mentée sans erreurs, devient-elle parfaitement sûre ? Pas encore. Il reste les attaques indirectes, qui sont comparables à la découverte de la combinaison numérique protégeant un coffre-fort grâce à l'écoute au stéthoscope des bruits de son mécanisme de verrouillage… C'est ce que l'on nomme les attaques par canaux auxiliaires ou canaux cachés.

Ce type d'attaque n'est malheureusement pas l'apanage des films d'espionnage : même un cryptosystème parfaitement réalisé peut être sujet à des fuites d'informations parasites qui, bien analysées et exploitées, suffisent parfois pour le casser. Il arrive ainsi qu'un processeur émette des rayonnemen ts exploitables, qu'il consomme de l'énergie suivant un profil traduisant les tâches de transmissions qu'il est en train d'accomplir, qu'un algorithme de chiffrement ait des temps d'exécution plus ou moins longs selon la valeur de la clef, etc.

De telles fuites ont déjà été utilisées pour casser certaines implémentations du cryptage RSA. Comment ? Les opérations d'élévation à la puissance nécessaires dans cet algorithme peuvent être réalisées soit par multiplication, soit par élévation au carré ; or ces deux opérations induisent des profils de courant électrique différents, lesquels, mesurés et rapportés à leur cause, ont *in fine* permis de déterminer la clef !

Attention aux fuites !

Les failles de ce type sont les plus difficiles à déceler. Les modèles de cryptosystèmes, qui ignorent toute la couche technique sous-jacente au fonctionnement des ordinateurs et des réseaux, sont inutiles pour les repérer. Une fois qu'elles sont connues cependant, il arrive que des contre-mesures simples suffisent à les faire disparaître, par exemple un brouillage supplémentaire de données, une égalisation des signaux ou de la consommation électrique, le blindage des puces pour les rendre opaques à l'espionnage électromagnétique, etc. Autant de défenses qui tiendront jusqu'à ce qu'un autre attaquant découvre encore une fuite insoupçonnée d'information…

La conception et la mise en œuvre d'un cryptosystème impliquent ainsi une science subtile. Cependant, toute la subtilité du monde ne peut rien contre l'utilisateur d'un coffre-fort qui laisse traîner les clefs ou contre la maladresse d'un utilisateur de communications chiffrées qui emploie le même mot de passe trivial (son prénom) pour se connecter au site de son groupe de fans de football, puis à sa banque et ensuite à sa messagerie d'entreprise…

De plus, les révélations du *Guardian* et du *New-York Times* suggèrent que le secret le mieux gardé de l'Internet aujourd'hui serait que les agences de renseignement anglo-saxonnes ont secrètement imposé aux « industries de l'Internet américaines et étrangères » d'insérer des « vulnérabilités dans leurs systèmes de cryptage ». Ces portes dérobées leur assureraient des entrées dans les protocoles de communication sécurisée largement utilisés en ligne, tels HTTPS (le protocole de transfert hypertexte sécurisé), la voix sur IP (téléphonie par Internet) ou encore SSL. Un coffre-fort doté d'une trappe indécelable, mais connue des espions, protégera bien mal nos secrets, c'est clair... ■

La quête d'un support numérique durable

Les disques durs, les CD, les DVD, les mémoires flash... se dégradent en quelques années. Cette situation est largement sous-estimée, mais diverses stratégies et de nouveaux dispositifs sont mis en place.

Franck Laloë et Erich Spitz

Les greniers des vieilles maisons abritent souvent des trésors. Quand on fouille dans des tiroirs ou dans des coffres, il n'est pas rare de retrouver de vieilles photos, des films huit millimètres ou des lettres manuscrites. Aujourd'hui, la plupart des documents, aussi bien les images que les vidéos et les textes, sont enregistrés sur des supports numériques, tels les disques durs des ordinateurs, les CD et les DVD.

À quoi ressembleront les chasses au trésor dans les greniers d'ici 50 ans ? Comment les supports numériques résistent-ils au passage du temps ? Les constructeurs ont annoncé que les supports numériques avaient une longue durée de vie. Les rares études réalisées montrent, au contraire, que ces supports se détériorent en quelques années. Dès lors, se pose la question de la pérennité des données numériques. Comment les conserver ? Ce problème est méconnu du grand public. Pourtant, il touche les documents personnels ayant une valeur sentimentale (les photos et les lettres), mais aussi les informations plus sensibles, telles les données administratives, financières, médicales ou scientifiques.

Aujourd'hui, le numérique est le véhicule dominant de l'information. Il a supplanté les supports analogiques parce qu'il présente de nombreux avantages : une grande densité d'information permettant un gain de place considérable, la facilité des recherches dans l'information stockée, un accès rapide, une circulation quasi instantanée et son universalité – la même clef USB peut contenir des textes, du son, des vidéos, des images et des tableaux de données. La question de la pérennité des données a été reléguée au second plan. Pour l'industrie, la priorité était de développer des dispositifs rapides et contenant des quantités d'informations toujours plus importantes.

