SPAM, SPAM, SPAM, LOVELY SPAM

Brian Hayes

A reprint from

American Scientist

the magazine of Sigma Xi, the Scientific Research Society

Volume 91, Number 3
May–June, 2003
pages 200–204

SPAM, SPAM, SPAM, LOVELY SPAM

Brian Hayes

I used to feel forlorn whenever I checked my e-mail and found nothing waiting for me. Not much risk of that these days. There's always someone who wants to help me lose 32 pounds, or clean up my tarnished credit report, or get me that college degree I skipped in my reckless youth. Absolute strangers send me tips on the stock market and ideas for starting a new business that I could run with a laptop computer from my new condo in Hawaii. Then there are the ill-considered schemes to share the ill-gotten gains of African dictators, as well as a voluminous and varied stream of messages that are best described as indecent proposals.

When the very first of these missives began appearing in my mailbox, some years ago, they were mildly intriguing, like messages in bottles washed up on the beach. I had to wonder—in the millisecond before I hit the delete button—who had sent them and from where and why. Most of all I wondered for whom they were meant, since they were clearly of no use or interest to me. By now, of course, the sense of mystery is long gone. The occasional message in a bottle has become a daily tide of wrack and flotsam; it's as if whole cargoes are being dumped overboard to litter our shores; an unstoppable oil slick of oleaginous marketing sludge slops into every e-mail inbox around the world, wave after wave and day after day.

The very efficiency and convenience of electronic communication gets some of the blame for this flood of unwanted and thoroughly unloved junk e-mail. By dramatically reducing costs, the Internet makes it economically feasible to blanket the globe with boring sales-pitch messages, even if only the tiniest percentage of the recipients respond. But if technology created this problem, maybe it can also contribute to the solution. Or are social and legal remedies more promising?

Green Cards and Spam

It is worth remembering, in this era of Web pages festooned with blinking and bleeping banner ads, and accompanied by pop-ups and pop-unders, that once upon a time the Internet was an advertising-free zone. As long as the U.S. government controlled a major part of the backbone, most forms of commercial activity were forbidden. The occasional violations of this rule attracted swift and severe retribution. For example, in 1978 Digital Equipment Corporation (since absorbed into Compaq) sent a notice about a new computer system to 600 subscribers on the ARPANET, one of the ancestors of the Internet. The message was immediately labeled a "flagrant violation" of government policy, with the assurance that "appropriate action is being taken to preclude its occurrence again."

The rules changed in 1993, as the Net was privatized, but social strictures on indiscriminate advertising remained powerful for some years. In April of 1994 a message with the subject heading "Green Card Lottery- Final One?" was posted simultaneously to 6,000 Usenet news groups. The advertisement, signed by the Phoenix law firm of Canter & Siegel, offered information and legal services to immigrants. Thousands of Usenet regulars—incensed not only by the commercial nature of the message but also by the waste of bandwidth and the breach of "netiquette"—hounded Canter & Siegel by e-mail and fax and telephone. The lawyers' Internet access was cut off, and eventually the firm went out of business; Canter was disbarred. There have not been many such victories in the fight against spam.

As it happens, the Canter & Siegel incident was the event that first popularized the term "spam." According to Brad Templeton, a Usenet pioneer and current chair of the Electronic Frontier Foundation, certain small online communities had used the word earlier to describe various kinds of unwelcome verbiage, but it was the Green Card affair that made it widely known. The ultimate source was a 1970 skit on the British television show *Monty Python's Flying Circus*, about a restaurant with a limited menu and a chorus of Vikings chanting "Spam, spam, spam, spam, lovely spam, lovely spam." (Incidentally, Hormel Foods, who make SPAM rather than spam, attempted a defense of their brand name, then decided to have a sense of humor about it.)

Although spam today is mainly a plague of e-mail, the Canter & Siegel ad and several other

Brian Hayes is Senior Writer for American Scientist. *Address: 211 Dacian Avenue, Durham, NC 27701; bhayes@amsci.org*

early examples were never sent as mail; instead they were posted to news groups. The Usenet news service is especially vulnerable to spamming because the complete list of groups is freely available to everyone, unlike e-mail addresses, for which there is no central directory. Furthermore, because a single news group can have many readers, and a single reader may look at many groups, posting to a few thousand groups annoys millions. E-mail, in contrast, is generally one-on-one, and it takes more effort to cause the same amount of consternation.

