



A DEEPER LOOK AT LOOKALIKE ATTACKS

NEW STUDY REVEALS
LATEST THREAT VECTORS

April 2023





LOOKALIKE DOMAINS TARGET EVERYONE

TABLE OF CONTENTS

EXECUTIVE SUMMARY	3
BACKGROUND	5
Homographs (née Homoglyphs)	6
Typosquats	7
Combosquatting	8
Soundsquatting	9
Other Forms of Lookalikes	10
EVERYONE IS A TARGET	11
They Target Us!	12
They Target Employees	14
They Target Do-Gooders	16
They Target Crypto	17
They Target Social Media and Mobile Users	20
They Target Everyone	22
HOW ARE LOOKALIKES USED?	23
They Send Texts	24
They Use Old-Fashioned Phone Calls	27
They Send Spam	28
They Use QR Codes	30
They Use DNS	31
WHY ARE THEY EFFECTIVE?	34
Psycholinguistics	35
Punycode support: hits and misses	36
To err is human	38
INFOBLOX SOLUTIONS	39
REFERENCES	40

EXECUTIVE SUMMARY

Threat actors have used visually similar domains to deceive users into visiting malicious websites since the advent of the internet. These domains, called lookalike domains, are so synonymous with phishing attacks that security awareness training includes learning to inspect links for them.


However, in spite of awareness campaigns and advances in technology, lookalike domains represent a persistent threat to consumers and organizations, one that actors continually adapt. Everyone is a target; from consumers to governments, from major retail brands to small restaurants, from world-renowned technology companies to lesser-known ones like ours. In this paper, you'll see that "everyone is a target" with examples of real domains and campaigns. As a modest-sized company in a fairly niche industry, even we are targeted.

This report describes the current threat landscape by showcasing real world examples across industries and user groups. Infoblox has been detecting lookalike domains for years and analyzes over 70 billion domain name system (DNS) events daily to find new and potential threats. For this paper, we focused on detections from January 2022 to March 2023. From over 300,000 lookalike domains, we've curated a set that highlights the challenges and risks associated with these attacks.

Lookalike domains are often associated with broad, untargeted attacks on consumers through email spam, advertising, social media, and SMS messages. Every day there are thousands of new domains registered that mimic popular software, financial institutions, and package delivery services. Phishing attacks that aim to steal user credentials or infect machines with malware are so prevalent, and often so unsophisticated, that they have become a source of numerous memes including "can't fall for phishing scams, if you don't check your email." While often portrayed as comical, phishing is a serious industry. The Anti-Phishing Working Group (APWG) reports phishing reached a record level in the third quarter of 2022.¹



All indicators in this paper have been defanged, regardless of their status as malicious or legitimate. We have defanged the indicators by placing brackets around the periods [.] and thereby preventing it from becoming a clickable link.



70+
BILLION

Infoblox analyses over 70 billion DNS events daily to identify new threats.

300K+

lookalike domains, have been curated for this report to highlight the challenge and risk of these attacks.



AN EXAMPLE OF A PHISHING MEME.

One example is this tweet from 2019.²

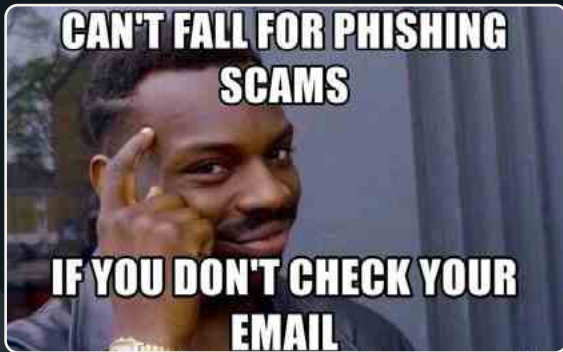


Image credit: The origin of this meme is unknown.

But lookalike domains are not just a threat to consumers — they are used to gain access to corporate networks.

Recent disclosures have revealed targeted attacks in which malicious actors deceived employees into providing their multi-factor authentication (MFA) credentials. In most cases, the lookalike domains not only mimicked the company but also included MFA keywords, further enhancing the illusion for the employees that the connection was secure. We found that actors have targeted businesses large and small, across many verticals, including internet service providers, banking and cryptocurrency, software and services, and insurance companies globally. These attacks began in early 2022 and gained momentum over time.

The use of lookalike domains is profitable because it is an asymmetric attack. Users must be ever vigilant to protect their personal finances and the information of their employers. Cheap domain registration prices and the ability to distribute large-scale attacks give actors the upper hand. Attackers have the advantage of scale, and while techniques to identify malicious activity have improved over the years, defenders struggle to keep pace.

Not only is lookalike phishing thriving, but the use of lookalikes has become more complex in a way that is revealed most clearly in DNS records. Our research shows that lookalike domains are being leveraged beyond the traditional phishing and typosquatting purposes. They are also being used in ways not previously reported: for example, as nameservers and for spear phishing mail distribution. There are large resilient networks that serve only lookalike domains and that are targeting both consumers and government employees.

Infoblox has multiple algorithms to identify lookalike domains. We use a combination of methods, including: watching for variants of common targets in the shopping, banking, software, and financial sectors; watching for variants of customer-specified domains; and watching DNS infrastructure actors who specialize in lookalike domains. This multifaceted approach gives us broad coverage of the threat landscape.



IMPORTANT NOTE: *This report contains a number of examples that illustrate the breadth and depth of lookalike domains in the wild; they are not intended to imply successful attacks or breaches of any entity.*

BACKGROUND

Like all good research papers, we're going to start with some background information. This is mostly vocabulary. We know most readers skip the background section, so we've kept it short.

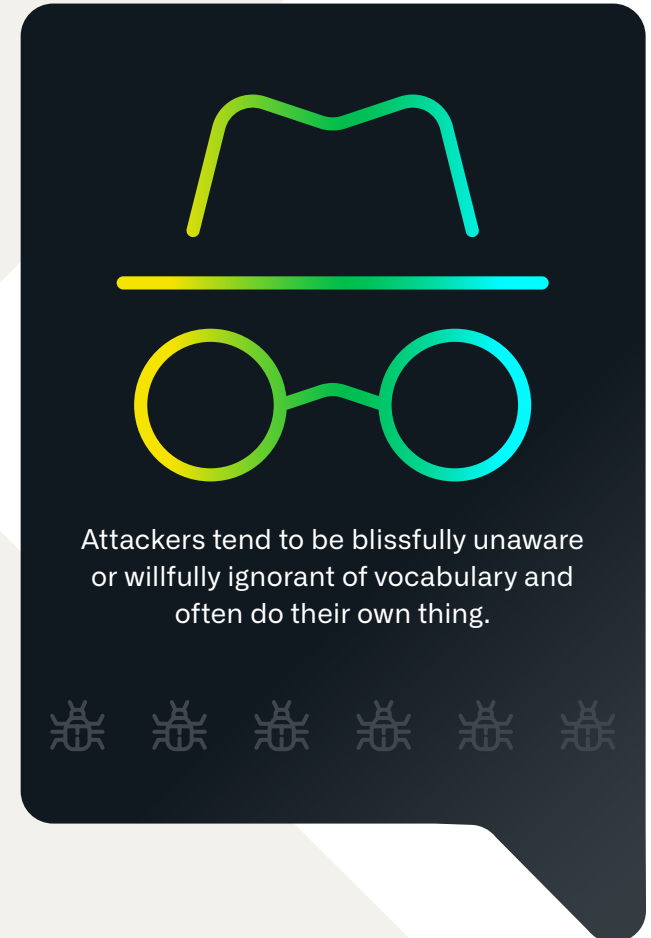
Malicious lookalikes — attacker-registered domains that look the same or very similar to a known domain — are a well-known, persistent threat in the cyber landscape. Generally speaking, lookalikes have both offensive and defensive applications. In the offensive sense, lookalikes are used for deceit wherever there could be human eyes. Actors use lookalikes to steal money, gain credentials or access, gather personally identifiable information, distribute malware, or earn ad revenue. They are also used for political purposes and to tarnish brand reputation. In short, they are a means to an end for cybercriminals. In the defensive sense, many organizations proactively register domains similar to their own to prevent attackers from claiming and using them.

Lookalikes take different forms. In the DNS space, domains can be:

- **Homographs**
- **Combosquats**
- **Typosquats**
- **Soundsquats**

They can be almost indistinguishable from the original target domain or objectively quite distinct. Much of the success of lookalike domains as an attack vector is due to the burden placed on individuals.

As we'll see, lookalikes can be found in every element of an attack, from email sender addresses, to phishing URLs, and malware command and control (C2). Although usually associated with address records (A/AAAA), we have even found lookalikes used for nameserver (NS), pointer (PTR) and canonical name (CNAME) records. They can be deployed through emails, SMS or text messages, compromised websites, malvertising networks, and phone calls. In the following section, we briefly describe the different forms of lookalikes and give examples of each.



BLAME THE TYPEWRITER

In fact, this modern issue can be traced all the way back to the early days of typewriters. On many older typewriters, there were no 0 or 1 keys, as typists were expected to use capital letter O's and lowercase letter L's to represent these digits.⁴

HOMOGRAPHS (NÉE HOMOGLYPHS)

Although the word homograph in English means “two words that are spelled the same, but not necessarily pronounced the same, and having different meanings,” the term homograph has been used for many years in security research literature to mean “two domains that appear visually the same.”³ A more accurate term is homoglyph. These domains look similar to one another and in some cases may be nearly indistinguishable. *For consistency with the research literature, we'll use the incorrect term homograph in this paper.*

This form of lookalike takes advantage of the fact that many characters in the same character set, or alphabet, look similar to each other. For example, 0 (the digit zero) and O (capital letter “o”), or “l” (lowercase letter “L”) and “I” (capital letter “i”). Some fonts accentuate this issue further. Classic examples of this are g0ogle.com and Infoblox.com, in which the “o” in Google is replaced with a zero (0) and the “i” in Infoblox is replaced with a lowercase “L,” respectively.

As the internet matured, and more non-English speakers began to log on to the World Wide Web, the need for internationalized domain names (IDNs) grew. An IDN is a domain that contains at least one character in non-Latin script; the introduction of Unicode enabled the rise of such domains. With IDNs came a new form of lookalike: the IDN homograph. It is still a homograph, but one that uses characters from other character sets or alphabets that look similar. Gabrilovich and Gontmakher showed the power of IDN homographs in their 2002 paper “The Homograph Attack.” The authors registered a lookalike of the authentic Microsoft domain microsoft[.]com which contained the Cyrillic letters “с” and “о.”⁵ The end result is a domain www.microsoft[.]com that is visually indistinguishable from the authentic Microsoft domain.

The Unicode Consortium has published a tool showing the vast number of confusable characters available for a given string.⁶ The string “hi” has 684 variations with Unicode characters; for a string like “infoblox” the number balloons to over 2.2 trillion variations. Some variations are less effective for a lookalike than others. For example, the Unicode Consortium lists “٥” (extended Arabic-Indic digit five) as a potential confusable character for “o” (Latin lowercase letter “O”).

