



SPAM WARS: EPISODE II

Spam 2.0 - SPAM WARS: EPISODE II

By Charles Taite, CTO, GWAVA

NOTE: This article is intended to be a non-technical and perhaps overly simplified description of the new wave of spam on the Internet for those of us who want to grasp the basic problem without getting too technical.

A few years ago, Bill Gates stated that the Spam problem would be solved by 2006. And if you can think back to March or April of 2006, you might have thought that Mr. Gates' claim was within reach. Sure, there was still spam, but with GWAVA or another anti-spam filter in place and properly configured, spam was probably limited to 1 or 2 messages in your inbox per day or less.

Fast-forward to December 2006, and it's clear that Act 1 of a little drama called "Spam Wars" is over. Act 2 (or Spam 2.0) has begun and it looks like the villain has the upper-hand. So in case the end of Act 1 was so boring that you fell asleep and have now woken up in the middle of Act 2 saying "what did I miss?", let me bring you up to speed: In short, the scale of spam has changed.

Worldwide spam volumes have increased by several factors since last year. Let us assume that your anti-spam system, catches 97% of your inbound spam. If you were used to receiving 2 or 3 pieces of spam in your inbox per day back in March 2006, that's because you were probably receiving 100 spam messages per day (I apologize if I'm stating the obvious: 97% catch-rate means 97 blocked and 3 reaching your inbox).

Today, if you feel that you are being "flooded" with 10 or more spam messages per day, check your anti-spam server's statistics and you'll likely find that's because you are receiving closer to 400 spam messages per day (97% catch-rate means 388 caught and 12 in your inbox).

How Has This Happened?

As I said the scales have changed. Spammers have been successfully using viruses to transform the computers they infect into a giant network of spam pumps or "Botnets" (short for Robot Networks). It is estimated that the number of infected PCs that form these Botnets number in the hundreds of thousands and growing fast! This obviously explains how a few spammers are capable of producing such a dramatic increase in spam.

Defeating Spam 2.0 requires Anti-Spam 2.0

Botnets are here to stay. From the Spammer's point of view, they are cheap and highly effective. For every Botnet PC that is discovered and cleaned or blacklisted, 2 more PCs are infected and incorporated into the Botnet. We must accept that this is a storm that shows no signs of subsiding. The elevated spam traffic of Spam 2.0 is simply the new reality. What is equally clear is that your current Anti-Spam systems (Anti-Spam 1.0) need to adapt in order to more effectively face this new challenge.

Anti-Spam 1.0 probably relied on "rule updates" or some sort of semi-automatic training to stay current with Spam 1.0. These approaches were quite effective in the past, but the problem with Spam 2.0, again, is scale. For example, with Spam 1.0, a single spammer may have been able to pump 2 Million spam messages per day. At the beginning of said day, the spammer in question crafted a clever message that would take many Anti-Spam 1.0 systems by surprise. At a rate of 83,000 spam messages per hour (that's how much you have to send to do 2 Million in a day), this new spam message starts reaching mailboxes. 4 hours and 300,000 spam messages later, the spammer winds up blacklisted, and if your anti-spam provider sends you rule updates, your updates



arrive and the issue is solved. That probably translates into 1 or 2 spam messages reaching your mailbox and the rest being successfully blocked. The unblocked spam in your inbox represents the time it takes for Blacklists and Anti-Spam 1.0 to react to the appearance of new spam. If you've ever wondered why your inbox sometimes contained spam messages similar to ones that were blocked and quarantined, you now know why.

So if we look at the same scenario with Spam 2.0, we no longer start with a 2 Million-spam-per-day capacity. Instead, this spammer's Botnet has a 10 Million-spam-per-day capacity, and instead of pumping out 83,000 spams per hour, he now has a 415,000 spam per-hour capability. Of course, in the same 4-hour period before getting blacklisted, he's delivered 1.66 Million spam messages. The sheer scale of Spam 2.0 exploits the reactionary nature Anti-Spam 1.0.

