Challenges in Anti-Spam Efforts
by Dave Crocker, Brandenburg InternetWorking

Reprinted from The Internet Protocol Journal (IPJ), Volume 8, No. 4, December, 2005
IPJ is a quarterly technical journal published by Cisco Systems. See www.cisco.com/ipj

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—when 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. Imagine a small town changing into a large, undisciplined 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. So, spam is merely an unfortunate—but frankly predictable—example of the Internet's success, not its failure.

This article explores the system-level complexities of the spam problem, as the intersection of social diversity, complexity of e-mail technology and operations, and specific lines of attack that seek to control spam. On the question of control methodologies, most prior work has been on analytic tools that are used by sites receiving spam, to evaluate the mail content, associated addresses or traffic flow. Recent efforts focus on assignment and assessment of an accountable identity that is responsible for individual messages or for the transit of aggregate message traffic.

The Nature of Spam

People agree that spam is a serious problem, but they have difficulty agreeing on its definition. *Unsolicited Bulk E-mail* (UBE) is probably the most useful. [1] A spammer sends a large number of messages to many different recipients who have not requested the content. (Interestingly most spammers do not care whether a particular addressee receives the message; they merely seek to get a sufficient percent of their postings delivered to some of the addressees.)

Spam can conform to Internet technical standards and can contain no technical differences from legitimate—desired—messages. Hence, spam that violates standards or has other peculiarities might be common today, but detection efforts that are based on these anomalies offer no long-term benefits. Spammers are highly adaptable and use the easiest method that works. However what spam *always* violates are our *social* conventions. Therefore, any long-term, proactive, technical responses to it, such as formulation of standards, must follow, rather than lead our

social decisions about it.

Like other social problems, we probably can control spam, even if we cannot eliminate it. This means that we must adjust to having spam as a permanent part of our social landscape, even as we seek to limit it to tolerable levels. Efforts to detect and eliminate spam have been underway for quite a few years. Some techniques have shown useful, localized results, but most only for a short time. In other words, none of the many spam control attempts, over the years, has yet reduced the amount of global spam! So we must be cautious about our expectations for any new anti-spam proposal. It also is likely that controlling spam requires an array of complementary techniques and continued efforts to adapt them, as spammers continue to adapt their own methods. This means that we need to assess any new proposal in terms of its likely *incremental* benefit, rather than as a candidate to be the *Final Ultimate Solution to Solve Spam* (FUSSP).

Changing a global infrastructure takes a long time and is very expensive. Some proposals require complex technology, while others require substantial, on-going administrative effort. Worse, some impose onerous requirements on end-users. Therefore we need to ensure that the mechanisms we deploy will have significant, long-term benefit, even after spammers try to adapt to their presence. They also must have reasonable development cost, require limited, on-going administration and be sufficiently easy to use. In evaluating the likely efficacy of a proposal, a useful heuristic is to ask whether it would be desired even if spam were not a problem. If the answer is yes, then it provides general, strategic benefit, so that counteracting spam merely adds urgency to its adoption.

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. Our efforts to control email's problems need to be made cautiously, lest we also reduce its benefits. Worse, our efforts need to limit the damage that might be done to innovative benefits that we have not yet envisioned.

The sender of spam incurs almost no incremental cost for a single message. It is easy to think that we should simply make e-mail 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 e-mail 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 are based on offsets to direct, real expenses. By contrast, e-mail is a highly decentralized service, with correspondents' private systems contacting each other directly, rather than having to be mediated by state-regulated utilities. If additional fees are charged, they also need to be based on the costs of real services; an arbitrary "tax" will simply create its own problems. For example, who gets the money, and why?

To retain its flexibility and its ability to support new human communication uses, we must retain the current, open model of spontaneous email exchanges. Therefore, over time, it is likely that Internet mail will evolve into two logical subsets. One comprises

trusted, accountable participants and the other includes everyone else. Trusted participants may be subject to less stringent checks and filtering. Perhaps more importantly when there is a problem, it is likely that mail from a trusted identity will still be delivered, while the origination agent is consulted, rather than rejecting the mail automatically.

E-mail Architecture

Internet mail is based on a simple model. It distinguishes the world of users from the world of transmission. Anyone may send a message to anyone else. The basic service does not have a central authority and does not require authentication by the Originator, the Recipient or the operators. (It is worth noting that the telephone and postal services usually do not authenticate those sending letters or making calls.)

