Chapitre 10:

Analyse des charges utiles des courriers électroniques

Le chapitre 9 se concentre sur l'analyse des e-mails en eux-mêmes et ce chapitre vous enseignera quelques bases pour analyser les pièces jointes et les liens malveillants (aussi appelés « charge utile ») dans les e-mails. La première partie couvre les bacs à sable, qui sont un moyen pratique d'analyser des fichiers et des liens potentiellement malveillants, et la deuxième partie est une introduction à l'analyse manuelle des fichiers. Bien que, comme nous l'expliquions dans le chapitre, il s'agisse d'une solution que vous devriez éviter dans la mesure du possible.

Vu que ce chapitre traite des logiciels malveillants, assurez-vous d'avoir bien lu le chapitre 4.

Bacs à sable

Un bac à sable est un programme qui fournit un environnement sûr dans lequel exécuter un programme informatique. Un type particulier de bac à sable est le bac à sable pour les logiciels malveillants, dans lequel les logiciels malveillants peuvent être exécutés de manière contenue et leurs actions peuvent être analysées.

Le bac à sable recommandé dans ce chapitre vous permettra de téléverser un fichier à exécuter ou de vous connecter à une URL à charger dans un navigateur. Après un certain temps d'analyse, le bac à sable vous fournira un rapport concernant les activités ayant eu lieu sur la machine, en soulignant souvent, le cas échéant, les activités qui indiquent un comportement malveillant.

Dans la mesure où les logiciels malveillants utilisent souvent plusieurs étapes, le brouillage et d'autres techniques perturbant l'analyse, l'utilisation d'un bac à sable bien exécuté est souvent beaucoup plus rapide que la réalisation d'une analyse manuelle. Il est également plus facile d'obtenir une compréhension de haut niveau du fonctionnement des logiciels malveillants.

Vous pouvez, en théorie, configurer votre propre environnement de bac à sable en utilisant des machines virtuelles : vous créez une machine virtuelle (ou VM), généralement exécutant une version récente de Windows, et prenez un instantané, puis exécutez le logiciel malveillant sur cette VM. Après un certain temps, vous regardez ce qui s'est passé et revenez à l'instantané.

Cependant, cette approche présente de nombreux inconvénients :

Bien qu'une VM vous aide à éviter tout dommage causé par des changements apportés au système d'exploitation, vous l'exécutez toujours sur votre propre réseau. Les logiciels malveillants peuvent trouver des moyens de se propager d'une machine à l'autre. Ils peuvent également se connecter à Internet, et si ces connexions sont détectées, votre adresse IP peut se retrouver sur une liste noire. Certains services vous interdisent tout accès si vous vous connectez à partir d'une adresse IP sur liste

noire.

- Vous devrez vérifier manuellement toutes les modifications apportées à la VM et décider si elles indiquent un comportement malveillant. Vous devrez aussi capturer le trafic réseau et voir quelles connexions sont faites.
- Les techniques de perturbation de l'analyse recherchent souvent la présence d'environnements virtuels. Auquel cas, le logiciel malveillant ne s'exécutera pas s'il en détecte un. Certains exemples de ce que les logiciels malveillants pourraient rechercher dans une machine virtuelle sont les CPU typiques des machines virtuelles, une image d'arrière-plan qui est la version par défaut de Windows, un disque dur déraisonnablement petit ou très peu de fichiers présents sur le disque dur à part ceux présents après l'installation.
- Certains programmes malveillants établissent une connexion avec leur serveur de contrôle, qui vérifie s'ils ont déjà été exécutés à partir de cette adresse IP particulière. Si c'est le cas, ils ne s'exécuteront plus.

Certains de ces problèmes peuvent être relativement facilement résolus. D'autres sont plus compliqués et la liste ci-dessus n'est certainement pas exhaustive.

Heureusement, il existe des bacs à sable qui sont déjà en place pour atténuer ces défis. L'un des plus populaires est <u>Cuckoo</u>, qui est open source. Si vous voulez expérimenter un bac à sable, Cuckoo est un bon point de départ. Il a été écrit à l'origine par des personnes ayant des liens avec la communauté de la liberté sur Internet. Il est donc conçu pour la société civile.

Cependant, il convient de souligner que la version originale de Cuckoo, dont la dernière version a été publiée en 2018, n'est pas en développement actif¹ et donc moins susceptible de pouvoir gérer correctement les menaces actuelles. Un groupe de personnes du CERT-EE, le CERT national estonien, réécrit Cuckoo pour Python 3. Vous pouvez suivre leur progression sur <u>GitHub</u> et si vous vous sentez aventureux, vous pouvez installer cette version de Cuckoo, mais attendez-vous à faire beaucoup de dépannage pour la faire fonctionner.

Une option beaucoup plus simple consiste à utiliser des bacs à sable basés sur le cloud. Plusieurs options sont disponibles, telles que <u>ANY.RUN</u>, <u>Hybrid Analysis</u>, <u>Joe Sandbox</u>, <u>Triage</u> et même une version en ligne de <u>Cuckoo</u>. Tous ces outils ont des versions gratuites qui vous permettent de téléverser des logiciels malveillants et des URL, bien que certains nécessitent une inscription². Nous devons toutefois vous avertir : comme avec VirusTotal, l'utilisation de l'un de ces bacs à sable rend essentiellement l'analyse accessible au public. S'il semble probable que le logiciel malveillant vous a ciblé spécifiquement, signaler publiquement que vous l'analysez pourrait révéler au cybercriminel que vous enquêtez sur lui.

Chaque bac à sable a ses propres caractéristiques, mais ils partagent également de

¹ Le problème est que Cuckoo a été écrit à l'origine pour Python 2, une version déconseillée de Python

² Au moment de la rédaction du présent document, l'analyse hybride et Cuckoo peuvent être utilisés sans inscription, les services Any.run et Triage nécessitent une inscription gratuite et Joe Sandbox nécessite une inscription et l'approbation du compte

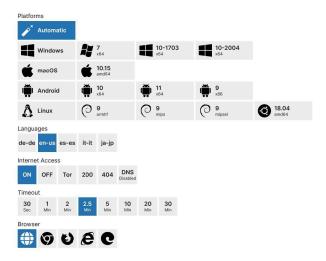
nombreuses fonctionnalités. Ce qui suit est une description de Triage, mais nous vous encourageons à essayer d'autres bacs à sable et, s'il y en a un que vous aimez particulièrement, à vous familiariser avec son fonctionnement.

