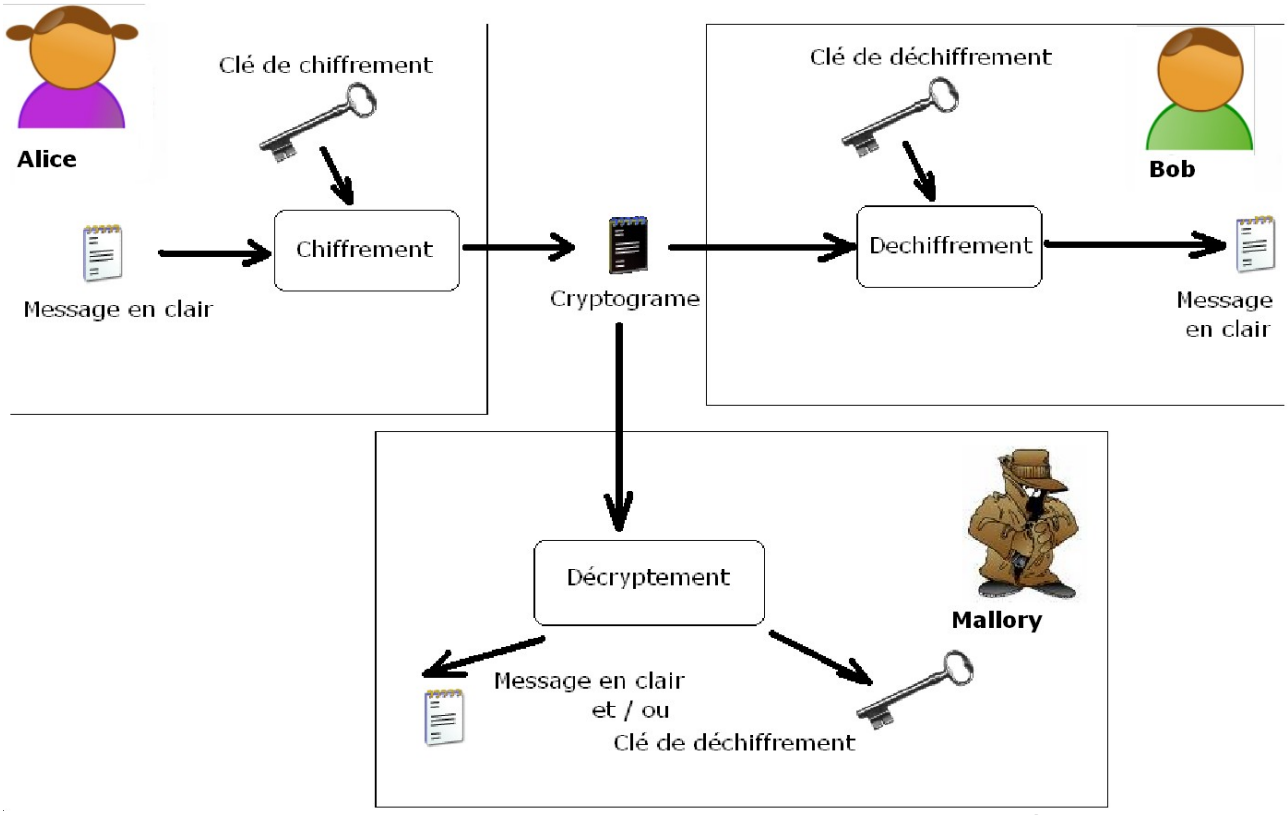
Initiation à la Cryptologie

Mars 2009

Table des matières

1 Vocabulaire



Coder : Transformer un texte, une information en remplaçant les **mots** dans une écriture faite de signes prédéfinis.

Chiffrer : Transformer un texte, une information en remplaçant les **lettres** dans une écriture faite de signes prédéfinis.[1]

Chiffrement : Transformation à l'aide d'une clé de chiffrement d'un message en clair en un message incompréhensible si on ne dispose pas d'une clé de déchiffrement (en anglais *encryption*)

Cryptogramme : Message chiffré

Clé : Une clé est un paramètre utilisé en entrée d'une opération cryptographique (chiffrement, déchiffrement, ...)

Décrypter : Retrouver le message clair correspondant à un message chiffré sans posséder la clé de déchiffrement (terme que ne possèdent pas les anglophones, qui eux « cassent » des codes secrets)

Cryptographie : Étymologiquement « écriture secrète », devenue par extension l'étude de cet art (donc aujourd'hui la science visant à créer des cryptogrammes, c'est-à-dire à chiffrer)

Cryptanalyse : Science analysant les cryptogrammes en vue de les décrypter ;

Cryptologie : Science regroupant la cryptographie et la cryptanalyse.

Alice, Bob & Mallory : Ce sont les personnages fictifs, des exemples de cryptographie. Alice veut écrire à Bob, et Mallory essaye d'intercepter le message (ou de le modifier, ou de se faire passer pour Alice ...)

Vu le sens des mots chiffrer; déchiffrer et décrypter, le terme «crypter» n'a pas de raison d'être (l'Académie française précise que le mot est à bannir et celui-ci ne figure pas dans son dictionnaire), en tout cas pas dans le sens où on le trouve en général utilisé.

Pas plus que le terme « décoder »[2].

1 Toute la différence entre coder et chiffrer : Le nombre de combinaisons possibles.
2 Sauf si vous vous destinez à une carrière d'installateur Canal+

2 Les enjeux

La cryptographie sert à assurer la confidentialité d'un document (seuls les possesseurs de la clé peuvent le lire) mais aussi a assurer l'intégrité d'un document : Je suis sur que le document que je lit est celui qui a été émis.

L'histoire de la cryptologie est très été liée à l'histoire de l'humanité. Surtout sur les plans militaires et informatique.

Eh oui, le premier ordinateur a été inventé pour cryptanaliser un message.

Des batailles ont été gagné grâce à la cryptanalyse[3]

On considère que la seconde guerre mondiale a été raccourcie d'au moins un an grâce à la connaissance des communications allemandes par les alliés.

De tout temps l'usage même de la cryptographie a été sensible

2.1 Le secret postal

Le «cabinet noir» a sévi pendant des siècles dans la plupart des États européens. Il s'est manifesté depuis l'ouverture des postes royales aux particuliers et l'institution du monopole. Le véritable mobile à été de placer la circulation des correspondances sous le contrôle royal, car elle a permis de mettre fin aux diverses postes particulières des grands seigneurs, des prélats et des universités.

Les agents des postes royales pouvaient ainsi lire ces lettres et en transmettre alors au gouvernement les extraits les plus compromettants. Ceux-ci étaient alors examinés par le monarque en son « cabinet noir », ce qui fut à l'origine de nombreuses disgrâces et condamnations.

Cette pratique jugée contraire à la légitimité monarchique, était très impopulaire et soulevait l'hostilité de toutes les classes de la population. En France en 1789, de nombreux cahiers de doléances réclamèrent son abolition. Cette revendication était si universellement partagée que, dans son rapport de synthèse du 27 juillet 1789 devant les États Généraux, le comte de Clermont-Tonnerre avait mis l'inquisition postale sur le même plan que les lettres de cachet, en ces termes :

« La Nation française s'élève avec indignation contre les lettres de cachet, qui disposaient arbitrairement des personnes, et contre la violation du secret de la poste, l'une des plus absurdes et des plus infâmes inventions du despotisme.»

Mais l'inquisition postale n'en fut pas moins rétablie par le Comité de salut public de la Convention, au nom de «la patrie en danger», pour lutter contre les conspirateurs, puis maintenue par les gouvernants ultérieurs.

2.2 Cryptologie et droit français

Longtemps les différents gouvernements, y compris le gouvernement français, interdirent la cryptographie.

Jusqu'en 1999, utiliser la cryptographie était assimilé à l'usage d'une arme de guerre ! (Sauf si vous utilisiez des clé suffisamment faible pour que le gouvernement puisse les casser ...)

Pour favoriser le développement du commerce électronique, le gouvernement a été contraint de libéraliser son utilisation. La cryptographie étant le seul moyen de protéger, de façon fiable, la circulation d'information sur l'internet (Protocole https).

Mais les lois sécuritaires du début du siècle (LSQ en 2001 puis LSI) considèrent que l'utilisation d'outils de cryptographie serait désormais considérée comme une "circonstance aggravante", et que ses utilisateurs seraient passibles de deux ans de prison, et 30 000 euros d'amende[4], s'ils refusaient de déchiffrer les messages chiffrés échangés.

Pire encore : La loi autorise également les juges à recourir aux "moyens de l'État soumis au secret de la Défense nationale" pour décrypter des informations chiffrées. Les rapports d'expertise sont donc classifiés et ne peuvent faire l'objet d'aucun recours; ce qui avait d'ailleurs été perçu, dans les milieux du renseignement français,

3 Ou « perdues à cause de » ... c'est très subjectif !
4 Porté à 3 ans et 45.000 € par la loi LSI

comme un excellent moyen de pouvoir fabriquer des preuves sans possibilité, pour l'accuser, de les contester.

Tout utilisateur d'outils de cryptographie se retrouve donc, sinon potentiellement suspect, tout du moins placé sous une épée de Damoclès dont il ne pourrait s'extirper si les autorités, pour quelque raison que ce soit, décidaient de s'intéresser à lui.

Les logiciels de cryptographie doivent être approuvé par le gouvernement avant de pouvoir être utilisé en France . C'est le cas de la plupart d'entre eux.

2.3 La NSA

Aujourd'hui, aux USA, une agence à, a elle toute seule, un budget supérieur à ceux de la CIA, du FBI et de la NASA réunis : La NSA (National Security Agency).

La NSA[5] est un organisme gouvernemental des États-Unis, responsables de la collecte et de l'analyse de toutes formes de communications, aussi bien militaires et gouvernementales que commerciales ou même personnelles, par radiodiffusion, par Internet ou par tout autre mode de transmission. Les agences ont aussi pour mission d'assurer la sécurité des communications (et donc des ordinateurs) du gouvernement américain.

En dépit du fait qu'elle soit le plus grand employeur de mathématiciens, d'informaticiens et d'électroniciens au monde, qu'elle possède un grand nombre d'ordinateurs, et un budget colossal (évalué à environ 4 milliards de dollars en 1997), l'agence a été remarquablement discrète jusqu'à récemment.

Selon une estimation de 2002, le quartier général de la NSA[6] utilise à lui seul assez d'électricité pour alimenter quatre Earth Simulators (l'ordinateur le plus puissant de l'époque)..

Le principal outil de la NSA est le réseau Échelon

2.4 Échelon

Échelon est un nom de code utilisé pendant de nombreuses années par les services de renseignements des États-Unis pour désigner une base d'interception des satellites commerciaux. Par extension, le Réseau Échelon désigne le système mondial d'interception des communications privées et publiques, élaboré par les États-Unis, le Royaume-Uni, le Canada, l'Australie et la Nouvelle-Zélande.

C'est un réseau global, appuyé par des satellites, de vastes bases d'écoutes situées aux États-Unis, au Canada, au Royaume-Uni, en Allemagne, en Australie et en Nouvelle-Zélande, des petites stations d'interception dans les ambassades, et le sous-marin USS Jimmy Carter, entré en service en 2005 pour écouter les câbles sous-marins de télécommunications.

Il intercepte les télécopies, les communications téléphoniques, les courriels et, grâce à un puissant réseau d'ordinateurs, est capable de trier en fonction de certains termes les communications écrites et, à partir de l'intonation de la voix, les communications orales.

Ces réseaux peuvent être utilisés pour des actions militaires, politiques ou commerciales. Il aurait été utilisé pour faire gagner des contrats à des compagnies américaines, face à ses concurrents, comme Boeing contre Airbus.

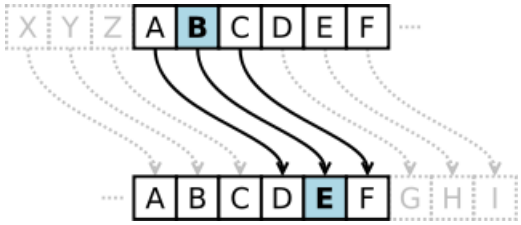
Toutes les informations récoltées par le réseau Echelon sont analysées au quartier général de la NSA à Fort Meade.

5 **N**o **S**uch **A**gency (une telle agence n'existe pas)
6 **N**ever **S**ay **A**nything (ne jamais rien dire)

3 Algorithmes de chiffrement faibles

Les premiers algorithmes utilisés pour le chiffrement d'une information étaient assez rudimentaires dans leur ensemble. Ils consistaient notamment au remplacement de caractères par d'autres. La confidentialité de l'algorithme de chiffrement était donc la pierre angulaire de ce système pour éviter un décryptage rapide.

3.1 Chiffre de César



le chiffre de César est une des plus simple méthodes d'encryptage connues.

C'est une technique de codage par substitution, c'est-à-dire que chaque lettre du texte en clair est remplacée par une autre lettre à distance fixe dans l'alphabet. Par exemple, si l'on utilise un décalage de 3, A serait remplacé par D, B deviendrait E, et ainsi de suite. Cette méthode doit son nom à Jules César, qui utilisait cette technique pour certaines de ses correspondances.

3.1.1 Exercice

Cryptogramme : QJAJE QF RFNS JY INYJX "UWJRNJW".

Déchiffrez !

3.1.2 ROT13

Le ROT13 ("rotate by 13 places") (une variante du chiffrier de César) est un algorithme très simple de chiffrement de texte. Comme son nom l'indique, il s'agit d'un décalage de 13 caractères de chaque lettre du texte à chiffrer.

L'avantage de ROT13, c'est le fait que le décalage soit de 13 : Comme l'alphabet comporte 26 lettres le même « algorithme » sert à la fois pour chiffrer et pour déchiffrer.

Il est encore utilisé dans les logiciels de messagerie comme Outlook express ou, dans un forum, on souhaite cacher une réponse à une question posé.

3.2 Chiffrement monoalphabétique

L'histoire n'a pas retenu le nom de l'inventeur de la première méthode de chiffrement à clé. Elle date probablement du début de notre ère.

L'intérêt de cette méthode est que pour que le cryptogramme soit déchiffré, il suffit de faire parvenir une clé (qui peut être un simple mot) à son (ou ses) destinataires légitimes.

Comme il existe des millions de mots clé possible, il existe des millions de combinaisons possible.

Étape 1 : Choisir un mot clé.

Ex : informatique

Étape 2 : Le "nettoyer" en enlevant tout les doubles et les accents :

INFORMATQUE

Étape 3 : Reporter ce mot dans une grille :

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

Étape 4 : Compléter l'alphabet

- En prenant soin de ne pas utiliser 2 fois une lettre
- En partant de la dernière lettre écrite

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

Étape 5 : On peut maintenant crypter un message par substitution

Texte en clair :	Aujourd'hui, j'ai mangé des nouilles[7]				
Texte préparé :	AUJOU	RDHUI	JAIMA	NGEDE	SNOUI	LLES
Cryptogramme :	IXUKX	SOTXQ	UIQHI	JAROR	VJKXQ	GGRV

3.2.1 Exercice

Déchiffrez le cryptogramme monoalphabétique suivant, sachant que le clé est « Il était une fois » :

AYUOY JYAWA KAIKA EJYEA DKAIJ

3.3 Cryptanalyse par étude de fréquence

Par sa simplicité et par sa force, la substitution monoalphabétique a dominé la technique des écritures secrètes pendant tout le premier millénaire. Il a résisté aux cryptanalystes jusqu'à ce que le savant arabe Abu Yusuf Ya'qub ibn Is-haq ibn as-Sabbah Oòmran ibn Ismaïl **al-Kindi** mette au point, au IX^{ème} siècle, une technique appelée analyse des fréquences.

Al-Kindi (801-873) rédige sa méthode dans un traité intitulé « Manuscrit sur le déchiffrement des messages cryptographiques ».

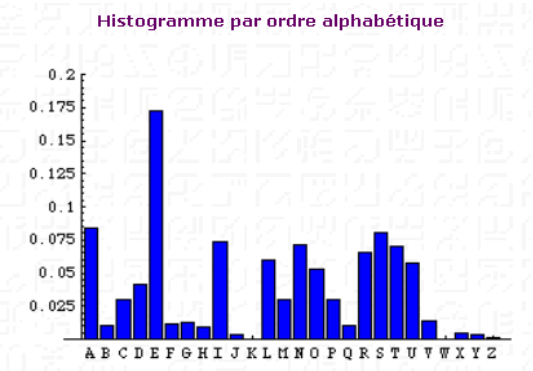
Il explique que « la façon d'élucider un message crypté, si nous savons dans quelle langue il est écrit, est de nous procurer un autre texte en clair dans la même langue, de la longueur d'un feuillet environ, et de compter alors les apparitions de chaque lettre.

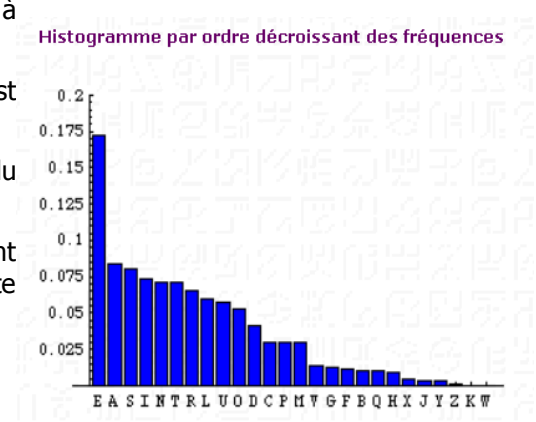
Ensuite, nous nous reportons au texte chiffré que nous voulons éclaircir et relevons de même ses symboles. Nous remplaçons le symbole le plus fréquent par la lettre première (la plus fréquente du texte clair), le suivant par la deuxième, le suivant par la troisième, et ainsi de suite jusqu'à ce que nous soyons venus à bout de tous les symboles du cryptogramme à résoudre ».

Cette technique ne fonctionne bien que si le cryptogramme est suffisamment long pour avoir des moyennes significatives.

Voici, ci contre, les diagrammes de fréquence des lettres du Français.

Il est « amusant » de constater que plusieurs nations ignoraient les travaux d'al-kindi et on utilisé des messages chiffrés par cette méthode jusqu'au milieu du XV^{eme} siècle.[8]





3.3.1 Casser un code

On considère qu'un code (ou un chiffre) a été cassé si il existe une méthode qui permet de le cryptanaliser plus rapidement que par une attaque en force brute.

Un attaque par force brute consiste à essayer successivement toutes les combinaisons possibles.

7 Eh oui, la cryptographie sert à dissimuler les plus grands secrets de l'univers ...
8 Ne comptez pas sur moi pour vous donner le nom de ces nations. C'est pas mon genre de me moquer de l'Espagne et du Portugal...

3.4 Attaque par mot probable

Nous avons vu que nous pouvions décrypter un chiffre monoalphabétique par une analyse des fréquences. Une autre technique consiste à deviner un mot qui doit, ou peut, apparaître dans le texte clair. Dans une substitution monoalphabétique, un mot chiffré a le même aspect que le mot clair. Prenons par exemple le mot anglais "AMMUNITION"; dans le texte chiffré apparaîtra une séquence de dix lettres avec les caractéristiques suivantes:

- Les 2ème et 3ème lettres sont les mêmes
- Les 5ème et 10ème lettres sont les mêmes (et différentes de la 2ème lettre)
- Les 6ème et 8ème lettres sont les mêmes (et différentes des 2ème et 5ème lettres)
- Toutes les autres lettres sont différentes.

Une fois que l'on a trouvé toutes les correspondances possibles, on peut utiliser une statistique chi-carré pour déterminer laquelle est la plus probable.

3.5 Chiffre de Vigenère

Le Chiffre de Vigenère est un système de chiffrement, élaboré par Blaise de Vigenère (1523-1596), diplomate français du XVIe siècle.

C'est un système de substitution polyalphabétique. Cela signifie qu'il permet de remplacer une lettre par une autre qui n'est pas toujours la même, contrairement aux chiffres vu précédemment qui se contentaient d'utiliser la même lettre de substitution. C'est donc un système relativement plus « solide ».

3.5.1 La table de Vigenère

L'outil indispensable du chiffrement de Vigenère est : « La table de Vigenère » ou « carré de Vigenère ».

On l'obtient en écrivant 26 fois l'alphabet, et en décalant chaque ligne d'une lettre.

Pour la 1er ligne : ABCDE ... XYZ

Pour la 2eme : BCDEF... YZA

La 3eme : CDEFG...ZAB

Etc

3.5.2 Chiffrement

Pour chaque lettre en clair, on sélectionne la colonne correspondante et pour une lettre de la clé on sélectionne la ligne adéquate, puis au croisement de la ligne et de la colonne on trouve la lettre codée. La lettre de la clé est à prendre dans l'ordre dans laquelle elle se présente et on répète la clé en boucle autant que nécessaire.

3.5.3 Exemple

Clé : Musique

Texte : j'adore écouter la radio toute la journée

On répète la clé jusqu'à ce qu'elle soit aussi longue que le texte à chiffrer :

Clé : **MUSIQU EMUSIQU EM USIQU EMUSI QU EMUSIQU**

Texte : **JADORE ECOUTER LA RADIO TOUTE LA JOURNEE**

Pour la 1ere lettre : On prend dans la table, la colonne de la clé (M) et la ligne de la lettre (J) : V
Pour la 2eme lettre : On prend la colonne de la clé (U) et la ligne de la lettre (A) : U
Pour la 3eme lettre : On prend la colonne de la clé (S) et la ligne de la lettre (D) : V
Pour la 4eme lettre : On prend la colonne de la clé (I) et la ligne de la lettre (O) : W

Etc

Le texte chiffré est alors :

VUVWHY IOIMBUL PM LSLYI XAOLM BU NAOJVUY.

