

ANONYMISATION

Mettre en place au sein de l'entreprise des mécanismes visant à limiter l'analyse et la provenance des communications sortant vers internet.

Olivier Singer
Réunion IE Clusir Lyon
16 Février 2011

L'action visant à rendre anonyme des données informatiques implique la dissimulation d'informations permettant la reconnaissance d'une signature distincte lors de l'utilisation d'une ressource ou d'un service distant.

Le renforcement de cette notion se fait au travers de l'impossibilité de pouvoir relier cette signature à l'utilisation qui en est faite, et ce sans qu'aucune corrélation ni observation ne soit possible. [1] et [2]

Pour se faire il faut à la fois que le groupe d'utilisateurs que l'on veut rendre anonyme soit le plus vaste et le plus hétérogène possible, et que les informations transmises ne puissent pas permettre de distinguer ou de corréler les informations relatives aux requêtes effectuées par un programme ou un utilisateur particulier.

Nous avons donc 2 domaines à couvrir :

1. La dissimulation de l'adresse IP de la machine réalisant effectivement la requête. (Couche routage réseau - OSI 3)
2. L'assainissement des données échangées par le navigateur avec le site distant. (HTTP: couche application OSI 7)

Le réseau de proxies Tor est une réponse intéressante au point 1 : sa communauté planétaire ne se limite pas à un pays, à une société. Il dispose d'une architecture assez complexe pour permettre un niveau intéressant de dissimulation de l'adresse IP source et les requêtes le traversant, n'empruntent pas toujours le même chemin. De plus les connexions sont chiffrées entre le client Tor et le dernier noeud de l'infrastructure, permettant un certain niveau de confidentialité.

Le nettoyage/filtrage des caractéristiques du client peuvent être effectués par différents moyens au niveau du poste à protéger et au niveau d'une infrastructure de proxies internes placés avant le relais vers l'infrastructure Tor, afin de rationaliser l'administration de l'anonymisation des requêtes sortantes vers internet.

Au niveau du navigateur Web :

- Installation d'outils ou de modules additionnels interceptant les échanges de certains contenus (headers, cookies, referrers...)
- Limitation des modules additionnels activés, des langages dynamiques supportés (javascript, java).
- Choix délibéré d'utiliser un client historique avec une surface d'attaque moins importante (Lynx en mode texte par exemple).
- Limitation des impacts en cas d'attaque (sandboxing, virtualisation...) [11].

Au niveau du proxy de filtrage :

- Assainissement par filtrage ou nettoyage du protocole HTTP et du code HTML.

Au niveau de l'infrastructure Tor :

- Camouflage de l'adresse IP du client.
- Augmentation de la confidentialité du trafic web émis côté client (chiffrement des communications sortantes).

Que pouvons nous faire avec les informations transmises par le client au serveur ?

Une corrélation entre :

- L'adresse IP
- Le referrer (la page qui contenait le liens vers la nouvelle page). Dans le cas ou cette page est une page de recherche google, le site a donc accès à la requête de recherche.
- Les informations sur la configuration du client internet et, suivant la technique utilisée, sur son fonctionnement historique.

Exploiter, consolider ces informations qui pourront par la suite être fouillées afin de permettre l'extraction d'informations plus exploitables.

exemple: analyse d'une attaque et remontée grâce au referrer à la requête google de détermination des cibles [14]
Google Analytics, cookies [19]

Le réseau de proxies TOR

Tor est un réseau de tunnels virtuels composé de proxies chaînés, qui permettent d'améliorer l'anonymisation des échanges TCP/IP, en rendant plus difficile les analyses du trafic réseau. Historiquement il est issu d'un projet du laboratoire de recherche Naval des états-unis qui avait comme but premier de protéger les communications du gouvernement.

