

联合建模异构社交和内容信息的活动推荐模型*

王绍卿^{1,2,3}, 王征^{1,2}, 李翠平^{1,2}, 赵衍衍^{1,2}, 陈红^{1,2}



¹(数据工程与知识工程教育部重点实验室(中国人民大学), 北京 100872)

²(中国人民大学 信息学院, 北京 100872)

³(山东理工大学 计算机科学与技术学院, 山东 淄博 255091)

通讯作者: 李翠平, E-mail: licuiping@ruc.edu.cn

摘要: 随着基于活动的社交网络的迅速发展,活动推荐已成为一个重要的工具,帮助人们在线上发现有趣的活
动,并在线下面对面地参与活动.但是,相对于传统的推荐系统,活动推荐面临着很多挑战.(1) 用户只能参与很少的
活动,这就导致一个非常稀疏的用户-活动矩阵;(2) 用户对活动的响应是隐性反馈;(3) 活动本身有生命周期,已经过
期的活动不能再向用户推荐;(4) 每天会有很多新的活动产生,需要及时向用户推荐.为了应对这些挑战,提出一个联
合建模异构社交和内容信息的活动推荐模型.该模型可同时探索用户的线上和线下社交活动,并结合活动内容建模
用户对活动的决策行为.在 Meetup 数据集上做实验以评估所提出模型的性能.实验结果表明,提出的模型优于其他
方法.

关键词: 基于活动的社交网络;社交网络分析;活动推荐;联合建模;泊松因子分解

中图法分类号: TP311

中文引用格式: 王绍卿,王征,李翠平,赵衍衍,陈红.联合建模异构社交和内容信息的活动推荐模型.软件学报,2018,29(10):
3134-3149. <http://www.jos.org.cn/1000-9825/5284.htm>

英文引用格式: Wang SQ, Wang Z, Li CP, Zhao KK, Chen H. Jointly modeling heterogeneous social and content information for
event recommendation. Ruan Jian Xue Bao/Journal of Software, 2018,29(10):3134-3149 (in Chinese). <http://www.jos.org.cn/1000-9825/5284.htm>

Jointly Modeling Heterogeneous Social and Content Information for Event Recommendation

WANG Shao-Qing^{1,2,3}, WANG Zheng^{1,2}, LI Cui-Ping^{1,2}, ZHAO Kan-Kan^{1,2}, CHEN Hong^{1,2}

¹(Key Laboratory of Data Engineering and Knowledge Engineering, Ministry of Education (Renmin University of China), Beijing 100872,
China)

²(Information School, Renmin University of China, Beijing 100872, China)

³(School of Computer Science and Technology, Shandong University of Technology, Zibo 255091, China)

Abstract: Event-Based social networks (EBSNs) have experienced rapid growth in people's daily life. Hence, event recommendation
plays an important role in helping people discover interesting online events and attend offline activities face to face in the real world.
However, event recommendation is quite different from traditional recommender systems, and there are several challenges: (1) One user
can only attend a scarce number of events, leading to a very sparse user-event matrix; (2) The response data of users is implicit feedback;

* 基金项目: 国家重点研发计划(2016YFB1000702); 国家基础研究发展计划(973)(2014CB340402); 国家自然科学基金
(61772537, 61772536, 61702522); 中国人民大学科学研究基金(中央高校基本科研业务费专项资金资助)(15XNLQ06)

Foundation item: National Key Research & Develop Plan (2016YFB1000702); National Basic Research Program of China (973)
(2014CB340402); National Natural Science Foundation of China (61772537, 61772536, 61702522); The Fundamental Research Funds for
the Central Universities, and the Research Funds of Renmin University of China (15XNLQ06)

收稿时间: 2016-11-16; 修改时间: 2017-01-14, 2017-03-16; 采用时间: 2017-03-24; jos 在线出版时间: 2017-07-12

CNKI 网络优先出版: 2017-07-12 15:33:31, <http://kns.cnki.net/kcms/detail/11.2560.TP.20170712.1533.005.html>

(3) Events have their life cycles, so outdated events should not be recommended to users; (4) A large number of new events which are created every day need to be recommended to users in time. To cope with these challenges, this article proposes to jointly model heterogeneous social and content information for event recommendation. This approach explores both the online and offline social interactions and fuses the content of events to model their joint effect on users' decision-making for events. Extensive experiments are conducted to evaluate the performance of the proposed model on Meetup dataset. The experimental results demonstrate that the proposed model outperforms state-of-the-art methods.

Key words: event based social network; social network analysis; event recommendation; jointly modeling; Poisson factorization

1 引言

随着第二代互联网、无线通信技术的迅速发展,以及用户产生内容的急剧增长,基于活动的社交网络(event-based social networks,简称 EBSNs)得到迅速发展,例如 Meetup(<http://www.meetup.com>)、豆瓣同城(<http://bj.douban.com>)。用户在 EBSNs 可以扮演两种角色:组织者和参与者。组织者发布一些活动(event),参与者在在线上找到感兴趣的活动进行响应,然后在线下面对面地参与活动。用户在线上加入兴趣小组,可以分享、评论活动,构建了线上社交网络。用户在线下面对面地参与活动,结交一些新的朋友,构建了线下社交网络。EBSNs 把线上服务和线下活动链接在一起,是一个异构的社交网络^[1]。用户之间具有异构的社交关系。

Meetup 官网上的资料(<http://www.meetup.com/about>)显示,当前有大约 2 837 万注册用户,每月产生大约 62 万个活动。一方面,组织者希望更多的用户来参与活动,另一方面,参与者希望快速地找到感兴趣的活动。有效地进行活动推荐可满足组织者和参与者的需求^[2]。通过利用所有参与者的历史记录及其他信息,为用户准确地推荐感兴趣的活动。若没有特殊说明,本文后面的内容中提到的用户都特指参与者这一角色。

活动推荐与电影推荐^[3,4]、音乐推荐^[5,6]、新闻推荐^[7]、POI 推荐^[8]等传统的推荐系统相比,面临着更加严峻的挑战:(1) 大部分用户只会参与很少的活动,大部分活动只会被很少的用户参与。根据 Meetup 官方资料显示,每月只有约 394 万人次的用户响应,这就导致一个极其稀疏的用户-活动响应矩阵。(2) 用户对活动的响应是隐性反馈。用户会对感兴趣的活动进行响应,而没有进行响应的活动也可能是感兴趣的。(3) 活动本身有生命周期,已经过期的活动不能再向用户推荐。(4) 每天会有很多新的活动产生,需要及时向用户推荐。

为应对活动推荐所面临的挑战,本文提出一个基于泊松因子分解的联合概率产生模型。(1) 探索和利用用户之间的异构社交关系来应对数据极其稀疏的挑战。社会关系理论中的同质性^[9]表明,有相似偏好的用户更可能成为社交好友。同时,社会影响^[10]启示社交好友更可能有相似的偏好^[11]。直觉上,用户对活动的响应,不仅受其自身兴趣的影响,也受用户的线上和线下社交好友的影响。在某些情形下,用户可能更加信任某种社交关系下的用户,而在另外一些情况下,用户可能更加信任具有多种社交关系的用户。仅依赖于一种社交关系的推荐模型必然会存在信息缺失问题。(2) 使用贝叶斯泊松因子分解模型来建模用户的隐性反馈。传统的基于高斯似然的矩阵因子分解模型对于用户响应的活动和未响应的活动赋予相同的权重。对于极其稀疏的用户-活动响应矩阵,传统的矩阵因子分解模型过分强调未响应的用户-活动对。而泊松因子分解模型可以更好地捕获用户的行为^[12]。(3) 把已经过期的活动过滤掉,然后再生成推荐列表。(4) 探索利用活动的内容信息来提升推荐的性能^[13]。

图 1 展示了在基于活动的社交网络上进行活动推荐的一个例子。给定每个用户所属的兴趣小组、每个用户的历史响应行为以及活动的内容,向用户推荐一些正在进行的且符合用户偏好的活动。

本文提出一个基于泊松因子分解的联合概率产生模型,可同时探索和利用用户的线上和线下社交活动。模型强调好友之间不同的信任强度会对用户带来不同的影响,同时结合活动内容建模用户对活动的决策行为。采用坐标上升算法推断模型参数。最后,在真实数据集上做实验以评估所提出模型的性能。实验结果表明,本文提出的模型优于其他现有方法。

本文第 2 节主要介绍社会化推荐、隐性反馈和活动推荐的相关工作。第 3 节是预备知识。第 4 节引入基于异构社交和内容信息的泊松因子分解模型。第 5 节对模型的参数推断进行详细推导,并给出活动推荐方法。第 6 节给出数据集、实验方法和指标,介绍对比模型,对实验结果进行分析,并讨论模型中不同因素的贡献。最后,在第 7 节给出研究结论以及下一步的研究展望。

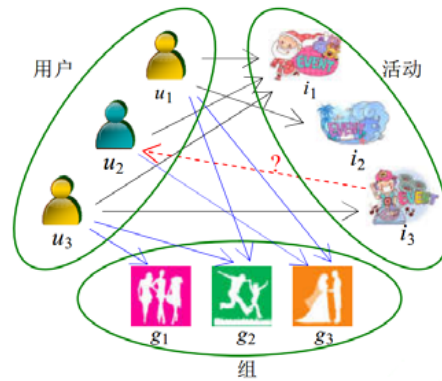


Fig.1 An example of event recommendation at an event-based social network

图1 在基于活动的社交网络上进行活动推荐的一个例子

2 相关工作

本文的研究内容是在基于活动的社交网络中进行活动推荐,所以,本节首先介绍社会化推荐的研究成果,然后引入基于隐性反馈的研究方法,最后给出活动推荐的研究现状.

