Interested in learning more?

Check out the list of upcoming events offering
"Security Essentials: Network, Endpoint, and Cloud (Security 401)"
at http://www.giac.org/registration/gsec

Technologies to Combat Spam

Thomas A. Knox
GIAC Security Essentials Certification (GSEC) Practical Assignment
Version 1.4b
Option 1
June 16, 2003

Abstract

Spam is an ever-increasing issue for anyone with an email account.   For the enterprise it leads to lost productivity, storage issues, bandwidth constraints, virus and malware intrusions, and quite possibly legal concerns.   In this paper I will give you some background on spam and its proliferation over the last few

years, and some of the issues that spam creates. I will explain the different technologies that are available to identify and remove spam, and lastly I will give you information about product offerings from vendors that you may choose to use in your fight against spam. It is not within the scope of this paper to recommend one technology or vendor over another. It is also not in the scope of this paper to explain the intricacies of email systems.   My intent is to give you the information about what is available to help eliminate spam in order to help you make a decision as to what best suits your particular situation.

Why is Spam an issue?

Spam has been seen for quite some time now and could be considered the junk mail of the 21st century. It is growing at an alarming rate.   Numbers vary depending on where you get them, but the percentage of emails that are spam appears to have quadrupled between 2001 and 2003 and now accounts for somewhere around 40% of all email. It is also expected that spam will increase to be over 50% of all email sent by the end of 2003. It can be very benign in nature like the advertisements for low mortgage rates or sales on the latest electronic devices. It can be offensive like advertisements for libido increasing drugs or pornographic websites. It can also be hostile and contain viruses, Trojan Horses, or other malware.

In the case of the benign spam it is more of a nuisance, but in sufficient volume it can present problems affecting productivity, bandwidth, and storage.

> It is clear that as spam rises the value of email as a business tool within corporate institutions will diminish. Assuming 10% of total mail is spam and each employee spends 30 seconds/day deleting that spam, the estimated annual cost of spam to 10,000-person company is $675,000. (Brightmail, p. 8)

This is assuming that only 10% of email received is spam. If you bump this number up to 40% the costs involved also increase. Please keep in mind that these costs are purely derived from lost productivity and do not include the costs to increase storage capacity nor the need to purchase more bandwidth to keep network traffic flowing.

The offensive spam may affect different people in different ways. Some may ignore it, while others may be deeply offended by it. Employers could be held liable when an employee sues based on a hostile work environment, if the company was aware of the issue and has not acted on it. Since spam originates from outside of your company consider it in the same venue as a vendor or client harassing one of your employees. If you are aware that it is occurring, you are responsible to take steps to remove it.

Employers face serious penalties if they don't remove such things from the working environment. People who have been subjected to harmful work settings can sue for up to $300,000 in compensatory and punitive damages, if the company has more than 500 employees. Damages are scaled back to $200,000 if the company has between 200 and 500 employees, with lower fees paid out by still smaller companies. If an employee leaves because of an environment judged hostile, they can ask for reinstatement, back pay and back benefits. (McCullagh p. 2)

In my opinion the most significant risk is the spam that would be considered hostile. These messages may contain viruses, Trojan Horses, worms, and web bugs among other things. The senders may try to fool recipients into thinking the emails are safe or from a trusted source by using names gleaned from an address book.   Without the proper precautions in place (virus and spam protection) this malware can spread like wildfire in an enterprise environment and bring messaging and network infrastructure to it's knees.   Take for example a Microsoft Exchange messaging infrastructure with 5000 employees. Introduce one worm on one workstation and it begins sending itself to all 5000 employees in the global address list.   A few more of the worms get installed on other workstations and start replicating in the same manner. In a very short time the messaging load can clog messaging queues and network segments leading to slow network response or a DOS (Denial of Service). Server storage space may be depleted resulting in legitimate email being lost or returned as undeliverable.

Technologies used to identify and remove spam

Vendors employ many different technologies in their products to catch spam. Some solutions may only employ one of the following technologies while others may employ some or all of them. You will need to decide what best fits your needs and also what fits within your current security policies as well as any legal requirements that you need to follow.

Blacklists/Whitelists

A blacklist is essentially a listing of domain names, mail servers, or specific email accounts that mail will not be allowed from. When using a blacklist incoming mail is checked against the list to see if it meets any of the criteria on the list and if it does, it is blocked, deleted, or whatever you have chosen to do with it. Blacklists can be created by you, or your spam solution vendor.   There are numerous blacklists available on the Internet that you can use freely or pay a fee to subscribe to.

