DOI: 10.3785/j.issn.1008-973X.2020.08.013

基于动态暗网的互联网扫描行为分析

武秋韵，丁伟
(东南大学 网络空间安全学院，江苏 南京 211189)

摘 要：为了对互联网上的扫描行为进行观测，采用基于动态暗网的互联网背景辐射 (IBR) 流量实时采集算法实现对 IBR 流量的采集，并对采集到的 IBR 流量进行分析；设计算法过滤出扫描流量，进行面向端口的扫描行为观测. 该动态暗网是相对稳定且分散的，不易被定位，通过其获取到的 IBR 流量是进行扫描分析的可靠数据源. IBR 流量主要由传输控制协议（TCP）、用户数据报协议（UDP）、Internet 控制消息协议（ICMP）这 3 种协议组成，其中 TCP 流量占 90% 以上，与正常流量中 3 种协议的分布不同. IBR 流量得到的 TCP、UDP、ICMP 流量都以扫描流量为主，且广泛采用水平扫描的形式. TCP、UDP 的热门扫描端口都是危险端口，证明面向端口的扫描行为分析对于发现互联网中新出现的漏洞有重要作用. TCP 端口扫描行为较分散，UDP 端口扫描行为较集中.

关键词：互联网背景辐射 (IBR)；暗网；扫描检测；扫描行为分析；端口扫描
中图分类号：TP 393.07　　**文献标志码：**A　　**文章编号：**1008−973X(2020)08−1550−07

Analysis of Internet scanning behavior based on dynamic dark network

WU Qiu-yun, DING Wei
(*College of Cyberspace Security*, *Southeast University*, *Nanjing* 211189, *China*)

Abstract: A real-time Internet background radiation (IBR) traffic acquisition algorithm based on the dynamic dark network was used to collect IBR traffic and the collected IBR traffic was analyzed, in order to observe the scanning behavior on the Internet. An algorithm was designed to filter out the scanning traffic to observe the port-oriented scanning behavior. The dynamic dark network is relatively stable and scattered, thus it is not easily to be located. The IBR traffic obtained through it is a reliable data source for scanning analysis. IBR traffic is mainly composed of transmission control protocol (TCP), user datagram protocol (UDP) and Internet control message protocol (ICMP) protocols, of which TCP traffic accounts for more than 90%. It is different from the distribution of the three protocols in normal traffic. The TCP, UDP and ICMP traffic obtained by IBR traffic are mainly scanning traffic, of which horizontal scanning is widely used. The popular scanning ports for both TCP and UDP are dangerous ports, which proves that the port-oriented scanning behavior analysis plays an important role in discovering new vulnerabilities on the Internet. The TCP port scanning behavior is more dispersed, while the UDP port scanning behavior is more concentrated.

Key words: Internet background radiation (IBR); dark network; scanning detection; scanning behavior analysis; port scan

互联网背景辐射 (Internet background radiation, IBR) 流量是指未经请求的单向流量 (unsolicited one-way traffic)[1-2]，概念最早由 Pang 等[3] 提出. IBR 流量广泛存在于互联网中，包括蠕虫传播、分布式拒绝服务 (distributed denial of service, Ddos) 攻击、扫描、错误配置和其他未经请求的活动[1,3].

收稿日期：2019−09−20.　　**网址：**www.zjujournals.com/eng/article/2020/1008-973X/202008013.shtml
基金项目：国家重点研发计划资助项目（2018YFB1800200）.
作者简介：武秋韵（1996—），女，硕士生，从事互联网管理和安全研究. orcid.org/0000-0001-7716-8870. E-mail：qywu@njnet.edu.cn
通信联系人：丁伟，女，教授. orcid.org/0000-0003-4182-1617. E-mail：wding@njnet.edu.cn

对 IBR 流量成分进行分类是相关研究领域的重要研究方法. Wustrow 等[1] 将 IBR 流量分为扫描流量、反向散射流量和错误配置流量. Dainotti 等[2] 将 IBR 流量分为反向散射流量、Conficker 扫描流量和其他. Glatz 等[4] 根据本地、远端和通信主机行为及 IBR 流记录特征, 将 IBR 流量分为恶意扫描、反向散射、无法到达的服务、良性 P2P 扫描、可能为良性的流量、bogon 和其他. 综合这些研究工作, 可以认为 IBR 流量主要由扫描、反向散射和其他三部分构成. 因此, 当获取足够多的 IBR 流量时可以对扫描行为进行观测和分析.

