Éléments de cryptologie

Travail Préparatoire

Dans ce travail préparatoire vous devrez programmer en python une attaque en brute force *intelligente*. Vous trouverez sur `ataraxy.info/Modelisation` dans la quatrième période un fichier `.zip` contenant :

- Un répertoire de dictionnaires de différentes langues (français, anglais, espagnol, italien, allemand, danois, suisse, néerlandais et suédois). Il s'agit de document texte avec les milliers de mots de la langue.
- Un fichier `BonusCrypto.py` contenant un ensemble de fonction simplifiant votre travail :
  - une fonction `MonDico` qui va prendre en paramètre la langue (`'FR'`, `'ANG'` etc) et qui va renvoyer un *arbre dictionnaire*. Les détails de la construction se trouve dans les commentaires introduisant ce fichier python.
  - Une fonction `pertinence` qui prend en paramètre une phrase et un arbre dictionnaire et renvoie le nombre de mot du dictionnaire contenu dans la phrase.
  - Une fonction `affichageSympathique` offrant un affichage sympathique.
- Un fichier `TravailPreparatoire.py` qui est un fichier python que vous devrez compléter. Vous pourrez, si vous estimez cela nécessaire, rajouter des fonction mais *en aucun cas en enlever*. Pour vous permettre de comprendre ce qui est attendu (et l'utilisation des différente fonction), la fonction d'attaque est donnée.

L'objectif est donc de compléter le programme pour le rendre fonctionnel et de transposer cela au chiffrement affine pour programmer une attaque en brute force sur ce type de chiffrement.

Le détail de ce principe de chiffrement est donnée ci dessous. Pour étrainner votre programme, vous trouverez dans le fichier `TravPrepPerso.pdf` un message personnel que vous devrez attaquer.

ATTENTION ! Ce devoir nécessite un travail personnel non négligeable. Vous avez deux semaines pour le faire parce que c'est nécessaire. N'attendez pas le dernier moment.

La fonction de chiffrement par la **méthode affine** est une généralisation de la méthode de César. Au lieu de prendre comme fonction de chiffrement une fonction *linéaire* de la forme $C(x) \equiv_{26} x + k$ on va considérer une fonction *affine* comme par exemple $C(x) \equiv_{26} 2x + 3$.

Chiffrons le message *BONJOUR*.

B	O	N	J	O	U	R
1	14	13	9	14	20	17
5	5	3	21	5	17	11
F	F	D	V	F	R	L

La première observation que nous pouvons faire est que la lettre B est cryptée en F et que la lettre O est également cryptée en F. Si nous recevons ce message comment pourrons-nous faire la différence entre le F qui est le cryptogramme de la lettre B et celui qui est le cryptogramme de la lettre O ?

Quelle est la fonction de déchiffrement ? Si nous travaillions avec nombres réels (on "oublie" le modulo 26), on vérifierai rapidement que la fonction $D(x) = \frac{1}{2}(x-3)$ permettrait de déchiffrer le message. Quel est l'équivalent du $\frac{1}{2}$ modulo 26 ? Pour répondre à cette question nous allons avoir besoin de définir l'inverse modulaire.

Définition

Soit $a \in \mathbb{Z}$. On note $D(a)$ l'ensemble des diviseurs positifs de a.

$$D(a) = \left\{ x \in \mathbb{N} \middle| \; x \mid a \right\}$$

Par exemple $D(132) = \{1, 2, 3, 4, 6, 12, 11, 22, 33, 44, 66, 132\}$.

Proposition

Soient a et b deux diviseurs positifs d'un entier n tel que $n = ab$.

$$(a \leqslant \sqrt{n}) \; \vee \; (b \leqslant \sqrt{n})$$

Démonstration. Raisonnons par l'absurde. S'il existe deux diviseurs a et b de n tel que $a > \sqrt{n}$ et $b > \sqrt{n}$ alors $ab > \sqrt{n}^2 = n$ et $n > n$ ce qui est impossible. □

Dans la pratique, pour la recherche des diviseurs d'un entiers, on va chercher à le factoriser par des entiers allant de 1 jusqu'à sa racine carré (sa partie entière précisement).

Par exemple déterminons $D(108)$. Puisque $\sqrt{108} \simeq 10.3923$. On va chercher à factoriser 108 par des entiers entre 1 et 10.

$$\begin{aligned} 108 &= 1 \times 108 \\ &= 2 \times 54 \\ &= 3 \times 36 \\ &= 4 \times 27 \\ &= 6 \times 18 \\ &= 9 \times 12 \end{aligned}$$

Ainsi $D(108) = \{1, 2, 3, 4, 6, 9, 12, 18, 27, 36, 54, 108\}$.

