

# Everything you always wanted to know about web-based device fingerprinting

(but were afraid to ask)

Nick Nikiforakis



# echo `whoami`

- Postdoctoral researcher at KU Leuven
- Working, mainly, on web security and privacy
- Identify online ecosystems
  - Players
  - Interactions
  - Common patterns
- Search for systematic problems

The New York Times

Business Day

## Media &amp; Advertising

Search All NYTimes.com

Go

WORLD U.S. N.Y. / REGION BUSINESS TECHNOLOGY SCIENCE HEALTH SPORTS OPINION ARTS STYLE TRAVEL JOBS REAL ESTATE AUTOS

Search

Global

DealBook

Markets

Economy

Energy

Media

Personal Tech

Small Business

Your Money

60% korting



» Klik hier

Advertise on NYTimes.com

## Study Finds News Sites Fail to Aim Ads at Users

By TANZINA VEGA

Published: February 13, 2012

Web sites for newspapers, magazines and television stations might be hungry to make money with digital advertising, but you wouldn't know it by the way some of them do business online.

Enlarge This Image



The Economist's home page is not unusual in displaying ads for the company's products.

A new study released Monday by the Pew Research Center Project for Excellence in Journalism looked at 22 news Web sites and more than 5,300 digital ads. It found that many of the sites had not attracted the same advertisers online as they did on other platforms.

In part, these sites were failing to attract online ads because they were not using technology that would customize ads based on their users' online behavior. For example, a user searching for tickets to a Broadway show might see ads for that show.

The study, which looked at Web sites for 11 newspapers, four magazines and six television outlets, as well as two online-only sites, focused on premium digital ad placements on home pages or at the top of article pages, which have generally cost more to buy.

"One of the great challenges that faces the financial future of journalism is, how can you begin to charge more for digital advertising?" said Tim Rasmussen, the director of the center. "The

f RECOMMEND

TWITTER

LINKEDIN

SIGN IN TO E-MAIL

PRINT

REPRINTS

+ SHARE

Log in to see what your friends are sharing on nytimes.com.  
Privacy Policy | What's This?

f Log In With Facebook

## What's Popular Now f

A Senate in the Gun Lobby's Grip



Messing With the Wrong City



Advertise on NYTimes.com

MOST E-MAILED

MOST VIEWED



# 3<sup>rd</sup> Party Tracking

- “Suddenly” all sorts of websites that you’ve never heard about, can create a browsing profile of you and sell it to advertising companies
  - [quantserve.com](http://quantserve.com)
  - [scorecardresearch.com](http://scorecardresearch.com)
  - [addthis.com](http://addthis.com)

# Motivation & Contributions

- Tracking involves more than just 3<sup>rd</sup> party cookies
- Fingerprinting: Telling users apart based on their browsing environments, without extra stateful identifiers
- Thorough study of current fingerprinting practices on the web
- Difficulty of hiding the true nature of a user's browsing environment



# Users reacted...

- 1/3 of users delete first & third-party cookies within a month after they've been setup
- Multiple extensions revealing hidden trackers
  - Ghostery
  - Lightbeam
- Private mode of browsers used to avoid traces of cookies from certain websites

# However...

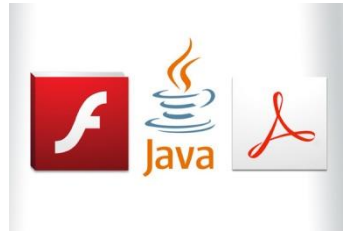
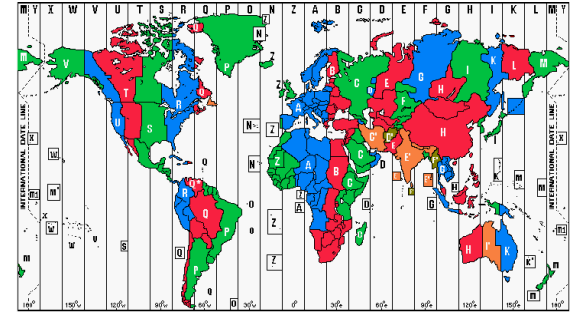
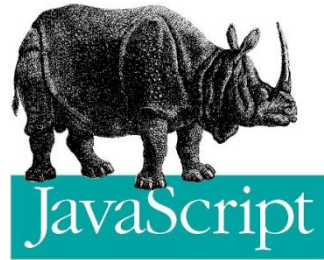


- What if you could track users without the need of cookies or any other stateful client-side identifier?
  - Hidden from users
  - Hard to avoid it / opt-out

## **Web-based device fingerprinting**

- Eckersley showed in 2010 that certain attributes of your browsing environment can be used to accurately track you
- These attributes, when combined, created a quite unique fingerprint of your system?
  - How?