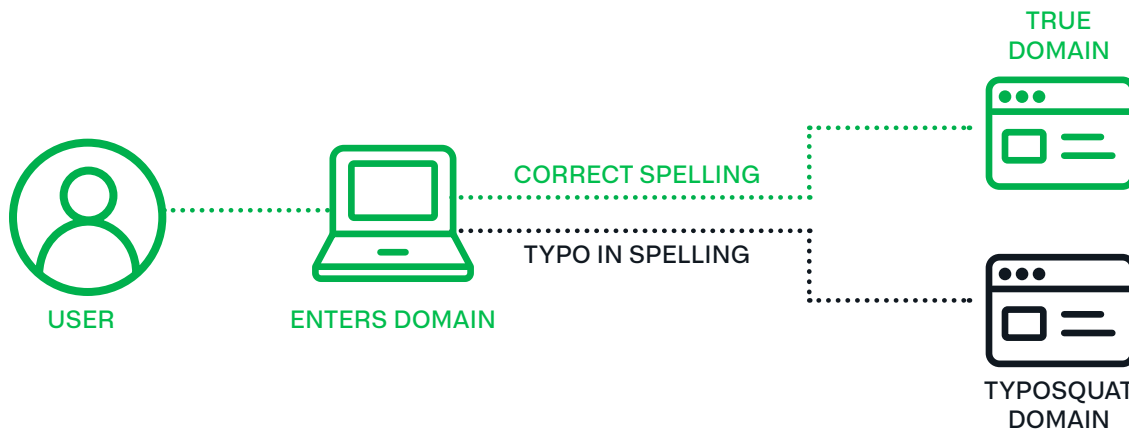
Clearly, inf٥blox[.]com is not a very effective lookalike; but can you tell the difference, when shown in the commonly-used Arial font, between the proper domain {infoblox[.]com} and {infoblox[.]com} (containing a Byelorussian or Ukrainian lowercase “i” and Armenian lowercase letter “vo,” written as “n”)? We can't either.

TYPOSQUATS

Typosquat domains capitalize on popular domain names and typing errors that users make, or that are caused by typing on broken keyboards. This term is usually associated with domains registered, but left unused, for the purpose of drawing advertising money. For example, one of the authors was recently trying to pay rent via their property management group's online portal, hosted via appfolio[.]com (a well-known software company that offers SaaS solutions to property management groups and landlords). Instead, they fat-fingered and almost visited appfollio[.]com, which was registered in 2013 but is currently parked.

Interestingly, another apparent typosquat domain for Appfolio, apfolio[.]com, seems to be owned by Appfolio. It redirects to the proper domain and has the same registrant, registrant organization, and registrar, and was registered just one month after the legitimate domain appfolio[.]com. This is an example of the defensive usage of lookalikes. Unfortunately, bad actors have the upper hand because there are simply too many possibilities for organizations to register all lookalike variations.

Typosquats are primarily perceived as a monetization method, but they can have a nefarious purpose. While they are used to sell 3rd party advertisements or to sell to the legitimate domain owner, they can also be used for “blackhat” affiliate marketing programs and as malware C2 domains, as we'll show later. Brands and companies do have civil protection against typosquatting under the Anticybersquatting Consumer Protection Act. Because of this threat of legal action, typosquatting is seen as a “blackhat” form of monetization in the domain flipping/parking community, and serious domain flippers such as iGoldrush recommend against typosquatting for profit.⁷



TYPOSQUAT EXAMPLES

gikthub[.]com
5whatsapp[.]com
Hdfcbank[.]vip
royalbsank[.]com
spartybet[.]city
bangkokbank[.]com
1337x[.]asia
moneycont5rol[.]com



of abusive combosquatting domains are active for more than 1,000 days



of abusive combosquatting domains appear on at least one public blocklist 100 days after initial resolutions

COMBOSQUATTING

Combosquatting is a form of lookalike that combines popular brand or company names with other keywords. Terms like support, help, security, and mail are common. Consider, for example, `wordpresssupport[.]ru`, `wordpresssupport[.]store`, and `wordpress-security[.]cloud`. These domains are all hosted on the same, Russia-based IP address and look like WordPress, the popular web content software. The inclusion of support and security in the domain name indicates that these are intended for WordPress users. They might be used to gather credentials to hijack WordPress sites or collect payment and personally identifiable information (PII) details.

In addition to generating combosquat domains themselves, actors also have the ability to use dictionary domain generation algorithms (DDGAs) to create lookalikes. In seconds, thousands of candidate domains can be generated for a multitude of brands or companies. By sheer luck, the algorithm can create candidate domains with just the right keywords for the domain to be effective. The user community of Steam, a top gaming platform, is a common target for actors using combosquat DDGAs. Some examples of domains within a recently observed set are: `steamcommiunity[.]com[.]ru`, `steamcommucnity[.]com[.]ru`, `steamcommunityjp[.]top`, and `steamcommunityiq[.]top`. Note the overlap between typosquatting and combosquatting in this domain set.

Kitsin et al. performed a longitudinal study of combosquatting in 2017, analyzing about 468 billion DNS records (sourced from both active and passive datasets), and found disturbing results:

- **Combosquat domains are 100 times more prevalent than typosquatting domains**
- **60% of abusive combosquatting domains are active for more than 1,000 days**
- **20% of abusive combosquatting domains appear on at least one public blocklist 100 days after initial resolutions**
- **Combosquat domain resolution increased year-over-year⁸**

We concur with the authors' finding regarding the prevalence of combosquat domains. We find more combosquat domains than we do pure typosquats or pure homoglyphs (IDN or otherwise) through our analytics.

SOUNDSQUATTING

Soundsquat domains leverage the use of homophones, words that sound the same but have a different spelling. Soundsquatting is the most recently identified form of lookalike, first appearing in the literature in 2014.⁹ Soundsquatting has gained more attention from researchers recently due to the proliferation of smart speakers á la Alexa, Siri, and Google Voice.¹⁰ Soundsquat domains overlap with other lookalike domain types, in that they might both sound and look similar. We have found pure soundsquatting domains, that is ones that don't appear visually similar but sound alike, to be rare; generally these domains can also be found by text-based similarity techniques.

It is important to note that lookalikes in the wild often do not fit into neat buckets as we've laid out here. A combination of forms is used to maximize the effectiveness of a lookalike domain. Many of the combosquat domains we see have elements of typosquats and homographs (IDN or otherwise). Typosquats utilize elements of homographs, soundsquats utilize elements of typosquats, and so on. The end result is an asymmetric threat landscape in which attackers can leave defenders gasping for breath.



SOUND THE ATTACK

The prevalence of soundsquatting has taken off with the advent of voice activated technology like Alexa, Siri and Google Voice.



OTHER FORMS OF LOOKALIKES

While the focus of this paper is on lookalike domains and their role in the current threat landscape, other types of lookalikes exist that can exploit vulnerable users. One notable example of these was recently found in Python PyPi packages.



<https://infosec.exchange/@tweedeg@cybersecurity.theater/109846797159938702>

Package managers for popular programming languages such as Python are subject to the same weaknesses that domains are. Anyone can upload a package with any name (so long as that name is not already taken) containing code that may or may not be free of security risks. In 2016, security researcher Nikolai Tschacher employed typosquatting in this manner to force more than 17,000 distinct hosts to execute arbitrary code.¹¹ Then, in 2021, security researcher Alex Birsan took Tschacher's idea and expanded upon it, coining the term "dependency confusion."¹²

Birsan found major companies' private, internal package names through various open sources. This included exploring source code on websites, hunting for packages on GitHub, or even finding package names on public forums. Then, he uploaded packages with the same name as private, internal packages to public package managers. Finally, Birsan utilized automated CI/CD pipelines, "confusing" the public packages for the private, internal packages. Rather than importing and installing the private packages, the automated pipelines found and imported Birsan's public packages instead. Birsan then used DNS exfiltration to notify him that his arbitrary code, and not the intended, private package, had been executed. Birsan's lookalike technique allowed him to breach 35 organizations, sometimes within hours of uploading his packages.

Regardless of the type of lookalike or the bailiwick in which a lookalike is used, lookalikes are a persistent threat. Part of the challenge of studying lookalikes is that they are undefined — there are more possibilities than can be computed and everything is a target. In the following sections, we show specific examples of these various forms of lookalikes in the wild, including targets, deployment methods, infrastructure, why they are effective, challenges, and Infoblox's solutions to the problem.



EVERYONE IS A TARGET

We think you'll find at least one surprising target in our examples.

One of the most powerful findings from our review of lookalike domains in DNS was that everyone is a target: we found lookalikes for all the expected targets, but also for smaller companies and services. These domains are used by bad actors to prey on individuals at work and at home.

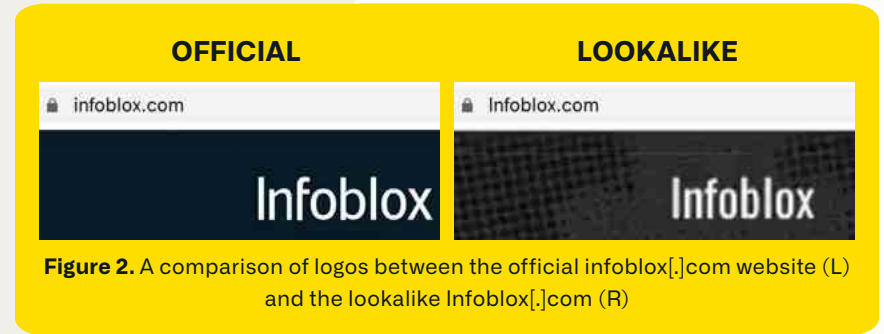
As Akamai recently noted, most lookalike campaigns only get press once a large target is affected.¹³ Our aim is to shed light on those underreported and overlooked targets alongside the “typical” targets. A few select examples are shown here to demonstrate this point, but we will also highlight the impact on different industries and the use of various methodologies in more detail later.

THEY TARGET US!

Infoblox is a modest-sized company with fewer than 2000 employees worldwide.

While we have a large share of the DNS, Dynamic Host Configuration Protocol (DHCP), and IP Address Management (IPAM) market—collectively known as DDI—this industry is fairly specific, and Infoblox is hardly a household name. One might be surprised that malicious actors would even be aware of us, much less that they’d actively target us with lookalike domains. Nevertheless, we found many domains designed to fool both our employees and our customers. Lookalikes of internal services, including our benefits portal, as well as our product names have been registered in the past year.

Some registered domains that are not owned by Infoblox include:



Homograph infoblox[.]com

Using a lowercase “L” to impersonate a capital “i” was registered in July 2022, and although it is offered for sale, the site shows in the upper left corner a rendering that is almost indistinguishable from that on our corporate website. See a comparison in Figure 2.

Typosquat infobloxbenefits[.]com

This domain was registered in China in April 2022 and is a slight typo from our employee benefits portal. This domain is currently parked with Bodis.

TLD Squat infoblox[.]info

Different top level domain, or TLD was registered in August 2022 through the highly abused registrar Sav[.]com. It is parked on dan[.]com, which allows users to sell domains.

Combosquat infobloxgrid[.]com

A combosquat lookalike to our flagship on-prem product used by thousands of customers around the world. Our patented Grid technology enables network administrators to combine diverse network applications into one single system. This domain is also available at dan[.]com and was registered in April 2022.

Combosquat infoblox-updater[.]com

An example of the technique of using common software words within the domain like “update” or “support.” In this case, a customer may be deceived into connecting with a false system thinking it was related to Infoblox system updates. Names or products of technology companies are frequently leveraged for this type of combosquat domain, which might be used as a phishing domain or as malware C2. Other examples include dev[.]gitlabs[.]me and jira[.]atlas-sian[.]net, both used by the advanced persistent threat (APT) actor Iron Tiger in their SysUpdate malware.¹⁴

In addition to targeting small technology companies like our own, we've seen a wide range of lookalikes that are deceptive variants of restaurants, law firms, and other small businesses.

Moreover, a single actor may use both well-known brands and small businesses as lures.

One actor that Infoblox has been tracking for some time has created lookalike domains for the New York City restaurant Cotenna and copied their website, presumably to lure visitors to make online reservations with their credit cards.¹⁵ The site cotenna[.]nyc was registered in April 2022 and is a lookalike to the restaurant website cotenna[.]com. This same actor has lookalike domains targeting large social media companies like Twitter.

In the sections that follow, we'll go into more depth about the industries that are most prevalently targeted today, as well as some of the many ways domains can be used for a successful attack. Because everyone is a target, we will highlight those areas in which we've seen the most malicious activity, based on a review of 300,000 lookalike domains.



