北京邮电大学

网络空间安全学院



**测试及结果展示报告**

**项目： 基于DNS流量分析的僵尸网络检测工具**

**组员： 王硕、彭致远、李懿飞、王晨旭**

**2020年12月10日**

目录

[1 引言 3](#_Toc58503482)

[1.1 目的 3](#_Toc58503483)

[1.2 背景及范围 3](#_Toc58503484)

[1.3 定义、术语和缩写 3](#_Toc58503485)

[2 测试概述 4](#_Toc58503486)

[2.1 测试环境与配置 4](#_Toc58503487)

[2.2 测试内容 4](#_Toc58503488)

[2.2.1 机器学习 4](#_Toc58503489)

[2.2.2 文件检测 4](#_Toc58503490)

[2.2.3 实时检测 5](#_Toc58503491)

[2.2.4 用户交互 5](#_Toc58503492)

[3 运行结果与分析 5](#_Toc58503493)

[3.1 机器学习 5](#_Toc58503494)

[3.1.1 DNS特征的重要度分析 5](#_Toc58503495)

[3.1.2 主机分类器准确率 6](#_Toc58503496)

[3.1.3 模型的保存与加载 7](#_Toc58503497)

[3.2 文件检测 7](#_Toc58503498)

[3.2.1 文件检测结果输出 7](#_Toc58503499)

[3.2.2 查看生成的csv备份文件 9](#_Toc58503500)

[3.3 实时检测 10](#_Toc58503501)

[3.3.2 准备工作 10](#_Toc58503502)

[3.3.2 正常上网报告正常 11](#_Toc58503503)

[3.3.3 搭建信道报告异常主机 12](#_Toc58503504)

[3.1 用户交互 14](#_Toc58503505)

[3.2.1 用户输入安全检查 14](#_Toc58503506)

[3.1.2 命令行交互 16](#_Toc58503507)

[4 结论 18](#_Toc58503508)

# 1 引言

### 1.1 目的

本测试及结果展示报告为大三上学期网络分析实践课程的“基于DNS流量分析的僵尸网络检测工具”的测试及展示报告，目的在于测试工具的基本功能并分析测试结果，判断系统是否符合需求，并将运行结果附在本报告中

### 1.2 背景及范围

* + 项目名称：基于DNS流量分析的僵尸网络检测工具
  + 项目成员：北京邮电大学网络空间安全学院“网络安全分析实践”课程开发小组
    - 王硕（组长）：2018213641
    - 彭致远：2018213646
    - 李懿飞：2018213632
    - 王晨旭：2018213636
  + 系统范围：具有python2.7.9环境的Windows系统计算机
  + 用户：无限制
  + 实现项目的计算机网络：校园网

本测试报告预期参考人员包括测试工具的同学、开发工具的同学、验收工具的老师。

### 1.3 定义、术语和缩写

|  |  |  |
| --- | --- | --- |
| 序号 | 术语或缩写 | 解释 |
| 1 | DNS | 域名系统服务协议，是一种分布式网络目录服务，主要用于域名与 IP 地址的相互转换，以及控制因特网的电子邮件的发送 |
| 2 | DNS特征 | 用来衡量主机进行DNS通信过程的行为 |
| 3 | DNS指纹 | 根据特征工程，提取主机DNS通信特征，对主机行为进行了多维度标识 |
| 4 | Botnet | 僵尸网络，是指采用一种或多种传播手段，将大量主机感染bot程序（僵尸程序）病毒，从而在控制者和被感染主机之间所形成的一个可一对多控制的网络 |
| 5 | 僵尸主机 | 本文表示处于僵尸网络中的主机 |
| 6 | 域/域名 | 域名（英语：Domain Name），是由一串用点分隔的名字组成的Internet上某一台计算机或计算机组的名称，用于在数据传输时对计算机的定位标识 |

# 2 测试概述

### 2.1 测试环境与配置

* + 操作系统：Windows 10
  + Python环境：Python2.7.9
  + Python安装依赖包：ipaddr、dpkt、geoip2、matplotlib、win\_inet\_pton、gephistreamer、pandas、numpy、seaborn、sklearn、itertools
  + 计算机网络：校园网
  + 虚拟机：kali linux 2020、Ubuntu 18

### 2.2 测试内容

#### 2.2.1 机器学习

|  |  |  |
| --- | --- | --- |
| 序号 | 功能 | 要求 |
| 1 | DNS特征的重要度分析 | 特征重要度排名图 |
| 2 | 机器学习分类 | 准确率达到95%以上 |
| 3 | 模型的保存与加载 | 查看是否生成模型 |

#### 2.2.2 文件检测

|  |  |  |
| --- | --- | --- |
| 序号 | 功能 | 要求 |
| 1 | 检测结果输出 | 列出可疑主机并人为判断 |
| 2 | Pcap文件解析到csv | 查看生成的csv文件是否符合要求 |

#### 2.2.3 实时检测

|  |  |  |
| --- | --- | --- |
| 序号 | 功能 | 要求 |
| 1 | 实时检测是否有异常主机 | 正常上网，没有检测出僵尸主机  搭建好dns隐蔽信道，检测出来并报告 |
| 2 | 用户可实时终止 | 验证多线程是否正常运行 |

#### 2.2.4 用户交互

|  |  |  |
| --- | --- | --- |
| 序号 | 功能 | 要求 |
| 1 | 错误输入过滤 | 输入命令不合法，看是否有回显 |
| 2 | 文件安全检查 | 文件不存在，提示用户，程序不报错 |
| 3 | 用户可处理检测结果 | 按照流程走一遍不报错 |

