

DOI:10.3969/j.issn.1673-4785.201403068  
网络出版地址: <http://www.cnki.net/kcms/doi/10.3969/j.issn.16734785.201403068.html>

# 社交媒体多模态、多层次资源推荐技术研究

王大玲<sup>1,2</sup>, 冯时<sup>1,2</sup>, 张一飞<sup>1,2</sup>, 于戈<sup>1,2</sup>

(东北大学 信息科学与工程学院, 辽宁 沈阳 110819; 2. 东北大学 医学影像计算教育部重点实验室, 辽宁 沈阳 110819)

**摘要:** 社交媒体中多模态和多层次的信息资源和基于各种关系构建的用户社群为推荐系统提供了更广阔的分析 and 选择空间, 同时也带来了更多的问题与挑战。分析了当前社交媒体中用户与资源的关系以及社交媒体资源推荐的特点, 分别从社交媒体资源推荐策略和相关支撑技术两方面综述了相关工作, 将其概括为“社交媒体中用户角色的变化构成了更加复杂的用户关系”、“社交媒体资源表示形式呈现多模态特点”以及“社交媒体资源推荐应该满足多层次的用户需求”, 并从多模态、多层次资源推荐方面提出进一步的研究方向。

**关键词:** 社交媒体; 推荐; 多模态信息; 多层次资源; 用户社群

**中图分类号:** TP301 **文献标志码:** A **文章编号:** 1673-4785(2014)03-0265-11

中文引用格式: 王大玲, 冯时, 张一飞, 等. 社交媒体多模态、多层次资源推荐技术研究[J]. 智能系统学报, 2014, 9(3): 265-275.  
英文引用格式: WANG Daling, FENG Shi, ZHANG Yifei, et al. Study on the recommendations of multi-modal and multi-level resources in social media[J]. CAAI Transactions on Intelligent Systems, 2014, 9(3): 265-275.

## Study on the recommendations of multi-modal and multi-level resources in social media

WANG Daling<sup>1,2</sup>, FENG Shi<sup>1,2</sup>, ZHANG Yifei<sup>1,2</sup>, YU Ge<sup>1,2</sup>

(1. School of Information Science and Engineering, Northeastern University, Shenyang 110819, China; 2. Key Laboratory of Medical Image Computing (Northeastern University), Ministry of Education, Shenyang 110819, China)

**Abstract:** The multi-modal and multi-level information resources and user communities based on various relationships in social media provide a broader space for recommenders to analyze and select the resources, but at the same time more problems and challenges develop. In this paper, the relationships between users and resources, and the characteristics of resource recommendations in current social media are analyzed. Related work is surveyed from social media resource recommendations as well as its corresponding support techniques, which can be summarized as follows: more complex relationships among users formed by changes of user roles, multi-modal social media resource expressions, and social media resource recommendations that satisfy the requirements of multi-level users. Finally, further research directions in multi-modal and multi-level resource recommendations are proposed.

**Keywords:** social media; recommendation; multi-modal information; multi-level resource; user community

“推荐”是系统为用户提供的一种主动的信息推送方式, 它不同于搜索那样具有直接而明确的用户需求, 因此涉及用户兴趣分析与推断、推荐信息资源选择 2 个主要部分。在用户兴趣推断方面, 传统的推荐技术分为“基于内容的推荐”、“基于协同过滤的推荐”及其混合技术<sup>[1]</sup>。近年来, Web 2.0 技术的发展促进了社交媒体 (social media) 这一新技术

的诞生和迅速发展。与传统的媒体相比, 社交媒体中用户角色的变化、用户关系的形成、特别是信息资源形式的丰富, 均对推荐系统提出了新的要求。基于社交媒体的推荐中, 无论用户兴趣的推断、还是推荐信息资源的选择, 均被赋予了新的内容。本文将首先分析社交媒体中的用户与资源的关系, 以及由此引出的社交媒体资源推荐的特点, 综述当前该领域研究的相关工作并进行讨论。在此基础上, 从多模态、多层次资源推荐方面提出进一步的研究方向。

## 1 社交媒体用户、资源及推荐的特点

### 1.1 社交媒体中的用户与资源

社交媒体作为一类新的在线交互平台,允许用户自发地撰写、传播和获取信息,具有广泛的用户参与性。本文中,将“资源”定义为用户在社交媒体中发布的原始信息单元及其组合或划分。与传统媒体相比,社交媒体中的“用户”和“资源”呈现出许多新的特点。

首先,从社交媒体用户的角度,呈现以下特点:

#### 1) 用户角色双重性。

社交媒体中信息的传播是“众对众”方式,用户不仅是信息的接收者,也是信息的发布者,发布时不仅提供信息,而且在社交网站上建立详细的个人档案并分享这些信息<sup>[2]</sup>。

#### 2) 社会网络社群性。

社交媒体的用户(包括信息的发布者和接收者)在社交媒体中通过信息交流可以表达关注同一事件或人、具有相同观点以及结为好友等,这些关系构成各种网络社群。

#### 3) 信息网络异构性。

将用户与资源作为对象并相互链接起来,即构成信息网络<sup>[3]</sup>。在此情况下,社交媒体中的信息网络节点是异构的,表现为用户与信息是不同类型的对象,而且不同媒体形式的信息内容也是异构的。

#### 4) 用户关系多元性。

前面1)中所述的用户角色双重性表现为信息网络中的显式链接即为“用户节点→资源节点”、或者相反。此外,除显式链接外,还存在更多的隐式链接关系,如信息之间的相关性链接以及由此而导致的用户之间的相关性链接等,构成了多元化的用户关系。

第二,从社交媒体资源的角度,呈现以下特点:

#### 1) 信息的多模态性。

在相关研究中,“模态”表现为不同的媒体形式(如文本、图像、视频等)<sup>[4]</sup>,又可表现为同一媒体形式的不同特征(如图像的颜色、纹理等特征)<sup>[5]</sup>。在本文中,从推荐的角度,网络社群作为一种资源,其本身也是一种模态。

#### 2) 资源的多层次性。

由于社交媒体允许用户自发地传播和获取信息,因此导致了社交媒体资源的多层次性。例如,在新浪博客中,用户A撰写了一篇关于三亚旅游的图文并茂的博文,用户B转发了其中的部分文字并加上了评论,用户C又配发一些照片,……,这些博文、

评论、图片等信息构成了同一资源的多个层次。同时,社交媒体信息之间由于背景、事件、位置、人物等实体的联系,同样会构成资源的不同层次。

### 1.2 社交媒体推荐的特点

如前所述,推荐系统涉及用户兴趣的分析和推荐资源的选择2个主要部分。由于社交媒体中用户和资源新的特点,因而对基于社交媒体的推荐系统提出了新的要求,从而导致了社交媒体推荐呈现出新的特点:

#### 1) 多模态的用户兴趣内容分析。

在分析用户兴趣方面,传统的推荐技术之一是“基于内容的推荐”。即:对于一个用户,推荐系统根据该用户自己对资源的访问历史推断其兴趣。在社交媒体推荐中,“基于内容的推荐”同样适用,但却赋予了新的要求。首先,用户对资源的访问历史记录中,其内容可能包括不同的媒体形式;其次,对于某一种媒体形式(特别是非文本形式),用户可能也会关注不同的方面。例如,对于用户访问历史中的一幅“裙子”的图片,用户兴趣可能是裙子的颜色、质地、图案、款式等不同方面,而非这条裙子本身。用户访问历史内容中不同的媒体形式和同一媒体形式的不同特征,构成了多模态的用户兴趣。

