**电子科技大学计算机科学与工程学院**

**标 准 实 验 报 告**

**（实验）课程名称 数据库新技术**

**电子科技大学教务处制表**

**电 子 科 技 大 学**

**实 验 报 告**

**工作方式：（个人/小组）（ 小组 ） 指导教师：牛新征**

**实验地点： 主楼A2-412 实验时间：2018.11.22**

**设计者：（组长） 王洋洋 ，学号： 201852081022**

**（组员） 李斌 ，学号： 201822080137**

**（组员） 何宗奎 ，学号： 201822080111**

**（组员） 杨俊 ，学号： 201822080119**

**（组员） 姜斯文 ，学号： 201722060131**

**一、实验室名称： 计算机学院实验示范中心**

**二、实验项目名称：数据库需求设计训练**

**三、实验学时：4学时**

**四、实验原理：**

1. 为需求人员明确目的；
2. 为编码人员提供依据；
3. 为修改、维护提供条件；
4. 项目负责人将按计划书的要求布置和控制开发工作全过程；
5. 项目质量保证组将按此计划书做阶段性和总结性的质量验证和确认。

**五、实验目的：**

给出电影数据分析系统（包括数据库）的软件需求，包括最终实现的系统必须满足的功能、系统整体组织结构、系统功能模块划分、系统内部接口设计和用户界面设计以及设计约束。

# 第一章引言

## 目的

本文档主要是根据目前电影数量繁多，依据用户期望获取更好、更适应自己的电影推荐系统的迫切需求，经过详细分析后整理出来的，描述开发电影推荐系统的需求相关的文档。本文档一方面的目的是指导项目成员了解最终的业务需求，另外一方面是作为后续开发设计的指导文档。

## 背景

在互联网技术不断成熟、移动终端快速普及的背景下，人民开始追求精神上的享受，精神需要在人民的需求中所占的比重越来越大，电影作为一种重要的精神娱乐方式，极大的受到人民的欢迎。当今，互联网上的各类视频网站比比皆是，这些站点都提供数量巨大的电影视频供用户观看，用户可以较为容易的获取电影资源。

虽然用户阅览电影的要求降低，但是用户总是需要花费大量的时间对电影进行挑选。从海量的电影中获取当下流行电影的信息，筛选搜索自己喜欢类型的电影，具有很大的实用价值。

## 术语

本文档中所涉及的一些专业术语如下：

MongoDB：一个基于分布式文件存储的数据库。旨在为WEB应用提供可扩展的高性能数据存储解决方案。

nodejs:开源Web服务器，用于解释JSP、Servlet文件，提供Web服务。

## 期望读者

本产品需求分析报告所针对的预期读者包括：

1.用户：只需关注项目分析、系统功能描述、项目运行环境，根据项目硬件和软件要求配置自己的计算机以达到运行项目的。

2.开发人员：严格按照此文档进行实施，需要关注项目的系统功能详细分析，系统流程图，数据流图，以及数据库设计功能，了解项目中的各个子系统之间的关系，各功能之间的处理以及实现要求。

3.测试人员：此文档作为测试的依据，需要关注项目的具体功能实现，了解项目报告中的系统功能要求，根据测试结果判断是否达到了系统的基本要求。

4.项目经理：根据此文档安排项目进度以及人员安排。

5.文档编写员：需关注项目报告中的每一项任务，针对报告中的每个模块，仔细查看，力争将系统的功能，项目分析，项目参与人员以及整个项目的架构清晰的抽象出来以展示给不同的预读人员。

## 1.5 参考文献

1. 《计算机软件工程规范国家标准汇编》，中国标准出版社，2000.7.包含以下标准：
2. GB/T 11457-1995《软件工程术语》；
3. GB/T 8566-1995《信息技术软件生存期过程》；
4. GB/T 8567-1988 《计算机软件产品开发文件编制指南》；
5. GB/T 1526-1989 《信息处理产品数据流程图的文件编制符号及约定》；
6. GB/T 9385-1988 《计算机软件需求说明编制指南》；
7. 教育部教育信息化标准:
8. JY/T 1001-2012《教育管理信息教育管理基础代码》。

