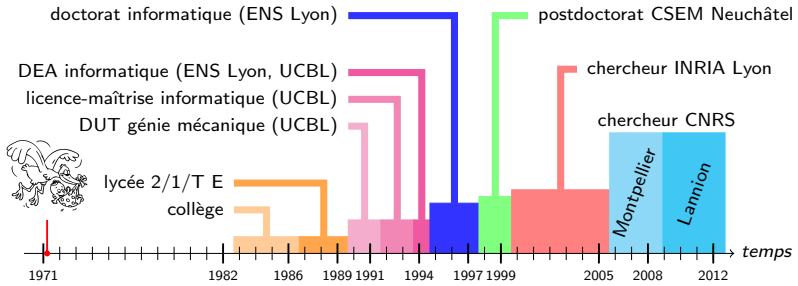
Activités autour de la cryptologie

Arnaud Tisserand

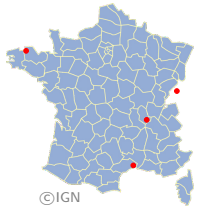
CNRS, IRISA, équipe-projet CAIRN

Réunion avec les Professeurs de Mathématiques
Lanmeur, février 2012

Présentation personnelle



- Université Claude Bernard de Lyon
- École normale supérieure de Lyon
- Centre suisse d'électronique et microtechnique
- Institut national de recherche en informatique et automatique
- Centre national de la recherche scientifique



Centre national de la recherche scientifique

`www.cnrs.fr`

créé en 1939

- organisme public de recherche (EPST)
- objectifs : *produire du savoir et mettre ce savoir au service de la société* (source : site web)
- personnels :
  - 11 450 chercheurs
  - 14 180 ingénieurs, techniciens et administratifs
  - 8 900 agents non permanents (p. ex. doctorants)
  - 1 000+ laboratoires en France
  - coopération scientifique avec 60 pays
  - environ 600 entreprises créées depuis 1999

Domaines de recherche : biologie, chimie, écologie et environnement, sciences humaines et sociales, sciences de l'information et de leurs interactions, sciences de l'ingénierie et des systèmes, sciences mathématiques et de leurs interactions, physique, physique nucléaire et physique des particules, sciences de l'univers

Remarque : EPST = établissement public à caractère scientifique et technologique

Laboratoire IRISA

Institut de recherche en informatique et systèmes aléatoires

`www.irisa.fr`

créé en 1975

- unité mixte de recherche (UMR 6074) entre :
  - CNRS
  - Université Rennes 1
  - INSA de Rennes
  - ENS Cachan antenne de Bretagne
- associé à l'INRIA Rennes Bretagne Atlantique
- environ 600 personnes
- environ 30 équipes de recherche
- sur les campus de Rennes et Lannion
- recherches dans le domaine des STIC
- formation par la recherche (masters, doctorat...)

Remarque : STIC = sciences et techniques de l'information et de communication

ENSSAT Lannion

École nationale supérieure des sciences appliquées et de technologie

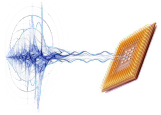
ENSSAT LANNION

`www.enssat.fr`

créée en 1986

- rentrée 2011-2012 : 120 nouveaux élèves
- 1788 ingénieurs diplômés depuis 1986
- spécialités :
  - électronique et informatique industrielle
  - logiciel et système informatique
  - optronique
  - informatique, multimédia et réseaux (apprentissage)
- pôle de recherche en télécoms et technologies émergentes
- 180 personnes (hors élèves)
- devenir des ingénieurs (enquête mai 2011) :
  - 100 % des diplômés 2010 en activité
  - 86 % ingénieurs
  - 14 % poursuite d'études

Équipe de recherche CAIRN



`www.irisa.fr/cairn`

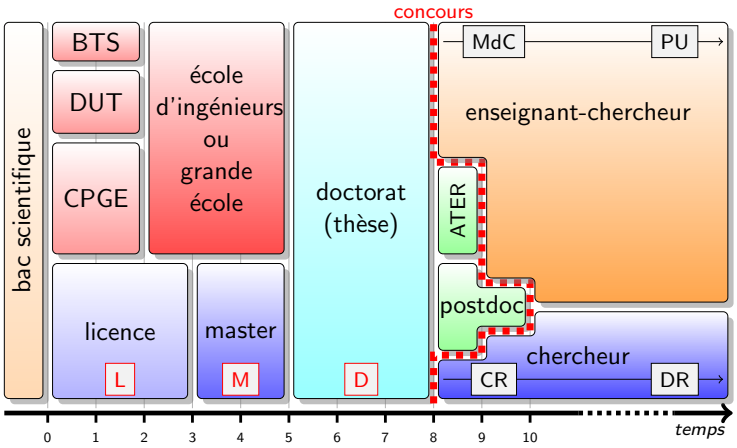
- personnels :
  - 19 permanents (Univ. Rennes 1, CNRS, INRIA)
  - 30 doctorants
  - 4 ingénieurs et 4 postdocs/ATER
  - 3 assistantes
- collaborations : 15+ pays et 15+ labos français
- 20+ projets nationaux/européens/internationaux

Thématiques de recherche théoriques et appliquées :

- puces électroniques (reconfigurables)
- outils informatiques pour programmer ces puces
- basse consommation d'énergie
- tolérance aux pannes
- sécurité contre les attaques
- applications : télécommunications, multimédia, sécurité...

Comment devenir (enseignant-)chercheur ?

