

一种基于协同矩阵分解的用户冷启动推荐算法

高玉凯¹ 王新华¹ 郭磊² 陈竹敏³

¹(山东师范大学信息科学与工程学院 济南 250358)

²(山东师范大学管理科学与工程学院 济南 250358)

³(山东大学计算机科学与技术学院 济南 250101)

(ykgao.cs@gmail.com)

Learning to Recommend with Collaborative Matrix Factorization for New Users

Gao Yukai¹, Wang Xinhua¹, Guo Lei², and Chen Zhumin³

¹(School of Information Science & Engineering, Shandong Normal University, Jinan 250358)

²(School of Management Science & Engineering, Shandong Normal University, Jinan 250358)

³(School of Computer Science and Technology, Shandong University, Jinan 250101)

Abstract Location-based service (LBS) as an information sharing platform can help people obtain more useful information. But with the increasing number of users, LBS is faced with a serious problem of information overload. Using the recommender system to filter information and help users to find valuable information has become a hot research topic in recent years. In LBS, only positive implicit feedback is available and user cold-start problem in this scenario is not well studied. Based on the observations, we consider the characteristics of location-based services platform and propose a recommender algorithm, which combines collaborative PMF (probabilistic matrix factorization) with GBDT (gradient boosting decision tree), to solve the cold start problem. The algorithm first use multi probabilistic matrix factorization to learn user latent feature in different dimension, and then use gradient boosting decision tree to train the factor and label to learn the user's preference, finally use the improved top- N recommender which considers the budget problem to produce the recommendation list. The experimental results on the real data show that the proposed algorithm can achieve better results in accuracy and $F1$ than other popular methods, and can solve the cold-start problem in LBS recommendation.

Key words recommender system; location-based service; probabilistic matrix factorization; cold-start problem; budget

摘要 位置服务作为一种信息共享平台,在方便人们交流和共享信息的同时,也因为用户数量的不断增加,而面临着严重的信息过载问题.如何利用推荐技术对信息进行过滤和筛选,帮助用户在位置服务中发现有价值的信息成为近年来研究的热点.但目前已有的推荐算法,在只有消费记录这种隐性数据情况下,针对用户较少活动区域或新用户的推荐效率较低,无法最大化挖掘隐性数据所带的信息.针对以上问题,结合位置服务平台的特点,针对用户冷启动问题,提出了一种结合协同概率矩阵分解与迭代

收稿日期:2017-03-16;修回日期:2017-06-08

基金项目:国家自然科学基金项目(61602282,61602284);中国博士后科学基金项目(2016M602181);国家社会科学基金项目(14BTQ049)

This work was supported by the National Natural Science Foundation of China (61602282, 61602284), the Postdoctoral Science Foundation of China (2016M602181), and the National Social Science Foundation of China (14BTQ049).

通信作者:郭磊(guolei@sdu.edu.cn)

决策树(gradient boosting decision tree, GBDT)的推荐算法. 该方法首先使用多层协同概率矩阵分解在多个维度上得到用户潜在特征, 然后使用 GBDT 学习算法对特征和标签进行训练得到用户对项目的偏好, 最后使用考虑约束问题的 top- N 推荐产生推荐列表. 在真实数据集上的实验结果表明, 与目前较为流行的方法相比, 提出的方法能在准确率、F1 值上取得较好的结果, 能更好地缓解位置服务中的冷启动问题.

关键词 推荐系统; 位置服务; 概率矩阵分解; 冷启动问题; 约束

中图法分类号 TP391

随着移动设备和位置定位技术的发展, 能够联系线上和线下社会, 方便人们共享和交流位置信息的位置服务(location based service, LBS)^[1]受到人们的广泛关注. 越来越多的人愿意通过位置服务平台来分享他们的实时位置、共享信息, 使用位置服务提升自己的生活(如导航、餐饮消费、购电影票等)水平. 但由于用户数量不断增加, 所产生的信息不断增多, 位置服务平台同样面临着严重的信息过载问题. 推荐系统作为一种有效的信息过滤技术, 由于使我们可以更好地探索用户的偏好, 帮助人们发现他们想要的服务和需求, 受到研究者的广泛关注^[2].

与传统的推荐算法相比, LBS 中的推荐问题面临严重的用户冷启动问题. 报告显示^①国内 LBS 平台美团与大众点评 2014 年新用户的增长率都超过 300%. 如何在用户消费的商家或位置信息很少时(用户冷启动问题), 为他们推荐感兴趣的商家已经成为当前 LBS 推荐中亟需解决的关键问题. 目前已有许多研究人员针对此问题进行了广泛的研究. 例如, Zhang 等人^[3]提出上下文感知的半监督协同训练模型(CSEL), 算法通过不同的上下文信息建立不同的弱预测模型, 然后利用协同训练策略, 每个模型学习其他模型的结果. Vairachilai 等人^[4]基于社区检测算法提出解决冷启动问题的方法, 通过社区检测算法在社交网络上区分用户的相似度. Zhou 等人^[5]使用基于用户的协同过滤和决策树算法预测新用户偏好. 但目前针对用户冷启动问题的研究大多是针对通用算法的优化, 很少直接研究 LBS 中的用户冷启动问题^[6]. 在 LBS 推荐系统中, 用户位置信息的极度稀疏和显式数据的缺乏, 导致传统基于协同过滤的方法推荐效果较差. 因此研究针对 LBS 的用户冷启动问题还存在较大挑战.

另一方面, 在现实生活中, LBS 提供商(例如一个景点或者餐馆)的服务能力通常受到其自身规模的限制, 当消费者的数量超过其服务能力时其服务

质量会下降, 用户也会有较低的满意度^[7]. 因此, 为了进一步提升用户满意度和推荐效果, 需要将商家的服务能力(约束)也同时考虑到推荐系统中, 研究基于实际约束的推荐问题.

针对上述存在的问题, 本文提出一种结合多层协同概率矩阵分解与迭代决策树(gradient boosting decision tree, GBDT)学习算法的推荐算法, 使用用户在其他系统的线上消费记录来丰富用户信息, 缓解冷启动问题. 该方法首先利用多层协同概率矩阵分解将 LBS 数据、线上消费数据相结合, 综合使用 LBS 商家、线上商家和线上商家类别 3 种维度的信息来学习用户的潜在特征. 然后使用 GBDT 学习算法对特征和标签进行训练得到用户对项目的偏好. 在生成推荐列表时, 考虑实际环境下的商家约束问题, 通过改进 top- N 推荐来获得较好的推荐效果. 具体来说, 商家每被推荐一次, 其约束值要根据预测结果进行相应地调整, 当其约束值小于 0 时, 则不能被推荐.

本文的主要贡献有 3 方面:

1) 提出了一种多层概率矩阵分解算法(multi probabilistic matrix factorization, MPMF). 该算法同时对多个矩阵进行分解, 共享用户的潜在特征矩阵, 从多个信息维度更加准确地对用户潜在特征进行估计. 同时该方法作为一种通用的特征学习方法, 可以在不依赖专业知识的情况下获取用户特征, 因此具有较强的可扩展性;

2) 针对现实社会中的“约束”问题, 将约束信息考虑到推荐结果中, 使算法在推荐过程中能同时考虑到 LBS 商家的服务能力, 提升用户的满意度;

3) 在多层概率矩阵分解算法的基础上, 通过将用户特征与 GBDT 学习算法相结合, 提出一种有效缓解用户冷启动的推荐模型, 并在真实数据集上验证了算法的有效性.