2.1 社会化推荐

社会化推荐是在传统推荐系统的基础上,把用户之间的社会关系信息(比如微博数据中的关注关系,科技论文中的合作关系等)融入到推荐模型中,从而提高推荐系统的性能^[14,15].近年来,社会化推荐已成为推荐系统研究领域较为活跃的研究方向之一.

文献[16]通过共享用户潜在特征空间的方式来连接社交网络结构和用户-物品评分矩阵;文献[17,18]使用组合信任的方式来同时建模个人兴趣和好友的偏好;文献[19]建模社交网络信息为正则化项来约束矩阵因子分解框架;但它们没有考虑用户之间的异构信任关系.

文献[20]把用户的种类信息融入用户-项目矩阵,并结合社交网络结构来提升推荐性能;文献[21]通过建模用户为信任者和信托者两种不同角色的方式来把社交网络结构整合到推荐模型;文献[22]探索将用户作为商品评论的作者和评分者两种角色,结合多种社会上下文来进行商品评论评分的预测;文献[11]利用局部和全局社交上下文进行推荐;文献[23]认为社交好友之间有显性的信任关系,评分相似的用户之间具有隐性信任关系;文献[24,25]综合利用用户之间的强依赖和弱依赖关系来提升推荐性能;但它们都使用了基于高斯似然的矩阵因子分解模型,不能很好地捕获活动推荐中的隐性反馈.

文献[26]提出了一种基于主题的张量分解的用户信任推荐算法,用来挖掘用户选取不同的物品时对不同朋友的信任程度;文献[27]提出一种融合项目特征和移动用户信任关系的推荐算法;文献[28]认为,用户购买商品的行为不但受社交好友的影响,也受网络购物平台上商品之间关联规则的影响,用超图把二者结合到一个框架,从而推导用户将来的购买行为;但这些模型的复杂度偏高,且不能利用内容信息.

文献[29]使用泊松因子分解的方式建模社交网络结构,但没有考虑用户之间的异构社交关系,且没有探索利用内容信息来提升推荐性能.

2.2 隐性反馈

用户对物品的偏好表现为显性评分和隐性反馈这两种方式^[30].在基于活动的社交网络(EBSNs)上,用户不能对活动进行显性的评分.在EBSNs的相关数据中只存在用户响应活动的历史行为,也就是只有正例,用于记录用户偏好的隐性反馈.响应的活动可以认为是符合用户兴趣的,而未响应的活动也可能与用户的兴趣一致.

文献[30]提出 SVD++模型,用于探索利用矩阵因子分解模型同时建模显性评分和隐性反馈,预测用户对物品的评分,根据评分的高低进行推荐.文献[31]指出,用户的偏好和物品的流行度是随着时间动态变化的,进而加

入了时序,并提出了 timeSVD++模型.但这些模型并不适合只包含隐性反馈的 EBSNs 数据.

文献[32]针对用户的隐性反馈,提出用物品对(item pair)作为训练数据,以优化物品的排序为最终目标.文献[33]提出一个 Maa-BPR 模型,探索组合多个属性和隐性反馈信息来学习各个实体的隐因子和属性的映射.文献[34]在因子分解机模型的基础上,把隐性反馈和内容信息结合起来,并提出一个个性化的排序模型.

文献[12]指出,在传统的基于高斯似然的模型中,存在用户对所有的物品是以相同的权重进行建模这一缺点,进而提出适合隐性反馈的贝叶斯泊松因子模型.文献[44]提出 CTPF 模型,探索把内容信息整合到泊松因子分解模型中.

2.3 活动推荐

基于活动的社交网络是一个相对新兴的社交网络.文献[1]首先指出 EBSNs 链接线上和线下社交世界,并用网络拓扑结构来探索活动推荐;文献[35]提出基于传统的奇异值分解的多因子近邻算法以预测用户对活动的响应;文献[36]提出一个图模型的方式,把 EBSNs 中向用户推荐活动、向用户推荐兴趣小组和向小组推荐标签整合到一个框架;文献[37]分别基于社会关系、内容、位置和时间建立推荐模型,然后把得到的评分和活动的响应数量组合为一个向量,进一步利用排序学习的方式训练用户对活动的偏好;文献[13]提出协同贝叶斯泊松因子分解模型,同时建模社交关系、内容、组织和位置信息;这些文献主要考虑到要尽可能多地融合多种因素(如线上好友、组织者、地点、时间等)进行活动推荐,而忽略了对异构的社交关系的分析和利用,不能对用户进行更准确的建模.文献[2]组合异构社交和位置信息进行活动推荐,但它使用了基于高斯似然的矩阵因子分解模型,不能很好地捕获活动推荐中的隐性反馈,也没有利用活动的内容来提升推荐的性能.文献[38]探索了 EBSNs 中异构的社交关系,但只是简单地使用用户共同加入的兴趣小组的数量来表示线上信任强度,用用户共同参与的活动的数量来表示线下信任强度,对于好友数量较多的用户,模型训练过程中容易导致数据溢出.另外,所提出的 HSPF 模型也没有利用活动的内容信息.

3 预备知识

Gopalan 提出贝叶斯泊松因子分解(Bayesian Poisson factorization,简称 BPF)模型^[12].BPF 模型假设用户 u 对项目 i 的响应(如,点击、收藏、购买、关注、评分、评论、转发等) y_{ui} 服从泊松分布:

$$y_{ui} \sim \text{Poisson}(\theta_u^T \beta_i),$$

其中, θ_u 表示用户 u 的偏好,是一个 K 维非负向量; β_i 表示项目 i 的属性,也是一个 K 维非负向量. θ_u 和 β_i 是未知的变量,需要根据观察的响应数据来估计.进一步地,假设 θ_u 和 β_i 中的每一个元素服从 Gamma 分布,即

$$\theta_{uk} \sim \text{Gamma}(\lambda_{ua}, \lambda_{ub}),$$

$$\beta_{ik} \sim \text{Gamma}(\lambda_{ia}, \lambda_{ib}),$$

其中, λ_{ua} 和 λ_{ia} 是 Gamma 分布的形状参数, λ_{ub} 和 λ_{ib} 是 Gamma 分布的尺度参数.可以利用吉布斯抽样或者变分推导的方式求出 θ_u 和 β_i 的估计值.

BPF 模型把泊松因子分解(Poisson factorization,简称 PF)^[39-41]和贝叶斯学习相结合,能够有效地处理稀疏数据,避免过拟合,且尤其适合处理隐性反馈数据^[12].

4 基于异构社交和内容信息的泊松因子分解模型——PF-HSC 模型

在基于活动的社交网络上,若用户 u 在线上对活动 i 做出积极的响应,即 $y_{ui}=1$,则认为用户 u 会在线下参加或参与活动 i .用户 u 是否对活动 i 做出积极的响应不仅受个人兴趣的影响,也受信任好友的影响.也就是说,如果活动 i 和用户 u 的历史兴趣偏好不一致,但用户 u 的信任好友对活动 i 做出了响应,那么用户 u 也可能参与活动 i .直觉上,对于用户 u ,不同的信任好友对其影响是不一样的.用户 u 与好友之间的信任关系越强,好友对其影响也会越大.

给定稀疏的用户-活动响应数据,PF-HSC 模型探索和利用用户之间的各种信任关系数据和活动的内容信

息来提高推荐或预测(哪些用户将来会对哪些活动做出积极的响应)的准确性.PF-HSC(Poisson factorization based on heterogeneous social and content information)模型是一个隐变量模型.PF-HSC 模型使用贝叶斯泊松因子分解的方式来同时建模用户-活动响应、用户之间的各种信任关系和活动的信息.

本文第 4.1 节定义了不同社交关系中用户之间的影响权重,具体来说,影响权重可具体分为信任系数和信任强度,其中,信任强度可根据用户的历史行为预计算出来,而信任系数则需要通过模型来推断.第 4.2 节给出了 PF-HSC 模型的细节.PF-HSC 模型充分利用了泊松分布的一个重要特性:多个独立的泊松分布之和仍是一个泊松分布.探索利用泊松因子分解建模异构的社交关系和活动的信息.

4.1 信任系数和信任强度

用户 u 和用户 v 在线上共同加入了一个兴趣小组(group),则他们具有线上信任关系,信任系数记作 δ_{uv}^{on} . 而用户 u 和用户 f 在线下共同参与了一个活动(event),则用户 u 和用户 f 构建了隐式的信任关系,称为线下信任关系,信任系数记作 τ_{uf}^{off} . δ_{uv}^{on} 和 τ_{uf}^{off} 是未知的,需要根据观察数据进行估计.

信任关系中,不同用户之间具有不一样的影响力,称为信任强度.对于线上信任网络,用户 u 和用户 v 共同加入的兴趣小组越多,则他们之间的线上信任强度 r_{uv}^{on} 越大.对于线下信任网络,用户 u 和用户 f 共同响应的活动越多,则他们之间的线下信任强度 r_{uf}^{off} 越大.本文使用杰卡德相似系数来计算用户之间的信任强度.即,

$$r_{uv}^{on} = \frac{|G(u) \cap G(v)|}{|G(u) \cup G(v)|} \quad (1)$$

$$r_{uf}^{off} = \frac{|E(u) \cap E(f)|}{|E(u) \cup E(f)|} \quad (2)$$

其中, $G(u)$ 和 $G(v)$ 分别表示用户 u 和用户 v 在线上所加入的兴趣小组的集合, $E(u)$ 和 $E(f)$ 分别表示用户 u 和用户 f 所响应的活动的集合.

