Anti-forensic

Laurent Roger

DGA/DET/CELAR
laurent.roger@dga.defense.gouv.fr

1 Introduction

D'après mes recherches, ce n'est qu'à partir de 2001 que le terme « antiforensic » est véritablement introduit dans la littérature du domaine. Sous ce terme, sont désignées les procédures et techniques ayant pour objectif de limiter les moyens d 'enquête ou d 'examen d 'un système. Plusieurs moyens seront donc utilisés à cette fin : détruire, camoufler, modifier des traces , prévenir la création de traces.

Ce domaine particulier n'a jamais été examiné de façon systématique, le présent article propose un début d'exploration au travers d'une analyse des publications abordant le thème, d'un examen contradictoire du processus post-incident du point de vue de l'attaquant.

Tout complément d'information (articles ou publications non mentionnés, moyens non abordés etc ...) sera le bienvenue afin d'améliorer la pertinence ou la complétude des analyses dans l'objectif d'une publication ultérieure plus complète.

2 Analyse des publications

Remarque préalable : les informations de ce paragraphe sont directement issues des publications, elles ne peuvent pas être considérées comme validées par cette simple transcription (et traduction) et devraient être soumises à des compléments d'investigation.

La première publication à ma connaissance qui traite du sujet est publiée en Juillet 1999 [1] par un professeur de l'université de Georgetown (D.Denning) et un ancien du Bureau Fédéral d'Investigation (FBI – W.Baugh). Elle met en exergue l'utilisation du chiffrement, de l'anonymat et d'autres techniques pour masquer des activités criminelles : les mots de passe, la compression, la stéganographie, le stockage à distance des informations, l'inactivation ou la suppression des enregistrements d'audit .

En 2001, le Digital Forensic Research Workshop [2] met à l'ordre du jour la détection et la récupération des données cachées, son rapport signale que ce problème n'est sans rappeler celui des canaux cachés car il est tout aussi impossible à résoudre dans le but d'une éradication totale. Un tableau liste les endroits où peuvent être cachées des données : au niveau graphique (bits de poids faible, stéganographie), des signaux (séquencement, altération d'algorithme), des applications (formats de fichier, métadonnées, espace libre de

fichier), de la géométrie du disque (mauvais clusters, pistes supplémentaires, partitions cachées), des systèmes de fichiers (espace non alloués, secteur de démarrage), des structures de communications (offsets réservés de paquets, messages non sollicités, protocoles), de la mémoire (BIOS, CMOS, RAM), des structures des données (pile), du système d'exploitation (rootkit, appels systèmes), des informations non numériques (noms de fichiers). La recherche s'est concentrée sur la stéganographie et met en évidence le manque d'outils et de techniques pour la capture des données réseaux, l'analyse des fichiers et des exécutables, la cryptanalyse etc ...

En 2001 , le projet Honeynet [3] propose dans ses challenges deux cas :

- Scan 15 : il s'agit pour le premier de trouver , identifier et recouvrer un rootkit sur un système Linux, le rootkit a été simplement effacé après son installation.
- Scan 24 : proposé par le DFRWS, il s'agit d'analyser une disquette dans le cadre d'une enquête fictive. Outre la suppression de fichier , les techniques de camouflage utilisées sont : modification des secteurs, stockage du mot de passe dans l'espace inutilisé d'une image JPG, modification du nom des extensions de fichier.

Toujours en 2001, le livre de W.G.Kruse [4] comporte un chapitre de 23 pages (sur 392) sur le camouflage des données : les techniques suivantes sont examinées : le chiffrement et les mots de passe, la compression, l'utilisation de code, la stéganographie, les noms invisibles au système d'exploitation, la falsification de noms de fichiers, le stockage aux endroits inhabituels (streams NT), l'absence de nom (zero link files Unix), les espaces libres dans le système de fichier, les supports de stockage externe (amovible ou réseau), le changement de comportement ou la modification du système d'exploitation (rootkit).

En 2002, la publication de The Grugq (en anonyme à cette date) dans Phrack 59 a pour motivation de montrer que les outils forensic peuvent être trompés, qu'il n'y a pas de documentation sur les techniques anti-forensic, et d'étude des contre mesures (anti-anti-forensics !). Il définit l'antiforensic comme la tentative d'altération de la qualité et de la quantité d'information qu'un investigateur peut examiner. Les mécanismes illustrés dans l'article sont la destruction et le camouflage des données. Une implémentation nommée The Defiler's Toolkit (TDT) accompagne l'article.

