# Introduction à la cryptographie et au chiffrement par substitution mono-alphabétique

par

Date de publication : 01/02/08

Dernière mise à jour :

La cryptographie, ou l'art de brouiller le sens d'un message aux indiscrets, est une science unique à l'histoire étroitement liée à celle de l'humanité. Ce tutoriel est une modeste introduction à cette discipline.

I - Introduction
II - La cryptographie par substitution mono-alphabétique
III - Cryptanalyse, ou comment casser le chiffre de César.
IV - Conclusion
V - Aller plus loin
VI - Remerciements

# I - Introduction

Depuis l'apparition de l'écriture, la communication entre les hommes a pris une nouvelle forme, simplifiant la persistance des connaissances et les échanges.

Plus tard, au fil de l'évolution humaine, l'écriture prit de plus en plus d'importance, jusqu'à devenir vitale, par exemple, en cas de guerre.

De tout temps, les chefs de guerre et diplomates avaient besoins de transmettre des informations stratégiques à leurs troupes ou leurs alliés. Cependant, il suffisait à l'ennemi d'intercepter le message pour obtenir un avantage déterminant.

Ceci n'est qu'un exemple où la cryptographie prend tout son sens.

La cryptographie est une discipline qui a pour but d'assurer la confidentialité, intégrité, et authenticité de messages. Elle fait partie, avec la cryptanalyse, de la cryptologie, ou « science des secrets ».

Le principe de la cryptographie est de « brouiller » les messages à l'aide d'une « clé » afin de protéger l'information d'un indiscret, et que seul le destinataire légitime puisse accéder au texte original.

Il ne faut pas confondre cryptographie et stéganographie, discipline qui consiste à cacher un message dans un lot d'informations, mais en aucun cas de « brouiller » l'information importante. On peut néanmoins mêler cryptographie et stéganographie.

## II - La cryptographie par substitution mono-alphabétique

C'est une ancienne forme de cryptographie apparue à l'antiquité, autrement appelée « Chiffre de César » du fait de son utilisation par l'empereur romain du même nom.

Elle consiste à une simple permutation des lettres de l'alphabet courant. On associe une lettre de l'alphabet « clair » à une autre lettre qui devient la lettre chiffrée. On obtient ainsi un alphabet chiffré.

Soit l'alphabet que l'on utilise tous les jours : a b c d e f g h i j k l m n o p q r s t u v w x y z. Décalons le tout de 2 lettres vers la droite.

On obtient ainsi le tableau suivant :

| a | b | c | d | e | f | g | h | i | j | k | l | m | n | o | p | q | r | s | t | u | v | w | x | y | z |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| C | D | E | F | G | H | U | J | K | L | M | N | O | P | Q | R | S | T | U | V | W | X | Y | Z | A | B |

Généralement, le texte «clair » (non chiffré, le message original donc) est en minuscule tandis que le message chiffré est en majuscule.

Dans l'exemple précédent, il n'y a qu'un simple décalage. On peut très bien imaginer un mélange complet de l'alphabet :

| a | b | c | d | e | f | g | h | i | j | k | l | m | n | o | p | q | r | s | t | u | v | w | x | y | z |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| I | Y | H | G | R | F | D | E | Z | S | Q | A | W | N | K | L | O | M | P | B | V | J | U | T | X | C |

Exemple d'utilisation : cryptons le message suivant « Attaquez les positions ennemies ! » avec l'alphabet suivant :

| a | b | c | d | e | f | g | h | i | j | k | l | m | n | o | p | q | r | s | t | u | v | w | x | y | z |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| I | Y | H | G | R | F | D | E | Z | S | Q | A | W | N | K | L | O | M | P | B | V | J | U | T | X | C |

Tout d'abord, il convient de supprimer les ponctuations et espaces, puis de mettre l'ensemble du message en minuscule. Après l'avoir fait, nous obtenons : « attaquezlespositionsennemies ».

Remplaçons simplement chaque lettre par son équivalent dans le tableau. Les 'a' deviennent 'I', les 't' deviennent 'B', ainsi de suite.