本研究围绕 NJNET_IBR 系统获取的 IBR 流量展开, 根据获取到的 IBR 流量分析其中各种协议的占比, 并对每个协议的报文进行分类; 对其中的传输控制协议 (transmission control protocol, TCP) 和用户数据报协议 (user datagram protocol, UDP) 端口扫描行为进行分析.

1　相关背景

目前 IBR 流量大多是基于暗网获取的. 暗网是指配置了路由但是未被使用的网络空间 (IP 地址段)[3], 因此暗网收到的流量均为 IBR 流量. 传统暗网是由一个或多个固定大小的地址块构成的, 比较著名的有 CAIDA 的 UCSD Network Telescope[5]、美国密歇根大学的 Internet Motion Sensor (IMS)[6]、威斯康星大学麦迪逊分校的 Internet Sink 系统[7]、Cymru 团队的暗网项目 (darknet project)[8] 等, 这些暗网均位于美国. 固定的暗网确保了收到 IBR 流量的质量, 但是也有 2 个较大的缺陷：1）为了保证能有足够规模的 IBR 流量用于研究, 须使用充分大的地址空间, 这样的暗网在 IP 地址相对匮乏的区域较难部署; 2）暗网地址是固定的, 在长时间的运行后这些地址会逐渐暴露给外界, 扫描者会避开这些暗网, 导致这些暗网收到的 IBR 流量的成分中扫描流量偏少.以 UCSD Network Telescope 为例, Jonker 等[9] 分析通过该暗网收集到的反向散射流量, 而不再对扫描流量进行研究.

NJNET_IBR 系统[10-11] 是位于 CERNET 南京主节点网络边界的实时 IBR 流量采集系统, 该系统基于运行网络中一个具有动态隐蔽属性的暗网[12-13] 对 IBR 流量进行实时采集.采用互联网背景辐射实时采集 (real-time Internet background radiation measurement, RIBRM) [10] 算法, 该算法的主要原理是对被管网地址活跃性进行实时测量, 根据地址历史活跃性信息过滤掉活跃地址块, 剩下的地址即为运行网络中的暗网, 从该暗网上获得的所有单向流量均为 IBR 流量.这样的暗网是“流动”且不易定位的, 因而采集到的 IBR 流量更加真实, 从中得到的扫描流量也更具有分析价值.

本研究选取 2019 年 6 月 20 日—6 月 26 日、7 月 4 日—7 月 10 日这 2 周的数据进行分析, 不连续时间段可以有效防止数据偏差.

2　IBR 流量的基本情况

2019 年 6 月 20 日—26 日、7 月 4 日—10 日这 2 个时间段的暗网 IP 地址数都约为 85 万, 其中近 98% 的 IP 地址是相同的, 分布在 328 个网段中. 这样的动态暗网相对稳定且隐蔽, 所收到的 IBR 流量也较可靠.

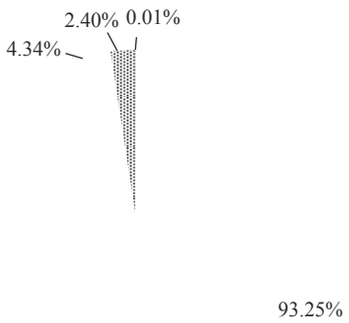
如图 1 所示为这 2 周 IBR 流量的协议分布情况. 图中, P_{TCP}、P_{UDP}、P_{ICMP}、P_{Other} 分别为 TCP、UDP、Internet 控制消息协议 (Internet control message protocol, ICMP) 和其他报文数在 IBR 报文数中的占比.可以看出, TCP、UDP、ICMP 报文数占 IBR 流量的 99.8% 以上, 其他协议占比较小, TCP 协议占据主导地位. 统计这 2 周采集到的正常流量的协议分布情况, 结果如图 2 所示. 可以看出, UDP 流量在 IBR 流量中的占比要小于其在正常流量中的占比.

