INF FAKULTÄT FÜR INFORMATIK

Der Webauftritt durchleuchtet:

das privatere, datensparsamere, nachhaltigere und IT-sicherere Web

SITSEC: DR2
14.06.2023

Jonas Morawietz
Bernhard Birnbaum
Meryem Lasri

Inhalt

1. Projektfortschritte
   a. Evaluation der EDPS-Erweiterung
   b. Website-Analysen
   c. Dynamische Analyse
2. Tabellarische Zusammenfassung
   a. Tools
   b. Referenzen
3. Aussicht
   a. Draft
   b. Nächste Schritte

1. Projektfortschritte: Eval. d. EDPS-Erweiterung

Webseite	Vergleichsmetrik	Ergebnisse ohne Akzeptieren des Cookie-Banners	Ergebnisse mit Akzeptieren des Cookie-Banners
https://www.docinsider.de	Third-Party Hosts	13 einzigartige Hosts	13 einzigartige Hosts
	Third-Party Beacons	5 einzigartige Beacons	5 einzigartige Beacons
	First-Party Cookies	2 Cookies	3 Cookies
	Third-Party Cookies	6 Cookies	5 Cookies
	Local Storage	2 Einträge	2 Einträge
https://www.gesundheit.de	Third-Party Hosts	12 einzigartige Hosts	81 einzigartige Hosts
	Third-Party Beacons	4 einzigartige Beacons	34 einzigartige Beacons
	First-Party Cookies	0 Cookies	3 Cookies
	Third-Party Cookies	3 Cookies	35 Cookies
	Local Storage	2 Einträge	39 Einträge
https://www.jameda.de	Third-Party Hosts	17 einzigartige Hosts	41 einzigartige Hosts
	Third-Party Beacons	5 einzigartige Beacons	14 einzigartige Beacons
	First-Party Cookies	1 Cookie	4 Cookies
	Third-Party Cookies	2 Cookies	21 Cookies
https://www.kliniken.de	Third-Party Hosts	1 einzigartiger Host	3 einzigartige Hosts
	Third-Party Beacons	0 einzigartige Beacons	2 einzigartige Beacons
	First-Party Cookies	0 Cookies	3 Cookies
	Third-Party Cookies	0 Cookies	2 Cookies
https://www.sanego.de	Third-Party Hosts	9 einzigartige Hosts	179 einzigartige Hosts
	Third-Party Beacons	4 einzigartige Beacons	65 einzigartige Beacons
	First-Party Cookies	2 Cookies	7 Cookies
	Third-Party Cookies	5 Cookies	220 Cookies
	Local Storage	3 Einträge	34 Einträge
https://www.seniorenportal.de	Third-Party Hosts	11 einzigartige Hosts	126 einzigartige Hosts
	Third-Party Beacons	4 einzigartige Beacons	52 einzigartige Beacons
	First-Party Cookies	10 Cookies	12 Cookies
	Third-Party Cookies	1 Cookie	126 Cookies
	Local Storage	2 Einträge	34 Einträge

- wenn Consent-Cookies übergeben:
  - durchschnittlich 7.42x mehr angefragte einzigartige Hosts
  - durchschnittlich 2.55x mehr First-Party-Cookies
  - durchschnittlich 38.6x mehr Third-Party-Cookies
  - durchschnittlich 8.31x mehr Web Beacons (Tracker)
- durch Erweiterung werden wesentlich mehr Analysedaten gesammelt

→ Akzeptieren des Cookie-Banners sollte bei einem realen Besuch von Websites in jedem Fall vermieden werden

