Wi-Fi told me everything about you

Mathieu Cunche

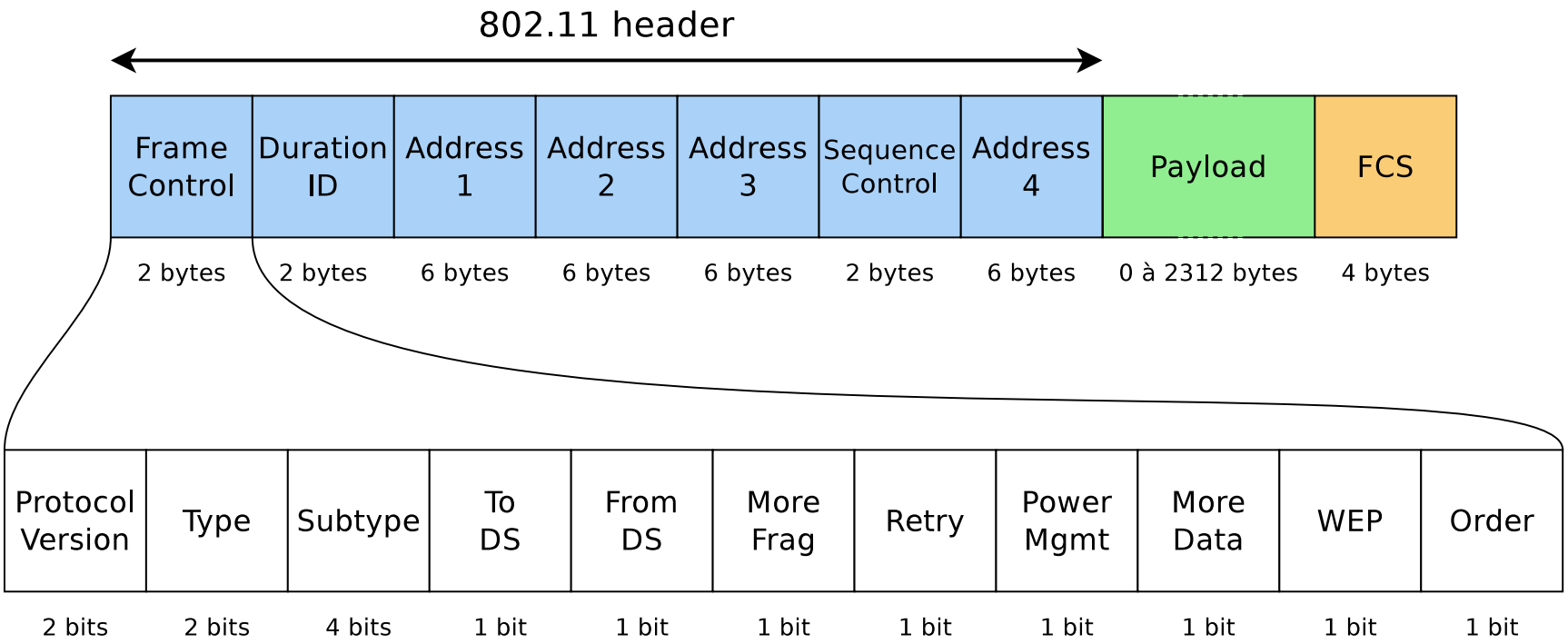
INSA-Lyon CITI, INRIA-Privatics

6 mars 2014

Wi-Fi networking

- IEEE 802.11 standard
  - Specifications for MAC and Physical layers
- Information transmitted by frames
  - **Data**: upper layer datagrams
  - **Management**: beacon, probe request/response, ...
  - **Control**: acknowledgement, ready to send, ...

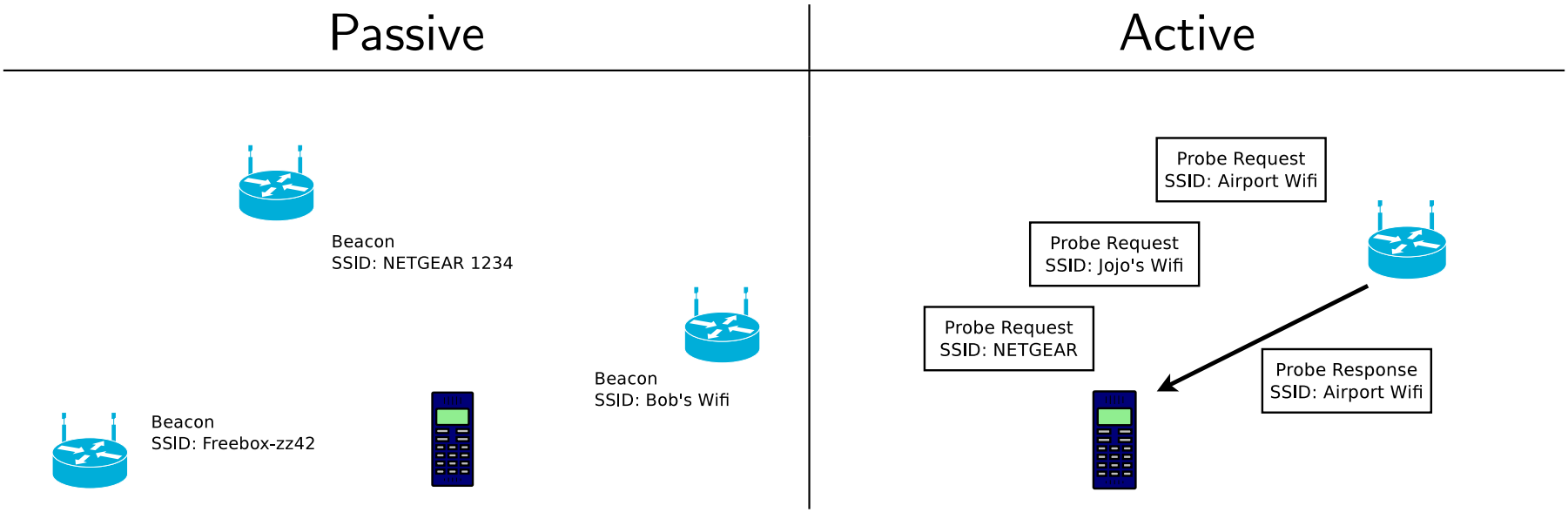
802.11 frame



- Address fields contain MAC addresses (src., dest., ...)
- MAC address: a unique identifier allocated to a network interface

Wi-Fi service discovery I

- Discover surrounding APs and Networks
  - Passive mode: Wi-Fi Beacons
  - Active mode: Probe requests and Probe Responses
  - Probe requests contain an SSID field to specify the searched network
- Active is less costly in energy
  - Preferred mode for mobile devices



Active service discovery

John's WiFi

Hotel d'Alsace

Lyon airport

Smith Family

SSID of configured networks

0C:56:4F:00:72:F5

MAC address

- Information available in cleartext (headers are not encrypted)
- Broadcasted: dest. Addr. = `FF:FF:FF:FF:FF:FF`

Active service discovery

- Probing frequency
  - Depends on model, OS version, ...
  - Several cycles per minutes (every 20/30 secs)

