Systèmes et sécurité : Introduction à la cryptographie

Fangan-Yssouf Dosso

dosso@univ-tln.fr

IMATH, Université de Toulon

April 7, 2021

Introduction: objectifs du cours

- Sensibilisation aux notions de chiffrement, de déchiffrement et de cryptanalyse.
- Présentation des schémas de chiffrement historiques : César, Vigénère, Hill, ...
- Étude du standard actuel de chiffrement à clé sécrète : l'AES.
- * AES : Advanced Encryption Standard

Déroulement du cours

Les séances :

- Cours : 7 séances de 1h30, en distanciel.
- TP : 2 séances de 3h (par groupe), en présentiel.

Évaluations :

- 1 contrôle terminal.
- 1 contrôle de TP.

Chiffrement et déchiffrement : idée principale

Un schéma de chiffrement est défini par quatre ensembles \mathfrak{M} , \mathfrak{C} , \mathfrak{K} et \mathfrak{K}' et deux applications \mathbf{e} et \mathbf{d} , tels que :

- $e: \mathcal{M} \times \mathcal{K} \longrightarrow \mathcal{C}$
- $d: \mathfrak{C} \times \mathfrak{K}' \longrightarrow \mathfrak{M}$
- Soit $m \in \mathcal{M}$. Pour tout $k \in \mathcal{K}$, il existe $k' \in \mathcal{K}'$, tel que : $\mathbf{d}(\mathbf{e}(m, k), k') = m$.

Quelques terminologies:

- e et d sont respectivement les fonctions de chiffrement et de déchiffrement.
- M et C sont respectivement les espaces des messages clairs et chiffrés.
- \mathcal{K} et \mathcal{K}' sont respectivement les espaces des clés de chiffrements et de déchiffrements.

Cryptographie à clé sécrète

Soit $m \in \mathcal{M}$.

Définition:

Un schéma de chiffrement est dit à clé sécrète si :

- $\mathfrak{K}=\mathfrak{K}'$ et
- pour tout $k \in \mathcal{K}$,

$$\mathbf{d}(\mathbf{e}(m, k), k) = m.$$

Exemples : Le code de César, DES, AES

Remarque:

Pour utiliser un tel schéma, il faut donc **au préalable** avoir échangé de façon confidentiel la **clé sécrète** *k*. Intéressant ..., n'est-ce pas ?

Cryptographie à clé publique

Définition

Un schéma de chiffrement qui n'impose pas la même clé pour le chiffrement et le déchiffrement est dit à clé publique. Dans ce cas,

- k est appelée la clé publique,
- k' est appelée la clé sécrète.

Exemples : RSA, Cryptosystème ElGamal

Remarque:

Le problème d'échange de clé précédent ne se pose plus.

Dans ce cas, pourquoi s'intéresse-t-on encore aux schémas de chiffrement à clé sécrète ?

Pour information :

Ce cours ne portera que sur les schémas de chiffrement à clé sécrète.

Un schéma de chiffrement historique : le code de César

Le code de César

- Un schéma de chiffrement par substitution mono-alphabétique.
- On identifie l'ensemble des 26 lettres de l'alphabet à l'ensemble $\mathbb{Z}/26\mathbb{Z}$, i.e. $a=0,\ b=1,\ ...,\ z=25.$
- $\mathbf{M} = \mathbf{C} = \{ \text{l'ensemble des textes possibles avec les 26 lettres de l'alphabet} \} = \{0,1,2,\ldots,25\}^*$
- $\mathcal{K} = \mathcal{K}' = \mathbb{Z}/26\mathbb{Z}$.

Quel est le nombre de clés ?

Code de César : le chiffrement

Le principe

Soit $M=m_1m_2\dots m_n\in (\mathbb{Z}/26\mathbb{Z})^n$, un message de longueur n et $k\in \mathbb{Z}/26\mathbb{Z}$ une clé. Le chiffré $C=c_1c_2\dots c_n$ de M est tel que :

$$c_i = (m_i + k)\%26$$

Code de César : le chiffrement

Le principe

Soit $M=m_1m_2\dots m_n\in (\mathbb{Z}/26\mathbb{Z})^n$, un message de longueur n et $k\in \mathbb{Z}/26\mathbb{Z}$ une clé. Le chiffré $C=c_1c_2\dots c_n$ de M est tel que :

$$c_i = (m_i + k)\%26$$

Exemple: chiffrement d'un texte

On prend M = "Un texte à chiffrer" et <math>k = 3. Calculer le chiffré C de M.

