Browser Fingerprinting

CorkSec, 2017-05-09

Johannes Ahlmann

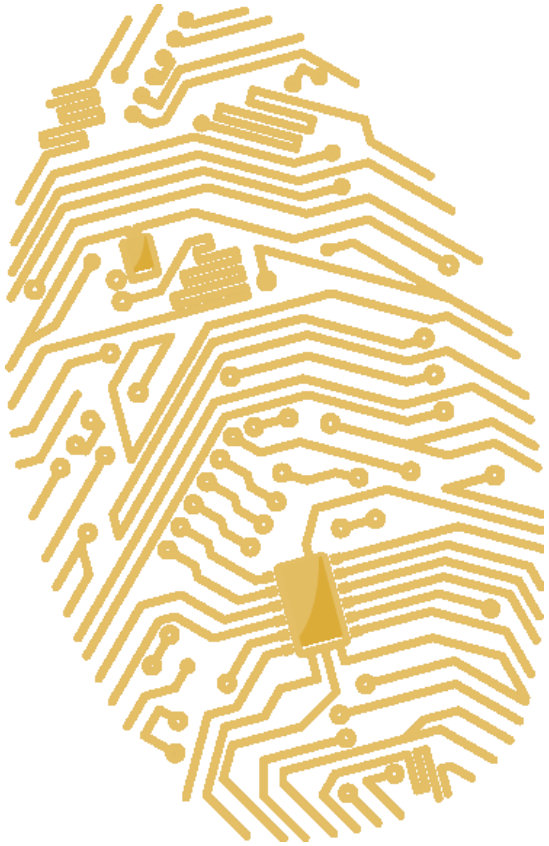


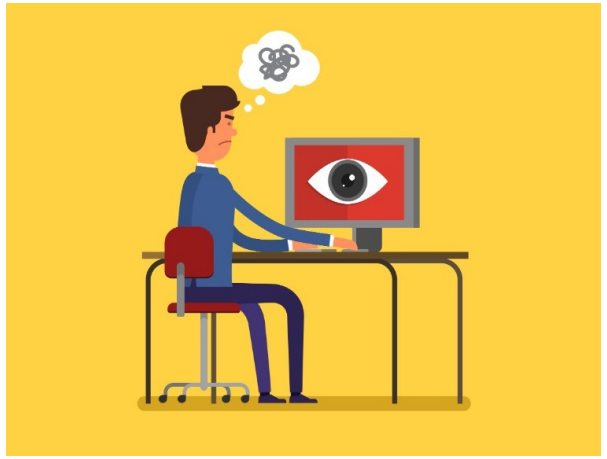
Image: http://bit.ly/2zn96Bx

About Me

- Johannes Ahlmann
- Recently started Fluquid Ltd.
  - Machine Learning (NLP, DL, etc.)
  - Information Extraction
  - Gathering and Enriching Web Data
- Slides + Code
  - Github: fluquid/browser_fingerprint
- Contact: johannes@fluquid.com

Relevance

- Privacy
  - shadow profiles
  - cross browser tracking
  - persistent tracking
- Security
  - journalism sources
  - dissidents
  - incognito de-anonimization
- De-automation
  - anti-fraud
  - anti-bot



Images:
https://weechookeong.files.wordpress.com/2013/07/spying-on-computer.jpg?h=379&w=500
http://i2.cdn.turner.com/money/dam/assets/170329111154-internet-privacy-outrage-1024x576.jpg
https://media.npr.org/assets/img/2016/09/26/gettyimages-173291681-c39025b4d516aa3ccb850e830efcce1527cb1d57.jpg?s=300
http://images.techhive.com/images/idge/imported/imageapi/2014/08/20/15/slide_scrape-100386264-orig.jpg

History

- Has been around forever, but given limited attention
- Particular interest by EFF and in academia since around 2009
- 2010 – EFF releases panopticlick
- Focus often on high-tech aspects
  - TCP stack
  - CPU fingerprint
  - GPU fingerprint
- Obviously “fingerprinting” is used in addition to classic techniques (IP, cookies, LocalStorage, etc.)

