**班 级 1718012**

**学 号 17180120056**

****

本科毕业设计论文



**题 目** 基于差分隐私保护技术的

协同过滤推荐系统设计与实现

**学 院** 网络与信息安全学院

**专 业**  信息安全

**学 生 姓 名**  王屹晨

**导 师 姓 名** 李兴华

**院内导师姓名** 李兴华

摘要

关键词:

Abstract

key words :

目录

第一章 绪论

1.1 研究背景及意义

当今社会，互联网与移动互联网的普及与发展十分迅速，各类网络应用与移动App已经融入到了人们日常工作和生活的各个方面，比如即时通讯、社交网络、电子商务与电子支付等，人们的日常工作与生活已经离不开移动互联网。在互联网化趋势下，各个行业数据量呈爆炸式增长，人们逐渐从信息匮乏的时代走入了信息过载的时代。在2011年，IDC的研究报告中显示，全球信息总量每过两年，就会增长一倍。而调查显示，这一增速仍在持续加快。10年前，该机构预测2020年全世界的数据量将达40ZB。而中国信息通信研究院2020年12月发布的大数据白皮书（2020年）（http://www.caict.ac.cn/kxyj/qwfb/bps/202012/t20201229\_367255.htm）中指出，根据国际权威机构Statista的统计和预测，2020年全球数据产生量预计达到47ZB，比十年前的预测有一个重量级的增长。2018年11月IDC白皮书《DataAge2025》的预测表示，这一数字将在2025年来到175ZB。这使得人们在获取丰富多彩的信息内容的同时，沉浸于信息海洋而难以及时、准确地获得满足其自身需要的信息。

信息过载的主要解决方法有两种：一是搜索，二是推荐。搜索是一种良好的解决方案，但搜索引擎解决的只是人们主动搜索数据的请求，而且需要人们对自己的需求有明确的认识，但是人们通常没有办法十分准确地描述出自己的需求和喜好，因此搜索没有办法满足人们日益增长的对个性化信息的需求。在这个背景下，推荐应运而生。凭借可以主动提供符合用户需求的信息这一优点，应用推荐算法的推荐系统已经被大量运用到各行各业中，成为了互联网中不可或缺的一个环节。我们平时在各个购物网站、新闻网站和影音网站都可以见到推荐系统的身影，比如淘宝、今日头条和Netflix，这些公司的APP都会根据用户的兴趣来进行推送。

