基于 TCP 和 DNS 流量特征的 APT 攻击检测

摘 要： 在对真实 APT 攻击流量进行分析的基础上，深入挖掘 APT 恶意软件与控制器通信时，在 TCP 协议、DNS 协议以及端口使用上表现出的规律特征，提出了一种基于周期性通信特征的 APT 攻击检测方法（简称 APDM，APT Periodicity Detection Method）。该方法对数据包中的元数据进行提取并聚类，对 TCP SYN 数据、PSH ACK 数据、DNS 查询数据的周期性以及端口使用的异常进行分析，以实现对 APT 攻击的检测。实验表明，该方法能够有对含有周期性通信特征的 APT 攻击进行检测。

关键词： 高级持续性威胁；周期性；深度包解析；流量特征

中图分类号： TP393 **文献标志码：** A **文章编号：** 0372-2112 (xxxx) xx-xxxx-xx

电子学报 URL： http://www.ejournal.org.cn **DOI：**10.3969/j.issn.0372-2112.xxxx.xx.xxx

APT Detection Based on TCP and DNS Traffic Features

Abstract: Based on the analysis of the APT attack traffic from real world, we thoroughly explored the feature of TCP protocol usage, DNS protocol usage and port usage from APT malwares' communication with the controller, and proposed a periodic communication feature based APT detection method (APDM, APT Periodicity Detection Method). This method extracts and clusters the metadata of the packets, analyzes the periodic feature of TCP SYN data, PSH ACK data, DNS query data, and the abnormal usage of ports to detect APT. Experiments show that this method is able to detect APT attacks with the feature of periodical communication.

Key words: advanced persistent threat; periodicity; deep packet inspection; traffic features

1 引言

1.1 背景介绍

APT 的定义从字面上理解：高级，代表攻击者利用未知漏洞进行入侵、采用专门定制的木马进行控制，难于防范；持续性，代表入侵之后并不立即爆发，而是长期潜伏[1]、逐步渗透，秘密进行窃密[2]活动，难以发现；威胁，代表攻击目的是获取高价值的机密信息并回传，危害极大。APT 攻击的目标通常为政府、教育机构、能源企业、军事系统、工业系统等[3]，它们含有重要机密文件和关键设施信息[4]，具有极大的价值。

APT 攻击不是全新的攻击类型，而是现有技术的复合。它与软件更新有类似的行为模式：长期在后台活动，定期进行程序下载和信息上传。不同之处在于，软件更新受到用户信任和授权，更新目的在于优化软件体验，而 APT 恶意软件在用户未知的情况下被植入系统，定期地接收指令进行未经授权的访问和机密资料的搜集、回传。与 APT 攻击在命令控制机制上具有高度相似性的一种安全威胁是僵尸网络[5]，两者的不同之处如表 1 所示：

表 1 僵尸网络和 APT 攻击的不同之处

	APT	Botnet
工作原理	入侵目标系统关键主机进行窃密活动	通过控制联网肉鸡协同工作
恶意行为	机密资料窃取	DDOS 攻击、垃圾邮件、隐私窃取
信道属性	对少数关键设备进行隐蔽、持续控制，信道趋于易用和隐蔽[5]	对对大规模肉鸡进行无差别控制，信道具有良好的健壮性[5]
攻击目标	政府、教育机构、能源企业、军事系统、工业系统等	互联网上的用户或者机构
传播方式	利用未知漏洞入侵、邮件附件、网页挂马、社会工程学	间谍软件、网页挂马、邮件附件
规模	对少数关键设备进行感染、控制	对大规模的主机进行感染、控制

1.2 相关研究

APT 攻击的检测需要从多层次、全方位的角度[6], [7]对攻击事件进行关联[8]，以发现攻击者的意图和最终目标。在检测架构方面，文献[9]提出了 APT 攻击的金字塔模型，攻击目标处在塔顶，由下至塔顶是各条可到达攻击目标的攻击路径，文献着重于考察攻击事件的环境因素和关联性。文献[10]提出了基于异常发现的 APT 安全体系架构，从恶意程序的源头、途径和终端 3 个层面进行监测，通过异常发现和预警响应定位 APT 攻击。APT 攻击检测根据研究方向的不同分为基于主机的异常检测和基于通信流量的异常检测：前者主要监测主机的连接、进程、文件访问、内存、系统环境变化等[11]，以发现异常行为[12]，定位恶意软件；后者通过对通信流量进行采集、分析，来实现对 APT 攻击的检测。基于通信流量分析的异常检测又可以分为：基于入侵检测技术的检测方法，如文献[13]通过对入侵检测系统报警日志的切分来寻找具有持续性特征的高优先级威胁；基于数据包 DPI 分析的检测方法，

