Sono finiti i biscotti

Device fingerprinting e tracciamento dell'utente senza l'utilizzo dei cookie ai tempi del GDPR

Autore: Edoardo Prandin (Università degli studi di Milano Bicocca)
e.prandin@campus.unimib.it

Sommario

I Cookies 1
Implicazioni in materia di protezione dei dati personali 4
GDPR e device fingerprinting 5
Rischi e misure 5
Conclusioni 6

Abstract

La relazione analizza i profili giuridici dell'utilizzo del device fingerprinting al fine di profilare e tracciare gli utenti, come tecnica alternativa all'utilizzo dei cookie, con riferimento all'attuale quadro normativo in materia di protezione dei dati personali, costituito dal nuovo regolamento generale in materia di protezione dei dati personali (Regolamento (EU) 679/206 cd. GDPR), dalla direttiva e-privacy (Direttiva 2002/58/CE) e dal progetto di regolamento che andrà a sostituirla, con considerazione delle pronunce del gruppo di lavoro articolo 29.

I Cookies

Il nuovo regolamento in materia di protezione dei dati personali[1] e i testi normativi ad esso antecedenti fanno spesso espressa menzione dell'uso dello strumento dei *cookies*. Questa tecnica di tracciamento degli utenti permette, con il salvataggio di semplici stringhe di testo sui dispositivi, di "riconoscerli" durante le diverse sessioni di navigazione. Ciò avviene per diverse finalità, quella più diffusa ad oggi è sicuramente quella di marketing. L'invio di messaggi pubblicitari mirati garantisce un'accuratezza ed efficacia maggiori rispetto alle pubblicità rivolte ad un pubblico indeterminato.

Con l'uso dei cookies è possibile se l'utente ha ad esempio visitato siti web che trattano orologi, proporre pubblicità mirate di vendita di siti di vendita o di produttori di orologi. I cookies non sono utilizzati unicamente per finalità di marketing ma anzi nascono per una finalità del tutto diversa, per adattare l'esperienza di navigazione alle esigenze dell'utente. Prima dell'invenzione dei cookies da parte di Lou

[1] Regolamento (UE) 2016/679 del Parlamento europeo e del Consiglio, del 27 aprile 2016, relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati e che abroga la direttiva 95/46/CE (regolamento generale sulla protezione dei dati). Conosciuto anche come GDPR (*General Data Protection Regulation*).

Montulli, all'epoca dipendente di *Netscape*[2], ogni visita ad un sito web era come se fosse la prima[3]. Montulli ebbe l'idea di risolvere questo problema salvando appunto sul dispositivo dell'utente una stringa di testo, costituita da caratteri univoci che ne permettesse il riconoscimento, così da poter identificare lo stesso utente ad una visita successiva rileggendo la stringa di testo univoca salvata sul suo dispositivo. Grazie ai cookies è possibile quindi stabilire se un utente ha visitato in precedenza un sito web. Se i siti web non facessero uso dei cookie sarebbe ad esempio necessario reinserire manualmente ad ogni pagamento su siti di e-commerce come Amazon i dati di pagamento e di fatturazione. Allo stesso modo sarebbe necessario reinserire i dati di login ad una piattaforma come ad esempio Facebook ad ogni riapertura di una nuova sessione del browser, senza avere la possibilità di "restare connessi" ed essere riconosciuti ad una visita successiva. L'invenzione dei cookies ha permesso così di trasformare la navigazione nel *world wide web,* e in generale l'uso degli applicativi che usufruiscono della rete, in un'esperienza ricca e dinamica, esperienza che prima era in realtà discontinua e statica. Da allora i cookies hanno subito una rapida e intensa evoluzione, portando alla nascita anche dei cosiddetti *extremely persistent cookies,* i cookies estremamente persistenti o resilienti, cioè difficilmente cancellabili, come ad esempio *evercookie*[4]*.* È seguito così un rapido sviluppo dei servizi fruibili attraverso la rete che basano il loro funzionamento proprio sull'utilizzo dei cookie.

