THE WEB CONFERENCE ACM

Research Track | Security

Phishing Vs. Legit: Comparative Analysis of Client-Side Resources of Phishing and Target Brand Websites

Kyungchan Lim Jaehwan Park Doowon Kim

University of Tennessee, Knoxville

Motivation

We aim to conduct a comparative analysis of phishing websites and their corresponding legitimate target brand websites to examine the utilization of client-side resources (e.g., JavaScript) in phishing attacks.
1) Gain insights into the construction and techniques of phishing websites
2) Suggest potential recommendations or mitigation against phishing attacks

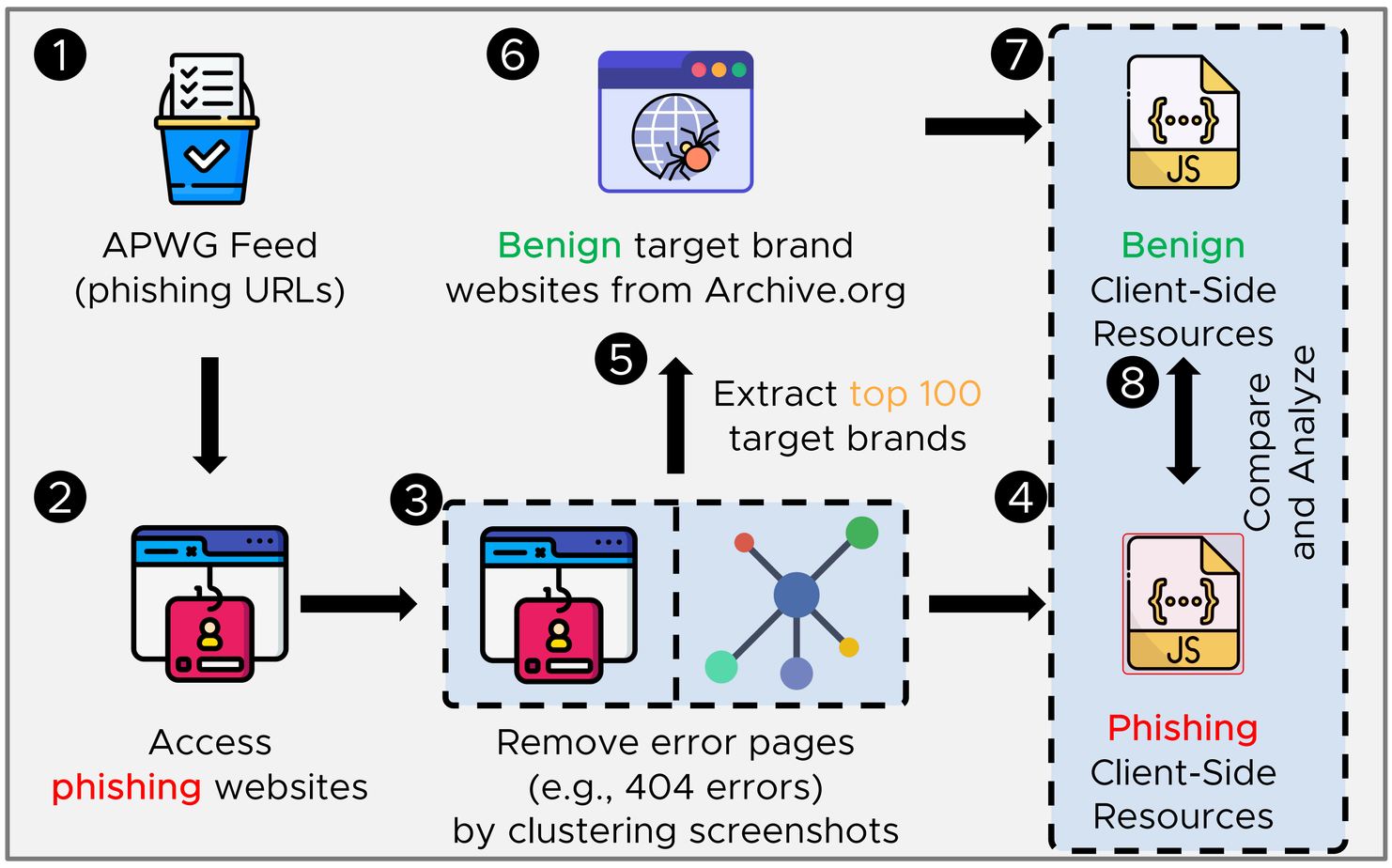
Research Questions

RQ1) What kind of client-side resources are employed in phishing websites?
RQ2) Which JavaScript libraries are widely prevalent in phishing websites?
RQ3) How similar are phishing websites and their corresponding legitimate target brand websites in terms of HTML structures?

