CRYPTO

Introduction à la cryptographie

Loïc Rouquette

2025-02-10

Présentation du cours

Les enseignants

- Cours / TD / TP : Loïc Rouquette loic.rouquette@epita.fr

Sources

- Cours de l'Université de Lorraine
- Introduction à la Cryptographie

Objectifs du Cours

Acquis d'apprentissage visés

À l'issue de ce cours vous serez capables de :

- **Utiliser** des primitives cryptographiques ;
- **Citer** les différentes catégories d'outils cryptographiques, leurs cas d'application et les standards associes ;
- **Employer** les outils cryptographiques standards ;
- **Identifier** des constructions cryptographiques faibles face aux ordinateurs quantiques ;
- **Reformuler** les 5 principes de la sécurité informatiques.

Objectifs du Cours(ii)

Modalités d'évaluation des connaissances

- QCMs (15%)
- TP (35%)
- Devoir Surveillé (Examen final) (50%)

Durant le cours

Ce que nous allons voir

- Les fonctions de hashage
- Les chiffrements symétriques
- Les chiffrements asymétriques
- Les générateurs pseudo-aléatoires
- Certains protocoles et les infrastructure à clé publiques
- Une initiation aux chiffrements post-quantique
- La construction des chiffrements post-quantique

Ce que nous n'allons pas voir

- La cryptanalyse avancée
- Le fonctionnement complet des chiffrements post-quantique
- La sécurité informatique dans le cadre général

Plan du cours

Partie 1

CM : Introduction

TD : Fonction de hashage

TP : Chiffrements historiques

Partie 2

CM : Cryptographie symétrique et PRNGs

TD :

TP :

Partie 3

CM : Cryptographie asymétrique

TD : Logarithme discret & décomposition en facteurs premiers

TP : RSA : implémentation et attaques

Partie 4

CM : Post-Quantique, Protocoles & PKI

TD :

TP :

Wooclap !

Figure 1. – https://app.wooclap.com/ESMECRYPTO1

Introduction à la sécurité informatique

Positionnement du problème

Caractéristiques des **Systèmes d'information** :

- Information numérique ;
- Communication par **canal public** ;
- Machines reliées par réseau ;
- Multi-utilisateurs.

Que voulons-nous protéger ?

> ⚠ **Avertissement**
>
> Il ne faut pas confondre **information** et **donnée** !

Donnée

Une **donnée** est la représentation d'une **information** sous une forme conventionnelle destinée à faciliter son traitement.

- On sait faire subir un traitement à une donnée ;
- Une **donnée** peut être porteuse de plusieurs **informations**, même involontaires (dépend du **contexte** que peut connaître l'attaquant) ;
- Protéger une **donnée** ne suffit pas toujours à protéger l'**information**.

Les besoins génériques de la sécurité

wooclap

Les besoins génériques de la sécurité(ii)

Les 5 grands principes de la sécurité informatique :

Disponibilité

Dans le domaine de l'ingénierie de fiabilité, la **disponibilité** d'un équipement ou d'un système est une mesure de performance obtenue en divisant la durée pendant laquelle ledit équipement ou système est opérationnel par la durée totale pendant laquelle on aurait souhaité qu'il le soit.

wooclap

Les besoins génériques de la sécurité(iii)

Intégrité

L'**intégrité** des données est l'assurance que les données de l'entreprise sont *exactes*, *complètes* et *cohérentes* tout au long de leur cycle de vie.

wooclap

Les besoins génériques de la sécurité(iv)

Confidentialité

La **confidentialité** est le fait de s'assurer que l'information n'est accessible qu'à ceux dont l'accès est autorisé.

wooclap

Les besoins génériques de la sécurité(v)

Non-répudiation

Aucun utilisateur ne doit pouvoir contester les opérations qu'il a réalisées dans le cadre de ses actions autorisées et aucun tiers ne doit pouvoir s'attribuer les actions d'un autre utilisateur.

wooclap

Les besoins génériques de la sécurité(vi)

Authentification

L'**authentification** est le processus visant à confirmer qu'un commettant est bien légitime pour accéder au système. Il existe quatre facteurs d'authentification classiques qui peuvent être utilisés dans le processus d'authentification d'un commettant.

- utiliser une information que seul le commettant connaît (ce que l'on connaît) ;
- utiliser une information unique que seul le commettant possède (ce que l'on possède) ;
- utiliser une information qui caractérise le commettant dans un contexte donné (ce que l'on est) ;
- utiliser une information que seul le commettant peut produire (ce que l'on sait faire).

wooclap

Les besoins génériques de la sécurité(vii)

> ⚠ **Avertissement**
>
> Il ne faut pas confondre **authentification** et **identification**.