As shown in Figure 1, this model has grown to distinguish:

- *Mail User Agents* (MUA), which represent end-users
- The *Mail Transfer Service* (MTS) comprising a sequence of one or more *Mail Transfer Agents* (MTA), using the *Simple Message Transfer Protocol* (SMTP) [2,3]
- Posting new mail via a *Message Submission Agent* (MSA) [7]
- A *Notification Handler* or *Bounce Handler,* is an MUA that processes returned transmission reports such as a notice about failure. The Handler's address is specified by the MSA, during message posting. [11]
- Delivering mail via a *Message Delivery Agent* (MDA), possibly with user-specific delivery behaviors [8, 9]

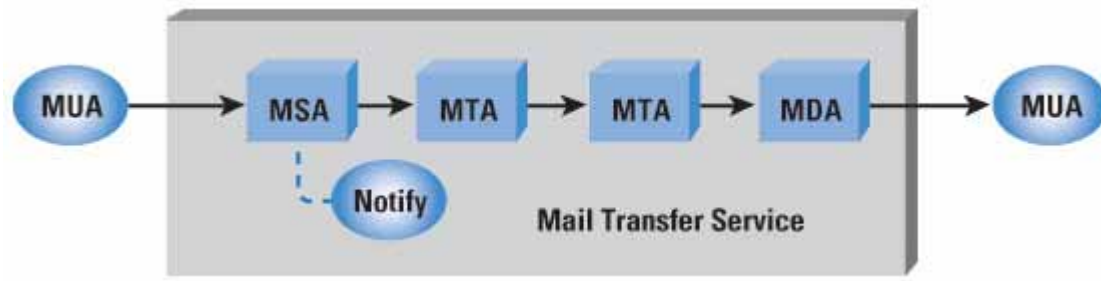

Figure 1: Internet Mail Architecture

The purpose of e-mail is to exchange messages among MUAs. For users, their e-mail client—the MUA—is all they directly experience. For most network administrators, the MTS software is their scope of concern.

The core e-mail message object also has a simple framework. Its *content* comprises:

- Structured, textual meta-information, called the *header*, including *fields* for addressing, posting date, unique message identifier and a free-form description of the content [4,5]
- Lines of free-form ASCII text, called the *body*, which has evolved to support

a potentially complex, structured set of multi-media, multicharacter set attachments [12]

Figure 2 demonstrates a simple user-to-user example, with a message sent to three addressees, one of which is a special MUA that re-mails it to two additional recipients. The purpose of the Figure is to emphasize the user-to-user nature of e-mail and to provide a basis for considering the combinatorial explosion that marks the aggregate interactions of Internet mail components even in very simple uses. It further introduces another architectural construct:

- A *Mediator* is an MUA that re-posts messages, such as for a mailing list. [10] It preserves much or all of the original message, including author address, but can make substantial changes or additions to the content, which an MTA cannot. Therefore, a Mediator's role is user-level content responsibility, rather than MTS-level transit responsibility.

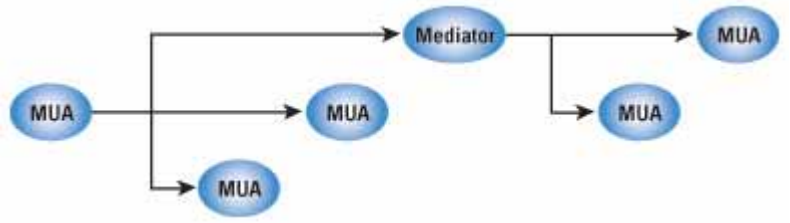

Figure 2: Simple Multi-Recipient Scenario

Spamming Architecture

Some spammers are legitimate businesses, engaged in overly aggressive marketing efforts, because there are no formal limits on their actions. In spite of the challenges created by needing to work at an international level, there is a reasonable expectation that legal strictures, both laws and contracts, will constrain in these businesses to a tolerable level. In contrast, *rogue* spammers actively seek to avoid accountability, to subvert barriers to their traffic, and to acquire unwitting and unwilling participation of machines owned by others. Independent of the legal details, 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.

Unfortunately, the technical and operational 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, rogue spammers control vast armies of compromised systems, called *zombies*, as shown in Figure 3. Zombies are owned by legitimate users who are unaware that their system has been compromised and is being used for spamming.

