Compte rendu du T.I.P.E

sur une Etude Bibliographique

Sujet n°14 :

« Cryptage et le problème du sac à dos »

ORANCI Sevan
POURROY Louis
2ème Année Ingénieur

E. Chanthery
2005-2006
09/06/2006

Tables de Matières

Résume de l'étude Bibliographique

Le sujet de notre étude bibliographique porte sur un thème bien connu de tout temps par les informaticiens et les historiens : il s'agit de l'univers du **Cryptage**.

Le sujet de cette étude s'étend sur trois grandes parties relatant du cryptage sous trois angles : historique, technique et applicatif. En effet, les origines et les évolutions du cryptage sont le sujet de notre première partie, la notion de cryptage par la technique du « Sac à dos » constitue notre seconde partie et enfin l'application de certain algorithme de cryptage au service de la confidentialité des informations épilogue notre étude.

La notion de cryptographie remonte à l'Antiquité durant laquelle les Grecs utilisaient des outils primaires pour envoyer des messages codés. La nécessité de vouloir crypter un message s'est manifestée par la volonté d'empêcher une tierce personne autre que le destinataire à décrypter le message. Plus tard, les outils de cryptographie se sont perfectionnés ainsi que les méthodes de chiffrement dans le but de transmettre plus rapidement et plus efficacement un message. Chaque guerre apporta son lot de perfectionnement au service de la stratégie militaire, car un message crypté attire une plus grande attention qu'un message en clair. Tôt ou tard, suivant la technologie et le génie d'une personne, il sera décrypté.

De nos jours, la cryptographie est utilisé dans de nombreux domaines : militaire, informatique, financier, sécurité… et permet de sécuriser les flux d'informations. Les exemples de programmes écrits par des étudiants montrent à quel point l'avenir de la cryptographie est prometteur.

Plan Prévisionnel et Organisationnel

Nous allons, ici, vous présenter les différentes étapes d'élaboration de ce dossier, en commençant par le choix du sujet, jusqu'à l'impression finale de ce document.

Une fois le binôme constitué, nous avons commencé à effectuer très rapidement nos premières recherches sur les différents sujets bibliographiques proposés dans le cadre du T.I.P.E. Chaque sujet a été soigneusement étudié en fonction de plusieurs paramètres : intérêt du sujet, antécédents individuels, préférences personnelles…

Tout deux passionnés par le monde des nouvelles technologies, en particulier celui de l'informatique, nous avons choisi un sujet d'étude ayant un impact direct sur cet univers. Le cryptage est aujourd'hui omniprésent dans nos communications : réseaux sans fils, transaction bancaire, confidentialité… et il nous a paru intéressant d'expliquer ce phénomène pour en établir un travail soigneux et bien ordonné. L'autre aspect qui émane de ce sujet est le patrimoine historique de la cryptographie ; ses origines sont nombreuses et l'étude des différentes évolutions de cryptographie fut passionnante.

Ainsi, à partir du 1er Avril 2006, et ce, jusqu'au 31 Avril 2006, nous avons individuellement entamés des recherches approfondies concernant le monde de la cryptographie pour nous imprégner du sujet et commencer à réfléchir à l'organisation de l'étude. Nous avons trouvés différentes sources exploitables concernant l'aspect technique et historique du sujet. De nombreux sites internet ont déjà profondément étudié le thème de la cryptographie sous tous les angles, et certains livres nous ont aidés à dégager des idées directrices de notre rapport (l'objectif étant de poser des questions sur le sujet et de pouvoir y répondre à l'aide de différents supports).

La première esquisse du plan s'effectua mi-Mai, ce qui nous permit de partager le travail de manière équitable jusqu'au 1er Juin : Mr. Pourroy s'est penché sur l'aspect technique du sujet et Mr. Oranci, sur l'aspect historique. Concernant les exemples d'applications, nous avons tout deux cherchés des sites proposant, soit des applets java (animations web), soit des codes sources de programmes permettant de crypter un message.

Une fois ces trois parties dactylographiées, Mr. Oranci a rédigé le résumé de l'étude, le plan organisationnel ainsi que l'introduction et Mr. Pourroy s'est chargé de rédiger la conclusion ainsi que les annexes.

Introduisons à présent notre sujet...

INTRODUCTION

La cryptologie est née avec l'apparition de l'écriture et fut justifiée par le besoin de protéger tout message écrit afin d'éviter que l'ennemi ne puisse, en se l'appropriant, exploiter les renseignements qu'il contenait. Littéralement « science du secret », elle a longtemps été associée à de mystérieux enjeux d'espionnage militaire et diplomatique bien éloignés des préoccupations scientifiques. La cryptographie est quant à elle, l'ensemble des techniques constituant la cryptologie.

Ses premières formes furent le plus souvent basiques, mais avec l'évolution des peuples des technologies et des « casseurs de codes », son développement est en constante recherche de progression. Comment la cryptologie émergea ? Où en est elle de nos jours ? Quel avenir a-t-elle ?

A l'époque de Jules César, des méthodes rudimentaires permettaient de rendre ses ordres « incompréhensibles » à ses adversaires. Néanmoins, l'absence de rigueur des concepteurs de ces systèmes mena à des failles qui permettaient à leurs adversaires de comprendre les messages malgré tout.

De nos jours l'information sous toutes ses formes : voix, images, oeuvres musicales, textes et autres, circule au format numérique à travers le monde en une fraction de seconde. Que ce soit par le téléphone, le câble, les fibres optiques ou par satellite, cette information est chaque jour échangée d'un point à un autre et se trouve susceptible d'être lue, copiée, supprimée, altérée ou falsifiée. La cryptologie répond aujourd'hui aux besoins du marché et constitue un domaine scientifique en pleine activité. Elle intervient dans de multiples applications et représente l'élément essentiel de la sécurisation du commerce électronique et du réseau Internet.

La transmission d'un message codé qui se veut sûr doit satisfaire les trois conditions suivantes :

- ✓ « Confidentialité » : la technique de cryptage se doit de garantir le secret de l'information, aucun tiers ne doit pouvoir lire le message.
- ✓ « Intégrité » : le cryptosystème ne doit engendrer aucune absence de modification de l'information, et le message ne doit pas pouvoir être modifié durant son « transport ».
- ✓ « Authenticité »: le codage du message doit garantir l'origine de l'information, le message doit annoncer son émetteur sans le trahir.

Nous allons dans un premier temps vous présenter les origines de la cryptographie, ses évolutions au cours du temps : de l'Antiquité à aujourd'hui, son utilité et son impact sur les communications.

Puis, nous allons tenter, à travers la deuxième partie, analyser un type de cryptage bien particulier nommé « Sac à Dos ». Cela se réalisera par des explications mathématiques ainsi que des schémas d'explications. Nous essayerons également de décrire ses techniques et de comprendre ce que les méthodes des mathématiques modernes ont apporté de plus que les autres techniques cryptographiques.

Enfin, nous vous présenterons des exemples de programmation permettant aux informaticiens de crypter leurs données. Nous nous attarderons sur les avantages et inconvénients des différents types de langages utilisés pour coder un message électronique.

A quand remonte le premier document crypté ?

I. Histoire et origine du Cryptage

La cryptologie, étymologiquement la **science du secret**, ne peut être vraiment considérée comme une science que depuis peu de temps. Cette science englobe la cryptographie — l'écriture secrète — et la cryptanalyse — l'analyse de cette dernière. On peut dire que la cryptologie est un **art ancien et une science nouvelle** : un art ancien car Jules César l'utilisait déjà ; une science nouvelle parce que ce n'est que depuis les années 1970 que c'est un thème de recherche scientifique académique (universitaire). Cette discipline est liée à beaucoup d'autres, par exemple la théorie des nombres, l'algèbre, la théorie de la complexité, la théorie de l'information, ou encore les codes correcteurs.

Nous allons vous présenter, dans cette première partie, l'origine réelle de la cryptographie en remontant jusqu'à l'Antiquité, l'évolution des différentes techniques de cryptographies à travers le temps, et enfin, plus proche de nous, l'impact de son utilisation durant la Seconde Guerre Mondiale.

A) Dans l'Antiquité

1) Les premières méthodes de chiffrement

A ce jour, le premier document chiffré connu remonte à l'époque de l'Antiquité (juste après la Préhistoire). Il s'agit d'une tablette d'argile, retrouvée en Orient (Irak), datant du XVI^{ème} siècle avant Jésus-Christ. Les nombreuses études scientifiques ont démontrées qu'un potier y avait gravé sa recette de cuisine secrète en supprimant des consonnes et en modifiant l'orthographe des mots.

Plus tard (X^{ème} et VII^{ème} siècle av. J.-C, les Grecs ont utilisés une technique de **chiffrement dit « par transposition »** à des fins militaires. Cela consiste à effectuer des changements de positions de lettres dans le message. L'outil employé est connu sous le nom de « **Scytale** » (Figure 1) également appelée « *bâton de Plutarque* », autour duquel ils enroulent en spires jointives une bande de cuir et y inscrivent le message. Une fois déroulé, le message est envoyé au destinataire qui doit posséder un bâton identique (diamètre) nécessaire au déchiffrage. Cette technique fut vite obsolète car son seul avantage réside dans le procédé de chiffrement : un bâton de diamètre quasi-identique suffit à déchiffrer le texte.