On ignorera les espaces et accents; les majuscules et minuscules sont prises équivalentes.

Code de César : le chiffrement

Le principe

Soit $M=m_1m_2\dots m_n\in (\mathbb{Z}/26\mathbb{Z})^n$, un message de longueur n et $k\in \mathbb{Z}/26\mathbb{Z}$ une clé. Le chiffré $C=c_1c_2\dots c_n$ de M est tel que :

$$c_i = (m_i + k)\%26$$

Exemple: chiffrement d'un texte

On prend M = "Un texte à chiffrer" et k = 3. Calculer le chiffré C de M.

On ignorera les espaces et accents; les majuscules et minuscules sont prises équivalentes.

On a :
$$u = 20$$
, $n = 13$, etc.. Donc, $c_1 = (20 + 3)\%26 = 23 = x$, $c_2 = (13 + 3)\%26 = 16 = q$, etc.

Le chiffré est donc :

Code de César : le déchiffrement

Le principe

Soit $C = c_1 c_2 \dots c_n \in (\mathbb{Z}/26\mathbb{Z})^n$, un chiffré de longueur n et $k \in \mathbb{Z}/26\mathbb{Z}$ une clé. Le message clair $M = m_1 m_2 \dots m_n$ correspondant est tel que :

$$m_i = (c_i - k)\%26$$

Code de César : le déchiffrement

Le principe

Soit $C = c_1 c_2 \dots c_n \in (\mathbb{Z}/26\mathbb{Z})^n$, un chiffré de longueur n et $k \in \mathbb{Z}/26\mathbb{Z}$ une clé. Le message clair $M = m_1 m_2 \dots m_n$ correspondant est tel que :

$$m_i = (c_i - k)\%26$$

Exemple : déchiffrement d'un texte

On prend C = "Un texte à déchiffrer" et k=5. Calculer le message clair correspondant à C.

On ignorera les espaces et accents; les majuscules et minuscules sont prises équivalentes.

Réponse :

Le Chiffre affine : une généralisation du Code de César

Le principe

- Un schéma de chiffrement par substitution mono-alphabétique.
- On identifie l'ensemble des 26 lettres de l'alphabet à l'ensemble $\mathbb{Z}/26\mathbb{Z}$, i.e. $a=0,\ b=1,\ ...,\ z=25.$
- $\mathcal{M}=\mathcal{C}=$ {l'ensemble des textes possibles avec les 26 lettres de l'alphabet} = $\{0,1,2,\ldots,25\}^*$
- $\mathcal{K} = \mathcal{K}' = (\mathbb{Z}/26\mathbb{Z})^* \times \mathbb{Z}/26\mathbb{Z}$.

 $(\mathbb{Z}/26\mathbb{Z})^*$: l'ensemble des éléments inversibles de $(\mathbb{Z}/26\mathbb{Z}\setminus\{0\},\times)$.

Le Chiffre affine : une généralisation du Code de César

Le principe

- Un schéma de chiffrement par substitution mono-alphabétique.
- On identifie l'ensemble des 26 lettres de l'alphabet à l'ensemble $\mathbb{Z}/26\mathbb{Z}$, i.e. $a=0,\ b=1,\ ...,\ z=25.$
- $\mathbf{M}=\mathbf{C}=\{$ l'ensemble des textes possibles avec les 26 lettres de l'alphabet $\}=\{0,1,2,\ldots,25\}^*$
- $\mathcal{K} = \mathcal{K}' = (\mathbb{Z}/26\mathbb{Z})^* \times \mathbb{Z}/26\mathbb{Z}$.

 $(\mathbb{Z}/26\mathbb{Z})^*$: l'ensemble des éléments inversibles de $(\mathbb{Z}/26\mathbb{Z}\setminus\{0\},\times)$.