Échelles : années en abscisse, aucune en ordonnée



BTS : brevet de technicien supérieur
DUT : diplôme universitaire de technologie
CPGE : classe préparatoire aux grandes écoles
ATER : attaché temporaire d'enseignement et de recherche

MdC : maître de conférences
PU : professeur des universités
CR : chargé de recherche
DR : directeur de recherche

Protection des données et des communications

Applications :


internet


signature électronique


carte bleue


vote électronique

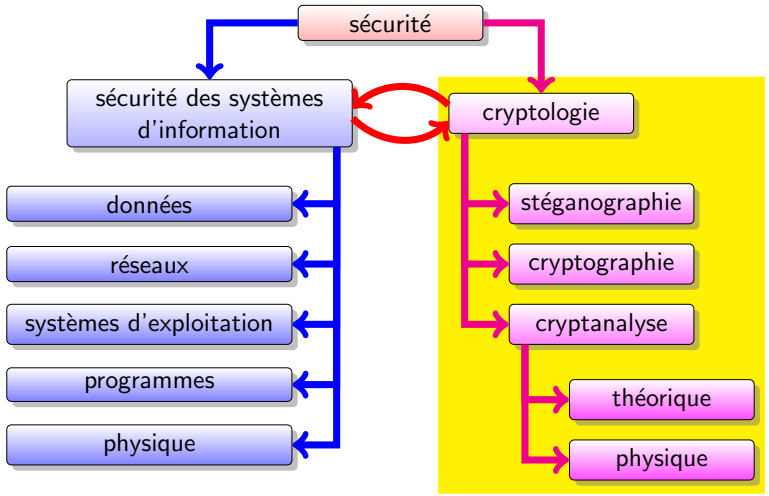

décodeur

...

Besoins :

- Confidentialité ⇝ les informations sont non compréhensibles
- Authentification ⇝ garantie de l'identité de l'émetteur
- Intégrité ⇝ garantie que les informations sont complètes
- Non-répudiation ⇝ l'émetteur ne peut pas se dédire

Quelques facettes de la sécurité



Terminologie (1/3)

Cryptographie : étymologiquement « écriture secrète », méthodes et outils pour rendre les messages incompréhensibles sans avoir la clé

⇝ art du secret

Cryptanalyse : analyse des cryptogrammes pour retrouver les informations secrètes (clé ou message en clair)

Cryptologie : cryptographie + cryptanalyse

Stéganographie : cacher un message dans un autre message

⇝ art de la dissimulation

Terminologie (2/3)

Vulnérabilité : faiblesse/faille dans un système

Attaque : méthode pour exploiter une vulnérabilité

Menace :

- utilisateur étourdi/insouciant
- sinistre : vol, inondation, incendie...
- programme malveillant (virus, vers, trojan...)
- personnes/groupes malveillants
  - individu seul (p. ex. *hacker*)
  - société concurrente
  - gouvernement étranger

puissance d'attaque ↓

Protection ou contre-mesure : méthode ou dispositif qui contre une attaque ou qui supprime/limite une vulnérabilité

Terminologie (3/3)

Confidentialité : garantit que les données sont compréhensibles seulement par les personnes autorisées

⇝ les données semblent totalement aléatoires à toute autre personne

Authentification : garantit que l'origine des données est la bonne

⇝ authentification d'une source ou d'un tiers (identité)

Intégrité : garantit que les données ne sont pas altérées

⇝ les données sont complètes et sans aucune modification

Non répudiation : garantit qu'une action ne peut être niée

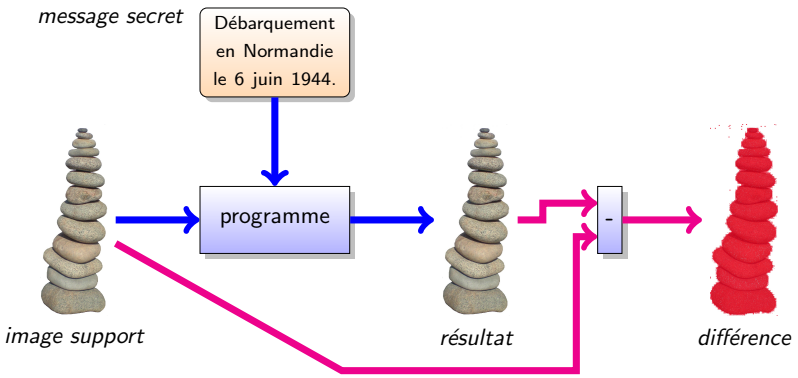
⇝ impossibilité de se dédire

Autres caractéristiques souvent utilisées : disponibilité, identification, contrôle d'accès...

Stéganographie

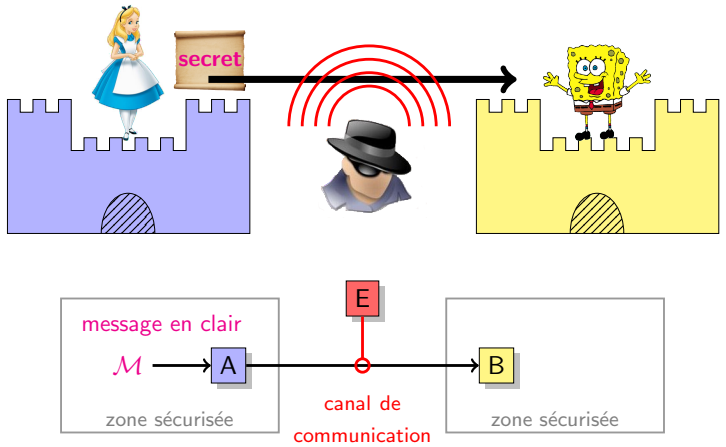
Cryptographie : art du secret
Stéganographie : art de la dissimulation

Principe : cacher un message (secret) dans un autre message (support)



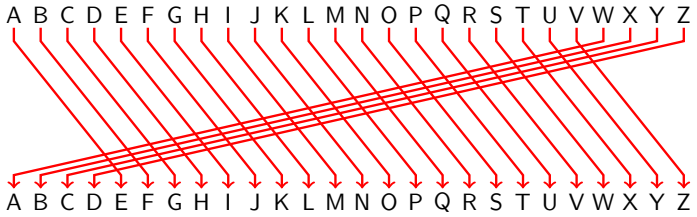
Cadre cryptographique : transmission d'un message secret

