Detect Possible Domain Spoofing and Homograph Attacks with Typosquatting Data Feed

Posted on July 2, 2020

Charles Caleb Colton once said that imitation was the sincerest form of flattery. This proverbial expression finds its origins in the 19th century and other historical writings before that. What likely wasn't foreseen at the time, however, was that certain forms of imitation in the 21st century could give organizations terrible headaches. We are talking about domain spoofing and homograph attacks.

Imitators in our contemporary context can register one or several domain names highly similar to that of an established brand and use these to deceive people and trick them into sharing sensitive information or even transfering funds to fraudulent bank accounts.

Registering copycat domain names of known brands and organizations isn't the only way to fool victims, though. At the height of coronavirus-themed attacks, the Typosquatting Data Feed proved useful in spotting potentially dangerous footprints containing thousands of domain names with word strings such as "covid" and "coronavirus" combined with "mask," "vaccine," "donation," "lawsuit," and plenty of others.

In this post, we put the feed's capabilities to the test to detect spoofed domain names, including Punycode domains, that could be used to abuse employees, customers, and other parties who regularly interact with Lloyds Bank and Apple. We will also show how other sources of intelligence can help learn more about possible impersonators and the infrastructure they use.

What Are Punycode Domains? How Can They be Used in a Homograph Attack?

Punycode domains are those that use Unicode or non-Latin characters. Their introduction was allowed as part of an effort to internationalize the Domain Name System (DNS). As a result, internationalized domain names (IDNs) let people worldwide use their local languages and scripts in domains.

IDNs are formed by using characters from scripts like Arabic, Chinese, Cyrillic, or Devanagari. But for systems to read them and point users to the right websites, these are encoded in Unicode based on approved IDN protocols.

An example would be ??[.]com. ?? is "apple" in English. In Unicode, ??[.]com translates to xn--

btvx9d[.]com. So, if Apple were to register the domain, Japanese users could easily use the local version of apple[.]com to reach the vendor's website.

While IDNs were allowed for a good cause, cybercriminals have learned to use these for their illicit gain. Malicious domains whose names include a mix of Latin and non-Latin characters are often used in phishing and other homograph attacks. Examples of such domains would include ?icrosoft[.]com (xn--icrosoft-93d[.]com) and ?nstagram[.]com (xn--nstagram-f80d[.]com).