3.5.4 Déchiffrement

Si on veut déchiffrer ce texte, on regarde pour chaque lettre de la clé répétée la ligne correspondante, et on y cherche la lettre codée. La première lettre de la colonne que l'on trouve ainsi est la lettre décodée.

```
Texte codé     :   VUVWHY IOIMBUL PM LSLYI XAOLM BU NAOJVUY
Clé répétée    :   MUSIQU EMUSIQU EM USIQU EMUSI QU EMUSIQU
                   ^^^
                   |||Ligne I, on cherche W: on trouve la colonne O.
                   ||Ligne S, on cherche V: on trouve la colonne D.
                   |Ligne U, on cherche U: on trouve la colonne A.
                   Ligne M, on cherche V: on trouve la colonne J.
```

3.5.5 Exercice

Texte codé : DHMZG HXUMY UUAMT UHLTG QTMAK

Clé : BTSIG

Déchiffrez ce cryptogramme.

3.6 Attaque par indice de coïncidence

La figure la plus étonnante de la cryptanalyse au XIXème siècle est celle de Charles Babbage (1792-1871).

Babbage était un touche à tout de génie : il fut le premier à comprendre que dans un tronc d'arbre, la largeur d'un anneau dépend du temps qu'il a fait dans l'année. Il s'intéressa aux statistiques (premières tables de mortalité). Il proposa un prix unique pour l'affranchissement d'une lettre. Après s'être rendu compte que les éphémérides nautiques pour trouver la latitude et la longitude en mer contenaient plus de mille erreurs, il échoua faute de financement à construire une machine mécanique capable de faire des calculs avec un haut degré de précision.

Cette machine (aujourd'hui construite d'après les plans de l'époque) fonctionne parfaitement : C'est l'ancêtre des ordinateurs.

Il apporta une contribution importante à la cryptanalyse: il réussit à casser le chiffre de Vigenère, probablement en 1854 car sa découverte resta ignorée en l'absence d'écrit. Pendant ce temps, un officier prussien à la retraite, Friedrich Wilhelm Kasiski (1805-1881), parvint au même résultat et publia en 1863 "Die Geheimschriftren und die Dechiffrir-kunst".

3.6.1 Cryptanalyse commentée d'un chiffre de Vigenère

```
KQOWE FVJPU JUUNU KGLME KJINM WUXFQ MKJBG WRLFN FGHUD WUUMB SVLPS NCMUE KQCTE SWREE
KOYSS IWCTU AXYOT APXPL WPNTC GOJBG FQHTD WXIZA YGFFN SXCSE YNCTS SPNTU JNYTG GWZGR
WUUNE JUUQE APYME KQHUI DUXFP GUYTS MTFFS HNUOC ZGMRU WEYTR GKMEE DCTVR ECFBD JQCUS
WVBPN LGOYL SKMTE FVJJT WWMFM WPNME MTMHR SPXFS SKFFS TNUOC ZGMDO EOYEE KCPJR GPMUR
SKHFR SEIUE VGOYC WXIZA YGOSA ANYDO EOYJL WUNHA MEBFE LXYVL WNOJN SIOFR WUCCE SWKVI
DGMUC GOCRU WGNMA AFFVN SIUDE KQHCE UCPFC MPVSU DGAVE MNYMA MVLFM AOYFN TQCUA FVFJN
XKLNE IWCWO DCCUL WRIFT WGMUS WOVMA TNYBU HTCOC WFYTN MGYTQ MKBBN LGFBT WOJFT WGNTE
JKNEE DCLDH WTVBU VGFBI JG
```

Phase 1 : Trouver la longueur de la clé

Étape 1 : Soulignez les répétitions de 3 caractères ou plus :

```
KQOWE FVJPU JUUNU KGLME KJINM WUXFQ MKJBG WRLFN FGHUD WUUMB SVLPS NCMUE KQCTE SWREE
KOYSS IWCTU AXYOT APXPL WPNTC GOJBG FQHTD WXIZA YGFFN SXCSE YNCTS SPNTU JNYTG GWZGR
WUUNE JUUQE APYME KQHUI DUXFP GUYTS MTFFS HNUOC ZGMRU WEYTR GKMEE DCTVR ECFBD JQCUS
WVBPN LGOYL SKMTE FVJJT WWMFM WPNME MTMHR SPXFS SKFFS TNUOC ZGMDO EOYEE KCPJR GPMUR
SKHFR SEIUE VGOYC WXIZA YGOSA ANYDO EOYJL WUNHA MEBFE LXYVL WNOJN SIOFR WUCCE SWKVI
DGMUC GOCRU WGNMA AFFVN SIUDE KQHCE UCPFC MPVSU DGAVE MNYMA MVLFM AOYFN TQCUA FVFJN
XKLNE IWCWO DCCUL WRIFT WGMUS WOVMA TNYBU HTCOC WFYTN MGYTQ MKBBN LGFBT WOJFT WGNTE
JKNEE DCLDH WTVBU VGFBI JG
```

Étape 2 : Pour chaque répétition, mesurer la période

Séquence répétée	Distance
WUU	95
EEK	200
WXIZAYG	190
NUOCZGM	80
DOEOY	45
GMU	90

Étape 3 : Pour chaque période, décomposer en facteurs premiers et regarder quel facteur est commun à tous :

		Longueurs de clef possibles			
Séquence répétée	Espace de répétition	2	3	5	19
WUU	95			x	x
EEK	200	x		x	
WXIZAYG	190	x		x	x
NUOCZGM	80	x		x	
DOEOY	45		x	x	
GMU	90	x	x	x	

La clé est ici longue de 5 caractères.

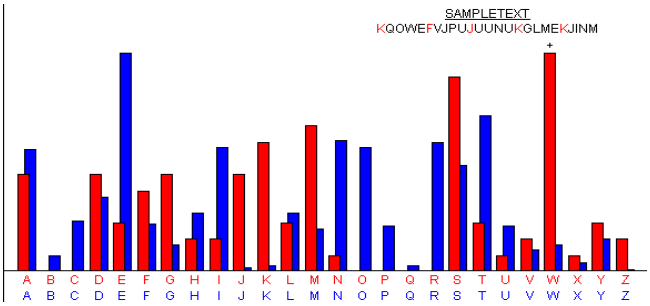
Phase 2 : Trouver la 1er lettre du mot clé

Étape 1 : Faire une analyse de fréquence seulement sur les caractères 1, 6, 11, ...

On obtient ici :

En rouge, l'analyse de fréquence « modulo 5 »

En bleu le diagramme de fréquence des lettres en français.



Étape 2 : On décale pour faire correspondre

On décale les diagrammes pour mettre le pic du **W** sur le **E**

... L'ensemble correspond à peut prés : On à la première lettre de la clé.

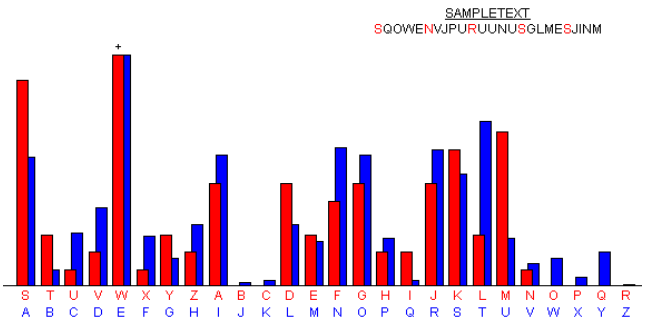
Avec **W** = 23 et **E** = 5, c'est la (23 – 5 + 1= 19eme) soit **S**

Phase 3, 4, 5 et 6 : On recommence pour avoir les 5 lettres du mot clé.

Le mot clé est **SCUBA**

On peut déchiffrer le cryptogramme :

Soit :




```
SOUVE NTPOU RSAMU SERLE SHOMM ESDEQ UIPAG EPREN NENTD ESALB ATROS VASTE SOISE AUXDE SMERS
QUISU IVENT INDOL ENTSC OMPAG NONSD EVOYA GELEN AVIRE GLISS ANTSU RLESG OUFFR ESAME RSAPE
INELE SONTI LSDEP OSESS URLES PLANC HESQU ECESR OISDE LAZUR MALAD ROITS ETHON TEUXL AISSE
NTPIT EUSEM ENTLE URSGR ANDES AILES BLANC HESCO MMEDE SAVIR ONSTR AINER ACOTE DEUXC EVOYA
GEURA ILECO MMEIL ESTGA UCHEE TVEUL ELUIN AGUER ESIBE AUQUI LESTC OMIQU EETLA IDLUN AGACE
SONBE CAVEC UNBRU LEGUE ULELA UTREM IMEEN BOITA NTLIN FIRME QUIVO LAITL EPOET EESTS EMBLA
BLEAU PRINC EDESN UEESQ UIHAN TELAT EMPET EETSE RITDE LARCH ERBAU DELAI RE
```

Soit encore :

Souvent pour s'amuser les hommes d'équipage prennent des albatros, vastes oiseaux des mers, qui suivent, indolents compagnons de voyage, le navire glissant sur les gouffres amers.

A peine les ont-ils déposés sur les planches que ces rois de l'azur, maladroits et honteux, laissent piteusement leurs grandes ailes blanches, comme des avirons, traîner à côté d'eux.

Ce voyageur ailé, comme il est gauche et veule, lui naguère si beau, qu'il est comique et laid. L'un agace son bec avec un brûle-gueule, l'autre mime en boitant l'infirme qui volait.

Le poète est semblable au prince des nuées, qui hante la tempête et se rit de l'archer.

Charles Baudelaire

Au 19eme siècle, les cryptanalyses reprennent l'avantage.

4 Algorithmes de cryptographie symétrique

Les algorithmes de chiffrement symétrique se fondent sur une même clé pour chiffrer et déchiffrer un message. Le problème de cette technique est que la clé, qui doit rester totalement confidentielle, doit être transmise au correspondant de façon sûre.

Ces algorithmes sont dit aussi « à clé secrète »

Les trois algorithmes étudiés précédemment sont, bien sur des algorithmes à clé secrète.

4.1 Clé et sécurité

L'un des concepts fondamentaux de la cryptographie symétrique est la clé, qui est une information devant permettre de chiffrer et de déchiffrer un message et sur laquelle peut reposer toute la sécurité de la communication. Un algorithme comme le ROT13 par exemple n'a pas de clé, il suffit de savoir que cette méthode a été utilisée pour chiffrer un message et on peut avoir accès au texte clair. En d'autres termes, ici, le secret réside dans la méthode utilisée. Ce type de secret ne satisfait pas les utilisateurs de chiffrement, car la conception d'un bon algorithme est très difficile, une fois découvert il n'offre plus de sécurité, et tous les messages qui ont été chiffré par lui deviennent accessibles.

4.1.1 Sécurité calculatoire

Auguste Kerckhoffs (La cryptographie militaire, 1883) est certainement l'un des premiers à avoir pleinement compris cela : pour espérer être sûr, l'algorithme doit pouvoir être divulgué.

Dit autrement, la sécurité ne doit reposer que que la connaissance de la clé : C'est ce que l'on appelle le principe de Kerckhoffs.

Il faut ajouter que cette clé doit pouvoir prendre suffisamment de valeurs pour qu'une attaque en force brute — essai systématique de toutes les clés — ne puisse être menée à bien car trop longue. On parle de sécurité calculatoire.

Cette sécurité calculatoire est bien évidemment dépendante du temps, les performances des moyens de calcul allant croissant, un système de chiffrement est confronté à une adversité toujours plus forte, alors que le message chiffré ne change plus.

L'illustration de ce problème est le DES, ce système est devenu obsolète à cause du trop petit nombre de clés qu'il peut utiliser (pourtant 2^{56}). On pense que, actuellement, 2^{80} est un strict minimum.

À titre indicatif, le dernier standard choisi par les Américains en décembre 2001, l'AES, utilise des clés dont la taille est au moins de 128 bits, autrement dit il y en a 2^{128}.

Pour donner un ordre de grandeur sur ce nombre, cela fait environ $3{,}4*10^{38}$ clés possibles; l'âge de l'univers étant de 10^{10} années, si on suppose qu'il est possible de tester 1 000 milliards de clés par seconde (soit $3.2*10^{19}$ clés par an) il faudra encore plus d'un milliard de fois l'âge de l'univers pour les essayer toutes.

Dans un tel cas on peut raisonnablement penser que notre algorithme est sûr. Cependant, c'est faire une hypothèse très forte sur l'algorithme que de supposer que le seul moyen de le casser est de mener une attaque par force brute : nombreuses sont les failles que peuvent receler un algorithme et encore plus nombreuses sont les mauvaises utilisations d'un algorithme.

4.2 Chiffre de Vernam

Le chiffre de Vernam, également appelé masque jetable est un algorithme de cryptographie inventé par Gilbert Vernam en 1917. Bien que simple, ce chiffrement est le seul qui soit théoriquement impossible à casser, même s'il présente d'importantes difficultés de mise en œuvre pratique.

4.2.1 Principe

Le chiffrement par la méthode du masque jetable consiste à combiner le message en clair avec une clé présentant les caractéristiques très particulières suivantes :

- La clé doit être une suite de caractères aussi longue que le message à chiffrer.
- Les caractères composant la clé doivent être choisis de façon totalement aléatoire.
- Chaque clé, ou « masque », ne doit être utilisée qu'une seule fois (d'où le nom de masque jetable).

L'intérêt considérable de cette méthode de chiffrement, c'est que si les trois règles ci-dessus sont respectées strictement, le système offre une sécurité théorique absolue.

4.2.2 Chiffrement et déchiffrement à la main

Le chiffrement à la main par la méthode du masque jetable fut notamment utilisée par Che Guevara[9] pour communiquer avec Fidel Castro.

Exemple commenté :

On veut chiffrer le message « **HELLO** ».

On choisit la clé : **X M C K L**

Pour cela, on attribue un nombre à chaque lettre, par exemple le rang dans l'alphabet, de 0 à 25. Ensuite on additionne la valeur de chaque lettre avec la valeur correspondante dans le masque; enfin si le résultat est supérieur à 25 on soustrait 26 (calcul dit "modulo 26") :

9 Oui, le célèbre fabriquant de T-shirts.

```
   7 (H)   4 (E)  11 (L)  11 (L)  14 (O) message
+ 23 (X)  12 (M)   2 (C)  10 (K)  11 (L) masque
= 30      16      13      21      25     masque + message
=  4 (E)  16 (Q)  13 (N)  21 (V)  25 (Z) masque + message modulo 26
```

Le texte reçu par le destinataire est « **EQNVZ** ».

Le déchiffrement s'effectue de manière similaire, sauf que l'on soustrait le masque au texte chiffré au lieu de l'additionner. Ici encore on ajoute éventuellement 26 au résultat pour obtenir des nombres compris entre 0 et 25 :

```
    4 (E)  16 (Q)  13 (N)  21 (V)  25 (Z) message chiffré
-  23 (X)  12 (M)   2 (C)  10 (K)  11 (L) masque
= -19       4      11      11      14     message chiffré - masque
=   7 (H)   4 (E)  11 (L)  11 (L)  14 (O) message chiffré - masque modulo 26
```

On retrouve bien le message initial « **HELLO** ».

4.2.3 Problème de la transmission des clés

La seule méthode sûre pour transmettre les clés au correspondant est le transport physique, typiquement dans une valise diplomatique. Aucune transmission sur un réseau n'est acceptable, une interception ne pouvant jamais être totalement exclue.

Le transport de la clé devient le point faible de Graal cryptographique qu'est un chiffre incassable.

4.2.4 Difficulté de produire une clé parfaitement aléatoire

Le fait que la clé soit constituée d'une suite de caractères (ou de bits) totalement aléatoires est une condition essentielle de la sécurité du chiffre de Vernam. Un défaut du masque sur ce point peut suffire pour que le cryptanalyse retrouve le message en clair.

Des clés parfaitement aléatoires ne peuvent pas être produites par un ordinateur : En effet ce dernier est une machine fondamentalement déterministe, dont le résultat est totalement prévisible quand on connaît les calculs programmés et leurs données initiales. Pourtant de nombreux algorithmes ont été proposés dans ce but : On les appelle des générateurs de nombres pseudo-aléatoires. Leur résultat est utile dans beaucoup de situations, mais il ne répond pas à la condition du parfait aléa, qui seul garantit la sécurité absolue

Dit autrement, il n'est pas possible de prédire quel sera le prochain nombre tiré au sort par un ordinateur. Mais si vous me donnez un série de 10.000 nombres, je trouve facilement le 10.001 eme

4.2.5 Problème de l'utilisation unique de chaque clé

Si le chiffre de Vernam est inviolable, c'est parce qu'une attaque en force brute à la même probabilité de trouver toute les combinaisons de n lettres. Dans notre exemple (`HELLO`, codé `EQNVZ`) La cryptanalyse donnerait la liste complète des mots de 5 lettres.

Mais si le même masque est utilisé pour deux messages différents, la solution devient triviale.

En pratique l'utilisation sûre du masque jetable demande une organisation rigoureuse : Chaque clé doit être précisément identifiée et soigneusement tracée, d'autant qu'elle est toujours fournie en deux exemplaires, à deux correspondants géographiquement distants. Imaginez qu'une chancellerie communique par cette méthode avec ses dizaines d'ambassades réparties dans des pays du monde, chacune d'elle envoyant et recevant plusieurs messages par jour, pouvant comporter un grand nombre de pages, et ceci pendant des années : Il faut une logistique lourde pour garantir la sécurité absolue. Mais cette méthode a été et est encore largement utilisée par certains états.

Pour garantir l'utilisation unique des clés, les agents du KGB utilisaient souvent des masques qui étaient imprimés sur un papier spécial, celui-ci brûlait presque instantanément et sans laisser de cendres.

4.2.6 Exercice

Cryptogramme :	E	U	U	E	Q	E	J	S	Y	K	Y	T	U	U	E	I	K	O	W	K	O	B
Clé :	K	H	U	S	I	O	P	K	D	W	E	B	Z	Q	K	P	H	U	V	C	K	O

Déchiffrez !

5 L'arrivée de l'informatique

5.1 La machine Enigma

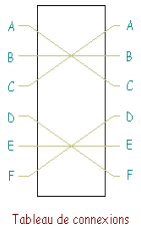
Après la défaite de 1918, en bonne partie due à des « revers cryptologiques »[10] l'Allemagne pense avoir trouvé la parade avec la machine Enigma.

L'histoire commence en 1919, quand un ingénieur hollandais, Hugo Alexander, dépose un brevet de machine à chiffrer électromécanique. Ses idées sont reprises par le Dr Arthur Scherbius, qui crée à Berlin une société destinée à fabriquer et à commercialiser une machine à crypter civile : l'Enigma. Cette société fait un fiasco, mais la machine Enigma a attiré l'attention des militaires.

5.1.1 Le fonctionnement

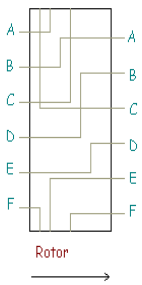
Le codage effectué par la machine Enigma est à la fois simple et astucieux. Chaque lettre est remplacée par une autre, l'astuce est que la substitution change d'une lettre à l'autre.

La machine Enigma est alimentée par une pile électrique. Quand on appuie sur une touche du clavier, une lampe s'allume qui indique quelle lettre codée l'on substitue.



Concrètement, le circuit électrique est constitué de plusieurs éléments en chaîne :

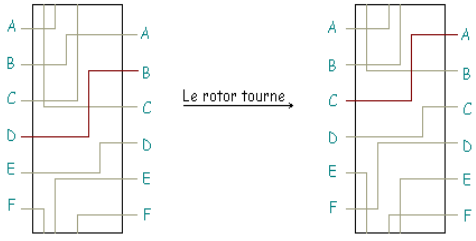
- le **tableau de connexions** : il permet d'échanger des paires de l'alphabet, deux à deux, au moyen de fiches. Il y a 6 fiches qui permettent donc d'échanger 12 lettres. Par exemple, dans le tableau suivant (avec simplement 6 lettres), on a échangé A et C, D et F, tandis que B et D restent invariants.
- **les rotors** : un rotor est un cylindre qui fait correspondre, a chaque lettre en entrée une autre lettre.



Entrée	Sortie
A	E
B	A
C	F
D	B
E	D
F	F

Les rotors sont montés à la suite les uns des autres. La machine Enigma disposera, au gré de ses évolutions successives, de 3 à 6 rotors. On a le choix de les placer dans l'ordre que l'on souhaite (ce qui constituera une partie de la clé).

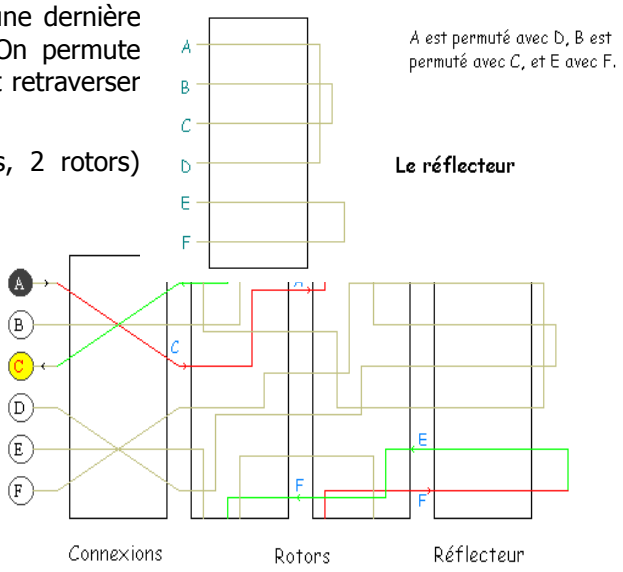
Surtout, les rotors sont mobiles. Ainsi, à chaque fois qu'on a tapé une lettre, le premier rotor tourne d'un cran, et la permutation qu'il engendre est changée. Sur la figure suivante : le rotor transforme initialement D en B. Lorsqu'il tourne d'un cran, cette liaison électrique D--->B se retrouve remontée en C--->A.



Chaque rotor possède donc 26 positions. A chaque fois qu'une lettre est tapée, le premier rotor tourne d'un cran. Après 26 lettres, il est revenu à sa position initiale, et le second rotor tourne alors d'un cran. On recommence à tourner le premier rotor, et ainsi de suite... Quand le second rotor a retrouvé sa position initiale, c'est le troisième rotor qui tourne d'un cran.