In view of the vulnerability of the news system, it's encouraging to report that the Usenet community was able to organize itself to cope with the problem, if not solve it. Both manual and automated controls have been put in place, allowing a message that can be identified as spam to be removed or at least flagged as suspect before it reaches the reader. (One of the anti-spam protocols is called NoCeM, pronounced "no see 'em.") Of course there is the potential for abuse by individuals who maliciously cancel legitimate postings, but safeguards are available, and the system seems generally to be working. Some spam still gets through to news groups, but the noise level peaked several years ago, and in most groups it is now quite tolerable. On the other hand, part of the reason that spammers leave Usenet in peace these days may be that they just don't care: Only a small fraction of Internet users ever look at news groups.

The Spam on My Plate

In the case of e-mail, the magnitude of the spam crisis is hard to pin down, but there is certainly a widespread perception that the amount has mushroomed in the past year or two. Lately the problem has been getting frequent attention in both the popular and the technical press, and there have been several recent conferences and workshops to explore remedies. The Internet Engineering Task Force has just formed an Anti-Spam Research Group. At the governmental level, the Federal Trade Commission has taken an interest (although not, as yet, much action). About half the states have enacted laws on spam; federal legislation is the subject of intense lobbying efforts on both sides. Elsewhere, the European parliament has adopted a stern policy, but their success in enforcing it is not yet known.

The statistics on e-mail spam that appear in news reports come mainly from companies that sell products and services for combatting spam. One of these vendors, Brightmail, Inc., says that spam made up 8 percent of all e-mail traffic in 2001 but had grown to 36 percent by the middle of 2002 and was over 40 percent by the end of last year. Postini, Inc., another company providing anti-spam services, reports that the proportion of spam in the mail they monitor climbed from 20 percent in January 2002 to 60 percent by December. I have no reason to doubt the accuracy of these figures, but it is only fair to point out that the companies' own interest lies in emphasizing the severity of the crisis.

A few consulting firms and foundations have also surveyed the volume of spam. Jupiter Research estimates that the average e-mail user gets about 2,200 spams a year, and the Gartner Group says that corporate e-mail is 25 to 35 percent spam. But a discordant note comes from the Pew Internet & American Life Project, which surveyed 2,500 Internet users, asking only about e-mail they receive at work. Half said they get no spam at all in their workplace accounts, and 71 percent reported no more than "a little."

Out of curiosity, I have been keeping track of spam in my own in-box for the past few months. From November 2002 through March 2003 I received 1,571 items that I would unequivocally classify as spam; another 287 are doubtful cases. The total number of messages received in the period was 6,028. Even giving the benefit of the doubt to all the doubtful ones, this tally suggests that I'm getting more than my fair share of spam when measured in absolute numbers; at this rate, I can expect almost 3,800 spams a year rather than the 2,200 predicted by Jupiter Research. Yet the proportion of spam in my mail is only 26 percent, less than the averages reported by Brightmail and Postini, and near the low end of the Gartner estimate. (Obviously, percentage measurements are sensitive to the amount of both spam and nonspam mail.)

Spammers are said to harvest most of their addresses on the World Wide Web. My address has been posted on the *American Scientist* Web site for eight years, which may have something to do with my popularity among the spammers. But there's more to the story. Another of my e-mail accounts has never been published on the Web or anywhere else, yet it attracted more than 70 spams.