Après l'encryptage, vous obtenez le message crypté : « IBBIOVRCARPLKPZBZKNPRNNRWZRP ».

Pour décrypter, le destinataire, munis de la clé, n'a qu'à reproduire l'opération en sens opposé, prenant tour à tour la lettre cryptée pour trouver la lettre claire correspondante.

## III - Cryptanalyse, ou comment casser le chiffre de César.

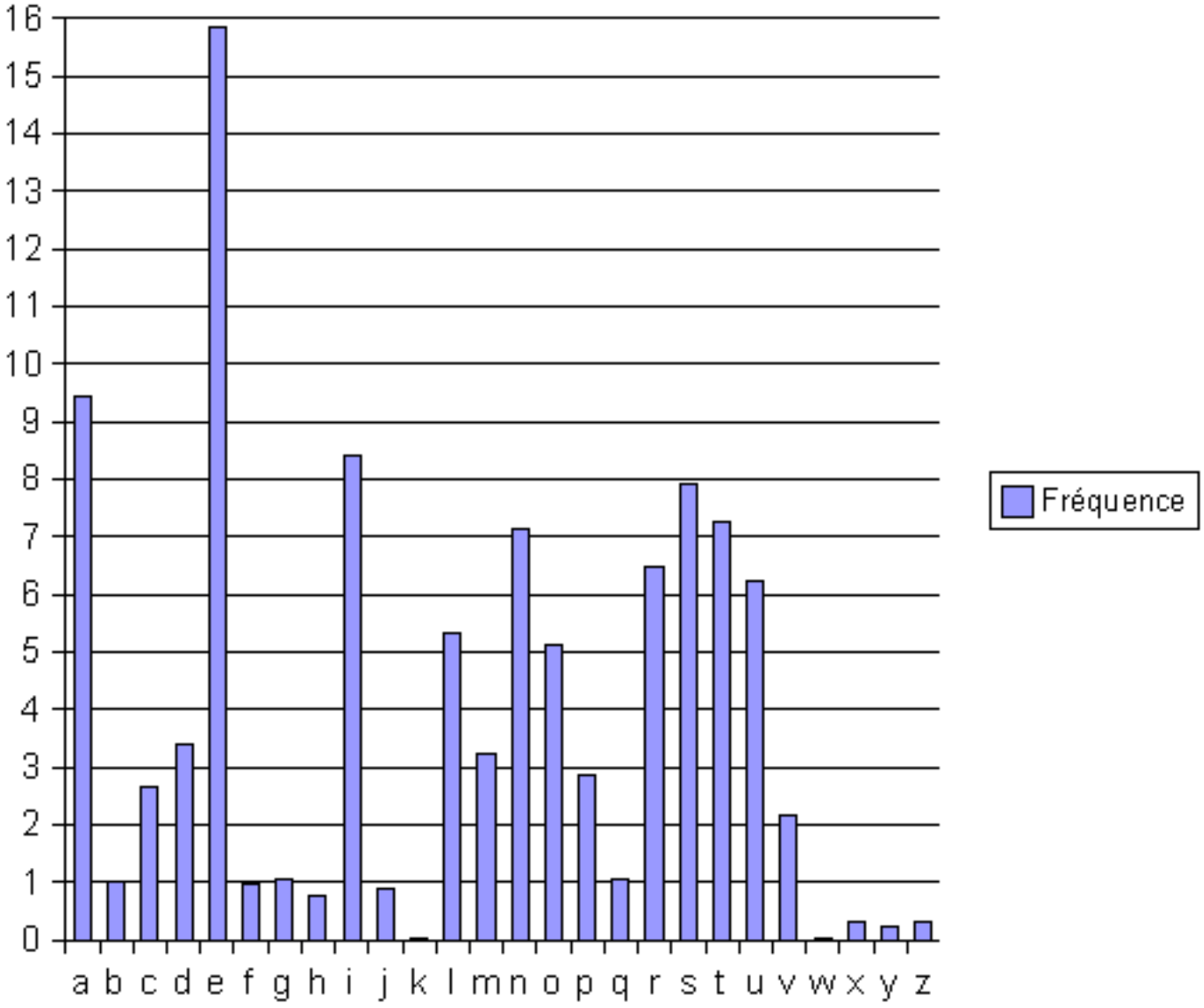
Le « Chiffre de César » est très simple, rapide à mettre en oeuvre et il était relativement efficace pendant la période de l'antiquité.