4.2 PF-HSC模型

在 PF-HSC 模型中,用户 u 对应一个 K 维向量 θ_u ,活动 i 对应一个 K 维向量 β_i , $N^{on}(u)$ 和 $N^{off}(u)$ 分别表示用户 u 的线上和线下信任网络中的好友集合.

观察到的 y_{ui} 的分布可表示为

$$y_{ui} | y_{-ui} \sim \text{Poisson} \left(\theta_u^T \beta_i + \sum_{v \in N^{on}(u)} \delta_{uv}^{on} r_{uv}^{on} y_{vi} + \sum_{f \in N^{off}(u)} \tau_{uf}^{off} r_{uf}^{off} y_{fi} \right) \quad (3)$$

其中, y_{-ui} 表示除用户 u 之外的其他用户对活动 i 的响应.为了完成对未知变量的估计,本文假设用户的偏好 θ_u 、活动的属性 β_i 、线上信任系数 δ_{uv}^{on} 和线下信任系数 τ_{uf}^{off} 均服从其各自形状和尺度超参数的 Gamma 分布.这是因为, Gamma 分布和 Poisson 分布同属于一个分布族,而 Gamma 分布是 Poisson 分布的共轭先验,这使得用贝叶斯的方式学习模型参数非常方便.另外, Gamma 先验非常适用于用户、活动和信任系数的稀疏表示.具体来说,通过设置形状参数(即, λ_{ua} , λ_{ia} , $\lambda_{\delta a}$ 和 λ_{ta}) 为一个很小的数值(在本文中为 0.3),绝大多数产生值会接近于 0. λ_{ub} , λ_{ib} , $\lambda_{\delta b}$ 和 λ_{tb} 为对应的尺度参数.

活动内容也是吸引用户进行参与的重要因素.假设每个关键词 w 对应一个 K 维向量 ξ_w ,服从参数为 λ_{wa} 和 λ_{wb} 的 Gamma 先验分布.用户 u 的所有历史响应活动中关键词 w 出现的频率 c_{uw} 服从泊松分布,即:

$$c_{uw} \sim \text{Poisson}(\theta_u^T \xi_w) \quad (4)$$

其中,公式(3)和公式(4)共享用户潜在特征向量 θ_u .

基于前面的描述,PF-HSC 模型的概率图模型表示如图 2 所示.

PF-HSC 模型的产生过程描述如下.

Step 1. 对用户 $u=1, \dots, U$ 和组件 $k=1, \dots, K$, 抽样潜在因子 $\theta_{uk} \sim \text{Gamma}(\lambda_{ua}, \lambda_{ub})$.

Step 2. 对活动 $i=1, \dots, I$ 和组件 $k=1, \dots, K$, 抽样潜在因子 $\beta_{ik} \sim \text{Gamma}(\lambda_{ia}, \lambda_{ib})$.

Step 3. 对用户 u 的线上信任网络中的好友 $v=1, \dots, V$, 抽样潜在因子 $\delta_{uv}^{on} \sim \text{Gamma}(\lambda_{\delta a}, \lambda_{\delta b})$.

- Step 4. 对用户 u 的线下信任网络中的好友 $f=1, \dots, F$, 抽样潜在因子 $\tau_{uf}^{off} \sim \text{Gamma}(\lambda_{\tau_a}, \lambda_{\tau_b})$.
 - Step 5. 对观察到的(用户 u , 活动 i), 用公式(3)抽样响应 y_{ui} .
 - Step 6. 对观察到的(用户 u , 词 w), 用公式(4)抽样词频 c_{uw} .
- PF-HSC 模型中使用的主要符号见表 1.

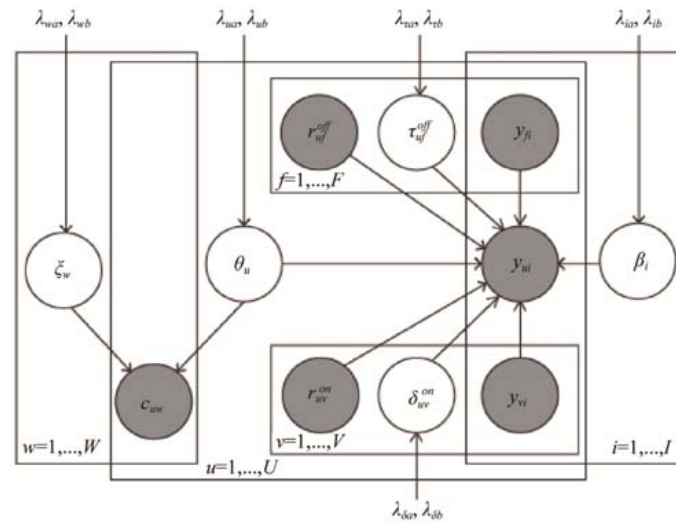


Fig.2 Probabilistic graphical model of the PF-HSC model

图 2 PF-HSC 模型的概率图模型表示

Table 1 The key notations used in PF-HSC model

表 1 PF-HSC 模型中使用的主要符号

符号	描述
u, U	用户, 用户的数量
i, I	活动, 活动的数量
v, V	线上好友, 线上好友的数量
f, F	线下好友, 线下好友的数量
w, W	词, 词典的长度
θ_u, β_i	用户 u 的 K 维偏好向量, 活动 i 的 K 维属性向量
$\delta_{uv}^{on}, r_{uv}^{on}$	线上信任网络中用户 u 和好友 v 的信任系数, 信任强度
$\tau_{uf}^{off}, r_{uf}^{off}$	线下信任网络中用户 u 和好友 f 的信任系数, 信任强度
y_{ui}	用户 u 对活动 i 的响应
$\lambda_{*a}, \lambda_{*b}$	Gamma 分布中的形状(shape)参数和尺度(rate)参数

5 参数推断和活动推荐

为了使用 PF-HSC 模型进行推荐, 需要解决的关键问题是后验推断. 所有潜在因子的 Gamma 先验记作 λ . 给定 $\Theta = \{\theta_u, \beta_i, \delta_{uv}^{on}, \tau_{uf}^{off}, \xi_w\}$ 和 $Y = \{y_{ui}, r_{uv}^{on}, r_{uf}^{off}, c_{uw}\}$, 那么后验分布为

$$p(\Theta | Y, \lambda) = \frac{p(\Theta, Y | \lambda)}{p(Y | \lambda)} = \frac{p(\Theta, Y | \lambda)}{\int p(\Theta, Y | \lambda) d\Theta} \quad (5)$$

但公式(5)不能被准确地推断. 本文采用变分法推断来近似后验分布. 因为变分法相对于马尔可夫链蒙特卡罗方法具有较好的扩展性. 变分法推断首先定义一个隐变量上的有参分布族, 即 $q(\Theta | \gamma)$, 来逼近后验, 然后估计近似于 $p(\Theta | Y, \lambda)$ 的 $q(\Theta | \gamma)$ 的参数.

基于凸性的变分法推导的基本思想是利用 Jensen 不等式来获取对数似然的证据下限^[42].

$$\begin{aligned}
 \log p(Y|\lambda) &= \log \int \frac{q(\Theta|\gamma)}{q(\Theta|\gamma)} p(\Theta, Y|\lambda) d\Theta \\
 &\geq \int q(\Theta|\gamma) \log p(\Theta, Y|\lambda) d\Theta - \int q(\Theta|\gamma) \log q(\Theta|\gamma) d\Theta \\
 &= E_q[\log p(\Theta, Y|\lambda)] - E_q[\log q(\Theta|\gamma)] \\
 &= \Delta(\gamma; \lambda)
 \end{aligned} \tag{6}$$

公式(6)中, $\log p(Y|\lambda)$ 和 $\Delta(\gamma; \lambda)$ 的差别是 $q(\Theta|\gamma)$ 和 $p(\Theta|Y, \lambda)$ 的 KL 散度, 即,

$$\log p(Y|\lambda) = \Delta(\gamma; \lambda) + D(q(\Theta|\gamma) \| p(\Theta|Y, \lambda)) \tag{7}$$

公式(7)表明, 最大化证据下限 $\Delta(\gamma; \lambda)$ 等价于最小化 $q(\Theta|\gamma)$ 和 $p(\Theta|Y, \lambda)$ 的 KL 散度, 这样, 后验推断问题即转化为一个优化问题.

5.1 辅助变量

本文利用条件共轭模型的结论来进行 PF-HSC 模型的推导^[43]. 为便于公式推导和算法描述, 需要增加一些辅助变量, 从而产生了一个 PF-HSC 模型的变形公式. 没有这些辅助变量, PF-HSC 模型不是一个条件共轭模型.