Définition 0.0.1

Le **plus grand commun diviseur** entre deux entiers a et b, noté $PGCD(a, b)$, est le plus grand entier de l'ensemble $D(a) \cap D(b)$.

Par exemple $PGCD(132, 108) = 12$ car $D(132) \cap D(108) = \{1, 2, 3, 4, 6, 12\}$.

Théorème (Algorithme d'Euclide)

Soit $a = bq + r$ une division euclidienne de deux entiers a et b.

$$\text{PGCD}(a, b) = \text{PGCD}(b, r)$$

Dans la pratique, lorsque l'on veut déterminer le PGCD entre deux entiers, on va réaliser des divisions euclidiennes successives en remplaçant à chaque fois, dividende et diviseur par diviseur et reste jusqu'à obtenir un reste nul, ce qui sera toujours possible par la construction de $\mathbb{N}$ (toute partie non vide de $\mathbb{N}$ admet un plus petit élément). Le PGCD est le dernier reste non nul.

Pour ordonner les idées, on place ces données dans un tableau, chaque ligne représentant une division euclidienne $a = bq + r$.

a	b	r	q
132	108	24	1
108	24	12	4
24	12	0	2

Ainsi le PGCD est 12.

Définition

Soient a, b et n des entiers tels que $n \geqslant 2$. On dira que b **est l'inverse de** a **modulo** n si $ab \equiv_n 1$

Par exemple 3 est l'inverse de 7 modulo 20 car $7 \times 3 = 21 \equiv_{20} 1$.

La question naturelle est de savoir si tous les nombres modulaires ont un inverse et accessoirement comment le trouver.

Définition

On dira que deux entiers a et b sont **premiers entre eux** si $\text{PGCD}(a, b) = 1$.

Par exemple 24 et 25 sont premiers entre eux.

Théorème

Soient a et n des entiers tels que $n \geqslant 2$. L'entier a admet un inverse modulo n si et seulement si a et n sont premiers entre eux.

Pour déterminer l'inverse modulaire d'un entier, on applique l'algorithme d'Euclide étendue. Détaillons un exemple et cherchons l'inverse de 382 modulo 2365. Pour que cela soit au moins possible, nous devons déterminer le PGCD de ces deux entiers. Appliquons l'algorithme d'Euclide.

a	b	r	q
2365	382	73	6
382	73	17	5
73	17	5	4
17	5	2	3
5	2	1	2
2	1	0	2

Puisque le dernier reste non nul est 1, on en déduit que 2365 et 382 sont premiers entre eux. D'après le corollaire précédent, il existe un inverse modulaire. Pour le trouver, il faut déterminer u et v tels que $2365u + 382v = 1$. Pur cela nous allons rajouter au tableau de l'algorithme d'Euclide deux colonnes u et v.

On va remplir ce tableau par le bas en initialisant u et v par des valeurs évidentes 0 et 1 respectivement. On peut vérifier, qu'à cette dernière ligne, on a bien $au + bv = 1$ ($2 \times 0 + 1 \times 1 = 1$).

a	b	r	q	u	v
2365	382	73	6		
382	73	17	5		
73	17	5	4		
17	5	2	3		
5	2	1	2		
2	1	0	2	0	1

On va remplir la ligne du dessus. On va mettre la valeur de v dans la nouvelle case de u.

Pour la nouvelle valeur de v on va mettre $-q \times \underset{\text{NEW}}{u} + \underset{\text{OLD}}{u}$ où $\underset{\text{NEW}}{u}$ représente le nouveau u (ici 1) et $\underset{\text{OLD}}{u}$ l'ancienne valeur (ici 0).

a	b	r	q	u	v
2365	382	73	6		
382	73	17	5		
73	17	5	4		
17	5	2	3		
5	2	1	2	1	-2
2	1	0	2	0	1

On recommence : à la ligne de dessus, on place l'ancienne valeur de v comme nouvelle valeur de u et pour la nouvelle valeur de u le résultat de $-q \times \underset{\text{NEW}}{u} + \underset{\text{OLD}}{u}$. Ici $\underset{\text{NEW}}{u} = -2$ et $\underset{\text{OLD}}{u} = 1$.

a	b	r	q	u	v
2365	382	73	6		
382	73	17	5		
73	17	5	4		
17	5	2	3	-2	7
5	2	1	2	1	-2
2	1	0	2	0	1

On peut vérifier qu'à chaque ligne $au + bv = 1$ (par exemple, notre dernier calcul permet d'aboutir à $17 \times (-2) + 5 \times 7 = 1$).

