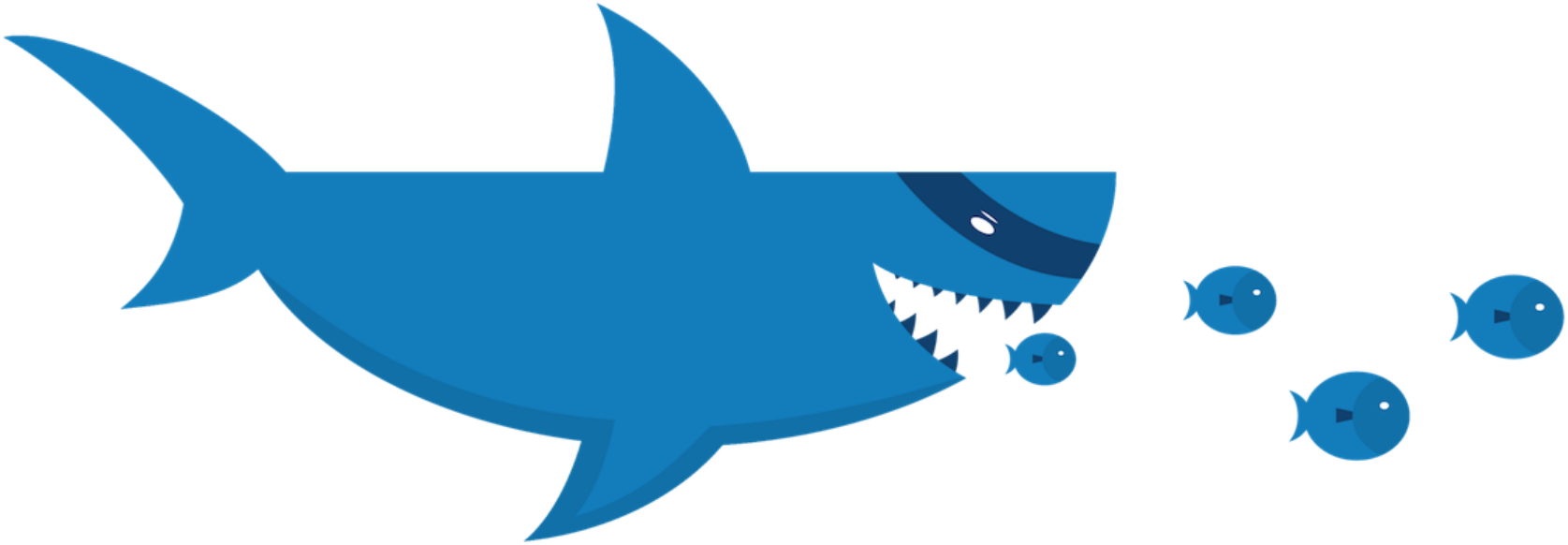
The Art of Subdomain Enumeration

24 APRIL 2017 on Technical, Subdomain enumeration

Introduction

In this blog post we will set you up with all you need to know about a dangerous art! Subdomain enumeration is an essential part of the reconnaissance phase in the cyber kill chain. Cyber attackers map out the digital footprint of the target in order to find weak spots to gain for example access to an internal network. Already know enough? Check out Sweepatic.com for powerful reconnaissance solutions, or subscribe to our newsletter on the button below. Still curious? Read on!



Subdomain enumeration is the process of finding valid (resolvable) subdomains for one or more domain(s). Unless the DNS server exposes a full DNS zone (via AFXR), it is really hard to obtain a list of existing subdomains. The common practice is to use

a dictionary of common names, trying to resolve them. While this method is effective in some cases, it doesn't include subdomains that have strange names. Another approach is to crawl the second-level domain in order to find links to subdomains (faster approach is to use a search engine directly).

Example:
After completing the subdomain enumeration process, the attacker finds `blog.example.com` as one of the subdomains in the target's DNS zone. The attacker is enriching this finding up to the web application layer and finds out that the blog is using Wordpress as a content management system. The attacker then runs wpscan in order to find Wordpress vulnerabilities. Fortunately, the target's Wordpress instance uses a vulnerable plugin which an attacker is able to exploit, gain access to the environment and pivot further into the network. This example might seem a bit exaggerated, however, this is exactly what happened in The Panama Papers case.

