Gen™

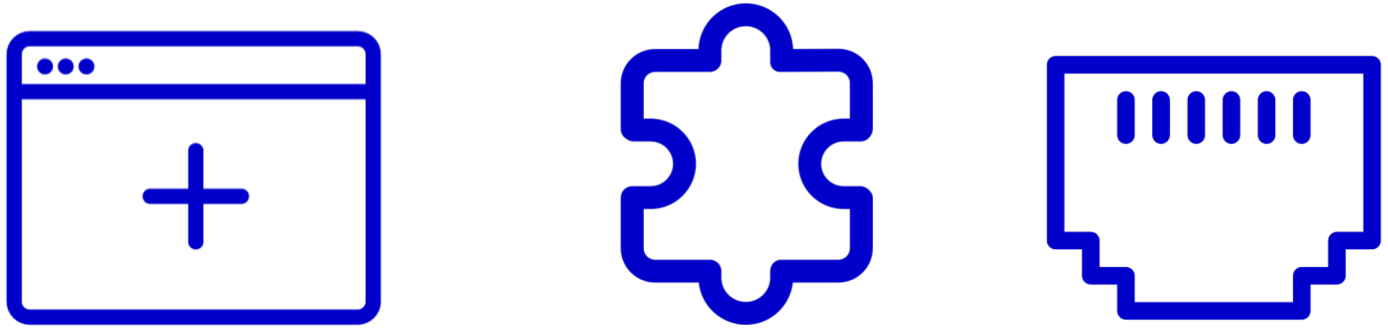
Rods with Laser Beams: Understanding Browser Fingerprinting on Phishing Pages

Iskander Sanchez-Rola, Leyla Bilge,
Davide Balzarotti, Armin Buescher,
Petros Efstathopoulos

32nd USENIX Security Symposium
August 9–11, 2023

Browser Fingerprinting

Websites have access to **information exposed** by the browser, such as the type and version, IP address, plugins, and the list of available fonts. By **combining these elements**, websites can generate fingerprints.



Browser Fingerprinting

This presents an attack surface for those using fingerprints for **privacy-invasive** purposes, but it has also been used for benign purposes—such as providing additional elements of (“zero-trust”) **authentication**.

Digital Fingerprint Theft

Complementing a user’s stolen credentials with the corresponding fingerprint significantly **increases the value** of the stolen assets.

Our Analysis

1.7M phishing websites were analyzed between Dec 2020 and Aug 2022

Our Analysis

1.7M phishing websites were analyzed between Dec 2020 and Aug 2022

We queried a large **phishing feed** and accessed the phishing websites immediately after they appeared in the feed

Our Analysis

1.7M phishing websites were analyzed between Dec 2020 and Aug 2022

We queried a large **phishing feed** and accessed the phishing websites immediately after they appeared in the feed

We focus on scripts that invoke fingerprinting-related API calls and later **harvest** the data

