

Comparison of Privacy Preserving Tools in Web Browsers and Extensions

Bc. Vojtěch Fiala

Supervisor: Ing. Libor Polčák, Ph.D.



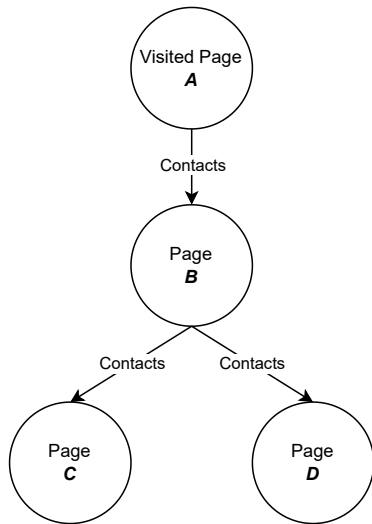
June 24, 2025

- Over 90 % (1) of all websites track their visitors.
- Requests to trackers can be blocked by content blockers.
- Multiple content blockers exist.
- How do they compare?
- Which one to choose?

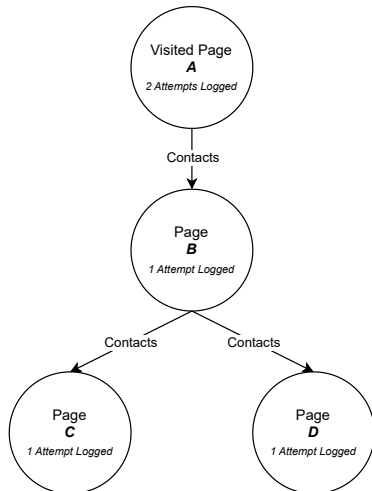


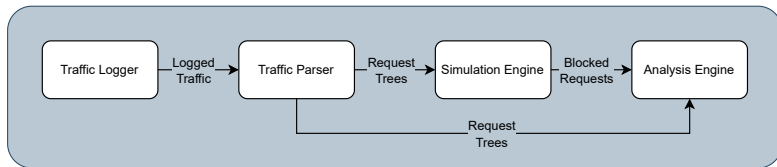
- Design **repeatable** and **deterministic** testing methods to allow for a **fair** comparison of content blockers.
- Ensure that the testing system can support **additional** content-blocking tools in the future.
- Compare **selected** content blockers and provide insights.

- **Determine which requests and how many of them are blocked by each tool.**
- **Monitor** and **record** network traffic from visits to **real-world websites**.
- Use the collected logs to reconstruct **request trees**.
- **Replay** traffic with the content-blocking tool **enabled** and log all **blocked requests**.
- **Propagate** the blocking effects throughout the request tree to identify indirect impacts.



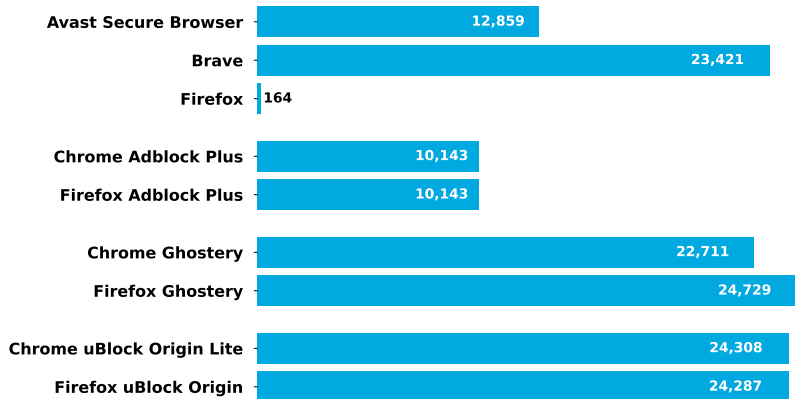
- Trackers use browser APIs to compute a **fingerprint** which **uniquely** identifies the visitor.
- **Determine how many API calls commonly exploited for fingerprinting are blocked by each tool as blocking a request blocks all its API calls.**
- Collect API usage data using a **customized version** of the privacy-oriented **JShelter** browser extension.
- Reuse the previously constructed **request trees** to evaluate content-blocking tools in this context.

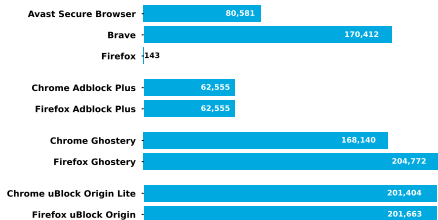




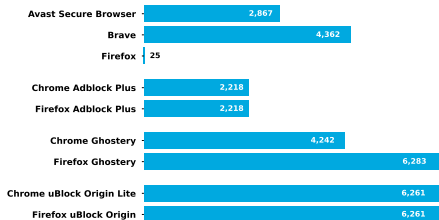
- Data were collected from 937 real pages.
- A total of 100,753 requests were analyzed.
- API calls often misused for fingerprinting were divided into three categories.
 - Browser Properties: 307,226 calls observed
 - Algorithmic Methods: 16,011 calls observed
 - Crawl Fp Inspector (Fingerprinting the Fingerprinters (2)): 7,460 calls observed
- The implementation was verified with unit tests, 100 % code coverage was reached.