# Properties fingerprinted by Panopticklick

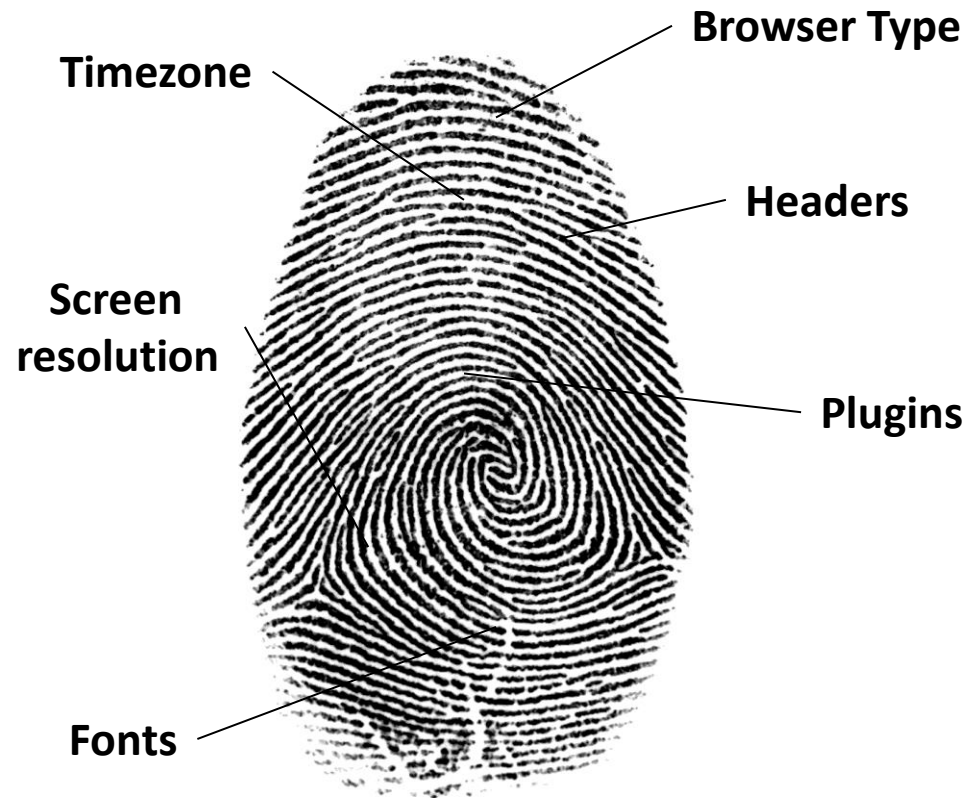


*Maverick*  
Ocean Front Villas  
mandarin tea  
**Regency**  
Sassafras & Ginger  
Dollhouse  
Athletics Dept.





# Resulting fingerprints



- 94.2% of the users with Flash/Java could be uniquely identified
- Simple heuristic algorithms could track updates of the same browser

Daywall

## Feds Are Suspects in New Malware That Attacks Tor Anonymity

BY KEVIN POULSEN 08.05.13 3:57 AM

[Follow @kpoulsen](#)

[Share](#) 8.7k

[Tweet](#) 1,384

[+1](#) 708

[Share](#) 183



# Fast forward 2 years

- In mid 2012, all we knew is that fingerprinting is possible and that a small number of companies offer it as a service
- Questions that begged answering:
  - How are they doing it?
  - Could they do more?
  - Who is using them?
  - How are users trying to hide?
    - Is it working?

