LES RISQUES PRÉSENTÉS PAR LES SITES ILLICITES

Mesures directes

30 juin 2017

Hadopi

www.hadopi.fr

TABLE DES MATIÈRES

Pratiques illicites, pratiques à risques ?

Plusieurs cas de cyberattaques, vols de données, fuites de fichiers confidentiels ont nourri l'actualité ces derniers mois, sensibilisant aux enjeux de la sécurité sur Internet tant les institutions et entreprises que les internautes. Dans le cadre de sa mission d'observation, l'Hadopi s'est intéressé aux risques potentiels pour la sécurité informatique et les données personnelles des internautes utilisant des sites proposant des biens culturels dématérialisés manifestement contrefaisants.

Selon les résultats d'une étude INCOPRO[1] pour la MPA (Motion Picture Association), environ un tiers des publicités sur des sites considérés comme contrefaisants seraient trompeuses ou inciteraient à installer des logiciels potentiellement indésirables (LPI, ou PUP « *potentially unwanted program* », par exemple des publiciels / *adware*).

L'Hadopi a souhaité identifier les risques potentiels pour les utilisateurs de ces sites et a mis en place un protocole complet de recherche en deux phases :

- une phase d'observation directe de 62 sites manifestement illicites menée par l'Hadopi entre décembre 2016 et janvier 2017 dont les résultats sont présentés dans le présent rapport ;
- une étude quantitative sur les perceptions des internautes, réalisée en ligne du 25 janvier au 2 février 2017 par l'institut Ifop[2], permettant de recueillir les expériences des internautes ayant fréquentés ces types de sites.

Cette phase d'observation fait suite à une première expérimentation réalisée fin 2015, dont les principaux résultats ont été publiés dans le rapport annuel 2014-2015 de l'Hadopi. Ceux-ci sont repris ici à titre de comparaison.

[1] Étude « The revenue sources for websites making available copyright content without consent in the EU », INCOPRO, Mars 2015. En avril 2014 INCOPRO avait publié les résultats d'une étude sur ce sujet : http://www.incopro.co.uk/case_studies/advertising-and-malware/

[2] Étude sur les risques encours sur les sites illicites, Ifop pour Hadopi, 2017.

1 | Protocole d'observation

1 - 1 | Principes généraux

La première étape a consisté à déterminer un échantillon de sites proposant des biens culturels dématérialisés. Ces sites ont été sélectionnés selon les critères suivants :

- ils n'hébergent pas eux-mêmes des fichiers (annuaires de liens) ;
- ils référencent des liens vers des contenus vraisemblablement mis à disposition sans l'autorisation des ayants droit et qui ne remplissent pas les conditions pour être considérés comme légaux[3] ;
- ils bénéficient d'une certaine notoriété en tant que sites manifestement contrefaisants (évaluée par rapport à leur fréquentation, leur évocation spontanée par les internautes dans les travaux d'études et de recherche menés par l'Hadopi ou d'autres organismes ou leur positionnement dans des moteurs de recherche).

L'échantillon devait également couvrir plusieurs secteurs culturels, et différents modes de consommation : annuaires de liens pour téléchargement direct, pair-à-pair (P2P) et *streaming*.

Les sites ont ensuite analysés manuellement un à un sur un poste de travail dédié et configuré comme suit :

- système d'exploitation : Windows 10 ;
- navigateur : Firefox (dernière version à jour) ;
- plugins : Acrobat Reader, Flash, Java, Silverlight (dernières versions à jour) ;
- antivirus installé et activé.

1 - 2 | Grille d'observation

Les informations recherchées étaient les suivantes :

- informations générales sur le site, par exemple types de contenus proposés, modes d'accès aux contenus etc. ;
- informations sur les modes d'ouverture ou d'affichage des publicités, par exemple présence de liens trompeurs[4], ouverture intempestive de fenêtres etc. ;
- informations sur la présence de contenus ou publicités potentiellement problématiques pour les mineurs ;
- informations sur la présence de messages ou annonces trompeurs, par exemple du type « iPhone à 1€ », « gagner de l'argent rapidement », fausses alertes de sécurité etc. ;

[3] Cf. la liste de sites référencés sur www.offrelegale.fr
[4] Liens qui pointent vers un contenu autre que celui annoncé.

