

黄仁 孟婷婷 (重庆大学计算机学院)

黄仁 孟婷婷 (重庆大学计算机学院)

借鉴。

关键词:个性化推荐系统 推荐算法 社交网络

1 概述

随着互联网的发展,大量的信息系统、资源、门户网站的建设,使得人们在享受到便利的信息获取方式的同时,也对如何快速准确地获取自己需要的信息越来越关注,可以说谁能帮助用户快速地获取信息,谁就取得了成功的通行证。这一方面的典型例子就是百度搜索引擎的出现,百度公司从初创不足 10 人到 5 年内实现美国纳斯达克上市,就充分说明了这一点。

搜索引擎在解决信息过载问题方面发挥了非常重要的作用,并具有不可替代性,但是搜索引擎也存在一些固有的缺点,首先,它不具有个性化搜索功能,对于相关的关键词,每次的搜索结果都是相同的,其次,在商业购买行为上,搜索引擎并不能对用户有效的购买建议,最后,由于竞价排名的问题,搜索引擎查询出来的排名靠前的信息可能并不是自己需要的信息。

由于搜索引擎的这些缺点,个性化推荐系统应运而生。个性化推荐系统根据用户的历史浏览信息,对用户兴趣进行建模,将用户可能感兴趣而又未曾浏览过的产品、信息等推荐给用户。和搜索引擎相比,推荐系统一般不需要用户自主输入查询关键字,就能完成信息推荐,另一方面,对于不同的用户,推荐系统推荐的信息并不完全相同,总体上符合用户的个性化需求。

个性化推荐系统现在已广泛应用于很多领域,包括电子商务、电影和视频网站、个性化音乐电台、社交网络、个性化阅读等等。其中由于商业的获利需求,电子商务领域的个性化推荐发展迅速,目前国内的大多数电商平台都已具备基本的个性化推荐功能,比如京东商城、淘宝网等。同时业界对于如何提高个性化推荐的准确性研究热情一直很高。本文论述的重点就是对推荐系统的关键技术进行综述。

2 推荐系统概念和定义

目前对推荐系统的定义有很多,被广泛认可和采用的是 Resnick 和 Varian 在文献[1]中给出的推荐系统的概念性描述:“它是以电子商务网站为平台,为消费者提供商品的信息和建议,协助他们决定应该购买什么产品,模拟推销人员协助消费者完成购买过程”。由此可以推断出一个完整的推荐系统应该包括用户、物品和推荐方法三个要素。通用的推荐系统模型流程图如图 1 所示,推荐系统根据用户的隐含信息或显示信息,对用户进行建模,同时对物品信息进行建模,推荐系统根据不同的推荐算法对用户兴趣和物品特征进行匹配筛选,找到用户可能感兴趣的推荐物品,最后推荐给用户。

文献[2-3]对推荐系统的定义如下: 设 C 是所有用户的集合, S 是所有可以推荐给用户的对象的集合, 设效应

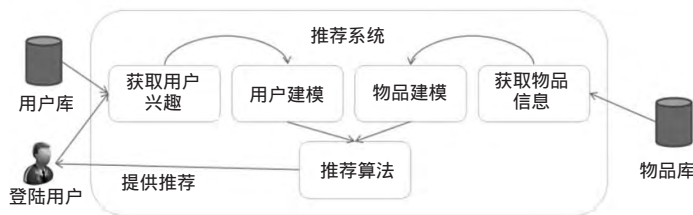


图 1 推荐系统模型

函数 $u()$ 可以计算对象 s 对用户 c 的推荐度和产品的可用性, 即 R 是一定范围内的全序非负实数, 推荐要研究的问题就是找到推荐度 R 最大的哪些对象 S^* , 如公式(1)

$$\forall c \in S^* = \operatorname{argmax}_{s \in S} u(c, s) \quad (1)$$

从上述概念和定义中可以看出,推荐系统关键是效用函数 u 的计算,即推荐算法的计算,下文将对流行的推荐算法进行描述。

3 推荐算法综述

推荐算法是整个推荐系统中最关键和核心的部分,在很大程度上决定了推荐系统的性能和推荐效果,目前很多学者对推荐系统的研究重点都在如何提高推荐效率和推荐效果方面,不同的推荐算法推荐效率不同,应用环境也不同,目前主流的推荐算法大致可以分为以下几种:协同过滤推荐和基于内容的推荐。下一节将讨论这几类算法,并对一些新的推荐算法给出描述。

