# 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: <u>基于 DNS 流量分析的僵尸网络检测工具</u>

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
	1.1 目的	3
	1.2 背景及范围	
	1.3 定义、术语和缩写	
2	· 测试概述	4
	2.1 测试环境与配置	4
	2.2 测试内容	4
	2.2.1 机器学习	4
	2.2.2 文件检测	4
	2.2.3 实时检测	5
	2.2.4 用户交互	5
3	'运行结果与分析	5
	3.1 机器学习	5
	3.1.1 DNS 特征的重要度分析	5
	3.1.2 主机分类器准确率	6
	3.1.3 模型的保存与加载	7
	3.2 文件检测	7
	3.2.1 文件检测结果输出	7
	3.2.2 查看生成的 csv 备份文件	9
	3.3 实时检测	10
	3.3.2 准备工作	10
	3.3.2 正常上网报告正常	11
	3.3.3 搭建信道报告异常主机	12
	3.1 用户交互	14
	3.2.1 用户输入安全检查	14
	3.1.2 命令行交互	16
4	· 结论	18

# 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
		域名系统服务协议,是一种分布式网络目录服务,主
1	DNS	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
0	DNS 指纹	根据特征工程,提取主机 DNS 通信特征,对主机行
3		为进行了多维度标识
		僵尸网络,是指采用一种或多种传播手段,将大量主
4	Botnet	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

		域名(英语:Domain Name),是由一串用点分隔的
6	域/域名	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 10Python 环境: Python 2.7.9

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	正常上网,没有检测出僵尸主机				
	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

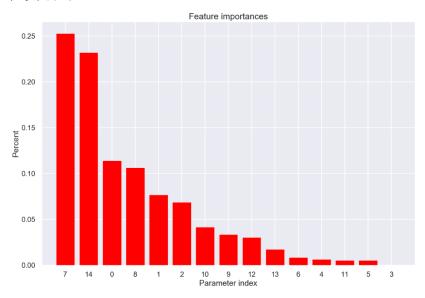
# 3.1.1 DNS 特征的重要度分析

运行结果如下图:

1、输出每个特征的重要程度:

排名	特征值	重要性		
1.	Nbr. of Distinct DNS Servers		(P8)	(0.252851)
2.	Flux ratio per hour		(P15)	(0.232253)
3.	Nbr. of DNS requests per hour		(P1)	(0.114255)
4.	Nbr. of Distinct TLD Queried		(P9)	(0.106292)
5.	Nbr. of Distinct DNS requests		(P2)	(0.076914)
6.	Highest Nbr. of requests(single	domain)	(P3)	(0.068819)
7.	Uniqueness ratio		(P11)	(0.041702)
8.	Nbr. of Distinct SLD Queried		(P10)	(0.033586)
9.	Nbr. of Distinct Cities		(P13)	(0.030138)
10.	Nbr. of Distinct Countries		(P14)	(0.017357)
11.	Nbr. of PTR Record Queries		(P7)	(0.008403)
12.	Highest Nbr. of requests		(P5)	(0.006382)
13.	Nbr. of Failed Queries		(P12)	(0.005466)
14.	Nbr. of MX Record Queries		(P6)	(0.005447)
15.	Average Nbr. of requests		(P4)	(0.000136)

#### 2、柱状图表示:

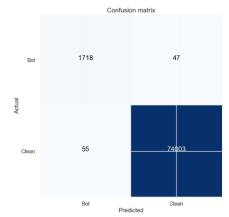


### 3.1.2 主机分类器准确率

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:

[随机森林算法]训练集分数: 0.9999773906555579 [随机森林算法]测试集分数: 0.9986547617477547

- 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

```
■ Project ▼ ⊕ 😤 🌣 — 🐔 ML\MachineLearning.py × 🐔 main.py × 🐔 Src\MachineLearning.py
                                           # plt.subplots_adjust(left=0.52, bottom=0.0
BotNetworkDetector [BotDAI]

✓ InputFile

      dns2tcp_02.pcap
  ∨ 🖿 ML
                                            # self.feature_plot_array(self.feature_x
      ₫ DNS_FP.csv
      ■ DNS_FP - 副本.csv
      a export.csv
                                       def save_model(self):
      MachineLearning.py
      MLTest.py
                                               cPickle.dump(self.rf_classifier, f)
      🖆 model
      test.py
                                       def load_model(self):
      test1.csv
  🗸 🖿 Output
                                               self.rf_classifier = cPickle.load(f)
```

