CHAPITRE 2

CRYPTOGRAPHIE CLASSIQUE

2.1 Introduction

La cryptographie classique décrit la période avant les ordinateurs. Elle traite des systèmes reposant sur les lettres et les caractères d'une langue naturelle. Les principales opérations sur lesquelles se base la cryptographie classique sont :

- Chiffrement par substitution
- Chiffrement par transposition

2.2 Chiffrement par substitution

Un chiffrement par substitution est un algorithme par lequel chaque caractère du message en clair (écrit dans un alphabet donné) est substitué (remplacé) par un autre caractère dans le message chiffré (qui peut être écrit dans un alphabet différent de celui du message clair) selon une règle convenue. En cryptographie classique, trois types de chiffrement par substitution sont distingués :

- **Substitution mono-alphabétique :** Un caractère du message clair est substitué par un caractère unique du message chiffré ;
- **Substitution poly-alphabétique :** Un caractère du message clair correspond à plusieurs caractères du message chiffré. Le principe est qu'à chaque caractère de l'alphabet des messages clairs est associé une liste de lettres dans l'alphabet des messages chiffrés, l'ensemble de ces listes formant une partition de l'alphabet des messages chiffrés.
- **Substitution poly-gramique :** Le principe est de substituer des blocs de caractères (deux ou trois généralement), au lieu d'un seul caractère. Par exemple, dans une substitution bigrammique, deux lettres du texte clair sont transformées en deux lettres du cryptogramme.

2.2.1 Chiffrement de César

Le chiffrement utilisé par Jules César est la méthode de cryptographie la plus ancienne communément admise par l'histoire. Il consiste en une substitution mono-alphabétique. Chaque lettre est remplacée ("substitution") par une seule autre, selon un certain décalage dans l'alphabet. Il consiste à décaler les lettres de l'alphabet d'un certain rang k, qui représente le nombre de lettres à décaler. Le rang k est alors considéré comme la clé de chiffrement. Le chiffrement de César peut également être défini par une fonction E_k de

l'ensemble $Z/26$.

Pour chaque lettre L du texte en clair qui subit un décalage de k lettres :

$$E_k(L) = (L + k) mod\ 26$$

À l'inverse, pour chaque lettre C du texte chiffré résultant d'un décalage de k lettres :

$$D_k(C) = (C - k) mod\ 26$$

$D_k(C)$ est la fonction de déchiffrement correspondant à E_k et le calcul commence par 0, c'est-à-dire; la lettre A=0.

2.2.1.1 Exemple

Chiffrement du message en clair "Bonjour" pour k= 3.

Texte en clair	B	O	N	J	O	U	R
Clé	3	3	3	3	3	3	3
(L+K=)mod 26	4	17	16	12	17	22	20
Texte chiffré	E	R	Q	M	R	X	U

2.2.1.2 Cryptanalyse

Le chiffre de César peut être cassé très facilement, même à l'aide du seul texte chiffré. On peut distinguer deux cas :

- **Par recherche de la valeur du décalage : Force brute :** Comme il n'y a qu'un nombre limité de décalages (vingt-six dont un inutile), il suffit de tester tous les chiffrements possibles jusqu'à trouver le bon. C'est ce qu'on appelle une attaque par force brute, technique de test de toutes les combinaisons possibles.

- **Par Analyse fréquentielle :** Méthode d'Al-Kinidi

  L'analyse fréquentielle, est une méthode de cryptanalyse dont la description la plus ancienne est réalisée par Al-Kindi au ixe siècle. Elle consiste à examiner la fréquence des lettres employées dans un message chiffré. L'analyse fréquentielle est basée sur le fait que, dans chaque langue, certaines lettres ou combinaisons de lettres apparaissent avec une certaine fréquence. Par exemple, en français, le e est la lettre la plus utilisée, suivie du a et du s. Inversement, le w est peu utilisé.

  Ces informations permettent aux cryptanalystes de mettre en relation la fréquence des lettres du message codé avec ces statistiques. Ainsi, il peut alors connaitre la plupart des lettres du message, mais pas toutes (certaines ayant des fréquences trop similaires). Cependant, la découverte des lettres principales permet de percer le reste du message. Cette technique nécessite en outre que le texte ait une longueur suffisante et implique que le cryptanalyste attaquant connaisse la langue d'origine du message crypté.

