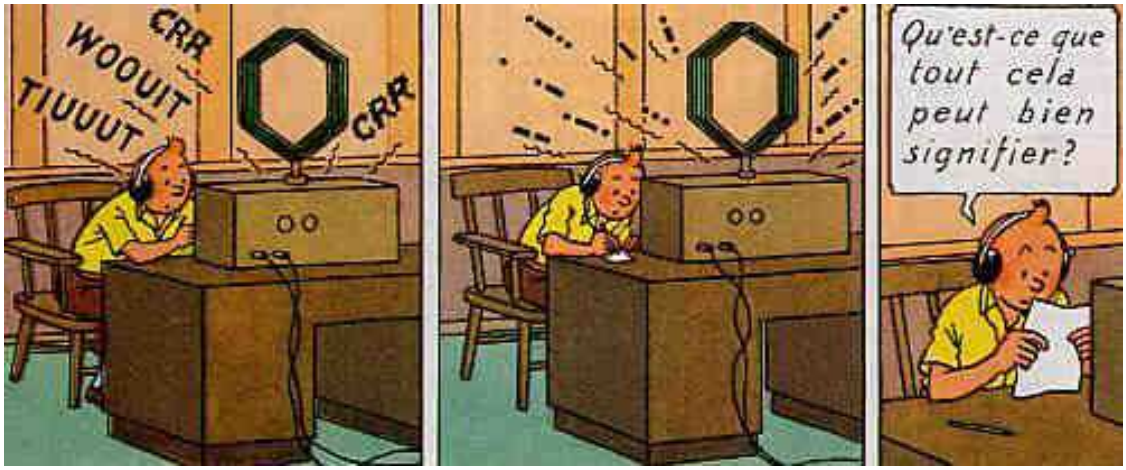
Cryptographie et codes secrets

La cryptographie est une des disciplines de la cryptologie s'attachant à protéger des messages (assurant confidentialité, authenticité et intégrité) en s'aidant souvent de secrets ou clés.

Elle est utilisée depuis l'Antiquité, mais certaines de ses méthodes les plus importantes datent de la fin du XXe siècle.



Partie 1 : L’art de cacher

- **La stéganographie**

Si la cryptographie est l'art du secret, la stéganographie est l'art de la dissimulation : l'objet de la stéganographie est de faire passer inaperçu un message dans un autre message et non de rendre un message inintelligible à autre que qui-de-droit. Pour prendre une métaphore, la sténographie consisterait à enterrer son argent dans son jardin là où la cryptographie consisterait à l'enfermer dans un coffre-fort — cela dit, rien n'empêche de combiner les deux techniques, de même qu'on peut enterrer un coffre dans son jardin.

C'est un mot issu du grec Stéganô, signifiant Je couvre et Graphô, signifiant J'écris. On retrouve une idée de dissimulation dans le nom de la technique.

Pouvez vous aider Tintin dans son problème ?

Dans les années 40, voici un tract apparemment collaborateur ... Qu'en pensez-vous ?

Aimons et admirons le chancelier Hitler
L'éternelle Angleterre est indigne de vivre
Maudissons, écrasons le peuple d'outre-mer
Le Nazi sur la terre sera seul à survivre
Soyons donc le soutien du Führer allemand
De ces navigateurs la race soit maudite
A eux seuls appartient ce juste châtiment
La palme du vainqueur répond au vrai mérite

L'acrostiche, c'est l'art de pouvoir partir d'un mot, et d'en déduire ce que l'on ressent par une série d'autres mots.

Quand je mets à vos pieds un éternel hommage,
Voulez-vous qu'un instant je change de visage ?
Vous avez capturé les sentiments d'un cœur
Que pour vous adorer forma le créateur.
Je vous chéris, amour, et ma plume en délire
Couche sur le papier ce que je n'ose dire.
Avec soin de mes vers lisez les premiers mots :
Vous saurez quel remède apporter à mes maux.

Alfred de Musset

Cette insigne faveur que votre cœur réclame
Nuit à ma renommée et répugne à mon âme.

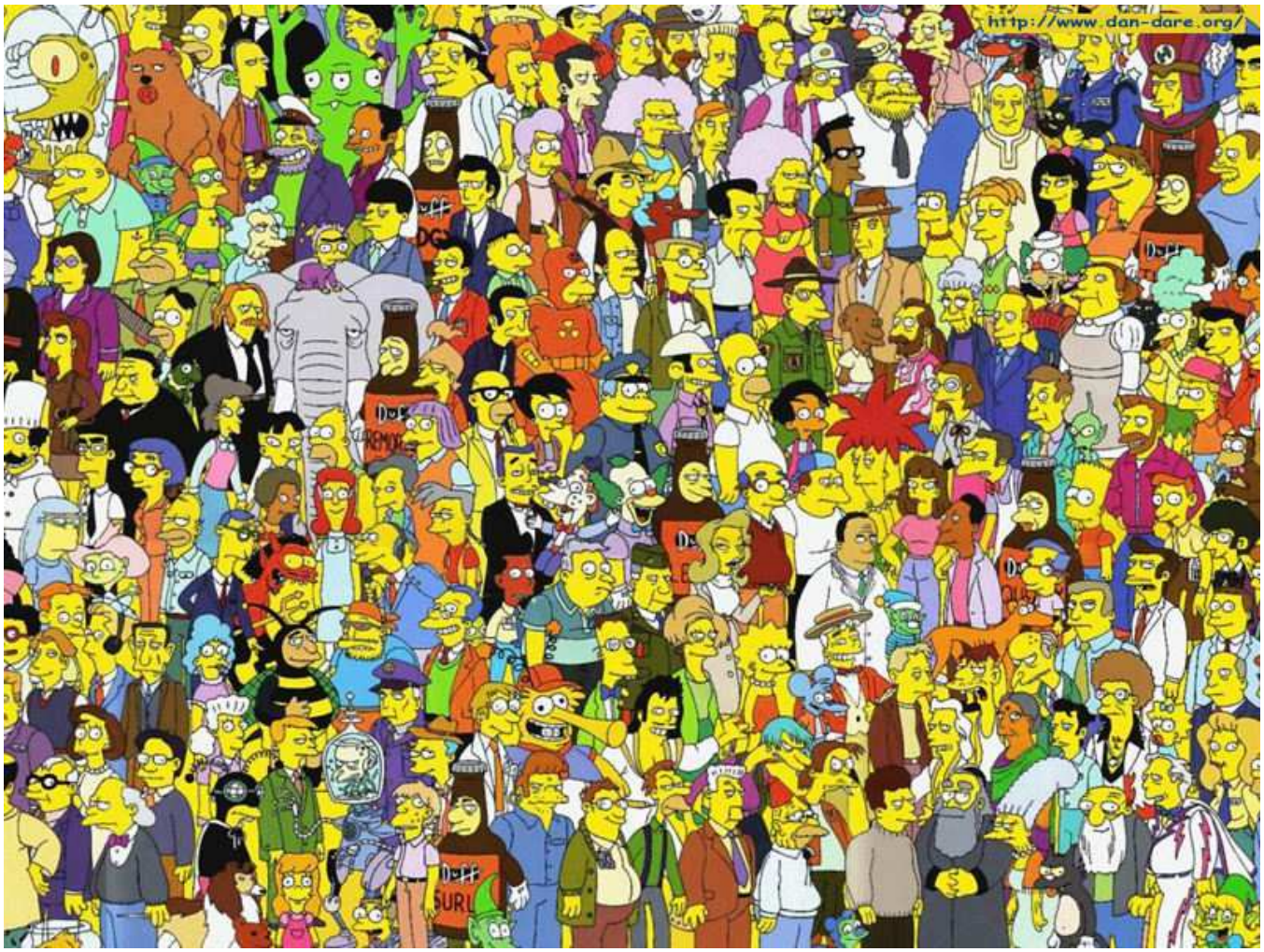
George Sand

Ou encore, ce double acrostiche célèbre dévoilant deux fois le prénom de l'être aimé ....

