1 Adapting Global Email for Controlling Spam

D. Crocker
dcrocker@bbiw.net
Brandenburg InternetWorking

for:
IPSJ Magazine, Vol.46, No.7, pp.741-746,July 2005
Special issue on Anti-Spam

It is said that the Internet teaches us one lesson. That lesson is "scaling." The Internet comprises perhaps one billion users, millions of machines and many tens or hundreds of thousands of independent service operators. It operates in, and between, virtually every country on the planet. It is used for personal, organizational and governmental services. Therefore, it must be compatible with many different cultures, many different styles of communication and many different methods of administration. The Internet has no central point of control and operates according to no set schedule. Hence, changes must be gradual and voluntary… once we agree on what those changes should be.

In the early 1990s, the Internet grew from a small research community into a global mass market. This is like having a small town change into a large city. In a large city, most people are strangers, and the strangers have a diverse range of values and behaviors. Hence, people must use much more caution with each other. In other words, the problems are not with the original way the town operated, but with changing requirements.

Everyone agrees that spam is a problem, but we have difficulty agreeing on its definition, although "unsolicited bulk email" (UBE) is probably the most useful [UBE]. Most spam conforms to Internet technical standards. What it violates are our social conventions. Therefore, our technical responses to it must follow, rather than lead, social decisions about it.

Like other social problems, we probably can control spam, even if we cannot eliminate it. Spam has become a permanent part of our social landscape. Some techniques have shown useful local results, but only for a short time. This means that we must be cautious about our expectations for any new proposal. It also is likely that controlling spam requires an array of complementary techniques and continued effort to adapt them, as spammers continue to adapt their own methods. Making changes to a global infrastructure takes a long time and is very expensive. Some proposals require complex technology, while others require substantial, on-going administrative effort. Therefore, we need to ensure that the mechanisms we deploy will have significant, long-term benefit. They also must have reasonable development cost, require limited, ongoing administrative effort, and are sufficiently easy to use. A simple heuristic is to look for mechanisms that would be desired even if spam were not a problem. Such mechanisms provide core, strategic benefit.

The Internet provides us all with vastly better access to each other. For collaboration, or the formation of specialized communities or for personal interaction, this is wonderful. For intrusions into our privacy and threats to our online security, this is problematic. Unfortunately, the benefits and the detriments are tightly coupled. So, our efforts to control the problems need to be made cautiously, lest they reduce the benefits. Worse, our efforts need to limit the damage they might do to innovative future benefits that we do not yet envision.

The sender of spam incurs almost no incremental cost for a single message. It is easy to think that we should simply make email be the same as sending letters or making phone calls, by directly charging the sender for every message. This cost provides a barrier against abusive, bulk use. In reality, email is a different kind of service, with an extensive history, and it is subject to different choices. Telephones and postal service have highly centralized, formal operational authorities, and the fees charged for their use offset direct, real expenses. By contrast, email is a highly decentralized service, with correspondents' private systems contacting each other directly, rather than having to be mediated by a state-regulated utility. If additional fees are charged, they also need to be for real services; an arbitrary "tax" will simply create its own problems.