如文献[14]提出对 DNS 数据进行解析，挖掘其流量特征用于检测；基于流记录分析的检测方法，如文献[15]提出了基于 NetFlow 的流记录对流量进行检测以发现异常数据的方法；基于机器学习分类的检测方法，如文献[16]提出了利用 KNN 和关联分形维数算法对 TCP 会话进行分类，以发现 APT 攻击。

未知漏洞的利用，使得基于签名规则的入侵检测能够有效对抗已知攻击，却对 APT 效果有限[17]。在通信过程中，APT 恶意软件为规避检测，通过多个不同端口建立连接进行通信，使得基于 IP、端口、传输层协议的流记录检测方法无法有效的将同一个攻击的流量进行汇聚和检测。而基于机器学习的检测方法，由于缺少足够的 APT 样本，容易导致数据不平衡问题，影响检测效果。此外，机器学习的分类器效果依赖于特征的选取，如何选取有效的特征进行 APT 恶意流量分类，也是目前研究的难点。

1.3 论文工作

APT 攻击的生命周期分为六个阶段[18],[19]：情报搜集阶段[20]；入侵阶段；远程控制阶段；逐步渗透阶段；资料搜集阶段；资料回传阶段。对前两个阶段检测的困难在于，攻击者可以在不直接接触目标系统的情况下，获取其系统信息和内部人员信息[21], [22]，而 0day 漏洞和定制木马的使用更加难以检测。此外，逐步渗透、资料搜集阶段发生在系统内部主机之间，而一般系统内部设备之间的通信网络十分复杂、流量巨大[23]，难以进行有效的检测和追踪。因此，远程控制和资料回传阶段暴露出的通信特征，成为 APT 攻击检测的重要研究对象。图 1 显示了远程控制和资料回传阶段，内部主机和外部控制器[24], [25]的通信过程。受害主机通过 DNS 请求获取控制器 IP 地址，与之建立连接，进行木马下载、指令获取，在搜集一定数量的机密数据之后进行回传。

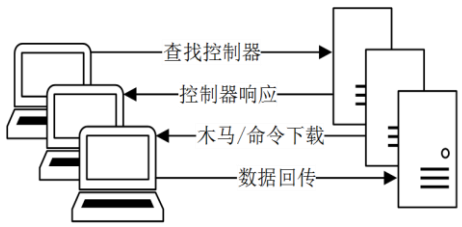

图 1 受害主机与控制器的通信模型

对真实 APT 攻击样本的分析表明，大多数 APT 恶意软件与外部控制器保持着周期性通信的特征。论文基于此特征，采取 DPI 分析的方法，结合上述通信模型，考察数据报文的流量特征，提出了 APDM 检测方法。论文的贡献如下：

1.对真实 APT 样本进行分析，归纳总结出 APT 攻击在 TCP 协议、DNS 协议和端口使用上表现出的特征；

2.依据挖掘出的特征，提出基于 TCP SYN 数据和 PSH ACK 数据的周期性、DNS 查询数据的周期性以及端口使用异常的 APDM 检测方法；

论文的章节安排如下：第一节是引言部分；第二节将对 APT 样本流量进行分析，归纳总结其通信特征；第三节将介绍 APDM 检测方法；第四节是实验结果及对比分析；最后一节将进行总结并对后续工作进行展望。

2 APT 通信特征分析

论文中使用的 APT 样本流量来自于 Contagio 恶意软件数据库[26]，样本的数量为 36 个，其名称和大小如表 2 所示，数据的封装格式是 PCAP 格式。

