Visualiser et contrôler ses traces sur le web

7 avril 2015 - cryptoparty

Chloé Lailic, bibliothèque de l'INSA Rennes
Damien Belvèze, bibliothèque de l'INSA Rennes

Cet atelier, destiné à des débutants, a pour objectif une prise de conscience : Quelles traces laissons-nous sur le web ? Qui en veut à nos données personnelles ? Comment reprendre le contrôle de ces données à l'ère du numérique et du tout connecté ?
Nous accompagnerons les participants dans l'installation et la manipulation de quelques outils libres et simples d'utilisation pour contrôler les traces que nous disséminons sur la toile.

Sources/outils :
http://www.controle-tes-donnees.net/
Reprenez le contrôle de vos données. La Quadrature du Net

https://www.cryptoparty.in/documentation/handbook
Cryptoparty Handbook. 2013/08/21

Introduction : pourquoi c'est important de prendre conscience de la manière dont on est pisté sur le web

"*It is not only our right to preserve our privacy but also our responsability to defend the right against the intrusions of governments, corporations and anyone who attempts to dispossess us. If we do not exercise thoses rights today, we deserve whatever happens tomorrow*."
(Cryptoparty Handbook, page 62)

(Ce n'est pas seulement notre droit de préserver notre vie privée, mais aussi notre responsabilité de défendre ce droit contre les intrusions des gouvernements, firmes commerciales et toute entité ou personne qui cherche à nous en déposséder. Si nous n'exerçons pas ces droits aujourd'hui, nous méritons tout ce qui peut nous arriver demain.)

Sur le point précis de la surveillance de masse exercée par les gouvernements, voir aussi la vidéo de la Quadrature du Net

Écosystème de la collecte de données personnelles : qui et comment ? Big Data et Big Brother ?

"*Tous les messages transmis sur le Net ainsi que les sites et vidéos consultés par tous les internautes sont analysés par les géants de l'Internet (Google, Facebook, Apple, eBay, Amazon, Microsoft). Quelles informations ces entreprises en tirent-elles ?*
Une étude de l'université de Cambridge en donne un aperçu : 58.000 personnes ont répondu à un test de personnalité, puis ce test a été recoupé à tous les « j'aime » que ces personnes avaient laissés sur Facebook. En repartant de leurs seuls « j'aime », l'université a alors pu déterminer leur couleur de peau (avec 95% de certitude), leurs orientations politique (85%) et sexuelle (80%), leur confession religieuse (82%), s'ils fumaient (73%), buvaient (70%) ou consommaient de la drogue (65%). Cette étude aurait été évidemment bien plus précise en se basant sur les 1.15 milliards d'inscrits (en mai 2013) de Facebook. En outre, Facebook est aussi en mesure de surveiller tout internaute visitant un site Internet proposant un bouton « j'aime », et cela même si cet internaute ne clique pas sur le bouton en question ou n'est pas inscrit sur Facebook. De la même manière, Google a la possibilité d'observer le milliard d'individus utilisant son moteur de recherche chaque mois. Cette entreprise peut également surveiller toute personne qui visite un site affichant l'une des 45 milliards de publicités qu'elle vend chaque jour (voir : ce qu'en dit la CNIL). Enfin, tous les messages envoyés ou reçus depuis Gmail sont analysés. Et ils ne sont pas les seuls : eBay, Amazon, Apple, Microsoft, Yahoo, etc. surveillent aussi leurs millions de visiteurs quotidiens."
Source :: http://www.controle-tes-donnees.net/

Qui en veut à nos données ?

- Intérêts commerciaux des services "gratuits" du web : "si vous ne payez pas, c'est vous le produit" & Intérêts commerciaux des sites tiers.

