Sécurité des Données en se basant sur le Chiffre de Hill

Naima HADJ-SAID[1]–A. ALI-PACHA[1] – A. M'HAMED[2]–A. BELGORAF[1]

[1]Université des Sciences et de la Technologie d'Oran –USTO, BP 1505 El M'Naouer Oran 31036 ALGERIE
[2]Institut National des Télécommunications Evry- Paris
Tél. /Fax : 213 – 041 / 46 26 85
E.Mail : nim_hadj@yahoo.fr

Résume :
La cryptologie, véritable science régissant le codage de l'information, a connu une réelle explosion avec le développement des systèmes informatiques, passant d'une ère artisanale et confidentielle à des systèmes de très hautes technologies nécessitant une importante puissance de calcul. Elle a connu un plus large essor encore avec l'arrivée des systèmes de communications modernes (Internet, etc...) où il y a une nécessité absolue de protéger les données échangées des individus.

Dans cette communication on essaye d'étudier Le chiffre Lester S. Hill, ce chiffre regroupe deux thèmes principaux : le calcul matriciel et le calcul modulo 26. Le contexte est celui des messages secrets. C'est un sujet qui est très riche en applications mathématiques et qui intéresse beaucoup les communications et le commerce électroniques.

Mots Clés : Cryptographie, Hill, Matrices, Polygraphiques.

1. Introduction :

Dans un contexte où les échanges d'information dématérialisées se développent, il est indispensable de pouvoir bénéficier de systèmes sécurisés pour protéger les données a caractère personnel ou confidentiel. Il est donc, nécessaire d'avoir accès à des outils permettant une meilleure protection contre les intrusions arbitraires de la confidentialité des données. Le chiffrement est souvent le seul moyen efficace pour répondre à ces exigences.
La cryptographie est l'ensemble des processus de verrouillage visant à protéger l'accès à certaines données afin de les rendre incompréhensible aux personnes non autorisées. Les technologies cryptographiques sont ainsi reconnues comme étant des outils essentiels de sécurité des données et de la confiance dans les communications et le commerce électroniques.
Le **chiffre de Hill** que nous allons étudier afin de l'implémenter pour sécuriser les données a été publié par *Lester S. Hill* en 1929. C'est un chiffre polygraphique, c'est-à-dire qu'on ne (dé) chiffre pas les lettres les unes après les autres, mais par paquets. Nous étudierons la version bigraphique, c'est-à-dire que nous grouperons les lettres deux par deux, mais on peut imaginer des paquets plus grands. C'est une méthode de chiffrement qui utilise des matrices carrées. Cet algorithme est, plus au moins, performant pour une petite entreprise qui veux se doter d'un moyen software sécurisant ses données et qui n'est pas cher.

2. Notion sur les Matrices

En mathématiques, les matrices servent à interpréter en termes calculatoires et donc opérationnels les résultats théoriques de l'algèbre linéaire et même de l'algèbre bilinéaire. Toutes les disciplines étudiant des phénomènes linéaires utilisent les matrices. Quant aux phénomènes non linéaires, on en donne souvent des approximations linéaires comme c'est le cas en optique géométrique avec les approximations de Gauss.

2.1 Matrice :

Soient A un ensemble et (m,n) un couple d'entiers positifs. On appelle matrice à coefficients dans A, corps commutatif quelconque, de dimension (ou taille) (m,n) ie à m lignes et n colonnes, une famille $(a_{i,j})$ d'éléments de A indexée par le produit cartésien des ensembles de nombres entiers [1,m] et [1,n]. La matrice M pourra être notée par :

$$M = (a_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n}$$

On représente généralement une matrice sous la forme d'un tableau rectangulaire. Par exemple, est représentée ci-dessous une matrice M, à coefficients entiers, et de dimension (3,4) :

$$M = \begin{pmatrix} 0 & 1 & 2 & 3 \\ 4 & 5 & 6 & 7 \\ 8 & 9 & 10 & 11 \end{pmatrix}$$

