La Cryptographie: Art, Science ou Magie...

David Pointcheval

CNRS/Ecole normale supérieure

19 février 2018

Outline

Outline

Sécurité des communications

On a toujours voulu échanger de façon confidentielle

Mais dans ce monde du “tout numérique”,
les moindres chuchotements sont interceptés

Sécurité des communications

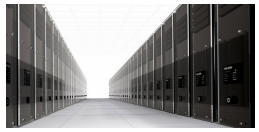
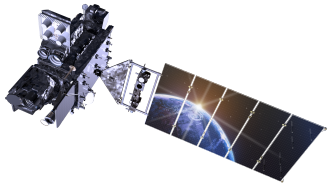
La cryptographie s'est donc immiscée partout, à notre issue

Dans votre poche



A la maison


Dans les grands systèmes

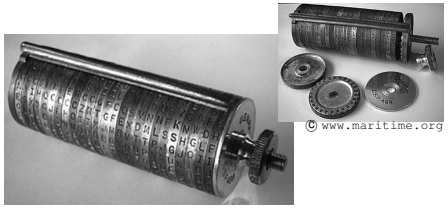
Premiers mécanismes de chiffrement

Le but du chiffrement est de cacher le message

Scytale (V^e siècle av. J.-C.)
Permutation

Substitutions et permutations
La **sécurité** repose sur
le secret du mécanisme

Disque d'Alberti (XV^e siècle)
Substitution mono-alphabétique



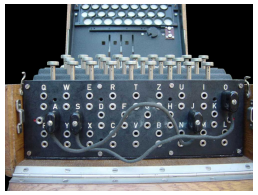
Cylindre de Jefferson (XVIII^e siècle)
Substitution poly-alphabétique

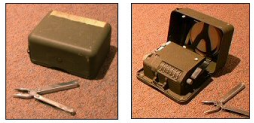
Utilisation d'un secret partagé : clé secrète

Une information partagée (clé secrète) entre l'émetteur et le récepteur sert de paramètre au mécanisme de chiffrement

Enigma (Sept. 1918) – Des variantes seront utilisées par l'armée allemande pendant la seconde guerre mondiale

Attaques statistiques par Alan Turing

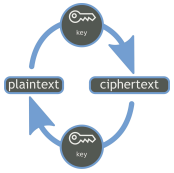



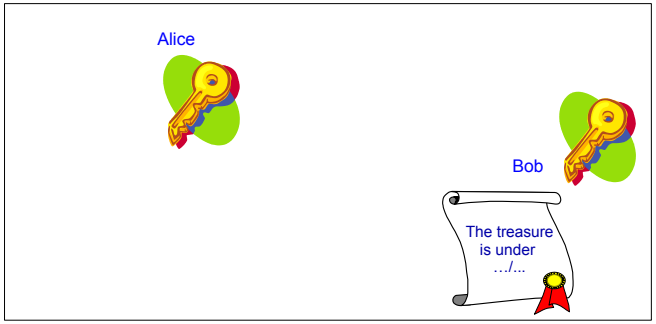


HAGELIN M-209 CIPHER MACHINE (GVG / PD)


Hagelin (Oct. 1919)

Hagelin M-209 sera utilisée par l'armée américaine

Chiffrement symétrique : clé commune partagée





Chiffrement symétrique

- Alice et Bob partagent une clé secrète commune
- Sécurité raisonnable, mais *heuristique*
- Grande efficacité

Mais comment (ou pourquoi) partager un secret préalable ?