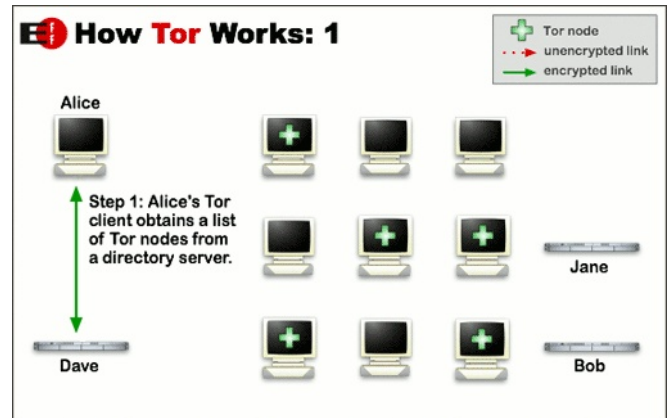
Fonctionnalités de Tor [4]

- Topologie évolutive permettant de sortir du réseau de n'importe quel point de routage
- Contrôle de congestion réalisé par les routeurs périphériques, réduisant la quantité d'informations transmises lorsqu'un problème de congestion ou de flood est détecté.
- Maintien des informations sur l'état de l'infrastructure grâce à des serveurs catalogue.
- Chaque noeud maintient une connexion TLS avec les autres noeuds en utilisant des clés éphémères.
- Contrôle de l'intégrité des données transmises au travers de l'infrastructure, avant leur émission en clair.

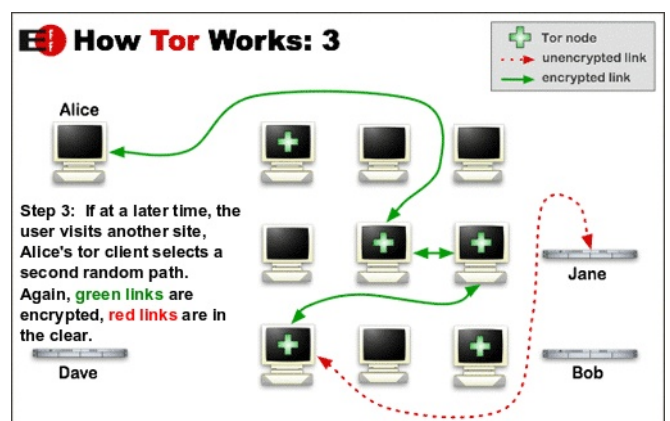
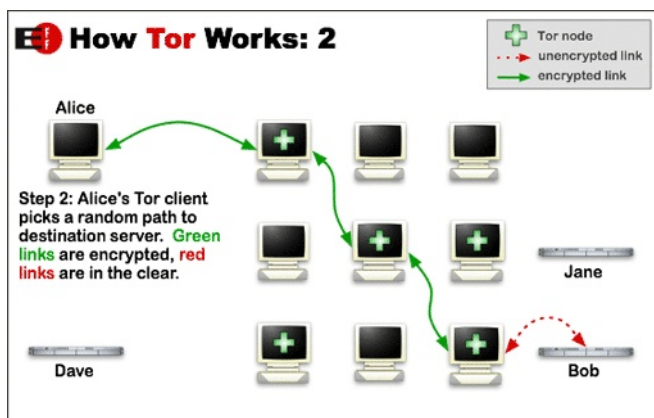
Fonctionnement de TOR

Tor utilise une infrastructure de noms spécifiques car le réseau est caché des serveurs DNS root. C'est le serveur annuaire de l'infrastructure Tor qui indique au client les noms et adresses des serveurs proxies. Ces informations étant primordiales, c'est le premier échange qui aura lieu entre le client Tor et son infrastructure.

Ensuite le client choisit au hasard un chemin complexe qu'il est le seul à connaître, qui passera par certains de ces serveurs, et négociera des clés de chiffrement symétrique pour chacun des noeuds. Chaque noeud traversé aura connaissance uniquement de la source et de la destination du paquet.



Toutes les communications TCP sont chiffrées et multiplexées entre le client et l'infrastructure Tor. C'est au niveau du dernier noeud que les informations passent en clair : la requête au serveur DNS internet pour la résolution du nom et la connexion au serveur de destination.



Afin de limiter les latences lors de l'utilisation de l'infrastructure Tor, le circuit de communication est construit préventivement afin d'être utilisé dès que cela est nécessaire. Un nouveau circuit est ainsi créé toutes les minutes, ou lorsqu'une nouvelle connexion TCP rend nécessaire la construction d'un nouveau chemin de communication.

