

**本科毕业论文**

**视频平台的多样性推荐算法研究**

院 系\_\_\_\_\_\_管理学院\_\_\_\_\_\_\_\_\_

专业班级\_管理学创新实验1801班\_

姓 名\_\_\_\_\_\_\_\_张晴\_\_\_\_\_\_\_\_\_\_\_

学 号\_\_\_\_\_U201815999\_\_\_\_\_

指导教师\_\_\_\_\_\_\_程海芳\_\_\_\_\_\_\_\_\_\_

2022年 6 月 5 日

**学位论文原创性声明**

本人郑重声明：所呈交的论文是本人在导师的指导下独立进行研究所取得的研究成果。除了文中特别加以标注引用的内容外，本论文不包括任何其他个人或集体已经发表或撰写的成果作品。本人完全意识到本声明的法律后果由本人承担。

作者签名： 年 月 日

**学位论文版权使用授权书**

本学位论文作者完全了解学校有关保障、使用学位论文的规定，同意学校保留并向有关学位论文管理部门或机构送交论文的复印件和电子版，允许论文被查阅和借阅。本人授权省级优秀学士论文评选机构将本学位论文的全部或部分内容编入有关数据进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。

本学位论文属于 1、保密 囗 ，在 年解密后适用本授权书。

2、不保密 囗 。

（请在以上相应方框内打“√”）

作者签名： 年 月 日

导师签名： 年 月 日

摘 要

在哔哩哔哩等用户端创作的视频平台，每个人都能无门槛地成为内容的创作者，信息过载问题在这些平台上显得尤为严重，推荐系统成为整个平台的核心。

现有的视频平台推荐系统存在推荐内容同质化严重的问题，忽略了用户对多样化内容的需求，降低了用户对推荐系统的使用体验。但是当前关于提高视频平台推荐算法多样性的研究还不多，研究重点仍然在提高准确度上面。准确度和多样性是两个不同方向的目标，两个指标的平衡一直是推荐算法的重点，而用户在多样性和准确度的偏好上展现出了个体差异和动态变化，简单的提高推荐系统在准确性和多样性的表现，很难广泛提高用户的使用体验。

为了解决以上问题，展开研究工作，本文研究内容和主要创新点如下：

（1）本文提出一种基于多臂赌博机的多样性优化推荐算法，通过建立多目标函数，引入准确度指标之外的列表多样性和时间多样性指标，作为向用户推荐视频时的重要参考指标，来提高单次推荐与多次推荐的多样性。

（2）算法为不同指标设定权重，并根据用户反馈对权重模型动态更新，通过多次迭代使权重模型更加符合用户偏好，在用户偏好改变时，跟踪用户偏好变化，将提高准确度和多样性的问题转化为直接满足用户个性化需求的问题。

（3）将中短视频平台中广泛的标签转化为对视频的描述，利用多臂赌博机不断测试各个标签的收益，并使用视频标签来计算视频的期望收益；将时间衰减函数与视频相似度结合，作为时间多样性指标，评价视频在时间多样性上的表现。

关键词：视频推荐算法； 多样性； 时间多样性； 多臂赌博机

Abstract

On video platforms created by users such as Bilibili, everyone can become a content creator without any barriers. The problem of information overload is particularly serious on these platforms. The recommendation system has become the core of the entire platform.

The existing video platform recommendation system has the serious problem of homogeneity of the recommended content, ignoring the user's demand for diversified content, and reducing the user's experience of using the recommendation system. However, there are not many studies on improving the diversity of video platform recommendation algorithms, and the focus of research is still on improving the precision. Precision and diversity are two goals in different directions. The balance of the two indicators has always been the focus of the recommendation algorithm, and users show individual differences and dynamic changes in their preferences for diversity and precision. Simply improving the recommendation system on precision or variety is difficult to widely improve the user experience.

In order to solve the above problems, this paper makes the following research and contributions:

(1) This paper proposes a multi-armed gambling machine-based diversity optimization recommendation algorithm. By establishing multi-objective functions, list diversity and temporal diversity indicators other than precision indicators are introduced as an important reference when recommending videos to users. indicators to improve the diversity of single recommendation and multiple recommendation.

(2) The algorithm sets weights for different indicators, and dynamically updates the weight model according to user feedback. Through multiple iterations, the weight model is more in line with user preferences. When user preferences change, tracking user preferences changes will improve precision and diversity. The problem of sexuality has turned into a problem of directly meeting the individual needs of users.

(3) Convert the extensive tags in the short and medium video platform into the description of the video, use the multi-arm gambling machine to continuously test the user's preference for each tag, and use the video tag to estimate the user's interest in the video; time decay function Combined with video similarity, it is used as a temporal diversity indicator to evaluate the performance of videos on temporal diversity

Key Words**:** Video recommendation algorithm; Diversity; Temporal diversity; Multi-armed gambling machine

目 录

[目 录](#_Toc"806247) 1

[摘 要](#_Toc104397068) 2

[Abstract](#_Toc104397069) 3

[1 绪论](#_Toc104397070) 5

[1.1 研究背景与意义](#_Toc686379293) 5

[1.2 研究现状](#_Toc104397072) 6

[1.3 研究内容与论文结构](#_Toc104397073) 8

[2 相关理论基础](#_Toc104397074) 10

[2.1 协同过滤推荐](#_Toc104397075) 10

[2.1.1 基于用户的协同过滤推荐](#_Toc104397076) 10

[2.1.2 基于项目的协同过滤推荐](#_Toc104397077) 11

[2.2 多臂赌博机主要算法](#_Toc104397078) 12

[2.2.1 贪心算法](#_Toc104397079) 12

[2.2.2 置信区间上界算法](#_Toc104397080) 12

[2.2.3 线性上行置信区间边界算法](#_Toc686379303) 12

[2.2.4 协同过滤和多臂赌博机算法](#_Toc104397082) 13

[2.2.5 汤普森抽样算法](#_Toc104397083) 13

[3 基于多臂赌博机的多样性优化推荐算法](#_Toc104397084) 15

[3.1 视频平台推荐系统问题分析](#_Toc104397085) 15

[3.2 多样性推荐算法的基本原理](#_Toc104397086) 16

[3.3 提高多样性的目标及推荐算法](#_Toc104397087) 17

[3.3.1 提高多样性的三个目标](#_Toc104397088) 17

[3.3.2 提高多样性的推荐算法](#_Toc104397089) 22

[3.4 提高多样性的推荐算法改进策略](#_Toc104397090) 25

[3.5 算法流程与描述](#_Toc104397091) 26

[3.5.1 算法流程](#_Toc104397092) 26

[3.5.2 算法描述](#_Toc104397093) 28

[4 实验与结果分析](#_Toc103165687) 29

[4.1 实验数据收集与预处理](#_Toc103165688) 29

[4.2 实验设置](#_Toc103165691) 30

[4.2.1 实验算法与对照算法](#_Toc103165692) 30

[4.2.2 评价指标](#_Toc103165693) 31

[4.3 实验结果与分析](#_Toc103165694) 32

[4.3.1 算法效果评估](#_Toc104397100) 32

[4.3.2 参数探索](#_Toc104397101) 34

[4.3.3 用户偏好分析](#_Toc104397102) 35

[5 总结与展望](#_Toc104397103) 36

[参考文献](#_Toc104397104) 38

# 1 绪论

## 1.1 研究背景与意义

人们的生活方式，生产方式，与信息获取方式被不断涌现的新技术快速改变着，其中互联网彻底颠覆了人们的日常生活。一方面，人们可以随时随地通过各种渠道快速获得互联网上的广泛信息，另一方面，每个用户都可以成为信息源和网络内容的创作者，通过智能设备上传信息。于是互联网上每天都被上传海量信息，这些信息在网上由用户进一步传播，进一步创作，被广泛使用，导致了信息过载问题（Informatica Overload）的出现。信息过载问题给人们从海量信息中搜寻有效信息带来了更大的困难和更高的成本，尤其在中短视频、新闻、和电子商务等供给端生产者众多的领域，信息过载问题更为严重。为了满足用户快速获取有效信息的需要，减少信息过载问题给人们带来的不便影响，个性化推荐系统成为了解决这一问题的最重要且有效的技术[1]。

推荐系统（Recommender System）旨在挖掘用户的兴趣并为每个用户推荐与其兴趣匹配的内容。推荐系统的研究和开发大多都集中在预测和如何提高推荐系统的准确度上，然而，人们越来越认识到，多样性在评价推荐系统性能上的重要性[2]。用户对不同种类或不同标签的兴趣往往存在个体化差异，同时，用户的兴趣一般也存在多样性，而不是单一的。因此，使用推荐系统向用户推送内容时，不能仅仅向用户推荐那些用户表现出兴趣的内容，还需要注重内容的多样性，使推荐内容符合用户兴趣的同时尽量多样。然而，提高多样性多样性往往会带来准确性的损失，如何平衡推荐算法的准确度和多样性就成为了一个关键挑战。

推荐系统最初是针对电商平台提出的[3]，后来逐渐扩展到新闻推荐和音乐推荐等其他领域，视频推荐是个性化推荐系统的分支之一。在中国，抖音和快手作为短视频平台的代表快速发展，B站，西瓜视频等中视频发展也面临新机遇，新挑战。视频平台凭借广泛的内容，发展正在不断成熟，在用户生活中逐渐占据更大的空间，成为用户生活中极为重要的一部分。推荐系统是整个视频平台的核心，对用户的使用体验起着决定性作用，即使一个视频平台内容再丰富精彩，没有合理的推荐算法也很难占用用户的更多时间。为用户推送符合用户兴趣的个性化内容可以提高用户粘度，延长用户的使用时间。但过度重视准确度，而忽视推送视频的多样性，也会损害用户体验，因此提升视频推荐的多样性是一项有价值的技术研究。

与此同时，当今的视频推荐在研究往往忽略时间多样性（新颖度）问题，时间多样性是多样性的一个研究方向，与新颖度问题的研究存在重叠部分，时间多样性强调的是为用户推荐内容在时间上的多样性，由Lathia[4]等初次提出并定义。实验表明在一段较短时间内，多次为用户推荐重复内容，或大量的相似内容，这种情况也会降低用户对推荐内容和推荐系统的评价[4]。

综上所述，除了准确度，多样性也是推荐系统的重要评价指标，如何平衡准确度与多样性一直是推荐系统考虑的重要问题，本文设计一种适用于视频平台的多样性优化推荐算法，通过设置多目标函数来提高推荐的列表多样性和时间多样性，通过对准确度指标和多样性指标的权重进行迭代更新来满足用户的个性化多样性偏好。

## 1.2 研究现状

个性化推荐系统的研究重点长期以来一直是提高推荐准确度，但近十年来关于推荐系统多样性的研究逐渐受到更多关注，也有许多专家和学者不断提出提高推荐系统多样性的方法，电子商务，新闻推荐，视频推荐等不同平台的推荐系统多样性研究实例也在不断增多。本节主要介绍国内外在提高推荐算法的多样性上的研究现状，以及在视频平台的多样性推荐算法的相关研究。

推荐系统的多样性优化方法大体上可以分为两类方法，第一类为非交互式的优化方法，非交互式方法通常使用用户与项目的历史交互信息训练模型，然后根据训练后的模型从内容池中选择并推荐项目。

Ziegler等[5]在2005年使用主题多样化来平衡推荐列表的个性化和多样化，将列表内相似性度量来作为评估推荐列表局部多样性的指标，并使用了主题多样化方法来减少列表内的相似性。Zhang等[6]在2009年基于用户历史记录表现出的兴趣进行推荐，将用户兴趣划分为不同维度，然后为每个维度的兴趣形成推荐列表，再将多个推荐列表进行集成，使推荐内容满足用户兴趣且尽量多样，在平衡准确度和多样性上取得了较好效果。Adomavicius[7]在2011年提出一种提高总体多样性的基于商品流行度的对列表重排序方法，但是会造成准确度的损失。Boim等[8]在2011年比较不同用户给项目的排名来估计项目的多样性，利用皮尔逊相关系数来度量物品之间的相似度，通过构建优先覆盖树来提高推荐列表的多样性，并且可适用于没有可用项目语义信息的场景。Javari等[9]在2015年提出用概率结构实现多样性优化，通过建立相关性和多样性的混合模型，解决推荐系统中多样性和准确度指标相互矛盾的情况。Cheng等[10]在2017年将每个用户视为一个训练实例，由一组经验确定的相关和不同的物品标记，提出了一种自动启发式标记方法来权衡相关性和多样性。Li等[11]在2017年提出了一种新的排序模型，该模型利用项目的分解类别特征来进行多样化的推荐，准确度和多样性结合在一个统一的函数中，并创建一个最佳列表。