Email Architecture

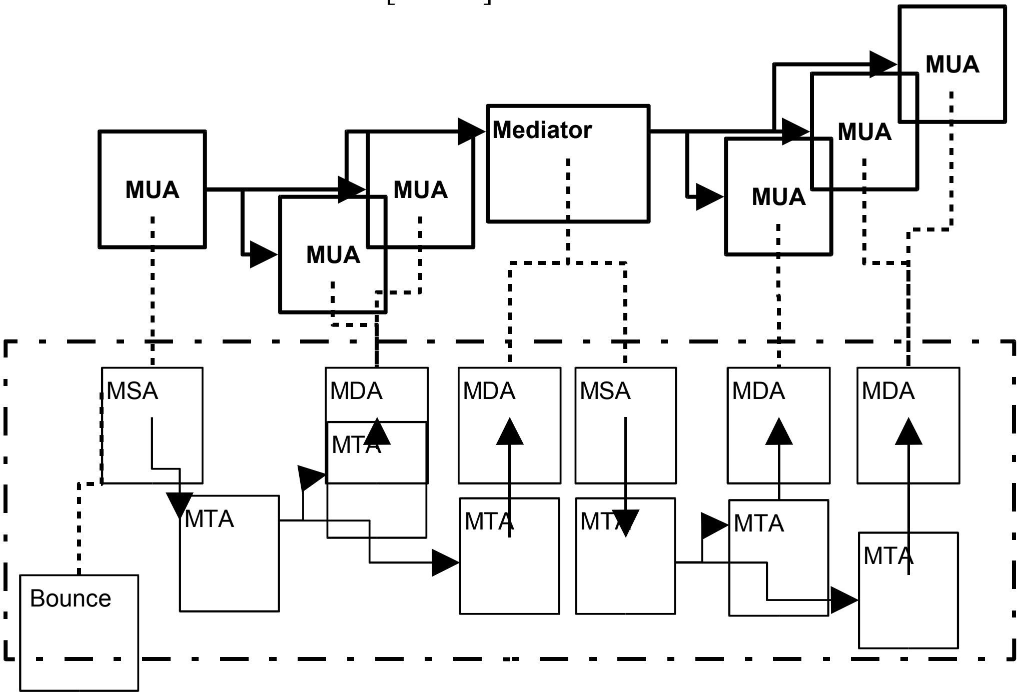
The core Internet Mail service follows a simple model. It distinguishes the world of users from the world of the transmission. Users are represented by Mail User Agents (MUA). Transmission is the Mail Handling Service (MHS) comprising a sequence of Mail Transfer Agents (MTA). The MHS transfers mail, from an originating user to one or more recipient users, using the Simple Message Transfer Protocol (SMTP) [RFC0821, RFC2821]. Anyone may send a message to anyone else. The basic service does not require authentication by the sender or the operators. (It is worth noting that this is the same as for telephone calls and postal mail.) For users, their email client – the MUA – is all they directly experience. For most network administrators, the MHS software is their scope of concern.

The core email message object is also simple. It begins with structured, textual meta-information, called the header, and is followed by lines of free-form ASCII text, called the body. The header includes such things as addressing, posting date, unique message identification and a brief description of the content [RFC0822, RFC2822].

As shown in the Figure, the current architecture is significantly more elaborate. It further separates functions, making more precise assignment of responsibilities, and it permits transfer of complex, multi-media content in any language. [ARCH]

These architectural and standards enhancements distinguish:

- Posting new mail via a Message Submission Agent (MSA) [SUBMIT]
- Delivering it via a Message Delivery Agent (MDA), possibly with user-specific delivery behaviors [POP, IMAP]
- User-level re-posting, such as for a mailing list (Mediator) [LIST]
- Separate designation of an agent to handle transmission reports such as a notice about failure (Bounce) [DSN]
- Hierarchically structured content attachments (MIME), also permitting content in alternate character sets [MIME]



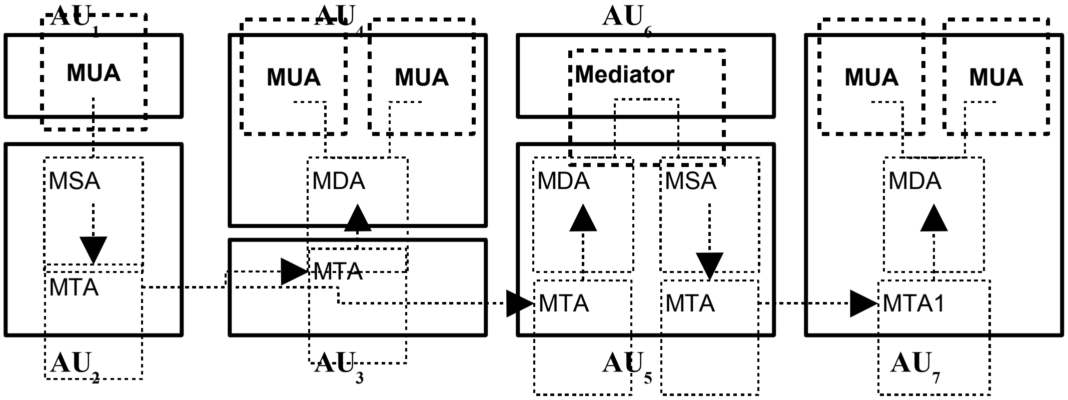
Controlling spam raises the question of accountability. Who is a responsible agent? Unfortunately, many identities are involved in creation or transmission:

Type	**Provided by**	**Meaning**
MTA IP Address	IP network	SMTP client
EHLO Domain Name	SMTP command	SMTP client
Provider IP Address	IP network	Site of SMTP client
Mail-From Mail Address	SMTP command	Bounces address
From Mail Address	RFC2822 message header	Author
Sender Mail Address	RFC2822 message header	Posting agent
Received Domain Name	RFC2822 message header	Relaying MTA sites

The SMTP Client is an agent of the previous hop's operator. Since the email operator might be different from the operator of the IP access network hosting that service, it might entail a different identity. This highlights an interesting aspect of the above table: Most of the above entries can be called "the sender". Consequently, the term has become nearly meaningless, in anti-spam discussion.

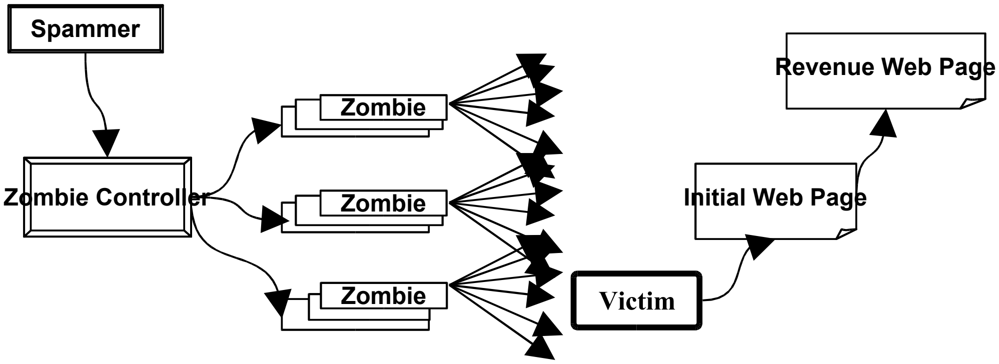
Internet mail requires a clear sense of the operational boundaries between independent operational authorities. These Administrative Units (AU) mark trust boundaries. They distinguish technical and operational variations, as discussed in [TUSSLE] as shown in the Figure.

Increasingly diverse participation in the Internet means that there is more behavior we call "abuse." This highlights the need to make changes in the nature of the trust between AUs and the way that that trust is enforced.



Spamming architecture

Unfortunately, the world of spamming has also developed in scale and sophistication.



Spamming used to entail one sender and one sending machine. Its performance was limited by the capacity of that machine and the bandwidth of its Internet connection. Today, spammers control vast armies of compromised systems, called "zombies," as shown in the following Figure:

The community of spammers is remarkably well organized; it has become an extensive, underground economy. Some participants specialize in developing methods for breaking through filters. Others take over machines, called zombies. Others sell the use of them for periods of spamming. The estimated number of zombie systems is in the many tens of millions. After spam delivery, recipients often "click" to a transaction Web page. Web hosting is provided at multiple levels, in order to obscure the server side of the process and reduce accountability.

Some spammers are legitimate businesses, engaged in aggressive efforts because there are no formal limits. Legal strictures, both laws and contracts, will rein them in. In contrast, some spammers seek to avoid accountability, to subvert barriers to their traffic, and to acquire unwitting and unwilling participation of machines owned by others. The best social model to use for analyzing this latter group is crime. Often, the activities do not violate particular laws, but what is most important is that the style of a spammer's conduct is the same as that of a criminal.

