

电影数据可视化分析系统（PC端）

星辰电影

软件需求规格说明书

名字	学号	贡献描述
肖安琪	SZ160110130	（用户界面编写，界面后台响应编写） 1、5 部分，审核
陈筱雯	SZ160110104	（爬虫程序编写） 4 部分
闫宏书	SZ160110118	（用户界面编写，界面后台响应编写） 3 部分
李昕阳	SZ160110123	（电影数据处理与图表绘制） 3 部分
祝清麟	SZ160110211	（后台功能编写，爬虫程序改进） 2 部分

文档变更历史记录

[illegible]

目录

1. 引言	4
1.1 编写目的	4
1.2 读者对象	4
1.3 软件项目概述	5
1.4 文档概述	5
1.5 定义	6
1.6 参考资料	7
2. 软件的一般性描述	8
2.1 软件产品与其环境之间的关系	8
2.2 限制与约束	8
2.3 假设与前提条件	9
3. 功能需求描述	10
3.1 系统的划分	10
3.2 各子系统的功能	11
3.2.1 系统的用例图	11
3.2.2 各子系统用例的描述	11
4. 其它需求描述	21
4.1 性能要求	21
4.1.1 响应速率	21
4.1.2 运行效率	21
4.2 设计约束	21
4.3 界面要求	22
4.4 进度要求	26
4.5 交付要求	26
4.6 验收要求	27
5. 软件原型	28

1. 引言

1.1 编写目的

本文档旨在对本小组开发的（“电影数据可视化分析系统（PC 端）”）“星辰电影”软件项目进行定义，在该文档中详尽说明整个软件项目的软件需求以及功能，包括修正记录。

1.2 读者对象

本文档的读者对象如下：

1. （“电影数据可视化分析系统（PC 端）”）“星辰电影”软件的开发、测试、维护人员
2. 要求开发（“电影数据可视化分析系统（PC 端）”）“星辰电影”软件的客户
3. （“电影数据可视化分析系统（PC 端）”）“星辰电影”软件的使用者
4. 负责协调（“电影数据可视化分析系统（PC 端）”）“星辰电影”软件引起的纠纷的司法机构工作者

针对开发项目组的人员（包含测试、维护），本文档有利于开发人员明确软件架构和开发目标，能指导各参与开发的成员所完成的部分的接口的对接，并且在交付时能将本文档作为验收标准，避免客户提出额外的需求而引起纠纷。

针对用户群体，本文档有利于用户明确软件实现的功能是否满足其需求，对不符合期望的地方能及时提出整改意见，并且能对软件所实现的功能有一个系统而全面的认知，方便交付时对各个功能是否全部实现进行检查。

针对司法机构成员，本文档中涉及到的软件功能即为开发小组与客户之间合同保护范围内需要实现的功能，其可作为利益纠纷产生时的佐证。

1.3 软件项目概述

- 项目名称：
 - 星辰电影（电影数据可视化分析系统（PC 端））
- 用户单位：
 - 哈尔滨工业大学（深圳）2016 级软件工程课程负责组
- 开发单位：
 - 哈尔滨工业大学（深圳）2016 级软件工程课程项目编写小组（2 组）
- 大致功能和用途：
 - 功能：

软件实现对 2015（.1.1）-2018（.11.30）时间段内于中国大陆上映且记录在猫眼电影上的各电影的信息（标题、题材、导演、主要演员、累计票房等）的分析，以图表的形式将数据可视化，即能够根据时间、题材、演员等信息提取符合条件的信息进行详细的绘图，并能对分析出的结果生成报表保存下来。
 - 用途：

本软件旨在：

 - 对电影界业内人士，软件能指导其了解当下电影的市场行情，指导其未来的影视工作活动。
 - 对一般用户，软件能提供其感兴趣的电影数据，帮助其淘取感兴趣的电影、了解知名演员等，满足娱乐方面的需求。

1.4 文档概述

本文档主要介绍本小组开发的软件（“电影数据可视化分析系统（PC 端）”）“星辰电影”的用户群体、开发约束、基本功能与各功能间的连接方式、非功能性需求以及开发原型，并且记录有本文档的编写者以及修改记录。

本文档总体共分为五个部分：引言、软件的一般性描述、功能需求描述、其他需求描述和软件原型，对整个（“电影数据可视化分析系统（PC 端）”）“星辰电影”进行介绍。

引言部分主要介绍本文档的撰写目的、结构、内容、读者对象、参考资料。

软件的一般性描述部分主要介绍软件的开发环境、条件限制、假设和前提。

功能需求描述部分主要介绍软件系统架构、各个子系统用例图与描述。

其他需求描述部分主要介绍软件的性能、界面、进度要求与约束，并给出交付和验收的准则。

软件原型模块给出开发过程中所产生的界面各个版本的外观及描述。

除上述模块之外，在引言部分之前有本开发小组各成员的贡献，以及本文档的修改记录。

1.5 定义

【术语 1】用例：软件工程或系统工程中对系统如何反应外界请求的描述，是一种通过用户的使用场景来获取需求的技术

【术语 2】用例图：由参与者、用例、边界以及它们之间的关系构成的用于描述系统功能的视图

【术语 3】泳道图：泳道图，一种 UML 活动图，能够清晰体现出某个动作发生在哪个部门。泳道图在纵向上是部门职能，横向是岗位（有时候横向上不区分岗位）。绘图元素与传统流程图类似，但在业务流程主体上，通过泳道（纵向条）区分出执行主体，即部门和岗位来。

【术语 4】数据流图：简称 DFD，它从数据传递和加工角度，以图形方式来表达系统的逻辑功能、数据在系统内部的逻辑流向和逻辑变换过程，是结构化系统分析方法的主要表达工具及用于表示软件模型的一种图示方法。

【术语 5】（网络）爬虫：使用程序，按照一定的规则，自动地抓取万维网上的信息的手段

【术语 6】服务器：提供计算服务的设备。由于服务器需要响应服务请求，并进行处理，因此一般来说服务器应具备承担服务并且保障服务的能力

1.6 参考资料

- [1] Guide to Software Requirements Specifications, IEEE.
- [2] 王新萍. 如何编写软件需求说明书[J]. 山西煤炭管理干部学院学报, 2005(02):117-118.
- [3] 刘瑜. Python 编程从零基础到项目实战 中国水利水电出版社
- [4] 刘忆智. Linux 从入门到精通 清华大学出版社

