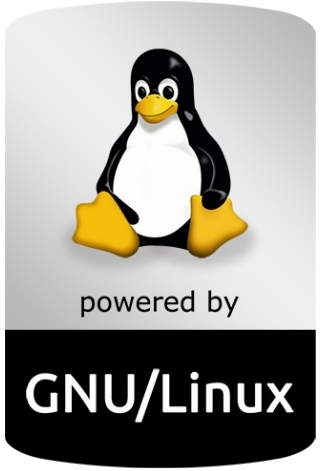
FINGERPRINTING

TECNICHE DI PROFILATURA DELL'UTENTE COME LIMITE ALLA LIBERTÀ INDIVIDUALE



Relatore:
Costantino Pastore

LINUX DAY MILANO 2018

FINGERPRINTING: UNA PRIMA AMPIA DEFINIZIONE

Il fingerprinting (o impronta digitale) é l'attività per mezzo della quale una serie complessa di dati relativi all'utente viene elaborata da specifici programmi /server per generare alla fine la segmentazione della utenza in gruppi omogenei di comportamento. I dati che possono essere presi in considerazione per la profilazione sono molteplici. Tra questi citiamo:

- la serie delle scelte di navigazione effettuate sul sito in esame dagli utenti unici identificati;
- le preferenze e interessi degli utenti
- la raccolta di dati demografici;
- la risposta degli utenti identificati a promozioni o a contenuti particolari.

I sistemi di profilazione più avanzati mettono a disposizione del settore commerciale di un'azienda la possibilità di segmentare in gruppi la propria utenza sia manualmente, scegliendo i parametri da prendere in considerazione, sia automaticamente, in base alle capacità native del software utilizzato. In entrambi i casi, il valore aggiunto è dato dalle molteplici correlazioni che è possibile istituire tra i dati raccolti, al fine di ricavarne informazioni commercialmente utili. Ecco a titolo esemplificativo alcune di queste correlazioni:

IL FINGERPRINTING: ATTIVITA' E FINALITA'

Successivamente vedremo quali sono le tecniche di profilazione che vengono utilizzate.
E' bene fin da subito sapere che il fingerprinting limita fortemente la libertà individuale non solo perché l'individuo viene in qualche modo costantemente monitorato, ma viene anche determinato in modo eterodiretto nelle sue scelte e quindi nella stessa manifestazione della sua volontà.
Non ultimo anche il fatto che questa attività viene **quasi sempre effettuata all'insaputa dell'utente** (**c.d. profilatura implicita**) sia con finalità in primo luogo legali (es. profilazione commerciale) sia con finalità illegali (intrusione nella vita della persona e/o hacking etico) o addirittura sovversive (es. controllo del diritto di voto).
Potremmo definire il *fingerprinting* come **qualsiasi attività o dato che un soggetto terzo sia in grado di raccogliere in rete a nostra insaputa al fine di identificarci, (il più possibile univocamente).**
Per chiarezza, la **profilatura esplicita**, invece, sussiste quando il soggetto é informato di essere oggetto di uno studio di mercato o di un sondaggio o comunque informato di attività relative alla sua riservatezza.
Tralasceremo volutamente il concetto di fingerprinting legato a normali attività informatiche che richiedono l'identità di un determinato software come le funzioni di *hashing*

LE FINALITA' LEGALI DEL FINGERPRINTING

Come detto il fingerprinting ossia l'identificazione univoca dell'individuo/utente sul web può avvenire anche con finalità *prima facie* legali, come garantire un più agevole accesso ad un certo sito o proporre contenuti relativi agli interessi precedentemente espressi dal soggetto (es. siti di e-commerce come Ebay o Amazon).
Lo scopo può anche essere quello di fornire notizie o approfondimenti contigui alla persona (es. Google news o Youtube) o fornire notizie su persone che si conoscono (es. Facebook) o che vorresti conoscere (es. Tinder).
A tal fine si utilizzano fondamentalmente oltre a userid e password anche e soprattutto tecniche di profilazione attraverso i c.d. cookies.
Queste finalità sono solo in prima istanza legali perché accedendo alla "cookie area" posso prendere visione dei siti che il soggetto legato a quell'IP ha visitato (e che conosco con nome e cognome per essersi loggato e aver lasciato le sue credenziali nel mio o in altro cookie) e sapere quali servizi visita (es. accesso ad ospedali) o quali interessi può avere (es. filosofia) e quindi ampliare le conoscenze sul medesimo per offrirgli contenuti sempre più mirati.

