

[旅游信息攻略查询搜索引擎]

需求规格说明书

拟 制 人_____G19_____

审 核 人_____

批 准 人_____

[二零一八年七月十二日]

目 录

1. 引言	4
1.1 目的	4
1.2 文本约定	4
1.3 项目范围	5
1.4 项目背景	5
1.5 定义	6
1.6 参考资料	6
2. 整体描述	7
2.1 产品前景	7
2.2 用户类及其特征	7
2.3 运行环境	7
2.4 设计和实现上的约束	8
2.5 假设和依赖	8
3. 系统特性	9
3.1 用例	9
3.2 IPO 图	22
3.3 状态图	24
3.3.1 信息查询用户	24
3.3.3 管理员	24
3.4 CRC 卡	25
4. 数据需求	29
4.1 数据流图	29
4.2 数据字典	33
4.2.1 用户、管理员登录数据	33
4.2.2 旅游线路数据	33
4.2.3 用户发表评论	33
4.2.4 用户搜索的历史记录	34
4.2.5 搜索过程的索引文档	34

4.3 报告	34
4.3.1 用户搜索记录日志	34
4.3.2 旅游路线对比数据	34
4.3.3 用户评论报告	34
4.3.4 网站流量表	35
4.4 数据获取、整合、保存和处理	35
4.4.1 数据获取与整合	35
4.4.2 数据保存与处理	35
5. 外部接口需求	36
5.1 用户界面	36
5.1.1 界面风格	36
5.1.2 主要接口分布	36
5.2 软件接口	37
5.3 硬件接口	37
5.4 通信接口	37
6. 质量属性	37
6.1 易用性	38
6.2 安全性	38
6.3 性能	38
6.4 其他质量属性需求	38

1. 引言

1.1 目的

本需求说明书的编写目的,主要在于规范旅游信息攻略查询搜索引擎的开发。它说明了本系统的各项功能和性能需求,明确标识各功能的实现过程,阐述实用范围及背景,提供客户解决问题或达到目标所需的条件或权能,提供一个度量和遵循的基准。另一方面,本说明书也是为了提高软件开发过程中的能见度,便于对软件开发过程中的控制与管理,同时也便于程序员和客户之间的交流、协作,并可以作为工作成果的原始依据保存下来。它是用户和开发者对软件初始规定的共同理解,是整个开发工作的基础,它对本系统的功能、性能、用户界面及运行环境等作出详细的描述,以便于系统的进一步开发工作。

本说明书的编写目的可以概括如下:

- 作为旅游信息攻略查询搜索引擎系统开发内容及其约束的参考依据,为开发者提供参考。
- 根据旅游信息攻略查询搜索引擎的运营特点和业务流程,为系统开发者进行设计和编程提供基础。
- 提供测试和验收的依据,即为选取测试用例和进行验收的依据。

本说明书的预期读者包括:

- 项目经理
- 项目开发人员
- 软件质量分析员
- 项目验收人员

1.2 文本约定

本文用 word 编写,约定正文字体为宋体小四号、1.5 倍行距,一级标题为宋体二号加粗,二级标题为等线三号加粗,三级标题为宋体小三号,四级标题为等

线小四号加粗，段首缩进二字符，除此外无其他限制。

1.3 项目范围

最初版本的范围在本文档后续内容会有详细说明，我们注重从功能的实用性上面去考虑，实现了用户旅游信息检索的基本需求，同时根据用户的使用过程进行一些个性化的推荐和功能，推送个性化旅游路线，对于一些更加精细化的需求在后续版本中会一一得到补充和完善。同时架构上保证了产品的质量，为后来的版本更新和迭代打下了良好的基础。

1.4 项目背景

- 待开发的软件系统的名称：
旅游信息攻略查询搜索引擎
- 项目的任务提出：
浙江大学软件工程项目实训课题组
- 开发者：
软件工程项目实训 2 班第 19 组 4 人
- 用户：
搜集旅游攻略的旅游者
旅游信息检索人员
- 实现该软件的计算中心或计算机网络：
基于阿里云的云服务器，可能需要用到多服务器节点
- 该软件系统同其他系统或其他机构的基本的相互来往关系：
可需要从多个旅游门户、攻略网站爬去旅游攻略及信息
由项目实训所有小组及老师共同进行评审工作
- 相关背景介绍
旅游信息攻略查询搜索引擎的构想主要来自于“垂直搜索引擎”的课程要求，同时由于当前市场上多个旅游网站信息比较混杂，内容也比较凌乱，提供一个较为统一的搜索引擎，同时对信息进行一定的归类、关联和分析，可以方便旅游信息检索人员、做旅游攻略的人员检索旅游信息，提升工作的效率，获得使用上的方便。

1.5 定义

认证：防范信息的损坏和泄露。

权限：指用户职能的范围，即各种用户所登录界面、所接触数据、所进行操作等的范围。

用例：用户场景通常称为用例，它提供了系统将如何被使用的描述。

IPO 图：是输入/处理/输出图的简称，描述输入数据、对数据的处理和输出数据之间的关系。

数据流图：是 SA 方法中用于表示系统逻辑模型的一种工具，它以图形的方式描述数据在系统中流动和处理的过程，它是一种功能模型。

CRC 卡：CRC 模型实际上是表示类的标准索引卡片的集合。这些卡片被分为三部分，顶部写类名，下面左侧部分列出类的职责，右侧部分列出类的协作关系。

数据词典：是对所有与系统相关的数据元素的一个有组织的列表、以及精确严格的定义，使得用户和系统分析员对于输入、输出、存储成分和中间计算有共同的理解。

系统集成：将不同的系统，根据应用需要，有机地组合成一个一体化的、功能更加强大的新型系统的过程和方法。

1.6 参考资料

- 《项目实训工作描述(案例二)》
- 《CKCEST 重大工程实践分享》
- 《软件需求（第三版）》
- 《软件工程：实践者的研究方法》

2. 整体描述

2.1 产品前景

旅游信息攻略查询搜索引擎的构想主要来自于“垂直搜索引擎”的课程要求，同时由于当前市场上多个旅游网站信息比较混杂，内容也比较凌乱，提供一个较为统一的搜索引擎，同时对信息进行一定的归类、关联和分析，可以方便旅游信息检索人员、做旅游攻略的人员检索旅游信息，提升工作的效率，获得使用上的方便。

2.2 用户类及其特征

用户类	描述
信息查询者	可以进行简单的关键词查询，查询包含对应名词的旅游路线，另外也可以通过模糊查询根据某些条件来进行旅游路线的筛选，另外可以根据网站提供的综合统计数据进行检查和分析对比，同时会对用户的信息查询行为进行搜集，构建出个性化推荐/推送的内容策略。
管理员	对网站动态爬取到的数据进行查看和管理，监测网站的运行状态，管理网站的内容和服务器运行状态。调整网站的各项功能参数等。

2.3 运行环境

- 开发中需要的计算机系统支持：Linux、Windows 10、MacOS 等
- 运行时所需的计算机系统支持：服务器环境、MongoDB、Apache、SpringCloud

- 对服务器的要求：
 - 服务器的中央处理部件应能满足 300 人同时访问负载的需求
 - 服务器内存必须使用服务器专用内存环境监测
 - 为了防止服务器不可预测的故障，或者服务器的定期维护对整个网站运行的影响，应该有服务器的备用方案
 - 其他性能在经济条件允许的情况下，应该尽量使用高速并且稳定的部件

2.4 设计和实现上的约束

系统的设计、编码和维护文档将遵照 Process Impact 公司内联网开发标准版本 1.3，Java 等 Spring Cloud 框架，所有的 HTML 代码将遵照 HTML5.0 标准，所有脚本都用 Javascript 来编写。