# 第二章 任务分析

## 2.1 开发背景

互联网在最近几十年出现了爆发性的增长，伴随着用户数量迅速增长的背后是用户数据量的指数级增长，面对着海量的信息，用户往往会感到束手无策，这就是互联网中所谓的信息过载问题。如何帮助用户从海量的信息中筛选获取用户最感兴趣的信息已经成为了一项热门的研究工作。

在2018年，就有37部电影票房过亿，更有10部票房超过10亿，面对如此众多的电影信息，用户总是希望从中进行一定的筛选，以便挑选中比较好的电影进行观看。

## 2.2 项目目标

系统的总体目标是：利用先进的计算机网络及爬虫技术，对2018年的猫眼电影数据进行爬取，并且进行分析，将数据按照各种要求进行最优化的寻找并且通过可视化的处理，使得用户能够根据需要，获得这一条件下的电影的相关信息，并可以根据相关信息进行电影推荐的操作。

## 2.3 项目介绍

本系统设计了几大子系统：数据爬取、数据维护、电影信息展示。

数据爬取业务主要包括从猫眼电影数据进行爬取，并且将其存入数据库中。该业务获取的基础数据是整个系统的根基，关系到后续功能的实现。

数据维护主要是对其中的数据进行管理，能够保证数据能个得到及时且有效的管理。

电影信息展示主要是按照用户的需求对所筛选出的电影信息进行对于的展示，方便用户进行阅读以及进行后续的电影推荐。

## 2.4 用户类型与特性

本系统的主要用户为以下几种：

1.管理员：管理各种电影的信息对系统进行设置和维护。并负责从网络中爬取即时的电影信息，以保证其中数据的时效性。

管理员作为该系统总体的管理人员，负责系统的正常运行和管理。

2.观影用户：能够在线对电影数据按照自己的需求进行筛选，并查看筛选后的电影信息，进行电影推荐。他们并不一定具备良好的计算机知识，且对系统的使用频率较高，因为要求前端功能使用设计简单易操作的同时，系统具备较高的并发性和稳定性。

表 2-1 用户类型和特性表

|  |  |  |
| --- | --- | --- |
| 编号 | 用户类别 | 特征描述 |
| 1 | 管理员 | 对系统进行设置管理，并且保证数据的时效性与正确性 |
| 2 | 观影用户 | 使用系统进行筛选出自己感兴趣的电影的信息同时可以进行行电影推荐的操作 |

