The Spam Problem and Brightmail's Solution

BRIGHTMAIL

February 15, 2002

The Spam Problem and Brightmail's Solution

Brightmail Incorporated

Document Version: 2.0

The Spam Problem and Brightmail's Solution

Brightmail Incorporated
301 Howard Street, 18th Floor
San Francisco, California 94105
U.S.A.
Phone +1(415) 365-6180
Fax +1(415) 348-9636
www.brightmail.com

Table of Contents

Introduction

Email is here to stay. Despite the disappointing revelation that the Internet revolution is not capable of ending the cyclical nature of the global economy, the Internet is alive and well. The most frequently used Internet application—email—is well established as an essential tool for businesses of all types, and for personal communications as well. Email addresses are replacing phone numbers as the prime contact method for friends, family, and colleagues. According to IDC (September 2000), email messages sent per year will increase from 9.7 billion in 2000 to 35 billion in 2005.

As with other innovations, the Internet is not without its unpleasant side effects. With respect to the most popular and important application of the Internet, email, this dark side comes in the form of unsolicited bulk email (UBE), commonly known as junk email or spam.

In contrast to other esoteric and technical problems on the Internet, spam is a topic that a broad cross-section of the Internet population can relate to. The system administrators of corporations and Internet service providers (ISPs) are on the front lines. They must regularly fight to defend their resources and try to stay one step ahead of the spammers who attempt to hijack mail servers and computers. And when these efforts fail, recovery is costly and time consuming. Other victims are parents, who typically want to prevent their kids from viewing objectionable content, but must contend with the salacious offers and links to pornographic Web sites that wind up in their children's email inboxes everyday. From a spammer's perspective, the most sought-after targets are those who use email every day for work and personal reasons. Many members of this group are simply resigned to the chore of sorting out and deleting junk mail from legitimate mail, a necessary evil that they must endure every time they check mail.

To say that we are under attack is not hyperbole. Brightmail Inc. estimated that the average spammer sends 250,000 junk emails a day. As troubling as this figure sounds, it is probably quite conservative. The tools and techniques that facilitate the delivery of spam are always improving, allowing spammers to continually update their arsenal with more efficient and productive tools. Additionally, the concept of “Internet time,” the phrase coined to describe the exponential growth and rate of change of technology related to the Internet, raises the stakes. As the processing power and bandwidth of the Internet as a whole increases, the amount of spam can keep pace and clog networks and mail systems at the same rate. Left unchecked, this costly “scourge of the Internet” will continue to impede the development of email as a powerful tool of business, communication, leisure, and entertainment.

This white paper describes the costs and effects of the spam problem in detail. In addition, this paper provides some insight into the challenge of identifying spam and distinguishing it from legitimate mail. Finally, this paper describes Brightmail's unique approach to stemming the tide of spam.

Spam and Who Sends It: Some Quick Definitions

For the purposes of this paper, spam is any email message, regardless of its content, that is sent to multiple recipients who have not specifically requested the mail.

A wide range of individuals and organizations send spam. On one end of the spectrum are those who knowingly and purposefully send unsolicited email. Members of this group indiscriminately blanket networks with messages, use purloined resources, employ devious and fraudulent tactics, and know full well that the recipients have not requested such communication. The business model of this type of spammer is to off load the costs of sending the messages entirely on the recipient and other unrelated parties. Because these spammers have no real out-of-pocket expenses, all they need to do is dupe a few unsuspecting recipients to make a profit.

Other people who send spam are much more innocent in their intent and approach. Everyday email users, by engaging in such seemingly harmless acts as forwarding on a chain letter to multiple recipients, often send spam without even realizing it. In addition, legitimate and ethical companies often send email to large lists of potential customers, taking advantage of a cost-effective medium to quickly reach a wide market. However, if such companies do not take adequate precautions to ensure that the recipients have explicitly requested or agreed to receive the email, they too are sending spam. Members of this group, while certainly not malicious, are also contributing to the spam problem. As we will see, intentional or not, sending unsolicited email has many costs.

Why Spam is a Problem

Defenders and purveyors of large-scale UBE practices claim that the resulting spam is harmless. After all, recipients of unwanted mail can quickly delete messages. In fact, proponents of UBE frequently compare their approach with traditional bulk mail delivered by the post office. On first glance, perhaps, the practices are similar: advertisers get quick access to a large advertising base and the burden of time and effort is placed on the recipient to sift and sort through the unsolicited communication. However, beyond this, the analogy quickly breaks downs. Postal bulk mail is paid for by advertisers and these revenues theoretically support the U.S. Post Office and its delivery infrastructure. It is also illegal to place mail in a mailbox without postage or to tamper with any aspect of the delivery process. Realistically, there are currently no such proscriptions and punishments with email. Postal customers can also choose to "opt out" and not receive bulk mail. While not foolproof, opting out generally cuts down on unwanted postal mail. In the online world, which lacks a central governing body and a track record with such measures, email users have little confidence that any similar online global opt-out list will have any positive effect.

Indeed, UBE bears much more of a resemblance to the phenomena of “junk faxing,” where the recipients pay—in terms of paper costs and wear and tear on machines—to receive advertisements. Although the analogy is closer, email is part of a world unto itself. Given the nature of the Internet, by definition a dynamic network of interconnected computers that cooperate, the aggregate costs of abuse can become astronomical. This section describes some of these costs in more detail, in some cases by uncovering tactics used by some less scrupulous spammers. “Just click delete” is hardly a solution.

Cost Shifting and Resource Abuse

Implicit in the above comparison between postal bulk mail and spam is the inherent cost-shifting nature of spam. By design or as an unintended side effect, the sender of spam bears a relatively small proportion of the total cost of delivery. For Internet service providers (ISPs), application service providers (ASPs), wireless service providers (WSPs), corporations, and other resource operators who own and manage the crucial hardware links on the Internet, the resultant costs are high. WSPs face these costs and also the potential for customer backlash, as customers generally pay for incoming messages.

Beyond the direct hardware, software, and bandwidth resources consumed by spam, and the customer retention issues, there are the storage costs and the personnel costs to deal with spam-related customer complaints and infrastructure demands. The final recipient ultimately pays for all of these costs. How does the spammer avoid paying these costs? Given the mechanics of how spam is purveyed, and of the operations of many mail systems and Internet protocols, this is a surprisingly easy goal to achieve.

