Les traitements invisibles sur Internet

Dinant, Jean-Marc

Published in:
Droit des technologies de l'information. Regards prospectifs : à l'occasion des vingt ans du C.R.I.D

Publication date:
1999

Les traitements invisibles sur Internet (Invisible processings on the Internet)

Outside Belgium, this article is mentionned namely by Cdt (US) and Juriscom (CA)

par Jean-Marc DINANT,
Informaticien,
Chercheur au Centre de Recherches Informatique et Droit (CRID)
Consultant auprès de la Commission Belge de protection des données
Expert auprès du Groupe de travail 29 sur la protection des données

Résumé

Aujourd'hui, le simple fait d'allumer un ordinateur branché sur Internet met en route plusieurs processeurs qui exécutent subrepticement des centaines de programmes sans que l'utilisateur moyen en soit informé ni puisse avoir le moindre contrôle sérieux sur les données qui y sont traitées.

Historiquement, cette tendance va en s'accentuant et constitue une menace majeure de plus en plus sérieuse pour la protection de données personnelles.

Ce risque se situe tout d'abord au niveau du profilage extraordinaire qu'il est aujourd'hui possible de mettre en œuvre en enregistrant la navigation d'une personne particulière sur le réseau Internet. Mais cette menace se situe aussi au niveau de la sécurité des données à caractère personnel qui sont stockées sur l'ordinateur contenant le navigateur. Un troisième danger est la conséquence des deux premiers. Le profilage pointu du comportement des internautes associé à la possibilité du contrôle des informations se trouvant sur le poste du client permettent dès à présent de mettre sur pied des techniques de simili-web (web-spoofing) ou plus précisément site polymorphique. Malgré une apparence classique et statique, les sites commerciaux d'aujourd'hui se présentent de manière tout à fait différente au regard d'un cybernaute particulier.

1. Introduction : l'inversion du paradigme client-serveur

Depuis que le livre est livre, l'information a toujours été "passive", chaque mot imprimé une fois pour toute, éventuellement reproduit dans un index, attendant patiemment parfois pendant très longtemps le regard d'un lecteur hypothétique. Depuis que le livre est livre, on n'a jamais vu un bouquin jaillir d'un rayonnage et s'ouvrir spontanément à une de ses pages pour attirer l'attention d'un lecteur. Depuis que les bibliothèques existent, on n'a jamais vu un ouvrage s'automutiler, s'arrachant certaines pages et s'en greffant d'autres selon le type de lecteur qui l'approcherait.

En quoi l'informatique et plus particulièrement Internet bouleverse-t-il cette conception séculaire de la bibliothèque ? Entre le papier et l'écran la différence serait-elle tellement énorme qu'elle changerait le rapport entre l'individu et l'information ?

Le modèle client-serveur a longtemps été la base régissant les banques de données : un ordinateur serveur possède l'information qu'il délivre à un ordinateur client. Il s'agit d'une "relation" maître-esclave, postulant la docilité du serveur et la stabilité de l'information. Dans ce monde il n'y a pas, dans le chef de l'ordinateur serveur de discrimination. Chaque client,

quelqu'il soit, effectuant une requête identique reçoit la même réponse. L'information n'est pas altérée. Ce modèle de consultation n'est pas très différent de ce qui se passe lorsqu'un lecteur se rend dans une bibliothèque afin de trouver et de consulter certains ouvrages.

Pour effectuer l'interface entre une machine serveur située quelque part sur le réseau et un utilisateur humain, l'ingénierie logicielle a mis au point à un niveau planétaire un ensemble de protocoles réparti en plusieurs couches. Cette manière de procéder permet de maîtriser la complexité du processus de communication en hiérarchisant la programmation en couche, chaque couche ayant un rôle particulier à jouer. Plus concrètement, chaque couche se compose d'une série impressionnante de plusieurs centaines de sous-programmes informatiques réalisant certaines fonctions à l'intérieur d'une couche particulière. Ces sous-programmes sont exécutés d'une manière invisible pour l'utilisateur moyen.

En fait chaque couche et les programmes qui s'y exécutent possèdent un degré historiquement croissant

- d'indépendance par rapport à l'utilisateur (et bien souvent le propriétaire) de la machine : les programmes exécutent certaines actions que l'utilisateur n'a pas demandé,
- d'opacité par rapport à l'utilisateur : non seulement les centaines de sous-programmes effectuent des actions non demandés par l'utilisateur, mais l'utilisateur n'en est pas informé et n'a pas de moyens fiables de savoir ce qui se passe,
- de dépendance ou d'asservissement par rapport au réseau, ou plus précisément par rapport aux machines installées sur le réseau et censées délivrer passivement l'information. Ces machines peuvent envoyer certaines requêtes aux machines clients du réseau en vue d'obtenir certains renseignements sur l'utilisateur,
- de transparence par rapport au réseau. Le réseau peut obtenir des informations précieuses sur le "naviguant" (c'est à dessein que nous emploierons ce terme pour le différencier du terme navigateur qui désigne le programme utilisé par le naviguant).

Le but de cet exposé est de démonter quelques-uns uns des mécanismes techniques qui permettent au réseau de transformer l'ordinateur de l'internaute en un serveur d'informations personnelles au service du réseau et d'aboutir ainsi à cette inversion du paradigme où l'ordinateur client devient subrepticement sur le terrain le serveur des ordinateurs distants du réseau.

2. Quelques risques liés au protocole TCP/IP

2.1. Le routage d'un courrier électronique "national"

Le réseau Internet utilise le protocole TCP-IP. Si le grand public a en général une assez bonne connaissance de ce que peut-être une adresse IP, plus rares sont les utilisateurs possédant une connaissance d'autres aspects de ce protocole.

Il existe cependant une couche inférieure à la couche TCP/IP. C'est la couche de procédures qui assurent le routage des informations. En effet il faut bien que, "quelque part" dans le réseau, certaines machines puissent savoir où se trouve une adresse IP en particulier. Des outils assez techniques permettent de connaître le chemin que parcourent les paquets IP pour arriver à destination. Le traçage de l'itinéraire d'un courrier électronique envoyé de l'université de Namur vers un fournisseur d'accès indépendant situé à Bruxelles peut surprendre

l'utilisateur néophyte Celui-ci pourrait être tenté de croire qu'un courrier électronique envoyé de Namur à Bruxelles (65 kilomètres) emprunte le chemin le plus court et en tous cas reste en Belgique. Il n'en est rien. Le tableau ci-dessous illustre le chemin de ce courrier (nom du routeur emprunté et adresse IP correspondante).

Itinéraire emprunté par un courrier électronique de Namur à Bruxelles (Avril 1997)

cr1.det.fundp.ac.be	138.48.38.1
leuven.belnet.be	193.190.196.30
brussels.belnet.be	193.190.196.5
BE-s2.dante.bt.net	194.72.26.177
BE-f0-0.eurocore.bt.net	194.72.25.17
CH-s1-3.eurocore.bt.net	194.72.24.77
CH-f0.global.bt.net	194.72.24.67
cern-ebs1.ebone.net	194.72.26.142
Paris-EBS2.Ebone.net	192.121.156.122
stockholm-ebs-s9-0.ebone.net	192.121.154.45
bel1-s5-0-1.datanet.tele.fi *(telecom **Finland**)*	192.130.130.130
brufrgw1.bel.tfi.net *(telecom **Finland**)*	193.209.45.226
mail.ib.be	193.210.152.13

Ce courrier sera passé par la Suisse, la France, la Suède et la Finlande. Plusieurs opérateurs étrangers auront en outre assuré une partie de l'acheminement du message. Cela leur permettrait, éventuellement, de tenir à jour des tables de correspondance entre les adresses IP et les adresses électroniques (dans laquelle figure souvent le nom en clair), de conserver une copie du courrier en vue d'un traitement ultérieur, et...