Whitelists are the opposite of blacklists. They contain email addresses that are allowed through even if the email matches all of the other criteria of being spam. The whitelist allows you to make sure that mail from specific addresses or domains will get through no matter what. If you are using other technologies to stop spam and want to make sure that email from your business partners, vendors, clients, etc isn't blocked then you can use a whitelist. Many spam solutions will give your end users the ability to create their own whitelists so they can configure who they want to be able to receive mail from.

Integrity Check

Mail can be checked to see if it has the characteristics of spam. The headers can be analyzed to see if they make sense. When mail is sent the From: header can be anything that you want, but when the mail server receives the mail it can check to see if where it is receiving mail from matches what is being reported by the headers.

Heuristics

Heuristics is a difficult concept to explain, but it generally means to apply knowledge gained previously to a new problem. In order to fool a spam solution that uses keyword filtering someone might take the word "badword" and change it somewhat to be b-a-d-w-o-r-d. Since this obfuscation might not be in the list of keywords the spam solution may not trigger on this word. Heuristics would come into play here and allow the spam solution to see past the literal typing of the word and see it for what it is.

Content/keyword filtering

Email can be checked for particular keywords. These keywords can be in any part of the email, the header, the subject line, or the body of the email text. Mail can then be blocked if it contains specific words or combinations of words. Email can also be filtered for specific content. You may wish to block emails that contain executable files (ending in .exe, .com, or .bat for example). You may also wish to block emails that contain attachments with extensions that are commonly associated with viruses, trojans, and other malware (examples would include .pif, .scr, .vbs, etc.) For a listing of other extensions that could be used to deliver malware please see http://www.cknow.com/vtutor/vtextensions.htm. Depending on your business needs it may be possible to block all file attachments.

<u>Reverse DNS lookup</u>

When an email is sent from one server to another a TCP/IP connection is made between the two servers. The mail server that is receiving the email can take the IP address of the sending server and do a DNS lookup on that address to see if it matches what is in the header information of the email. This is a means of finding out if the sender is attempting to spoof where the mail is actually originating.

<u>What are my options?</u>

Now that you know how prevalent spam is and what issues can be attributed to it, what do you do about it? There are many vendors supplying many different products to control spam, but they can be broken down into a few major types.

<u>Outsource</u>

One option would be to outsource your spam protection. This solution may be a good choice for small and medium sized businesses. There are no hardware or software investments and you won't need to assign responsibility for upkeep and monitoring to internal staff. The vendors will typically also be using the latest and greatest hardware and software, which can be very expensive to purchase yourself. You pay a monthly or yearly fee to a vendor such as MessageLabs (http://www.messagelabs.com) and they do the monitoring for you. The way it works is relatively simple. Mail destined for your domain is re-routed to the vendor instead. There the mail is checked for viruses, spam, and offensive material. If the mail is identified as any of these it can be dealt with in whatever manner you choose. It can be flagged and sent on to the recipient. It can be sent to a different recipient (possibly a designated mailbox that you could use to recover the mail if someone indeed wanted it). Lastly it could be deleted. If nothing is found to be wrong with the email it is sent on to your mail system. One of the wonderful things about using the outsourcer would be that none of the mail ever reaches your network until it has already been screened. This eliminates the bandwidth consumption associated with the spam. This solution is also relatively simple and quick to set up. All that really needs to be done is pick a vendor, have your inbound mail routed to them, and choose what you want done with the mail that is found to be questionable. This option also relieves your IT staff from having to deal with updating and maintaining the solution. The downside to using such a provider is that your mail is flowing through a 3rd party and is somewhat out of your control. If your email contains sensitive or confidential information this is a risk that you will need to consider.

Two other solutions would be considered "in-house" or ones that would be implemented within the confines of your network: Network Appliances and Software Solutions.

Network Appliance

The network appliance is placed on your network to scan your inbound mail before sending it on to your messaging infrastructure. There are many vendors of these appliances, but let's take a look at one from Ciphertrust (http://www.ciphertrust.com/).

The appliance is called IronMail and offers a myriad of features. It is an "inline" appliance meaning that you install it and route your mail through it. It offers spam detection, content filtering, virus detection, and will also work as a mail proxy. This last feature is of interest since many companies have implemented webmail such as Microsoft Outlook Web Access or Novell GroupWise Web Access. This provides secure authenticated access to webmail servers without making the servers directly accessible to the Internet. Without such a mail proxy it is necessary to make your webmail servers accessible to the Internet. This usually means placing them in a DMZ. With IronMail it is possible to place the webmail servers inside your enterprise and proxy connections to them through the IronMail appliance, thus making them less vulnerable to attack or compromise.