L'**identification** est le processus de reconnaissance d'un utilisateur. Il s'agit essentiellement d'une déclaration d'identité, c'est-à-dire qu'un utilisateur déclare « C'est qui je suis ».

Qu'est-ce que la cryptographie ?

La Cryptographie

Étymologie : crypto- (du grec ancien $\kappa\rho\upsilon\pi\tau\sigma\varsigma$, « caché », « secret ») et -graphie (du grec ancien $\gamma\rho\alpha\varphi\epsilon\iota\nu$, « secret »).

Art d'écrire en language codé, secret, chiffré.

– Dictionnaire de l'Académie française, 9ème édition.

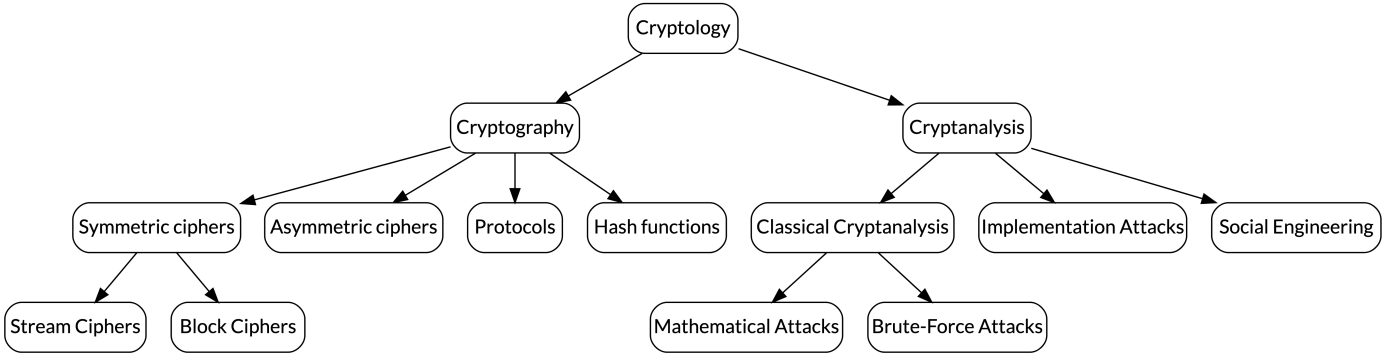
De nombreuses applications dans la vie courante :

- SSL/TLS (https), SSH, GnuPG, etc.
- Carte bleue, téléphonie cellulaire, wifi, bluetooth
- etc.

Cryptologie

Cryptologie

Étude de la **protection** (algorithmique) **de l'information sous forme numérique** contre l'accès ou les manipulations non-autorisées.



Overview on cryptology C. Paar et J. Pelzl [1].

Histoire de la cryptographie

Cryptographie artisanale

Antiquité - XIXème siècle

Date	
Ier siècle a.v. J.C.	Chiffrement de César
1586	Chiffrement de Vigenère

Constructions par **permutations** et **substitutions**

Cryptographie mécanique

XIXème siècle - milieu du XXème siècle

Date	
1883	La Cryptographie Militaire
1914 - 1918	Première guerre mondiale
1926	Cipher Printing Telegraph Systems for Secret Wire and Radio Telegraphic Communications
1939 - 1945	Seconde guerre mondiale, Enigma et les « bombes » de Bletchley Park, machines de Lorenz
1950 - 1960	Machines Hagelin de Boris Hagelin, Crypto AG

Cryptographie industrielle

milieu du XXème siècle - maintenant

Date	
1949	Communication Theory of Secret Systems
	• Introduction de la sécurité inconditionnelle
1973 - 1977	Standardisation de Data Encryption Standard (DES)
1976	New Directions in Cryptography
	• Création de la cryptographie à clé publique
1978	A Method for Obtaining Digital Signatures and Public-Key Cryptosystems
	• Invention de RSA

Cryptographie industrielle(ii)

Date	
1997 - 2000	Standardisation de l'Advanced Encryption Standard (AES)
2007 - 2012	Standardisation de Secure Hash Algorithm 3 (SHA-3)
2017 - 2024	Standardisation des chiffrements post-quantiques Module-Lattice-Based Key Encapsulation Mechanism (ML-KEM) et Module-Lattice-Based Digital Signature Algorithm (ML-DSA)

Les modèles de Communication

Modèle de Communication

Figure 2. – Modèle d'un système de communication général

Modèle de Communication(ii)