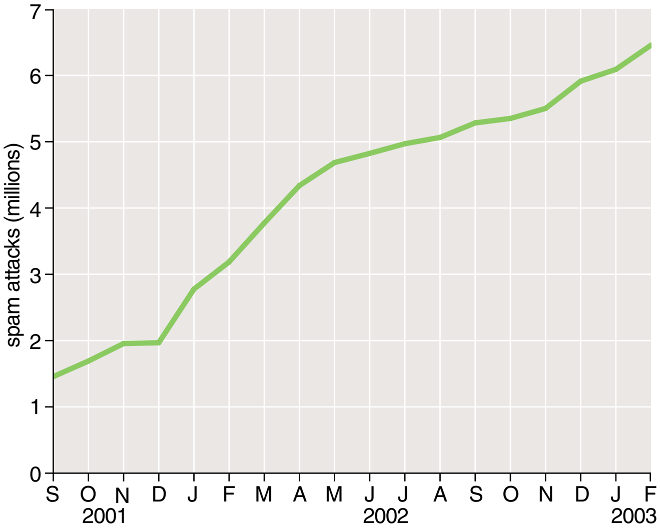

Figure 1. Number of spam attacks has quadrupled in the past year and a half. Each attack sends out up to several million pieces of unwanted e-mail. The data were collected by Brightmail, Inc., using a network of "honeypot" mail accounts set up deliberately to attract spam.

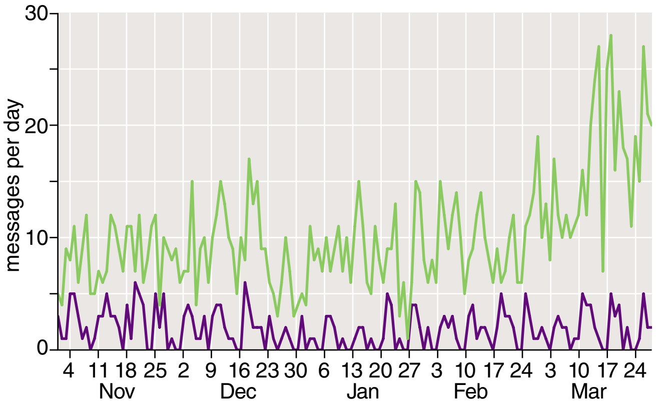

Figure 2. Daily tally of spam received in my own mailbox shows a strong uptick in the most recent weeks. The green curve records anonymous spam—unwanted mail from unknown senders. The purple curve traces messages that many would also classify as spam but that fall in a different category because they come from sources known to me. In the five months studied, there was not a day without spam.

Classifying my mail as spam or nonspam took more thought than I expected. Of course there are many echt spams that I could recognize in an instant, without glancing beyond the subject line: "Eat pizza, watch TV ... and lose 22 pounds," "Absolutely FREE - Receive $610.00!!!" I have a hard time believing that *anyone* would actually want to receive some of these messages. The worst of them can only be understood as "trolls," deliberately meant to be annoying, and thereby to provoke the recipient into responding—thus verifying that the e-mail address is valid.

Alongside these blackest of spams, however, there are also shades of gray—mail that I don't want and that I didn't ask for (or at least I don't remember asking for it) but that might conceivably interest someone else and that comes from a source I know. For example, when a certain scientific society of which I'm a member (*not* Sigma Xi!) sends me seven invitations to register for its annual meeting, is that spam? What about a catalogue merchant from whom I've made purchases in the past and who, unbidden, sends me weekly sales flyers? Or an employment recruiter looking for candidates to fill a job? Would it make a difference if the job might interest me? Then there's my cousin Larry, who sends everyone in the family a continuing stream of chain letters, urban legends and unfunny jokes. Is Larry a spammer?

I find it worthwhile to distinguish between mail that comes from a known source (my bank, my travel agency, my cousin) and mail whose sender is unknown and perhaps unknowable. Some of the mail from identified sources may well be unwanted and indeed may qualify as spam, but at least there are direct ways of dealing with the issue. I can call my bank, or have a word with Cousin Larry. The anonymous spam is a harder problem. I have no reliable means of throttling back the glut of e-mail I get from YourShoppingRewards and Freddys-Fabulous-Finds.

Spam is often defined as "unsolicited commercial e-mail," and indeed almost all of it that I'm receiving currently appears to have some commercial purpose. Whoever is sending this junk is trying to make money out of it (although it's not always obvious how). Still, lots of noncommercial bulk mail would be equally unwelcome. An early Usenet spam—even before Canter & Siegel—had the subject header "Global Alert For All: Jesus is Coming Soon"; the aim was to save souls, not earn bucks, but readers were just as unhappy with the author. Spam wars can also break out between political factions or over issues such as abortion or the death penalty. (If I were a spammer, I would make 10 percent of my mailings political, just to bolster the freedom-of-speech argument against regulation of bulk e-mail. I hope I haven't just given someone an idea.) In any case, plans for controlling spam should probably not be tied too closely to its commercial nature.