1. Projektfortschritte: Website-Analysen

Website	**Ergebnisse: PrivacyScore, webbkoll, EDPS Website Evidence Collector** [PS/wk/EDPS]
https://www.kliniken.de/	einzigartige Third-Party-Hosts: 1/0/3, davon Tracker: 0/0/0, First-Party-Cookies: 0/0/3, Third-Party-Cookies: 0/0/2 **PrivacyScore** (*potentielle Schwachstellen*, da nicht gesetzt): HSTS-preloading, CSP-Header Referrer-Policy **webbkoll** (*potentielle Schwachstellen*, da inkorrekt/unvollständig implementiert): CSP, Referrer-Policy; *SRI*: 29 Objekte von unsicheren Quellen
https://www.docinsider.de/	einzigartige Third-Party-Hosts: 13/13/13, davon Tracker: 6/9/14, First-Party-Cookies: 6/2/3, Third-Party-Cookies: 1/2/5 **PrivacyScore** (*potentielle Schwachstellen*, da nicht gesetzt):HSTS-preloading, CSP-Header, XFO-Header, Referrer-Policy **webbkoll** (*potentielle Schwachstellen*, da inkorrekt/unvollständig implementiert): CSP, Referrer-Policy; *SRI*: 29 Objekte von unsicheren Quellen
https://www.jameda.de/	einzigartige Third-Party-Hosts: 17/16/41, davon Tracker: 6/3/14, First-Party-Cookies: 2/2/4, Third-Party-Cookies: 1/1/21 **PrivacyScore** (*potentielle Schwachstellen*, da nicht gesetzt): XFO-Header, X-XSS-Protection-Header, X-Content-Type-Header, Referrer-Policy **webbkoll** (*potentielle Schwachstellen*, da inkorrekt/unvollständig implementiert): CSP, Referrer-Policy; *SRI*: 29 Objekte von unsicheren Quellen 2 “Tracking-Pixel” (GIF) von verschiedenen Hosts
https://www.sanego.de/	einzigartige Third-Party-Hosts: 9/10/179, davon Tracker: 3/4/65, First-Party-Cookies: 9/8/7, Third-Party-Cookies: 0/0/220 **PrivacyScore** (*potentielle Schwachstellen*, da nicht gesetzt): XFO-Header, X-XSS-Protection-Header, X-Content-Type-Header, Referrer-Policy **webbkoll** (*potentielle Schwachstellen*, da inkorrekt/unvollständig implementiert): CSP, Referrer-Policy; *SRI*: 8 Objekte von unsicheren Quellen mehrere (10+) “Tracking-Pixel” (GIF, Jpeg) von verschiedenen Hosts, teilweise erst per Redirect
https://www.seniorenportal.de/	einzigartige Third-Party-Hosts: 10/10/126, davon Tracker: 3/2/52, First-Party-Cookies: 11/11/12, Third-Party-Cookies: 0/0/126, **PrivacyScore** (*potentielle Schwachstellen*, da nicht gesetzt): HSTS-preloading, CSP-Header Referrer-Policy,XFO-Header, X-Content-Type-Header **webbkoll** (*potentielle Schwachstellen*, da inkorrekt/unvollständig implementiert): CSP, Referrer-Policy; *SRI*: 5 Objekte von unsicheren Quellen
https://www.gesundheit.de/	einzigartige Third-Party-Hosts: 9/9/81, davon Tracker: 2/0/34, First-Party-Cookies: 2/2/3, Third-Party-Cookies: 1/1/35, **PrivacyScore** (*potentielle Schwachstellen*, da nicht gesetzt): CSP-Header, Referrer-Policy **webbkoll** (*potentielle Schwachstellen*, da inkorrekt/unvollständig implementiert): CSP, Referrer-Policy; *SRI*: 49 Objekte von unsicheren Quellen

1. Projektfortschritte: Website-Analysen

Website	Ergebnisse: Firefox ESR, NoScript, uBlockOrigin
https://www.kliniken.de/	ohne Blockiermaßnahmen: 1.69 MB bei 31 Anfragen mit Blockiermaßnahmen*: 1.69 MB bei 31 Anfragen NoScript blockiert alles außer “kliniken.de”, um Kernfunktionalität zu erhalten
https://www.docinsider.de/	ohne Blockiermaßnahmen: 4.87 MB bei 84 Anfragen mit Blockiermaßnahmen*: 3.50 MB bei 46 Anfragen NoScript blockiert alles außer “docinsider.de”, um Kernfunktionalität zu erhalten
https://www.jameda.de/	ohne Blockiermaßnahmen: 4.46 MB bei 66 Anfragen mit Blockiermaßnahmen*: 2.33 MB bei 45 Anfragen NoScript blockiert alles außer “jameda.de”, “docplanner.com”, “sentry.io”, “maps.googleapis.com”, um Kernfunktionalität zu erhalten
https://www.sanego.de/	ohne Blockiermaßnahmen: 11.96 MB bei 370 Anfragen mit Blockiermaßnahmen*: 1.11 MB bei 11 Anfragen NoScript blockiert alles außer “sanego.de”, um Kernfunktionalität zu erhalten
https://www.seniorenportal.de/	ohne Blockiermaßnahmen: 1.88 MB bei 59 Anfragen mit Blockiermaßnahmen*: 0.99 MB bei 45 Anfragen NoScript blockiert alles außer “seniorenportal.de”
https://www.gesundheit.de/	ohne Blockiermaßnahmen: 5.72 MB bei 160 Anfragen mit Blockiermaßnahmen*: 625 KB bei 21 Anfragen NoScript blockiert alles außer “gesundheit.de” um Kernfunktionalität zu erhalten