3　IBR 流量分类

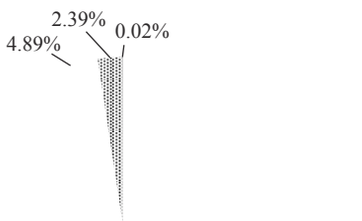
IBR 流量可以简单分为扫描、反向散射和其他三部分, 本研究的主要目的是对扫描流量进行观测, 因此须将扫描流量从 IBR 流量中分离出来. IBR 流量主要由 TCP、UDP、ICMP 这 3 种协议构成, 须分别讨论 3 种协议中扫描流量所占的比例. 对 ICMP、TCP 报文采用现有方法[1,9,11] 进行分类, 对 UDP 报文设计简单算法进行分类.

3.1　ICMP 报文分类

对于 ICMP 报文而言, 将 Ping 请求报文归类为扫描报文, 将响应报文归类为反向散射报文, 剩余报文为其他. 分类结果如图 3 所示, $P_{\mathrm{s}}^{\mathrm{ICMP}}$、$P_{\mathrm{b}}^{\mathrm{ICMP}}$、$P_{\mathrm{o}}^{\mathrm{ICMP}}$ 分别为扫描、反向散射和其他报文数在 ICMP 报文数中的占比. 可以看出, 在 ICMP 报



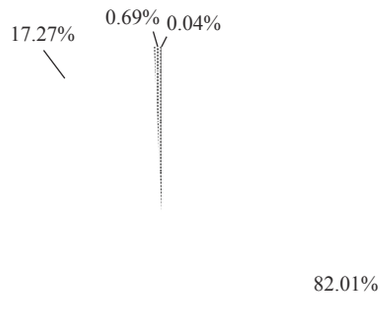
(a) 2019年6月20日—26日



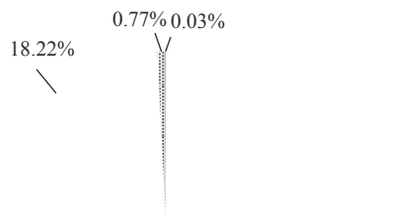
(b) 2019年7月4日—10日

图 1 IBR 流量的协议分布情况

Fig.1 Protocol distribution of IBR traffic



(a) 2019年6月20日—26日



(b) 2019年7月4日—10日

图 2 正常流量的协议分布情况

Fig.2 Protocol distribution of normal traffic

文中有约 97% 的扫描报文、约 2% 的反向散射报文，其他报文不到 1%，由此可见 ICMP 报文以扫描报文为主.

3.2 TCP 报文分类

类似 ICMP，将 TCP 请求报文 (TCP SYN 报文) 归为扫描报文，将 TCP 响应报文 (TCP SYN+ACK、TCP ACK、TCP ACK+RST、TCP RST 报文) 归为反向散射报文，剩余报文为其他报文.对于扫描报文，根据王力[14]提出的扫描检测算法将其分为水平扫描报文和除水平扫描报文外的其他扫描报文.

如图 4 所示为分类结果. 图中，P_{b}^{TCP}、$P_{h\_s}^{TCP}$、$P_{o\_s}^{TCP}$、P_{o}^{TCP}分别为反向散射、水平扫描、其他扫描和其他报文数在 TCP 报文数中的占比. 可以看出，TCP 水平扫描报文数占 TCP 报文总数的 94% 以上，其他扫描报文数不到 2%，因而 TCP 报文也是以扫描报文为主，且 TCP 扫描广泛采用水平扫描方式.