## 2.5 运行环境

### 2.5.1 服务器环境

硬件环境：Intel Pentuim4 Xeon 2.5GHz CPU、160GB Raid 5 磁盘阵列、16GB DDR266 内存

操作系统：Microsoft Windows 2003 Server

支持环境：nodejs web

数 据 库：MongoDB 2.4.8

### 2.5.2 用户环境

硬件环境：CPU 1.0G以上

内存256M以上

硬盘20G以上

操作系统：Microsoft Windows 2000，XP，VISTA，win7，win8，win10

支持环境：建议使用IE8.0以上及Chrome浏览器

## 2.6 设计和实现上的限制

开发人员需掌握的技术：

1.MongoDB 2.4.8的应用

2.Python 3.6的使用

3.nodejs web本地服务器调试技巧

# 

# 第三章 外部接口需求

## 3.1 用户界面

根据用户操作需要，设计简洁美观的、友好的Web用户界面，为用户提供简单、直观的操作界面。

主要遵循以下原则：

1、界面元素符合用户习惯；

2、操作简易性；

3、界面一致化；

4、以用户体验至上为原则。

## 3.2 软件接口

服务器程序通过接口对MongoDB进行访问。

# 第四章 系统功能需求

## 4.1 系统功能结构

### 4.1.1 管理员子系统

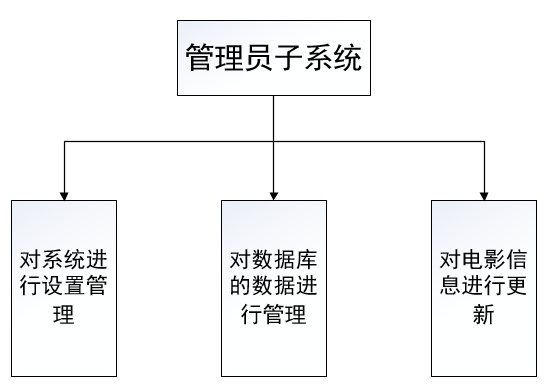


图4-1 管理员子系统图

1.对系统进行设置管理

对系统的各种配置情况进行调整，以适应当前情况。

2.对数据库中的数据进行管理

对数据库中的数据进行各种管理，包括对影片信息的更改，对影片信息的删除，对影片信息的查询。

3.对电影信息进行更新

通过爬虫，保持数据库中数据的及时更新。

### 4.1.2 观影用户子系统

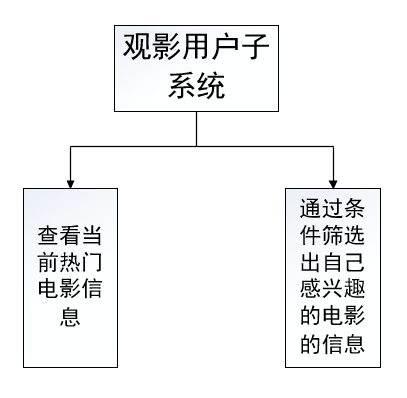


图4-2 观影用户子系统

1.查看当前热门电影信息

用户可以查看票房当前最高的十部电影，每月上映电影的数量，每月电影票房情况，按电影类型查看电影信息，并进行可以电影推荐的操作。

2.通过条件筛选出自己感兴趣电影的信息

用户可以设置筛选条件，从而筛选出满足条件的电影的信息。

## 4.2 分析模型

以数据流图对本系统进行分析。

## 4.3 数据说明

### 4.3.1 相关的数据流说明

电影信息=编号+电影名+类型+出版地区+电影时长+电影首映日期+电影评分+观影人数+电影大小

评论信息=编号+评论者+评论时间+点击量+得分+评论内容

### 4.3.2 数据流图

1.顶层数据流图

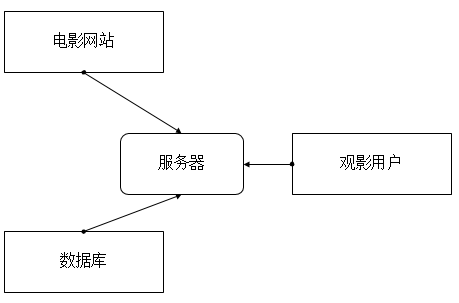


图4-3 顶层数据流图

2.一层数据流图

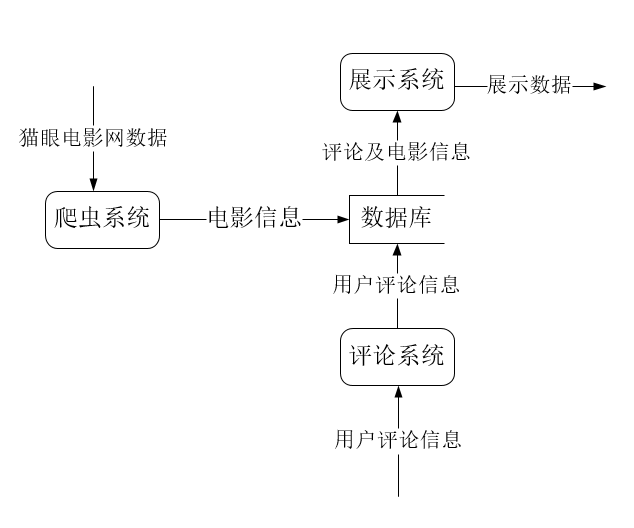


图4-4 一层数据流图

# 第五章 非功能性需求

## 5.1 安全性要求

在本平台中，主要涉及到身份安全认证、数据传输安全、数据备份恢复机制以及容灾方案的设计等。

## 5.2 美观性要求

界面要求简介美观，避免过于繁乱复杂。

## 5.3 简洁性要求

根据用户操作的需要，设计简洁美观、友好的用户界面，为用户提供简单、直观的操作界面。

## 5.4 数据库设计需求

数据库中的表要满足3NF减少数据的冗余度。

## 5.5 性能需求

1、在最大并发量访问下，系统仍能提供服务，不出现致命性错误。

2、一般操作的响应时间应在1秒内，响应时间也应在可接受的时间内完成，具体情况以网络速度为限制。