- informations sur la présence de logiciels ou extensions de navigateur potentiellement indésirables (LPI), ou de logiciels malveillants, que ce soit directement sur le site ou par l'intermédiaire des liens et publicités qu'on peut y trouver ;
- informations sur la demande de données personnelles et/ou bancaires ;

S'agissant des logiciels potentiellement indésirables (LPI) ou malveillants, leur détection a été faite sur la base de l'analyse de fichiers exécutables proposés à l'installation à l'aide de l'outil Virustotal[5], en vérifiant la présence d'alertes pour un site dans Virustotal ou Google Safe Browsing[6]. Ces deux outils ont également été utilisés pour vérifier d'éventuelles alertes au sujet des « *drive-by download* »[7].

Et par ailleurs la limitation à l'analyse de fichiers exécutables excluait les fichiers contenant *a priori* des œuvres culturelles (ex. extensions .avi, .mp3 etc., mais aussi des .zip). Les risques potentiels liés au téléchargement de tels fichiers, pouvant contenir du code exploitant des vulnérabilités informatiques, n'étaient pas pris en compte.

Les alertes de Google Safe Browsing peuvent être de plusieurs niveaux, un site pouvant être « *dangereux* », « *en partie dangereux* », « *non dangereux* » etc. Ces indications étaient complétées par des informations de sécurité relatives à un site, tel que « *les internautes sont redirigés via certaines pages de ce site vers des sites Web dangereux.* », « *Sur ce site, des pirates informatiques pourraient vous inciter à télécharger un logiciel ou vous dérober des informations (telles que vos mots de passe, vos messages ou les informations de votre carte de crédit).* », etc.

La classification comme LPI de certaines extensions de navigateur proposées à l'installation a été faite après recherche dans Google ou sur la base de raisons de plausibilité (ex. une mise à jour Java ne nécessitant pas l'installation d'une extension de navigateur, on peut légitimement conclure que l'extension présente un risque potentiel).

Une utilisation moyenne simulée

L'observation conduite avait pour vocation à simuler le comportement naturel d'un internaute ayant recours à ces sites, avec une recherche circonscrite à deux ou trois biens culturels dématérialisés par site. Le parcours de l'utilisateur commençait sur la page d'accueil du site et se terminait soit au lancement d'un stream, soit à l'initiation d'un téléchargement d'un fichier contenant l'œuvre ou d'un fichier .torrent, soit à un écran demandant l'ouverture d'un compte sur un cyberlocker pour accéder au contenu recherché. Dans le cas où plusieurs sources de streaming ou téléchargement étaient proposées, deux à trois de ces sources ont été vérifiées.

[5] https://www.virustotal.com/ ; VirusTotal est un service gratuit qui analyse les fichiers et URL suspects, et facilite la détection rapide des virus, vers, chevaux de Troie et tous types de malwares.

[6] https://www.google.com/transparencyreport/safebrowsing/?hl=fr : Google Safe Browsing est un service offert par Google qui fournit des listes noires de ressources web (sites) qui potentiellement présentent des dangers pour la sécurité informatique ou les données des utilisateurs.

[7] Un *drive-by download* désigne l'installation d'un logiciel malveillant ou potentiellement indésirable sans qu'une action de l'internaute ne soit nécessaire. Une telle attaque exploite des failles de sécurité d'un navigateur ou d'un système et peut avoir lieu simplement en naviguant sur le web.

1 - 3 | Critères retenus pour classifier les sites

Les observations ont permis de classifier les sites selon plusieurs critères[8] :

- sites présentant un **risque potentiel pour la sécurité informatique des utilisateurs** : sites exposants les utilisateurs à au moins un logiciel malveillant ou potentiellement indésirable (au moins un résultat dans Virustotal pour les fichiers exécutables analysés ou au moins une incitation à l'installation d'une extension de navigateur potentiellement indésirable, ou pour lesquels une recherche sur l'URL dans Virustotal ou Google Safe Browsing fait remonter une alerte[9] en cours au moment de l'observation) ;
- sites présentant des **messages ou annonces trompeurs** : sites sur lesquels un visiteur est exposé à au moins un message ou une annonce de ce type ;
- sites **recueillant ou demandant des informations personnelles ou bancaires** : sites proposant aux utilisateurs de s'inscrire, de laisser des commentaires, etc. ;
- sites contenant des **publicités pouvant être considérées comme intrusives** : sites contenant des liens (ou boutons) trompeurs, des fenêtres de publicité qui s'ouvrent intempestivement, des *pop-under,* etc. ;
- Sites proposant du **contenu inapproprié pour des mineurs** (contenus pornographiques) : sites proposant des liens vers des contenus pour adultes sans vérification d'âge préalable, ou qui affichent des publicités pornographiques ou qui incluent des liens publicitaires vers des sites avec du contenu pornographique.