Dans cette représentation, le premier coefficient de la dimension est le nombre de lignes, et le deuxième, le nombre de colonnes du tableau. Une matrice pour laquelle le nombre m de lignes est égal au nombre n de colonnes sera dite matrice carrée de taille n. Une matrice ne comportant qu'une seule ligne et n colonnes est appelée matrice ligne de taille n. Une matrice ne comportant m lignes et une seule colonne est appelée matrice colonne de taille m.
Pour repérer un coefficient d'une matrice, on indique son indice de ligne puis son indice de

colonne, les lignes se comptant du haut vers le bas et les colonnes de la gauche vers la droite. Par exemple, on notera $a_{i,j}$, les coefficients de la matrice *M*, pour 1≤ i≤ 3 désignant le numéro de la ligne sur laquelle figure le coefficient envisagé, et 1≤ j ≤ 4 désignant son numéro de colonne ; ainsi $a_{2,4}=7$. La disposition générale des coefficients d'une matrice *M* de taille *(m,n)* est donc la suivante :

$$M = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix}$$

Pour effectuer certaines opérations, il peut être utile de travailler sur le système des lignes ou des colonnes d'une matrice. On pourra alors l'écrire sous une des formes suivantes :

$$M = \begin{pmatrix} L_1 \\ L_2 \\ \vdots \\ L_m \end{pmatrix} \quad \text{ou} \quad M = \begin{pmatrix} C_1 & C_2 & \dots & C_n \end{pmatrix}.$$

L'ensemble des matrices à coefficients dans *A* possédant *m* lignes et *n* colonnes est noté $M_{m,n}(A)$. Lorsque *m=n on note plus simplement* $M_n(A)$.Soit :

$$M = (a_{i,j})_{1\leq i\leq m,\ 1\leq j\leq n} \in M_{m,n}(A)$$

On appelle transposée de *A* la matrice :

$${}^tM = (a_{j,i})_{1\leq j\leq n,\ 1\leq i\leq m}.$$

Remarquons que :

$${}^tM \in M_{n,m}(A).$$

Par exemple, avec la matrice *M* des exemples précédents, on a :

$${}^tM = \begin{pmatrix} 0 & 4 & 8 \\ 1 & 5 & 9 \\ 2 & 6 & 10 \\ 3 & 7 & 11 \end{pmatrix}$$

L'opération de transposition est involutive, c'est-à-dire que ${}^t({}^tM) = M$

2.2 Matrice inverse :

2.2.1 Méthodes d'inversion

Avant de décrire les méthodes usuelles d'inversion, notons qu'en pratique, il n'est pas nécessaire de calculer l'inverse d'une matrice pour résoudre un système d'équations linéaires. Il est toutefois nécessaire que la matrice considérée soit inversible. Des méthodes de décomposition comme la décomposition LU sont *beaucoup plus rapide* que l'inversion.

2.2.2 Méthode des cofacteurs

L'inverse d'une matrice *A* s'écrit sous une forme très simple à l'aide de la matrice complémentaire ${}^t\mathrm{com}A$

$$A^{-1} = \frac{1}{\det A}\,{}^t\mathrm{com}A = \frac{1}{\det A}\,{}^t(C_{ij}) = \frac{1}{\det A}\begin{pmatrix} C_{11} & C_{21} & \cdots & C_{j1} \\ C_{12} & \ddots & & C_{j2} \\ \vdots & & \ddots & \vdots \\ C_{1i} & \cdots & \cdots & C_{ji} \end{pmatrix}$$

Où det*A* est le déterminant de *A*, com*A* est la comatrice de *A* et ${}^tA$ est la matrice transposée de *A*. Cette écriture permet un calcul aisé de l'inverse d'une matrice de petite dimension. Pour des matrices de plus grandes dimensions, cette méthode essentiellement récursive devient inefficace.

2.2.3 Inversion des matrices 2 x 2

L'*équation des cofacteurs* ci-dessus permet de calculer l'inverse des matrices de dimensions 2x2 :

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix},\ A^{-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

$$\det A = ad\text{-}bc \neq 0$$

3. Chiffre de Hill

Lester Hill, mathématicien cryptographe (1891-1961) publie en 1929 dans la revue American Mathematical Monthly un article intitulé *Cryptography in an algebraic alphabet*, où il détaille un nouveau type d'algorithme de chiffrement. Son idée n'est plus de coder lettres par lettres, mais de coder simultanément des groupes de m lettres! Bien sûr, plus m est grand, plus les analyses statistiques deviennent difficiles!. Le chiffre de Hill est une méthode de chiffrement qui utilise des matrices carrées. On désigne par 2-chiffre de Hill, le chiffre obtenu en codant les lettres par blocs de deux, par 3-chiffre de Hill celui obtenu en codant les lettres par blocs de trois, et ainsi de suite.

