Spam and the CAN-SPAM Act

Matt Bishop

Department of Computer Science[1]
University of California, Davis
1 Shields Ave.
Davis, CA 95616-8562

bishop@cs.ucdavis.edu
phone: (530) 752-8060

1. Introduction

The goal of the CAN-SPAM Act of 2003 [13] (called "Act" or "CAN-SPAM Act" in this paper) is to prevent senders of commercial electronic mail (called "email") from misleading recipients about the origin and content of email, and to give recipients the ability to decline to receive additional commercial email from the same source. The Act provides penalties for violating its provisions, and the provisions describe what senders of commercial email must, and must not, do. The Act also provides several definitions related to these provisions.

The Federal Trade Commission is required to report to Congress on the effectiveness of the provisions of the Act, of its enforcement, and recommend any appropriate changes to the provisions of the Act ([13], §10(a)). This paper is part of the technical input for that review.

This paper examines three questions:

1. How does international email affect the effectiveness of the Act?
2. How can consumers and ISPs deal with spam? What technologies are available? In particular, how can people be protected, or protect themselves, from receiving unwanted sexually explicit material? Part of the problem here is that those companies that can be identified will likely comply with the CAN-SPAM Act, because it is the law; but those companies that cannot be identified may not comply. Hence the enforcement provisions of the Act are not enough, because not every non-compliant company can be caught and prosecuted. Technological means are a necessary adjunct.
3. How effective and enforceable are the provisions? Because this paper focuses on the technology, and not legal or public policy matters, we discuss the state of technology that can be used to comply with, and enforce, the provisions. As the reader will see, many technological and scientific impediments make the effectiveness difficult to measure.

To summarize the conclusions of this paper briefly, the Act improves the ability of consumers and ISPs to block unwanted spam, provided the companies sending the spam comply with the Act. Technology can reduce the amount of spam received from companies that do not comply with the Act, but technology is of limited effectiveness in those cases where the senders take steps to evade or bypass this technology.

---

1. Affiliation included for identification purposes only. The opinions expressed in this paper are those of the author alone, and not necessarily the opinions of any other person, entity, agency, or organization.

The next section presents a quick summary of background material; the appendices cover this material in far more technical detail. The following three sections discuss the three questions posed above. The conclusion summarizes the key points of this paper and evaluates how the technical issues affect the enforcement of the Act. Several appendices discuss some common techniques for sending spam, countermeasures, and difficulty of measuring the effectiveness of the Act.

2. Background

This section summarizes key aspects of electronic mail, and presents the technologies used to send spam, and countermeasures.

Throughout this paper, the term "email" refers specifically to messages sent using electronic mail. In particular, messages sent using instant messaging protocols such as AIM are not considered. Further, the term "header information" refers to the sender and recipient information, and any routing information, contained in the header lines of the message. It does not include other header lines such as the subject line. See section 8.1 for more details.

2.1. Key Aspects of Electronic Mail

The way that electronic mail works is described in section A.1 of Part III of the FTC's report on the Do Not Email Registry [5]. Section 8.2 also discusses this process, using slightly different terminology.

The following are key points about the Internet, and electronic mail, that are relevant to the analysis of the Act. The reader interested in more details should see section 8.

- The term "MTA"[2] refers to the mail server, and the term "UA"[3] refers to the program that the user uses to compose or read the mail. The UA sends the message to the first MTA, and receives it from the last MTA, and the MTAs forward the message from the originating address (called the "source") to the mail server for the recipient's address (called the "destination"). The MTAs communicate using a protocol called SMTP[4] [10]; the UAs use this protocol to communicate with the MTAs. The receiving MTA is called a "mail server"; the UA, or other MTA, passing the email to the mail server is called a "client." See section 8.2 for more details.
- When an MTA is asked to accept an email, the email will contain an originating address. The MTA can accept the address as correct without any checking, can check that the client is on the same computer as the originating address, and/or check that the originating address is authorized to send email to the MTA.
- Email headers can be forged or altered.
- Determining the computer from which a particular email was sent is difficult because the originating computer's address may have been forged, or it may have been reassigned since the email was sent, or a zombie (see Sections 2.2 and 9.5) may have sent the email from a compromised system.

---

2. "MTA" stands for "Message Transfer Agent" or "Message Transport Agent."
3. "UA" stands for "User Agent."
4. "SMTP" stands for "Simple Mail Transfer Protocol."

2.2. Technologies for Sending Spam

The following table summarizes the technologies spammers use to send spam. Each technology is discussed in more detail in Section 9 (see the section named in the table entry).

Table 1: Technologies Used to Send Spam

Type of Technology	*What It Is*
bulk email	send out large quantities of spam (§9.1)
joe job	send spam with existing, but incorrect, return addresses (9.2)
relaying through third parties	send spam through an intermediate MTA that accepts messages from untrusted parties (§9.3)
break-ins to send spam	compromise a computer system, and send spam from it (§9.4)
zombie	a program placed on a system that sends spam, in conjunction with hundreds or thousands of other copies on other computers (§9.5)

2.3. Technologies for Controlling Spam

The following table summarizes technologies for blocking or inhibiting spam. Each technology is discussed in more detail in section 10 (see the section named in the table entry).

Table 2: Anti-Spam Technologies

Type of Technology	*What It Does*
authentication services	authenticate the user, computer, or domain of origin (§10.1)
authorization services	confirm that the sending computer is authorized to send email for that domain (§10.2)
reputation services	determine whether the sending user, computer, or domain is trustworthy enough to accept mail from (§10.3)
challenge-response	request the sender of email to verify that he or she intended to send email to the putative recipient (§10.4)
port blocking and rate limiting	disallow a set of hosts from sending email through an MTA, restrict the amount of email that a host can send through an MTA (§10.5)
preventing relaying	allow the MTA to accept email only from trusted computers (§10.6)
spam filtering	analyze the incoming email to determine if it is spam (§10.7)
image blocking	block images that email references until reader requests their downloading (§10.8)
keywording	look for a specific word or words in the message, and if found allow the message through; this is used to enable trusted senders to force delivery of messages, even if they are flagged as spam (§10.9)

2.4. Measuring Effectiveness of the Act

The term "effective" could have many definitions, all of which are measured differently. This section outlines several such definitions; the reader should see section A for a more detailed discussion.

Underlying all these definitions is the observation that the goal of the Act is, apparently, not to *reduce* the volume of spam or to *eliminate* spam.[5] It is to make the flow of spam more manageable for ISPs and consumers. We shall call spam that complies with the Act "commercial email," and spam that does not coply with the Act "non-compliant commercial email.". "Spammers" refers to people who send non-compliant commercial email, and "anti-spam technologies" refers to technologies designed to handle both non-compliant commercial email and compliant but unwanted commercial email.

1. "Effective" means "companies are able to comply with the requirements of the Act." A simple analysis of each requirement will confirm this.
2. "Effective" means "companies complying with the Act make their messages easily detectable as commercial email." Here, the requirements of the Act must be compared to anti-spam technologies, to determine if those requirements result in making commercial email more easily identifiable. In fact, the requirements of the Act have this effect.

Taking these two definitions together, one clear benefit of the Act is that it serves as a "best practices" guide to sending bulk commercial email. This means that companies who wish to send commercial email, and who plan to identify themselves as the senders, can comply with the requirements of the Act, and consumers and ISPs who wish to block this type of email can do so.

An interesting question is whether the companies *do* comply with the Act. As noted above, companies that can easily be identified will likely comply with the Act, if only to avoid legal consequences; but companies that are difficult to identify may well not comply (and in fact many do not). Part of the problem is that the effectiveness of the Act in this sense relies upon effective enforcement procedures, and those procedures are legal, not technical. They are beyond the purview of this report. The most that can be said is that the Act clearly lays out requirements that commercial and transactional email must meet to be legal, providing a basis for determining whether someone is, or is not, complying with the Act.