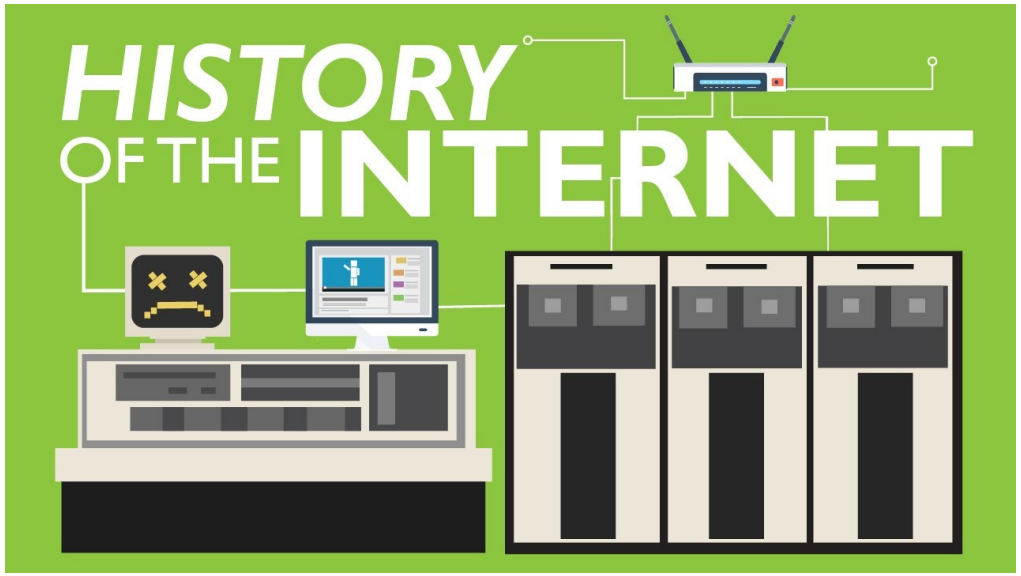


Image: https://img.youtube.com/vi/h8K49dD52WA/mqdefault.jpg

Aspects

1. Vanilla browser, plugins, battery
2. Canvas, WebGL
3. Audio Stack
4. WebRTC
5. Countermeasures

Image: https://www.airsassociation.org/media/k2/items/cache/3b31c9fc44003dace54e8f637805f850_XL.jpg

1. Vanilla Browser

- User agent
- Fonts, font metrics
- Plugins
- Mime-types
- Languages
- Screen, colors, touch
- CPU, cores, OS, timezone

Browser Characteristic	bits of identifying information	one in x browsers have this value	value
Limited supercookie test	0.37	1.29	DOM localStorage: Yes, DOM sessionStorage: Yes, IE userData: No
Hash of canvas fingerprint	14.64	25597.4	604fb19587fd403fafc46fe8917cd73f
Screen Size and Color Depth	5.07	33.51	1536x864x24
Browser Plugin Details	4.17	18.01	Plugin 0: Chrome PDF Viewer; ; mhjfbmdgcfjbbpaeojofohoefgiehjai; (; application/pdf; pdf). Plugin 1: Chrome PDF Viewer; Portable Document Format; internal-pdf-viewer; (Portable Document Format; application/x-google-chrome-pdf; pdf). Plugin 2: Native Client; ; internal-nacl-plugin; (Native Client Executable; application/x-nacl;) (Portable Native Client Executable; application/x-pnacl;). Plugin 3: Widevine Content Decryption Module; Enables Widevine licenses for playback of HTML audio/video content. (version: 1.4.8.970); widevinecdmadapter.dll; (Widevine Content Decryption Module; application/x-ppapi-widevine-cdm;).
Time Zone	3.16	8.94	-60
DNT Header Enabled?	0.75	1.69	True
HTTP_ACCEPT Headers	2.59	6.03	text/html, */*; q=0.01 gzip, deflate, br en-US,en;q=0.8
Hash of WebGL fingerprint	7.68	204.78	3b16fb37ecb998ad4c62786516f83678
Language	0.79	1.73	en-US
System Fonts	5.37	41.26	Arial, Arial Black, Arial Narrow, Book Antiqua, Bookman Old Style, Calibri, Cambria, Cambria Math, Century, Century Gothic, Century Schoolbook, Comic Sans MS, Consolas, Courier, Courier New, Garamond, Georgia, Helvetica, Impact, Lucida Bright, Lucida Calligraphy, Lucida Console, Lucida Fax, Lucida Handwriting, Lucida Sans, Lucida Sans Typewriter, Lucida Sans Unicode, Microsoft Sans Serif, Monotype Corsiva, MS Gothic, MS Outlook, MS PGothic, MS Reference Sans Serif, MS Sans Serif, MS Serif, Palatino Linotype, Segoe Print, Segoe Script, Segoe UI, Segoe UI Light, Segoe UI Semibold, Segoe UI Symbol, Tahoma, Times, Times New Roman, Trebuchet MS, Verdana, Wingdings, Wingdings 2, Wingdings 3 (via javascript)
Platform	1.23	2.34	Win32
User Agent	9.3	630.48	Mozilla/5.0 (Windows NT 10.0; WOW64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/58.0.3029.81 Safari/537.36
Touch Support	5.68	51.31	Max touchpoints: 10; TouchEvent supported: true; onTouchStart supported: true
Are Cookies Enabled?	0.18	1.13	Yes