2. 软件的一般性描述

2.1 软件产品与其环境之间的关系

本系统是使用 python 3 开发，支持 Windows 7 及以上（，Mac os）Windows 操作系统。开发工具是 pycharm。需要使用 Python scrapy, pandas, matplotlib, kivy 等库函数。

硬件接口：无特殊要求

软件接口：需要 Python 相关库函数的支持

通信接口：无特殊要求

2.2 限制与约束

2.2.1 硬件的限制：

目前为止软件仅限在 Windows7 及以上的平台运行

2.2.2 技术限制：

项目团队成员的技术水平不够成熟，需要在开发中并发学习多种技术和能力。若技术支持出现故障或疑难问题无法解决、程序开发出现偏差，会延误工程进度，影响工程的按期完工。爬虫系统受到网站反扒机制的限制，会出现不稳定和爬取时间较长的现象。

2.2.3 管理限制：

如果管理者没有协调好各个部分的关系，对出向的问题无法采用可行的解决手段，都会影响开发模块之间的互动，从而影响工程的顺利开展，导致工程无法按期完工。

2.2.4 人力、时间的约束：

软件由五位成员组成的小组开发；开发周期较短，约两个月的时间，需要开发成员积极合作配合，合理规划时间，按照预定要求完成任务。

2.2.5 性能的约束：

要求软件能完成所有预定的功能，软件运行流畅，除爬虫功能需要耗费一定的时间，不占用过多的系统资源。