A la même époque, un message de la liste de diffusion SecurityFocus Forensics [5] peut retenir notre attention car il signale deux rootkits sous Linux « dans la nature » employant des méthodes antiforensics (de type vfs).

Le Computer Emergency Response Team/Coordination Center [7] observe les activités d'intrusion depuis 1998, il signale l'augmentation de la sophistication des outils d'attaques : nature anti-forensic, comportement dynamique et modularité des outils. Les attaquants utilisent des techniques qui masquent la nature des outils d'attaques, en vue de rendre plus difficile et plus longue leur analyse et la compréhension de la menace.

Anton Chuvakin illustre brièvement dans un article [8] comment effacer, recouvrer les données sous Linux et comment les données peuvent être cachées dans le système de fichier.

Le document de Christian Johansson du Blekinge Tekniska Högskola – Suède [9], diffusé en décembre 2002, complète l'approche de destruction et de camouflage des données (s'il faut effacer les traces , il faut également effacer les preuves de l'effacement en supprimant les métadonnées), avec le ciblage de la phase d'analyse forensic par l'attaquant en recourant au chiffrement ou mieux aux modifications subtiles des données (la seule différence avant et après une attaque serait une légère augmentation de la taille de tous les fichiers Word). L'intérêt croissant pour le forensique informatique n'a pas fait croître l'intérêt pour l'anti-forensic, l'auteur suppose que cela est dû au fait qu'un effort minimum appliqué à l'anti-forensic peut ralentir conséquemment l'examen forensic. Comme chaque coté d'une pièce de monnaie, forensic et anti-forensic doivent être étudiés avec attention.

La page dédiée aux outils « anti-forensic » du site `www.networkintrusion.co.uk` [10], liste de nombreux outils libres ou commerciaux : effacement sécurisé de fichiers (srm, fwipe, grind, secureIT, cryptomite) ou de disques (wipe, overwrite, dban, diskzapper, bcwipe, stealthdisk, declasfy), Defiler's Toolkit, effacement de log ou traces (ntsecurity, evidence eliminator, tracks eraser), camouflage (hiderman, steganos security suite, cloak, invisible secrets). On peut noter que cette page a été publiée (d'après le site) après un avis du UK National High Tech Crime Unit (NHTCU).

Une présentation en allemand de sensibilisation réalisée par la société Internet Security AG [11] met en évidence dans le processus d'attaque (reconnaissance, furetage, intrusion, installation, nettoyage) le terme anti-forensic en s'appuyant sur l'article de The Grugq.

Le projet Honeynet [12] poursuit ses challenges avec le Scan 26 nouveau problème sur disquette : modification de la structure (effacement par formatage rapide), camouflage de fichier et de données, stéganographie, stockage externe (sur site Internet).

The Grugq introduit avec son article dans Phrack 62 [13] la stratégie dite de contraception consistant à exécuter un binaire sur un système sans créer de fichier sur le disque. Deux principes sont donc mis à l'épreuve : prévenir que les données n'atteignent le disque (en restant en mémoire) et utiliser des utilitaires usuels plutôt que des outils spécifiques (exemple d'un squelette de backdoor utilisant « awk »). Une implémentation d'un tel outil est proposé (remote exec).

Le guide de Sapphire Ltd [14] après avoir décrit le processus forensic , le profil de l'attaquant, aborde les outils et techniques antiforensic selon 2 catégories :

- celles qui fonctionnent : effacement et remplacement par un fichier de même nom et taille, formatage bas niveau, logiciels anti-forensic (effacement des traces de fichiers supprimés), chiffrement
- celles qui ne fonctionnent pas : effacement simple et récent, formattage haut niveau, defragmentation.

Kevin Mandia, rappelle lors de la conférence CEIC 2004 [15], que des centaines voire des milliers d'outils spécialisés permettent d'effacer les données, de chiffrer, de modifier les dates et heures, les métadonnées, de contrer le processus forensic. Même s'ils ne sont plus installés, la détection de l'installation et de l'utilisation de ce type de logiciels peut jouer un rôle important dans l'investigation.

