



闽江学院

本科毕业论文（设计）

题目	基于大数据技术的旅游攻略平台“旅行家” 的设计与实现
学生姓名	陈弘毅
学号	3197103142
学院	数学与数据科学学院（软件学院）
年级	2019
专业	软件工程（闽台合作）
指导教师	曹永忠
职称	教授
完成日期	2023年4月

闽江学院毕业论文（设计）诚信声明书

本人郑重声明：

兹提交的毕业论文（设计）《基于大数据技术的旅游攻略平台“旅行家”的设计与实现》，是本人在指导老师 曹永忠 的指导下独立研究、撰写的成果；论文（设计）未剽窃、抄袭他人的学术观点、思想和成果，未篡改研究数据，论文（设计）中所引用的文字、研究成果均已在论文（设计）中以明确的方式标明；在毕业论文（设计）工作过程中，本人恪守学术规范，遵守学校有关规定，依法享有和承担由此论文（设计）产生的权利和责任。

声明人（签名）：

年 月 日

摘要

随着旅游行业的大力发展，自由行越来越受欢迎，人们在网络平台上探索旅游信息，制定旅行线路并查看当地美食、景点。但一个非专业的平台会让用户感兴趣的信息与冗余信息杂乱无章的混合在一起，这给用户带来诸多不便。为了向用户提供高质量信息，推荐系统走进人们的视野，让用户更了解目的地。

本文使用 Python 语言，设计并实现了一款旅游攻略系统。该系统提供用户历史行为兴趣分析挖掘的个性化旅游攻略平台，结合用户行为兴趣特征，为用户推荐感兴趣的旅游景点。利用 Python 的 Flask 框架、MySQL 数据库以及 JavaScript 等前端技术，实现用户登陆、景点推荐、预订景点等功能。实验结果表明，该系统具有较高的推荐准确率和用户满意度，具有一定的实用价值和应用前景。

关键词：旅游攻略平台；协同过滤推荐；Python 爬虫

Abstract

With the vigorous development of the tourism industry, free travel is becoming more and more popular. People explore tourism information, make travel routes and check local food and scenic spots on online platforms. However, a non-professional platform will leave users with a messy mix of information of interest and redundant information, which will bring users a lot of inconvenience. In order to provide users with high quality information, the recommendation system comes into people's view, so that users know more about the destination.

This paper designs and implements a travel recommendation system based on Python language. The system adopts collaborative filtering algorithm and combines user behavior and interest characteristics to provide personalized travel recommendation service for users. The collaborative filtering algorithm based on user behavior is adopted to analyze and mine users' historical behaviors and interests, so as to recommend tourist attractions that users are interested in. The Flask framework of Python and front-end technologies such as HTML, CSS, JavaScript are used to achieve user login, scenic spot recommendation, scenic spot reservation and other functions. The experimental results show that the system has high recommendation accuracy and user satisfaction, and has certain practical value and application prospect.

.

Key words: tourism recommendation; Collaborative filtering recommendation; Python crawler

目 录

1 引言	1
2 研究目的	1
2.1 研究背景	1
2.2 研究意义	2
3 研究设计	2
3.1 可行性分析	2
3.1.1 技术可行性	2
3.1.2 操作可行性	2
3.2 功能需求分析	3
3.3 技术选型	4
3.3.1 Python 语言	4
3.3.2 Flask 框架	4
3.3.3 My SQL 简介	4
3.4 基于协同过滤的推荐模型创建	5
3.4.1 算法介绍	5
3.4.2 评价指标的建立	5
3.4.3 构建评分矩阵	6
3.4.4 计算相似度	6
3.4.5 推荐算法的相似度融合及预测用户评分	6
3.5 功能模块设计	7
3.5.1 系统架构设计	7
图 3.1 系统结构图	7
3.5.2 功能结构设计	7
3.6 系统流程设计	8
3.6.1 用户注册流程	8
3.6.2 景点购票流程	9
3.6.3 管理员流程	9
3.6.4 协同过滤算法流程	9
3.7 数据库设计	10
3.7.1 数据库实体设计	10
3.7.2 数据表设计	12
4 系统实现	17
4.1 个性化推荐算法模块的实现	17

4.2 前台功能模块	18
4.2.1 网站首页	18
4.1.2 会员注册登录界面	19
4.2.3 景点推荐页面	20
4.2.4 景点查看页面	20
4.2.5 门票购买页面	21
4.2.6 在线留言页面	22
4.2 后台管理模块	22
4.2.1 管理员登录页面	22
4.2.2 购票信息管理页面	23
4.2.3 景点管理页面	23
4.2.3 景点门票管理页面	24
5 系统测试	25
5.1 系统测试的定义	25
5.2 系统测试的目的	25
5.3 测试用例方法	25
5.4 测试验收结果	26
6 结论与分析	29
致谢	30
参考文献	32

1 引言

随着旅游行业的发展，越来越多的人开始热爱旅游。然而，旅游攻略的质量参差不齐，有的攻略过于简单，有的攻略过于繁琐，让人不知道从何入手。基于这种情况，设计并实现了一款基于大数据的旅游攻略平台——“旅行家”。

“旅行家”旨在为旅游爱好者提供一站式的旅游攻略服务，帮助用户规划旅游路线、寻找景点、选择住宿、了解特产等。平台采用大数据技术，通过分析用户的历史行为、偏好、评价等信息，为用户推荐个性化的旅游攻略，提高用户的旅游体验。

本文将详细介绍“旅行家”系统的设计。首先，将阐述“旅行家”系统的背景和意义。接着，分析了“旅行家”系统的需求，包括用户需求和系统需求。然后，将介绍“旅行家”系统的架构和实现。最后，将对系统进行用户体验测试，以验证系统的可行性和有效性。

旅游推荐的设计是一个复杂的过程，需要综合考虑用户需求、数据处理和推荐算法等多方面因素。通过本文的介绍，读者可以更深入地了解旅游推荐系统的设计原理和实现方法，为旅游推荐系统的开发提供参考。

2 研究目的

2.1 研究背景

旅游业是一项十分依赖于信息的产业。旅游景区能否及时获取旅游相关信息，并将其整合成具有吸引力的旅游产品或服务，决定着该景区旅游业能否在当前信息快速更新的时代顺利发展。随着旅游业的不断发展，游客的心态日趋成熟和理性^[1]。因此，为了更好地满足游客的需求，提供准确、全面、个性化的旅游信息尤为关键，是保持可持续发展的必要条件之一。

随着互联网技术的不断发展，旅游行业也在不断地发展壮大^[2]。越来越多的人选择利用互联网寻找旅游攻略和旅游信息，然而，旅游攻略的质量和数量却存在着较大的差异性，旅游者面临着信息不对称、行程安排不合理、景点选择不当等问题。

为了解决这些问题，基于协同过滤的旅游攻略推荐平台应运而生^[3]。该平台利用协同过滤算法将用户分成不同的小组，通过收集用户的收集记录等信息，判断用户可能感兴趣的旅行攻略，并根据该小组的历史行为进行推荐。

2.2 研究意义

基于协同过滤算法的旅游推荐系统的设计是为给旅游爱好者推荐合适的旅游景点和旅游线路，而且给旅游者提供推荐信息以帮助他们决定到哪里旅游是目前最合适的，或者是最迎合个性化的旅游选择^[4]。许多大型网站都采用个性化推荐技术。例如，淘宝的收藏夹功能允许用户收藏商品，同时显示其他用户收藏了哪些相似商品^[5]。数据量的增多导致了个性化推荐质量的提高变得尤为重要。然而，要保证推荐结果的准确和满意，现有的个性化推荐技术和实现方法还有待于深入研究和改进。为了实现这一目标，开发本系统是必要的，同时也可以研究协同过滤算法在旅游行业中的应用。相较于其他算法，协同过滤算法在旅游领域中具有广泛的应用前景，因此值得深入探索。

3 研究设计

3.1 可行性分析

为了保证新系统研制成功，进行技术、经济和社会方面的分析，是很有必要的。以较小的成本，来确定解决问题所耗费的精力及资源。而可行性研究主要集中在以下两个方面：

3.1.1 技术可行性

Python 有着丰富的第三方库和框架支持，可以方便地实现数据获取、处理和分析，同时也具备 Web 开发的能力，能够实现用户界面和交互等功能。因此对于此次开发的技术，并没有太大困难。

