Policy Responses to Spam
PORTIA Reading Group
March 9, 2005

James Grimmelmann and Becky Bolin

DEFINING SPAM

Though almost everyone agrees spam is bad, even experts disagree about the correct definition of "spam." Common formulations include one or more of the elements "unsolicited," "commercial," "bulk," and "email." But all of these elements create problems for someone, and there seem be examples of communications that are considered "spam" despite not satisfying one or more of these elements. While arriving at an exact definition may not be possible, working through the difficulties should at least provide both a better sense of what is objectionable about spam and also a sense of the complexities involved.

"Unsolicited bulk commercial email" is more of a paradigm case of spam than an explicit definition. "Unsolicited" is critical to the definition, but is also hard to define precisely at its edges. "Bulk" is also central to the idea, but has odd corners of over- and under-inclusiveness related to the question of where one measures the flow. "Commercial" is perhaps a proxy for amount, given that the heart of the spam flood is commercial, but may also reflect hesitation to stop charitable or political messages to have relatively free rein. And "email" is not so important to the definition: it is just the medium currently most afflicted by spam, which is spreading to mobile phone text-messaging (SMS spam), instant messaging (spim), and VOIP (spit).

"Unsolicited"

"Unsolicited" raises three issues: who is authorized to consent to receiving email, what actions of theirs constitute consent, and how far does the consent extend?

Whose Consent?

Many online newspapers have 'mail this story to a friend' features; online greeting card systems generate notification emails to anyone with an email address; your mother forwards you another obnoxious joke. Few would call these emails 'spam,' but they are technically unsolicited by the recipient. Compounding the problem, many services allow users to generate large mailing lists. Plaxo, for example, sends advertisements to everyone in a user's contact book, urging them to join Plaxo; Yahoo Groups enables list moderators to add email addresses in bulk and attaches an advertisement to the bottom of every message.

What Constitutes Consent?

This is a hotly-debated policy problem: most agree that an express request not to receive particular emails should be honored ('opting out'), but they then part company. Many consumer activists believe that some act of affirmative consent is necessary ('opting in'), while many industry groups believe that consent can be fairly implied from some prior business dealing or that a first message is critical to free speech rights and the value of advertising.

How Far Does Consent Extend?

Large companies often deal with customers through several different divisions: for example, one may have both a checking account and a mortgage through the same bank, which may be handled by entirely separate IT departments. Does consent to receive notices of new banking account offerings through email constitute consent to receive offers of home equity loans? But, conversely, must the bank treat a refusal to receive home equity offers as a blanket refusal of all email, including savings account offers? At least at the level of consumer expectations, no categorical answer seems acceptable: the same emails will be considered spam by some and not by others. Much depends on the individual's history of dealings with the

company and the extent to which the individual distinguishes among the company's divisions. Given the full menagerie of modern corporate structures—e.g. local franchisors of central offices, wholly-owned subsidiaries in conglomerates, and independent contractors—permission or refusal of solicitation is rarely a simple yes or no matter.

The Legal Status Quo

The United States avoided the consent problems altogether by requiring an "opt-out" in CAN-SPAM. That is, anyone can send you anything (with some exceptions) until you tell them to stop. This system is based on the default in both physical mail (you can opt out of any mailing by notifying your postmaster) and telephone (you can opt out using the Do-Not-Call list). In the E.U., member states must use an "opt-in" regime, but can define this as they see fit. The Australian opt-in regime uses a circular opt-in definition that provides no help whatsoever.

"Bulk"

There are at least two troublesome definitional issues when it comes to defining "bulk." Any set threshold may be underinclusive for particularly offensive mailings, but overinclusive for mailings that are large but appropriately targeted. Further, it is not clear whether one should count at the sender's end or the recipient's end. These issues raises the question of whether "bulk" is a proxy for some other, more pertinent criterion.

Non-Bulky Spam

One can imagine scenarios in which even individually-addressed emails are fiercely unwanted—certainly a single email in violation of a restraining order is actionable. It is probably the case that the wrong involved in such cases is best of not being classified as "spam," but we should not rule out the possibility: LawMeme receives an article submission approximately twice a month from the same Moldovan company, despite repeated requests for them to cease, and one of the authors of this paper receives several emails a day from the same individual alleging various corruption on the part of high officials in the Indian government. These communications do not rise to the level of bulk; the ISP at issue, Yale, likely has no other correspondence bogging the system down. This mail has small effect assuming there is no volume to aggregate, but is equally annoying.

Bulky Non-Spam

Consider the Hamidi problem: a disgruntled ex-Intel employee sent a series of six emails to current Intel employees, urging them to join a group of disgruntled Intel employees. He honored requests from individuals not to be included in future emails, but his initial mailing was sent to every current Intel employee, thanks to his use of a leaked email list. While the California Supreme Court ruled in his favor on Intel's attempt to sue him for "trespass to chattels," his case poses spam policy questions. His list was carefully selected, but it was large, and the manner in which he employed it led to a large collective burden on a single entity: Intel. If Hamidi had instead driven up to Intel headquarters to deliver thousands of paper letters to Intel employees or to leave orange flyers on their cars, would his behavior have been more reprehensible, more spammy than his similar actions online?

Counting Problems

On the other hand, consider the action-center problem. Many grassroots groups have turned extensively to 'action centers' to mobilize their constituencies. On important issues, the groups will provide a web site through which constituencies may email and fax government officials or corporate executives with a single click. These centers streamline the traditional process of sending in thousands of letters and flooding offices with phone calls. Are these activities spam? The Forest Service may have thought so: in 2002 and 2003 it reportedly installed spam filters on email sent to an address used to take public comments, and many grassroots groups believed that those filters were set to reject email coming from action centers. Here, each of the senders was sending non-bulk email, but the effect on the recipient clearly was bulk—perhaps magnified by the agency's statutory duty to file and respond to public comments.

"Commercial"

"Commercial" is almost certainly underinclusive; much unwanted bulk email is not, on its surface, commercial. But it is hard to find a more inclusive adjective that does not also sweep in some email we do not consider "spam."

Charitable Emails

For many Americans in an age of a marketers' do-not-mail registry, the single largest source of mail is charities asking for donations. Similarly, the federal Do-Not-Call list exempts calls by non-profits: the reasoning is based on increased First Amendment respect for charitable and political speech, but it is unclear whether the recipients consider these communications any less irritating. Almost immediately after the Indian Ocean tsunami, spammers took to sending out floods of fake solicitations for contributions for the victims. These heartless schemes are illegal because of their fraudulence, but had the spammers been taking merely a 10% 'administrative' fee and forwarding the rest to real charities, should that have insulated these schemes from all spam regulation? If not, on what basis can their appeals be distinguished from appeals by the Red Cross? (One possible answer: on the basis of no preexisting relationship.) If so, are we comfortable telling other charities not to send bulk email, even in cases of great emergency?

Political Emails

Political speech is also a category that has had enormous trouble with spam regulation. The poster child for legitimate political email is MoveOn, which has a mailing list of millions. Unfortunately, its mailings to its active membership have often been filtered out by ISPs' anti-spam measures. Mailing lists and other legitimate large-scale email systems have also been crippled by challenge-response, sender-identification, and pay-to-send technical anti-spam measures. Lest we give political speech a complete pass, however, there have also been scandals involving political non-profits sharing their membership rosters surprisingly widely. Many people may well have received much mail they considered spam as a consequence of having given their email address and permission-to-email to the ACLU, which has used the same data-mining and list-sharing techniques beloved by commercial marketers and other professional fundraisers.