10 Voir chapitre 12.3 Le télégraphe de Zimmerman et 12.4 le chiffre ADFGVX

- **Le réflecteur :** Au bout des 3 rotors se situe une dernière permutation qui permet de revenir en arrière. On permute une dernière fois les lettres 2 par 2, et on les fait retraverser les rotors, et le tableau de connexion.



Résumons sur la machine simplifiée suivante (6 lettres, 2 rotors) comment est codée la lettre A :

- On traverse le tableau de connexions : on obtient C.
- On traverse les 2 rotors : on obtient successivement A et F.
- On traverse le réflecteur où on obtient E, puis on renvoie dans les rotors pour obtenir F, A et finalement C après le tableau de connexions.

Remarquons que si on avait tapé C, le courant aurait circulé dans l'autre sens et on aurait obtenu A.

5.1.2 Nombre de clés possibles

Il y a trois éléments à connaitre pour pouvoir coder un message avec la machine Enigma.

1. la position des 6 fiches du tableau de connexion : (12 lettres parmi 26, 6 paires de lettres parmi 12) Soit 100.391.791.500 possibilités.
2. l'ordre des rotors : il y a autant d'ordre que de façons d'ordonner 3 éléments : 3!=6.
3. la position initiale des rotors : chaque rotor ayant 26 éléments, il y a 26^3=17.576 choix.

On multiplie tout cela, et on obtient plus de 10^{16} possibilités, ce qui est énorme pour l'époque!

Il est important de remarquer que les permutations employées dans les rotors et les réflecteurs ne peuvent pas être considérées comme faisant partie du secret. En effet, toutes les machines utilisent les mêmes, et il suffit donc d'en avoir une à disposition. Les Anglais, par exemple, en ont récupéré une pendant la guerre dans un sous-marin coulé. Ceci est une illustration du principe de Kerckhoffs, qui veut que tout le secret doit résider dans la clé secrète de chiffrement et de déchiffrement, et pas dans une quelconque confidentialité de l'algorithme (ici de la machine) qui ne peut être raisonnablement garantie.

Les allemands ont une confiance totale en la machine Enigma, dont ils fabriqueront 100.000 exemplaires. Au su et au vu de tous, ils s'échangeront des communications radios cryptées, persuadés que jamais les Alliés ne les comprendront.

5.2 La cryptanalyse de la machine Enigma

« C'était impossible. Tout le monde le savait. Jusqu'au jour ou est arrivé un imbécile qui ne le savait pas ... et qui l'a fait ! »

Sir Winston Churchill

5.2.1 Point forts et faiblesses

L'une des failles de la machine Enigma est que jamais la lettre A ne sera codée par un A. Cela élimine un certain nombre de cas à inspecter.

Une des autres faiblesse 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 anglais 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 !

5.2.2 Le travail des Polonais

Dès 1933 et jusqu'au début de la guerre, grâce aux renseignements recueillis par un militaire français (Gustave Bertrand) et au travail de trois mathématiciens polonais (Marian Rejewski, Jerzy Rózicki et Henryk Zygalski), le "Polski Biuro Szyfrów" sait décrypter les messages allemands, chiffrés avec la machine Enigma, exploitant une faille dans la procédure de début de transmission (Les opérateurs allemand saisissaient deux fois les trois premiéres lettres du message).


Marian Rejewski


Jerzy Rózicki


Henryk Zygalski

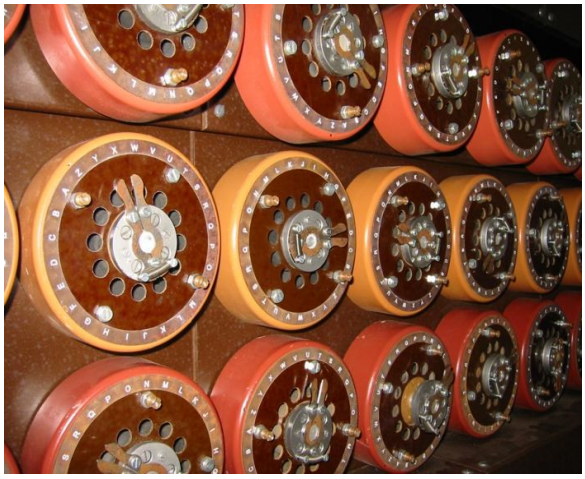
Même si ces trois lettres sont inconnues, le nombre de câblages qui peuvent transformer ces trois lettres en une séquence particulière sont limités. Rejewski les appelle des « chaînes ».

5.2.3 Trouver les bonnes chaînes

Le nombre de chaînes possibles se monte à 105 456 et, à l'époque, en l'absence d'une puissance de calcul suffisante, la recherche est quasi-impossible. L'équipe de Rejewski découvre plusieurs techniques pour accélérer les calculs. L'une d'entre elles fait appel à des feuilles transparentes, avec les schémas des rotors. Les feuilles sont superposées et les lettres composant une chaîne dite « impossible » sont rayées. À la fin de l'opération, les trois lettres restantes donnent le code utilisé pour l'en-tête.

Malgré cette performance, la recherche manuelle n'en demeure pas moins très pénible. Les Polonais construisent alors une « bombe cryptologique », véritable calculateur électromécanique qui date de 1938. Six exemplaires sont montés à Varsovie dans le bureau du chiffre juste avant le début de la Seconde Guerre mondiale. Chacune contient l'équivalent de six machines Enigma alimentées électriquement. Son volume est par contre considérable, l'équivalent d'un atelier pour 100 personnes, mais les gains sont significatifs : il suffit de deux heures pour obtenir la clé.

Les Polonais seront ainsi capables de déchiffrer une bonne partie des transmissions de l'armée allemande dès 1933, durant la Guerre civile espagnole et ceci jusqu'à l'aube de la Seconde Guerre mondiale.



5.2.4 L'invasion allemande

En 1939, les Allemands ont complexifié la structure de leur machine : Jusqu'alors, seuls trois rotors étaient employés mais l'armée allemande introduit deux rotors supplémentaires. De plus, le début de la guerre marque l'arrêt de la procédure de répétition du code au début des messages et tous les efforts des Polonais sont réduits à néant puisque leur cryptanalyse repose sur cette redondance.

Il est probable que le service cryptographique de l'armée allemande eut vent des progrès polonais ou tout au moins soupçonne ou même découvre une faille dans leur procédé de chiffrement.

Les Polonais partagent alors le résultat de leurs travaux avec la France et les Britanniques, devenus leurs alliés. Durant l'été 1939, ils fournissent à la France et au Royaume-Uni des copies d'Enigma, la description précise de l'attaque, la technique de la grille et les plans de la bombe cryptographique.

En septembre 1939, l'invasion de la Pologne par les Nazis débute. Les cryptologues polonais évacuent dans

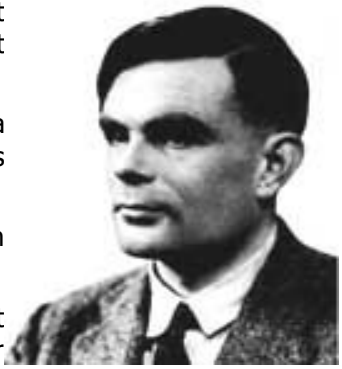
l'urgence. Ils atteignent la France. Peu avant l'invasion de la France en mai 1940, le mathématicien britannique Alan Turing demeure quelques jours au PC Parisien où il sera briffé par ses confrères polonais.

Après l'armistice, les cryptologues se déplacent dans le sud de la France et en Algérie. Rejewski et Zygalski fuient travers l'Espagne où ils seront temporairement emprisonnés, le Portugal et finalement Gibraltar d'où ils pourront gagner le Royaume-Uni. Różycki aura beaucoup moins de chance puisqu'il meurt noyé lors d'un naufrage en 1942 au sud de la France, après un voyage en Algérie.

5.2.5 Bletchley Park

La cryptanalyse d'Enigma était devenue entre temps une affaire britannique et américaine.

C'est Alan Turing qui va s'occuper de l'analyse de l'Enigma. Turing est le chef de la huitième section à Bletchley Park, un manoir proche de Londres où se sont retranchés tous les cryptologues et mathématiciens.



Les attaques Britanniques font appel à l'analyse des mots probables. Les messages avaient de forte chance de contenir des termes comme « Heil Hitler », « Panzer », « Führer », « Stuka », etc.

Les cryptologues pouvaient a posteriori deviner le contenu des messages en fonction de l'actualité et des assauts ennemis.

En faisant quelques hypothèses sur le contenu et sachant qu'une lettre est obligatoirement modifiée lors du chiffrement, il n'était pas impossible de retrouver une partie du texte chiffré en essayant tous les alignements possibles. Cette attaque ressemblait à celle des Polonais qui tentaient de deviner l'en-tête des messages. Turing avait découvert des « clicks », c'est-à-dire des paires de lettres qui apparaissaient plusieurs fois entre le message chiffré et sa version déchiffrée (n'oublions pas qu'il avait des machines Enigma à sa disposition pour tester). Comme Enigma est réversible, une correspondance A->J est équivalente à J->A.

Pour illustrer ce principe, prenons la phrase suivante que l'on suppose présente dans le message original :

« INTEROSURPRISELASEMAINEPROCHAINE ».

On intercepte le message chiffré :

« YAOPWMKLTBFZLVCXKTRTOMMYFLOWERS ».

Effectuons une première comparaison :

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

L'attaque de Turing se base sur la recherche de boucles entre le texte en clair et chiffré. La troisième lettre du message en clair « T » donne un « O » chiffré.

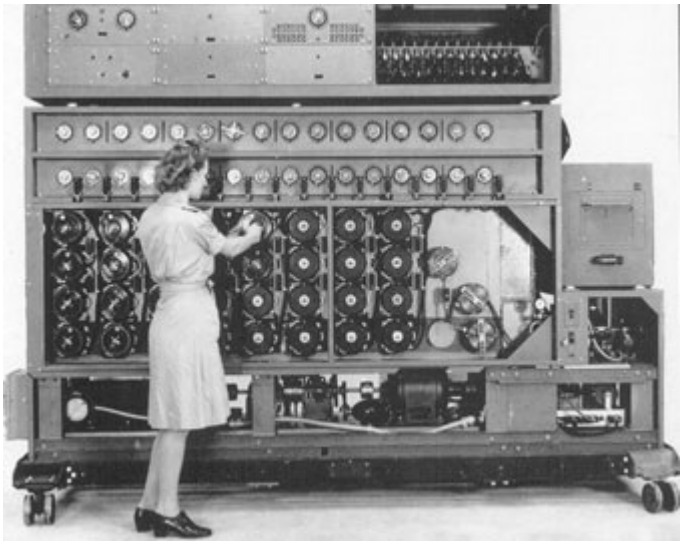
La 6e lettre du message en clair est un « O » qui se transforme en « M ».

La 19e lettre du message en clair est un « M » qui se transforme en « R ».

Finalement, la 9e lettre du crible est un « R », il se transforme en « T » et nous voilà donc avec une boucle car elle commence et se termine avec la même lettre.

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

La bombe testait en fait les configurations qui permettaient d'obtenir des boucles.



Une bombe de Turing[11]

Pour toutes les possibilités, on cherchait si le crible était compatible avec la boucle observée. Si ce n'était pas le cas, on continuait avec la configuration suivante.

Le nombre de possibilités se montait à 26*26*26*60 = 1.054.560, impossible à la main mais pas impossible pour la bombe de Turing. Pour calculer ces combinaisons, on ne pouvait se contenter d'une machine. Les Britanniques vont introduire une parallélisation astucieuse de la machine Enigma.

Un énorme travail devait être fait en amont pour trier les messages et retenir ceux qui étaient intéressants. Avec suffisamment de messages, il était possible de déterminer les paramètres journaliers. D'autres opérateurs ajoutaient un en-tête constante selon le type de message, par exemple WET (de Wetter, temps / météo en allemand) s'il s'agissait d'un rapport météo.

Plus tard dans la guerre, ces bulletins d'informations météorologiques furent une pièce maîtresse de la cryptanalyse : les météorologues en mer rédigeaient les messages et les envoyaient en Allemagne à l'aide d'un système de chiffrement moins robuste qu'Enigma (la météo n'étant pas une information véritablement secrète). Une fois arrivés dans les quartiers généraux, ces messages étaient expédiés quasiment sans modifications aux U-Boot, mais cette fois-ci en chiffrant avec Enigma. Les Alliés disposaient de textes clairs qu'ils pouvaient ainsi mettre en rapport avec les textes chiffrés d'Enigma dans le but d'établir des cribles.

5.2.6 Alan Turing

Le travail d'Alan Turing pour déchiffrer les messages allemands a profondément changé le cours de la seconde guerre mondiale.

11 La machine, pas la secrétaire !

L'homosexualité de Turing lui valut d'être persécuté et brisa sa carrière. En 1952, accusé d'«indécence manifeste et de perversion sexuelle» il est inculpé. Alors qu'il avait été consacré en 1951, en devenant membre de la Royal Society, à partir de 1952 il sera écarté des plus grands projets scientifiques.

En 1954, il meurt d'empoisonnement en mangeant une pomme contenant du cyanure. Beaucoup de gens pensent que cette mort est intentionnelle et elle fut présentée comme telle (« Blanche-Neige » de Walt Disney était son film préféré).

Parmi les nombreuses hypothèses circulant sur l'origine du logo d'Apple, l'une d'elle est que la pomme serait celle mordue par Turing[12]

5.2.7 Prix Turing

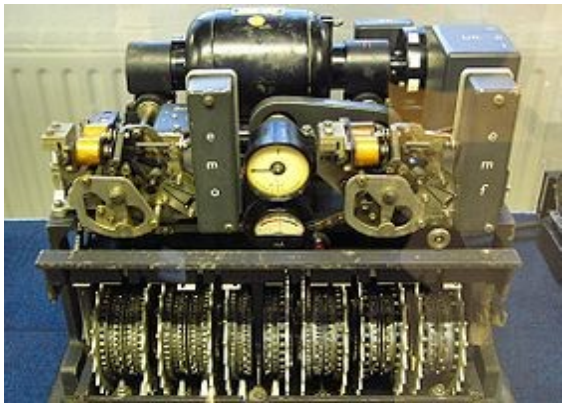
Depuis 1966, le prix Turing (Turing Award en anglais) est annuellement décerné par l'Association for Computing Machinery à des personnes ayant apporté une contribution scientifique significative à la science de l'informatique. Cette récompense est souvent considérée comme le prix Nobel de l'informatique.

Le titulaire du prix 2007 est le Grenoblois Joseph Sifakis

5.3 Lorenz Vs Colossus

Une autre histoire, peut-être encore plus impressionnante est celle des machines de Lorenz.

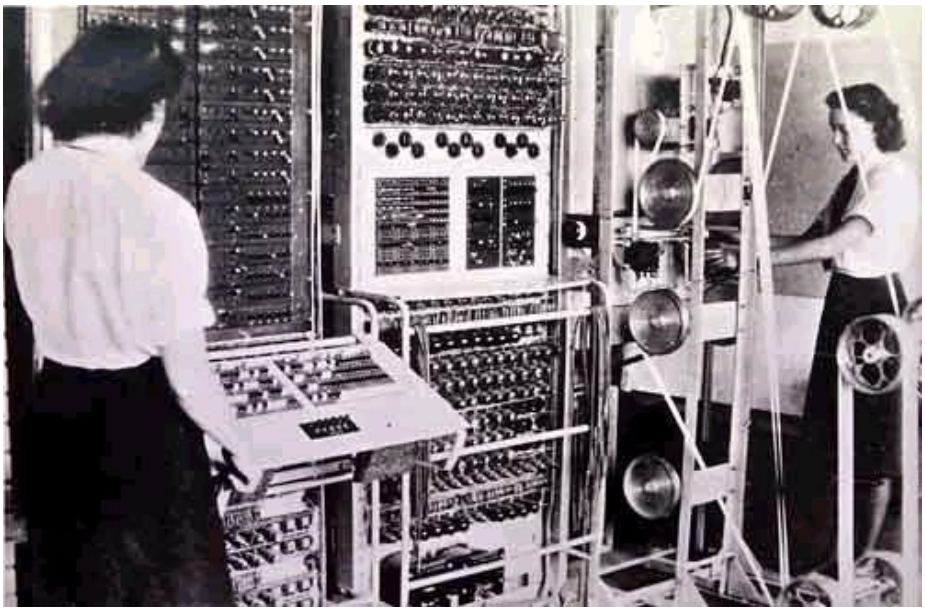
5.3.1 SZ40 et SZ42



Les machines de Lorenz SZ 40 et SZ 42 (Schlüsselzusatz, signifiant « pièce jointe chiffrée ») étaient des machines allemandes de chiffrement utilisées pendant la Seconde Guerre mondiale pour les envois par téléscripteur. Les cryptographes britanniques, qui se référaient de façon générale au flux des messages chiffrés allemands envoyés par téléscripteur sous l'appellation Fish (Poissons), ont nommé la machine et ses messages Tunny (Thons). Pendant que la renommée machine Enigma servait à l'armée, la machine de Lorenz était destinée aux communications de haut niveau entre le quartier-général du Führer et les quartiers-généraux des groupes d'armées, qui pouvaient s'appuyer sur cet appareil lourd.

Les cryptanalystes de Bletchley Park ont compris le fonctionnement de la machine dès janvier 1942 sans jamais en avoir vu un seul exemplaire. Cela fut possible à cause d'une erreur commise par un opérateur allemand. Le 30 août 1941, un message de 4 000 caractères fut transmis; cependant, le message n'ayant pas été reçu correctement à l'autre bout, celui-ci fut retransmis avec la même clé (une pratique formellement interdite par la procédure). De plus, la seconde fois le message fut transmis avec quelques modifications, comme l'utilisation de certaines abréviations. À partir de ces deux textes chiffrés, John Tiltman a été en mesure de reconstituer à la fois le texte en clair et le chiffrement. D'après le chiffrement, toute la structure de la machine fut reconstruite par W. T. Tutte.

12 L'hommage présumé a toujours été démenti par Apple (la pomme faisant référence à Newton).

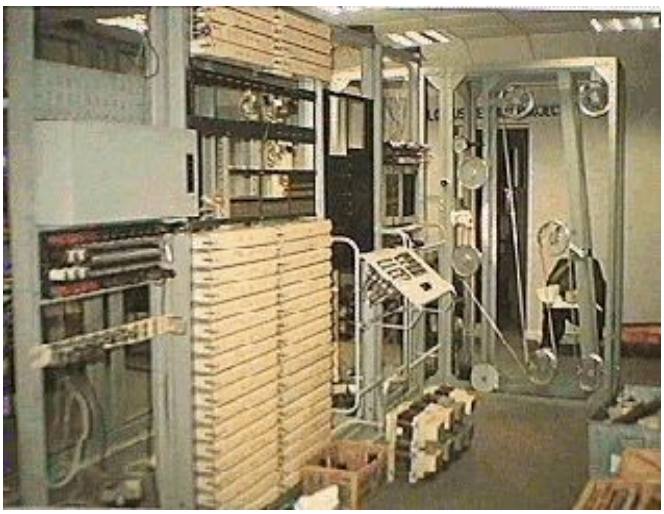


Plusieurs machines complexes furent élaborées par les Britanniques pour s'attaquer à ce type de messages. La première, de la famille connue sous le nom de "Heath Robinsons", utilisait des bandes de papier circulant rapidement le long de circuits électroniques logiques, pour décrypter le flux chiffré.

La suivante fut l'ordinateur Colossus, le premier ordinateur électronique numérique du monde (cependant, comme ENIAC, il ne comportait aucun logiciel embarqué et était programmé par l'intermédiaire de cartes enfichables, de commutateurs et de panneaux de connexion). Il était à la fois plus rapide et plus fiable que les "Heath Robinsons" ; son utilisation permit aux Britanniques de lire une grande part des communications de type "Tunny".

5.3.2 Le Colossus

Contrairement à Enigma, qui a été vaincue par par force brute, Lorenz a été cassé. Ainsi, le texte clair était re-calculé depuis le texte chiffré, sans récupérer la clé. Colossus a été conçu pour réaliser cette opération. Étant donné qu'il ne travaillait pas comme la machine Lorenz, tout le concept de Colossus était différent de celui de la machine d'encodage-décodage originale.



5.3.3 Reconstitution de la machine originale

Les 10 machines Colossus originales furent détruites après la guerre mondiale afin que leur fonctionnement reste secret. Sur la base d'une poignée de photographies et de quelques schémas électriques, le britannique Tony Sale conduisit un projet de reconstruction d'un Colossus. Ce projet aboutit en novembre 2007, après 14 ans de travail.

5.4 Conclusion

Mr X (Inventeur du 1er ordinateur)

Eh oui, le premier ordinateur au monde a été construit pour cryptanaliser un message

Et il est resté secret pendant plus de 30 ans.

Au fait, vous ne savez toujours pas le nom du génie qui à construit le Colossus ?

<- C'est lui.

Son nom est caché dans le paragraphe 5.2.5

Les travaux d'Alan Turing ont été déterminant pour la mise au point des ordinateurs (Machines de Turing).

On voit à travers ces deux histoires combien l'informatique doit à la cryptologie !

6 La cryptographie moderne

A partir de ce point, la cryptographie entre dans son ère moderne avec l'utilisation intensive des ordinateurs, c'est-à-dire à partir des années septante. Dans la cryptographie moderne, les textes sont remplacés par des chiffres. Via l'utilisation de la table ASCII, par exemple. Les problèmes sont de plus en plus mathématiques.

6.1 Les chiffrements par blocs

Inventé par Horst Feistel (1915 – 1990) et son chiffrement Lucifer[13].

C'est un chiffre à clé privé symétrique.