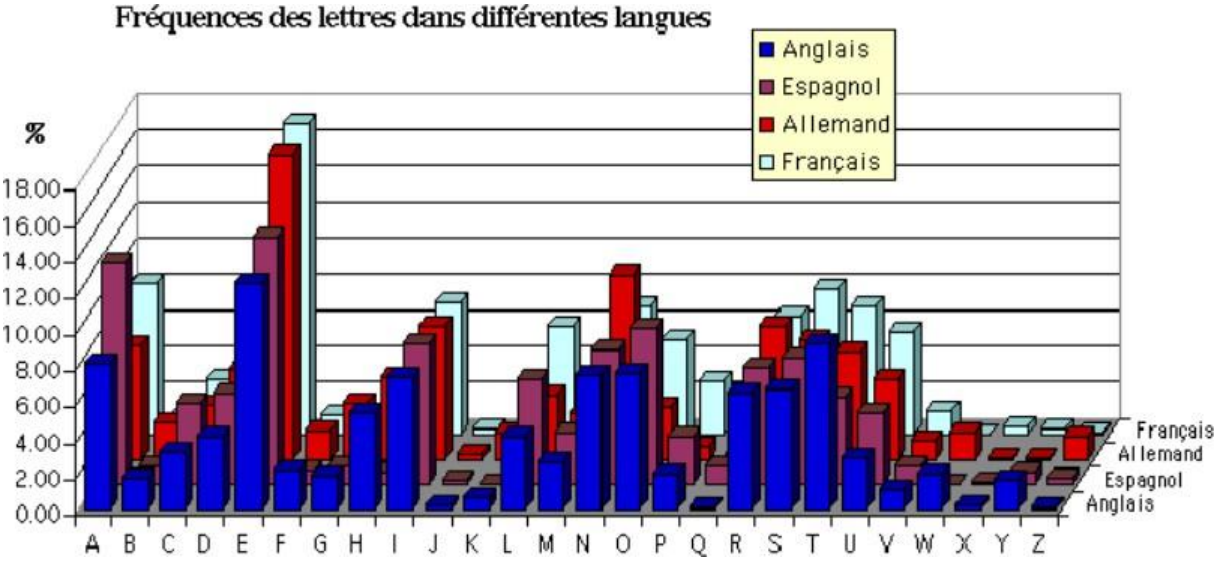

Figure 2.1: Fréquence d'utilisation des lettres.

2.2.1 Chiffrement d'Affine

2.2.2.1 Principe

Le chiffrement affine est une méthode de chiffrement par substitution mono-alphabétique. Il s'agit d'une version améliorée du chiffre de césar. Ce chiffrement est réaliser à l'aide d'une fonction appelée Fonction d'affine :

$$F(x) = a * x + b$$

On choisit deux entiers a et b, de tel sorte : (a,b) $\in$ [0.25] et PGCD (a,26) =1. Le couple (a, b) représente la clé de chiffrement.

2.2.2.1 Chiffrement

$C = (a * L + b) \ mod\, 26$ On remarque que l'on retrouve le chiffre de césar dans le cas ou $a = 1$

2.2.2.2 Déchiffrement

Pour que l'on puisse déchiffrer, il faut que l'opération inverse soit possible. C'est à dire que l'équation $y = a * x + b mod\, 26$ ait une solution unique (cela revient à dire que la fonction f doit être injective.

Théorème : l'équation $y = a * x + b mod\, 26$ admet une solution unique $x \in Z/26$ pour tout $b \in Z/26$ si $pgcd(a, 26) = 1$. On a donc 12 possibilités pour choisir a (1,3,5,7,9,11,15,17,19,21,23,25). Par contre b doit être quelconque. On a donc 12*26 clés possible pour le chiffrement d'affine. Considérons maintenant la fonction f du chiffrement et m=26. On suppose que le pgcd(a,26) =1. L'équation $y = a * x + b mod\, 26$ admet alors une solution. Pour la déterminer il faut faire appel à la notion d'inverse modulaire définie par :

Soit $a \in Z/Z\,26$ l'inverse de a est un élément $a^{-1} \in Z/Z\,26$ tel que : $a * a^{-1} = 1 mod\, 26$.

