Internet Background Radiation Revisited

Eric Wustrow
Networking Research and
Development
Merit Network Inc.
Ann Arbor, MI 48104 USA
ewust@merit.edu

Manish Karir
Networking Research and
Development
Merit Network Inc.
Ann Arbor, MI 48104 USA
mkarir@merit.edu

Michael Bailey
Department of EECS
University of Michigan
Ann Arbor, MI 48109 USA
mibailey@eecs.umich.edu

Farnam Jahanian
Department of EECS
University of Michigan
Ann Arbor, MI 48109 USA
farnam@eecs.umich.edu

Geoff Huston
Asia Pacific Network
Information Centre
Brisbane QLD 4064 Australia
gih@apnic.net

ABSTRACT

The monitoring of packets destined for routeable, yet unused, Internet addresses has proved to be a useful technique for measuring a variety of specific Internet phenomenon (e.g., worms, DDoS). In 2004, Pang et al. stepped beyond these targeted uses and provided one of the first generic characterizations of this non-productive traffic, demonstrating both its significant size and diversity. However, the six years that followed this study have seen tremendous changes in both the types of malicious activity on the Internet and the quantity and quality of unused address space. In this paper, we revisit the state of Internet "background radiation" through the lens of two unique data-sets: a five-year collection from a single unused /8 network block, and week-long collections from three recently allocated /8 network blocks. Through the longitudinal study of the long-lived block, comparisons between blocks, and extensive case studies of traffic in these blocks, we characterize the current state of background radiation specifically highlighting those features that remain invariant from previous measurements and those which exhibit significant differences. Of particular interest in this work is the exploration of address space pollution, in which significant non uniform behavior is observed. However, unlike previous observations of differences between unused blocks, we show that increasingly these differences are the result of environmental factors (e.g., misconfiguration, location), rather than algorithmic factors. Where feasible, we offer suggestions for clean up of these polluted blocks and identify those blocks whose allocations should be withheld.

Categories and Subject Descriptors

C.2 [**Information Systems Applications**]: Network Operations

General Terms

Measurement

Keywords

Darknet, Internet background traffic, Network pollution

1. INTRODUCTION

The monitoring of allocated, globally routeable, but unused Internet address blocks has been widely used by the security, operations, and research communities to study a wide range of interesting Internet phenomenon. As there are no active hosts in these unused blocks, packets destined to these IP addresses must be the result of worm propagation[1, 2, 3], DDoS attacks[4], misconfiguration, or other unsolicited activity. Systems that monitor unused address spaces have a variety of names, including darknets [5], network telescopes [6], blackhole monitors [7], network sinks[8], and network motion sensors [9].

While this monitoring technique had seen heavy use in the measurement of specific phenomena, it wasn't until 2004 when Pang et. al [10] published their seminal paper "Characteristics of Internet Background Radiation" that a detailed characterization of this incessant non-productive traffic was available. Through passive measurement and active elicitation of connection payloads over several large unused blocks, the authors characterized the behavior of sources and the activities prevalent in Internet background radiation. Most notable in their analysis was the ubiquity of Internet background radiation, its scale, its rich variegation in targeted services, and the extreme dynamism in many aspects of the observed traffic.

The six years since this landmark paper have seen significant changes both in the size, shape, and traffic carried by the Internet as well as the methods and motivations of malicious traffic that makes up Internet background radiation. While both scanning as a reconnaissance activity and
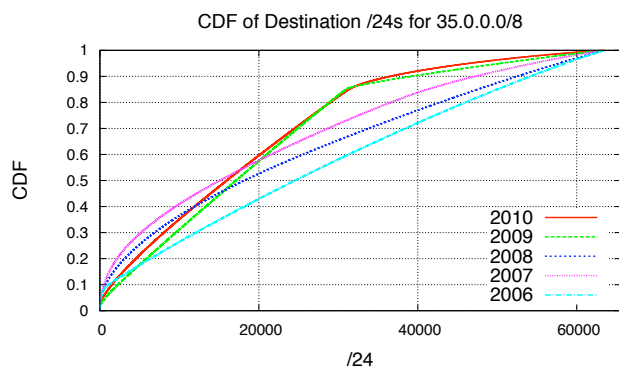
(a) The distribution of destination /24s targeted in the unused address block.
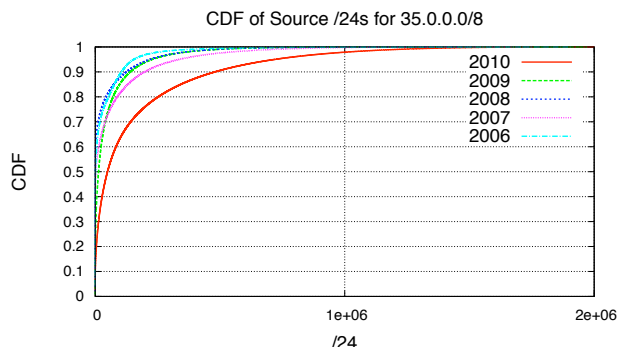
(b) The distribution of source /24s targeted in the unused address block.

Figure 3: Changes in source and destination behavior from 2006-2010 using datasets D-1, D-2, D-3, D-4, D-5.

TCP Port	A1 - A2	B1 - B2	C1 - C2
21	40.3	—	—
25	1.7	—	—
80	8.7	—	—
443	1.6	—	—
445	-75.0	-2.0	-7.5
143	32.5	—	—
1024	—	—	-1.1
5022	1.6	—	—
6112	2.2	—	—

Table 7: The most significant changes in the contribution of a TCP destination port when compared between blocks. Only those ports whose contribution to total traffic at a block that were different by more than %1 are shown.

after 2008, we observe roughly 3 times less traffic for destination IPs with a second or fourth octet of 128 or greater.

4.2 Spatial Analysis of Internet Background Radiation

Figures 4, 6, 8, 9, and 7 represent our analysis of datasets A-[1,2], B-[1,2], and C-[1,2]. The stacked graphs represent data collected from the 35/8 darknet (A2, B2, and C2) on the top row, while the bottom row of graphs represents data collected from 1/8(A1), 50/8(B1), and 107/8(C1).

