SANS ISC: Hunting phishing websites with favicon hashes - SANS Internet Storm Center SANS Site Network Current Site SANS Internet Storm Center Other SANS Sites Help Graduate Degree Programs Security Training Security Certification Security Awareness Training Penetration Testing Industrial Control Systems Cyber Defense Foundations DFIR Software Security Government OnSite Training SANS ISC InfoSec Forums

isc.sans.edu/forums/diary/Hunting+phishing+websites+with+favicon+hashes/27326/

Hunting phishing websites with favicon hashes

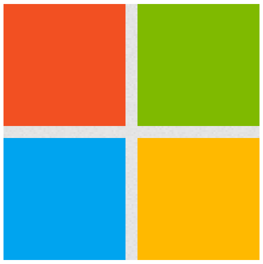
Jan
73
Posts
ISC
Handler
Apr
19th
2021

HTTP favicons are often used by bug bounty hunters and red teamers to discover vulnerable services in a target AS or IP range. It makes sense – since different tools (and sometimes even different versions of the same tool) use different favicons[1] and services such as Shodan calculate MurmurHash values[2] for all favicons they discover and let us search through them, it can be quite easy to find specific services and devices this way.

But while the use of favicon hashes is common in the “red” community, significant number of blue teamers don’t use them at all. Which is unfortunate, given that – among their other uses – they can provide us with a simple way of identifying IPs hosting phishing kits. After all, this was the reason why searches using HTTP favicon hashes have been introduced into Shodan in the first place[3].

As an example, we will show how to detect IPs hosting phishing pages by looking for sites that try to pass themselves of as login portals for O365 and other Microsoft services, however the same principle would work for any other service as well. One could therefore for example calculate hashes of unique favicons used by systems specific to a company one is trying to protect (e.g. favicon from a company website) and use periodical lookups of these on Shodan and other services in order to implement a – admittedly fairly simple – phishing detection/brand protection mechanism...

So how would one look for fake Microsoft login portals? First, we need to calculate a MurmurHash value of a favicon that we expect might be reused on a phishing website to make it look more trustworthy. Looking at official Microsoft websites, it seems that they use the favicon located at https://c.s-microsoft.com/favicon.ico.

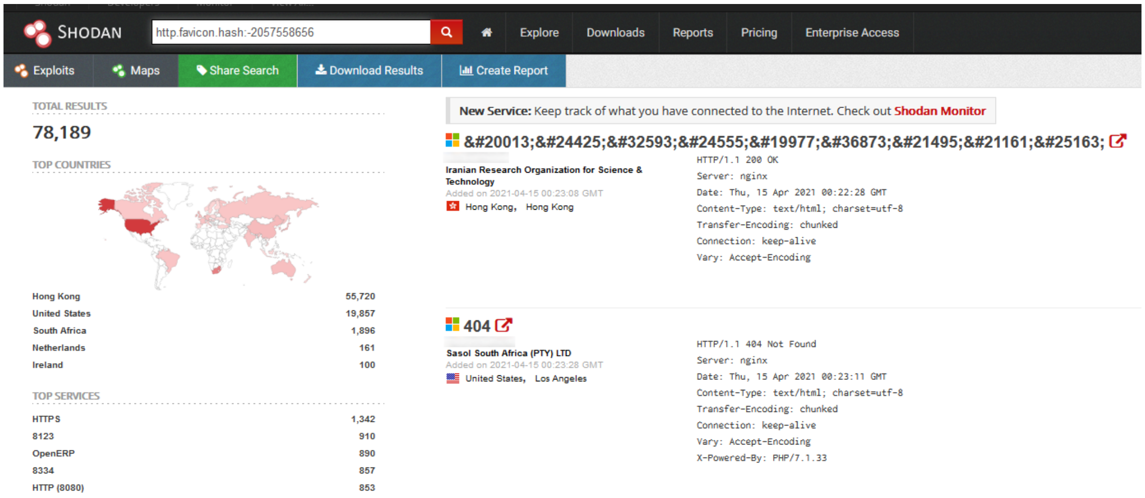


Its hash can be easily calculated using Python code that may be found on GitHub[4]:

```
import requests,mmh3,base64
response = requests.get('https://c.s-microsoft.com/favicon.ico')
favicon = base64.encodebytes(response.content)
hash = mmh3.hash(favicon)
print(hash)
```

If we run this script, we will get the hash -2057558656.

Now that we have a hash to look for, we can query Shodan to get the list of all IP addresses where it found a favicon with the same one. We may use the filter http.favicon.hash to do so.



