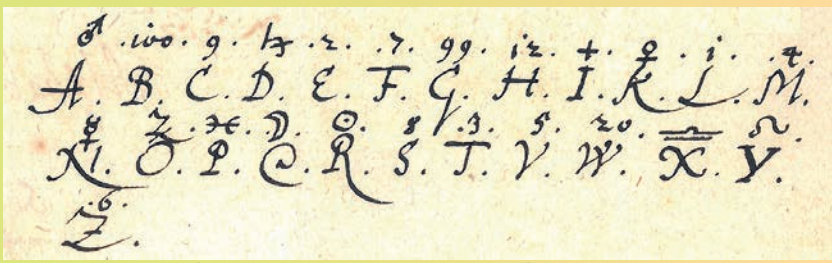
Langages chiffrés : ce qu'en disent les mathématiques

Hervé Lehning

Agrégé de mathématiques, écrivain scientifique, membre de l'ARCSI et commandant de réserve

Dans certains domaines, comme la diplomatie ou l'armée mais aussi les relations commerciales, mieux vaut garder ses communications secrètes. Le décryptement d'un seul message peut décider du sort d'une bataille ou d'une négociation. Ce fut le cas en 1626 quand les troupes du prince de Condé assiégeant Réalmont interceptèrent un individu sortant de la ville, porteur d'un message incompréhensible. Condé fit venir un jeune professeur de mathématiques de la région, Antoine Rossignol des Roches, qui en trouva le sens. Le message annonçait que la ville était à cours de munition. Condé fit porter le message décrypté à la ville, qui se rendit.



Alphabet chiffré contemporain de la prise de Réalmont. Chaque lettre est remplacée par un chiffre différent (I et J ainsi que U et V sont confondues).

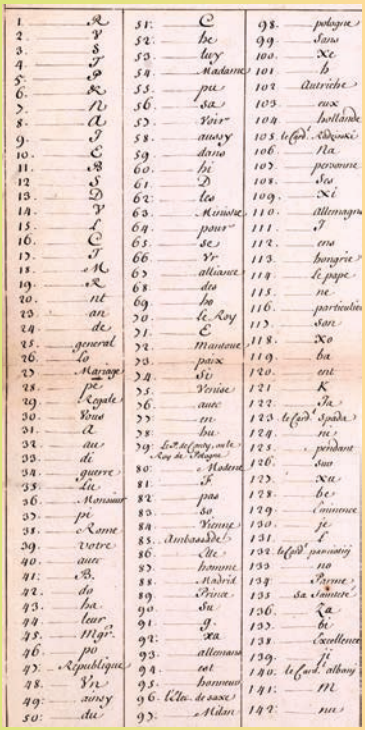
© Archives de Strasbourg

Des méthodes encore trop sensibles à l'espionnage

Le décryptement des messages chiffrés ainsi repose à la fois sur les mathématiques et sur la linguistique. Côté mathématique, on utilise la *méthode des fréquences*. En français, le *e* est la lettre la plus fréquente ; ensuite viennent les lettres *sainturlo*, ce qui est facile à retenir en pensant que Urlo est le saint patron des cryptologues. Côté linguistique, il s'agit

de la *méthode du mot probable*. Par exemple, dans un message sortant d'une place forte assiégée, il est logique de penser à la présence d'un mot comme *munition*. Il peut alors être cherché à travers les répétitions qu'il contient, soit ***xy***y***x* où *x* et *y* sont deux symboles distincts et les étoiles, d'autres symboles.

La faiblesse des alphabets chiffrés, même améliorés en chiffrant de plusieurs façons différentes les lettres fréquentes et en ajoutant des *nulles* (des symboles ne signifiant rien), amena Rossignol à créer des dictionnaires chiffrés, c'est-à-dire des dictionnaires bilingues dont l'une des langues est le français et la seconde, des nombres. Ainsi, on chiffre non seulement des lettres (et ce de plusieurs manières), comme auparavant, mais aussi des syllabes et des mots. La méthode des fréquences n'a alors plus aucun sens et celle du mot probable devient difficile à utiliser.