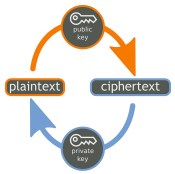
Chiffrement asymétrique : intuition

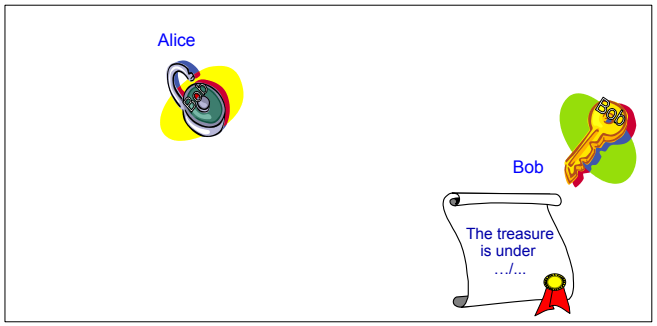
Confidentialité (Diffie-Hellman – 1976)

- Seul le destinataire peut prendre connaissance du contenu
- Aucune contrainte sur l'émetteur

Pourquoi avoir besoin d'un secret pour chiffrer un message ?

Diffie et Hellman, Prix Turing 2015





Chiffrement RSA

RSA = Rivest-Shamir-Adleman – 1978

- Clés de Bob
  - Clé publique : $n = p \times q$, exposant e
  - Clé privée : p et q
- Algorithmes
  - Chiffrement pour Bob : $c = m^e \bmod n$
  - Décryptement de c : racine e-ième modulaire
    Difficile sans la factorisation du module
  - Déchiffrement par Bob : racine e-ième modulaire
    Facile avec la factorisation du module

Rivest, Shamir et Adleman, Prix Turing 2002

Recommandations : 310 chiffres (soit 1024 bits) strict minimum
voire 617 chiffres (soit 2048 bits) pour 10 ans

Cryptographie asymétrique : garanties apportées

Chiffrement = confidentialité de données

Nul ne peut apprendre un bit d'information sur m à la vue de c,
même s'il a pu demander le déchiffrement de tout chiffré $c' \neq c$

Signature = authentification de données

Nul ne peut générer une nouvelle signature valide,
même s'il a pu demander les signatures de messages de son choix

Sécurité prouvée

- Si un adversaire peut mettre en défaut ces notions
- Alors on peut résoudre un problème difficile sous-jacent
  tel que le problème de la factorisation

Outline

Respect de la vie privée : besoins

Confidentialité des données

Seules les personnes légitimes ont accès aux données
En contradiction avec l'**intégrité des données** ?

Anonymat des individus

Nul ne sait qui est connecté
En contradiction avec le **contrôle d'accès** ?

Secret des requêtes

Nul ne connaît les questions posées
En contradiction avec la **pertinence des réponses** ?

Comment répondre à des questions ?

Contrôle d'accès et anonymat

Respect de la vie privée

Pourquoi fournir son identité pour accéder à un service ?

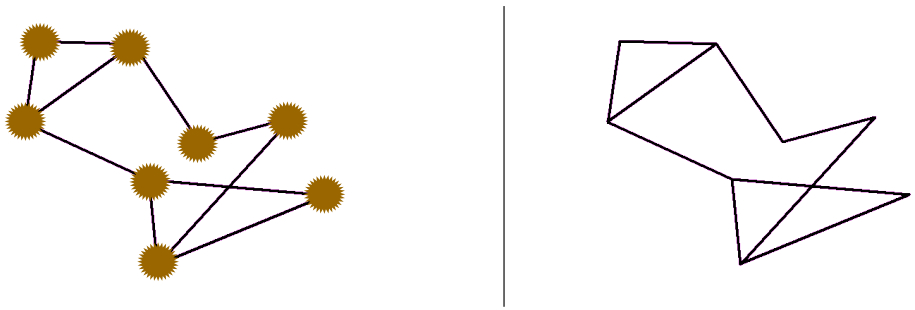
Prouver que l'on a droit à cet accès est suffisant

- Inscription : le service remet un certificat personnel (et secret) de droit d'accès
- Accès au service : on prouve la possession/connaissance d'un certificat

Peut-on prouver sa connaissance d'un certificat secret sans révéler ce certificat (anonymat) ?

Preuves *Zero-Knowledge*

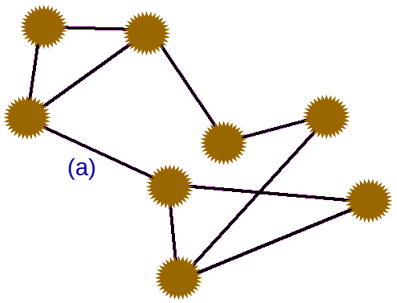
Comment prouver que je connais ce 3-coloriage, sans le révéler ?