Alice souhaite envoyer un message secret à Bob sans que l'espion puisse avoir des informations sur ce message



Méthode de chiffrement très simple (mais peu fiable)

Idée : remplacer les lettres de l'alphabet par une version décalée de l'alphabet (chiffre de César)



La clé secrète est la valeur du décalage

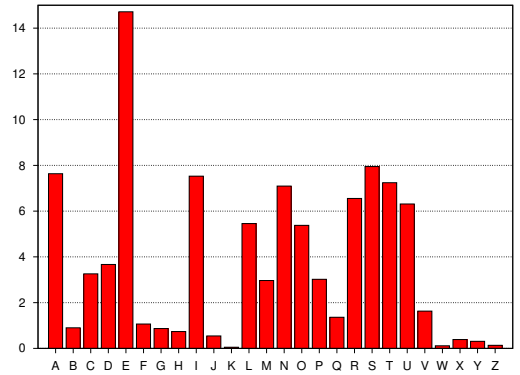
Exemple :

F	E	T	E		D	E		L	A		S	C	I	E	N	C	E
↓	↓	↓	↓		↓	↓		↓	↓		↓	↓	↓	↓	↓	↓	↓
J	I	X	I		H	I		P	E		W	G	M	I	R	G	I

Fréquence d'apparition des lettres en français

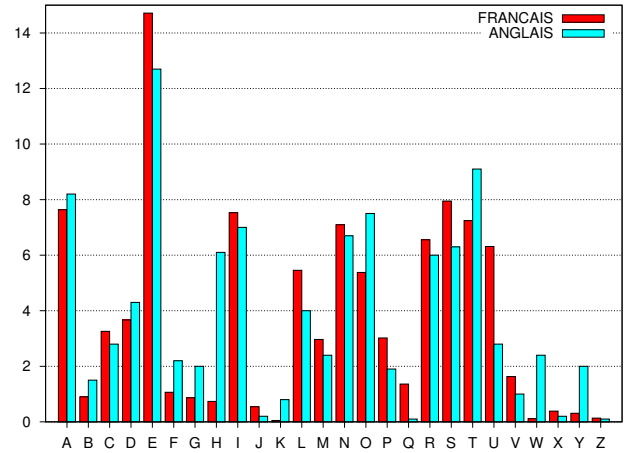
Constat : il y a plus de e ou de a que de z dans nos textes

Pourcentage mesuré des apparitions des lettres en français :



Application : les valeurs des lettres du jeu Scrabble

Comparaison des fréquences d'apparition des lettres



On parle de distribution des fréquences d'apparition des lettres

Attaque par analyse fréquentielle des lettres (1/4)

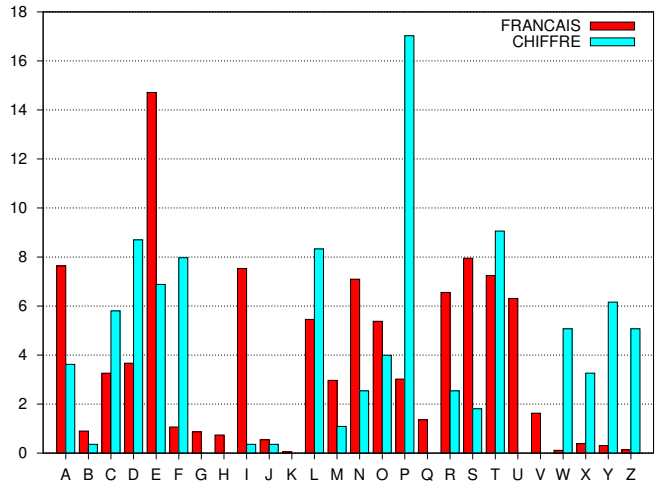
Exemple de texte chiffré :

> Szwxpd pelte lddtd opaftd awfdtpfcd spfcpd dtw-pyntpfdpxpye ; dzy ozd wzyr pe xtynp pelte nzfcmp dfc fy mznlw op nstxtp, olyd wpbfpw tw xpwlyrplte opd aczofted o'fyp zopfc ylfdplmzyop ezfep alcetn-fwtpcp. Dl epep pelte apynspp dfc dl aztectyp pe tw cpacpdpyelte, l xpd jpfi, w'txlrp o'fy rclyo ztdplf xltrcp lf awfxlrp rctd, l wl szfaap yztcp.

But : trouver la clé secrète du décalage

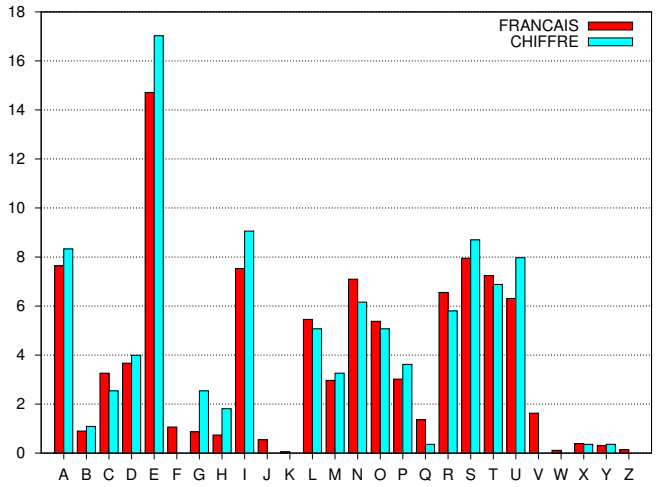
Attaque par analyse fréquentielle des lettres (2/4)

Les distributions ne semblent pas proches



Attaque par analyse fréquentielle des lettres (3/4)