Quel est le nombre de clés ?

Le Chiffre affine : une généralisation du Code de César

Le principe

- Un schéma de chiffrement par substitution mono-alphabétique.
- On identifie l'ensemble des 26 lettres de l'alphabet à l'ensemble $\mathbb{Z}/26\mathbb{Z}$, i.e. $a=0,\ b=1,\ ...,\ z=25.$
- $\mathcal{M} = \mathcal{C} = \{ \text{l'ensemble des textes possibles avec les 26 lettres de l'alphabet} \} = \{0, 1, 2, \dots, 25\}^*$
- $\mathcal{K} = \mathcal{K}' = (\mathbb{Z}/26\mathbb{Z})^* \times \mathbb{Z}/26\mathbb{Z}$.

 $(\mathbb{Z}/26\mathbb{Z})^*$: l'ensemble des éléments inversibles de $(\mathbb{Z}/26\mathbb{Z}\setminus\{0\},\times)$.

Quel est le nombre de clés ?

Exemples de clés : (3, 18), (9, 2).

Chiffre affine : le chiffrement

Le principe

Soit $M=m_1m_2\dots m_n\in (\mathbb{Z}/26\mathbb{Z})^n$, un message de longueur n et $(a,b)\in (\mathbb{Z}/26\mathbb{Z})^*\times \mathbb{Z}/26\mathbb{Z}$ une clé. Le chiffré $C=c_1c_2\dots c_n$ de M est tel que :

$$c_i = (am_i + b)\%26$$

Que remarque-t-on pour a = 1?

Chiffre affine : le chiffrement

Le principe

Soit $M=m_1m_2\dots m_n\in (\mathbb{Z}/26\mathbb{Z})^n$, un message de longueur n et $(a,b)\in (\mathbb{Z}/26\mathbb{Z})^*\times \mathbb{Z}/26\mathbb{Z}$ une clé. Le chiffré $C=c_1c_2\dots c_n$ de M est tel que :

$$c_i = (am_i + b)\%26$$

Que remarque-t-on pour a = 1?

Exemple: chiffrement d'un texte

On prend M = "Un texte à chiffrer" et <math>(a, b) = (5, 1). Calculer le chiffré C de M.

On ignorera les espaces et accents; les majuscules et minuscules sont prises équivalentes.

Réponse :

Chiffre affine : le déchiffrement

Le principe

Soit $C = c_1 c_2 \dots c_n \in (\mathbb{Z}/26\mathbb{Z})^n$, un chiffré de longueur n et $(a,b) \in (\mathbb{Z}/26\mathbb{Z})^* \times \mathbb{Z}/26\mathbb{Z}$, une clé. Le message clair $M = m_1 m_2 \dots m_n$ correspondant est tel que :

$$m_i = a^{-1}(c_i - b)\%26$$

Chiffre affine : le déchiffrement

Le principe

Soit $C = c_1 c_2 \dots c_n \in (\mathbb{Z}/26\mathbb{Z})^n$, un chiffré de longueur n et $(a,b) \in (\mathbb{Z}/26\mathbb{Z})^* \times \mathbb{Z}/26\mathbb{Z}$, une clé.

Le message clair $M=m_1m_2\dots m_n$ correspondant est tel que :

$$m_i = a^{-1}(c_i - b)\%26$$

Exemple : déchiffrement d'un texte

On prend C = "Un texte à déchiffrer" et <math>(a, b) = (17, 4). Calculer le message clair correspondant à C.

On ignorera les espaces et accents; les majuscules et minuscules sont prises équivalentes.

Réponse :

Cryptanalyse du Chiffre affine

- On teste toutes les clés possibles. Il y en a 312.
- On peut effectuer une analyse fréquentielle.

Cryptanalyse du Chiffre affine

- On teste toutes les clés possibles. Il y en a 312.
- On peut effectuer une analyse fréquentielle.

Analyse fréquentielle :

- Rappel : un schéma de chiffrement par substitution mono-alphabétique.
- Donc, pour une clé donnée, un symbole est toujours remplacé par le même symbole.
- Donc, le Chiffre affine est vulnérable à l'analyse fréquentielle.