I cookie, consistendo in file di testo salvati sul dispositivo dell'utente, possono essere facilmente cancellati manualmente dall'utente o direttamente dal browser web o dall'applicativo utilizzato. Il discorso si complica nel caso i siti web facciano uso di cookie estremamente persistenti, di fatto anche questi seppur con un po' di difficoltà si possono cancellare. Grazie alle recenti normative a livello europeo, in primis grazie alla cd. *Direttiva europea e-privacy*[5]*,* oggi quando visitiamo un sito web siamo abituati ad incontrare un'informativa o un banner che illustra il funzionamento dei cookie utilizzati dal sito web visitato. Abbiamo inoltre la possibilità di effettuare una scelta rispetto a quali cookie accettare che siano "lasciati" o meno sul nostro dispositivo in virtù del fatto che questa libertà di scelta è imposta dalla stessa direttiva. Per assurdo, se non ci fossero i cookies stessi, ad ogni visita di un sito web anche già visitato, sarebbe necessario visualizzare ed accettare ogni volta l'informativa stessa.

Per fare un esperimento è possibile ad esempio visitare in modalità incognito[6] il sito web www.google.com; all'apertura della pagina principale Google sottoporrà all'utente l'informativa sul trattamento dei dati personali e sull'utilizzo dei cookies, una volta chiusa la sessione, se il browser viene riavviato e, sempre in modalità incognito, si visita nuovamente il sito www.google.com, verrà riproposta comunque l'informativa anche se già visualizzata. Ciò accade proprio perché questo tipo di modalità di navigazione impedisce al browser di memorizzare cookie sul dispositivo, o se memorizzati li cancella al termine della sessione. Google non è quindi in grado di riconoscere l'utente alla successiva sessione e sottopone nuovamente l'informativa non conoscendo chi si trova di fronte.[7]

---

[2] Netscape fu uno dei primi web browser grafici, ebbe grande successo negli anni novanta salvo poi fallire a seguito di alcune sfortunate scelte di mercato che portarono al fallimento della società che ne curava lo sviluppo.

[3] Per un approfondimento: Englehardt, S., & Narayanan, A. (2016). Online Tracking. *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security - CCS16*. doi:10.1145/2976749.2978313

[4] Per un approfondimento su evercookie: https://samy.pl/evercookie/

[5] 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)

[6] Questa modalità può anche avere un nome diverso come ad esempio "anti-tracciamento" a seconda del browser utilizzato, si tratta di una modalità di navigazione che permette di evitare di memorizzare dati di navigazione sul dispositivo al termine dell'attività.

[7] L'utilizzo di modalità di navigazione in incognito implementate nel browser non rende anonima la navigazione dell'utente, impediscono il salvataggio dei dati di navigazione in locale sul dispositivo ma non nascondono, al *service provider* e a chi è in grado di monitorare l'attività di rete, i siti web navigati. Per nascondere gli indirizzi visitati è

Se quindi l'utente rifiuta il salvataggio di cookie che ne traccino l'attività di navigazione, il fornitore del servizio o gestore del sito web visitato non dovrebbe essere in grado di riconoscere l'utente ad una seconda visita, né di tracciarlo su di un altro sito o piattaforma. Quest'affermazione in realtà è in parte falsa, perché, per come oggi è strutturato il *world wide web* esistono metodi alternativi per tracciare l'utente anche senza l'utilizzo dei cookies. Questi metodi sono definiti *stateless,* in contrapposizione a quelli definiti *stateful*. I sistemi di tracciamento di tipo statless si basano sul fingerprinting dei dispositivi utilizzati per la navigazione e non necessitano in assoluto dell'utilizzo di cookies. Il fingerprinting consiste nella rilevazione di alcune caratteristiche dei dispositivi o degli applicativi utilizzati dall'utente al fine di raccogliere informazioni che ne permettano l'identificazione univoca, anche multipiattaforma e multisessione. Di fatto queste tecniche basate sulla rilevazione dell'"impronta digitale" lasciata dal dispositivo (cd. *fingerprint*) durante la navigazione, può essere persino più efficiente ed efficace dell'utilizzo dei cookies e può avvenire ad insaputa dell'utente. Ciò è possibile grazie a come è strutturata la rete e come essa è utilizzata. Essendo la rete dinamica, per poter adattare l'esperienza di navigazione all'esigenze dell'utente, i siti web comunicano con i dispositivi al fine di ricevere istruzioni sulle modalità in cui devono trasmettere il contenuto che propongo.