3.1 Chiffrement

Dans une première phase, chaque lettre du texte à chiffrer est remplacée par une valeur numérique, celle de son rang dans l'alphabet, c'est-à-dire, nous remplaçons chaque lettre par son ordre dans l'alphabet : A devient 1, B devient 2,..., Z devient 26. On groupe les nombres ainsi obtenus par m (prenons par exemple m=2).
Les lettres P_k et P_{k+1} du texte clair seront chiffrés C_k et C_{k+1} avec la formule ci-dessous:

$$\begin{pmatrix} C_k \\ C_{k+1} \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} P_k \\ P_{k+1} \end{pmatrix} \pmod{26}$$

Ce qui signifie, pour fixer les idées, que les deux premières lettres du message clair (P_1 et P_2) seront chiffrées (C_1 et C_2) selon les deux équations suivantes:

$$C_1 \equiv aP_1 + bP_2 \pmod{26}$$
$$C_2 \equiv cP_1 + dP_2 \pmod{26}$$

a,b,c,d sont des entiers, C_1 et C_2 seront aussi des entiers. Le choix de la clé correspond ici au choix d'un nombre m, et au choix des combinaisons linéaires à effectuer (ce sont toujours les mêmes de blocs en blocs).

Remarque

a	**b**	**c**	**d**	**e**	**f**	**g**
1	2	3	4	5	6	7
h	**i**	**j**	**k**	**l**	**m**	**n**
8	9	10	11	12	13	14
o	**p**	**q**	**r**	**s**	**t**	**u**
15	16	17	18	19	20	21
v	**w**	**x**	**y**	**z**		
22	23	24	25	0		

La valeur 0 est attribuée à la lettre Z afin de travailler modulo 26. Cependant, d'autres auteurs posent "A"=0, "B"=1, ..., "Z"=25. Les deux conventions se défendent. L'essentiel est que les protagonistes se mettent d'accord avant d'échanger des messages. On pourrait même imaginer de prendre un alphabet désordonné, par exemple "A"=15, "B"=6, etc., ce qui constituerait un surchiffrement.

3.2 Exemples de chiffrement

1) On souhaite coder le message "je vous aime" en prenant comme clef de cryptage la matrice $P_1=\begin{pmatrix}9 & 4\\5 & 7\end{pmatrix}$. Après avoir remplacé les lettres par leur rang dans l'alphabet (a=1, b=2, etc.), elle obtiendra:

$$C_1 \equiv 9\cdot 10+4\cdot 5 \pmod{26} = 110 \pmod{26} = 6$$
$$C_2 \equiv 5\cdot 10+7\cdot 5 \pmod{26} = 85 \pmod{26} = 7$$

Elle fera de même avec les 3e et 4e lettres, 5e et 6e, etc. Elle obtiendra finalement:

Lettres	j	e	v	o	u	s	a	i	m	e
Rangs (P_k)	10	5	22	15	21	19	1	9	13	5
Rangs chiffrés (C_k)	6	7	24	7	5	4	19	16	7	22
Lettres chiffrées	F	G	X	G	E	D	S	P	G	V

2) Soit la matrice $P_2=\begin{pmatrix}3 & 2\\1 & 3\end{pmatrix}$, si on veut crypter le message "Rendez-vous ce soir", on obtient le cryptogramme : "UDZBI WLOSR VSSAY STAIE AM"

3.3 Déchiffrement :

Pour déchiffrer, le principe est le même que pour le chiffrement: on prend les lettres deux par deux, puis on les multiplie par une matrice.

$$\begin{pmatrix}P_1\\P_2\end{pmatrix} = \begin{pmatrix}a & b\\c & d\end{pmatrix}^{-1}\begin{pmatrix}C_1\\C_2\end{pmatrix} \pmod{26}$$

Cette matrice doit être l'inverse de matrice de chiffrement (modulo 26). Ordinairement l'inverse de la matrice $\begin{pmatrix}a & b\\c & d\end{pmatrix}$ est:

$$\begin{pmatrix}a & b\\c & d\end{pmatrix}^{-1} = \frac{1}{ad-bc}\begin{pmatrix}d & -b\\-c & a\end{pmatrix}$$

Mais que cela signifie-t-il dans le contexte Z_{26}? Reprenons notre exemple.