Les captures d'écran suivantes sont prises à partir d'un document malveillant <u>découvert</u> (sha256:

8404d3dc32b0555bc3b076d7fc080d2a341508b4a2c84805a1d5ffc0057e2b39) sur Malware Bazaar, où vous pouvez le télécharger gratuitement si vous le souhaitez. L'analyse en bac à sable est <u>disponible</u> dans la version publique de Triage.

Dans Triage, utilisez l'option « Soumettre » en haut à gauche de la page pour téléverser un fichier. L'outil est très utile et vous permet de téléverser des fichiers protégés par mot de passe (le mot de passe par défaut qui sera essayé est « infected » : rappelez-vous que c'est la norme de l'industrie à utiliser !). Ainsi, vous n'avez pas à vous soucier de garder un fichier potentiellement malveillant sur votre ordinateur avant de le soumettre.

Vous pouvez également soumettre des URL, soit pour télécharger un logiciel malveillant directement à partir d'une URL et ensuite l'analyser, ou pour analyser une URL dans un navigateur.



Sur la page suivante, vous pouvez choisir différents paramètres pour le bac à sable. Le plus important est le système d'exploitation. Comme vous pouvez le voir, vous pouvez choisir parmi différentes versions de Windows, macOS, Android ou Linux. Si vous ne choisissez rien ici, vous pouvez demander à Triage de choisir un ou plusieurs systèmes d'exploitation pour vous, et cela fonctionne généralement.

Vous pouvez également choisir les paramètres de langue de la machine. Le paramètre par défaut est en-us : anglais des États-Unis. Parfois, les logiciels malveillants qui ciblent une région ou un groupe de personnes particulier ne fonctionnent que sur des machines avec un langage particulier. Ensuite, vous pouvez choisir de garder l'accès à Internet activé (par défaut) ou désactivé, ou de vous connecter via Tor. Beaucoup de logiciels malveillants se connectent à un serveur contrôlé par les cybercriminels, et si cela vous préoccupe (peut-être que vous ne voulez pas alerter le cybercriminel que vous effectuez une analyse), vous

pouvez tenter d'exécuter le bac à sable en désactivant la connexion à Internet ou de le configurer pour que les requêtes Web émulent des codes de retour 200 ou 404. Certains logiciels malveillants vérifient rapidement la présence de certains sites Web pour voir s'ils sont exécutés dans un environnement sans connexion à Internet, et cela pourrait servir à simuler une telle connexion. Certains programmes malveillants peuvent toutefois ne pas s'exécuter.

Vous pouvez également définir un délai d'attente : l'exécution s'arrêtera après une période d'inactivité (2,5 minutes par défaut). Une technique antianalyse courante consiste à attendre un certain temps avant d'effectuer toute activité malveillante, de sorte que dans certains cas, vous pouvez augmenter le délai d'attente.

Enfin, vous pouvez changer le navigateur que le bac à sable utilise pour ouvrir les sites Web. Cela peut être très utile dans le cas où la menace potentielle que vous analysez ne fonctionne que dans un navigateur particulier.

Vous pouvez maintenant commencer à analyser le fichier (ou l'URL), ce qui l'ouvre dans le bac à sable.

Une chose intéressante à propos de Triage est que vous pouvez voir ce qui se passe sur le système d'exploitation virtuel du bac à sable (ou les systèmes d'exploitation : il peut exécuter plusieurs analyses en parallèle!) et aussi interagir avec ce système d'exploitation. Parfois, cela peut s'avérer utile, par exemple si un logiciel malveillant surveille votre système et enregistre le clic d'un bouton. Dans la plupart des cas, il suffit que le bac à sable soit exécuté en arrière-plan.

Lorsque le bac à sable est terminé, vous pouvez afficher l'analyse. C'est ce qui vous intéresse le plus.



La page « Overview » est l'endroit que vous voudrez probablement regarder en premier. Elle vous donne un résumé de l'analyse du bac à sable. Dans ce cas, nous voyons que Triage donne à notre échantillon une note de 10 sur 10 et le marque avec le nom « agenttesla ». Agent Tesla est un logiciel malveillant commun utilisé pour voler des informations, telles que les identifiants, à partir d'une machine infectée. Vous noterez également des étiquettes telles que « stealer », « keylogger » et « spyware » qui indiquent encore plus clairement les menaces associées.

Dans la plupart des cas, cela suffira : vous savez que le fichier que vous vouliez analyser est l'Agent Tesla. Il y a suffisamment d'analyses de l'Agent Tesla disponibles sur Internet pour vous donner une idée de ce qu'il peut faire. Comme l'Agent Tesla est assez courant, Triage sait également comment extraire les informations de configuration du logiciel malveillant et affiche cela dans la section « Configuration du logiciel malveillant ».

Question 10.1. Sur la base de la configuration extraite du logiciel malveillant, comment cette variante de l'Agent Tesla extrait-elle les informations d'une machine infectée ? (Consultez l'annexe pour connaître la réponse.)

La section « Cibles » vous montre à nouveau qu'il s'agit de l'Agent Tesla, mais elle montre également ses activités malveillantes. Par exemple, un processus bloqué fait une demande réseau et un fichier³ « MZ/PE » est téléchargé. Rappelez-vous que le fichier que vous analysez est un document Office. Le fait qu'il tente de télécharger des fichiers est très suspect!

Cette section montre également les autres choses que le logiciel malveillant tente de faire, comme l'accès aux fichiers de configuration FTP et aux données des clients de messagerie et des navigateurs Web. Cette activité est courante pour les logiciels malveillants voleurs d'informations. Même si vous essayez d'analyser des logiciels malveillants que le bac à sable n'a pas reconnus, une telle activité est hautement suspecte.

Il y a aussi une section pour la « MITRE ATT&CK Matrix ». ATT&CK est un « framework » que les analystes des menaces utilisent comme langage commun, et plus facile à comprendre pour décrire et documenter les tactiques et techniques utilisées par les cybercriminels. Dans ce cas, cet onglet vous montre les techniques (telles que la collecte d'e-mails ou la modification du registre) utilisées par ce logiciel malveillant. ATT&CK a été conçu pour les entreprises et se concentre sur les menaces, et non sur des échantillons de logiciels malveillants individuels, bien qu'il soit toujours utile d'examiner les techniques enregistrées dans cette section.

