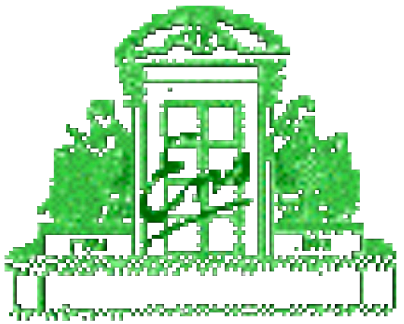
Initiation à la cryptographie

Chiffrement asymétrique et échange de clé

Pierre-Alain FOUQUE
Équipe de Cryptographie
École normale supérieure



Introduction

- Confidentialité: cacher à un tiers le contenu d'une communication ou d'un fichier
  - Notion d'indistinguabilité
  - Bits difficiles et applications au chiffrement
  - Générateur aléatoire
  - Sécurité polynomiale
  - Exemple de schéma
  - Chiffrement hybride
  - Échange de clé

But de chiffrement

- Le chiffrement doit garantir la confidentialité des communications
- On ne veut pas qu'un adversaire puisse apprendre un bit du message clair
- La difficulté d'inverser (OW) ne suffit pas car les fonctions à sens unique ne disent rien si on connaît une partie de l'image (cf. RSA)

Indistinguabilité

- 2 ensembles de distributions de probabilité $(X_n)_{n\in N}$ et $(Y_n)_{n\in N}$ sont indistinguables
  - Parfaitement: $\forall n\in N\ \forall x\in D,\ Pr(X_n=x) = Pr(Y_n=x)$
  - Statistiquement: $\Delta(n) \leq negl(n)$ où
    $$\Delta(n) = (1/2)\times(\Sigma_{x\in D}\ |Pr(X_n=x) - Pr(Y_n=x)|)$$
  - Calculatoirement: $\forall D$ PPT, $\forall p$ poly., $\exists n_0\in N$, $\forall n>n_0$, $|Pr(D(X_n)=1) - Pr(D(Y_n)=1)| < 1/p(n)$

    (indistinguabilité en temps polynomial). Les proba sont prises sur les choix de D et les va.

Bits difficiles

- Tous les bits d'une OWF ne sont pas difficile à calculer
  - $f(x)=g(x_1)||x_2$ ou LSB(x) pour $f(x)=g^x \bmod p$

Définition (Hardcore bit): *Une fonction h de $\{0,1\}^n$ vers $\{0,1\}$ est un bit difficile pour f (OWF) si*

- *h(x) est calculable en temps polynomial*
- *Il n'existe pas d'algo. PPT qui prédit h(x) étant donné f(x) mieux qu'au hasard*

$\forall$PPT A, $\Pr[A(f(x))=h(x) : x \leftarrow \{0,1\}^n] \leq 1/2+\varepsilon(n)$

Bit difficile d'une fonction concrète

- p premier, p=3 mod 4, et g générateur de Z^*_p
- $f(x)=g^x \bmod p$
- MSB(x)=0 si x<(p-1)/2 et 1 sinon
- Rmq: $MSB(x) \neq x_1$ le bit physique de poids fort de x car p n'est pas une puissance de 2
- Th: Si $f(x)=g^x \bmod p$ est une OWF, alors MSB(x) est un bit difficile de f

Bit difficile du Log Discret

Th: S'il existe un PPT algo. A qui calcule toujours de manière correcte MSB(x) à partir de f(x), alors il existe un PPT algo. qui inverse f (calcule le log discret de y)

Rmq: résultat moins fort car ne prédit avec proba>1/2, mais trouve le résultat avec proba=1

Algo: 1. Trouver x_k=LSB(x)

2. Calculer racine carrée de y/g^{xk} mod p

($y_0=g^{[x1...xk-1]}$ et $y_1=g^{(p-1/2)+[x1...xk-1]}$) $A(y_0)$?

3. Trouver le LSB de y_0 etc.

Bit difficile ∀OWF et chiffrement

- Th (Goldreich-Levin): Il existe un bit difficile pour toute OWF f
- Si f est OWF, $g_f(x,r)=f(x)||r$ OWF
- $h(x,r)=\oplus_i x_i.r_i$ avec $|x|=|r|$
- Bob veut envoyer un bit b à Alice
- Alice connaît une TOWP f
- Bob: x aléatoire et envoie $(f(x),h(x)\oplus b)$
- Est-ce qu'on peut extraire plus d'un bit difficile ?