① TalkingData-2015 年餐饮 O2O 移动应用行业报告(<http://mi.talkingdata.com/>)

1 相关工作

传统的推荐算法根据使用的信息类型主要可以分为:基于内容的推荐(content-based recommendation)和基于协同过滤的推荐(collaborative filtering-based recommendation).其中,基于内容过滤的推荐,通过发现项目内容之间的相似性,然后根据用户以往的喜好记录,推荐给用户相似的项目.基于协同过滤的推荐主要可以分为3种:

1) 基于用户的协同过滤,使用用户对物品的偏好找到其相似邻居用户,然后将邻居用户喜欢的推

荐给当前用户;

2) 基于项目的协同过滤,利用用户对物品的偏好找到相似的物品,然后根据用户的历史偏好,推荐相似的物品;

3) 基于模型的协同过滤,利用样本的用户喜好信息,训练一个推荐模型,然后根据用户喜好的信息进行预测,计算推荐.

2种方法的优缺点比较如表1所示.基于协同过滤的推荐简单、高效,并且应用广泛,但其难以为没有任何记录的用户或者历史记录稀疏的用户发现相似用户,导致不能为其给出准确的推荐,称为用户冷启动问题^[8].

Table 1 Comparison of Recommendation Methods

表1 推荐方法优缺点比较

Recommendation Methods	Advantages	Disadvantages
Content-based Recommendation	the recommendations are intuitive and easy to interpret; require little field knowledge	sparsity problem; cold start problem; hard handle Complex attributes; require plenty of data
Collaborative Filtering-based Recommendation	find new interest; require little field knowledge and performance improves over time; personalization and automation	sparsity problem; cold start problem; scalability problem; rely on historical data

在LBS推荐系统中,由于用户位置信息的极度稀疏性,LBS推荐面临严重的数据稀疏问题.同时由于LBS的发展迅速产生大量的新用户,面临严重的用户冷启动问题.本文主要针对用户冷启动问题,引入额外用户数据与LBS数据相结合,提出结合多层协同概率矩阵分解与GBDT学习算法的推荐算法.

已有的解决用户冷启动问题的方法,大致可以分为3类:

1) 使用额外数据源,如用户属性数据、用户的观点、社会标签等,从而更好地选择新用户的邻居.例如,Lin等人^[9]使用社交网络数据解决APP推荐中冷启动问题;Lika等人^[10]利用人口统计学信息,首先对用户进行聚类,得到用户分组,在分组中计算用户在不同属性上的相似度,得到整体相似度,得到用户的邻居集,然后将相似度加入到预测评分进行推荐;Zhang等人^[3]提出上下文感知的半监督协同训练模型(CSEL),算法通过不同的上下文信息建立不同的弱预测模型,然后利用协同训练策略,每个模型学习其他模型的结果;Meng等人^[11]使用子社区和实体决策模型等额外信息,帮助解决冷启动情况下的推荐问题.

2) 选择最具可信度的相似邻居,其核心思想是

改进确定用户邻居的方法.例如Zhang等人^[12]利用评分可信度来对物品评分矩阵进行降维.其中物品和用户被共同聚类,处于每个用户簇中的评分被分散处理,从而克服数据稀疏性问题,基于物品和用户簇来预测用户偏好;Vairachilai等人^[4]基于社区检测算法提出解决冷启动问题的方法,通过社区检测算法在该社交网络上区分用户的相似度.

3) 使用混合方法改进预测评分,其核心思想是在确定和新用户相似的用户之后,运用混合方法计算相似性或者产生预测评分.例如Le等人^[13]提出新颖的混合推荐方法HU-FCF++,包含2个过程:①通过人口统计学的信息计算出一个评分;②计算在列表中的评分,通过对人口统计学信息的聚类将用户分组,并找出分组中的新用户.Wang等人^[14]提出一个混合推荐框架来解决用户冷启动问题.首先利用当前用户特征、用户上下文和操作记录用于将用户进行分类.然后,根据用户的类型动态的选择合适的推荐算法,产生推荐列表.

使用额外数据源是一种简单高效解决用户冷启动的方法,但是如何使用额外的数据或者将额外的数据与原有数据相结合,是仍需研究的问题.著名研究者Singh和Gordon^[15]提出协同矩阵分解的方法,

即同时对 2 个矩阵进行分解,共享用户潜在特征矩阵 U ,在 2 个信息维度上求用户的潜在特征. Ji 等人^[16]提出一种包含 3 种因素的矩阵分解模型,用于提高推荐准确度和解决冷启动问题;Forsati 等人^[17]提出一种矩阵分解模型,将信任关系与不信任关系结合起来,提高推荐效果的同时缓解冷启动问题;Ma 等人^[18]提出了 SoRec 算法,该算法将评分矩阵 R 与社会关系矩阵 D 联系起来,实验取得较好结果,特别是在用户很少评分甚至没有评分的情况下. 上述研究表明协同矩阵分解技术可以有效结合 2 方面信息,使用共享用户特征矩阵的方式,得到用户在两个信息维度上的特征.

2 推荐模型

本文算法首先使用多层概率矩阵分解算法学习用户特征,然后使用 GBDT 学习算法训练用户特征得到用户对项目的预测评分,最后采用考虑“约束”问题的 top- N 推荐方法生成推荐列表.

2.1 问题描述

在本文中,我们以真实业务场景“口碑”平台中的商家推荐问题作为主要研究对象.“口碑”平台是一个新兴的 LBS 平台,由于存在较多的新用户与新商家,而面临严重的用户冷启动问题.其中多数用户直接由成熟线上网购平台“淘宝”转化而来,而这些用户在“淘宝”中均具有较多的线上消费记录,这为解决新用户冷启动问题提供了新的思路.我们可以使用丰富的淘宝信息学习用户的偏好,然后应用到口碑系统的推荐中.我们将淘宝数据称为线上数据,口碑数据成为线下数据.用户的线上记录形式如表 2 所示,例如其一条数据为 $(u_1, s_1, \text{apple, fruits, 1, } t)$ 表示用户 u_1 在时间 t 购买了商家 s_1 中种类为 *fruits* 的商品 *apple*.在口碑中,用户通过线上购买商家服务线下体验的方式与商家产生关系.其日志记录形式为表 3 所示.令 $U = \{u_1, u_2, \dots, u_m\}$, $M = \{m_1, m_2, \dots, m_n\}$, $L = \{l_1, l_2, \dots, l_p\}$ 分别表示用户的集合、商家的集合和位置的集合.当用户 u_1 购买并消费 l_1 位置的商家 m_1 ,则系统会产生 (u_1, m_1, l_1, t) 的日志记录, t 为购买时间.

我们设计的推荐算法根据用户的数据信息将用户分为 3 类:1)老用户(old user),也就是存在口碑数据的用户;2)新用户(new user),也就是不存在口碑数据但存在淘宝数据的用户;3)完全的新用户(absolutely new user),也就是既不存在口碑数据也

不存在淘宝数据的用户.对上述 3 类用户分别采用适合的推荐方法,从而提升总体的推荐效率.第 1 类与第 2 类用户采用多层协同概率矩阵分解与 GBDT 结合的方法,第 3 类用户采用基于商家流行度的方法.本文模型主要针对 2 个问题:1)探究如何利用大量的线上数据来丰富用户信息促进线下商家的推荐效果.2)根据实际的业务场景,需要考虑跟 LBS 相关的实际约束,比如商家的服务能力、商家能够提供的折扣数量等.