3. "Effective" means "the amount of commercial email that violates the Act has been reduced." The problem with measuring this aspect is that one cannot obtain a representative sample of email from before the Act was passed, nor could one obtain such a sample of email from the current Internet. Basically, any such samples will be biased and therefore would provide biased, incorrect answers to this question. Section 11 discusses the problem in more detail.
4. "Effective" means "as a result of the Act, anti-spam technology has advanced." While anti-spam technology has advanced, it is not clear that the Act drove the advances. But it is clear that anti-spam technology has been more widely deployed, and improved, since the Act was passed.

---

5. Were that the goal, the Act could simply have banned spam outright, or banned certain types of spam. Instead, it specifically legalizes spam that meets the criteria laid out in the Act.

3. International Considerations

This section discusses how international email affects the effectiveness of the Act.

First, the Act does not define "international email." It could be electronic mail that originates in a foreign country, that transits MTAs in one or more foreign countries, or that has the final MTA in a foreign country[6].

Secondly, regardless of which definition is selected, the problems that plague determining effectiveness also arise here. How does one measure the proportion of commercial email, and non-compliant commercial email, that is "international?" Exactly the same issues of sampling as discussed in definition 3 of section 2.4, and in section 11, apply here. One common methodology used to gather statistics determines the IP address of the origin (or of transited MTAs) of the email, and then to which country that IP address belongs. These point to ISPs that fail to restrict the flow of non-compliant commercial email. This methodology does not show how much commercial email or non-compliant commercial email comes from those foreign countries.

It can be difficult to determine whether email originated in (or transited) foreign countries. A sender in a foreign country may forge an IP address assigned to a computer in the United States; similarly, a sender in the United States may forge an IP address of a foreign computer. Email originating in the Unites States may be relayed through a foreign open relay. In these cases, the origin of the email is likely to be classified incorrectly.

This creates a problem for determining how international electronic mail affects the effectiveness of the Act. Section 2.4 discusses the difficulty of determining the effectiveness of the Act. Complicating the issues discussed in that section is the problem of cataloguing all emails as "international" or "domestic". As discussed above, the origin of the email cannot be identified with complete accuracy. The *purported* point of origin can be determined. This may, or may not, cause the classification of the email as "domestic" or "international" to be correct.

Hence, from a technical point of view, it is not possible to answer the question of the effectiveness of the Act with respect to international email in a meaningful way.

Complicating the matter is that, even when a foreign spammer is identified, the hosting ISP may decline to remove or block the user. Presuming the ISP is foreign, the difficulty of enforcing the Act increases. Also, foreign registrars who have assigned an IP address (or set of IP addresses) to the spammer's computers may decline to remove the assignment, or may ignore requests to do so, or local authorities may also decline to act.[7] Non-technical remedies, such as international treaties, seem to be the best manner of handling this situation.

4. Sexually Explicit Material

This section considers the question of how effective the Act has been in protecting people from unwanted sexually explicit material.

First, if the Act is followed, the Act should be effective in blocking pornographic or obscene commercial email. According to the Act, §5(d)(3), one need only block emails with the clearly identifiable marks or notices that inform the recipient that the commercial electronic mail

6. That is, the "last hop" of the email is from an MTA in a foreign country to a UA in the United States.
7. This is sometimes called "bulletproof hosting" because the United States authorities cannot force the registrar to remove the IP address assignment, and in some instances the local authorities will not act.

contains sexually oriented material. Further, under the FTC's power to make regulations to support the Act, the "brown paper wrapper rule" requires that the email must be constructed in such a way that the recipient must take some affirmative action—such as clicking on a link—before the image is shown. If these rules are not followed, other techniques must be used.

Two techniques are common. The first is filtering on content as described in section 10.7. The second uses the fact that sexually oriented material may contain images. The ISP, or the mail reading program, should be configured to block images in electronic mail as discussed in section 10.8. Then the recipient may read words, but will not see sexually explicit images unless he or she requests that the images may be shown. Both these technologies have improved over the last two years, but (as stated above) whether the passing of the Act influenced their improvement cannot be determined.

It is difficult to evaluate the effectiveness of the Act in protecting people from unwanted pornography. The problems of measuring effectiveness discussed in section 2.4 arise here. Unlike the question of international emails, the definition of "sexually explicit material" can be applied to the email without having to deduce the email's origin. So the contents of the available email messages can be analyzed to draw conclusions about the sample, and by inference about the effectiveness of the Act.

5. The Civil Provisions

This section reviews the CAN-SPAM Act's civil provisions.[8]

Measuring effectiveness of these provisions has the same problems as measuring the effectiveness of the Act in general. Section 2.4 summarizes these, and section 11 discusses these at length. We frame the discussion around three aspects of effectiveness. First, how hard is it for companies to comply with the Act? Second, can anti-spam technology detect and handle commercial email that complies with the Act? Third, given a sample of email, could one determine whether the email in the sample complies with the requirements of the Act? The first two follow definitions of "effectiveness," and the third tells how difficult measuring effectiveness from an ideal sample of email messages would be.

Whether companies do comply with the act requires an analysis of the data gathered by the FTC staff, and is not within the scope of this report.

5.1. Section 5(a)(1): Prohibition of false or misleading transmission information

This provision speaks to the contents of the header fields.

If the sender wishes to comply with the Act, the sender need only send the email from a computer that the sender is authorized to use for that purpose, and put the sender's email address (or that of someone who has authorized use of their email address) in the "from" line. So senders can comply with the Act.

The Act also requires the sender to include some information that enables anti-spam technology to detect the commercial email, and discard it if appropriate. This is more problematic. Only simple filtering on the "from" line can be effective, because most email addresses cannot be

8. At the direction of FTC staff, this report focuses on the civil provisions only. However, much of what is said applies to the criminal provisions as well.

verified as accurate. If the sender is complying with the Act, filtering based on domain name or email address will work once that domain or address is found to be a sender of commercial email. However, the user may see some compliant commercial email before making this determination. If the sender is not complying with the Act, the "from" field can be forged to pass filters. Using blocking technology based on IP address is also problematic, because there is no way to tell from a specific message whether access to an IP address was obtained illegitimately (in the sense of §5(a)(1)(A)). Relaying messages through protected computers may, or may not, be visible from the headers; thus, one may or may not be able to tell whether the forwarding occurs. Thus, anti-spam technology can be effective with respect to filtering on "from" addresses, but absent external information cannot determine whether the address is materially misleading.

As stated in section 2.1, for many of the headers it is not possible to check for false or misleading information. One can look for inconsistencies in the headers that may indicate falsity. For example, consider the "Received:" lines added by MTAs as the message is sent from its origin to its destination. The computers named in successive lines should match, as discussed in section 8.4; if not, that may indicate a forgery. Similarly, if the letter contains a "Sender:" header line, its value can be compared to that of the "From:" line. However, there are legitimate reasons why these values may differ, so mismatches do not mean the data in the headers is false. Thus, one may not be able to determine whether a sample email message complies with this provision of the Act.

In summary, anti-spam technology can help check that this requirement is met, and filter compliant messages, but requires external information. But the requirement that the header information not be misleading is one that technology is ill-suited to check.

5.2. Section 5(a)(2): Prohibition of deceptive subject headings

This provision forbids making the subject header deceptive. A subject header is deemed deceptive if the sender knows, or should know, that the subject heading would be likely to mislead the recipient about the contents of the message.

If the sender wishes to comply with the Act, the sender need only supply a non-deceptive subject header. So a sender can easily comply with this provision of the Act.

