Sécurité
Mécanismes & Vie privée

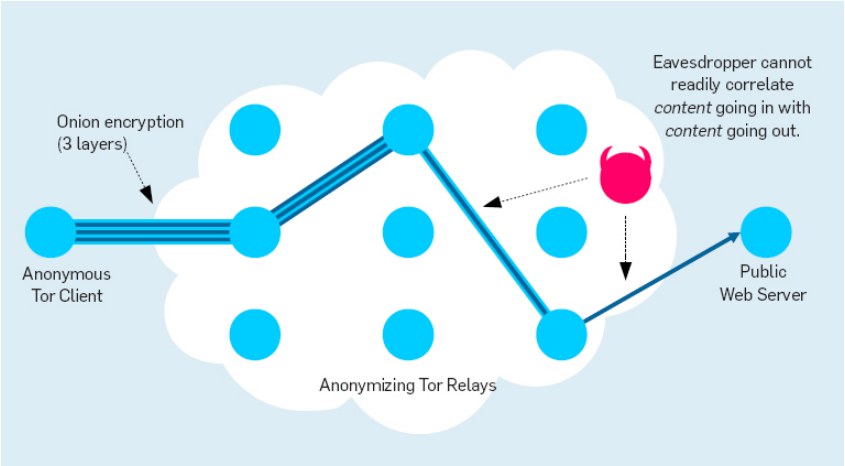
Bruno Martin

Université Côte d'Azur

M1 Informatique & Interactions

Plan

Onion routing/TOR



Onion routing/TOR

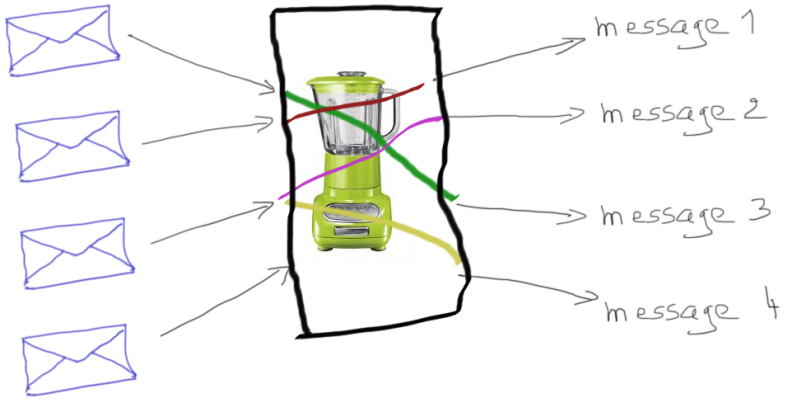
Sur Internet (public), les en-têtes des paquets identifient les parties qui communiquent.
Même en chiffrant la charge utile des paquets, les en-têtes sont accessibles et ne masquent pas l'information de routage.
Cette information fait partie de la vie privée :

- qui communique avec qui ?
- quels sites web regardez-vous ?
- où travaillez-vous, où habitez-vous ?
- vous achats, vos médecins, vos loisirs ?

Deux approches pour anonymiser les communications :

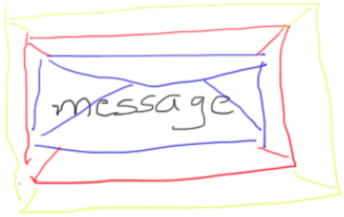
- mixees / mélangeur
- proxies / serveur mandataire

