EUIPO
OFFICE DE L'UNION EUROPÉENNE POUR LA PROPRIÉTÉ INTELLECTUELLE

IDENTIFICATION ET ANALYSE DES LOGICIELS MALVEILLANTS SUR UNE SELECTION DE SITES WEB SOUPÇONNES DE PORTER ATTEINTE AU DROIT D'AUTEUR

SYNTHÈSE

Septembre 2018

Résumé

Les contenus soupçonnés de porter atteinte au droit d'auteur constituent une sérieuse atteinte aux droits de propriété intellectuelle. Il existe certains sites web qui partagent ces contenus publiquement, parfois même gratuitement, sans aucune inscription. Outre ces contenus, il est fréquent que les sites web diffusent divers types de logiciels malveillants, ainsi que des programmes potentiellement indésirables (PPI), incitant les utilisateurs par la ruse à télécharger et à lancer ces fichiers. L'étude fournit une vue d'ensemble des exemples les plus récents de logiciels malveillants et PPI présents sur des sites web soupçonnés de porter atteinte au droit d'auteur. Ces programmes ont recours à des techniques trompeuses et l'ingénierie sociale — installations de jeux vides et logiciels apparemment «utiles», par exemple — pour inciter les utilisateurs finaux à divulguer des informations sensibles. Au cours de l'étude, divers PPI ont été découverts, tels que des logiciels «utiles», de faux installateurs de jeux et des clients pour plateformes de lecture vidéo en transit. Ces logiciels ne présentent pas nécessairement de danger direct pour le logiciel ou le matériel de l'utilisateur. Toutefois, par des ruses d'ingénierie sociale, ils permettraient de convaincre un utilisateur de divulguer des informations à caractère personnel jugées sensibles ou des renseignements sur ses cartes de paiement. En outre, des informations concernant l'ordinateur lui-même pourraient être divulguées à d'autres parties sans le consentement explicite de l'utilisateur.

Équipe de recherche

L'équipe de recherche était composée de Francesca Bosco, chargée de programme à l'UNICRI, et d'Andrii Shalaginov, chargé de recherche en sécurité de l'information au département des technologies de sécurité de l'information et de la communication (groupe de criminalistique informatique), à la faculté de technologie de l'information et d'ingénierie électrique de l'Université norvégienne des sciences et de technologie.

Clause de non-responsabilité

Dans ce contexte, il convient de souligner que la recherche avait pour seul objectif de déterminer les caractéristiques techniques des logiciels malveillants et des PPI rencontrés au cours de l'étude et susceptibles d'être rencontrés par des internautes à la recherche de contenus soupçonnés de porter atteinte au droit d'auteur. Les échantillons de logiciels malveillants et de PPI documentés ne sauraient être considérés comme exhaustifs et l'étude (ou ses résultats) n'avait pas pour objectif de fournir une évaluation générale de la probabilité ou du risque de contamination par des logiciels malveillants et des PPI que rencontrerait un internaute à la recherche de matériel soupçonné de porter atteinte au droit d'auteur.

Avant-propos

Les activités en ligne soupçonnées de porter atteinte au droit d'auteur peuvent être financées de diverses façons: frais d'abonnement, donations, paiement de services auxiliaires et revenus provenant de l'affichage publicitaire en ligne, etc.

Toutefois, tous les moyens de financement ne sont pas aussi bénins que les exemples cités. Depuis des années, la dissémination d'infections par le biais de logiciels malveillants et d'autres programmes potentiellement indésirables (PPI) revêt une importance capitale pour le financement des activités réalisées sur l'internet et soupçonnées de porter atteinte au droit d'auteur.

Les internautes ordinaires commencent à prendre conscience des risques d'infection lorsqu'ils accèdent à des sites web ou à des applications mobiles soupçonnés de porter atteinte au droit d'auteur.

Le tableau de bord 2015 de l'EUIPO sur les jeunes et la PI a montré que 52 % des jeunes considèrent que la sécurité sur un site web est importante lorsqu'ils accèdent à des contenus en ligne. Au total, 78 % des jeunes ont déclaré qu'ils y réfléchiraient à deux fois s'ils savaient que l'ordinateur ou l'appareil risquait d'être infectés par des virus ou des logiciels malveillants. Au total, 84 % ont déclaré qu'ils y réfléchiraient à deux fois s'ils savaient que les données d'une carte de crédit risquaient d'être volées.

Dans les recherches menées dans le cadre de cette étude, l'Office s'est lancé dans une tâche très complexe sur le plan technique, à savoir détecter et documenter des exemples de logiciels malveillants et de PPI qu'un internaute pourrait rencontrer en tentant d'accéder à des films, de la musique, des jeux vidéo et des émissions télévisées, tous ces produits à succès ayant été piratés.

