LES TECHNIQUES DE CRYPTOGRAPHIE

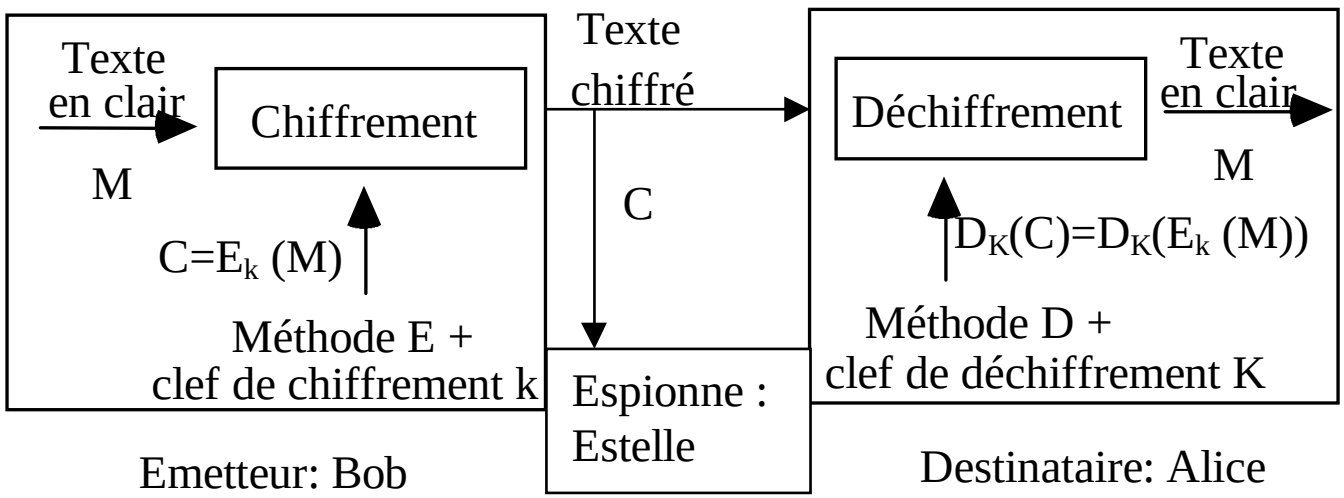
G. Florin

S. Natkin

Mars 2003

Généralités

Définition


Texte
en clair
M
Chiffrement
C=E_k (M)
Méthode E +
clef de chiffrement k
Emetteur: Bob
Texte
chiffré
C
Espionne :
Estelle
Déchiffrement
Texte
en clair
M
D_K(C)=D_K(E_k (M))
Méthode D +
clef de déchiffrement K
Destinataire: Alice

Chiffrement

Bob, doit transmettre à Alice, un message M $\in$ MESSAGES_A_ENVOYER.
M est dit “en clair“.

Estelle, une espionne, d’écouter la voie de communication pour connaître M.

Bob, construit un texte chiffré C$\in$ MESSAGES_CHIFFRES.

$$C = E_k(M). \quad \text{ou} \quad C = \{M\}_k^E$$

La fonction E_k dépend d’un paramètre k appelé clef de chiffrement.

Le **chiffrement** est donc une transformation d'un texte pour en cacher le sens

La possibilité de chiffrer repose donc sur la connaissance de l’algorithme de chiffrement E et de la clef k de chiffrement.

Déchiffrement

Le **déchiffrement** est l'opération inverse permettant de récupérer le texte en clair à partir du texte C chiffré.
Il repose sur la fonction D_K de MESSAGES_CHIFFRES dans MESSAGES_A_ENVOYER telle que

$$M= D_{K'}(C) \text{ ou } C = \{M\}_K^D$$

On doit avoir

$$D_K (E_k(M))=M$$

D_K est donc une fonction inverse à gauche de E_k.

Pour un couple cr = (E,D) donné de famille de fonction de chiffrement et de déchiffrement, l'ensemble des couples (k, K) vérifiant cette propriété est noté CLE(cr).

Crypto-systèmes

Pour que ces opérations assurent la confidentialité du transfert entre Alice et Bob, il est nécessaire qu'au moins une partie des informations E, D, k, K soit ignorée du reste du monde.

Décrypter ou casser un code c'est parvenir au texte en clair sans posséder au départ ces informations secrètes. C'est l'opération que doit réaliser Estelle pour retrouver M.

L'art de définir des codes est la cryptographie. Un spécialiste en cryptographie est appelé cryptographe.

L'art de casser des codes est appelé cryptanalyse ou cryptologie. Un spécialiste en cryptanalyse est appelé cryptanalyste.

Un crypto-système est l'ensemble des deux méthodes de chiffrement et de déchiffrement utilisable en sécurité.