3.4 Exemple de déchiffrement

Pour déchiffrer le message qu'on a envoyé par la clef P_1 doit calculer :

$$\begin{pmatrix}9 & 4\\5 & 7\end{pmatrix}^{-1} = \frac{1}{43}\begin{pmatrix}7 & -4\\-5 & 9\end{pmatrix} \pmod{26} = (43)^{-1}\begin{pmatrix}7 & -4\\-5 & 9\end{pmatrix} \pmod{26}$$

Comme pgdc(43,26)=1, $(43)^{-1}$ existe dans Z_{26} et $(43)^{-1}$ égale 23. Le récepteur a donc maintenant la matrice de déchiffrement:

$$\begin{pmatrix}9 & 4\\5 & 7\end{pmatrix}^{-1} = 23\begin{pmatrix}7 & -4\\-5 & 9\end{pmatrix} \pmod{26} = \begin{pmatrix}161 & -92\\-115 & 207\end{pmatrix} \pmod{26} = \begin{pmatrix}5 & 12\\15 & 25\end{pmatrix} \pmod{26}$$

Il prend donc la matrice $\begin{pmatrix}5 & 12\\15 & 25\end{pmatrix}$ pour déchiffrer le message "FGXGE DSPGV". Après avoir remplacé les lettres par leur rang dans l'alphabet (A=1, B=2, etc.), il obtiendra:

$$P_1 \equiv 5\cdot 6+12\cdot 7 \pmod{26} = 114 \pmod{26} = 10$$
$$P_2 \equiv 15\cdot 6+25\cdot 7 \pmod{26} = 265 \pmod{26} = 5$$

Il fera de même avec les 3e et 4e lettres, 5e et 6e, etc. Il obtiendra finalement:

Lettres chiffrées	F	G	X	G	E	D	S	P	G	V
Rangs chiffrés (C_k)	6	7	24	7	5	4	19	16	7	22
Rangs (P_k)	10	5	22	15	21	19	1	9	13	5
Lettres	j	e	v	o	u	s	a	i	m	e

3.5 Chiffre de Hill avec des mathématiques

Le chiffre de Hill est à l'intersection de l'arithmétique et de l'algèbre linéaire. En remplaçant les lettres par des nombres (A-->0,...), on ne traite plus que des entiers compris entre 0 et 25. En outre, un nombre n est identifié avec tous les nombres n+26k, où k est un entier (en clair, si 1 représente B, 27,53,-25... aussi!).

Quand les calculs faits par les combinaisons linéaires sortent des entiers de 0 à 25, on s'y ramène en prenant le reste dans la division par 26. On dit que l'on travaille dans **Z|26Z**.

Chaque groupe de 2 lettres, ou par identification de 2 nombres x_1-x_2, est représenté par un vecteur colonne $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. Les relations de dépendance linéaire sont, comme souvent, représentés par une matrice $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. On a, dans **Z|26Z**, la relation

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

où $\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$ est le bloc codé, et $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ est le bloc clair.

3.6 Réversibilité :

Toute matrice de chiffrement ne convient pas! Par exemple, si on prend la matrice A suivante:

$$\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$$

Elle n'est pas une bonne matrice de chiffrement, car par exemple,

$$A\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \text{ et } A\begin{pmatrix} 2 \\ 25 \end{pmatrix} = \begin{pmatrix} 52 \\ 104 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

(Où on a ramené les calculs dans **Z|26Z**). Ainsi, deux vecteurs (ou encore deux couples de deux lettres) différents sont codés de la même façon. Il est donc impossible, même en connaissant la clé, de décrypter.

