informations quand on utilise la fonction « copier » ou « couper »)…

Pour dupliquer la mémoire vive, on se sert d'un programme spécifique, nommé imageur de mémoire. Tous les types d'appareils électroniques (ordinateurs, téléphones, équipements du réseau tels les routeurs, etc.) ont une RAM, mais ils utilisent des logiciels différents, si bien que les programmes conçus pour analyser certains systèmes ne fonctionnent pas toujours sur les autres. Aujourd'hui, aucune méthode universelle n'existe.

Dans un nombre croissant d'enquêtes, les données à analyser ne sont pas celles enregistrées sur une machine individuelle, mais celles échangées *via* une connexion réseau. Cette connexion est alors mise sur écoute.

La sauvegarde des données n'est que la première étape. On doit ensuite identifier les informations pertinentes pour l'enquête. Certains outils peuvent extraire les fichiers de l'image disque, rechercher ceux qui contiennent un mot ou une phrase particuliers (dans des langues variées), détecter des données cryptées...

Sur un ordinateur, les fichiers dits alloués peuvent être visualisés grâce au système de gestion des fichiers (par exemple dans des répertoires arborescents). Le terme alloué se réfère aux secteurs du disque où le contenu du fichier est stocké : ils sont dédiés à ce fichier particulier et ne peuvent pas être affectés à d'autres. De nombreux outils d'informatique judiciaire permettent de voir les fichiers alloués d'une image disque sans utiliser le système d'exploitation de l'ordinateur, ce qui préserve l'intégrité judiciaire des preuves.

L'AUTEUR

Simson GARFINKEL est maître de conférences à la *Naval Postgraduate School*, à Monterey, en Californie.

Article publié avec l'aimable autorisation de la revue *American Scientist*.

À la recherche des fichiers perdus

Quand on supprime un fichier, c'est-à-dire quand on le place dans la poubelle de l'ordinateur et qu'on vide cette dernière, l'espace qu'il occupait est dit désalloué. Son contenu peut rester en mémoire, mais le fichier n'est plus visible et les métadonnées permettant de le localiser sont perdues. Un des principaux progrès de l'informatique judiciaire au cours des 15 dernières années concerne la récupération de ces fichiers « supprimés ». On utilise une technique nommée *file carving* (littéralement « sculpture de fichier »), inventée vers 1999 par Dan Farmer, un chercheur indépendant en sécurité.

De nombreux types de fichiers commencent et se terminent par des séquences caractéristiques d'octets, respectivement nommées en-tête et pied de fichiers. Ainsi, les fichiers JPEG créés par les appareils photo numériques commencent en général par la séquence FF D8 FF E0 et se terminent par FF D9. Les premiers dispositifs de *carving* passaient en revue l'image disque à la recherche de ces séquences

Le hachage, technique clef de la cybersécurité et de l'informatique judiciaire

Le hachage a été inventé au début des années 1950 par l'informaticien allemand Hans Peter Luhn, chez IBM. Une fonction de hachage de x bits associe un nombre binaire de taille x, c'est-à-dire comportant x bits, à une chaîne de caractères, tel un extrait de texte. Le terme hachage vient de ce que, en pratique, les fonctions utilisées segmentent les données avant de les mélanger. Une fonction de hachage de 32 bits peut ainsi produire $2^{32} = 4\,294\,967\,296$ valeurs. Ces fonctions sont conçues de façon à obtenir un résultat complètement différent quand on change un seul caractère de la chaîne d'entrée. Certaines chaînes ont la même valeur de hachage (on parle de collision de hachage), mais plus celle-ci est longue, moins cela risque de se produire.

En 1979, Ralph Merkle, alors doctorant à l'Université de Stanford, aux États-Unis, a appliqué le hachage au domaine de la sécurité informatique. Il a utilisé une fonction de hachage produisant plus de 100 bits en sortie et dite à sens unique : il est facile de calculer la valeur de hachage d'une chaîne, mais presque impossible, à partir d'une valeur de hachage, de trouver la chaîne correspondante. L'idée de R. Merkle était d'utiliser la valeur de hachage d'un document comme signature numérique. Le nombre considérable de valeurs de hachage possibles (2^{100}, soit environ 10^{30}) devait empêcher les pirates de trouver celle associée à un document cible. Aujourd'hui, les signatures électroniques obtenues par hachage sont à la base de nombreux systèmes de cybersécurité, protégeant par exemple les numéros de cartes bancaires envoyés par Internet.

Dans le domaine de l'informatique judiciaire, le hachage sert notamment à vérifier que les copies des données sont fidèles à l'original. Pour un même système informatique, on réalise souvent deux images disques, auxquelles on applique une fonction de hachage. Si les images ont la même valeur de hachage, on suppose qu'elles sont toutes deux une copie fidèle des données d'origine. Tout enquêteur en possession d'une copie ultérieure des données peut calculer la valeur de hachage et vérifier qu'elle correspond à la valeur initiale. De nombreux outils d'informatique judiciaire font automatiquement cette comparaison.

Le hachage est aussi utilisé pour étiqueter des fichiers ou des segments de fichiers connus comme étant « bons », tels les programmes fournis avec les systèmes d'exploitation, ou « mauvais », tels des virus informatiques, des documents volés ou de la pédopornographie. Aujourd'hui, les experts judiciaires distribuent les valeurs de hachage de ces fichiers, afin de faciliter leur détection. Grâce à l'identification de fichiers complets ou partiels par hachage, on peut également déterminer automatiquement si l'un des millions de fichiers que contient un disque dur apparaît dans une base de données renfermant des centaines de millions de fichiers, et ce en un temps assez court, de l'ordre de quelques heures.