Mélangeur



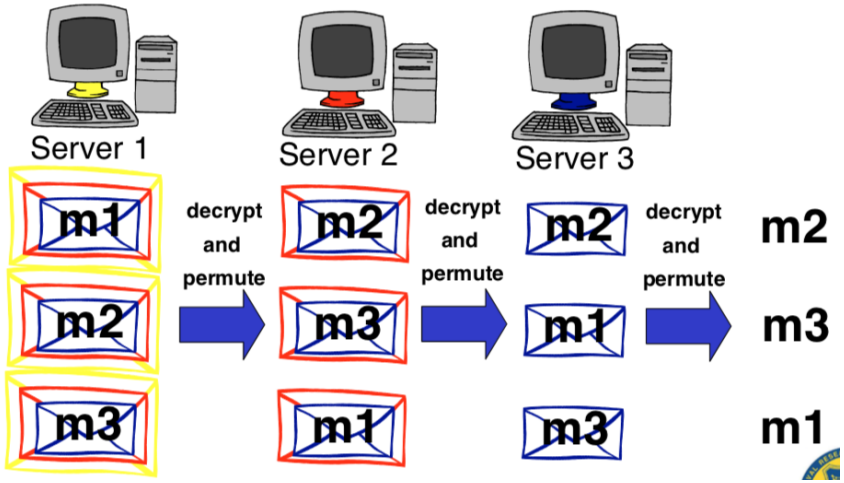
Déchiffre et permute les entrées.
Propriété principale : un adversaire ne peut pas associer un chiffré (une enveloppe) à un des messages.

Chiffrement d'un message



Avec Chiffré=$\{\{\{\text{message}\}_{PK3}\}_{PK2}\}_{PK1}$

Mélangeur de Chaum de base



Un seul serveur honnête préserve l'anonymat.

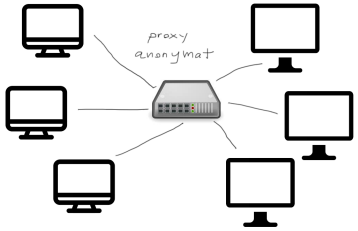
Inconvénients des mélangeurs...

Tout ce qui demande une interaction rapide :

- navigation web, connexion distante, chat, etc.
- mélangeurs introduits pour le mail et d'autres applications de latence élevée
- chaque enveloppe autour du message demande beaucoup de cryptographie à clé publique

Serveur mandataire d'anonymat

- liens semblent venir du proxy, pas de l'émetteur
- approprié pour communications à faible latence
- chiffrement symétrique
- **avantage** : simple ; permet d'anonymiser bcp de trafic
- **inconvénient** : un seul point de panne, compromission, attaque



Routage en oignon

But : Construire une infrastructure robuste à l'analyse de trafic.

- Combiner les avantages des mélangeurs et des serveurs mandataires
- Utiliser PKC (coûteuse) pour établir des routes
- Utiliser crypto symétrique pour échanger des données
  - analogue à des proxies basés sur TLS
- Confiance distribuée par les mélangeurs
- Pas mal de travaux, implémentation, POC
  - ISDN mixes
  - Crowds, JAP webmixes, Freedom net
  - Tarzan, Morphmix
- https://www.torproject.org/docs/onion-services.html.en
- https://gitweb.torproject.org/torspec.git/tree/tor-spec.txt
- https://www.onion-router.net/Publications/tor-design.pdf

Plan

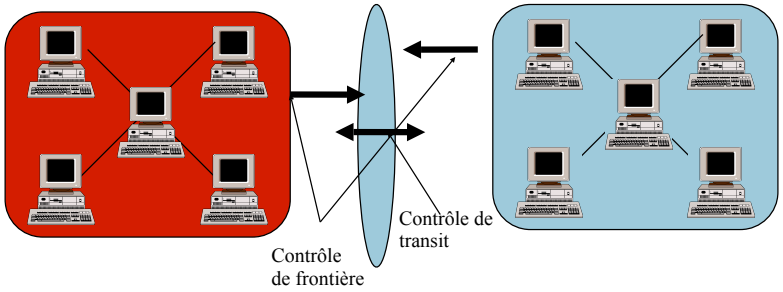
Contrôle des frontières

- Problème principal des LAN connectés à Internet
- **Solution** : utiliser des coupe-feux (firewall) avec :
  - filtres de paquets
  - proxy (serveur mandataire)
  - mécanismes cryptographiques
- tout en diminuant le nombre de points d'entrée

Filtres de paquets

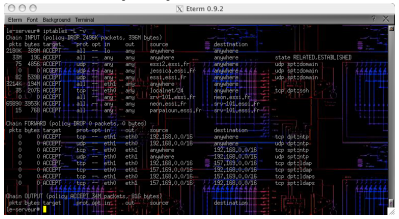
- Fonction assurée par routeurs ou hôtes dédiés.
- Principe du contrôle :
  - redistribuer, effacer et/ou tracer chaque paquet
  - selon sur les informations d'en-tête des paquets
    - adresse source et destination
    - direction (entrée/sortie) par rapport au LAN
    - type d'application par numéro de port
  - en conjonction avec TCP/IP
  - généralement au niveau du noyau de l'OS

