

The Rise of Image-Based Spam

No matter how you slice it - the spam problem is getting worse. In 2004, it was sufficient to use simple scoring mechanisms to determine whether email was spam or not because it was primarily text-based. Techniques such as Heuristics (weighted scoring), Bayesian filters (probability analysis), and reputation lists (RBL's) were widely adopted and incorporated in solutions from leading vendors at the time. They also became the core techniques for open source solutions like Spam Assassin, which spammers have full access to. More recently, the sheer quantities of spam have increased by over 100%, and most of that growth is attributed to an increase in more sophisticated methods -like image-based spam.

While image-based spam has been around for years, it became much more sophisticated in 2006. Image-based spam messages initially consisted of images only, with no text, URL hyperlinks or other identifying characteristics. Because it has no text included with the message, it rendered text-centric anti-spam technology virtually useless. Making matters worse, spammers often surrounded the image with random "innocent" text so that the message could not be blocked based on a simple "image-only" filter rule. The extra text was also used to corrupt or confuse Bayesian filters.

Most spam fighters, including the open-source community, responded with two basic methods for blocking image-based spam: fingerprinting and OCR (Optical Character Recognition).

Fingerprinting identifies a specific graphic through a set of characteristics such as an MD5 checksum. However, the counter-measure to this technique was quite simple. By modifying a few pixels in the graphic (Figure 1), the fingerprint can be easily changed. By randomizing the "noise" in the image, each image fingerprint is unique and the simple fingerprint filter that is coded becomes severely compromised and ineffective against the spam.

Another simple countermeasure to image fingerprinting is to break the single image into multiple images pieced together to appear as one. This technique is effective because spammers send out the same baseline image, but slice it randomly to create unique messages with variable number of jigsaw puzzle pieces of varying size.

The second method of blocking image-based spam is OCR. OCR attempts to convert the text within the image to characters and then filter them using the traditional Bayesian and Heuristics methodologies. OCR works well under stable conditions like traditional black text on a white background, but it's easy to make an OCR algorithm confused by adding variability into the image. As illustrated in Figure 2, background colors, patterns, font size, font color, text layout and text super/subscripting are all used to randomize the images and cause the OCR algorithm to fail. If the OCR algorithm doesn't find any recognizable text, the traditional scoring filters are unable to block the spam.