La connexion à l'infrastructure TOR peut être réalisée de différentes manières : D'un navigateur embarquant tous les outils permettant une connexion immédiate au réseau Tor, pouvant s'exécuter d'une clef USB, ou au travers d'un outil de type proxy (privoxy, polipo), sous la forme d'un service ou daemon réalisant l'encapsulation des requêtes vers le client Tor avec le protocole SOCKS v4/5. Il existe également des distributions linux particulières embarquant un large pannel d'outils de sécurité et d'anonymisation (Distribution linux backtrack).

Limitations connues à Tor [4, 5, 6, 7]

- Faible bande passante

Attaques passives :

- Observation statistique de la signature du trafic réseau en bordure de Tor (entrée/sortie)
- Pas d'anonymisation au niveau 6/7 : les fonctionnalités doivent être couvertes par des produits spécifiques comme Privoxy.

Attaques actives :

- Compromission des clefs TLS
- Compromission d'un Onion proxy.
- Compromission d'un serveur catalogue.

PRIVOXY

"Privacy Enhancing Proxy"

L'outil privoxy est un proxy sans fonctionnalités de cache possédant des capacités de filtrage du contenu du protocole HTTP et du langage HTML.

Il permet de compléter la fonction de protection de l'adresse IP par des fonctionnalités de contrôle (filtrage, assainissement) et de communication (chaînage avec d'autres proxies, SOCKS 4/4a/5).

Il comprends un mécanisme simple de configuration permettant de définir le niveau de protection (Cautious / medium/ advanced) afin de définir le niveau des actions et filtrages suivants :

- Ad-blocking
- Ad-filtering by size
- Ad-filtering by link
- Pop-up blocking
- Privacy Features
- Cookie handling (session-handling / Kill)
- Referer forging
- GIF de-animation
- Fast redirects
- HTML taming
- JavaScript taming
- Web-bug killing
- Image tag reordering
- ACLs

Capacités principales de Privoxy

- Fonctionnement en mode transparent possible afin d'intercepter et filtrer les flux HTTP
- Filtrages conditionnel de contenu Web, de sites par ACL
- Chaînage conditionnel avec d'autres proxies, ou un client Tor.
- Supporte les expressions régulières de type Perl
- Fonctionnalités d'anonymisation (Définition d'actions et de filtres permettant de travailler le contenu du flux HTTP, particulièrement des headers client/serveur)
- Contrôle (filtrage, blocage, assainissement) du code javascript, HTML (pop-ups, comportements, caractéristiques, évènements, cookies, abus, bannières, exploits...).
- Filtrage du Flash.
- Support tagging (Client/server Header behaviour) : A la capacité de réaliser des actions spécifiques éventuellement chaînées, à partir des informations HTML reconnues grâce à des expressions régulières

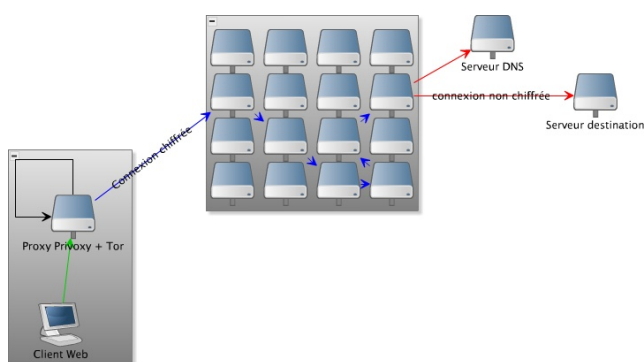
Les possibilités offertes par le système de configuration puissant de privoxy permettent de réaliser beaucoup d'opérations sur les flux HTTP, avec le risque de devoir passer plus de temps dans la maintenance des cas particuliers si ceux-ci n'ont pas été anticipés au niveau de la configuration.

NAVIGATION ANONYME AVEC FIREFOX PRIVOXY ET TOR

Mise en place d'une navigation anonyme avec les outils opensource Firefox et modules additionnels, privoxy (Anonymisation HTTP/HTML), le client Tor sur liaison internet standard.