The Grugq fait les conférences BlackHat [16], HITBSecConf 2004 puis [27].

Bram Shirani lors de conférences HTCIA [17] passe en revue les méthodes pour cacher de l'information : rootkits, chiffrement, fichiers cachés (data stream, espace non alloué, œufs de pâques, stéganographie), modules de noyau, modification de log, sniffers, utilitaires de porte dérobée.

Le challenge « *Lord of the Ring-Zero Challenge* » d'Ed Skoudis est expliqué par Raul Siles [18], il permet de mettre en évidence les techniques de camouflage de fichier et répertoire par un rootkit.

Clemens Fruhwirth [19] prend en compte les capacités de recouvrement de données sur disque dur pour améliorer les mécanismes de gestion des clés pour le chiffrement de disque.

Un article de magazine [20] met également en évidence la fragilité des implémentations logicielles de procédé de chiffrement : stockage en mémoire, pagination sur disque, fichier en clair avant chiffrement, fichiers temporaires ...

Harlan Carvey [21] présente des techniques de camouflage sous NT : altération des fichiers (nom, extension, signature), modification des attributs DOS, découpage de fichier, modification des dates et heures, conteneurs OLE, alternate data streams, chiffrement, stéganogaphie, nettoyage du registre, rootkits.

Ces points sont repris dans son livre [26] Windows Forensics and Incident Recovery au chapitre 3 (camouflage de données) et au chapitre 7 (savoir ce qu'il faut regarder).

Le sujet est pris en compte au niveau d'université comme l'illustre le poster de Matthew Geiger [22] qui étudie les outils antiforensic de protection de la vie privée.

Dans son étude du processus d'attaque, [23] l'auteur du site trapkit.de, prend en compte la phase de furtivité basée sur les techniques qui permettent à l'attaquant d'être intraçable pendant le déroulement de l'exploit et dans la phase post incident.

Samuel Dralet dans [24] nous illustre en français les travaux de The Grugq et propose deux possibilités de camouflage dans le système Linux.

Lors des conférences [25], Michael Legaris propose une méthode de détection des rootkits camouflés sous FreeBsd, et met bien en évidence que l'attaquant va d'abord chercher à camoufler son activité afin d'empêcher sa détection.

Le livre Forensic Discovery de Dan Farmer et Wietse Venema [28], auteurs des outils SATAN et Coroner's toolkit recèle de nombreux exemples de techniques anti-forensic, mais ce terme n'apparaît pas dans l'index.

La figure 1 illustre l'augmentation significative des publications traitant du sujet à partir de 2004 et le bon début de l'année 2005 (sans compter cette publication).

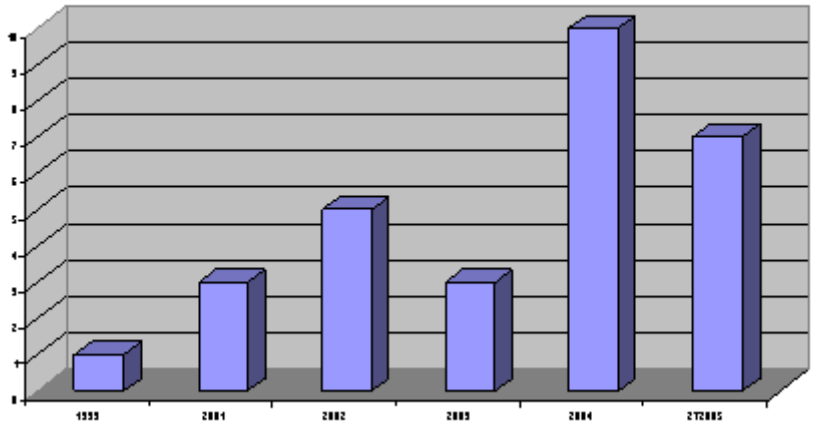

Fig. 1. – nombre de publications relatives au domaine anti-forensic depuis 1999

3 Processus post-incident

Nous allons mettre en parallèle les activités de l'attaquant et en particulier celles qualifiées d'anti-forensic avec celle de l'analyste post-incident.