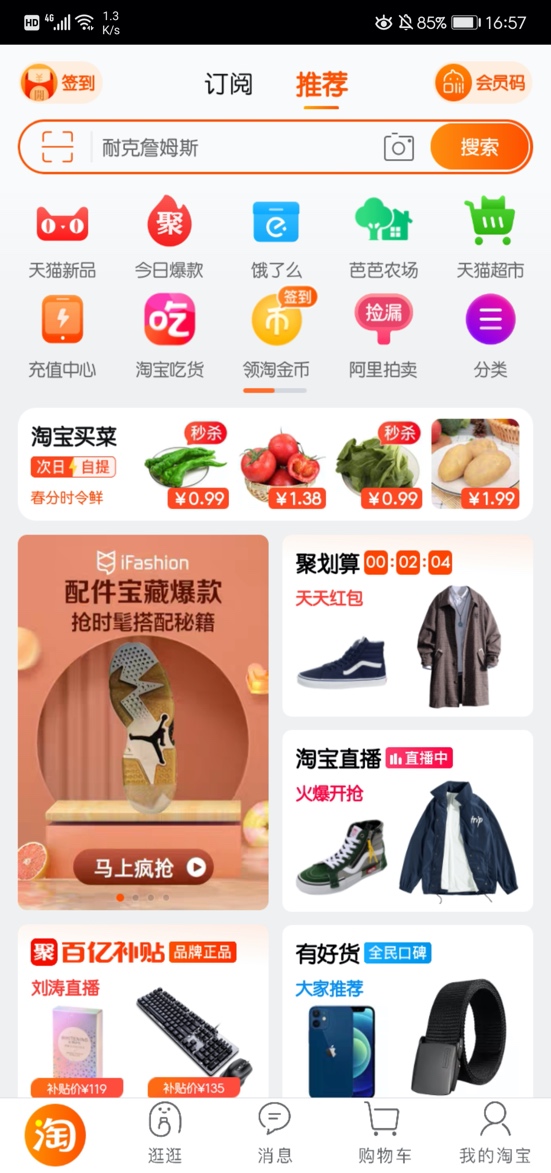


图1 手机淘宝的推荐页

协同过滤是迄今为止较为成功、使用最广泛的推荐算法之一。亚马逊和Netflix都是应用协同过滤而大获成功的网站。从思路上来说，推荐系统在为用户建立定制化服务时，首先获取并分析用户的各类数据，然后学习用户的偏好和兴趣从而为用户推荐他所需要的信息和服务。近年来，帮助用户快速找到自己所需信息的推荐技术成为计算机领域研究的热点。然而，因为推荐系统进行推荐，首先需要收集大量的用户行为和使用习惯，例如用户的评分记录、浏览记录和购买信息等。而且用户行为数据越丰富，得到的推荐结果也就相对越准确。也正是因此，这些信息中存在着泄露用户个人隐私的风险。2006年，AOL发布了web搜索日志数据集，其中包含了650000用户三个月内的搜索记录。AOL将用户的user id隐去，以随机数字代替。然而No.4417749，一位妇人Arnold仍被找了出来。搜索日志中包含她的姓，附近地址的名字，生活信息等，攻击者将这些信息与phonebook结合找出了这位No.4417749的真实身份。2009年，Netflix发布了电影评分数据集，用于一项数据挖掘竞赛。该竞赛希望能够通过用户的历史评分来预测用户对电影的偏好。数据包括用户对一系列电影的评分记录，经匿名化处理抹去了用户id。然而若攻击者有某个用户的部分评分记录，就能够很快判断出这个用户是否出现在了数据集中。攻击者通过IMDB用户的若干评分记录来定位它在数据集中的位置，从而该用户的整条记录都被发掘。

随着使用互联网应用的用户人数不断增加和互联网应用使用频率的不断提高，隐私保护成为了一个非常重要的研究方向。尽管推荐系统有其便捷性和高效性，但是对用户隐私的威胁成为了该系统在实际应用中的一个很大的障碍。如果用户想要获得更加有效的推荐，则需要提供他们自己的评分数据或者浏览记录，目前还没有一个具体的政策来保护用户隐私问题。前面说到一些公司会把某些数据集公开，这是一种侵害用户隐私的行为。更恶劣的，一些恶意的推荐系统在收集到用户数据之后，会将其出售以获得高额利润。有时候，即使不是一个恶意的推荐系统，数据的意外泄露或者第三方应用的恶意插件也会入侵到数据库中盗取用户的隐私数据。笔者在写此篇论文时，个人数据似乎也遭到了泄露：笔者用电脑在百度搜索了一个旅游景点，不久后手机端的一个视频软件就推送了该旅游景点的攻略。这样的隐私威胁使用户不愿意再交出个人的数据或制造虚假数据，从而降低了数据集的量级和准确度。若想鼓励用户更广泛地参与推荐系统，需要在推荐系统中考虑隐私保护使得用户隐私数据不被泄露。

在隐去敏感数据、传统加密技术和改变数据内容都无法起到显著效果时，Cynthia Dwork在2006年提出的差分隐私机制解决了这些传统方法的缺陷。它定义了一个极为严格的攻击模型，通过对数据集中的原始信息或者统计数据添加噪声来实现对隐私的保护。因此，即便攻击者拥有除去目标隐私信息外的所有背景知识，数据依然可以得到有效的保护。差分隐私的优点使得它受到国内外学者的广泛关注和研究。但是目前国内外所研究的差分隐私方向大都还只是基础理论或者数据发布领域，在数据挖掘方向的应用研究还不是很多，特别是涉及推荐系统的隐私保护研究领域，目前国内外相关的研究还很少。由于差分隐私保护在实际的应用过程中，大多是通过对数据集或者算法的输出结果中添加噪声来实现，如果不恰当的使用，则会带来不必要的开销，造成数据可用性降低和推荐精确度降低的情况。因此差分隐私在实际应用中，通常会和实际的算法本身相结合，才能在提供差分隐私保护的同时，不影响数据本身的分析与使用。

因此，研究差分隐私保护技术与协同过滤推荐算法的有机结合，均衡隐私保护和推荐精度，在能提供隐私保护的情况下，降低推荐算法的准确度的损失，在理论与应用层面都具有十分重要的意义。

1.2 国内外研究现状

推荐系统已经经历了几十年的风雨，能有今天的成就，要归功于研究者们的辛勤耕耘。

自从协同过滤于 1992 年被正式提出以后，通过不断地被赋予更多的含义和技术创新，已经成为最流行的推荐算法之一，但一些典型问题如数据稀疏问题等阻碍着其更进一步地发展。当数据非常稀疏时，用户实际的相似度很难用如今的相似度计算方法来衡量，有时会将不相似的用户选为目标用户的邻居用户，进而影响推荐效果。Lyle等人提出了基于聚类的协同过滤方法，以用户为例，首先将用户分组到某个集群中，然后将目标用户也分配至最相似的那个集群中，最后在该集群中为目标用户寻找邻居进行推荐。对于维度较高的大规模数据集，使用这类方法之前还需要进行降维或采样。之后，有些研究者通过采用不同的聚类策略来达到更优的推荐效果，但是在可扩展性上表现得不尽如人意。

自 NetFlix大赛以来，矩阵分解(MF)成为协同过滤算法中最受欢迎的方法之一。为了 缓解评分的稀疏性，有时需要将额外推荐信息整合到MF中，Pagare等人引入用户社交关系到MF中，并重新设计损失函数；Jamali等人根据用户之间的历史交互记录来获取信任度，并与 MF相结合，提升了推荐效果；此外，Ma等人通过研究社交网络中的上下文信息来挖掘影响用户兴趣的更深层次的原因，并取得了不错的推荐效果。