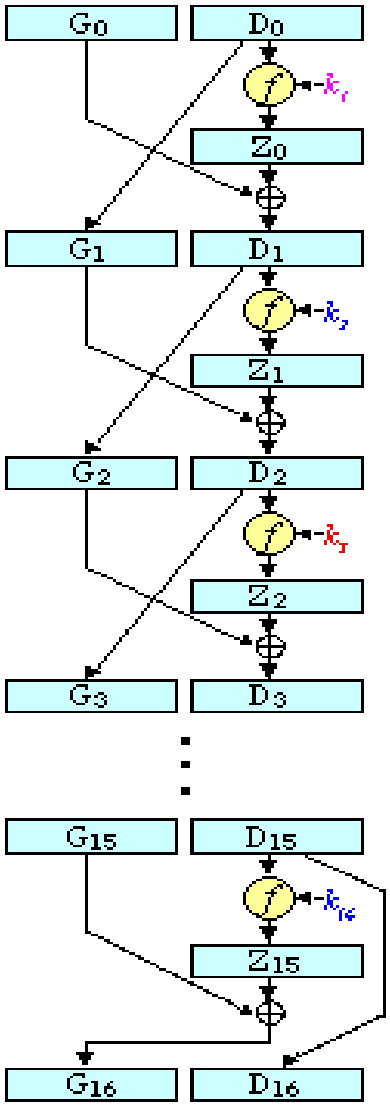
C'est à dire que non seulement on utilise la même clé pour chiffrer et déchiffrer (Clé privé), mais on utilise le même algorithme pour chiffrer et déchiffrer (Comme le ROT13)

6.1.1 Fonctionnement

Principe du schéma de Feistel avec un texte de 8 bits, une clé de 4 bits et 4 tours ou rondes.

Soit le « texte en clair » **0100 1001**

Et la clé : **1010**



Ronde N° 1

Étape 1 : On découpe le bloc à chiffrer en deux parties.

G_0 et D_0 (pour Gauche, ronde N°0 et Droite N°0)

G_0 : **0100** D_0 : **1001**

Étape 2 : On calcul Z_0 un ou exclusif entre la clé et D_0 :

1001
1010

0011

Étape 3 : On calcul D_1 un ou exclusif entre Z_0 et G_0 : **0111**

Étape 4 : G_1 vaut D_0 : **1001**

Et voila, la 1ere ronde est finie.

Ronde N° 2

On recommence les mêmes manipulations,

On obtient Z_1 : **1101**
G_2 : **0111**
D_2 : **0100**

Ronde N° 3

Idem,

On obtient Z_2 : **1110**
G_3 : **0100**
D_3 : **1001**

Ronde N° 4 :

attention, la dernière ronde utilise un schéma différent :

Étape 1 : On calcul Z_4 un ou exclusif entre la clé et D_3 (ça , ca na change pas) : **0011**

Étape 2 : On calcul G_4 un ou exclusif entre Z_4 et G_3 : **0111**

Étape 3 : D_4 vaut D_3 : **1001**

13 Le nom de « Lucifer » vient de « Demon » qui était obtenu en tronquant « Démonstration ». Le système d'exploitation sur lequel travaillait Feistel ne pouvant pas supporter des noms aussi longs ...

Et voila. On à le cryptogramme : **0111 1001**

6.1.2 Exercice : Déchiffrez le cryptogramme suivant

Sur le principe du schéma de Feistel avec un texte de 8 bits, une clé de 4 bits et 4 tours ou rondes.

Soit le « Cryptogramme » **0111 1001**

Et la clé : **1010**

6.2 D.E.S (*Data Encryption Standart*)

6.2.1 L'histoire

En mai 1973, le National Bureau of Standards américain demande la création d'un chiffrement utilisable par les entreprises. À cette époque, IBM dispose déjà de Lucifer, l'algorithme d'Horst Feistel.

En bonne logique, cet algorithme aurait dû être sélectionné par le NBS. En pratique, ce fut presque le cas : la NSA demanda à ce que Lucifer soit modifié, par ses soins. Ainsi fut créé le DES (Data Encryption Standart), qui fut adopté comme standard en novembre 1976.

Cela suscita des rumeurs selon lesquelles la NSA aurait volontairement affaibli l'algorithme, dans le but de pouvoir le casser. (L'algorithme initialement conçu par IBM utilisait une clé de 112 bits. L'intervention de la NSA a ramené la taille de clé à 56 bits.)

6.2.2 Utilisation

DES a été l'algorithme « officiel » de l'administration américaine jusqu'en 1999. C'est à dire que tout les document confidentiel défense et secret défense étaient cryptée DES.

Le système Unix qui date de cette période utilise encore le DES pour crypter les mots de passe.

6.2.3 La fin du DES

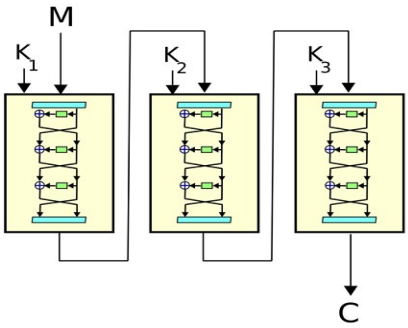
Le DES à fait l'objet de très nombreuses attaques. Mais c'est l'augmentation de la puissance des ordinateurs qui l'a rendu obsolète.

Un chiffrage DES offre 2^{54} possibilités de clés. Ce qui représente un nombre de combinaisons « réaliste » pour une grosse puissance de calcul. En 1998. Deep Crack un ordinateur conçu par IBM pour cet usage pouvait casser lune clé DES en moins d'une semaine.

Plus tard, le calcul distribué sut Internet, en utilisant les ordinateurs des particuliers, a prouvé son efficacité en cassant une clé en moins de 24 heures.

De plus, des attaques permettent de réduire le nombre de combinaisons à « seulement » 2^{43}. soit plusieurs milliards de fois moins que 2^{54}

6.2.4 Le triple DES



Le Triple DES (aussi appelé 3DES) est un algorithme de chiffrement symétrique enchaînant 3 applications successives de l'algorithme DES sur le même bloc de données de 64 bits, avec 2 ou 3 clés DES différentes. Cette utilisation de trois chiffrements DES a été développée par Walter Tuchman.

Même quand 3 clés de 56 bits différentes sont utilisées, la force effective de l'algorithme n'est que de 112 bits et non 168 bits, à cause d'une attaque type « rencontre au milieu ».

Bien que normalisé, bien connu, et assez simple à implémenter, il est assez lent.

6.3 A.E.S (Advanced Encryption Standard)

AES est un algorithme de chiffrement symétrique, choisi en octobre 2000 par le NIST pour être le nouveau standard de chiffrement pour les organisations du gouvernement des États-Unis.

6.3.1 Origine

Il est issu d'un appel à candidatures international lancé en janvier 1997 et ayant reçu 15 propositions. Au bout de cette évaluation, ce fut le candidat Rijndael (prononcer "Rayndal"), du nom de ses deux concepteurs Joan Daemen et Vincent Rijmen (tous les deux de nationalité belge) qui a été choisi.

6.3.2 Principe de fonctionnement

L'algorithme prend en entrée un bloc de 128 bits (la clé fait 128, 192 ou 256 bits. Les 128 bits en entrée sont « mélangés » selon une table définie au préalable.

Ces octets sont ensuite placés dans une matrice de 4x4 éléments et ses lignes subissent une rotation vers la droite.

L'incrément pour la rotation varie selon le numéro de la ligne.

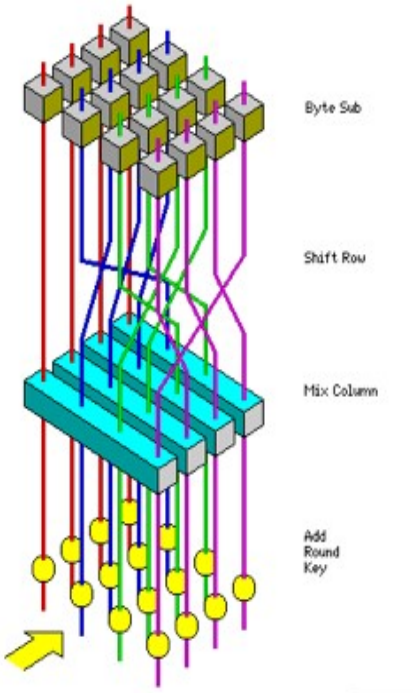
Une transformation est ensuite appliquée sur la matrice par un XOr avec une matrice clé.

Finalement, un XOr entre la matrice et une autre matrice permet d'obtenir une matrice intermédiaire.

Ces différentes opérations sont répétées plusieurs fois et définissent un «tour».

Pour une clé de 128, 192 ou 256, AES nécessite respectivement 10, 12 ou 14 tours.

L'algorithme AES n'est pas cassé à la date d'aujourd'hui.



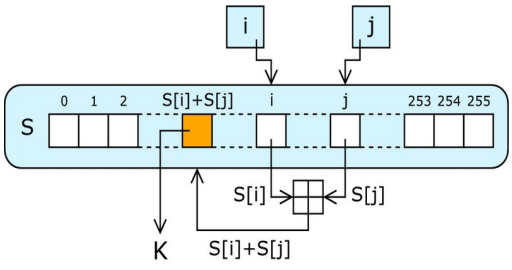
6.4 RC4

6.4.1 Histoire

RC4 a été conçu par Ronald Rivest (Le 'R' de RSA) en 1987. Officiellement nommé Rivest Cipher 4, l'acronyme RC est aussi surnommé Ron's Code

Il est utilisé dans des protocoles comme WEP, WPA ainsi que TLS. Les raisons de son succès sont liées à sa grande simplicité et à sa vitesse de chiffrement. Les implémentations matérielles ou logicielles étant faciles à mettre en œuvre.

6.4.2 Principe



La clef RC4 permet d'initialiser un tableau de 256 octets en répétant la clef autant de fois que nécessaire pour remplir le tableau.

Par la suite, des opérations très simples sont effectuées : les octets sont déplacés dans le tableau, des additions sont effectuées, etc. Le but est de mélanger autant que possible le tableau.

Au final on obtient une suite de bits pseudo-aléatoires qui peuvent être utilisés pour chiffrer le message via un XOR (Comme dans le cas d'un masque jetable).

6.4.3 Attaque de Fluhrer, Mantin et Shamir (attaque FMS)

En 2001, Scott Fluhrer, Itsik Mantin et Adi Shamir (Le 'S' de RSA, donc le collègue de Rivest) ont présenté une

nouvelle attaque.

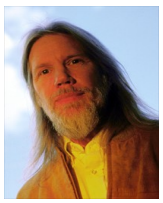
Les premiers octets du flux utilisé pour le chiffrement ne sont pas aléatoires. Si l'on se contente de concaténer la clé et le vecteur d'initialisation pour produire la nouvelle clé, alors il est possible de découvrir des informations en utilisant un grand nombre de messages chiffrés avec cette clé augmentée. C'est ce type d'attaque qui a été utilisée pour casser le WEP des réseaux sans fil, une action qui a abouti à la mise en place d'une version améliorée du chiffrement, à savoir WPA. Dans le cas du WEP, un vecteur d'initialisation de 24 bits est utilisé, ce qui permet de produire environ 16,8 millions clés supplémentaires à partir d'une clé principale (celle du point d'accès). Ce nombre est insuffisant eu égard les possibilités de l'attaque FMS.

7 Algorithmes de cryptographie asymétrique

7.1 Principe

Dans les années 1970, la cryptographie n'est plus seulement l'apanage des militaires. Les banques, pour la sécurité de leurs transactions, sont devenues de grandes consommatrices de messages codés. Les chiffrements disponibles alors, comme le DES, sont sûres, eu égard aux possibilités d'attaque contemporaines. Le problème essentiel est alors la distribution des clés, ce secret que l'envoyeur et le destinataire doivent partager pour pouvoir respectivement chiffrer et déchiffrer. Les armées et les états ont recours aux valises diplomatiques pour ces échanges, mais ceci n'est pas accessible aux civils...

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.



Whitfield Diffie

Martin Hellman

Whitfield.diffie@sun.com[14]

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.

- Vous envoyez 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 et surtout, la clé n'a pas voyagé !

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.**R**ivest, A.**S**hamir et L.**A**dleman trouvent une solution possible, la meilleure et la plus utilisée à ce jour, la cryptographie RSA. Le RSA repose sur la dichotomie suivante :

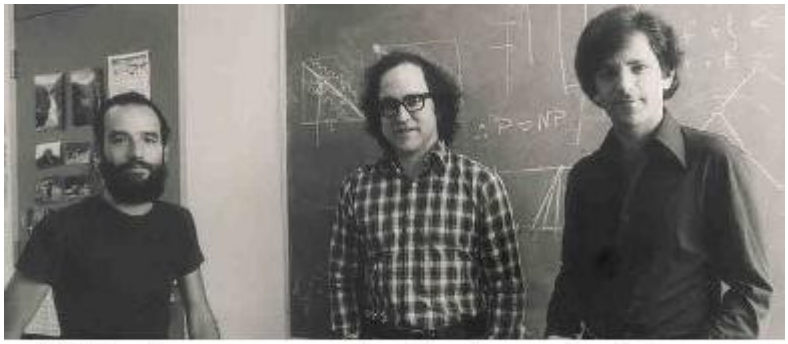
14 On fait quand même un boulot formidable : On peut communiquer avec les pionniers. Imaginez ce que donnerait un physicien pour avoir le mail de Newton

- 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.

7.2 RSA



Adi Shamir Ron Rivest Len Adleman

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.

7.2.1 Fonctionnement

1. **Création des clés :** Bob crée 5 nombres p,q,n,e et d :
   - p et q sont deux grands nombres premiers distincts. Leur génération se fait au hasard.
   - n est un entier tel que n = pq.
   - e est un entier premier et d un entier tel que ed=1 modulo [(p-1)(q-1)]. Autrement dit, ed-1 est un multiple de (p-1)(q-1).
- Exemple simplifié
  - Deux petits nombres 1er : p = 5 et q = 7
  - n = 5 * 7 = 35
  - n' = (5 – 1) * (7 – 1) = 24
  - On cherche e et d tels que :
    - e est premier
    - et ed = 1 modulo 24
      - ed = 1 Non, trop petit
      - ed = 25 Ok, mais e = d = 5 et alors Clé privé = clé publique => Faille de sécuritée.
      - ed = 49 Pareil, e = d
      - ed = 73 73 est 1er, raté
      - ed = 97 97 est 1er
      - ed = 121 11 au caré, encore raté
      - ed = 165 165 = 5 * 33, et 5 est 1er : Ok
  - On à Clé publique = RSA(35, 5) et clé privée = RSA(35, 33).
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 $C=M^e$ modulo 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 $D=C^d$ (modulo n) Il a donc reconstitué le message initial.

7.2.2 Exemple commenté

1. Choix de la clef

Alice choisit deux entiers premiers **p** et **q** (Ici on prend des nombres à 7 bits soit inférieurs à 2^7 (128)) et fait leur produit **n = p·q**. Puis elle choisit un entier **e** premier avec **(p-1)·(q-1)**. Enfin, elle publie dans un annuaire, par exemple sur le web, **sa clef publique**: **(RSA, n, e)**.

p = 53 q=97 donc n = 5141 e = 7 et d = 4279

2. Chiffrement

Bob veut donc envoyer un message à **Alice**. Il cherche dans l'annuaire la clef de chiffrement qu'elle a publiée. Il sait maintenant qu'il doit utiliser le système RSA avec les deux entiers **5141** et **7**. Il transforme en nombres son message en remplaçant par exemple chaque lettre par son rang dans l'alphabet.

"JEVOUSAIME" devient : "10 05 22 15 21 19 01 09 13 05".

Puis il découpe son message chiffré en blocs de même longueur (En partant de la droite) représentant chacun un nombre le plus grand possible tout en restant plus petit que **n**. Cette opération est essentielle, car si on ne faisait pas des blocs assez longs (par exemple si on laissait des blocs de 2 dans notre exemple), on retomberait sur un simple chiffre de substitution que l'on pourrait attaquer par l'analyse des fréquences.

Son message devient : "010 052 215 211 901 091 305"

Un bloc **B** est chiffré par la formule $\mathbf{C = B^e \bmod n}$, où **C** est un bloc du message chiffré que **Bob** enverra à **Alice**.

- 010 $(10^7) \equiv 5141 = 755$ 0755 (Sur 4 positions)
- 052 $(52^7) \equiv 5141 = 1324$ 1324
- 215 $(215^7) \equiv 5141 = 1324$ 2823
- ...

Après avoir chiffré chaque bloc, le message chiffré s'écrit : "0755 1324 2823 3550 3763 2237 2052".

3. Déchiffrement

Alice calcule à partir de **p** et **q**, **qu'elle a gardés secrets**, la clef **d** de déchiffrage (c'est sa **clef privée**). Celle-ci doit satisfaire l'équation $\mathbf{e \cdot d \bmod ((p-1)(q-1)) = 1}$. Ici, **d**=4279

Chacun des blocs **C** du message chiffré sera déchiffré par la formule $\mathbf{B = C^d \bmod n}$.

- 0755 $(755^{4279}) \equiv 5141 = 10$ 010
- 1324 $(1324^{4279}) \equiv 5141 = 52$ 052
- 2823 $(2823^{4279}) \equiv 5141 = 215$ 215
- ...

Elle retrouve : "010 052 215 211 901 091 305"

En regroupant les chiffres deux par deux et en remplaçant les nombres ainsi obtenus par les lettres correspondantes, elle sait enfin que Bob l'aime secrètement, sans que personne d'autre ne puisse le savoir.

7.2.3 Exercice

Connaissant la clé publique (119, 5) de ce cryptogramme RSA 7 bits

1. Calculez (par tout les moyens à votre disposition) p et q

2. Calculez d
3. Déchiffrez le cryptogramme suivant :

```
090 086 036 067 032 001 003 031 059 031
```

7.2.4 Le RSA est-il sûr?

La solidité du RAS repose uniquement sur la difficulté de décomposer le nombre n (public) en deux nombres premiers !

Dit autrement, il suffit de résoudre l'équation n = pq

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}).

```
1094173864157052742180970732204035761200373294544920599091384213147634998428893478471799725789126733249762575289978183379707653724402714674353159335433389 7=
1026395928297411057720541965739916759007165678080380668033419335217907113077 79
x 1066034883801684548209272203600128786792079585759892915222706082371930628086 43
```

Factorisation d'un entier à 155 chiffres.

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 jusqu'en ... 2079.

7.3 Signature électronique : être sûr de l'expéditeur.

La cryptographie à clé publique permet de s'affranchir du problème de l'échange de la clé, facilitant le travail de l'expéditeur. Mais comment s'assurer de l'authenticité de l'envoi? Comment être sûr que personne n'usurpe l'identité d'Alice pour vous envoyer un message? Comment être sûr qu'Alice ne va pas nier vous avoir envoyé ce message?

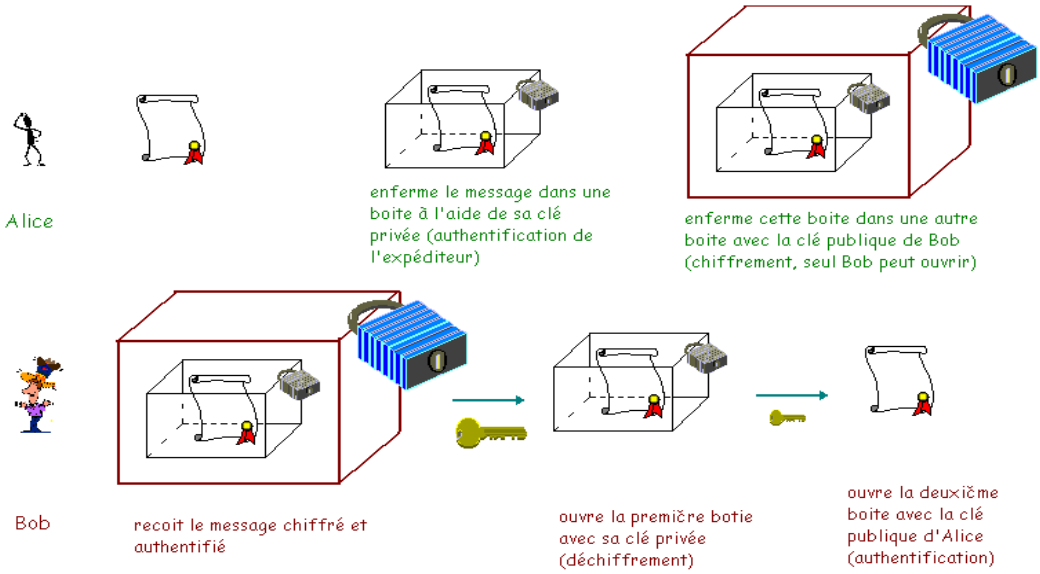
Là encore, la cryptographie à clé publique peut résoudre ce problème. Alice veut donc envoyer un message crypté à Bob, mais Bob veut s'assurer que ce message provient bien d'Alice.

Ici on va appeler Pa la clé publique d'Alice, Sa sa clé privée. Pb et Sb pour Bob. Et M le message à envoyer par Alice.

- **Phase d'envoi :** Alice calcule $S_A(M)$, à l'aide de sa clé secrète, puis $P_B(S_A(M))$, à l'aide de la clé publique de Bob.
- **Phase de réception :** A l'aide de sa clé privée, Bob calcule $S_B(P_B(S_A(M)))=S_A(M)$. Seul lui peut effectuer ce calcul (=sécurité de l'envoi). Puis il calcule $P_A(S_A(M))=M$. Il est alors sûr que c'est Alice qui lui a envoyé ce message, car elle-seule a pu calculer $S_A(M)$.

Contrairement à certains utilisateurs, vous ne confondrez plus « signature électronique » et « image d'une signature ajoutée en bas de page »

Remarquons pour terminer qu'une signature numérique est plus sûre qu'une signature papier, car elle est infalsifiable, inimitable : la signature change en effet à chaque message!



7.4 Certificat électronique : être sûr du destinataire.

Les algorithmes de chiffrement asymétrique sont basés sur le partage entre les différents utilisateurs d'une clé publique. Généralement le partage de cette clé se fait au travers d'un annuaire électronique (généralement au format LDAP) ou bien d'un site web.

Toutefois ce mode de partage a une grande lacune : rien ne garantit que la clé est bien celle de l'utilisateur a qui elle est associée. En effet un pirate peut corrompre la clé publique présente dans l'annuaire en la remplaçant par sa clé publique. Ainsi, le pirate sera en mesure de déchiffrer tous les messages ayant été chiffrés avec la clé présente dans l'annuaire.