3.1.2 操作可行性

操作可行性是系统可用性的一个重要方面。在开发系统时，必须关注界面设计、操作流程和功能模块的放置。系统应当拥有简洁美观的界面，将功能模块有序的排列

在导航栏中，操作流程不能过于复杂，要求使用起来易于操作，让人容易理解。开发好的系统通过研究运行的可行性，能够使用户使用起来顺畅、简便。这有助于提高系统的可用性和用户体验，从而提高系统的销售量和用户满意度。因此，在开发系统时，必须注重操作可行性，以确保系统具备良好的可用性。

3.2 功能需求分析

该系统主要实现用户管理功能、系统权限管理功能、景点推荐、门票预定功能，不过只有具有某种功能权限的人才有执行相应的操作。详细描述如下：

前台用户模块

1. 登录模块：该模块设计用于用户登录器功能的实现。登入模块要求提供作为资料库查询和匹配依据的用户名称和密码。
2. 旅游景点推荐：该模块用于实现通过算法分析，向用户推荐旅游景点。
3. 行程攻略查看：该模块用于行程攻略查看及攻略点评的实现。
4. 旅游攻略查看：该模块用于实现旅游攻略的查看，并对攻略进行评论。
5. 景区门票查看：该模块用于实现景点门票信息的查看，并对门票信息进行点评。
6. 在线留言：用户发布留言的功能，利用该模块实现
7. 旅游资讯：该模块用于实现旅游资讯信息的查看。
8. 个人中心：该模块用于实现查看本人购票信息，支付票款，修改本人信息等功能。

后台管理员模块

1. 用户管理：该模块用于实现添加、查询、修改、删除等用户信息的操作。
2. 攻略类目管理：该模块用于实现添加、查询、修改、删除等攻略类目的操作。
3. 旅游攻略管理：该模块用于实现添加、查询、修改、删除、查看用户点评等旅游攻略操作。
4. 景点管理：该模块用于实现添加、查询、修改、删除、查看景点点评等景点管理操作。
5. 票务信息管理：查询票务信息和对票务信息的审核都是通过这个模块来实现的。
6. 景区门票管理：可查询景区门票信息，并利用该模块实现爬取操作。

7. 留言管理：留言查询操作通过模块实现，对留言内容进行回复。

8. 数据可视化：该模块用于实现对景点的个性化分析和统计，主要从景点门票、价格、点评、位置等方式进行统计，通过图表方式进行展示。

3.3 技术选型

3.3.1 Python 语言

Python 和 Java 都是面向对象的语言，但是它们有一些不同之处。Python 是一种直译式计算机程序设计语言，同时也是一种功能强大的多平台语言^[6]。

Python 支持采用命令的编程方式、面向对象编程等多种程序设计方式^[7]。Python 虚拟机基本上可以在所有的操作系统中运行^[8]。

Python 也有多种版本的实现，通常所说的 Python 实际上就是 C 语言实现 Python 的 CPython。而随着 Python 越来越受到开发者的欢迎，现在 Python 已经拥有了 Jython, IronPython 等其他实现^[9]。Python 已经深入到了传统的企业级应用开发领域，这些实现已经为传统应用使用 Python 进行改进铺平了道路。

3.3.2 Flask 框架

MVC 分层结构有助于视图设计能够从业务逻辑中独立出来，从而使开发者能够更加集中精力进行优化改进。其中 Flask 是一种非常强大的框架，它同时具有了其他框架的一些特性^[10]。Flask 以“轻”来概括它的特性毫无问题。它微小的特性更适用于我们开发一些小项目，因此被称之为“微框架”^[10]。Flask 框架让开发者可以自由地选择最适合自己的项目的数据库插件。

3.3.3 MySQL 数据库

作为一种中小型的关系数据库系统，MySQL 具有高性能、可靠性和可扩展性，因此被广泛应用于各种领域。MySQL 建立在客户端/服务器架构之上，相对于其他数据库管理系统，MySQL 在一些中小型系统上的使用优势很多，比如更轻便，功能齐全，使用起来也很顺手。结合本文的系统，考虑到系统功能和成本建设等多方面的语言，因此选择了 MySQL 数据库^[11]。

3.4 基于协同过滤的推荐模型创建

3.4.1 算法介绍

旅行推荐系统的主要功能是在计算方面主要涉及用户和景点的信息处理与分析，为用户提供个性化的旅行推荐。协同过滤算法的原理和思想是基于用户历史信息来进行相似度计算，找到与目标用户兴趣相投的历史用户，最终为目标用户推荐相似的物品或服务。^[12]。为此，建立了旅游景区评价指标体系，并引入景点属性。具体地说，我们考虑把景点属性作为计算景点间相似度的特征，在协同过滤算法中加以考虑。这样就可以更好的满足用户的需求，由景点属性得出推荐列表。

本系统使用协同过滤推荐算法，基本公式为：

$$P_{ij} = \sum_{j \in N_u \cap S(i,K)} W_{ji} r_{uj} \quad (\text{公式 3-1})$$

该方法的基本原理是为了找到用户偏好的景点集合，根据用户的收藏等行为进行分析。然后，这些历史记录再通过计算相似度来确定。最后，针对相似度高的新景点，将其推荐给用户。

3.4.2 评价指标的建立

评价指标是分析旅游竞争力不可或缺的要素，对于景点相似度计算结果的提升具有重要作用^[13]。在对文献资料的考核整理基础上，我们将这些指标体系加以完善。根据下表 3.1 所示，本评价指标体系分为三档。

表 3-1 评价指标体系

一级指标	二级指标	三级指标	备注
旅游 景点 评价 指	市场竞 争力 B ₁	可进入性 C ₁	景区距离机场、火车站及高速路口的距离
		景区面积 C ₂	景区的占地总面积
		景区质量等级 C ₃	依据国家标准评定的景区质量等级
		收费标准 C ₄	景区的价格
	核心竞 争力 B ₂	餐饮住宿服务 C ₅	景区周围两公里范围内的餐馆、酒店数量
		景区热度 C ₆	各大旅游网站上景区热度的平均值
	城市支持 竞争力 B ₃	接待游客量 C ₇	景区所在城市每年所接待的游客总量
		旅游收入 C ₈	景区所在城市每年旅游总收入
		航空客运吞吐量 C ₉	景区所在城市机场每年的吞吐量

标		旅行社数量 C_{10}	景区所在城市的旅行社数量
	其他竞争力 B_4	用户评价 C_{11}	其他各旅游网站上的用户对景区的评价

3.4.3 构建评分矩阵

用户的偏好是算法的数据基础，了解用户的景点偏好可以使用用户对景点进行收藏等方式。假设系统中有 m 个用户和 n 个景点，用 $u = \{u_1, u_2, \dots, u_m\}$ 表示用户集合，用 $I = \{I_1, I_2, \dots, I_n\}$ 表示景点集合。则用户对景点的偏好矩阵为：

$$R = \begin{Bmatrix} C_{1,1} \dots C_{1,n} \\ \vdots \\ C_{m,1} \dots C_{m,n} \end{Bmatrix} \quad (\text{公式 3-6})$$

3.4.4 计算相似度

不同的偏好表示了用户对不同景点的不同喜好程度，因此本文使用修正余弦相似度公式来计算景点之间的相似度，以便更好地反映用户的喜好和行为^[14]：

$$sim_c(I_i, I_j) = \frac{\sum_{u,i,j} (C_{u,i} - \bar{C}_u) \cdot (C_{u,j} - \bar{C}_u)}{\sqrt{\sum_{U_i} (C_{u,i} - \bar{C}_u)^2} \cdot \sqrt{\sum_{U_j} (C_{u,j} - \bar{C}_u)^2}} \quad (\text{公式 3-7})$$

3.4.5 推荐算法的相似度融合及预测用户评分

记 sim_{kg} 为景点相似度， sim_c 为基于用户行为的景点相似度。本文通过加权融合的方式对二者进行融合，融合公式如下：

$$sim(I_i, I_j) = \alpha \cdot sim_{kg}(I_i, I_j) + (1 - \alpha) sim_c(I_i, I_j)$$

$$= \alpha \frac{1}{1 + \sqrt{\sum_{k=1}^d (E_{ki} - E_{kj})^2}} + (1 - \alpha) \frac{\sum_{U_{i,j}} (C_{u,j} - \bar{C}_u) \cdot (C_{u,i} - \bar{C}_u)}{\sqrt{\sum_{U_i} (C_{u,i} - \bar{C}_u)^2} \cdot \sqrt{\sum_{U_j} (C_{u,j} - \bar{C}_u)^2}} \quad (\text{公式 3-8})$$