针对上述 2 个问题,我们使用以下方法:1)采用迁移学习的思想,使用特征学习算法在线上数据中学习用户特征,并使用 GBDT 算法对学习到的特征和标签进行训练,得到用户对商家的偏好信息;2)将约束信息考虑在推荐生成过程,采用改进的 top- N 推荐,不在是将用户预测评分最高的 N 个商家推荐给用户,而是将所有用户对其所在位置的所有商家的预测评分进行排序,商家每被推荐一次其实际约束值减 \hat{r} (\hat{r} 为当前用户对商家的预测评分),当值为零时则不能被推荐,这样既能保证准确度又能满足约束条件.

Table 2 Online User Behavior

表 2 用户线上数据

Field	Description
<i>User_id</i>	unique user id
<i>Seller_id</i>	unique online seller id
<i>Item_id</i>	unique item id
<i>Category_id</i>	unique category id
<i>Online_Action_id</i>	“0” denotes “click” while “1” for “buy”
<i>Time_Stamp</i>	date of the format “yyyymmdd”

Table 3 Users' Shopping Records at Brick-and-Mortar Stores

表 3 用户线下历史数据格式

Field	Description
<i>User_id</i>	unique user id
<i>Merchant_id</i>	unique merchant id
<i>Location_id</i>	unique location id
<i>Time_Stamp</i>	date of the format “yyyymmdd”

2.2 评分预测方法

为了有效地利用信息,缓解冷启动问题,我们采用准确度较高的迭代决策树(gradient boosting decision tree, GBDT)^[19]作为基本评分预测方法. GBDT 是一种迭代的决策树算法,其输出由所有的决策树结论累加产生.在测试集中训练用户特征与标签,学习得到特征与标签的关系,然后预测测试集

中用户特征对应的标签. GBDT 方法具有良好的准确性和扩展性,并且其使用了迭代计算方式,所需的特征向量维度较低,计算时间复杂度较低. 使用 GBDT 学习算法需要对用户数据进行特征提取. 根据文献[20]与生活经验,主要从 4 个维度提取特征:

1) 用户特征. 描述用户的特征,主要包括消费间隔、重复购买行为、最近购买行为等.

2) 商家特征. 描述商家的特征,主要包括生命周期、实际限制、日均销量等.

3) 用户-商家特征. 描述用户对特定商家的偏好特征,主要包括最近购买、生命周期、在该用户的购买排名等.

4) 商家-位置特征. 描述在商家在当前位置的特征,主要包括被购买次数、该商家占当前位置的比例、该商家在当前位置的排名等.

Table 4 Merchant Information

表 4 线下商家数据格式

Field	Description
<i>Merchant_id</i>	unique merchant id
<i>Budget</i>	budget constraints imposed on the merchant
<i>Location_id_list</i>	available location list, e. g. 1:356:89

当维度上升之后训练效率较低,耗时较长,同时 GBDT 不需要太大的维度提升准确度. 我们利用随机森林算法(random forest, RF)^[21]对提取的 32 维特征进行选择,保留 10 维最有效的特征. 虽然 GBDT 方法可以获得比较好的结果,但是只能用于老用户. 新用户没有历史数据无法获得其相关特征,导致方法无法使用. 第 2.3 节和第 3 节主要介绍如何在大量的淘宝数据中学习用户的偏好.

2.3 用户特征学习

针对上述对新用户进行推荐时产生的冷启动问题,本文采用了迁移学习的思想进行解决. 即在线上数据学习用户特征,然后作为 GBDT 算法的输入,为新用户进行推荐.

虽然淘宝商家与口碑商家具有相似之处,但仍然存在不可忽视的差异. 用户在 LBS 的线下偏好与纯网络购物的线上偏好存在差异. 因此直接使用人工提取特征的方式在线上数据中提取用户特征应用到线下数据的推荐中是不合适的. 并且人工选取出来的特征依赖人力和专业知识,不利于推广. 于是我们通过机器学习算法来学习特征,促进特征工程的工作更加快速. 结合本文问题,综合考虑多种有监督学习算法,最终我们选择概率矩阵分解算法 PMF

作为基础的特征学习算法. 在协同矩阵分解的基础上,考虑本文问题,需要在线上 and 线下数据中学习用户潜在特征,提出多层协同概率矩阵分解算法(multi probabilistic matrix factorization, MPMF). MPMF 算法作为本文的主要创新点,其具体细节将在第 3 节进行详细描述.

2.4 生成推荐列表

本文研究的问题,存在‘约束’这一独特的因素. 商家的约束信息主要指服务能力,即其约束值就是可接待用户数量. 我们希望用户量尽可能地接近或等于“约束”值,但不能超过. 为了提高推荐准确度,并满足约束条件,我们改进 top-N 推荐用于推荐列表的生成环节. 生成推荐包括如下 4 个步骤:

1) 根据 GBDT 产生的预测评分 \hat{r} , 将所有的 (*user-id*, *location-id*, *merchant-id*, \hat{r}) 按照 \hat{r} 由大到小排序.

2) 按行执行步骤 1 产生的结果,当预测评分 \hat{r} 大于阈值 α 时,执行步骤 3; 否则计算下一条数据.

3) 用 $L_{user-id}$ 表示当前用户拟推荐列表的长度,初值为 0. 按照第 1 步产生的序列逐条执行下面的运算,在商家“约束”值大于 0 并且 $L_{user-id} < 4$ 的情况下将 *merchant-id* 加入到当前用户的拟推荐列表. 然后 $L_{user-id}$ 加 1, 商家“约束”值减 \hat{r} . 转到步骤 2.

4) 上述步骤执行完毕时,遍历 *user-id* 确定用户的推荐列表.

经过上面的步骤便能得到每个用户的推荐列表,并满足 3 个条件: 1) 预测评分大于阈值,这样做是为了获得较高的准确度; 2) 用户的推荐列表长度小于 4, 这是我们数据中得到的结果,在实验部分会解释; 3) 商家被推荐的次数与其“约束”值与用户对其预测评分有关,每次商家被添加到推荐列表时都要保证用户对其预测评分在当前序列是最高的.

3 用户特征学习

我们采用协同矩阵分解的方式在多维数据上学习用户潜在特征,缓解用户冷启动问题. 提出一种多层协同概率矩阵(MPMF)学习算法,该算法基于协同矩阵分解与概率矩阵分解. 本节首先介绍 PMF 算法的定义与原理. 然后介绍一个协同矩阵分解在推荐系统上的经典应用——SoRec 算法,该算法将评分矩阵 \mathbf{R} 与社会关系矩阵 \mathbf{D} 联系起来,提高推荐效果. 最后给出 MPMF 算法的图模型、目标函数以及最优化求解方式.

3.1 PMF 算法介绍

我们采用概率矩阵分解(PMF)^[22]方法对用户-项目矩阵进行分解,得到用户潜在特征矩阵和项目潜在特征矩阵,其中的每一行或每一列都可以表示一个用户或项目的潜在特征.另外,PMF 一般产生的特征向量维度较低,计算复杂度较低,可以适用于大规模数据集^[23].

