Overview

Image Spam: The Email Epidemic of 2006

SECURITY TRENDS

OVERVIEW

End-users around the world are reporting an increase in spam. Much of this increase can be attributed to a resurgence of spam in 2006 — driven by the emergence of new, more sophisticated forms of image spam.

Image spam is a technique with which spammers advertise the "call to action" of their message as part of an embedded file attachment (like a .gif or .jpeg) rather than in the body of the email. These images are automatically displayed to end-users, yet the content of the image itself remains hidden from most spam filters.

The increase in more complex image spam attacks has caused spam capture rates across the email security industry to decline, resulting in wasted productivity and end-user frustration as more spam gets delivered to their inboxes. The sheer increase in the volume of spam, combined with a higher percentage of larger-sized spam, is also clogging the email infrastructure as many mail systems are unable to keep up with these spam volumes.

This document summarizes (1) the recent trend in image spam, (2) why it is difficult to detect and (3) how IronPort® protects customers from this increasing threat.

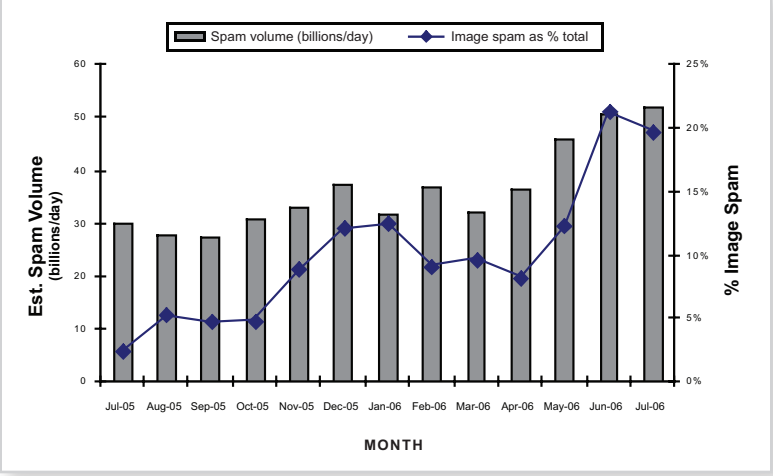
TRENDS & SOLUTIONS

According to IronPort's SenderBase® Network, spam volumes leveled off in 2005, but surged again in the second quarter of 2006. As illustrated on the left axis in the chart below, SenderBase shows that worldwide spam volumes grew from approximately 30 billion messages per day to over 50 billion over the last 12 months. IronPort saw a 40 percent increase in spam volumes during Q2 alone. This means that, even if the spam capture rate is held constant, the average end-user will have noticed 40 percent more spam in their inbox since April.

Much of this increase in overall spam volume can be attributed to the growth in image spam. As illustrated by the right-hand axis in the chart below, image spam rose from around 3 percent of spam a year ago to over 20 percent today. When overall spam volumes spiked in Q4 '05 and Q2 '06, image spam was fueling the increase.

FIGURE 1.

Fueled by a worldwide increase in image spam, overall spam volumes surged in the second quarter of 2006.



TRENDS & SOLUTIONS (CONTINUED)

The root cause behind this sharp increase in spam volumes is money. Spammers are single-minded: they send spam to make money. The more messages that are delivered to inboxes, the better the chances recipients take action on the messages, resulting in more income for spammers.

As illustrated in the next section, randomized image spam is especially difficult for most spam filters to detect — causing more of the spam to get delivered. Spammers can also make their images appear quite normal and compelling to users, resulting in higher response rates. Since neither of these factors is likely to change in the near-term, IronPort expects image spam to remain a problem for the foreseeable future. IronPort has also seen spammers innovate rapidly in their use of image spam, suggesting that image spam will soon become even more challenging to detect.

Why Image Spam Is Difficult To Detect

Image spam has been around for years. It was originally created in order to get past "heuristic" filters, which block messages containing words and phrases commonly found in spam. Since image files are in an entirely different format than the text found in an email, heuristic filters never "see" the content of the message. Therefore, these filters were easily defeated by this type of spam.

To deal with this problem, anti-spam vendors developed "fuzzy signature" technologies. These signature-based technologies collect samples of known spam and then classify "near-identical" messages as spam. These signatures were sometimes written against just the message attachment, so that messages with different content but the same attachment would still be marked as spam.