La fonction de déchiffrement est :

$$L = a^{-1} * (C - b) \ mod\, 26$$

2.2.2.3 Application

Pour une clé de chiffrement $a = 7$ et $b = 3$. Chiffrer le message en clair M="BONJOUR"

Texte en clair	B	O	N	J	O	U	R
Rang L	1	14	13	9	14	20	17
C=a*L+b	10	23	16	14	10	13	18
Texte chiffré	K	X	Q	O	X	N	S

2.2.2.4 Cryptanalyse

On peut se servir de la méthode d'Al-Kindi, pour établir la fréquence relative à chaque lettre Chiffrée, puis identifier les chiffres des deux lettres les plus fréquentes. Enfin, on résout le système d'équations à deux inconnus :

$$F(L1) = C1 = (a * L1 + b) mod\ 26$$
$$F(L2) = C2 = (a * L2 + b) mod\ 26$$

2.2.3 Chiffrement de Vigenère

C'est un système de chiffrement par substitution poly-alphabétique. La grande force du chiffre de Vigenère est que la même lettre sera chiffrée de différentes manières (une même lettre du message clair peut, suivant sa position dans celui-ci, être remplacée par des lettres différentes) Il consiste à utiliser 26 alphabets décalés pour chiffrer un message. Les 26 alphabets décalés sont représentés dans ce qu'on appelle un carré de Vigenère. Ce chiffre utilise une clé qui définit le décalage pour chaque lettre du message. Une clé se présente généralement sous la forme d'un mot ou d'une phrase.

2.2.3.1 Application de la table de Vigènère

- **Principe de chiffrement :** A chaque lettre en clair, on sélectionne la colonne correspondante tandis que la lettre de la clé se sélectionne par ligne, au croisement de la ligne et de la colonne on trouve la lettre chiffrée.
- **Principe de Déchiffrement :** On regarde pour chaque lettre de la clé répétée, la ligne correspondante sur laquelle on cherche la lettre chiffrée. La première lettre de la colonne que l'on trouve ainsi est la lettre déchiffrée

2.2.3.2 Formellement

Les opérations de chiffrement et de déchiffrement sont, pour chaque lettre, celles du chiffre de César. En désignant la i^e lettre du texte clair par Texte[i], la i^e du chiffré par Chiffré[i], et la i^e lettre de la clé, répétée suffisamment de fois, par Clés[i], elle se formalise par :

$$Chiffré[i] = (Texte[i] + Clés[i]) mod\ 26$$
$$Texte[i] = (Chiffré[i] - Clés[i]) mod\ 26$$

Pour le chiffrement, il suffit d'effectuer l'addition des deux lettres puis de soustraire 26 si le résultat dépasse 26. Pour le déchiffrement il suffit d'effectuer la soustraction et d'additionner 26 si le résultat est négatif.

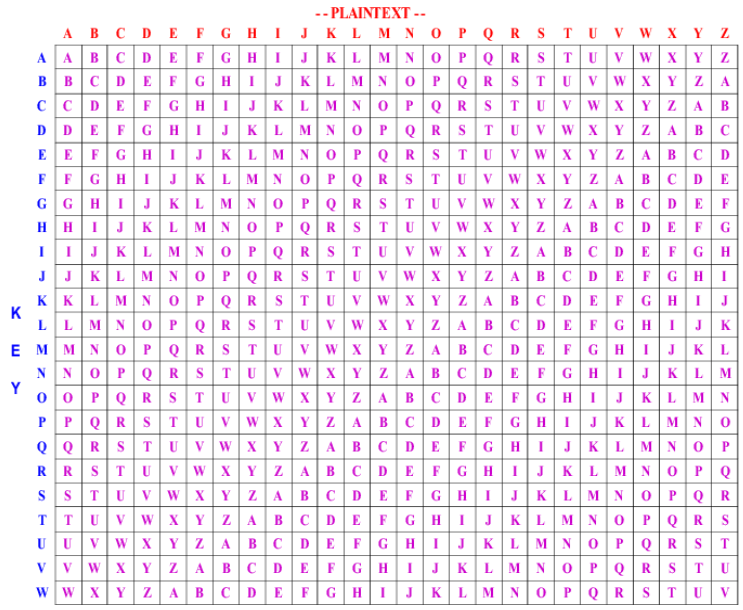

Figure 2.2: Carré de Vigenère.

