Ph.D. proposal

Leveraging Browser Fingerprinting to Fight Fraud on the Web

Supervisors

- *Walter* **Rudametkin** (Associate professor, Spirals) <Walter.Rudametkin@univ-lille.fr>
- *Romain* **Rouvoy** (Professor, Spirals) <Romain.Rouvoy@univ-lille.fr>

Research team

The Ph.D. student will join the Spirals project-team led by Lionel Seinturier (Professor, Spirals) <lionel.seinturier@univ-lille.fr> Spirals is a joint project-team between Inria and the University of Lille, within UMR CRIStAL.

Spirals project-team
https://team.inria.fr/spirals

Université de Lille UMR CRIStAL Bâtiment M3, Université de Lille 1, 59655 Villeneuve d'Ascq – FRANCE	**Inria Lille - Nord Europe** Parc Scientifique de la Haute Borne 40, avenue Halley - Bat. B, Park Plaza 59650 Villeneuve d'Ascq – FRANCE

Scientific Context

Browsers and web technologies, such as HTML5, are redefining the limits of what web applications can do. At the same time, concerned web users are becoming aware of practices that jeopardize their privacy, security and comfort, as can be seen by the immense popularity of browser extensions like AdBlock and Ghostery, as well as new legislation concerning the use of cookies. However, a new threat to privacy that leaves no trace on users' devices has emerged. *Browser fingerprinting* [Eckerseley10, Laperdrix16] exploits modern web technologies, protocols and APIs to uniquely identify users. The collected data is stored on servers the user has no control over it. Encryption does little to limit fingerprinting because it is performed by the website you visit; it is not a sniffing nor man-in-the-middle attack. Moreover, it is becoming widespread [Englehardt16], used to complement or even replace cookies for tracking purposes. And new research shows it can be used to track people for extended periods of time [Vastel18]. This is an important threat to privacy.

However, browser fingerprinting also has a positive side since it can be used for security purposes and as a deterrent to digital identity theft. In this domain, it has currently been used in the following applications:

- As a way to re-identify fraudsters before they commit online transactions [Vasilyev15],
- Detect bots or classes of devices [Bursztein16], most often to avoid website scraping,
- Enhance authentication and mitigate session hijacking [Alaca13, Preuveneers15].

Indeed, users are even more tied to their online identities, thus having their online accounts stolen can be disastrous. Websites also wish to protect against fraudsters that often use bots to steal their content, or that buy articles using stolen accounts/credit cards. As both bots and online fraud keep evolving, we need new mechanisms to automatically detect and prevent it.

Ph.D. Project

Positioned in the context of web security and privacy, this Ph.D. project will focus on numerous issues website developers face when attempting to detect and prevent fraud online.

This Ph.D. will benefit from our browser fingerprints dataset, collected through the AmIUnique.org website and browser extensions for over 3 years. The dataset will enable the study of fingerprints diversity, the way they evolve, as well as the impact of fingerprint countermeasures on collected fingerprints. This Ph.D. will focus on providing developers tools to enhance security based on browser fingerprinting. In particular, security to protect against (i) changes in fingerprint identity, (ii) malicious bots, in particular crawlers that either steal content or look for website vulnerabilities, (iii) online account theft and session hijacking.

The objective of this Ph.D. is to define and implement new mechanisms capable of detecting bots and malicious users who change their fingerprints, all while ensuring that normal users are not negatively impacted by these mechanisms.

In order to do so, we propose to apply the following methodology:

1. Evaluate and classify the state of the art of browser fingerprinting for security uses ;
2. Evaluate the state of the art of countermeasures used to circumvent systems that rely on browser fingerprinting to enhance security (*e.g.*, browser extensions, virtual machines, container technologies) ;
3. Build a tool capable of detecting if a fingerprint has been altered. Current simplistic techniques already exist in FingerprintJS2 [Vasilyev15]. However, these do not detect advanced countermeasure techniques, such as FPRandom [Laperdrix17] or Blink [Laperdrix15]. Indeed, because fingerprints are collected client-side, this allows fraudsters to alter them using several different mechanisms ;
4. Analyse the impact of virtualization and containers on fingerprints. Bot farms and fraudsters may use virtual machines and containers, possibly hosted in the cloud ;
5. Extend the approach to tracking algorithms based on machine learning techniques, for example FP-Stalker [Vastel18], may be useful in finding fingerprint abnormalities. Adapt such techniques to focus on fingerprint inconsistencies to detect fraudsters ;
6. Extend the approach to new families of bots are now being based on Google Chrome headless and Firefox headless, or may even operate directly in browser extensions and run while users browse the web. These bots may potentially expose "*normal*" fingerprints, thus fingerprinting may not be capable of detecting them ;
7. Extend the approach to bots assisted by humans are also being used to solve captchas. This is a case of Artificial Intelligence with a human in the loop ;
8. Study the use of behavioral analysis, which may provide a means of detection of *"non fingerprintable"* bots.