For dishonest spammers, an easy way to misappropriates resources and services is to exploit third party relay systems. An email message does not always go directly from the originating mail server to the mail server that handles final delivery. There can be many stops at different mail servers along the way, a feature that allows for flexible email delivery and provides safeguards if one link on the Internet is down or congested. These intermediary machines in the delivery process are known as relays.

By altering and forging headers—the parts of the email that provide tracing information about messages—spammers can relay spam messages off the mail server of an innocent third party. In doing so, the ultimate receiving system and the innocent relay system are deluged with spam. And because the identifying information of the spam has been altered, the resulting flood of complaints is sent back to an innocent site, which has been made to look like the origin of the spam. Frequently, the spammer has sent mail from a “throwaway” email account, one that was created solely to launch a spam attack, never to be accessed again. After initially sending out their stream of messages, spammers bear no significant costs or repercussions for their advertising.

On the other hand, the resulting complaints and spam have significant economic repercussions for the service providers involved. First, there are the direct costs and maintenance of message storage as complaints, bounced mail, and spam fills up user mailboxes and clogs abuse desks at the ISP or corporation. Even if this onslaught doesn't cause disks to crash or hardware to fail, it can impact the service to legitimate

users. Because computers can only process a finite number of actions per second, and because only a certain amount of data can be passed along the bandwidth of an Internet connection, massive quantities of email can severely disrupt an ISP's service to its other subscribers.

There is also the inevitable cost of investigating and replying to complaints. Due to the indiscriminate nature of spamming, many of the messages have incorrect delivery information. When the mail arrives for email boxes that are no longer valid, the Postmaster of the ISP suddenly must contend with thousands of additional messages. Add to this the damage to brand for organizations identified as the source of unsolicited mail. Any subsequent proactive measures taken by the victim organization, in terms of strategy or purchasing of additional hardware, are additional direct costs that spammers have shifted to them because the actions are solely in response to the attack.

Although most administrators of ISPs and corporations can configure their mail servers to disallow relay abuse, others cannot. In some cases, the technical limitations of their mail server software and support contracts prohibit such changes. In other cases, such a restriction might cause a negative relationship with their customers, some of whom may need to dial in using a relay system. In any event, even if virtually all mail server administrators on the Internet closed up relay holes, all it would take is exploitation of the few remaining sites to cause significant damage.

What about those who don't exploit relays or hack into other systems to deliver unsolicited email? Even when a company uses its own resources, cost shifting persists. Many end recipients feel the effects of cost shifting first hand. If email customers belong to an ISP that charges for access on a metered basis, customers are directly paying for the privilege of reading or deleting spam. Many email accounts have limited mailbox storage available, and spam can quickly consume this scarce resource. Finally, users can suffer through higher prices or lower service levels due to their ISP's preoccupation with fighting spam.

Fraud

Spam and the practice of disseminating it have many elements of fraud. The exploitation of open relays described in the previous section demonstrates fraud on several levels. First, by corrupting headers and routing information, spammers resort to devious methods in order to conceal their identity and location. This practice, along with sending huge amounts of messages, is generally prohibited by the "term of service" agreements that govern the use of email accounts with ISPs. Spammers make use of software that effectively conceals their tracks, allowing them to violate the terms of service agreement with relative impunity. The practice of using free or throwaway accounts to disseminate the mail represents additional fraud, with the spammers often providing false names and billing data in order to open new accounts after previous accounts have been terminated.

For the email user who receives spam, the actual content of the messages often gives rise to fraud. According to the spam watchdog group CAUCE, the most prevalent spam messages are:

- Chain letters

- Pyramid schemes (including multilevel marketing)
- Other “get rich quick” or “make money fast” schemes
- Offers of phone sex lines and ads for pornographic web sites
- Offers of software for collecting e-mail addresses and sending spam
- Offers of bulk emailing services for sending spam
- Stock offerings for unknown start-up corporations
- Questionable health products and remedies
- Illegally pirated software

Finally, in order to get around filters and to force people to actually read the messages, spammers employ a wide range of deceitful strategies. These can include anything from inserting misleading subject lines, adding official-sounding content to the message, or other tricks to entice recipients to open the message.

Loss of Productivity

Lost productivity is another negative effect of spam. The cumulative costs add up quickly when email users spend a few minutes a day dealing with and disposing of spam. In a business setting, where the adage, “time is money” holds true, any time employees are forced to spend dispensing with spam is time spent away from legitimate work. Organizations need to examine what percentage of their labor costs are lost because employees are sifting through junk email, not to mention the diversion of attention of data center and MIS staff. There are other productivity drains as well: on a legal front, there have been many instances of lawsuits as a result of pornographic and other messages circulated via email in the workplace.

Social Costs

Another set of “costs” are perhaps more difficult to explore fully at this point: the social costs of spam. Social costs refer to the extent to which spam discourages people from appreciating and getting involved in the Internet. In other words:

- Does the specter of getting flooded with spam and having their privacy invaded make people more reluctant to participate in chats or newsgroups?
- Does the fear of address harvesting, whereby spammers cull email addresses from Internet locations such as auction sites, Web pages, and magazines, cause people to bow out of the greater communication potential of the Internet?
- Are people tempted to provide fake email addresses, which in itself compounds the problem, as more mail gets “bounced,” which means more resources are wasted?
- Is spam a considerable factor that causes people to distrust the Internet in general?

A truly unfortunate side effect of spam would be if it prevented the Internet from realizing its potential as a truly vibrant, participatory, and inclusive mode of communication.

Quantifying Effects of Spam: The GartnerGroup Study

The preceding section described some of the general costs of spam and how these costs are incurred. For some hard data, we next look to a major study produced by Gartner Consulting, a division of GartnerGroup. This effort, commissioned by Brightmail Inc., culminated in the study titled *ISPs and Spam: The Impact of Spam on Customer Retention and Acquisition*. The study was published in March, 1999. This section presents some of the essential findings of the GartnerGroup study.