直至 2000年左右，推荐系统才引起国内的关注，并逐渐成为计算机领城研究的热点。 1999年，清华大学路海明等人提出基于多代理技术的混合智能个性化推荐服务。2000年，北京大学余锦凤等人提出了个性化定制服务。2001 年，南京大学研发了个性化信息检索智能系统DOLTRL-Agemt。2003年邓艾琳等人的《基于项目评分预测的协同过滤推荐算法》；2004年余力等人的《电子商务个性化推荐研究》，2007年彭玉等人的 《基于属性相似的Iter-based协同过速算法》；2009年许海玲等人的《互联网推荐系统比较研究)》；2013年孟样武等人的《移动推荐系统及其应用》；2015年张玉洁等人的《组推荐系统及其应用》。这些优秀的系统或者论文标志着个性化推荐技术的理论研究在学术界逐渐丰富起来。

近几年，随着互联网电商的快速发展，以及Amazon个性化推荐系统的应用成功，国内电子商务网站也纷纷构建其推荐系统。2006年，当当网开始提供个性化推荐服务，用以向客户推荐书籍。2008年，淘宝网推出了个性化推荐系统，用于帮助用户从大量的商品中找到符合自己偏好的商品。2011年，百度推出了个性化推荐首页，根据用户的行为向其推荐符合需求的信息。2014年，阿里巴巴开始举办“天猫”推荐算法大赛，吸引了国内外众多研究者的参加，促进了个性化推荐系统的发展。

差分隐私(Diferential Privacy)的概念最早由微软研究院的研究员Cynthia Dwork在 2006年至 2008年陆续提出,他将隐私保护的数据分析建立在严格的数学基础之上，使得数据即使在高精度的数据分析下仍能保证隐私性。差分隐私保护基于数据失真技术，通过向原始数据添加一定量的随机噪声来实现隐私保护的目的。

2016 年，苹果首次宣布在iOS 10中采用差分隐私技术，这种技术允许苹果收集用户信息，同时让用户保持数据的完全私有，并不会对用户隐私造成损害。差分隐私已经Mac和iOS设备上使用，包括emoji的使用、搜索预测、文本预测和其他使用机器学习的小功能等等。苹果的差分隐私实现包含了每笔捐款隐私预算的概念（由参数𝜀量化），并对用户的贡献数量设定了严格的限制，以保护用户的隐私。原因是差分隐私中使用的略带失真的噪音往往在大量的贡献中被平均化，使得理论上有可能在来自单个用户的大量观察中确定关于用户活动的信息。

应用差分隐私保护的一些主流研究领域有推荐系统领域、位置隐私领域和数据发布领域。

MeShery和Mironor是最早对差分隐私在推荐系统特别是在协同过滤中的应用进行研究的学者。他们利用拉普拉斯机制对输人评分生成的带噪声进行计数与求和，并计算产品协方差矩阵具有差分隐私特性的版本。接下来，带噪声的协方差矩阵可以用来生成具有差分隐私的k近邻和SVD推荐。

Zhu等人则采用了一种不同的方法来实现具有差分隐私的基于邻居的协同过滤推荐算法。该算法主要针对的目标是Calandrino等人提出的女巫攻击(sybilatack)。 作者提出了一种具备差分隐私的k近邻算法，其分为两步执行：邻居选择以及基于邻居的评分预测。该算法依赖于相似度函数的平滑敏感性，相较于拉普拉斯机制所需的噪声而言，允许引入程度更低的噪声。他们将随机化加入k近邻的选择过程中，同时保证选中的邻居拥有相似分数的可能性非常大。

Machanavajhala等人基于链接用户和产品的图结构分析了隐私保护的社交推荐。如果合定一个这样的图，那么他们可以得出反映产品与用户效用的效用向量，以归纳出一种可以最大化用户效用并保证效用向量隐私的概率分布。作者对该问题提供了理论分析，并得出结论，认为好的推荐仅在较弱的隐私参数下才能够得到，或者仅有少数用户可以获得较好的推荐。他们强调，隐私与准确率之间的平衡依然存在于基于差分隐私的方法之中。

对于差分隐私推荐系统的研究表明，在某些情况下(如社交推荐)，可能无法同时保证隐私和准确率，而在另一些情况下，隐私保护推荐系统可以达到合理的准确率。然而，到目前为止的研究工作都假设计算是一次性的，因而当更多数据变得可用时则需要对推荐进行重新计算，这又会引人额外的隐私风险。因此，对多重计算或者数据发布提供隐私保护则需要提高引人噪声的程度，当然，这会降低推荐的准确率。笔者利用拉普拉斯机制对推荐算法发布的物品因子矩阵加入噪声，实现推荐算法的差分隐私保护。

1.3 论文的主要研究内容

从上面的介绍可以看出，推荐系统已经在日常生活中得到了十分广泛的应用，是解决信息过载问题的一个重要手段。我们总是希望得到更精准的推荐，从而帮助我们做出更好的决策。但是潜在的攻击者会希望得到推荐系统中记录的敏感信息从而获利。所以，在保证推荐系统推荐结果有效性的同时也要保护用户的隐私数据，防止攻击者对这些数据进行非法操作。