Un certificat permet d'associer une clé publique à une entité (une personne, une machine, ...) afin d'en assurer la validité. Le certificat est en quelque sorte la carte d'identité de la clé publique, délivré par un organisme appelé autorité de certification (souvent notée CA pour Certification Authority).

L'autorité de certification est chargée de délivrer les certificats, de leur assigner une date de validité (équivalent à la date limite de péremption des produits alimentaires), ainsi que de révoquer éventuellement des certificats avant cette date en cas de compromission de la clé (ou du propriétaire).

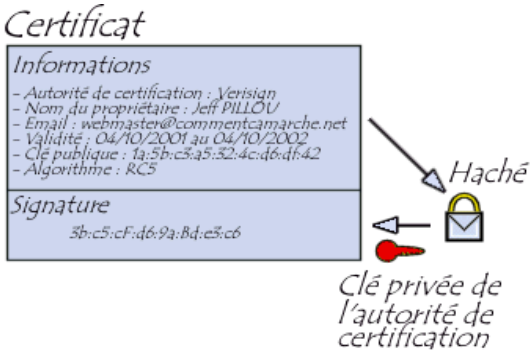
7.4.1 Structure d'un certificat

Les certificats sont des petits fichiers divisés en deux parties :

- La partie contenant les informations
- La partie contenant la signature de l'autorité de certification

La structure des certificats est normalisée par le standard X.509 de l'UIT, qui définit les informations contenues dans le certificat :

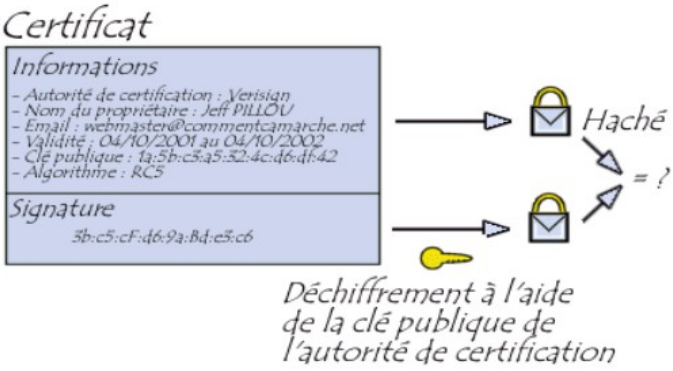
- La version de X.509 à laquelle le certificat correspond
- Le numéro de série du certificat
- L'algorithme de chiffrement utilisé pour signer le certificat
- Le nom (DN, pour Distinguished Name) de l'autorité de certification émettrice
- La date de début de validité du certificat
- La date de fin de validité du certificat
- L'objet de l'utilisation de la clé publique
- La clé publique du propriétaire du certificat
- La signature de l'émetteur du certificat (thumbprint)



L'ensemble de ces informations (informations + clé publique du demandeur) est signé par l'autorité de certification, cela signifie qu'une fonction de hachage crée une empreinte de ces informations, puis ce condensé est chiffré à l'aide de la clé privée de l'autorité de certification; la clé publique ayant été préalablement

largement diffusée afin de permettre aux utilisateurs de vérifier la signature avec la clé publique de l'autorité de certification.

Lorsqu'un utilisateur désire communiquer avec une autre personne, il lui suffit de se procurer le certificat du destinataire. Ce certificat contient le nom du destinataire, ainsi que sa clé publique et est signé par l'autorité de certification. Il est donc possible de vérifier la validité du message en appliquant d'une part la fonction de hachage aux informations contenues dans le certificat, en déchiffrant d'autre part la signature de l'autorité de certification avec la clé publique de cette dernière et en comparant ces deux résultats.



7.4.2 Signatures de certificats

On distingue différents types de certificats selon le niveau de signature :

- Les certificats auto-signés sont des certificats à usage interne. Signés par un serveur local, ce type de certificat permet de garantir la confidentialité des échanges au sein d'une organisation, par exemple pour le besoin d'un intranet. Il est ainsi possible d'effectuer une authentification des utilisateurs grâce à des certificats auto-signés.
- Les certificats signés par un organisme de certification sont nécessaires lorsqu'il s'agit d'assurer la sécurité des échanges avec des utilisateurs anonymes, par exemple dans le cas d'un site web sécurisé accessible au grand public. Le certificateur tiers permet d'assurer à l'utilisateur que le certificat appartient bien à l'organisation à laquelle il est déclaré appartenir.

7.4.3 Types d'usages

Les certificats servent principalement dans trois types de contextes :

- Le certificat client, stocké sur le poste de travail de l'utilisateur ou embarqué dans un conteneur tel qu'une carte à puce, permet d'identifier un utilisateur et de lui associer des droits. Dans la plupart des scénarios il est transmis au serveur lors d'une connexion, qui affecte des droits en fonction de l'accréditation de l'utilisateur. Il s'agit d'une véritable carte d'identité numérique utilisant une paire de clé asymétrique d'une longueur de 512 à 1024 bits.
- Le certificat serveur installé sur un serveur web permet d'assurer le lien entre le service et le propriétaire du service. Dans le cas d'un site web, il permet de garantir que l'URL et en particulier le domaine de la page web appartiennent bien à telle ou telle entreprise. Par ailleurs il permet de sécuriser les transactions avec les utilisateurs grâce au protocole SSL.
- Le certificat VPN est un type de certificat installé dans les équipement réseaux, permettant de chiffrer les flux de communication de bout en bout entre deux points (par exemple deux sites d'une entreprise). Dans ce type de scénario, les utilisateurs possèdent un certificat client, les serveurs mettent en oeuvre un certificat serveur et les équipements de communication utilisent un certificat particulier (Type IPSec).

7.5 Infrastructure à clé publique (PKI)

Une Infrastructure à clés publiques (ICP) ou Infrastructure de Gestion de Clefs (IGC) ou encore Public Key Infrastructure (PKI), est un ensemble de composants physiques (des ordinateurs, des équipements cryptographiques, des cartes à puces), de procédures humaines (vérifications, validation) et de logiciels (système et application) en vue de gérer le cycle de vie des certificats électroniques).

Une infrastructure à clés publiques délivre un ensemble de services pour le compte de ses utilisateurs.

En résumé, ces services sont les suivants :

- Enregistrement des utilisateurs (ou équipement informatique),

- Génération de certificats,
- Renouvellement de certificats,
- Révocation de certificats,
- Publication des certificats,
- Publication des listes de révocation (comprenant la liste des certificats révoqués),
- Identification et authentification des utilisateurs (administrateurs ou utilisateurs qui accèdent à la PKI),
- Archivage, séquestre et recouvrement des certificats (option).

7.5.1 Rôle d'une PKI

Une PKI délivre des certificats numériques. Ces certificats permettent d'effectuer des opérations cryptographiques, comme le chiffrement et la signature numérique qui offrent les garanties suivantes lors des transactions électroniques :

- confidentialité : seul le destinataire (ou le possesseur) légitime d'un bloc de données ou d'un message pourra en avoir une vision intelligible
- authentification : lors de l'envoi d'un bloc de données ou d'un message ou lors de la connexion à un système, on connaît sûrement l'identité de l'émetteur ou l'identité de l'utilisateur qui s'est connecté
- intégrité : on a la garantie qu'un bloc de données ou un message expédié n'a pas été altéré, accidentellement ou intentionnellement
- non-répudiation : l'auteur d'un bloc de données ou d'un message ne peut pas renier son œuvre.

7.5.2 Composants d'une PKI

Les PKI se scindent en 4 entités distinctes :

- L'Autorité de Certification (CA) qui a pour mission de signer les demandes de certificat (CSR : Certificate Signing Request) et de signer les listes de révocation (CRL : Certificate Revocation List).
- L'Autorité d'Enregistrement (AE ou RA) qui a pour mission de générer les certificats, et d'effectuer les vérifications d'usage sur l'identité de l'utilisateur final.
- L'Autorité de Dépôt (Repository) qui a pour mission de stocker les certificats numériques ainsi que les listes de révocation (CRL).
- L'Entité Finale (EE : End Entity) L'utilisateur ou le système qui est le sujet d'un certificat (En général, le terme "entité d'extrémité" (EE) est préféré au terme "sujet" afin d'éviter la confusion avec le champ Subject.)

Plus une 5eme entité, propre au droit Français (et pas définie spécifiquement par l'IETF) :

- L'Autorité de Séquestre (Key Escrow), cette entité a un rôle particulier, en effet lorsqu'on génère des certificats de chiffrement, on a l'obligation légale (en France) de fournir aux autorités un moyen de déchiffrer les données chiffrées pour un utilisateur de la PKI. C'est là qu'intervient le Séquestre, cette entité a pour mission de stocker de façon sécurisée les clés de chiffrement qui ont été générées, pour pouvoir les restaurer le cas échéant.

8 Fonctions de hachage

8.1 Principe

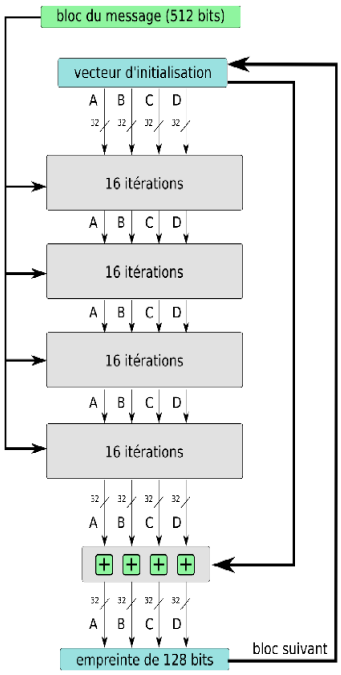
Une fonction de hachage est une fonction qui fait subir une succession de traitements à une donnée quelconque fournie en entrée pour en produire une « empreinte » servant à identifier la donnée initiale sans que l'opération inverse de décryptage soit possible.

Le terme hachage évite l'emploi de l'anglicisme hash. Le résultat de cette fonction est par ailleurs aussi appelé somme de contrôle, empreinte, résumé de message, condensé, condensat ou encore empreinte

cryptographique.

Les fonctions de hachage sont conçues pour effectuer un traitement de données rapide : calculer l'empreinte d'une donnée ne doit couter qu'un temps négligeable. Une fonction de hachage doit aussi éviter le plus possible les collisions (deux empreintes identiques alors que les données différaient)

Selon l'emploi de la fonction de hachage, il peut être souhaitable qu'un infime changement de la donnée en entrée (un seul bit, par exemple) entraine une perturbation conséquente de l'empreinte correspondante, rendant une recherche inverse par approximations successives impossible : on parlera d'effet avalanche.



8.2 MD5

En 1991, Ronald Rivest améliore l'architecture de MD4 et crée MD5 (Message Digest 5).

C'est une fonction de hachage cryptographique qui permet d'obtenir pour chaque message une chaine de 32 caractères (soit 128 bits) hexadécimaux avec une probabilité très forte que, pour deux messages différents, leurs empreintes soient différentes.

Ce quelle que soit la taille de l'information en entrée (de 0 octets à plusieurs gigas).

Cette transformation est donc irréversible (Dans le sens ou on ne peut pas trouver l'information en entrée à partir d'une somme MD5).

En 1996, une faille grave (possibilité de créer des collisions à la demande) est découverte. En 2004, une équipe chinoise découvre des collisions complètes. MD5 n'est donc plus considéré comme sûr au sens cryptographique.

MD5 reste encore utilisé comme outil de vérification lors des téléchargements (par exemple, en FTP). Les sites affichent encore souvent la signature en MD5 de leurs fichiers.

Le programme John the ripper permet de casser les MD5 triviaux par force brute. Des serveurs de "tables inverses" (à accès direct, et qui font parfois plusieurs gigaoctets) permettent de les craquer souvent en moins d'une seconde.

Aujourd'hui, il est par exemple possible de créer des pages HTML aux contenus très différents et ayant pourtant le même MD5. La présence de codes de "bourrage" placés en commentaires, visibles seulement dans la source de la page web, trahit toutefois les pages modifiées pour usurper le MD5 d'une autre.

8.3 SHA-1

SHA-1 (Secure Hash Algorithm) est une fonction de hachage cryptographique conçue par la NSA (1995), et publiée par le gouvernement des États-Unis comme un standard fédéral de traitement de l'information. Elle produit un résultat de 160 bits (40 caractères).

Même si on arrive à générer des collisions avec SHA-1. C'est-à-dire que l'on peut trouver deux messages au contenu aléatoire qui produisent la même signature. On ne sait toujours pas, à partir d'une signature donnée, forger un second message qui génère la même valeur. Or, c'est ce type d'attaque qui pourrait mettre en péril les applications comme PGP et l'authenticité des données.

Des versions offrant plus de sécurité sont également disponible : SHA-256, SHA-384 et SHA-512. Comme leur nom l'indique, ces versions fournissent des signatures de 256, 384 et 512 bits.

Exemple :

```
SHA-1(Les poules se sont échappées dés qu'on avait ouvert la porte) :
a187da360890f111566557aa6c197aa238dc533a
```

```
SHA-1(Les poules se sont échappées des con avait ouvert la porte) :
15166056114addee4bd717a1c45ae51389920a61
```

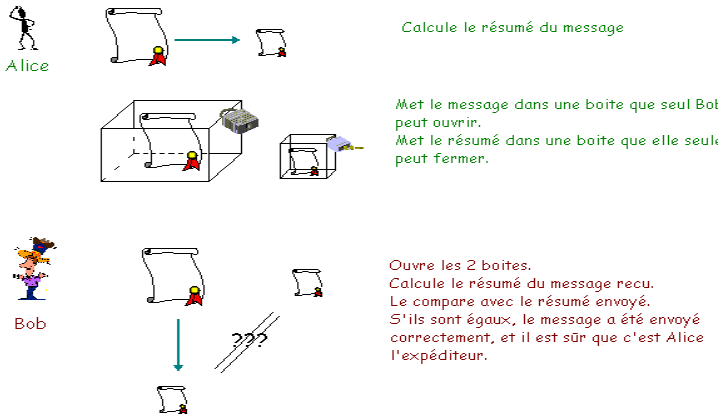
Comme vous le constatez, les deux phrases n'ont rien à voir entre elles ...

8.4 Retour sur les signatures électroniques

Le protocole vu au chapitre 7.3, s'il est fiable, est lent puisque deux fois plus lent qu'un algorithme à clé publique (lui-même déjà très lent!). En outre, il ne garantit pas l'intégrité du message, c'est-à-dire que celui-ci n'est pas altéré par des erreurs de transmission. L'utilisation des fonctions de hachage résout ces problèmes.

L'utilisation d'une fonction de hachage permet un gain de temps (et de place) pour le même effet :

- **Phase d'envoi :** Alice calcule h(M) - le résumé - et envoie à Bob $P_B(M)$ (calculé à l'aide de la clé publique de Bob) accompagné de $S_A(h(M))$.
- **Phase de réception :** Bob calcule $S_B(P_B(M))=M'$. Puis il calcule $P_A(S_A(h(M)))$, qu'il compare à h(M'). Si les quantités sont égales, il est sûr que c'est bien Alice qui a envoyé le message, et que celui-ci a été correctement transmis.



9 Application

9.1 Les cartes bancaires

Lorsqu'on introduit sa carte bleue dans un distributeur automatique, on imagine assez mal tout ce qui se passe. Chacun sait qu'il faut rentrer son code secret pour pouvoir débloquer le paiement, mais ceci n'est que la face visible de la sécurité des cartes bleues.

9.1.1 La carte à puce

La carte à puce a été créée par deux français, Roland Moreno et Michel Ugon, à la fin des années 1970. La puce est un petit ordinateur, avec un processeur (peu puissant) qui permet d'effectuer des calculs, une mémoire dont une partie est accessible en écriture (enregistrement de l'historique des transactions), une autre en lecture seule, et enfin une dernière en lecture cachée.

9.1.2 Mécanisme de paiement par carte bleue

Lorsque l'on introduit sa carte dans un terminal, il se déroule un processus en plusieurs étapes :

1. **Authentification de la carte :** elle se fait hors-ligne (sans appeler un centre de paiement de CB). Sur la carte sont inscrites certaines informations relatives au propriétaire (nom, numéro de carte, date de validité...), et une valeur de signature (VS). La VS est calculée une fois pour toute lors de la fabrication de la carte. On calcule d'abord Y, qui est une valeur numérique déduite des informations écrites dans la carte (par une fonction de hachage). Notons Y=f(info). La VS est alors calculée en utilisant la clé secrète S du groupement des cartes bancaires[15] (le GIE carte bancaire) : VS=S(Y). Lorsque la carte est introduite dans le terminal, celui lit les informations portées par la carte, et la valeur de signature VS. Il calcule alors Y1=f(info), et Y2=P(VS)=P(S(Y)), P étant la clé publique du GIE. Puis il compare Y1 et Y2 : pour qu'une carte soit valide, il faut que Y1=Y2.
2. **Code confidentiel :** Le code secret est stocké (sous forme chiffrée) à la fois dans la puce et sur la piste magnétique de la carte. Dans la premier cas, c'est la puce de la carte qui elle-même vérifie si le code entré est le bon, et transmet sa réponse au terminal.
3. **Authentification en ligne** (par le DES) : Cette étape n'est pas réalisée pour toutes les transactions, mais seulement pour celles dépassant un certain montant (avec affichage de "Autorisation" su l'écran du

15 Ces fonctions à clé secrète et à clé publique sont basées sur le RSA. Le modulo public français est un nombre connu entre 768 et 1024 bits, produit de 2 premiers inconnus. L'exposant est e=3.

terminal). Le terminal interroge un centre de contrôle à distance, qui envoie à la carte une valeur aléatoire x. La carte calcule y=f(x,K), où K est une clé secrète, inscrite dans la partie illisible de la carte, et f est la fonction de chiffrement du DES (ou du triple DES depuis 1999). La valeur y est retransmise au centre, qui lui-même calcule f(x,K), et donne ou non l'autorisation. Remarquons que ceci nécessite que le centre connaisse la clé secrète de toutes les cartes.

9.1.3 L'affaire Humpich

En 1998, l'affaire Serge Humpich fait la une des journaux. Cet informaticien a montré qu'il était possible de fabriquer de toute pièce une fausse carte qui permettait de payer chez un comerçant. Serge Humpich avait contourné deux systèmes de sécurité :

1. D'une part, il a fabriqué des "yes card", c'est-à-dire des cartes à puce qui, quel que soit le code secret entré, renvoie "code bon".
2. D'autre part, il a contourné l'authentification hors-ligne RSA. La sécurité de ce système repose, rappelons-le, sur la difficulté à factoriser un "grand' entier. Or, en 1998, le n utilisé par le GIE avait pour taille 320 bits (taille inchangée depuis 1990). A cette époque, factoriser un tel entier n'était plus impossible (le record se situait à 512 bits), et Humpich, en utilisant simplement un logiciel japonais de factorisation, a réussi à factoriser le n du GIE, et à découvrir la clé secrète S.

La 3ème fonction de sécurité, elle, est toujours restée valide. Dans cet affaire, le GIE a pêché d'abord par excès de confiance (les 320 bits étaient suffisants en 1990, plus en 1998) et aussi par manque de communication. Depuis, le tir a été rectifié : le n a changé, et est désormais long de 768 bits, l'authentification en ligne est passée du DES au 3DES.

9.1.4 Autres types de fraude

La fraude la plus répandue avec les C.B. est beaucoup plus simple que la faille exploitée par Humpich. Les 16 chiffres de votre C.B. suffisent pour commander sur Internet ou par correspondance. Ils ne sont pas choisis au hasard (tous les nombres à 16 chiffres ne donnent pas un numéro de CB valide). Avant 2001, les terminaux inscrivaient sur les facturettes les numéros de CB. Voila pourquoi les voleurs raffolaient de ces petits tickets.

Signalons que les risques de fraude en utilisant simplement le numéro de carte bleue (sans code secret) sont normalement couverts par les banques : elles doivent rembourser le client qui conteste un prélèvement de ce type. Signalons enfin qu'il existe des sites Internet où on explique comment calculer un numéro de CB valable.

9.2 SSL et TLS

9.2.1 SSL

SSL (Secure Sockets Layers, ou couche de sockets sécurisée) est un procédé de sécurisation des transactions effectuées via Internet. Le standard SSL a été mis au point par Netscape, en collaboration avec Mastercard, Bank of America, MCI et Silicon Graphics. Il repose sur un procédé de cryptographie par clef publique afin de garantir la sécurité de la transmission de données sur internet. Son principe consiste à établir un canal de communication chiffré entre deux machines (un client et un serveur) après une étape d'authentification.

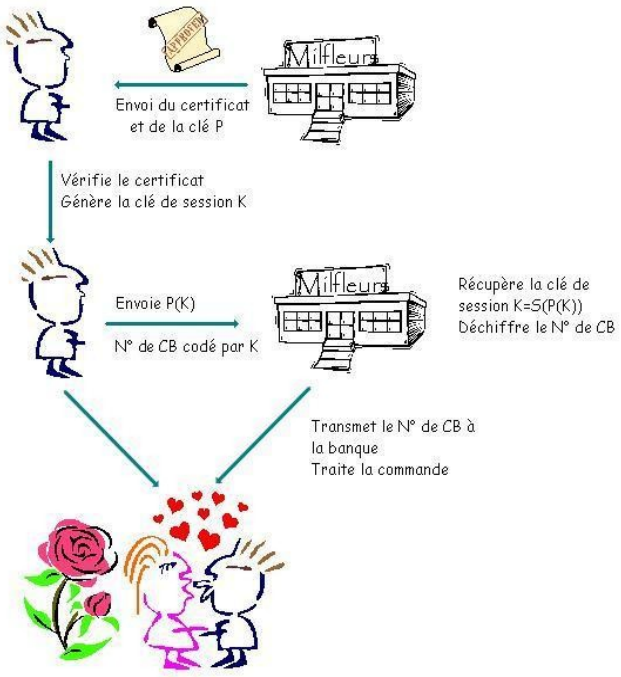
Le système SSL est indépendant du protocole utilisé, ce qui signifie qu'il peut aussi bien sécuriser des transactions faites sur le Web par le protocole HTTP que des connexions via le protocole FTP, POP ou IMAP. En effet, SSL agit telle une couche supplémentaire, permettant d'assurer la sécurité des données, située entre la couche application et la couche transport (protocole TCP par exemple).