表 2 样本的信息

编号	名称	大小（KB）	编号	名称	大小（KB）
1	BIN_9002	4086	19	BIN_NJRatLV_6fd86	45
2	XTremeRAT	3588	20	BIN_Gh0st-gif	43
3	BIN_TrojanCookies	905	21	BIN_Taleret	42
4	BIN_Sanny-Daws	869	22	BIN_Nettravler_DA583	39
5	8202_tbd	628	23	BIN_RssFeeder	37
6	BIN_Lagulon	627	24	BIN_OnionDuke	35
7	PDF_CVE	341	25	BIN_Taidoor_40D79	31
8	BIN_XtremeRat	308	26	BIN_Nettravler_1f26	29
9	BIN_LetsGo	306	27	BIN_PlugX	29
10	BIN_Taidoor_46ef9	196	28	RTF_Mongall	15
11	BIN_Mediana	169	29	BIN_Enfal	12
12	Scieron_sandbox1/2	147	30	BIN_njRAT.LV_66070	12
13	Pingbed	119	31	BIN_Tapaoux	12
14	BIN_LURK	104	32	BIN_Likseput	10
15	BIN_DNSWatch	94	33	Mswab_Yayih	10
16	BIN_TrojanPage	90	34	BIN_Gh0st_variant	6
17	Xinmic	48	35	BIN_IXESHE	6
18	BIN_Hupigon	46	36	Darkcomet	4

2.1 基于 TCP 协议的特征分析

APT 恶意样本在 TCP 协议使用上表现出的异常为，每隔时间 t 秒，向对端同一端口发送净荷大小相同的 SYN 报文，用于建立 1~n 个 TCP 连接。时间间隔 t 从 2 秒到 600 秒不等，每个周期内建立的连接数量在 1~3 个不等。含有 SYN 周期性的样本占比为 52.78%。对 SYN 数据时间戳按照 IP 地址对、宿端口、净荷大小进行聚类，计算相邻时间戳的差值，得到 APT 样本和正常数据在 SYN 数据发送时间间隔上的对比，如图 2 所示。

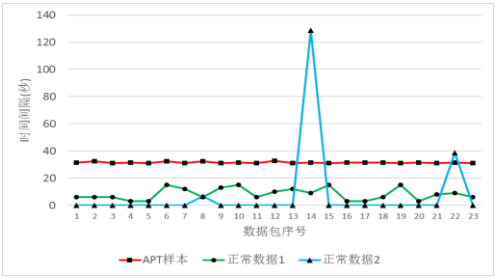

图 2 APT 样本与正常数据的 SYN 数据包发送时间间隔

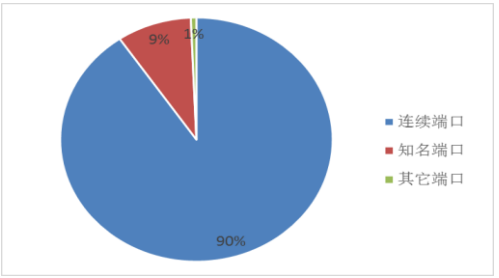

图 3 源端的端口使用情况

纵坐标代表相邻 SYN 数据包的时间间隔，横坐标代表数据包的序号。从图 2 的红线可以看出，APT 样本的连接建立时间间隔基本相同，在图中表现为趋于笔直的水平线；正常数据表现出两种情况，一是连接的建立具有随机性，如绿色曲线所示；一是连接的建立具有突发性，如蓝色曲线所示。此外，69.44%的样本在 PSH ACK 数据上表现出了周期性，其产生原因很大一部分是因为恶意软件周期性的建立连接，并传输标记为 PSH ACK 的数据，因而可以将 PSH ACK 数据周期性结合 SYN 数据周期性共同分析。

2.2 基于 DNS 协议的特征分析

APT 恶意样本中的 DNS 数据异常表现在周期性的 DNS 查询以及查询的域名特征。有 13.89%的样本在 DNS 请求数据上表现出了周期性。DNS 查询数据周期性存在的原因是，受害主机周期性的请求控制器的 IP 地址而形成的。对于 DNS 查询报文中的域名信息，部分 APT 恶意软件请求的域名存在一定的钓鱼网站特征。例如，样本中查询的域名如 nasa.usnewssite.com，域名中采用了关键字 nasa 和 usnews 来混淆，查询的域名如 document.myPicture.info，容易在字面上混淆这些恶意域名和正常域名。

2.3 基于端口的特征分析

APT 恶意软件通过使用多个连续的或形成一定等差规律的端口和对端同一端口建立连接进行通信，以规避检测。图 3 显示了某一样本的源端的端口使用情况。论文将 APT 恶意软件的端口使用分为三类：知名端口、连续端口和其它端口。其中，连续端口指的是连续或形成等差规律的端口，知名端口指的是如 53、80、443 等常见端口，而其他端口指的是非知名、且非连续的普通端口。

