



## 硕士学位论文

关于在线视频系统中网络资源部署及广告投放的研究

Research on the Deployment of Network Resources and  
Advertisement Insertion in Online Video System

作者：张琛玥

导师：李纯喜

北京交通大学

2021年6月

## 学位论文版权使用授权书

本学位论文作者完全了解北京交通大学有关保留、使用学位论文的规定。特授权北京交通大学可以将学位论文的全部或部分内容编入有关数据库进行检索，提供阅览服务，并采用影印、缩印或扫描等复制手段保存、汇编以供查阅和借阅。同意学校向国家有关部门或机构送交论文的复印件和磁盘。学校可以为存在馆际合作关系的兄弟高校用户提供文献传递服务和交换服务。

(保密的学位论文在解密后适用本授权说明)

学位论文作者签名：

张课桐

导师签名：

李健

签字日期：2021年6月1日

签字日期：2021年6月1日

学校代码：10004

密级：公开

# 北京交通大学

## 硕士学位论文

关于在线视频系统中网络资源部署及广告投放的研究

Research on the Deployment of Network Resources and  
Advertisement Insertion in Online Video System

作者姓名：张琛玥

学 号：18120172

导师姓名：李纯喜

职 称：副教授

学位类别：工学

学位级别：硕士

学科专业：通信与信息系统

研究方向：信息网络

北京交通大学

2021年6月

## 致谢

本论文是在导师李纯喜老师的悉心指导，家人的默默支持，一众朋友的鼓励陪伴下完成的。在此我想对完成论文工作期间以及三年读研期间所有给予我帮助和磨砺的人和事表示感谢和感恩。

这三年来，最大的荣幸莫过于可以在交大，在有线组度过三年的研究生时光。在这里，我遇到了亦师亦友的老师，李纯喜、赵永祥和郑宏云等老师对我的帮助不仅是在学术方面悉心的指导，更有生活和成长方面的指引。在读研期间，非常感谢李纯喜老师的信任和鼓励，使得我有机会可以参加学校的国际交换项目，感受国内外校园生活的异同，经历不同的求学体验，拥有更加丰富的学习生活感受。这三年里，我收获的不仅是科研工作能力，更是一份面对生活的淡然和从容。

在有线组，最大的幸运是认识了实验室的一众好友，在这三年中的每一个阶段，都有他们的陪伴和鼓励，尤其是完成最终的论文期间，感谢高志朋学长、李小乐、张虎信、李文雯、赵映南和张庆贺的陪伴，九教南 309 的时光将是研究生记忆中最难忘、最充实的一部分。

求学十九载，最感恩的是父母家人的支持，无论是在国内的求学之路，亦或是在英国的留学历程，都离不开父母为我创造的学习条件，离不开家人给予我的关爱和支持，尤其是在 2020 年疫情期间，家人的鼓励和支持给了我面对种种困难的勇气，正是他们的理解和陪伴使得我可以专心顺利的完成学业。

这三年的研究生生活因为一场疫情显得较为特殊，因为选择出国交换变的更加与众不同。感谢交大为我们提供的学习环境以及交流机会，使得我不仅有会接触前沿的科研内容，更有机会拥有异国的学习体验；感恩祖国的强大与努力，让我可以在疫情期间顺利的回到校园，让研究生生活不因疫情的影响留下太多遗憾。

此外还要诚挚的感谢国家自然科学基金（No. 61872031，基于熵理论的信息匹配网络测量与建模）的资助。

## 摘要

在线视频已经成为当前互联网上主流的产业系统，占据绝大部分网络流量。在这个产业系统中，内容商通过租用带宽和缓存向用户提供视频内容进而赚取用户（订阅）收益、通过投放视频广告赚取广告收益。如何租用缓存和带宽，以及如何投放视频广告，以最大化内容商的总体收益是视频内容商关注的热点问题。

本文在现有研究基础上，从以上两个角度分别研究网络资源租用策略及视频广告投放策略。本文首先研究最优预算分配问题，根据缓存和带宽对用户体验贡献的不同，优化预算分配以租用缓存和带宽，在考虑用户观看体验的同时，最大化内容商的用户订阅收益。其次，本文研究视频广告插入问题，根据广告投放时长和位置的对用户离开概率影响的不同，分析广告投放策略，对内容商总体广告收益的影响，从而确定最佳投放位置及时长。

本文对以上两个问题分别进行优化建模，揭示系统内各个因素对内容商收益的影响，并对每个问题进行分析和求解，以得到最优的预算分配方案和广告投放策略。具体来说，论文的主要贡献包括以下两个方面：

(1) 针对预算分配问题，本文根据带宽和缓存分别与用户观看时延之间的非线性负相关性，以及时延与用户视频观看量的负相关性，建立有预算约束的资源优化分配模型，以寻找最佳带宽和缓存分配策略来最大化内容商的视频收益。本文设计最大可行面算法，根据最优缓存和最优带宽的正相关性，求解最佳带宽和缓存分配策略。数值结果验证了在不同系统参数设置下算法的有效性。

(2) 针对广告投放问题，本文在用户观看视频的自然离开概率的基础上，联合广告的插入位置和时长对用户离开概率的影响，建立广告投放模型，以寻找最佳的广告投放时长和位置，最大化内容商的广告收益。本文对给定广告插入位置下的最优广告时长推导了理论最优解，对广告位置和时长两个因素共同作用下的广告收益进行了趋势性分析，数值分析发现了在给定用户自然离开概率和广告造成的用户离开概率共同作用下的最佳广告投放策略。

**关键词：**缓存；带宽；在线视频；资源分配；广告投放

## ABSTRACT

Online video has become the mainstream industrial system on the Internet, occupying most of the network traffic. In this industrial system, content providers rent bandwidth and cache resources from the internet service provider to provide video content to end-users. Content providers' revenue consists of the users' subscription fee and advertising revenue by inserting video advertisements. To maximize the overall revenue, how to rent cache and bandwidth, and how to insert video advertisements are important issues that content providers focus on.

Based on the existing research, in this paper, we study the network resource renting strategy and the video advertisement insertion strategy from the above two perspectives respectively. First, we study the optimal budget allocation problem of content providers. According to the different contributions of cache and bandwidth to user experience, the strategy of budget allocation to rent cache and bandwidth is optimized. The objective is to maximize the user subscription revenue of content providers with the consideration of user watching experience. Secondly, we optimize the video advertisement insertion problem. According to the different effects of ad insertion proportion and position on the probability of users leaving, we analyze the impact of ad insertion strategy on the overall advertising revenue of content providers, to determine the best advertisement insertion strategy.

In this paper, we model the above two problems as corresponding optimization problems. Then we find out the impact of various factors in the system on the revenue of content providers. We analyze and solve the problems and find the optimal budget allocation strategy and the optimal advertisement insertion strategy. Specifically, the main contributions of the dissertation include the following two aspects:

- (1) Aiming at the problem of budget allocation, we establish a budget-constrained resource optimization allocation model. The model is based on two correlations in the system. One is the non-linear negative correlation between network resources (bandwidth and cache capacity, respectively) and video transmission delay. The other one is the negative correlation between user required video data volume and the transmission delay. The objective of the model is to find out the best bandwidth and cache resource deployment strategy to maximize the overall revenue of content providers. In this paper, based on the positive correlation between the optimal cache and the optimal bandwidth, we design the feasible-plane algorithm to find out the best allocation strategy. The numerical results verify the effectiveness of

the algorithm under different system parameter settings.

(2) Aiming at the problem of ad insertion, we establish an ad insertion model to find the best ad insertion proportion and position. This model is based on the natural departure probability of the user watching behavior and combined with the influence of the ad on the users' departure probability. The objective of this model is to maximize the advertising revenue of content providers. In this paper, we derive the theoretical optimal solution for the advertising proportion under a given insertion position, and analyze the trend of advertising revenue under the combined effect of the two factors: insertion proportion and position. Numerical analysis shows the best ads insertion strategy under combined effect of the given user's natural departure probability and the probability of users leaving due to ads insertion.

**KEYWORDS:** Cache; Bandwidth; Video Delivery System; Network Resources Allocation; Advertisement Insertion

## 目录

摘要 .....	iii
ABSTRACT .....	iv
1 引言 .....	1
1.1 研究背景及意义 .....	1
1.2 国内外研究现状 .....	2
1.2.1 视频传输系统资源利用相关研究 .....	3
1.2.2 视频系统广告投放相关研究 .....	4
1.3 本文研究内容 .....	6
1.4 本文工作内容安排 .....	7
2 相关研究及技术 .....	8
2.1 典型的视频缓存传输系统 .....	8
2.1.1 系统中的主要参与对象 .....	8
2.1.2 系统中网络资源和各参与方收益 .....	8
2.2 视频系统中的广告投放 .....	10
2.2.1 视频中插入的广告形式 .....	10
2.2.2 商业广告投放策略 .....	11
2.3 在线网络视频的用户行为相关研究 .....	12
2.4 本章小结 .....	13
3 含预算约束模型的网络资源分配 .....	14
3.1 研究场景 .....	14
3.2 系统模型 .....	15
3.2.1 参与者之间的关系 .....	15
3.2.2 服务时延 .....	17
3.2.3 供需平衡 .....	17
3.2.4 预算限制 .....	18
3.3 优化建模及理论分析 .....	19
3.3.1 优化建模 .....	19
3.3.2 理论分析 .....	20
3.3.3 Feasible-plane 算法设计 .....	21
3.4 数值仿真 .....	21
3.4.1 价格函数 .....	21
3.4.2 缓存命中率函数 .....	23
3.4.3 仿真结果分析 .....	23

3.5 本章小结 .....	27
4 含广告占比及位置影响的效益优化 .....	28
4.1 研究场景 .....	28
4.2 系统模型及优化建模 .....	29
4.2.1 商业广告投放比例 .....	29
4.2.2 商业广告投放位置 .....	32
4.3 数值仿真 .....	35
4.3.1 广告在视频中占比的影响 .....	35
4.3.2 内容占比及位置联合影响 .....	40
4.4 本章小结 .....	43
5 结论 .....	44
5.1 本文工作总结 .....	44
5.2 存在问题与展望 .....	45
参考文献 .....	46
附录 A .....	51
作者简历及攻读硕士学位期间取得的研究成果 .....	52
独创性声明 .....	53
学位论文数据集 .....	54

# 1 引言

## 1.1 研究背景及意义

进入信息时代以来，互联网使得人们的生活发生了巨大的改变，从搜索引擎到网络购物以及在线医疗、教育等，互联网融入了现代人们生活的方方面面。在线视频是人们休闲娱乐主要方式的一种，截至 2020 年 12 月，中国网民网络视频及短视频应用的用户规模达 9.23 亿，使用率高达 93.7%<sup>[1]</sup>。随着互联网基础设施的不断完善和升级，用户在观看在线网络视频时，不仅希望可以观看到高质量的视频内容，同时也希望可以享受到更进一步的视频服务，如高清晰度、流畅度的视频播放过程，拥有更加舒适的观看体验，这给在线视频服务提供商提出了更高的要求和挑战。

视频内容提供商拥有大量的视频资源，通过向终端用户提供视频点播服务并在其中投放一定量等的商业广告，从而获得用户订阅费用或广告主商业推广费用。网络服务提供商拥有的网络传输和存储的基础设施，如骨干核心网链路、边缘无线基站等，为视频内容商及终端用户提供视频传输服务。视频内容商需要向互联网服务提供商支付网络基础设施资源的使用费用，得以使用其基础设施向用户提供服务。

由于在线视频的流行度服从一定的分布，少量视频内容占据了大量的请求，因此内容商为用户传输的数据中存在大量的重复传输<sup>[2,3]</sup>，这给核心网带来了很大的传输压力。为了减少重复文件在核心网的传输，同时提高网络边缘用户请求的响应速度，内容分发网络 (Content Delivery Network, CDN) 被提出并得到大规模的应用<sup>[4]</sup>。CDN 通过在核心网的关键节点及网络的边缘设立缓存服务器，通过一定的缓存规则，将一些可能重复传输的文件存储在缓存服务器中，当用户请求这些文件时，可以在网络边缘响应用户请求，缓解核心网的传输压力，同时距离用户更近的网络边缘可以更快速的响应用户请求，提高服务质量，因此缓存服务器的设置对于网络提供商、视频内容商及终端用户均是有益的。

为了向终端用户提供高质量的服务，内容商不仅需要提升其视频内容的质量，同时需要合理的购买网络服务商的资源，包括核心网传输带宽资源和边缘基站缓存服务器空间，以保证源服务器和边缘服务器可以协作高效的为用户传输视频内容。传输带宽资源决定源服务器向终端用户传输数据的能力，边缘基站缓存空间决定缓存服务器响应用户请求的概率，两者的数量均与用户感受到的服务质量呈正相关，因此为了提升用户感知的服务质量，内容商希望配置的传输系统中两类资源量均多多益善。在实际的传输系统中，资源配置是先于服务提供的，即内容商需要在开始向用户提供服务之前决定其向网络上购买的两类资源量，且决策过程

中会有一些的预算、资源最大使用量等的约束条件，这就使得内容商需要对其提供服务区域的用户需求量及购买资源的预计使用率有一定的先验认知或预测，并通过合理的配置其预算分配策略实现效益的最大化。

在线网络视频商的盈利模式有多种，包括用户订阅付费、商业推广收益以及部分单独付费视频等。用户可以通过订阅内容商视频或者付费成为会员，享受跳过商业广告、提前观看视频内容或者享受会员音效等服务。部分视频服务网站免费为用户提供视频，通过向用户播放一定量的商业广告或推广视频内容，根据内容的曝光量及播放量向推广方收取费用，用户需要观看商业广告以继续观看后续视频内容。

商业广告的投放可以为内容商带来收益，但其投放长度及位置会影响用户观看在线视频的体验，降低终端用户对视频商内容的需求量。用户对视频中广告的时长有一定的接受度，当广告的投放比例较低时，用户会观看完整的广告并观看后续的视频内容；当广告的投放比例较高时，会较大程度的影响用户体验，并造成用户离开视频观看系统。因此内容商需要合理的决定其商业广告投放比例，在保证系统内用户体验的同时提高商业收益。

为了提高商业广告的观看次数、完播率等，内容商可以将商业广告投放在视频播放过程中的不同位置。由于用户观看在线视频的过程中并非每次均会观看完整的视频内容，根据个人的观看习惯、兴趣及使用的播放设备等，用户的观看行为存在自然离开的概率<sup>[5,6]</sup>。当广告投放在视频的中间位置时，广告的前后均为视频内容，其观看体验与广播电视非常相似，随着播放进程的增加，用户在观看视频的专注度也会有所改变<sup>[7]</sup>。内容商可以根据用户的观看习惯，选择不同的位置进行广告投放，提高广告内容的播放率。

本文关注在预算约束下的最佳资源分配策略问题和视频播放过程中的广告投放问题。针对这两类问题分别建立优化模型，并通过仿真分析其中各个参数在系统中的影响。对于资源分配问题，分析随着预算金额的变化，最佳资源策略中缓存和带宽资源购买量的变化。对于广告投放问题，仿真分析用户群体平均对广告接受程度及广告投放位置对于广告系统的数据量及内容商效益的影响。

## 1.2 国内外研究现状

在线网络视频的流行及其数据量的迅速增长给传输网络带来了很大的挑战，如何利用有限的网络资源为庞大的用户群体提供高质量低时延的视频传输服务成为学术界和工业界的研究热点。针对含有缓存的视频传输网络的研究主要包括缓存算法的研究、系统网络资源的分配，也包括视频传输网络结构的设计等，均旨在提高系统的服务质量以吸引更多的用户，从而提高资源利用率创造更高的价值。基于在线网络视频的庞大用户群体及其高速的传播效率，在线广告推广可以更加

快捷的实现潜在用户的触达，提升广告宣传的效率，因此在线广告的投放问题同样具有很高的关注度。在线广告的投放策略制定过程被细化为多方面的子问题进行研究，旨在进行更加用户友好、准确高效的广告投放决策。

### 1.2.1 视频传输系统资源利用相关研究

网络资源对于视频传输系统中的每一个参与者都非常重要，既关乎终端用户观看过程中的清晰度和流畅度，又影响视频内容商和网络服务提供商的利益，因此合理地利用有限的网络资源以实现与系统收益相关的目标是一个值得探究的问题。现有关于网络资源利用的研究主要从以下几个角度出发：用户感知的观看体验、网络资源的利用率以及系统中各方的收益。这三个出发的角度之间并非完全独立，它们之间的关系相辅相成，可在同一系统中进行综合考虑。

部分研究从用户感知的服务质量的角度，通过缓存部署策略及缓存内容更新算法，在为用户提供高质量的视频传输服务的同时提高网络资源的利用率<sup>[8]</sup>。文献[9]提出通过有效的带宽资源分配方案控制服务器的发送速率，且在客户端设置三个队列：播放、下载和等待播放，用来减少网络速率抖动带来的影响，同时将观看视频的用户区分优先级，提出加权带宽分配方案为等待中的用户传输视频，最大程度的减少用户的等待时间。在这样的传输系统中，多个基站的协同合作，提升文件的命中概率，从而降低请求的平均响应时间，通过存储多个文件副本降低平均传输时延，提升用户体验，但是存在基站缓存中文件冗余。部分研究通过分析其对数据包传输时延及丢包率的影响，通过有效分配带资源，降低视频传输过程中时延的波动，从而提升用户体验<sup>[10,11]</sup>。有研究通过优化资源的利用策略以在基站之间分配传输带宽传输视频文件，执行缓存空间管理，最大化在给定时间内响应的视频请求数量<sup>[12,13]</sup>。用户的偏好会影响网络资源的分配决策，因此有研究将网络拓扑结构、用户请求模式的多样性和动态性、视频流行的生命周期等因素组合考虑，设计缓存策略，保证视频播放的时延，提升用户观看体验<sup>[14-19]</sup>。上述研究均是在网络资源已经给定的情况下，通过带宽和缓存空间的调度使用方式，提高现有资源的利用率，进而提高用户体验或内容商的收益，本文则从最初的部署阶段进行分析，根据网络资源对用户体验和内容商收益的影响，作出资源的部署决策，对现有研究的优化过程进行补充。

部分研究则将两类网络资源的优化进行整合，在系统中同时考虑带宽和缓存的影响，综合分析两者在系统中部署的策略。文献[20]中在多用户系统中，将缓存配置和带宽分配问题建立为一个混合优化问题，联合优化缓存的内容放置和带宽分配。在保证用户公平性的约束下，采用模拟退火算法，固定一类网络资源优化另一类，迭代的进行整体的资源配置优化。文中主要的优化目标为用户之间的公平性，但是没有考虑实现公平性的资源开销及内容商的收益情况。文献[21]中通过时延将缓存和带宽对用户体验带来的影响联系起来，分析在不同参数影响下