Crypto-systèmes symétriques

Tels que soit k=K, soit la connaissance d'une des deux clefs permet d'en déduire facilement l'autre.

Conséquences :
Dichotomie du monde : les bons et les mauvais

Multiplication des clefs (un secret n'est partagé que par 2 interlocuteurs), donc pour N interlocuteurs N.(N-1)/2 couples

La qualité d'un crypto système symétrique s'analyse par rapport à des propriétés statistiques des textes chiffrés et la résistance aux classes d'attaques connues.

En pratique tant qu'un crypto système symétrique n'a pas été cassé, il est bon, après il est mauvais.

Crypto-systèmes asymétriques (a clefs publiques)

Tels que la connaissance de k (la clef de chiffrement) ne permet pas d'en déduire celle de K(la clef de déchiffrement).

Un tel crypto-système est dit asymétrique, la clef k est appelée la **clef publique**, la clef K est appelée la **clef privée**.

Fondement théorique : montrer que la recherche de K à partir de k revient à résoudre un problème mathématique notoirement très compliqué, c'est à dire demandant un grand nombre d'opérations et beaucoup de mémoire pour effectuer les calculs.

RSA (l'algorithme le plus utilisé à l'heure actuel) la déduction de K à partir de k revient à résoudre le problème de factorisation d'un grand nombre un problème sur lequel travaille les mathématiciens depuis plus de 2000 ans,

On estime que le plus rapide ordinateur que l'on puisse construire utilisant la meilleure méthode connue met plus de 1000 ans pour retrouver la clef privée d'un système RSA utilisant un modulo de 1024 bits (ordre de grandeur de la taille des clefs).

Asymétrie de l ’usage des clefs

BANQUE_MODERNE, désire autoriser ses clients à envoyer des ordres de virement chiffrés
Elle publie dans un annuaire infalsifiable

Nom = BANQUE_MODERNE, Algorithme de Chiffrement = E, Clef Publique = k

La banque conserve secrète la clef privée K.

Tout client peut calculer $C= E_k(M)$.

Seul la banque peut déchiffrer le message $M = D_K(C)$.

L ’algorithme d ’Estelle (Cryptanalyste)
Etape 1) Recherche des crypto -systèmes possibles

Hypothèses
Estelle veut décrypter C= E_k(M).
Estelle ne connaît ni D, ni E ni k, ni K.
Elle peut connaître des informations sur sa syntaxe et la sémantique de M.

Etape 1) Recherche des crypto systèmes possibles
CR={cr=(D,E)}

L ’algorithme d ’Estelle
Etape 2) Réduction de l ’espace des clefs

Pour tout cr déterminer le plus petit ensemble CLE_REDUIT⊂CLE(cr) contenant la clef utilisée.
Si card (CLE_REDUIT={(k,K)})=1, M= $D_K(C)$. Fin

A priori, tous les couples (k,K) sont équiprobable sur CLE(cr)
Estelle doit acquérir une connaissance soit déterministe (clefs impossibles) ou probabiliste (clefs improbables)qui facilite ses essais (réduit l ’entropie)

Exemples
Estelle possède M’ et C’ chiffrée avec le même cryptosystème et les mêmes clefs déduction de propriétés des clefs: **attaque à texte en clair**.
Estelle peut chiffrer des messages M avec E_k (sans connaître k) **attaque à texte en clair choisi**.

Estelle connaît des propriétés de l ’algorithme de génération de (k,K).

L ’algorithme d ’Estelle
Etape 3) Analyse syntaxique

Déterminer le plus petit ensemble MESSAGES_SYNTAXIQUEMENT_CORRECTS ⊃ MESSAGES_A_ENVOYER qui vérifie des propriétés de syntaxe connues d’Estelle

Objectif : Construire un test d’arrêt simple pour le calcul mené à l’étape 4
Une règle syntaxique est sous une forme ou une autre un invariant d’un langage. Elle implique donc une certaine redondance de l’information.
Exemples
Le plus grand mot de la langue française a 25 lettres (anticonstitutionnellement). Possibilité d’écrire 10^{34} mots de 25 lettres ou moins 80000 mots dans le dictionnaire Hachette
"le" est nécessairement suivi d’un nom masculin

- Règles logique classique (toutes les suites de huit bits à partir du début du texte appartiennent à l’alphabet ASCII)
- Règles résultant de l’application d’un test statistique permettant d’accepter ou de rejeter une hypothèse (la répartition des caractères ASCII dans le texte en clair suit la même loi que la répartition des lettres dans la langue française). Fréquences d'apparition (en anglais)

Lettres	Digrammes	Trigrammes
E 13,05	TH 3,16	THE 4,72
T9,02	IN 1,54	ING 1,42

L ’algorithme d ’Estelle
Etape 3) Analyse syntaxique 2

Deux cas possibles

Etape 3.1 Construction de l’espace des messages

Informations très précise sur la syntaxe de M (M est un mot de passe sur 8 caractères qui est très probablement composé d’un mot ou de deux mots français concatènes).