Typically, spammers have the classic goal of selling products. However, they also can have political or religious motivations or even blatantly criminal intent, such as extortion. The ability to send very large number of messages allows spammers to threaten to flood a target network with traffic, in a denial of service attack.

Efforts to control spammers and spamming have, so far, failed to reduce the amount of global spam. At best, control mechanisms have had localized, short-term benefit.

Technical Control Choices

It is tempting to believe that spam is an easy problem to solve, but history teaches us to be cautious. A web page, at <http://craphound.com/spamsolutions.txt>, takes an irreverent approach in challenging simplistic proposals, by providing a checklist for the common weaknesses. It is a surprisingly useful way to screen proposals quickly.

The large installed base of email users and operators makes major or rapid change unlikely. In addition, every technical characteristic of spam is also a characteristic of some legitimate mail. This means that tools for evaluating the message content, or the aggregate traffic flow of messages, might have some transient utility, but they cannot be effective, long-term tools. Any attempt to perform real-time content or traffic assessment has two problems. The first is one of "false positives" in which legitimate mail is incorrectly labeled spam. The second is in the nature of an "arms race" between spammers and anti-spammers who must each constantly adapt techniques.

The common point of spam control is the "filter", named for its conditionally permitting mail to flow through it. The most common filters are at points of reception, such as the boundary MTA for an AU or at the MDA [SIEVE]. However they may be placed anywhere along the path, notably including the MSA, as well as MTAs of the outbound AU. Filters at the reception side cannot reduce Internet spam traffic. At the outbound side, they can. Filters have choices in the way they treat suspect messages. They can:

- Add a special annotation to the message
- Divert it into special storage
- Reject it back to its Bounce address or the Client SMTP during the session
- Simply erase it
- Accept it slowly, with “traffic shaping”, to control the rate of SMTP transmission

The difficult question is: What are the criteria that a filter should use? The difficult answer is: many. Hence, a filtering engine is really a general, extensible platform for spam control. There are two classes of criteria, although each is complex:

- Content analysis, such as Bayesian statistics tracking of vocabulary
- Source identity assessment, either for permission (whitelist) or rejection (blacklist)

Content analysis is always a matter of partial success (and partial failure.) It is usually statistical and depends upon a database of training messages, to establish norms. Spammers are constantly developing techniques for bypassing the current filtering technology.

Identity assessment seeks to hold an entity accountable for problematic email. There are two broad classes of accountable entities:

- Content agents. Authors (From) and those posting mail (Sender) who are responsible for individual messages. If the content agent is validated for a message, then the content probably reflects their intent. That is, it is unlikely that some other entity changed the content. Because the Bounce address (MailFrom) appears in the SMTP protocol but is associated with the posting agent, it is often considered useful for analysis. Unfortunately the address often has no obvious relationship to the From field author or the Sender field posting agent. However spammers often specify false bounce addresses, in order to direct the mass of failed deliveries elsewhere. Consequently, it can be useful to validate the MailFrom address.
- Operations agents. Those operating MTAs or underlying networks, are often held accountable for bulk traffic. Although they do not create the content, it is possible for them to enforce strict rules on their customers and to detect patterns of violations among them. Recommended practices for operators are beginning to obtain some consensus [SPAMOPS].

Assessment of identities can be proactive or reactive, essentially acting as an agent of the sender or an agent of the receiver:

- "Accreditation" is registration by the sender; for these, senders align with a registry that extracts quality assurance commitments
- "Reputation" refers to assessments made about the sender's prior postings; for these, independent third parties evaluate the sender's history.