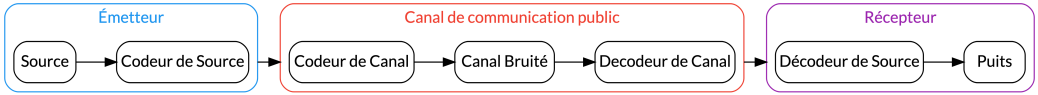


Figure 3. – Modèle d'un système de communication sans bruit

Modèle de Communication(iii)

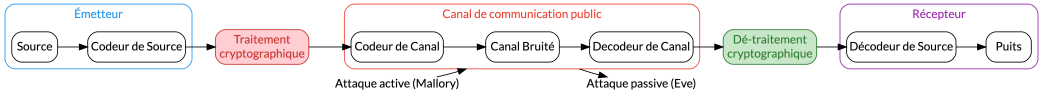


Figure 4. – Modèle d'un système de communication cryptographique

Modèle de Communication(iv)

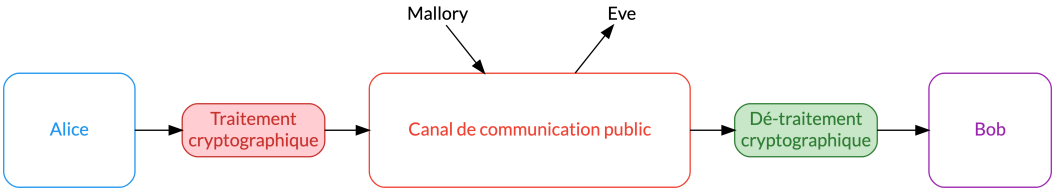


Figure 5. – Modèle d'un système de communication cryptographique

Différentes menaces

Une attaque peut être

- **passive** : espionnage ;
- **active** :
  - Usurpation d'identité ;
  - Altération des données ;
  - Répudiation des messages ;
  - Répétition des messages ;
  - Retardement de la transmission ;
  - Destruction des messages.

Les réponses cryptographiques

La Cryptographique : un empilement de couches

Table 1. – Représentation des différentes couches de la cryptographie.

<table>
<tr><td colspan="4">Mise en œuvre globale</td></tr>
<tr><td rowspan="4">Lois</td><td>Protocoles</td><td rowspan="4">Implémentations</td><td rowspan="4">Matériels</td></tr>
<tr><td>Mécanismes Cryptographiques</td></tr>
<tr><td>Modes / Paddings / Encapsulations</td></tr>
<tr><td>Primitives</td></tr>
</table>

Les mécanismes fondamentaux

Algorithmes fournissant une fonctionnalité cryptographique fondamentale :

- Contrôle d'intégrité cryptographique : **fonction de hashage** ;
- Génération des clés, de vecteur d'initalisation (Init Vector, IV), d'aléa : **générateur d'aléa cryptographique** ;
- Authentification de l'origine des messages : **Code d'Authentification de Message (MAC)**, **algorithmes de signature** ;
- Confidentialité : **chiffrements**.

Cas n°1 : Intégrité des données

Améliorer un canal authentifié

Contrôle d'intégrité avec fonction de hashage

- Un canal public pour transmettre des messages de grande taille ;
- Un canal authentifié pour transmettre un **contrôle d'intégrité** de petite taille.

Exemple

Télécharger depuis un site mirroir, calculer le haché du fichier, vérifier qu'il correspond au haché sur le site officiel, canal « authentifié » via la confrance dans les DNS et leur sécurité.

Améliorer un canal authentifié(ii)

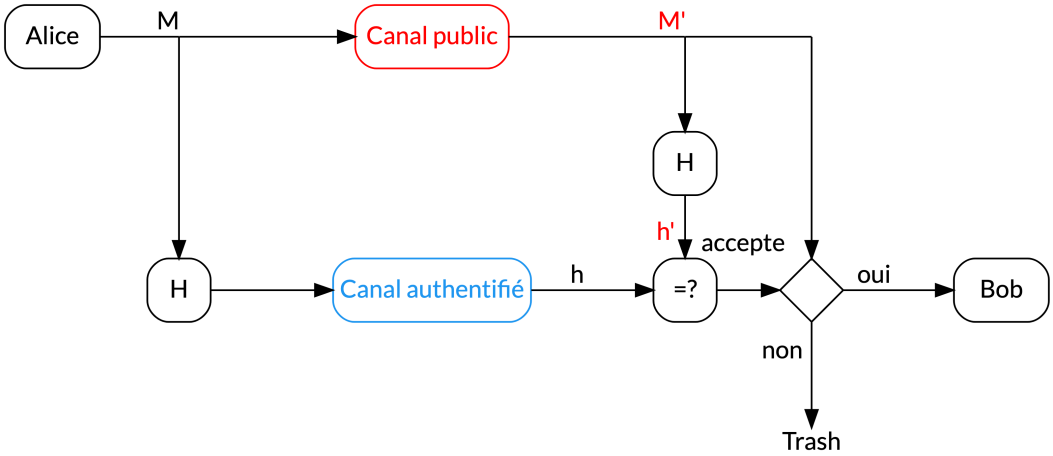


Figure 6. – Canal d'authentification amélioré.