The overall traffic volume in bytes and packets is shown in Figure 4. One of the most dramatic features is the enormous volume of traffic in 1/8. The 1/8 network sees Internet Background Radiation rates as high as 150Mbps. As we discuss in the following section, most of this traffic is directed toward a small number of destinations in 1/8, due to misconfiguration in a wide range of Internet devices. Both 50/8 and 107/8 traffic rates show a significant diurnal pattern with almost similar data rates. The overall darknet traffic volume ranges from 20-40Mbps or 40-60Kps. One puzzling feature visible in these figures is the clipped nature of the 35/8 graphs. We believe this is caused by a rate limit on a device that is present in the path of our data collector. While we have been able to verify that such a limit is not present in oCDFown collection network, we have so far not

been able to verify that there is no such setting at our upstream provider. The traffic volume by protocol is shown in Figure 5. The traffic volume in Figures 5 and 4 shows a sharp dip on day 7 of the A-1 dataset. This was caused by a temporary duplicate BGP announcement by APNIC.

The first column of Figures 6 and 7 show the cumulative distribution function (CDF) of the cumulative contributions of traffic with destination and source in each /24 network. The 1/8 graphs show extremely high hotspot activity in both these figures as evidenced by the extremely sharp knee in these graphs. The second and third columns correspond with datasets B-1 and C-1 respectively. Both of these display moderate hotspot activity in the destination CDF but the source CDF graphs are virtually identical across A-2, B-1, B2, C-1, C-2. We describe some of this hotspot activity in detail in the next section.

Datasets A-2, B-1, B-2, C-1, and C-2 all display remarkable similarity in the TCP destination port distributions. Table 7 summarizes the differences between these datasets. It shows ports whose contribution was different by more than 1% when compared to the A2, B2, and C2. The most interesting features that we discovered during our analysis of the UDP destination ports was some unusual activity on port 514 in dataset B-1, which is the port associated with syslog, as well as activity on port 15206, which represents SIP traffic in dataset A-1. Figure 8 shows the traffic volume contributed by these features and we discuss some of them in the following section.

The source Operating System estimate obtained by observing the TTL values in the TCP packets shows that the relative volume of traffic generated from the various sources appears to be the same for 35/8, 50/8, and 107/8 (datasets A-2, B-1, B-2, C-1, C-2). Recall that the default TTL values for Windows, Linux and Solaris are 128, 64 and 255 respectively. Windows hosts tend to dominate the total traffic volume by various sources in all except the 1/8 darknet block, where Linux sources are responsible for a majority of the traffic. Analysis of the UDP TTL values displays similar distribution for all darknets except once again the 1/8 where we see Windows, Linux, Solaris and perhaps some embedded devices as possible contributors to the pollution. We were, for example, able to identify some pollution at this network
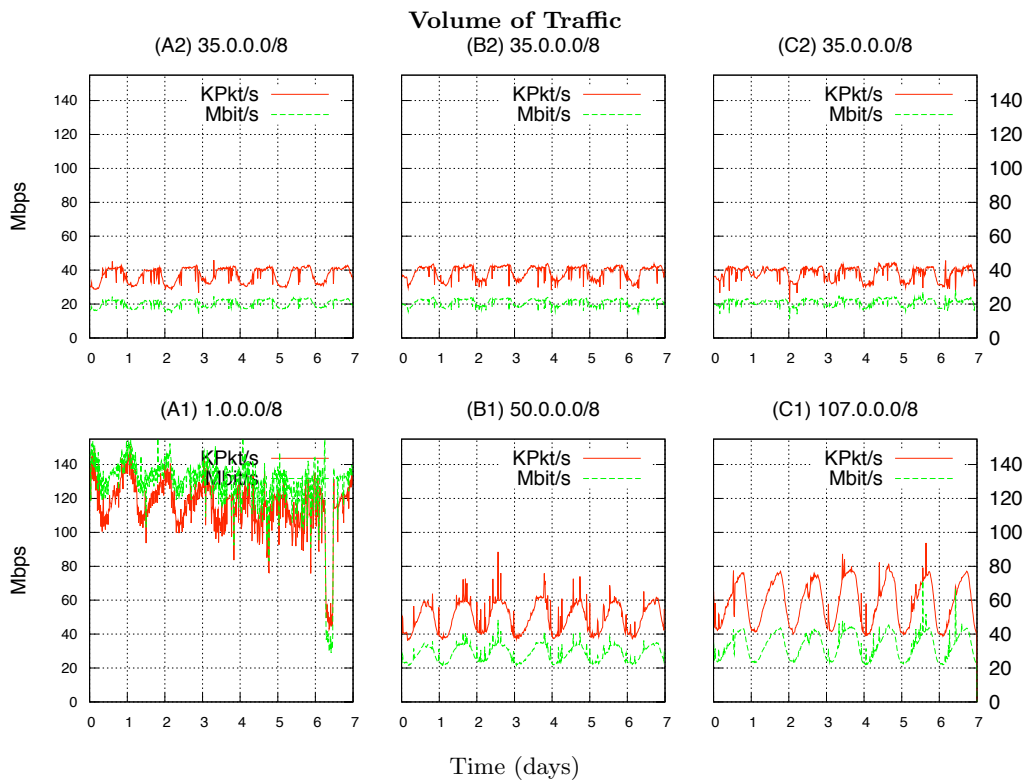
Volume of Traffic

Figure 4: Spatial analysis of Internet Background Radiation. Overall measured traffic (bytes and packets) is shown for datasets A-1, A-2, B-1, B-2, C-1, C-2