2.2. L'utilisation des DNS

C'est pour permettre une utilisation plus facile des adresses IP que l'on a inventé les noms de domaines (DNS). Ici encore, cela est concrétisé par une série de sous-programmes qui sont chargés sur la machine hébergeant le navigateur. Un sous-programme aura pour objet d'obtenir l'adresse IP équivalente au nom DNS. Ainsi lorsque l'utilisateur tape une adresse du type *http://www.belgium.fgov.be* il peut croire naïvement qu'il accède directement au site officiel du gouvernement belge. La réalité est toute autre.

Etapes de traduction d'un DNS		
commande visible : http://www.belgium.fgov.be/		
Origine du message	**Contenu du message**	**Destinataire du message**
Navigateur	http://www.belgium.fgov.be/page	Protocole de communication
Protocole de communication	www.belgium.fgov.be	**Serveur DNS**
serveur DNS	adresse IP est 138.190.198.32	Protocole de communication
Protocole de communication	138.190.198.32	Réseau Internet
résultat visible : affichage de la page de bienvenue du site du gouvernement belge		

L'utilisateur peut naïvement croire qu'il a demandé l'accès au site du gouvernement belge. Dans les faits, il a demandé d'accéder à un serveur dont l'adresse IP est considérée par son serveur DNS comme étant celle du gouvernement belge. Autrement dit, préalablement à la connexion au site voulu, le navigateur a effectué une transmission invisible avec un ordinateur du réseau qui effectue la traduction entre l'adresse DNS et l'adresse TCP/IP. La personne qui contrôle le serveur DNS peut donc connaître la liste des sites visités par un internaute particulier.

2.3. La commande PING

Cette commande est relativement connue car elle sert à s'assurer qu'une adresse IP est en cours d'utilisation, quel que soit l'endroit ou elle se trouve dans le monde. Ainsi, il est possible de savoir si l'ordinateur qui sert de serveur au gouvernement belge est connecté.

Etapes d'une interrogation sur le fonctionnement d'une adresse IP

- ping 138.190.198.32
- si l'ordinateur est connecté la réponse sera *réponse de 138.190.198.32*
- si l'ordinateur n'est pas connecté : *pas de réponse*

Concrètement, cela signifie que lorsqu'un ordinateur est allumé, les sous-programmes gérant le protocole TCP/IP vont répondre à un appel du programme Ping issu de n'importe quel ordinateur connecté au réseau Internet. Cela se fait toujours à l'insu de l'utilisateur qui ignore que quelqu'un, dans le réseau, a voulu savoir s'il était connecté au réseau.

3. Le bavardage des programmes de navigation

Lors de la navigation sur Internet, l'internaute utilise un programme appelé navigateur. Les deux marques les plus répandues sont Microsoft Explorer et Netscape Navigator. Ce programme permet de visualiser les pages écrites dans le langage HTML qui se trouvent sur les millions de serveurs rendus accessibles via Internet. Ces programmes de navigation envoient et reçoivent et stockent des informations personnelles relatives à l'internaute via le protocole HTTP qui est utilisé pour le transfert des documents écrits en HTML.

3.1. HTTP, HTML : quelles différences et quels enjeux ?

Le langage HTML est un langage de codage d'un document hypertexte sur Internet. Il permet la mise en page d'un contenu se composant non seulement de texte richement mis en forme (gras, italique, souligné, listes diverses, etc.) mais aussi d'images codées selon des standards largement répandus (p.e. GIF, JPEG,...). Un document stocké sur un support classique (disquette, disque dur, CD-ROM) peut donc être codé en langage HTML. Le langage HTML possède la particularité d'être indépendant d'une société particulière et repose sur des standards définis par une organisation internationale : le World Wide Web Consortium (W3C).

Pour qu'un document écrit en HTML sur un serveur situé en Amérique puisse être visualisé par un internaute européen, il doit d'abord être transféré de ce serveur vers l'utilisateur via le réseau. Préalablement il est bien entendu nécessaire que ce serveur puisse connaître quel est le document concerné et à qui il doit l'envoyer. C'est le but du protocole HTTP. Au sein de ce protocole, on distingue la requête faite par l'internaute et la réponse produite par le serveur. Si

la demande est cohérente, la réponse se composera notamment (mais pas seulement) du code HTML du document demandé. Dans le cas contraire, l'ordinateur serveur générera un message d'erreur au format HTML.

Lorsqu'il souhaite accéder à un site, l'utilisateur tapera une commande du type *http://serveur/document.* Le programme de navigation va alors produire une requête HTTP (c'est le sens du premier paramètre) au serveur désigné par le deuxième paramètre en lui demandant de transmettre un document dont le nom est le troisième paramètre. Il est important pour notre propos de noter que, concrètement, ce n'est pas l'utilisateur qui effectue matériellement la requête HTTP sur le réseau. Il donne à son programme de navigation les éléments nécessaires (serveur et document) pour effectuer cette requête en son nom.

Ce point est important car nous verrons plus loin que le programme de navigation est bien plus bavard que l'internaute ne l'est lors qu'il s'agit de parler au réseau Internet mais qu'il devient subitement plus discret lors qu'il s'agit de montrer à l'utilisateur les données qu'il reçoit effectivement du site visité.

3.2. Hyperliens explicites et visibles

HTTP et HTML commencent tous deux par HyperTexte. Cette notion permet d'introduire la notion d'interactivité au sein d'un document. Le document n'est plus plat et statique mais possède en puissance plusieurs dimensions représentées par les hyperliens qu'il propose. La notion d'hyperlien est familière aux utilisateurs d'Internet. Dans une page apparaissant à l'écran, des groupes de mots sont mis en évidence (typiquement par un soulignement et une couleur différente du corps du texte) et, par un simple clic sur ces mots spéciaux, l'internaute se verra transporté dans un autre site du réseau ou dans une autre page du site visité. Avant de cliquer sur ce mot hyperlié, l'internaute peut observer dans les programmes de navigation classiques l'adresse du site qu'il se prépare à visiter. Ici aussi, ce dernier point est important puisqu'on peut en déduire que l'utilisateur peut raisonnablement savoir quel est le site qu'il s'apprête à visiter. Si l'adresse du site ne l'inspire pas (p.e. www.marketing.com ou www.nazi.org), il peut choisir de ne pas aller visiter ce site particulier. Notons que lorsque l'utilisateur tape une première adresse à visiter sur son programme de navigation (il peut l'avoir par exemple lue dans un journal), le même raisonnement peut être développé. Dans ces deux cas, il s'agit de liens visibles (l'utilisateur se rend compte qu'il va visiter un nouveau site) et explicites (l'utilisateur pose un acte positif pour accéder au site : soit en cliquant sur un hyperlien, soit en tapant en toutes lettres l'adresse du site à visiter).

