Early Typosquatting Detection Made Possible: A Short Illustration in the Financial Sector

On March 12, 2020, IBM X-Force Exchange published an early warning on a Wells Fargo Squatting Campaign:

"We observed 3 Squatting Domain registrations related to a victim in the finance and insurance sector. The campaign was identified starting with the registration on 2020-03-07 01:15:57 up to the latest registration on 2020-03-12 11:13:30.

For all registered domains we could identify NameCheap, Inc. as the registrar based in Panama. The email address used for registering the domains was anonymized.

The registered domains could not be resolved to any hosting IPs throughout our analysis.

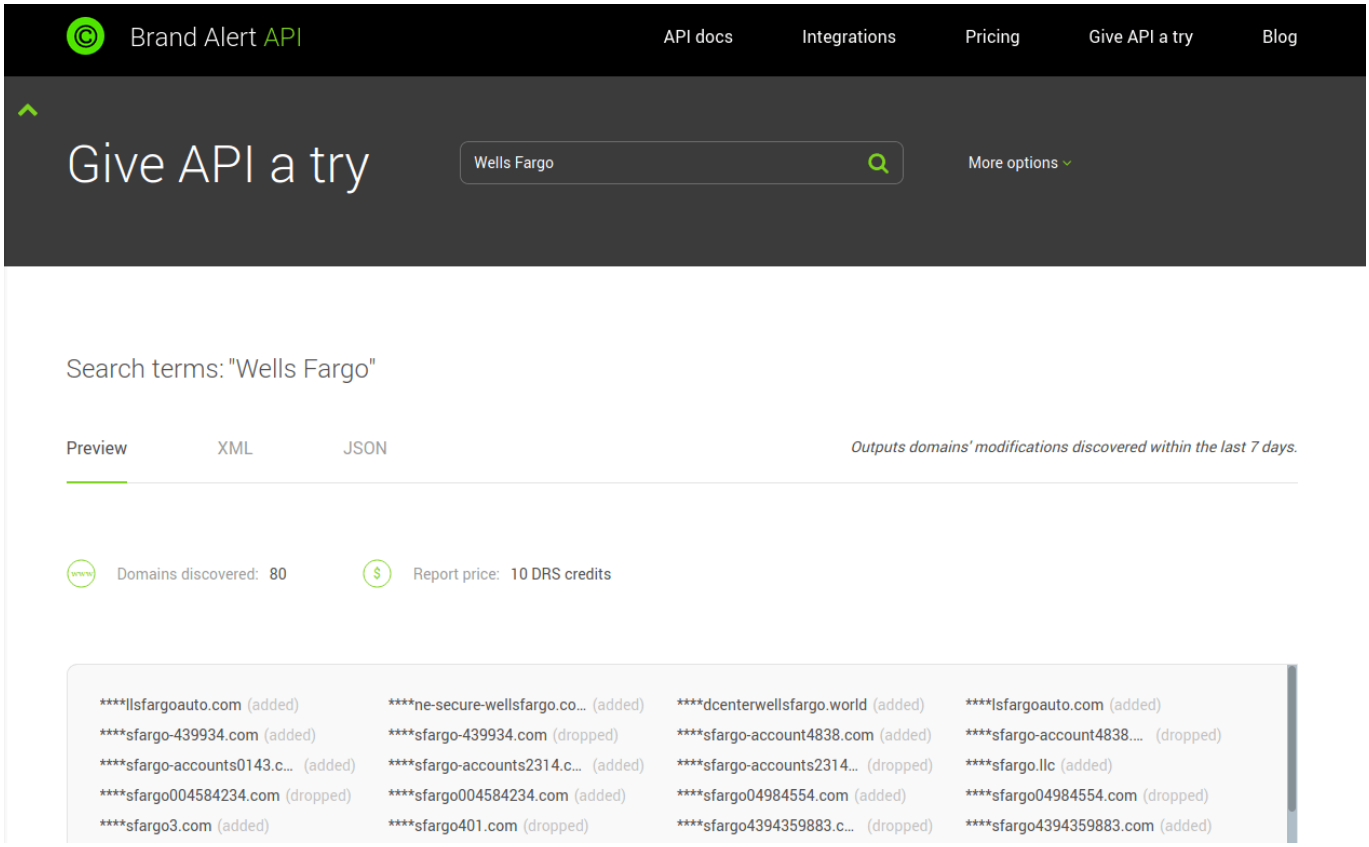
However the registrar NameCheap, Inc. covers a pool of 38.312.478 domains where at least 0.43% can be considered as potentially malicious."

To those who keep an eye on trends in IT security threats, notably phishing and typosquatting attacks, the name Wells Fargo is not unfamiliar, not even to those who have no business relation whatsoever with this multinational financial services company. In fact, all financial companies are likely targets for phishing campaigns, and Wells Fargo had TCPA settlement cases which are amongst the greatest attractors of these kinds of threats. So, rather unsurprisingly, there has been a continuous and significant malicious activity against this company.