2.3 假设与前提条件

本软件的使用用户为电影的相关从业人员与电影爱好者。该用户群体具有一定的学习和适应能力，并具有基础的电脑操作水平。能够看懂相关报表并具有报表的分析能力。

3. 功能需求描述

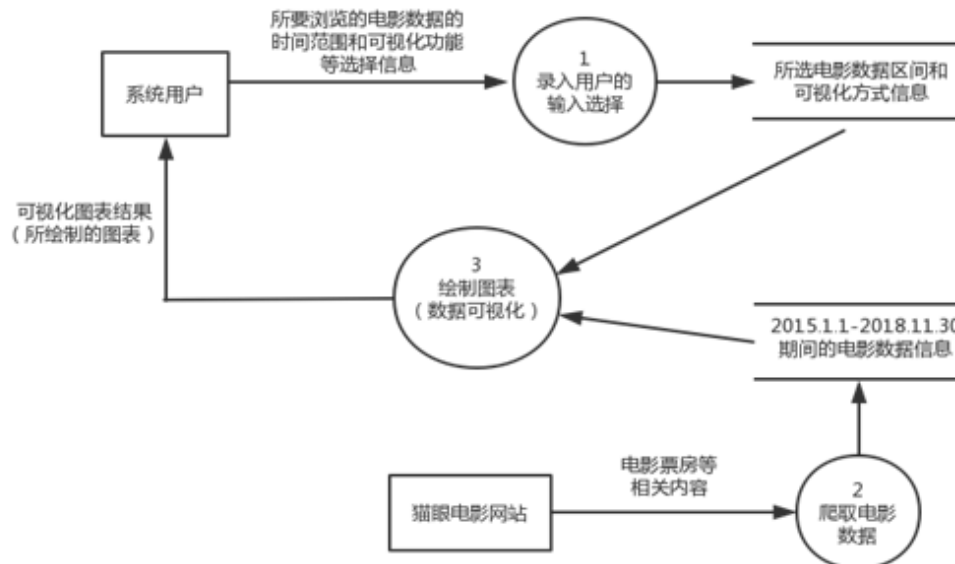
3.1 系统的划分

系统顶层数据流图：



顶层的 DFD

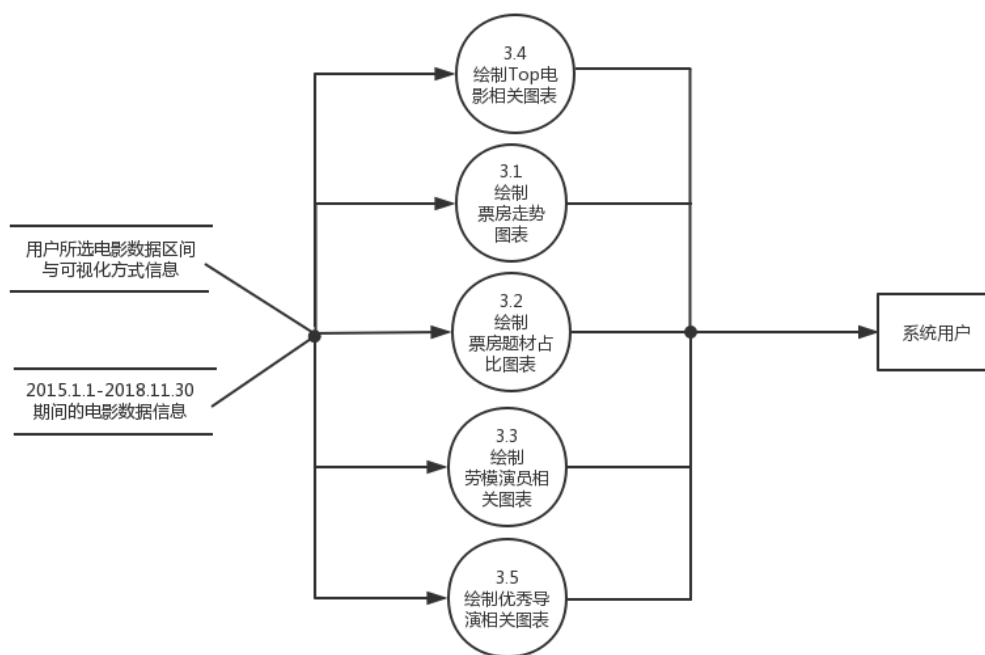
零层数据流图：



0 层 DFD 图

(新添加的) 0 层 DFD 图

一层数据流图：

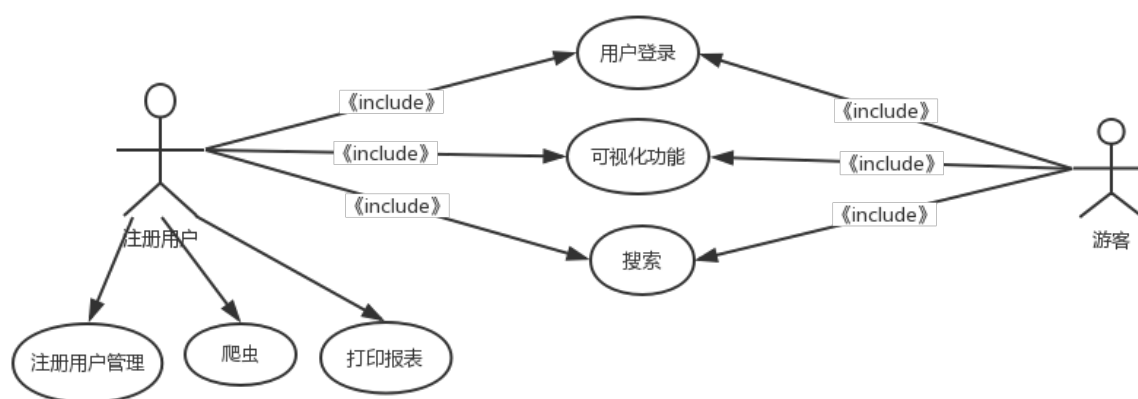


1 层 DFD 图

3.2 各子系统的功能

3.2.1 系统的用例图

更改后的用例图如下：

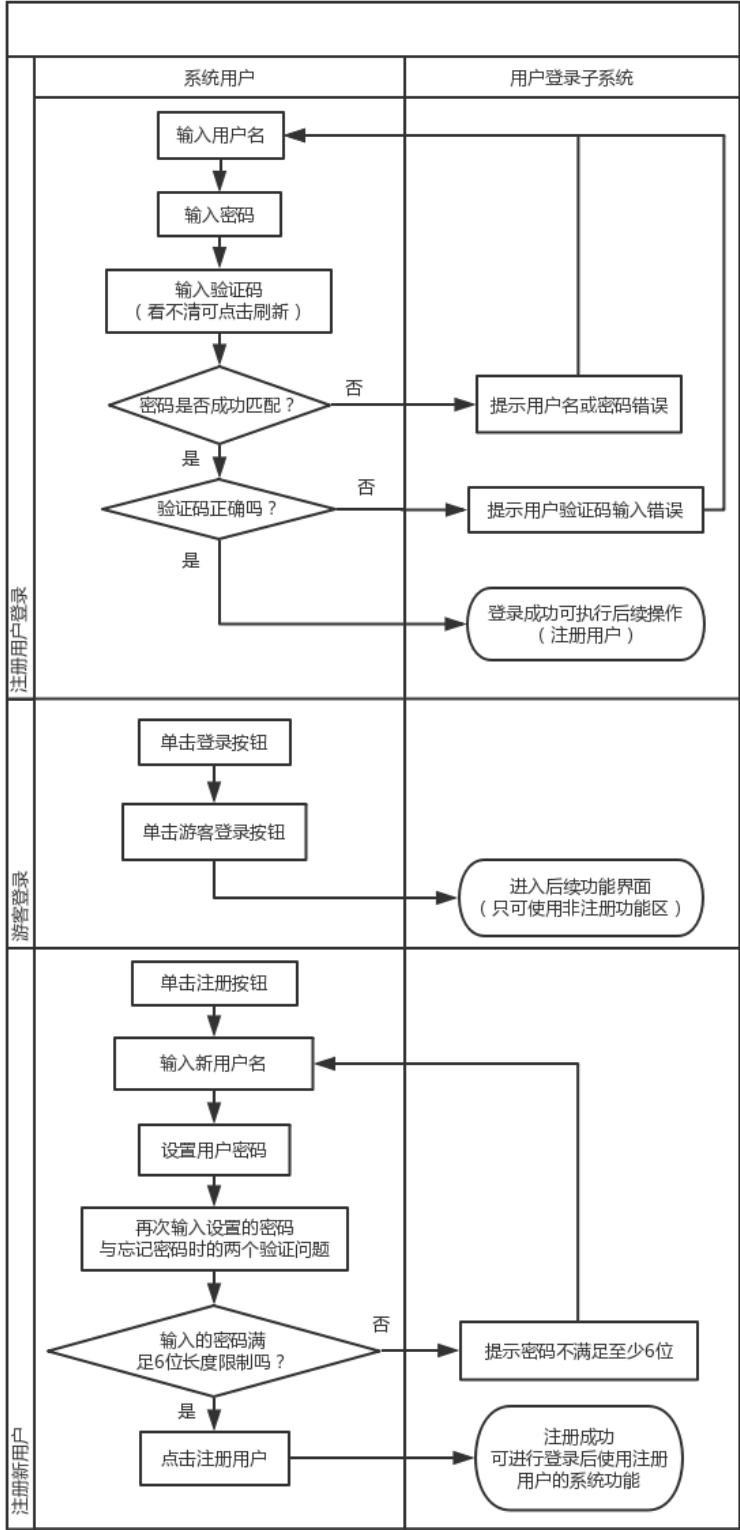


3.2.2 各子系统用例的描述

一、用例名称：用户登录与注册子系统

参与者与关注点：系统用户希望浏览近几年的电影数据及电影市场的统计数据，不必自行查找琐碎的信息。

1. 用户登录与注册子系统泳道图



2. 子用例描述

事件流：

主成功场景（常规流）：

- （1）注册用户进入电影可视化系统的登录界面，输入用户名与密码；
- （2）注册用户输入验证码，可选择密码回显或隐藏；
- （3）用户信息匹配后，进入并开始使用系统。

