PROTECTING EXCHANGE ENVIRONMENTS FROM SPAM

Abstract

“Two years from now, spam will be solved." So Bill Gates confidently informed members of the World Economic Forum in 2004. Unfortunately, he was wrong - spam has become a global pandemic that will cost businesses $100 billion in 2007.

Spam's constantly evolving and mutating nature makes it extremely difficult to separate the good e-mail from the bad. This means that solutions which rely on static detection methods often fail to block an unacceptably high percentage of junk e-mails - or misidentify and incorrectly block an unacceptably high percentage of non-spam e-mails.

Like other mail servers, Microsoft Exchange Server, today’s dominant corporate e-mail platform, includes mechanisms to combat spam, but even these are not completely effective and must be supplemented with a third party solution if a business is to be able to eradicate spam from its network.

This paper will discuss the shortcomings of some established filtering methods and outline how SPAMfighter's community-based approach to filtering can result in it performing much better than competing products and delivering a far superior return on investment.

SPAMfighter
Nattergalevej 6, 2
2400 Copenhagen NV
Denmark

Contents

Introduction

In 2003, spam cost businesses $10 billion. By 2005, the cost had risen to $50 billion. In 2007, Ferris Research estimates that spam will cost business $100 billion. Why is spam costing businesses so much more than in previous years? Simply because there's so much more spam than in previous years. In 2001, about 10% of the e-mails sent were spam. Today, more than 50% of the 50 billion e-mails sent each day are spam. Spam is no longer simply a minor annoyance - it's become a major problem which costs businesses almost $2 billion per week. And the situation is unlikely to improve any time soon. Government and industry efforts to find a solution to the problem have had next to no impact – the spammers simply keep on spamming.

Protecting e-mail systems from the exponentially increasing volume of spam has become a job which is ever more important. E-mail has become a mission critical communication channel to a constantly escalating number of businesses. In 2005, there were about 675 million business e-mail users, but it's predicted that that number will have increased to about 935 million by 2010 – and most of those users rely on e-mail each and every day to send and receive vital business-related communications. E-mail is not a luxury, it's a business necessity and any interruption to service can be both extremely inconvenient and extremely costly.

While protecting e-mail systems has become increasingly important, it's also become increasingly difficult. Spammers need their e-mails to reach people and so are constantly looking for new ways to get their messages past spam filters and onto desktops. Unfortunately, some filtering methods are simply not sufficiently adaptive to be able to cope with spam's constantly evolving form and so fail to block an unacceptably high percentage of junk e-mail. And what's the point in a spam filter that doesn't block spam?

Most modern mail servers include mechanisms to block and filter spam. Exchange Server, with approximately 100 million seats, is the most widely used mail server today and deploys a variety of techniques to stop spam from reaching end user desktops. But, because it's the most widely used server, it's also the server which spammers focus most on attempting to beat: if they manage to get their spam past Exchange's filters, they'll be able to get it onto 100 million desktops. And, for reasons that will be outlined later in this document, spammers are often able to do just that. Therefore, businesses that wish to block a high percentage of spam cannot simply rely on their mail server's built-in filters.

This paper will explain why businesses need to block spam, why some methods of filtering do not work and how SPAMfighter's community-based approach to filtering enables it to detect spam more accurately and deliver a much better return on investment than other products.

Why businesses need to stop spam

Why do businesses need to stop spam? To save money, that's why. The cost impact of spam can be considerable:

- **Lost productivity**

  Users need to examine their e-mail and sort the good (the ham) from the bad (the spam). That process takes time - and time is money. Should an employee receive 10 spams per day and spend 20 seconds on each, about 20 hours will be lost to spam during the course of a year. In a business with 100 employees who earn an average of $30 an hour, that would translate to an annual cost of $60,000. Spam also results in calls to the Help Desk – and, of course, each of those calls costs time and money.

» **IT costs**

Each spam consumes bandwidth and disk storage – and businesses foot the bill for both commodities. While the cost associated with the transmission and storage of each spam will be extremely small, the cost of transmitting and storing a high volume of spams will be substantial.

» **Sundry costs**

Spam results in some additional costs which are almost impossible to quantify. Users will take up Help Desk time by calling for advice as to how to deal with spam. Phishing scams may result in sensitive information being disclosed. Spam containing viruses or other forms of malware may result in expensive system maintenance or downtime. A business which allows its staff to be exposed to offensive material could be held legally liable.

Spam continues to cost businesses money even once an anti-spam solution has been installed. Firstly, it costs money to license an anti-spam solution – and most solutions do not come cheap. Installing, updating and maintaining the solution add to that cost. Finally, there is the cost associated with misidentified e-mails – ham misidentified as spam or spam misidentified as ham. Such errors can be extremely costly and can quickly erode the return on investment (ROI) which the solution was expected to deliver.

