Chapitre 7 :
Threat Intelligence et VirusTotal

Ce chapitre porte sur la Threat Intelligence (renseignements sur les menaces) et sur l'utilisation de VirusTotal.

La **Threat intelligence**(ou Cyber Threat Intelligence, souvent abrégée en « CTI ») vous aide à comprendre les attaques numériques et leur contexte, pour déterminer qui ou quoi se cache derrière ces attaques et quels liens existent entre les différentes attaques.

Par exemple, si vous assistez un journaliste qui a reçu un e-mail d'hameçonnage, vous voudrez probablement non seulement comprendre s'il s'agissait bien d'un e-mail d'hameçonnage, mais aussi comprendre de quel type d'e-mail d'hameçonnage il s'agit. Était-ce un e-mail d'hameçonnage ordinaire envoyé à des milliers ou des millions de cibles sans distinction ? Ou s'agissait-il d'un e-mail envoyé à une personne en particulier et peut-être à quelques autres par un cybercriminel qui les aurait pris pour cible ? L'assistance que vous apportez peut dépendre des réponses à ces questions.

Si une menace implique un fichier ou une application malveillante (« malware »), vous voudrez également comprendre qu'elle est capable de faire. Cela peut vous aider à atténuer la menace au cas où le logiciel malveillant serait exécuté, ainsi que de comprendre comment vous défendre contre elle à l'avenir.

La Threat Intelligence est un domaine immense, et il y a de nombreuses choses à apprendre à ce sujet, certainement plus que ce guide peut fournir. Si vous souhaitez approfondir le sujet, l'analyste du renseignement sur les menaces Katie Nickels a rédigé un plan d'auto-apprentissage en deux parties (partie 1 machine à remonter le temps ; partie 2 machine à remonter le temps).

VirusTotal

VirusTotal est un service en ligne populaire qui peut être utilisé pour fournir des renseignements de base et parfois plus avancés sur les menaces. VirusTotal a été fondé en 2004 en Espagne et a été acquis par Google en 2012. Au moment de la rédaction (novembre 2022), le front-end n'est pas particulièrement intégré avec d'autres services Google. Par exemple, les comptes sur le site ne sont pas directement liés à un compte Google.

La fonction initiale de VirusTotal était de servir de dépôt de logiciels malveillants où les fichiers pouvaient être vérifiés par de nombreux produits antivirus, et c'est probablement encore sa fonctionnalité la plus connue. Compte tenu des millions de morceaux de logiciels malveillants téléversés par les utilisateurs et les organisations au fil des ans, VirusTotal est probablement le plus grand dépôt de logiciels malveillants dans le monde, ce qui en fait un excellent endroit pour rechercher des liens entre différents types de logiciels malveillants.

Ce qui le rend encore plus utile, c'est que VirusTotal ne se concentre pas seulement sur les échantillons de logiciels malveillants, mais explore également les différents éléments avec lesquels les échantillons interagissent. VirusTotal analyse et agrège également les adresses IP, les noms de domaine et les URL, et trouve des liens entre eux. Vous pouvez donc téléverser une pièce jointe trouvée dans un e-mail, apprendre qu'il s'agit d'un logiciel

malveillant, apprendre qu'il est lié à un domaine qui, à un moment donné, pointait vers une adresse IP à laquelle un autre fichier de logiciel malveillant était connecté lors de l'exécution, et ensuite conclure que la pièce jointe originale que vous avez téléversée est liée à ce logiciel malveillant connu.

Avant d'ouvrir VirusTotal, il convient de souligner deux choses importantes. Tout d'abord, *ne téléversez* aucun fichier sur VirusTotal sans comprendre ce que vous faites. Deuxièmement, l'établissement de liens comme celui-ci n'est pas une science exacte et il est facile de tirer des conclusions erronées. Par exemple, les adresses IP sont fréquemment réassignées à de nouvelles organisations, de sorte que ce qui était une adresse IP malveillante la semaine dernière pourrait être bénin cette semaine. Méfiez-vous de tirer des conclusions hâtives, surtout si ces conclusions peuvent avoir des conséquences.

Concepts importants pour le renseignement sur les menaces

Cette section présente quelques concepts importants que vous devrez connaître pour le reste du chapitre et dans les chapitres à venir.

Ce chapitre offre un bref aperçu de ces concepts. Cependant, nous vous encourageons à utiliser votre moteur de recherche préféré pour explorer chaque sujet plus en profondeur. Les moteurs de recherche seront peut-être l'un des outils les plus importants dans votre travail de renseignement sur les menaces. Peu importe vos compétences, vous trouverez toujours le besoin de rechercher des informations, qui pourront souvent sembler étonnamment basiques. Se sentir capable et à l'aise de le faire est une compétence importante pour tout analyste des renseignements sur les menaces.

Indicateurs de compromission

Le terme **indicateur de compromission** (ou IoC pour faire court) désigne tout artefact ou élément de preuve trouvé lors d'une recherche sur un appareil ou un réseau compromis. Cela fait généralement référence aux adresses IP ou aux domaines auxquels le logiciel malveillant ou le système compromis se connecte, ou à tous les fichiers malveillants que le cybercriminel laisse sur votre système. Parfois, cela fait également référence à des adresses e-mail, des adresses Bitcoin, des phrases courantes utilisées par les cybercriminels ou des

artefacts plus ésotériques.

Si vous voulez publier un incident ou une menace que vous avez analysée, il est recommandé d'inclure un résumé distinct des IoC que vous avez trouvés. Cela aidera les autres utilisateurs à vérifier leurs propres systèmes et réseaux, par exemple en vérifiant les journaux du système ou du serveur pour détecter des preuves d'une attaque similaire.

Voici un aperçu plus détaillé de certains IoC que vous pourriez rencontrer.