2. Canvas, WebGL

- Canvas Fonts
- Renders information using
  - 2D Canvas
  - 3D WebGL
- Each hardware/system will render information slightly differently
- In principle works across browsers, operating systems
- Possibly quite difficult to fake well

Your Fingerprint :

Signature	✓ AC2475F8
Uniqueness	99.82% (325 of 176794 user agents have the same signature)

Image File Details : BrowserLeaks.com <canvas> 1.0

File Size	6089 bytes
Number of Colors	659
PNG Hash	5C679989F3F3EB5CEB5EFB2B181E4D43

WebGL Context Info :

Supported Context Name(s)	**webgl2**, webgl, experimental-webgl
GL Version	WebGL 2.0 (OpenGL ES 3.0 Chromium)
Shading Language Version	WebGL GLSL ES 3.00 (OpenGL ES GLSL ES 3.0 Chromium)
Vendor	WebKit
Renderer	WebKit WebGL
Antialiasing	True
ANGLE	True, Direct3D 9
Major Performance Caveat	True

Debug Renderer Info :

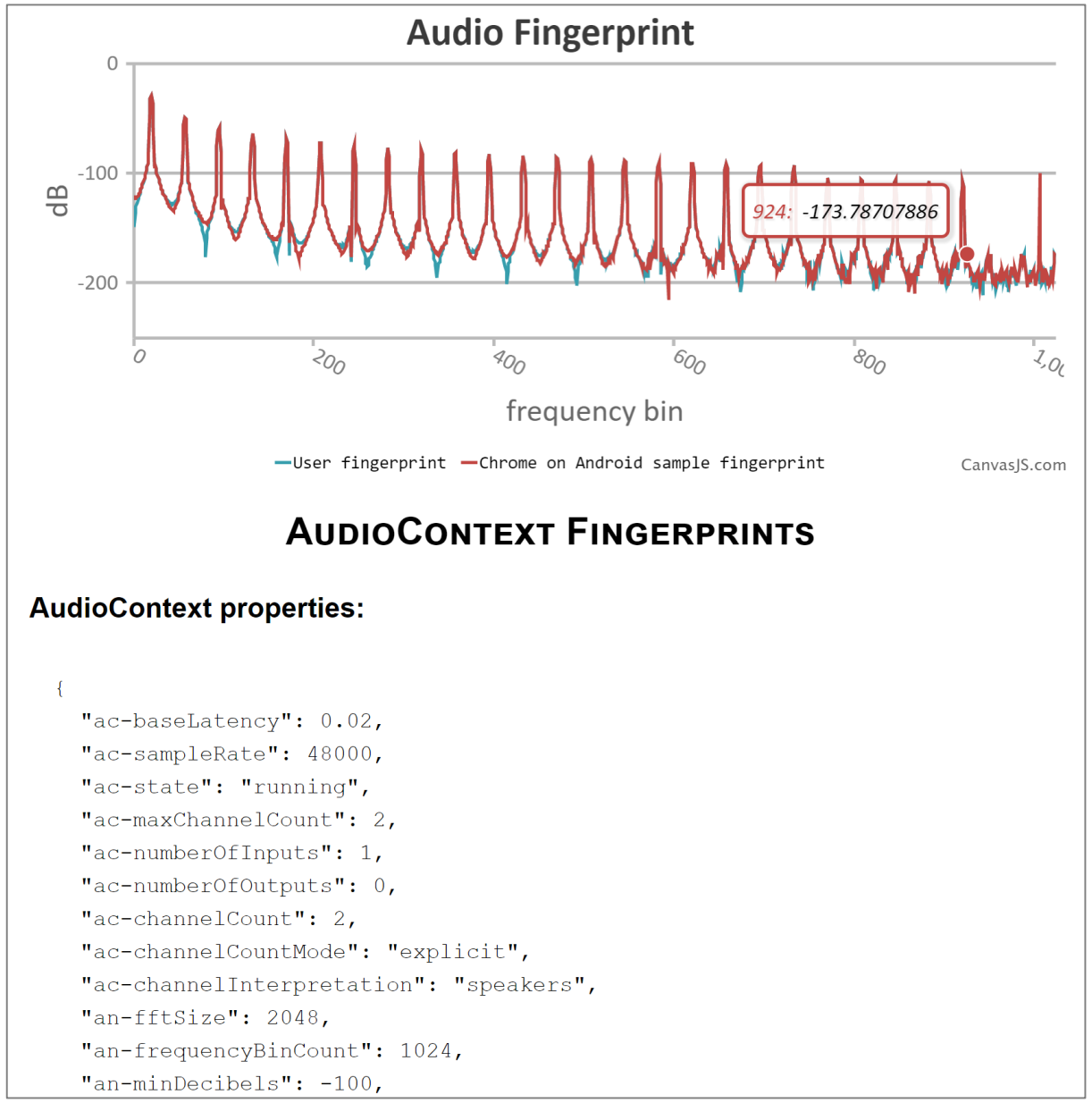
Unmasked Vendor	! Google Inc.
Unmasked Renderer	! Google SwiftShader

