## 抖音短视频爬取实战 Fillder+Mitmproxy+appium

**【开发环境】**

* Python3
* Android environment
* Appium
* Mitmproxy

**【环境安装】**

1.下载adt-bundle-windows-x86\_64

ANDROID\_HOME 值是Android SDK的解压路径

Path (%ANDROID\_HOME%\platform-tools; %ANDROID\_HOME%\tools;)

2.下载nodejs

nodejs下载地址：<https://nodejs.org/en/download/>可直接下载免安装版.zip

解压nodejs，并配置环境变量

Path D:\app tools\nodeV8.7.0Winx64; （也就是node.exe文件所在的目录）  
测试：CMD中输入node-v，能看到版本号

3.安装Appium[下载地址](https://bitbucket.org/appium/appium.app/downloads/" \t "https://www.jianshu.com/p/_blank)<https://bitbucket.org/appium/appium.app/downloads/>

**方法1安装Appium。**

打开CMD窗口，执行 npm install -g appium 命令开始安装Appium.

Appium默认被安装到C:\Users\Administrator\AppData\Roaming\npm\ 下。

把路径C:\Users\Administrator\AppData\Roaming\npm\ (appium可执行文件所在的目录）添加到环境变量Path

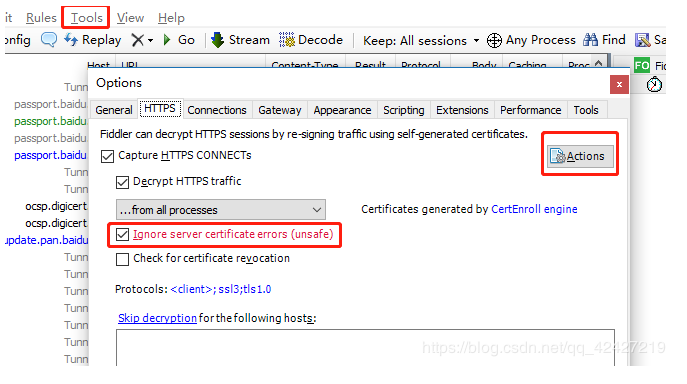
**方法2安装Appium。**

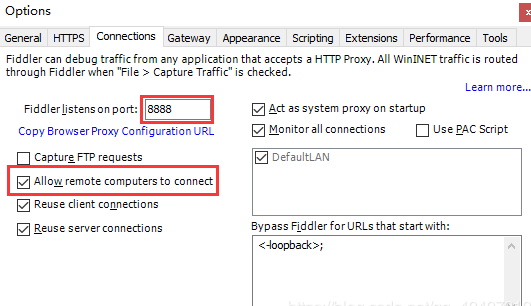
下载Appium压缩包<https://jingyan.baidu.com/article/3f16e003ea626e2590c1035c.html>

解压后配置appium环境变量  
Path D:\tools\appium\Appium\node\_modules.bin;  
测试：CMD中输入appium-doctor，能看到All Checks were successful

1. Fillder配置 下载地址（<https://www.telerik.com/download/fiddler> ）

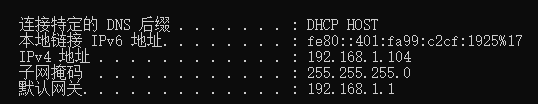
在tools中的options中，按照图中勾选后点击Actions



配置远程链接：  
选择允许监控远程链接，端口可以随意设置，只要别重复就行，默认8888

然后：重启fiddler！！！这样配置才能生效。

1. 夜神模拟器/手机配置

\*首先查看本机的IP：在cmd中输入ipconfig，记住这个IP

\*手机确保和电脑在同一局域网下。模拟器配置：设置中长按已连接wifi，代理选择手动，然后输入上图ip，端口号为8888



\*代理设置好后，在浏览器中输入你设置的ip：端口，例如10.10.16.194：8888，就会打开fiddler的页面。然后点击fiddlerRoot certificate安装证书，要不手机会认为环境不安全。



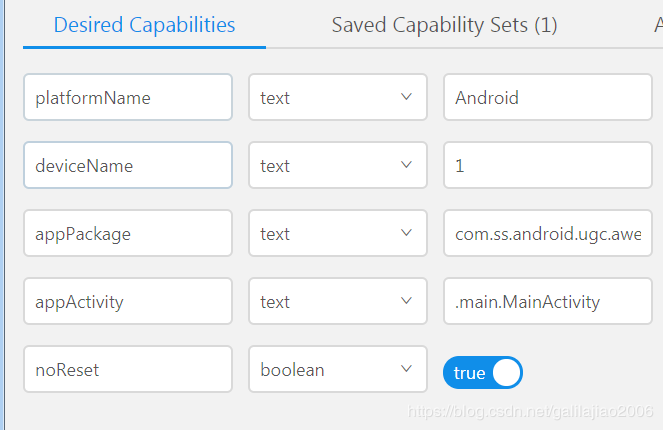
**【程序思想】**

1 启动弄mitmproxy：mitmdump -p 8888 -s xxx.py

2 利用appium自动化测试工具，驱动app模拟人的动作（滑动、点击等）

3 mitmproxy中间人代理的方式抓取数据包，并保存我们需要的数据。

**appium实现抖音app自动滑动实现粉丝数据包的处理**

1 测试是否连通,代码如下



from appium import webdriver

#等待元素加载

from selenium.webdriver.support.ui import WebDriverWait

#定义设备参数

cap={

"platformName": "Android",

"platformVersion": "5.1.1",

"deviceName": "127.0.0.1:62025",

"appPackage": "com.ss.android.ugc.aweme",

"appActivity": "com.ss.android.ugc.aweme.main.MainActivity",

"noReset": True,

'unicodekeyboard':True, #使用unicode输入法

'resetkeyboard':True, #还原输入法

}

driver =webdriver.Remote('http://localhost:4723/wd/hub',cap)

2 连通后使用uiautomator viewer,获取空间的id,xpath...表达式

注释事项:抓取截屏数据的时候,必须等待数据加载完毕,或者把视频暂停,否则报错此过程使用了appium和uiautomator viewer交替定位元素.