2.5 假设和依赖

系统的使用寿命最短为一个学年，开发的经费来源小组内部解决，硬件条件有 PC 机、云服务器即可。

假定在开发过程中，得到学校，学院和老师同学们的认可和肯定，以及在开发过程中老师同学的积极配合。此外，还需要团队制定良好的计划，组员之间的协调配合。

网站的范围是：1.信息获取，2.信息查询，3.外部链接，4.信息统计。

3. 系统特性

3.1 用例

表 1 用例-用户注册

用例	用户注册
主要参与者	注册用户、游客
目标	用户注册一个账户
前提条件	用户拥有一台可以联网的设备
触发器	用户决定注册一个账户
场景	<ol style="list-style-type: none"> 1. 用户打开浏览器 2. 用户输入网址 3. 用户进入旅游信息查询系统的首页 4. 点击页面上方状态栏中的注册按钮 5. 浏览器跳转至注册页面 6. 用户设置用户名和密码 7. 用户点击注册按钮 8. 账号注册成功
异常	<ol style="list-style-type: none"> 1. 网址无法打开 2. 用户名或密码为空或存在非法字符 3. 用户名已存在 4. 点击注册按钮后界面无反应或显示 404 page not found
优先级	必须的，必须被实现的
使用频率	比较频繁
使用方式	通过浏览器
次要参与者	无
次要参与者使用方式	无
未解决问题	<ol style="list-style-type: none"> 1. 是否需要验证码

	2. 是否需要设置密码的安全等级
	3. 是否需要加入手机/邮箱/密保问题绑定以提高安全性

表 2 用例-用户登录

用例	用户登录
主要参与者	注册用户
目标	用户登录
前提条件	1. 用户已有平台账号 2. 成功进入旅游信息查询系统任意页面 3. 用户未登录
触发器	用户决定登录旅游信息查询系统
场景	1. 用户点击界面上方状态栏登录按钮 2. 用户跳转至登录界面 3. 用户输入用户名和密码 4. 用户点击登录按钮 5. 登录成功后进入跳转前的页面
异常	1. 网址无法打开 2. 用户名或密码错误 3. 点击登录按钮后界面无反应或显示 404 page not found
优先级	必须的，必须被实现的
使用频率	频繁
使用方式	通过浏览器
次要参与者	无
次要参与者使用方式	无
未解决问题	1. 是否需要验证码 2. 是否需要安全证书的认证 3. 如何防止暴力破解

表 3 用例-用户注销

用例	用户注销
主要参与者	注册用户
目标	用户注销
前提条件	1. 用户已登录 2. 用户处于平台某一页面
触发器	用户决定注销登录状态
场景	1. 用户将鼠标移至页面中账号名的上方 2. 出现包含注销选项的下拉菜单 3. 点击注销 4. 注销成功
异常	1. 因太长时间无操作，服务器断开了与浏览器的连接
优先级	必须的，必须被实现的
使用频率	频繁
使用方式	通过浏览器
次要参与者	无
次要参与者使用方式	无
未解决问题	1. 点击注销后是否需要操作确认，以防止误操作

表 4 用例-用户搜索关键词

用例	用户搜索关键词
主要参与者	注册用户、游客
目标	用户根据关键词搜索得到相关联的内容
前提条件	用户进入到搜索主页面
触发器	用户决定搜索某个或某些关键词
场景	1. 用户在输入栏输入搜索关键词 2. 用户点击搜索按钮 3. 出现搜索结果

异常	<ol style="list-style-type: none"> 1. 网址无法打开 2. 搜索字段超过限制 3. 搜索结果不存在 4. 点击搜索按钮后界面无反应或显示 404 page not found
优先级	必须的，必须被实现的
使用频率	频繁
使用方式	通过浏览器
次要参与者	无
次要参与者使用方式	无
未解决问题	<ol style="list-style-type: none"> 1. 如何提高搜索准确性 2. 关于字段最大长度的限制 3. 是否需要加入除文本之外的搜索方式，如图片

表 5 用例-指定类型搜索

用例	指定结果类型搜索
主要参与者	注册用户、游客
目标	得到指定内容类型的搜索结果
前提条件	用户进入搜索主界面
触发器	用户决定获得指定类型的搜索结果
场景	<ol style="list-style-type: none"> 1. 用户点击展开类型选择的下拉菜单 2. 指定一个内容类型（如景点、攻略、路线等） 3. 输入搜索关键字 4. 点击搜索按钮 5. 显示指定类型的搜索结果
异常	<ol style="list-style-type: none"> 1. 网址无法打开 2. 搜索字段超过限制 3. 指定类型的结果不存在 4. 点击搜索按钮后界面无反应或显示 404 page not found

优先级	必须的，必须被实现的
使用频率	频繁
使用方式	通过浏览器
次要参与者	无
次要参与者使用方式	无
未解决问题	1. 是否需要支持类型多选

表 6 用例-根据日期筛选搜索结果

用例	根据日期筛选搜索结果
主要参与者	注册用户、游客
目标	用户得到指定时间段内的搜索结果
前提条件	用户已完成一次基本的搜索
触发器	用户决定查找某时间段内的结果
场景	<ol style="list-style-type: none"> 1. 用户已得到一次基本搜索的结果 2. 用户点击时间选择器选择时间段 3. 单击确定按钮 4. 刷新指定时间段的结果
异常	<ol style="list-style-type: none"> 1. 网址无法打开 2. 起始时间晚于结束时间 3. 指定时间段内的结果为空 4. 点击按钮后界面无反应或显示 404 page not found
优先级	必须的，必须被实现的
使用频率	频繁
使用方式	通过浏览器
次要参与者	无
次要参与者使用方式	无
未解决问题	1. 是否只需提供比较模糊的时间段选择，如一天之内、一个

	月之内等；
--	-------

表 7 用例-根据热门标签筛选搜索结果

用例	根据热门标签筛选搜索结果
主要参与者	注册用户、游客
目标	用户得到指定标签的搜索结果
前提条件	用户已完成一次基本的搜索
触发器	用户决定查找某时间段内的结果
场景	<ol style="list-style-type: none"> 1. 用户已得到基础的搜索结果 2. 用户点击选择内容标签 3. 刷新为指定标签的搜索结果
异常	<ol style="list-style-type: none"> 1. 网址无法打开 2. 指定标签的结果为空 3. 点击按钮后界面无反应或显示 404 page not found
优先级	必须的，必须被实现的
使用频率	频繁
使用方式	通过浏览器
次要参与者	无
次要参与者使用方式	无
未解决问题	<ol style="list-style-type: none"> 1. 是否需要支持多标签搜索 2. 如何在大量标签中选取若干热门标签

表 8 用例-用户查看结果详情

用例	用户查看结果详情
主要参与者	注册用户、游客
目标	用户查看某个结果的详细内容
前提条件	用户已得到搜索结果列表
触发器	用户决定查看某一结果的详细内容

场景	<ol style="list-style-type: none"> 1. 用户已得到一个搜索结果列表 2. 用户点击某一结果 3. 用户进入到该结果的详细内容界面
异常	<ol style="list-style-type: none"> 1. 网址无法打开 2. 点击搜索按钮后界面无反应或显示 404 page not found
优先级	必须的，必须被实现的
使用频率	频繁
使用方式	通过浏览器
次要参与者	无
次要参与者使用方式	无
未解决问题	<ol style="list-style-type: none"> 1. 是否需要支持多标签搜索 2. 如何在大量标签中选取若干热门标签