#### 2) 基于社会关系的协同过滤。

传统推荐技术中,用户兴趣分析的另一主要方法是“基于协同过滤的推荐”。即:对于一个用户,推荐系统根据与该用户兴趣相似的其他相关用户对资源的访问情况推断其兴趣,所依据的是“用户-项目”点击矩阵的相似性度量。社交媒体中,前述的“网络社群性”使这种依据可以扩展为基于用户社会关系的推荐,从而解决“点击矩阵稀疏”等问题,而“用户角色双重性”、“信息网络异构性”及“用户关系多元性”等用户特点使这一问题变成异构信息网络中的社群发现及社群的用户兴趣推断问题。

#### 3) 面向用户的多层次资源推荐。

传统的推荐系统中,无论“基于内容的推荐”、“基于协同过滤的推荐”、还是其混合技术,推荐结果的粒度大多是原始信息,即便是多媒体推荐,也较少考虑用户不同粒度、不同层次的需求。这里重提前述“社交媒体资源多层次性”中的例子,即:用户A撰写了一篇关于三亚旅游的图文并茂的博文,用户B转发了其中的部分文字并加上了评论,用户C又配发了一些照片,……。在推荐时,通过对这些博文、评论、图片等原始信息的分析,可以推荐:1)与博文中的图片或某部分文字相关的信息;2)由多幅图片和文字构成的三亚某一景点的介绍信息;3)诸如

“三亚旅游最佳路线”这样的综合信息。这里,推荐1)是原始信息的子集,推荐2)是若干原始信息基于某种方式的组合,推荐3)则是原始信息基础上高度的分析、处理、总结和重组。上述推荐项目的不同粒度表明了从原始信息的子集、原始信息、到它们某种形式的组合、再到更高层次的综合这样不同的资源层次。

综上,对社交媒体推荐中涉及的“资源”和“用户”给出如下定义:

**定义1** 社交媒体用户。社交媒体资源的发布者和使用者(一般需要注册),包括上传、转载、共享、浏览、关注等一切对社交媒体资源进行过操作的用户。

**定义2** 社交媒体资源。即用户在社会媒体中发布的原始信息单元及其组合或抽象。原始信息单

元如一篇博文、一条微博、一幅图片、一段视频、一首乐曲等,称之为单一资源;单一资源中的某个部分称为子资源,如微博中的图片、视频中的音乐等;单一资源的有序组合称为复合资源,如不同用户上传的关于同一景点的多幅图片及若干文本;通过对单一、复合资源的分析和挖掘,可望得到综合资源,例如,对于某个景点,对用户上传的图片、撰写的博文、发布的微博、微信等进行深入挖掘可构成“旅游综合资源”;针对某个学术问题,对相关社群用户的讨论、发表的文章、乃至该社群本身的挖掘可构成“学术综合资源”。由于用户社群因与资源的关系而形成,并通过分析和挖掘而发现,因此,用户社群也是一种综合资源。

根据定义2,本文提出的社交媒体推荐中的多模态、多层次资源如图1所示。

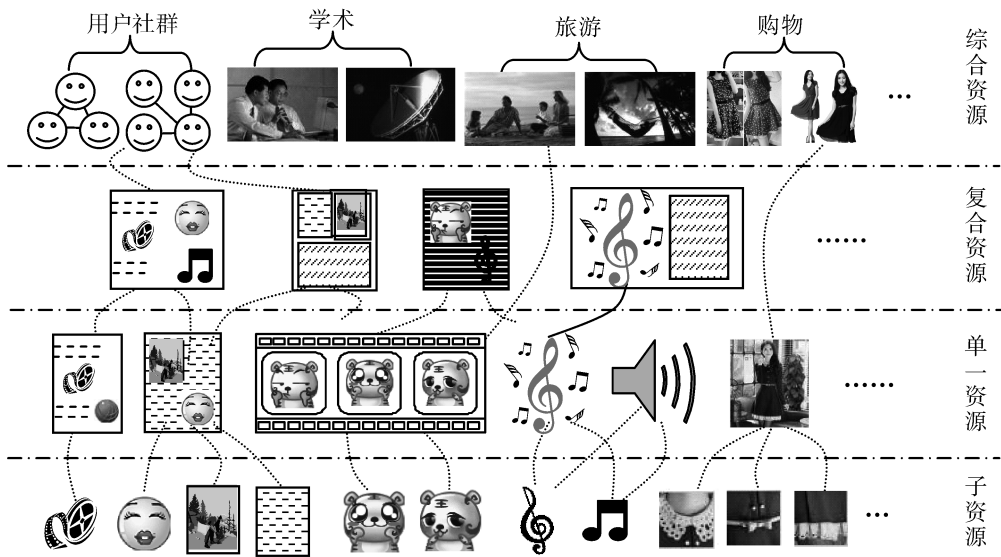


图1 社交媒体推荐中的多模态、多层次资源

Fig.1 Multi-modal and multi-level resource in social media for recommendation

基于社交媒体平台,在用户与相关信息构成的异构信息网络中挖掘用户社群并进而推断用户兴趣,在推荐内容上实施不同粒度、分层次、高质量的信息资源推荐,是用户的潜在需求,更是推荐系统应该、并且通过应用相关技术可望实现的功能。

当前社交媒体资源推荐的特点,实际上可以视为传统的推荐技术被赋予了社交媒体新的内容。在目前的相关工作中,作为最终目标的“面向用户的多层次资源推荐”主要表现为基于用户社群、媒体内容以及多种因素的推荐,而作为用户兴趣分析方法的“多模态的用户兴趣内容分析”和“基于社会关系的协同过滤”则需要社交媒体用户关系分析、社交媒体多模态信息挖掘以及社交媒体中多模态信息相似性度量等相关技术的支撑。文中对上述涉及的

内容进行归类,其结果如图2所示(图中虚线意为当前工作并未完全具有的功能)。基于图2,下文分别从社交媒体资源推荐策略和社交媒体推荐的支撑技术两方面对相关工作进行综述,并在第4节进行总结和讨论。

