



# 物联网 DDoS 僵尸网络 C&C 通信流量检测分析研究

贺余盛

(国家计算机网络应急技术处理协调中心江西分中心, 江西 南昌 330038)

**摘要:** 物联网 DDoS 僵尸网络 C&C 通信检测是识别 DDoS 僵尸网络的重要组成部分。通过分析僵尸网络 C&C 通信流量总结出 C&C 通信数据分组较小和 DNS 通信的周期性的特点, 从而提出基于网络通信流量的特性识别僵尸网络 C&C 通信的方法, 提升了僵尸网络 C&C 通信识别的查准率。

**关键词:** C&C 通信; DDoS; 僵尸网络; DNS; 数据分组

**中图分类号:** TP393

**文献标识码:** A

**doi:** 10.11959/j.issn.1000-0801.2019152

## Research on C&C traffic detection of DDoS BotNet in internet of things

HE Yusheng

Jiangxi Branch of National Internet Emergency Center, Nanchang 330038, China

**Abstract:** C&C communication detection used in IoT DDoS BotNet is an important part to identify DDoS BotNet. By analyzing C&C communication traffic in the BotNet, the characteristics of the smaller C&C communication data packet and the periodicity of DNS communication were summarized. Based on these characteristics of the network communication traffic, a method to identify C&C communication in the BotNet was proposed, which helped to improve the accuracy rate of identifying C&C communication in the BotNet.

**Key words:** C&C communication, DDoS, BotNet, DNS, data packet

### 1 引言

随着信息安全技术的不断发展,物联网 DDoS 僵尸网络现在已成为信息安全领域内最危险的安全威胁之一。在 F5 实验室最新发布的物联网安全报告中,2018 年 1—6 月期间,全球 IoT (internet of things, 物联网) 设备受攻击的数据,涵盖物联网设备使用的主流服务和 20 个端口的分析数据,

其中,巴西、中国、日本、波兰和美国是主要的攻击来源国。黑客在非法入侵大量物联网设备之后,通过网络控制节点(C2),利用物联网设备形成“僵尸大军”,以此作为攻击跳板,聚集大量网络资源发起有针对性的大规模 DDoS 攻击,给个人或集体造成巨大的损失。为对物联网 DDoS 僵尸网络的有效打击,需对现有僵尸网络进行快速检测和准确追踪<sup>[1]</sup>。本文从物联网 DDoS 僵尸网络检测

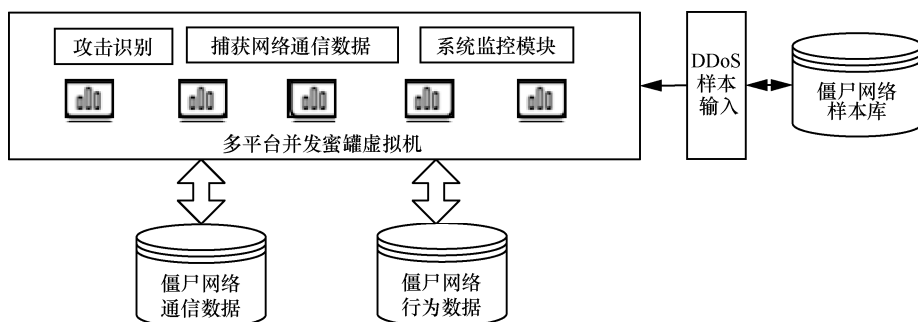


图1 僵尸网络通信数据捕获示意图

研究实际出发,重点从僵尸网络中的 C&C 通信流量的采集及检测两个方面进行研究分析。

## 2 僵尸网络 C&C 通信流量采集

对物联网 DDoS 僵尸网络 C&C 的通信流量的采集是物联网 DDoS 僵尸网络 C&C 通信检测的基础。通信流量的采集可通过两种方式实现:一是通过构建高低交互蜜罐获取;二是在网络的核心路由器或交换机上采用镜像或分光的方式采集互联网流量。在此,以构建高低交互蜜罐实现 C&C 通信流量采集为主进行介绍。图 1 为僵尸网络通信数据捕获示意图,通过高低交互蜜罐的方式捕获僵尸网络产生的通信数据和行为数据。