De cette manière, SSL est transparent pour l'utilisateur (entendez par là qu'il peut ignorer qu'il utilise SSL). Par exemple un utilisateur utilisant un navigateur internet pour se connecter à un site de commerce électronique sécurisé par SSL enverra des données chiffrées sans aucune manipulation nécessaire de sa part.

Un serveur web sécurisé par SSL possède une URL commençant par https://, où le "s" signifie bien évidemment secured (sécurisé).

Ssl utilise la cryptographie à clé publique, la cryptographie à clé secrète, et les certificats électroniques. Supposons que Bob commande un bouquet à la boutique en ligne "Milfleurs". Nous assistons à l'échange suivant :

- Bob se connecte au site sécurisé de Milfleurs. Ce dernier lui envoie un certificat électronique, qui donne sa clé publique d'échange P, ainsi que le certificat qui prouve qu'il s'agit bien de lui.
- Bob (en fait, son navigateur) vérifie le certificat. Il se met d'accord avec le serveur distant sur un système cryptographique commun à clé secrète à utiliser (en pratique, le plus sûr qu'ils ont en commun). Puis il choisit au hasard une clé pour cet algorithme, la clé de session.
- Bob transmet cette clé au serveur, en utilisant la clé publique de celui-ci. Les 2 protagonistes sont alors en possession d'une même clé de session qu'ils sont les seuls à posséder. Bob peut envoyer son numéro de carte bancaire en toute sécurité.



9.2.2 TLS

Transport Layer Security (TLS), est l'évolution de SSL (SSL version 3).

Il utilise à la fois un chiffrement asymétrique (pour l'authentification), c'est l'algorithme **RSA**, et à la fois un chiffrement symétrique pour la transmission des informations (le **AES**). On y adjoint une fonction de hachage, (le **MD5)**, pour s'assurer que les données sont transmises sans être corrompues.

En 2008, **TLS** est utilisé par la plupart des navigateurs Web. On reconnaît qu'une transaction est sécurisée lorsqu'une clé ou un cadenas fermé s'affiche dans un coin inférieur de l'écran ainsi que dans la barre d'adresse, l'adresse commence par http**s**://...

9.3 PGP

Le PGP (Pretty Good Privacy) est un algorithme de chiffrement à destination des particuliers. Il est surtout utilisé pour chiffrer des messages envoyés par courrier électronique, même s'il peut aussi être utilisé pour chiffrer tous les fichiers. PGP a été mis au point en 1991 par Philip Zimmermann, et ceci lui valut divers problèmes avec la justice.

D'une part, le PGP utilise l'algorithme RSA, qui est breveté aux Etats-Unis. D'autre part, la NSA a tout fait pour tenter d'empêcher la diffusion du PGP.

La plainte de la NSA a été classée sans suite par le gouvernement début 1996 sous la pression des internautes qui se sont mobilisés pour défendre Zimmermann.

«If privacy is outlawed, only outlaws will have privacy», soit, en français : « Si l'intimité est mise hors la loi, seuls les hors-la-loi auront une intimité. »


Philip Zimmermann[16]

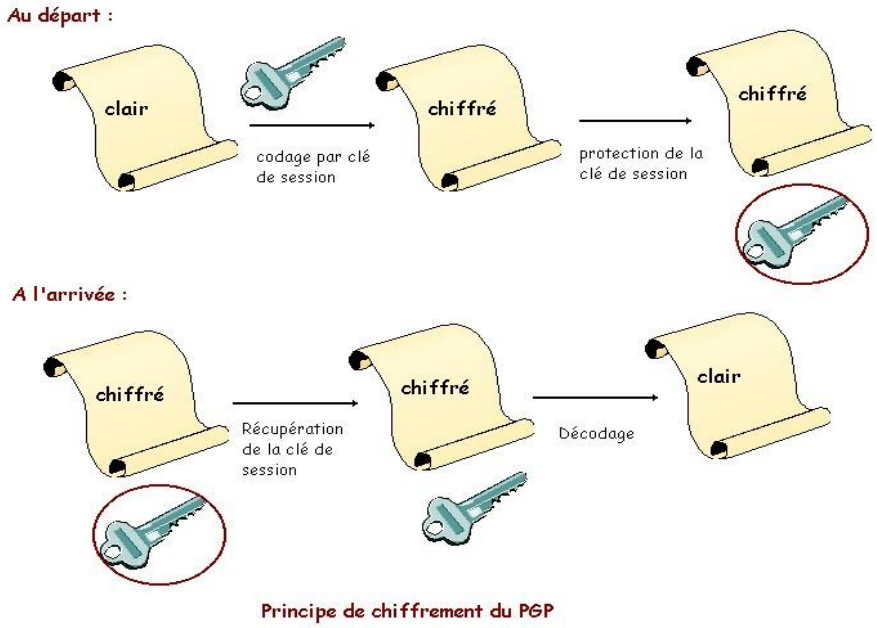
PGP utilise le meilleur de la cryptographie symétrique (rapidité du chiffrement) et de la cryptographie asymétrique (sécurité de l'échange de clés). Il fonctionne suivant le principe suivant :

1. Compression : le texte à envoyer est compressé. Cette étape permet de réduire le temps de transmission des données, et améliore également la sécurité. En effet, la compression détruit les modèles du texte (fréquences des lettres, mots répétés). Et on sait que ces modèles sont souvent utilisés dans les analyses cryptographiques.
2. Chiffrement du message : un mot de passe aléatoire est générée (On parle ici de clé de session), et le message est

16 Lire (en Français) Pourquoi j'ai écrit PGP par Philip Zimmermann : http://biblioweb.samizdat.net/article.php3?id_article=4

chiffré par un algorithme symétrique à l'aide de cette clé de session. L'algorithme utilisé a varié au cours du temps : il s'agissait au début de DES, puis de CAST.

3. Chiffrement de la clé de session : la clé de session est chiffrée en utilisant la clé publique du Destinataire (et l'algorithme RSA).
4. Envoi et réception du message : l'expéditeur envoie le couple message chiffré / clé de session chiffrée au Destinataire. Celui récupère d'abord la clé de session, en utilisant sa clé privée, puis il déchiffre le message grâce à la clé de session.



Principe de chiffrement du PGP

Le PGP implémente aussi les fonctions de certificat et de signature électroniques. Il n'y a pas d'autorité centrale de certification, mais un grand rôle joué à la proximité sociale : les amis de mes amis sont mes amis !

En 2006, Zimmermann a créé Zfone, un logiciel de chiffrement de communication de téléphonie sur Internet au standard ouvert SIP, fonctionnant en P2P.

10 Cryptanalyse

La cryptanalyse s'oppose, en quelque sorte, à la cryptographie. En effet, si déchiffrer consiste à retrouver le clair au moyen d'une clé, cryptanalyser c'est tenter de se passer de cette dernière.

Même si on décrit les cryptanalyses comme des «briseurs de codes», il convient de remarquer qu'un algorithme est considéré comme cassé lorsqu'une attaque permet de retrouver la clé en effectuant moins d'opérations que via une attaque par force brute. L'algorithme ainsi cassé ne devient pas inutile pour autant, mais son degré de sécurité, c'est-à-dire le nombre moyen d'opérations nécessaires pour le déchiffrer, s'affaiblit.

10.1 Familles d'attaques cryptanalytiques

10.1.1 L'attaque par force brute

L'attaque par force brute consiste à tester toutes les solutions possibles de mots de passe ou de clés.

Nombre de secondes dans un jour	86 400 secondes (~ 10^5)
Nombre de secondes dans une année	31 536 000 secondes (~ 10^8)
Nombre de secondes depuis la création de l'univers (13,7 milliard d'années)	432 043 200 000 000 000 secondes (~ 10^{17})
Durée de vie d'un proton	~ 10^{30} années

Masse du soleil	~ 10^{33} grammes
Nombre d'atomes dans un gramme de matière	~ 10^{24} atomes
Nombre d'atomes dans l'univers	~ 10^{80} atomes

Le but de ces quelques nombres est de montrer l'inutilité des attaques type force brute.

En effet, un cryptanalyste en herbe pourrait se dire "Ce message est chiffré avec l'algorithme AES-256, il y a 2^{256} clés possibles (soient environ 10^{76} clés), testons les toutes pour déchiffrer le message".

A raison de 100 milliards de tentatives par seconde (ce qui est énorme), il faudrait 10^{58} secondes pour tester toutes les clés.

Soit beaucoup plus que l'âge de l'univers...

10.1.2 L'analyse fréquentielle

Voir chapitre 3.3

10.1.3 L'indice de coïncidence

Voir chapitre 3.6

10.1.4 L'attaque par mot probable

Voir chapitre 3.4.

10.1.5 L'attaque par dictionnaire

L'attaque par dictionnaire consiste à tester tous les mots d'une liste comme mot clé. Elle est souvent couplée à l'attaque par force brute.

10.1.6 Attaque par paradoxe des anniversaires

Le paradoxe des anniversaires est un résultat probabiliste qui est utilisé dans les attaques contre les fonctions de hachage. Ce paradoxe permet de donner une borne supérieure de résistance aux collisions d'une telle fonction. Cette limite est de l'ordre de la racine de la taille de la sortie, ce qui signifie que, pour un algorithme comme MD5 (empreinte sur 128 bits), trouver une collision quelconque avec 50% de chance nécessite 264 hachages d'entrées distinctes.

Le paradoxe des anniversaires, est à l'origine, une estimation probabiliste du nombre de personnes que l'on doit réunir pour avoir une chance sur deux que deux personnes de ce groupe aient leur anniversaire le même jour. Il se trouve que ce nombre est 23, ce qui choque un peu l'intuition. À partir d'un groupe de 57 personnes, la probabilité est supérieure à 99 %.

Cependant, il ne s'agit pas d'un paradoxe dans le sens de contradiction logique ; c'est un paradoxe, dans le sens où c'est une vérité mathématique qui contredit l'intuition : la plupart des gens estiment que cette probabilité est très inférieure à 50 %.

10.2 Cryptanalyse moderne

Dès les années 70 apparaissent les méthodes de chiffrement modernes par blocs comme DES. Il sera passablement étudié et attaqué ce qui mènera à des attaques majeures dans le monde de la cryptographie. Les méthodes présentées ci-dessous ne sont pas vraiment génériques et des modifications sont nécessaires pour attaquer un type de chiffrement donné.

Souvent, on ne s'attaque pas à une version complète de l'algorithme de chiffrement mais une variante avec moins de tours (dans le cas des schémas de type Feistel ou les fonctions de hachage). Cette analyse

préliminaire, si elle permet de déceler des vulnérabilités, laisse entrevoir une attaque sur l'algorithme complet.

10.2.1 Familles d'attaques répertoriées :

- Cryptanalyse linéaire
- Cryptanalyse différentielle
- Cryptanalyse différentielle tronquée
- Cryptanalyse différentielle d'ordre supérieur
- Cryptanalyse par différentielles impossibles
- L'attaque boomerang
- L'attaque rectangle
- Cryptanalyse différentielle-linéaire
- Cryptanalyse χ^2
- Cryptanalyse quadratique
- Cryptanalyse modulo n
- Compromis temps/mémoire
- Attaques sur les modes opératoires
- Attaque par rencontre au milieu
- Attaques sur les systèmes asymétriques
- Attaques par canaux auxiliaires

10.3 Attaques par canaux auxiliaires

Revenons sur cette famille d'attaque, dans la quelle on range plein de protocoles qui s'apparentent parfois plus à la bidouille qu'a la cryptanalyse

10.3.1 Un peu d'histoire : Cap'tain Crunch

Dans les années 60, aux USA, une ligne longue distance inoccupée émet en permanence une tonalité de 2600 Hz, indiquant à un central téléphonique qu'elle est prête à recevoir un appel.

Or, un électronicien, John Draper à remarqué que le sifflet pour enfants que Quaker Oats offrait avec ses céréales était accordé sur le la aigu et permettait de reproduire cette tonalité.

Cette propriété découverte par hasard a été exploitée par les phreakers[17] [18] pour passer gratuitement des appels longue distance.

Son surnom provenait des boîtes de céréales Cap'n Crunch[19].

Bref, tout ça pour montrer que parfois, l'attaque vient d'où on ne l'attend pas.

10.3.2 Attaque temporelle

Une attaque temporelle consiste à estimer et analyser le temps mis pour effectuer certaines opérations cryptographiques dans le but de découvrir des informations secrètes. Certaines opérations peuvent prendre plus de temps que d'autres et l'étude de ces informations temporelles peut être précieuse pour le cryptanalyste. La mise en œuvre de ce genre d'attaque est intimement liée au matériel ou au logiciel attaqué.

Attaques sur la cryptographie asymétrique

Les algorithmes d'exponentation modulaire sont coûteux, le temps d'exécution dépend linéairement du nombre de bits à '1' dans la clé. Si connaître le nombre de '1' n'est pas une information toujours suffisante pour trouver la clé, le recoupement statistique entre plusieurs chiffrements avec cette clé peut offrir de nouvelles possibilités au cryptanalyste.

17 Phreaker : Hacker de téléphonne (*phone* + *freak*)
18 Hacker : Littéralement "Bidouilleur" Il n'y a aucune connotation négative dans l'expression « Hacker »
19 Les céréales, c'est dangereux : Cap'tain Crunch a été condamné à deux mois de prison en 1976

Attaques sur un réseau

En 2003, Boneh et Brumley ont démontré une attaque pratique contre des serveurs SSL. Leur cryptanalyse est basée sur des vulnérabilités découvertes dans les implémentations du théorème des restes chinois. L'attaque fut toutefois menée à travers un réseau de taille limitée mais elle montrait que ce type d'attaque était sérieuse et praticable en l'espace de quelques heures. Les implémentations furent améliorées pour limiter les corrélations entre la clé et le temps de chiffrement.

Attaque sur les chiffrements par bloc

Les chiffrements par bloc sont en général moins sensibles aux attaques temporelles, la corrélation entre la clé et les opérations étant plus limitées, mais celles-ci existent quand même. La plupart reposent sur les temps mis pour accéder aux différentes tables (par exemple les S-Boxes).

En 2005, Daniel Bernstein a démontré qu'une attaque contre une implémentation vulnérable d'AES était possible à partir du cache des processeurs modernes des PC (AMD ou Intel). Bernstein reproche au NIST d'avoir négligé ces problèmes lors du concours AES, il ajoute que le NIST s'est trompé en partant du principe que le temps d'accès aux tables était constant.

10.3.3 Attaque par sondage

Une attaque par sondage est ce que l'on qualifie d'attaque invasive, c'est-à-dire que la mise en œuvre de celle-ci peut détériorer, voire détruire le circuit à analyser.

Le principe d'une attaque par sondage (appelée probing attack) est d'espionner l'activité électrique d'un composant électronique du circuit en positionnant une sonde suffisamment proche dudit composant.

En récoltant des données de cette manière, l'attaquant peut être en mesure de déduire tout ou une partie du secret du circuit cryptographique.

Tout d'abord il faut préparer le circuit à analyser. Il faut souvent le tremper dans l'acétone, puis « gratter » sa surface (généralement couverte d'un enduit chimique) pour mettre à nu les couches supérieures de métal.

Pour cela il faut approcher très près de l'équipotentielle à espionner une pointe métallique (typiquement en tungstène) qui réagit au passage d'un bit sur celle-ci (en fait un changement ou non d'état). Avec un oscilloscope suffisamment précis et un chronométrage très rigoureux, on peut ainsi déterminer les bits transitant par le bus.

10.3.4 Analyse de consommation

En cryptanalyse de matériel cryptographique, l'analyse de consommation consiste à étudier les courants et tensions entrants et sortants d'un circuit dans le but de découvrir des informations secrètes comme la clé de chiffrement. Certaines opérations, plus coûteuses, augmentent la consommation électrique du circuit, notamment par l'utilisation de plus de composants (analogiques ou logiques). Cette analyse des variations et des pics permet de tirer des informations précieuses pour le cryptanalyste.

Une attaque basée sur les temps de réponse a été menée par Serge Vaudenay sur TLS/SSL, ce qui a forcé les concepteurs du standard à faire une mise à jour critique. Les constructeurs de puces de chiffrement visent à aplanir la courbe de consommation électrique pour dissimuler les opérations sous-jacentes.

10.3.5 Attaque par faute

les attaques par faute sont une famille de techniques qui consistent à produire volontairement des erreurs dans le cryptosystème. Ces attaques peuvent porter sur des composants matériels (cryptoprocesseur) ou logiciels. Elles ont pour but de provoquer un comportement inhabituel des opérations cryptographiques dans le but d'en extraire des informations secrètes (comme une clé de chiffrement). Une attaque par faute peut être couplée à d'autres méthodes comme l'analyse de la consommation ou une attaque temporelle.

Les attaques sont possibles sous l'hypothèse que l'attaquant peut affecter l'état interne du système en écrivant des valeurs par exemple en mémoire ou sur un bus informatique.

Un exemple classique d'attaque par faute concerne RSA et en particulier le produit des deux nombres premiers (p, q) qui composent en partie la clé (le secret, donc) du système. Le principe est de faire en sorte qu'un de ces deux nombres soit modifié juste avant leur produit p * q. Il est en effet très difficile de retrouver p ou q en fonction de p * q. Or, si on arrive à transformer p en p' (non premier), on peut retrouver beaucoup plus

facilement q en calculant le pgcd de p' * q.

11 Stéganographie

Si la cryptographie est l'art du secret, la stéganographie est l'art de la dissimulation : l'objet de la stéganographie n'est pas de rendre un message inintelligible à autre que qui de droit mais de le faire passer inaperçu. Si on utilise le coffre-fort pour symboliser la cryptographie, la stéganographie revient à enterrer son argent dans son jardin. Bien sûr, l'un n'empêche pas l'autre, on peut enterrer son coffre dans son jardin.

11.1 Histoire

11.1.1 301

Dans son Enquête, l'historien grec Hérodote (484-445 av. J.-C.) rapporte ainsi une anecdote qui eut lieu au moment de la seconde guerre médique. En 484 avant l'ère chrétienne, Xerxès, fils de Darius, roi des Perses, décide de préparer une armée gigantesque pour envahir la Grèce). Quatre ans plus tard, lorsqu'il lance l'offensive, les Grecs sont depuis longtemps au courant de ses intentions. C'est que Démarate, ancien roi de Sparte réfugié auprès de Xerxès, a appris l'existence de ce projet et décide de transmettre l'information à Sparte : « il prit une tablette double, en gratta la cire, puis écrivit sur le bois même les projets de Xerxès ; ensuite il recouvrit de cire son message : ainsi le porteur d'une tablette vierge ne risquait pas d'ennuis ».

Un autre passage de la même œuvre fait également référence à la stéganographie : Histiée incite son gendre Aristagoras, gouverneur de Milet, à se révolter contre son roi, Darius, et pour ce faire, « il fit raser la tête de son esclave le plus fidèle, lui tatoua son message sur le crâne et attendit que les cheveux eussent repoussé ; quand la chevelure fut redevenue normale, il fit partir l'esclave pour Milet ».

En Chine, on écrivait le message sur de la soie, qui ensuite était placée dans une petite boule recouverte de cire. Le messager avalait ensuite cette boule.

Dès le Ier siècle av. J.-C., Pline l'Ancien décrit comment réaliser de l'encre invisible (ou "encre sympathique"). Les enfants de tous les pays s'amusent à le faire en écrivant avec du lait ou du jus de citron : le passage de la feuille écrite sous un fer à repasser chaud révèle le message.

Durant la Seconde Guerre mondiale, les agents allemands utilisaient la technique du micropoint de Zapp, qui consiste à réduire la photo d'une page en un point d'un millimètre ou même moins. Ce point est ensuite placé dans un texte normal. Le procédé est évoqué dans une aventure de Blake et Mortimer, SOS météores.

11.2 Techniques rendues possibles par l'ordinateur

11.2.1 Usage des bits de poids faible d'une image

L'idée est de prendre un message et de le modifier de manière aussi discrète que possible afin d'y dissimuler l'information à transmettre. Le message original est le plus souvent une image. La technique de base (dite LSB pour Least Significant Bit) consiste à modifier le bit de poids faible des pixels codant l'image : une image numérique est une suite de points, que l'on appelle pixel, et dont on code la couleur à l'aide d'un triplet d'octets par exemple pour une couleur RGB sur 24 bits. Chaque octet indique l'intensité de la couleur correspondante (rouge, vert ou bleu) par un niveau parmi 256. Passer d'un niveau n au niveau immédiatement supérieur (n+1) ou inférieur (n-1) ne modifie que peu la teinte du pixel, or c'est ce que l'on fait en modifiant le bit de poids faible de l'octet.

Exemple

Considérons l'image de 2 pixels :

Chaque pixel d'une image est représenté par 3 nombres codés sur 8 bits : R représente l'intensité du rouge (un entier entre 0 et 255), G celle du vert, B celle du bleu. Si l'on modifie les 2 bits de droite de R, on modifie très peu sa valeur (au plus, de 3), et cela est imperceptible à l'oeil humain. On remplace alors les 2 bits de droite de R par les 2 premiers bits du message. Puis on continue pour les composantes G,R, puis pour le 2ème pixel,etc... Il est impossible, à l'oeil, de distinguer l'image qui cache le message, et l'image initiale.

Image initiale	R1=01001110 R2=01110011	G1=01101111 G2=01110110	B1=11111111 B2=10101010
Message	101100011011		
Image qui cache le message	R1=01001110 R2=01110001	G1=01101111 G2=01110110	B1=11111100 B2=10101011

Cette technique de stéganographie très basique s'applique tout particulièrement au format d'image BMP, format sans compression destructive, avec codage des pixels entrelacé sur 3 octets comme énoncé ci-dessus. Réciproquement, tout procédé de compression-décompression d'images avec pertes est susceptible de détruire un message stéganographique codé de cette façon. On parle alors de stérilisation. Un pays totalitaire pourrait stériliser à tout hasard toute image BMP entrant ou sortant de son territoire, moyennant les ressources techniques nécessaires.