3.3 UDP 报文分类

与 TCP、ICMP 报文不同，UDP 报文无标志位，无法从报文头直接判断报文是否为扫描报文，因此设计简单算法，根据主机行为来判断报文是否为扫描报文. UDP 水平扫描主机、垂直扫描主机和随机扫描主机定义如下.

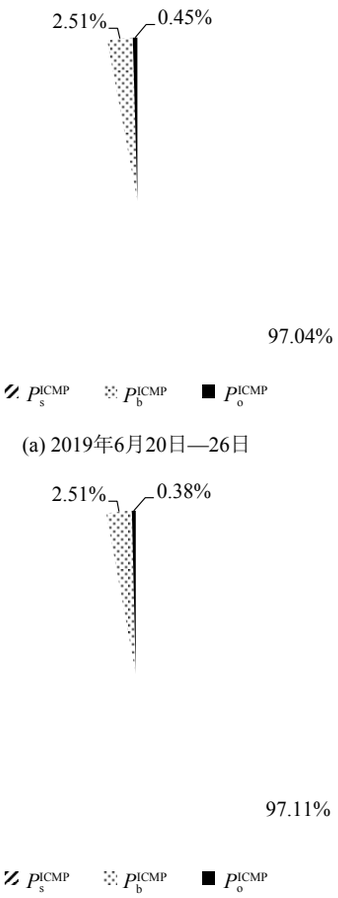
1）UDP 水平扫描主机. 若一个主机在 T 时间内向 O 个不同主机的同一端口发送相同字节数的报文，认为该主机为水平扫描主机，其向该端口发出的所有报文为水平扫描报文.

2）UDP 垂直扫描主机. 若一个主机在 T 时间内向同一主机的 P 个不同端口发送报文，则认为该主机为垂直扫描主机，其向该目的主机发送的所有报文为垂直扫描报文.

3）UDP 随机扫描主机. 若一个主机在 T 时间内向至少 Q 个不同的 (宿 IP，宿端口) 对发送报文，且对其发送的报文计算熵值：

$$H = -\sum P(X)\ln P(X). \tag{1}$$

式中：$P(X)$ 为该主机向第 X 个 (宿 IP，宿端口) 对发送的报文数与其发送的报文总数的比值. 若

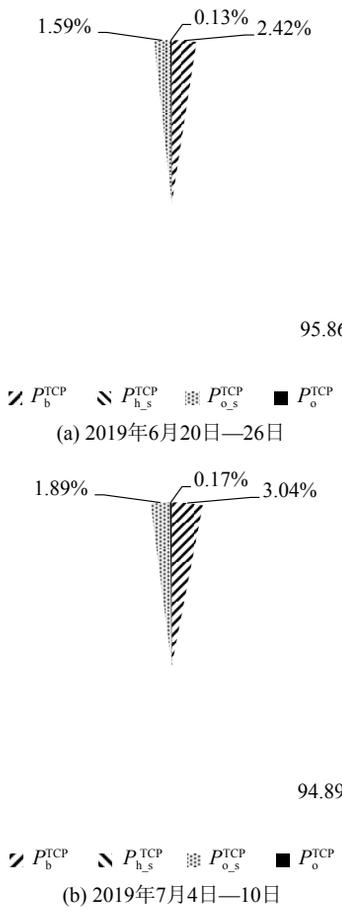


(a) 2019年6月20日—26日

(b) 2019年7月4日—10日

图 3　ICMP 报文分类结果

Fig.3　ICMP classification results



(a) 2019年6月20日—26日

(b) 2019年7月4日—10日

图 4　TCP 报文分类结果

Fig.4　TCP classification results

$H \geqslant \ln Q$, 则认为该主机为随机扫描主机, 其发送的所有报文为随机扫描报文.

根据上述定义, 算法设计如下.

输入: T 时间内 IBR 中的所有 UDP 报文.

输出: 水平扫描报文集合 H_1、垂直扫描报文集合 H_2、随机扫描报文集合 H_3 和其他报文集合 H_4.

操作:

1) 对所有 UDP 报文按 (源 IP, 宿端口, 字节数) 进行分类, 对每一类报文, 若其不同的宿地址数大于等于 O, 则认为该类报文属于集合 H_1, 否则进行下一步分类.