La section « Replay Monitor » vous permet de reproduire ce que le logiciel malveillant a fait au système d'exploitation, ce qui s'avère très utile pour comprendre visuellement ce qui s'est passé. Enfin, si le logiciel malveillant a essayé de télécharger de nouveaux fichiers à partir d'Internet, vous pouvez les consulter dans la section « Download » et éventuellement les télécharger vous-même pour effectuer une analyse supplémentaire. Manipulez ces fichiers avec prudence, car ils peuvent être malveillants!

En plus de la page « Overview », vous pourrez aussi jeter un coup d'œil aux onglets du bac à sable correspondant au comportement du fichier. Dans ce cas, le fichier a été exécuté sur Windows 7 et Windows 10.

Examinons d'abord Windows 7. Vous trouverez les informations les plus intéressantes dans l'onglet « Rapport ». Une partie de ces informations étaient également disponibles sur la

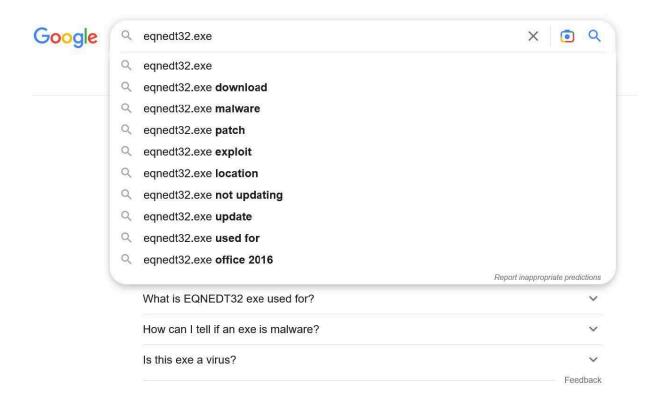
³ PE signifie « portable executable », tandis que MZ sont les deux premiers octets d'un fichier PE. Les exécutables Windows .exe et les bibliothèques Windows .dll sont des exemples de tels fichiers.

page « Overview », mais vous remarquerez quelques informations supplémentaires. Par exemple, dans la section « Signatures », il est indiqué que le fichier lance l'éditeur d'équations et que « Equation Editor est un ancien composant Office souvent ciblé par des exploitations de failles telles que CVE-2017-11882 ». Il est intéressant de noter que cette vulnérabilité est couramment exploitée pour cibler la société civile wayback machine.

Dans la section « Processus », vous pouvez voir quels processus le logiciel malveillant a démarré après l'ouverture du fichier. Vous pouvez voir que quatre processus ont été lancés : WINWORD.EXE, splwow64.exe, EQNEDT32.EXE et arnolded4874.exe, ce dernier ayant été lancé deux fois. Vous ne devez pas forcément savoir ce que tout cela signifie, mais ces processus peuvent être utiles pour essayer de comprendre ce qui se passe.

Un moteur de recherche vous dira que WINWORD.EXE est Microsoft Word, ce qui est logique, car il s'agit d'un fichier Word: Word s'exécute lorsque vous ouvrez le fichier malveillant. Le deuxième processus, splwow64.exe, est également un programme légitime qui concerne l'exécution de programmes 32 bits sur un système d'exploitation 64 bits.

EQNEDT32.EXE est également légitime, mais en regardant les suggestions des moteurs de recherche pour ce fichier, vous verrez « malware » et « exploit » mentionnés fréquemment, suggérant qu'il est couramment utilisé pour exécuter des logiciels malveillants. En effet, il s'agit de l'éditeur d'équations mentionné précédemment.



Enfin, il y a très peu de résultats pour arnolded4874.exe dans Google : au moment de la rédaction du présent article, il n'y en a qu'un, et c'est notre propre échantillon dans Malware Bazaar. Un fichier qui démarre un processus nommé de façon unique est extrêmement suspect !

En effet, dans la section suivante, « Réseau », nous remarquons une requête HTTP qui a été faite à une URL contenant ce même fichier. Le logiciel malveillant a donc téléchargé ce fichier à partir d'Internet, puis l'a exécuté sur votre système virtuel. Même en ignorant tous les signaux d'alerte précédents, cela devrait être très important pour un document Office.

Si vous ne trouvez pas suffisamment d'informations dans ces onglets, vous pouvez jeter un coup d'œil aux autres onglets dans lesquels vous trouverez beaucoup plus d'activités, comme l'ensemble de l'activité du réseau, les fichiers auxquels le programme a accédé et les clés de registre qui ont été lues et définies. Les clés de registre peuvent être liées aux paramètres globaux sur Windows. Elles contiennent beaucoup d'informations et il faut de l'expérience pour en comprendre tous les détails, mais vous trouverez quelques indices ici pour mieux comprendre les comportements potentiellement malveillants. Souvent, l'onglet « Report » vous donne toutes les informations dont vous avez besoin.

Rappelez-vous, l'échantillon a également été exécuté sur Windows 10. Il est intéressant de noter que rien de malveillant ne se produit dans ce cas : notez le score de 1, par opposition à un score de 10 (sur 10) pour Windows 7.



Il y a plusieurs raisons possibles à cela. La raison la plus plausible dans ce cas est que l'exploitation de faille n'a pas fonctionné : CVE-2017-11882 a été corrigé à la fin de l'année 2017. Windows 10 date de 2020 et il inclut le correctif qui a corrigé cette vulnérabilité.

Exercice 10.2. Téléchargez l'échantillon

37419d3a8a50d2e5bc0eef676a37d6757ba43a64eff868edb4af5c386900235f de Malware Bazaar et exécutez-le dans Triage. Partagez autant d'indicateurs de comportement malveillant que vous pouvez trouver dans l'onglet « Rapports ».

Exercice 10.3. Téléchargez l'échantillon

a43e0864905fe7afd6d8dbf26bd27d898a2effd386e81cfbc08cae9cf94ed968 de Malware Bazaar et lancez-le dans Triage. Recherchez des indicateurs malveillants. Si l'analyse ne donne pas un score de 10 sur 10, exécutez-la à nouveau, mais cette fois-ci, interagissez avec la machine et cliquez sur « Suivant » dans la fenêtre d'avertissement, tel que demandé par OneNote. Cela change-t-il le comportement détecté par le bac à sable et le score final ?

Exercice 10.4. (facultatif) Si vous utilisez régulièrement un autre bac à sable en ligne, exécutez-y également les échantillons des deux exercices précédents. Quels indicateurs de comportement malveillant trouvez-vous ?