» **Spam misidentified as ham (false negatives)**

Each and every spam that a filter fails to block will incur the costs outlined previously: productivity, bandwidth and storage will all be impacted and the infrastructure will be exposed to threats such as viruses and malware. Should a solution allow too many spams to slip through the net, its value will be substantially diminished.

» **Ham misidentified as spam (false positives)**

According to Ferris Research, it costs about $3.50 of an employee's time to recover an e-mail that has been erroneously deleted or quarantined by a spam filter – so just a few false positives per employee per month could result in a substantial cost.

The main risk associated with false positives, however, is that of missed opportunity – and a recent court case highlighted just how costly such missed opportunities can be. A Colorado-based law firm adjusted the settings on their spam filter in order to block junk mails that had been slipping through the net. The adjustment resulted in not only the spam being blocked, but also e-mails from the United States District Court for the District of Colorado – including a notification of the date of an upcoming hearing. The law firm missed the hearing and the judge ordered that they pay opposing counsel's costs – a decision which left them facing a bill of several thousand dollars.

Some anti-spam solutions exacerbate the risks associated with false positives by failing to provide users with an easy way to review suspected spam – with such solutions, suspected spam is not redirected to a special folder in users' mailboxes, but instead must be accessed via a cumbersome web interface in which users must open and delete e-mails one-by-one. This increases the temptation for users delete messages en masse without proper review – or, worse yet, acts as disincentive to them even attempting to review suspected spam.

No anti-spam product can deliver 100% accuracy, 100% of the time – any vendor that claims otherwise is simply not being honest. To minimize the chance of users overlooking or missing the inevitable false positives, solution must provide an easy and speedy review mechanism – but not all do. In any business, missed e-mails can be costly. Sales can be lost. Contracts can be lost. Customers can be lost. Who wants to deal with a business that does not respond to its e-mails? To be able to deliver the best possible ROI, an anti-spam solution must block a very high percentage of spam while only misidentifying a very small percentage of ham – but, unfortunately, that is becoming increasingly difficult to do.

Why some methods of filtering are not effective

Many people mistakenly believe that spammers are nothing more than small-time, nickel-and-dime operators – kids working out of their bedroom trying to make a few bucks commission by directing people to a website that sells counterfeit Rolexes. The reality, however, is very different. Spammers are highly organized and motivated by enormous potential profits. Complex pump-and-dump scams can net them millions of dollars (see "SEC v. Michael Saquella, a.k.a. Michael Paloma, and Lawrence Kaplan" in References). Phishing scams enable identity theft – something which cost the US economy more than $55 billion in 2006. Given the stakes, it's hardly surprising that spammers constantly look for ways to get their messages past spam filters and onto desktops - and unfortunately, it is something they have been quite successful at doing. Many methods of filtering are simply not sufficiently adaptive to be able to cope with spammers' constantly shifting tactics.

» **Bayesian filtering**

Bayesian filters use statistical probability to determine whether an e-mail is more likely to be spam than ham. Should the word "Viagra" appear in an e-mail while not being frequently used in other e-mails, then that e-mail is probably spam. Bayesian filters are self-learning. They create a database of words from known ham e-mails and known spam e-mails and use that database to calculate the probability of other messages being spam.

Bayesian filtering can be an extremely effective method of detecting spam, but it's far from perfect. Because Bayesian filters examine the words in an e-mail, they are ineffective against image-based spam - which is why spammers started embedding their messages in pictures. To combat image-based spam, some vendors updated their products with Optical Character Recognition (OCR) capabilities. The spammers then began to obfuscate the images with noise and colours to render them unreadable by OCR while still being readable by humans. While vendors attempt to find a solution, businesses are being bombarded with image-based spams that are double the size of most spam e-mails and so consume double the bandwidth and disk storage.

Bayesian filters can also be poisoned or confused by word salads – a collection of random words that is included in an e-mail for the sole purpose of confusing the filter (see References).

» **Blacklist filtering**

Blacklists are used to block e-mail from IP addresses, domains or ISP's from which spam is known to originate or e-mail that contains links to websites that are known to be spam-advertised. Such blacklists are run by various bodies, both for-profit and not-for-profit, and have been highly contentious (see "Busting Blocks: Appropriate Legal Remedies for Wrongful Inclusion in Spam Filters under U.S. Law" in References).