Figure 1: Une scytale

A partir du V^{ème} siècle av. J.-C., l'une des **premières techniques** de chiffrements est utilisée dans les textes religieux par les Hébreux. Parmi les différents procédés élaborés par les Hébreux, celui-ci est appelé « ***Atbash*** » : c'est une méthode de substitution alphabétique inversée (Figure 2). Son origine est tirée des premières et dernières lettres de l'alphabet hébreux : alph, tau, beth, shin… Sa technique est très simple : elle consiste à remplacer chaque lettre du texte en clair par une autre lettre de l'alphabet choisie de la manière suivante : A devient Z, B devient Y, etc…

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

Figure 2: Chiffrement Atbash

Un procédé **original et très peu technique** est celle employée par le roi de Babylone, Nabuchodonosor, qui, aux alentours de -600 av. J.-C. écrivait sur le crâne rasé de ses esclaves, attendait que leurs cheveux aient repoussé, et il les envoyait à ses généraux. Pour déchiffrer le message, il suffisait de raser l'individu. Ce procédé était certes fiable (l'interception du message par un tiers est tout de suite remarqué), mais son temps d'exécution était long (il faut attendre plusieurs semaines pour la pousse des cheveux).

Il faut attendre -200 av. J.-C. pour voir apparaître les **premiers « vrais » systèmes de cryptographie**. Ce sont essentiellement des chiffrements par substitution.

Il existe 4 types de substitutions:

➔ « ***Mono-alphabétique*** » : remplace chaque lettre du message par une autre lettre de l'alphabet.

➔ « ***Poly-alphabétique*** » : utilise une suite de chiffres mono-alphabétiques (la clé) réutilisée périodiquement.

➔ « ***Homophonique*** » : fait correspondre à chaque lettre du message en clair un ensemble possible d'autres caractères.

➔ « ***Polygrammes*** » : substitue un groupe de caractères dans le message par un autre groupe de caractères.

Nous allons à présent nous attarder sur un code bien connu de l'épopée Romaine, celui de César…

2) Le fameux code de César

Le code de **César** est réputé car c'est la méthode cryptographique par substitution mono-alphabétique la plus ancienne (I^{er} av. J.-C.). Pour coder un message il effectua un décalage de 3 lettres dans l'alphabet : la lettre A était remplacé par D, B par E, etc…

C'est un code, en définitive, très peu sûr, car il n'y a que 26 façons différentes de crypter un message, néanmoins, la faible alphabétisation de la population la rend suffisamment efficace. Plus tard, durant la Guerre de Sécession, les sudistes n'hésitèrent pas à employer cette technique, tout comme les Russes en 1915.

Voici (Figure 3) le décalage effectué par César :

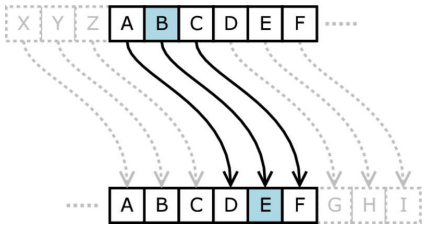


Figure 3 : Principe du chiffre de César

Aujourd'hui ce code est encore présent dans de nombreux forums de discussions concernant la programmation informatique de cryptage. Connu sous le nom de « ROT13 » (rot-ation de 13 lettres ou A➔N), son seul objectif est d'empêcher la lecture involontaire d'une réponse à une devinette, ou de l'intrigue d'un film. Il n'a pas pour but de rendre strictement confidentiel un texte.

Intéressons nous au code d'un historien Grec, qui révolutionna la cryptographie, tout comme César…

3) L'originalité du « carré » de Polybe

Après César, ce fût un historien Grec nomme **Polybe** qui écrivit le premier procédé de chiffrement par substitution **homophonique** en l'an -165 av. J.-C. Il donna le nom de « carré » à son procédé de cryptage car celui-ci est basé sur un carré de 25 (ou plus) cases.

Prenons un tableau de 25 cases (Figure 4), disposons notre alphabet comme suit :

	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

Figure 4: Carré de Polybe

Nous sommes obliger d'amalgamer deux lettres (ici c'est la lettre I et J). Chaque lettre est donc représentée par deux coordonnées : celui de sa ligne et celui de sa colonne.

Voici un exemple de codage par le carré de Polybe :

Sevan-et-louis-travaillent ➔ 4315511133-15-3134452443-4442115111243131153344

Les cryptologues modernes ont vu dans le « carré de 25 » plusieurs caractéristiques extrêmement intéressantes:

- ✓ La conversion des lettres en chiffres,
- ✓ La réduction de nombres de symboles,
- ✓ La représentation de chaque lettre par deux éléments séparés.

Ce système de chiffrement peut être compliqué avec un mot de passe. Par exemple, si le mot de passe est « salut », on commencera à remplir le carré avec les lettres de ce mot, après avoir supprimé les lettres identiques, puis on complètera le tableau avec les lettres inutilisées.

Voici (Figure 5) le même procédé mais avec une clé (mot de passe) :

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

Figure 5: Carré de Polybe avec une clé

Concernant son moyen de transmission, Polybe proposait d'utiliser deux torches lumineuses : l'une placée au devant du bateau et l'autre à l'arrière de ce dernier. Ainsi, pour transmettre la lettre E, il suffisait d'allumer une fois une des deux torches et cinq fois l'autre : tout dépend de l'orientation du bateau.

Après l'Antiquité et les débuts du cryptage,
voyons la période précédant la Seconde Guerre Mondiale…

B) De l'Antiquité à la guerre

Nous n'allons, bien évidemment, pas passer en revue toutes les évolutions – siècle par siècle – de la cryptographie. Seul certain tournant ou procédé exceptionnel (révolutionnaire) seront expliqués.

1) Chiffrage par disques

En 1412, le savant italien **Leone Battista Alberti** expose pour la première fois le chiffrement par substitution « polyalphabétique » qu'il applique à l'aide d'un disque à chiffrer (Figure 6) et d'un mot clé. Ce dernier est constitué de deux disques composés de deux alphabets. Le petit disque étant mobile, et le second fixe.

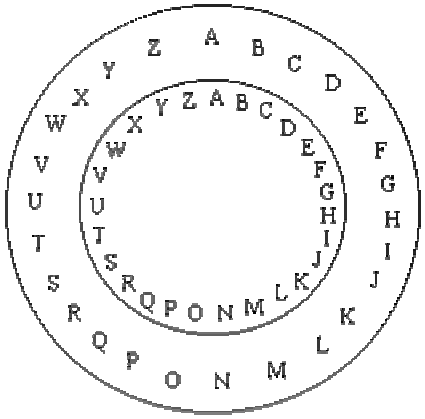


Figure 6: Disque à chiffrer

Méthode de chiffrement :

➔ Prenons le mot clé « Louis », et cryptons le mot « Serein »

➔ Pour coder la première lettre, il faut tourner le disque mobile de manière à ce que sa lettre « A » soit en face de la lettre « L » du second disque. On remplace alors la lettre « S » par la lettre correspondante sur le grand disque.

➔ Pour coder la deuxième lettre, il faut tourner encore une fois le disque, de manière à faire correspondre la lettre « A » du petit disque avec la lettre « O » du grand disque. On remplace la deuxième lettre du message (« E ») par celle correspondante sur le grand disque.

➔ On fait de même jusqu'à avoir passé toutes les lettres de la clé. On reprend alors la clé au début et ainsi de suite.

De nombreuses variantes concernant la clé sont apparues par la suite (utilisation de majuscules, changement d'alphabet...).

2) Système Della Porta

En 1563, l'Italien **Giovanni Battista Della Porta** expose dans son livre *« De Furtivis Literarum Notis, vulgo de ziferis »* les connaissances en cryptologie connues jusqu'à cette époque. Il présente un procédé de chiffrement par substitution polyalphabétique utilisant 13 alphabets différents (Figure 7) qui restera efficace pendant trois siècles et expose une technique de substitution digrammatique (Figure 8) consistant à remplacer chaque couple de lettres du texte en clair par un symbole.

Méthode de chiffrement polyalphabétique :

➔ Partageons notre alphabet traditionnel par groupe de deux lettres (AB, CD, EF…)

➔ Pour chacun de ces groupes, établissons un tableau de 13 colonnes et de 2 lignes

➔ Dans chacun des tableaux, nous disposons 13 lettres de l'alphabet (le choix est libre au codeur) sur la première ligne et 13 autres sur la seconde ligne. Ainsi, chaque lettre de la première ligne aura sa lettre de cryptage respectif.

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

Figure 7: Chiffrage Poly-alphabétique

➔ Convenons d'un mot de passe, par exemple « TIPE ». Ce mot de passe indiquera, pour chaque lettre, l'alphabet à utiliser. Pour crypter un mot, nous allons donc voir pour chaque lettre, l'alphabet à utiliser. Une fois qu'on a l'alphabet, on remplace chaque lettre du mot à chiffrer par celle qui se trouve juste au-dessus ou en dessous.

➔ Cryptons le mot « sevan » avec la clé « tipe ».

➔ Pour la lettre « s » on utilise l'alphabet « t », et au dessus du « s » il y a « b ».
➔ Pour la lettre « e », on utilise l'alphabet « i », et en dessous du « e » il y a « n ».
➔ Pour la lettre « v », on utilise l'alphabet « p », et au dessus du « v » il y a « c ».
➔ Pour la lettre « a », on utilise l'alphabet « e », et en dessous du « a » il y a « y ».
➔ Pour la lettre « n », on utilise l'alphabet « t », et au dessus du « n » il y a « j ».

Voici le résultat :

Texte clair :	S	E	V	A	N
Clé :	T	I	P	E	T
Cryptogramme :	B	N	C	Y	J

Méthode de chiffrement polyalphabétique :

➔ Ici, nous utilisons le tableau (Figure 8) inventé par Della Porta dont le procédé de cryptage consiste à la représentation d'un symbole pour un ensemble de deux caractères de l'alphabet.