## 2 社交媒体资源推荐策略

根据图2的归纳,将从基于用户社群的推荐、基于媒体内容的推荐以及基于多种因素的推荐3方面综述社交媒体资源推荐的相关工作。

### 2.1 基于用户社群的推荐

用户的网络社群性是社交媒体的主要特点之一,挖掘社交媒体中的社群关系、并基于此进行推荐,是社交媒体推荐的一个主要途径。



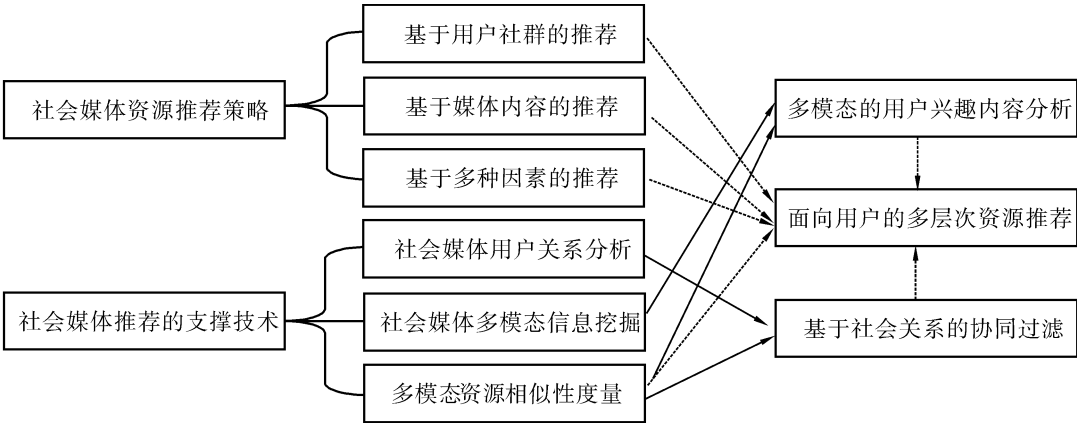


图 2 社交媒体资源推荐的特点与相关工作

Fig.2 Characteristics and related work of social media resources recommendation

在基于用户社群或社会关系的推荐方面,Pavlidis 等<sup>[6]</sup>基于用户个人信息和朋友信息等社交媒体信息推断用户的兴趣、用户朋友的兴趣以及朋友的亲密度,找到推荐礼物的重要时机,进行礼物推荐;Popescu 等<sup>[7]</sup>把个性化旅游推荐视为协同过滤问题,挖掘在网上用户的旅行数据、记录数据,并利用这些数据构建一个用户-用户的相似度矩阵,为一个想去某地旅游的用户推荐景点;Schirru<sup>[8]</sup>基于用户在 Web2.0 资源分享平台上与其同事交换的感兴趣话题的内容,建立一个基于话题的推荐系统,利用话题追踪检测算法,检测用户短期或长期感兴趣话题,提供基于话题的推荐,满足用户的需求和偏好;贾大文等<sup>[9]</sup>把用户对具体媒体对象的偏好转化为用户对媒体对象所蕴含兴趣元素的偏好,将具有相同偏好的用户聚合成“共同偏好组”,基于共同偏好组进行社会媒体的共享与推荐;Koohborfardhaghighi 等<sup>[10]</sup>在朋友网中向处在转接点位置和具有高特征向量中心值的用户实施推荐,认为前者具有重要的信息,后者与其他用户具有更多的联系或者被重要的用户所联系;Li 等<sup>[11]</sup>通过发现面向用户和面向社群的话题来捕获用户兴趣和社群焦点,从而推荐有影响力的用户和相关于话题的社群;Zhao 等<sup>[12]</sup>在 Twitte 风格的社会网络中根据关注与被关注关系发现社群,并基于社群进行粉丝推荐,从而改进基于内容推荐的低准确率和协同过滤推荐的稀疏性问题。

在此领域,与上述工作相似的还有很多,这里不多赘述。总体上,这些工作的主要特点是基于用户的社会关系及由于共同话题、关注点等原因形成的社群,发现当前用户的兴趣偏好,从而进行推荐。推荐的内容包括社会媒体的原始信息<sup>[6-8]</sup>,也包括用户的相关信息,如社群<sup>[9,11]</sup>、朋友<sup>[10,12]</sup>等。

2.2 基于媒体内容的推荐

无论是传统的推荐系统还是社交媒体资源的推

荐,基于内容的推荐均是一个重要推荐技术。不同的是,社交媒体中的内容已呈现多模态的信息表现形式,因此,用户的访问记录信息将由多模态的信息构成,推荐的内容也将是多模态、甚至是跨媒体(即基于一种媒体形式的访问历史推荐其他媒体形式的资源)的。

在基于媒体内容的推荐中,Zhang 等<sup>[13]</sup>、Bu 等<sup>[14]</sup>结合多种类型的社交媒体信息以及音乐原声信号的多资源媒体信息,用超图对社交媒体信息高阶关系建模,利用各种多媒体信息数据以及基于音乐原声的内容进行音乐推荐;Tan 等<sup>[15]</sup>采用图模型对评论间的关系以及评论与原始新闻间的关系进行建模,捕捉用户关注点的动态变化,抽取话题模式,综合考虑读者和作者的观点,进行相关信息推荐;Hu 等<sup>[16]</sup>将在线社交网站 OSNs 和视频分享网站 VSSes 的度量标准结合起来,提出了一种基于相似度的方法来为 Youtube 这样的视频分享网站 VSSes 增加推荐视频列表;Ma 等<sup>[17]</sup>将用户在社会媒体中发表和查看的数据视为社会流,提出了一个新的在线协同过滤框架和流排名矩阵分解,通过分析社会流来进行个性化主题推荐及发现;Ernesto 等<sup>[18]</sup>结合用户评论、Li 等<sup>[19]</sup>使用用户产生的评论作为数据源,建模评论之间的关系,进行新闻推荐。

与上述工作相似的工作还有很多,这里不多赘述。就社交媒体资源推荐而言,除了新闻、评论等文本形式的社交媒体资源外,其他媒体形式的推荐主要还是如传统多媒体推荐方法那样以资源的原始粒度作为推荐单元,而缺少更小或更大粒度、特别是综合资源的推荐。

2.3 基于多种因素的推荐

除前述基于用户关系和资源内容分析的推荐外,还有一些基于社交媒体多因素分析的推荐。

在基于多因素的推荐中,Messenger 等<sup>[20]</sup>基于

用户在社交网络中与其他成员的关联关系、社会数据以及词相关因素进行个性化书籍推荐;Pera等<sup>[21]</sup>研究基于用户和标签的社会媒体推荐,表明将两者结合起来的性能会更为优异;Guy等<sup>[22]</sup>利用包括帖子内容、帖子间相似度、帖子间或者其与博客间的链接以及与外部网站的链接在内的多种特征训练了一个SVM排序分类器,进行用户预测和链接预测,应用预测结果进行推荐;Wu等<sup>[23]</sup>基于旅行者使用搜索引擎查询旅游计划的上下文信息,生成一些有代表性的旅游目的地;Xiang<sup>[24]</sup>构建了相同类型的实体(如用户与用户)和不同类型的实体(如用户与图片)之间的关系图,将推荐任务描述为一个在关系图上的随机游走过程;Chidlovskii等<sup>[25]</sup>挖掘Tweet内容,并用其哈希标签和内容对youtube、Flickr、Photobucket、Dailymotion和SoundCloud进行标注,实现对大量的未被充分标注的用户生成的Web资源进行语义上的标注,以此实现标签推荐;Liu等<sup>[26]</sup>获取社交媒体中各种女性面部和发式图像,提取其中化妆品特征,构建一个化妆推荐系统,提出一个多树结构的超图模型来探索各种化妆面部图像的高层美丽属性、中层关于美丽的属性以及低层图像特征,基于用户输入的短发、素颜正脸图像,为其推荐最合适的发式和化妆模式,并展示合成结果。与上文论述的工作相比,本节中大部分推荐是基于社会媒体的多种因素,包括用户关系和媒体相关内容,但其中的“内容”主要是以标签、链接等元数据或语义信息为主。此外,日常生活中经常接触的如“淘宝”、“当当”这类购物网站,均具有根据用户访问历史、好友关系等因素推荐商品的功能,但这种推荐也主要基于标签等语义、文本信息的分析,因此可能出现这样的例子:对于一个用户,如果本人或其好友关注一条裙子的图片,推荐的也往往就是各种裙子图片,而不去考虑该用户到底是喜欢裙子的款式、花色、进而据此进行推荐。这个例子表明当前的研究在针对不同模态和层次细节内容方面的推荐受到一定限制。

