Cryptographie à clé publique
Cours 7

Julien Lavauzelle

Université Paris 8

Master 1 mathématiques et applications – parcours ACC

11/03/2024

Plan

Plan

Schémas d'identification : motivation

Objectif d'un schéma d'identification : prouver à l'autre son identité.

Motivations :

- essentiellement pour du **droit d'accès** : à un réseau, à un site web, à des ressources, ou même à un lieu physique (passeport)
- brique de base pour d'autres protocoles (ex : signature)

Essentiellement 3 manières de procéder à une **vérification d'identité**.

1. Par ce que l'on **est**.
   → Exemple : biométrie (empreintes digitales, reconnaissance faciale, etc.)
2. Par ce que l'on **possède**.
   → Exemple : documents d'identité, clés, etc.
3. Par ce que l'on **sait**.
   → Exemple : mots de passe.

Schémas d'identification : définitions

Définition. Un **schéma d'identification** implique un **prouveur** $\mathcal{P}$ et un **vérifieur** $\mathcal{V}$.
Il est constitué d'un **algorithme** de génération de clefs et d'un **protocole** de vérification.

- Lors de la **génération de clefs** KeyGen, le prouveur engendre une paire de clefs publique/privée et émet la clé publique.
- Lors du **protocole de vérification** noté $[\mathcal{P} \longleftrightarrow \mathcal{V}]$, le vérifieur retourne un booléen $b \in \{\mathsf{true}, \mathsf{false}\}$ suivant s'il est convaincu ou non de l'identité du prouveur

Remarque. Contrairement à la signature, on a ici un **protocole** de vérification, c'est-à-dire une échange de données **interactif**.

Définition. On dit qu'un prouveur est **honnête** s'il il possède la clé privée et s'il suit le protocole.

Définition. Un schéma d'identification est **valide** si tout prouveur honnête convainc le vérifieur de son identité avec probabilité 1.

Sauf si le contraire est indiqué : dans ces slides, Alice = prouveur et Bob = vérifieur.

Schémas d'identification : modèle de sécurité

Pour la **sécurité**, le but d'une attaque est l'**imposture**, *i.e.* se faire passer pour le prouveur $\mathcal{P}$ auprès d'un vérifieur.

Il y a différents modèles d'attaques, mais tous se caractérisent par deux étapes.

1. **Étape d'apprentissage.** L'attaquant observe des échanges entre un prouveur $\mathcal{P}$ et un vérifieur $\mathcal{V}$ lors de plusieurs itérations du protocole de vérification $[\mathcal{P} \longleftrightarrow \mathcal{V}]$.
2. **Étape d'imposture.** L'attaquant produit une simulation de prouveur $\widetilde{\mathcal{P}}$. L'attaque est réussie si, lors d'une exécution de $[\widetilde{\mathcal{P}} \longleftrightarrow \mathcal{V}']$ avec un vérifieur $\mathcal{V}'$, la probabilité que le vérifieur $\mathcal{V}'$ soit convaincu d'interagir avec $\mathcal{P}$ est non-négligeable.

Selon l'activité de l'attaquant dans l'**étape d'apprentissage**, on différencie deux moyens d'attaque :

- Les attaques **passives** sont celles où l'attaquant ne fait qu'observer les échanges entre le prouveur et un vérifieur externe.
- Les attaques **actives** sont celles où l'attaquant est également autorisé à jouer le rôle d'un vérifieur (il va donc « conduire » les exécutions du protocole de vérification).

Plan

Une tentative de construction

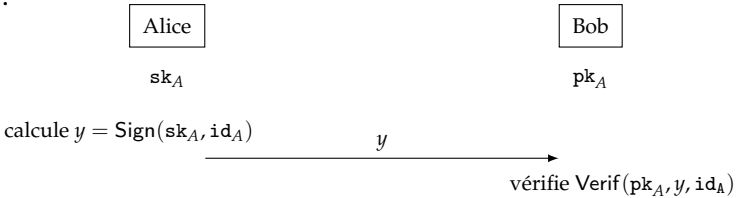
Théorème (informel). Il est possible de construire un schéma d'identification sûr contre des adversaires actifs, **à partir** d'une signature numérique EUF-CMA.

Essayons !

Mise en place. On suppose que Alice est munie d'une paire de clefs $(\mathtt{pk}_A, \mathtt{sk}_A)$ certifiée, pour un schéma de signature numérique quelconque.