候选场景（扩展流）：

（1）用户未注册：在主流程中，若用户未注册，（则视为游客）可使用“游客”使用本软件，但此时用户仅能使用系统的公共功能区。

（2）输入的用户名或密码不正确：在主流程中，若注册用户输入的用户名或密码与存储的用户信息不匹配时，会提示出错的信息（弹出错误提示框），并要求用户再次输入。

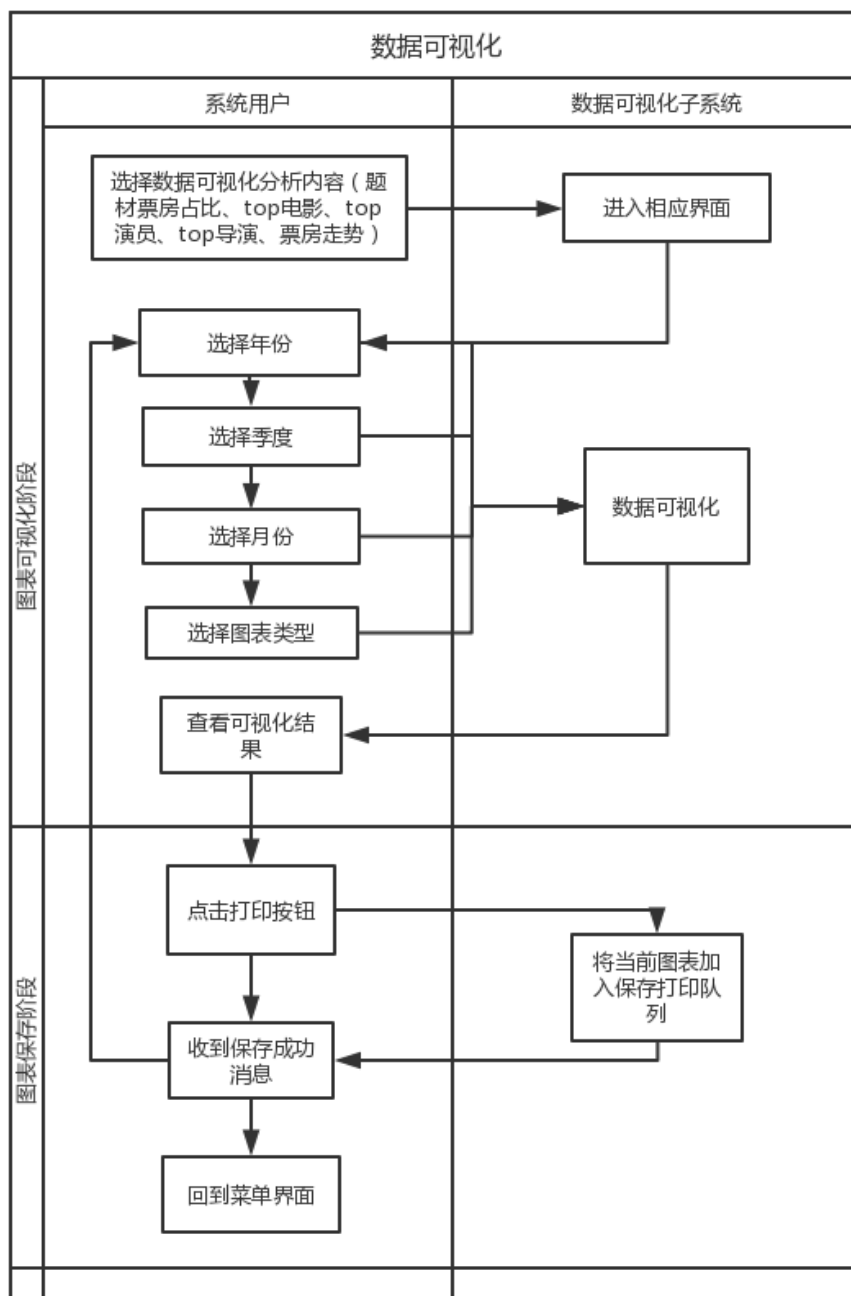
（3）输入的验证码不匹配：在主流程中，若用户输入的验证码不正确，会要求重新输入，若当前图片无法识别出来，可点击刷新按钮，对验证码进行更新。

前置条件：用户拥有 PC 端设备，并打开了该软件。

后置条件：成功登录进入系统，开始享用本系统提供的服务。

二、用例名称： 数据可视化

1. 数据可视化子系统泳道图



2. 参与者与关注点： 软件用户希望获得一定时间范围内、所选电影相关数据的可视化呈现。

3. 子用例描述

事件流：

主成功场景（常规流）：

- (1) 用户在功能界面的通用功能选择区域内，按下所希望实现功能的按钮
- (2) 进入相应功能的界面后，依次通过按钮选择限定数据的起始年份，终止年份，季节以及月份
- (3) 通过按钮选择希望的可视化呈现方式，包括折线图、条形图、饼图、词云呈现效果

(4) 点击 apply 按钮，触发绘图

(5) 用户期待得到的可视化结果呈现在屏幕指定的区域内

候选场景（扩展流）：

(1) 用户尚未选择时间范围，就选择可视化呈现方式、并点击了 apply 按钮：

系统弹出界面提示：“尚未限定时间范围”