Amour parfait dans mon cœur imprima
Nom très heureux que j'aime bien Non !
Non jamais, cet amoureux lien
Autre que mort défaire ne pourra.

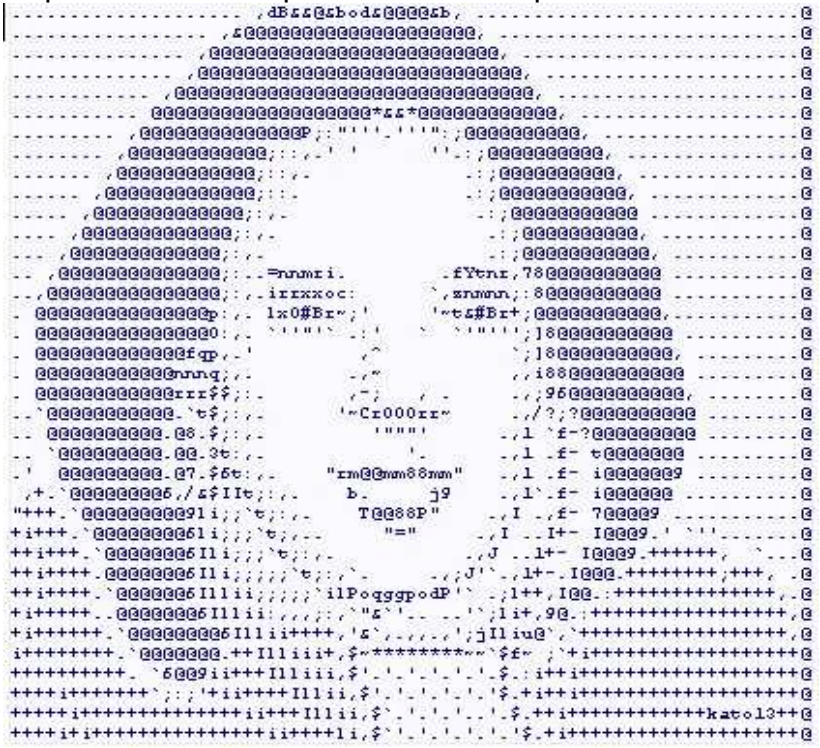
La dissimulation graphique

Pour cacher beaucoup, on peut montrer beaucoup, mais ou est Bart ?



Comment désigner précisément quelqu'un avec quelques caractères comme @ &% § ... ?

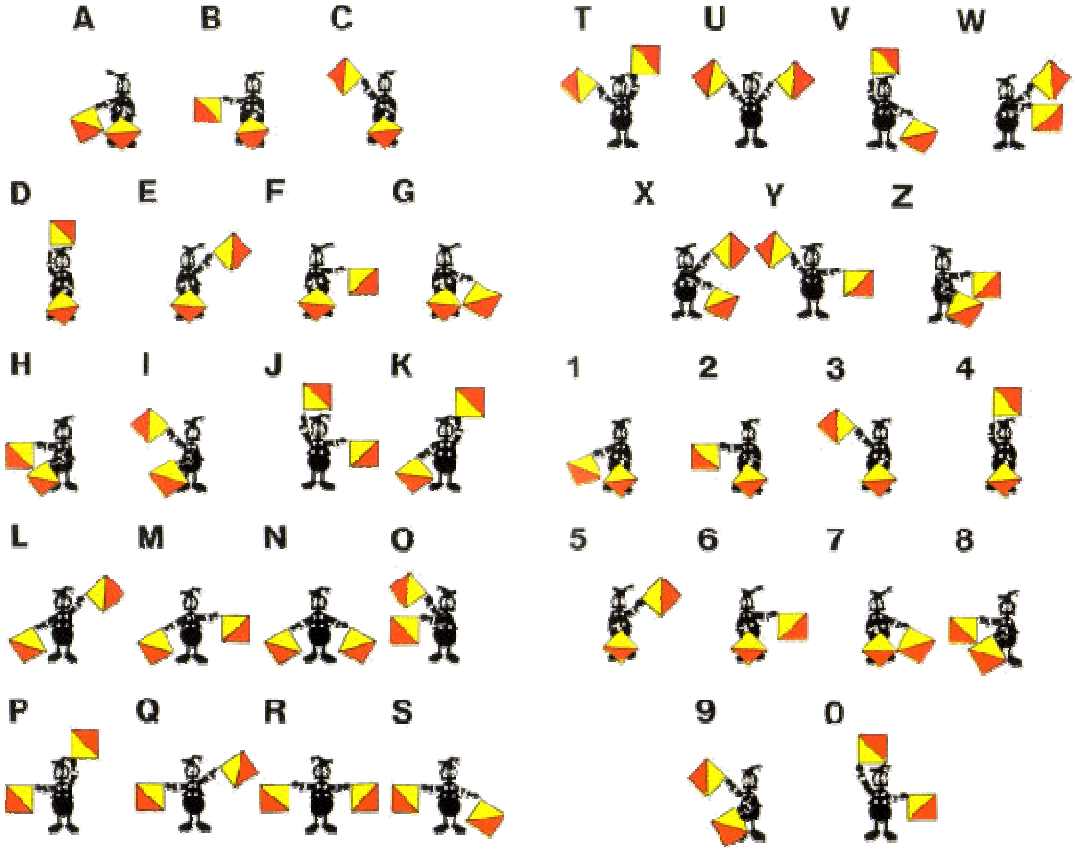
Réponse : Avec un peu d'ordre et de patience !



Quel est le mot caché dans ce billet de banque de l'île des Seychelles ? (officiel et véridique !)