假设用户的评分矩阵 \mathbf{R} 中有 m 个用户, n 个推荐对象,其中 $r_{ij} \in [0, 1]$ 表示用户 u_i 对推荐对象 v_j 的偏好. \mathbf{U} 和 \mathbf{V} 分别表示分解得到的与用户和推荐对象相关的 l 维特征矩阵,其列向量 \mathbf{U}_i 和 \mathbf{V}_j 则分别表示相对应的潜在特征向量.由于 PMF 假设可观测评分是由概率线性模型 $\mathbf{U}_i \mathbf{V}_j$ 和高斯观测噪声组成的,因此评分矩阵 \mathbf{R} 的条件概率分布可以定义为

$$p(\mathbf{R} | \mathbf{U}, \mathbf{V}, \sigma_{\mathbf{R}}^2) = \prod_{i=1}^m \prod_{j=1}^n [N(r_{ij} | g(\mathbf{U}_i^T \mathbf{V}_j), \sigma_{\mathbf{R}}^2)]^{I_{ij}^{\mathbf{R}}}, \quad (1)$$

其中, $N(x | \mu, \sigma^2)$ 表示 x 服从均值为 μ 、方差为 σ^2 的高斯分布, $I_{ij}^{\mathbf{R}}$ 是指示函数,表示如果用户 u_i 对项目 v_j 进行了评分,它的值为 1, 否则为 0. 函数 $g(x)$ 表示逻辑函数 $g(x) = 1/(1 + e^{-x})$, 其作用为将 $\mathbf{U}_i^T \mathbf{V}_j$ 的值限制在 $[0, 1]$ 之内. 同时,假设 \mathbf{U} 和 \mathbf{V} 同样服从均值为 0 的高斯先验:

$$p(\mathbf{U} | \sigma_{\mathbf{U}}^2) = \prod_{i=1}^m N(\mathbf{U}_i | 0, \sigma_{\mathbf{U}}^2 \mathbf{I}), \quad (2)$$

$$p(\mathbf{V} | \sigma_{\mathbf{V}}^2) = \prod_{j=1}^n N(\mathbf{V}_j | 0, \sigma_{\mathbf{V}}^2 \mathbf{I}), \quad (3)$$

经过贝叶斯推断,可得到 \mathbf{U} 和 \mathbf{V} 的联合后验概率分布为

$$\begin{aligned} & p(\mathbf{U}, \mathbf{V} | \mathbf{R}, \sigma_{\mathbf{R}}^2, \sigma_{\mathbf{U}}^2, \sigma_{\mathbf{V}}^2) \\ & \propto p(\mathbf{R} | \mathbf{U}, \mathbf{V}, \sigma_{\mathbf{R}}^2) p(\mathbf{U} | \sigma_{\mathbf{U}}^2) p(\mathbf{V} | \sigma_{\mathbf{V}}^2) = \\ & \prod_{i=1}^m \prod_{j=1}^n [N(r_{ij} | g(\mathbf{U}_i^T \mathbf{V}_j), \sigma_{\mathbf{R}}^2)]^{I_{ij}^{\mathbf{R}}} \times \\ & \prod_{i=1}^m N(\mathbf{U}_i | 0, \sigma_{\mathbf{U}}^2 \mathbf{I}) \times \prod_{j=1}^n N(\mathbf{V}_j | 0, \sigma_{\mathbf{V}}^2 \mathbf{I}). \quad (4) \end{aligned}$$

PMF 算法虽然准确得到用户和项目的潜在特征,但只能利用评分矩阵一方面信息,无法满足我们在多维信息上学习用户特征的要求. Ma 等人提出的 SoRec 算法是利用协同矩阵分解的经典算法,该算法结合评分数据与用户间的社交关系数据,实验表明可以提升推荐准确度. 该算法适用于本文问题,下面进行详细介绍.

3.2 基于协同概率矩阵分解的 SoRec 算法

Ma 等人提出的 SoRec 算法通过对用户-项目评分矩阵 \mathbf{R} 与用户社会关系矩阵 \mathbf{D} 的协同分解,得

到用户的潜在特征矩阵 \mathbf{U}_i . 与前面提到的评分矩阵 \mathbf{R} 的概率矩阵分解相似,社会关系矩阵 \mathbf{D} 的条件概率分布定义为

$$p(\mathbf{D} | \mathbf{U}, \mathbf{Q}, \sigma_{\mathbf{D}}^2) = \prod_{i=1}^m \prod_{k=1}^m [N[d_{ik} | g(\mathbf{U}_i^T \mathbf{Q}_k), \sigma_{\mathbf{D}}^2)]^{I_{ik}^{\mathbf{D}}}, \quad (5)$$

其中, $I_{ik}^{\mathbf{D}}$ 为指示函数,当 i 与 k 之间存在直接的社会关系时,取值为 1, 否则为 0. 并假设 \mathbf{U}, \mathbf{Q} 服从均值为 0 的高斯先验:

$$p(\mathbf{U} | \sigma_{\mathbf{U}}^2) = \prod_{i=1}^m N(\mathbf{U}_i | 0, \sigma_{\mathbf{U}}^2 \mathbf{I}), \quad (6)$$

$$p(\mathbf{Q} | \sigma_{\mathbf{Q}}^2) = \prod_{k=1}^n N(\mathbf{Q}_k | 0, \sigma_{\mathbf{Q}}^2 \mathbf{I}). \quad (7)$$

通过对 2 部分信息的联合分解,将 2 部分信息结合起来. 概率图模型如图 1 所示, $\mathbf{U}, \mathbf{V}, \mathbf{Q}$ 的联合后验概率分布可以表示为

$$\begin{aligned} & p(\mathbf{U}, \mathbf{V}, \mathbf{Q} | \mathbf{R}, \mathbf{D}, \sigma_{\mathbf{R}}^2, \sigma_{\mathbf{D}}^2, \sigma_{\mathbf{U}}^2, \sigma_{\mathbf{V}}^2, \sigma_{\mathbf{Q}}^2) \propto \\ & p(\mathbf{R} | \mathbf{U}, \mathbf{V}, \sigma_{\mathbf{R}}^2) p(\mathbf{D} | \mathbf{U}, \mathbf{Q}, \sigma_{\mathbf{D}}^2) \times \\ & p(\mathbf{U} | \sigma_{\mathbf{U}}^2) p(\mathbf{V} | \sigma_{\mathbf{V}}^2) p(\mathbf{Q} | \sigma_{\mathbf{Q}}^2) = \\ & \prod_{i=1}^m \prod_{j=1}^n [N(r_{ij} | g(\mathbf{U}_i^T \mathbf{V}_j), \sigma_{\mathbf{R}}^2)]^{I_{ij}^{\mathbf{R}}} \times \\ & \prod_{i=1}^m \prod_{k=1}^m [N[d_{ik} | g(\mathbf{U}_i^T \mathbf{Q}_k), \sigma_{\mathbf{D}}^2)]^{I_{ik}^{\mathbf{D}}} \times \\ & \prod_{i=1}^m N(\mathbf{U}_i | 0, \sigma_{\mathbf{U}}^2 \mathbf{I}) \times \prod_{j=1}^n N(\mathbf{V}_j | 0, \sigma_{\mathbf{V}}^2 \mathbf{I}) \times \\ & \prod_{k=1}^m N(\mathbf{Q}_k | 0, \sigma_{\mathbf{Q}}^2 \mathbf{I}). \quad (8) \end{aligned}$$