Tout d'abord, il est nécessaire de **déceler une activité de l'attaquant** : n'importe quelle étape est susceptible d'être détectée :

- par manque de discrétion , l'*exploration* par l'utilisation des scans réseau trop verbeux ou un utilisateur plus suspicieux que les autres signalera la tentative d'ingénierie sociale, l'attaquant a comme recours la légitimité en étant ou en cherchant à être comme à l'intérieur de la cible visée.
- l'*espionnage* oblige l'attaquant à prendre des risques et peut être découvert à cette étape, il peut utiliser des requêtes légitimes du système pour récupérer l'information nécessaire à cette étape.
- l'*intrusion* va changer l'état du système cible, de nouvelles traces de cette activité doivent être évitées grâce au camouflage.
- l'*installation* nécessite de recourir aux techniques de contraception pour limiter l'empreinte sur le système.
- l'*action*, véritable objectif de l'attaquant, se traduit dans certains cas par une activité détectable : par exemple dans le cas d'un deni de service, le service n'est plus fourni ce qui peut être détecté, sous réserve qu'il soit sous surveillance et que la surveillance soit intègre ! On voit dans ce cas que non seulement la connaissance approfondie du système est nécessaire pour l'attaquant mais aussi les modalités de protection du système.
- le *camouflage*, partie intégrante de l'anti-forensic, peut aussi révèler la présence d'une attaque : un fichier de log vide ou avec des incohérences de

analyste post-incident	**attaquant**	Anti-forensic
Identification d'activité	Exploration Espionnage Intrusion Installation Action Camouflage	Légitimité Camouflage Contraception
Acquisition données volatiles mémoire traffic réseau données non volatiles	Exploration Espionnage Intrusion Installation	Destruction Camouflage Contraception
Préservation copie bas niveau empreinte (hash)	Installation Intrusion Action	Installation/Action Destruction Camouflage
Analyse extraction des données Identification des preuves pertinentes pour le cas Présentation	Exploration Espionnage Intrusion Installation Action Camouflage	Destruction Camouflage

date est en effet suspect, encore faut-il que quelqu'un l'examine de temps à autre ...

L'activité d'**acquisition** vise à collecter les données dûes :

- à l'exploration et à l'espionnage : il s'agit usuellement du trafic réseau
- à l'intrusion et à l'installation : il s'agit usuellement des données stockées en mémoire ou sur disque

En complément du camouflage et de la contraception, la destruction par exemple physique des supports, l'extinction des systèmes, l'utilisation de supports de stockage inhabituels (qui ne seront donc pas saisis) vont contribuer à contrer cette activité.

L'activité de **préservation** permet de s'assurer de l'intégrité et de la fidélité des données recueillies. La prise d'empreinte des différents fichiers du système permet de filtrer rapidement entre les fichiers connus pour un système d'exploitation donné et des données spécifiques ou issues d'un package d'attaque connu. Les travaux sur les collisions des empreintes [29] couramment utilisées doivent être pris en compte afin de maîtriser les conséquences d'une attaque de haut niveau visant à éliminer les fichiers suspects du ciblage réalisés (automatiquement) par les outils forensics en les faisant passer pour des fichiers légitimes du système. La connaissance technique de ces outils forensics permet également , comme pour tout produit informatique, d'exploiter dans la phase d'installation ou d'action des attaques spécifiques sur ces produits ou du système support (comme leur arrêt lors du processus d'acquisition).

Les procédures de copie sont consommatrices de ressources de stockage et de temps, et donc inutiles en présence de mesures d'effacement effective des données. Il est donc important de vérifier préalablement à cette activité que les données ne sont pas toutes nulles en calculant rapidement un CRC par exemple.

Les activités **d'analyse et d'identification des preuves** ont pour objectif d'extraire les données afin de rechercher les traces pertinentes pour le cas étudié. Les actions spécifiques de camouflage et de destruction de ces traces sont souvent révélatrices du niveau de maîtrise de l'attaquant. Il faut cependant différencier le niveau nécessaire pour la mise au point d'outils automatiques de type rookit de celui de leur utilisation (dans le but de constituer des zombies en vue d'une attaque DDOS ultérieure).