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:

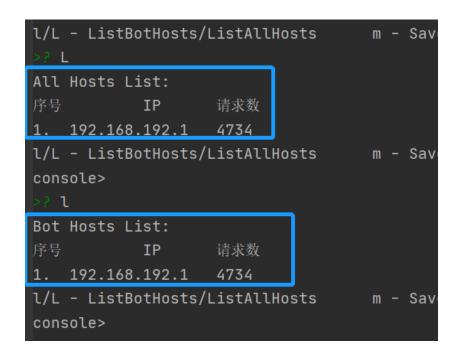
```
文件(F) 编辑(E) 视图(V) 跳转(G) 捕获(C) 分析(A) 统计(S) 电话(Y) 无线(W) 工具(T) 帮助(H)
# ■ ② ● | ■ ② ② | Q ← ← ← ● ● ▼ ③ □ □ ■ Q Q Q Ⅲ
| 使用显示过滤器 --- (Ctri-/>
       | Time | Source | Destination | Protocol | Length | Info |
| 10.000000 | 192.168.192.128 | 192.168.192.1 | DNS | 110 Standard query response 0x3769 TXT 9C0AjgCeBA.dnsc2.
| 20.005658 | 192.168.192.1 | 192.168.192.128 | DNS | 85 Standard query 0x6247 TXT 9C0AjwCfBA.dnsc2.test.com |
| 30.507615 | 192.168.192.128 | 192.168.192.1 | DNS | 110 Standard query response 0x6247 TXT 9C0AjwCfBA.dnsc2.test.com |
| 40.511193 | 192.168.192.1 | 192.168.192.128 | DNS | 85 Standard query 0xf222 TXT 9C0AkACgBA.dnsc2.test.com |
             5 1.012725
                                        192.168.192.128
                                                                              192,168,192,1
                                                                                                                   DNS
                                                                                                                                    110 Standard query response 0xf222 TXT 9C0AkACgBA.dnsc2.
85 Standard query 0x515d TXT 9C0AkQChBA.dnsc2.test.com
                                       192.168.192.1
192.168.192.128
                                                                                                                                    110 Standard query response 0x515d TXT 9C0AkQChBA.dnsc2.
             7 1.518848
                                                                              192.168.192.1
                                                                                                                   DNS
              8 1.525815
                                        192.168.192.1
                                                                              192.168.192.128
                                                                                                                                      85 Standard query 0xa576 TXT 9C0AkgCiBA.dnsc2.test.com

    Wireshark - 分组 1 - dns2tcp_02.pcap

       Frame 1: 110 bytes on wire (880 bits), 110 bytes captured (880 bits)
Encapsulation type: Ethernet (1)
Arrival Time: Nov 15, 2020 21:03:15.636625000 中国标准时间
                [Time shift for this packet: 0.0000000000 seconds]
Epoch Time: 1605445395.636625000 seconds
                [Time delta from previous captured frame: 0.000000000 seconds]
[Time delta from previous displayed frame: 0.000000000 seconds]
               [Time since reference or first frame: 0.000000000 seconds]
Frame Number: 1
Frame Length: 110 bytes (880 bits)
                Capture Length: 110 bytes (880 bits)
                [Frame is marked: False]
                [Frame is ignored: False]
```

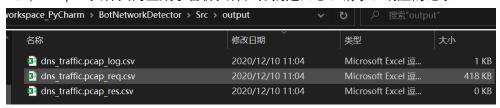
#### 2、执行文件检测功能,输出结果如下:

#### 3、列出所有主机和受害主机,成功检测出来

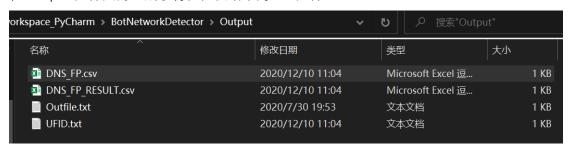


## 3.2.2 查看生成的 csv 备份文件

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:

Α	В	С
1	PCAP Processing Started at	04:24.0
2	Processing completed at	04:24.5
3	Total number of Packets Processed	9474
4	Total number of DNS Query	4737
5	Total number of DNS Responses:	4737
6	Total number of Unknown Response F	0
7	Total number of Failed Responses	0
8	Total Time taken	00:00.4