Deux machines :

- Un PC utilisateur sur lequel est installé la dernière version de firefox et certains modules additionnels (voir encadré).
- Une passerelle sur laquelle est installée le proxy Privoxy et le client Tor, reliée à internet.



Configuration :

- Tor : configuration standard, daemon exécuté en tant qu'utilisateur particulier, écoute sur la loopback.
- Privoxy : anonymisation activée, chaîné avec Tor, SOCKS5.
- Firefox : navigation Tor activée, configuration proxy de privoxy renseignée, modules additionnels installés et configurés.

Test :

- Surf sur internet : lent mais fonctionnel.
- Observation des requêtes HTTP échangées entre le client et le serveur.
- Vérification du fonctionnement du camouflage de l'adresse IP et des caractéristiques principales du navigateur.

Modules additionnels FireFox :

- Torbutton (permet de basculer sur Tor ou naviguer normalement. augmente l'anonymisation, limite les fonctionnalités représentant un risque pour l'anonymité)
- httpfox (Extension Firefox pour observer les headers HTTP)
- NoScript (permet de limiter l'accès à JavaScript, Java, Flash et d'autres plug-ins, Anti-XSS protection.)
- PrivacySuite (gestion des cookies, informations sur les données conservées...)

Test de connexion à google.com : il nous redirige vers <http://www.google.co.uk/>, preuve qu'il n'arrive plus à détecter notre arrivée de France.

autre exemple :

```
Hostname: tor-exit-router38-  
readme.formlessnetworking.net  
Proxy: No Proxy or Invisible Proxy Used  
Internal (LAN) IP: Checking... Sorry You  
Need Java For This To Work
```

Attaques contre le navigateur Web permettant d'isoler les caractéristiques privées : [5, 6, 7, 8, 14, 15, 16, 17] utilisant des techniques :

- VBscript / ActiveX
- Flash
- Javascript
- Java
- Fingerprinting web client (SSL..), telnet, SSH par analyse réseau.
- HTML5 (éléments multimedia référencés au travers d'un FTP).
- Certains protocoles connus (FTP..)
- Fuites de requêtes DNS (normalement émises au niveau du proxy Tor sortant).

Test d'anonymisation réalisée sur le site <http://www.plinko.net/404/supersleuth.asp> :

ALL HTTP

HTTP KEEP_ALIVE:115

HTTP_ACCEPT:text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8

HTTP_ACCEPT_CHARSET:ISO-8859-1,utf-8;q=0.7,*;q=0.7

HTTP_ACCEPT_ENCODING:gzip,deflate HTTP_ACCEPT_LANGUAGE:en-us,en;q=0.5

HTTP_HOST:www.plinko.net HTTP_USER_AGENT:LYNX (VT100; I; Searching-for-the-light) HTTP_DNT:1

ALL RAW

Keep-Alive: 115 Accept: text/html,application/xhtml+xml,application/xml;q=0.9,*/*;q=0.8

Accept-Charset: ISO-8859-1,utf-8;q=0.7,*;q=0.7 Accept-Encoding: gzip,deflate Accept-

Language: en-us,en;q=0.5 Host: www.plinko.net User-Agent: LYNX (VT100; I;

Searching-for-the-light) DNT: 1

LOCAL_ADDR

72.167.183.8

LOGON_USER

PATH_INFO

/404/supersleuth.asp

REMOTE_ADDR

192.251.226.206

REMOTE_HOST

192.251.226.206

REMOTE_USER

HTTP_USER_AGENT

LYNX (VT100; I; Searching-for-the-light)

REFERENCES

Anonymisation :

[1]: ISO/IEC 15408-2:2008 Information technology -- Security techniques -- Evaluation criteria for IT security -- Part 2: Security functional components
<http://standards.iso.org/ittf/PubliclyAvailableStandards/index.html>

[2]: Anonymity, Unobservability, and Pseudonymity– A Proposal for Terminology / Andreas Pfitzmann and Marit Köhntopp

[3]: On the Economics of Anonymity / Alessandro Acquisti, Roger Dingledine, and Paul Syverson

Tor :