WebGL Fingerprint :

WebGL Report Hash	42514FA46404C122B71A608CCEC08971
WebGL Image Hash	15C2CD008A6BFD6DD3D02FD2B7EB3F14
WebGL Image	

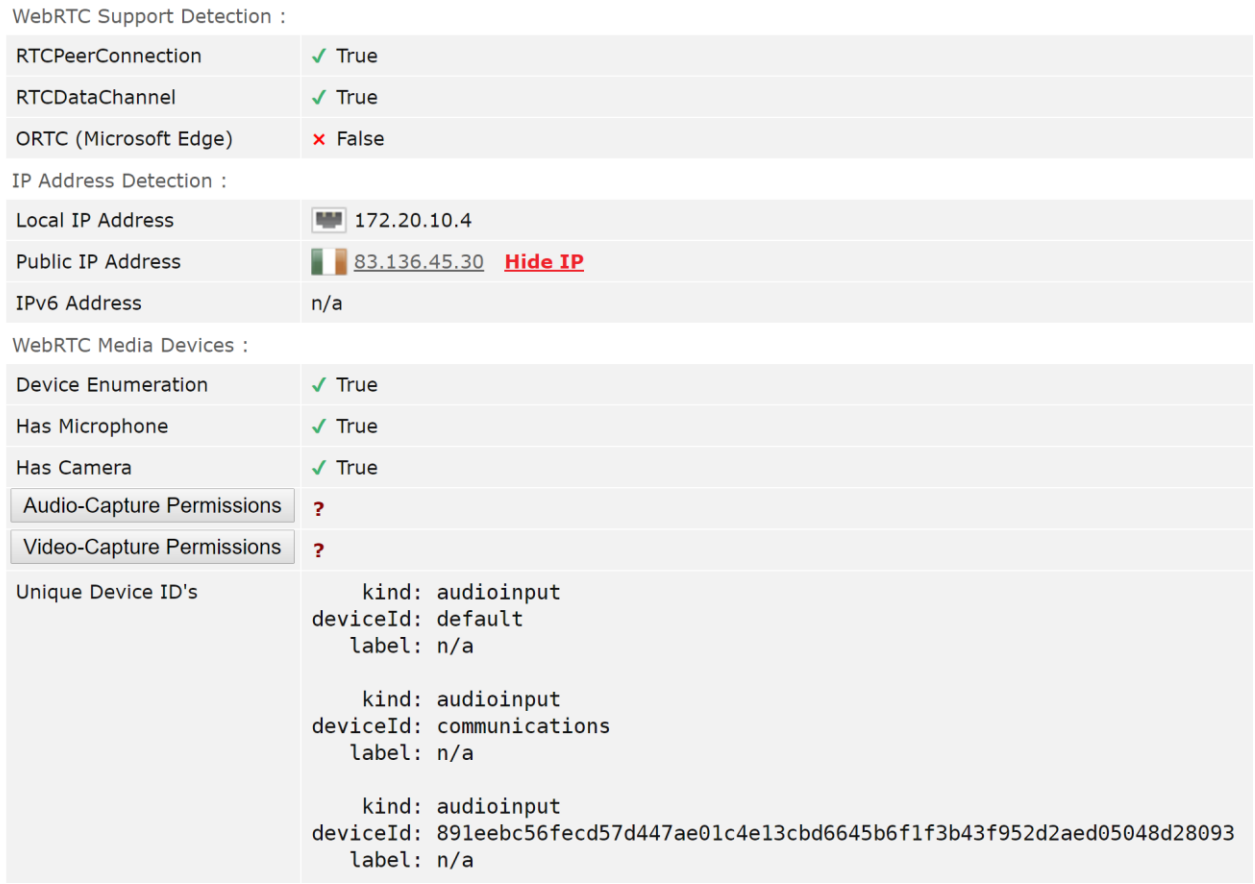
3. Audio Stack

- Does not record/play audio
- Captures oscillation/compression properties of your machine's audio stack itself
- OscillatorNode -> DynamicsCompressorNode -> OfflineAudioContext



4. WebRTC

- Can leak local IP address
- Input/Output
  Device Enumeration
  (At least hashes are self-generated)
- “collects all available candidate addresses, including on local interfaces and makes them available to the web application without explicit permission from the user.” (source)



WebRTC Support Detection :	
RTCPeerConnection	✓ True
RTCDataChannel	✓ True
ORTC (Microsoft Edge)	× False
IP Address Detection :	
Local IP Address	172.20.10.4
Public IP Address	83.136.45.30 Hide IP
IPv6 Address	n/a
WebRTC Media Devices :	
Device Enumeration	✓ True
Has Microphone	✓ True
Has Camera	✓ True
Audio-Capture Permissions	?
Video-Capture Permissions	?
Unique Device ID's	kind: audioinput deviceId: default label: n/a kind: audioinput deviceId: communications label: n/a kind: audioinput deviceId: 891eebc56fecd57d447ae01c4e13cbd6645b6f1f3b43f952d2aed05048d28093 label: n/a

4. Countermeasures

- Proper Counter-Intelligence
- Act of blocking, protecting, faking responses can lead to track-ability itself
- If you are the only person using a particular technique/solution that's a 100% detection rate ;)

Network Filters Detection :	
HTTP Proxy	✓ not detected
Tor Browser Detection :	
TOR Relay IP	✓ not detected