(2) 用户选择的终止年份小于起始年份：系统弹出界面提示：“终止年份小于起始年份，请重新选择时间范围

(2) 系统自动会根据用户选取的起始年份，限定终止结束年份选择范围，不会出现终止年份小于起始年份的情况。

(3) 用户选择的月份不在相应的季节当中，比如说：季节选择了 autumn 但月份选择了 2 月，此时系统弹出界面提示：请重新选择时间范围。

(3) 用户选择了季节之后，系统会自动限定月份选择范围，例如选择了第三季度之后，月份的选择范围会自动限定为 7，8，9，全部，不会出现月份不属于季节的情况。

(4) 用户尚未选择可视化呈现方式，就点击了 apply 按钮进行绘图，系统弹出界面提示：请选择可视化呈现方式。

(5) 用户点击 apply 按钮时，若有时间范围尚未选定（为空），系统弹出界面提示：请选择时间范围。

前置条件

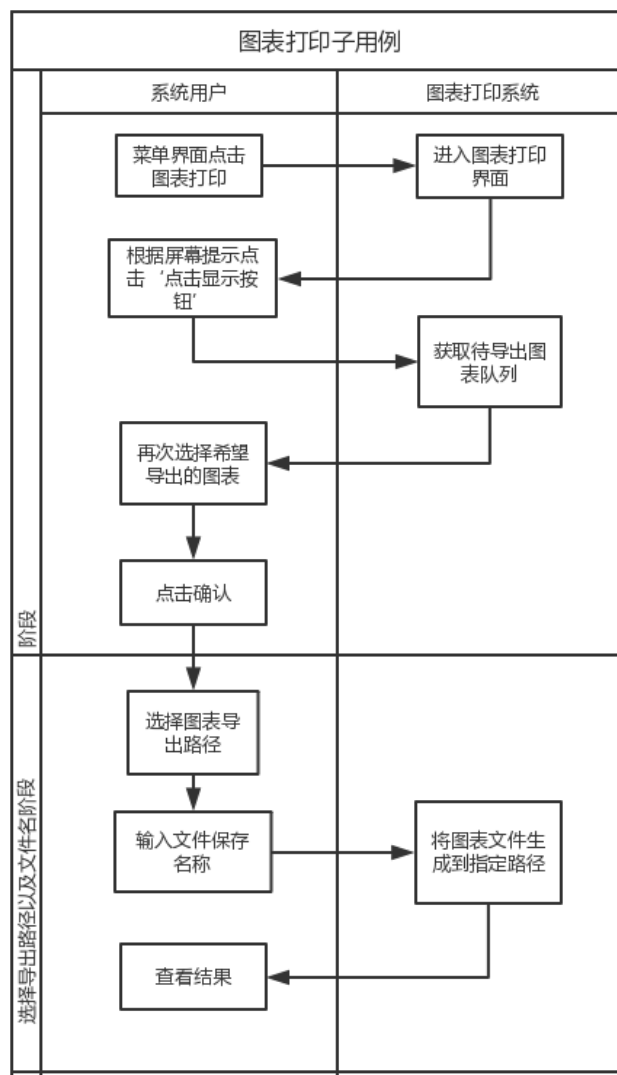
用户已经完成了登录流程，进入到了功能选择界面。

后置条件

屏幕显示用户选定的可视化结果，并加载出是否需要将该图片导出的选择框。

三、用例名称： 图表导出打印

1. 图表导出打印子用例泳道图：



2. 参与者关注点：系统注册用户希望将选定的可视化结果批量导出到指定路径的 pdf 文件当中。

3. 子用例描述

事件流：

主成功场景（常规流）：

- （1）用户完成登录之后选定导出图表的路径
- （2）每次得到一个可视化结果之后，可以选择点击屏幕上的图表导出按钮。
- （3）当用户退出系统的时候，存有选定可视化图表的 pdf 文件已经保存在了指定的路径当中。

- （1） 每次得到一个可视化结果之后，可以选择点击屏幕上的打印按钮。
- （2） 在主菜单上点击“打印”按钮，进入图表导出界面，在界面中会显示已经选择导出的图表名称列表。
- （3） 再次选择确认导出的图表，点击确认按钮。
- （4） 进入文件路径选择界面，可以选择图表导出的文件的路径。
- （5） 输入文件名后，点击“确认”按钮，图片成功以 pdf 形式导出到指定的路径中。

候选场景（扩展流）：

- （1）以游客身份进入软件的用户，点击图表导出按钮，系统弹出页面提示：“游客用户无权限”。

前置条件：

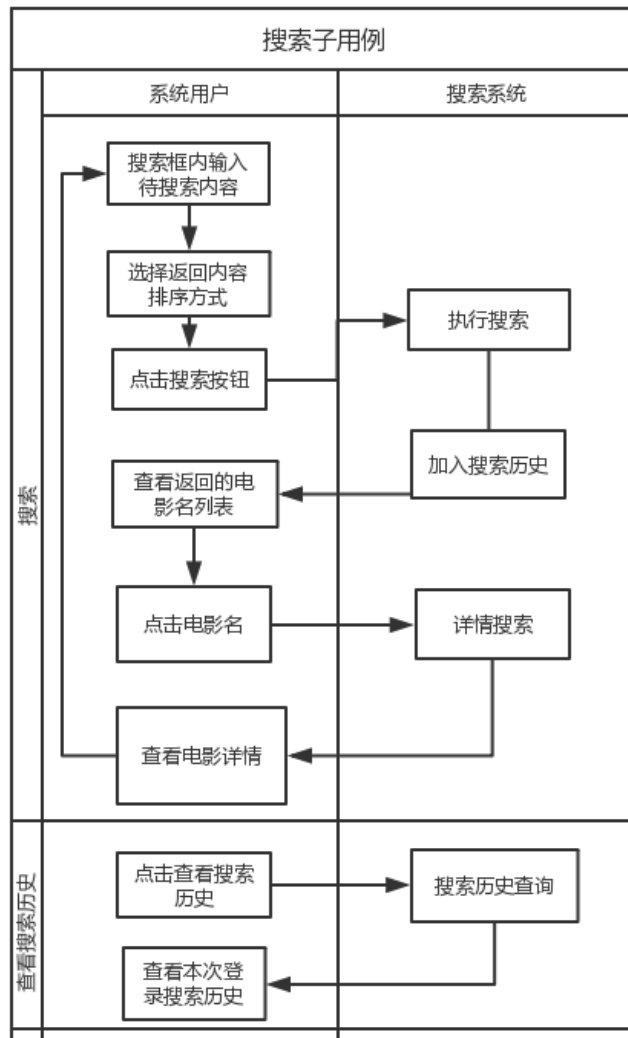
用户已经以注册用户身份完成了登录。

后置条件：

用户所选的可视化图表 pdf 文件已经导入到了指定路径当中

四、用例名称：搜索功能

1. 搜索子用例泳道图



2. 参与者与关注点：游客和注册用户希望查找关于电影的特定信息，不必逐一浏览

3. 子用例描述

事件流：

主成功场景（常规流）：

- （1）系统用户点击搜索框并输入想要查看信息的关键字；
- （2）输入完毕后，点击搜索按钮，系统开始搜索相关信息；
- （3）搜索完成后，系统返回包含该关键字的所有信息，呈现给用户。
- （4）用户可选择返回电影信息的排序方式（分为按票房降序和按评分降序两种）