如图 1 所示,僵尸网络通信数据的捕获基于多平台的虚拟技术实现,即以 KVM (kernel-based virtual machine) 技术为基础的高并发 Wine 环境<sup>[2]</sup>。KVM 是内核里的虚拟机,操作简便,仅需要加载相应的 KVM 模块即可在后台待命。Wine 是类 UNIX 系统下运行微软 Windows 程序的“兼容层”。在 Wine 中运行的 Windows 程序,如同运行原生 Linux 程序,不会存在模拟器的性能问题。在操作系统的基础上,部署 libpcap 抓取分组程序实现流量采集,采集的内容包括进出流量、进程、文件等系统资源,同时,还可根据不同的攻击场景需要,安装相应的应用程序,伪装成个人电脑或服务器,处理针对不同应用程序的攻击报文。僵尸网络流量采集的具体流程如图 2 所示。

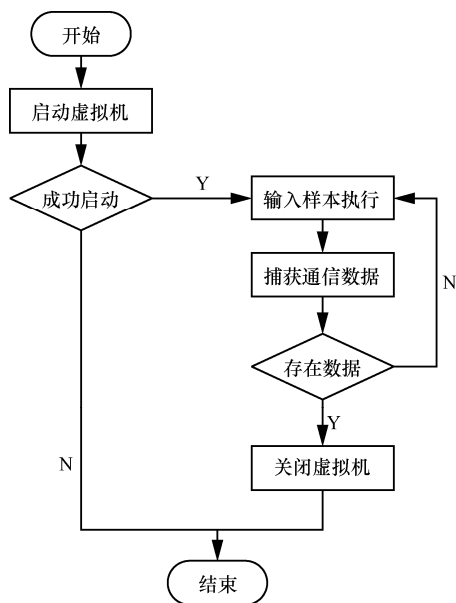


图2 僵尸网络通信数据采集流程

(1) 访问僵尸样本数据库,获取可执行的僵尸程序样本;

(2) 启动虚拟机,将可执行的僵尸样本放置在虚拟机中;

(3) 运行僵尸样本,捕获并存储僵尸网络的通信数据并获取僵尸网络通信所对应的端口号;

(4) 将捕获的相关通信数据上传到数据库,关闭虚拟机。

## 3 僵尸网络 C&C 通信流量检测分析

### 3.1 C&C 通信特征分析

本文从僵尸网络的通信流量数据分组长度及 DNS 请求的周期性两个方面进行特征分析。



### (1) 僵尸网络通信数据分组较小

僵尸网络被控节点会不间断地向 C&C 服务器发送查询或连接的请求指令, 该请求信息由多个类似的 TCP 或 UDP 数据分组构成, 且数据分组规模较小。在实际研究中发现, 僵尸网络通信数据分组具有通信数据量小且连接时间短的特点, 这符合僵尸网络通信隐蔽性的需求, 因此, 既要保证通信连接的可信, 又要尽可能将其隐藏。由此, 在僵尸网络的全生命周期中, 僵尸网络的通信需实时存在, 因此, 可将僵尸网络通信的数据分组规模小作为检测僵尸网络的一个重要特征。

通过阅读参考文献[3]及相关的流量数据分类统计分析, 不同类型的数据分组的大小分布存在一定的差异性, 在此, 将 Spyeye、Storm 两类僵尸与正常的网络通信数据分组进行对比, 结果如图 3~图 5 所示。