表 9 用例-用户点赞

用例	用户点赞
主要参与者	注册用户
目标	用户为某个内容点赞
前提条件	<ol style="list-style-type: none"> 1. 用户进入到内容详情界面 2. 用户已登录
触发器	用户决定为某一内容点赞
场景	<ol style="list-style-type: none"> 1. 用户已处在某一内容的详情界面 2. 用户点击点赞按钮 3. 点赞成功，内容的点赞数加一
异常	<ol style="list-style-type: none"> 1. 用户未登录 2. 用户已为该内容点赞
优先级	必须的，必须被实现的
使用频率	比较频繁
使用方式	通过浏览器

次要参与者	无
次要参与者使用方式	无
未解决问题	1. 是否需要使用评分机制来代替点赞功能 2. 点赞效果的设置与按钮的放置位置

表 10 用例-用户取消点赞

用例	用户取消点赞
主要参与者	注册用户
目标	用户将某个已点赞的内容取消点赞
前提条件	1. 用户进入到内容详情界面 2. 该内容已被该用户点赞 3. 用户已登录
触发器	用户决定将某一内容的点赞取消
场景	1. 用户已处在某一已点赞内容的详情界面 2. 用户再次点击点赞按钮 3. 取消点赞成功，内容的点赞数减一
异常	1. 用户尚未给此内容点赞
优先级	必须的，必须被实现的
使用频率	比较频繁
使用方式	通过浏览器
次要参与者	无
次要参与者使用方式	无
未解决问题	

表 11 用例-根据标签查看内容

用例	根据标签查看内容
主要参与者	注册用户、游客
目标	用户通过标签查看内容
前提条件	用户已处于存在标签的页面
触发器	用户决定查看指定标签相关的内容
场景	1. 用户点击某一标签的按钮 2. 显示出打有该标签的内容列表
异常	1. 服务器无响应
优先级	必须的，必须被实现的
使用频率	比较频繁
使用方式	通过浏览器
次要参与者	无
次要参与者使用方式	无
未解决问题	1. 标签的具体设置方式 2. 一个内容是否需要打上多个标签，数量上限该如何设置

表 12 用例-用户查看原文

用例	用户查看原文
主要参与者	注册用户、游客
目标	用户查看内容原文
前提条件	用户已处于存内容详情界面
触发器	用户决定查看原文
场景	1. 用户点击详情界面的原文地址链接 2. 浏览器在新标签页打开原文地址
异常	1. 服务器无响应 2. 原文链接已失效
优先级	必须的，必须被实现的

使用频率	频繁
使用方式	通过浏览器
次要参与者	无
次要参与者使用方式	无
未解决问题	1. 如何尽量确保原文链接的有效性

表 13 用例-查看相似内容列表

用例	查看相似内容列表
主要参与者	注册用户、游客
目标	用户查看相似内容列表
前提条件	用户已处于内容详情界面
触发器	用户滑动页面到页面底部
场景	1. 用户滑动到详情页面的底部 2. 底部出现与该内容相似的其他内容列表
异常	1. 服务器失去响应
优先级	必须的，必须被实现的
使用频率	频繁
使用方式	通过浏览器
次要参与者	无
次要参与者使用方式	无
未解决问题	1. 如何选择更合适的推荐算法 2. 在初始情况下，样本数量、用户数据太少，如何保证推荐结果尽可能合理 3. 相似内容列表刷新频率、显示数量的确定 4. 是否允许用户关闭相似内容推荐功能

表 14 用例-查看相似内容详情

用例	查看相似内容详情
主要参与者	注册用户、游客
目标	用户查看相似内容详情
前提条件	用户获得相似内容列表
触发器	用户决定查看某一相似内容详情
场景	1. 用户点击相似内容列表中的某一选项 2. 浏览器跳转至该内容的详情界面
异常	1. 页面无响应或出现 404 状态提示
优先级	必须的，必须被实现的
使用频率	比较频繁
使用方式	通过浏览器
次要参与者	无
次要参与者使用方式	无
未解决问题	1. 相似内容列表中的每个内容需要如何呈现

表 15 用例-查看推荐内容

用例	查看推荐内容列表
主要参与者	注册用户
目标	用户查看推荐内容列表
前提条件	1. 用户已打开本搜索平台 2. 用户已登录
触发器	用户决定查看自己的个性化推荐列表
场景	1. 用户点击状态栏中的推荐选项 2. 显示系统为该用户推荐的内容列表
异常	1. 页面无响应或出现 404 提示 2. 用户未登录

优先级	必须的，必须被实现的
使用频率	频繁
使用方式	通过浏览器
次要参与者	无
次要参与者使用方式	无
未解决问题	<ol style="list-style-type: none"> 1. 如何选择推荐算法 2. 如何为新用户推荐尽量准确的内容 3. 在平台上线初期用户数据较少时如何通过调整参数、优化算法来尽量保证推荐准确性 4. 推荐内容的数量及刷新频率的确定