2 Potential Domain Spoofing and Homograph Attacks Under Study

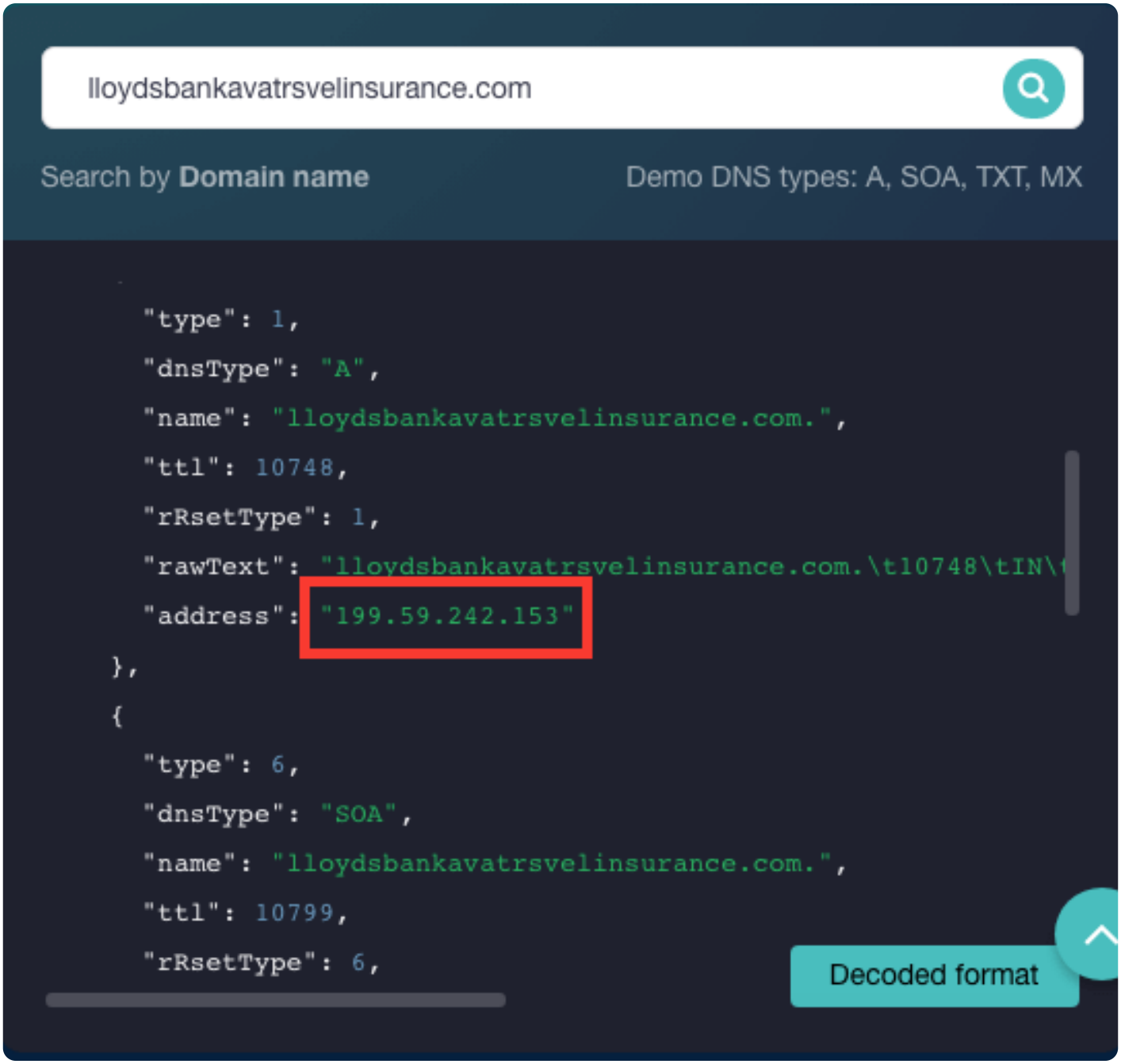
With a greater understanding of what a Punycode domain used in a homograph attack may look like, let's review some real-life examples where lookalike domain names—both Latin-alphabet and IDNs—might be confused with organizations' legitimate domains. These analyses also include the use of cyber threat intelligence tools to study the suspicious online properties further.

Case #1: Lloyds Bank

IBM X-Force warned users on 6 May 2020 about an ongoing Lloyds Bank cybersquatting campaign. The vendor advised the bank's customers to steer clear of 68 domains and two nameserver addresses.