In addition to being contentious, blacklists are also becoming less effective. In the past, the majority of spam was routed through mail servers that had been configured to allow anybody to send e-mail through them (open mail relays) and it was a relatively easy task to keep track of these servers. Today, however, a high percentage of spam is routed through botnets: a collection of home computers that have been compromised by malware and are under the control of somebody other than the owner. Because the computers that comprise the botnets are spread across multiple networks, blocking by IP, domain or ISP has become impractical: it's simply not possible to block the spammers without also blocking the hammers.

Blocking on the basis of web addresses included in the body of a message (SURBL checking) has also become more difficult as those addresses are now often hidden within unreadable images.

Blacklist checking has become a hit and miss affair with either too little spam being blocked or too much ham being blocked (see "Blacklist Statistics Center" in References).

» **Keyword filtering**

Keyword filtering is impractical. Creating and maintaining a list of keywords is simply too time consuming and the filter can be defeated simply by changing "Viagra" to "V!agra". Furthermore, keyword filtering offers no protection from image-based spam.

» **Challenge/response filtering**

Challenge/response filters send a message to unknown senders informing them that must take some form of action (often to reply to the message) before the original e-mail will be delivered . While challenge/response filters certainly block spam, they also block or delay ham. In today's fast paced world, delaying potentially mission-critical e-mail may simply not be acceptable to many businesses. Additionally, there is the risk that a challenge may be blocked by the recipient's own spam filter or that the recipient may simply not bother responding to the challenge - either of which could result in lost sales or lost contracts.

Because no single filtering method is able to block spam with sufficient accuracy, many solutions use a combination of methods. While this certainly results in improved accuracy, it also results in increased development costs - which leads to more expensive products - and more complex solutions that have a higher overhead. For example, OCR is computationally expensive: extracting words from text is a resource-intensive process which may slow mail delivery to a point that would be unacceptable to many businesses – especially if the extracted words are then run through a Bayesian classifier. But a filter which leverages multiple detection methods can still be defeated – a spam that consists of a blurred image and which was sent via a compromised home computer with a non-blacklisted IP way well not be detected.

Unless a spam filter can accurately block an extremely high percentage of spam, it is practically worthless – as outlined previously, both misidentifications can be extremely costly.

Why your business should not rely on Exchange Server's spam filters

Most businesses understand the need to keep spam off their networks and most already use some form of spam filter – and, more often than not, they use whatever came bundled with their mail server. Unfortunately, such solutions tend to be basic or not to produce sufficiently accurate results.

Exchange Server is the dominant corporate e-mail platform and leverages a combination of SPAM detection mechanisms including content filtering, Protocol Data Analysis Gathering and blacklist checking. But, as outlined previously, problems exist with each of these mechanisms - they simply are not sufficiently adaptive to be able to accurately deal with new forms of spam. For example, Exchange Server still does not have a mechanism to combat image spam and, until Microsoft come up with a solution to the problem, administrators will need to find some other way to keep it off their networks.

Configuring Exchange's built-in filters is not an easy job. To use Microsoft's own words, "Your strategy for how to configure the anti-spam features and establish the aggressiveness of your anti-spam agent settings requires that you plan and calculate carefully. If you set all anti-spam features filters to their most aggressive levels and configure all anti-spam features to reject all suspicious messages, you are more likely to reject messages that are not spam. On the other hand, if you do not set the anti-spam filters at a sufficiently aggressive level and do not set the SCL threshold low enough, you probably won't see a reduction in the spam that enters your organization."

Configuring the filters to block most of the spam without also blocking the ham can be a hit and miss process that takes some considerable time – and that's time that administrators could be using far more productively.

To deliver the maximum ROI, an anti-spam solution must be able to accurately block spam with only minimal configuration – and that is something which in-built filters, such as Exchange's, simply cannot do.

SPAMfighter: maximum protection for your Exchange environment

The SPAMfighter Exchange Module works differently to other anti-spam solutions. SPAMfighter does not use Bayesian filtering, it does not check blacklists, it does not check for keywords, and it certainly does not challenge senders.

So, how does SPAMfighter know what's spam and what's ham? Because a global network of more than 4 million users tell it, that's how. SPAMfighter maintains a central database of the fingerprints of e-mails which its users have reported as spam. Whenever a user receives an e-mail, its fingerprint is compared to the fingerprints in the database. Should a match be found, the e-mail is classified as spam and deleted or moved to the junk mail folder (depending on what options the user or administrator has selected). Should no match be found, the e-mail is classified as ham and moved to the user's mailbox. Should a spam manage to bypass SPAMfighter, the user can add its fingerprint to the database simply by clicking the SPAMfighter button in his or her e-mail client. To avoid the possibility of, say, a newsletter being misidentified as spam on the basis of an erroneous report, SPAMfighter only classifies an e-mail as spam once it has been reported as such by a certain number of users.

