An Empirical Study of Spam Traffic and the Use of DNS Black Lists

Jaeyeon Jung
jyjung@csail.mit.edu

Emil Sit
sit@csail.mit.edu

MIT Computer Science and Artificial Intelligence Laboratory
32 Vassar Street, Cambridge, MA 02139, USA

ABSTRACT

This paper presents quantitative data about SMTP traffic to MIT's Computer Science and Artificial Intelligence Laboratory (CSAIL) based on packet traces taken in December 2000 and February 2004. These traces show that the volume of email has increased by 866% between 2000 and 2004. Local mail hosts utilizing black lists generated over 470,000 DNS lookups, which accounts for 14% of all DNS lookups that were observed on the border gateway of CSAIL on a given day in 2004. In comparison, DNS black list lookups accounted for merely 0.4% of lookups in December 2000.

The distribution of the number of connections per remote spam source is Zipf-like in 2004, but not so in 2000. This suggests that black lists may be ineffective at fully stemming the tide of spam. We examined seven popular black lists and found that 80% of spam sources we identified are listed in some DNS black list. Some DNS black lists appear to be well-correlated with others, which should be considered when estimating the likelihood that a host is a spam source.

Categories and Subject Descriptors

C.2.2 [**COMPUTER-COMMUNICATION NETWORKS**]: Network Protocols

General Terms

Measurement

Keywords

Spam traffic, DNS black lists, Zipf-like distribution

1. INTRODUCTION

Unsolicited bulk email ("spam") is a major fraction of mail sent and received on the Internet [9]. One of the commonly deployed techniques to block spam is address-based filtering, using which one can refuse to accept mail from hosts that are believed to send spam. Once identified, the IP address of a host engaged in spam delivery is registered in centrally maintained databases. This database is made available via the Internet DNS and hence often called *domain name system black lists* (DNSBLs). Mail recipients can query this database using standard DNS lookups and deny any mails from the hosts that are listed in the database.

Over the past several years, an increasing number of DNSBLs are offering various lists of IP addresses based on some criteria — for example, each IP address may be an open relay, a virus source, or an actual spam source caught by a spam trap. In this paper, we first provide a classification of DNSBLs, which shows that DNSBLs vary widely in terms of their focus and the coverage of IP addresses and one should factor in this difference when using DNSBLs to assign the mail's spam score.

We then analyze SMTP and DNSBL traffic, based on a pair of packet traces taken at the MIT Computer Science and Artificial Intelligence Laboratory (CSAIL). Data from two traces, taken in 2000 and 2004, are used to provide a comparison of traffic patterns over time. We extensively analyzed the 2000 trace in our previous work [11]. This paper re-analyzes the 2000 trace for SMTP specific activity and compares it with a trace taken in February 2004.

In analyzing these traces, this paper makes the following contributions:

- We characterize the incoming SMTP workload and resulting outgoing DNSBL workload between 2000 and 2004. In our February 2004 (2000) trace, 73,601 (4,553) remote hosts initiated 776,363 (28,909) incoming SMTP connections to 183 (151) hosts at CSAIL per day. In response, local mail hosts utilizing some form of black lists issued 472,644 (2,153) DNS lookups, which accounts for 14.09% (0.4%) of all DNS lookups recorded at the border gateway of CSAIL.
- We observed the existence of machines that receive many SMTP connection attempts despite not running mail transfer agents; we use these machines as spam traps and show that the distribution of spam sources is Zipf-like in 2004 but not in 2000. The heavy-tailedness of Zipf-like distribution results from the significant increase of the number of spam sources that made less than 10 connections to the monitoring network in 2004. The considerable increase of these low-profile spam sources may cause DNSBLs difficulty identifying spam sources.
- We show that a black list that aggressively blocks "open relays" includes 80% of hosts that sent mail to our spam traps. Using conditional probabilities, we show that some lists, like SORBS [5] and DSBL [3], are highly correlated to each other, which should be factored in estimating the likelihood that a given host is a spam source.