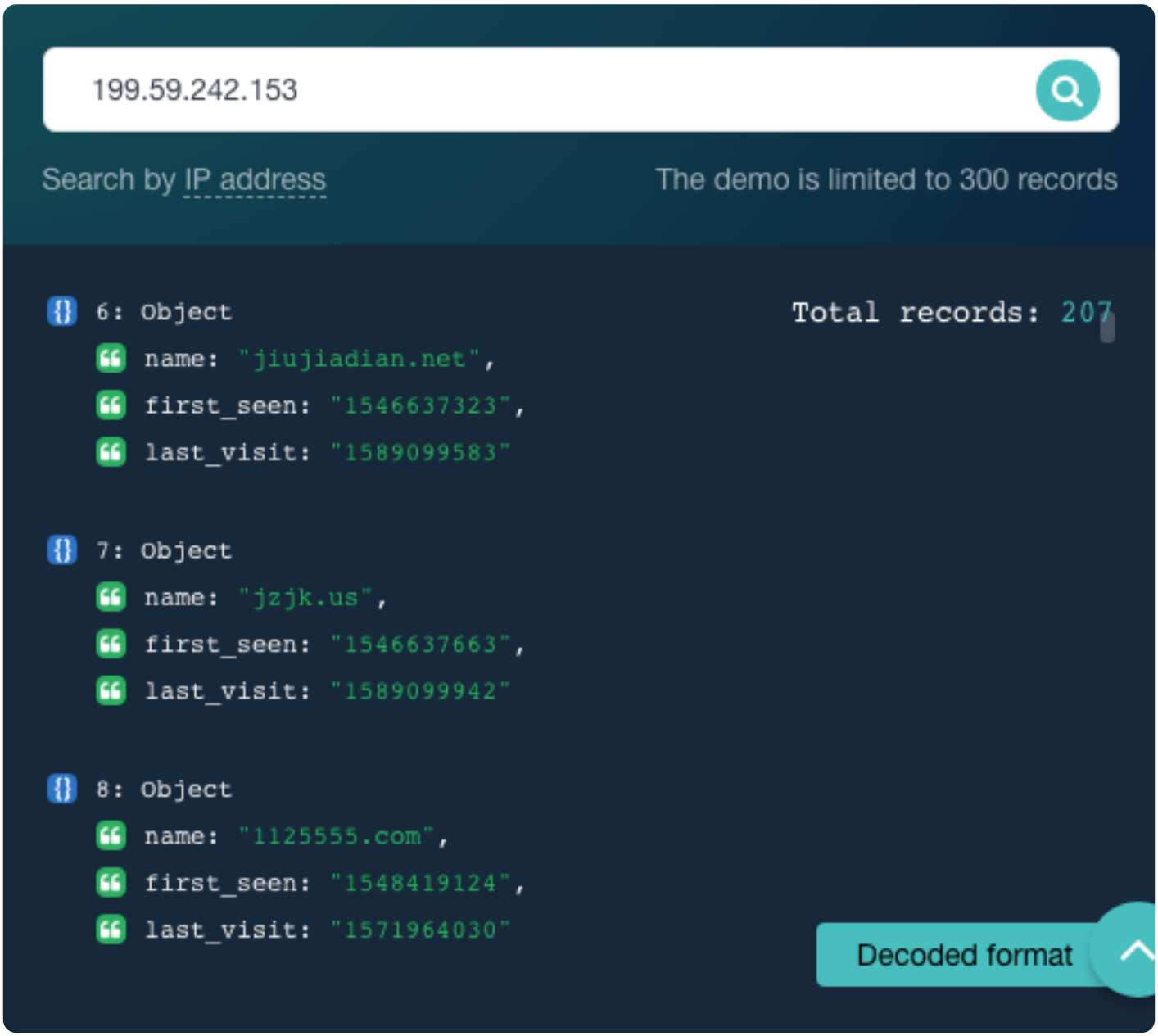
The Typosquatting Data Feed detected several of the malicious domains shortly after they appeared in the DNS. We also found 64 more domains registered between 9 November 2019 and 30 April 2020. Apparently, Lloyds Bank did not own any of these domains (based on a Bulk WHOIS Lookup query we conducted).

Looking more closely at one of these domains (lloydsbankavatrsvelinsurance[.]com), we found that it was registered on 28 April 2020 and was flagged as potentially dangerous in our threat intelligence feeds. DNS Lookup API showed that it resolved to the IP address 199[.]59[.]242[.]153.