2.2.3.3 Application

Chiffrement du message en clair "Bonjour" pour le mot clé "mot".

Texte en clair	B	O	N	J	O	U	R
Clé	M	O	T	M	O	t	M
$(L[i]+Cl\acute{e}[i])mod\,26$	13	2	6	21	2	13	3
Texte chiffré	N	C	G	V	C	N	D

2.2.4 Chiffrement de Playfair

Le chiffrement Playfair est un chiffrement poly-grammique. Dans le chiffrement de Playfair, on dispose les 25 lettres de l'alphabet (W exclu, car inutile, on utilise V à la place) dans une grille de 5x5. La variante anglaise consiste à garder le W et à fusionner I et J. On remplit la matrice par les lettres du mot clé. Si une lettre se répète, on l'écrit qu'une seule fois. Puis on complète par les lettres restantes de l'alphabet.

2.2.4.1 Chiffrement

Le chiffrement se fait par groupes de deux lettres (des bigrammes) en appliquant les règles suivantes :

- Si les deux lettres sont sur les coins d'un rectangle, alors les lettres chiffrées sont sur les deux autres coins. Autrement dit, chaque groupe de deux lettres est codé par la lettre à l'intersection de la ligne de la première et la colonne de la deuxième. Puis à l'intersection de la ligne de la deuxième et de la colonne de la première ;
- Si deux lettres sont sur la même ligne, on prend les deux lettres qui les suivent immédiatement à leur droite ;
- Si deux lettres sont sur la même colonne, on prend les deux lettres qui les suivent immédiatement en dessous.

- Si le bigramme est composé de deux fois la même lettre, on insère une nulle (usuellement le X) entre les deux pour éliminer ce doublon ; S'il y a une seule lettre, on complète par la lettre X.

2.2.4.2 Déchiffrement

Pour le déchiffrement, on applique les règles de chiffrement à l'envers :

- Si les deux lettres sont sur les coins d'un rectangle, Chaque groupe de deux lettres est codé par la lettre à l'intersection de la ligne de la première et la colonne de la deuxième puis à l'intersection de la ligne de la deuxième et de la colonne de la première ;
- Si deux lettres sont sur la même ligne, on prend les deux lettres qui les suivent immédiatement à leur gauche ;
- Si deux lettres sont sur la même colonne, on prend les deux lettres qui les suivent immédiatement en dessus.

2.2.4.3 Application

Nous avoir le message en clair "BONJOUR" à chiffrer avec le mot clé "SALUT"

S	A	L	U	T
B	C	D	E	F
G	H	I	J	K
M	N	O	P	Q
R	V	X	Y	Z

Le cryptogramme C= "DMPHPLVY

2.2.5 Chiffrement de Hill

Il consiste à chiffrer le message en substituant les lettres du message, non plus lettre par lettre, mais par groupe de lettres. Il permet ainsi de rendre plus difficile de le casser par observation des fréquences. C'est un chiffrement à base de l'algèbre matricielle, la substitution se fait à l'aide de *m* équations linéaires. L'algorithme remplace *m* lettre successives par m lettre chiffrées.

2.2.5.1 Chiffrement

Les lettres sont remplacées par leur rang suivant l'alphabet. On choisit une clé *k* sous forme d'une matrice de 2X2 telle que $PGCD\ (det\ (k), 26) = 1$ Chaque paire de lettres L_k et L_{k+1} du message en clair sont remplacées par C_k et C_{k+1} :

$$\begin{pmatrix} C_k \\ C_{k+1} \end{pmatrix} = k * \begin{pmatrix} L_k \\ L_{k+1} \end{pmatrix} \bmod 26 = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, * \begin{pmatrix} L_k \\ L_{k+1} \end{pmatrix} \bmod 26$$

Remarque : s'il y a une seule lettre, on remplace par la lettre "X".