Extrait de la partie déchiffrante d'un dictionnaire chiffré du XVII^{e} siècle.

© Archives de Strasbourg

Leur inconvénient principal est leur sensibilité à l'espionnage ou aux hasards de la guerre. Pour cette raison, ce type de chiffrement n'était plus utilisé dans les armées de terre mais seulement dans la diplomatie et la marine dès la Première Guerre mondiale. Le chiffrement évolua donc vers des méthodes modifiables régulièrement grâce à des clefs selon le principe énoncé par Auguste Kerckhoffs en 1883 et résumé en 1949 par Claude Shannon : *« L'ennemi connaît le système. »* Dès la Renaissance, un diplomate, Blaise de Vigenère, avait conçu un système répondant à ce principe. Il consiste en une substitution alphabétique variant selon une clef. Par exemple, la clef ABC transforme « secret » en « sferfv », la première lettre étant conservée (A), la deuxième décalée d'un cran (B), la troisième de deux (C), *etc*. Une autre méthode consiste à permuter les lettres d'un message. Pour ce faire, on peut disposer le texte à chiffrer sur plusieurs colonnes, en ajoutant quelques lettres pour obtenir un rectangle :

ledecryptemen
tdunmessagech
iffrepartrans
positionestpl
usdelicatquec
eluidunmessag
echiffreparsu
bstitutionxxx

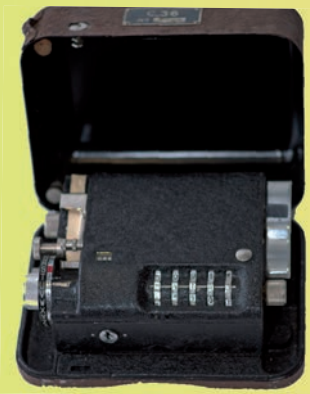
On modifie alors l'ordre des colonnes puis on écrit le texte obtenu par lignes. La clef de ce chiffrement est donc le nouvel ordre des colonnes. Par exemple, si l'on écrit d'abord les colonnes paires puis les impaires, on obtient « eerpeeldcytmndne […] tttoxx ». Le décryptement sans connaître la clef est délicat, mais loin d'être impossible : les mathématiques permettent de trouver le rectangle utilisé ; un raisonnement linguistique est ensuite nécessaire pour déterminer la transposition.

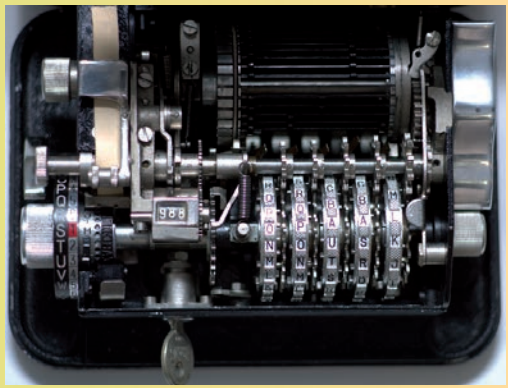
La Seconde Guerre mondiale écourtée… grâce aux maths

Les chiffres de l'armée allemande de la Première Guerre mondiale alliaient ainsi transpositions et substitutions. Ils furent cependant régulièrement décryptés par l'armée française. Ainsi, elle sut souvent à temps où allaient se situer les attaques allemandes, ce qui permettait de disposer les réserves aux bons endroits. Après guerre, le chiffre est progressivement dominé par les machines, dont la célèbre *Enigma*, qui, comme le chiffre de Vigenère, fabriquent des substitutions poly-alphabétiques. Le premier décryptement d'Enigma fut un succès conjoint de l'espionnage français et du génie de trois mathématiciens polonais, Marian Rejewski, Jerzy Rozycki et Henryck Zygalski. L'espionnage fournit les tables de chiffrement de l'armée allemande de 1931 à 1938. Les mathématiques permirent, grâce à ce renseignement, de reconstituer les câblages de la version militaire de l'Enigma et d'en fabriquer des répliques. Les messages furent alors régulièrement décryptés. Les mathématiciens polonais cherchèrent à pouvoir se passer des tables de chiffrement, ce qu'ils réussirent à faire, en créant une machine, la *bomba*. Elle permettait de trouver la clef du jour en quelques minutes. Après la défaite de la Pologne puis celle de la France, les résultats polonais furent transmis aux Britanniques. Alan Turing et son équipe de Bletchley Park reprirent avec succès le décryptement en l'améliorant et en l'adaptant aux complexifications successives d'Enigma.

