

Applications de l'arithmétique en cryptographie*

Robert BROUZET

04/12/2004

1 La cryptographie en deux mots

La cryptographie a pour but d'assurer la confidentialité d'un échange d'informations entre un émetteur A, généralement appelé Alice et un récepteur B, généralement appelé Bob lorsque ces informations transitent via un canal non sûr (poste, téléphone, email etc.) Le principe général consiste à brouiller (crypter) l'information (message en clair) en suivant une certaine "recette" (algorithme) sur laquelle Alice et Bob se sont mis d'accord. En règle générale, la recette utilisée fait partie de toute une famille de mêmes recettes qui diffèrent seulement par la valeur de certains ingrédients (la clé). D'autres personnes peuvent connaître quel type de recette est utilisé par Alice et Bob mais s'ils ne connaissent pas les ingrédients exacts ils ne pourront pas, en cas d'interception du message crypté, reconstituer le message en clair (du moins si le système utilisé est performant!). Il est clair aussi que le cryptage doit être réversible en ce sens que lorsque Bob reçoit le message il puisse avoir l'"antidote" lui permettant de retrouver le message en clair.

Un exemple

Prenons l'exemple où la recette utilisée est le décalage des lettres qui consiste à faire tourner les lettres de l'alphabet : ça c'est la recette générale. Celle-ci donne lieu à 25 sous-recettes particulières : on peut décaler A sur B ou A sur C ou ... A sur Z. On a donc 25 ingrédients différents à notre disposition, autrement dit 25 clés possibles. Ce procédé de cryptage est connu comme le système de Jules César, ce dernier l'ayant beaucoup utilisé avec la clé consistant à décaler A sur D. Pour Jules César le mot ESSAI devient une fois crypté HVVDL.

*Mini-conférence dans la classe de TS1 du lycée Daudet, le 17 décembre 2004.

2 Codage et décodage

Suivant la célèbre expression de Pythagore, « Tout est nombre » !

Les informations qui doivent être cryptées se présentent au départ sous forme de textes ou de nombres (par exemple numéro de carte bancaire). Quoiqu'il en soit, on aura très vite intérêt à manipuler exclusivement des nombres car ce sera plus pratique. Par exemple, même avec le système simple de Jules César, on peut "coder" l'alphabet en faisant $A=0, B=1, \dots, Z=25$. Le mot ZUT devient alors le "mot" 25, 20, 19. Notons qu'il est important de mettre des virgules pour séparer les nombres car sinon le codage serait ambigu, le nombre 252019 pouvant tout aussi bien représenter le mot ZUBJ ou d'autres (lesquels?)! Avec ce codage, le mot 25, 20, 19 serait crypté par Jules César en ajoutant 3 à chaque nombre, soit 28, 23, 22; ce faisant on obtient une lettre numéro 28 qui est en fait la lettre numéro 2, c'est-à-dire C, d'où le mot crypté CXW. Ainsi pointent le nez, les congruences modulo 26.

Maintenant, pour coder un texte par un seul nombre et non une suite de nombres, on a déjà remarqué que la simple concaténation des numéros des lettres était ambigu. Pour enlever cette ambiguïté, on peut recourir à une écriture en base b où b est le nombre de lettres de l'alphabet utilisé (par exemple 26). Pour revenir à ZUT et à son codage 25, 20, 19, on pourra le considérer comme le nombre dont les chiffres en base 26 sont 25, 20, 19 c'est-à-dire le nombre $25 \times 26^2 + 20 \times 26 + 19$; ce dernier nombre pourra être écrit alors en base 10, soit 17439 ou en base 2, soit 100010000011111, ou tout autre base selon les besoins. Évidemment, si vous écrivez un texte en le codant suivant le procédé qui envoie ZUT sur 17439 ou pire sur 100010000011111, vous aurez déjà un texte assez incompréhensible pour votre petite soeur (ou petit frère) mais on ne peut pas qualifier cela de cryptage (sérieux)! Il s'agit d'un cryptage de niveau zéro, d'un codage, uniquement destiné à transformer la matière première en quelque chose d'exploitable par les mathématiques.