Voir Spécial Investigation : Big Data, les nouveaux devins / Canal+ : (13m40 -> 16m24)

- Surveillance gouvernementale -> Les géants du web collaborent main dans la main avec des institutions comme la NSA : *'La NSA a accès à toutes les informations que Google, Facebook, YouTube, Microsoft, Yahoo!, Skype, AOL et Apple ont réunies sur leurs utilisateurs : leurs messages privés, leurs recherches, leurs fréquentations...'*
  Pour...combattre le terrorisme mais il y a surtout un véritable objectif politique et économique. L'une des suites des révélations de Snowden rapportées par Glenn Greenwald ont mis à jour l'espionnage économique auquel se livraient les Etats Unis sur le géant pétrolier brésilien Petrobras : on est très loin du terrorisme et très près de l'espionnage industriel entre pays alliés qui plus est (cf. Les Echos.fr 17/09/2013).

1. Comment naviguer avec des navigateurs grands publics en minimisant ses traces

Quelles données ?
Quand on s'inscrit à des services, on est susceptibles de renseigner adresse/date de naissance/sexe, etc.
+
Au cours d'une navigation, vous laissez les informations suivantes :
- IP
- Profil de navigation (nom du navigateur, du système d'exploitation, marque du portable)
- Données de navigation (temps passé sur chaque page, clics, requêtes)
- Géolocalisation

1. Pour une navigation sécurisée il est préférable d'opter pour un **navigateur libre** comme Mozilla Firefox. De manière générale, les outils libres (système d'exploitation Linux, navigateurs comme Firefox, Searx de la Quadrature du Net, Duckduckgo, référentiels de messagerie sécurisée comme Open SSL) sont des garants de sécurité, même s'ils ne sont pas à l'abri de failles qui seront corrigées d'autant plus vite que le nombre et la compétence des internautes relecteurs du code source sera grande, cf. le cas Heartbleed)

2. Il est préférable d'opter pour des protocoles alternatifs au HTTP-IP comme HTTPS

Sécuriser un maximum de connexions avec **Https Everywhere** (module qu'on retrouve dans le navigateur de navigation anonyme Tor)
Voici comment fonctionne le protocole HTTPS par opposition au protocole HTTP :

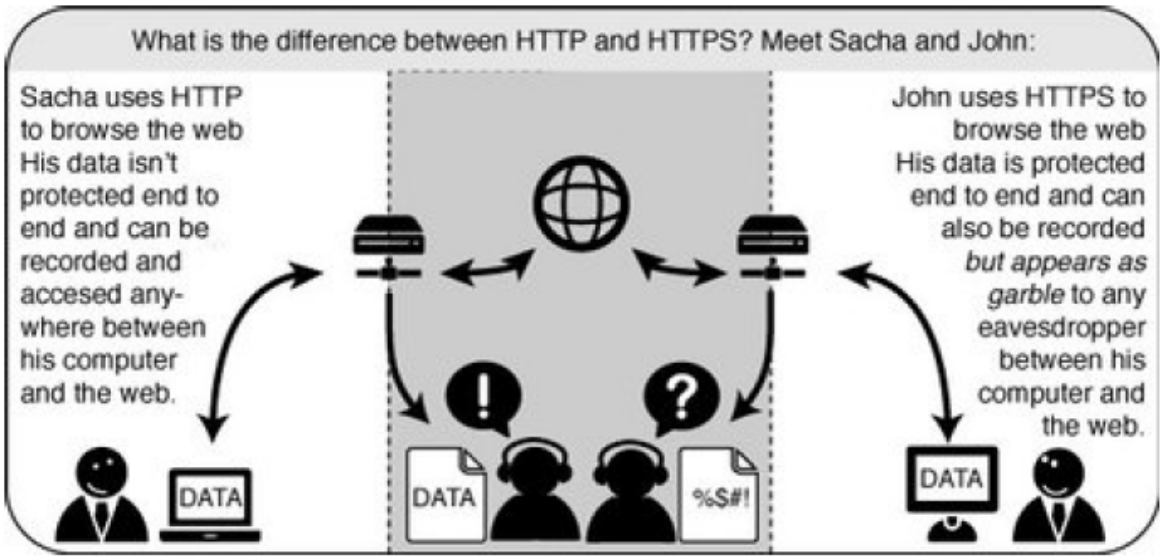


Figure 8.13: HTTPS Schema

(Source : Cryptoparty handbook)

Les communications de John (à droite) comme celles de Sacha (à gauche) peuvent être enregistrées par un tiers, mais dans le deuxième cas, elles ne sont pas lisibles.
Le module HTTPS everywhere consiste à proposer ce type de communication avec le serveur du site visité dans tous les cas, même quand le site ne propose pas par défaut de communication via le serveur HTTPS.