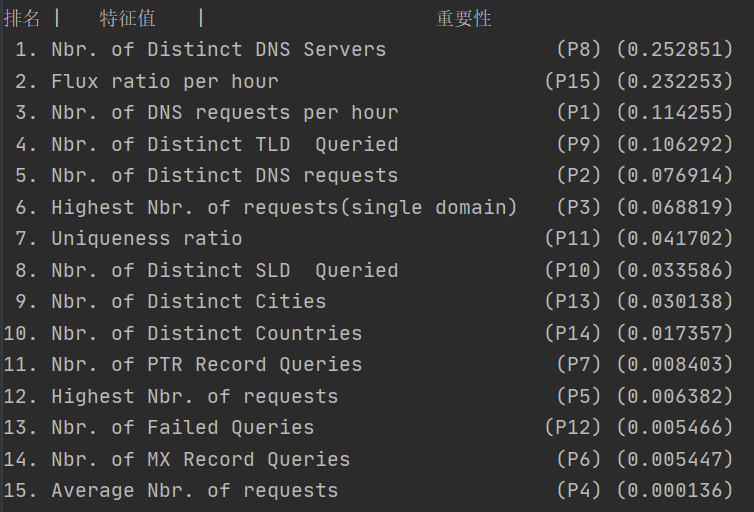
# 3 运行结果与分析

### 3.1 机器学习

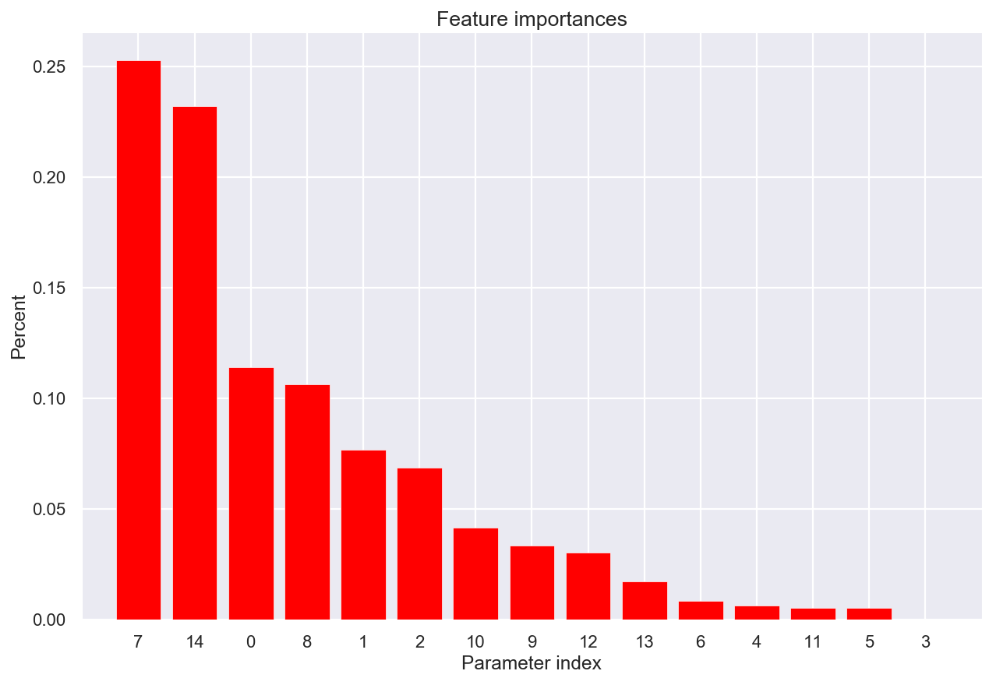
#### 3.1.1 DNS特征的重要度分析

运行结果如下图：

1. 输出每个特征的重要程度：

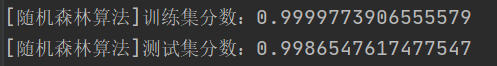


1. 柱状图表示：



#### 3.1.2 主机分类器准确率

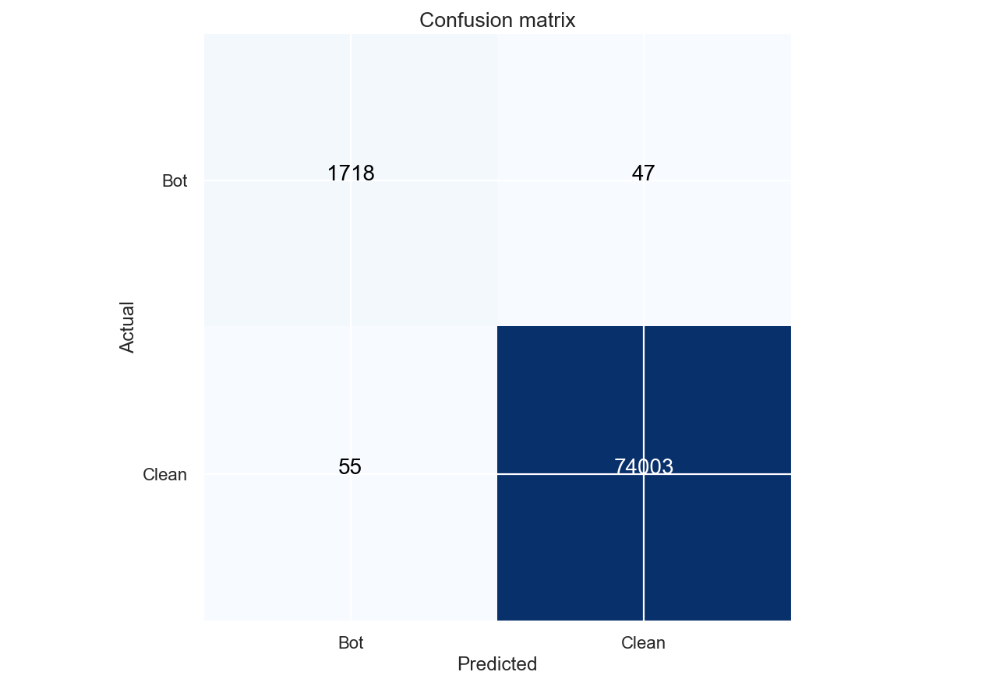
1、以训练集：测试集 = 6:4的比例进行机器学习，输出随机森林算法分类模型的准确率如下：



2、引入混淆矩阵，四类结果的准确率如下：

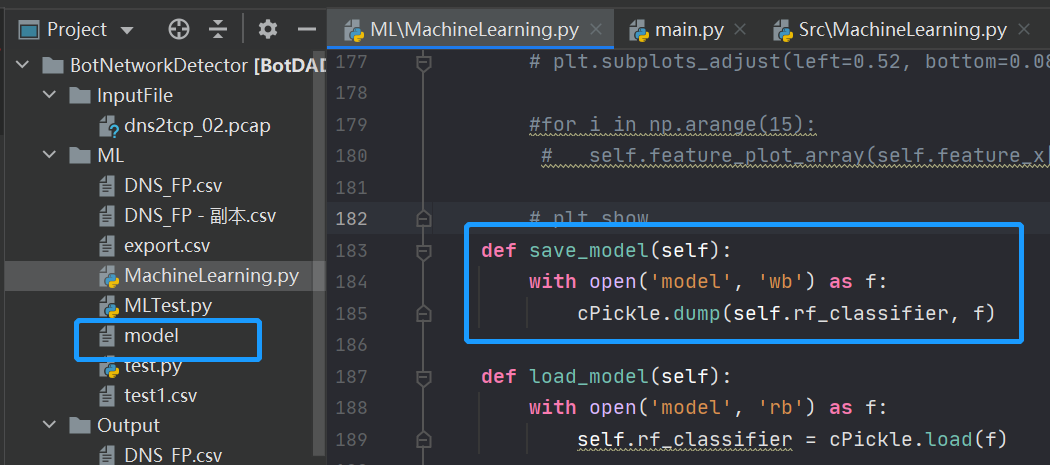
* 将负类预测为负类数,真实为Bot，预测为Bot
* 将正类预测为负类数,真实为Clean，预测为Bot
* 将正类预测为正类数,真实为Clean，预测为Clean
* 将负类预测为正类数,真实为Bot，预测为Clean





