



WHITEPAPER

Search Engine Link Spam: Risks, Threats, Solution

Nick Johnston, Software Engineer, MessageLabs

Old Enemy, New Weapon

Viruses, trojans, spyware and phishing may sound more sinister, but spam remains the biggest email-borne threat to businesses. Incredibly, unsolicited emails now account for almost three-quarters of all electronic traffic heading for corporate gateways around the world.

Spam's relatively innocuous name should never blind businesses to the serious harm it can cause. Quite simply, unless adequate defenses are in place, electronic junk mail will inevitably lead to overburdened inboxes, creaking networks and wasted bandwidth. An organization's efficiency, productivity and profitability will all be prominent, immediate casualties wherever anti-spam protection is not up to the job.

Of course, spam has a very long pedigree and, over the years, email security vendors have responded by strengthening the detection and filtering techniques they offer their clients. But this is a war which escalates remorselessly. Spammers continually devise new, cunning and increasingly sophisticated ways of evading spam defenses and achieving their objectives.

In autumn 2007, MessageLabs detected the emergence of a new "smart" weapon in the spammers' arsenal — so-called "redirector" or "search engine" spam. By early 2008, this had grown into a significant threat — one that businesses need to be aware of and take effective measures to combat.

This MessageLabs whitepaper puts redirector/search engine spam under the spotlight. It explains why the phenomenon evolved and how it works. But it also pinpoints a proven, cost-effective solution to this latest manifestation of spammers' never-ending ingenuity. The information presented here is based on MessageLabs hands-on experience of providing proven messaging and web security management services for over 17,000 clients worldwide, with around 2.5 billion attempted Simple Mail Transfer Protocol (SMTP) connections processed every day on their behalf.

URLs — A Key Battleground

Almost without exception, a spam email will target the recipient with some sort of call to action. In most cases, this will consist of a URL (Universal Resource Locator — an Internet address) accompanied by text saying "visit our online store!" or something similar. In other cases, the call to action might revolve around a phone number or a stock ticker symbol (a series of characters representing a particular listed or publicly traded stock).

But including a URL is by far the most popular technique preferred by spammers. It's easy to see why. URLs are quick and simple to insert into emails. If clicked on, they will take the recipient of the email directly to the spammer's website. Unlike the spam email itself — which has to be designed in a way which maximizes its chances of evading anti-spam filters — spammers are not restricted in what they can include on their websites.

Little wonder, then, that analysis of URLs contained in emails now plays a key role in efforts to identify spam and stop it from reaching its destination. For example, many security vendors now use "honeypot" systems designed with the specific intention of attracting spam. The messages captured by these honeypots can be analyzed and all "bad" or suspicious URLs extracted. (Often, this is achieved by identifying instances where the same URL appears in thousands or even hundreds of thousands of emails — a telltale sign that those emails constitute a spam run). Any email subsequently identified as containing such a URL can then be blocked and prevented from reaching its target. This anti-spam technique is now well-tested and has proved both efficient and reliable.

To counter URL blocking, spammers have tried a number of different tactics:

- One approach is to add random hostnames and gibberish to a URL. Take the URL <http://lhlgca.globren.info/?83217971&men>, for example. "lhlgca" and "/?83217971&men" are not part of the core URL. But by changing them slightly in every message sent out, the spammer aims to make the messages more difficult to block. However, security vendors can counter this tactic by focusing on the actual domain part of the URL ("globren.info" in this instance).

Redirector spam is a real threat to businesses.

Spammers have tried to counter URL blocking.

- Spammers are also adept at abusing legitimate, free website-hosting services like Geocities, Blogspot and Google Pages, as well as URL-shortening websites like TinyURL. Indeed, they have been known to set up hundreds of thousands of cheap, disposable, unmonitored accounts by using such services. Spam forums on the Internet have even advertised software, such as Geocities Account Creator, which automatically produce accounts of this sort, allegedly at rates of up to several accounts per second.
- Another technique involves camouflaging and muddying URLs by inserting superfluous characters or spaces and instructing recipients to remove them. The problem with this, from the spammer's viewpoint, is that it removes much of the simplicity of using URLs. Instead of simply clicking on a link, the recipient will have to copy the URL to their clipboard and modify it before visiting the relevant website. This "hassle factor" will undoubtedly reduce the spammer's ultimate conversion rate.
- To try to evade anti-spam filters, spammers like to change the domains they use — perhaps as often as several times an hour. As a result they have been quick to exploit a practice known as "domain tasting". This enables a spammer to register a large number of domains, use them in a spam run, get their registration fees refunded within a five-day "grace period", and then repeat the process all over again. Until the Internet Corporation for Assigned Names and Numbers (ICANN) recently closed this loophole for top-level domains, at least 90% of all domain name registrations were part of various domain-tasting schemes.