- The content blockers tested included browsers and extensions for Chrome and Firefox.

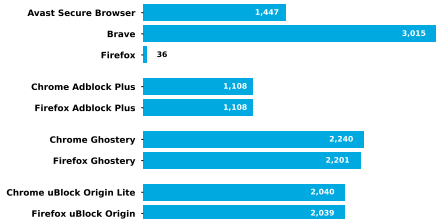




(a) Blocked *Browser Properties* Calls



(b) Blocked *Algorithmic Methods* Calls



(c) Blocked *Crawl Fp Inspector* Calls

- Many content blockers blocked over 20 % of all requests.
- Some extensions performed differently on the Chrome and Firefox browsers.
- Ghostery in Firefox performed best in the most categories.
- While Ghostery in Firefox blocked 24.5 % of all requests, it blocked a lot more API calls:
 - 66.6 % of Browser Properties calls
 - 39.2 % of Algorithmic Methods calls
 - 29.5 % of Crawl Fp Inspector calls
- The blocking ratio was similar even for other content blockers.

- (1) DAMBRA, S.; SANCHEZ ROLA, I.; BILGE, L. and BALZAROTTI, D. *When Sally Met Trackers: Web Tracking From the Users' Perspective*. In: *31st USENIX Security Symposium (USENIX Security 22)*. Boston, MA: USENIX Association, August 2022, p. 2189–2206. ISBN 978-1-939133-31-1.
- (2) IQBAL, U.; ENGLEHARDT, S. and SHAFIQ, Z. Fingerprinting the Fingerprinters: Learning to Detect Browser Fingerprinting Behaviors. In: *2021 IEEE Symposium on Security and Privacy (SP)*. 2021, p. 1143–1161. ISBN 978-1-7281-8934-5.

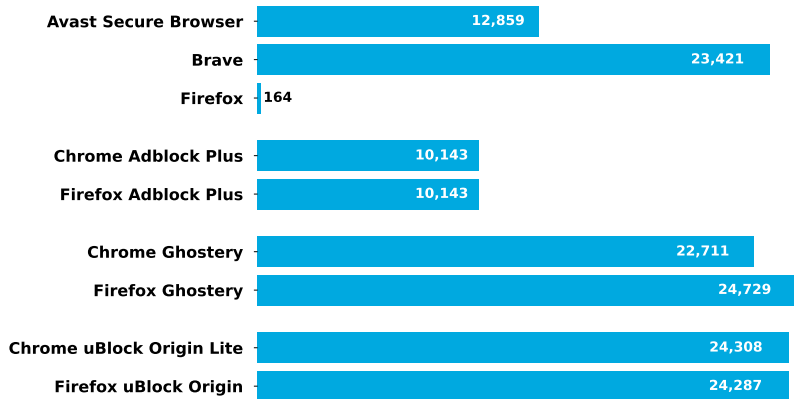


Figure: Total Blocked Requests per Tool

Otázka:

- Jaké výhody a nevýhody má Vaše deterministická metoda testování oproti jiným nedeterministickým metodám (např. oproti online navštívení jedné stránky vícekrát a průměrování výsledků metrik z každé návštěvy)?

Odpověď:

- Výhodou determinismu je zaručená férovost, neboť nástroje jsou testovány na stejných datech, což při návštěvách stránek zajistit nelze. Další výhodou je možnost měřit tranzitivně zablokované požadavky. Nevýhodou je, že blokace požadavku může vyvolat další požadavky, které se takto neprojeví.

Otázka:

- V předložené práci zmiňujete, že reálný webový provoz nelze vždy reprezentovat stromovou strukturou, protože obsahuje cykly. Bylo by reálné ve Vašich metodách upustit od stromové struktury a využít pro zachycení webového provozu obecnější grafy s cykly a s jedním vstupním vrcholem?

Odpověď:

- Stromová struktura umožňuje snadný výpočet měřených metrik, ale jiné vlastnosti (např. unikátní kontaktované domény) může být obtížnější získat. Použité metody by bylo možné upravit pro práci s obecným grafem, ale vyžadovalo by to jiný způsob uchovávání informací o návaznosti požadavků, které jsou ve stromovém grafu reprezentovány implicitně.