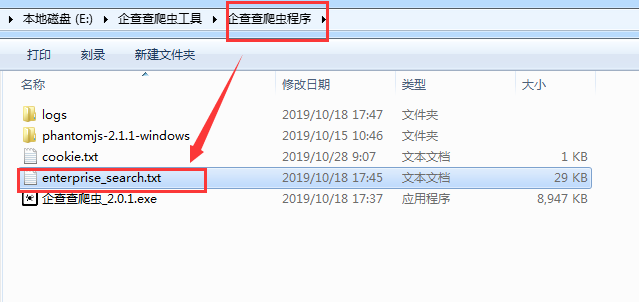
**企查查爬虫使用手册**

本爬虫软件为一个半自动化的数据采集软件，需要用户模拟登录“企查查”网站后，复制到cookie值才能进行程序后续的自动化采集工作。而复制cookie时的界面、快捷键由于浏览器不同会有所差异，下面截图中的浏览器是Chrome（谷歌浏览器），也推荐使用Chrome。步骤如下：

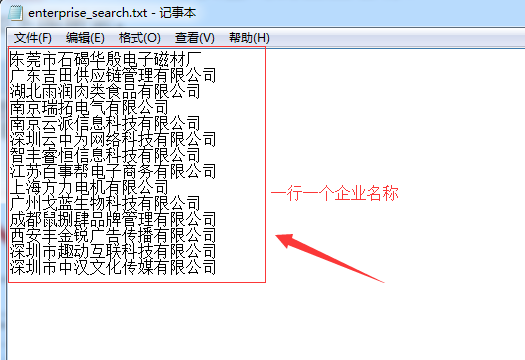
#### **添加企业名称**

找到爬虫程序文件中的“enterprise\_search.txt”，如下图所示：



记事本打开enterprise\_search.txt文件，加入你要查询的企业名称，

注意每个名称必须是一行，如下图所示：



#### **获取cookie**

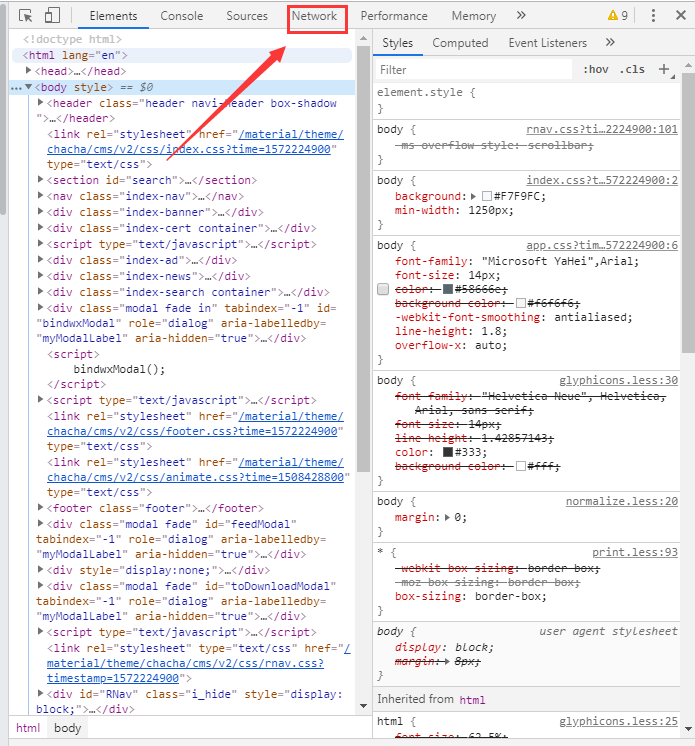
手动进入企查查网站www.qichacha.com，用自己的账号登录系统。如下图：