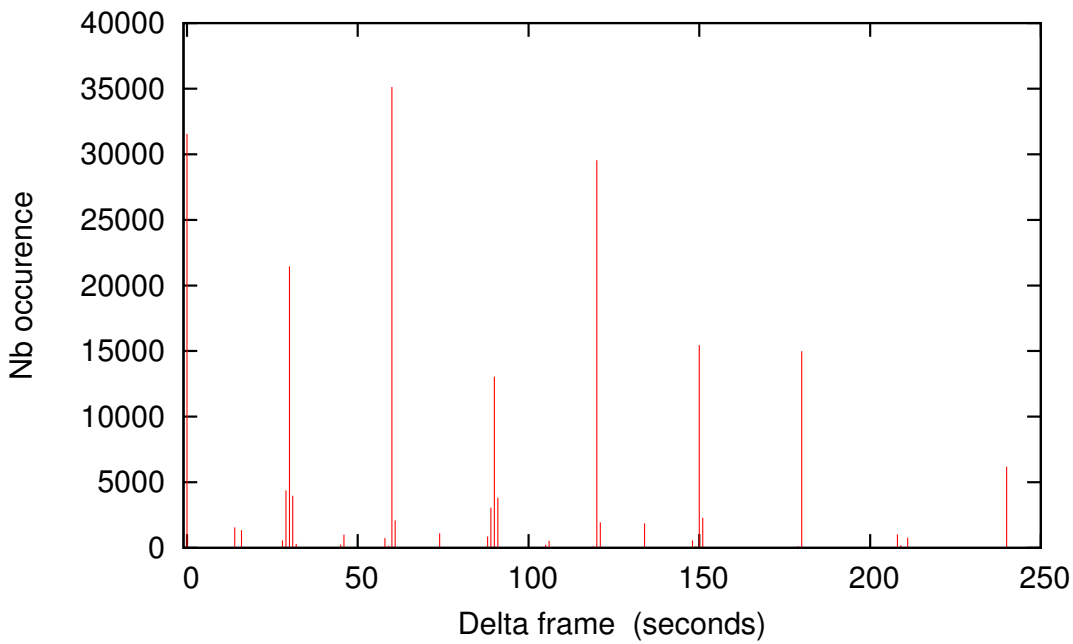


Figure: Delta between probes of a Samsung phone.

Wi-Fi Fingerprint

Source MAC Address	Destination MAC Address	Signal strength	SSID
00:24:d7:20:4e:45	ff:ff:ff:ff:ff:ff	-70	TECOM-AH4222-561ABC
00:24:d7:20:4e:45	ff:ff:ff:ff:ff:ff	-68	TP-LINK
00:24:d7:20:4e:45	ff:ff:ff:ff:ff:ff	-72	wireless
00:24:d7:20:4e:45	ff:ff:ff:ff:ff:ff	-80	ACCESS-StarHub
00:1f:3b:a2:be:39	ff:ff:ff:ff:ff:ff	-79	A-Company Ltd
00:1f:3b:a2:be:39	ff:ff:ff:ff:ff:ff	-75	Apple Store
00:1f:3b:a2:be:39	ff:ff:ff:ff:ff:ff	-79	dd-wrt
00:19:d2:64:5f:7f	ff:ff:ff:ff:ff:ff	-81	INRIA-guest
00:19:d2:64:5f:7f	ff:ff:ff:ff:ff:ff	-75	INRIA-grenoble
04:46:65:53:8d:ac	ff:ff:ff:ff:ff:ff	-78	A-Company Ltd
04:46:65:53:8d:ac	ff:ff:ff:ff:ff:ff	-77	McDonald's FREE WiFi
04:46:65:53:8d:ac	ff:ff:ff:ff:ff:ff	-74	Cafe_Bello
04:46:65:53:8d:ac	ff:ff:ff:ff:ff:ff	-59	Quality Inn
04:46:65:53:8d:ac	ff:ff:ff:ff:ff:ff	-45	BigPond9568

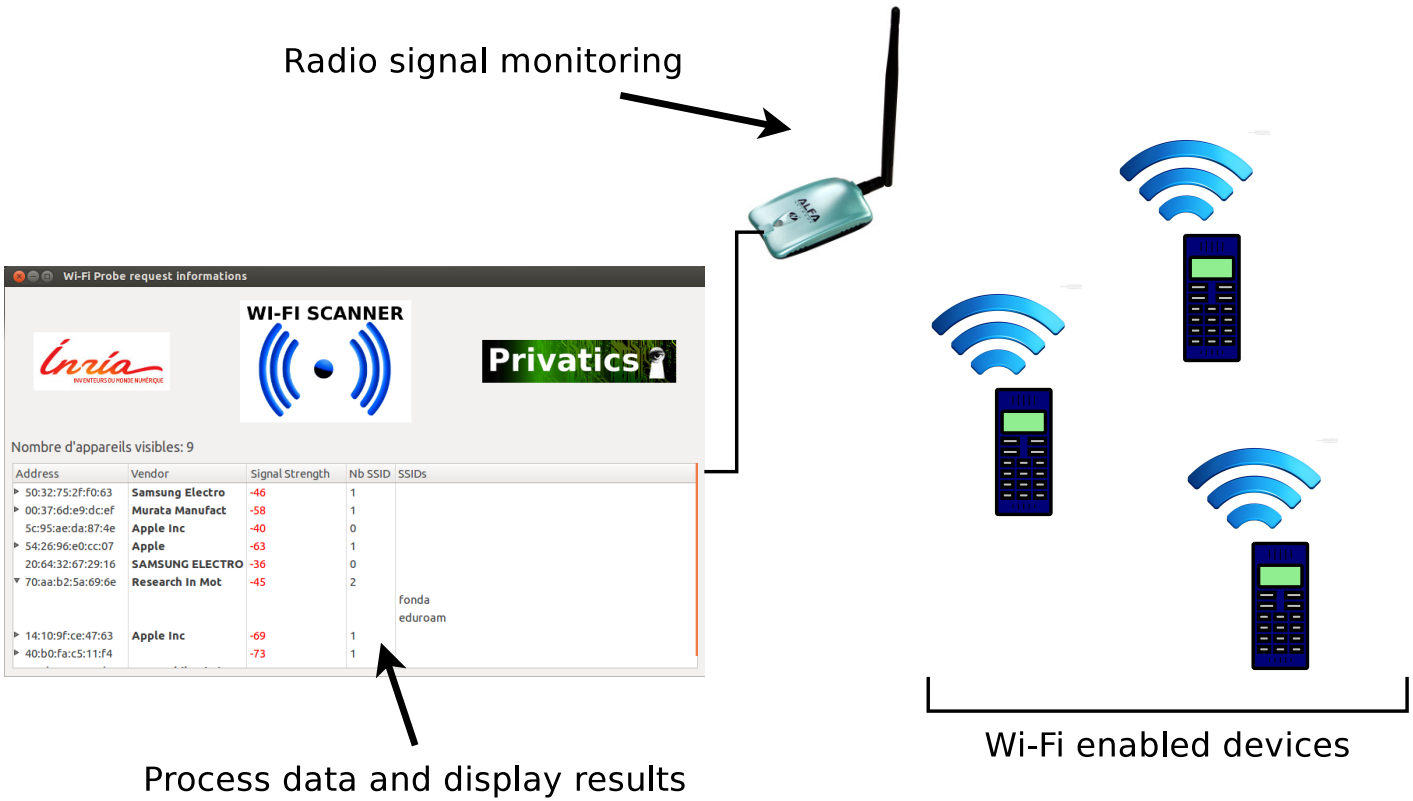
- Wi-Fi Fingerprint = List of SSIDs broadcasted by a device

Monitoring probe requests

- What about encryption (WPA,WPA2, ...)?
  - Only payload of DATA frames are is encrypted
  - Header are not encrypted
  - Management and Control frame are not encrypted (Probe Requests)

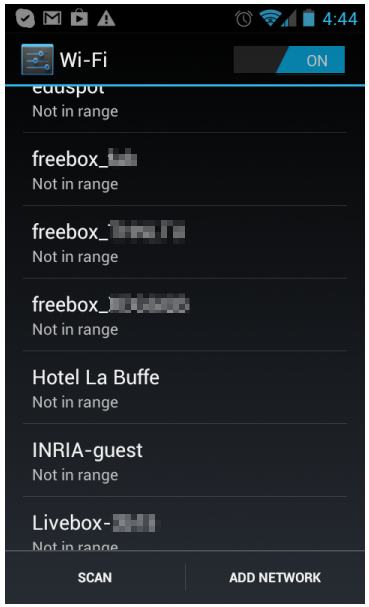
Monitoring probe requests (Demo.)