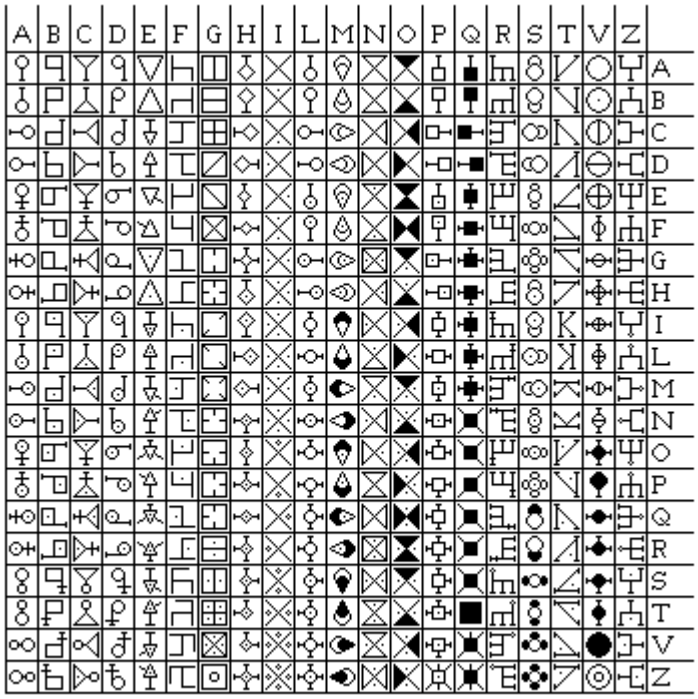


Figure 8: Substitution Digrammatique

Le souci de ce procédé réside dans le fait que si l'on observe le tableau original de Porta, on remarque tout de suite des répétitions gênantes: par exemple, la colonne G est constituée uniquement de carrés, la colonne I uniquement de X, autant de renseignements que pourrait utiliser un décrypteur.

3) Le chiffre de Vigenère

En 1586, le diplomate Français **Blaise de Vigenère** élabore un système de substitution polyalphabétique. Le procédé consiste à remplacer une lettre par une autre qui n'est pas toujours la même. C'est un système bien plus solide que le code de César (rotation) car elle nécessite une clé de (dé) cryptage. En effet, pour pouvoir chiffrer un texte, nous avons besoin d'une clé, dont les caractères sont utilisés pour effectuer la substitution. Évidemment, plus la clé sera longue et variée et mieux le texte sera chiffré.

Méthode de chiffrement :

➔ Tout d'abord il est indispensable de disposer de la table de Vigenère présentée ci-après (Figure 9).

➔ Choisir une clé de (dé) cryptage (exemple : sevan)

➔ Avoir un texte à coder : « bonjour monsieur »

➔ Faire correspondre la clé avec le texte à coder autant de fois que nécessaire :

B	O	N	J	O	U	R		M	O	N	S	I	E	U	R
S	E	V	A	N	S	E		V	A	N	S	E	V	A	N

➔ Situer la colonne de la lettre du texte en clair (lettre « B » ➔ colonne B)

➔ Situer la ligne de la lettre de la clé (lettre « S » ➔ ligne S)

➔ Le croisement de la ligne et la colonne vous donne la lettre codée : « T »

➔ Au final, nous obtenons le texte crypté suivant : « TSIJAMV HOAKMZUE »

Méthode de déchiffrement :

➔ Faire correspondre le texte codé avec la clé autant de fois nécessaire pour la clé :

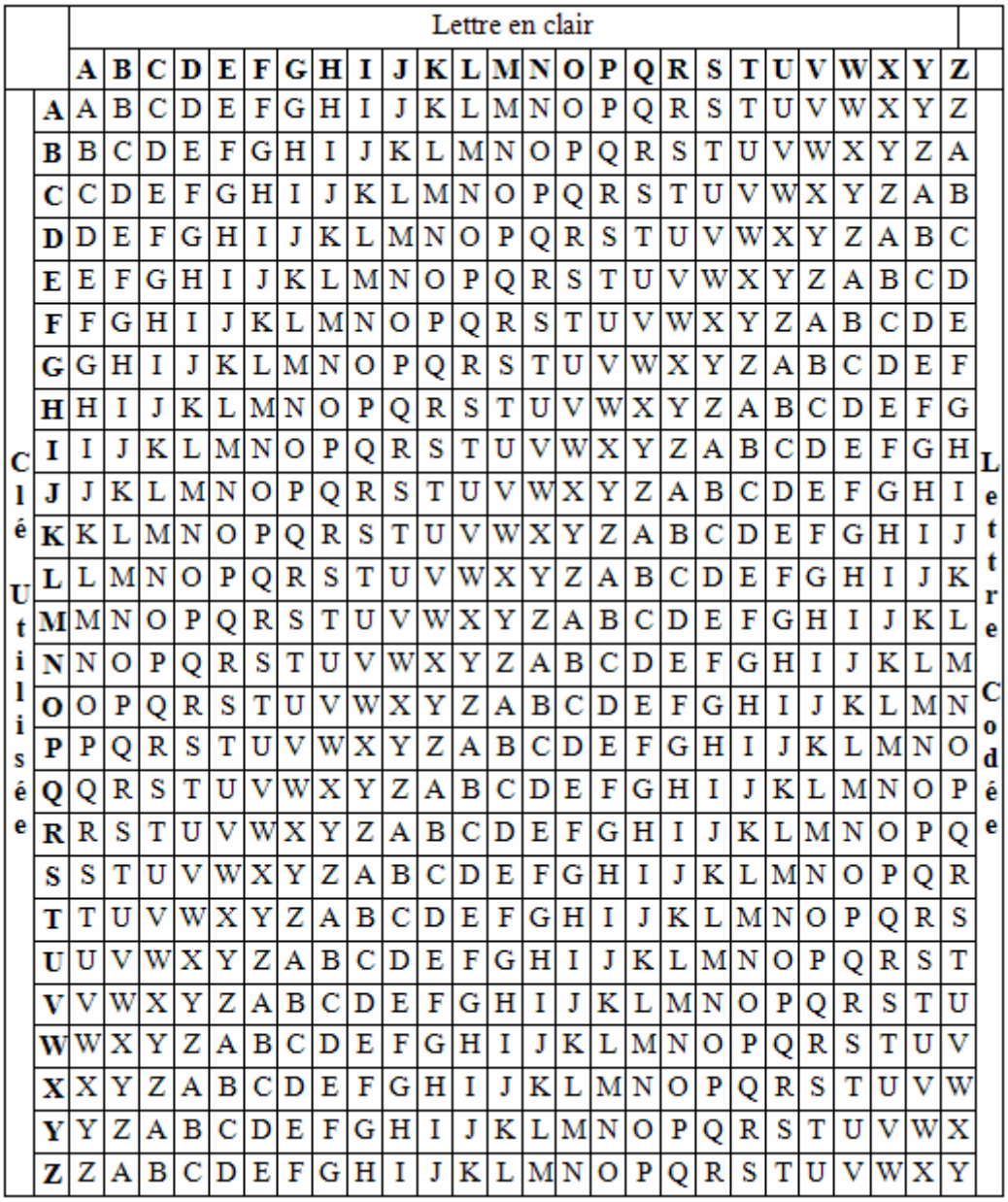
T	S	I	J	A	M	V		H	O	A	K	M	Z	U	E
S	E	V	A	N	S	E		V	A	N	S	E	V	A	N

➔ Situer la ligne de la lettre de la clé (lettre « S » ➔ ligne S)

➔ Sur cette ligne, chercher la colonne de la lettre du texte codé (on cherche « T » ➔ colonne B). On obtient ainsi la première lettre.

➔ Le décodage du texte donne : « Bonjour Monsieur ».

Pour améliorer le cryptage, on peut choisir plusieurs clés de (dé) cryptage, et les associer les uns après les autres pour le chiffrement.



Clé Utilisée \ Lettre en 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	Lettre Codée
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	

Figure 9 : Tableau de Vigenère

→ En 1623, le philosophe **Francis Baron** expose dans un ouvrage « *De dignitate et augmentis scientiarum* » une technique **stéganographique** de codage. Ce procédé consiste à représenter chaque lettre du texte en clair par un groupe de cinq lettres A ou B. Le texte crypté est ainsi constitué d'une succession de ces deux lettres. C'est l'équivalent à un codage binaire des lettres de l'alphabet sur cinq bits, avant gardiste du code ASCII, actuellement utilisé en informatique mais sur huit bits.

→ En 1854, **Charles Wheatstone**, pionnier du télégraphe, apporte sa contribution à la cryptologie en inventant le chiffrement de **Playfair**, du nom de celui qui l'a fait connaître. Cette technique est basée sur une méthode de substitution digrammatique consistant à remplacer un couple de lettres adjacentes par un autre couple choisi dans une grille qui constitue la clé.

→ En 1883, un cryptologue militaire néerlandais, appelé **Auguste Kerckhoffs** publie un livre sur la cryptologie « *La cryptographie militaire* ». Il y expose notamment quelques règles à respecter pour concevoir un bon système cryptographique, toujours valable actuellement, dont la principale est la suivante : la sécurité d'un système ne doit pas reposer que sur le secret de la méthode de cryptage.

4) Le chiffre de Delastelle

L'inventeur de ce système est un français nommé **Félix-Marie Delastelle**. Il utilise une grille de chiffrement/déchiffrement (Figure 10) analogue à celle du chiffre de Polybe. C'est un mélange de codage par substitution et par transposition.

Voici la méthode de chiffrement :

➔ Tout d'abord, il faut commencer par établir le tableau de Polybe : chaque lettre possède deux coordonnées : A(1:1), M(3:2)…

	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

Figure 10 : Exploitation du tableau de Polybe

➔ Si l'on souhaite écrire « Sevan et Louis », on regroupe les lettres du message clair 5 par 5 (au besoin, on rajoute des nulles pour arriver à un multiple de cinq).