本文的主要研究内容如下：

（1）分析了在推荐系统中加入差分隐私的重要性和必要性，介绍了推荐系统隐私保护的研究背景和目前国内外推荐系统、差分隐私技术以及二者结合产物的研究现状。

（2）推荐系统概述。介绍推荐系统的主要分类方法和对于协同过滤推荐算法的研究；介绍了协同过滤算法的主要步骤：收集用户偏好、找到相似的用户或者物品、计算并推荐。

（3）差分隐私概述。分析了差分隐私的概念和该模型相对于传统安全模型的优势，研究了差分隐私的性质和常用的实现机制。

（4）基于差分隐私的协同过滤推荐系统的设计、实现与测试。基于前面的理论知识，设计了一个使用基于用户的协同过滤的推荐系统,其中相似度采用两种方法计算，并使用差分隐私将推荐结果进行加密。然后将设计出来的系统应用于MovieLens的两个不同规模的数据集上进行实验比较和分析，通过测试，找到可以相对较好地平衡推荐结果准确度和隐私保护程度的隐私预算。

1.4 论文结构

本篇论文共包含六个章节，具体各个章节的内容介绍如下：

第一章为绪论。首先介绍了本文的研究背景，其次分析了推荐系统和差分隐私的研究现状及存在的问题，最后总结了本文主要的研究内容。

第二章介绍了相关的背景知识。首先对推荐系统进行概述，引出协同过滤算法，并根据协同过滤推荐系统的三个主要步骤对其进行详细介绍。然后介绍本文的另一主题——差分隐私。主要介绍其定义、概念和实现机制。

第三章介绍了基于差分隐私保护的协同过滤推荐系统的具体实现。在本部分，首先介绍了攻击者基于协同过滤的推荐系统挖掘用户隐私数据的攻击模型，总结了要应对该种手段，差分隐私对于传统数据加密方式的优越性。然后，提出了本文涉及系统的代码实现过程。

第四章主要分析了在不同相似度计算方法和不同隐私预算的情况下，对于推荐精度和隐私保护程度的平衡。找到了最佳的相似度计算方法和隐私预算组合。

第五章主要总结了本文的主要工作和成果，并对后续的研究工作方向进行展望。

第二章 基础知识

2.1 推荐系统

2.2.1 推荐系统概述

推荐系统(RS)是一种向目标用户建议可能感兴趣物品的软件工具和技术。推荐系统的建议可用于多种决策过程，如购买什么物品、听什么音乐以及在网上浏览什么新闻等。推荐系统中的“物品”指系统向用户所推荐内容的总称。某个推荐系统通常专注于一个特定类型的物品(如CD或新闻)，因此它的设计、图形用户界面以及用于生成推荐结果的核心技术都是为特定类型的物品提供有用且有效的建议而定制的。

推荐系统主要服务于缺乏经验和能力的用户，他们通常无法从大量可供选择的物品中选取感兴趣的物品，如无法从某网站中选取感兴趣的商品。推荐系统中一个典型的例子是图书推荐，系统可以帮助用户挑选一木可能感兴趣的书来读。知名网站亚马逊中，系统采用个性化推荐技术为每个客户进行推荐。通常所说的推存具有个性化，即不同的用户或用户组所接收的建议是不同的。当然也存在非个性化推荐，但它们大都非常简单，通常出现在报纸或杂志上。最简单的个性化推荐是提供一个排序好的物品列表。通过该排序列表，系统试图根据用户的偏好和某些约束条件来预测最合适的产品或服务。为了完成上述计算任务，推荐系统通常需要收集用户的喜好信息，这种喜好信息或是显式的，如物品的评分信息，或通过解释用户的行为做出推断所得。

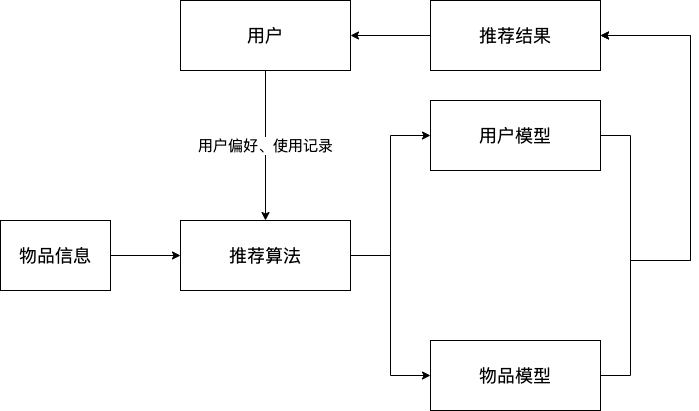


图2.1 推荐系统模型

推荐系统最初起源于一个相当简单的现象：人们在做日常工作和决策时总是依赖于他人提供的建议。例如，要选择读本书时，通常依靠朋友的推荐；雇主依靠推荐信做招募的决定；当选择观看影片时，人们倾向于阅读并且依赖影评家在报纸上所写的影评。

为了模拟上述行为，推荐系统通过算法将社区用户的建议推荐给一个活跃用户。系统向用户所推存的物品是相似用户喜欢的。这种方法称为协同过滤，它的理论依据是：如果活跃用户的历史评分信息与某些用户有相似之处，那么由这些相似用户所得出的推荐结果应与该活跃用户相关，且这些推荐结果也是该活跃用户所感兴趣的。

