



## FIGHT BACK AGAINST SPAM

### The evolution of anti-spam technology

By [Ron Herardian](#)

Unsolicited Commercial email (or UCE or "spam") has been a growing problem for the Internet since the mid 1990s. At that time, the Internet was rapidly becoming commercialized as well as accessible to consumers. Initially perceived merely as a nuisance created by a handful of unethical advertisers, spam currently accounts for the majority of email traffic over the Internet. Spam has a negative impact on the ability of businesses and consumers to benefit from the use of email technology and the Internet, and it poses a serious and growing threat to the reliability, efficiency, and security of corporate electronic messaging systems and the Internet.

### Spam not caused by technology

Most of the discussions about spam focus on technology issues. However, the fundamental factors driving the increase in spam are economic. The technology of Internet email is such that the recipient and intermediate service providers pay the vast majority of the cost involved in delivering spam messages.

Analyst firm Ferris Research has estimated that, in 2004, spam cost U.S. companies over \$10 billion per year. At the same time, spam represents the least expensive way of advertising to literally millions of businesses and consumers. The cost to spammers of sending millions of email messages over the Internet can be quickly recovered through a small number of sales, thus even if the response rate for spam is extremely low compared with legitimate advertising media, (e.g., 100 sales in response to 10 million email messages), it remains profitable to send spam.

As long as the economics of spam hold true, and as long as it is technologically possible to advertise through spam without being held accountable for the true costs, the volume of spam will certainly continue its increase.

### Messaging industry slow to respond

By the late 1990s, spam had become a major problem for Internet Service Providers (ISPs) and businesses. ISPs and enterprises were forced to take steps to stop spam from overwhelming their email servers and flooding computer networks.

Although the problem was widely recognized in the messaging industry, there were relatively few anti-spam tools and technologies. The messaging industry, comprising vendors of email infrastructure software and related products, as well as standards bodies, was slow to respond to the spam problem--having initially underestimated the scale and technical complexity of the spam problem.

In particular, standards bodies such as the Electronic Messaging Association (now the OpenGroup Messaging Forum) and the Internet Engineering Task Forces (IETF) failed to effectively address the spam problem, although the IETF did eventually create the Anti-spam Research Group (ASRG) in 2003.

At the same time, spammers rapidly multiplied, became far more sophisticated, and began to operate through quasi-legitimate ISPs and to move operations offshore (substantially as a result of legal risks made apparent through high profile law suits such as CompuServe Incorporated vs. Cyber Promotions Inc. and Sanford Wallace in 1997).

In the mid 1990s, the methods used to slow the flow of spam were based upon crude technological capabilities, such as restricting the "mail relay" feature of Internet email servers, technology fixes that were not designed

specifically to address the spam problem.

In the late 1990s, the first true anti-spam technologies emerged, such as the unsolicited bulk email filter built into the Netscape Messaging Server product. However, most messaging infrastructure products did not have such capabilities and, by 2000, it was clear that businesses would have to play a role controlling spam in order to protect their own server networks.

It was also clear at that time that the existing capabilities of messaging infrastructure software and associated products, such as email anti-virus gateways, were inadequate and that that activities of industry standards bodies were not keeping up with the problem. This situation, combined with the fact that U.S. court cases related to spam, such as the 1997 court case mentioned above, had no apparent deterrent effect on spammers, presented a potential business opportunity for anti-spam technologies.

## **Precursors of anti-spam technology**

Virtually every Simple Mail Transfer Protocol (SMTP) Message Transfer Agent (MTA) has some native ability to control communications with other MTAs over the Internet. For example, sendmail, the most widely used SMTP MTA in the world, has the ability to disallow relaying and to limit the networks from which it accepts messages.

These features, referred to below as precursors and first-generation technologies, were not created specifically to address the spam problem, nor do they constitute an anti-spam technology. In the past few years, messaging product vendors seeking to address the spam problem have implemented rudimentary capabilities like these and crude features such as simple "white lists" and "black lists" into basic anti-spam feature sets. The latter approach, which is still evident in many products that include SMTP MTAs such as IBM Lotus Domino, does not truly represent an anti-spam technology.