Adresses IP

Les adresses IP sont des adresses lisibles par ordinateur composées de chiffres (et parfois de lettres) pour chaque appareil connecté à Internet. Contrairement à votre adresse ou numéro de téléphone, elles sont souvent rapidement réassignées par le réseau pour lui permettre d'accueillir plus d'appareils : par exemple, votre réseau Wi-Fi domestique peut retransmettre la même adresse IP à votre ordinateur portable, votre téléphone ou les appareils de vos amis en fonction de la disponibilité. Les noms de domaine, qui sont la base des adresses lisibles par l'homme que vous pouvez utiliser pour visiter des sites Web, « résolvent » ou sont connectés à une certaine plage d'adresses IP.

Quand nous parlons d'adresses IP, nous entendons généralement **IPv4** (adresses IP version 4). Il existe également des adresses **IPv6**, mais nous nous contenterons des adresses IPv4 dans ce guide, sauf indication contraire. Il n'y a pas d'adresses IP autres que les versions 4 et 6.

Vous n'avez pas à comprendre tous les détails du fonctionnement des adresses IP. Toutefois, il est important que vous puissiez au moins faire ce qui suit :

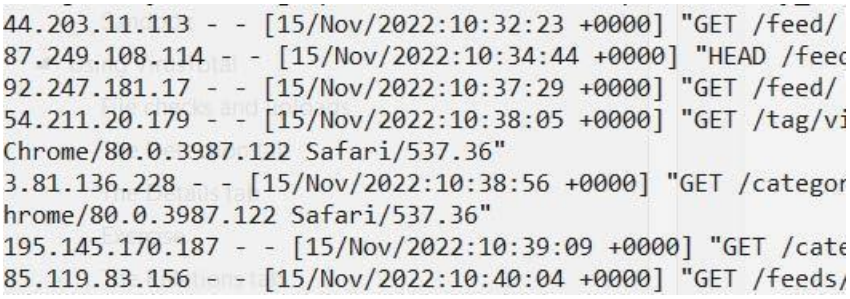
- Reconnaître à quoi ressemblent les adresses IPv4 et IPv6.
- Comprendre ce qu'est une adresse IP privée (v4) telle que `192.168.1[.]2`[1]
- Comprendre ce qu'est Tor et comment il cache l'adresse IP de l'utilisateur
- Comprendre ce qu'est un VPN et comment il cache l'adresse IP de l'utilisateur

Une façon de trouver une adresse IP qui crée une connexion suspecte consisterait à trouver dans le journal de votre site Web que l'adresse IP a demandé une URL qui semblait exploiter une vulnérabilité[2]. L'adresse IP que vous voyez pourrait être l'adresse IP publique de l'acteur malveillant. Cependant, de nombreux acteurs malveillants utilisent un type de proxy, tel que Tor ou un VPN, pour cacher leur adresse IP. Dans ce cas, l'adresse IP n'est peut-être pas pertinente pour votre recherche. Cependant, « l'acteur malveillant utilisait Tor pour se connecter » peut également être un élément de données utile à suivre ou à partager.

---

[1] Ignorez les crochets pour l'instant ; ils sont expliqués dans la section « désactivation » plus loin dans ce chapitre

[2] Cela est extrêmement courant, et la plupart de ces requêtes sont faites entièrement automatiquement et ne sont pas ciblées. Cela ne signifie pas que le serveur a la vulnérabilité particulière ou même qu'il exécute ce logiciel particulier !



Adresses IP dans un journal de serveur Web

Si vous voulez étudier l'adresse IP, la première question à laquelle vous devrez répondre est :

- Est-elle entièrement contrôlée par l'acteur qui gère la page d'hameçonnage (son propre site et non un site d'hébergement public, comme par exemple imgur ou YouTube) ? Dans ce cas, vous pouvez raisonnablement envisager que d'autres contenus hébergés sur le site sont liés au même acteur.
- Est-elle partagée avec d'autres contenus ? De nombreuses adresses IP hébergent plusieurs services non liés, tels qu'un serveur Web qui dessert plusieurs sites Web ou un serveur de messagerie qui dessert plusieurs domaines. Ces services ne sont pas nécessairement liés les uns aux autres. Notez également que du contenu légitime peut parfois partager la même adresse IP avec plusieurs types de contenus malveillants. Sans aucun autre type de preuve (par exemple, des pages d'hameçonnage qui ont un aspect très similaire), vous ne pouvez pas lier en toute confiance le contenu hébergé sur la même adresse IP au même acteur.
- Est-elle gérée par une entité légitime, mais quelque peu compromise ? Par exemple, l'acteur malveillant peut avoir piraté le serveur. Dans ce cas, vous devez comprendre le piratage et quand il s'est produit avant de pouvoir tirer des conclusions relatives à l'adresse IP.

  Il n'est souvent pas facile de savoir dans laquelle de ces trois situations vous vous trouvez, mais un service comme VirusTotal peut vous aider.

Un **nom de domaine** est la partie d'une adresse Web lisible par l'humain qui est achetée à partir d'un registre de domaine, principalement le domaine de premier niveau (`.com`, `.org`, `.co.uk`, `.in`, `.br`, etc.) et ce qui précède immédiatement ce point (`google`, `amazon`, `internews`, etc.). Un **nom d'hôte** fait référence à cette adresse ainsi qu'à tout sous-domaine que le propriétaire du domaine peut configurer dans son domaine. Donc internews[.]org ou google.co[.]uk sont des noms de domaine aussi bien que des noms d'hôte, mais `www.internews[.]org` et `mail.google.co[.]uk` ne sont que des noms d'hôte. Les termes « nom de domaine » et « nom d'hôte » sont souvent utilisés de façon interchangeable. Vous n'avez pas à comprendre tous les détails du fonctionnement des noms de domaine, des noms d'hôte et des DNS. Toutefois, il est important que vous puissiez au moins faire ce qui suit :

- distinguer les domaines de premier niveau à code de pays des autres domaines de premier niveau ;
- savoir ce qu'est un sous-domaine ;
- savoir comment utiliser la commande `dig` sur Linux (par exemple votre instance REMnux !) pour effectuer des recherches DNS (voici une bonne introduction)

- savoir ce que sont les enregistrements A, AAAA et MX ;
- savoir trouver la date d'enregistrement d'un domaine en utilisant la commande « whois » ;
- comprendre ce qu'est un registraire de domaine.