What can be done?

The answer is simple and challenging: your anti-spam systems cannot get by on 97% or 98% or even 99% catch-rates. Even at 99%, you are still looking at something like 4 or 5 spam messages in your mailbox (based on the above-mentioned sample numbers). In order to return to the 1 or 2 spam messages per day "Nirvana" of last Spring, you'll need at least a 99.5% catch rate. Sound impossible? It is for Anti-Spam 1.0, but not for Anti-Spam 2.0.

Anti-Spam 2.0 needs to offer more proactive anti-spam technologies in order to reach and maintain a 99.5% or higher spam catch rate. It cannot wait for "rule updates" or rely on blacklists. To rely solely on reactionary technologies opens the door to these high-volume Botnets. Anti-Spam providers can go from hourly "rule updates" to by-the-minute updates, but that will only delay the inevitable as Botnets grow in size and deliver more spam to your inbox even with a shorter window of opportunity.

As CTO at GWAVA, I can tell you what are we delivering in our Anti-Spam 2.0 framework called GWAVA 4:

Automatic Training: GWAVA 4 can be set to automatically collect Spam and Ham (non-spam) from a variety of sources in order to train itself. This is significant and different from previous versions of GWAVA, because the training is perpetual. GWAVA 4 never stops collecting spam/ham, training and optimizing itself. This allows it to easily keep pace with the Botnets. This also eliminates the reliance on "rule updates" because GWAVA 4's self-training can deal with new spam much faster than it could receive updates. GWAVA 4 still has the capability to receive rule updates, but that is not (and should not be) its primary weapon against spammers. GWAVA 4 becomes, essentially, self-reliant.

"Ham" Detection: I believe that too much emphasis has been placed on Spam detection and not enough on Ham (non-spam) detection. GWAVA 4 spends most of its time learning about your Ham. Think about this for a moment: If your anti-spam system can reliably identify your Ham, then it can assume that any message that does not look like ham **MUST** be spam! This solves a lot of problems and allows GWAVA 4 to not only deal with today's spam scourge, it will also deal easily (and autonomously) with whatever the spammers and their Botnets throw at us next. No need to wait for updates while you are pummeled with spam.

Your Ham doesn't change very much. Most of the trouble in dealing with spam is that it changes so radically from month-to-month. GWAVA 4 automatically learns about your Ham, so the more radically spam changes, the easier it is for GWAVA 4 to detect non-Ham. This completely turns conventional anti-spam technology on its head with very beneficial results.

GWAVA 4 will receive frequent updates automatically, but those frequent updates are in support of some highly automated intelligence that allow it to meet the challenge of Spam 2.0 and beyond.



GWAVA 4 for Linux is already shipping, with GWAVA 4 for NetWare scheduled to ship in January. Existing GWAVA 3.6 customers should expect a GWAVA 3.7 release shortly that will provide them with some of the GWAVA 4 Anti-Spam technology to assist them in the short term.

END NOTE: This article was intended to offer a brief overview of the current problem of increased spam volume. Yes, in addition to pumping out record amounts of spam, spammers are also using message composition techniques to evade your current anti-spam filters. That is nothing new. Be it simple character replacement or HTML tricks or the so-called "image spam" or Bayesian Poisoning, spammers have always, and will always be trying to circumvent your anti-spam system. The current techniques in use by spammers is an interesting topic for another article but not relevant to this one, since spammers are always going to be employing newer and newer anti-spam evasion tactics.

Author Bio.

Charles Taite is CTO and Co-Founder of GWAVA, Novell's largest GroupWise partner. Based in Montreal, Quebec, Canada, GWAVA has offices in Europe, North America, and Asia Pacific. The company's success has come from a tight focus on solving real issues for Novell GroupWise customers. For more information about Charles or GWAVA visit www.gwava.com