Let's present the most popular open-source tools and techniques for performing subdomain enumeration. Here we go:

Zone transfer

The most simple and basic technique is to try an AXFR request directly on the DNS server:

```
dig @ns.example.com example=.com AXFR
```

A zone transfer is used to copy the content of the zone across primary and secondary DNS servers. The best practice advises administrators to allow AXFR requests only from authorized DNS servers, so the above technique will probably not work. But if it does, you have found a goldmine.

Similar to zone transfer, there is a so called NSEC walking attack, which enumerates DNSSEC-signed zones.

Google Dorking

Time to use Google! Luckily, you can use various operators to refine you search queries (we also call these queries "Google dorks"). As mentioned previously, many subdomains can be found using crawling the target. Google (and also other search engines like Bing) does it as a byproduct of its primary intention. We can use the `site` operator to find all subdomains that Google has found:

```
site:example.com
```

Rapid7 DNS dataset

Rapid7 publicly provides its Forward DNS study/dataset on scans.io repository. The DNS dataset aims to discover all domains found on the Internet. While they do a very good job, the list is definitely not complete. You can read more about how they compile their dataset here. After downloading the latest snapshot, we can run jq on it to find subdomains:

```
zcat snapshop.json.gz |
jq -r 'if (.name | test("\\.example\\.com$")) then .name else empty end'
```

jq tests the regular expression *"ending with .example.com"* to find all subdomains in the dataset.

DNSDumpster is a free online service which is using exactly this technique.

Subject Alternative Name

Subject Alternative Name (SAN) is an extension in x.509 certificates to provide different names of the subject in one certificate. Companies often generate one certificate for multiple subdomains to save money.

We can look into certificates to hunt for subdomains in SAN's using two different sources:

Censys.io

Censys.io is an interface to a subset of data published by scans.io. The good part is, that it allows to search for keywords in certificates and thus potentially reveal new subdomains:

```
https://censys.io/certificates?q=.example.com
```

Crt.sh

Crt.sh is an online service for certificate search provided by COMODO. It uses a different dataset than Censys, but the principle is the same: find subdomains in certificates.

```
https://crt.sh/?q=%25.example.com
```

It is good to note, that although some domains respond with *NXDOMAIN*, they still might exist on the internal network. Administrators sometimes reuse the certificates for public servers on their intranet servers...

Sublist3r

One of the most popular open source tools for subdomain enumeration, is called Sublist3r. It aggregates output from many different sources, including:

- Google
- Bing
- Virustotal
- crt.sh
- ...

While the data is correct in most cases, you might encounter non-resolvable subdomains (domains responding with *NXDOMAIN*). This is because Sublist3r relies heavily on *passive data* and it doesn't validate whether the found subdomains really exist.
Sublist3r also uses a standalone project called subbrute. Subbrute is using the dictionary of common subdomain names in order to find a subset of subdomains that are resolvable.

To use it, simple run:

```
python sublist3r.py -d example.com
```

and the list of subdomains of *example.com* will be presented to you.

theHarvester

Another Open Source intelligence gathering tool, is called theHarvester and finds e-mail addresses on target domains as well as subdomains and virtual hosts. However, compared to Sublist3r, it provides fewer subdomain results. You can run theHarvester using the following command:

```
python theHarvester.py -d example.com -b all
```

Smart DNS Brute-Forcer (SDBF)

Subdomain enumeration tools often include a list of common subdomains that they try to resolve. This approach can be extended by using Markov chains in order to discover a subdomain name structure (e.g. you have www1, it is likely that www2 will exist and so on). There is a research paper from Cynthia Wagner et al explaining this technique in greater detail. The results produced by SDBF are far better than simple keyword enumeration of subdomains.

Conclusion

Periodically checking what subdomains can be discovered by your cyber adversaries provides a great input towards vulnerability assessment teams and other tactical cyber teams in the organization.

Running frequent reconnaissance against your environment will provide you with greater visibility to find forgotten subdomains. The latter can expose your environment and the company to a wide range of threats like subdomain takeover or even full compromise as seen in the example at the beginning of this blogpost.