Methodology

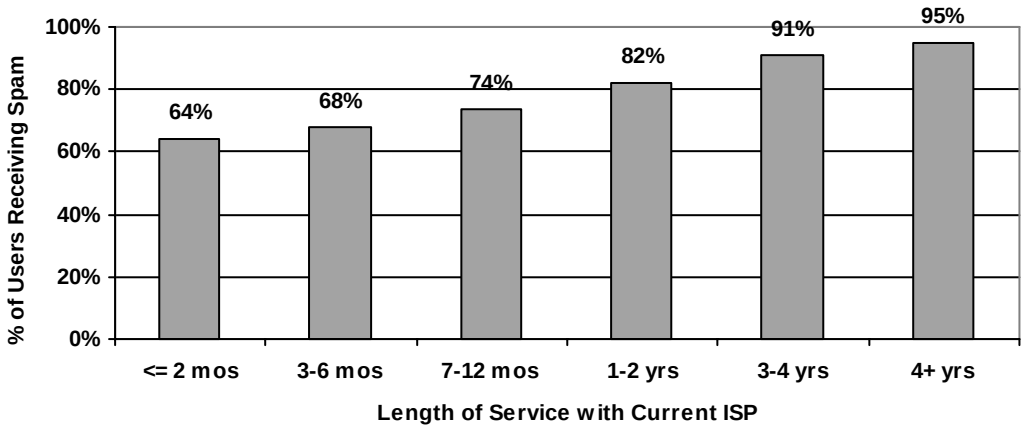
In November 1998, GartnerGroup began an online survey of Internet users. Users were invited by their ISPs to participate. Upon acceptance, they were sent to a unique and secure Web site to fill out the survey. ISPs either sent invitations by email or posted invitations at their Web site. GartnerGroup recruited ISP participants from a list of major ISPs, including AOL, EarthLink Networks, Juno, and others. Data collection ended in March 1999 with more than 13,100 responses collected.

Respondents were asked which ISPs they use and, if they subscribed to more than one, which they consider their primary ISP. ISPs with the heaviest representation included America Online, AT&T WorldNet, Concentric, Juno, MSN, and Netcom (now Mindspring). Thirty-eight percent had been with their ISP for one year or less; most of those had switched from another ISP. Thirty percent had been with their ISP for three years or more.

Frequency of Spam Mailings

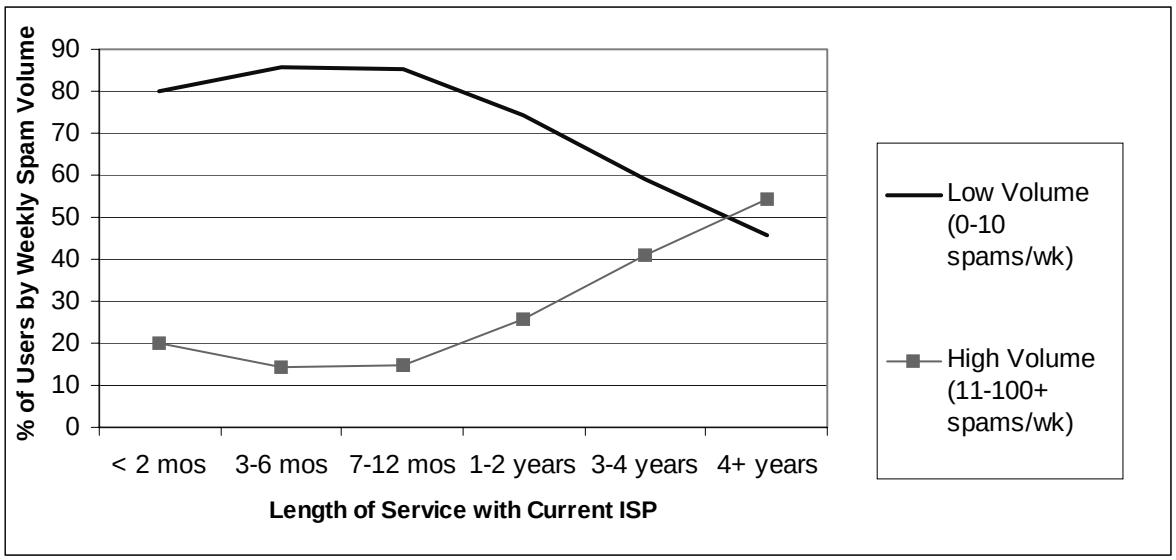
One significant finding was the direct relationship between length of time with an ISP and the likelihood of receiving spam messages. The longer a subscriber stays with an ISP, the study found, the greater the probability of getting spammed, as indicated in Figure 1, “Spam and Length of Time with ISP”.

Figure 1. Spam and Length of Time with ISP



Furthermore, the volume of weekly spam messages increased with length of ISP service, as seen in Figure 2, "Figure 2. Spam Volume and Length of ISP Service."

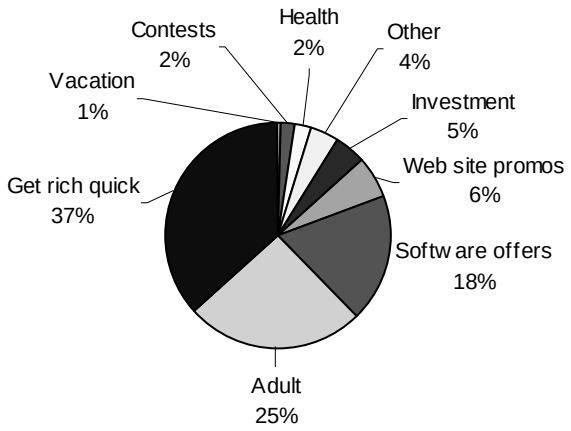
Figure 2. Figure 2. Spam Volume and Length of ISP Service



Types of Spam

The two most frequent types of spam received were characterized as “get rich schemes” and “adult.” Other types included health promotions, software offers, investment information, and Web site promotions. Figure 3, “Figure 3. Types of Spam” shows the breakdown. There is particular adverse sensitivity by females to mailings falling into the adult category. When asked why they dislike spam, 23 percent of females, compared to 13 percent of males, stated that they found it offensive.

Figure 3. Figure 3. Types of Spam



Attitudes Toward Spam

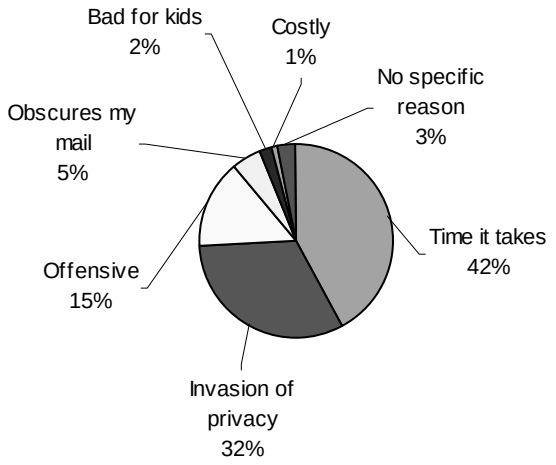
Clearly, respondents were familiar with spam. Given the fact that many were receiving significant numbers of these mailings, what did they think about it? Not surprisingly, they were overwhelmingly unhappy with it. 83 percent disliked spam. Table 1 on page 8 shows respondents' attitudes toward spam.

