



Cookieless Monster

Or Lavi
Yonatan Bitton





מפלצת חסרת עוגיות – אתם אף פעם לא לבד...

Cookieless Monster -
Exploring the Ecosystem of Web-based Device Fingerprinting
2013 IEEE Symposium on Security and Privacy





מי אנחנו



אור לביא

מפתח ביחידה מסווגת

סטודנט במכללה למנהל שנה ג'

מתמחה סייבר

@klavior

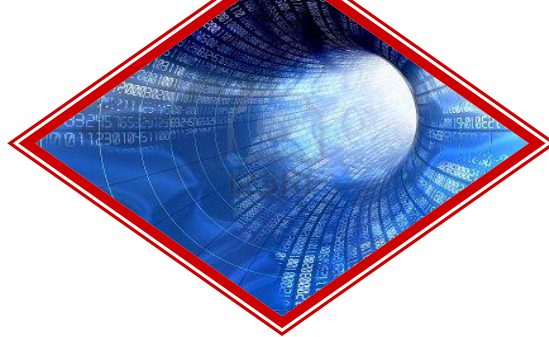
יונתן ביטון

רש"צ פיתוח ביחידה מסווגת

סטודנט במכללה למנהל שנה ג'

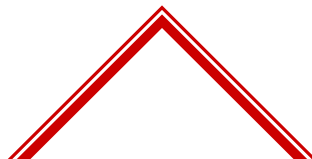
מתמחה סייבר

@bityob



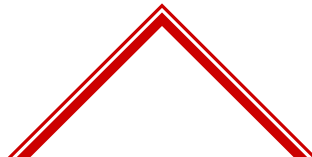
רקע

- מיליארדי משתמשים סורקים כל יום את האינטרנט
- יש אתרים עם מיליארדי משתמשים
- מעקב אחרי משתמשים וההרגלים שלהם –
- משתלם לחברות הפרסום
- פוגע בפרטיות של המשתמשים
- האם זה אפשרי לשמור על פרטיות ברחבי האינטרנט?



מהלך השיעור הקרוב

- מושגי יסוד והיסטוריה
- נבין למה כל אחד מאתנו מיוחד ואין שני לו
- נחקור מימושי fingerprint של חברות מסחריות
- נבחן את השימוש של fingerprint ברחבי האינטרנט
- נסביר על אפשרויות חדשות למציאת fingerprint שלא קיימות בשוק
- תוספי דפדפן ש"שומרים" על פרטיות המשתמשים
- הדגמה והוכחה שלכל אחד מאתנו יש fingerprint ששייך רק לו



http is a stateless protocol -

הפרוטוקול אינו מספק שום אמצעי לאחסון מידע מהמשתמש בין הבקשות.

מטפל בכל בקשה כטרנזקציה נפרדת.

- no state <- אין פרטים אישיים -
אין מייל אישי, קניות, שירותי בנק וכדומה.



1993



"On the Internet, nobody knows you're a dog."

©The New Yorker Collection 1993 Peter Steiner
From cartoonbank.com. All rights reserved.

מה הפתרון? Cookies !

כמעט כל האתרים
ברחבי האינטרנט
משתמשים ב
.Cookies



stonybrook.edu	username=nick; Date=30/09/2014;
google.com	g1=190233213; g2=afasfdioujewf;
slashdot.org	adheses_count=2; bcn=e4f5d957-00;

עוגייה (לפי ויקיפדיה) - היא מחרוזת אותיות ומספרים, המשמשת לאימות, למעקב ולאגירת מידע על אודות גולש באתר אינטרנט, כגון שמירת העדפות המשתמש.

How cookies work



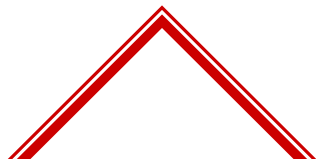
Source: www.flatworldknowledge.com

3rd Party Tracking – מעקב בעזרת שירותי צד שלישי (שירותים זרים)

"פתאום" כל מיני אתרים שאף פעם לא שמענו עליהם
הצליחו לבנות פרופיל דפדפן של משתמשים ולמכור אותם
לחברות פרסום.



איך הם עושים את זה?