（还有推荐系统的种类没有加进去）

2.2.2 协同过滤推荐算法

协同过滤(colaboratile fitering)。这种方法是找到与用户有相同品味的用户，然后将相似用户过去喜欢的物品推荐给该用户，两用户之间的品位相似性是通过计算用户历史评分记录的相似性所得。协同过滤被认为是推荐系统中最流行和最广泛实现的技术。

设计这种模型的主要挑战是底层评分矩阵是稀疏的。例如在某个用户可以具体给出评分表达自己喜爱程度的APP中，大多数用户可能只看了浩如烟海的众多影片中的一小部分，因此评分大多是未知的。

由于已知评分常常是与用户和物品密切相关的，因此协同过滤法的基本思想是由已知评分估计未知评分。例如，有两个用户分别叫Alice和Bob，他们具有相似品味。当两人都给出具体评分时，给出的评分应当是十分相似的，这种相似度可以被底层算法检测出来。在这种情况下，对某个物品，两人中如果仅有一人做出了评分，另一个人的评分可能与该评分十分接近。多数协同过滤模型着重于借助物品之间或是用户之间内在的关联性做出预测，还有些模型两者都考虑了。更进一步，部分模型采用经过仔细设计的优化算法来创建训练模型。然后这个模型被用来估计矩阵中缺失的值。有两种方法在协同过滤算法中经常用到：基于记忆的方法以及基于模型的方法。

1)基于记忆的方法：基于记忆的方法也被称为基于近邻的协同过滤算法。这是最早的协同过滤算法之一，其中用户—物品组合的评分是在“近邻”的基础上进行预测。这些“近邻”可以用以下两种方式之一来定义：

1.基于用户的协同过滤。其基本思想为：给用户推荐和他兴趣相似的用户感兴趣的物品。当需要为一个用户A进行推荐时，首先，找到和A兴趣相似的用户集合，用U表示，然后，把集合U中用户感兴趣而A没有听说过，即未进行过操作的物品推荐给A。算法分为两个步骤：首先计算用户之间的相似度，选取最相似的N个用户，然后根据相似度计算用户评分。

用户相似度的计算是基于用户的协同过滤算法的重要内容，主要可以通过余弦相似度、Pearson系数等方式进行计算。

假设：u和v是两名用户，就是u看过的电影的集合，是v对看过的电影的集合，是用户u对电影k的评分，同理，𝜇是对于某一用户所有打过分的电影所求出的平均分，则用户u和v之间的相似度可以通过如下方式计算：

余弦相似度：

Pearson系数：

得到用户相似度后，可以根据如下公式计算用户评分：

其中代表用户u对物品i的评分，为与用户u最相似的N个用户，为对物品i进行过操作的用户集合，为用户u与用户v的相似度，为用户v对物品i的评分。

基于用户的协同过滤的推荐结果反映了用户所在的一个兴趣群体中的热门物品，更加社会化但缺乏个性化，能够满足物品的时效性，在新闻推荐领域能够发挥很大的作用。用户的兴趣在一段时间内是相对固定的，因此用户相似度矩阵不会实时进行更新，存在新用户的冷启动问题。

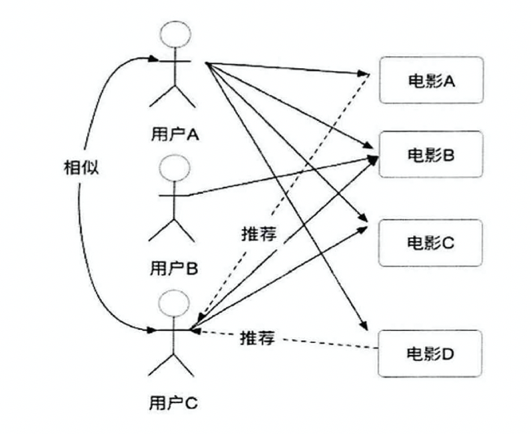


图2.2 基于用户的协同过滤推荐系统模型

2.基于物品的协同过滤。该算法的基本思想为：给用户推荐和他们以前喜欢的物品相似的物品，这里所说的相似并非从物品的内容角度出发，而是基于一种假设：喜欢物品A的用户大多也喜欢物品B代表着物品A和物品B相似。基于物品的协同过滤算法能够为推荐结果做出合理的解释，比如电子商务网站中的“购买该物品的用户还购买了…”。基于物品的协同过滤的计算步骤和基于用户的协同过滤大致相同：首先，计算物品相似度，选出最相似的N个物品，然后根据相似度计算用户评分。

基于物品的协同过滤的推荐结果更加个性化，反映了用户的个人兴趣，对挖掘长尾物品有很大帮助，被广泛应用于电子商务系统。在物品数较多时，物品相似度计算效率较差，因此通常以一定的时间间隔离线进行计算，然后将物品相似度数据缓存在内存中，这样一来，便可以根据用户的新行为实时向用户做出推荐。但基于物品的协同过滤同样存在新用户冷启动问题。