- Wi-Fi interface supporting monitoring mode
- Traffic capture and analysis tools



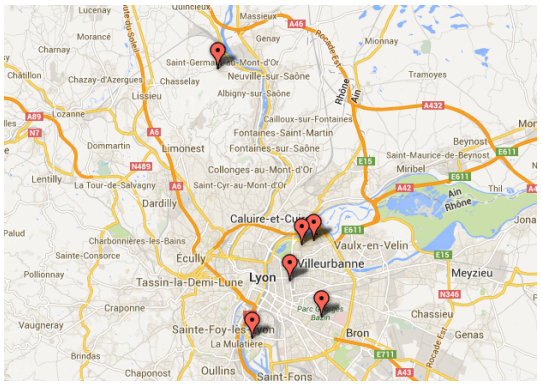
Personal information from SSIDs

- SSIDs: name of the previously connected networks
  - Stored in the Configured Network List (CNL)
  - Observed up to 80 configured networks !

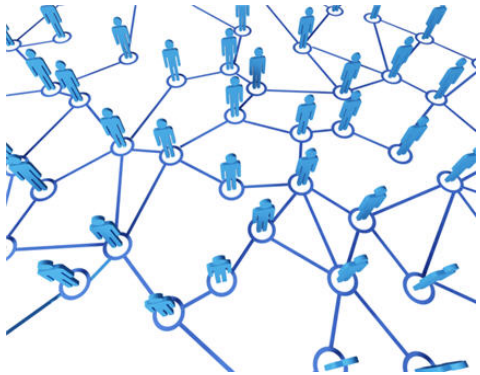


- SSIDs: personal data

Travel history



GPS coordinates

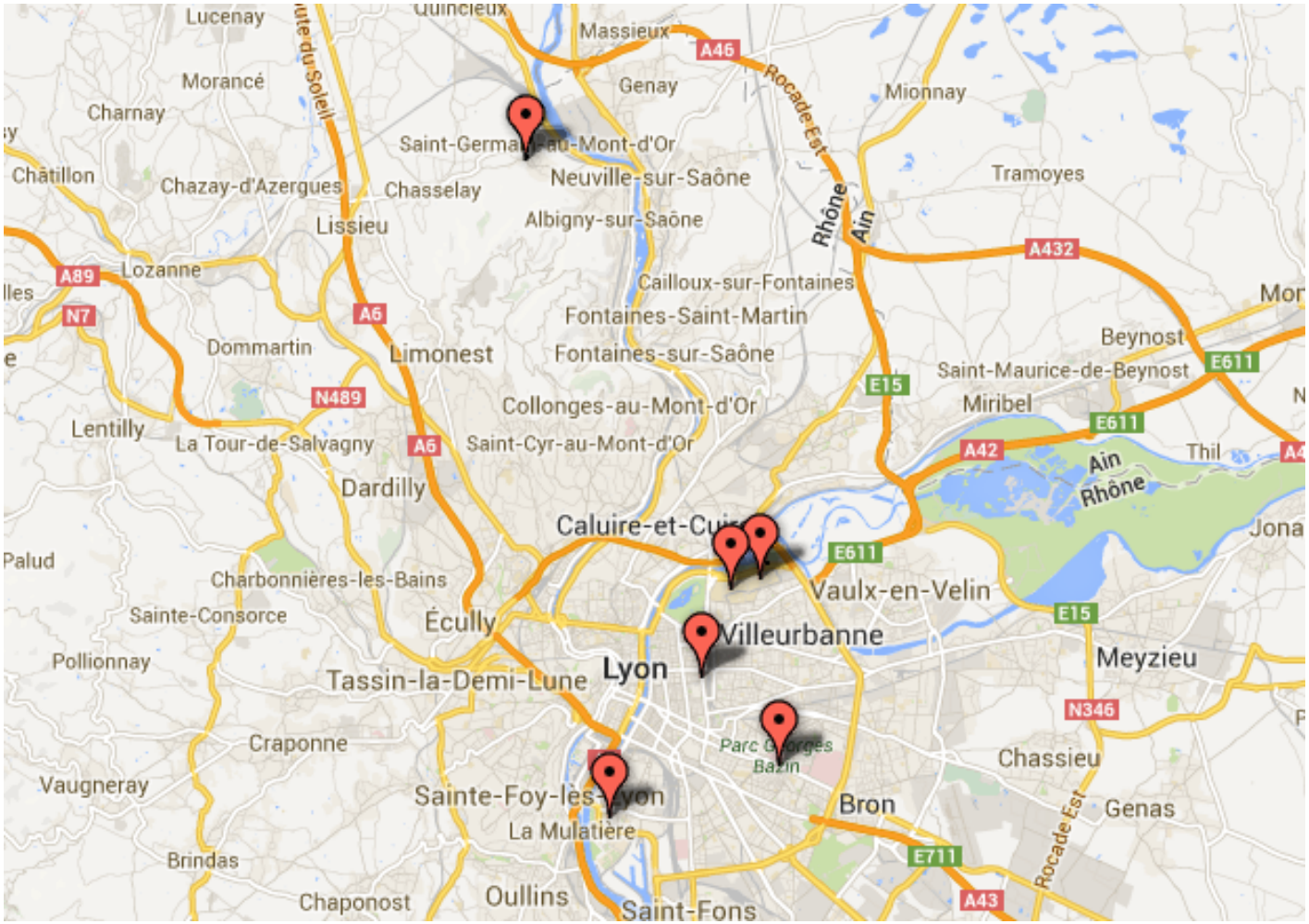


Social links

- Personal information found in SSIDs
  - Link with a company/university/organisation
    - INRIA-interne, INSA-INVITE, GlobalCorp Ltd.
  - Attended conferences
    - WiSec14, PETs, CCS
  - Visited places (hotel, restaurant, coffee-shop, airport)
    - Hilton-NY WiFi, Aloha Hotel WiFi, Brasserie de l'Est, Sydney-airport-WiFi
  - Individual's identity
    - Marc Dupont's iPhone, Bob Fhisher's Network

Precise geolocation information

- From SSIDs to precise geolocation[1]



[1] Ben Greenstein et al. "Can ferris bueller still have his day off? protecting privacy in an era of wireless devices". In: *In HotOS XI*. 2007.