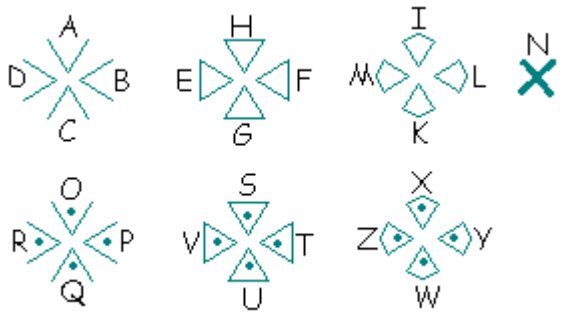
- **Les codages par signes**

Sémaphore vient du grec " SEMA " signe et " PHOROS " qui porte. Il désigne "un poste de signalisation établi sur une côte pour communiquer par signaux optiques avec les navires en vue".



Plutôt que de remplacer chaque lettre par une autre, certains (essentiellement des membres de société plus ou moins secrète) ont préféré remplacer les lettres par des symboles pour diffuser leurs messages codés. C'est ainsi qu'ont procédé les Templiers en utilisant un code fondé sur la croix des huits béatitudes

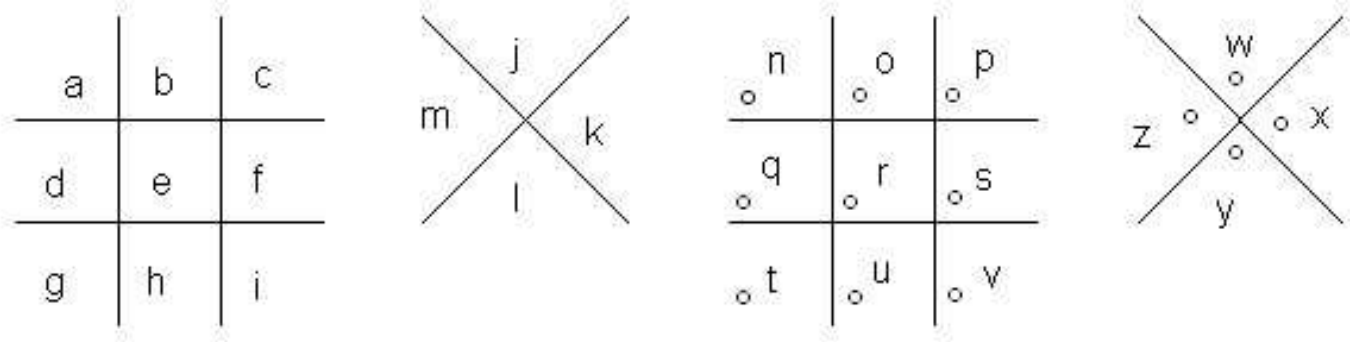
Le Temple était un ordre de moines fondé au XII ième siècle, dont la mission première était d'assurer la sécurité des pélerins en Terre Sainte. Rapidement, les templiers se détournèrent de cet objet, et s'enrichirent considérablement au point de devenir les trésoriers du roi et du Pape. Pour coder les lettres de crédit qu'ils s'échangeaient, ils remplaçaient chaque lettre par un symbole, suivant la substitution suivante :



Décoder le titre suivant :

Ecrivez votre nom de famille en code Templiers :

Les Francs-maçons ont suivi une idée similaire à celle des Templiers jusqu'au XIX ieme siècle. Leur code nommé "pig pen" était fondé sur des dessins en forme de petites cages.



Voici une petite énigme à décoder :

- **Partie 2 : les méthodes de cryptographie anciennes**

La plupart des méthodes de chiffrement reposent sur deux principes essentiels : **la substitution et la transposition.** Substituer signifie qu'on remplace certaines lettres par d'autres, ou par des symboles. Transposition signifie qu'on permute les lettres du message afin de le rendre inintelligible. Au cours des siècles, de nombreux systèmes cryptographiques ont été mis au point, de plus en plus perfectionnés, de plus en plus astucieux !

Le code de César

Le code de César est la méthode de cryptographie la plus ancienne communément admise par l'histoire. Il consiste en une substitution mono-alphabétique, où la substitution est définie par un décalage de lettres. Par exemple, si on remplace A par D, on remplace B par E, C par F, D par G, etc... Donnons un exemple sur à partir de ce décalage de 3 lettres :

Texte clair	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
Texte codé	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

Codez le texte suivant :

Mentir ? Jamais, la vérité est bien trop amusante (Steven Spielberg)

Codage :

Il n'y a que … façons différentes de crypter un message avec le code de César. Cela en fait donc un code très peu sûr, puisqu'il est très facile de tester de façon exhaustive toutes les possibilités.

Pourtant, en raison de sa grande simplicité, le code de César fut encore employé par les officiers sudistes pendant la guerre de Sécession, et même par l'armée russe en 1915.

Le carré de Polybe

Polybe est un historien grec qui vécut environ de -205 avant JC jusque -125 av. JC. A 40 ans, il est emmené parmi 1000 otages par les Romains suite à la bataille de Pydna en Macédoine et à la victoire de Paul-Émile sur les Grecs. Polybe tomba en admiration devant la civilisation romaine de l'époque, et d'otage il devint même ami de la famille de Paul-Émile.

Polybe est à l'origine d'une méthode très originale pour chiffrer, et qui est même antérieure au code de César. Pour cela, il dispose les lettres dans un tableau 5x5 :

	1	2	3	4	5
1	A	B	C	D	E
2	F	G	H	I,J	K
3	L	M	N	O	P
4	Q	R	S	T	U
5	V	W	X	Y	Z

On remplace alors chaque lettre par ses coordonnées dans le tableau, en écrivant d'abord la ligne, puis la colonne. Par exemple, le A est remplacé par 11, le B est remplace par 12, le F par 21, le M par 32.... Notons qu'il subsiste une ambigüité pour les lettres I et J

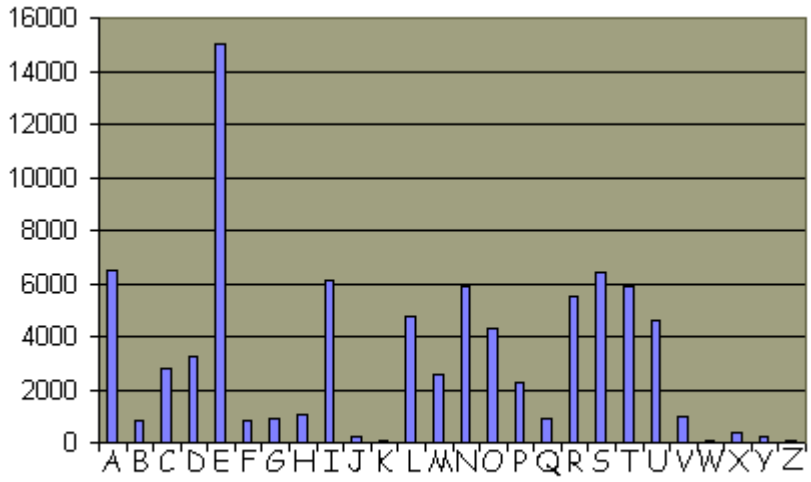
Codez grâce au carré de Polybe le texte suivant :

La politique est éphémère mais une équation est éternelle (Albert Einstein)

Codage :

Les faiblesses de la méthode par substitution :