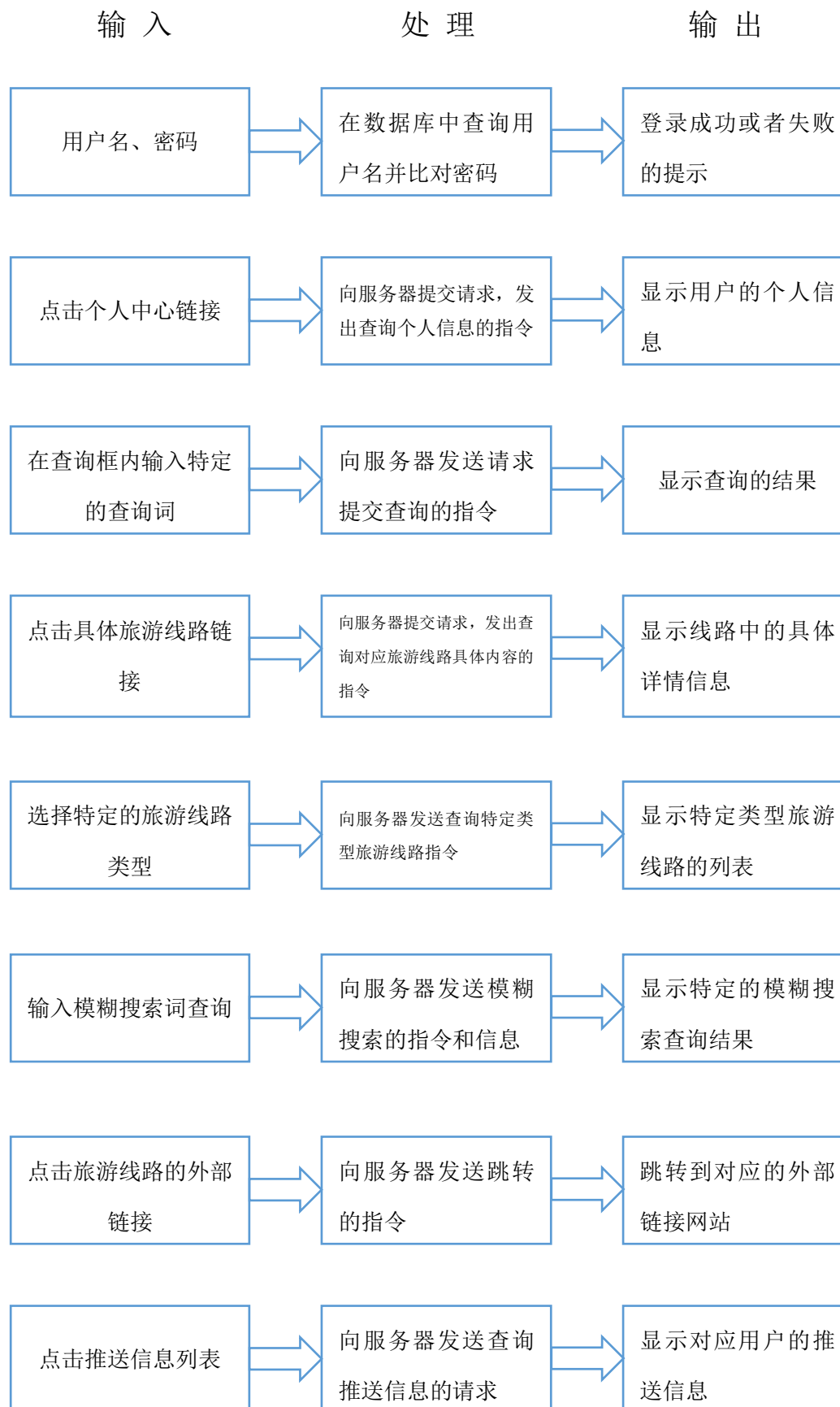
表 16 用例-查看推荐内容详情

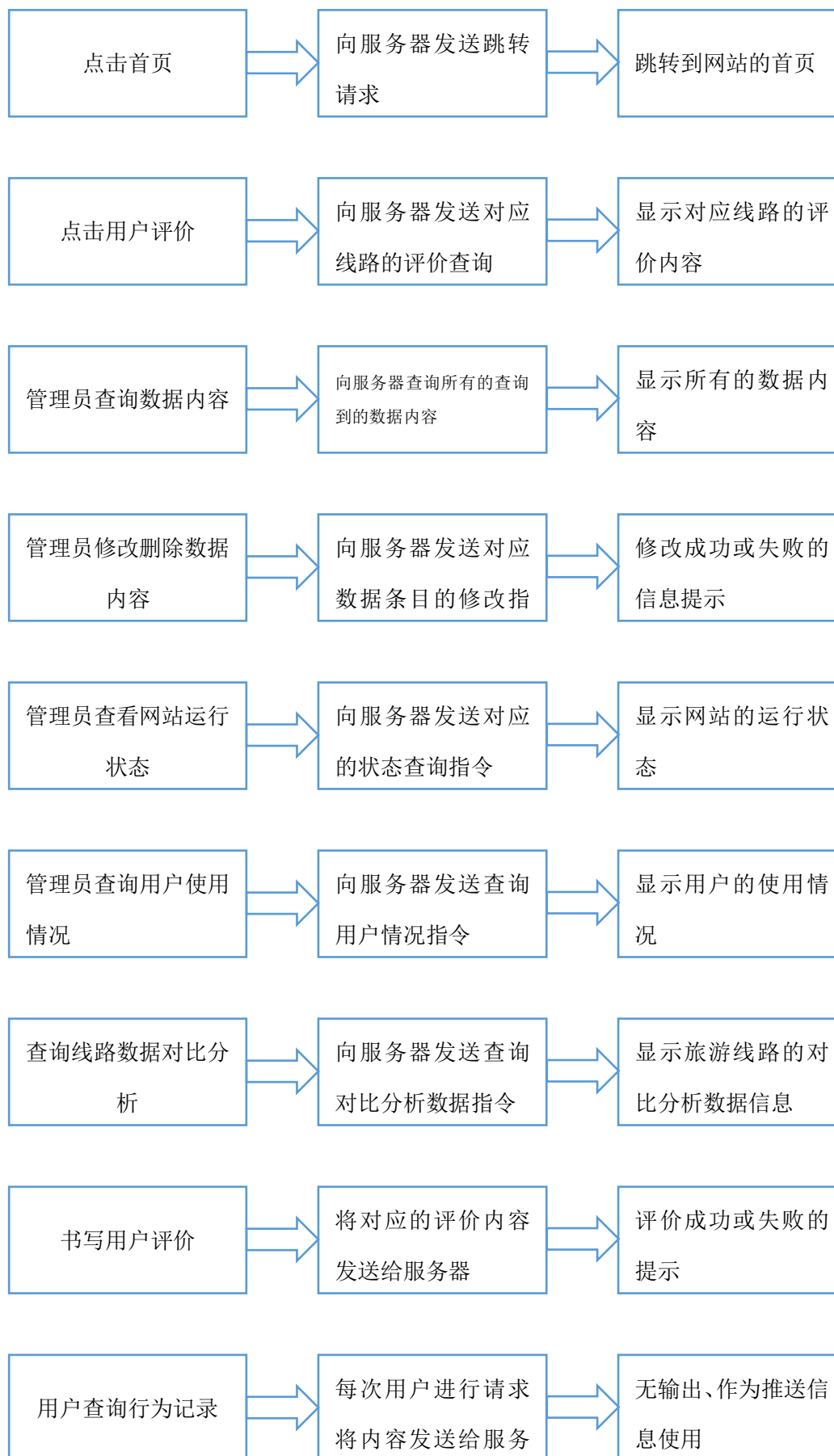
用例	查看推荐内容详情
主要参与者	注册用户
目标	用户查看推荐内容详情
前提条件	用户已处于推荐页面
触发器	用户决定查看某一推荐内容
场景	<ol style="list-style-type: none"> 1. 用户点击某一推荐内容链接 2. 浏览器跳转至该内容详情界面
异常	<ol style="list-style-type: none"> 1. 服务器失去响应
优先级	必须的，必须被实现的
使用频率	比较频繁
使用方式	通过浏览器
次要参与者	无
次要参与者使用方式	无
未解决问题	

表 17 用例-用户评论内容

用例	用户评论内容
主要参与者	注册用户
目标	用户评论内容
前提条件	1. 用户已处于内容详情界面 2. 用户已登录
触发器	用户决评论该内容
场景	1. 用户在评论框中输入评论内容 2. 点击评论按钮 3. 提示评论成功，评论栏中新增评论
异常	1. 用户未登录 2. 服务器无响应
优先级	必须的，必须被实现的
使用频率	比较频繁
使用方式	通过浏览器
次要参与者	无
次要参与者使用方式	无
未解决问题	1. 是否需要限制评论的频率防止灌水 2. 是否需要增加敏感词过滤的功能

