

基于异构社交网络信息和内容信息的事件推荐*

尚燕敏, 曹亚男, 刘燕兵



(中国科学院 信息工程研究所, 北京 100093)

通讯作者: 曹亚男, E-mail: shangyanmin@iie.ac.cn

摘要: 基于事件的社交网络使得事件推荐受到越来越多的关注. 不同于其他推荐问题(如电影推荐等), 事件推荐具有 3 类不同信息: 用户构成的异构社交网络关系信息(在线社交网络和离线社交网络)、用户/事件的内容信息、用户对事件的隐式反馈信息. 如何有效融合这些信息进行事件推荐是该领域学者普遍关注的问题. 提出一种新的混合事件推荐方法 CHS-BPR, 该方法以贝叶斯潜在因子模型为基本框架来处理用户对事件的隐式反馈信息, 同时考虑用户/事件的内容信息和用户之间的异构社交网络信息, 首次实现了同时使用 3 种信息来做事件推荐, 并以真实数据集验证了所提方法的有效性.

关键词: 事件推荐; 异构社交网络; 内容信息; 正则项; 贝叶斯潜在因子模型

中图法分类号: TP311

中文引用格式: 尚燕敏, 曹亚男, 刘燕兵. 基于异构社交网络信息和内容信息的事件推荐. 软件学报, 2020, 31(4): 1212–1224. <http://www.jos.org.cn/1000-9825/5544.htm>

英文引用格式: Shang YM, Cao YN, Liu YB. CHS-BPR: Combining content-aware and heterogeneous-aware for event recommendation. Ruan Jian Xue Bao/Journal of Software, 2020, 31(4): 1212–1224 (in Chinese). <http://www.jos.org.cn/1000-9825/5544.htm>

CHS-BPR: Combining Content-aware and Heterogeneous-aware for Event Recommendation

SHANG Yan-Min, CAO Ya-Nan, LIU Yan-Bing

(Institute of Information Engineering, Chinese Academy of Sciences, Beijing 10093, China)

Abstract: The Web has grown into one of the most important channels to communicate social events nowadays. However, the sheer volume of events available in event-based social networks (EBSNs) often undermines the users' ability to choose the events that best fit their interests. Recommender systems appear as a natural solution for this problem. Different from classic recommendation problems (e.g. movies), event recommendation generally faces three complex problems: Heterogeneous social relationships (online and offline) among users, the implicit feedback data and the content-context information of users/events. How to effectively fuse this information for event recommendation is a common concern for scholars in this field. This work presents a Bayesian latent factor model that combines users/items content-context information and heterogeneous social information for event recommendation. Experimental results on several real-world datasets demonstrate the proposed method can efficiently tackle with implicit feedback characteristic for event recommendation.

Key words: event recommendation; heterogeneous social network; content information; regularization; Bayesian latent factor model

新出现的基于事件的在线网络服务(如 EBSNs、Plancast)为人们发布、组织事件提供了便利的平台,在这些平台中用户发布/组织的事件可以从非正式的聚会(如看电影和外出就餐等)到正式的活动(如技术会议和商务会议等). 如何精准地为用户推荐与其兴趣契合的事件对于用户本身和服务提供商来说都至关重要. 面对用户和

* 基金项目: 国家自然科学基金(61602466, 61403369); 国家重点研发计划(2016YFB0801300)

Foundation item: National Natural Science Foundation of China (61602466, 61403369); National Key R&D Program of China (2016YFB0801300)

收稿时间: 2017-09-01; 修改时间: 2017-11-08; 采用时间: 2017-12-21

事件数量的显著增长,如何从众多的事件中找到用户感兴趣的事件正是事件推荐系统关注的问题。

不同于电影、音乐等传统推荐问题,在 EBSNs 上的事件推荐有其独有的特性。EBSNs 的第 1 个独特性是社交网络的异构性。众所周知,EBSNs 上有两种类型的社交网络:在线社交网络和离线社交网络。在线社交网络由兴趣相近的用户组成的不同在线 groups 连接而成;离线社交网络由兴趣相似的用户共同参加线下事件而构成。EBSNs 的第 2 个独特性是用户对事件的隐式反馈。不同于传统的推荐问题,事件推荐不具有从 1~5 这种清晰的打分反馈信息;相反,事件推荐(以 Meetup 为例)只能获得如(yes,no)这种隐式的用户反馈信息。众所周知,对级排序算法,如贝叶斯个性化排序(BPR)及其扩展^[1-3]非常适用于处理隐式反馈信息。对级排序算法的原理是用户对那些被选中项目的兴趣度高于未被选中的项目。本文使用对级排序算法来处理用户/事件隐式反馈信息。

除了上述独特性,EBSNs 也面临着与传统推荐同样的难题——冷启动问题。在传统推荐系统中,冷启动是指某些用户或者被推荐项目先验知识不足的情况。具体到事件推荐,冷启动是指用户初次接触 EBSNs,其对事件的隐式反馈信息为 0 或者非常少;在这种情况下做事件推荐就需要用户和事件的附加信息,如用户的历史行为信息和社交网络信息、事件的描述信息等。近几年,基于混合信息的推荐成为研究热点。本文关注个性化事件推荐,研究如何利用用户的历史行为信息、事件描述信息和异构社交网络信息这些混合信息来推荐 Meetup 事件。上述混合信息的使用基于两个事实:第 1 个事实是用户兴趣可以从用户的历史行为信息中获取,事件主题可以从事件描述信息中获取。如果一个事件的主题与某用户的兴趣非常契合,则该用户可能以较高概率加入该事件;第 2 个事实是推荐系统中用户对被推荐项目的打分通常受其朋友影响,这种影响可以通过社交网络进行度量。

本文提出一种新的混合贝叶斯潜在因子模型,该模型以 BPR 为基本框架处理用户对事件的隐式反馈信息,同时结合用户/事件的内容信息(用户历史行为信息、事件描述信息)和异构社交网络信息进行事件推荐。具体来说,对用户/事件的内容信息(用户历史行为信息、事件描述信息)建立内容正则项;对异构社交网络信息建立社会正则项。然后将两种正则项加入到 BPR 框架。本文将这种混合的事件推荐模型称为基于内容信息和异构社交网络信息的贝叶斯个性化排序模型(CHS-BPR)。

本文的创新之处在于:

- (1) 提出一种新的混合事件推荐方法 CHS-BPR,该方法首次同时使用了用户的历史信息、事件描述信息和异构社交网络信息;
- (2) 使用正则项来分别建模用户/事件内容信息、异构社交网络信息,处理方法更直观;
- (3) 在真实数据集上进行了一系列实验来验证 CHS-BPR 模型的有效性。结果表明,这种同时考虑 3 种信息的多源数据融合事件推荐方法能够显著提升事件推荐的精度。

1 相关工作

本节从 4 个方面介绍有关事件推荐的相关工作:传统推荐、基于社交网络信息的事件推荐、基于内容信息的事件推荐以及处理隐式反馈信息的因子模型。

传统推荐。推荐系统旨在为用户推荐其感兴趣的项目,此目标的实现通过挖掘用户的历史打分信息完成。早期的推荐方法主要包括基于内容的方法、协同过滤方法和混合方法^[4]。近年来,随着社交网络的发展,基于社会信任的推荐提升了早期推荐方法的性能,此类推荐方法的原理是用户的兴趣会受到社交网络上他所信任的朋友的影响。早期的推荐方法和基于社会信任的推荐方法多用于电影推荐、音乐推荐等,是适用于多种个性化服务的传统推荐方法。然而,事件推荐与音乐、电影等推荐服务不同,它所涉及的信息更复杂,不仅具有用户的社会关系信息,还具有用户/事件的内容信息,因此传统的推荐方法并不适用于事件推荐。

基于社交网络信息的事件推荐。EBSNs 具有异构的社交网络信息,不仅具有在线社交网络,还具有离线社交网络。目前 EBSNs 上已有的事件推荐方法仅关注一种社交网络信息:在线信息或者离线信息。在使用在线社交网络信息方面, Lee 提出了一种基于信任关系的文化事件推荐模型^[5]; Simon 等人提出了一种基于在线用户中心的文化事件推荐方法。在使用离线社交网络信息方面,已有的相关工作几乎没有,但是同时使用在线和离线社交信息的事件推荐工作已经得到关注。Liu 等人首先提出了 EBSNs 同时具有在线社交网络和离线社交网络这种异构

社交网络信息的特性^[6].在此工作的基础上,Qiao 等人提出将这种异构的社交网络信息以社会正则项的方式用于事件推荐^[7].

基于内容信息的事件推荐.有些研究学者从用户/事件的内容信息出发进行事件推荐,如文献[1,8,9]等.总体来说,这类方法关注用户/事件的内容信息,如用户的历史行为信息、事件描述信息,并从这些信息中学习用户/事件的潜在向量进行事件推荐.目前使用用户/事件内容信息进行事件推荐的方法主要分为两类:(1) 第 1 类方法使用因子模型建模用户/事件的内容信息,如 Rendle 提出的 FM 方法^[8],该方法将所有属性信息表示为一个特征矩阵,但并未考虑冷启动问题.Map-BPR 可以解决推荐系统的冷启动问题,该方法通过进行内容信息到潜在向量空间的映射来学习新实体(用户/事件)的潜在向量.然而,Map-BPR 将学习到的潜在向量划分为两个独立的部分,这限制了在隐式反馈推荐背景下的潜在向量含义,使得潜在向量仅能表示实体(用户/事件)的协同属性,忽略了其内容属性.为了得到更可靠的潜在向量,Guo 等人通过同时结合实体协同信息和实体内容信息来学习潜在向量^[10].(2) 第 2 类方法使用主题模型来建模实体(用户/事件)的内容信息.Hsieh 等人提出一种新的用户中心推荐模型,该模型将不同类型的源信息和多种用户个人信息加入到推荐过程,不仅可以建模用户的兴趣,还可以建模事件的主题^[11].

处理隐式反馈信息的因子模型.因子模型在个性化推荐系统中应用广泛,它可以用来处理不同类型的数据,如隐式反馈信息^[3,12]、被推荐项目的属性信息^[1,8]、用户信息^[7]和社交网络信息^[7].根据处理信息类型的不同,推荐系统中使用因子模型的研究工作大致分为两类:基于协同的方法和基于内容的方法.无论哪种方法,贝叶斯个性化排序方法(BPR)及其扩展方法在因子模型中的地位都至关重要.基于内容的 BPR 使用用户/项目的内容信息,基于协同的 BPR 使用隐式协同信息.BPR 方法的基本理论是基于成对排序算法,它适用于大规模训练集,通过 SGD 完成求解.

2 CHS-BPR 模型

2.1 概念定义

在介绍本文工作之前,首先明确文中出现的几个术语概念,详见表 1.

Table 1 Terminology concept

表 1 术语概念

术语	概念
用户历史信息	全称为用户历史行为信息,是指用户在事件推荐网络服务上发表搜索词、点击事件、留言等一系列用户行为
用户内容信息	本文从内容语义的角度挖掘用户历史信息,因此又将用户历史信息称为用户内容信息
用户内容特征	从用户内容信息中提取的用户特征,称为用户内容特征
事件描述信息	指事件关键词、事件描述语句、事件发起者等与事件有关的信息
事件内容信息	本文从内容语义的角度挖掘事件描述信息,因此又将事件描述信息称为事件内容信息
事件内容特征	从事件内容信息中抽取的事件特征,称为事件内容特征

2.2 问题定义

在事件推荐中,有 4 种类型的实体,分别是用户 S (user)、用户历史信息 P^S (user profile)、事件 E (event)和事件描述信息 P^E (event profile);两种社交网络,分别是在线社交网络 G^{On} (online social network)和离线社交网络 G^{Off} (offline social network).用户集合表示为 $S=\{u_1, u_2, \dots, u_m\}$,对于其中每一个用户 $u_i, i=1, 2, \dots, m$,我们使用 Con_i^u 来表示用户 u_i 的内容信息.事件集合表示为 $E=\{v_1, v_2, \dots, v_n\}$,对于其中每一个事件 $v_j, j=1, 2, \dots, n$,我们使用 Con_j^v 来表示事件 v_j 的内容信息.在线社交网络 G^{On} 表示用户之间的在线关系,离线社交网络 G^{Off} 表示用户之间的离线关系.事件推荐问题定义为排序问题,即对每个用户 u ,根据用户 u 对每个事件 v 的打分 $r(u, v)$ 高低将事件排序,其中, $r(u, v)$ 代表了用户 u 对事件 v 的偏好.因此,准确预测 $r(u, v)$ 是事件推荐的核心问题.