Filtrage de paquets

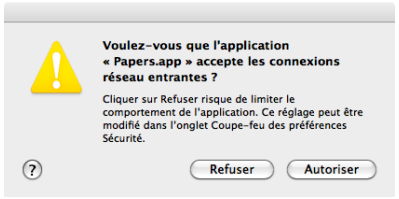


Différents types de pare-feu

Filtre de paquets



Filtrage applicatif



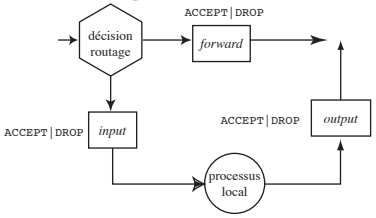
Fonction d'un pare-feu
Un pare-feu désigne un logiciel et/ou un matériel (appliance), qui a pour fonction de faire respecter la politique de sécurité du réseau. Celle-ci définit quels sont les types de communications autorisés ou interdits.

Différents selon l'OS

- **Linux** : IPtables/netfilter/nftables/ufw : firewall de paquets disponible sous linux. Composé de règles de filtrage
  - chaîne : liste ordonnée de règles
  - chaque règle exprime une condition
  - si règle i ne s'applique pas, consulter règle $i + 1$
  - une fois épuisé l'ensemble des règles, appliquer la politique par défaut de la chaîne (`ACCEPT`,`DROP`)
- **BSD** : 3 mécanismes différents :
  - **IPFILTER** : commande `ipf`
  - **IPFIREWALL** : commande `ipfw`
  - **PacketFilter** : commande `pf`
- **Windows** : Windows defender (ou de tierce partie)
- **MacOS** : 2 mécanismes intégrés, un provenant de BSD et un firewall applicatif.

IPtables – chaînes prédéfinies

`iptables` filtre les paquets qui traversent une machine au moyen des chaînes : `INPUT`, `OUTPUT` et `FORWARD`. Le noyau examine le paquet entrant et :



- Destiné à la machine, il traverse la chaîne `INPUT`. S'il est autorisé à poursuivre son chemin (par un `ACCEPT`), il est traité par le processus local auquel il est destiné. Si la décision est `DROP`, le paquet est supprimé.
- Destiné à une autre interface, le paquet traverse la chaîne `FORWARD` et, s'il est accepté, il poursuit son chemin. Si le `forwarding` n'est pas activé ou si on ne sait pas comment transmettre ce paquet, le paquet est supprimé.
- Un processus local exécuté par la machine peut également envoyer des paquets traités par la chaîne `OUTPUT`.

UFW

Les distributions après 2016 basées sur `Debian` intègrent Uncomplicated Firewall, une surcouche à `iptables` ou `nftables`. Fonctionne avec le même principe que `iptables`. Spécifier une politique par défaut :

```
# ufw default deny incoming
# ufw default allow outgoing
```

Autoriser l'accès à un service (`ssh`, 22) (refuser par `deny`)

```
# ufw allow ssh
```

Liste des règles actives : `# ufw status numbered`

```
Status: active
     To                         Action      From
     -                          ---      --
[ 1] 22                         ALLOW IN    Anywhere
```

Supprimer une règle par son numéro : `# ufw delete 1`

Limites d'un pare-feu

- toutes les communications doivent passer par le pare-feu
- le pare-feu doit être convenablement configuré
- éviter le contournement (modem, gsm...)
- éviter l'utilisation de clés usb, ordinateurs portables
- tenir un journal (logs)
- détecter les anomalies et/ou les intrusions

Plan

Qu’est-ce qu’un virus ?