3rd Party Tracking – מעקב בעזרת שירותי צד שלישי (שירותים זרים)

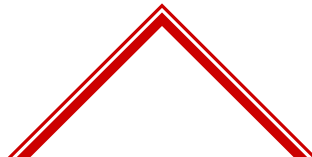


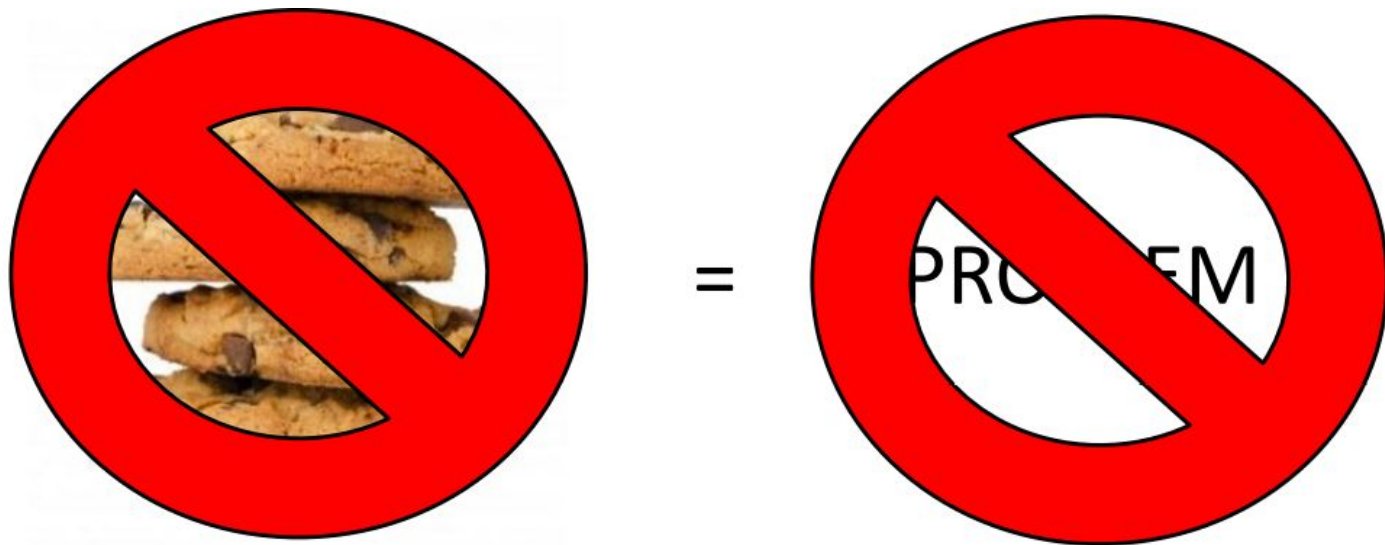
תגובת המשתמשים לשירותי המעקב -

1. $\frac{1}{3}$ מהמשתמשים מוחקים את cookies שלהם פעם בחודש.
2. נבנו תוספים שמגלים אתרי third party נסתרים שעוקבים אחרי משתמשים.
3. הומצאה גלישה בסתר של דפדפנים.
4. חוק אירופאי - שמירה של cookies מותרת רק באישור המשתמש.

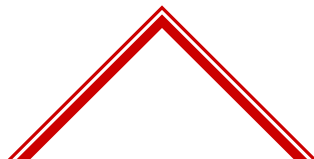


= PROBLEM





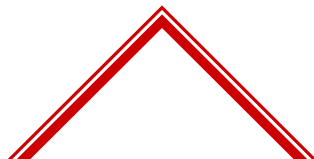
נכון?



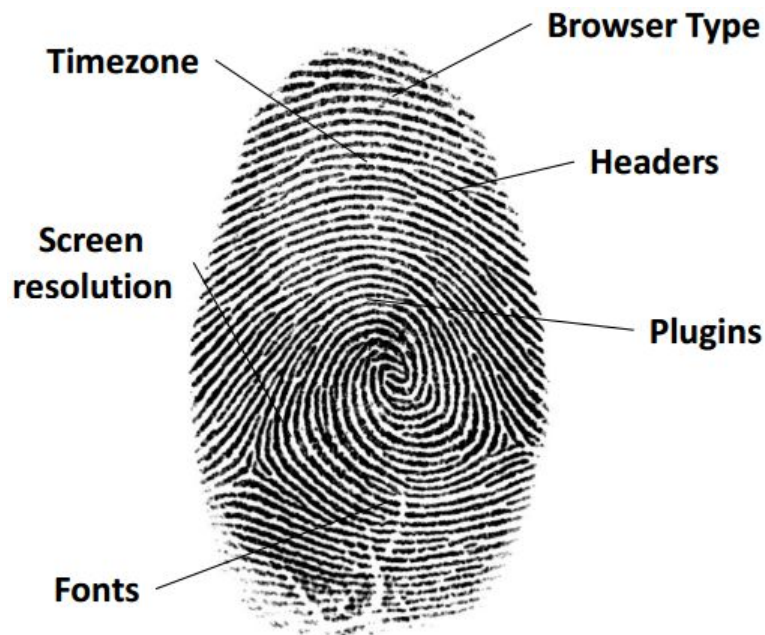
מה אם הייתי אומר לכם לכם ש....
לא חייבים cookies כדי לעקוב אחרי משתמשים

Web-Based Device fingerprinting

- הוצג לראשונה ע"י Eckersley ב-2010.
- מעקב מוסתר מהמשתמשים
- מאוד קשה להימנע ממנו
- משלב את כל המאפיינים של המשתמש ל-fingerprint ייחודי לכל משתמש.