在事件推荐中,利用传统的矩阵分解方法,可以从用户-事件矩阵得到用户对事件的偏好,我们称此偏好为一般偏好.除了一般偏好以外,从用户和事件的内容信息还可以挖掘出用户对事件的兴趣偏好.本文提出一种融合用户一般偏好和兴趣偏好的事件推荐模型.具体来讲,该模型通过贝叶斯基本框架来处理用户对事件的隐式

反馈信息(一般偏好),同时将用户历史信息 and 事件描述信息(兴趣偏好)以内容正则项的方式加入贝叶斯基本框架,而对于异构社交网络信息则以社会正则项的方式加入到贝叶斯框架,以提升事件推荐的精度.基于上述想法,本文所提方法是基于贝叶斯基本框架融入了内容信息和异构社交网络信息的事件方法,称为 CHS-BPR.

2.3 异构社交网络

(1) 在线社交网络

在基于事件的社交网络服务 EBSNs 中,组群(group)反映了其成员用户的在线社交网络关系(如图 1 所示).以 Meetup 为例,在同一族群中和 u_i 之间的亲密程度可以用边的权重信息 ω_{ij}^{om} 来表示,如式(1):

$$\omega_{ij}^{om} = \frac{|G(u_i) \cap G(u_j)|}{|G(u_i) \cup G(u_j)|} \quad (1)$$

这里, $G(u_i)$ 表示用户 u_i 加入的所有群组, $|G(u_i) \cap G(u_j)|$ 表示用户 u_i 和 u_j 共同加入的群组数量, $|G(u_i) \cup G(u_j)|$ 表示用户 u_i 加入的群组数量与用户 u_j 加入的群组数量之和.

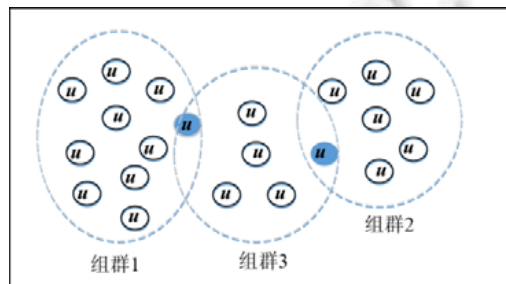


Fig.1 Online social network

图 1 在线社交网络

(2) 离线社交网络

离线社交网络由用户在线下共同参加事件活动而产生(如图 2 所示).如果用户 u_i 和 u_j 线下共同参加一个社会事件,则他们之间存在离线链接.同理, ω_{ij}^{off} 表示用户 u_i 和 u_j 离线社会关系的强弱,表示为

$$\omega_{ij}^{off} = \frac{|E(u_i) \cap E(u_j)|}{|E(u_i) \cup E(u_j)|} \quad (2)$$

这里, $E(u_i)$ 表示用户 u_i 参加的事件集合, $|E(u_i) \cap E(u_j)|$ 表示用户 u_i 和 u_j 共同参加的事件数量, $|E(u_i) \cup E(u_j)|$ 表示用户 u_i 参加的事件数量与用户 u_j 参加的事件数量之和.

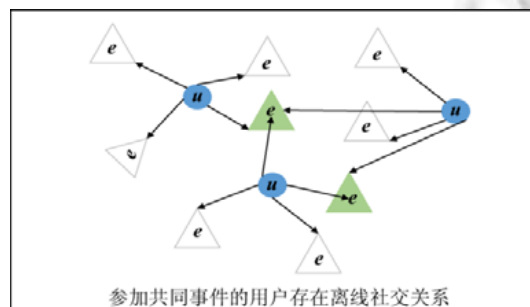


Fig.2 Offline social network

图 2 离线社交网络

2.4 一般偏好

用户-事件矩阵可以反映用户对事件的一般偏好,这个一般偏好可以通过矩阵分解潜在因子模型获得,基本

思想是将用户 u_i 和事件 v_j 分别以低维向量 $U_i \in R^k, V_j \in R^k$ 表示 (U_i 表示一般偏好中的用户潜在向量, V_j 表示一般偏好中的事件潜在向量), 那么用户 u_i 对事件 v_j 的一般偏好打分 $r'(u_i, v_j)$ 表示为

$$r'(u_i, v_j) = U_i^T V_j \quad (3)$$

2.5 兴趣偏好

(1) 用户/事件的内容信息特征

用户的历史信息和事件的描述性信息中都蕴含着语义信息, 也称为内容信息. 本文使用经典的语义处理方法——词袋子模型将用户与事件的内容信息提取出来. 对于每一个用户, 从该用户过去参加过的事件描述词语中抽取用户的 TF-IDF 向量表示, 其中, 该用户参加过的每一个事件都以距离当前推荐时刻的时间长短来表示其重要程度. 基于此, 用户 $u_i \in U$ 的内容信息特征 A_i^U 表示为

$$A_i^U = \sum_{v \in E(u_i)} \frac{1}{1 + \alpha^{T(v)}} \bar{e} \quad (4)$$

这里, \bar{e} 表示从事件 v 的描述词中抽取的 TF-IDF 向量, α 表示时间衰减因子, $T(v)$ 代表事件 v 从被推荐给用户 u_i 时刻起到用户 u_i 真正参加事件 v 的时间差.

对于每一个事件 $v_j \in V$, 它的内容信息特征 A_j^V 表示为

$$A_j^V = \bar{e}_j \quad (5)$$

这里, \bar{e}_j 表示事件 v_j 的 TF-IDF 向量.