LOOKALIKE DOMAINS TARGET EVERYONE

américafirst[.]com
instagram[.]dev,
caterpillarespaña[.]com
steamcommuntly.net[.]ru
boatairbuds[.]in
secure1-scotiabank[.]com
saveukraine[.]xyz
expressvpn-app[.]com



10K+ ORGS

In July 2022 Microsoft warned that over 10,000 organizations were the target of AitM attacks designed to steal MFA credentials from users in real time.

1,600+

Our research found over 1,600 domains contained a combination of corporate and MFA lookalike features.



THEY TARGET EMPLOYEES



Until recently, many corporations felt that the use of multi-factor authentication (MFA) protected their internal networks from phishing attacks.

But early in 2023, Coinbase revealed that their employees had been targeted by spear phishing attacks that used lookalike domains to the company's internal MFA login.

This reveal was quickly followed by corroborating reports from other companies who had been targeted by similar attacks. Based on victim reporting, we know that malicious actors sent employees SMS messages, as well as emails, urging them to sign into internal systems. In some cases, phone calls were also involved, during which the attacker provided a domain name for the employee to visit in their web browser. The attackers used adversary-in-the-middle (AitM) techniques to reassure employees that they were interacting with the company's real network. Employees were prompted for an MFA code, which was then captured by the attacker and used to gain access to internal systems.

Microsoft had warned in July 2022 that over 10,000 organizations were the target of AitM attacks designed to steal MFA credentials from users in real time.¹⁶ Those attacks were specific to the use of Outlook 365 authentication, but Microsoft further reported in February 2023 that a phishing kit enabling MFA attacks was available for sale in July 2022 and was widely used.¹⁷ Other companies, including Twilio, had disclosed similar attacks in the Summer of 2022, but the breadth of attack wasn't well publicized until the Coinbase revelations.¹⁸

To investigate this incident, we performed a retrospective analysis of lookalike domains that mimicked MFA by using keywords like "mfa," "okta," and "2fa." Our research found a wide array of targets and a distinct uptick in activity starting in July 2022, although there were a significant number of lookalike domains utilized for these attacks earlier in the year. Over 1,600 domains contained a combination of corporate and MFA lookalike features. Targets ranged from the reported large corporations like Coinbase, Reddit, and Twilio, to major banks, software companies, internet service providers, government entities, and gaming platforms worldwide. Also targeted, but underreported, were smaller technology companies, grocery stores, and retailers.



As an example of lesser-known targets, multiple MFA lookalike domains mimicked the Western Electricity Coordinating Council (WECC).

The WECC promotes Bulk Electrical System reliability for a large portion of the Western United States. The lookalikes included wecc-okta[.]org, wecc-oktc[.]org, and wecc-okta[.]com. All were registered in February 2023 and share an IP address.



Another surprising example is Feldman Auto Group, which consists of several car dealerships in the United States.

While the company has a branding relationship with American actor Mark Wahlberg, it is otherwise a moderate-sized company with 18 locations in the midwest.¹⁹ An MFA lookalike to this domain, feldmanauto-okta[.]com, was registered in late January 2023.



Some of the company targets of the MFA lookalikes are more uncertain.

The domain frb-okta[.]com shows a login prompt with a nondescript FRBOkta logo that could be the Federal Reserve Bank, First Reserve Bank, or a lookalike to a site like the Polish clothing company, Farbokta.²⁰ In many cases, we can't be sure what the target was, and the phishing kit may have been active for only a short time. *We've included a screenshot of the login in Figure 3 so you can guess for yourself.*



These AitM attacks were also used against consumers in 2022, particularly those in the gaming community who use MFA to protect their in-game purchases.

In one case known to the authors, the victim was lured to visit a site from a Twitch livestream of a popular online game. After entering their MFA credentials they experienced a brief denial of service (DoS) attack against their home network, creating an internet outage for several minutes. When they were able to return to their game account, all of their purchases had been stolen. *We might think of gamers as teenagers living in their parent's basement, but the amount of money spent on in-app purchases makes gaming, and their players, from Roblox to Counter-Strike, a lucrative target set.*

FRBOKTA.COM MFA LOOKALIKE

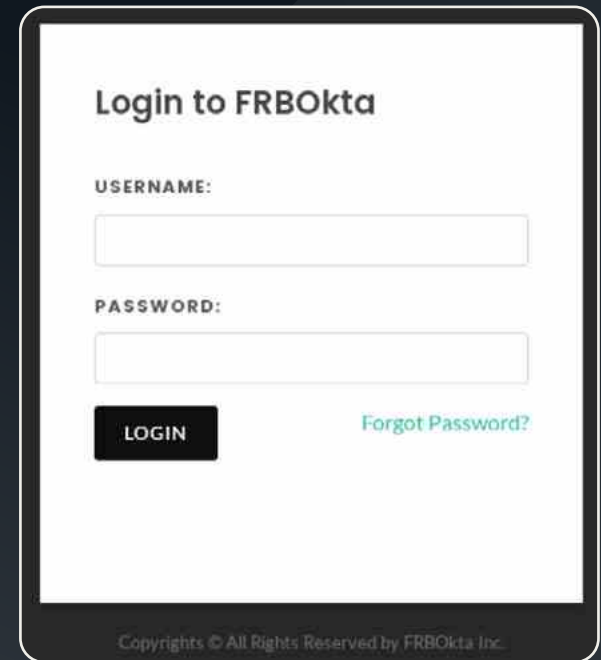


Figure 3. The website at frb-okta[.]com shows a nondescript login page with a reference to FRBOkta. Image credit: URLScan.²¹

TURKISH MINISTRY LOOKALIKE PAGE

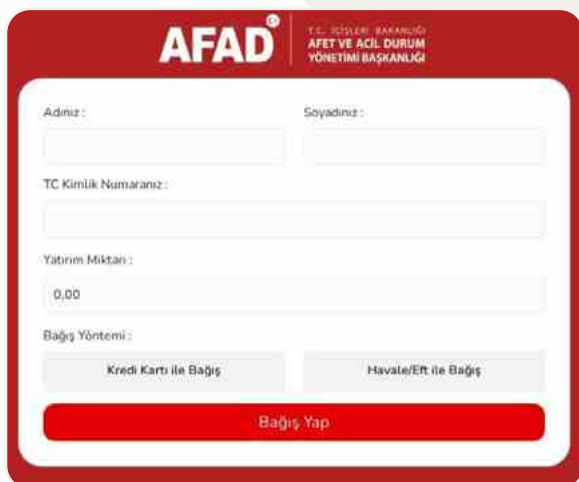


Figure 4. AFAD lookalike afadestek[.]net
Image credit: DomainTools.

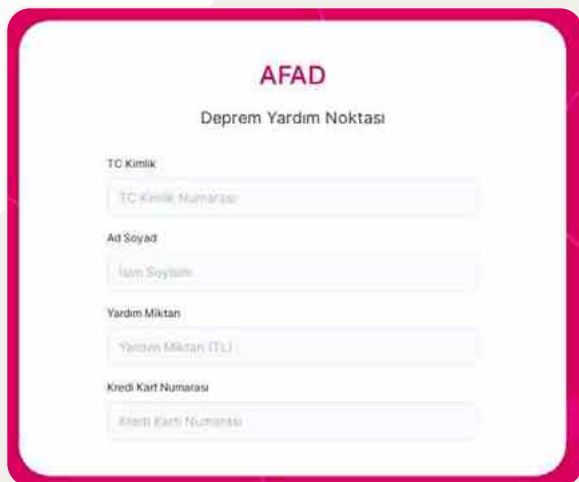


Figure 5. AFAD lookalike domain afadbagislari[.]net
Image credit: DomainTools.

THEY TARGET DO-GOODERS

Scammers looking to steal money are often “first responders” when it comes to using world events and disasters for ill-gotten gains.

Infoblox has found that scammers are quick to take advantage of any event in the news, such as health crises like COVID-19 or the Russian invasion of Ukraine. Unfortunately, 2023 brought a humanitarian crisis in the form of the Turkish-Syrian earthquake in early February.²² After the initial earthquake on February 6th, several fraudulent domains sought to imitate websites of the Turkish Ministry of the Interior’s Disaster and Emergency Management Authority (AFAD). These domains leveraged ‘AFAD’ in the fully qualified domain name, attempting to look like the legitimate domain afad[.]gov[.]tr. The below examples are domains that were newly registered, and while they have a lengthy fully qualified domain name (FQDN), they all begin with ‘AFAD.’

The use of longer FQDNs offer fraudsters more permutations of the legitimate domain for use in multiple AFAD-themed campaigns:

- afad-kizilay[.]yardim-yap[.]net
- afad-online-odeme-bagis[.]net
- afad-kizilay[.]yardimbagis[.]net
- afadtr[.]bagislama[.]net

In addition to combosquatting, some of these sites use the legitimate AFAD logo to help trick visitors into donating to the sites. For example, the fraudulent site afadestek[.]net was registered on February 7th, and displayed a web design similar to that of the legitimate Turkish AFAD site, as shown in *Figure 4*. According to machine translation, it appears to collect donations by credit card or money order via electronic funds transfer, as well as collecting PII such as first and last names and national identity numbers.

Other fraudulent domains didn’t bother to use the official AFAD logo and were quickly thrown together to maximize the amount of money they could extract from donors. Two examples are afadbagislari[.]net and afadyardim yap[.]net, both hosted at the same IP address. Dedicated infrastructure for lookalikes is common and will be discussed in further detail later. Both sites feature the same layout and content, shown in *Figure 5*, asking for donations for earthquake relief via credit card payments.

THEY TARGET CRYPTO

Apart from scammers looking to make a quick buck, lookalikes are heavily used to steal credentials.

A lookalike domain is probably what most laypeople think of when they think of a generic “phishing” website attempting to gain credentials from users. With the rise in popularity of cryptocurrencies, attackers target these financial services, including marketplaces, wallets, and exchanges. We found a number of very convincing lookalikes for the popular U.S.-based exchange Coinbase. One such site is shown in *Figure 6*.²³

The domains in the table below, for example, were registered in January 2023:

Table 1. Examples of Coinbase cryptocurrency exchange lookalike domains.

securefinancialcoinbase[.]com	reconfirmfocoinbase[.]com
secureaccountreverify-coinbase[.]com	reconfirmaccount-coinbase[.]com
secure4-coinbase[.]com	kyc-reverifycoinbase[.]com
secure2reconfirm-accountcoinbase[.]com	ap-coinbase[.]com
secure2financial-coinbase[.]com	accountupdate-financialcoibase[.]com
secure2-financialcoinbase[.]com	2farecoverycoinbase[.]com
secure-2faupdatecoinbase[.]com	recovery-financialcoinbase[.]com
2fa-accountupdatecoinbase[.]com	2fa-updatecoinbase[.]com

With the growth in non-fungible tokens (NFTs) - trades of which reached over \$2B in February 2023 - actors were quick to expand beyond traditional cryptocurrency in their efforts to steal money from investors.²⁴

As an example, the Blur marketplace opened in October 2022 and the Blur token launched a few months later, driving a record investment in NFTs since May 2022.²⁵ We began seeing Blur lookalikes soon after the product launch, and then saw a dramatic increase in lookalikes as the platform increased in popularity.