### 3 社交媒体推荐的支撑技术

作为用户兴趣分析方法的“多模态的用户兴趣内容分析”和“基于社会关系的协同过滤”需要许多相关技术的支撑。根据图2,这些相关技术主要包括社交媒体用户关系分析、社交媒体多模态信息挖掘、以及社交媒体中多模态资源相似性度量等,本节将从这3个方面综述相关工作。

#### 3.1 社交媒体用户关系分析

社交媒体中的用户关系分析和用户社群发现是

社会网络研究中的一个重要组成部分。在社交媒体资源推荐中,由于用户对资源的各种操作而表现出的“兴趣相投”、“观点相似”、“关注”、“跟随”等关系构成了各种网络社群,这些社群所表达的用户之间的关系成为“基于协同过滤推荐”的一个有效的扩展。

在用户关系分析和用户社群发现研究方面,Sun等<sup>[27]</sup>在多类型对象构成的异构信息网络中(如学术网络中的术语、作者、文章等),提出基于元路径与用户指导聚类的方法对社会网络中的对象进行聚类;Guv等<sup>[28]</sup>根据“认识同一个人或者被同一个人认识”、“对同样的事情感兴趣”、“都在相同的地点”发现社交媒体中的用户关系;Barbier等<sup>[29]</sup>基于社交媒体存在大量群、数据有噪音、群不断变化的问题,提出发现群、根据群建立用户profile、理解群对其他个人或群的影响、探索跨社交媒体群以补充群成员评价的社会媒体在线群的理解;Yang等<sup>[30]</sup>针对电子商务社会网络中“许多具有相似兴趣的社会媒体用户可能没有直接的交互或购买相同商品、从而导致社交媒体中显式链接稀疏”的问题,提出识别用户之间这种隐式关系来丰富社交媒体结构的方法;Jorge等<sup>[31]</sup>认为链接预测对于社会网络成员未来关系的发现至关重要,针对此前的“链接预测基于结构信息、而结构信息在大规模社会网络中不足以获得好的预测性能”的问题,研究将“Twitter”用户的行为、兴趣等非结构信息引入用户社群分析,改进链接预测性能。

总体上,社会网络挖掘中的社群发现与推荐均属此范畴,还有很多的研究工作,这里不多赘述。上述工作中,有只考虑用户关系或者文档相似关系的同构网络,也有考虑用户、资源及其关系的异构网络,链接关系包括显式链接,也包括隐式链接,其最终目标还是发现各种用户社群。

#### 3.2 社交媒体多模态信息挖掘

社交媒体内容丰富,为推荐系统提供了巨大的信息资源。但无论是用户兴趣分析、还是推荐资源的选择,以及社交媒体信息具有的海量性、有噪音等特点,均需要对社交媒体内容(包括不同层次、粒度的内容)进行分析。由于当前社交媒体信息呈现多模态特征,因此多模态信息挖掘成为这种分析的主要技术手段。

在社交媒体多模态信息挖掘以及相关的信息检索方面,Tang等<sup>[32]</sup>提出了社交媒体分析中去除噪音、多媒体数据及用户的建模、标签代表性与资源相关性、大规模社交媒体数据的挖掘与检索的扩展性、

数据的衡量标准等问题;Lee<sup>[33]</sup>构建了在线突发事件发展评价模型,提出了离线事件评价衡量指标,用以支持事件间隐性关系的分析;Jin 等<sup>[34]</sup>提出社交媒体中的垃圾信息检测问题,通过图像、文本、社会网络中用户历史及其行为等特征的提取和分析,以识别垃圾内容;Lee 等<sup>[35]</sup>通过 Facebook、Twitter、DocTrack、MemRecap 等文本集和查询集等社交数据的测试,研究社交媒体个性化搜索和评价问题;Choudhury 等<sup>[36]</sup>提出一种对信息单元加权的维度表示,来描述大规模社交媒体空间,并通过一种抽样方法来降低这种大型社交网络空间的维度;Tsai 等<sup>[37]</sup>应用社交媒体并根据出现概率最大的话题对其进行分类,进而构建分类框架;Kaschesky 等<sup>[38]</sup>提出在社交媒体中进行观点挖掘的过程,包括:社交媒体内容分析,意见层建模、模拟和预测,社交网络的可视化、接口设计,系统架构和实时运行平台;Jin 等<sup>[39]</sup>提出了一个 LikeMiner 系统以便在社交媒体网路中挖掘“like”的内容,针对社交媒体引入“like”异构网络模型,建立可视和文本话题空间,提出挖掘算法评价社会媒体的表现力和影响力;Cui 等<sup>[40]</sup>提出社交媒体搜索中融合多特征及其相关性的方法进行相似性度量,其特征涉及正文特征、视觉内容特征和用户特征,分别以这些特征为节点构建特征交互图,并考虑同模态节点边和跨模态节点边的定义和构建。

社交媒体多模态信息挖掘与传统的多媒体挖掘技术相关,因此还有很多研究成果,这里不多赘述。就支持社交媒体资源推荐而言,由于社交媒体中用户发布、浏览信息时可能会为其赋予具有一定语义信息的标签,众包<sup>[41]</sup>、分类分众<sup>[42]</sup>技术又使标签进一步丰富和规范,加之用户关系等线索,目前很多推荐策略是基于标签和用户关系而避开内容(特别是非文本内容)分析的。但是,针对前述“面向用户的多层次资源推荐”的特点和需求,不同粒度的多模态内容分析和信息挖掘是必不可少的。

### 3.3 多模态资源相似性度量

第 2 节所论述的各种社交媒体推荐策略中,均需要根据用户兴趣(通过用户社群、媒体内容及其他各种因素等渠道获得)找到最相关的推荐资源,相似性度量是必不可少的。此外,前文论述的相关支撑技术中,分析用户的相关或相似性,特别是多模态信息挖掘的很多算法,更需要进行资源之间、特别是不同模态形式的资源之间的相似性度量。

在多模态资源相似性度量方面,Liu 等<sup>[43]</sup>研究社交媒体用户所构成的社会网络中一个节点与对等节点连接方式的多样性问题,提出捕获多样性语义

的度量标准,在社交媒体网站中获得各种类型的朋友、合作等关系;Ling 等<sup>[44]</sup>提出一种支持跨媒体信息检索的异构媒体对象的相似性度量方法,该方法探索一种结合了原始的低层特征空间和第三公共空间特点的 tri 空间,基于该空间进行不同媒体对象的相似性度量;Zhai 等<sup>[45]</sup>提出一种使用最近邻方法的异构相似性度量方法,通过计算 2 个不同媒体对象属于同一语义类别的概率来获得这两个媒体对象之间的相似性;Jia 等<sup>[46]</sup>构建一个话题模型的 Markov 随机域,将文本和与其松散相关的图像结合起来,对这两种不同模态之间的联系进行编码,用以进行跨模态的话题学习;Zhu 等<sup>[4]</sup>提出一种跨模态哈希方法,在对每种模态的数据进行聚类基础上,将得到的数据表达转换成普通二进制子空间,使所有模态的二进制编码是“一致”和可比较的,同时输出针对所有模态的哈希函数,用于将未知数据转换成二进制代码,从而实现跨模态搜索;Wu 等<sup>[5]</sup>提出一种在线多模态深度相似性学习框架,针对每种单一模态学习一种非线性转换函数,在此基础上去学习发现多种模态的最优组合,应用于多模态图像检索任务,与文献[4]不同的是,文献[5]中所述的模态是图像的不同类型的特征;Wu 等<sup>[47]</sup>提出一种 Bi-CMSRM 的跨媒体特征表示方法,将图像和文本 2 种不同媒体对象映射到相同的公共空间,支持“图像→文本”和“文本→图像”的双向查询。