*Blockiermaßnahmen, die die Funktionalität **nicht** einschränken

1. Projektfortschritte: Website-Analysen

Website	**Ergebnisse: Wireshark** [S = Server, C = Client]
https://www.kliniken.de/	ohne Blockiermaßnahmen: 15 einzigartige DNS-Anfragen, S→C: 354 Pakete (645 KB), C→S: 322 Pakete (35 KB) mit Blockiermaßnahmen*: 15 einzigartige DNS-Anfragen, S→C: 354 Pakete (645 KB), C→S: 322 Pakete (35 KB)
https://www.docinsider.de/	ohne Blockiermaßnahmen: 25 einzigartige DNS-Anfragen, S→C: 1322 Pakete (2918 KB), C→S: 922 Pakete (129 KB) mit Blockiermaßnahmen*: 15 einzigartige DNS-Anfragen, S→C: 661 Pakete (2155 KB), C→S: 574 Pakete (53 KB)
https://www.jameda.de/	ohne Blockiermaßnahmen: 60 einzigartige DNS-Anfragen, S→C: 4376 Pakete (5050 KB), C→S: 2291 Pakete (457 KB) mit Blockiermaßnahmen*: 21 einzigartige DNS-Anfragen, S→C: 2228 Pakete (3018 KB), C→S: 819 Pakete (128 KB)
https://www.sanego.de/	ohne Blockiermaßnahmen: 153 einzigartige DNS-Anfragen, S→C: 8369 Pakete (13 MB), C→S: 7073 Pakete (1735 KB) mit Blockiermaßnahmen*: 12 einzigartige DNS-Anfragen, S→C: 171 Pakete (528 KB), C→S: 153 Pakete (17 KB)
https://www.seniorenportal.de/	ohne Blockiermaßnahmen: 116 einzigartige DNS-Anfragen, S→C: 3647 Pakete (2153KB), C→S: 2273 Pakete (1978 KB) mit Blockiermaßnahmen*: 42 einzigartige DNS-Anfragen, S→C: 1880 Pakete (1999 KB), C→S: 745 Pakete (203 KB)
https://www.gesundheit.de/	ohne Blockiermaßnahmen: 162 einzigartige DNS-Anfragen, S→C: 3144 Pakete (3502KB), C→S: 2910 Pakete (747 KB) mit Blockiermaßnahmen*: 22 einzigartige DNS-Anfragen, S→C: 1574 Pakete (2345 KB), C→S: 1388 Pakete (162 KB)

*Blockiermaßnahmen, die die Funktionalität **nicht** einschränken

1. Projektfortschritte: Dynamische Analyse

- Ziel:
  -Erstellung eines Programms, um die Überprüfung von Trackern zu automatisieren, indem Netzwerk Aufzeichnungen (Wireshark) mit der DuckDuckGo Tracker Radar Liste abgeglichen werden.
  Der Hauptfokus liegt darauf, die Privatsphäre der Benutzer zu schützen.

- Fortschritte:
  -Die verwendete methode in SPASS2 untersuchen
  -Programm für die Automatisierung der dynamischen Analyse erstellen

1. Projektfortschritte: Dynamische Analyse

Dynamische Analyse in SPASS2:

- Verwendung von Tshark zur Erfassung von Netzwerkdaten über das Terminal
- Filtern von Source Hosts und Destination Hosts mit:
  - Auflösung der IP-Adresse in Domain Names
  - Auflösung der MAC Adresse und Network Adresse
- Ergebnis dieser Analyse durchgehen und mit den Tracker Liste (DisconnectMe Liste,..) vergleichen
- Gefundene Tracker speichern

1. Projektfortschritte: Dynamische Analyse

Dynamische Analyse in SPASS1:

- Automatisierte Öffnen eines Browsers, von dem aus Wireshark den Netzwerkverkehr erfassen wird
- im geöffneten Browser surfen
- wenn nicht mehr aufgezeichnet werden soll, Browser schließen (Aufzeichnung mit Tshark)
- Ergebnis der Analyse speichern
- Vergleich der Tshark Analyse mit der DuckDuckGo Liste
- Gefundene Tracker speichern