COINBASE LOOKALIKE

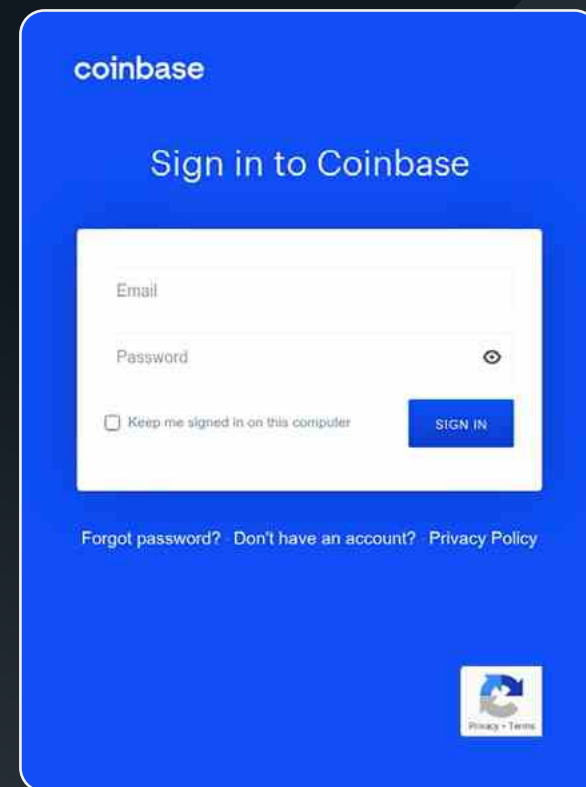


Figure 6. Coinbase lookalike click-coinbase[.]com
Image credit: DomainTools.

BLUR NFT LOOKALIKE

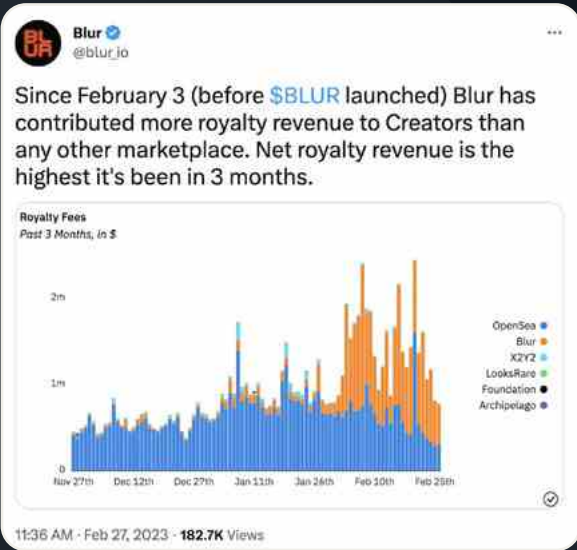


Figure 7. Blur NFT market place is among the leading drivers of the \$2B in NFT trades seen in February 2023.²⁶ Image credit: Infoblox

In the lead up to the release of the Blur Token on February 14th, 2023, we saw a five-to-six fold increase in the number of Blur-related lookalikes. Even with the amount dropping somewhat in March 2023, this pattern demonstrates actors' willingness to keep up with trends in the crypto world in order to scam a quick buck.

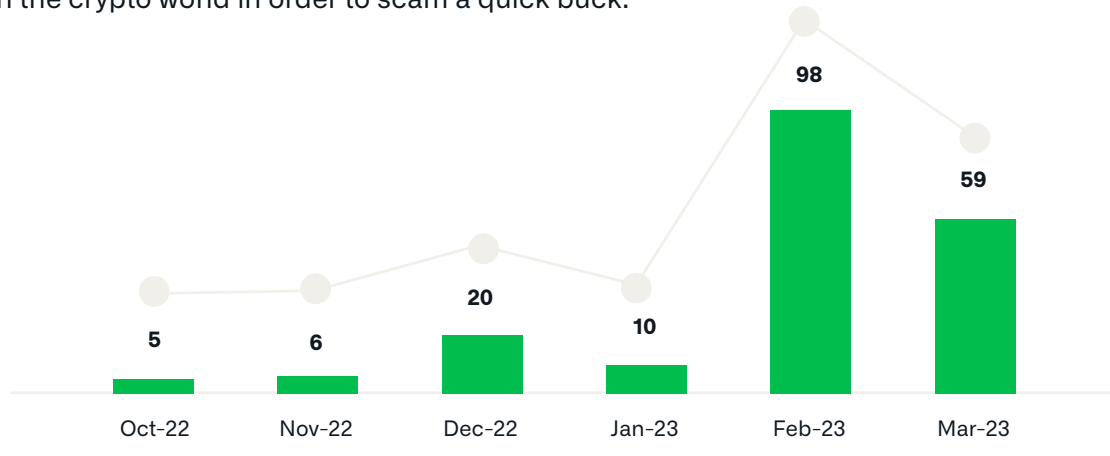


Figure 8. Drastic increase in Blur-related lookalikes since the marketplace's announcement in October 2022.

Infoblox tracks multiple actors that specialize in cryptocurrency-related lookalikes. These actors target all major entities in the market, including Blur and their competitor Yuga Labs, the owner of ApeCoin and the popular NFT Bored Ape Collection. In the table below, we provide a small sample of these domains. Techniques used by these actors include simple changes in the top level domain (TLD), the addition of a single letter, and Unicode domain names, which can be particularly challenging to recognize. Notice that in the table below, there is an accent over the "i" in apecoins[.]com. In DNS this domain looks like xn--apecons-cza[.]com, which is somewhat difficult to recognize as a lookalike, but in a web browser it would be virtually indistinguishable from the original.

Table 2. Examples of Blur token and Yuga Labs lookalike domains.

Blur lookalike domains [blur.io]	Yuga Labs lookalike domains [yuga.com]
blurclaim[.]com	yugaslabs[.]com
blurdrop[.]com	apecoins[.]com
blurnft[.]pw	apecoinstake[.]world
blur-nft[.]org	yugas[.]app
blur-coin[.]com	ape-claim[.]com

There are also less traditional cryptocurrency-related lookalikes that use YouTube as a vector to lure targets to their domains.



These schemes begin with threat actors spearphishing popular YouTube creators using fake sponsorship offers that appear to be related to legitimate products.²⁷

The emails prompt the creator to download and open a file that is allegedly related to the sponsorship offer, such as a copy of the software being promoted or a PDF file containing a sponsorship contract.²⁸ In reality, these files are malware payloads that, when opened, steal session cookies from the victim's browser. The stolen cookies allow the attacker to gain access to the victim's YouTube account, even if multi-factor authentication is enabled.



Once the attacker has access to the creator's YouTube account, they try to obfuscate the fact that the channel has been hacked by changing its name and profile photo to match the theme of their attack, which is often something related to Elon Musk or one of his companies.²⁹

The attacker may also delete or hide the channel's existing videos to further cover their tracks. The attacker then begins streaming an edited version of a cryptocurrency-related video, such as Elon Musk's Ark Invest speech, in order to lure in the channel's existing subscribers.



These edited videos include a text overlay directing users to visit the attacker's cryptocurrency-related lookalike domain, and a link to the domain is also included in the description of the stream.

The domains themselves are standard "double your money" scams that prompt victims to send a certain amount of cryptocurrency to a specific wallet address with the promise that they'll receive twice that amount in return. In these attacks, the purpose of the lookalike domain is to enhance the believability of the offer by matching its theme with the edited video and rebranded YouTube channel.

TESLA LOOKALIKE



Figure 9. Cryptocurrency-related Tesla lookalike domain tesla-online[.]net prompting users to send cryptocurrency to specific addresses in order to receive twice as much in return. Image credit: Infoblox.

THEY TARGET SOCIAL MEDIA AND MOBILE USERS

Social media platforms, like Instagram and Twitter, alongside major brands like Apple are also popular targets for phishing lookalikes.

Every popular brand and service is continually targeted in these attacks, but we will use just a few examples from these three brands as an illustration of the current threat. Credential gathering is nothing new; before social media and universal ID platforms like Apple ID appeared, bad actors were trying to get into your email account. However, with how deeply entwined social media and universal ID platforms are now with our lives, these lookalikes pose a persistent threat.

Threat actors will go after anyone's social media account, not just the accounts of influencers and celebrities. There are many lookalikes for Instagram — some combosquats, others homographs. Often such domains appeared in clusters of simultaneously registered domains, suggesting that they were a part of a coordinated campaign created using a DDGA. The examples below are all part of an Instagram set that combines the brand with words like help and feedback.

Table 3. Examples of Instagram support lookalike domains.

help-instagram-notice[.]com	help-instagram-about[.]com
feedback-instagram[.]com	help-Instagram-notice[.]com
help-Instagram-about[.]com	help-Instagram-notice[.]gq

The content on these domains claim the user has violated Instagram's copyright rules and asks the user to enter their username to appeal the verdict; see Figures 10 and 11.

INSTAGRAM LOOKALIKE



Figure 10. The Instagram lookalike help-instagram-notice[.]com displays a copyright infringement appeal call to action. Image credit: DomainTools.^{30z}

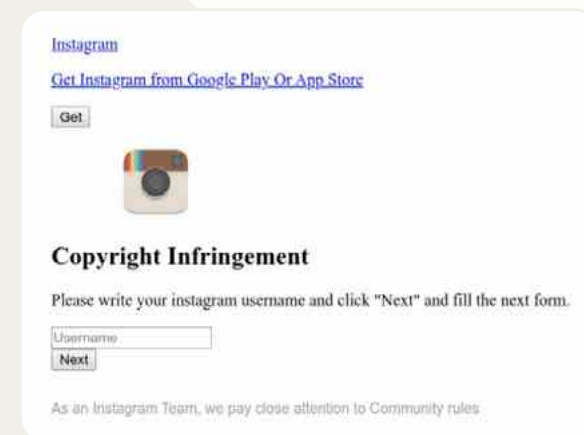


Figure 11. Instagram lookalike help-instagram-about[.]com, showing another copyright infringement appeal call to action. Image credit: URLScan.³¹

TWITTER LOOKALIKE

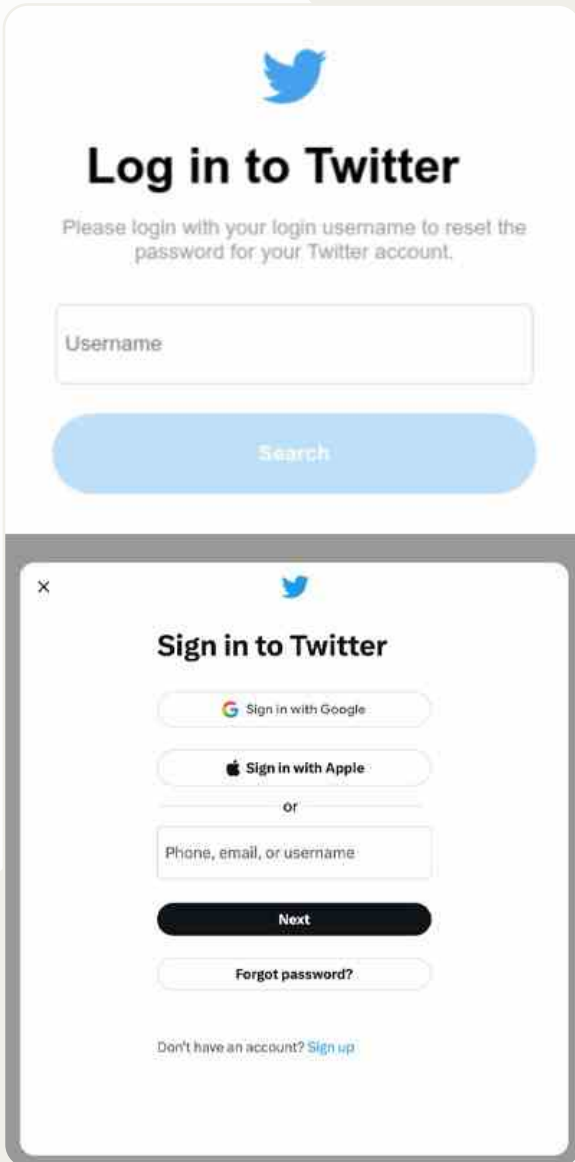


Figure 12. Convincing password reset portal on Twitter lookalike help-twitter-centre[.]net. The phishing image is on top, the legitimate one is on the bottom. Image credit: DomainTools.³²

Other Instagram lookalikes target the coveted “blue checkmark” (Instagram’s approach to verification as a public figure), by using a lowercase “l” in place of an uppercase “i.” Ironically, Instagram introduced the blue checkmark for well-known personalities or companies as a way to combat impersonation. *Don’t put it past bad actors to use lookalikes targeting anti-lookalike solutions.*

Some examples are:

Table 4. Examples of Instagram verification lookalike domains.

