D, LIU C, et al. Recommender systems with social regularization[C]// *Forth International Conference on Web Search and Web Data Mining (WSDM 2011)*. Hong Kong, China, 2011.

- [13] RAWASHDEH M, AIHAMID M F, ALJAAM J M, et al. Tag-based personalized recommendation in social media services[J]. *Multimedia Tools & Applications*, 2016, 75(21):13299-13315.
- [14] BAI J. Collaborative filtering recommendation algorithm based on semantic similarity of item[C]// *IEEE Fifth International Conference on Advanced Computational Intelligence*. 2012.
- [15] LI Y M, HSIAO H W. Recommender Service for Social Network based Applications[C]// *Proceedings of the 11th International Conference on Electronic*. 2009.
- [16] MA H, ZHOU T C, LYU M R, et al. Improving Recommender Systems by Incorporating Social Contextual Information[J]. *Acm Transactions on Information Systems*, 2011, 29(2):219-230.
- [17] TANG J, HU X, LIU H. Social recommendation: a review[J]. *Social Network Analysis and Mining*, 2013, 3(4):1113-1133.
- [18] JANNACH D, GEYER W, DUGAN C, et al. Workshop on recommender systems and the social Web[J]. *Recommender Systems & the Social Web*, 2009:421-422.
- [19] FREYNE J, ANAND S S, GUY I, et al. 3rd workshop on recommender systems and the social Web[C]// *ACM Conference on Recommender Systems*. 2011.
- [20] MOBASHER B, JANNACH D, GEYER W, et al. The fifth ACM RecSys workshop on recommender systems and the social Web[C]// *Recsys*. 2013:60-70.
- [21] YANG X, GUO Y, LIU Y, et al. A survey of collaborative filtering based social recommender systems[J]. *Computer Communications*, 2014, 41(5):1-10.
- [22] ZANDA A, EIBE S, MENASALVAS E. SOMAR: A SOcial Mobile Activity Recommender[J]. *Expert Systems with Applications*, 2012, 39(9):8423-8429.
- [23] SYMEONIDIS P, PAPADIMITRIOU A, MANOLOPOULOS Y, et al. Geosocial recommendations based on incremental tensor reduction and local path traversal[C]// *ACM Sigspatial International Workshop on Location-Based Social Networks*. 2011.
- [24] LIU J, WU C, XIONG Y, et al. List-wise probabilistic matrix factorization for recommendation [J]. *Information Sciences*, 2014, 278:434-447.
- [25] CAI Y, LEUNG H, LI Q, et al. Typicality-Based Collaborative Filtering Recommendation[J]. *IEEE Transactions on Knowledge & Data Engineering*, 2014, 26(3):766-779.
- [26] ARORA G, KUMAR A, DEVER G S, et al. Movie Recommendation System Based on Users' Similarity [J]. *International Journal of Computer Science and Mobile Computing*, 2014, 3(4):765-770.
- [27] 孙光福, 吴乐, 刘淇, 等. 基于时序行为的协同过滤推荐算法[J]. *软件学报*, 2013(11):2721-2733.
- [28] 刘鲁, 任晓丽. 推荐系统研究进展及展望[J]. *信息系统学报*, 2008(1):82-90.
- [29] LI H. *Learning to Rank for Information Retrieval and Natural Language Processing* [M]. Morgan & Claypool Publishers, 2011.
- [30] SHI Y, LARSON M, HANJALIC A. List-wise learning to rank with matrix factorization for collaborative filtering[C]// *ACM Conference on Recommender Systems(Recsys 2010)*. Barcelona, Spain, 2010.
- [31] KOREN Y, BELL R, VOLINSKY C. Matrix Factorization Techniques for Recommender Systems[J]. *Computer*, 2009, 42(8):30-37.
- [32] CAO H, VERMA R, NENKOVA A. Speaker-sensitive emotion recognition via ranking: Studies on acted and spontaneous speech [J]. *Computer Speech & Language*, 2015, 28(1):186.
- [33] RENDLE S, FREUDENTHALER C, GANTNER Z, et al. BPR: Bayesian personalized ranking from implicit feedback[C]// *Conference on Uncertainty in Artificial Intelligence*. 2009.
- [34] SHI Y, LARSON M, HANJALIC A. Unifying rating-oriented and ranking-oriented collaborative filtering for improved recommendation[J]. *Information Sciences*, 2013, 229(6):29-39.
- [35] ZHAO Z, LU H, CAI D, et al. User Preference Learning for Online Social Recommendation[J]. *IEEE Transactions on Knowledge & Data Engineering*, 2016, 28(9):1.
- [36] GOLBECK J. Generating Predictive Movie Recommendations from Trust in Social Networks[M]// *Trust Management*. Berlin: Springer, 2006:93-104.
- [37] MASSA P, AVESANI P. Trust-aware recommender systems [C]// *ACM Conference on Recommender Systems*. 2007.