Quelle que soit la méthode de cryptographie par substitution mono-alphabétique, il y a a priori 26! clés possibles, ce qui en soi est déjà un chiffre énorme. En fait, ce nombre de clés est illusoire, car la cryptographie par substitution possède une grosse faiblesse structurelle : dans les langues, toutes les lettres n'ont pas la même fréquence d'apparition. Dans un texte français, il y a presque toujours beaucoup plus de E que de W. Or, le E est toujours remplacé par la même lettre et le W aussi. Donc, si dans votre texte, la lettre qui apparait le plus fréquemment est un L, il y a de fortes chances que ce soit un E. En revanche, si il n'y a presque pas de D, on peut se dire que c'est probablement un W, ou un K, un X,etc... Voici par exemple l'analyse de la fréquence d'apparition des différentes lettres dans une série de textes variés en langue française :



Fréquence des lettres dans l'alphabet

On voit clairement apparaître sur le graphique précédent plusieurs groupes de lettres qui ont la même fréquence :

- le E est de loin le plus fréquent. Il y a donc de fortes chances que la lettre la plus fréquente du texte codé est en fait un E.
- Ensuite, les A,I,N,R,S,T.
- Puis les O,U,L,...

Pour déterminer dans le deuxième groupe quel lettre est un A, une méthode possible est d'étudier les lettres qui apparaissent isolées (c'est-à-dire dans un mot à une seule lettre du texte) : la lettre isolée la plus fréquente a de fortes chances d'être un A.

Ensuite, on étudie les groupes de 2 lettres. On peut étudier leur fréquence d'apparition dans le texte, et comparer aux fréquences possibles du texte. Par exemple, les EN,NE,TE,SE sont bien plus fréquents que les ST,NR,..

Menez l'enquête : saurez vous décrypter ce message bien mystérieux reçu dans une boite mail?

♍♒♏❒ ♓■♦♏❒■♋◆♦♏
❖□◆⬩ ❖♏■♏⌘ ♎♏ ❒♏♍♏❖□♓❒ ◆■ ❖♓❒◆⬩ ♓■➚□❒❍♋♦♓❑◆♏ ♌♏●♑♏
♍□❍❍♏ ■□◆⬩ ■♏ ⬩□❍❍♏⬩ □♋⬩ ♦❒♏⬩ ♋❖♋■♍♏⬩ ♦♏♍♒■□●□♑♓❑◆♏❍
♏■♦ ♍♏♍♓ ♏⬩♦ ◆■ ❖♓❒◆⬩ ❍♋■◆♏●
♠❼♓● ❖□◆⬩ □●♋➘♦🖃 ♏➚➚♋♍♏⌘ ♦□◆⬩ ●♏⬩ ➚♓♍♒♓♏❒⬩ ♎♏ ❖□♦❒♏
♎♓⬩❑◆♏ ♎◆❒ ♏♦ ♏■❖□🖂♏⌘ ♍♏ ❍♋♓● → ♦□◆⬩ ♍♏◆☒ ❑◆♏ ❖□◆⬩
♍□■■♋♓⬩⬩♏⌘
❍♏❒♍♓ ♌♏♋◆♍□◆◘ ♎♏ ■□◆⬩ ♋♓♎♏❒

Le chiffre de Vigenère : la cryptographie par substitution poly-alphabétique

Même si l'on connaissait depuis fort longtemps les faiblesses de la cryptographie par substitution, il n'y eut pas entre César et le XVI è s. de véritable nouveau procédé cryptographique, à la fois sûr (pour les moyens de l'époque) et facile à utiliser! Blaise de Vigenère, né en 1523, fut l'initiateur d'une nouvelle façon de chiffrer les messages qui domina 3 siècles durant. Vigenère était quelqu'un de très hétéroclite, tantôt alchimiste, écrivain, historien, il était aussi diplomate au service des ducs de Nevers et des rois de France. C'est en 1586 qu'il publie son Traicté des chiffres ou Secrètes manières d'écrire, qui explique son nouveau chiffre (le texte intégral est disponible sur le site de la Bibliothèque Nationale de France).

L'idée de Vigenère est d'utiliser un chiffre de César, mais où le décalage utilisé change de lettres en lettres. Pour cela, on utilise une table composée de 26 alphabets, écrits dans l'ordre, mais décalés de ligne en ligne d'un caractère. On écrit encore en haut un alphabet complet, pour la clé, et à gauche, verticalement, un dernier alphabet, pour le texte à coder :

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

Pour coder un message, on choisit une clé qui sera un mot de longueur arbitraire. On écrit ensuite cette clé sous le message à coder, en la répétant aussi souvent que nécessaire pour que sous chaque lettre du message à coder, on trouve une lettre de la clé. Pour coder, on regarde dans le tableau l'intersection de la ligne de la lettre à coder avec la colonne de la lettre de la clé.

Exemple : On veut coder le texte "CRYPTOGRAPHIE DE VIGENERE" avec la clé "MATHWEB". On commence par écrire la clef sous le texte à coder :

C	R	Y	P	T	O	G	R	A	P	H	I	E	D	E	V	I	G	E	N	E	R	E
M	A	T	H	W	E	B	M	A	T	H	W	E	B	M	A	T	H	W	E	B	M	A

Pour coder la lettre C, la clé est donnée par la lettre M. On regarde dans le tableau l'intersection de la ligne donnée par le C, et de la colonne donnée par le M.

On trouve O. Puis on continue. On trouve : ORRWPSHDAIOEI EQ VBNARFDE.

Avec la clé « Rempart », cryptez la citation suivante :

Le cercle n'est qu'une ligne droite revenue à son point de départ (Fred. Dard)

Codage :

Enigme : testez votre expérience grandissante en cryptographie …

GNV WBXPCMQNTW WKTPRFIWAUCP PWF UEW IFAPVUEEUI YFT GANKJILTE P'QXGWVQPYGQ CFEQCYP

Réponse :

- # Partie 3 : Approches de la cryptographie moderne

La cryptographie est depuis le milieu du siècle dernier, un secteur scientifique de première importance. Le monde moderne diffuse un nombre toujours croissant d'informations de plus en plus confidentielles, principalement par le biais de l'Internet. Les applications civiles du chiffrement (banques, télécommunications, informatique, cartes bleues...) sont un moteur fondamental de progrès.
Bien entendu, la cryptographie moderne a largement besoin des mathématiques, notamment de sa branche arithmétique, commençons par quelques notions simples qui seront fondamentales pour la compréhension de la suite de la partie 3 :

Glossaire arithmétique de la cryptographie

Les nombres entiers : en arithmétique, si vous lisez « nombre entier », comprenez nombre entier relatif, c'est-à-dire positif ou négatif.

Division euclidienne : combien de groupes de 10 moutons peut on réunir dans un troupeau de 63 bêtes ? La réponse est 6 et il restera 3 moutons esseulés, ceci s'écrit 63 = 6 x 10 + 3. Nous avons procédé à la division euclidienne (ou « avec reste ») de 63 par 10. Le résultat 6 est appelé le quotient et 3 est le reste de la division de 63 par 10.

