A novel approach to alert website owners for defacement

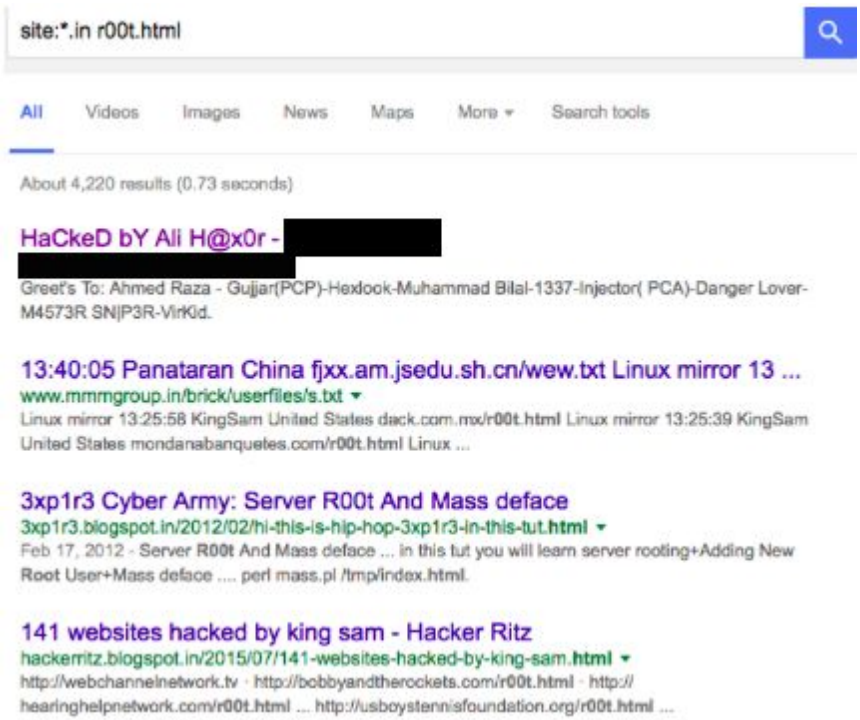
Background:
News like the website defacement of an Indian university, government organization or a private company is becoming common place.

In the recent keynote of Mr. Muktesh Chander which was shared on youtube. It was mentioned that last year more than 25,000 websites were defaced.

Website defacement is similar in terms to computer virus industry by looking for a website that can be exploited for a vulnerability and instead of a virus, the payload is a harmless webpage that is inserted. However, the action is more politically motivated. So, while bigger companies like Google were able to enhance their browsers by identifying and preventing innocent users from opening malware-distributing websites. They are probably not looking at the website defacement as an area where they can make impact.

Google as an alert mechanism:
We all know the bots of Google are crawling the web and indexing them to give us search results all the time.

For example a query like: " site:*.in r00t.html" returns the following results:



As a small explanation to the above query. It is searching for websites that are with **.in** domain and have a mention of '**r00t.html**' *(our intention is to look for this webpage in sitemap, we found **r00t.html** in the defacement done by a particular pakistani hacker group's unique attribute. Hence, a signature to find the websites defaced by them.)*

Here the top result is a website (masked with black color to protect the identity of the business) of a pune based business whose website was defaced by a Pakistani hacker and they were not aware of the compromise until we last checked with them.

As we have done this analysis again on 13th June 2016, the pune based business has not yet rectified it. The reasons for pune based business not aware of the compromise we assume are two:

a. Maybe, the business is not auditing their own website frequently
b. The page inserted is not in the main menu of the website. While, the main page (home) being intact such a defacement might have gone unnoticed.

So, when Google was indexing the web, it also indexed the defaced page, which enabled us to find it. Hence, we can not just find a defacement that happened in past with a time filter like websites that are defaced in last 24 hours or a week. But also in future! As, Google allows us to set email alerts for search queries they can alert us for any such defacement as they find it. ***Hence, Google can do our monitoring work for free!***

Past success stories of building the crowdsource security mechanisms
When in 2006 web applications like Web of Trust showed a possibility to crowdsource information through a firefox plugin about malware distributing websites. Companies like Google, which were involved in the business of making their own browsers latched on to the idea of making the service available to third parties and which is also built into their own browsers as a protection mechanism. The Safe Browsing API provided by Google is covering all kind of websites, which are either involved in phishing and malware distribution but not any mechanism that could be of use in tackling website defacement.

Next Steps:
Taking the cue from the above methodology we propose two possibilities. The authorities can build a robust crowdsourced repository of known signatures of defacement and:

a. Provide the database to a company like Google and request the search engine for providing alerts to website owners on real time basis. Hence, recognizing defacement as a crime at the same level as malware distribution for the browser and search engine.
b. Build their own crawler to find these known signatures in almost real time in government websites or .in websites.

Further, a process can be established to study the infected website servers for the known vulnerabilities and establish the vulnerabilities that can enable a hacker to deface a website. Maybe, an online service can then scan the prospective website for possibility of being defaced.

Contact Details of the Author:
Sachin Gaur,
sachin@mixorg.com
+91 99999 79349