La finalità del fingerprinting legale é sia quella di favorire i commerci attraverso il marketing (come quando vado al supermercato), sia quella di "aiutare" il povero utente nella ricerca sul web, oltre a quella di favorirgli il mantenimento o la costruzione di relazioni interindividuali.

LA TRADIZIONALE TECNICA DI PROFILAZIONE: I COOKIES

Come anticipato, il fingerprinting si é lungamente basato sui c.d. cookies.

Il concetto e il termine cookie, che letteralmente significa "biscotto" , derivano dal magic cookie (biscotto magico) una tecnica nota in ambiente UNIX già negli anni '80 e tipicamente utilizzata per implementare meccanismi di identificazione di un client presso un server, come ad esempio l'autenticazione del server X Window System.

Il primo uso di cookie in HTTP risale al 1994. I cookie HTTP (più comunemente denominati cookie web, o per antonomasia cookie) sono un tipo particolare di magic cookie (una sorta di gettone identificativo). Vengono utilizzati dalle applicazioni web lato server che archiviano e recuperano informazioni a lungo termine sul lato client.

I server inviano i cookie nella risposta HTTP e ci si aspetta che i browser salvino e inviino i cookie al server, ogni qual volta si facciano richieste aggiuntive al server.

Nel dettaglio, un cookie è una stringa di testo (ma non necessariamente) contenuta in un file di piccole dimensioni inviata da un web server ad un web client (di solito un browser) e poi rimandata indietro dal client al server (senza subire modifiche) ogni volta che il client accede alla stessa porzione dello stesso dominio web. I cookie sono stati originariamente introdotti per fornire un modo agli utenti di memorizzare gli oggetti che volevano acquistare, mentre navigavano nel sito web (il cosiddetto "carrello della spesa").

Tale riconoscimento permette di realizzare meccanismi di autenticazione, usati ad esempio per i login; di memorizzare dati utili alla sessione di navigazione, come le preferenze sull'aspetto grafico o linguistico del sito; di associare dati memorizzati dal server, ad esempio il contenuto del carrello di un negozio elettronico; di tracciare la navigazione dell'utente, ad esempio per fini statistici o pubblicitari.

La sicurezza di un cookie di autenticazione dipende generalmente dalla sicurezza del sito che lo emette, dal browser web dell'utente, e dipende dal fatto che il cookie sia criptato o meno. Le vulnerabilità di sicurezza possono permettere agli hacker di leggere i dati del cookie, che potrebbe essere usato per ottenere l'accesso ai dati degli utenti, o per ottenere l'accesso (con le credenziali dell'utente) al sito web a cui il cookie appartiene

I cookie, e in particolare i cookie di terza parte, sono comunemente usati per memorizzare le ricerche di navigazione degli utenti; questi dati sensibili, possono essere una potenziale minaccia alla privacy degli utenti e sono stati oggetto di specifica regolamentazione normativa.

ALCUNE TIPOLOGIE DI COOKIES

- **Cookie di sessione:** questi cookie non vengono memorizzati in modo persistente sul dispositivo dell'utente e vengono cancellati alla chiusura del browser[6]. A differenza di altri cookie, i cookie di sessione non hanno una data di scadenza, ed in base a questo il browser riesce ad identificarli come tali.
- **Cookie persistenti:** invece di svanire alla chiusura del browser, come vale per i cookie di sessione, i cookie persistenti scadono ad una data specifica o dopo un determinato periodo di tempo. Ciò significa che, per l'intera durata di vita del cookie (che può essere lunga o breve a seconda della data di scadenza decisa dai suoi creatori), le sue informazioni verranno trasmesse al server ogni volta che l'utente visita il sito web, o ogni volta che l'utente visualizza una risorsa appartenente a tale sito da un altro sito (ad esempio un annuncio pubblicitario). Per questo motivo, i cookie persistenti possono essere utilizzati dagli inserzionisti per registrare le informazioni sulle abitudini di navigazione web di un utente per un periodo prolungato di tempo. Tuttavia, essi sono utilizzati anche per motivi "legittimi " (come ad esempio mantenere gli utenti registrati nel loro account sui siti web, al fine di evitare, ad ogni visita, l'inserimento delle credenziali per l'accesso ai siti web).
- **Zombie cookie:** Gli Zombie cookie sono cookie che vengono ricreati automaticamente dopo essere stati eliminati. Questo si ottiene attraverso la memorizzazione dei contenuti del cookie in più posizioni, come la flash local storage, HTML5 storage, e attraverso altri meccanismi di archiviazione sia da parte del client che da parte del server. Quando viene rilevata l'assenza del cookie, quest'ultimo viene ricreato utilizzando i dati memorizzati in queste posizioni.