En suivant cette approche, on démontre que de façon générale, pour tous nombres entiers *a* et *b* (*b* strictement positif), il existe deux entiers *q* et *r* tels que : $\boldsymbol{a = bq + r}$ **avec** $\boldsymbol{0 \leq r < b}$

Divisibilité de a par b ≠ 0 : on dit que l'entier b (non nul) divise a si le reste de la division euclidienne de a par b est nul. On dit aussi que a est divisible par b ou enfin que a est un multiple de b.

Exemple : 8 divise 56 car 56 = 8 x 7 + 0. Par contre 10 ne divise pas 63 car le reste est 3 ≠ 0

Pgcd de deux nombres entiers : chaque nombre entier possède des diviseurs (au moins deux qui sont 1 et lui-même). L'ensemble des diviseurs communs à deux entiers a et b possède un plus grand élément appelé «plus grand commun diviseur » entre a et b et noté pgcd(a,b). Si celui-ci est égal à 1, on dit que ***a et b sont premiers entre eux***. Cette dernière notion est très importante.

Exemple 1: plus grand commun diviseur de 18 et 24 = pgcd(18 , 24)

Les diviseurs de 18 sont 1, 2, 3, 6, 9 et 18
Les diviseurs de 24 sont 1, 2, 3, 4, 6, 8 et 12
L'ensemble des diviseurs communs de 18 et 24 est {1,2,3,6}, son plus grand élément étant 6 on conclut : pgcd(18,24) = 6.

Exemple 2 : 9 et 16 sont premiers entre eux car pgcd(9,16) = 1.

En effet les diviseurs de 9 sont 1,3, 9 et ceux de 16 sont 1, 2, 4, 8,12. Le seul diviseur commun est bien 1.

Algorithme d'Euclide : pour déterminer le pgcd de deux nombres, le plus simple est d'effectuer une suite de divisions euclidiennes. Calculons par exemple pgcd (168,45).

Nous commençons par diviser 168 par 45 ce qui donne : 168 = 3 x 45 + 33.
On écrit alors que 33 = 168 – 3x45 ce qui montre que tout diviseur commun à 168 et 45 est aussi un diviseur commun à 33 et 45 et donc que **pgcd(168,45) = pgcd(33,45)**.

On généralise ce résultat : **si** ***a = bq + r*** **avec** ***0 ≤ r < b*** ***alors pgcd(a,b) = pgcd(r,b)***
On ré-itère le procédé en divisant 45 par 33 : 45 = 1x33 + 12 donc pgcd(33,45) = pgcd(33,12)
Et on poursuit : 33 = 2x12 + 9 donc pgcd(33,12) = pgcd(12,9)
12 = 1x9 + 3 donc pgcd(12,9) = pgcd(3,9)
9 = 3x3 + 0 stop : pgcd(3,9) = 3

Ainsi, en remontant la « cascade », on détermine pgcd(168,45) = 3.

L'algorithme d'Euclide que nous venons d'appliquer donne toujours le pgcd de deux nombres a et b en un nombre fini d'étapes, car la suite des restes des divisions effectuées dans ce cadre est strictement décroissante, ainsi l'un d'entre eux est nul ce qui implique que le *précédent* est le pgcd recherché.

Congruence modulo n. C'est une notion essentielle en cryptographie, soit un entier a que l'on souhaite diviser par l'entier n non nul, on peut écrit a = nq + r qui se note a = r [n] ou encore a ≡ r [n]

Vocabulaire : a ≡ r [n] se lit « a égale r modulo n » ou encore « a est congru a n modulo n »

Proposons la définition suivante : ***trouver la congruence d'un nombre a modulo n, c'est remplacer ce nombre a par son reste r dans la division euclidienne de a par n.***

Exemple 1 : prenons a = 23 et n = 6, on a 23 = 3x6 + 4 donc 23 ≡ 4 [6]

Exemple 2 : prenons a = 13 et n = 15, on a cette fois 13 = 0x15 + 13 et donc 13 ≡ 13 [15] !

Théorème de Bezout : on démontre facilement grâce à l'algorithme d'Euclide, que si a et b sont deux nombres entiers et d leur pgcd, **il existe deux entiers u et v tels que** ***a x u + b x v = d***

En particulier, si a et b sont premier entre eux (ie pgcd(a,b) = 1), ***a x u + b x v = 1***

Nombres premiers : un nombre entier strictement supérieur à 1 est dit premier s'il n'est divisible que par 1 et lui-même. Les nombres premiers inférieurs à 100 sont :

2	3	5	7	11	13	17	19	23	29	31	37	41
43	47	53	59	61	67	71	73	79	83	89	97	

Le principe de la cryptographie à clé secrète :

Les données lisibles et compréhensibles sans intervention spécifique sont considérées comme du *texte en clair.* La méthode permettant de dissimuler du texte en clair en masquant son contenu est appelée le *cryptage.* Le cryptage consiste à transformer un texte normal en charabia inintelligible appelé *texte chiffré.* Cette opération permet de s'assurer que seules les personnes auxquelles les informations sont destinées pourront y accéder. Le processus inverse de transformation du texte chiffré vers le texte d'origine est appelé le *décryptage.* La Figure suivante illustre ce processus.

Un outil important : l'utilisation des nombres modulo n

L'avantage des nombres modulo n, c'est qu'ils permettent de travailler avec un ensemble de n nombres qui vont « rester dans cette ensemble » par les opérations courantes (additions, multiplications, élévation à la puissance ...). Expliquons cette idée :

En cryptographie moderne, lorsqu'on voudra crypter une expression, la première étape sera de transformer chaque caractère de l'expression en nombre, par exemple en utilisant le code ASCII par exemple. La codification la plus simple est d'attribuer à chaque lettre de l'alphabet un nombre, par exemple celui de son rang en partant de zéro : A se code 00, B se code 01,C se code 02 … Z se code 25 . Ainsi le mot « SECRET » se codera 18 04 02 17 04 19. Cette démarche est nécessaire car vous savez qu'un ordinateur ne travaille qu'avec des chiffres et encore, pas avec tous …

L'expression 180402170419 ne résisterait pas longtemps à un pirate qui aurait intercepté votre message, le codage utilisé est trop trivial ! Il va falloir la crypter en utilisant par exemple une clé secrète, c'est-à-dire un procédé de transformation des précédents nombres, connus de vous seul et de votre correspondant. Disons pour faire simple que la clé secrète est la multiplication de chaque nombre par 3…

L'expression devient 54 12 06 51 12 57 qui est déjà plus énigmatique si on ne connaît pas la clé secrète. Seulement, j'aimerai transmettre à mon correspondant autre chose qu'une suite de chiffres comme 541206511257, très pénible à lire : je pars d'un mot en lettres, je souhaite transmettre un mot en lettres, crypté mais en lettres … Problème : 54, 51 ou 57 ne correspondent pas à des lettres de l'alphabet. L'utilisation des nombres modulo 26 vont apporter une solution à mon problème.