(2) 兴趣偏好驱动的打分

给定用户 u_i 的内容信息特征 A_i^U 和事件 v_j 的内容信息特征 A_j^V , 可以确定用户 u_i 对事件 v_j 的兴趣偏好打分, 记为 $r''(u_i, v_j)$, 其公式表示如下:

$$r''(u_i, v_j) \sim N((A_i^U)^T A_j^V, c_{ij}^{-1}) \quad (6)$$

这里, A_i^U 表示用户 u_i 的内容信息特征(用户兴趣)向量, A_j^V 表示事件 v_j 的内容信息特征(事件主题)向量, $(A_i^U)^T A_j^V$ 表示用户兴趣与事件主题的契合程度. 参数 c_{ij} 表示对这种兴趣偏好打分的信任程度, 如果我们越信任 $r''(u_i, v_j)$, 则 c_{ij} 参数设置得越大, 反之 c_{ij} 设置得越小. 信任参数非常适用于事件推荐这种隐式打分情况. 例如, 在事件推荐服务 Meetup 中, 用户参加某个事件表示他对该事件感兴趣, 那么用户对事件的兴趣打分偏好设为 $r''(u_i, v_j) = 1$, 且设置较大的 c_{ij} . 用户没有参加某个事件则可能有两种原因: 其一是用户对此事件主题不感兴趣; 其二是用户对此事件主题感兴趣, 但并没有注意到这个事件. 本文忽略上述第 2 个原因, 对用户没有参加的事件, 只假定为用户对事件不感兴趣, 那么用户对事件的兴趣偏好打分为 $r''(u_i, v_j) = 0$, 且设置较小的 c_{ij} . 有关信任参数 c_{ij} 的具体设定如下:

$$c_{ij} = \begin{cases} a, & \text{if } r''(u_i, v_j) = 1 \\ b, & \text{if } r''(u_i, v_j) = 0 \end{cases} \quad (7)$$

这里, a 和 b 是调整参数, 且 $a > b > 0$, 本文中, $a = 0.9, b = 0.5$.

2.6 混合偏好

一般偏好和兴趣偏好决定了用户 u_i 对事件 v_j 的混合偏好, 混合偏好驱动的打分表示为

$$r(u_i, v_j) = r'(u_i, v_j) + \beta r''(u_i, v_j) \quad (8)$$

这里, β 作为系数用来调整一般偏好和兴趣偏好所占比重.

2.7 贝叶斯个性化排序

用户对事件呈现隐式打分, 即用 0/1 表示用户是否参加某事件. 贝叶斯个性化排序(BPR^[12])则非常适用于这种隐式打分情况, 它赋予用户参加过的事件比用户未参加过的事件更高的分数, 并按得分将用户从高到低排序. 基于 BPR 的基本思想, 我们将用户 u_i 参加过的事件集合记为正样本集, 用 $P_{u_i}^+$ 表示; 用户 u_i 未参加过的事件集合

记为负样本集,用 $N_{u_i}^l$ 表示.我们期望使得用户 u_i 的正样本集 $P_{u_i}^l$ 排序高于负样本集 $N_{u_i}^l$ 排序,公式表示为

$$\max_{\Theta} \prod_{(u_i, v_j, v_k) \in (U, P_{u_i}^l, N_{u_i}^l)} P(r(u_i, v_j) > r(u_i, v_k) | \Theta) \quad (9)$$

这里,参数 $\Theta = (U, V), P(r(u_i, v_j) > r(u_i, v_k) | \Theta) = \ell(r(u_i, v_j) - r(u_i, v_k))$, 函数 $\ell(\cdot)$ 是逻辑斯特函数.

至此,融合用户一般偏好和兴趣偏好的事件推荐用户打分模型构造完成,表示为

$$\min_{\Theta} L_{feedback} = \min_{\Theta} \left[-\sum_{(u_i, v_j, v_k) \in (U, P_{u_i}^l, N_{u_i}^l)} \ln(r(u_i, v_j) - r(u_i, v_k)) \right] \quad (10)$$

这里, $L_{feedback} = -\sum_{(u_i, v_j, v_k) \in (U, P_{u_i}^l, N_{u_i}^l)} \ln(r(u_i, v_j) - r(u_i, v_k))$.

2.8 异构的社会正则

在 EBSNs 的在线社交网络中,用户兴趣受用户在线链接关系的影响.本文采用社会正则项的方式处理这种影响,依据的基本思想是用户兴趣接近与其有链接关系的朋友兴趣的平均值,且认为这种兴趣接近程度服从高斯先验,这种在线社会正则项用公式表示为

$$P\left(U_i - \sum_{j=1}^n \frac{\omega_{ij}}{\sum_{k=1}^n \omega_{ik}} U_j \mid W^{on}\right) \propto N\left(U_i - \sum_{j=1}^n \frac{\omega_{ij}}{\sum_{k=1}^n \omega_{ik}} U_j \mid 0, \sigma^2 I\right) \quad (11)$$

另外,事件推荐通过组群来发布事件消息,一个组群可以发布多个事件,一个事件也可以同时发布在多个组群,因此组群和事件之间存在隐式关系,这种隐式关系可以通过如下方式定义:

$$f_{ij}^H = \frac{|G(u_i) \cap G(u_j) \oplus E(u_i) \cap E(u_j)|}{|G(u_i) \cap G(u_j) \otimes E(u_i) \cap E(u_j)|} \quad (12)$$

这里, $|G(u_i) \cap G(u_j) \oplus E(u_i) \cap E(u_j)|$ 表示用户 u_i 和用户 u_j 同时参加的组群中共同参加的事件数目; $|G(u_i) \cap G(u_j) \otimes E(u_i) \cap E(u_j)|$ 表示用户 u_i 和用户 u_j 同时参加的组群数目与用户 u_i 和用户 u_j 同时参加的事件数目之和.

用户共同参加的事件可以构造离线社交网络,结合离线社交网络和组群事件之间的隐式关系信息,将公式(11)中的在线社会正则项进行扩展,得到异构社会正则项:

$$L_{social} = g(U_i \mid W^{on}, W^{off}, F^H) \propto N\left(U_i - \sum_{j=1}^n \frac{f_{ij}^H (\omega_{ij}^{on} + \omega_{ij}^{off})}{\sum_{k=1}^n [f_{ik}^H (\omega_{ik}^{on} + \omega_{ik}^{off})]} U_j \mid 0, \sigma^2 I\right) \quad (13)$$

这里,令 $\gamma_{ij} = \frac{f_{ij}^H (\omega_{ij}^{on} + \omega_{ij}^{off})}{\sum_{k=1}^n [f_{ik}^H (\omega_{ik}^{on} + \omega_{ik}^{off})]}$, 因此有 $g(U_i \mid W^{on}, W^{off}, F^H) \propto N\left(U_i - \sum_{j=1}^n \gamma_{ij} U_j \mid 0, \sigma^2 I\right)$.