Table 1. Attitudes Toward Spam

Attitude	Percent of Responses
Like it a lot	1
Like it somewhat	2
Neutral	14
Dislike it somewhat	20
Dislike it a lot	63

Why such strong reactions? Figure 4, "Primary Reasons for Disliking Spam" helps explain the sources of these reactions. Interestingly, unlike the email service providers, who are burdened with significant direct costs in dealing with spam, the end-user apparently does not perceive a direct financial impact—only one percent complained about the cost. This makes sense because almost all respondents were paying fixed monthly subscription fees.

Figure 4. Primary Reasons for Disliking Spam



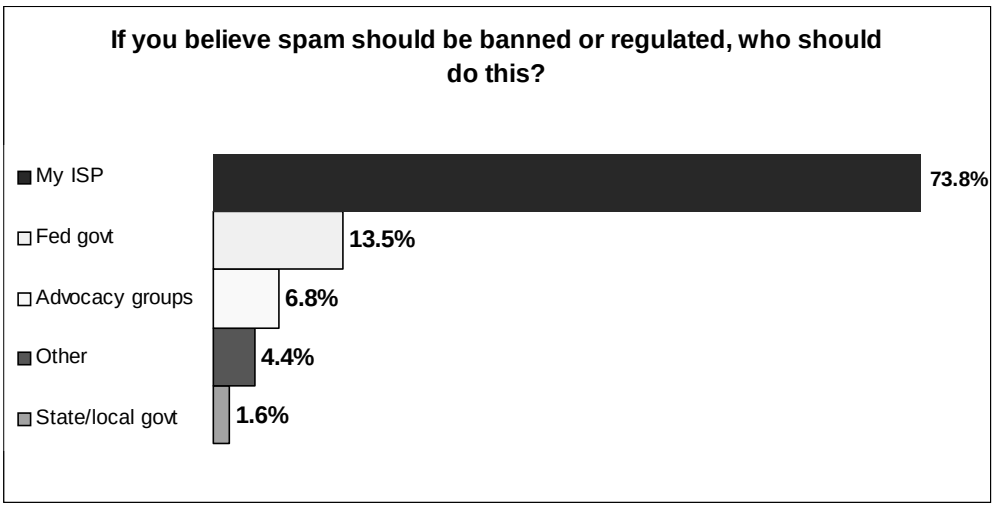
However, 42 percent said that the time it takes them to read and discard spam was their main reason for disliking it. This could be considered an indirect cost to the end-user ("time is money"). In addition, over 30 percent complained that spam constitutes a significant invasion of their privacy, and 15 percent found it offensive.

Role of the ISP

Although one-quarter of the respondents felt nothing could be done to control spam-other than deleting it oneself-40 percent favored banning it and 25 percent wanted it regulated.

Those favoring banning or regulation overwhelmingly (almost 74%) looked to ISPs as the principal spam regulator, as opposed to government agencies or advocacy groups. Figure 5, "Who Should Regulate Spam?" shows their preferences.

Figure 5. Who Should Regulate Spam?



The Bottom Line: Churn Rate, Infrastructure, Personnel Costs

The final section of the study analyzed the financial impact for a hypothetical ISP. There are three areas of an ISP's business that are affected by spam and that potentially affect the ISP's profitability:

- Churn rate costs: Lost revenue due to customer defections and new customer acquisition costs incurred to replace customers who defected.
- Infrastructure costs: Hardware costs and software development costs.
- Personnel costs: Customer support personnel and system administrators.

In order to understand the quantitative impact on ISPs, GartnerGroup made some assumptions based on discussions with numerous ISPs. A representative ISP had:

- A subscriber base of 1 million.
- 4.5 percent of total churn/month, or 45,000 subscribers/month lost to churn.
- Of the total monthly churn, 7 percent is due to spam (3,150 subscribers/month lost because of spam).
- Acquisition costs of $75/customer. The monthly acquisition cost is the amount that must be spent to replace a lost customer.

- Average subscription revenue of $19.95/month.
- Defecting customers leave at the end of the month.

Table 2, "Financial Impact of Churn," sums up the losses that a one-million-subscriber ISP experiences as a function of the sum of lost subscriber revenues due to spam, and the total acquisition costs to replace those subscribers.

Table 2. Financial Impact of Churn

Month	Cumulative Lost Subscriptions	* Monthly Fee	Lost =Subscription Revenue	Monthly +Acquisition Costs	= Total Impact
January	0	19.95	0	236,250	236,250
February	3,150	19.95	62,843	236,250	299,093
March	6,300	19.95	125,685	236,250	361,935
April	9,450	19.95	188,528	236,250	424,778
May	12,600	19.95	251,370	236,250	487,620
June	15,750	19.95	314,213	236,250	550,463
July	18,900	19.95	377,055	236,250	613,305
August	22,050	19.95	439,898	236,250	676,148
September	25,200	19.95	502,740	236,250	738,990
October	28,350	19.95	565,583	236,250	801,833
November	31,500	19.95	628,425	236,250	864,675
December	34,650	19.95	691,268	236,250	927,518
TOTAL			$4,147,608	$2,835,000	$6,982,608

Source: GartnerGroup

On the revenue side, ISPs that offer comprehensive and effective filtering services would seem to have a competitive advantage compared to ISPs that don't offer such services. The survey indicated that current spam-sensitive subscribers would be drawn to ISPs with these services. New users accessing the Internet select their ISP based on a variety of factors, one of which is certain to be the promise of a spam-free environment. The study suggests that ISPs could realize increased revenues from attracting "incremental" new customers they would not have ordinarily gained.

Personnel and infrastructure costs related to fighting spam were also significant. Based on the representative 1 million-subscriber ISP, and assuming a total of between $5 million and $10 million of telecommunications and computer hardware, the hardware-related incremental costs of managing spam were approximately 2.5 percent, or between $125,000 and $250,000 per year.