The rest of this paper is structured as follows. In Section 2, we provide a brief overview of the different classes of DNS black lists and how they are used to combat spam. We describe our collection

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.
IMC'04, October 25–27, 2004, Taormina, Sicily, Italy.
Copyright 2004 ACM 1-58113-821-0/04/0010 ...$5.00.

methodology and data set in Section 3. We analyze the membership and relationships of several black lists in Section 4. We briefly summarize our results and future work in Section 5.

2. BACKGROUND

While "spam" itself is difficult to define precisely, most people know what spam is when presented with it. Loosely, it is any unsolicited email that is sent to a large number of recipients. In some cases, this mail may be commercially oriented (*e.g.*, an advertisement), but much of mail termed "spam" is actually not commercial, but rather some scam or virus.

2.1 Mail Life-cycle

Spam, as a subset of email, is technically delivered just like any other piece of email. After it is composed, the sender hands off the mail to a local *mail transfer agent* (MTA) that assumes responsibility for delivering the mail to the final destination. This MTA is known as the injection point. If the sender is part of a large network, the local MTA may transfer the mail to additional hosts within the same administrative domain, allowing aggregation of mail within an organization. Each host is said to *relay* the mail.

Eventually, one of the MTAs in the sender's organization (possibly the local injection MTA) will identify a host responsible for receiving mail for the recipient's domain, and relay the mail to that host. This host is known as the domain's *mail exchanger*, and is specified via a DNS `MX` record. At this point, the MTA on the recipient's mail exchanger may relay the message internally within the organization, ultimately arriving at an MTA that performs *delivery*, storing the message in a local mailbox for the recipient. In some cases, the recipient may choose to have his local MTA automatically perform some filtering to select the appropriate local mailbox for a given message.

MTAs relay mail between each other using the Simple Mail Transfer Protocol (SMTP) [12]. This protocol allows the receiving MTA to weakly authenticate and authorize the sending MTA and reject the transaction if desired. Additionally, MTAs annotate each mail message with a *received header* which identifies both parties involved in the exchange, and the time at which it occurred. This header adds accountability and aids in debugging mail problems.

2.2 Black Lists

Originally, spammers sent spam just like regular email from their own accounts. However, with a limited number of actual senders [16], it became easy to block the addresses of the senders and also to trace them back to origin server and request that the account of the sender be revoked. To get around this "problem" and to get free bandwidth and CPU cycles, spammers started relaying mail through hosts that would accept responsibility for delivering anyone's mail — these hosts are called *open relays*. More recently, spammers are believed to be using compromised client machines to directly deliver mail [13]. We call any machine used to deliver spam a *spam source*.

Various lists of spam sources are maintained, with data culled from dedicated and casual participants, and are centrally administered and disseminated. These lists are called *black lists* because people typically choose to refuse all mail originating from a listed host: recipients or mail transfer agents consult one or more of these databases to determine if the remote peer is believed to be a spam source.

We classify black lists based on two axes. First, we identify the focus of the list (*e.g.*, RFC compliance, open relays, actual spam sources, country/ISP net blocks, virus sources, verified senders/spammers, etc.). Second, we consider how aggressively addresses are added to the lists. The level varies from active probing (*e.g.*, testing for open relays), to manual entry (*e.g.*, verifying hosts before adding), and to passive monitoring (*e.g.*, only adding hosts that send mail to spam traps). We summarize 31 widely used black lists in Table 1. Compared to in 2000 when there were only 5 black lists actively used within CSAIL, there are now many more black lists in use.

While some of these lists can be downloaded locally and used with tools such as OpenBSD's `spamd` [10], most people access these lists via the Domain Name System (DNS). For this reason, these black lists are typically called DNS black lists. The DNS provides an efficient and scalable way for providing geographically distributed clients with the ability to test whether a given host is a member of the set of black-listed hosts. The test for whether a host $a.b.c.d$ is on a given list is simple: a DNS lookup is done for the `A` record of the name $d.c.b.a$`.blacklist-name`. This is similar to the `in-addr.arpa` style of reverse DNS lookups.