2) 对第 1)步所有剩余报文按 (源 IP, 宿 IP) 进行分类, 对每一类报文, 若其不同的端口号大于等于 P, 则认为该类报文属于集合 H_2, 否则进行下一步分类.

3) 对第 2)步所有剩余报文按源 IP 进行分类, 若其 (宿 IP, 宿端口) 对数目小于 Q, 则将该类报文归为 H_4; 否则按式 (1) 计算熵值, 若熵值大于等于 $\ln Q$, 则将报文归类为 H_3, 否则归类为 H_4.

UDP 报文分类算法中的参数设置参考文献 [14]、[15], T=24 h, O、P、Q=5, 分类结果如图 5 所示. 图中, $P_{h\_s}^{UDP}$、$P_{r\_s}^{UDP}$、$P_{v\_s}^{UDP}$、P_o^{UDP}分别为 UDP 水平扫描、随机扫描、垂直扫描和其他报文数在 UDP 报文数中的占比. 可以看出, UDP 扫描报文数约占 UDP 报文总数 97%, 与 TCP 一样, 也以水平扫描为主.

4　面向端口的扫描行为观测

通过对面向端口的扫描行为进行观测可以发现最受扫描者们青睐的端口, 这是互联网上扫描现状的直观反映, 有利于观测者们发现互联网上新出现的漏洞, 进而及时采取针对措施. ICMP 协议是网络层协议, 其报文没有端口号, 所以只对 TCP、UDP 端口扫描行为进行分析.

4.1　TCP 端口扫描行为

如表 1、2 所示分别为 2019 年 6 月 20 日—6 月 26 日、7 月 4 日—7 月 10 日 TCP 扫描报文数排名前 10 的端口, port 为端口号, n 为 7 d 内扫描

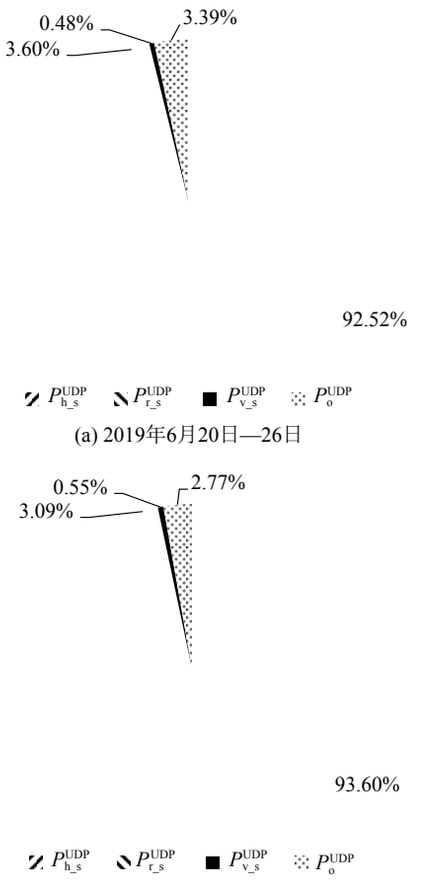


图 5 UDP 报文分类结果

Fig.5 UDP classification results

该端口的报文总数，P_n 为 n 与这 7 d 该类型端口扫描报文总数的比值，P_c 为 P_n 的累加值，Service 为端口号对应的服务或漏洞. 可以看出，这 2 个星期的热门扫描端口较一致，其中扫描报文数排名前 8 的端口完全一样，不同的是 8 545 端口在 6 月 20 日—26 日排名第 11 位，5038 端口在 7 月 4 日—10 日排名第 15 位，但其扫描报文数在 2 个星期中并没有太大差别. 除此以外，在这 2 个星期中，65 536 个端口都收到了扫描报文，其中排名前 10 的端口的扫描报文数只占总扫描报文数的不到 16%，说明 TCP 端口上的扫描行为是分散的.