Je choisis une permutation sur les couleurs
et l'applique aux sommets
Puis je masque les sommets

Preuves *Zero-Knowledge*

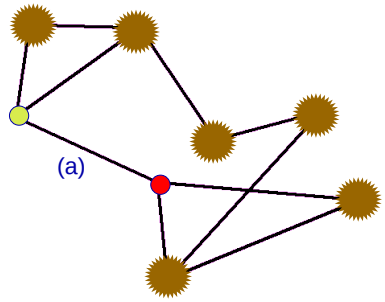
Comment prouver que je connais ce 3-coloriage, sans le révéler ?



Je communique le graphe masqué au vérifieur
Le vérifieur désigne une arête (2 sommets adjacents)

Preuves *Zero-Knowledge*

Comment prouver que je connais ce 3-coloriage, sans le révéler ?



J'enlève les masques
Le vérifieur vérifie qu'il y a 2 couleurs différentes
Il est alors convaincu...

Preuves *Zero-Knowledge*

Preuves de connaissance *Zero-Knowledge*
Goldwasser-Micali-Rackoff –1985

Ces protocoles garantissent que

- si on ne connaît pas de solution, la probabilité d'être accepté après plusieurs itérations est négligeable
- le vérifieur n'apprend aucune information

Goldwasser et Micali, Prix Turing 2012

Applications

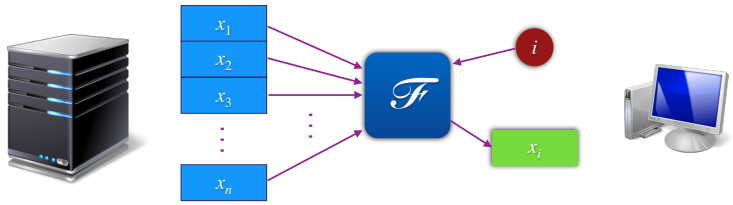
Ces preuves ont de nombreuses applications à l'anonymat :

- signature de groupe
- contrôle de droit d'accès anonyme
- vote électronique universellement vérifiable
- etc...

Interrogation confidentielle de bases de données

Respect de la vie privée

Pourquoi révéler nos questions pour avoir des réponses ?



Chiffrement homomorphe

- A partir de chiffrés de m_1 et m_2 $\rightarrow$ chiffré de $m_1 + m_2$
- A partir de chiffrés de m_1 et m_2, et des scalaires a, b $\rightarrow$ chiffré de $a \times m_1 + b \times m_2$

$$\mathcal{E}(x_i) = \sum_j x_j \times \mathcal{E}(q_j), \text{ où } q_i = 1 \text{ et } q_j = 0 \text{ pour } j \neq i$$

Exécution confidentielle d'un programme

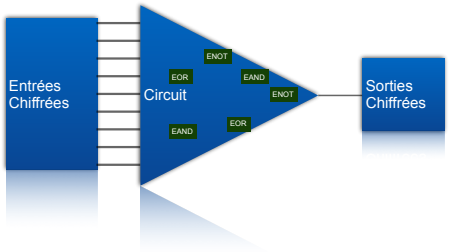
Respect de la vie privée

Pourquoi révéler nos questions pour avoir des réponses ?

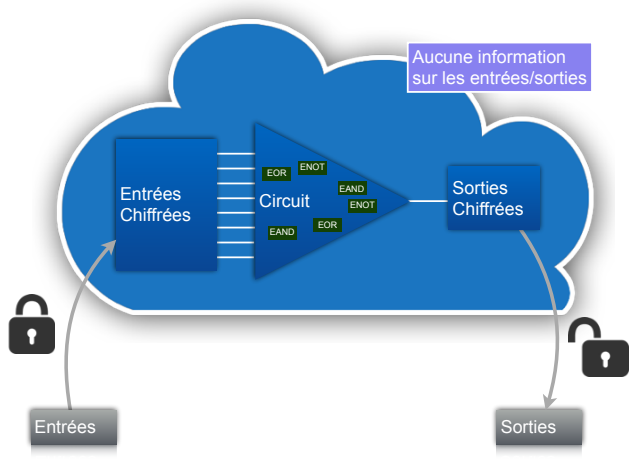
Pourquoi fournir ses entrées en clair à un programme ?