Exercice 10.5. (facultatif) Recherchez des logiciels malveillants récents et exécutez-les dans un bac à sable de votre choix pour vous familiariser avec le fonctionnement du bac à sable. Malware Bazaar est un excellent endroit où découvrir de nouveaux logiciels malveillants,

mais il est encore plus utile de trouver un article qui analyse le logiciel malveillant en la mettant en ligne : vous pouvez alors comparer le résultat du bac à sable avec l'analyse fournie dans l'article pour comprendre les résultats de votre bac à sable.

Analyse manuelle des pièces jointes

Dans la section précédente, vous avez appris à quel point les bacs à sable sont utiles pour analyser les pièces jointes ou les liens trouvés dans les e-mails. Dans la plupart des cas, un bac à sable vous donnera toutes les informations dont vous avez besoin. Mais parfois, vous voudrez analyser une pièce jointe potentiellement malveillante de façon plus approfondie, parce qu'elle vous aidera dans votre travail ou simplement pour satisfaire votre curiosité. Cette section vous aide à comprendre comment effectuer une analyse manuelle pour explorer le fichier de façon plus approfondie.

Le courrier électronique est un vecteur commun de la propagation des logiciels malveillants depuis la fin des années 1990 (lorsque les logiciels malveillants étaient encore appelés virus⁴). Il suffit d'ouvrir une pièce jointe pour infecter votre ordinateur, après quoi le logiciel malveillant pourra procéder à ses méfaits et utilisera votre compte de messagerie pour envoyer une copie de lui-même à tous vos contacts.

Pour les auteurs de logiciels malveillants, les choses sont beaucoup plus difficiles à l'heure actuelle. Les filtres anti-spam efficaces compliquent fortement la livraison des fichiers exécutables⁵ de toutes sortes, y compris les logiciels malveillants⁶, à une cible. Les ordinateurs sont également mieux protégés, par exemple grâce à un antivirus intégré, contre les fichiers malveillants téléchargés sur Internet ou inclus dans des archives comme les fichiers .zip.

Ainsi, les auteurs du logiciel malveillant tentent de trouver des moyens pour contourner ces complications. Presque tous ces moyens impliquent de convaincre le destinataire d'agir, que ce soit en cliquant sur un lien, en activant des macros ou en contournant les règles de sécurité qui empêchent l'infection automatique par des logiciels malveillants.

Cela transforme la livraison de logiciels malveillants en un jeu du chat et de la souris. Les fournisseurs de sécurité améliorent constamment leur détection des nouveaux types de fichiers, et les fournisseurs de logiciels comme Microsoft compliquent encore l'utilisation abusive de leurs logiciels par les programmes malveillants. Mais les auteurs de logiciels malveillants continuent de trouver de nouvelles façons de livrer leurs créations.

Par conséquent, votre objectif en tant que personne qui réagit aux incidents de sécurité (et

⁴ Les premiers virus informatiques se comportaient de la même manière que les virus biologiques, en ce sens qu'ils infectaient des fichiers légitimes et s'exécutaient une fois le fichier « hôte » exécuté. Ces virus sont extrêmement rares de nos jours, mais le terme « virus » est encore souvent utilisé pour désigner les logiciels malveillants.

⁵ Fichier qui s'exécute sur l'ordinateur. Cela peut inclure les programmes légitimes, mais aussi beaucoup de logiciels malveillants. Sous Windows, les fichiers exécutables ont souvent l'extension .exe.

⁶ Dans certains cas, les logiciels malveillants ne sont même pas redirigés vers le dossier du courrier indésirable. Les filtres anti-spam rejettent beaucoup d'e-mails sans en informer l'utilisateur, surtout ceux qui sont malveillants. Ils peuvent également supprimer automatiquement les pièces jointes malveillantes.

l'objectif de ce module de formation) ne devrait pas être de savoir comment analyser tous les types possibles de charges utiles, mais plutôt de comprendre les bases et de savoir où chercher si vous trouvez un nouveau type de charge utile.

Les pièces jointes et les liens sont deux façons différentes pour un cybercriminel d'ajouter une charge utile (qui peut être malveillante) à un e-mail. Cependant, il est courant qu'une pièce jointe ne contienne rien d'autre qu'un lien vers la charge utile réelle ou un lien de téléchargement d'un fichier malveillant qui contient la véritable charge utile. Parfois, une pièce jointe contient un lien qui télécharge un autre fichier.

Documents Office malveillants

Comme mentionné précédemment, les fichiers Microsoft Office tels que les documents Word, les feuilles de calcul Excel et les présentations PowerPoint peuvent contenir une ou plusieurs macros : des morceaux de code qui sont automatisés. Il existe des raisons légitimes pour lesquelles les organisations utilisent des macros dans de tels documents, mais les macros sont depuis longtemps populaires parmi les auteurs de logiciels malveillants.

Vers le début du siècle, les macros s'exécutaient automatiquement dès l'ouverture des fichiers. Cela a conduit à l'apparition des vers dans les courriers électroniques : des e-mails comprenant des pièces jointes qui, lorsque celles-ci étaient ouvertes, exécutaient une macro qui envoyait une copie de la pièce jointe par e-mail à chaque personne reprise dans le carnet d'adresses. Sans surprise, Microsoft a désactivé l'exécution automatique des macros. Pendant une dizaine d'années, les logiciels malveillants basés sur les macros semblaient un phénomène relégué au passé.

Cependant, vers 2014, des auteurs de logiciels malveillants ont adopté une nouvelle tactique qui utilisait l'ingénierie sociale pour faire en sorte que l'utilisateur active les macros : le fichier affichait par exemple une page floue et les macros devaient être activées pour que le contenu soit affiché, en indiquant souvent un message tel que « pour des raisons de sécurité ». Cela est devenu un vecteur d'infection majeur pour de nombreux types de logiciels malveillants.

Récemment, Microsoft a apporté des <u>modifications</u> qui rendent plus difficile pour les auteurs de logiciels malveillants de faire en sorte que les utilisateurs exécutent les macros. C'est pourquoi ces logiciels malveillants sont désormais utilisés moins fréquemment. Cependant, vous pouvez toujours les rencontrer dans le cadre de votre travail.

Le meilleur outil pour analyser les documents Office est <u>oledump.py</u>, un outil écrit par le chercheur belge en sécurité Didier Stevens (il a également écrit emldump.py, que nous avons utilisé précédemment). Il est inclus dans REMnux, que vous avez probablement configuré au chapitre 6. Si ce n'est pas le cas, la vidéo de l'exercice suivant explique comment l'installer.

