目 录

[第1章 相关技术 1](#_Toc500280420)

[1.1 推荐算法相关技术 1](#_Toc500280421)

[1.1.1 基于用户的协同过滤 1](#_Toc500280422)

[1.1.2 基于物品的协同过滤 2](#_Toc500280423)

[1.1.3 概率矩阵分解 3](#_Toc500280424)

[1.2 统计方法相关技术 3](#_Toc500280425)

[1.2.1 核密度估计 3](#_Toc500280427)

[1.2.2 分布之间相似度度量方法 3](#_Toc500280428)

[1.3 推荐系统相关技术 3](#_Toc500280429)

[1.3.1 移动客户端Android平台概述 3](#_Toc500280431)

[1.3.2 推荐系统服务端概述 4](#_Toc500280432)

[1.4 本章小结 4](#_Toc500280437)

# 相关技术

个性化推荐系统是信息时代解决信息过载问题【】的关键技术之一。在过去的几十年里，推荐系统领域技术不断革新，发展迅猛，不论在科研教学、商业运作上，还是在企业决策、信息管理上，都扮演者极其重要的角色。本章将从推荐算法、统计方法和推荐系统三个角度简要介绍本文研究的推荐系统相关的核心技术。

## 推荐算法相关技术

个性化推荐系统在我们日常生活中随处可见，比如电商网站的“猜你喜欢”、新闻网站的推荐阅读，影视频网站的视频推荐等等。在这些推荐系统中都会使用到推荐系统中非常有效的算法来提供相应的推荐服务。在推荐系统中，有很多推荐算法，比如：协同过滤算法、隐语义模型方法、图模型方法、基于上下文的推荐算法、基于社交网络的推荐算法等等。这些算法广泛的应用在各大推荐系统中，给企业带来巨大利润。作为示例，本节要讨论现在主流的推荐系统中所广泛使用的几个核心算法：基于用户的协同过滤算法、基于物品的协同过滤算法以及概率矩阵分算法。这几种算法具有的典型思想将为本文提出的算法提供丰富的参考价值。

### 基于用户的协同过滤(700字，一个图，已完成)

基于用户的协同过滤算法算是推荐系统中最古老的算法【9】，该算法可以被看做是推荐系统诞生的标志【9】。其首先被用于电子邮件的过滤。基于用户的协同过滤算法其实是一种考虑用户的相关性而把相关用户偏好的物品推荐给目标用户的方法，属于基于邻域的推荐算法。

由此可见，基于用户的协同过滤算法主要包含着两个步骤：

1. 搜索与目标用户具有相似偏好的用户；
2. 在搜索的用户所偏好的物品中找出目标用户不了解的目标物品推荐给目标用户。

这两个步骤就是基于用户的协同过滤算法的基础步骤。在第一步中，计算相似性就有较多的方法，比如Jaccard相似度：



其中，为物品集合，为用户喜欢的物品集合，为用户喜欢的物品集合。用户和用户偏好的Jaccard相似度就可以定义为式。类似的还有余弦相似度：



用户与其喜欢的商品构成了用户-物品列表，利用用户-商品列表数据建立物品到用户的列表，如【图】所示。然后建立用户的相似矩阵，对物品到用户的列表遍历，认为喜欢同一物品的用户都是类似的，他们的相似的，更新他们之间的关系值。例如对于物品，有和两个用户喜欢它，则和都自增1.将物品-用户列表扫描结束后，就可以得出一个庞大的用户相似关系的矩阵了。然后利用衡量相似度的方法式或式来计算出用户之间的偏好相似度。



基于用户的协同过滤算法会将与目标用户最相似的Top-k个用户的喜欢的物品推荐给目标用户。用户对物品的感兴趣程度可以定义为：



其中，是用户偏好最相似的Top-k个用户集合，表示对物品表示喜欢的用户集合，表示用户和用户的偏好相似度，是用户对物品的偏好。据此即可给出推荐列表。

基于用户的协同过滤算法思路简单，但是性能略差。可以针对用户相似度计算进行改进，是计算出来的用户相似度更加逼真，文献【39】就是利用这种思路提升了推荐质量。

### 基于物品的协同过滤(500字，已完成)

基于物品的协同过滤算法是目前工业上应用最广泛的推荐算法，首先由亚马逊提出【40】。由于基于用户的协同过滤算法存在一些天生的劣势，如：用户数量增加，产生用户偏好的相似度矩阵的计算时间会非线性提升，另外，也无法对产生的推荐结果做出合理解释。基于物品的协同过滤算法通过分析用户的行为来计算物品之间的相似度，用户对两种物品都喜欢则两种物品的相似度就越高。

该算法也主要分为两个步骤：

1. 计算物品的相似度；
2. 利用计算的相似度与用户发生的行为对目标用户做出推荐。

计算物品相似度的方法如下：



计算好物品间的相似度后，与式类似，基于物品的协同过滤算法通过式计算用户对物品的偏好：



这里，是用户喜欢的物品集合，是与物品最相似的Top-k个物品集合，是物品与物品的相似度。

基于物品的协同过滤算法也可以从相似度的角度来优化，文献【41】是通过分析用户活跃的程度来对物品相似度度量从而提高推荐质量。另外，还可以对物品相似度进行正规化来提高覆盖率与多样性，文献【42】给出相关的研究结果。

### 概率矩阵分解(700字，一个图)

（定义）

由于传统的协同过滤算法对大数据量的处理性能较差，文献【27】提出了概率矩阵分解理论。概率矩阵分解的核心思想是线性因子分析模型，用户的偏好建模成几个向量的线性组合，然后使用梯度下降法迭代求出特征向量。

（基本步骤）

（特点）

## 统计方法相关技术

（统计方法在推荐系统中的应用简介）



### 核密度估计

（定义）

### 分布之间相似度度量方法

（分布间相似度度量的作用、常用方法介绍、各种方法的特点）

## 推荐系统相关技术

（将推荐算法应用到实际的项目场景中去的实施化方案）



### 移动客户端Android平台概述

（简要介绍Android平台）

### 推荐系统服务端概述

（服务端架构）

（服务端应用的主流技术框架）



## 本章小结

# 