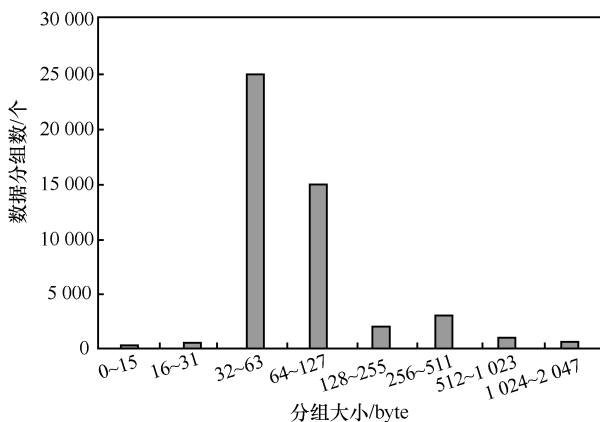


图 3 Spyeye 数据分组大小统计

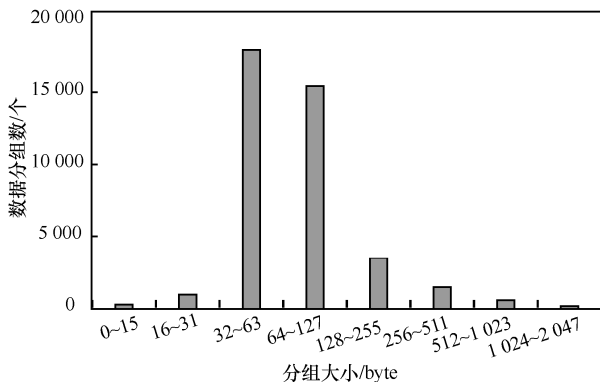


图 4 Storm 数据分组大小统计

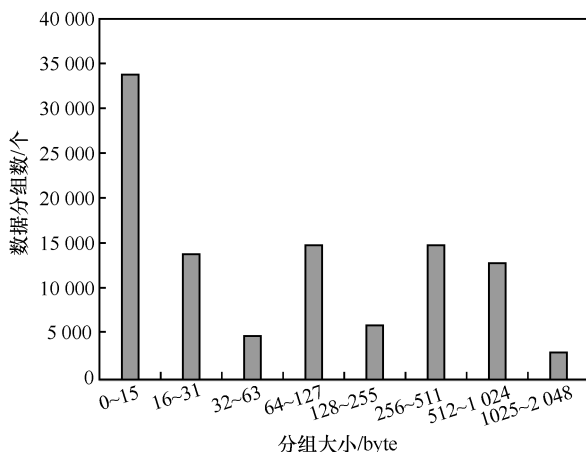


图 5 正常访问数据分组大小统计

通常情况下, 将长度为 32~255 byte 的数据分组定义为小数据分组, 且经过统计分析发现小数据分组的主要作用在于各对端的通信连接, 基本上与实际的通信内容数据传输关系较小。通过对物联网 DDoS 僵尸网络通信数据分组的分析, 流量较小的数据分组在数量、报文长度、发送频率等方面与正常的互联网访问存在较大的差异, 因此, 可将此作为僵尸网络 C&C 通信的特征。

### (2) 僵尸网络通信 DNS 请求的周期性

僵尸网络被控主机为保证实时接收 C&C 服务器发送的控制指令, 需定期向 C&C 服务器发送 HTTP 请求, 通过 DNS 查询保证与 C&C 服务器的连接, 因此, 可将僵尸网络被控端对 DNS 的周期性访问作为僵尸网络被控主机的识别方式。

根据对僵尸网络样本的长期分析可知, 僵尸网络程序的设计者已将 C&C 服务器的 IP 地址由原先的硬编码至程序体内转变为请求域名配置的方式, 此外, 为避免僵尸网络控制端的打击, C&C 服务器采用了 DDNS (dynamic domain name server, 动态域名服务)、DGA (domain generate algorithm, 域名生成算法) 等新技术实现 IP 地址或域名的周期性变化。

(1) 利用 DNS 技术能够动态改变 C&C 服务器的 IP 地址, DNS 请求信息呈现出 TTL (time to live, 生存时间) 值较小及查询频繁的特征<sup>[4]</sup>。

(2) 利用 DGA 技术能够为 C&C 服务器生成数以千计的域名, 然而, 只有极少部分的域名指向真实的 C&C 服务器, 因此, 僵尸网络被控节点会产生大量的 DNS 请求进行 C&C 服务器地址查询。表 1 展示了某僵尸网络样本向 C&C 服务器发送 DNS 信息查询的请求信息。

表 1 僵尸网络被控节点 DNS 查询信息

| 序号  | 查询域名             | 时间                  | 时间间隔/s |
|-----|------------------|---------------------|--------|
| 1   | 35732.camdvr.org | 2018-07-24 12:58:48 | 58.21  |
| 2   | 104.168.94.130   | 2018-07-24 12:59:46 | 57.34  |
| 3   | 35732.giize.com  | 2018-07-24 13:00:43 | 61.31  |
| 4   | 35732.ooguy.com  | 2018-07-24 13:01:44 | 53.44  |
| 5   | 170.130.15.198   | 2018-07-24 13:02:37 | 58.24  |
| 6   | 35732.jdkfi.com  | 2018-07-24 13:03:35 | 57.17  |
| ... | ...              | ...                 | ...    |

根据表 1 所示, 僵尸网络被控节点向 C&C 服务器发送 DNS 查询的时间间隔大约为 58 s, 具有一定的周期性。由此, 僵尸网络被控节点产生 DNS 请求信息具有两个特点: 一是具有较高的频繁性;