Premier groupe de cinq : Sevan

S	E	V	A	N
4	1	5	1	3
3	5	1	1	3

Deuxième groupe de cinq : et

E	T			
1	4	0	0	0
5	4	0	0	0

Troisième groupe de cinq : louis

L	O	U	I	S
3	3	4	2	4
1	4	5	4	3

➔ Après transcription on obtient : 4315511133 1544 3134452443

➔ Ensuite on change l'ordre des chiffres (un sur deux) : 4151335113 1454 3342412543

En définitive, le chiffre de Delastelle diffère peu de celui de Polybe, il est plus qualifié de simple adaptation. Il est plus difficile à décrypter que les autres codes par substitution. Mais cette difficulté repose pour une bonne part sur la connaissance de l'algorithme employé.

C) L’utilisation d’Enigma lors de la Seconde Guerre Mondiale

Dans ce paragraphe, nous allons aborder les aspects de la « guerre de l’intelligence », concernant le monde de cryptographie entre alliés et Allemands au cours de la Seconde Guerre Mondiale. Nous évoquerons l’utilisation de la machine **Enigma** mise à profit par les Allemands.

1) Origine d’Enigma

Lors de la Seconde Guerre Mondiale, **Enigma** bouleversa le monde de la cryptographie en assurant la confidentialité des communications allemandes, sur laquelle le III^{ème} Reich comptait dans la mise en œuvre de sa stratégie. Cette machine permettait aux Allemands de crypter les messages militaires, en se basant sur un dispositif électromécanique de chiffrement. Les Allemands s’étaient équipés de 30 000 machines au début de la guerre, pour un total de 200 000 machines construites à la fin de la guerre.

Arthur Scherbius, inventeur Allemand, créa Enigma en 1918 (déposant un brevet par la même occasion (Figure 11)). Il ne put proposer sa machine en raison de l’Armistice de 1918 mais créa à Berlin une société destinée à fabriquer et à commercialiser sa machine. Sa société fut un réel fiasco mais intéressa tout particulièrement le monde de la finance et celui des militaires.

Figure 11: Logo et machine Enigma

2) Caractéristiques techniques d’Enigma

Enigma est une **machine électromécanique portable**, ce qui signifie qu’elle utilise une combinaison de parties mécaniques et électriques. La partie mécanique est composée d'un clavier (Figure 12), d'un jeu de disques rotatifs adjacents appelés *rotors* arrangés le long d'un axe, et d'un mécanisme entraînant en rotation un ou plusieurs des rotors chaque fois qu'une touche est pressée (Figure 13).

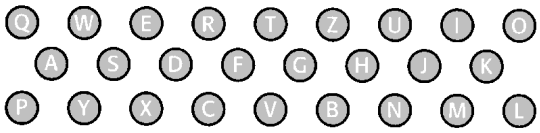


Figure 12: Le clavier d'Enigma

Les **rotors** forment le cœur de la machine Enigma. D'un diamètre d'environ 10 cm, chaque rotor est un disque fabriqué à partir de caoutchouc dur ou de bakélite. Sur une face sont disposés en cercle des contacts électriques à aiguilles, donc équipés de ressorts. Sur l'autre face, le même nombre de contacts plats sont disposés. Les contacts plats et à aiguilles représentent l'alphabet — généralement les 26 lettres (on fera cette supposition dans la suite de cette description, même si parfois des chiffres pouvaient être utilisés suivant le même principe).

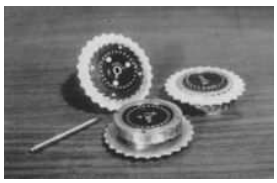


Figure 13 : Trois rotors et l'axe sur lequel ils sont assemblés

Une fois les rotors assemblés, les contacts à aiguilles d'un rotor se positionnent en face des contacts plats du rotor voisin, formant ainsi la connexion électrique. À l'intérieur du rotor, un ensemble de 26 câbles électriques assurent les connexions entre les contacts à aiguilles et les contacts plats suivant un schéma compliqué, qui permet l'enchiffrement des lettres.

Le rotor utilisé tout seul ne réalise en fait qu'un enchiffrement très simple — le chiffrage par substitution. **Par exemple**, le contact à aiguille correspondant à la lettre E peut être connecté sur la face opposée au contact plat de la lettre T. La complexité de la machine Enigma provient de l'utilisation de plusieurs rotors en série, généralement trois ou quatre, ainsi que le mouvement régulier de ces rotors.

Après avoir pressés 26 lettres, une came entraîne le rotor de la fente suivante et le fait tourner. L'alphabet de substitution est ainsi modifié avec chaque pression de touche. Ces deux mécanismes forment un type d'enchiffrement beaucoup plus performant. **Par exemple**, lorsqu'un message commence par la séquence ANX..., l'opérateur devra presser la touche A et la lampe Z pourrait s'allumer ; dans ce cas, Z sera la première lettre du texte chiffré. L'opérateur procède alors au chiffrage de la lettre N de la même manière, et ainsi de suite pour les lettres suivantes.

3) Méthode de chiffrement

Enigma dispose de 26 lettres de l'alphabet, les signes de ponctuations sont codés par des lettres peu fréquentes en Allemand comme X et Y (X correspondant à un point et Y à une virgule). Ensuite, les nombres sont écrits en toutes lettres, et les caractères accentués ä, ö et ü, fréquents dans la langue Allemande sont remplacés par AE, OE et UE, et CH est remplacé par Q.

Le cryptage des textes requérait **deux opérateurs** de manière à éviter les erreurs de saisie et à gagner du temps.

- ✓ L'un chargé du **cryptage lettre à lettre** en ayant préalablement réglé la machine sur une clé de trois lettres qu'il choisit et garde en mémoire.
- ✓ L'autre chargé du **cryptage de la clé** qu'il a gardé en mémoire à son tour, en se servant d'une clef du jour partagée par tous les opérateurs de son armée, et répertoriée sur des documents d'instructions mensuels pour le réglage de la machine.

4) Points forts et faiblesse

L’avantage d’Enigma réside sur le nombre de clés à crypter et de la réversibilité du message codé. Avec la même clé secrète initiale, on tape le message clair, on obtient le message codé, et avec le message codé, on obtient le message clair. Rapide, Enigma était plus simple à utiliser que d’autres procédés de cryptages puisqu’elle évitait aux opérateurs l’effort qu’aurait nécessité un cryptage avec crayon et papier. Petite, elle était transportée aisément, et n’entravait pas la mobilité des unités militaires.

L'une des failles de la machine Enigma est que jamais la lettre A n’est codée par un A. Cela élimine un certain nombre de cas à inspecter. Une des autres faiblesses dépend plutôt du protocole utilisé par les Allemands : certains opérateurs - par exemple, ceux qui informaient de la météo - prenaient peu de précautions et commençaient toujours leurs messages par les mêmes mots (typiquement « Mon général... »). Les Britanniques connaissaient ainsi pour une partie du message à la fois le texte clair et le texte codé, ce qui aide à retrouver la clé. Et comme c'est la même clé qui sert pour toutes les machines Enigma de l'armée allemande pour un jour donné, une erreur de protocole dans un message peut compromettre la sécurité de tous les autres!

5) Fin d’Enigma

Le code d’Enigma a en fait été cassé dès 1933 par des mathématiciens polonais aidés de moyens électromécaniques (« bombes »). Les Français achètent les clés mensuelles d'Enigma, aussitôt transmises aux Polonais.

Il ne suffisait pas de décrypter toutes les communications de l'ennemi : il fallait aussi que celui-ci l'ignore, faute de quoi il aurait progressivement changé de matériel codant. La destruction de chaque navire allemand dont la position était connue était donc précédée de l'envoi d'un avion de reconnaissance censé passer là par hasard. Celui-ci prenait soin de bien se faire voir, et l'attaque pouvait ensuite être lancée sans alerter l'État-Major ennemi (qui devait se dire que les Alliés avaient soit beaucoup de chance, soit beaucoup d'avions !). En revanche, là où l'on ne pouvait envoyer ces avions pour des raisons d'autonomie, les Alliés devaient laisser partir des convois dont ils savaient qu'ils n'arriveraient jamais : les annuler eût été alerter immédiatement l'ennemi.

6) Utilisation moderne d’Enigma

Un étudiant s'amusa un jour à programmer en langage C la simulation du fonctionnement d'une machine Enigma. Ce programme fut inclus dans les distributions UNIX sous le nom de **crypt** (utilisable comme une commande UNIX). Jusqu'à la déclassification des travaux du groupe de Bletchley Park, les bureaux d'études d'engineering croyaient ce codage très sûr et l'utilisaient pour s'échanger leurs informations confidentielles. Pour la plus grande joie, n'en doutons pas, de la National Security Agency, qui en voyait son travail considérablement facilité !

Le codage effectué par la machine Enigma est à la fois simple et astucieux. Chaque lettre est remplacée par une autre, l'astuce est que la substitution change d'une lettre à l'autre.

D) L’utilisation de la Cryptographie de nos jours

De nos jours, la cryptologie est omniprésente lorsqu’il s’agit de protéger des informations. Son emploi devient indispensable lorque l’on souhaite faire transiter des informations confidentielles (identité, authentification, informations bancaires…) mais également lorsque l’on souhaite crypter des communications (le cas des militaires). Maîtriser les concepts liés à la cryptologie devient dorénavant indispensable pour toute personne travaillant dans l’informatique. Nous allons, ici, vous présenter les différentes situations où le cryptage d’information est indispensable.