Exercice 10.6. Regardez l'<u>atelier YouTube</u> de Didier sur l'analyse des documents malveillants. Comme vous avez probablement installé REMnux, vous pouvez sauter la

première vidéo concernant la configuration d'oledump, mais gardez à l'esprit les points suivants si vous voulez suivre ce que fait Didier :

- Si vous voulez exécuter oledump sur REMnux, exécutez simplement oledump.py suivi des arguments, et non ./oledump.py comme le fait Didier⁷
- Faites précéder les noms des extensions avec /opt/oledump-files/, car elles sont stockées dans ce répertoire.
- Vu qu'il s'agit d'un atelier relativement long, nous vous recommandons de vous asseoir, de regarder la vidéo (peut-être à une vitesse légèrement plus élevée, car il est parfois un peu lent), et de ne pas vous concentrer sur tous les détails. Ce n'est pas un problème si vous ne comprenez pas ou ne vous souvenez pas de tout. Assurez-vous simplement qu'au terme de l'atelier, vous compreniez :
- Comment utiliser oledump pour afficher les différentes parties d'un fichier Office (exercice 1)
- Comment utiliser oledump pour afficher les macros VBA intégrées dans un document (exercice 6)
- Que les macros VBA malveillantes utilisent souvent le brouillage du code (exercice 17) et des chaînes/URL (exercice 20)
- Les commandes Linux less et head. Ces dernières ne sont pas expliquées explicitement dans l'atelier. Ce <u>site Web</u> wayback machine est une bonne introduction si elles ne vous sont pas familières.

Question 10.7. Didier explique pourquoi les downloaders sont généralement plus avantageux pour les auteurs de logiciels malveillants que les droppers. Pouvez-vous penser à un avantage que les droppers peuvent avoir sur les downloaders pour les auteurs de logiciels malveillants ? (Consultez l'annexe pour connaître la réponse.)

Maintenant, utilisons oledump dans la pratique. Nous allons le faire sur le fichier ayant le hachage sha256

3d76f59c4dceb13546eb9c72a7c0f03fd335093583d326de9a314f3dbd5a77cc qui a été <u>téléversé</u> sur MalwareBazaar juste avant la rédaction de ce guide. Comme c'est courant, le téléchargement est un fichier zip protégé par le mot de passe « infected ».

MalwareBazaar fournit déjà beaucoup d'autres informations sur le fichier qui peuvent être utiles si vous voulez l'analyser. Mais imaginons que nous avons trouvé ce fichier joint à un e-mail et qu'il n'y a aucune information publique disponible à son sujet.

Comme Didier l'explique dans ses vidéos, oledump peut traiter le fichier zip. Nous pouvons l'utiliser pour trouver si le fichier contient des macros VBA, ce qu'il fait à la partie 8 :

⁷ Si vous exécutez une commande sur la ligne de commande, Linux recherche un fichier exécutable avec ce nom dans l'un des répertoires (echo \$PATH vous montre lesquels). Si vous voulez exécuter un fichier à partir du répertoire courant, vous devez ajouter ce répertoire, et dans Linux, le répertoire courant est désigné par un point unique. Cela explique pourquoi dans la vidéo, où oledump est installé dans le répertoire local, la commande est précédée par ./

```
remnux@remnux:-$ oledump.py 3d76f59c4dceb13546eb9c72a7c0f03fd335093583c326de9a314f3dbd5a77cc.zip

1: 114 '\x3lCompObj'
2: 4696 '\x3bSbucunentSummaryInformation'
3: 4696 '\x3bSbucunentSummaryInformation'
4: 7401 'lTable'
5: 15599 'Data'
5: 441 'Macros/PROJECT'
7: 41 'Macros/PROJECT'
9: 3304 'Macros/VBA/ThisDocument'
9: 3304 'Macros/VBA/YBAP PROJECT'
10: 523 'Macros/VBA/YBAP PROJECT'
11: 4696 'WorcDocument'
remnux@remnux:-$
```

Nous pouvons maintenant afficher les macros :

```
Temnux@remnux:-$ oledump.py -$ 8 -v 3d76f59c4dceb13546eb9c72a7c0f03fd335093583d326de9a314f3dbd5a77cc.zip
Attribute VB Name = "ThisDocument"
Attribute VB Name = "ThisDocument"
Attribute VB Base = "INOrmal.ThisDocument"
Attribute VB GlobalNameSpace = False
Attribute VB ColbalNameSpace = False
Attribute VB PredeclaredId = True
Attribute VB PredeclaredId = True
Attribute VB Exposed = True
Attribute VB Exposed = True
Attribute VB TemplateDerived = True
Attribute VB TemplateDerived = True
Attribute VB TemplateDerived = True
Attribute VB Customizable = True
Private Declare Function szerajHBjvcsiduhusigwyqggfefevbbwsgfrbsjgbfvhjvJGYhjvgvIUGFyviygibfohdfugdkYGvhkvhjvGKME
et Lib "shell3z.adt." Alias
"shell5z.cuteA" (ByVal dgxWZECk As Long, _
ByVal Vtkboyh As String,
ByVal NuWxEQtHSPiuifYrolv As String, _
ByVal NuWxEQtHSPiuifYrolv As String, _
ByVal LibraphApArEoyxIdywhydugk As Long) As Long

Private Declare Function QuHGWbdUuJB Lib "urlmon" Alias _
"URLDownloadToFileA" (ByVal mTVqSXmlQeDEfoAdw As Long, _
ByVal ZhxXWn As Long) As Long

Sub tKwAbBMGYYsonfbnDSoc()
Dim mWDRLJ0zszLXJNk BBjbvCgIUGuyVcgHCGfxcGFDfcG As String
Dim YQUIGGGGGGGG SA Long,
ByVal ZhxXWn As Long) As Long

Sub tKwAbBMGYYsonfbnDSoc()
Dim mPMRLJ0zszLXJNk BBjbvCgIUGuyVcgHCGfxcGFDfcG As String
Dim YQUIGGFQCDFSXxcgyJbHbjbkhhiBHlGyvCDRxxrxdcyGhk) As String
Dim YQUIGGFQCDFSXxcgyJbHbjbkhhiBHlGyvCDRxxrxdcyGhk) As String
Dim QuIGGAddwbZRrRdybxZyJjnYklGjobfoylogfyviguteVoylyBgVgyylyBgYGyvhj As String
Dim UygujpsmbCzfCloijiBBhyycfeederSECytfTphGVGhgddsxSrErXfz0xfd As String
Dim UygujpsmbCzfCloijiBBhyycfeederSECytfTphGVGhgddsxSrE
```

Toutes les macros ne s'affichent pas à l'écran. Si vous exécutez cela vous-même, vous pouvez ajouter | less à la fin de la commande pour parcourir l'ensemble du résultat. Vous pouvez utiliser la barre de défilement à droite de la fenêtre du terminal pour afficher plus d'informations.