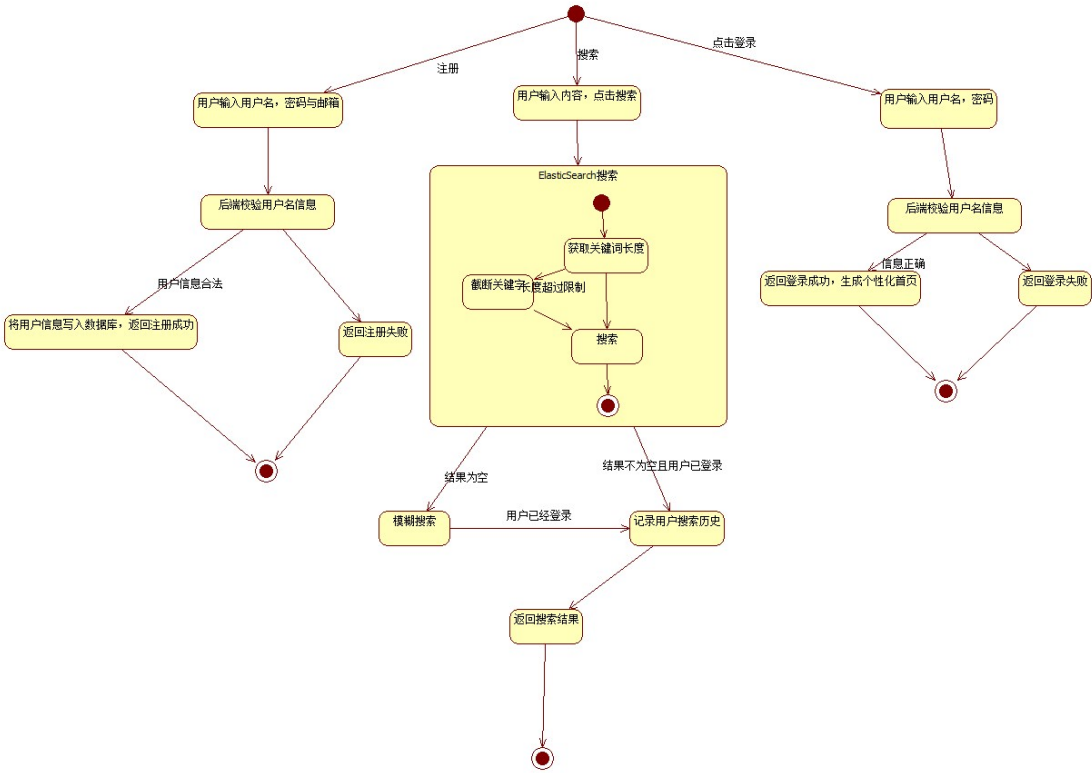
3.2 IPO 图



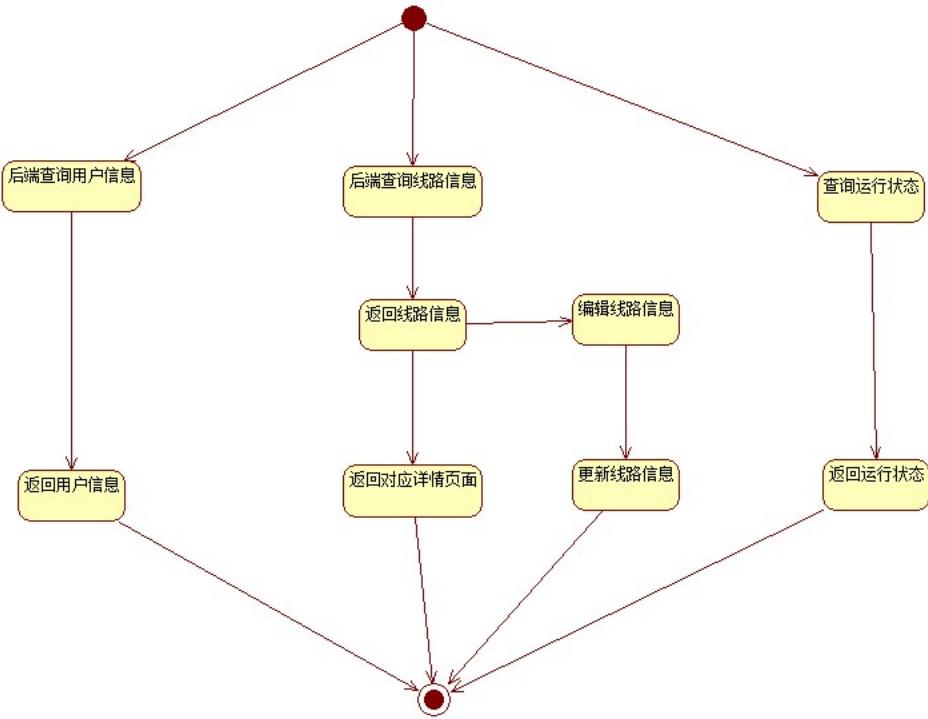


3.3 状态图

3.3.1 信息查询用户



3.3.3 管理员



3.4 CRC 卡

CRC 卡

表 18 CRC 卡——Account 类

类：Account	
说明：	
记录用户账号信息，用于管理用户账号	
职责：	协作者：
创建账号	
验证账号密码	

表 19 CRC 卡——User_State 类

类：User_State	
说明：	
管理用户登录状态	
职责：	协作者：
管理用户登录	Account
注册新账号	Account
注销登录状态	

表 20 CRC 卡——Content 类

类：Content	
说明：	
用于管理具体内容，如攻略、地点信息	
职责：	协作者：
管理内容的具体信息 (包括增删改)	
为上层提供查询接口	

表 21 CRC 卡——Crawler 类

类：Crawler	
说明：	
用于获取数据	
职责：	协作者：
获取原始数据	
对数据进行整理	
将数据存储	Content

表 22 CRC 卡——User_Item 类

类：User_Item	
说明：	
用于生成并储存 User-Item 的评分矩阵	
职责：	协作者：
获取并保存点赞信息	
获取并保存点击信息	
根据点赞和点击信息计算评分矩阵	

表 23 CRC 卡——Similar_Contents 类

类：Similar_Contents	
说明：	
用于生成相似内容列表	
职责：	协作者：
确定相似列表	User_Item
返回相似列表的具体内容	Content

表 24 CRC 卡——Recommended_Contents 类

类：Recommended_Contents	
说明：	
用于生成推荐内容列表	
职责：	协作者：
确定推荐列表	User_Item
返回推荐列表的具体内容	Content

