

# 项目使用说明书

Arshtyi

2024 年 10 月 31 日

## 摘要

本文为本人开发的山东大学网页通知爬虫和山东大学青岛校区寝室剩余电费查询脚本使用说明书.

项目仓库地址为 <https://github.com/Arshtyi/T4>

目录

<b>1</b>	<b>项目简介</b>	<b>1</b>
1.1	地址及依赖 . . . . .	1
1.2	简述 . . . . .	1
1.2.1	网页通知爬虫 . . . . .	1
1.2.2	电费查询 . . . . .	1
1.3	下载与使用 . . . . .	1
<b>2</b>	<b>网页通知爬虫</b>	<b>2</b>
2.1	简述 . . . . .	2
2.1.1	依赖 . . . . .	2
2.1.2	爬取内容 . . . . .	2
2.2	操作 . . . . .	2
2.2.1	操作选项 . . . . .	2
2.2.2	爬取操作 . . . . .	3
2.2.3	功能操作 . . . . .	3
2.3	输入与输出 . . . . .	3
2.3.1	输入 . . . . .	3
2.3.2	输出 . . . . .	3
2.4	定时自动爬取功能 . . . . .	3
2.4.1	爬取时间 . . . . .	3
2.4.2	操作 . . . . .	3
<b>3</b>	<b>电费查询</b>	<b>4</b>
3.1	简述 . . . . .	4
3.1.1	依赖 . . . . .	4
3.1.2	功能 . . . . .	4
3.2	操作 . . . . .	4
3.2.1	查询 . . . . .	4
3.2.2	定时查询与邮件提醒 . . . . .	4
3.3	其他 . . . . .	4

## 1 项目简介

### 1.1 地址及依赖

项目仓库地址为 <https://github.com/Arshtyi/T4>

依赖于 Python12.3 及以上版本. 各依赖项见.py 文件.

每个项目的程序都被切割为 main.py 文件和一个对应文件, 请将两个文件下载到同一目录下再运行 main.py 文件.

### 1.2 简述

#### 1.2.1 网页通知爬虫

是否想要打好校园信息战? 是否想要第一时间获得学校的通知内容? 是否想要了解更多的校园通知? 来试试本项目吧.

#### 1.2.2 电费查询

众所周知, 山东大学青岛校区的电费系统非常抽象, 它的充值接口只开放到十点, 于是经常有忘记充电费导致半夜断电的情况, 该项目正是为了方便查询寝室剩余电量. 不需要使用山大 V 卡通查询了. 同时还有定时查询、邮件提醒的功能.

### 1.3 下载与使用

到项目地址 <https://github.com/Arshtyi/T4> 下载你所需的脚本.

注意, 每个项目的程序都由同一目录下的一个 main.py 文件和一个对应的功能文件组成, 二者缺一不可.

下载需要的.py 文件后运行 main.py 文件并按照给出的提示操作即可.

## 2 网页通知爬虫

### 2.1 简述

爬取几个山东大学相关网站发布的通知，让你获得最新的校园新闻，不必再频繁刷新查看网页通知。爬取内容将输出在当前目录下便于查看。同时能够实现灵活的定时爬取功能。

#### 2.1.1 依赖

依赖项详见各.py 文件，请确保正常连接。

#### 2.1.2 爬取内容

该脚本暂时能够爬取以下三个网站发布的通知：

- (1) 山大视点-山大要闻
- (2) 本科生院-工作通知
- (3) 计算机学院-本科教育

获取内容包括网站通知的来源，链接，发布时间，标题，简述。这些信息将存储于当前目录下的 ret.xlsx 文件内。

### 2.2 操作

#### 2.2.1 操作选项

运行 main.py 文件后，将给出对应一定的提示信息供你选择：

- (1) 山大视点-山大要闻
- (2) 本科生院-工作通知
- (3) 计算机学院-本科教育
- (4) 在已定时间自动进行三个网站内容的爬取
- (5) 立刻停止，同时格式化输出表格
- (6) 删除原有表格，新建输出表格

### 2.2.2 爬取操作

选择 1 – 3 后将提示选择期望的爬取页数，之后立即对指定网站进行爬取. 结果输出为当前目录下的 ret.xlsx 文件.

### 2.2.3 功能操作

- 4 操作使程序进入定时自动爬取.
- 5 操作使程序对输出的表格做格式处理后立刻停止.
- 6 操作将会删除原有表格，新建一个输出表格.

## 2.3 输入与输出

### 2.3.1 输入

输入只需要依据提示选择需要的功能即可，特别地，定时自动爬取中 Ctrl+C 会让程序在下一次爬取时终止.

### 2.3.2 输出

每次启动程序都会重置 ret.xlsx 文件，默认完成爬取后不会进行格式处理，只有进行 5 操作停止后才会对表格做格式处理.

输出的 ret.xlsx 文件在当前目录下，推荐使用 Excel 查看.

## 2.4 定时自动爬取功能

### 2.4.1 爬取时间

自动爬取时间默认为 7、12、18、22 四个时间点.

### 2.4.2 操作

选择 4 后，将进入自动爬取状态，到达已定时间点后，程序将进行爬取，默认每个网站爬取一页通知.

此时按下 Ctrl+C 将在下一次爬取时退出.

## 3 电费查询

### 3.1 简述

该脚本能够查询山东大学青岛校区的宿舍剩余电量情况. 解放山大卡通. 同时能够实现定时查询、邮件提醒功能.

#### 3.1.1 依赖

依赖项详见.py 文件, 请确保正常连接以及输入正确.

#### 3.1.2 功能

该脚本将会接收校园卡号、宿舍楼、房间号后输出寝室剩余电量信息. 结束后输入一定信息设置定时自动查询功能.

### 3.2 操作

#### 3.2.1 查询

运行 main.py 文件后, 输入校园卡号、宿舍楼名 (如 S2, 字母需要大写)、房间号 (如 b114, 字母大小写均可), 将会输出房间电量剩余情况, 同时一个包含各宿舍楼 buildingid 的表格 Buildings.xlsx 将保存在当前目录下, 推荐使用 Excel 打开查看.

#### 3.2.2 定时查询与邮件提醒

查询结束后, 按要求输入即可进入定时查询模式, 请确保自己输入正整数的小时数以完成任务. 同时需要你的收件邮箱地址. 期间 Ctrl+C 即可退出.

邮件的发送者默认为作者的个人 QQ 邮箱, 可对 sender\_qq 作修改以改变 (这样做时也记得修改 password).

邮件发送成功将会有个提示.

### 3.3 其他

该程序不会如上面的爬取通知一样对输出的表格做格式处理, 也不存在主动的停止选项 (定时任务中可以 Ctrl+C 退出)