Precise geolocation information

- WiGLE: *Wireless Geographic Logging Engine: Making maps of wireless networks since 2001*
  - SSID, BSSID, channel, security, GPS coordinates, ...

map it	netid	ssid	comment	name	type	freenet	paynet	firsttime	lasttime	flags	wep	trilat	trilong	lastupdt	channel	bcninterval	qos	userfound
Get Map	F4:CA:E5:84:00:28	Freebox-5F9191			infra	?	?	2013-12-26 11:46:39	2013-12-26 07:11:10		W	45.71317673	4.85502815	20131226071318	3		0	N
Get Map	F4:CA:E5:84:00:29	FreeWifi			infra	?	?	2013-12-26 11:46:37	2013-12-26 07:11:14		?	45.71334839	4.85501671	20131226071423	3		0	N
Get Map	F4:CA:E5:84:00:2A	FreeWifi_secure			infra	?	?	2013-12-26 11:46:39	2013-12-26 07:11:05		2	45.71317673	4.85502815	20131226071247	3		0	N
Get Map	F4:CA:E5:84:00:44	gaelle			infra	?	?	2014-01-06 11:31:32	2014-01-06 08:21:05		W	45.75151825	4.85475159	20140106082414	11		0	N
Get Map	F4:CA:E5:84:00:45	FreeWifi			infra	?	?	2014-01-06 11:31:32	2014-01-06 08:21:08		?	45.75151825	4.85475159	20140106082547	11		0	N
Get Map	F4:CA:E5:84:00:46	FreeWifi_secure			infra	?	?	2014-01-06 11:31:32	2014-01-06 08:21:03		2	45.75151825	4.85475159	20140106082517	11		0	N
Get Map	F4:CA:E5:84:00:88	freebox_alsica			infra	?	?	2013-07-13 16:34:42	2013-07-14 06:45:52		W	48.10209274	-1.68101156	20130714064705	11		0	N
Get Map	F4:CA:E5:84:00:8A	FreeWifi_secure			infra	?	?	2013-07-13 16:34:42	2013-07-14 06:45:49		2	48.10209274	-1.68101156	20130714064707	11		0	N
Get Map	F4:CA:E5:84:00:94	freebox_OJAAMJ			infra	?	?	2013-04-28 23:04:20	2013-10-07 11:45:34		W	47.90141678	1.93103909	20131007114629	11		3	N

- Other databases exist (CIA, Google, Apple, ...)

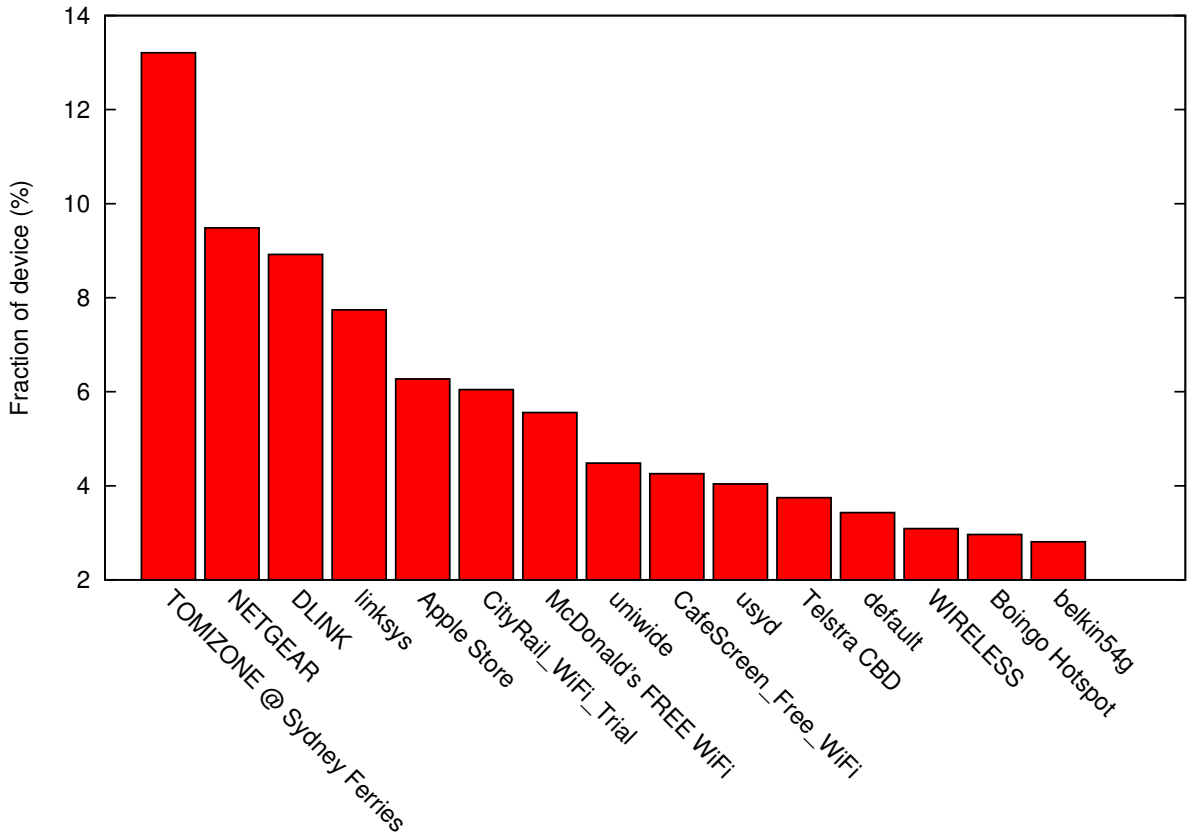
Inferring social links I

- Hypothesis: similarity between Wi-Fi fingerprint can betray social links
  - People tends to share their Wi-Fi network with people who are close
- The experiment: "I know who you will meet this evening"[2]
  - A wild dataset: fingerprints of 8000+ devices
  - A control dataset: fingerprint with 30 existing social links

[2] Mathieu Cunche, Mohamed-Ali Kaafar, and Roksana Boreli. "Linking wireless devices using information contained in Wi-Fi probe requests". In: *Pervasive and Mobile Computing* (2013), pp. –.

Inferring social links I

- Frequency of SSIDs
  - Some are frequent (ex. NETGEAR)
  - Other are rare (ex. Freebox_YTC689)



Inferring social links I

- Quantifying the similarity between fingerprints
  - Metric considering size and rarity of the intersection
- Cosine-IDF and Jaccard index

$$\text{Cosine-idf}(X,Y) = \frac{\sum_{x \in X \cap Y} \text{idf}_x{}^2}{\sqrt{\sum_{x \in X} \text{idf}_x{}^2}\sqrt{\sum_{y \in Y} \text{idf}_y{}^2}} \qquad \text{J}(X,Y) = \frac{|X \cap Y|}{|X \cup Y|}$$

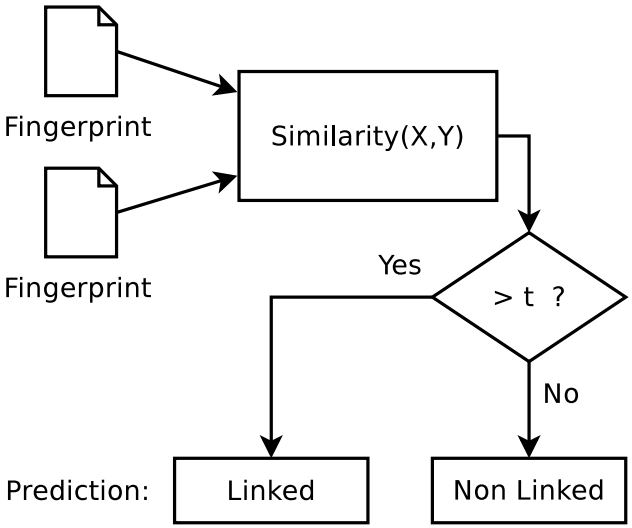
where idf_X : inverse document frequency of x