For personnel costs, assuming between three and five additional customer support representatives are needed to handle spam questions and complaints, at fully burdened salaries of $50,000 per year, the total costs were between $150,000 and

$250,000. Furthermore, adding an additional system administrator to handle spam-related problems increased the cost by another $75,000. Adding a dedicated software engineer for software development resulted in an additional cost of $150,000.

Quantifying Spam: A Brightmail Update

Since the GartnerGroup study was completed, indications are that the spam threat has grown more quickly than the Internet itself has. Spam is getting more and more attention as more and more types of organizations are affected by spam. Particularly among corporations, the size of the spam threat has grown in the last few years.

Brightmail is in a unique position to report on the magnitude of the spam threat and the patterns of its growth. In a three-month period, from November 2001 through January 2002, Brightmail's networks experienced a 14% growth in email traffic, and a 46% growth in spam.

Identifying Spam: A Key Challenge

Sometimes it's quite easy to determine if a message is spam, based on the obvious "spam-like" content of a given message or the name of the sender. Many spam filters work simply by searching for the most common words and names used by spammers. However, things are rarely that easy. A definition we've been moving towards in this white paper is that spam is "unsolicited bulk, commercial, or objectionable email, often sent using stolen resources." Once we unpack this definition, it becomes clear that spam identification is problematic and requires a systematic approach, one that cannot be completely automated. For example, it takes a certain amount of research and analysis to determine whether headers have been forged and at what step in the delivery process. In addition, determining whether a message is truly "unsolicited" opens up another level of complexity, where certain qualitative decisions need to be made.

At Brightmail, a crucial part of our mission is spam identification. This section briefly summarizes how we evaluate a questionable message. In short, the following are the general questions that we ask when judging whether a message is spam.

Is there a prior relationship between the sender and the recipient?

From our perspective, determining if a message is unsolicited is the key goal. To this end, it helps to verify the existence of a prior relationship between the sender and the recipient. If you receive bulk email from a person or company that you never heard of, it is unlikely that you requested to receive the email. An analysis of the message content and of the business that's advertising can often rule out or confirm a prior relationship.

However, companies and individuals who have had relationships with the victims can still send messages unsolicited, and it's still spam.

Is there a legitimate removal option?

Another important clue is a removal option. Removal, or “opt-out,” options typically come in the form of an email address or Web site link within the content of the email. The recipient can then theoretically follow the removal instructions to cease delivery of further mailings. Research on the particulars of the removal option is necessary to distinguish spam from legitimate bulk email that recipients may have subscribed to in the past.

The presence of an effective removal option in the email message does not by itself mean the message is not spam. Some less legitimate senders of email actually sell removal response messages to spammers, who then send further unsolicited emails to these “confirmed live” email accounts. Thus, a removal option becomes yet another tool in the spammer toolbox. Further investigation is required to determine whether the removal option works and does not lead to more unsolicited email messages.

Was there an attempt to conceal header information?

The next bit of detective work involves examining the headers. The headers provide information such as the sender of the message, the recipient, the mailer that was used to send the mail, the names of the different servers that processed the message along the way, and so on. Header information provides a good summary of the path that a piece of mail took. Unfortunately, spammers can forge header information easily; it is a trivial matter to insert arbitrary information to cover their tracks. The last thing that many spammers want is to reveal their identities and whereabouts.

Various tools allow you to trace the paths of messages. Detailed and systematic analysis of the headers is often necessary to sort out what doesn't make sense and spot the inconsistencies or impossibilities.

Was the message sent by bulk methods?

Another important clue is whether the email message is bulk email. Were multiple copies of this email message sent? If you received multiple copies of the same message, this often indicates that the message was delivered in bulk by an automated tool. In most cases, however, it is difficult to tell, as you receive only one copy, and cannot know if one or one million such messages were sent. Brightmail's unique architecture, however, enables Brightmail staff to quickly verify if certain messages were delivered in bulk. For more information, see “The Brightmail Probe Network” on page 17.

Bulk mail delivery alone does not identify a message as spam, as many legitimate, solicited email messages are sent in bulk. However, it is a worthwhile clue to consider.

What is the content of the email

Finally, the content of an email may provide clues to whether it was unsolicited. Regarding content, the first giveaway is usually the curious grammar and word choice that spammers seem to employ. Certain patterns, such as liberal amounts of ALL CAPS and multiple exclamation (!!!) points are often favored by spammers.

As we discussed earlier, spam as advertising attracts certain businesses more than others. Many of these common types of email, such as multilevel marketing