Etape 3.2 Construction d’une règle syntaxique

$\exists$SYN tel que alors $\forall$ M

$\in$ MESSAGES_SYNTAXIQUEMENT_CORRECTS SYN(M)=vrai.

L ’algorithme d ’Estelle
Etape 4) Recherche Exhaustive

Construire MESSAGES_POSSIBLES$\subset$ MESSAGES_SYNTAXIQUEMENT_CORRECTS tel que mes $\in$ MESSAGES_POSSIBLES

Soit Etape 4.1 : Recherche sur l’espace des clefs de chiffrement
mes $\in$ MESSAGES_SYNTAXIQUEMENT_CORRECTS et $\exists$ (k, K) $\in$ CLE_REDUIT(cr) et $E_k(mes)=C$

Soit Etape 4.2 : Recherche sur l’espace des clefs de déchiffrement
$\exists$(k, K) $\in$ CLE_REDUIT(cr)et $D_K(C)$=mes et SYN(mes).

Si card(MESSAGES_POSSIBLES)=1, M=mes, Fin

Attaque à texte chiffré en parcourant itérativement soit l’espace des clefs de chiffrement soit celui des clefs de déchiffrement.

L ’algorithme d ’Estelle
Etape 5) Analyse sémantique

Trouver une règle sémantique SEM
(le message porte sur la cocaïne ou les fausses factures)
telle que :
card ({X MESSAGES_POSSIBLES tel que SEM(X)})=1
Si une telle règle existe M=X, Fin

Sinon Estelle a échouée

Point de vue du cryptographe
Etape 1

Opération autrefois difficile,
devenue simple: standard de cryptographie, systèmes commercialisés.
la sécurité d ’un crypto-système ne repose plus que sur le secret des clefs
(sauf dans le domaine militaire).

Point de vue du cryptographe
Etape 2

Choisir un crypto système cr dont l'espace des clefs est très grand.

Choix des clefs est le plus imprédictible possible
(éviter les mots d'un dictionnaire , nombres pseudo aléatoires à grain de génération difficile à deviner)

Limiter l 'usage des clefs

Choisir un bon crypto asymétrique tel que le calcul de K à partir de k,
ou même de la réduction des K possibles connaissant k est un problème reconnu scientifiquement comme très difficile.
Si par un hasard extraordinaire, Estelle arrive à résoudre ce problème,
elle devient célèbre, riche et par conséquent heureuse en amour.
Elle n'a donc plus aucune raison d'embêter Bob et Alice.

