

Web ブラウザ毎に分類した時間経過に伴う Web Browser Fingerprint の変化の分析

石川 貴之[†] 磯 侑斗[‡] 桐生 直輝[‡] 塚本 耕司[‡] 高須 航[‡]

山田 智隆[‡] 武居 直樹[‡] 細井 理央[†] 齋藤 孝道[†]

明治大学[†] 明治大学大学院[‡]

1 はじめに

HTTP ヘッダや JavaScript により採取可能な端末や Web ブラウザの特徴を組み合わせた Web Browser Fingerprint (以降, Fingerprint と呼ぶ) を用いた端末識別手法が考案され, 一部のターゲティング広告やリスクベース認証に利用されている. 一方, Fingerprint に利用される端末情報の一部は時間経過に伴い変化するので, Fingerprint による識別能力の低下をもたらすことが我々の研究[1]により明らかとなった. [1]では収集した Fingerprint を Web ブラウザの区別なく扱っていた, 本論文では Web ブラウザの違いに着目する. Fingerprint は Web ブラウザから採取できる特徴を利用しているので, Web ブラウザ毎に端末情報を分析した場合に, その情報が変化する時期に違いが生じ, 識別能力の低下が生じにくい Web ブラウザが存在すると思われる. これは, 長期間にわたる同一端末の識別が可能となることを示す. 本論文では, Google Chrome, Mozilla Firefox, Internet Explorer (以降, IE と呼ぶ) に対して, 我々が運用する Fingerprint の採取を行う Web サイト[3] (以降, Fingerprint 採取サイトと呼ぶ) から得られたデータを Web ブラウザ毎の差異に着目し分析することで, Web ブラウザ毎に分類した時間経過に伴う Fingerprint の変化を示す.

2 Web Browser Fingerprint

2.1 概要

Web ブラウザが Web サーバへアクセスした際に, JavaScript などを用いて, Web サーバが Web ブラウザから採取できる端末の情報を特徴点と呼ぶ. 特に, UserAgent やインストール済みプラグインなど端末の特定に繋がる情報を指す. 特徴点を1つ以上組み合わせたものを Fingerprint と呼ぶ. Fingerprint は端末ごとに異なるため, アクセスした利用者の端末の識別に利用できる.

既存の研究として, Eckersley は HTTP ヘッダ, JavaScript 及び Flash から採取できる特徴点を用い

Analysis of Changes over Time in the Web Browser Fingerprint

[†]Takayuki ISHIKAWA [‡]Yuto ISO

[‡]Naoki KIRYU [†]Koji TSUKAMOTO

[†]Ko TAKASU [†]Tomotaka YAMADA

[‡]Naoki TAKEI [†]Rio HOSOI [†]Takamichi SAITO

[†]Meiji University

[‡]Graduate School of Meiji University

て, Web サイト閲覧者の Fingerprint を収集し, それらの分析を行った. その結果, 収集した Fingerprint の 94.2% が一意な値をとり, Fingerprint によって Web サイト閲覧者の端末を識別できることを示した[2].

2.2 Fingerprint 採取サイト

我々が運用する Fingerprint 採取サイトでは, 我々が独自に導入した特徴点を含め, 端末に関する様々な情報を採取している[1]. また, 同一端末からのアクセスであることを確認するために, Web ブラウザ識別用の HTTP クッキー (以降, UID と呼ぶ) を生成し, Fingerprint 採取サイトのデータベースに保存している. UID により同一とみなせる端末を分析の対象とした. 同時に, 時間経過に伴う Fingerprint の変化を確認するために, アクセス時刻も保存している.

3 採取データの分析

3.1 データセット

Fingerprint 採取サイトにおいて, 2013 年 12 月 6 日から 2014 年 10 月 13 日までの期間で, 正常に採取が完了した 1,814 個の Fingerprint を分析用のサンプルとして用いた. サンプルには 561 個の UID が含まれる. Web ブラウザ毎のサンプル数, UID 数を表 1 に示す.

表 1 Web ブラウザ毎のサンプル数と UID 数

| Web ブラウザ | サンプル数 | UID 数 |
|-------------------|-------|-------|
| Google Chrome | 919 | 296 |
| Mozilla Firefox | 496 | 164 |
| Internet Explorer | 267 | 118 |
| Safari | 109 | 50 |
| Opera | 15 | 9 |
| その他 | 7 | 3 |
| 総計 | 1,814 | 561 |

3.2 エントロピー

エントロピーとは, ある事象がどの程度情報を持つかを示す値である. エントロピーが高い特徴点の中から, Web ブラウザによって大きな差異が生じた特徴点を表 2 に示す. なお, IE ではプライベート IP アドレスを採取することができないので空欄とする.

表 2 の “6 個の特徴点” (以降, FP6 と呼ぶ) とは, 表 2 に示した 6 個の特徴点を組み合わせた

Fingerprint である. また, “3 個の特徴点” (以降, FP3 と呼ぶ) とは, FP6 から 3.3 節で説明する 3 個の変化しやすい特徴点を除いた Fingerprint である.