Lorsque vous rencontrez un nom d'hôte lors de l'analyse d'une menace, c'est généralement parce qu'il héberge du contenu ou qu'une connexion est établie avec le nom d'hôte. Si vous voulez étudier le nom de domaine ou le nom d'hôte plus en détail, déterminez si vous êtes dans l'une des situations suivantes :

- Le domaine est légitime et ne peut pas être utilisé à des fins malveillantes. C'est par exemple le cas pour `internews[.]org`. Les logiciels malveillants peuvent toujours se connecter à des domaines légitimes : parfois pour une raison explicite (ils peuvent, par exemple, utiliser un site comme `whatismyipaddress[.]com` pour déterminer l'adresse IP publique de l'appareil), parfois comme un leurre.
- Le domaine est légitime, mais aussi utilisé à des fins malveillantes. Par exemple, `drive.google[.]com` est clairement un site légitime, mais beaucoup de logiciels malveillants y sont hébergés.
- Le domaine est légitime, mais a été compromis et utilisé à des fins malveillantes. La compromission aurait pu se produire au niveau du serveur vers lequel pointe le domaine, mais aussi au niveau du DNS. Dans ce dernier cas, l'utilisation d'un site comme VirusTotal peut vous aider à vérifier l'historique DNS.
- Le domaine a été enregistré et créé à des fins malveillantes.

  Lors de l'enregistrement d'un domaine, les acteurs malveillants n'indiquent pas qu'ils vont l'utiliser pour mener des actes d'hameçonnage ou des logiciels malveillants. Donc, si vous voulez déterminer si un domaine a été créé à des fins malveillantes, vous devrez faire une enquête.

Les domaines enregistrés à des fins malveillantes :

- sont généralement enregistrés depuis peu ;
- ont souvent l'air très aléatoire (p. ex., `vniopquiopvqr[.]com`) ou ressemble à des domaines légitimes (p. ex., `internews-official[.]org`) ;
- utilisent le plus souvent des domaines de premier niveau inhabituels, comme .top ou .surf ;
- n'ont souvent pas d'enregistrement MX configuré ou utilisent l'enregistrement par défaut du registraire.

Les recherches DNS pour un domaine ne sont pas visibles par l'entité qui gère le domaine[3], mais c'est le cas des recherches actives d'un serveur Web ou de messagerie via une connexion à ce domaine. Gardez cela à l'esprit si vous ne voulez pas que votre enquête soit visible, et utilisez un VPN ou Tor pour dissimuler votre identité !

Désactivation

Lorsque vous traitez avec des noms de domaine ou des adresses IP potentiellement malveillantes, vous ne voulez pas que quelqu'un clique accidentellement dessus (notez que

[3] Ils peuvent remarquer une recherche si elle est faite pour un sous-domaine très spécifique, mais c'est un cas particulier.

très souvent, ils sont automatiquement transformés en noms). Vous ne voulez pas non plus qu'un logiciel de sécurité analyse le contexte dans lequel les artefacts sont partagés (par exemple, un produit de sécurité pour le courrier électronique) pour les détecter comme étant malveillants et déclencher une alerte.

Par conséquent, il est considéré de bonne pratique de **désactiver** les domaines et les adresses IP en plaçant tous les points ou au moins le dernier entre crochets. En pratique, la dernière option est généralement suffisante, c'est pourquoi nous l'avons appliquée dans ce guide. Donc, au lieu d'écrire `internews.org`, nous écrivons `internews[.]org` et au lieu d'écrire `127.0.0.1`, nous écrivons `127.0.0[.]1`.

Les artefacts du paragraphe précédent et de nombreux autres dans ce guide ne sont pas malveillants et pourraient, par conséquent, être omis. Cependant, il est courant de les désactiver tous.

De nombreux outils et services permettant de saisir des domaines et des adresses IP, y compris VirusTotal, supprimeront automatiquement les crochets. Vous n'avez donc pas à le faire avant de chercher quelque chose.

Question 7.1. Pourquoi n'avez-vous pas besoin de désactiver les noms de fichiers ? (Consultez l'annexe pour connaître la réponse.)

Hachages

Une **fonction de hachage cryptographique** est un algorithme mathématique qui prend des données de taille arbitraire (par exemple, un mot de passe ou le contenu d'un fichier) et les transforme en un nombre fixe d'octets : la **valeur de hachage** ou tout simplement **hachage**. Les hachages sont très pratiques pour de nombreuses raisons.

La première est la sécurité. Vous ne voulez pas partager des logiciels malveillants d'une manière qui pourrait être accidentellement les exécuter et causer des dommages. Le partage du hachage d'un échantillon est plus sûr.

La deuxième est la taille. Un hachage est petit, vous pouvez donc le partager dans un e-mail, un message Signal ou sur les réseaux sociaux sans avoir à joindre le fichier original.

La troisième raison est que le hachage sert d'empreinte digitale pour l'échantillon, ce qui permet de trouver facilement d'autres instances du hachage dans une collection d'échantillons de logiciels malveillants. Cela peut s'avérer très utile. Parfois, vous voudrez éviter de rendre publiques des informations de l'échantillon en plus de celles qui le sont déjà. Lorsque vous partagez que vous avez observé une pièce jointe malveillante avec un hachage particulier, les autres utilisateurs peuvent vérifier leur propre collection de pièces jointes ou les dépôts d'échantillons de logiciels malveillants comme VirusTotal, et voir s'ils ont déjà vu le même fichier.

Lorsqu'un acteur malveillant envoie plusieurs fichiers de logiciels malveillants dans une campagne, il lui est assez facile de modifier chacun d'entre eux de manière minime afin que les hachages soient complètement différents. Cela n'arrive toutefois pas souvent dans la pratique et les hachages continuent d'être un bon moyen pour aider les gens à déterminer s'ils ont affaire aux mêmes fichiers malveillants (et donc les mêmes cybercriminels) dans une

campagne d'attaques.