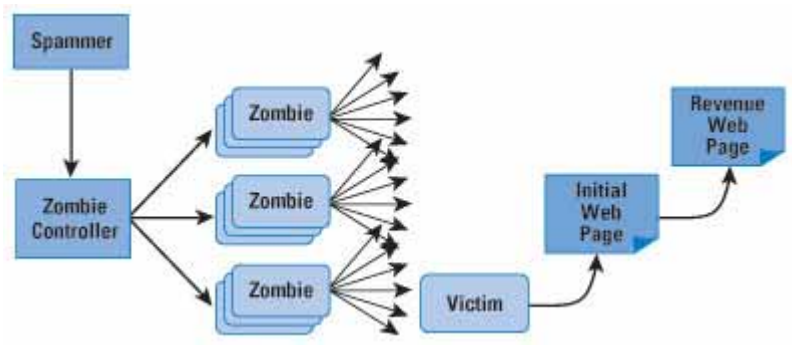

Figure 3: Rogue Spammer Control Network

The community of rogue 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 and turn them into zombies. Others sell the use of a zombie collection 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, further reducing accountability.

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 to a specific destination gives spammers a tool that can be used to threaten an organization with a denial of service attack on their network.

Practical Efforts at Spam Control

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

The most common mechanism for spam control is a localized mechanism, the “filter” [14] , named for its conditionally permitting mail to flow through it. Filters typically are used within the recipient’s network (or Administrative Management Domain, as described later in this article.) However they may be placed anywhere along the path, notably including the MSA. 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 Handling Notification (RFC 2821 MailFrom) address or to the Client SMTP during the transfer session
- Simply delete 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. This need to support a wide, and changing, variety of decision criteria has caused filtering engines to evolve into extensible platforms for spam detection and handling modules. As the mixture and complexity of filtering algorithms become more sophisticated, the overhead they entail has grown substantially larger.

It is convenient to divide techniques into three, basic classes of criteria, although each is complex:

- *Content analysis*, such as Bayesian statistics tracking of vocabulary and content hashing, to detect bulk duplication
- *Responsible Agent assessment*, either for permission (whitelist) or rejection (blacklist)
- *Traffic analysis*, such as rates at which messages come from the same author address or IP Host Address

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 vocabulary norms. Spammers are constantly developing techniques for bypassing the current analysis technologies. Further, different recipients on the same e-mail service can have wildly different statistical patterns of acceptable content. This makes fine-grained filtering by their service provider problematic.

It is clear that these tools for evaluating individual messages, or aggregate traffic flow, can have significant transient utility. However they cannot be effective, long-term tools, even with continuing enhancement. Notably they have little or no effect at reducing spam at its source. These post-hoc analysis tools have two inherent deficiencies, both of which are coupled to their using heuristics, rather than reliable, accurate and objective rules. The first is one of “false positives” in which legitimate mail is incorrectly labeled as spam. As an example, this could mean that an essential business transaction is not delivered, instead being classed as junk mail. Perhaps the most insidious example of this problem occurs when spammers send mail that purports to be from a well-known, legitimate business. This is called *phishing* and results in making *all* mail with the address suspect, so that legitimate postings of essential mail are not delivered.

The second problem with using heuristics is in the nature of an “arms race” between spammers and anti-spammers who must each constantly adapt techniques, consume more resources, and yet never win. It does not help that those fighting spam have been losing the war, since spammers have tended to be more aggressive, more innovative and better organized...

A different line of effort is based on the social assessment that the sender of an e-mail should be held accountable for it. The goal is to identify such an agent and then evaluate the agent’s acceptability. This approach requires three enhancements to Internet mail:

- A clear sense of the boundaries between independent operational authorities
- A means of verifying an accountable identity that is associated with the message
- A means of formulating and sharing assessment information about accountable identities

Although e-mail operators often refer to *boundary* MTAs that face the open Internet, there is no accepted term for a region of e-mail components under unified authority. This article suggests a term derived from the OSI X.400 e-mail effort: *Administrative Management Domain* (ADMD) to mark these trust boundaries. They distinguish a collection of operational components subject to the same administrative policies, as discussed in [13].

An example of ADMDs is shown in Figure 4, and is derived from the scenario shown in Figure 2.