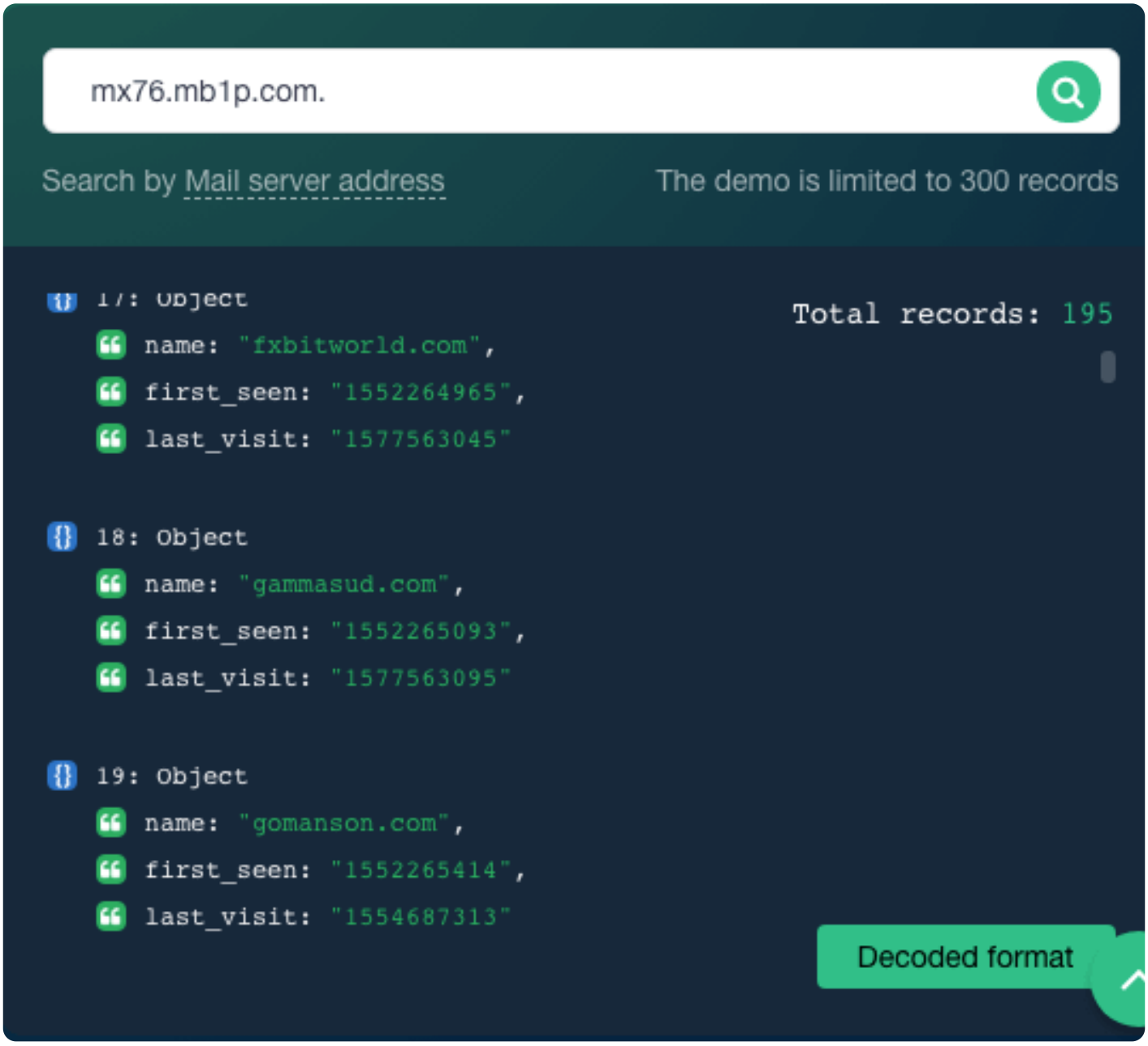
We then looked into the other domains that resolved to the IP address at one point and got a list of 207 results. Of these, 1125555[.]com was cited for malicious activity, too (as per a Threat

Intelligence report).



Lloydsbankavatrsvelinsurance[.]com's DNS lookup results tell us as well that the domain is hosted on the nameserver ns1[.]bodis[.]com[.]—one of the indicators of compromise (IoCs) indicated in the IBM X-Force report. We also saw that it was connected to the mail server

mx76[.]mb1p[.]com[.], which hosts 195 domains, including fxbitworld[.]com, which has malicious ties.



Apart from the domains mentioned above, we harvested 158 others that had misspellings, and top-level domains (TLDs) that differed from those used by the target organization.