3 APDM 检测方法介绍

3.1 APDM 检测方法的架构

APDM 方法的架构如图 4 所示。该方法支持在线和离线检测模式。对输入数据按照传输层协议和端口号进行分类，TCP 处理模块主要负责 TCP 数据的解析、过滤、周期性判定以及端口使用异常的检测；DNS 处理模块负责对 TCP/UDP 53 端口数据的解析、白名单过滤以及 DNS 查询数据的周期性判定。最后，根据各模块的结果对原始数据进行二次扫描，将可疑的数据进行筛选、进一步的分析。

3.2 周期性检测

对于 TCP SYN 数据，按照 IP 地址对、宿端口和净荷大小进行聚类。设观测时间为 T，将其切分为 K 个大小为 t 的时间窗口，即 T=Kt。将聚类后的数据在每个时间窗口 t 进行计数，得到序列 S_{syn}。而对于 PSH ACK 数据，则按照 IP 地址对、端口对和净荷大小进行聚类，以同样的方法进行处理，得到序列 S_{pa}。对于 DNS 查询数据按照原 IP 地址和查询的域名进行聚类，以同样的方法处理得到序列 S_{dns}。以 DNS 查询数据为例，内部 IP 地址 192.168.248.135 周期性的查询了恶意域名 www.prettylikeher.com，设观测时间 T=6000 秒，时间窗口大小 t=600 秒，对其计数得到的查询序列分布情况如表 3 所示。表中，$t1$~$t10$ 代表大小为 600 秒的 10 个时间窗口。

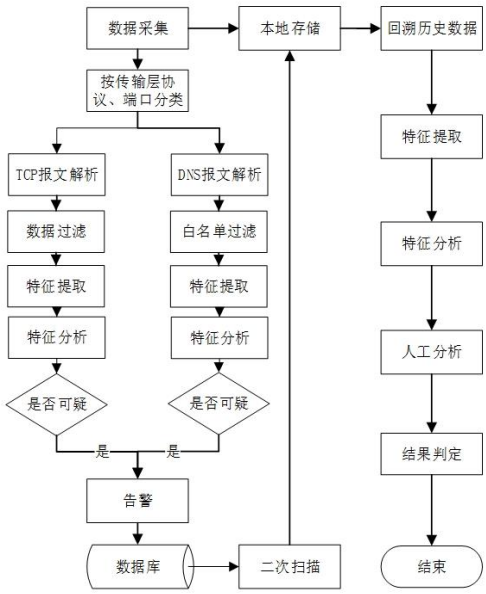

图 4 APDM 方法的架构

表 3 192.168.248.135_www.prettylikeher.com DNS 查询序列

时间窗口	*t1*	*t2*	*t3*	*t4*	*t5*	*t6*	*t7*	*t8*	*t9*	*t10*
计数（个）	2	0	2	2	0	2	2	0	2	0

在统计得到上述的查询序列之后，采用循环自相关方法对其进行计算[26]，其公式如下：

$$r(k)=\sum_{n=1}^{N} f(n)\,f(n+k) \tag{1}$$

其中，$f(n)$表示观测时间 T 内第 n 个时间窗口中聚类条目的计数值，$f(n+k)$表示将 $f(n)$循环右移 k 次后所得的值。步长 k 的变化，会使得 $r(k)$的值呈现出周期性的变化，当 k 的取值为序列周期的整数倍的时候，$r(k)$将取得一个极大值 $r(k)_{\max}$。由此可估计出序列的大致周期。对于任意自然数 k，都有：

$$r(k)\le r(0)(0\le k\le K-1,k\in N) \tag{2}$$

且有：

$$0\le \alpha(k)=\frac{r(k)}{r(0)}\le 1 \tag{3}$$

通过公式(3)，将结果转换为 0 到 1 小数来表示周期性的强弱。通过设置阈值 α_0，则周期性的检测转换为如下判定式，当且仅当 $\alpha(k)_{\max}$ 满足：

$$\alpha(k)_{\max}=\frac{r(k)_{\max}}{r(0)}\ge \alpha_0 \tag{4}$$

则认为此序列具有周期性。

3.3 端口异常检测

论文中的端口异常检测主要是对含有等差规律的端口进行检测。对数据中的源端口按照原宿IP、宿端口三个元组进行聚类，并检测聚类后的原端口分布情况。对端口序列等差性的检测标准是，对于至少三个连续端口，其相邻端口在数值上的差值相同。论文中将端口使用异常和 TCP 周期性检测进行共同判断，当两者均满足时，输出其结果。