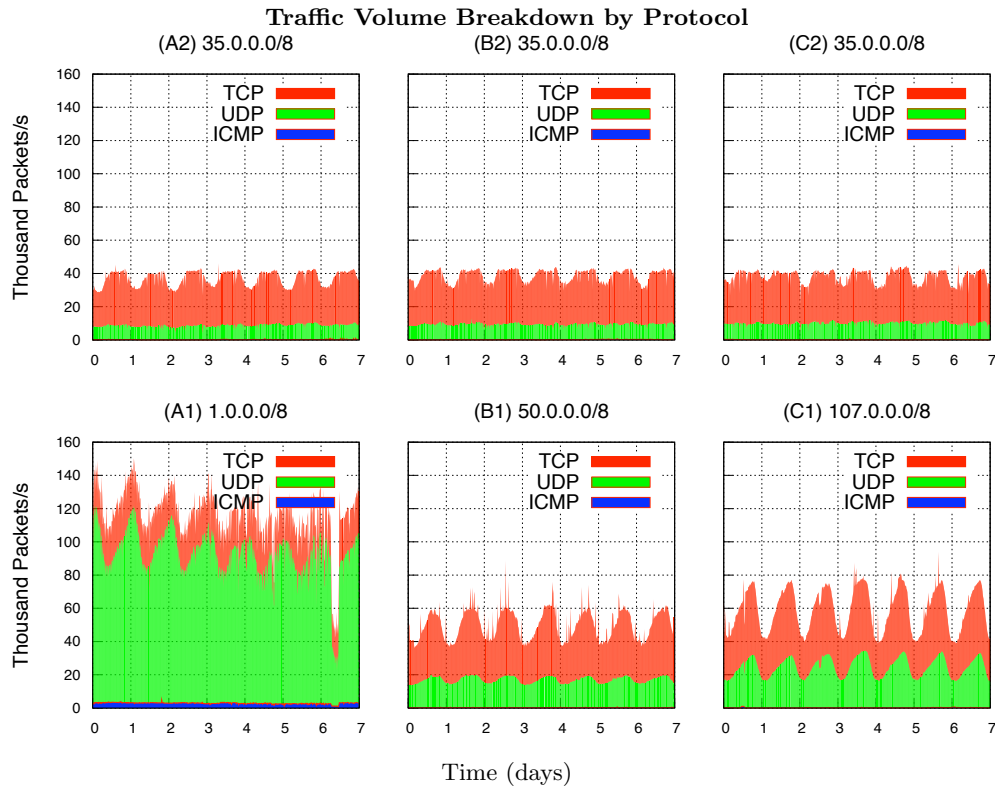## Traffic Volume Breakdown by Protocol

Figure 5: Spatial analysis of Internet Background Radiation. Overall measured traffic by protocols is shown for datasets A-1, A-2, B-1, B-2, C-1, C-2

Destination IP Address CDF

(A2) 35.0.0.0/8

(B2) 35.0.0.0/8

(C2) 35.0.0.0/8

(A1) 1.0.0.0/8

(B1) 50.0.0.0/8
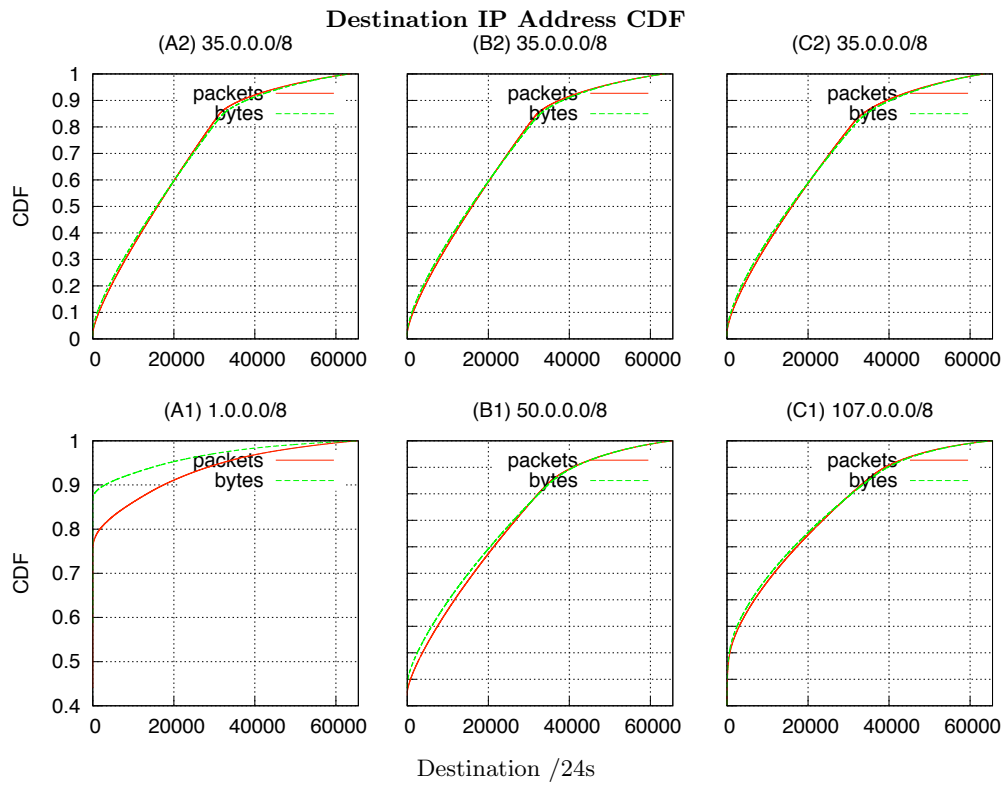

(C1) 107.0.0.0/8

Destination /24s

Figure 6: Spatial analysis of Internet Background Radiation. The CDF representing the cumulative contribution of individual /24 destination using datasets A-[1,2], B-[1,2], C-[1,2].