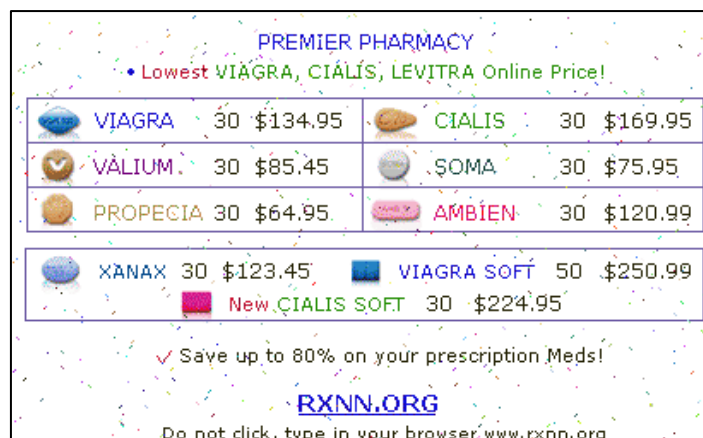


Figure 1 – "Specked" image

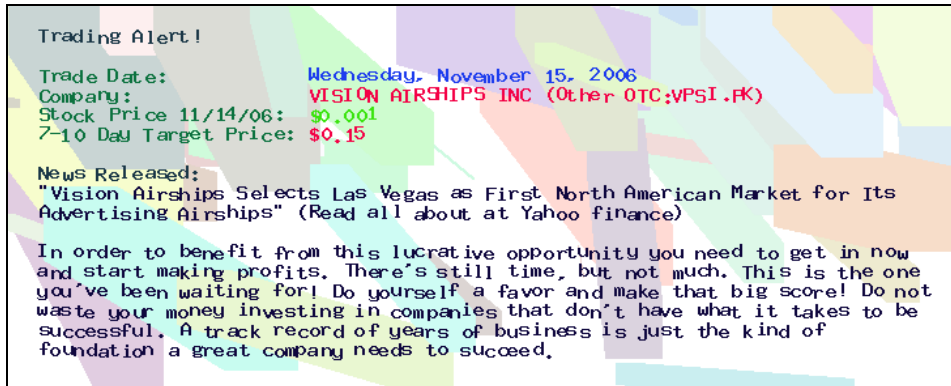


Figure 2 – Randomized image designed to circumvent OCR

More Tricks

At the end of 2006, spammers adopted spam techniques that significantly crippled fingerprinting, OCR and earlier scoring-based technologies. In addition to the almost 100% randomization of the images as described above, spammers also adopted animation techniques that hide the image's call to action.

One animated GIF technique places the “money image” within a series of frames. Each frame is randomized, and the number, animation timing, and sequence of frames are randomized within the series. Figure 3 illustrates a typical animation sequence. This technique is effective because many simple filter technologies only examine the first image in the animation sequence. Also, the animation sequence is variable, making it difficult to determine which frame contains the call to action or spam.

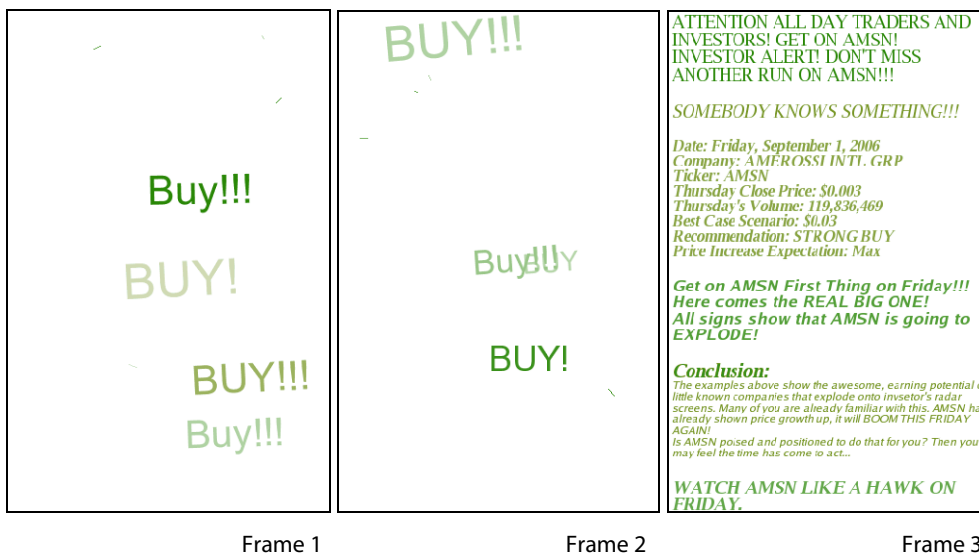


Figure 3 – Multiple image animated GIF

Another animation technique involves slicing the image into layers. When animated, these slides appear to the user as a single flat image. Again, the base image is randomized, and the slices are unique to each message, often slicing through lines of text and making it impossible to analyze using OCR.



Figure 4 – Sliced image animated GIF

The Future of Spam

New obfuscation methods are implemented relentlessly and in real-time, and early generation technologies are not able to keep up with the shape-shifting nature of the attacks. Spammers do not rely on any one technique for long, so spam fighting research continues to develop new techniques to identify spam message types that are on the horizon. In 2007, Red Condor anticipates new iterations to continue to grow.

Attachment-based spam: Spammers have discovered that embedding images in attachments gets malicious email past many spam filtering solutions, and are now exploiting this new vulnerability. Embed advertisements in attached Adobe Portable Document Format (PDF) files and zipped (compressed) Microsoft Excel documents were launched and quickly locked by filters in 2007.

Red Condor: A More Effective Solution

Red Condor's proprietary anti-spam technology keeps organizations ahead of the spammers by leveraging behavior-based reputation merit with general and advanced image analysis filtering. Behavior-based rules for each new spam campaign are custom built, and filters are modified in real-time as spammers change tactics.