Chiffre affine : l'analyse fréquentielle

Efficace pour des messages suffisamment longs.

Le principe :

- On classe les symboles du langage cible par fréquence d'apparition (d'utilisation).
- On identifie le symbole le plus fréquent dans le message chiffré.
- On suppose que ce symbole est le chiffré du symbole le plus fréquent du langage utilisé. On peut faire de même pour d'autres symboles.
- On en déduit une ou plusieurs équation(s) assez simple(s) qu'on résout.

Exemple d'analyse fréquentielle

Fréquence des lettres les plus utilisées dans la langue française :

E (15.87), A (9.42), I (8.41), S (7.90), T (7.26). (Wikipédia)

Décrypter ce message, chiffré avec le Code de César :

Hjhnjxyzsywjxgjfzyjcyj

Exemple d'analyse fréquentielle

Fréquence des lettres les plus utilisées dans la langue française :

E (15.87), A (9.42), I (8.41), S (7.90), T (7.26). (Wikipédia)

Décrypter ce message, chiffré avec le Code de César :

Hjhnjxyzsywjxgjfzyjcyj

Réponse :

- Clé: 5
- Message clair : Ceci est un très beau texte.

Exercice : message (chiffré avec le Code de César) à décrypter

 $fypgztefcpgpylteolcctgpcopwlfecpnzepopwlmlcctpcpfypapcdzyypdzcetefy\\ xtwteltcptwpeltenzxxpolydwpdqtwxdoprfpccpapydlolgtowpdopnzcletzyd\\ cpxawtddltpyewlglyeopdlgpdeptwdlaacznslopwlgztefcpzfdpeczfglteolgtowp\\ nslfqqpfczfgctewlqpypecp$

Chiffre de Vigenère

Le principe

- Un schéma de chiffrement par substitution polyalphabétique.
- On identifie l'ensemble des 26 lettres de l'alphabet à l'ensemble $\mathbb{Z}/26\mathbb{Z}$, i.e. $a=0,\ b=1,\ ...,\ z=25.$
- $\mathcal{M}=\mathcal{C}=\{$ l'ensemble des textes possibles avec les 26 lettres de l'alphabet $\}=\{0,1,2,\ldots,25\}^*$
- $\mathfrak{K} = \mathfrak{K}' = (\mathbb{Z}/26\mathbb{Z})^n$, avec n > 0 un entier.
- n est la longueur de clé.

substitution polyalphabétique : une même lettre peut être remplacée par des lettres différentes.

Chiffre de Vigenère : chiffrement

Le principe

- On commence par découper le texte en blocs de taille n, qui seront chiffrés indépendamment.
- Pour chiffrer un bloc, on effectue une addition (par colonne) de ce dernier et de la clé. Donc, la *i*-ième lettre du bloc est additionnée avec la *i*-ième lettre de la clé.

Les additions sont faites dans $\mathbb{Z}/26\mathbb{Z}$.

Exemple

- Texte en clair : Ceci est un rappel sur le chiffre de Vigenère
- Clé : lundi

Exemple

- Texte en clair : Ceci est un rappel sur le chiffre de Vigenère
- Clé : lundi

Découpage

Texte en clair découpé : cecie stunr appel surle chiff redev igene re

Clé répétée : lundi lun

Exemple

- Texte en clair : Ceci est un rappel sur le chiffre de Vigenère
- Clé : lundi

Découpage

- Texte en clair découpé : cecie stunr appel surle chiff redev igene re
- Clé répétée : lundi lun

Chiffrement

- Texte en clair découpé : cecie stunr appel surle chiff redev igene re
- Clé répétée : lundi lundi lundi lundi lundi lundi lundi lu
- Texte chiffré découpé : nyplm dnhqz ljcht doeom nbvin cyqhd tarqm cy

Exemple

- Texte en clair : Ceci est un rappel sur le chiffre de Vigenère
- Clé : lundi

Découpage

- Texte en clair découpé : cecie stunr appel surle chiff redev igene re
- Clé répétée : lundi lun

Chiffrement

- Texte en clair découpé : cecie stunr appel surle chiff redev igene re
- Clé répétée : lundi lundi lundi lundi lundi lundi lundi lu
- Texte chiffré découpé : nyplm dnhqz ljcht doeom nbvin cyqhd tarqm cy

Le texte chiffré est donc : nyplmdnhqzljchtdoeomnbvincyqhdtarqmcy

Chiffre de Vigenère : déchiffrement

Le principe est le même que celui du chiffrement. L'unique différence est qu'on effectue ici des soustractions (et non des additions) avec les lettres de la clé pour le déchiffrement des blocs.