En effet considérons l'ensemble fini {0, 1, 2, 3, …,24, 25} , (NB : cet ensemble de 26 éléments se note Z/26Z et est ce que l'on appelle un anneau d'entiers), cet ensemble a la particularité d'être « stable » pour l'addition et la multiplication pour des résultats modulo 26, la stabilité indiquant que le résultat obtenu ne « sort pas » de l'ensemble {0, 1, 2, 3, …,24, 25}

Par exemple 12 + 20 = 32 dans l'addition traditionnelle : on sort de {0, 1, 2, 3, …,24, 25}, mais si on prend le résultat congru à 26, on obtient $12 + 20 \equiv 6\ [26]$. En effet $32 = 1 \times 26 + 6$ donc 6 est bien le reste de la division de 32 par 26.

Ainsi dans notre ensemble {0, 1, 2, 3, …,24, 25} , si on parle en « modulo 26 » on a $12 + 20 \equiv 6$

De même : $17 + 23 \equiv 14$ et par exemple $3 \times 18 \equiv 2$ … tous les résultats obtenus « restent » dans {0, 1, 2, 3, …,24, 25} , c'est le principe de la stabilité.

On choisira donc au lieu de l'expression 54 12 06 51 12 57, la même mais modulo 26 c'est-à-dire 02 12 06 25 12 05 qui correspond au mot CMGZMF, qui est le mot crypté.

Allons plus loin :Comment choisir une clé de cryptage ? Dans notre exemple, on multiplie chaque nombre par 3 modulo 26, réfléchissez à ce qui se serait produit si on avait choisi une multiplication par 2 modulo 26 ?

Allons encore plus loin … Oui, car il s'agit maintenant pour le récipiendaire du message de le décrypter afin de revenir au message initial, c'est-à-dire le mot « SECRET » … si la clé de cryptage est le procédé « x 3 modulo 26 », quel est le procédé de décryptage ? Vous trouvez ? … Ce n'est pas évident … Voyons dans ce qui suit, le cas général de la théorie.

Cryptage et décryptage par transformation affine

Ce procédé reprend celui expliqué ci-dessus en le complexifiant peu : on ne va pas juste transformer chaque nombre n issu du message en le multipliant par un entier a mais on va lui faire subir une transformation affine a x n + b. Nous sommes donc soumis à deux exigences :

- Comment choisir a et b pour que chaque nombre n possède une image unique n' modulo 26 par la transformation n' = a x n + b ?
- Comment trouver la transformation inverse pour décrypter le message obtenu ?

Ce procédé, quoique déjà assez complexe comme nous allons le voir, n'est pas utilisé pour les transferts de données sensibles sur Internet, il est encore trop simple pour les pirates chevronnés, par exemple, le fait d'obtenir un mot crypté avec des lettres ayant la même occurrence que celles du mot de départ (puisque chaque lettre du mot de départ correspond à une lettre unique du mot crypté), un test de fréquences sur le message intercepté pourrait facilement venir à bout du cryptage !

Néanmoins, ce procédé montre dans quel esprit est réalisé un cryptage beaucoup plus complexe appelé code RSA, qui est le plus utilisé à ce jour …

Traitement d'un exemple : supposons que nous voulions crypter le message suivant

« Le rire est le propre de l'homme »

Comme précédemment, numérotons de 0 à 25 chaque lettre de l'alphabet et choisissons comme le couple (a=5 ; b=3) comme clé de cryptage affine, c'est-à-dire que chaque nombre n du message codé sera égal modulo 26 à n' = 5 x n + 3.

Commençons par remplir **le tableau de cryptage** (5,3) suivant

Lettre en clair	A	B	C	D	E	F	G	H	I	J	K	L	M
n	0	1	2	3	4	5	6	7	8	9	10	11	12
n'					23								
Lettre codée					X								
Lettre en clair	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
n	13	14	15	16	17	18	19	20	21	22	23	24	25
n'													
Lettre codée													

Le message crypté est donc : ……

……

Comment choisir les nombres a et b ?

Pour que le codage puisse remplir sa fonction, il faut qu'à chaque entier n dans [0,25] soit associé une ***unique*** image n' = a x n + b dans [0,25] ce qui ne sera pas le cas pour tous les couples (a,b) .

En effet, si par exemple on choisit a = 4 et b = 1, on voit que pour n = 0 on trouve n' = 4 x 0 + 1 ≡ 1 [26] mais si n = 13 alors n' = 4 x 13 + 1 ≡ 1 [26] , on a donc une ambiguité car 0 et 13 ont la même image par la transformation affine.

Théorème : soit n est un entier dans [0,25], soit n' son image modulo 26 définie par n' = a x n + b

Lorsque n décrit [0,25], les valeurs de n' seront toutes différentes les unes des autres si et seulement si a est un entier premier avec 26 dans [0,25].
b peut être librement choisi parmi les 26 entiers de [0,25]

Remarque : cela laisse comme choix pour a un des nombres **: 1,3 ,5, 7, 9, 11, 15, 17, 19, 21, 23, 25**.

Question : combien y a-t-il de transformations affines utilisables pour crypter un texte ?

Réponse :

La preuve mathématique du précédent théorème (plutôt niveau 1 ère S)

Soit a premier avec 26 dans [0,25] . supposons qu'il existe m et n dans [0,25] tels que am + b ≡ an + b modulo [26] . On a alors a(m – n) ≡ 0 ce qui signifie que 26 divise a(m-n) mais comme 26 et a sont premiers entre eux, on conclut que 26 divise (m-n), ce qui s'écrit m-n ≡ 0 [26] et donc m ≡ n [26]

Comment décrypter ?

Car ce n'est pas tout, il faut penser aussi au « voyage retour », c'est-à-dire au décryptage du message par notre correspondant. Il semble évident que la transformation réciproque de notre transformation affine sera aussi une transformation affine … c'est le cas et nous l'admettrons.

Il s'agit donc de trouver désormais a' et b' tels que par la transformation a' x n' + b' on retrouve l'entier n de départ (celui qui a donné n !)

Par exemple, si comme dans notre exemple a = 5 et b = 3 , nous pourrons décrypter avec le couple a' = -5 et b' = 15. Voyons sur un exemple comment ça fonctionne.

Je souhaite crypter le mot «OM »

1. On code O = 14 et M = 12, donc OM = 14 12
2. on crypte avec la clé de cryptage (5,3), on obtient 21 11 qui correspond au mot VL
3. J'envoie VL à mon destinataire qui le recode lui aussi en 21 11
4. on décrypte 21 11 avec la clé de décryptage (-5 , 15). Voici les calculs
   Décryptage de 21 : 21 x -5 + 15 = -90 = -4 x 26 + 14 donc 21 x -5 + 15 ≡ 14 [26]
   Décryptage de 11 : 11 x -5 + 15 = -40 = -2 x 26 + 12 donc 11 x -5 + 15 ≡ 12 [26]

On est bien revenu sur le code 14 12 qui signifie OM.