1) Apparition d’un nouveau commerce

L'apparition d'un réseau mondial (Internet) facilement accessible permet d'envisager de nombreuses nouvelles applications et en particulier le développement d'une nouvelle forme de commerce que l'on qualifie d'électronique. L'Internet n'est pas à l'abri d'actions d'espionnage ou de malveillance d'où l'urgence actuelle pour les informaticiens de mettre en place des procédures de sécurisation.

Il faut ajouter à cela la sécurité des terminaux mobiles, qu’il s’agisse d’ordinateurs portables, de smartphone, PDA... En effet, si la perte ou le vol d’un terminal est problématique, c’est bien plus en raison des données qu’il peut contenir que pour le prix de ce dernier. Le cryptage apparaît là, comme une des solutions pour se prémunir contre le vol matériel ou numérique de ces informations souvent confidentielles.

Grâce à la puissance de calcul offerte par les actuels et futurs processeurs des micro ordinateurs, la cryptologie s’oriente sur une phase numérique grâce a l’apparition entre autre du code ASCII. Le cryptosystème s’oriente désormais sur la puissance des nombres très grands et la mise en place de clés pour ouvrir les codes.

2) Sécurisation des réseaux Informatiques

L’avènement du monde « Internet » a apporté une multitude de différents types de flux d’informations parmi les réseaux privés et publics. Lorsque l’on fait transiter des données sur Internet il est nécessaire de les crypter, d’authentifier les interlocuteurs, et de vérifier l’intégralité des informations transmises.

Le cryptage peut s’effectuer sur deux niveaux : il peut se réaliser par des **logiciels de chiffrement** ou par des **plateformes matérielles**. La solution logicielle est certes fiables mais présente certains inconvénients dont notamment la complexité de leur gestion (politique de mot de passe rigoureuse). La solution matérielle est quant à elle moins complexe mais nécessite la présence d’un tiers ordinateur s’occupant des algorithmes de cryptage.

Le cas des réseaux **WiFi** nouvellement développés a forcé les informaticiens à élaborer une méthode de chiffrage fiable et ne nécessitant que très peu de ressources matérielles. Les clés WEP et WPA actuellement utilisées pour rendre privée des accès à Internet ont été normalisées afin de respecter une harmonisation Mondiale en terme de sûreté.

3) Cryptage des communications militaires

Les militaires font la distinction parmi deux types de conversations : celle pour un échange momentané de lettres entre quelques personnes isolées, et celle pour la correspondance des différents chefs d’armée entre eux pour une durée illimité. Ceux-ci, ne peuvent, à leur gré modifier leurs conventions de cryptage ; de plus, ils ne doivent jamais garder sur eux aucun objet ou écrit qui soit de nature à éclairer l’ennemi sur le sens des dépêches secrètes qui pourraient tomber entre ses mains.

Il faut que certaines conditions soient appliquées pour le cryptage de message militaire :

- ✓ Le système doit être matériellement, sinon mathématiquement indéchiffrable
- ✓ Il faut qu'il n'exige pas le secret, et qu'il puisse tomber entre les mains de l'ennemi sans inconvénients
- ✓ La clé doit pouvoir en être communiquée et retenue sans le secours de notes écrites, et être changée ou modifiée au gré des correspondants
- ✓ Il faut qu'il soit applicable à la correspondance télégraphique
- ✓ Il faut qu'il soit portatif, et que son maniement ou son fonctionnement n'exige pas le concours de plusieurs personnes
- ✓ Il ne doit n'exiger aucune tension d'esprit, ni connaissance d'une longue série de règles à observer.

Il suffit en un mot que le système soit combiné de telle façon que la traduction d'un cryptogramme exige au moins trois à quatre heures de travail.

4) Vente par correspondance

La généralisation des offres commerciales sur Internet par l'intermédiaire des transactions bancaires et autres identifications électroniques a amené les informaticiens à rendre la communication « client-entreprise » très secrète. A chaque nouveau système de cryptage, apparaît un nouveau système de décryptage, car il faut le rappeler la sécurité des réseaux et télécommunications n'existe que pour contrecarrer les individus malveillants.

5) Signature électronique

« La signature électronique permet, à l'aide d'un procédé cryptographique, de garantir l'intégrité du document signé et l'identité du signataire. » (DCSSI : Direction Centrale de la Sécurité des Systèmes d'Information). En d'autres termes, elle permet de garantir qu'un document électronique est, comme pour un original papier, le document authentique considéré.

Pour être fiable, un document doit apporter certaines garanties : l'expéditeur doit être authentifié, le document ne doit être ni modifié ni altéré, s'assurer de la date et de l'heure de création du document et vérifier que le destinataire a bien reçu le document. C'est la signature électronique qui répond à ces besoins.

Nous allons à présent nous pencher sur les deux façons actuelles de crypter un message en fonction du type de clé. De plus, nous nous intéresserons à une technique de cryptage en particulier qui se présente sous la dénomination : « Sac à dos »...

III. Notions de cryptographie et problème du « Sac à dos »

A. Chiffrement asymétrique ou à clef publique

Avec les algorithmes asymétriques, **les clefs de chiffrement et de déchiffrement sont distinctes et ne peuvent se déduire l'une de l'autre**. On peut donc rendre l'une des deux publique tandis que l'autre reste privée. C'est pourquoi l'on parle de chiffrement à clef publique. Si la clef publique sert au chiffrement, tout le monde peut chiffrer un message que seul le propriétaire de la *clef privée* pourra déchiffrer. On assure ainsi la confidentialité. Certains algorithmes permettent d'utiliser la clef privée pour chiffrer. Dans ce cas, n'importe qui pourra déchiffrer, mais seul le possesseur de la clef privée peut chiffrer. Cela permet donc la signature de messages.

Le concept de cryptographie à clef publique fut inventé par **Whitfield Diffie** et **Martin Hellman** en 1976, dans le but de résoudre le problème de distribution des clefs posé par la cryptographie à clef secrète. De nombreux algorithmes permettant de réaliser un cryptosystème à clef publique ont été proposés depuis. Ils sont le plus souvent **basés sur des problèmes mathématiques difficiles à résoudre**, donc leur sécurité est conditionnée par ces problèmes, sur lesquels on a maintenant une vaste expertise. Mais, si quelqu'un trouve un jour le moyen de simplifier la résolution d'un de ces problèmes, l'algorithme correspondant s'écroulera.

Le nombre des algorithmes proposés pour la cryptographie à clef publique se sont révélés rapidement non sûrs, ou non réalisables sur le plan pratique. Tous les algorithmes actuels présentent l'inconvénient d'être **bien plus lents que les algorithmes à clef secrète**. Certains algorithmes asymétriques ne sont adaptés qu'au chiffrement, tandis que d'autres ne permettent que la signature. Seuls trois algorithmes sont utilisables à la fois pour le chiffrement et pour la signature : RSA, ElGamal et Rabin.

RSA (Rivest, Shamir, Adleman)

Inventé par Rivest, Shamir et Adleman en 1978, RSA permet le chiffrement et la signature. Il est aujourd'hui encore très largement utilisé. Cet algorithme repose sur la difficulté de factoriser des grands nombres.

Voici comment se fait la génération des paires de clefs :

1. On commence par choisir deux grands nombres premiers, p et q, et on calcule $n = pq$. n est rendu public ; p et q doivent rester secrets.
2. On choisit ensuite aléatoirement une clef publique e telle que e et $(p-1)(q-1)$ soient premiers entre eux.
3. La clef privée d est obtenue grâce à l'algorithme d'Euclide : $ed \equiv 1 \bmod (p-1)(q-1)$.

Soit m le message en clair et c le cryptogramme. La fonction de chiffrement est, de façon simplifiée, $c = m^e \bmod n$ (si m est plus grand que n, il est séparé en morceaux de valeur inférieure à n et chaque morceau est chiffré séparément suivant cette formule). Du fait de la relation entre e et d, la fonction de déchiffrement correspondante est $m = c^d \bmod n$. La signature se fait de manière similaire, en inversant e et d, c'est-à-dire en chiffrant avec une clef privée et en déchiffrant avec la clef publique correspondante : $s = m^d \bmod n$ et $m = s^e \bmod n$.

Pour un cryptanalyste, retrouver la clef privée à partir de la clef publique nécessite de connaître $(p-1)(q-1) = pq-p-q+1 = n+1-p-q$, donc de connaître p et q. Pour cela, il doit factoriser le grand nombre n. Donc n doit être suffisamment grand pour que cela ne soit pas possible dans un temps raisonnable par rapport au niveau de sécurité requis. Actuellement, la longueur du module n varie généralement de 512 à 2048 bits suivant les utilisations. Compte tenu de l'augmentation des vitesses de calcul des ordinateurs et des avancées mathématiques en matière de factorisation des grands nombres, la longueur minimale des clefs (Figure 144: RSA) doit augmenter au cours du temps.

Clefs

Clef publique	$n = pq$, où p et q sont deux grands nombres premiers tenus secrets e telle que e et $(p-1)(q-1)$ soient premiers entre eux		
Clef privée	$d \equiv e^{-1} \bmod (p-1)(q-1)$		
Algorithmes			
Chiffrement	$c = m^e \bmod n$	Déchiffrement	$m = c^d \bmod n$
Signature	$s = m^d \bmod n$	Vérification	$m = s^e \bmod n$

Figure 144: RSA

B. Générateurs aléatoires et pseudo-aléatoires

La cryptographie a souvent recours à des nombres aléatoires. Ainsi, lorsqu'une personne génère une clef secrète ou privée, elle doit faire intervenir le hasard de façon à empêcher que la clef puisse être deviné.

