

A Reputation Clearinghouse

mengwong@pobox.com
20040729

Reputation information is born when a human decides if a message is spam and then communicates that information to somebody else, perhaps through a “this is junk mail” button in an MUA. Mail systems that absorb thousands of these clicks can do interesting things with that data. AOL’s SCOMP program offers feedback to sender domains indicating subscriber dissatisfaction with received mail. Cloudmark’s SpamNet builds a theory of Who On The Internet Is A Spammer. Reputation information is useful to many people.

Accreditation: a third party vouches for a sender. Often, this is because the sender has paid them to, and they have conducted due diligence on the sender. Accreditation operates on behalf of senders.
Reputation: a receiver community shares an opinion about a sender, based on that sender’s past actions. Reputation operates on behalf of receivers.

Good domains send a lot of mail and get very few spam complaints.
Bad domains send a lot of mail and get very many spam complaints.
Domains in between may send a lot or a little mail, and get many or few complaints.

Reputation systems help good mail senders get their mail through.

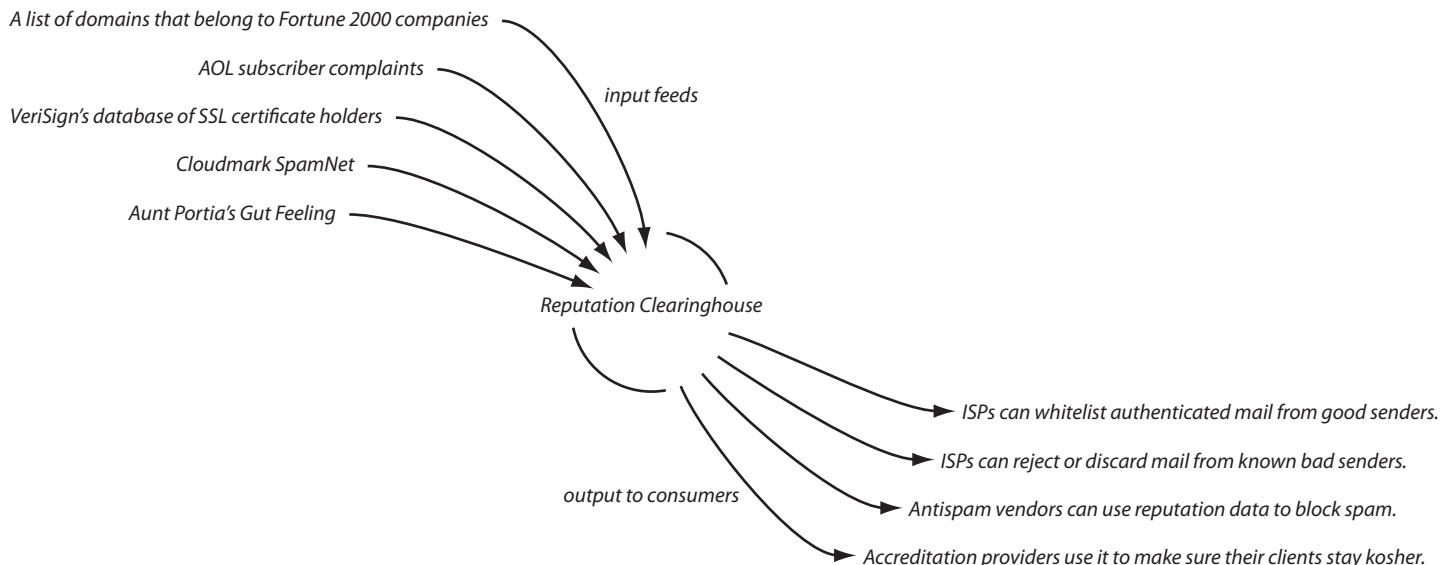
Reputation systems also tell mail senders whether they would benefit from obtaining accreditation. Domains that are new to the Net, or who have not appeared in a reputation system for whatever reason, are good candidates for accreditation services.

Reputation systems help accreditation services keep an eye on their clients. If a client goes bad, the accreditor needs to react quickly or their own reputation might suffer.

Reputation systems help mail receivers do three things:

- 1) whitelist authenticated mail from known good domains to avoid false positives
- 2) block authenticated mail from known bad domains to avoid spam
- 3) greylist or subject to further scrutiny mail that’s somewhere in between.

Reputation data can come from many sources. Some sources may be more dynamic than others. Some sources will be more comprehensive than others. Some sources will be more reliable and accurate than others. Some sources might only say who’s bad. Some sources might only say who’s good.



Reputation data has many consumers.

It is inefficient for each consumer to approach each provider and attempt to obtain a feed separately. A reputation clearinghouse solves that problem, reducing N^2 connections to $2N$. A reputation clearinghouse should follow the Visa model and be jointly owned by participating consumers and producers. It can optimize different views of its input feeds for different consumers. The inputs can be anonymized to protect privacy concerns. Money should flow from the consumers of reputation to producers of reputation.