这里用的是谷歌浏览器，这时按“F12”打开浏览器页面管理界面，如下图（红色边框区域就是管理界面）



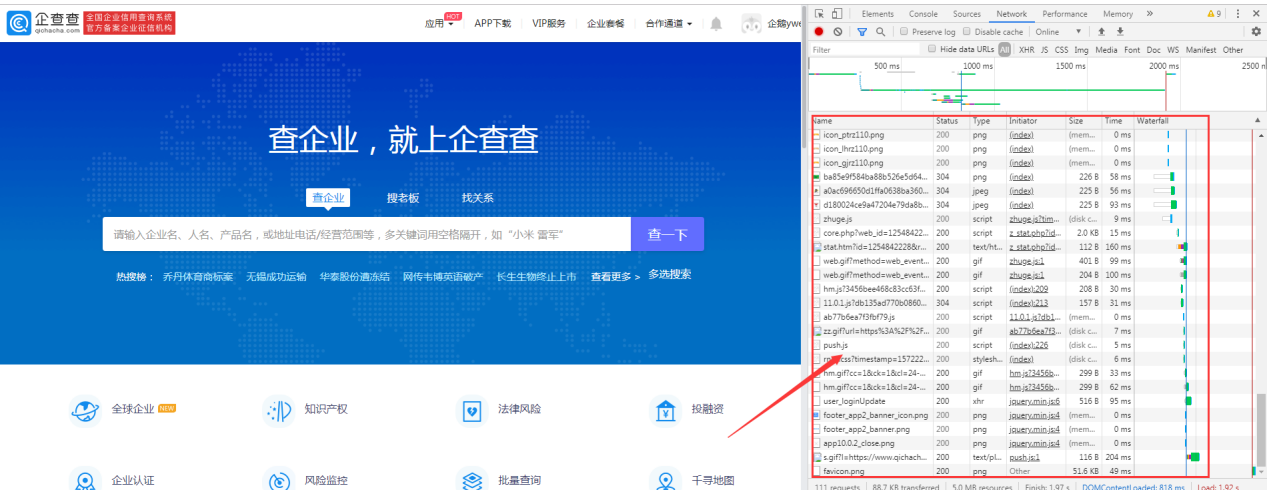
点击右上方按钮条中的“Network”如下图：



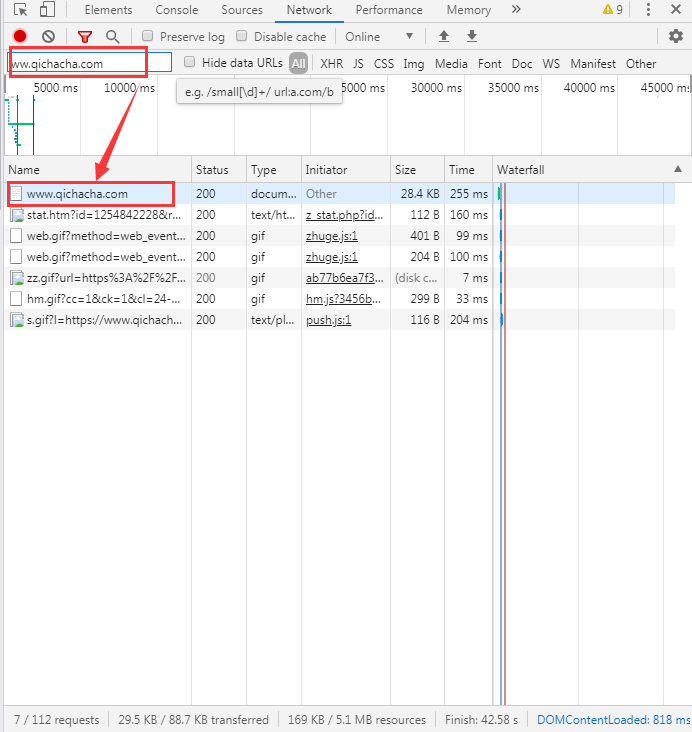
然后如下图，注意保证红色圆圈内的“All”是灰色（选中）的



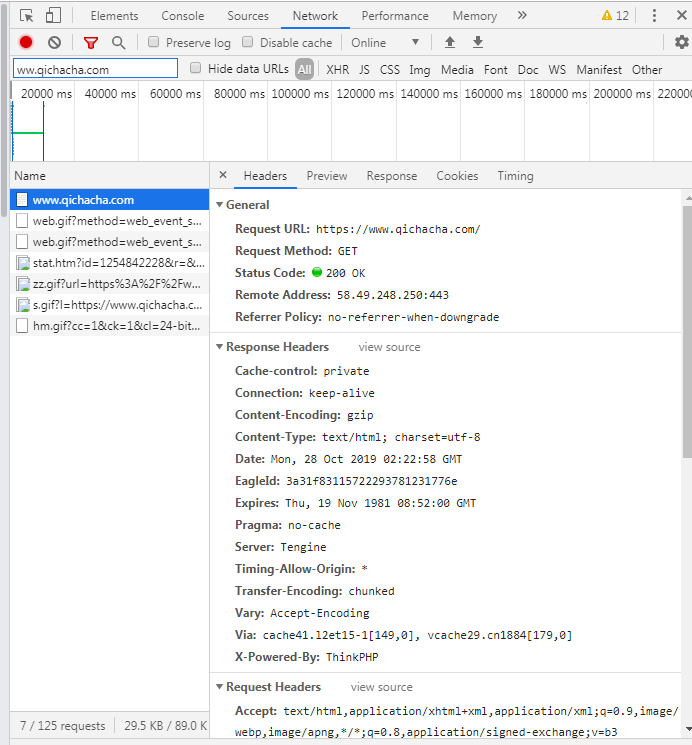
这时刷新（快捷键F5或点击刷新）页面，下方会出现很多条数据，如下图：