Lorsque vous utilisez un navigateur, celui-ci communique votre IP au serveur afin que ce dernier puisse répondre à votre requête (protocole TCP/IP)
Vous ne donnez pas votre adresse postale à tout le monde. Pour la même raison, vous devez être prudent dans votre navigation : votre IP sera conservée par le serveur

3. Pendant que vous consultez un site, vous laissez des traces de votre navigation au serveur -> historique de votre navigation

> Effacer l'historique ou allouer 0 espace pour l'historique

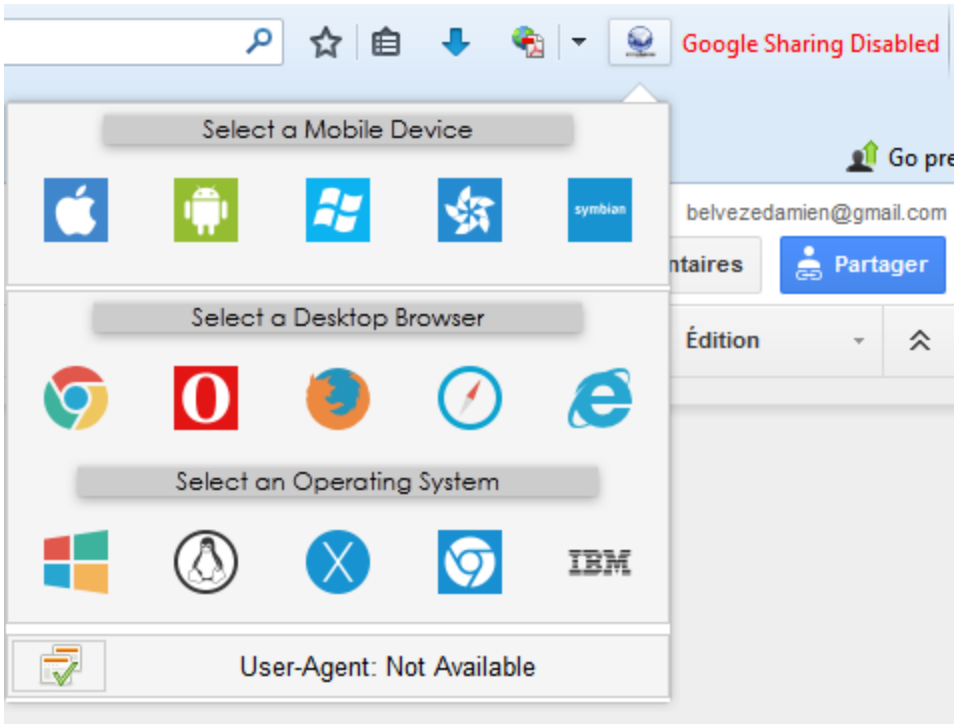
- Données dans le cache

> Effacer le cache de manière régulière ou mieux au moment de chaque fermeture de votre navigateur

- Données sur votre navigateur (User Agent)

> changer régulièrement par un simple clic les informations que vous transmettez automatiquement. L'extension **User Agent Switcher** vous permet de faire cela. En trois clics, vous pouvez changer radicalement le profil de votre appareil tel que ce profil sera conservé

en mémoire dans le serveur que vous interrogez. Ce profil constitue la base de ce qu'on appelle la signature numérique grâce à laquelle on estime aujourd'hui pouvoir identifier 84% des ordinateurs connectés alors même qu'on en ignore l'IP (voir atelier de Pierre La Perdrix sur le sujet)