True anti-spam technologies, referred to below as second and third generation technologies, are a relatively new development that followed the advent of spam in the mid 1990s. Anti-spam technologies are entirely new and are sharply differentiated from basic MTA controls. In particular, Bayesian algorithms and more recent advances in text analysis, for example, using Artificial Intelligence, are a radical departure from pre-existing mechanisms characterized or repurposed as anti-spam tools.

A short description of the successive technologies that have used to combat spam would cover the following generations.

## **Anti-spam technology - the first generation**

It's difficult to sharply distinguish between pre-existing features available in most MTAs and first-generation anti-spam features because pre-existing features were used to combat spam together with new features developed largely to address the spam problem. For example, message header and envelope tests as well as simple DNS tests were motivated primarily by the need to combat spam but cannot directly distinguish between spam messages and other messages.

### **Basic MTA controls**

MTA controls that restrict communications based on networks and domain names, e.g., to prevent "relaying" are not an anti-spam technology. Some of these features predate the spam problem and exist primarily for the purposes of security.

### **White lists and black lists**

White lists and black lists in an MTA are a logical extension of basic MTA controls used primarily as a crude (and ultimately ineffective) tool to control spam. Although white lists and black lists are an important anti-spam feature, they are not a new technology because they merely extend rather than depart from historical MTA controls. MTAs in email anti-virus and security/compliance products had this feature before it was characterized as an anti-spam capability.

### **Simple keyword searching**

One of the first methods used to combat spam was simple keyword searching. This functionality existed prior to spam becoming a major problem on the Internet, as a part of content filtering and compliance solutions and server-based email anti-virus products.

This approach can be used to identify a subset of spam messages, but is not effective because it fails to recognize word variations or context and can result in many false positives (legitimate email misidentified as spam).

MTAs in email anti-virus and security/compliance products had this feature prior to its use as an anti-spam tool. Simple keyword searching is trivial for spammers to defeat through simple variations in spelling or alternate word choices.

### Message header and envelope tests

Message envelope testing means that the MTA checks the information passed through the SMTP protocol, for example the sender and recipient information, used when transferring a message and rejects messages if they are not transmitted with valid information. Message header tests give an MTA the ability to check information stored inside of messages such as the To, From, and Date fields and to reject messages if the header is not properly formed or contains invalid information.

While these capabilities are obviously useful to eliminate spam, they ultimately guarantee only that messages are correctly transmitted and constructed, not whether they are sent by a spammer or if the message contents are spam -- thus they are not an anti-spam technology in their own right. Also problems can occur in message headers and envelopes for reasons not necessarily indicative of spam.

### Simple DNS tests

Looking up sender information exchanged during the SMTP protocol using the Internet Domain Name System can be used to validate information exchanged during the SMTP protocol -- for example to check for the existence of the sender's Internet domain or the name of a machine sending messages (by looking up the name associated with the sender's Internet address).

Simple DNS tests help to prevent "spoofing" (when a machine masquerades as another by using its name). Although simple DNS tests are a significant tool to combat spam, they are not an anti-spam technology as such because they only check a sending machine's name and address, not whether the sender is a spammer or if the message contents are spam. Simple DNS tests are a weak technology because that can indicate a problem for a variety of completely legitimate reasons thus are not necessarily indicative of spam.

First generation	Second generation	Third generation
Basic MTA controls	Realtime Black Lists (RBLs)	Artificial intelligence
While lists and black lists	Signature networks	Machine learning
Keyword searching	Bayesian filtering	-
Message envelope tests	-	-
Message header tests	-	-
Simple DNS tests	-	-

### Anti-spam technology - the second generation

Realtime Black Lists and signature networks are second-generation anti-spam technologies, because they are not a simple reuse or extension of pre-existing MTA features for anti-spam purposes. Second generation anti-spam technologies exist purely for the purpose of stopping spam and are vastly improved compared with prior methods.

#### Realtime Black Lists (RBLs)

Although they are technically a DNS test when based on Internet addresses rather than domain names, RBLs were the first true anti-spam technology. The concept behind RBLs is simply to maintain a list of the Internet addresses that send spam and block them from further transmission.

The technology has some effect but can be easily circumvented through a variety of means such as changing IP addresses or relaying messages through a third party not previously identified as a spammer. Similarly, domain names can be easily acquired, spoofed, or fabricated so that a spammer's sending domain cannot be depended upon to detect spam.