2.9 内容信息正则

对于用户历史信息 and 事件描述信息,我们以内容正则项的方式来处理. $A^U = \{A_1^U, A_2^U, \dots, A_n^U, \dots, A_m^U\}$ 表示所有用户的内容信息特征集合, $A^V = \{A_1^V, A_2^V, \dots, A_j^V, \dots, A_n^V\}$ 表示所有事件的内容信息特征集合.我们认为用户 u_i 的内容信息特征 A_i^U 与其一般偏好潜在向量 U_i 某种程度上接近,表示如下:

$$L_{content}^{U_i} = \|A_i^U W^U - U_i\|_F^2 \quad (14)$$

对于所有用户而言,则有:

$$L_{content}^U = \|A^U W^U - U\|_F^2 \quad (15)$$

这里, $W^U \in R^{d^U \times k}$ 是参数, k 为潜在向量 U_i 的维度.

相似地,对于事件 v_j ,它的信息内容特征 A_j^V 与其一般偏好潜在向量 V_j 接近,表示为

$$L_{content}^V = \|A^V W^V - V\|_F^2 \quad (16)$$

对于所有事件,则有:

$$L_{content}^U = \|A^U W^U - U\|_F^2 \quad (17)$$

同样, $W^V \in R^{d^V \times k}$ 是参数, k 为潜在向量 V_j 的维度.

2.10 构造CHS-BPR模型

至此,给定用户事件隐式反馈矩阵 R 、由 R 得到的用户潜在向量矩阵 U 、事件潜在向量矩阵 V 、在线社交关系 W^{on} 、离线社交关系 W^{off} 、组群/事件关系 F^H 、用户内容特征 A^U 、事件内容特征 A^V ,那么融合用户/事件内容信息和异构社交网络信息的事件推荐模型 CHS-BPR 构建完成,表示为

$$\begin{aligned} \arg \min_{\Omega} [L_{feedback} + L_{social} + L_{content}] = \arg \min_{\Omega} & \left[- \sum_{(u_i, v_j, v_k) \in (U, P_{u_i}^I, N_{u_i}^I)} \ln(r(u_i, v_j) - r(u_i, v_k)) + \right. \\ & \sum_i \ln g(U_i | W^{on}, W^{off}, F^H, \sigma) + \\ & \sum_i P(U_i | \sigma_u) + \sum_j P(V_j | \sigma_v) + \|A^U W^U - U\|_F^2 + \\ & \frac{1}{2} \sum \lambda^U \|W^U\|_F^2 + \|A^V W^V - V\|_F^2 + \\ & \left. \frac{1}{2} \sum \lambda^V \|W^V\|_F^2 \right] \quad (18) \end{aligned}$$

上式中,参数 $\Omega = \{U, V, W^U, W^V\}$ 是模型需要学习的.

3 参数学习

对于 CHS-BPR 模型的求解,我们采用 SGD 方法学习参数 $\Omega = \{U, V, W^U, W^V\}$,求解步骤如下.

学习参数 U 和 V .

$$\begin{aligned} \frac{\partial F}{\partial U_i} = & - \sum_{V_j \in P_{u_i}^I} \sum_{V_k \in N_{u_i}^I} \beta \frac{e^{r(u_i, v_k) - r(u_i, v_j)}}{1 + e^{r(u_i, v_k) - r(u_i, v_j)}} I(i, j, k) (v_k - v_j) - \frac{1}{\sigma_u^2} \left(2U_i - \sum_{j=1}^{|U|} \Upsilon_{ij} U_j \right) + \\ & \frac{\Upsilon_{ii}}{\sigma_u^2} \sum_{t=1, t \neq i}^{|U|} \left(U_t - \sum_{j=1}^n \Upsilon_{ij} U_j \right) + U_i - A^U W^U \quad (19) \end{aligned}$$

根据 SGD 的迭代规则:

$$U_i = U_i + \gamma \frac{\partial F}{\partial U_i} \quad (20)$$

这里, γ 是预先定义的步长.

$$\frac{\partial F}{\partial U_i} = \sum_{i=1}^{|U|} \sum_{k=1}^{|V|} \beta \left\{ \left[\frac{e^{r(u_i, v_k) - r(u_i, v_j)}}{1 + e^{r(u_i, v_k) - r(u_i, v_j)}} U_i \right] I(i, j, k) - \left[\frac{e^{r(u_i, v_k) - r(u_i, v_j)}}{1 + e^{r(u_i, v_k) - r(u_i, v_j)}} U_i \right] I(i, k, j) \right\} + V_j - A^V W^V \quad (21)$$

$$V_j = V_j + \gamma \frac{\partial F}{\partial V_j} \quad (22)$$

学习参数 W^U 和 W^V .

$$\frac{\partial F}{\partial W_i^U} = (A_i^U)^T (A_i^U W_i^U - U_i) + \lambda_i^U W_i^U \quad (23)$$

令 $\frac{\partial F}{\partial W_i^U} = 0$,按下式更新 W_i^U :

$$W_i^U = ((A_i^U)^T A_i^U + \lambda_i^U O) A_i^U U_i \quad (24)$$

这里, O 为单位矩阵.

$$\frac{\partial F}{\partial W_j^V} = (A_j^V)^T (A_j^V W_j^V - V_j) + \lambda_j^V W_j^V \quad (25)$$

令 $\frac{\partial F}{\partial W_j^V} = 0$, 按下式更新 W_j^V :

$$W_j^V = ((A_j^V)^T A_j^V + \lambda_j^V O) A_j^V V_j \quad (26)$$

算法 1 给出了参数学习的细节.

算法 1. CHS-BPR 模型求解.

输入: 用户事件隐式反馈矩阵 R , 在线社交关系 W^{on} , 离线社交关系 W^{off} , 组群/事件关系 F^H , 用户内容特征 A^U , 事件内容特征 A^V , 预先定义的收敛矩阵 Λ , 先验参数 $\sigma_u, \sigma_v, \sigma, \alpha, \beta$,

输出: 用户潜在向量矩阵 U , 事件潜在向量矩阵 V , 参数矩阵 W^U 和 W^V .