IronMail uses many of the standard techniques to detect spam including whitelists, blacklists, reverse DNS lookup, heuristics, keyword content filtering, and header analysis. One thing that it does differently from the other products is to create a hash of each email and send that hash back to a Ciphertrust server where it is compared to millions of other hashes. If it finds matching hashes, there is a good probability that the email may be spam. After applying the different analysis engines to the email, it is given a score that is used to determine the likelihood that the email is spam.

You can configure the system to take different actions depending on what the score is such as quarantine the email, delete it, or send it on normally. It is possible to tune the appliance using this scoring system to find an acceptable false positive/false negative level for your needs.

The advantages of this type of solution are that you can configure and tune the spam detection to a level that works best for you particular needs, and you can also filter the problematic emails before they reach your internal messaging infrastructure. This also provides you with defense in depth by affording you another level of antivirus protection.

Disadvantages of this solution would be that spam email would still be consuming bandwidth as it flows inbound to the device. There is also a small amount of bandwidth consumed by relaying the hashes of the emails back to ciphertrust and the corresponding responses.

Software based solution

The second "in-house" option would be a software spam solution. Once again there are numerous vendors supplying spam software, but let's take a look at the solution provided by McAfee.

McAfee has a product available called Spamkiller for Microsoft Exchange Small Business that is designed to run on servers running Microsoft Exchange 2000 with less than 500 mailboxes per server. Later this summer they will be releasing a version for servers with more than 500 mailboxes. Spamkiller uses a rating system to give emails a "score" in order to decide whether an email is spam or not. Rules are applied to each email and each rule carries a related score. The scores can be negative or positive and when added together they give an overall probability of the email being spam. Negative numbers would be associated with a low probability and positive numbers with a higher probability. The higher the number the more likely it is that the email in question is spam.

Spamkiller employs the following technologies; Integrity analysis, Heuristics, Content filtering, Black and white lists, and it is also self tuning. Spamkiller learns over time and is able to adjust the overall spam scores for new emails from known senders. Spamkiller also affords you the ability to handle mail in different ways depending on the overall scores. You may choose to move emails with very high scores into a system junk folder immediately while emails with lower scores may be routed to a users junk mail folder. The whitelist automatically synchronizes with a users contact list in outlook allowing mail from anyone in the users contact list. Additions and subtractions from the users contact list are automatically synchronized also. Rules can be customized and you can also write your own rules.

The advantages of this solution are that no hardware infrastructure changes need to be made. It can be installed right on your Exchange 2000 server. There is no need to re-route mail through a third party or a hardware appliance. It is also typically less expensive than an appliance. You also have complete control and your mail is not being routed through a third party as in the case of outsourcing.

The disadvantages are that the mail is not checked until it is at your exchange server. If your mail is routing through a sendmail server or something similar between your perimeter and the exchange server these systems will still be

handling the spam email. Spamkiller also doesn't contain any virus scanning like the outsourced solutions and appliances. This means that it doesn't add that extra layer of defense against malware.