Mais avec un décalage de 11 vers la gauche on a :



Attaque par analyse fréquentielle des lettres (4/4)

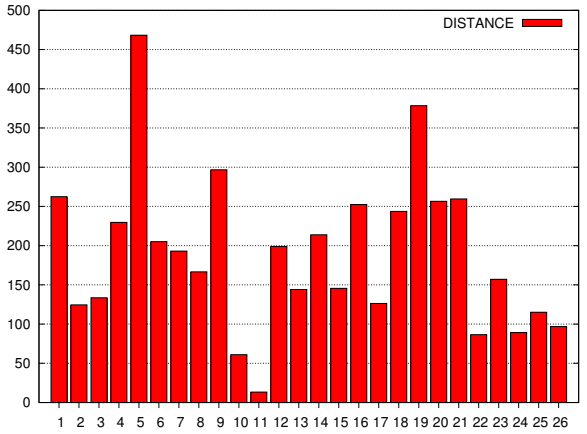
On a alors le text déchiffré suivant :

> Holmes etait assis depuis plusieurs heures silencieusement ; son dos long et mince etait courbe sur un bocal de chimie, dans lequel il melangeait des produits d'une odeur nauseabonde toute particuliere. Sa tete etait penchee sur sa poitrine et il representait, a mes yeux, l'image d'un grand oiseau maigre au plumage gris, a la houppe noire.

Remarque : extrait d'une nouvelle sur les aventures de Sherlock Holmes écrite par Arthur Conan Doyle (1903) et intitulée *Les Hommes dansants* (*The Adventure of the Dancing Men*)

Pouvons-nous trouver le secret par le calcul ?

Oui, la preuve :



Et c'est très simple et très très rapide ! ! !

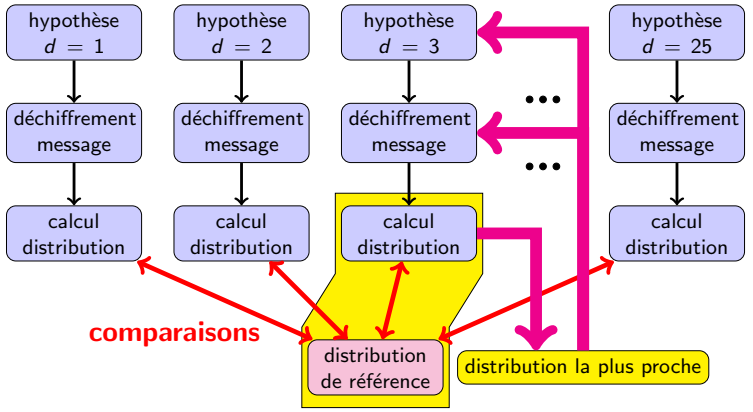
Attaque par calcul : choix du décalage

décalage d	A	B	C	D	E	···	X	Y	Z
				↑	$\xrightarrow{d}$	↓			
1	B	C	D	E	F	···	Y	Z	A
2	C	D	E	F	G	···	Z	A	B
3	D	E	F	G	H	···	A	B	C
⋮					⋮				
24	Y	Z	A	B	C	···	V	W	X
25	Z	A	B	C	D	···	W	X	Y

D	est le chiffré de	C	B	A	E
	quand	$d = 1$	$d = 2$	$d = 3$	$d = 25$

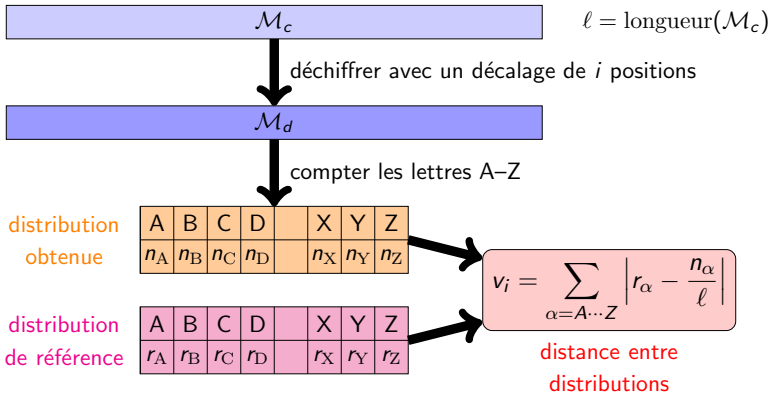
Mais une seule de ces possibilités est la bonne !

Attaque par calcul : idée



Attaque par calcul : les détails

Pour chaque valeur (hypothèse) de i dans $\{1, 2, 3, \cdots, 24, 25\}$ faire :



La « bonne » valeur du décalage est $d = \min_i v_i$

Attaques par force brute

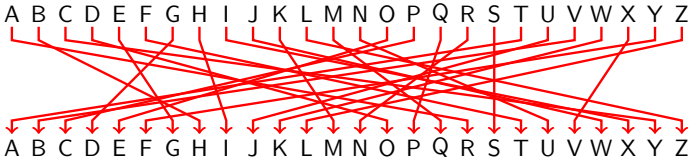
Principe : tester toutes les clés possibles

Limite : possible que si le nombre de clés possibles est raisonnable

Exemple du chiffrement par décalage de l'alphabet :

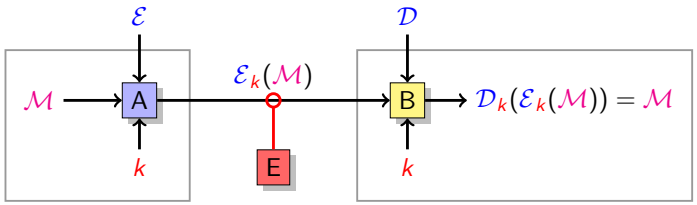
seulement 26 clés possibles

Exemple du chiffrement par permutation des lettres de l'alphabet :