对这 11 个端口分别进行分析，其中 22、23、80、445、1 433、3 389、8 080 这几个端口是一直以来被扫描者们持续关注的危险端口，说明这些端口上的漏洞仍有较大威胁. 华为路由器 HG532 在 37 215 端口上存在 CVE-2017-17215 漏洞，该漏洞允许远程执行任意代码，在 2017 年便有报道表明在该端口上有 Mirai 变种 Satori 类似蠕虫式传播. 5 555 端口允许通过 Android 调试桥 (Android debug bridge, ADB) 管理设备，这是一种 Android SDK 功能，允许开发人员与设备通信并在其上运行命令或完全控制它们，自 2018 年起针对该端口的攻击逐渐增多. 5 038 端口是 Asterisk 服务器侦听端口. 8 545 端口是以太坊通信端口，有黑客利用该端口上的漏洞窃取以太币. 这 4 个端口都是近几年新出现的具有漏洞的端口. 可以发现，所有的热门扫描端口均为危险端口，说明对端口扫描行为进行持续观测对发现新漏洞有重要作用. 除此以外，须关注的是，443 端口并不在这 11 个

表 1 2019 年 6 月 20 日—26 日 TCP 热门扫描端口

Tab.1 TCP popular scanning ports from June 20 to June 26, 2019

port	n	P_n / %	P_c / %	Service
23	2 020 697 594	6.212	6.212	Telnet
445	1 120 403 460	3.444	9.656	SMB
22	501 441 369	1.541	11.197	SSH
3 389	313 051 184	0.962	12.159	RDP
80	259 926 595	0.799	12.958	HTTP
37 215	225 917 093	0.694	13.652	华为路由器 HG532 CVE-2017-17215漏洞
1 433	208 795 986	0.642	14.294	SQL Server
8 080	204 845 408	0.630	14.924	Alt-HTTP
5 555	203 975 022	0.627	15.551	ADB
5 038	143 645 366	0.442	15.993	Asterisk服务器侦听端口

表 2 2019 年 7 月 4 日—10 日 TCP 热门扫描端口

Tab.2 TCP popular scanning ports from July 4 to July 10, 2019

port	n	P_n / %	P_c / %	Service
23	1 744 165 990	5.035	5.035	Telnet
445	1 039 652 805	3.001	8.036	SMB
80	775 669 513	2.239	10.275	HTTP
22	549 943 948	1.587	11.862	SSH
3 389	347 211 413	1.002	12.864	RDP
37 215	277 886 064	0.802	13.666	华为路由器 HG532 CVE-2017-17215漏洞
8 080	221 624 050	0.640	14.306	Alt-HTTP
1 433	209 356 000	0.604	14.910	SQL Server
8 545	192 044 920	0.554	15.464	以太坊通信端口
5 555	179 283 575	0.518	15.982	ADB

端口之中，它在 6 月 20 日—26 日排第 15 位，在 7 月 4 日—7 月 10 日排第 14 位，扫描报文数约为 1 亿条，远远小于 80 端口，在 HTTPS 协议应用愈加广泛的现在，足以说明 HTTPS 协议相比 HTTP 有更高的安全性.

4.2 UDP 端口扫描行为

如表 3、4 所示分别为 2019 年 6 月 20 日—26 日、7 月 4 日—10 日 UDP 扫描报文数排名前 10 的端口. 和 TCP 一样，这 2 周的 UDP 扫描端口也较一致，不同的是 UDP 端口上的扫描行为更加集中，虽然也是全部 65 536 个端口都收到了扫描报文，但是排名前 10 的端口汇聚了超过 1/3 的 UDP 扫描报文.其中，19、53、123、137、161、1 900、11 211 都是著名的具有放大器漏洞的端口. 5 060 端口上的会话发起协议 (session initiation protocol, SIP) 协议是信令控制协议，用于创建、修改和释放一个或多个参与者的会话. 53 413 端口被 Netcore 路由器使用，早在 2014 年便被爆出在该端口上存在严重的后门漏洞，攻击者可以通过此漏洞获取路由器 Root 权限. 389 端口被 LDAP、ILS 协议共用，其中 LDAP 协议是轻量级目录访问协议，因为是轻量级而不包含安全措施，易受到恶意攻击和篡改，存在较大的安全隐患. 111 端口是 Sun 公司的远程过程调用 (remote procedure call, RPC) 服务，其存在远程缓冲溢出漏洞. 这 11 个端口都是危险端口. 从扫描报文数占比上可以看出，反射攻击的危险程度有所降低，扫描更多集中在 5 060 端口和 53 413 端口，表明近些年对反射攻击的防范更加到位，同时在今后须加强对 5 060、53 413 端口的关注.