On associe également à Alice son identité $\mathtt{id}_A$.

Une première idée :



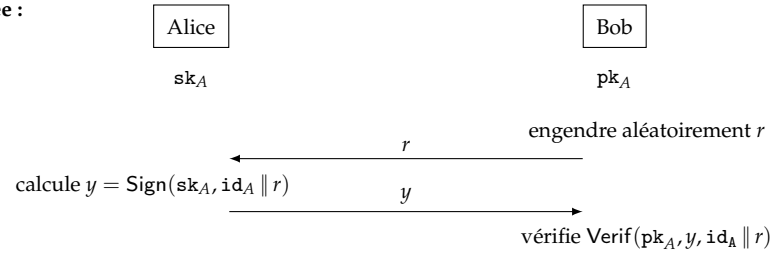
Problème. Attaque par rejeu.

→ Oscar, qui observe le canal de transmission, peut réutiliser la signature y (ailleurs) afin d'être identifié comme Alice.

Identification à partir d'une signature

Mise en place. On suppose que Alice est munie d'une paire de clefs $(\mathtt{pk}_A, \mathtt{sk}_A)$ certifiée, pour un schéma de signature numérique quelconque. On associe également à Alice son identité $\mathtt{id}_A$.

Une seconde idée :

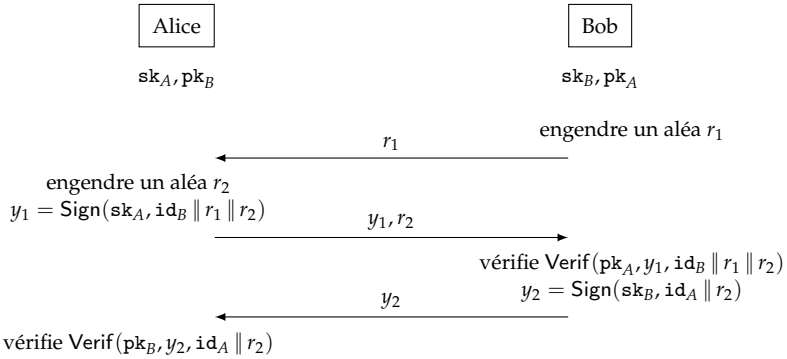


La valeur de y ne peut pas être réutilisée, car elle dépend du "challenge" r engendré par Bob.

Théorème (informel). Il est possible de construire un schéma d'identification sûr contre des adversaires actifs, **à partir** d'une signature numérique EUF-CMA.

Identification mutuelle

Remarque. On peut également construire un protocole d'identification **mutuelle**, où Alice et Bob sont mutuellement convaincus de leurs identités réciproques.



Plan

Schéma d'identification de Schnorr

But : construire un schéma d'identification sans signature préalable.

On suppose qu'Alice a émis une clé publique $\alpha = g^{-a}$ qui est **certifiée**, où
- $g \in \mathbb{F}_p$ est d'ordre premier q dans $\mathbb{F}_p^\times$
- $a \in \{0, \dots, q-1\}$ est gardé secrètement par Alice.

SCHÉMA D'IDENTIFICATION DE SCHNORR

Alice		Bob
a		$\alpha = g^{-a}$
engendre un aléa $k \in \{0, \dots, q-1\}$		
calcule $b = g^k \mod p$		
	$\xrightarrow{b}$	
		engendre un aléa $r \in \{1, \dots, 2^t\}$
	$\xleftarrow{r}$	
calcule $c = k + ar \mod q$		
	$\xrightarrow{c}$	
		vérifie $b \equiv g^c \alpha^r \mod p$

Sécurité du schéma d'identification de Schnorr

Validité. Si le protocole est respecté, alors on a bien

$$g^c a^r \equiv g^{k+ar} g^{-ar} \equiv g^k \equiv b \mod p\,.$$

Sécurité. On peut démontrer le résultat suivant.

Théorème (informel). S'il existe un algorithme polynomial qui réalise une imposture sur le schéma d'identification de Schnorr avec probabilité $\varepsilon > 2^{-t+1}$, alors il existe un algorithme probabiliste qui résout le logarithme discret avec complexité $O(1/\varepsilon)$.

Ce type de schéma d'identification est appelé « **identification en trois passes** ».