Redirecting the Danger

With some of their existing techniques struggling to outfox URL blocking, spammers have naturally focused their energies on developing alternative ways of achieving their required results. And it's in this context that they have pinpointed an opportunity presented by Internet search engines themselves.

Search engines, such as Google and Yahoo!, are designed to respond to user queries by displaying a list of the top results for that query.

However, a user clicking on one of these results is not sent directly to the website concerned. Instead, the search engine redirects them to the chosen destination using a "redirector script". This enables the search engine to track and record which links are being clicked on in response to a particular query. The data can then be used, for example, to optimize the search engine's response to similar queries in future. A large number of websites other than search engines also use techniques of this kind — which are completely legitimate — to monitor what advertisements visitors are clicking on, for example.

From a spam perspective, the key point is that the destination URL (i.e. the website the user/visitor wants to access) is specified as part of the URL contained in the redirector script. Take the following example:

www.google.com/pagead/ickk?sa=l&ai=bdialA&num=17119&adurl=http://legitsite.com

This URL will redirect the user, via Google, to the address specified in the "adurl" parameter at the end, i.e. "http://legitsite.com". However, a problem arises because these scripts have to be publicly available and publicly accessible. This means spammers can access them too. All they have to do is take a redirecting URL like the one above and replace the final destination with one of their own, for example:

www.google.com/pagead/ickk?sa=l&ai=bdialA&num=17119&adurl=http://imortgage.tw?tee

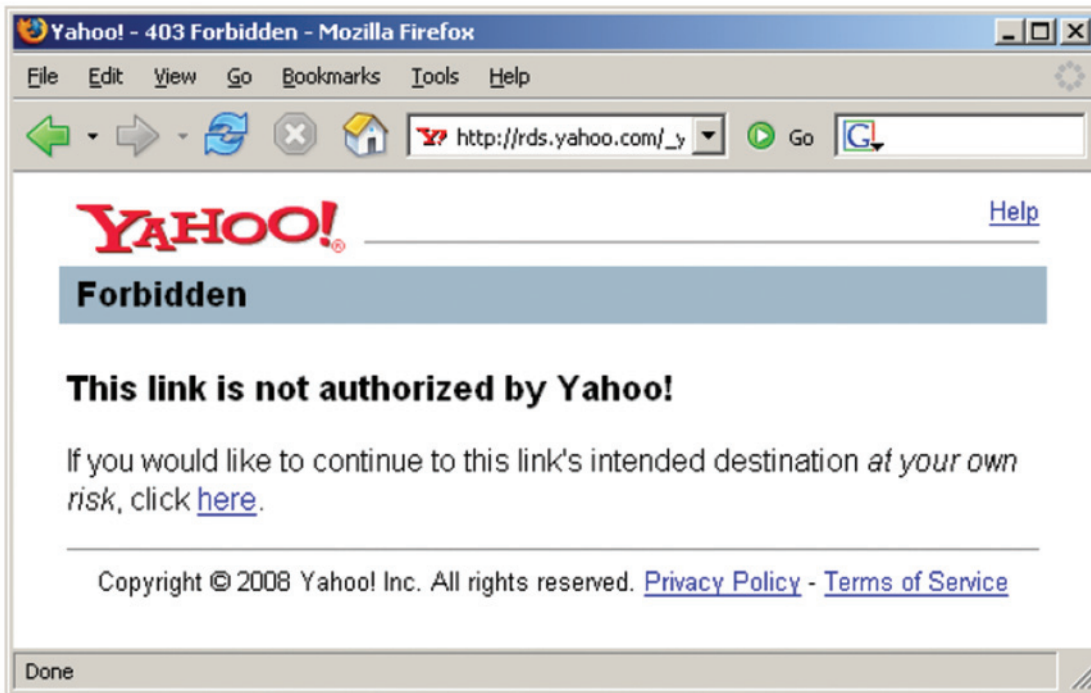
The spammer then simply inserts this URL in the spam email they send out. Any recipient clicking on the link will thus be directed, via Google, to the spammer's website.

From the spammer's viewpoint, this technique is simple and effective. It has the potential benefit that recipients may be more inclined to believe the spam message containing the modified URL is actually legitimate, because the URL incorporates the domain name of a well-known search engine, website or company. Furthermore, the incorporation of such a domain name may improve the message's chances of evading anti-spam defenses based on URL blocking.