Ciò avviene ad esempio con la comunicazione da parte del dispositivo al server al quale ci si collega delle caratteristiche dello *user agent*. Lo user agent è un'applicazione installata sul computer dell'utente che si connette ad un processo server, come ad esempio un web browser (Chrome, Internet Explorer, Edge, Firefox etc.) o programmi client per l'invio e la ricezione di posta elettronica (*Mail User Agent* come ad esempio Outlook o Mozilla Thunderbird). Accade che, quando gli utenti visitano un sito web o si collegano alla rete attraverso l'applicativo, una stringa di testo viene inviata per fare identificare al server lo user agent che sta cercando di comunicare con esso. Ciò fa parte delle stesse richieste http e include generalmente informazioni, a titolo esemplificativo e non esaustivo, quali: il nome dell'applicazione client e la versione di quest'ultimo; eventuali plug installati; il sistema operativo utilizzato; la lingua di sistema; i font installati nel sistema operativo.

Per fare un esempio, nel momento in cui ci si collega a siti che propongono contenuti multimediali come ad esempio Netflix, il nostro dispositivo comunica alla piattaforma la qualità del video da riprodurre in punti immagine come ad esempio 1080p o 720p sulla base della risoluzione dello schermo e il formato audio da riprodurre (stero o sorround) sulla base delle periferiche audio collegate. Nel caso di siti di informazione come ad esempio quello del giornale Repubblica, il dispositivo può ad esempio comunicare le dimensioni della finestra di navigazione così da poter mostrare i testi formattati correttamente in maniera che questi siano leggibili, come anche le dimensioni del font da proporre o addirittura tutti i font installati sul sistema operativo così da evitare errori di visualizzazione.

Le informazioni trasmesse e raccolte di per sé non forniscono grandi informazioni sull'utente dato che si tratta di dati impersonali che non permettono singolarmente di identificare l'utente. L'alto numero di interazioni che avvengono fra il dispositivo e la rete e il livello di personalizzazione di esperienza richiesta dall'utente può però aumentare il numero di informazioni e quindi l'entropia. Maggiore è il numero di informazioni che si hanno a disposizione, maggiori sono le probabilità di effettuare il cd. *single out* dell'utente che sta visitando il sito o che sta fruendo del servizio, tutto ciò senza l'utilizzo dei cookies.[8]

necessario utilizzare tecniche di *Virtual Private Networking* o *Tunneling* o appoggiarsi a reti particolari come la rete TOR. Anche con questo genere di strumenti non è sempre garantito l'anonimato.

[8] Questi sono solo alcuni dei dati potenzialmente comunicati, ulteriori esempi sono: la risoluzione in pixel del dispositivo; le dimensioni della finestra aperta dall'utente; configurazioni dell'utente; caratteristiche ambientali (eventuali parti hardware come fotocamera, lettore di impronta digitale, flash e così via); comportamento dell'utente; user agent mobile o desktop; plugin installati; time-zone; ecc..

Fingerprinting e protezione dei dati personali

Come appena illustrato è possibile osservare passivamente le caratteristiche di un dispositivo per tracciarne un'impronta digitale, ed effettuarne il cosiddetto fingerprinting. Al contempo attivamente, è possibile invece identificarne caratteristiche attraverso l'osservazione delle interazioni per mezzo di Script Java o altre forme di codice eseguiti dal client locale, oppure con tecniche quali la rilevazione delle dimensioni della finestra, della risoluzione dello schermo e simili.