1 - 4 | Composition de l'échantillon et durée d'observation

L'échantillon retenu se composait de 62 sites (contre 72 lors de la première observation en 2015) :

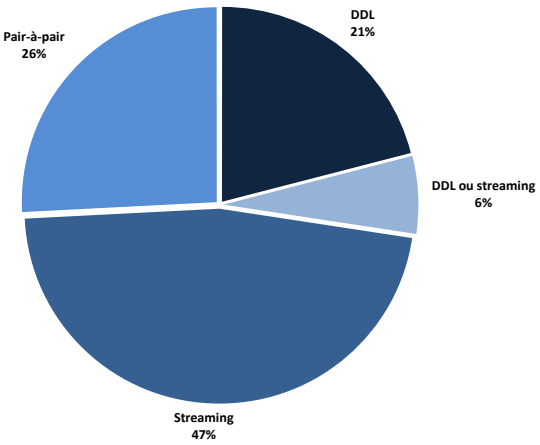
- 13 sites proposant des liens pour téléchargement direct (DDL) ;
- 4 sites proposant des liens pour téléchargement direct (DDL) ou streaming ;
- 29 sites proposant du streaming ;
- 16 sites proposant des liens pour téléchargement en pair à pair (P2P).

Ces 62 sites couvraient des œuvres et contenus de quatre secteurs culturels : audiovisuel, musique, jeu vidéo et livres.

L'observation a eu lieu entre décembre 2016 et janvier 2017.

[8] La classification ayant évoluée par rapport à celle de l'observation 2015, il n'est pas possible de comparer l'ensemble des résultats.
[9] Statut autre que « non dangereux » et aucune information de sécurité particulière relative au site dans le cas de Google Safe Browsing.

Figure 1 : Répartition des sites observés selon les types de service proposés

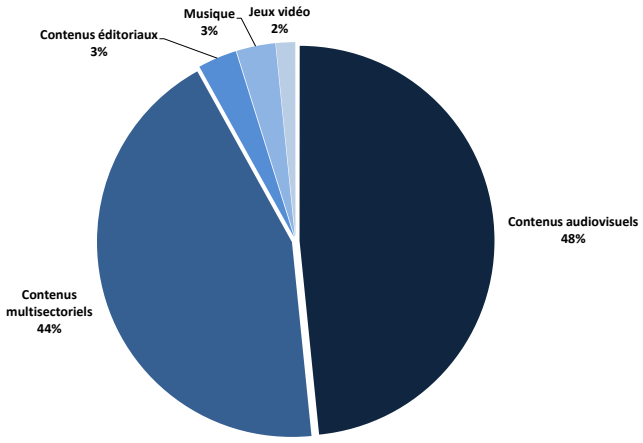


Source : Hadopi

L'échantillon se répartissait comme suit en termes de contenus proposés :

- 30 sites proposant uniquement des contenus audiovisuels ;
- 2 sites proposant uniquement du contenu éditorial (livres, magazines etc.) ;
- 2 sites proposant uniquement de la musique ;
- 1 site proposant uniquement des jeux vidéo ;
- 27 sites proposant des contenus de plusieurs secteurs culturels.

Figure 2 : Répartition des sites observés selon les contenus proposés



Source : Hadopi

2 | Résultats

L'observation a eu lieu entre décembre 2016 et janvier 2017.

2 - 1 | Sites présentant un risque potentiel pour la sécurité informatique des utilisateurs

Au cours de l'observation, Google Safe Browsing a signalé 9 sites, soit environ 15 %, avec des alertes relatives au danger de *drive-by download*.[10]