Cependant, à l'apogée de la civilisation musulmane, les mathématiciens arabes trouvèrent un moyen de casser le cryptage des textes cryptés avec une substitution alphabétique.

Comment ont-il fait ? Ils ont utilisés une caractéristique propre à chaque langue : la fréquence d'utilisation de chaque lettre de l'alphabet.

Par exemple, pour la langue française, la lettre 'e' est de loin la plus utilisée, elle a une fréquence d'utilisation d'environ 15%, suit la lettre 'a' avec 9%, 'i' avec 8%, puis 's', 't', 'n' avec environ 7%.

Voici une table de fréquences pour l'alphabet français (les résultats sont tirés du livre « Histoire des codes secrets » de Simon Singh, que je vous encourage fortement à lire si la cryptographie vous intéresse) :




*Histogramme des fréquences.*

Ces résultats sont tirés d'un échantillonnage de 10 000 journaux ou romans rédigés en langue française. Ces données peuvent varier quelque peu d'une étude à l'autre, mais la marge de différence reste convenable.

C'est dans son « Manuscrit sur le déchiffrement des messages cryptographique » que le savant arabe « Al-Kindi » expliqua pour la première fois comment venir à bout d'un texte crypté.

Il démontra ainsi qu'en connaissant la langue originelle du message crypté, et en ayant une table de fréquences de cette langue, on pouvait associer à chaque symbole crypté une lettre.

Il faut tout d'abord compter le nombre d'apparition de chaque symbole différent dans le message crypté, puis de comparer les résultats avec la table de fréquences.

Si la lettre V apparaît le plus souvent de fois dans le texte, c'est sans doute la lettre 'e' de notre alphabet. Si la deuxième lettre qui apparaît le plus souvent est le O, nous pouvons penser que c'est en réalité un A. Et ainsi de suite.

La faiblesse de cette méthode réside dans le fait que la table de fréquence n'est qu'une moyenne. Or, sur un texte court, les résultats peuvent être complètement différents, et une tentative de transposition serait un échec.

## IV - Conclusion

Utilisé pendant des siècles par les chefs de guerre et diplomates, le chiffre de substitution mono-alphabétique, ou « Chiffre de César », s'est finalement révélé « faible » face aux mathématiciens arabes.

Malgré sa faiblesse, ce chiffre reste amusant et une de ses qualités est qu'il permet de crypter un message rapidement.

Il ne sera plus utilisé par les gouvernements bien sûr, mais rien ne vous empêche de l'utiliser pour cacher vos petits secrets des indiscrets.

La cryptographie est un monde merveilleux où se rencontrent toutes sortes de disciplines : les mathématiques bien sûr, mais aussi la linguistique par exemple.

L'utilisation de la cryptographie encourage à penser vite, à être astucieux. Ce n'est finalement qu'un perpétuel combat entre cryptographes et cryptanalistes.

## V - Aller plus loin

Ce tutoriel n'est qu'une courte introduction au monde très vaste de la cryptographie. Vous avez maintenant une petite idée à quoi il ressemble.

Si la cryptologie vous intéresse, vous intrigue, je vous conseille fortement de lire le livre « Histoire des codes secrets » de Simon Singh.

Ce livre est tout simplement incontournable, présentant de façon très ludique la cryptographie et son histoire, il retrace l'évolution des codes secrets au cours de l'histoire.

C'est un point de passage je dirais presque obligatoire pour toute personne souhaitant s'initier à la cryptographie.

"Autres ressources, liste tutoriels"

# VI - Remerciements