Source IP Address CDF (/24 networks)

(A2) 35.0.0.0/8

(B2) 35.0.0.0/8

(C2) 35.0.0.0/8

(A1) 1.0.0.0/8

(B1) 50.0.0.0/8

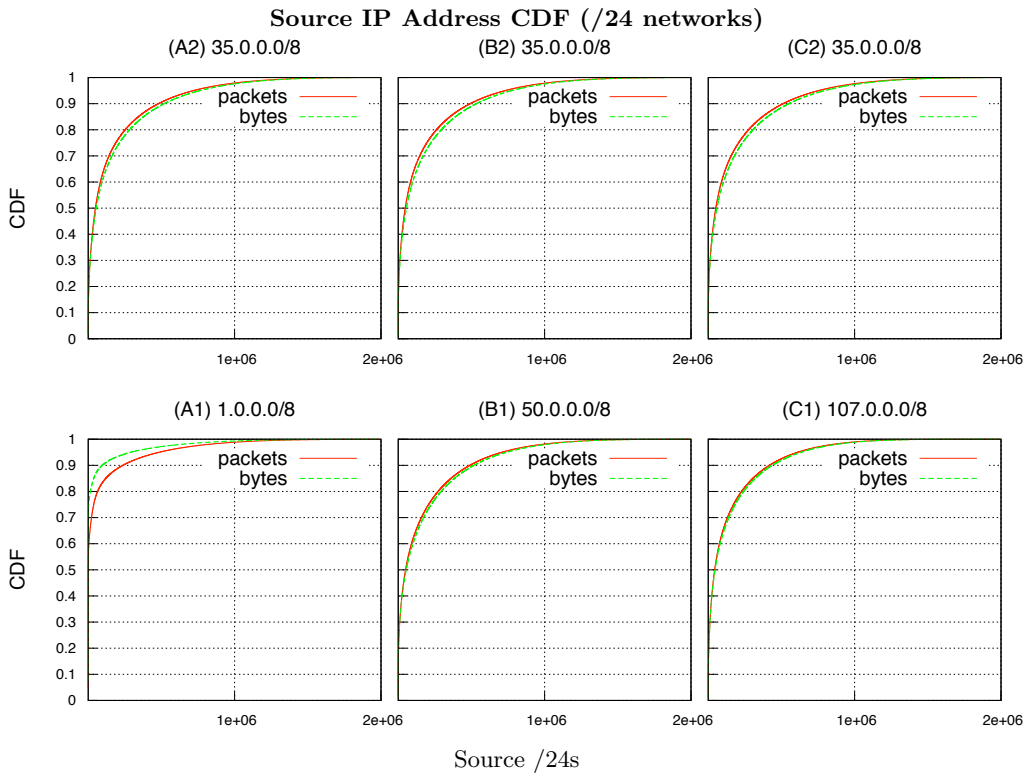**(C1) 107.0.0.0/8**

Source /24s

Figure 7: Spatial analysis of Internet Background Radiation. The cumulative distribution function (CDF) representing cumulative contribution of individual /24 source network blocks for both total packets and bytes are shown using datasets A-[1,2], B-[1,2], C-[1,2]. Sorted with highest contributors on the left.
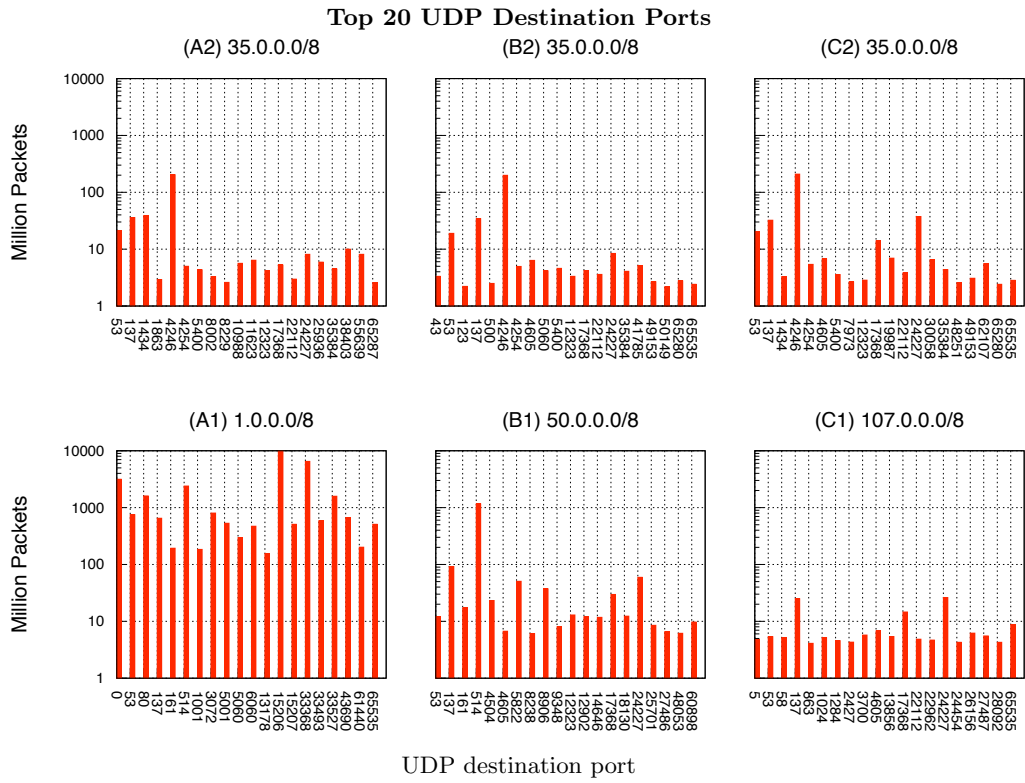
Top 20 UDP Destination Ports