视频内容商效益最大化的缓存和带宽的最优值的取值情况。其仿真结果表明，在视频内容的价值、网络资源价格及用户需求在一定的区间内时，缓存和带宽的最佳值呈正相关，即两者存在相互促进的关系，一方的增加使得系统内另一方的最优值也有相应的提升，但是在实际网络资源购买决策时，两类资源的购买量并非可以无限增大，往往是在一定预算约束条件下进行决策，则两者之间会形成竞争有限预算的关系。

由于多个内容商提供视频传输的服务，因此其业务策略及资源利用策略非常重要，尤其是在预算有限时。在内容商制定自身的策略时很难了解竞争对手的策略选择，作者将内容商之间的竞争建模为 Blotto 上校博弈，其中每个区域的收益均随分配给它的预算而变化，作者表明内容商的最佳预算分配策略在其纳什均衡点上，且存储资源对于内容商的收益起关键作用，同时建议在预算较少时，将预算资源集中在特定的几个区域<sup>[22]</sup>。

根据现有的主要研究可知，关于含有缓存的视频传输网络的研究均与用户感知的服务质量及提供服务方的效益相关，而这两个方面通过系统中的网络资源作为桥梁联系起来，因此关于网络资源的部署和相关的使用策略显得尤为重要。

### 1.2.2 视频系统广告投放相关研究

在线广告投放以其高效的潜在用户触达率及巨大的变现潜力成为网络时代的一种重要宣传推广方式和盈利方式。由于在线视频拥有庞大的用户群体和日益增长的数据量，在其中投放广告成为在线广告投放领域内的研究热点。关于在线多媒体广告的研究主要集中在以下几个方面：广告插入位置、时长、内容等的选择，以及广告的插入对于用户观看过程的影响及内容商收益的影响。这两类研究之间存在一定的因果关系，前者影响广告投放的准确性，从而影响用户的观看体验和内容商的收益。

在线视频及广告的播放情况及用户感知的质量（Quality of Experience, QoE）对于内容提供商而言非常重要，如谷歌使用 YoMoApp 等工具进行视频流播放情况等检测，并进行用户反馈的收集，使用这些数据对于用户的参与度、QoE 等质量评估指标进行分析研究<sup>[23]</sup>。有研究在投放过程中可以根据一定的算法预测网络服务的响应时间、吞吐量和质量等，结合上下文数据，分析用户的 QoE 信息<sup>[24]</sup>。

不同的用户群体对于广告的兴趣和接受程度是不尽相同的，内容商可以根据用户群特性，定制化地选择广告投放策略。文献 [25] 指出在 IPTV 系统中传输信道切换时，向用户播放广告的用户体验较显示黑屏好，但是在切换时间较短时，用户倾向于黑屏。从中可知，在线视频的某些播放场景下，用户对于广告插入情况体验不同，但是广告的插入并非一定会影响用户体验。文献 [26] 将 IPTV 系统中用户划分为多个群体，根据特定接受程度的用户，采用不同的广告插入方法，同时兼容传统的电视节目广告放置方法，实验表明其策略可以最大程度地提高从广

告中获得的收入。文献 [27] 表明只有小部分用户非常明确的拒绝观看插入的广告，在视频内容与广告相关或者对广告内容感兴趣时用户会欣赏并完整观看广告。

影响在线视频系统中用户对于广告接受程度的原因是多方面的。文献 [28] 总结了互联网多媒体广告的发展趋势，并归纳出多媒体广告投放模式中的主要问题，包括上下文相关性：用于选择与视频最佳的相关广告；上下文介入性：即视频等多媒体广告候选插入位置的选择；广告插入优化：即候选广告和位置之间的最佳关联，以提高广告的有效性和上下文相关性，并减小广告插入的介入性。针对上下文的相关性，现有研究考虑了意、场景、情感、对象、音频和颜色等多个方面的内容，并结合广告相关的产品、服务信息与视频内容相关的信息，提高在线广告投放的用户友好性<sup>[29-32]</sup>。有研究提出根据视频上下文的关系选择广告的候选投放位置，然后根据广告候选列表投放至匹配的位置中，最小化其对用户的干扰<sup>[33,34]</sup>，并且在此基础上实时地调整广告顺序，以更好的适应用户的观看喜好及视频的上下文情况<sup>[35]</sup>。

用户对于不同视频的广告插入时长同样会感受到一定的 QoE 差异。在文献 [7] 中介绍了现有的多种视频中插入广告的方式，设计并实现了含有广告插入的视频网站，通过邀请用户观看视频获取反馈分析不同类型的广告插入对于用户观看体验的影响，分析在长视频（60 分钟以上）、中长视频（约 20 分钟）、短视频（约 3 分钟）中以不同方式插入广告时对用户的 QoE 带来的影响，其实验结果表明，短时间的广告或者可跳过的广告（5 秒）对于用户 QoE 的影响较小，而在长视频中插入广告则对用户 QoE 的影响稍小于短视频。

投放在线广告的收益是网络商和内容商关注的一个重点内容，文献 [36] 中，作者通过调整视频速率和插入广告时长两个向量最大化无线网络提供商的利润，同时提出将总的优化问题分解为多个子优化问题，使用分层贪婪算法来获得最佳利润点。文献 [37] 中，作者提出了一种在线视频播放过程中的多目标优化方案，在给定广告主预算和用户覆盖率的情况下提高广告的曝光量和用户参与度（观看时间）。用户参与度是在线视频关注的另外一个重要因素，用户观看完整的视频是对其内容和质量的肯定。为了提高在线广告的用户参与度，除考虑位置和上下文相关性外，将计算性情感视频引入视频广告投放策略，考虑视频的情感影响和一组待选广告，使用非线性整数规划识别广告插入点和合适的广告内容，通过用户调研和用户眼动追踪实验进行效果分析，实验表明其可以增强用户体验，增加用户对广告的参与度，帮助增加广告内容的获利潜力<sup>[38]</sup>。

综上所述，目前关于在线广告投放的研究主要包含两大方面：从用户的角度，如何进行广告投放决策可以使得对其的干扰较小，且可以吸引其观看广告内容，现有的研究包括投放位置、长度、内容等的选择，考虑了多方面的因素；从投放者的角度，制定合适的策略在不造成其用户流失的情况下最大化其收益是关注的重点。这两类研究问题相互联系，互相影响，一方面广告投放的准确性会直接影响用户的感受，良好的广告投放准确性既可以减小对用户的影响又可以帮助投放者提高

收益；另一方面大量的广告投放可以提高投放收益，但是会严重的影响用户体验。这两方面是商业广告投放的主要关注点。

### 1.3 本文研究内容

内容商在现有的视频传输系统中部署其网络资源时，有限的预算无法支持其购买无限的网络资源，随着用户需求的不断增长，缓存资源可以使得网络边缘的服务器响应更多的用户请求，传输带宽使得远端源服务器可以为用户提供高质量的服务，两者形成争用预算的关系，如何进行预算划分对内容商实现收益最大化的目标非常重要。

在不同的计费模式下，内容商的收益来自于不同的数据量。当用户向内容商付费时，其收益取决于用户观看的视频数据量，而当内容商通过投放广告赚取推广费用时，其收益取决于用户观看视频中广告内容的占比。由于广告内容投放的比例及位置均会影响用户的观看体验，造成用户离开视频观看系统，因此制定合理的广告投放策略对内容商的收益具有重要意义。

内容商的预算分配策略，即网络资源购买策略，从其支出的角度出发，在考虑服务质量对用户观看视频数据量的影响下，通过优化其资源配置，提升其收益；而在线视频广告则将内容商所提供的视频数据量划分为两部分，高质量的视频内容用于吸引用户，而插入的广告内容则使得内容商获得收益，因此广告投放策略直接影响内容商的整体收入。本文根据现有的研究工作，将以上两个问题分别进行建模为优化问题，并分析系统内各个因素对内容商收益的影响，主要分为以下两个方面：

(1) 针对预算分配问题，本文分析带宽和缓存资源对于用户观看时延的影响，依据观看时延建立用户观看内容商视频量之间的关系，进而建立有预算约束下的资源分配模型。在给定预算的情况下，根据用户观看时延、用户观看的数据量、系统内传输的总数据量及内容商收益之间的关系，寻找最佳的带宽和缓存资源分配策略，最大化内容商的收益。根据系统中最优策略时带宽和缓存之间存在争用有限预算的关系，设计最大可行面算法，求解最优的带宽和缓存分配方案，并通过仿真验证在不同参数设置下，算法的有效性及其对内容商最佳策略和收益的影响。

(2) 针对广告投放问题，本文在用户观看视频的自然离开概率的基础上，联合广告投放比例及位置对于用户离开概率的影响，分析在含有广告投放的系统中用户观看视频的数据量的变化和内容商的收益的变化。通过建立广告投放为内容商收益的影响的模型，寻找最佳的广告投放时长和位置，最大化内容商的收益。本文对给定广告投放位置下的最优广告投放时长进行了理论推导，对广告位置和时长两个因素共同作用下的系统内用户需求的数据量和内容商的收益进行了趋势分析，数值分析发现了在给定用户自然离开概率和广告投放因素造成用户离开概率的共

同作用时的最佳广告投放策略及其随投放策略的变化趋势。

## 1.4 本文工作内容安排

本文对有预算约束下的视频缓存系统中的网络资源部署策略进行建模研究，并对含有商业广告投放系统的投放策略进行建模分析。主要包括五个章节，每个章节的内容安排如下：

第一章对本工作的研究背景和研究意义进行介绍，简要的介绍互联网视频的发展状况，视频传输系统及互联网视频广告系统的研究现状。

第二章对相关研究背景及技术进行介绍。首先介绍典型的含有缓存的多级视频传输系统，分析各参与方在系统中的收益；然后对在线视频的广告投放研究及技术进行介绍；最后介绍用户观看在线网络视频的付费及观看行为习惯。

第三章对含有预算约束的视频缓存系统的网络资源部署策略进行分析。首先建立有内容提供商、网络提供商和终端用户三方参与的系统模型；然后分析三者之前的资源供求关系，并建立优化模型分析内容商在给定预算条件下的最佳网络资源购买策略；最后通过数值仿真分析系统内各因素对于内容商的最佳策略和收益的影响。

第四章对含有商业广告投放的视频传输系统进行建模分析。首先对系统的模型进行明确，分析商业广告的投放比例、投放位置和在线网络视频的用户观看行为对系统的影响，并建立优化模型分析最佳投放比例及位置；最后通过数值仿真分析商业广告的投放比例、位置对于内容商收益的影响。

第五章对本文的工作进行总结，分析本文的主要成果，并提出本文的下一步可能的工作内容。

## 2 相关研究及技术

本章首先对典型的带有缓存的视频传输系统、以及对缓存系统中网络资源的相关研究进行介绍；然后介绍现有在线视频广告的类别和广告投放相关的工作；最后介绍用户观看在线视频离开行为的相关研究。

### 2.1 典型的视频缓存传输系统

本节首先介绍含有缓存视频传输系统中的三类主要参与对象，对其在系统中的关系和收益进行分析，为本文的主要研究系统建立概览。

#### 2.1.1 系统中的主要参与对象

本文主要的研究系统是移动网络环境下的视频缓存传输系统，其中的三类主要参与对象为视频内容提供商、网络服务提供商和终端用户。三类参与者分别提供不同的服务或价值，构成完整的资源和效益的流动链。

典型的视频缓存传输系统的结构如图2-1所示，从视频文件的传输路径，系统可以自上向下划分为三段：内容提供商、网络提供商和终端用户。内容提供商维护视频源服务器，存储其拥有的所有视频文件。网络提供商搭建起连接服务端和客户端的网络基础设施，包括连接内容提供商的接入网、远距离通信的传输网、为终端提供服务的基站和相应的缓存服务器。终端用户同时享受内容提供商和网络提供商的服务，内容商提供视频文件内容，网络商提供内容传输服务，共同完成终端用户的观看请求。

#### 2.1.2 系统中网络资源和各参与方收益

典型的带有缓存的视频传输系统由上一节介绍的三个主要参与方组成，其中内容商和网络商都拥有各自的资源以提供服务，从中获取相应的收益。

内容商服务的对象是终端用户，内容商端的源服务器存储着其全部的视频资源，用户通过搜索请求或者订阅等多种方式发起视频文件请求，内容商通过相应的计费策略收取用户的观看费用。常见的收费策略有以下两种：

(1) **用户付费模式**：用户通过内容商制定的计费策略，支付相应的订阅费用观看其拥有的视频内容。

(2) **免费内容 + 商业广告模式**：用户免费观看内容商的视频，在其观看过程中，内容商插入一定量的广告从而获得收益，即广告主向内容商支付的广告费用。



图 2-1 视频缓存传输系统

Fig 2-1 Cache-assisted Video Delivery System

用户在观看内容商的视频时，可以选择其中一种计费模式。无论选择哪一种计费模式，内容商均需保证其提供的视频的质量，以吸引用户订阅其内容，或者通过观看商业广告盈利。同时，商业广告出现的占比、频次及位置，均会对用户的观看体验造成影响，从而导致观看的中断，因此内容商也需要制定合理的商业广告投放策略。

在为用户提供视频内容时，数据从远端的源服务器发送到边缘用户，内容商需要租用网络商所提供的网络基础设施，传输数据需要租用一定量的带宽资源。租用的传输带宽越多，则传输的速率越快，可以尽可能的为用户提供高分辨率、低中断概率的视频播放服务。在用户量较大时，边缘用户所请求的视频会出现一定程度的集中，即大量的用户会重复请求少量的视频文件，而通过骨干网重复传输相同的文件内容会造成网络资源的浪费，为解决这样的问题，网络商通过在距离用户较近的基站端设置缓存服务器，为内容商临时存储请求量较高的文件提供服务。内容商则通过将此类文件存储在网络边缘，减小传输带宽租用量的同时，距离用户较近的资源可以为用户提供更高效的服务，使得缓存服务器的设置成为一种共赢的选择。综上所述，为了保证服务质量的同时吸引更多的用户，内容商需要设置合理的网络资源租用策略和视频缓存更新策略。本文的研究内容主要关注网络资源租用策略。

## 2.2 视频系统中的广告投放

本节中首先对现有在线网络视频中插入广告的形式进行简要的介绍，然后简述关于在线视频中投放商业广告的主要研究内容。

### 2.2.1 视频中插入的广告形式

随着网络传输技术的不断发展，人们的观看习惯逐渐从广播电视转向了在线视频，用户可以在任何地点、任何时刻通过移动终端设备请求观看内容商在其服务器上存储的视频内容。内容商在视频播放的过程中通过播放商业广告内容赚取广告主的推广费用。在线视频观看过程中的广告投放存在多种形式，可以投放的广告内容包括文字、图片、视频等多种广告类型。<sup>[39]</sup>

文字和图片类广告可用于投放展示广告和重叠式广告，主要在网站观看在线视频时显示于台式机或笔记本电脑上。如图2-2、图2-3所示，展示广告显示在视频的右侧播放列表处，重叠式广告为半透明广告，显示在视频下方的区域内。拥有广告位的在线视频网站对于这一类的广告有着规格、质量、类型等的要求，以方便与其网页的设计进行匹配，从而设计合理的算法进行广告投放。



图 2-2 展示广告  
Fig 2-2 Display ads

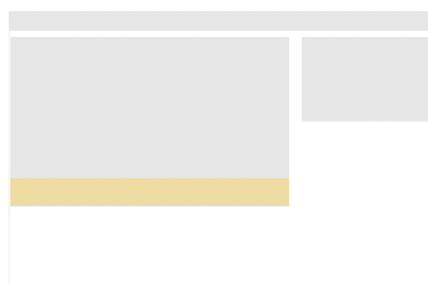


图 2-3 重叠广告  
Fig 2-3 Overlay ads

视频类广告主要包括可跳过广告、不可跳过广告和导视广告，在台式机、笔记本电脑及移动设备上均可投放。可跳过广告和不可跳过广告可以投放在视频播放前、播放中或播放后。可跳过广告一般较长，用户观看一定时长后可以选择跳过该广告，继续观看后续视频内容。例如，在 YouTube 平台上，可跳过广告时长最长为 3 分钟，用户观看 5 秒后可以选择跳过该广告，也可继续观看完整的广告内容<sup>[40]</sup>。不可跳过广告是指在用户观看过程中插入的一定长度的广告，用户必须观看完整的广告内容才可以观看后续的内容。例如，在 iQiYi 平台，视频前的不可跳过广告时长约为 2 分钟，用户必须观看完整广告后才可继续观看请求的视频内容<sup>[41]</sup>。导视广告是视频播放前的短视频广告，通过较短的时间内（YouTube 平台为 6 秒）紧凑的向用户传递一件事情，该短视频导时广告不可跳过，但是可以带

来较好的广告效果，通过较短的时间给用户留下较为深刻的印象<sup>[40]</sup>。

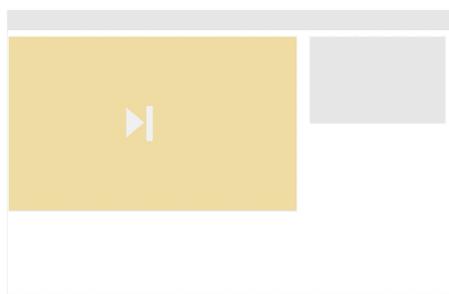


图 2-4 网页端视频广告  
Fig 2-4 Website video ads

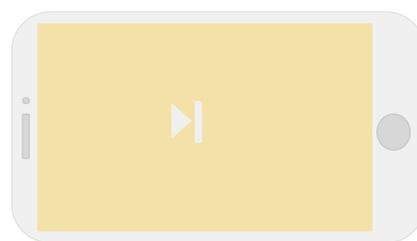


图 2-5 移动端视频广告  
Fig 2-5 Mobile video ads

### 2.2.2 商业广告投放策略

在线广告存在非常高的传播效率，若广告主希望其推广内容触达 4000 万潜在用户，传统的电视广告需要 4 年，基于 PageRank 的算法需要 6 个月，而基于机器学习刻画的用户特征模型进行广告投放则仅需 2 个月即可触达目标用户<sup>[7]</sup>。在线视频网站通过向用户提供视频时播放商业广告赚取广告主的推广费用，因此用户观看情况直接影响其收益，内容商则将注意力集中在广告的投放策略上，以最大化自身收益。

现有的广告投放策略不仅需要考虑向用户播放广告的形式，还需要考虑其他多方面因素，从而避免过多影响视频观看的用户体验。关于广告投放策略的研究包含以下几个方面：广告内容与视频内容上下文的联系性、广告时长与视频时长的关系以及广告的目标受众及其他信息，同时也将不同用户群体对商业广告的态度纳入考虑。