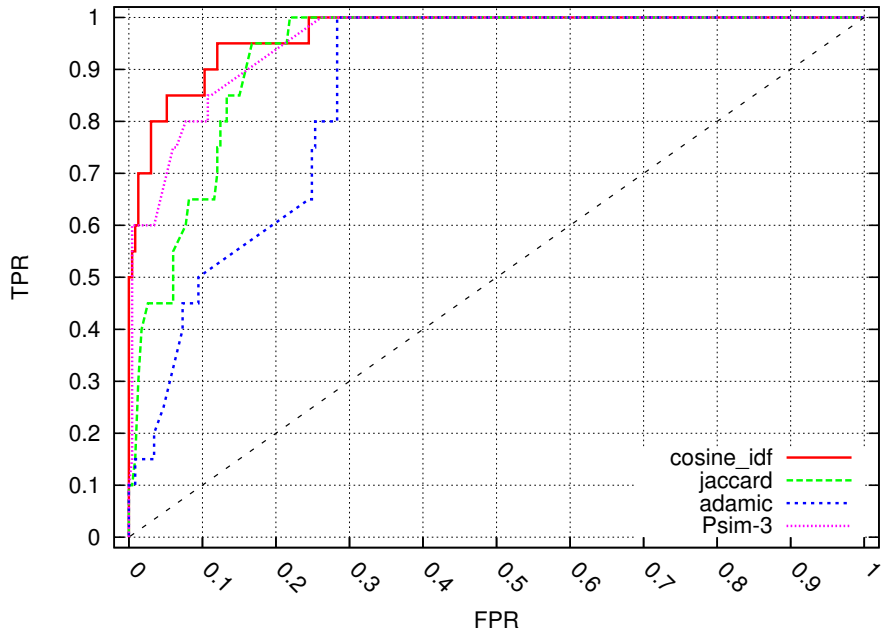
- Adamic, modified Adamic

$$\text{Adamic}(X,Y) = \sum_{x \in X \cap Y} \frac{1}{\log f_x} \qquad \text{Psim-}q(X,Y) = \sum_{x \in X \cap Y} \frac{1}{f_x^q}$$

where f_X : document frequency of x

Inferring social links I





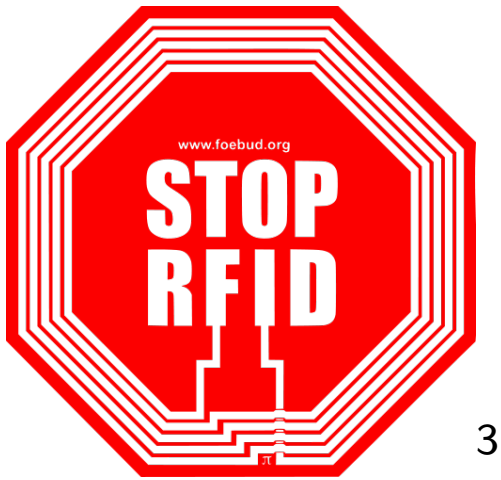
- Performances: detects 80% of social links with less than 8% of error.

The end of broadcasted SSIDs ?

- The good news: *Broadcast* Probe Requests
  - SSID field is left empty
  - AP must responds to all Broadcast Probe Requests
  - Adopted by major vendors to reduce privacy risks
- The bad news: Hidden Wi-Fi networks
  - Hidden: not broadcasting beacons
  - Probing with SSID is the only way to discover
  - Device continuously broadcast SSID of the network

- A short parenthesis on RFID
  - Privacy concerns over RFID
  - Chip embedded in goods (clothes)
  - A combination of RFIDs can constitute a unique ID

"How would you like it if, for instance, one day you realized your underwear was reporting on your whereabouts?"
– US Senator Bowen on RFID chips. 2003.

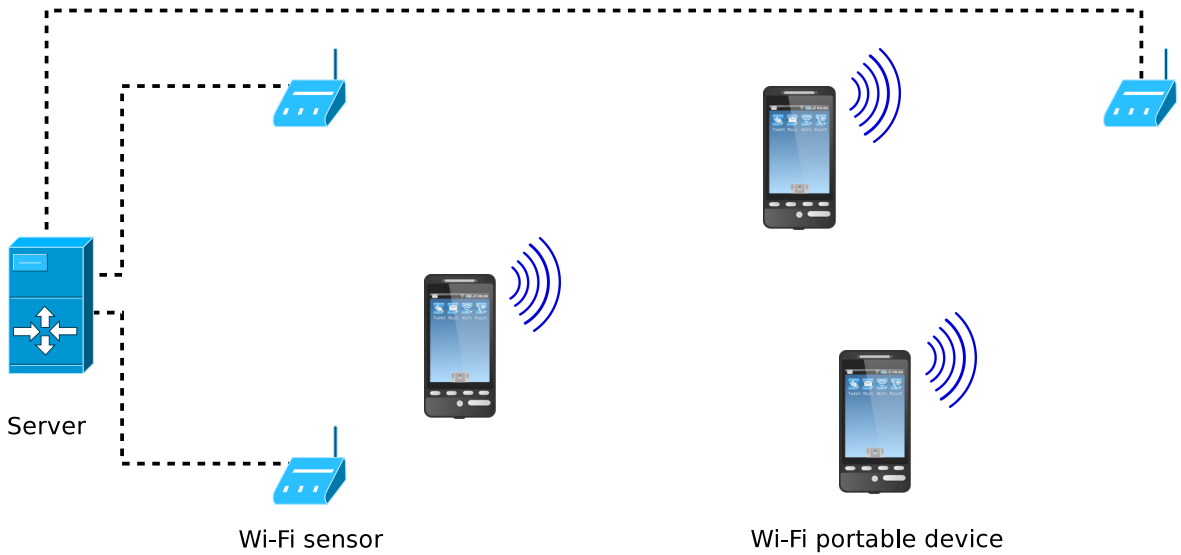


3

[3] `http://digitalcourage.de/`

Wi-Fi tracking

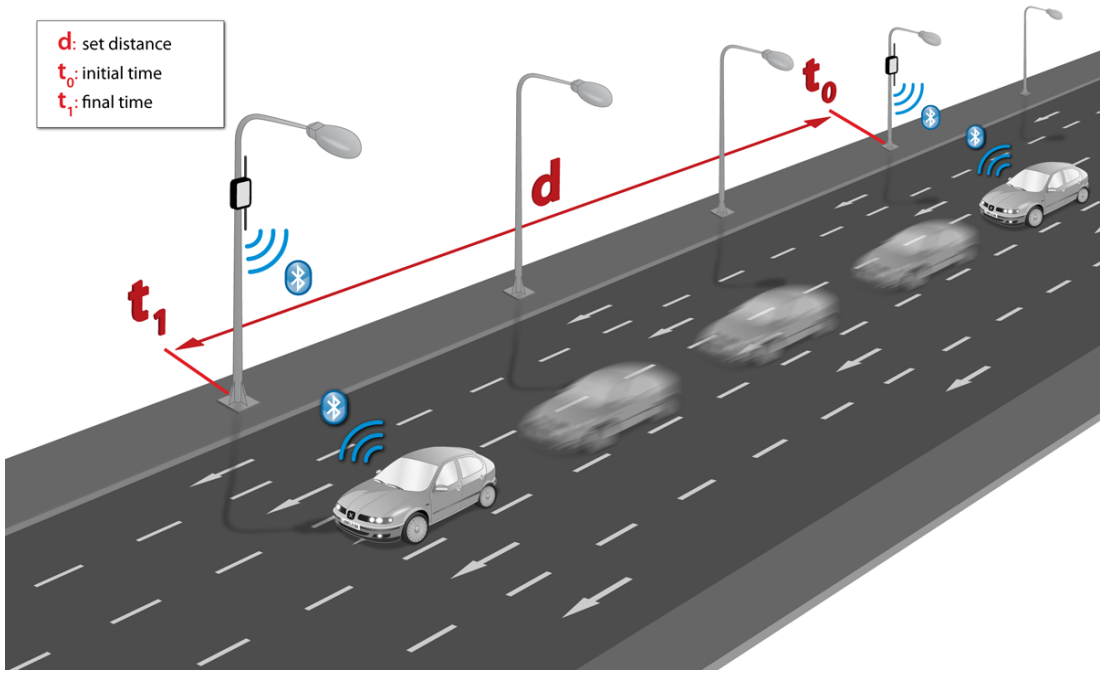
- Wi-Fi enabled smartphone: portable personal beacon
  - Broadcast a unique ID
  - Several 10s meters range
- Wi-Fi tracking system[4]
  - Set of sensors collect Wi-Fi signal
  - Detect and track Wi-Fi devices and their owners