# Cookieless Monster: Exploring the Ecosystem of Web-based Device Fingerprinting

Nick Nikiforakis\*, Alexandros Kapravelos<sup>†</sup>, Wouter Joosen\*, Christopher Kruegel<sup>†</sup>, Frank Piessens\*, Giovanni Vigna<sup>†</sup>

*\*iMinds-DistriNet, KU Leuven, 3001 Leuven, Belgium*

*{firstname.lastname}@cs.kuleuven.be*

*<sup>†</sup>University of California, Santa Barbara, CA, USA*

*{kapravel,chris,vigna}@cs.ucsb.edu*

**Abstract**—The web has become an essential part of our society and is currently the main medium of information delivery. Billions of users browse the web on a daily basis, and there are single websites that have reached over one billion user accounts. In this environment, the ability to track users and their online habits can be very lucrative for advertising companies, yet very intrusive for the privacy of users.

In this paper, we examine how web-based device fingerprinting currently works on the Internet. By analyzing the code of three popular browser-fingerprinting code providers, we reveal the techniques that allow websites to track users without the need of client-side identifiers. Among these techniques, we show how current commercial fingerprinting approaches use questionable practices, such as the circumvention of HTTP proxies to discover a user’s real IP address and the installation of intrusive browser plugins.

At the same time, we show how fragile the browser ecosystem is against fingerprinting through the use of novel browser-identifying techniques. With so many different vendors involved in browser development, we demonstrate how one can use diversions in the browsers’ implementation to distinguish

servers. With every request toward a third-party website, that website has the ability to set and read previously-set cookies on a user’s browser. For instance, suppose that a user browses to *travel.com*, whose homepage includes a remote image from *tracking.com*. Therefore, as part of the process of rendering *travel.com*’s homepage, the user’s browser will request the image from *tracking.com*. The web server of *tracking.com* sends the image along with an HTTP Set-Cookie header, setting a cookie on the user’s machine, under the *tracking.com* domain. Later, when the user browses to other websites affiliated with *tracking.com*, e.g., *buy.com*, the tracking website receives its previously-set cookies, recognizes the user, and creates a profile of the user’s browsing habits. These *third-party cookies*, due to the adverse effects on a user’s privacy and their direct connection with online behavioral advertising, captured the attention of both the research community [2], [3], [4] and the popular media outlets [5] and, more since, caused the public’s

Cookieless Monster:  
Exploring the Ecosystem of Web-based Device Fingerprinting

## FPDetective: Dusting the Web for Fingerprinters

Gunes Acar<sup>1</sup>, Marc Juarez<sup>1,2</sup>, Nick Nikiforakis<sup>3</sup>, Claudia Diaz<sup>1</sup>, Seda Gürses<sup>1,4</sup>,  
Frank Piessens<sup>3</sup> and Bart Preneel<sup>1</sup>

<sup>1</sup>KU Leuven, Dept. of Electrical Engineering (ESAT), COSIC, iMinds, Leuven, Belgium  
{gunes.acar,marc.juarez,claudia.diaz,seda.guerses,bart.preneel}@esat.kuleuven.be

<sup>2</sup>IIIA-CSIC, Bellaterra, Spain  
mjuarez@iiia.csic.es

<sup>3</sup>KU Leuven, Dept. of Computer Science, iMinds-DistriNet, Leuven, Belgium  
{nick.nikiforakis,frank.piessens}@cs.kuleuven.be

<sup>4</sup>New York University, Dept. of Media, Culture, and Communication, NY, USA

### ABSTRACT

In the modern web, the browser has emerged as the vehicle of choice, which users are to trust, customize, and use, to access a wealth of information and online services. However, recent studies show that the browser can also be used to invisibly fingerprint the user: a practice that may have serious privacy and security implications.