#### 3.1.3 模型的保存与加载

1、执行保存模型，在文件夹内保存结果model

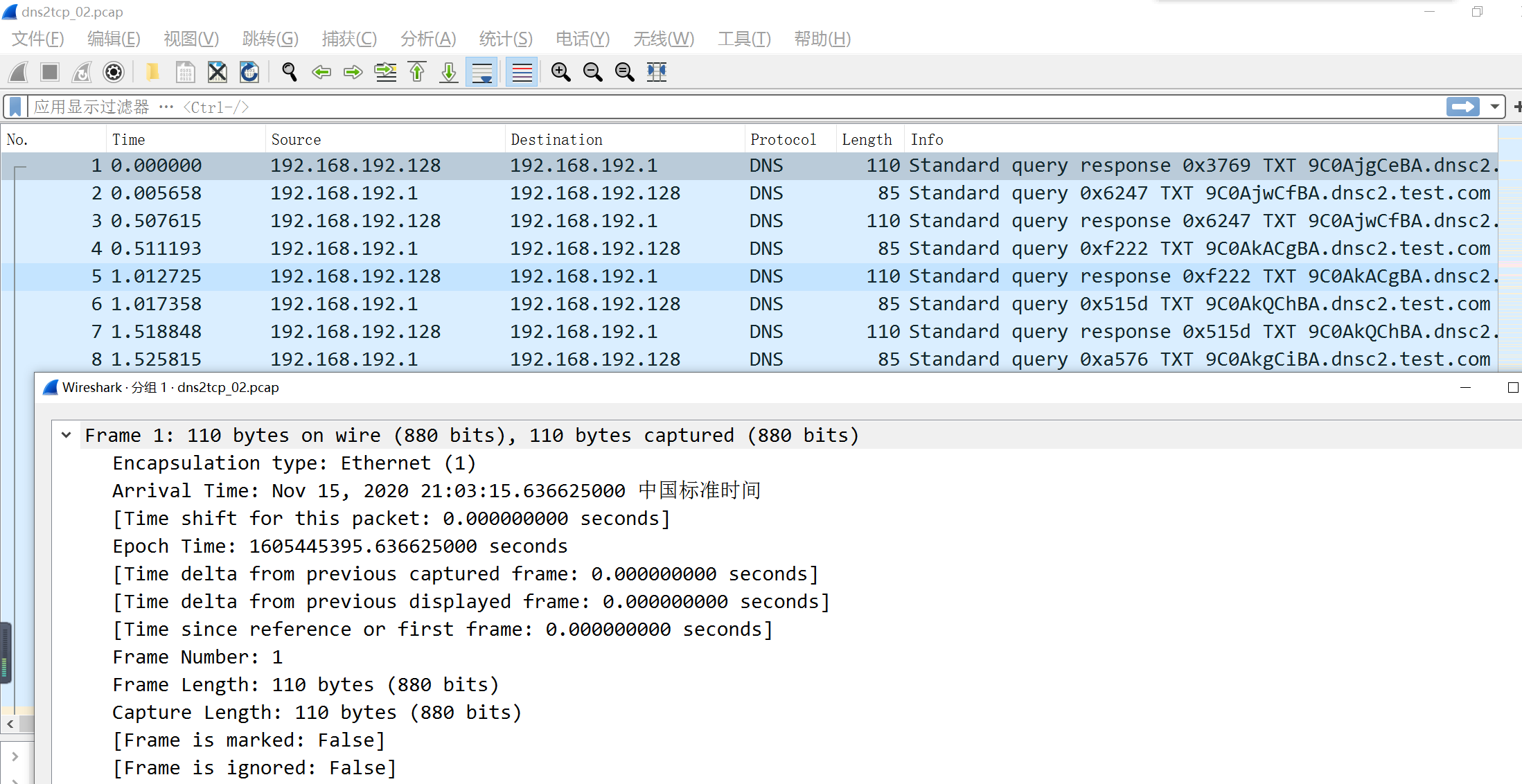


2、执行load\_model可以加载模型，运行结果同上。

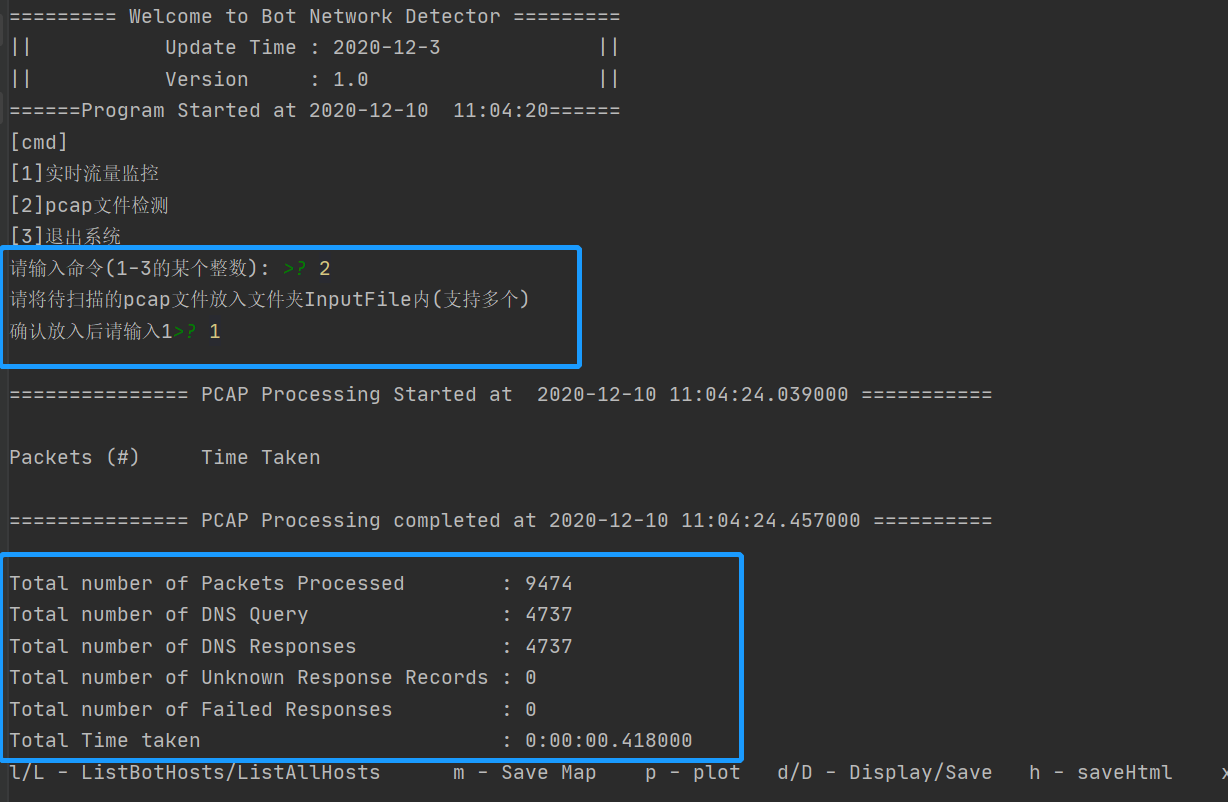
### 3.2 文件检测

#### 3.2.1 文件检测结果输出

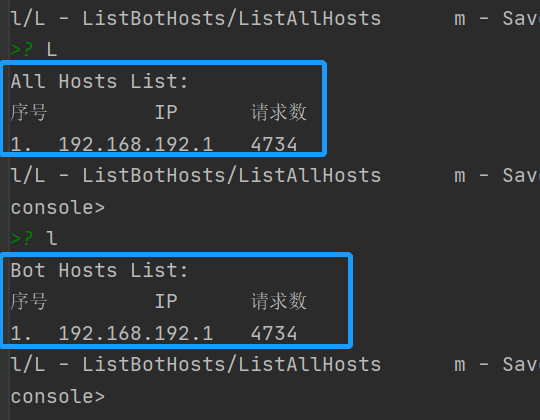
1、对于自己搭建的信道，pcap文件如下：



2、执行文件检测功能，输出结果如下：

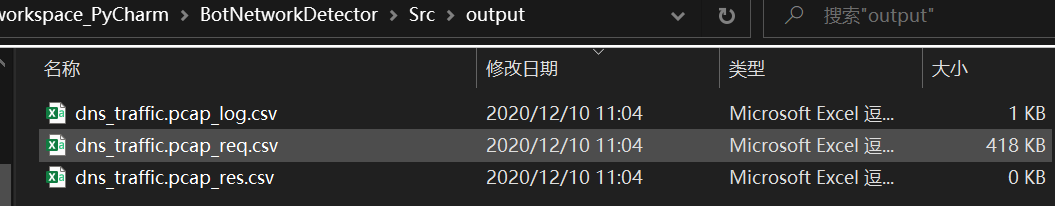


1. 列出所有主机和受害主机，成功检测出来

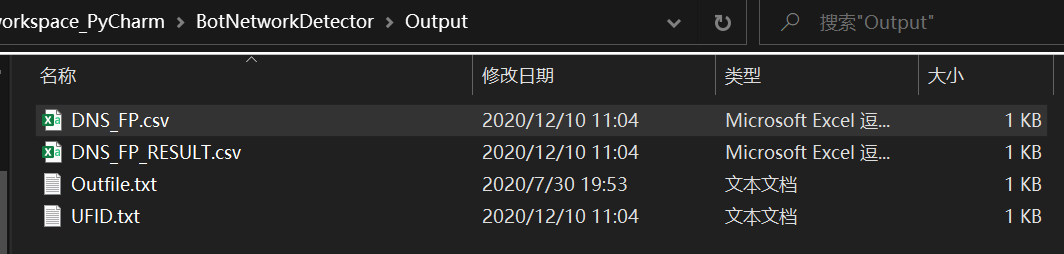


