

基于多源异构数据的移动社交网络个性化推荐研究

许秀

(湖南应用技术学院计算机应用研究中心 湖南 常德 41500)

[摘要]随着信息技术的发展,互联网使用人数大规模增加,网络数据也得到了大量的积累,解决网络信息过载问题成为一个巨大的挑战。社交网络是人们日常生活中重要的交流互动平台,丰富移动社交网络的个性化推荐,有利于满足人们的需求。由于单一的推荐模型已经无法高效处理异构数据,所以通过构建实现数据共享的多源异构数据融合知识库,结合多源异构数据对用户情景偏好进行分析,设计一种融合多源异构数据的移动社交网络个性化推荐框架及算法,来提高移动社交网络推荐的个性化。通过对基于多源异构数据的移动社交网络个性化推荐研究和分析,为此提供借鉴和帮助。

[关键词]多源异构数据;移动社交网络;个性化推荐;研究

[DOI] 10.12252/j.issn.2096-627X.2021.11.024

移动用户通常会受到社交网络上好友的推荐和分享的直接或间接影响,好友的推荐会增加移动用户对此的信任程度。基于移动社交网络的融合多源异构数据,将社交关系、推荐评分、分享评论等多源异构数据进行融合,利用有效数据分析出用户需求和服务偏好,提高移动社交平台中个性化服务推荐的准确度。移动社交网络推荐的个性化发展、推荐系统智能化革新,有利于帮助移动用户快速获得感兴趣的信息,提升用户的体验感受,增加用户的活跃度,有效提高了电商的经济利益。

一、基于多源异构数据的移动社交网络个性化研究内容

1.1对多源异构数据的融合方法进行研究

由于单一的推荐算法无法满足大规模的移动社交网络数据,所以将推荐算法融合成混合推荐算法模式为移动用户提高个性化推荐。这种混合推荐算法主要是从用户特征上进行融合,将不同的多源异构数据放在统一的框架内进行分析,从而得到推荐结果。首先,从不同情景下获取用户数据,包括从平台系统内部直接获取的个人信息、历史信息以及从系统外部得到的位置信息、交互信息等,使用微格式对数据进行描述、集成等相关处理,得到初步的用户数据。其次,将获取的用户数据进行预处理,过滤掉语义不明、作用较小的用户数据。最后,对预处理后的数据进行融合操作,使用改进的RBF神经网络设置权值,利用D-S证据理论融合用户多源异构数据,从而构建用户知识库。

1.2融合情景的移动社交网络用户偏好分析

根据用户多源数据融合知识库,通过贝叶斯网络来分析用户所处情景和其偏好特征的复杂关系,获取用户在情景状态下的特定偏好。根据用户已经选择的物品来获得用户的爱好,从而判断用户是否会喜欢新的物品。同时可以根据用户买过的物品,分析该物品的特征和属性,进而对用户偏好进行分析。除此之外,还可以结合信息熵理论与情景因素来计算各情景的信息熵值,筛选出对用户偏好具有较大影响的情景实例,根据获得的关键情景实例来进一步分析用户在该情景下对商品的偏好程度。比如说A、B、C三个用户分别购买了A、B、C三种款式的衣服,对A、B、C三款衣服特征进行分析,A、C两款衣服风格类型很相似,由于A用户喜欢A款风格的衣服,因此就可以得到A用户对A款运动风格的衣服类型感

兴趣,从而推荐系统可以将A衣服在风格类型上相似的C衣服推荐给用户A。根据用户的喜好进行物品推荐,提高了移动社交网络的个性化水平。

1.3构建基于多源异构数据融合的移动社交网络个性化推荐框架

根据用户偏好及关键情景,建立一个结合情景及社会网络关系的移动社交网络个性化推荐框架。首先,计算关键情景下的用户相似度以及社会网络关系下的用户信任度;由于大多数用户的偏好与好友的偏好极为相似,用户也会受到社交好友偏好的影响。将用户的社交好友关系也引入到推荐系统中来,在社交网络的数据中计算出用户之间的相似度,将评论、评分等多源异构数据进行统一并得到最后的推荐结果,更准确地预测用户的选择,提高了系统推荐的准确性。其次,通过对两者数据的融合操作推理出用户的综合情景相似度,进而搜寻用户的最近邻居。计算用户之间的户距,来判断用户之间的相似性程度,得出用户的最近邻居。直接好友与间接好友的影响是不同的,用户对直接好友的用户信任度更高,相似偏好相似度也更高。最后,依据用户最近邻数据,计算用户对未评分项目的潜在偏好值,基于用户偏好度来为用户进行服务推荐。

二、基于多源异构数据的移动社交网络个性化研究过程

首先,通过各种智能终端、传感器设备对情景因素(自然情景、社会情景和个人信息特征等)下的移动社交网络用户多源数据进行收集,包括用户在网站门户存入的基本信息、个人偏好等特征数据,和用户在移动平台进行的浏览、交易等操作数据以及在社交软件上的评论、收藏、转发等行为数据。再对收集到的用户数据进行预处理,将来源、格式、语义逻辑各不相同的数据资源排除噪音、去冗余杂,留下对推荐系统有价值的用户数据最为推荐证据,减少无用数据对系统分析的干扰。并采取改进的证据理论进行数据融合,利用基于可信度的证据体修改和基于局部冲突分配的合成规则来融合证据,得到逻辑清晰、语义明确的用户多源数据融合知识库。其次,利用知识库中的情景数据、用户偏好信息以及商品属性类型来构建贝叶斯网络模型,通过贝叶斯网络框架来系统地表示各种变量的信息,利用条件概率以图形化的方式描述各个信息要素之间的因果关系,根据不同情