3.1 协同过滤算法

协同过滤算法是研究最早并且应用最广的推荐算法之一,该算法的研究极大地推动了推荐系统的发展,在推荐算法体系中占有重要的研究地位,迄今为止仍有大量的论文对算法的性能进行研究。

该算法的基本思想借鉴了人们日常社交生活中购买商品的方法。如果自己的亲戚朋友购买了某件商品并且对该商品评价很好时,那么自己在很大程度上也会尝试购买该商品,协同过滤正是把该思想应用到了推荐系统中。该算法根据相似度比较对象的不同可以分为基于用户的推荐和基于物品的推荐。

3.1.1 基于用户的最近邻推荐

基于用户的协同推荐的基本思想是:首先,给定一个评分数据集和一个用户 c ,在评分数据集中找出与当前用户过去有相同偏好的用户 c_i ,即最近邻计算,然后,对当前用户没有购买过的每个物品 o ,利用其近邻对 o 的评分计算预测值。

此算法包含两个方面,一是寻找当前用户的最近邻(即计算用户之间的相似性),二是计算当前用户对物品 o 的评分值。其中寻找最近邻是该算法的核心。推荐系统中一般采用基于关联和基于余弦两种算法计算。近年来,为了提高该算法的性能,一些学者同时也提出了其他方法,如缺省投票、用户倒排评分、实例扩展和主流加权预测^[4]等。

3.1.2 基于物品的最近邻推荐

基于用户的协同过滤方法对于用户较少的情况执行效率和预测值较好,但对于一些有着数以百万计的用户和物品的大型购物网站来说,执行效率并不理想,基于物品的最近邻推荐,由于其离线计算的优越性,已经逐渐取代基于用户的计算方法。

该算法的核心是计算物品两两之间的相似度,由于不涉及用户,因此可以在数据库中提前计算好,当需要对用户 c 推荐物品 o 时,通过确定与 o 最相似的物品 oi ,计算用户 c 对这些近邻物品 oi 评分的加权总和来得到用户 c 对物品 o 的预测评分。其中近邻数量受限于用户评过分的物品个数,由于这样的物品数量一般都比较少,因此计算预测值可以在线上段时间内完成^[9]。

3.1.3 其他协同推荐方法

基于用户的协同过滤算法因为评分数据保存在内存中,运行时直接生成推荐结果,属于基于内存的推荐方法,基于物品的协同过滤由于离线计算原始数据,运行时使用预计算或学习过的模型预测,属于基于模型的推荐方法,近年来使用数据统计或机器学习方法设计用户模型成为协同过滤的一个研究重点,常用的方法有神经网络方法、矩阵分解方法、概率分析方法、线性回归方法和最大熵方法^[10]等。

3.2 基于内容的推荐

基于内容的推荐与协同过滤推荐不同,不需要用户对物品的评分数据,且不需要比较多个用户或多个物品之间的相似度,该算法的基本思想是根据用户的历史兴趣数据,建立用户模型,然后针对推荐物品的特征描述进行特征提取,最后将物品特征与用户模型相比较,相似度较高的物品就可以得到推荐。其中物品特征提取目前的研究主要集中于文档特征提取,特征提取算法一般采用词频-倒排文档频率法,即 TF-IDF 算法,而建立用户模型的算法一般使用机器学习领域的算法,比如决策树分类算法、贝叶斯分类算法、神经网络、基于概率模型的方法和线性分类器等。对象特征和用户模型相似度比较的算法,最简单的可以采用向量夹角余弦法计算^[7]。

基于内容的推荐目前主要应用于文本的推荐,比如网页、新闻、文档等,但目前网页等超文本文档中越来越多的包含有多媒体信息,比如图片、音频和视频等,而多媒体内容的特征提取的研究仍处于研究阶段。目前基于标签的推荐是一个新的研究方向,该方法使用人工对多媒体内容进行标注的方法对多媒体进行分类,通过对标签的分类达到超文本推荐的目的。

3.3 基于社交网络的推荐

社交网络是现实社会中人与人之间关系的延续与扩展,在网络上人们不但可以把同学、朋友加为好友,还可以关注不认识,但兴趣爱好相同的人,并加为好友,社交网络极大地扩展了人与人之间的交际圈,将不同行业、不同职业、不同年龄的人们联系起来,因此基于社交网络的营销具有巨大的商业价值和商业潜力,目前微商的流行就是一个证明。