Exercice 1

a) Coder les mots MATHS, CRYPTOGRAPHIE et VACANCES suivant le procédé précédent (on donnera leurs codages décimal et binaire).

b) Décoder 6767203 et 1011100110110 sachant que le premier est le codage décimal et le second le codage binaire d'un mot selon le procédé précédent.

3 Cryptage affine par blocs; procédé de Hill

On a vu dans la précédente section que si l'on code l'alphabet A, B, \dots, Z par $0, 1, \dots, 25$, le cryptage de Jules César revient à faire l'opération $x \mapsto$

$x + c \pmod{26}$, où c est la clé choisie, par exemple $c = 3$. On peut généraliser cette méthode en se permettant de multiplier x par un nombre; on fera alors la transformation affine $x \mapsto ax + c \pmod{26}$. Pour que ce cryptage soit sans ambiguïté et donc puisse être décrypté par Bob, on ne peut pas utiliser n'importe quel nombre a . En effet, par exemple, si on effectue la transformation $x \mapsto 2x + 1 \pmod{26}$, on obtient pour $x = 1$ et $x = 14$ le même résultat 3 c'est-à-dire que les lettres B et O sont toutes les deux envoyées sur la lettre D.

Exercice 2 Montrer que ce cryptage est sans ambiguïté si, et seulement si a est premier avec 26.

Exercice 3 Crypter le mot JULES selon ce procédé avec $a = 5$ et $c = 2$ puis décrypter le mot VUFQ.

On peut encore généraliser ce procédé de la manière suivante. Étant donné un message, découpons le en tranches de longueur l donnée, par exemple pour commencer $l = 2$. Le mot MATHEMATIQUE est ainsi scindé en

MA TH EM AT IQ UE

Choisissons un tableau de nombres entiers $\mathcal{A} = \begin{pmatrix} a, b \\ c, d \end{pmatrix}$, appelé *matrice* et un couple (x_0, y_0) d'entiers. Si (x, y) est le codage d'un groupe de deux lettres, par exemple $(12, 0)$ pour MA, on lui fait correspondre le couple

$$(ax + by + x_0 \pmod{26}, cx + dy + y_0 \pmod{26})$$

On obtient ainsi un cryptage de chaque bloc de longueur 2 donc un cryptage de tout le message. Là encore on ne peut pas choisir n'importe quelle matrice, sans quoi le cryptage peut être ambigu.

Exercice 4

a) Montrer que ce type de cryptage (avec $l = 2$) est possible (sans ambiguïté) si, et seulement si, les nombres entiers a, b, c, d sont tels que $ad - bc$ est premier avec 26. La quantité $ad - bc$ est appelée le *déterminant* de la matrice \mathcal{A} .

b) On choisit $a = 1, b = c = 0, d = 1$ et $x_0 = 3, y_0 = 14$. Crypter le mot MATHEMATIQUE et interpréter ce système en termes non mathématiques (procédé de Vigenère).

c) On choisit cette fois $a = 4, b = 5, c = 1, d = 2$ et $x_0 = y_0 = 0$. Crypter à nouveau le mot MATHEMATIQUE puis décrypter le mot GVICKRCO.

d) Généraliser ce procédé à un découpage en tranches de longueur $l \geq 3$. Ce procédé consistant à "multiplier" des tranches de longueur l du message par une matrice à coefficients entiers de taille $l \times l$, dont le déterminant est un entier premier avec le nombre de lettres de l'alphabet utilisé (ici 26), date des années 1930 et est dû à Lester Hill.

4 Le système RSA

Les systèmes cryptographiques se scindent en deux grandes familles, ceux à clé secrète ou symétrique, ceux à clé publique ou asymétrique.