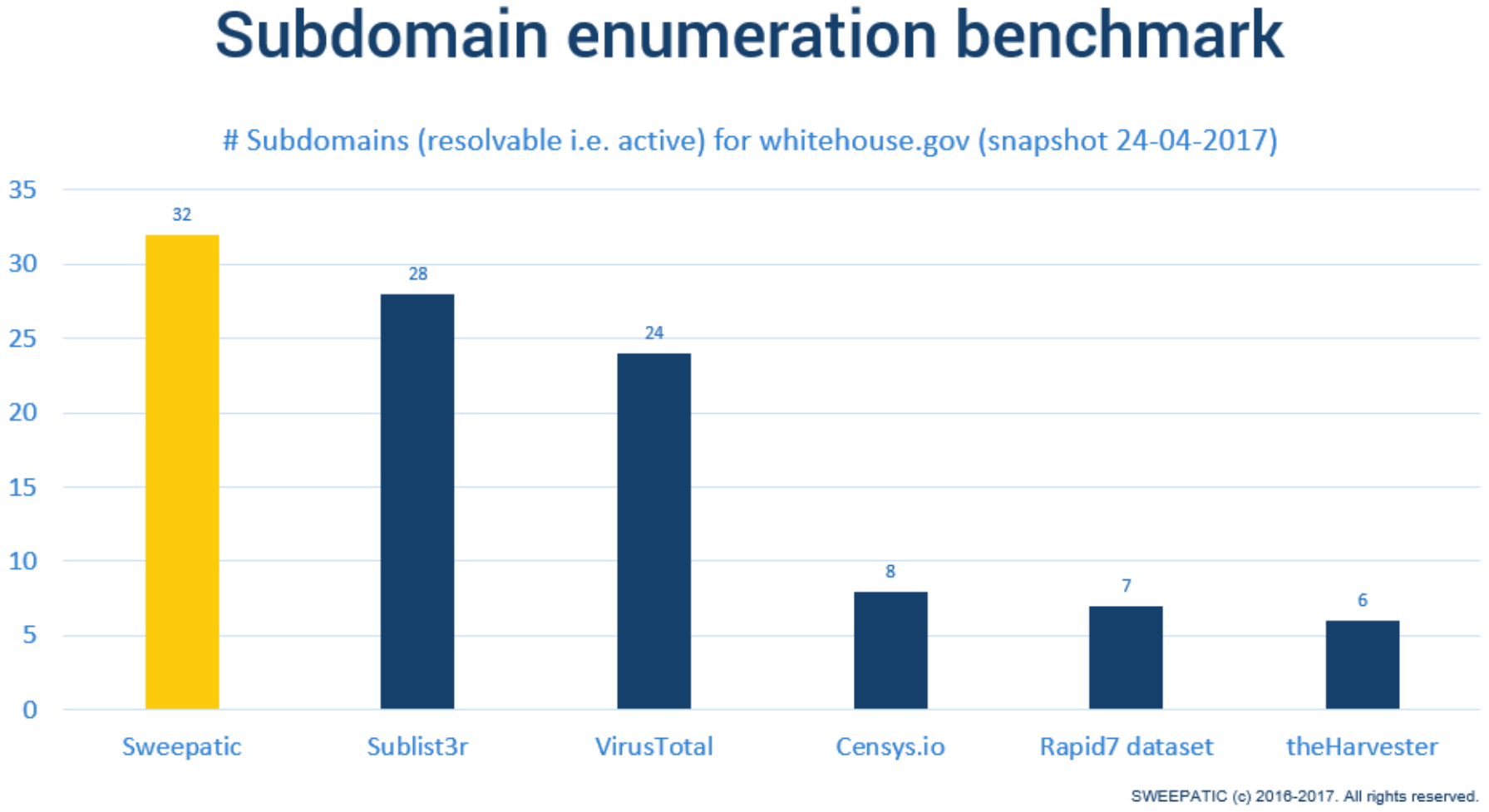
Even still, too many times have we witnessed a false feeling of security in cyber teams. By only partially understanding the dimension of the company's digital footprint, internal teams are not always aware of their complete exposure and fail to minimize the risks.

But don't worry, we are here to help!

Through our Sweepatic reconnaissance platform, we offer you and your company a unique and easy to use solution, that will help you understand, monitor and reduce your dynamic digital footprint to minimize the risk of compromise.

The table below shows a benchmark we recently (data from 24-04-2017) did and contains subdomain enumeration results from Sweepatic and various other tools on

the domain: *whitehouse.gov*



Sweepatic has by far the best result. So if your company has a complex footprint, come and get your complete report at Sweepatic.com, or subscribe to our newsletter

for regular updates.

That's all for now, more coming soon!
Until next time!

Subscribe to our Newsletter

Email Address

Subscribe

Patrik Hudak

Read more posts by this author.

READ THIS NEXT

The Principles of a Subdomain Takeover

YOU MIGHT ENJOY

Loading up to Infosecurity.be Brussels 2017, featuring the White House.