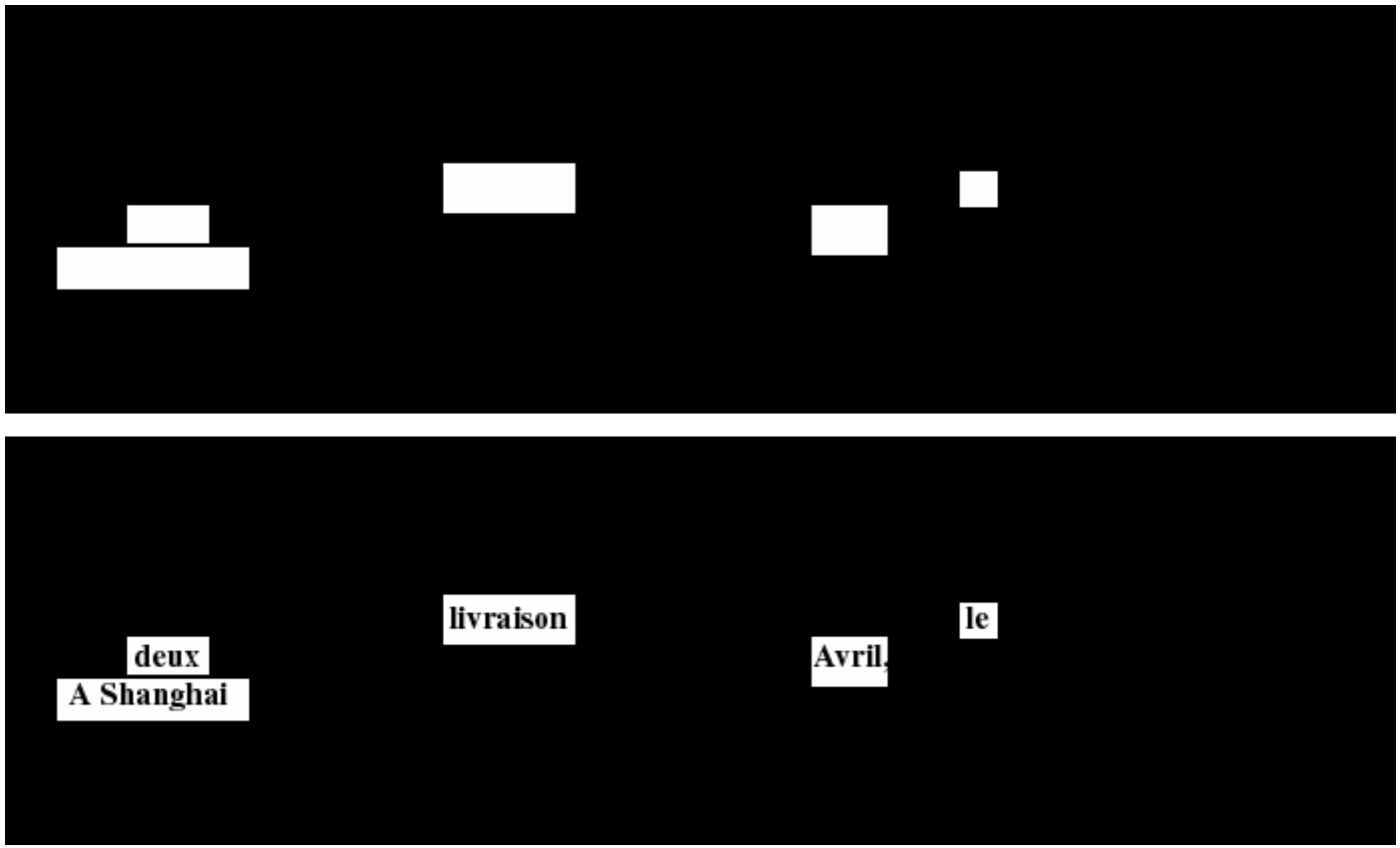
Point de vue du cryptographe
Etape 3

Deux stratégies :

Limiter la redondance (compression à un niveau syntaxique bas)

Augmenter la redondance en donnant plusieurs syntaxes possibles pour un même message (texte caché dans une image)

Masque classique



Cher ami
Je pense pouvoir assurer la livraison des 3O tonnes de blé prévue le 10 mars.
Mes deux assistants ne pourront venir vous voir avant Avril, mais je vous attends au Lotus Bleu
A Shanghai
Rastapopoulos

Point de vue du cryptographe
Etape 4
le masque jetable

Méthode imparable : le **chiffre parfait** ou masque jetable.

M sous forme d'une suite de n bits. Clef k_M de n bits, parfaitement aléatoire (suite uniforme de bits) utilisée qu'une seule fois (l'étape de réduction de l'espace des clefs n'a apportée aucune information)
$\Rightarrow$ Essai de toutes les clefs de CLE(cr) qui est l'ensemble des suites de n bits.

Chiffrement: $C=E_{k_M}(M)=M \oplus k_M$
Ou $\oplus$ représente le ou exclusif

Déchiffrement: $M=D_{k_M}(C)=C \oplus k_M$

très rapide et sans faille.

Point de vue du cryptographe
Etape 4
le masque jetable 2

$\forall Mess \in MESSAGES\_SYNTAXIQUEMENT\_CORRECTS \; \exists X \in CLE(cr)$ tel que

$$C = E_X(Mess)$$

$X = C \oplus Mess \Rightarrow$

$$C = X \oplus Mess = C \oplus Mess \oplus Mess = C$$

En balayant tout l'espace des clefs on trouvera tous les messages de MESSAGES_SYNTAXIQUEMENT_CORRECTS.
Le fait d'avoir espionné pour connaître C n' a apporté aucune information.

Point de vue du cryptographe
Etape 4
le masque jetable 3

Notons :
A l'événement : C a été reçu par Estelle
B l' événement : M a été émis par Bob

Information apportée par la réception de :

$$I=-\log_2(\text{Probabilité(B sachant A)}/\text{Probabilité(B)})$$

Or comme toutes les clefs sont équiprobables, C a pu être construit avec à partir de n'importe quel message possible M'.Donc A et B sont indépendants.

$$\text{Probabilité(B sachant A)}=$$
$$\text{Probabilité(A).Probabilité(B)}/\text{Probabilité(A)}=\text{Probabilité(B)}$$

$$I=-\log_2(\text{Probabilité(B)}/\text{Probabilité(B)})=-\log_2(1)=0$$

Conclusion

- Le problème de cryptanalyse est souvent posé en terme de complexité de la recherche exhaustive
- Pourtant le problème du test d'arrêt est souvent beaucoup plus complexe: la découverte de la syntaxe et à fortiori du sens d'un message par un ordinateur relève de la reconnaissance des formes
- Une technique de cryptographie élémentaire rend un système d'espionnage systématique (Echelon) inefficace