1. 随机初始化 U, V, W^U 和 W^V
2. for 每个用户潜在向量 $U_i \in U$, do
3. $U_i = U_i + \gamma \frac{\partial F}{\partial U_i}$
4. end for
5. if $U(new) - U(old) < \Lambda$, then U 收敛
6. for 每个事件潜在向量 $V_j \in V$, do
7. $V_j = V_j + \gamma \frac{\partial F}{\partial V_j}$
8. end for
9. if $V(new) - V(old) < \Lambda$, then V 收敛
10. for 每个用户内容的参数矩阵 $W_i^U \in W^U$, do
11. 按照公式(24)更新 W_i^U
12. end for
13. If $W^U(new) - W^U(old) < \Lambda$, then W^U 收敛
14. for m 每个事件内容的参数矩阵 $W_j^V \in W^V$, do
15. 按照公式(26)更新 W_j^V
16. end for
17. if $W^V(new) - W^V(old) < \Lambda$, then W^V 收敛

4 实验及评价

4.1 数据集

本文使用的数据集来源于 Macedo 等人的工作^[13], 选取 Phoenix、Chicago 和 San Jose 这 3 座城市的 Meetup 事件数据, 数据集的详细信息见表 2.

Table 2 Statistics of the data set

表 2 数据集统计信息

城市	组群数	用户数	事件数	RSVPs
Chicago	2 321	207 649	190 927	1 375 154
Phoenix	1 661	117 458	222 632	1 209 324
San Jose	2 589	242 143	206 682	1 607 985

4.2 评价方法

本文使用 3 种评价方法来评估事件推荐的结果: AUC、 $P@k$ 和 MAP.AUC 用来度量分类的整体结果, 它非

常适用于正负样本数量不平衡的情况.在事件推荐中,推荐给用户的事件为正样本,未推荐给用户的事件为负样本,负样本占有非常高的比例.AUC 的值越高,则代表事件推荐的结果越好.本文在成对测试集上使用 AUC 度量 CHS-BPR 算法的性能.

$$AUC = \frac{\sum_{i=1}^{|U|} \sum_{v_j \in P'_i} \sum_{v_k \in N'_i} \delta(\gamma(u_i, v_j) > \gamma(u_i, v_k))}{\sum_{i=1}^{|U|} |P'_i| |N'_i|} \quad (27)$$

这里, $\delta(\cdot)$ 是一个指示函数,如果用户 u_i 对推荐事件 v_j 的打分高于对未被推荐事件 v_k 的打分(即 $\gamma(u_i, v_j) > \gamma(u_i, v_k)$),则 $\delta(\cdot)=1$, 否则, $\delta(\cdot)=0$.

$P@k$ 和 MAP 在排序问题中使用较多.本文在列表测试集上使用 $P@k$ 和 MAP 评估 CHS-BPR 算法性能.对每个用户 u , 平均准确率 AP 表示为公式(28).

$$AP_u = \frac{\sum_{k=1}^m P@k \cdot \delta(L^k(u) \in P'_u)}{|P'_u|} \quad (28)$$

在上述公式中, $P@k$ 表示在推荐列表中前 k 个事件的精度, m 表示事件数量, $\delta(\cdot)$ 与公式(27)的含义相同, $L^k(u)$ 表示 CHS-BPR 模型推荐给用户 u 的事件列表中的第 k 个事件, P'_u 表示用户 u 参与的事件数目.最后,我们通过对所有用户 AP_u 的均值来获得 MAP.

4.3 对比模型

本文使用 4 种经典推荐方法来验证 CHS-BPR 模型的性能.

- MF 是一种经典的因子模型^[12], 经常用于处理协同信息.
 - BPR-MF 是一种在 BPR 学习框架下的矩阵分解模型^[3].
 - CA-BPR 是一种结合了隐式反馈信息和内容信息的 BPR 扩展模型, 考虑了用户和项目内容的相似性^[13].
 - Hesig 同时结合 3 种信息进行事件推荐: 用户位置信息、异构社交网络信息和用户/事件隐式反馈信息^[7].
- 上述 4 种模型使用的用户/事件内容信息、社交网络信息、用户/事件反馈信息各有侧重, 详见表 3.

Table 3 Characteristic of compared methods

表 3 对比模型使用的信息类型

对比模型/使用信息	内容信息	社交网络信息	隐式反馈信息
MF	No	No	Yes
BPR-MF	No	No	Yes
CA-BPR	Yes	No	Yes
Hesig	No	Yes	Yes
CHS-BPR	Yes	Yes	Yes

4.4 参数设置

1) 先验参数和学习率

(1) 模型中的先验参数 σ 、 σ_u 和 σ_v 设置为经验值 0.001.

(2) 学习率控制 CHS-BPR 模型训练的速度, 学习率太大或太小都会影响模型收敛. 本文将 CHS-BPR 模型的多个学习率 (γ , λ_i^u 和 λ_j^v) 均设为经验值 0.001.

2) 参数 α 和 β

(1) 公式(4)中的 α 为时间衰减系数, 本文将其设为 0.005.

(2) 公式(8)中的相对权重 β 控制着用户对事件的混合偏好打分中一般偏好打分和兴趣偏好打分各自所占比例, 本文根据 AUC 的变化来选择合适的 β 值. 如图 3 所示, 当 $\beta \in [0.64, 0.72]$ 时, CHS-BPR 模型的 AUC 值较大, 这里我们选择 0.68 作为最终的 β 值.

3) 潜在向量的维度

除了上述参数,我们还考虑了潜在向量维度对模型性能的影响;实验结果表明,潜在向量维度对模型性能的影响并不大.

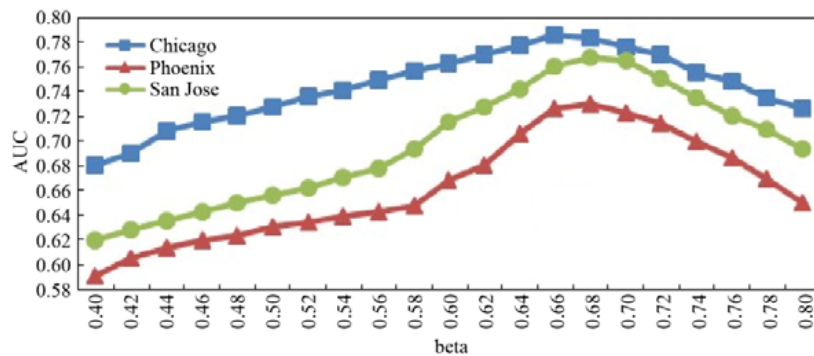


Fig.3 Test on the parameter β with pair data sets
图3 参数 β 对 AUC 的影响

4.5 结果分析

本节介绍 CHS-BPR 模型在 Meetup 数据集上的实验结果,并通过回答两个问题对实验结果进行分析.