In this paper, we report on the design, implementation

*services*; K.4.4 [Computers and Society]: Electronic Commerce — *Security*

### Keywords

Web security; privacy; device fingerprinting; tracking; dynamic analysis; JavaScript; Flash

## 1. INTRODUCTION

# Manual analysis of 3 fingerprinting companies



1. Find the domains that they use to serve their fingerprinting scripts
2. Find some websites that use them and extract the code
3. De-obfuscate and analyze
4. Compare and classify



# Step 3 took a while...

```
return;}var _i_b=_i_aa.getElementById(window.io_bbout_element_id);_i_b["value"]=_if_fa;}func
window.io_bb_callback:__if_d;_i_c(_if_fa,_if_fb);}var _i_d={__if_p:function(_if_fc){return _if_
(_if_fc.getUTCDate(),2)+" "+this.__if_ad(_if_fc.getUTCHours(),2)+":"+this.__if_ad(_if_fc.get
__i_e=_if_fd.toString(16);return(_i_m)?this.__if_ad(_i_e,_i_m):_i_e;},__if_u:function(_i_bz)
deAt(_i_g);if(_i_h>=56320&&_i_h<57344)continue;if(_i_h>=55296&&_i_h<56320){if(_i_g+1>=_i_bz.
nue;_i_h=((_i_h-55296)<<10)+(s-56320)+65536;}if(_i_h<128)_i_f+=String.fromCharCode(_i_h);els
f+=String.fromCharCode(224+((_i_h>>12),128+((_i_h>>6)&63),128+(_i_h&63));else _i_f+=String.fr
rn _i_f;},__if_y:function(_if_fe){if(typeof(encodeURIComponent)=="function")return encodeURI
length;_i_g++){var _i_k=_i_j.charAt(_i_g);var _i_l=new RegExp("[a-zA-Z0-9-_.!~*'()]");_i_f+=
nction(_i_bz,_if_ff){var _i_m="";var _i_n=_if_ff-_i_bz.length;while(_i_m.length<_i_n)_i_m+="
JKLMNOPQRSTUVWXYZabcdefghijklmnopqrstuvwxyz0123456789+/",__if_aj:function(_i_bz){var _i_e="
_i_bz.charCodeAtAt(_i_g+1);var _i_r=_i_bz.charCodeAtAt(_i_g+2);var _i_s=_i_p>>2;var _i_t=((_i_p&3
=64;}else if(isNaN(_i_r)){_i_v=64;}_i_e=_i_e+this._i_ej.charAt(_i_s)+this._i_ej.charAt(_i_t)
nction(_i_bz){var _i_w="";var _i_x,chr2,chr3="";var _i_s,_i_t,_i_u,_i_v="";var _i_g=0;var _i
indexOf(_i_bz.charAt(_i_g++));_i_t=this._i_ej.indexOf(_i_bz.charAt(_i_g++));_i_u=this._i_ej.
t)|(_i_t>>4);chr2=((_i_t&15)<<4)|(_i_u>>2);chr3=((_i_u&3)<<6)|_i_v;_i_w=_i_w+String.fromCharCode
ring.fromCharCode(chr3);_i_x=chr2=chr3="";_i_s=_i_t=_i_u=_i_v="";}while(_i_g<_i_bz.length);re
el:12,_i_em:false,_i_en:"",_i_eo:"",_i_ep:true};if(typeof(window.io_install_stm)!="boolean")w
.io_install_flash=_i_z._i_em;if(typeof(window.io_exclude_stm)!="number")window.io_exclude_st
b_url===undefined)window.io_stm_cab_url=_i_o.__if_aq("aHR0cHM6Ly9tcHNuYXJLLmllc25hcmUuY29t")
l_stm_error_handler===undefined)window.io_install_stm_error_handler=_i_z._i_en;if
needs_update_handler===undefined)window.io_flash_needs_update_handler=_i_z._i_eo;if(typeof(w
function(_if_fg){if(_if_fg===undefined)return null;if(typeof(_if_fg)=="object"&&_if_fg.tagName
tElementsByName(_if_fg);for(var _i_g=0;_i_g<_i_ab.length;_i_g++){if(_i_ab[_i_g]._i_dc&&_i_ab[
```

# Results

- After extracting all features, we created a taxonomy of all fingerprinted features, and compared each company to Panopticlick
- Collectively, Panopticlick was fully covered

Browser customizations	<b>ActiveX + CLSIDs</b>
Browser-level User Conf.	<b>DNT Choice</b>
Browser Family & Version	<b>Math constants</b>
OS & Applications	<b>Windows Registry</b>
Hardware & Network	<b>TCP/IP Parameters</b>