So let us take a closer look at the present early warning. IBM nominates three domains:

- mobil-support-wells-fargo[.]work
- support-team-wells-fargo[.]work
- support-wells-fargo-team[.]work

With [WhoisXML API's](#) tools and data, we shall soon find that this is just the tip of the iceberg. There is tremendous activity going on around Wells Fargo in the domain name registration space, with numerous suspicious instances. As a first insight, let's use the Brand Alert API. Even without signing up for this service, we can just give the API a try by typing in "Wells Fargo" to the appropriate entry box. And here is what we find with the default settings:



We have as many as 80 domains that had been added or dropped in the last 7 days. And it is very unlikely for a domain named like "wellsfargo-account..." to serve benign purposes. Subscribing to this API opens even more possibilities: apart from providing complete reports, one can search for longer time spans, include typo variants, that is, misspelled versions of the search term, etc. In addition, this can be done programmatically via a RESTful API, that enables the integration of

the findings to e-mail filters, firewalls, etc., beefing up protective measures against such threats. Moreover, one can set up daily warnings about the registration activity to continuously follow what is going on.

But what if we want to proactively apply similar security measures but don't know the victim in advance? Alright, in the present example we were given "Wells Fargo" as a search term, but what if we just want to detect potential typosquatting and phishing campaigns in advance without any initial information? For instance, when operating a mail service for a community, we cannot know if one of our users is being tricked by the name of "Wells Fargo" or the name of some other bank... Also, although highly recommended, it is not always viable to have someone who reads IT security reports and keeps an eye on what is going on. Can we do anything in an automated manner? Yes, we can.

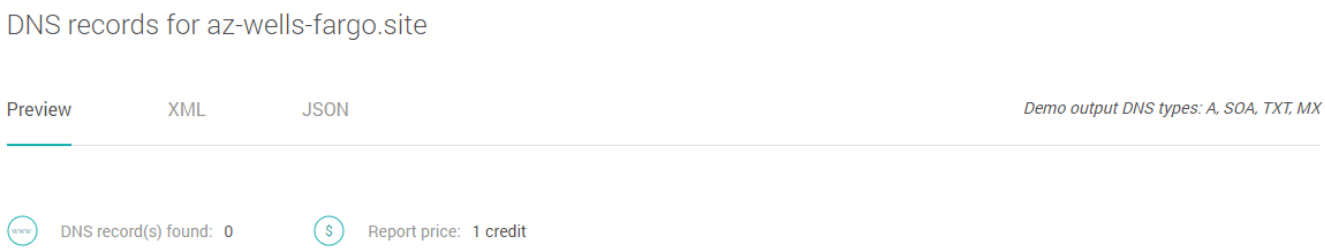
The right tool is the "Typosquatting Data Feed". The idea behind it comes from the observation that many malicious domains are registered in bursts, that is, many similarly-named domains (e.g. typo variants) on the same date. So if we had a list of all groups of domains that were registered on the same day, and their names were similar to each other within the group, we could pay particular attention to these. Even though some of them could be benign (as some companies register potential typosquatting domains themselves as a countermeasure to prevent the attacks in the argument), it is nevertheless better to be safe than sorry.

To get such a list, one needs, first of all, the names of all registered domains on a given day. Secondly, one needs sufficient mathematical expertise to be able to filter out "similar" domain names, and of course, enough computational resources to do the calculation regularly. WhoisXML API, Inc. has all these ingredients at your disposal, so such a list is now published every day as the "Typosquatting Data Feed". It lists up to a few ten thousand domains daily, up to a few thousand similarly named domains that meet the above criteria. This may sound like a big number at first, but when compared to the several hundred thousand domains registered a day, it is not really that big. Meanwhile, it illustrates that the trend of registering domains in bursts is indeed relevant. And it is a piece of cake then to check for a domain name's presence in such a list.

Let’s see what Typosquatting Data Feed says about our “Wells Fargo” case. Imagine that we don’t know anything about the earlier situation and are searching for a domain name, e.g. “us-wells-fargo[.].site”, because one of our users received an e-mail from it. Using “Wells” and “Fargo” as search terms, we find that they have appeared in a group of 3 domains on January 28, 2020; the other two were “on-wells-fargo[.]site” and “az-wells-fargo[.]site”. Note that the date is far before the early warning (i.e, March 12) which was our starting point. While none of the found domains are in the list of the early warning, they share a number of common features with those:

- Not resolving in the DNS
- Registered at NameCheap
- Registered in Panama

The feature can be also verified with the DNS Lookup API, and there are no DNS records for these domains, e.g.:



The other two features be found out by looking up the WHOIS record of the domains with the WHOIS API. A closer look of the WHOIS data will in fact give us an even more detailed insight. We find that the status of the XForce’s early warning includes “clientHold”. It explains why the domains do not resolve: this is an unusual status set by the Registrar and it reflects that the domain has some (possibly legal) issues. The domains we have just found in the data feed are on “serverHold”, a similar status set by the registry. Both prevent the domain from resolving in the DNS.

So indeed, although a domain listed in the Typosquatting Data Feed could be

benign, it is better not to trust it immediately before any verification. In total, the Typosquatting Data Feed lists a total of 39 domains which contain both the “wells” and “fargo” strings, so there could be several other examples listed when considering one string at a time.

Typosquatting and phishing are very harmful forms of cybercrime, and organizations dealing with sensitive information such as financial services companies are especially prone to it. For IT security experts it is a must to follow the trends reported, e.g., by IBM Xforce Exchange. Meanwhile, WhoisXML API provides data sources that can serve as foundations for efficient and automated proactive security measures, as illustrated in the case described above.