Fonction de hashage

Définition

Une **fonction de hachage** est un algorithme (*efficace*) qui calcule une **valeur de taille fixe**, appelée **empreinte** ou **haché**, à partir de **messages de taille quelconque**.

Plus formellement une fonction de hashage est définie par :

$$H : \{0,1\}^* \rightarrow \{0,1\}^n, H(m) = h$$

Fonction de hashage cryptographique

Pour que H soit qualifiée de cryptographique, il faut que H résiste aux attaques par calcul de :

- **premier antécédant** : étant donné y, il est difficile de trouver x tel que $H(x) = y$;
- **deuxième antécédant** : étant donnés x et $H(x)$, il est difficile de trouver $x' \neq x$ tel que $H(x\prime) = H(x)$;
- **collision** : il est difficile de trouver x et x' tel que $H(x) = H(x\prime)$.

La pertinance des attaques dépend des applications.

Le contrôle d'intégrité est assuré sur le canal authentifié par la vérification de l'égalité des hachés.

Sécurité calculatoire

Les systèmes utilisés dans la pratique sont **théoriquement cassables**.

Sécurité pratique ou calculatoire

Un attaquant possédant les spécifications de l'algorithme, autant de données que possible à sa disposition et une grande puissance de calcul ne peut pas casser le système cryptographique en un **temps humainement raisonnable**.

Complexité d'une attaque : ordres de grandeur

$$Complexité = \mathcal{O}(Temps + Mémoire + Données)$$

Table 2. – Exemple de complexités.

n	$2^n = 10^x$	**Exemples**
32	$2^{32} = 10^{9.6}$	Nombre d'êtres humains sur Terre
46	$2^{46} = 10^{13.8}$	Distance Terre – Soleil en millimètres
		Nombre d'opérations effectuées par **jour** par un ordinateur à 1 Ghz
55	$2^{55} = 10^{16.6}$	Nombre d'opérations effectuées par **an** par un ordinateur à 1 Ghz
82	$2^{82} = 10^{24.7}$	Masse de la Terre en kilogrammes
90	$2^{90} = 10^{27.1}$	Nombre d'opérations effectuées en **15 milliards d'années** par un ordinateur à 1 Ghz
155	$2^{155} = 10^{46.7}$	Nombre de molécules d'eau sur Terre
256	$2^{256} = 10^{77.1}$	Nombre d'électrons dans l'univers

Sécurité des fonctions de hashage

Les collisions sont **inévitables**, on veut qu'elles soient **difficiles à trouver** !

Exemples :

- MD5 (128 bits) ;
- SHA-1 (160 bits) ;
- SHA-2 (224, 256, 384 et 512 bits) ;
- SHA-3 (224, 256, 384 et 512 bits).

Attaque générique par paradoxe des anniversaires

S'il y a 2^l valeurs de hashés possibles, on trouve une collision avec une probabilité supérieure à $1/2$ dès qu'on teste plus de :

$$\sqrt{\frac{2^l \pi}{2}} = \mathcal{O}\left(2^{\frac{l}{2}}\right) \text{ entrées aléatoires}$$

Exemples de complexités et standards

- 2^{64} pour MD5 -> Trop faible
- 2^{80} pour SHA-1 -> Trop faible

Compétition SHA-3 du National Institute of Standards and Technology (NIST)

- 24 janvier 2007 : appel à propositions pour les fonctions de hachage ;
- 2 octobre 2012 : sélection de Keccak pour le nouveau standard ;
- 5 août 2015 : publication du standard SHA-3 FIPS 202

Programme des séances pratiques (TD + TP)

TD

Complexité et fonctions de hashage

TP

Les chiffrements historiques, implémentation et cryptanalyse

Bibliographie

[1] C. Paar et J. Pelzl, *Understanding Cryptography: A Textbook for Students and Practitioners*. Springer Berlin Heidelberg, 2010. doi: 10.1007/978-3-642-04101-3.