Anti-spam technology can filter compliant but unwanted commercial emails based upon the subject header. The only difficulty here might be that the subject, and contents of the message, are written in a natural language such as English, French, Chinese, or Russian—and natural languages are ambiguous. For example, one commercial email might have a subject header of "Have you heard the news?" and a body describing a sexually explicit act. The sender might argue that the subject is not deceptive because some aspect of the sexually explicit act might be news to some recipients; but the vast majority of people would consider the subject deceptive as nowhere does it mention that the body of the message discusses a sexually explicit act. However, adaptive spam filters have proven effective, and while they are not perfect, if the user obtains a preconfigured spam filter, few such messages should get through.

If the sender does not comply with the Act, anti-spam technology may be able to detect the non-compliant commercial email based on the subject header. If the subject header is innocuous, though (for example, "Meeting tomorrow" as the subject header for a non-compliant commercial email suggesting the recipient meet to discuss buying medicine from a web site), filtering based solely on the subject line will not work.

Because of the ambiguity of natural languages and the potential for differing views on the descriptiveness of subject headers, it is unlikely that sample email messages could be automatically processed to detect violations of this provision.

In summary, anti-spam technology can filter compliant messages based on subject headers, but may not be able to detect that non-compliant messages have deceptive subject headers.

5.3. Section 5(a)(3): Inclusion of return address or comparable mechanism in commercial electronic mail

This provision requires that the mail message include a mechanism for opting out of future messages from that sender at the electronic mail address to which the letter was sent. This may be a functioning return address or a menu or list that allows the user to opt out.

If the sender wishes to comply with the Act, the sender need only supply a return address or menu allowing the recipient to opt out. So a sender can easily comply with this provision of the Act.

Email messages can be automatically checked to see if they include a return address. The return address can be validated by sending a message to it; however, that merely shows that email to the return address does not bounce, and would not validate that an "opt out" message was received and honored. For example, if the return address is forged (a "joe job;" see section 9.2), the address would pass the above test yet not comply with the provision. The mail could similarly be scanned for some phrase indicating a list or menu to opt out, but again, the list or menu would need to be tested. So anti-spam technologies could check for the presence (or absence) of return addresses, but may not be able to check that the return address is valid in a timely manner.

A similar problem arises for automated checking of emails taken from a sample.

Hence anti-spam technology can detect emails that are obviously non-compliant because they omit a return address. But other methods, involving human analysis, must be applied to validate compliance with this requirement.

5.4. Section 5(a)(4): Prohibition of transmission of commercial email after objection

This provision requires that the sender of the mail message honor a request by the recipient not to receive any more email from that sender at that recipient's address. The sender is also forbidden to have anyone send email to that address on his or her behalf.

If the sender wishes to comply with the Act, the sender will have to delete the recipient's address from the sender's database, and not distribute that address any further. So a sender can easily comply with this provision of the Act.

As complying with this provision means no further commercial email will be sent, compliant companies will not send any more commercial email from that email address, and anti-spam technology will not be involved. For non-compliant companies, anti-spam technology can inhibit non-compliant commercial email by adding the email address to a blacklist (or deleting it from a whitelist). This assumes the anti-spam technology detects the opting out; if it does not, then the anti-spam technology will allow future commercial email from that email address through. Further, there is no way to determine whether the email comes from an address that the user opted out of from the mail message itself, because any attempt to opt out does not affect the current message. A honeypot technique could be used to determine whether the sender honors the opt out request, and whether such a request causes non-compliant commercial email from another source

to be sent to the recipient (the implication being the spammer gave away, or sold, the address in the opt out request).

A honeypot is a system that is monitored but takes no authorized actions. The hope is that an attacker will attack that system. The attacker's activity is monitored and recorded for later analysis. In this context, the honeypot would be created by creating a series of mailboxes, called "decoy addresses". These addresses should be random names that a spammer is unlikely to guess. The owner then waits for the first commercial email to arrive; if necessary, the owner may visit a web site and "seed" one decoy address in one or two places.

When any of these mailboxes receives commercial email, the owner immediately sends an "opt out" message. If the same sender sends another commercial email message after the opt out period (10 business days; Act, §7704(a)(4)(A)(I)), then this spammer is violating the Act. If the number of commercial email messages climbs dramatically, then the spammer may have sold the address to others.

A variant of this test is to opt out using one of the other decoy addresses. This has the advantage that the address used to opt out will not have been exposed anywhere, and so if commercial email messages arrive then the spammer has almost certainly passed the address on to another person.

In the absence of external information (specifically, whether the recipient opted out of commercial email from the sending email address, and when the opt out occurred), there is no way for automatic analysis of a sample of email to detect violations of this provision.

Thus, anti-spam technology can help inhibit non-compliant commercial email sent in violation of this provision of the Act, and can help determine whether a sender is violating this provision of the Act.

5.5. Section 5(a)(5)(A) Inclusion of identifier, opt-out, and physical address in commercial electronic mail

This provision requires that commercial email include an identifier indicating the message is an advertisement, an opt out notice, and a valid postal address for the sender.

If the sender wishes to comply with the Act, the sender need only include the required information. As with the other provisions discussed so far, complying with this provision of the Act is straightforward.

If commercial email complies with this provision, anti-spam technology can detect the identifier, and block or filter the message. Hence anti-spam technology can be effective with compliant emails. But if the email message does not comply with this provision of the Act, anti-spam technology may or may not filter the message appropriately.

Determining whether an email message complies with this provision requires checking that it includes all three elements. The main obstacle is the "valid postal address" requirement. Determining whether there is a postal address in the message is straightforward. Determining whether it belongs to the sender is more complicated, and automated tools may not be able to do this. Determining whether it is valid requires an investigation to determine whether the sender can receive email at that address. This cannot be done without physical investigation of the address. Thus, validating that a commercial email meets this provision requires investigators to check out the physical address. This is time-consuming, and will require considerable resources, and almost certainly cannot be done using an automated mechanism.

To summarize, anti-spam technology can detect and handle messages that comply with this provision. But automated analysis of messages to determine compliance is impractical.

5.6. Section 5(b)(1) Address harvesting and dictionary attacks

This provision forbids using an automated means of generating possible email addresses for recipients of commercial email using permutations of names, letters, or numbers, and from obtaining addresses from sites with a notice that the service at that site will not give addresses to another party. This provision applies only if the commercial email being sent does not comply with the requirements in section 5(a) of the Act, so for the purposes of this discussion, we assume the commercial email does not meet those requirements.

If the sender wishes to comply with this provision of the Act, the sender need only refrain from using automated means to generate possible email addresses for recipients of non-compliant commercial email using permutations of names, letters, or numbers, and from obtaining addresses from sites with a notice that the service at that site will not give addresses to another party. The sender should send commercial email that does comply with section 5(a), or should obtain email addresses in a manual fashion.

Because this provision requires knowing how email addresses are generated or obtained, anti-spam technology cannot distinguish between compliance and non-compliance with this provision. The best it could do is determine that many emails have recipient addresses that appear to be randomly generated, but even that assumes the return address on those suspicious emails are the same—which is highly unlikely in this type of spamming.

5.7. Section 5(b)(2) Automated creation of multiple email accounts

This provision bars using automated means to obtain multiple email accounts to send commercial email that does not comply with the requirements in section 5(a). This provision applies only if the commercial email being sent does not comply with the requirements in section 5(a) of the Act, so for the purposes of this discussion, we assume the commercial email does not meet those requirements.

If the sender wishes to comply with this provision of the Act, the sender need only refrain from using automated means to obtain multiple email accounts to send non-compliant commercial email that does not comply with the requirements in section 5(a). The sender should send commercial email that does comply with section 5(a), or should obtain email addresses in a manual fashion.

Because this provision requires knowing whether a number of email accounts are obtained through automated means, and by the same person, anti-spam technology cannot distinguish between compliance and non-compliance with this provision.

