Un point de vue technique sur la loi "Internet et création"

Fabrice Le Fessant

Expert des systèmes distribués et des réseaux pair-à-pair
Chercheur à l'INRIA – Saclay - Île-de-France
(Institut National de Recherche en Informatique et Automatique)
Enseignant en informatique à l'École Polytechnique

16 février 2009

La loi "Internet et Création" introduit dans la loi française le système dit "de riposte graduée", qui permettra aux ayants-droit de collecter automatiquement sur les réseaux pair-à-pair les adresses Internet des internautes opérant des téléchargement illégaux, puis de les transmettre à l'Hadopi, organisme chargé de contacter les particuliers et, en dernier recours, de les sanctionner.

Cette loi est inadaptée à un certain nombre de caractéristiques techniques des réseaux et d'Internet, et va introduire de nouveaux problèmes, sans résoudre ceux qu'elle vise.

Dans ce document, trois de ces problèmes vont être détaillés :

1. Le système de la riposte graduée suppose que l'internaute sait qu'il commet un acte illégal en téléchargeant une œuvre protégée. Or, la nature même de l'Internet cache souvent le caractère illégal de l'acte de téléchargement, l'internaute ne pouvant découvrir qu'il vient de commettre un délit qu'après l'exécution de ce délit.
2. Le système d'identification du coupable, fondé sur l'adresse de la connexion Internet, implique une sécurisation de l'installation informatique, hors de portée des simples particuliers, comme d'ailleurs de la plupart des entreprises.
3. Le système de repérage automatique des téléchargeurs illégaux deviendra rapidement obsolète, en poussant les internautes vers des systèmes pair-à-pair qu'il sera impossible de contrôler.

1 L'impossibilité de différentier *a priori* transfert légal et transfert illégal

Sous une interface élémentaire, une monstrueuse complexité

Internet est un univers complexe, fait de l'interconnexion de milliers de réseaux permettant à des millions d'ordinateurs de discuter entre eux. Internet n'aurait pas connu son succès actuel sans un effort considérable de ses concepteurs (chercheurs, ingénieurs et administrateurs) pour cacher cette complexité et offrir aux particuliers, même débutants en informatique, des interfaces simples et intuitives pour naviguer dans cet univers et y dénicher les informations recherchées.

Une navigation puissante, mais imprévisible

Un tel travail de simplification ne va pas sans désavantages. Ainsi, il est aujourd'hui difficile, voire impossible, même pour les experts du domaine, de prédire le résultat d'un simple clic sur une page web, sans avoir analysé les sources de celle-ci. Un tel travail ne peut être exigé des simples internautes.

Avec le web (ou World Wide Web) inventé en 1992 au CERN, l'internaute a pris l'habitude de se promener de document en document, d'un simple clic sur un lien hypertexte, le menant par exemple de son journal quotidien en ligne préféré vers un autre site web à l'autre bout de la planète, de ce dernier à un autre, et ainsi de suite. Seule l'analyse des liens hypertextes peut permettre, *a priori*, de deviner où l'utilisateur sera transféré, et requiert une expertise en informatique peu commune. Avec l'arrivée du Web 2.0 et de pages web beaucoup plus interactives, cette analyse des liens est devenue impossible, même pour les experts.

Or, un clic sur une page web est suffisant pour déclencher le téléchargement d'un document protégé, aussi bien directement sur le web que par l'entremise d'un logiciel pair-à-pair.

Agir suivant le nom, être sanctionné suivant l'identifiant

Ce problème provient de la nature même d'Internet, qui fait à la fois communiquer entre eux des ordinateurs et des humains. Pour désigner un document, qu'il soit protégé ou pas, les ordinateurs utilisent entre eux un identifiant, souvent basé sur le contenu du document. Les humains utilisent eux des noms, sans nécessairement de rapport avec le contenu du document. Aussi, les systèmes de repérage utilisés par la riposte graduée détecteront

les infractions suivant l'identifiant du document téléchargé, alors que les internautes téléchargeront en fonction du nom du document.