第二类是交互式的优化方法，交互式方法通常根据用户反馈动态更新模型，主要过程可以概括为：推荐，反馈，更新等三个环节。深度学习和多臂赌博机是两种常见的交互式方法。

Zheng等[12]在2018年设计了基于深度Q-learning的推荐框架，将用户返回模式作为用户是否点击推荐内容的补充，以获取更多的用户反馈信息，采用一种决斗赌博机梯度下降算法进行探索，在多样性上取得了较好的效果。Qin等[13]把熵正则化器纳入奖励函数（最初仅由准确度构成），以促进所选臂的多样性，并在实验中证明提高了推荐系统的性能。帖军等[14]在2018年对LinUCB进行改进，提出基于LinUCB的多目标优化推荐算法，引入基尼系数作为衡量推荐商品覆盖率的指标，提高对长尾商品的挖掘能力，同时保证了推荐内容的准确度。Hidayatullah等[15]在2018年使用排列赌博机形成商品推荐列表，以开发概率选择准确度，多样性和新颖性三个指标的加权期望最大的商品，根据用户反馈进行指标权重的更新。何炜俊等[16]在Hidayatullah的基础上进一步研究，基于理想点法构建多目标函数，以开发概率选择加权遗憾值最小的商品形成商品推荐列表。

上述介绍的多样性推荐算法大多是为一般平台设计出的一般性推荐算法，也有针对新闻推荐等与视频推荐存在较大差异的推荐算法。视频平台和电子商务平台，音乐平台等其他平台有许多不同的特点，推荐系统也存在特殊性，现在针对视频推荐系统的多样化优化的方法还不多。张艳红等[17]在2019将信息熵与用户配置信息长度作为评估项目各个属性的多样性依据，得到用户多样性对于两个指标组合而成的4个象限的隶属度，通过隶属度对推荐列表进行重新排序，提高推荐的多样性，且未损失准确度。

## 1.3 研究内容与论文结构

本文的研究重点是设计适用于视频平台的基于多臂赌博机的多样性优化推荐算法。为解决当前视频推荐系统中推荐内容同质化严重的问题，采取多目标函数构建的方式，在多臂赌博机进行探索开发过程中，引入列表多样性和时间多样性指标，提高视频平台推荐系统的多样性。通过权重更新，为用户设置符合用户偏好的指标权重，基于用户需求来平衡多样性与准确度。

研究内容包括：

（1）构建多目标函数用于多臂赌博机挑选视频。在多臂赌博机进行探索开发过程中，将对视频准确度收益的探索转化为对标签收益的探索，再使用标签收益估计视频的准确度收益，作为视频的准确度指标的评分；同时引入列表多样性指标，设计并引入时间多样性指标，来提高单次推荐的多样性和推荐内容在时间上的多样性。

（2）权重模型动态更新。不同用户在准确度，列表多样性和时间多样性这三个指标上有不同的偏好，在交互式推荐中使用用户对推荐列表的反馈来更新权重模型，为用户构建符合用户指标偏好的权重模型。

（3）设计实验，进行结果分析。将本文设计的算法与固定权重模型，协同过滤算法等对比，进行实验。通过比较算法在准确度，列表多样性，和时间多样性等不同指标上的表现，验证本文算法的效果，并为本文算法找到合适的探索概率参数。

在结构上，本文主要包括五个章节。

第一章：绪论。先对视频平台的多样性推荐算法的研究背景进行了介绍，说明研究意义；之后介绍了多样性推荐算法的研究现状；最后介绍了文章的主要研究内容和文章的组织结构。

第二章：相关理论基础。介绍了协同过滤算法的类型，原理，实现的过程，与适用的场景；之后介绍了几种常见的多臂赌博机算法的原理与各自的特点。

第三章：基于多臂赌博机的多样性优化推荐算法。首先，对当前视频平台推荐系统存在的问题进行分析；之后介绍了基于多臂赌博机的多样性优化推荐算法的基本原理；然后构建多目标函数，并对视频在不同指标上评估方法进行说明；之后，介绍了各指标权重的动态更新策略；最后，介绍了总的算法流程与算法描述。

第四章：实验与结果分析。首先对实验的数据进行了介绍，介绍了对数据处理过程，以及实验中各算法效果的三个评价指标：准确率，列表多样性和时间多样性。将基于多臂赌博机的多样性优化推荐算法与协同过滤算法，固定权重模型等算法在三个指标上的结果进行对比，判断算法的效果，进行参数探索，为算法寻找最优探索概率。

第五章：总结与展望。总结本算法对视频推荐的贡献，并分析现存的不足，提出下一步的改进方向。

# 2 相关理论基础

## 2.1 协同过滤推荐

协同过滤推荐作为推荐算法被各种平台广泛使用，最早由Goldberg等提出。协同过滤推荐的核心是相似性，它利用用户群在喜好上的相似性与用户喜爱内容与其他内容的相似性来从内容池中筛选用户可能感兴趣的内容，并推荐给用户。协同过滤推荐通常可分为两种方法：基于用户的协同过滤推荐方法[18]（User CF）和基于项目的协同过滤推荐方法[19](Item CF)。

### 2.1.1 基于用户的协同过滤推荐

User CF是最常见的一种推荐算法，User CF利用用户与项目的交互信息及用户对这些项目的评价，计算用户间相似度，使用相似度找到目标用户的最近邻用户集。一般认为，相似用户对同样内容的评价也是相似的，则可以向目标用户推荐最近邻用户集评价较高的内容。

杰卡德系数与余弦相似度是用来计算用户间相似度的两种常见方式[20]，计算公式如和所示。N（a）表示用户喜欢的项目集合。表示集合的元素个数。

 ( SEQ MTChap \c \\* Arabic \\* MERGEFORMAT 2- SEQ MTEqn \c \\* Arabic \\* MERGEFORMAT 1)

 ( SEQ MTChap \c \\* Arabic \\* MERGEFORMAT 2- SEQ MTEqn \c \\* Arabic \\* MERGEFORMAT 2)

计算出目标用户和其他用户的相似度之后，利用相似度对用户进行排名，选择相似度最高的K个用户，构建出目标用户的最近邻用户集合，然后将近邻用户与目标用户的相似度和其对目标用户未评价项目的评价结合起来，预测用户对未评分项目的评价，如所示。

 ( SEQ MTChap \c \\* Arabic \\* MERGEFORMAT 2- SEQ MTEqn \c \\* Arabic \\* MERGEFORMAT 3)

表示目标用户的最近邻集合，邻近用户个数一般根据平台特点和用户个数进行自主设定，表示邻近用户对项目的评价或反馈。在中短视频平台，可以使用用户是否对项目进行点击/完播等反馈信息。

用户对项目的评价信息很大程度上决定了User CF的推荐效果，密集且准确的评价信息，能为用户构建更准确相似的最近邻用户群，并利用用户群的交互信息与评级，对目标用户进行较为准确的推荐。由于User CF是从所有的用户中搜寻目标用户的相似用户，当用户数量较少时，User CF效率较高，在用户较多的情况下表现出较低的效率。

### 2.1.2 基于项目的协同过滤推荐

Item CF利用项目之间的相似性进行推荐，基于用户评价较高的项目集，推荐与其类似的项目。如果某些项目之间具有极高的相似性，那么用户在喜欢某项目时，很大可能也会喜欢与它相似的项目，那么可以通过用户喜欢的项目，从内容池搜寻其相似项目，推荐给用户。

项目之间的相似度可以用余弦相似度来计算，如所示，表示喜欢项目i的用户集合。

 ( SEQ MTChap \c \\* Arabic \\* MERGEFORMAT 2- SEQ MTEqn \c \\* Arabic \\* MERGEFORMAT 4)

利用用户对交互过的项目的评价与项目间的相似度，可以估计用户对未接触过的项目的评价，并基于估计的评价选择项目推荐给用户。计算用户对项目的预测评价如所示

 ( SEQ MTChap \c \\* Arabic \\* MERGEFORMAT 2- SEQ MTEqn \c \\* Arabic \\* MERGEFORMAT 5)

表示与项目的最近邻集合，邻近项目个数一般根据平台特点和项目个数进行自主设定，表示用户对项目的评价或其他反馈信息。

用户和项目的交互信息是协同过滤技术的关键，User CF利用交互信息判断用户的相似度，再将最相似的K个用户的交互信息经过处理，作为为目标用户推荐的依据，但Item CF直接寻找与用户评价较高的项目的相似项目，其假设是用户会喜欢相似项目。基于项目的协同过滤适用于项目较少的场景。

## 2.2 多臂赌博机主要算法

多臂赌博机问题可以描述为：一个赌博机有多个臂，每个臂的收益的概率分布不同，其他条件相同，但用户并不知道每个臂收益的概率分布情况，用户要得到最大收益，只能通过实验的方式测试每个臂。实验过程并不是盲目的，而应该是遵循一定的实验策略，Bandit算法就是这些实验策略，依照实验策略的不同，可以将多臂赌博机算法分成不同的类型。

常见的应用于推荐系统的多臂赌博机算法有贪心算法（Greedy）、置信区间上界算法（）、线性上行置信区间边界算法（）、协同过滤和多臂赌博机算法、汤普森抽样算法（Thompson sampling）等。

### 2.2.1 贪心算法

贪心算法每次都选择历史收益最大的臂，但是这种情况容易落入“局部最优解”，缺少探索性，贪心算法在贪心算法上发展起来，与贪心算法不同的是，它设定了一个探索概率，每次挑选臂时，会以的概率进行探索：从所有臂中随机挑选一个；以的概率进行开发：选择当前表现最好的臂。的值越大，探索性越强，风险也越高；越小，表明越保守，探索性越弱。当为固定值时，前期可能起到较好的探索效果，但在后期累计遗憾值会出现持续增长。

### 2.2.2 置信区间上界算法

置信区间上界算法（ ）利用每个臂的实验结果来估计各个臂的期望收益和置信区间，然后根据各个臂置信区间上界的值对所有臂进行排序，选择置信区间上界的值最大的臂。

在实验过程中，一个臂被选择的次数越多，得到的平均收益越接近臂的真实收益，同时，臂的置信区间也会越来越窄。一方面，当臂被选择的次数较少时，它的置信区间会越宽，置信上界也会越大，这些不确定性较强的臂将会被挑选，另一方面，当臂的真实收益越高时，它的置信区间上界往往也会越大，这些收益较高的臂将会被挑选，就是通过挑选置信区间上界较大的臂来平衡探索和开发的问题。

### 2.2.3 线性上行置信区间边界算法

线性置信区间上界算法（）是基于产生的改进算法，由Yahoo!的Li等科学家在2010年提出[21]，用于新闻推荐。算法考虑的是每个臂的不同，而算法引入上下文信息（），用户和每一个臂的组合都可以构成一个上下文特征，并假设每个臂的期望收益与它的相关特征存在线性关系。则在挑选臂的过程中，可以利用臂的相关特征来估计臂的期望收益的置信区间，并选择置信区间上界较大的臂。同时，利用每次选择臂的收益来更新臂的期望收益与相关特征线性关系的参数。

给用户和臂构建上下文特征是进行有效挑选和推荐的前提，但在很多现实场景中，不具有构建上下文信息的条件，这大大限制了的使用的广泛性[22]。

### 2.2.4 协同过滤和多臂赌博机算法

协同过滤和多臂赌博机算法是将协同过滤算法与多臂赌博机进行结合的一种推荐算法，使用协同过滤算法的原理，根据用户的相似度划分出不同的用户群，根据项目的相似性将项目划分成不同的项目群。在挑选臂时，仍然使用多臂赌博机算法的原理，但臂的期望收益不再是仅限于一个用户的实验得出的估计值，而是对用户所在的用户群，即与用户相似的其他用户的集体的实验结果，计算出来的期望收益[23, 24, 25]；在使用基于项目的协同过滤算法时，计算单个臂的期望收益时，也不仅仅考虑单个臂的实验结果，还考虑与臂相似的其他臂的实验结果[23]。挑选臂后，得到臂的收益值，不仅对多臂赌博机对应的参数进行更新，还要对用户群，项目群的划分进行更新。

王宇琛等[25]将协同过滤算法与算法结合，引入邻近用户建立协同偏置向量，用邻近用户与目标用户的加权平均偏好代替用户偏好，其中邻近用户的权重由与目标用户的相似度决定，为目标用户进行推荐，有效提高了推荐性能。

### 2.2.5 汤普森抽样算法

汤普森抽样算法不考虑每个臂的具体收益，而是假设每个臂产生收益的概率满足概率分布[26]。设臂产生收益的概率为，的初始分布为，每次挑选臂时，汤普森抽样算法从每个臂的的概率分布中随机抽选一个随机数，并选择对应随机数最大的臂，并在收到臂的期望收益后，对臂产生收益的概率的概率分布进行更新，如果臂产生收益，则，更新后的的概率分布为；如果臂不产生收益，则，更新后的的概率分布为。假设臂进行推荐了次，其中产生收益的为次，不产生收益的次数为次，则次推荐后分布为。

通过多轮推荐，汤普森抽样算法可以得到的较为准确的概率分布，并利用的概率分布挑选有较大概率获得收益的臂，同时，因为汤普森抽样算法并不直接计算，而是利用的概率分布来挑选臂，能较好地平衡探索与开发问题。与其他算法相比，汤普森抽样算法的累计遗憾值更小，表现更好[27]。

# 3 基于多臂赌博机的多样性优化推荐算法

## 3.1 视频平台推荐系统问题分析

为了设计更符合用户需求的多样性推荐算法，本文通过对用户关于视频平台推荐系统使用体验进行问卷调查，来了解当今视频平台推荐系统存在的问题，并确定多样性推荐算法的改进方向。在此次调查中选择中视频平台哔哩哔哩（B站）作为研究和调查的对象。B站使用推荐页面作为主页面，为用户推荐视频，每次推荐形成推荐列表，因此推荐列表的多样性是B站推荐系统的重要维度，用户对于每次推荐内容的多样性会有更清晰的评价。本次调查共收集到有效问卷250份，主要针对偶尔使用或经常使用B站推荐页面的用户进行调查，收集这些用户对B站推荐系统的使用体验评价。