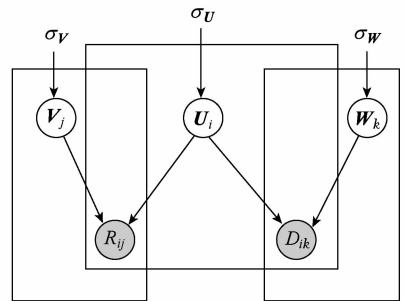


Fig. 1 Graphic model for SoRec

图 1 SoRec 概率图模型

3.3 多层的协同概率矩阵分解算法 MPMF

本文在协同概率矩阵分解的基础上,提出多层的协同概率矩阵分解. 结合本文研究问题,参考相关文献^[24-25]在用户-项目矩阵 \mathbf{R} 、用户-类别矩阵 \mathbf{C} 和用户-店铺矩阵 \mathbf{S} 三个维度信息上学习用户特征. 该方法通过将上述 3 个矩阵协同时分解,得到的用户潜在特征结合 3 方面信息,并且更加准确,通过实验证明可以提升推荐效果. 其概率图模型如图 2 所示:

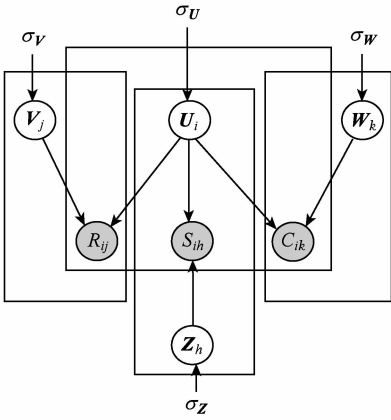


Fig. 2 Graphic model for MPMF

图2 MPMF图模型

$\mathbf{R}, \mathbf{C}, \mathbf{S}$ 的对数联合后验概率分布表示为

$$\ln p(\mathbf{U}, \mathbf{V}, \mathbf{W}, \mathbf{Z} \mid \mathbf{R}, \mathbf{C}, \mathbf{S}, \sigma_{\mathbf{R}}^2, \sigma_{\mathbf{C}}^2, \sigma_{\mathbf{S}}^2, \sigma_{\mathbf{U}}^2, \sigma_{\mathbf{V}}^2, \sigma_{\mathbf{W}}^2, \sigma_{\mathbf{Z}}^2) =$$

$$-\frac{1}{2\sigma_{\mathbf{R}}^2} \sum_{i=1}^m \sum_{j=1}^n I_{ij}^{\mathbf{R}} (r_{ij} - g(\mathbf{U}_i^{\mathbf{T}} \mathbf{V}_j))^2 -$$

$$\frac{1}{2\sigma_{\mathbf{C}}^2} \sum_{i=1}^m \sum_{k=1}^q I_{ik}^{\mathbf{C}} (c_{ik} - g(\mathbf{U}_i^{\mathbf{T}} \mathbf{W}_k))^2 -$$

$$\frac{1}{2\sigma_{\mathbf{S}}^2} \sum_{i=1}^m \sum_{h=1}^t I_{ih}^{\mathbf{S}} (s_{ih} - g(\mathbf{U}_i^{\mathbf{T}} \mathbf{Z}_h))^2 -$$

$$\frac{1}{2\sigma_{\mathbf{U}}^2} \sum_{i=1}^m \mathbf{U}_i^{\mathbf{T}} \mathbf{U}_i - \frac{1}{2\sigma_{\mathbf{V}}^2} \sum_{j=1}^n \mathbf{V}_j^{\mathbf{T}} \mathbf{V}_j - \frac{1}{2\sigma_{\mathbf{W}}^2} \sum_{k=1}^q \mathbf{W}_k^{\mathbf{T}} \mathbf{W}_k -$$

$$\frac{1}{2\sigma_{\mathbf{Z}}^2} \sum_{h=1}^t \mathbf{Z}_h^{\mathbf{T}} \mathbf{Z}_h - \frac{1}{2} \left(\sum_{i=1}^m \sum_{j=1}^n I_{ij}^{\mathbf{R}} \right) \ln \sigma_{\mathbf{R}}^2 -$$

$$\frac{1}{2} \left(\sum_{i=1}^m \sum_{k=1}^q I_{ik}^{\mathbf{C}} \right) \ln \sigma_{\mathbf{C}}^2 - \frac{1}{2} \left(\sum_{i=1}^m \sum_{h=1}^t I_{ih}^{\mathbf{S}} \right) \ln \sigma_{\mathbf{S}}^2 -$$

$$\frac{1}{2} ml \ln \sigma_{\mathbf{U}}^2 - \frac{1}{2} nl \ln \sigma_{\mathbf{V}}^2 -$$

$$\frac{1}{2} ql \ln \sigma_{\mathbf{W}}^2 - \frac{1}{2} tl \ln \sigma_{\mathbf{Z}}^2 + a, \quad (9)$$

其中, a 为常量, 当参数确定时, 最大化式(9), 相当于最小化损失函数式为

$$E = \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^n I_{ij}^{\mathbf{R}} (r_{ij} - g(\mathbf{U}_i^{\mathbf{T}} \mathbf{V}_j))^2 +$$

$$\frac{\lambda_{\mathbf{C}}}{2} \sum_{i=1}^m \sum_{k=1}^q I_{ik}^{\mathbf{C}} (c_{ik} - g(\mathbf{U}_i^{\mathbf{T}} \mathbf{W}_k))^2 +$$

$$\frac{\lambda_{\mathbf{S}}}{2} \sum_{i=1}^m \sum_{h=1}^t I_{ih}^{\mathbf{S}} (s_{ih} - g(\mathbf{U}_i^{\mathbf{T}} \mathbf{Z}_h))^2 +$$

$$\frac{\lambda_{\mathbf{U}}}{2} \|\mathbf{U}\|_F^2 + \frac{\lambda_{\mathbf{V}}}{2} \|\mathbf{V}\|_F^2 + \frac{\lambda_{\mathbf{W}}}{2} \|\mathbf{W}\|_F^2 + \frac{\lambda_{\mathbf{Z}}}{2} \|\mathbf{Z}\|_F^2,$$

$$(10)$$

其中, $\lambda_{\mathbf{C}} = \sigma_{\mathbf{R}}^2 / \sigma_{\mathbf{C}}^2$, $\lambda_{\mathbf{S}} = \sigma_{\mathbf{R}}^2 / \sigma_{\mathbf{S}}^2$, $\lambda_{\mathbf{U}} = \sigma_{\mathbf{R}}^2 / \sigma_{\mathbf{U}}^2$, $\lambda_{\mathbf{V}} = \sigma_{\mathbf{R}}^2 / \sigma_{\mathbf{V}}^2$,

$\lambda_{\mathbf{W}} = \sigma_{\mathbf{R}}^2 / \sigma_{\mathbf{W}}^2$, $\lambda_{\mathbf{Z}} = \sigma_{\mathbf{R}}^2 / \sigma_{\mathbf{Z}}^2$, $\|\cdot\|_F$ 表示 Frobenius 范数. 通过在 $\mathbf{U}, \mathbf{V}, \mathbf{W}, \mathbf{Z}$ 上使用梯度下降算法求目标函数 E 的局部极小值. $\mathbf{U}, \mathbf{V}, \mathbf{W}, \mathbf{Z}$ 的梯度更新公式为