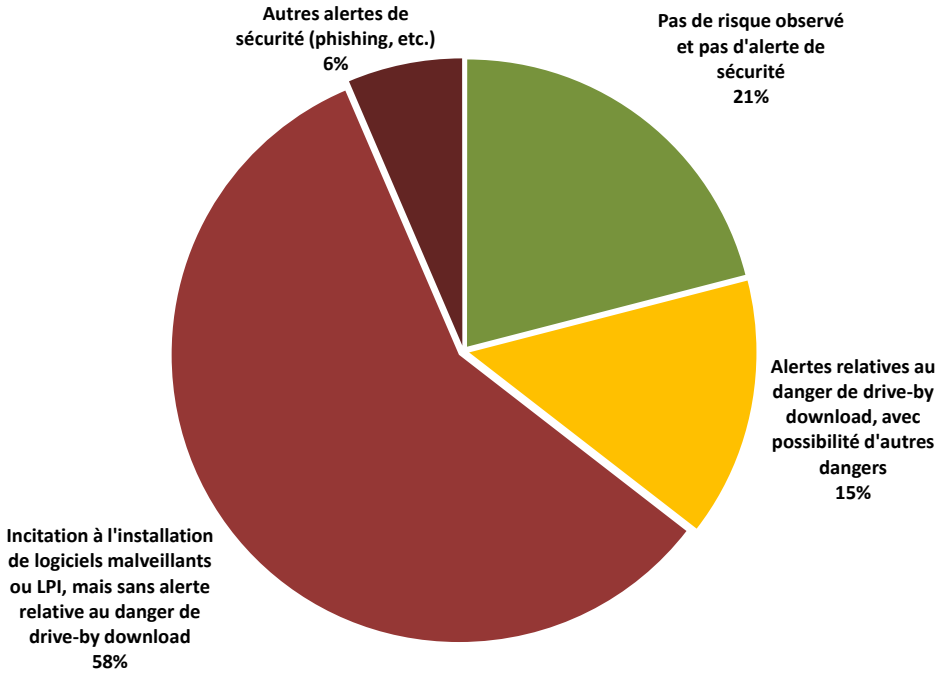
L'incitation à l'installation de logiciels ou extensions potentiellement indésirables semble plus répandue que le *drive-by download*. Cette incitation se présente principalement sous forme de demande d'installation de mise à jour de lecteurs média (Flash, Media Player) ou de mise à jour de Java, donc sur la base d'une information trompeuse. Ce type d'incitation était présent sur 41 sites, soit 66 % des sites observés (2015 : 51 %). 19 de ces sites faisaient également l'objet d'une alerte sur Google Safe Browsing et/ou sur Virustotal, dont 5 avec des alertes relatives au danger de *drive-by download*.

Sur les 21 sites restants, 8 faisaient l'objet d'une alerte sur Google Safe Browsing et/ou Virustotal, dont 4 avec des alertes relatives au danger de *drive-by download*.

En tout, environ 79 % des sites observés (49 sur 62) présentaient, au moment de l'observation, un risque potentiel pour la sécurité informatique des utilisateurs, ou faisaient l'objet d'une alerte de sécurité (2015 : 60 %).

[10] Les résultats de l'étude « Digital Bait », publiée en décembre 2015 par Digital Citizens Alliance et RiskIQ, montrent que sur un échantillon de 800 sites considérés comme contrefaisants 33% pouvaient exposer les utilisateurs à l'installation de logiciels malveillants, dont 55% avec et 45% sans nécessité d'action de l'utilisateur. (http://www.digitalcitizensalliance.org/cac/alliance/content.aspx?page=digitalbait)

Figure 3 : Risques potentiels pour la sécurité informatique des utilisateurs sur les sites observés



Source : Hadopi

2 - 2 | Sites présentant des messages ou annonces trompeurs

Environ 71 % des sites observés (44 sur 62) exposaient les utilisateurs à un ou plusieurs contenus pouvant être considérés comme des messages ou annonces trompeurs. Les types de messages ou annonces les plus fréquemment observés étaient des propositions pour gagner rapidement de l'argent en spéculant, la souscription « cachée » d'un abonnement (cas par exemple dans le « gain » d'un iPhone à 1€ ou lors de la vérification d'âge à l'aide d'une carte bancaire), ou des fausses alertes de sécurité. L'annonce trompeuse du type « souscription cachée d'abonnement » a été observée sur 20 sites (32 % ; 2015 : 31 %).

2 - 3 | Sites recueillant ou demandant des données personnelles ou des informations bancaires

Sur les 26 sites recueillant des données personnelles, soit 42 % de l'échantillon (contre 63 % en 2015), seulement un seul site exigeait de fournir des données personnelles pour pouvoir accéder à son contenu. Les 25 autres sites demandaient des données personnelles, notamment une adresse mail, pour s'inscrire, laisser des commentaires ou s'abonner à une liste de diffusion, mais la fourniture de ces données par l'internaute n'était pas obligatoire pour accéder au site.

Dans trois cas des sites demandaient des données de carte bancaire, sous prétexte de vérification d'âge, pour accéder à du contenu pour adultes. Dans les trois cas, soit 5 % des sites observés (contre 13 % en 2015), il s'agissait d'une arnaque visant en réalité à faire souscrire un abonnement à l'utilisateur.

Des données de carte bancaire peuvent également être demandées pour payer certains services d'hébergement (*cyberlockers*, Usenet) qui proposent des abonnements. Il ne s'agit toutefois pas de souscriptions cachées.