Our Analysis

1.7M phishing websites were analyzed between Dec 2020 and Aug 2022

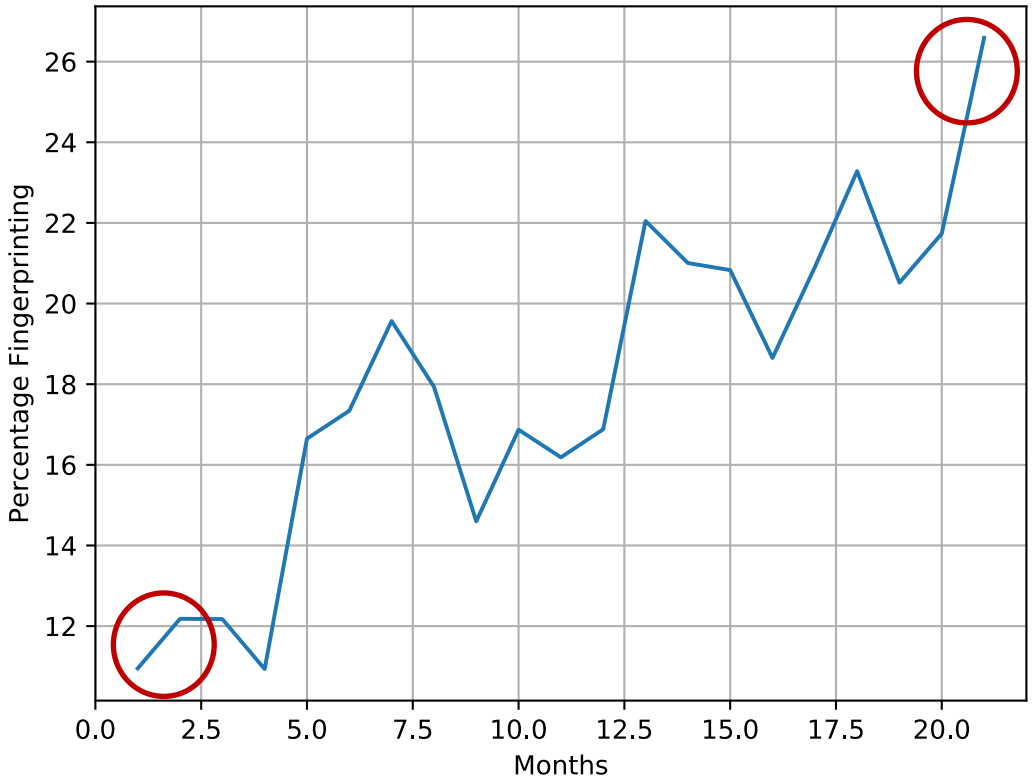
We queried a large **phishing feed** and accessed the phishing websites immediately after they appeared in the feed

We focus on scripts that invoke fingerprinting-related API calls and later **harvest** the data

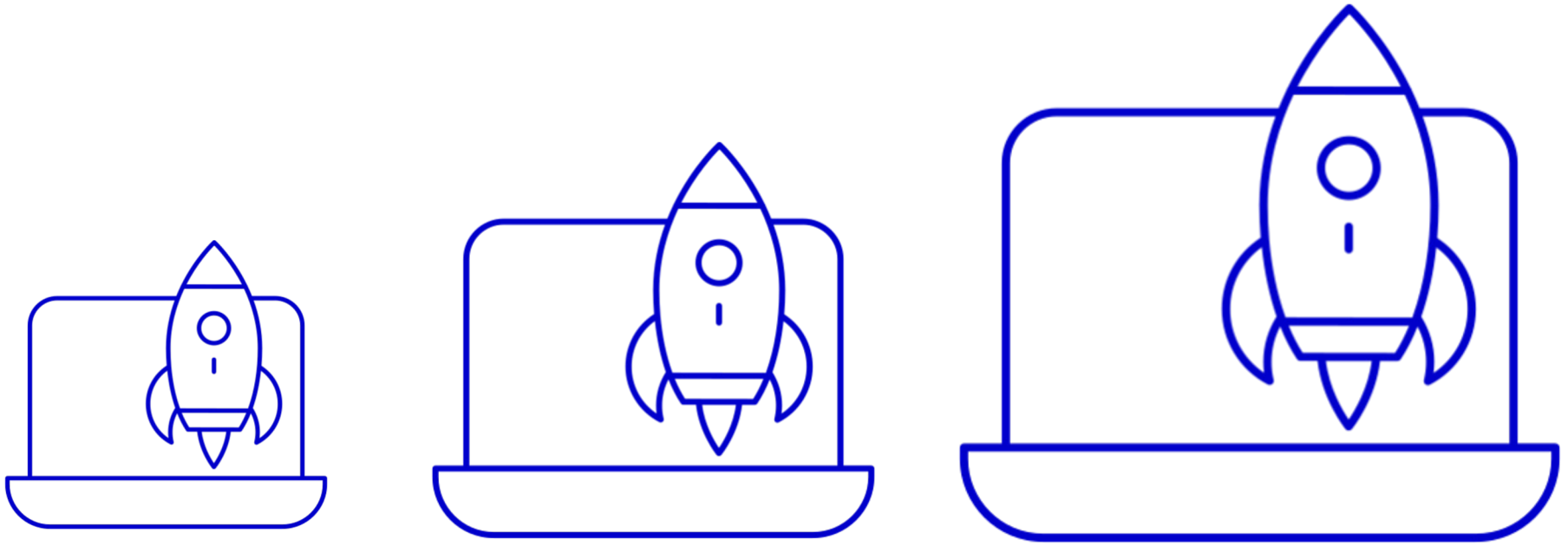
To assess the actual impact on real users, we used our **telemetry** data

The General Picture

We observe a dramatic increase on the number of phishing sites that adopt fingerprinting techniques, **from around 12% to over 26%.** This emphasizes the ever-growing interest of phishers in using fingerprinting functions.