Expected Deliverables

As part of this PhD thesis, the candidate is expected to produce the following deliverables:

Id	Description	Type	Due
D1	State of the art of browser fingerprinting threats and countermeasures	Report	T_0+3

D2	Analysis of the AmIUnique datasets	Report	T_0+6
D3	Proof-of-concept fingerprinting software infrastructure (client/server)	Code	T_0+12
D4	Double-blind experimentation on traffic classification	Code	T_0+18

Skills Summary

The Ph.D. candidate will develop her/his skills in Javascript, Python, web security, as well as machine learning and statistical data analysis, among many other technologies.

As is a common practice in the Spirals research team, all source code is expected to be open sourced. The student should publish high-level academic papers, as well as participate in related open source communities. This should assist in the technological transfer from academic prototypes to industry-ready tools.

References

[Acar13] G. Acar, M. Juarez, N. Nikiforakis, C. Diaz, S. Gurses, F. Piessens, and B. Preneel, "**FPDetective: Dusting the web for fingerprinters**". *ACM SIGSAC Conf. on Computer and Communications Security* (CCS'13).

[Bursztein16] E. Bursztein, A. Malyshev, T. Pietraszek and K. Thomas, "**Picasso: Lightweight Device Class Fingerprinting for Web Clients**". *Workshop on Security and Privacy in Smartphones and Mobile Devices* (SPSM'16).

[Eckersley10] P. Eckersley. "**How unique is your web browser?**". *Int. Conf. on Privacy Enhancing Technologies* (PETS'10).

[Englehardt16] S. Englehardt and A. Narayanan, "**Online tracking: A 1-million-site measurement and analysis**". *ACM SIGSAC Conf. on Computer and Communications Security* (CCS'16).

[Laperdrix15] P. Laperdrix, W. Rudametkin and B. Baudry. "**Mitigating browser fingerprint tracking: multi-level reconfiguration and diversification**". *Int. Symp. on Software Engineering for Adaptive and Self-Managing Systems* (SEAMS'15).

[Laperdrix16] P. Laperdrix, W. Rudametkin, B. Baudry. "**Beauty and the Beast: Diverting modern web browsers to build unique browser fingerprints**". *IEEE Symp. on Security and Privacy* (S&P'16).

[Vastel18] A. Vastel, P. Laperdrix, W. Rudametkin, R. Rouvoy. "**FP-STALKER: Tracking Browser Fingerprint Evolutions**". *IEEE Symp. on Security and Privacy* (S&P'18).

[Mowery12] K. Mowery and H. Shacham, "**Pixel perfect: Fingerprinting canvas in html5**", 2012.

[Nikiforakis13] N. Nikiforakis, A. Kapravelos, W. Joosen, C. Kruegel, F. Piessens, and G. Vigna. "**Cookieless monster: Exploring the ecosystem of web-based device fingerprinting**". *IEEE Symp. on Security and Privacy* (S&P'13).

[Alaca13] F. Alaca and P.C. Van Oorschot, "**Device fingerprinting for augmenting web authentication: Classification and analysis of methods**". *Annual Conf. on Computer Security Applications* (ACSAC'16).

[Laperdrix17] P. Laperdrix., B. Baudry and V. Mishra, "**FPRandom: Randomizing core browser objects to break advanced device fingerprinting techniques**". *Int. Symp. on Engineering Secure Software and Systems* (ESSoS'17).

[Preuveneers15] D. Preuveneers and W. Joosen, "**Smartauth, Dynamic context fingerprinting for continuous user authentication**". *Annual ACM Symp. on Applied Computing* (SAC'15).

[Vasilyev15] V. Vasilyev. "**fingerprintjs2: Modern & flexible browser fingerprinting library**", Aug. 2017.

[Iovation] iovation, "**Multifactor Authentication and Online Fraud Prevention Solutions**".