

3月10号:

完成了自动登陆微博, 抓取数据。

步骤:

1、新浪登陆的步骤采用的多次ajax访问服务器获取数据, 对用户名密码进行加密

2、首先访问[https://login.sina.com.cn/sso/prelogin.php?](https://login.sina.com.cn/sso/prelogin.php?entry=openapi&callback=sinaSSOController.preloginCallBack&su=uname&rsakt=mod&checkpin=1&client=ssologin.js(v1.4.15))

[entry=openapi&callback=sinaSSOController.preloginCallBack&su=uname&rsakt=mod&checkpin=1&client=ssologin.js\(v1.4.15\)](https://login.sina.com.cn/sso/prelogin.php?entry=openapi&callback=sinaSSOController.preloginCallBack&su=uname&rsakt=mod&checkpin=1&client=ssologin.js(v1.4.15)), su=后面接自己的用户名, 会返回数据:

```
sinaSSOController.preloginCallBack({"retcode":0,"servertime":1456313380,"pcid":"gz-79a8f7e02d9275e24e3485f463862010236b","nonce":"QNH5YS","pubkey":"EB2A38568661887FA180BDDDB5CABD5F21C7BFD59C090CB2D245A87AC253062882729293E5506350508E7F9AA3BB77F4333231490F915F6D63C55FE2F08A49B353F444AD3993CACC02DB784ABBB8E42A9B1BBFFFB38BE18D78E87A0E41B9B8F73A928EE0CCEE1F6739884B9777E4FE9E88A1BBE495927AC4A799B3181D6442443","rsakv":"1330428213","exectime":7})
```

其中有用的数据为servertime, pubkey, rsakv, nonce

3、对用户名进行base64编码

```
su = base64.b64encode(bytes((urllib2.quote(uname).encode('utf-8'))))
```

4、对公钥进行转换16位int, 并做rsa公钥运算

```
rsaPublickey = int(pubkey, 16)
```

```
key = rsa.PublicKey(rsaPublickey, 65537)
```

5、将servertime, nonce和自己的密码联合起来, 与公钥做rsa加密, 转换成ascii码

```
message = bytes((str(servertime) + '\t' + str(nonce) + '\n' + str(pwd)).encode('utf-8'))
```

```
sp = binascii.b2a_hex(rsa.encrypt(message, key))
```

6、根据手动登陆时抓取的包, 获取的头信息, 写一个param, 装载入相应的信息, 例如加密后的密码, servertime等, 将数据提交到

[http://login.sina.com.cn/sso/login.php?client=ssologin.js\(v1.4.15\)](http://login.sina.com.cn/sso/login.php?client=ssologin.js(v1.4.15))

通过urllib2.build_opener建立含cookie的请求对象

完成登陆, 使用cookielib的LWPCookieJar格式的cookie, 通过urllib2的

HTTPCookieProcessor记录下登陆后的cookie, 使用install_opener绑定好

cookie, 然后使用该对象访问其他网页时, 自动登陆。

组员: 丁涵宇 李旦荣 瞿靖坤 单伶俐