(5) 用户点击显示搜索历史按钮，显示本次登录时间内的搜索内容。

候选场景（扩展流）：

(1) 输入的信息没有匹配项：在主流程中，若没有与用户键入的关键字相匹配的信息，系统给提示未找到指定项。

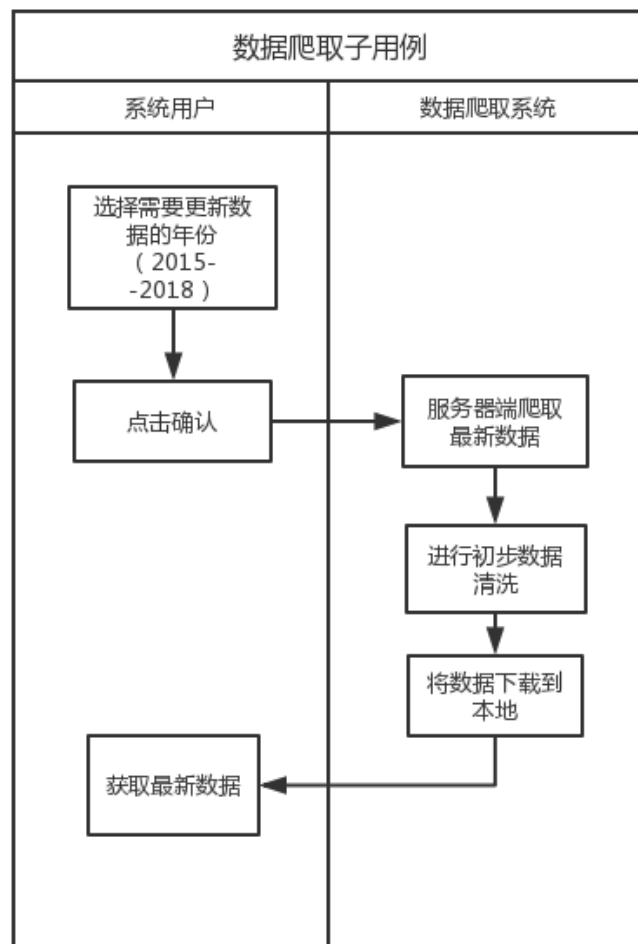
前置条件：用户已经打开该软件，并进入系统内部。

后置条件：系统完成搜索任务后，用户可查看其想要浏览的信息。

(2) 若选择的电影名系统不存有对应的电影海报，电影详情栏内的电影海报为空。

五、用例名称： 数据爬取

1. 数据爬取子用例泳道图



2. 参与者关注点： 系统可以更新用户所选时间段内的所有电影数据，并且数据爬取过程中，屏幕可以显示目前数据爬取的进度条。

3. 子用例描述

事件流：

主成功场景（常规事件流）：

- （1）系统注册用户的功能选择界面当中，点击（“数据爬取”）“爬虫”按钮
- （2）在数据爬取功能界面当中，选择需要更新数据的（开始以及终止）时间（以年为单位）
- （3）用户按下“确认”按钮
- （4）数据爬取开始，（屏幕加载进度条，显示数据爬取的进度），用户等待界面的提示
- （5）数据爬取完成，屏幕提示数据爬取完毕

候选场景（扩展流）：

- （1）用户尚未选择时间范围，就按下确认按钮开始数据爬取，系统弹出界面提示：“尚未选择时间范围”
- （2）用户选择的时间范围，终止年份小于起始年份，系统弹出界面提示：“请重新输入时间范围”
- （2）没有实现，换为选择某些年份，对指定的年份数据进行爬取
- （3）以游客身份进入系统的用户选择数据爬取功能，系统弹出界面提示：“游客用户无权限”

前置条件：

用户已经通过了用户登录界面，进入到了功能选择界面

后置条件：

系统注册用户所选时间范围内的数据完成了爬取，并且已经读入到了系统当中。

4. 其它需求描述

4.1 性能要求

4.1.1 响应速率

- (1) 当用户打开软件时，软件可以在 30 秒内启动并显示主界面。

当用户打开软件时，软件可以在 15 秒内启动并显示主界面。

- (2) 当用户点击软件的绘图及登录相关的按钮进行操作时，系统可以在 3 秒内返回结果。
- (3) 当注册用户点击爬虫按钮获取实时信息时，系统可以 5 分钟内返回爬取结果。
- (4) 当用户点击一般功能按钮时，系统能立刻做出响应

4.1.2 运行效率

- (1) 当用户点击软件的绘图及登录相关的按钮进行操作时，系统可以在 3 秒内返回结果。
- (2) 当注册用户点击爬虫按钮获取实时信息时，系统可以 5 分钟内返回爬取结果。
- (3) 当用户点击一般功能按钮时，系统能立刻做出响应

4.2 设计约束

- (1) 软件使用 scrapy 爬虫框架进行电影数据的爬取，使用 kv 进行 UI 界面开发。
- (2) （软件系统可以在任意操作系统中运行。）

软件系统只能在 Windows 7 及以上的版本中运行。并且在 Windows 10 系

统中使用搜索功能时，使用输入法输入中文，无法回显候选词，但在 Windows 7 上该搜索功能可以正常显示。

- (3) 用户要求软件可以保存打印的报表记录，并且用户的打印记录不会被其他用户看见。

（（3）功能暂未实现）

- (4) 用户要求软件返回的信息是实时可靠的。

(5) 安全性

- a. 通讯安全：保证数据和信息在传输过程中的基本安全。
- b. 认证安全：保证用户系统在认证上的安全性。
- c. 数据安全：保证密码等私密信息在服务端 / 数据库持久存储上的安全。

(6) 可靠性：

- a. 服务端：对于服务端短期的内存耗尽或者被系统杀死导致的应用崩溃一一类的可恢复错误，可被监控系统在较短且适当的时间内恢复。
- b. 客户端：在服务端因特殊情况而无法工作情况下，能够检测并主动告知用户，而非浪费用户宝贵的时间

4.3 界面要求

设计原则：

在感知和认知中，尽量提取自然元素加强产品体验，减少用户额外操作，让人机交互更加自然。

以下为软件原型，具体实现可能与图片不符。

登录界面如下所示：用户输入匹配的用户名和密码后，输入正确的验证码即可登录软件。

电影分析系统

登 录

用户名:

密 码:

验证码: 402F

数据分析界面如下所示：用户可选择要生成的图表的时间、题材内容和图表形式。用户选好后点击确认，系统便返回生成的图表。

用户在主界面中选择要分析的数据内容，点击按钮之后，软件会跳转到相应的界面。用户在界面中选择年份、图表形式等信息，系统便返回相应的图表。



爬虫界面如下所示：（用户点击按钮后，系统会对 2018 年的电影数据进行爬取，并弹出进度条，实时显示爬取进度）。

用户可选择要更新的年份，点击确认后，系统便更新相应年份的信息。



报表打印功能如下所示：界面内会显示用户的历史打印记录。（没有历史打印记录）

用户需要点击图表显示界面的打印按钮，打印报表界面会显示用户想要打印的报表。用户选择要打印的图表之后，系统会将图表以 pdf 形式打印存放在电脑中。



电影查询界面如下所示：用户按下搜索键后，系统将显示搜索结果。



查询历史界面如下所示：系统会显示用户查询过的内容以及用户查询该内容的时间。(没有查询内容的时间)



用户管理界面如下：用户可在该界面下修改密码。(用户需要答对至少一个验证问题才能修改密码)



4.4 进度要求

- (1) 2018 年 12 月前，完成软件登录界面的初步设计。
- (2) 2018 年 12 月 9 日，完成软件系统的爬虫功能，实现系统界面的基础框架，实现对数据进行初步的处理。
- (3) 2018 年 12 月 20 日，完成初版软件系统，要求该软件系统能够实现基本功能。
- (4) 2018 年 12 月 25 日，交付软件系统，要求该软件系统功能完善，且能够处理错误的输入和一些异常情况。

