

物联网僵尸网络的恶意域名检测技术研究

李雪妍,陈 伟,杜俊雄

(南京邮电大学 计算机学院,江苏 南京 210023)

摘 要:随着物联网智能设备的普及,所带来的社会安全隐患也越来越多。正如 2016 年爆发的 Mirai 恶意软件,它正是由物联网智能设备中漏洞的入侵和渗透形成的一个大型僵尸网络。其变种内置的域名生成算法大大增强了自身的健壮性,极大程度上延长了其自身的生命周期。域名系统作为互联网重要资源,也带来了很大的安全威胁。文中分析研究了现有的恶意域名识别技术,并提出一种基于信誉评分体制的全新检测系统。选取了基于域名维度与 IP 维度的特征集,同时设计并实现了异常值自动评分算法,算法可以自动选择最可疑的恶意域名事件且无需已标记数据集。实验结果表明,将文中采用的自动评分技术与标准异常检测技术相比较,误报率低至 0.003%,该系统的准确率比标准检测技术平均提升 5 ~ 10 倍。

关键词:物联网;僵尸网络;恶意域名;自动评分算法;信誉特征

中图分类号:TP301

文献标识码:A

文章编号:1673-629X(2019)08-0113-06

doi:10.3969/j.issn.1673-629X.2019.08.022

Research on Malicious Domain Name Detection Technology in IoT Botnet

LI Xue-yan, CHEN Wei, DU Jun-xiong

(School of Computer, Nanjing University of Posts and Telecommunications, Nanjing 210023, China)

Abstract: With the popularization of Internet of things devices, there exists more and more security risks. Like the Mirai malware outbreak in 2016, it is a large Botnet created by the intrusion and penetration of vulnerabilities in smart devices in the Internet of things. The Mirai variant built-in domain name generation algorithm greatly enhances its robustness and extends its life cycle. As an important resource of Internet, DNS (domain name system) also brings great security threat. We analyze the existing malicious domain name recognition technology, and propose a new detection system based on the credit rating system. The feature set based on domain name dimension and IP dimension is selected, and the outliers automatic scoring algorithm is designed and implemented, which can automatically select the most suspect malicious domain name events through unmarked datasets. The experiment shows that compared with the standard abnormal detection technology, the false alarm rate of the proposed automatic scoring technology is as low as 0.003%. The accuracy of the system is 5 ~ 10 times higher than that of the standard detection technology.

Key words: IoT; Botnet; malicious domain name; automatic scoring algorithm; reputation

0 引言

近年来,随着多种接入网络以及人工智能等计算机技术的快速发展,国际电信联盟电信标准化组织发表了研究报告《泛在传感器网络》。报告中提出,传感器网络在日常生活中已经无处不在,它可以通过各种模式各种方式存在。物联网可以把感应器、处理器和无线通信模块嵌入或装备到隧道、公路、电网、铁路、桥梁、建筑等各种物体中,使它们相互连接,构成物联网^[1]。

物联网作为当前互联网的主要组成部分,是通过

互联网连接各类智能设备的。恶意域名作为互联网中的重要安全威胁,同样影响物联网中的各类设备。域名解析系统是互联网中的基础设施,主要负责域名解析,将域名转换为 IP 地址。域名解析系统本身具有脆弱性,容易受到攻击,易成为恶意网络行为攻击的隐蔽通道^[2]。通过恶意域名,可以直接或间接完成分布式拒绝服务(DDOS)等恶意网络行为,并可以为后续进一步提高系统控制权提供踏板和条件^[3]。

目前,分布式拒绝服务攻击几乎都是由僵尸网络发动^[4],并且除了传统的 PC 和服务器作为僵尸主机

收稿日期:2018-09-27

修回日期:2019-01-30

网络出版时间:2019-03-27

基金项目:国家自然科学基金(61602258,61702283)

