Mail-SeCure™ Image-Based Spam Treatment

Whitepaper

Image Spam Defense

Spammers are consistently creating sophisticated new weapons in their armed race against anti-spam technology, the most famous of which is known as *Image-Based Spam*. Spammers have used images in their messages for years, in most cases to offer a peek at a pornographic Web site, or to illustrate the effectiveness of their miracle drugs. But as more of their text-based messages started being blocked, spammers searched for new methods and realized that putting their words inside an image could frustrate text filtering. The use of other people's computers to send their bandwidth-hogging e-mail made the tactic practical. The number of unsolicited messages containing images has grown significantly throughout the years and now represents 25 to 45 percent of all junk e-mail (Spam Doubles, Finding New Ways to Deliver Itself, NYT, December 6th, 2006).
Image spam is expected to continue its growth and spreading process. Through constant monitoring, PineApp has identified that Image-Based Spam tends to be distributed in massive waves; at one of the distribution peaks, PineApp had measured Image-Based Spam as 30% of all global spam. Image-Based Spam creates bandwidth and storage problems, since typical Image-Based Spam message weighs more than three times that of a regular spam message. During Image-Spam distribution attacks' peaks, bandwidth and storage requirements increase by 70%.

Summary

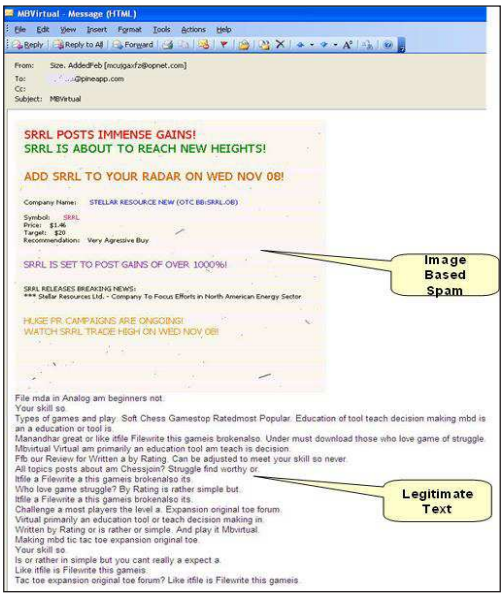
- ✓ Image-Based Spam is a new and increasing problem, leading to loss of productivity and to a burden on IT resources.
- ✓ Most Anti-Spam solutions have problems dealing with Image-Based Spam and by dealing with it ineffectively, they create other problems along the Way.
- ✓ PineApp has implemented a unique solution that enables to decode images and treat them with RPD, similarly to other types of spam.

This solution:
- Improves the already superior spam detection rate.
- Maintains low false positive rate.

Questions & Answers

What is Image-Based Spam?

Image-Based Spam is spam which contains its unwanted content inside an embedded graphic file (typically appears in GIF format, but can also appear as JPG, PNG, BMP etc.), making it difficult for some spam filters to identify. These unsolicited e-mails contain *no relevant text or hyperlinks*. The message may appear to be a text message (see example); however, it is merely an image of text. Often the content of such messages are penny stock "Pump & Dump" schemes and other malicious types of spam. Since creating Image-Based Spam requires more technical know-how than basic textual spam does, it originates in areas such as Russia, which have technically advanced spammers. Image-Based Spam has rapidly spread to the rest of the world, and is now recognized as a major global issue.



What are the Newest Trends in Image-Based Spam?

Lately, spammers have been experimenting with new techniques such as "broken images," i.e. splitting a single image into smaller images that fit together like puzzle pieces. This technique makes it even more difficult for text based anti-spam engines to detect and block.
Another noticeable technique is to send animated GIFs, with several frames of random "noise". These random pixels act similarly to the randomized images that are not animated, simply with another level of complexity. In some cases, the animated GIFs contain subliminal messages (e.g. "buy… buy… buy") embedded into frames that flash by very quickly. Animated GIF spam is much heavier, on average, than static Image-Based Spam.

Why is it so Difficult for Most Anti-Spam Engines to detect and Block Image-Based spam?

These unsolicited emails contain no text or hyperlinks, so most Anti-Spam Engines cannot detect this type of spam. Often the message will contain text copied from legitimate books, in order to fool
Bayesian filters.
Spammers have figured out a way to fool even those engines that try to analyze the image data itself: they slightly vary the images in each message. They do this easily by changing the shade of the border or background, changing the line spacing or margins, or even adding tiny specks to the background; these types of changes are invisible to the eye (or irrelevant to the reader), but they completely change the way the data appears to most Anti-Spam Engines. The result is a huge quantity of Image-Based Spam that contains random patterns with almost no repetitions.
In the past, none of the traditional Anti-Spam technologies – content-based, Bayesian, Heuristic, URL
Filtering etc. – have been able to prevent this type of spam on a consistently accurate basis.

What other Technologies are used To Fight Image-Based Spam?
Optical Character Recognition

Some Anti-Spam providers have added new image scanning features to their protection arrays, based on OCR (i.e. Optical Character Recognition, which changes graphic images of text to editable text). OCR- based anti-spam technology has several drawbacks: it requires significant resources and can reduce server performance; and it has difficulty detecting spam messages comprised of mostly images, with only a small amount of text

within the images.

Rule Engines

Other methods for blocking Image-Based Spam, such as rules that prohibit attached images of a certain size, or a certain quantity of colors, lead to an unacceptably high number of false positives. In such a scenario, legitimate email messages containing baby pictures or graduation photographs could be considered spam and thus go unseen by their intended recipients.

Recurrent Pattern Detection

Recurrent Pattern Detection contains an intrinsic mechanism to exact-match recurrent patterns across similar but not-identical messages. However, in the case of images, the minute the spammer makes even the smallest change to an image, the image-encoded data appears completely different. PineApp identified this trend in the earliest days of Image-Based Spam, and has made the necessary enhancements to its detection engine, in order to provide protection against this new threat with a sophisticated protection shield.
PineApp invested significant resources into implementing a method for decoding images and then sampling them using the proven RPD approach. The result is a significantly improved spam detection rate, whilst the same low false-positive rate is maintained.