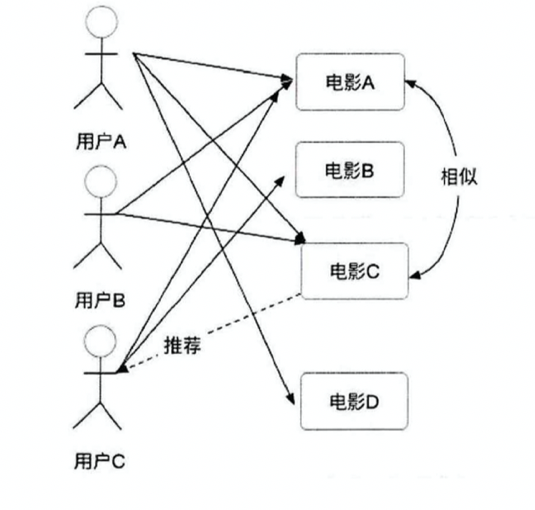


图2.3 基于物品的协同过滤推荐系统模型

基于记忆的方法的优点在于容易实现，并且其生成的推荐易于解释。然而，基于记忆的方法并不适用于稀疏的评分矩阵。例如，它很难找到与Bob足够相似并且评价过《Gladiator》的用户，这种情况下很难准确地预测Bob对《Gladiator》的评分。换句话说，这种方法缺少对评分预测的全面覆盖。但如果只需要预测出top-k个最相似的物品，覆盖面不全也没关系。

2)基于模型的方法：在基于模型的方法中会用到机器学习和数据挖掘技术。这是因为模型的参数需要通过一一个优化框架学习得到。基于模型的方法包括决策树、基于规则的模型、贝叶斯方法和潜在因子模型。包括潜在因子模型在内的许多方法，即使对稀疏的评分矩阵也能有较高的覆盖率。

基于记忆的协同过滤算法很简洁，但却是启发式的，并不适用于所有环境。基于模型的方法与基于记忆的方法之间的区别有点人为因素，因为基于记忆的方法实际上可以被认为是基于相似性的方法。由于Netflix大奖赛的影响，潜在因子模型近几年得到了推广，实际上，在不完整数据集上与其相似的算法很早就被提出了。最近的研究表明，一些基于记忆和基于模型的方法的结合体能提供非常准确的结果。

要实现协同过滤，需要进行如下几个步骤：

1.收集用户偏好

2.找到相似的用户或者物品

3.计算并推荐

收集用户偏好是指从用户的行为和偏好中发现规律，并基于此进行推荐，所以如何收集用户的偏好信息成为系统推荐效果最基础的决定因素。用户有很多种方式向系统提供自己的偏好信息，比如：评分，投票，转发，保存书签，购买，点击流，页面停留时间等等。以上的用户行为都是通用的，在实际推荐引擎设计中可以自己多添加一些特定的用户行为，并用它们表示用户对物品的喜好程度。通常情况下，在一个推荐系统中，用户行为都会多于一种，那么如何组合这些不同的用户行为呢？基本上有如下两种方式：

1.将不同的行为分组。一般可以分为查看和购买，然后基于不同的用户行为，计算不同用户或者物品的相似度。类似于亚马逊给出的“购买了该书的人还购买了”，“查看了该书的人还查看了”等等。

2.不同行为产生的用户喜好对它们进行加权。对不同行为产生的用户喜好进行加权，然后求出用户对物品的总体喜好。

当收集好用户的行为数据后，还要对数据进行预处理，最核心的工作就是减噪和归一化。减噪是指因为用户数据在使用过程中可能存在大量噪音和误操作，所以需要过滤掉这些噪音。归一化是指不同行为数据的取值相差可能很大，例如用户的查看数据肯定比购买数据大得多。通过归一化，才能使数据更加准确。

通过上述步骤的处理，就得到了一张二维表，其中一维是用户列表，另一维是物品列表，值是用户对物品的喜好。还是以电影推荐为例，如下表：

|  |  |  |
| --- | --- | --- |
|  | Tatanic | Avatar |
| User1 | 5 | 3 |
| User2 | 4 | 3 |
| User3 | 4 | ？ |

表2.1 不同用户对电影的评分

然后是找到相似的用户或者物品。对用户的行为分析得到用户的喜好后，可以根据用户的喜好计算相似用户和物品，然后可以基于计算结果进行推荐。在计算用户之间的相似度时，是将一个用户对所有物品的偏好作为一个向量，而在计算物品之间的相似度时，是将所有用户对某个物品的偏好作为一个向量。求出相似度后，接下来可以求相似邻居了。

接下来是计算并推荐。在上一步求相似邻居的时候，通常是求出TOP-k邻居，然后根据邻居的相似度权重以及他们对物品的偏好，对当前用户没有评分的未涉及物品生成一个预测评分，并通过对这个预测评分从高到低排序，得到一个物品列表，然后推荐列表中的前几个物品。表2.2和表2.3就是在电影推荐系统中，对一个用户进行推荐之后，利用Pearson系数计算得出的TOP-k邻居和对被推荐用户未看过部分电影的预测评分。

|  |  |
| --- | --- |
| 用户编号 | 相似度 |
| 366 | 0.30 |
| 417 | 0.28 |
| 378 | 0.27 |
| 550 | 0.25 |
| 528 | 0.24 |