1. Phase d'**engagement** (*commitment*) : Alice s'engage sur une valeur aléatoire (k) qu'elle transmet de manière cachée ($g^k \mod p$) à Bob.
2. Phase de **défi**/**challenge** : Bob construit un défi aléatoire (c) auquel Alice doit répondre. La résolution de ce défi dépend de l'engagement d'Alice.
3. Phase de **réponse** : Alice répond au défi aléatoire de Bob, qui vérifie ensuite si la réponse est correcte.

Remarque. Il existe des schémas d'identification à 5 passes, comportant deux challenges et deux réponses.

Exemple-jouet pour Schnorr

Un exemple-jouet (petites valeurs) pour le schéma d'identification de Schnorr.

Paramètres. Soit $p = 88667$ et $q = 1031$. On fixe $t = 10$ (taille des défis). Un sous-groupe de $\mathbb{F}_p^\times$ d'ordre q est engendré par $g = 70322$ par exemple.

Clés. La clé privée d'Alice est $a = 755$. La clé publique correspondante est :

$$\alpha = g^{-a} = 70322^{1031-755} \equiv 13136 \mod 88667$$

Engagement. Alice choisit $k = 543$ et envoie

$$b = g^k \equiv 70322^{543} \equiv 84109 \mod 88667$$

Défi. Bob envoie le défi $r = 1000 < 2^t$.

Réponse. Alice calcule

$$c = k + ar = 543 + 755 \times 1000 \equiv 851 \mod 1031$$

Vérification. Bob vérifie que $b \equiv g^c \alpha^r \mod p$, c'est-à-dire

$$84109 \equiv 70322^{851} 13136^{1000} \mod 88667\,.$$

Plan

Transformation de Fiat–Shamir

La **transformation** (ou **heuristique**) de **Fiat–Shamir** permet de convertir un protocole de vérification interactif en une preuve non-interactive.

Informellement, l'idée est la suivante : l'aléa provenant du vérifieur est **remplacé** par l'utilisation d'une **fonction de hachage**.

Si le protocole initial est un **schéma d'identification**, alors la transformation de Fiat-Shamir permet d'obtenir un **schéma de signature numérique**.

Pour des protocoles en trois passes (engagement, défi, réponse), l'idée est la suivante :

- le prouveur s'engage sur un élément x
- plutôt que de recevoir un défi aléatoire r du vérifieur, le prouveur hache x et le message $\boldsymbol{m}$
- le prouveur calcule ensuite une réponse c au défi $r = H(x, \boldsymbol{m})$, qui constitue alors la signature de $\boldsymbol{m}$.

Si l'on instancie cette idée avec le protocole d'identification de Schnorr, on obtient la **signature de Schnorr**.

Remarque. Dans certains cas, la transformation de Fiat-Shamir permet également de préserver une propriété de **non-divulgation** du protocole interactif.

Signature de Schnorr

SIGNATURE DE SCHNORR : KeyGen

1. Choisir a aléatoirement dans $\mathbb{Z}/q\mathbb{Z}^\times$.
2. Calculer $\alpha = g^{-a}$.
3. La clé publique est $\mathtt{pk} = \alpha$, la clé privée est $\mathtt{sk} = a$.

SIGNATURE DE SCHNORR : Sign$(\boldsymbol{m}, \mathtt{sk})$

1. Choisir $k \in \{1, \dots, q-1\}$ aléatoirement.
2. Calculer $b = (g^k \mod p) \mod q$.
3. Calculer $r = H(b \,\|\, \boldsymbol{m})$ (r correspond à un "défi" aléatoire issu de la fonction de hachage).
4. Calculer $c = k + ar \mod q$.
5. Retourner $s = (r, c)$.

SIGNATURE DE SCHNORR : Verif$(\boldsymbol{m}, s, \mathtt{pk})$

1. Calculer $r' = g^c \alpha^r$
2. Calculer $b' = H(r' \,\|\, \boldsymbol{m})$
3. Faire le test $b' \equiv b \mod q$ et retourner le booléen associé.

Performances. $\simeq$ DSA : clés courtes et adaptable sur les courbes elliptiques.

Plan

Plan

IBE : motivation

Une nouvelle primitive : **chiffrement basé sur l'identité** (*identity-based encryption*, IBE)

Motivation. Dans un système de chiffrement, on souhaite utiliser l'**identité d'un utilisateur comme clé publique** (par exemple, son adresse email).

Avantages.