TIPOLOGIE DI COOKIES STABILITI DALLA LEGGE

Nella Cookie Law, il Garante ha individuato 3 tipi di cookie: i "cookie tecnici", i "cookie di profilazione" e i "cookie di terze parti" e per ognuno di essi ha prescritto regole di informativa e di uso differenti che gli editori di siti web devono rispettare.
I cookie tecnici sono quelli utilizzati al solo fine di "effettuare la trasmissione di una comunicazione su una rete di comunicazione elettronica, o nella misura strettamente necessaria al fornitore di un servizio della società dell'informazione esplicitamente richiesto dall'abbonato o dall'utente a erogare tale servizio" (cfr. art. 122, comma 1, del Codice).

Essi non vengono utilizzati per scopi ulteriori e sono normalmente installati direttamente dal titolare o gestore del sito web.

I cookie di profilazione sono volti a creare profili relativi all'utente e vengono utilizzati al fine di inviare messaggi pubblicitari in linea con le preferenze manifestate dallo stesso nell'ambito della navigazione in rete.

Anche in questo caso, le difficoltà dovrebbero essere minime: la distinzione tra "cookie di terze parti" e gli altri due tipi di cookie si fonda sul "chi" installa tali cookie: se il cookie lo installa il tuo sito direttamente, allora è un "cookie tecnico" o un "cookie di profilazione"; se, invece, il cookie lo installa un servizio terzo usando il tuo sito, allora è un "cookie di terza parte".

LA REGOLAMENTAZIONE DEI COOKIES

Quella che volgarmente chiamiamo Cookie Law, e che prende il nome di Direttiva ePrivacy, il cui nome completo è Direttiva 2002/58/CE del Parlamento europeo e del Consiglio, del 12 luglio 2002, relativa al trattamento dei dati personali e alla tutela della vita privata nel settore delle comunicazioni elettroniche (direttiva relativa alla vita privata e alle comunicazioni elettroniche) non fa altro che dare indicazioni e linee guida su come trattare i dati personali tramite mezzi elettronici, quindi newsletter e tutto ciò che è email marketing, cookie nei siti web, ma anche spyware e simili. **Essa, quindi, non viene abrogata, come si può pensare, dal GDPR, ma si affianca ad esso, abbiamo così GDPR e Cookie che sono complementari per aumentare il livello di privacy delle persone fisiche.**
La Cookie Policy non è altro che una pagina del sito web in cui viene descritto in dettaglio quali sono le finalità di installazione ed uso dei cookie. In essa vengono indicati quali cookie siano necessari al funzionamento del sito web, i cosiddetti cookie tecnici, quali sono i cookie utilizzati per azioni di marketing, i cosidetti cookie di profilazione, ed infine quali sono i cookie installati da altri siti, i cosiddetti cookie di terze parti. Inoltre, devono essere ben definiti sia quali siano le terze parti che installano, o almeno potrebbero installare cookie tramite il sito, sia quale è la privacy policy adottata da questi servizi ed eventualmente come procedere alla revoca del consenso. Molto importante, ed è l'unica cosa che al momento veramente si trae dal GDPR (Reg. 2016/679UE), è che la revoca del consenso deve essere facile allo stesso modo di come il consenso possa essere dato, cioè se con un click si può dare il consenso con uno si dovrebbe poter revocare. Il "si dovrebbe" è un obbligo in quanto non sempre è possibile, ad esempio per i cookie di terze parti, ed in quel caso si rimanda ai riferimenti delle informative di terzi.

Da tutto ciò si deduce cosa serve per far andare d'accordo GDPR e Cookie:

- blocco preventivo dei cookie, fino ad accettazione della cookie policy;

- consenso informato, esplicito **e preventivo** (prima dell'installazione dei cookie), il tutto da ricevere chiaramente con una azione di opt-in, cioè di accettazione positiva;

- registrazione del suddetto consenso che sia in qualche modo provabile alle autorità competenti;

- **possibilità di revoca del consenso ed indicazione su come effettuare tale operazione**;

- informativa chiara e dettagliata sull'utilizzo dei cookie e sugli installatori di terze parti;

- collegamento alla pagina di privacy policy del sito web.