Malheureusement, il est impossible de produire des suites aléatoires à l'aide uniquement d'un ordinateur : le générateur sera toujours périodique, donc prévisible. On a donc recours à des générateurs dits pseudo-aléatoires et cryptographiquement sûrs. Un tel générateur doit présenter les caractéristiques suivantes :

1. La période de la suite doit être suffisamment grande pour que les sous-suites finies utilisées avec l'algorithme ou le protocole cryptographique ne soient pas périodiques.
2. Ces sous-suites doivent, sur le plan statistique, sembler aléatoires.
3. Le générateur doit être imprévisible, au sens où il doit être impossible de prédire le prochain aléa à partir des aléas précédents.

La plupart des générateurs pseudo-aléatoires sont construits en utilisant des registres à décalage (*shift registers* en anglais) et, en particulier, les registres à décalage à rétroaction linéaire (*Linear Feedback Shift Registers, LFSR)*. Ces derniers présentent l'inconvénient de générer des suites linéaires, si bien que des grands nombres générés à partir de sous-suites sont fortement corrélés. C'est pourquoi les générateurs pseudo-aléatoires sont généralement construits en combinant, à l'aide d'une fonction non linéaire, plusieurs registres à décalage de tailles différentes. Ce type de générateur est très utilisé par les algorithmes de chiffrement en continu.

Si l'on veut vraiment générer des suites aléatoires, au sens où ces suites sont de plus non reproductibles, on a généralement recours à des éléments extérieurs comme les déplacements de la souris, la vitesse de frappe, l'entrée d'un micro enregistrant le bruit atmosphérique,...

C. Fonctions de hachage à sens unique

Aussi appelée fonction de condensation, une **fonction de hachage** est une fonction qui convertit une chaîne de longueur quelconque en une chaîne de taille inférieure et généralement fixe ; la chaîne résultante est appelée *empreinte* (*digest* en anglais) ou condensé de la chaîne initiale.

Une **fonction à sens unique** est une fonction facile à calculer mais difficile à inverser. La cryptographie à clef publique repose sur l'utilisation de fonctions à sens unique à brèche secrète : pour qui connaît le secret (i.e. la clef privée), la fonction devient facile à inverser.

Une **fonction de hachage à sens unique** est une fonction de hachage qui est en plus une fonction à sens unique : il est aisé de calculer l'empreinte d'une chaîne donnée, mais il est difficile d'engendrer des chaînes qui ont une empreinte donnée, et donc de déduire la chaîne initiale à partir de l'empreinte. On demande généralement en plus à une telle fonction d'être **sans collision**, c'est-à-dire qu'il soit impossible de trouver deux messages ayant la même empreinte. On utilise souvent le terme fonction de hachage pour désigner, en fait, une fonction de hachage à sens unique sans collision.

La plupart des fonctions de hachage à sens unique sans collision sont construites par itération d'une fonction de compression : le message *M* est décomposé en *n* blocs $m_1,...,m_n$, puis une fonction de compression *f* est appliquée à chaque bloc et au résultat de la compression du bloc précédent ; l'empreinte *h(M)* est le résultat de la dernière compression.

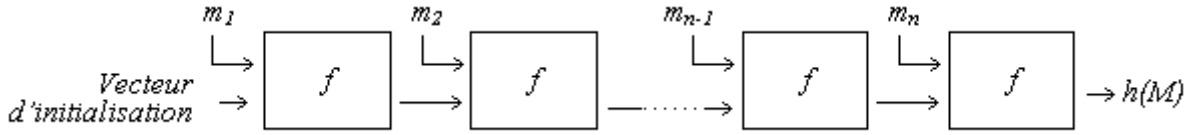


Figure 155: Fonction de hachage itérative

Des exemples de fonctions ainsi conçues sont MD5, SHA et RIPE-MD.

MD5 (Message Digest 5)

Développé par Rivest en 1991, MD5 produit une empreinte de 128 bits à partir d'un texte de taille arbitraire. MD5 manipule le texte d'entrée par blocs de 512 bits.

SHA (Secure Hash Algorithm)

SHA est la fonction de hachage utilisée par SHS (*Secure Hash Standard*), la norme du gouvernement Américain pour le hachage. SHA-1 est une amélioration de SHA publiée en 1994. SHA-1 produit une empreinte de 160 bits à partir d'un message de longueur maximale 2^{64} bits. Tout comme MD5, SHA-1 travaille sur des blocs de 512 bits.

RIPE-MD

Développée dans le cadre du projet RIPE (*RACE Integrity Primitives Evaluation*) de la communauté Européenne, RIPE-MD fournit une empreinte de 128 bits. RIPE-MD-160 est une version renforcée de RIPE-MD qui fournit une empreinte de 160 bits.

C. Le problème du sac à dos

1) Problème

En 1978, **Ralph Merkle** et **Martin Hellman** proposèrent un cryptosystème à clé publique basé sur un problème célèbre, le problème du sac à dos (*Knapsack problem*).

Imaginons une collection de cailloux de poids $a_1, a_2, ... , a_n$ connus.

Supposons que l'on place certains de ces cailloux dans un sac à dos et que l'on pèse le tout.

Est-il possible, connaissant ce poids total, de savoir quels cailloux sont dans le sac ?

Le problème peut s'exprimer ainsi :

Etant donné une suite d'entiers positifs $S=(a_1, a_2, ... , a_n)$ et un nombre entier s,

existe-t-il $(x_1, x_2, ... , x_n)$ tel que $\sum_{i=1}^{n} x_i a_i = s$, avec x_i égal à 0 ou à 1 ?

Si n est grand, le problème du sac à dos s'avère très difficile à résoudre et la fonction qui à un ensemble (x_1, x_2, ... , x_n) associe s est une fonction à sens unique, s est facile à calculer et retrouver (x_1, x_2, ... , x_n) à partir de s est difficile.

Martin Hellman propose l'exemple suivant :
S = (14, 28, 56, 82, 90, 132, 197, 284, 341, 455)
s = 516 ne peut s'obtenir comme somme d'élément de S.
s = 515 peut s'obtenir trois fois (dont 515=14+28+132+341=1·14+1·28+0·56+0·82+0·90+1·132+0·197+0·284+1·341+0·455)
Il ne semble pas exister d'algorithme rapide permettant de résoudre le problème du sac à dos.

2) Le cas d'une suite super-croissante

Dans un cas cependant, le problème du sac à dos ne présente aucune difficulté, celui où les éléments de S forment une suite super-croissante (chaque élément est supérieur à la somme des éléments précédents).
Prenons, par exemple, la suite super-croissante S = (2, 5, 9, 20, 42, 90, 250) et *s* = 56.
42 est le plus grand élément de S inférieur à s.
42 doit intervenir dans le calcul de *s*.
Sinon, ou la somme comporte un élément supérieur à *s* et elle est supérieur à 56, ou la somme ne comporte que des éléments inférieurs à 42 et elle est inférieure à 56.

Le solde vaut 14.
9 étant le plus grand élément de S inférieur à 14, il doit, selon le même raisonnement, intervenir dans le calcul de *s*.
Le nouveau solde étant 5, le problème est résolu : 56 = 5 + 9 + 42.

3) Le chiffre de Merkle-Hellman

L'idée de base du système consiste à construire une suite non super-croissante à partir d'une suite super-croissante, en conservant une clé secrète permettant de retrouver la suite initiale.
Alice choisit une suite super-croissante S= (a_1, a_2, ... , a_n),

un nombre m supérieur à $\sum_{i=1}^{n} a_i$ et un entier e, 1 < e < m, premier avec m.
Pour chaque élément a_i de S, Alice calcule $b_i = a_i \cdot e \mod m$.
Elle ordonne les éléments b_i dans l'ordre croissant pour obtenir une nouvelle suite S' = (b_1, b_2, ... , b_n) qui n'est plus super-croissante.

Elle publie cette suite en conservant comme clés secrètes m, l'inverse d de e modulo m, la suite super-croissante S et la permutation $\pi \in$ ayant permis d'ordonner S'.

Pour chiffrer un message, Bernard le représente en code binaire et le décompose en blocs de longueur n au plus.

Pour chaque bloc m_1 m_2 .... m_n, il calcule M = $\sum_{i=1}^{n} m_i b_i$.

Pour déchiffrer le message M reçu, Alice calcule M' = M · d modulo m et détermine $x_1, x_2, \ldots, x_n$ tels que $\sum_{i=1}^{n} x_i a_i = M'$ (Il s'agit d'un problème simple de sac à dos, la suite $(a_1, a_2, \ldots, a_n)$ étant super-croissante).

Elle retrouve finalement le message $m_1 m_2 \ldots m_n$ en appliquant à $x_1, x_2, \ldots, x_n$ la permutation [1].

Exemple :

Alice choisit S= (1, 3, 5, 11, 25, 53, 101, 205, 512), m = 960 et e = 143.
(l'inverse d de 143 mod 960 est 47)
Pour chaque élément a_i de S, Alice calcule $b_i = a_i \cdot e \mod m$, ce qui donne
(143, 429, 715, 613, 695, 859, 43, 515, 256).

En ordonnant b_i, elle obtient la clé publique S' = (43, 143, 256, 429, 515, 613, 695, 715, 859) .

Pour exprimer le message « RAS » en code binaire, Bernard peut, par exemple, utiliser le code ASCII à 8 bits.

R correspond à 01010010, A à 01000001 et S à 01010011.

Le message à coder est 0101001 0010000 0101010 011.

Il le décompose en blocs de longueur convenue (7 par exemple) et chiffre chacun des blocs :

0101001 se code 43 + 429 + 613 = 1085,

0010000 se code 515,

0101010 se code 143 + 429 + 613 = 1185,

011 s'écrit 0110000 et se code 515 + 613 = 1128.