问卷调查结果中，对B站推荐系统非常满意的用户占7.27%，满意的用户占44.85%，还有4.24%的用户选择了非常不满意，4.85%的用户选择了不满意。B站推荐系统做不到让大多数用户满意，还有小部分用户认为B站推荐系统的使用体验很差，因此B站推荐系统还存在不足榷待改进。

在用户使用推荐系统的过程中，47.88%的用户认为B站推荐了自己不感兴趣的视频，55.15%的用户被多次推荐同一视频，63.03%的用户曾被B站密集推送同一类型视频。用户的使用体验会在不同程度上被这些经历影响，43%的用户认为被推送完全不感兴趣的内容会影响自己的使用体验，在被推荐了不感兴趣的视频的用户中这一比例为59.49%；41.82%的用户认为被密集推送同一类型内容会影响自己的使用体验，而在经历过被密集推送同一类型的内容的用户中这一比例为有50.96%；51.52%的用户认为多次被推送同一视频会影响自己的使用体验，多次被推荐同一视频的用户中这一比例为58.24%。

不同的用户对于推荐系统的三个评价指标：准确性，列表多样性，时间多样性有不同的偏好特点。但总的来说，在这三个指标上敏感的用户都分别达到了40%，在某些指标上体验很差的用户对对应的指标存在更高的敏感性。因此，在设计推荐算法时，除了准确度，列表多样性和时间多样性也该作为重要指标，添加到推荐算法中。同时考虑不同用户在这三个指标上的不同偏好，通过动态交互测试出用户对不同指标的偏好程度，反馈到推荐过程中，达到基于用户需求的列表多样性，时间多样性和准确度的平衡。

用户对于被推送未接触领域的视频的态度是：4.25%认为完全能接受，35.15%认为能接受，44.85%的用户保持中立，9.09%的用户不能接受，6.67%的用户认为完全不能接受。大多数用户对被推送新领域视频保持良好态度，因此将新颖内容推送给用户进行测试，并挖掘用户的新兴趣点是可行的。也有部分用户对这一情况非常抵触，因此在提高推荐内容的多样性时，也需要考虑用户对新颖程度的接受程度存在个体上的差异，通过用户反馈进行快速更改推荐策略。

用户对B站推荐系统的个性化和多样性的满意度的平均分数都在3-4之间，即用户对B站推荐系统的个性化和多样性的评价介于中立-满意之间，其中，多样性推荐的评分为3.55，明显低于个性化推荐的评分3.76。关于用户对B站推荐系统的期望，64.20%的用户希望推荐内容更符合个人兴趣，67.28%的用户希望推荐的内容更加多样，69.14%的用户希望内容更加新鲜，挖掘自己的新兴趣点。用户在推荐内容更个性化和提高列表多样性，提高时间多样性方面都有很大的期望，尤其对提高内容的列表多样性和时间多样性抱有很大的期望，因此推荐算法的设计考虑将这两个指标作为推荐系统为用户进行推荐的重要指标，加强推荐视频时，这两个指标的参考作用。

通过前期调查，提高推荐内容的多样性是可行的，可探索用户对未接触过的领域视频的兴趣点，提高推送的多样性。同时应及时应对用户反馈，减少用户对不感兴趣视频的接触与反感度。在后续的推荐算法设计中，应综合考虑推荐内容对用户的个性化需求与多样性需求的平衡与满足，对用户反馈进行及时准确的处理。

## 3.2 多样性推荐算法的基本原理

本文算法在协同过滤算法形成的Top-N视频推荐列表基础上，使用多臂赌博机模型，构建包含准确度，列表多样性，时间多样性这三个指标的多目标函数。通过多目标函数，计算候选视频在三个指标上的期望收益与各指标上的最高期望收益的差，通过加权的方式计算候选视频的加权遗憾值，然后以概率在候选视频中随机挑选视频，以概率选择加权遗憾值最小的视频，作为推荐视频。经过逐个挑选，形成包含j个视频的视频推荐列表，具体过程如图3-1。

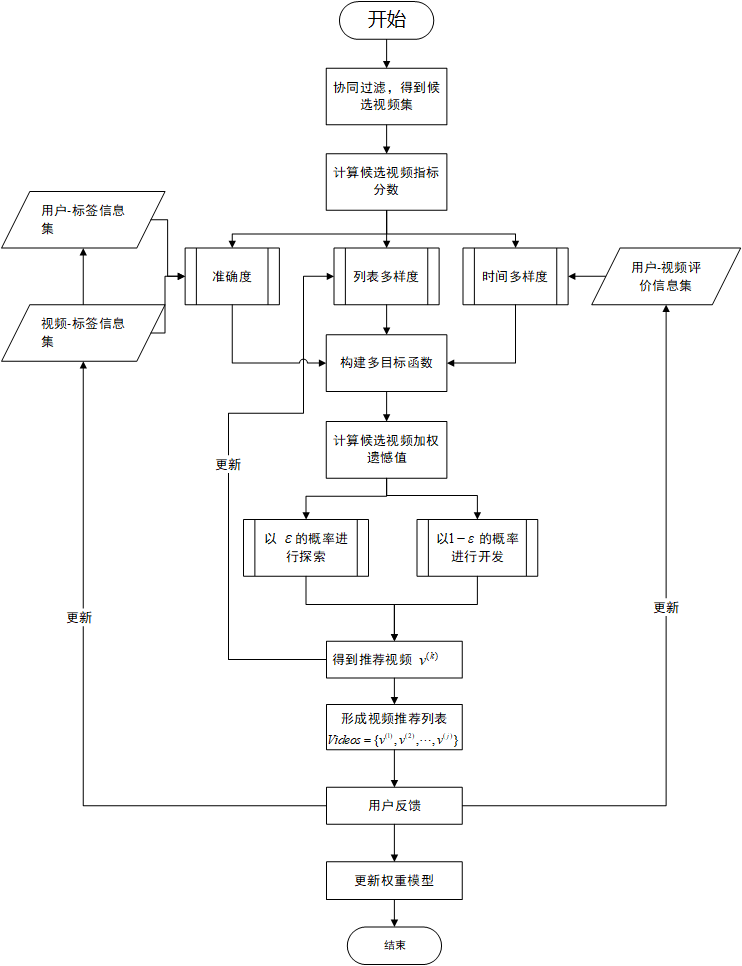


图3-1 用户推荐视频过程

经典多臂赌博机问题中包含h个臂，玩家每次只能选择其中一个臂，每个臂都有一定概率获得收益，每次尝试过后，玩家记录当次选择的臂的收益，并根据已有的臂收益记录来计算各臂收益，在下次选择臂时根据每个臂的期望收益，进行选择。使用-贪心策略，玩家可以的概率进行探索，随机选择一个臂，也可以的概率进行开发，选择使期望收益最大，或遗憾值最小的臂，玩家的目标是获得最大收益。使用多臂赌博机进行视频推荐过程中，视频即可以视为多臂赌博机的不同臂，但向用户进行重复推荐将影响用户的使用体验。

哔哩哔哩，快手等中短视频平台都包含广泛大量的标签，标签作为视频的描述能一定程度上反映视频内容，并且用来描述用户兴趣，构建用户兴趣模型。因此本文将多臂赌博机测试视频收益的过程转化为多臂赌博机测试视频标签收益的过程，利用推荐系统向用户推荐视频产生的收益，以及视频与标签信息集中视频与标签的对应关系，可以计算出推荐系统向用户推荐标签的收益。假设标签能很好地描述视频内容，且视频所含标签的收益能互相补偿，则可将视频所含标签期望收益的平均值作为视频的期望收益。通过多轮推荐，测试出在某用户处，标签的期望收益。

但在以上的测试期望收益的过程中，仅有准确度指标的收益能依托点击率，完播率等计算出来，列表多样性和时间多样性指标的收益不能直接通过测试得出。列表多样性和时间多样性仍需人为设定评价标准来评价。

但通过构建多目标函数，并为各个指标设定权重，使目标函数成为准确度，列表多样性，时间多样性三个指标的结合，根据候选视频的加权遗憾值来挑选视频，可以提高多臂赌博机算法的多样性。此外，指标权重利用用户反馈不断更新，在下一次推荐时，直接使用已更新的权重来计算候选视频加权遗憾值，并挑选视频，在视频推荐中，用户反馈表现为用户是否点击/完播推荐视频，通过多次迭代来确定用户对指标的偏好，达到满足用户个性化需求的目标。

## 3.3 提高多样性的目标及推荐算法

本节介绍候选视频在准确度指标，列表多样性指标，时间多样性指标上的得分的具体计算方法，与多目标模型的构建方法。

### 3.3.1 提高多样性的三个目标

#### （1）准确度

本算法使用视频的期望点击率/完播率作为视频在准确度指标上的得分。

多臂赌博机主要是通过多次实验来测试各个臂的收益，从而得到收益最大的臂，在使用多臂赌博机进行视频推荐时，考虑到对用户多次推荐视频将损害用户的使用体验，使用标签来描述视频，在多臂赌博机推送视频给用户时，相当于将视频所含标签推荐给用户，用户对视频进行点击/不点击，完播/切换的反馈也是用户对标签的反馈。通过多次向用户推荐含标签的视频，得到用户反馈，可以测试出用户对于标签感兴趣程度，使用标签描述视频，可以将用户对于标签的感兴趣程度用来估计用户对视频的兴趣程度。

在算法中，用户对视频所含标签的平均点击率作为视频的期望点击率，通过多臂赌博机不断推荐视频标签，测试出用户对标签的喜爱程度，为用户进行视频推荐时，用户对标签的平均点击率可以用式表示。

 ( SEQ MTChap \c \\* Arabic \\* MERGEFORMAT 3- SEQ MTEqn \c \\* Arabic \\* MERGEFORMAT 1)

表示包含标签的视频集合，表示已经被推荐给用户，且用户点击/完播的视频集合，表示被推送给用户的视频集合。

使用标签来作为视频内容的描述，并用用户对视频标签的平均兴趣程度估计用户对视频的兴趣程度。设视频的标签集为，利用标签估计视频的期望点击率（完播率）可用式计算得出。

 ( SEQ MTChap \c \\* Arabic \\* MERGEFORMAT 3- SEQ MTEqn \c \\* Arabic \\* MERGEFORMAT 2)

用算法1表示使用历史信息计算用户对标签点击率的策略。

|  |
| --- |
| 算法1.计算用户对标签点击率的策略 |
| 输入：标签，用户  输出：用户对标签的平均点击率  ①初始化标签点击率计算函数  ②  ③  ④  ⑤  ⑥  ⑦  ⑧ if =0 do:  ⑨ =0  ⑩ else do  ⑪  ⑫ end if  ⑬返回} |

用算法2表示使用标签计算视频期望点击率的策略。

|  |
| --- |
| 算法2.使用标签计算视频期望点击率的策略 |
| 输入：视频；视频标签集；用户  输出：用户对视频的期望点击率  ①for  do  ② CTR(u,d)#调用算法2  ③end for  ④  ⑤ |

#### （2）列表多样性

本文使用视频的列表多样度作为视频在列表多样性指标上的分值。个多臂赌博机一共挑选个视频，在逐个挑选视频时，计算候选视频集中，每个候选视频和已被选择的视频集中的视频的余弦相似度，并经过一定的处理，如式作为候选视频的多样度。

视频的余弦相似度在前面已经介绍过，候选视频与视频集中视频的相似度用式表示。

 ( SEQ MTChap \c \\* Arabic \\* MERGEFORMAT 3- SEQ MTEqn \c \\* Arabic \\* MERGEFORMAT 3)

在挑选第个位置上的视频时，候选视频的列表多样性分数用式计算得到。

 ( SEQ MTChap \c \\* Arabic \\* MERGEFORMAT 3- SEQ MTEqn \c \\* Arabic \\* MERGEFORMAT 4)

表示挑选推荐列表上第个位置上的视频时，已被挑选的个视频的视频集，表示中视频的个数 。

用算法3表示挑选第个位置上的视频时，计算候选视频列表多样性分数的策略。

|  |
| --- |
| 算法3.计算候选视频列表多样性分数的策略 |
| 输入：视频集Videos；候选视频  输出：候选视频的列表多样性分数  ①  ②  ③ if do  ④ for (=1,2,,) do  ⑤  ⑥  ⑦ if do  ⑧  ⑨  ⑩ else do:  ⑪  ⑫ end if  ⑬ end for  ⑭  ⑮  ⑯else do  ⑰  ⑱end if |

每次选择第位置的视频之后，都会更新，因此下一次选择视频时，候选视频的列表多样性指标分值也将如图3-1展示的过程进行更新。

#### （3）时间多样性

时间多样性解决相同的内容一次又一次推荐给用户的问题，或者被推荐的商品是否呈现一定程度的新颖性。本文希望引入时间多样性指标来解决视频推荐系统推荐内容同质化严重的问题，提高多次推送的视频内容的多样性，将视频的时间多样度作为候选视频在时间多样性指标上的得分。