#### 3.2.2 查看生成的csv备份文件

1、在output文件夹内生成了备份文件，分别是日志、请求、响应的记录：

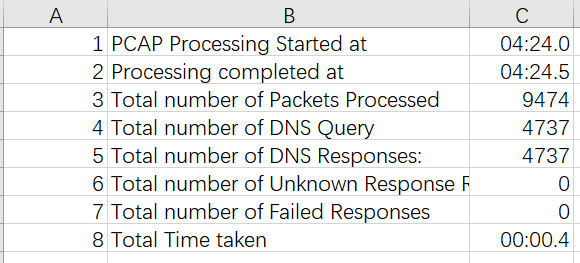


在Output文件夹内生成了特征值及结果的csv文件：

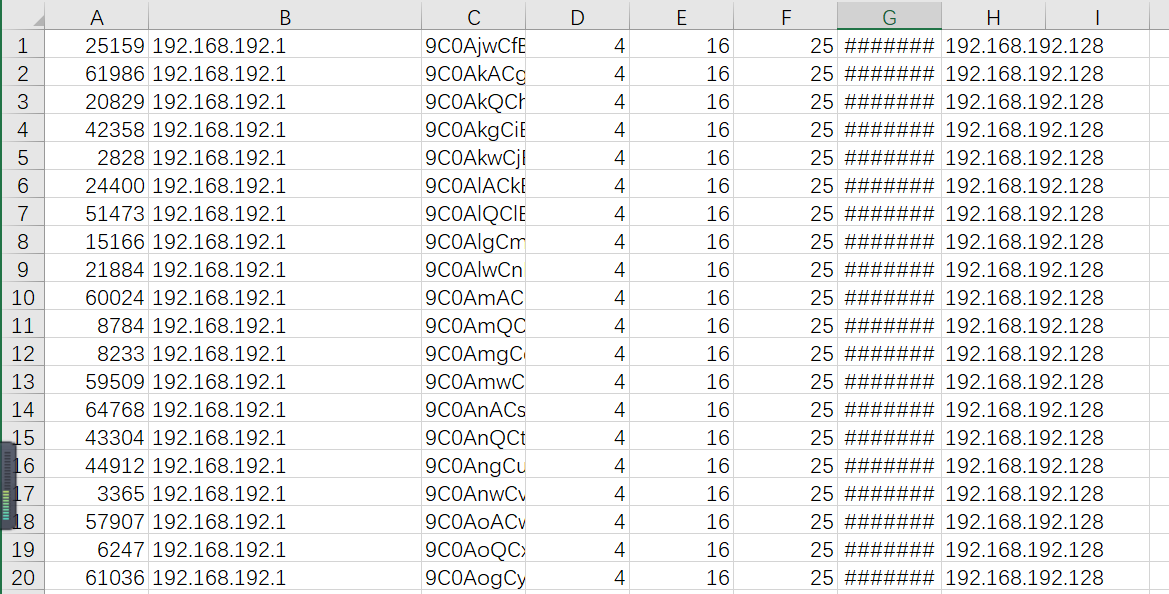


打开查看效果如下：

（1）日志记录：



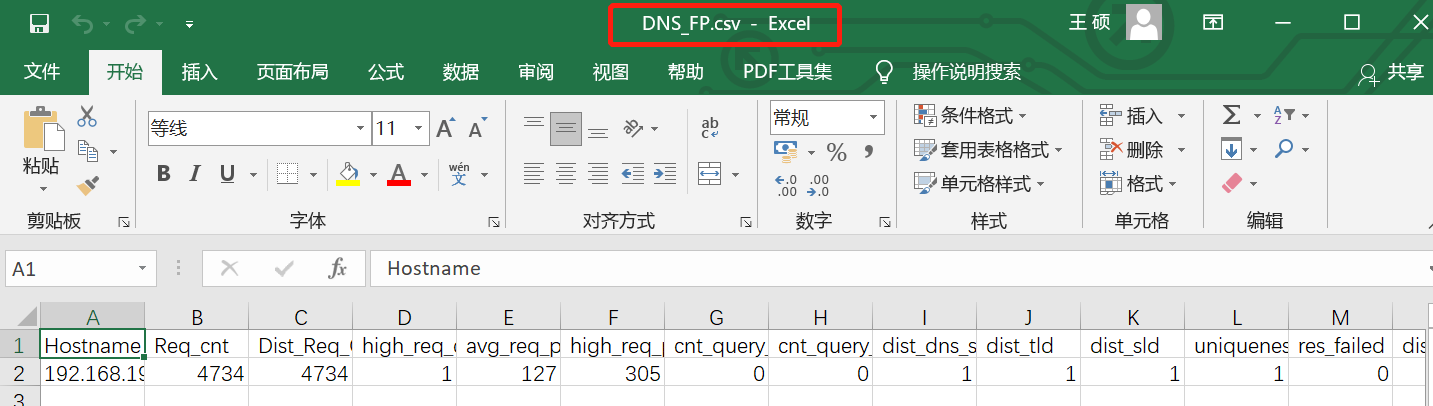
（2）request：



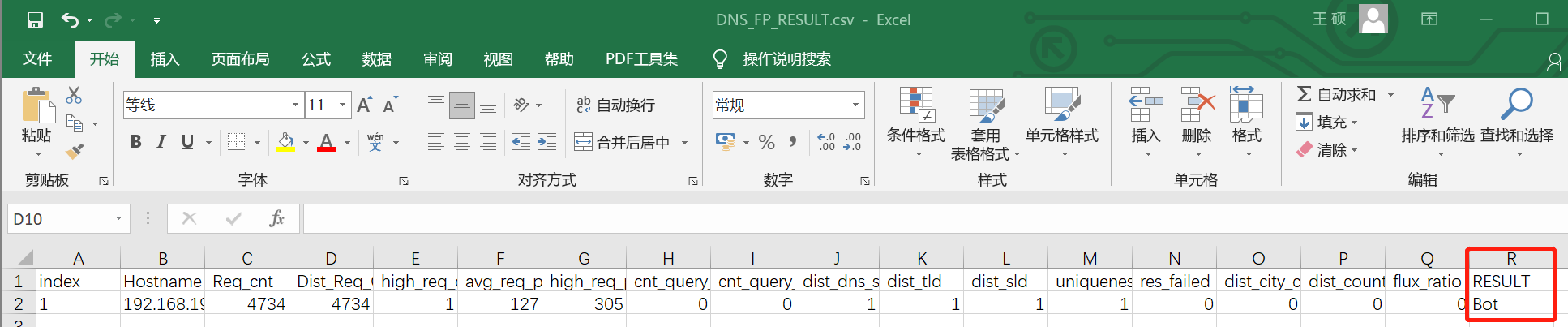
（3）response：

为空，符合实际要求

（4）特征值记录：



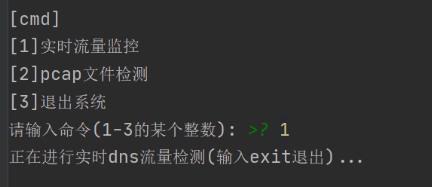
（5）结果：