# Non-trivial extras

- Non-plugin font detection
  - Comparison of text's width & height
- Native Fingerprinting plugins
  - Accessing highly-specific registry value
- Fingerprint delivery mechanisms
- Proxy detection

# Font Detection through JavaScript

String

Dimensions

I\_DO\_NOT\_NEED\_FLASH

500 x 84

I\_DO\_NOT\_NEED\_FLASH

520 x 84

I\_DO\_NOT\_NEED\_FLASH

580 x 87

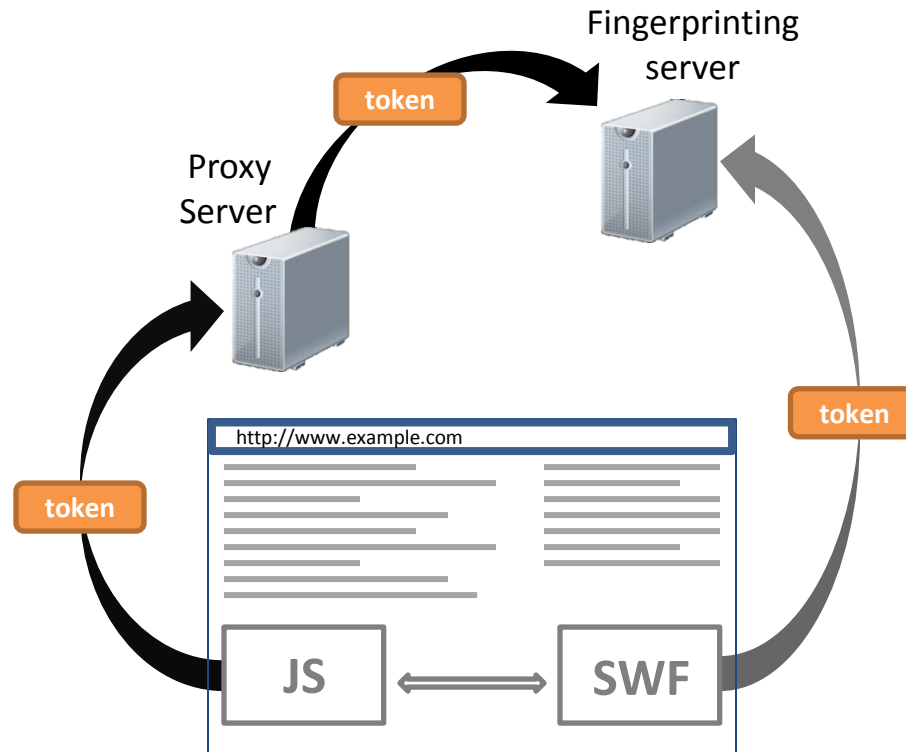
I\_DO\_NOT\_NEED\_FLASH

399 x 82

# Non-trivial extras

- Non-plugin font detection
  - Comparison of text's width & height
- Native Fingerprinting plugins
  - Accessing highly-specific registry values
- Fingerprint delivery mechanisms
- Proxy detection