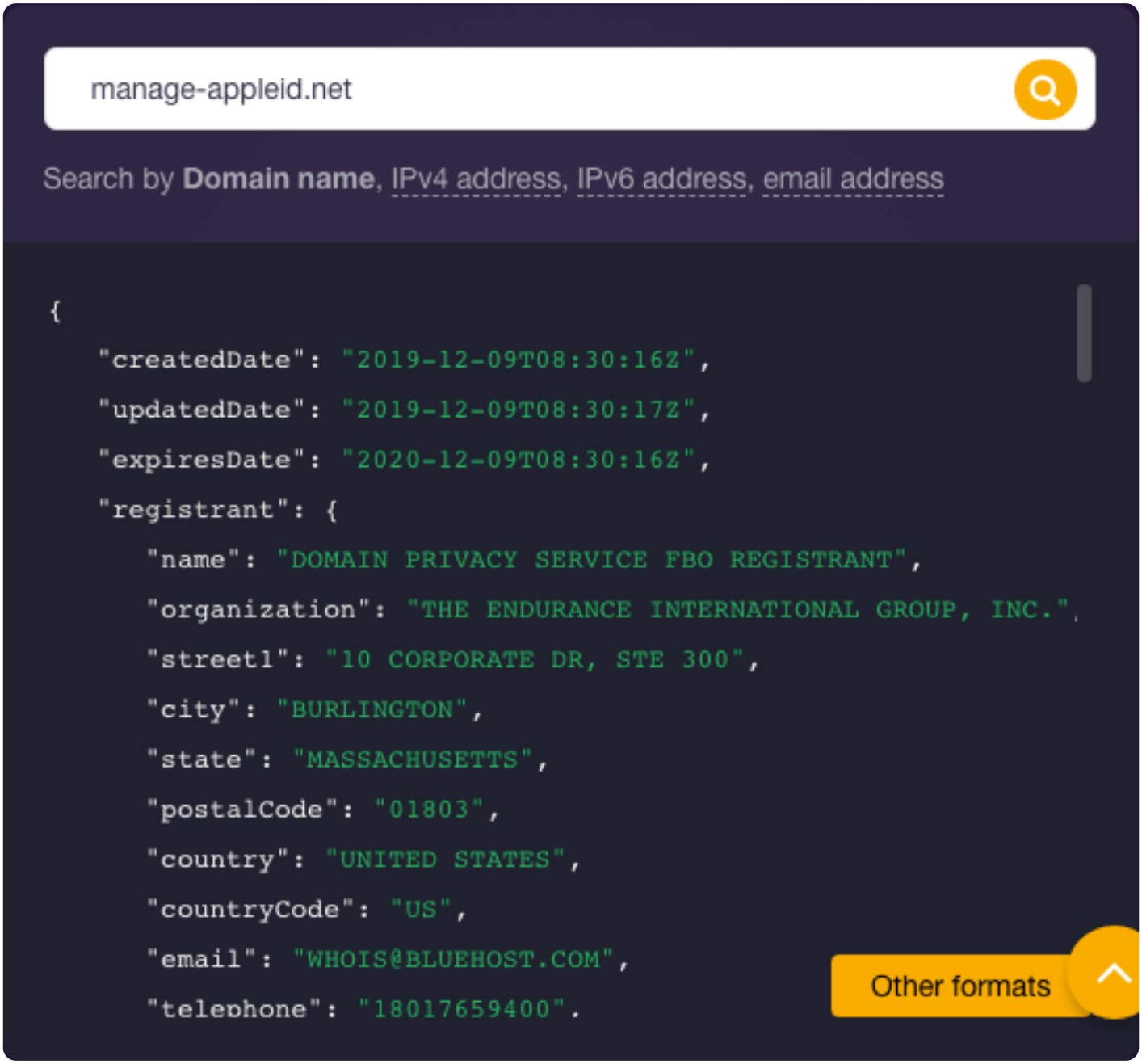
Of the 158 domains, we checked a Punycode one (xn--lloydsbnk-ccb[.]com). In Unicode, this translates to lloydsb?nk[.]com. Notice the macron or line on top of the vowel “a.” While the domain has no detected malicious ties, a check in our WHOIS database for data about its owner revealed that it is not likely owned by Lloyds Bank PLC.