1) 问题 1:CHS-BPR 模型在成对测试集上的性能如何?

表 4 展示了在成对测试集上 CHS-BPR 模型与 4 个对比模型的 AUC 值.从该结果对比中我们发现:

- (1) 所有模型都适用于处理非平衡隐式反馈数据;
- (2) 异构社交网络和内容信息确实可以提升事件推荐的性能;
- (3) CA-BPR 模型使用用户的内容信息来建模用户兴趣偏好,其性能高于 MF 和 BPR-MF 但低于 CHS-BPR,原因在于 CA-BPR 忽略了异构社交网络信息;
- (4) Hesig 使用用户异构社交网络信息建模朋友对用户的影响,但其性能仍然低于 CHS-BPR,原因在于 Hesig 忽略了用户/事件的内容信息;
- (5) CHS-BPR 同时结合用户/事件内容信息和异构社交网络信息进行事件推荐,在一定程度上弥补了 CA-BPR 和 Hesig 的不足,提高了事件推荐精度.

Table 4 AUC of comparison methods

表 4 CHS-BPR 与对比模型的 AUC 值

城市/方法	MF	BPR-MF	CA-BPR	Hesig	CHS-BPR
Chicago	0.533	0.692	0.701	0.712	0.720
Phoenix	0.510	0.635	0.637	0.634	0.640
San Jose	0.532	0.618	0.633	0.637	0.645

2) 问题 2:CHS-BPR 模型在列表测试集上的性能如何?

对于推荐给用户的事件列表,用户往往关注排在前面位置的事件,所以本文考察 top1、top2、top3 位置的事件推荐精度,使用 MAP 来度量整体精度.

(1) 整体分析:图 4~图 6 分别展示了在 Chicago、Phoenix 和 San Jose 这 3 个城市 CHS-BPR 模型与 4 个对比模型的 $P@1$ 、 $P@2$ 、 $P@3$ 和 MAP.实验结果表明,CHS-BPR 的事件推荐精度要高于 4 个对比模型.

(2) $P@k$ 性能对比:当推荐列表的 top k 的 k 值越大时,如果 CHS-BPR 模型推荐准确率高于其他对比模型,则说明 CHS-BPR 模型性能越好.综合 3 个城市的实验结果发现,CHS-BPR 模型在 $P@3$ 时的准确率与在 $P@2$ 和 $P@1$ 时差距最大,这正验证了 CHS-BPR 模型的性能.

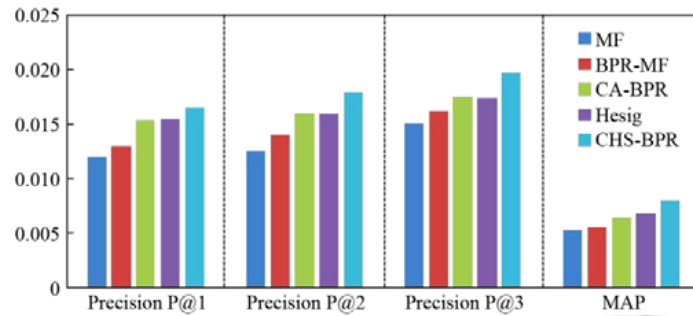


Fig.4 Comparisons on the Chicago's list test sets

图4 Chicago 列表测试集上的模型性能对比

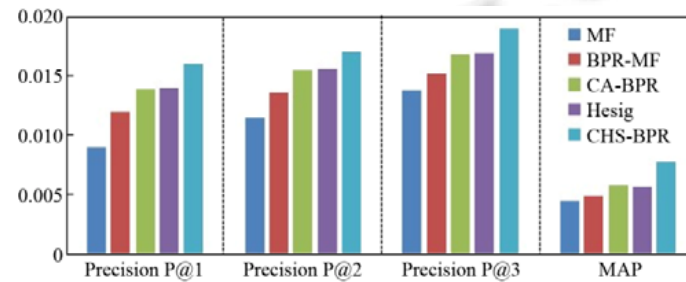


Fig.5 Comparisons on the Phoenix's list test sets

图5 Phoenix 列表测试集上的模型性能对比

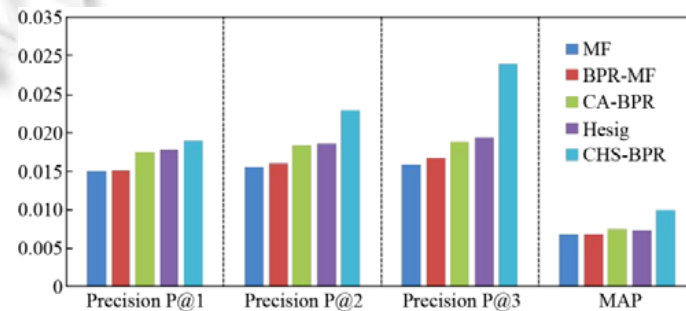


Fig.6 Comparisons on the San Jose's list test sets

图6 San Jose 列表测试集上的模型性能对比

(3) 不同城市性能对比:对比 Chicago、Phoenix 和 San Jose 这 3 个城市的模型性能发现:

- 4 个对比模型和 CHS-BPR 模型在 San Jose 上的准确率均略高于另外两个城市,产生这种现象的原因在于 San Jose 数据集包含的用户数、事件数和 RSVP 数量均多于另两个数据集;
- 在 San Jose 数据集上,对比 $P@1$ 和 $P@2$,当 $P@3$ 时 CHS-BPR 模型性能提升明显,这也得益于 San Jose 丰富的用户信息和事件信息,证明了内容信息和社交网络信息的确有助于提升事件推荐的精度.

3) 问题 3:内容信息、社交网络信息对 CHS-BPR 模型效果的提升度如何?

- 内容信息对 CHS-BPR 模型效果的提升度:由于 Hesig 使用了社交网络信息,CHS-BPR 使用了社交网络信息和内容信息,对比 Hesig 与 CHS-BPR 可知内容信息对 CHS-BPR 模型性能确有提升,AUC 平均提升了 0.007,详见表 5;

• 社交网络信息对 CHS-BPR 模型效果的提升度:由于 CA-BPR 使用了内容信息,CHS-BPR 使用了内容信息和社交网络信息,对比 CA-BPR 与 CHS-BPR 可知社交网络信息对 CHS-BPR 模型性能确有提升,AUC 提升了 0.011,详见表 5.