Figure 8: **Spatial analysis of Internet Background Radiation. The top 20 UDP destination ports are shown using datasets A-[1,2], B-[1,2], C-[1,2].**
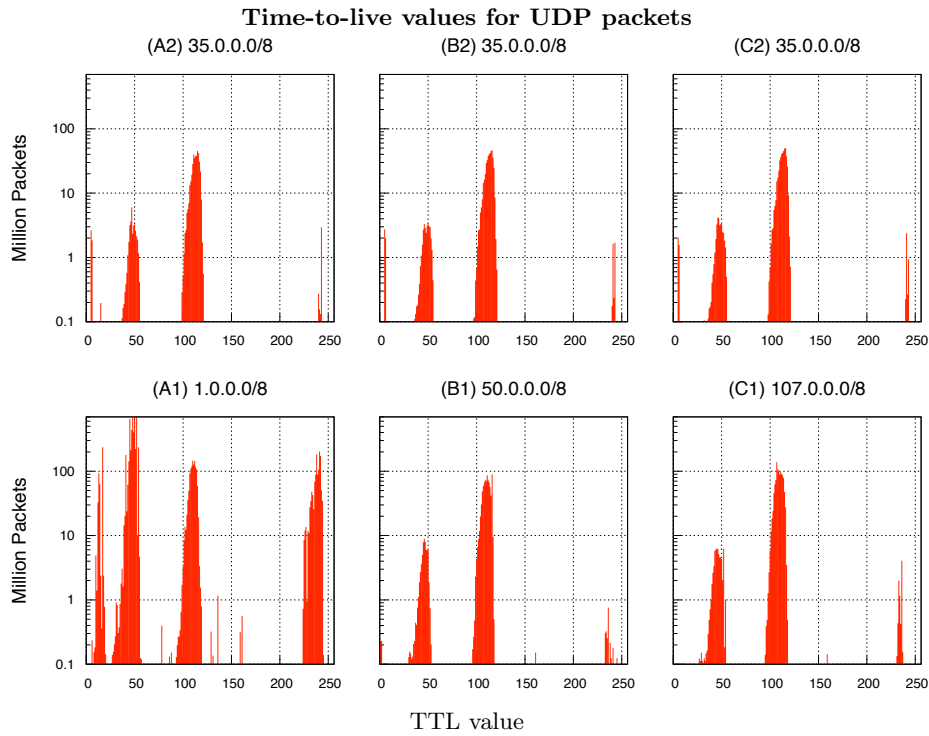
Time-to-live values for UDP packets

Figure 9: **Spatial analysis of Internet Background Radiation. The distribution of TTL values for UDP traffic is shown using datasets A-[1,2], B-[1,2], C-[1,2].**

subnet (/24)	%pkts	%bytes
1.1.1.0	44.0	58.7
1.4.0.0	16.7	9.4
1.0.0.0	10.6	6.2
1.2.3.0	2.0	8.6
1.1.168.0	0.6	0.3
1.10.10.0	0.3	2.4
1.1.0.0	0.2	0.1
1.0.168.0	0.2	0.1
1.0.1.0	0.1	0.1
1.2.168.0	0.1	0.1
total	74.8	86.0

Table 8: Top 10 /24 destinations in 1.0.0.0/8

2-byte prefix	packets(M)	bytes(M)	% pkts	%bytes
0x8000	17093	3658152	71.0	75.8
0x8008	4213	901639	17.5	18.6
0x8004	1791	138671	7.4	2.8
0x8012	605	51917	2.5	1.0
0x8080	334	71540	1.3	1.4
0x8088	5	1283	0.0	0.0
0x8003	2	232	0.0	0.0
0xa012	0.5	28	0.0	0.0

Table 9: Top 8 RTP Payloads

dport	packets(M)	bytes(M)	%pkts	%bytes
33368	6511	515323	55.7	55.8
514	2114	165388	18.0	17.9
33527	1582	124775	13.5	13.5
3072	803	63827	6.8	6.9
33493	588	46752	5.0	5.0
721	50	3974	0.4	0.4
17055	18	1418	0.1	0.1
33437	7	517	0.0	0.0
570	4	303	0.0	0.0
58689	3	232	0.0	0.0

Table 10: Top 10 destination ports for UDP traffic to 1.4.0.0

block due to a specific model of a DSL modem. Figure 9 shows this source OS distribution.

5. POLLUTION

In the previous section we discussed that many of the large scale differences observed between announced blocks and our reference unused block were clustered in a small portion of destination or source address space. We call this significant nonuniform behavior *address space pollution*. Unlike previous observations of differences between unused blocks [21] we show that increasingly these differences are the result of environmental factors (e.g., misconfiguration, location), rather than algorithmic factors.

For example, in analyzing the significant difference between A-1 and A-2 we find that the top 10 /24 destinations receive 75% of the total packets. Table 8 shows the 10 /24 destination subnets that received the most packets in 1.0.0.0/8. These blocks observe significant non uniform traffic that is primarily the result of misconfiguration. Example classes of these misconfigurations include: network protocol vulnerabilities, misconfigured network servers, services, and devices, misconfigured attack tools, misconfigured peer-to-peer network software, and various other software programming bugs. In the following sections we explore these interesting sources of pollution in more depth.

5.1 UDP SIP Traffic to 1.1.1.1

In analyzing the destination IP addresses of traffic to 1/8, we discovered a relatively high amount of hot-spots, compared to other /8s we analyzed during roughly the same time. We found that packets with a destination address in the 1.1.1.0/24 subnet made up 44.0% of packets and 58.7% of bytes in the 1.0.0.0/8 traffic over the entire week analyzed. Further analysis of this traffic showed that the vast majority of this traffic was UDP packets to 1.1.1.1, port 15206. This highly specific subset of traffic made up 34.2% of packets and 49.3% of bytes to the entire 1.0.0.0/8 subnet. We found that 71.0% of packets (75.8% of bytes) of traffic to UDP 1.1.1.1:15206 started with a payload of 0x8000. An additional 17.5% of packets (18.6% of bytes) started with 0x8008, and 7.4% of packets (2.8% of bytes) started with 0x8004. Previous analysis by RIPE suggested that this traffic was a trojan, however a couple of SIP blogs [28] [29] revealed that this traffic was likely RTP streams resulting from malicious SIP INVITE packets sent to vulnerable servers. These INVITE packets request that the server dial a telephone number, and send the resulting audio stream back to an IP address and port specified in the INVITE packet. In

this case, the packets were created to have the stream sent to 1.1.1.1:15206. Since these RTP streams are done over connectionless UDP, no response from 1.1.1.1 was necessary for us to receive these RTP streams in our capture.