Web-Based Device fingerprinting



- **לכל משתמש יש מאפיינים שיש לו אבל אין לכולם - למשל, לX אחוז מהמשתמשים יש CHROME בגרסה Y.**
- **מכיוון שהדפדפן הוא דבר מאוד מורכב אז לכל משתמש יש המון מאפיינים.**
- **אם נאחד יחדיו את כל המאפיינים של המשתמש ניצור fingerprint חד ערכי שמתאים רק לאותו משתמש.**

Are you unique?

Yes! (You can be tracked!)

36.94 % of observed browsers are **Chrome**, as yours.

2.73 % of observed browsers are **Chrome 56.0**, as yours.

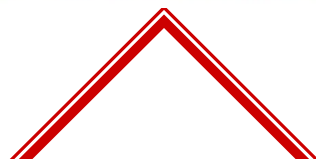
56.81 % of observed browsers run **Windows**, as yours.

13.75 % of observed browsers run **Windows 10**, as yours.

0.09 % of observed browsers have set "**he**" as their primary language, as yours.

3.38 % of observed browsers have **UTC+3** as their timezone, as yours.

However, your full fingerprint is unique among the 344552 collected so far. Want to know why?



מסחרי "Fingerprinting"





Fingerprinting של חברות מסחריות

בדקנו את היכולות של 3 חברות מסחריות

מצאנו את הדומיינים שהחברות הללו משתמשות בהם בשביל לעשות



Fingerprinting


מצאנו אתרים שמשתמשים בהם, בשביל חילוץ הקוד



ביצענו "הנדסה לאחור" (Reversing) של הקוד

ועשינו השוואה בין החברות עצמן





Fingerprinting Category	Panopticklick	BlueCava	Iovation ReputationManager	ThreatMetrix
<i>Browser customizations</i>	Plugin enumeration _(JS) Mime-type enumeration _(JS) ActiveX + 8 CLSIDs _(JS)	Plugin enumeration _(JS) ActiveX + 53 CLSIDs _(JS) Google Gears Detection _(JS)		Plugin enumeration _(JS) Mime-type enumeration _(JS) ActiveX + 6 CLSIDs _(JS) Flash Manufacturer _(FLASH)
<i>Browser-level user configurations</i>	Cookies enabled _(HTTP) Timezone _(JS) Flash enabled _(JS)	System/Browser/User Language _(JS) Timezone _(JS) Flash enabled _(JS) Do-Not-Track User Choice _(JS) MSIE Security Policy _(JS)	Browser Language _(HTTP, JS) Timezone _(JS) Flash enabled _(JS) Date & time _(JS) Proxy Detection _(FLASH)	Browser Language _(FLASH) Timezone _(JS, FLASH) Flash enabled _(JS) Proxy Detection _(FLASH)
<i>Browser family & version</i>	User-agent _(HTTP) ACCEPT-Header _(HTTP) Partial S.Cookie test _(JS)	User-agent _(JS) Math constants _(JS) AJAX Implementation _(JS)	User-agent _(HTTP, JS)	User-agent _(JS)
<i>Operating System & Applications</i>	User-agent _(HTTP) Font Detection _(FLASH, JAVA)	User-agent _(JS) Font Detection _(JS, FLASH) Windows Registry _(SFP)	User-agent _(HTTP, JS) Windows Registry _(SFP) MSIE Product key _(SFP)	User-agent _(JS) Font Detection _(FLASH) OS+Kernel version _(FLASH)
<i>Hardware & Network</i>	Screen Resolution _(JS)	Screen Resolution _(JS) Driver Enumeration _(SFP) IP Address _(HTTP) TCP/IP Parameters _(SFP)	Screen Resolution _(JS) Device Identifiers _(SFP) TCP/IP Parameters _(SFP)	Screen Resolution _(JS, FLASH)

Any questions? □



Fingerprinting דרך שימוש בפלאגינים פופולאריים

Flash

- משתמשים בו בשביל להציג תוכן Media שבעבר לא יכלו בעזרת Html
- כל החברות בחרו להשתמש בFlash
- למרות שFlash ידוע בבעיות הביצועים שלו ושטכנולוגיות חדשות כמו Html 5 יכולות להחליף אותו, הוא עדיין [2013] נמצא בשימוש רב
- הופתענו לגלות שה-API של Flash חושף יותר מידע על המערכת הפעלה מאשר פונקציות מקבילות וזהות בJavaScript



Fingerprinting דרך שימוש בפלאגינים פופולאריים

Flash

● דוגמא -

- בפיירפוקס על שרת לינוקס 64 ביט -
כאשר "שואלים" את הדפדפן בנוגע למערכת ההפעלה של המשתמש,
הוא מחזיר: *Linux x86_64*
לעומת זאת, ב-Flash התשובה היא: *Linux 3.2.0-26-generic*
○ בעצם מספק את גירסת הקרנל הספציפית, בעייתי מבחינת פרטיות ואבטחה