Dans la première famille, on trouve par exemple les systèmes de Jules César, de Vigenère ou leur généralisation du type cryptage de Hill vus à la section 2. Pour pouvoir communiquer Alice et Bob ont dû au préalable, non seulement se mettre d'accord sur la méthode utilisée mais encore sur la clé (le nombre de décalages pour Jules César, la matrice \mathcal{A} et le couple (x_0, y_0) pour le cryptage de Hill. Cette clé ne doit être connue que d'eux (elle est secrète) et leur connaissance de la clé est symétrique en ce sens que tous les deux connaissent exactement la même chose ou du moins quelque chose qui leur permet de savoir ce que sait l'autre (la clé de cryptage permet de connaître la clé de décryptage par un calcul simple).

Dans la seconde famille, on trouve des algorithmes qui ont vu le jour, pour les premiers d'entre eux, au milieu des années 1970, donc assez récemment et dont le plus célèbre et le plus utilisé est l'algorithme RSA (des initiales de ses inventeurs, Rivest, Shamir et Adleman). Il pallie un défaut de la première famille qui nécessite qu'Alice et Bob se soient au préalable mis d'accord sur la clé et donc aient échangé une première information (mais comment crypter cette information initiale si Alice et Bob n'ont pu se rencontrer et se murmurer la clé à l'oreille ce qui est le cas lorsque Alice et Bob ne sont autres que votre ordinateur qui communique avec un ordinateur distant !) L'idée est que la clé sera composée d'une partie secrète qui sera propre à chacun des protagonistes et ne nécessitera donc pas de partage préalable et d'une partie dite *publique* car on se souciera peu qu'elle soit connue de tous et qui seule fera l'objet d'un échange préalable entre Alice et Bob. Cela justifie les termes de *publique* et *asymétrique* utilisés.

Les deux exercices suivants propose la théorie de la méthode ainsi qu'un exemple d'utilisation.

Exercice 5 Un système cryptographique moderne : l'algorithme RSA

Alice doit transmettre à Bob des informations confidentielles sur un canal non sûr. Pour cela Bob choisit deux entiers premiers distincts p et q qu'il garde secrets et envoie à Alice, sans se soucier de l'interception éventuelle de cette donnée, le produit $n = pq$ de ces deux nombres. Il choisit ensuite un nombre e *inversible modulo* $\phi(n) = (p - 1)(q - 1)$, c'est-à-dire tel qu'il existe un entier d vérifiant $ed \equiv 1 \pmod{\phi(n)}$ et le communique aussi à Alice (en revanche il garde d secret). On peut vérifier que l'inversibilité de e modulo ϕ équivaut au fait que e soit premier avec $\phi(n)$. En pratique p et q et donc aussi n sont très grands et c'est sur ce fait qu'est basée la sécurité du système car alors connaissant seulement e et n , la découverte de p , q , d prendrait un temps considérable. La donnée que veut

transmettre Alice est un entier naturel $M < n$. Pour ce faire, elle envoie à Bob la donnée cryptée $C = M^e \pmod{n}$.

Le “miracle” RSA consiste dans le fait que Bob peut reconstituer le message initial M d’Alice car on peut montrer (le faire) que

$$C^d \equiv M \pmod{n}.$$

Exercice 6 Paiement sécurisé par RSA

Vous vendez des produits en ligne sur internet et proposez à vos clients un paiement sécurisé par carte bancaire. Pour cela, lorsqu’un client saisit son numéro de carte bancaire, qui consiste en 4 groupes de 4 chiffres, les trois premiers groupes sont transmis en clair tandis que le dernier groupe (noté M) est crypté. Pour ce cryptage, vous avez choisi le procédé RSA. En pratique, pour chaque client votre ordinateur tire au hasard des nombres premiers p et q (secrets) et transmet à l’ordinateur du client $n = pq$ et e inversible modulo $(p-1)(q-1)$ pour que celui-ci puisse lui transmettre $M^e \pmod{n}$. Pour le client Alice, les clés RSA sont les suivantes : $p = 53$ et $q = 71$ (secrets), $n = pq = 3763$ et $e = 331$.