作者简介:李雪妍(1993-),女,硕士研究生,研究方向为网络安全;陈 伟,博士,教授,CCF 会员(E200025391M),研究方向为网络安全。

网络出版地址: <http://kns.cnki.net/kcms/detail/61.1450.TP.20190327.1629.060.html>

之外,智能设备也逐渐成为攻击者的重点关注对象,成为新的僵尸节点。Mirai 恶意程序正是通过对物联网智能设备^[5]中的漏洞进行入侵渗透,从而形成的庞大的僵尸网络。360 网络安全研究院在 2017 年 11 月发布了 2 个新型感染载体(TCP 端口 7547/TCP 端口 5555),可以用于传播 Mirai 恶意软件。360 安全研究院的 Ye Genshen 通过设置一些蜜罐,到目前为止已经从 6 个被托管的服务器上获取了 53 个独立恶意样本。在分析其中一个新样本时,发现一些类似 DGA(domain generation algorithm,域名生成算法)的代码,并猜测变种中可能包含有 DGA 功能。

综上所述,恶意域名对于物联网造成的危害不容忽视。文中的主要贡献在于:针对 DNS 查询数据不同阶段进行特征分析,从域名特征、查询次数等不同角度提出 DNS 查询两个维度的七个不同特征;提出基于自动评分算法的恶意域名异常检测系统,设计异常值自动评分算法,并实时生成异常警报,可以从未标记的数据集中自动选择可疑的恶意域名;利用与已知恶意域名集合的域名共性特征属性,计算感染主机查询过的其他域名与已知恶意域名集合的相似程度来推断恶意域名,并且通过离线以及实时数据进行系统评估检测,误报率在 0.003% 以下,在安全领域有良好的应用前景。

1 相关工作

目前很大一部分恶意软件检测的研究是基于 DNS 协议展开的,通过 DNS 数据分析恶意软件有明显的优势。目前绝大多数恶意软件都是通过 DNS 解析其网络控制服务器地址。DNS 数据检测与 C&C 协议不同,可以用于检测加密通信的恶意软件。同时,DNS 流量只占网络总流量的很小一部分,基于 DNS 数据分析的开销远远低于分析网络中所有流量的开销,因此可以用于超大型网络系统。如今通过对僵尸网络恶意域名的特征进行分析提取,利用属性判断进行恶意域名识别的方式更为通用。柳厅文等提出基于多元属性的恶意域名识别技术^[6],分析恶意域名的多元属性来识别恶意域名。周昌令^[7],Antonakakis^[8]等分别采用不同的机器学习算法,提取多种恶意域名特征和属性,建立恶意域名识别模型,从而进行识别。Engin Kirda 等实现的 EXPOSURE 系统^[9],对域名请求时间特征、TTL 特征、回答特征和域名特征进行机器学习分类,可以对恶意软件以及钓鱼网站等恶意域名进行检测识别。

2 自动评分异常检测系统设计

鉴于传统检测技术的局限性,文中引入一种简单

而通用的技术。该系统采用的数据集无需进行标记,并且可以进行可疑事件自动选择。之前的一些学术研究尝试将域名测试特征结合到恶意软件攻击检测的识别中,然而这些系统的误报率为 1% 或更高^[10],当数据量大时这将导致数百万的误报,这个误报的绝对数量在实际使用中是难以接受的。

文中设计的检测器由三个阶段组成(见图 1):特征提取阶段、自动评分阶段和实时警报生成阶段。从概念上说,该设计引入了两个关键技术,使探测器能够检测到恶意域名攻击,同时比以前的方法降低 5~10 倍的实际误报率。