ALTRE TECNICHE DI PROFILAZIONE “IN AND OUT OF WEB”: I PLUGIN DEI CMS, METADATI, TRIANGOLAZIONI CELLULARI, GPS.

Possiamo essere “controllati” non solo dai cookie di grandi aziende che visitiamo sul web, ma anche da qualsiasi piccolo sito che visitiamo o che non visitiamo (c.d. cookies di terze parti).
Generalmente ciò avviene, non solo attraverso l’inserimento di codice proprio, ma anche tramite alcuni piccoli programmi aggiuntivi (c.d. plugin) installabili su qualsiasi piattaforma di publishing CMS (wordpress, joomla, drupal, ecc..)
Esistono anche alcuni plugin di Google Ads (pubblicità mirata) e Google Analytics (monitoraggio online dei visitatori, a fine statistico).

Ogni volta che pubblichiamo una foto sul web o archiviamo una foto dobbiamo essere consapevoli che essa potrebbe suggerire inconsapevolmente anche dati personali come il luogo dove la foto é stata scattata, le caratteristiche del cellulare o della macchina fotografica utilizzati, la data e ora dello scatto (c.d. metadati).

Ovviamente se utilizziamo il nostro cellulare con il GPS attivo sarà molto facile profilarci e sapere dove e perché ci troviamo in un determinato luogo. A volte é possibile sapere dove si trovava un soggetto semplicemente attraverso i dati di triangolazione GSM.

IL CASO DI GOOGLE

Anche nel caso in cui non concedeste o revocaste il consenso all'accesso alla cronologia delle posizioni sembra che Big G tenga comunque traccia della vostra posizione.

Inoltre, sembrerebbe che chi abbia installato Google Chrome su device Android, detto browser in background invii la posizione del dispositivo al server ogni 4 minuti.

Google giustifica il suo operato affermando che il tutto é rivolto a garantire una migliore efficacia dei propri servizi.

Fonte: PC Professionale – Editoriale – Settembre 2018

BIG BROTHER, BIG DATA E DATA MINING

“Se non paghi un servizio la merce sei tu”

E’ vero che in cambio della gratuità di molti servizi, siamo disposti a concedere l’uso dei nostri dati personali ed é anche vero che molto spesso l’uso dei nostri dati é ben specificato nella privacy policy. Tuttavia non possiamo sapere quale altra mole di dati venga raccolta a nostra insaputa da aziende, governi o criminali.

Tutti questi dati sono talmente tanti che vengono definiti “Big data” sui quali viene operato una attenta cernita attraverso il c.d. “data mining” ovvero la raccolta, la preordinazione e l’analisi dei dati raccolti.

I risultati a cui può giungere l’attività di data mining sono molteplici:

1) associazioni – due eventi si verificano spesso insieme (ad esempio chi compra delle scarpe tende ad acquistare anche dei calzini);
2) sequenze – due eventi successivi sembrano legati da una relazione di causa-effetto (chi compra un mouse su Internet, tempo dopo acquista anche un tappetino);
3) classificazioni – il riconoscimento di un ordine in una serie di eventi, con la conseguente riorganizzazione dei dati in proprio possesso;
4) raggruppamenti – la ricerca e la presentazione di gruppi di fatti non precedentemente noti;
previsioni – lo studio della probabile evoluzione futura della propria attività in base alle risultanze dei dati raccolti.

SCANDALO CAMBRIDGE ANALYTICA E FACEBOOK MARZO 2018

Facebook dichiara che i profili social coinvolti sono 87 milioni
In Italia gli utenti coinvolti sono 214.134, a fronte di 31 milioni di account registrati.
Facebook raccoglieva anche i dati di chiamate e sms
In Italia la violazione dei dati è partita da 57 utenti che avevano scaricato la app Thisisyourdigitallife, creata da Alexander Kogan, il ricercatore dell'Università di Cambridge che ha scaricato i dati da Facebook per passarli poi a Cambridge Analytica.