- L'expéditeur du message peut s'implement utiliser l'identité publique de son destinataire.
- Évite une prédistribution des clefs publiques (pas de PKI).

Inconvénient. Une autorité centrale, possédant des **clés maîtresses**, doit délivrer des clés privées pour le déchiffrement.

Historique. En 2001 :

- Cocks propose un IBE fondé sur le problème de la résiduosité quadratique, mais peu efficace.
- Boneh et Franklin proposent un IBE fondé sur le problème de Diffie-Hellman bilinéaire, qui est alors un problème peu étudié.

IBE : définitions

Définition. Un schéma de **chiffrement basé sur l'identité** (*identity-based encryption*, IBE) est constitué de 4 algorithmes :

1. Un algorithme de génération de **clés maîtresses** `MasterKeyGen()`, qui engendre `mpk/msk`, où la clé maîtresse privée `msk` est détenue par un tiers de confiance.
2. Un algorithme de génération de **clés personnelles** $\texttt{KeyGen}(\texttt{msk}, \texttt{id}_U)$ qui produit une clé privée $\texttt{sk}_U$ destiné à un utilisateur U, à partir de son identité (publique).
3. Un algorithme de chiffrement $\mathsf{Enc}(\texttt{id}_U, \boldsymbol{m})$, avec un fonctionnement similaire au chiffrement à clef publique « classique », mais avec une clé publique correspodant à l'indentité de U (parfois, on notera $\texttt{pk}_U$ une clé publique dérivée de $\texttt{id}_U$),
4. Un algorithme de déchiffrement $\mathsf{Dec}(\texttt{sk}_U, \boldsymbol{c})$, avec un fonctionnement similaire au déchiffrement à clef publique « classique ».

Les **modèles de sécurité** sont similaires à ceux du chiffrement à clef publique vus dans les cours précédents.

Plan

Rappels de résidus quadratiques

Soit $n = pq$ où p et q sont deux nombres premiers distincts tels que $p \equiv q \equiv 3 \mod 4$.
On note $\mathrm{QR}_n^\times$ les résidus quadratiques inversibles modulo n, c'est-à-dire

$$\mathrm{QR}_n^\times \coloneqq \left\{ x \in (\mathbb{Z}/n\mathbb{Z})^\times \;\middle|\; \left(\frac{x}{p}\right) = \left(\frac{x}{q}\right) = 1 \right\}.$$

Les **pseudo-résidus quadratiques** modulo n sont

$$\overline{\mathrm{QR}_n^\times} \coloneqq \left\{ x \in (\mathbb{Z}/n\mathbb{Z})^\times \;\middle|\; \left(\frac{x}{p}\right) = \left(\frac{x}{q}\right) = -1 \right\}.$$

On note enfin

$$\mathcal{Q} \coloneqq \mathrm{QR}_n^\times \cup \overline{\mathrm{QR}_n^\times} = \left\{ x \in (\mathbb{Z}/n\mathbb{Z})^\times \mid \left(\frac{x}{n}\right) = 1 \right\}$$

Description de l'IBE de Cocks. On suppose qu'on a à notre disposition une fonction de hachage $H : \{0,1\}^* \to \mathcal{Q}$.

Clés maîtresses. La clé publique est $\texttt{mpk} = n$ et la clé privée est $\texttt{msk} = (p, q)$.

Génération de clés pour l'utilisateur U d'identité $\texttt{id}_U$.

1. La clé publique $\texttt{pk}_U$ est $H(\texttt{id}_U) \in \mathcal{Q}$.
2. La clé privée $\texttt{sk}_U$ est une racine carrée de $\begin{cases} \texttt{pk}_U & \text{si } \texttt{pk}_U \in \mathrm{QR}_n^\times, \\ -\texttt{pk}_U & \text{si } \texttt{pk}_U \in \overline{\mathrm{QR}_n^\times}. \end{cases}$

Remarque. Seul le détenteur de la factorisation (p, q) de n est capable de décider si $\texttt{pk}_U$ est dans $\mathrm{QR}_n^\times$ ou dans $\overline{\mathrm{QR}_n^\times}$. C'est le problème de la **résiduosité quadratique** qui est supposé difficile.

IBE de Cocks

L'espace des clairs est $\mathcal{M} = \{-1, 1\}$, et celui des chiffrés est $\mathcal{C} = (\mathbb{Z}/n\mathbb{Z})^2$.

Chiffrement. On veut chiffrer un élément $m \in \{-1, 1\}$.