Instagram-blueticket-form[.]ml	Instagram-contactbluebadge[.]ga
Instagram-verification-badges-service[.]com	Instagrambluetickverification[.]cf
Instagramverifybadge-contact[.]cf	Instagram-badgecentre[.]gq

In tracking Instagram lookalikes, we found that actors didn’t put all their eggs into one social media basket.

Lookalikes for Twitter were also hosted alongside the Instagram “copyright infringement” lookalikes. These Twitter lookalikes were combosquat domains phishing for users’ credentials, and the landing pages appear to be a legitimate password reset portal; see *Figure 12*.

In addition to social media lookalikes, during our research we often saw lookalikes for iCloud, Apple’s cloud service that offers cloud storage and synchronization across Apple devices. These domains leveraged a relatively small number of keywords; we most frequently observed “apple,” “findmy,” “id,” and “icloud.” There were no shortage of Apple-related lookalike domains.

Below are a few examples, including some that appear to target Spanish-speaking users:

Table 5. Lookalike domains targeting Apple-related services.

supportid-apple[.]com	sopport-apple[.]com
soporte-latam[.]us	soporte-appleid[.]com
icloud-web-app[.]com	icloud-fndmy[.]com

THEY TARGET EVERYONE



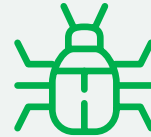
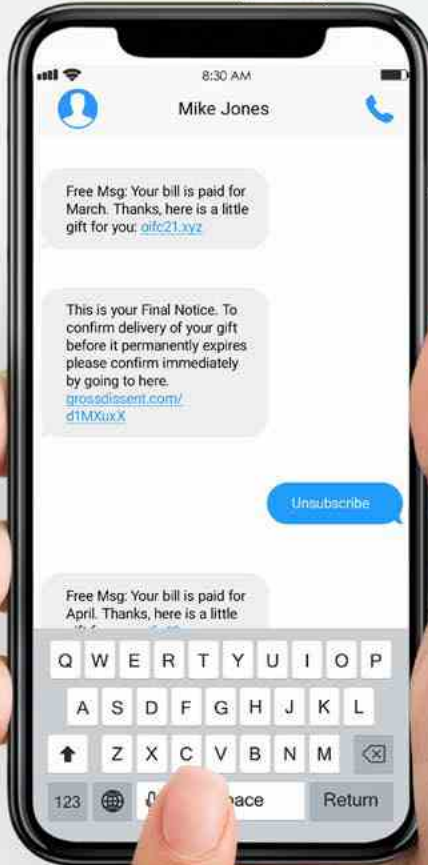
Our detection algorithms identify thousands of new lookalike domains every day.

Any company or service, big or small, where malicious actors can steal money or identities will be targeted. We'll close out this section with an assortment of lookalike domains we've observed in the wild and their target.

Table 6. Lookalike domains and their targets.

Lookalike domains	Lookalike target
mee6bot[.]ru	Discord bot, Mee6
vulcan[.]pm	Discord bot, Vulcan
o365-outlook[.]com, ms-o365[.]com, o365-outlook[.]com, https-o365[.]com	Microsoft Office 365
myato-refund[.]online	Australian Tax Office
checkscam22[.]com, checkscams[.]online, checkscammer[.]xyz	Scam checking websites
xpressvpn[.]business, expressvpn-app[.]com, expressvpn-okta[.]com	Express VPN
anpost-paymentduty[.]com, ups-pay-deliveryfee[.]info, caddeliverypostca[.]com	Postal and delivery services
crarebate-info[.]com	Canadian tax rebate
ebl-ch[.]com	Swiss energy company EBL
op-fi-palvelut[.]co, op-fi-io[.]in	Op[.]fi, Finnish digital banking and insurance service
boatairbuds[.]in, boatbudsmusc[.]in, boatflashsale[.]in, boatmusicairbud[.]in	Indian technology company BoAt
pumauaeshoes[.]com, pumanzsale[.]com, pumaireland[.]com, vejaoutletcanada[.]ca	Shoe companies
secure1-scotiabank[.]com, r-scotiabank[.]com, chasebank-jpm[.]com, thetrustnationalbank[.]com, americafirst[.]com	Banks
sprint-ldg[.]com, tds-telecom[.]com, teistra[.]ne, 1111systems-okta[.]com, t-mobile-okta[.]us, vzw-ss0[.]com	Internet and cloud service providers
ss0-authentication[.]de, ss0-securelogin[.]com, service-sys-2fa[.]com	Multi-factor authentication and single sign on domains





HOW ARE LOOKALIKES USED?

Now that we've covered what lookalikes are and some example targets, let's talk about how they're used.

By "how," we mean their deployment methods. Infoblox saw lookalikes deployed in a variety of manners, such as:

- **SMS messages**
- **Phone calls**
- **Direct messages on social media sites**
- **Emails**
- **Embedded in QR codes**
- **Domains on the World Wide Web**

THEY SEND TEXTS



In spite of improvements in spam filters for mobile phone text messages (SMS), the use of SMS to deliver phishing messages, often called smishing, continues to rise.

Actors are able to quickly distribute a large number of messages and avoid some of the security mechanisms that are put into place to protect against email phishing attacks. SMS is used both in broad consumer attacks and in narrow spearphishing attacks against organization employees. In this section we'll describe two threat actors who've used SMS and lookalike domains to attack consumers and government employees.

For nearly a year, Infoblox has been tracking a persistent lookalike smishing actor we call OpenTangle. To our knowledge, this actor has not been reported elsewhere. OpenTangle initially targeted Western consumers by using lookalikes to financial institutions, internet providers, and online retailers. The actor recently began targeting government employees and contractors. We are aware of over 1500 lookalike domains controlled by OpenTangle since they began operating about two years ago. Some of OpenTangle's domains include mtbsupportz0610[.]com, americafirstOnline[.]com, and mygov03-ato[.]com.



Notice their use of different lookalike techniques.

One of this paper's authors has received multiple texts from OpenTangle, including lookalikes to M&T Bank, with which the author has no affiliation. Early in their campaigns, OpenTangle included shortened URL links in their smishing texts, perhaps hoping the obfuscation would be successful. However, by May 2022, they converted to lookalike domains. *Figure 13* shows an example of one of their banking campaigns in which they request the user's credentials.

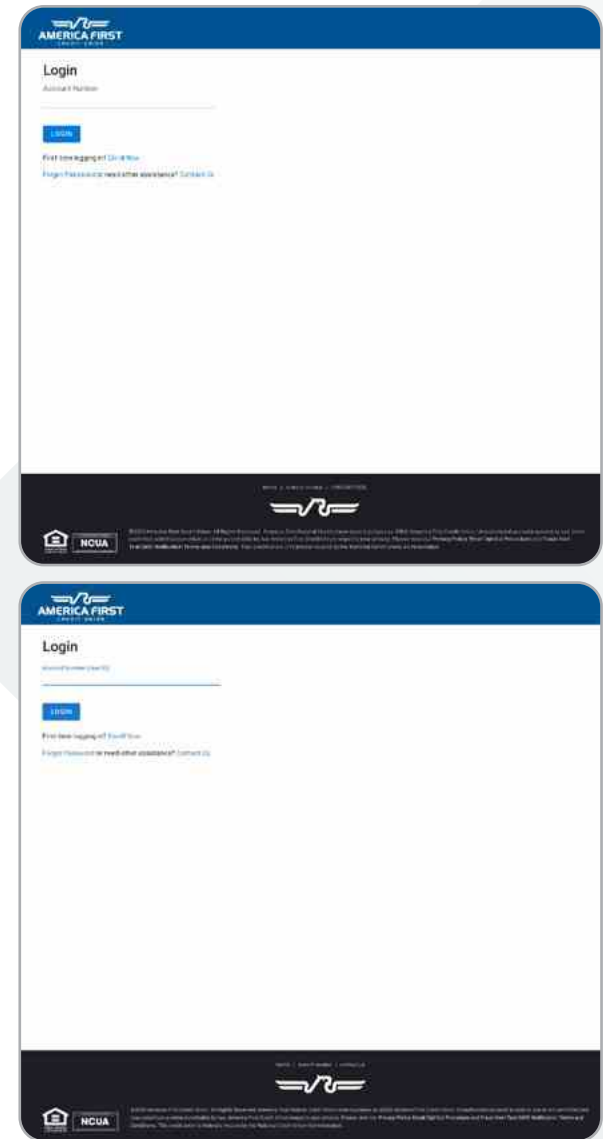


Figure 13. A phishing page at the domain americafirstOnline[.]com targeting America First Credit Union account holders. The image on the top is the phishing page, the image on the bottom is the legitimate page. Image credit: URLScan.³³



OpenTangle began exploiting MFA using AitM phishing kits within the last year.

While their earlier campaigns used standard phishing login pages and generally targeted consumers, *Figure 14* shows an example of how they've advanced their campaigns. In this case they are targeting Australian Government myGov account holders and requesting an MFA code, rather than a simple login. They also included a link to call the helpdesk, another technique that emerged in 2022 as a means to convince users to visit malicious websites.