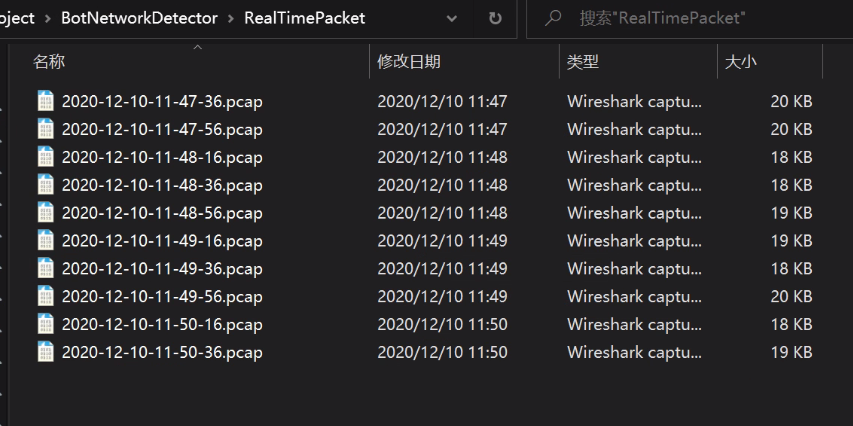
### 3.3 实时检测

#### 3.3.2 准备工作

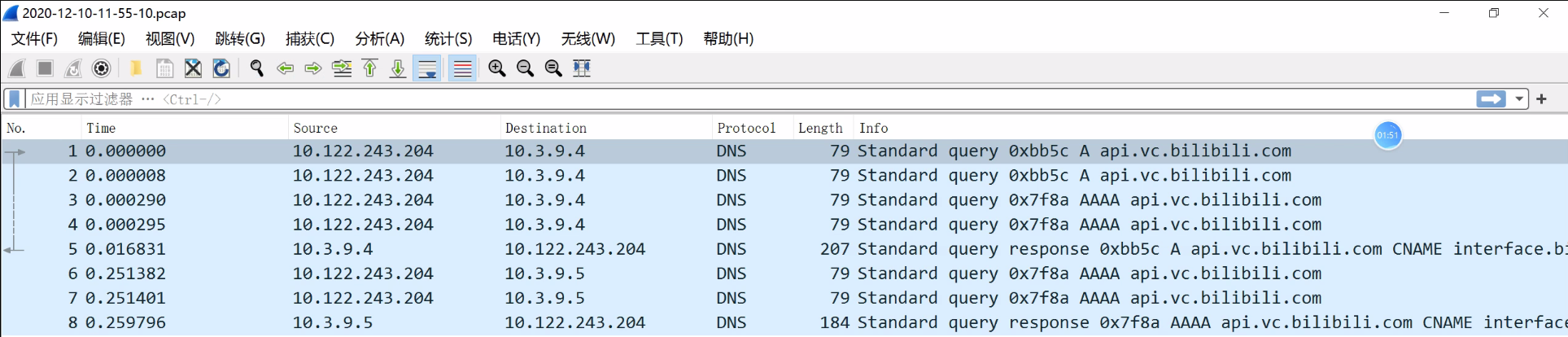
1、启动系统，并启动实时检测模块：



2、可以在RealTimePacket文件夹内看到抓到的流量包：



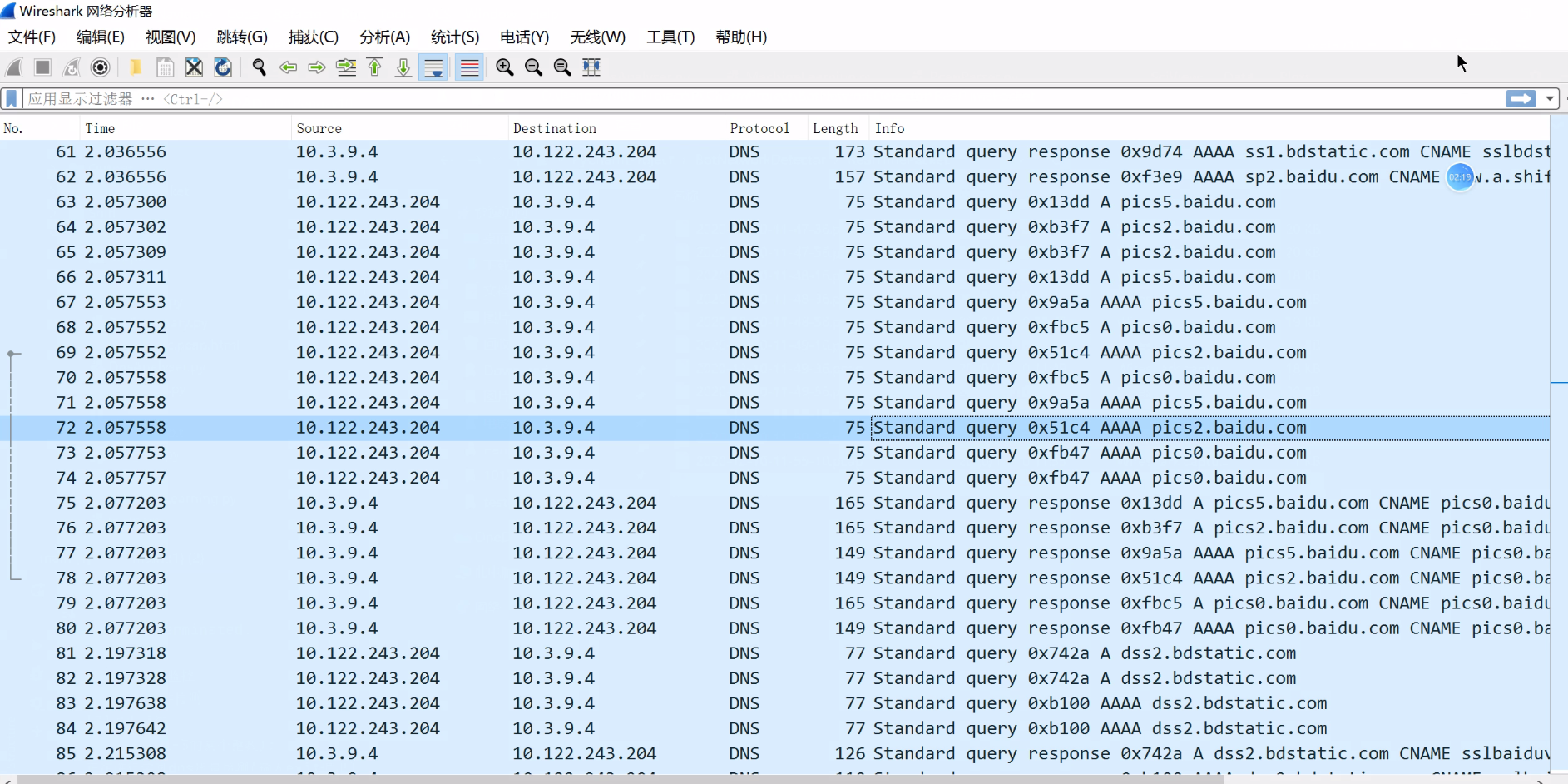
打开查看：



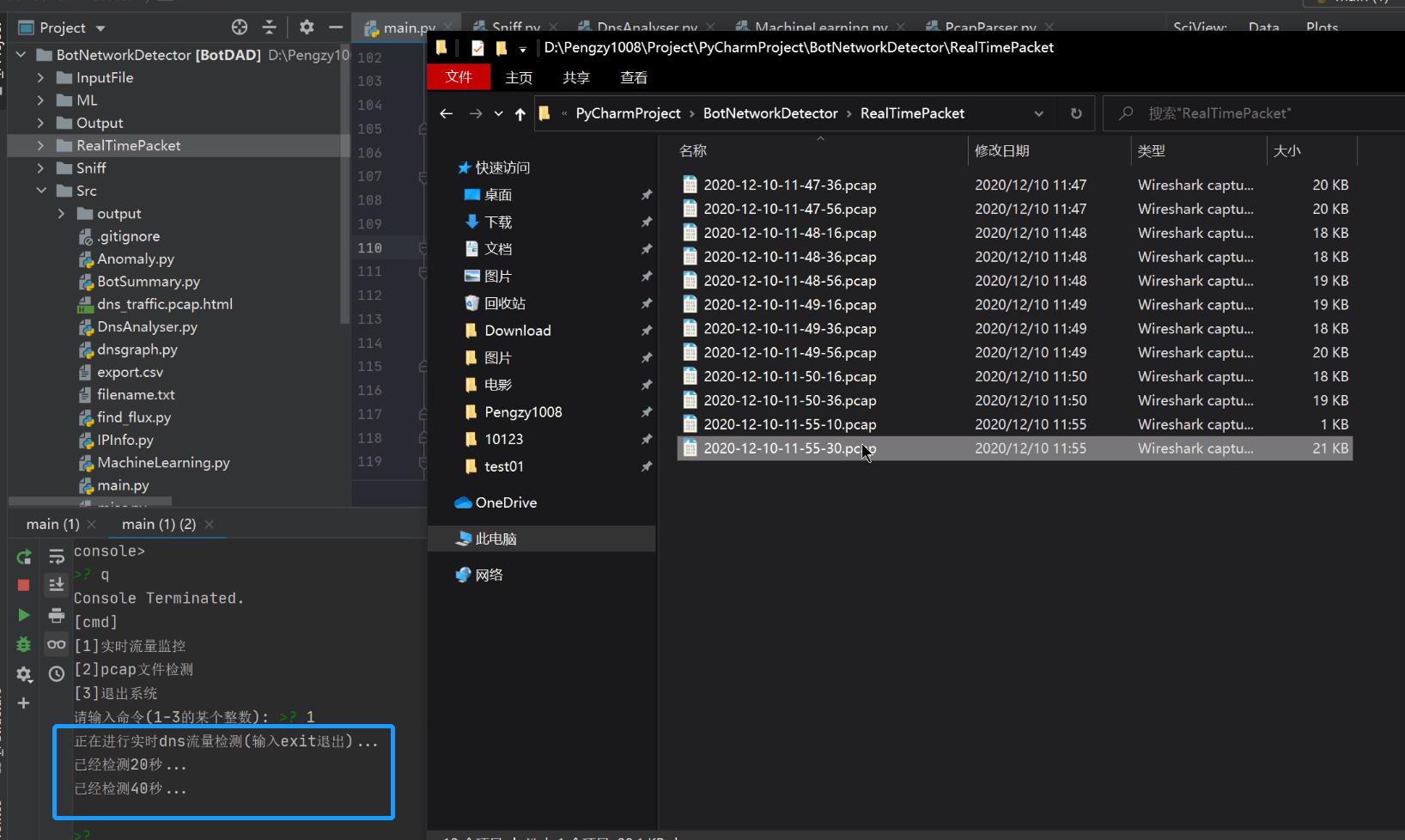
#### 3.3.2 正常上网报告正常

1、正常上网，网页及抓取的流量包展示：





2、每20秒报告一次：

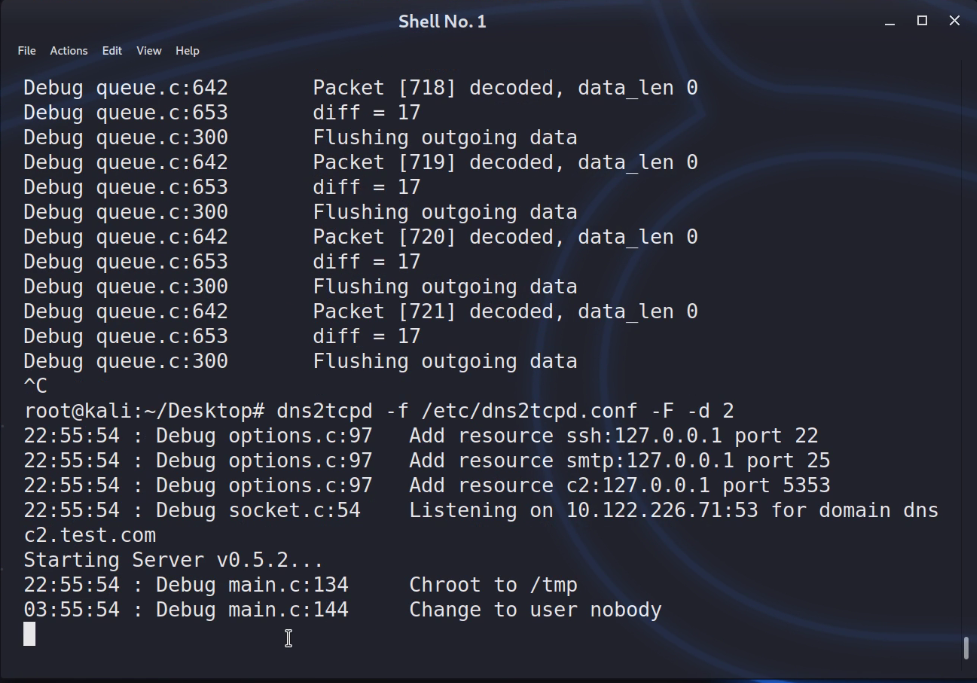


3、未发现异常。

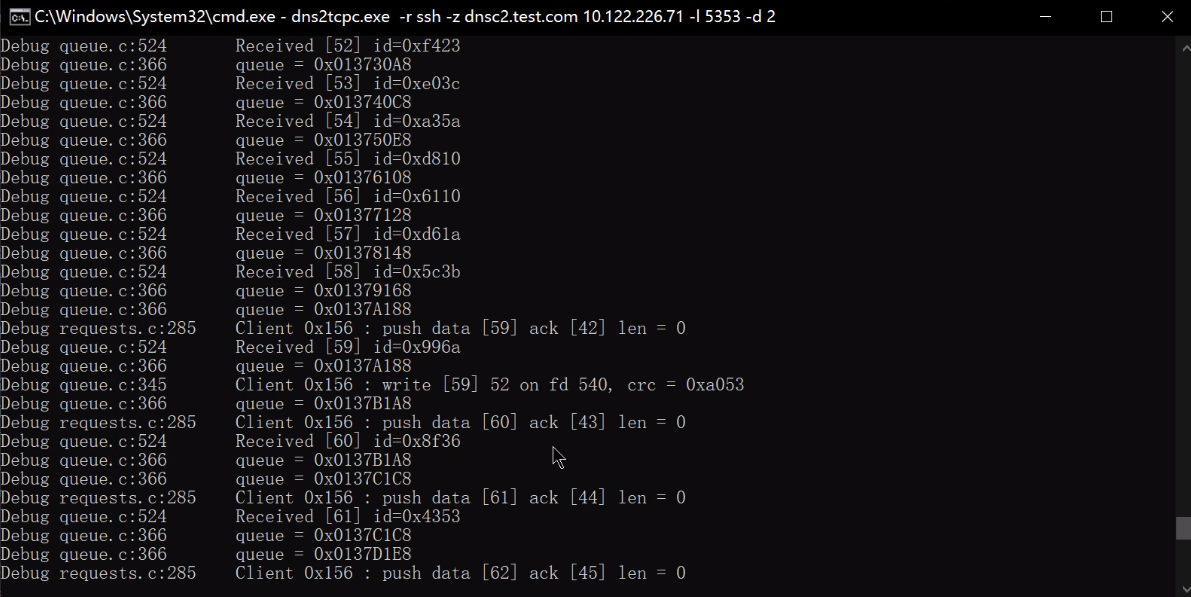