$$\begin{aligned} \frac{\partial E}{\partial \mathbf{U}_i} &= \sum_{j=1}^n I_{ij}^{\mathbf{R}} g'(\mathbf{U}_i^{\mathbf{T}} \mathbf{V}_j) (g(\mathbf{U}_i^{\mathbf{T}} \mathbf{V}_j) - r_{ij}) \mathbf{V}_j + \\ &\lambda_{\mathbf{C}} \sum_{k=1}^q I_{ik}^{\mathbf{C}} g'(\mathbf{U}_i^{\mathbf{T}} \mathbf{W}_k) (g(\mathbf{U}_i^{\mathbf{T}} \mathbf{W}_k) - c_{ik}) \mathbf{W}_k + \\ &\lambda_{\mathbf{S}} \sum_{h=1}^t I_{ih}^{\mathbf{S}} g'(\mathbf{U}_i^{\mathbf{T}} \mathbf{Z}_h) (g(\mathbf{U}_i^{\mathbf{T}} \mathbf{Z}_h) - s_{ih}) \mathbf{Z}_h + \lambda_{\mathbf{U}} \mathbf{U}_i, \end{aligned} \quad (11)$$

$$\begin{aligned} \frac{\partial E}{\partial \mathbf{V}_j} &= \sum_{i=1}^m I_{ij}^{\mathbf{R}} g'(\mathbf{U}_i^{\mathbf{T}} \mathbf{V}_j) (g(\mathbf{U}_i^{\mathbf{T}} \mathbf{V}_j) - r_{ij}) \mathbf{U}_i + \\ &\lambda_{\mathbf{V}} \mathbf{V}_j, \end{aligned} \quad (12)$$

$$\begin{aligned} \frac{\partial E}{\partial \mathbf{W}_k} &= \lambda_{\mathbf{C}} \sum_{i=1}^m I_{ik}^{\mathbf{C}} g'(\mathbf{U}_i^{\mathbf{T}} \mathbf{W}_k) (g(\mathbf{U}_i^{\mathbf{T}} \mathbf{W}_k) - c_{ik}) \mathbf{U}_i + \\ &\lambda_{\mathbf{W}} \mathbf{W}_k, \end{aligned} \quad (13)$$

$$\begin{aligned} \frac{\partial E}{\partial \mathbf{Z}_h} &= \lambda_{\mathbf{S}} \sum_{i=1}^m I_{ih}^{\mathbf{S}} g'(\mathbf{U}_i^{\mathbf{T}} \mathbf{Z}_h) (g(\mathbf{U}_i^{\mathbf{T}} \mathbf{Z}_h) - s_{ih}) \mathbf{U}_i + \\ &\lambda_{\mathbf{Z}} \mathbf{Z}_h, \end{aligned} \quad (14)$$

其中, $g'(x) = \exp(x) / (1 + \exp(x))^2$ 是回归函数 $g(x)$ 的导数.

4 实 验

在实验中, 为验证我们所提出算法的有效性, 主要验证了 2 个问题: 1) 比较 MPMF 方法与已有推荐算法对冷启动用户的推荐效果. 2) 方法中参数对实验结果的影响.

4.1 数据集

数据为由阿里天池大数据科研平台提供^①. 该数据集共包括 3 部分: 1) 线上(淘宝)部分交易记录, 共 963 923 用户的 44 528 127 条记录; 2) 线下(口碑)数据, 包括测试集与训练集两部分. 其中训练集中包括 1 081 724 条数据、230 783 用户、6 039 商家、426 个位置, 测试集包括 473 533 用户-位置对, 但是仅有 24.5% 出现在训练集中即测试集中的老用户仅有 24.5%; 3) 口碑商家数据, 包括商家实际“约束”与商家的分布位置.

经过简单的数据分析显示, 用户平均访问 1.05 个位置与 1.25 个商家, 表明用户趋向于消费同一商家. 由此我们确定推荐列表的长度不超过 4. 口碑数

① <https://tianchi.aliyun.com>

据中新用户的比例为 75.5%，用户占比、商家周活跃度、用户周活跃度如图 3~5 所示. 上述数据均表明口碑网是快速发展的, 存在严重的冷启动问题.

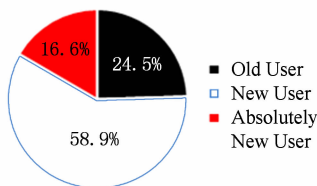


Fig. 3 User Scale

图 3 用户比例

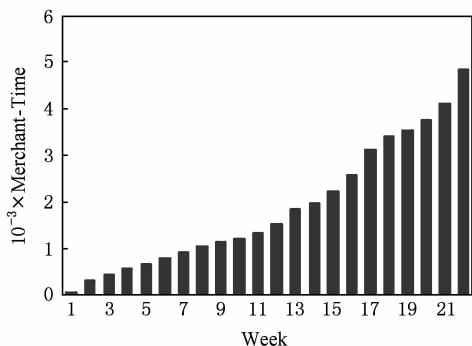


Fig. 4 Merchant week activity

图 4 商家周活跃度

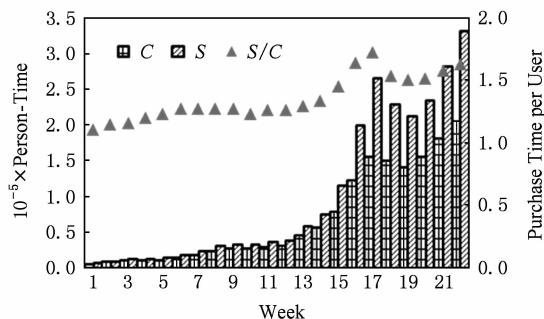


Fig. 5 User week activity

图 5 用户周活跃度

图 4 表明商家的数量不断增加, 且增幅较大. 图 5 中 C 表示一个周的有记录用户数, 即口碑网的周活跃用户量; S 表示一个周的记录数 (左垂直坐标轴), 即口碑网的销量; 则 S/C 代表平均一个用户的购买数 (右垂直坐标轴). 横坐标表示时间, 距离数据最早时间 (2015.07.01) 的周数. 通过图 5 可以看出销量增长迅速.

图 6 为一个位置所含商家的数量的关系, 如图 6 所示超过 75% 的位置所含商家数量少于 20, 所以我们预测用户对其当前所在位置的所有商家的评分是可行的.

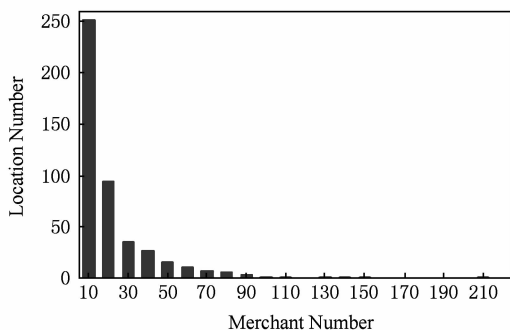


Fig. 6 Merchant number in location

图 6 位置所含商家数量分布

4.2 评价方法

本文使用准确率、召回率和 $F1$ 值作为评价指标. 本文采用的 $F1$ 值计算方法与传统有差异, 主要是为了考虑具体问题中的商家“budget”因素, 计算为

$$P = \frac{\sum_i \|\min(|S_i \cap \hat{S}_i|, b_i)\|}{\sum_i |\hat{S}_i|}, \quad (15)$$

$$R = \frac{\sum_i \|\min(|S_i \cap \hat{S}_i|, b_i)\|}{\sum_i \| \min(|S_i|, b_i) \|}, \quad (16)$$

$$F1 = \frac{2 \times P \times R}{P + R}, \quad (17)$$

其中, b_i 表示商家 i 的“budget”, S_i 表示真实的 $(user-id, location-id, merchant-id)$, \hat{S}_i 则表示预测的 $(user-id, location-id, merchant-id)$.

4.3 结果比较

为了验证提出方法的有效性, 我们比较了 UserCF, GBDT, PMF, SoRec 等主流的推荐算法. 由于数据集的格式为日志记录, 我们首先对数据进行了预处理, 将其转化为评分形式.