5.8. Section 5(b)(3) Relaying email through unauthorized access

This provision bars relaying or retransmitting commercial email that does not comply with section 5(a) of the Act, without authorization to use the system doing the relaying or resending. This provision applies only if the commercial email being sent does not comply with the requirements in section 5(a) of the Act, so for the purposes of this discussion, we assume the commercial email does not meet those requirements.

If the sender wishes to comply with this provision of the Act, the sender need only obtain authorization to use the system from which the non-compliant commercial email is relayed or resent). As with the other provisions, this one is easy to obey: either send commercial email that complies with section 5(a), or obtain authorization before relaying or resending non-compliant commercial email.

If email complies with this provision, anti-spam technology can detect whether the email was resent or relayed, and act accordingly. Hence anti-spam technology can be effective with compliant emails. But, as explained in the next paragraph, anti-spam technology cannot distinguish between compliant and non-compliant emails, and so will either allow both through or filter both.

Determining whether an email meets this requirement requires checking that any addresses from which it was resent belong to people who did not authorize the resending, or whether any relaying computers were used without authorization. As the authorizations are not reflected in host names or email addresses, anti-spam technology cannot help determine compliance with this provision. However, it can determine whether commercial email was resent or relayed through hosts by examining the headers of the message, assuming the headers are not forged.

Hence anti-spam technology can determine whether an email was resent or relayed. Whether that is a violation of this provision requires investigation beyond what the anti-spam technology can do.

5.9. Section 5(d)(1) Warning labels on commercial emails containing sexually oriented material, or require the recipient to go elsewhere for the sexually oriented material

This provision requires that the sender of commercial email containing sexually oriented material either include in the subject header a notice that the email contains sexually oriented material, or write the email so that the sexually oriented material is not visible in the email and the recipient must take affirmative action to see it.

From the technical point of view, this requirement is similar to those discussed in sections A and B. If the subject header includes notice that the email is sexually oriented, then the comments under section 5.2 apply because the notice is simple to put in the subject line and is something that the anti-spam technology can filter. That part of section 5.5 that deals with the inclusion of an identifier applies here, too. Thus, the anti-spam technology can be effective in handling compliant messages of this type. Assuming the warning label is not obscure or difficult to understand, analyzing sample email messages can be automated to a large extent.

If the email points the user elsewhere for the sexually oriented material, anti-spam technology can filter the unwanted but compliant commercial email based on the message body. Similarly, non-compliant commercial email can be filtered with a high degree of accuracy because of the nature of the contents. Hence anti-spam technology can block or filter this type of unwanted commercial email.

In both cases, automated analysis of emails can flag non-compliant emails. Hence samples of emails can be analyzed automatically, in most cases.

5.10. Section 6(a) Selling using email in violation of section 5(a)(1)

This provision bars the promotion of merchandise or services if the seller knows that the email sent for that purpose was sent in violation of section 5(a)(1), received or expected an economic benefit from the emails, and took no action to prevent or report the sending of the email.

This provision adds non-technical issues to the questions raised by section 5(a)(1) and discussed in section 5.1 above. Thus, anti-spam technology plays no role in the detection or prevention of non-compliant emails.

6. Conclusion

One of the goals of this work was to discuss how to measure the effectiveness of the CAN-SPAM Act. A precise quantitative evaluation is not possible for the reasons discussed above. However, under several other definitions of "effectiveness," it appears the Act does require legitimate companies to take steps to make filtering commercial email more effective.

Of course, illegitimate companies avoid these steps, and so the provisions of the Act do not mandate technology that will limit those companies. Such technology does not yet exist, nor could it be feasibly deployed without making fundamental changes to the Internet. This would require an international effort, from all segments of all societies that use the Internet, and would in all likelihood cripple the exchange of information that is the Internet's greatest strength and, paradoxically, its greatest weakness.

Contrary to widespread belief, the CAN-SPAM Act does not ban spam; it regulates spam. It places restrictions upon commercial email, for example requiring a mechanism for opting out of future commercial emails from the same source. An interesting question is whether the CAN-SPAM Act would noticeably reduce the amount of spam sent through the Internet, assuming all senders of commercial email followed all its provisions. The answer is far from clear.

Moreover, as should be clear from this paper, the CAN-SPAM Act is a legislative solution to a technical problem. The nature of the spam problem precludes either a technological solution or a legislative solution. Further, the Internet being international affects the nature of any approaches to solving the problem. The most effective approaches will combine legislation and technology in ways that deal with the spam problem as an international, and not merely a national, problem.

7. References

[1] D. Balenson, "Privacy Enhancement for Internet Electronic Mail: Part III: Algorithms, Modes, and Identifiers," RFC 1423 (Feb. 1993).

[2] J. Chan, "SURBL—Spam URI Realtime Blocklists" (Oct. 5, 2005); available at http://www.surbl.org.

[3] J. Doll, "Spam Attack!" (undated), available at http://www.joes.com/spammed.html.

[4] R. Everett-Church, "The Spam That Started It All," *Wired News* (Apr. 13, 1999); available at http://www.wired.com/news/politics/0,1283,19098,00.html.

[5] Federal Trade Commission, "National Do Not Email Registry: A Report to Congress" (June 2004).

[6] S. Garfinkel, *PGP: Pretty Good Privacy*, O'Reilly and Associates, Sebastopol, CA (1994).

[7] S. Kent, "Privacy Enhancement for Internet Electronic Mail: Part II: Certificate-Based Key Management," RFC 1422 (Feb. 1993).

[8] J. Klensin, *Simple Mail Transfer Protocol*, RFC 2821 (Apr. 2001).

[9] J. Linn, "Privacy Enhancement for Internet Electronic Mail: Part I: Message Encryption and Authentication Procedures," RFC 1421 (Feb. 1993).

[10] J. Postel, *Simple Mail Transfer Protocol*, RFC 821 (Aug. 1982).

[11] Y. Rekhter, B. Moskowitz, D. Karrenberg, G. J. de Groot, and E. Lear, *Address Allocation for Private Internets*, RFC 1918 (Feb. 1996).

[12] G. Wittel and S. Wu, "On Attacking Statistical Spam Filters," *Proceedings of the First Conference on Email and Anti-Spam (CEAS)* (July 2004).

[13] "Controlling the Assault of Non-Solicited Pornography and Marketing Act of 2003 (CAN-SPAM Act of 2003)," Public Law 108–187, 117 STAT. 2699–2719 (Dec. 16, 2003).

[14] "DomainKeys: Proving and Protecting Email Sender Identity" (undated); available at http://antispam.yahoo.com/domainkeys.

[15] "How Do SPF and SenderID Work?" (undated); available at http://spf.pobox.com/how-works.html.

8. Appendix I: Details of the Email System

This appendix discusses technical aspects of some key words. It then describes how mail is handled, in order to provide a basis for certain terms such as relaying. After that, we discuss mail headers, and origin and IP addresses.

8.1. Definitions

In §3(6) of the Act, the term "electronic mail message" is defined to be "a message sent to a unique mail address." In this paper, treat this as meaning messages sent using electronic mail protocols such as SMTP. We explicitly exclude messages sent via AIM or other mechanisms that do not involve the use of message transfer agents or mail user agents.

In §3(8) of the Act, the term "header information" is defined as "the source, destination, and routing information attached to an electronic mail message". The standard definition of "header lines" is simply all lines from the first line of the mail message to the blank line dividing the header lines from the body of the message. In what follows, we use "header information" to refer to the sender address, recipient address, and any ancillary information indicating the routing of the message (usually contained in "Received:" lines). We explicitly *exclude* the subject header line unless noted otherwise.