#### 3.3.3 搭建信道报告异常主机

1、搭建好DNS隐蔽信道：

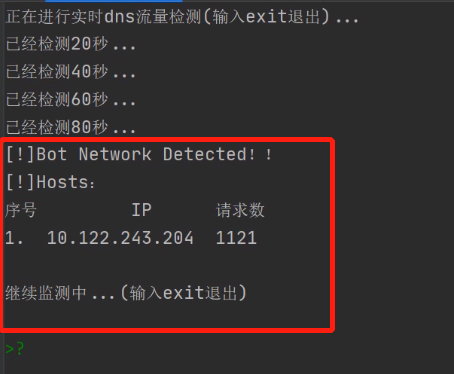
（1）攻击端：



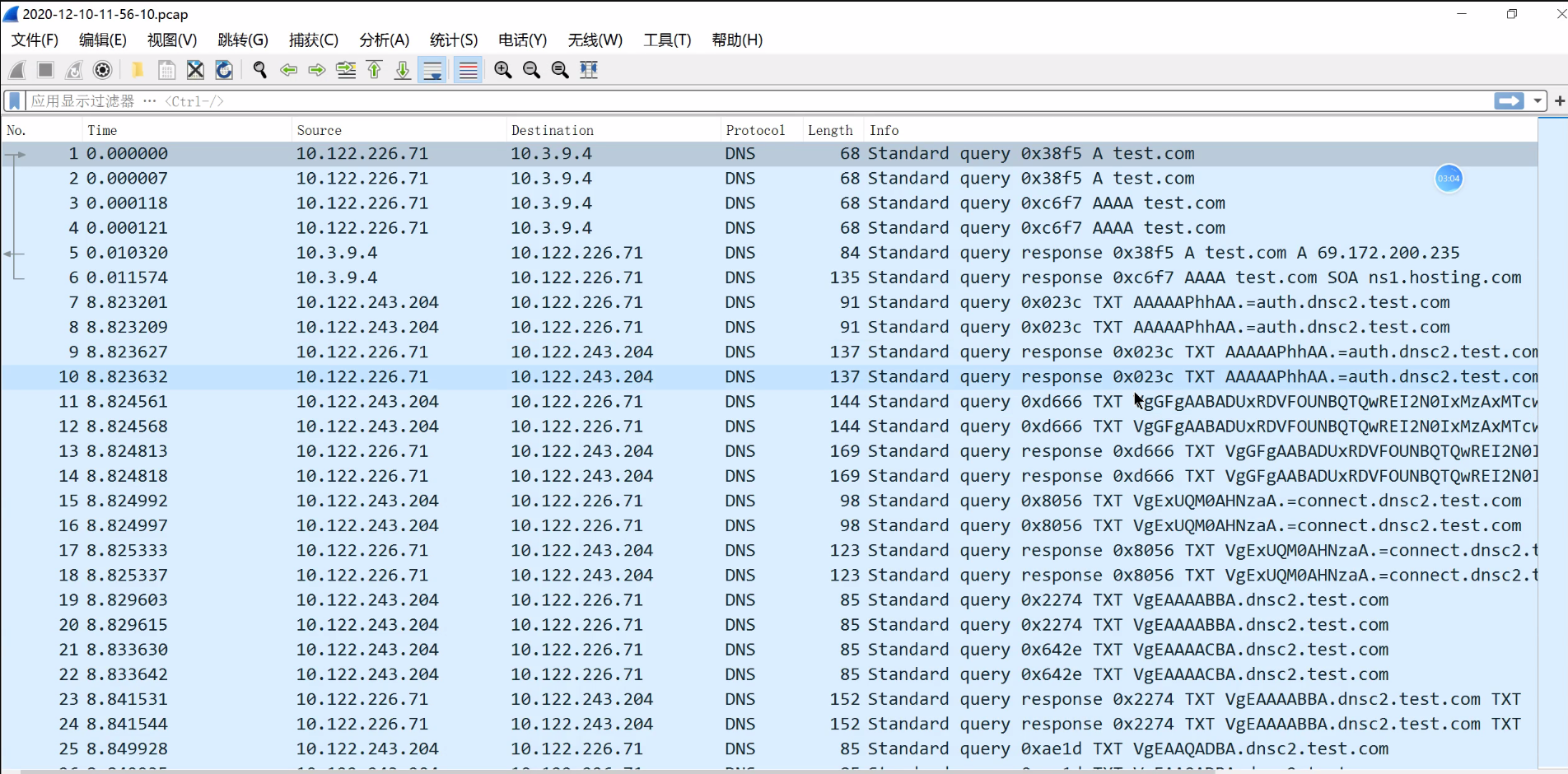
（2）客户端：



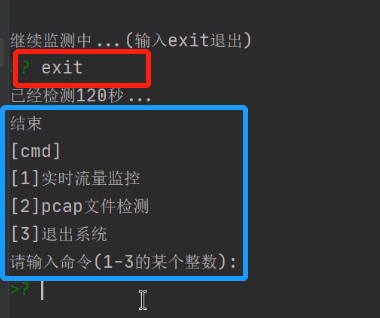
2、检测一会后，下一次报告显示发现了僵尸主机：



查看抓到的数据包，确实是搭建信道的隐蔽通信数据包：



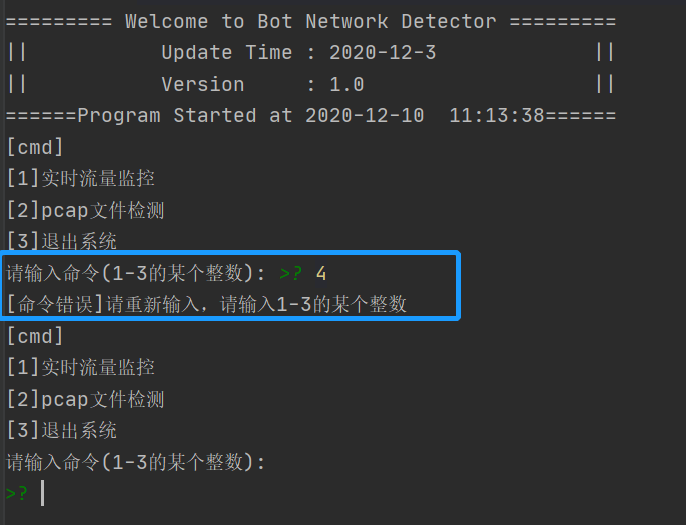
3、输入exit退出检测：

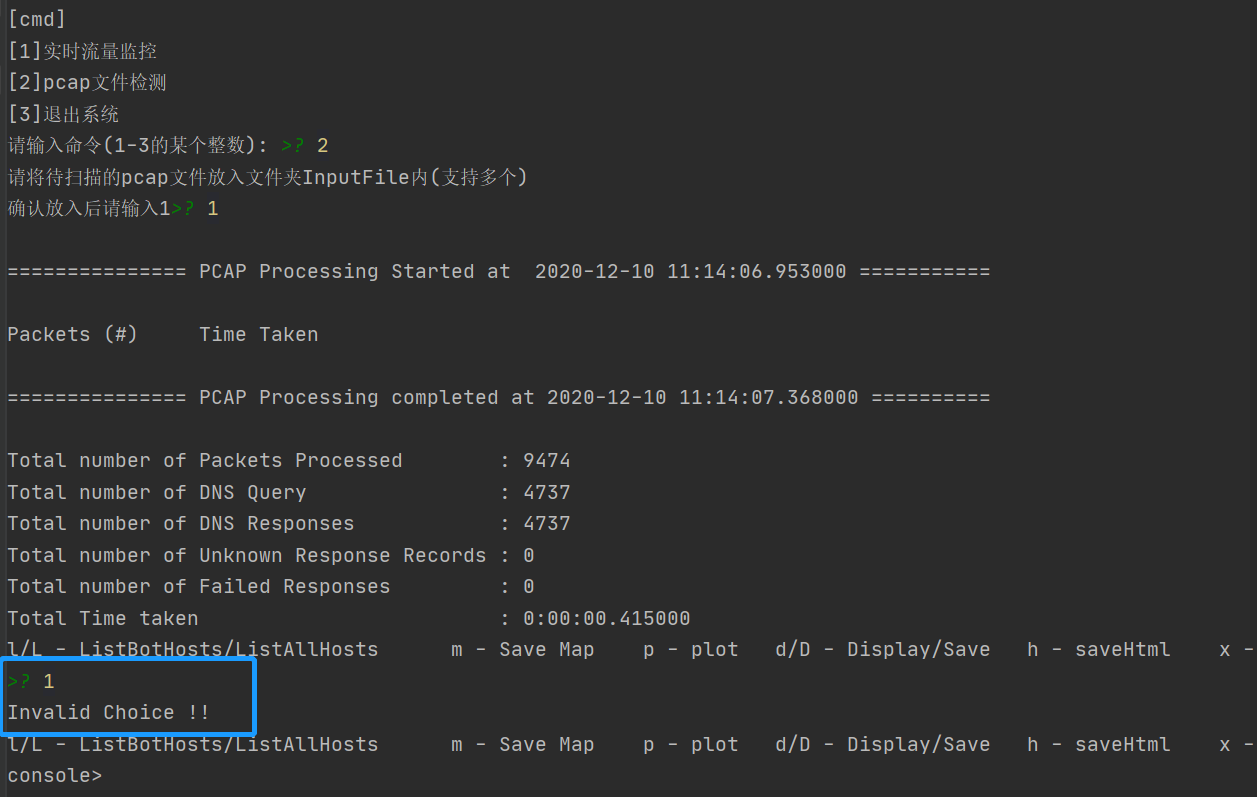


### 3.1 用户交互

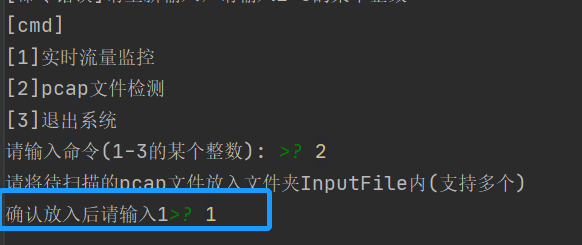