Nutjob Emails

Finally, there is nutjob speech: bulk emails sent out by people in pursuit of some strange private agenda but with no apparent commercial motivation. The same author who has received

repeated emails about corruption in the Indian government has also received plain-text emails saying "Call out Gouranga be happy!!! Gouranga Gouranga Gouranga .... That which brings the highest happiness!!" and emails purporting to be from a time traveler stuck in our age and desperately in search of a "dimensional warp generator." Both of these are relatively well-known: many others have received the same emails, strongly suggesting that they are being sent out in great bulk. While the great bulk of spam is driven by grey market advertising, outright fraud, and other crass commercialism, there is also a residue of inexplicable, perhaps undeterrable spam. (Although one notes that even these spammers need access to large lists of addresses, and that one of the principal techniques by which such lists propagate is . . . spam advertisements selling lists of addresses.)

The Legal Status Quo

In the United States, the FTC is the current center of attention for defining what constitutes "commercial" email. It had been expected to issue rules to flesh out the CAN-SPAM standards of a commercial "primary purpose," since CAN-SPAM defines spam as unsolicited (mail whose primary purpose is commercial. After missing the CAN-SPAM-mandated deadline, the FTC claimed this rulemaking was just too difficult.

Indeed, a mailing with a commercial sponsor (like a traditional newspaper) may not have a commercial primary purpose. A page excerpted from the New York Times with only an ad may be. Examples from the Senate CAN-SPAM report—a "bank account statement" and a "product recall notice"—are not especially helpful, but the report does note that messages like these are exempted "even if the message includes at the bottom some promotional information about the sender's other products." A product recall e-mail with a suggestion for a replacement may or may not be primarily commercial.

"Email"

Although we often think of spam as an "email" problem, spam both predates email and will long outlive it. With apologies to Arthur C. Clarke and Clay Shirky, any sufficiently advanced technology is indistinguishable from a spam vector.

Spam Before Email

We are familiar with the problems of postal spam, acknowledged both in one's ability to ask the post office to discard mail from certain senders, and in the Direct Marketing Association's opt-out Mail Preference Service. Telephone spam led first to the MPS's telephone analogue—the Telephone Preference Service—and then to the national Do-Not-Call list. And online spam is generally acknowledged to have started not with email spam but with USENET spam, in the infamous 1994 "Green Card Lottery" spam.

Web-Based Spam (Herein of Search Engines)

The hottest fronts in the spam wars today may not even be email. The centrality of search engines to many people's Internet experiences has created a lucrative niche for those who can manipulate search engine rankings to drive hapless searchers to their own advertising-laden pages. When this technique was confined to creating fake sites whose meta tags or link structure made them attractive to search engine ranking algorithms, the damage was largely visible only in the distorted rankings. But the proliferation of online communities and applications that allow users to edit or add content has led to attempts to plant fake links in these communities. Blogs with comment features fall prey to software bots that post link-heavy comments. Open mailing

lists with public archives receive link-heavy emails. Even public indexes of web site statistics are vulnerable: referrer log spam involves sending requests for web pages with faked HTTP referrer fields so that the forged URLs show up in the displays of site statistics. LawMeme's comment fields are riddled with "comment spam:" fake comments riddled with links to sites run by spammers, designed to trick Google into siphoning some of LawMeme's high PageRank over to the spammers.

Next-Generation Spam

New communications technologies are also being spammed directly. Instant messenger spam (spim), while not epidemic yet, is growing. Initially used by renegade viruses, it is now the domain of true spammers, and is starting to generate criminal prosecutions. Text-message spam is becoming increasingly common, and clever scammers have figured out how to manipulate the SMS system to have users call back the equivalent of a 900 number. Abroad, experience with SMS spam has lead to innovation in technology and varyingly draconian laws. Voice-over-IP systems, by making phone calls costless at the margin, threaten to make telephone spam to VoIP users (spit) costless at the margin. Even the domain name system itself has been the subject of a sort of spam: bulk registrations of domains with the intent of showing ads to hapless typo-typers or selling the domains to those who realize too late that they want a particular domain.

POLICY VALUES

Many important social values are at stake in the spam wars. They include civil liberties, such as freedom of speech and privacy, economic efficiency, and the freedom- and innovation-enhancing architecture of the Internet.

Freedom of Speech

Freedom of speech is a human right recognized in the United States in the form of the First Amendment. Email, the so-called killer app of the Internet, has been an incredible engine for freedom of speech. It enables cheap, near-instantaneous communication by anyone with access to a computer. The power of email communications has been important to activists fighting oppressive regimes, to groups debating pressing political issues, to grandparents receiving updates and photos, to friends keeping up their existing social ties, to new social networks building a democratic culture, and to many others. Email has given the world a wonderful conduit for free speech . . .and for spam.

Spam reduces the value of email; in so doing, it undermines email's free-speech value. Email's ability both to narrowcast like phones and to broadcast like traditional mass media has made it a paradoxical creature. Many of the solutions that would restrict spam would also undermine the power of email as a forum for free speech. Finding a baby-preserving mechanism for discarding the bathwater is not easy: senders, recipients, and third parties all have legitimate free speech interests.

Right to Solicit

The most obvious free speech interest implicated by the spam debate is the right of senders to communicate. Doctrinally, the United States Supreme Court has held that "commercial" speech—speech that merely proposes a transaction—requires less protection than more important forms of speech such as political speech. Nonetheless, commercial speech is protected under the First Amendment, and cannot simply be banned. Under the *Central Hudson* test, courts ask four questions about possible restrictions on commercial speech:

First, if the speech itself is illegal or misleading, it is unprotected. (Note that a significant portion of spam probably falls within this category: over 20% according to some studies.) Second, the restriction must serve a "substantial" governmental interest, and third, it must directly advance that interest. Fourth, there must not be a "more limited restriction" available that would serve the governmental interest just as well. If any of these final three tests fails, then the court will strike down the restriction. In practice, it is the final test—the availability of more limited but equally effective restrictions—that is often determinative.

One sometimes sees the argument that laws targeting commercial speech but leaving non-commercial speech alone discriminate among speech on the basis of its content, and therefore must be subjected to the stricter analysis used to examine "content-based restrictions." A legal challenge of this sort was raised to the Do-Not-Call list, for example. Courts routinely reject such challenges; they rest on a deliberate misreading of what "content-based" means and if they were allowed, they would undermine the Supreme Court's holding that commercial speech need not be protected as strongly as "core" speech.

Several arguments are typically made for giving commercial speech some protection. First, restraints on government's power to restrict commercial speech are restraints on government's power over the marketplace; they keep legislatures from certain excesses of economic favoritism. Second, they promote healthy competition: they allow businesses to

advertise and drum up business in a relatively unfettered way, helping new companies enter markets and increasing consumers' choices and education about possible choices. Third, because commercial speech often shades into political or expressive speech (think of advertisements announcing a business's commitment to Black History Month), protecting commercial speech has the collateral effect of protecting these other forms of speech. All of these arguments are important in the spam context. Recall in particular the difficulty of drawing a "commercial" boundary around spam.

Collateral Damage

Another critical free speech interest is that of non-spammers who use email. Spammers indirectly undermine the speech value of non-spam email by degrading the usefulness of email systems. Unfortunately, many of the technical solutions in use to abate spam problems have caused serious disruption to non-spam email. The cure is often worse than the disease, as far as non-spam freedom of speech in email is concerned.

Political and social groups that rely on mailing lists to communicate have been hard-hit by over-zealous filtering. So too have individual Internet users from countries with unfortunate reputations for spam, or even from ISPs that have allowed spam to pass in the past, or ISPs with gullible customers, or even just large quantities of subsribers. Today, AOL filters about 80% of incoming mail before it reaches any sort of mailbox. Any email user not running her own mail server has had some of her messages filtered out by anti-spam systems, and had mail sent to her filtered. ISP actions sometimes go even further: AOL now blocks HTTP access to websites it believes to be run by spammers.

This collateral damage is in many cases severe; were the filtering being done by a government, it might be illegal as an "overbroad" restriction on freedom of speech. Of course, most filtering is being done by private parties, who can point to contracts consenting to filtering, and to whom the First Amendment does not directly apply. But the problems of collateral damage both indicate what can go wrong, on a policy level, even with solutions created through private ordering, and also may place limits on the techniques governments can use to fight spam.