相似性度量一直是数据挖掘、信息检索、资源推荐等领域的一个研究热点,包括相同模态信息、跨模态资源之间的相似性度量,涉及文本、图像、视频等不同媒体信息,还涉及数值、向量、树、图等不同的数据结构,更涉及距离与非距离等不同的相似性表达机制,因此相关工作还有很多。本节所述的工作更多侧重于非文本和跨媒体的相似性度量,途径是将不同模态特征在相同的新特征空间表示并度量,其余的不多赘述。

## 4 社交媒体推荐面临的问题与挑战

纵观前述第 2、3 节的相关工作,可以总结出以下的特点和问题,并进而引出社交媒体推荐面临的挑战。

### 4.1 目前工作的特点和问题

前文论述了社交媒体资源推荐策略和支撑技术方面的相关工作,针对其中 3 方面推荐策略和 3 方面支撑技术进行归纳,可以总结出如下特点。

1)在社交媒体用户社会关系分析方面,考虑网络的异构性、并进行相关分析和挖掘,是一个新的研究热



点。在异构信息网络中如何更好地发现隐式社群并最终应用于社交媒体资源推荐,仍有许多待解决的问题;

2)在基于用户社群的推荐方面,相关工作主要是基于社群内的用户兴趣或用户关系本身进行推荐。在社交媒体资源推荐中,还应考虑如何在用户社群发现的基础上、进一步推演用户兴趣并应用于资源推荐;

3)在社交媒体资源挖掘方面,除针对社交媒体海量、噪音等研究外,一方面基于传统的多媒体挖掘技术,另一方面基于社交媒体中丰富的语义、元数据信息;

4)在社交媒体资源推荐方面,大多以资源原始粒度作为推荐单元,一些具有综合资源推荐功能的工作也主要局限在“旅游”这样特定需求方面,因而尚缺少更小或更大粒度、特别是综合资源的推荐;

5)在多因素分析的社交媒体资源推荐方面,相关工作结合了用户关系和媒体内容,但其中的内容分析仍主要以文本形式的元数据和语义信息为主。

实际上,内容分析在传统的推荐、特别是基于内容的推荐中曾起过重要作用。在社交媒体中,丰富的标签、元数据等语义信息以及用户关系等线索,使得很多推荐工作在未进行内容分析的情况下也能取得不错的推荐效果,甚至在 ACM Multimedia 2012 国际会议上还引发了在社交媒体环境下“内容已死”与“内容万岁”的讨论<sup>[48]</sup>。但是,如果考虑前述异构信息网络的社群发现以及多层次、不同粒度的信息推荐,内容分析是必不可少的:① 在构建异构信

息网络并分析用户社群时,信息节点的相关性分析与信息净化(去噪、去重等)、以及隐式链接关系分析均离不开内容分析;② 如果要推荐的信息比原始信息粒度更小,则需要对原始信息进行内容分割;③ 如果要推荐的是粒度更大的综合信息,则更需要文本、图像、视频的内容分析和挖掘。由于用户通过社交媒体传播和交流的信息形式的多媒体化,内容分析更需要多模态内容分析技术。

可见,为实现社会媒体的多层次推荐,内容分析仍然有其他方法不可替代的作用。当然,内容分析在进行跨媒体分析时有其局限性,除众所周知的高维性以外,不同类型媒体之间的相似性度量仅依赖于内容分析是难以完成的。考虑一幅图片与一段音乐是否相似,不借助于其他信息几乎无法实现,因而也就限制了不同媒体之间(即跨媒体)的推荐。

综上,内容分析能够提供针对单一媒体精细的分析和比较结果,在推荐系统中的作用是其他分析方法不可替代的。但是,当这种分析和比较涉及不同类型媒体时,内容分析则不能直接应用。此时,内容信息之外的语义信息、元数据信息等,成为不同媒体之间的“桥梁”。基于此,可以认为,结合社交媒体中的用户信息,标签、媒体上下文、评论等用户生成内容,以及这些内容中的术语、视觉、听觉等特征于一体的分析方法,能够更好地支持社会媒体的多模态、多层次资源推荐和用户关系挖掘。表 1 归纳了上述讨论的结果,该表解释了图 2 中虚线的原因。

表 1 相关工作总结  
Table 1 Summary of related work

相关工作		应用的技术	实现的功能	存在的问题
社 会 媒 体 资 源 推 荐 策 略	用户社群	好友挖掘、话题挖掘、偏好挖掘、关注点挖掘、链接挖掘……	礼物推荐、话题推荐、社群推荐、朋友推荐、粉丝推荐……	1) 推荐的主要是原始粒度的资源; 2) 综合资源推荐领域单一; 3) 缺少多层次资源推荐。
	媒体内容	音频挖掘、文本挖掘、视频挖掘、图像挖掘、数据流流分析等……	音乐推荐、新闻推荐、视频推荐、主题推荐、网站推荐……	
	多种因素	用户、标签、文本、图像之间的关联分析、分类、聚类……	书籍推荐、链接推荐、旅游推荐、标签推荐、化妆推荐……	
社 会 媒 体 推 荐 支 撑 技 术	用户关系	同构或异构图聚类、显式或隐式链接分析、频繁图发现……	发现具有共同话题、偏好、潜在合作关系的用户社群……	缺少支持多层次资源推荐的视觉特征、语义信息、元数据等信息的综合数据挖掘技术和方法。
	多模态信息挖掘	链接挖掘、文本挖掘、多媒体数据挖掘、情感分析……	支持多(跨)媒体信息搜索、各种不同模态资源的推荐……	
	多模态资源度量	映射各模态特征到同一公共特征空间、学习潜在语义空间、排序……	支持相同或不同模态资源、同构或异构信息资源的相似性计算……	

4.2 社交媒体多模态、多层次资源推荐面临的挑战

根据图 2、表 1 及上节的讨论,社交媒体多模态、多层次资源推荐仍面临诸多挑战。

1)异构信息网络构建与用户社群发现问题。  
传统的推荐系统中,基于内容的推荐不考虑用户关系,基于协同过滤的推荐主要依据“用户-项

目”的点击情况,因而存在“用户-项目矩阵稀疏”等问题。社交媒体中,“网络社群性”可望更好地解决上述问题。但这里用户与资源的关系已不仅是“点击”操作,在“用户角色双重性”及在多种媒体形式构成的社交媒体环境中,如何构建“用户-资源”异构信息网络并在其中挖掘用户社群,并最终映射为用户兴趣,则是一个挑战。