徐元萍等[28]中介绍了基于新兴趣主题的新颖性指标：计算节点与用户曾点击过的项目的相似度，再经过处理，得到节点对于用户的新颖性指标的得分。而事实上，一个视频被推荐给用户，无论用户是否喜欢被推荐的视频，进行点击与否，在后面的推送中，此视频都会对用户对推荐内容的新颖度的感受产生影响。因此在计算候选视频的时间多样性指标上的值时，应将所有被推送给用户的视频考虑在内。

此外，人对事物的遗忘是在前期是快速的，但遗忘速度会随着时间逐渐减慢，视频时间多样度依靠候选视频与往期推荐视频的相似度来确定，在一段时间内密集推荐相似视频会降低用户对推荐系统的评价，且时间越短，推送越密集，用户的感受会越强烈，在这里引入时间衰减函数的概念，认为用户对一个已经浏览过的视频的记忆的衰减过程可以用式表示。

 ( SEQ MTChap \c \\* Arabic \\* MERGEFORMAT 3- SEQ MTEqn \c \\* Arabic \\* MERGEFORMAT 5)

表示记忆经过衰减的残留程度，为当前日期，为视频被推送给用户的日期，为自主设置的参数，决定用户对该视频的记忆衰减的速度。

候选视频在时间多样性指标上的分值，不仅由与已被推荐给用户的视频的相似度决定，还会受时间的影响。将与相似度的乘积作为候选视频与已被推荐的视频的相似度对用户来说经过时间衰减后的值，用式表示。

 ( SEQ MTChap \c \\* Arabic \\* MERGEFORMAT 3- SEQ MTEqn \c \\* Arabic \\* MERGEFORMAT 6)

用表示已被推送给用户的视频集，表示中视频的个数，候选视频的时间多样度可以用式计算得到。

 ( SEQ MTChap \c \\* Arabic \\* MERGEFORMAT 3- SEQ MTEqn \c \\* Arabic \\* MERGEFORMAT 7)

用算法4表示计算候选视频时间多样性分数的策略。

|  |
| --- |
| 算法4.计算候选视频时间多样性分数的策略 |
| 输入：被推荐给用户的视频集，用户与视频交互时间信息集；候选视频；推荐当天日期  输出：候选视频的时间多样性分数  ①设定参数  ②  ③if do  ④ for 与用户交互的视频 do  ⑤  ⑥  ⑦ If do  ⑧  ⑨  ⑩ end if  ⑪  ⑫  ⑬ end for  ⑭  ⑮  ⑯else do  ⑰ =1  ⑱end if |

### 3.3.2 提高多样性的推荐算法

#### （1）归一化处理

对于候选集中的视频，将所有视频的准确度，列表多样性和时间多样性指标的得分全部计算完成后，应分别将准确度，列表多样性和时间多样性分数列表进行归一化处理，作为视频在各指标上的期望收益。归一化策略可表示为：

 ( SEQ MTChap \c \\* Arabic \\* MERGEFORMAT 3- SEQ MTEqn \c \\* Arabic \\* MERGEFORMAT 8)

表示第个指标的分数列表，表示候选视频在第个指标的得分。表示候选视频在第个指标上的期望收益。

用算法5表示对指标分数进行归一化处理的策略

|  |
| --- |
| 算法5.指标分数的归一化处理 |
| 输入：候选视频集某指标分数列表  输出：归一化后的分数列表  ①初始化归一化函数ToNorm  ②ToNorm  ③  ④  ⑤ for #遍历列表中所有值  ⑥ if  do  ⑦  ⑧ else do  ⑨ =1  ⑩ end if  ⑪ end for  ⑫ 返回} |

经过处理后，在第轮推荐时，得到候选视频集中每个视频对应的各个指标的期望收益值，用期望收益向量表示，候选视频的期望收益向量的集合为。

#### （2）基于理想点法的推荐算法

准确度，列表多样性，和时间多样性是三个不同维度的目标，尤其是准确度和两个多样性指标，目标是相反的。采取标量函数构建的方法，对一个视频的三个指标值进行综合，为不同的指标设定不同的权重值，通过理想值点法，计算每个视频的加权遗憾值，并利用加权遗憾值挑选视频。以下是标量函数的构建过程。

本文设定准确度，列表多样性，时间多样性三个指标分别对应的权重为。

基于理想点法首先应分别计算出三个指标的最优值，由于前期已进行归一化处理，每个指标的最优值都为1。则理想点的期望收益为，然后以理想点为基准，计算候选集中每个视频的加权遗憾值，视频的加权遗憾值可以用式表示。

 ( SEQ MTChap \c \\* Arabic \\* MERGEFORMAT 3- SEQ MTEqn \c \\* Arabic \\* MERGEFORMAT 9)

只要，则作为非劣解，可以达到帕累托最优[16]。在推荐列表的第个位置，赌博机将以的概率选取使标量函数值最小，即加权遗憾值最小的视频作为位置的推荐视频，其他位置的视频挑选与此相同。即

 ( SEQ MTChap \c \\* Arabic \\* MERGEFORMAT 3- SEQ MTEqn \c \\* Arabic \\* MERGEFORMAT 10)

用算法6表示多臂赌博机基于多目标函数为用户形成视频推荐列表的策略。

|  |
| --- |
| 算法6.基于多目标函数的推荐列表的形成 |
| 输入：权重向量；候选视频集；候选视频期望收益集合；标量函数  输出：视频推荐列表  ①初始化：视频挑选函数；视频遗憾列表；  ②  ③ for 每个视频,(do  ④  ⑤ end for  ⑥  ⑦ 返回 }  ⑧for each多臂赌博机()：#多臂赌博机挑选位置上的视频  ⑨  #调用函数  ⑩  ⑪ del from #将被选择的视频从候选视频集中删除  ⑫end for |

## 3.4 提高多样性的推荐算法改进策略

在视频推荐列表中视频选择的过程，权重是非常关键的一部分，不同视频可能在不同指标上表现出不同优势，而最终是各指标的权重决定哪些视频能被推送给用户。决定权重的是用户表现出的对不同指标的偏好，这种偏好反映在用户对推荐列表的反馈中，本文利用用户的反馈对权重进行更新，满足用户的个性化需求，提高用户对推荐列表的满意度。

将包含个位置的视频列表推送给用户，用户会对各个位置的视频进行反馈，在哔哩哔哩等平台中，使用形成推荐列表的方式进行推荐，用户反馈的方式是是否点击视频，而在使用自动播放机制的视频平台，如快手等，用户反馈的方式是是否将视频完播，使用记录用户反馈，表示用户对第次推荐的反馈，用户点击/完播在第位置上的视频，则=1，反之，则=0。

参考何炜俊等[16]和Hidayatullah[15]对权重的更新方式：建立状态空间模型、使用卡尔曼滤波器计算数字滤波，构建权重模型更新机制，并引入权重更新速度调节系数调节权重更新速度，在进行第T次推荐后，以式的方式更新权重。

 ( SEQ MTChap \c \\* Arabic \\* MERGEFORMAT 3- SEQ MTEqn \c \\* Arabic \\* MERGEFORMAT 11)

为在第T次推荐的视频的期望收益向量经标准化处理之后的值，为推荐的视频对应的的方差，为推荐的位于第位置上的视频对应的用户反馈，用户点击/完播则为1，否则为0。

在权重更新过程中，记，则应满足：



通过计算，得到的上界（），设定，在保证更新得到的权重符合权重规范的同时，使权重沿通过反馈计算出的方向更新；此外，使每个指标的权重的单次变化幅度绝对小于0.25，维持权重更新速度的均衡，减小权重更新波动，从而调节权重更新速度。

随着权重模型的不断迭代更新，权重会逐渐接近用户对不同指标真实的偏好，如果用户在指标偏好上发生动态变化，比如说，一般新用户对比较喜欢看自己感兴趣的东西，但随着使用时间增多，用户会逐渐增强对陌生领域内容的探索兴趣，权重也会随之动态更新，跟踪用户偏好的改变。

|  |
| --- |
| 算法7.权重模型的更新 |
| 输入：权重向量；第次推荐视频收益向量集合；用户对推荐列表反馈向量  输出：权重向量  ①初始化：权重更新函数  ②  ③ for each do:  ④  ⑤  ⑥ end for  ⑦  ⑧  ⑨ for q=1,2,3 do  ⑩ if  do  ⑪  ⑫ elif  do  ⑬  ⑭ if  not null,do  ⑮  ⑯  ⑰ else do  ⑱  ⑲  ⑳ 返回 } |

## 3.5 算法流程与描述

### 3.5.1 算法流程

总结以上各算法，可以将为用户推荐视频的算法表示为图3-2。

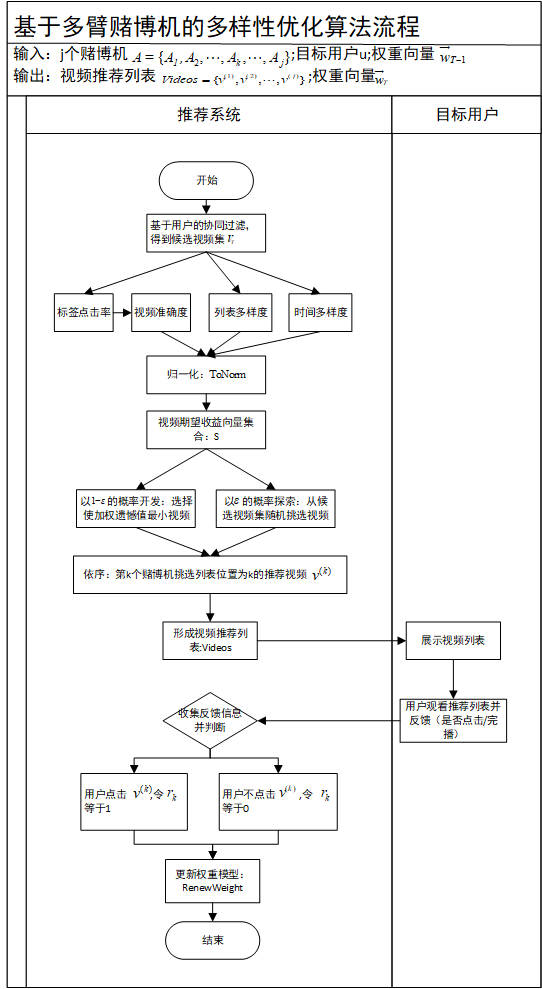


图3-2 算法流程图

### 3.5.2 算法描述

如算法8使用伪代码对基于多臂赌博机的多样性优化推荐算法作完整描述。

|  |
| --- |
| 算法8：基于多臂赌博机的多样性优化推荐算法 |
| 输入：个赌博机，即；目标用户；候选视频集合；权重；用户反馈  输出：视频推荐列表，权重。  ①计算视频各指标得分  ②归一化处理指标得分，得到视频期望收益向量集合  ③依序用第k个赌博机挑选列表位置为k的视频：重复k=1,2,j  ④ {生成一个（0,1）的随机数y:  ⑤ if ，开发：，#挑选加权遗憾值最小的视频  ⑥ elif，探索： random()}#从候选集中随机挑选  ⑦ del from  #将已被选择的视频从候选视频集删除}  ⑧ 为u展示推荐列表，并记录用户反馈：  ⑨ 若用户点击/完播，令，反之，则为0  ⑩ 更新权重模型：  ⑪end |

# 4 实验与结果分析

本章主要通过与其他算法进行对比实验，验证分析基于多臂赌博机模型的多样性优化推荐算法在实际为用户推荐视频过程中提高多样性的效果。同时，进行参数探索，找到最优的探索概率参数。本文在来自快手的kuairec真实数据集上进行实验，将本文算法，User CF和普通的固定权重多目标推荐算法进行对比，使用准确率，列表多样性，和时间多样性为评价指标，对实验结果进行比较，验证基于多臂赌博机模型的多样性优化推荐算法的效果。

## 4.1 实验数据收集与预处理

本次实验在数据集KuaiRec[29]中的small matric上进行。KuaiRec数据集是从中国短视频app快手的推荐日志中收集的一个真实数据集，除了small matric记录了用户和视频的交互信息：视频播放率（watch\_ratio），交互日期（date）；KuaiRec中还有item\_categories数据集，记录了每个视频的标签列表。small matric数据集的信息密度达到了99.6%，用户-项目矩阵中几乎没有缺失的值，也就是说，每个用户已经观看了每个视频，然后留下反馈。

Kuairec中的small metric数据集一共包含1411个用户和3327条视频的相关数据。本文使用数据集中用户与视频交互的“watch\_ratio”信息，经过处理，作为用户对推荐系统推荐视频的反馈。“watch\_ratio”表示的是用户对视频的播放率，如果watch\_ratio<1，则用户未观看完视频，如果watch\_ratio1，则用户至少播放了一遍视频。短视频采用自动推送的机制，视频单个被推送到用户界面，用户如果对视频没兴趣则会快速划到下一个视频，在本实验中认为视频完播等同于视频用户点击视频。在本实验中，如果watch\_ratio1，认为用户对视频满意，相应的收益是1，否则是0。