Recipients' Speech Interests

Although free speech is usually thought of in terms of a right to speak, every communication involves both a speaker and a recipient. The right to receive information is often important in First Amendment law. For example, it may give recipients a right to insist that governmental intermediaries, like the Post Office, not intercept messages intended for them. Any anti-spam "solution" that prevents email recipients from receiving emails they want to receive is both bad policy and will encounter significant legal hurdles, if done by government.

On the other hand, recipients may also have a speech interest in *not* receiving communications they do not wish to receive. One may, for example (unless one is a member of Congress), instruct the Post Office not to deliver any future mail from a given sender. Various free speech-implicating legal doctrines—such as the land-use rules that govern offensive billboards or the rules that govern door-to-door solicitation—take some account of the wishes of the potential recipient of a communication not to be subjected to it. This interest is often balanced against speakers' interest in being able to try to attract the attention of a supposedly unwilling listener.

The rising flood of spam has not been kind to recipients' interest in being free from unwanted speech. In many cases, that speech is profoundly offensive. Its sheer volume makes it

harder to listen to legitimate speech. Much of the anger directed at spam may be derived from spam's inroads on this speech interest.

Freedom to Receive and Agency Problems

There is a speech interest in being able to receive speech that others consider unwanted. About 14% of email users read their spam for content, and about 4% have bought products from spam. Spam, like billboards, may be annnoying, but does provide some people with information, and they some kind of right to receive it.

That said, virtually all email users have subcontracted out this right. When you sign up for a Yale account, you agree that Yale can delete any mail it wants for any cause with no notice. There is nothing stopping religious universities from screening mail for curse words or blocking foreign mail, and no legal reason your ISP can't just arbitrarily block mail, even in ways otherwise offensive. Yale, for example, could choose to block all mail sent from names that sound feminine.

Filtering on the server side creates a transparency problem—the algorithm used to determine what items to filter may be a trade secret, indeed a quite valuable one. Often mail isn't even bounced so senders don't even know they are being deleted. There is, in theory, a robust market for mail providers; you could always change, but good luck figuring out how your new ISP's policy is different before it is too late.

We can imagine other scenarios in which the right to receive information is more valuable than the convenience of a third party. In fact, rather than imagine, we could simply look to South Africa. After a history of filtering during the apartheid era, it now provides a constitutional right both to send and to receive content. Thus, it is unconstitutional to delete someone else's mail, a doctrine codified in recent law. Such a law protects speech values in some sense, but it easy to picture doomsday scenarios for ISPs and an even worse deluge of spam. If anything, a law like this may just force explicit contracting-around, resulting in the same filtering processes.

Privacy

Free speech is not the only civil liberty at stake in the spam wars. Privacy values are also implicated, as they are in almost every major cyberlaw controversy. If one views the problem of spam as a problem of unaccountable senders, then increasing accountability should reduce spam. The problem, from a privacy point of view, is that this increased accountability will come at the expense of various features that contribute to sender privacy. Spam is hardly unique in this respect: one sees a similar dynamic at work with copyright infringement, hate speech, sexually-explicit speech, offshore gambling, terrorist communications, cracking, and many other hot-button issues of online behavior. But there are several additional twists on privacy issues that spam raises in troubling ways.

Anonymity

The anonymity debates are by now almost trite. Any measure which facilitates the location of email senders in the real world will be good insofar as it helps good people find bad senders, bad insofar as it helps bad people find good senders, and highly controversial insofar as "good" and "bad" are contested concepts with respect to particular behavior. Further, as cyberlaw experts frequently note, architectural or legal changes that reduce anonymity in one online context are likely to be leveraged extremely quickly to reduce anonymity in others. Commercial entities that have anonymity-piercing user data will find that data subpoenaed by

law enforcement and by other commercial entities; nations that deploy anonymity-reducing measures to combat spammers will see those measures imitated by other nations to combat dissidents. All online swords are double-edged.

Anonymity is, as this discussion suggests, linked to free speech values. In the law, the right to speak anonymously is considered an important strand of the right to speak (and has suggested natural outgrowths, such as Julie Cohen's proposed right to read anonymously). But the classic Warren-Brandeis definition of privacy as a general right to be let alone extends further. A system that restricts senders' anonymity is also likely to make it harder to receive email anonymously. In an age in which Internet access is increasingly becoming a necessity to all sorts of social and economic interactions, tying that access to offline identity implicates more and more of one's "private" identity. Danny O'Brien has said that the category of the "private" is collapsing the categories of the "secret" and the "public;" policy responses to spam may be a significant part of this trend.

Note that some anti-spam schemes are more compatible with the preservation of anonymity than others, and that many can be implemented in ways that are more or less anonymity-preserving. A pay-to-send system that required payment in cash, for example, could allow anonymous payments or require the use of more traceable financial instruments. Sender authentication at the ISP level is less anonymity-piercing that sender authentication at the user level. Ex post attempts to catch known spammers raise more questions of governmental abuse of identity unmasking, but ex ante technical mandates raise more questions of private abuse. Anonymity is not univalent, and different schemes shape the landscape of privacy in different ways.

Pseudonymity is often suggested as a workable intermediate alternative: actions would be anonymous unless some technical or legal threshold was crossed that would permit revelation of identity. Such a system would be both more and less anonymous than today's email infrastructure. It would be more anonymous in that it would have to build in an infrastructure in which peripherally identifying details were obscured (contrast today's system of explicit SMTP delivery headers on every message), but less anonymous in that its effectiveness would depend on the availability on techniques for unmasking senders. These debates are strikingly reminiscent of the debates over the Clipper chip.

Cryptography

Another way in which many spam debates recall the Clipper debates is the mooting of cryptography as both an identity-revealing and identity-concealing device. Many of the sender-identification or pay-to-send systems depend on rich cryptographic infrastructures. The design of those infrastructures establishes an expected set of power relations around identity and content: one's name and one's words are known to certain parties, and not to others. Different designs privilege different parties.

But that is not all; cryptosystems also frequently have failure modes. When the cryptography involved in an identity architecture fails, the power relations shift. Clipper was meant to privilege government actors—subject to judicial review—while limiting third-party access to the sensitive content of identity. But many computer scientists feared that hackers could combine government databases to spy on any communication, and cryptographer Matt Blaze discovered a simple hash-collision flaw that would have undercut the utility of the keys held in escrow. Introducing cryptosystems to the anti-spam equation introduces a new set of risks that the privacy "guarantees" in the system will be too strong or too weak in an unanticipated way.

Database Accountability

Privacy issues also enter the spam debate through another current hot-button cyberpolicy front: the assembly of large databases containing personal or personally-identifiable information. These databases raise consumer-protection concerns against the companies which typically create them. They raise government-surveillance concerns—since the government might ask for access to them. And they raise identity-theft and stalking concerns, since crackers might steal them. It is impossible to say at the moment what the right boundaries on such databases are or should be; both public opinion and policymaker opinion are in flux. Whether and how data collection or use should be regulated is a contentious question, which various activists taking positions ranging from outright prohibition to demanding strong intellectual property protections for databases.

Such databases enter the spam world in many places. Most obviously, spammers themselves depend on huge databases of addresses (and, recursively, databases of addresses are frequently offered for sale through spam). Some might say that the harvesting of spammable email addresses without recipient consent—and the subsequent marketing of those databases—is offensive over and above their use. Effectively dealing with such practices reduces to the known hard problem of dealing with spammers, however.

At the moment, many anti-spam activists are themselves creating large databases, most obviously of email addresses, IP addresses, and domain names used by spammers. Being mistakenly put in such a database can be devastating; there is arguably just as strong a privacy/personal integrity interest in having accurate information (or no information) recorded about one's emailing habits as about one's credit history. The DNS system is increasingly being used as a very rough database about senders.