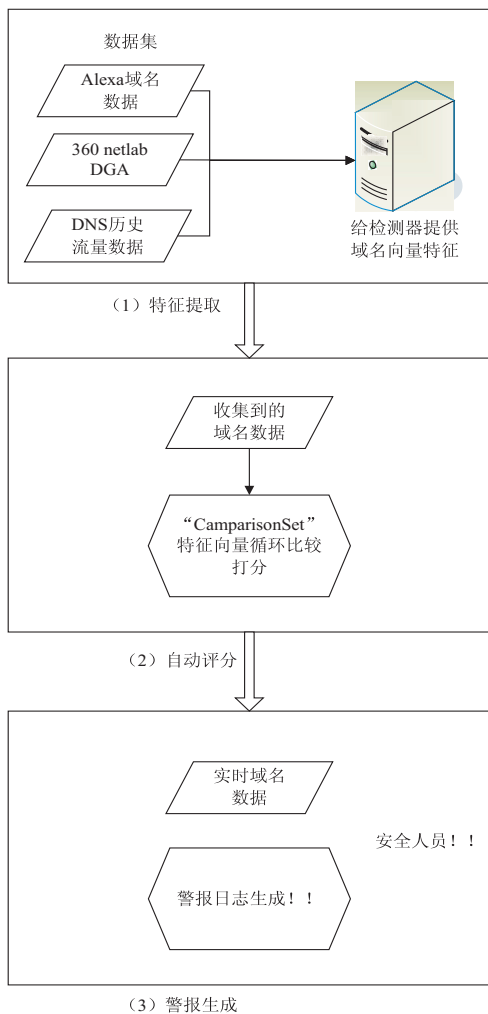


图 1 系统流程

2.1 实验环境部署及数据采集

实验采集的良性样本从 Alexa 网站靠前排名域名中进行获取。Alexa 每天爬取互联网上超过 1 TB 的信息,其中包括数十亿网址链接,并对其按照热门程度进行排序。研究人员常常把 Alexa 排名靠前的域名作为良性样本进行处理。恶意样本来自于 360 data.netlab 中已公布的 GDA 数据,360 data.netlab 公示了近 10 年 39 个 DGA 家族,表 1 所示为部分样本数,恶意样本数约为 5 G。

表 1 DGA 生成域名

conficker	cryptolocker	gameover
gfedo. info	nvjwoofansjbh. ru	14dtuor1aubbmjhgup7915tlinc. net
.....
ydqtkptuwsa. org	eqmbcmgemghxbcj. co. uk	uhjmkml1i7oih11i3wxl71kcf7x6. org

在实验环境中采用树莓派模拟物联网设备使用场景,复现 Mirai 恶意软件攻击,收集其恶意软件攻击场景下的 DNS 流量包。Mirai 恶意软件主要攻击步骤如下:

- (1)IoT 设备通过下载恶意软件感染 Mirai 病毒;
- (2)IoT 设备与 DNS 服务器进行交互,目的是得到 CNC 服务器的 IP 地址;
- (3)IoT 设备与 CNC 服务器进行周期性交互;
- (4)IoT 设备扫描存在弱密码的其他 IoT 设备;
- (5)扫描成功后汇报给 Loader 服务器;
- (6)Loader 对此设备开始进行攻击。

为了复现 Mirai 恶意软件攻击流程,实验室采用树莓派模拟 IoT 设备,通过主机 1 连接至相应的局域网,手机和主机 2 作为用户使用的控制器接入点,通过连接至无线局域网后和树莓派进行通信。针对主机 1,在控制器和树莓派进行通信时,通过 Linux 系统下进行相应 DNS 数据包抓取,以便后续进行分析。

2.2 特征选取

检测器使用域名解析阶段的两个维度进行特征选

择,分别为域名信誉维度^[11]以及 IP 信誉维度。使用一共包含七个标量值的特征向量,其中五个用于域名信誉,两个用于 IP 信誉。

2.2.1 域名信誉特征

(1)元音字母个数占比。

根据 Alexa 域名数据表统计分析图表可以看出,正常域名元音字母个数占比远大于域名生成算法随机生成的恶意域名。这是因为正常域名在命名时,为了方便记忆域名,通常带有“好读”的属性。以元音字母比例为横轴,域名长度为纵轴,绘制以下图表。根据图 2 也可以清晰地看出,良性域名和恶意域名有很明显的聚类效果。

(2)去重后字母数字个数占比。

去重后数字字母个数指的是域名中去重后的字母数字个数与域名长度所计算出的比值。从某种程度上反映了域名字符组成的统计特征。获取僵尸网络 DGA 以及 Alexa 域名数据,分别计算去重后字母数字个数占比,良性域名去重后字母数字个数占比较高。