表 2 Web ブラウザ毎の特徴点のエントロピー

| 略称 | 特徴点 | エントロピー | | |
|-----|----------------------|---------|---------|---------|
| | | Chrome | Firefox | IE |
| PIP | プライベート IP アドレス | 7.80425 | 6.90985 | N/A |
| PGN | インストール済み プラグイン | 7.25398 | 6.57249 | 3.51932 |
| GIP | グローバル IP アドレス | 6.39927 | 4.33697 | 4.16612 |
| FNT | インストール済み フォント | 5.83638 | 4.96499 | 4.50831 |
| HUA | HTTP UserAgent | 5.45492 | 3.21011 | 2.37218 |
| JUA | JavaScript UserAgent | 5.44569 | 3.18816 | 4.50872 |
| FP6 | 6 個の特徴点 | 8.61254 | 7.82105 | 5.92207 |
| FP3 | 3 個の特徴点 | 6.77074 | 5.75121 | 5.14756 |

HTTP ヘッダから採取した UserAgent と JavaScript により採取した UserAgent を比較したところ, IE では他の Web ブラウザと異なりエントロピーに大きな差が生じた. これは, IE で JavaScript により UserAgent を採取した場合には, HTTP ヘッダから採取した場合と異なり, .NET Framework のバージョンやインストールされたツールバーなど, より多くの情報を採取できるためである.

3.3 時間経過に伴う Fingerprint の変化

本節では, 同一 UID を持つサンプルの組み合わせに対して, 時間経過に伴う Fingerprint の変化と, 各 Web ブラウザにおける特徴点の変化する時期の違いについて述べる.

今回の分析では, 2 つの Fingerprint のアクセス時刻の差を求め, それを 4 つの期間に分類し, それぞれの期間における特徴点の一致した割合 (以降, 一致率と呼ぶ) を Web ブラウザ毎に求めた. その結果を表 3 に示す. 表 3 の期間はアクセス時刻の差を日数で分類したものであり, 特徴点の略称は表 2 に対応している.

すべての Web ブラウザにおいて, 21 日経過するまでに一致率が 20%以上減少している特徴点を変化しやすい特徴点とする. 表 3 から, プライベート

IP アドレス, インストール済みプラグイン, グローバル IP アドレスがこれに該当し, FP3 はこれらの特徴点を除いた Fingerprint である. また, IP アドレスは, すべての Web ブラウザにおいて, 1 日経過時点での一致率が 30%以上減少している. これは, モバイル端末の普及により, 複数地点からのアクセスが多いためであると考えられる.

表 3 より, FP3 の一致率が 80%未満に低下する期間は Google Chrome が最も短く, Mozilla Firefox, IE の順に長期化している. 特に Google Chrome では FP3 の一致率は 1 日経過した時点で 69%まで低下している. これらの一致率の低下は, アプリケーションなどのインストールやバージョンアップに伴い, 特徴点が変わることが主な理由として挙げられる. 一方 IE での FP3 の一致率は 14 日経過した時点においても 90%以上を保っており, 他の Web ブラウザと比較しても変化が生じにくいことがわかる. また, IE では他の Web ブラウザでは変化しやすい特徴点であるプラグインリストにおいても 7 日経過時点で 80%以上の一致率を保っていた. これらは IE での Web ブラウザやプラグインのバージョンアップの頻度が, 他の Web ブラウザに比べて少ないためであると考えられる.

4 まとめ

本論文では, Web ブラウザ毎に特徴点を分析することにより, 時間経過による Fingerprint の変化は Web ブラウザによって差が生じることを示した. 分析の結果から, 時間が経過しても Fingerprint に変化が生じにくい Web ブラウザが IE であることが判明した.

5 参考文献

- [1] 磯侑斗, 桐生直輝, 塚本耕司, 高須航, 山田智隆, 武居直樹, 齋藤孝道, “Web Browser Fingerprint を採取する Web サイトの構築と採取データの分析”, コンピュータセキュリティシンポジウム 2014 論文集 p.378-p.385
- [2] P. Eckersley, “How Unique is Your Web Browser?”, In Proc. Privacy Enhancing Technologies Symposium (2010), LNCS vol. 6205
- [3] <https://www.saitolab.org/fingerprint/>

表 3 Web ブラウザ毎の時間経過に伴う特徴点の一致率の変化(一致率が 80%以上は灰色)

| ブラウザ | Google Chrome | | | | Mozilla Firefox | | | | Internet Explorer | | | | |
|-------------|---------------|-------|-------|-------|-----------------|-------|-------|-------|-------------------|-------|-------|-------|-------|
| | 期間(日) | 0~1 | 1~7 | 7~14 | 14~21 | 0~1 | 1~7 | 7~14 | 14~21 | 0~1 | 1~7 | 7~14 | 14~21 |
| 特 徴 点 | PIP | 96.5% | 66.0% | 40.9% | 47.1% | 95.0% | 32.9% | 16.4% | 5.4% | N/A | | | |
| | PGN | 88.3% | 56.4% | 56.5% | 20.1% | 97.7% | 82.6% | 76.5% | 23.4% | 98.3% | 88.2% | 88.3% | 5.5% |
| | GIP | 96.1% | 61.0% | 40.4% | 50.7% | 96.3% | 40.5% | 37.9% | 26.8% | 99.4% | 22.0% | 14.2% | 19.5% |
| | FNT | 93.9% | 84.4% | 71.4% | 64.4% | 98.1% | 94.0% | 99.4% | 97.5% | 98.3% | 94.1% | 98.5% | 94.5% |
| | HUA | 98.2% | 87.6% | 80.3% | 67.1% | 98.9% | 94.8% | 89.9% | 79.9% | 100% | 98.9% | 100% | 100% |
| | JUA | 98.2% | 87.8% | 80.3% | 67.1% | 99.3% | 95.7% | 89.9% | 79.9% | 100% | 98.9% | 100% | 98.4% |
| | FP6 | 79.9% | 20.7% | 6.2% | 2.6% | 92.6% | 7.1% | 4.1% | 1.2% | 97.3% | 19.4% | 12.1% | 6.8% |
| FP3 | 90.0% | 69.4% | 56.7% | 34.0% | 97.2% | 86.2% | 88.0% | 77.3% | 97.8% | 98.5% | 99.7% | 96.2% | |