二是具有一定的周期性。

### 3.2 C&C 通信检测方案

物联网僵尸网络 C&C 通信信道是维系僵尸网络的重要通信渠道, 也是识别和检测物联网僵尸网络的重要途径。通过深入地分析 C&C 通信过程能够更清晰地了解僵尸网络间的通信及控制机制。目前, C&C 通信检测方式主要分为主动式和被动式 C&C 通信检测。本文采用高低交互混合的蜜罐的方式实现对互联网流量的采集及 C&C 通信检测, 从物联网僵尸网络的 C&C 监控中获取 C&C 服务器向各僵尸节点发送的控制指令, 主要包括传播的 payload、URL 等。由于主动式 C&C 通信检测是模拟被控节点, 所获取数据的时效性要明显高于通过探针方式获取的数据。

为减少不必要的报文数据分析, 在检测前需配置报文的过滤规则, 主要如下。

(1) 考虑到僵尸网络数据请求的发起方为被控节点, 因此, 忽略内部主机间的交互数据及外部主机主动发起的请求信息。

(2) 将部分合法域名加入白名单中。

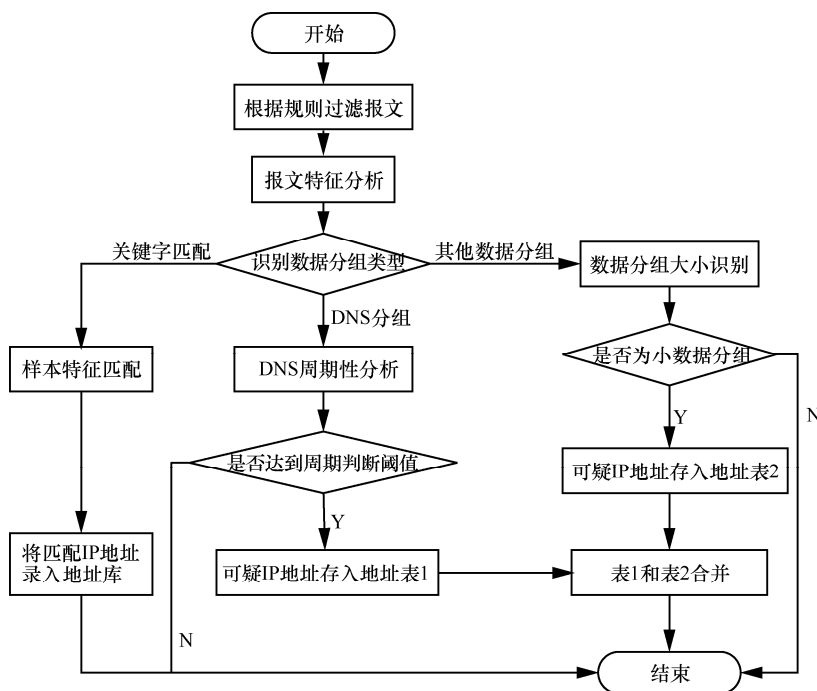


图 6 C&C 通信检测的流程



对于 C&C 服务器的检测主要通过 DNS 查询数据的周期性分析及数据分组的大小进行甄别,图 6 为 C&C 通信检测的流程。

## 4 检测结果与分析

### 4.1 检测数据集

为验证 C&C 服务器通信检测效果,下面对检测实验中所调用的检测数据集进行介绍。

利用单位已构建的僵木蠕监测系统检索已感染僵尸网络的被控主机的 IP 地址,然后,再按照已感染的 IP 地址段引流近 1 GB 数据流量,时间段为 2018 年 10 月 28—30 日。引流的数据流量中不仅包含僵尸网络的数据流量,还包含网页浏览、即时通信、在线游戏等正常互联网访问流量。表 2 统计了各个 IP 地址及对应的互联网通信的流量信息。

表 2 各 IP 地址对应的流量信息

| 序号  | IP 地址           | 大小/MB | 报文个数    |
|-----|-----------------|-------|---------|
| 1   | 223.84.192.134  | 121.7 | 82 417  |
| 2   | 218.64.91.93    | 89.4  | 71 425  |
| 3   | 223.83.210.72   | 11.24 | 9 245   |
| 4   | 113.195.161.105 | 191.5 | 125 932 |
| ... | ...             | ...   | ...     |

### 4.2 检测结果分析

本次检测以 HTTP 类的僵尸网络流量数据为主要检测对象。在引流的通信流量中,能够识别出的数据流共 8 992 条,其中,不仅包括僵尸网络的报文数据,还包括互联网正常访问数据。经过筛选和比对,共得到 1 399 条僵尸网络的访问记录,其统计结果如图 7 所示。

原有已知的感染僵尸网络的被控主机共有 1 102 条记录。根据僵尸网络检测结果,在获得的 1 399 条僵尸网络的访问记录中,与原有已知的感染僵尸网络的被控主机吻合的记录共 982 条,不吻合的记录为 417 条(可视为错误条数),具体见表 3。