La guerre en fut probablement écourtée de deux ans !

Au cours des années 1930, l'armée française s'équipa de machines comparables, une machine mécanique légère, la C-36, destinée aux communications des petites unités tactiques, et une machine électromécanique lourde, la B-211, destinée aux communications au niveau des grosses unités. La C-36 fut perfectionnée pour devenir la M-209 de l'armée américaine. Cette machine était assez solide pour son niveau d'utilisation. Cependant, dès 1943, les Allemands étaient capables de décrypter ses messages en quelques heures.





À gauche : la C-36 fermée. Les cinq lettres visibles sur le capot (MNTRJ) constituent une clef modifiable à chaque message. La machine est actionnée par le levier de droite. Les lettres à chiffrer sont choisies sur la roue de gauche, le chiffrement sort sur la bande de papier à gauche.

À droite : la C-36 ouverte. On voit les ergots des rotors, chacun étant poussé vers la droite ou la gauche, ce qui constitue la clef principale avec la position de l'élément rouge sur la roue de gauche (ici devant le S).

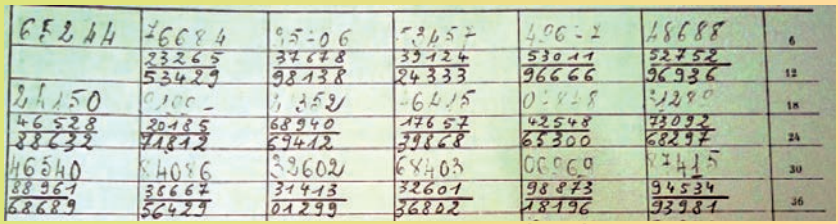
Photos : H. Lehning

Pendant ce temps, la méthode de Vigenère resta en usage, en particulier dans l'espionnage. Son décryptement est d'autant plus difficile que la clef est longue et choisie aléatoirement. L'idéal est qu'elle soit aussi longue que le message et qu'on ne l'utilise qu'une fois ; on parle alors de *masque jetable*. Claude Shannon a prouvé que, sous ces hypothèses, un message ainsi chiffré est indécryptable.

C'est le chiffre utilisé dans le fameux téléphone rouge, qui relie Washington et Moscou depuis 1963, et dans des matériels comme le Tarec en service dans l'armée française du milieu des années 1950 à la fin des années 1970. Le masque jetable est très probablement utilisé aussi dans les *stations de nombres*, des radios qui n'émettent que des suites de nombres, très courantes pendant la guerre froide et qui n'ont pas disparu depuis.

Le principe est simple. On chiffre d'abord le message en une suite de nombres (par exemple, A = 01, B = 02…) puis on ajoute au message obtenu un nombre aléatoire de même longueur sans effectuer de retenue. Ainsi, « secret » se traduit d'abord en « 19 05 03 18 05 20 ». Si l'on dispose de la suite de nombre aléatoires 90689 91275 03682, on obtient par ce procédé le chiffrement « 09639 22070 23 ».

Pour que le destinataire puisse déchiffrer le message, il faut qu'il dispose de la même suite de nombres aléatoires et qu'il fasse des soustractions. Cette méthode, qui fut utilisé par Che Guevara et des espions soviétiques comme Georges Beaufils, est toujours opérationnelle ; sa seule faiblesse reste l'échange des clefs.