## 2) 多层次资源质量与条件相关性判别问题。

社交媒体的资源是用户发布的,因此存在大量噪音、重复数据等信息质量问题,不同用户发布的相同内容的资源质量也不尽相同,信息质量问题较传统媒体愈发严重。此外,就推荐而言,在传统的推荐系统中,如果以项目或网页序列表达用户兴趣,那么序列“A”与“AB”对应的推荐应该是不同的。在社交媒体中,这一特点同样存在,然而不同的是,A、B资源的媒体形式更加多样。而在多层次资源推荐中,A、B的粒度则可为子资源、资源和复合资源,推荐结果还可能是综合资源,条件相关性的分析变得更加困难。

## 3) 用户兴趣与多层次资源的描述问题。

无论多层次资源模型、还是用户兴趣模型,均涉及用户、各层次资源及其关系的描述,而资源之间关系的获取需要不同模态的内容分析,这种分析又将落实到特征分析,因此带来如下问题:首先,表现这些资源的数据本身大多是非结构化的;其次,不同资源数据之间是内容异构的;第三,这些非结构化数据的特征是高维的。若准确地合理地表达不同资源之间的关系,并有效地支持综合资源的挖掘以及多层次资源推荐时的搜索,这种描述非简单的数据结构所能胜任。

## 4) 综合资源的发现与推荐问题。

前述定义2中,多层次资源的“子资源”、“资源”和“复合资源”一般以显式方式存在,而“综合资源”则是隐式的,需要在资源和复合资源基础上,通过分析和挖掘才能获得。如何在多(跨)模态的社交媒体资源中,通过针对各种模态的内容分析,发现并表达综合资源、进而向所需用户推荐综合资源,将涉及多(跨)模态挖掘以及推荐时的搜索与匹配等若干新问题。

## 5) 多层次资源模型中的信息度量问题。

如本节内容3)所述,表达社交媒体多层次资源推荐模型已非简单的数据结构,推荐时用户兴趣与资源的匹配又涉及条件相关性判别。在结构化和纯文本内容中,由于特征形式单一,传统的距离度量机制一般尚可奏效。但在社交媒体多模态内容和多层

次资源中,用户兴趣与资源之间、资源与资源之间的相似或相关性度量无法采用传统的“距离”度量机制予以实现。

# 5 进一步的研究方向

可以认为,社交媒体中用户角色的变化构成了更加复杂的用户关系,社交媒体资源表现形式呈现多模态特点,社交媒体资源推荐应该、并可以满足多层次的用户需求。基于此,提出进一步的研究方向。

## 1) 社交媒体资源与用户兴趣模型结构研究。

令社交媒体资源集合为  $R = \{r_1, r_2, \dots, r_n\}$ , 每个  $r_i (i=1, 2, \dots, n)$  为一个资源,  $n$  为资源总数。对一用户  $u$ , 其兴趣模型  $Mu = \{mu_1, mu_2, \dots, mu_m\}$  是根据该用户对资源的操作历史及其社群关系得到的资源序列集合, 其中每个  $mu \in Mu$  为用户  $u$  感兴趣的一个资源序列。不失一般性, 将  $mu$  表示为  $mu = r_1 r_2 \dots r_m$  ( $m$  为该序列长度, 即用户  $u$  感兴趣的一组相关资源数目,  $m < n$ )。就兴趣模型存储本身, 将  $r \in R$ , 表示为  $\langle r, w_r \rangle$ ,  $w_r$  为用户对  $r$  的兴趣度, 基于用户对  $r$  的操作(发布、转发、...、以及来源于社群兴趣等)赋予不同的权值。而在推荐时, 将根据兴趣度对资源序列进行筛选。基于指定的  $mu \in Mu$  的推荐, 实际上需要在  $R$  中发现一个最适合推荐的资源序列  $R_r \subseteq R$ , 使该序列中任一项  $r \in R_r$  满足  $\arg\max_{rec} (r | r_1 r_2 \dots r_m)$ , 其中  $f_{rec}$  为推荐指数,  $f_{rec}(r | r_1 r_2 \dots r_m)$  即用户  $u$  对  $r_1 r_2 \dots r_m$  感兴趣的条件下、推荐资源  $r$  的指数。由于社交媒体推荐中“资源”已非原始粒度的信息, 因此, 资源的识别和描述将是一个新的问题。同时, 如何对多层次资源和用户兴趣建模, 如何进行推荐条件判断, 乃至模型的维护和支持查询的有效索引均为需要研究的问题。

## 2) 社交媒体多模态内容与资源质量分析基础研究。

由于社交媒体中存在大量噪音、重复资源等质量问题, 无论在构建资源模型与用户兴趣模型、还是构建用户-资源信息网络时, 均存在资源的质量分析问题, 噪音和重复资源的检测依赖于多模态内容分析。此外, 资源模型的表达是以计算两资源  $r_a, r_b \in R$  的  $f_{rec}(r_a | r_b)$  为基础的, 而  $f_{rec}(r_a | r_b)$  正比于资源内容相关性, 内容相关性的度量同样依赖于多模态内容分析。如前所述, 这里的多模态内容除包括媒体上下文线索、标签等元数据等语义信息外, 视觉、听觉等内容分析的基础是特征提取, 这需要将特征相关性问题首先在不同类别的特征中解决, 再进一步在多类特征中融合。这样, 分层次、分类别的特征提取将是一个重要的研究问题。



3) 社交媒体异构信息网络社群挖掘技术研究。

本节内容 1) 中涉及的用户兴趣模型结构,是将用户兴趣模型表示为资源序列集合  $Mu = \{mu\}_m$ , 一个用户的  $Mu$  可以从该用户对资源的发布和使用历史获得。在社交媒体中,一个更显著的特点是基于用户的网络社会关系、也就是其所在的社群中其他用户的兴趣予以推断。在本文中,社群本身也被视为一种综合资源,通过社群挖掘获得。前面提及,社交媒体中用户和资源的关系构成异构信息网络,网络中除显式的链接(用户对资源的操作)外,还包括资源内容相关性以及由此导致的用户相关性等隐式链接。

4) 社交媒体资源条件相关性挖掘技术研究。

本节内容 1)、2) 分别研究模型的结构和各层次特征抽取及上下文线索的获取问题。而更为复杂的问题是  $f_{rec}(r_i|r_{i-1}, \dots, r_1)$  关系的构建,它涉及同一资源的不同模态、相同模态的不同资源以及不同模态的不同资源之间的条件相关性的计算问题,多模态信息挖掘技术可望解决上述问题。

5) 社交媒体综合资源发现与推荐策略研究。

综合资源并非像单一资源或复合资源那样显式地存在,它是通过对系统资源挖掘并结合用户兴趣模型中的资源分析后发现的,是由若干相同或不同来源的、相同或不同媒体类型的资源构成的集合。此外,前面提及,用户社群也是一种综合资源。

6) 社交媒体资源与用户兴趣匹配机制研究。

本节 1)~5) 的研究内容将资源模型表示为一系列  $f_{rec}(r_i|r_{i-1}, \dots, r_1)$  ( $i>1$ ) 关系及这些资源不同层次的特征,将用户模型表示成一系列资源访问序列  $r_1r_2\dots r_m$  及其中每个资源不同层次的特征。构建该模型的最终目的是在系统资源中发现满足用户兴趣的资源予以推荐。因此,应用该模型推荐时,首先应在资源模型  $f_{rec}(r_i|r_{i-1}, \dots, r_1)$  的条件“ $r_{i-1}, \dots, r_1$ ”中找到与用户模型中的兴趣资源序列“ $r_1r_2\dots r_m$ ”最匹配者,计算满足该条件的所有  $f_{rec}(r|r_1r_2\dots r_m)$ , 并将满足  $\operatorname{argmax}_{f_{rec}}(r|r_1r_2\dots r_m)$  的资源  $r$  (即具有最大推荐指数的资源)作为推荐候选。