New and emerging technology

There is a new technology that is becoming available that I want to touch on. It is called Challenge/Response or "CR" for short. The concept behind it is rather straightforward. If you want to stop spam make sure that there is a person sending you the email and make them prove it. The way that it works is this. Joe wants to send John an email. Joe creates the email and sends it to John. Before the email is delivered to John's inbox the mail server sends an email back to Joe asking him to enter a code in a box (or something similar, there are different ways to accomplish the same thing) and submit this back to the email server. What this accomplishes is in order for the email to be delivered there has to be that human interaction that a bulk emailer would find too time consuming to pursue. Emails that are sent out by an automated process wouldn't be delivered because they wouldn't be able to accomplish the "puzzle" of filling in the blank. The following screenshot shows what a Challenge email might look like:
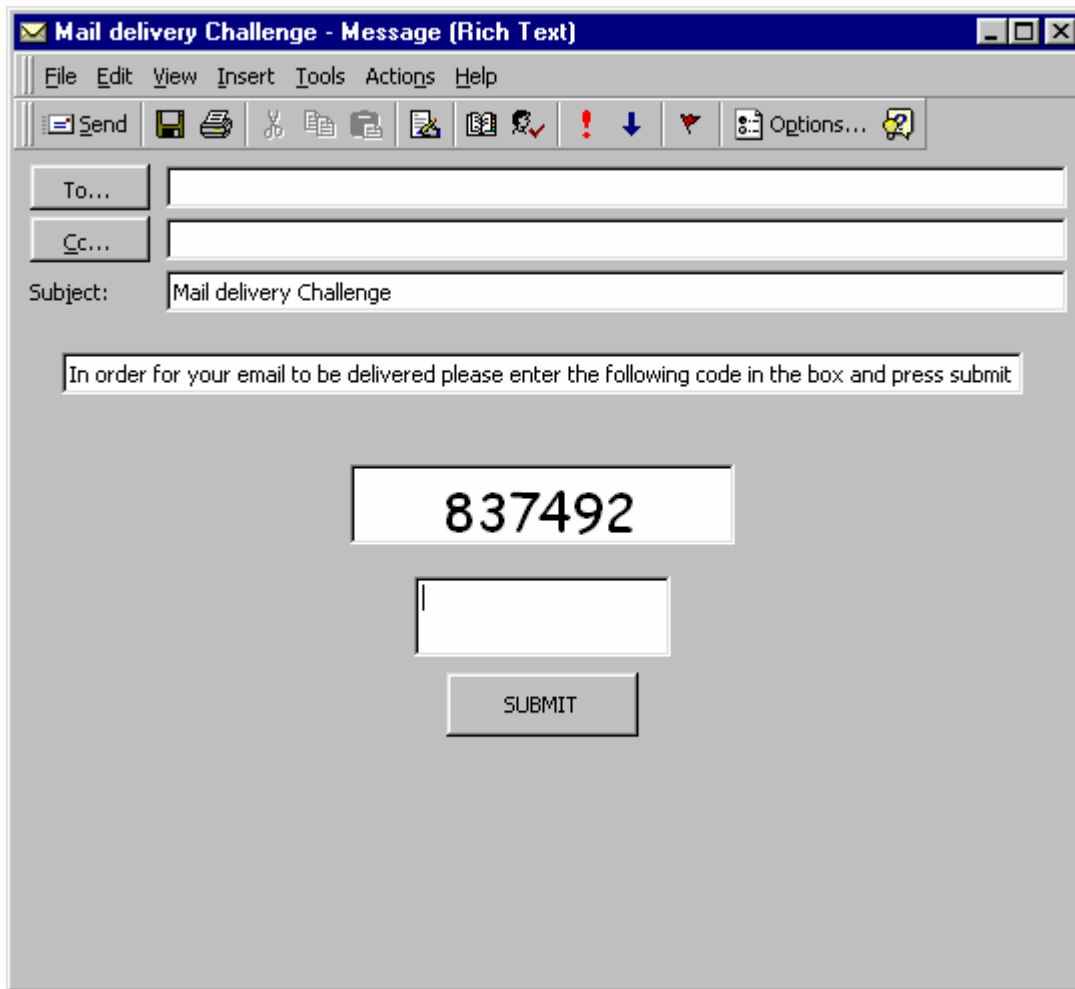
There are many differing opinions on the Challenge/Response concept. It definitely creates some overhead on a mailsystem, and also makes it more difficult for legitimate individuals to send an email to you. On the positive side it would cut down drastically the amount of spam that you would receive.

Conclusion

Spam is becoming an ever-increasing burden on employee time, network resources, system administrators, and HR departments. It has become the delivery method of choice for viruses, Trojan Horses, and other malware. If you are not doing anything currently to control the amount of spam getting into your mail system you will need to in the near future. In this paper I have given you information about technologies that are available today that you can employ in the battle against spam. As in any technology each has it's own pros and cons and you will need to weigh those in making your decision about which to employ. Whether you decide to go with a solution that you manage yourself or outsource your spam protection everyone in your organization should be happy that you've put something in place to ease the burden of unsolicited email.

References

McCullagh, Declan. CNET News.com. "Porn Spam—Legal minefield for employers. April 7, 2003. URL: http://rss.com.com/2100-1032-995658.html (May 20, 2003).

Brightmail. "The State of Spam: Impact & Solutions". January 2003. URL: http://www.brightmail.com/press/state_of_spam.pdf (May 20, 2003).

Krim, Jonathan. Washington Post. "Spam's Cost To Business Escalates. March 13, 2003. URL: http://www.washingtonpost.com/ac2/wp-dyn?pagename=article&node=&contentId=A17754-2003Mar12&notFound=true (May 21, 2003)

http://www.messagelabs.com

http://www.ciphertrust.com

http://spamcop.net/

http://www.stopspam.org/email/email.html

http://www.cknow.com/vtutor/vtextensions.htm

http://www.networkassociates.com/us/products/mcafee/antivirus/antispam/spk_ms_exchange_smallbusiness.htm