Collected Datasets

Type	# of URLs	# of Domains
APWG URLs	15,747,193	1,545,253
Accessed URLs	7,067,778	1,135,264
Screenshots	6,125,810	939,103
Refined Dataset	3,388,997	**757,421**
Collection Period	July '21 – July '23 (25 months)	

Overview of Our Study



Results

1

Client-side Resources	Average Usage	Average Usage (%)
JavaScript	626,719	82.7%
CSS	547,660	72.3%
Favicon	265,182	35%
SVC	124,734	16.5%
CMS	55,135	7.3%

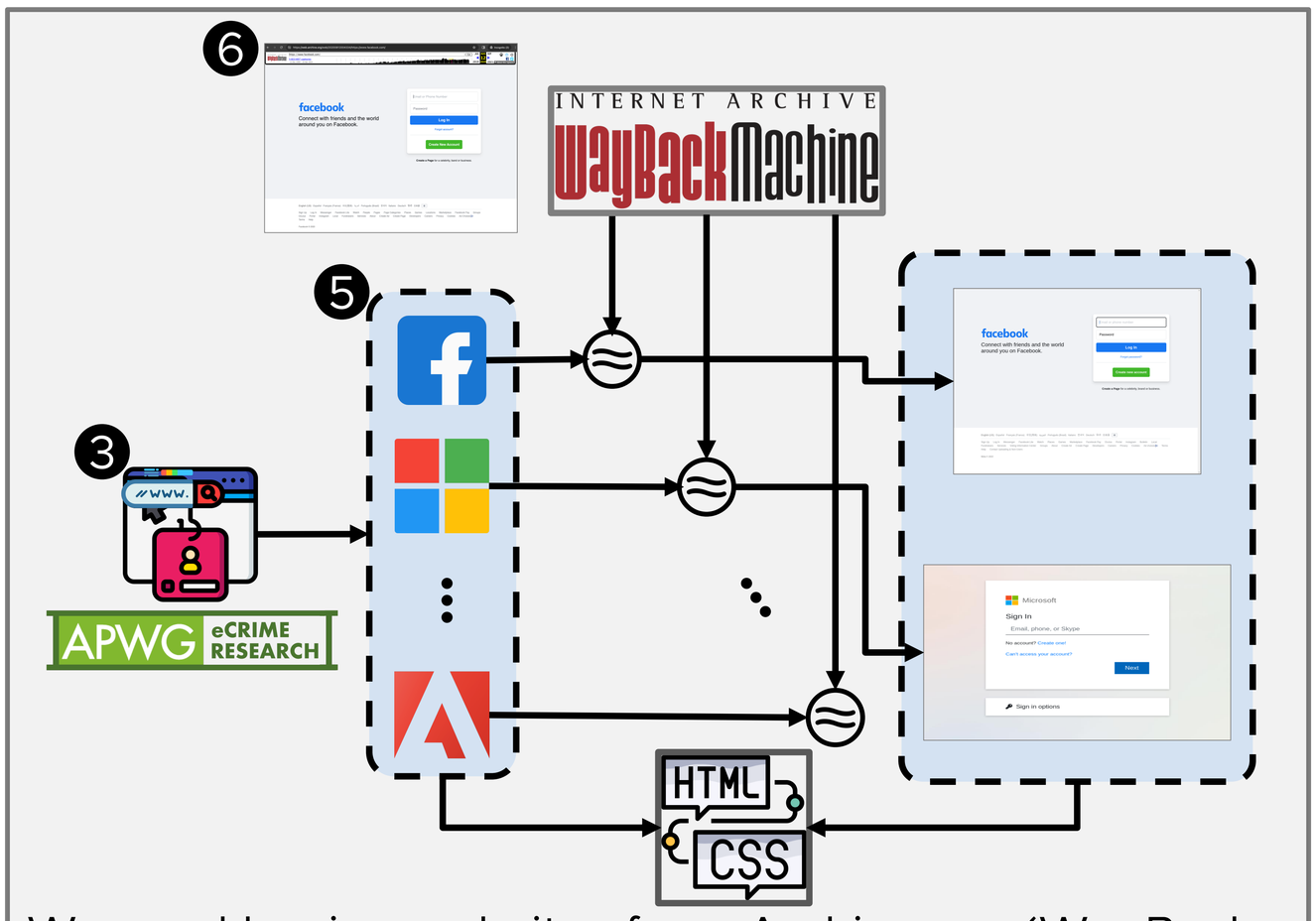
Client-resources used in phishing websites

2

Library	Remark
Clipboard.js	Popular in phishing
Select2	Popular in phishing
SweetAlert2	Popular in phishing
Axios	Only shown in Phishing
Socket.IO	Only shown in Phishing
Hammer.js	Only shown in Phishing

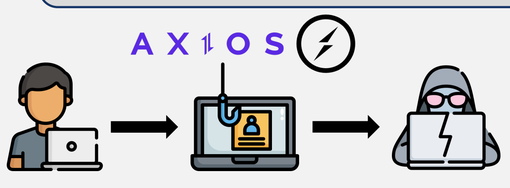
There are libraries that are popular on phishing and libraries only shown in phishing websites.

Comparison of Resources between Phishing and Target Benign Websites



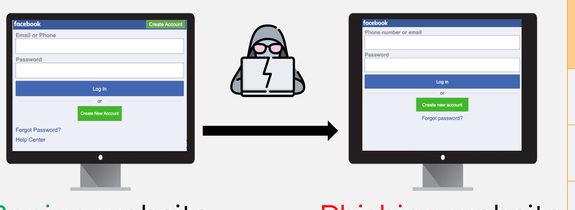
We used benign websites from Archive.org (WayBack Machine) then compare with our phishing websites.

```
d a t a . append ('email') ;
d a t a . append ('password');
```



Attackers utilize specific libraries to illicitly extract victims' information.

3



Benign website on 01/03/2018 — Phishing website on 07/14/2022

Brand	First Seen	Mimicked -Date	Diff.
Facebook	2021-07-11	2020-08-12	333
Microsoft	2021-07-11	2018-01-03	1,285
...			

Phishing sites mimic target brands from an average of 585.5 days older versions.

THE UNIVERSITY OF TENNESSEE KNOXVILLE