

ISPadmin

In this installment of ISPadmin, I interview Paul Graham, the man behind one approach to Bayesian spam filtering. Paul has a new book coming out in May titled *Hackers and Painters*. I interviewed him via email in October and November of 2003.

RH: You are well known throughout the community for your outstanding anti-spam work, but less so for anything previous to that. Please tell us a little about yourself before you invented the “Bayesian” anti-spam concept.

PG: I went to grad school in the '80s intending to study AI. AI turned out to be a lost cause, but Lisp seemed worth salvaging from the wreckage, so I concentrated on that. While I was in grad school I got interested in painting, so afterward I went to art school – first to the Accademia in Florence, and then to RISD [Rhode Island School of Design]. Unfortunately, art school is a joke; mostly what you learn is how to act like an artist.

After a few years of being a starving artist in New York, I decided it would be a good idea to make a lot of money, so I dragged my friends Robert Morris and Trevor Blackwell into starting a startup with me. We were, as far as I know, the first ASP. We sold the company to Yahoo in 1998, and it's now Yahoo Store.

After that I could work on what I wanted. One thing I'd always wanted to do was design a new, better Lisp. And to test this new language, Arc, I wrote a spam filter in it.

I didn't invent the concept of Bayesian spam filtering, by the way. I just invented a variant of it that worked well and was easy to implement.

RH: I suppose that is strictly true, given the Microsoft patents and other work prior to yours. Why did it take your essay “A Plan for Spam” in August 2002 to kick-start the idea into what it is today? Was it simply an issue of publicity?

PG: Earlier Bayesian filters didn't work very well, so they tended to lead to the opposite conclusion, that filtering wasn't a viable option. What was new in “A Plan for Spam” was not the idea of Bayesian filtering, but that Bayesian filtering could be done in a way that worked. And the algorithm was so simple that anyone who was skeptical could try it and see for themselves.

There was one other good statistical filter at the time, Bill Yerazunis's CRM114, but I don't think he'd written anything about it then. His algorithm is very ingenious, and very effective, too, because it looks at multiword phrases.

RH: In your first answer, you mentioned that you wrote a Bayesian implementation to test the Arc language. Why haven't you released any publicly available Bayesian anti-spam Lisp (or related) implementations you have written? Also, are you partial to any particular publicly available Bayesian anti-spam implementation(s)?

PG: I haven't released my filter because I haven't released Arc itself yet. No one would be able to run the code.

As far as I know, the two most effective filters right now are CRM114 and Brian Burton's SpamProbe. Both have filtering rates around 99.9%.

RH: One of the greatest benefits of the Bayesian approach is how it adapts to changing data (i.e., spam). Spammers have adapted their methodology to get around any filtering attempts. Do you think the Bayesian approach will be able to keep ahead of the spammers in this proverbial “arms race”? Are there any other promising anti-spam methods out there which have merit?

by Robert Haskins

Robert D. Haskins is currently employed by Renesys Corporation in Hanover, NH.



rhaskins@usenix.org

It would be a great thing if Congress passed an airtight spam law. But I don't have much hope of that; this is the same Congress that gave us the DMCA.

PG: A Bayesian filter works by comparing incoming email to the spam and legitimate mail you've received in the past, to see which it's more like. So the only way to spoof a Bayesian filter is to make spams sound more like real email. You can't get away with saying things like "ACT NOW!" because people never say that in real email.

The spam of the future will consist of some more or less neutral text, plus a link. The question is, what will the response rate be for such low-key spam? If it's low enough, spam will stop being profitable, and we win. And if not, we just follow the link and run the filter on whatever's waiting there. In this new kind of spam, the sales pitch is pushed one step back, but if the user can get to it, so can a filter.

I think there is also room for other anti-spam methods. You don't have to rely on just one. For example, it would be a great thing if Congress passed an airtight spam law. But I don't have much hope of that; this is the same Congress that gave us the DMCA.

RH: You mentioned that you have done a lot of work with Lisp. While many people are superficially familiar with the Lisp "family" of languages, can you give us a (short) introduction to the language? What makes Lisp unique?

PG: Its origins: Lisp grew out of an effort by John McCarthy in the late 1950s to answer the question, What is the smallest number of operators you need in order to write an interpreter for a language in itself? His answer was seven.

