

## The State of the Spam Problem

Every day, when people open their e-mail inboxes, they find numerous messages from unknown parties offering a range of services and products. These unwanted messages have come to be referred to as “spam.” Just a few years ago, spam was considered a minor nuisance. The increase in spam over the last few years, however, has led many to focus on this problem. The scale and the effects of the spam epidemic suggest that spam is no longer simply a nuisance but is a large-scale network problem.

### Defining Spam

The definition of spam is neither clear nor consistent across different individuals or organizations. Loosely speaking, we can describe spam as unwanted e-mail messages. These types of messages are often referred to as “unsolicited commercial e-mail.” However, over the years there has been an increase in unsolicited mail that is not necessarily commercial in nature. Therefore, some have begun to refer to these types of messages as “unsolicited bulk e-mail.” This column will focus on commercial spam, which makes up the majority of all unsolicited e-mail.

There are also other types of messages that people consider spam but that do not fit this definition. For example, many people receive forwarded mail such as jokes from family and friends. Some consider these to be spam, whereas others eagerly await such messages. Also, many people inadvertently sign up for mailing lists and newsletters but later may not remember having signed up or may not actually want to receive them. Since these people “opted-in” (specifically signed up

to receive them), the messages are not necessarily unsolicited but may still be classified by some as spam.

The problem of defining spam becomes more complex across different types of organizations. For example, organizations with more liberal network-access policies allow users to receive personal e-mail and mailing lists; other organizations restrict users to receiving only business-related messages and therefore describe all other messages as spam. A good approach to this problem is to define the different categories of messages that may be deemed spam and allow organizations or individuals to create an appropriate definition for their environment.

### Why People Spam

Many wonder why spam activity has increased over the last couple of years. Is it because more people want to be nuisances to society? Spamming is not a pastime but is an actual business process. Spammers are in business, and like most others in business, they have a goal of making a profit. This fact is useful in understanding the swift growth in the use of spam.

As in any other business, spammers must perform a few essential activities in order to create a profit:

1. *Find potential customers.* For spammers, this involves obtaining a list of e-mail addresses. There are two main methods that can be used to obtain these lists: address harvesting and list purchasing.
2. *Offer a product or service to the potential customers.* This involves sending information or an offer to the list of e-mail addresses.

3. *Sell and deliver the product or service to some percentage of the potential customers.*

The success of spam as a business is based on the low cost of #1 and #2, allowing a low response rate to still lead to a profit.

Sending spam can cost \$0.0005 per recipient; direct mail can cost \$1.21 per recipient, or about 2,400 times more. Direct mailers usually require a response rate of about 2 percent; spammers, on the other hand, can break even with response rates as low as 0.001 percent—about 2,000 times lower. For example, a spammer can send 500,000 messages and still be pleased and profitable with five responses.<sup>1</sup>

### The Volume of Spam

Just one year ago, spam accounted for only 10 percent of inbound e-mail traffic; today, spam accounts for over 60 percent of inbound e-mail traffic on average.<sup>2</sup> Consequently, an average user now has more unwanted messages than wanted ones in his or her inbox. This influx of messages has introduced a burden not only on end-users but also on administrators and the infrastructure.

The cost of the spam problem includes lost productivity from the users who must deal with spam messages and from the computing resources that must be used to handle these messages. The computing resources include additional bandwidth, storage, and e-mail servers. Ferris Research reports that spam cost U.S. corporations \$8.9 billion in 2002.<sup>3</sup> IT resources accounted for 44 percent of the cost, lost user productivity accounted for 39 percent, and help-desk resources accounted for the remaining 17 percent. A report by the European Union estimates

the global bandwidth costs of spam at \$8–10 billion annually.<sup>4</sup>

### Solving the Problem

Dealing with spam requires a long-term cure as well as some immediate symptom relief. Currently, anti-spam products that are available for e-mail gateways and for desktops can detect and block over 90 percent of spam. These solutions are effective at relieving the spam problem for organizations that have decided to deploy them. Although colleges and universities, corporations, and ISPs are rapidly deploying such tools to protect their networks, many mailboxes across the Internet remain unprotected.

Stopping spam globally requires removing the incentive to spam. The overall solution to spam thus must aim to reduce the profitability of sending unwanted e-mail. Just as in any other business, the profit in spamming is equal to revenues minus expenses. In spamming, expenses include the cost of obtaining the lists of e-mail addresses and the cost of sending the messages. Revenue is equal to the number of spam messages actually received by the intended recipients times the response rate times the profit per item. Both expenses and revenue can be affected by user education, which influences the recipients' response rate as well as the spammer's difficulty and costs involved in obtaining e-mail addresses.