3.3. Hyperliens invisibles et implicites

Il existe une autre catégorie d'hyperliens qui se distingue des précédents par le fait qu'ils sont invisibles et implicites (automatiques). Il est possible d'inclure, au sein d'un document HTML, un appel HTTP de téléchargement d'une image située sur un site quelconque, donc éventuellement différent du site contenant la dite page HTML. A l'inverse des hyperliens classiques définis supra, ces hyperliens externes ne nécessitent pas un acte positif de la part de l'internaute pour être activés. D'autre part leur exécution n'est guère visible que par l'apparition fugace du nom du site visité. Au niveau de l'affichage de la page elle-même, l'internaute ne peut qu'avoir l'impression que l'image affichée est issue du site qu'il est en train de visiter.

Au niveau technique, cette possibilité découle de certains choix techniques posés par les auteurs de programme de navigation. Par construction, lorsqu'un programme de navigation rencontre un appel HTTP au sein du document HTML qu'il est en train de visualiser, il ouvre une nouvelle session HTTP vers le site indiqué et gère le transfert (et l'affichage) de l'image parallèlement avec le chargement de la page en cours.

Il va de soit que ce chargement d'image via le protocole HTTP suit les règles protocolaires décrites supra et donc que certaines données invisibles seront transmises tant dans l'entête de requête issu du programme de navigation que dans l'entête HTTP.

3.4. Contenu invisible d'une requête HTTP

Cette requête est l'instrument classique de toute navigation sur Internet. Elle est générée par le programme de navigation dans trois cas de figure :

Types et caractéristiques des hyperliens sur Internet

Type de Lien	Explicite	Automatique
1. Frappe, par l'internaute de l'adresse d'une page à visiter	Oui	Non
2. Activation, par l'internaute, d'un hyperlien visible à l'écran	Oui	Non
3. Chargement, par le programme de navigation, d'une image incluse	Non	Oui

L'entête de la requête HTTP varie selon le type et la version du programme de navigation. Le tableau ci-après ventile le nom des données transmises selon quelques versions des navigateurs courants.

	Version du navigateur	Référant	Langue acceptée	Système d'exploitation	Cookie
Microsoft Explorer 3.0, 4.0 Fr	Oui	Oui	Oui	Oui	Oui
Netscape Navigator 3.01 Gold et 4.03	Oui	Oui	Oui	Oui	Oui

A titre d'exemple, le lecteur pourra visualiser les variables transmises par son programme de navigation sur Internet à l'adresse suivante :

"http://www.droit.fundp.ac.be/crid/privacy/strenv.exe?full".

Ceci provoque l'exécution d'un programme situé sur le site de l'université de Namur. Ce programme extrait les variables qu'il reçoit dans la requête HTTP pour les remettre dans le code HTML de la réponse, permettant ainsi leur visualisation par l'internaute. Ceci met également en évidence que le contenu d'une page HTML n'est pas nécessairement statique mais peut s'adapter, en temps réel, aux particularités de l'internaute, transmises via le navigateur, avant la construction de la réponse par le serveur. On a arrive donc à la notion de site polymorphique, c-à-d dont le contenu peut s'adapter, en temps réel, au profil détecté de l'utilisateur. Concrètement, il est actuellement réalisable de construire des sites renvoyant des pages de contenu différent à deux requêtes HTTP dont le contenu visible (HTTP ://serveur/document) est identique. Nous verrons plus loin comment cette potentialité est exploitée des millions de fois par jour par les entreprises de cybermarketing.

Certaines variables méritent quelques explications

- HTTP_USER_AGENT : il s'agit non seulement de la marque du navigateur, de son nom mais aussi du numéro de version principal (par exemple version 3 ou 4) mais aussi du numéro de version secondaire (par exemple version 3.01 ou 4.03). Dans ce nom de version secondaire se trouve aussi souvent mentionné la langue utilisée par le programme de navigation et la marque du processeur principal de la machine utilisée.
- HTTP_REFERRER : le référant est une variable crée par le programme de navigation dans lors de la mise en œuvre des deux derniers types d'hyperlien tels que décrits supra. Cette variable n'est donc pas transmise dans l'entête HTTP dans le cas où l'utilisateur tape lui-même l'adresse complète d'un nouveau site à visiter. Dans les deux autres cas cette variable transmise par la majorité des navigateurs contient l'adresse complète de la page ou se trouve situé l'hyperlien, qu'il s'agisse d'un hyperlien visible ou explicite sur lequel l'utilisateur a cliqué ou qu'il s'agisse d'un hyperlien implicite et invisible de téléchargement d'une image à partir d'un site éventuellement externe à celui en cours de visite.
- HTTP_LANGUAGE : Dans certains navigateurs, il est possible d'encoder les langues acceptées par l'internaute. Ces renseignements seront systématiquement transmis via le programme de navigation lors de chaque requête HTTP.
- HTTP_COOKIES : lorsqu'un navigateur reçoit la variable HTTP_COOKIES d'un site Internet, il stocke par défaut systématiquement son contenu sur le disque dur de l'internaute. Par la suite, cette valeur du cookie sera automatiquement transmise dans l'entête HTTP lors de chaque requête (explicite ou non) faite au site qui a stocké le cookie dans un premier temps. Le site peut alors modifier à nouveau cette valeur.. Mieux qu'une adresse IP dont la valeur peut être attribuée dynamiquement par les fournisseurs d'accès, les cookies permettent de marquer d'une manière invisible mais très efficace les machines utilisées par les internautes. Le phénomène des cookies est détaillé dans le chapitre suivant.

Notons que ces entêtes sont définies par des normes élaborées par le W3C qui contiennent des avertissements et des recommandations pour que le bavardage des navigateurs de nuise pas à la protection des données à caractère personnel de l'utilisateur.

4. Les cookies

Le phénomène des cookies a fait couler beaucoup d'encre. A mon sens pas parce qu'il représente le danger majeur d'atteinte aux données personnelles mais parce qu'il cristallise symboliquement dans l'imaginaire social cette fameuse inversion du paradigme client-serveur. Le phénomène des cookies demeure socialement dur à accepter, parce qu'il permet à un site distant et éventuellement inconnu (c'est le cas des millions de fois par jour sur le réseau grâce à des firmes de cybermarketing (cfr infra)) d'utiliser le disque dur du naviguant à son insu, en y stockant les données personnelles codées de l'internaute et en allant les y retrouver et les modifier à sa guise. En fait tout ce passe comme si chaque ordinateur du réseau possédait en puissance la libre disposition gratuite d'un espace disque chez chaque internaute alors que le surfeur moyen a conscience de louer un espace disque chez son fournisseur d'accès.

4.1. De quoi s'agit-il ?

Les cookies sont des informations persistantes enregistrées sur la machine du client Internet. Leur apparition n'a rien de "vicieux" (au sens juridique du terme) en soi : le serveur a besoin,

dans un certain nombre de cas, de pouvoir identifier qu'il a "en face" de lui, la même personne. Pour ce faire, il peut envoyer un cookie (ou plusieurs) au programme de navigation. Il s'agit d'un paquet d'informations contenant

1. le nom du cookie. Un même serveur peut stocker plusieurs cookies de nom différent;
2. sa valeur (souvent incompréhensible; peut être un index vers le parcours de l'internaute stocké chez le cybermarketeur);
3. sa date d'expiration. Dans la pratique cette date est souvent lointaine;
4. le nom du domaine auquel appartient l'ordinateur envoyant le cookie. Par la suite seuls les ordinateurs du même domaine pourront réaccéder à ce cookie.