Chiffrement

- Même idée que One Time Password
- On calcule $f^n(x)=y$
- $G'(x)=h(f^{n-1}(x))||h(f^{n-2}(x))...||h(x)$
- Bob envoie $c=(f^n(x), G'(x)\oplus m)$
- Pour Adv.: $p_1=h(f^{n-1}(x))$, …, $p_1=h(x)$ OTP.
  - p_1 aléa car aléa étant donné $f^n(x)$, donc m_1 sûr
  - $p_2=h(f^{n-2}(x))$ aléa étant donné $f^n(x)$ et p_1, car on peut calculer les 2 à partir de $f^{n-1}(x)$ et p_2 est sûr même si $f^{n-1}(x)$ est connu en entier

Chiffrement symétrique

- Sécurité du bit suivant:
  - Étant donné les (i-1) premiers bits de G', calculer le i-ième
- Si G' a cette propriété, on peut construire un chiffrement symétrique
- x est la clé secrète
- $E_x(m)=G'(x)\oplus m$ et le déchiffrement est similaire

Générateur aléatoire

- G fonction déterministe calculable en temps polynomiale $\{0,1\}^k \rightarrow \{0,1\}^{p(k)}$, est un générateur pseudoaléatoire (PRG) si
  - $p(k)>k$
  - $\forall$ PPT D, 1 «pseudorandom», 0 «random»

    $P[D(G(x))=1 : x \leftarrow \{0,1\}^k] - P[D(R)=1 : R \leftarrow \{0,1\}^{p(k)}] < negl(k)$

- Blum-Micali: $x_{i+1}=g^{xi} \bmod p$ et MSB(x)
- Blum-Blum-Shub: $x_{i+1}=x_i^2 \bmod N$ et LSB(x)

Schéma de chiffrement

- $G(1^n)$: PPT algo. de génération de clé
- $E_{pk}(.)$: PPT algo. de chiffrement
- $D_{sk}(.)$: algo. de déchiffrement
- Pour tout $m \in M_n$, $D_{sk}(E_{pk}(m;r))=m$

Sécurité polynomiale

- Un cryptosystème (G,E,D) est IND-sûr (contre adv. CPA) si ∀PPT $B=(B_1,B_2)$

$P[b=b' : (pk,sk)\leftarrow G(1^k);$
$(m0,m1,s)\leftarrow B_1(pk);$
$b\leftarrow\{0,1\};$
$c'\leftarrow E_{pk}(m_b,r);$
$b'\leftarrow B_2(c',s,pk)] \leq 1/2+negl(n)$

- Av(B) = |Pr(b=b')-1/2| avantage de B

Types d'adversaires et sécurité

- Types d'attaques:
  - CPA: attaque à clairs choisis
  - CCA: attaque à chiffrés choisis
- Notions de sécurité:
  - OW-CPA : One-Way CPA
  - IND-CPA :
  - IND-CCA : notion la plus forte

