TEU00311
What is the Internet doing to me? (witidtm)

Stephen Farrell
stephen.farrell@cs.tcd.ie

https://github.com/sftcd/witidtm
https://down.dsg.cs.tcd.ie/witidtm

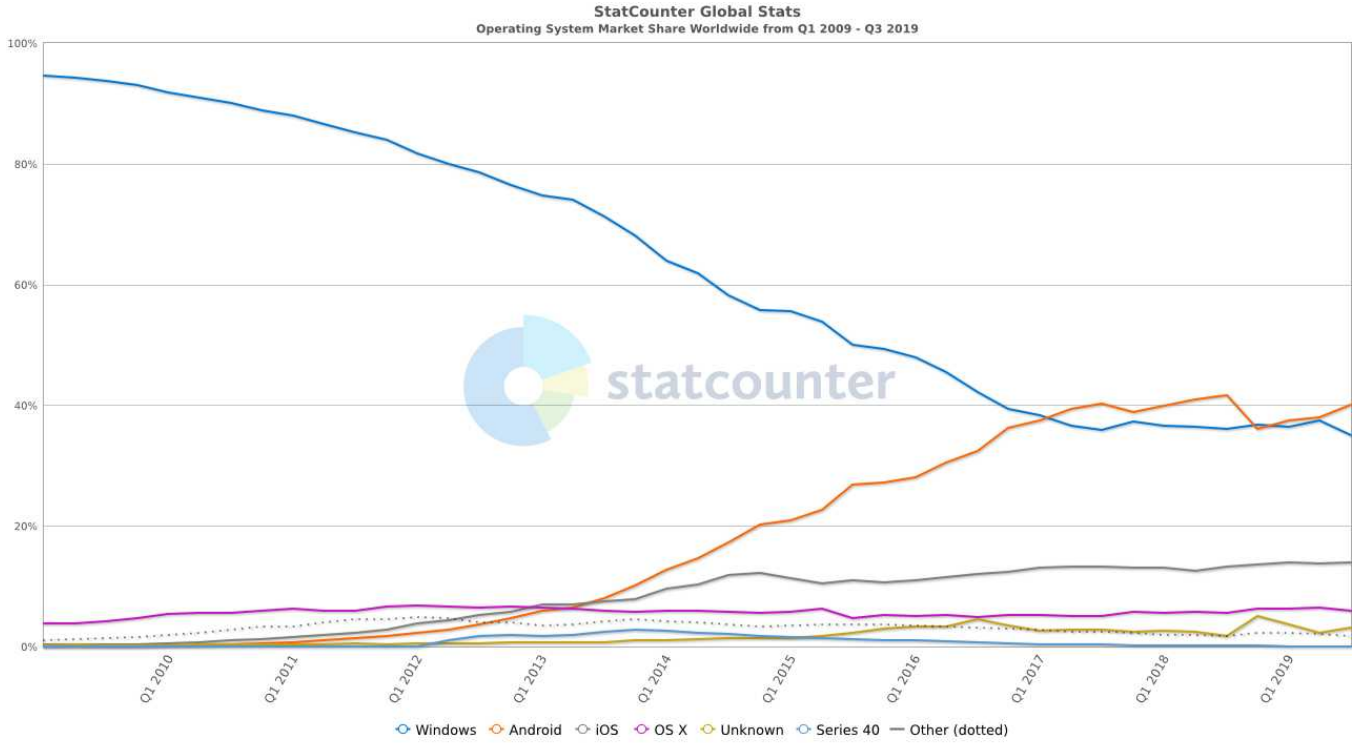
URLs accessed 20190907

Computers, Operating Systems and some example risks

Overall landscape

- Desktops, laptops, tablets, phones, “smart” speakers, home routers, raspberry-Pis, “fitness” devices…are all computers running some operating system
- Differently popular, open, reliable, invasive and general purpose or not
- Your choices should not only be driven by money and ease-of-use, though those factors tend to dominate

Operating Systems Market Shares



Graph: last 10 years

Android
Windoze
iOS
OS X
...
Desktop Linuxes
e.g Ubuntu

https://gs.statcounter.com/os-market-share#quarterly-200901-201903

Linux (desktop) now <1%

Machine/OS issues to consider (1)

- Price, ease of use, updates/support, backup
  - Insurance? Meh;-(
- Security: probability of attack, probability of success
  - Probability of attack relates to market share
- Privacy: hiding in crowds vs. standing out by being more privacy-aware
  - Tension vs. Probability of attack?
- Having backups is your #1 protection – if you do nothing else set that up!
  - But only if you sometimes test restore!
- Automated updates should be considered mandatory, or else you'll eventually suffer
  - Consider how well update system works, some OSes take ages to update