Dans ce contexte, il convient de souligner que la recherche avait pour seul objectif de déterminer les caractéristiques techniques des logiciels malveillants et des PPI rencontrés au cours de l'étude et susceptibles d'être rencontrés par des internautes à la recherche de contenus soupçonnés de porter atteinte au droit d'auteur. Les échantillons de logiciels malveillants et de PPI documentés ne sauraient être considérés comme exhaustifs et l'étude (ou ses résultats) n'avait pas pour objectif de fournir une évaluation générale de la probabilité ou du risque de contamination par des logiciels malveillants et des PPI que rencontrerait un internaute à la recherche de matériel soupçonné de porter atteinte au droit d'auteur.

La recherche a été menée en plusieurs étapes, en étroite coopération avec le Centre européen de lutte contre la cybercriminalité (EC3) au sein d'Europol.

Les résultats font ressortir divers risques posés par des logiciels malveillants et des PPI qu'un internaute peut rencontrer en recherchant des contenus soupçonnés de porter atteinte au droit d'auteur. La plupart des logiciels malveillants et des PPI documentés peuvent être décrits comme des chevaux de Troie ou comme tout logiciel indésirable capable d'accéder indûment aux données à caractère personnel des internautes. Ces exemples seront utiles et intéressants non seulement pour l'ensemble des titulaires de droits de propriété intellectuelle, mais aussi pour les autorités répressives et, enfin, et surtout, pour les consommateurs préoccupés par le fait qu'il soit possible d'accéder à leurs données à caractère personnel sans leur autorisation.

Synthèse

L'étude fournit une vue d'ensemble des exemples les plus récents de logiciels malveillants et programmes potentiellement indésirables (PPI) trouvés sur des sites web soupçonnés de porter atteinte au droit d'auteur. Ces programmes ont recours à des techniques trompeuses et à l'ingénierie sociale — installations de jeux vides et logiciels apparemment «utiles», par exemple — pour inciter les utilisateurs finaux à divulguer des informations sensibles.

L'objectif de cette étude est de découvrir et de documenter des logiciels malveillants ou indésirables diffusés sur certains sites web soupçonnés de porter atteinte au droit d'auteur et de classer par catégories les échantillons trouvés en ayant recours à diverses taxonomies de logiciels malveillants. Dans ce contexte, il convient de souligner que l'étude avait pour seul objectif de déterminer les caractéristiques techniques des logiciels malveillants et des PPI rencontrés au cours de la recherche et susceptibles d'être rencontrés par des internautes en quête de contenus soupçonnés de porter atteinte au droit d'auteur. Les échantillons de logiciels malveillants et de PPI documentés ne sauraient être considérés comme exhaustifs et la recherche (ou ses résultats) n'a pas pour objectif de fournir une évaluation générale de la probabilité ou du risque de contamination par des logiciels malveillants et des PPI que rencontrerait un internaute à la recherche de matériel soupçonné de porter atteinte au droit d'auteur. Aux fins de cette étude, les émissions télévisées, les films, la musique et les jeux vidéo sont considérés comme des contenus protégés par le droit d'auteur.

Conclusions de l'étude

Les contenus soupçonnés de porter atteinte au droit d'auteur représentent une violation importante des droits de propriété intellectuelle. Il existe certains sites web qui partagent ces contenus publiquement, parfois même gratuitement, sans aucune inscription. Outre ces contenus, les sites web diffusent généralement divers types de logiciels malveillants et PPI, incitant les utilisateurs par la ruse à télécharger et à lancer ces fichiers. Au cours de l'identification des sites web sur la base du classement des 500 principaux sites réalisé par Alexa, outre une simulation des recherches effectuées par l'utilisateur moyen à l'aide de moteurs de recherche bien connus, tels que Google, Yahoo et Bing, il a été constaté que l'ensemble de sites web avait changé entre les deux cycles d'étude. Ce changement est probablement le résultat des efforts déployés par les moteurs de recherche pour supprimer les liens vers des sites web soupçonnés de porter atteinte au droit d'auteur, tandis que de nouveaux sites web suspects continuent d'apparaître. En ce qui concerne l'identification des sites web, une constatation intéressante a trait au fait que l'écrasante majorité des sites web sont hébergés aux États-Unis ou ont des noms de domaine liés à l'hébergement dans ce pays. En revanche, seuls quelques sites sont situés sur des serveurs localisés dans l'UE. Par ailleurs, .com et .net sont les noms de domaines de premier niveau les plus fréquemment utilisés sur les sites web soupçonnés de porter atteinte au droit d'auteur. Cela peut s'expliquer par le fait que, contrairement à des domaines spécifiques à un pays, ceux-ci ne peuvent exiger une identification de l'utilisateur au moyen d'un passeport ou d'autres documents d'identification. En moyenne, 20 % de nouveaux sites web ont été ajoutés et 20 % d'anciens sites web ont été supprimés entre les deux cycles d'identification. En outre, près de 8 % des sites identifiés au cours des deux cycles ont été qualifiés de malveillants par la plateforme VirusTotal. À l'aide de divers systèmes de gestion de contenus, il est maintenant devenu possible de créer presque sans effort un site web et de fournir du contenu aux utilisateurs, voire des applications malveillantes.