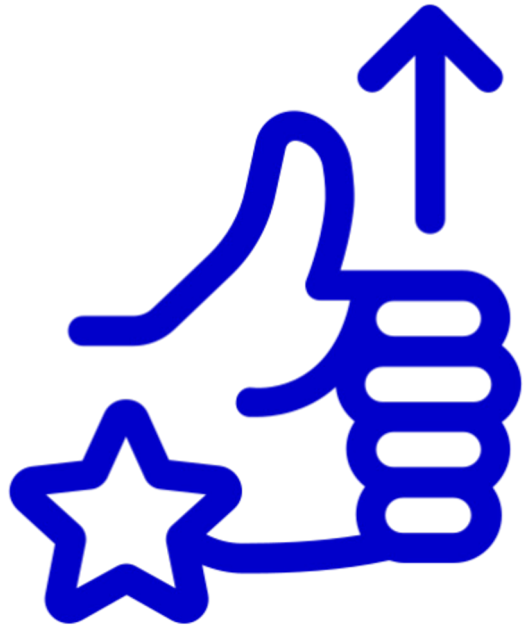
The General Picture



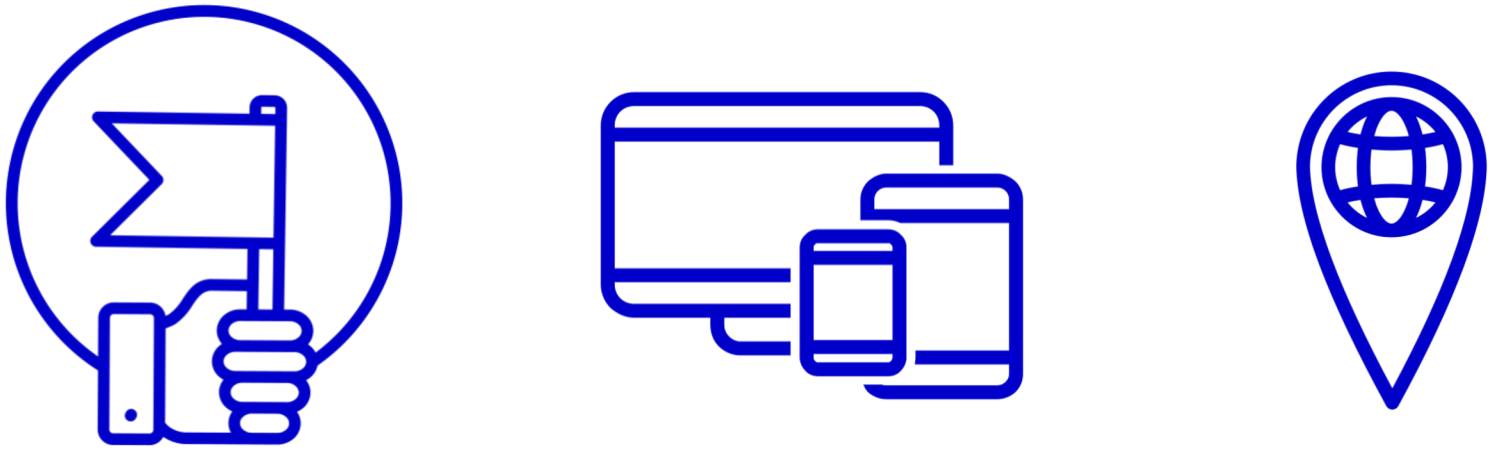
The fraction of phishing websites that use fingerprinting has **more than doubled** over the past two years. Today, **one in four** phishing sites invoke fingerprinting APIs.

Real-World Impact

Restricting the analysis to **most popular** phishing pages – computed according to the number of victims in the telemetry – we observe that the percentage of those using fingerprinting APIs significantly increases to **44.5%** if we consider the **top 50K**.