Le navigateur recevant un cookie le stocke dans un fichier particulier situé sur la machine de l'utilisateur.

Par la suite, si la date d'expiration du cookie n'est pas atteinte, le navigateur le communiquera systématiquement lors de chaque requête HTTP (lien visible ou invisible) si le serveur appartient au même domaine (au niveau DNS) que le serveur ayant transmis la valeur initiale du cookie. Dans sa réponse l'ordinateur serveur peut modifier la valeur du cookie ou sa date d'expiration ou renvoyer des cookies supplémentaires. Il peut aussi le supprimer (en lui donnant une valeur nulle).

Une spécification précise des cookies peut être trouvée chez son inventeur.

4.2. Quels sont les risques ?

Certains risques ont été rapportés par le CERT : *" Les cookies sont des informations que le serveur stocke sur le disque de l'utilisateur sans que celui-ci soit prévenu. Le serveur peut également interroger le fichier de cookies qui se trouve sur le disque dur de l'utilisateur. Cette fonctionnalité s'apparente un peu à la petite mémoire que possèdent les Minitels. Mais, le serveur peut également récupérer l'historique des pages Web que vous avez consulter. Cette fonctionnalité est utilisée par les sociétés de marketing directe et "one-to-one marketing" afin de cibler des utilisateurs et d'enregistrer dans des bases des profils précis d'habitudes, réflexes, goûts et centres d'intérêts dans le but de générer dans les pages Web que consulte l'utilisateur des publicités très ciblées. Cela pose un important problème de respect de la vie privée."*

Bien que ces cookies existent depuis plusieurs années, les troisièmes versions des deux principaux programmes de navigation (Netscape Explorer et Microsoft Explorer) permettent de signaler "en temps réel" le passage de cookies. A ce moment, l'utilisateur peut accepter ou refuser l'enregistrement du cookie. C'est un premier pas vers un début de protection des données. Il n'a valeur que de symbole et n'est pas réellement efficace pour un certain nombre de raisons.

1. Par défaut, la protection n'est pas activée et les connaissances techniques nécessaires pour l'activation ne sont pas évidentes. Pour être averti du passage des cookies, l'utilisateur peut effectuer un paramétrage du navigateur.
2. Les avertissements du passage de cookies deviennent à la longue épuisants : l'utilisateur doit chaque fois cliquer "oui" ou "annuler" dans une fenêtre qui vient interrompre sa navigation.
3. Les cookies demeurent généralement inintelligibles car codés.

4. Le choix d'accepter ou de refuser est mince. Il faudrait, par exemple, que l'utilisateur puisse déterminer une durée de vie limité, c-à-d modifier le cookie.

La technique des cookies permet donc de marquer un utilisateur particulier avec certaines données qui le concernent et dont la signification est tout à fait hermétique. Il est techniquement possible d'inclure dans ces cookies des données sensibles que l'on aurait pu déduire de certaines réponses à des formulaires envoyés précédemment. En d'autres termes, si un site Internet révisionniste (moyennant une programmation adéquate) arrive à la conclusion qu'un utilisateur est juif, il pourra coller une étoile codée sur le dos de cet utilisateur de telle manière que chaque site de sa famille DNS puisse avoir vent de cette caractéristique, avant d'afficher une quelconque de ses pages

4.3. Les parades techniques.

Il est tout d'abord intéressant de voir comment l'industrie informatique, après avoir créé les cookies, a timidement tenté de limiter les risques liés à leur utilisation.

Type de navigateur / Version	Microsoft Explorer	Netscape Navigator
Version 1	Pas de cookies	
Version 2	Première implémentation	
Version 3	Droit d'opposition de base (impraticable)	
Version 4	• accepter chaque cookie • refuser chaque cookie • voir au cas par cas	• accepter chaque cookie • refuser chaque cookie • refuser dans des hyperliens invisibles • être averti ou non de l'arrivée d'un cookie

Notez l'amélioration du navigateur de Netscape qui permet de s'opposer systématiquement au cookies issus de firmes de cybermarketing (hyperliens invisibles). Le mécanisme d'opposition des cookies ne joue que dans un sens : celui de la réception. Autrement dit, l'internaute ne sait pas et ne peut s'opposer à ce qu'un cookie préalablement reçu soit systématiquement renvoyé au site émetteur.

D'autres parades existent.

- Récemment, on a vu apparaître des programmes tueurs de cookies et il est aussi possible, pour un utilisateur averti de supprimer le fichier qui contient tous les cookies.
- Une autre solution réside dans l'utilisation d'un proxy serveur. Il s'agit d'un serveur pare-feu qui sert d'intermédiaire entre l'internaute et le réseau, effectue les requêtes HTTP en son nom et lui communique les résultats. Certains proxy serveurs ont cette capacité de modifier l'entête HTTP en se faisant passer pour un navigateur de tel ou tel type et de filtrer les cookies. Cette solution est séduisante parce que rien ne s'oppose à ce que le proxy lui-même soit atteint par les cookies et devienne l'objet d'un profilage qui devient le profilage moyen des utilisateurs du proxy qui reste anonymes, ménageant ainsi, dans une certaine mesure, les intérêts des firmes de cybermarketing.

Toutefois, le filtrage par l'intermédiaire de proxy est aujourd'hui battu en brèche par la présence du langage JavaScript qui permet l'utilisation de cookies dans le corps du langage HTML (voir infra) et non plus dans l'entête HTTP.

4.4. Pour en finir avec les cookies

Toute utilisation de cookies n'est pas à proscrire. Certains cookies peuvent être extrêmement précieux pour faire gagner du temps à l'utilisateur (par exemple pour le guider directement dans une page de sa langue), toutefois leur usage, dans le chef des responsables de traitement devrait se référer aux grands principes de la directive 95/46 CE et, en particulier

1. Informer lors de la collecte : avant de stocker un cookie sur le disque dur de l'internaute le site doit
    - dévoiler son identité (pas seulement virtuelle (adresse internet) mais aussi légale)
    - décliner les finalités pour lesquelles ce cookie sera utilisé
    - signaler l'existence d'un droit d'accès

1. Permettre un droit d'accès : il est particulièrement aisé, vu le mécanisme des cookies de prévoir une page particulière du site où un internaute pourra connaître la signification du cookie qu'il transporte puisque ce cookie est par définition envoyé systématiquement au site visité.
2. Permettre un droit d'opposition : ce point est plus difficile à trancher dans la mesure où la finalité n'est pas nécessairement une finalité de marketing et on peut se demander s'il est légitime que certains sites fonctionnent uniquement sur base de cookies.

De son côté l'industrie informatique doit fournir des produits sécurisés qui protègent de façon raisonnable l'intimité de consommateur surfeur et notamment ses préférences en matière de protection des données à caractère personnel. En particulier et par défaut, les programmes de navigation devraient permettre un filtrage et une modification interactive des différents cookies, sans en informer les sites en envoyant. Par exemple, les programmes de navigation devraient permettre une date de péremption interne alors que le cookie renvoyé porterait toujours la date limite fixée par le site émetteur; dès le dépassement de la date interne fixée par défaut par le navigateur (éventuellement affinée par l'internaute), le cookie se verrait effacé. Un tel comportement se révèle particulièrement efficace et reste invisible au réseau.