Exercice

Déchiffrer le message suivant :

ykscetktgipiwvkopriosxurmcjezwqe

La clé utilisée est : exofacile

Chiffre de Vigenère : cryptanalyse

Le principe : deux étapes

- Étape 1 : Estimation de la longueur de la clé.
- Étape 2 : Recherche de la valeur de la clé.

Chiffre de Vigenère : cryptanalyse

Le principe : deux étapes

- Étape 1 : Estimation de la longueur de la clé.
- Étape 2 : Recherche de la valeur de la clé.

Remarque essentielle pour l'étape 1

Si deux groupes de lettres sont chiffrés avec le même morceau de la clé répétée, alors la distance entre ces deux groupes de lettres est un multiple de la longueur de la clé.

Chiffre de Vigenère : cryptanalyse

Le principe : deux étapes

- Étape 1 : Estimation de la longueur de la clé.
- Étape 2 : Recherche de la valeur de la clé.

Remarque essentielle pour l'étape 1

Si deux groupes de lettres sont chiffrés avec le même morceau de la clé répétée, alors la distance entre ces deux groupes de lettres est un multiple de la longueur de la clé.

Exemple

- Texte en clair découpé : cecie stunr appel surle chiff redev igene re
- Clé répétée : lundi lundi lundi lundi lundi lundi lundi lu
- Texte chiffré découpé : nyplm dnhqz ljcht doeom nbvin cyqhd tarqm cy

La méthode présentée ici est appelée le test de Kasiski.

Le principe:

- Déterminer le groupe de lettres qui se répète le plus souvent dans le message chiffré.
- Ensuite, avec l'hypothèse que chaque occurrence de ce groupe de lettres a été chiffrée avec le même morceau de la clé, on déduit que la longueur probable de la clé est un diviseur du pgcd des distances entre ces occurrences.

La méthode présentée ici est appelée le test de Kasiski.

Le principe:

- Déterminer le groupe de lettres qui se répète le plus souvent dans le message chiffré.
- Ensuite, avec l'hypothèse que chaque occurrence de ce groupe de lettres a été chiffrée avec le même morceau de la clé, on déduit que la longueur probable de la clé est un diviseur du pgcd des distances entre ces occurrences.

Simplifications pour la suite du cours :

- On se focalisera sur le triplet qui se répète le plus souvent dans le chiffré.
- On supposera que la longueur de la clé est un nombre premier, ce qui limitera le nombre de valeurs possibles une fois qu'on aura calculé le pgcd.

Exemple

Quelle est la longueur de la clé utilisée pour obtenir le chiffré suivant :

mbtgmrgylrmbtntynfsydskxdmbtzfk

Exemple

Quelle est la longueur de la clé utilisée pour obtenir le chiffré suivant :

mbtgmrgylrmbtntynfsydskxdmbtzfk

- Le triplet le plus fréquent est : mbt
- En mettant ce triplet en couleur dans le chiffré, on obtient :

mbtgmrgylrmbtntynfsydskxdmbtzfk

Exemple

Quelle est la longueur de la clé utilisée pour obtenir le chiffré suivant :

mbtgmrgylrmbtntynfsydskxdmbtzfk

- Le triplet le plus fréquent est : mbt
- En mettant ce triplet en couleur dans le chiffré, on obtient :

$\color{red}mbt gmrgylrmbt ntynfsydskxdmbtzfk$

- Les distances entre ces occurrences de **mbt** sont : 10 (pour la première et la deuxième), 25 (pour la première et la troisième) et 15 (pour la deuxième et la troisième).
- La longueur probable de la clé est un diviseur premier de pgcd(10,15,25). Donc, 5 ici.