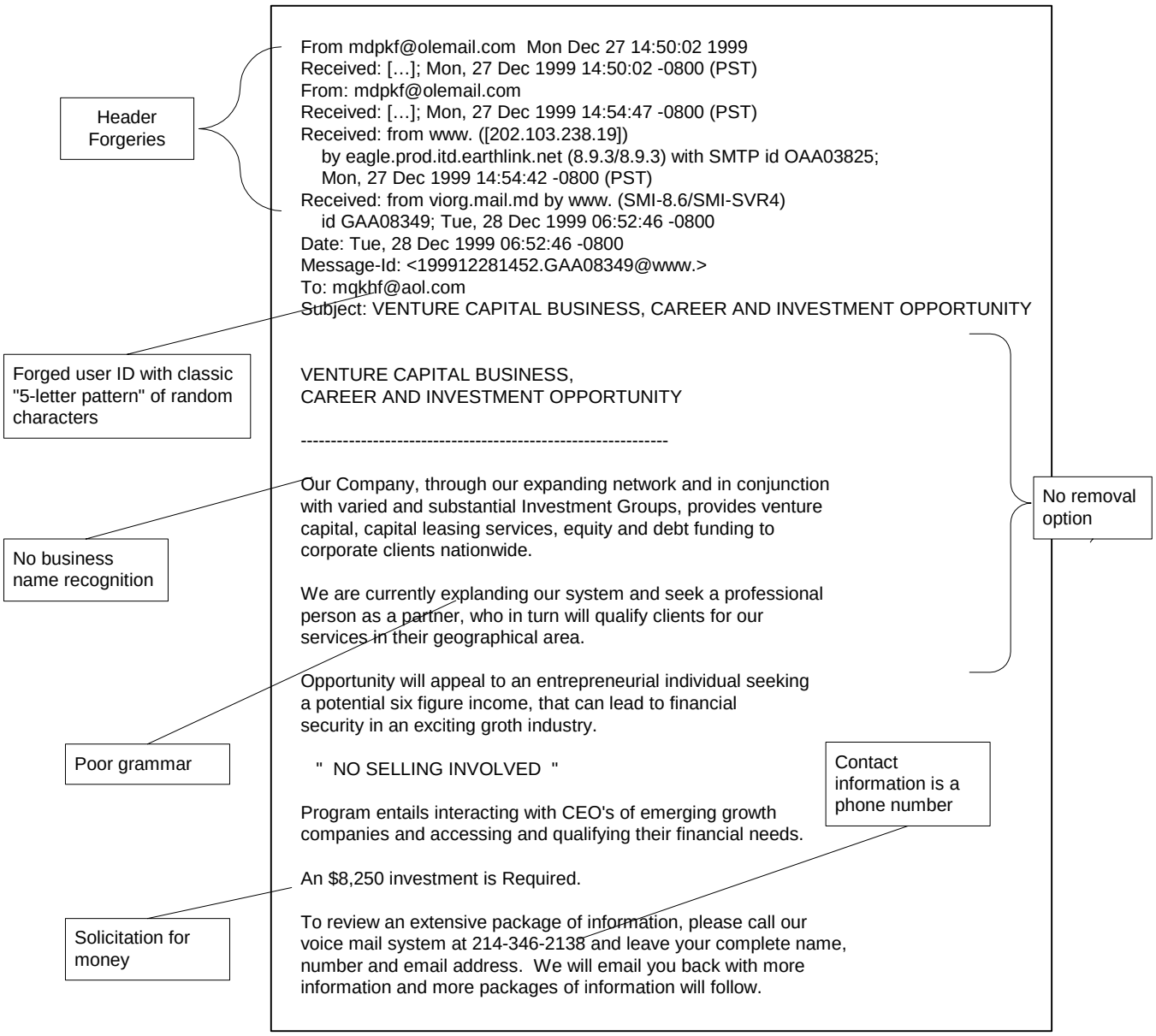
messages, are often less than fully legal attempts to involve consumers in schemes that may be completely fraudulent. An analysis of the content involves verifying whether the email message:

- Advertises something for sale
- Offers money-making opportunities
- Advertises pornographic web sites or products
- Contains offensive material
- Otherwise follows patterns typical of many other already-identified spam messages
- Contains or attaches suspicious software code

Another thing to check is whether the apparent “legitimate business” referred to in the content of the message has a Web-based email account while advertising a business domain within the message body.

The clues discussed in this section are just some of those used by the Brightmail staff to determine whether a message is spam. No single clue alone causes Brightmail to treat a piece of email as spam. If Brightmail could feasibly ask each user whether they had requested the email, we would not need to use these indirect clues. A certain amount of judgment is required to judge whether a specific email message is unsolicited and whether it's spam. Figure 6, “Anatomy of a Spam Message” shows the basic parts of a spam message, illustrating the discussion in this section.

Figure 6. Anatomy of a Spam Message



The Brightmail Solution

This section provides a brief overview of traditional spam-fighting methods, and then describes Brightmail Anti-Spam, part of the Brightmail Solution Suite.

Traditional Anti-Spam Methods

The most logical and practical places to filter email for spam are in the mail user agent (MUA) or mail transfer agent (MTA), but the two are by no means equally effective. MUAs are the client applications that allow users to retrieve and send mail from their computers. Common MUAs include Netscape Messenger, Microsoft Outlook, and Eudora. MTAs are like post offices; they are programs that reside on mail servers and are responsible for routing and sometimes delivering mail. MTA and MUA-based filtering is usually based on the header information, the mailer type, or the IP address or domain name of the sender.

To filter at the MUA level requires that email users explicitly create anti-spam filters on their machines. This approach has a number of shortcomings. First, the onus of the anti-spam work is placed on the recipient. This is not only time-consuming, but also largely ineffective. Email users typically do not have the expertise to create effective filters, nor do they have access to the most current spam. Filters based on past spam will generally be ineffective in blocking current spam, as spammers constantly change their messages to avoid such filters.

Any attempt to combat the flood of spam must itself leverage the power of the same networks that the spammers exploit, and must operate in real time, around the clock. The filters users create quickly become outdated because they are fighting yesterday’s spam. In addition, by the time spam hits the user’s machine, much of the damage is already done, in terms of storage costs on the mail server.

Filtering in the MTA, on the other hand, is often accomplished by adding rules to the configuration for the specific mail system running on the server. MTA level filtering is more effective than MUA filtering because it enables filtering for a larger number of mail accounts from a central point for administration. The drawback in this case is that users need to provide spam messages and other information to the email administrators so that current information can be incorporated into an organization-wide filtering list. This method requires continuous maintenance to keep the filter list current and effective, because it is built in reaction to spamming activity. The filters are drawn from only one ISP, and lack input on the types of spam circulating in the rest of the Internet. Another problem is the tendency to identify “false positives,” cases in which legitimate mail is incorrectly identified and filtered as spam. If the filter list is not made with care, or if domains are incorrectly blocked, valid email messages are discarded along with the spam.

Whether it's MUA filtering or MTA filtering, the same essential problems exist. For individual ISPs and email users, the available information about current spam attacks is limited, and the Internet represents a huge playing field. Traditional measures both block legitimate email and reduce productivity because service providers' staff and the user community need to continually devote time to fighting this problem. It's a never-ending battle because spammers' techniques and tools are always changing.

Additionally, once a particular domain is blocked, it is trivial for a spammer to obtain another one and resume spam attacks. Because persistent spammers can easily obtain new IP addresses and new domain names on a daily basis, reactive blocking and filtering is futile, like trying to hit a moving target.

How Brightmail Differs

Brightmail Anti-Spam is a server-side, Internet-wide, anti-spam solution that actively seeks out, identifies, analyzes, and ultimately diffuses spam attacks before they can overwhelm networks and irritate email users. Furthermore, it is part of the Brightmail Solution Suite, a comprehensive mailwall solution that can also be used to block viruses and other threats that arrive via email.

> **mailwall solution** — An analysis and filtering system, analogous to a firewall, that protects the integrity and security of electronic mail systems, and protects individual users, from email-borne threats. These threats can include virus invasions, spam attacks, and other content-related risks. A mailwall solution uses filters that are based on human and/or machine analysis to determine if email messages should be routed normally, sidelined, or modified.