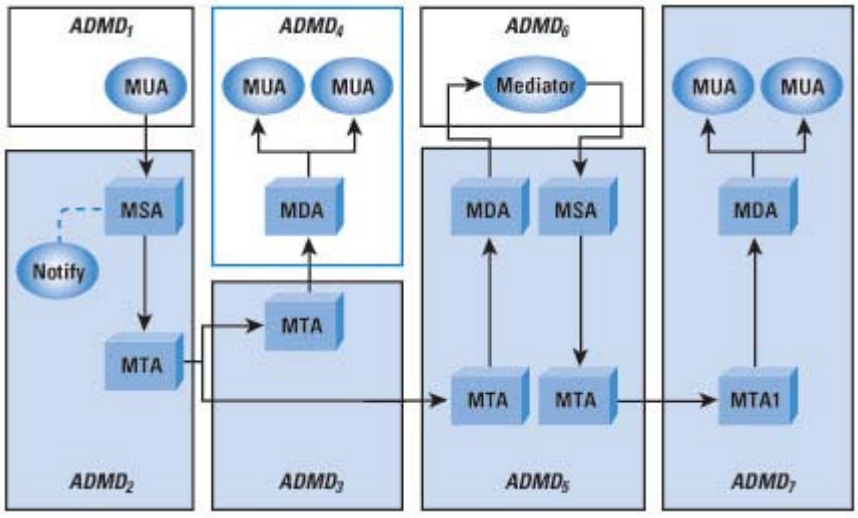

Figure 4: Independent Administrative Management Domains (ADMD)

The implied complexity of responsibilities and interactions is striking, even for this relatively modest case. For simplicity, think of the ADMDs labeled at the top of the Figure as representing users or value-added services, whereas the ADMDs labeled at the bottom could be a variety of classic Internet service (access) providers. The "boundary" agents are the ones with lines connecting over to another ADMD.

The increased diversity among Internet participants and ADMDs results in abuses such as spam. Proactive efforts to deal with these abuses require that we make changes in the nature of the trust between ADMDs and the way that that trust is enforced.

Accountability

Agent assessment seeks to hold an entity (agent) accountable for problematic e-mail. Who is a responsible agent for the content or for injecting the message into the MTS, and are they assessed as trusted or problematic?

There are two broad classes of accountable entities:

- *Content agents* comprise authors (RFC 2822 From) and those who are responsible for posting individual messages, as specified in the RFC 2822 Sender field. 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 Notification Handler address (RFC 2821 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, so its use for filtering can be problematic. However spammers often specify false Handling Notices addresses, in order to direct the mass of failed deliveries elsewhere. Consequently, it can be useful to validate the MailFrom address.
- *Operations agents* provide MTA or basic Internet access services. They are often held accountable for the impact of the bulk traffic their systems generate. 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, such as with [15]. More are needed.

Assessment of agents can be proactive or reactive:

- *Accreditation* is the proactive registration by a sender, who aligns with a registry that extracts quality assurance commitments; any trust of the sender is therefore inherited from trust of the accreditation agency.
- *Reputation* refers to reactive evaluation of a sender's prior postings; for these, independent third parties evaluate the sender's history.

The functions that are combined, to establish useful accountability, comprise:

Identification: An identity label provides a unique reference to an entity.

Authentication: Validates the use of the identity label.

Authorization: Determines that the user associated with the identity is authorized to perform a particular function.

Assessment: Obtains an analysis of the trustworthiness or "quality" of the agency that is providing the authorization, or of the validated entity itself.

Unfortunately, many identities are involved in e-mail creation or transmission, as shown in Table 1.

Table 1: Roles for Internet Mail Identities

Type	Provided by	Identity of
MTA IP Host Address	Network-level service	SMTP client
EHLO Domain Name	RFC 2821 SMTP command	SMTP client

MTA Provider's IP Network Address	Network-level service	Site of SMTP client
Mail-From Mail Address	RFC 2821 SMTP command	Handling notices
From Mail Address	RFC 2822 header field	Author
Sender Mail Address	RFC 2822 header field	Posting agent
Received Domain Name	RFC 2822 header field	Relaying MTA site

Relative to an SMTP Server that is being asked to accept a message, the SMTP Client is an agent of the operator of the previous hop. Since the e-mail operator might be different from the operator of the IP access network that is hosting the e-mail service, it might entail a different identity. This highlights an interesting aspect of Table 1: Most of the identities associated with e-mail handling can be called "the sender." Consequently, that term has become nearly meaningless, in anti-spam discussions.