Avant la collecte de logiciels malveillants, cette étude s'est livrée à un examen documentaire des menaces posées par les logiciels malveillants en 2017 et à un classement par catégories en rapport avec l'état de la technique. Cet ensemble de connaissances a ensuite été utilisé lors de l'analyse des logiciels malveillants afin de suivre les principes reconnus par l'ensemble des acteurs dans ce domaine en ce qui concerne les types et l'identification des familles de logiciels malveillants. Au total, 106 dossiers ont été rassemblés au cours des deux cycles de collecte de données. Il s'agit notamment des fichiers téléchargés directement sur des sites web soupçonnés de porter atteinte au droit d'auteur, ainsi que des fichiers créés au cours de l'exécution des fichiers téléchargés. Lors de l'étude, divers PPI ont été découverts, tels que des logiciels «utiles», de faux installateurs de jeux et des clients pour plateformes de lecture vidéo en continu. Ces logiciels ne constituent pas nécessairement un danger immédiat pour le logiciel ou le matériel de l'utilisateur. Toutefois, par des ruses d'ingénierie sociale, ils permettraient de convaincre un utilisateur de divulguer des informations à caractère personnel jugées sensibles ou des renseignements sur ses cartes de paiement. En outre, des informations concernant l'ordinateur lui-même pourraient être divulguées à d'autres parties sans le consentement explicite de l'utilisateur.

Les logiciels malveillants collectés ont été analysés dans un premier temps à l'aide d'outils de source ouverte pour comprendre la logique interne, détecter d'éventuelles activités malveillantes et évaluer leur pertinence pour la présente étude sur les logiciels malveillants. Outre l'analyse préliminaire réalisée au moyen d'outils de source ouverte, les échantillons de logiciels malveillants collectés ont été analysés par la plateforme des solutions d'analyse des logiciels malveillants d'Europol (EMAS), ce qui a permis la détection d'un grand nombre de différents artefacts et d'activités malveillantes. Les rapports EMAS comprennent une analyse complète des fichiers à l'aide de quatre versions de MS Windows, dans lesquelles le trafic réseau, les appels de fonctions et les activités des disques sont consignés de façon détaillée dans un journal en vue d'une analyse plus approfondie. En outre, la plateforme met en évidence toutes les activités suspectes détectées lors des routines d'exécution des fichiers. Après analyse de tous les rapports, 35 types d'activités malveillantes ont été relevées par EMAS, lesquelles sont regroupées dans 17 classes d'événements malveillants. Ces cas vont d'anomalies générales (comme le lancement de processus système ou la recherche de processus dans les mémoires) à des actions indéniablement malveillantes telles que l'utilisation d'un enregistreur de frappe, d'un outil de dissimulation d'activité («rootkit») et d'un outil d'altération du trafic réseau.

Les échantillons binaires de logiciels malveillants et PPI collectés ont généralement révélé plusieurs modèles d'activité généraux: des programmes «utiles» prétendant nettoyer les anciens fichiers sur l'ordinateur de l'utilisateur contre un abonnement payant, des simulateurs d'installation de jeux nécessitant des données à caractère personnel de l'utilisateur, et des programmes gratuits offrant un accès à des plateformes qui distribuent du contenu piraté, par exemple au moyen du traqueur BitTorrent. Les deux cycles d'identification de sites web et de collecte de logiciels malveillants ont donné des résultats prometteurs en termes de compréhension des méthodes de dissémination des logiciels malveillants et d'ingénierie sociale visant à obtenir par la ruse des informations personnelles et d'identification jugées sensibles. Par ailleurs, la détection d'un grand nombre de PPI pour le système d'exploitation Android OS, disponibles sur des plateformes de distribution de contenus soupçonnés de porter atteinte au droit d'auteur, atteste de façon évidente de l'augmentation de la popularité des appareils mobiles ces dernières années. Une corrélation des analyses a permis de conclure que l'éventail des menaces liées aux logiciels malveillants diffusés par des sites web portant atteinte au droit d'auteur est plus sophistiqué qu'il n'y paraît à première vue. Parmi les logiciels découverts, certains peuvent également être classés dans les catégories des chevaux de Troie, des logiciels de publicité, des portes dérobées et des agents. À cela s'ajoute le fait que de nombreuses familles de logiciels malveillants spécifiques, telles que WisdomEyes, DealPly et FileRepMalware, ont également été découvertes. En outre, un tel classement complet par catégories est également valable pour la plateforme Android, et pas uniquement pour Microsoft Windows. Il existe un large éventail de menaces pour les biens des utilisateurs, dont, entre autres, le vol d'informations d'identification jugées sensibles, de données à caractère personnel et d'informations de configuration matérielle, ainsi que la modification du trafic réseau. Par conséquent, même s'il est possible que les logiciels identifiés soient des PPI, ceux-ci peuvent néanmoins avoir une incidence sur les utilisateurs, en particulier dans les cas