8.2. The Structure of Electronic Mail Processing

This section briefly reviews how electronic mail works [8,10]. A program, called the *user agent* (abbreviated "UA" here), enables the user to compose the email. The UA then contacts a *message transport agent* (abbreviated "MTA"), and passes the email to the MTA. The MTA determines the next MTA to pass the email to, and does so. This continues until an MTA that can deliver the email to the recipient's mailbox is reached. That MTA then invokes a second UA to place the email into the recipient's mailbox. Figure 1 shows this process.

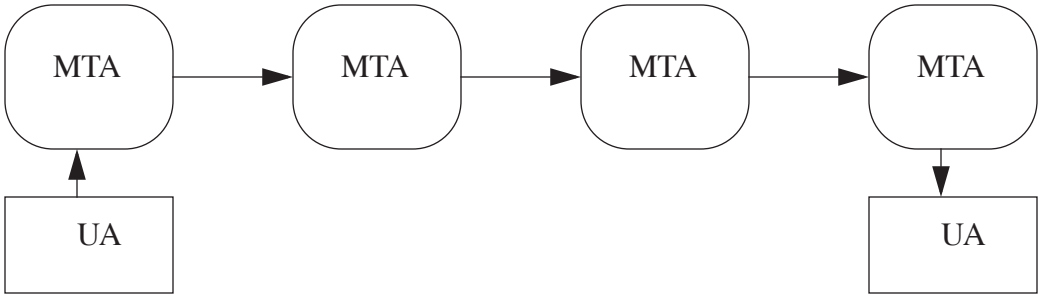

Figure 1. Programs involved in delivering electronic mail.

For example, suppose Alice is working on a PC named gnome, and she wishes to send a email to Bob. Bob works on the opposite coast, and his email address is bob@example.xyz. Alice composes her email using Outlook™. Outlook is a user agent. Outlook contacts the MTA at Alice's mail server, and passes the email to that server. The server's MTA contacts the MTA at example.xyz, and forwards Alice's email to it. That MTA delivers the email to Bob's mailbox. Figure 2 shows this process.

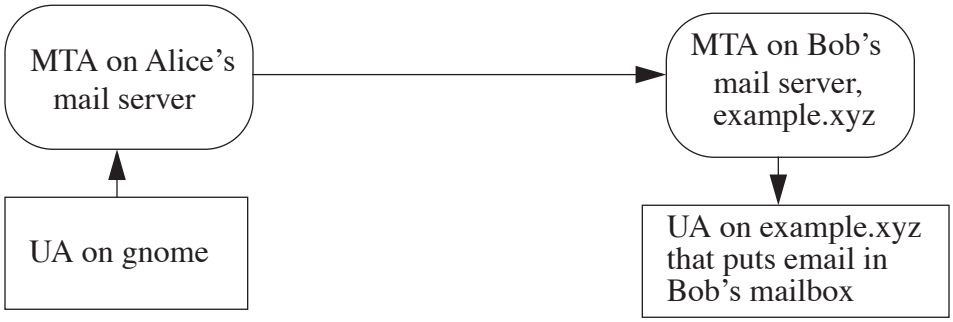

Figure 2. How Alice's mail gets to Bob

Sometimes the first MTA cannot contact the recipient's MTA directly. In this case, it forwards the email to an intermediate MTA. This intermediate MTA serves as a *mail relay*, forwarding the mail to the recipient's MTA, or to another intermediate MTA. In Figure 1, there are two intermediate relays along the path that the email takes.

Information about which computers have mail servers to which electronic mail can be sent is contained in the Domain Name System (DNS)[9]. Each entry in the DNS contains several records. The record involved in sending electronic mail is the MX record, which gives the IP address of the computer to which mail for a given computer is to be forwarded. For example, if "example.xyz" has an MX record with 192.168.4.210, then when Alice sends her email to Bob in the above example, Alice's mail server forwards the email to the computer with the IP address in the MX record corresponding to example.xyz, which is 192.168.4.210.

When Alice's MTA communicates with Bob's MTA, they use a protocol called the Simple Mail Transport Protocol (SMTP). As part of this protocol, Alice's MTA sends Bob's MTA the origin ("alice@gnome.fuzzy.abc") and the destination ("bob@example.xyz"). Then comes the email message (as discussed in section 8.3). It is important to realize that Alice's MTA can supply *any*

9. The DNS is a distributed database that maps IP addresses to computer names, and *vice versa*.

origin, whether or not that origin is accurate. Bob's MTA may, or may not, attempt to verify the accuracy of the origin address (see section 8.5).

Under some conditions, Bob's MTA may determine that the email should be sent to a different address, possibly via a different MTA. This process, called *forwarding*, typically arises when someone moves from one company to another, or when a message is sent to a mailing list that forwards each message to a (possibly large) number of recipients.

8.3. Electronic Mail Headers

Electronic mail headers are described in Section A.2 of Part III of the FTC's report on the Do Not Email Registry [5].

To summarize briefly, an electronic mail email is divided into two parts. The *headers* consist of a sequence of lines beginning with the first line of the message, and ending with the first blank line in the message. Each header is one or more lines[10]. The first line of each header begins with a string of characters followed by a colon. This string is the name of the header. The rest of the line, and any subsequent lines in that header, form the value of that header. For example, the email

```
From: bishop@example.xyz
To: matt@sample.pqr
Date: Tue Sep 15 1959 13:14:15 -0700
Subject: sample

Example email body
```

has four headers. The "From" header has the value "bishop@example.xyz". The "Subject" header has the value "sample". The body consists of one line, "Example email body". The blank line immediately preceding the body divides the headers from the body.

Electronic mail headers are intended to indicate origin (sender), intended recipient, and subject, as well as ancillary processing information. As the Act refers to specific header fields in several places, a brief discussion of the trustworthiness of the header values is necessary.

Electronic mail is simply data sent over a network. When an MTA is passing a email to another (remote) MTA, it engages in a dialogue with the remote MTA. As part of this dialogue, the entire mail message (headers and body) are sent to the remote MTA. Note that there is no difference between the headers and the body as the message is sent to the remote MTA.

8.4. Origin and IP Address

The Act makes numerous references to "originate" and "initiate," and in §5(a)(1)(A) refers to an IP address explicitly. A critical question is how to determine the computer (or network) from which the message "originated" or at which the sending was "initiated." The Act defines "initiate" as "to originate or transmit [a] message"(§3(9)). From a technical perspective, determining the origin poses several problems.

The first is that an IP address can be forged. An IP address in a network communication is simply a set of bits on the network. One can output *any* such set. Typically, the IP address is set

10.A multi-line header has every line except the first line indented.

within a computer, and the computer automatically includes that set of bits as the origin (more commonly called the *source*) address of the message. Thus, any electronic mail message can be traced back to a purported origin, but that origin may not be the actual point of origination.[11]

A second problem is that mail headers can be forged. The sender creates a set of bogus headers, connects to an intermediate MTA, and sends the email to that computer. If the MTA does *not* place IP addresses in its "Received:" lines, the email will appear to have originated elsewhere.

A third problem arises because a computer can be accessed without authorization. Say I am an authorized user of computer A. From that computer I break into computer B, and send out non-compliant commercial email. Is the originating computer A or B? This question is a practical one, as spammers use *zombie programs* (see section 9.5) to send out non-compliant commercial email. The zombie is placed on a compromised system. When the spamming is discovered, it can be traced back to the compromised system. Going to the system from which the compromise was made requires access to logs (of that system or, more likely, the ISP), and may require considerable forensic investigation. The attacker can "fuzz her trail" by going through several computers before compromising the one on which she places the zombies.

A fourth problem comes from an attack called "IP address block hijacking." Several blocks of unassigned IP addresses exist at any given time. A spammer can take advantage of this by simply finding such a block, assigning it to one or more systems, and send non-compliant commercial email from those systems. Here, the IP addresses are "correct" in the sense that they are not forged, but are "incorrect" in the sense that they were not assigned to the systems using them.