CSS Fonts Protection	✓ not detected
HTML5 Canvas Protection	✓ not detected
WebGL Blocking (NoScript)	✓ not detected
TBB Banned Ports	✓ not detected
Adblock Detection :	
ABP Type	✓ Adblock Plus not detected
Subscriptions	0
Local Content Filters :	
Privoxy	✓ not detected
Proxomitron	✓ not detected
Adguard	✓ not detected
Ad Muncher	✓ not detected

Solutions

- No real end-to-end solution
- Don’t stick out!
  - You can’t not communicate
  - Use as common a setup as possible
  - If you fake a profile, fake it consistently and choose a low-tech target
- Privacy browsers
  - Epic
  - Brave
  - Comodo Dragon/Ice Dragon
  - Tor Browser
- Privoxy
- Browser Plugins
  - Privacy Badger
  - Disconnect
  - uBlock
  - uMatrix
- Ideally we’d want to instrument JS engine to intercept calls (electron, nw.js)
- Hopefully fingerprinting is brittle, so that small perturbations cause false negatives

Resources

Online

- browserleaks.com
- amiunique.org
- panopticlick.eff.org
- uniquemachine.org
- browserspy.dk
- The web never forgets
- OpenWPM Tracking Study
- Intro to NAT/Firewall problem
- Princeton Web Census

Show me the Code

- https://github.com/Valve/fingerprintjs2
- https://github.com/jackspirou/clientjs
- https://github.com/RobinLinus/ubercookie
- https://github.com/qqTYXn7/browserprint
- https://github.com/ghostwords/chameleon
- https://github.com/AlexanderSelzer/BeaverBird
- https://github.com/dillbyrne/random-agent-spoofer
- https://github.com/efforg/panopticlick-python
- https://github.com/Song-Li/cross_browser
- https://github.com/citp/TheWebNeverForgets