Table 5 The enhancement of the effect of different information on the CHS-BPR model

表 5 不同信息对 CHS-BPR 模型效果的提升度

城市/信息种类	内容信息(CHS-BPR 对比 Hesig)	社交网络信息(CHS-BPR 对比 CA-BPR)
Chicago	0.008	0.019
Phoenix	0.006	0.003
San Jose	0.008	0.012
平均值	0.007	0.011

5 总结与未来工作

本文提出了一种混合的事件推荐模型 CHS-BPR,该方法同时考虑了异构社交网络信息和用户/事件内容信息,并在真实数据集上验证了所提模型的有效性.通过与 4 种经典推荐方法进行对比,发现本文所提 CHS-BPR 模型在一定程度上弥补了已有事件推荐模型的不足,进一步提升了事件推荐精度.

近年来,深度学习为传统的推荐系统带来了新的活力.目前深度学习在推荐算法方面的研究主要包括两个方面:一是基于推荐系统的输入数据,分为使用用户内容信息的方法^[14,15]和不使用用户内容信息的方法^[16];二是基于推荐系统的输出数据,分为预测 Item 排序的方法^[17,18]和预测用户对 Item 打分的方法^[19,20].在这些已有的基于深度学习的推荐算法研究工作中,循环神经网络和前向反馈神经网络应用较多,卷积神经网络很少用在推荐算法中,未来工作将开展卷积神经网络在推荐算法方面的研究.

References:

- [1] Cohen D, Aharon M, Koren Y, Somekh O, Nissim R. Expediting exploration by attribute-to-feature mapping for cold-start recommendations. In: Proc. of the 11th ACM Conf. on Recommender Systems. New York: ACM, 2017. 184–192.
- [2] Pan WK, Chen L. GBPR: Group preference based bayesian personalized ranking for one class collaborative filtering. In: Proc. of the 23rd Int'l Joint Conf. on Artificial Intelligence. Beijing: AAAI Press, 2013. 2691–2697.
- [3] Rendle S, Freudenthaler C, Gantner Z, Schmidt-Thieme L. BPR: Bayesian personalized ranking from implicit feedback. In: Proc. of the Conf. on Uncertainty in Artificial Intelligence. Montreal: Association for Uncertainty in Artificial Intelligence, 2009. 452–461.
- [4] Bao J, Zheng Y, Wilkie D, Mokbel M. Recommendations in location-based social networks: A survey. *GeoInformatica*, 2015, 19(3): 525–565.
- [5] Lee DH. Pittcult: Trust-based cultural event recommender. In: Proc. of the 2008 ACM Conf. on Recommender systems. New York: ACM, 2008. 311–314.
- [6] Ma H, Zhou DY, Liu C, *et al.* Recommender systems with social regularization. In: Proc. of the 11th Int'l Conf. on Data Mining. New York: ACM, 2011. 287–296.
- [7] Qiao Z, Zhang P, Cao YN, Zhou C, Guo L, Fang BX. Combining heterogenous social and geographical information for event recommendation. In: Proc. of the 28th AAAI Conf. on Artificial Intelligence. Québec: AAAI Press, 2014. 145–151.
- [8] Rendle S. Factorization machines with LIBFM. *ACM Trans. on Intelligent Systems and Technology*, 2012, 3(3): 57:1–57:22.
- [9] Hong LJ, Doumith AS, Davison BD. Co-factorization machines: Modeling user interests and predicting individual decisions in Twitter. In: Proc. of the 6th ACM Int'l Conf. on Web Search and Data Mining. New York: ACM, 2013. 57–566.
- [10] Guo W, Wu S, Wang L, Tan T. Adaptive pairwise learning for personalized ranking with content and implicit feedback. In: Proc. of the IEEE/WIC/ACM Web Intelligence. Singapore: IEEE, 2015. 369–376.
- [11] Hsieh CK, Yang LQ, Wei HH. Immersive recommendation: News and event recommendations using personal digital traces. In: Proc. of the 25th Int'l World Wide Web Conf. 2016. 51–62.
- [12] Hu YF, Koren Y, Volinsky C. Collaborative filtering for implicit feedback datasets. In: Proc. of the 8th Int'l Conf. on Data Mining. New York: ACM, 2008. 263–272.

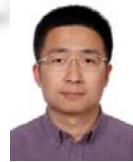
- [13] Macedo AQ, Marinho LB, Santos RLT. Context-aware event recommendation in event-based social networks. In: Proc. of the 9th ACM Conf. on Recommender Systems. New York: ACM, 2015. 123–130.
- [14] van den Oord A, Dieleman S, Schrauwen B. Deep content-based music recommendation. In: Proc. of the 26th Int'l Conf. on Neural Information Processing Systems. Curran Associates Inc., 2013,2:2643–2651.
- [15] Hamel P, Lemieux S, Bengio Y, Eck D. Temporal pooling and multiscale learning for automatic annotation and ranking of music audio. In: Proc. of the ISMIR. 2011. 729–734.
- [16] Zuo Y, Zeng J, Gong M, Jiao L. Tag-aware recommender systems based on deep neural networks. Neurocomputing, 2016,204: 51–60.
- [17] Hidasi B, Karatzoglou A, Baltrunas L, Tikk D. Session-based recommendations with recurrent neural networks. In: Proc. of the ICLR. Computer Science, 2016.
- [18] Caihua W, Wang J, Liu J, Liu W. Recurrent neural network based recommendation for time heterogeneous feedback. Knowledge-Based Systems, 2016,109:90–103.
- [19] Larochelle H, Murray I. The neural autoregressive distribution estimator. Journal of Machine Learning Research, 2011,15:29–37.
- [20] Salakhutdinov R, Mnih A, Hinton G. Restricted Boltzmann machines for collaborative filtering. In: Proc. of the 24th Int'l Conf. on Machine Learning. New York: ACM, 2007. 791–798.



尚燕敏(1982—),女,河北保定人,博士,助理研究员,主要研究领域为用户画像.



曹亚男(1985—),女,博士,副研究员,CCF专业会员,主要研究领域为自然语言处理,用户建模.



刘燕兵(1981—),男,博士,副研究员,CCF高级会员,主要研究领域为模式串匹配,信息过滤,图数据分析挖掘.