综上所述,本文 4.2 节提出的挑战与本节提出的研究方向之间的对应关系如图 3 所示。

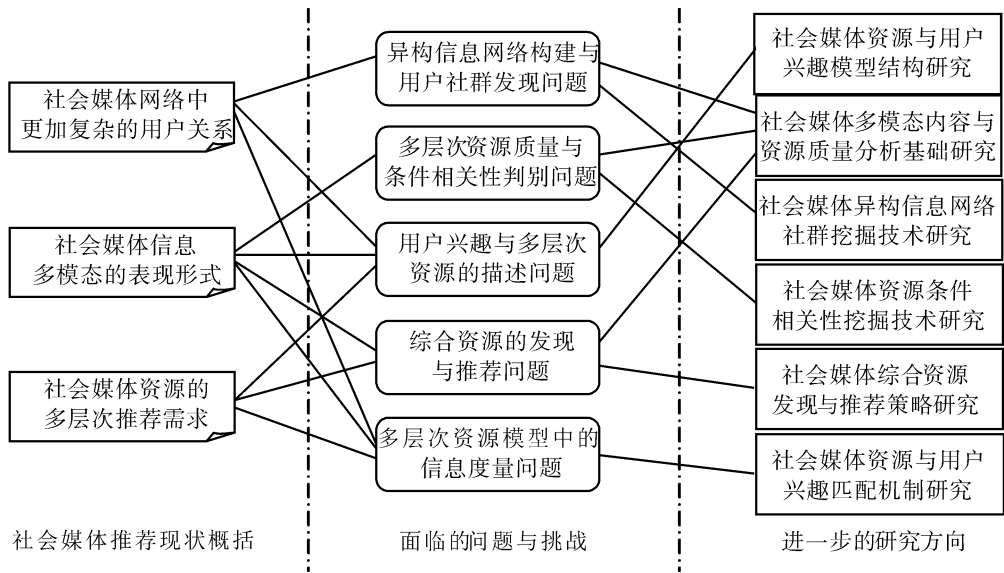


图3 社交媒体推荐现状、问题挑战与研究方向的对应关系

Fig.3 Relationship among present situation, challenges, and research direction in social media

6 结束语

本文分析了社交媒体中的用户和资源所呈现的新特点,提出社交媒体资源的推荐是对传统的“基于内容的推荐”、“基于协同过滤的推荐”及其混合模式的扩展,具有“多模态的用户兴趣内容分析”、“基于社会关系的协同过滤”和“面向用户的多层次资源推荐”的特点,进而以此展开,综述了基于社群的推荐、基于媒体内容的推荐、基于多种因素的推荐

等3种推荐策略的研究现状,综述了社交媒体用户关系分析、社交媒体多模态信息挖掘、多模态信息相似性度量等3种社交媒体推荐的支撑技术的相关工作。在对研究现状和相关工作讨论的基础上,将其概括为“社交媒体中用户角色的变化构成了更加复杂的用户关系”、“社交媒体资源表现形式呈现多模态特点”、以及“社交媒体资源推荐应该、并可以满足多层次的用户需求”,基于此提出了当前社交媒体资源推荐面临的挑战及进一步的研究方向。

社交媒体是一个巨大的数据和信息资源,涉及众多研究领域,基于社会媒体的资源推荐仅仅是其中的一个领域。本文仅从社交媒体资源推荐、特别是多模态与多层次资源推荐方面进行了讨论,其中所述的推荐策略和相关支撑技术本身均可作为一个研究方向并扩展出更多新的研究和应用,社交媒体与数据质量、隐私保护、大数据分析处理等技术的结合,均为社交媒体研究的一些新领域。

## 参考文献:

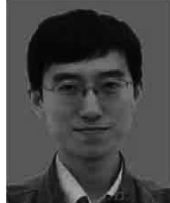
- [1] ADOMAVICIUS G, TUZHILIN A. Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2005, 17(6): 734-749.
- [2] 黄立威, 李德毅. 社交媒体中的信息推荐[J]. *智能系统学报*, 2012, 7(1): 1-8.  
HUANG Liwei, LI Deyi. A review of information recommendation in social media[J]. *CAAI Transactions on Intelligent Systems*, 2012, 7(1): 1-8.
- [3] SUN Y, HAN J. Mining heterogeneous information networks: a structural analysis approach[J]. *SIGKDD Explorations*, 2012, 14(2): 20-28.
- [4] ZHU X, HUANG Z, SHEN H, et al. Linear cross-modal hashing for efficient multimedia search[C]//*ACM Multimedia*. Barcelona, Spain, 2013: 143-152.
- [5] WU P, HOI S, XIA H, et al. Online multimodal deep similarity learning with application to image retrieval[C]//*ACM Multimedia*. Barcelona, Spain, 2013: 153-162.
- [6] PAVLIDIS Y, MATHIHALLI M, CHAKRAVARTY I, et al. Anatomy of a gift recommendation engine powered by social media[C]//*ACM SIGMOD Conference*. Scottsdale, AZ, USA, 2012: 757-764.
- [7] POPESCU A, GREFFENSTETTE G. Mining social media to create personalized recommendations for tourist visits[C]//*2nd International Conference and Exhibition on Computing for Geospatial Research & Application*. Washington, DC, USA, 2011: 37.
- [8] SCHIRRU R. Topic-based recommendations in enterprise social media sharing platforms[C]//*ACM Conference on Recommender Systems*. Barcelona, Spain, 2010: 369-372.
- [9] 贾大文, 曾承, 彭智勇, 等. 一种基于用户偏好自动分类的社交媒体共享和推荐方法[J]. *计算机学报*. 2012, 35(11): 2381-2391.  
JIA Dawen, ZENG Cheng, PENG Zhiyong, et al. A user preference based automatic potential group generation method for social media sharing and recommendation[J]. *Chinese Journal of Computer*, 2012, 35(11): 2381-2391.
- [10] KOOHBORFARDHAGHIGHI S, KIM J. Using structural information for distributed recommendation in a social network[J]. *Applied Intelligence*, 2013, 38(2): 255-266.
- [11] LI L, PENG W, KATARIA S, et al. FRec: a novel framework of recommending users and communities in social media[C]//*ACM International Conference on Information and Knowledge Management*. San Francisco, CA, USA, 2013: 1765-1770.
- [12] ZHAO G, LEE M, HSU W, et al. Community-based user recommendation in uni-directional social networks[C]//*ACM International Conference on Information and Knowledge Management*. San Francisco, CA, USA, 2013: 189-191.
- [13] ZHANG H, ZHA Z, YANG Y, et al. Attribute-augmented semantic hierarchy: towards bridging semantic gap and intention gap in image retrieval[C]//*ACM Multimedia*. Barcelona, Spain, 2013: 33-42.
- [14] BU J, TAN S, CHEN C, et al. Music recommendation by unified hypergraph: combining social media information and music content[C]//*ACM Multimedia*. Firenze, Italy, 2010: 391-400.
- [15] TAN S, BU J, CHEN C, et al. Using rich social media information for music recommendation via hypergraph model[C]//*Social Media Modeling and Computing*. London, UK, 2011: 213-237.
- [16] HU C, ZHANG C, WANG T, et al. An adaptive recommendation system in social media[C]//*45th Hawaii International Conference on System Sciences*. Maui, USA, 2012: 1759-1767.
- [17] MA X, WANG H, LI H, et al. Enhancing recommended video lists for youtube-like social media[C]//*IEEE International Workshop on Multimedia Signal Processing*. Banff, AB, Canada, 2012: 244-249.
- [18] ERNESTO D, LUCAS D, LARS S, et al. Real-time top-n recommendation in social streams[C]//*ACM Conference on Recommender Systems*. Dublin, Ireland, 2012: 59-66.
- [19] LI Q, WANG J, CHEN Y, et al. User comments for news recommendation in forum-based social media[J]. *Information Science*, 2013, 180(24): 4929-4939.
- [20] MESSENGER A, WHITTLE J. Recommendations based on user-generated comments in social media[C]//*IEEE Third International Conference on Social Computing*. Boston, MA, USA, 2011: 505-508.
- [21] PERA M, CONDIE N, NG Y. Personalized book recommendations created by using social media data[C]//*WISE 2010 International Symposium WISS, and International Workshops CISE, MBC*, Hong Kong, China, 2010: 390-403.
- [22] GUY I, ZWERDLING N, RONEN I, et al. Social media recommendation based on people and tags[C]//*ACM SIGIR Conference on Research and Development in Information Retrieval*. Geneva, Switzerland, 2010: 194-201.
- [23] WU S, RAND W, RASCHID L. Recommendations in social media for brand monitoring[C]//*ACM Conference on Recommender Systems*. Chicago, IL, USA, 2011: 345-348.
- [24] XIANG Z. Dynamic social media in online travel information search: a preliminary analysis[C]//*International Conference in Innsbruck*. Innsbruck, Austria, 2011: 343-353.
- [25] Chidlovskii B. Learning recommendations in social media