淘宝数据的处理方法为: 1) 用户对项目 (商品、商家、类别) 有过点击行为, 但没有购买行为, 其评分为 0. 2) 有过购买行为, 则用用户对项目的购买次数与该项目被单一用户购买次数最大值的比值作为用户对项目的评分. 经过上述处理, 评分的取值区间为 $[0, 1]$.

口碑数据的处理方法: 1) 与淘宝数据相似, 用户对项目的购买次数与该项目被单一用户购买次数最大值的比值作为用户对项目的评分. 2) 口碑数据为购买记录, 可以认为全为正例, 因此需要生成负例 (GBDT 算法需要). 方法为: 用户到过的位置的销量前 3 的商家中, 如果存在用户未对其有购买行为, 则代表用户可能不喜欢, 其评分为 0. 我们选择销量前 3

的原因为销量大被用户知道的可能性越大. 在用户知道商家存在的前提而未购买, 则很可能是不喜欢的.

本文方法按照数据来源将用户分为 3 类, 其中第 3 类用户不具有任何历史数据, 所以采用基于商家流行度的方法. 具体为将用户当前所在位置的最流行的前 3 名商家作为候选集. 我们选择前 3 的数据依据, 历史数据显示某一位置销量前 3 名的商家的总销量占当前位置总销量比率超过 65% 的几率为 82%, 这说明绝大多数位置用户选择前 3 流行商家的可能性大于 65%. 商家流行度的计算即为商家总销量比上当前位置的总销量. SoRec 方法的实现在本文有 2 种选择: 1) 用户-项目矩阵 R 加用户-类别矩阵 C ; 2) 用户-项目矩阵 R 加用户-店铺矩阵 S . 在淘宝数据集上经过实验证明, 第 2 种 R 加 S 效果优于第 1 种. 下面涉及的 SoRec 算法均为 R 加 S 的协同概率矩阵分解.

GBDT 算法的参数设置: 树的个数为 100, 步长为 0.01, 深度为 8. MPMF 的参数设置: $\lambda_s = 10$, $\lambda_u = \lambda_v = \lambda_w = \lambda_z = 0.01$. 在 MPMF 算法中参数 λ_c 非常重要, 它决定了矩阵 C 与矩阵 S 对用户特征的影响权重. 当 $\lambda_c = 0$ 时, 该方法仅使用 R 加 S , 即成为 SoRec 方法; 当 $\lambda_c \gg \lambda_s$ 时, 相当于仅使用 R 加 C , 其结果差于 SoRec 方法. 下面将对 λ_c 的取值进行讨论, 以获得较好的结果. 图 7 为 λ_c 取值与 $F1$ 值的关系, 分别取 λ_c 的值为 0, 1, 5, 10, 15, 当 $\lambda_c = 5$ 时结果最优.

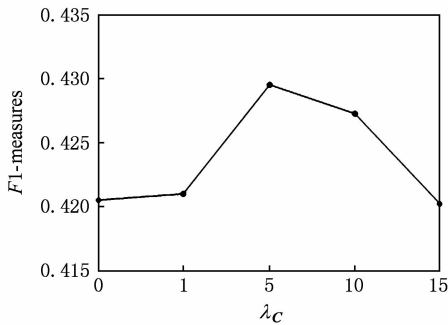


Fig. 7 The relationship of λ_c

图 7 MPMF 结果与 λ_c 取值的关系

由于最终的推荐结果要经过“约束”信息与阈值的筛选, 导致针对某一用户的可推荐商家数目不定, 所以我们不固定推荐列表的长度, 但是设置其长度的最大值, 保证推荐的准确度. 根据前文所述的数据分析结果, 我们将推荐列表长度的最大值设置为 4. 表 5 表示了当推荐列表长度最大值为 4, 阈值为 0.2 时各种方法的结果 ($\alpha = 0.2$ 为多次试验取最好结果时的选择, 实验中我们将 α 初值设置为 0, 每次加

0.05, 最大值为 0.5). 结果表明: 本文提出的方法 GBDT+MPMF 在准确度和 $F1$ 值上均优于对比算法, 证明其有效性. 本文方法对用户进行分类, 针对不同类别采用合适的方法, 不同方法取得的实验结果如表 6 所示.

Table 5 The Results of Method

表 5 结果比较 (推荐列表的长度 $L=4$, $\alpha=0.2$)

Data Set	Method	Precision	Recall	F1-measures
Koubei	User-CF	0.1173	0.6012	0.1963
Koubei and taobao	GBDT	0.2607	0.7634	0.3872
Koubei and taobao	GBDT+PMF	0.2751	0.7497	0.4025
Koubei and taobao	GBDT+SoRec	0.2936	0.7412	0.4206
Koubei and taobao	GBDT+MPMF	0.3027	0.7397	0.4296

Table 6 The Results of Diffient User Categories

表 6 不同用户分类方法的结果

User category	Method	Precision	Recall	F1-measures
Old User	MMF+GBDT	0.7229	0.8798	0.7937
New User	SoRec+GBDT	0.2793	0.7136	0.4015
Absolutely New User	Merchant Popularity Based Method	0.1622	0.2677	0.2017

由于商家“约束”信息的存在, 推荐系统尽量推荐用户最可能消费的商家. 使用 2.4 节描述的方法, 将所有候选推荐集合按照预测概率由大到小排序, 并经过阈值的筛选. 这样可以在保证一定准确度的情况下, 平衡准确度与召回率, 以获得较好的 $F1$ 值. “约束”信息会筛掉一部分候选集对召回率影响较大. 不考虑“约束”信息时我们的方法效果如表 7 所示. 与不考虑“约束”信息相比, 考虑约束信息时, 算法的准确率、召回率、 $F1$ 值均下降, 算法的整体性能降低了接近 15%.

Table 7 The Influence of Budget

表 7 Budget 对推荐效果的影响

Budget	Precision	Recall	F1-measures
With budget	0.3027	0.7397	0.4296
Without budget	0.3517	0.8506	0.4976

5 总 结

本文提出了结合多维信息的推荐方法, 该方法结合多层概率矩阵分解与 GBDT 算法, 并给出了将多层信息结合到一起的推荐框架. 同时, 考虑实际问题中的“实际限制”. 在真实数据集上的实验结果表明,

该方法比已有方法取得更好的推荐效果. 虽然该算法是在具体问题中得到启发, 但使用特征学习算法学习用户特征, 不依赖专业的行业知识和 GBDT 学习算法本身的通用性, 使得本文的方法具有较强的可扩展性.

本文研究问题中的“约束”因素具有很强研究性, 在未来的工作中我们将会更多考虑这个因素, 比如设计在算法的损失函数中.