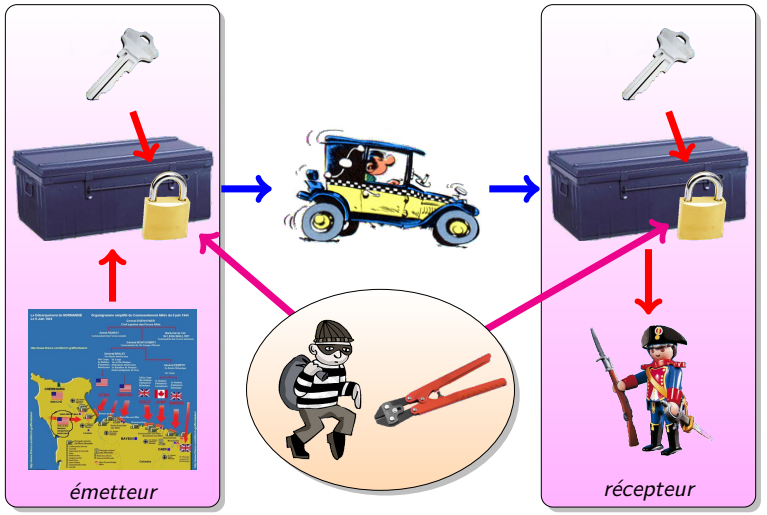
403 291 461 126 605 635 584 000 000 clés possibles

Cryptographie symétrique ou à clé secrète

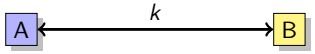
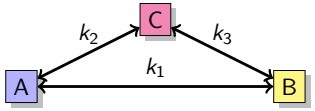
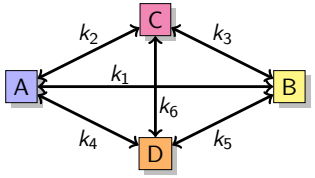


- A est Alice, B est Bob
- $\mathcal{M}$ est le message en clair que A souhaite envoyer à B
- $\mathcal{E}$ est l'algorithme de chiffrement, $\mathcal{D}$ celui de déchiffrement
- k est la cle secrète partagée entre A et B
- $\mathcal{E}_k(\mathcal{M})$ est le message chiffré
- $\mathcal{D}_k(\mathcal{E}_k(\mathcal{M}))$ est le message déchiffré
- E est l'espion qui écoute toutes les communications

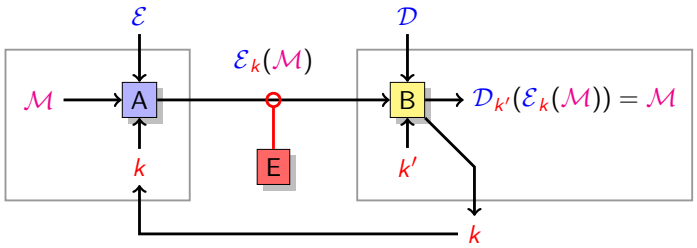
Analogie avec un coffre et un cadenas



Cryptographie symétrique : problème pour n personnes

personnes		clés nécessaires	
n	liste	liste	nombre
2	A, B		1
3	A, B, C		3
4	A, B, C, D		6
n	A,...		$\frac{n\times(n-1)}{2}$

Cryptographie asymétrique ou à clé publique



- k est la clé publique de B (connue de tout le monde, y compris E)
- $\mathcal{E}_k(\mathcal{M})$ est le message chiffré
- k' est la clé privée de B (que B doit garder secrète)
- $\mathcal{D}_{k'}(\mathcal{E}_k(\mathcal{M}))$ est le message déchiffré

Cryptographie symétrique ou asymétrique ?

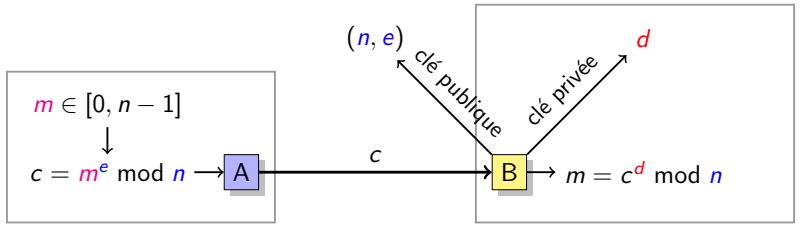
Cryptographie à clé privée ou symétrique :

- algorithmes assez simples
  - $\Longrightarrow$ calculs rapides
  - $\Longrightarrow$ coût modéré (surface de circuit, énergie)
- nécessite un échange de clé
- distribution des clés problématique pour n personnes

Cryptographie à clé publique ou asymétrique :

- pas besoin d'échange de clé
- 2 clés seulement par utilisateur (privée, publique)
- permet de faire des signatures numériques
- algorithmes plus complexes
  - $\Longrightarrow$ calculs plus lents
  - $\Longrightarrow$ coût plus élevé

Chiffrement RSA (Rivest, Shamir, Adleman 1977)



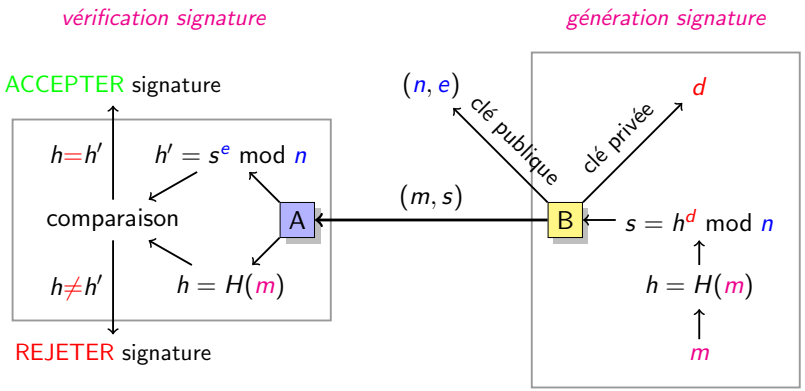
RSA : génération des clés pour B :