Pour que le processus soit inversible, il est nécessaire et suffisant que A soit inversible, mais attention, inversible dans **Z|26Z**. On montre que cela est vérifié si, et seulement si, det A=ad-bc est inversible dans **Z|26Z** (c'est-à-dire qu'il existe k un entier tel que k×det A=1+26n, où n est un entier - cela est équivalent à dire que det A est premier avec 26). Dans ce cas, l'inverse est donné par :

$$A^{-1} = (\det A)^{-1} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

On décrypte en utilisant le même procédé, mais en utilisant A^{-1}.

Il est facile de montrer que A est régulière dans Z_{26} si son déterminant D est inversible modulo 26 (i.e. D et 26 sont premiers entre eux). Le calcul de A^{-1} s'effectue selon les règles usuelles en travaillant dans Z_{26} .

On ne peut pas prendre n'importe quoi comme matrice de chiffrement. Ses composantes doivent tout d'abord être des **nombres entiers positifs**. Il faut aussi qu'elle ait une matrice inverse dans Z_{26}.

$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} \pmod{26}$ existe si $(ad-bc)^{-1}$ (mod 26) existe. Or $(ad-bc)^{-1}$ (mod 26) existe si (ad-bc) et 26 sont premiers entre eux. Il faut donc contrôler que **(ad-bc) est impair et n'est pas multiple de 13** (voir à ce sujet le chiffre affine).

Exemple : Soit k=43. Trouvons k^{-1} (mod 26).

43·1 (mod 26) = 43 (mod 26) = 17 (mod 26).	Donc 1 n'est pas le nombre cherché.
43·3 (mod 26) = 129 (mod 26) = 25 (mod 26).	Donc 3 n'est pas le nombre cherché.
43·5 (mod 26) = 215 (mod 26) = 7 (mod 26).	Donc 5 n'est pas le nombre cherché.
43·7 (mod 26) = 301 (mod 26) = 15 (mod 26).	Donc 7 n'est pas le nombre cherché.
43·9 (mod 26) = 387 (mod 26) = 23 (mod 26).	Donc 9 n'est pas le nombre cherché.
43·11 (mod 26) = 473 (mod 26) = 5 (mod 26).	Donc 11 n'est pas le nombre cherché.
43·15 (mod 26) = 645 (mod 26) = 21 (mod 26).	Donc 15 n'est pas le nombre cherché.
43·17 (mod 26) = 731 (mod 26) = 3 (mod 26).	Donc 17 n'est pas le nombre cherché.
43·19 (mod 26) = 817 (mod 26) = 11 (mod 26).	Donc 19 n'est pas le nombre cherché.
43·21 (mod 26) = 903 (mod 26) = 19 (mod 26).	Donc 21 n'est pas le nombre cherché.
43·23 (mod 26) = 989 (mod 26) = 1 (mod 26).	**Donc 23 est le nombre cherché. STOP**.

La table suivante donne les inverse modulo 26:

1	3	5	7	9	11	15	17	19	21	23	25
1	9	21	15	3	19	7	23	11	5	17	25

3.7 Cas d'image

Nous avons pris comme exemple une image paysage bitmap [256*256] pixels et nous avons crypté cette image par les deux clés de chiffrement P_1 et P_2, voici les résultats obtenus :

Image originale

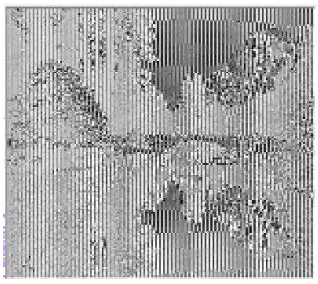

Image cryptée par le P_1

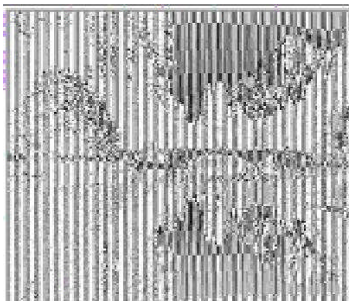

Image cryptée par P_2

Les images cryptées ont la même taille que l'image claire.

3.8. Attaque à texte clair connu : chiffrement de Hill

Lorsque l'on cherche à déterminer la clé de chiffrement d'un adversaire, on peut se situer à plusieurs niveaux d'information. On peut n'avoir à sa disposition que le message chiffré. Parfois, et cela apporte beaucoup d'informations, on dispose à la fois du message chiffré et de sa traduction en clair, ou au moins une partie de celle-ci. Cela n'est pas si saugrenu : bien des messages protocolaires (et dans l'armée, les protocoles...) comportent le même début ou la même fin, et c'est ainsi par exemple que Türing a procédé pour la machine Enigma. C'est ce que l'on appelle l'attaque à texte clair connu!