Prenons un peu de recul. Le créateur du fichier a utilisé des macros VBA et du brouillage. Même si vous n'êtes pas programmeur, vous devriez en déduire que ce fichier est malveillant. Si vous effectuez votre analyse pour vérifier s'il est malveillant, vous pouvez vous arrêter ici et tirer vos conclusions. Mais vous pouvez continuer votre analyse pour en savoir plus sur ce fichier.

Étant donné que les downloaders sont plus courants que les droppers, recherchons des URL. L'extension http_heuristics de Didier ne fonctionne pas ici (essayez vous-même pour confirmer), et ce n'est pas très surprenant. S'il était aussi facile d'extraire une URL du document, les produits de sécurité l'auraient déjà fait et pourraient vérifier la présence du domaine ou de l'adresse URL dans les listes noires. En conséquence, les auteurs de logiciels malveillants cachent assez bien les URL et continuent à trouver de nouvelles façons de le faire.

Donc, regardons les chaînes présentes dans le document : tout ce qui est entre guillemets ("..."). Il y en a une douzaine, mais la plupart sont trop courtes pour contenir une URL brouillée. La seule exception est la chaîne fyf/higehsvj0tuofuopdeffg0npd/ujmjlufn/xxx00;tquui.

À ce stade, vous pouvez faire deux choses. Vous pouvez noter que cette chaîne est l'argument de la fonction Uubhuyfbhf, qui est définie plus loin dans le code. Si vous avez quelques connaissances en programmation, vous pouvez en déduire ce que cette fonction fait et noter qu'elle enlève effectivement le brouillage de la chaîne d'une URL.

Vous pouvez également regarder la chaîne plus attentivement et noter qu'une URL commence par http:// ou https://, qui comprennent un double t et un double /. À la fin de notre chaîne, nous trouvons un double u et un double 0. Si vous regardez un peu plus attentivement, vous remarquerez que tquui orthographié à l'envers donne iuuqt, et que ces lettres donnent https si elles sont décalées d'une lettre dans l'alphabet.

Si vous remarquez que : est suivi de ; dans la <u>table ASCII</u> et que 0 est suivi de /, alors vous avez compris l'encodage : pour décoder la chaîne, nous devons l'inverser et ensuite, pour chaque caractère, prendre le précédent dans la table ASCII. Si vous pouvez programmer, vous pouvez écrire un script rapide qui peut le faire, ou vous pouvez simplement déchiffrer manuellement la chaîne et trouver ce qu'elle donne⁸

https://www.metkilit[.]com/feedcontents/iurgdfhg.exe

En effet, si nous avions cherché notre échantillon sur <u>VirusTotal</u>, nous aurions remarqué qu'il se connecte au domaine www.metkilit[.]com. Mais rappelez-vous, dans cette expérience, nous avons prétendu qu'il n'y avait rien de connu sur le fichier.

Maintenant que nous avons une URL qui est probablement l'endroit où le logiciel malveillant obtient ses téléchargements, nous pouvons enquêter de façon plus approfondie si nous le souhaitons. Plus loin dans ce quide, nous allons brièvement analyser les URL.

Comprendre le fonctionnement du brouillage des URL est une chose. C'est une autre chose de pouvoir le trouver par vous-même. Deux choses sont nécessaires pour y parvenir : la chance et l'expérience. La chance parce que le brouillage, dans ce cas, était assez simple et parce qu'il fallait tout de même le voir. Lorsque vous effectuez une analyse comme celle-ci, il est toujours bénéfique de travailler avec d'autres personnes : deux (ou plus) personnes sont beaucoup plus susceptibles de repérer un motif qu'une seule personne.

Examinons maintenant un autre logiciel malveillant :

8404d3dc32b0555bc3b076d7fc080d2a341508b4a2c84805a1d5ffc0057e2b39, l'échantillon que nous avons analysé dans Triage ci-dessus. Si nous exécutons oledump sur le fichier zip, nous obtenons une erreur indiquant qu'il ne s'agit pas d'un fichier zip valide. Vous pensez peut-être que la décompression du fichier est utile⁹, mais l'exécution d'oledump sur le fichier .doc donne la même erreur.

Heureusement, Linux dispose d'une commande utile file, qui nous indique qu'il s'agit d'un fichier au format texte enrichi :

⁸ Rappelez-vous la bonne pratique qui consiste à ajouter des crochets autour du dernier point d'un nom de domaine pour le désactiver. Ce point étant disponible au chapitre 7.

⁹ Comme expliqué dans le chapitre 7, vous devrez probablement exécuter 7z x [fichier] plutôt que décompresser unzip [fichier]

```
remnux@remnux:~/Desktop/Playground/0015/onedump × remnux@remnux:~ × 

remnux@remnux:~$ file 8404d3dc32b0555bc3b076d7fc080d2a341508b4a2c84805a1d5ffc0057e2b39.doc
8404d3dc32b0555bc3b076d7fc080d2a341508b4a2c84805a1d5ffc0057e2b39.doc: Rich Text Format data, version 1, unknown character set
remnux@remnux:-$
```

Ces fichiers RTF (qui ont souvent l'extension .rtf, mais qui peuvent aussi avoir la forme.doc, comme vous pouvez le voir) ont un format différent par rapport aux fichiers Word ordinaires. Heureusement, il existe un autre outil que Didier a écrit et qui est également inclus dans REMnux : rtfdump.py.