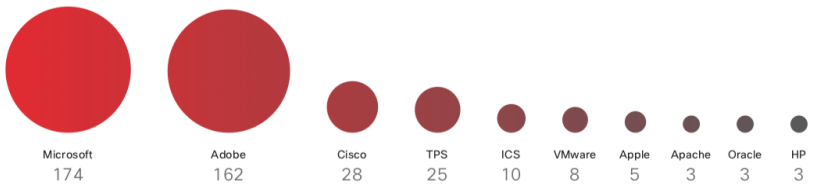
Un automate autoréplicatif à la base non malicieux, mais aujourd’hui souvent additionné de code malveillant, conçu pour se propager à d’autres ordinateurs en s’insérant dans des logiciels légitimes, appelés « hôtes ». Il perturbe plus ou moins gravement le fonctionnement de l’ordinateur infecté. Il peut se répandre par tout moyen d’échange de données numériques comme les mails les réseaux informatiques et les cédéroms, les clefs USB, les disques durs, etc.

Programmes souvent assimilés aux virus : le cheval de troie (qui héberge souvent des ransomware), les vers (qui n’infectent pas les fichiers mais se propagent aux autres ordinateurs)

Programme anti-virus

Protège l’ordinateur contre les virus et programmes assimilés. Ces programmes contiennent un catalogue de miliers de virus connus qu’ils peuvent détecter et éradiquer.
Ils détectent les modifications standard apportées au fichiers par les virus pour rechercher de nouveaux virus.
On trouve des programmes anti-virus gratuits comme Avast ou ClamAV, proposés par l’OS comme Windows defender ou proposés comme applications tierces payantes par Symantec, McAfee, Norton,...

Cibles favorites



Bonne pratiques

En plus d’avoir un bon logiciel antivirus, il faut ajouter de bonnes pratiques :

- sauvegardes fréquentes : en cas d’attaque fructueuse, récupérer son travail
- installer des logiciels d’origine : ne pas faire l’économie du prix d’un logiciel et être sûr de sa source
- scanner son ordinateur fréquemment, surtout après le passage d’un réparateur ou d’un consultant
- ne jamais ouvrir les pièces jointes à un mail ou un lien sur un réseau social (mail, .doc, image, ...)
- vérifier tout ce qui est connecté à un système (physiquement ou logiquement)

Plan

Services
- Onion routing
- Firewalls
- Anti-virus
- Spam
- Détection d'intrusion

Vie Privée
- RGPD

Web tracking

Qu'est-ce que le spam ?

Du mail non sollicité, généralement pour vendre un produit, parfois sans ciblage, parfois avec (voir vie privée). Le spam est une généralisation des courriers (papier) publicitaires, surtout à cause de leur coût quasi nul pour inonder les destinataires.
Le spam a un coût :

- pour les destinataires : il faut trier les bons mails des mauvais
- pour les fournisseurs d'accès : les spams sont transportés, acheminés et traités comme des mails "légitimes" avant d'être filtrés in fine avant leur (non)-distribution.

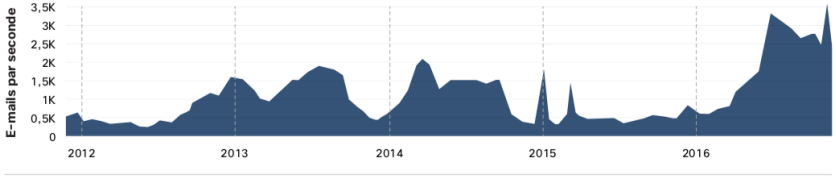
Une estimation : 75% des mails acheminés sont du spam.
Le spam présente aussi des risques car ils acheminent souvent des virus ou des liens vers des logiciels malicieux.

Catégories de spams

- **Publicitaires :** provenant de companies inconnues qui proposent des produits qui ne vous concernent pas.
- **Phishing :** plus dangereux, le phishing vous invite à divulguer des informations privées (numéro de carte de crédit, informations de connexion à des réseaux sociaux, des sites,..). Cela peut être ravageur, même avec des utilisateurs avertis
- **scams :** promettent de vous offrir quelque chose (de l'argent, un prix, une reconnaissance). Vous pensez qu'on peut gagner quelque chose sur Internet ?