Machine/OS issues to consider (2)

- Control – how much does a machine and operating system offer you real control?
  - Telemetry, settings/preferences, how AppStores/installs are managed
  - Too many knobs != real control
- Invasiveness – to what extent is hardware you bought spying on you?
  - Do you care? Even if not always, are there times when you might care?
- Openness – to what extent can someone find out what's going on under the hood?
  - Even if you can't/don't do that, whether or not others can makes a difference
  - It's a big world and many of the tech issues you face will have been overcome by someone else sometime
    - And likely being nerdy, they won't have been able to resist telling the world how they fixed stuff:-)
- To be fair, there are also some benefits in a more complete but closed ecosystem, like Apple's
  - https://www.apple.com/business/docs/site/iOS_Security_Guide.pdf
  - Note: the above document is written for someone like me, it may be a tough read;-)

End of Life

- Devices do not last forever
  - Ultimately the battery has a limited lifetime in terms of charge cycles
- When (when, not if!) a device fails or is lost/stolen, what will you lose?
  - See “backup” point earlier
- When (when, not if!) a vendor turns off support, or a service (e.g. OS updates, some photo-sharing feature), what willl happen to your stuff?
- When you’re done with the hardware, what do you do?
  - “Two-thirds of used disc drives on Craigslist and eBay contain sensitive data” is a 2016 article, no reason to expect that’ll change v. soon

    https://www.itpro.co.uk/security/26814/two-thirds-of-used-disc-drives-on-craigslist-and-ebay-contain-sensitive-data

Device-based tracking/surveillance

- Anything with “location services” can be unexpectedly dodgy
  - **Image metadata inside the image file can include location**
  - “Fitness” trackers
- Cell-tower history within mobile operators
- List of Wi-Fi networks to which you’ve attached sometime
- **Bluetooth Tracking**
- “Smart” speakers
- “Security” cameras
- Telemetry from applications or OS
- Software update!
  - But didn’t we want that? Yes. But can’t it help tracking? Yes. Hmm...

Blackboard Assignment AS1:

- Research and report on the ways in which a real device/OS you possess or use exposes you to potential device-based tracking, as you normally use the device/OS
- Report: ~1 page, more if needed but keep it short, use URLs/references for tech detail
- Content: say how that tracking could work, describe your exposure, say what you think about that and what you will or will not do about it - justify that last part, esp. If you conclude you'll do nothing
- Deadline: September 20
- Marks: 5%, bonuses for the unexpected and for thorough research!
- We'll talk about two examples, you can't use those for the assignment