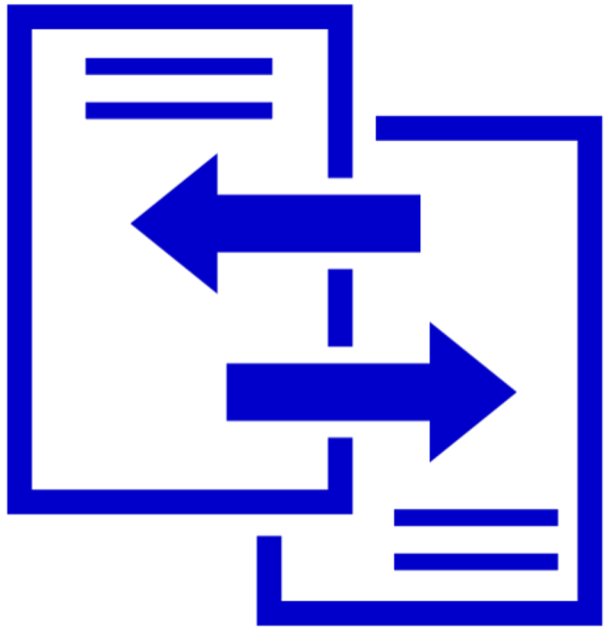
Real-World Impact



Phishing websites that are more successful at **attracting users**, target **more devices**, or whose victims span **more countries**, tend also to be the ones that use **fingerprinting the most**.

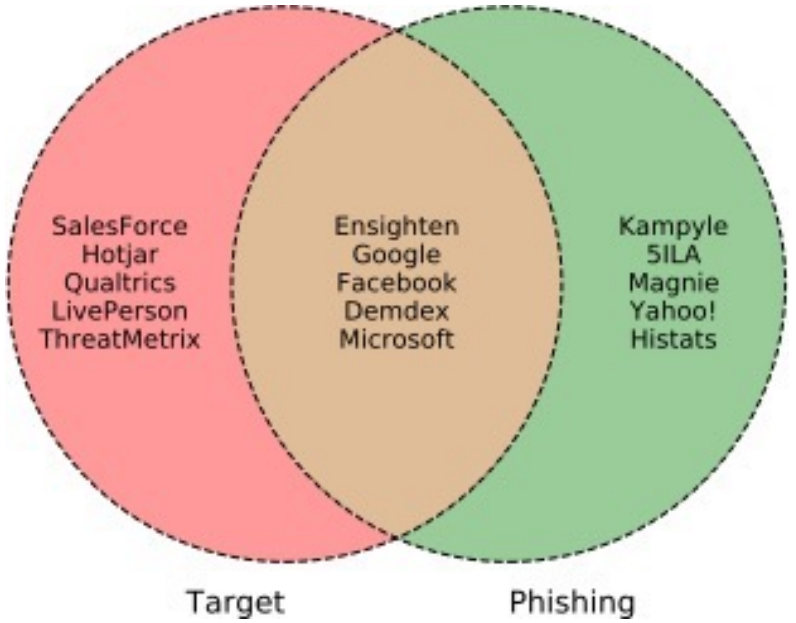
Additional Fingerprinting

On **53.7%** of the phishing pages, fingerprinting content is sent directly by the first-party. Moreover, **77.8%** of the websites which use third parties for fingerprinting **do not copy them** from the legitimate website they impersonate.



Additional Fingerprinting

While **top trackers** observed on legitimate websites are present in both cases, we also observe **less popular entries** that are not among the common trackers in legitimate websites. This shows that phishers tend to use **different trackers** than legitimate pages.