本实验模拟在线情况下对用户进行交互式推荐，且考虑时间因素的影响，small matric数据集中用户和视频交互信息的跨度为两个月，交互信息的时间于2020.07.05-2020.09.05之间，且存在互动时间缺失的数据，选择small matric中互动时间为2020.07.05-2020.0715的六十万条数据作为训练集，交互时间在2020.0716-2020.09.05的数据和交互时间缺失的数据作为测试集。

从small matric数据集中提取：用户集，视频集。

从训练集中提取出各个用户的已被推荐视频列表与对应的交互时间，经过对“watch\_ratio”处理得到各个用户的喜爱视频集，每个视频对应的用户集。

从item\_categories提取出每个视频对应的标签集。

在本次实验中选择Python为开发语言。

## 4.2 实验设置

### 4.2.1 实验算法与对照算法

为验证本文提出的基于多臂赌博机的多样性优化算法在提高列表多样性和时间多样性方面的有效性，并进行参数探索，寻找使多臂赌博机在开发与探索问题上表现最好的探索概率，设置以下算法进行实验。

表4-1 实验算法与对照算法

| 算法名称 | 参数 |
| --- | --- |
| 基于用户的协同过滤算法（User CF） | 无 |
| 本文算法(MAB0.05) | 初始权重 |
| 本文算法（MAB0.20） | 初始权重 |
| 本文算法（MAB0.35） | 初始权重 |
| 固定权重模型（M1） |  |
| 固定权重模型（M2） |  |
| 固定权重模型（M3） |  |
| 固定权重模型（M4） |  |

因为基于多臂赌博机的多样性优化算法是在基于用户的协同过滤推荐算法形成的Top-100候选视频集中挑选视频，目的是提高推荐的列表多样性和时间多样性，所以选择基于用户的协同过滤推荐算法（User CF）作为对照算法。

基于多臂赌博机的多样性优化推荐算法与固定权重模型相比，不同点在于它对用户的指标权重进行动态更新，为了进一步判断能够动态更新指标的本文算法相比于固定权重模型的效果，如表4-1增加不同权重分布的固定权重模型作为对照算法。固定权重模型的探索概率为0.2，缺少基于多臂赌博机的多样性优化算法进行权重更新的环节，其他环节与其相同。

本文使用贪婪算法的多臂赌博机模型，因为不同的探索概率会如何影响算法在准确率，多样性和时间多样性上的表现需要进一步探索，以找到最优参数，同时需要观察不同探索概率下，权重更新的情况是否会存在不同，所以如表4-1，设计了三种不同探索概率的基于多臂赌博机的多样性优化推荐算法进行实验。

### 4.2.2 评价指标

#### （1）准确率

在实验过程中每个算法进行50轮推荐，每轮对随机抽取的同一批100名用户进行一次推荐，一个推荐列表包含5个视频，用准确率来衡量推荐系统满足用户个性化兴趣的性能。参考文献[30]中对准确率的定义，本实验在第轮进行推荐后，计算用户截止到第轮所有推送内容的平均点击率作为算法为用户推荐的准确率，并将参与测试的100位用户的准确率进行平均，作为前轮推荐时，推荐系统的推荐准确率。

 ( SEQ MTChap \c \\* Arabic \\* MERGEFORMAT 4- SEQ MTEqn \c \\* Arabic \\* MERGEFORMAT 1)

表示截止到第轮推荐，推荐系统为用户推荐的视频的视频集，表示截止到第轮推荐，用户喜欢的视频的视频集。

#### （2）列表多样性

本文采用列表多样性，即推荐系统为用户构建的推荐列表的多样性作为推荐算法内容多样性的评价指标，计算前轮推荐时，用户的平均列表多样性，再由用户多样性的平均值衡量推荐系统列表多样性指标上的性能。

推荐算法为用户在第轮形成的推荐列表的列表相似性为列表内视频的平均相似度：

 ( SEQ MTChap \c \\* Arabic \\* MERGEFORMAT 4- SEQ MTEqn \c \\* Arabic \\* MERGEFORMAT 2)

用户u在截止第轮推荐的列表多样性的平均得分可用表示。

 ( SEQ MTChap \c \\* Arabic \\* MERGEFORMAT 4- SEQ MTEqn \c \\* Arabic \\* MERGEFORMAT 3)

推荐系统在前轮推荐的列表多样性得分为所有测试用户的列表多样性得分的平均值，可用式表示。

 ( SEQ MTChap \c \\* Arabic \\* MERGEFORMAT 4- SEQ MTEqn \c \\* Arabic \\* MERGEFORMAT 4)

#### （3）时间多样性

在3.3.1节中，本文已经介绍了新的时间多样性指标的概念，在实验效果评估阶段，定义用户在第轮推荐时的时间多样性分数为：截止第轮推荐时，所有推荐视频的平均时间多样度，具体定义如式。

 ( SEQ MTChap \c \\* Arabic \\* MERGEFORMAT 4- SEQ MTEqn \c \\* Arabic \\* MERGEFORMAT 5)

推荐系统在第轮推荐的时间多样性得分为所有测试用户平均时间多样性得分的平均值，具体定义如式。

 ( SEQ MTChap \c \\* Arabic \\* MERGEFORMAT 4- SEQ MTEqn \c \\* Arabic \\* MERGEFORMAT 6)

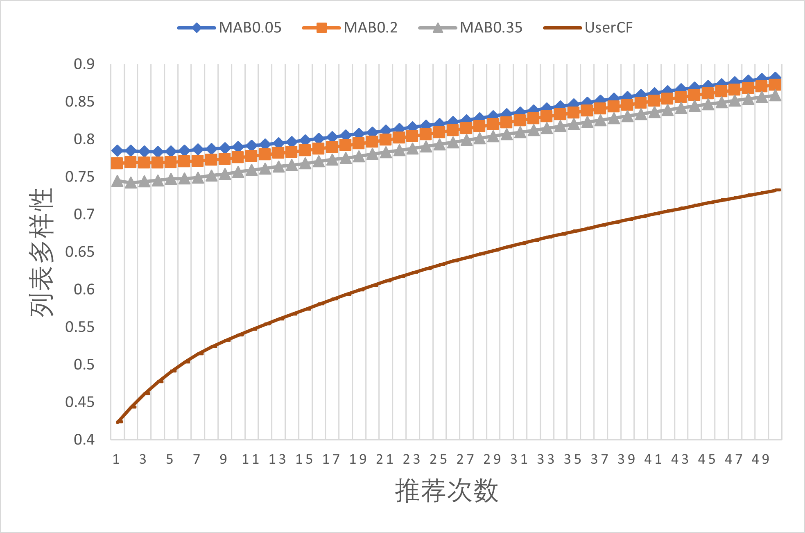
## 4.3 实验结果与分析

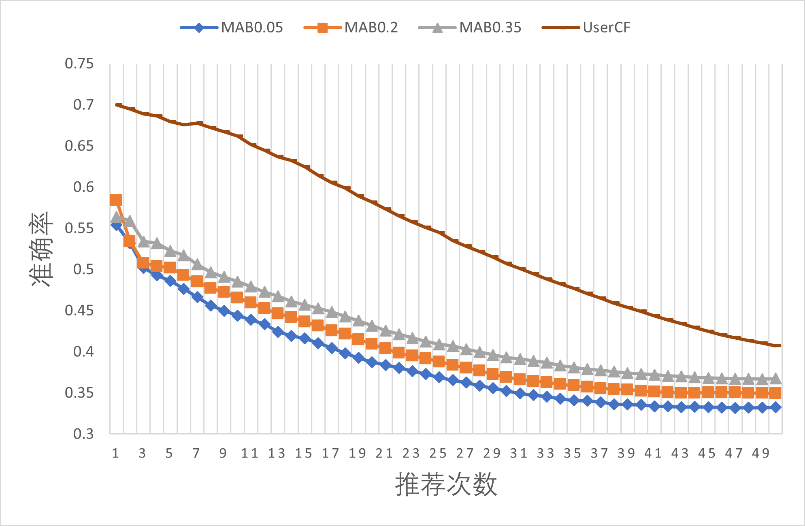
### 4.3.1 算法效果评估

本节根据设定指标分值计算方式评估推荐算法在准确度，列表多样性，时间多样性指标的表现，对不同算法的运行结果进行分析对比。

#### （1）MAB0.05/0.2/0.35与User CF对比分析

MAB0.05/0.2/0.35与User CF算法在准确率，列表多样性和时间多样性指标上的表现如图4‑1所示。在准确率的表现上，User CF显著高于MAB0.05/0.2/0.35，说明加入列表多样性和时间多样性指标构建多目标函数，从协同过滤形成的Top-N视频集中挑选视频，将不可避免地导致整体准确率的损失。在列表多样性的表现上，User CF显著低于MAB0.05/0.2/0.35，说明MAB0.05/0.2/0.35确实可以有效提高推荐列表的多样性。算法在时间多样性的表现与在列表多样性上的表现基本一致，User CF显著低于MAB0.35/0.2/0.05，说明MAB0.35/0.2/0.05确实可以有效提高推荐内容的时间多样性。





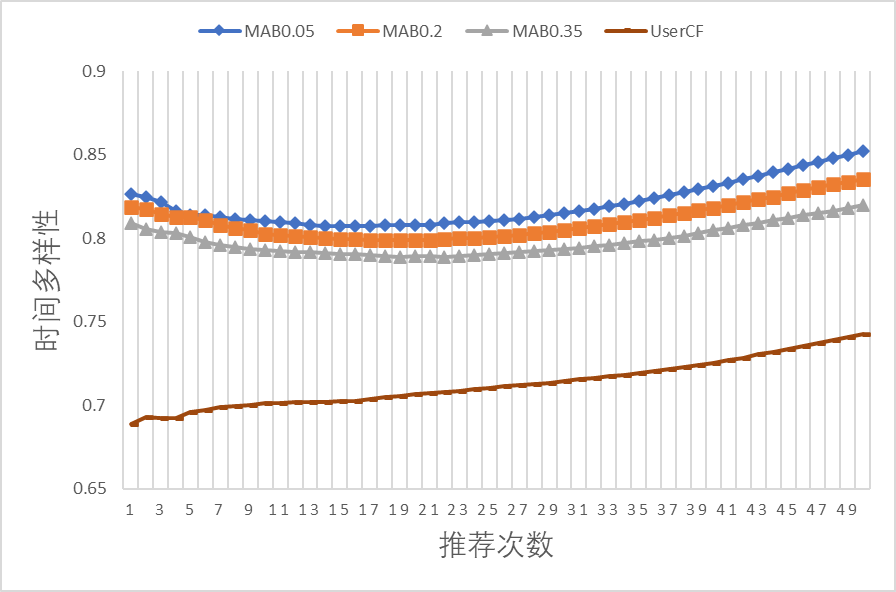


图 4‑1 MAB0.05/0.2/0.35与User CF对比

#### （2）MAB0.2与固定权重模型对比分析

将探索概率相同的MAB0.2与固定权重模型在准确率，列表多样性和时间多样性指标上的表现如图4‑2所示。除了准确度权重较高的M2在准确率上明显优于MAB0.2，MAB0.2优于M3、M4，与M1表现相近。从MAB0.2与固定权重模型在列表多样性上的表现来看，MAB0.2表现出较高的列表多样性，稍稍劣于列表多样性权重较高的M3，在列表多样性上的表现与M1接近，优于除M3外的其他固定权重模型。从MAB0.2与固定权重模型在时间多样性上的表现来看，MAB0.2与M1，M3较为接近，劣于时间多样性指标权重较高的M4，明显优于M2。