5. La captation des mots-clés tapés sur les moteurs de recherche

Bien souvent, les défenseurs de la privacy peuvent être perçus comme des paranoïaques, mélangeant ce qui est possible et ce qui est effectivement réalisé sur le terrain. L'objet de cette section n'est dès lors pas de se perdre dans le techniquement réalisable mais d'analyser le matériellement réalisé. Pour ce faire nous détaillerons, pas à pas, la technique de cybermarketing utilisée aujourd'hui par la société Double Click située aux Etats-Unis. Cette société se vante sur son site Internet (" http ://www.doubleclick.net ") de fournir plusieurs millions de bannières publicitaires par jour, d'une manière ciblée. Cela est-il réellement possible ? Posons la question différemment.

Est-il possible qu'une société puisse centraliser le " clickstream " d'un internaute particulier ? Que cette même société puisse adapter les bannières envoyés à cet internaute selon le profil

détecté, puisse sans cesse affiner ce profil et ce à l'insu de l'internaute moyen ? La réponse est oui. Le but de cette section est de montrer comment cela est effectivement réalisé aujourd'hui sur le réseau Internet, plusieurs millions de fois par jour. Pour ce faire nous détaillerons, pas à pas une interrogation effectué par un internaute à l'aide du moteur d'AltaVista.

5.1. Connexion à un moteur de recherche

L'utilisateur tape dans la fenêtre du programme de navigation : " http ://www.altavista.com ". Dans la requête effectuée par le programme de navigation, les données suivantes sont transmises vers le site d'AltaVista :

- type de navigateur et système d'exploitation (pourquoi faire ?)
- langue acceptée (si communiquées au programme de navigation)
- adresse TCP/IP (indispensable pour la réponse)

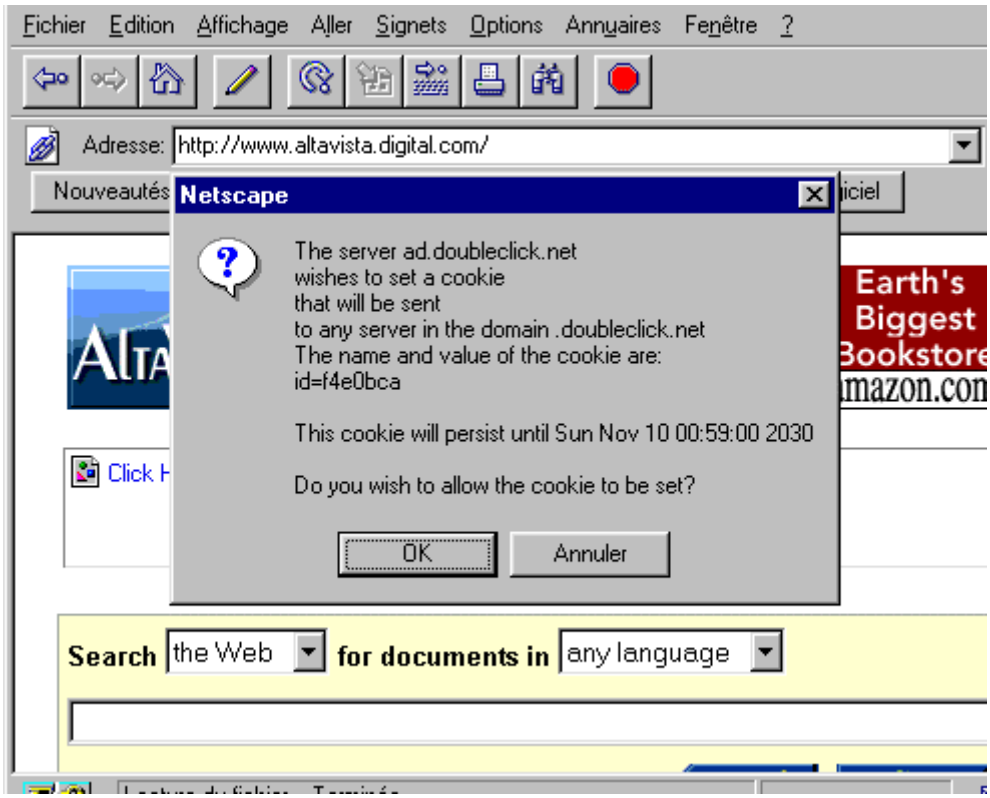
Le serveur d'Altavista répond en transmettant l'entête HTTP et le code HTML de la home page d'AltaVista

Le programme de navigation commence à dessiner à l'écran le code HTML de la page qu'il reçoit d'AltaVista.

Dans cette page se trouve un hyperlien de type 3 (implicite et automatique) vers le site de DoubleClick. Le programme de navigation effectue alors une requête HTTP, adressée au site Double Click et lui communique donc :

- type de navigateur et de système d'exploitation
- cookies stockées sur le disque dur (si le site de DoubleClick a déjà été visité par la machine)
- langues acceptées (si communiquées au programme de navigation)
- le référant, c-à-d, l'adresse de la page en cours (" http ://www.altavista.digital.com ")

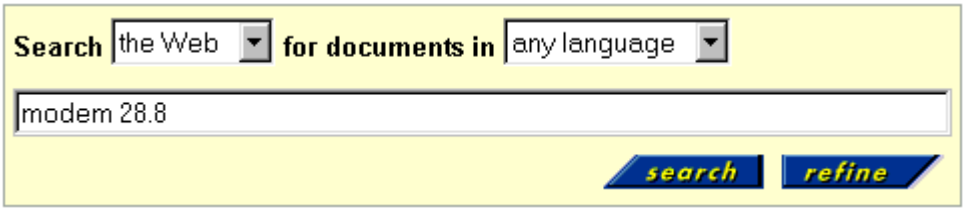
Si l'avertissement de la réception de cookies est activé, on peut observer l'écran suivant :



Cliquez sur OK ou sur Annuler. Peu après, l'affichage de la page est complet (y compris la bannière publicitaire issue téléchargée du site de Double Click).

5.2. Recherche de certains mots-clés.

Imaginons qu'à ce moment, l'utilisateur souhaite effectuer une recherche sur les mots-clés suivants : " modem " et " 28.8 ". Il lui suffit de taper ces deux mots dans la zone prévue à cet effet et d'appuyer sur le bouton recherche.



A ce moment le programme de navigation effectue la requête suivante à AltaVista :

- http://www.altavista.digital.com/cgi-bin/query?pg=q&what=web&kl=XX&q=modem+28.8&search.x=53&search.y=10
- ...

Comme vu précédemment, le code HTML de la réponse d'Altavista comporte une image incluse téléchargée du site de Double Click par le programme de navigation de l'internaute. Lors de l'affichage de la réponse, le programme de navigation effectue une requête invisible chez Double Click avec, notamment, la référence précise de la page principale en cours de chargement. Dans l'entête HHTP de la requête faite par le programme de navigation, on trouvera :

- HTTP_REFERRER = http://www.altavista.digital.com/cgi-bin/query?pg=q&what=web&kl=XX&q=modem+28.8&search.x=53&search.y=10

Cela signifie, qu'avant de renvoyer une bannière publicitaire, le site de Double Click est informé d'une recherche basée sur les mots-clés " modem " et " 28.8 ". Cela permet à cette société de marketing de renvoyer au navigateur une première bannière véritablement personnalisée (en l'espèce une publicité de Texas Instrument pour des modems 56 Kbps). Cette bannière est affichée en même temps que les résultats de la recherche.