其中， α 为融合比例，取值范围为 $[0, 1]$ 。

在该系统中，如果用户对某个景点的评价越高，这个景点与另一景点的相似度相对较高，那么就会有更高的预测分数。为了实现这一目标，提出了以下的计算公式：

$$P_{ui} = \frac{\sum_{j \in N(u), S(i,k)} sim(I_i, I_j) \times C_{u,j}}{\sum_{j \in N(u), S(i,k)} sim(I_i, I_j)} \quad (\text{公式 3-9})$$

其中， $sim(I_i, I_j)$ 是景点 I_i 和景点 I_j 的相似度， $C_{u,j}$ 表示用户对景点 I_j 的评分， $N(u)$

表示用户评分的集合。在推荐系统中，系统为每个用户生成由 N 个景点组成的推荐列表^[15]。据预测评分给出推荐，系统会将用户偏好的景点由高到低进行推荐^[16]。

3.5 功能模块设计

3.5.1 系统架构设计

这套系统需要对用户和景点的属性信息分别建立两个记录数据库表。通过计算这些属性信息，确定用户与景点的相似度，并生成推荐列表。如图 3-1 所示的系统架构图。

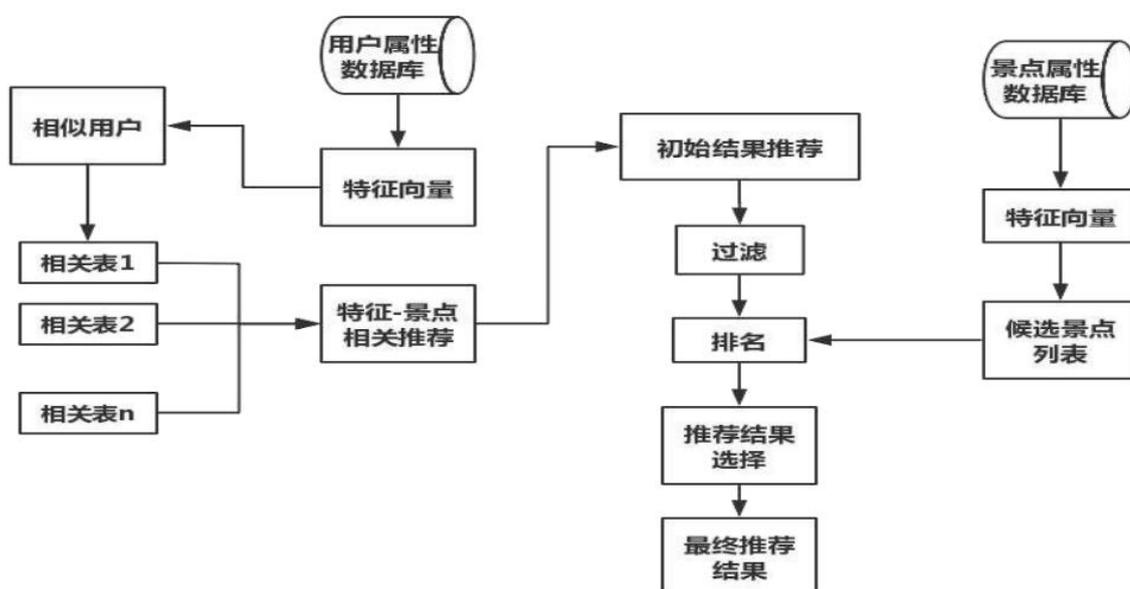


图 3-1 系统结构图

3.5.2 功能结构设计

本旅游攻略推荐系统的主要功能模块分为管理员模块和用户模块，用户功能模块有注册登录、旅游资讯查看、景点推荐查看、旅游攻略推荐、景点购票、景点门票查看、在线留言、个人中心等。管理员功能模块登录需要管理员权限的账号，有景点管理，攻略类目管理，旅游攻略管理，景点门票管理，购票信息管理，系统管理等功能模块。系统的主要功能模块如图 3-2 所示。

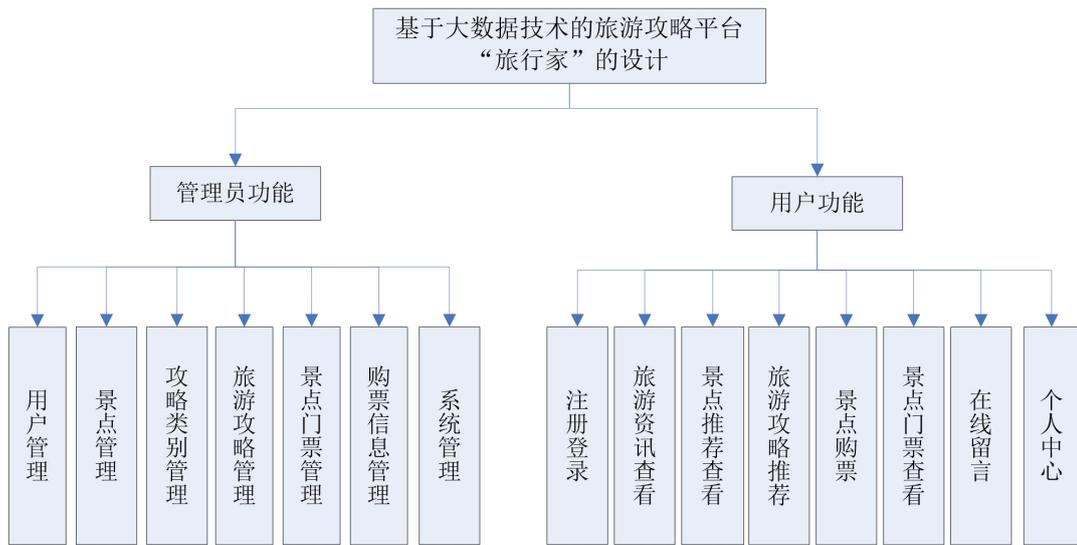


图 3-2 功能模块图

3.6 系统流程设计

3.6.1 用户注册流程

在注册时，需要注意新用户姓名此前不能存在，必须是唯一的注册用户名；要求填写的必填内容，不可以为空。用户注册的具体流程如图3-3所示：

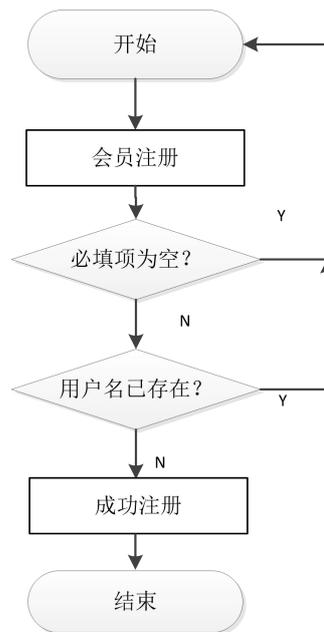


图 3-3 会员注册流程

3.6.2 景点购票流程

在进行景点购票时，需要在登录状态下操作，可以先对旅游景点的信息进行浏览，再完成景点购票的操作。景点购票的具体流程如图3-4所示：

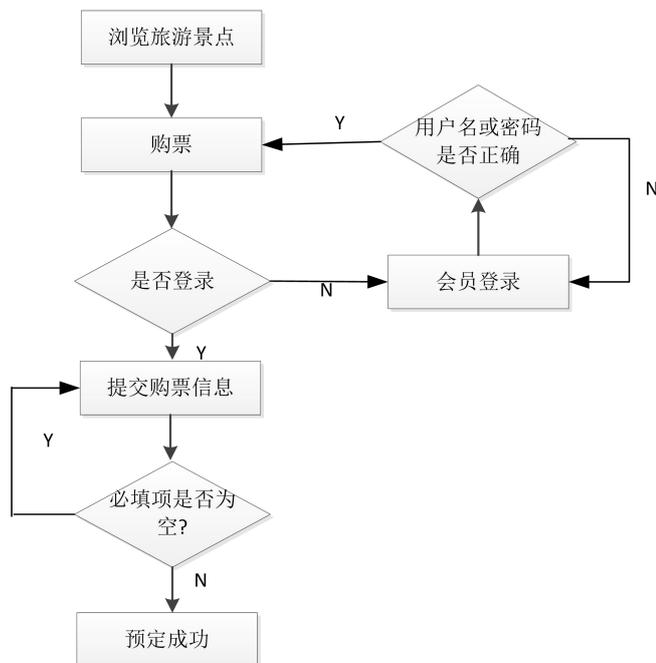


图 3-4 景点购票流程

3.6.3 管理员流程

用户端的添加、修改、查询或删除等功能模块，只有管理员成功登录后，才能进行管理。图3-5是管理人员具体的操作流程：

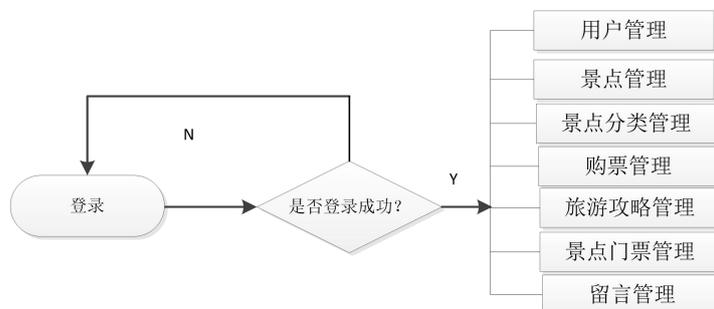


图 3-5 管理员后台管理流程

3.6.4 协同过滤算法流程

基于用户的协同过滤推荐算法，把是用户进行分类，找到相近类型的用户，然后将其浏览过的景点推荐给这个用户。这种算法主要的好处是可以个性化地推荐给每一位用户。相比之下，基于内容则是通过分析所有用户对景点的收藏等操作，找到相似的景点，然后根据用户的偏好信息，将类似景点推荐给用户^[17]。本系统将结合使用协同过滤推荐技术。