Finally, even if the IP address is correct, it can be misleading. Specifically, if the IP address is assigned dynamically, as is done by many Internet service providers, the IP address may be reassigned in the interval following the sending of the message but before the tracing of the route the message tool. For that reason, any investigation into the computer of origin must check the assignment of the IP address *at the time of the sending of the message*.

For these reasons, identifying the actual computer of origin can be problematic. For the purposes of this paper, we assume that all IP addresses are correct (that is, none have been forged or reassigned). We take the computer of origin to be the one indicated in the latter, as follows.

A typical email header includes the following routing information:

```
Received: from x.des.com (x.des.com [10.3.1.4])
     by y.des.com (8.12.11/8.12.11) with ESMTP id j8GK9Jq5012345
     for <xyzzy@z.des.com>; Fri, 29 Jul 2005 13:15:19 -0700 (PDT)
Received: from smtp.ori.com (smtp.ori.com [192.168.33.1])
     by x.des.com (8.13.3/8.13.3) with ESMTP id j8GK8miw021524
     for <xyzzy@des.com>; Fri, 29 Jul 2005 13:14:18 -0700 (PDT)
Received: from mysystem [192.168.32.7] by smtp.ori.com
  (SMTPD-8.21) id A7260204; Fri, 29 Jul 2005 13:12:22 -0700
```

The routing information is in the "Received:" lines. The route of the message is obtained by reading from the bottom "Received:" line up, because as the message transits each intermediate computer, that computer adds a "Received:" line atop the previously added "Received:" line.

---

11. There are complications to this type of forgery. Some routers will reject communications with obviously bogus IP addresses. If the communication is bidirectional, replies will generally not go to the forger (although there are ways to force this that may work, depending on router settings).

So, the last line shows that this message was created on the computer "mysystem" with IP address 192.168.32.7.[12] It was sent to the computer "smtp.ori.com" (which is, presumably, the mail hub for that organization). The next line message was then sent to the computer "x.des.com." This line gives the IP address of the sending computer (here, "x.ori.com", with IP address 192.168.33.1) and the mailbox to which the email was sent ("xyzzy@des.com"). The top line shows that "x.des.com" forwarded the message to another computer in the des.com domain ("y.des.com"). It also gives the IP address of the forwarding computer, here "x.des.com" (IP address 10.3.1.4), and the destination mailbox ("xyzzy@z.des.com").

These are simply lines in the message. If the message is saved on a system (for example, queued for later delivery), anyone with access to the system and sufficient privileges can alter the message headers. So the origin computer can be altered.

But, for our purposes, the origin computer is the one in the last "Received:" line; in this example, it is "mysystem", with IP address 192.168.32.7.

8.5. Checking the Origin of Electronic Mail Messages

Bob's MTA is a server that waits for requests to accept electronic mail. These requests typically come from other MTAs or from UAs. For brevity, we refer to the MTA or UA contacting Bob's MTA as a "client", the computer on which the client is running as the "client's computer", and Bob's MTA as a "mail server".

As noted above, when Bob's mail server is asked to accept a email, the client will supply an originating address. Bob's mail server then can take one of several actions.

1. Accept the address as correct without any further checking.
2. Determine whether the client is on the same computer as the originating address. If not, Bob's mail server can take appropriate action. It could reject the email, but the client may have legitimate reasons for being on a different computer than is named in the sender's address. Usually, the mail server will simply insert the IP address of the client's computer into a header line.
3. Determine whether the client's computer is authorized to send email to Bob's server. If not, the mail is refused, or is accepted and marked as suspicious. There are a number of protocols under development for this purpose.

9. Appendix II: Technologies for Sending Spam

The first recorded instance of commercial electronic spam occurred on USENET, not through electronic mail. Two lawyers posted a note to over 6,000 USENET newsgroups offering to enroll respondents in a green card lottery that would allow the winners to stay in the United States and work. They charged a fee for this service. This was considered spam because it was posted to newsgroups unrelated to immigration in the United States [4].

---

12. All computer, domain, and user names are fictitious; all IP addresses are chosen from the set of addresses that should not appear on the public Internet [11].

9.1. Bulk Email

Early spamming methods involved using automated bulk mailers to generate and send spam. Using one's own system to do this created problems, because recipients could easily trace the messages back to their origin. As service providers (ISPs) usually had clauses banning spamming in their contracts or Authorized User Policies, they could (and did) delete users who sent spam. Spammers began using techniques to protect themselves. The rest of this section describes four common techniques.

9.2. Joe Jobs

A "joe job" is spam that has as its sender some innocent third party, occasionally the mailbox of the recipient himself or herself. The idea is that people reporting the spam will blame the innocent "sender", leaving the spammer unobstructed as more spam is sent.

A "joe job" may damage the reputation of the (purported) sender, and possibly cause threats of legal action against that party. They attempt to turn people's distaste for spam against an innocent third party.

The name "joe job" comes from an incident in which a user at Joe's Cyberpost, a mailbox provider, had his account deleted for sending spam. The user sent another spam, but with the return address set to the address of the webmaster. As a result, the webmaster received many threats and attacks, including denial of service attacks directed against his website [3].

9.3. Relaying Through Third Parties

One common approach to sending spam is to route it through an innocent third party's MTA. An open mail relay accepts email from any source, and forwards it to the given destination (usually a recipient at another MTA).

Recall from section 8.2 that electronic mail may be forwarded from one MTA to another. To relay spam through an innocent domain, say exploited.def, a spammer uses software that acts as another MTA. The software simply relays messages to the open relay, asking it to forward the message to the destination. As the open relay accepts mail from any source, the spams are simply forwarded.

This is usually accompanied by forged sender addresses and (sometimes) modified "Received:" header lines in order to hide the origins of the spam.

9.4. Break-Ins To Send Spam

Rather than try to hide the origin of the spam, some spammers have resorted to compromising systems and sending spam from those systems. The spammers, or people acting on their behalf, break into an innocent user's system, and use a bulk mailer to send spam from that system. This has the advantage of not requiring any forging of mail headers; the mail appears to come from the user whose account was compromised, and from the system that was compromised. This is often used in conjunction with the next technique.

9.5. Zombies

A "zombie" is a program that acts on behalf of another user without manual intervention. The term has the connotation of illegitimacy. In the context of spam, a zombie is a program placed on a compromised system in a variety of ways. Zombies typically are placed on hundreds or thou-

sands of systems. Then each zombie sends out some spam. As the spam will come from a very large number of systems, it effectively defeats the use of blacklists.

Zombies are especially pernicious because they are often downloaded by unsuspecting users when the users visit a web site, or when the user runs an attachment to email. Thus the system on which the zombie is run need not be broken into.

10. Appendix III: Technologies for Controlling Spam

This section discusses several technologies that have been proposed to control spam, or to assist users and service providers (ISPs and mail service providers) in controlling spam. It also provides context to assist the reader in evaluating the effectiveness of the technology.

10.1. Authentication Services

An *authentication server* validates the identity of a sender. This may happen at any of three levels.

A *user authentication server* validates the identity of the sender of a message. When a user sends a message, he or she digitally signs the message. When a receiving program (MTA or other program) receives the letter, it determines the appropriate user authentication server to use based upon the sender's address. It then queries that server, and using information from that server determines whether the user digitally signed the message. Examples of this technology are PGP [6] and PEM [1,7,9].

User authentication servers tie the origin of the letter to a user. The problem here is that users may be transient; for example, Charlie could obtain a mailbox from the company mailboxes.xyz, which offers free email. Charlie then creates an entry in a user authentication server, tying back to charlie@mailboxes.xyz. Charlie then sends out messages with false and misleading header information. The user authentication server will validate the message (after all, Charlie did send it). By the time the company, mailboxes.xyz, is notified of Charlie's action, Charlie is long gone.

