avec un paramètre commun (le nombre aléatoire tiré au début).

Les mécanismes distribués constituent un pas majeur vers la sécurisation du processus d'assainissement, mais leur mise en œuvre pose des problèmes de passage à l'échelle, car ils nécessitent des calculs importants et une forte connectivité des participants. Ils sont pour l'instant relégués au traitement de jeux de données peu volumineux.

Avec notre équipe, nous nous sommes attaqués à ce problème. Nous avons développé une architecture de gestion de données personnelles fondée sur des dispositifs individuels sécurisés (telles des clefs USB dotées de processeurs sécurisés, insérables dans un smartphone, une tablette ou un ordinateur). Nous avons montré que les calculs nécessaires aux algorithmes d'assainissement peuvent être répartis entre ces dispositifs et une entité centrale sans que celle-ci ne dispose des données personnelles non chiffrées. La puissance de calcul de l'entité centrale assure l'applicabilité à grande échelle et, grâce au fait qu'elle coordonne les dispositifs personnels, ceux-ci n'ont pas besoin d'être interconnectés en permanence. Cette architecture est capable d'appliquer des algorithmes d'assainissement à des jeux de données massifs, correspondant à plusieurs millions d'individus.

Une telle architecture décentralisée pourrait assurer la gestion des dossiers médicaux, des factures ou de tout autre type de dossier personnel. En pratique, les données seraient stockées localement sur les appareils de l'utilisateur et ne seraient jamais exportées sous une forme non perturbée ou non chiffrée. Aucun serveur ne les regrouperait toutes, mais les échanges entre l'entité centrale et les appareils des utilisateurs permettraient tout de même de collecter certaines informations, d'une façon respectueuse de la vie privée.

Pendant la dernière décennie, une grande diversité de modèles et d'algorithmes d'assainissement de données ont été élaborés, afin de mieux prendre en compte les capacités de l'attaquant. Aujourd'hui, la pseudonymisation reste massivement utilisée, alors qu'elle ne garantit pas une protection suffisante ; dans le reste des cas, la *k*-anonymisation est le plus souvent appliquée, et les techniques plus complexes sont encore exotiques. Ces techniques doivent donc continuer à se répandre et à se perfectionner.

Pour autant, la science de l'assainissement reste la recherche du meilleur compromis entre utilité et confidentialité des données, dont la protection absolue est illusoire. Seule la destruction des données peut garantir qu'un attaquant n'en tirera aucune information sensible (cette destruction est d'ailleurs compliquée, en raison de leur présence sur de nombreux serveurs). L'analyse des nouveaux gisements de données personnelles pouvant apporter des bénéfices notables, l'assainissement de données est une question sociétale avant d'être un défi scientifique. Il faudra y apporter des réponses multidisciplinaires, impliquant informaticiens, juristes et analystes. ■

La protection des informations

Rémy Chrétien et Stéphanie Delaune

Un bon cryptage des données sensibles ne suffit pas à les protéger. Il faut aussi que les protocoles de communication utilisés pour coder ces données soient dépourvus de failles, dans leur logique comme dans leur mise en œuvre.

Le 6 septembre 2013, le monde apprenait avec stupéfaction l'existence de *Bullrun*. Ce programme de l'Agence de sécurité américaine, la NSA *(National Security Agency)*, vise à décoder un grand nombre d'informations cryptées circulant sur Internet. À partir des documents secrets que leur a livrés le lanceur d'alerte Edward Snowden, le *New York Times* et le *Guardian* révélaient qu'une percée technique avait multiplié la puissance de *Bullrun* en 2010, de sorte que les transactions et communications en ligne protégées des banques et de la plupart des grands fournisseurs de service sur Internet étaient désormais accessibles « en grands volumes » aux agences de renseignement *(voir la figure 1)* ; on apprenait aussi que ces agences disposent de superordinateurs pour décrypter par la « force brute » les données qui résistent…

sensibles

© Domaine public

Aujourd'hui, sans vraiment le savoir, nous réalisons tous quotidiennement des opérations de cryptage. Des chiffrements permettent, par exemple, de protéger notre numéro de carte bancaire lors d'un achat en ligne ; on les retrouve dans les protocoles (les procédures de communication) utilisés en téléphonie mobile ou pour lire les passeports électroniques. Pour autant, toutes ces données sensibles, nos données, sont-elles vraiment protégées ? Où sont les points faibles ? Nous allons analyser dans cet article ce qui conditionne la qualité des systèmes de sécurisation des données utilisés de façon routinière.

Sur un réseau, tout se passe comme si les informations protégées ne circulaient qu'à l'intérieur de coffres-forts réalisés dans un matériau solide : le chiffre (le cryptage). Toutefois, quelle que soit la solidité mathématique du chiffrement opéré, encore faut-il que l'assemblage des pièces qui constituent les coffres-forts en circulation ne laisse pas apparaître de fente exploitable par une personne mal intentionnée… Finalement, l'expédition de coffres-forts d'un côté et de leurs clefs de l'autre n'est pas une pratique sans risques, et il arrive qu'on y commette des erreurs, voire qu'on laisse traîner les clefs…

Chiffre et compagnie

Le chiffrement, nous l'avons dit, est l'analogue du matériau constituant les parois du coffre-fort. Il en existe de nombreux types, chacun caractérisé par des contraintes d'utilisation et des niveaux de sécurité différents. Quoi qu'il en soit, la qualité d'une primitive cryptographique, c'est-à-dire d'un algorithme de cryptage particulier, se mesure à sa robustesse et à sa facilité de mise en œuvre. Tout comme les parois d'un coffre-fort peuvent être plus ou moins épaisses, la robustesse d'une primitive cryptographique peut être plus ou moins importante. Une primitive sera considérée comme faible quand il est facile de découvrir comment inverser son chiffrement, inversion qui permet de révéler les données protégées *(voir l'encadré page suivante)*.

1. CES DÔMES RECOUVRENT LES ANTENNES utilisées (ici au Japon) par le système mondial d'interception des communications privées et publiques des agences de renseignement des États-Unis, du Royaume-Uni, du Canada, de l'Australie et de la Nouvelle-Zélande.