2.3 DNSBL Clients

DNS black list clients are typically either built into MTAs or separate programs invoked when mail is filtered into mailboxes during final delivery.

MTAs check DNS black lists to determine whether the party relaying the mail is listed. Multiple lists may be checked, as each black list is targeted differently. At this point, the MTA can choose to refuse to accept the mail, end the transaction with an appropriate error code, and terminate the connection. This activity happens *before* the mail is accepted locally.

Sometimes, filtering is performed *after* mail is accepted; the entire body has been received and programs like SpamAssassin [14] run a variety of tests to determine if the mail is spam. Each test influences the mail's "spam score"; when performing final delivery, the user's filter can use this score to determine if the mail is spam. SpamAssassin contains tests that parse the "Received" headers of a given piece of mail and check if the mail has transited any blacklisted hosts. These tests have been available since the 0.3 release of SpamAssassin in June 2001, and have expanded from checking three black lists to checking thirty-five lists hosted at eleven sites (in the 2.6.3 release).

3. DATA

We collected traces in December 2000 and in February 2004 at a link connecting MIT Computer Science and Artificial Intelligent Laboratory (CSAIL) to the rest of the world. The trace contains bidirectional TCP SYN/FIN/RST packets, outgoing DNS query packets and incoming DNS response packets. The details of the collection methodology are described in our previous paper [11]; in the intervening time, the network topology remained unchanged but the access link was upgraded from a single symmetric 100Mbps link to two parallel 100Mbps links.

3.1 DNS Traffic

DNSBL related traffic now is a considerable source of DNS lookups from within CSAIL. As in our previous work [11], we define a DNS lookup as a series of packets, related to a single domain name from the initial query to the final answer. We identify DNSBL related lookups using the following two-step process.

- We first obtain a set of DNSBLs that our local hosts queried by examining all the DNS `A` queries for a host of the form `[1-255].[1-255].[1-255].[1-255].name`. We manually checked each name and excluded those that were not black lists.

Table 1: Classification of DNSBLs used within CSAIL: BLs marked with † only appeared in 2000; those with ∗ appeared in both 2000 and 2004. All others appeared only in 2004

Focus	Maintenance	Blacklists
Known spammers	Conservative	`sbl.spamhaus.org`,
Open proxies	Conservative	`opm.blitzed.org`,
Open relays	Conservative	`rbl.maps.vix.com`∗, `list.dsbl.org`, `multihop.dsbl.org`, `relays.mail-abuse.org`∗, `relays.osirusoft.com`, `relays.visi.com`, `relays.orbs.org`†, `relays.ordb.org`,
Open relays	Aggressive	`unconfirmed.dsbl.org`, `dnsbl.sorbs.net`,
Virus/Exploit attackers	Aggressive	`xbl.spamhaus.org`, `cbl.abuseat.org`,
Country/ISP netblocks	Aggressive	`{argentina,att,...}.blackholes.us`, `dul.maps.vix.com`†, `dul.dnsbl.sorbs.net`, `dynablock.easynet.nl`, `blackholes.easynet.nl`, `dialups.mail-abuse.org`,
RFC violators	Mix	`{dsn,ipwhois,whois,abuse,postmaster,bogusmx}.rfc-ignorant.org`,
Mix	Mix	`sbl-xbl.spamhaus.org`, `bl.spamcop.net`, `dnsbl.njabl.org`,
Commercial	Commercial	`hil.habeas.com`, `sa-hil.habeas.com`, `query.bondedsender.org`, `sa-other.bondedsender.org`, `sa-trusted.bondedsender.org`,
Unknown	Unknown	`rbl.dorkslayers.com`†, `rbl.debian.net`,