多个独立的泊松随机变量的和仍服从泊松分布, 且其尺度参数等于各个独立的泊松分布的尺度值之和. 分别为(用户 u , 活动 i)、(用户 u , 线上好友 v)、(用户 u , 线下好友 f)和(用户 u , 词 w), 引入辅助变量 z_{uik}^M , z_{uiv}^{on} , z_{uif}^{off} 和 z_{uwk}^W , 则公式(3)和公式(4)可分别变换为

$$y_{ui} | y_{-ui} = \sum_{k=1}^K z_{uik}^M + \sum_{v=1}^V z_{uiv}^{on} + \sum_{f=1}^F z_{uif}^{off} \tag{8}$$

$$c_{uw} = \sum_{k=1}^K z_{uwk}^W \tag{9}$$

其中, $z_{uik}^M \sim \text{Poisson}(\theta_{uk} \beta_{ik})$, $z_{uiv}^{on} \sim \text{Poisson}(\delta_{uv}^{on} r_{uv}^{on} y_{vi})$, $z_{uif}^{off} \sim \text{Poisson}(\tau_{uf}^{off} r_{uf}^{off} y_{fi})$ 和 $z_{uwk}^W \sim \text{Poisson}(\theta_{uk} \xi_{wk})$, 且 $V = |N^{on}(u)|$, $F = |N^{off}(u)|$.

增加辅助变量之后, 变分分布 $q(\Theta|\gamma)$ 转变成 $q(\Theta, Z|\gamma, \phi)$, 其中, Z 和 ϕ 分别表示所有增加的辅助变量和参数.

5.2 平均场变分族

在平均场变分族中, 每个隐变量是独立的, 并且仅受其自有变分参数的控制. 为便于描述, 这里忽略了参数 γ 和 ϕ , 平均场分布族为

$$q(\Theta, Z) = \prod_{u,k} q(\theta_{uk}) \prod_{i,k} q(\beta_{ik}) \prod_{u,v} q(\delta_{uv}^{on}) \prod_{u,f} q(\tau_{uf}^{off}) \prod_{w,k} q(\xi_{wk}) \prod_{u,i,k} q(z_{uik}^M) \prod_{u,i,v} q(z_{uiv}^{on}) \prod_{u,i,f} q(z_{uif}^{off}) \prod_{u,w,k} q(z_{uwk}^W) \tag{10}$$

平均场分布族中的每个因子的分布被设置为与其完全条件(complete conditional)的分布一致^[12,44]. 首先计算完全条件^[45], 然后推导出变分参数. 完全条件的计算公式参见附录. 因为 θ_{uk} , β_{ik} , δ_{uv}^{on} , τ_{uf}^{off} 和 ξ_{wk} 的完全条件是 Gamma 分布, 所以, 它们的变分参数是对应的 Gamma 分布的参数, 即, $\theta_{uk} \sim \text{Gamma}(\gamma_{uk}^{shp}, \gamma_{uk}^{rte})$, $\beta_{ik} \sim \text{Gamma}(\gamma_{ik}^{shp}, \gamma_{ik}^{rte})$, $\delta_{uv}^{on} \sim \text{Gamma}(\gamma_{uv}^{shp}, \gamma_{uv}^{rte})$, $\tau_{uf}^{off} \sim \text{Gamma}(\gamma_{uf}^{shp}, \gamma_{uf}^{rte})$ 和 $\xi_{wk} \sim \text{Gamma}(\gamma_{wk}^{shp}, \gamma_{wk}^{rte})$. 相似地, 可以推导出 z_{uwk}^W 和 z_{ui} 的变分分布, 即 $z_{uwk}^W \sim \text{Mult}(c_{uw}, \phi_{uw}^W)$ 和 $z_{ui} \sim \text{Mult}(y_{ui}, \phi_{ui})$, 其中, $\phi_{ui} \propto \langle \phi_{ui}^M, \phi_{ui}^{on}, \phi_{ui}^{off} \rangle$.

根据条件共轭模型, 每个变分参数等于完全条件中对应参数的期望^[43]. 对于变分 Gamma 分布, 本文以 θ_{uk} 为例推导其变分参数的闭包更新公式.

$$\gamma_{uk}^{shp} = E_q \left[\lambda_{ua} + \sum_i z_{uik}^M + \sum_w z_{uwk}^W \right] = \lambda_{ua} + \sum_i y_{ui} \phi_{ui}^M + \sum_w c_{uw} \phi_{uw}^W \tag{11}$$

$$\gamma_{uk}^{rte} = E_q \left[\lambda_{ub} + \sum_i \beta_{ik} + \sum_w \xi_{wk} \right] = \lambda_{ub} + \sum_i \frac{\gamma_{ik}^{shp}}{\gamma_{ik}^{rte}} + \sum_w \frac{\gamma_{wk}^{shp}}{\gamma_{wk}^{rte}} \tag{12}$$

β_{ik} , δ_{uv}^{on} , τ_{uf}^{off} 和 ξ_{wk} 的变分参数更新公式可以相似地推导出来, 结果如下:

$$\gamma_{ik}^{shp} = \lambda_{ia} + \sum_u y_{ui} \phi_{ui}^M \tag{13}$$

$$\gamma_{ik}^{rte} = \lambda_{ib} + \sum_u \frac{\gamma_{uk}^{shp}}{\gamma_{uk}^{rte}} \tag{14}$$

$$\gamma_{uv}^{shp} = \lambda_{va} + \sum_i \gamma_{ui} \phi_{uiv}^{on} \tag{15}$$

$$\gamma_{uv}^{rte} = \lambda_{vb} + \sum_i r_{uv}^{on} y_{vi} \tag{16}$$

$$\gamma_{uv}^{rte} = \lambda_{vb} + \sum_i r_{uv}^{on} y_{vi} \tag{17}$$

$$\gamma_{uf}^{rte} = \lambda_{fb} + \sum_i r_{uf}^{off} y_{fi} \tag{18}$$

$$\gamma_{wk}^{shp} = \lambda_{va} + \sum_u c_{uw} \phi_{uwk}^W \tag{19}$$

$$\gamma_{wk}^{rte} = \lambda_{wb} + \sum_u \frac{\gamma_{uk}^{shp}}{\gamma_{uk}^{rte}} \tag{20}$$

对于变分多项式分布 $z_{uw}^W \sim Mult(c_{uw}, \phi_{uw}^W)$ 的参数 ϕ_{uw}^W 和 $z_{ui} \sim Mult(y_{ui}, \phi_{ui})$ 的参数 ϕ_{ui} :

$$\phi_{ui} \propto \langle \phi_{ui}^M, \phi_{ui}^{on}, \phi_{ui}^{off} \rangle \tag{21}$$

本文以 $\phi_{ui}^M \propto \langle \phi_{ui1}^M, \dots, \phi_{uiK}^M \rangle$ 为例推导其变分参数的更新公式,推导如下:

$$\left. \begin{aligned} \phi_{uik}^M &= G_q[\theta_{ik} \beta_{ik}] \\ &= \exp\{E_q[\log \theta_{ik} + \log \beta_{ik}]\} \\ &= \exp\{\Psi(\gamma_{uk}^{shp}) - \log \gamma_{uk}^{rte} + \Psi(\gamma_{ik}^{shp}) - \log \gamma_{ik}^{rte}\} \end{aligned} \right\} \tag{22}$$

其中, $u=1, \dots, U; i=1, \dots, I; k=1, \dots, K$; $\Psi(\bullet)$ 是 digamma 函数, $G_q[\bullet] = \exp\{E_q[\log(\bullet)]\}$ 表示几何期望^[46].同理,

$$\phi_{uiv}^{on} = \exp\{\Psi(\gamma_{uv}^{shp}) - \log \gamma_{uv}^{rte} + \log(\gamma_{uv}^{on} y_{vi})\} \tag{23}$$

$$\phi_{uif}^{off} = \exp\{\Psi(\gamma_{uf}^{shp}) - \log \gamma_{uf}^{rte} + \log(\gamma_{uf}^{off} y_{fi})\} \tag{24}$$

$$\phi_{uwk}^W = \exp\{\Psi(\gamma_{uk}^{shp}) - \log \gamma_{uk}^{rte} + \Psi(\gamma_{wk}^{shp}) - \log \gamma_{wk}^{rte}\} \tag{25}$$

其中, $v=1, \dots, V; f=1, \dots, F; w=1, \dots, W$. $\phi_{ui} \propto \langle \phi_{ui}^M, \phi_{ui}^{on}, \phi_{ui}^{off} \rangle$ 和 ϕ_{uw}^W 应该分别进行归一化处理.

5.3 坐标上升算法

令 $Y = \{y_{ui}\}$ 表示用户对活动的响应数据, $R^{on} = \{r_{uv}^{on}\}$ 表示预计算的用户之间的线上信任强度, $R^{off} = \{r_{uf}^{off}\}$ 表示预计算的用户之间的线下信任强度, $C = (c_{uw})$ 表示预统计出的用户响应活动中关于活动内容的关键词词频, 则使用坐标上升算法更新变分参数, 从而获得局部最优解. 如算法 1 所示.

算法 1. PF-HSC 模型的变分推断算法.

输入: Y, R^{on}, R^{off}, C , 超参数 $\lambda = \{\lambda_{ua}, \lambda_{ub}, \lambda_{ia}, \lambda_{ib}, \lambda_{\delta a}, \lambda_{\delta b}, \lambda_{\tau a}, \lambda_{\tau b}, \lambda_{wa}, \lambda_{wb}\}$;

输出: 变分参数 $\gamma = \{\gamma_{uk}^{shp}, \gamma_{uk}^{rte}, \gamma_{ik}^{shp}, \gamma_{ik}^{rte}, \gamma_{uv}^{shp}, \gamma_{uv}^{rte}, \gamma_{uf}^{shp}, \gamma_{uf}^{rte}, \gamma_{wk}^{shp}, \gamma_{wk}^{rte}\}$.