表2.2 被推荐用户的TOP-5邻居

|  |  |
| --- | --- |
| 电影 | 预测分数 |
| Caddyshack (1980) | 5.41 |
| Bull Durham (1988) | 5.40 |
| Goods: Live Hard | 5.28 |
| Texas Chainsaw Massacre | 5.27 |
| Larry David | 5.26 |
| Descendants | 5.21 |

表2.3 对某个用户得出的推荐结果

2.2.3 协同过滤算法的相关问题

协同过滤与基于内容的过滤相比，优势就是可以在根据各个用户的历史信息推荐项目，跟项目本身的内容属性无关。尽管协同过滤技术取得了一定成功，但仍然存在一些问题：

1.冷启动问题：在产品刚刚上线、新用户到来的时候，如果没有用户在应用上的行为数据，也无法预测其兴趣爱好。另外，当新商品上架也会遇到冷启动的问题，没有收集到任何一个用户对其浏览，点击或者购买的行为，也无从对商品进行推荐。

2.数据稀疏性问题：当用户仅对数据库中可用的项目中的一小部分进行评分时，就会导致这种问题。

3.可扩展性问题：这是与推荐算法相关的另一个问题，因为计算通常随着用户和项目的数量线性增长。当数据集的量有限时，推荐技术是有效可行的，但当数据集的量增加时，生成推荐的量就不太好。在这种情况下，用于解决可扩展性问题和加速推荐生成的方法会基于降维技术，例如奇异值分解（SVD）。

2.3 差分隐私

差分隐私是一种基于该理论的隐私模型——某个计算的输出不应该让任何输入数据推测出来。为了达到这个目标，必须保证所有计算结果的概率分布不会因为在输入当中加入或者删除任何特别的记录而发生显著变化。举个例子，假设现在有一个数据库可以查到有多少人单身。刚开始的时候查询发现，2个人单身；现在张三跑去登记了自己婚姻状况，这个时候再一查，发现有3个人单身，所以得出张三就是单身。现在如果我们不想让攻击者知道张三是否处于单身，就要用差分隐私来保护他的信息。

差分隐私的数学概念为：设有一个随机算法A，Range（A）为A所有可能的输出构成的集合。对于任意两个邻近数据集 和以及Range（A）的任何子集t，若算法A满足：，则称算法A提供了 — 差分隐私保护，其中，参数称为隐私保护预算，Pr表示发生某一事件的概率。一般而言， 越小，隐私保护程度越好。

差分隐私为缓解通过推荐系统的输出来推测用户隐私数据的风险提供了手段。一个实现差分隐私常用的方法是通过拉普拉斯(Laplace)机制，对服从拉普拉斯分布的噪声数据进行仔细调整并将其加人到计算当中，而加入噪声的多少就取决于敏感度和隐私预算的取值。这些噪声屏蔽了在某条记录中任何变化对计算结果所造成的影响。

Laplace分布是统计学中的概念，是一种连续的概率分布。如果随机变量的概率密度函数分布为：

, 𝜇 均为常数

图2.4 Laplace概率密度函数

那么它就是拉普拉斯分布。其中，𝜇 是位置参数， 是尺度参数。在实现的时候一般把 𝜇 设为0，由得到。是差分隐私的隐私预算， 则称为敏感度。在差分隐私中，总共有两种敏感度：全局敏感度和局部敏感度。本文中所实现系统只用到了全局敏感度，故只介绍前者。它代表的意思是对于两个相邻数据集，一个查询函数 最大的变化范围。其数学公式为：

表示函数 在数据集上的查询结果。矩阵范数和下标1一起代表相邻数据集的所有查询结果之差的最大值，即这两个数据集查询结果之差最大是1。

对原始数据加入服从拉普拉斯分布的噪声，则称为引入了拉普拉斯机制。拉普拉斯机制的定义是给定查询函数 ，满足

其中，表示最终的一个确定的查询结果 加上一个不确定的随机噪声 得到的最终结果，是独立同分布的变量，为 )。则该机制满足 — 差分隐私，证明过程如下：

假设表示 的pdf（probability density function），表示 的pdf，则对于某个输出，我们有

在刚才使用差分攻击成功获得张三是单身的例子里，本来两次查询结果是确定的2个单身和3个单身，现在加入拉普拉斯噪声后，确定的结果变成了两个随机变量，在多次查询之后，将查询结果进行统计，统计结果为期望分别为2和3拉普拉斯分布，如下图所示。现在，如果张三不在数据库的话，得到结果可能是2.5；张三在的话，得到的结果也可能是2.5；两个数据集查询得到某一个结果的概率很接近，以至于我们根本分不清这个结果来自于哪一个数据集。这样攻击者就没有办法在单次查询中得到具体的数值，就有效防止了个人的隐私泄露。

