Cryptographie à clé secrète

9 juin 2023

Sommaire

Rappel : but de la cryptographie

L'objectif principal de la cryptographie c'est d' obtenir des communications sécurisées sur des canaux non sûrs. Cet objectif est classiquement divisé en 4 services pour la sécurité.

1. confidentialité,
2. intégrité des données,
3. authentification (origine des entités et des données),
4. non repudiation (signature numérique).

Definition

Un cryptosystème est un terme général qui se référe à un ensemble de primitives cryptographiques utilisées pour assurer les services de la sécurité informatique. Ces routines (les primitives cryptographiques) concernent entre autres les fonctions de chiffrement.

Formalisme des schémas de chiffrement

Terminologie

- $\mathcal{A}$ dénote un ensemble fini appelé **alphabet de définition**. $\mathcal{A}^*$ dénote l'ensemble des mots sur $\mathcal{A}$.
- $\mathcal{P}$ ou $\mathcal{M}$ dénote un ensemble appelé **espace des messages**. $\mathcal{P} \subset \mathcal{A}^*$. Un élément de $\mathcal{P}$ est appelé **texte clair**.
- $\mathcal{C}$ dénote un ensemble appelé **espace des chiffrés**. Tout comme $\mathcal{P}$, $\mathcal{C} \subset \mathcal{A}'^*$ pour un alphabet $\mathcal{A}'$. Un élément de $\mathcal{C}$ est appelé **texte chiffré**.
- $\mathcal{K}$ dénote un ensemble appelé **espace des clés**.
- Les éléments (e, d) de $\mathcal{K} \times \mathcal{K}$ pour lesquels il existe deux fonctions

$$E_e : \mathcal{P} \longrightarrow \mathcal{C} \text{ et } D_d : \mathcal{C} \longrightarrow \mathcal{P}$$

telles que $D_d \circ E_e(m) = m \; \forall m \in \mathcal{P}$,(**avec une forte probabilité**) sont appelé **pair de clé**.