- | | |
|---------|--|
| Step 1. | repeat |
| Step 2. | For each response y_{ui} |
| Step 3. | 分别用公式(22)~公式(24)更新 ϕ_{uik}^M 、 ϕ_{uiv}^{on} 和 ϕ_{uif}^{off} |
| Step 4. | 对公式(21)中的 ϕ_{ui} 进行归一化处理 |
| Step 5. | For each term frequency c_{uw} |
| Step 6. | 用公式(25)更新 ϕ_{uwk}^W |
| Step 7. | 对 ϕ_{uwk}^W 进行归一化处理 |
| Step 8. | For each user u |

- Step 9. 分别用公式(11)和公式(12)更新 γ_{uk}^{shp} 和 γ_{uk}^{rte}
- Step 10. 分别用公式(15)和公式(16)更新 γ_{uv}^{shp} 和 γ_{uv}^{rte}
- Step 11. 分别用公式(17)和公式(18)更新 γ_{uf}^{shp} 和 γ_{uf}^{rte}
- Step 12. For each item i
- Step 13. 分别用公式(13)和公式(14)更新 γ_{ik}^{shp} 和 γ_{ik}^{rte}
- Step 14. For each term w
- Step 15. 分别用公式(19)和公式(20)更新 γ_{wk}^{shp} 和 γ_{wk}^{rte}
- Step 16. until Converge

算法 1 的时间复杂度主要取决于用户对活动的响应数量 $N(N \ll U \times I$ 且 $N > W)$ 、向量维度 K 、用户线上好友的数量 V 以及用户线下好友的数量 F 。最后,令 $M = \max\{K, V, F\}$,则算法 1 的时间复杂度为 $O(N \times M)$ 。算法 1 对于稀疏的响应数据是非常有效的。

5.4 活动推荐

PF-HSC 模型是一个基于隐参数 $(\theta_u, \beta_i, \delta_{uv}^{on}, \tau_{uf}^{off}$ 和 $\xi_w)$ 和观测数据 $(c_{uv}, r_{uv}^{on}, r_{uf}^{off}$ 和 $y_{ui})$ 的贝叶斯概率模型。为了给用户推荐活动,需要根据给定的观测数据来估计隐参数的后验分布。一旦后验分布被确定,就可以用公式(26)来预测用户 u 对每一个未响应的活动 i 的评分,然后按照评分进行排序,最后向用户推荐 top- n 个活动。

$$\hat{y}_{ui} = E_q \left[\theta_u^T \beta_i + \sum_{v \in N^{on}(u)} \delta_{uv}^{on} r_{uv}^{on} y_{vi} + \sum_{f \in N^{off}(u)} \tau_{uf}^{off} r_{uf}^{off} y_{fi} \right] = \sum_{k=1}^K \frac{\gamma_{uk}^{shp}}{\gamma_{uk}^{rte}} \frac{\gamma_{ik}^{shp}}{\gamma_{ik}^{rte}} + \sum_{v \in N^{on}(u)} \frac{\gamma_{uv}^{shp}}{\gamma_{uv}^{rte}} r_{uv}^{on} y_{vi} + \sum_{f \in N^{off}(u)} \frac{\gamma_{uf}^{shp}}{\gamma_{uf}^{rte}} r_{uf}^{off} y_{fi} \quad (26)$$

6 实验

本节首先介绍所选用的数据集,之后给出实验方法和指标,然后介绍现有的模型,最后给出实验结果,并对结果进行分析。

6.1 数据集

本文选择 Meetup 数据集^[37]来评估模型的性能。Meetup 数据集的时间跨度为 2010 年 1 月~2014 年 4 月。包含用户所加入的组信息、活动的内容、用户对活动的响应等数据。如图 3 所示,大部分用户只参与很少的活动,大部分活动只会被很少的用户参与。

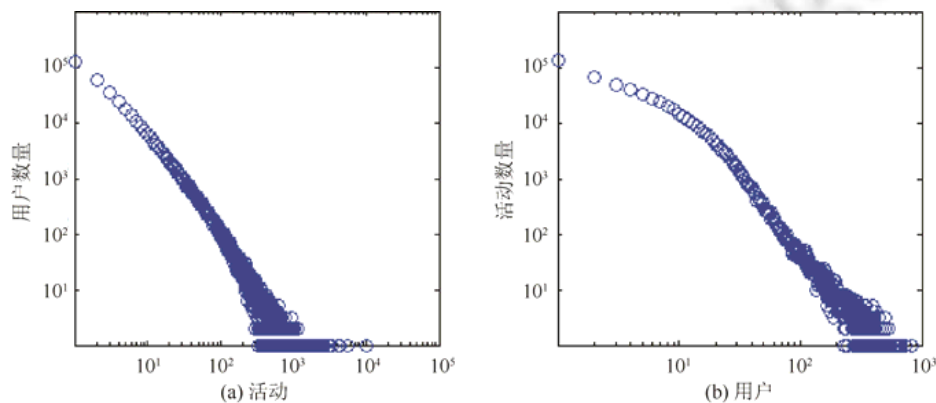


Fig.3 Events and users distributions in Meetup dataset

图 3 Meetup 数据集中活动和用户的分布

如图 4 所示,大部分用户只参与很少的兴趣小组,大部分兴趣小组只会被很少的用户加入。

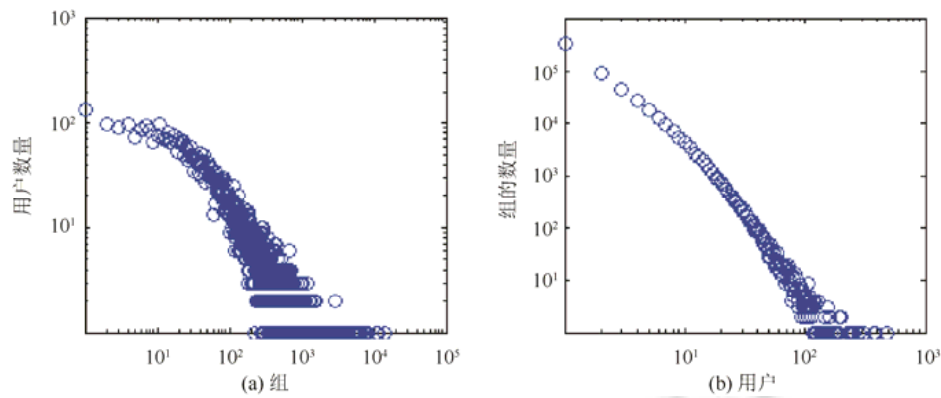


Fig.4 Groups and users distributions in Meetup dataset

图4 Meetup 数据集中兴趣小组和用户的分布

因为用户倾向于参加本地的活动^[13],所以,本文根据用户对本地活动的历史响应数据,向用户推荐本地正在进行的其他活动.选择数据量相对较大的3个城市(Chicago、Phoenix 和 San Jose)评估模型的性能.相关的统计信息见表2.

Table 2 Statistics of the Meetup dataset

表2 Meetup 数据集的统计信息

城市	组的数量	用户的数量	活动的数量	响应的数量
Chicago	2 138	133 357	100 701	810 213
Phoenix	842	43 112	64 255	326 913
San Jose	779	40 219	31 670	219 475

6.2 实验方法和指标

本文分隔每个城市的数据为10份训练集和测试集.对于每个城市,具体做法是,最后10个月的数据分别作为测试集,并从测试集中去除了冷启动活动,对应测试集起始时间的前6个月的数据作为训练集.首先计算每一个用户对每一个活动的评分,然后根据评分值的大小按用户对活动进行排序,最后向每个用户推荐top- n 个正在进行的活动.本文选择在top- n 推荐中最常用的准确率(precision)和归一化折损累积增益(NDCG)作为评价指标,即, $P@3$ 、 $P@5$ 、 $NDCG@3$ 和 $NDCG@5$.

$$P@n = \frac{\#Hits}{n},$$

其中, $\#Hits$ 是top- n 个推荐的活动中用户实际参与的活动的个数.

$$NDCG@n = \frac{1}{IDCG} \times \sum_{i=1}^n \frac{2^{r_i} - 1}{\log(i+1)},$$

其中,如果用户实际参与了第 i 个活动,则 r_i 被置为1,否则, r_i 被置为0.IDCG用于归一化,使得在最优的排序结果下, $NDCG@n$ 的值为1.

6.3 对比模型

MP:根据活动的流行度向用户推荐本地最热门的活动.具体来说,统计训练集中正在进行的每个活动获得的响应数量,对活动按获得的响应数量进行降序排序,然后向用户推荐top- n 个活动.

BPR^[32]:BPR模型应用于活动推荐,则强调用户响应活动的顺序关系,即,用户对于任意一个响应过的活动的偏好大于未响应活动的偏好,从而得到所有活动的偏好关系.

BPF^[12]:BPF建模用户对活动的响应为一个泊松分布,是一个贝叶斯泊松因子分解模型.能够有效地处理稀疏数据,避免过拟合,且尤其适合处理隐性反馈数据.BPF模型只对用户和活动建模,没有混入社交网络、内容等上下文.

SPF^[29]:SPF模型可以把用户对活动的响应以及用户的社交网络用贝叶斯泊松因子分解的方式建模到一个

框架.但 SPF 模型只能建模一种社交关系.在本文后面的实验中,只把线上社交关系混入 SPF 模型.另外,SPF 没有探索信任强度和内容信息.

FM^[47]:FM 是一个通用的上下文感知推荐模型,可以很方便地对各种上下文进行建模.本文中把用户、活动、社交网络好友及信任强度、内容信息中的词频等作为上下文混入到 FM 模型中.