In general, there is little that operators of redirector scripts can do to protect against the hijacking of these scripts. Victories have been achieved in

some areas, though — by Yahoo!, for instance. Some of Yahoo’s URLs now include a one-way, irreversible cryptographic hash or “fingerprint” based on parts of the URL, including an expiration time and destination URL. Now, when someone clicks on any URL containing the Yahoo! domain name, Yahoo! automatically recalculates the hash and works out whether the URL has been tampered with. If it has, Yahoo! displays the following page:

There is little that script operators can do to protect against this.



Feeling Lucky?

But many avenues still remain open to spammers intent on exploiting publicly available redirector scripts. One clever redirection exploit uses Google’s “I’m Feeling Lucky” feature. This feature is designed to allow users searching on Google to be redirected immediately to the top search result for their query. No list of search results is shown, so the user doesn’t have to click on a subsequent link.

Consider this URL included in a spam email:

google.com/search?hl=en&q=inurl:rneskimo.com%2BVPXL%2BMade%2BEasy&btnI=6904018

The crucial detail here is the URL’s last parameter (“&btnI=6904018”). It’s equivalent to the user clicking the “I’m Feeling Lucky” button instead of the normal “Google Search” button (they appear beside each other on Google’s homepage).

Many redirector scripts are available to spammers.

By clicking the URL, the spam recipient will immediately be redirected, via Google, to the spammer's website, in just the same way as the "I'm Feeling Lucky" button would take them straight to a top search result. Another spammer, meanwhile, has discovered that Google has registered several misspelled variants of the domain name "google.com" (e.g. "goooogle.com"). Google's aim in doing this was presumably to combat so-called "typo squatters" who register domain names that are slightly misspelled variations of well-known trade names. The spammer has therefore tried to use these variants as part of their redirector script abuse strategy, probably in an attempt to evade tighter checks on google.com URLs.

The Best Defense

Redirector or search engine spam simply represents the latest tactic creative and versatile spammers are adopting in order to deliver their payloads into the heart of businesses. However, traditional appliance-based and software-based anti-spam defenses — and even some managed email security services — are finding it increasingly difficult to keep pace with this creativity and versatility.

Fortunately, a readymade solution is at hand. MessageLabs Email Anti-Spam service is a fully managed service that provides unparalleled ability in keeping all kinds of spam away from corporate networks. Delivering over 99% spam capture and near-zero false positives, it provides the reassurance businesses need that they will not fall victim to redirector — or any other type of — spam.

Incorporating advanced honeypot technology and managed 24/7 by a technical support team that constantly monitors spam activities and identifies newly emerging techniques, the MessageLabs service guns down spam before it becomes a problem. It's affordable. It's effective. And it's guaranteed to provide the armour-plated anti-spam protection your business needs.

For more information about MessageLabs Email Anti-Spam service or for a free trial, visit www.messagelabs.com.

MessageLabs provides
armour-plated anti-spam protection.



Americas
AMERICAS HEADQUARTERS

512 Seventh Avenue
6th Floor
New York, NY 10018
USA
T +1 646 519 8100
F +1 646 452 6570

CENTRAL REGION
7760 France Avenue South
Suite 1100
Bloomington, MN 55435
USA
T +1 952 830 1000
F +1 952 831 8118

Asia Pacific
HONG KONG
1601
Tower II
89 Queensway
Admiralty
Hong Kong
T +852 2111 3650
F +852 2111 9061

AUSTRALIA
Level 14
90 Arthur Street
North Sydney
NSW 2060
Australia
T +61 2 9409 4360
F +61 2 9955 5458

SINGAPORE
Level 14
Prudential Tower
30 Cecil Street
Singapore 049712
T +65 6232 2855
F +65 6232 2300

Europe
HEADQUARTERS
1270 Lansdowne Court
Gloucester Business Park
Gloucester, GL3 4AB
United Kingdom
T +44 (0) 1452 627 627
F +44 (0) 1452 627 628

LONDON
3rd Floor
1 Great Portland Street
London, W1W 8PZ
United Kingdom
T +44 (0) 207 291 1960
F +44 (0) 207 291 1937

NETHERLANDS
Teleport Towers
Kingsfordweg 151
1043 GR
Amsterdam
Netherlands
T +31 (0) 20 491 9600
F +31 (0) 20 491 7354

BELGIUM / LUXEMBOURG
Culliganlaan 1B
B-1831 Diegem
Belgium
T +32 (0) 2 403 12 61
F +32 (0) 2 403 12 12

DACH
Feringastrasse 9
85774 Unterföhring
Munich
Germany
T +49 (0) 89 189 43 990
F +49 (0) 89 189 43 999