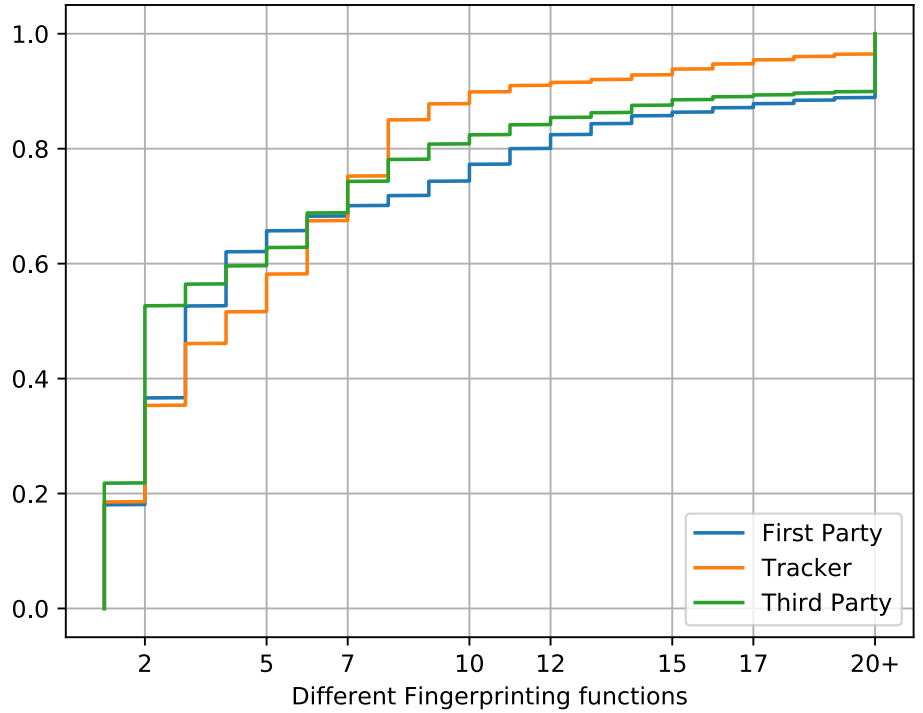


Additional Fingerprinting

Overall, **90%** of the phishing websites that send or store fingerprints, compute them from **newly included code** (not present in their target site).

Fingerprinting Functions

If we break down the number of functions called by first-party, tracker and third-party scripts, the average number of **fingerprinting functions** we observed in phishing pages is between **5.4** and **6.7**.



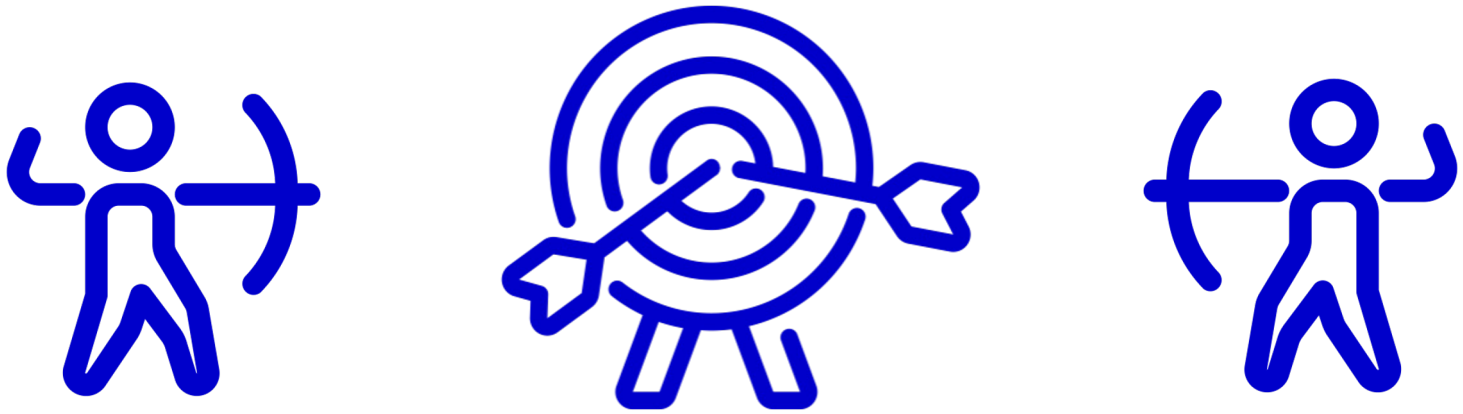
Fingerprinting Functions

Phishers often go beyond using basic approaches, and are starting to **implement advanced methods (24.6%)** such as webGL, canvas or audio fingerprinting, to refine the information needed for operations.

Target Brand Analysis

Financial Services, Social Networking and Technology/Internet are the most targeted categories with between half and three-quarters adopting **complex fingerprinting** techniques.