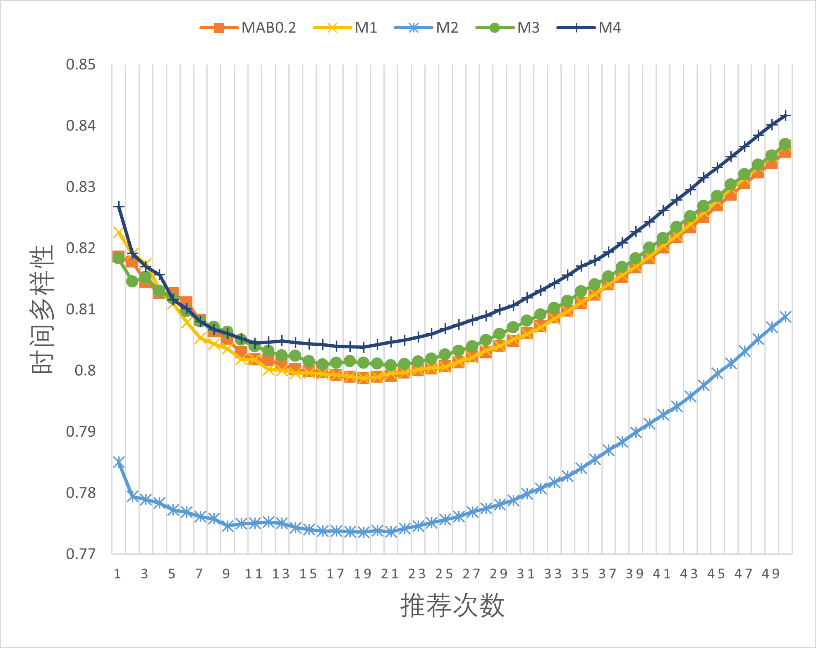
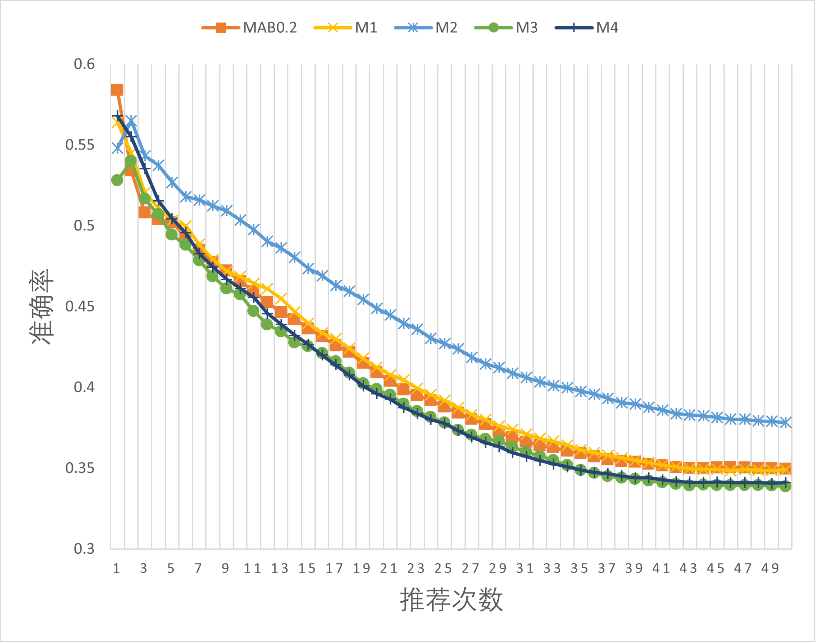
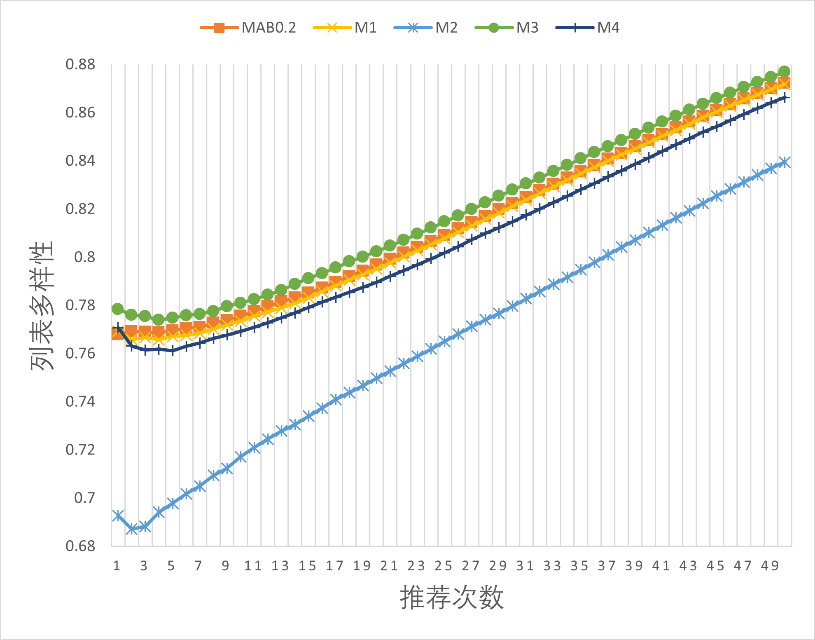


图 4‑2 MAB0.2与固定权重模型对比

总的来说，固定权重模型因指标权重存在较大差异，在不同指标上表现差别很大，M2因为准确率权重较高，在准确率方面表现突出，但在列表多样性和时间多样性上表现很差；M3在权重较高的指标上未表现出明显优势；M4在时间多样性指标上表现突出，但在准确率和列表多样性上表现略差；M1与能够进行权重更新的MAB0.2综合表现最好。固定权重模型和本文算法都能有效提高推荐列表多样性和时间多样性，但是固定权重模型中，各指标权重是固定的，并不考虑用户偏好的差异，某些指标上的损失可能不是用户期望的，会损害用户体验。但在MAB0.35/0.2/0.05中，虽然准确率和多样性指标的收益仍然是冲突的，但按照用户偏好设置权重，能较好地达到三个指标间的平衡。

因此，可以认为，基于多臂赌博机的多样性优化算法能够达到提高多样性和时间多样性的效果，满足用户在准确度，多样性和时间多样性上的个性化偏好需求。

### 4.3.2 参数探索

在图4‑1中，不同探索概率的算法在不同指标的表现上表现出明显差异。MAB0.5在列表多样性和时间多样性上表现较好，MAB0.35在准确率上表现较好。总的来说，由于算法是在协同过滤算法形成的Top-N候选集中进行挑选，当探索概率更大时，算法的准确率指标表现更好，当探索概率更小时，算法的多样性指标表现更好。而综合考虑算法在三个指标上的表现，MAB0.2在三个指标上的综合表现最好，能较好地平衡多样性和准确率，因此，本实验认为，探索概率为0.2是基于多臂赌博机的多样性优化算法的最优参数。

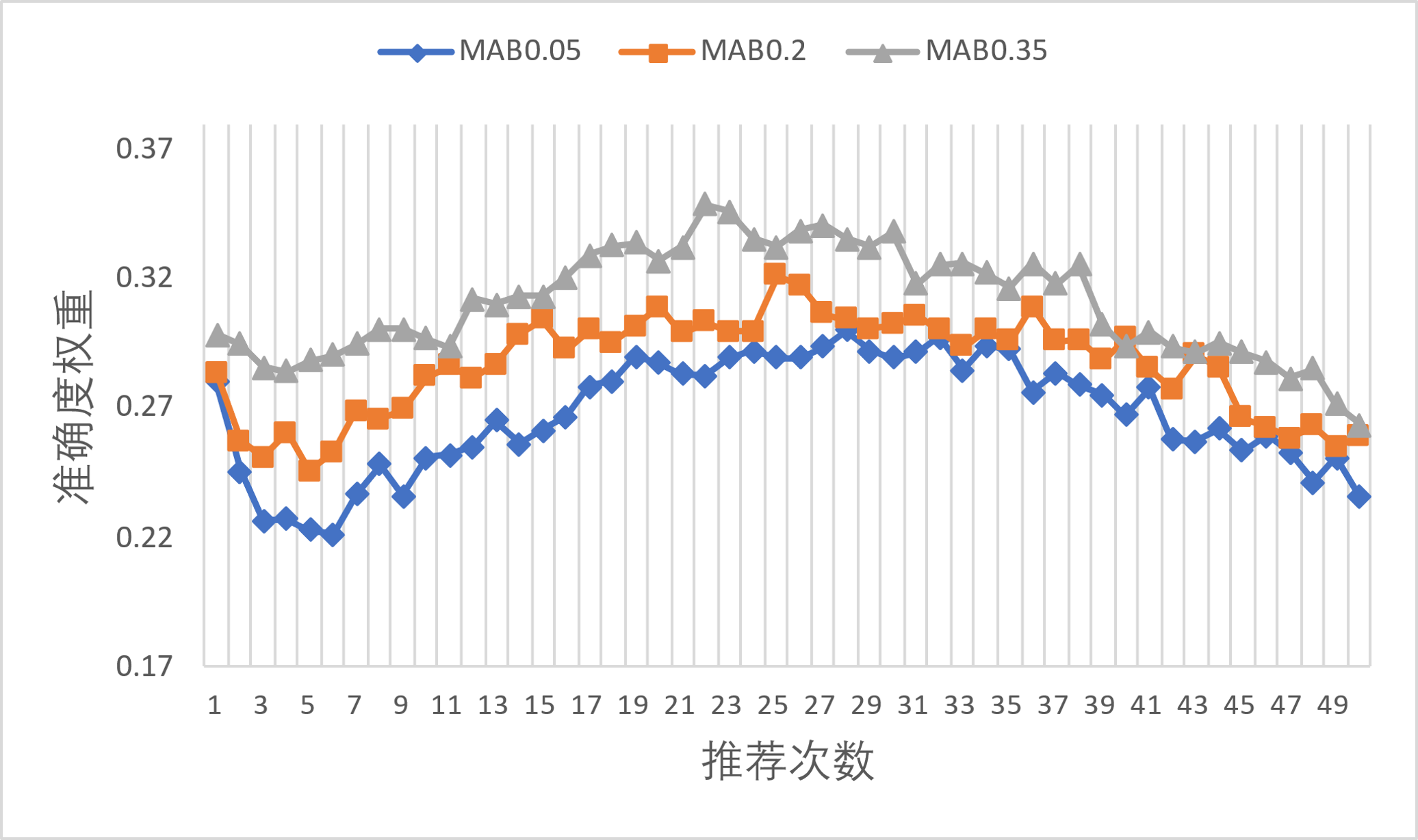
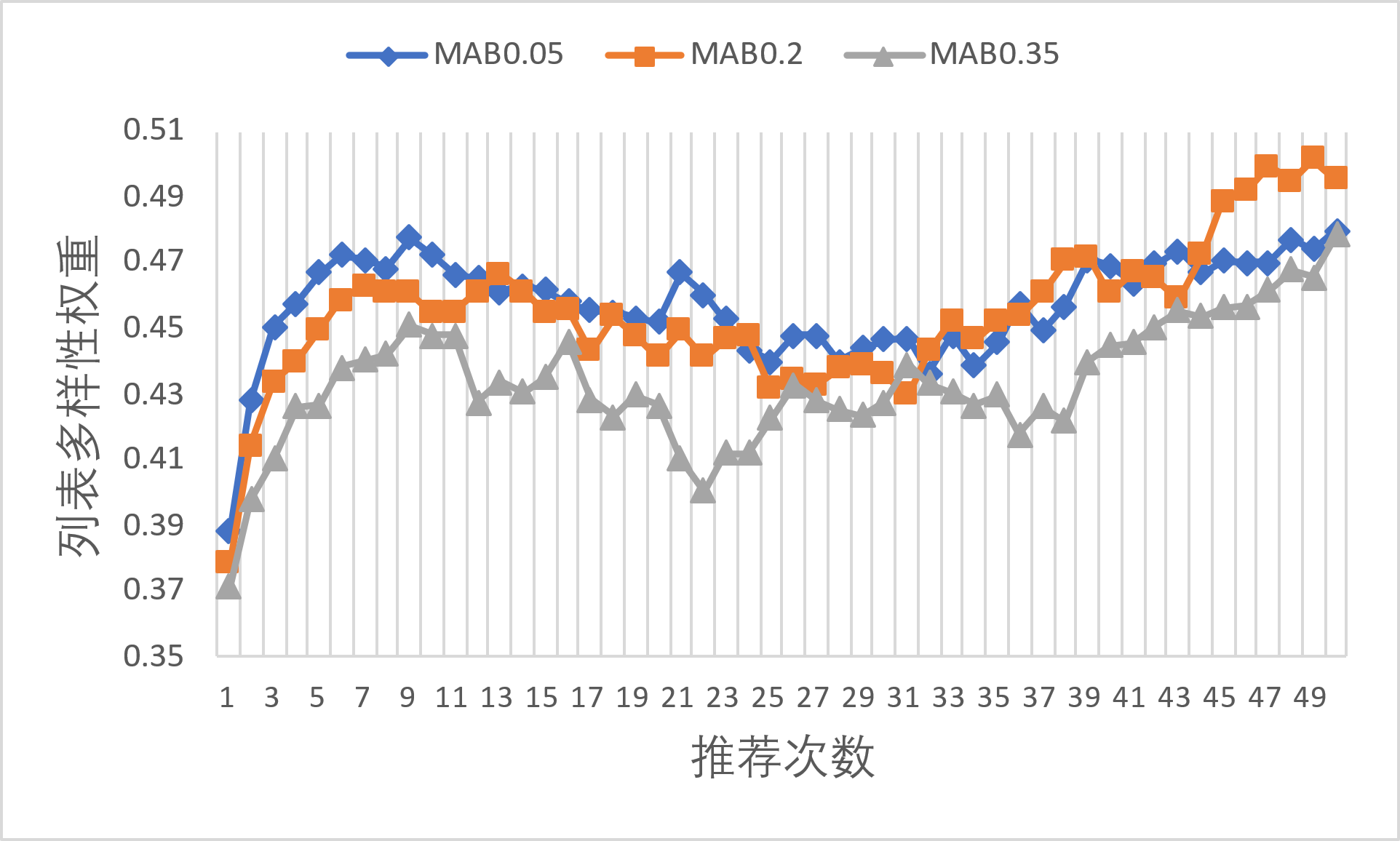
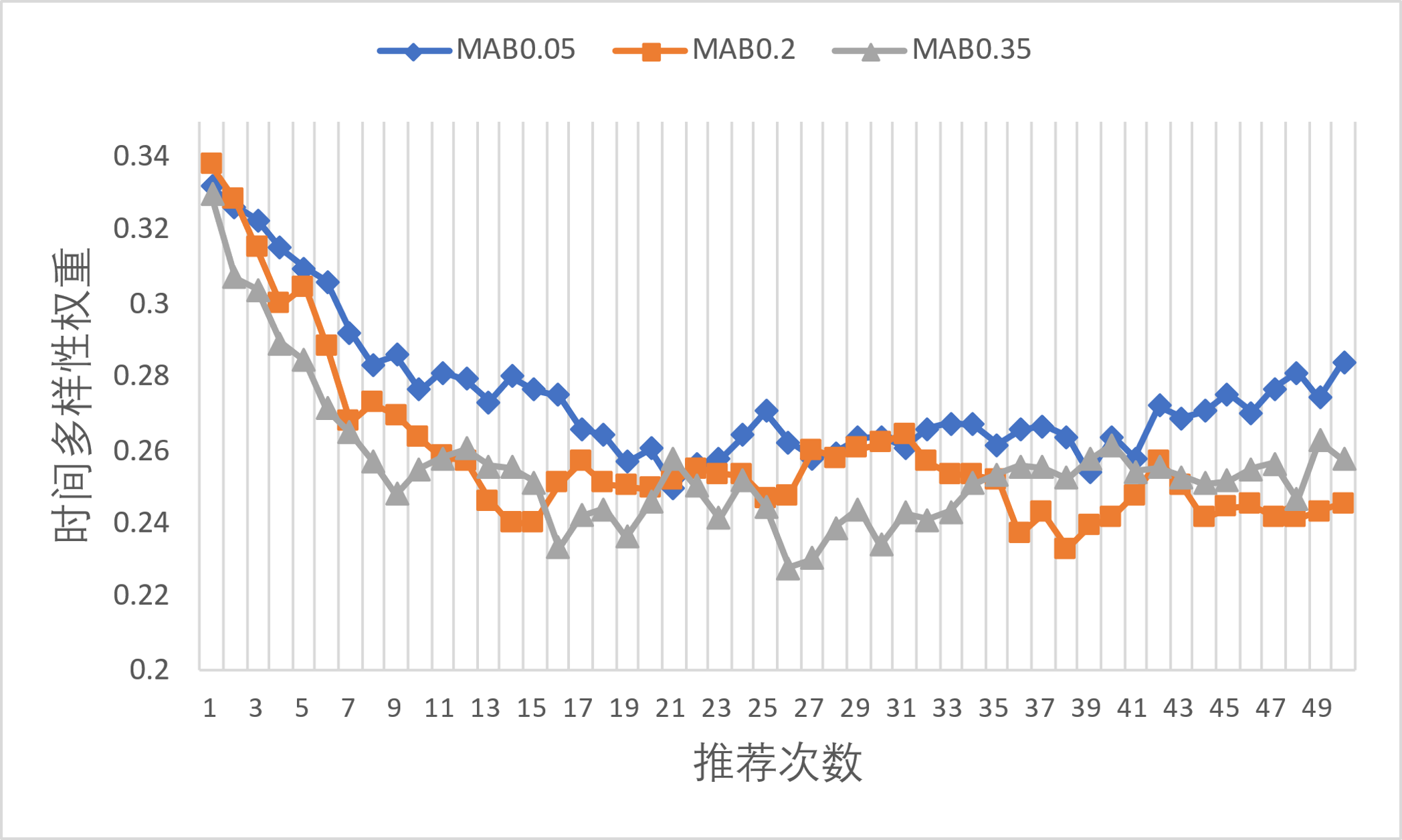
除了参数对算法在不同指标上的表现效果存在影响，接下来进一步探索参数在权重更新上的影响。