4.5 交付要求

用户希望软件系统交付时，交付可运行使用的软件包以及软件报告。

其中软件报告中需要包括软件的（使用手册）需求规格说明书、异常情况处理方法，以及相关的人员的联系方式，以便遇到突发情况时，可以询问解决。

4.6 验收要求

项目验收时，软件应实现以下功能：

- (1) 有完整的登录界面。注册用户可通过登录界面登录自己的账号，而未注册用户可以跳过登录手续，但在使用软件时会受到部分限制。
- (2) 有美观实用的界面。用户根据界面提示，点击相应的按钮，软件便会返回相应的结果。要求软件返回的结果正确，且响应时间必须不超过 3 秒。
- (3) 注册用户和非注册用户之间有区分。注册用户可以使用软件的全部功能，而非注册用户不能使用软件实时更新数据、打印报表和账户管理的功能。
- (4) 爬虫爬取功能正常。用户按下爬虫按钮后，要求软件能够在 5 分钟内返回实时爬取的结果，并且要求爬虫爬取的数据正确，且可以在软件中进行加工。
- (5) 能生成报表。用户所感兴趣的图表可以打印在同一张报表中，以 .pdf 的形式导出到电脑中。

5. 软件原型

所有的界面中的箭头图标表示“返回”，带小三角形的图标点击后有下拉框，可以选择不同的选项。

(1) 用户登录界面：



上图为用户界面 1.0，实现基本的登录界面功能，但密码部分为明文显示，且验证码无法刷新，且成功时显示的颜色为红色，且点击按钮后不会清空已经输入的内容。



上图为用户界面 2.0，在 1.0 的基础上完善了界面，密码默认为“*”回显，可根据按钮“回显/隐藏”改变其为明文回显，按下“刷新”得到新的验证码。注册成功时后提示信息（位置同 1.0 版）为绿色，否则为红色。

(2) 分析数据界面

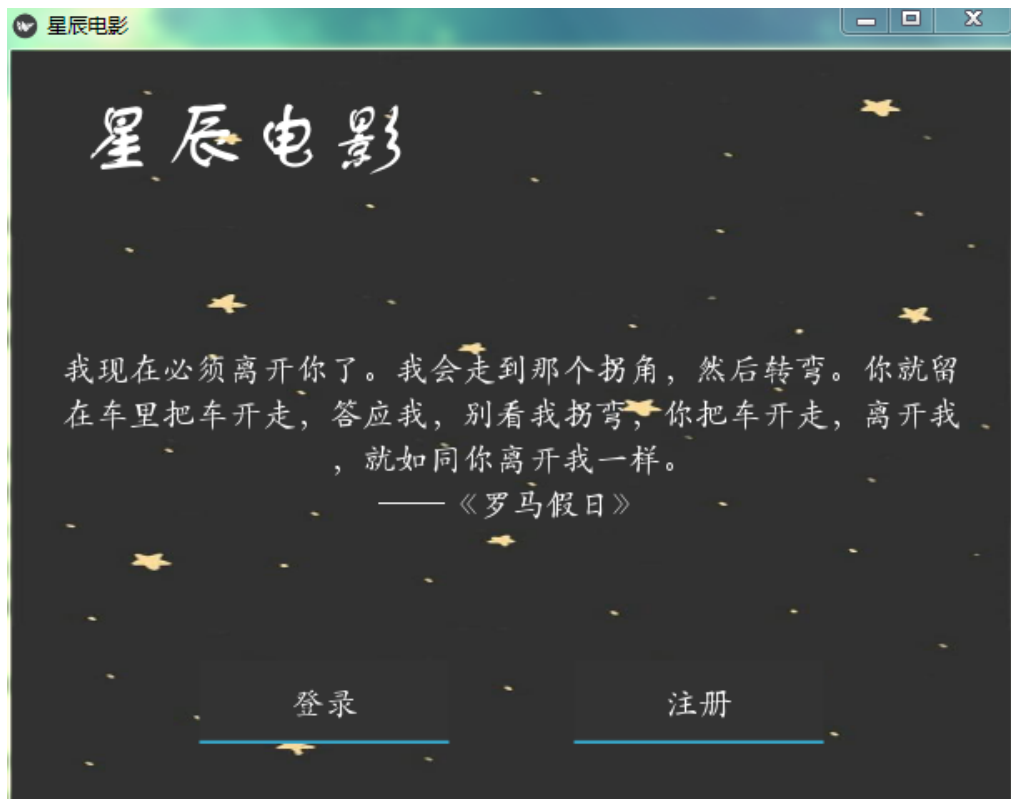
计划完成的界面：数据可视化界面，搜索界面，爬虫界面，用户管理界面。)

(1) 初始界面



初始界面显示本软件的名称以及一句随机的电影中的经典台词。包括登录和注册两个按钮。

修改后的初始界面：



(2) 登录界面



登录界面中的身份有“注册用户”和“游客”两种，注册用户可以使用全部的功能，而游客的功能受限。

修改后的登录界面：



(3) 功能目录



通过改界面可以进入各种功能中。

新版界面引入音乐开/关功能，背景音乐每次启动软件的时候随机选择音乐库中的一首，单曲循环，无法调节音量。

修改后的功能目录：



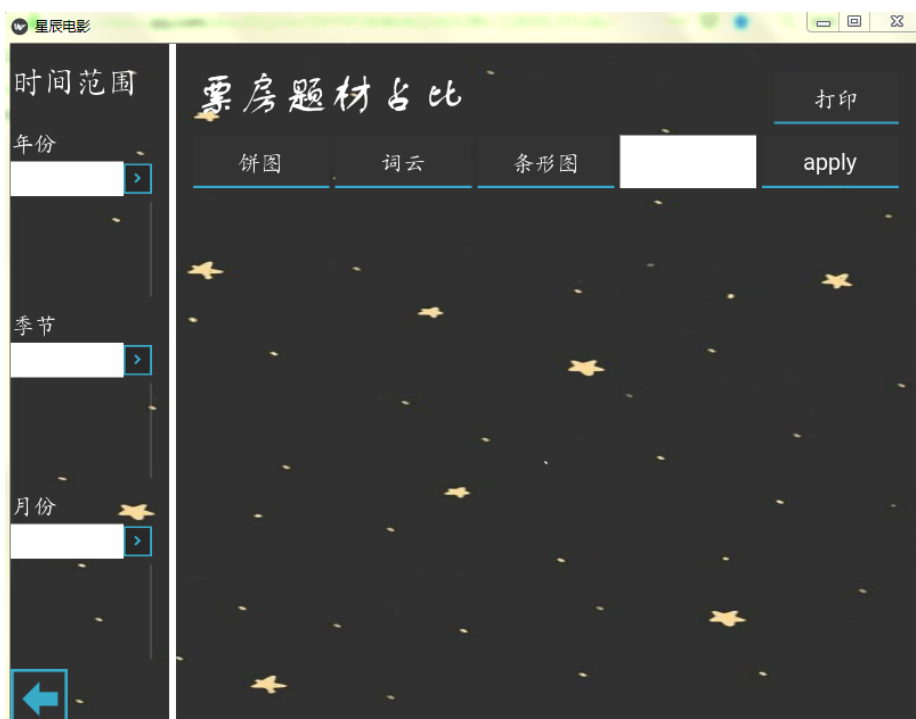
(4) 可视化功能



用户选择某项可视化功能后可从左侧选择时间、年份等信息，然后从右上侧选择所绘制的图片类型，然后点击 apply 执行绘图。点击“打印”按钮会将绘制出来的图片加入打印列表中。仅有注册用户可以查看打印列表。