表 25 CRC 卡——Give_a_like 类

类：Give_a_like	
说明：	
用于点赞相关功能的实现	
职责：	协作者：
点赞	User_Item
取消点赞	User_Item

表 26 CRC 卡——Comments 类

类：Comments	
说明：	
用于评论功能的实现	
职责：	协作者：
存储评论	
获取评论列表	

表 27 CRC 卡——Search 类

类：Give_a_like	
说明：	
用于搜索功能的实现	

职责：	协作者：
返回搜索结果	Content

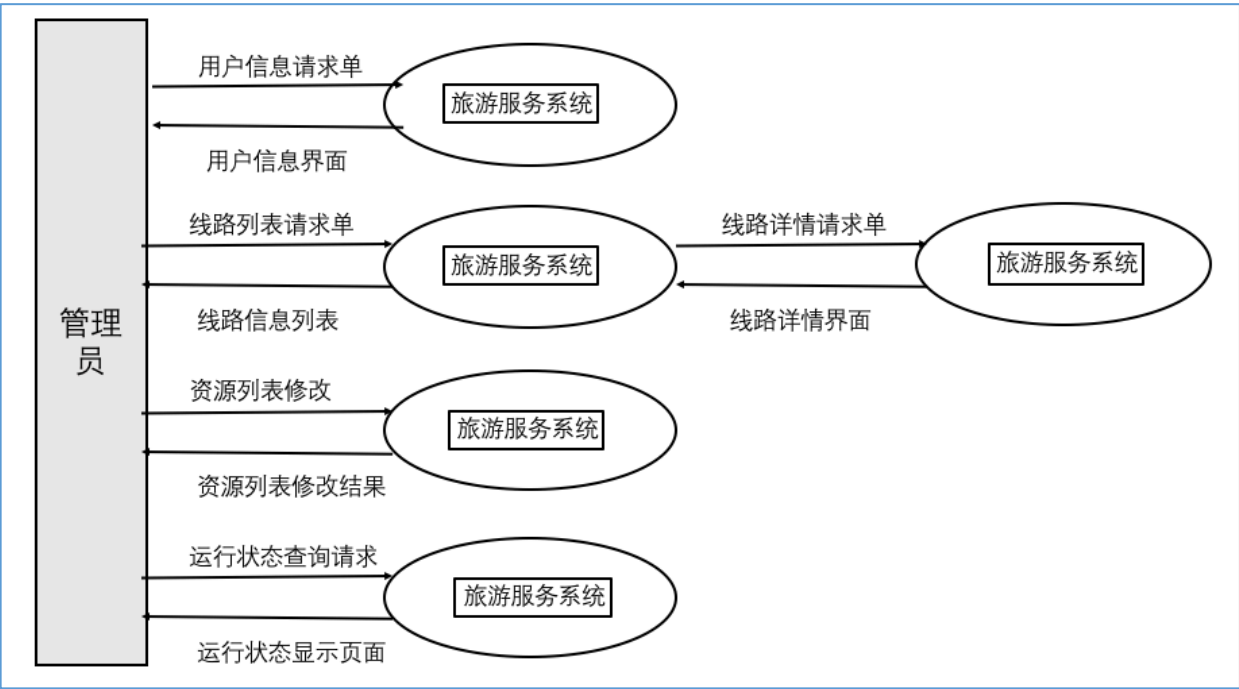
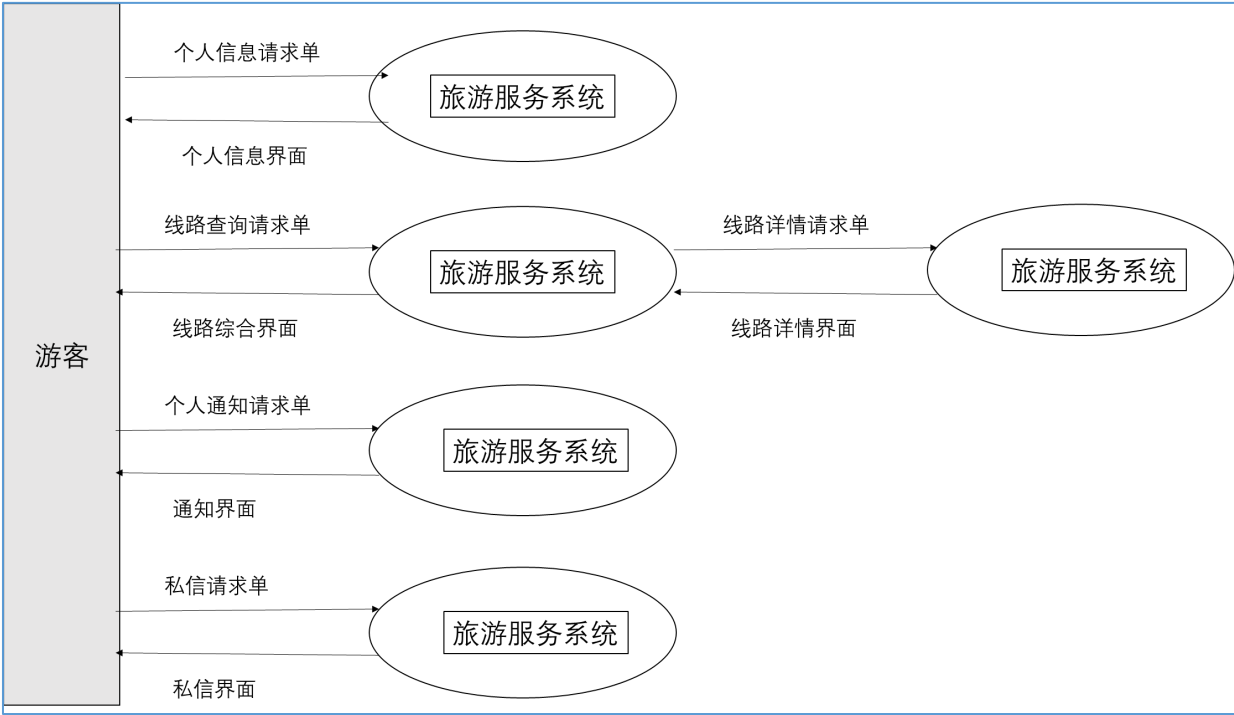
表 28 CRC 卡——Label 类