Many anti-spam proposals would create new databases. A do-not-email list, most obviously, would be a tempting, extraordinarily lucrative target for crackers. Many sender-authentication technologies would require credentialing databases; shared blacklists and whitelists also raise accuracy and security concerns.

Other anti-spam proposals try to limit databases. For example, Australia bans several kinds of spamming software, including sophisticated bots or "scrapers" which harvest email from the Internet. This policy would almost certainly be overinclusive in the United States. It raises many of the same innovation- and free-speech-damping concerns that the DMCA does.

Economics

In a broad sense, economic efficiency is an important criterion in almost any debate about Internet policy. More narrowly, spam policy is economic policy in several linked ways. Most obviously, enormous quantities of unwanted email create enormous inefficiencies. The spam wars are also generating inefficient technological arms races.

Advertising, Resources, and Attention

One way of thinking about spam is as advertising. Spammers offer consumers information about possible products. The enormous volume of spam suggests that this advertising is profitable, that there are many people—somewhere between 4 and 15% of Internet users—who do in fact find the advertised products appealing enough to purchase. A much larger number read some portion of their spam. Spam is valuable economic activity, albeit with massive externalities. One important counter-argument is that many consumers who respond to spam are being defrauded and later regret their gullibility. Another is that spam is in fact

artificially subsidized by an email structure in which spammers do not pay the full costs that their emails generate.

Some of those costs are concrete. Spam consumes bandwidth, cycles, and storage space, and though all three of these commodities are falling in price, none is free. Even on the most optimistic view of spam, most of these costs are pure waste. Most are borne by ISPs which hire experts, implement software, and buy storage and routing facilities. These costs may be as high as $5 per mailbox. A system which better internalizes the costs of spam or one which improves the accuracy with which messages go to people who want them will reduce this waste.

Other costs are more intangible but no less real. Spam consumes attention. Every minute spent deleting spams from one's inbox is a wasted minute. Advertisers (including spammers) understand that attention (sometimes "eyeballs") is a valuable commodity. The act of deleting offensive and annoying spam wastes that commodity in enormous quantities. Effective anti-spam solutions will reduce this waste, as well.

Some extremely graphic spam most likely only has value to an extremely small portion of recipients—the rest are offended, perhaps made even more heistant to use email. Fraudulent mail has no redeeming social value, except perhaps in a dystopian Social Darwinist vision in which spam helps weed out the gullible.

Destructive Self-Help and Arms Races

A major concern of legal policy is to prevent mutually destructive behavior. Thus, for example, landlords are often forbidden from locking tenants out except through designated court process out of a fear that self-help lockouts will cause tenants to break down doors or engage in violence. The law prefers to divert these potential fights into the legal system early, rather than take the risk that parties will take self-help too far and start damaging property or attacking each other.

A related idea is the law's desire to prevent arms races. Here, the opponents are not hurting each other in legally-cognizable ways, but they are engaged in pointless effort trying to outdo each other. These efforts are pointless because each new investment renders the other side's previous investments worthless; the parties often spend enormous sums just to maintain the status quo. It would be better for all concerned if the law decreed a solution (often, for example, fixing the status quo) so that they could both spend their money elsewhere, rather than wasting it fighting each other. Of course, this response assumes, perhaps incorrectly, that spam is a regulable market in the first place.

Spammers currently engage in enormous amounts of destructive behavior. One common spam technique involves using a virus or worm to infect a home computer with a broadband connection and use that computer to send out spam; these computer are called "zombies." Such infections undercut Internet security and degrade the usefulness of consumers' computers. Spammers have also been known to launch enormous denial-of-service attacks on anti-spam sites and to use spam itself as a bludgeon against those they dislike. For their part, anti-spam activists have sometimes turned to the same tactics; Lycos announced (and then withdrew in shame) a vigilante campaign of dubious legality in which its users could user their computers to engage in denial-of-service attacks on known spammers. Within days, spammers were sending out virus-laden "add-ons" to the troublesome program to wreak havoc on vigilante machines.

The spam wars are also generating worrisome arms races. The most obvious is the race between spammers and filter writers. As filters learn to recognize classes of spam, spammers mutate their spam to evade filters; v1agr^ is just the beginning. Spammers now lie about their locations, their ISPs, their route to your computer. Most of this effort is pure waste; it has the

collateral fallout effect of making spams messier, harder to spot, and more annoying to read when they do get through. Although filter-writers have scored some large temporary victories—some filters when first released are more statistically effective than modern birth control—there is little reason to expect that they will ever be able to defeat spam on their own. Spammers will always be in the position of the attackers: all they need to do is find a way to foil the current generation of filters, poke a tiny hole in the armor. The filter-writers, by responding to a new spam technique, will force the spammers to upgrade, but the long-term trend is likely to be that the filters are usually one step behind (and occasionally half-a-step or a step-and-a-half). A good approach to spam will not encourage similar arms races to develop along some other front.

Architecture

Loosely speaking, "architectural" values have to do with the technical and institutional structures of the computer networks we currently have. To the extent that the spam wars make the Internet function better or worse by changing its topology, tolerance of innovation, or universality, we might call those effects "architectural."

End-to-End Values

Since Saltzer, Reed, and Clark's famous "End-to-End Arguments in System Design," computer scientists have used the term "end-to-end" to describe the virtues of a specific form of network architecture. An end-to-end network supplies consistent and simple connectivity between its nodes; it does not generally discriminate among content, applications, or endpoints. Any "intelligence" in the network is left up to specific applications and layered on top of the simple connectivity provided by the network. The Internet is generally considered to be largely end-to-end; the old AT&T switched telephone network was much less end-to-end.

An end-to-end network tends to have a number of good properties. First, it encourages decentralized innovation. Two users can agree on a new communications application and try it out without needing to seek permission from other users or network authorities. Second, it promotes fairness: an end-to-end network does not block communications for invidious reasons or favor some kinds of content. Third, it promotes simplicity and reliability. An end-to-end network typically involves simpler, more-easily implementable components and is therefore also easier to troubleshoot when those components malfunction.

Many anti-spam proposals, including almost every server-side solution, violate end-to-end principles in one way or another. Server-side filtering breaks end-to-end delivery guarantees for email. Challenge-response systems put sophisticated anti-spam intelligence into the network itself, as do sender-identification systems. A significant variation in legal regimes governing spam is also an end-to-end violation; it means that messages will be regarded differentially based on properties that may be very difficult for a user to observe or to specify. These changes may reduce the utility of end-to-end virtues both for email and for the Internet in general.

Transition Issues

Another related issue is the degree to which a proposed solution depends on coordination among a large number of responsible parties. The Internet has succeeded, in significant part, because of its decentralized nature. A change which will only be effective if every Internet participant agrees will be ineffective; there are still, after all, gopher servers in regular use. Broadly speaking, the more parties who will have to coordinate their upgrades to implement a solution, the harder such a solution will be to arrange and the more disruptive the transition will be. Proposals which require large-scale upgrades face an uphill battle.

As an illustration, one reason that SPF has been more widely-adopted than more comprehensive sender-identification techniques is that it is easier to engraft onto existing technologies. At the recipient-server end, the upgrade is voluntary; at the sending-server end, the upgrade can be accomplished by updating DNS records, rather than requiring the deployment of significant new software.

One axis along which to evaluate anti-spam proposals is to ask which pieces of software would need to be rewritten to make it work. Would it require a change to all mailers? To all email clients? To mail transport agents? To IP routers? To all DNS servers? The powerful network effects that accompany the use of common standards act as a powerful check on changes to those standards. Precisely because they are successful as standards, upgrades are expensive, and especially so if backwards compatibility cannot be maintained.

TECHNICAL RESPONSES

Because many legal approaches involve specific technology mandates or prohibitions, it is helpful to discuss technical considerations first. Technical approaches, in turn, can be roughly divided into three large groups: filtering systems, sender verification systems, and pay-to-send systems. We begin this section by discussing several basic tools and concepts that apply to many different technical solutions.

