**天软众创空间模拟用户签到程序**

目录

[一、 开发的思路 3](#_Toc536610755)

[1.1 提出要求 3](#_Toc536610756)

[1.2 配置开发的环境 3](#_Toc536610757)

[1.3 代码编写 4](#_Toc536610758)

[1.3.1设置访问请求 4](#_Toc536610759)

[1.3.2 模拟用户 4](#_Toc536610760)

[1.3.3导入用户 4](#_Toc536610761)

[1.4 windows上的运行 4](#_Toc536610762)

[二、 服务器上的部署 6](#_Toc536610763)

[2.1 服务器的搭建 6](#_Toc536610764)

[2.2远程控制 6](#_Toc536610765)

[2.3在Linux上运行程序 6](#_Toc536610766)

[2.3.1 配置运行环境 6](#_Toc536610767)

[2.3.2 运行程序 6](#_Toc536610768)

# 开发的思路

## 提出要求

一开始的要求是,制作一个，模拟用户登录的请求，并访问签到页面模拟签到的小爬虫。（需求简单，但是由于本人没有经验，在制作过程中遇到了很多问题。后一一说明。）

## 配置开发的环境

这个程序一开始打算用python+scrapy来写，但是scrapy一直配置不上，所以就单用了python。

Python：

在选择python版本的时候，有两个版本可以选择python2和python3。两个版本不互相兼容，根据个人喜好，我就选择了python3.7。安装包从https://www.python.org/上下载即可，安装过程比较简单。最后利用cmd检测一下版本如图1.1

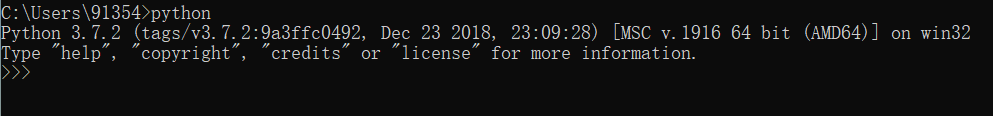


图 1.1

编译器：

VScode是我一开始选用的的IDE，但在使用过程中发现了一个问题，我在程序中引用了python的request和cookiejar的库。VScode再找这些库的时候有些麻烦。

又找了找发现Pycharm这款IDE主要用于python的编译，而且对于上述的以及其他的库支持简易安装。配置成功后，开发环境就OK了。

## 代码编写

### 1.3.1设置访问请求

根据网页headers，摘取 request URL，如图1.2。并设置useragent，制作request。根据response的返回值判断是否成功。



图 1.2

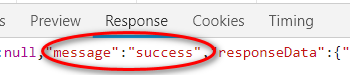


图 1.3

### 1.3.2 模拟用户

需要设置休眠时间，因为人工的访问速度与机器差异很大。



图 1.4

### 1.3.3导入用户

在user.txt中可随时更改用户名和密码。

## windows上的运行

利用cmd命令，并把python配置到环境变量中，进入所在文件夹后，直接用python XXX.py 即可执行程序。如图 1.2所示

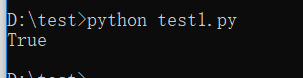


图 1.4

# 服务器上的部署

## 2.1 服务器的搭建

由于最终是要在公司的服务器上运行。所以在自己的PC上先试验一下。

首先需要在自己的PC上建一个虚拟机：

VMware 下利用Ubuntu server的镜像建立虚拟机（需开启BIOS中的Intel Virtualization Technology）

## 2.2远程控制

虚拟机搭建成功后，进行远程控制，因为是Linux系统，所以需要下载putty。还有winscp进行文件的远程传输。

## 2.3在Linux上运行程序

### 2.3.1 配置运行环境

Linux上自带了python2和python3（以我的为例，是3.5的版本），我编译的环境是3.7所以还是需要更新一下。

Python更新：

首先需要安装gcc，如果没有gcc，一些依赖c的程序就不能运行

安装zlib，再安装python3.7。

安装完成后也检测下版本，如图2.1

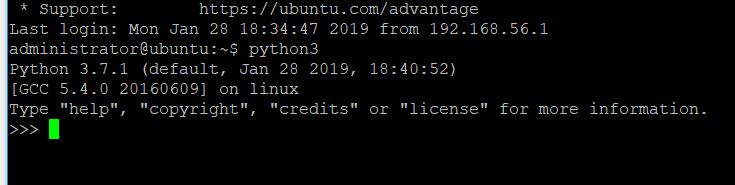


图2.1

### 2.3.2 运行程序

确定文件位置（注意文件内的引用要用绝对路径，否则很容易出错）

输入命令python3 /路径/XXX.py即可运行程序，如图2.2

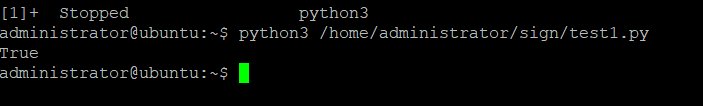


图2.2