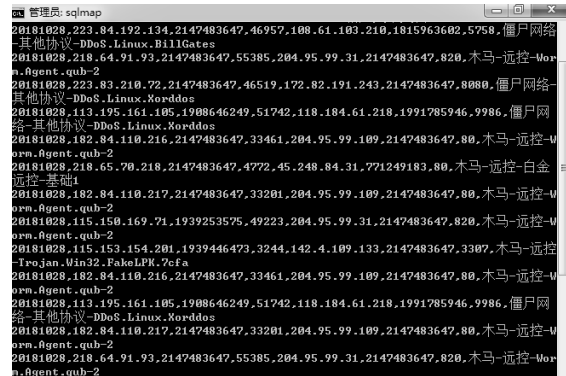


图 7 僵尸网络检测结果截图

表 3 僵尸网络检测结果统计

|    | 正确  | 错误  |
|----|-----|-----|
| 数量 | 982 | 417 |

主要采用查准率、查全率和综合评价 3 个指标判定僵尸网络检测方案的性能<sup>[5]</sup>,此次检测结果的查准率为 89.11%,查全率为 70.10%,综合评价为 79.61%。

现有的检测方案仅根据通信数据分组的大小进行单维度判断,共获得 1 156 条僵尸网络的访问记录,与原有已知的感染僵尸网络的被控主机吻合的记录共 665 条,不吻合的记录为 491 条(可视为错误条数),具体结果统计见表 4。

表 4 僵尸网络检测结果统计

|    | 正确  | 错误  |
|----|-----|-----|
| 数量 | 665 | 491 |

根据检测结果统计,查全率为 57.52%,查准率为 60.33%,综合评价为 58.93%。

从以上 3 个指标可以看出,本文所阐述的僵尸网络 C&C 通信检测方法在查准率方面较高,查询出的僵尸网络通信数据准确度较高,在查全率方面虽有提升,但还需进一步改进。

## 5 结束语

本文提出的 C&C 通信检测方法能够比较有效地检测引流中的僵尸主机与 C&C 服务器的通

信,虽然误报率较高,但能够较为全面地检测僵尸网络通信,再通过后续的僵尸网络特征匹配及威胁智能识别办法,能够有效地提升物联网僵尸网络通信的准确率,物联网僵尸网络 C&C 通信检测是一个多维度相互协同的过程,在后期的研究中,将采用卧底养殖方式监听物联网僵尸网络中 C&C 服务的最新攻击指令和流量数据。

### 参考文献:

- [1] 周大,刘亚.一种识别 DDoS 僵尸网络通信协议的方法及装置:CN201410196838.2[P].2017-02-22.  
ZHOU D, LIU Y. Method and device for identifying DDoS BotNet communication protocol: CN201410196838.2[P]. 2017-02-22.
- [2] 胡渊博,程静,陈渤.基于多特征的 SMS 僵尸网络检测模型[J].中国电子科学研究院学报,2018(4):384-388.  
HU Y B, CHENG J, CHEN B. Multi-feature based SMS BotNet detection model[J]. Journal of China Academy of Electronics and Information Technology, 2018(4): 384-388.
- [3] 李传煌,孙正君,袁小雍,等.基于深度学习的实时 DDoS 攻

击检测[J].电信科学,2017,33(7):53-65.

LI C H, SUN Z J, YUAN X Y, et al. Real-time DDoS attack detection based on deep learning[J]. Telecommunications Science, 2017, 33(7): 53-65.

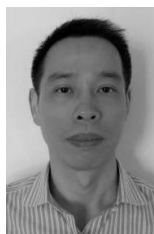
- [4] 滕建,陈骏君,赵英.基于用户网络行为模型的 DDoS 攻击检测[Z].2018.

TENG J, CHEN J J, ZHAO Y. DDoS attack detection based on user network behavior model[Z]. 2018.

- [5] 任斌,周亦敏.大规模网络环境下僵尸网络检测研究[J].软件导刊,2017,16(10):189-191.

REN B, ZHOU Y M. Research on BotNet detection in large-scale network environment[J]. Software Guide, 2017, 16(10): 189-191.

### [作者简介]



贺余盛(1971—),男,国家计算机网络应急技术处理协调中心江西分中心副主任兼总工程师、高级工程师,主要从事互联网及数据通信相关系统建设、维护、经营管理和网络安全监管工作,主要研究方向为互联网网络安全应急响应。