图 4‑3 不同参数下平均权重变化趋势

不同参数的基于多臂赌博机的多样性优化算法的用户平均准确度权重，平均列表多样性权重和平均时间多样性权重随推荐次数的变化趋势如图4‑3。探索概率为0.05, 0.2, 0.35时，用户的平均准确率权重，列表多样性权重和时间多样性权重都相差不大，且呈现出了相似的变化趋势。说明不同的探索参数不会改变用户的权重更新情况，用户的权重仍然是由用户的个性化偏好决定。

### 4.3.3 用户偏好分析

使用MAB0.2算法对100位用户进行50轮推荐，进行最后一次权重更新后，得到不同用户的指标权重如图4-1。

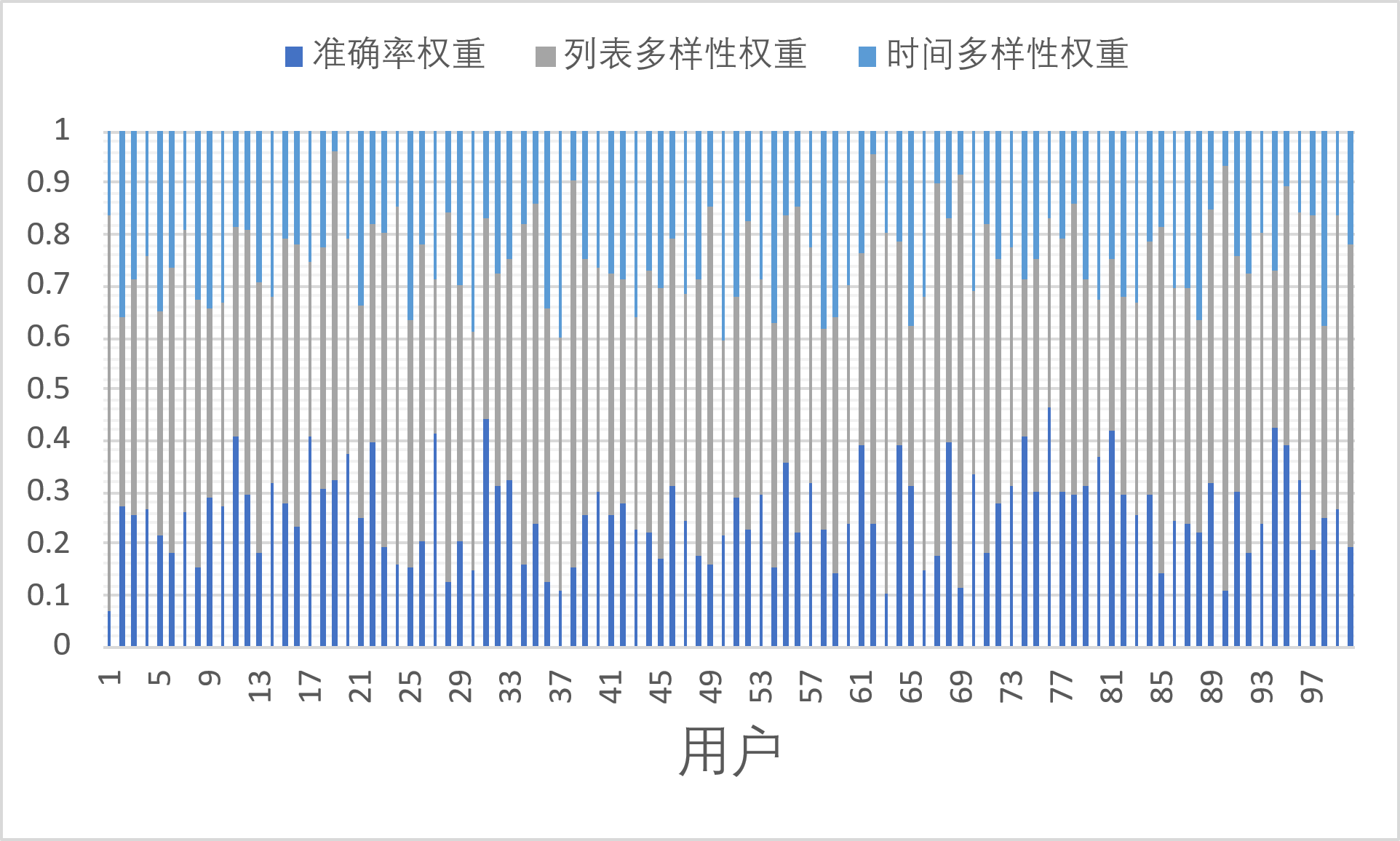


图4-1 用户指标权重

从图4-1可以看出大多数用户在准确度，列表多样性和时间多样性权重上存在很高的相似性，但仍存在个体区别，且有少数用户在某指标上表现出极大的偏好，与其他用户表现出了很大的差异性。

用户在指标偏好上表现出的区别证明了基于多臂赌博机的多样性推荐算法进行权重更新的意义与价值，固定的指标权重无法满足用户的个性化多样性需求，单一地提高推荐系统在某指标上的表现不一定能提高所有用户的使用体验，基于用户在指标上的不同偏好进行推荐更能满足用户的个性化需求。

# 5 总结与展望

本文设计了一种适用于视频平台的基于多臂赌博机的多样性优化算法，将列表多样性指标和时间多样性指标引入多臂赌博机挑选视频时构建的遗憾值中，而让多臂赌博机不再只是将准确度作为评价标准，提高了推荐算法在列表多样性和时间多样性上的表现，为用户形成兼顾准确度，多样性的视频推荐列表。

为多指标的权重模型设置权重更新策略，利用视频推荐列表的收益集合，与用户对视频推荐列表的反馈，动态更新权重，并应用于下一次的推荐之中，在多次迭代后，权重会逐渐逼近于用户对各指标的真实偏好，使推荐列表满足用户的个性化多样性需求。如果用户对指标的偏好发生动态改变，权重模型也会随之更新，满足用户在不同时期的个性化推荐需求。

在准确度，列表多样性，时间多样性三个指标中，使用视频标签的平均点击率表示视频在准确度指标上的收益，时间多样性考虑推荐内容在时间上的多样性，但在评价候选视频在时间多样性上的表现时，仅仅用候选视频与往期视频的相似度来评价，忽略了时间导致的用户对旧视频的印象减弱，候选视频与不同时间被推荐给用户的视频的相似度在影响用户时间多样性的使用体验上的效果是不同的。因此引入时间衰减函数，与相似度结合评价候选视频在时间多样性上的表现。

本文使用用户对视频所含标签的平均点击率来估算用户对未观看过的视频的期望点击率，将多臂赌博机探索视频收益的问题转化为多臂赌博机探索标签收益的问题，避免了重复推荐为用户带来的不良体验。但是基于用户创作分享的中短视频平台视频标签的添加带有随意性，由视频创作者自主添加，很难达到完全准确描述视频内容的效果，视频平台在用户上传视频时，会给出标签推荐建议，标签推荐的效果也会影响后期利用用户对标签喜好来估计用户对视频喜好的准确度。同时因为权重更新策略需要准确度上的期望收益，当准确度上的期望收益不准确时，也会影响权重更新。

此外，用户端创作的标签池也是数量庞大且关系复杂的，标签之间具有复杂结构关系，上下级标签如何联动更新需要进一步研究，标签收益可能并不是线性补偿的，某些标签的结合可能给用户带来的效果，同时，用户对某些标签的恶感可能很难被其他标签抵消掉。如何挖掘标签池中标签的特定关系，并在使用标签描述视频时考虑这些关系，是非常值得研究的。

在实验过程中，使用了列表多样性指标和时间多样性指标来评价算法的提高列表多样性和时间多样性的效果，但实际上，用户对于多样性的体验是直接的，只有通过对用户的直接调查才能作出更为准确的判断。多样性对于用户使用体验的影响更多的是影响用户的使用时间等，在多样性和准确度达到平衡的情况下，用户对单次推荐列表可能点击率一直不高，但可以通过频繁刷新，来得到每次推荐列表中含有的个别非常感兴趣的视频，来延长用户的使用时间，在使用数据集进行实验的过程中，无法对多样性和时间多样性提高带来的这一效果进行验证，因此，在真实场景下的实验更有说服力。

在本文中使用了简单的-贪心算法，多臂赌博机算法进行探索和开发的概率是确定的，但在后期，模型对用户的拟合效果较好时，这种固定概率的探索可能会带来很大的累计遗憾值的损失。如何按照模型对用户的拟合效果来更改后期进行探索的概率，使探索概率随算法效果进行适应性改变也是有研究意义的。

参考文献

[1] Bobadilla J, Ortega F, Hernando A, Gutiérrez A. Recommender systems survey[J]. Knowledge-based systems, 2013, 46: 109-132.

[2] Castells P, Wang J, Lara R, Zhang D. Workshop on novelty and diversity in recommender systems-DiveRS 2011[C] // Proceedings of the fifth ACM conference on Recommender systems. 393-394.

[3] Resnick P, Varian H R. Recommender systems[J]. Communications of the ACM, 1997, 40(3): 56-58.

[4] Lathia N, Hailes S, Capra L, Amatriain X. Temporal diversity in recommender systems[C] // Proceedings of the 33rd international ACM SIGIR conference on Research and development in information retrieval. 210-217.

[5] Ziegler C-N, Mcnee S M, Konstan J A, Lausen G. Improving recommendation lists through topic diversification[C] // Proceedings of the 14th international conference on World Wide Web. 22-32.

[6] Zhang M, Hurley N. Novel item recommendation by user profile partitioning[C] // 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology. IEEE, 1: 508-515.

[7] Adomavicius G, Kwon Y. Improving aggregate recommendation diversity using ranking-based techniques[J]. IEEE Transactions on Knowledge and Data Engineering, 2011, 24(5): 896-911.

[8] Boim R, Milo T, Novgorodov S. Diversification and refinement in collaborative filtering recommender[C] // Proceedings of the 20th ACM international conference on Information and knowledge management. 739-744.