4 实验结果分析与对比

4.1 检测结果分析

实验所使用的数据包括 APT 样本流量[26]，以及在教育网某核心节点边界上采集的离线数据。由于该节点边界流量非常大，因此根据 APT 攻击的特点，筛选出 7 个重要的 IP 地址进行监控。这些 IP 地址对应的服务器功能作用如表 4 所示。数据存储的格式为 PCAP 格式，采集时间从 2016 年 1 月 5 日到 2016 年 5 月 31 日，数据总量在 723GB 左右。

表 4 监控 IP 地址以及对应的服务器功能

目标	IP 地址	功能作用
某大学 1	***.***.***.111	技术转移中心
	..***.93	教务处、科技处、研究生院
某大学 2	***.***.***.16	技术转移中心
	..***.143	科技处
	..***.5	教务处
某大学 3	***.***.***.60	科技处
	..***.83	教务处

在实验中，将 APT 样本数据和离线数据进行混合检测，检测结果如表 5 所示。

表 5 样本的检测结果

检测特征	结果数量	结果中为 APT 的数量	结果涉及的样本数量	样本检出率
DNS 查询周期性	8	5	2	40%
SYN 周期性	22	21	18	94.73%
PSH ACK 周期性	2	1	1	-

定义检出率和误报率如下：

$$\text{检出率}=\frac{\text{检测出的APT样本数量}}{\text{含有此特征的样本数量}}\times 100\% \tag{5}$$

$$\text{误报率}=\frac{\text{非APT攻击的结果数量}}{\text{总的检测结果数量}}\times 100\% \tag{6}$$

APDM 方法首先检测 SYN 数据的周期性，如果判定其具有周期性，则不再检测 PSH ACK 数据的周期性，因此在表 5 中计算 PSH ACK 的检出率并没有意义。对于 SYN 数据，19 个含有 SYN 周期性特征的样本中检测出 18 个，检出率为 94.73%；对于 DNS 查询数据，5 个含有 DNS 周期性查询特征的样本中检测出 2 个，检出率为 40%。检测结果分布在 19 个不同的样本文件中，覆盖了原样本的 52.78%。通过对原始数据进行筛选和观察，发现检测结果中的另外 5 个可疑结果并不是 APT 攻击，因

而在当前实验数据下得出的误报率为 15.62%。

4.2 对比实验

论文中对比实验选取的比较对象是 Siddiqui 的 APT 检测算法[16]。该算法的思想是，对 TCP 数据进行组流，提取 TCP 流的持续时间和流数据包数目作为特征，并使用 KNN 算法对其进行分类。论文从 VirusShare 等恶意软件数据库下载了多个 APT 恶意程序在沙盒环境中运行，采集其通信流量。选取其中的 18 份数据，随机替换之前用于分析的样本中的 18 份数据，并将其与正常数据进行混合，组成 5 份独立的实验数据，数据情况如表 6 所示。

表 6 对比实验所用的数据

数据编号	1	2	3	4	5
正常数据大小(MB)	100.81	201.84	404.27	858.51	1215.38
APT 数据大小(MB)	8.20	4.88	8.14	3.85	4.50
总数据量大小(MB)	109.01	206.72	412.41	862.36	1219.88

设置 40%的数据用于训练，60%的数据用于分类。实验结果发现，Siddiqui 的算法对于 TCP 流的分类具有较好的效果，准确率平均达到 93%以上。在分类的结果中，含有四种不同的类型：TP(True Positives)，FP(False Positives)，FN(False Negatives)和 TN(True Negatives)。而在 APT 攻击的检测中，更多关注的是有多少 APT 攻击被检测出来(即检出率，和 TP、FN 相关) 以及误报了多少（即误报率，和 FP 相关），而对于 TN 并不关注。此外，Siddiqui 的算法处理对象是 TCP 流，而论文的处理对象是 IP 地址对。因此，论文对实验结果做了如下处理：