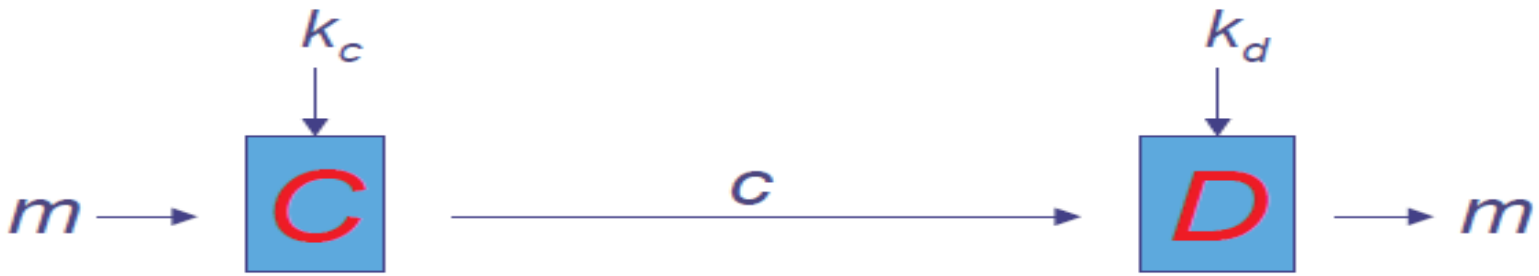
Formalisme des schémas de chiffrement

Pour une **pair de clés** $(e, d) = (k_c, k_d)$,

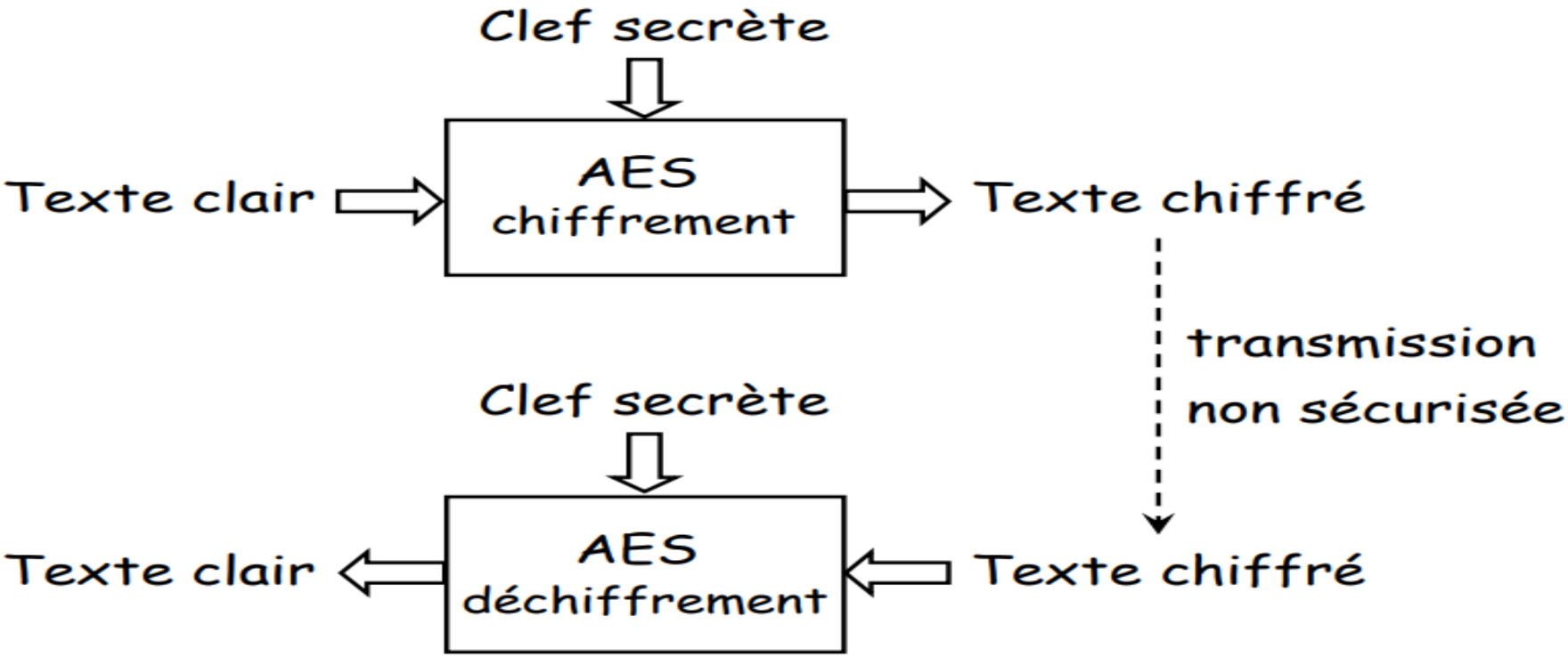
- la fonction E_e est injective (inversible à gauche) et est appelée la **fonction de chiffrement**
- la fonction D_e est appelée la **fonction de déchiffrement**

En pratique, pour une communication sécurisée entre Alice et Bob au niveau d'un canal non sécurisé

- l'expéditeur (Alice) chiffre un texte clair m en calculant le texte chiffré $y = E_e(m)$, et
- le destinataire (Bob) décrypte le texte chiffré y en calculant le texte clair $m = D_d(y) = D_d \circ E_e(m)$ (**avec une forte probabilité**)



Exemple standard de cryptosystème : AES



Nous verrons le système AES, en détails, dans le prochain chapitre.

Notions de cryptanalyse

Objectif de la cryptanalyse

La cryptanalyse a pour principal objectif d'étudier les faiblesses des outils de sécurité produits par la cryptographie afin de les corriger ou nuire au système de communication.

Définition

Attaquant : Entité susceptible d'agir sur un schémas de communication dans le but :

1. de violer la confidentialité des données,
2. de détourner et de modifier des données,
3. d'usurper l'identitité de l'un des partenaires de la communication.

Types d'attaques

Il y'a deux classes d'attaques :

1. **Attaques passives** : l'attaquant écoute seulement. Donc, c'est une attaque qui cible la confidentialité.

2. **Attaques actives** : l'attaquant agit concrètement sur le système de commnucation. Il est susceptible de détourner, modifier des données ou récupérer des informations. Les attaques actives ciblent toutes les fonctions de sécurité : confidentialité, intégrité, authentification-identification et non répudiation.