On réitère pour aboutir à :

a	b	r	q	u	v
2365	382	73	6	157	-972
382	73	17	5	-30	157
73	17	5	4	7	-30
17	5	2	3	-2	7
5	2	1	2	1	-2
2	1	0	2	0	1

Nous avons donc trouvé : $2365 \times 157 + 382 \times (-972) = 1$.

Pour finir (la question de départ était de trouver l'inverse de 382 modulo 2365), regardons cette égalité modulo 2365 (on rappel que $2365 \equiv_{2365} 0$) : $382 \times (-972) \equiv_{2365} 1$ et -972 est l'inverse modulaire de 382. On peut choisir un représentant positif : $-972 \equiv_{2365} -972 + 2365 = 1393$ et on peut donc conclure que l'inverse de 382 modulo 2365 est 1393.

Cryptosystème affine

Dans l'exemple d'introduction de ce chapitre nous avions crypté le message *BONJOUR* avec $2x+3$ comme fonction de chiffrement. Nous avions obtenue le message *FFDVFRL* et nous avions d'ailleurs remarqué qu'il allait être difficile de distinguer le *F* qui est un *B* du *F* qui est un *O*. En fait la méthode de chiffrement utilisée n'est pas bonne car 2 n'est pas inversible modulo 26 (car $\mathrm{PGCD}(26,2) = 2$). Pour pouvoir chiffrer avec la méthode affine il faut que le a (lorsque la fonction de chiffrement est $C(x) \equiv_{26} ax+b$) soit inversible modulo 26, c'est à dire que a et 26 soit premier entre eux.

Dans le cryptosystème affine une **clef** est donc de la forme (a,b) où a est un nombre premier à 26.

Le message suivant est chiffré par la méthode affine avec $(3,2)$ comme clef (cette clef est valide puisque 3 est premier avec 26 et admet donc un inverse modulaire) : *NSAIAKPMOEECUOCRRAPO*.

Pour déchiffrer ce message, déterminons l'inverse modulaire de 3.

a	b	r	q	u	v
26	3	2	8	-1	9
3	2	1	1	1	-1
2	1	0	2	0	1

Ainsi, $26 \times (-1) + 3 \times 9 = 1$. En regardant cette égalité modulo 26, on arrive à $3 \times 9 \equiv_{26} 1$ et 9 est l'inverse modulaire de 3. La fonction de déchiffrement est alors $D_{(3,2)}(x) \equiv_{26} 9(x-2)$. Déchiffrons le message.

	N	S	A	I	A	K	P	M	O	E	E	C	U	O	C	R	R	A	P	O
	13	18	0	8	0	10	15	12	14	4	4	2	20	14	2	17	17	0	15	14
-2	11	16	-2	6	-2	8	13	10	12	2	2	0	18	12	0	15	15	-2	13	12
$\times 9$	99	144	-18	54	-18	72	117	90	108	18	18	0	162	108	0	135	135	-18	117	108
$\equiv_{26}$	21	14	8	2	8	20	13	12	4	18	18	0	6	4	0	5	5	8	13	4
	V	O	I	C	I	U	N	M	E	S	S	A	G	E	A	F	F	I	N	E

Et *VOICI UN MESSAGE AFFINE*

Cryptanalyse

Dans le cas de la méthode affine, une attaque en force brute permet en général de casser le message crypté. On peut montrer que la cardinalité de l'espace des clefs pour la méthode affine (par paquet de 1) est 312. C'est un nombre important sur le papier mais dérisoire du point de vu de l'informatique.

Voici un exemple (nous ne présentons pas les 312 clefs possibles). Le message reçu est *TNCYGA*.

