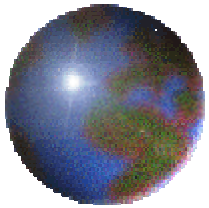
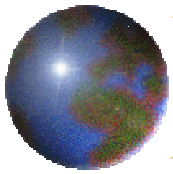


The Anti-Spam Landscape

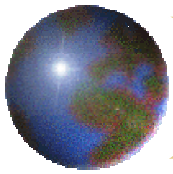
Hal Pomeranz
Deer Run Associates



In the Beginning…

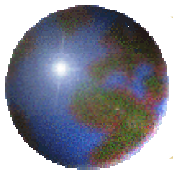
- Sendmail 8.9 (May 19, 1998) declares no more open relays by default
- Also validated sender domain
- Would not accept unqualified senders

Broke a lot of existing email configurations at the time…



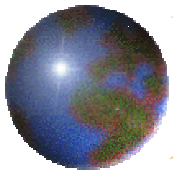
Also the Access Database

- Manual blacklist of sender addresses, domains, IP addresses
- Can also be used to "white list" sender addrs that would otherwise be rejected
- Recent releases allow overrides for specific recipient addresses as well
- Check out spamlist.org for an *extremely* aggressive access DB (dangerous!)



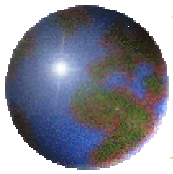
Automated Black Lists

- `FEATURE(dnsbl)` allows you to subscribe to real-time DNS-based lists
- Most effective lists lately:
  - spamcop.net
  - spamhaus.org
- Historically, dnsbls have tended to move around due to "enemy action"



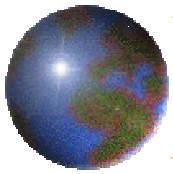
Other Early Heuristics

- Reject if more than "N" recipients
  - *Problem for mailing lists and aliases*
- Reject if recipient address is not in "To:" or "Cc:" line
  - *Breaks "Bcc:" functionality, mailing lists*
- String matching on "Subject:" line
  - *Easily thwarted by spammers*



Early Methods Were Effective…

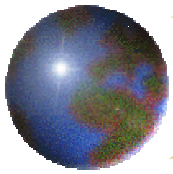
- Relay spam tailed off as open relays were eliminated or black-listed
- "Spam farms" were charted and black-listed, as were spam-friendly countries
- Spammers now resorting to viruses and worms to set up "spam-bot" armies



New Weaponry

- More Advanced Filtering
- Content Signature Databases
- SPF (Sender Policy Framework)
- Greylisting

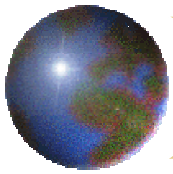
In the Sendmail universe, most of these were made possible via the "milter" interface (Sendmail v8.11 and above)



More Advanced Filtering

- Drop HTML and other "bad" content
  - Kind of draconian, don't you think?
- Server-side regular expression filtering
  - See milter-regexp
- Bayesian (statistical) filtering
  - See Spamassassin and Bogofilter

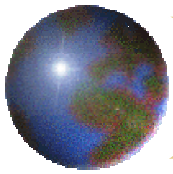
Last two require regular "training", but have proven quite effective…



Content Signature Databases

- Basically, maintain spam traps to create databases of known spam
- Discard any future appearance of similar messages
- Both free (Vipul's Razor) and commercial (e.g. Brightmail) solutions

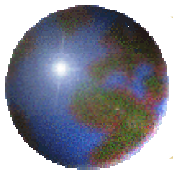
At least it prevents you from getting spammed with the same thing twice…



SPF (Sender Policy Framework)

- Domains publish "reverse MX" records listing their outgoing mail servers
- Can now validate email from sender domain originates from "correct" server

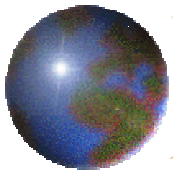
Solves the "forged domain" email problem, but at some cost…



SPF Problems

- Breaks forwarding
  - Have to use "remailing" instead
- Hard on mobile users
  - Requires SMTP Auth, MSAs, etc.
- Encourages use of "throwaway domains"
  - Have to wait and see on this one…

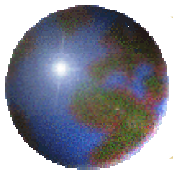
Passions running very high on the SPF issue at the moment…



Greylisting

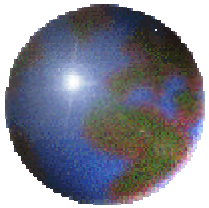
- Greylist automatically tracks "triplets" of *(sender IP, sender addr, recipient addr)*
- Messages from new "triplets" are deferred with temporary failure codes
- Message finally accepted after 1 hour and "triplet" added to auto whitelist

Totally eradicates spam from today's most prevalent spam sources, but…

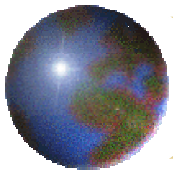


Greylisting Problems

- Delayed delivery on first message
  - Get over it…
- "Farms" of outgoing mail servers
  - Simple heuristics cover most cases
- Broken outgoing mail and list servers
  - White list the ones you have to talk to
- Spammers are going to get smarter
  - Yep, it's an arms race…

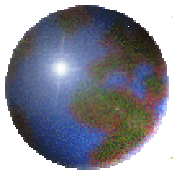


The View from Deer Run



Hal's Prejudices

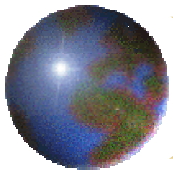
- I "pay by the byte" for incoming email– want to reject spam w/o accepting msg
  - "Filtering" options less attractive to me
- Don't want to spend time "training"
  - Looking for fully automated solutions
- Don't want to waste processing resources dealing with spam
  - Buy new servers just to filter more spam?



Hal's Solution for Deer Run

- Subscribe to every dnsbl I can find
- Use spamlist.org blacklist
  - Have some home-brew automation scripts
- Greylisting with milter-greylist
  - No external package requirements
  - May not be appropriate for large sites

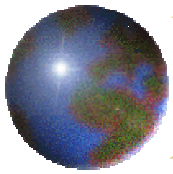
Have to maintain a substantial "white list" because of aggressive filters…



Numbers (as of Mon, 6/28/04)

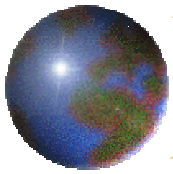
- Messages relayed in past week: 649
- Messages blocked in past week: 1715
  - spamlist.org: 891
  - Manual blacklist: 24
  - spamcop.net: 476
  - spamhaus.org: 127
  - Other dnsbls: 50
  - Invalid domains: 137
  - Relay attempts: 10

Also 180 currently greylisted "triplets"...



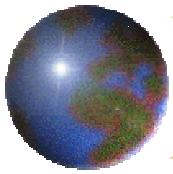
The Human Story

- We simply don't get spam anymore…
  - … at least not *@deer-run.com*
- Still have concerns about the future:
  - Expect greylisting to become less effective
  - Dislike increasing balkanization of the Internet due to black lists, et al



On-Line References (1)

- Good milter info
  http://milter.free.fr/intro/all.htm
- Spamassassin
  http://spamassassin.org/
- Bogofilter
  http://bogofilter.sourceforge.net/
- Greylisting
  http://projects.puremagic.com/greylisting/



On-Line References (2)

- SPF (primary site)

  http://spf.pobox.com/

- SPF (arguments against)

  http://spf.pobox.com/objections.html

  Also see smtp-spf-is-harmful.html under

  http://homepages.tesco.net./~J.deBoynePollard/FGA/