Brightmail's mailwall solution incorporates service and software components, automated and human-directed functions to forge the best defense against spam. The main service components are the Brightmail Probe Network and the Brightmail Logistics and Operations Center (BLOC). The main software component is the Brightmail Server.

Together these components add up to a dynamic and effective solution to the spam problem, one that takes the guesswork out of spam identification.

The Brightmail Probe Network

The Probe Network is a large collection of email accounts with a statistical reach of over 100 million email addresses. The email accounts in this pool consist of those owned by Brightmail and those owned by Brightmail's Probe Partners, who include some of the largest ISPs in the world. The email accounts that are owned by Brightmail are called probe accounts.

Probe accounts are the first step in the real-time detection and analysis of spam. They attract spam. As mentioned earlier, spammers are quite resourceful in their harvesting of email addresses. Many of the probe accounts, therefore, are strategically seeded to attract and catch large quantities of spam. At Brightmail, we know where spammers go to collect email addresses, and we use this knowledge to strengthen the Probe Network. As a result, spammers never know if they are sending mail to an unsuspecting recipient or to a Brightmail probe account.

The structure of the Probe Network also provides powerful evidence that helps to judge if a message is spam. This virtual "net" of numerous accounts spread all over the Internet makes it easy for us to quickly verify that a given message was sent using bulk methods. When the same questionable message is caught by different probes, alarms go off and we can take action.

As we continue to add Probe Partners, the breadth and scope of the Brightmail Probe Network constantly increases. As the size of our pool grows, our ability to receive early warning information about current spam attacks increases exponentially.

The BLOC

When a Brightmail probe detects a possible spam attack on the Internet, the probe immediately routes the message to the Brightmail Logistics and Operations Center (BLOC), a spam-analysis center staffed round-the-clock, 365 days a year.

The BLOC consists of a dedicated team of email experts whose mission is to provide swift, accurate responses to spam threats, and pro-actively research and develop technologies that eliminate future threats. Their duties include:

- Analyzing incoming email from the Probe Network
- Developing, validating, and transmitting anti-spam rules to Brightmail Servers, Brightmail software housed in customer mail servers
- Managing and seeding the accounts in the Probe Network
- Researching spam attacks
- Collecting statistics and information to evaluate the effectiveness of customer Brightmail Servers

The experts at the BLOC are another example of what sets Brightmail apart from automated filtering schemes. As we saw earlier, certain qualitative skills are essential to accurately distinguish spam from legitimate email. Most email users won't tolerate losing legitimate mail to the fight against spam. Brightmail's extremely low false positive rate is a direct result of the incorporation of the BLOC into the anti-spam process. The BLOC serves as an intelligent buffer between the spammer and the unwilling recipient of spam.

This added intelligence, however, doesn't come at the expense of privacy. The BLOC only has access to mail addressed to the probe accounts Brightmail owns. The specialists at the BLOC have no access to email users' personal email.

In the end, the email user has final say. Email users can use a simple Web-based interface to access the mail that the Brightmail system has caught. We refer to these suspected spam messages as *gray mail*. Users can always access and choose to view their gray mail at their leisure. The difference is that they are not bombarded by spam every time they check for mail. And if users find messages in gray mail that they want to keep, they can redirect those messages to their regular inbox with one click.

The Brightmail Server

The cornerstone of the Brightmail system is the Brightmail software at the ISP, ASP, WSP, or corporate site. The key component of this software is the Brightmail Server, which integrates with mail server software. Brightmail Servers perform the actual filtering of incoming messages. When a message passes through the mail system, the Brightmail Server acts upon the message. Using updated anti-spam rules transmitted from the BLOC, the Brightmail Server checks the headers, contents, and other information in each message and identifies gray mail (suspected spam). The gray mail is routed to a special storage area where email users can review it.

Summary: The 3-Step Process

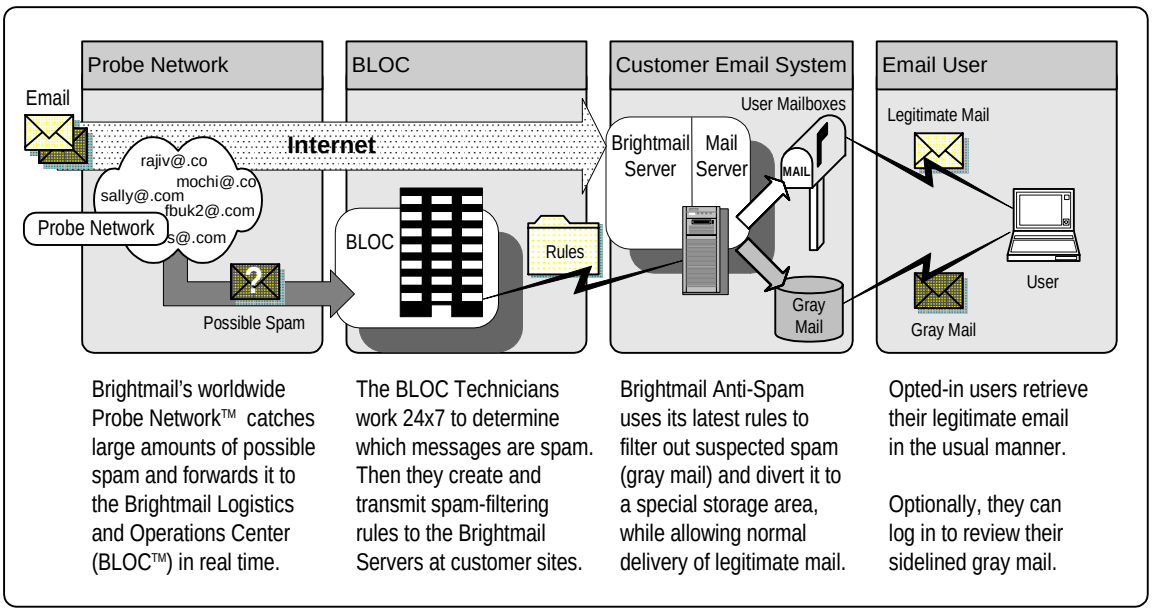
Brightmail's unique anti-spam solution can be summarized in three steps:

1. **We Find Spam** — First, we actively seek spam using our Probe Network, the extensive array of dedicated email accounts with a statistical reach of 100 million Internet addresses. We use these accounts to attract spam.
2. **We Identify Spam —** When the Probe Network finds possible spam, it forwards that email to the BLOC. There, spam experts verify that the email is spam and write rules to block it. They send those rules to Brightmail Servers.
3. **We Stop Spam** — Using updated rules from the BLOC, Brightmail Servers identify and filter spam messages from incoming email. Gray mail is diverted to a special storage area, where users can check caught spam at their leisure.