2. Tabellarische Zusammenfassung: Tools

Ausgewählte Werkzeuge:

Titel	Link	Beschreibung
Testerstick	intern	rauscharme Untersuchungsumgebung
Website-Evidence-Collector (EDPS)	https://github.com/EU-EDPS/website-evidence-collector	automatisierte Analyse von Speicherung und Transfer von persönlichen Daten
ungoogled Chromium	https://github.com/ungoogled-software/ungoogled-chromium	Google Chrome without Google
PrivacyScore	https://privacyscore.org/	bewertet Websites hinsichtlich einer Reihe von Sicherheits- und Datenschutzfunktionen
webbkoll	https://webbkoll.dataskydd.net/	prüft, welche Datenschutzmaßnahmen eine Website ergriffen hat
Firefox ESR	https://www.mozilla.org/de/firefox/all/#product-desktop-esr	Webbrowser mit Netzwerkanalyse
NoScript	https://noscript.net/	Blockierung bzw. kontrollierte Freigabe von JavaScript
uBlockOrigin	https://ublockorigin.com/de	Open-Source-Werbeblocker mit Zusatzfunktionen
Wireshark	https://www.wireshark.org/	Netzwerkprotokoll-Analyse und Entschlüsselung von Traffic (SSLKEYLOGFILE)

2. Tabellarische Zusammenfassung: Referenzen

Wissensbasis:

Titel	Quelle
[AKL+20] Introduction to Being a Privacy Detective: Investigating and Comparing Potential Privacy Violations in Mobile Apps Using Forensic Methods	ISBN 978-1-61208-821-1, pp 60-68, 2020
[Rie23] GitHub - EU-EDPS/website-evidence-collector: The tool Website Evidence Collector (WEC) automates the website evidence collection of storage and transfer of personal data.	https://edps.europa.eu/edps-inspection-software_en
tshark(1) Manual Page	https://www.wireshark.org/docs/man-pages/tshark.html
How to convert .img to usable VirtualBox format	https://superuser.com/questions/554862/how-to-convert-img-to-usable-virtualbox-format

3. Aussicht: Draft

- Hauptgliederung und Template übernommen
- Unterpunkte für einzelne Sub-Tasks erstellt und mit Bullet-Points befüllt
- Tabellen mit Ergebnissen pro Website im Anhang
- genutzte Consent-Cookies im Anhang
- Notwendigkeit vollständiger Host- bzw. Cookie-Listen?

3. Aussicht: nächste Schritte

- Interpretation der Ergebnisse
- Schreiben des Abschlussberichts
- “Dynamische Analyse” erweitern:
  -Verknüpfe jede Domain mit ihrem Unternehmen
  -Ermögliche dem Programm, nicht nur die Netzwerk-Analyse mit DuckDuckGo zu vergleichen, sondern auch mit jeder anderen Liste, solange sie ein Repository hat.

3. Code

```
#!/bin/bash
clear
read -p 'Give a name to the folder: ' folder
path=$HOME/AppDynamics/$folder

echo A browser will open now, to capture traffic use only this
echo browser. After you are "done" with capturing please close
echo the browser
sleep 5

 if [ -d "$path" ]; then rm -Rf $path; fi
mkdir $path
tshark -w $path/recording.pcapng &
export SSLKEYLOGFILE=$path/sslkeys.txt
/home/tester/firefox/firefox -profile /home/tester/.mozilla/firefox/r8l23gfa.default-esr

killall tshark
chmod -R a+rwx $path

tshark -r $path/recording.pcapng -N dmnNtv -T fields -e ip.src_host > $path/recording.txt
tshark -r $path/recording.pcapng -N dmnNtv -T fields -e ip.dst_host >> $path/recording.txt

sed 's/www[[:punct:]]//' $path/recording.txt > $path/recording_.txt
List=$HOME/AppDynamics/Tracker/DDGList_domain.txt
Host=$path/recording_.txt
Date=`date +"%Y-%m-%d %T"`
path2=$HOME/AppDynamics/Results
mkdir -p $path2
mkdir -p $HOME/AppDynamics/Results/$folder

perl -i -ne 'print if ! $a{$_}++' $Host



counter=0
>"Results/$folder/$Date.txt"

echo "The trackers found : " >>"Results/$folder/$Date.txt"

for line in `cat $Host`
do
      if grep -q $line $List;then
          counter=$((counter+1))
          echo $line>>"Results/$folder/$Date.txt"
      fi
done
```

Danke für Ihre Aufmerksamkeit!