A *host authentication server* validates that a message came from the sending computer. When a message is sent from a computer, the computer's MTA digitally signs the letter. When a receiving program receives the letter, it determines the appropriate computer authentication server to use based upon the sender's address. It then queries that server, and using information from that server determines whether the computer digitally signed the message.

Host authentication servers will enable the MTA to verify that the letter came from the purported computer. A "throwaway" computer is a computer that is registered for the purpose of sending email, and once the mail is sent, it can be removed from the network. If such a computer registers with an authentication server, the server will enable to recipient to validate that the letter did indeed come from that computer—and if the letter is spam, the computer (and spammer) are gone by the time the authorities are notified.[13]

A *domain authentication server* validates that a message came from the sending domain. When a message is sent from a computer in a domain, the message is digitally signed. When a receiving program receives the letter, it determines the appropriate domain authentication server

13.In fact, the computer may remain. But the spammer cannot be traced from it.

to use based upon the sender's address. It then queries that server, and using information from that server determines whether the message was digitally signed by the domain[14]. Domain Keys [14] is an example of this technology.

Domain authentication servers simply verify that the message originated from within a domain, and suffer from problems similar to computer authentication servers such as "throwaway domains."

"Throwaway" domains serve the same function as throwaway computers. Here, registration requires a domain registrar. An obvious suggestion to eliminate throwaway domains is to require that registrars tie domain ownership to a specific, established individual or company. The problem is that such a requirement eliminates anonymity, which under some circumstances is not merely desirable but necessary, and further that specific criteria must be established to assure that the claimed identity is in fact the right one. This requires building a superstructure of validation and verification similar to that used by certificate authorities. Such a structure would require widespread agreement on what is acceptable for validation and verification, and for that reason alone is impractical.

10.2. Authorization Services

An *authorization server* identifies those computers authorized to send electronic mail from a given domain. For example, if an MTA receives an email from post.mailisus.stu, the MTA asks the authorization server for domain mailisus.stu if the computer called post.mailisus.stu is authorized to send mail from that domain. If so, the email is accepted. If not, the email is either rejected, or it is labeled as suspect. The Sender Policy Framework (SPF) [15] is an example of a protocol for authorization servers.

The problem is that if the message is sent from an authorized computer, it could still be spam. This is especially true if the computer allows messages to be relayed through it, or if a spammer sets up the domain. In the latter case, the authorization server will simply confirm that the system used by the spammer is authorized to send email, defeating the purpose of the authorization server.

10.3. Reputation Services

Reputation services come in several forms. The best-known are whitelists and blacklists.

A *blacklist*, sometimes called an RBL (Realtime Black List) is a list of addresses or domains that have a poor reputation, such as for sending spam. Electronic mail from these addresses or domains is discarded or marked as potential spam. An example of a blacklist is the Spamhaus Block List (SBL)[15]. The effectiveness of blocking spam based on these lists depends upon the quality of the lists. Some lists are aggressive, in that it is easy to be added to the lists by accident; others are relatively conservative. All have mechanisms to remove entries added in error, although the methods differ wildly.

A *whitelist* takes the opposite approach. Electronic mail from these addresses or domains is accepted; all other electronic mail is discarded or marked as potential spam. An example of a whitelist is Spamhaus' ".mail" top level domain[16]. Spamhaus plans to vet sites in this domain to

---

14.Actually, by a computer in the domain.

15.http://www.spamhaus.org/sbl/index.lasso

eliminate any registrants who might send spam, the goal being that email from any computer in that domain may be accepted without fear that the mail is spam.

A third type of reputation server tracks a sender's reputation. When a message is received (either by an MTA or a UA), the sender is identified and the appropriate reputation server queried. If the sender's reputation is good enough (the precise level of acceptability is determined by the policy of the receiving system, or the recipient), the message is accepted. This approach is similar to that used by eBay to rate the reliability of its users.

Using a reputation server has two problems. The first is that the reputation is tied to the (purported) identity of the user and not the user himself or herself. Thus, a spammer need only establish a good reputation for one identity and then send spam. That identity now has a bad reputation (in the best case). The spammer now establishes a second identity, and proceeds to build a new reputation. This process can be repeated indefinitely.

It is tempting to require the user of a reputation server to establish reputation through some method that would hurt the user were the reputation to be betrayed (for example, by sending spam). Spamhaus' ".mail" domain approaches this problem by charging a considerable sum, and imposing time delays, before one can use the domain. The economics of this approach seem to be effective, based on logic, but have not yet been tried in practice. But this approach assumes the recipient of messages will use the reputation server; therein lies the second problem.

Many users work with the public or are often approached by unknown parties. For example, the author is a professor who often receives email from prospective students and other people who wish information. If the author only accepted email from people with established reputations, letters from most of these correspondents would be discarded. But the nature of teaching and research in a university environment preclude this. Hence a reputation server would be of questionable usefulness to the author in screening correspondents, especially unknown ones.

The proposed "Do Not Email" registry is an example of a blacklist because the addresses on it are those of people not to be sent spam. Any authorized user of at registry can check for valid addresses. This is necessary in order to eliminate addresses that are on a mailing list. But this also means that any addresses that are to be kept secret cannot be on such a list. Hence such a list could be used to garner valid email addresses that could then be targets of spam.[17]

10.4. Challenge-Response

Challenge-response refers to a class of authentication techniques in which a user asserts an identity, is presented with a question (the challenge), and must provide the correct answer (the response) to validate the claimed identity. In the context of handling spam, challenge-response techniques require that the sender of electronic mail affirm that the email is not spam.

When an MTA receives a letter, it places the letter into a temporary queue. It (or some other program acting on its behalf) contacts the sender, requesting that the sender reply to a challenge of some sort. The challenge may be as simple as responding to the request. It may be more complicated, requiring the sender to solve a simple puzzle. When the response is received (and

---

16.See http://www.spamhaus.org/faq/answers.lasso?section=The%20.mail%20TLD.

17.The FTC report on the "Do Not Email" registry [5] goes into more detail. See for example §IV.B and especially §IV.B.2.b, which explains why one-way hashing of addresses on such a list does not protect the privacy of those email addresses adequately.

possibly validated), the email is delivered. The sender may also be placed on a whitelist, to avoid future challenges. If no response is received, the letter is marked suspect or discarded.

Challenge-response techniques have two problems. The first is that the sender may not respond in a timely manner, causing the message to be delayed or discarded erroneously. The second is that challenge-response techniques in this context are designed to inhibit mass mailings from reaching the recipient. But some mass mailings are legitimate and not spam, such as mailing lists to which the recipient has subscribed. For this reason, challenge-response techniques employ a whitelist to which a user may add senders. Emails from senders on the whitelist are accepted without challenges.

10.5. Port Blocking and Rate Limiting

Sending spam requires that the computer from which the spam is sent be able to send large quantities of electronic mail, usually over a (relatively) short period of time. This suggests two approaches to inhibiting spam. Both should be implemented by the service provider (ISP), not by the local computer.

The first is to disallow some computers from sending electronic mail. This is referred to as "port blocking" or (more specifically) "blocking port 25 traffic"[18]. When this is done, the ISP will refuse to accept any email from computers so blocked. This is often combined with blacklists or whitelists, so email from blacklisted computers or domains is simply not accepted. Another common approach is for service users to designate a system as their *mail host*, or computer at which they will receive email. If no user designates a system as a mail host, no mail is accepted from, or sent to, that computer.

ISPs in particular often designate some set of systems as *mail servers*, and allow outgoing mail only if the mail originates at one of those servers. Thus, zombies implanted on a home user's system will be blocked, as those systems are not designated as mail servers.