用户推荐流程如图 3-6 所示。

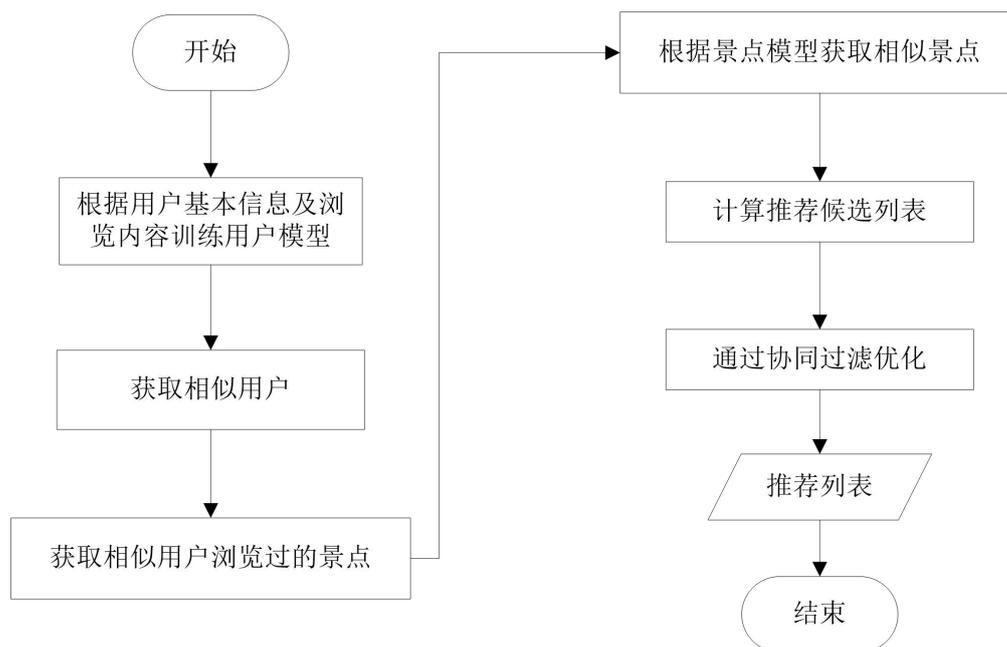


图 3-6 用户推荐流程图

3.7 数据库设计

3.7.1 数据库实体设计

数据库实体设计是一个非常重要的阶段。这个阶段需要通过对业务领域的深入了解，实体和关系都会为后面的设计埋下伏笔。主要实体的实体属性图及其相互关系，如管理员实体、使用者实体、景点实体、旅行攻略实体、景点实体等，说明如下：

1. 管理员实体：

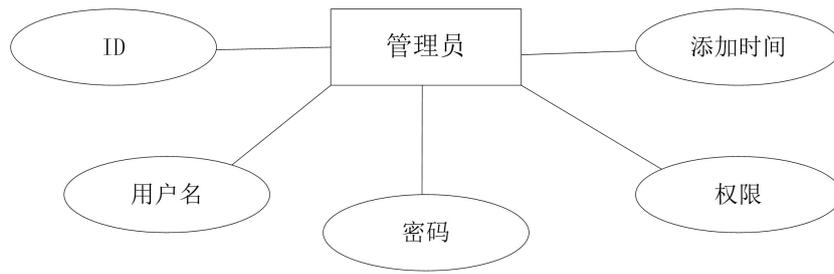


图 3-7 管理员 E-R 图

2. 用户实体:

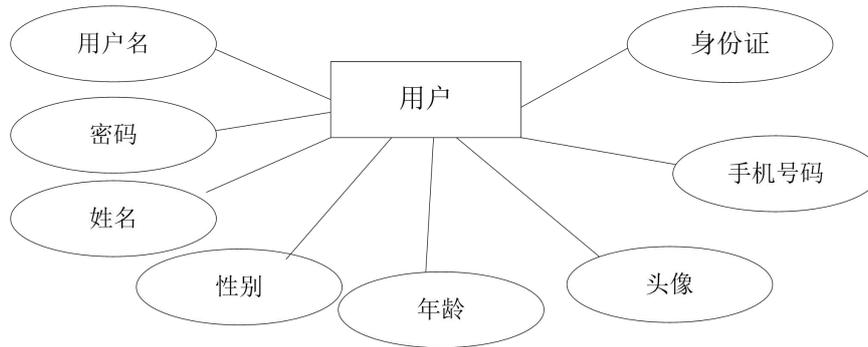


图 3-8 用户 E-R 图

3. 景点实体:

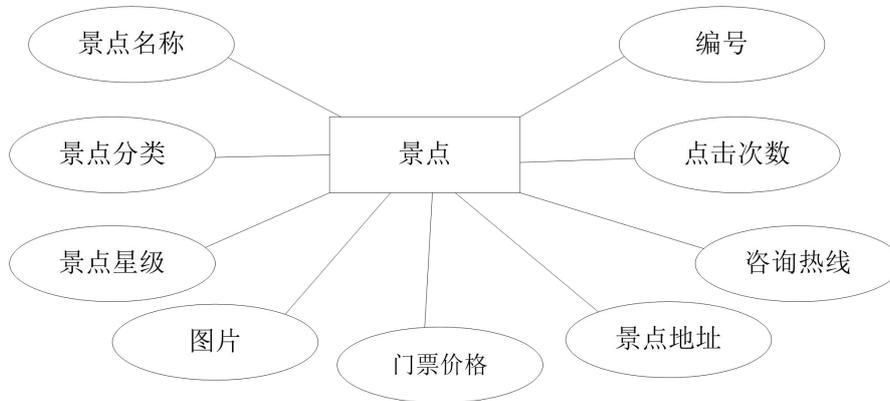


图 3-9 景点 E-R 图

4. 旅游攻略实体:

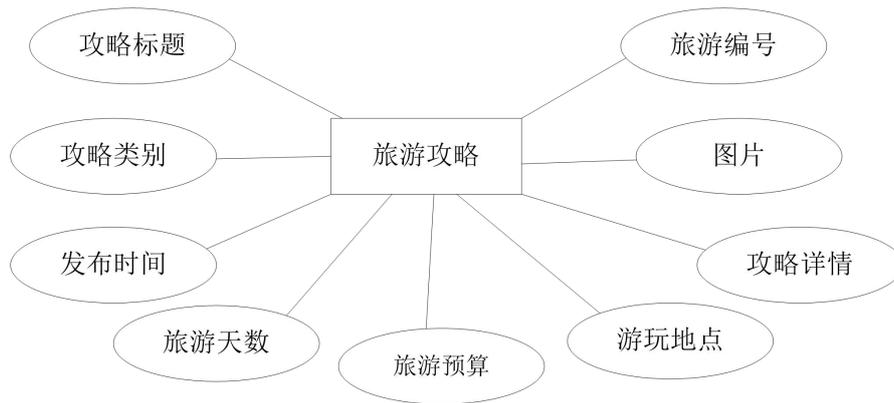


图 3-10 旅游攻略 E-R 图

5. 预定实体:

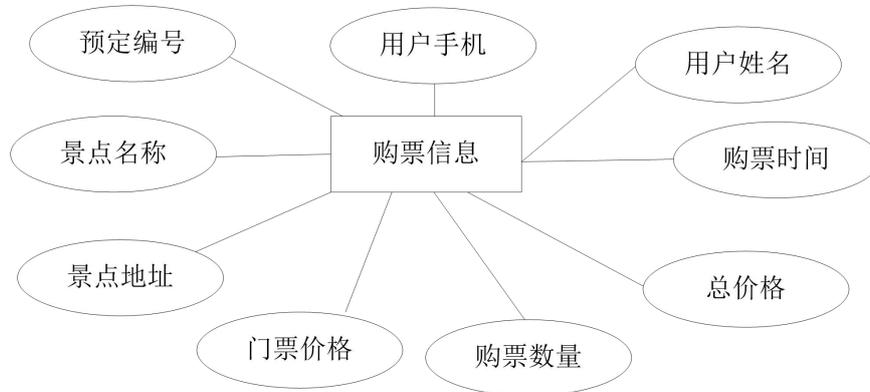


图 3-11 预定 E-R 图

3.7.2 数据表设计

本旅游推荐系统的数据库表可以根据上文的分析建立起来，以下是各表格的详细信息说明：

表 3-2 用户表

字段名称	类型	长度	字段说明	主键	默认值
id	bigint		主键	主键	
addtime	timestamp		创建时间		CURRENT_TIMESTAMP
yonghuzhanghao	varchar	200	用户账号		
mima	varchar	200	密码		
yonghuxingming	varchar	200	用户姓名		
xingbie	varchar	200	性别		
yonghudianhua	varchar	200	用户电话		
touxiang	longtext	4294967295	头像		

表 3-3 收藏表

字段名称	类型	长度	字段说明	主键	默认值
id	bigint		主键	主键	
addtime	timestamp		创建时间		CURRENT_TIMESTAMP
userid	bigint		用户 id		
refid	bigint		商品 id		
tablename	varchar	200	表名		
name	varchar	200	名称		
picture	longtext	4294967295	图片		
type	varchar	200	类型(1:收藏,21:赞,22:踩,31:竞拍参与,41:关注)		1
inteltype	varchar	200	推荐类型		
remark	varchar	200	备注		

表 3-4 用户表

字段名称	类型	长度	字段说明	主键	默认值
id	bigint		主键	主键	
username	varchar	100	用户名		
password	varchar	100	密码		
role	varchar	100	角色		管理员
addtime	timestamp		新增时间		CURRENT_TIMESTAMP

表 3-5 在线留言表

字段名称	类型	长度	字段说明	主键	默认值
id	bigint		主键	主键	
addtime	timestamp		创建时间		CURRENT_TIMESTAMP
userid	bigint		用户 id		
adminid	bigint		管理员 id		
ask	longtext	4294967295	提问		
reply	longtext	4294967295	回复		
isreply	int		是否回复		

表 3-6 景点资讯表

字段名称	类型	长度	字段说明	主键	默认值
id	bigint		主键	主键	
addtime	timestamp		创建时间		CURRENT_TIMESTAMP
title	varchar	200	标题		
introduction	longtext	4294967295	简介		
picture	longtext	4294967295	图片		
content	longtext	4294967295	内容		

表 3-7 旅游攻略表

字段名称	类型	长度	字段说明	主键	默认值
id	bigint		主键	主键	
addtime	timestamp		创建时间		CURRENT_TIMESTAMP
gonglvbianhao	varchar	200	攻略编号		
gonglvbiaoti	varchar	200	攻略标题		
gonglvleibie	varchar	200	攻略类别		
fengmian	longtext	4294967295	封面		
fabushijian	date		发布时间		
lvyoutianshu	varchar	200	旅游天数		
lvyoyusuan	varchar	200	旅游预算		
youwandidian	varchar	200	游玩地点		
gonglvxiangqing	longtext	4294967295	攻略详情		
clicktime	datetime		最近点击时间		
clicknum	int		点击次数		0

表 3-8 景点信息表

字段名称	类型	长度	字段说明	主键	默认值
id	bigint		主键	主键	
addtime	timestamp		创建时间		CURRENT_TIMESTAMP
jingdianbianhao	varchar	200	景点编号		
jingdianmingcheng	varchar	200	景点名称		
jingdianzhaopian	longtext	4294967295	景点照片		
jingdiandengji	varchar	200	景点等级		
menpiaojiage	float		门票价格		
jingdiandizhi	varchar	200	景点地址		
kaifangshijian	varchar	200	开放时间		
goupiaoxuzhi	longtext	4294967295	购票须知		
jingdianxiangqing	longtext	4294967295	景点详情		

clicktime	datetime	最近点击时间	
clicknum	int	点击次数	0

表 3-9 景点门票表

字段名称	类型	长度	字段说明	主键	默认值
id	bigint		主键	主键	
addtime	timestamp		创建时间		CURRENT_TIMESTAMP
biaoti	varchar	200	标题		
fengmian	longtext	4294967295	封面		
weizhi	varchar	200	位置		
dianping	varchar	200	点评		
pinglun	varchar	200	评论		
jiage	float		价格		
laiyuan	longtext	4294967295	来源		
tese	longtext	4294967295	特色		
kaifangshijian	longtext	4294967295	开放时间		
miaoshu	longtext	4294967295	描述		
clicktime	datetime		最近点击时间		
clicknum	int		点击次数		0

表 3-10 购票信息表

字段名称	类型	长度	字段说明	主键	默认值
id	bigint		主键	主键	
addtime	timestamp		创建时间		CURRENT_TIMESTAMP
goupiaobianhao	varchar	200	购票编号		
jingdianmingcheng	varchar	200	景点名称		
goupiaoshijian	datetime		购票时间		
menpiaojiage	float		门票价格		
shuliang	float		数量		
zongjia	float		总价		
goupiaobeizhu	longtext	4294967295	购票备注		
yonghuzhanghao	varchar	200	用户账号		
yonghuxingming	varchar	200	用户姓名		
sfsh	varchar	200	是否审核		待审核
shhf	longtext	4294967295	审核回复		
ispay	varchar	200	是否支付		未支付

表 3-11 攻略类别表

字段名称	类型	长度	字段说明	主键	默认值
id	bigint		主键	主键	
addtime	timestamp		创建时间		CURRENT_TIMESTAMP
gonglveleibie	varchar	200	攻略类别		

表 3-12 旅游攻略评论表

字段名称	类型	长度	字段说明	主键	默认值
id	bigint		主键	主键	
addtime	timestamp		创建时间		CURRENT_TIMESTAMP
refid	bigint		关联表 id		
userid	bigint		用户 id		
avatarurl	longtext	4294967295	头像		
nickname	varchar	200	用户名		
content	longtext	4294967295	评论内容		
reply	longtext	4294967295	回复内容		

表 3-13 景点信息评论表

字段名称	类型	长度	字段说明	主键	默认值
id	bigint		主键	主键	
addtime	timestamp		创建时间		CURRENT_TIMESTAMP
refid	bigint		关联表 id		
userid	bigint		用户 id		
avatarurl	longtext	4294967295	头像		
nickname	varchar	200	用户名		
content	longtext	4294967295	评论内容		
reply	longtext	4294967295	回复内容		

表 3-14 景点门票评论表

字段名称	类型	长度	字段说明	主键	默认值
id	bigint		主键	主键	
addtime	timestamp		创建时间		CURRENT_TIMESTAMP

refid	bigint		关联表 id
userid	bigint		用户 id
avatarurl	longtext	4294967295	头像
nickname	varchar	200	用户名
content	longtext	4294967295	评论内容
reply	longtext	4294967295	回复内容

4 系统实现

4.1 个性化推荐算法模块的实现

本旅行推荐系统的核心思想就是个性化推荐算法。该算法主要是为了实现对每一位用户的个性化服务，而对不同用户进行个性化的景点推荐。为了实现这个目标，核心策略是识别当前用户的相似用户，并向目标用户推荐可能喜爱但未曾听说过的景点^[18]。这个基本思想基于对用户行为的分析，通过找到相似用户，从而向目标用户推荐他们可能感兴趣的景点，增进用户体验。