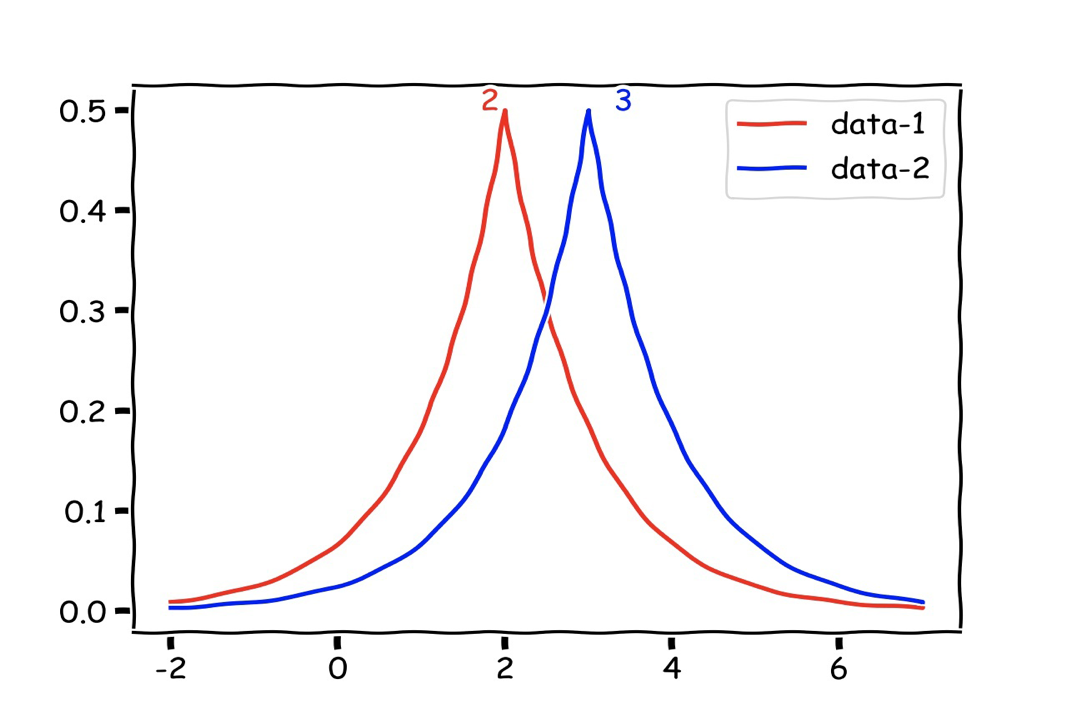


图2.5 加入噪声之后的查询结果（多次）

还有一种机制是指数机制。指数机制与拉普拉斯机制不同，前者是简单地对输出的数值结果加入噪声实现差分隐私。而对于非数值型数据而言，它的输出是一组离散数据中的元素。

指数机制整体的思想就是，当接收到一个查询之后，不是确定性地输出一个 结果，而是以一定的概率值返回结果，从而实现差分隐私。而这个概率值则是由打分函数确定，得分高的输出概率高，得分低的输出概率低。

差分隐私具有两个最重要的优点：

1.严格定义了攻击者的背景知识：除了某一条记录，攻击者知晓原数据中的所有信息——这样的攻击者几乎是最强大的，而差分隐私在这种情况下依然能有效保护隐私信息；

2.差分隐私拥有严谨的统计学模型，极大地方便了数学工具的使用以及定量分析和证明。

正是由于差分隐私的诸多优势，使其一出现便迅速取代了之前的隐私模型，成为隐私研究的核心，并引起理论计算机科学、数据库与数据挖掘、机器学习等多个领域的关注。

但差分隐私并不完美，也有一些缺点。最主要的就是由于对于背景知识的假设过于强，需要在查询结果中加入大量的随机化，导致数据的可用性急剧下降。特别对于那些复杂的查询，有时候随机化结果几乎掩盖了真实结果。这也是导致目前应用不多的一个原因。为了平衡数据可用性和隐私保护，有时候人们采用较为松弛的高斯机制来代替拉普拉斯机制。

2.4 本章小结

本章介绍了与作品相关的背景知识。首先对推荐系统进行概述，介绍了其概念和功能，然后介绍了推荐系统的常用实现方法——协同过滤算法，主要及分析并介绍了基于用户的协同过滤推荐系统的实现原理和主要步骤。然后介绍了本文的另一主题——差分隐私。对其定义、概念和实现机制进行了介绍，最后分析了差分隐私的优缺点。

第三章 基于差分隐私保护的协同过滤推荐系统

3.1 对于推荐系统的攻击方式

在文献[1]中，作者等人可以结合辅助信息推断出用户的历史评分和行为。辅助信息的定义是：对于任意用户，他的历史信息为，则定义其辅助信息为可以背攻击者所获得到的历史信息子集。辅助信息的的常见获取方式有以下三种：

1.许多网站会直接将用户对某些商品的评分和评价进行公开，有部分网站还会询问用户是否对外公开自己的交易记录；

2.用户有时候自己会在某些A网站上泄露有关于自己的隐私信息。这个时候在使用A的账号第三方登录B网站时，B网站有时候会拉取该用户在A网站的数据；

3.现实生活中，用户的一些行为习惯和部分信息会在社交聚会中被泄露。

基于掌握的辅助信息，文献[2]中的一个算法可以推断用户的非辅助信息部分。

第四章 实验结果与分析

第五章 总结与展望

结束语

致谢

参考文献

附录

正文中文

The word in English

图名表名