1. générer les premiers p et q (taille $l/2$)
2. calculer $n = pq$ et $\phi = (p-1)(q-1)$
3. sélectionner e tel que $1 < e < \phi$ que $\gcd(e, \phi) = 1$
4. calculer d qui satisfait $1 < d < \phi$ et $ed \equiv 1 \bmod \phi$

Sécurité :

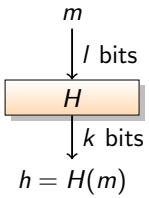
- problème de factorisation des grands entiers : calculer (p,q) à partir de n est très difficile
- taille minimale de clé recommendée : 1024 bits

Signature RSA



H est une fonction de hachage cryptographique

Fonction de hachage cryptographique (1/2)



- message m
- taille variable l
- haché ou *digest* h
- taille fixe k (en pratique $k \ll l$)

Propriétés de sécurité des fonctions de hachage cryptographique :

- résiste à la préimage (fct. sens unique) : $h \nrightarrow m \mid h = H(m)$
- résiste à la seconde préimage : $m_1 \nrightarrow m_2 \neq m_1 \mid H(m_1) = H(m_2)$
- résiste aux collisions : trouver (m_1, m_2) tel que $m_1 \neq m_2$ et $H(m_1) = H(m_2)$ est très difficile

Exemples : MD5, WHIRLPOOL, SHA-1, SHA-2, SHA-3 (sélection 2012)

Fonction de hachage cryptographique (2/2)

Exemples avec `openssl` :

```
> echo "string" | openssl dgst -sha224 -hex
```

texte	digest
0123456789	95ae4be607f065743ac1c81d1180591a919d08b8d4765e176b26f214
1123456789	5c527cd1341a4338f09086e71d1a0d69f818d74a828c974b9433524a
0123446789	94e5aa1d275dc3a21c76d28b011f4ea6121fa228af3ec7fa329da44f
0123456788	b04c6b0b1d663ad0c00d749441747cc6df211ea6c98f4fd2dbf283ff

Fonction à sens unique [à trappe]

Fonction à sens unique : $f : x \mapsto y = f(x)$

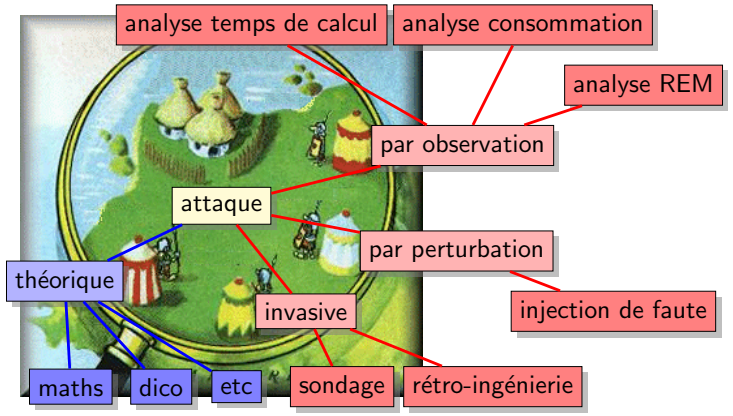
- étant donné x, calculer y est simple
- étant donné y, calculer x est très difficile

Fonction à sens unique à trappe : $f : x \mapsto y = f(x)$

- étant donné x, calculer y est simple
- étant donné y, calculer x est très difficile
- étant donnés une information (secrète) et y, calculer x is simple

Exemple : p et q premiers, calculer $n = pq$ est simple mais trouver (p, q) à partir seulement de n est très difficile

Quelques types d'attaques



REM = rayonnement électromagnétique

Attaques théoriques

Tout à fait réalisable en pratique avec :

- maths
- algorithmique de compétition
- programmation hyper hyper optimisée

Exemple : factorisation de grands entiers pour casser RSA

- RSA 512 bits, août 1999
- RSA 576 bits, décembre 2003
- RSA 640 bits, novembre 2005, équipe allemande avec 80 μP Opteron à 2.2 GHz
- RSA 768 bits, décembre 2009

Attaque RSA 768 en décembre 2009

6 mois de calcul sur 80 machines avec du parallélisme (équiv. 1 500 ans de calcul sur une seule machine simple !

RSA-768 =
33478071698956898786044169848212690817704794983713768568912431388982883793878002287614711652531743087737814467999489
×
36746043666799590428244633799627952632279158164343087642676032283815739666511279233373417143396810270092798736308917

Source : article (en anglais) `http://eprint.iacr.org/2010/006.pdf`
Factorization of a 768-bit RSA modulus. Thorsten Kleinjung, Kazumaro Aoki, Jens Franke, Arjen K. Lenstra, Emmanuel Thome, Joppe W. Bos, Pierrick Gaudry, Alexander Kruppa, Peter L. Montgomery, Dag Arne Osvik, Herman te Riele, Andrey Timofeev, and Paul Zimmermann

Pour éviter les attaques théoriques

Suivre les recommandations des spécialistes...