- [38] JAMALI M, ESTER M. TrustWalker: a random walk model for combining trust-based and item-based recommendation [C]// *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. 2009.
- [39] MA H, KING I, LYU M R. Learning to recommend with social trust ensemble[C]// *International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2009*. Boston, MA, USA, 2009.
- [40] TANG J, GAO H, LIU H. mTrust: Discerning Multi-Faceted Trust in a Connected World[C]// *International Conference on Web Search and Web Data Mining(WSDM 2012)*. Seattle, WA, USA, 2012.
- [41] JAMALI M, ESTER M. A matrix factorization technique with trust propagation for recommendation in social networks[C]// *ACM Conference on Recommender Systems(Recsys 2010)*. Barcelona, Spain, 2010.
- [42] CAGLIERO L, FIORI A, GRIMAUDO L. Personalized tag recommendation based on generalized rules[J]. *Acm Transactions on Intelligent Systems & Technology*, 2013, 5(1):328-328.
- [43] LI Q, WANG J, CHEN Y P, et al. User comments for news recommendation in forum-based social media[J]. *Information Sciences*, 2010, 180(24):4929-4939.
- [44] JIANG B. Improving collaborative tag recommendation by using local lexicon in social comment context[C]// *International Conference on Computer Supported Cooperative Work in Design (CSCWD 2011)*. Lausanne, Switzerland, 2011.
- [45] MAO J, LU K, LI G, et al. Profiling users with tag networks in diffusion-based personalized recommendation[J]. *Journal of Information Science*, 2016, 42(5):711-722.
- [46] LEI X, QIAN X, ZHAO G. Rating Prediction based on Social Sentiment from Textual Reviews [J]. *IEEE Transactions on Multimedia*, 2016, 18(9):1.
- [47] 何军. 国外群推荐聚集策略研究综述[J]. *图书情报工作*, 2013(7):127-133, 88.
- [48] WANG W, ZHANG G Q, LU J. Member contribution-based group recommender system [J]. *Decision Support Systems*, 2016, 87:80-93.
- [49] ORTEGA F, et al. Recommending items to group of users using Matrix Factorization based Collaborative Filtering[J]. *Information Sciences*, 2016, 345(C):313-324.

- [J]. 中国集体经济, 2016(4):61-62.
- [2] 孙晔, 杨照东, 陈德华, 等. 大数据用户画像技术在商业银行的应用[J]. 数字通信世界, 2016(9):86-88.
- [3] 刘海, 卢慧, 阮金花, 等. 基于“用户画像”挖掘的精准营销细分模型研究[J]. 丝绸, 2015, 52(12):37-42, 47.
- [4] 陈志明, 胡震云. UGC网站用户画像研究[J]. 计算机系统应用, 2017, 26(1):24-30.
- [5] 吕辉, 许道强, 仲春林, 等. 基于电力大数据的标签画像技术与应用研究[J]. 电力信息与通信技术, 2017, 15(2):43-48.
- [6] 陈一平. 今品指数:用大数据为企业画像[J]. 华夏时报, 2015-07-06(021).
- [7] 王磊, 郑国勇. 大连国税:大数据为企业精准画像[N]. 中国税务报, 2016-10-17(A04).
- [8] 王雪. 基于企业画像的公示数据查询系统设计与实现[D]. 大连:大连海事大学, 2017.
- [9] ZHONG C J. “Picture” for domestic enterprises to go to the western “circle”[N]. NanJing Times, 2016-04-25(A02).
- [10] 李代伟, 谢丽艳, 钱慎一, 等. 基于 Scrapy 的分布式爬虫系统的设计与实现[J]. 湖北民族学院学报(自然科学版), 2017, 35(3):317-322.
- [11] 胡萍瑞, 李石君. 基于 URL 模式集的主题爬虫[J]. 计算机应用研究, 2018(3):1-2.
- [12] GOUNARIS A, TORRES J. A Methodology for Spark Parameter Tuning[J]. Big Data Research, 2018, 11:22-32.
- [13] 李映坤. 大数据背景下用户画像的统计方法实践研究[D]. 北京:首都经济贸易大学, 2016.
- [14] DING W, WANG T, LIU X H, et al. Research on mobile phone user’s portrait and credit based on large data technology[J]. DesigTechn Post Telecommun, 2016(3):64-69.
- [15] 刘鹏. 基于 Spark 的数据管理平台的设计与实现[D]. 杭州:浙江大学, 2016.
- [16] 徐凯. 基于主题抽取演化模型的实时新闻推荐系统[D]. 广州:广东工业大学, 2016.
- [17] SEVERYN A, MOSCHITTI A. Learning to rank short text pairs with convolutional deep neural networks[C]// Proceedings of the 38th International ACM SIGIR Conference. 2015:373-382.