Il est important de noter que le nom du fichier ne fait pas partie des données hachées. Donc le hachage ne changera pas si vous renommez le fichier. Cependant, si vous changez ne serait-ce qu'un octet du fichier, le hachage change.

Les hachages cryptographiques ont les propriétés suivantes :

- le hachage peut être calculé rapidement, même pour des entrées très volumineuses ;
- il est en pratique impossible de calculer les données originales avec une valeur de hachage ;
- quelles que soient les données qui permettent de calculer la valeur de hachage, il est en pratique impossible de trouver d'autres données ayant la même valeur de hachage ;
- la plus petite des modifications apportées aux données donne une valeur de hachage complètement différente.

Il existe d'autres types de hachages qui jouent un rôle dans le renseignement sur les menaces et qui ne partagent pas cette dernière propriété. Il s'agit des hachages flous. Vous pouvez, par exemple, remarquer SSDEEP sur VirusTotal. Il s'agit d'un hachage flou.

En pratique, il existe trois fonctions de hachage pertinentes : md5, sha1 et sha256. Les deux premières sont plus âgées et légèrement imparfaites, mais il est peu probable que cela affecte votre travail. Lorsque vous pouvez choisir la fonction de hachage, il est recommandé d'opter pour sha256.

Un hachage sha256 est composé de 32 octets (ou 256 bits, d'où le nom) et écrit en 64 caractères avec les chiffres 0 à 9 et les lettres a à f. En raison des propriétés ci-dessus, les caractères particuliers d'un hachage n'ont pas d'importance : dans votre recherche, il s'agit simplement d'une chaîne de caractères.

Question 7.2. Vous ne pouvez pas simplement « inverser » le hachage pour trouver les données originales, étant donné la deuxième propriété ci-dessus. Cependant, si vous trouvez un hachage et n'avez aucune idée de ce que c'est, vous pouvez toujours utiliser un moteur de recherche et compter sur la chance. Quelles sont les données originales qui ont donné lieu aux hachages suivants ?

```
md5:d41d8cd98f00b204e9800998ecf8427e
sha1:da39a3ee5e6b4b0d3255bfef95601890afd80709
sha256:e3b0c44298fc1c149afbf4c8996fb92427ae41e4649b934ca495991b7852
b855
```

(La notation des deux points est commune pour indiquer le type de hachage, mais comme vous le voyez, les trois hachages ont également des longueurs différentes.)

Exercice 7.3. Sur votre instance REMnux, créez un fichier texte, saisissez le texte « `hello world` » (sans guillemets ou saut de ligne) et enregistrez le fichier en le nommant `test.txt`.

1 Ensuite, utilisez la commande `sha256sum` pour calculer le hachage sha256 du fichier. Notez la valeur de hachage.

Utilisez `md5sum` et `sha1sum` pour calculer également les hachages md5 et sha1.

2 Renommez maintenant le fichier en `newtest.txt` en utilisant la commande `mv`. Calculez le hachage sha256 du nouveau fichier. Confirmez qu'il s'agit du même fichier.

3 Maintenant, ouvrez le nouveau fichier pour l'éditer et transformez le « h » en un « H » (majuscule). Enregistrez-le et calculez le nouveau hachage sha256. Confirmez qu'il est différent et sans rapport avec le hachage d'origine.

4 Accédez à Malware Bazaar (assurez-vous de le faire dans un navigateur au sein de REMnux), trouvez sa base de données de logiciels malveillants et cliquez sur n'importe quel fichier. Téléchargez l'échantillon et décompressez-le (rappelez-vous qu'il est de pratique courante de mettre les logiciels malveillants dans un fichier zip avec le mot de passe « infected »).

La décompression d'un fichier zip fonctionne normalement avec la commande unzip, mais vous pouvez obtenir une erreur. Si c'est le cas, installez la commande 7z en exécutant

```
sudo apt-get install p7zip-full
```

et décompressez le fichier en exécutant

```
7z x [file]
```

Puis calculez le hachage sha256 de l'échantillon de logiciel malveillant.

Si vous faites cela correctement, vous remarquerez que le hachage est identique au nom du fichier. Cela est assez commun parmi les bases de données de logiciels malveillants !

Utiliser VirusTotal

Avec toutes ces informations de base, regardons VirusTotal, où vous pouvez comparer les domaines, adresses IP et URL de vos fichiers avec ceux téléversés par des analystes des menaces du monde entier.

Avant d'utiliser VirusTotal, vous devez comprendre comment le site fonctionne. Sa fonctionnalité principale est gratuite et ne nécessite même pas d'inscription. La création d'un compte gratuit vous donnerait des options supplémentaires, mais nous ne les utiliserons pas dans ce guide.

VirusTotal propose également des comptes payants, ce qui facilite la recherche dans l'énorme collection de fichiers pour certaines propriétés, comme un nom, et leur téléchargement. Voici pourquoi cela peut être important : supposons que vos adversaires, surtout les plus puissants, ont payé des comptes et qu'ils les utilisent pour rechercher des fichiers pertinents pour leurs cibles. Faites donc preuve de prudence lorsque vous téléversez des fichiers sur VirusTotal : *envoyez uniquement des fichiers que vous pouvez rendre publics*.

Vérification et téléversement de fichiers

Il y a deux raisons pour lesquelles vous pouvez vouloir téléverser un fichier. La première est de vérifier son contexte et d'en apprendre plus à son sujet : est-il malveillant, à quels domaines se connecte-t-il quand il est exécuté, etc. ?

La deuxième est de le partager avec l'industrie de la sécurité. La plupart des entreprises de sécurité sont des clients de VirusTotal, et le téléversement d'un fichier est une façon simple de partager le fichier avec toutes les parties intéressées. Cela pourrait les aider à détecter l'attaque spécifique à laquelle vous faites face et à empêcher qu'elle affecte d'autres personnes.