Mail Flow with Brightmail Anti-Spam

Figure 7, "Mail Flow with Brightmail Anti-Spam" shows how Brightmail software is typically integrated with mail servers at ISPs and corporations.

Figure 7. Mail Flow with Brightmail Anti-Spam



For more information on Brightmail's solutions for corporations, ISPs, ASPs, and WSPs, see http://www.brightmail.com.

Conclusion

Used responsibly, email is a powerful and invaluable tool for business, marketing, and everyday communication. The related spam problem is really part of a larger social phenomenon, one that is accompanying the maturation of the Internet. How people deal, for example, with a medium that allows people to send thousands of messages with the same effort and expense it takes to send one, will be unpredictable for some time. But as this white paper attempted to highlight and quantify, the costs associated with email abuse are real, documented, and on the whole, staggering.

Education, increased public awareness, and legal measures are certainly useful approaches to controlling spam. However, owing to the dynamic nature of the Internet and the spam problem, a true solution is one that is just as dynamic, one that instantly responds to changes and new scenarios. Technical solutions such as Brightmail Anti-Spam provide a more immediate, adaptable, and effective response to the spam problem. With Brightmail Anti-Spam, we can move towards a more responsible stage in the evolution of email and the Internet, where email users can truly control and trust what comes into their inbox.

About Brightmail

Brightmail provides advanced message management and content-filtering solutions for xSPs, mobile network operators, and enterprise companies, protecting more than 40 million mailboxes worldwide. The Brightmail Solution Suite is a comprehensive mailwall solution that detects and filters spam and viruses from the message flow and requires no action by the end-user. The Brightmail Logistics and Operations Center (BLOC) continuously analyzes millions of email messages each day, deploying real-time abuse countermeasures to customer messaging systems. Brightmail customers include Critical Path, Earthlink, MSN, and many others. For more information about Brightmail, visit the company's web site at http://www.brightmail.com.

Glossary

ASP – Application service provider. Third-party entities that manage and distribute software-based services and solutions to customers across a wide area network from a central data center.

BLOC™ – *See Brightmail Logistics and Operations Center.*

Brightmail™ Anti-Spam – Brightmail's system for spam detection and filtering. This includes the *Brightmail Probe Network*, the *BLOC*, the *Rule Base*, and the *Brightmail Anti-Spam Software*.

Brightmail™ Anti-Virus – Brightmail's system for anti-virus detection and filtering. For virus detection and cleaning, Brightmail and Symantec Corporation have teamed up to provide the most comprehensive protection in the industry.

Brightmail Logistics and Operations Center (BLOC™) – When a new spam attack is detected via the Probe Network, specialists in the 24x7 operation center (BLOC) generate new rules to detect and catch the spam, and distribute them to all Brightmail Servers at customer sites.

Brightmail Server – The Brightmail Server, integrated into the customer mail server, uses the installed filtering modules and active rules to filter suspected spam out of incoming mail and divert, reject, discard or disinfect it.

Brightmail™ Solution Suite – Brightmail's industry-leading mailwall solution, which includes Brightmail Anti-Spam and can include other products as well.

Email – Electronic mail.

Filtering Module – The module in a *Brightmail Server* that checks incoming mail against an active *Rule Set* to determine which messages are suspected to be spam. Each Brightmail Server may be configured to use one or more filtering modules.

Gray mail – Messages classified as suspected spam by *Brightmail Anti-Spam*. There are many definitions of *spam*, or junk email. For this reason *Brightmail Anti-Spam* never labels any user's mail as "spam."

Header – First part of an email message, containing information such as the address of the recipient, the address of the sender, message type, routing, and time sent.

ISP – Internet Service Provider. A company that specializes in providing connections to the Internet, including Web access and email accounts.

Mail Clients – Also known as MUAs (mail user agents.) Programs like the Netscape mail reader and Eudora client that enable end recipients to view and edit their mail messages and folders.

Mailwall Solution – An analysis and filtering system, analogous to a firewall, that protects the integrity and security of electronic mail systems, and protects individual users from email-borne threats. These threats can include virus invasions, junk email attacks, and other content-related risks. A mailwall solution uses filters that are based on human and/or machine analysis to determine if email messages should be routed normally, sidelined, or modified. The Brightmail Solution Suite is a mailwall.

MDA – Message Delivery Agent, a general term for a program that delivers mail.

MTA – Mail Transfer Agent, a generic term for programs such as Sendmail or qmail that send and receive mail between servers.

MUA – Mail User Agent, *see Mail Clients*.

Opt-in – An action by which end users choose to have their email checked for spam.

Probe Accounts – Email addresses assigned to Brightmail by our Probe Network Partners, and used by Brightmail Anti-Spam to detect spam.

Probe Network – The entire installed base of email accounts owned by Brightmail's Probe Network Partners. Used by Brightmail Anti-Spam for the detection of *spam*, the Probe Network has a statistical reach of over 100 million email addresses.

Probe Network Partners – ISPs or corporations that participate in the *Brightmail Probe Network*.

Reporter – A client process of the Brightmail Server. It aggregates statistics collected by the *Brightmail Server* and sends them to the BLOC.

Rule – An instruction that a filtering Module uses to determine whether a message is spam, or contains one or more viruses. Rules are written at the BLOC on the basis of information gathered from the Brightmail Probe Network, and then distributed to all Brightmail Servers.

SMTP – Simple Mail Transfer Protocol, a server-to-server mail transfer protocol used by many mail systems, such as Sendmail. It is based on TCP/IP.

Spam – Unwanted, unsolicited commercial bulk email. Brightmail Anti-Spam uses the term *gray mail* to identify email messages that are suspected to be spam, according to its filtering rules.

Spammer – A person or company that sends spam. Many spammers are professionals who produce mass e-mailings for companies as a service.

Suspected Spam – See *gray mail*.

Trojan Horse – A destructive program disguised as a game, utility, or application. When run, the Trojan horse does something harmful to the computer system while appearing to do something useful.