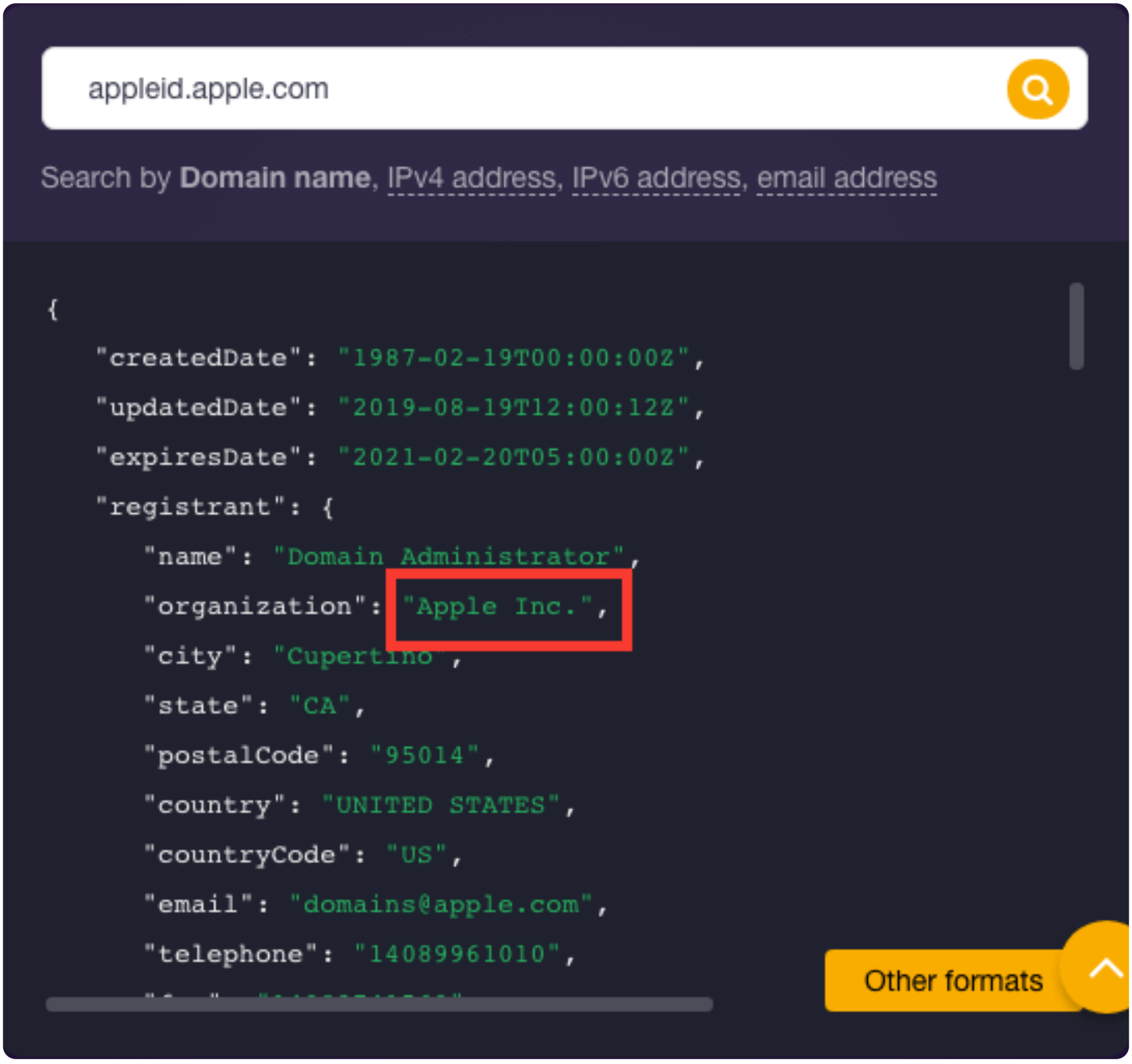
Case #2: Apple

On 29 April 2020, IBM X-Force also reported seeing an AppleID typosquatting campaign targeting media practitioners. The vendor warned AppleID users to stay away from three malicious domains.

In parallel, Typosquatting Data Feed found 160 more domains containing the term “appleid” registered in bulk between 1 October 2019 and 3 April 2020. Of these, 45 were cited for malicious ties such as appleidinformation[.]com, customer-support-appleid00145[.]com, and appleid-doom[.]business.