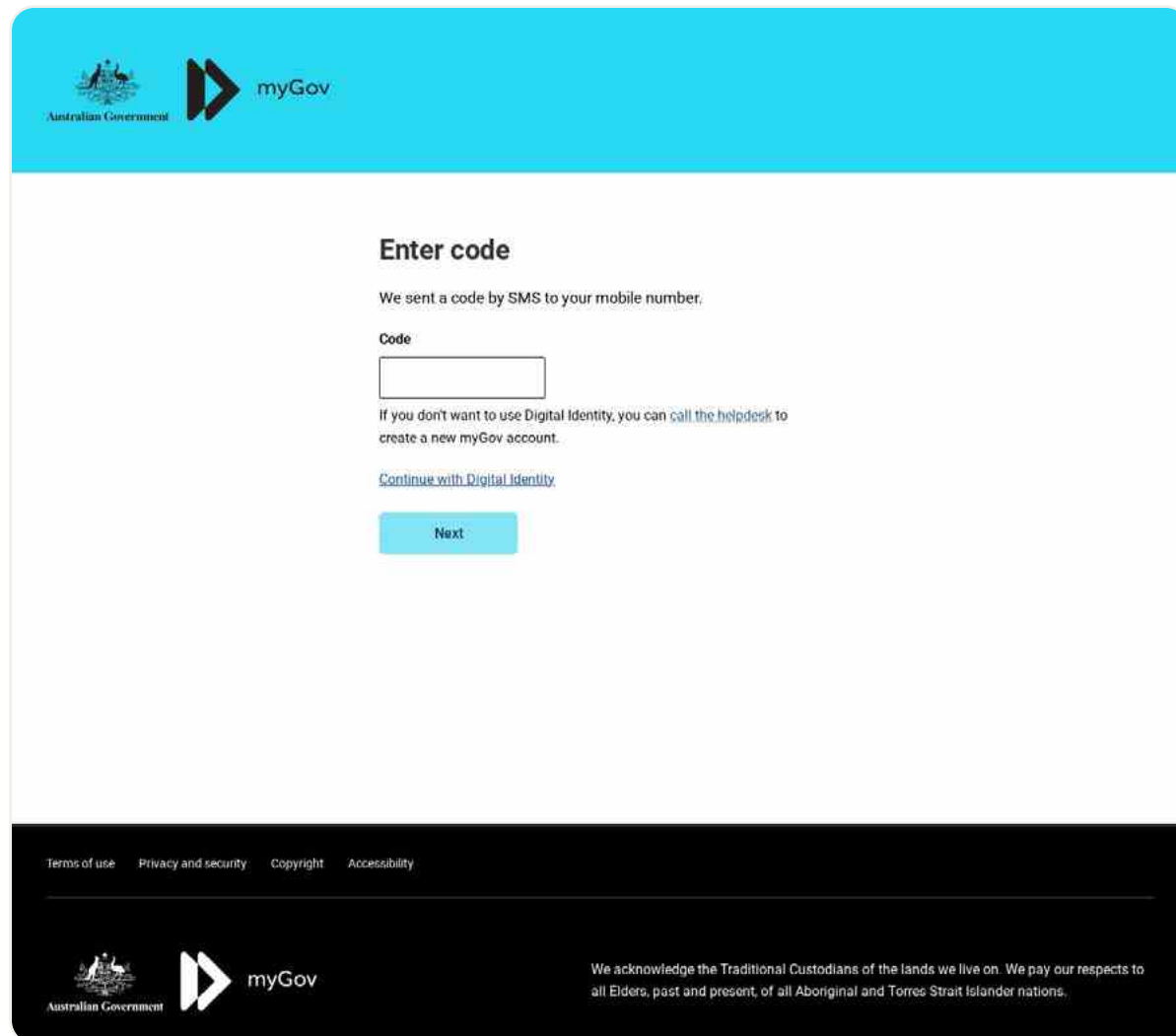


Figure 14. OpenTangle lookalike domain [www.mygovsupport-ato\[.\]com](http://www.mygovsupport-ato[.]com), imitating myGov, the Australian Government's online portal for the government cloud. Image credit: URLScan.³⁴

Scamélie is another example of an actor using smishing messages to spread lookalikes.

The actor we call Scamélie is a collection of loosely affiliated groups and individuals involved in a long list of scams coming from and primarily targeting French-speaking countries. We've also seen them engaged in more general targeting across Europe and the UAE. Scamélie's lookalike domains primarily impersonate ISPs, banks, government services, and delivery companies. Due to the loose affiliation of the group, we've also seen scams for less expected companies, such as travel companies, sports apparel companies, and grocery stores.

Scamélie's lookalike domains are often hosted on large cloud providers or "bulletproof" hosting companies. In some cases the scammers have set up their own or use hosting providers set up by other, unaffiliated scammers. We saw both targeted domains as well as general-purpose domains (my-account, resolve-an-issue, etc.) registered through stolen identities and paid for with virtual credit cards or cryptocurrencies.



Once the actors have collected credit card information, they call the victim, posing as an employee of the victim's bank or credit card issuer.

They explain that the victim's credit card information has been stolen but that they will help remedy the issue. The caller then says that the victim will receive two MFA codes that will have to be read back to the caller for account security. In reality, the attacker needs the MFA codes to steal money from the victim in real time. The first MFA code increases the wire transfer amount and the second enables the transaction to go through. To increase the effectiveness of their calls, the actor employs callers who are ideally young women and/or individuals who speak French in a manner that won't raise the suspicion of a native speaker.

As an unorganized group, Scamélie is difficult to track and analyze. They often smish during their victims' nighttime and take their domains down after just a couple of hours or days. They use anti-bot and anti-scraping scripts to further obstruct security researchers.

SCAMÉLIE LOOKALIKE EXAMPLES



Figure 15. A Scamélie lookalike amende.gov.fr, imitating a French government service portal. Image credit: Infoblox.

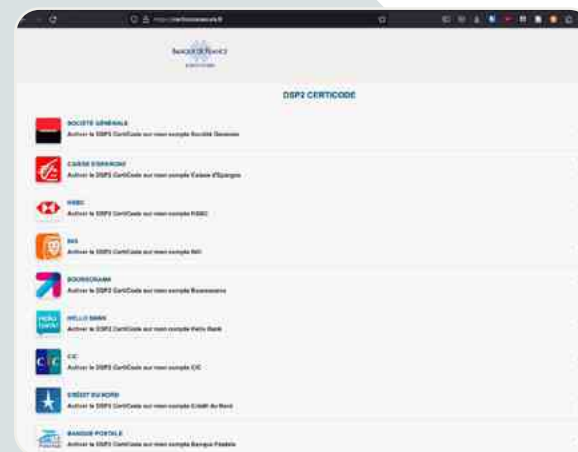


Figure 16. A Scamélie lookalike site certicodesecure.fr, spoofing a French banking service and enticing victims to link their bank account information. Image credit: Infoblox.



THEY USE OLD-FASHIONED PHONE CALLS



The Cybersecurity and Infrastructure Security Agency (CISA) released a Cybersecurity Advisory (CSA) on January 26th, 2023 about the malicious use of remote monitoring and management software (RMM).³⁵

CISA identified a campaign in October 2022, in which bad actors sent phishing emails containing a phone number and prompted users to call. The email was designed to pass as a customer support message, and when users called the phone number, actors prompted them to visit a malicious domain. When the user did so, an executable file was downloaded and then contacted a second malicious domain, from which additional RMM software was downloaded. This software — AnyDesk and ScreenConnect — was legitimate, but preconfigured to connect to the actor's RMM server for persistence.



The domains used are lookalikes of well-known services; the likelihood of accepting the domain is even higher for those victims who were given it over the phone due to the additional social engineering used to craft the scripts and callers' personas. We performed a retroactive review of our data and found evidence that the actor has been active for longer than the CSA indicates.³⁶ These campaigns were active since at least Spring 2021, over a year before the incidents that CISA and Silent Push, in a separate article, described. We also saw some domain re-use. For example, the domain amzsupport[.]live, an Amazon lookalike, was part of an active campaign in April 2020 and then used again in October 2021.

As attacks against MFA protection of internal corporate systems came to light in early 2023, it was revealed that in some cases the actors phoned the victim, pretending to be their IT department. This was done after the victim had not responded to the initial prompt and was used to provide further legitimacy to the need for the user to visit the lookalike domain. Users who complied enabled the actor to steal their corporate credentials.

THEY SEND SPAM

While we've seen crafty actors use smishing and phone calls to distribute lookalikes and ensnare victims, the phishing email has never gone out of style.

Infoblox analyzes tens of thousands of malspam emails every day, revealing a seemingly unending stream of campaigns distributing lookalike domains. We will highlight a few of these campaigns but stress the importance of organizations maintaining diligent monitoring for phishing emails.

One such campaign targets Xfinity, a major American telecommunications company. These lookalikes have DGA-like characteristics and are of the form xfnity<short or partial word>.com. Note that "Xfinity" is misspelled because it is missing the first "i." The actor also ensured the sender name appeared legitimate, showing as "Xfinity Mobile," which uses a Cyrillic capital letter "X." The sender emails used their own domains and appeared to also have DGA-like characteristics in the username as well, consisting of the pattern noreply-<keyword>, such as noreply-corporate@xfnitycard[.]com. The actors did not use unique domains for each email. In some cases, the domains were repeated, but the keyword was changed, as in: noreply-corporate@xfnitycard[.]com and noreply-active@xfnitycard[.]com.

Table 7. Xfinity lookalike domains.

xfnitykuri[.]com	xfnitycomp[.]com
xfnitystarter[.]com	xfnityhlaty[.]com
xfnityersa[.]com	xfnityothie[.]com
xfnitykaris[.]com	xfnityrkles[.]com
xfnityrayton[.]com	xfnitycard[.]com

The domains identified in the campaign utilize a technique we have coined decoy parking: when a domain is visited directly and it appears to be parked, but in reality, the domain's mail server is active and sending malicious emails. We have found decoy parking to be fairly common and not reported on by other vendors. See *Figure 17* for an example of a decoy parking page.

XFINITY LOOKALIKE



Figure 17. Decoy parking page exhibited by Xfinity lookalike xfnityrayton[.]com. Image credit: URLScan.³⁷

WEDO MACHINERY LOOKALIKE

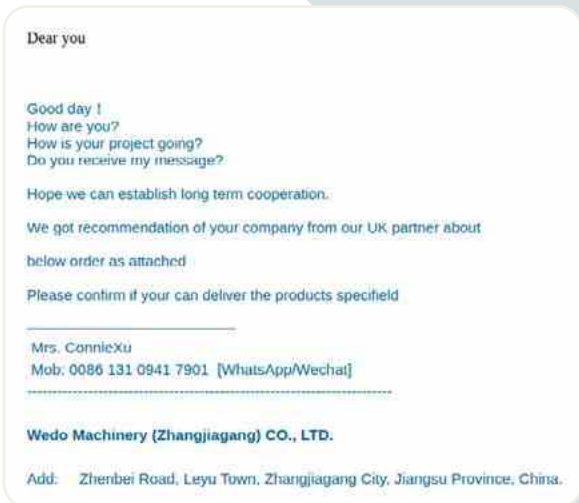


Figure 18. Body of malspam campaign using Wedo Machinery as a lure and acrobat-adobe[.]com lookalike domain as a malware C2. Image credit: Infoblox

Our analysis found these Xfinity lookalikes in distributed malicious Word documents.

Campaign subjects doubled as a call to action and centered around payment being refused or a threat of service termination, such as “[Announcement] Your service is at risk of being terminated” or “[Need Action] We can’t charge your card, fix this error.” The body of these emails was framed as coming from customer support, asking recipients to “see attachment for case details.”

Another campaign Infoblox identified used a Chinese recycling company, Wedo Machinery, to drop a ransomware loader. We identified 176 emails within this campaign, each with a .zip file containing a single executable identified as Zmutzy. See *Figure 18* as an example of an email within the campaign. We saw two file names within the campaign: PO-0097(1).zip and PO-29862K.zip. The Zmutzy loader uses the lookalike domain acrobat-adobe[.]com to download additional payloads.



THEY USE QR CODES

In addition to direct cryptocurrency lookalikes, we observed the use of QR phishing — where a QR code is used to obfuscate a URL destination and deliver malicious content — in conjunction with lookalike domains created to entice users to claim free prizes and provide crypto wallet account information.

In one example, the QR code redirected the victim to a `bridge[.]walletconnect[.]com` link, a mechanism used to steal funds. In this scam, the actors set up a Twitter account, `@adidas_weare`, to build credibility and share their lookalike domains; see *Figure 19*. The account amassed 16,000 followers as of February 21st, 2023; fortunately, the account has now been deleted or taken down.

The actors advertised fake giveaways of different items including Porsche cars and Adidas clothing or shoes. The domains are predominantly combosquats containing the keywords “adidas” or “porsche.” Upon visiting the lookalike domains, such as is shown below in *Figure 20*, users were asked to scan a QR code that would allow them to claim the item being given away, then redirect them to the decentralized application, WalletConnect, which provides the actor access to the user’s funds.

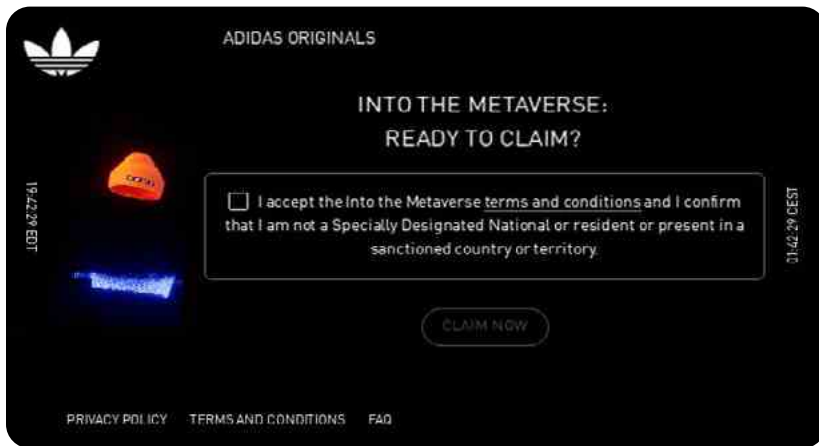


Figure 20. Adidas lookalike domain `adidas-go[.]com` enticing users to click to claim a free item. Image credit: URLScan.³⁹

If users scan the QR code and link their cryptocurrency wallets to the decentralized application, the actors are able to extort cryptocurrencies from the user. These domains use shared nameservers and are hosted on a Russian-resolving IP address, `185[.]149[.]120[.]83`, which is completely actor-controlled and contains other lookalikes to Blur as well as Arbitrum, a solution to improve the speed and scalability of Ethereum smart contracts.