#### (2) request:

	Α	В	С	D	E	F	G	Н	1
1	25159	192.168.192.1	9C0AjwCfE	4	16	25	#######	192.168.1	92.128
2	61986	192.168.192.1	9C0AkACg	4	16	25	#######	192.168.1	92.128
3	20829	192.168.192.1	9C0AkQCh	4	16	25	#######	192.168.1	92.128
4	42358	192.168.192.1	9C0AkgCiE	4	16	25	#######	192.168.1	92.128
5	2828	192.168.192.1	9C0AkwCjI	4	16	25	#######	192.168.1	92.128
6	24400	192.168.192.1	9C0AlACkI	4	16	25	#######	192.168.1	92.128
7	51473	192.168.192.1	9C0AlQCIE	4	16	25	#######	192.168.1	92.128
8	15166	192.168.192.1	9C0AlgCm	4	16	25	#######	192.168.1	92.128
9	21884	192.168.192.1	9C0AlwCn	4	16	25	#######	192.168.1	92.128
10	60024	192.168.192.1	9C0AmAC	4	16	25	#######	192.168.1	92.128
11	8784	192.168.192.1	9C0AmQC	4	16	25	#######	192.168.1	92.128
12	8233	192.168.192.1	9C0AmgC	4	16	25	#######	192.168.1	92.128
13	59509	192.168.192.1	9C0AmwC	4	16	25	#######	192.168.1	92.128
14	64768	192.168.192.1	9C0AnACs	4	16	25	#######	192.168.1	92.128
15	43304	192.168.192.1	9C0AnQCt	4	16	25	#######	192.168.1	92.128
16	44912	192.168.192.1	9C0AngCu	4	16	25	#######	192.168.1	92.128
17	3365	192.168.192.1	9C0AnwCv	4	16	25	#######	192.168.1	92.128
18	57907	192.168.192.1	9C0AoACv	4	16	25	#######	192.168.1	92.128
19	6247	192.168.192.1	9C0AoQC	4	16	25	#######	192.168.1	92.128
20	61036	192.168.192.1	9C0AogCy	4	16	25	#######	192.168.1	92.128

(3) response:

#### 为空,符合实际要求

(4) 特征值记录:



#### (5) 结果:



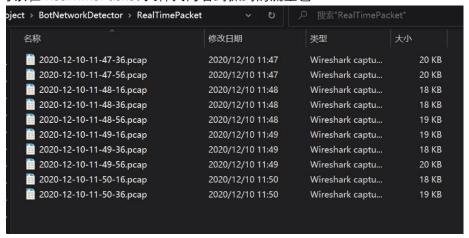
## 3.3 实时检测

## 3.3.2 准备工作

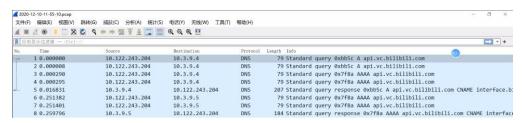
1、启动系统,并启动实时检测模块:

```
[cmd]
[1]实时流量监控
[2]pcap文件检测
[3]退出系统
请输入命令(1-3的某个整数): > 2 1
正在进行实时dns流量检测(输入exit退出)...
```

2、可以在 RealTimePacket 文件夹内看到抓到的流量包:



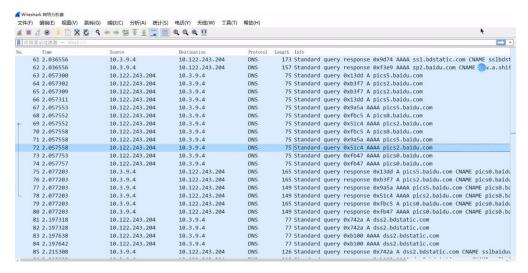
#### 打开查看:



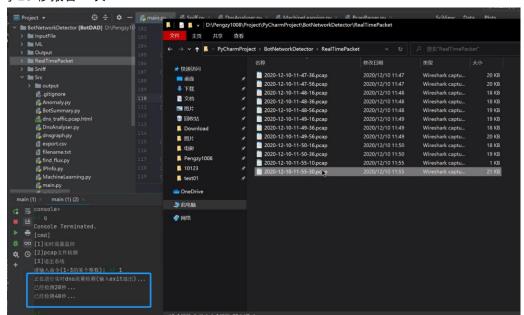
## 3.3.2 正常上网报告正常

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:

```
File Actions Edit View HMG

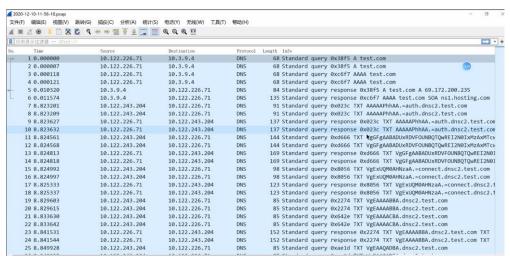
Debug queue.c:642 Packet [718] decoded, data_len 0
Debug queue.c:653 diff = 17
Debug queue.c:300 Flushing outgoing data
Debug queue.c:300 Flushing outgoing data
Debug queue.c:653 diff = 17
Debug queue.c:654 Packet [720] decoded, data_len 0
Debug queue.c:653 diff = 17
Debug queue.c:654 Packet [721] decoded, data_len 0
Debug queue.c:654 Packet [721] decoded, data_len 0
Debug queue.c:654 diff = 17
Debug queue.c:654 Add resource sst:127.0.0.1 port 25
22:55:54 : Debug options.c:97 Add resource sst:127.0.0.1 port 25
22:55:54 : Debug options.c:97 Add resource smp:127.0.0.1 port 25
22:55:54 : Debug options.c:97 Listening on 10.122.226.71:53 for domain dns
c2.test.com
Starting Server v0.5.2...
22:55:54 : Debug main.c:134 Chroot to /tmp
O3:55:54 : Debug main.c:134 Chroot to /tmp
Change to user nobody
```

#### (2) 客户端:

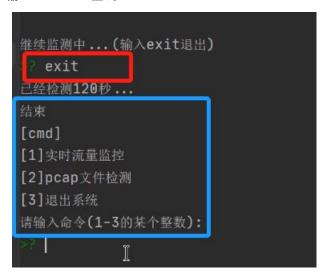
#### 2、检测一会后,下一次报告显示发现了僵尸主机:



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:

```
[cmd]
[1]实时流量监控
[2]pcap文件检测
[3]退出系统
请输入命令(1-3的某个整数): > 2
请将待扫描的pcap文件放入文件夹InputFile内(支持多个)
确认放入后请输入1>? 1
```

### 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:

```
l/L - ListBotHosts/ListAllHosts m - Save Map p - plot d/D - Di

2 L

All Hosts List:

序号 IP 请求数

1. 172.31.157.164 35

2. 172.31.55.40 21

3. 172.31.157.166 9715

4. 172.31.111.237 168

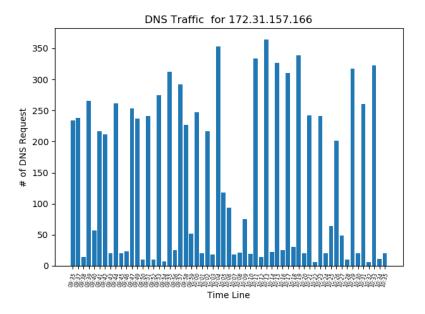
5. 172.31.157.161 181

6. 172 31 157 165 155
```

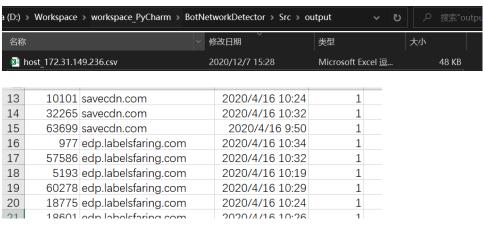
(2) 输入 I, 列出所有感染了僵尸网络的主机信息:

(3) 画出主机这段时间的流量图, p 命令:

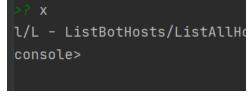
可以在 plot 区看到绘制的流量图:



(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:



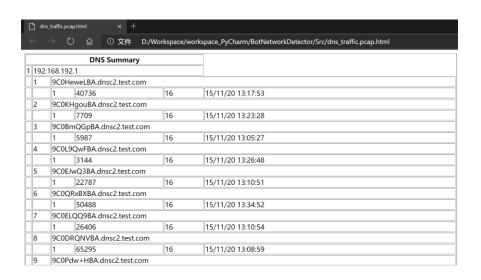
(5) 输入 x 保存为 csv, 保存本次扫描的结果



							,	IX.	L
nbr_requ	esnbr_uniqu	eavg_req_m	max_req_r	failed_cnt	ratio	nbr_countr	req_type	sum_url	sum_token
34	0 4734	0	0	0	0	0	0	0	0
7	nbr_requ 734								

(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