Per gli utenti è difficile determinare se il server al quale sono collegati sta tracciando un fingerprint del dispositivo utilizzato poiché ciò può avvenire ad insaputa dell'utente. Se anche fosse possibile affermare che ciò avviene, nulla è dato sapere circa quali dati sono raccolti e quali sono poi consevati, per quali finalità ciò avviene e a quali soggetti questi dati sono comunicati. È inoltre molto difficile evitare questo tipo di profilazione o identificazione, in alcuni casi questa è necessaria proprio per riceve i dati in forma corretta. Se anche questa esigenza tecnica può giustificare in parte l'utilizzo di tecniche di device fingerprinting, di fatto nella maggior parte dei casi ne viene fatto un uso promiscuo con secondi fini.

Dal punto di vista normativo a livello europeo, lo stesso gruppo di lavoro articolo 29 ha espresso un parere in materia[9], con il quale affronta il tema del device fingerprinting e afferma che l'art. 5, paragrafo 3 della direttiva 2002/58/CE[10], relativa alla vita privata e alle comunicazioni elettronica, si applica nel caso di tracciamenti di questo tipo. L'articolo in questione stabilisce che gli Stati membri debbano assicurare che "l'archiviazione di informazioni oppure l'accesso a informazioni già archiviate nell'apparecchiatura terminale di un abbonato o di un utente" siano consentiti unicamente a condizione che l'abbonato o l'utente in questione abbia espresso preliminarmente il proprio consenso, dopo essere stato informato in modo chiaro e completo, a norma della direttiva 95/46/CE6, sugli scopi del trattamento di questi dati.

In un precedente parere, il gruppo di lavoro si era espresso in materia *cookies*, affermando che il suddetto articolo si applica non esclusivamente ai *cookies*, bensì anche a "tecnologie simili". È quindi assodato che per effettuare il tracciamento e la profilazione degli utenti attraverso tecniche quali il fingerprinting è necessario informare l'utente e raccoglierne il consenso. Il gruppo di lavoro sottolinea inoltre che questa tecnica può costituire trattamento di dati personali.

Il quadro normativo attuale è in realtà mutato rispetto al momento storico nel quale è stata espresso questo parere, ad oggi il testo di riferimento in materia di protezione dei dati personali è costituito dal GDPR, il Regolamento (EU) 679/2016, e la Direttiva 2002/58/CE, sarà a breve sostituita da un regolamento europeo. Di fatto facendo un'analisi della normativa precedente, delle pronunce del WP29 e del nuovo regolamento, appare evidente che anche quest'ultimo tutela direttamente i diritti dell'interessato in tema di profilazione e quindi anche nel caso di tecniche di fingerprinting, con la necessità, per chi effettua il trattamento, di ottenere il consenso dell'utente e informalo del trattamento stesso.

[9] Gruppo di lavoro è stato istituito in virtù dell'articolo 29 della direttiva 95/46/CE. È l'organo consultivo indipendente dell'UE per la protezione dei dati personali e della vita privata. I suoi compiti sono fissati all'articolo 30 della direttiva 95/46/CE e all'articolo 15 della direttiva 2002/58/CE. Ad oggi è stato sostituito dall'EDPB, *european data protection board,* istituito dal Regolamento (EU) 679/2916.

[10] La direttiva sarà a breve sostituita da un Regolamento, al fine di uniformare la disciplina, al vaglio alla commissione europea. I regolamenti a differenza delle direttive sono direttamente applicabili nei paesi europei senza la necessità di un atto di recepimento.

GDPR e device fingerprinting

Il nuovo regolamento in materia di protezione dei dati personali (regolamento (EU) 2016/679, conosciuto anche come GDPR) non menziona espressamente le tecniche di fingerprinting. Proprio sulla base di una scelta ponderata sull'esperienza pregressa, il legislatore europeo ha infatti scelto di rimanere tecnologicamente neutrale. La genericità delle norme permette di garantire che nuove tecnologie non rimangano escluse dalla disciplina, evitando la necessità di ricomprenderle con pareri e pronunce *ad hoc*. Nonostane il considerando n. 30 del GDPR citi espressamente i cookie, il testo normativo rimane generico, tecnologicamente neutro, flessibile e in linea con l'idea di progresso tecnologico, non fornendo elenchi esaustivi o tassativi di tecnologie di trattamento dei dati.