基于社交网络的推荐是协同过滤推荐在社交网络中的延伸,同时又具有基于网络结构推荐的特点,因此大体上可以分为两类,即基于邻域的社交网络推荐和基于网络结构的社交网络推荐。

3.3.1 基于邻域的社交网络推荐

基于邻域的社交网络推荐,其基本思想是查询社交网络中当前用户所有的好友,根据好友的兴趣数据,向当前用户推荐好友喜欢的物品集合。不过,即使是用户 u 的好友,不同的好友和用户 u 的熟悉程度和兴趣相似度也是不同的^[9],因此算法中还要考虑好友和用户的熟悉程度和兴趣相似度。用户 u 对物品的兴趣 pui 可以使用以下公式表示^[9]:

$$pui = \sum_{v \in f(u)} (f_{uv} + s_{uv}) r_{vi} \quad (2)$$

其中 $f(u)$ 是用户 u 的好友集合, f_{uv} 表示用户 u 和 v 之间的熟悉程度, s_{uv} 表示用户 u 和 v 之间兴趣爱好的相似度,这二者的计算方法可以参考协同过滤算法中近邻用户相似度的计算方法。 r_{vi} 表示用户 v 是否喜欢物品 i ,如果用户 v 喜欢物品 i ,则 $r_{vi}=1$,否则 $r_{vi}=0$ 。

3.3.2 基于网络结构的社交网络推荐

该算法把用户和用户的好友,以及用户和好友喜欢的物品抽象为节点,如果用户之间存在好友关系,就会在用户之间形成连线,同时如果用户喜欢某一物品,也会在用户和物品之间形成连线。通过获取社交网络数据,就可以绘制出用户和物品之间的网络结构图,如图 2 所示,该图表示用户之间的好友关系和用户喜欢物品的列表,图 3,将二者之间的关系表示为网络结构图的形式,图中用户使用圆圈表示,物品使用方块表示。



图2 用户之间和与物品之间的关系

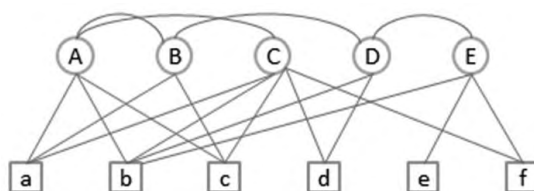


图3 用户之间和与物品之间的网络关系

此外,根据用户之间的熟悉程度和兴趣相似度,可以定义用户之间边的权重,根据用户对物品的喜欢程度,可以定义用户和物品之间边的权重。

在定义好节点、边和边的权重后,对用户 u 的推荐问题,就转化为计算用户节点和所有物品节点之间的相关

基于蓝牙技术的智能传感器的研究

万义星 (江西水利职业学院)

摘要:网络技术的不断发展,尤其是无线移动通信技术的发展,为我国工业和生活自动化的发展带来了革命性技术。将蓝牙技术应用于传感器领域,不仅能够减少现场的布线作业,避免现场施工环境对有线连接的影响,而且还可以为人们带来巨大的生活、工作便利。

关键词:蓝牙技术 传感器 构成

传感技术是现代科技的产物,为人们的生产与生活带来巨大的便利,随着网络技术的发展,传感技术在经历了独立感知阶段之后,到目前为止已经进入智能型传感器发展阶段。智能型传感器可依靠无线通信方式进行信号的传输。因此本文就基于蓝牙技术的智能传感器的功能实现技术进行详细的剖析。

1 蓝牙技术的概述

蓝牙技术是一种比较适用于短距离无线数据与语音通信的开放性全球规范技术,蓝牙技术的实现是通过近距离的无线连接,为移动通信设备提供通信链路,进而实现近距离信息设备资源的共享。

蓝牙技术具有以下特征:

1.1 具有使用全球通用的频段 蓝牙技术工作使用的全球通用频段是 2.4GHz 频段,该频段属于工业科学医疗频段。工业科学医疗频段是向所有无线电系统开放的频段,因此蓝牙技术基于其无线电的功能能够使用该频段。为了避免蓝牙技术与该频段其他设备之间发生相互干扰的情况,影响蓝牙系统的工作,蓝牙技术设计了相应的调频方案,以此保证链路的稳定,实现蓝牙系统运行的稳定性。