表 3 2019 年 6 月 20 日—26 日 UDP 热门扫描端口

Tab.3 UDP popular scanning ports from June 20 to June 26, 2019

port	n	P_n / %	P_c / %	Service
5 060	182 714 745	12.180	12.180	SIP
53 413	142 737 514	9.515	21.696	Netcore(Netis)路由器后门漏洞
53	57 359 489	3.824	25.520	DNS
1 900	54 719 415	3.648	29.167	SSDP
123	47 214 569	3.147	32.315	NTP
161	33 744 445	2.250	34.564	SNMP
389	31 243 000	2.083	36.647	LDAP、ILS
137	23 160 513	1.544	38.191	NetBIOS
11 211	15 765 197	1.051	39.242	Memcached
19	15 421 089	1.028	40.270	Chargen

表 4 2019 年 7 月 4 日—10 日 UDP 热门扫描端口

Tab.4 UDP popular scanning ports from July 4 to July 10, 2019

port	n	P_n / %	P_c / %	Service
5 060	188 249 763	10.250	10.250	SIP
53 413	81 318 222	4.428	14.677	Netcore(Netis)路由器后门漏洞
1 900	81 186 397	4.420	19.098	SSDP
123	61 200 710	3.332	22.430	NTP
53	58 530 676	3.187	25.617	DNS
389	47 914 908	2.609	28.226	LDAP、ILS
161	30 607 468	1.667	29.892	SNMP
137	22 735 052	1.238	31.130	NetBIOS
19	19 809 086	1.079	32.209	Chargen
111	15 924 692	0.867	33.076	Sun RPC

5 结 语

本研究基于动态隐蔽暗网实时获取 IBR 流量，并对获取的 IBR 流量的协议分布进行分析，同时设法从中分离出扫描流量，进行面向端口的扫描行为观测. 本研究针对的动态暗网的稳定 IP 数约为 85 万个，规模小于著名暗网 UCSD、Network、Telescope 等，无法从中收到足够规模的反向散射流量，因而较难观测反射攻击行为. 该动态暗网的 IP 地址不是完全固定的，且分散在许多不同的地址块中，较难被扫描者定位并避开，所以可以用采集到的 IBR 流量进行扫描行为观测.

根据采集到的 IBR 流量，可以发现其中 TCP 报文数占 90% 以上，UDP 报文数约占 5%，而在正常流量中 TCP、UDP 分别占 80%、18%，有明显差别.在分别对 TCP、UDP、ICMP 报文进行分类时，发现三者的扫描报文数都占其各自总报文数约 95%，且 TCP、UDP 的扫描形式都以水平扫描为主. 对 TCP 和 UDP 扫描报文进行面向端口的扫描行为观测，发现 TCP 端口上的扫描行为较分散，但 2 周的热门扫描端口高度一致且均为具有漏洞的端口. 须注意的是 443 端口并不在热门扫描端口中，证实 HTTPS 协议比 HTTP 协议更具安全性.

UDP 的热门扫描端口也高度一致且均为具有漏洞的端口，与 TCP 不同的是，UDP 端口上的扫描行为更加集中.除此以外，相比具有放大器协议的端口，5 060 端口和 53 413 端口获得了更大的流量，说明近年来反射攻击的强度有所降低，但须加强对 5 060 和 53 413 端口的观测.