- systems by weighting multiple relations [C]//Machine Learning and Knowledge Discovery in Databases- European Conference. Athens, Greece, 2011: 328-334.
- [26] LIU L, XU H, XING J, et al. "Wow! You are so beautiful today!" [C]//ACM Multimedia, Barcelona, Spain, 2013: 3-12.
- [27] SUN Y, NORICK B, HAN J, et al. Integrating meta-path selection with user-guided object clustering in heterogeneous information networks[C]//ACM SIGKDD Conference. Beijing, China, 2012: 1348-1356.
- [28] GUY I, JACOVI M, PERER A, et al. Same places, same things, same people?: Mining user similarity on social media[C]//ACM Conference on Computer Supported Cooperative Work. Savannah, Georgia, USA, 2010: 41-50.
- [29] BARBIER G, TANG L, LIU H. Understanding online groups through social media[J]. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 2011, 1(4): 330-338.
- [30] YANG C, YANG H, TANG X, et al. Identifying implicit relationships between social media users to support social commerce [C]//International Conference on Electronic Commerce. Singapore 2012: 41-47.
- [31] JORGE C, ALNEU L. Exploiting behaviors of communities of twitter users for link prediction[J]. Social Network Analysis and Mining, 2013, 3(4): 1063-1074.
- [32] TANG J, WANG M, HUA X, et al. Social media mining and search[J]. Multimedia Tools and Applications, 2012, 56(1): 1-7.
- [33] LEE C. Unsupervised and supervised learning to evaluate event relatedness based on content mining from social-media streams [J]. Expert Systems with Applications, 2012, 39(18): 13338-13356.
- [34] JIN X, LIN C, LUO J, et al. SocialSpamGuard: A data mining-based spam detection system for social media networks[J]. Proceedings of the VLDB Endowment, 2011, 4(12): 1458-1461.
- [35] LEE C, CROFT W, KIM J. Evaluating search in personal social media collections[C]//International Conference on Web Search and Web Data Mining. Seattle, WA, USA, 2012: 683-692.
- [36] CHOUDHURY M, COUNTS S, CZERWINSKI M. Find me the right content! diversity-based sampling of social media spaces for topic-centric search[C]//International Conference on Weblogs and Social Media. Barcelona, Catalonia, Spain 2011.
- [37] TSAI F. Probabilistic models for social media mining[J]. International Journal of Information Technology and Web Engineering, 2011, 6(1): 13-24.
- [38] KASCHESKY M, SOBKOWICZ P, BOUCHARD G. Opinion mining in social media: modeling, simulating, and visualizing political opinion formation in the Web[C]//12th Annual International Conference on Digital Government Research. MD, USA, 2011: 317-326.
- [39] JIN X, WANG C, LUO J, et al. LikeMiner: a system for mining the power of 'like' in social media networks[C]//International Conference on Knowledge Discovery and Data Mining. San Diego, CA, USA, 2011: 753-756.
- [40] CUI B, TUNG A, ZHANG C, et al. Multiple feature fusion for social media applications [C]//Proc ACM SIGMOD, Conference. Indianapolis, Indiana, USA, 2010: 435-446.
- [41] SAXTON G, OH O, KISHORE R. Rules of crowdsourcing: models, issues, and systems of control[J]. Information Systems Management, 2013, 30(1): 2-20.
- [42] RAWASHDEH M, KIM H, EL-SADDIK A. Social media annotation and tagging based on folksonomy link prediction in a tripartite graph[C]//Advances in Multimedia Modeling. Huangshan, China, 2013: 24-35.
- [43] LIU L, ZHU F, JIANG M, et al. Mining diversity on social media networks[J]. Multimedia Tools and Applications, 2012, 56(1): 179-205.
- [44] LING L, ZHAI X, PENG Y. Tri-space and ranking based heterogeneous similarity measure for cross-media retrieval [C]//International Conference on Pattern Recognition. Tsukuba, Japan, 2012: 230-233.
- [45] ZHAI X, PENG Y, XIAO J. Effective heterogeneous similarity measure with nearest neighbors for cross-media retrieval [C]//Advances in Multimedia Modeling. Klagenfurt, Austria, 2012: 312-322.
- [46] JIA Y, SALZMANN M, DARRELL T. Learning cross-modality similarity for multinomial data[C]//IEEE International Conference on Computer Vision. Barcelona, Spain, 2011: 2407-2414.
- [47] WU F, LU X, ZHANG Z, et al. Cross-media semantic representation via bi-directional learning to rank [C]//ACM Multimedia. Barcelona, Spain, 2013: 877-886.
- [48] XIE L, SHAMMA D, SNOEK C. Content is dead: long-live content! [C]//ACM Multimedia. Nara, Japan, 2012: 7-8.

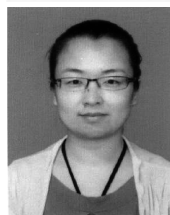
#### 作者简介:



王大玲,女,1962年生,教授,中国计算机学会高级会员,中国计算机学会中文信息技术专业委员会委员,主要研究方向为数据挖掘、信息检索、Web推荐等,主持国家自然科学基金项目3项,参与国家自然科学基金项目2项。



冯时,男,1981年生,讲师,主要研究方向情感挖掘、网络舆情分析。



张一飞,女,1977年生,讲师,主要研究方向为图像处理、多媒体数据挖掘。