The Economics of Spam

Eric Allman, Sendmail

You know what I hate about spam filtering? Most of what we do today hurts the people who are already being hurt the most. Think about it: Who pays in the spam game? The recipients. That's what's wrong in the first place—the wrong folks pay for this scourge.

The recipients have to actually store the data, which takes disk space, plus the majority of the I/O bandwidth costs: at least once on the network to receive it, and then at least twice on the disk to keep it (storing and then retrieving the space on disk), not to mention the cost of the disk space itself. You might say that “bandwidth is free” and/or “disk space is free,” and it's true that they are much cheaper than they were when I was a wee lad (we used to say that disks were shipped full from the factory, given that they never seemed to last long—but then I also remember when having an entire gigabyte of space on one multi-rack system was a really big deal). But bandwidth and disk space have never been free, and when you are a large company or ISP, these things can make a difference. In fact, even if you aren't a large company, spam can pretty trivially saturate your external bandwidth.

So what do we do to address this problem? We add spam filtering, usually based on content. And who pays for that? The recipients. You remember, the guy getting !#\$@%& in the first place. What happens to the spammer? Nothing.

Let's drop back to freshman economics. For all the absurdity of what they taught us back then (“assume a rational buyer”—any good marketer knows that this is a lousy assumption), there are still some good points to take away. Perhaps most important, people optimize locally—that is, they do what's best for themselves. Actually, the “irrational” folks sometimes do what's good for the greater goal, assuming that it will do them or their ancestors good; hence, ecologic thought. And thank (insert deity of your choice) for the “irrational” people. We need more of them, especially in the computer business today.

So assuming the irrational folks are a minority (alas), what happens? Well, the seller tries to maximize profits. This is done by maximizing sales and minimizing costs (a very crude approximation, but it will do for now). One of

How did we

END UP WITH RECIPIENTS

PAYING THE PRICE?

the costs is informing folks about your product—after all, they can't buy it if they don't know it exists,

right? You can do this in a number of ways. If you're selling a large-ticket item (a private jet, for example), you use a direct sales force that goes out and gets to know each buyer personally, wines them and dines them, even becomes friends with them—they develop long-term relationships that work for both parties.

But suppose you're not selling private jets. Suppose you're selling kitchen appliances or organ enhancements or economic success or an opportunity to participate in a multimillion-dollar fraud involving overthrown African dictators (30 percent for you, of course, with only the tiniest commitment). OK, for the last one, if legit, I would use a direct (but discreet) sales force. But ignoring that, how do you work then?

Well, how about advertising? An honest profession, billions poured into it, with all kinds of creative people with cool ponytails doing cutting-edge work. Oh wait, that “billions” means that I'll have to pay for it. There must be a better way.

But there is! Direct marketing! Now this can happen in many ways. You can call people up during the dinner hour and inform them of how your product will make them richer/prettier/healthier. You can send them pieces of paper that contain that same message, which they of course are eagerly waiting for. But wait! Both of those cost you money. If you didn't have to spend that money, you could of course offer a better deal to your buyers. (If you were unethical you could take it as profit, of course, but who would do something as nasty as that?)

So back to economics. How does this work in the real world? Well, sellers try to maximize profits by maximizing price and minimizing expense. Advertising is an expense, so keeping this cheap is a good thing. And there's the rub.

Calling people up, even with those annoying predictive dialers, costs a significant amount of money per call. Sending a paper solicitation through the mail costs real

Continued on page 78

Continued from page 80

money. I've seen claims that acquiring a single lead (not a customer) through direct mail costs close to \$10. Sellers limit the amount of mail they send out because every piece costs money, and they want to maximize profit. So they mail only to the people who are most likely to buy their products. That increases response rate, which is a good thing.

But suppose it costs you essentially nothing to send out a mailing. Then your best strategy to maximize profits is to send to any and every address you can find. After all, if you're selling mortgage financing, there might actually be some renters who own property somewhere (there are). You would miss those potential buyers if you trimmed your list. Perhaps some folks who have expressed interest in designer beer mugs are also interested in antique dolls. If you did the "rational" thing you would miss them, and it costs you nothing, right?

The sad point of all of this is that I'm going to (sort of) defend the spammers and point out that they are responding to basic economic forces that we all respond to at one level or another. As long as spammers can take in more money than it costs them, they will continue to spam. This is "rational" behavior in the economic sense.