(1) 对 TCP 流按照分类结果进行标记，正常数据标记为 0，APT 数据标记为 1；
(2) 将所有已标记的 TCP 流按照 IP 地址对进行聚类；
(3) 如果一个 IP 地址对中有一条 TCP 流被标记为 1，则将此 IP 地址对标记为 1；否则，当且仅当一个 IP 地址对的所有 TCP 流都被标记为 0，才将此 IP 地址对标记为 0；

处理过后，得到的结果如表 7 所示。

表 7 处理后的结果

数据编号	检出数量（个）	漏报数量（个）	误报数量(个)
1	17	27	406
2	30	28	462
3	17	30	284
4	21	35	387
5	18	41	341

由此，得到了 Siddiqui 的算法对 IP 地址的分类结果。同时，使用 APDM 方法对这 5 份数据进行了检测，对检测结果在检出率和误报率上进行了对比，检出率和误报率的对比如图 5 和图 6 所示。

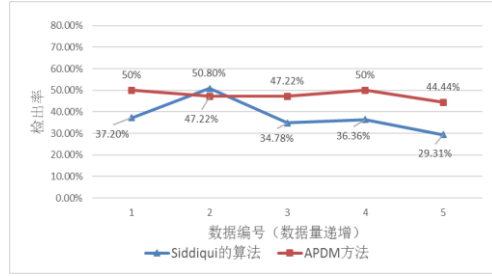

图 5 检出率对比

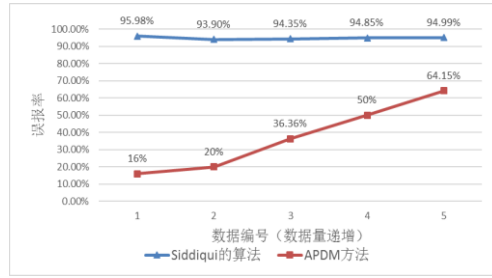

图 6 误报率对比

这两个方法处理时间的对比如图 7 所示。

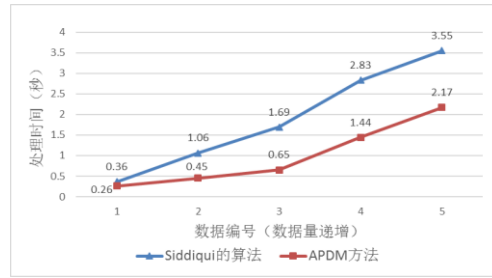

图 7 处理时间对比

从图 5 和图 6 中可以看出，APDM 方法较 Siddiqui 的算法在当前的数据集上具有更高的检出率和更低的误报率。从图 7 可以看出，APDM 方法具有更高的处理性能。

5 结论与展望

论文通过对真实 APT 样本流量的分析，归纳总结出 APT 恶意软件在通信过程中存在的周期性特征，并基于这些特征提出了 APDM 检测方法，通过实验结果分析和对比验证了此检测方法的效果。

在 APT 攻防博弈过程中，双方的技术都在不断提高和进步，不排除恶意软件利用随机性或者模拟人类行为的方式将周期性特征减弱甚至混淆掉的情况，论文也将对检测方法作适当调整，引入域名的语义特征、域名的用户 IP 分布熵等难以改变的特征来进行改进。时域上的随机化特征无法掩藏恶意软件需要定期访问 C&C 服务器的刚需，下一步研究的重点是在时域上加入一些抗随机干扰的处理

以对抗 APT 攻击。

参考文献

[1] 林龙成, 陈波, 郭向民. 传统网络安全防御面临的新威胁: APT 攻击[J]. 信息安全与技术, 2013, 4(3): 20-25.

[2] Moore J W. From phishing to advanced persistent threats: The application of cybercrime risk to the enterprise risk management model[J]. Review of Business Information Systems (RBIS), 2010, 14(4).

[3] Chandran S, Hrudya P, Poornachandran P. An efficient classification model for detecting advanced persistent threat[C]//Advances in Computing, Communications and Informatics (ICACCI), 2015 International Conference on. IEEE, 2015: 2001-2009.

[4] Marchetti M, Pierazzi F, Guido A, et al. Countering Advanced Persistent Threats through security intelligence and big data analytics[C]//Cyber Conflict (CyCon), 2016 8th International Conference on. IEEE, 2016: 243-261.

[5] 李可, 方滨兴, 崔翔, 等. 僵尸网络发展研究[J]. 计算机研究与发展, 2016, 53(10): 2189-2206.