广告内容与视频内容的上下文联系不仅是指广告内容之间的关系，同时也包含广告视频的质量、分辨率、色彩、风格等内容。良好的视频和广告的相关性可以吸引用户的注意力，使得用户可以观看完整的广告内容，达到预期的商业推广、品牌建立等商业投放目的。制作粗糙、质量较差的广告不仅会影响用户观看体验，退出观看系统，甚至可能影响用户对于内容商平台视频的需求，影响其整体的用户信任度及品牌形象。

广告时长及广告插入位置会直接影响用户观看视频的时间成本及观看连续性，因此其对用户观看体验的影响非常直接，且不当的投放策略会造成用户直接离开观看系统。现有的广告插入时长及位置选择重要考虑以下几方面因素。

首先为视频时长因素，当视频时长在 3-5 分钟内时，用户倾向于迅速跳过广告，观看视频内容，因此在线短视频根据其性质，当广告投放时长与其内容时长相当时，将会严重影响用户体验；而长视频则可插入较长的广告，不同的用户群

体对于广告时长的接受程度存在差异，其投放策略需要根据用户特性进行调整。

其次为插入位置及频率因素，短视频根据其性质，广告一般投放在视频内容之前；而长视频则可以参考广播电视的模式，在视频中的不同位置插入多段广告视频，但是同样长度的视频，若多次插入时间较短的广告会较大程度的影响用户体验，现有的长视频倾向于每十分钟插入一个广告，广告时间会在播放的时间轴上显示<sup>[7]</sup>。广告插入位置则将视频转场、镜头切换等信息纳入考虑，在这类时间节点上插入广告可以减少对用户观看连续性的影响<sup>[25,29]</sup>，而在内容播放过程中，如视频中人物说话中，插入广告则会严重影响用户观看过程，带来较差的用户体验<sup>[42]</sup>。

最后为广告的可跳过性，视频广告根据其特点及付费策略，可以设置其播放过程中的商业广告是否可以跳过，并设置跳过广告的必须观看时长，对于较长的广告内容，可跳过性和强制观看时长对用户的观看体验会带来一定程度的影响，因此其设置决策需要根据用户特征进行选择。

### 2.3 在线网络视频的用户行为相关研究

用户观看在线网络视频时，可以根据其观看偏好及对内容的兴趣，随时离开视频系统结束观看。大量的视频曝光量并非表示视频的参与度高，即用户并非观看完整的视频内容，而是随着播放过程的持续，不断会有用户离开视频播放系统<sup>[43]</sup>。根据文献 [6] 和 [7]，用户观看视频过程中的自然离开概率可以用一个被截断的指数分布函数表示，其离开概率的累积概率分布如图2-6所示。

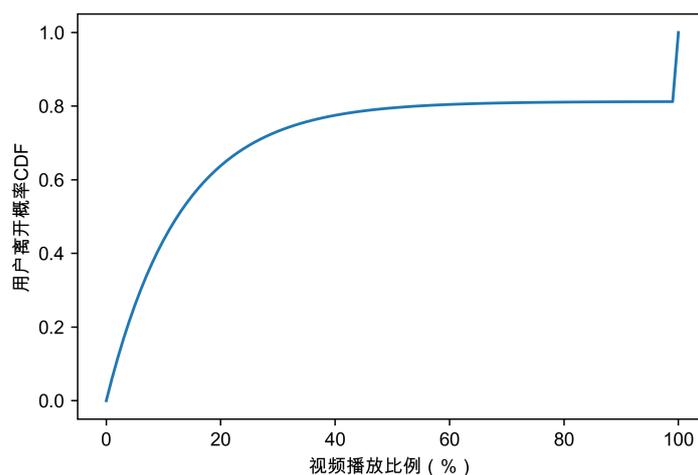


图 2-6 用户自然离开概率的累积分布

Fig 2-6 CDF of user natural leave rate

图2-6表示随着视频播放进度的增加，用户离开系统的累积概率分布逐渐趋紧于稳定，在最终时刻跳变为 1，其表示的含义为：用户在视频播放初始时期结束观

看的概率较高，即大部分用户会在视频播放初始阶段根据其兴趣等原因决定是否继续观看；随着播放进度的增加，用户离开系统的概率逐渐下降并趋于稳定，即随着播放时长的继续增加，若用户继续观看视频，则其以较高概率观看完整的视频内容；用户离开概率的累积分布在最终时刻跳变为 1，即当视频播放结束时，所有用户均会离开视频播放系统。

根据文献 [6]，用户在不同设备上观看线视频时，其自然离开概率模式相似，但是其曲线并非完全相同。用户在移动端和 PC 端观看视频离开系统的概率随着视频的播放进度存在不同的峰值点，这被认为和不同的内容流行度相关。数据分析表明 PC 端的视频流行度并非如移动端一般陡峭，即 PC 端用户具有更加广泛的兴趣，因为其包含更多的流行视频内容，且其分布也表明移动端流行的视频在 PC 端并非同样流行，这被归因为两者的视频源并非完全相同，即视频数据库存在差异。两类终端的观看数据均表现出大量的视频观看集中在少量的视频内容上，为部署 CDN 服务器及优化缓存算法提供了根据。

## 2.4 本章小结

本章对本工作中的主要研究场景进行了简要介绍，即带有缓存的视频传输系统。该系统主要包括三类参与者：内容提供商、网络服务提供商和终端用户，同时介绍了内容商常见的两类计费模式：“付费内容”模式和“免费内容 + 商业广告”模式。随后对现有的商业广告形式及现有关于视频广告投放的主要研究内容进行介绍。最后介绍了用户在线网络视频时的离开行为特征及其随播放时间的离开概率累积分布。

### 3 含预算约束模型的网络资源分配

上一章节中对研究系统及其中的主要参与者进行了简单的介绍，本章将会对各个参与者进行更加深入的刻画，并对系统中涉及的各类资源、影响因素进行参数化，细化系统中各环节的因素。根据视频传输系统内各因素对用户观看视频量的影响，从内容商支出的角度，分析其在有限预算下的网络资源部署策略对内容商收益的影响。

本章的内容主要划分为以下内容，第一节对研究场景及本章节的研究内容进行明确，第二节对研究场景中的各个环节、资源、影响因素进行参数化，并建立优化问题模型，第三节对本章的研究问题进行理论分析，第四节对优化模型设置合理的参数进行数值仿真，最后对本章内容进行总结。

#### 3.1 研究场景

在含有缓存的传输系统中，内容商租用网络商的基础设施向终端用户提供视频服务，通过收取用户的订阅费用或者投放一定比例的广告获取利润。在租用网络资源时主要包括两类资源：网络传输带宽和基站缓存空间。通过一定的缓存算法将部分请求频率较高的文件存放在距离用户较近的基站缓存服务器中，内容商可以降低用户请求的响应时间，提升用户体验；同时对于网络服务提供商而言，可以减少重复文件的传输，降低骨干网传输压力，这被认为是一个共赢的措施。内容商通过制定合适的缓存策略，在发生缓存遗漏时或者设置更新时间，对缓存服务器中的内容进行更新，以提高缓存命中率。除租用缓存空间外，内容商仍需租用一定量的带宽资源，当用户请求缓存服务器中未存储的视频资源时，响应用户的请求，向用户传输视频数据。在网络中立性原则被投票取消后<sup>[44]</sup>，网络商通过与内容商签订合约，内容商可以在购买网络带宽时要求网络商保证相应的服务质量。由于购买的带宽量越大，对网络商的要求越高，所以其单位价格随着购买量而变化。在为用户提供视频传输服务时，缓存空间和带宽量均会影响视频的服务质量，且两者之间的关系呈正相关<sup>[21]</sup>。内容商在制定资源的租用策略时，两类资源均可以使得其服务质量上升。当内容商的预算有限时，必须合理的分配两类资源的预算额，使得在预算范围内实现效益的最大化，因此本章将对系统中三方参与者之间的关系进行细化分析，并在预算固定的情况下内容商的资源分配策略进行研究，分析系统中不同的因素对其策略及收益的影响。

## 3.2 系统模型

在本节中,将对系统中三方之间的关系进行详细的阐述,并对主要的变量、影响因素进行参数化。视频服务系统中,主要有以下三个参与者:内容提供商(Content Provider, CP)、网络服务提供商(Internet Service Provider, ISP)和终端用户(End Users, EU)。内容提供商通过向终端提供视频内容,根据用户观看的视频数据量获得相应的收益,网络服务提供商为内容商提供数据的转发服务和缓存服务,内容商支付相应的网络基础设施租赁费用。用户通过网络服务商搭建的基础设施得以观看到相应的视频内容,其观看质量取决于各类网络资源的使用策略。在本节中将会对这三个参与者之间的关系进行分析。

### 3.2.1 参与者之间的关系

在本小节中对在线视频服务系统中的三类参与者之间的关系进行明确,并将其与网络资源之间的关系进行参数化,以便进行下文模型的刻画。

#### (1) 内容提供商和网络服务提供商

内容商(CP)拥有着大量的视频数据资源,并将其存储在距离用户较远端的源服务器上,通过向用户提供视频服务,获得相应的收益。网络商(ISP)拥有网络基础设施,并向CP提供视频数据的转发和缓存服务,将视频内容传输到网络边缘的用户。ISP提供免费的基础网络传输服务,该服务实现尽最大努力的交付(A best-effort service),同时也提供保证服务质量的数据传输服务,需要CP支付额外的购买费用。由于视频的流行度符合一定的分布,当大量用户请求CP的视频内容时,网络中传输的视频内容存在大量的重复,为了减少重复内容的传输,缓解骨干网的压力,ISP在距离用户较近的基站处设置缓存服务器,根据缓存更新策略,将一些高请求量的内容存储在网络的边缘响应用户的请求,既提高请求的响应速度,又提升CP的服务质量。由于CP的服务质量会影响用户观看视频的决定,为了可以向用户提供高质量的服务,CP需要向ISP购买相应的网络资源,即带宽资源和缓存资源,以提升自身提供服务的能力和服务的质量。

当CP选择购买两类网络资源的量时,其组成的资源组合成为一种策略。将购买的带宽资源记为 $\bar{b} = b_0 + b$ ,其中 $b_0$ 表示ISP提供的基础服务,实现尽最大努力的交付, $b$ 表示CP购买的额外资源,以实现与终端用户之间链路上较好质量的传输。将购买的缓存资源量记为 $c$ ,缓存服务器位于边缘基站,根据CP的更新策略对存储的内容进行替换,为边缘用户提供高质量的请求响应服务。将带宽-缓存组合 $(b, c)$ 称为CP的资源组合策略,CP可以根据需求决定两种资源的购买量,因此资源量满足非负约束,即 $b \geq 0, c \geq 0$ 。

ISP为CP提供免费的基础服务以及收费的额外资源服务。由于在额外服务

部分要求 ISP 保证质量的交付, 所以随着购买的量的增大, ISP 提供服务的难度随之增长, 因此 CP 购买的带宽的额外服务的单价随购买量的上升而上升。将  $p_b(b)$  记为单位流量的价格, 当用户确定购买量后, 其产生的费用为单位时间内使用的总流量与单位价格的乘积。在网络边缘基站, CP 可购买缓存空间, 其单位容量的价格记为  $p_c(c)$ 。购买缓存空间的价格受购买容量的影响, 但是不受实际服务吞吐量的影响。

## (2) 网络服务提供商和终端用户

在视频服务系统中, 内容商所能提供的服务质量会直接影响到用户是否选择其视频服务。内容商在为用户提供视频内容选择, 同时也会提供视频不同的清晰度、分辨率、音效等, 而更高质量的视频内容的数据量相应会更大。如果内容商要求保证高分辨率内容的传输速率, 即保证不同内容的服务质量, 则其需要更大的网络带宽来保证传输的速率, 或者购买足够多的缓存空间, 并制定缓存策略将请求较多的高分辨率文件存储在网络边缘, 以保证满足尽可能多的用户请求。在相应的资源策略下, 用户通过 ISP 网络观看视频会感受到一个整体的平均传输时延, 记为  $\bar{d}$ 。平均传输时延表示平均每个数据包传输到用户的时延, 用户接收到的数据包的传输起点可能是内容商的源服务器, 也可能是网络边缘的缓存服务器, 但是影响其观看体验的并非单个的数据包, 而是整体时延的平均值。

## (3) 内容提供商和终端用户

内容商通过高质量的视频内容和服务质量吸引用户, 从而获得用户的订阅费用。假设对内容商提供的视频内容感兴趣的用户数为  $n$ , 用户决定订阅其内容的概率取决于其提供的传输服务质量, 当用户提供服务的平均时延为  $\bar{d}$  时, 平均每个用户的订阅概率记为  $\eta(\bar{d})$ 。当用户订阅该内容商时, 每个用户的平均视频服务需求量记为  $r(\bar{d})$ , 也就是内容商给每个用户的平均视频发送速率, 则系统内平均每个用户的视频需求可以记为  $f(\bar{d}) = \eta(\bar{d}) \cdot r(\bar{d})$ , 总的需求可以记为  $T = n \cdot f(\bar{d})$ 。当  $\bar{d}$  发生变化时, 即系统内的数据传输时延变化时, 用户对内容商的内容订阅概率及需求量随之发生改变, 从而影响系统内整体的传输数据量。当其传输情况恶化时, 也就是  $\bar{d}$  迅速上升时, 为防止其导致的用户订阅概率的下降导致的系统总需求下降, 此时内容商倾向于通过调整其向用户发送数据的速率, 防止整体服务质量的进一步恶化。

终端用户通过支付订阅费用, 观看内容商提供的视频内容。从系统整体的角度分析, 假设用户支付固定的订阅费用, 且其平均视频需求已知, 则可得系统中内容商向用户传输单位数据量的视频可以获得的收益。假设单位流量的视频价值为  $p_v$ , 则内容商的总收入即可记为  $p_v \cdot T$ 。

### 3.2.2 服务时延

在视频传输网络系统中，内容商为用户提供的视频数据可以存储于其远端源服务器，也可以存储于网络商部署在边缘基站的缓存服务器。由于缓存服务器容量的限制，只有部分视频文件可以存储在缓存服务器中，文件存储的位置根据其请求的分布情况和缓存更新算法决定。当用户发起一个视频观看请求时，如果缓存服务器中已经存储该文件，称其为缓存命中 (Cache hit)，并由离用户相对较近的缓存服务器进行服务，此时产生的传输时延为边缘缓存服务器到用户的传输时延，即  $d_0$ 。将缓存命中的概率记为  $H(c)$ ， $c$  为内容商购买的缓存空间。当用户请求的文件未存储在缓存中时，称之为缓存遗漏 (Cache miss)，由远端的源服务器提供视频文件，此时的传输时延包括两部分，从源服务器到边缘基站的传输时延  $d$  及从基站到用户的传输时延  $d_0$ 。将缓存未命中的概率记为  $M(c) = 1 - H(c)$ 。当用户请求到达时，源服务器将文件发送至缓存服务器和用户，并根据缓存算法决定是否将该文件存储在缓存服务器中。在这样的服务流程中，内容商向网络商购买的两类网络资源共同影响着内容商为用户提供服务的能力和质量。

得益于内容商购买的缓存空间，为满足边缘用户对视频内容的需求，从其源服务器端传输到网络边缘的数据量可以有所下降，其下降的程度取决于缓存服务器的命中概率。由于缓存服务器的分担，当系统中总的总数据需求量为  $T$  时，由源服务器发送至用户的数据量为缓存未命中的部分  $\tilde{T} = T \cdot M(c)$ ，其余部分则从缓存服务器发送至用户。当网络商需要传输  $\tilde{T}$  的数据量时，其传输时延取决于数据量和内容商购买的传输带宽，即  $d = D(\bar{b}, \tilde{T})$ ，其中  $d$  与  $b$  呈负相关，与  $\tilde{T}$  呈正相关，即当购买的传输带宽越大或者传输的数据量越小时，传输时延越小。由于传输到用户的平均时延包含从源服务器和缓存服务器传输两部分，因此有整体的传输时延如公式 (3-1) 所示

$$\bar{d} = d_0 \cdot H(c) + (d_0 + d) \cdot M(c) = d_0 + M(c) \cdot D(\bar{b}, \tilde{T}) \quad (3-1)$$

### 3.2.3 供需平衡

当内容商选择资源组合后，其提供视频传输的服务能力会逐渐与用户的视频需求形成动态平衡。当系统内的订阅用户较少，需求量偏低时，内容商可以保证提高质量，提供低时延的传输服务，此时会吸引更多的用户转化为订阅用户，系统内的总需求会随之上升；当系统内的需求量上升时，平均时延会增加，内容商为克服该问题可以自适应的调整其传输视频的分辨率，时延和分辨率的变化导致用户感知到的服务质量逐渐下降，随之导致用户流失，系统内的总需求下降。系统内的需求最终与内容商在相应资源组合的服务能力之间达到动态平衡。<sup>[21]</sup> 由时延、用户需求及资源组合之间的关系可知，最终达到平衡时各参数之间的关系满足公

式 (3-2)。

$$T = n \cdot f(d_0 + M(c) \cdot D(\bar{b}, \tilde{T})) \quad (3-2)$$

根据公式 (3-2) 可知, 系统内总的吞吐量可以记为内容商资源策略  $(b, c)$  的方程, 即  $T = T(b, c)$ 。为方便后续分析, 根据缓存空间和带宽对时延和需求量的影响, 引入用户需求和时延与资源量之间的关系式。

首先根据文献 [21] 引入用户需求与时延之间的关系式 (3-3), 平均用户需求量与时延呈负相关。

$$f(\bar{d}) = \beta \cdot \frac{1}{\bar{d}}, \quad \beta > 0 \quad (3-3)$$

当系统内传输时延很小时, 用户的需求接近于无穷; 随着时延的增加, 用户需求迅速下降; 当时延逐渐增大至无穷时, 系统内的用户需求下降至零。

其次根据文献 [45] 并参考 M/M/1 排队论系统模型, 引入从源服务器到基站边缘的传输时延之间的关系, 如式 (3-4) 所示, 平均传输时延与内容商购买的带宽容量呈负相关, 与系统内单位时间的数据需求量呈正相关。

$$D(\bar{b}, \tilde{T}) = \frac{1}{\bar{b} - \tilde{T}}, \quad \bar{b} \geq \tilde{T} \quad (3-4)$$

系统内单位时间内的需求量较小时, 系统内的传输时延逐渐接近其下限, 取决于内容商购买的带宽容量; 同时, 单位时间需求量不能超过带宽容量, 且当需求量逐渐接近带宽容量时, 时延迅速上升, 导致服务质量下降。