Basic Tools

Clients and Servers

The same policy issues recur with every technical solution when choosing between client-side and server-side implementations. A client-side approach is more accountable to an individual user. In cases of false positives, it is much easier for the affected user to discover, and if necessary, modify, the details of the implementation. It can be customized to her preferences about wanted and unwanted emails and can learn from patterns specific to her. It also has an early-adopter advantage in that individuals users can experiment with systems and spontaneously adopt them in a bottom-up fashion.

A server-side approach, on the other hand, has advantages of significant aggregation. It can draw on a much larger corpus of data about spam and non-spam emails and can recognize patterns in which multiple users are receiving a coordinated spam barrage. Because it operates earlier in the process, it also can avoid spending server time and storage space on spam emails that client-side filtering would have rejected anyway. Finally, it has an institutional-adopter advantage: an ISP can deploy measures on a coordinated basis and raise a large number of users at once to full capability.

Challenge-Response

A challenge-response system attempts to verify the legitimacy of a message by asking a question of its sender that only a legitimate sender will be able to answer. These challenges take many forms: some are mean for machines, while others are meant for humans. They may make either trivial or quite substantial demands.

Systems with a challenge-response architecture consistently raise a problem of backwards compatibility. The current SMTP architecture does not support an extensive two-way communication system, and neither does the typical email application level (at a granularity finer than exchanging two complete email messages). Thus, attempts to create a challenge-response system face a consistent problem of what to do with legitimate senders who are not capable of understanding the challenge. In a sense, all new demands placed on the sender pose an analogous problem; challenge-response just puts the sender on the spot very late in the process, when it is actually attempting to send a message.

Data Pooling

Many current and proposed anti-spam efforts involve some kind of collective data-gathering and aggregation. Honeypots, spam reporting systems, vigilante blacklistings, friend-of-a-friend systems, and many more depend upon many email users providing their individual data for the purposes of a collective anti-spam effort.

A generic problem in data pooling is that of false data, since almost any pooling effort will include some data source that is unreliable or actively malicious. Spammers have been known to engage in joe-jobs: trying to insert anti-spam services onto blocklists in the hopes that

automated blocking services will turn on each other. Even without active malice, pooling unreliable data creates a need for self-correcting systems and practices.

Blacklisting and Whitelisting

Blacklisting is the process of marking all communication from a particular sender as spam. One can blacklist within many different namespaces: email addresses, hostnames, IP blocks. Whitelisting is the opposite: marking all communication from a particular sender as not being spam—and immediately bypassing all other spam checks. One sometimes sees the term greylisting to describe some practice of flagging particular emails for heightened scrutiny: one will still let them on the plane, but only after a pat-down search and bag inspection.

Many anti-spam groups publish their own blacklists using secret complaints, deciding which mail they consider abusive. United States-based ones have been the targets of legal assaults by spammers, and are now almost entirely gone. Most such lists are now run out of places with little to no accountability—e.g., Irkutsk, Russia. The blacklist Yale is using is run by an advocacy group in the U.K., Spamhaus, which is constantly involved in litigation over the list. Spamhaus itself is not even a mail server. It merely publishes a list and ISPs voluntarily block the mail servers. These solutions are in a sense extralegal; the rely on the norms of the relevant watchdog group.

One immediate question about any colorlisting system is the choice of namespace. Both the ease of verifying the correctness of the supplied name and the ease of obtaining a new name in the given space are issues. The easier it is to verify the correctness of a name and the harder it is to create a new name, the more effective colorlisting will be at identifying a stable attribute of a sender (sender-identification systems aim to enhance these values). Also, the fewer senders associated with a given name, the more effectively targeted colorlisting will be: the same hostname may send out spam and non-spam mail from different users, but most email addresses are either spammers or not spammers. Blacklisting at the hostname or IP level raises perennial issues of collateral damage to non-spammers' email. For example, Spamhaus once blocked all of Yahoo! Shops because of the actions of a few abusive mailers. Such responses were devastating to law-abiding businesses, but certainly forced Yahoo! to police its network.

Note an important asymmetry: whitelists tend to be more reliable than blacklists, because spammers tend not to send out most of their mail impersonating actual, particular individuals. While some spams and viruses do use large databases of harvested email addresses (or build such databases themselves from the address books on infected computers) as faked "From:" lines, improvements in sender-identification technology have made such attacks less fruitful. Spammers more frequently attempt to impersonate wholly unknown senders, rather than particular senders.

Finally, note that colorlisting is almost never considered a complete solution. It is an important technique, but it always leaves open the question of what to do with email from a sender whose name is unrecognized. Indeed, because colorlisting takes care of recognized names, the typical spam problem reduces to the problem of how to handle messages from unknown senders.

Sender Identification

Sender identification systems attempt to improve the reliability of information about the sender of a message. First, such reliability can make colorlisting more effective by making it harder to forge messages from a given sender. Note that this technique, by itself, makes only

whitelists effective, because it deals only with the impersonation of a known sender. . A digital signature doesn't add much to one's knowledge when dealing with a heretofore unknown sender.

The more significant role of sender identification comes when the identification also serves an accountability purpose: the identification provides some guarantee that a sender can be held accountable if messages from them turn out to be spam. For this technique to work, the identification credential needs to be one that only a sufficiently accountable sender will be able to present. Wrapped up in this definition of "accountable" is the idea that other institutions—e.g., the legal system, social norms, or vigilante retaliation—can deal with spammers once they have been traced through the identification technology.

Identifying Clients

The simplest identification in an email is its From: header. Digital signatures and other cryptographic techniques can be used quite directly to prevent forgery of the sender's email address. This plan is being advocated by several major ISPs, including Earthlink and Yahoo!.

To date, digital signatures have not made any substantial progress in the email authentication space, although they are extensively in use for HTTPS server authentication. Their widespread adoption there is in part thanks to their coupling with a related technique: third-party certification, in which an institution with substantial credibility agrees to vouch for a given HTTPS server. In the HTTPS world, this voucher takes the form of participation in a digital signature scheme that is leveraged onto users' desktops through their web browsers. There have been similar proposals for email, along with proposals that use the third-party certification even without the cryptographic aspect. The idea is that some institution credentials the email—possibly by passing it through its servers or through offering specified content that only trusted senders are allowed to use—and one trusts all email coming through that institution.

Some interesting schemes use third-party certification without cryptography. The Habeas mark, for example, is a copyrighted work inserted into email headers by users. Habeas aims to prevent spammers from using the work by suing them for copyright infringement, and it has even been successful in enforcing it in the courtroom occasionally.

Another interesting variation on third-party certification is shared whitelisting, which combines third-party certification with data pooling based on social networks. Here, one's friends are the network: each friend vouches for the other members in her address book. In effect, groups of users can cooperate to produce a shared whitelist of known-non-spammers.

Identifying Servers

The other major form of sender identity attached to most emails today consists of the headers which indicate the logical chain of SMTP servers that the message followed en route to the recipient. In theory, identifying the originating SMTP server for a message promotes substantial accountability, because that server corresponds to a physical machine on a physical network, and the ISP that runs that network is both an intermediary that can be held accountable and can be asked to hold the server's operator accountable.

The first issue is the accuracy of the chain of SMTP headers. For overall identification to work, every server along the way must tell the truth about itself and the previous servers in the chain. Thus, there is a logically prior problem to preventing spam itself, that of forcing honest practices on all SMTP servers. Blacklists based on pooled information have been used at this left to enforce compliance with honest practices, along with social sanctions and occasional vigilantism. For example, "open relays," which accept inbound STMP traffic and pass it along without checking the sender's identity, are systematically hounded off the network through

blacklisting and social pressure. However, because these blacklists operate both at the server level and not even directly on spam-producing (but merely spam-facilitating) servers, problems of overinclusion have caused repeated controversy, of which Verizon's apparent blocking of all emails from large parts of Europe is only the most recent.