IL FINGERPRINTING OGGI: IL DEVICE FINGERPRINT

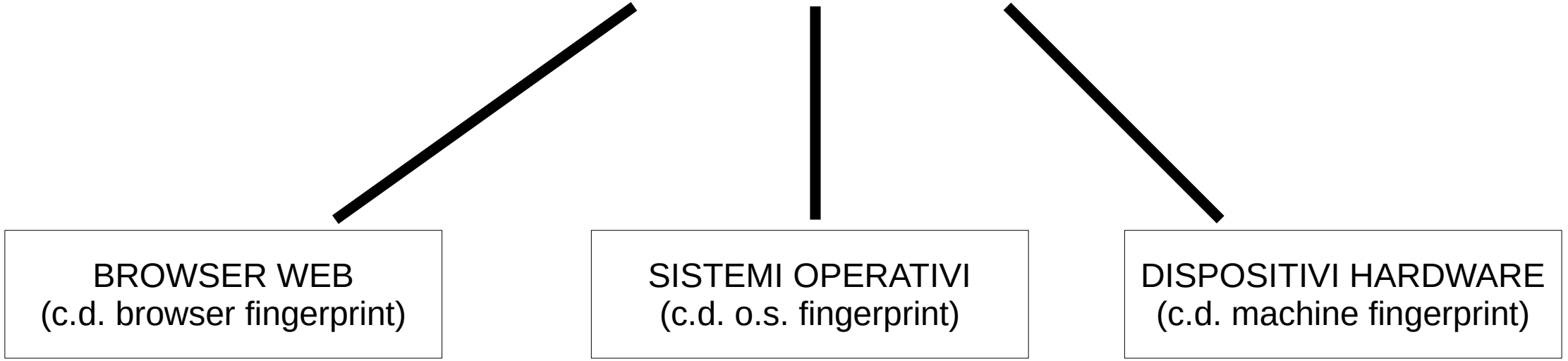
Il **device fingerprint** (letteralmente "impronta digitale del dispositivo") in informatica è l'informazione raccolta su di un dispositivo di elaborazione remoto a scopo di identificazione.

La device fingerprint permette di identificare in tutto o in parte i singoli utenti o dispositivi **anche quando i cookie sono disattivati**, di fatto superando i limiti imposti dalla normativa sui cookie e sulla privacy vigente.

Più in dettaglio, mediante il device fingerprint l'utente può essere monitorato attraverso la tracciatura e raccolta di dati tecnici e proprietà del suo dispositivo desktop o mobile connesso a Internet. Le informazioni ottenibili spaziano dalla dimensione dello schermo, alle versioni di software e plugin installati, alla lista dei caratteri installati. Oppure ancora è possibile ottenere indicazioni riguardo alle configurazioni TCP/IP, alle informazioni tratte dall'orologio del dispositivo, al sistema operativo utilizzato, alle impostazioni delle connessioni wireless, fino ad arrivare all'indirizzo IP originale dell'utente.

Lo scopo é osservare le abitudini di navigazione dell'utente sia a fini commerciali quanto a fini di possibile controllo sociale.

DEVICE FINGERPRINT



TIPOLOGIE DI DEVICE FINGERPRINT

Esistono due differenti modalità con cui tale tecnica può essere utilizzata:

Fingerprint passivo: si tratta di tecniche che implicano un'**analisi passiva del traffico di rete**, al fine di identificare il sistema operativo di un computer o la configurazione TCP/IP del client. Il traffico di rete infatti, può fornire oltre ai dati effettivamente scambiati, una serie di informazioni e parametri del sistema dell'utente. Pertanto l'utilizzo del fingerprint passivo può risultare utile anche per compiti di ricognizione attraverso l'utilizzo di tecniche quali lo sniffer. Gli scanner passivi generalmente risultano meno precisi rispetto a quelli attivi in quanto hanno meno controllo sui dati che stanno analizzando, tuttavia permettono di esaminare in modo del tutto anonimo tutte le varie informazioni.

Fingerprint attivo: presuppongono che l'**agente operi attivamente agendo sul dispositivo del soggetto di cui si vogliano raccogliere informazioni**. Il metodo più attivo è l'installazione di codice eseguibile direttamente sul computer client. Tale codice può avere accesso ad attributi non tipicamente disponibili con altri mezzi, come l'indirizzo MAC, o altri numeri di serie univoci assegnati all'hardware della macchina. Questa tecnica risulta essere utile ed efficace, se implementata nell'utilizzo di tecniche quali il port scanning; tuttavia rispetto a quella passiva, presenta un alto rischio nell'essere scoperti.