TOTAL RESULTS

78,189

TOP COUNTRIES

Hong Kong	55,720
United States	19,857
South Africa	1,896
Netherlands	161
Ireland	100

TOP SERVICES

HTTPS	1,342
8123	910
OpenERP	890
8334	857
HTTP (8080)	853

New Service: Keep track of what you have connected to the Internet. Check out **Shodan Monitor**

&#20013;&#24425;&#32593;&#24555;&#19977;&#36873;&#21495;&#21161;&#25163;

Iranian Research Organization for Science & Technology
Added on 2021-04-15 00:23:08 GMT
Hong Kong, Hong Kong

```
HTTP/1.1 200 OK
Server: nginx
Date: Thu, 15 Apr 2021 00:22:28 GMT
Content-Type: text/html; charset=utf-8
Transfer-Encoding: chunked
Connection: keep-alive
Vary: Accept-Encoding
```

404

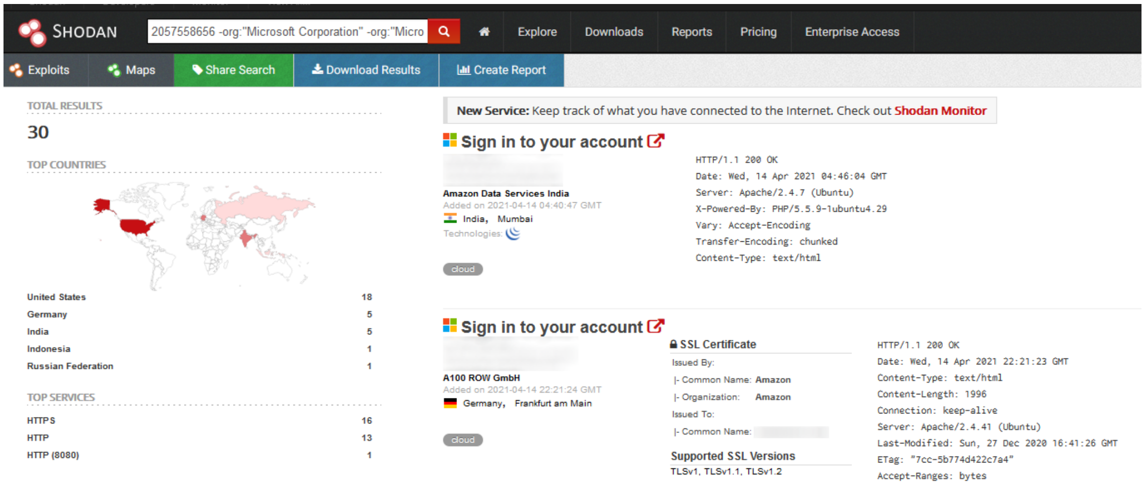
Sasol South Africa (PTY) LTD
Added on 2021-04-15 00:23:28 GMT
United States, Los Angeles

```
HTTP/1.1 404 Not Found
Server: nginx
Date: Thu, 15 Apr 2021 00:23:11 GMT
Content-Type: text/html; charset=utf-8
Transfer-Encoding: chunked
Connection: keep-alive
Vary: Accept-Encoding
X-Powered-By: PHP/7.1.33
```

As we can see, the number of results is quite high. This is hardly surprising though, given that they conain all servers – malicious as well as legitimate – where the favicon is used. In order to discover only the suspicious ones, we would need to further refine the search. One would do this differently for one's own favicons, but in order to search for suspicious Microsoft login portals, we could extend our search to look only for IPs with web pages looking like log in portals (http.html:"Sign in") and that are not hosted on Microsoft infrastructure (-org:"Microsoft Corporation" -org:"Microsoft Azure") but are running an Apache web server (product:"Apache httpd"). Taken together, our search might look like this:

```
http.favicon.hash:-2057558656 -org:"Microsoft Corporation" -org:"Microsoft
Azure" product:"Apache httpd" http.html:"Sign in"
```

If we ran this updated search, the number of results would be significantly lower.



2057558656 -org:"Microsoft Corporation" -org:"Micro

New Service: Keep track of what you have connected to the Internet. Check out **Shodan Monitor**

TOTAL RESULTS

30

TOP COUNTRIES

United States	18
Germany	5
India	5
Indonesia	1
Russian Federation	1

TOP SERVICES

HTTPS	16
HTTP	13
HTTP (8080)	1

Sign in to your account

[redacted]

Amazon Data Services India
Added on 2021-04-14 04:40:47 GMT
India, Mumbai
Technologies:

```
HTTP/1.1 200 OK
Date: Wed, 14 Apr 2021 04:46:04 GMT
Server: Apache/2.4.7 (Ubuntu)
X-Powered-By: PHP/5.5.9-1ubuntu4.29
Vary: Accept-Encoding
Transfer-Encoding: chunked
Content-Type: text/html
```

Sign in to your account

[redacted]