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 and a listing might be inaccurate. Still, there are significant challenges with the use of identity-based filtering:

- Which identity should be used and how does it relate to spamming behaviors? Note that Table 1 listed quite a few choices. In addition an author can create bad content, but the identity listed in the RFC 2822 From field of that content might not be the actual author, even if that field is validated. The message might have originated on a compromised machine and used the identity associated with it, unbeknown to the owner of the machine. Also the operator of the mail-sending network might have nothing to do with creating content, but it might be reasonable to hold the operator 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, be tricked?
- How is an identity determined to be a spammer or non-spammer? What entity is vouching for the quality of that identity and why is the vouching entity trusted?

Authentication Standards

Accountability requires having an accurate, reliable identity of the agent that is to be accountable. Authenticating an identity is, therefore, a prerequisite for assessment efforts. However it does not, by itself, ensure a positive assessment. Spammers can register and authenticate their identities, too.

Early anti-spam identity schemes use the IP Address of the client SMTP MTA that is sending directly to the server running the filter. The Address is provided by the underlying network service, and therefore has been trusted. However, spammers are

becoming proficient at stealing IP Address space, such as by advertising routes that use allocated-but-unused blocks of IP Addresses! Also an IP Address changes as the host changes its attachment to the Internet, and it is affiliated with operators, not authors. This makes the IP Address obscure and unreliable, when attempting to assess e-mail.

A more recent focus is on the use of Domain Names, for references that are more stable and align better with the authority boundaries of Administrative Management Domains. Broadly there are two lines of effort at using Domain Names for validating messages being relayed. One associates the identity with the systems that handle the message along its path. These "path registration" schemes include Sender Policy Framework, Sender-ID, and Certified Server Validation. The other schemes tie a Domain Name identity to the message object. These include Domain-Keys Identified Mail, and Bounce-Address Tag Validation.

The *Sender Policy Framework* (SPF) [16] has evolved over time, attempting to encompass multiple identities. It primarily uses the Domain Name in the RFC 2821 MailFrom command. It queries the *Domain Name System* (DNS) with that name and determines whether the IP address of the previous-hop MTA is registered under that name. Since any SMTP server along the transit path may choose to perform this query, SPF requires that the Domain Name contain a registration for every MTA along every delivery path for a message. (A common simplification for this model is to use it only between boundary MTAs, but this considerable constraint is not specified in SPF. Rather, its use is usually characterized as being more general.) Although the software overhead for SPF is quite small, the administrative overhead can become substantial, as the number of paths increase and as paths change. In addition, some sender SPF DNS configurations can trigger a very large number of queries per addressee. Lastly, the role of the RFC 2821 MailFrom command is to specify the Notification Handler address. This address might be entirely different from other origination information, making registration of all of the MTAs in the path problematic. SPF therefore has significant administrative problems with redirected traffic, such as when going through a third-party forwarding service.

Sender-ID (SID) [17] uses a model similar to SPF, but it is based on the posting address Domain Name in the RFC 2822 Sender field (or RFC 2822 From field, if no Sender field is present.) Both SID and SPF sought IETF standardization in 2004 but the working group effort failed, due to lack of rough consensus convergence among participants and due to concerns over intellectual property claims.

Certified Server Validation (CSV) [18] covers only the current client/server SMTP hop. The client specifies an operator's Domain Name in the RFC 2821 EHLO command. The server uses this name to query the DNS. It then validates the IP Address of the SMTP client and determines whether the Domain Name administrator has authorized the client to send mail. CSV also specifies a standard mechanism for querying an assessment service about the client's Domain Name.

DomainKeys Identified Mail (DKIM) [19] specifies an accountable Domain Name that applies to a message during transit. It uses public key cryptography to digitally sign the message and provides guidance when the signing Domain Name differs from the Domain Name in the RFC 2822 From field.

DKIM Domain Name validation represents a significantly different goal from that of the strong authentication methods, such as [20, 21] which focus on long-term protection of message content. Also DKIM places its parametric information in a special RFC 2822 header field, rather than in the message body, so that it does not have any impact on recipient user agents that do not support DKIM. Although public key cryptography has relatively high computational cost, e-mail processing is usually i/o-bound, so that the real-world use of DKIM appears to have little impact on the aggregate message-handling capacity of a server.