Comme oledump, rtfdump vous donne les différentes parties du document.

```
ⅎ
                                                                                        ۹ ≡
                                               remnux@remnux: ~
   remnux@remnux: ~/Desktop/Playground/0015/onedump
                                                                        remnux@remnux: ~
         nux:~$ rtfdump.py 8404d3dc32b0555bc3b076d7fc080d2a3415<u>0</u>8b4a2c84805a1d5ffc00<u>5</u>7e2b39.doc
                             2 p=00000000 l=
                                                27068 h=
                                                                                                6872 \rtf1
                                                             5104;
                                                                                     0 u=
                            0 p=0000000e l=
                                                   20 h=
                                                                                                   0 \*\dbrun
44865257
   3 Level 2
                            1 p=00000025 l=
                                                27030 h=
                                                             5104;
                                                                                                6872
                            2 p=0000287f l=
                                                                                                    -
0 \*\objda
a933725
        Level 4
                            0 p=00002890 l=
                                                   44 h=
                                                                         18 b=
                                                                                                   0 \emspace
75834615
 6 Level 4
nnux@remnux:~$
                            0 p=000028bf l=
                                                                                                   7 \*\line
                                                   50 h=
                                                               28:
                                                                         18 b=
```

Comme dans oledump, nous pouvons vider le contenu de chaque section. Malheureusement, les choses ne sont plus aussi simples à partir d'ici. Lorsqu'un document Office contient des macros VBA malveillantes, le logiciel malveillant est une fonctionnalité d'Office, même s'il s'agit d'une fonctionnalité plutôt indésirable. Mais dans ce cas, le logiciel malveillant agit comme un bug. Plus précisément, il utilise une vulnérabilité d'Office nommée CVE-2017-11882. Cela rend le logiciel malveillant plus difficile à trouver, en plus des difficultés causées par le brouillage dans le code malveillant.

Pour décoder le logiciel malveillant, un analyste devrait probablement comprendre le fonctionnement interne des fichiers RTF et de l'exploitation de faille CVE-2017-11882, et pourrait même avoir à adapter l'outil rtfdump. En effet, Didier continue d'ajouter de nouvelles fonctionnalités à ses outils en réponse aux nouveaux défis comme celui-ci.

L'incapacité à décoder les logiciels malveillants peut s'avérer frustrant et un peu gênant ! C'est aussi la réalité de l'analyse des logiciels malveillants. Une leçon importante de cet exemple est que nous avons pu analyser le fichier dans Triage, qui nous a donné tout ce que nous devions savoir. Ceci fournit un argument supplémentaire pour utiliser des bacs à sable au lieu de se fier uniquement à des analyses manuelles.

Exercice 10.8. (facultatif) À titre de référence, voici une <u>analyse</u> d'un autre fichier RTF qui pourrait être analysé en utilisant rtfdump. Pouvez-vous trouver le code malveillant à l'intérieur

du document RTF en utilisant les outils de REMnux?

L'expérience est acquise par la pratique et en lisant ce que font les autres. Ce processus est long, et cela n'empêche pas les analystes de logiciels malveillants avancés de se retrouver parfois coincés. Ne vous découragez pas et ne supposez pas que vous devez être capable de gérer toutes les pièces jointes. Même les analystes les plus expérimentés se retrouvent régulièrement bloqués.

Pièces jointes PDF

Les PDF sont d'autres pièces jointes courantes. Il fut un temps où les vulnérabilités dans Adobe Reader, le logiciel PDF le plus populaire, étaient très courantes. L'utilisation de fichiers PDF malveillants était un moyen courant d'infecter les appareils exécutant une version légèrement obsolète de Reader.

Cela est moins courant de nos jours, mais les PDF malveillants existent toujours. Parfois, un autre document est intégré dans le PDF, comme un document Office malveillant. D'autres fois, les PDF contiennent simplement un lien qui télécharge la charge utile.

Encore une fois, Didier Stevens a développé plusieurs <u>outils</u> pour faciliter l'analyse des PDF. Il a également créé un <u>atelier vidéo</u> pour démontrer certains de ces outils. Cependant, cet atelier a 11 ans et se concentre principalement sur le type de logiciel malveillant PDF qui était habituel à l'époque. Si vous travaillez régulièrement avec des fichiers PDF, vous pouvez toujours profiter de l'atelier, mais considérez-le comme facultatif.

L'outil principal de Didier est pdf-parse.py, qui comme son nom l'indique, analyse les fichiers PDF.

Prenons à nouveau un échantillon téléchargé à partir de Malware Bazaar :

26509fa876a07966824bed8e5b0a6e0626b37d68355871fac49b2fa636d7fe6e.

Pour commencer, nous exécutons :

```
pdf-parser.py
26509fa876a07966824bed8e5b0a6e0626b37d68355871fac49b2fa636d7fe6
e.pdf |less
```

sur une seule ligne pour parcourir les différentes composantes du fichier. Notez comment un fichier PDF est constitué de divers objets numérotés. Si nous parcourons ces objets, nous remarquons une adresse URL dans l'objet 3, hébergé sur l'espace de stockage de Google Firebase. Il s'agit d'un service légitime (si vous ne le saviez pas, une recherche rapide sur Internet vous le confirmerait), mais il est couramment utilisé pour héberger des logiciels malveillants.

Et c'est tout ! Nous avons extrait l'adresse URL du fichier PDF, et c'est tout ce que vous devez savoir. (Pour le vérifier, vous pouvez examiner les autres objets présents dans le fichier pour confirmer qu'il n'y a rien d'autre d'inquiétant.)

Voici un autre fichier PDF à consulter : 907e75030b0e09cec6524f612f1c7439b5260b57b43d515968f81ba69278ba77.

Pour parcourir les différents objets, nous pouvons lancer pdf-parser.py sur ce fichier, suivi par | less. Malheureusement, cela ne produit pas d'adresse URL, nous devons donc étudier les objets plus attentivement. Il peut s'avérer utile de regarder la longueur des objets : beaucoup sont longs de 10 octets, ce qui signifie qu'ils ne contiennent probablement rien de malveillant.

Nous pouvons voir des objets plus grands, en commençant par l'objet 61, qui indique / Subtype / Image. Comme vous pouvez probablement le deviner, ces fichiers contiennent des images intégrées dans le PDF. Mais il y a aussi l'objet 82, qui est un fichier incorporé (/ Type / EmbeddedFile). L'ajout de l'option $- \circ$ à la fin de la commande que nous venons de saisir nous permet d'afficher cet objet :

L'outil pdf-parser extrait utilement l'objet en utilisant l'option -d (signifiant « dump ») suivie d'un nom de fichier. Comme l'objet est codé (voir /Filter /FlateDecode), nous pouvons aussi utiliser -f pour le décoder.