Unlike solutions which guess at the probability of an e-mail being spam, SPAMfighter knows what is spam and it's therefore extremely accurate.

Like all anti-spam solutions, SPAMfighter may produce a very small number of false positives – but, unlike some solutions, the SPAMfighter Exchange Module makes it extremely easy for users to review their suspected spam. Should a message be determined to be spam, it's moved to a special mailbox folder. Users do not need to review such messages via a clunky web interface – they can do it easily and speedily from within their mail client.

SPAMfighter's community-based approach to filtering makes it extremely adaptable. SPAMfighter can block new forms of spam as soon as they appear. With some solutions, such as Exchange, administrators will need to wait for an update from the vendor before being able to keep the latest image-based spam off their network – but with SPAMfighter, it will be blocked automatically and almost instantaneously. SPAMfighter's community-based filter also makes it extremely easy to configure and manage. With SPAMfighter installed, administrators will be able to focus on more important jobs, while leaving the spam fighting to SPAMfighter.

The SPAMfighter Exchange Module can be used Exchange Server 2000, 2003 and 2007 and features full support for active/passive clusters.

To find out more about the SPAMfighter Exchange Module, please visit:
http://www.spamfighter.com/product_sem.asp.

Other versions of SPAMfighter

In addition to the Exchange Module, there are a number of other versions of SPAMfighter available:

- **SPAMfighter Hosted Spam filter**

  The most economical solution for businesses or individuals. SPAMfighter Hosted Spam filter intercepts e-mail sent to your domain, filters out the spam and then relays the ham back to your mail server. SPAMfighter Hosted Spam filter eliminates more than just spam – it also eliminates the expense of managing and maintaining an anti-spam solution.

- **SPAMfighter SMTP Anti Spam Server**

  SPAMfighter SMTP Anti Spam Server is easy to install, easy to maintain and will integrate seamlessly with the existing infrastructure.

- **SPAMfighter for Outlook, Outlook Express and Windows Mail**

  The perfect solution for both home users and small businesses, this version of SPAMfighter integrates with the mail client and is offered in Pro and Standard editions (the Standard edition is free for non-commercial use).

Each version of SPAMfighter uses the same powerful community-based filtering – and each user of SPAMfighter helps in the fight against spam.

To find out more about the SPAMfighter line of products, please visit: www.spamfighter.com.

About SPAMfighter

Headquartered in Copenhagen, Denmark, SPAMfighter is Europe’s leading anti-spam developer. SPAMfighter has a very simple mission: to eradicate spam.

SPAMfighter is a Microsoft Gold Certified Partner with products that are used and trusted by more than 4 million people worldwide.

www.spamfighter.com

References

Gates: Spam To Be Canned By 2006
http://www.cbsnews.com/stories/2004/01/24/tech/main595595.shtml

Industry Statistics (Ferris Research)
http://www.ferris.com/research-library/industry-statistics/

The Email Security Market, 2005-2010 (Ferris Research)
http://www.ferris.com/?p=310118

Porn spam could land EU firms in hot water
http://www.silicon.com/research/specialreports/thespamreport/0,39025001,39120300,00.htm

Spam filter costs lawyers their day in court
http://www.washingtonpost.com/wp-dyn/content/article/2007/07/13/AR2007071300606.htm

Pace v. United Services Automobile Association, Case No. 05-cv-01562-LTB-MJW, D. Co., 2007 U.S. Dist. LEXIS 49425, July 9, 2007
http://spamnotes.com/files/31236-29497/Pace.pdf

Microcap Stock Fraud
http://en.wikipedia.org/wiki/Microcap_stock_fraud

SEC v. Michael Saquella, a.k.a. Michael Paloma, and Lawrence Kaplan, Civil Action No. 1:07CV895 (BRP) (E.D.Va.)
http://www.sec.gov/litigation/litreleases/2007/lr20269.htm

Does Bayesian poisoning exist?
http://www.virusbtn.com/spambulletin/archive/2006/02/sb200602-poison

Word salad (Wikipedia)
http://en.wikipedia.org/wiki/Word_salad_%28computer_science%29

Busting Blocks: Appropriate Legal Remedies for Wrongful Inclusion in Spam Filters under U.S. Law
http://papers.ssrn.com/sol3/papers.cfm?abstract_id=944551

Blacklist Statistics Center
http://stats.dnsbl.com/

Can DNSBased Blacklists Keep Up with Bots?
http://www.cc.gatech.edu/~avr/publications/ramachandran-ceas06.pdf

Configuring Anti-Spam Features to Reduce the Volume of Spam (Exchange Server)
http://technet.microsoft.com/en-us/library/bb124696.aspx