impliquant un utilisateur moyen susceptible de ne pas avoir pleinement connaissance des pratiques et des mesures élémentaires en matière de sécurité en ligne.

Un exemple des conclusions de l'étude est présenté ci-dessous.

Site web 03	Site web 09	Site web 08
Le site web incite par la ruse les utilisateurs à utiliser une fausse installation de jeu; l'ensemble du processus d'obtention des informations sensibles d'un utilisateur a changé entre le premier et le second cycle de collecte de logiciels malveillants. L'utilisateur de ce service télécharge une archive contenant du contenu masqué sous forme de fichiers liés au jeu et non un fichier exécutable explicitement binaire que tous les antivirus peuvent identifier comme malveillant. Les archives	Le site web donne accès à tout type de contenu vidéo disponible au moyen de traqueurs «torrent» à l'aide d'un outil logiciel. Cet outil nécessite moins d'interactions avec l'utilisateur par rapport à d'autres traqueurs BitTorrent. Seuls quelques clics sont nécessaires pour télécharger des contenus à partir de sources inconnues, et l'utilisateur n'est pas protégé pas plus qu'il n'a de contrôle sur ce qui est téléchargé.	(Android) Le site web donne accès, sans qu'il soit nécessaire de s'inscrire, à une série d'applications mobiles gratuites. Une application donne un accès illimité à la diffusion en continu de programmes télévisés et de films. Il n'est pas explicitement demandé à l'utilisateur de fournir des informations sensibles ou des données de paiement pour acheter un accès à des vidéos protégées par le droit d'auteur. Toutefois, cet utilisateur doit désactiver les paramètres de sécurité qui permettront d'installer d'autres applications que celles

Méthodologie

Pour mener à bien la recherche, une méthodologie rigoureuse a dû être adoptée pour la sélection des titres et des sites web, ainsi que pour la tâche difficile du point de vue technique consistant à détecter et à documenter les exemples de logiciels malveillants et de PPI découverts. Un bref aperçu de la méthodologie est présenté ci-dessous:

1. Au cours de la phase I de la recherche de l'UNICRI, en collaboration avec l'Observatoire européen des atteintes aux droits de propriété intellectuelle (l'«Observatoire»), un groupe d'appui composé d'experts a été mis en place afin de fournir des conseils concernant la méthodologie de recherche et la sélection des sites web utilisés à des fins d'analyse, et d'évaluer les recherches entreprises lors de chaque phase de mise en œuvre du projet. Ce groupe comprenait des représentants des parties prenantes de l'Observatoire, d'organisations de titulaires de droits, d'universités, de services répressifs et d'agences de l'UE.

2. En parallèle, l'équipe de recherche a été sélectionnée. Dans le cadre de ce rapport, il n'était pas techniquement possible[1] de procéder à une recherche dans tous les États membres de l'UE; par conséquent, un échantillon de 10 pays a été sélectionné de manière aléatoire parmi les 28 États membres de l'UE au cours de la phase II.

3. Au cours de la phase III, des films, des programmes télévisés, des chansons et des jeux vidéo populaires ont été identifiés. Cette popularité comprenait aussi bien la popularité au niveau mondial que la popularité dans un seul ou plusieurs des 10 pays retenus dans l'échantillon au 23 juin 2017, date du début de la période de collecte des données. Dans les phases suivantes de l'étude, cet échantillon de titres a été systématiquement utilisé dans les recherches en ligne afin

[1] Le nombre de pays sélectionnés aura une incidence directe (augmentation) sur le nombre de sites web soupçonnés de porter atteinte au droit d'auteur sélectionnés et sur les fichiers binaires à analyser. Par conséquent, il a été décidé de ne se concentrer que sur un échantillon de pays afin de pouvoir mener à bien la partie pratique de l'étude dans un délai donné.

de trouver des sites web et des applications mobiles portant atteinte au droit d'auteur. Chaque titre répondait à deux ou plusieurs des critères suivants:

- populaire au moment de la collecte des données au sein des États membres de l'UE;
- populaire au moment de la collecte des données à l'échelle mondiale;
- historiquement populaire à l'échelle mondiale; et
- classé dans la catégorie des films, programmes télévisés, chansons ou jeux vidéo.