[There are also some vigilante RBLs out there that will block a domain or IP address with a single complaint. We've found that some RBLs actively block email from IBM, Lotus, Microsoft, and even DominoPower simply because they don't agree with our message or, in the case of Microsoft, their product directions and choices. -- Editor]

### **Signature networks**

These are a significant and relatively new technology to combat spam. The concept of a signature network is to collect and identify spam messages by generating a unique 'signature' that can be used to identify a given message. Since spam messages are sent in bulk (many copies of the same message) this approach, if supported by a sufficient sample of spam messages, can stop a significant percentage of spam. Nonetheless, the concept depends upon the timeliness of the signature network's operation since spam must first be received somewhere in order for a signature to be generated.

### **Anti-spam technology - the new generation**

Identifying spam (signatures) and spammers (RBLs) after the fact is an approach doomed to fail. Spammers can easily circumvent RBLs and even the best signature network takes time to identify spam messages and can never detect one hundred percent of them. Circa 2002, a radically new approach gained prominence in the Internet community and software industry.

### **Bayesian filtering**

Bayesian filtering, which is a statistical approach to spam detection based on the spam probabilities of individual words, was the first breakthrough anti-spam technology. Premised upon the idea that the commercial content of spam messages is the Achilles heel of spam, the development of Bayesian filters fundamentally changed the focus of anti-spam efforts from networks and protocols to the content of messages.

Simple Bayesian filters, while effective on most spam, are easily circumvented. The approach relies on 'training' filters by processing known spam to generate a scoring system based on 'spam words' that is then used to evaluate new messages.

Spammers quickly learned to continuously vary the contents of messages by adding neutral words or word variations (such as substituting numbers for letters as in the number "0" for the letter "O"). By constantly varying the neutral words and other message contents and by creating new word variations, Bayesian filters can be consistently circumvented as they are always one step behind the most recent spam.

### **Artificial intelligence and machine learning**

While improved techniques based on Bayesian filtering (e.g., taking into account word proximity, stripping HTML tags used by spammers to break up words, etc.) continue to be widely used, a more intelligent approach was needed.

Around 2003, the need for new, specialized anti-spam technologies became clear and a number of high tech startups came into existence to develop these technologies, among them Corvigo (acquired by Tumbleweed in March 2004), Proofpoint, and others. The technological approaches of these companies were based on software algorithms originating in the field of artificial intelligence, a branch of computer science.

Fundamentally, these technologies perform text classification using non-Bayesian techniques. However, they must also adapt automatically to the changing character of spam. To do this machine learning techniques have been applied.

### **The future of anti-spam technology**

Spam exists partly because, when SMTP was originally created, only legitimate government and military, university and industry entities were connected to the Internet. There was no concept of illegitimate use or abuse of the Internet or email because it was a closed system.

With the commercialization of the Internet in the 1990s, the situation changed but the technology of the Internet did not. Standards bodies are today specifying new technologies that will help to eliminate spam by enforcing a greater degree of legitimacy on message senders. Early efforts along these lines, sometimes promoted as "anti-spam" technologies, were actually attempts to legitimize bulk commercial email considered by

many to be spam.

In contrast, current standards-based efforts to establish a sender authentication standard using digital certificates to as domain keys to generate digital signatures for messages can be employed by all enterprises, rather than exclusively by senders of bulk commercial email.

New standards, however, will take several years to become incorporated in a majority of Internet MTAs. It is likely that third-generation anti-spam technologies will persist as a permanent feature of the messaging landscape.

In coming months, we'll take a look at the Anti-Spam Technical Alliance, a group lead by Microsoft and several major ISPs including AOL, Comcast, Yahoo!, Earthlink, and British Telecom. Microsoft's Microsoft's proposed CallerID approach and the Sender Profile Framework proposed by anti-spam researcher Meng Wong are two components of what will certainly become a new anti-spam Internet standard.

### **Product availability and resources**

For more information on Corvigo, visit <http://www.corvigo.com>.

For more information on Proofpoint, visit <http://www.proofpoint.com>

Ron Herardian is CEO & Chief Systems Architect at Global System Services (GSS). You can reach Ron via email at [rherardi@gssnet.com](mailto:rherardi@gssnet.com), or via his web page at <http://www.gssnet.com>.

Copyright © 1998-2008, [ZATZ Publishing](#). All rights reserved worldwide.