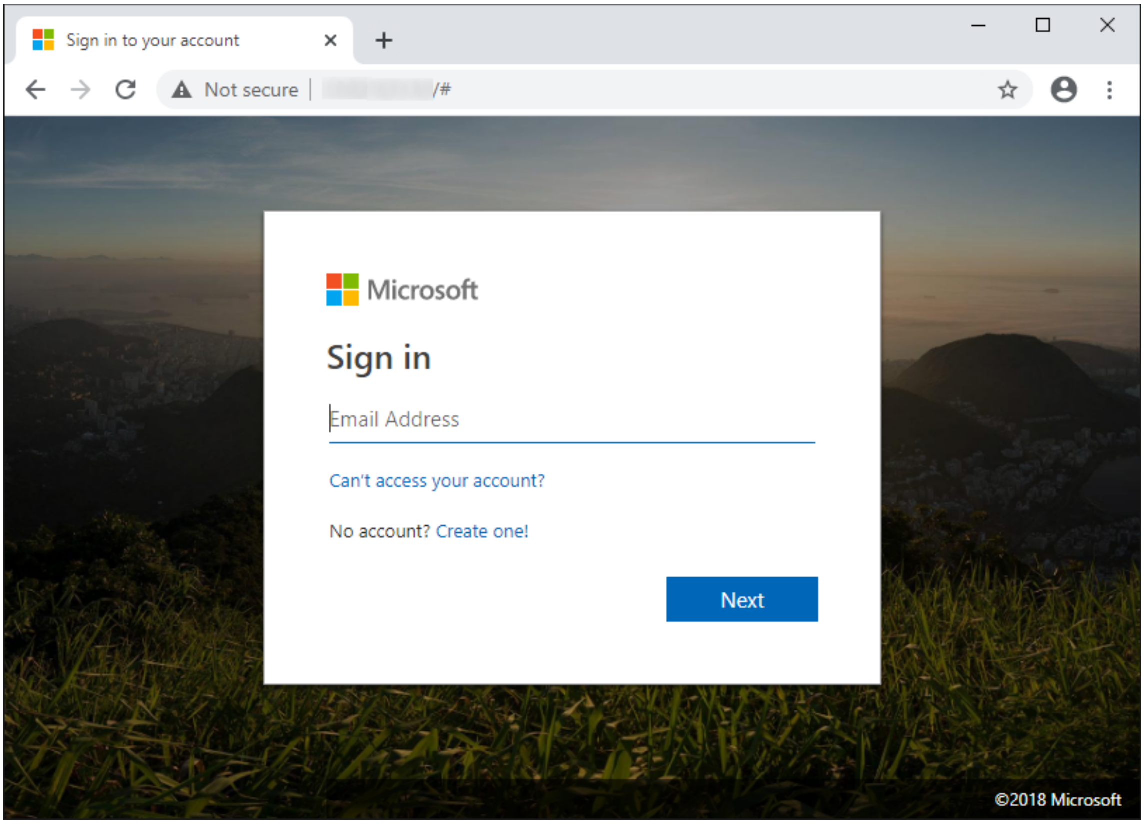
A100 ROW GmbH
Added on 2021-04-14 22:21:24 GMT
Germany, Frankfurt am Main

SSL Certificate
Issued By:
|- Common Name: Amazon
|- Organization: Amazon
Issued To:
|- Common Name: [redacted]

Supported SSL Versions
TLSv1, TLSv1.1, TLSv1.2

```
HTTP/1.1 200 OK
Date: Wed, 14 Apr 2021 22:21:23 GMT
Content-Type: text/html
Content-Length: 1996
Connection: keep-alive
Server: Apache/2.4.41 (Ubuntu)
Last-Modified: Sun, 27 Dec 2020 16:41:26 GMT
ETag: "7cc-5b774d422c7a4"
Accept-Ranges: bytes
```

Not all IPs identified in this way would necessarily turn out to host a phishing website, but most of them almost certainly would (or would at least turn out to have done so recently). In any case, all of them would unquestionably be worth investigating, and it probably wouldn't take too long to discover something interesting. In our search, for example, the following web site was hosted on the very first IP that Shodan returned.



Sign in

Email Address

As we've mentioned, the same approach can be used to identify phishing web sites using any other favicon as well.

Given how easy it is to implement automatic periodic lookups (for example against Shodan API) for a list of specific hashes (e.g. the ones that are used on our company log in pages/in our products), favicons can provide a cheap and simple way to detect phishing sites targeting either one's company or its customers. Even if one decided not to automate them, favicon hash lookups can still provide us with additional information useful, for example, for “long tail” threat hunting or enrichment of other data.

In any case, if you are on the “blue” side and don’t use favicon hashes in any way, consider whether they might not provide you with at least some value.

[1] https://github.com/sansatart/scrapts/blob/master/shodan-favicon-hashes.csv
[2] https://en.wikipedia.org/wiki/MurmurHash
[3] https://twitter.com/shodanhq/status/1280247570586099719
[4] https://gist.github.com/yehgdotnet/b9dfc618108d2f05845c4d8e28c5fc6a

----
Jan Kopriva
@jk0pr
Alef Nula