Exemple : courbe elliptique P-521 sur un corps premier recommandée par le NIST (cf. FIPS 186-2)

$p =$ 6864797660130609714981900799081393217269435300143305409394463459185543183397656052122559406614545549772963113914808580371219879997166438125740282911150571510000000001

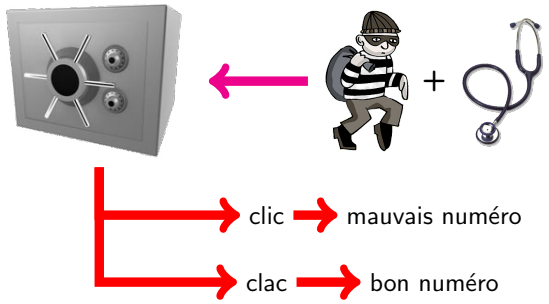
$r =$ 6864797660130609714981900799081393217269435300143305409394463459185543183397655394245057746333217197532963996371363321113864768612440380340372808892707005449

$s =$ d09e8800 291cb853 96cc6717 393284aa a0da64ba

$c =$ 0b4 8bfa5f42
0a349495 39d2bdfc 264eeeeb 077688e4 4fbf0ad8
f6d0edb3 7bd6b533 28100051 8e19f1b9 ffbe0fe9
ed8a3c22 00b8f875 e523868c 70c1e5bf 55bad637

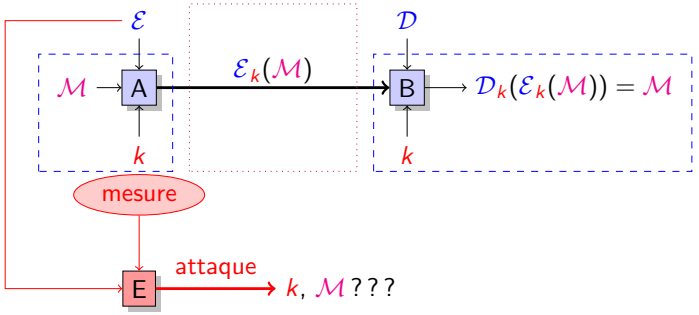
Attaques par canaux cachés, une nouvelle technique ?

Pas vraiment ! Vous connaissez tous des attaques par canaux cachés...



Attaque physique par canaux cachés

En anglais : *Side Channel Analysis/Attacks (SCA)*



Principe : mesurer des paramètres externes du dispositif pendant son fonctionnement pour en déduire des informations internes

Que mesurer ?

Réponse : tout ce qui peut "entrer" et/ou "sortir" dans le/du dispositif

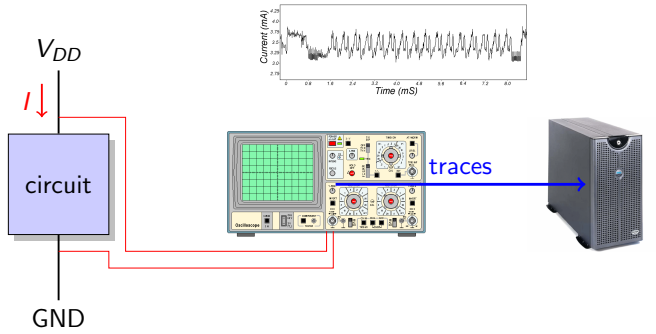
- la consommation d'énergie
- le rayonnement électromagnétique
- la température
- le bruit (son)
- le temps de calcul
- le nombre de défauts de cache
- le nombre et le type de messages d'erreur
- ...

Ce que l'on mesure peut donner des informations sur le comportement :

- global (température, consommation du circuit, bruit...)
- local (REM, nb défauts de cache...)

Attaque par mesure de la consommation d'énergie

Principe : mesurer le courant I qui alimente le dispositif

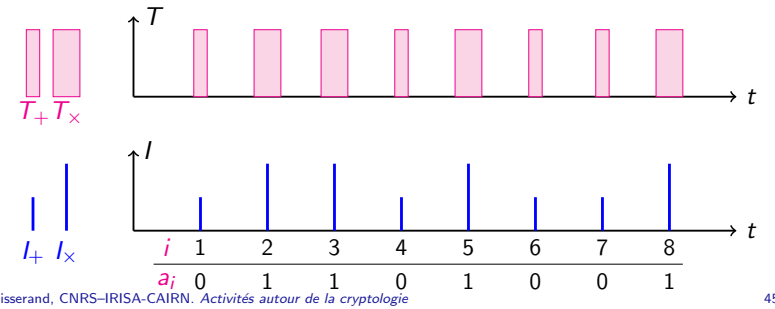


Notations : V_{DD} tension d'alimentation (5, 3, 2.5, 1.2 V), GND la masse

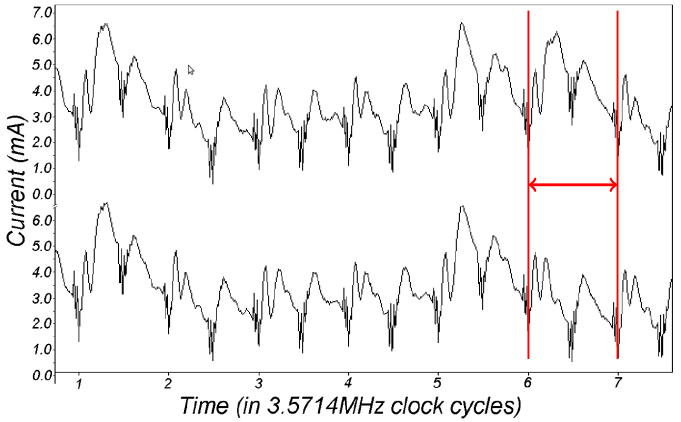
Exploiter les différences

Un algorithme a une signature en courant et en temps de calcul :

```
r = c_0
pour i allant de 1 à n faire
  si a_i = 0 alors
    r = r + c_1
  sinon
    r = r × c_2
```

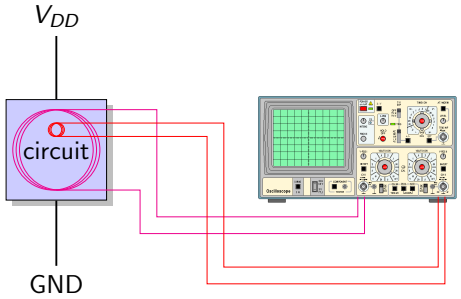


Analyse simple de la consommation



Analyse du rayonnement électromagnétique (1/2)

Principe : utiliser une sonde qui va capter le REM.



