1F-C6, No. 1, Lishing1st Rd.,
Science-Based Industrial Park,
Hsinchu City 30078, Taiwan
TEL: 886 3 578 9399
FAX: 886 3 578 0707

Detection of Black Hat SEO Links

A Lionic White Paper | Mar 2011

Lionic Corp.
1F-C6, No. 1, Lising 1st Rd.,
Science-Based Industrial Park,
Hsinchu City 300, Taiwan
www.Lionic.com

What is Black Hat SEO and How to Detect Them?

To steal credit card numbers, passwords, and other sensitive personal information, malware writers always try to get their virus reach more people. They used to propagate virus via USB drives, spam and drive-by-download web site. Lately, they have increasingly been targeting the links served up by search engines. It is reported that 10 percent or more of the results returned by one-third of popular search terms led to malware [1-3]. What's worse, downloaded malwares have very low detection rate among AV vendors [4].

Google was aware of the threat and had a solution for this [5]. However, the threat is still very severe. For example, on May 21, 2010, we searched for the then trending topic "adam wheeler Harvard" in Google. In the first 250 links, we found 131 links are redirected to hostile web sites. It is very obvious that attackers can easily occupy over half the top results of searching for trending topics.

The attack is called Black Hat SEO (Search Engine Optimization) or search engine poisoning. The steps to launch this attack are as followed:

1. Compromise legitimate web sites. Attackers compromise as many web sites as possible. Leveraging the high page rank of these sites, they can launch more effective SEO poisoning attacks.

2. Create SEO-friendly fake pages. These pages are easily for search engine bot to parse. They are linked with each other via popular keywords to boost page rankings. Hundreds of linked pages can ensure that their malicious sites reach the top of the page rankings for given search keywords.

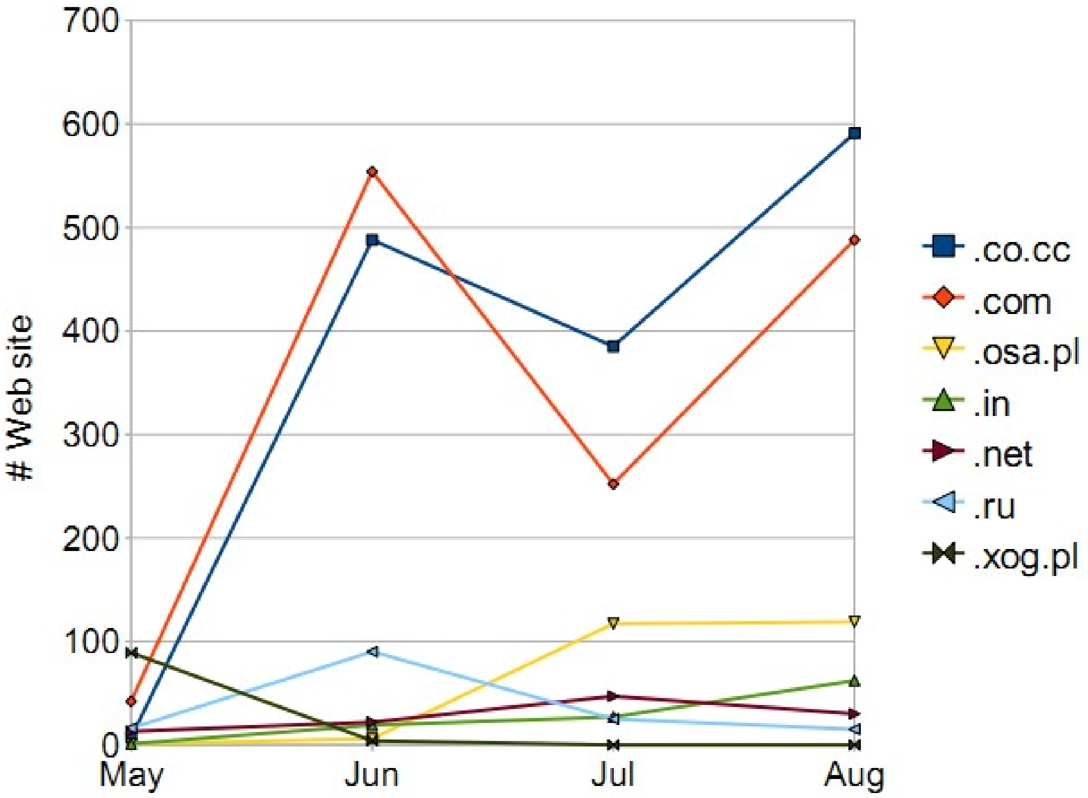
3. Hide malicious content from search bots and security analysts. SEO poisoning attacks have been difficult to eliminate, because the attackers are shielding their fake pages from search engine detection and security analysts. Poisoning pages serve up alternative non-malicious pages with relevant keywords and links to other poisoning pages.

4. Deliver payload using social engineering technique. If traffic does come from a search engine, poisoning pages will serve up the bad content. Recently, researchers report that the bulk of them are used to send users to a fake AV scan page to convince them to install Fake AV binary [6].

We tested the above-mentioned 131 SEO poisoning pages. There were total 9 hopping sites when clicking these 131 web pages. If we can block these 9 hopping sites, the users can be free from this threat. Google Safe Browsing [7] provides this blacklist that are used in Firefox and Google Chrome browser. However, in this case, Google Safe Browsing blocks only 4 out of 9 hopping sites. Thus, we design a new method to collect the latest Black Hat SEO links.

Malicious Domains Distribution

The following figure shows the top seven domains we found in each month. The .xorg.pl domain is abused in May 2010 to service Black Hat SEO sites. After this domain is listed in many black lists [8-9] in May, we no longer found .xorg.pl sites in the field. The .co.cc domain became popular since Jun 2010, and the trend is still on rise. Recently, we observe there are more and more malicious sites in the .osa.pl domain. We keep tracking the most abused domains to help the classification algorithm.



References

1. **Barracuda Labs: Barracuda labs 2010 midyear security report. Tech. rep. (2010)**

2. **Ferraris, R.: Internet searches under attack: next in series, http://comunity.ca.com/blogs/securityadvisor/archive/2008/01/15/internet-searches-under-attack-next-in-series.aspx**

3. **Howard, F., Komili, O.: Poisoned search results: How hackers have automated search engine poisoning attacks to distribute malware. Tech.rep., SophosLabs (2010), http://www.sophos.com/security/technical-papers/sophos-seo-insights.html**

4. **Julien: Nearly 3 million "hot video" pages pushing fake av are undetected. http://research.zscaler.com/2010/08/nearly-3-millions-hot-video-pages.html (august 2010), http://research.zscaler.com/2010/08/nearly-3-millions-hot-video-pages.html**

5. **Rajab, M.A., Ballard, L., Mavrommatis, P., Provos, N., Zhao, X.: The nocebo effect on the web: An analysis of fake anti-virus distribution. In: Workshop on Large-Scale Exploits and Emergent Threats (April 2010), http://www.usenix.org/event/leet10/tech/**

6. **Cova, M., Leita, C., Thonnard, O., Keromytis, A., Dacier, M.: An analysis of rogue av campaigns. In: the Symposium on Recent Advances in Intrusion Detection (RAID) (September 2010)**

7. **Google: Google safe browsing API. http://code.google.com/apis/safebrowsing/**

8. **McAfee: Mcafee siteadvisor. http://www.siteadvisor.com/**

9. **Norton: Norton safe web. http://safeweb.norton.com/**