Ainsi, si un logiciel pair-à-pair demande à un humain de télécharger l'un des deux documents suivant :

- ed2k ://|file|Bétisier de mes vacances au ski 2009.avi |723302400|81CD3B9A6DBDB72455CF3335E2BC9C1B|/
- ed2k ://|file|Jean.Paul.Belmondo.-.Une.chance.sur.deux.-. DVD.Rip.-.5.02pro.-.By.Timal.avi |723302400|81CD3B9A6DBDB72455CF3335E2BC9C1B|/

un utilisateur averti pourra se douter que le second document (dont le nom contient le titre d'un film récent de Jean-Paul Belmondo) est probablement protégé et donc son téléchargement illégal, tandis que le premier document est probablement l'œuvre d'un particulier. Un ordinateur, lui, saura que ces deux documents sont semblables, car leur identifiant est identique (81CD3B9A6DBDB72455CF3335E2BC9C1B) , et que soit le téléchargement des deux est légal, soit le téléchargement des deux est illégal.

Ce problème est encore plus exacerbé sur le web. Ainsi, entre deux liens indiquant "mes dernières chansons" et pointant vers les adresses suivantes :

- http ://www.piratebay.com/429874297429/madonna-last-hit.mp3
- http ://85.17.142.41/429874297429/4987AB868BEF.mp3

l'internaute peut se douter que le premier aboutira à un téléchargement illégal, mais sera incapable de juger de la légalité du second, jusqu'à ce que le téléchargement soit terminé et que l'internaute puisse écouter le contenu du document.

Conclusion

La riposte graduée établira donc le caractère délictueux de l'acte commis par l'internaute, mais ne prouvera aucunement l'intention de l'internaute de commettre cet acte délictueux, ce dernier pouvant, en toute bonne foi, croire ce téléchargement parfaitement légal.

2 L'impossibilité de sécuriser les installations informatiques domestiques

Le rêve inatteignable de la sécurité informatique

Pour identifier l'internaute à l'origine d'un téléchargement délictueux, le système de la riposte graduée utilise l'adresse (IP) de la connexion Internet, correspondant au point de raccordement à Internet de l'ordinateur sur lequel le téléchargement a eu lieu.

Or, relier l'observation d'un téléchargement illicite au possesseur de la connexion Internet suppose une confiance totale dans la capacité d'un particulier à sécuriser complètement son installation informatique domestique. Quand nombre d'entreprises emploient à plein temps un expert (le RSSI, Responsable de la Sécurité des Systèmes d'Information) pour sécuriser leur réseau informatique et ne parviennent pas, malgré cela, à une sécurité totale, une telle supposition dans le cas des particuliers est absurde et inatteignable.

Le logiciel fiable n'existe pas

La première faille de sécurité est souvent le logiciel. Avec l'augmentation de la puissance de calcul des ordinateurs, les logiciels se sont considérablement complexifiés. Certains logiciels atteignent des millions de lignes de code source, suites d'instructions expliquant à l'ordinateur comment combiner des milliers de méthodes différentes (les algorithmes) pour répondre aux désirs de l'utilisateur. Il n'est jamais inutile de rappeler qu'il s'agit de la raison pour laquelle l'idée du brevet logiciel est absurde, tant l'implantation d'une idée dans un logiciel ne peut se faire qu'en combinaison de milliers d'autres idées, dont il serait impossible de gérer la propriété intellectuelle.

Depuis quarante ans, des milliers d'ingénieurs et chercheurs travaillent à améliorer la fiabilité du code écrit par les programmeurs, au moyen d'outils de vérification extrêmement performants. Pourtant, le logiciel est encore bien loin d'avoir atteint la fiabilité qui le protégerait des attaques malveillantes. Ainsi, l'éditeur Microsoft a diffusé, en 2006, 49 mises-à-jour critiques de sécurité pour son système d'exploitation Windows, le plus utilisé, et 43 mises-à-jour en 2007. Chacune corrige, un peu tard hélas, une erreur dans le logiciel qui permettait jusqu'alors de prendre le contrôle de l'ordinateur.

De telles prises de contrôles ne sont pas anecdotiques : 100 milliards de spams (courriels publicitaires indésirables) sont émis chaque jour par les milliers d'ordinateurs victimes de telles prises de contrôle, sans que leur propriétaires en soient conscients. Dans le cas de la riposte graduée, ces ordinateurs pourraient être utilisés pour télécharger illicitement les œuvres protégées, mettant leurs propriétaires sous le feu des sanctions tandis que les véritables commanditaires resteraient inconnus. Pire, les internautes seront souvent incapables de corriger la source du problème, tant que celui-ci n'aura pas été identifié par l'éditeur du logiciel attaqué. C'est donc, au final, l'internaute qui subira, une fois de plus, les erreurs de programmation des éditeurs informatiques.