Cible des attaques

Ces attaques peuvent être dirigées :

- sur les modèles mathématiques qu'utilise le cryptosystème ;
- sur l'implémentation matériel et/ou logiciel ;
- sur les entités légitimes d'un scénarios de communication ;
- sur la gestion (distribution, stockage, tests de validité, ...) de données ;

Définition (Cryptosystème sûr)

Un système de chiffrement est dit sûr si la probabilité d'obtenir une information sur le texte clair ou la clé de déchiffrement à partir du chiffré est presque nulle.

Taille des données secrètes en cryptographie

Problème crucial

on travaille avec des ensembles finis et publiques dans lesquels on doit chosir des valeurs secrètes.

En cryptographie, la première faiblesse qu'il faut prendre en compte c'est le problème de la taille des données. En effet il représente le premier niveau de sécurité. Par exemple sans cette mesure l'attaque par force brute qui consiste à tester toutes les valeurs possibles pourra découvrir le secret.

Ainsi il est necessaire d'avoir une idée assez précise sur la puissance de calcul des machines.

Puissance de calcul des machines

Un PC à $1GHz = 10^9 Hz$ effectue 10^9 opérations élémentaires (affectation, instructions de contrôle, calcul binaire,...) par seconde.

Temps	Nbre operations / 1 PC	Nbre operat / 10^{18} PCs
1s	10^9	10^{27}
1 an	$3,1.10^{16}$	$3,1.10^{34}$
1000 ans	$3,1.10^{19}$	$3,1.10^{37}$
10^9 ans	...	...
15.10^9 ans	$46,5.10^{25}$	$46,5.10^{43}$

Idée sur la puissance de calcul des machines

La vitesse de la lumière est de $3 \times 10^8 m/s$ (environ). Comparer le temps que mettra cette lumière pour traverser une pièce de 3 métres de largeur et le nombre d'opérations élémentaires qu'effectue un PC de $1GHz$. Qu'en est-il pour un PC de $2GHz$.

Puissance de calcul des machines

Pratique : voir la série d'exercices.

Sécurité des mots de passe

Mots de passe de 8 caractères

Alphabet	Alphabet et chiffre	Alphanumerique
$26^8 \sim 2^{38}$	$36^8 \sim 2^{42}$	$256^8 \sim 2^{64}$

Mots de passe de 22 caractères

Alphabet	Alphabet et chiffre	Alphanumerique
$26^{22} \sim 2^{104}$	$36^{22} \sim 2^{118}$	$256^{22} \sim 2^{176}$

En résumé

- Un mot de passe qui protège certaines données doit au moins être de 8 caractères alphanumériques.
- Un mot de passe qui joue le role d'une clé de chiffrement doit au moins être de 22 caractères alphanumériques.

Ces deux points sont en fonction des tableaux établis précédemment.

La communication symétrique

Cryptographie à clé symétrique

Definition

Pour une **paire de clé** (e, d), **quand** $e = d$ **ou** e **est facile à calculer à partir de** d **et inversement**, on dit que le processus de chiffrement est symétrique et dans ce cas E_e et D_d sont bijectives. **Pour la sécurité**, il est nécessaire de garder e et d secrètes.