Il est à noter ici que la plupart des entreprises de sécurité voient des millions d'échantillons de logiciels malveillants par jour, et la plupart des analyses (et l'ajout de détection) sont entièrement automatiques. Cela est bien souvent suffisant, mais si vous voulez qu'ils prêtent une attention particulière au fichier pour une raison quelconque, assurez-vous de les contacter directement !

Lors de la vérification d'un fichier sur VirusTotal, effectuez les quatre étapes suivantes :

1. Calculez le hachage sha256 du fichier comme décrit ci-dessus. Vérifiez le hachage sur VirusTotal. S'il existe, le fichier a déjà été téléversé.
2. Si ce n'est pas le cas, décidez si vous voulez envoyer le fichier. Même si le fichier est malveillant, il peut avoir été conçu spécifiquement pour la cible et donc contenir des informations personnelles. Dans le cas d'un fichier ciblé, le téléversement sur VirusTotal indique qu'il a été ouvert et analysé. Il est également judicieux de décider de ne pas téléverser le fichier si vous avez simplement des doutes !
3. Si vous voulez le téléverser, considérez si le nom du fichier est un élément que vous souhaitez partager. Si ce n'est pas le cas ou si vous avez des doutes, renommez le fichier en `[sha256].[extension]` : conservez l'extension d'origine (.exe, .docx, etc.), mais renommez la partie située avant le point au hachage sha256. Il s'agit d'une bonne pratique (nous l'avons vu avec Malware Bazaar ci-dessus) et ne révèle rien concernant le fichier.
4. Téléversez le fichier à l'aide de l'interface Web de VirusTotal. Assurez-vous de ne téléverser que le fichier lui-même et non un dossier entier ou un fichier zip qui le contient !

L'onglet Détection

Dans les sections suivantes, nous utiliserons l'exemple 971c5b5396ee37827635badea90d26d395b08d17cbe9e8027dc87b120f8bc0a2[4], qui est un fichier .exe malveillant utilisé dans des attaques ciblées par un cybercriminel lié à l'Iran appelé APT42. Il a été mentionné dans un rapport de machine à remonter le temps de la société de sécurité Mandiant (maintenant détenue par Google).

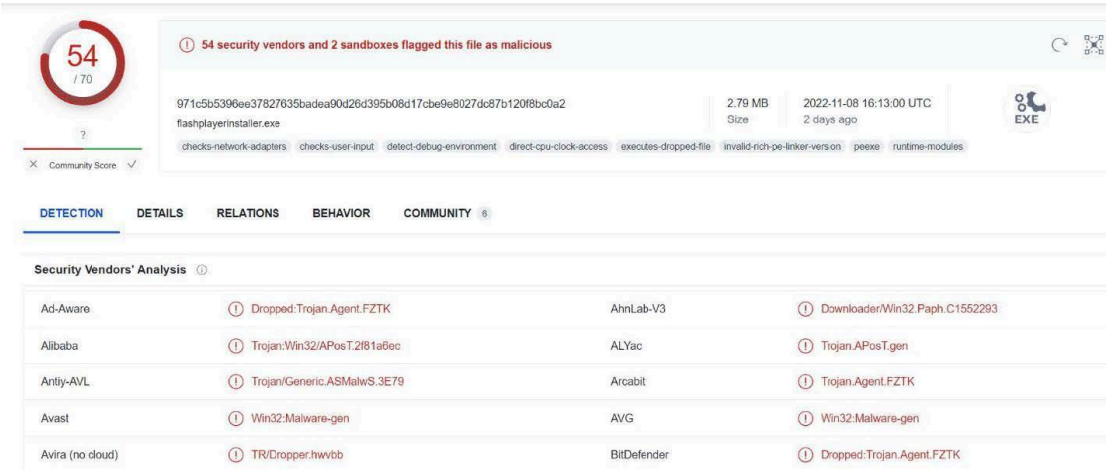
VirusTotal vous montre un en-tête comprenant plusieurs onglets. L'en-tête inclut le hachage sha256, la taille du fichier et la date de sa dernière analyse.

En dessous, le premier onglet est appelé détection. Il montre les résultats de l'analyse de

[4] Un exemple de référence à un fichier par son hachage !

dizaines de moteurs antivirus. Le mot « moteur » est utilisé ici délibérément : cette partie du produit n'examine généralement que le fichier lui-même plutôt que ce qu'il fait.

Il convient de noter que les faibles taux de détection sont assez courants, surtout pour les nouveaux fichiers, et que le nombre de détections devrait également être considéré comme un indicateur du degré de connaissance du fichier dans la communauté de la sécurité, et non pas de la façon dont un produit particulier aurait arrêté la menace dans une situation réelle[5].



Les fausses alertes : des fichiers inoffensifs ou légitimes détectés par erreur comme logiciels malveillants. Ces fausses alertes se produisent rarement. En pratique, tout ce qui dépasse quelques détections suggère qu'un fichier est malveillant.

Un conseil : si le fichier a été scanné pour la dernière fois il y a quelque temps, vous pouvez cliquer sur la flèche incurvée dans le coin supérieur gauche pour scanner à nouveau le fichier. La détection s'améliore généralement avec le temps, ce qui peut vous procurer une image plus précise.

Enfin, il y a deux choses à savoir à propos des moteurs. Certains sont en fait les mêmes. Par exemple, AVG a été acquis par Avast il y a quelque temps et les produits ont longtemps été exactement les mêmes. Ils sont uniquement répertoriés séparément parce que les marques existent toujours. Les noms de détection peuvent généralement être ignorés : ils sont souvent génériques, et quand ce n'est pas le cas, ils sont généralement faux.

Enfin, notez que si vous vous connectez à VirusTotal, vous verrez certaines règles de détection listées au-dessus des résultats antivirus. Ce sont d'autres règles de détection, et contrairement aux noms de détection antivirus, elles peuvent être utiles pour comprendre quel type de logiciel malveillant il s'agit.