DEVICE FINGERPRINT PASSIVO

Si intende la raccolta delle informazioni di rete che vengono acquisite senza l'uso di un'applicazione in particolare. In questo caso si tratta di informazioni che sono contenute come da standard nei dati dell'intestazione dei pacchetti IP e raggiungono così il web server. Rientrano tra queste informazioni anche l'indirizzo IP, la porta utilizzata e il tipo di browser. Ma si annoverano anche le configurazioni essenziali, come i tipi di dati desiderati (HTML, XHTML, XML), i set di caratteri (ad esempio UTF-8) e le lingue (di solito la lingua del browser o del sistema operativo). Inoltre l'header HTTP fornisce in alcuni casi anche informazioni sul sistema operativo utilizzato e la pagina di provenienza.

INFORMAZIONI RILEVABILI

Nome del browser e versione

Versione S.O.

Cookie attivati/disattivati

Lingua locale del browser o del s.o.

Piattaforma Hardware / dispositivi mobili

Dati dettagliati del Browser

Informazioni delle schermo / e dell‘ orario macchina

I dati dell'header HTTP forniscono le informazioni rilevanti

L'header del protocollo HTTP che, come già menzionato, viene utilizzato per la trasmissione dei contenuti web, non ha una dimensione fissa, al contrario dell'header della porta TCP e dell'indirizzo IP. Oltre alla capacità di comprendere input personalizzati, vengono prescritti diversi campi standardizzati, alcuni dei quali sono di vitale importanza per la creazione del browser fingerprint. Si tratta nello specifico dei seguenti dati dell'header:

"Referer" (pagina di provenienza): se un utente giunge tramite link da una pagina A a una B, l'URL della pagina A viene trasmesso come referer al server dalla pagina B. A volte può succedere che un utente giunga alla pagina di destinazione sempre da una determinata pagina. Allo scopo della creazione del fingerprint, il referer risulta quindi tanto utile quanto il parametro GET contenuto nell'URL.

"User-Agent" (descrizione del client): per ogni richiesta HTTP il relativo client fornisce solitamente anche una descrizione di sé nel campo "User-Agent". Lì, oltre alla denominazione e al numero di versione, l'header HTTP offre anche spazio per un commento, che molti browser utilizzano per eseguire la piattaforma o il sistema operativo sottostanti.

“Accept“ (formati di output consentiti): tramite il campo “Accept” il browser comunica al server quali tipi di contenuti possa elaborare e quindi quali possibili formati di output desideri. Oltre all’HTML sono richiesti soprattutto l’XHTML (Extensible Hypertext Markup Language) e l’XML (Extensible Markup Language). Se il campo è mancante, il client supporta tutti i tipi di contenuti.

“Accept-Charset“ (set di caratteri ammessi): in aggiunta al formato di output il client può anche definire il set di caratteri desiderati, che il server deve utilizzare per la sua risposta. In questo caso si tratta soprattutto del set UTF-8 e dello standard ISO ISO/IEC 8859-1.

“Accept-Encoding“ (formati di compressione accettati): per ottimizzare il tempo di caricamento di un sito web è uso comune comprimerne i contenuti. Il browser deve di conseguenza estrarre i dati compressi prima di poterli riprodurre. Nel campo “Accept-Encoding” comunica al server contattato quali formati di compressione supporta. La lista dei possibili procedimenti, gestita da IANA, comprende anche gzip, deflate, exi e br.

“Accept-Language“ (lingue accettate): tramite la voce HTTP “Accept-Language” i client comunicano le preferenze relative alla lingua di visualizzazione. Quando queste sono presenti per il sito richiamato, il web server le fornisce. La scelta automatica della lingua, cosiddetta “forzata“, si basa sulle impostazioni del browser o del sistema operativo. Le impostazioni di alcuni browser offrono inoltre la possibilità di indicare ulteriori preferenze di lingua.

Ad esempio, le proprietà del browser dell'utente che vengono richieste sono pressoché comuni al device fingerprint attivo. Il tracking avviene tramite l'oggetto navigator, il quale è una possibile proprietà degli oggetti window, cioè della finestra che si apre nel browser. Anche se per l'oggetto navigator non è definito uno standard generale, viene tuttavia supportato da tutti i browser comuni e si occupa di inoltrare anche le seguenti informazioni al web server.