Par ailleurs, si le programme de navigation a été configuré pour détecter la réception de cookies dans l'entête HTTP, on peut observer que le site de Double Click renvoie un nouveau cookie.

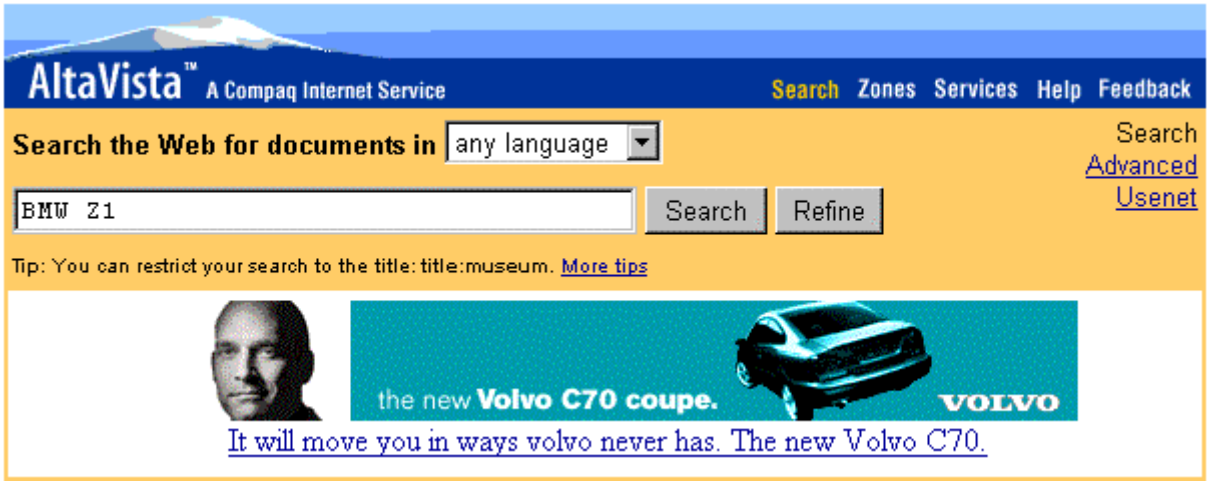
Ce cookie permet d'affiner le profilage de l'internaute et renvoie à un nouveau numéro de référence. Notons au passage la date de péremption de cette identification qui persistera jusqu'en l'an 2030.

5.3. Sélection d'une publicité

Imaginons un bref instant que l'internaute clique sur la publicité des modems rapides proposés par Texas Instrument. L'analyse préalable de l'hyperlien lié à la bannière publicitaire de TI montre le lien suivant :

- http://ad.doubleclick.net/jump/altavista.digital.com/result_front;kw=modem+28+8;ord=173757616

Cela signifie que l'activation de ce lien par l'internaute va avoir pour effet de l'envoyer sur le site de Double Click. Or la pratique montre que le clic provoque le chargement de la page http://www.ti.com/sc/x2 . Ceci est du à une dernière caractéristique cachée dans le fonctionnement des navigateurs. Il est possible pour un site d'incorporer dans le code HTML d'une page les instructions pour que le navigateur charge automatiquement une autre page, éventuellement d'un autre site. L'avantage de Double Click dans cette manière de procéder est de savoir que l'internaute a manifesté un intérêt pour une bannière publicitaire particulière. Cela permet d'affiner son profilage et de proposer des produits dans la même gamme.



Ce profilage peut aussi poser des problèmes déontologiques importants. Le cybermarketeur ne proposera jamais que des bannières pour ses clients. Ainsi on peut se trouver dans des situations étranges où un consommateur s'adressant à un moteur de recherche pour trouver un produit X (BMW) pourrait se voir proposer un produit Y (Volvo) concurrent. Cette concurrence est-elle loyale ?

5.4. Conclusion

Grâce à la complicité silencieuse des programmes de navigation, une société comme Double Click peut savoir, pour chacune des centaines de pages appartenant aux milliers de sites abonnés, le parcours de n'importe quel internaute au sein de ces pages. Ceci est causé par le fait que les programmes de navigation transmettent un certain nombre de caractéristiques sur le réseau, non seulement à l'insu de l'utilisateur mais sans que celui-ci puisse réellement s'y

opposer. Plus précisément, Double Click peut collecter le click stream des utilisateurs des sites affiliés :

- Adresse IP,
- Click stream antérieur (cookie) de l'internaute,
- Mots-clés tapés sur les moteurs de recherche affiliés,
- Bannières publicitaires déjà envoyées,
- Bannières publicitaires ayant reçus une réponse positive.

Tous ces renseignements peuvent évidement être datés à la minute près. L'envoi d'un cookie sur la machine de l'internaute permet en outre de pouvoir reconnaître un internaute particulier même si celui-ci se connecte à des heures différentes avec plusieurs fournisseurs.

6. Applets Java et JavaScript

6.1. Les Scripts JavaScript

6.1.1. De quoi s'agit-il ?

Lorsqu'un navigateur reçoit une page d'un site, cette page peut contenir certaines instructions en JavaScript. Ces instructions sont interprétées en temps réel par un interpréter JAVA intégré dans ou appelé par le navigateur. Typiquement l'exécution de ces microprogrammes écrits en langage JavaScript a pour effet de dynamiser la page affichée en effectuant certaines animations ou en liant l'exécution de certains microprogrammes à certains boutons particuliers présent dans la page.

6.1.2. Quels sont les risques ?

L'apparition du langage JavaScript, automatiquement téléchargé et exécuté sur les navigateurs standards du marché est une étape supplémentaire de l'inversion du paradigme client-serveur. Dans le phénomène des cookies nous avons vu qu'un utilisateur distant et anonyme dans le réseau pouvait inscrire sur le disque dur de l'internaute certaines données personnelles le concernant. Ici une étape supplémentaire est franchie puisqu'il ne s'agit plus d'écrire sur le disque dur de l'internaute mais de le contraindre à exécuter sur sa propre machine et toujours à son insu un programme conçu et stocké sur un site extérieur.

Il semble certain à l'heure actuelle que plusieurs trous de sécurité existaient dans le langage JavaScript lorsqu'il était interprété par les versions 2.0 et 2.01 du navigateur Netscape. Il semblerait que certaines pages auraient eu pour effet de transmettre au serveur appelé des fichiers se trouvant sur le poste client (à l'insu, bien sûr de l'utilisateur), voire de les modifier. Toutefois aucune attaque de ce type n'a été rapportée, bien que la possibilité technique ne semble plus faire aucun doute au niveau théorique. Si certains problèmes de protection des données ont été résolus dans les versions ultérieures, il n'en demeure pas moins que de nouveaux problèmes surgissent sans cesse.

Il n'en reste pas moins que les scripts Java s'exécutant sur la machine de l'internaute peuvent évidement être beaucoup plus "privacides" que les simples documents HTML transmis via HTTP. Notamment, le langage JavaScript permettrait de s'affranchir du filtrage opéré par les proxy serveurs au niveau des cookies et du masquage de l'adresse IP.