L'onglet Détails

L'onglet Détails affiche plus d'informations concernant un fichier, comme divers hachages : sha256, md5, etc. Notre fichier est un exécutable Windows (vous verrez le type de fichier indiqué). Pour les autres types de fichiers, les informations sont différentes. Par exemple,

[5] Cela revient à VirusTotal en utilisant uniquement un moteur dépouillé du fichier, et non le produit complet. Comparez cela à un agent de sécurité qui n'aurait pas reconnu un voleur en fonction de son apparence, mais qui l'aurait quand même empêché de voler quelque chose.

pour les fichiers Android, elles montrent les autorisations demandées, ce qui peut s'avérer très utile.

L'onglet Détails vous indique quand le hachage a été soumis pour la première fois à VirusTotal, ce qui peut être un bon indicateur. Il vous indique également quand le fichier a été créé et quand il a été vu pour la première fois. Ces informations ne sont pas particulièrement précises, alors ne leur prêtez pas trop d'attention.

Vous remarquerez également les différents noms sous lesquels les fichiers ont été soumis. Parfois, cela vous aide à comprendre comment le fichier a été utilisé, comme on peut le voir dans l'exercice suivant.

Question 7.4. Une personne que vous assistez reçoit un fichier malveillant nommé `tax_information.docx` et vous soupçonnez qu'un cybercriminel la cible spécifiquement avec une campagne qui semble provenir du bureau local des impôts. Sur VirusTotal, vous trouverez le même fichier qui a également été soumis comme package_receipt.docx et birthday_gift.docx. Comment ces informations peuvent-elles vous aider ? (Consultez l'annexe pour connaître la réponse.)

L'onglet Relations

L'onglet Relations montre les relations entre le fichier que vous avez téléversé et d'autres IoC sur VirusTotal : URL, domaines et adresses IP contactés lorsque le fichier a été exécuté dans un bac à sable (vous en apprendrez plus sur ce sujet dans un chapitre ultérieur) ainsi que les fichiers téléchargés sur Internet ou créés directement lors de l'exécution du fichier.

Vous pouvez cliquer sur ces objets pour en apprendre davantage à leur sujet et éventuellement trouver des liens vers d'autres menaces. Cette méthode s'appelle « pivot ». Nous en discuterons plus loin dans le guide.

Exercice 7.5. Le domaine `update-driversonline[.]bid` se démarque parmi ceux qui sont connectés à ce fichier (rappelez-vous, nous parlons toujours de 971c5b5396ee37827635badea90d26d395b08d17cbe9e8027dc87b120f8bc0a2). Il y a de nombreuses raisons de soupçonner que cela est malveillant et certaines raisons de penser que ce n'est pas le cas. Quelles sont les raisons qui vous viennent à l'esprit ?

L'onglet Relations peut également contenir d'autres contextes, tels que les URL à partir desquelles ce fichier a été téléchargé ou les archives (comme un fichier zip) qui le contiennent. Toutes ces informations peuvent être utiles pour comprendre la menace !

Les onglets Comportement et Communauté

L'onglet Comportement décrit ce que fait le fichier. Pour un exécutable Windows, cela peut être très utile. Cependant, pour un analyste débutant, les informations présentées ici peuvent être assez déroutantes. Sans comprendre plus en détail ce que font les fichiers (pour lesquels vous pourriez, par exemple, utiliser un bac à sable, qui sera traité au chapitre 10), évitez de tirer des conclusions hâtives uniquement sur la base de cet onglet.

L'onglet Communauté contient des commentaires sur le fichier. De nombreux commentaires sont générés automatiquement et montrent plus de résultats du bac à sable. Les commentaires ajoutés par des humains sont particulièrement utiles. Si vous le souhaitez, vous pouvez créer un compte gratuit sur VirusTotal et fournir un contexte sur le fichier, et peut-être même une suggestion pour vous contacter si quelqu'un en sait plus. Dans ce dernier cas, gardez à l'esprit que cela est publiquement disponible, y compris par les cybercriminels. Veillez donc à ne rien ajouter qui puisse être lié à vous en tant que personne si vous considérez que ce n'est pas sûr de le faire.

Autres types de fichiers

VirusTotal peut gérer de nombreux types de fichiers différents. Les onglets et leur contenu peuvent être légèrement différents selon le type de fichier.

Consultez les onglets VirusTotal pour les trois exemples suivants et prenez quelques minutes pour vous familiariser avec chacun d'entre eux :

Un document Word malveillant :
2382d4957569aed12896aa8ca2cc9d2698217e53c9ab5d52799e4ea0920aa9b9

Un fichier de paquet Android (APK) :
86acaac2a95d0b7ebf60e56bca3ce400ef2f9080dbc463d6b408314c265cb523

Un exécutable macOS :
483b2f45a06516439b1dbfedda52f135a4ccdeafd91192e64250305644e5ff48

Question 7.6. Dans les exemples ci-dessus, consultez les onglets et essayez de répondre aux questions suivantes :

1. Le document Word contacte-t-il des domaines suspects ? Sinon, crée-t-il des liens suspects d'une autre manière ?
2. Le fichier Android demande-t-il l'autorisation d'activer le microphone ?
3. Sans utiliser un moteur de recherche autre que VirusTotal, pouvez-vous trouver un rapport sur l'exécutable macOS ?

(Consultez l'annexe pour connaître la réponse.)

Noms d'hôte et domaines

Maintenant, recherchez VirusTotal le domaine `update-driversonline[.]bid` que nous avons vu dans notre échantillon original.
Vous pouvez également accéder à cet exemple sur VirusTotal et cliquer sur ce lien dans l'onglet des relations.

Vous verrez que VirusTotal comprend également beaucoup d'informations concernant les noms de domaine. L'onglet Détection vous semblera familier. La détection de domaines est une science moins complexe que la détection de fichiers. Cependant, il n'y a pas toujours de distinction claire entre les domaines « bons » et « mauvais » (rappelez-vous les quatre types de domaines dont nous avons discuté dans ce chapitre, dans la section relative aux concepts importants pour le renseignement sur les menaces). Cette situation est similaire :

plus les fournisseurs le détectent comme étant malveillant, plus il est probable qu'il le soit.