Il transmet à Alice le message 108 515 1185 1128

Alice va déchiffrer ce message élément par élément, en calculant M · d mod m et en déterminant la solution du problème du sac à dos correspondant.

1085 · 47 mod 960 = 115

515 · 47 mod 960 = 205

1185 · 47 mod 960 = 15

1128 · 47 mod 960 = 216

115 = 101+11+3 correspond à 0000001 + 0100000 + 0001000 = 0101001

205 = 205 correspond à 0010000

15 = 11+3+1 correspond à 0100000 + 0001000 + 0000010 = 0101010

216 = 205+11 correspond à 001000 + 0100000 = 0110000

Alice retrouve le message 0101001001000001010100110000 : RAS.

Remarquons que si la longueur des blocs de chiffrement est égale à celle des caractères en code ASCII à 8 bits, chaque lettre sera codée par le même nombre. Le système est alors vulnérable à une attaque à l'aide d'une analyse de fréquence. Il convient donc de choisir des blocs de chiffrement de longueur inférieure à celle de la clé. Le système a été cassé en 1982 par Adi Shamir et n'est plus utilisé de nos jours.

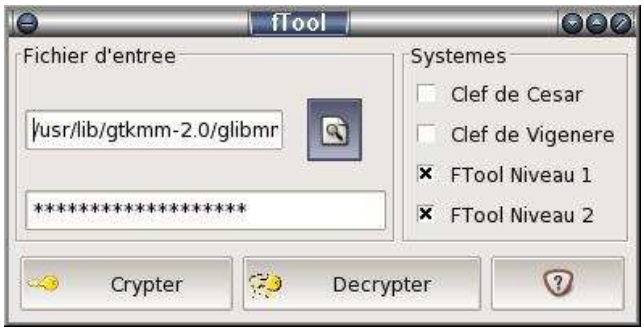
III. Exemples d'algorithmes de cryptage

Nous allons dans cette dernière partie vous montrer l'ingéniosité des programmeurs quant au cryptage de message à l'aide de l'outil informatique. Différents programmes vous seront présentés, accompagnés d'interface pour la plupart et de référencement des auteurs. Concernant les codes sources (l'écriture même des programmes), ils ne seront pas montrés car généralement très long, en partie à cause de la complexité des fonctions de cryptages.

1) Cryptage en langage C

Le Langage C est le langage de programmation qui a révolutionné le genre et qui est utilisé par des millions de personnes dans le monde. Il permet de développer des applications à la fois complexes et rapides, tout en générant par compilation un code machine efficace. L'utilisation de librairies permet d'avoir le contrôle quasi-total de la machine et du système d'exploitation.

Un programmeur nommé **Naopic Tastalian** a crée un site Web (http://tastalian.free.fr/) mettant en avant des cours de C, de C++ et de programmes dans les mêmes langages en libre service pour toute personne intéressé. Nous avons trouvé un programme nommé **FTOOL rédigé en langage C** qui est un outil de cryptographie de fichier : il permet d'encrypter des fichiers écrits sur un support accessible en utilisant des algorithmes simples et connus (clef de césar, vigenère...) ou des algorithmes crées par **Tastalian** (fTool lv1, lv2...) Il est également possible de cumuler l'utilisation de ces différents systèmes et crypter/décrypter plusieurs fois les mêmes fichiers pour plus de sécurité. Muni d'une interface convivial () il est estampillé à la version 0.2. Il est compatible Windows et Linux.



L'utilisation de ce « mini-logiciel » est simple. Il faut commencer par spécifier un fichier à crypter/décrypter, puis de spécifier un mot de passe et enfin de choir la méthode de (dé) chiffrement (ou de les cumuler).

Voici les fichiers nécessaires au fonctionnement de l'application :

Nom	Taille	Type
fTool.exe	506 Ko	Application
main.cpp	2 Ko	Fichier CPP
fTool.dev	2 Ko	Fichier DEV
codec.hpp	8 Ko	Fichier HPP
debug.hpp	1 Ko	Fichier HPP
interface.hpp	12 Ko	Fichier HPP
crypt.png	1 Ko	Image PNG
decrypt.png	1 Ko	Image PNG
ftool.png	2 Ko	Image PNG

Le premier fichier est un fichier dit « exécutable », celui à lancer pour voir l'interface proposant le cryptage/décryptage. Le fichier « **main.cpp** » est le programme principal (main) faisant appel à d'autres sous-programmes (« **interface.cpp** » pour l'interface, « **codec.hpp** » pour effectuer les fonctions de cryptage pures, « **debug.hpp** » pour avertir d'un bug dans le processus de (dé) cryptage). Le fichier « **ftool.dev** » contient toutes les informations relatives à FTOOL (version, nombre de fichiers, date d'écriture...) compilé ici, il n'est d'aucune utilité pour notre programme. Enfin, les fichiers « **.png** » sont des images rendant l'interface agréable à voir.

A l'utilisation ce type de cryptage est efficace pour des fichiers relativement faible en taille, sa facilité d'utilisation et son niveau de cryptage (en tant qu'utilisateur lambda) en font un programme convaincant en langage C.

2) Cryptage en QuickBasic

Nous allons vous présenter un programme écrit en **QuickBasic** qui est Quick BASIC est une version professionnelle de l'environnement de développement intégré de Microsoft permettant l'édition, le débogage, l'exécution et la compilation du langage BASIC sous MS-DOS.

Le site Internet http://www.mon-pc.net/ propose un petit tutorial (manuel d'utilisation) pour l'écriture de programme en QuickBasic. L'un de ces derniers est un programme de cryptage pas très évolué qui ne permettra pas de crypter des données de manière infaillible (d'ailleurs pratiquement aucun programme ne le fait) mais il permettra plus vraisemblablement de coder des données d'importance moyenne et de comprendre les bases de la cryptographie.

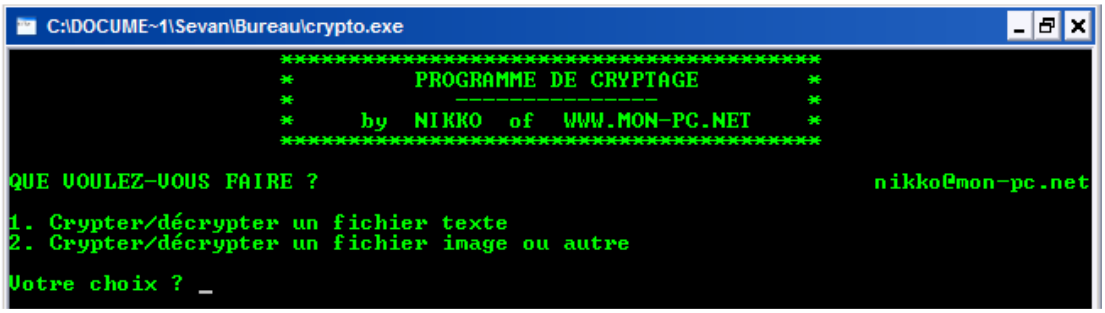
Ce programme en Quick Basic utilise une « clé » : un nombre qui permet le cryptage et le décryptage. Ainsi, le destinataire et le receveur du fichier devront tous deux connaître ce code. Cela permet donc de multiples possibilités pour le codage et une grande difficulté pour « casser » la protection.

Le programme se nomme « Crypto » et se présente en un seul fichier exécutable sous MS-DOS et pèse 48 Ko. Une fois téléchargé, et quelques manipulations de dossiers pour préparer la destination des fichiers cryptés et décrypter, le programme s'exécuté et une fenêtre () s'ouvre à vous et propose deux façons d'effectuer du cryptage :

- ✓ Pour les fichiers textes, il prend le code ASCII de chaque caractère et ajoute à ce code la valeur de la « clé » que vous aurez définie. Ainsi, le texte « aaabbbccc » devient « bbbcccddd » avec la clé 1 (car code ASCII de A + 1 = code ASCII de B ...).

- ✓ Pour les fichiers images, exécutables ou autres, il effectue sur chaque octet du fichier un XOR (ou exclusif) avec la clé que vous avez choisi.

Pour le décryptage, l'opération inverse est effectuée.



Une fois que l'on choisi de crypter un texte, par exemple, il nous demande de saisir une clé (nombre de 1 à 100) puis le message à crypter. Le résultat est enregistré dans un fichier texte de notre choix et de notre nom.