1. Choisir aléatoirement $t_1, t_2 \in \mathbb{Z}/n\mathbb{Z}$, tels que $\left(\frac{t_1}{n}\right) = \left(\frac{t_2}{n}\right) = m$.
2. Calculer
$$\begin{cases} y_1 &= t_1 + \mathtt{pk}_U \times t_1^{-1} \mod n \\ y_2 &= t_2 - \mathtt{pk}_U \times t_2^{-1} \mod n \end{cases}$$
3. Le chiffré est $y = (y_1, y_2)$.

Déchiffrement. On veut déchiffrer $y = (y_1, y_2) \in (\mathbb{Z}/n\mathbb{Z})^2$.

1. Tester si $\mathtt{pk}_U = (\mathtt{sk}_U)^2$:
   - si c'est vrai, définir $s = y_1$,
   - sinon, définir $s = y_2$.
2. Calculer et retourner
$$m' = \left(\frac{s + 2\mathtt{sk}_U}{n}\right)$$

Remarque. On ne chiffre qu'un bit à la fois, mais on peut réitérer le procédé.

Validité. Supposons que $(\mathtt{sk}_U)^2 = \mathtt{pk}_U$ (par exemple). Il faut démontrer que $\left(\frac{s+2\mathtt{sk}_U}{n}\right) = m$ où $s = t_1 + \mathtt{pk}_U\,(t_1^{-1})$ mod n et $\left(\frac{t_1}{n}\right) = m$.
On a :

$$\left(\frac{s+2\mathtt{sk}_U}{n}\right) = \left(\frac{t_1 + \mathtt{pk}_U\,(t_1^{-1}) + 2\mathtt{sk}_U}{n}\right) = \left(\frac{t_1 + (\mathtt{sk}_U)^2\,(t_1^{-1}) + 2\mathtt{sk}_U}{n}\right)$$
$$= \left(\frac{t_1(1 + 2\mathtt{sk}_U\,(t_1^{-1}) + (\mathtt{sk}_U\,t_1^{-1})^2)}{n}\right) = \left(\frac{t_1(1+\mathtt{sk}_U\,t_1^{-1})^2}{n}\right)$$
$$= \left(\frac{t_1}{n}\right)\left(\left(\frac{1+\mathtt{sk}_U\,t_1^{-1}}{n}\right)\right)^2 = \left(\frac{t_1}{n}\right) = m$$

Exercices.

- Pourquoi $\left(\frac{1+\mathtt{sk}_U\,t_1^{-1}}{n}\right) \neq 0$?
- Démontrer la validité pour le cas $(\mathtt{sk}_U)^2 = -\mathtt{pk}_U$.

Sécurité de l'IBE de Cocks

Sécurité. On peut démontrer que la sécurité du système se réduit au **problème du résidu quadratique** (problème QR).

Éléments de preuve. Supposons que l'on ait `ALGO` qui décrypte les chiffrés de Cocks.

Entrée : a et n tels que $\left(\frac{a}{n}\right) = 1$
Sortie : $a \in \mathrm{QR}_n^\times$?

1. Choisir $x \in \{-1, 1\}$ aléatoirement
2. Calculer un t tel que $\left(\frac{t}{n}\right) = x$
3. Calculer $y_1 = t + at^{-1} \mod n$ et choisir y_2 aléatoirement
4. Définir $y = (y_1, y_2)$ et appeler $x' \leftarrow \texttt{ALGO}((n, a), y)$
5. Si $x' = x$, alors retourner $a \in \mathrm{QR}_n^\times$
6. Sinon, retourner $a \notin \mathrm{QR}_n^\times$

L'algorithme ci-dessus retourne une réponse correcte au problème QR avec

- probabilité 1 si $a \in \mathrm{QR}_n^\times$
- probabilité $1/2$ si $a \notin \mathrm{QR}_n^\times$

Plan

Couplages

Pour construire l'IBE de Boneh-Franklin, on a besoin de la notion de **couplage** (*pairing*).