Table 2: Statistics for DNSBL-related DNS traffic

	7 Dec 2000	8 Dec 2000	19 Feb 2004	20 Feb 2004
Total DNS lookups	572,936	519,422	3,428,816	3,282,231
DNSBL-related lookups	2,238 (0.39%)	2,069 (0.40%)	487,077 (14.21%)	458,211 (13.96%)
Total root lookups	38,184	35,369	301,042	286,324
DNSBL-related root lookups	2 (0.01%)	0 (0.00%)	45 (0.01%)	49 (0.02%)
Total DNS packets	1,196,042	1,091,130	5,618,333	5,245,229
DNSBL-related packets	2,396 (0.20%)	2,169 (0.20%)	519,415 (9.25%)	488,859 (9.32%)

- Using the resulting 31 DNSBLs, we identify a lookup as DNSBL related if a DNS query name is of the form listed above or ends with `dsn.rfc-ignorant.org` (a black list that operates on domain names).

Table 2 summarizes DNSBL related DNS lookups seen in each trace.[1] While DNSBL related lookups are only less than 0.4% of the total number of lookups in December 2000, they are now about 14% of all DNS lookups, which makes up about 9% of the entire DNS traffic in February 2004.

DNSBLs do not affect the root servers much — most DNSBLs have well replicated name servers, and each client has cached contact information of an authoritative server for each DNSBL. As shown in Table 2, DNSBL related root lookups account for less than 0.02% of the total root lookups in 2004 even though DNSBL related lookups are almost 14% of total DNS lookups. The disproportionate increases between DNSBL related root lookups and DNSBL lookups is due to name server caching, which we discussed in detail in our previous paper [11]. In fact, the time-to-live (TTL) for most DNSBLs' authoritative name server records is on the order of 1 to 2 days. This long TTL allows clients to directly connect to one of the DNSBL's authoritative name servers in resolving DNSBL queries, bypassing the root-servers.

3.2 Inbound SMTP Traffic

We use Bro [15] to pre-process raw TCP packet traces, which provides a succinct summary of each TCP connection including a source IP address, a destination IP address, bytes transmitted by each end point, and the connection status. We consider only established or rejected inbound Simple Mail Transfer Protocol (SMTP) connections for this analysis. While our trace includes unanswered SYN packets, those are most likely generated by port scanners, and not by the remote mail servers that we wish to characterize in this study.

3.2.1 SMTP Growth

Inbound SMTP traffic to CSAIL increased significantly between December 2000 and February 2004. Comparing daily inbound SMTP traffic to CSAIL between those two collection periods, we see that the email traffic volume increased by 8.7 times (row 6 in Table 3). The number of successful inbound SMTP connections[2] increased by 11.9 times (row 5 in Table 3). We expect that most of these inbound connections are likely spam-related. This hypothesis is supported by statistics from a major anti-spam company indicating that as of April 2004, 64% of all mail sent over the Internet is spam, up from 48% in May 2003 [9].

We find that more than 70,000 distinct remote hosts attempted to make an SMTP connection to less than 200 local CSAIL hosts in February 2004 (row 1 in Table 3). This is 15.2 times more remote hosts than observed in December 2000 when fewer than 5,000 remote hosts attempted to deliver emails through the local hosts at CSAIL. Figure 1 shows the number of inbound connection attempts (including rejected connection attempts) per each remote host initiating them. Both plots from December 2000 and February 2004 traces show a very long tail; over 80% of remote hosts made less than 10 connection attempts across all datasets, 40% of which made only one connection attempt per day. We note that this large number of distinct remote hosts initiating *inbound* SMTP connections results in significant amount of *outbound* DNS traffic to DNS black lists.

[1]We noticed that IPv6 related lookups increased tremendously in February 2004 datasets, most of which failed to elicit an answer. Given the query names of these IPv6 lookups, we believe that they are "sysqueries" from BIND8/9 asking for IPv6 addresses of authoritative name servers.

[2]We consider a connection attempt successful if the connection is established and then later terminated with FIN packets.