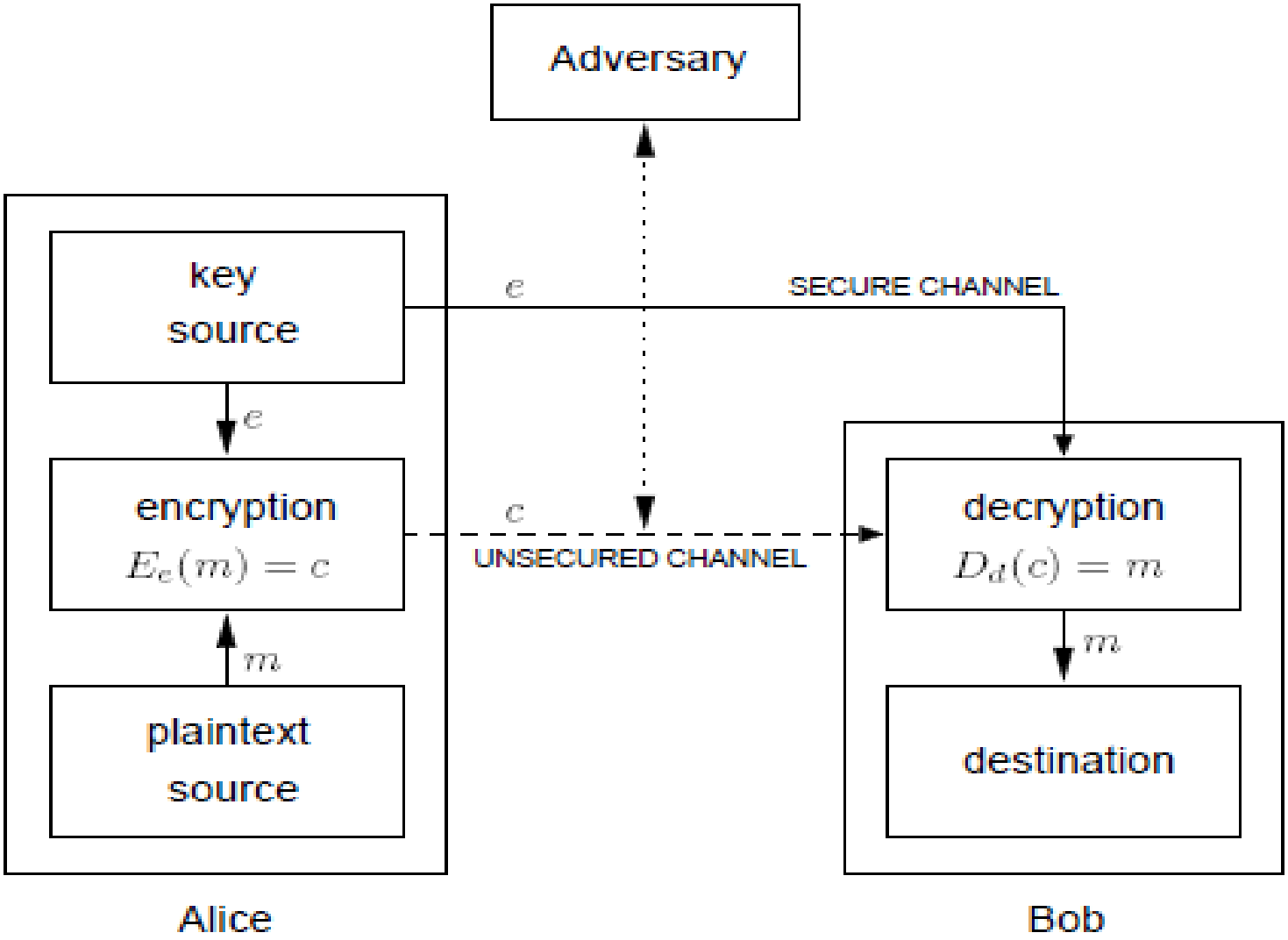
Le problème principal avec les systèmes à clé symétrique est de trouver une méthode efficace et sûre d'échange de clés. Ce problème est appelé **problème de distribution des clés.**

Hypothèse de sécurité

Pour évaluer la sécurité d'un chiffrement, il est usuel d'assumer qu'un adversaire

1. a accès à toutes les données transmises sur un canal à texte chiffré et
2. (**principe de Kerckhoff**) connait tous les détails sur la fonction de chiffrement mis à part la clé secrète.

Une communication (deux-à-deux) à travers un système de chiffrement symétrique peut être décrite par la figure suivante.



Exemples de chiffrement symétrique

Algorithme de chiffrement affine

On prend $\mathcal{P} = \mathcal{C} = \frac{\mathbb{Z}}{26\mathbb{Z}}$ et $\mathcal{K} = \left(\frac{\mathbb{Z}}{26\mathbb{Z}}\right)^* \times \frac{\mathbb{Z}}{26\mathbb{Z}}$; $k = (a, b)$ et

$$C_k(x) = ax + b$$

comme fonction de chiffrement.