Bounce Address Tag Validation (BATV) [22] attacks the problem of misdirected handling notices, such as bounces. It permits the creator of an RFC 2821 MailFrom bounce address to digitally sign it. When the bounce agent of that creator receives a message purporting to be a bounce, the agent can validate the address. Standardization of its format is needed so that e-mail intermediaries—such as some mailing list software—can determine the "core" of the mailbox portion. Since the creator of the signature semantics is the only consumer of the signature semantics, any signature algorithm can be used, including one based on symmetric keys. For convenience—and an existence proof—the BATV specification provides an example algorithm already in use.

Collaboration Support

Fighting spam must be a collaborative effort, which will benefit from using tools and standards that aid in exchanging information and performing coordination. To this end, standard methods of reporting spamming events, of characterizing particular spam, and of sending spam control data can be helpful. Some work in that direction is already underway. [23] Fighting spam requires global operations collaboration; this will be aided by services to facilitate interactions between network administrators speaking different languages. It is also likely that there should be standards for the syntax and semantics of whitelists and blacklists.

Acknowledgement

The author wishes to express particular appreciation for the unusual amount of dialogue that took place with the reviewers of this article. It produced a substantially clearer and more concise article. It also highlighted the extraordinary diversity of views on the topic, in case one had had any doubt. In fact, the article by John Klensin which follows this one is a direct result of the dialog.

References

[1] 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.

[2] Postel, J. B., "Simple Mail Transfer Protocol," STD 10, RFC 821, August 1982.

[3] Klensin, J., "Simple Mail Transfer Protocol," RFC 2821, April 2001.

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

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

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

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

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

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

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

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

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

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

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

[15] Hutzler, C., Crocker, D., Resnick, P., Sanderson, R., and E. Allman, “Email Submission: Access and Accountability,” Internet-Draft, draft-hutzlerspamops-05, October 2005.

[16] Wong M., Schlitt M., “Sender Policy Framework (SPF) for Authorizing Use of Domains in EMAIL, version 1,” Internet Draft, draft-schlitt-spf-classic-02, June 2005. *Now available as RFC 4408, April 2006.*

[17] Lyon J., Wong M., “Sender ID: Authenticating Email,” Internet Draft, draft-lyon-senderid-core-01.txt, May 2005. *Now avilable as RFC 4406, April 2006.*

[18] Crocker D., Leslie J., Otis D., “Certified Server Validation (CSV),” Internet Draft, draft-ietf-marid-csv-intro-02, February 2005. Also see: http://mipassoc.org/csv

[19] Allman E., Callas J., Delany M., Libbey M., Fenton J., Thomas M., “DomainKeys Identified Mail (DKIM),” Internet Draft, draft-allman-dkim-base-00, July 2005. *Also see* http://dkim.org. *Recently approved by the IETF as a Proposed Standard.*

[20] Ramsdell B. (ed.), “Secure/Multipurpose Internet Mail Extensions (S/MIME) Version 3.1 Message Specification,” RFC 3851, July 2004.

[21] Elkins M., Del Torto D., Levien R., Roessler T., “MIME Security with OpenPGP,” RFC 3156, August 2001.

[22] Levine J., Crocker D., Silberman S., Finch T., “Bounce Address Tag Validation (BATV),” Internet Draft, draft-levine-massbatv-00, September 2004. Also see http://mipassoc.org/batv. *Now available as Internet Draft, draft-levine-smtp-batv-oo, January 2006.*

[23] Shafranovich, Y., “An Extensible Format for Email Feedback Reports,” Internet Draft, draft-shafranovich-feedbackreport-01.txt, May 2005. *Also see* http://mipassoc.org/arf.

Ed.: This article is a revision of “Adapting Global Email for Controlling Spam,” in *Information Processing Society of Japan (IPSJ) Magazine—Special issue on Anti-Spam*, Japanese/English, Volume 46, No. 7, pp. 741–746, July 2005.

DAVE CROCKER is a principal with Brandenburg InternetWorking. He has authored or contributed to most Internet mail standards, and an assortment of e-mail products and businesses, as well as working on facsimile, security, ecommerce and EDI. He received the 2004 *IEEE Internet Award* for his work on e-mail. Dave is a contributor to the development efforts for DKIM, CSV and BATV, motivated by a strong desire to protect more than 30 years of professional investment that is being threatened by spamming. E-mail: dcrocker@bbiw.net