We were able to isolate a handful of these streams, and using Wireshark, extract the unencrypted audio encoded in these streams. The audio file consists of a series of reorder tones (fast busy), followed by an automated voice stating: "The number you have dialed is not in service. Please check the number and try again." Each RTP stream contributes about 40-50 packets per second (80kbit/s), with an average of more than 5000 streams sending to 1.1.1.1:15206 simultaneously at any time.

Analysis of the TTL values for this traffic reveals 4 separate default TTL values that appear to be the original TTL for the packet. Due to the nature of TTL values decreasing at every hop on its way from the source to our darknet, we expect the received TTL value to be approximately 10-20 less than the starting value. From this, we infer 4 distinct starting values, of 32, 64, 128 and 255. As different operating systems choose different default TTL values, we can conclude that this traffic is likely coming from a cross-platform software application.

5.2 DNS Traffic to 1.4.0.0

The second highest hot-spot in 1.0.0.0/8 is 1.4.0.0/24, receiving 16.6% of packets and 9.4% of bytes over the week-long capture. We observed that almost all of this traffic was UDP packets to 1.4.0.0, on a handful of destination ports.

Inspection of these packets reveals them to be validly constructed DNS queries. Over the week-long capture, 6,536,254 unique source IPs contributed to this traffic, mostly from a handful of ASNs. Using nmap, we were able to determine that most of these hosts were in fact ASUS DSL modems. We suspect that these modems have either a hardware or software misconfiguration that causes them to use 1.4.0.0 on certain non-standard ports as a DNS server. We are un-

Domains
hotelnikkohimeji.co.jp
x.myspacecdn.com
wirelessdigest.typepad.com
th411.photobucket.com
www.google.com

Table 11: Example A record lookups to 1.4.0.0

sure why these modems would send DNS queries on such non-standard, yet concentrated few ports. Analysis of the domains that are being looked up to 1.4.0.0 shows a mix of domains that users are not likely to look up directly — for example, content distribution network domains from popular sites like myspace or youtube. This suggests that 1.4.0.0 is not the sole DNS server for the misconfigured box, as these CDN domains are looked up upon retrieval of the main site — which requires a successful DNS resolution in the first place.

5.3 Iperf Traffic to 1.2.3.4

Roughly 1.8% of all packets measured to the 1.0.0.0/8 network were UDP packets with a destination port of 5001. This port is the default port commonly used by the network testing application *iperf*. These packets were all sent to the 1.2.3.4 destination IP address. This pollution by itself accounted for roughly 10Mbps of traffic and was observed to be originating from fewer than 100 unique sources.

5.4 IP Address Byte-Order Misconfiguration

Though not singly a high contributor of packets, there are three seemingly out-of-place /24s in the top 10 destination subnets for traffic captured to 1.0.0.0/8. These are 1.1.168.0/24, 1.0.168.0/24 and 1.2.168.0/24. Further analysis of packets with 168 in the third octet reveals almost all of these packets to be to 1.x.168.192. Interestingly, this is the popular RFC1918-space gateway address 192.168.x.1 in host-byte order for little-endian machines. Furthermore, these packets are UDP to destination port 80, and contain the same data. The UDP length field specifies 1 byte of payload data (9 bytes - 8 byte UDP header), and the data that follows the UDP packet is always 0x31. While we do not know of a specific device or program that would produce such packets, it is possible a program is sending raw packets, and not doing a proper htonl() on its destination IP. Another explanation is an embedded device (or other platform) using a big-endian architecture is running an incorrectly ported network application from a little-endian system, and could still be performing the byte-ordering switch. We also see the same UDP packets (destination port 80, same payload data) sent to 1.1.0.10, and 1.0.0.10, which are 10.0.1.1 and 10.0.0.1 (other popular RFC1918 gateway addresses) in little endian.

5.5 Syslog to 50.153.199.194

In dataset B-1 (50/8), an interesting hotspot on UDP destination port 514 caught our attention. Subsequent analysis revealed that the hotspot destination /24 is 50.153.199.0/24, receiving 3.8% of packets, and 6.7% of bytes for the entire 1 week capture. Almost all of these were to a single IP, 50.153.199.194, UDP port 514. Closer examination revealed that these packets are all syslog messages originating mainly from IPs in the .de (Germany) TLD. Many of these mes-

```
<31>Mar 11 23:59:57 Muck-TS.CheckUserDir -
check snapshot in
\\muck-ts\david\archive\user\
10088000\todo
<31>Mar 11 23:59:59 srv-tobit Creating
Watchdog (C:\Programme\COSYNUS\
BlackBerry4Dv\TXEngine4BB.
watchdog.txt)
<31>Mar 12 00:02:01 vm-eco_cosynus
archive \\eco-online-serv\david\
archive\system\cosynus\bb4dv\bcc\
archive.dat is empty
```

Table 12: Three example syslog messages received on 50.153.199.194 UDP/514. The PRI part (<31>) corresponds to security/authorization messages with a debug (lowest) level severity.

sages contain combinations of the strings "david", "tobit" and "cosynus".

We contacted Cosynus, a German software publishing company, that owns the Cosynus Blackberry connector, to verify the source of this traffic. This software allows Tobit software's David to run on the blackberry. Tobit's David software is a multimedia application available for Windows that consists of E-mail, speech, fax, RSS and instant messaging.

Cosynus confirmed that some of their customers entered "062" as the first octet of an IP address during configuration. This is interpreted as octal, resulting in the traffic being sent to 50.153.199.194 instead of 62.153.199.194. Cosynus offered to firewall the victim IP address in future client updates, which should allow this /24 netblock to become useable. Table 12 shown an example of the messages received.

5.6 eMule to 35.206.63.212

In each week-long capture of 35/8 (datasets A-2, B-2, C-2), 1.0% of packets and 0.9% of bytes were for a specific destination IP 35.206.63.212. Of this, roughly 83% was UDP packets, mostly with 18 bytes of data. The first two bytes of these packets are the same (e3 9a), followed by 16 bytes of varied data. These packets, when interpreted with the eDonkey protocol, a peer-to-peer file sharing protocol, indicate that these are "Get Sources" packets, used to fetch a seed-list for a given file hash.