Mesure du REM :

- global avec une grande sonde
- local avec une microseconde

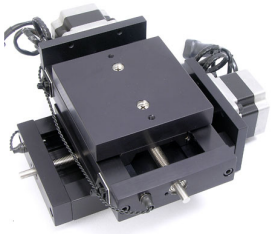
Analyse du rayonnement électromagnétique (2/2)

Types d'analyse du REM :

- simple : SEMA (*simple electromagnetic analysis*)
- différentielle : DEMA (*differential electromagnetic analysis*)

Le caractère local de l'analyse du REM peut être vraiment intéressant pour essayer de déterminer l'architecture du circuit, puis d'attaquer des endroits bien précis.

$\Longrightarrow$ table X-Y



Addition modulo M

Entrées :

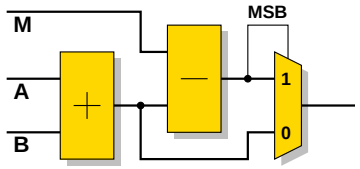
$$A, B \in \{0, 1, 2, 3, \ldots, M-1\}$$

Sortie [1] :

$$(A + B) \bmod M$$

Algorithme :

$$(A + B) \bmod M = \begin{cases} A + B & \text{si } A + B < M \\ A + B - M & \text{si } A + B \geq M \end{cases}$$



Contre-mesures

Empêcher une (ou des) attaques par :

- un nouveau dispositif de protection
- la modification/sécurisation du dispositif original

Exemples :

- blindage
- uniformiser les temps de calcul
- uniformiser la consommation d'énergie
- introduire du bruit (instructions inutiles)
- reconfigurer le circuit
  - changer le codage des données
  - changer les algorithmes
- ...

Système de représentation en base double

DBNS : en anglais *double-base number system*

Représentation redondante des nombres basée sur la somme de puissances de 2 ET de 3 :

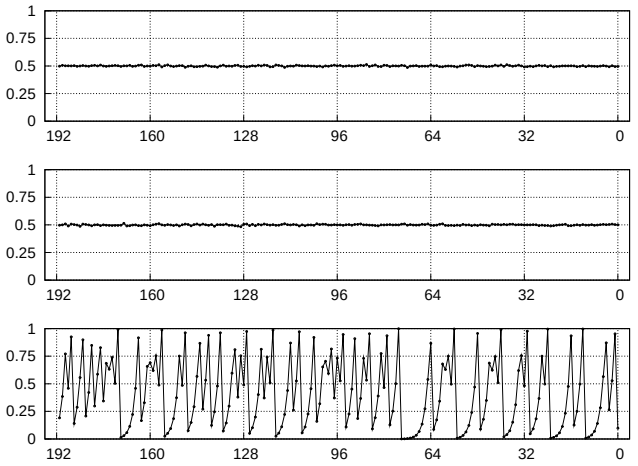
$$x = \sum_{i=1}^{n} x_i 2^{a_i} 3^{b_i}, \text{ with } x_i \in \{-1, 1\},\ a_i, b_i \geq 0$$

Exemple : $127 = 108 + 16 + 3 = 72 + 54 + 1 = \ldots$

	1	2	4	8	16
1					1
3	1				
9					
27			1		

	1	2	4	8
1	1			
3				
9				1
27		1		

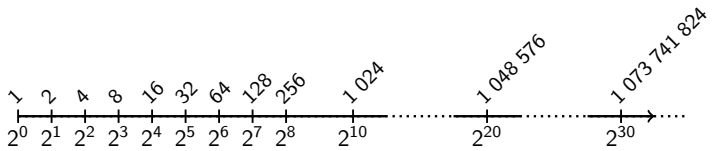
Activité électrique dans les opérateurs arithmétiques

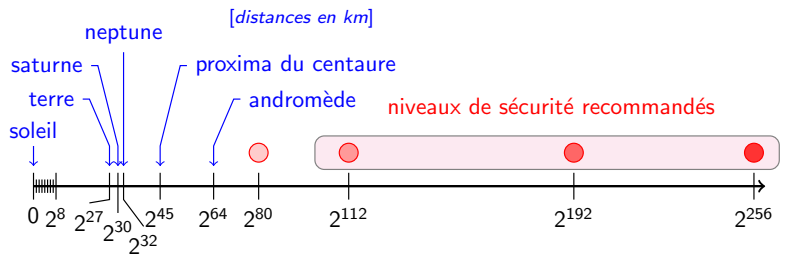


haut : addition, milieu : multiplication, bas : addition avec une constante

Niveau de sécurité

Nombre d'opérations à effectuer pour ≪ casser ≫ le système





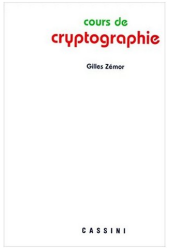
Livres sur le sujet

Histoire des codes secrets
Simon Singh
1999
Livre de poche

Mathématiques, espionnage et piratage informatique
Joan Gomez
2010
Le monde est mathématique, RBA

Livres sur le sujet

Cryptographie appliquée
Bruce Schneier
1997, 2ème édition
Wiley
ISBN: 2–84180–036–9



Cours de cryptographie
Gilles Zémor
2000
Cassini
ISBN: 2–84225–020–6

Fin, des questions ?

Contact:

- `mailto:arnaud.tisserand@irisa.fr`
- `http://www.irisa.fr/prive/Arnaud.Tisserand/`
- Equipe-projet CAIRN `http://www.irisa.fr/cairn/`
- Laboratoire IRISA, CNRS–INRIA–Univ. Rennes 1
  6 rue Kérampont, BP 80518, F-22305 Lannion cedex, France

Merci