11.2.2 Manipulation de la palette de couleurs d'une image

Certains formats graphiques tel que GIF ou PNG permettent le stockage des couleurs de l'image par référence à une palette de couleurs insérée dans le même fichier.

Ainsi au lieu de stocker Bleu,Blanc,Rouge dans une image du drapeau français, on trouve dans un format de fichier la description de l'objet la suite Couleur1,Couleur2,Couleur3 ainsi qu'une palette qui définit que Couleur1 est le Bleu, Couleur2 le Blanc et Couleur3 le Rouge.

La même image peut-être stockée de la façon suivante: Couleur2,Couleur3,Couleur1 avec une palette qui définit que Couleur2 est le Bleu, Couleur3 est le Blanc et Couleur1 est le Rouge.

Ces deux images sont visuellement identiques mais le stockage de celles-ci est différent. Pour une image contenant 256 couleurs uniques dans sa palette, on a factoriel de 256 façons de stocker cette image. En utilisant un code connu entre l'émetteur et le récepteur de l'image, on peut donc communiquer un message de petite taille caché dans la permutation des couleurs dans la palette de l'image.

11.2.3 Message caché dans les choix de compression d'une image

Tout semble indiquer que l'on ne peut cacher un message dans un format d'image utilisant une compression avec perte. En réalité la plupart des programmes de stéganographie sérieux s'attaquent justement au format JPEG qui utilise ce type de compression.

L'idée n'est pas de cacher une information dans les couleurs ou dans la palette (puisqu'il n'y en a pas) mais dans les choix de compression. En effet, tout algorithme de compression nécessite une succession de choix.

Avec des algorithmes de compression tel que Zip ou Gzip, on peut choisir la puissance de compression. En consommant plus de temps calcul et/ou plus de mémoire pour les opérations intermédiaires, on peut obtenir de meilleurs résultats de compression. Ainsi deux fichiers compressés de tailles différentes peuvent être décompressés en deux fichiers identiques.

La compression dans le format JPEG est double. La première compression consiste à découper l'image en bloc de 8 fois 8 pixel et de transformer ce carré sous une forme mathématique simplifiée. Cette compression introduit des pertes et la version mathématique peut être légèrement différente du carré original tout en étant visuellement très semblable. Une fois tous les blocs compressés, il faut coder les formes mathématiques en consommant le moins possible d'espace. Cette deuxième compression n'introduit pas de perte et est similaire dans les principes à ce que l'on peut retrouver dans Zip ou Gzip. C'est en introduisant dans cette phase des bits d'informations que l'on arrive à transporter un message caché.

11.2.4 Modulation fine d'un texte écrit

Décaler une lettre de quelques pixels ne pose aucun problème sur une imprimante à laser et est pratiquement invisible à l'œil nu. En jouant sur les interlettrages d'un texte très long et à raison de deux valeurs d'espacement correspondant à 1 et 0, il est possible de transmettre un message sous forme papier, qui ne révèlera son vrai sens qu'une fois analysé par un scanner ayant une bonne précision.

Historiquement, le procédé fut utilisé dès les années 70 en utilisant non pas des imprimantes laser, mais des imprimantes à marguerite Diablo, qui permettaient de jouer sur l'espacement des caractères au 1/120e de

pouce près.

11.2.5 Marquage de caractères

Une technique similaire — mais plus facilement détectable — consiste à marquer certains caractères d'un document. Des points peuvent par exemple être placés sous les lettres d'un texte afin de dissimuler un message. Étalées sur un texte de plusieurs pages, ces marques peuvent s'avérer relativement efficaces vis-à-vis d'un œil non-averti. Un ordinateur n'est pas indispensable à la mise en œuvre de cette technique.

En guise d'exemple, aviez-vous remarqué le message caché dans le chapitre 11.2.4 ?

11.2.6 Message transporté dans un son

Dans les formats sonores, il existe à peu près les mêmes possibilités de cacher des messages que dans les images.

Dans un fichier sonore au format MIDI, il n'existe pas de palette de couleurs mais bien différentes pistes qui peuvent être permutées.

Dans un fichier sonore avec compression sans perte, on peut cacher de l'information dans des variations imperceptibles du son, les bits faiblement significatifs.

Dans un fichier sonore avec compression avec perte, on peut cacher de l'information dans les choix de compression.

11.2.7 Autres possibilités

Il est aussi possible de cacher des informations dans bien d'autres types de fichiers couramment échangés sur des réseaux tel la vidéo ou bien dans des textes (ce fut une des premières formes de la stéganographie) ou encore dans des zones d'un disque dur inutilisées par le système de fichiers.

11.3 Usage

Après les évènements du 11 septembre 2001, on a prétendu qu'Oussama Ben Laden transmettait ses ordres en les cachant par des procédés stéganographiques dans des images transmises ou hébergées sur internet (ces suppositions n'ont jamais été étayées par des éléments concrets).

Il faut noter que, si la cryptographie, qui permet de protéger la vie privée et l'activité industrielle sans cacher cette protection, est souvent maltraitée par les états totalitaires et les sociétés démocratiques à tendance sécuritaire, il n'en va pas nécessairement de même pour la stéganographie, qui est pourtant une technique beaucoup mieux adaptée à une activité criminelle éventuelle.

12 Conclusion

12.1 Aujourd'hui

On sait faire des chiffres inviolables : Vernam

On sait faire des chiffres sans échanges de clé : RSA

Alors, c'est la fin de l'histoire ?

Même si le RSA à 512 bits est aujourd'hui sur. Cette certitude repose sur une infaillibilité du système. Quand on arrivera (si on arrive ?) à générer des nombres premiers facilement, le RSA ne sera pas plus sécurisé qu'un chiffre de Vigenère ...

Qui peut dire si aujourd'hui la NSA ne sait pas déjà résoudre l'équation n=pq ?

A de nombreuse reprises, des puissances se sont effondrées parce qu'elles pensaient avoir un système cryptographique sur. L'histoire est un éternel recommencement ...

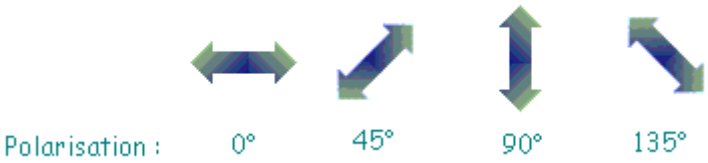
12.2 Demain : La cryptographie quantique

La prochaine génération de chiffres est déjà prête. Nous sommes aujourd'hui dans la situation de Diffie et Hellman en 1976 : Nous avons un outil parfait ... sur le papier : La cryptographie quantique.

L'utilisation de la mécanique quantique va permettre une nouvelle avancée : Cette fois, la sécurité est garantie non par des théorèmes mathématiques, mais par les lois fondamentales de la physique.

Dans le transport de « clé quantique », l'information est transportée par les photons. Chaque photon peut être polarisé, c'est-à-dire que l'on impose une direction à son champ électrique. La polarisation est mesurée par un angle qui varie de 0° à 180°.

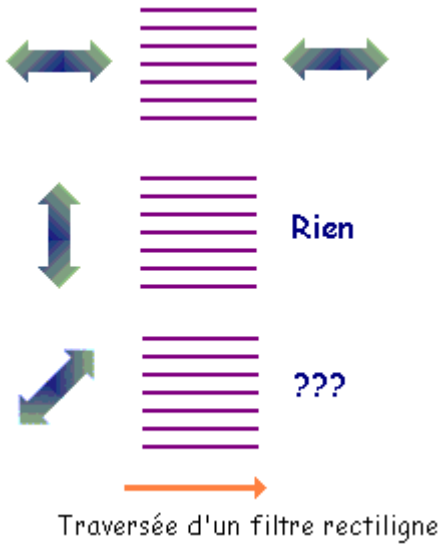
Dans le protocole que nous décrivons, dû aux canadiens CH.Bennett et G.Brassard, la polarisation peut prendre 4 valeurs : 0°, 45°, 90°, 135°.



Il nous faut pouvoir détecter la polarisation des photons. Pour cela, on utilise un filtre polarisant suivi d'un détecteur de photons. Si un photon polarisé à 0° rencontre un filtre polarisant orienté à 0°, il traverse ce filtre polarisant et est enregistré par le détecteur placé juste après. Si un photon polarisé à 90° rencontre le même filtre, il est immédiatement stoppé, et le détecteur n'enregistre rien.

Maintenant, si le photon est polarisé diagonalement (45° ou 135°), une fois sur deux, il traverse le filtre, et une fois sur deux, il est stoppé.

C'est dans cette incertitude de 50% que réside la force de la cryptographie quantique.



12.2.1 Alice & Bob dans le futur ...

Décrivons alors le protocole qu'Alice et Bob doivent respecter pour qu'Alice envoie à Bob une clé secrète constituée de 0 et de 1; ils disposent de 2 canaux d'échange : un canal quantique, où ils peuvent s'échanger des photons polarisés, et un canal non protégé (radio, téléphone, ...), où ils peuvent discuter.

Ils conviennent que les photons polarisés à 0° ou 45° représentent 0, et ceux polarisés à 90° ou 135° représentent 1.

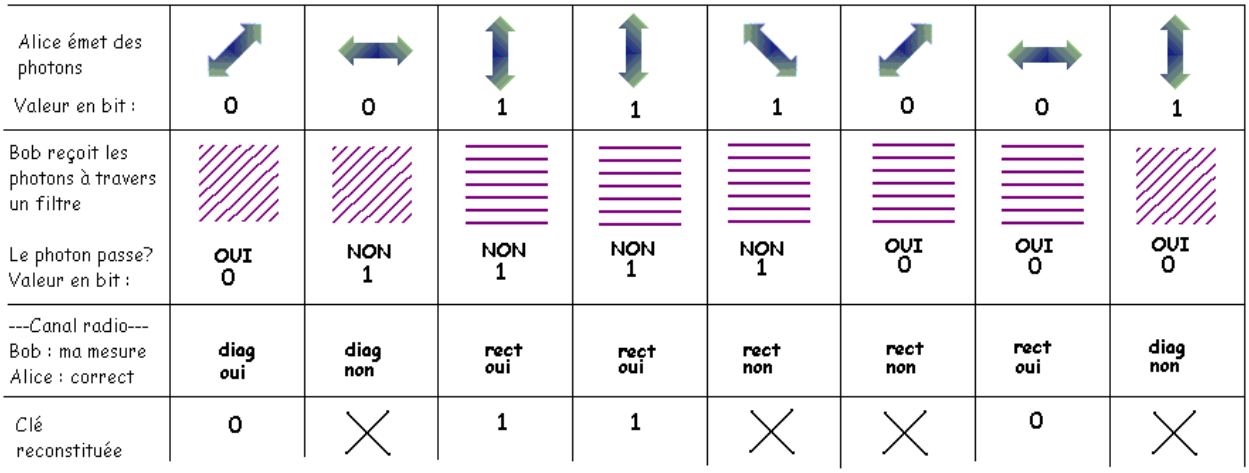
Alice émet, sur le canal quantique, une suite de photons polarisés au hasard parmi 0°, 45°, 90° et 135°. A l'autre bout, Bob reçoit les photons et mesure aléatoirement ou leur polarisation rectiligne (filtre placé à 0°), ou leur polarisation diagonale (filtre placé à 45°). Si le photon traverse le filtre, Bob note 0, sinon il note 1.

Bien sûr, certaines mesures de Bob (en moyenne, une sur deux) n'ont pas d'intérêt : il a pu essayer de mesurer la polarisation rectiligne d'un photon polarisé à 45°, ce qui n'a pas de sens et donne un résultat aléatoire (par exemple, le photon a été bloqué par le filtre, Bob note donc 1 alors qu'Alice avait envoyé 0).

Pour éliminer ces bits sans sens, il indique à Alice, par le canal radio, quelle type de mesure (rectiligne ou diagonale) il a faite pour chaque photon. Par le même canal radio, Alice lui indique quelles sont les mesures correctes (photon polarisé à 0° ou 90° avec filtre rectiligne, photon à 45° ou 135° avec filtre diagonal), dans l'exemple ci-dessous la 1, la 3, la 4, et la 7. Les bits 1,3,4,7 sont désormais connus à la fois de Bob et d'Alice, et constituent leur clé secrète commune.

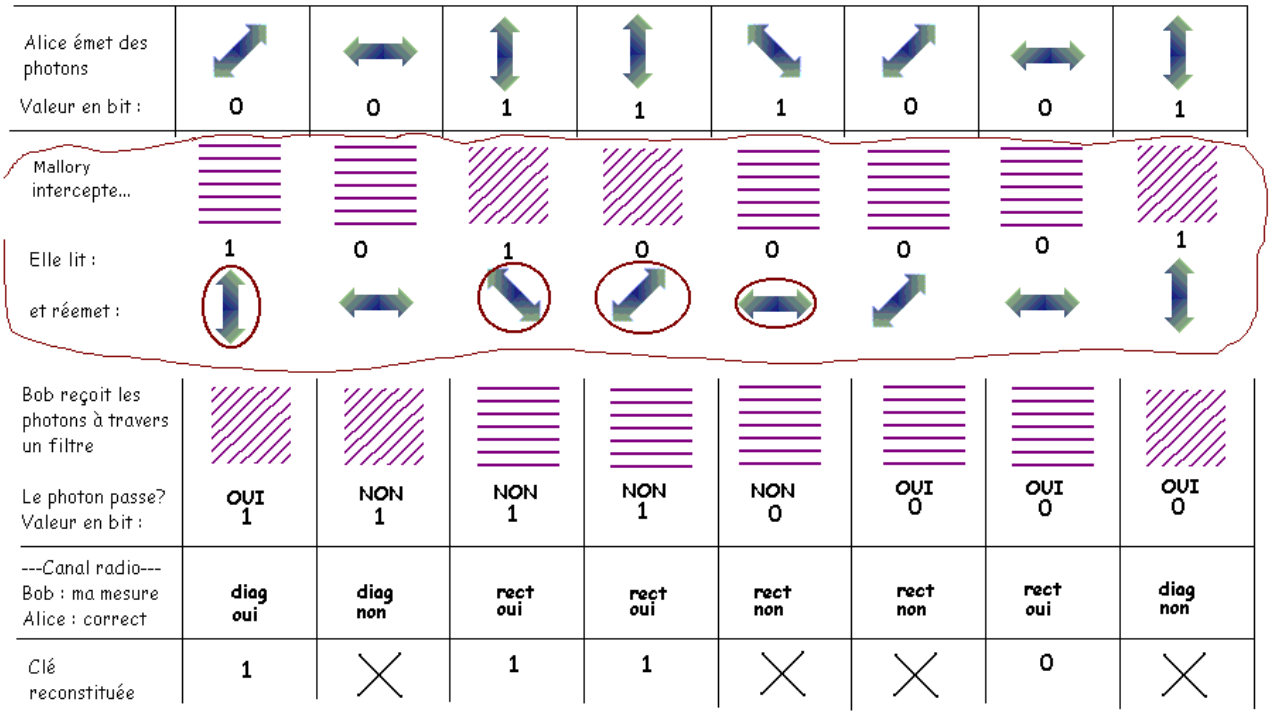
Il faut encore vérifier que ce protocole est sûr.

Si Mallory écoute le canal quantique, elle peut faire la même chose que Bob, c.a.d intercepter les photons en plaçant un filtre polarisant tantôt rectiligne, tantôt diagonal. Pour que Bob ne se doute de rien, elle doit réémettre un photon polarisé. (Car le fait de « lire » un photon le détruit.)



Elle va essayer d'envoyer le même photon qu'Alice, mais comme elle a une chance sur deux d'avoir choisi le mauvais filtre, elle a une chance sur deux de se tromper. Quand Bob reçoit le photon, s'il est mal polarisé par Mallory, il a une chance sur deux d'avoir un résultat différent d'avec le photon original, et finalement, pour chaque photon intercepté par Mallory, il y a une chance sur 4 que Bob reçoive une information erronée.

Alice et Bob décident alors de "sacrifier" une partie de leur clé commune. Parmi tous les bits qu'ils ont en commun, ils en choisissent quelques-uns au hasard, et les compare publiquement par le canal radio : s'ils sont différents, ils ont une preuve qu'ils ont été écoutés, et ils oublient vite cette clé. En comparant suffisamment de bits, ils ont une garantie presque absolue de ne pas avoir écouté.



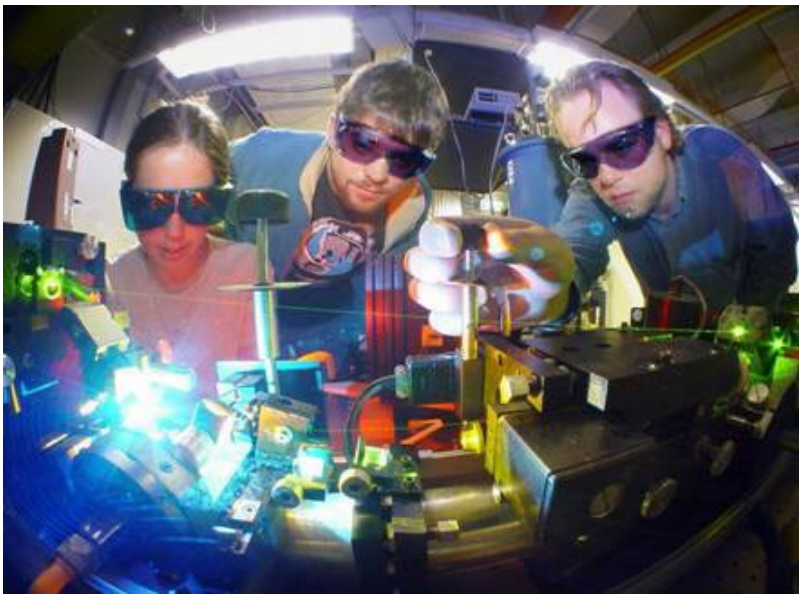
Puis... Bob : j'ai peur que nous ayons été espionné, sacrifions le premier bit de notre clé - j'obtiens 1

Alice : je t'avais envoyé 0, nous avons été espionnés...

Remarquons pour terminer que même non repérée, Mallory n'avait pas la bonne clé, puisque le troisième bit de la clé qu'elle obtient est 0 alors qu'Alice avait envoyé 1!

Bien sûr, il reste des problèmes pratiques à résoudre : émettre des photons un par un, conserver la polarisation sur de grandes distances... mais les physiciens sont au travail !

12.3 Demain aussi : La cryptanalyse quantique



Chercheurs australiens développant la nouvelle génération d'ordinateur quantique

En 1982, le prix Nobel de physique Richard Feynman imagina un modèle théorique illustrant comment un système quantique pourrait être utilisé pour faire des calculs.

Rapidement, en 1985, David Deutsch, montra que les idées de Feynman pouvaient mener à un ordinateur quantique, un ordinateur qui effectuerait n'importe quelle tâche, mais serait capable de tirer avantage des propriétés quantiques de la matière, principalement du principe de superposition des états.

Cette recherche s'avéra plus difficile que prévu, et il fallut attendre le milieu des années 90 pour qu'un chercheur des laboratoires Bell, Peter Shor, invente des opérations mathématiques élémentaires propres aux ordinateurs quantiques et les applique pour créer un algorithme quantique de factorisation.

12.3.1 L'ordinateur quantique

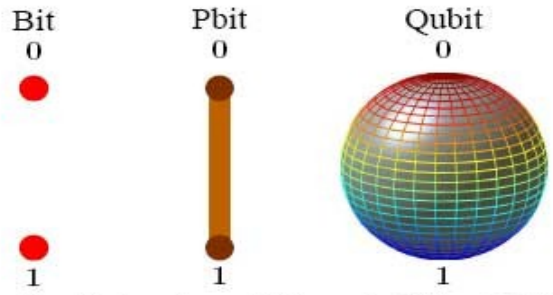
Nous savons qu'un ordinateur classique traite des informations élémentaires, des bits, qui ne peuvent présenter qu'un état parmi deux possibles : 0 ou 1. C'est le langage binaire.

La révolution que propose l'informatique quantique est de remplacer ces bits par des bits quantiques, ou q(u)bits en abrégé, pouvant prendre un ensemble de valeurs beaucoup plus large. En effet, la physique quantique, avec son principe de superposition, permet à un état d'être un "mélange" d'autres états. Ainsi, un qbit peut prendre les valeurs 0 ou 1, mais aussi un état constitue de 10% de 0 et 90% de 1, ou toute autre combinaison.

Ceci signifie que quand on mesure la valeur du qbit, on a 10% de chances de trouver 0 et 90% de trouver 1.

Un peu plus concrètement, avec 4 bits, un ordinateur classique peut traiter un état parmi 2^4 soit 16 états différents : 0000, 0001, 0010, 0011, etc.

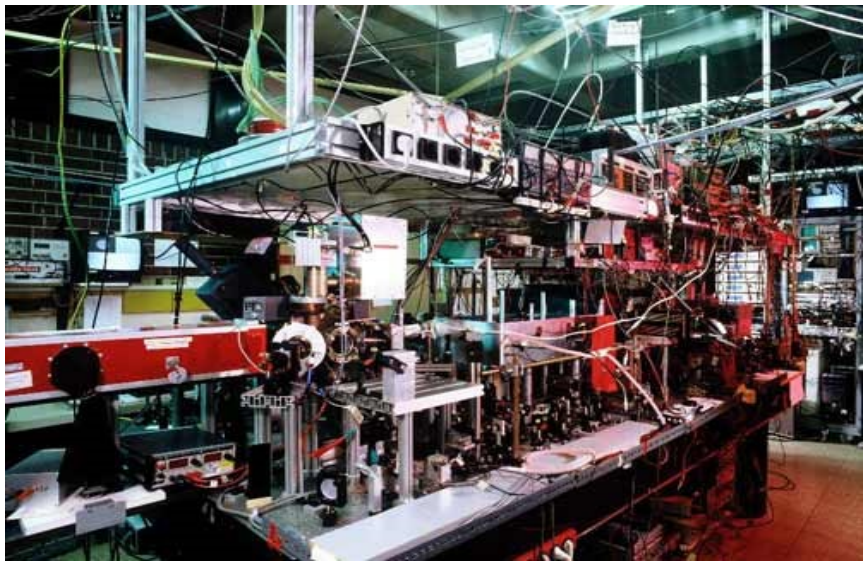
Dans un ordinateur quantique, les quatre qbits pourraient être dans une superposition de tous ces états. Dans cette situation, l'avantage de l'ordinateur quantique est de pouvoir traiter simultanément les 16 états.