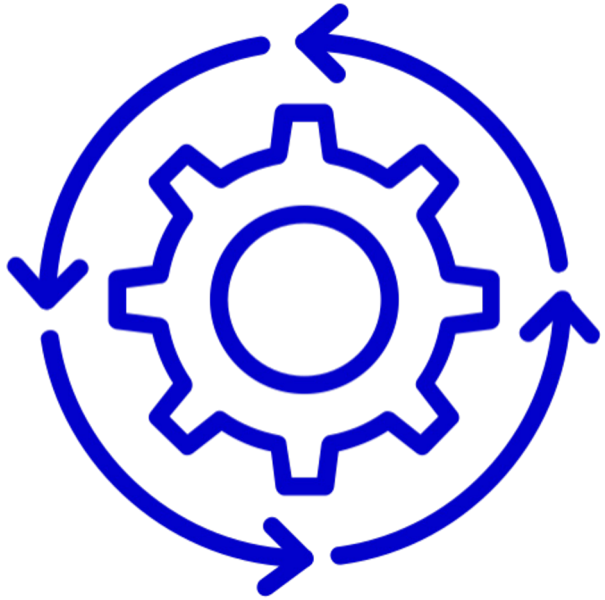
Target Brand Analysis



On average, over **80%** of highly targeted brands, categories or countries, implement **multiple fingerprinting** functions, as their benign counterparts do.

Comparing Phishing with Their Targets

One might expect phishers to simply **copy the exact same** fingerprinting functionality used by their targets. However, this would be a **poor and very time-consuming**. Instead, a better strategy would be to deploy **generic** fingerprinting code snippets, which could be **easily reused.**



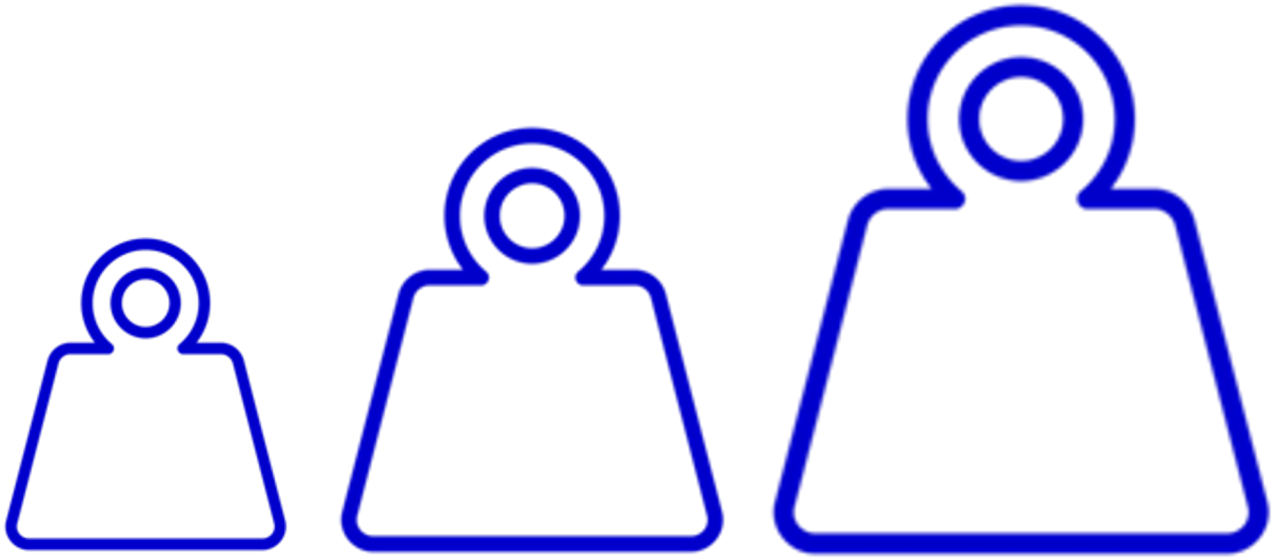
Comparing Phishing with Their Targets

Nearly **30%** of the targets have at least one phishing page that **matches** their fingerprinting functions, and **65%** of the phishers invoke **additional functions**.

Who are the Heavy Fingerprinters?

We computed a **signature** based on the list of **fingerprinting functions** used and their interactions. Through a manual analysis of some such cases, we discovered clusters that suggest **specific attacker groups** with distinct modes of operations.

Who are the Heavy Fingerprinters?



We were able to extract **32** unique **signatures** impersonating more than one target brand, matching up to **15 targets** from **7** different **categories**.

Thank you

Iskander Sanchez-Rola

Director Of Privacy Innovation

Iskander.Sanchez@GenDigital.com

GenDigital.com

Gen™