#### 3.2.1 用户输入安全检查

1、当用户输入命令错误时给予提示：





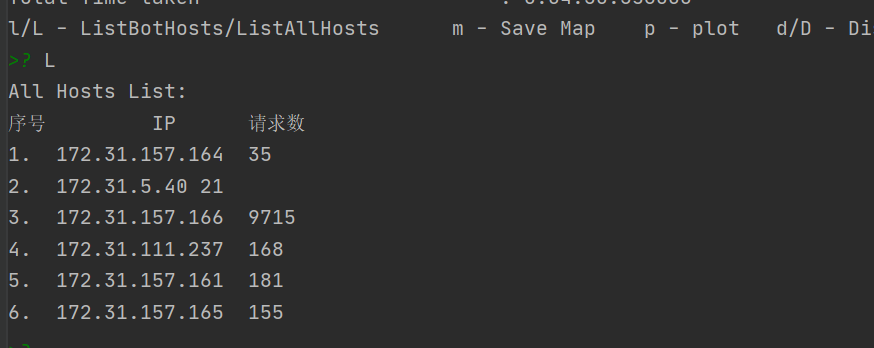
2、检查用户是否放入文件，并给用户提供时间：



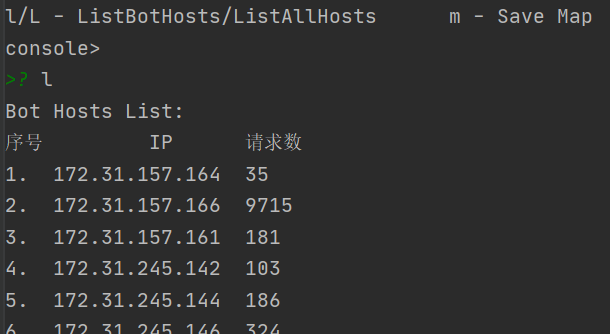
#### 3.1.2 命令行交互

此处扫描来自国外数据集内的流量，比较大，扫描结果如下：

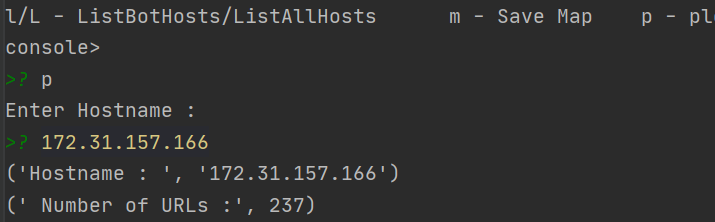
（1）扫描结束，输入L，列出所有主机信息：



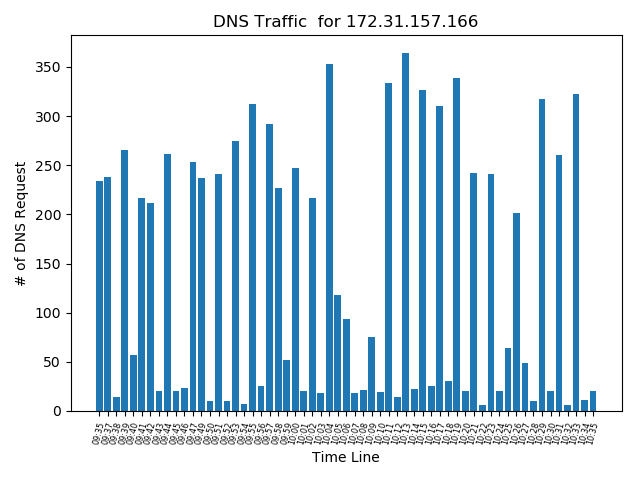
（2）输入l，列出所有感染了僵尸网络的主机信息：



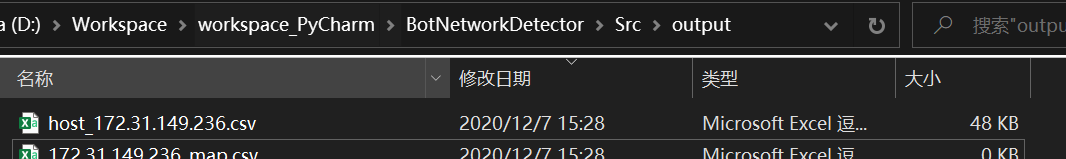
（3）画出主机这段时间的流量图，p命令：