IL DEVICE FINGERPRINT ATTIVO

Il fingerprint attivo prevede, come già intuibile dal nome, che il gestore di un progetto web richieda attivamente informazioni sul client. Le proprietà e i dati richiesti non risultano quindi dai dati dell'header forniti dai pacchetti del client. Visto che devono essere eseguiti a questo scopo applicazioni sul browser, l'utente può teoricamente verificare in qualsiasi momento se sia o meno in corso un processo di fingerprinting, semplicemente analizzando i pacchetti in uscita o il codice sorgente HTML o JavaScript. Nella maggior parte dei casi il processo viene però tenuto nascosto ai visitatori, come succede anche nei procedimenti di tracking.

Molto più spesso il fingerprint device attivo si può declinare, nella sua variante illecita, anche con l'installazione sul device vittima di malware (rootkit, spyware, keylogger, trojan) o come già accennato con attività come port scanning, accesso remoto, attacchi mitm, dns cache poisoning, cross-site scripting.

IL FINGERPRINT ATTIVO CON JAVASCRIPT

Per uno scambio di dati facile e veloce tra client e server di solito si implementa il browser fingerprint attivo tramite elementi AJAX (Asynchronous JavaScript and XML). Grazie a questa tecnologia i visitatori possono interagire con una pagina, senza che la stessa debba essere ricaricata completamente a ogni richiesta HTTP. Per questo motivo vengono caricate in background solamente le risorse richieste, mentre l'utente può continuare a vedere e a utilizzare tutti gli altri elementi. Le informazioni, di cui si viene a conoscenza tramite i relativi script, si possono suddividere in due categorie: informazioni del browser e dello schermo. Inoltre fingerprinting può essere ampliato grazie a JavaScript anche di indicazioni circa il fuso orario e i colori

IL FINGERPRINT ATTIVO CON CANVAS: IL SUCCESSORE DEI COOKIES

Per Canvas si intende una estensione dell'HTML standard che permette il rendering dinamico di immagini bitmap gestibili attraverso un linguaggio di scripting.

Con l'aiuto degli elementi canvas introdotti nell'HTML5, può essere generata senza problemi un'impronta digitale individuale sulla base della configurazione di sistema dell'utente web

Negli elementi canvas citati ci sono campi ben definibili (altezza e larghezza), che possono essere specificati con JavaScript, per creare ad esempio grafici, loghi e pulsanti, inclusivi di testo. Tuttavia, usando il canvas fingerprinting, la presentazione del testo avviene in maniera diversa a seconda:

- del sistema operativo
- del browser
- della scheda grafica
- dei driver della scheda grafica
- dei font del client installati.

l gestore del sito web ha bisogno per questo scopo solo di un codice canvas fingerprint specifico, che induce il browser nell'apertura della pagina a visualizzare in background un testo nascosto via JavaScript e, per mezzo di ciò, inoltrare le informazioni ottenute al server web. Grazie al grande numero di fattori presi in considerazione, l'impronta digitale creata in questo modo è unica in oltre l'80% dei casi, per cui può essere riconosciuta in qualunque momento, a patto che l'utente non apporti alcuna modifica a uno dei sistemi di configurazione sopra elencati.

l gestore del sito web ha bisogno per questo scopo solo di un codice canvas fingerprinting specifico, che induce il browser nell'apertura della pagina a visualizzare in background un testo nascosto via JavaScript e, per mezzo di ciò, inoltrare le informazioni ottenute al server web. Grazie al grande numero di fattori presi in considerazione, l'impronta digitale creata in questo modo è unica in oltre l'80% dei casi, per cui può essere riconosciuta in qualunque momento, a patto che l'utente non apporti alcuna modifica a uno dei sistemi di configurazione sopra elencati.

IL VALORE DEL CANVAS FINGERPRINT PER L'ANALISI WEB

Il canvas fingerprinting ottiene di fatto solo le informazioni già citate dal sistema e dal browser. Ma queste offrono già la possibilità di identificare il visitatore di un sito web come singolo individuo, per determinare infine il suo comportamento di navigazione. Può trattarsi in contemporanea delle attività di uno o più siti web, se lo script è implementato su diverse pagine. Di conseguenza il metodo è molto interessante sia nell'ottimizzazione del sito web che anche nell'ideazione della pubblicità mirata ad uno specifico target. Uno dei vantaggi maggiori del moderno metodo di tracciamento degli utenti è che non raccoglie alcun dato personale. Per gli analisti web questo rende il canvas fingerprinting un'alternativa da prendere seriamente in considerazione rispetto ai cookie, che creano sempre problemi a livello legale e vengono bloccati ed eliminati da molti utenti.