More recently, attention has focused on ways to use this chain of attribution to enforce stronger guarantees of identity, often by linking the IP information to the email address information. One approach seeing increased adoption is Sender ID, already in use by Microsoft's Hotmail, in which the DNS system is used to provide a rough-and-ready guarantee that an email comes from a computer authorized to send email from a given domain. An ISP's DNS records are tailored to state that only certain IP addresses are authorized to send emails that claim to have return addresses in that domain; computers adhering to the Sender ID standard, on seeing an email address, check the DNS record for the domain of that address. This system promotes accountability goals by forcing a would-be spammer who wants to send email "from" a domain to compromise the accountability systems put in place by that domain's owner—or to take the risky and relatively traceable step of tampering with DNS records.

Unfortunately, many domains' internal accountability policies fall well short of the assumptions made by the Sender ID standard: many ISPs that provide high-speed Internet access, for example, simply pass along all email messages that originate with a computer registered with them. A spammer who captures a home user's computer through a virus or worm, thus, can simply send mass emails as that user without triggering any Sender ID alarms. More sophisticated systems, still being proposed and standardized, would add cryptography along the lines of digital signatures to the mix.

Filtering

Email filtering is conceptually simple and often easy to integrate with existing systems. In general, a filter is a function that splits an input stream into one or more output streams, possibly removing some unwanted inputs entirely. An email filter filters a stream of email messages, trying to distinguish spam messages from messages the recipient wants to see.

Headers

A trivial form of filtering is rejecting emails whose To: field consists of a nonexistent email address. Similarly, certain common Subject: and X-Mailer: headers are associated with known spams and spam programs, and can fairly safely be rejected as spam on sight. More subtly, there are a number of computationally inexpensive checks that a filter can perform on the headers of a received an email: a message whose headers are inconsistent, malformed, or incoherent is likely to be spam. Poorly configured bulk-mail programs, for example, may create emails with an incorrectly formatted Date: header or Content-Type: header. Ultimately, however, spam mailers have been quite successful in imitating the header structure of legitimate email.

Keywords

Commercial spam must ultimately contain either an advertisement or point to a one, or there is no economic use to sending it. This fact, combined with the pure bulk of spam and its recurrent subjects, means that it is possible for filters to look for the telltale traces of such advertisements. Some phrases are classic spam giveaways. "Viagra," "Be your own BOSS!," and "Greetings, I am long lost" are all recognized by most email users on sight as being likely to be associated with spam. Computer programs can easily be created to scan the text of a message for such red flags. For spams which rely on external images, which link to order forms on web

servers, or which contain web bugs, one can also look for URLs that contain the names of known spam servers. Spam marketers, however, have developed effective countermeasures to such forms of filtering. Instead of sending a message text containing known spam keywords, they send a message containing a single GIF or JPG image—a black box to any keyword parser—or they use unrelated words to a questionable website. They change HTTP servers frequently, to avoid blacklisting based on known bad URLs. And keyword obfuscation is also a well-known technique, from pen1s to L*3*V*I*T*R*A.

Fingerprinting

Another approach based on message contents is fingerprinting. One applies a hash function to each known-spam message to generate a database of hash values of messages which indicate a very high likelihood of spamminess. Any new message with a hash value that matches one in the database is considered presumptively spam. More sophisticated variations generate partial hashes based on subsets of the input: they can catch messages which contain a forbidden sequence, even if that sequence is less than the whole message. This approach is moderately effective in stopping computer virii: the binary executable forms of most virii don't change much from one infected host to the next, so anti-virus scanners can often spot known virii from their fingerprints. It has proven less successful in dealing with spam: spammers have learned to use randomization quite extensively, never sending the exact same message twice and varying quite substantially the text of their messages.

Machine Learning

Increasingly, anti-spam technologists have turned to sophisticated algorithms drawn from machine learning and artificial intelligence. Instead of trying to specify an explicit algorithm that decides whether a message is spam or not, they 'train' a computer program to learn the characteristics of spam messages. These techniques are inductive, not deductive: the computer is shown a large number of emails (a "corpus") and told which ones are and aren't spam. It then generates its own hypotheses as to what criteria distinguish spam messages; those hypotheses are tested out against more emails, the computer learns from its mistakes and generates new hypotheses, and so on.

While the class of possible hypotheses must be specified in advance by the programmer—messages containing certain texts, messages with certain kinds of formatting, etc.—there are huge possible numbers of them, and the computer can evolve a rule that combines many hypotheses in a very subtle manner. Neural nets, Bayesian learning, genetic algorithms—there are many different specific techniques in this category, and some of them have been reasonably successful so far.

Pay-to-Send

The third major category of technical solutions to spam consists of pay-to-send systems. These approaches try to force spammers to internalize some of the costs of sending large quantities of email: an email sender must expend some resource on each email. If that resource has a value equal the costs of processing the email, then economic logic predicts that the socially optimal amount of email will be sent.

Such systems, like other technical solutions, can be implemented either client-side or server-side; the differences tend not to affect the details of the implementation significantly in ways distinctive to pay-to-send systems. There are three main candidates for the appropriate resource to use as email currency: money, computer processing time, and human attention.

Pay-to-send is usually proposed in connection with explicit whiteisting: senders on an approved list are exempted from the need to pay. Since most emails are between parties who already know each other, this tweak reduces overall costs in exchange for a slight increase in the risk of forgery.

Attention

Pay-to-send using attention is conceptually simplest. Upon receiving an email, the recipient generates a question which can, in theory, only be answered by a human. Such CAPTCHAs often involve displaying an image that contains text that has been distorted or overlaid with other images or perhaps a challenge to count the number of puppies or kittens. Since human letter- and picture- recognition abilities currently far exceed the corresponding abilities of computers, a person can type in the message while a computer would find the image file impenetrable, even if it was known to be a CAPTCHA.

A technique now apparently being used to defeat CAPTCHAs is to have the spamming computer serve as a man-in-the-middle between the querying computer and a person, for example by requiring users of (often bootlegged) pornography sites to provide the answer to a CAPTCHA in order to see the site's content. Critics have also suggested that spammers may use cheap human labor in the developing world to overcome this burden.

Cycles

Pay-to-send using cycles requires the sender to exhibit proof that it has carried out some computationally difficult task; the scheme works because checking such proofs can be much shorter than generating them. The most common cycle-based pay-to-send proposals are challenge-response: the recipient proposes a computation to which it already knows the answer, which the sender then carries out, and transmits the result back. Choosing an appropriate computation is an exercise in computational complexity and depends on the state of the art in the design of algorithmic cryptographic primitives.

A related, but simpler, system is degradation: the recipient slows its receipt of messages down, or sends temporary unavailability notices. A true spammer will be unwilling to wait and will abandon the connection or choose another target, while, in theory, a genuine sender will be willing to wait. Yet another technique involves opening a simultaneous connection to *send* an email back to the sender. If that connection succeeds, then the sender has indicated its bona fides of running an actual SMTP server that accepts mail as well as generating it, and is, arguably, more trustworthy.

The above techniques are explicitly challenge-response. This is not entirely necessary for pay-to-send cycle schemes: one might also insist upon a proof that can be generated offline but is in some way specific to the message. It might either be a difficult search problem based on the message (e.g., for a hash collision in a significantly large secure-hash keyspace) or some certificate issued by a trusted third party upon proof of sufficient computational energy.

Money

This last technique—a third party certifies that one has expended the necessary computational energy—can easily be adapted to condition the certification on financial, rather than computational, payment. Much of the infrastructure here resembles that for third-party certification for sender identification, but the focus is different. There, it is the ex post real-life relationship that matters; here it is the ex ante fact of payment. The third party must be trusted, but it can hand out certificates to anyone who pays their bills.