2 - 4 | Sites contenant des publicités pouvant être considérées comme intrusives ou trompeuses

L'observation a permis de constater que la quasi-totalité des sites (59 sur 62, soit 95 %) contenait, à des degrés plus ou moins élevés, des publicités pouvant être considérées comme intrusives, et ce dans des proportions comparables à l'étude 2015 (93 %). Il s'agit plus particulièrement d'ouvertures intempestives d'une ou plusieurs fenêtres de publicité, parfois avec son, et de liens trompeurs (le lien ou l'image cliquables suggèrent une action alors que ce sont des liens publicitaires).

2 - 5 | Sites proposant du contenu inapproprié (pornographique) pour des mineurs

Environ 31 % des sites observés, soit 19 sur 62, affichaient ou proposaient des liens vers du contenu ou des publicités pour adultes sans vérification préalable de l'âge des utilisateurs (2015 : 40 %).

3 | Conclusions

79 % des sites observés présentaient un risque potentiel pour la sécurité informatique des utilisateurs ou faisaient l'objet d'une alerte de sécurité. En comparaison avec l'étude de 2015, ce chiffre est en augmentation de 19 points (2015 : 60 %).

Tableau 1 : Récapitulatif des dangers et risques potentiels rencontrés sur les sites observés

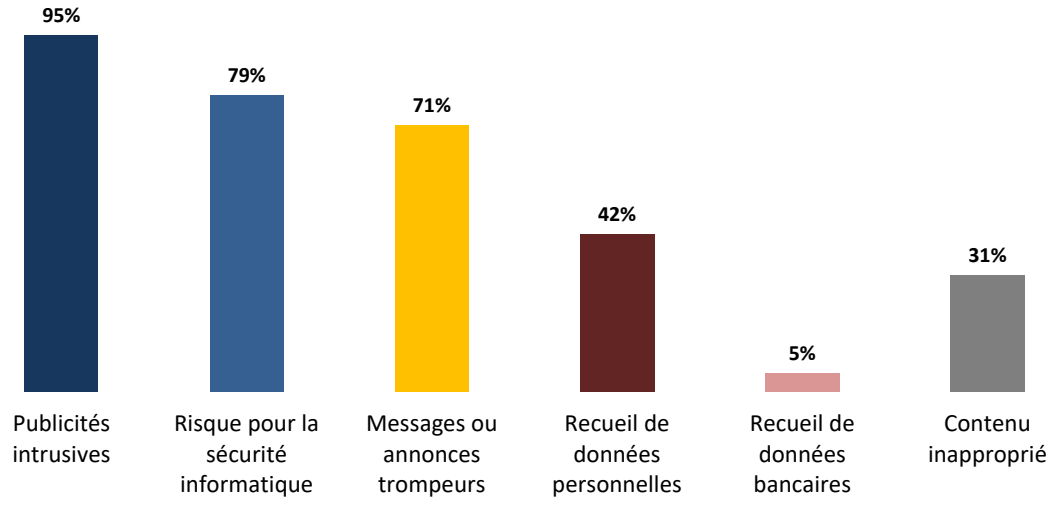
Catégories de sites	Répartition au sein de l'échantillon
Sites contenant des **publicités pouvant être considérées comme intrusives**	95 %
Sites présentant un **risque potentiel pour la sécurité informatique des utilisateurs**	79 %
Sites présentant des **messages ou annonces trompeurs**	71 %
Sites **recueillant ou demandant des données personnelles ou des informations carte bancaire**	Données personnelles : 42 % Informations carte bancaire : 5 %
Sites proposant du **contenu inapproprié pour des mineurs** (contenus pornographiques)	31 %

Source : Hadopi

Les risques observés se manifestaient principalement sous forme d'incitation, à partir de liens trompeurs, à installer des logiciels potentiellement indésirables. Un pourcentage non négligeable de sites faisait l'objet d'alertes sur des techniques de *drive-by download*, particulièrement dangereux pour les utilisateurs car l'installation de logiciels malveillants peut se faire sans action de leur part, notamment sur des systèmes informatiques qui ne sont pas à jour.

Les résultats montrent que la navigation sur des sites manifestement contrefaisants expose considérablement les utilisateurs, en particulier ceux qui sont les moins familiers des nouvelles technologies, à des risques.

Figure 4 : Récapitulatif des dangers et risques potentiels rencontrés sur les sites observés



Source : Hadopi

Hadopi

Haute Autorité pour la diffusion des œuvres
et la protection des droits sur internet

4, rue du Texel – 75014 Paris

presse@hadopi.fr

www.hadopi.fr

www.twitter.com/InsidOpi