Voyons un exemple à partir du chiffrement de Hill (cette méthode s'applique à tout algorithme fonctionnant à partir de combinaisons linéaires). On suppose qu'on a le texte codé suivant : COR ZZETMDW...., qui correspond au début de MON GENERAL... Un espion dans les bases ennemies nous a permis de savoir que nos adversaires utilisent le chiffre de Hill, avec une longueur de clé égale à 2. On note A la matrice 2×2 de chiffrement à coefficients dans **Z|26Z**. La première paire CO s'obtient en appliquant la matrice A à partir de la paire MO, la seconde paire RZ s'obtient en appliquant la matrice A à partir de la paire NG. Cela se traduit matriciellement par la relation :

$$\begin{pmatrix} 2 & 17 \\ 14 & 25 \end{pmatrix} = A \begin{pmatrix} 12 & 13 \\ 14 & 6 \end{pmatrix}$$

que nous écrivons sous la forme B=AC. Si la matrice C est inversible dans **Z|26Z**, on obtient en multipliant à droite $A=BC^{-1}$. On s'empresse de calculer le déterminant de C : il vaut -110, qui n'est pas premier avec 26. C n'est pas inversible dans **Z|26Z**. C'est raté! On recommence avec les deuxième et troisième paires :

$$\begin{pmatrix} 17 & 25 \\ 25 & 4 \end{pmatrix} = A \begin{pmatrix} 13 & 4 \\ 6 & 13 \end{pmatrix}$$

qu'on écrit en D=AE. Le déterminant de E vaut 145, il est premier avec 26, et E est inversible dans **Z|26Z**. On obtient alors :

$$E^{-1} = \begin{pmatrix} 13 & 24 \\ 10 & 13 \end{pmatrix} \text{ et } A = \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix}.$$

On peut vérifier la matrice de chiffrement sur les autres paires.

3.9 Autre exemple : décryptement d'un chiffre de Hill (2x2)

Le chiffre de Hill est sensible aux attaques de type texte clair connu. Nous allons décrypter le message ci-dessous, sachant qu'il contient le nom GEORGE PAPANDREOU:

CMYPZ GTAYO EQBYQ JLAOW INELN NECNN UESZT YTFRU OWYXH KYADM NJRUK CUFZP YPNNM XWSQQ OJMGO JZQZQ FLVAY XGIPR OPUFJ WTSVA ATQU

Remarquons tout d'abord que "GEORGE PAPANDREOU" contient des **répétitions de bigrammes à des intervalles pairs**:

1. GEORGE PAPANDREOU
2. GEORGE PAPANDREOU

On retrouvera ces répétitions de bigrammes dans le texte chiffré si le rang de la première lettre du bigramme est impair. Avec GEORGE PAPANDREOU, on est sûr de retrouver soit GE...GE et PAPA, soit EO...EO, car soit c'est G qui est de rang impair soit c'est E.

On remarque dans le cryptogramme le segment ZQZQ qui pourrait correspondre à PAPA; le premier Z de ce bigramme est la 77ème lettre du cryptogramme. Essayons cette correspondance (tableau de correspondance) :

Cela semble coïncider parfaitement. Si notre hypothèse est exacte, alors, après avoir remplacé les lettres par leur rang dans l'alphabet (a=1, b=2, ..., y=25, z=0), le couple chiffré (15;10) a été obtenu à partir du couple clair (7;5), (13;7) à partir de (15;18), (0;17) à partir de (16;1), etc. Il s'agit maintenant de trouver la matrice de déchiffrement (D) à partir de ces couples.

Chiffré	OJ	MG	OJ	ZQ	ZQ	FL	VA	YX
Couples chiffrés	(15;10)	(13;7)	(15;10)	(0;17)	(0;17)	(6;12)	(22;1)	(25;24)
Couples clairs	(7;5)	(15;18)	(7;5)	(16;1)	(16;1)	(14;4)	(18;5)	(15;21)
Clair	GE	OR	GE	PA	PA	ND	RE	OU

Tableau de correspondance)

Dans notre tableau, prenons les premiers et les quatrièmes couples chiffrés et formons une matrice (A) en disposant verticalement ces valeurs. Prenons les premiers et les quatrièmes couples clairs de notre tableau pour former une matrice (B). On aurait pu choisir n'importe quel couple de colonnes du tableau pourvu que la matrice A formée soit inversible modulo 26.