根据公式 (3-3) 和 (3-4), 以及  $\tilde{T} = T \cdot M(c)$ , 通过求解公式 (3-2), 可以得到系统内达到平衡时的总数据量如下:

$$T(b, c) = \frac{\beta n \cdot (b_0 + b)}{(1 + \beta n) \cdot M(c)} \quad (3-5)$$

由公式 (3-5) 易知, 平衡时数据量与带宽容量呈正相关, 与缓存未命中概率呈负相关, 即与命中率呈正相关。

### 3.2.4 预算限制

为了提供更高质量的服务, 两类网络资源的购买量越大越好。当带宽容量非常大时, 内容商的源服务器可以提供高质量的服务; 当缓存容量非常大时, 缓存命中率的提升会使得用户请求在更近处得到响应。由于一般情况下, 源服务器与用户之间的距离较大, 所以这种物理距离造成的传输时延不可避免, 为了提高服务质量, 内容商一定需要购买缓存空间; 而内容商的源服务器中户不断添加新的视频内容, 所以无论使用怎样的缓存算法, 当用户第一次请求新的文件时,

缓存必然无法命中，内容商一定需要购买带宽容量，用于传输源服务器增加的新文件。

基于以上分析可知，当内容商在给定预算的情况下做出网络资源购买决策时，两类网络资源为互相竞争的关系。在制定资源组合策略时，内容商需要考虑两类网络资源在为用户提供服务量及创造利润方面的影响，合理分配资源组合中两类资源花费的占比，以实现收益的最大化。基于资源组合、内容商服务能力、用户感知的平均时延、资源价格和视频收益之间的关系，在预算有限时内容商的资源策略制定问题可以总结为优化问题，其优化目标为最大化内容商收益。

### 3.3 优化建模及理论分析

在本节中，首先对系统中内容商的效益进行明确，并建立最优化模型，然后根据约束条件对最优化问题的可行解进行推导，并分析在约束条件下资源分配之间的关系。

#### 3.3.1 优化建模

在视频传输系统中，内容商从网络商处购买缓存和带宽两类网络资源，为终端用户提供视频服务，从而获得用户的订阅收益。当内容商的资源组合  $(b, c)$  确定后，系统在达到平衡时的总数据量可通过公式 (3-5) 获得，而通过源服务器传输至用户的流量需支付带宽费用，从而其资源购买费用可由缓存和带宽的单价获得，内容商总支出如式 (3-6) 所示。

$$E(b, c) = p_b(b) \cdot M(c) \cdot T(b, c) + p_c(c) \cdot c \quad (3-6)$$

内容商的收入则可由单位视频的价值和系统内总数据量获得，即  $p_v \cdot T(b, c)$ 。由此可知，当给定内容商资源策略时，对应的效益为其策略的函数，如公式 (3-7) 所示。

$$\begin{aligned} U(b, c) &= p_v \cdot T(b, c) - (p_b(b) \cdot M(c) \cdot T(b, c) + p_c(c) \cdot c) \\ &= (p_v - p_b(b) \cdot M(c)) \cdot T(b, c) - p_c(c) \cdot c \end{aligned} \quad (3-7)$$

当内容商在有预算约束条件下制定资源策略时，其策略问题可以建模为最优化问题，如 (3-8) 所示，假设内容商的预算约束为  $B$ 。

$$\text{Maximize}_{b,c} U(b, c) \quad (3-8a)$$

$$\text{Subject to } b \geq 0 \quad (3-8b)$$

$$c \geq 0 \quad (3-8c)$$

$$B \geq p_b(b) \cdot M(c) \cdot T(b, c) + p_c(c) \cdot c \quad (3-8d)$$

其中约束 (3-8a) 和约束 (3-8b) 为可行解的容量约束, 约束 (3-8c) 为内容商的预算约束。

在实际的策略组合中, 优化问题的可行解可能存在一些其他的约束, 例如缓存和带宽的容量具有上限约束, 两者的购买量不可持续增加, 或者带宽和缓存存在最小可划分单元, 优化问题则变为整数规划问题。本文在分析中, 为简化问题以应用优化问题的通用解决方法, 将分析该优化问题的一般形式, 将其作为连续变量问题进行分析。

### 3.3.2 理论分析

根据式 (3-8) 所建模的最优化问题, 可以通过拉格朗日乘子法求解其最优解满足的 KKT 条件。引入拉格朗日乘子  $l_1 \geq 0$ ,  $l_2 \geq 0$ ,  $l_3 \geq 0$ , 可得优化问题的拉格朗日函数为

$$\begin{aligned}
 L(b, c, l_1, l_2, l_3) &= U(b, c) + l_1 \cdot b + l_2 \cdot c + l_3 \cdot (B - p_b(b) \cdot M(c) \cdot T(b, c) - p_c(c) \cdot c) \\
 &= (p_v - p_b(b) \cdot M(c)) \cdot T(b, c) - p_c(c) \cdot c + l_1 \cdot b + l_2 \cdot c \\
 &\quad + l_3 \cdot (B - p_b(b) \cdot M(c) \cdot T(b, c) - p_c(c) \cdot c) \\
 &= (p_v - (l_3 + 1) \cdot p_b(b) \cdot M(c) + l_3 B) \cdot T(b, c) - (l_3 + 1) \cdot p_c(c) \cdot c \\
 &\quad + l_1 \cdot b + l_2 \cdot c
 \end{aligned} \tag{3-9}$$

根据式 (3-9) 可得原问题的 KKT 条件, 分别包括稳定性条件、互补松弛条件、原问题可行性条件和对偶问题可行性条件, 如下所示:

$$\left\{ \begin{array}{l}
 (p_v - (l_3 + 1) \cdot p_b(b) \cdot M(c) + l_3 B) \frac{\partial T(b, c)}{\partial b} - (l_3 + 1) \cdot p'_b(b) \cdot M(c) \cdot T(b, c) + l_1 = 0, \\
 (p_v - (l_3 + 1) \cdot p_b(b) \cdot M(c) + l_3 B) \frac{\partial T(b, c)}{\partial c} - (l_3 + 1) \cdot p_b(b) \cdot M'(c) \cdot T(b, c) \\
 + (l_3 + 1) \cdot p'_c(c) \cdot c + (l_3 + 1) \cdot p_c(c) + l_2 = 0, \\
 l_1 b = 0, \quad l_2 c = 0, \quad l_3 (B - (p_b(b) \cdot M(c) \cdot T(b, c) + p_c(c) \cdot c)) = 0, \\
 b \geq 0, \quad c \geq 0, \quad B - (p_b(b) \cdot M(c) \cdot T(b, c) + p_c(c) \cdot c) \geq 0, \\
 l_1 \geq 0, \quad l_2 \geq 0, \quad l_3 \geq 0
 \end{array} \right. \tag{3-10}$$

当  $l_1 = 0$ ,  $l_2 = 0$ ,  $l_3 = 0$  时, 根据互补松弛条件可知, 求得的最优解满足优化问题的三个不等式条件, 为资源可行区间的内部解, 且预算并非恰好完全使用。当预算  $B$  恰好完全使用时, 根据互补松弛条件可知  $l_3 \neq 0$ , 满足 KKT 条件的资源组合为优化问题的边界解。为求得优化问题的最优解, 需计算其内部解和边界解的目标函数值进行比较, 进而求得全局最优解。

当资源组合  $(b, c)$  确定时, 根据系统内总流量  $T(b, c)$ , 缓存命中率  $M(c)$ , 网络资源单价及视频的单位收益, 可以求得单位网络资源为内容商所带来的利润, 进而可以判断在给定预算时, 在不同的资源组合下收益的变化曲线。

由文献 [21] 可知, 在内容的流行度分布、网络资源的单价满足一定的区间时, 最佳缓存资源和最佳带宽资源呈正相关, 即增加两者的购买量均会使得系统内达到平衡时的流量值增加。在本章节后续的分析中, 将各类资源的参数设置在其呈正相关的区间中, 此时两类网络资源的增加均会给内容商带来收益的提升, 形成两者之间对有限预算的争用关系。

### 3.3.3 Feasible-plane 算法设计

由以上分析可知, 在预算允许的范围内, 固定一类网络资源, 另一类资源越多, 内容商的收益越多。根据网络资源策略的这一性质, 在求解内容商的最佳资源策略时, 可以通过求解给定预算时, 最大程度利用预算的可行解面上各个资源组合所能实现的最大效益值点, 其对应的资源组合即为所求的最佳策略组合, 具体算法如下。

算法 1 的主要流程为, 对于给定的预算、缓存和带宽的最大容量限制, 首先计算可行解曲线 (第 3-11 行), 即对于每一个给定的可行带宽, 寻找缓存空间的最大可行解, 然后对可行解曲线上的每一个点, 计算并寻找其中内容商效益最大的点及其对应的效益值 (第 12-18 行), 所得点及对应值即为所求目标值。

## 3.4 数值仿真

在本节中, 首先根据两类网络资源的价格特点对其价格曲线进行假设, 对视频内容的流行度进行介绍并说明数值仿真过程中采用的缓存命中率假设。然后对不同参数影响下的最佳资源策略及其对应的内容商收益进行分析。

### 3.4.1 价格函数

在制定资源策略时, 两类网络资源的单价会直接影响内容商的购买量及其整体系统的收益, 因此必须在预算范围内根据单价选择合适的资源策略, 以最大化内容商的收益。

#### (1) 带宽价格

当内容商与网络商制定购买合约时, 则要求网络商保证为内容商提供服务的质量, 即保证数据传输到达终端用户, 而非尽最大努力的交付。在保证传输质量

**Algorithm 1** Feasible-plane Algorithm**Input:** Budget, max\_cache, max\_bandwidth

---

```

1: Initialization: feasible_plane = {}, step_size = 1,  $p_b(b)$ ,  $p_c(c)$ ,  $b_0$ ,  $p_v$ ,  $M(c)$ 
2:  $c = 0$ 
3: repeat
4: Find maximum feasible  $b$ :
5: for  $b = \text{max\_bandwidth}$ ;  $b \geq 0$ ;  $b = b - \text{step\_size}$  do
6:   if expenditure based on Equation (3-6)  $\leq$  Budget
7:     feasible_plane  $\leftarrow (b, c)$ 
8:     break
9: end for
10:  $c = c + \text{step\_size}$ 
11: until  $c > \text{max\_cache}$ 
12: tmp_solution =  $(b, c)$ , max_profit = 0
13: for sol in feasible_plane do
14:   tmp_profit based on Equation (3-7)
15:   if tmp_profit  $>$  max_profit
16:     max_profit = tmp_profit
17:     tmp_solution = sol
18: end for
19: The best feasible solution  $(b, c) = \text{tmp\_solution}$ 

```

**Output:** The best feasible solution  $(b, c)$ , maximum profit max\_profit

---

的要求下，随着带宽容量的上升，网络商提供服务的难度随之上升，因此会收取单位数据流量更高的价格，故假设流量的单价计算公式如下：

$$p_b(b) = p_1 \cdot b^\gamma \quad (3-11)$$

其中  $p_1$  表示正常数， $\gamma > 0$ 。

当给定资源策略下，系统内需要为带宽支付的费用时，首先需要计算系统达到平衡时系统内的总流量及给定缓存空间下的缓存命中率，进而得到通过网络商的网络为用户传输的数据，并计算相应的传输费用，如下式所示。

$$o_b = p_1 \cdot b^\gamma \cdot T(b, c) \cdot M(c) \quad (3-12)$$

**(2) 缓存价格**

在本文中，假设缓存单价与购买的缓存空间大小无关，则用于购买缓存的总

价与购买容量之间呈线性关系。假设缓存为常数，

$$p_c(c) = p_2 \quad (3-13)$$

即  $p_2$  表示正常数。缓存的价格随着购买容量的增加而线性增加。

$$o_c(c) = p_2 \cdot c \quad (3-14)$$

### 3.4.2 缓存命中率函数

在内容商购买缓存容量时，边缘缓存命中率会随着其容量的增加而提升，但是命中率的增加过程并非根据容量呈线性提升。用户对内容商的内容请求并非均匀分布的，而是有明显的流行度偏向且流行度是随时间变化的，视频文件的流行度分布符合 Zipf 分布。<sup>[46]</sup> 假设文件集中共有  $N$  个视频文件，Zipf 分布给出了文件集中第  $i$  流行的文件的请求概率。

$$P(i) = \frac{1}{i^\kappa \cdot \sum_{i=1}^N \frac{1}{i^\kappa}} \quad (3-15)$$

其中  $0 < \kappa < 1$ ，在实际系统中  $\kappa$  的取值在 0.7-0.9 之间。

内容商可以通过制定不同的缓存更新策略，决定存储在网络边缘的文件，以提高缓存命中率，使得缓存文件尽可能多的向用户提供服务。在本文中假设内容商采用最不经常使用 (Least Frequently Used, LFU) 缓存更新策略，即存储用户最常请求的文件，当缓存空间用尽时，移除其中最不常被请求的文件。当内容商购买的缓存空间为  $c$  时，则用户最新请求过的  $c$  个文件将会被存储在缓存服务器中，此时缓存命中概率为

$$H(c) = \left(\frac{c}{N}\right)^{(1-\kappa)} \quad (3-16)$$

缓存未命中概率为

$$M(c) = 1 - H(c) = 1 - \left(\frac{c}{N}\right)^{(1-\kappa)} \quad (3-17)$$

### 3.4.3 仿真结果分析

根据上一小节中的网络资源价格函数及缓存命中率曲线的假设，本小节对内容商的最佳资源购买策略进行仿真分析。首先给定各个函数的具体参数，在给定的不同预算的情况下的最佳策略的总体趋势进行分析，然后对影响内容商策略的因素进行具体分析，包括资源价格、内容流行度分布，观察在不同情况下最佳策略组合的变化情况。

### (1) 总体趋势分析

根据优化问题 (3-9)，网络资源价格 (3-11)、(3-13)，缓存命中率 (3-16) 作出如下参数设置： $p_v = 1$ ， $p_1 = 0.01$ ， $\gamma = 0.9$ ， $p_2 = 1$ ， $\kappa = 0.75$ ， $N = 1000$ 。设置缓存空间和带宽的最大购买容量分别为 200，内容商的预算范围为  $[0, 600]$ 。根据 Feasible-plane 算法，可以得到在不同预算设置下，内容商实现最优效益的资源购买策略如图3-1所示。

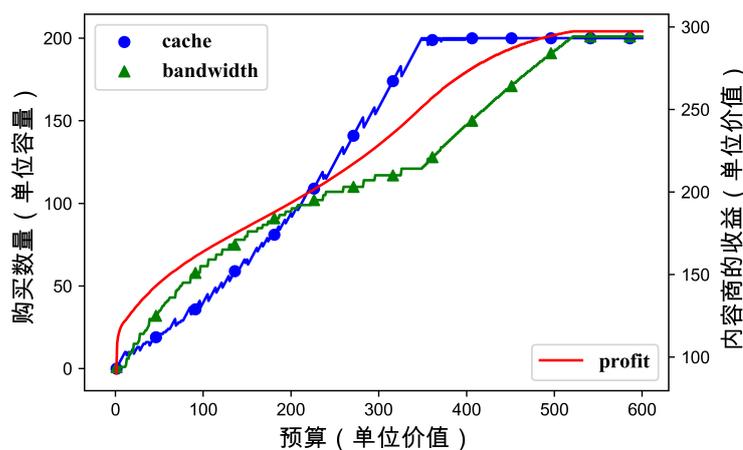


图 3-1 内容商的收益

Fig 3-1 Profit of content provider

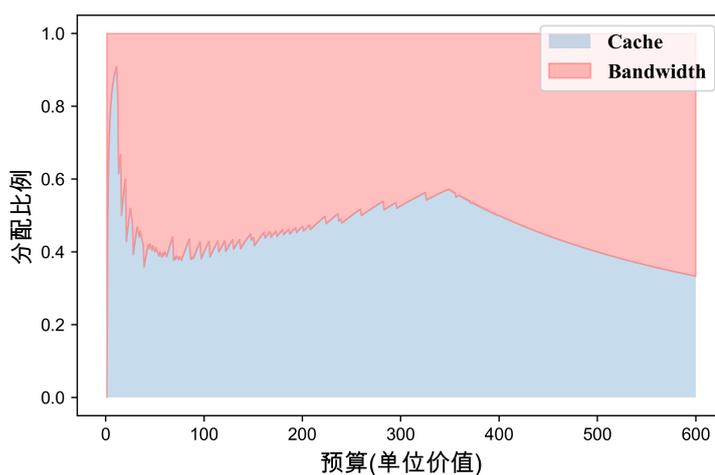


图 3-2 预算分配比例

Fig 3-2 Rate of budget allocation

从图3-1和图3-2中可知，随着预算的线性增长，两类网络资源的最佳购买量呈非线性增加，且其增长趋势存在差异。随着预算的增加，内容商以较高的速度增

加购买缓存的量，并且持续增加至最大值，即内容商倾向于购买尽可能多的缓存资源为终端用户在网络边缘提供视频传输服务。当缓存购买量达到最大值，预算金额继续增加时，内容商的最优策略则继续购买更多的带宽资源为用户提供视频传输服务，进一步提高系统收益，即当视频价值及网络资源价格在相应的可盈利区间内时，内容商倾向于购买尽可能多的网络自资源提供视频传输服务，获取更多的利润。

内容商最佳的资源策略会受到网络内各变量的影响，接下来将分析内容流行度和网络资源价格对于不同预算时最佳分配策略及内容商收益的影响。

## (2) 内容流行度影响

内容商的视频流行度决定存储在缓存服务器中的内容可以响应终端用户请求的概率，当内容流行度分布集中在少量的流行文件时，即用户请求的内容相对集中时，缓存服务器中的内容可以响应更高比例的用户请求，从而使得缓存容量产生更多的收益。