2.2.5.2 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. Cette matrice doit être l'inverse de matrice de chiffrement (modulo 26).

$$\begin{pmatrix} L_k \\ L_{k+1} \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} * \begin{pmatrix} C_k \\ C_{k+1} \end{pmatrix} \bmod 26$$

$$\begin{pmatrix} L_k \\ L_{k+1} \end{pmatrix} = dek(k)^{-1} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} * \begin{pmatrix} C_k \\ C_{k+1} \end{pmatrix} \bmod 26$$

Pour calculer $dek(k)^{-1}$, on utilise l'inverse modulaire. En effet, pour calculer l'inverse de *bmodn*. Rappelons que l'inverse mod n de b est le nombre entier $(b)^{-1}$ tel que $b * (b)^{-1}(\text{mod n}) = 1$. Donc $dek(k)^{-1}$ est calculé tel que $dek(k)^{-1} * dek(k) \bmod 26 = 1$. Pour cela, on utilise algorithme d'Euclide étendu (les équations diophantiennes). C'est-à-dire : $dek(k)^{-1} * dek(k) * \bmod 26 = 1$ est équivalente à : $dek(k)^{-1} * dek(k)\det(K) + y * 26 = 1$

2.2.5.3 Outil mathématique nécessaire au déchiffrement

Description de l'algorithme d'Euclide étendu (les équations diophantiennes)

- On commence par descendre avec l'algorithme d'Euclide pour le PGCD, en notant les quotients dans la colonne de gauche et les restes successifs dans la seconde colonne. Pour deux éléments successifs descendants, on calcule le troisième par division : a divisé par b = q et reste c.
- Le processus s'arrête lorsqu'on obtient un reste nul (0), qui permet d'écrire dans la colonne 3, en face des deux derniers nombres obtenus, les nombres 0 et 1, dont les produits croisés avec la deuxième colonne ont une différence de 1. On peut les lire comme des déterminants valant alternativement 1 et -1.
- On remonte ensuite en construisant la troisième colonne avec l'égalité caractéristique de la division euclidienne : $x' = y' * q + z'$, parallèlement à la première colonne. On note des traits obliques alternés pour indiquer dans quel sens la soustraction donne 1.
- A chaque ligne, on a par construction une solution d'une équation de la forme $ay - bx = +/-1$, avec une alternance des signes. La ligne d'en haut fournit ainsi la solution de l'équation diophantienne initiale, les lignes suivantes celles d'équations diophantiennes réduites.

2.2.5.4 Application : soit l'équation suivante : $95x + 14y = 1$

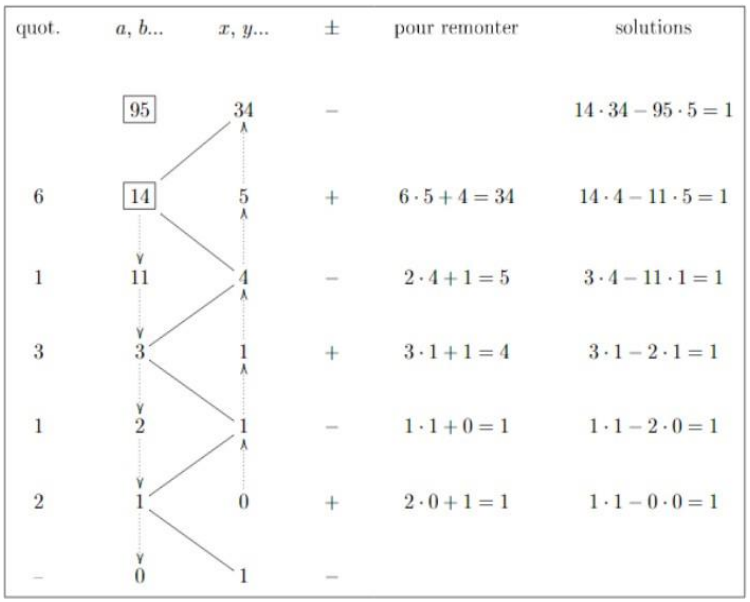

Figure 2.3: exemple de la division euclidienne.