1) Vérifier que e est inversible modulo $(p-1)(q-1)$ et calculer son inverse d .

2) Le numéro transmis par Alice est 1234 8765 1973 * * * * et vous recevez 1234 8765 1973 **2286** c’est-à-dire que le nombre à quatre chiffres M qui est crypté est transmis sous la forme $M^e \equiv 2286 \pmod{n}$.

Quel est le numéro de carte bancaire de votre cliente Alice ?

Souvent, en pratique on fait un mixage des méthodes à clé secrètes et à clés publiques. En effet, les premières sont en général moins coûteuses en temps de calcul et sont utilisées pour crypter l’essentiel de la transaction, la méthode à clé publique n’intervenant alors que dans la phase initiale pour permettre de crypter l’échange des clés secrètes qui vont être utilisées ensuite. Dans l’exercice suivant, nous simulons un tel mixage tout à fait fantaisiste pour mieux faire comprendre les lignes précédentes.

Exercice 7 Transmission d’un secret par RSA

Vous êtes Bob et décidez avec Alice que celle-ci vous transmettra des messages cryptés par une méthode classique à clé secrète. Pour cela il est nécessaire qu’Alice vous transmette la clé secrète qu’elle a choisie. Cette clé est un mot qu’elle va vous faire parvenir sous forme cryptée par le système à clé publique RSA. Pour réaliser cela, vous devez choisir des nombres premiers p et q (secrets) et envoyer à Alice $n = pq$ et un entier e inversible modulo $\phi(n) = (p-1)(q-1)$ afin que celle-ci puisse vous transmettre $M^e \pmod{n}$, où M est son message écrit sous forme d’un entier $< n$. Vous choisissez $p = 487$, $q = 491$ (secrets), $n = pq = 239117$ et $e = 216491$.

1) Calculer $\phi(n) = (p-1)(q-1)$.

2) Vérifier que e est inversible modulo $\phi(n)$ et calculer son inverse d .

3) Pour vous envoyer son mot secret Alice commence par le coder en un nombre décimal par le procédé décrit dans la section codage (passer en base 26).

Alice vous envoie $M^e \equiv 39178 \pmod{n}$

- a) Combien vaut M ?
- b) Quel est le mot clé choisi par Alice ?

Enfin un dernier exemple de procédé mixte (pour les gourmands uniquement !) :

Exercice 8 Vigenère et RSA

Vous êtes Bob, et Alice doit vous envoyer un message chiffré (=crypté) selon la méthode de Vigenère (évoquée à l'exercice 4). Elle choisit un mot clé, le code toujours suivant le même procédé, ce qui lui fournit un nombre décimal K et vous le transmet sous forme cryptée en utilisant le système RSA. Pour cela, elle choisit les entiers premiers $p = 7$ et $q = 617$ et publie $n = pq = 4319$ et $e = 2609$ (N.B. Alice a fait en sorte que $K < n$).

a) Vérifier que e est inversible modulo $(p - 1)(q - 1)$ et calculer son inverse d modulo $(p - 1)(q - 1)$.

b) Alice vous transmet alors le couple $(\mathcal{C}, K^e \pmod{n})$ où \mathcal{C} est le message crypté suivant Vigenère et $K^e \pmod{n}$ la clé cryptée suivant RSA. Dans le cas présent, Alice vous a transmis, pour le message $\mathcal{C} =$ crypté par Vigenère,

« UZR XCIUTIZ Y IAE TID PICKM OY BZYB XSV OICIMMXI VP
QM OMB AEA WE DPVQEI UZR BZYB PWB FR OPRMCET CSULMV »

et pour la clé cryptée par RSA, $K^e \equiv 2467 \pmod{n}$.

- i) Combien vaut K ?
- ii) Quel est le mot clé choisi par Alice pour le procédé de Vigenère ?
- iii) Quel est le message envoyé par Alice ?
- iv) Sauriez-vous répondre au problème posé par Alice ? !