This technique also suffers from the same issues that any digital cash system does—for that is what it is. The incentives to hack the third party or skim off the top are extremely high, and the cryptography involved must be carefully designed to prevent the forgery of unlimited email tokens. In theory, a pay-to-send system that uses actual cash values could be entirely distributed, but in practice most serious proposals have pointed to some trusted central payment and authentication infrastructure. There is currently no system that could run a sufficiently massive level of micropayments (currently over a billion emails a day). The largest similar system in practice is Bonded Sender, used by Hotmail, which has a small number of senders who pledge bonds for good mail practices. If complaints exceed a set level, the bond is cashed by a third-party charity.

LEGAL RESPONSES

Basic Concerns

Anti-spam legislation raises a number of conflict-of-laws issues that are common to many fields of cyber-regulation.

Jurisdictional Limits

First, there is the familiar problem of how far a territorial sovereign's laws do or should extend when applied to cyberspace-based conduct. A spam recipient may have no idea whether the email came from nearby or from across the world and there may be no way to find out for sure without an extensive investigation. Businesses advertised may be situated anywhere in the world. That investigation itself may become substantially more difficult as it crosses boundaries and the number of different jurisdictions in which intermediate servers are located. Emailers may argue that it offends notions of due process or international law to subject them to liability for email when they couldn't know in advance what jurisdictions it would enter or cross through.

Spam may be simpler than the general problem of Internet jurisdiction in the sense that there exists a broad international consensus that spam is unethical conduct that should be treated as illegal. Whether or not different states' authorities are willing to cooperate in investigations or in coordinating technical measures, they are not likely to have major disagreements on substantive question. Spam is a more unitary phenomenon than hate speech and gambling, in which differing international standards have led to significant friction.

Spam could also be regulated in a way similar to Lessig's Internet zoning by using domains as proxies for various information—e.g. a kids.yahoo.au email addresscould indicate a minor in Australia. Currently, domain groupings have little-to-no value (except perhaps .edu), but we can imagine systems which make them a jurisdictional hook. In the United States, the jurisdictional hook is long-established: wherever a spam lands has jurisdiction.

Offshoring of even domestic spam operations, like the offshoring of other forms of IT-intensive business, seems to be an accelerating trend. U.S.-based spammers, for example, use servers both in the U.S. and abroad to send out their email. Jurisdictions which do not prove responsive in combating spam become increasing spam havens, which gives email from those jurisdictions a bad reputation (in a pattern that mirrors that for ISPs which do not combat spam). Spam will not be seriously solved without some measure of international coordination.

Conflicting Definitions

Another familiar problem from cyberjurisdiction is outright conflict of laws. Even without disagreement on whether spam is good or bad, having anti-spam regulations from many jurisdictions creates difficult coordination problems.

Technical mandates, in particular, raise a serious danger of incompatibility. For example, if half a dozen countries have different labeling requirements, then the required labels may overwhelm the Subject: header itself, ADV ADVERTISEMENT COMMERCIAL. Similarly, pay-to-send systems of different types do not play well together, because a key feature of any such system is that the sender know what sort of payment the recipient is expecting. Too many technical mandates could create an anti-commons, in which anyone working with email must realistically implement many different protocols and markings, fragmenting the end-to-end cleanliness of email. (Note that this anti-commons could also come about from purely private action as walled gardens and burbclaves fence themselves apart.)

But there are also problems with inconsistent behavioral commands. One jurisdiction may wish to offer a safe harbor for conduct that falls within another's prohibited set; both jurisdictions might be justified in these decisions based on the empirical economics of the spam they see. A proliferation of spam laws creates an ugly mess for marketers and mailers, and less confidence in the long-term sustainability of any legitimate strategy.

Deference or Preemption?

A special form of these two sets of conflicts is posed when the jurisdictional entities involved have a hierarchical structure. The United States faces these issues under the name of "federalism," but the same issues arise in the E.U., within various trading treaty organizations, and at the subnational level in other large countries. In each case, the larger entity must decide how its own anti-spam policy will interact with the policies of its subunits. This decision poses difficult questions of institutional competence.

One possibility is complete preemption: spam is declared a collective problem, not a local one, and local regulation is displaced. This approach favors simplicity and creates more predictability and uniformity. But it also requires the attention of political coalitions with much else to think about, and may result in suboptimal tailoring to the conditions of local spam problems. In particular, it may not be a good idea to enforce full preemption across jurisdictions with substantially different telecommunications regulations or with major language differences.

Another is complete deference: the federated entity adopts no anti-spam policy and lets the local entities regulate freely. This opposite approach maximizes flexibility, but at the cost of creating uncertainty and potential conflicts in local laws. In between are a whole range of intermediate approaches. One could displace some local laws (e.g. technological mandates), but leave others (e.g. criminal penalties) in place. One could supply resources to local entities wishing to do investigatory work to find spammers and create institutions to enable investigation across jurisdictional lines. One could add federal penalties and regulations on top of those generated by local jurisdictions.

These questions, of course, are of special concern to the United States, which has seen first an era of regulation of spam by states, followed by federal regulation that substantially, but not completely, preempts those laws.

Taxonomy

Punishment of Spammers

Punishments of spammers are the most fundamental legal tool available. These punishments could be either civil fines or criminal penalties, including fines, imprisonment, or various alternative punishments. The law is used to administering such penalties and has the administrative and institutional apparatus already available to prosecute spammers. The two key questions are how much in resources to devote to prosecuting them and what level of punishments are appropriate.

Spam is interesting as a misdeed. It is an offense against property, not against the person, and as such, seems obviously and appropriately lower-priority for prosecutors and police than crimes such as murder, assault, and rape. Because of its intangibility, it fits into a white-collar mold, like stock fraud and copyright infringement, rather than burglary or auto theft.

The sheer lack of knowledge about the structure of the spam industry is frustrating. It is hard to tell how much good each individual arrest does, because it is hard to tell how much spam is generated by each participant in the industry. Without this sense, it is also hard to know how

seriously to go after individual spammers and how much to punish them. If one person were causing the entire world's spam problem, for example, a long prison term would probably be justified. But such terms seem seriously disproportionate against lower-level operators who are largely interchangeable from the spam operation's perspective.

Clouding matters is the extreme frustration many people feel over spam. Spam is a subject that causes a red mist to descend over the eyes of many otherwise quite forgiving and coolly rational people. It might not be an exaggeration to say that spam is the technologist's morals crime: something that causes outrage because it perverts central tenets of some people's lives.

Finally, in any discussion of criminal penalties for spam, it is worth remembering that many spam practices would be illegal even without *sui generis* anti-spam statutes. Traditional fraud and intellectual-property statutes, in particular, have teeth against many classes of spam. Many internet drug ads are false advertisements for imitation goods, and many more commit trademark and patent infringement. Nigerian 419 emails are paradigmatic frauds. Discount software ads are typically secondary copyright infringement. And so on and so forth: the close connections between spam and scam mean that the former can often be punished as the latter.

Currently, spam qualifies for civil punishment by ISP lawsuits or attorney-general lawsuits and criminal prosecution in extreme cases, usually involving fraud, use of a "protected system," or hacking into a computer. The latter hook has been used, for example in the arrest of a spammer accused of "wardriving"—accessing unprotected wireless networks from his car to send huge quantities of spam.

Tracking Down Spammers

Another avenue along which the legal system could invest more resources is the police work involved in locating and arresting spammers. The decentralized structure of the Internet and the frequent use of zombie computers make spam, like much online crime, difficult to investigate. There are few telltale identifiers like fingerprints or DNA fragments available, and the effort is often dependent on the recordkeeping of systems which keep terrible records.

Computer crime has been something of a backwater for law enforcement: typically only child pornography, truly massive hacking, or major virus releases have sufficed to interest investigators. There are some signs that spam is starting to interest them more; the number of spam-related arrests has picked up sharply in recent years, though it remains quite small in absolute terms. Again, increased investment here may only be justified in tracking down relatively big fish in the spam world.