1) 为了实现这个目标，该系统需要收集用户的行为信息，存储用户的偏好。可用于提高推荐准确率和用户体验，提供个性化推荐服务。如图 4-1 所示。

2) 与搜索推荐系统相类似，但是处理方法不同的是，智能推荐系统更加灵活和适应性强，能够根据用户的历史行为和兴趣偏好，利用模糊属性和协同过滤数据进行推荐。

3) 用户访问阶段的输出推荐列表包括两种：预测列表和推荐列表。前者基于用户需求预测评分并输出，后者则针对用户由高到低提供最感兴趣的项目列表^[19]。

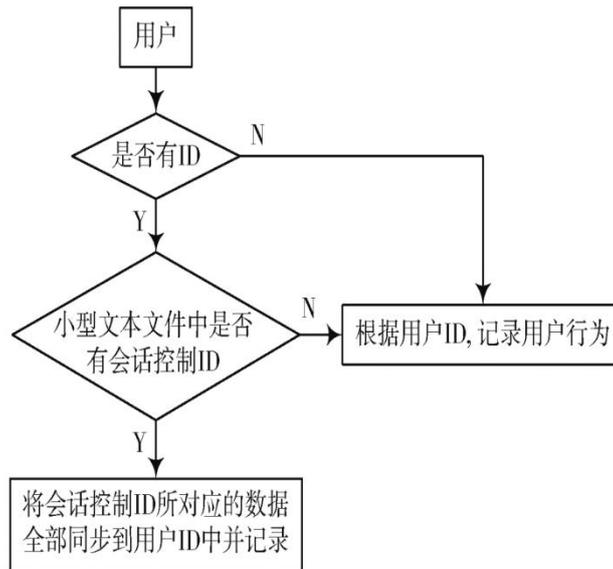


图 4-1 用户信息收集流程图

4.2 前台功能模块

4.2.1 网站首页

该页面是旅游攻略网站的首页，旨在向用户展示网站的全面模块。其中包括了旅游攻略、景点信息、景点门票、景点资讯和在线留言。将轮播图放置在导航栏下方，展示给用户不同的景点推荐信息。中间版块为网站旅游景点推荐展示，向用户推荐不同类型的景点，以提升用户浏览体验。详见图4-2所示。



图 4-2 网站首页

4.1.2 用户注册登录界面

此页面实现用户注册，数据库中以用户名作为Home键的唯一标识，若出现重复会弹出相应提示，则不能重复使用。用户注册时必须填写的内容，包括用户名、密码、电话等在内的报名资料不能为空。详见图4-2所示。

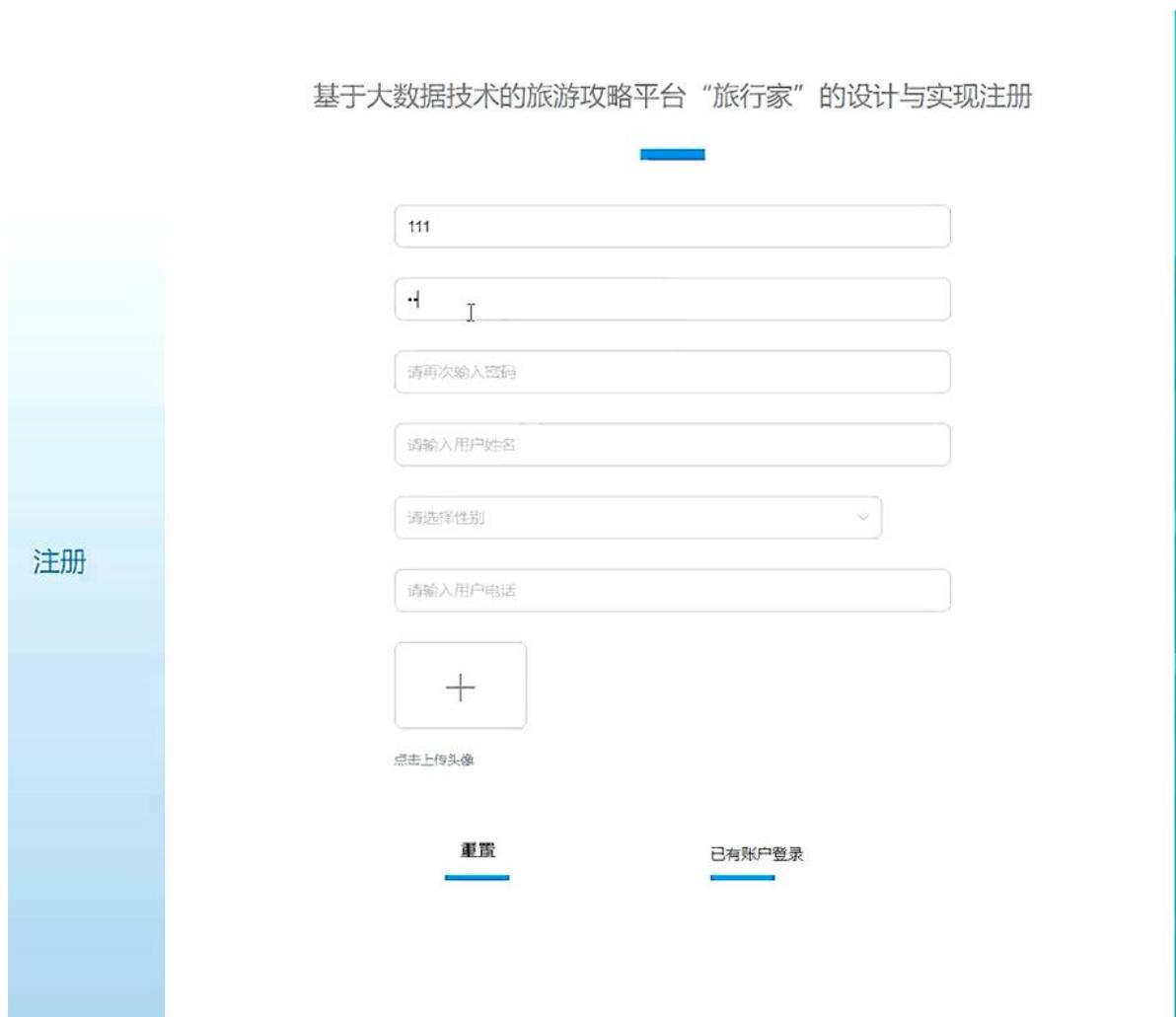


图 4-2 会员注册页面

当注册成为新用户时，需要录入相应的内容，系统会将这些信息封装到一个用户类别中并检测，如果用户名已被占用则系统就会提示“注册用户已存在”；否则会提示用户注册成功。如图4-3所示。