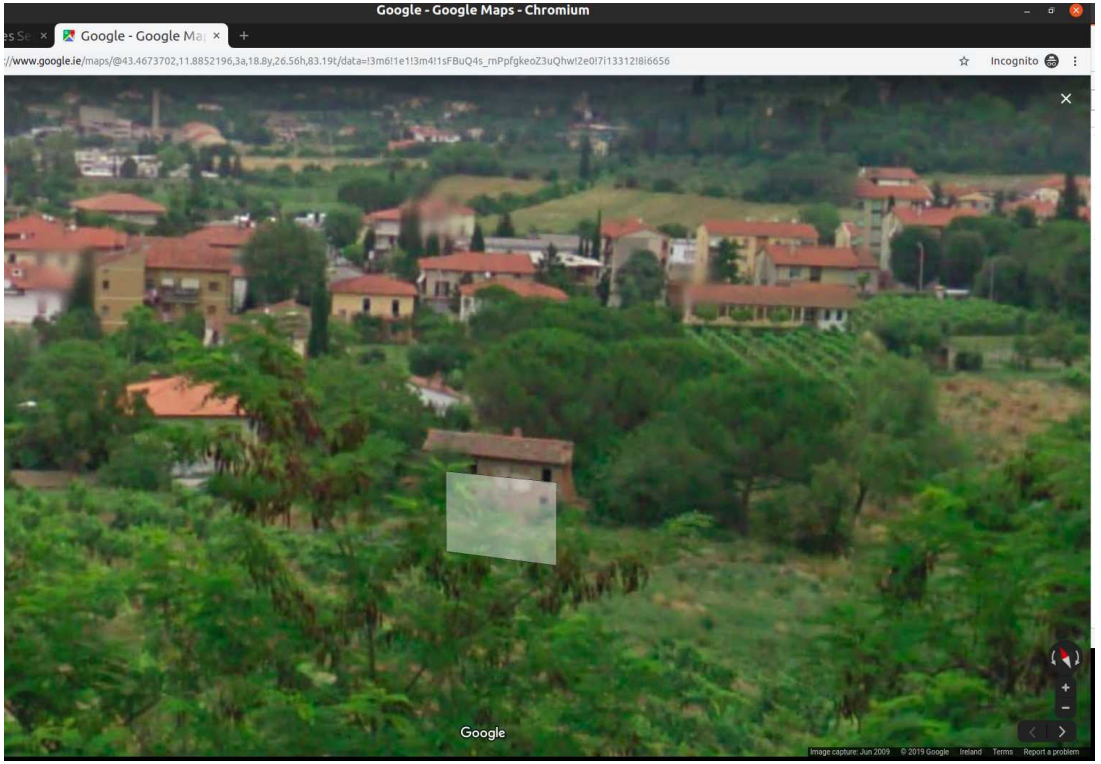
Image Metadata: EXIF

https://photographylife.com/what-is-exif-data
https://helpdeskgeek.com/how-to/how-to-remove-exif-data-from-your-photos/

```
exif DSCN0010.jpg
EXIF tags in 'DSCN0010.jpg' ('Intel' byte order):
--------------------
+-------------------------------------------------------------
Tag                |Value
--------------------
+-------------------------------------------------------------
Image Description  |
Manufacturer       |NIKON
Model              |COOLPIX P6000
Orientation        |Top-left
X-Resolution       |300
Y-Resolution       |300
Resolution Unit    |Inch
Software           |Nikon Transfer 1.1 W
Date and Time      |2008:11:01 21:15:07
YCbCr Positioning  |Centered
Compression        |JPEG compression
X-Resolution       |72
Y-Resolution       |72
Resolution Unit    |Inch
Exposure Time      |1/75 sec.
F-Number           |f/5.9
Exposure Program   |Normal program
ISO Speed Ratings  |64
Exif Version       |Exif Version 2.2
Date and Time (Origi|2008:10:22 16:28:39
Date and Time (Digit|2008:10:22 16:28:39
Components Configura|Y Cb Cr -
Exposure Bias      |0.00 EV
Maximum Aperture Val|2.90 EV (f/2.7)
Metering Mode      |Pattern
Light Source       |Unknown
Flash              |Flash did not fire, compulsory flash mode
Focal Length       |24.0 mm
Maker Note         |3298 bytes undefined data
User Comment       |
FlashPixVersion    |FlashPix Version 1.0
Color Space        |sRGB
Pixel X Dimension  |640
Pixel Y Dimension  |480
File Source        |DSC
Scene Type         |Directly photographed
Custom Rendered    |Normal process
Exposure Mode      |Auto exposure
White Balance      |Auto white balance
Digital Zoom Ratio |0.00
Focal Length in 35mm|112
Scene Capture Type |Standard
Gain Control       |Normal
Contrast           |Normal
Saturation         |Normal
Sharpness          |Normal
Subject Distance Ran|Unknown
North or South Latit|N
Latitude           |43, 28, 2.81400000
East or West Longitu|E
Longitude          |11, 53, 6.45599999
Altitude Reference |Sea level
GPS Time (Atomic Clo|14:27:07.24
GPS Satellites     |06
GPS Image Direction |
Geodetic Survey Data|WGS-84
GPS Date           |2008:10:23
Interoperability Ind|R98
Interoperability Ver|0100
-------------------+---------------------------------------------------------
EXIF data contains a thumbnail (6702 bytes).
```

https://raw.githubusercontent.com/ianare/exif-samples/master/jpg/gps/DSCN0010.jpg

52100 Arezzo, Province of Arezzo, Italy



- Took about 5 minutes to find this in Google street view
- Most of that was finding a way to map degree, minutes, seconds to fractional Lat,Long
- All **automatable**, could easily produce location history from a set of images
- How could such a “leak” be damaging to you or to someone else in your images?

https://www.google.ie/maps/@43.4673702,11.8852196,3a,75y,23.31h,83.05t/data=!3m6!1e1!3m4!1sFBuQ4s_rnPpfgkeoZ3uQhw!2e0!7i13312!8i6656