  - OW-CPA < IND-CPA < IND-CCA

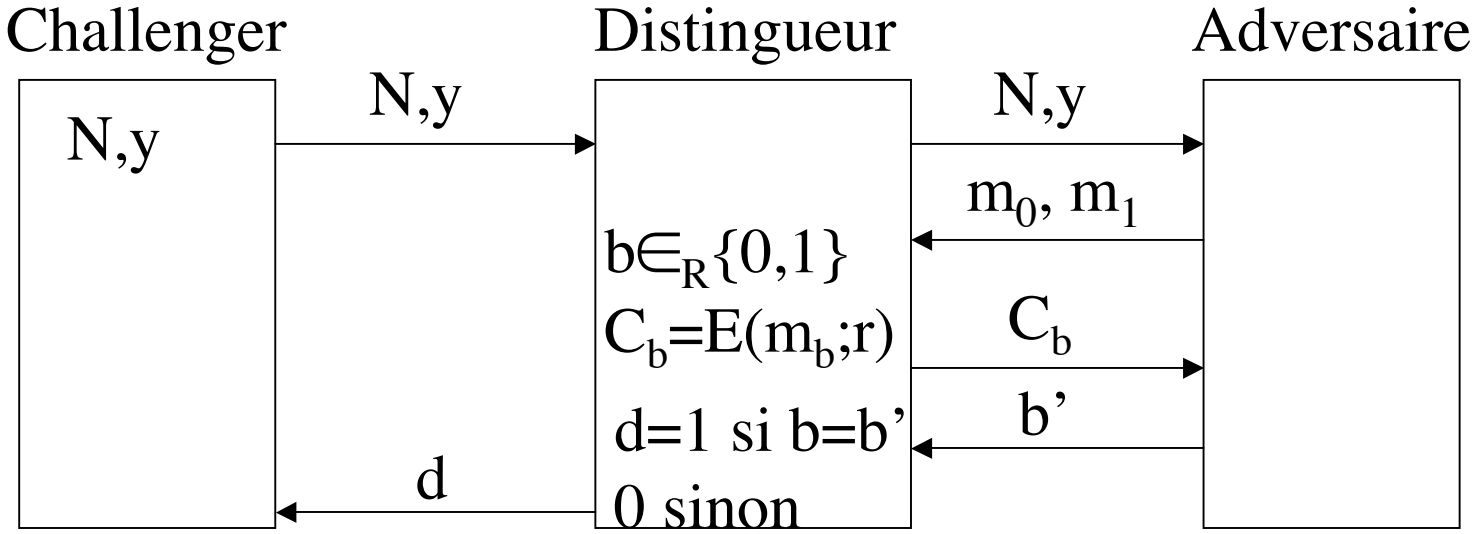
Sécurité Goldwasser-Micali

- Hypothèse de Résiduosité Quadratique: Il est difficile de distinguer des carrés de non carrés pour les éléments x tq (x|N)=1 si N=pq

- Théorème: Sous l'hypothèse de la résiduosité quadratique, le schéma de chiffrement GM est IND-CPA.

Chiffrement Goldwasser-Micali

- pk=(N,y) tq N=pq, (y|N)=1 et (y|p)=-1
- sk =(p,q)
- $E(m;r)=r_1^2 y^{m_1}$ [N]…$r_n^2 y^{m_n}$ [N] si $m=m_1…m_n$, $r=r_1…r_n$
- $D(c) = (c_1|p) … (c_n|p)$ où $c=c_1…c_n$



Analyse de sécurité de GM

- Par l'absurde, supposons que A soit un adversaire avec avantage non négligeable, ε
- $Av(D)=Pr[D=1:y\in NQR]-Pr[D=1:y\in QR]$
- Si $D=1$ correspond à $y\in NQR$.
- $Pr[D=1: y\in NQR] = Pr[b=b': E=GM]$ $=Pr[A \text{ réussit}]=Av(A)+1/2$

Analyse de sécurité de GM (2)

- $Av(D) = Pr[D=1: y \in NQR] - Pr[D=1: y \in QR]$
- $Pr[D=1: y \in NQR] = Av(A) + 1/2$

- $Pr[D=1: y \in QR] = Pr[b=b': E \neq GM] = 1/2$ car dans le cas où $y \in QR$, $E(m_0)=E(m_1)$ car on voit que des carrées et A ne peut donc pas décider s'il s'agit de m_0 ou m_1 !

- $Av(D) = \varepsilon$. D'après l'hypothèse RQ, il n'existe pas de distingueur avec proba non négligeable, donc A n'existe pas.

Schéma de chiffrement ElGamal

- $G(1^n)$: (p,q,g,y,x) où $p=2q+1$ et q sont premiers, g un générateur du groupe G_q d'ordre q et $y=g^x \bmod p$ où x est aléatoire dans Z_q
- $M=G_q$
- $E_{(p,q,g,y)}(m;r)=(g^r,m.y^r)$
- $D_{(p,q,g,x)}(A,B)=B/A^x$
- Pour tout $m \in G_q$, $m.y^r/g^{rx} = m$ car $y^r=(g^x)^r=g^{xr}$

Diffie-Hellman Calculatoire (CDH) et Décisionnel (DDH)

- g générateur du groupe G_q, d'ordre q et sous-groupe de Z^*_p p et q premier tq q|p-1 de taille n
- Hypothèse CDH:

Pour tout PPT (ε,t)-algo. A dépendant de n,

$Pr[A(g^a,g^b)=g^{ab} | a,b \leftarrow Z_q] \leq \varepsilon(n)$

Un triplet aléatoire choisis dans la distribution $\{(g^a,g^b,g^{ab}): a,b \leftarrow Z_q\}$ est dit un *triplet DDH*

Sinon $\{(g^a,g^b,g^c): a,b,c \leftarrow Z_q\}$ est un *triplet aléatoire*

- Hypothèse DDH:

Pour tout PPT (ε,t)-algo. D dépendant de n,

$|Pr[D(g^a,g^b,g^{ab})=1 | a,b \leftarrow Z_q] - Pr[D(g^a,g^b,g^c)=1 | a,b,c \leftarrow Z_q]| \leq \varepsilon(n)$

Chiffrement ElGamal

Th: L'inversion de ElGamal (OW-CPA) est équivalente au CDH

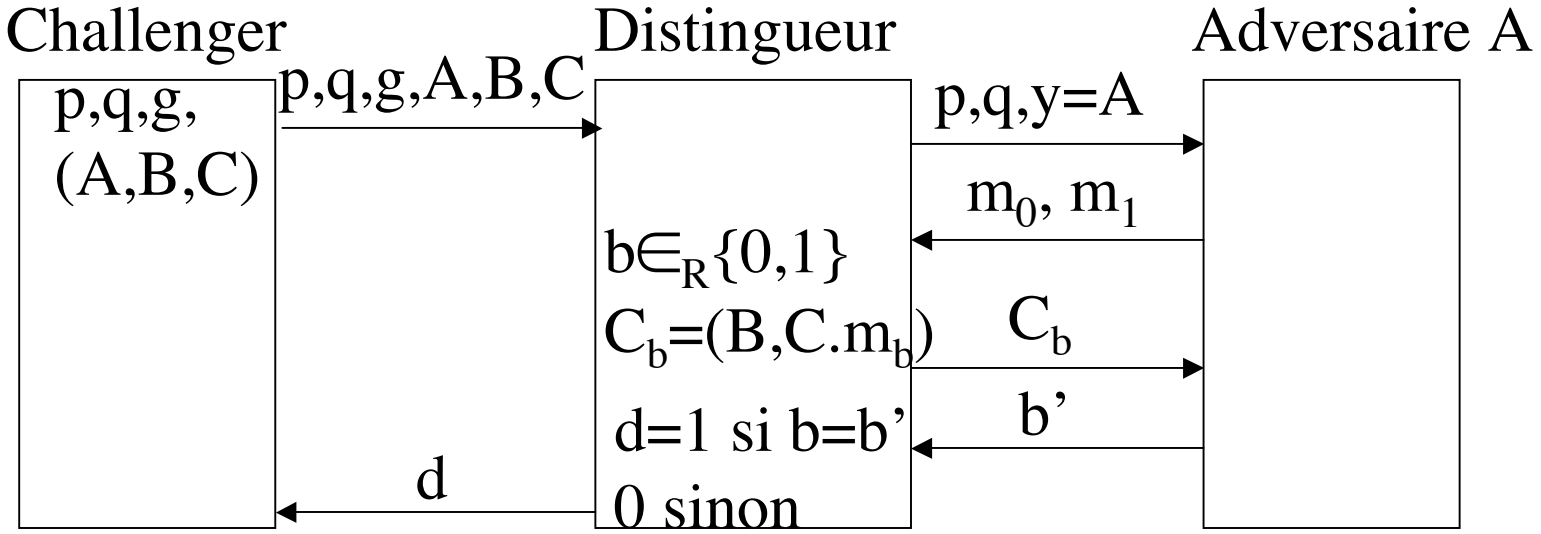
Preuve: Par l'absurde. Supposons qu'il existe un adversaire qui inverse El Gamal, et montrons qu'on peut exhiber un adversaire qui résout CDH.

Sécurité de ElGamal

- Th: Sous l'hypothèse DDH, le schéma de chiffrement El Gamal est sémantiquement sûr contre les attaques à clairs choisis (IND-CPA)

Chiffrement ElGamal

- Si on sait résoudre DDH, alors on peut casser IND-CPA
- Etant donné la clé publique $y=g^a$ et un chiffré (A,B) de m_0 ou m_1, alors $(y,A,B/m_0)$ ou $(y,A,B/m_1)$ est un triplet DDH



Analyse de sécurité d'ElGamal

- Par l'absurde, supposons que A soit un adversaire avec avantage non négligeable, ε
- Av(D) =Pr[D=1: DDH]-Pr[D=1: Rand]
- Si D=1 correspond à (A,B,C) $\in$ DDH.
- Pr[D=1: DDH] = Pr[$\alpha=\alpha'$| E=EG] = Pr[A réussit] = Av(A)+1/2