LDA^[48]:计算训练集中用户参与的活动和测试集中活动的内容相似度.具体来说,应用 LDA 模型把每个活动映射为面向主题的分布,然后应用 Jensen-Shannon 散度计算测试集中的活动与训练集中用户参与的活动内容的相似度,最后向用户推荐相似度最高的 top- n 个活动.

HSPF^[38]:HSPF 模型探索了 EBSNs 中异构的社交关系.但 HSPF 模型中简单地使用用户共同加入的兴趣组数量来表示线上信任强度,用用户共同参与的活动的数量来表示线下信任强度.另外,HSPF 模型没有利用活动的内容信息.

CBPF^[13]:CBPF 模型用协同因子分解的方式建模社交网络、组织者、位置和内容信息.其中,用户-好友和用户-活动共享用户潜在向量.活动的组织者、位置和内容信息加权后得到活动的潜在向量.

PF-HSC\H:本文提出的 PF-HSC 模型的第 1 个变种.PF-HSC\H 模型表示从 PF-HSC 模型中去除了异构社交关系及相应的用户之间的信任强度.仅保留了用户-活动响应和用户-词交互,其中,用户是共享的潜在向量.类似于 CMF^[49],但并不基于高斯似然.

PF-HSC\S:本文提出的 PF-HSC 模型的第 2 个变种.相对于 PF-HSC 模型,PF-HSC\S 模型去除了好友之间的信任强度.

PF-HSC\C:本文提出的 PF-HSC 模型的第 3 个变种.PF-HSC\C 模型只对异构的社交网络进行建模,去除了活动的内容.类似于 HSPF^[38],但是信任强度的计算方式有所不同.

6.4 实验结果

根据文献[44]的设置,将 Gamma 先验的形状(λ_a)和尺度(λ_b)超参数都固定为 0.3.对应的变分参数 γ^{shp} 和 γ^{re} 都初始化为在 0.3 上加上一个很小的随机数.对于所有的模型,本文选择潜在向量的维度 $K=50$.

6.4.1 模型的有效性

模型之间性能对比的实验结果如图 5 所示.从图 5 可以观察到:

基于 3 个城市的数据,本文提出 PF-HSC 模型在所有的指标上都取得了最佳的性能.这说明,PF-HSC 模型在基于活动的社交网络中进行活动推荐具有一定的优势.

PF-HSC、HSPF 和 FM 模型战胜了 BPF、SPF 等模型,因为 PF-HSC、HSPF 和 FM 模型都利用了线上和线下的异构社交关系.而 BPF 仅探索了用户和活动二者之间的交互,SPF 在 BPF 的基础上探索了线上社交关系,没有探索线下社交关系、信任强度、内容等信息.

PF-HSC 相对于 HSPF 取得了更好的性能,这是因为:(1) PF-HSC 模型使用了更为合理的信任强度计算方法;(2) PF-HSC 模型融入了内容信息.而由于数据的极其稀疏及 FM 模型不能很好地建模用户的隐性反馈,使得 FM 模型取得了相对较差的性能.

尽管 CBPF 模型同时探索了用户对活动的隐性反馈,用户之间线上的社交网络,活动的内容、组织方和位置等信息,但本文提出的 PF-HSC 模型依然战胜了 CBPF 模型.这是因为,相对于 CBPF 模型,PF-HSC 利用了异构的社交网络,并把用户之间的信任强度融入到框架中.另外,根据文献[13]中的研究结论,相对于活动的内容和位置,活动的组织者在活动推荐中起着更重要的作用.而活动的组织者经常会周期性地发布一些相同或相似的活动,比如,某健身俱乐部成员会在每个周末组织瑜伽活动,关于活动的内容信息几乎是一样的.也就是说,活动的内容能在一定程度上体现活动组织者的态度和思想.不同于 CBPF 模型中把活动的内容作为活动因子的一个加权项,PF-HSC 模型把用户作为用户-活动交互和用户-词交互的共享因子,更加突出了内容信息对提升推荐性能所起的作用.

BPR 模型的性能稍逊于 BPF 模型,这是因为,在 Meetup 数据集中正例和负例的比例极其不平衡.同时也说

明,贝叶斯泊松分解模型更适合于建模用户对活动的隐性反馈.而 MP 导致了最差的性能,这说明,小范围的热门活动并不能吸引大量用户的兴趣.而 LDA 的性能优于 MP,也说明,在 PF-HSC 模型中融入内容信息有助于提高模型的性能.

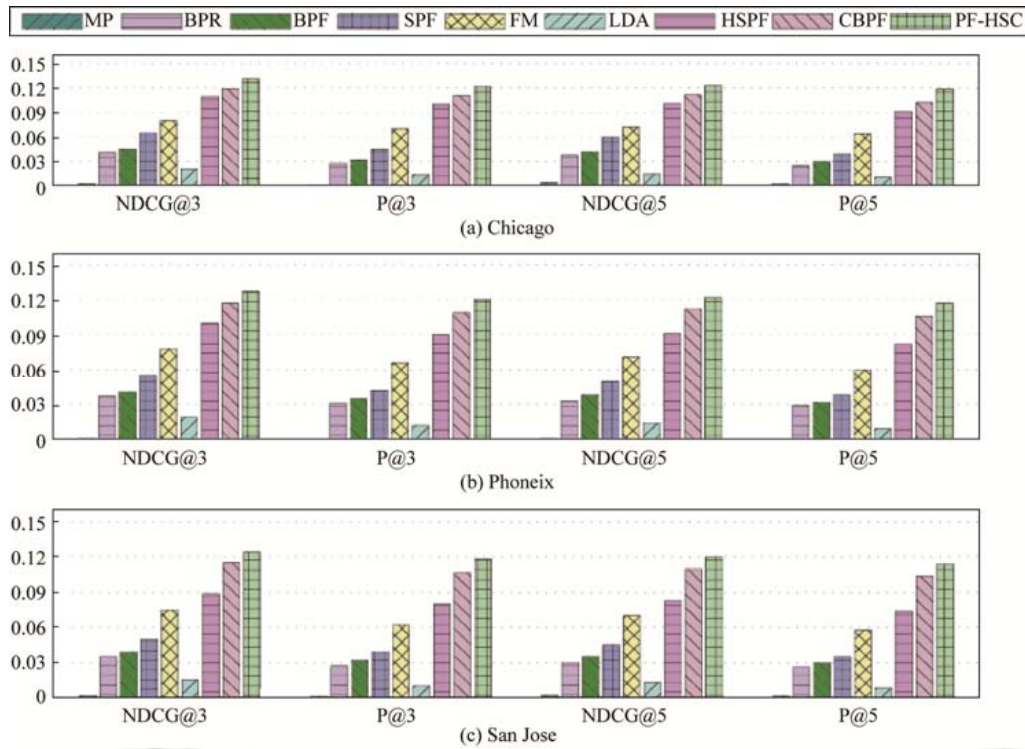


Fig.5 Comparisons of different methods on NDCG@n and P@n

图5 不同模型关于 NDCG@n 和 P@n 指标的性能对比

6.4.2 模型性能影响因素分析

本文提出的 PF-HSC 模型把线上和线下社交网络、信任强度、活动内容融合到一个框架中,相对于其他的模型,取得了最优的性能.每一个因素对模型性能的影响如图 6 所示.

从图 6 可以观察到:

PF-HSC\H 模型的性能最差.这说明,在基于活动的社交网络中进行活动推荐,用户之间异构的社交关系及信任强度是最重要的因素.

PF-HSC\S 模型稍微优于 PF-HSC\C 模型.这说明,相对于用户之间的信任强度,活动的内容信息更有助于提升推荐的性能.

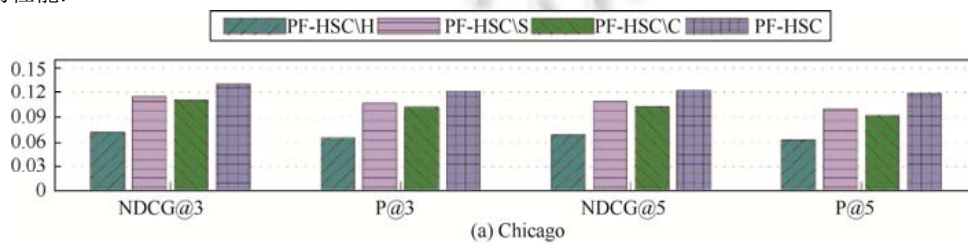


Fig.6 Factor contribution of the PF-HSC model

图6 PF-HSC 模型的影响因素

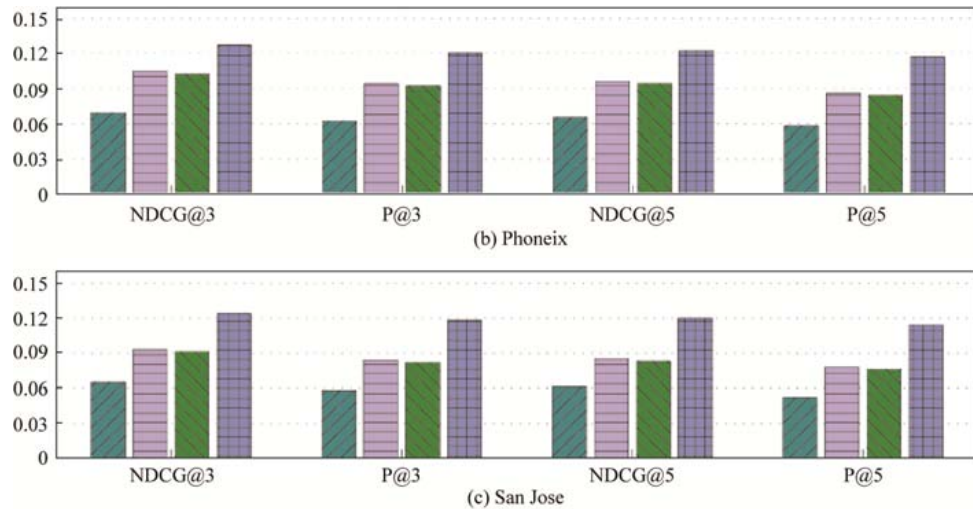


Fig.6 Factor contribution of the PF-HSC model (Continued)

图 6 PF-HSC 模型的影响因素(续)

6.4.3 不同维度下模型性能的比较与分析

潜在因子向量的维度 K 是 PF-HSC 模型的一个重要参数,不同的取值会导致不同的性能.通过变化 K 的取值,可以测试其对模型的性能的影响. K 的取值分别为 30,40,50,60 和 70.实验结果如图 7 所示.从图 7 可以观察到,PF-HSC 模型的性能随着 K 值的变大逐渐提升.然而,当 K 的取值大于某个阈值时,模型性能的提升变得不再显著.这是因为, K 值对应于模型的复杂度,当 K 值太小时,模型不能很好地表示数据;当 K 值达到一个阈值后,模型就可以很好地表示数据.