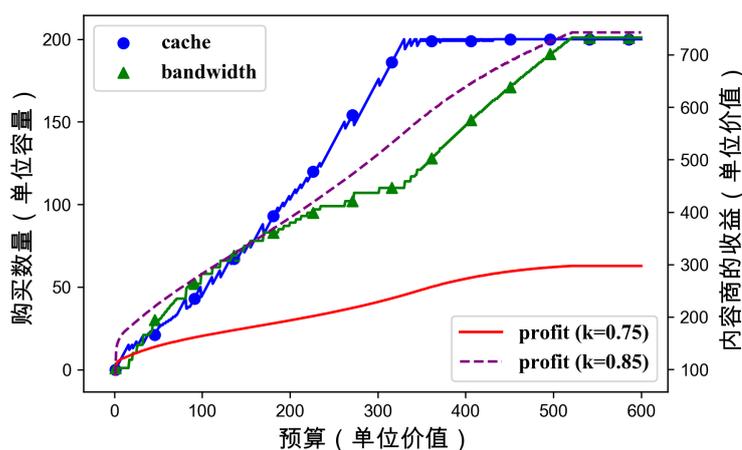


图 3-3 流行度对最佳策略的影响

Fig 3-3 Affect of content popularity on the best strategy

在仿真中，选取不同的流行度分布函数的参数，即  $\kappa = 0.75$  和  $\kappa = 0.85$  两种情况，进行仿真分析随着预算的变化，内容商最佳资源策略的变化情况，如图3-3所示。图3-3为  $\kappa = 0.85$  时的最佳资源策略变化曲线图，将其与  $\kappa = 0.75$  时的情况进行比较，可以看出随着视频内容更加集中的较为流行的文件上时，在预算有限时，内容商首先倾向于购买较多的缓存资源，便于在网络边缘存储流行度较高的文件为用户提供服务。当购买的缓存容量达到最大值时，随着预算的增加，内容商则继续购买带宽资源直至达到容量的最大值。图3-3中对比了在  $\kappa$  取不同值时，在最佳策略下内容商的收益随预算值的变化情况，从图中可以看出，随着预算的增加，内容商购买了更多的网络资源为用户提供服务，内容商的收益不断增加。当内容流行

度更为集中时, 购买缓存空间可以在边缘响应更多的用户请求, 因此随着缓存空间的不断增加, 内容商收益增长速度较快, 其收益远高于流行度分布稍为分散的系统。

### (3) 网络资源价格影响

单位网络资源的价格是影响内容商收益的主要因素之一, 缓存和带宽的购买量不仅影响为用户提供服务的能力, 同时决定内容商的开支。当单位容量的资源价格发生变化时, 内容商的利润率将会随之波动。在接下来的仿真中, 分别比较了缓存和带宽单价对于最佳购买策略及其对应的收益的影响。

图3-4包含当单位缓存价格为  $p_2 = 1.5$  的最佳资源策略随预算的变化情况, 及其对应收益与  $p_2 = 1$  时相应预算下收益的对比, 即分析单位缓存空间的价格更加昂贵时的情况。从图3-4与图3-1的对比可以看出, 由于缓存单价的变化, 导致单位缓存资源带来的收益的变化, 随着预算的增加, 缓存空间购买量的变化趋势稍缓于带宽购买量的变化趋势, 即在预算较低时, 最佳策略首先偏向于购买较多的带宽资源。随着预算的增加, 最佳策略逐渐偏向于购买缓存, 缓存较先达到购买量的最大值, 最终随着预算的持续增加, 带宽的购买量逐渐达到最大值。随着网络资源容量的增加, 它们为内容商带来收益的增长速度不同, 缓存的收益增长速度略高于带宽资源, 因此其最佳购买策略先于带宽资源达到最大值。从图3-4可以看出, 由于单位缓存资源价格的的增长, 内容商为用户提供服务的成本增加, 而单位视频的收益没有提高, 故随着预算的增加, 内容商收益的增长趋势较缓存价格较低时的增长趋势有明显下降。

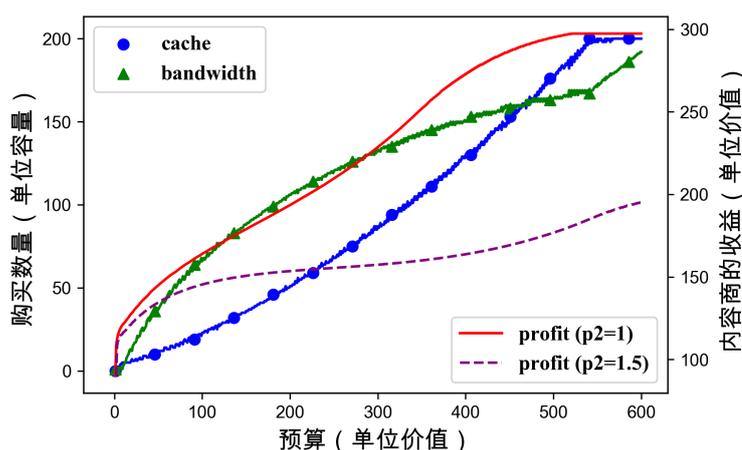


图 3-4 缓存价格对最佳策略的影响

Fig 3-4 Affect of cache price on the best strategy

图3-5包含当单位缓存资源的价格参数  $\gamma = 0.75$  时的最佳资源策略随预算的变化情况, 及其收益与  $\gamma = 0.9$  时相应预算下的收益的对比情况, 即分析单位带宽价格随购买带宽量的变化速率稍慢一些的情况。从图3-5中可以看出, 由于单位带

宽的价格的变化趋势变缓，在同样的预算情况下，内容商可以购买更多的带宽资源，从而提高为用户提供服务的能力，即图3-5所示带宽购买量随预算增加而增加的速度较图3-1快。图3-5所示为相较于  $\gamma = 0.9$  的带宽价格函数的内容商收益的比较曲线，由于带宽价格较为便宜时，内容商可以在对应预算的情况下购买更多的网络资源，提高服务能力，提升传输视频流量带来的收益，因此其收益的曲线增长速度较带宽单位价格高的系统快。

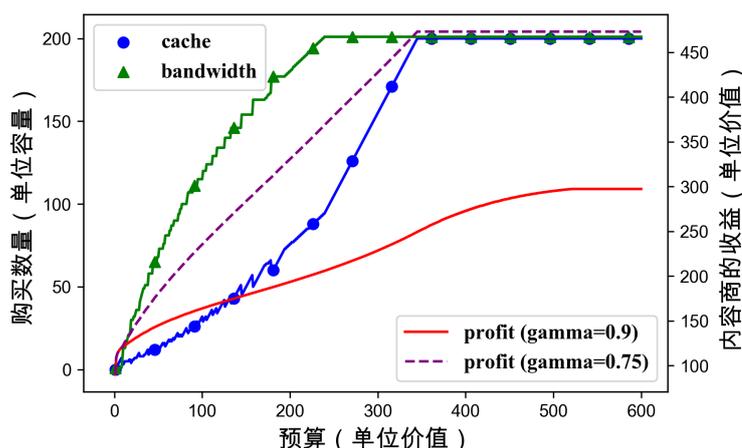


图 3-5 带宽价格对最佳策略的影响

Fig 3-5 Affect of bandwidth price on best strategy

### 3.5 本章小结

本章中首先对研究的带有缓存的视频传输系统进行介绍，对其主要的三个参与者：内容提供商、网络提供商和终端用户及其之间的关系进行梳理和建模。然后对内容商在有预算限制时的资源购买方案和效益最大化问题进行数学建模分析。随后根据现有的研究及其资源购买特性，提出最佳资源购买策略求解算法，并通过仿真分析在不同预算时，随着内容流行度及网络资源价格的变化，内容商的最佳资源购买策略及其对应的收益的变化趋势。

## 4 含广告占比及位置影响的效益优化

在现有的内容商的盈利策略中，“免费内容 + 商业广告”的形式仍在广泛使用并被观众接受。商业广告的投放比例、投放位置均会对观众的体验造成一定程度的影响。在这种模式下，内容商的视频内容及其提供的视频传输质量影响用户观看的视频数据量，而其中广告的投放为其主要收入来源，因此广告投放策略对于内容商提升整体收益非常重要。本章主要分析当采用该计费机制时的广告内容投放比例及投放位置对与内容商收益的影响。

本章的主要内容如下，4.1 节对研究场景进行介绍，4.2 节介绍引入商业广告对用户的观看行为以及内容商收益的影响，4.3 节将该问题建模为最大化内容商收益的优化问题，并对投放比例及投放位置进行理论分析，4.4 节对优化问题分析的结果进行数值仿真验证，最后在 4.5 节对本章内容进行总结。

### 4.1 研究场景

在现有的视频内容商收费策略中，主要由“免费内容 + 商业广告”与“付费订阅免广告”两种模式相结合。内容商除了向终端用户提供视频资源外，还会投放一定比例的商业广告以获取广告利润。免费的视频内容及良好的视频播放质量可以吸引观众观看内容商的视频，在视频播放的起始位置或者中间位置插入一定量的广告可以获得广告利润，该部分费用由广告主向内容商支付。

在这样的系统中，用户将可以划分为以下三类：

- (1) 因为视频中的广告离开观看；
- (2) 观看广告并继续观看免费的视频内容；
- (3) 为视频内容付费成为内容商的订阅用户，从而跳过广告。

不同类型的用户对内容商收入的影响是不同，其中因为广告投放而离开的用户不会给内容商带来收益；而观看广告后观看视频的用户将会为由广告主给内容商支付广告投放费用，从而带来收益，其收益取决于用户观看的视频总数据量以及其中的广告内容的占比；而内容商的订阅用户则不会观看广告，直接为内容商的视频内容进行付费。在本章中，主要关注的是前两类用户对内容商的效益的影响，其观看内容商视频的概率不仅与视频内容的质量相关，同时与广告的投放策略相关。

当商业广告的投放比例及投放位置不同时，对用户整体观看过程的影响是不同的，所造成的用户离开概率同样会发生变化，而内容商的广告收益随之收到影响。本章主要对广告投放比例及位置对于用户离开概率及系统内用户的视频总需求量的影响进行分析，从而分析广告投放策略对内容商的收益的影响，并通过仿

真实实验对分析结果进行验证讨论。

## 4.2 系统模型及优化建模

在投放广告的视频播放系统中，用户根据内容商提供的视频内容及其播放质量决定是否选择观看该内容商的视频。在用户向内容商请求某视频内容后，内容商在视频的播放过程中投放一定比例的广告内容，用户可以选择观看广告后继续观看其请求的视频，或者直接离开视频播放系统。用户请求视频内容后离开系统的原因包括两类：视频内容或其传输质量或者广告内容的投放，因此产生用户有效观看需要满足两个条件，即内容商资源策略所能实现的视频传输质量和内容商广告投放策略对用户造成的影响均未导致用户离开视频播放系统，只有当两者均满足用户的要求时，才会产生可以为内容商带来利润的有效观看。当商业广告投放放在视频的中间位置时，用户的离开行为可能发生在视频播放过程的中间，在“免费内容 + 商业广告”的计费模式下，用户中途离开系统，内容商无法对用户离开之前的观看内容进行收费，因此其系统中的流量及盈利会受到广告内容投放比例及位置这两个因素的影响。下面将分别分析广告投放比例和广告投放位置对于用户行为及内容商收益的影响。

### 4.2.1 商业广告投放比例

在“免费内容 + 商业广告”的盈利模式下，广告主可以购买内容商在视频内容开始播放过程中的广告位，并根据其播放时长、曝光次数、点击转化次数等相关指标向内容商支付费用。内容商则根据自身的用户特征及资源策略制定相应的广告投放方案，以最大化自身的效益。在这样的模式下，用户可以通过观看广告内容而免费观看到其请求的视频内容，但是用户的观看体验会受到广告投放量的影响，进而导致其离开观看系统。内容商在根据其用户特征进行投放时，既希望尽可能多的展示广告主的内容，提高曝光率以获得更多的广告收益，又希望不至于造成太多的用户流失，失去其原有的用户基础。因此内容商在选择投放比例时，需要对影响其收益的因素进行综合考虑。

#### (1) 系统平衡时的总流量

首先分析含有广告投放的系统达到平衡时传输的总数据量的变化。将系统内的用户作为一个整体进行分析时，可以通过用户因为广告内容离开概率分布、累积概率分布、平均值等统计信息分析和制定广告投放策略。首先对用户因为广告投放比例离开的概率作出假设。若系统内的广告投放比例记为  $\tau$ ， $\tau \in [0, \tau_{max}]$ ， $\tau_{max} \leq 1$  时，将用户因广告离开的累积概率记为  $l(\tau)$ ，其中  $\tau_{max}$  表示系统内最大的广告投放比例。

**假设 1:** 随着广告占比的上升, 用户的累积离开概率先缓慢上升, 然后迅速上升, 最终逐渐趋近于 1, 即  $l(\tau) \in [0, 1], l'(\tau) > 0$ 。

假设 1 表示, 当系统中广告的比例逐渐上升时, 用户起初的平均离开概率增长速度较慢, 因为其观看广告后的预期收益较高, 即观看广告后可观看的视频内容较大; 而随着广告占比继续增大时, 用户的累积离开概率迅速增加并逐渐趋近于 1, 即随着广告投放占比的逐渐增大, 用户后续可观看视频量减小, 故造成用户离开概率增加, 即用户的离开概率与内容中商业广告占比量呈正相关。

播放过程中广告的投放会影响用户在系统内的平均视频需求量, 即在相应的资源策略  $(b, c)$  下, 系统内吞吐量达到稳定时的总流量  $T(b, c)$  同样会随着广告占比  $\tau$  的变化而变化, 因此在计算系统稳定时的流量时需要考虑商业内容占比因素, 将其记为  $T(b, c, \tau)$ 。

当商业内容投放在视频内容开始播放之前时, 用户观看视频内容需要满足两个条件: 对内容商所能提供的网络服务质量满意, 且未因为视频中广告投放的占比离开观看系统, 即假设用户因为服务质量选择内容商与因为商业内容离开系统两者之间无关。用户在发送视频请求后, 最终观看的概率为  $\eta(\bar{d}) \cdot (1 - l(\tau))$ 。假设用户选择观看后的平均需求量仍为  $r(\bar{d})$ , 则系统中平均每个用户的视频需求则为  $f(\bar{d}, \tau) = \eta(\bar{d}) \cdot (1 - l(\tau)) \cdot r(\bar{d})$ 。根据公式 (3-1)-(3-5) 可得含有商业广告的系统在达到平衡时的系统内总流量如公式 (4-1) 所示:

$$T(b, c, \tau) = \frac{\beta n \cdot (1 - l(\tau)) \cdot (b_0 + b)}{(1 + \beta n \cdot (1 - l(\tau))) \cdot M(c)} \quad (4-1)$$

## (2) 内容商的收益

在“免费内容 + 商业广告”的模式下, 内容商的盈利来自于广告主支付的推广费用, 故其收入取决于数据量中的广告占比、广告的单位价格及总的流量。将单位流量的广告价格记为  $p_a$ , 则广告收益为投放占比的函数。系统达到平衡时的广告收入如公式 (4-2) 所示。

$$I(b, c, \tau) = p_a \cdot \tau \cdot T(b, c, \tau) \quad (4-2)$$

## (3) 优化问题建模

当内容商的资源策略  $(b, c)$  给定时, 根据其收益函数, 可以将寻找最优广告投放比例的问题建模为优化问题, 其优化目标为最大化内容商收益, 优化变量为广告占比, 其中需给出用户因为广告占比离开视频观看系统的累积概率分布。在给定资源策略  $(b, c)$  时内容商的效益为其广告收入减去购买资源的支出, 结合如式 (4-1)-(4-2) 可得, 其对应的效益为广告插入比例  $\tau$  的函数, 如公式 (4-3) 所示。

$$U(b, c, \tau) = p_a \cdot \tau \cdot T(b, c, \tau) - (p_b(b) \cdot M(c) \cdot T(b, c, \tau) + p_c(c) \cdot c) \quad (4-3)$$

根据内容商的收益方程，广告插入比例的优化问题可以建模为如下形式，假设资源策略  $(b, c)$  已经给出。

$$\text{Maximize}_{\tau} U(b, c, \tau) \quad (4-4a)$$

$$\text{Subject to } l(\tau) \geq 0 \quad (4-4b)$$

$$0 \leq \tau \leq 1 \quad (4-4c)$$

其中约束 (4-4b) 为广告占比造成的用户离开概率曲线，约束 (4-4c) 为广告在视频内容中的占比。

在实际的广告投放方式中，广告的时长有一些常见的标准长度，如 15s、30s、60s 等，广告的投放长度是这些标准长度的组合<sup>[41]</sup>。在视频前面播放的广告长度通常较短，因此占比趋近于 1 的情况较少，这一影响因素体现在在用户离开概率的约束中，即当  $\tau$  趋近于 1 的时候， $l(\tau)$  趋近于 0。为了简化优化问题的求解过程，在分析过程中将广告的占比约束上限设置为 1，并使用连续变量刻画其具体占比，便于求解优化问题的最优解。

#### (4) 理论分析

在本部分中，根据公式 (4-4) 刻画的优化问题，并结合其资源约束、用户累积离开概率等约束条件，对于优化问题的求解进行进一步分析。将内容商收益公式 (4-3)、系统内的流量公式 (4-1) 带入优化目标函数可得下式 (4-5)：

$$\begin{aligned} U(b, c, \tau) &= (p_a \cdot \tau - p_b(b) \cdot M(c)) \cdot T(b, c, \tau) - p_c(c) \cdot c \\ &= (p_a \cdot \tau - p_b(b) \cdot M(c)) \cdot \frac{\beta n \cdot (1 - l(\tau)) \cdot (b_0 + b)}{(1 + \beta n \cdot (1 - l(\tau))) \cdot M(c)} - p_c(c) \cdot c \end{aligned} \quad (4-5)$$

当内容商的资源策略已经给出时， $U(b, c, \tau)$  可以记为  $\tau$  的函数，即  $U(\tau)$ 。将公式 (4-5) 中与  $b, c$  相关的变量及函数用常数表示。根据 (4-5) 作出如下简化表示，令  $C_1 = p_b(b) \cdot M(c)$ ， $C_2 = \frac{b_0 + b}{M(c)}$ ， $C_3 = p_c(c) \cdot c$ ，并记  $T_1(\tau) = \frac{\beta n \cdot (1 - l(\tau))}{1 + \beta n \cdot (1 - l(\tau))}$  则优化目标函数可以记为公式 (4-6)：

$$\begin{aligned} U(\tau) &= (p_a \cdot \tau - C_1) \cdot \frac{\beta n \cdot (1 - l(\tau))}{1 + \beta n \cdot (1 - l(\tau))} \cdot C_2 - C_3 \\ &= (p_a \cdot \tau - C_1) \cdot T_1(\tau) \cdot C_2 - C_3 \end{aligned} \quad (4-6)$$

根据公式 (4-4) 中的约束变量及化简后的优化目标，引入  $l_1, l_2$  为求解引入的拉格朗日乘子，可得优化问题的拉格朗日函数如下：

$$L(\tau, l_1, l_2) = (C_1 - p_a \cdot \tau) \cdot T_1(\tau) \cdot C_2 + C_3 + l_1(\tau - 1) - l_2 \tau \quad (4-7)$$