景之间的因果关系进而捕获情景化的用户偏好。并利用信息熵理论计算用户在某情景因素所包含的不同情景实例下对所选择商品的熵值,采用情景信息熵来表达该情景条件下选择不同属性类别的商品混乱程度,进而衡量该情景因素所包含的各个实例对用户选择结果的贡献程度,识别出个性化推荐中对用户选择行为或偏好产生影响的关键情景因素。根据得到的关键情景不断调整用户偏好贝叶斯模型,去除网络模型中对目标用户的偏好不起作用的情景因素,并计算用户偏好度。最后,采用协同过滤的推荐思想,利用关键情景用户偏好向量和社会网络关系矩阵,对基于关键情景的用户相似度和社会网络关系的用户信任度进行计算,然后对两者加权融合,获得用户的综合情景相似度,从而发现用户的最近邻,通过计算用户在情景约束下的对未评分项目的潜在偏好预测来为用户推荐Top-N偏好值最大的服务项目。目前,个性化推荐在移动社交网络平台上的应用已非常广泛,通过研究基于多源异构数据的个性化推荐系统,将有效地刻画用户偏好,为用户定制个性化的具有层次体系的信息,运用现有的J2EE结构可以实现整个系统的跨平台、分布式、协同化、智能化。此外,本课题由高校和企业共同研发和推广,具备成熟的项目开发团队及有利于项目开发的管理经验,并且研发基地已具备研究实验需要的复杂应用环境和完善的网络环境。

三、基于多源异构数据的移动社交网络个性化研究创新和问题解决

3.1 基于多源异构数据的移动社交网络个性化研究需解决的问题

3.1.1 构建多源异构数据融合知识库

将基于情景状态下移动社交网络用户的位置信息、互动信息、网页操作信息和基本信息作为证据体进行融合,采用优化的RBF神经网络技术并根据D-S证据理论来实现多源异构数据的融合,构建实现数据共享和推理的多源异构数据融合知识库,以有效解决大数据时代数据异构、杂乱和共享的问题,同时为后续研究提供数据支持。

3.1.2 提高用户偏好计算的精确度

采用贝叶斯网络建模技术来分析用户在不同情景下对服务类别的偏好,综合考虑不同情景对用户偏好产生的影响以确定关键情景因素,基于信息熵理论来计算用户在关键情景下的偏好度,提高用户偏好计算模型的准确度。

3.1.3 建立用户好友信任度模型

采用结合情景和协同过滤的移动推荐算法,将影响用户偏好的关键情景与其社会网络关系相结合,通过关键情景下用户间的相似度和社会网络关系下的用户信任度计算用户综合情景相似度,依据搜索得到的当前用户最近邻居推断出用户对未购买商品的情景化偏好,以实现移动社交网络服务的精准推荐。

3.2 基于多源异构数据的移动社交网络个性化研究创新要点

3.2.1 基于D-S证据理论和优化的RBF神经网络分析法提出多源异构数据的融合策略模型,构建用于数据共享和推理的多源数据融合知识库。充分挖掘了各个方位的信息为用户提供更准确个性化的推荐方案,提高系统的可解释性和可扩展性,解决冷启动问题。

3.2.2 利用贝叶斯网络处理不确定信息和概率推理的方法来分析不同情景下的用户偏好,提出情景信息熵和情景重要度概念并通过情景信息熵算法来识别关键情景因素,提出一种更精确地偏好获取方法。

3.2.3 提出一种融合情景的协同过滤推荐算法和基于多源异构数据融合的个性化推荐框架,提高个性化推荐准确度。协同过滤推荐算法可以分为基于内存的协同过滤和基于模型的协同过滤两大类。基于用户的协同过滤通过计算用户之间的相似值,来寻找与目标用户相似度高的,再预测目标用户没购买过而相似用户购买过的物品的打分,最后基于物品打分分数的高低,选择分数高的进行推荐。基于物品的协同过滤通过计算用户购买过和未购买过物品之间的相似度,再预测用户对信物品的喜爱度,最后将预测结果分数高的物品推荐给用户。将不同的异构数据放在统一的框架中进行预测得出推荐结果,提高了推荐个性化的准确度。

四、结束语

互联网信息的快速发展,既带来了机遇,也带来了巨大的挑战。人们可以在互联网社交网络上获得自己感兴趣的信息内容,但由于网络上信息量大且复杂,导致信息的利用率达不到预想的要求。所以本文研究分析了基于多源异构数据的移动社交网络个性化推荐并进行改进,将其他多源异构数据融入推荐系统中来,更精准地提供用户的偏好需要和高质量的推荐内容。在提高推荐精准度的同时,也要提供推荐的效率,推动我国互联网信息社会可持续、高质量发展。

参考文献

- [1] 电力系统多源异构数据混合式存储方式研究[J]. 黄欣, 吴伟杰, 赵伊宁. 机电工程技术, 2021
- [2] 电网建设智慧前期平台多源异构数据融合模型[J]. 陈星, 孙博洋, 赵文渊. 电力学报. 2021
- [3] 多源异构数据时空融合关键技术研究与应用[J]. 李仕峰. 地理空间信息, 2021

作者简介:

许秀(1990年9月),女,汉,湖南常德人,讲师,研究方向为信息技术、数字经济。

项目基金:湖南省教育厅科学研究项目(基于多源异构数据的移动社交网络个性化推荐研究,编号:19C1368),湖南省普通高等学校教学改革研究项目(大数据背景下信息工程专业教育与双创教育融合模式研究,编号:HNJG-2020-1273),湖南应用技术学院课程思政综合研究类(理工类高校专业课融入课程思政价值的创新研究,编号:HYSZYB202022)。