Canning Spam

In the early years of the Internet, dealing with miscreants was remarkably straightforward. Although Net lore and legend celebrate the lack of any central governing body, the network was actually run by a cohesive community with shared aims and values. Any serious breach of the rules could be punished by expulsion—by canceling the violator's account. The situation is different now. If Hotmail kicks you out, you just move over to AOL or MSN. Furthermore, Net pariahs with enough resources can set up their own Internet service provider.

Yet the power to isolate renegade sites has not been lost entirely. You cannot easily reach out and unplug an offending node of the network, but you can set up your own system to ignore any information coming from that node. In particular, a system administrator can configure an Internet router so that it refuses traffic from selected sites. Some years ago Paul Vixie, who conceived several early Internet protocols, began publishing a list of network nodes from which spam was emanating. This Realtime Blackhole List, or RBL, is now maintained by a nonprofit organization called MAPS. For subscribers to the RBL, the listed sites become black holes—no e-mail can get out, and in some cases traffic of all kinds is blocked. The weapon is quite blunt, in that it stops not only the spam but also other innocent communication. The rationale for this policy is that legitimate users of a blacklisted service will exert pressure to shut down the spammer so that they can again reach the outside world. It's rather like keeping the whole class after school until someone turns in the naughty child. Not everyone approves of this strategy, and there have been several lawsuits against Vixie and MAPS. Meanwhile, spammers cope by keeping on the move and by disguising their whereabouts.

Other approaches to filtering out spam try to block only the offending mail. The filter can be installed on the individual user's computer, on a mail server or farther upstream. Much ingenuity has been brought to bear on designing filters. Also on evading them.

The simplest kind of filtering uses static criteria to sort incoming mail into various folders or directories. For example, a filter might reject any mail that comes from "Bargain Blizzard" or that has "inkjet" in the subject line. But the spammer's response is all too easy: The sender becomes "Blizzard of Bargains" and the subject becomes "i-n-k-j-e-t" or one of a thousand other variations. There is also the problem that a friend writing for advice on an inkjet printer may have a hard time getting through.

Large-scale services such as Brightmail and Postini cannot hope to keep up with the evolving spam ecosystem by hand-crafting filter rules. The key to their methodology is to set up thousands of "honeypots"—e-mail accounts whose only purpose is to attract spam. Since these addresses should have no legitimate e-mail sent to them, messages collected there can serve as templates for filtering the stream of mail going to the service's subscribers. In this way one of the essential, defining characteristics of spam—the fact that it goes simultaneously to thousands of addresses—is turned into a weapon against it.

Another mechanism for building filters is based on the collaborative effort of thousands of people performing the routine daily chore of sorting their e-mail. If you are participating in such a cooperative network, then every time you mark a spam message for deletion, a copy of the e-mail is sent to a central repository; there many such reports are gathered and compiled into filter criteria. When the same message arrives again, addressed either to you or to another participant in the cooperative, the mail is automatically shunted to the spam bin. This idea originated with Vipul Ved Prakash, a San Francisco programmer, who created a public-domain program called Vipul's Razor. A commercial version called SpamNet has been in testing for the past year. The SpamNet cooperative has more than 300,000 members.

Schemes that filter out only identical copies of a known exemplar message have a serious weakness: The spammer can overcome them by making each copy of a mailing slightly different, perhaps by adding a few random characters to the text. (Presumably, this explains subject headers such as "Married, Lonely, and home alone ! 2563SEpT0-115eltW64-18.") Brightmail reports that 90 percent of all spam messages are now unique, and so more-elaborate algorithms are needed to establish a match between a template and a target. The SpamNet technique is vulnerable to the same countermeasure, and so again the filter cannot rely on a simple, exact match. Given many copies to examine, however, an algorithm can determine which regions of the message are constant and which are variable, and thereafter focus only on the stable, identifying features. But already a counter-countermeasure has appeared. It is "scramblespam," where random characters are not a minor addition but make up most of the message, typically with the actual content of the ad embedded in an image. Cloudmark, the company behind SpamNet, reports it has recently devised an algorithm for identifying scramblespam.