[4]: Tor: The Second-Generation Onion Router / Roger Dingledine, Nick Mathewson, Paul Syverson

<https://svn.torproject.org/svn/projects/design-paper/tor-design.pdf>

[5]: Compromising Tor Anonymity - Exploiting P2P Information Leakage / Pere Manils, Abdelberi Chaabane, Stevens Le Blond, Mohamed Ali Kaafar, Claude Castelluccia, Arnaud Legout, Walid Dabbous

<http://hal.inria.fr/docs/00/47/15/56/PDF/TorBT.pdf>

[6]: Peeling the Onion: Unmasking TOR Users / Andrew_Christensen & Dan_Faerch

http://www.fortconsult.net/images/pdf/tpr_100506.pdf

[7]: Breaking Tor Sessions with HTML5 / Marco Bonetti

<http://sid77.slackware.it/tor/BreakingTor.pdf>

[8]: Metasploit Decloaking Engine

<http://decloak.net/>

Histoire :

[9]: Electronic Privacy Information Center : CARNIVORE
(Système d'observation des communications internet mis en place au niveau de fournisseurs d'accès internet américains par le FBI)

[10]: Scandale des données d'AOL :

http://en.wikipedia.org/wiki/AOL_search_data_scandal

Protection des données du poste utilisateur et du navigateur Web:

[11]: Damage Control for Network Applications / Magnus Melin - 2007-11-18
(3.1 Application Confinement And Sandboxing)
http://www.tml.tkk.fi/Publications/C/25/papers/Melin_final.pdf

[12]: BlackHat-DC-09 : Dissecting Web Attacks (20/01/2009) / Val Smith, Colin Ames, Delchi
<http://www.blackhat.com/presentations/bh-dc-09/ValSmith/BlackHat-DC-09-valsmith-colin-Dissecting-Web-Attacks.pdf>

[13]: iPhone Applications & Privacy Issues: An Analysis of Application Transmission of iPhone Unique Device Identifiers (UDIDs) - Eric Smith

[14]: The Pen Test Perfect Storm Part 5 / Ed Skoudis, Kevin Johnson,& Joshua Wright (Web Client Attack Tools and Techniques)

[15]: Browser Exploitation for Fun and Profit - SANS Special Webcast / Raúl Siles - November 2, 2010
http://www.taddong.com/docs/Browser_Exploitation_for_Fun&Profit_Taddong-RaulSiles_Nov2010_v1.1.pdf

[16]: Cross Context Scripting with Firefox / Nick Freeman - 21 April 2010
http://www.security-assessment.com/files/whitepapers/Cross_Context_Scripting_with_Firefox.pdf
http://www.security-assessment.com/files/whitepapers/Exploiting_Cross_Context_Scripting_vulnerabilities_in_Firefox.pdf

[17]: Social Networking: Good for Business or Security Nightmare - TraceSecurity

[18]: A Dynamic End-to-End Security for Coordinating Multiple Protections within a Linux Desktop / Jeremy Briffaut, Martin Peres, Christian Toinard (2010-05-21)
<http://mupuf.org/media/files/cts2010-ensib.pdf>

Ressources :

- <http://www.googlesharing.net/>
- <https://check.torproject.org/>
- <https://www.torproject.org>
- <http://www.privoxy.org/faq/misc.html#TOR>
- <http://www.whatsmyip.org/more/>
- HTML v5 : <http://html5.org/>
- Social Engineering Toolkit : [http://www social engineer org /](http://www.social-engineer.org/) Manuel de l'utilisateur : http://svn.thepentest.com/social_engineering_toolkit/readme/User_Manual.pdf
- BeEF –Browser Exploitation Framework
- Metasploit Framework : <http://metasploit.com>
- xss attacks informations : <http://xssed.com/>
- Tests d'anonymisation basic : <http://www.plinko.net/404/supersleuth.asp>

Emission télévisée:

[19]: M6 Capital du 6 février 2011
(Quand votre vie privée vaut de l'or / Les dangers de l'internet) :
http://www.m6.fr/emission-capital/06-02-2011-quand_votre_vie_privée_vaut_de_l_or-22767620.html