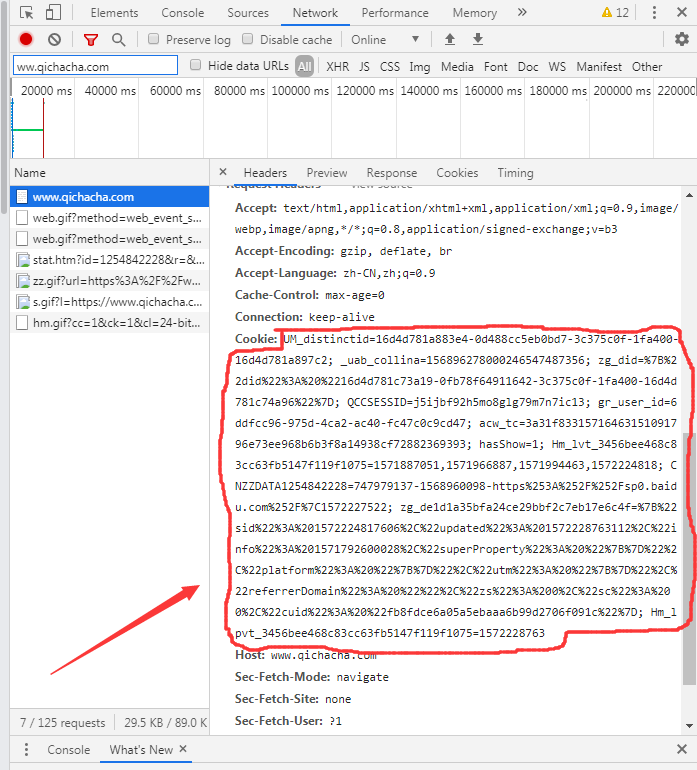
这时我们要找到“www.qichacha.com”这条数据，如果没有找到，可以通过上图右侧滑块向上滑动，一般都是放在第一个，也可以通过上面的“Filter”框搜索。如下图：



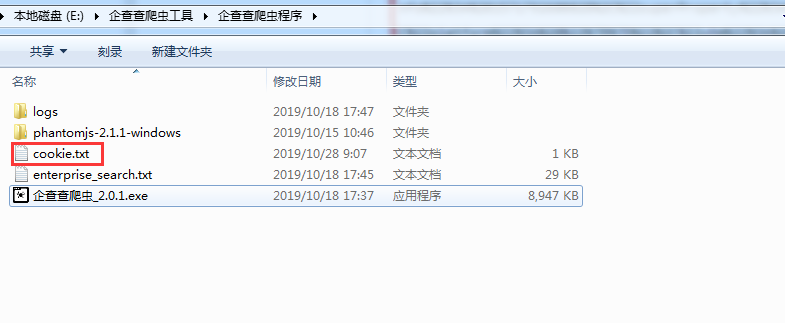
点开这个“www.qichacha.com”以后就看到如下图：



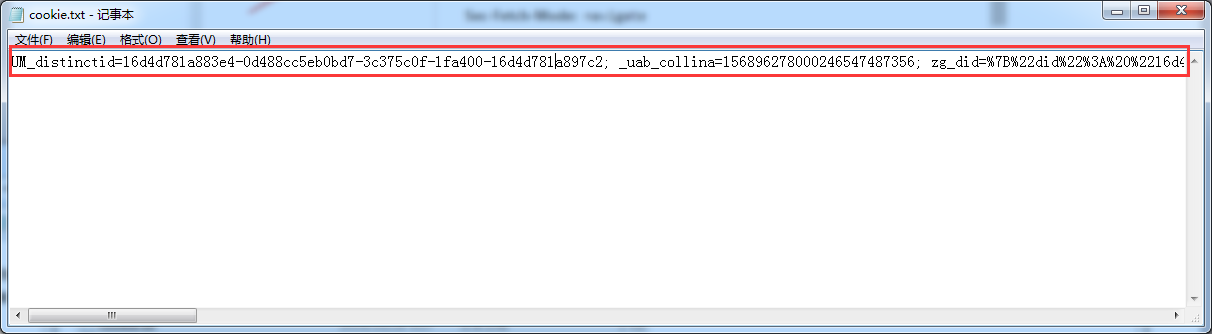
然后向下滑动上图右侧的滑块，找到Cookie:开头的这项，把它后面的数据复制，如下图红色框中部分：