On obtient l'équation matricielle:

$$D\ A = B$$

Dans notre exemple, on a:

$$D\begin{pmatrix}15 & 0\\ 10 & 17\end{pmatrix} = \begin{pmatrix}7 & 16\\ 5 & 1\end{pmatrix}$$

Pour trouver D, il faut calculer A^{-1}, puis multiplier à droite les deux termes de l'équation avec A^{-1} (tant que A n'est pas inversible modulo 26, il faut essayer d'autres matrices A et B):

$$D\ A\ A^{-1} = B\ A^{-1}, \quad \text{d'où } D = B\ A^{-1}$$

A est inversible modulo 26:

$$\begin{pmatrix}15 & 0\\ 10 & 17\end{pmatrix}^{-1} = \begin{pmatrix}7 & 0\\ 2 & 23\end{pmatrix} \text{ d'où}$$

$$D = \underbrace{\begin{pmatrix}7 & 16\\ 5 & 1\end{pmatrix}}_{B}\underbrace{\begin{pmatrix}7 & 0\\ 2 & 23\end{pmatrix}}_{A^{-1}} = \begin{pmatrix}3 & 4\\ 11 & 23\end{pmatrix} \pmod{26}$$

On obtient alors facilement le message décrypté:

IT IS BELIEVED BY MANY GREEKS THAT THE HEAD OF THE GROUP CALLED THE SHIELD IS THE SON OF GEORGE PAPANDREOU EX PREMIER OF GREECE (T).

4. Conclusion

On peut classer les algorithmes cryptographiques en deux grandes familles. Certains opèrent sur le message en clairs par groupes de bits (blocs). Et d'autres opèrent sur le message en clair un bit à la fois (continue).

Les algorithmes de bloc sont souvent coûteux en termes de temps de calcul, et ne permettent que malaisément le chiffrage d'information de type synchrone (parole, images animées, …). Le chiffre de Hill fait partie de cette catégorie d'algorithme en bloc.

Pour rendre le chiffre de Hill plus difficile à casser, on peut penser aux améliorations suivantes:

- Transposer les lettres avant de les chiffrer.
- Modifier les nombres associés aux lettres. Par exemple, au lieu de dire "A"=1, "B"=2, etc., on pourrait dire "A"=12, "B"=5, etc. Cela revient en fait à surchiffrer le chiffre de Hill avec un alphabet désordonné.
- Agrandir la taille de la matrice de chiffrement. On peut utiliser des matrices 3x3, 4x4, 5x5, etc. La seule limitation est d'ordre pratique: plus les matrices sont grandes, plus le temps de calcul est élevé, et plus le risque d'erreur augmente.

5. Bibliographies

[1] Bruce Schneir, "Applied Cryptography, Protocols, Algorithms, and source Code in C" , edition John Wiley & Sons Inc., 1994.

[2] Beckett Brian :'' Introduction aux méthodes de la cryptologie'', Editions Masson, 1990.

[3] Gilles Zémor, «Cours de Cryptographie», CASSINI, 2000.

[4] Didier Müller, « Une application intéressante des matrices : le chiffre de Hill « , Article paru dans le bulletin No 90 de la SSPMP (www.vsmp.ch), octobre 2002.

[5] http://www.apprendre-en-ligne.net/crypto/rabin/euclide.html

[6] Eastaway R. et Wyndham J., Pourquoi les bus arrivent-il toujours par trois ? Flammarion, 2001, pp. 95-107

[7] Hill Lester S., « Cryptography in an Algebraic Alphabet », American Mathematical Monthly, 36, 1929, pp. 306-312

[8] Lewand Robert Edward, Cryptological Mathematics, published by The Mathematical Association of America, 2000, pp. 124-140

[9] Stinson Douglas, Cryptographie, Théorie et pratique, Vuibert, 2001, pp. 12-16