根据拉格朗日函数可以求的广告占比优化问题的 KKT 条件，包含其原问题

及对偶问题的可行性条件、互补对称性条件及稳定性条件如下式所示：

$$\begin{cases} -p_a \cdot T_1(\tau) + C_2(C_1 - p_a \cdot \tau)T_1'(\tau) + l_1 - l_2 = 0 \\ T_1(\tau) = \frac{\beta n \cdot (1 - l(\tau))}{1 + \beta n \cdot (1 - l(\tau))}, T_1'(\tau) = \frac{-\beta n \cdot l'(\tau)}{(1 + \beta n \cdot (1 - l(\tau)))^2} \\ l_1(\tau - 1) = 0, l_2\tau = 0 \\ \tau \leq 1, \tau \geq 0 \\ l_1 \geq 0, l_2 \geq 0 \end{cases} \quad (4-8)$$

当求解该优化问题的最优解时，满足 KKT 条件是可行解为最优解的必要条件，即需要求得该问题的边界解及内部解，通过比较其目标函数值判断最优解的具体值。当  $l_1 \neq 0, l_2 \neq 0$  不同时成立时，可得优化问题的两个边界解，此时系统中不含广告或者视频内容均为广告，通过公式 (4-5) 可得此时内容商的收益。当  $l_1 = 0, l_2 = 0$  时，根据互补松弛条件可知，求得的最优解满足优化问题的广告占比的不等式，即视频播放过程中有广告插入，且并非全部视频内容均为广告，因此在求解广告投放比例的最优内部解时，有  $l_1 = 0, l_2 = 0$ ，满足下式即可满足稳定性条件。

$$-p_a \cdot (1 - l(\tau))(1 + \beta n(1 - l(\tau))) + (p_a \cdot \tau - C_1) \cdot l'(\tau) = 0 \quad (4-9)$$

当给定单位广告内容的价值和用户因为广告投放的累积离开概率曲线时，求解最优的广告投放比例的内部解即为求解使得公式 (4-9) 成立的  $\tau$  值，进而可以通过公式 (4-5) 计算得到内容商在相应的资源策略及广告投放方案下所能获得的收益，将其收益与边界收益进行比较可以得到优化问题的最优解及目标函数的最优值。

#### 4.2.2 商业广告投放位置

当商业广告投放在视频播放的不同位置时，用户的离开概率会根据位置发生改变。当商业广告的投放位置位于视频的中间时，用户可以在请求视频内容后直接开始观看，而在其视频播放过程中插入广告，用户因为广告的离开概率则会大幅降低，因为视频内容具有连贯性，前面的内容播放使得用户更倾向于去观看完整的视频，因此用户会观看完整的广告，从而继续观看剩余的视频<sup>[7]</sup>。用户可以在观看广告内容前提前观看部分视频内容，使得其观看内容商视频的机会成本有所下降，该设置在一定程度上可以提升用户体验，增加内容商系统的用户粘性。因此商业广告内容投放位置的选择对于内容商的服务体验及广告收益均会带来影响。

##### (1) 对系统内流量的影响

由于广告投放位置不再位于视频开始播放前的位置，用户已经观看了部分的视频内容，根据用户观看视频的自然离开概率，用户已经表现出其是否会继续观

看后续视频内容，因此用户观看完广告内容并继续观看后续视频内容的概率会增加。

**假设 2:** 随着广告插入在视频中位置的比例  $t$  的提升，用户对广告插入时长的敏感度有所下降，即用户可以接受的平均广告时长上升，且随着广告时长增加离开视频观看系统概率的下降速率逐渐下降。

根据文献 [5,6]，用户在观看视频时，并非所有的用户均会观看完整的视频内容，用户在可能在观看过程中离开视频观看系统，因此一段视频内容起始片段的播放量远高于其结尾片段的播放量。用户离开观看视频的概率会随着播放时长而逐渐下降，最终趋近于 0。

**假设 3:** 随着播放比例  $t$  的提升，用户自然离开视频播放系统的概率逐渐下降，最终逐渐下降为 0，且下降的速率逐渐缓慢，即  $v(t) \in [0, 1]$ ， $v'(t) < 0$ 。

当将商业广告投放在视频内容的中间位置时，虽然用户因为广告内容离开系统的概率有很大程度的下降，但是由于用户观看过程中的离开，使得观看到广告开始阶段的总用户数也会有相应的下降，并且在这样的模式下，用户可以直接观看到视频内容，在系统达到平衡时，在广告投放位置之前传输到视频内容并没有收益，也就是说，仅有用户观看到视频内容投放位置之后的内容时，内容商才会得到相应的广告收益，因此内容商在不同位置投放商业广告时，需要在播放的免费内容的占比及广告收益之间作出折中选择。

当系统内用户因为网络质量对内容商的视频需求量给出时，假设因为广告投放离开观看系统的用户会观看广告投放位置之前的内容，而并未因广告投放离开系统的用户则会观看其请求的所有视频内容。由于用户的自然离开概率并非播放量的线形函数，因此在广告插入前播放的视频量并非随着其在视频播放过程中的位置线性增加，故而可以根据用户对视频的总需求量及观看视频的自然离开概率，计算在考虑位置因素时，系统达到平衡时系统内总的视频内容需求量。根据上述分析，系统达到平衡时的总数据量将划分为两部分，观看完整视频及广告的用户需求的数据量和仅广告播放前视频的用户，综合可得系统达到平衡时所需的视频量如下式 (4-10) 所示：

$$T(b, c, \tau, t) = \frac{(\beta n \cdot (1 - l(\tau, t)) + \beta n \cdot l(\tau, t) \cdot \int_0^t v(\epsilon) d\epsilon) \cdot (b_0 + b)}{(1 + \beta n \cdot (1 - l(\tau, t)) + \beta n \cdot l(\tau, t) \cdot \int_0^t v(\epsilon) d\epsilon) \cdot M(c)} \quad (4-10)$$

其中  $v(\epsilon)$  表示用户观看到视频的比例  $\epsilon$  时的离开概率，其概率分布满足归一性，即  $\int_0^1 v(\epsilon) d\epsilon = 1$ 。

## (2) 内容商收益分析

当商业广告的内容不再投放在视频内容开始播放之前时，用户在观看广告之前仍可以观看部分视频内容，即表示并非平衡时系统内所有的视频数据量均可以获得相应的收益，仅观看完整的视频和广告内容部分的用户所产生的数据量中含

有相应比例的广告内容，故根据公式 (4-10) 计算所得的系统平衡时流量、广告占比及单位商业广告内容的收益可计算得到在考虑投放位置因素影响下的内容商总收益如公式 (4-11) 所示：

$$I(b, c, \tau, t) = p_a \tau \cdot \frac{\beta n \cdot (1 - l(\tau, t)) \cdot (b_0 + b)}{(1 + \beta n \cdot (1 - l(\tau, t)) + \beta n \cdot l(\tau, t) \cdot \int_0^t v(\epsilon) d\epsilon) \cdot M(c)} \quad (4-11)$$

其中  $p_a$  为商业广告的单位价格。

### (3) 优化问题建模

当内容商的资源策略、用户的自然离开概率及因为广告离开的概率给出时，求解最优广告投放位置及投放比例的问题同样可以刻画为优化问题，其优化目标为最大化内容商的收益，优化变量为广告的投放占比及投放位置，结合公式 (4-3) 可得，该模型下内容商的效益函数为：

$$U(b, c, \tau, t) = I(b, c, \tau, t) - [p_b(b) \cdot M(c) \cdot T(b, c, \tau, t) + p_c(c) \cdot c] \quad (4-12)$$

结合资源策略、效益函数及用户的离开概率，商业广告投放位置及比例的问题可以建模为如下形式：

$$\text{Maximize}_{\tau, t} U(b, c, \tau, t) \quad (4-13a)$$

$$\text{Subject to } l(\tau) \geq 0 \quad (4-13b)$$

$$v(\epsilon) \geq 0 \quad (4-13c)$$

$$0 \leq \tau \leq 1 \quad (4-13d)$$

$$0 \leq t \leq 1 - \tau \quad (4-13e)$$

其中约束 (4-13b) 为广告占比造成的用户离开概率曲线，约束 (4-13c) 用户的自然离开概率曲线，约束 (4-13d) 为广告在视频内容中的占比，约束 (4-13e) 为商业广告在视频播放过程中的投放位置。

在广告投放位置的约束中，其表示的为商业广告开始播放的起点，为保证必须在整个视频播放结束前完成整个广告内容的播放，根据广告占比变量的值，其播放位置必须早于  $1 - \tau$  的值，以确保内容播放的完整性。

### (4) 理论分析

在本部分中，对公式 (4-13) 建模的优化问题进行分析和求解。与公式 (4-4) 优化问题的目标函数整理和化简方式相同，将优化目标变量中的资源策略  $(b, c)$  固定，分析广告位置  $t$ ，广告占比  $\tau$  的效益函数可得：

$$U(\tau, t) = \frac{(C_2 p_a \tau - C_1 C_2) \beta n (1 - l(\tau, t)) - C_1 C_2 \beta n l(\tau, t) \int_0^t v(\epsilon) d\epsilon}{1 + \beta n (1 - l(\tau, t)) + \beta n l(\tau, t) \int_0^t v(\epsilon) d\epsilon} - C_3 \quad (4-14)$$

其中,  $C_1 = p_b(b) \cdot M(c)$ ,  $C_2 = \frac{b_0+b}{M(c)}$ ,  $C_3 = p_c(c) \cdot c$ 。

根据优化问题建模公式 (4-13) 及目标函数公式 (4-14), 引入拉格朗日乘子  $l_1$ 、 $l_2$ 、 $l_3$ 、 $l_4$ , 可得优化问题的拉格朗日函数如下式:

$$L(\tau, t, l_1, l_2, l_3, l_4) = -U(\tau, t) + l_1(\tau - 1) - l_2\tau + l_3(t + \tau - 1) - l_4t \quad (4-15)$$

从而可推导得到最优解应满足的 KKT 条件:

$$\begin{cases} \frac{\partial U(\tau, t)}{\partial \tau} - l_1 + l_2 - l_3 = 0 \\ \frac{\partial U(\tau, t)}{\partial t} - l_3 + l_4 = 0 \\ l_1(\tau - 1) = 0, l_2\tau = 0, l_3(t + \tau - 1) = 0, l_4t = 0 \\ \tau \leq 1, \tau \geq 0, t + \tau \leq 1, t \geq 0 \\ l_1 \geq 0, l_2 \geq 0, l_3 \geq 0, l_4 \geq 0 \end{cases} \quad (4-16)$$

当  $l_1, l_2, l_3, l_4$  均取值为 0 时, 求解 KKT 方程所得的解为该优化问题的内部解, 此时系统内有商业广告投放, 且并非所有的视频内容均为广告, 且其投放位置位于视频播放过程中, 并非开头或结尾位置。当  $l_1, l_2, l_3, l_4$  的取值不全为 0 时, 其对应的解为优化问题的边界解。当求解系统内的最优解时, 通过将内部最优解及边界解的目标函数值进行比较, 从而可获得全局最优解。

在实际优化投放位置和占比的广告插入问题中, 其投放位置并非可以在视频中任意位置进行连续变化, 而是根据视频播放过程中的场景转换、情节变化进行投放位置的选择。当投放位置确定后, 根据用户对于广告接受程度等因素, 选择广告插入时长。因此在求解上述优化问题的内部最优解时, 可以根据给定的投放位置, 求解 KKT 方程得到最佳投放比例, 并将不同投放策略时的内容商收益进行比较, 从而获得最优解。

### 4.3 数值仿真

在本节中, 首先考虑在视频开始时插入广告内容的数值仿真, 分析优化广告内容占比过程中内容商的收益变化; 然后将广告内容播放位置因素的影响纳入考虑, 进一步对广告占比和位置共同对内容商收益的影响进行分析。

#### 4.3.1 广告在视频中占比的影响

在本小节中, 首先为不同广告占比时用户离开概率的曲线进行建模; 然后根据不同用户离开概率曲线分析达到平衡时系统内的数据量及其相对应的内容商的收益; 最后分析用户需求对广告最优投放比例的影响, 并分析在不同用户需求情

况下系统采用最优广告占比时，达到平衡时系统内数据量及内容商收益的变化。

### (1) 用户因广告的离开概率

根据用户的离开特性假设 1，对用户的离开概率进行数学建模假设，采用一个拟 sigmoid 函数表示随着广告比例增加离开观看视频的累积概率分布，即公式 (4-17)。

$$l_{\alpha,\delta}(\tau) = \frac{1}{1 + e^{-\frac{(\tau-\alpha)}{\delta}}} \quad (4-17)$$

其中含有  $\alpha, \delta$  两个超参数，以调整适应不同的离开函数曲线。在公式 (4-17)， $x$  表示视频中的广告占比， $\alpha, \delta$  表示用户群体对于视频中广告比例的平均接受程度及接受程度的方差。

当用户离开概率曲线中参数  $\alpha$  和  $\delta$  设置为不同的参数时，用户离开概率随着广告占比的提升呈现不同的曲线，如图 4-1 所示。图中采用三组不同的超参数绘制了相应的用户离开概率曲线，横轴表示视频中投放的商业广告内容的比例，纵轴表示用户在播放视频根据广告比例的累积离开概率分布。

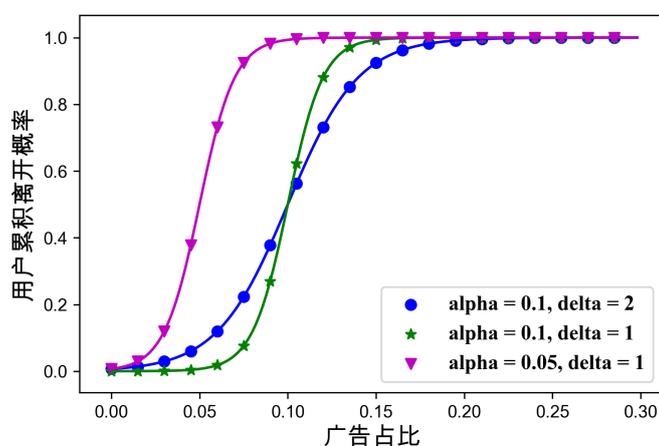


图 4-1 用户离开概率

Fig 4-1 User leave rate

对于每一条曲线，随着广告内容占比的提升，用户的累积离开概率不断提升。在开始阶段提升较慢，超过一定门限值后，离开概率迅速提升，最后缓慢趋近于 1。用户离开概率曲线表示用户对于少量的广告内容可以接受，随着广告内容的增加，用户的离开概率增加，但仍有小部分用户为观看后续的视频内容会坚持观看广告内容。

图中所示曲线间，当  $\alpha$  相同， $\delta$  不同时，表示两个用户群体对于广告占比的平均接受程度相近，但是用户对于不同比例的接受程度的分布不同，其中  $\delta$  越小，表示，用户的接受程度更多的集中在平均值附近；当  $\alpha$  不相同， $\delta$  相同时，表示

两个用户群的对于广告内容占比的平均接受程度不同，但是接受程度在均值附近的集中程度相似。

## (2) 系统平衡时的数据量及内容商收益

根据公式 (4-1) 和公式 (4-3) 当给定内容商的资源策略时，可以计算得到当系统达到平衡时传输的数据量。当给定内容商的资源策略为缓存容量  $c = 200$ ，带宽容量  $b = 200$  时，根据不同的离开概率特性计算可得在不同广告占比时系统达到平衡时的数据量及对应的内容商收益分别如图4-2及图4-3所示。其他仿真参数设置如下， $q_a(\tau) = 20$ ， $p_c(c) = 1$ ， $p_0 = 0.01$ ， $\gamma = 0.9$ ， $\alpha = 0.75$ ， $\beta n = 3$ 。

在图4-2中，将“付费内容”和“商业广告 + 免费内容”两种模式下系统达到平衡时的数据量进行比较。从仿真结果中可以看出，当内容商的计费策略为订阅付费模式时，给定其网络资源的购买策略，系统达到平衡时系统内的流量为常数，且与有广告插入模式的上限相同，该上限取决于内容商此时的服务能力。随着视频中广告占比的提升，系统达到平衡时的数据量逐渐下降，其下降速度与用户因为广告的离开速率相关。随着  $\alpha$  的增大，平衡时数据量快速下降的起点逐渐后移，而随着  $\delta$  的增大，商业广告的可投放比例区间逐渐变大，且迅速下降期的速率变缓，这是由于用户对于广告占比的接受程度的分布较为分散，并非集中在平均接受区间附近。

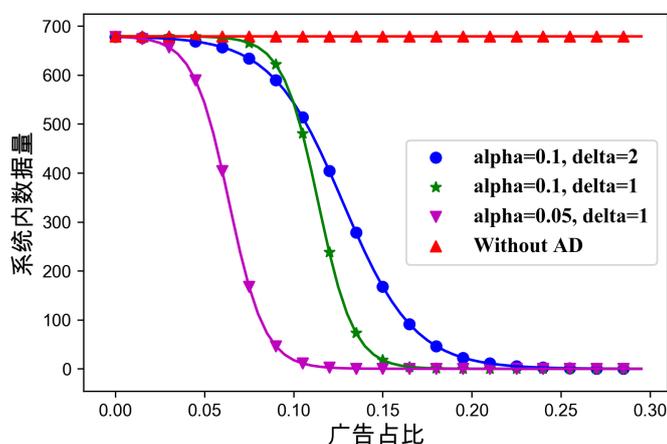


图 4-2 系统平衡时数据量

Fig 4-2 Traffic in the system

在图4-3中，在“付费内容”模式下的内容商收益保持为常数，因为当给定内容商购买的缓存空间和带宽容量时，系统的服务能力固定，系统内的数据量稳定为常数，故其收益值固定。在“商业广告 + 免费内容”模式下，随着广告占比的提升，内容商的收益首先开始逐渐提升，到达峰值后，随后其收益逐渐下降，最终逐渐变为负值，对应为购买网络资源的支出。如图4-3所示，当  $\alpha$  和  $\delta$  取不同的

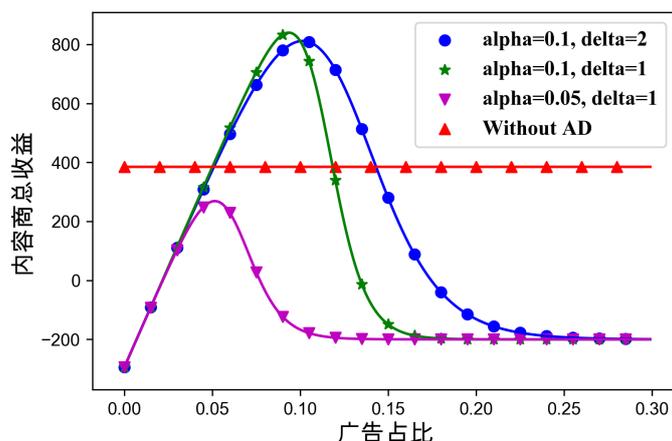


图 4-3 内容商收益

Fig 4-3 Profit of content provider

值时，随着广告占比的提升，内容商的收益在初始阶段的增长趋势相似，因为在不同的离开概率分布下，用户在广告占比增长的初始阶段的离开概率相近，但是由于此时广告占比较小，广告内容盈利较少；随着广告占比的进一步提升，在用户的离开概率未快速上升之前，内容商的收益随着广告比例的提升得到逐步提升；随后，用户离开概率随着广告比例的提升而迅速增加，内容商的收益开始下降，最终用户离开概率逐渐趋近于 1，内容商的广告收入趋近于 0，其利润为负，等于购买网络资源的支出。