Ici, nous stockons les données extraites dans un fichier, test, puis nous utilisons la commande file pour trouver la nature de l'objet. Il s'avère qu'il s'agit d'un document Excel! Si nous le voulions, nous pourrions analyser cela de façon plus approfondie avec oledump.

```
\odot
                                                                                Q
                                            remnux@remnux: ~
                                                                                     Ξ
         remnux@remnux:~9
o 82 -f -d test
This program has not been tested with this version of Python (3.8.10)
Should you encounter problems, please use Python version 3.7.5
obj 82 0
Type: /EmbeddedFile
Referencina:
Contains stream
   /Filter /FlateDecode
   /Type /EmbeddedFile
/Length 181498
emnux@remnux:~$ file test
test: Composite Document File V2 Document, Little Endian, Os: Windows, Version 6.2, Code page: 1252,
Name of Creating Application: Microsoft Excel, Create Time/Date: Sat Sep 16 01:00:00 2006, Last Saved
Time/Date: Fri Jan 27 01:34:39 2023, Security: 0
```

Exercice 10.9. Utilisez l'option -d dans pdfdump.py pour extraire une image de l'ODF.

Exercice 10.10. Utilisez l'outil pdf-parser pour analyser le fichier PDF

ad517cb885ee279ec6ca95cd7402da998ec5461461f745c2f075085ef49b4eb6.

Exercice 10.11. (facultatif) Ouvrez les deux fichiers PDF dans Triage ou un autre bac à sable de votre choix et confirmez ce que l'analyse manuelle a montré.

Autres types de pièces jointes

Il existe de nombreux autres types de pièces jointes utilisées pour propager les logiciels malveillants, et les acteurs malveillants en trouvent toujours de nouveaux à utiliser. Comme mentionné précédemment, vous ne devriez pas vous attendre à savoir comment les analyser toutes. Si vous rencontrez un nouveau type de fichier que vous ne savez pas comment gérer, cherchez la solution sur Internet. Il est très probable que quelqu'un ait écrit un outil et/ou un guide sur la façon d'analyser ce type de fichier particulier.

Il se peut que ce soit Didier Stevens. Il est utile de suivre son <u>blog</u> et ses publications sur le <u>blog</u> de l'<u>Internet Storm Center</u>. Par exemple, au début de l'année 2023, il a écrit comment analyser <u>les fichiers OneNote</u> <u>wavback machine</u> et <u>les fichiers HTA</u> <u>wavback machine</u>.

Analyse des liens

Contrairement à un document malveillant, un lien Web (ou une adresse URL) ne contient rien, il renvoie simplement à une ressource située quelque part ailleurs. Il peut s'agir d'une page HTML, d'un logiciel malveillant ou même d'une page donnant une erreur 404. Ce qui est renvoyé par le lien dépend parfois de la façon et du moment où votre appareil l'ouvre. De nombreuses URL cessent de fonctionner après un certain temps, ce qui est particulièrement vrai pour les adresses malveillantes. Cependant, même dans ces circonstances, l'analyse est encore possible.

Tout d'abord, il y a le nom de domaine qui a été utilisé. Il s'agit souvent d'un indice important si vous tentez de déterminer si le lien est ou était malveillant. Reportez-vous au chapitre 7 pour avoir un rappel sur la façon d'utiliser VirusTotal et obtenir plus d'informations sur un nom de domaine potentiellement malveillant.

Deuxièmement, vous pouvez regarder ce qui est chargé lorsque vous tentez d'accéder au lien. Ici, vous aurez besoin de chance : la ressource peut avoir été supprimée ou modifiée. Même si la ressource est toujours là, le serveur malveillant peut ne pas vous la renvoyer en fonction de votre adresse IP, navigateur Web ou d'autres caractéristiques. Et, bien sûr, vous devrez la télécharger dans un environnement sécurisé, tel que REMnux. L'exercice facultatif 10.12 ci-dessous vous donne des conseils sur la façon de le faire.

En général, il est plus sûr d'ouvrir l'URL dans un bac à sable. Cela ne vous aidera pas si la ressource a été supprimée ou modifiée, mais si elle est toujours là, il est beaucoup moins dangereux de l'ouvrir dans un bac à sable. Rappelez-vous que dans Triage, la section « Téléchargements » vous donne accès aux fichiers téléchargés au cas où vous voudriez les analyser plus en profondeur.

Au début des années 2010, les téléchargements furtifs étaient un moyen courant d'infecter les ordinateurs. Les sites Web malveillants ou infectés détectaient que votre navigateur ou une extension de votre navigateur (telle que Flash Player ou Java) était vulnérable et utilisaient cette vulnérabilité pour installer des logiciels malveillants sur votre ordinateur. Aujourd'hui, les navigateurs modernes bloquent les extensions à risque et installent généralement automatiquement les mises à jour de sécurité, ce qui rend ces téléchargements furtifs beaucoup plus rares. Ils existent toujours, cependant, et dans certains cas, ils exploitent les vulnérabilités zero-day dans les navigateurs.

Pour cette raison, l'ouverture de liens suspects dans un navigateur n'est toujours pas judicieuse, même si le risque est beaucoup plus faible qu'il ne l'était autrefois.

Exercice 10.12. Recherchez une URL malveillante récente dans <u>URLhaus</u> (un site Web créé par les mêmes personnes qui gèrent également Malware Bazaar) et ouvrez-la dans Triage ou dans un autre bac à sable de votre choix. Vérifiez que le contenu téléchargé est visible dans l'onglet « Téléchargements » de Triage. Si rien ne se passe, essayez une autre URL. (Cet exercice vous pousse délibérément à chercher des URL par vous-même plutôt que de vous les suggérer.)

Exercice 10.13. (facultatif) Lisez cet wayback machine article de blog sur l'utilisation de curl

pour télécharger le contenu à partir d'une URL potentiellement malveillante et l'essayer avec des URL récentes de URLhaus.

Clients de messagerie vulnérables

Enfin, il existe un autre type d'e-mail malveillant rare qui mérite d'être mentionné : celui dans lequel la vulnérabilité d'un client de messagerie est exploitée directement par un e-mail spécialement conçu. En mars 2023, il a été découvert wayback machine qu'un groupe de piratage lié à la Russie, connu sous le nom de Fancy Bear ou APT28 (ou plusieurs autres noms), avait utilisé une vulnérabilité zero-day dans Outlook pour infecter des ordinateurs de cette manière.

Il n'existe pas de méthode générique pour analyser ces e-mails. Les cybercriminels peuvent utiliser différents types de vulnérabilités. Ils sont très prudents sur l'utilisation des vulnérabilités zero-day. Souvent, les cybercriminels ne veulent pas révéler leur capacité à les exploiter, au risque que les développeurs du logiciel de messagerie l'apprennent et corrigent la faille. Heureusement, ces situations sont très rares. Mais cela souligne l'importance de s'assurer que les clients de messagerie sont tenus à jour. C'est quelque chose que les clients de messagerie Web font automatiquement.