6.2. Les applets Java

6.2.1. De quoi s'agit-il ?

Il s'agit d'une évolution des scripts JAVA à ne pas confondre avec ces derniers. Dans le cas des applets, les instructions JAVA sont préalablement traduites dans un pseudocode par les soins d'un précompilateur installé sur le serveur. Lors de l'exécution d'un applet, le navigateur charge le Pcode lié à l'applet et l'exécute par le biais d'un interpréteur.

6.2.2. Quels sont les risques ?

La différence fondamentale entre les scripts et les applets est que les premiers sont transmis en langage clair au navigateur tandis que les seconds sont préalablement converti en PCODE et donc inintelligible, même pour un programmeur moyen. Dans les deux cas de figure, la transmission s'effectue de manière souterraine mais, ici encore, la possibilité existe, dans les navigateurs récents d'inhiber l'exécution des scripts ou des applets.

Toutefois une telle solution a pour effet de bord de désactiver toute l'animation possible des pages HTML, rendant ainsi leur consultation plus morne, voire impossible.

7. Conclusion et questions

7.1. L'inversion progressive du paradigme client-serveur

L'ordre dans lequel ces sections (risques liés à TCP/IP, bavardage du navigateur, hyperliens invisibles, captation des mots-clés tapés sur les moteurs de recherche, Java) ont été présentées n'est pas innocent. Il s'agit d'un ordre plus ou moins chronologique mais, aussi et surtout, d'une gradation dans l'opacité des traitements effectués et, parallèlement, dans la perte, par l'utilisateur du navigateur, de la maîtrise tant des traces qu'il laisse lors de la consultation que de l'information elle-même à laquelle il accède.

Le bavardage des navigateurs, les hyperliens invisibles, les cookies, la captation des mots-clés tapés sur les moteurs de recherche, les scripts et applets Java permettent une dynamisation et un polymorphisme de l'espace informationnel dans lequel l'utilisateur évolue.

Par ailleurs, ces techniques ne sont exclusives l'une de l'autre. Un navigateur montrant une même page HTML peut, par exemple, télécharger des hyperliens invisibles avec stockage et réception de cookies, exécuter quelques instructions en JavaScript et un applet JAVA.

L'imaginaire social des débuts d'Internet pourrait probablement s'apparenter à une gigantesque bibliothèque dont les pages des ouvrages seraient reliées entre elles par des hyperliens. L'utilisateur naviguait alors dans ce cyberespace à l'aide de sa souris.

Si l'imaginaire social n'a guère évolué, la réalité technique est aujourd'hui très différente de cette image. Il ne s'agit plus de naviguer dynamiquement dans un nuage d'informations statiques stockée sur des serveurs passifs. Aujourd'hui les sites Internet peuvent tracer le comportement d'un utilisateur particulier et via diverses techniques officielles ou officieuses, injecter des données et même des programmes sur les ordinateurs individuels reliés au réseau, à l'insu de la personne connectée. Ces programmes et ces données acquièrent alors une

existence propre, un fonctionnement autonome et clandestin au service d'intérêts parfois obscurs mais souvent mercantiles de sociétés inconnues.

7.2. L'adresse IP est-elle une donnée à caractère personnel sur Internet ?

Selon la directive 95/46 CE, il s'agit d'informations se rapportant à une personne physique identifiée ou identifiable. Une question fondamentale se pose : le profilage d'un internaute est-il une donnée à caractère personnelle ?

Un récent arrêt de la cour de cassation française en juin 1995 se révèle extrêmement éclairant dans le cas qui nous occupe. Les faits sont les suivants. Une banque recourt à la segmentation comportementale de ses clients et cette segmentation comprend des catégories telles que "ne s'améliorera pas avec le temps", "laxistes", "modernistes", "difficiles à convaincre", "méfiants". Chaque client se voit attribuer un code et le personnel reçoit des instructions quant à la manière d'aborder les clients d'un segment particulier. La banque a soutenu devant le conseil d'Etat que le segment n'est pas une donnée à caractère personnel ("nominative"). Le conseil d'Etat a répondu que "*si le segment ne constitue pas à lui seul une information nominative, il le devient dès lors qu'il est associé à une personne identifiée ou identifiable".* En l'espèce le programme informatique était fait de telle sorte que la frappe d'un numéro de compte particulier au guichet provoquait l'apparition simultanée du code de segmentation lié à la personne titulaire de ce compte.

L'argument des firmes de cybermarketing est semblable. Elles soutiennent qu'elles ne sont pas en mesure d'identifier les internautes puisqu'elles ne collectent pas leur adresse électronique. Toutefois, il y a lieu d'apprécier ce critère d'identifiabilité, d'une part, au regard du monde virtuel dans lequel l'internaute évolue et, d'autre part, en se référant à l'objectif de la loi.

Dans le monde virtuel, l'internaute s'identifie par une adresse IP. Celle-ci peut être comparée à un numéro de téléphone à cette nuance près que l'identification d'appel n'est pas désactivable. Cette caractéristique découle du fonctionnement par paquet du réseau TCP/IP sur lequel Internet est basé. Certains internautes, connectés par ligne louée (typiquement c'est le cas des moyennes et grosses entreprises) possèdent une adresse IP fixe. Toutefois, d'autres internautes (principalement les utilisateurs résidentiels) louent un accès à Internet (une cabine téléphonique) chez leur fournisseur d'accès, ce qui signifie que, à chaque connexion, un internaute particulier se voit attribuer un numéro IP unique qui ne changera pas durant toute la durée de la connexion même si un même numéro IP peut successivement être attribué à plusieurs internautes différents lors de sessions différentes. Il n'en reste pas moins, -et ce point s'avère décisif- qu'une adresse IP est un identifiant unique au moment du traitement des données sur le réseau et notamment au moment où l'internaute reçoit des bannières publicitaires ciblées sur base des données récoltées sur lui depuis l'ouverture de se session.

Si l'adresse IP ne permet pas, dans tous les cas de figure, d'identifier à coup sûr un internaute particulier à des moments différents, en est-il de même pour le cookie ? Contrairement à l'adresse IP d'un utilisateur résidentiel le cookie est stable, de connexion en connexion. A l'opposé de l'adresse IP, il reste néanmoins possible que le cookie ne reflète que l'itinéraire particulier d'un internaute et ses intérêts pour un certain nombre de produits et qu'il ne soit pas le seul dans une catégorie donnée; en d'autres termes, le cookie n'est pas nécessairement identifiant au sens où il ne permet pas de distinguer un internaute particulier de tous les autres. Analysons ce dernier cas, le moins favorable pour les défenseurs de la protection des données.

Un internaute bardé de cookies se présente à un guichet électronique (un site internet). Ces cookies racontent à son insu son histoire de surfeur, les sites qu'il a été voir, les bannières publicitaires qu'il a reçues, celles auxquelles il a réagi favorablement, les mots-clés qu'il a tapé sur les moteurs de recherche,... Derrière le guichet électronique le site Internet connaît toutes ses caractéristiques. Cet internaute est-il identifiable ?