Cinq titres de films, cinq titres de programmes télévisés, cinq titres musicaux et cinq titres de jeux vidéo ont été sélectionnés, soit un échantillon total de 20 titres. Une attention particulière a été accordée aux sources utilisées pour déterminer la popularité d'un titre en particulier, ce qui a impliqué la mise en œuvre d'une procédure de sélection systématique afin de garantir que des données sources soient disponibles pour l'ensemble ou la plupart des États membres.

4. La phase IV a mis en évidence des sites web soupçonnés de fournir un accès illégal à des contenus protégés par le droit d'auteur, qui étaient populaires au niveau mondial ou parmi les 10 pays de l'échantillon au 26 juin 2017 (premier cycle de collecte de logiciels malveillants). Dans une phase ultérieure de l'étude, ces sites web ont été analysés aux fins de détecter la présence de logiciels malveillants et de programmes potentiellement indésirables.

   La méthode utilisée pour identifier les sites web soupçonnés de porter atteinte au droit d'auteur a été mise au point avec la contribution du groupe d'appui identifié au cours de la phase I, et à la suite d'un examen de la littérature existante par l'UNICRI. Elle a été spécialement conçue pour générer un échantillon de sites web qui:

   - étaient populaires dans différents États membres de l'UE, ce qui garantit un balayage géographique étendu;
   - représentaient différents types de sites web soupçonnés de porter atteinte au droit d'auteur, notamment les sites web de diffusion en continu, les sites web établissant des liens, les sites web d'hébergement, les sites d'hébergement de fichiers («cyberlockers») et les sites web de torrent;
   - représentaient un large éventail de contenus soupçonnés de porter atteinte au droit d'auteur, y compris des films, des programmes télévisés, de la musique et des jeux vidéo; et
   - représentaient des sites web que l'internaute moyen rencontrerait en tentant d'accéder à des contenus soupçonnés de porter atteinte au droit d'auteur.

   Cinq étapes ont été suivies pour sélectionner les sites web soupçonnés de porter atteinte au droit d'auteur. Les trois premières étapes ont été conçues de manière à identifier les sites web soupçonnés de porter atteinte au droit d'auteur les plus populaires dans les États membres de l'UE. Cette méthode a reproduit les scénarios dans lesquels un utilisateur moyen était susceptible de rechercher de tels sites web sans préciser, par exemple, le titre d'un film ou d'une chanson. Les deux dernières étapes ont été conçues de manière à identifier les sites web soupçonnés de porter atteinte au droit d'auteur qu'un utilisateur moyen était susceptible de rencontrer en cherchant des moyens de télécharger un titre populaire en particulier sans préciser de site web. Cette étape a été particulièrement importante, compte tenu de la présence de sites web malveillants suspects qui se livrent à une contamination des résultats de recherche, grâce auquel ils exploitent les sujets les plus tendance par une optimisation des moteurs de recherche. Ensemble, les deux approches couvraient les différentes façons dont un internaute moyen tente de trouver en ligne des contenus soupçonnés de porter atteinte au droit d'auteur.

   L'accent a été placé sur l'analyse simultanée de la présence de logiciels malveillants et de PPI spécifiques aux applications mobiles sur des appareils tels que les terminaux de poche et les tablettes, présence qui représenterait l'une des principales menaces de cybercriminalité émergentes. L'analyse a été limitée aux appareils Android en raison d'indications figurant dans la littérature existante d'une présence plus importante de logiciels malveillants dans les magasins

d'applications Android (à savoir Google Play) que dans le magasin iTunes d'Apple. La méthodologie a été spécialement conçue pour générer un échantillon d'applications mobiles qui:

- étaient populaires au moment de la collecte des données à l'échelle mondiale;
- représentaient différents types d'applications (pour y inclure des applications permettant la diffusion en continu, des applications de torrent et des applications d'hébergement);
- contenaient un large éventail de contenus soupçonnés de porter atteinte au droit d'auteur (y compris des films, des programmes télévisés, de la musique et des jeux mobiles) ou y donnaient accès; et
- représentaient ce qu'un utilisateur moyen d'un appareil mobile serait susceptible de rencontrer en tentant de télécharger ou d'utiliser une application permettant d'accéder à des contenus soupçonnés de porter atteinte au droit d'auteur.