Uno dei punti di forza di questo Regolamento è l'ampia definizione di dati personali data dall'art.4:

> **«dato personale»**: qualsiasi informazione riguardante una persona fisica identificata o identificabile («interessato»); si considera identificabile la persona fisica che può essere identificata, direttamente o indirettamente, con particolare riferimento a un identificativo come il nome, un numero di identificazione, dati relativi all'ubicazione, un identificativo online o a uno o più elementi caratteristici della sua identità fisica, fisiologica, genetica, psichica, economica, culturale o sociale.

Dati personali sono tutti quei dati che possono essere riferiti ad un individuo identificabile. Dal punto di vista tecnologico sono quindi incluse informazioni quali il *MAC address* di un dispositivo, l'indirizzo IP, o il codice identificativo per pubblicità contenuto all'interno di un cookie. Sono da includere anche le informazioni e caratteristiche che, raccolte in un insieme e correlate e combinate tra loro, permettono di identificare e individuare il fingerprint di un dispositivo, client o applicativo. Maggiori sono questi dati, più alte sono le possibilità di identificare un individuo poiché aumenta l'entropia, queste informazioni riguardano inoltre direttamente o indirettamente una persona fisica identificata o identificabile. È importante sottolineare che "identificare" non implica ai sensi del GDPR stabilire effettivamente l'identità di un individuo. È sufficiente che, elaborando i dati si possa identificare, anche indirettamente, un individuo, anche sulla base di dati pseudonimi, al fine di eseguire determinate azioni basate su tale identificazione, come ad esempio proporre pubblicità e annunci personalizzati sulla base del profilo dell'utente.

Se lo scopo del device fingerprinting è dunque quello di tracciare gli utenti, costituisce quindi "trattamento di dati personali" ai sensi del regolamento. Un'attività di questo tipo, considerata la giusta base legale, è da considerarsi accettabile, ma necessita del consenso espresso dell'utente e di una chiara informativa circa l'attività di trattamento. Nella pratica dunque è poi necessario, per raccogliere il consenso dell'utente, un'azione informata e non ambigua, come ad esempio l'operare una scelta cambiando delle impostazioni da "no" a "si" con un click. Per fare ciò è necessario che i titolari o responsabili delle attività di trattamento, in primo luogo, manifestino il fingerprinting prima che questo sia effettuato (come già accade nel caso dei cookie) e attendano il consenso informato e libero dell'utente prima di procedere al tracciamento.

Rischi e misure

Il device fingerprinting nasce da un'esigenza tecnica, si è poi evoluto ed è finito per essere sfruttato come forma celata di tracciamento e profilazione finalizzata al marketing ad insaputa dell'utente. Il concetto di legittimo interesse, così come inteso dal regolamento, consiste nel bilanciamento e nel compromesso fra i diritti dell'interessato e del titolare del trattamento, un concetto molto più ambiguo e vago rispetto alle altre basi legali previste dal GDPR. Potenzialmente il rischio che titolari e responsabili giustifichino trattamenti di questo genere con l'interesse legittimo è molto alto. Se il titolare è consapevole, o si presume

lo sia, di adottare tecniche di fingerprinting, nella maggior parte dei casi così non è per l'utente, inconsapevole di essere oggetto di trattamento. Per gli utenti le cose si complicano poiché combattere ed evitare il fingerprinting è complesso e difficile e questi non hanno né gli strumenti né la conoscenza per opporsi ad esso.[11]