基于大数据技术的旅游攻略平台“旅行家”的设计与实现登录



图 4-3 会员登录页面

4.2.3 景点推荐页面

此页面实现景点推荐信息的查看。详见图4-3所示。



图 4-3 景点推荐页面

4.2.4 景点查看页面

该模块实现了可以让用户对景点进行搜索和查看，可以查询详细的景点信息。如图 4-4 所示。

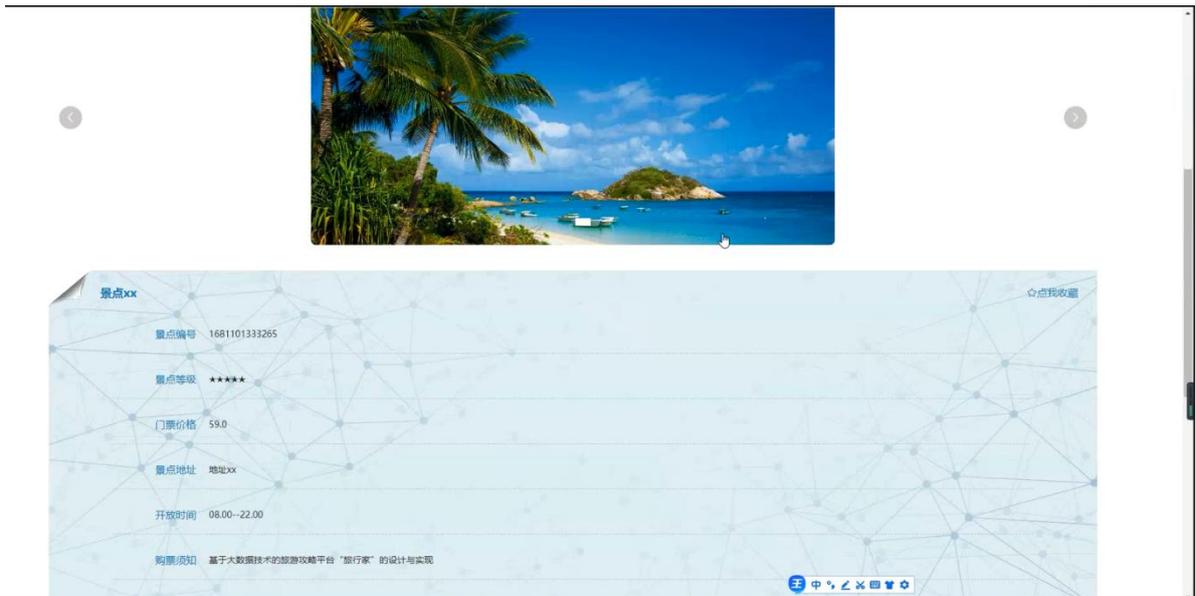


图 4-4 景点页面

4.2.5 门票购买页面

用户进入景点详情页，点击购票，即可进入购票界面，输入购票信息，即可成功购买门票，如图 4-5 所示。

图 4-5 景点预定页面

用户提交购票之后，进入管理中心，可以查看自己的购票信息，进行门票支付，如图 4-6 所示。

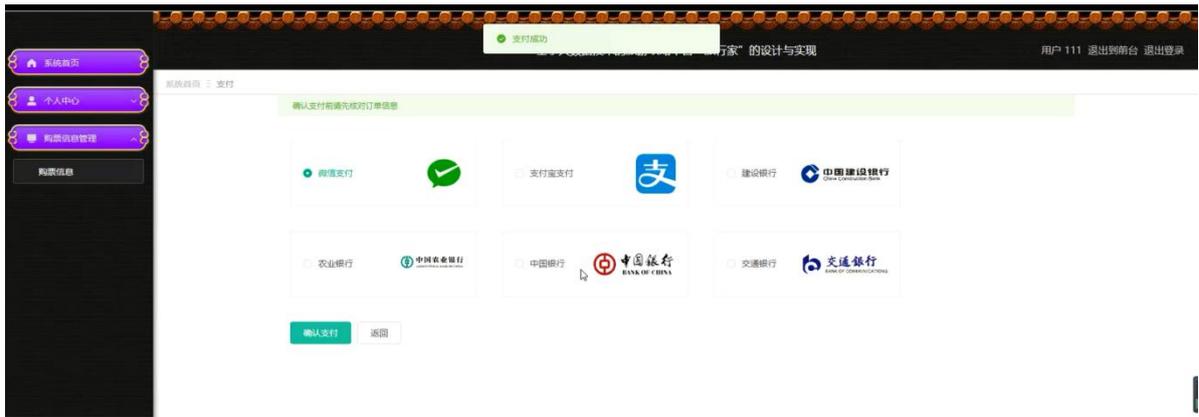


图 4-6 门票支付页面

4.2.6 在线留言页面

进入留言模块，用户可以对留言进行查看，留言内容可由使用者自行公布。图 4-7 所示的留言发布界面。



图 4-7 在线留言页面

4.2 后台管理模块

4.2.1 管理员登录页面

本页面实现管理员登录，在正确选择并填入相应信息后进入。具体如图4-8所示。

基于大数据技术的旅游攻略平台“旅行家”的设计与实现登录

用户名

密码:

管理员 用户

图 4-8 管理员登录页面

4.2.2 购票信息管理页面

本页面为管理员功能，通过查找购票信息，实现管理员对购票信息的查询，以及对购票信息的审核和处理等功能。如图4-9所示。

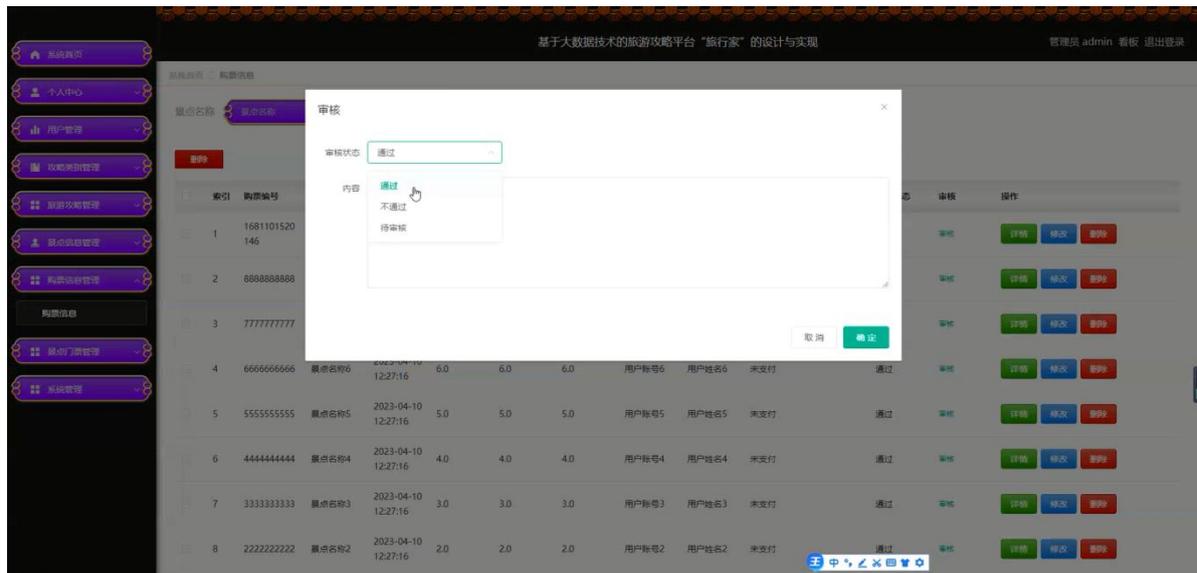


图 4-9 购票审核界面

4.2.3 景点管理页面

该页面实现管理员对景点信息的添加、搜索、查看、修改或删除等操作。如图 4-10 所示。

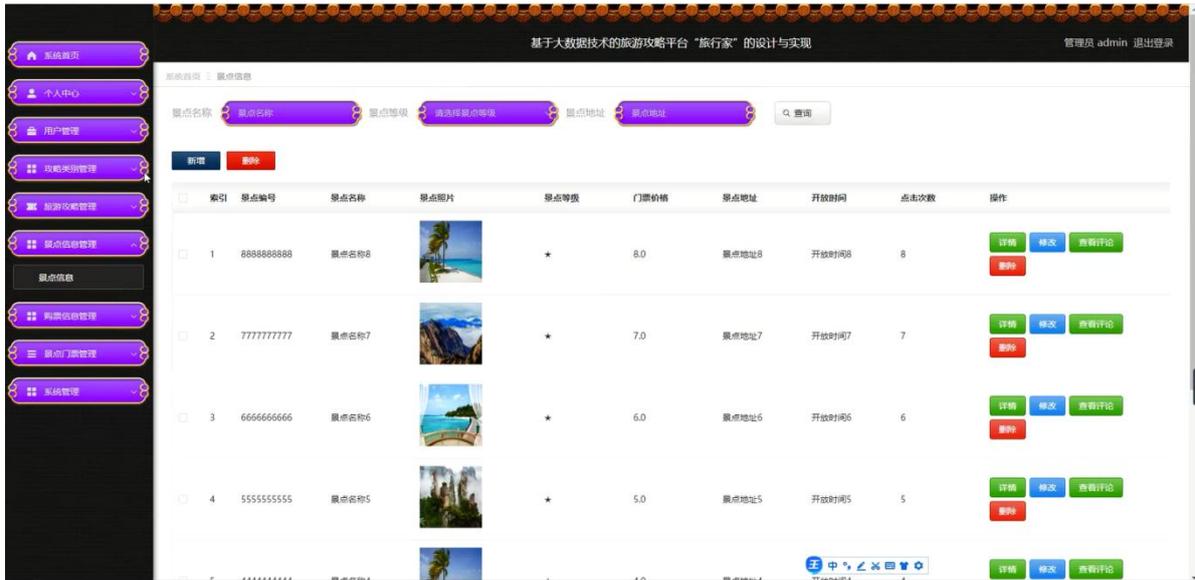


图 4-9 景点管理页面

4.2.3 景点门票管理页面

管理员登陆后台管理系统后，进入景点门票管理模块，可以通过 Python 爬虫技术从多个平台上获取景点门票信息。资料爬取模组主要是将后续模组需要分析处理的相关资料从网页中提取出来。