(3)Jaccard 系数。

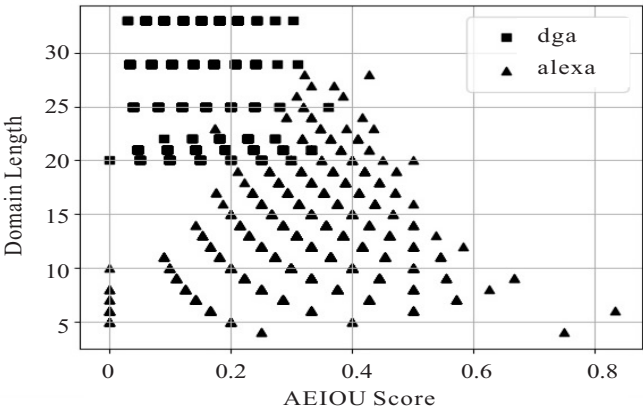


图 2 元音字母占比数

Jaccard 系数又称 Jaccard 相似系数 (Jaccard similarity coefficient),用于比较样本集之间的相似性与差异性^[12]。其定义为两个集合交集与并集元素个数的比值,表示为:

$$J(A,B)=\frac{|A\cap B|}{|A\cup B|}$$

(1)

在该检测器中基于 2-gram 系数计算两个域名之间的 Jaccard 系数,Jaccard 系数越大,样本相似度越高。获取僵尸网络 DGA 以及 Alexa 域名数据,分别计算 Jaccard 系数,样本相似性较大的 DGA 恶意域名

Jaccard 系数较高,良性域名 Jaccard 系数较低。

(4)HMM 系数。

隐马尔可夫模型(hidden Markov model,HMM)是统计模型,用来描述一个含有隐含未知参数的马尔可夫过程。对于 HMM 模型,假设 Q 是所有可能的隐藏状态的集合, V 是所有可能的观测状态的集合,即: $Q=\{q_1,q_2,q_3\}$, $V=\{v_1,v_2,v_3\}$ 。其中, N 是可能的隐藏状态数, M 是所有可能的观察状态数。

正常人取域名的时候,都会偏向于选取常见的几个单词的组合,抽象成数学可以理解的语言。因此以

常见英文单词建立 HMM 模型,正常域名的 HMM 系数偏高,僵尸网络 DGA 域名是随机生成的,所以 HMM 系数偏低。

(5)完全限定域名(fully qualified domain name, FQDN)访问次数。

该特征从域名访问次数进行可疑程度判断,即在流量中显示访问次数越少的域名,其可疑程度越高。在采集的 DNS 数据集中统计每一个域名访问频次,若统计次数越高,可疑程度越低。

2.2.2 IP 信誉特征

(1)Cname 返回值信息熵。

引入香农提出的“信息熵”的概念,基于域名解析记录来建立信息熵,熵值用来评估域名稳定性^[13]。在对域名发起查询时,将返回信息根据事件排列,域名解析记录的结果一般以“.”或者空格进行分隔,旨在计算同字段序列的稳定性。例如在 A 记录中,对域名的解析结果返回的是 IP 地址,IP 地址是一个以“.”划分的四个字段,每个字段中是一个范围在 0 ~ 255 的数值。

在 Cname 记录中,对域名的解析结果返回的也是域名,表示多个域名指向同一个服务器 IP,即表示与查询域名共享一个 IP 地址的域名,其形式符合域名的规范^[14]。文中将信息熵形式化定义如下:

SeriesSet = { Info_{1,1}, Info_{1,2}, ..., Info_{n,k} }, 其中 Info_{n,k} 表示第 n 次 DNS 请求返回的第 k 个信息。这是由于一次解析返回的内容存在多种情况,若只存在单个返回内容,则可手动设置 k 值为 1,同时,对于多次返回相同的信息,集合中不会进行删除。然后分别计算每个集合中信息的概率:

$$P(B) = \frac{|\{B \mid \forall B \in \text{SeriesSet}\}|}{M} \quad (2)$$

其中, M 为该序列集合中所有的信息个数。

最终将该序列的信息熵定义如下:

$$\text{EntropyOfInfo} = - \sum P(B) \log P(B) \quad (3)$$

整个域名的信息熵表示为:

$$\text{InfoValue} = \sum \text{EntropyOfInfo} \quad (4)$$

通过域名信息熵评估域名稳定性,域名信息熵值越高,在时间范围内其变化程度越大,若其信息熵值越小则域名越稳定。

(2)域名对应的 IP 个数。

DNS 被广泛应用于各种常见网络,主要提供域名解析功能。由于域名注册的灵活性和域名对应 IP 地址的动态性,黑客常采用域名而非传统 IP 地址的形式标识大量参与网络攻击行为的服务器。所以在该部分特征中,文中选取域名对应的 IP 指标作为评分特征,

域名对应的 IP 个数越高即可疑程度越高。

2.3 异常值自动评分算法

鉴于传统检测技术的局限性,引入了一种简单适用的技术,用于从未标记的数据集中自动选择最可疑的事件,称之为基于自动评分算法的异常检测系统。该系统通过比较每个事件相对于所有其他事件的可疑程度来对所有事件进行排名。一旦所有事件被排序,异常评分系统只选择 N 个最可疑(排名最高)的事件,其中 N 是安全团队的预警期望值^[15]。

算法:异常值自动评分与警报生成算法。

GetSubScore($E'D_1, E'D$):

1:for each $E'D_1$ in $E'D$ do:

2:if $E'D_1$ is more suspicious than $E'D$:

3:Increment score of $E'D_1$ by one

GetTotalScore(E, L)

1:for each E in L do:

2:if E is more suspicious than X in at least β dimension

3:Sum SubScore(E)

TopToFileAlert(L (a list of events), N):

1:for each event E in L do:

2:GetTotalScore(E, L)

3:Sort L by each event's score

4:return the first N events from L with the highest scores

算法显示了使用基于信誉特征的异常自动评分和生成警报的过程。算法中使用特征向量 E 来识别每个待评分事件, L 表示待评分事件集合, $E'D_1$ 表示待评分事件 E 的某一个特征维度的子特征向量, $E'D$ 表示待评分事件 E 的所有特征向量集合。

根据上述特征计算 E 的特征向量,实验中选择域名以及 IP 两个维度特征,每一个维度依次有若干子维度特征,根据算法中 GetSubScore() 得出每一个域名特征的子分数;

将事件 E 与 L 中的每个事件根据特征维度分数进行横向对比,当且仅当域名维度中至少 β 特征(实验中 β 取 80%,即域名维度 7 个特征中至少有 5 个特征并且 IP 维度至少满足 1 个特征)都比其他待比较事件特征可疑时,累加得出 E 的最终异常值得分,因此,得分越高表示该域名事件越可疑。选择超过 β 特征而非全部,是因为若匹配全部,攻击者任意避开其中一项即可能产生漏报的情况;

在对每一个事件进行评分排序后,算法将所有事件按其异常值得分进行排序,并最终输出 N 个得分最高的事件(N 是安全团队预警期望值)。同时,算法在此处添加异常处理操作,如若设置的预警值为 N , N 的异常得分与 $N+1$ 事件的异常得分相同时,即可让安

全人员选择是否包含 N 相同得分的所有事件。

2.4 实时监测架构

综合前述章节的讨论,在本节中详细说明如何利用检测器生成警报。首先,该检测器可以随时访问实时网络流量(如校园 DNS 流量等),并且可以获取到安全团队可接受的每日预警数。当收到一个域名请求时,检测器就会自动提取该事件的二维特征向量,并将其存放在域名事件索引表中。

在评分过程中,系统采用文中设计的批处理自动评分算法,在实时检测运行过程中,该检测器每天都会收集当天内所有的域名事件,并计算提取上文提出的特征向量,并在系统算法的 ComparisonSet 集合中与所有同维度的特征向量进行评分比较。在两个维度的每一个特征向量都评分完毕后,将针对安全团队可接受的警报值定时输出警报。当然,如果某一天中大批量