从图4-3可以看出，不同的用户离开特性曲线所对应的内容商收益均存在最大值，其最大值与广告投放比例和用户离开特性有关，当用户对广告的平均接受度较低时，内容商所能达到的最大利润有限；当用户对广告的平均接受度高，且接受程度越集中在平均值，内容商所能实现的最大收益越大，最佳投放比例离接受度平均值越近；当用户对广告接受程度较为分散时，内容商所能实现较高收益的投放区间较大，此时内容商的投放策略可以将更多的因素纳入考虑，从而将利润值与广告内容相关的因素相结合，实现多重投放目标的均衡。

### (3) 用户需求对于最优投放比例的影响

从公式 (4-5) - (4-9) 可知，在给定内容商资源策略、资源和视频单位流量价格的情况下，内容商的收益主要受用户群体对于内容商视频的需求量的影响，接下来对于不同用户需求下内容商的最佳投放比例的变化趋势，及相应的最佳投放比例下的系统平衡时数据量及内容商收益进行分析。

在仿真中，各项参数选择与前述实验相同，主要分析  $\alpha = 0.05/0.1$ ， $\delta = 1$  时的最佳投放比例、系统内数据量及内容商收益。当仅考虑服务质量时的用户需求  $\beta n \in [0, 20]$  时，依次求得不同用户需求下的最佳广告投放比例如图4-4所示。

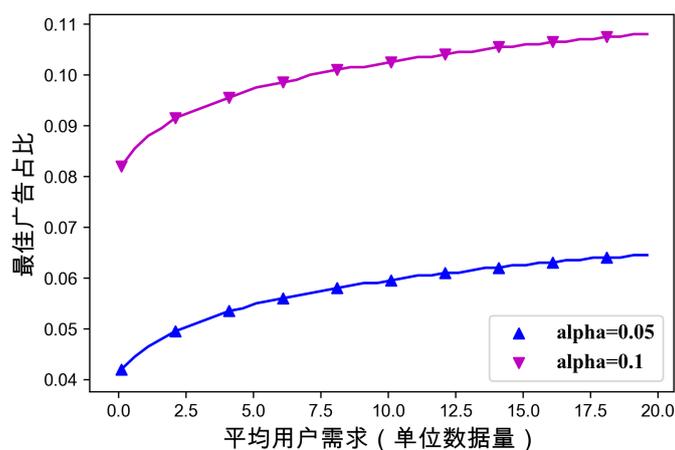


图 4-4 最佳投放比例

Fig 4-4 Best AD proportion

从图4-4可以看出，随着用户需求的逐渐提升，最佳的广告投放比例也逐渐提升，但是上升速率逐渐下降，且用户对广告平均接受程度高的群体的最佳投放比例明显高于平均接受程度低的群体，当用户群体的平均接受度较高时，内容商可以通过投放较多的广告以获取更多的利润；而当用户对于内容商的视频需求量较高时，虽然会有部分用户会因为广告投放而离开观看系统，但是留下的用户会观看更多的视频内容，从而较高的广告投放比例也可以使得内容商获得更多收益。

随着用户需求的提升，在最佳广告投放比例下，平衡时系统内的数据量和相对应的内容商收益如图4-5、4-6所示。从图4-5中可知，在两种计费模式下，随着用户需求的提升，达到平衡时系统内的数据量的变化趋势相同，均随着用户需求的提升呈非线性增长，且增长速度逐渐变慢。含有广告投放的系统中平衡时的流量总是低于不含广告的系统，其原因是在同样的用户需求的情况下，部分用户会因为广告内容离开观看系统；在相同的用户需求下，用户对广告平均接受度高的系统中平衡时数据量较高于用户接受度较低的，这是由于在各自系统的最优广告投放比例下，两个系统内的用户因为广告离开系统的概率不同。

图4-6中为在不同用户需求下，平衡时的数据量下内容商所能获得的利润。从图中可知，随着用户需求和数据量的提升，内容商的收益随之增长，但是不同系统中所表现出的增长速度不同。当计费模式为“付费内容”时，内容商收益随着用户需求量的增加首先迅速增加，随后逐渐趋于稳定，这是由于收到购买的网络资源量的限制，系统内达到平衡时的数据量趋于稳定。当用户群体广告接受度较低时，内容商在用户需求量较小时的最佳收益小于内容计费模式，因为在内容商传输的数据量中所含的广告内容量较小，在支付网络资源费用后的盈利较少，随着用户需求的逐渐增加，内容商的收益逐渐增加，但是受限于用户所能接受的广告投放量，因此其利润的增长幅度有限；当用户对广告接受程度较高时，随着用户需

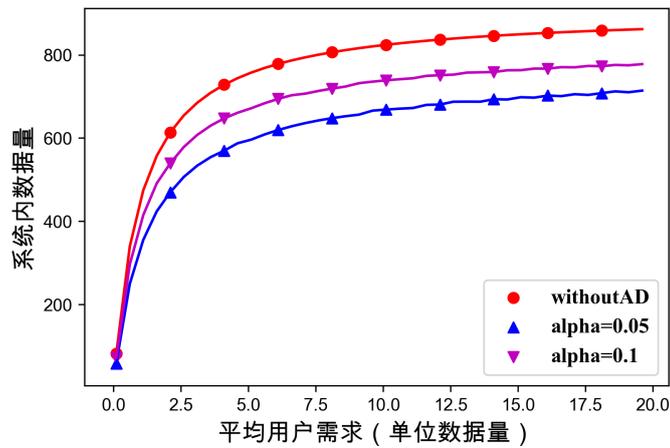


图 4-5 系统内的数据量 (用户需求)

Fig 4-5 Traffic in the system (user demand)

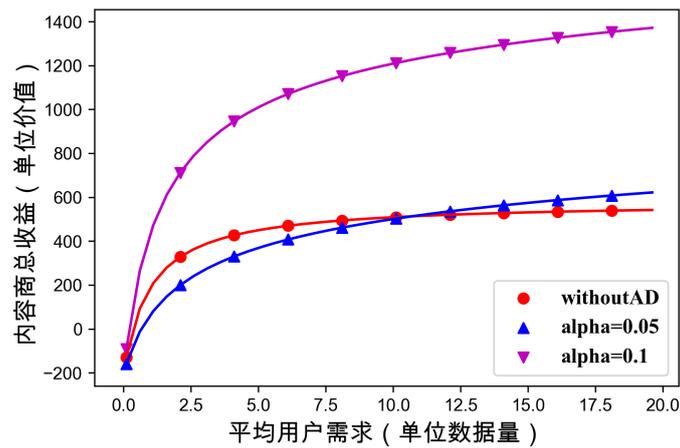


图 4-6 内容商的收益 (用户需求)

Fig 4-6 Profit of content provider (user demand)

求量的增加，其利润的增速均高于内容计费模式，由于随着用户需求的增加，内容商可以投入更高比例的广告以获取利润，此时有较小比例的用户，坚持观看广告及后续的视频内容，从而使得内容商的单位数据量的视频可以获得更多的收益。

#### 4.3.2 内容占比及位置联合影响

在本小节中考虑将不同比例的商业广告投放在视频播放过程中的不同位置，通过仿真分析其对系统内容的数据量和内容商的收益的影响。在仿真中固定内容商购买的资源策略、用户需求和资源价格，参数值设置与上一小节相同，改变视频内容中的广告占比及投放位置，并设置用户因为广告离开观看系统的累积概率分

布的参数为  $\alpha = 0.1$ ,  $\delta = 1$ , 用户观看视频的自然离开概率设置为  $v(\epsilon) = \frac{\exp\{-\frac{\epsilon}{13}\}}{13}$ , 仿真结果如下图4-7 - 图4-9。

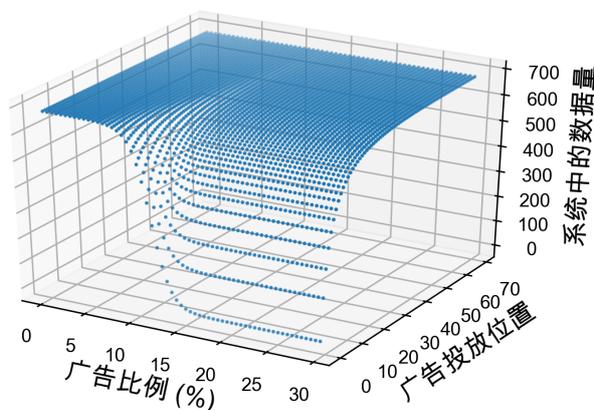


图 4-7 系统内的流量

Fig 4-7 Traffic in the system)

当在视频中不同的位置投放不同比例的广告时, 系统内平衡时的数据量如图4-7所示, 其中  $X$  轴为广告投放比例,  $Y$  轴为广告投放位置, 即视频播放的比例,  $Z$  轴为系统达到稳定是单位时间内传输的数据量。

从图中可以看出, 当固定投放位置不变时, 随着投放比例的变化, 系统内数据量的下降趋势与用户因为广告比例的离开概率相关。例如, 当广告投放位置为视频内容播放前, 即  $y = 0$  时, 系统内的数据量与4-2中所示的数据量变化曲线相同。当固定广告的投放比例, 分析不同广告播放位置对于系统稳定时数据量时, 其数据量可以分为两部分, 仅观看广告播放前视频的用户及观看含有广告的完整视频的用户。例如, 当固定广告投放量为 0, 即  $x = 0$  时, 系统内的数据量不受位置因素的影响, 值为稳定值; 而固定广告投放量为 30%, 即  $x = 0.3$  时, 用户因为广告比例的离开概率趋近于 1, 所以仅观看广告播放前的视频, 根据用户观看视频的自然离开概率, 系统内的数据量的变化曲线即为随着播放进度用户离开概率的累积分布, 最终基本趋近于用户观看完整视频的总数据量。虽然广告放置在视频播放过程中较后位置时, 系统内传输的数据量保持在较高的水平, 但是并非所有的数据量均可以为内容商带来收益。由于广告投放位置靠后, 在其播放前的内容传输用户吸引用户, 因此这部分内容并没有为内容商带来收益。只有用户观看完内容商投放的广告内容所产生的数据量才可以创造利润。

随着广告投放比例及插入位置的改变, 系统内可以产生利润的数据量如图4-8所示。从图4-8可以看出, 随着视频中广告占比的增加, 导致用户的离开概率增加, 系统内产生收益的数据量最终趋近于 0; 而随着广告播放位置的后移, 用户对广

告比例的接受度有一定程度的提升，从而系统中可行的投放广告比例也随之提升，然而，用户群体对于广告接受程度并非可以无限上升，随着广告占比的不断增加，用户离开视频观看系统的概率随之增加，从而使得系统内的数据量迅速恶化。

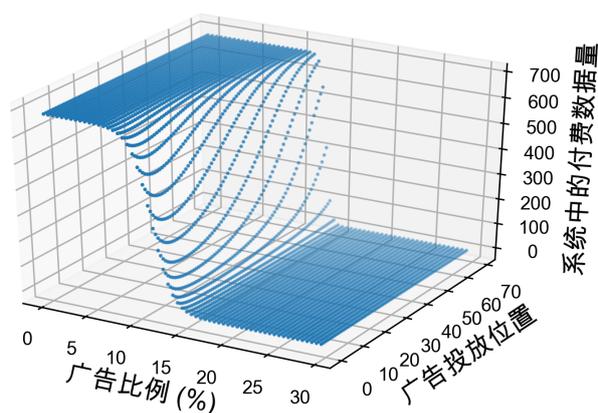


图 4-8 系统内的付费流量

Fig 4-8 Paid traffic in the system

在不同的广告占比及播放位置时对应的内容商收益如图4-9所示，其  $Z$  轴为对应的内容商收益。从图中可以看出，当固定投放比例为 0，即  $x = 0$  时，内容商的收益相同，因为此时视频播放过程中没有广告插入，用户的视频观看量完全取决于系统在购买相应网络资源下的服务质量。对于每一个投放位置，均存在一个最佳的广告投放比例，在这组最佳的（位置，比例）的组合下，系统内稳定时的数据量可能不是最大的，但是其对应的收益为相应投放位置所能实现的最大值。随着投放位置的后移，最佳的投放比例逐渐增大，由于用户对于广告比例的接受程度稍有提升，因此可以在广告投放比例和用户的离开概率之间取得折中，从而进一步提升内容商收益。根据用户的离开概率曲线，广告比例并非可以无限提升，随着其占比的不断增加，用户的离开概率会迅速增加，从而使得系统内可计费的流量迅速下降，而内容商的收益随之恶化，其收益可能差于投放在视频开始播放之前的位置。例如，当投放比例为 30% 时，用户因为广告比例离开观看系统的概率接近于 1，随着播放位置的后移，用户在广告播放位置之前观看的数据量逐渐增大，并在广告播放位置之前离开，因此系统中用户观看的免费数据量不断增大，而计费数据量非常小，因此随着投放位置的持续后移，内容商的亏损不断增大。

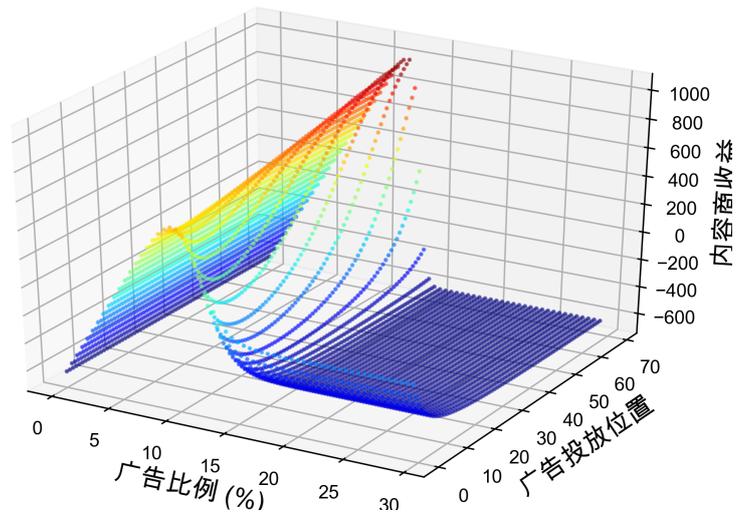


图 4-9 内容商的收益 (广告位置)

Fig 4-9 Profit of content provider (AD position)

#### 4.4 本章小结

在本章中，首先对商业广告投放比例及其投放位置对于视频播放系统达到稳定时的总数据量进行分析，推导在给定内容商的资源购买策略时系统内稳定时数据量的变化。然后将最大化内容商收益的问题建模为优化问题，通过求解 KKT 方程的分析方法求解最优解。最后通过仿真分析广告比例对于用户离开视频播放系统的概率、用户需求量对于最优广告投放比例的影响，并分析广告投放位置对于系统内容的计费流量及最佳投放比例的影响。

## 5 结论

### 5.1 本文工作总结

随着在线网络视频传输技术的不断发展，视频内容商可以通过现有的网络基础设施向用户提供视频播放服务赚取用户订阅费用或广告主的推广费用，同时内容商需要向网络服务提供商支付其使用网络资源的费用。内容商在制定网络资源购买策略及商业广告投放策略时，需要考虑两方面的内容，一方面是网络资源的购买量决定其可以为用户提供服务的质量，另一方面是广告投放会影响部分用户的观看决策，从而影响其系统内整体的用户需求。以上两方面内容均会影响内容商在视频传输系统中的效益，因此其策略制定对系统整体有重要意义。

本文根据现有关于含有缓存的视频传输系统的研究，对于内容商有预算约束情况下的网络资源部署及商业广告投放策略这两个场景进行研究，通过建立用户体验与网络资源之间的关系和广告投放策略与用户离开概率之间的关系，将系统中内容商的收益建模为优化问题，通过理论推导和仿真分析系统中不同参数对内容商收益的影响。

本文的主要工作主要包括以下三个部分：

(1) 建立有预算约束时内容商收益的优化问题，寻找在给定预算时内容商资源购买的最佳策略。根据现有研究，在视频收益与系统内各参数在一定范围内时，内容商的缓存购买量与带宽购买量呈正相关，因此可以在求解过程中缩小最优解的求解范围。仿真表明，随着预算的增加，两类资源的购买量均会逐渐上升至最大值，但是其上升的速度不同，这会受到多方面原因的影响。本文中就对视频内容的流行度及两类网络资源的价格进行仿真，当视频内容的流行度更加集中时，随着预算增加，内容商购买缓存资源的增长速度快于带宽资源；当两类网络资源的价格发生变化时，最佳策略则偏向于价格较便宜的一方，这与购买资源的利润率是一致的。

(2) 建立含有商业广告投放的系统中内容商收益的模型并分析不同广告占比对内容商收益的影响。该系统中的广告内容均投放于视频播放之前，用户离开视频播放系统的概率随广告比例的提高而上升，且不同用户群体的离开概率曲线不同。在含有广告投放的视频系统中，内容商的收益取决于广告的播放量。本文通过仿真分析，对于不同的用户群体，系统达到平衡时其传输的数据量、内容商的收益随着其中投放的广告比例的变化情况，并将其与“内容计费”模式下的数据量和收益进行比较。仿真结果表明，随着广告比例的提升，系统内平衡时流量首先缓慢下降，随后迅速下降，其下降趋势取决于用户因为广告的离开概率；而内容商的收益则随广告比例，首先上升达到一定值后开始下降，即用户的大量离开

使得系统中的收益下降。

(3) 仿真分析包含广告投放位置因素的系统达到平衡时传输的数据量和内容商收益的变化。当商业广告的广告投放位置发生变化时，用户可以在广告开始之前观看部分视频内容，因此其因为广告离开视频播放系统的概率稍有下降，但是由于用户观看在线视频的自然离开概率，使得观看到广告投放位置的用户数下降，因此将广告投放位置后移会使得部分流量为这部分用户传输视频，但是并没有相应的广告收益。该系统中的流量可以被划分为两部分：含广告的付费流量部分和不含广告的免费流量部分。随着广告投放位置的不断后移，系统达到稳定时的流量并未下降很多，但是其中可以为内容商带来收益的付费流量在不断下降。仿真结果表明，系统中的内容商收益同时随广告投放比例和投放位置两者发生变化。