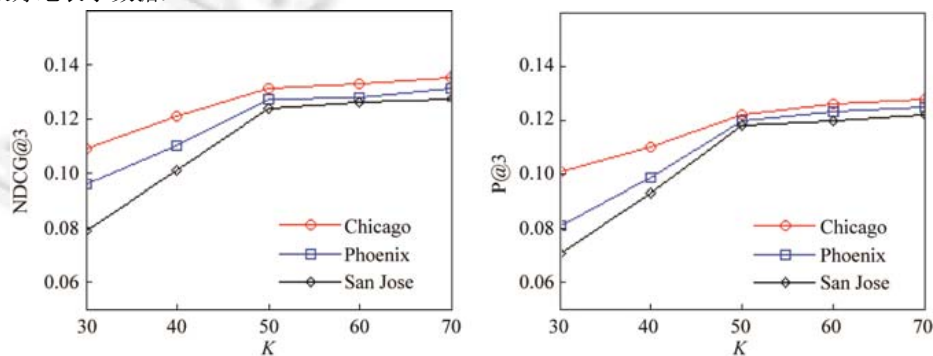


Fig.7 Impact of the latent dimensionality of the PF-HSC model

图 7 PF-HSC 模型中不同维度的影响

7 结论与下一步研究

针对于基于活动的社交网络下的活动推荐问题,本文提出一个基于异构社交和内容信息的泊松因子分解模型(PF-HSC).PF-HSC 模型联合建模异构的社交网络和内容信息到一个框架中,并且混入了用户之间的信任强度.在 Meetup 数据集上的实验结果表明,本文提出的 PF-HSC 模型的性能优于其他模型.当然,本模型也有不足之处,例如,不能处理完全冷启动问题.

针对于活动推荐,在未来还有很多可以进一步研究的问题.例如,(1) 探索把时间、地点、组织者等信息融入到框架中,使得推荐的性能进一步提高,并期望较好地应对冷启动问题;(2) 利用迁移学习^[50]或主题模型^[51]的方式探索跨城市的活动推荐;(3) 在分布式平台(如 Spark)上实现快速的活动推荐.

References:

- [1] Liu X, He Q, Tian Y, *et al.* Event-Based social networks: Linking the online and offline social worlds. In: Proc. of the KDD. ACM Press, 2012. 1032–1040.
- [2] Qiao Z, Zhang P, Cao Y, Zhou C, Guo L, Fang B. Combining heterogenous social and geographical information for event recommendation. In: Proc. of the AAAI. AAAI Press, 2014. 145–151.
- [3] Koren Y, Bell R, Volinsky C. Matrix factorization techniques for recommender systems. *Computer*, 2009,42(8):30–37.
- [4] Liu Q, Wu S, Wang L. COT: Contextual operating tensor for context-aware recommender systems. In: Proc. of the AAAI. AAAI Press, 2015. 203–209.
- [5] Chen T, Zheng Z, Lu Q, *et al.* Informative ensemble of multi-resolution dynamic factorization models. In: Proc. of the KDD-Cup Workshop. ACM Press, 2011.
- [6] Craw S, Horsburgh B, Massie S. Music recommenders: User evaluation without real users. In: Proc. of the IJCAI. AAAI Press, 2015.
- [7] Das S, Datar M, Garg A, Rajaram S. Google news personalization: Scalable online collaborative filtering. In: Proc. of the 16th Int'l Conf. on World Wide Web (WWW 2007). ACM Press, 2007. 271–280.
- [8] Wang WQ, Yin HZ, Chen L, *et al.* Geo-Sage: A geographical sparse additive generative model for spatial item recommendation. In: Proc. of the 21st ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining (KDD 2015). 2015. 1255–1264.
- [9] McPherson M, Smith-Lovin L, Cook J. Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 2001,27(1): 415–444.
- [10] Marsden P, Friedkin N. Network studies of social influence. *Sociological Methods and Research*, 1993,22(1):127–151.
- [11] Tang JL, Hu X, Gao HJ, Liu H. Exploiting local and global social context for recommendation. In: Proc. of the IJCAI. 2013. Morgan Kaufmann Publishers, 264–269.
- [12] Gopalan P, Hofman JM, Blei DM. Scalable recommendation with poisson factorization. arXiv Preprint arXiv, 2013:1311.1704.
- [13] Zhang W, Wang J. A collective Bayesian Poisson factorization model for coldstart local event recommendation. In: Proc. of the 21st ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining (KDD 2015). ACM Press, 2015. 1455–1464.
- [14] 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. <http://www.jos.org.cn/1000-9825/4831.htm> [doi: 10.13328/j.cnki.jos.004831]
- [15] Yang X, Guo Y, Liu Y, Steck H. A survey of collaborative filtering based social recommender systems. *Computer Communications*, 2014,(41):1–10.
- [16] Ma H, Yang H, Lyu MR, King I. Sorec: Social recommendation using probabilistic matrix factorization. In: Proc. of the 17th ACM Conf. on Information and Knowledge Management (CIKM 2008). ACM Press, 2008. 931–940.
- [17] Ma H, King I, Lyu M. Learning to recommend with social trust ensemble. In: Proc. of the SIGIR. ACM Press, 2009. 203–210.
- [18] Jamali M, Ester M. A matrix factorization technique with trust propagation for recommendation in social networks. In: Proc. of the 4th ACM Conf. on Recommender Systems (RecSys 2010). ACM Press, 2010. 135–142.
- [19] Ma H, Zhou D, Liu C, *et al.* Recommender systems with social regularization. In: Proc. of the 4th ACM Int'l Conf. on Web Search and Data Mining (WSDM 2011). ACM Press, 2011. 287–296.
- [20] Guo L, Ma J, Chen Z. Learning to recommend with multi-faceted trust in social networks. In: Proc. of the 22nd Int'l Conf. on World Wide Web (WWW 2013). ACM Press, 2013. 205–206.
- [21] Park C, Kim D, Oh J, Yu H. Improving top-K recommendation with truster and trustee relationship in user trust network. *Information Sciences*, 2016,(374):100–114.
- [22] Tang JL, Gao HJ, Hu X, Liu H. Context-Aware review helpfulness rating prediction. In: Proc. of the RecSys. ACM Press, 2013. 1–8.
- [23] Ma H, King I, Lyu MR. Learning to recommend with explicit and implicit social relations. *ACM Trans. on Intelligent Systems and Technology (TIST)*, 2011,(2):1–19.
- [24] Tang JL, Wang SH, Hu X, *et al.* Recommendation with social dimensions. In: Proc. of the AAAI. AAAI Press, 2016. 251–257.
- [25] Xiang RJ, Neville J, Rogati M. Modeling relationship strength in online social networks. In: Proc. of the 19th Int'l Conf. on World Wide Web (WWW 2010). ACM Press, 2010. 981–990.
- [26] Zou BY, Li CP, Tan LW, Chen H, Wang SQ. Social recommendations based on user trust and tensor factorization. *Ruan Jian Xue Bao/Journal of Software*, 2014,25(12):2852–2864 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/4725.htm> [doi: 10.13328/j.cnki.jos.004725]
- [27] Hu X, Meng XW, Zhang YJ, Shi YC. Recommendation algorithm combing item features and trust relationship of mobile users. *Ruan Jian Xue Bao/Journal of Software*, 2014,24(8):1817–1830 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/4491.htm> [doi: 10.13328/j.cnki.jos.004491]