We found that 35.206.63.212 was listed as a fake server on the official eMule forum [30], confirming that this traffic is in fact peer-to-peer traffic using the eDonkey protocol.

5.7 μTorrent traffic

Recently, newer versions of the popular bittorrent client μTorrent have implemented a new protocol called Micro Transport Protocol (μTP), to provide better congestion control for bittorrent connections. μTP runs on top of UDP, allowing the μTorrent application to perform congestion control on its bittorrent streams independent of TCP congestion control.

In each of our datasets we were able to observe approximately 4 MBit/s of μTP traffic. This traffic consists primarily of 33 byte packets, to various UDP ports. The destination IPs seem to be varied as well, though the source IP for a given destination IP appears to send packets only to that destination IP. In other words, each source IP sends several packets to only a single destination IP.

Each 33 byte UDP packet starts with 12 bytes of changing

data, followed by 21 bytes of data that is the same in all packets: 7f ff ff ff ab 02 04 00 01 00 00 00 08 00 00 00 00 00 00 00 00. We were able to confirm that this 21 byte sequence occurs in packets generated during a torrent download with a recent version of μTorrent.

5.8 Responding to Pollution

In responding to these unique forms of pollution, we adopt the philosophy of the original authors of [10] to build classifiers for the removal of the unwanted traffic. Our purpose, however, is different. Unlike [10], we are not attempting to reduce traffic in order to build a scalable active responder, our goal instead is to determine the usability of a network block and the utility of any cleanup effort. Given the concentrated nature of much of this pollution, these classifiers need not be complicated and may often simply filter based on net-blocks. Table 8 lists the top 10 most polluted /24 network blocks in 1/8. Together these account for 75% of all packets and 85% of all bytes of pollution traffic to the 1/8 network block. Based on this, APNIC has already proposed the following filters for one of their recently allocated blocks based on this philosophy: "The following /24s should be withheld from general allocation by APNIC: 1.0.0.0/24, 1.1.1.0/24, 1.2.3.0/24, 1.4.0.0/24, 1.10.10.0/24. If further investigation reveals that the traffic to any of these /24s abates to a normal background level in the future, then these addresses would be returned to the APNIC unallocated address pool at that time."

We also know that a cleanup of some of the pollution to 50/8 is possible, as we have contacted the software vendors responsible for the misconfiguration. Their response was helpful in eliminating the pollution being caused by their software. The pollution traffic which we were able to identify as SIP was originating from only a few thousands of sources and the iperf pollution traffic was originating from only a few tens of unique sources. These as well as other similar types of pollution can likely be minimized with a sustained cleanup effort.

6. CONCLUSIONS AND FUTURE WORK

In this paper we have taken a fresh look at Internet background radiation. While today's Internet radiation continues to be as ubiquitous, variegated, and dynamic as was uncovered in the initial study [10], we note several important changes since the last study including: rapid growth outpacing the growth in productive network traffic, reduced contribution from the exploit ports reported in previous work, and trends toward increasing SYN and decreasing SYN-ACK traffic. In examining traffic across address blocks, we note that significant differences exist, but often these differences are clustered in a handful of network blocks. We use the term *Internet address pollution* to refer to this significant nonuniform behavior that is primarily the result of environmental rather than algorithmic factors. We examine several case studies in Internet address pollution and offer some specific suggestions for filtering the most egregious of these blocks.

We would like to develop a more systematic process that passes all newly allocated network blocks through an evaluation phase, where potential usability of the block is assessed. Any recipient of a tainted network block will be unfairly penalized in terms of bandwidth costs and actual usability of their address space. Therefore, it is important that any such network blocks that are found be placed on a well-known temporary watch list, publicized to the network operator community. Furthermore, address block tainting is not simply a result of Internet background radiation, but can also occur when address space is re-allocated from one user to another. Prior ownership of an address block might have resulted in that network block being placed on various spam or botnet blacklists, thereby affecting its usability. We would like to work with the RIRs and the Internet operator community to understand how a simple process can be put in place which attempts to mitigate the affect of address block pollution, either via the dissemination of widespread filters or via a sustained cleanup effort. A related issue is of the study address pollution beyond unallocated or re-allocated blocks, including the problem of discovering and removing pollution in already allocated and routed blocks. While some work has been done on examining what is normal for background radiation [7] and how to maximize visibility to unused space within a network [31], we believe this is a very interesting future area of study.

We view this work as yet another step in understanding the interesting phenomenon of Internet background radiation. As such, we hope to encourage follow-on work in this area by making the datasets used here available via the PREDICT [16] dataset archive. We acknowledge that the current policy for PREDICT limits the use of these datasets to research carried out at US institutions. We will continue to work with existing internal PREDICT efforts to broaden participation beyond its initial scope. Until these efforts bear fruit, the authors encourage interested non-US researchers to contact them directly to investigate alternative paths to data sharing, including jointly seeking institutionally approved research studies (e.g., via research oversight organizations such as the US institutional review board system and similar systems abroad) and/or through entering into researcher-specific data sharing agreements.

7. REFERENCES

[1] D. Moore, V. Paxon, S. Savage, and Shannon C. Inside the Slammer Worm. *In Proceedings of IEEE Security and Privacy*, Jun 2003.

[2] D. Moore, C. Shannon, and J. Brown. A Case Study on the Spread and Victims of an Internet Worm. *In Proceedings of ACM SIGCOMM Internet Measurement Workshop*, Nov 2002.

[3] Michael Bailey, Evan Cooke, David Watson, Farnam

Jahanian, and Jose Nazario. The Blaster Worm: Then and Now. *IEEE Security & Privacy*, 3(4):26–31, 2005.

[4] D. Moore, G. Voelker, and S. Savage. Inferring Internet Denial of Service Activity. *In Proceedings of the 2001 USENIX Security Symposium*, Aug 2001.