Yet another filtering strategy abandons the whole idea of matching messages to templates and simply looks at the statistical properties distinguishing desirable e-mails from spams. This idea began to gain momentum last summer when Paul Graham, a computer scientist best known for his books on the Lisp programming language, circulated an article titled "A Plan for Spam." It soon emerged that similar principles were already familiar and well-developed in other fields, such as automated text analysis and computational learning theory. By the time of a conference on spam held at MIT in January, several variations of the algorithm were being actively explored.

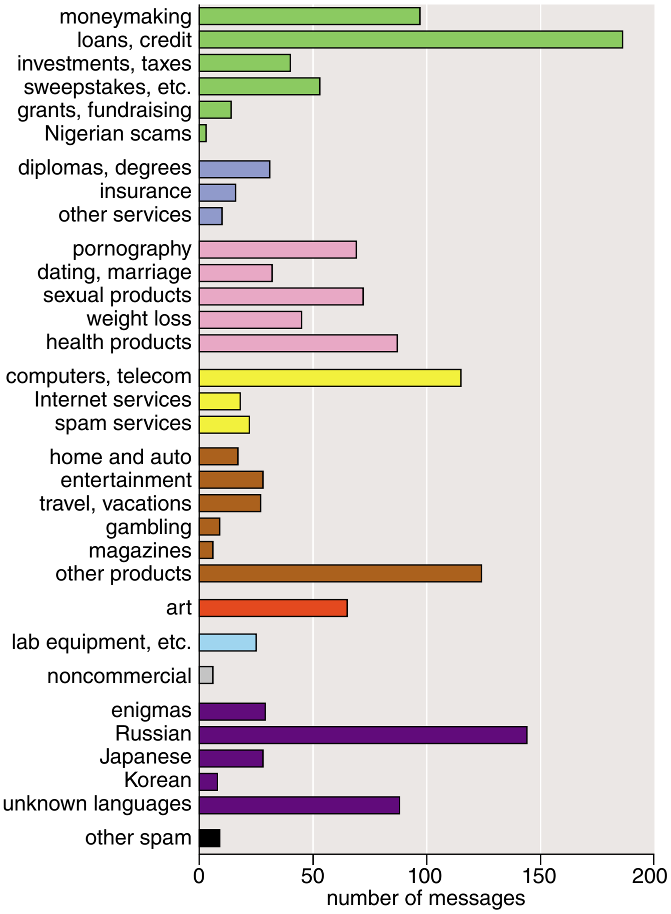

Figure 3. Money and sex are leading topics in the spam that lands in my account, as they are in broader surveys; but my spam sample also has a few surprises. Some 65 messages deal with the art market; all of them may come from a single mailer. Two dozen advertisements for laboratory equipment and other items of possible interest to people in the sciences suggest that some spam may be targeted to a particular audience rather than broadcast to the entire world. On the other hand, more than 250 messages are clearly *not* meant for me, since they are written in alphabets my computer cannot display. Another 29 messages, labeled "enigmas," are written in English, but I haven't a clue what they're about. Six political and religious tracts are classified as noncommercial.

The statistical filtering process requires a reasonably large corpus of e-mail messages, already divided into spam and nonspam categories so that they can serve as a training set. The program breaks messages down into individual words and other "tokens," recording the number of appearances of each token in the two groups of messages. The resulting frequency tables give the probability that any spam or nonspam message contains a specified token. Furthermore, from the same information it is also possible to calculate the inverse probability: Given any token, the tables determine the probability that a message containing the token is either spam or nonspam. New messages are classified by finding the most "interesting" tokens—those whose probabilities are closest either to 0 or to 1—and then computing an overall composite probability that the e-mail is spam. In Graham's experiments, the probability distribution turned out to be strongly bimodal: Most messages were either close to 0 or close to 1, with few in the middle. His reported error rate is about five per 1,000 for false negatives (spams that squeak by as legitimate mail), with no false positives (legitimate mail misidentified as spam).