Quand des soustractions servent de pièces à conviction.
Extrait du déchiffrement d'un message par Georges Beaufils, espion de l'URSS.

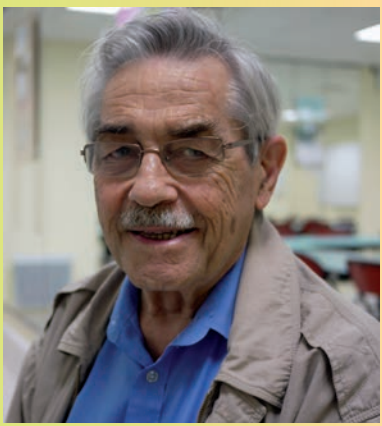
© Ministère de la Défense

Clef publique : la révolution des chiffres asymétriques

La connaissance de la clef de chiffrement semble suffire pour déchiffrer. Cependant, il existe des chiffres où savoir chiffrer n'implique pas que l'on sache déchiffrer ! Ces chiffres sont dits *asymétriques* (par opposition, les autres sont *symétriques*). La clef de chiffrement est dite *publique*, celle de déchiffrement, *privée* ou *secrète*. Le premier chiffre asymétrique est dû à trois mathématiciens, Ronald Rivest, Adi Shamir et Leonard Adleman, qui le proposèrent en 1977. Cette méthode, nommée RSA, est fondée sur la difficulté d'un problème mathématique : la factorisation en nombres premiers. Ainsi, la clef publique est un nombre, produit de deux grands nombres premiers, et la clef secrète ces deux nombres. Tout un chacun peut chiffrer, mais seul le détenteur de la clef secrète peut déchiffrer ! Ce système est assez lourd en calcul (il correspond à des élévations de grands nombres à une puissance élevée, alors que les systèmes symétriques correspondent à des additions). C'est pourquoi il est limité à l'authentification et à l'échange de clefs symétriques. Il est utilisé dans les cartes bancaires, les messages sur Internet, et dans le téléphone sécurisé Teorem de Thales, réservé aux hautes autorités de l'État.

La furtivité pour échapper aux surveillances de masse

Une autre méthode ancienne pour protéger ses communications est la *stéganographie*, où l'on cache l'existence même du message et non pas son seul contenu. On a pu utiliser les encres sympathiques, ou les microfilms. Aujourd'hui, on utilise les mathématiques… Par exemple, une image est une suite de *bits*. Si la résolution est bonne, changer quelques *bits* concernant la couleur ne modifiera pas énormément l'image. Une méthode possible consiste à convenir que les *bits* dont le numéro d'ordre est multiple de 50 contiennent le message caché. Ceci modifie légèrement l'image, mais personne ne semble actuellement capable de détecter une telle manipulation. Les mêmes principes sont applicables à la musique. De faibles variations, imperceptibles pour l'oreille, dans les basses fréquences, peuvent contenir beaucoup d'informations. Un grésillement infime peut cacher des secrets ! Il serait étonnant que ces méthodes difficilement détectables ne soient pas utilisées par toutes sortes de réseaux malveillants, mafias, pédophiles ou terroristes. On peut ainsi imaginer qu'un site de vente aux enchères ou votre site de partage de vidéos préféré serve à déposer des messages de toutes sortes. La faiblesse de ces méthodes reste l'importance du secret ; elles sont très sensibles à l'espionnage.



Louis Pouzin, l'un des inspirateurs d'Internet, président d'OpenRoot, une solution furtive pour communiquer sur Internet à l'abri des surveillances de masse.

Pour finir, une autre méthode pour protéger ses secrets est la furtivité, de ne pas apparaître sur les écrans radars de la surveillance. C'est ce que propose la société OpenRoot, créée par Louis Pouzin, l'un des inventeurs d'Internet. Avec un site sur OpenRoot, vous ne figurez pas sur les serveurs de DNS (Domain Name System) de Google, qui sont des dictionnaires traduisant les noms de domaines (comme cijm.org par exemple) en adresses IP, et échappez ainsi aux surveillances de masse !

H. L.