类：Label	
说明：	
提供基于标签分类的内容展示	
职责：	协作者：
返回指定标签的内容列表	Content

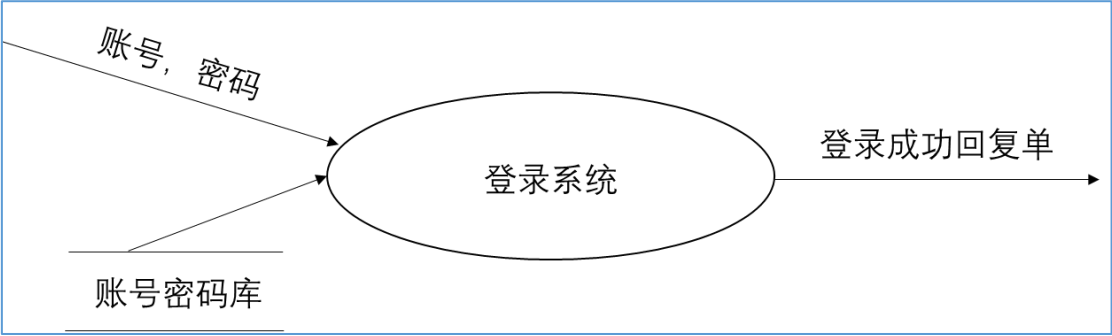
4. 数据需求

4.1 数据流图

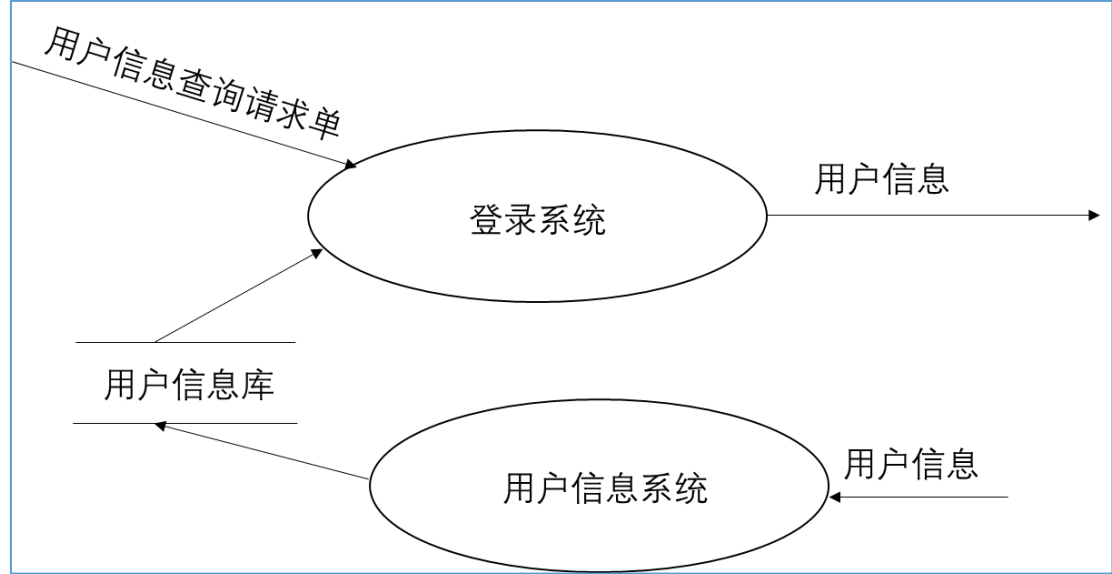
顶层：



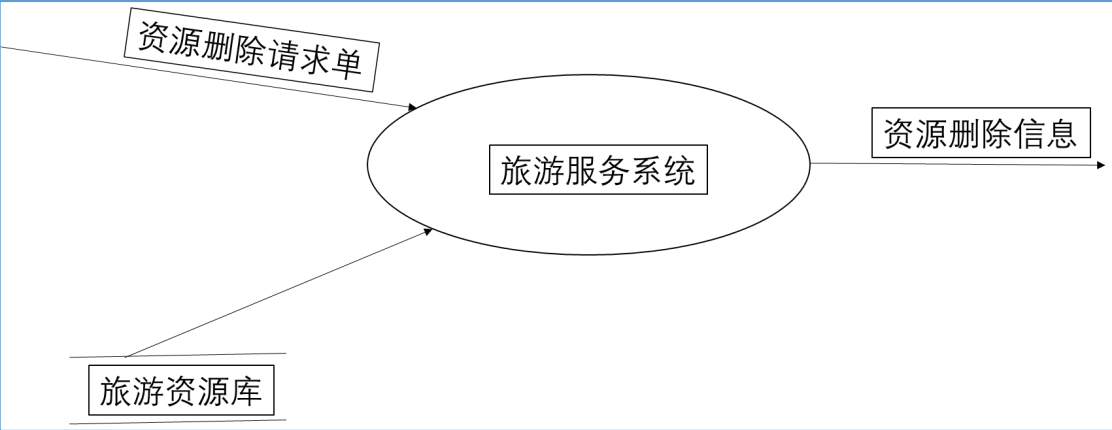
登录第 0 层



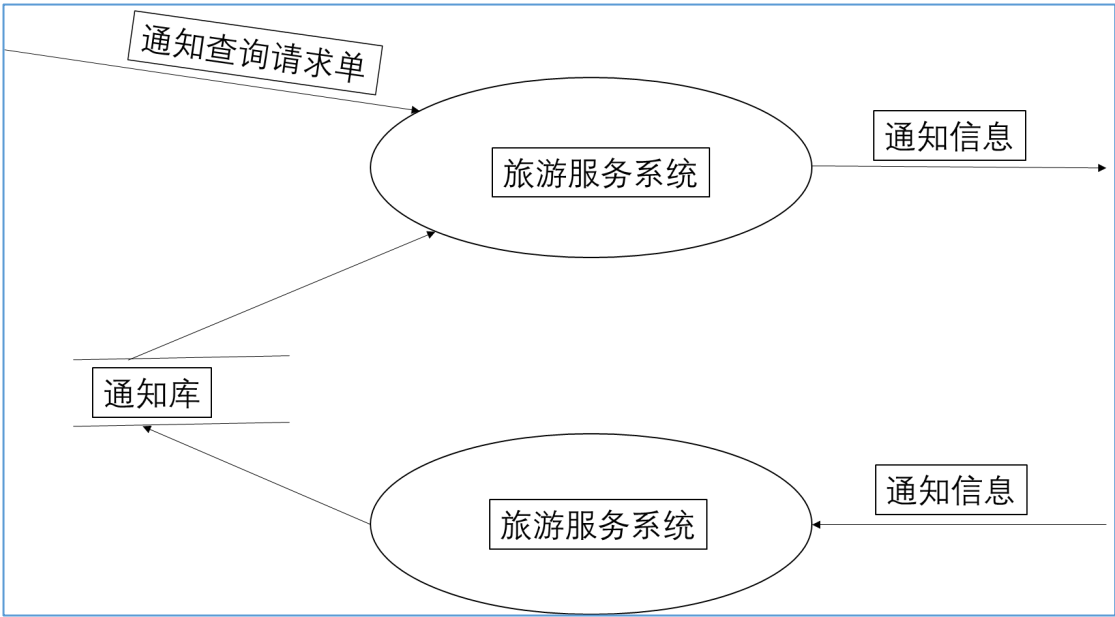
用户信息查询第 0 层



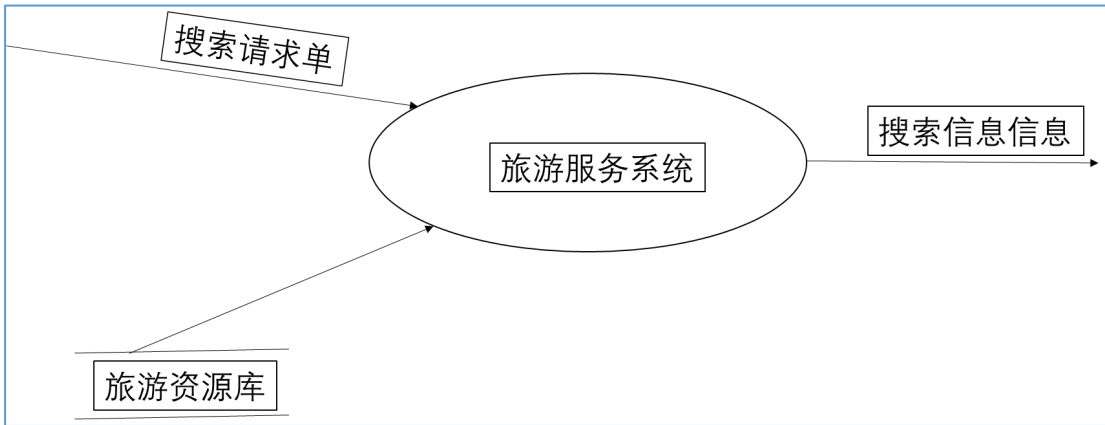
资源删除第 0 层



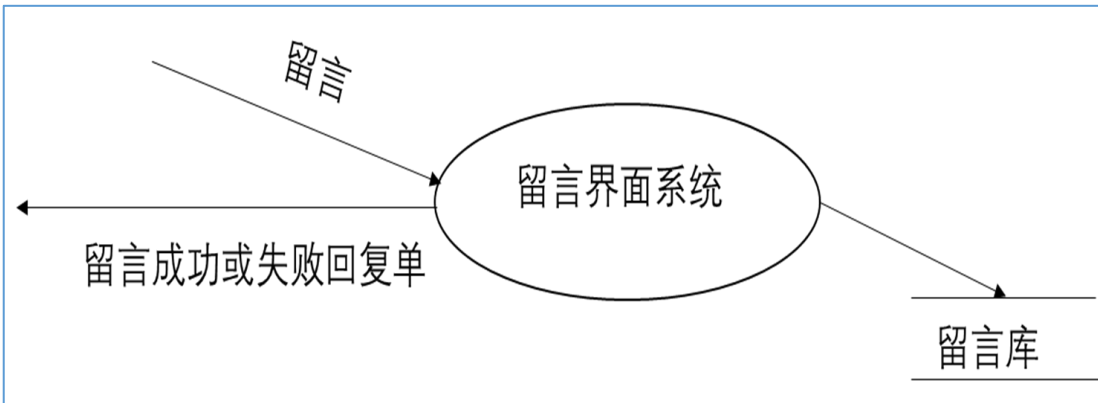
通知查询第 0 层



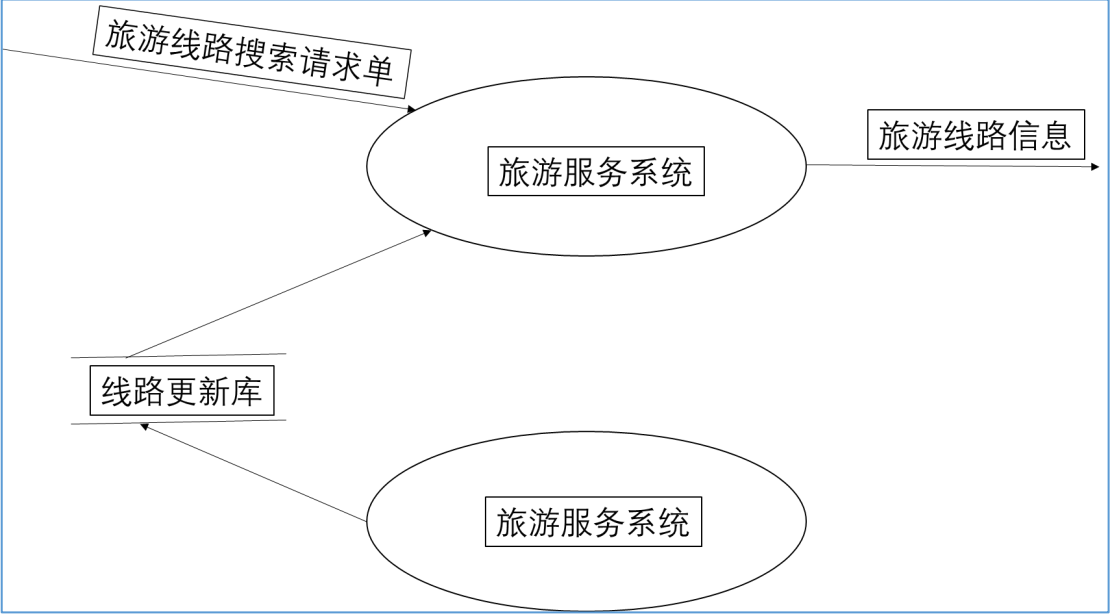
搜索第 0 层



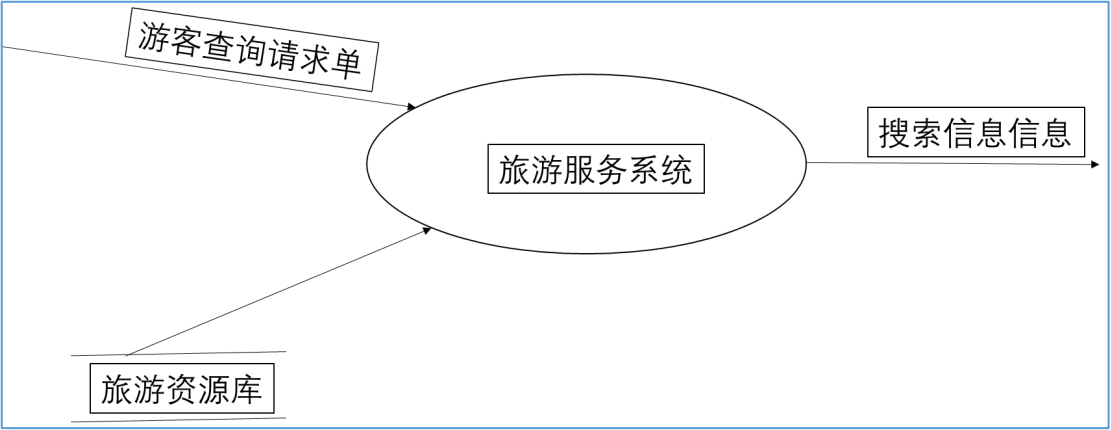
留言第 0 层



旅游线路更新第 0 层



游客查询请求



4.2 数据字典

4.2.1 用户、管理员登录数据

用户信息

数据元素	描述	数据类型	数据长度	数据取值
账号	成员用户名	数字，字母，下划线	3-14 位	0~9,a~z,A~Z,_
邮箱	注册时的邮箱账号	合法的邮箱账户	边长	0~9,a~z,.,@
密码	成员登录密码	数字，字母，下划线	变长，至少为 6 位	0~9, A~Z,,a~z, _

4.2.2 旅游线路数据

数据元素	描述	数据类型	数据长度	数据取值
旅游路线	用于存储网站爬取到的每一条旅游路线的信息	路线名称、路线类型 路线具体内容、来源链接、时间等		文本数据、类型、时间数据

4.2.3 用户发表评论

数据元素	描述	数据类型	数据长度	数据取值
评论	搜索用户发表评论	评论时间 评论 ID 评论内容	有些场合限定一定的数字	文本数据、时间数据

4.2.4 用户搜索的历史记录

数据元素	描述	数据类型	数据长度	数据取值
搜索历史	用户之前搜索过的关键字	字符串	动态	0~9,a~z,A~Z, 所有汉字

4.2.5 搜索过程的索引文档

数据元素	描述	数据类型	数据长度	数据取值
JSON 文档	含有所需属性的文档	字符串	动态	0~9,a~z,A~Z, 所有汉字

4.3 报告

通过网站的运行，将会形成下述报告

4.3.1 用户搜索记录日志

通过对用户搜索历史记录的记录，形成日志性的搜索记录文档，一方面用于网站功能性的 bug 记录及恢复，另一方面用于用户的行为分析，最终形成个性化的推荐服务策略。

4.3.2 旅游路线对比数据

通过多家公司旅游路线数据的收集，形成一定的旅游路线数据积累，用于旅游路线的横向及纵向对比，得出最终的旅游路线对比数据图/报表。

4.3.3 用户评论报告

对用户评论数据进行收集，并形成报告，进而用于搜索结果之中，对于评论较差的路线可以适当降低其优先级。

4.3.4 网站流量表

通过对网站流量的监控，明了一般网站的高峰访问时间，知晓访问人员的规律，方便网站的压力及流量控制。

4.4 数据获取、整合、保存和处理