ADIDAS LOOKALIKE



Figure 19. The lookalike Twitter account `@adidas_weare` of Adidas Originals `@adidasoriginals`. Image credit: Infoblox.

THEY USE DNS

Lookalikes don't only occur as website domains.

We have found them to be used in several DNS capacities, including:

- **Nameserver**
- **Mail server**
- **CNAME records**
- **PTR records**

In most cases, these domains will not have a typical A record or website presence and may often appear parked — an implementation of decoy parking that we described in an earlier section. Attackers also use lookalike domains for redirection and C2 communication in DNS.

NAMESERVERS

As an example of lookalike nameservers, the domains `bitkeep[.]dev` and `flutter[.]direct` were registered in November 2022. Both of these are lookalikes to different domains, but they share an infrastructure. BitKeep is a decentralized multi chain crypto wallet that aims to be a single hub for all cryptocurrency transactions. The official domain for BitKeep is `bitkeep[.]com` and the company has been in operation for five years with over 8 million users.⁴⁰ Flutter is Google's portable user interface (UI) toolkit for crafting natively-compiled applications for mobile, web, and desktop from a single codebase. The official domain for Flutter is `flutter[.]dev`.⁴¹

Both of the legitimate domains host web content at the primary domain, but neither of the lookalike domains do. When initially registered, both of the domains were acting as a nameserver for one other domain, `get-flutter[.]com`, which is another lookalike of Flutter. At that time, the domains were hosted on Swiss offshore hosting provider Private Layer. This network also hosted `flutter[.]vision`. While we can't definitively attribute these domains to malicious activity, they demonstrate a pattern of leveraging lookalike domains for nontraditional purposes. They prove to be quite challenging to analyze even for experienced researchers and are unlikely to trigger many threat intelligence algorithms.

MAIL SERVERS

In addition to nameservers, we've seen lookalikes being used as mail servers. The domains `whirlpoolmxonline[.]com` and `whirlpoolservicesmx[.]com` target the major appliance brand Whirlpool and share common infrastructure. They're hosted on the same IP address, owned by Lyra Hosting, a low-quality VPS and hosting provider located in the Seychelles, and share common nameservers.

While they target Whirlpool directly with the second level domain (SLD) name, we've also identified characteristics within each domain that show they are targeting other major appliance brands as well. The SLD `whirlpoolmxonline[.]com` has three subdomains: `mabe-onlinemx[.]whirlpoolmxonline[.]com`, `samsung-onlinemx[.]whirlpoolmxonline[.]com`, and `lg-onlinemx[.]whirlpoolmxonline[.]com`. Mabe is a Mexican appliance company. The SLD `whirlpoolservicesmx[.]com` has no subdomains, but the historical chain of SSL certificates associated with the domain point to the targeting of similar appliance brands as `whirlpoolmxonline[.]com`: `www[.]lgservicesmx[.]mabeservice[.]com` and `*.lgservicesmx[.]com`.

Using lookalikes as mail servers offers an additional challenge for detecting phishing emails on an endpoint due to the appearance of legitimacy upon a first glance at email headers.

MALWARE C2s

In the email deployment section earlier, we mentioned how a malspam campaign we identified that was dropping the Zmutzy ransomware loader, used the lookalike domain `acrobat-adobe[.]com` as a malware C2 server. Lookalikes are perfect for malware C2s because they can easily blend into network traffic alongside legitimate domains. Researchers at ESET, a Slovak security software company, identified malware C2s for FatalRAT (remote access trojan) posing as Telegram, the messaging application, in February 2023.⁴²

Table 8. Telegram lookalikes functioning as malware C2s.

<code>12-03.telegramxe[.]com</code>	<code>12-25.telegraem[.]org</code>
<code>12-25.telegraxm[.]org</code>	<code>12-25.telegraem[.]org</code>

The domains hosting the malicious .exe files were also lookalikes to Telegram, as well as WhatsApp, Skype, Google Chrome, and Firefox.





REDIRECTS

Lookalikes can also be employed as redirects. We've identified a large network of typosquat domains that redirect visitors to choto[.]xyz, a C2 domain that conditionally redirects victims to the landing domain lotto60[.]com. The actor uses reverse proxy services and Cloudflare bot protection on choto[.]xyz, presumably to prevent detection and exploration by security researchers. The landing domain appears to be running a fraudulent affiliate marketing program. By analyzing the document object model (DOM), we can see that the HTML contains an inline gtag() function that sends visitor data to Google Analytics with the analytics ID G-DT4YWT5VP8. In addition to inflating the actor's affiliate marketing numbers, we've seen lotto60[.]com being requested over HTTP by potentially malicious files that match file signatures confirmed to be the remote access trojan Nighthawk.⁴³

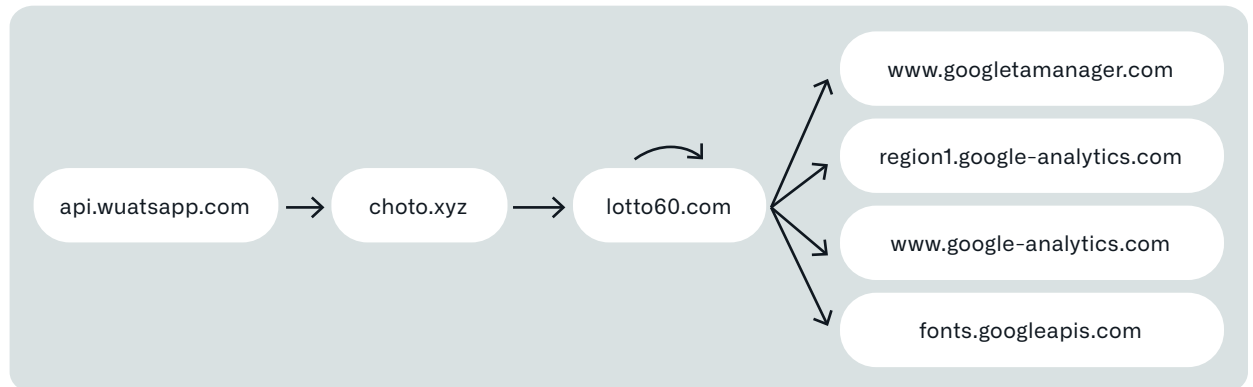


Figure 21. Example redirect chain from a typosquat domain to Google Analytics. Image credit: URLQuery.⁴⁴

The first-stage typosquat domains imitate a variety of companies. Some examples include →

These typosquats are typically parked for one to three months before being used as redirects. The actor has shown great care in crafting these typosquat domains. Each incorrect character is directly adjacent to the correct character on a U.S. English, QWERTY keyboard. These are mistakes that any average typist could make multiple times in a single day – save for those users who still “hunt and peck.”

Table 9. Lookalikes functioning as redirects in a fraudulent affiliate marketing campaign.

gi6hub[.]com	whatysapp[.]com
bankofamegica[.]com	babgkokbank[.]com
intuhit[.]com	scotiasbank[.]com

WHY ARE THEY EFFECTIVE?



Dear Reader, did you notice the 19 lookalike words we sprinkled through this paper so far? Some of them are very challenging to see!

Hint: There are 6 more. See if you can find them.

So far we've covered some specific targets as well as the deployment methods' infrastructure of lookalike domains. But why are they so effective? What makes them such a persistent threat?

The answer is complicated and involves aspects of psychology, technical implementations, and simple human mistakes — **that's what makes us human, after all!**





PSYCHOLINGUISTICS

Psychologically, the human brain short-circuits (in this case, we mean the literal definition of a current taking an unintended path of least resistance) while reading. You have probably seen a meme reading something like:

Aoccdrnig to rscheearch at Cmabrigde Uinervtisy, it deosn't mtttaer in waht oredr the ltteers in a wrod are, the olny iprmoetnt tihng is taht the frist and lsat ltteer be at the rghit pclae. The rset can be a toatl mses and you can sitll raed it wouthit porbelm. Tihs is bcuseae the huamn mnid deos not raed ervey lteter by istlef, but the wrod as a wlohe.

While the claim is unfounded in the sense that no such research at Cambridge was ever published, the underlying concept seems to have merit. For example, recent actual research suggests that “viewing a jumbled word activates a visual representation that is compared to known words.”⁴⁵ While proving or disproving fundamental questions of psycholinguistics is beyond the scope of this paper, we do want to show how psycholinguistics plays an important part in the effectiveness of lookalikes.

Specifically, the human brain’s short-circuiting plays a role when it comes to homographs and typosquats. When you see a domain like Infoblox[.]com, your brain doesn’t necessarily parse each individual letter in that domain name, and so you may never notice that the first character is actually a lowercase “l” and not an uppercase “i.”

For similar reasons, when you see the domain google[.]com, your brain may not stop to recognize that there are three of the letter “o” rather than the proper two... at least, not until it’s too late and you’ve already clicked on it.

PUNYCODE SUPPORT: HITS AND MISSES

Web browsers have ways to defend users against internationalized domain name (IDN) homograph attacks. The first and most prominent line of defense is to “translate” the Unicode domain into Punycode, which can be recognized by its leading “xn--” and appears to be gibberish to the naked eye. This is because Punycode maps Unicode characters to the far more limited subset of American Standard Code for Information Interchange (ASCII) characters containing only letters, digits, and hyphens. Each major browser has support for Punycode domains. Google gives a detailed description of the heuristics involved in the algorithm determining whether to show the internationalized or the Punycode version of a domain in Chromium.⁴⁶ Mozilla gives a similar description.⁴⁷

Mozilla also offers this inspiring bit of text in the description of their IDN display algorithm:

Our response to this issue is that in the end, it is up to registries to make sure that their customers cannot rip each other off. Browsers can put some technical restrictions in place, but we are not in a position to do this job for them while still maintaining a level playing field for non-Latin scripts on the web. The registries are the only people in a position to implement the proper checking here. For our part, we want to make sure we don't treat non-Latin scripts as second-class citizens.

In 2017, security researcher Xudong Zheng registered a domain already in Punycode, xn--80ak6aa92e[.]com, which translates to “apple[.]com,” containing Cyrillic characters that mimic the appearance of the Latin characters in “apple.”⁴⁸ At the time, Internet Explorer, Microsoft Edge, Safari, Brave, and Vivaldi web browsers were not vulnerable, but Chrome, Firefox, and Opera were. At this time, only Firefox continues to translate the Punycode, leaving users vulnerable to the attack (we did not recently test the domain on Internet Explorer or Microsoft Edge).

WHAT IS PUNYCODE?

Punycode is a special encoding used to convert Unicode characters to ASCII, which is a smaller, restricted character set. Punycode is used to encode internationalized domain names (IDNs).



iMESSAGE SMISHING USING IDN HOMOGRAPHS



Figure 22. Top Image sourced from Tyler Butler showing a real New York Times article sent via iMessage. Bottom Image sourced via Tyler Butler showing a spoofed NYT article on an IDN homograph domain. Image credit: Tyler Butler.