L'onglet Détails contient plus d'informations au sujet du domaine, y compris les informations whois, qui vous indiquent quand il a été enregistré et, parfois, par qui. Les onglets Communauté et Relations vous sembleront familiers. Ce dernier est particulièrement utile. Vous pouvez l'utiliser pour pivoter et trouver des sous-domaines.

Question 7.7. Le domaine a été contacté par trois fichiers, y compris notre échantillon original. Pensez-vous que les trois fichiers sont liés ? Pourquoi pensez-vous cela ? (Consultez l'annexe pour connaître la réponse.)

Adresses IP

VirusTotal contient également des informations sur les adresses IP. L'onglet Relations est ici très utile : il montre que les domaines ont eu leur enregistrement A pointant vers une adresse IP. Cela peut être très utile lorsque vous essayez de trouver d'autres domaines liés à un domaine particulier.

Question 7.8. Utilisez l'entrée VirusTotal pour `update-driversonline[.]bid` et examinez les adresses IP associées. Pouvez-vous trouver d'autres domaines que vous pensez avoir été utilisés dans la même campagne ? (Consultez l'annexe pour connaître la réponse.)

URL

Enfin, VirusTotal propose également des pages sur les URL, mais elles ajoutent rarement des informations utiles concernant ce que le domaine ou les pages de nom d'hôte correspondantes affichent. Étant donné que les URL sont souvent uniques, leur ajout à VirusTotal peut fournir des informations sur votre analyse. Par conséquent, il est préférable de chercher le nom d'hôte.

La vérification d'un nom d'hôte ou d'un nom de domaine est presque toujours « correcte » : VirusTotal analyse activement les noms de domaine et vous n'avez pas à vous soucier de « téléverser » un nouveau domaine. Le seul cas dans lequel vous devriez faire preuve de prudence est celui où vous avez un nom d'hôte très spécifique, comme `[long string].domaine`, qui peut être unique à cette cible particulière. Cette situation est toutefois rare.

Détection des menaces

La détection des menaces est une activité proactive où vous recherchez les menaces plutôt que d'analyser celles qui existent déjà. Parfois, la détection des menaces vise une organisation ou une communauté spécifique sans aucun objectif précis, par exemple en recherchant des domaines nouvellement enregistrés qui utilisent le nom de votre organisation, ce qui pourrait signifier que l'auteur de la menace les enregistre pour mener une campagne d'hameçonnage.

Plus souvent, la détection de menaces commence par une menace existante, comme un fichier malveillant, puis cherche des artefacts connexes (autres fichiers, domaines, etc.) pour avoir une meilleure idée de la menace et de l'acteur qui se cache derrière elle.

VirusTotal est un excellent outil pour la détection des menaces, et dans la section précédente, nous en avons fait un peu l'exemple en recherchant des fichiers et des domaines connexes. Les comptes payants de VirusTotal sont excellents pour cela, car ils offrent beaucoup plus de possibilités pour la détection des menaces. Vous pouvez faire pivoter votre recherche pour examiner d'autres indicateurs que vous trouvez et utiliser également les « règles YARA », une méthode de recherche de fichiers dans une grande collection qui est particulièrement utile lors de la recherche de logiciels malveillants connexes.

Vous serez peut-être surpris(e) de voir à quel point vous pouvez obtenir des résultats avec des outils en ligne gratuits sans savoir comment faire de rétro-ingénierie des logiciels malveillants[6] ou écrire des règles YARA. Nous devons toutefois vous fournir deux mises en garde importantes.

La première mise en garde est de se méfier des conclusions audacieuses. Il est tentant de s'enthousiasmer[7] si vous reliez un logiciel malveillant à un acteur connu, surtout lorsqu'il s'agit d'un acteur avancé lié à un gouvernement. Parfois, ces liens sont bien là. Dans bien d'autres cas, ce n'est pas le cas. Par exemple, les deux acteurs peuvent avoir partagé une infrastructure tierce ou il y a eu une tentative délibérée de tromper les chercheurs. Une autre possibilité est que les domaines utilisés par les deux menaces ont fini par pointer vers le même sinkhole (nous en apprendrons plus sur les sinkholes prochainement).

La deuxième mise en garde est de ne pas oublier l'objectif de votre enquête. Vous n'avez pas à rechercher des menaces similaires pour chaque menace que vous examinez. Si la personne qui signale l'événement n'a pas ouvert la pièce jointe malveillante ou que la pièce jointe ne visait pas spécifiquement cette personne ou son organisation, il est normal de fermer l'enquête et de se concentrer sur d'autres choses.

Menaces ciblées et non ciblées

Les menaces numériques telles que les logiciels malveillants et l'hameçonnage ont traditionnellement été divisées en menaces ciblées et non ciblées. Dans le premier scénario, la menace vise un utilisateur ou une organisation individuelle ou peut-être un très petit nombre d'entre eux. Dans le deuxième scénario, n'importe quel utilisateur peut recevoir la menace et les cybercriminels ne font qu'attendre que l'un d'entre eux la reçoive et tombe dans le piège.

Le défi est que, dans la pratique, de nombreuses menaces non ciblées peuvent encore apparaître comme étant légèrement ciblées. Premièrement, il est très courant qu'une campagne non ciblée ne cible qu'un pays ou une région en particulier. Cela permet à son

[6] La rétro-ingénierie est le processus d'apprentissage de ce qu'un logiciel, souvent un logiciel malveillant, peut faire à partir du bytecode compilé

[7] Il est normal de s'enthousiasmer à ce sujet, même si les menaces sont sérieuses et affectent des personnes réelles de manière souvent grave ! Ne laissez pas votre désir d'aider les gens et les groupes être guidé par la menace qui vous passionne le plus. La grande majorité des menaces sont très banales

auteur d'écrire des messages dans la langue locale et d'ajouter un contexte local qui la rend plus crédible.