仅展示其中一个界面的图示。

修改改后的可视化功能界面：



(5) 搜索界面



搜索界面可以保留用户本次使用（退出软件或者注销）前的搜索记录。

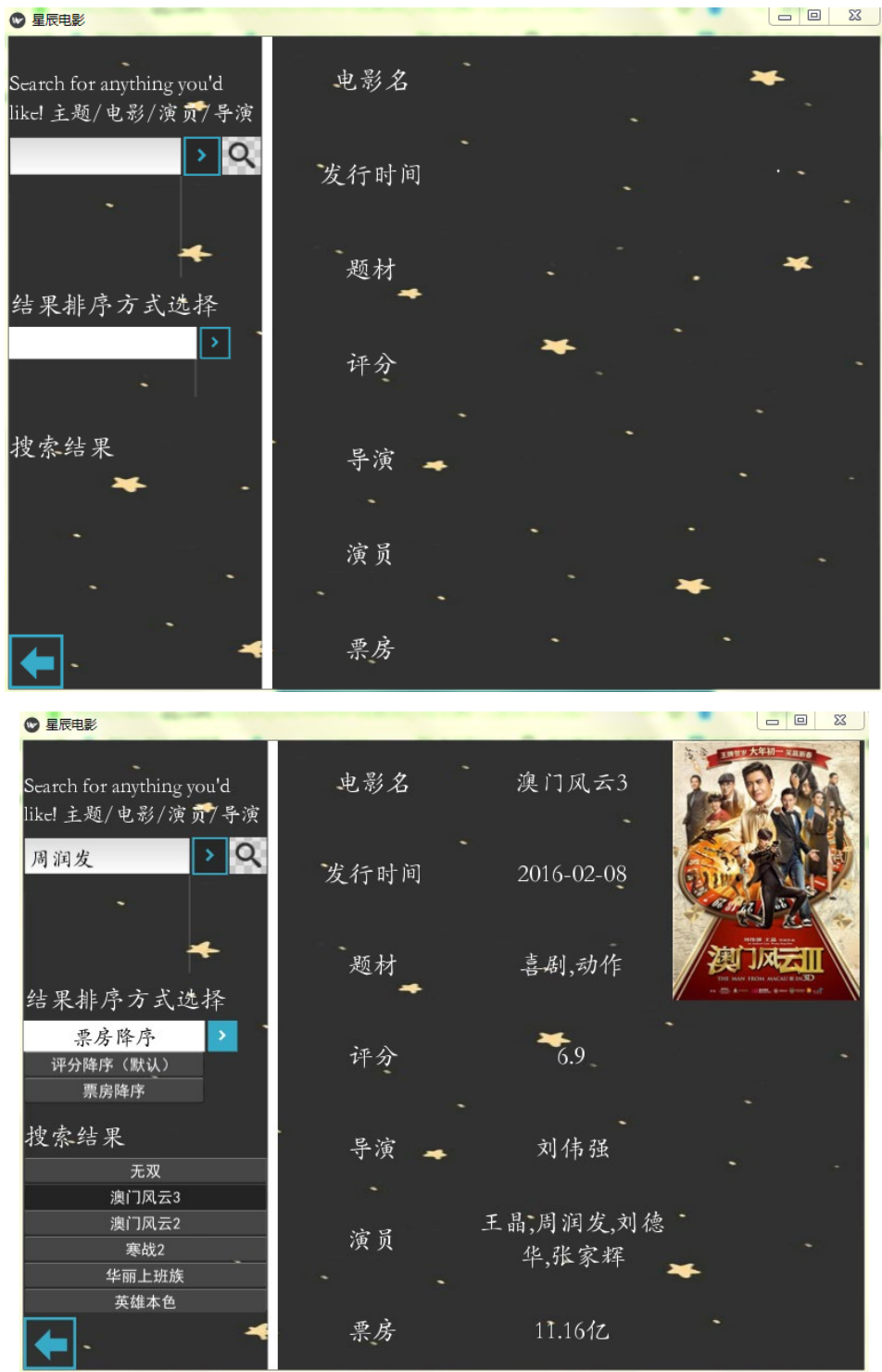
Windows10 系统下，搜索框不能中文显示候选词，但可以使用外部的剪切板（复制粘贴）完成搜索。

仅当按下放大镜的时候才会执行搜索。

放大镜前的按钮点击后可以看到最近的搜索历史。

对于部分影片右上角可能没有图片显示，是由于其海报数据缺失引起的。

修改后的搜索界面：



(6) 打印



注册用户选择打印报表时，会先进入以上页面，被选中的选项会加深颜色，作为确认要导出的图表，然后点击“确认”选择路径。



默认路径为用户桌面，单击进入目录。

修改后的打印报表界面：





(7) 爬虫



点击不同年份获取不同年份的最新数据。

修改后的爬虫界面：



修改后，单击爬虫图片即可进行数据爬取。

(8) 注册用户管理

星辰电影

用户名: 刷新 按下"刷新"获取用户名

以下两个验证问题(Q1 Q2) 仅需回答正确一个即可

Q1 你的生日

Q2 你的名字

新密码:

再次输入:

← 确认修改密码

在本界面上可以修改注册用户的密码，使用验证问题进行身份验证，用户可依据这两个验证问题修改自己的密码。

修改后的注册用户管理界面：

星辰电影

用户名: 刷新 点击"刷新"显示用户名

以下两个验证问题(Q1 Q2) 仅需回答正确一个即可

Q1 你的生日

Q2 你的名字

新密码:

再次输入:

← 确认修改密码