Fingerprinting דרך שימוש בפלאגינים פופולאריים

Java

- במפתיע אף אחת מהחברות לא השתמשה בJava, לצורך איסוף המידע
- כנראה הסיבה היא עקב המודעות הגבוהה לחולשות אבטחה של Java בדפדפן, שממילא הוביל לירידה חדה בשימוש בJava אצל המשתמשים



Fingerprinting דרך זיהוי *Font*-ים בדפדפן

- רשימת ה-Fontים של המערכת הפעלה יכולה גם לשמש כחלק מה-Fingerprint של

המשתמש

Archer ARCTIC Arctic *Santa Fe* **SANTAFE**
Chicago **CHICAGO** *New City* *Furistuff* Andy **ANDY**
 Kedzie **KEDZIE** Comic **COMIC** Massey **MASSEY** *MATISSE*
Darwycke **KEYSTER** *Spiffy* MEAD Mead *War Bird*
Squirt *Blizzard* *Anaze* Tango Normal *Brisk* Normal Engraved
French Script Koffee *Sherwood* *Swing* Bold Jester *Phyllis*
 Tekton, **TEKTON** *Boulevard* *Fire* Normal *Vladimir* *Pepita*
Brush, **BRUSH** *Brush Script* Freeze, **FREEZE** *Lynda* Cursive Chancery
Script Bold **PAPER** *Formal Script* *Mythic*, **MYSTIC** *Lucy* Normal
SNYDER SPEED *Surfer*, **SURFER** **Marker**, **MARKER** *Cezanne*
 Calligrapher *Liberate* *Shell* Normal *Sweden* Normal *Park* *Place*
Broach, **BROACH** **LITHOS** *Blew*, **BLEW** **BARBEDOR**, *Barbedor*
 Caxton, **CAXTON** **COPPERPLATE** **DEXTOR** *Garmond*, **GARMOND**
GOUDY, *Goudy* **Serpentine** **WIDE Latin**
 Avant Garde **ARIAL**, arial *Bobo*, **BOBO** **BRITANNIC**, *Britannic*
AUREA *Bordeaux*, **BORDEAUX** *Compacta*, **COMPA** *Formata*, **FORMATA**
Futura, **FUTURA** **F Gothic**, **F GOTHIC** *Tech*, **TECH** *ANNA*



Fingerprinting דרך זיהוי Font-ים בדפדפן

Plugin-based detection

- הדפדפן עצמו לא מספק את רשימת הFont-ים כמובן
- אבל לFlash (זכור לטוב) יש מתודה לקבלת רשימת הFont-ים...
- בנוסף, חוץ מהרשימה עצמה, גם סדר הפונטים הוא דינאמי



Fingerprinting דרך זיהוי *Font*-ים בדפדפן

Side-channel inference

- כל החברות הכילו מתודה ספציפית עבור font detection
- המתודה מקבלת שם פונט ויוצרת Div ובתוכו טקסט עם הפונט, ואח"כ בודקת מה הגודל שלו
- את התוצאה משווים מול רשימה קבועה של גדלים, בשביל לדעת-
 - האם הפונט קיים (אם לא קיים, הדפדפן משתמש בפונט דיפולטי)
 - באיזה דפדפן מדובר - כל דפדפן מציג קצת באופן שונה את אותו פונט
- כל הבדיקות האלה נעשות ב-Iframe מוסתר, כך שהמשתמש כלל לא מודע לבדיקה הזו