Table 3: Summary of inbound SMTP traffic

		7 Dec 2000	8 Dec 2000	19 Feb 2004	20 Feb 2004
1	Remote hosts initiating inbound connections	4,334	4,773	76,676	70,526
2	Local hosts answering inbound connections	117	186	193	173
3	Total attempted inbound connections	29,303	28,515	787,231	765,496
4	Total rejected inbound connections	4,513	2,962	463,097	437,143
5	Total successful inbound connections	24,790	25,553	324,134	328,353
6	Total Kbytes transmitted	270,655	207,826	2,203,239	2,417,132

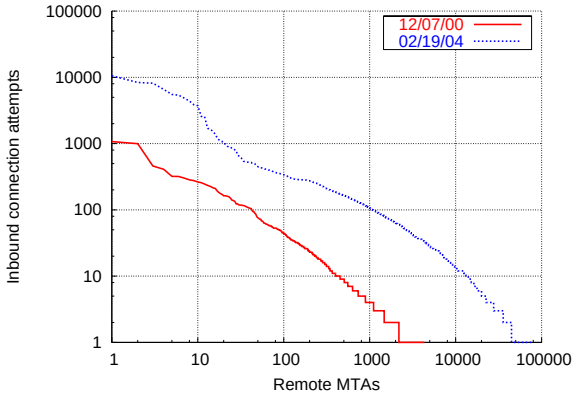

Figure 1: Number of inbound connection attempts that each remote MTAs made to the monitoring network. Those remote MTAs are sorted by the decreasing number of connection attempts. Both X and Y axes are in log scale

3.2.2 *Spam Sources*

To our surprise, we observe that approximately 56% of inbound SMTP connection attempts seen in February 2004 dataset are rejected by local hosts; that is, 56% of all inbound SMTP connections are to local hosts that are not running a Mail Transfer Agent (MTA) and hence immediately reset inbound SMTP connections. This percentage is up from less than 13% in 2000.

We believe that the remote hosts initiating these connections are likely to be spammers since legitimate email users do not send emails unless there is reason to believe that the recipient host exists and is willing to accept their emails.[3] Indeed, we find that among the local hosts receiving inbound SMTP connection requests, 38% of them had not accepted any inbound SMTP connection requests nor initiated any outgoing SMTP connections. These local hosts thus function effectively serve as "spam traps", allowing us to identify likely spam sources. This is supported by anecdotal evidence: we know that one of those hosts had formerly been relaying emails for a certain domain that is not in use any more. Despite now no longer running an MTA, that host received 141,917 SMTP connections from 7,732 remote hosts on 20 February 2004. Setting up a server speaking SMTP, recording all the email messages coming to the host during 24 hours, and checking them using SpamAssassin [14], we corroborated that all the messages sent to the host were spam.

We classify a remote host as a spam source if the remote host attempted to connect local spam traps and there were no successful mail transfers initiated from any local host to that remote host. The second criterion was needed to filter out a few mailing lists that may have been led to the spam traps from outdated subscriber information. Table 4 summarizes the results.

[3]One possibility for legitimate attempts to these hosts would be misconfigured `MX` records. We did not investigate this possibility.

Table 4: We classify a remote host as a *spam source* if the remote host attempted to make an SMTP connection to a local host that was not running a Mail Transfer Agent and hence rejected the connection attempt. Such a local host effectively serves as a *spam trap* to catch spammers walking through a list of recipients including obsolete ones

	7 Dec 2000	8 Dec 2000	19 Feb 2004	20 Feb 2004
Local spam traps	19	89	90	59
Remote spam sources	79	66	7,970	8,780

One interesting observation is that connection attempts from this set of remote hosts is Zipf-like in 2004 but not so in 2000. The distribution of the number of connections from remote spam hosts is plotted in Figure 2. The straight line, a/x^{α}, fitting the tail of the distribution in 2004 shows that the tail can be well approximated with Zipf-like distribution [8], with $\alpha = 1.03$.

The heavy-tailedness of Zipf-like distribution results from the dramatic increase of the number of spam sources that made less than 10 connections to the monitoring network: In 2000, there are only 27 such spam hosts, contributing 2.6% to the all incoming spam traffic, while now the number increased to 4,832 spam hosts making up 12.5% of the all incoming spam traffic. Assuming that those low-profile spam sources are hard to identify, the significant increase of low-profile spam sources can affect the effectiveness of DNSBLs in reducing spam traffic.