4.4.1 数据获取与整合

4.4.1.1 旅游路线数据获取

通过定时的动态爬虫，获取旅游网站上的旅游路线数据，并加以整合存储。

4.4.1.2 用户搜索数据获取

通过对用户搜索记录和评论记录的获取，储存用户的个性化数据，用于推荐和推送。

4.4.2 数据保存与处理

爬取和获取的数据保存至网站服务器，由数据处理对象对数据进行处理如

4.4.2.1 旅游路线数据处理

对爬取到的旅游路线数据进行一定的过滤，并且分类聚合，最终形成搜索引擎所展示的旅游路线数据。

4.4.2.2 搜索用户数据处理

对用户数据简单记录后，通过程序化处理或机器学习的方法进行用户行为分

析，并最终得出个性化推荐的方案

4.4.2.3 管理员处理数据

管理员定期对网站的数据进行删除，管理，监视等

5. 外部接口需求

5.1 用户界面

5.1.1 界面风格

- 采用常见的旅游搜索平台的界面风格，整体风格清晰明亮；
- 整个系统所有界面的色彩搭配、字体风格、窗口布局保持一致；
- 每个页面的功能分布合理，图片文字搭配适当；

5.1.2 主要接口分布

一级链接	二级链接	三级链接	四级链接
搜索引擎主页	登录页面	注册页面	
搜索页面	搜索结果列表	详细线路信息	查看外部链接
			评论
			查看相似路线
		同类路线对比	进入某一具体路线
			按照某一标准进行排序

推送列表	个人推送列表	某一具体路线	推送旅游
			查看具体路线
		删除推送记录	/
管理员界面	用户信息查看	用户添加、删除	
	网站运行状态	查看具体条目	
	资源列表查看	修改资源内容	
		删除资源条目	
		管理评论信息	评论删除

5.2 软件接口

该系统通过 Elastic Search 提供的 API 实现与数据库的连接，支持各种操作系统，IE11 及以上的浏览器，用户通过浏览器访问该系统。

5.3 硬件接口

没有与硬件的直接连接。

5.4 通信接口

将系统部署在 web 服务器的指定目录下即可；用户可通过浏览器访问该平台。

6. 质量属性

本节主要规定非功能性需求和外部界面需求。下面根据用户访谈来考虑质量属性并按照优先级逐个进行描述是相关的质量需求。

6.1 易用性

界面风格统一，主色调不超过三四种为宜，导航目录条理清晰，排列有序，界面文字排列有序。整体画面简洁清晰，让人一目了然。

对用户的操作耗时比较长的，要有反馈，如下载文件要有下载进度显示。对用户一般的操作，要有一定的容错能力以及处理用户不当输入的能力。

可以提供联机帮助或使用手册，不做刚需。

6.2 安全性

权限控制，根据不同的用户角色设置相应的权限。用户重要的操作应该做相应的日志进行记录，没有相应权限的用户禁止使用相关的功能，如学生只能看自己的成绩，教师只能查看自己的学生的成绩。

系统登陆需要密码。重要的数据进行加密处理，如登陆名以及密码，应经过加密。

数据备份，系统应进行常规的数据备份，以防止以外情况导致的数据丢失。

6.3 性能

客户端一般的用户操作响应时间不超过一秒。支持教学班容量的数量的学生和老师以及额外加上游客并发使用。

6.4 其他质量属性需求

系统能长时间工作，除了例行维护以外。编码符合规范，方便后期维护。对系统模块化，方便对于用户的需求修改做出响应。