# Proxy-detection



# Adoption

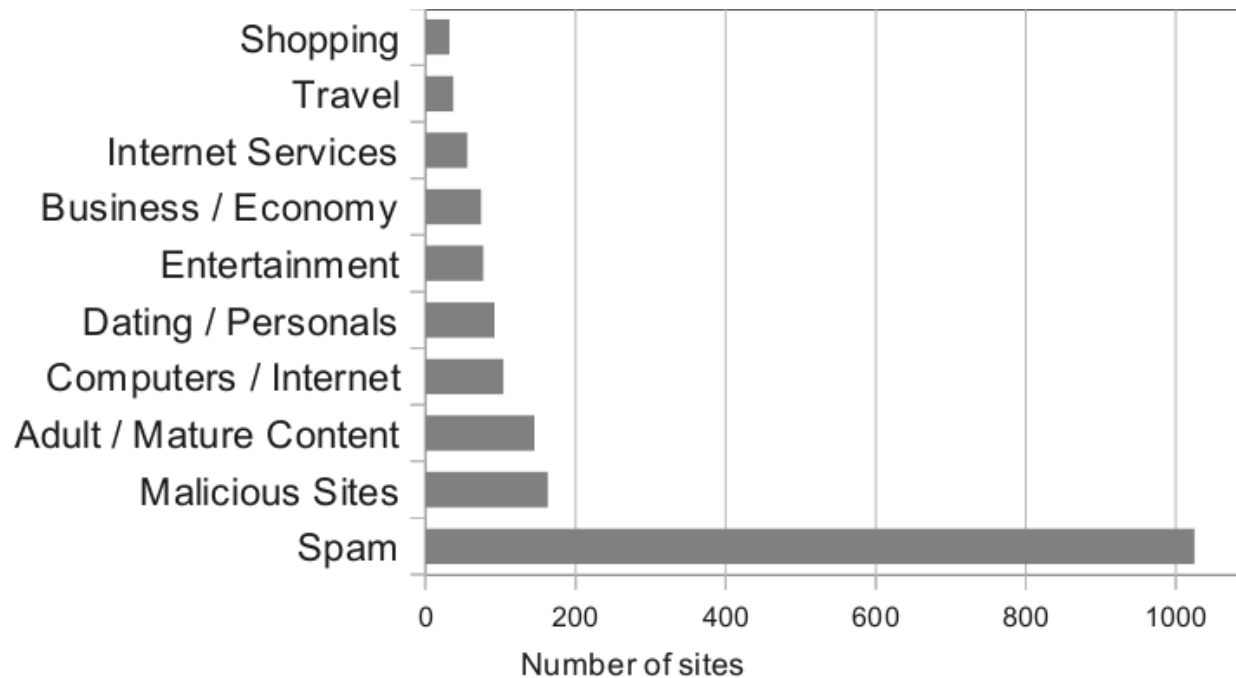
## Dataset A

- Crawled top 10,000 sites, searching for inclusions from the 3 fingerprint providers
- 40 sites discovered
  - Porn & dating sites most prominent
    - Shared credentials & Sybil attacks
  - skype.com the highest ranking one

# Adoption

## Dataset B

– 3,804 domains from Wepawet



# But wait... there's more!

- Can we find unknown fingerprinting parties?
  - How do we separate a fingerprinting script from a generic analytics script?
- Fonts!
  - Separating feature between normal analytics and fingerprinting
  - Second most identifying feature according to Eckersley



# FPDetective

- Fingerprinting-sensitive crawler
  - If fonts are touched, record site
- Detection of font snooping
  - JS-based font probing (Modified browser)
  - Flash-based font probing (decompilation of Flash)





# Adoption (revisited)

## Dataset A

- Crawled top 10,000 sites, searching for inclusions from the 3 fingerprint providers
- 40 sites discovered
  - Porn & dating sites most prominent
    - Shared credentials & Sybil attacks
  - skype.com the highest ranking one

# Adoption (revisited)

# Dataset A

- **145 fingerprinting sites in the top Alexa 10K**
- **DNT does not matter**

# Status

- Fingerprinting is out there
  - Quite a number of new techniques over Panopticlick
- Large and popular sites are using them
- Could they be doing more?
  - How do the browser internals relate to a browser's identity?

# DIY Fingerprinting



- We decided to try some fingerprinting of our own
- Focus on the two special JS objects that fingerprinters probe the most:
  - navigator
  - screen
- Perform a series of everyday operations and search for differences across browsers
  - Add properties
  - Remove properties
  - Modify properties

# Novel methods discovered

- E.g., Natural ordering of properties can give away a browser family, and occasionally, a browser version

navigator.geolocation  
navigator.onLine  
navigator.cookieEnabled  
navigator.vendorSub  
navigator.vendor

navigator.appCodeName  
navigator.appName  
navigator.appVersion  
navigator.language  
navigator.mimeType

↔ navigator.appCodeName  
↔ navigator.appName  
navigator.appMinorVersion  
navigator.cpuClass  
navigator.platform



# Status

- Fingerprinting is out there
  - Quite a number of new techniques over Panopticlick
- Large and popular sites are using them
- There could be more fingerprinting done by the companies
- How could a user react?