That said, there are many known spammers who continue to laugh at state judgments and state laws, and the New York Times has even reported on spammers openly flouting CAN-SPAM. Some spammers will be hard or even impossible to find—but then again, some, like the one who started a "Spam King" clothing line until Hormel caught up with him, will be easy.

Hit the Beneficiaries

The complex relationship between legitimate and illegitimate business is an interesting aspect of modern society. Spam is no exception: many companies that do not themselves spam profit from spam. X10, makers of a notorious line of wireless cameras, were also responsible for the largest popup marketing operations of all time. X10 was a real company with a real address and a real product, at least until all of its questionable practices caught up with it at once. Had X10 been shut down earlier, the popups it was buying would have disappeared, too.

While a large part of the spam sector is entirely illegitimate, as described above, a large part of it is probably dependent on legitimate enterprises, companies that can be located and have enough attachable assets to be held accountable. These companies buy ads that are sent out as spam; they sell spammers services. Carefully-worded statutes might impose vicarious liability on these companies and hold their feet to the fire as a way of suppressing spam. One could, for example, make it a civil offense knowingly to benefit from spammed advertisements for one's products without taking appropriate steps to prevent the spam; the McCain amendment to CAN-SPAM does just this, with some restrictions. Such techniques may not work against spam that is illegitimate from start to finish, but they could dry up a key source of revenue for spammers.

Technical Mandates

The law could make any of the above technical approaches mandatory. That is, it could specify one of the above approaches and then punish anyone who failed to comply with that approach in their email use. The idea is to take a reasonable solution and use the law to speed its adoption. A technical mandate is a reasonable idea if the added incentive the law could provide would induce sufficiently many people to adopt the technology that the effectiveness of that solution would be sufficiently improved to make a real difference. A technical mandate could also be part of a more comprehensive legal solution; for example, a sender-identification system, if widely-used, would give legal sanctions against spammers more weight.

On the downside, technical mandates have a serious innovation-damping effect. They divert legal resources from the direct prosecution of spammer; they could challenge the values of anonymous speech paradigmatic to the Internet. They potentially interfere with liberty, economic efficiency, and free speech interests. If they are poorly chosen, the damage could be considerable. And they are especially vulnerable to jurisdictional boundaries: a pay-to-send system has little value if overseas systems can avoid its requirements with impunity.

One additional technical solution that the law could mandate is labeling. Advertising messages might be required to have a Subject: header prefixed with "ADV:," for example, as in South Korea. These labels would be designed to serve as a quick key for immediate filtering.

Lawrence Lessig famously proposed coupling labelling with a bounty system for vigilantes—and promised to resign from his job if such a system were implemented and proved ineffective. As his suggestion suggests, disregard of labeling requirements could be used to open mailers up to increased penalties while perhaps providing a safe harbor for legitimate mailers who respected it. Some have suggested that labeling could be enforced as a matter of industry self-regulation, although that possibility seems to assume a greater degree of legitimate industry control over spammers than may actually be realistic.

Recruit Private-Sector Allies

Many of the above techniques could be implemented with an attempt to convince private-sector parties to carry them out. For example, spam bounties could be paid to those who managed to locate spammers who were then successfully prosecuted—thereby having the private sector do the work of tracking down spammers. The FTC has announced that this kind of bounties, as implemented by CAN-SPAM, will be in the six figures and awarded to insiders with valuable information.

Similarly, the government could fund research or adoption of technical solutions, rather than requiring them directly. One might suggest that the marketing industry might adopt a set of best practices for email solicitation and then self-regulate, but one would not be taken seriously:

No one seriously pretends that any industry association has significant leverage over most sources of spam.

Do-Not-Email

One idea that is regularly floated is a do-not-email registry akin to the federal Do-Not-Call registry that has proven remarkably successful in curbing unwanted telemarketing phone calls. This idea is a notably poor one, at least in unmodified form. First, phone numbers are already largely a matter of public record; consider any phone book. Email addresses, on the other hand, are typically not advertised to the entire world. Similarly, the average email address cannot be guessed by brute force search of a relevant space, while a brute-force robo-call can eventually assemble a list of valid phone numbers anyway.

These two points, taken together, suggest that compiling a list of email addresses not to be spammed is equivalent to compiling a master list of spammable addresses, a list that would be golden to any spammer. And no one has seriously proposed a technical scheme capable of keeping such a list secret that does not require passing all emails through a centralized government server. The end result is to deliver into the hands of illegitimate spammers—ones whose actions are illegal anyway even without special penalties for violating do-not-email—a list of precisely those addresses they should spam. There may be intelligent variations on a do-not-email registry, but a list by itself is not an effective anti-spam measure.

CAN-SPAM

Federal spam legislation was first proposed in 1999; but spam was annoying long before then. Because the feds were late to the scene, states had already passed a huge variety of laws with varying definitions of spam (including extreme variations in their definitions of bulk, commercial, and unsolicited) and contradictory labeling provisions, consent requirements, and defenses.

The impetus for federal action was perhaps a California spam law with million dollar liquidated damages. The bill that would later become CAN-SPAM passed the Senate 97-0. It contains a strong preemption clause, preventing further conflict with state law. It does, however allow state regulation for fraudulent mail, which states are pursuing wholeheartedly.

CAN-SPAM requires a few things from senders. All are intuitive; none are too difficult. All mail must be labeled as commercial (though not in the subject line). Pornographic email ("sexually oriented material") must be labeled "SEXUALLY-EXPLICIT." All mail must contain the business name and physical address. It must have a valid link or address to opt-out for 30 days. Headers and routing information cannot be forged. This last one is important normatively because previously, sending mislabeled and misleading information in email was not a federal crime.

Most importantly, CAN-SPAM imposes an opt-out regime. Mail is legal to send until a user opts out. This last provision is why many advocacy groups have called CAN-SPAM the "YOU-CAN-SPAM Act:" it makes that first message entirely legal. This charge isn't quite true, as CAN-SPAM also prohibits dictionary attacks, zombie networks, and harvesting—all signatures of spammers. The opt-out requirement is also based in First Amendment speech rights. Like the Do-Not-Call list, it presumes speech to be wanted—or at least not unwanted—until a user opts out.

CAN-SPAM also authorizes the FTC to offer bounties in the six-figure range to insiders with valuable information about spammers. It authorizes a study of SMS spam, and authorized

the FTC to investigate the possibility of a Do-Not-Email registry (the FTC investigated the idea, and decided against it). Just because users must opt-out for to make email illegal does not mean that the mail must be delivered. ISPs have the right to filter as they wish and they also have a safe harbor for mail sent via their systems.

CAN-SPAM enforcement is limited to three parties: state attorney generals, ISPs, and federal agencies (the FTC and FBI). This is actually the most innovative part and important of the bill, as few other spam laws give ISPs any standing. Australia and the U.K., for example, give standing only to a horribly underfunded FTC equivalent. ISPs bear most of the costs of spamming, from consumer complaints to physical infrastructure, so they have been the 800 pound gorillas of spam enforcement actions. Considering that they were already filing lawsuits under state laws with little to no chance of collecting, it is no surprise that three months after CAN-SPAM's enactment, several ISPs announced hundreds of lawsuits in a major display of cooperation. Many of these cases have already settled; some for six figures or more, some even overseas. The Massachusetts Attorney general has filed a state complaint, and the FBI is expected to arrest about 50 notorious spammers by the end of the year. About a dozen have already been arrested on criminal charges, including a New Zealand spammer.

The one group left out of the spam lawsuit party is users. The Junk Fax Act gave recipients standing to sue their tormentors directly, and many believed that CAN-SPAM should give users a similar right to sue spammer directly. But that was not to be: CAN-SPAM allows ISPs to speak for their users, if you will.