打开程序文件中的cookie.txt，把红色框中的内容复制到文件中，并且保存。如下图：

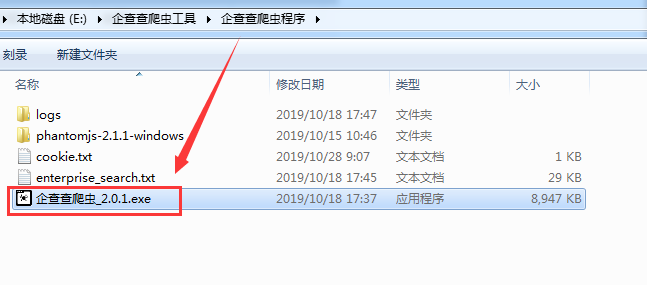


注意：一个cookie只能放一行，不允许换行；如果多个cookie可以换行保存。



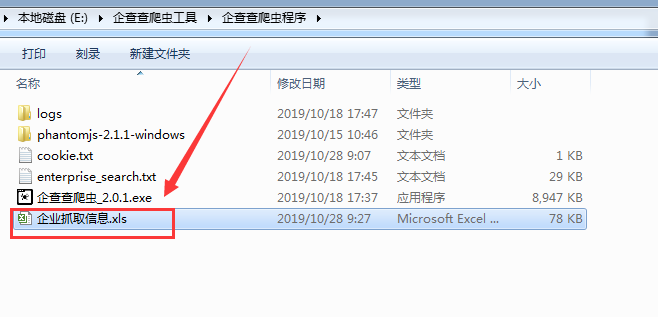
#### **启动程序**

双击企查查爬虫程序“企查查爬虫\_2.0.1.exe”启动爬虫程序。如下图：

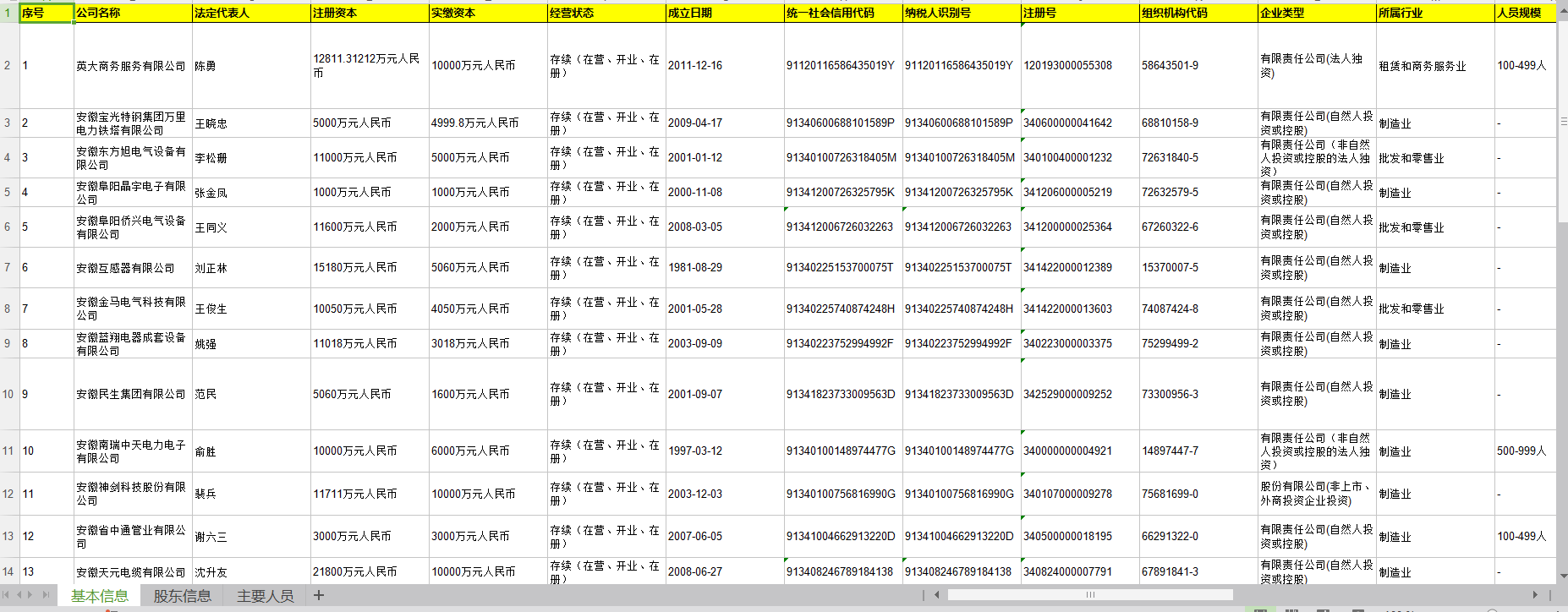


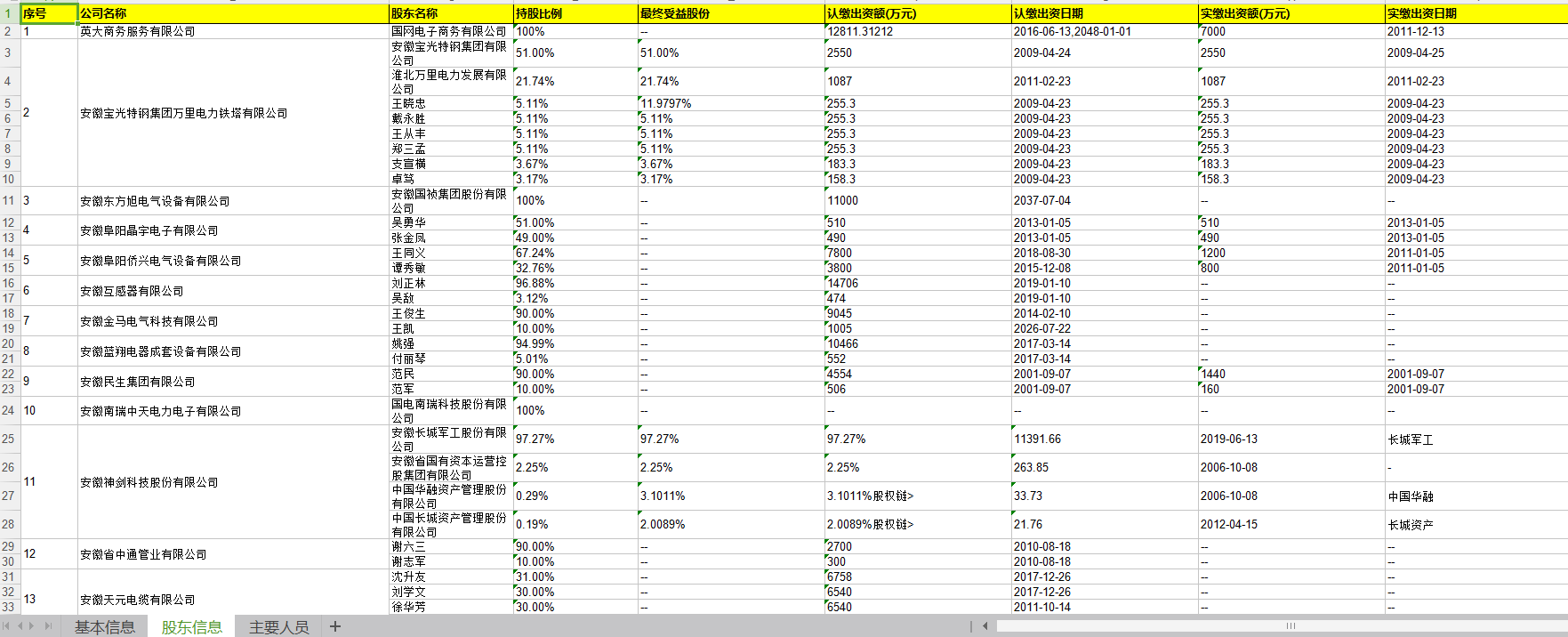


程序运行完毕会输出一个excel文件，包含爬取的企业信息结果。如下图所示：



文件内容如下图：







#### **注意事项**

1. **模拟登录复制cookie的操作不需要用一次软件复制一次，这个cookie的有效期很长，应该可以用好几天。然后保证不要对Chrome浏览器进行清空，特别是不要选清除cookie选项，否则cookie立即失效。如下图：**



1. 不要手动创建“企业抓取信息.xls”文件，程序会自动创建。
2. 数据是追加到这个excel表格的，即关闭了程序，第二次再打开程序，不会覆盖之前采集好的数据，而是紧随其后写入。
3. 请勿在爬取过程中打开excel表格，否则会中断，而下次再爬这个企业的时候，数据就会重复了，如果不在乎那无所谓。