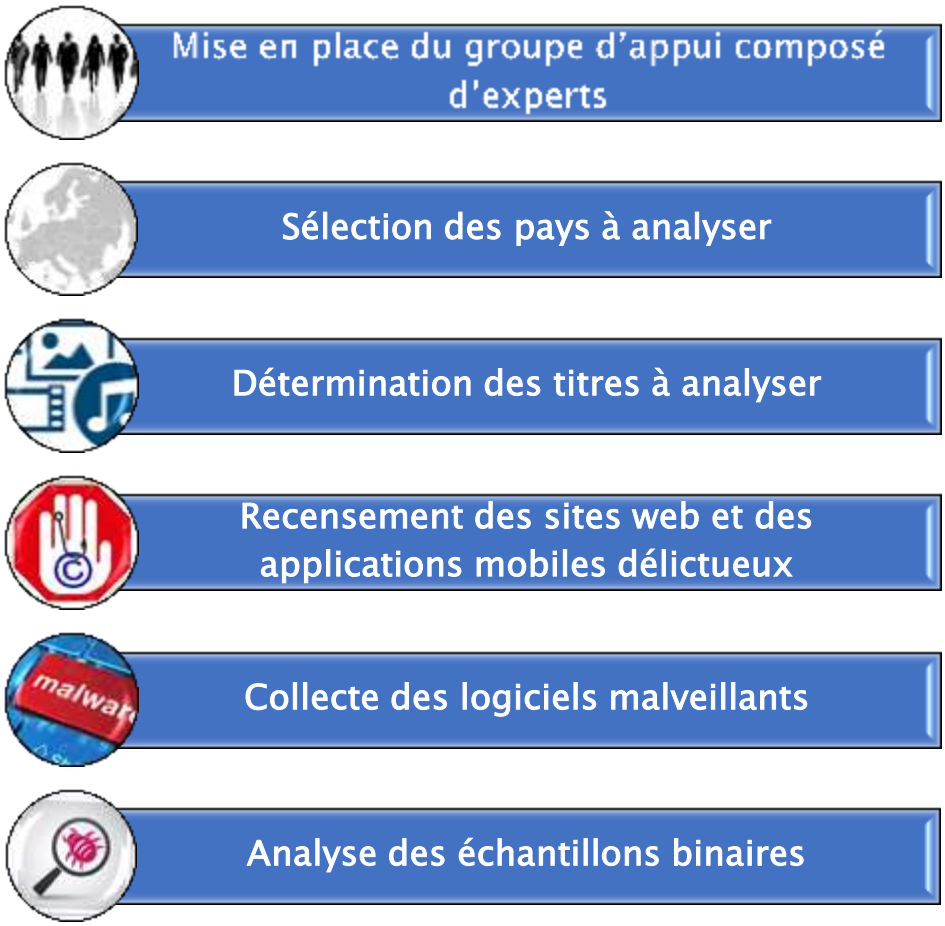
5. La phase V a consisté en la collecte de logiciels malveillants et de PPI en plus des applications mobiles sur les sites web identifiés, qui seraient examinés ultérieurement en vue d'un classement en catégories approprié. La phase d'acquisition des données a consisté en deux cycles de collecte et une analyse réalisée au cours de l'été 2017. Le premier cycle de collecte des logiciels malveillants a permis l'obtention de 1 054 noms de domaine uniques et le deuxième cycle de 1 057 noms de domaine uniques dans les 10 États membres de l'UE sélectionnés. Les logiciels malveillants ont été collectés de façon manuelle et de façon automatisée afin de simuler l'expérience d'un internaute moyen.

Collecte manuelle. Cette méthode impliquait un examen manuel des domaines identifiés lors de la phase précédente. Grâce à la collecte manuelle, l'expert a pu simuler l'expérience d'un internaute moyen en cliquant sur des annonces publicitaires et en interagissant avec les sites web qui nécessitaient des invites.

Collecte automatisée. Cette méthode a fait intervenir un robot d'indexation automatisé (web crawler), conçu par un expert, pour suivre tous les liens disponibles sur un site web soupçonné de porter atteinte au droit d'auteur désigné. Premièrement, sur un site web donné, le robot commençait par collecter des informations provenant des liens figurant sur la page d'accueil. Deuxièmement, le robot suivait chacun de ces liens menant vers des sites web secondaires. Troisièmement, le robot suivait chacun de ces liens menant vers des sites web tertiaires. À chaque étape, le robot récupérait des fichiers binaires susceptibles de présenter un intérêt pour l'analyse manuelle ultérieure, y compris les logiciels malveillants potentiels ou suspects et les programmes potentiellement indésirables. Ce processus a été réalisé sur un total de 1 000 liens au maximum par site web.

6. Une fois les fichiers binaires collectés, ils ont été analysés dans un environnement informatique sûr afin de comprendre leur fonctionnalité interne et en vue d'établir un classement par catégories adéquat. Une analyse préliminaire a été réalisée au moyen d'outils de source ouverte pour permettre la corrélation des conclusions et des rapports de cybermenaces. Les échantillons de logiciels collectés ont ensuite été livrés à EMAS pour analyse; l'analyse EMAS a alors été comparée aux résultats préliminaires.