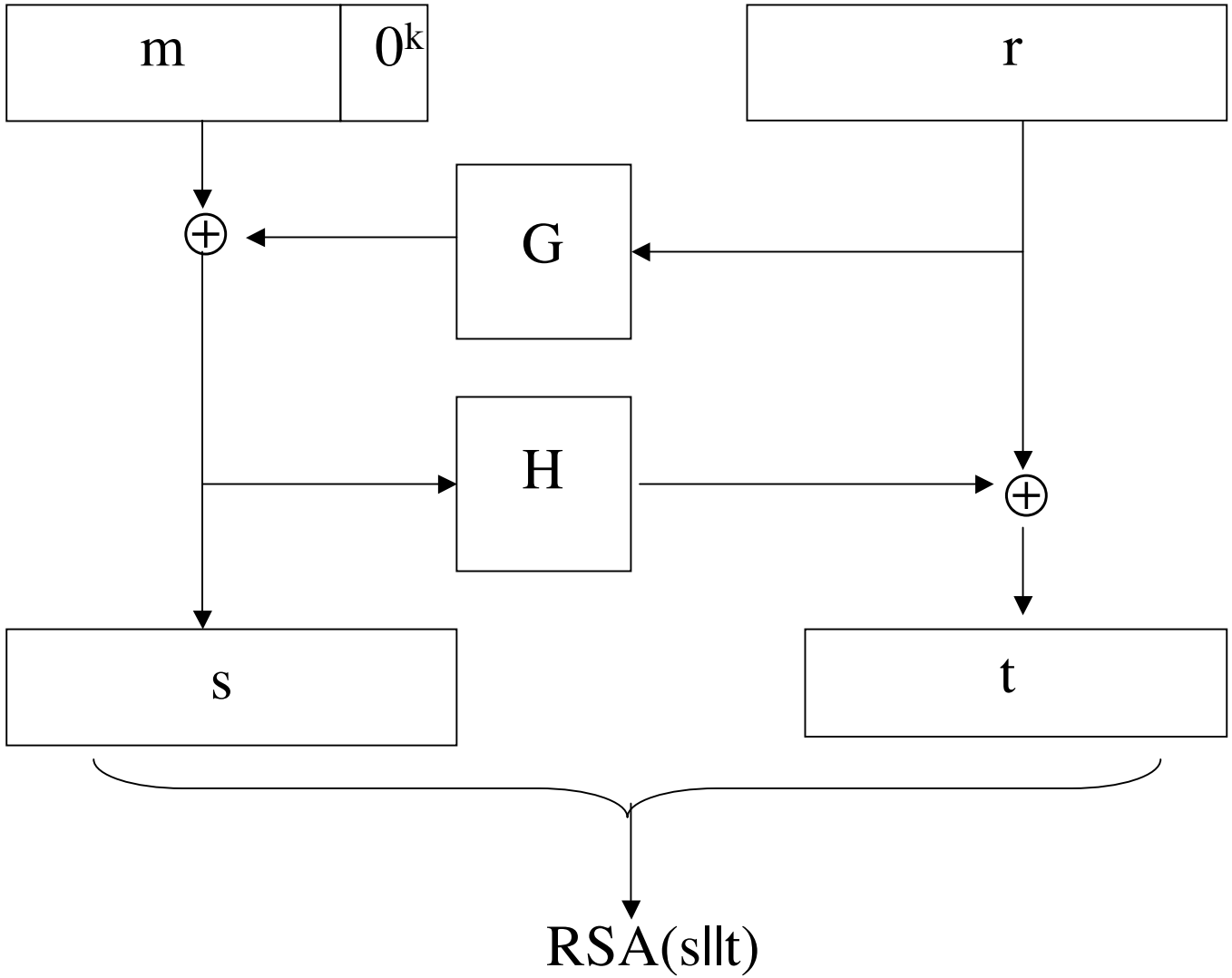
Analyse de sécurité d'ElGamal (2)

- Av(D) =Pr[D=1 : DDH]-Pr[D=1 : DDH]
- Pr[D=1: y∈NQR]= Av(A)+1/2

- Pr[D=1: Rand] = Pr[$\alpha=\alpha'$| E≠GM] =1/2 car dans le cas où C≠CDH(A,B), l'adversaire ne peut pas déterminer le bit α car il ne connaît pas C et ne peut le prédire ! C.m_α représente un OTP avec clé aléatoire.