Comment trouver a' et b' connaissant a et b ?

NB : Dans tout l'exposé les congruences s'entendent modulo 26

On peut écrire que $n \equiv a' \times (a \times n + b) + b'$ ce qui s'écrit $(aa' - 1) n + a'b + b' \equiv 0$ pour tout n, ceci n'est possible que si $aa' - 1 \equiv 0$ et $a'b + b' \equiv 0$.

Conclusion : **il est nécessaire et suffisant que $aa' \equiv 1$ et que $b' \equiv - a'b$.**

Un tel a' et b' existent-ils toujours ? oui si a premier avec 26 : en effet, le théorème de Bezout énonce qu'alors, il existe u et v entiers tels que $a \times u + 26 \times v = 1$, posons $a' = u$ et on obtient $aa' = -v \times 26 + 1$ ce qui s'écrit aussi $aa' \equiv 1$.

Pour trouver la valeur de a' (et de v mais dont on se moque …) tel que $a \times a' + 26 \times v = 1$, on utilise l'algorithme d'Euclide que l'on va « remonter » comme dans l'exemple ci-dessous.

Rappel : ici, a = 5 et b = 3.

Algorithme d'Euclide :

On divise 26 par 5 $\rightarrow 26 = 5 \times 5 + 1$

On divise 5 par 1 $\rightarrow 5 \times 1 + 0$ ***stop***

L'avant dernière égalité donne donc $1 \times 26 - 5 \times 5 = 1$ et on peut donc identifier $a' = -5$ et $v = 1$

Le calcul de b' est donc $b' = - (-5) \times 3 = 15$

Conclusion : la clé de décryptage est bien (-5 ,15)

Exercice : On choisit la clé de cryptage affine (9,4), calculer la clé de décryptage

Réponse : ……

Les faiblesses du cryptage à clés secrètes : le problème avec le cryptage affine, c'est que pour décrypter, le récipiendaire du message a besoin de connaître la clé initiale du cryptage, ce qui sous-entend que d'une manière ou d'une autre, cette cruciale information doit circuler aussi entre les correspondants … c'est donc un serpent qui se mord la queue : a quoi sert un cryptage complexe si il faut aussi envoyer par le même canal la façon dont on s'y est pris pour crypter l'information ?

Le cryptage à clés secrète peut se schématiser de la façon suivante : un message est un papier enfermé dans une boite muni d'un cadenas ne comportant qu'une seule clé .

A souhaite envoyer correspondre confidentiellement avec B
A écrit un message sur le papier, l'enferme dans la boite et ferme le cadenas dont il a la clé
A envoie la boite a B, le message est protégé car cadenassé, mais A doit aussi transmettre la clé pour B afin que celui-ci puisse ouvrir la boite ! C'est en ça que réside la faille de sécurité

Pour le pirate interceptant la boite, il y a deux solutions : casser le cadenas ou voler la clé ! Si la clé circule quelque part, ce n'est pas impossible à faire …

L'idéal serait donc un cryptage qui ne nécessiterait pas de la part de A de faire circuler une quelconque information sur la clé de cryptage utilisée : c'est le principe du cryptage par clé publique.

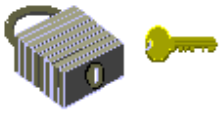
Le principe de la cryptographie à clé publique :

En 1976, Whitfield Diffie et Martin Hellman propose une nouvelle façon de chiffrer, qui contourne cet écueil. Commençons par expliquer leur procédé de façon imagée. Un ami doit vous faire parvenir un message très important par la poste, mais vous n'avez pas confiance en votre facteur que vous soupçonnez d'ouvrir vos lettres. Comment être sûr de recevoir ce message sans qu'il soit lu? Vous commencez par envoyer à votre ami un cadenas sans sa clé, mais en position ouverte. Celui-ci glisse alors le message dans une boite qu'il ferme à l'aide du cadenas, puis il vous envoie cette boite. Le facteur ne peut pas ouvrir cette boite, puisque vous qui possédez la clé pouvez le faire …

Cryptographie à clé publique :

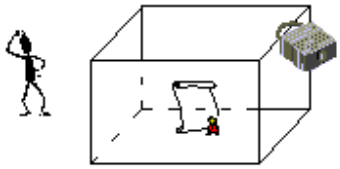
Alice Bob

Etape 1 : Fabrication des clés. Bob fabrique une clé publique qui permet de sceller le message codé dans la boite (ici : le cadenas), et une clé privée qui permet d'ouvrir le cadenas.

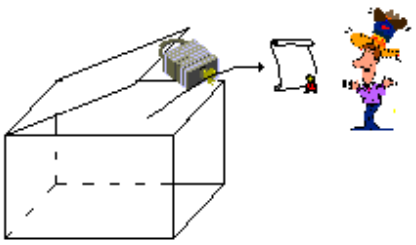


Etape 2 : Distribution des clés. Bob fait parvenir à Alice le cadenas, mais garde la clé pour lui.

Etape 3 : Envoi du message. Alice met son message dans une boite qu'elle ferme à l'aide du cadenas.



Etape 4 : Réception du message. Bob ouvre la boite a l'aide de sa clé, et récupère le message. Personne n'a pu l'intercepter puisque lui seul pouvait ouvrir la boite.



La cryptographie à clé publique repose exactement sur ce principe. On dispose d'une fonction P sur les entiers, qui possède un inverse S. On suppose qu'on peut fabriquer un tel couple (P,S), mais que connaissant uniquement P, il est impossible (ou au moins très difficile) de retrouver S.

- P est la clé publique, que vous pouvez révéler à quiconque. Si Louis veut vous envoyer un message, il vous transmet P(message).
- S est la clé secrète, elle reste en votre seul possession. Vous décodez le message en calculant S(P(message))=message.
- La connaissance de P par un tiers ne compromet pas la sécurité de l'envoi des messages codés, puisqu'elle ne permet pas de retrouver S. Il est possible de donner librement P, qui mérite bien son nom de clé publique.