Because identity listings are made explicitly in a database, they are capable of producing almost no false positives, although there might be many identities not listed. Still, there are significant challenges with the use of identity-based filtering:

- Which identity should be used and how does it relate to potential spamming? Note that the table, above, lists quite a few choices. In addition, an author can create bad content, but the identity listed in the From field of that content might not be the actual author. The operator of the mail-sending network might have nothing to do with creating content, but it might be reasonable to hold them accountable for aggregate traffic problems.
- How is the identity validated (authenticated)? What entity is doing the validation, how does it relate to the identity being validated and why is it trusted? Can the validation mechanism, itself, by tricked?
- How is an identity determined to be a spammer or non-spammer? What entity is vouching for the quality of the sender and why are they trusted?

Most current identity schemes use the IP Address of the client SMTP MTA that is sending directly to the server with the filter. It is provided by the underlying network, and therefore has been difficult to spoof. However, spammers are becoming proficient at stealing IP Address space!

An IP Address changes as the host changes its attachment to the Internet, and it is affiliated with operators, not authors. New schemes seek to use domain names, for more stable references.

Over time, it is likely that Internet mail will evolve into two logical subsets. One will include trusted, accountable participants and the other will include everyone else. Trusted participants might still be subject to less stringent checks and filtering. More importantly, when there is a problem, it is likely that mail received over the trusted channel will still be delivered, while the origination agent is consulted, rather than rejecting the mail automatically.

Support Mechanisms

Fighting spam is a collaborative effort. It can benefit from tools and standards that aid in exchanging information and performing coordination. To this end, standard methods of reporting spamming events, characterizing particular spam, and sending spam control data can be helpful. Given the global nature of spam and spam-fighting the collaboration also needs services to facilitate interactions between network administrators speaking different language. It is also likely that there will be standards for the syntax and semantics of whitelists and blacklists.

References

[ARCH] Crocker, D., “Internet Mail Architecture”, Internet Draft draft-crocker-email-arch, April 2005.

[DNS] Mockapetris, P., “Domain names - concepts and facilities”, STD 13, RFC 1034, November 1987.

[DSN] Moore, K., “Simple Mail Transfer Protocol (SMTP) Service Extension for Delivery Status Notifications (DSNs)”, RFC 3461, January 2003.

[SPAMOPS] Hutzler, C, Crocker, D, Resnick, P, Sanderson, R, and E Allman, “Email Submission Between Independent Networks”, Internet-Draft draft-spamops-00, March 2004.

[IMAP] Crispin, M., “Internet Message Access Protocol - Version 4rev1”, RFC 3501, March 2003.

[MIME] Freed, N. and N.S. Borenstein, “Multipurpose Internet Mail Extensions (MIME) Part One: Format of Internet Message Bodies”, RFC 2045, November 1996.

[POP] Myers, J.G. and M.T. Rose, “Post Office Protocol - Version 3”, STD 53, RFC 1939, May 1996.

[RFC0821] Postel, J.B., “Simple Mail Transfer Protocol”, STD 10, RFC 821, August 1982.

[RFC0822] Crocker, D.H., “Standard for the format of ARPA Internet text messages”, STD 11, RFC 822, August 1982.

[RFC2821] Klensin, J., “Simple Mail Transfer Protocol”, RFC 2821, April 2001.

[RFC2822] Resnick, P., “Internet Message Format”, RFC 2822, April 2001.

[LIST] Chandhok, R. and G. Wenger, “List-Id: A Structured Field and Namespace for the Identification of Mailing Lists”, RFC 2919, March 2001.

[SIEVE] Showalter, T., “Sieve: A Mail Filtering Language”, RFC 3028, January 2001.

[SUBMIT] Gellens, R. and J.C. Klensin, “Message Submission”, RFC 2476, December 1998.

[TUSSLE] Clark, D, Wroclawski, J, Sollins, K, and R Braden, “Tussle in Cyberspace: Defining Tomorrow’s Internet”, ACM SIGCOMM, 2002.

[UBE] Hoffman, P. and D. Crocker, **“Unsolicited Bulk Email: Mechanisms for Control”**, Internet Mail Consortium, UBE-SOL
IMCR-008, http://www.imc.org/ube-sol.html, revised May 4, 1998