Conclusioni

Osservando come sono state finora utilizzate tecniche di fingerprinting, è molto difficile immaginare che titolari e responsabili passino dall'oscurità intenzionale alla piena trasparenza e alla comunicazione diretta e trasparente con gli utenti. Titolari e responsabili del trattamento di dati ottenuti con tecniche di fingerprinting dovrebbero operare un radicale cambiamento e adeguare i trattamenti come già è avvenuto nel caso dei cookie. È ragionevole aspettarsi, in virtù dell'uso che viene fatto di queste tecniche di tracciamento, che molti opteranno per la totale occultazione. Se oggi il tema dei cookie è un argomento per il quale vi è una diffusa consapevolezza da entrambe le parti (interessati e titolari) ed è stato oggetto di un'intensa campagna di sensibilizzazione, ciò non accade per il device fingerprinting. Con l'entrata in vigore del GDPR sarà sicuramente più difficile e costoso effettuare trattamenti sfuggendo alle regole operando al di sotto di esse. Non ci si può aspettare che il fingerprinting scompaia a causa del nuovo regolamento, come a seguito dell'entrata in vigore della direttiva *e-privacy* non sono scomparsi i cookie ma appare quanto meno necessario intraprendere delle azioni per porre fine agli usi che ne vengono fatti che non sono legittimi. Seppure sia evidente che questo tipo di tecnica rientri nei casi regolamentati dalla normativa europea, ciò non appare sufficiente al fine di tutelare i diritti degli interessati, sarebbe necessaria un'opera più invasiva da parte delle autorità di controllo nazionali e del EDPB, come ad esempio l'inclusione dei trattamenti basati su tecniche di fingerprinting nel novero dei trattamenti soggetti obbligatoriamente a DPIA[12], e un controllo diretto. Seppur il regolamento miri a responsabilizzare titolari e responsabili in questo caso gli interessati possono fare affidamento nella maggior parte dei casi solo a strumenti di autotutela.

[11] Esistono alcuni strumenti per evitare o aggirare il problema del device fingerprintin, per un approfondimento: https://panopticlick.eff.org/self-defense; FaizKhademi, Amin & Zulkernine, Mohammad & Weldemariam, Komminist. (2015). FPGuard: Detection and Prevention of Browser Fingerprinting. 293-308. 10.1007/978-3-319-20810-7_21

[12] Anche se sul piano teorico potrebbero essere inclusi nei casi già previsti dal WP29 e dallo stesso regolamento (monitoraggi regolare e sistematico). Un'espressa menzione, esterna al regolamento che quindi non ne inficierebbe i caratteri di generalità e neutralità tecnologica, avrebbe sicuramente un'efficacia persuasiva e preventiva maggiore.

Fonti

Acar, G., Eubank, C., Englehardt, S., Juarez, M., Narayanan, A., & Diaz, C. (2014). The Web Never Forgets. *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security - CCS 14*. doi:10.1145/2660267.2660347

Am I unique? https://amiunique.org/

Boda, K. (n.d.). Cross-browser fingerprinting test 2.0. Retrieved from https://fingerprint.pet-portal.eu/

Englehardt, S., & Narayanan, A. (2016). Online Tracking: A 1-million-site Measurement and Analysis. *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security - CCS16*. doi:10.1145/2976749.2978313

Evercookie. https://samy.pl/evercookie/

Faizkhademi, A., Zulkernine, M., & Weldemariam, K. (2015). FPGuard: Detection and Prevention of Browser Fingerprinting. *Data and Applications Security and Privacy XXIX Lecture Notes in Computer Science,* 293-308. doi:10.1007/978-3-319-20810-7_21

Fingerprinting. https://wiki.mozilla.org/Fingerprinting

Panopticlick. https://panopticlick.eff.org/

Schwartz, J. (2001, September 04). Giving Web a Memory Cost Its Users Privacy. Retrieved from https://www.nytimes.com/2001/09/04/business/giving-web-a-memory-cost-its-users-privacy.html

Szymielewicz, K., & Budington, B. (2018, June 21). The GDPR and Browser Fingerprinting: How It Changes the Game for the Sneakiest Web Trackers. Retrieved from https://www.eff.org/it/deeplinks/2018/06/gdpr-and-browser-fingerprinting-how-it-changes-game-sneakiest-web-trackers

Technical analysis of client identification mechanisms - The Chromium Projects. (n.d.). Retrieved from https://sites.google.com/a/chromium.org/dev/Home/chromium-security/client-identification-mechanisms#TOC-Machine-specific-characteristics

Wang, T., & Goldberg, I. (2013). Improved website fingerprinting on Tor. *Proceedings of the 12th ACM Workshop on Workshop on Privacy in the Electronic Society - WPES 13*. doi:10.1145/2517840.2517851