Graham argues that a filter based on the entire content of a message cannot be evaded without altering the content itself. "It would not be enough for spammers to make their emails unique or to stop using individual naughty words," he writes. "They'd have to make their mails indistinguishable from your ordinary mail.… Spam is mostly sales pitches, so unless your regular mail is all sales pitches, spams will inevitably have a different character." The argument seems compelling and the results so far are impressive, but the real test will come when such filters are widely deployed, putting pressure on spam authors to invent countermeasures.

The Price of Spam

Purely technical remedies are not the only option for controlling spam. There are also legal and economic maneuvers, as well as a few tricks of social engineering.

At last count 26 states had enacted some form of law imposing sanctions on bulk e-mail, but many spam opponents are lukewarm about these remedies, worrying that regulating the spam business will tend to legitimize it. Most of the laws take an "opt-out" approach, meaning it's the recipient's responsibility to get off the mailing list rather than the sender's responsibility to obtain permission first. The guidelines adopted by the European Union require an "opt-in" mechanism. But whatever the details of the law, enforcement is always problematic because e-mail crosses jurisdictional boundaries so readily.

Another legal tactic is to forbid concealing or disguising the origin of e-mail, on the principle that spam can usually be stopped if it can be traced back to its true source. Proposed changes in the Internet protocols for transporting e-mail would have the same effect, making it more difficult to send messages with a forged identity.

Economic remedies would shift the cost of spam from the recipients back onto the senders. It now costs only $99 (according to some spam I received) to spew out a million e-mails. At that price it could well be profitable to irk and inconvenience 999,900 people in order to swindle the most gullible 100. A small tax or fee imposed on each message might restore the balance. Paying a penny per message, the ordinary user would hardly notice the charge, but a spammer sending 100 million e-mails would face a bill of $1 million. The catch again is enforcement and jurisdiction.

Scott E. Fahlman has proposed a variation in which the fee would be paid directly to the recipient. Under this plan, e-mail software would accept a message from unknown correspondents only if the sender agreed to pay for "interrupt rights." The charge could be waived retroactively at the recipient's option, so that nonspammers would never actually have to pay.

Most of the social and economic anti-spam schemes would work only if a large majority of e-mail users were to adopt them. Assembling that majority is the challenge. Indeed, if we could achieve universal agreement on the issue of spam, there would be no need for strategy at all. The very simplest approach would suffice: We could just say no. In the end, all that's needed to defeat spam is for everyone to ignore it. But that doesn't seem to be happening so far. Looking at the 1,571 tantalizing offers in my mailbox, I don't feel the slightest itch to buy anything, but *someone* must be taking the bait.

Even if all the filters, blackhole lists and other measures fail to abolish spam, efforts to control it are still worth making. They can reduce the volume. They might even bring us *better* spam: In order to get through to reluctant or jaded readers, advertisers will have make their spam more appetizing. Live with it long enough, and you might develop a taste for the stuff. Spam, spam, spam, spam, spam, spam, lovely spam, wonderful spam.

Bibliography

Fahlman, Scott E. 2002. Selling interrupt rights: A way to control unwanted e-mail and telephone calls. *IBM Systems Journal* 41:759–766.

Garcia, Dan. Dan Garcia's spam homepage. http://www.cs.berkeley.edu/~ddgarcia/spam.html

Gleick, James. 2003. Tangled up in spam. *The New York Times Magazine*, February 9, 2003.

Graham, Paul. 2002. A plan for spam. http://www.paulgraham.com/spam.html

Graham-Cumming, John. 2003. The spammer's compendium. http://spamconference.org/proceedings2003.html

Krim, Jonathan. 2003. Spam's cost to business escalates. *The Washington Post*, March 13, 2003, page A01. http://www.washingtonpost.com/ac2/wp-dyn/A17754-2003Mar12.html

Templeton, Brad. Origin of the term "spam" to mean net abuse. http://www.templetons.com/brad/spamterm.html