[4] A. B. M. Musa and Jakob Eriksson. "Tracking unmodified smartphones using Wi-Fi monitors". In: *Proceedings of the 10th ACM Conference on Embedded Network Sensor Systems*. 2012.

Wi-Fi tracking: applications

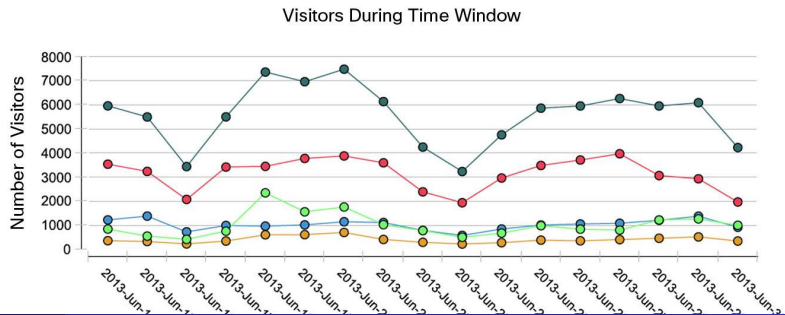
- Road monitoring
  - Measure point-to-point travel time
  - Detect traffic jam



Wi-Fi tracking: applications

- Retail, shopping center monitoring

- Physical analytics
  - Similar to Web Analytics
  - Frequency and length of visit, number of visitor, peak hour ....

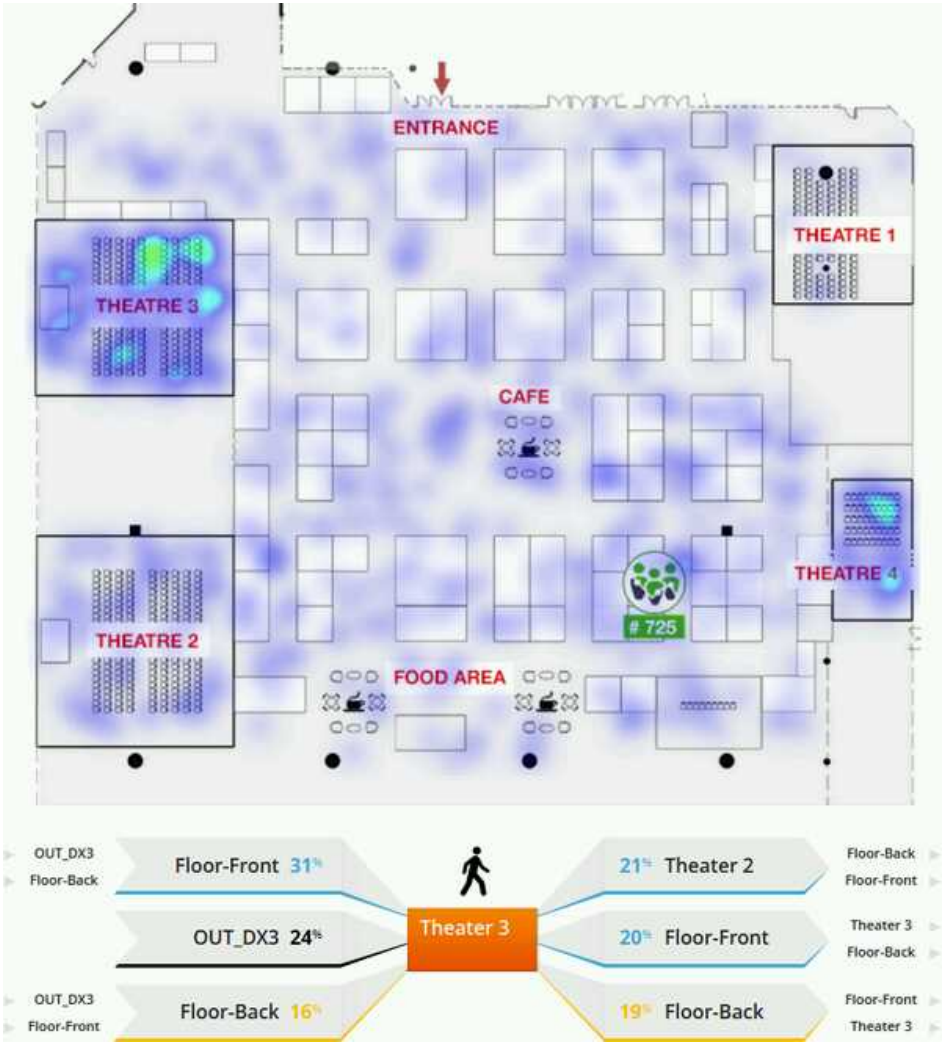


Wi-Fi tracking: applications

- Trajectory reconstruction
  - Triangulation based on signal strength

Wi-Fi tracking: applications

- Illustration: monitoring Dx3 2014[5]



[5] Credits: Aislelabs

Wi-Fi tracking: applications

- Current state of Wi-Fi tracking (in the US)
  - More than 12 tracker companies: Euclid, Navizon, ...
  - Major retailers are getting involved
  - 50 millions individual tracked by Euclid in less than 5 months of activity

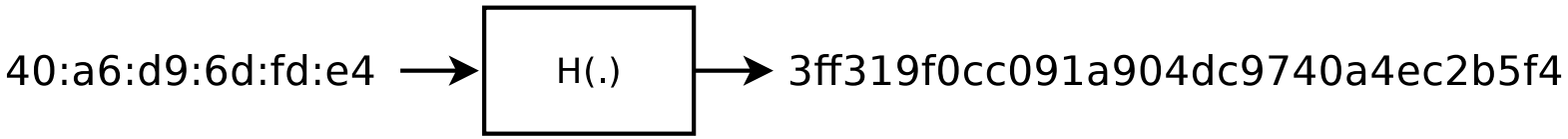
Wi-Fi tracking: privacy

- Privacy concerns

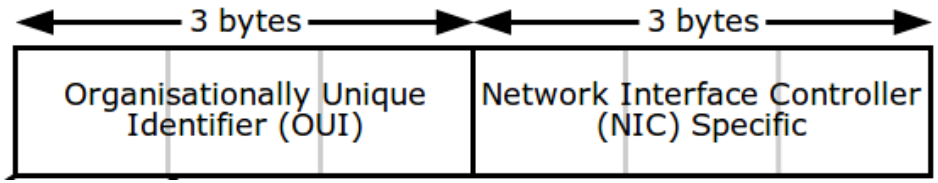
"People have a fundamental right to privacy, and I think neglecting to ask consumers for their permission to track them violates that right" – Senator Al Franken