# Browser extensions



- Reviewed 11 different browser extensions that spoof a browser's user-agent
  - 8 Firefox + 3 Chrome
  - More than 800,000 users
- Advice to use such extensions:
  - Previous research in web tracking
  - Underground hacking guides
- How do they stand-up against fingerprinting?

# Worse than nothing...



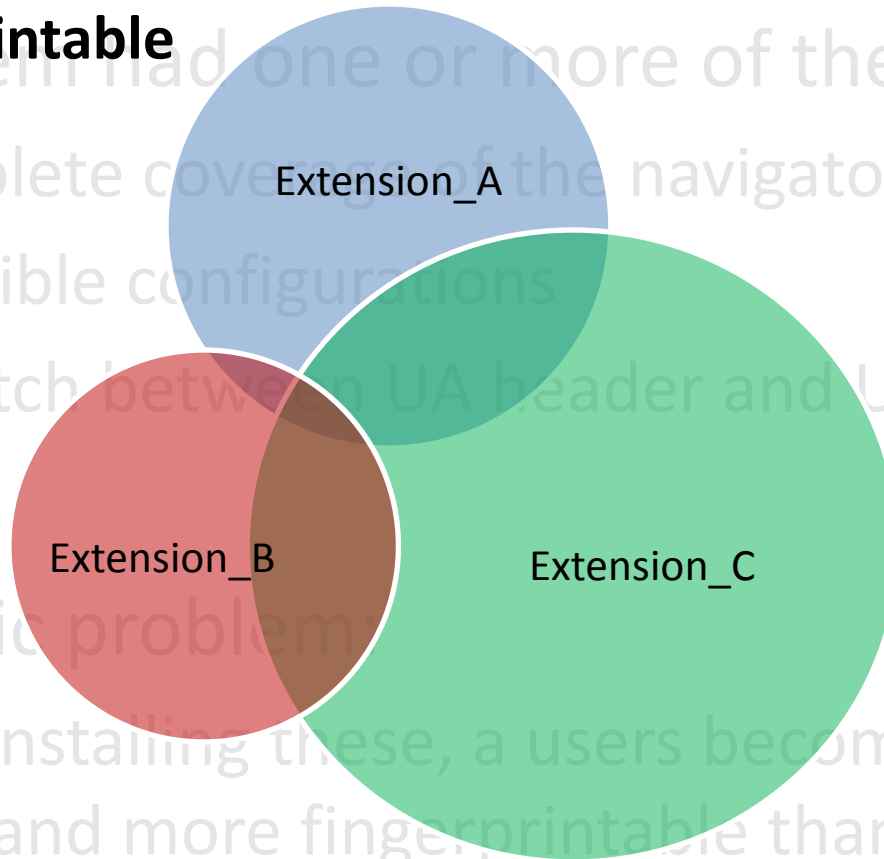
- All of them had one or more of the following:
  - Incomplete coverage of the navigator object
  - Impossible configurations
  - Mismatch between UA header and UA property
- Iatrogenic problem:
  - When installing these, a user becomes more visible and more fingerprintable than before





# Worse than nothing...


- **Fingerprintable Surface**
  - Incomplete coverage of the navigator object
  - Impossible configurations
  - Mismatch between UA header and UA property
- **Latrogenic problems**
  - When installing these, a users becomes more visible and more fingerprintable than before



# Conclusion

- Fingerprinting is a real problem
- Browsers are so complex that it is really hard to make them seem identical
- Current browser extensions should not be used for privacy reasons
- Long term solutions will most-likely not be pure technical ones
  - Legislation required, like in stateful tracking

# "I spy with my free fingerprint kit."



If you're going on a spying mission, you need a fingerprint kit you can hide down your sock.

Now you can get one free, with Trebor Double Agents.

It comes with fingerprint powder, a brush, magnifying glass, record cards and full instructions.

Everything you need to be a dab hand at catching spies.

To get your fingerprint kit, just send us the coupon with any four Double Agents

[nick.nikiforakis@cs.kuleuven.be](mailto:nick.nikiforakis@cs.kuleuven.be)

<http://www.securitee.org>

(for postage and packaging).  
before enemy agents  
foil your plans