出现可疑的域名时,该异常检测系统会产出比之前的目标预算更多的警报数。但是此类情况在实际检测实验过程中的发生概率几乎为零。

3 实验结果评估与分析

实验主要利用离线数据以及实时数据进行评估检测。

3.1 离线数据检测

采用 Alexa top100000 域名与 2007 年 9 月至 2018 年 6 月 360 data. netlab 公布的僵尸网络主机域名对设计的检测器精度进行评估,数据量约为 10 G。利用文中提出的检测系统对每一个家族进行逐一检测,区分于僵尸网络 DGA 家族。表 2 列出了多个 DGA 家族进行检测的结果,在所有 DGA 样本数据检测中,准确率为 0.003%。

表 2 DGA 生成域名检测精度

DGA 家族	域名数	正确检测数	误报数
gspy	177	177	0
xshellghost	231	231	0
Chinad	232	232	0
suppobox	246	246	0
proslikefan	189	189	0
...
Dircrypt	135	135	0
Dyre	231	230	1
Emotet	309	309	0

此外,将自动评分系统与标准异常检测技术进行对比,采取同样的特征值,选择 LSTM(long short-term memory,长短期记忆网络)机器学习算法进行建模预测。选择总数据的 80% 为训练集,20% 为测试集进行建模运算,准确率最高值可达 99.8%,最低值为

80.1%,平均值为 91.2%。图 3 为文中提出的异常检测系统与标准异常检测技术 LSTM 的结果对比。可以清晰看出,使用文中的异常值自动评分算法的检测准确率明显高于 LSTM 算法,同时该系统对于不同 DGA 家族检测准确率的影响也比较小。

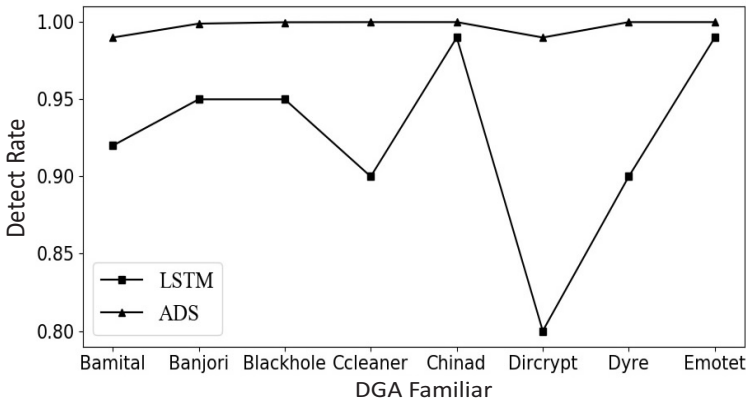


图 3 异常检测系统(ADS)与 LSTM 算法检测结果对比

3.2 实时数据检测

本节使用校园 DNS 流量对检测器实时性进行评估分析。针对实时性检测,在自动评分检测系统中利用 Tshark 实时抓包和 Pyshark 离线数据包分析模块,

对于实时网络流量进行分析。通过定时任务收集一天的网络数据包,并导入自动评分系统进行检测,当捕获到可疑事件发生时即触发一次警报。依据大数据经验,一天内校园网络中报警次数 30 次以内,视为安全

状态,超过 30 次但小于 100 次为可疑状态,大于等于 100 次为威胁状态。检测出的恶意域名通过后台数据库系统存储,并动态更新到已知恶意域名库中,以提高下一次自动评分系统的识别效率。

采集实验室历史数据,在测试电脑客户端输入已知恶意域名 972 个,通过文中提出的自动评分系统对其进行识别,历史数据已知恶意域名全部被成功检测,并且捕获到 8 个未被发现的新型恶意样本。