2.3 Chiffrement par transposition

Le chiffrement par transposition consiste à réarranger les lettres suivant un ordre différent de façon à les rendre incompréhensibles, c'est-à-dire ; construire des anagrammes. Les lettres du texte ne sont pas altérées, seul l'ordre des lettres est changé de façon à aboutir à un mélange sans cohérence. Chaque lettre reste inchangée, mais est mise à une autre place.

2.3.1 Transposition simple par colonnes

Dans ce type de chiffrement, on dispose les lettres du message dans un tableau de n colonnes. Pour chiffrer un message, on va disposer les lettres du message horizontalement sur la matrice de langueur n, puis on collecte les lettres verticalement. Pour un message de langueur m alors on aura deux cas de figure :

- $m \bmod n = 0$: dans ce cas, toutes les colonnes ont la même hauteur $[m/n]$.
- $m \bmod n = i$, tel que $i > 0$: dans ce cas, la hauteur des i premières colonnes est $[m/n] + 1$. La hauteur des $n - i$ colonnes restantes est $[m/n]$.

Pour le procède de déchiffrement, on dispose les lettres du chiffre verticalement, puis on collecte les lettres horizontalement. La construction de la matrice se fait de la même manière.

2.3.1.1 Application

Nous avons le message en clair M = "Transposition simple" à chiffrer sur une matrice de longueur $n = 5$. Le nombre de lettres du message en clair est 19. Nous Calculons la hauteur des colonnes comme suit : nous 19 mod 5 = 4, donc, les 4 premières colonnes auront la hauteur de [19/5] + 1 = 4, et la hauteur de la dernière colonne est [19/5] = 3.

Matrice :

T	R	A	N	S
P	O	S	I	T
I	O	N	S	I
M	P	L	E	

Le cryptogramme est : C ="TPIMROOPASNLNISESTI ".

2.3.2 Transposition complexe par colonnes

Une transposition complexe par colonnes s'effectue à partir d'une clé (mot ou expression) de la longueur souhaitée. On numérote ensuite les lettres dans l'ordre alphabétique. Si une même lettre apparaît plusieurs fois, elle est numérotée successivement de la gauche vers la droite.

Pour le procédé de chiffrement ; la clé du message donne à la fois le nombre de colonnes du tableau, et l'ordre de la récolte des lettres. De la même manière que la transposition simple, on dispose horizontalement les lettres du message sur la matrice de langueur équivalente à la longueur de la clé. On collecte verticalement les lettres suivant l'ordre croissant des lettres de la clé par rapport à l'alphabet. On effectue l'opération inverse pour déchiffrer le message : On dispose verticalement les lettres du message chiffre suivant l'ordre croissant des lettres du mot clé par rapport à l'alphabet ; Collecter horizontalement les lettres.

2.3.2.1 Application

Nous avons le message en claire M = "Transposition complexe" à chiffrer avec le mot clé « Ligne ». Nous avons le nombre de lettres du message en clair est 21. La hauteur des colonnes est calculée comme suit :

Nous avons 21 mod 5 = 1, donc, la première colonne aura la hauteur de [21/5] + 1 = 5, et la hauteur des quatre dernières colonnes est [21/5] = 4. Matrice :

L	I	G	N	E
4	3	2	5	1
T	R	A	N	S
P	O	S	I	T
I	O	N	C	O
M	P	L	E	X
E				

Le cryptogramme obtenu est : C =" STOXASNLROOPTPIMENICE"

2.4 Conclusion

La cryptographie classique permet de chiffrer des textes seulement, dans ce chapitre nous avant présenté certaines méthodes les plus connues. La cryptographie classique a assuré la confidentialité des messages pendant des siècles. Néanmoins, avec l'avènement de l'Informatique, ces méthodes ne sont pas efficaces face à cette technologie. Dans le chapitre suivant, nous allons présenter la cryptographie moderne.