So, hey Eric, you're getting boring. What's this flame all about? Well, I've tried to convince you, fair reader, that it's all about economics. And basically the economics of e-mail are severely distorted. Physical mail, phone solicitations, etc., have the essential property that "sender pays." But e-mail today uses a "receiver pays" model. By the way, this isn't terribly different from cellphone fee structures in the United States, where if someone calls *me* on *my* cellphone I get to pay the airtime costs whether I wanted to talk with the caller or not. In most places in Europe the caller pays the airtime costs when calling a mobile phone.

The economics are wrong. So our response? We make the economics even worse. We make the recipient pay more. Sometimes much more. Filtering, quarantining, redirecting, stripping, modifying, deleting—all of these costs are borne by the recipient. Also, the

recipient pays the costs of false positives (nospam that is mistakenly classified as spam) in a variety of ways.

When I was first getting into the e-mail business one of my main goals was to eliminate the excuse of "My dog ate my e-mail" (the dog, of course, being the e-mail system itself). For a few years we seemed to have succeeded. But today I hear, all too often, "My spam filter ate your e-mail." Not once a year, or once a month, but close to once a week. Ironically, the less often your spam filter eats your homework, the worse it is. If I've got a lousy filter, I'll check my spambox at least a couple of times a week. But if it works well (e.g., has a false positive only once a month), then I'll probably look at it only occasionally. But I've done this and found mail I really wanted to receive condemned to spam hell. Once they are a month or so old, they really start to stink—which is when I often find them.

When it comes right down to it, heuristics and Bayesian filters and challenge/response systems do improve things from the point of view of the recipient, but not from the point of view of the IT group that has to support all this overhead. Ultimately, e-postage is probably the right way to go, but the costs (implementing the micropayment overhead, plus protocol changes, plus the human frustration) are prohibitive in the short run. Don't look for this in the next couple of years. Besides, people

just hate the idea of paying for their e-mail.

Challenge/response systems actually do increase the cost to the sender, in addition to the recipient, and hence have some value as long as the cost to the sender is high enough to change their behavior. But that's surprisingly hard to do. Ultimately we have to reassign costs from the recipient back to the sender. Such costs can be artificial (e.g., e-postage) or fundamental (e.g., slowing down SMTP connections, perhaps by adding authentication overhead).

Personally, I believe that in this hostile world we may find ourselves limited to "permission-based" mail, where senders can transmit mail to me only if I've already given them permission. But I also hate that idea. If a friend of mine from university has just found me, do

A small number of people
are **polluting a**
great medium.

I really want to reject their mail just because I've never seen the e-mail address before? Some people will say, "Yes, that's just fine." I'm not sure I can for myself.

Fundamentally, the spam problem sucks eggs. A small number of greedy people are polluting a great medium for their own private benefit, regardless of the cost to the rest of the world. They run the risk of damaging their own medium (I think this falls into the class of "pissing into your own well") but they don't care—they are short-term thinkers. The problem is that our approach to the solution has also been short-term thinking. We have to think long-term. We have to make the spammers pay more than we do. Q

LOVE IT, HATE IT? LET US KNOW

feedback@acmqueue.com or www.acmqueue.com/forums

ERIC ALLMAN is the cofounder and chief technology officer of Sendmail, one of the first open source-based

companies. Allman was previously the lead programmer on the Mammoth Project at the University of California at Berkeley. This was his second incarnation at Berkeley, as he was the chief programmer on the INGRES database management project. In addition to his assigned tasks, he got involved with the early Unix effort at Berkeley. His first experiences with Unix were with 4th Edition. Over the years, he wrote a number of utilities that appeared with various releases of BSD, including the `-me` macros, `tset`, `trek`, `syslog`, `vacation`, and, of course, `sendmail`. Allman spent the years between the two Berkeley incarnations at Britton Lee (later Sharebase) doing database user and application interfaces, and at the International Computer Science Institute contributing to the Ring Array Processor project for neural-net-based speech recognition. He also coauthored the "C Advisor" column for *Unix Review* for several years. He was a member of the board of directors of the Usenix Association.

© 2003 ACM 1542-7730/03/1200 \$5.00

acm
queue
tomorrow's computing today

Game Development

- AI in games—Finally a reality?
- Making every CPU-cycle count
- Programming games: Harder than you think?

Also Next Month

- Beyond presence: Better than being there
- The hitchhiker's guide to biomorphic software

Coming in February