Dal momento che l’impronta digitale non è unica al 100% al contrario di quella reale, i risultati del canvas fingerprint non sono sempre completamente affidabili. Così, due visitatori di un sito web con, ad esempio, le stesse configurazioni ottengono solo un ID utente, aspetto che altera le ulteriori ricerche. Poiché la probabilità di una corrispondenza di questo tipo cresce a seconda del numero degli utenti tracciati, questa problematica aumenta insieme al traffico del sito web analizzato. Un ulteriore problema del fingerprint canvas emerge durante l’identificazione degli utenti che utilizzano i dispositivi mobili: gli hardware e i software dei tablet e degli smarphone sono di regola standard, ragion per cui sussistono pochi segni distintivi per creare un’impronta digitale unica.

Quando nel 2014 apparve la lista dei siti web che facevano uso del canvas fingerprinting, molti gestori di siti web si stupirono di trovarsi tra gli oltre 5000 casi, perché non avevano implementato loro stessi il processo di tracking ma si erano affidati, ad esempio, ad agenzie di performance marketing, come la tedesca Ligatus. Ma, stando a quanto sostenuto dall'agenzia, si trattava solo di una fase di test temporanea, nella quale furono raccolte esclusivamente informazioni anonime senza possibilità di risalire all'utente reale. Tuttavia, la maggior parte dei siti web usa, allo stesso modo senza saperlo, il code tracking dell'azienda statunitense AddThis, che è conosciuta soprattutto attraverso i pulsanti social media sui siti web.

COME E' POSSIBILE EVITARE IL CANVAS FINGERPRINT

Al contrario dei cookie, il canvas fingerprinting non può essere eliminato facilmente, perché i dati vengono direttamente trasmessi al server, quindi non avviene una memorizzazione lato client. Allo stesso modo, anche la modalità in incognito di un browser non impedisce agli script canvas di spiare le informazioni del sistema e del browser. Ma gli utenti non sono completamente indifesi contro questo metodo di tracking e possono impedire preventivamente l'esecuzione di script simili, ad esempio mettendo in atto le seguenti misure:

Disattivare JavaScript: senza JavaScript gli elementi canvas non possono essere caricati e quindi non si può richiamare nemmeno alcuna informazione tramite il client. Ma, dal momento che molti siti web contengono JavaScript, può capitare che questi non vengano più visualizzati in maniera corretta dopo la disattivazione di JavaScript.

Con **CanvasBlocker** gli utenti di Firefox bloccare il canvas fingerprinting. Ad esempio, è possibile ignorare tutte le richieste canvas o manipolare i dati recenti in modo che il fingerprinting ottenga sempre un altro valore.

LA FUNZIONALITA' DO NOT TRACK NON BASTA

Do not track è una funzionalità disponibile su tutti i principali browser e permette di bloccare il tracciamento della nostra navigazione online. Ma che cosa è veramente Do not track? Si tratta di un header HTTP, cioè di una direttiva di controllo delle pagine web che comunica al web server le preferenze degli utenti sulla raccolta dei dati della loro abitudini online. Lo scopo di Do not track è quello di sospendere la raccolta delle informazioni da parte delle aziende sulle abitudini online degli utenti. L'header accetta solamente due tipi di valore:

Valore 1: l'utente non vuole essere tracciato
Valore 0: l'utente vuole essere tracciato o non ha impostato la funzionalità.
Do not track non si attiva a meno che non sia l'utente a deciderlo. Quindi, se si vuole bloccare l'attività dei tracker online, sarà necessario entrare nelle impostazioni del proprio browser e attivare la funzione.

La funzionalità non riesce a bloccare tutti i sistemi utilizzati dalle aziende per tracciare le abitudini online degli utenti e quindi è probabile che qualche informazione arrivi comunque alle compagnie pubblicitarie. La funzionalità non è supportata da tutti i server e soprattutto non esistono delle norme giuridiche che impongano ai pubblicitari di rispettare la scelta espressa dagli utenti

COME EVITARE IL FINGERPRINTING

- Installare CyDec Platform Antifingerprint
- Evitare i like su Facebook
- Evitare di loggarsi su Google
- Evitare di usare Google Chrome
- Evitare di tenere attivo il GPS
- Evitare di inviare foto con metadati su Instagram
- Evitare di salvare le password
- Usare comunque do not track e bloccare pop-up
- Evitare reti wifi libere
- Non cliccare su pagine o canvas sospetti
- Usare TOR o cambiare IP