MAC Addresses

- Device-tracking often (ab)uses **long-term hard-coded identifiers** such as MAC addresses (or IMEI in mobile n/w)
- MAC address: layer 2 address (mostly) hardcoded to radio or other network chip
  - Same form of address used in WiFi and most other network protocols at layer 2, e.g. Bluetooth
  - Roughly how two devices on the same local area network (LAN) identify one another
- Looks like “6C:9C:ED:87:27:60” (48 bits) - 1st half is manufacturer ID (Cisco), 2nd half device-ID (a WiFi router in TCD SCSS)
  - You can look up the manuf ID in the registry: https://www.adminsub.net/mac-address-finder/84:C7 gives a list that includes Sony
  - 2nd half is often fixed, for the lifetime of the device; There is now a 64-bit version, not sure how widely used yet
- You can probably see these in the “about device” tab or similar
- MAC address randomisation is a good idea and starting to be deployed
  - Often, the MAC address only really needs to be stable for a session, so can be randomised
  - But – if you paid for the hotel WiFi that might be based on your MAC address, or an enterprise network might use MAC addresses to decide which machines are allowed on the local network, or the machine may be a switch/router/server where changing MAC address would break stuff or be inefficient
  - So you can’t always randomise

Bluetooth

- Early BT devices regularly broadcast their MAC addresses
- Later BT specs try to fix this, allowing support for randomised MAC addresses
- All good so? Nope!
  - Even if MAC addresses and payload identifiers change randomly, doing so out of sync enables tracking, sometimes indefinitely (read paper, and see next slide)
  - Becker, J., Li, D., & Starobinski, D. (2019). “Tracking Anonymized Bluetooth Devices,” Proceedings on Privacy Enhancing Technologies, 2019(3), 50-65. doi: https://doi.org/10.2478/popets-2019-0036
- Generalising: identifier correlation over time, devices and networks, is a very hard problem to mitigate – similar issues with DHCP leases, IP and email addresses, account names and web artefacts (cookies etc.)

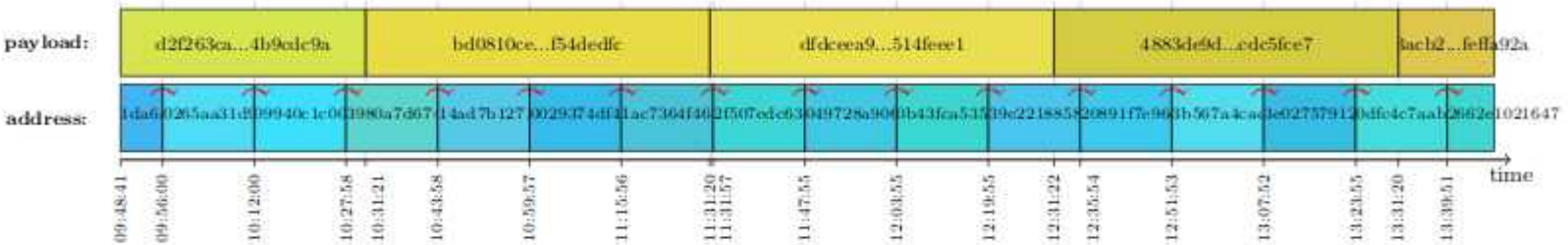


Fig. 9. An experiment illustrating the carry-over effect on Windows 10 devices. Asynchronous value changes allow updating the device identity whenever it changes.

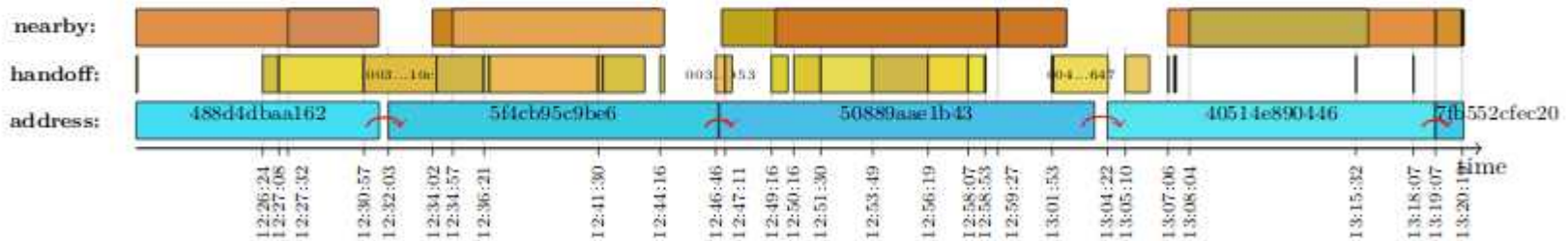


Fig. 10. This timeline shows address-carryover across 5 random addresses on iOS. The first three hops occur via the handoff identifying token, the last one occurs via the nearby token. Different colors denote different values.