## 5.2 存在问题与展望

本文存在的问题以及未来可开展的工作方向如下：

(1) 本文中所使用的内容价值、网络资源价值均根据相关文献进行数学模型假设并用于仿真分析，在实际的在线网络视频播放系统中，运营商于内容商之间的网络资源交易模型存在更加复杂的商业化信息，将这些信息应用于已有的网络资源分配模型中，将更加贴适于内容商制定资源分配策略时的应用场景。

(2) 本文在视频中商业广告投放系统所使用的用户离开概率模型根据假设的用户观看行为特性选择相应的数学模型，在真实的视频播放系统中，根据大量的用户观看行为及视频播放进度记录，可以进行更加真实的用户离开概率曲线拟合，将真实数据拟合得到函数应用于模型仿真中，将进一步提高模型的参考价值。

(3) 本文中广告的投放比例及位置分析模型中，将两个优化变量松弛为连续变量，若应用于真实系统，存在以下两种优化方式：a) 实际系统中的广告时长为某些固定时长的组合，投放位置为若干分割点的选择问题，则投放比例及位置问题进一步可以表示为整数规划问题，其求解方法更加灵活多变；b) 目前的在线网络视频播放系统可以为用户提供更加个性化、定制化的服务，不同的用户对于系统的需求、广告时长和质量的接受程度、观看习惯等存在较大的差异，得益于日渐进步的客户端计算能力，内容商的在线广告投放策略可以根据不同的用户实现个性化投放，从而进一步提高系统内的整体效益。

## 参考文献

- [1] 中国互联网网络信息中心. 《中国互联网络发展状况统计报告》, 第 42 次 [R]. 2021.
- [2] Zhang Y, Li P, Zhang Z, et al. Challenges and Chances for the Emerging Short Video Network [C]. In IEEE INFOCOM 2019 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), 2019: 1025–1026.
- [3] Goian H S, Al-Jarrah O Y, Muhaidat S, et al. Popularity-Based Video Caching Techniques for Cache-Enabled Networks: A Survey [J]. IEEE Access, 2019, 7: 27699–27719.
- [4] Hasan S, Gorinsky S, Dovrolis C, et al. Trade-offs in optimizing the cache deployments of CDNs [C]. In IEEE INFOCOM 2014 - IEEE Conference on Computer Communications, 2014: 460–468.
- [5] Zhou C, Guo Y, Chen Y, et al. Characterizing user watching behavior and video quality in mobile devices [C]. In 2014 23rd International Conference on Computer Communication and Networks (ICCCN), 2014: 1–6.
- [6] Chen Y, Liu Y, Zhang B, et al. On distribution of user movie watching time in a large-scale video streaming system [C]. In 2014 IEEE International Conference on Communications (ICC), 2014: 1825–1830.
- [7] Bulkan U, Dagiuklas T, Iqbal M. Modelling Quality of Experience for Online Video Advertisement Insertion [J]. IEEE Transactions on Broadcasting, 2020, 66 (4): 835–846.
- [8] Natarajan S, Das D. Demand Aware Edge Caching Architecture for evolved Multimedia Broadcast Multicast Service to Reduce Latency and bandwidth Savings [C]. In 2018 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), 2018: 1–6.
- [9] Han R, Zeng X, Sun P. Bandwidth allocation on video content distribution servers [C]. In 2010 IEEE Youth Conference on Information, Computing and Telecommunications, 2010: 247–250.
- [10] Zhang Y, Zhang Y, Qin S, et al. Delay-Bounded Priority-Driven Resource Allocation for Video Transmission Over Multihop Networks [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2014, 24 (7): 1184–1196.
- [11] Su G-M, Wu M. Efficient bandwidth resource allocation for low-delay multiuser video streaming [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2005, 15 (9): 1124–1137.
- [12] Li X, Kanonakis K, Cvijetic N, et al. Joint bandwidth provisioning and cache management for video distribution in software-defined Passive Optical Networks [C]. In OFC 2014, 2014: 1–3.
- [13] Zhong C, Gursoy M C, Velipasalar S. Deep Multi-Agent Reinforcement Learning Based Cooperative Edge Caching in Wireless Networks [C]. In ICC 2019 - 2019 IEEE Interna-

- tional Conference on Communications (ICC), 2019: 1–6.
- [14] Wang F, Wang F, Liu J, et al. Intelligent Video Caching at Network Edge: A Multi-Agent Deep Reinforcement Learning Approach [C]. In IEEE INFOCOM 2020 - IEEE Conference on Computer Communications, 2020: 2499–2508.
- [15] Zhang Y, Li P, Zhang Z, et al. AutoSight: Distributed Edge Caching in Short Video Network [J]. IEEE Network, 2020, 34 (3): 194–199.
- [16] Ayoub O, Musumeci F, Andreoletti D, et al. Optimal Cache Deployment for Video-on-Demand Delivery in Optical Metro-Area Networks [C]. In 2018 IEEE Global Communications Conference (GLOBECOM), 2018: 1–6.
- [17] Ahlehagh H, Dey S. Video-Aware Scheduling and Caching in the Radio Access Network [J]. IEEE/ACM Transactions on Networking, 2014, 22 (5): 1444–1462.
- [18] Wu D, Huang J, He J, et al. Toward Cost-Effective Mobile Video Streaming via Smart Cache With Adaptive Thresholding [J]. IEEE Transactions on Broadcasting, 2015, 61 (4): 639–650.
- [19] Xiao A, Huang X, Wu S, et al. User Preference Aware Resource Management for Wireless Communication Networks [J]. IEEE Network, 2020, 34 (3): 78–85.
- [20] Sun W, Li Y, Hu C, et al. Joint Optimization of Cache Placement and Bandwidth Allocation in Heterogeneous Networks [J]. IEEE Access, 2018, 6: 37250–37260.
- [21] Shi L, Wang X, Ma R T B. On Optimal Hybrid Premium Peering and Caching Purchasing Strategy of Internet Content Providers [C]. In IEEE INFOCOM 2019 - IEEE Conference on Computer Communications, 2019: 2107–2115.
- [22] Ghoreishi S E, Karamshuk D, Friderikos V, et al. A Cost-Driven Approach to Caching-as-a-Service in Cloud-Based 5G Mobile Networks [J]. IEEE Transactions on Mobile Computing, 2020, 19 (5): 997–1009.
- [23] Seufert M, Wehner N, Wamser F, et al. Unsupervised QoE field study for mobile YouTube video streaming with YoMoApp [C]. In 2017 Ninth International Conference on Quality of Multimedia Experience (QoMEX), 2017: 1–6.
- [24] Baraki H, Comes D, Geihs K. Context-Aware Prediction of QoS and QoE Properties for Web Services [C]. In 2013 Conference on Networked Systems, 2013: 102–109.
- [25] Godana B E, Kooij R E, Ahmed O K. Impact of Advertisements during Channel Zapping on Quality of Experience [C]. In 2009 Fifth International Conference on Networking and Services, 2009: 249–254.
- [26] Kodialam M, Lakshman T V, Mukherjee S, et al. Online Scheduling of Targeted Advertisements for IPTV [J]. IEEE/ACM Transactions on Networking, 2011, 19 (6): 1825–1834.
- [27] Zach O, Slanina M, Seufert M. Investigating the Impact of Advertisement Banners and Clips on Video QoE [C]. In 2018 IEEE 38th International Conference on Distributed Computing Systems (ICDCS), 2018: 1618–1623.
- [28] Mei T, Hua X-S. Contextual Internet Multimedia Advertising [J]. Proceedings of the

- IEEE, 2010, 98 (8): 1416–1433.
- [29] Song X, Xu B, Jiang Y-G. Predicting Content Similarity via Multimodal Modeling for Video-In-Video Advertising [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2021, 31 (2): 569–581.
- [30] Wang J, Duan L-y, Wang B, et al. Linking video ADS with product or service information by web search [C]. In *2009 IEEE International Conference on Multimedia and Expo, 2009*: 274–277.
- [31] Hua X-S, Mei T, Li S. When multimedia advertising meets the new Internet era [C]. In *2008 IEEE 10th Workshop on Multimedia Signal Processing, 2008*: 1–5.
- [32] Tapu R, Mocanu B, Zaharia T. DEEP-AD: A Multimodal Temporal Video Segmentation Framework for Online Video Advertising [J]. *IEEE Access*, 2020, 8: 99582–99597.
- [33] Mei T, Hua X-S, Li S. VideoSense: A Contextual In-Video Advertising System [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2009, 19 (12): 1866–1879.
- [34] Mei T, Li L, Tian X, et al. PageSense: Toward Stylewise Contextual Advertising via Visual Analysis of Web Pages [J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2018, 28 (1): 254–266.
- [35] Yim H-J, Kim S, Lim B-M, et al. Application-Based Targeted Advertisement System for ATSC 3.0 UHD Service [J]. *IEEE Transactions on Broadcasting*, 2021, 67 (1): 56–67.
- [36] Ji W, Chen Y, Chen M, et al. Profit Maximization through Online Advertising Scheduling for a Wireless Video Broadcast Network [J]. *IEEE Transactions on Mobile Computing*, 2016, 15 (8): 2064–2079.
- [37] Geyik S C, Faleev S, Shen J, et al. Joint Optimization of Multiple Performance Metrics in Online Video Advertising [C]. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, New York, NY, USA, 2016*: 471–480.
- [38] Yadati K, Katti H, Kankanhalli M. CAVVA: Computational Affective Video-in-Video Advertising [J]. *IEEE Transactions on Multimedia*, 2014, 16 (1): 15–23.
- [39] Academy Y C. Know the different ad formats. [https://creatoracademy.youtube.com/page/lesson/how-ads-work\\_ad-formats\\_image?cid=how-ads-work&hl=en](https://creatoracademy.youtube.com/page/lesson/how-ads-work_ad-formats_image?cid=how-ads-work&hl=en) Accessed April 23, 2021.
- [40] <https://www.youtube.com/> Accessed April 25, 2021.
- [41] <http://www.iqiyi.com/> Accessed April 25, 2021.
- [42] Academy Y C. Best practices for mid-roll ads. [https://creatoracademy.youtube.com/page/lesson/how-ads-work\\_mid-rolls\\_video?cid=how-ads-work&hl=en](https://creatoracademy.youtube.com/page/lesson/how-ads-work_mid-rolls_video?cid=how-ads-work&hl=en) Accessed April 25, 2021.
- [43] Academy Y C. How engaged is my audience? <https://creatoracademy.youtube.com/page/lesson/engagement-analytics?hl=en-GB> Accessed April 23, 2021.
- [44] Kang C. F.C.C. Repeals Net Neutrality Rules. <https://www.nytimes.com/2017/12/>

- [14/technology/net-neutrality-repeal-vote.html](#) Accessed April 22, 2021.
- [45] Chau C K, Wang Q, Chiu D M. Economic Viability of Paris Metro Pricing for Digital Services [J]. *Acm Transactions on Internet Technology*, 2015, 14 (2-3): 1–21.
- [46] Breslau L, Cao P, Fan L, et al. Web caching and Zipf-like distributions: evidence and implications [C]. In *IEEE INFOCOM '99. Conference on Computer Communications. Proceedings. Eighteenth Annual Joint Conference of the IEEE Computer and Communications Societies. The Future is Now (Cat. No.99CH36320)*, 1999: 126–134 vol.1.
- [47] Jedari B, Francesco M D. Auction-based Cache Trading for Scalable Videos in Multi-Provider Heterogeneous Networks [C]. In *IEEE INFOCOM 2019 - IEEE Conference on Computer Communications*, 2019: 1864–1872.
- [48] Kamiyama N, Hosokawa Y. Optimally Designing Virtualized CDN Maximizing Profit of Content Providers [C]. In *2019 16th IEEE Annual Consumer Communications Networking Conference (CCNC)*, 2019: 1–4.
- [49] Wang H, Wu K, Wang J, et al. Rldish: Edge-Assisted QoE Optimization of HTTP Live Streaming with Reinforcement Learning [C]. In *IEEE INFOCOM 2020 - IEEE Conference on Computer Communications*, 2020: 706–715.
- [50] Hao Y, Hu L, Qian Y, et al. Profit Maximization for Video Caching and Processing in Edge Cloud [J]. *IEEE Journal on Selected Areas in Communications*, 2019, 37 (7): 1632–1641.
- [51] Dong X, Zhou X, Zhao L, et al. Cloud Resource Provision of Competitive Content Providers: Models and Analysis [C]. In *2019 IEEE Intl Conf on Parallel Distributed Processing with Applications, Big Data Cloud Computing, Sustainable Computing Communications, Social Computing Networking (ISPA/BDCloud/SocialCom/SustainCom)*, 2019: 53–60.
- [52] Kaliski R, Wei H-Y. Dynamic Resource Allocation and Advertisement Revenue Optimization for TV Over eMBMS [J]. *IEEE Transactions on Broadcasting*, 2016, 62 (3): 579–597.
- [53] Seeliger R, Bassbouss L, Arbanowski S, et al. Towards Personalized Content Replacements in Hybrid Broadcast Broadband Environments [C]. In *2019 23rd International Computer Science and Engineering Conference (ICSEC)*, 2019: 385–389.
- [54] Khalek A A, Caramanis C, Heath R W. Delay-Constrained Video Transmission: Quality-Driven Resource Allocation and Scheduling [J]. *IEEE Journal of Selected Topics in Signal Processing*, 2015, 9 (1): 60–75.
- [55] Ghoreishi S E, Friderikos V, Karamshuk D, et al. Provisioning cost-effective mobile video caching [C]. In *2016 IEEE International Conference on Communications (ICC)*, 2016: 1–7.
- [56] Huang A, Li Y, Xiao Y, et al. Distributed Resource Allocation for Network Slicing of Bandwidth and Computational Resource [C]. In *ICC 2020 - 2020 IEEE International Conference on Communications (ICC)*, 2020: 1–6.

- [57] 黄楠栖. 在线广告的跨平台测量方法研究 [D]. [S. 1.]: 北京交通大学, 2018.
- [58] 刘鹏, 王超. 计算广告 [M]. 人民邮电出版社, 2015.
- [59] 陈宝林. 最优化理论与算法 [M]. 清华大学出版社, 2005.

## 附录 A

## 符号及函数的含义

表 0-1 符号和函数及其相应含义

Table 0-1 Symbols, functions and their definition

参数	含义
$b$	内容商购买的带宽资源
$c$	内容商购买的缓存资源
$b_0, \bar{b}$	基础服务带宽资源, 总带宽资源
$(b, c)$	内容商购买策略
$p_b(b), p_1, \gamma$	带宽的单位价格及参数
$p_c(c), p_2$	缓存的单位价格及参数
$\bar{d}$	平均传输时延
$n$	系统内用户总数
$\eta(\bar{d})$	根据时延, 用户的订阅概率
$r(\bar{d})$	根据时延, 订阅用户的平均需求数据量
$f(\bar{d})$	系统内所有用户的平均需求数据量
$T(b, c)$	系统内传输的总数据量
$\tilde{T}(b, c)$	通过带宽传输的数据量
$p_v$	平均单位视频数据量的价值
$d_0$	缓存服务器至用户的传输时延
$d$	源服务器至网络边缘的传输时延
$H(c)$	缓存命中概率
$M(c)$	缓存未命中概率
$E(b, c)$	内容商的总支出
$U(b, c)$	内容商的总收益
$B$	内容商的预算
$N$	内容商的文件总数
$\kappa$	文件流行度分布参数
$\tau$	视频中的广告占比
$l(\tau)$	用户因为广告比例离开视频观看的概率
$\alpha, \delta$	用户离开概率的参数
$t$	视频的播放比例
$v(t)$	用户随视频播放进度的自然离开概率

## 作者简介及攻读硕士学位期间取得的研究成果

### 一、作者简介

2014年9月至2018年6月，就读于北京交通大学，电信学院，通信工程专业，获学士学位；  
2019年9月至2020年9月，就读于英国伯明翰大学，计算机学院，计算机专业，获硕士学位；

2018年9月至2021年6月，就读于北京交通大学，电信学院，通信与信息系统专业，导师李纯喜副教授。

### 二、发表论文

[1] **Chenyue Zhang**, Chunxi Li, Yongxiang Zhao, Nanxi Huang and Baoxian Zhang, "Cross-Platform Measurement on Ad Exchanges", 2019 IEEE International Conference on Consumer Electronics - Taiwan (ICCE-TW), Yilan, Taiwan, 2019, pp. 1-2.

[2] Huxin Zhang, Chunxi Li, Yongxiang Zhao and **Chenyue Zhang**, "Caching Videos with Category Diversity based Recommendation: Model and Algorithm", 2020 IEEE 6th International Conference on Computer and Communications (ICCC), Chengdu, China, 2020, pp. 1840-1846.

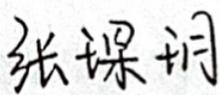
### 三、参与科研项目

[1] 国家自然科学基金面上项目，61872031，基于熵理论的信息匹配网络测量与建模，2019/01-2022/12，参加。

## 独创性声明

本人声明所提交的学位论文是本人在导师指导下进行的研究工作和取得的研究成果，除了文中特别加以标注和致谢之处外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得北京交通大学或其他教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

学位论文作者签名：



签字日期：2021年6月1日

## 学位论文数据集

表 1.1: 数据集页

关键词*	密级*	中图分类号	UDC	论文资助
缓存; 带宽; 在线视频; 资源配置; 广告投放	公开			国家自然科学基金 No. 61872031
学位授予单位名称*		学位授予单位代码*	学位类别*	学位级别*
北京交通大学		10004	工学	硕士
论文题名*		并列题名		论文语种*
关于在线视频系统中网络资源部署及广告投放的研究				中文
作者姓名*	张琛玥	学号*	18120172	
培养单位名称*	培养单位代码*	培养单位地址	邮编	
北京交通大学	10004	北京市海淀区西直门外上园村 3 号	100044	
学科专业*	研究方向*	学制*	学位授予年*	
通信与信息系统	信息网络	3	2021	
论文提交日期*	2021 年 6 月			
导师姓名*	李纯喜	职称*	副教授	
评阅人	答辩委员会主席*	答辩委员会成员		
	郭宇春	赵永祥、郑宏云、张立军、陈一帅		
电子版论文提交格式 文本 ( ) 图像 ( ) 视频 ( ) 音频 ( ) 多媒体 ( ) 其他 ( ) 推荐格式: application/msword; application/pdf				
电子版论文出版 (发布) 者		电子版论文出版 (发布) 地		权限声明
论文总页数*	51			
共 33 项, 其中带*为必填数据, 为 21 项。				