Content alerts

Defacement

Websites are the most public representation of any modern enterprise. This combination of high publicity with complex technology makes websites a prime target for political hacktivists who aim to embarrass a company and cause as much damage as possible.

Website defacements therefore “consist of hacking into a web server and replacing a web page with a new page bearing some sort of message.”
[Samuel, A. W. (2004). Hacktivism and the future of political participation, Harvard University, Cambridge, Massachusetts, p.8]

Technology-enabled political activism is often driven by major international and social disputes. These disputes are reflected in defacement waves that target a similar group of websites and communicate similar messages.

Examples for such “hacktivism” waves are:

- Anonymous with changing targets since 2003
- Syrian Electronic Army with pro-Assad messages aimed at new outlets, universities
- ISIS mass defacements in 2015

Our goal is zero false positives. Because of the advanced methods used here, we cannot guarantee 100% detection of a defacement, but we see the changes and tell if they are really suspicious.

Definition

A Nimbusec defacement alarm is defined by the following conditions:

- The visual appearance of a website was changed to distribute social, political or just for fun messages to the visitor.
- The website was changed in a way that the visitor will easily distinguish it from an official change. As a result, the main context of the website was replaced.
- The definition of a defacement is way to unspecific. @Nimbusec, **we also differentiate between a phishing attempt (placing links on the website) and a defacement like defined above**. You can find phishing attempts by the suspicious link alarm event (see blacklist category)

How is it detected

Nimbusec detects defacements typically from external, by its Cloud Scan.

Basically this is a fast approach that may detect defacements by signatures. Therefore the content of a website is extracted (plain text without tags and attributes). This content is then matched against an increasing set of rules which represent the content of typically seen defacements. Included are also bad words in multiple variations e.g. hacked by | h4cked | h4ck3d!!! | ...

To sum it up:

- Content is matched on rules, representing typical defacements
- Bad words are looked up in the content

We also gather data from external threat intelligence services. Defacement alarms are directly created from those data sources and are also used to write new signatures to improve the overall detection rate.

Alert levels

- *RED*: Identification of known defacement based on defacement signature.
- *YELLOW*: Possible defacement warnings detected by our external threat intelligence services that are not confirmed yet.

Red Alert

On red alerts a defacement was seen based on signatures. This can be held as defacements similar to the detection rate of anti virus products and should be taken really seriously. Automated reaction can be considered.

Yellow Alert

These alerts have to be taken serious as well, but we recommend to have a look at the page and verify the issue yourself.

Recommended action

The defacement result detail gives you a lot of information and has in mind to give you on first sight what went wrong.

Also screenshots of the landing page are rendered to see what our crawlers saw at the time of the scan.

The content result gives you the following information:

- Occured on: the time of the detection
- Reason: Description of the modules which voted for this issue
- Region: the region where the scan was performed from
- Viewport: a mobile or desktop browser client setting can be used
- Path: URL to the page/resource where the change was seen
- Change: Usually for this issue --> Defacement detected
- Name: Name of the defacement, if we have one (e.g. Hacker group). So you can identify the defacement better.

Try to verify the result by browsing the path, shown in the result. Request the website directly and also over a search engine. You may get different results. **Do this only from a secure environment, as defaced websites are also likely to spread malware!**

In case the website got defaced, and you don't have a incident response chain yourself, we prepared a short guide here.

1. It is always good to create a backup of the webspace before changing anything.
2. Redirect to a maintainance page, or in worst case to just a blank page, to not bring site visitors to danger.
3. Investigate to find the weak spot (Outdated CMS or other application, some plugin, ..)
4. Fix the vulnerabilities and remove evidence

The detected change might have been part of a major, but intentional content change. Contact the website's administrator for information.

When to mark as False Positive

If no problem shows and you are certain that no defacement took place, mark the alert as "false positive". This will create a rule that should prevent such alerts in the future. But every content is different and therefore hard to compare and analyse - If you have a specific alert that shows up unless you marked it as false positive, please contact us and we will work out a solution.

If you need in depth information about why this alert was generated, please contact nimbusec support.

Revision #6
Created 4 August 2021 08:19:05 by Patrick Wall
Updated 7 June 2023 10:00:54 by Lena