Hu et al. performed a longitudinal and quantitative analysis on the effectiveness of browser-based defenses against IDN homograph attacks.⁴⁹

They set out to answer three questions:

1. What policies do major browsers implement and how well do they enforce those policies?
2. Are there ways to systematically bypass the existing policies?
3. How well can web surfers recognize IDN homographs, and are those IDN homographs that bypass the browser policies more or less deceptive?

To answer the questions, the authors looked at five mainstream browsers (Chrome, Firefox, Safari, Microsoft Edge, and Internet Explorer) over five years (from January 2015 to April 2020). They generated 9,000 testing cases to answer their first two questions and ran a user study to answer the third. Chrome and Edge were most successful at displaying Punycode instead of their corresponding IDN homographs; both browsers had an overall failure rate (showed the IDN version instead of Punycode) of 20.62%. Safari and Firefox were much worse, with an overall failure rate of 42.91% and 44.46%, respectively. Each browser had differing failure rates depending on the category of IDN. Furthermore, the authors found that web surfers struggle to identify homograph IDNs, and those IDNs that browsers blocked were the most troublesome in determining authenticity: 48.8% of users thought they were, 48.5% of users thought they weren't, and 2.7% couldn't tell.

So far we've only focused on desktop browsers. But as we've seen with the lookalike smishing attacks described earlier in this paper, IDN homograph domains are also quite at home on mobile devices. In fact, they could be more pernicious. Smaller screen sizes, smaller address bars, and a general lack of link previewing can lead to more effective lookalike attacks. Even when there is link previewing, IDN homographs can still be effective on mobile devices. In 2021, security researcher Tyler Butler published on the plausibility of smishing using IDN homographs in iMessage.⁵⁰ iMessage offers rich previews of links, but a savvy attacker can get around this quite easily with a good-enough lookalike domain and a bit of styling work for the web page itself. As Mr. Butler notes, this form of attack can be used to spread misinformation, steal credentials, or deliver targeted malware or spyware.

Mr. Butler describes that Apple claimed they will not address the vulnerability due to the fact that the homographs are “visually distinguishable.” Given Figure 22, what do you think? Can you spot the difference?

TO ERR IS HUMAN, TO FORGIVE IS DIVINE... BUT TO AUTOMATE IS WISE

On the World Wide Web, some other humans aren't so forgiving of others' mistakes.

As we've mentioned, actors use typosquat domains to prey on the natural spelling mistakes of others. All an attacker has to do for a typosquat to be effective is register a plausible domain and wait. That's it. Sooner or later, a human will make that spelling mistake and land on a domain that they never intended to visit. Of course, bad actors don't just wait, they proactively entice people to click. And in our fast-moving world, many times we don't even realize that we made a mistake in the first place.

At the end of the day, lookalikes are called lookalikes for a reason: they look similar to known domains with the intent to deceive a human. As we've seen, some lookalikes are more effective than others, but the choice of domain name is only one part of a lookalike's effectiveness. The way a lookalike domain is deployed can also have a significant impact on the overall success of the campaign. Take, for example, an Okta or MFA lookalike like `okta[.]Infoblox[.]com`, or `okta-lfoblox[.]com`. A discerning person who triple-checks each domain name before visiting (good luck finding one of those folks) might be able to notice that the "i" in the second level domain (SLD) is actually a lowercase "L." But that lookalike, paired with a well-crafted SMS message to the phone number they have in their employer's online profile, for example, could be the difference-maker. Add to the equation a phone call with an urgent call to action, and it's game over. Of course, this is a fictional example (with all component parts being used) of spearphishing, and not a general campaign employing lookalikes, but the point remains: lookalike techniques can be effectively applied to domains in multiple ways and to multiple parts of DNS infrastructure.

All this to say that the oft-quoted proverb "fool me once, shame on you; fool me twice, shame on me" doesn't apply to lookalikes. Even the most hawk-eyed, security-aware individuals can fall prey to a lookalike — and do it again, and again. Bad actors have the upper hand in this war, but it's not lost yet. Infoblox has solutions at the DNS level to ensure that organizations have the capability to fight back and effectively defend themselves.

IOCs



The complete list for this paper can be found on GitHub at <https://github.com/infobloxopen/threat-intelligence>



INFOBLOX SOLUTIONS

Lookalike domains remain popular with attackers due to their effectiveness and the difficulty in detecting them at scale. The challenge is compounded by the difficulty in automatically identifying a suspect domain that is intended to mimic a legitimate target. This has resulted in businesses and government agencies becoming increasingly concerned about lookalike domains that impersonate their corporate domains or supply chain.

Infoblox BloxOne Threat Defense (B1TD) Advanced offers a uniquely broad and comprehensive solution against lookalike threats. Leveraging large-scale DNS, Infoblox is able to apply a series of analytics to hundreds of thousands of new SLDs every day. This includes multiple analytics for lookalike detection such as an automatic assessment of visual similarities in IDN Homographs.

Customers can select from commonly targeted domains or create a custom list for specialized lookalike monitoring and analysis. The results of this in-depth analysis can be accessed through the lookalike Reporting UI, which also flags instances where the detected lookalike is associated with suspicious or phishing activity. Overall, policies can be customized to suit the needs of a customer's specific environment and level of risk tolerance. And detailed domain data includes valuable annotations that are accessible through B1TD Advanced UIs and APIs, providing customers with context that can speed threat investigations and make incident responses more effective.

These lookalike threat detection capabilities are just one of many services offered by BloxOne Threat Defense that enables it to see threats that other solutions do not and stop attacks earlier in the threat lifecycle. Through pervasive automation and ecosystem integration, it can drive greater efficiencies in SecOps, uplift the effectiveness of the existing security stack, secure digital and work-from-anywhere efforts and lower the total cost for cybersecurity.

FOR MORE INFORMATION



Visit infoblox.com



Follow-us on LinkedIn



Follow-us on Twitter

REFERENCES

- ¹ https://docs.apwg.org/reports/apwg_trends_report_q3_2022.pdf
- ² <https://twitter.com/kgrouppcompanies/status/1188878363068391425>
- ³ https://en.wikipedia.org/wiki/IDN_homograph_attack
- ⁴ <https://i.imgur.com/68oL4U9.jpg>
- ⁵ https://www.researchgate.net/publication/220420915_The_Homograph_Attack
- ⁶ <https://util.unicode.org/UnicodeJsps/confusables.jsp>
- ⁷ <https://www.igoldrush.com/domain-guide/domain-issues/cybersquatting-and-typosquatting>
- ⁸ <https://dl.acm.org/doi/pdf/10.1145/3133956.3134002>
- ⁹ <https://core.ac.uk/download/pdf/34615371.pdf>
- ¹⁰ [https://iris.polito.it/retrieve/handle/11583/2970511/1dd2efbb-9eed-4db5-be4f-ecfcc362572c/_Workshop_Data_driven_Soundsquatting_Generation%20\(7\).pdf](https://iris.polito.it/retrieve/handle/11583/2970511/1dd2efbb-9eed-4db5-be4f-ecfcc362572c/_Workshop_Data_driven_Soundsquatting_Generation%20(7).pdf)
- ¹¹ <https://incolumitas.com/2016/06/08/typosquatting-package-managers/>
- ¹² <https://medium.com/@alex.birsan/dependency-confusion-4a5d60fec610>
- ¹³ <https://www.akamai.com/blog/security-research/combosquatting-keyword-analysis-support>
- ¹⁴ <https://www.trendmicro.com/content/dam/trendmicro/global/en/research/23/c/iron-tiger-sysupdate-reappears-adds-linux-targeting/LOCs-iron-tiger-sysupdate-reappears-adds-linux-targeting.txt>
- ¹⁵ <https://urlscan.io/result/41e8b29f-55cc-4887-9186-41a064ffb2ac/>
- ¹⁶ <https://thehackernews.com/2022/07/microsoft-warns-of-large-scale-aitm.html>
- ¹⁷ <https://thehackernews.com/2023/03/microsoft-warns-of-large-scale-use-of.html>
- ¹⁸ <https://www.hackread.com/hackers-employee-accounts-twilio-internal-system/>
- ¹⁹ <https://www.feldmanauto.com/>
- ²⁰ <https://urlscan.io/result/98f407d6-96b4-4652-bd38-aa44470b5b78/>
- ²¹ <https://urlscan.io/result/98f407d6-96b4-4652-bd38-aa44470b5b78/>
- ²² <https://blogs.infoblox.com/cyber-threat-intelligence/scammers-first-on-the-scene-for-turkiyes-disaster-of-the-century/>
- ²³ <https://urlscan.io/result/4f295f57-7d46-49e9-94f6-d90858a4cfcb/>
- ²⁴ <https://www.coindesk.com/web3/2023/03/02/nft-trading-volumes-hit-2b-in-february-highest-since-luna-crash-thanks-to-blur/>
- ²⁵ <https://nftnow.com/guides/blurs-token-just-dropped-heres-what-you-need-to-know/>
- ²⁶ https://twitter.com/blur_io/status/1630290782211981312/
- ²⁷ <https://www.wired.com/story/youtube-bitcoin-scam-account-hijacking-google-phishing/>
- ²⁸ <https://twitter.com/FoolishBB/status/1627059614654279682>
- ²⁹ <https://www.bleepingcomputer.com/news/security/fake-crypto-giveaways-steal-millions-using-elon-musk-ark-invest-video/>
- ³⁰ <https://www.domaintools.com/>
- ³¹ <https://urlscan.io/result/8e94bf31-7295-47e8-9de4-756743937f46/>
- ³² <https://www.domaintools.com/>
- ³³ <https://urlscan.io/result/7f3c8f83-1922-4570-a9b1-1542e32ccc89/>
- ³⁴ <https://urlscan.io/result/f60f5548-4b54-4a97-add5-1f37a89f4e7e/#summary>
- ³⁵ <https://www.cisa.gov/news-events/cybersecurity-advisories/aa23-025a>
- ³⁶ <https://blogs.infoblox.com/cyber-threat-intelligence/cyber-campaign-briefs/dont-dial-that-number-distribution-of-phishing-lookalikes-through-fake-support-calls/>
- ³⁷ <https://urlscan.io/result/41a6ef99-fef1-4d08-80e1-623123280b6a/>
- ³⁸ <https://walletconnect.com/>
- ³⁹ <https://urlscan.io/result/a79ba8e3-9f9a-4a9c-b54b-b26a300afc23/>
- ⁴⁰ <https://bitkeep.com/>
- ⁴¹ <https://docs.flutter.dev/>
- ⁴² <https://www.welivesecurity.com/2023/02/16/these-arent-apps-youre-looking-for-fake-installers/>
- ⁴³ <https://www.virustotal.com/gui/file/271229d5d007baf5324fb2705b7a0b3751bd228bbdb08a86e7b7e2856bbf9b08>
- ⁴⁴ <https://urlquery.net/report/ef86060b-39e3-4e41-a480-a2b138ee0a49>
- ⁴⁵ <https://elifesciences.org/articles/54846>
- ⁴⁶ <https://chromium.googlesource.com/chromium/src/+main/docs/idn.md>
- ⁴⁷ https://wiki.mozilla.org/IDN_Display_Algorithm
- ⁴⁸ <https://www.xudongz.com/blog/2017/idn-phishing/>
- ⁴⁹ <https://www.usenix.org/system/files/sec21-hu-hang.pdf>
- ⁵⁰ <https://tbutler.org/2021/04/16/considering-the-plausibility-of-idn-homograph-attacks>



Infoblox unites networking and security to deliver unmatched performance and protection. Trusted by Fortune 100 companies and emerging innovators, we provide real-time visibility and control over who and what connects to your network, so your organization runs faster and stops threats earlier.

Corporate Headquarters
2390 Mission College Blvd, Ste. 501
Santa Clara, CA 95054

+1.408.986.4000
www.infoblox.com