[6] Moon D, Im H, Lee J D, et al. MLDS: multi-layer defense system for preventing advanced persistent threats[J]. Symmetry, 2014, 6(4): 997-1010.

[7] Torii S, Morinaga M, Yoshioka T, et al. Multi-layered defense against advanced persistent threats (apt)[J]. FUJITSU Sci. Tech. J, 2014, 50(1): 52-59.

[8] 李杰, 楼芳, 金渝筌, 等. 面向 APT 攻击的关联分析检测模型研究[J]. 计算机工程与科学, 2015, 37(08): 1458-1464.

[9] Giura P, Wang W. A context-based detection framework for advanced persistent threats[C]. Cyber Security (CyberSecurity), 2012 International Conference on. IEEE, 2012: 69-74.

[10] 杜跃进, 翟立东, 李跃, 等. 一种应对 APT 攻击的安全架构: 异常发现[J]. 计算机研究与发展, 2014, 51(7): 1633-1645.

[11] 朱平, 史记, 杜彦辉. 基于文件, 进程和网络的 APT 检测模型[J]. 信息安全与通信保密, 2014 (3): 99-103.

[12] Auty M. Anatomy of an advanced persistent threat[J]. Network Security, 2015, 2015(4): 13-16.

[13] Quader F, Janeja V, Stauffer J. Persistent threat pattern discovery[C]. Intelligence and Security Informatics (ISI), 2015 IEEE International Conference on. IEEE, 2015: 179-181.

[14] Zhao G, Xu K, Xu L, et al. Detecting APT Malware Infections Based on Malicious DNS and Traffic Analysis[J]. Access, IEEE, 2015, 3: 1132-1142.

[15] Vance A. Flow based analysis of Advanced Persistent Threats detecting targeted attacks in cloud computing[C]//Problems of Infocommunications Science and Technology, 2014 First International Scientific-Practical Conference. IEEE, 2014: 173-176.

[16] Siddiqiu S, Khan M S, Ferens K. Detecting Advanced Persistent Threats using Fractal Dimension based Machine Learning Classification[C]. Proceedings of the 2016 ACM on International Workshop on Security and Privacy Analytics. LA, 2016, pp.64-69.

[17] Dube T E, Raines R A, Grimaila M R, et al. Malware target recognition of unknown threats[J]. IEEE Systems Journal, 2013, 7(3): 467-477.

[18] Bhatt P, Yano E T, Gustavsson P. Towards a framework to detect multi-stage advanced persistent threats attacks[C]//Service Oriented System Engineering (SOSE), 2014 IEEE 8th International Symposium on. IEEE, 2014: 390-395.

[19] Chen P, Desmet L, Huygens C. A study on advanced persistent threats[C]//IFIP International Conference on Communications and Multimedia Security. Springer Berlin Heidelberg, 2014: 63-72.

[20] Brogi G, Tong V V T. TerminAPTor: Highlighting Advanced Persistent Threats through Information Flow Tracking[C]//New Technologies, Mobility and Security (NTMS), 2016 8th IFIP International Conference on. IEEE, 2016: 1-5.

[21] Alavi R, Islam S, Mouratidis H. Human Factors of Social Engineering Attacks (SEAs) in Hybrid Cloud Environment: Threats and Risks[C]//International Conference on Global Security, Safety, and Sustainability. Springer International Publishing, 2015: 50-56.

[22] Molok N N A, Ahmad A, Chang S. Information leakage through online social networking: Opening the doorway for advanced persistence threats[J]. Journal of the Australian Institute of Professional Intelligence Officers, 2011, 19(2): 38.

[23] Yen T F, Juels A, Kuppa A, et al. Anomaly sensor framework for detecting advanced persistent threat attacks: U.S. Patent 9,378,361[P]. 2016-6-28.

[24] Moran N. Understanding advanced persistent threats: A case study[J]. USENIX August 2011, 2011, 36(4).

[25] Wang X, Zheng K, Niu X, et al. Detection of command and control in advanced persistent threat based on independent access[C]//Communications (ICC), 2016 IEEE International Conference on. IEEE, 2016: 1-6.

[26] Mila Parkour. (2013) Contagio malware database.[Online].https://www.mediafire.com/folder/c2az029ch6cke/TRAFFIC_PATTERNS_COLLECTION#734479hwy1b97

[27] 陈玉祥. 基于时空关联性的僵尸网络检测系统的研究与实现[D]. 东南大学, 2016.