The problem with port blocking is that a spammer may send mail from an authorized mail host, and that host cannot be blocked because it sends legitimate email. A service provider may implement a related technique, called "rate limiting," "port throttling" or "throttling port 25 traffic". This technique allows mail to go through, but limits the rate at which the computer can send mail. For example, the ISP may allow the computer to send no more than 100 letters a day. Considering that the best estimates of a typical spam are in the millions of emails, the effect of throttling port 25 traffic can materially inhibit the sending of spam.

10.6. Preventing Relaying

Most MTAs can be configured to accept email only from known, trusted computers. This requires spammers to take more complicated steps to forge the identity of a trusted computer, and is typically far too time-consuming; spammers simply find another MTA that is an open relay. Lists of open relays are available on the Internet, and such computers are often simply blacklisted (see section 10.3). This blocks spam.[19]

---

18. A *port* is a construct used in networking to indicate a type of network message. On the Internet, network traffic sent to port 25 is electronic mail.
19. Such lists also provide a set of MTAs that spammers may use.

10.7. Spam Filtering

By far, the most pervasive technique for blocking spam is to analyze the contents of a message to determine if it is spam. There are two variations of this, and they can be combined.

The first is statistical filtering. This technique builds a statistical model of messages manually identified as spam. As the number of spam messages grows, the accuracy of the model improves. As electronic mail arrives, the filter scans each message, and if the statistics of the message are "close enough" to those of the model, the message is labeled as spam. The model can be static (in which case it is not updated) or dynamic (in which case the statistics of the newly identified spam message are incorporated into the model). Types of filters include Bayesian filters, neural nets, and genetic algorithms. This technology is common, SpamAssassin[20] being one of the most widely used filtering tools.

Wittel and Wu [12] discuss ways that spammers can evade spam filters. These involve spammers obtaining copies of the filters and statistical models of spam messages (which are often distributed with the filters). Then the spammers attempt to craft messages that the spam filter will not flag as spam.

The second type of filtering method looks for specific keywords and other non-statistical indicators of spam in the message. One common indicator is a hyperlink in which the displayed URL does not match the hyperlinked URL, for example:

```
<A HREF="http://192.168.4.5/gotcha">http://www.bigbank.corp/security</A>
```

Here, if the user clicks on the seemingly-legitimate URL "http://www.bigbank.corp/security, she will be taken to the rogue URL "http://192.168.4.5/gotcha". This is an obvious attempt to hide the true web site from the recipient.

Another common indicator is that a URL is associated with spam. These filters scan the body of a message, and if a URL that is on a blacklist appears, the message is immediately labeled as spam. One example of such a list is the Spam URI Realtime Block List (SURBL) [2].

10.8. Image Blocking

Image blocking is an attempt to prevent consumers from seeing sexually explicit material without their consent. It works when the body of an email contains the address of an image. The image itself is located on a web server. When the user reads the email, the user agent contacts the web server, downloads the image, and displays it in the appropriate place in the body of the message. One problem with doing so is that the image may be sexually explicit, so the user will see it whether or not he or she wants to. The most widely used email reading programs now contain an option to allow the user to block this automatic downloading of images, and by default it is turned on.

10.9. Keywording

One problem with anti-spam technology is that legitimate mail may occasionally be blocked as spam. To enable a legitimate recipient to get around the spam block, some MTAs can be configured to accept as legitimate any email that contains a key word or password embedded in the body, or in the subject header. Then the MTA checks the letter for the keyword before it determines if the message contains spam. If the keyword is present, the spam check is not performed

20.See http://spamassassin.apache.org.

and the message is accepted. This is a form of whitelisting, but involves a keyword in the message rather than an originating address.

A key problem is how to inform legitimate recipients about the keyword. One technique is to post the keyword prominently on the recipient's web page; most spammers will not check there before sending spam. It is strongly recommended that any spam *not* be returned to the sender, because then a "joe job" could cause an innocent victim to receive the bounced message.

11. Appendix IV: Measuring Effectiveness of the Act

The question of measuring the effectiveness of the Act is critical to determining whether the Act is working. Unfortunately, any such measurements are very difficult, if not impossible, to determine precisely. This section explains why, and describes how to develop an approximate measure. The term "approximate" is emphasized.

The first problem is the definition of the term "effective." It could mean that, as a result of the Act, companies comply with its requirements. It could mean that the Act requires companies to insert information in their messages to enable recipients to discard any letters they consider spam. It could mean that the Act has caused anti-spam technology to advance, and become more effective at blocking spam. It could also mean that the amount of spam that violates the Act is reduced. We consider these definitions separately.

The first definition (that companies comply with the requirements of the Act) is discussed elsewhere. For this report, we simply note that the Act's goal is to make companies comply with the requirements of the Act. The Act provides enforcement mechanisms that are procedural, not technical, and so their ability to counter these offending spammers is outside the purview of this report.

The second definition (that companies who wish to comply with the Act can do so, and thereby mark their messages so recipients can detect spam and discard it) requires checking what the Act requires against the anti-spam measures discussed in section 10. It is discussed in detail in section 5 of this report. Briefly, in this sense, the Act is effective.

The third definition (that as a result of the Act, anti-spam technology has advanced) cannot be answered by considering only the technology, because it is *not* clear if the advances in technology were driven by the Act. But anti-spam technology has advanced in the past 2 years (the period of time since the Act was passed). Image blocking (see section 10.8) is perhaps the most obvious case, and was used before the Act was passed. However, back then, most user agents would download the images by default. But it is not possible to say whether the change in defaults (from downloading to asking the user whether to download) was due to the Act. Other technologies that have improved are the closing of open relays (see section 10.6), port blocking and throttling (see section 10.5), and spam filtering (see section 10.7).

The fourth definition (that the amount of spam that violates the Act is reduced) is the most difficult definition to analyze. The reduction, presumably, must be significant; if the Act caused 1 less offending spam email to be sent, it is unlikely that the Act would be considered effective. So, let us call the Act "effective" in this sense if it reduces offending emails by more than 80% (or some other percentage).

The problem with determining if the Act is effective in this sense is that one cannot do a controlled experiment to test that hypothesis. Consider the problem of collecting emails first. As

we cannot possibly obtain copies of all emails, we must choose a sample that represents all emails, and analyze it. Selection of such a sample is a difficult task because one must demonstrate that what you are measuring in the sample emails does indeed represent the same measurement of *all* emails. We are not aware of any such scientific analysis for any existing sample of email. One can use any of several collections of email available on the Internet, but there is no assurance that those collections provide an accurate model of the actual emails on the Internet.

This is true for other collections of emails. Honeypots, or addresses that do not send emails but only receive them, and the FTC's large sample of emails sent to it[21], represent biased samples. The honeypots are biased because they do not reflect the way users interact with the email system, and the FTC's sample is composed of emails sent to it when users report spam. Neither represents email traffic in general.

Thus, any answer we get from the analysis of such a sample may not generalize to all emails. Unfortunately, as we cannot capture all electronic mail sent over the Internet, use of these samples is the best that can be done. No other technique appears to work any better.

Even if one could obtain a sample that demonstrably represented all Internet email, other variables would affect the measurements of effectiveness of the Act. For example, as noted above, anti-spam technologies have evolved during the lifetime of the Act. If the amount of offending spam were shown to have decreased within the lifetime of the Act, one could not determine whether the Act was responsible, or better anti-spam technology was responsible, for this decrease.

An alternate interpretation of this notion of "effectiveness" is to determine what percentage of email messages comply with the Act. Again, the issues of sampling discussed above arise, the best one can do is determine the percentage of emails in a given collection of emails that comply with the Act's provisions. Again, this does not answer the question of the Act's effectiveness in a scientific sense, but as argued above, it is the best one can do.

21.This is called the "FTC refrigerator" in other documents.