Font זיהוי - ים בדפדפן Fingerprinting

String

Dimensions

I_DO_NOT_NEED_FLASH

500 x 84

I_DO_NOT_NEED_FLASH

520 x 84

I_DO_NOT_NEED_FLASH

580 x 87

I_DO_NOT_NEED_FLASH

399 x 82



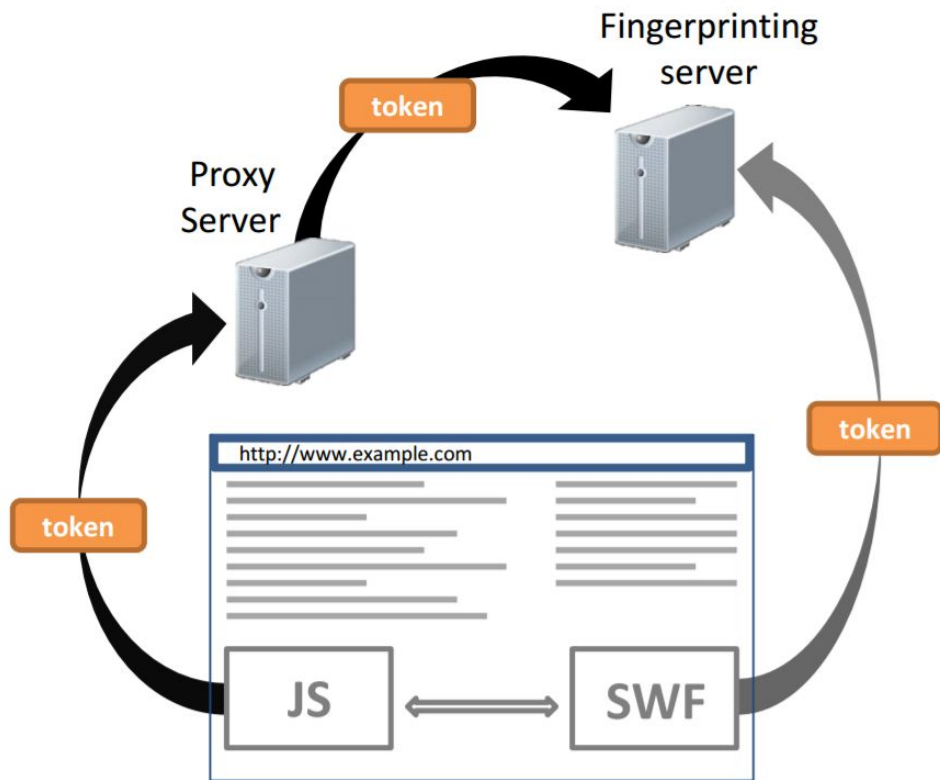
זיהוי של HTTP Proxies

- אחד מהמאפיינים שהכי קשה לזייף זה הכתובת IP של המשתמש
- המשתמש חייב להשתמש בכתובת אמיתית ותקינה בשביל שהתקשורת תתנהל כמו שצריך
- אחד הפתרונות הוא להשתמש בשרת פרוקסי שמנתב את התעבורה דרכו, וכך השרת Web מקבל את הכתובת IP של השרת פרוקסי
- בנוסף, קיימים גם תוספים לדפדפן, שמנתבים חלק מהתעבורה לכתובות מסוימות דרך פרוקסי לפי סט חוקים קבוע
- כך בעצם השרת מקבל כתובת של פרוקסי מסויים ללא יכולת לדעת מה הכתובת IP המקורית של המשתמש...

האמנם??



זיהוי של HTTP Proxies



- הסקריפט משתמש ב**Flash** ושולח בקשה ישירות לשרת Fingerprinting, תוך מעקף של הגדרות הפרוקסי של הדפדפן\המכונה
- במקביל הסקריפט שולח בקשה רגילה דרך הJS
- בשני הבקשות מצורף מזהה ייחודי (Token)
- כך בעצם השרת יכול לזהות את המשתמש, ולשייך את ה-IP של הפרוקסי ל-IP האמיתי של המשתמש



Native Fingerprinting plugins

- מצאנו שחלק מהסקריפטים חיפשו פלאגין מסוים ב-Internet Explorer
- הפלאגין הזה יכול לגשת לנתוני המערכת של המשתמש, או דרך אישור של ActiveX, או כחלק מחבילות אחרות שהמשמש כבר התקין
- הורדנו את הפלאגין ומצאנו שהוא ניגש לנתונים ישירות מה-Registry
- הרבה יותר חזק מכל פלאגין מבוסס Flash או JavaScript
- בנוסף, אף אחד מהאנטי-ווירוסים של VirusTotal, לא זיהה את הפלאגין הזה כבעייתי



על התקשורת בין השרת המקורי לשרת הFingerprint

● אפשרות א

○ הFingerprint נשלח לשרת Fingerprint, ושום מידע ממנו לא נשלח לשרת המקורי

● אפשרות ב

○ הFingerprint נשלח מצופן לשרת המקורי, שבתורו שולח אותו לשרת הFingerprint, בשביל לקבל ממנו מידע על המשתמש

● אפשרות ג

○ השרת המקורי מוסיף את הSession ID לקובץ Html שנשלח למשתמש. המזהה הזה נשלח בכל פניה לשרת Fingerprint



שימוש של *fingerprint* ברחבי האינטרנט

- **בצענו מחקר עבור -**
 - **אתרים פופולאריים** - התמקדנו ב-10,000 האתרים המדורגים הכי הרבה ב-Alexa
 - **אתרים לא פופולאריים**
- **עבור כל אחד מהם:**
 - **חפשנו את הדומיינים** שיש בהם סקריפטים שפונים אל הדומיינים של אותם חברות מסחריות שחקרנו
 - **קבלנו רשימה של דומיינים וחלקנו אותם לקטגוריות**