La question de la pérennité reste essentielle. Nous aborderons les différents aspects

L'ESSENTIEL

- La pérennité des données numériques est un problème peu étudié par l'industrie et les organismes publics.
- L'altération des supports est la difficulté principale pour l'archivage. Les dispositifs les plus courants ont des durées de vie courtes.
- La méthode d'archivage durable la moins onéreuse aujourd'hui consiste à graver des DVD en verre.

© R. MACKAY PHOTOGRAPHY, LLC/shutterstock.com

de la pérennité de l'information numérique et soulignerons que le problème le plus délicat est celui de l'altération des dispositifs de stockage. Des solutions existent, mais elles sont encore peu répandues.

Notons d'abord que sauvegarde et archivage à long terme des données ne sont pas synonymes. La première a pour but de prévenir les pertes soudaines de données (accidents divers, vols, etc.), la seconde, la disparition progressive des données utiles à long terme. En ce qui concerne la sauvegarde, comment garantir que la duplication d'un fichier est fidèle à l'original ? Cette difficulté trouve une solution efficace dans les codes correcteurs, des algorithmes qui identifient les erreurs de copie.

Dupliquer sans erreurs

Cet aspect des méthodes numériques passe souvent inaperçu des utilisateurs : la possibilité de copier un document numérique avec la garantie que la copie ne contient pas d'erreurs. En effet, quand vous enregistrez un texte, aucune altération accidentelle ne viendra s'y glisser et changer une lettre en une autre pendant la sauvegarde. Cette fiabilité ne se retrouve pas dans les supports analogiques. Il suffit de penser aux textes anciens qui nous sont parvenus par le travail séculaire des moines copistes qui, malgré leur soin extrême, ont introduit des erreurs ; il est humainement impossible de faire des copies identiques. C'est d'ailleurs valable pour toute technique analogique : quelle que soit sa qualité, chaque recopie introduit un peu de « bruit » qui s'accumule de façon exponentielle de copie en copie. Au bout d'un certain nombre de sauvegardes, l'information initiale devient inexploitable.

Derrière cette merveille de la copie sans erreurs se cache le travail de mathématiciens, tels les Américains Irving Reed et Gustave Solomon, qui, dans les années 1960, ont mis au point des codes de correction d'erreurs numériques permettant des millions de copies sans erreurs d'un fichier. L'algorithme est complexe, mais son principe est simple : il faut « suréchantillonner » l'information, c'est-à-dire donner plus d'éléments qu'il n'est nécessaire pour restituer l'information. La redondance permet de vérifier qu'il n'y a pas eu d'erreurs dans la copie. Les algorithmes sont si efficaces que si nous recopions un fichier légèrement endommagé, sur un CD de données un peu ancien par exemple, la copie éliminera les erreurs apparues entre-temps – tant que le taux d'erreurs apparues n'est pas trop grand. Le fichier sera restauré dans son état initial !

Le numérique semble avoir les mêmes vertus magiques que le Phénix renaissant de ses cendres. Et pourtant, il a un talon d'Achille, et pas des moindres : la préservation à long terme des données. Certes, il est de peu d'intérêt de conserver la plupart d'entre elles, dont la pertinence est éphémère ; mais il est essentiel d'en préserver la petite fraction qui garde un intérêt pendant des années, des décennies, parfois des siècles. L'intérêt de certaines données croît avec leur ancienneté, comme les données des satellites d'observation de la surface terrestre : en période de réchauffement climatique, il devient important de comparer les observations actuelles aux plus anciennes possibles, afin de discerner des évolutions à long terme. Les données des grands instruments scientifiques (telles celles du CERN, du satellite *Planck*, du radiotélescope SKA ou *Square Kilometer Array*) sont irremplaçables, rien ne garantissant qu'il sera possible un jour de répéter ces expériences très coûteuses. À plus petite échelle, les données du cadastre, de la comptabilité des entreprises, les données médicales, etc., doivent être conservées pendant longtemps. Enfin, à titre individuel, chacun d'entre nous souhaite conserver les souvenirs de sa vie personnelle et familiale.

La conservation de données numériques n'est donc pas une question anodine, puisqu'elle touche de nombreux domaines. Elle est confrontée à trois obstacles : documenter les fichiers pour savoir de quoi il s'agit, conserver le matériel et les logiciels nécessaires à la lecture du fichier et lutter contre le vieillissement des supports physiques.

Les deux premiers points sont ceux auxquels chacun songe avant tout. Toutes les formations de documentaliste enseignent l'importance des métadonnées. Celles-ci regroupent les informations qui accompagnent un document : sa nature, sa date, ses auteurs, des mots clefs associés, etc. Ces éléments facilitent une exploitation future.

Outre le problème des métadonnées, chacun a une anecdote à raconter de disquettes contenant des fichiers importants, mais illisibles faute de matériel adéquat.