A : Algorithme de génération de clés

Pour créer une clé secrète, Bob fait ce qui suit :

(1) Sélectionne a copremier avec 26 et $b < 26$ puis calcule $a^{-1} \mod 26$;

(2) La clé secrète de Bob est (a, b) qu'il envoie à Alice par un canal sûr.

B : Algorithme de chiffrement

Pour envoyer un texte chiffré à Bob, Alice fait ce qui suit

(1) Transforme chaque lettre "x" en un entier < 26 : $a = 0, ..., z = 25$;

(2) Chiffre chaque lettre "x" du texte clair en "$y = ax + b \mod 26$" ;

C : Algorithme de déchiffrement

Bob déchiffre chaque lettre "y" par "$x = a^{-1}(y - b) \mod 26$"

Exemple de chiffrement affine

Exemple 1 : Chiffrement affine sur $\mathbb{Z}/n\mathbb{Z}$

- $e = (e_1, e_2)$ et $y = e_1 x + e_2 \mod n$ où y est le chiffré de x dans $\frac{\mathbb{Z}}{n\mathbb{Z}}$ avec la clé secrète $e = (e_1, e_2)$
- – prendre $n = 26$, chiffre le texte

$$M = \text{"cryptography"}$$

avec $e = (7, 23)$ dans C avec l'identification

$$a = 0, b = 1, \ldots, z = 25 \text{ dans } \mathbb{Z}/n\mathbb{Z}$$

– Calculer d et déchiffrer C.

Exemple de chiffrement affine

Exemple 2 : Chiffrement affine sur $M_m(\mathbb{Z}/n\mathbb{Z})$

$M_m(\mathbb{Z}/n\mathbb{Z})$ est l'ensemble des matrices carrées d'ordre m sur $\mathbb{Z}/n\mathbb{Z}$

- $e = (A, B)$ et $Y = AX + B \mod n$ où Y est le chiffré de X dans $M_m(\frac{\mathbb{Z}}{n\mathbb{Z}})$ avec la clé secrète $e = (A, B)$
- – Prendre $m = 2, n = 26$, chiffrer le texte $X =$ "tdsi" avec $e = (A, B)$ dans C en utilisant l'identification

$$a = 0, b = 1, \ldots, z = 25 \text{ dans } \mathbb{Z}/n\mathbb{Z}$$

$A = \begin{pmatrix} 13 & 12 \\ 2 & 3 \end{pmatrix}$ et $B = \begin{pmatrix} 3 & 12 \\ 21 & 23 \end{pmatrix}$

– Calculer d et déchiffrer C.

Chiffrement par transposition

Une classe bien connue de chiffrement à clé symétrique est le chiffrement par transposition, qui permute simplement les symboles dans un bloc.

Definition

On considère un schéma de chiffrement à clé symétrique de longueur de bloc égale à t. Soit $\mathcal{K}$ l'ensemble des permutations sur $1, 2, \cdots, t$. Pour tout $e \in \mathcal{K}$ on définit la fonction de chiffrement

$$E_e(m) = (m_{e(1)}, m_{e(2)}, \cdots, m_{e(t)}).$$

L'ensemble de toutes ces transformations est appelé chiffrement à transposition simple.

La clé de déchiffrement correspondante à e est l'inverse $d = e^{-1}$.

Chiffrement par substitution

Substitution monoalphabétique

Definition

Les chiffrements par substitution sont des chiffrements qui remplacent des symboles (ou des groupes de symboles) par d'autres symboles ou des groupes de symboles.

Chiffrement par substitution monoalphabétique :

- La clé secréte est une permutation aléatoire π de l'alphabet français.
- Un message $m = m_1, m_2, m_3$ (où chaque m_i est une lettre) se chiffre une lettre à la fois, pour produire le texte chiffré $c = c_1, c_2, c_3$ où $c_i = \pi(m_i)$.
- Le déchiffrement s'effectue en appliquant la permutation inverse.

Le nombre total de clé possible est 26!, ce qui est suffisamment large.