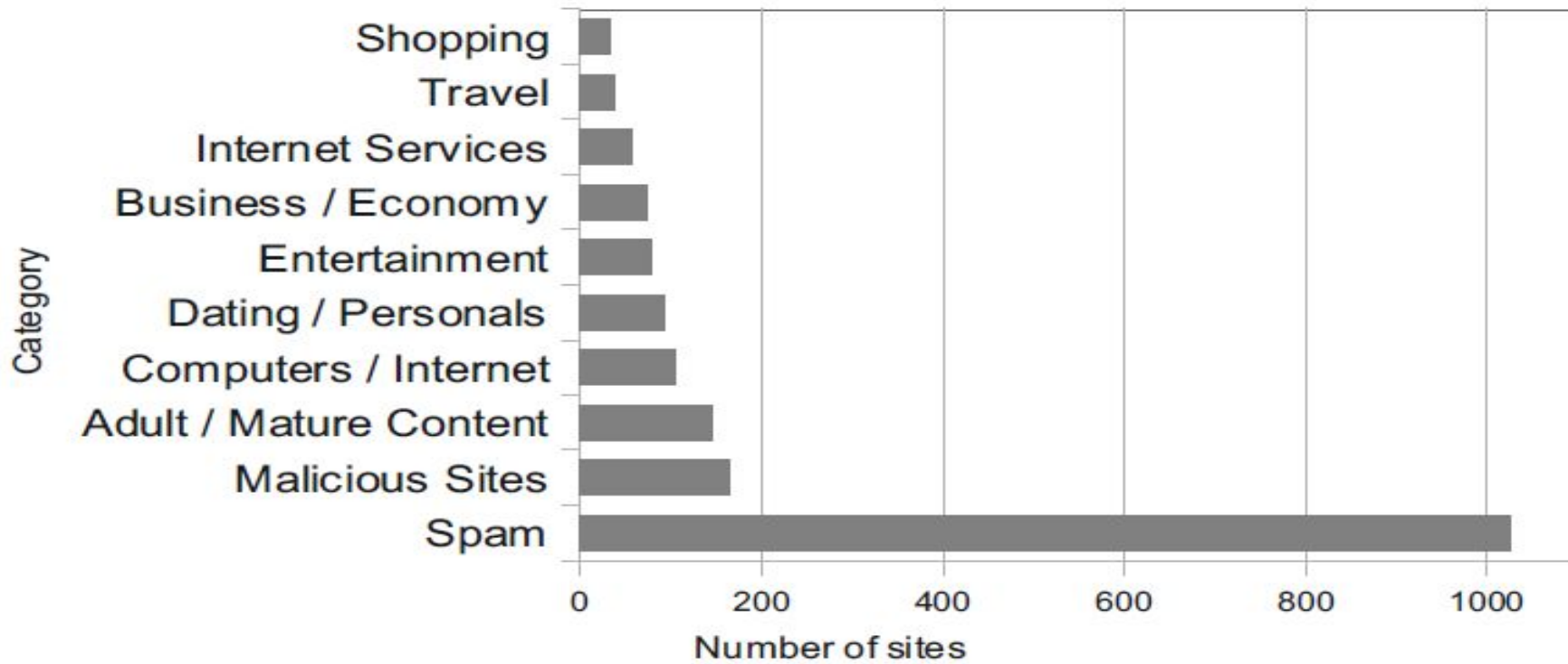


שימוש של *fingerprint* ברחבי האינטרנט

- 10,000 אתרים פופולאריים
- 40 אתרים (0.4%) משתמשים משתמשים ב*fingerprint* של אחת מהחברות המסחריות.
 - האתר הכי פופולרי שמשתמש בטביעות הוא **skype.com**.
 - הקטגוריות הכי נפוצות זה אתרי "פורנוגרפיה" (15%) ודייטים (12.5%).
- *fingerprints* הם כבר חלק מהאתרים הכי פופולריים באינטרנט וטביעות אצבע של מכשירים של אלפי משתמשים נלקחים מדי יום.
- שימו לב, שיכול להיות שאתרים פופולריים נוספים משתמשים בקוד שלהם כדי להשיג טביעות של משתמשים ולא בקוד של שלושת החברות שחקרנו.



שימוש של *fingerprint* ברחבי האינטרנט





שימוש של *fingerprint* ברחבי האינטרנט

אתרים לא פופולריים

- 8 מתוך 10 הקטגוריות כוללות אתרים שמשתמשים במנויים, רבים מתוכם כוללים מידע אישי ואולי פיננסי על המשתמשים - בדרך כלל משמשים כדי למנוע הונאות.
- בדומה לאתרים פופולריים - קטגוריית **Adult / Mature Content** מכילה שימוש רחב ב*fingerprinting*.
- 2 הקטגוריות המפתיעות שבראש הרשימה - אתרים זדוניים וספאם.

ככל הנראה, כנראה חברות המעקב עובדות עם אתרים מפוקפקים כדי להגדיל את מאגר המעקב שלהם ולרכוש עוד מידע על משתמשים.