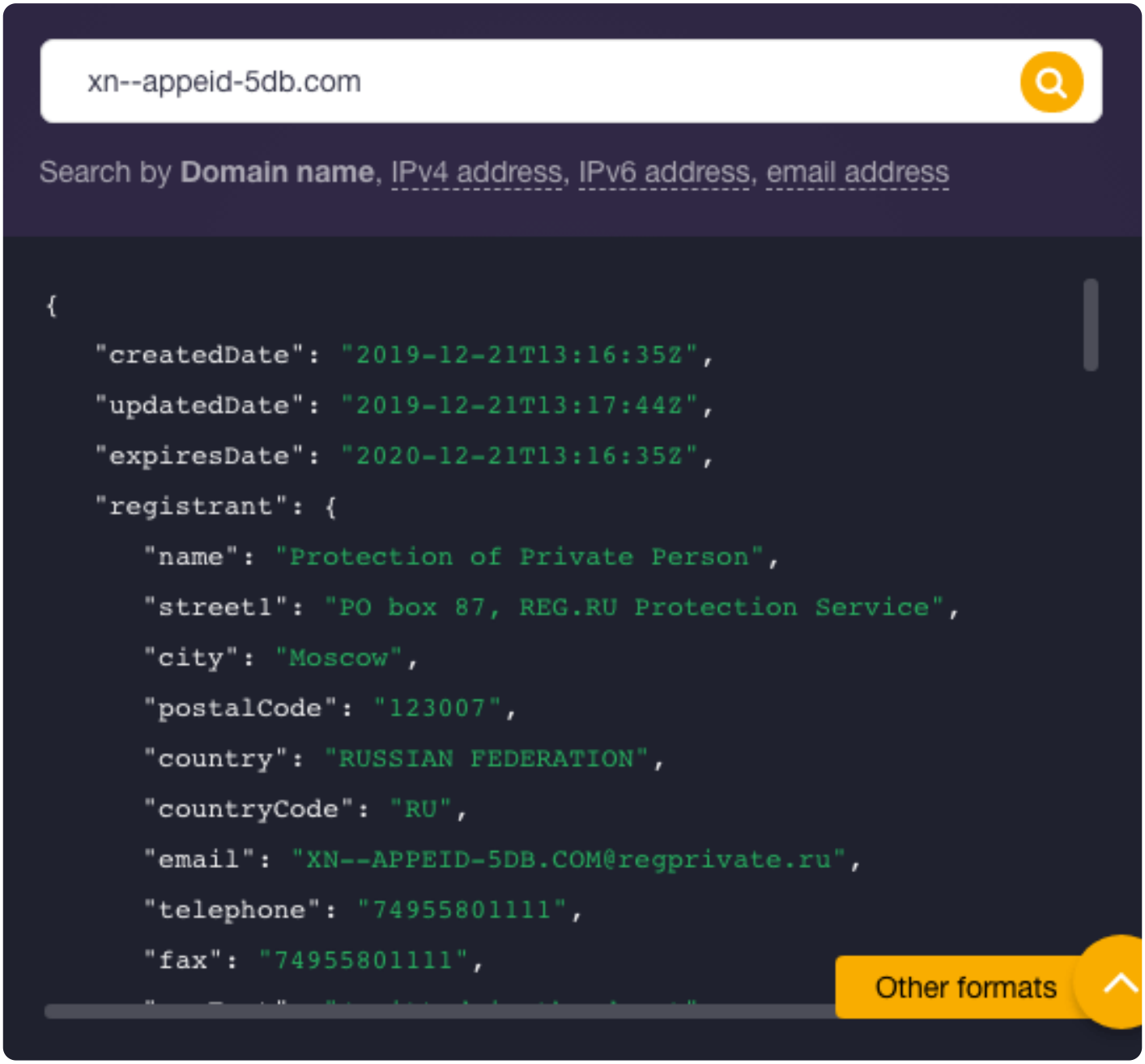
A Bulk WHOIS Lookup query revealed that none of the 160 lookalike domains seemed to belong to Apple. Considering one domain name (manage-appleid[.]net), in particular, we found in its WHOIS record details that it was registered on December 9. Registration information turned out to be heavily redacted and differed from that of the legitimate AppleID domain—appleid[.]apple[.]com—which clearly indicates “Apple Inc.” as its registrant.


manage-appleid.net
Search by Domain name, IPv4 address, IPv6 address, email address
{
"createdDate": "2019-12-09T08:30:16Z",
"updatedDate": "2019-12-09T08:30:17Z",
"expiresDate": "2020-12-09T08:30:16Z",
"registrant": {
"name": "DOMAIN PRIVACY SERVICE FBO REGISTRANT",
"organization": "THE ENDURANCE INTERNATIONAL GROUP, INC.",
"street1": "10 CORPORATE DR, STE 300",
"city": "BURLINGTON",
"state": "MASSACHUSETTS",
"postalCode": "01803",
"country": "UNITED STATES",
"countryCode": "US",
"email": "WHOIS@BLUEHOST.COM",
"telephone": "18017659400",
Other formats



The 160 domains mentioned above were not the only ones on our feed. We harvested another 115 domains that had typos and TLDs other than .com, which was customary of Apple's legitimate domain name.

We looked at a Punycode domain (xn--appeid-5db[.]com) on the additional list. In Unicode, the IDN translates into app?eid[.]com. The owner substituted the “l” in “apple” with a slashed lowercase “l,” which is the Polish character for the uppercase “L.” A threat intelligence query for the domain didn’t reveal any malicious connections but a WHOIS API query for the lookalike domain revealed that Apple Inc. was not its registrant. That said, users looking to communicate with the vendor should be wary of visiting the IDN.



Domain spoofing and homograph attacks are a concerning form of imitations and not that all flattering for established organizations with millions of customers or users to protect. Daily data monitoring of potential typosquatting domains using Typosquatting Data Feed, in addition to footprint expansion via domain and IP intelligence tools, may help achieve overall protection and

enhanced cybersecurity against such threats.