Signature-based defenses remained effective for several years. In 2006, however, spammers began randomizing images to appear the same to the human viewer but totally different to spam filters. For example, some spammers are sending messages advertising the purchase of stocks with an attached .gif file that has random "dots" inserted in the image and borders with subtly different color and width. The signatures that most anti-spam vendors rely on to detect these attacks vary dramatically, based on these small changes to the image. This means that anti-spam vendors may publish a rule that stops one instance, but this rule doesn't stop all the rest of the spam messages in the attack.

There is an almost infinite number of ways that spammers can randomize images. In addition to inserting dots, spammers have recently used techniques such as varying the colors used in an image, changing the width and pattern of the border, altering the font style, and "slicing" images down into smaller pieces (which are then reassembled to appear as a single image to the recipient). Page 3 includes two examples of the many techniques recently used by spammers to get past signature-based defenses.

EXAMPLE 1.

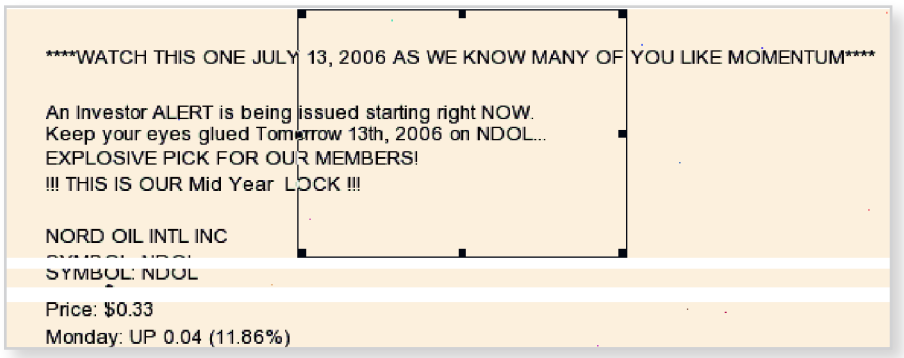
"POLKA DOTS"

An embedded .gif file containing all "text" with dots randomly inserted in the image to make every message appear unique to spam filters

EXAMPLE 2.

"SLICE & DICE"

Images are broken down into many smaller files of varying sizes and then reassembled in the mail client so as to appear as a single image to the email recipient. The rectangle highlighted represents the border of one of over a dozen image files used to construct this message. This technique is used to defeat signature-based defenses and break up words that could be found by OCR (described below).



TRENDS & SOLUTIONS (CONTINUED)

Some vendors have recently introduced Optical Character Recognition (OCR) as a means of detecting image spam. OCR is a technology used to extract typewritten text from an image. While more effective than signature-based solutions alone, OCR has several limitations. First, OCR is very computationally expensive. Fully rendering each message and then looking for word matches against different character set libraries can take as long as several seconds per message. This lowers system throughput below levels acceptable to most ISPs and enterprises. OCR is also extremely vulnerable to obfuscation. While modern OCR technology can reliably detect typed letters and numbers, it can be easily fooled by basic techniques used by spammers. For example, OCR is ineffective at detecting image spam that includes hand-written text, graphics or any abstract data.

Protecting Against Image-based Threats With IronPort Anti-Spam

IronPort Anti-Spam™ uses a unique, multi-layered approach that stops over 98 percent of image-based spam, with near-zero false-positives. The first layer of defense is powered by IronPort's Context Adaptive Scanning Engine™ (CASE). This is followed by an inner layer of image spam protection powered by IronPort's patent-pending Multidimensional Pattern Recognition™ (MPR) technology.

CONTEXT ADAPTIVE SCANNING

Most anti-spam filters depend heavily on content-analysis for stopping spam. This is like building a house on a weak foundation. These filters all share a common weakness — relying heavily on something that can easily be manipulated by spammers themselves. Image spam is just one instance where content-based filters fall short. As in the examples on page 3, the "content" of the spam is invisible to many filters because it is embedded in the image itself.

To detect image spam, IronPort has augmented traditional content-based techniques with techniques that analyze the full ***context*** in which the message was received. Specifically, CASE detects threats by analyzing four broad areas:

1. ***Who*** sent the message and what do we know about this sender?
2. ***Where*** does the call to action in the message take you?
3. ***What*** is the nature of the message content?
4. ***How*** was the message technically constructed?

Instead of generating a signature based on the content of the message, IronPort creates a specific spam profile for an image-based spam attack that combines the "who, where, what and how" of a message.