- Av(D) = ε. D'après l'hypothèse DDH, il n'existe pas de distingueur D avec proba non négligeable, donc A n'existe pas.

RSA-OAEP = PKCS #1 v2.0



Chiffrement hybride

- Avantage des systèmes asymétriques et symétriques

- E est un schéma de chiffrement à clé publique pk
- E' un schéma de chiffrement à clé secrète K
- Chiffrement: $c=(E_{pk}(K;r),E'_K(m))$

Sécurité

Construction: E^{asym} OW-CPA, E^{sym} IND, et G et H fonctions de hachage

$$E_{pk}{}^{hy}(r;m)=(E^{asym}{}_{pk}(r,H(r,m)), E^{sym}{}_{G(r)}(m))$$

Th:

- Si le schéma de chiffrement à clé publique f est OW-CPA et
- si le schéma de chiffrement à clé secrète est OTP,

alors le schéma hybride est IND-CCA dans le modèle de l'oracle aléatoire

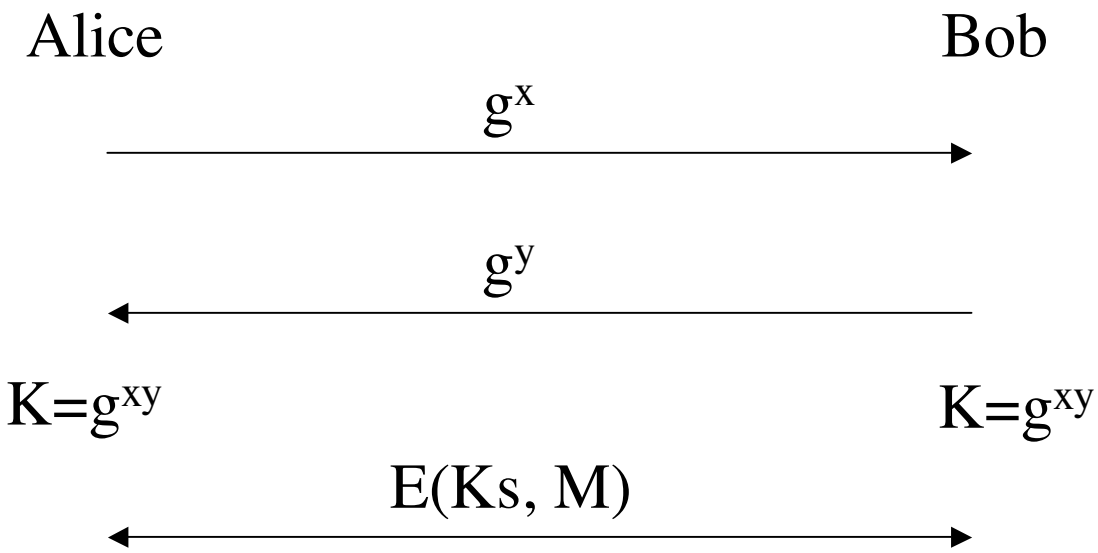
Échange de clé

- Propriétés garanties:
  - Authentification des correspondants
  - Echange d'une clé de session
- Première phase des protocoles de sécurité (SSL,IPSEC, …)
  - Échange de clé avec RSA
  - Échange de clé Diffie-Hellman (forward security)

Primitive DH

- Génération des paramètres
  - p premier avec factorisation de p-1
  - g un géné de Z^*_p d'odre q où q premier et q|(p-1)
  - Clé secrète: $a \in \{2,\ldots,q-1\}$
  - Clé publique: $g^a \bmod p$
- Problème CDH et DDH
  - Etant donné (g,g^a,g^b), calculer $g^{ab} \bmod p$
  - Etant donné (g,g^a,g^b,g^c), décider si $c=ab \bmod q$