A gauche un bit ordinaire est caractérisé par deux états, 0 ou 1. Au centre un pbit ou "bit probabiliste". Il représente la distribution des probabilités d'un bit. L'expression indiquée signifie que le pbit a une probabilité p d'être dans l'état 0 et 1-p d'être dans l'état 1. C'est l'exemple typique de la pièce de monnaie que l'on jete en l'air : elle a 1 chance sur 2 de tomber sur pile, 1 chance sur 2 de tomber sur face. A droite, le qubit opère dans un univers multidimensionnel, ses états propres correspondant à la surface d'une sphère dite de Bloch tandis que ses états logiques correspondent aux pôles de cette sphère.

Des ordinateurs quantiques équipés de processeurs de N qubits permettent donc de gérer 2^N informations différentes simultanément ! Ils calculent donc N fois plus vite qu'un ordinateur classique puisqu'ils sont capables

d'effectuer ces calculs en parallèle ! Le nombre de qubits augmente donc de manière exponentielle la puissance du travail en parallèle. Il est ainsi facile de calculer qu'un ordinateur quantique de 300 qbits pourraient gérer environ 10^{90} informations, soit plus que le nombre d'atomes dans l'Univers observable !



Mais les ordinateurs quantiques n'en sont encore qu'à leurs prémices, et leur record (automne 2001) est la factorisation de 15=3×5 avec 7 qbits.

Pas de nouvelles depuis.

Or, on à vu que casser le RSA revenait à factoriser n...

13 Exercices de cryptanalyse

13.1 César

Substitution polyalphabétique

Source : "Bien que les champs, les fleuves et les lieux" de Pierre de Ronsard

Difficultés : 1/5

Cryptogramme :

```
JQMVY CMTMA KPIUX ATMAN TMCDM AMBTM ATQMC FTMAU WVBAT MAJWQ AYCMR IQTIQ AAMAL MZZQM
ZMUMB QMVVM VBTWQ VLMUI LWCKM OCMZZ QMZMI ABZMN IBITL WCAMK WCTMU WVUQM CFYCM TYCML
MUWVX IZTMK WVOML MAKQM CFYCQ XZMAQ LIQBI UWVIZ LMCZX ZMUQM ZMKWV LCQBB WCRWC ZALCV
MIQTM KWCBC UQMZM AIJMT TMQUI OMICA MRWCZ LMUMA GMCFB WCBMA TMAVC QBAQU XIBQM VBLMP
IBMMV BZMUM AJZIA RMZMU JZIAA MMBZM BIBMA WVDIQ VXWZB ZIQBM VKMVB NWZUM ABZWU XMCZU
IQAYC IVLQT DWQBY CMKWV BMVBR MAWUU MQTTM UWYCI VBUMA JZIAQ TAMVN CQBMB UMDMQ TTMAM
CTMVU WVTQB XTMQV LMPWV BMMBL MXMCZ
```

13.2 Vigenère

Substitution polyalphabétique

Source : Un extrait de la "La bête humaine".

Difficultés : 3/5

Cryptogramme :

```
OSYDZ BEUMW YSSOY TQCFB ZIOSH BEEQS DSZQZ MOOCA MHDOM UPAMH LSZUL RDRFH ZJCEB VLQTS
QOHGB UHZGE MOTTC SWARC CTDDL SRSCU MXZUQ OAAQW DESEF IKRPS BSYDZ WECGS KLZAP RDJTC
SCTRD ZPMDH TEQZP RDDCE MOTTR CFSKO XAQEF IRSOE RUCAM RPSKW RNDGW AQFTV DSOUM HCAHB
OELOY TDGLV ZWEAM WXEKS DQTOT SDHTL RITVH HOERM PUWZL MZQSI MSOEL OYODI GRDIY EOSEI
SSXAB VTNDH PNCSC ATLER NWDRN IPSAO DSDGP TBCFP KSPSP ITCNA XEMQL ISZPD DPCAM QSELS
YTCIE RZWYA KSCTD PPSNU YETGP ELAPN ZBERD TZUKO YTKSD WZUZN RGFRK SDVNW PSCSC ELWDA
FSFND OFTQS XABVT NDDFI RGLNS SNEKZ PLZIY ELONH HBPDD LARDG DATLO ETLRR ZBOER FZUDG
OEUCC AMHPS RHLTH CYNZW ESDIW EKONH ZWEPZ FDABV PMHBP ETBPG QCDSD TFMDS YOHFP MNBEA
MHORN WETQS DLDBE ECOYS KOTRB OWMDA LIRHZ USSDO MOETD BEINB QUSDC IRSAA QZPTQ OTNCS
ERNWD HDICE RJTNF HNIME LDDGE IMOEI NBOEB OPNDA ALHRP JZRPS DGGOX ORETF DESEF IZHEE
MRLIS GLMZQ SIMST LMOAE QQPVZ WEPZG NEKZP CHOCR DHPEZ IOEKO OUOCY TCSWE TFZPD WWLDB
EEMRL ISGPU KSXEM HOELO YDDFW AUCTE ZZPGD FDCNI ASCSD IETWE SDCER GPSDB AEQGZ NMSBU
DZTMO OEIDB NEFOR ND
```

13.3 RSA

Difficulté : 4/5

RSA avec clé à 11 bits

Clé publique : RSA(1370477, 377)

Cryptogramme :

```
1278281 1180729 1168560 1105842 1168560 0827986 0231383 0042925 0167338 0002372
1349403 0726411 0902962 0277609 0363080 0794055 0565269 0129271 0060967 0827986
1274266 0827986 0420867 0633528 1274266 0709512 0794799 0792163
```

14 Anecdotes en vrac

14.1 Kama-sutra

Le Kama-sutra est un texte écrit au 5e siècle par le brahmane Vatsayayana, mais fondé sur des manuscrits du 4e siècle avant J.-C.

Le Kama-sutra recommande que les femmes apprennent 64 arts, entre autres cuisiner, s'habiller, masser et élaborer des parfums. La liste comprend aussi des domaines moins évidents, comme la prestidigitation, les échecs, la reliure et la tapisserie.

Le numéro 45 de la liste est le mlecchita-vikalpa, ou l'art de l'écriture secrète, qui doit leur permettre de dissimuler leurs liaisons.

14.2 Georges Sand et Alfred de Musset

Voici la lettre envoyée par Georges Sand à Alfred de Musset :

Cher ami,

Je suis toute émue de vous dire que j'ai bien compris l'autre jour que vous aviez toujours une envie folle de me faire danser. Je garde le souvenir de votre baiser et je voudrais bien que ce soit une preuve que je puisse être aimée par vous. Je suis prête à montrer mon affection toute désintéressée et sans calcul, et si vous voulez me voir ainsi vous dévoiler, sans artifice, mon âme toute nue, daignez me faire visite, nous causerons et en amis franchement je vous prouverai que je suis la femme sincère, capable de vous offrir l'affection la plus profonde, comme la plus étroite amitié, en un mot : la meilleure épouse dont vous puissiez rêver. Puisque votre âme est libre, pensez que l'abandon ou je vis est bien long, bien dur et souvent bien insupportable. Mon chagrin est trop gros. Accourrez bien vite et venez me le faire oublier. A vous je veux me soumettre entièrement.

Votre poupée

Georges Sand, dessinée par Alfred de Musset.

Classe non ? Et pourtant, si vous saviez lire entre les lignes ...

La réponse d'Alfred de Musset n'est pas mal non plus :

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 coeur
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.

Vous ne lirez plus jamais les auteurs classiques de la même façon...

14.3 Le télégramme de Zimmermann

Nous sommes en janvier 1917. La guerre entre l'Allemagne et les Alliés fait rage. Le conflit s'enlise, les tranchées ennemies se font face. De l'autre côté de l'Atlantique, les États-Unis sont prudemment restés neutres. En 1916, le président Woodrow Wilson a d'ailleurs été réélu avec le slogan *« He kept us out of the war »* (il nous a préservé de la guerre). Il a même déclaré que ce serait *« un crime contre la civilisation »* de laisser entraîner les États-Unis dans la guerre.

En janvier 1917, l'état-major allemand s'impatiente. Il propose au Kaiser de déclencher une guerre sous-marine totale, afin de couper les approvisionnements de l'Angleterre. Le problème pour l'Allemagne est qu'en coulant de nombreux bateaux civils américains, cette stratégie aurait pour probable conséquence l'entrée en guerre des États-Unis. Arthur Zimmermann, alors ministre allemand des affaires étrangères, a une idée afin de retarder l'envoi de renforts américains : il faut occuper les troupes américaines sur d'autres fronts, créés par le Mexique et le Japon. Ainsi, si le Mexique envahit le sud des États-Unis (les 2 pays ont alors un contentieux autour de certains territoires), avec le soutien logistique, financier et militaire allemand, l'armée américaine sera toute accaparée par la défense de son propre territoire, et ne pourra intervenir en Europe.

Le 16 janvier 1917, Zimmermann envoie sa proposition dans un télégramme codé à destination de Bernstoff, ambassadeur d'Allemagne aux États-Unis (Berlin n'a pas de liaisons directes avec le Mexique). A charge pour lui de le faire parvenir à son homologue mexicain, qui lui-même le transférera au Président mexicain. Bernstoff déchiffre le message, et l'envoie au représentant allemand au Mexique, Johann Eckardt, avec un code commun entre lui et Eckardt. Ce dernier télégramme est intercepté par les services secrets britanniques, et immédiatement confié au prestigieux Bureau 40 qui se consacre au chiffrement. L'équipe, dirigée par le révérend Montgomery, parvient après un mois de labeur le 22 février 1917 à décrypter ce message. Et voici ce qu'ils ont pu lire :

> *Nous avons l'intention de déclencher à partir du 1er février une guerre sous-marine totale. Malgré cela, nous tenterons de maintenir les États-Unis dans la neutralité. Si nous n'y parvenons pas, nous proposerons au Mexique une alliance sur les bases suivantes : faire la guerre ensemble, faire la paix ensemble, large soutien financier et accort de notre part pour la reconquête par le Mexique des territoires perdus du Texas, du Nouveau Mexique, et de l'Arizona. Le règlement des détails est laissé à vos soins. Vous informerez secrètement le Président du Mexique dès que l'entrée en guerre des États-Unis sera certaine, et vous lui suggérerez que, sous sa propre initiative, il peut immédiatement solliciter la participation du Japon, et en même temps servir de médiateur entre le Japon et nous-même. Prière d'attirer l'attention du Président sur le fait que l'emploi sans limites de nos sous-marins offre désormais la possibilité d'obliger l'Angleterre à faire la paix dans peu de mois.*

Les Britanniques envoient immédiatement le résultat de leurs travaux aux États-Unis, et le 1er mars, toute la presse nord-américaine répand l'information. L'opinion publique, désormais prévenue des intentions allemandes, change d'avis, et le Président Wilson fait voter par le Congrès, le 6 avril 1917, une déclaration officielle de guerre à l'Allemagne. C'est sans doute la première fois que la cryptanalyse d'un message codé a autant changé le cours de l'histoire.

14.4 Le chiffre ADFGVX

14.4.1 Le contexte

1918. Les armées allemandes et françaises sont exsangues. L'état-major français, dirigé par le maréchal Foch, redoute l'imminence d'une offensive massive de l'ennemi, qui enfoncerait les lignes de défense jusque Paris. Cinq points d'offensive sont possibles, mais les forces françaises de réserve ne permettent de se concentrer que sur un seul. Il est vital, pour l'issue de la guerre, de ne pas se tromper.

14.4.2 Le code

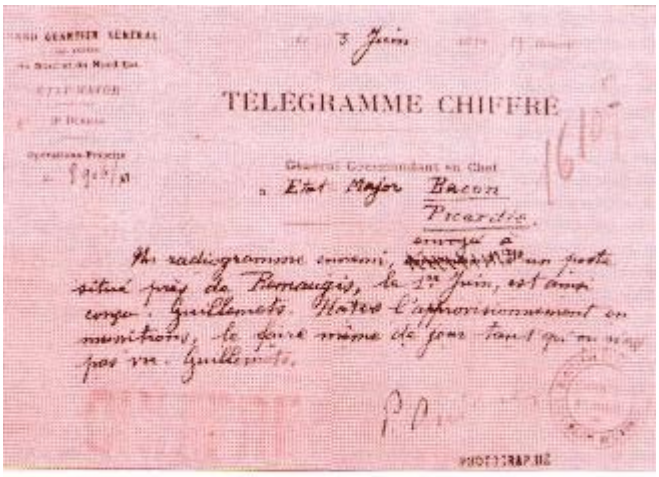
Depuis mars 1918, l'armée allemande utilise un nouveau code pour communiquer, le chiffre ADFGVX, ou GEDEFU 18 (GEheimschrift DEr FUnker 18, chiffre des télégraphistes 18). Ce chiffre est constitué d'une substitution de type carré de Polybe, suivie d'une transposition. Pour réaliser la substitution, les 26 lettres de l'alphabet et les 10 chiffres sont rangés dans un tableau 6×6, aux extrémités desquelles on a ajouté les lettres ADFGVX.

	A	D	F	G	V	X
A	Q	Y	A	L	S	E
D	Z	C	R	X	H	0
F	F	O	4	M	8	7
G	3	I	T	G	U	K
V	P	D	6	2	N	V
X	1	5	J	9	W	B

Chaque lettre est codé par le couple de lettres qui correspond à sa ligne et à sa colonne. Ainsi, R est codé DF, et le message RENFORT COMPIEGNE 16H10 devient :

DFAXV VFAFD DFGFD DFDFG VAGDA XGGVV AXXAV FDVXA DX

On choisit ensuite, pour faire la transposition, une clé qui est un mot courant, par exemple DEMAIN. On écrit le texte intermédiaire sous ce mot, puis on réordonne les colonnes par ordre alphabétique croissant :


TÉLÉGRAMME CHIFFRE

Message allemand d'origine : Munitionierung beschleunigen punkt soweit nicht gesehen auch bei Tag.

Traduction : Hâter l'approvisionnement en munitions, le faire même de jour tant qu'on n'est pas vu.

Il ne reste plus qu'à relire le tableau de gauche à droite, et de haut en bas :

XDFVA VDFAD FFDGF FDGDV AAGXV GGAVX FXADV VXXAD

D	E	M	A	I	N
D	F	A	X	V	V
F	A	F	D	D	F
G	F	D	D	F	G
V	A	G	D	A	X
G	G	V	V	A	X
X	A	V	F	D	V
X	A	D	X		

A	D	E	I	M	N
X	D	F	V	A	V
D	F	A	D	F	F
D	G	F	F	D	G
D	V	A	A	G	X
V	G	G	A	V	X
F	X	A	D	V	V
X	X	A		D	

Les lettres ADFGVX ont été choisies pour ce code car l'essentiel des télécommunications est transmis par radio, et les lettres ADFGVX ont des codes morses très différents. Les utiliser évite les confusions pendant la transmission.

14.4.3 Le travail des Français

L'équipe française de cryptographie est particulièrement talentueuse au cours de la Première Guerre Mondiale. Dès le début des hostilités, la majorité des messages chiffrés allemands sont compris, mais malheureusement ces exploits sont gâchés par des indiscrétions dans la presse. Avec l'apparition du chiffre ADFGVX, la situation se complique sérieusement...

Le cryptologue le plus doué de la Section du chiffre est le paléontologue Georges Painvin, ancien major de l'école Polytechnique. Après un travail acharné, il parvient le 2 juin 1918 à déchiffrer les fameux messages allemands, dont un radiogramme à destination d'une unité située au nord de Compiègne (voir ci-contre). Le maréchal Foch est immédiatement averti des intentions de l'ennemi, et fait masser les troupes au lieu adéquat. L'assaut allemand a lieu le 9 juin 1918. Il est stoppé net. La dynamique de la victoire s'enclenche. Pas étonnant que ce message chiffré ait reçu le nom de Radiogramme de la Victoire.

Alliés 2 – Allemagne 0

14.5 Olivier Levasseur

Olivier Levasseur plus connu sous le nom de "La Buse", surnommé ainsi en raison de sa rapidité à fondre sur sa proie est un authentique pirate.

La Buse, écuma l'océan Indien au début du 18ème siècle. Il aurait caché un trésor estimé à 4,5 milliards d'euros quelque part à La Réunion. Aujourd'hui encore, des chercheurs et des scientifiques se lancent à la recherche de ce trésor précieusement conservé depuis plus de 280 ans.

14.5.1 Son histoire :

Olivier Levasseur est né à Calais à la fin du XVIIè siècle. En 1721, La Buse est associé au pirate anglais Taylor. Ils se sont emparé au mois d'avril du riche vaisseau portugais de 72 canon La Vierge du Cap qui avait cherché

refuge contre les tempêtes dans le port de Saint-Denis (île Bourbon).

A bord du vaisseau se trouvaient le comte Ericeira, vice-roi des Indes et l'archevêque de Goa. La Buse fit main basse sur les objets d'inestimable valeur : rivières de diamants, bijoux, perles, barres d'or et d'argent, meubles, tissus, vases sacrés et cassettes de pierres précieuses, et la crosse d'or de Goa constellée de rubis pesant une centaine de kilos, le tout évalué à **4,5 milliards d'euros**.

La Vierge du Cap, radoubée et remise à neuf, devint le vaisseau de La Buse et prit le nom de Le Victorieux.

Mais l'année d'après, Duguay-Trouin et le commodore anglais Matthews vinrent se chercher querelle dans les parages. La Buse et Taylor se sont méfiés et ont préféré prendre "le large". Taylor s'enfuit aux Antilles et La Buse se retira à l'île Sainte-Marie près de la côte de Madagascar.

Il est certain qu'il a caché son trésor...mais où ?

On a avancé le nom de 6 îles : Maurice, La Réunion, Frigate, Mahé, Rodrigues, Sainte-Marie.

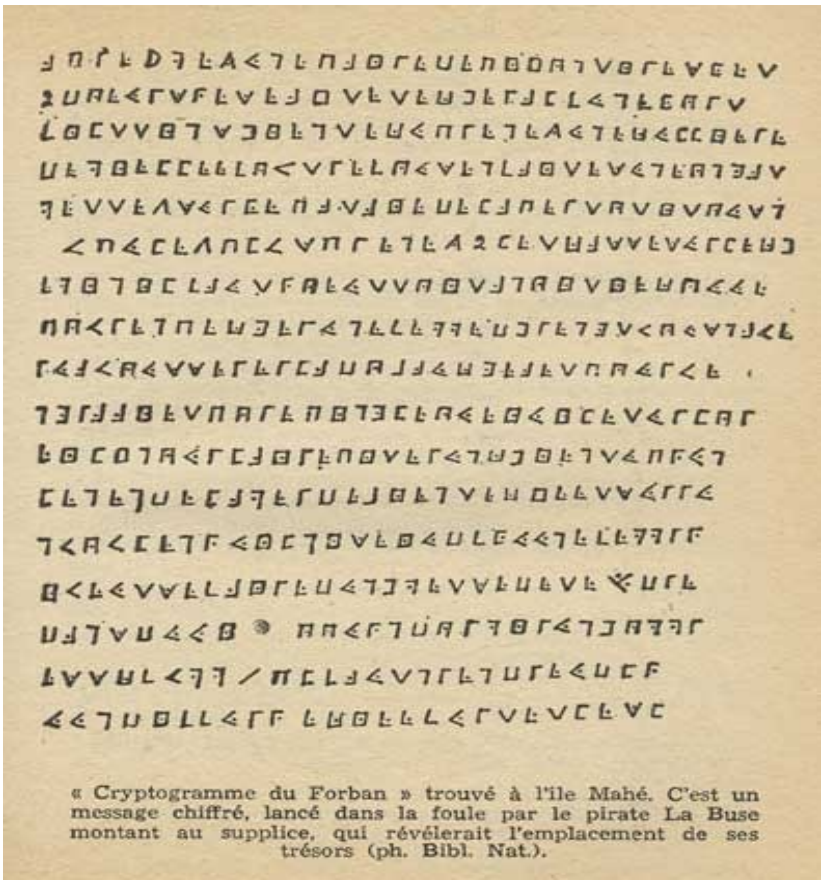
Vers 1729, exerçant le métier de pilote dans la baie d'Antongil (Madagascar), il offrit des services au vaisseau La Méduse, de la Compagnie des Indes, qui voulait entrer dans le port.

Le Capitaine d'Hermitte, commandant de bord, le reconnut, et se souvenant que le pirate avait maintes fois arraisonné des navires de sa compagnie, il l'arrêta.

Le 7 juillet 1730, La Buse était condamné à mort.

Quand il monta sur l'échafaud pour expier ses crimes de pirate, Olivier Levasseur, dit La Buse, lança dans la foule un cryptogramme et s'écria :

"Mes trésors à qui saura comprendre !"



« Cryptogramme du Forban » trouvé à l'île Mahé. C'est un message chiffré, lancé dans la foule par le pirate La Buse montant au supplice, qui révélerait l'emplacement de ses trésors (ph. Bibl. Nat.).

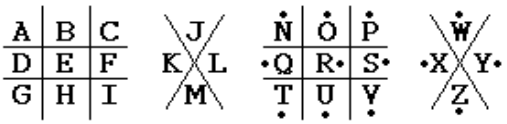
Qu'est devenu le trésor?

Nul ne saurait le dire, mais depuis plus de deux siècles, l'océan Indien, des îles Seychelles à la pointe de Madagascar, est le centre de recherches incessantes et foisonne de documents à clés, de rébus et de signes gravés qui tous, selon la tradition, se rapportent aux prodigieux trésors de La Buse.

14.5.2 Cryptanalyse

La Buse utilise le chiffre dit de Pig Pen (parc à cochons) qui est un simple chiffre de substitution monoalphabétique

Soit



14.5.3 Exercice

Trouvez le trésor et rapportez le au prof.