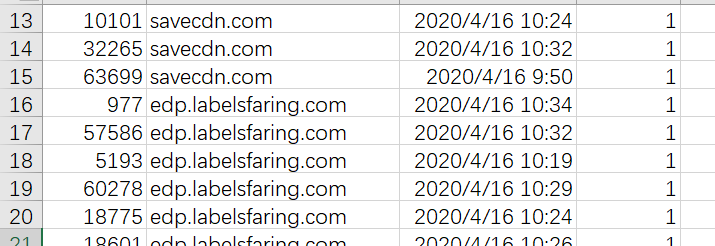


可以在plot区看到绘制的流量图：



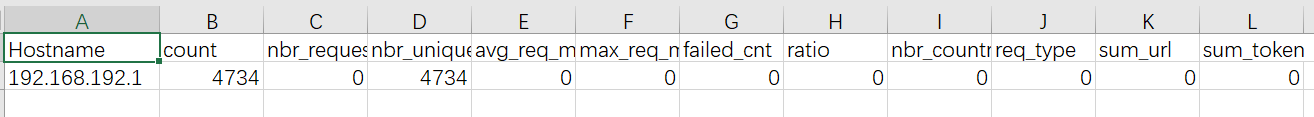
（4）输入D和主机ip=172.31.157.166，可以看到在output文件夹内生成了该主机的信息：



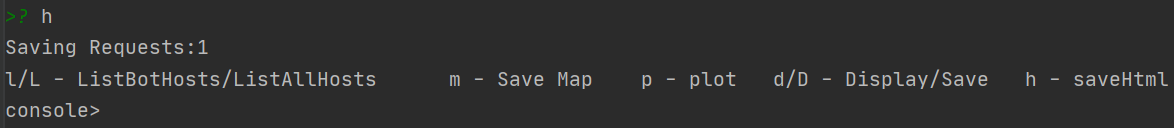


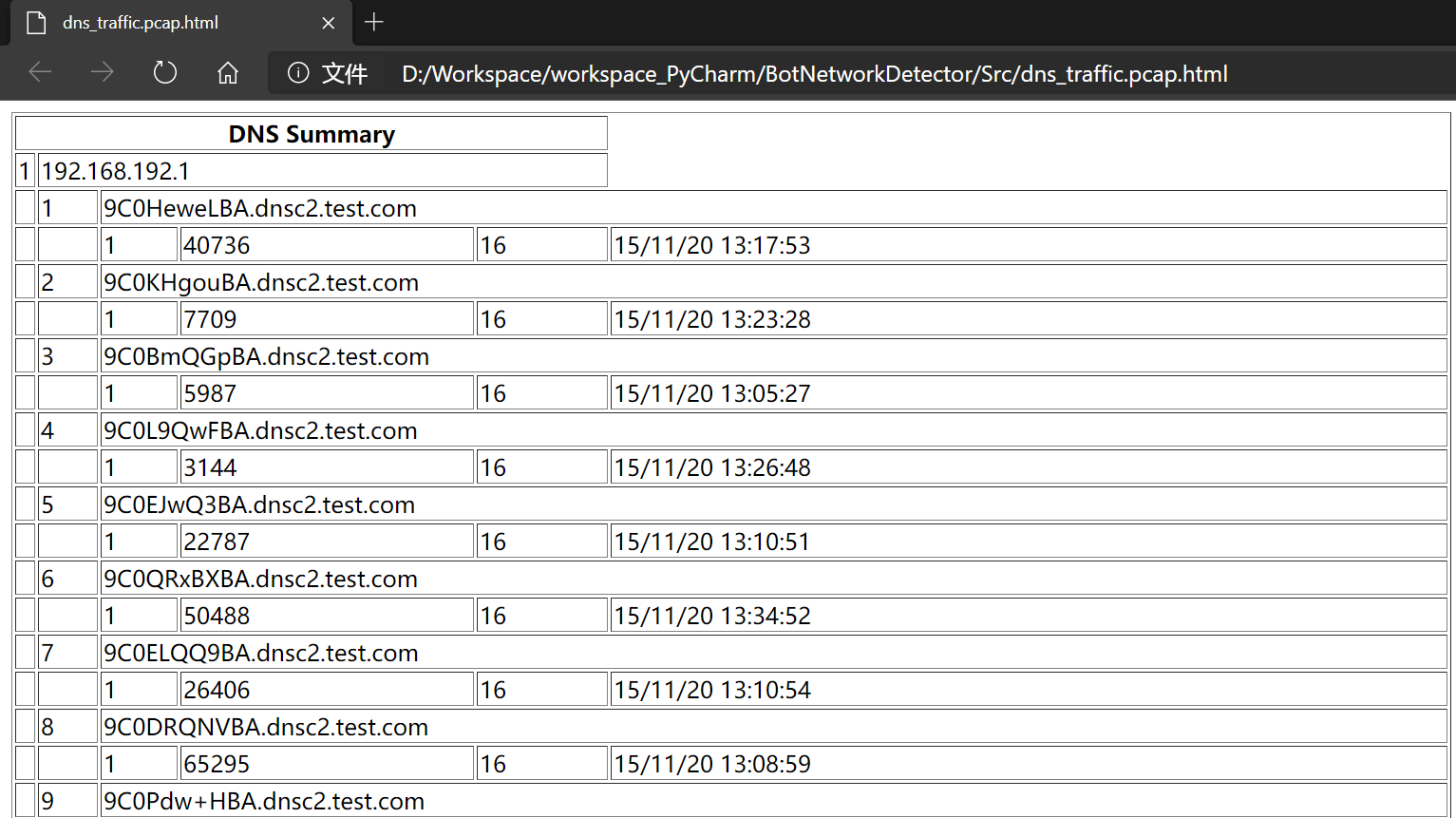
（5）输入x保存为csv，保存本次扫描的结果





（6）输入h保存为html，即保存DNS扫描的结果（Pcap文件的摘要）





# 4 结论

本工具测试基本通过，能够实现需求中的功能：

1. 功能较为全面
2. 检测准确率高、速度较快
3. 系统安全性较好
4. 用户友好程度高