La question est souvent posée de cette manière mais me semble mal posée. Dans le monde virtuel d'Internet, le simple fait de connaître l'adresse IP qui identifie au moment du traitement un internaute particulier et le différencie de tous les autres signifie que cet internaute est identifié, tout comme l'identification d'appel sur le réseau téléphonique permet d'identifier un abonné particulier. Identifier une personne pourrait donc signifier pouvoir la distinguer des autres au moment du traitement. Dès lors, ce n'est pas le cookie lui-même qui permettrait d'identifier la personne (il n'est d'ailleurs pas nécessairement identifiant puisque que l'on admet dans le cas présent que plusieurs personnes possèdent le même cookie). Le cookie ne sert qu'a révéler des données supplémentaires concernant un internaute particulier préalablement identifié par son adresse IP, unique dans le monde entier au moment du traitement.

Par ailleurs, si l'on se réfère à la "finalité" de la Directive 95/46/CE, un des objectifs de cette directive est de permettre à l'individu d'avoir une maîtrise sur son image informationnelle (savoir qui fait quoi avec quelles données, pouvoir y accéder, les rectifier ou s'opposer à leur utilisation). Si l'on accepte pas qu'un internaute particulier est identifié dès lors qu'il possède un numéro unique et public (en l'espèce son adresse TCP/IP, comparable à un numéro de téléphone dans un réseau où l'identification d'appel serait systématique), la Directive censée protéger les données personnelles d'un individu offrirait le flanc à des décisions discriminantes sans qu'aucun des concepts habituels de la protection des données ne puissent s'appliquer. Il deviendrait par exemple possible d'inscrire dans un cookie des données religieuses ou médicales d'un internaute particulier sans qu'aucun recours soit possible, par exemple contre la décision virtuelle d'octroi ou non d'une assurance-vie. Tout comme l'employé d'une banque adopte un comportement différent et discriminant suivant la personne qui se présente à son guichet, selon le profilage qui lui a été attribué.

En fait, on peut raisonnablement conclure, en regard de l'article 2 a) de la Directive que l'adresse IP est un *numéro d'identification* et que les cookies sont des *élements spécifiques propres à l'identité économique, culturelle et sociale.* Le bavardage du navigateur (type de navigateur, type de machine et de système d'exploitation (version et langue), langue, page préalablement visitée, ...) associé à l'adresse IP dans l'entête HTTP devient lui aussi un élément propre à l'identité économique (version récente du navigateur, processeur récent sont des indices de richesse) culturelle et sociale (quelles pages sont visitées, quel est l'ordinateur (PC ou Macintosh) utilisé, quel est le navigateur utilisé,...). Les hyperliens invisibles sont la technique par laquelle le bavardage des navigateurs est transmis à des tiers à la communication entre un internaute et le site qu'il désire visiter.

On pourrait s'étonner que l'identité dans un monde virtuel ne passe pas par la connaissance des nom et prénom et adresse, ni même par la connaissance de l'adresse électronique, ce dont se défendent les firmes de cybermarketing. Si, dans le monde physique, de tels renseignements sont indispensables pour pouvoir effectuer du marketing direct, la technologie actuelle et l'alchimie réalisée actuellement sur base du bavardage des navigateurs, des hyperliens invisibles et des cookies permettent de bombarder de bannières publicitaires ciblées un internaute particulier sur la seule base de son adresse IP. Le cookie n'est pas

indispensable et ne sert qu'à maintenir le profilage entre deux connexions du même internaute. L'adresse IP seule suffit à assurer le profilage pour une session particulière d'un internaute résidentiel. A fortiori suffit-elle pour assurer le profilage d'un internaute possédant une adresse IP fixe.

7.3. Les éditeurs de navigateurs responsables...juridiquement ?

S'il semble acquis que l'industrie informatique et plus particulièrement les éditeurs de programmes de navigation possèdent une large part de responsabilité effective quant au développement exponentiel des techniques de fichages des internautes, cette interprétation ouvre la porte à la responsabilité juridique de ces éditeurs dans la mesure où, selon 1 d) de la Directive précitée, le concepteur pourrait être considéré comme *"une personne morale qui seule ou conjointement avec d'autres détermine... les moyens du traitement"*...

7.4. Réglementer les autoroutes de l'information sans contrôler les véhicules ?

Il faut bien constater que la technologie d'Internet, telle qu'incorporée dans les programmes de navigation, n'est pas l'alliée de la protection des données personnelles de l'internaute sur le réseau. Bavardage du navigateur, cookies, hyperliens invisibles, java sont autant de techniques qui permettent aux sociétés de cybermarketing de capter silencieusement sur une échelle planétaire le comportement détaillé de chaque internaute en particulier. Ces techniques permettent l'inversion du paradigme client-serveur : malgré une apparence inchangée le poste de travail de l'internaute est devenu pour le réseau le serveur de ses données personnelles. Bien plus, les firmes de cybermarketing effectuent actuellement une forme de censure du réseau en décidant des publicités que l'internaute peut voir et de celles qui ne seront jamais montrées, sur base du profil détecte. Techniquement, rien ne s'oppose à ce que ce polymorphisme utilisé actuellement exclusivement à des fins commerciales de cybermarketing ne soit utilisé à d'autres fins. On peut imaginer un parti politique ou un site révisionniste qui se présenterait de manière tout à fait différente, selon le type d'internaute qui se présenterait.

En matière de cybermarketing, les grandes sociétés sont essentiellement américaines. Actuellement elles ne respectent pas ou peu les grands principes de protection des données : information lors de la collecte, possibilité d'un droit d'opposition, exercice d'un droit d'accès et de correction, légitimité,...

Si l'Europe a peu de prises contre ces multinationales étrangères, il apparaît que ces atteintes à la protection des données ne peut s'accomplir qu'avec la collaboration effective des programmes de navigation installés sur le territoire européen. La directive peut, selon nous, s'appliquer non seulement à ceux qui traitent des données en Europe mais aussi à ceux qui conçoivent les outils permettant de traiter des données personnelles et de les exporter subrepticement vers un territoire non européen. Encore faut-il qu'il existe une volonté politique de réglementer le marché des navigateurs au regard des exigences de la directive 95/46 et, plus largement au regard des exigences de la protection du consommateur qui est en droit de s'attendre à un produit loyal qui ne transfère pas à son insu sur le réseau les détails de son comportement.

Pour atteindre un niveau satisfaisant de sécurité sur les autoroutes, on a édicté un code de la routes pour les usagers, une réglementation précise sur la manière de concevoir les entrées et sorties ainsi que les connexions aux diverses échoppes qui se trouvent sur les autoroutes. Par

ailleurs les véhicules sont soumis à une réglementation de plus en plus stricte (ceintures de sécurité, tests de collision, airbag, ABS, ...). Ces règles s'appliquent également aux véhicules étrangers immatriculés en Europe.

Sur les autoroutes de l'information, il n'y a guère actuellement que les constructeurs de navigateurs (véhicules permettant de se déplacer sur les autoroutes de l'information) qui ne soient soumis à aucune réglementation. Tout comme il est naïf de croire que la sécurité des usagers des autoroutes peut être atteinte avec des véhicules non sécurisés, il semble aujourd'hui utopique de considérer que la protection des données personnelles sur Internet peut être atteinte sans un contrôle minimal de l'industrie informatique qui construit les outils manipulant de manière invisible les données à caractère personnel sur Internet.