Malicious Software (Malware)

- Malware on your devices could:
  - Exfiltrate your data (keyboard sniffers, ID theft, “wallet” attacks)
  - Work for someone else (cryptominers, clickfraud/adware)
  - Be part of a botnet (DDoS)
  - Encrypt your files (ransomware)
  - (Sometimes) survive factory-reset (though less commonly)
- Technical types: viruses, worms, trojans, ...
  - We’re less concerned here with how malware works or details of how to detect/prevent malware doing bad things
  - More interesting is: How can you avoid or recover?
- Note that anti-malware tools (e.g. anti-virus) can look similar to malware - they both want access the “innards” of the OS/network-interactions – that can make such tools attractive as a target

How Malware Arrives

- Pre-installed (accidentally or deliberately)
- Supply chain attacks (e.g., buy a popular npm library, infect that, or generate a new library with a typosquatted name)
- Appear to be an application or browser extension that someone may want (e.g., phone torch app)
- Mail attachment (pdf, zip, tar, ...) in spam or spear-phish
- Drive-by exploitation via web-site/browser vulnerability
- USB stick (watering hole attack)
- USB charger (how much do you trust Dublin Bus? :-)

Actors in the “AppStore” model

- The interested parties include the OS vendors, device manufacturers, mobile network operators, app developers and the services those apps use/enable
  - Your interests likely come last in that list, despite what most of the others may say, and even if they really mean what they say!
  - Their interests won’t always be aligned either (e.g. mobile network operator vs. over-the-top service provider)
  - And there are non-stupid bad actors in the game, esp. in the app developer and services-those-use/enable cohorts
- So what do you think are their interests?
  - OS makers: Google (android), Apple (IOS), Canonical (Ubuntu)
  - Device manufacturers: Samsung, Apple, Huawei, Sony, …
  - Mobile network operators: China Mobile, … Vodafone, … Telefonica, … Eir, Three
  - App developers: Google, Apple, device makers plus many others I’ve never heard of
  - Service providers: Google, Facebook, Twitter, … Netflix, Sky, ...

“AppStore” Permissions model

- Both iOS and android “app” install processes use a “permissions” model, windoze, chromebooks and MAC OS are heading that way more
  - As is Ubuntu, but only sort-of, it being a much more open OS
- Android used (before 2015) present “permissions needed” only at install time, now both it and iOS pop up permission requests when the app first tries to make use of something “sensitive”
  - But are those permission classes really meaningful? And do they protect you? TBH: I find it hard to tell
- Research in this space also seems less concentrated on how well all this protects users and more on how “well accepted” it is by users, e.g.:
  - Reinfelder, Lena, et al. "An Inquiry into Perception and Usage of Smartphone Permission Models." International Conference on Trust and Privacy in Digital Business. Springer, Cham, 2018.
    - https://www.researchgate.net/profile/Lena_Reinfelder/publication/326630389_An_Inquiry_into_Perception_and_Usage_of_Smartphone_Permission_Models_15th_International_Conference_TrustBus_2018_Regensburg_Germany_September_5-6_2018_Proceedings/links/5c0e58564585157ac1b73e03/An-Inquiry-into-Perception-and-Usage-of-Smartphone-Permission-Models-15th-International-Conference-TrustBus-2018-Regensburg-Germany-September-5-6-2018-Proceedings.pdf
  - Rajivan, Prashanth, and Jean Camp. "Influence of privacy attitude and privacy cue framing on android app choices." Twelfth Symposium on Usable Privacy and Security ({SOUPS} 2016). 2016.
    - https://www.usenix.org/system/files/conference/soups2016/wpi16_paper-rajivan.pdf
  - Could be I didn’t spend enough time looking though

What do I use?

- Main devices:
  - Home router: OpenWRT (Turris Omnia)
  - Laptop: Ubuntu (Lenovo)
  - Phone: Sailfish (Sony Xperia)
  - A cardboard box: with a pile of old phones and laptop hard-drives:-)
- Various others @ home and for work, mostly Ubuntu servers on hosted virtual machines (VMs), some real (“bare-metal”) servers and some old laptops
- Other people on my home network use android, iphone, mac, windoze
- Main living room laptop also runs Ubuntu and is entirely usable for non-nerds

What do you use?

What would you like to use?

What do you dislike?

If you wanna try Ubuntu...

- Ideally try with old laptop
  - Drivers for newer models can take a while to catch up with proprietary versions, esp for WiFi
- You can try before installing to hard drive by running from a “live” USB stick
- You can make system dual-boot if you wanna try some more but still keep your earlier OS
- If some of you have an old laptop we could do that in a hackathon session later – interested?