- [28] Hung HJ, Shuai HH, Yang DN, *et al.* When social influence meets item inference. In: Proc. of the KDD. 2016.
- [29] Chaney AJ, Blei DM, Eliassi-Rad T. A probabilistic model for using social networks in personalized item recommendation. In: Proc. of the 9th ACM Conf. on Recommender Systems (RecSys 2015). ACM Press, 2015. 43–50.
- [30] Koren Y. Factorization meets the neighborhood: A multifaceted collaborative filtering model. In: Proc. of the 14th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining (KDD 2008). ACM Press, 2008. 426–434.
- [31] Koren Y. Collaborative filtering with temporal dynamics. *Communications of the ACM*, 2010,53(4):89–97.
- [32] Rendle S, Freudenthaler C, Gantner Z, Schmidt-Thieme L. BPR: Bayesian personalized ranking from implicit feedback. In: Proc. of the UAI 2009. AUAI Press, 2009. 452–461.
- [33] Guo W, Wu S, Wang L, *et al.* Multiple attribute aware personalized ranking. In: Proc of the Asia-Pacific Web Conf. (APWeb 2015). Springer Int'l Publishing, 2015. 244–255.
- [34] Guo W, Wu S, Wang L, *et al.* Personalized ranking with pairwise factorization machines. *Neurocomputing*, 2016,214:191–200.
- [35] Du R, Yu Z, Mei T, *et al.* Predicting activity attendance in event-based social networks: Content, context and social influence. In: Proc. of the 2014 ACM Int'l Joint Conf. on Pervasive and Ubiquitous Computing. 2014. 425–434.
- [36] Pham T, Li X, Cong G, Zhang Z. A general graph-based model for recommendation in event-based social networks. In: Proc. of the 31st Int'l Conf. on Data Engineering (ICDE 2015). 2015. 567–578.
- [37] Macedo A, Marinho L, Santos R. Context-Aware event recommendation in event-based social networks. In: Proc. of the 9th ACM Conf. on Recommender Systems (RecSys 2015). ACM Press, 2015. 123–130.
- [38] Wang SQ, Wang Z, Li CP, Zhao KK, Chen H. Learn to recommend local event using heterogeneous social networks. In: Proc. of the 18th Asia Pacific Web Conf. (APWeb 2016). 2016. 169–182.
- [39] Yu YH, Gao Y, Wang H. A ranking based Poisson matrix factorization model for point-of-interest recommendation. *Journal of Computer Research and Development*, 2016,53(8):1651–1663 (in Chinese with English abstract).
- [40] Liu B, Fu YJ, Yao ZJ, Xiong H. Learning geographical preferences for point-of-interest recommendation. In: Proc. of the KDD. ACM Press, 2013. 1043–1051.
- [41] Ma H, Liu C, King I, Lyu M. Probabilistic factor models for Web site recommendation. In: Proc. of the SIGIR. ACM Press, 2011. 265–274.
- [42] Jordan M, Ghahramani Z, Jaakkola TS, Saul LK. An introduction to variational methods for graphical models. *Machine Learning*, 1999,37(2):183–233.
- [43] Hoffman M, Blei D, Wang C, Paisley J. Stochastic variational inference. *Journal of Machine Learning Research*, 2013,14(1):1303–1347.
- [44] Gopalan PK, Charlin L, Blei D. Content-Based recommendations with Poisson factorization. In: Proc. of the Advances in Neural Information Processing Systems (NIPS 2014). 2014. 3176–3184.
- [45] Cemgil AT. Bayesian inference for nonnegative matrix factorisation models. *Computational Intelligence and Neuroscience*, 2009.
- [46] Schein A, Paisley J, Blei DM, Wallach H. Bayesian Poisson tensor factorization for inferring multilateral relations from sparse dyadic event counts. In: Proc. of the 21st ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining (KDD 2015). ACM Press, 2015. 1045–1054.
- [47] Rendle S, Gantner Z, Freudenthaler C, Schmidt-Thieme L. Fast context-aware recommendations with factorization machines. In: Proc. of the 34th Int'l ACM SIGIR Conf. on Research and Development in Information Retrieval (SIGIR 2011). ACM Press, 2011. 635–644.
- [48] Blei D, Ng A, Jordan M. Latent Dirichlet allocation. *Journal of Machine Learning Research*, 2003,(3):993–1022.
- [49] Singh AP, Gordon GJ. Relational learning via collective matrix factorization. In: Proc. of the 14th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining (KDD 2008). ACM Press, 2008. 650–658.
- [50] Pan SJ, Yang Q. A survey on transfer learning. *IEEE Trans. on Knowledge and Data Engineering*, 2010,22(10):1345–1359.
- [51] Yin HZ, Sun YZ, Cui B, *et al.* LCARS: A location-content-aware recommender system. In: Proc. of the 19th ACM SIGKDD Int'l Conf. on Knowledge Discovery and Data Mining (KDD 2013). ACM Press, 2013. 221–229.

附中文参考文献:

- [14] 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> [doi: 10.13328/j.cnki.jos.004831]
- [26] 邹本友,李翠平,谭力文,陈红,王绍卿. 基于用户信任和张量分解的社会网络推荐. *软件学报*, 2014,25(12):2852–2864. <http://www.jos.org.cn/1000-9825/4725.htm> [doi: 10.13328/j.cnki.jos.004725]
- [27] 胡勋,孟祥武,张玉洁,史艳翠. 一种融合项目特征和移动用户信任关系的推荐算法. *软件学报*, 2014,25(8):1817–1830. <http://www.jos.org.cn/1000-9825/4491.htm> [doi: 10.13328/j.cnki.jos.004491]

[39] 余永红,高阳,王皓.基于 Ranking 的泊松矩阵分解兴趣点推荐算法.计算研究与发展,2016,53(8):1651-1663.

附录 1. 完全条件(complete conditional)的计算公式

在模型中,给定观察数据和其他隐变量,一个完全条件是一个隐变量的条件分布.模型中,所有隐变量的完全条件都可以被计算出来.在 PF-HSC 模型中,用户偏好因子 θ_{uk} 、活动的属性因子 β_{ik} 、线上信任系数 δ_{uv}^{on} 、线下信任系数 τ_{uf}^{off} 和每个关键词 w 的向量因子 ξ_{wk} 的完全条件分别是 Gamma 分布,如公式(27)~公式(31)所示.

$$\theta_{uk} | \lambda, \beta, \xi, \delta, \tau, z, y, c \sim \text{Gamma} \left(\lambda_{ua} + \sum_i z_{uik}^M + \sum_w z_{uwk}^W, \lambda_{ub} + \sum_i \beta_{ik} + \sum_w \xi_{wk} \right) \quad (27)$$

$$\beta_{ik} | \lambda, \theta, \xi, \delta, \tau, z, y, c \sim \text{Gamma} \left(\lambda_{ia} + \sum_u z_{uik}^M, \lambda_{ib} + \sum_u \theta_{uk} \right) \quad (28)$$

$$\delta_{uv}^{on} | \lambda, \theta, \beta, \xi, \delta, \tau, z, y, c \sim \text{Gamma} \left(\lambda_{va} + \sum_i z_{uiv}^{on}, \lambda_{vb} + \sum_i r_{uv}^{on} y_{vi} \right) \quad (29)$$

$$\tau_{uf}^{off} | \lambda, \theta, \beta, \xi, \delta, \tau, z, y, c \sim \text{Gamma} \left(\lambda_{fa} + \sum_i z_{uif}^{off}, \lambda_{fb} + \sum_i r_{uf}^{off} y_{fi} \right) \quad (30)$$

$$\xi_{wk} | \lambda, \theta, \beta, \delta, \tau, z, y, c \sim \text{Gamma} \left(\lambda_{wa} + \sum_u z_{uwk}^W, \lambda_{wb} + \sum_u \theta_{uk} \right) \quad (31)$$

因为,多个泊松之和的条件分布是一个多项分布,并且,这个多项分布的参数是对应的多个泊松分布尺度参数的归一化集合^[45].因此, z_{uw}^W 和 $z_{ui} = \langle z_{ui}^M, z_{ui}^{on}, z_{ui}^{off} \rangle$ 的完全条件服从多项分布.也就是说, $z_{uw}^W \sim \text{Mult}(c_{uw}, \psi_{uw}^W)$ 和 $z_{ui} \sim \text{Mult}(y_{ui}, \psi_{ui})$, 其中,

$$\psi_{uw}^W \propto \langle \theta_{u1} \xi_{w1}, \dots, \theta_{uk} \xi_{wk} \rangle \quad (32)$$

$$\psi_{ui} \propto \langle \psi_{ui}^M, \psi_{ui}^{on}, \psi_{ui}^{off} \rangle \quad (33)$$

公式(33)中元素的取值分别如公式(34)~公式(36)所示.

$$\psi_{ui}^M = \langle \theta_{u1} \beta_{i1}, \dots, \theta_{uk} \beta_{ik} \rangle \quad (34)$$

$$\psi_{ui}^{on} = \langle \delta_{u1}^{on} r_{u1}^{on} y_{1i}, \dots, \delta_{uv}^{on} r_{uv}^{on} y_{vi} \rangle \quad (35)$$

$$\psi_{ui}^{off} = \langle \tau_{u1}^{off} r_{u1}^{off} y_{1i}, \dots, \tau_{uf}^{off} r_{uf}^{off} y_{fi} \rangle \quad (36)$$

另外, ψ_{uw}^W 和 ψ_{ui} 需要分别进行归一化,使得 ψ_{uw}^W 和 ψ_{ui} 的各个分量之和分别为 1.



王绍卿(1981—),男,山东莘县人,讲师,CCF 会员,主要研究领域为推荐系统,社交网络挖掘.



赵衍衍(1991—),男,学士,主要研究领域为推荐系统,数据融合,大数据分析.



王征(1990—),男,学士,主要研究领域为推荐系统,大数据分析.



陈红(1965—),女,博士,教授,博士生导师,CCF 杰出会员,主要研究领域为数据仓库,无线传感器网络中的数据管理.



李翠平(1971—),女,博士,教授,博士生导师,CCF 杰出会员,主要研究领域为信息网络分析,数据挖掘.