In addition, the difference between the number of spam messages sent and the number received is a factor of the effectiveness and deployment rate of anti-spam technologies. These technologies are divided into three types of systems: spam prevention, spam detection, and spam response. The first type, spam-prevention systems, includes deterrence and protection approaches. Deterrence systems aim to discourage the act of spamming, whereas protection systems aim to shield systems from exposure to spam. The second type of technology, spam-detection systems, attempts to distinguish spam messages from non-spam messages. This distinction must be based on four questions asked about the message: (1) Who is it from? (2) What is in it? (3) How was it sent? (4) Where was it sent?

Spam-response systems make up the

## The overall solution to spam thus must aim to reduce the profitability of sending unwanted e-mail.

third type of anti-spam technology. Traditionally, the response to spam has been deletion. Current spam-response systems allow a range of responses including quarantining the message for review by an administrator or end-user. This response has grown in popularity as the community has become sensitive to false positives, or non-spam that is incorrectly detected as spam. Challenge-response systems are increasingly used in some situations. When a suspected spam message is received, these systems reply to the sender with a "challenge," an e-mail requesting some sort of response to verify that the sender is a person rather than bulk-mail software. Some proposed systems would respond to spam messages by charging the sender. Other systems involve a passive response that provides feedback to systems that track the behavior of mass mailers.

Finally, anti-spam legislation not only provides a strong deterrence but also introduces spammer overhead in the form of the expenses of litigation. The relationship between technology and legislation is a familiar one. For example, consider the problem of computer intrusions: there is a set of technology available—such as firewalls and intrusion-detection systems—to protect resources, and the technology is supported by legislation to allow prosecution of and litigation against those who are determined to circumvent the technology. This same relationship is seen in something as basic as property theft: burglar alarms and door locks provide protection, and they are supported by laws that provide a strong deterrence.

### Moving Forward

The recent effects of spam have led both lawmakers and technologists to focus on

solving the problem. On the legal front, several pieces of federal anti-spam legislation already have been proposed this year. On the technical front, the Internet Research Task Force formed the Anti-Spam Research Group (ASRG), which focuses on technical solutions to the problem.<sup>5</sup> This research group consists of hundreds of members representing computer researchers, anti-spam companies, e-mail server companies, e-mail service providers, ISPs, anti-spam advocacy groups, end-users, and administrators.

The type of collaboration and systematic focus provided by the ASRG presents a unique opportunity in the fight against spam and represents a different approach from the ad hoc efforts of the past. Participation from each member of the e-mail ecosystem sets the stage for a focus on interoperable and deployable solutions. Due to efforts like this, many believe the spam problem will be well under control within the next couple of years. Of course, we must not merely hope for the best; instead, we must take a range of active steps—in policy and laws, in new technologies, and in the education of both the public and industry on this issue.

### Notes

1. Direct mail costs for 10,000 recipients: design: \$4,000; printing: \$5,700; postage: \$2,400. On spam costs, see Mylene Mangalindan, "For Bulk E-Mailer, Pestering Millions Offers Path to Profit," *Wall Street Journal Online*, November 13, 2002, <[http://online.wsj.com/article\\_email/0,,SB1037138679220447148,00.html](http://online.wsj.com/article_email/0,,SB1037138679220447148,00.html)> (accessed June 12, 2003).
2. CipherTrust Research: <<http://www.ciphertrust.com>> (accessed July 10, 2003).
3. Marten Nelson, "Spam Control: Problems and Opportunities," *Ferris Research*, January 2003, <<http://www.ferris.com/url/spam.html>> (accessed July 10, 2003).
4. Serge Gauthronet and Etienne Drouard, "Unsolicited Commercial Communications and Data Protection," Commission of the European Communities, January 2001, <[http://europa.eu.int/comm/internal\\_market/privacy/docs/studies/spamstudy\\_en.pdf](http://europa.eu.int/comm/internal_market/privacy/docs/studies/spamstudy_en.pdf)> (accessed July 10, 2003).
5. See <<http://www.irtf.org/asrg>>.

**Dr. Paul Q. Judge is Chief Technology Officer at CipherTrust, the provider of anti-spam and e-mail security solutions for large enterprises. Judge also serves as chair of the Internet Research Task Force (IRTF) Anti-Spam Research Group (ASRG). He earned his Ph.D. in Computer Science from the Georgia Institute of Technology, where he currently is a Post-Doctoral Fellow.**