Wi-Fi tracking: privacy

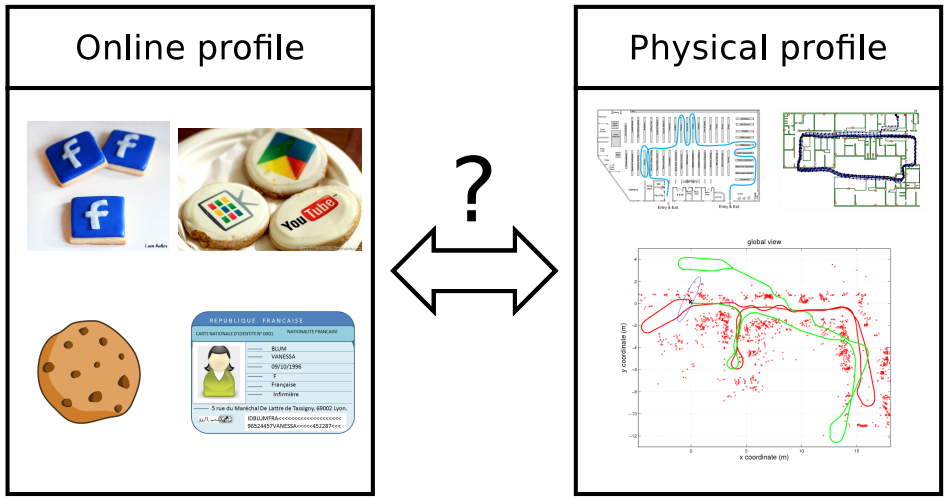
- Response to privacy concerns
  - MAC addr. does not contain personal information
  - User notification
  - Opt-out mechanisms
  - MAC addr. is "anonymized" (Hash function)



Wi-Fi tracking: privacy

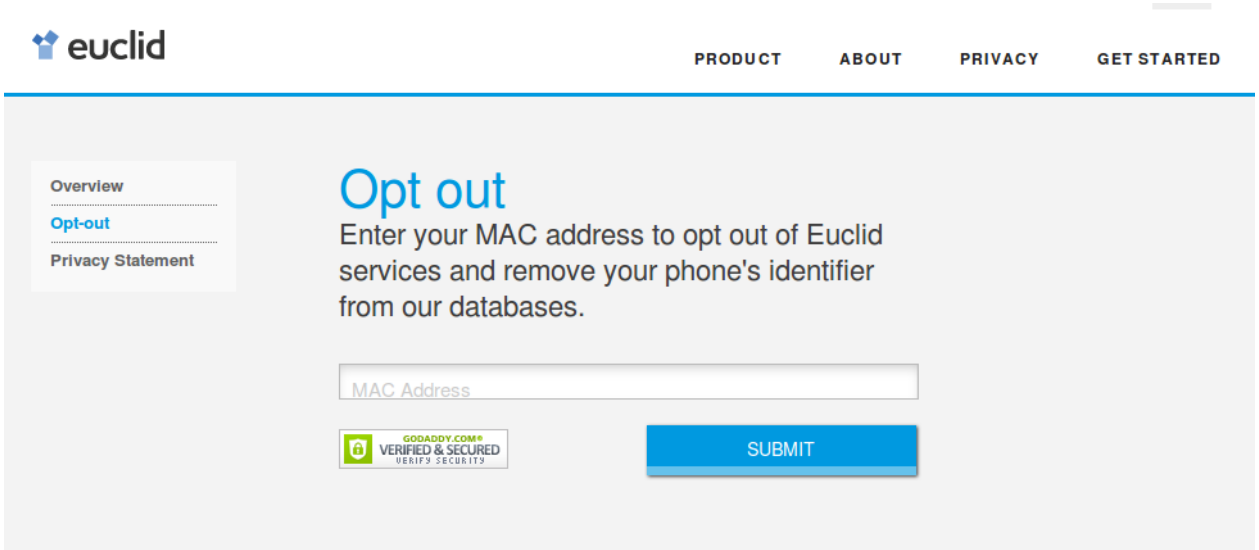


- The MAC address: not a personal information ?
  - Unique identifier
  - Collected by mobile applications
  - The missing link between physical and online profile



Wi-Fi tracking: privacy

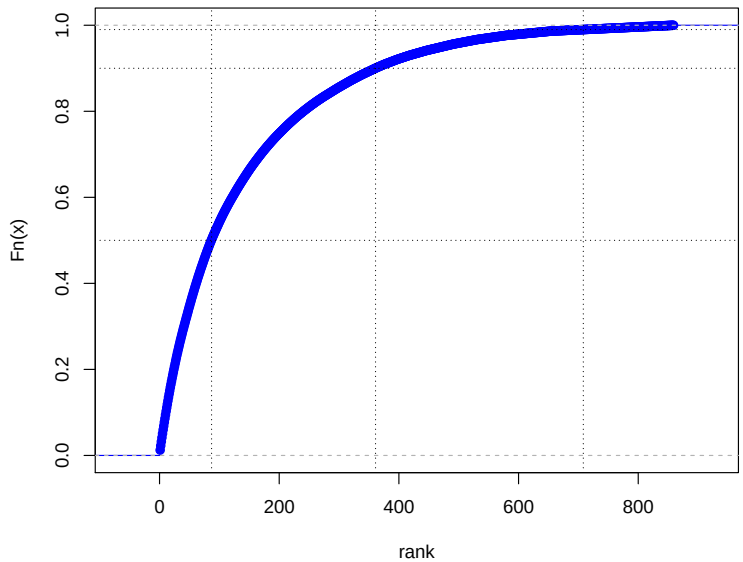
- Opt-Out mechanism
  - By default your activity is recorded
  - Opt-Out service: enter your MAC address



- Is this a good idea to give out your MAC addr. ?
- Can your grand-mother find her MAC addr. ?

Wi-Fi tracking: privacy

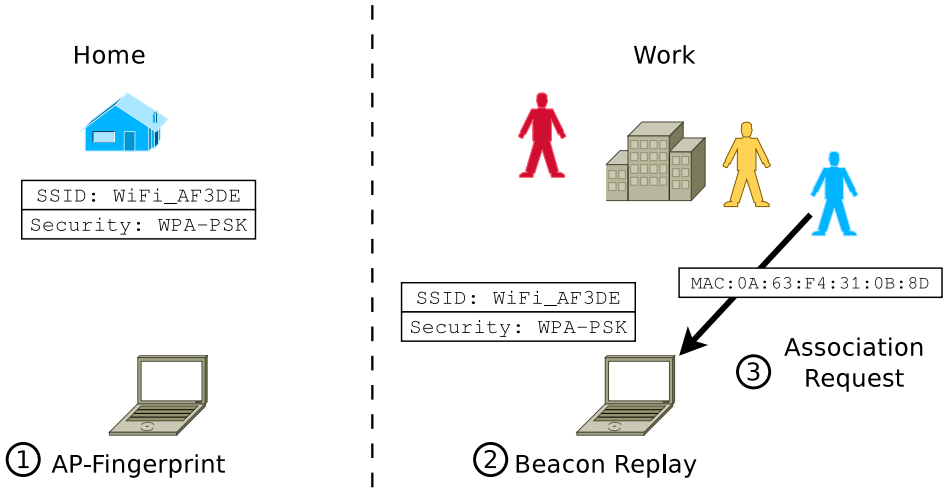
- The failure of hash-based anonymization of MAC addr.
  - Trackers: Don't worry we don't store the MAC in clear
  - *Irreversible* operation: hashing, scrambling, ...
  - Can be de-anonymized[6]
    - In $\sim$ 1 day using high-end GPU
    - In a handful of seconds by exploiting skewed MAC addr. distribution.