Exemple simple de chiffrement par substitution

Supposons que la clé e représente la permutation qui à une lettre donnée associe la lettre qui est à sa troisième position vers la droite, comme indiqué ci-dessous.

$$e = \begin{pmatrix} A & B & C & D & E & F & G & H & I & J & K & L & M & N & O & P & Q & R & S & T & U & V & W & X & Y & Z \\ D & E & F & G & H & I & J & K & L & M & N & O & P & Q & R & S & T & U & V & W & X & Y & Z & A & B & C \end{pmatrix}$$

Le message

"THIS CIPHER IS CERTAINLY NOT SECURE"

s'écrit en premier lieu comme suit

"THISC IPHER ISCER TAINL YNOTS ECURE"

puis chiffré comme suit

"WKLVF LSKHU LVFHU WDLQO BQRWV HFXUH"

Méthode de substitution simple : Problématique

Avec le chiffrement par substitution simple un adversaire qui intercepte un texte chiffré assez long (disons de longueur 500) peut facilement utiliser la connaissance statistique de l'alphabet de la langue utilisée (telle que la fréquence relative des lettres) pour déterminer la clé secréte.

Le tableau ci-dessus donne les fréquences des lettres de l'alphabet dans la langue française.

a	**0,0825**	n	0,0725
b	0,0125	o	0,0575
c	0,0325	p	0,0375
d	0,0425	q	0,0125
e	**0,1270**	r	0,0725
f	0,0223	**s**	**0,0825**
g	0,0202	t	0,0725
h	0,0609	u	0,0625
i	0,0697	v	0,0175
j	0,0150	w	0,0000
k	0,0077	x	0,0000
l	0,0403	y	0,0075
m	0,0241	z	0,0000

Cryptanalyse du chiffrement par substitution

Exercice

En se basant sur le tableau des fréquences des lettres essayer de décrypter le message (en anglais) chiffré ci dessous :

"WKLVF LSKHU LVFHU WDLQO BQRWV HFXUH"

On note qu'il existe une similarité de la fréquence de certaines lettres de l'alphabet français et l'anglais.

Chiffrement par substitution polyalphabétique

Definition

Un chiffrement par substitution polyalphabétique est un chiffrement à longueur de bloc t sur un alphabet $\mathcal{A}$ ayant les propriétés suivantes :

1. L'espace des clés $\mathcal{K}$ est constitué par l'ensemble des t permutations ordonnées $(p_1, \cdots, p_t)$, où chaque permutation est définie sur l'ensemble $\mathcal{A}$.
2. le chiffrement du message $m = (m_1 m_2 \cdots m_t)$ sous la clé $e = (p_1, \cdots, p_t)$ est donné par $E_e(m) = (p_1(m_1) \cdots p_t(m_t))$
3. la clé de déchiffrement associée à e est $d = (p_1^{-1}, \cdots, p_t^{-1})$

Exemple : Le chiffrement de Vigenère

Prendre $t = 3$ et $e = (p_1, p_2, p_3)$ où

- p_1 associe chaque lettre à la troisième lettre située à sa droite sur l'alphabet.
- p_2 à celle qui est à la position un-septième vers sa droite, et
- p_3 à celle qui est à la position 10 vers sa droite.

Exercice : En utilisant le chiffrement de Vigenère chiffrer le message suivant : "VIGENERE VOUS SALUE"

Les chiffrements polyalphabétiques ne sont pas d'une manière significative plus difficile à casser.

Exercice : Proposer une méthode pour la cryptanalyse du chiffrement polyalphabétique.

One-time pad

Le one-time pad est un chiffrement qui cache toutes les informations statistiques qui peuvent être présentent sur le message clair.

- La clé secrète est une chaine binaire $k = k_1, k_2, k_3, \cdots$ aléatoire.
- Le texte clair est écrit sous la forme d'une chaine binaire $m = m_1, m_2, m_3, \cdots$
- Le chiffrement se fait bit par bit : le texte chiffré est $c = c_1, c_2, c_3, \cdots$, où $ci = m_i \oplus k_i$.
- Le déchiffrement s'effectue de la même manière : $m_i = c_i \oplus k_i$.

Comme démontré par Shannon, le one-time pad assure une haute sécurité. Cependant le one-time pad a de sérieuses défaillances.

- La clé secrète doit avoir une même taille que le message clair.
- La clé secrète ne peut pas être réutilisée.

Exercice

Pour le One Time Pad

1. Pourquoi nous devons respecter les deux points précédents.
2. Proposer une méthode pour contourner les défaillances citées.