4. BLACK LISTS AND SPAM SOURCES

This section explores the effectiveness of DNSBLs in tracking active spam sources and how different black lists are quantitatively related to one another.

We selected seven popular DNSBLs to ascertain how many of the likely spam hosts in our trace appeared in black lists at the time of our analysis. While this analysis was performed after the traces were collected and may lead to some inaccuracies, we have no other way of definitively knowing if a host was listed since not all local MTAs used DNSBLs. We use the set of spam sources (described in §3.2) and check whether each host is listed in a database using DNS queries. Table 5 shows that over 80% of possible spam sources connecting to hosts at CSAIL in 2004 are still listed in one or more black lists in April 2004.

It is interesting to note from Table 5 that 34 of likely spam sources in December 2000 are still listed in one of the black lists queried in 2004. It is because that either the 34 listed hosts are still engaged in spam delivery (although we see no activity from those hosts in our 2004 trace), or black lists are delinquent in removing obsolete entries from database. The former indicates a difficulty in stopping spammers even after they have been identified. The lat-

Table 6: Conditional probability $\Pr[B|A]$ matrix for 7 black lists based on Feb 2004 datasets

A \ B	abuseat	dsbl	opm	rfc-ignorant	sorbs	spamcop	spamhaus
abuseat	1.00	0.88	0.06	0.16	0.77	0.26	0.02
dsbl	0.16	1.00	0.02	0.14	0.77	0.05	0.01
opm	0.69	0.99	1.00	0.27	0.75	0.39	0.00
rfc-ignorant	0.11	0.54	0.02	1.00	0.57	0.04	0.04
sorbs	0.13	0.69	0.01	0.14	1.00	0.04	0.06
spamcop	0.74	0.76	0.09	0.15	0.67	1.00	0.09
spamhaus	0.02	0.04	0.00	0.07	0.44	0.04	1.00

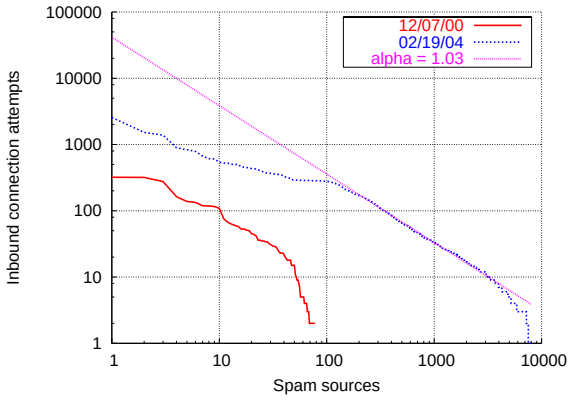

Figure 2: The number of inbound connection attempts made by spam sources; spam sources are sorted by decreasing number of connection attempts. The set of hosts observed in 2004 follows a Zipf-like distribution, but the 2000 set does not.

Table 5: Spam sources listed in DNSBLs

	Dec 2000	Feb 2004
`cbl.abuseat.org` [2]	0	1,401
`list.dsbl.org` [3]	5	7,624
`opm.blitzed.org` [1]	0	122
`ipwhois.rfc-ignorant.org` [4]	25	2,030
`dnsbl.sorbs.net` [5]	3	8,529
`bl.spamcop.net` [6]	0	496
`sbl.spamhaus.org` [7]	2	1,123
Total unique hosts black listed	34	11,521
Total spam sources	100	14,090

ter may prevent those hosts previously used as spam sources from legitimately sending mail.