אפשרויות חדשות למציאת Fingerprint





Fingerprint אפשרויות חדשות למציאת

- החלקים הקודמים, התבססו על אדיבות ליבו של הדפדפן בהספקת מידע לגבי המשתמש
○ אם ישירות בJavaScript או דרך דפדפן
- בחלק הזה נראה עד כמה רגיש הדפדפן, ובקלות ניתן לדלות מידע שיספק לצורך Fingerprinting של המשתמש
- השיטות הללו קשות למניעה, שכן הן מבוססות על המימושים הפנימים השונים של הדפדפן



Fingerprint אפשרויות חדשות למציאת

- אנחנו נתמקד בשני אובייקטי JS חשובים: screen ו-navigator
- בשונה מאובייקטים השונים הנוצרים על ידי המפתח של האתר, אלו מכילים מתודות ספציפיות של סביבת העבודה
- כתבנו סקריפט שמבצע פעולות בסיסיות מול האובייקטים המיוחדים הללו
- הסקריפט רץ על המאפיינים של האובייקטים, ניסה למחוק, לערוך, להוסיף, ובכל שלב, ההתנהגות של האובייקט נשמרה
- את הסקריפט הרצנו על כל סוגי הדפדפנים בגרסאות שונות (68 סוגים שונים)
- לאחר מכן, עשינו השוואה, ובדקנו האם ניתן לאפיין התנהגויות שונות ולזהות לפי זה את סוגי הדפדפן, גירסה ומערכת הפעלה



Fingerprint אפשרויות חדשות למציאת

סדר של המאפיינים באובייקט

- גילינו שהסדר של המאפיינים באובייקטים הנ"ל, שונה בין דפדפנים שונים, וכן בין גירסאות שונות של אותו דפדפן, ולעיתים אף בתלות של המערכת הפעלה
- בנוסף, בצורה הגיונית, גם **כמות** המאפיינים שונה בין הגרסאות, כך שככל שהגירסה חדשה יותר, יש יותר מאפיינים באובייקטים שנוספו
- בעצם, ניתן למפות את כל הפרמטרים הללו לטבלה מסודרת, ובעצם לגלות את גירסת הדפדפן המדויקת רק על ידי אבחון של אובייקטי הJS שלו



Fingerprint אפשרויות חדשות למציאת

navigator.geolocation
navigator.onLine
navigator.cookieEnabled
navigator.vendorSub
navigator.vendor

navigator.appCodeName
navigator.appName
navigator.appVersion
navigator.language
navigator.mimeTypes

↔ navigator.appCodeName
↔ navigator.appName
navigator.appMinorVersion
navigator.cpuClass
navigator.platform



What are we?!



Browsers!



Browsers!



Browsers!



What do
we want?!



More
speed!



More
speed!



More
speed!



And when do
we want it?!



Right
now!



Right
now!



Right
now!



Browsers!





Fingerprint אפשרויות חדשות למציאת

פיצ'רים ייחודיים

- בעבר באמצע שנות ה-90, במלחמת הדפדפנים, במאבק על נתח השוק, כל דפדפן הציג לראווה פיצ'רים ייחודיים שרק לו יש בשביל למשוך משתמשים
- למרות שכיום יש כבר סטנדרט של ארגון W3C של Html, עדיין נשארו וקיימים פיצ'רים שונים

פר דפדפן

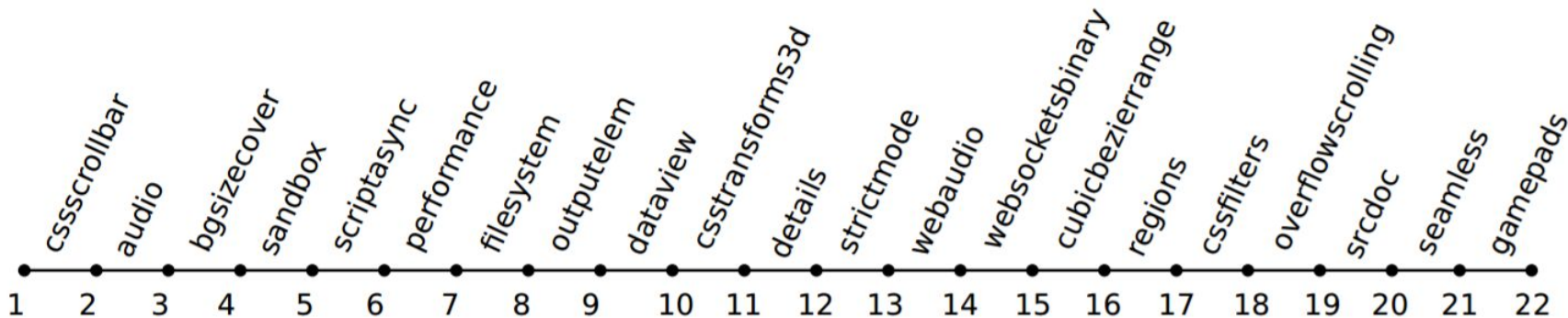
Browser	Unique methods & properties
Mozilla Firefox	screen.mozBrightness screen.mozEnabled navigator.mozSms + 10
Google Chrome	navigator.webkitStartActivity navigator.getStorageUpdates
Opera	navigator.browserLanguage navigator.getUserMedia
Microsoft IE	screen.logicalXDPI screen.fontSmoothingEnabled navigator.appMinorVersion +11