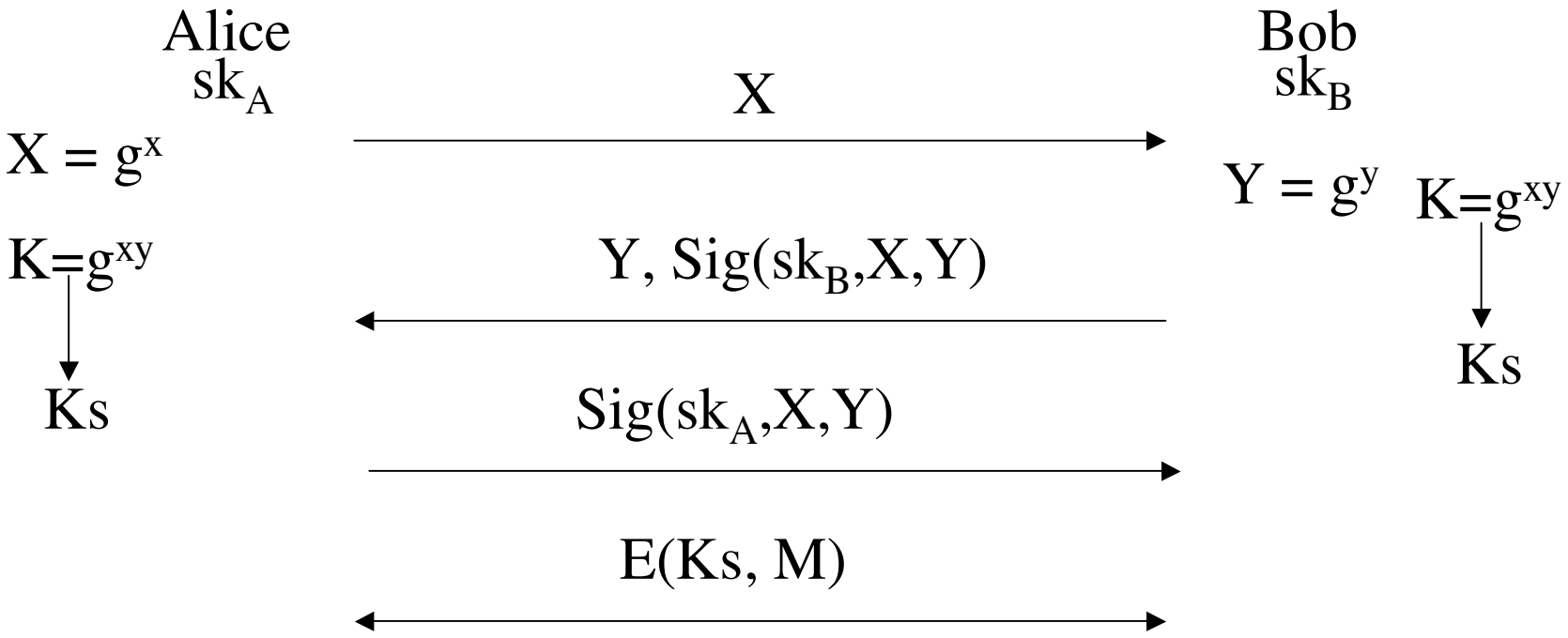
Diffie-Hellman



- Attaque passive : Résoudre le problème CDH
- Indistinguabilité de la clé: DDH
- Attaque active : man-in-the-middle

⇒ Authentifier les flux

Echange de clé authentifié Diffie-Hellman



- Comment passer de K à Ks ? En cryptographie symétrique, on veut que tous les bits de Ks soient aléatoires

Conclusion

- Chiffrement garantit la confidentialité
- Notion forte de confidentialité: on veut cacher tous les bits du message clair
- Schéma prouvé par réduction contre des adversaires forts (CCA)

Modèle de sécurité

- Moyens de l'adversaire:
  - Obtenir des échanges de clés (⇒ adv passifs)
  - Obtenir de vieilles clés (⇒ mauvaise utilisation)
  - Obtenir des flux particuliers (⇒ adv actifs)
- Buts:
  - Sécurité de la clé de session :

    Sur des sessions où la clé n'a pas été révélée, distinguer une clé aléatoire de la clé réelle

  - Authentification (implicite, unilatéral, mutuelle):

    Un adv ne peut pas se faire passer pour un autre utilisateur

Problème

- DDH: difficile de distinguer g^{xy} d'un élément aléatoirement de G
- *DDH ne dit pas que les bits de g^{xy} sont une chaîne de bits aléatoires (clé symétrique)*
- implique «seulement» que log(q) bits sont *calculatoirement indistinguables* dans g^{xy}

Extraction de bits

- Problème de l'extraction de bits :

  Comment choisir des bits aléatoires à partir de g^{xy}

- Si l'ordre du groupe est pair, le bit de poids faible du CDH peut être calculé à partir de g^x et g^y en utilisant le symbole de Legendre
- Oracle aléatoire: $H(g^{xy})$ avec H=SHA-1