Aperçu de la méthodologie



Échantillons des logiciels malveillants et PPI détectés

Au 28 juillet 2017, 5 240 sites web (1 054 noms de domaine uniques) avaient été automatiquement vérifiés au cours du premier cycle de collecte, avec 617 fichiers pertinents récupérés (musique, vidéos, fichiers de torrent et logiciels), dont la taille globale atteignait 47 Go. Ce lot de fichiers non triés a nécessité une analyse plus approfondie pour décider quels étaient les fichiers pertinents pour l'étude. Les échantillons de sites web portant atteinte au droit d'auteur étaient similaires dans les 10 pays de l'échantillon pour chacun des types de médias (programmes télévisés, films, musique et jeux vidéo). En conséquence, la Belgique a été choisie au hasard dans les pays de l'échantillon et tous les sites web identifiés comme portant atteinte au droit d'auteur pour la Belgique ont été vérifiés manuellement afin de détecter la présence de logiciels malveillants ou autrement indésirables. Le 10 août 2017, après le deuxième cycle de collecte, 3 665 fichiers au total ont été automatiquement récupérés à partir des sites web pour tous les pays, pour une taille totale de 167 Go. Étant donné que le nombre global d'URL uniques extraites pour tous les pays était de 1 057 sur les 5 606 sites web, il s'est avéré impossible de les vérifier toutes manuellement.

Après une analyse préliminaire des fichiers collectés, 106 fichiers binaires uniques pour MS Windows, Android et Mac OS ont été extraits à la suite des deux cycles de collecte de logiciels malveillants. Plus précisément, 41 dossiers ont été sélectionnés au cours du premier cycle et 65 au cours du deuxième — en particulier: 2 pour Mac, 15 pour Android et 89 pour MS Windows. Sur ces fichiers, 21 peuvent être considérés comme des programmes malveillants notoirement connus et signalés par plusieurs éditeurs d'antivirus comme provenant de la plateforme VirusTotal. Il s'agit notamment de fichiers téléchargés directement sur des sites web soupçonnés de porter atteinte au droit d'auteur, ainsi

que de fichiers créés au cours de l'exécution des fichiers téléchargés. Par la suite, les échantillons de logiciels collectés ont été analysés dans un bac à sable («sandbox»), puis livrés à EMAS en vue d'une analyse plus poussée des activités malveillantes éventuelles. Globalement, 821 événements malveillants distincts ont été découverts dans quatre rapports EMAS (Windows 7 SP1, Windows 7 SP1 64 bits, Windows 10 64 bits, Windows XP SP3) pour tous les fichiers binaires. Certains des rapports n'indiquaient aucune activité suspecte tandis que d'autres signalaient jusqu'à 10 activités malveillantes déjà connues. Au cours de la phase finale de l'étude, les résultats de l'analyse préliminaire et des rapports EMAS ont été mis en corrélation. La synthèse quantitative des résultats est présentée dans le tableau ci-dessous.

	Cycle 1	**Cycle 2**
Date	28 juillet 2017	jeudi 10 août 2017
Sites web découverts dans 10 pays de l'UE	5 240	5 606
Sites web uniques	1 054	1 057
Fichiers pertinents	617	3 665[2]
Taille des fichiers pertinents (Go)	47	167
Livrés à EMAS		
Android	3	12
Mac OS	2	–
MS Windows	36	53
Taille totale (octets)	175 600 117	522 991 095

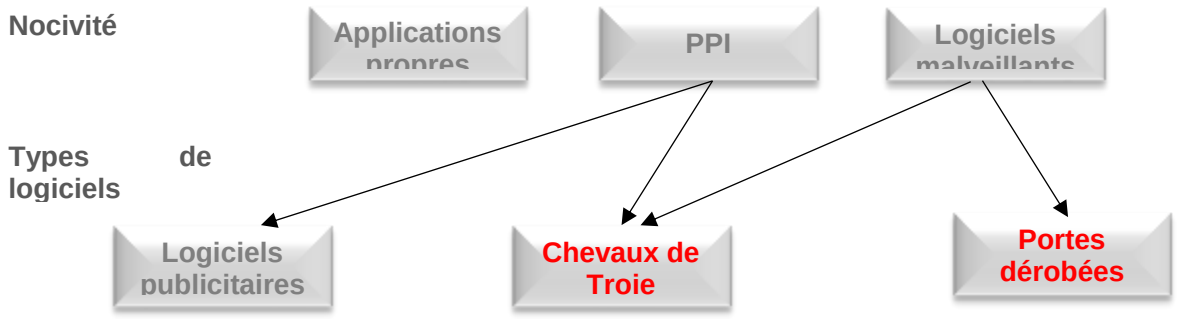
Solution d'analyse des logiciels malveillants d'Europol (EMAS)