Bien sûr, il reste une difficulté : comment trouver de telles fonctions P et S. Diffie et Hellman n'ont pas eux-même proposé de fonctions satisfaisantes, mais dès 1977, D.Rivest, A.Shamir et L.Adleman trouvent une solution possible, la meilleure et la plus utilisée à ce jour, la cryptographie RSA. Le RSA repose sur la dichotomie suivante :

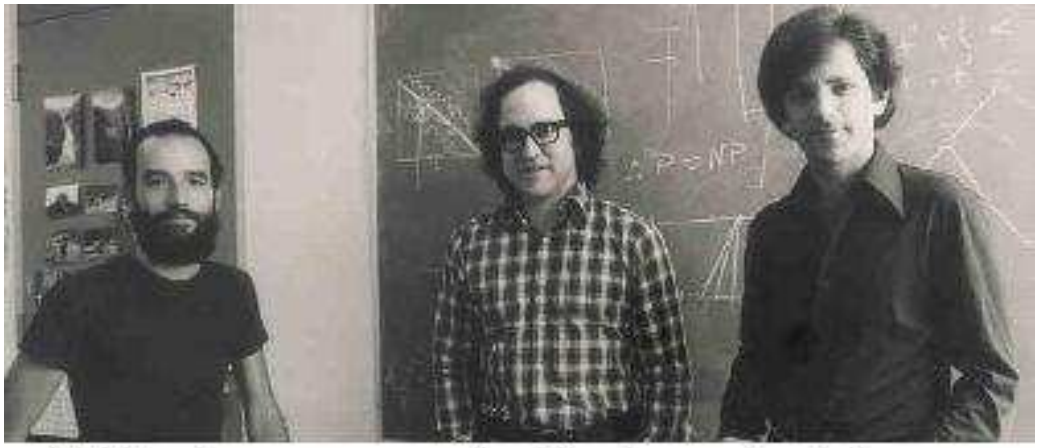
- il est facile de fabriquer de grands nombres premiers p et q (pour fixer les idées, 100 chiffres).
- étant donné un nombre entier n=pq produit de 2 grands nombres premiers, il est très difficile de retrouver les facteurs p et q.

La donnée de n est la clé publique : elle suffit pour chiffrer. Pour décrypter, il faut connaitre p et q, qui constituent la clé privée.

Les algorithmes à clé publique (on parle aussi de chiffrement asymétrique) ont pourtant un grave défaut : ils sont lents, beaucoup plus lents que leurs homologues symétriques. Pour des applications où il faut échanger de nombreuses données, ils sont inutilisables en pratique. On a alors recours à des cryptosystèmes hybrides. On échange des clés pour un chiffrement symétrique grâce à la cryptographie à clé publique, ce qui permet de sécuriser la communication de la clé. On utilise ensuite un algorithme de chiffrement symétrique. Le célèbre PGP, notamment utilisé pour chiffrer le courrier électronique, fonctionne sur ce principe

Le cryptage RSA

La méthode de cryptographie RSA a été inventée en 1977 par Ron Rivest, Adi Shamir et Len Adleman, à la suite de la découverte de la cryptographie à clé publique par Diffie et Hellman. Le RSA est encore le système cryptographique à clé publique le plus utilisé de nos jours. Il est intéressant de remarquer que son invention est fortuite : au départ, Rivest, Shamir et Adleman voulaient prouver que tout système à clé publique possède une faille.



Adi Shamir Ron Rivest Len Adleman

Principe de fonctionnement : Si Bob souhaite recevoir des messages en utilisant le RSA, il procède de la façon suivante :

1. Création des clés : Bob crée 4 nombres p,q, e et d :
   - p et q sont deux grands nombres premiers distincts. Leur génération se fait au hasard, en utilisant par exemple un algorithme de test de primalité probabiliste.
   - e est un entier premier avec le produit (p-1)(q-1).
   - d est tel que ed=1 modulo (p-1)(q-1). Autrement dit, ed-1 est un multiple de (p-1)(q-1). On peut fabriquer d à partir de e, p et q, en utilisant l'algorithme d'Euclide.
2. Distribution des clés : Le couple (n,e) constitue la clé publique de Bob. Il la rend disponible par exemple en la mettant dans un annuaire. Le couple (n,d) constitue sa clé privée. Il la garde secrète.
3. Envoi du message codé : Alice veut envoyer un message codé à Bob. Elle le représente sous la forme d'un ou plusieurs entiers M compris entre 0 et n-1. Alice possède la clé publique (n,e) de Bob. Elle calcule $C=M^e$ mod n. C'est ce dernier nombre qu'elle envoie à Bob.
4. Réception du message codé : Bob reçoit C, et il calcule grâce à sa clé privée $D=C^d$ (mod n). D'après un théorème du mathématicien Euler, $D=M^{de}=M$ (mod n). Il a donc reconstitué le message initial

Les démonstrations mathématiques du procédé RSA utilisent le théorème de Fermat et encore le théorème de Bezout pour la construction des clés …

Le RSA est il sûr ?

Les attaques actuelles du RSA se font essentiellement en factorisant l'entier n de la clé publique. La sécurité du RSA repose donc sur la difficulté de factoriser de grands entiers. Le record établi en 1999, avec l'algorithme le plus performant et des moyens matériels considérables, est la factorisation d'un entier à 155 chiffres (soit une clé de 512 bits, 2^{512} étant proche de 10^{155}). Il faut donc, pour garantir une certaine sécurité, choisir des clés plus grandes : les experts recommandent des clés de 768 bits pour un usage privé, et des clés de 1024, voire 2048 bits, pour un usage sensible. Si l'on admet que la puissance des ordinateurs double tous les 18 mois (loi de Moore), une clé de 2048 bits devrait tenir jusque ... 2079.

```
109417386415705274218097073220403576120037329454492059909138
421314763499842889347847179972578912673324976257528997818337
970765372440271467435315933543338975
1026395928297411057720541965739916759007165678080380668033341
933521790711307779
×106603488380168454820927220360012878679207958575989291522
70608237193062808643
```

Factorisation d'un entier à 155 chiffres.

Quoique... Il n'est pas interdit de penser que cela est illusoire. D'abord, les algorithmes de factorisation vont probablement être améliorés. Ensuite, rien ne dit que casser le RSA est aussi difficile que de factoriser n. Il existe peut-être un autre moyen d'inverser la clé publique sans passer par la factorisation de n. En particulier, une mauvaise utilisation de la cryptographie RSA (choix d'un exposant e trop petit, mauvaise complétion des blancs,...) la rend particulièrement vulnérable. Enfin, les progrès de la physique vont peut-être sonner le glas de la cryptographie mathématique. Il a été défini, du moins en théorie, un modèle d'ordinateur quantique qui, s'il était réalisé, permettrait de factoriser très rapidement des entiers. Les ordinateurs quantiques n'en sont encore qu'à leurs prémices, et leur record (automne 2001) est la factorisation de 15=3×5!

En conclusion, on peut dire que le RSA est probablement une méthode de chiffrement assez sûre; il y a très certainement plus de risques liés à une mauvaise utilisation qu'à une attaque. Il ne faut toutefois pas verser dans un optimisme béat. Rappelons que dans l'histoire, des dizaines de fois, les

armées ont cru posséder un chiffrement absolument sûr, et l'utilisaient avec une confiance aveugle; des dizaines de fois les adversaires ont rivalisé d'ingéniosité et ont réussi à décrypter ces messages. L'exemple de l'Enigma est en cela édifiant. Il n'est pas exclu qu'un service secret comme la NSA ait réussi une percée considérable dans un des axes cités précédemment, et que pour lui, le décryptage du RSA ne soit plus qu'un jeu d'enfant!

Sources utilisées pour ce document :

Site bibmath.net

Cryptographie et Codes secrets / Tangente édition Pole

Site wikipédia

Site casepubli.lu

Site apprendre-en-ligne.net