Les éléments suivants sont tirés du livre Forensic Discovery [28] pour diminuer la capacité d'analyse et d'identification :

- l'attaquant peut modifier l'intégrité du système en ne modifiant aucun fichier, en ajoutant simplement, un seul fichier utilisé lors du processus de démarrage. En effet, les analyses conduites par les auteurs montrent que les fichiers non existants au démarrage des systèmes RedHat et Solaris sont légions. Ce fichier sera très difficile à trouver puisqu'il n'est pas censé exister dans le processus normal de démarrage.
- le recours à l'obfuscation , au chiffrement , au polymorphisme des codes et tout autre astuce qui brouille la limite entre le code et les données.

L'activité **de présentation** doit reconstituer tous les éléments du puzzle afin de présenter les éléments auprès d'une autorité ou d'un tribunal. La complexité des attaques peut rendre cette activité difficile : comment expliquer à des personnes non techniques les moyens utilisés par l'attaquant, comment être certain de l'identité de l'attaquant. La présence d'activités spécifiquement anti-forensic peut-elle être un élément à charge ?

4 Conclusion

La connaissance approfondie des techniques anti-forensiques permet de mettre en place des contre-mesures spécifiques au niveau technique et organisationnel visant à déclencher des actions de maîtrise de la sécurité du système permettant ainsi de déceler au plus tôt des activités nuisibles au processus de réaction après incident.

Cependant, une recherche plus approfondie est nécessaire pour déterminer s'il s'agit comme pour l'activité anti-virale [30] d'un problème indécidable.

Références

1. D. Denning, W. Baugh - *Hiding Crimes in Cyberspace.*
2. Report From the First Digital Forensic Research Workshop - Workshop 3 - Detection and Recovery of Hidden Data.
3. Honeyney project - Scan of the month 15, 24.

4. W.G Kruse II - Computer Forensics - Incident Response Essentials - chapter 5 - Data Hiding.
5. Lawless, Tim - SecurityFocus Forensics - RE : An alternative method to check LKM backdoor/rootkit - Fri Apr 19 2002 - 09 :12 :49 CDT.
6. The Grugq -The art of defiling : Defeating Forensic Analysis on Unix.
7. Allen Householder, Kevin Houle, and Chad Dougherty, CERT Coordination Center - Computer Attack Trends Challenge Internet Security.
8. Anton Chuvakin - Linux Data Hiding and Recovery – linuxsecurity.com – 3/10/2002.
9. Christian Johansson - Forensic and Anti-Forensic Computing.
10. Talisker anti forensic Tools - `http://www.networkintrusion.co.uk/foranti.htm` - mise à jour du 28 mai 2003.
11. Internet Security AG - Internet Security Worum geht's eigentlich ?
12. Honeyney project - Scan of the month 26.
13. The Grugq - Remote Exec.
14. Sapphire Ltd - Computer Forensics – A Short Practitioners Guide - 5. Anti-Forensic Tools and Techniques.
15. Kevin Mandia - A Look at Some Challenges Facing Incident Response and Computer Forensics.
16. The Grugq - The art of defiling : Defeating Forensic Analysis on Unix File Systems.
17. Bram Shirani - Anti Forensics - HTCIA Spring Training June 2004.
18. Raul Siles - Lord of the Ring-Zero Challenge (March 2004) - Ed Skoudis.
19. - Clemens Fruhwirth - An anti-forensic, two level, and iterated key setup scheme - July 15, 2004.
20. Frater Ignotius - Anti-forensics : make secrets stay secrets.
21. Harlan Carvey - BlackHat Windows Security 2004 - Data Hiding on a Live System.
22. Matthew Geiger – a Forensic Evaluation of Anti-Forensic Privacy Tools - Privacy Policy, Law, and Technology class - poster session.
23. trapkit.de - IT-Defense security conference 2005 – advanced exploiting.
24. Samuel Dralet - Anti-forensics sur systèmes de fichiers ext2/ext3.
25. Michael Legary - Seccuris Labs - Digital Anti-Forensics : emerging trends in data transformation techniques.
26. Harlan Carvey - Windows Forensics and Incident Recovery - chapter 3 - Data Hiding, chapter 7 - Knowing what to look for.
27. The Grugq - The art of defiling : Defeating Forensic Analysis – BCSASIA.
28. Dan Farmer, Wietse Venema Forensic Discovery.
29. Dan Kaminsky - MD5 to be considered harmful someday.
30. Fred Cohen - Computer Viruses - Theory and Experiments.