La solution d'analyse des logiciels malveillants d'Europol (EMAS) est une solution d'analyse dynamique et automatisée de logiciels malveillants, fournie par Europol aux États membres de l'UE. EMAS donne la possibilité de créer des rapports d'analyse, mais sa caractéristique la plus révolutionnaire consiste à produire des renseignements à l'intention des enquêteurs de police. Les contrôles croisés automatisés peuvent faire apparaître des liens entre des attaques menées dans différents pays à l'aide des mêmes logiciels malveillants, ou par la même organisation criminelle à l'origine de la même famille de logiciels malveillants, qui se connectent aux mêmes domaines et font l'objetde différentes enquêtes dans l'UE ou dans des pays tiers. En 2015, la plateforme EMAS est devenue totalement automatisée pour permettre

Comme le montre le graphique ci-dessous, les fichiers binaires collectés peuvent généralement être classés en fonction de leur nocivité, en tant que fichiers bénins (fichiers n'occasionnant aucun dommage), PPI et logiciels malveillants nuisibles. En outre, des PPI ont été découverts non seulement pour Microsoft Windows, mais également pour Android et Mac OS, ce qui donne à penser que les développeurs de logiciels malveillants tentent d'affecter autant d'internautes que possible en utilisant différentes plateformes. Il est possible de différencier encore davantage les PPI et les logiciels malveillants en fonction des principaux types de logiciels malveillants, à savoir chevaux de Troie, logiciels publicitaires et portes dérobées. La plupart des logiciels découverts relevaient de la catégorie des PPI. Le fonctionnement des PPI peut être associé à l'un des modèles d'activité suivants: fausse installation de jeux nécessitant des données personnelles et bancaires; téléchargement de programmes «utiles» qui forcent les utilisateurs à acheter un abonnement à une version payante; ou

[2] La différence de chiffres entre le cycle 1 et le cycle 2 s'explique par le fait que, lors du cycle 2 de la collecte automatisée, certains sites web publiaient de multiples ensembles de fichiers sur chacune de leurs pages web.

installation de programmes gratuits pour accéder à des plateformes portant atteinte au droit d'auteur. Ces applications peuvent mettre en danger les données personnelles et la configuration de l'ordinateur des utilisateurs. Par des ruses d'ingénierie sociale, des données privées de différentes natures, telles que des informations sur les cartes de paiement, des informations d'identification personnelle et des données d'identification pour les comptes de réseaux sociaux, peuvent également être divulguées. De même, la recherche a permis d'identifier 15 applications Android provenant de marchés d'applications tiers et, après l'analyse préliminaire, il a été conclu que ces applications pouvaient participer à la distribution de contenus portant atteinte au droit d'auteur et à la divulgation de données à caractère personnel.



Menaces pour les utilisateurs finaux

Au cours de deux cycles d'identification de sites web et d'analyse de logiciels malveillants, aucun fichier binaire de rançongiciel («ransomware») n'a été découvert. De manière générale, la plupart des logiciels malveillants collectés peuvent être qualifiés de chevaux de Troie, ce qui signifie qu'ils pourraient être représentés sur les sites web comme des logiciels bénins couramment utilisés ou comme des logiciels populaires, alors qu'en réalité, ils peuvent dérober ou divulguer des informations privées. Il se pourrait qu'un internaute inexpérimenté place toute sa confiance dans le logiciel sans être en mesure de détecter une quelconque anomalie. En outre, l'analyse statique et les observations dynamiques du comportement de ces logiciels pourraient ne pas révéler toute leur fonctionnalité sans disposer d'un code source. À l'issue de l'analyse préliminaire des logiciels malveillants, l'analyse EMAS a mis en évidence des activités malveillantes plus spécifiques. L'installation de ces logiciels sur l'ordinateur d'un utilisateur final pourrait avoir des conséquences considérables, entraînant non seulement des pertes financières, mais aussi des vols de données à caractère personnel et d'autres risques d'accès et de contrôle indésirables. Des informations à caractère personnel pourraient ainsi être collectées et transmises à des tiers sous une forme cryptée ou dans un format de texte ouvert. Il pourrait s'agir, par exemple, de données d'identification d'un compte bancaire récupérées à partir du navigateur, de détails de la configuration matérielle ou logicielle de l'ordinateur ou, en substance, de toutes les informations saisies sur le clavier.

IDENTIFICATION ET ANALYSE DES LOGICIELS MALVEILLANTS SUR UNE SÉLECTION DE SITES WEB SOUPÇONNÉS DE PORTER ATTEINTE AU DROIT D'AUTEUR

SYNTHÈSE

Septembre 2018