[6] Mathieu Cunche, Levent Demir, and Cédric Lauradoux. "Anomymization for Small Domains: the case of MAC address". In: *Atelier sur la Protection de la Vie Privée - APVP 2013*. June 2013.

Wi-Fi tracking: privacy

- How to obtain the MAC addr. of an individual ?
  - Without a physical access
- Beacon replay attack
  - Home/work locations uniqueness

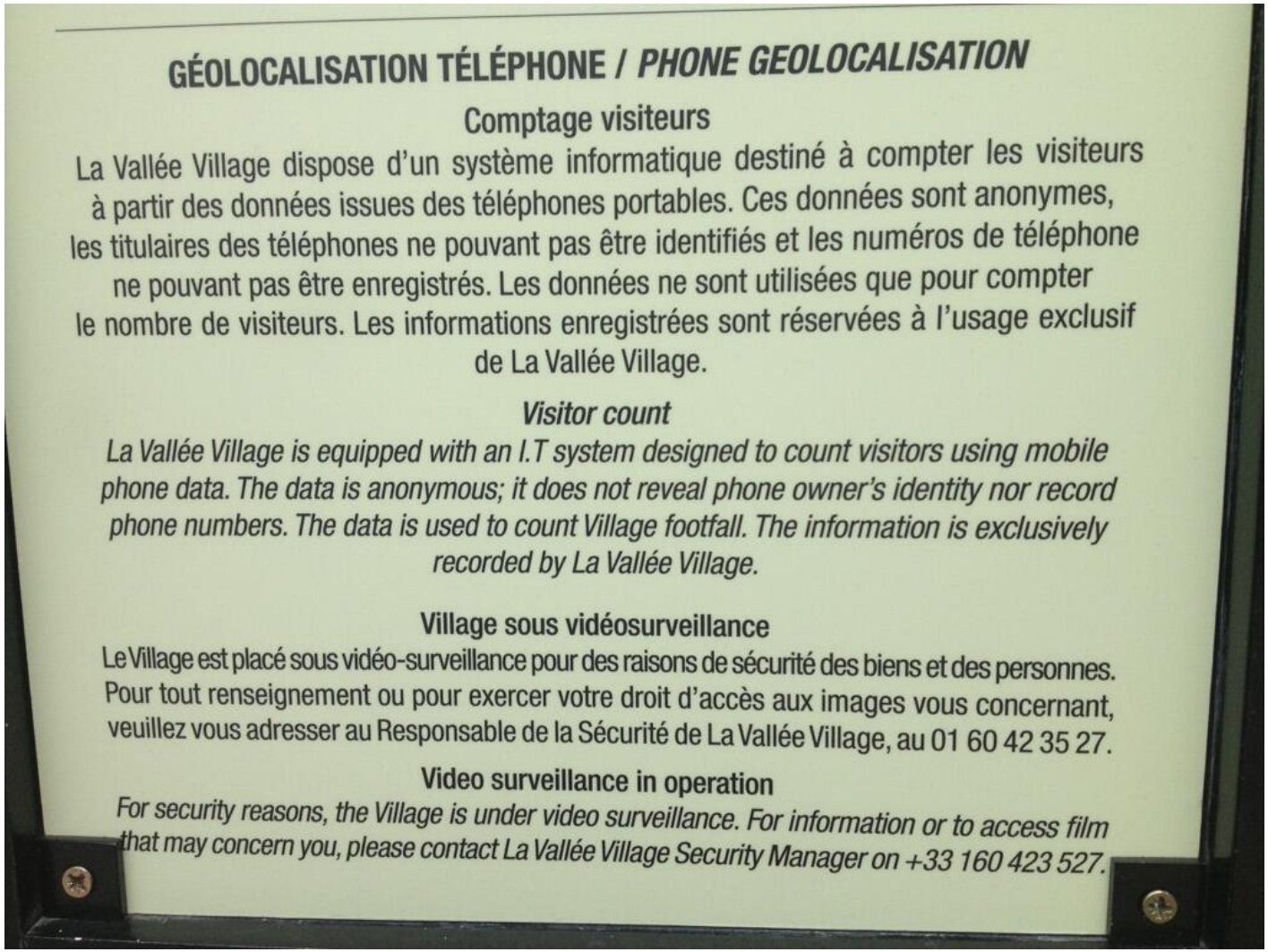


- *Stalker* attack
  - Simply follow the target in the streets

Wi-Fi tracking

- London's Wi-Fi bins
  - Detect individuals via Wi-Fi
  - Display targeted advertisement on screen
  - Based on a user profile: consuming habits, gender, ...

Wi-Fi tracking



GÉOLOCALISATION TÉLÉPHONE / *PHONE GEOLOCALISATION*

Comptage visiteurs

La Vallée Village dispose d'un système informatique destiné à compter les visiteurs à partir des données issues des téléphones portables. Ces données sont anonymes, les titulaires des téléphones ne pouvant pas être identifiés et les numéros de téléphone ne pouvant pas être enregistrés. Les données ne sont utilisées que pour compter le nombre de visiteurs. Les informations enregistrées sont réservées à l'usage exclusif de La Vallée Village.

Visitor count

La Vallée Village is equipped with an I.T system designed to count visitors using mobile phone data. The data is anonymous; it does not reveal phone owner's identity nor record phone numbers. The data is used to count Village footfall. The information is exclusively recorded by La Vallée Village.

Village sous vidéosurveillance

Le Village est placé sous vidéo-surveillance pour des raisons de sécurité des biens et des personnes. Pour tout renseignement ou pour exercer votre droit d'accès aux images vous concernant, veuillez vous adresser au Responsable de la Sécurité de La Vallée Village, au 01 60 42 35 27.

Video surveillance in operation

For security reasons, the Village is under video surveillance. For information or to access film that may concern you, please contact La Vallée Village Security Manager on +33 160 423 527.

Figure: Seen at *La Vallée Village* shopping center (near Paris).

Wi-Fi tracking: privacy

- Other field of application[7]
  - Surveillance: CIA, NSA, GCHQ, ...
  - Surveillance: private (stalkers)
  - Triggered "events"

CREEPYDOL

[7] Mathieu Cunche. "I know your MAC Address: Targeted tracking of individual using Wi-Fi". In: *International Symposium on Research in Grey-Hat Hacking - GreHack*. Grenoble, France, Nov. 2013.

Countermeasures (short-term)

- Geofencing
  - Wi-Fi only activated in trusted places (home, office, ...)
  - Apps: Wi-Fi Matic[8] and AVG Privacy Fix[9]

- MAC address Spoofing
  - Periodically change MAC address to a random value


10

[8] https://play.google.com/store/apps/details?id=org.cprados.wificellmanager

Countermeasures (long-term)

- Significant modification of the 802.11 protocols[11]
  - Encrypt/obfuscate all identifiers in the 802.11 protocol
  - Issues with retro-compatibility
  - Not before several years (decades ?)

[11] Janne Lindqvist et al. "Privacy-preserving 802.11 access-point discovery". In: WiSec '09. 2009.

Questions ?

Figure: Artist's interpretation[12].

[12] credit P. Treimany