1.2 成本低、消耗低以及辐射低 轻、薄、小是蓝牙技术发展的主要目标,蓝牙技术是将多种芯片等系统融为一

体,以 USB 或者 RS-232 接口与现有的设备相互连接,或者直接将蓝牙模块内嵌在信息设备中,处于实用性考虑,蓝牙技术的制造成本会越来越低、消耗也会越来越低,便于蓝牙设备的长期使用。

1.3 安全性 由于蓝牙系统具有开发性和移动性,因此其容易受到一些系统的干扰,导致蓝牙信号被截取,影响用户的安全,为此蓝牙系统提供了认证与加密措施,通过加入认证措施能够提高蓝牙系统的安全性,保证链接的安全稳定。一般蓝牙系统的加密是通过硬件实现的,而密钥则是由软件系统实现,从而有效地提高了蓝牙技术的安全性,同时蓝牙技术属于短距离传输,因此它的安全性相比有限传输要安全许多。

1.4 网络特性 蓝牙技术属于点对多点的通信协议,由一个主网与多个对应从节点构成,蓝牙组网时最多可以容纳 8 个蓝牙单元设备形成微网,实现蓝牙系统的广泛应用性,蓝牙技术实现了多用途,能够应用到不同的电子设备中,比如蓝牙遥感开门系统等。

2 蓝牙技术智能传感器系统的实现

蓝牙技术智能传感器系统的实现主要包括:硬件、软件以及网络三部分。

2.1 蓝牙智能传感器无线通信的硬件实现

2.1.1 蓝牙智能传感器硬件的构成。 蓝牙智能传感器无线通信的硬件主要包括传感器、单片机以及蓝牙模块。一是传感器。传感器的最大优势就是能够大大提高系统的抗干扰性,能够准确地感应出蓝牙智能系统所传输的各种信号信息;二是单片机是传感器设备的核心,是实现对传

(上接第 272 页)

性,然后取与用户没有直接边相连的物品,按照相关性的高低生成推荐列表,即网络结构图的排名问题。图的排名问题可以通过 PersonalRank^[9]图排序算法得到结果。

4 总结

推荐系统的研究,既有重大的社会意义,又有重大的经济意义,一方面它可以帮助互联网用户在知识过载的网络环境中查找自己感兴趣的知识,另一方面还可以帮助电子商务网站更好地给用户推荐商品,提高电子商务网站的销售额,而推荐算法的研究一直是推荐系统研究的重点。随着对推荐系统实时性和准确性要求的不断提高,算法实现的技术难度也越来越高,需要不断地对算法进行完善,以适应时代的发展。

参考文献:

- [1]ResnickP,Varian H R.Recommender systems [J].Communications of the ACM,1997,40(3):56-58.
- [2]AdomaviciusG,TuzhilinaA.Toward the next generation of recommender systems:A survey of the state-of-the-art and possible extensions [J].IEEE Trans on Knowledge and Data Engineering,2005,17(6):734-749.
- [3]Sung H H.Helping online customers decide through Web

personalization[J].IEEE Intelligent Systems 2002 9:34-43.

[4]赵良辉,熊作贞.电子商务推荐系统综述及发展研究[J].技术应用 2013.12.

[5]DietmarJannach Markus Zanker.Recommender Systems,北京:人民邮电出版社 2014 8-14.

[6]王国霞,刘贺平.个性化推荐系统综述[J].计算机工程与应用,2012 48(7) 66-76.

[7]许海玲,吴潇,李晓东,阎保平.互联网推荐系统比较研究[J].软件学报 2009(02).

[8]项亮.推荐系统实践[M].北京:人民邮电出版社 2012:151-153.

[9]庞秀丽,冯玉强,姜维.电子商务个性化文档推荐技术研究[J].中国管理科学 2008,16(专辑):581-586.

[10]Zhou T,Jiang LL,Su RQ,et al.Effect of initial configuration on network-based recommendation.Europhys Letters,2008,81.

[11]刘建国,周涛,汪秉宏.个性化推荐系统的研究进展[J].自然科学进展 2009(01).

[12]曹毅.基于内容和协同过滤的混合模式推荐技术研究[D].长沙:中南大学 2007.

作者简介:黄仁(1962-),男,四川成都人,副教授,主要研究方向:模式识别与知识工程、计算机控制及软件;孟婷婷(1986-),女,山西运城人,硕士研究生,主要研究方向:计算机控制及软件开发。