[9] Javari A, Jalili M. A probabilistic model to resolve diversity–accuracy challenge of recommendation systems[J]. Knowledge and Information Systems, 2015, 44(3): 609-627.

[10] Cheng P, Wang S, Ma J, Sun J, Xiong H. Learning to recommend accurate and diverse items[C] // Proceedings of the 26th international conference on World Wide Web. 183-192.

[11] Li S, Zhou Y, Zhang D, Zhang Y, Lan X. Learning to diversify recommendations based on matrix factorization[C] // 2017 IEEE 15th Intl Conf on Dependable, Autonomic and Secure Computing, 15th Intl Conf on Pervasive Intelligence and Computing, 3rd Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress (DASC/PiCom/DataCom/CyberSciTech). IEEE: 68-74.

[12] Zheng G, Zhang F, Zheng Z, Xiang Y, Yuan N J, Xie X, Li Z. DRN: A deep reinforcement learning framework for news recommendation[C] // Proceedings of the 2018 World Wide Web Conference. 167-176.

[13] Qin L, Chen S, Zhu X. Contextual combinatorial bandit and its application on diversified online recommendation[C] // Proceedings of the 2014 SIAM International Conference on Data Mining. SIAM: 461-469.

[14] 帖军, 孙荣苑, 孙翀, 郑禄. MOOB: 一种改进的基于Bandit模型的推荐算法[J]. 中南民族大学学报(自然科学版), 2018, 37(01): 114-119.

[15] Hidayatullah A, Anugerah M A. A Recommender System for E-Commerce Using Multi-objective Ranked Bandits Algorithm[C] // 2018 International Conference on Computing, Engineering, and Design (ICCED). IEEE: 170-174.

[16] 何炜俊, 艾丹祥. 以多臂赌博机建模的多目标互动式推荐系统[J]. 小型微型计算机系统, 2021, 42(06): 1192-1198.

[17] 张艳红, 张春光, 周湘贞, 王怡鸥. 项目多属性模糊联合的多样性视频推荐算法[J]. 计算机科学, 2019, 46(8): 78-83.

[18] Goldberg D, Nichols D, Oki B M, Terry D. Using collaborative filtering to weave an information tapestry[J]. Communications of the ACM, 1992, 35(12): 61-70.

[19] Sarwar B, Karypis G, Konstan J, Riedl J. Item-based collaborative filtering recommendation algorithms[C] // Proceedings of the 10th international conference on World Wide Web. 285-295.

[20] 何羽丰. 基于用户聚类和多臂赌博机算法的微博推荐模型[D]. 河北大学, 2021. ·doi: 10.27103/d. cnki. ghebu. 2021.001855.

[21] Li L, Chu W, Langford J, Schapire R E. A contextual-bandit approach to personalized news article recommendation[C] // Proceedings of the 19th international conference on World wide web. 661-670.

[22] 章晓芳, 周倩, 梁斌, 徐进. 一种自适应的多臂赌博机算法[J]. 计算机研究与发展, 2019, 56(3): 643.

[23] Li S, Karatzoglou A, Gentile C. Collaborative filtering bandits[C] // Proceedings of the 39th International ACM SIGIR conference on Research and Development in Information Retrieval. 539-548.

[24] Gentile C, Li S, Zappella G. Online clustering of bandits[C] // International Conference on Machine Learning. PMLR: 757-765.

[25] 王宇琛, 王宝亮, 侯永宏. 融合协同过滤与上下文信息的Bandits推荐算法[J]. 计算机科学与探索, 2019, 13(3): 361-373.

[26] 陈珂. 基于多臂赌博机算法的推荐系统研究[J]. 长江信息通信, 2021.

[27] Agrawal S, Goyal N. Analysis of thompson sampling for the multi-armed bandit problem[C] // Conference on learning theory. JMLR Workshop and Conference Proceedings: 39.31-39.26.

[28] 徐元萍, 陈翔. 推荐系统中的新颖性问题研究[J]. 计算机应用研究, 2020, 37(08): 2310-2314.

[29] Gao C, Li S, Lei W, Li B, Jiang P, Chen J, He X, Mao J, Chua T-S. KuaiRec: A Fully-observed Dataset for Recommender Systems[J]. arXiv preprint arXiv: 2202.10842, 2022.

[30] 厍向阳, 蔡院强, 董立红. 新的基于多目标优化的推荐算法[J]. 计算机应用, 2015, 35(01): 162-166.

致 谢

2018年的9月，还记得那个炎热的夏天，我的家人带着我从河南经过了6个多小时的车程来到了华中科技大学，那个时候我还认不清华科方方正正，两边种着梧桐树的道路，在堵塞的校园道路上缓慢行驶，想象着自己未来的大学生活。而转眼间，从18年到22年，我已在华科生活了四年。

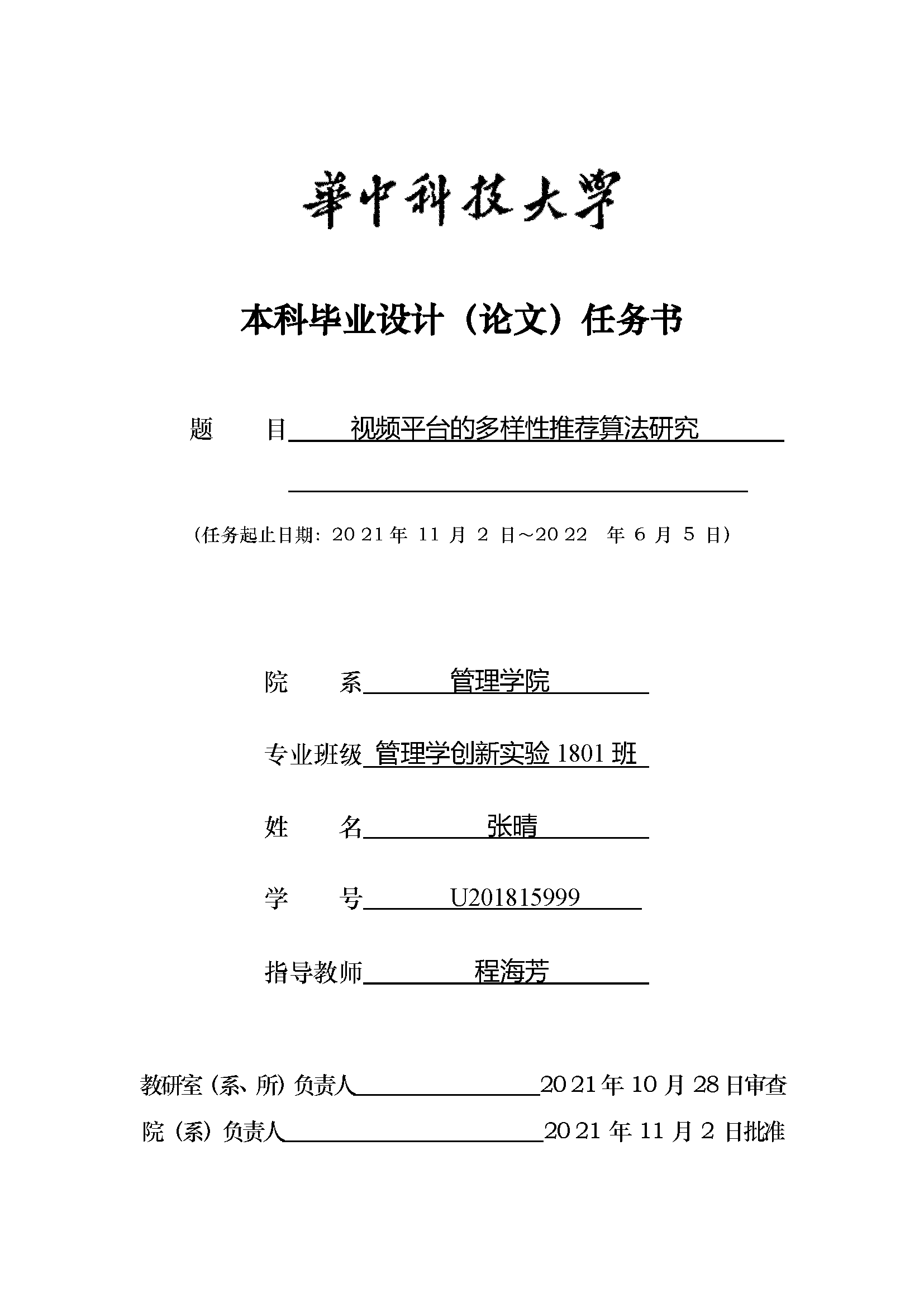
从A栋到D栋，不知道在东九楼多少个教室里上过课，上自习，还有许多次考试的前几天，在东九走廊上和许多同学一起背书。亲眼见证了从韵苑到东九那条一到上课与下课就异常拥挤的道路在四年里的改变，从最开始的路上人头攒动，到改造道路后也有了机动车行驶的空间。只是没办法再见证道路两旁新栽的梧桐树苗慢慢长大。经过了三个多月的学习和研究，完成我的毕业论文之后，就要与华科说再见了。

首先，我要感谢我的导师程海芳老师，从立题到最后的论文结构与内容修改，在我写毕业论文的每一个阶段都给了我很多建议，在我有不懂的问题时，程老师也会耐心为我解答，无论什么时候，程老师都是那么温柔和耐心地回复我。在我遇到瓶颈时，程老师也和我一起解决问题。正是因为程老师的帮助，我才顺利地完成了我的毕业论文。

其次，我要感谢的是我的学院管理学院，让我有机会在这里学习，让我在本科期间遇到如此多优秀的老师，有了和这些老师交流，向他们学习的机会；让我遇到如此多可爱的同学，和他们一起学习，得到了许多珍贵的朋友；还要感谢管院，对于本科生的细致帮助，尤其是我们的辅导员，在我们的大学四年中，给了很多及时的建议，帮助我们处理了很多事情；正是因为管院，我才有机会感受主校区春夏季绵里藏针的梧桐絮，感受武汉变幻莫测的天气，过往在华科，在管院遇到的每一件事情都是如此的珍贵，让我如此的不舍。

再其次，我要感谢我的父母，他们是我在学习与生活中最大的依靠，是他们的辛苦工作让我得到在经济上的支持，可以完成我的学业。在我最迷茫，最失落的时候，在我处于人生低谷的时候，我知道，我的父母会永远在后面支持我，给我鼓励，为我担心，他们永远是我最大的依靠。我要感谢本科期间遇到的朋友们，正是因为有他们的存在，这四年的大学时光才会如此快乐。我还要感谢国家助学金的帮助，在入学之前，我就申请了生源地贷款，这笔贷款缴纳了我读书四年的学费，减轻了我读书期间家里的经济压力。

毕业论文完成，我的大学生活也即将结束，我想祝我遇到的每一个人事业学业生活顺利，希望我们都能变得越来越好。我们不知道未来的生活会怎样，但我相信，所有的难关都将过去，未来仍然有无限可能！



|  |
| --- |
| 课题内容：  （1）研究个性化推荐算法及多样性推荐算法的优化方法。了解其相关的关键技术以及其应用领域；了解目前在推荐算法多样性优化方面的最新进展，了解其关键算法和原理。  （2）以B站作为具体研究对象，对B站推荐页面使用体验进行调查，对B站推荐页面现状进行分析，分析B站推荐系统目前存在的不足，并根据B站推荐页面的不足，选择确定的推荐算法进行研究，找到适用于视频网站的推荐系统多样性改进方法。  （3）针对B站推荐系统，寻找最适合B站的推荐系统多样性优化方案，设计提升推荐内容多样化和符合用户兴趣的推荐算法，考虑引入用户行为特征作为推荐系统反馈，基于用户的协同过滤与多臂赌博机融合进行交互式推荐算法设计。  （4）选择数据集，在数据上对改进后的算法与原算法进行实验，得出实验数据并分析，验证改进算法的效果。 |
| 课题任务要求：  1．毕业论文资料调研。要围绕选题搜集、阅读有关中英文文献资料。  2．根据指导老师给出的英文原文，完成英文翻译。所完成的译文要求，内容忠于原文，逻辑清楚，语句通顺。  3．撰写开题报告，并及时交由指导教师审定开题报告。  4.撰写毕业论文大纲和详细提纲。论文提纲应分为几个部分或几个层次。写明论文的中心、重点、  主要观点、结论等。  5．撰写论文，反复修改。写作过程中要继续搜集、补充资料，写作要层次分明，条理清楚，观点  明确，论证有理有据，具有说服能力。文章的文字要简洁、通顺、流畅、无错别字。凡引用文字、数据注明出处。以期达到学校要求的规范论文格式及内容要求。  6．将论文按规定格式打印定稿，递交指导教师。  7.按统一要求进行答辩准备。  8.毕业答辩。 |
| 主要参考文献（由指导教师选定）：  [1]多样性推荐算法的定义及优化方法  [2] Qi T , Wu F , Wu C , et al. HieRec: Hierarchical User Interest Modeling for Personalized News Recommendation[J]. 2021.  [3] Qi T , Wu F , Wu C , et al. PP-Rec: News Recommendation with Personalized User Interest and Time-aware News Popularity[J]. 2021. |
| 同组设计者： |
| 指导教师签名：  2022年 2 月 21 日 |