4 结束语

文中提出的自动评分系统能够达到高精度的检测准确率,通过在匿名数据库上的评估,异常值自动评分系统误报率低于 0.003%,检测出匿名数据库已标记恶意域名 972 个,未标记恶意域名 8 个。与文中的异常评分技术相比,在检测相同数据集时其他异常检测技术会产生 5~10 倍的误报率。但是该方法仍存在不足,在以下方面还需要进一步完善:

攻击者规避策略:检测器警报设置需要有数量上限设定,因此,如果在某一天内出现大规模恶意域名攻击,该检测器会出现漏报的可能性。为了解决该问题,安全人员在保留警报日志时,可以适当增加检测器的预警期望值。

另外,攻击者可能会学习良性域名的信誉特性。例如,攻击者可以手动生成恶意域名,匹配良性域名的行为特征,从而规避检测器特征匹配,降低自己的异常值得分而不被识别出。这种规避策略给攻击者带来了一定的成本与风险,因此在今后的工作中还需要对异常值自动评分算法进行完善。

参考文献:

- [1] 臧劲松. 物联网安全性能分析[J]. 计算机安全,2010(6): 51-52.
- [2] LIU Daiping, HAO Shuai, WANG Haining. All your DNS records point to us: understanding the security threats of dangling DNS records[C]//Proceedings of the 2016 ACM SIG-SAC conference on computer and communications security. Vienna, Austria: ACM, 2016:1414-1425.
- [3] NIRMAL K, JANET B, KUMAR R. Phishing - the threat that still exists[C]//International conference on computing and communications technologies. Chennai, India: IEEE, 2015:139-143.
- [4] 江 健, 诸葛建伟, 段海新, 等. 僵尸网络机理与防御技术[J]. 软件学报, 2012, 23(1): 82-96.
- [5] RONEN E, SHAMIR A. Extended functionality attacks on IOT devices: the case of smart lights[C]//2016 IEEE European symposium on security and privacy. Saarbrücken, Germany: IEEE, 2016: 3-12.
- [6] 张 洋, 柳厅文, 沙泓州, 等. 基于多元属性特征的恶意域名检测[J]. 计算机应用, 2016, 36(4): 941-944.
- [7] 周昌令, 陈 恺, 公绪晓, 等. 基于 Passive DNS 的速变域名检测[J]. 北京大学学报: 自然科学版, 2016, 52(3): 396-402.
- [8] ANTONAKAKIS M, PERDISCI R, LEE W, et al. Detecting malware domains at the upper DNS hierarchy[C]//Proceedings of the 20th USENIX conference on security. San Francisco, CA: USENIX Association, 2011: 27.
- [9] BILGE L, KIRDA E, KRUEGEL C, et al. EXPOSURE: finding malicious domains using passive DNS analysis[C]//Proceedings of the ISOC network and distributed system security symposium. San Diego, California, USA: [s. n.], 2011.
- [10] XIANG Cui, FANG Binxing, YIN Lihua, et al. Andbot: towards advanced mobile botnets[C]//Proceedings of the 4th USENIX conference on large-scale exploits and emergent threats. Boston, MA: USENIX Association, 2011: 11.
- [11] 刘 焱. Web 安全之机器学习入门[M]. 北京: 机械工业出版社, 2017.
- [12] 李建华. 网络空间威胁情报感知、共享与分析技术综述[J]. 网络与信息安全学报, 2016, 2(2): 16-29.
- [13] ANTONAKAKIS M, PERDISCI R, NADJI Y, et al. From throw-away traffic to bots: detecting the rise of DGA-based malware[C]//Proceedings of the 21th USENIX conference on security. Bellevue, WA: USENIX Association, 2012: 24.
- [14] KNY SZ M, HU Xin, SHIN K G. Good guys vs. bot guise: mimicry attacks against fast-flux detection systems[C]//2011 proceedings IEEE symposium. Shanghai: IEEE, 2011: 1844-1852.
- [15] HO G, SHARMA A, WAGNER D, et al. Detecting credential spearphishing attacks in enterprise settings[C]//2017 USENIX symposium on security. Canada: USENIX Association, 2017: 477-478.