Fingerprint אפשרויות חדשות למציאת

התקדמות של הפונקציונליות של הדפדפן

- בכל גירסה חדשה של דפדפן, נוספות תכולות חדשות שמבדילות את הגירסה מהגירסה הקודמת
- בדקנו 202 גרסאות שונות של כרום החל מ-1.0.1 עד 22.0
- מצאנו שישנם 109 מאפיינים שניתן בעזרת לאבחן את הגירסה הספציפית של הדפדפן





תוספי דפדפן – מנסים לשמור על הפרטיות של המשתמש

- יש **11 תוספים שונים** שטוענים שהם מזייפים user-agent של הדפדפן.
 - 8 תוספים לfirefox ו 3 תוספים לchrome
 - יש להם יותר מ 800,000 משתמשים
- מי ממליץ להשתמש בתוספים כאלה
 - מחקרים קודמים בנושא מעקב באינטרנט
 - מדריכי hacking מחתרתיים
- איך הם מתמודדים כנגד **fingerprinting**?



יותר טוב מכלום? ... לא

● לכולם היה לפחות אחד מהדברים הבאים:

○ חוסר תאימות וכיסוי של אובייקט navigator

○ הגדרות לא הגיוניות

○ חוסר תאימות בין User-Agent header לבין User-Agent property.

● בעיה אירונית

○ כאשר מתקינים את התוספים, המשתמש הופך להיות נראה ויותר קל לבנות לו

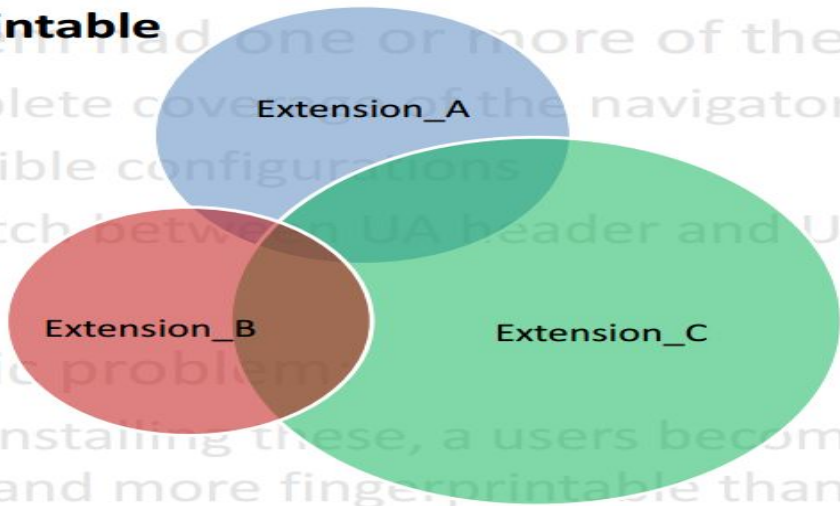
fingerprint ממקודם.



יותר טוב מכלום? ... לא

Fingerprintable Surface

- Incomplete coverage of the navigator
- Impossible configurations
- Mismatch between UA header and UA



- When installing these, a users become more visible and more fingerprintable than



סיכום

- מעקב בעזרת fingerprinting הרבה יותר משמעותי מcookies
- fingerprinting זו בעיה אמיתית
- הפתרון טמון בלגרום לדפדפנים בלהתנהג בצורה זוהה, הבעיה היא שעקב המורכבות שלהם זה בלתי אפשרי
- שימוש בתוספי "פרטיות" לא מונע בהכרח Fingerprinting, ויכול להיות שזה רק יותר בעייתי





Thanks!

Any questions?

You can find us at:
@bityob · @klavior



Credits And Resources

Credit

- ◆ Presentation template by [SlidesCarnival](#)
- ◆ Photographs by [Benedikt Geyer](#)
- ◆ [Wallpapers](#) - wallpapersafari

Resources

- [Browser Fingerprinting](#) - Presentation by Nick Nikiforakis
- [Cookieless Monster](#) - Source Article
- "Cookieless Monster"'s POC - [GitHub](#)
- [Am I Unique?](#) - Learn how identifiable you are on the Internet
- [Browserprint.info](#) - Fingerprint test
- [Fingerprintjs2](#) - Modern & flexible browser fingerprinting library (GitHub)