> activer la navigation privée (pas de sauvegarde des pages vues, pas de sauvegarde de la navigation : Firefox : menu > Start private browsing
Inconvénients : pas de possibilité de conserver les mots de passe, pas de social bookmarking possible

Navigation sociale : Paramètres de confidentialité sur Facebook / Google (de temps en temps : ces paramètres évoluent)

2. Focus sur les cookies

Ces derniers ont souvent pour but d'améliorer votre expérience de lecture :

- Ils augmentent les performances du navigateur (navigation plus facile)
- Ils permettent de réduire le téléchargement de données
- Ils permettent de garder en mémoire les mots de passe et vos préférences pour ce site (par exemple Réseau social)

Les cookies peuvent garder des traces de votre visite :

- de manière inoffensive pour mesurer l'audience du site et des pages qui le constituent : cf. Le tracker de Google Analytics sur les sites institutionnels par exemple

- Ils peuvent repérer des usages qui reviennent sur la page
- Certains peuvent collecter des informations sur les autres sites que vous fréquentez, notamment afin de vous envoyer de la publicité comportementale (behavioral advertising)

Ex. Vous visitez la page de Wikipedia consacrée à Majorque et dans la foulée, vous recevez des mails publicitaires pour des hôtels situés sur cette île des Baléares
A force le web devient un peu trop familier et vous enferme dans une de ces bulles de filtrage (filter bubble)

Ghostery fonctionne comme une machine à rayons X qui révèle toute la technologie destinée à vous surveiller qui est généralement implémentée dans la page web que vous lisez.
Comme d'autres services du même genre (Do not track PLus et Trackerblock) Ghostery vous permettra de bloquer les cookies par catégories ou individuellement

Genre	fonction	exemple
analytique	offrent des données de recherche ou d'analyse à des éditeurs de sites web	Google analytics tracker
balises	servent à assurer le suivi (balises, pixels de conversion, pixels de segmentation d'audience)	
confidentialité	avis de confidentialité et autres élements liés à la confidentialité	Cookie Q : permet d'afficher un bouton d'acceptation ou de refus de chargement d'autres cookies
publicité	cookie à vocation publicitaire destiné à orienter le marketing et à envoyer de la publicité contextualisée	AD4
widget	cookie proposant des fonctionnalités de pages (bouton de réseau social, formulaire de commentaire)	Addthis Facebook social plugin

D'après Ghostery

> Effacer les cookies régulièrement (Firefox : menu > Options > Vie privée > clear your recent history + Paramétrer Firefox pour ne pas qu'il conserve les cookies.
Option Do not track (options > Vie privée > Dire aux sites que je ne souhaite pas être pisté
Cette option est installée par défaut dans le navigateur Safari. Google a eu la mauvaise idée de la contourner en 2011-2012. Les autorités fédérales ont obligé la firme à verser 22,5 millions de dollars pour éviter les poursuites liées à cette intrusion.
http://www.numerama.com/magazine/32615-google-sera-bien-poursuivi-pour-violation-de-la-vie-privee-en-gb.html
Le logiciel Disconnect.me mis au point par des ingénieurs de Google et un magistrat engagé dans la défense des utilisateurs et de leur vie privée permet dans sa version gratuite de visualiser les trackers qui vous pistent et d'anonymiser les recherches en ligne.

Le cas des Spywares

Un cookie est un fichier texte conservé sur le navigateur, et même s'il livre des indications sur la navigation de l'internaute au propriétaire du site visité. Ce n'est pas un exécutable ce qui le distingue d'un logiciel espion (spyware). Pour lutter contre ces derniers, le plus pertinent est de se doter d'un (seul) logiciel antispyware ou bien, si on utilise Windows, de charger le logiciel Detekt, promu par Amnesty International qui détecte les spywares les plus couramment utilisés par les services de renseignement.