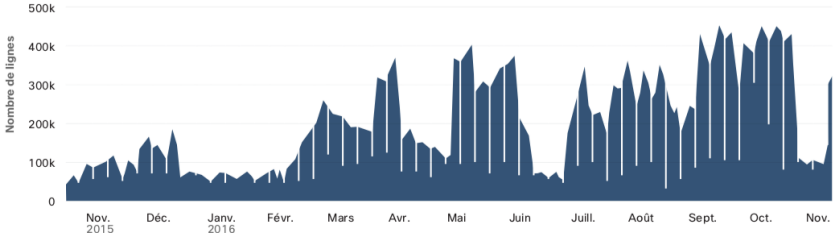
Spam en chiffres

Figure 15 Volume total de spams



Source : CBL

Figure 16 Taille de la liste de blocage SCBL (SpamCop Blocking List)



Lutte contre le spam

3 grandes catégories de lociels :

- **filtre anti-spam :** fonctionne en interaction avec le serveur mail de l'entreprise. A son arrivée le mail est filtré, le mail légitime délivré et le mail douteux rangé dans un répertoire spécifique ou effacé
- **boitiers anti-spam :** proposés par des constructeurs, sous la forme d'un boitier. Aucune configuration, tout est fait par le fournisseur qui met à jour des listes noires, grises, d'IP douteuses, filtre avec des règles plus complexes. Cher.
- **service dématérialisé dans le cloud**. Le mail est récupéré par le service dédié, filtré et redistribué à l'utilisateur. Fonctionne selon le principe de l'abonnement.

Plan

Détection d'intrusion

Un système de détection d'intrusion (IDS : Intrusion Detection System) est un mécanisme destiné à repérer des activités anormales ou suspectes sur la cible analysée (un réseau ou un hôte). Il permet ainsi d'avoir une connaissance sur les tentatives réussies comme échouées des intrusions.
Classés en plusieurs catégories :

- NIDS : Network based IDS qui surveillent l'activité au niveau du réseau `snort`
- HIDS : Host based IDS qui surveillent l'activité au niveau des hôtes `ossec`
- IDS hybrides : qui combinent les 2 précédents

https://www.comparitech.com/net-admin/network-intrusion-detection-tools/

Contenu

Définition Privacy

Concept abstrait et subjectif qui dépend de la discipline d'étude, de normes sociales et des attentes sociétales ainsi que du contexte.

- **Point de vue légal :** "droit à la tranquillité" ou the right to be let alone ; droit de chaque personne de décider quelle information personnelle peut être divulguée aux autres et dans quelles circonstances
- **Point de vue psychologique :** "la liberté de construire son identité propre dans un environnement contraint" car la construction de son identité dépend de la vision d'autrui ; cf. réseaux sociaux, profilage,....

Classification Privacy (Solove, 2006)

- collecte d'information
  - surveillance
  - interrogation
- traitement de l'information
  - agrégation
  - identification
  - insécurité
  - usage détourné
  - exclusion
- dissémination de l'information
  - perte de confidentialité
  - divulgation
  - exposition
  - accessibilité accrue
  - blackmail
  - appropriation
  - distortion
- invasion
  - intrusion
  - interférence décisionnelle

Protection des données

- 1995 : Directive Européenne sur la protection des données
- 2016 : Règlement Général sur la Protection des Données
  - s'applique aux "données personnelles" : toute information relative à un individu (pas d'application aux activités de sécurité nationale ou juridiques)
  - Règlement sur la protection des personnes physiques relatif au traitement des données personnelles et au mouvement de ces données.

RGPD : les principes

- loyauté : ce qui est traité doit correspondre à ce qui a été décrit
- transparence : les personnes ont le droit d'obtenir les infos nécessaires pour assurer un traitement loyal
- limitation des finalités : les données ne peuvent être obtenues que pour des finalités déterminées, explicites et légitimes
- minimisation des données : la quantité minimale de données doit être recueillie et conservée pour un traitement spécifique.
- exactitude : les détenteurs de données doivent créer des processus de rectification et suppression dans les BD.
- droit d'accès et de rectification :
- limitation de la conservation
- intégrité et confidentialité
- responsabilité : Le responsable du traitement des données doit être capable de démontrer sa conformité avec la totalité des autres principes

Protection des données : introduction

La CNIL en collaboration avec un Youtuber a présenté une introduction au RGPG : RGPD en video réponse aux questions en Video