参 考 文 献

- [1] Liu Shudong, Meng Xiangwu. Approach to network services recommendation based on mobile users' location [J]. *Journal of Software*, 2014, 25(11): 2556-2574 (in Chinese)
(刘树栋, 孟祥武. 一种基于移动用户位置的网络服务推荐方法[J]. *软件学报*, 2014(11): 2556-2574)
- [2] Kang Zhao, Peng Chong, Cheng Qiang. Top-N recommender system via matrix completion [C] //Proc of the 30th AAAI Conf on Artificial Intelligence (AAAI-16). Menlo Park, CA: AAAI, 2016: 179-185
- [3] Zhang Mi, Tang Jie, Zhang Xuchen, et al. Addressing cold start in recommender systems: A semi-supervised co-training algorithm [C] //Proc of the 37th Int ACM SIGIR Conf on Research & Development in Information Retrieval. New York: ACM, 2014: 73-82
- [4] Vairachilai S, Kavithadevi M, Raja M. Alleviating the cold start problem in recommender systems based on modularity maximization community detection algorithm [J]. *Circuits and Systems*, 2016, 7(8): 1268-1279
- [5] Zhou Ke, Yang Shuanghong, Zha Hongyuan. Functional matrix factorizations for cold-start recommendation [C] //Proc of the 34th Int ACM SIGIR Conf on Research and Development in Information Retrieval. New York: ACM, 2011: 315-324
- [6] Yao Yichen, Li Zhongjie. Cold-start solution to location-based entity shop recommender systems using online sales records [C] //Proc of the 2nd Int Workshop on Social Influence Analysis(SocInf 2016). Menlo Park, CA: AAAI, 2016: 57-67
- [7] Hu Ke, Li Xiangyang, Wu Chaotian. Cold start purchase prediction with budgets constraints [C] //Proc of the 2nd Int Workshop on Social Influence Analysis(SocInf 2016). Menlo Park, CA: AAAI, 2016: 68-80
- [8] Li xin, Liu Guiquan, Li Lin, et al. Circle-based and social connection embedded recommendation in LBSN [J]. *Journal of Computer Research and Development*, 2017, 54(2): 394-404 (in Chinese)
(李鑫, 刘贵全, 李琳, 等. LBSN上基于兴趣圈中社会关系挖掘的推荐算法[J]. *计算机研究与发展*, 2017, 54(2): 394-404)
- [9] Liu N N, Meng Xiangrui, Liu Chao, et al. Wisdom of the better few: Cold start recommendation via representative based rating elicitation [C] //Proc of the 5th ACM Conf on Recommender Systems. New York: ACM, 2011: 37-44
- [10] Lika B, Kolomvatsos K, Hadjiefthymiades S. Facing the cold start problem in recommender systems [J]. *Expert Systems with Applications*, 2014, 41(4): 2065-2073
- [11] Meng Chen, Yang Cheng, Chen Jiechao, et al. A method to solve cold-start problem in recommendation system based on social network sub-community and ontology decision model [J]. *Acta Neurochirurgica*, 2013, 156(3): 577-580
- [12] Zhang Daqing, Zou Qin, Xiong Haoyi. CRUC: Cold-start recommendations using collaborative filtering in Internet of things [J]. *Computer Science*, 2013, 13(6): 3454-3461
- [13] Le H S. HU-FCF++: A novel hybrid method for the new user cold-start problem in recommender systems [J]. *Engineering Applications of Artificial Intelligence*, 2015, 41(3): 207-222
- [14] Wang J H, Chen Yihao. A distributed hybrid recommendation framework to address the new-user cold-start problem [C] //Proc of the 13th IEEE Int Conf on Ubiquitous Intelligence and Computing. Piscataway, NJ: IEEE, 2015: 1686-1691
- [15] Singh A P, Gordon K. Relational learning via collective matrix factorization [C] //Proc of the 14th ACM SIGKDD Int Conf on Knowledge Discovery and Data Mining. New York: ACM, 2008: 650-658
- [16] Ji Ke, Shen Hong. Addressing cold-start: Scalable recommendation with tags and keywords [J]. *Knowledge-Based Systems*, 2015, 83(1): 42-50
- [17] Forsati R, Mahdavi M, Shamsfard M, et al. Matrix factorization with explicit trust and distrust side information for improved social recommendation [J]. *ACM Trans on Information Systems*, 2014, 32(4): No. 17
- [18] Ma Hao, Yang Haixuan, Lyu M R, et al. SoRec: Social recommendation using probabilistic matrix factorization [C] //Proc of the 18th ACM Conf on Information and Knowledge Management. New York: ACM, 2008: 931-940
- [19] Xie Jianjun, Coggeshall S. Prediction of transfers to tertiary care and hospital mortality: A gradient boosting decision tree approach [J]. *Statistical Analysis & Data Mining*, 2010, 3(4): 253-258
- [20] Wei Hao, Shi Bei, Chen Junwen. Location based services recommendation with budget constraints [C] //Proc of the 2nd Int Workshop on Social Influence Analysis (SocInf 2016). Menlo Park, CA: AAAI, 2016: 48-56
- [21] Breiman L. Random forest [J]. *Machine Learning*, 2001, 45(3): 5-32
- [22] Salakhutdinov R, Mnih A. Probabilistic matrix factorization [C] //Proc of the 20th Int Conf on Neural Information Processing Systems. New York: Curran Associates Inc, 2007: 1257-1264

- [23] Guo Lei, Ma Jun, Chen Zhumin, et al. Incorporating item relations for social recommendation [J]. Chinese Journal of Computers, 2014, 37(1): 219-228 (in Chinese)
(郭磊, 马军, 陈竹敏, 等. 一种结合推荐对象间关联关系的社会化推荐算法[J]. 计算机学报, 2014, 37(1): 219-228)
- [24] Zhang Weiyu, Wu Bin, Geng Yushui, et al. Joint rating and trust prediction based on collective matrix factorization [J]. Chinese Journal of Electronics, 2016, 44(7): 1581-1586 (in Chinese)
(张维玉, 吴斌, 耿玉水, 等. 基于协同矩阵分解的评分与信任联合预测[J]. 电子学报, 2016, 44(7): 1581-1586)
- [25] Hernando A, Bobadilla J, Ortega F, et al. A probabilistic model for recommending to new cold-start non-registered users [J]. Information Sciences, 2016, 376(1): 216-232



Gao Yukai, born in 1993. Master candidate of Shandong Normal University. Student member of CCF. His main research interests include recommender system and POI recommendation.



Wang Xinhua, born in 1970. Professor and master supervisor in Shandong Normal University. Received his master's degree from Dalian University of Technology. His main research interest include distributed network and recommendation system.



Guo Lei, born in 1983. Received his PhD degree in computer architecture from Shandong University, Jinan, in 2015. His main research interests include information retrieval, social network and recommender system.



Chen Zhumin, born in 1977. Associate professor and master supervisor in Shandong University. Senior member of CCF. His main research interests include Web information retrieval, data mining.

《信息安全研究》期刊简介

习近平总书记指出“没有网络安全就没有国家安全,没有信息化就没有现代化”。数字时代信息安全工具的大众化是不可阻挡的历史潮流。大众化的信息安全已经直接影响到我们每个人的利益,信息安全已成为国家、地方区域经济结构优化提升和转型发展的新机遇。在信息安全上升为国家战略、行业迎来崭新发展机遇形势下,《信息安全研究》期刊应时代而生。

《信息安全研究》是由国家发改委主管、国家信息中心主办的中文学术期刊,其宗旨是集中展示和报道国际、国内网络和信息安全研究领域研究成果及最新应用,传播信息安全基础理论和技术策略,服务国家信息安全形势发展需要。所刊登的论文均经过专家严格评审。

《信息安全研究》于2015年10月创刊发行。刊期为月刊,每期96页,由《信息安全研究》杂志社出版,国内外公开发行。

《信息安全研究》将以研究致以应用,搭建信息安全领域的学术交流平台,愿意和同行业及社会各界建立联系,友好合作,共赢美好未来。欢迎大家积极投稿、赐稿,洽谈合作。

投稿邮箱:ris@cei.gov.cn

编辑部联系人:崔先生(185 0008 6481)

马先生(158 1058 2450)