La sécurité défaillante des réseaux sans fil

Une deuxième faille importante dans les installations informatiques domestiques est la présence d'un réseau sans fil (wifi), aujourd'hui disponible sur l'ensemble des boitiers ADSL fournis par les fournisseurs d'accès Internet. Les protocoles utilisés par ces matériels sont la cible d'attaques multiples et de plus en plus efficaces.

Ainsi, il suffit aujourd'hui d'une dizaine de minutes pour pénétrer le réseau sans fil de son voisin, si celui-ci utilise le protocole WEP, l'un des protocoles les plus anciens, encore très souvent utilisé. Le protocole WPA, plus récent, peut lui aussi être attaqué, si le mot de passe utilisé n'est pas suffisament complexe. Or, il est souvent commode d'utiliser un mot de passe simple quand on veut facilement configurer les nombreux équipements informatiques qui composent aujourd'hui un foyer (ordinateurs portables, téléphones intelligents, stations de jeux, boitiers télévision, etc.).

Le risque ici est qu'il sera facile pour un téléchargeur de se connecter sur le réseau de son voisin pour y télécharger illicitement sans risque. Une telle opération ne nécessitera souvent que de télécharger les outils adaptés sur Internet, sans grande expertise dans le domaine. Au contraire, détecter l'intrusion d'un voisin sur son propre réseau domestique requiert des outils beaucoup plus complexes, car les attaquants imitent souvent parfaitement le comportement de l'un des composants de l'installation.

Conclusion

La riposte graduée automatique repose sur une hypothèse, la capacité de l'utilisateur à sécuriser totalement son installation informatique domestique, dont l'impossibilité est bien connue de tous les experts.

3 Le contournement rapide et massif des mesures anti-piratages

La riposte graduée profite de l'ouverture des réseaux pair-à-pair, une situation qui ne devrait pas durer

Les réseaux pair-à-pair les plus utilisés aujourd'hui – eDonkey2000/Emule et Kazaa – ont été conçus en 2000-2001. Ces protocoles sont relativement simples, et donc facilement observables. Ils n'utilisent pas ou peu de chiffrement. Ils sont vulnérables au blocage (la reconnaissance du type de protocole, et son interdiction) et au filtrage (la reconnaissance du contenu

téléchargé, aboutissant à l'interdiction de la communication si son téléchargement est illicite).

Pour la riposte graduée, le repérage automatique s'effectue en proposant au téléchargement, sur un ordinateur témoin, une œuvre dont le téléchargement est illicite. Les internautes qui se connectent sur l'ordinateur et téléchargent l'œuvre sont identifiés par leur adresse Internet et enregistrés dans une base qui sera transmise à l'Hadopi.

Or, un nouveau concept de réseau a fait récemment son apparition sur Internet : il s'agit du réseau social, rendu populaire par Facebook, dans lequel chaque utilisateur ne peut communiquer qu'avec ses amis. Plusieurs projets de réseaux pair-à-pair limités au réseau social sont en train de se développer. Dans ces réseaux, un internaute n'autorise son ordinateur qu'à se connecter aux ordinateurs de ses amis, en utilisant des protocoles d'identification et de chiffrement évolués. Les documents sont alors téléchargés uniquement depuis les amis, ou les amis des amis, rendant le système de repérage actuel inopérant. Il est d'ailleurs bien connu qu'aucun système n'arrive, aujourd'hui, à observer les échanges sur le réseau de téléphonie Skype, qui utilise les mêmes principes.

Bien que ces réseaux soient encore peu utilisés, la riposte graduée devrait paradoxalement accroître leur popularité dans les prochaines années, à la faveur de l'augmentation de la taille des disques durs (un disque dur peut stocker aujourd'hui un millier de films) et de la généralisation des connexions par fibre optique. Ainsi, un ordinateur connecté à 200 amis (une moyenne observée sur le réseau social Orkut de Google) aurait accès directement à près de 200 000 copies de films ou de 50 millions de copies de chansons en MP3, sans qu'il soit possible d'observer ces téléchargements.

Conclusion

Les mesures techniques mises en place dans le cadre de la riposte graduée ne permettront pas de continuer le repérage des téléchargements illicites avec l'avénement probable des réseaux pair-à-pair basés sur le réseau social.