Once you've implemented these seven operators, you can write the rest of the language on top of them. This is why Alan Kay, the inventor of Smalltalk, said, "Lisp isn't a language, it's a building material."

Kay also called Lisp "the greatest single programming language ever designed." A lot of people think that, and the reason they do is probably that Lisp's origins as an exercise in axiomatization forced it to be very elegant.

That's what professors see when they look at Lisp. When undergrads look at Lisp, what they see is that the syntax looks weird. But all those parentheses are there for a reason. Lisp code is made out of Lisp data structures. This was the trick that made McCarthy's Lisp so small, and it is also a lot of the reason Lisp is so powerful in practice. It means you can write programs that write programs. Once you've gotten used to that kind of power, it's hard to give it up.

RH: Let's say I ran a Web site and wanted to build some applications in a language like Lisp. What [open source] programs/environments are available for people to implement Lisp applications in a Web server?

PG: The two main ones are AllegroServe and PLT Scheme. PLT Scheme probably has more people working on it. And Scheme has continuations, which are extremely useful here because they let you transcend the statelessness of HTTP sessions. You can make a Web page that behaves like a subroutine call.

RH: How do you currently earn your livelihood? And do you have any sponsors for the Arc development work you are doing?

PG: A couple friends and I sold a startup to Yahoo in 1998. So there are no sponsors except me. I think that will help make Arc a better language, because it doesn't have to do anything except be good to program in.

RH: Do you have any books in the works? Any releasable software in the works?

PG: Both, actually. I'm in the final negotiations with a publisher about a new book. I'm also now finishing the Arc core, which, since I'm doing the McCarthy thing, will be both a language spec and runnable software.

RH: Some have criticized the 2003 Spam Conference at MIT, saying it should have been called the “2003 Spam Filtering Conference,” since there was limited coverage of anything besides filters (e.g., sender authentication). What is your response to those critics? Is the upcoming 2004 conference going to have more coverage of anti-spam non-filtering topics than the 2003 conference?

PG: My response is the text of the conference Web site at the time:

“Interested in spam filters? Come join us at a conference on spam filtering. While anyone will be welcome, we’re hoping most of all to make this conference an opportunity for hackers working on spam filters to get together and compare notes.”

Hard to claim this is unclear.

The 2004 conference site doesn’t refer to filtering specifically, but I expect that’s still what most of the talks will be about, because filtering is the most active area of research at the moment.

RH: You mentioned that a tough, enforceable law against spam would go a long way toward solving the issue. Do you think legislation will result in less spam, or will the (bad) spammers keep doing this regardless of the “social” pressures (like laws) which are applied? What are your thoughts on the topic of legislation as a means to reduce spam?

PG: I think a tough, enforceable (and enforced) law against spam would help. The question is, will Congress give us one? The laws they’re currently considering have been watered down by lobbyists. Perhaps it will require two steps: Congress passes a watered-down law, we find (surprise!) that it doesn’t work, and then, under pressure, they pass a fairly tough law that does work.

Even a wimpy, sporadically enforced law might help. People really start to look askance at a practice that has criminal penalties. In a famous *Wall Street Journal* article, one spammer said:

“You can call me spam queen, I don’t really care. As long as I’m not breaking any laws, you don’t have to love me or like what I do for a living.”

Many spammers say roughly this, if not to reporters, then to their friends and families. If spamming were a federal crime, they would have to be willing to become criminals to keep doing it. I think many wouldn’t.

RH: Recently, I was amused to find an online article about how Sanford (a.k.a. Spamford) Wallace is running a nightclub in New Hampshire (see http://www4.fosters.com/News2003/October2003/October_19/News/su_1019b.asp). Have you personally met any spammers? Have you ever had any sort of conversation with one?

PG: I’ve never met any spammers that I know of, though I believe a few came to the spam conference last year. I think the only times I’ve talked to spammers have been when I couldn’t tell whether some email was a spam or not. I need to know because I’m trying to keep track of filtering rates, so when I’m not sure, I try to ask the sender. Usually it is a spam, but occasionally it might be a friend of a friend whose name I didn’t remember.

RH: Is there anything you’d like to add before we wrap up the interview?

PG: Some URLs: Anyone who wants to learn more about the spam conference, which is in January, can learn more at <http://spamconference.org>, and there are (a few) more details about Arc at <http://paulgraham.com>.