Chiffrement complètement homomorphe **Gentry – 2009**

- A partir de chiffrés de m_1 et m_2 $\rightarrow$ chiffré de $m_1 + m_2$
- A partir de chiffrés de m_1 et m_2 $\rightarrow$ chiffré de $m_1 \times m_2$



Calculs externalisés sur des données secrètes



Données “manipulées” **chiffrées** et résultat retourné **chiffré**
⇒ seul l’utilisateur légitime peut retrouver le résultat en clair

Permet des recherches “**google**” sans révéler la question !

Diffusion contrôlée de données agrégées

L’accès à une base de données chiffrées doit contrôler les calculs pour éviter la diffusion de données personnelles
⇒ contrôle d’accès “by design”

Student Name	English		CS		Math	
	Written	Spoken	Theory	Practice	Algebra	Analysis
Year 1						
Year 2						
Year 3						

Class	English	CS	Math
Year 1			
Year 2			
Year 3			

Class	English		CS		Math	
	Written	Spoken	Theory	Practice	Algebra	Analysis
Total						

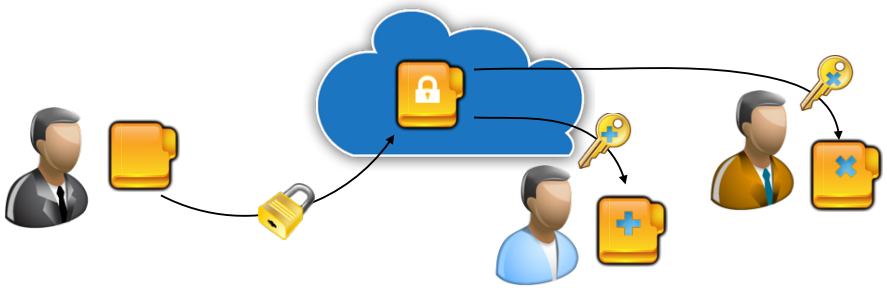
Class	Total
Year 1	
Year 2	
Year 3	

Class	Total
3Years	

Pour chaque nouvelle case, $v = \vec{x} \cdot \vec{y}$
($\vec{x}$ = vecteur des valeurs initiales, $\vec{y}$ = vecteur des pondérations)

Chiffrement fonctionnel

Chaque clé de déchiffrement n’autorise qu’un déchiffrement agrégé



- des solutions concrètes existent pour des moyennes pondérées
- exploitent le chiffrement *simplement* homomorphe
  (juste additif : Paillier, ElGamal, LWE)
- coût algorithmique raisonnable

Consultations de bases de données chiffrées

Chiffrement fonctionnel — **Boneh-Sahai-Waters – 2011**

Un tel système garantit que

- les utilisateurs n’obtiennent que les clairs agrégés autorisés
- sans besoin de contrôle d’accès interactif
- en protégeant les données personnelles

Applications

Un tel schéma de chiffrement permet :

- l’agrégation de données (calculs statistiques)
- l’interrogation de bases de données chiffrées

Calculs secrets entre plusieurs participants

Plusieurs individus veulent évaluer une fonction sur leurs données secrètes, sans rien apprendre/révéler plus que le résultat

Ex : Problème du Millionnaire Yao, Prix Turing 2000

Outline

Conclusion

La cryptographie permet d'atteindre des objectifs *a priori* **paradoxaux**

- envoyer des données secrètes sur un canal public
- s'authentifier anonymement
- prouver que l'on sait, sans ne rien révéler
- répondre sans connaître la question
- manipuler des données sans les voir
- contraindre la diffusion de données sous forme agrégée
- ...

Respect de la vie privée