[5] M. Bailey, E. Cooke, D. Watson, F. Jahanian, and N. Provos. Practical Darknet Measurement. *In Proceedings of the 40th Annual Conference on Information Sciences and Systems (CISS)*, Mar 2006.

[6] D. Moore, C. Shannon, G.M. Voelker, and S. Savage. Network Telescopes. *Cooperative Association for Internet Data Analysis - Technical Report*, 2004.

[7] M. Bailey, E. Cooke, D. Watson, F. Jahanian, and N. Provos. Towards Understanding Distributed Blackhole Placement. *In Proceedings of the 2nd Workshop on Rapid Malcode (WORM)*, Oct 2004.

[8] V. Yegneswaran, P. Barford, and D. Plonka. On the Design and Use of Internet Sinks for Network Abuse Monitoring. *In Proceedings of the Symposium on Recent Advances in Intrusion Detection*, Sep 2004.

[9] M. Bailey, E. Cooke, D. Watson, F. Jahanian, and N. Provos. The Internet Motion Sensor - A Distributed Blackhole Monitoring System. *In Proceedings of the 12th Annual Network and Distributed System Security Symposium (NDSS)*, Feb 2005.

[10] R. Pang, V. Yegneswaran, P. Barford, V. Paxson, and L. Peterson. Characteristics of Internet Background Radiation. *In Proceedings of the 4th ACM SIGCOMM Conference on Internet Measurement*, Oct 2004.

[11] Moheeb Abu Rajab, Jay Zarfoss, Fabian Monrose, and Andreas Terzis. A multifaceted approach to understanding the botnet phenomenon. In *IMC '06: Proceedings of the 6th ACM SIGCOMM on Internet measurement*, pages 41–52, New York, NY, USA, 2006. ACM Press.

[12] Evan Cooke, Farnam Jahanian, and Danny McPherson. The Zombie roundup: Understanding, detecting, and disrupting botnets. In *Proceedings of the Steps to Reducing Unwanted Traffic on the Internet (SRUTI 2005 Workshop)*, Cambridge, MA, July 2005.

[13] E. Eugene Schultz. Where have the worms and viruses gone?–new trends in malware. *Computer Fraud & Security*, 2006(7):4–8, 2006.

[14] Craig Labovitz, Scott Iekel-Johnson, Danny McPherson, Jon Oberheide, and Farnam Jahanian. Internet Inter-Domain Traffic. In *Proc. ACM SIGCOMM (To Appear)*, 2010.

[15] Geoff Huston. The changing Foundation of the Internet: confronting IPv4 Address Exhaustion. *The Internet Protocol Journal*, September 2008.

[16] Protected Repository for the Defense of Infrastructure Against Cyber Threats. `http://www.predict.org`.

[17] Michael Bailey, Evan Cooke, Farnam Jahanian, Niels Provos, Karl Rosaen, and David Watson. Data Reduction for the Scalable Automated Analysis of Distributed Darknet Traffic. *Proceedings of the USENIX/ACM Internet Measurement Conference*, October 2005.

[18] Sushant Sinha, Michael Bailey, and Farnam Jahanian. Shedding light on the configuration of dark addresses.

In *Proceedings of Network and Distributed System Security Symposium (NDSS '07)*, February 2007.

[19] John Bethencourt, Jason Franklin, and Mary Vernon. Mapping Internet sensors with probe response attacks. In *Proceedings of the 14th USENIX Security Symposium*, Baltimore, MD, August 2005.

[20] Moheeb Abu Rajab, Fabian Monrose, and Andreas Terzis. On the effectiveness of distributed worm monitoring. In *Proceedings of the 14th USENIX Security Symposium*, Baltimore, MD, August 2005.

[21] Evan Cooke, Z. Morley Mao, and Farnam Jahanian. Hotspots: The root causes of non-uniformity in self-propagating malware. In *Proceedings of the International Conference on Dependable Systems and Networks (DSN'2006)*, June 2006.

[22] Abhishek Kumar, Vern Paxson, and Nicholas Weaver. Exploiting underlying structure for detailed reconstruction of an internet-scale event. *Proceedings of the USENIX/ACM Internet Measurement Conference*, October 2005.

[23] Mark Allman, Vern Paxson, and Jeff Terrell. A brief history of scanning. In *IMC '07: Proceedings of the 7th ACM SIGCOMM conference on Internet measurement*, New York, NY, USA, 2007.

[24] Route Views Project. University of Oregon Route Views Project. *http://archive.routeviews.org/*, Aug 2010.

[25] RIPE (Rẽl'seaux IP Europẽl'ens). RIS Raw Data. *http://www.ripe.net/projects/ris/rawdata.html*, Aug 2010.

[26] RADb: Merit network inc. routing assets database. `http://www.radb.net/`.

[27] B. Kantor, S. Savage, R. Wesson, B. Enright, P. Porras, V. Yegneswaran, J. Wolfgang, and Castro S. Conficker/Conflicker/Downadup as seen from the UCSD Network Telescope - Feb 2009. *http://www.caida.org/research/security/ms08-067/conficker.xml*.

[28] S. Gauci. RTP Traffic to 1.1.1.1 - Feb 2010. *http://blog.sipvicious.org/2010/02/rtp-traffic-to-1111.html*.

[29] S. Eivind. usken.no - VoIP news! - Feb 2010. *http://www.usken.no/2010/02/sip-scanning-causes-ddos-on-ip-1-1-1-1/*.

[30] Adrian Mariẽśo. Fake Servers List - Official eMule-Board - Apr 2010. *http://forum.emule-project.net/index.php?showtopic=139609&st=60*.

[31] Evan Cooke, Michael Bailey, Farnam Jahanian, and Richard Mortier. The Dark Oracle: Perspective-Aware Unused and Unreachable Address Discovery. In *Proceedings of the 3rd Symposium on Networked Systems Design & Implementation (NSDI '06)*, pages 101–114, San Jose, California, USA, May 2006.