Monde offline → monde online

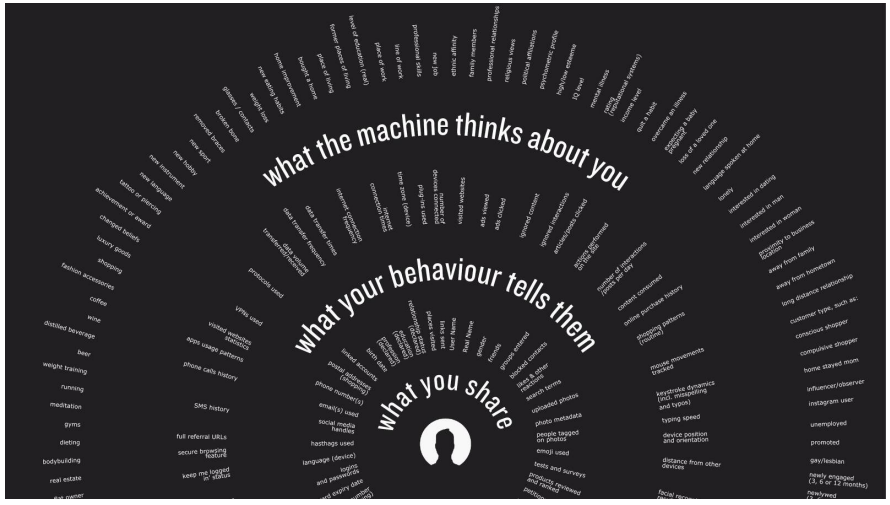
- information difficile à collecter, mémoriser, chercher et accéder
  - conversation F2F
  - documents papier
  - paiement liquide
  - filature
  - trouver vos relations
  - recherche dans dictionnaires
- difficile de copier, diffuser, facile à détruire
- difficile à agréger, profiler et inférer
- oubli dans le temps
- ...

- information facile à collecter, mémoriser, chercher et accéder
  - messagerie instantanée
  - mails
  - fichiers numériques
  - paiement par carte
  - géolocalisation
  - amis en ligne
  - requêtes google,. . .
- facile à copier, diffuser, dur à détruire
- facile à agréger, profiler et inférer
- information jamais perdue
- ...

Rien à cacher ?

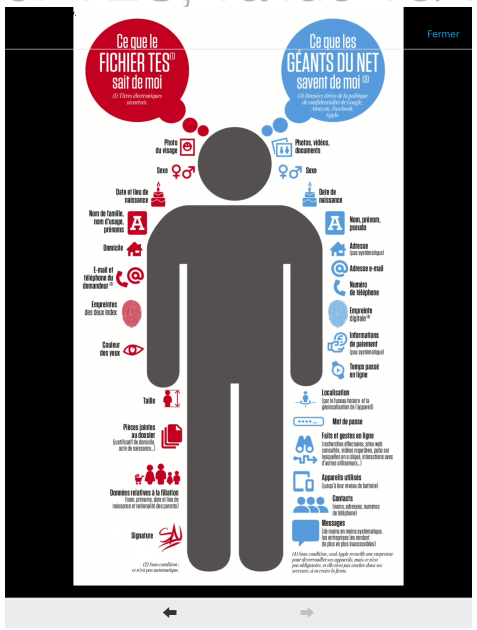
- L'argument "je n'ai rien à cacher" repose sur l'hypothèse que la vie privée est de cacher des actions négatives
- Ce qui rend une société agréable à vivre est sa capacité à protéger les individus de l'intrusion des autres dans la vie privée. Une société sans protection de la vie privée deviendrait vite étouffante.
- Ecart entre secret et privé :
  - vos actions quotidiennes, vos amis ou relations, vos propos dans une conversation, vos hobbies,...
  - tout ce qui précède n'est peut-être pas secret mais vous n'avez pas forcément envie de rendre ces activités publiques ou accessibles à d'autres qui peuvent les analyser et en tirer des conclusions

Niveaux de vie privée



Exemple pratique

Fichier TES, validé 18/10/18



Vu sur capital

Mieux vaut lire les conditions d'utilisation avant de s'inscrire sur Tinder. Le géant américain Match Group qui possède plus de 130 sites de rencontre à travers le monde comme Meetic ou OkCupid collecte la majorité des données de ses utilisateurs. Rien y échappe : position géographique, adresses mail, profession, loisirs, photos, mensurations et mêmes les préférences sexuelles.