- [18] 余本功, 张连彬. 基于 CP-CNN 的中文短文本分类研究[J]. 计算机应用研究, 2018, 35(4).
- [19] 李春梅. 基于 TF-IDF 的网页新闻分类的研究与应用[J]. 贵州师范大学学报(自然科学版), 2015, 33(6):106-109.
- [20] 谢慧志. 基于用户选购及在线点评行为的门店画像研究[D]. 广州:华南理工大学, 2016.
- [21] GLOROT X, BORDES A, BENGION Y. Deep spares rectifier neural network[C]// Proceedings of International Conference on Artificial Intelligence and Statistic. 2011:315-323.
- [22] KIM Y. Convolutional Neural Networks for Sentence Classification[C]// Proceedings of Empirical Methods in Natural Language Processing. 2014:1746-1751.
- [23] 张燕琴, 潘利强. Apriori 算法在 WEB 的图书销售智能系统中的应用[J]. 网络安全技术与应用, 2017(9):58-59.
- [24] 刘志刚. 基于标签关联规则的挖掘与研究[J]. 科技创新与应用, 2017(26):26.
- [25] 左爱文. 数据挖掘技术在气象数据中的应用[D]. 西安:西安电子科技大学, 2006.
- [26] 张玲玲, 周全亮, 唐广文, 等. 基于领域知识和聚类的关联规则深层知识发现研究[J]. 中国管理科学, 2015, 23(2):154-161.
- [27] 夏光虎, 贾宇波, 范红丹. 分类与预测挖掘在信用风险评估中的应用研究[J]. 工业控制计算机, 2012, 25(7):71-72.
- [28] SZEGED Y C, LIU W J, YANG Q. Going deeper With Convolutions[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition(CVPR). Boston, USA, 2015.
- [29] JOHNSON R, ZHANG T. Effective Use of Word Order for Text Categorization with Convolutional Neural Networks[C]// Proceedings of North American Chapter of the Association for Computational Linguistics. 2014:103-112.

(上接第 42 页)

- [50] 张玉洁, 杜雨露, 孟祥武. 组推荐系统及其应用研究[J]. 计算机学报, 2016(4):745-764.
- [51] ORTEGA F, BOBADILLA J, HERNANDO A, et al. Incorporating group recommendations to recommender systems: Alternatives and performance[J]. Information Processing & Management, 2013, 49(4):895-901.
- [52] MASTHOFF J. Group Recommender Systems[C]// Combining Individual Models. 2010:677-702.
- [53] KASSAK O, KOMPAN M, BIELIKOVA M. Personalized hybrid recommendation for group of users; Top-N multimedia recommender[J]. Information Processing & Management, 2016, 52(3):459-477.
- [54] PERA M S, NG Y K. A group recommender for movies based on content similarity and popularity[J]. Information Processing & Management An International Journal, 2013, 49(3):673-687.
- [55] BALTRUNAS L, MAKCINSKAS T, RICCI F. Group recommendations with rank aggregation and collaborative filtering[C]// ACM Conference on Recommender Systems(Recsys 2010). Barcelona, Spain, 2010.
- [56] GARCIA I, SEBASTIA L, ONAINDIA E. On the design of individual and group recommender systems for tourism[J]. Expert Systems with Applications, 2011, 38(6):7683-7692.
- [57] NAAMANI-DERY L, KALECH M, ROKACH L, et al. Preference elicitation for narrowing the recommended list for groups[C]// ACM Conference on Recommender Systems. 2014.
- [58] JAMESON A. More than the sum of its members; challenges for group recommender systems[C]// Working Conference on Advanced Visual Interfaces(AVI 2004). Gallipoli, Italy, 2004.
- [59] QUIJANO-SANCHEZ L, RECIO-GARCIA J A, DIAZ-AGUDO B, et al. Social factors in group recommender systems[J]. Acm Transactions on Intelligent Systems & Technology, 2013, 4(1):1199-1221.
- [60] QUEIROZ S R M, DE CARVALHO F A T, RAMALHO G L, et al. Making recommendations for groups using collaborative filtering and fuzzy majority[C]// 16th Brazilian Symposium on Artificial Intelligence(SBIA 2002). Bittencourt G, Ramalho G L, 2002:248-58.
- [61] GUO J, ZHU Y, LI A, et al. A Social Influence Approach for Group User Modeling in Group Recommendation Systems[J]. IEEE Intelligent Systems, 2016, 31(5):1.
- [62] ARDISSONO L, GOY A, PETRONE G, et al. Intrigue: Personalized recommendation of tourist attractions for desktop and hand held devices[J]. Applied Artificial Intelligence, 2003, 17(8-9):37-41.
- [63] GORLA J, LATHIA N, ROBERTSON S, et al. Probabilistic group recommendation via information matching[C]// International Conference on World Wide Web. 2013.