For example, one profile might be created for message that originated from a dynamic IP address, contains a certain header pattern, has an embedded image of a specific size-range and type and contains little or no text in the body of the email itself. None of these factors alone are likely to indicate with certainty that a message is spam, but they are highly accurate when combined. Context adaptive scanning allows IronPort to filter the majority of image-based spam attacks without decoding the image file. The second layer of protection is provided by Multidimensional Pattern Recognition (MPR).

TRENDS & SOLUTIONS (CONTINUED)

MULTIDIMENSIONAL PATTERN RECOGNITION

To the human eye, image spam is extremely recognizable. In fact, this is one of the properties of image spam that make it attractive to the spammer — they don't have to go to nearly the same lengths to obfuscate their content when sending image spam to avoid filtering as they do with traditional text spam. But, if this spam is so obvious to the end-user, why can't spam filters identify it?

The challenge is that humans interpret the content of messages using a much richer data set than just the text displayed. Attributes such as image color, shape, font size and type, graphics and many other characteristics also shape a reader's perception of a message. This information is entirely hidden from traditional content filters — and technologies like OCR only capture a fraction of this information.

IronPort Anti-Spam developed a patent-pending technology called Multidimensional Pattern Recognition (MPR) to address this problem. After decoding the binary image files, IronPort uses MPR to analyze the decompressed image data across over 13 dimensions to determine whether or not the message is spam.

Color is an example of a dimension that provides rich information about the content of a message. IronPort analyzes the distribution of colors found in each message to establish the likelihood that the message is spam. For example, MPR can scan a .gif file to look for pixel patterns indicating that the image file is displaying "all text" to the user, a pattern that is common in spam but rare in legitimate email (most legitimate .gif files contain pictures not text). MPR can also detect anomalous "dots" in images that don't fit the "smoother" gradients of light typically found in legitimate email (these dots may represent attempts by the spammer to defeat signatures).

To make this level of inspection possible, without compromising performance, IronPort applies the concept of "early exit". This means that the more intensive MPR process is only applied to messages with images that have already passed through the regular context adaptive scanning process. This same concept is applied within MPR as well. If part of the image file has been analyzed and there is sufficient data to determine that the message is spam, the full image file will never be analyzed. The end result is a process that is not only more accurate, but also several times faster than traditional OCR technologies. Critical to the effectiveness of this technology is the real-time nature of IronPort Anti-Spam. Updates to the system are made every five minutes, ensuring immediate and accurate protection from image-based threats.

SUMMARY

Image spam has exploded in 2006, as spammers have found it to be an effective means of bypassing traditional spam filters. The flood of image spam is frustrating end-users and taxing the already strained email infrastructures of many companies.

Spammers have rendered traditional anti-spam technologies ineffective by hiding content in embedded images and subtly randomizing these images so that each message appears unique to spam filters. Some anti-spam vendors are looking towards introducing OCR technology to stop this problem. Unfortunately, this technology is too slow for many customers and can easily be defeated by simple changes in spammer tactics.

IronPort has taken a fundamentally different approach to the problem. By interpreting image content more along the lines of how a human would interpret the image, using Multi-dimensional Pattern Recognition, IronPort has turned the spammers' own techniques against them. In their efforts to defeat traditional anti-spam systems, image spammers are leaving behind subtle traces that IronPort Anti-Spam is using to stop over 98 percent of their messages. IronPort Anti-Spam is available on IronPort's email security appliances. IronPort technology protects the infrastructures of organizations worldwide — not only from today's threats, but from those certain to evolve in the future.

CONTACT US

HOW TO GET STARTED WITH IRONPORT

IronPort sales representatives, channel partners and support engineers are ready to help you evaluate how IronPort products can make your email infrastructure secure, reliable and easier to manage. If you believe that your organization could benefit from IronPort's industry-leading products, please call 650-989-6530 or visit us on the Web at www.ironport.com/leader

IRONPORT®

IronPort Systems, Inc.
950 Elm Avenue, San Bruno, CA 94066
TEL 650.989.6500 FAX 650.989.6543
EMAIL info@ironport.com WEB www.ironport.com

IronPort Systems is the leading email and Web security products provider for organizations ranging from small businesses to the Global 2000. IronPort provides high-performance, easy-to-use and technically innovative products for those faced with the monumental task of managing and protecting their mission-critical networks from Internet threats.

IMAGE SPAM
09/06
DOC RELEASE