Deuxièmement, les auteurs de la menace utilisent souvent des données provenant de nombreux comptes de messagerie compromis lorsqu'ils envoient des e-mails malveillants. De cette façon, ils peuvent faire en sorte qu'un e-mail ressemble à une réponse à un e-mail que vous avez envoyé ou à un e-mail envoyé par l'un de vos contacts. L'e-mail concerné paraîtrait ainsi très ciblé, même si ce n'est pas le cas : les cybercriminels sont capables d'envoyer automatiquement un grand nombre d'e-mails, chacun d'entre eux apparaissant suffisamment personnel pour le destinataire.

Enfin, les auteurs de la menace, en particulier ceux qui se livrent à la cybercriminalité, ont trouvé un hybride entre les menaces ciblées et non ciblées, où une menace non ciblée entraîne la compromission d'un compte, souvent au sein d'une grande organisation. L'accès à ce compte est ensuite vendu à un autre acteur, qui utilise l'accès de manière plus ciblée, par exemple en déployant des rançongiciels sur tout le réseau. Même si la dernière menace était très ciblée, la menace originale ne l'était pas !

La grande majorité des menaces numériques ne sont pas ciblées, et cela est vrai même pour les menaces subies par ceux qui font également face à des menaces très ciblées.

Question 7.9. Un e-mail contenant une pièce jointe malveillante s'adresse au destinataire, un employé d'une organisation de la société civile, par son prénom et son nom de famille et semble provenir d'un autre employé de l'organisation. Pouvez-vous expliquer pourquoi il pourrait s'agir d'une menace automatisée non ciblée ? (Consultez l'annexe pour connaître la réponse.)

Sinkholes

La plupart des logiciels malveillants sont contrôlés par un serveur ou un ensemble de serveurs, qui sont contrôlés par l'auteur de la menace. Ce serveur ou ensemble de serveurs est appelé le serveur de commandement et de contrôle (souvent désigné comme C&C ou C2). La plupart des logiciels malveillants nécessitent une connexion à Internet pour pouvoir recevoir les commandes de leur serveur C&C. Les logiciels malveillants indiquent à votre appareil de se connecter à un ou plusieurs domaines qui pointent vers l'adresse IP du serveur C&C. Pointer vers un domaine au lieu d'une adresse IP permet aux acteurs de passer à un autre serveur si le serveur d'origine devenait indisponible.

Pour arrêter l'exécution d'un serveur C&C et du logiciel malveillant qu'il exploite, un organisme d'application de la loi ou une société de sécurité peut montrer à un registraire de domaine des preuves que le domaine a été utilisé à des fins malveillantes et convaincre le registraire de les laisser saisir le domaine. Ils peuvent alors pointer le domaine vers un serveur qu'ils contrôlent (souvent appelé « **sinkholing** » du domaine). Cela signifie que chaque fois que le logiciel malveillant tentera de se connecter à son serveur C&C, il se connectera au domaine sinkhole, en donnant ainsi à ses opérateurs une bonne idée de la menace. Parfois, l'organisme d'application de la loi ou l'entreprise peut alors envoyer des commandes pour neutraliser le logiciel malveillant si leur compréhension du fonctionnement du serveur C&C est suffisante.

Certains logiciels malveillants utilisent un algorithme générant des domaines (ou DGA) pour générer de nouveaux domaines, souvent quotidiennement. Cette manœuvre sournoise empêche l'analyste de déterminer facilement si plusieurs échantillons de logiciels malveillants pointent vers le même serveur C&C, car il ne pourra pas rechercher facilement des domaines identifiables. Les auteurs du logiciel malveillant enregistrent ces nouveaux domaines tant que leur campagne est en cours, et ils n'ont pas à trop se soucier des prises de contrôle ou des blocages de domaine par des produits de sécurité. Dans de nombreux cas, le DGA est « craqué », ce qui permet aux chercheurs de prédire quels domaines seront utilisés, de les enregistrer de manière proactive et de les diriger vers un sinkhole.

Lors de l'investigation d'un domaine malveillant, il est donc important de garder à l'esprit que le domaine que vous voyez est peut-être un sinkhole. Il n'existe pas de liste publique d'adresses IP de sinkhole. Si c'était le cas, les auteurs de logiciels malveillants feraient simplement en sorte que le logiciel malveillant ne se connecte à aucune adresse IP de cette liste. Et il n'est pas toujours évident de déterminer si une adresse IP est un sinkhole. Le pointage de nombreux domaines provenant de campagnes apparemment sans rapport vers l'adresse IP concernée est souvent un bon indice indiquant qu'il s'agit d'un sinkhole. En cas de doute, saisissez l'adresse IP dans un moteur de recherche, vérifiez l'onglet « Communauté » sur VirusTotal, ou demandez de l'aide autour de vous !

Question 7.10. L'utilisation d'un sinkhole pour envoyer des commandes afin de neutraliser les infections par des logiciels malveillants est quelque peu controversée. Pouvez-vous trouver des raisons qui soutiennent ce point de vue ? (Consultez l'annexe pour connaître la réponse.)

Autres outils

VirusTotal est un excellent outil pour l'analyse et la détection des menaces. Ce n'est toutefois pas le seul. Les moteurs de recherche sont peut-être les outils les plus utiles. Que vous préfériez Google ou l'une des nombreuses alternatives, vous serez peut-être surpris(e) de voir combien d'informations vous pouvez trouver en recherchant un nom de domaine, une adresse IP, un hachage de fichier ou un autre artefact.

Il en va de même pour les médias sociaux, notamment Twitter, bien qu'au moment de la rédaction (début décembre 2022), de nombreux chercheurs dans le domaine de la sécurité ont quitté Twitter pour migrer vers Mastodon. Le partage d'IoC était assez courant sur Twitter et pourrait le devenir sur Mastodon. L'un des avantages de la publication des IoC sur les médias sociaux est que vous pouvez répondre à la personne qui le publie et poser des questions ou ajouter vos propres commentaires.