接下来会建立完善的扫描行为分析系统，以及时发现新的扫描趋势，并建立模型实现对异常扫描行为的检测. 同时还会对扫描主机进行观测并对其扫描意图进行分析，因为扫描并不都是恶意的，像类似 Shodan 或 Zoomeye 这样的机构进行扫描只是为了发现服务，对扫描者进行意图识别有助于对扫描流量进行非恶意过滤.

参考文献 (References):

[1] WUSTROW E, KARIR M, BAILEY M, et al. Internet background radiation revisited [C]// **Proceedings of the 10th ACM SIGCOMM Conference on Internet Measurement 2010**. Melbourne: ACM, 2010: 62–74.

[2] DAINOTTI A, AMMAN R, ABEN E, et al. Extracting benefit from harm: using malware pollution to analyze the impact of political and geophysical events on the internet [J]. **Acm Sigcomm Computer Communication Review**, 2012, 42(1): 31–39.

[3] PANG R, YEGNESWARAN V, BARFORD P, et al. Characteristics of Internet background radiation [C]// **Proceedings of the 4th ACM SIGCOMM Conference on Internet Measurement 2004**. Sicily: ACM, 2004: 27–40.

[4] GLATZ E, DIMITROPOULOS X. Classifying Internet one-way traffic [J]. **ACM SIGMETRICS Performance Evaluation Review**, 2012, 40(1): 417.

[5] MOORE D, SHANNON C, VOELKER G M, et al. Network telescopes: technical report [R]. [s.l.]: Proceedings of the Cooperative Association for Internet Data Analysis, 2004.

[6] BAILEY M, COOKE E, JAHANIAN F, et al. The Internet motion sensor: a distributed blackhole monitoring system [C]// **Proceedings of the 12th Annual Network and Distributed System Security Symposium (NDSS 2005)**. San Diego: The Internet Society, 2005.

[7] YEGNESWARAN V, BARFORD P, PLONKA D. On the design and use of Internet sinks for network abuse monitoring [C]// **Proceedings of the Symposium on Recent Advances in Intrusion Detection (RAID 2004)**. Berlin: Springer-Verlag, 2004: 146–165.

[8] Team cymru darknet project [EB/OL]. (2005) [2019-07-23]. http://www.team-cymru.org/Services/darknets.html.

[9] JONKER M, KING A, KRUPP J, et al. Millions of targets under attack: a macroscopic characterization of the DoS ecosystem [C]// **Proceedings of the 2017 Internet Measurement Conference**. London: ACM, 2017: 100–113.

[10] 缪丽华, 丁伟, 杨望. 运行网络背景辐射的获取与分析 [J]. 软件学报, 2015, 26(3): 663–679.
MIAO Li-Hua, DING Wei, YANG Wang. Extracting and analyzing Internet background radiation in live networks [J]. **Journal of Software**, 2015, 26(3): 663–679.

[11] 杨扬. 互联网背景辐射流量的获取与统计分析 [D]. 南京: 东南大学, 2016.
YANG Yang. Obtaining and analyzing on Internet background radiation [D]. Nanjing: Southeast University, 2016.

[12] HARROP W, ARMITAGE G. Greynets: a definition and evaluation of sparsely populated darknets [C]// **Proceedings of the 2005 ACM SIGCOMM Workshop on Mining Network Data**. Philadelphia: ACM, 2005: 171–172.

[13] HARROP W, ARMITAGE G. Defining and evaluating greynets (sparse darknets) [C]// **Proceedings of the IEEE Conference on Local Computer Networks 30th Anniversary**. Sydney: IEEE Computer Society, 2005: 344–350.

[14] 王力. 互联网扫描行为研究 [D]. 南京: 东南大学, 2018.
WANG Li. R esearch of scanning behavior on Internet [D]. Nanjing: Southeast University, 2018.

[15] DURUMERIC Z, BAILEY M, HALDERMAN J A. An Internet-wide view of Internet-wide scanning [C]// **Proceedings of the 23rd USENIX Conference on Security Symposium**. San Diego: USENIX Association, 2014: 65–78.