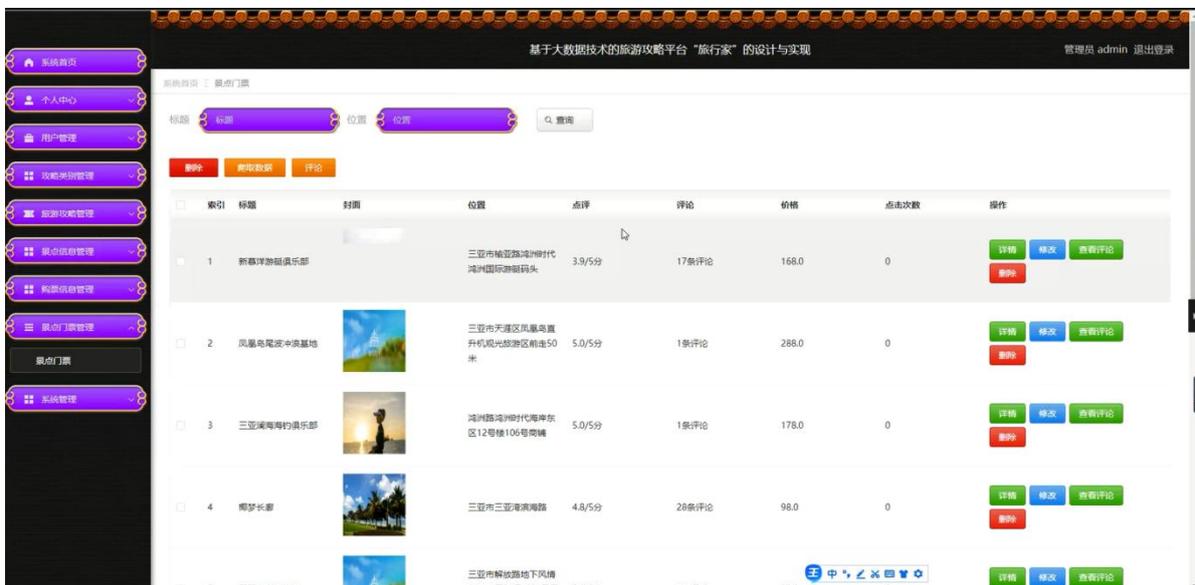


图 4-10 景点门票爬取页面

5 系统测试

5.1 系统测试的定义

软件测试是对软件需求分析、设计规格说明、编码等在软件投入运行前的最后复审，是确保软件品质的关键步骤^[20]。软件测试是一个去查找缺陷的经历，需要花费相当的时间成本来保障程序的正常运行，做到不留遗漏。总的来说为了检验是否满足设计需求。

5.2 系统测试的目的及要求

- 1.通过与系统的需求定义进行比较，发现软件与设计需求是否冲突
- 2.测试系统能否正常运行
- 3.用例应根据需求分析说明书来设计，并在实际使用环境中运行。

5.3 测试用例方法

对各子系统功能模块分别设置若干组数据，在规范的测试流程下，提高测试覆盖率，确保软件得到全面的测试，使测试过程更加高效，发现和修复问题的速度也更快，提高系统的质量、稳定性和效率。

注册测试：

表 6-1 用户注册

用例编号	用例名称	前提条件	测试步骤	预期结果	实际结果比对
TC_001	游客注册	输入的用户名不能是已经存在的用户名	1. 游客点击注册 2. 进入注册页面 3. 游客填写用户名和密码等相关信息，点击提交	注册成功提示并可以在后台数据库查询到该用户数据。	PASS
TC_002	游客注册	输入了不正确或者不完整的数据	1. 游客点击注册 2. 进入注册页面 3. 用户名为空或者用户填写有误，点击提交	系统提示注册失败	PASS
TC_003	游客注册	输入的用户名已经存在	1. 游客点击注册 2. 进入注册页面 3. 游客输入的用户	系统提示注册失败	Failed

和密码等信息已经存在，点击提交

系统登录测试：

表 6-2 系统登录

用例编号	用例名称	前提条件	测试步骤	预期结果	实际结果比对
TC_004	系统登录	输入正确的用户名和密码	用户进入登录页面填写用户名、密码，点击登录按钮	用户登录成功	PASS
TC_005	系统登录	为输入用户名或密码	用户进入登录页面未填写用户名和密码就点击登录按钮	提示用户名和密码错误	PASS
TC_006	游客注册	输入不存在的用户名和密码	用户进入登录页面输入错误的用户名和密码点击登录按钮	提示用户名和密码错误	PASS

购票测试：

表 6-3 景点购票

用例编号	用例名称	前提条件	测试步骤	预期结果	实际结果比对
TC_006	景点购买	用户登入系统	1.选择景点 2.进入购票界面 3.填写购票日期信息等，点击下单按钮	系统提示下单成功并显示购票信息	PASS
TC_007	用户购票	未有用户登入系统	1.选择景点 2.进入购票界面	提示未登入系统	PASS

5.4 测试验收结果

基于协同过滤的旅游推荐系统经过多次反复测试，功能均能正常使用，体验较为良好，后期也能顺利维护。

6 结论与分析

本文介绍了所设计的系统“旅行家”为旅游爱好者提供一站式的旅游攻略服务，可以帮助提升用户的出行旅游质量，解决旅游者面临着信息不对称、行程安排不合理、景点选择不当等问题。

通过大数据技术收集和分析用户的旅游信息，可以为旅游攻略平台提供更准确的数据支持，协同过滤算法是一种常用的个性化推荐算法，基于大数据技术的旅游攻略平台是一种非常有前途的发展方向，可以为旅游业的发展和创新做出重要贡献。提升用户体验和满意度可以通过算法的改进和推荐准确率的提高来推动旅行攻略平台的发展。同时，在旅游攻略平台上应用大数据技术，能够为促进旅游业可持续发展提供更精准、更个性化的推荐服务，为旅游行业的发展和创新做出更多贡献。

参考文献

- [1] 赵宇茹,师军,王祖正.国内旅游信息系统研究概述[J].桂林旅游高等专科学校学报,2006(06):737-740.
- [2] 阳晓萍.B2C旅游电子商务网站个性化服务研究[D].上海:东华大学,2017.
- [3] 朱木易洁,鲍秉坤,徐常胜.协同过滤发展与构建的研究进展[J].南京信息工程大学学报(自然科学版),2017(06):13-20.
- [4] 彭玉,程小平.一种基于自适应近邻选择的协同过滤推荐算法[J].电脑知识与技术:学术版,2016(6):4.
- [5] 程江飞.面向个性化推荐系统的协同过滤算法研究[D].苏州:苏州大学,2019.
- [6] 代金龙.协同过滤算法中数据稀疏性问题研究[D].重庆大学,2019.
- [7] 黄国言,李有超,高建培,等.基于项目属性的用户聚类协同过滤推荐算法[J].计算机工程与设计,2019,31(5):1038-1041.
- [8] Sarwar B M, Karypis G, Konstan J A, et al. Application of dimensionality reduction in recommender systems - a case study[R]. Minnesota Univ Minneapolis Dept of Computer Science,2019.
- [9] Burke R. Recommender Systems: An Introduction, by Dietmar Jannach, Markus Zanker,Alexander Felfernig, and Gerhard Friedrich[J]. International Journal of Human-Computer Interaction, 2020, 28(1):72-73.
- [10]Vora M N. Hadoop-HBase for large-scale data[C]// International Conference on Computer Science and Network Technology. IEEE, 2019:601-605.
- [11]江健锋,徐振平.Springboot最小系统的设计与实现[J].电脑知识与技术:学术版,2021.
- [12]杨庆虎.基于JavaScript的运动框架设计研究[J].计算机时代,2022,(09):123-126+131.
- [13]李艳杰.MySQL数据库下存储过程的设计与应用[J].信息技术与信息化,2021,(01):96-97.
- [14]贾美云.基于知识图谱的汽车交易智能推荐系统研究[D].中北大学,2020.

-
- [15]李占渊,肖蓉,缪仲凯,周双娥.基于改进的协同过滤相似性度量算法研究[J].计算机测量与控制,2017,25(09):287-290+294.
- [16]王一鸣. 基于知识图谱的推荐技术研究及应用[D].电子科技大学,2018.
- [17]楼剑锋. 基于研究性学习的大学生实践创新平台设计与实现[D].吉林大学,2014.
- [18]孙文胜. 基于个性化推荐算法的在线学习交流系统设计与实现[D].山东师范大学,2016.
- [19]郭嘉,代耀宗,沈建京.基于协同过滤的景区旅游智能推荐系统设计[J].现代电子技术,2018,41(23):140-144.
- [20]于敬业. 基于J2EE技术的自动化测试工具的设计与实现[D].北京邮电大学,2007.

致 谢

毕业设计虽然完成了，但是这当中离不开我的专业指导老师曹永忠教授和同学们的帮助。在设计开始时缺乏思路，后来冷静下来，进行查阅资料，咨询老师和同学们，因为有他们的指导，才有了一点头绪。在设计过程中，也是困难重重，遇到每一次的不足，都会请教老师和同学、朋友们帮助，他们不厌其烦的给予支持，在他们的帮助下我才得以克服难关，让我从他们身上学到了，做人做事都要有耐心，才会收获结果。

还要感谢我的父母，没有他们的支持和鼓励，我不可能坚持到现在，是他们不计付出的支持和帮助，让我安心学习，完成大学学业。当我遇到不如意的时候，是他们一次次的鼓励和尊重我的选择，以他们长辈的经验来为我排忧解难。

最后，再次感谢我的专业指导老师曹永忠教授和所有帮助我、陪伴我的朋友和长辈们。