Behavior-Based Reputation Scoring

Red Condor creates a merit-based reputation score for each email sender based on both sender history and message characteristics. When determining a sender's reputation score Red Condor considers such factors as:

- Has this sender been seen before?
- How much email has been received by the sender?
- What is the pattern of the sender's email history?
- Do the sender's messages exhibit specific behavior indicative of spammers?

Red Condor then uses this real-time reputation score to modulate how much email is accepted from each sender, and any interaction with this sender is immediately available to all Red Condor servers. In many cases, Red Condor can block over 90% of connections based purely on reputation data.

General Image Filtering

The Red Condor image filtering system includes both sophisticated image analysis as well as dynamic feedback of the message content. The image filters track specific components of the message. During initial message parsing, externally-linked images are downloaded for filter processing. General image filtering is applied to all images (downloaded and attached). An MD5 checksum fingerprint is calculated for the images and compared against all current fingerprint rules. The MD5 checksum is generated from the data portion of the image only, omitting the header and footer areas. This is necessary because spammers often randomize palettes and other header fields.

Advanced Image Filtering

Pre-processing

If the message passes the general images filters, advanced filtering is then applied to any attached images. Random coloring is very common so the first step of advanced image filtering is to convert the image to black-and-white. Several different approaches are used to attempt to determine the true foreground and background colors. Once the image data is available as a black-and-white bitmap, it can be efficiently processed by the advanced image filters. During the conversion process, a "de-speckling" algorithm is applied. This eliminates random dots added by the spammer to obfuscate the image.

Animated GIF Filter

When an animated gif is encountered, each of the individual frames is initially scanned and compared to determine which frame has the actual ad content. Typically there are a several randomization frames (usually sparse dot patterns) and one frame with the text of the ad. The "payload" frame is then made available for further processing.

Historically image campaigns utilizing animated GIF's have had a few frames of random garbage and one frame of content. Recently spammers have begun breaking up the "payload" image into multiple slices, each on a different animation frame. The resulting animated GIF is usually of a fairly large dimension and has 10 or more frames of animation. This makes the true image more difficult to

process but it is also a very distinctive characteristic present almost exclusively in spam. If the message has one of these images and also matches certain other criteria, it is blocked here.

Pixel Pattern Filter

Pixel Patterns are essentially cropped image areas selected from a captured spam sample. These patterns can be searched for anywhere within a target image. The Pixel Pattern filter uses advanced comparison techniques to quickly and efficiently apply all Pixel Patterns to the image. Additionally, the advanced pattern-matching algorithms allow for a certain amount of "fuzziness" to the match. This helps cope with image randomization.

Text Image Filter

The job of the Text Image filter is to determine if the image consists mainly of lines of text. This sophisticated filter is used instead of OCR to identify image-based spam and to prevent identification of non-spam images.

Red Condor does not rely on OCR or other simple image fingerprinting technologies to fight spam, and instead relies on a unique combination of techniques to get results. With spam campaigns changing every day, Red Condor uses our human-in-the-loop factor to help us combat spammers and keep organizations ahead in the game.

About Red Condor

Red Condor is revolutionizing spam fighting with its next generation technology. Red Condor's highly accurate email filter, hybrid architecture Vx Technology™, and fully managed appliances lead to a dramatic reduction in the cost of owning a premium spam filter. With solutions for small business, as well as ISPs with millions of email inboxes, Red Condor has a cost-effective, time-saving solution that is rapidly gaining market share. The system's design has built-in zero tolerance for lost email, and a near zero false-positive rate while achieving long-term spam block rates greater than 99%. This next-generation technology is backed by a 24x7 customer care center staffed by email security experts at Red Condor's headquarters. For more information, visit www.redcondor.com.

RedCondor

1300 Valley House Drive, Suite 115 Toll Free: 888-9NOSPAM
Rohnert Park, CA 94928 1-888-966-7726
info@redcondor.com
www.RedCondor.com