(a,b)	a^{-1}						
$(17,0)$	23	V	N	U	G	I	A
$(17,1)$	23	Y	Q	X	J	L	D
$(17,2)$	23	B	T	A	M	O	G
$(17,3)$	23	E	W	D	P	R	J
$(17,4)$	23	H	Z	G	S	U	M
$(17,5)$	23	K	C	J	V	X	P
$(17,6)$	23	N	F	M	Y	A	S
$(17,7)$	23	Q	I	P	B	D	V
$(17,8)$	23	T	L	S	E	G	Y
$(17,9)$	23	W	O	V	H	J	B
$(17,10)$	23	Z	R	Y	K	M	E
$(17,11)$	23	C	U	B	N	P	H
$(17,12)$	23	F	X	E	Q	S	K
$(17,13)$	23	I	A	H	T	V	N
$(17,14)$	23	L	D	K	W	Y	Q
$(17,15)$	23	O	G	N	Z	B	T
$(17,16)$	23	R	J	Q	C	E	W
$(17,17)$	23	U	M	T	F	H	Z
$(17,18)$	23	X	P	W	I	K	C
$(17,19)$	23	A	S	Z	L	N	F
$(17,20)$	23	D	V	C	O	Q	I
$(17,21)$	23	G	Y	F	R	T	L
$(17,22)$	23	J	B	I	U	W	O
$(17,23)$	23	M	E	L	X	Z	R
$(17,24)$	23	P	H	O	A	C	U
$(17,25)$	23	S	K	R	D	F	X

(a,b)	a^{-1}						
$(19,0)$	11	B	N	W	E	O	A
$(19,1)$	11	Q	C	L	T	D	P
$(19,2)$	11	F	R	A	I	S	E
$(19,3)$	11	U	G	P	X	H	T
$(19,4)$	11	J	V	E	M	W	I
$(19,5)$	11	Y	K	T	B	L	X
$(19,6)$	11	N	Z	I	Q	A	M
$(19,7)$	11	C	O	X	F	P	B
$(19,8)$	11	R	D	M	U	E	Q
$(19,9)$	11	G	S	B	J	T	F
$(19,10)$	11	V	H	Q	Y	I	U
$(19,11)$	11	K	W	F	N	X	J
$(19,12)$	11	Z	L	U	C	M	Y
$(19,13)$	11	O	A	J	R	B	N
$(19,14)$	11	D	P	Y	G	Q	C
$(19,15)$	11	S	E	N	V	F	R
$(19,16)$	11	H	T	C	K	U	G
$(19,17)$	11	W	I	R	Z	J	V
$(19,18)$	11	L	X	G	O	Y	K
$(19,19)$	11	A	M	V	D	N	Z
$(19,20)$	11	P	B	K	S	C	O
$(19,21)$	11	E	Q	Z	H	R	D
$(19,22)$	11	T	F	O	W	G	S
$(19,23)$	11	I	U	D	L	V	H
$(19,24)$	11	X	J	S	A	K	W
$(19,25)$	11	M	Y	H	P	Z	L

Et on gagne une bonne chose avec la fonction de chiffrement $19x+2$ qui admet $11(x-2)$ comme fonction de déchiffrement.

Pour compliquer un peu

Exactement de la même manière que pour la méthode de César, on peut raisonner en paquet.

Considérons par exemple, la clef $(2017, 123)$. Il faut d'abord s'assurer que 2017 est bien inversible modulo 2526. Pour ce faire, on applique l'algorithme d'Euclide étendu.

a	b	r	q	u	v
2526	2017	509	1	531	-665
2017	509	490	3	-134	531
509	490	19	1	129	-134
490	19	15	25	-5	129
19	15	4	1	4	-5
15	4	3	3	-1	4
4	3	1	1	1	-1
3	1	0	3	0	1

Ainsi $2526 \times 531 + 2017 \times (-665) = 1$. En regardant cette égalité modulo 2526, on arrive à $2017 \times (-665) \equiv_{2526} 1$ et $-665 \equiv_{2526} 1861$ est l'inverse de 2017 modulo 2526.

Ceci prouve que $(2017, 123)$ est bien une clef du cryptosystème affine. On obtient de plus que la fonction de déchiffrement est $D_{(2017,123)}(x) \equiv_{2526} 1861(x - 123)$.

Déchiffrons le message *701-211-1485-2369-1306-1215-852-816-861* obtenu en appliquant la méthode affine par paquet de 2 avec $(2017, 123)$ comme clef.

Message chiffré	701		211		1485		2369		1306		1215		852		816		861	
Déchiffrement	2108		2104		1104		1802		1417		1308		207		1413		1800	
Paquettage	21	08	21	04	11	04	18	02	14	17	13	08	02	07	14	13	18	00
Décodage	V	I	V	E	L	E	S	C	O	R	N	I	C	H	O	N	S	A