On suppose ici que la longueur / de la clé est connue.

Le principe :

 On commence par découper le message chiffré en *I* sous-textes : pour 0 ≤ *k* < *I*, chaque sous-texte *k* est la suite des lettres qui sont aux positions *j*, telles que *k* = *j* mod *I*.

Note : Avec ce découpage, le k-ième sous-texte correspond à un message chiffré avec le Code de César en utilisant le k-ième caractère de la clé.

• Ensuite, on utilise l'analyse fréquentielle pour déterminer les différents caractères de la clé à partir des sous-textes correspondants.

Exemple:

On reprend le texte chiffré de l'exemple précédent :

mbtgmrgylrmbtntynfsydskxdmbtzfk

On sait que la clé est de longueur 5.

Exemple:

On reprend le texte chiffré de l'exemple précédent :

mbtgmrgylrmbtntynfsydskxdmbtzfk

On sait que la clé est de longueur 5.

Les cinq sous-textes correspondants :

• Sous-texte[0] = "mrmydmk" (mbtgmrgylrmbtntynfsydskxdmbtzfk)

 $\bullet \ \, \mathsf{Sous\text{-}texte}[1] = \text{"}\,\mathsf{bgbnsb"} \qquad \quad \, \big(\mathsf{mbtgmrgylrmbtntynfsydskxdmbtzfk}\big)$

 $\bullet \ \, \mathsf{Sous\text{-}texte}[2] = \text{"tytfkt"} \qquad \qquad \mathsf{(mbtgmrgylrmbtntynfsydskxdmbtzfk)}$

 $\bullet \ \, \mathsf{Sous\text{-}texte}[3] = \text{"}\,\mathsf{gInsxz"} \qquad \qquad \mathsf{(mbtgmrgylrmbtntynfsydskxdmbtzfk)}$

 $\bullet \ \, \mathsf{Sous\text{-}texte[4]} = \text{"mrtydf"} \qquad \quad \, \big(\mathsf{mbtgmrgylrmbtntynfsydskxdmbtzfk}\big)$

Exemple:

On reprend le texte chiffré de l'exemple précédent :

mbtgmrgylrmbtntynfsydskxdmbtzfk

On sait que la clé est de longueur 5.

Les cinq sous-textes correspondants :

- $\bullet \ \, \mathsf{Sous\text{-}texte}[0] = "\mathsf{mrmydmk}" \quad \, \big(\mathsf{mbtgmrgylrmbtntynfsydskxdmbtzfk}\big)$
- $\bullet \ \ \mathsf{Sous\text{-}texte}[1] = \text{"} \mathsf{bgbnsb"} \qquad \qquad \mathsf{(mbtgmrgylrmbtntynfsydskxdmbtzfk)}$
- $\bullet \ \, \mathsf{Sous\text{-}texte}[2] = \text{"tytfkt"} \qquad \qquad \mathsf{(mbtgmrgylrmbtntynfsydskxdmbtzfk)}$
- $\bullet \ \, \mathsf{Sous\text{-}texte}[3] = \text{"}\,\mathsf{gInsxz"} \qquad \quad \, \big(\mathsf{mbtgmrgylrmbtntynfsydskxdmbtzfk}\big)$
- La valeur de la clé est :
- Le message clair est :

Chiffre de Vigenère : exercice

Le cryptogramme ci-dessous a été obtenu avec le Chiffre de Vigenère. Déterminer la clé qui a été utilisée ainsi que le message clair associé à ce cryptogramme.

Le cryptogramme :

pbcjbjcrgwisiixrbrgxwquplcptgfiaeeamcftxqbnvisiyceteadxxfqwxgbsvnrzhky joeoxrqpckwrbummquvbskmcbwrngfijenxxqrgwyjeuledeeeefrrxyqswbzxnvle moubxfopweksexpriebiqrgkijpntgbertvaeueiqttxwaihyioepmipcqgxoakkijepm ernurwqeoxhbcjbjcrgfiktohrlanilxbgmmnugvsjmgeizhkyjoefxgbsckurinnxflkli zerxraapmglmoxglmrhwxnvnraeetpxgg