Surtout, le groupe n'hésite pas à revendre ces données à qui le souhaite. L'ONG Tactical Tech et la chercheuse Joana Moli ont pu facilement acquérir ces informations concernant des millions de personnes pour environ 135 euros sur le site US Date. Un achat tout à fait légal. L'utilisateur donne son accord pour cette revente de données lorsqu'il accepte les conditions générales d'utilisation pendant son inscription. Il ne peut donc pas attaquer Match Group sur la vente de ces données.

Cette pratique est courante dans l'industrie selon Tactical Tech. L'ONG craint surtout que la récolte de ces informations soit utilisée à mauvaise escient par d'autres groupes privés ou administratifs. "L'utilisateur pourrait se voir imposer des restrictions sur son assurance maladie, ses demandes de crédits, son accès à l'éducation et bien plus encore. Exploiter ce genre de données très intimes peut causer des dégâts catastrophiques sur la vie des personnes concernées. Surtout si leurs profils est rendu public", conclut Tactical Tech.

Vie privée et technologie

- Ligne rouge : nos actions et interactions sont de plus en plus tracées par la technologie
  - on laisse des traces partout
  - ces traces sont combinées, agrégées et analysées pour déduire plus d'informations sur nous et de prendre des décisions qui auront un impact
  - nous n'avons pas de contrôle sur nos informations ou sur les déductions qui en découlent
- L'information n'est jamais oubliée
  - mais peut être parfois utilisée hors contexte

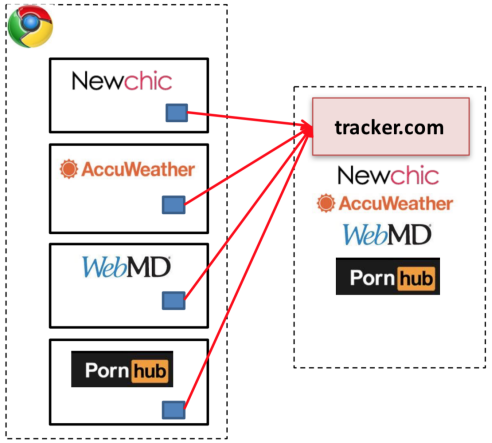
Plan

Exemple

Navigation sur un site de presse :

Avec 21 traceurs !
D'après Springer : "le contenu journalistique est seulement un prétexte pour que le lecteur regarde les publicités".

Web tracking (N. Bielova, INRIA)



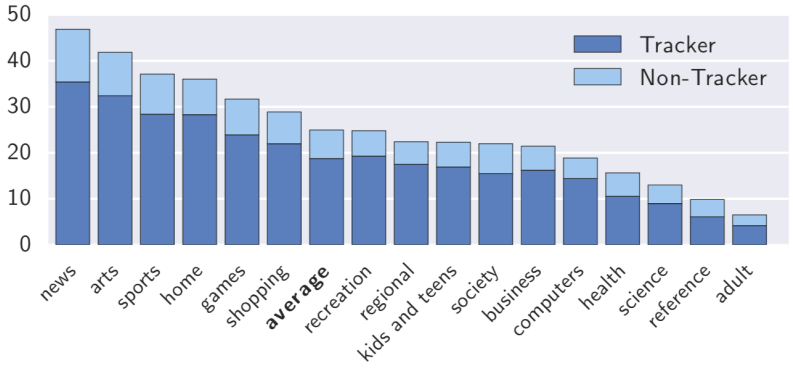
Bigger browsing profiles
= increased value for trackers
= reduced privacy for users

Pourquoi s'inquiéter ?

- Collecte de nos données sans qu'on le sache
  - sur des sites sensibles
  - mémoire de nos habitudes de navigation sur le web, de nos préférences, envies, humeurs
- utilisation de nos données
  - ciblage publicitaire
  - manipulation

Cambridge Analytica

Combien de traceurs par site ?



Comment ça marche ?

- Par les cookies
- par d'autres mécanismes de stockage et les cookies zombies
- par des mesures à large échelle
- avec des publicités ciblées et du cookie synching
- par l'empreinte de votre navigateur