We observe that the coverage of each black list widely varies; highly conservative lists, such as SpamHaus [7], list about 87% fewer hosts than the aggressive lists, like SORBS [5]. There is also little overlap between known spam sources and open relay lists — SpamHaus [7] lists only 6% of hosts that appear in SORBS [5]. In contrast, different open relay lists *are* highly correlated: 77% of hosts listed in DSBL [3] also appear in SORBS [5]. Table 6 shows conditional probabilities of each pair of black lists that we tested. For given two black lists A and B, $\Pr[B|A]$ is defined as:

$$\frac{\text{\# of spam sources that both B and A listed}}{\text{\# of the spam sources that A listed}}$$

This conditional probability matrix could be used to weigh the importance of different black lists in scoring spam (*e.g.*, in SpamAssassin [14]). For instance, if the remote host is listed in both SpamHaus [7] and SpamCop [6], which have very little overlap, it could be considered more likely to be spam. If false positive data were available, Table 6 could also be used to choose the member of a set of highly correlated lists with the least false positives.

5. SUMMARY AND FUTURE WORK

The work in this paper was motivated by a dramatic increase in the use of DNSBLs at CSAIL from December 2000 to February 2004. DNSBL lookups now account for nearly 14% of all DNS lookups from our network, up from 0.4% less than four years ago. This is due to an increase in the volume of mail traffic, as well as a large increase in the number of public black lists.

By using spam traps, we observed that the activity distribution of spam source hosts was *not* Zipf-like in 2000, but is now Zipf-like with $\alpha = 1.03$. 80% of these hosts were found to be listed in black lists, two months after the trace was collected.

Our future work includes better understanding the distribution of active spam sources, assessing a false positive rate and a miss rate of DNSBLs, and devising a prescriptive algorithm in using a set of DNSBLs to block spam.

Acknowledgments

We gratefully acknowledge Hari Balakrishnan and Robert Morris for their continued support and guidance. Dave Andersen was invaluable in helping classify black lists. We also thank Dave Andersen, Nick Feamster, Stuart Schechter, Alex Snoeren, and anonymous reviewers for their comments on the previous drafts that improved the paper.

6. REFERENCES

[1] Blitzed open proxy monitor list. `http://opm.blitzed.org/`. Last accessed 15 Aug 2004.

[2] Composite blocking list. `http://cbl.abuseat.org/`. Last accessed 15 Aug 2004.

[3] Distributed server boycott list. `http://dsbl.org/main/`. Last accessed 15 Aug 2004.

[4] Rfc-ignorant.org. `http://www.rfc-ignorant.org/`. Last accessed 15 Aug 2004.

[5] Spam & open relay blocking system. `http://www.us.sorbs.net/`. Last accessed 15 Aug 2004.

[6] Spamcop. `http://www.spamcop.net/`. Last accessed 15 Aug 2004.

[7] The spamhaus project. `http://www.spamhaus.org/`. Last accessed 15 Aug 2004.

[8] L. Breslau, P. Cao, L. Fan, G. Phillips, and S. Shenker. On the implications of Zipf's law for web caching. Technical Report CS-TR-1998-1371, University of Wisconsin, Madison, April 1998.

[9] Brightmail. Spam percentages and spam categories. `http://www.brightmail.com/spamstats.html`. Last accessed 15 Aug 2004.

[10] Theo de Raadt. OpenBSD `spamd`. `http://www.openbsd.org/cgi-bin/man.cgi?query=spamd`. Last accessed May 2004.

[11] Jaeyeon Jung, Emil Sit, Hari Balakrishnan, and Robert Morris. DNS performance and the effectiveness of caching. *IEEE/ACM Transactions on Networking*, 10(5), October 2002.

[12] J. Klensin. *Simple Mail Transfer Protocol*, April 2001. RFC 2821.

[13] Ben Laurie and Richard Clayton. "proof-of-work" proves not to work. In *Proceedings of the The Workshop on Economics and Information Security*, Minneapolis, MN, May 2004.

[14] Justin Mason. Spamassassin. `http://www.spamassassin.org/`. Last accessed April 2004.

[15] Vern Paxson. Bro: a system for detecting network intruders in real-time. *Computer Networks*, 31(23–24):2435–2463, 1999.

[16] Evan I. Schwartz. Spam wars. *Technology Review*, pages 32–39, July/August 2003.