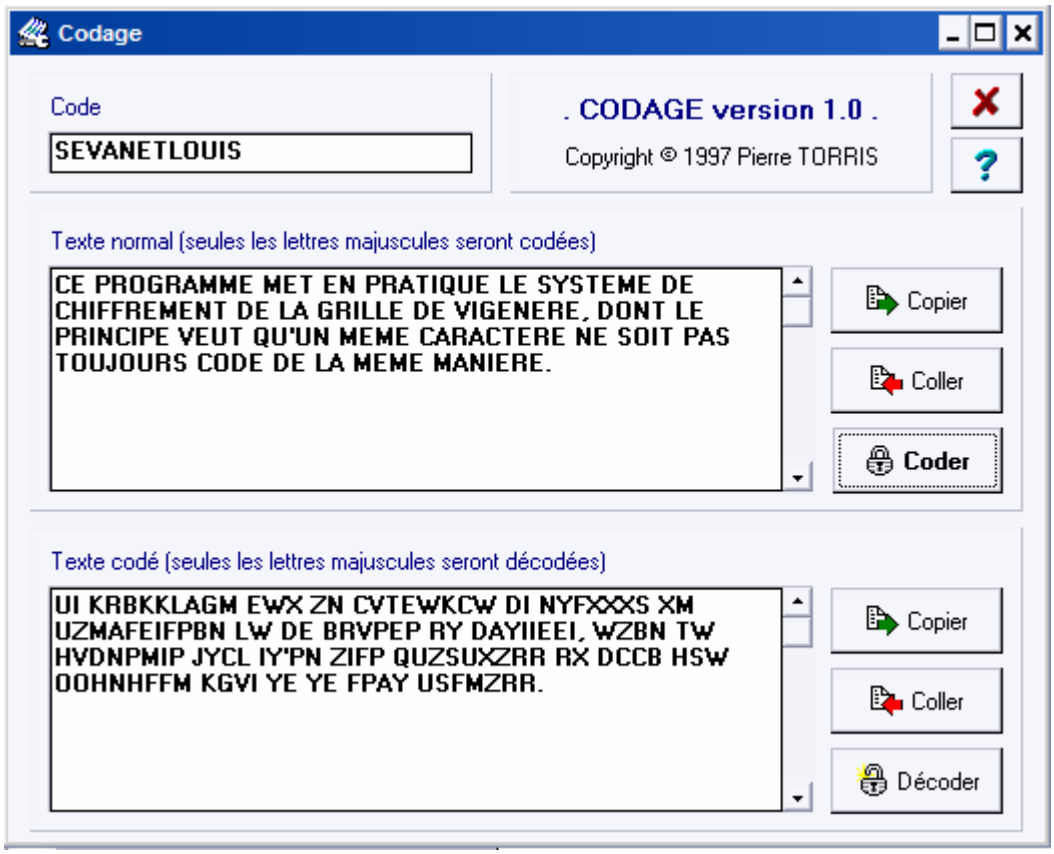
En définitive, ce petit utilitaire est assez pratique au niveau de son utilisation mais le langage utilisé est depuis fort longtemps dépassé et ne permet plus aujourd'hui de rivaliser avec le codage en C/C++ et celui en JAVA.

Il existe évidemment de nombreux langages permettant de crypter et de coder des messages ou fichiers quelconques, mais leur explications demanderaient un long travail.

3) Logiciel de cryptage

De nombreux éditeurs de logiciels mettent à disposition des simples utilisateurs, des programmes dont le code source n'est pas visible mais dont leur distribution est gratuite. Le site Web de **Pierre Torris** (http://www.ptorris.com/index.htm) regroupe de nombreux logiciels de cryptage pour différents systèmes d'exploitation (Windows 3.1X, 9X, Me, XP).

Le mini-logiciel (gratuit) « Codage » mis en ligne a pour rôle de crypter un message texte à l'aide du chiffre de Vigénère, expliqué en partie I. Une fois le fichier d'installation téléchargé (aucune installation n'est nécessaire), et exécuté, une fenêtre s'ouvre et deux champs doivent être remplis pour le fonctionnement du programme : le code et le message en clair ou codé. En effet, le programme effectue également le décryptage, à condition de connaître le mot de passe.



CONCLUSION

A travers cette étude, nous avons constaté que les techniques de cryptage ont beaucoup progressé au cours des âges. Cette évolution est la conséquence des décryptages successifs qui ont poussé l'homme à développer des techniques de plus en plus élaborées. La sensibilité des informations aidant, l'homme fut obliger après avoir longtemps fait confiance à de simples substitutions (disque à chiffrer, Scytale...) à passer à des systèmes plus complexes.

On a pu donc partager l'histoire du cryptage en deux parties l'avant et l'après Enigma. La complexité de la machine allemande a forcé les ingénieurs à créer le premier ordinateur (Colossus) afin de pouvoir décrypter les messages en sa provenance. Dès l'avènement de l'informatique les techniques de cryptage ont franchis un grand pas : l'apparition des premiers réseaux informatiques et la multiplication des échanges de données (identité, authentification) sous forme numérique ont forcé les mathématiciens a repensé entièrement les systèmes de cryptage (RSA, DES, AES).

Le problème du sac à dos est une preuve de la complexité des récentes techniques de cryptage dans le monde informatique. Ce sont tous les scientifiques et chercheurs qui se mettent au travail, il n'est plus question d'un génie qui trouve le miracle d'un cryptage inviolable mais de protocole et de systèmes entiers normalisés qui se mettent en place.

La démonstration de quelques codes de programmes nous également montre qu'aujourd'hui on se retrouve avec des codes accessibles au plus grand nombre et dont la complexité était inimaginable il y a encore cinquante ans. Ce n'est plus un sujet secret mais ce sont des recherches académiques lancées à travers le globe. C'est une démocratisation de la cryptographie qui ne se limite plus aujourd'hui à l'échange de données stratégiques. Malgré toutes les améliorations ces systèmes ne sont toujours pas inviolables. La bataille du cryptage entre ceux qui tentent de protéger les données et ceux qui tentent de les détourner a donc encore de beaux jours devant elle.

Nous restons ébahi par le bouleversement qu'à subit le monde la cryptologie en pensant que chaque évolution technologique en matière de communication affecte profondément l'univers de la cryptographie. Les disques à chiffrer ont remplacés les Scytales, la mécanique a remplacé les disques, l'électronique s'est mêlée à la mécanique, et aujourd'hui l'informatique a révolutionné le monde des réseaux. On peut extrapoler cette idée en se demandant quel sera l'avenir du cryptage sécurisé si un jour, l'informatique venait à être détrôné par une nouvelle technologie ? Les communications numériques cesseront-elles ? Vers quelles formes vont elle tendre ? Car son évolution impactera à jamais le monde la cryptographie...

Annexes

➔ Bibliographie

Références Web :
Cryptage de César : http://www.bibmath.net/crypto/substi/cryptcesar.php3
Cryptage_Poly-alphabétique : http://titan.glo.be/tsf/h6.html
Cryptage_Digrammatique : http://www.apprendre-en-ligne.net/crypto/porta/biographie.html
Playfair : http://www.apprendre-en-ligne.net/crypto/subst/playfair.html
Cryptage par disque à chiffrer : http://dblanc.ifrance.com/crypto/histoir2.htm
Cryptage par le chiffre de Vigenère : http://www.apprendre-en-ligne.net/crypto/vigenere/
Site Web du gouvernement : http://www.ssi.gouv.fr/fr/dcssi/
Site Web sur la cryptographie militaire :
http://www.petitcolas.net/fabien/kerckhoffs/la_cryptographie_militaire_i.htm
Voici un lien pour de plus ample information concernant la machine Enigma (Anglais) :
http://en.wikipedia.org/wiki/Enigma_rotor_details
Très bonne documentation sur Wikipédia :
http://fr.wikipedia.org/wiki/Cryptographie
De grandes pistes données par Ghislaine Labouret :
http://www.labouret.net/crypto/
De la documentation technique concernant le problème du sac à dos:
http://www.gymnase-yverdon.vd.ch/branches/mathematique/cryptographie/nombres/merkle_hellexpl.htm

Références littéraires :
La Cryptographie décryptée de Doris Baker, H.X. Mel, CampusPress.
Introduction à la cryptographie de Johannes A. Buchmann, Dunod.
Cours de cryptographie de Gilles Zemor, Cassini.
Cryptographie : théorie et pratique : de Douglas Stinson (traduction de Serge Vaudenay), Vuibert.

➔ Mots clés

CRYPTOGRAMME : message écrit en chiffre ou sous n'importe quelle forme secrète qui exige une clé pour percer sa signification.

CRIBLE : preuve (le plus souvent un livre de code capturé ou un texte en clair) qui donne des indices permettant de décrypter un pictogramme: « il va sans dire que le crible...est l'outil le plus essentiel de tout cryptologue ».

CAPTURER : dérober du matériel cryptographique à l'ennemi; CAPTURE: tout objet dérobé a l'ennemi qui renforce les chances de percer ses codes ou chiffres.

TEXTE EN CLAIR : texte original intelligible tel qu'il se présentait avant tout chiffrement, révélé après un décodage ou un décryptement réussi.

➔ **Les grands précurseurs en Cryptologie.**

Leone Battista Alberti : Né à Gênes le 18 février 1404. Mort à Rome le 25 avril 1472. D'origine Italienne, il fut peintre, architecte, savant, théoricien de la peinture et de la sculpture, et philosophe italien de la Renaissance. Créateur du disque à chiffre pour le cryptage polyalphabétique.

Giovanni Battista Della Porta : né à Naples vers 1535, mort le 4 février 1615) est un physicien et alchimiste italien.

Francis Baron : (1561-1626), Baron de Verulam, vicomte de St. Albans, homme d'État et philosophe anglais

Blaise de Vigenère, Charles Wheastone, Auguste Kerchkoffs, Félix-Marie Delastelle...

IXe siècle :	Al-Kindi, arabe
XIVe siècle :	Gabriele de Lavinde, italien
XVe siècle :	Leone Battista Alberti, italien
	Abd Allah al-Qalqashandi, égyptien
	Sicco Simonetta, italien
XVIe siècle :	Johannes Trithemius, allemand
	Giovan Batista Belaso, italien
	Giovanni Battista Della Porta, italien
	Blaise de Vigenère, français
XVIIe siècle :	Francis Bacon, anglais
XIXe siècle :	Charles Babbage, anglais
	William F. Friedman, américain
	Friedrich Kasiski, polonais
	Auguste Kerckhoffs, hollandais
	Charles Wheatstone, britannique
XXe siècle :	Eli Biham, israélien
	Joan Daemen, belge
	Whitfield Diffie, américain
	Taher Elgamal, égyptien
	Martin Hellman, américain
	Lester Hill, américain
	Joseph Mauborgne, américain
	Ralph Merkle, américain
	Marian Rejewski, polonais
	Vincent Rijmen, belge
	Ronald Rivest, américain
	Bruce Schneier, américain
	Adi Shamir, israélien
	Alan Mathison Turing, britannique
	Gilbert Vernam, américain
	David Wagner, américain