Automated software testing to improve the privacy of browsers

Clément Quinton, MCF, University of Lille (clement.quinton@univ-lille.fr)

Walter Rudametkin, MCF HDR, University of Lille (walter.rudametkin@univ-lille.fr)

Keywords : Browser fingerprinting, Cybersecurity, Web privacy, Software Product Lines, Software engineering

Context

Browser fingerprinting is the process of identifying devices by accessing a collection of relatively stable attributes through Web browsers. We call the identifiers browser fingerprints. Fingerprints are stateless; no information is stored on the client's device. Browser fingerprinting exploits the diversity of modern web configurations, technologies, protocols and APIs (Application Programming Interfaces) to uniquely identify devices. And contrary to tracking cookies that are stored on the device and can be erased, fingerprints are stored on servers the user has no control over. 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.

Browsers are some of the most complex software ever built. There are layers and layers of abstractions, optimizations and features that have taken thousands of man years to build. However, despite the importance of privacy and the risks that your browser's configuration may be used to identify you, almost all of the privacy improvements have relied on purely manual efforts. Browsers and their features are rapidly developed with little interest or caution to privacy issues, opening the door to fingerprinting and other side-channel attacks. Browsers are very configurable and different configurations can introduce many issues. Furthermore, browsers are extensible through extensions, and extensions can introduce privacy bugs of their own that make users identifiable.

Our work around the *Am I Unique* project [1] has shown that browser fingerprinting can identify desktop and mobile devices [2], that users can be tracked over time [3], and that countermeasures are lacking and often counterproductive [4]. We have also shown that browser fingerprinting can be also used for security, to identify bots [5] and to protect websites from fraudsters [6].

Proposed research project

The goal of this research project is to build an extensible browser testing platform that explores browser configurations, executes browsers and generates browser fingerprints that are analyzed to find privacy issues and the components at fault, at development-time.

Browsers are highly configurable software, with hundreds of unique parameters and options, and running on many different platforms and devices. This makes exhaustive exploration of every possible configuration impossible. **The first step** will consist in providing a means to automatically extract Feature Models from browser source code. Feature models [7] are a standard modeling tool from Software Product Lines that allow reasoning on complex configurations.

The second step, through the use of search-based or machine learning techniques, will consist in exploring the configuration space and identifying the privacy side-effects. More specifically, this would likely take the form of a platform that compiles, configures and launches browsers. Once the browser is started, a battery of tests, including collecting fingerprints, are run and analyzed. The objective is to find configurations, and browser components, that are being developed and introduce *privacy bugs*.

Open source

All results will be made open source. Where applicable, we will integrate any tests into the *Am I Unique* project https://amiunique.org, which we manage. *Am I Unique* has thousands of daily visitors and users of our extension, making it an excellent platform for our research studies.

Host laboratory

Spirals (Self-adaptation for distributed services and large software systems) is a project-team at Inria Lille – Nord Europe research centre. Our research program focuses on building distributed software systems and studying their properties. We are leading experts in Web Tracking, Web Privacy, Cloud Computing and Software Product Lines. Our research addresses both large scale empirical studies as well as formal methods to design, build and verify the properties of complex software systems.

The student will have access to our extensive internal infrastructures and datasets, including our test and production servers, as well as VMs for deploying their experiments. They will also be asked to collaborate with our international and national partners.

Required skills

The following skills are required for this project:

- Knowledge of Python
- Knowledge of JavaScript
- Experience with Linux
- Basic knowledge of systems administration and virtual machines
- Proficiency in scientific English (written / spoken)

Any of the following skills would be highly appreciated:

- Advanced JavaScript
- Experience with Java
- Experience writing browser extensions
- Formal methods (e.g. automata, predicate logics, temporal logics)
- Constraint programming
- Machine learning
- DevOps (contrinuous integration, continous deployment)
- Docker
- Experience in C++ (particularly if applied to browser development)

References

1. The Am I Unique project https://amiunique.org

2. Pierre Laperdrix, Walter Rudametkin, Benoit Baudry. **Beauty and the Beast: Diverting modern web browsers to build unique browser fingerprints**. IEEE Symposium on Security and Privacy (S&P 2016). Core Rank: A*. 2018 CNIL-Inria European Privacy Protection Award. https://hal.inria.fr/hal-01285470v2/ https://www.cnil.fr/fr/la-cnil-et-inria-decernent-le-prix-protection-de-la-vie-privee-2018

3. Antoine Vastel, Pierre Laperdrix, Walter Rudametkin, Romain Rouvoy. **FP-STALKER: Tracking Browser Fingerprint Evolutions**. IEEE Symposium on Security and Privacy (S&P 2018), Pages 728-741. Core Rank: A*. https://hal.inria.fr/hal-01652021

4. A. Vastel, P. Laperdrix, W. Rudametkin, R. Rouvoy. FP-Scanner: The Privacy Implications of **Browser Fingerprint Inconsistencies**. USENIX Security 2018. Core Rank: A*. https://hal.inria.fr/hal-01820197

5. Antoine Vastel, Walter Rudametkin, Romain Rouvoy, Xavier Blanc. **FP-Crawlers: Studying the Resilience of Browser Fingerprinting to Block Crawlers.** MADWeb'20, San Diego, California. Best Paper Award. https://hal.archives-ouvertes.fr/hal-02441653

6. Antonin Durey, Pierre Laperdrix, Walter Rudametkin, Romain Rouvoy. **FP-Redemption: Studying Browser Fingerprinting Adoption for the Sake of Web Security.** DIMVA'21, Lisboa, Portugal. https://hal.inria.fr/hal-03212726 Core Rank: B.

7. G. Sousa, W. Rudametkin, L. Duchien. **Extending Feature Models with Relative Cardinalities**. SPLC 2016. Core Rank: A. https://hal.inria.fr/hal-01312751