Définition. Soient $(\mathbb{G}_1, +)$, $(\mathbb{G}_2, +)$, $(\mathbb{G}_3, \cdot)$ trois groupes commutatifs. Un **couplage** est une application

$$e : \mathbb{G}_1 \times \mathbb{G}_2 \to \mathbb{G}_3 .$$

Un couplage est **bilinéaire** si, pour tout $P_1, P_2 \in \mathbb{G}_1$ et tout $Q_1, Q_2 \in \mathbb{G}_2$ on a :

$$e(P_1 + P_2, Q_1) = e(P_1, Q_1) \cdot e(P_2, Q_1) \quad \text{et} \quad e(P_1, Q_1 + Q_2) = e(P_1, Q_1) \cdot e(P_1, Q_2)$$

Remarque. Si $e : \mathbb{G}_1 \times \mathbb{G}_2 \to \mathbb{G}_3$ est un couplage bilinéaire, alors, $e(aP, bQ) = e(P, Q)^{ab}$.

Exemples (hors contexte cryptographique)

- Le produit scalaire $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, où $\mathbb{G}_3 = \mathbb{R}$ est noté additivement
- Le déterminant $\det : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$, où $\mathbb{G}_3 = \mathbb{R}$ est noté additivement.

En cryptographie, on utilise des couplages provenant de **courbes elliptiques**. Par exemple, si $E(\mathbb{F}_p)$ est le groupe de points d'un courbe E sur $\mathbb{F}_p$, on note $E(\mathbb{F}_p)[\ell]$ le sous-groupe des points d'ordre divisant ℓ. Alors, il existe un **couplage de Weil** (explicite)

$$e : \ E(\mathbb{F}_p)[\ell] \times E(\mathbb{F}_p)[\ell] \to \mathbb{F}_{p^k}^{\times}$$

Remarque : la description de ce couplage demande des notions avancées en courbes elliptiques.

IBE de Boneh-Franklin

Soit $e : \mathbb{G}_1 \times \mathbb{G}_2 \to \mathbb{G}_3$ un couplage bilinéaire, et P un générateur de $\mathbb{G}_2$, que l'on suppose d'ordre q. On note $\mathbb{G}_1$ et $\mathbb{G}_2$ additivement, et $\mathbb{G}_3$ multiplicativement. On se donne également deux fonctions de hachage

$$H_1 : \{0,1\}^* \to \mathbb{G}_1 \setminus \{1\}$$
$$H_2 : \mathbb{G}_2 \to \{0,1\}^n$$

où n est un paramètre (longueur du clair).

Boneh et Franklin (2001) ont proposé l'IBE suivant.

Clés maîtresses. $\mathtt{msk} = s \in \{1, \ldots, q-1\}$ et $\mathtt{mpk} = sP \in \mathbb{G}_2$

Génération de clés. Pour l'utilisateur U, la paire de clef $(\mathtt{pk}_U, \mathtt{sk}_U)$ est créée, où

$$\mathtt{pk}_U = H_1(\mathtt{id}_U) \in \mathbb{G}_1$$

et

$$\mathtt{sk}_U = s \cdot \mathtt{pk}_U \in \mathbb{G}_1$$

L'espace des clairs est $\mathcal{M} = \{0,1\}^n$ et l'espace des chiffrés est $\mathcal{C} = \mathbb{G}_2 \times \{0,1\}^n$.

Chiffrement. Pour chiffrer un message $\boldsymbol{m} \in \{0,1\}^n$:

1. Choisir $r \in \mathbb{Z}/q\mathbb{Z}$ inversible.
2. Calculer $\boldsymbol{y}_1 = \boldsymbol{m} \oplus H_2(\, e(\mathtt{pk}_U, \mathtt{mpk})^r \,) \in \{0,1\}^n$
3. Calculer $y_2 = rP \in \mathbb{G}_2$.
4. Le chiffré est $y = (\boldsymbol{y}_1, y_2)$

Déchiffrement. Pour déchiffrer $y = (\boldsymbol{y}_1, y_2)$:

1. Calculer $\boldsymbol{m}' = \boldsymbol{y}_1 \oplus H_2(e(\mathtt{sk}_U, y_2)) \in \{0,1\}^n$.
2. Retourner $\boldsymbol{m}'$.

Validité. Il suffit de vérifier que $e(\mathtt{sk}_U, y_2) = e(\mathtt{pk}_U, \mathtt{mpk})^r$. C'est vrai car :

$$e(\mathtt{sk}_U, y_2) = e(s\,\mathtt{pk}_U, rP) = e(\mathtt{pk}_U, P)^{rs} = e(\mathtt{pk}_U, sP)^r = e(\mathtt{pk}_U, \mathtt{mpk})^r .$$

Sécurité. Dans le modèle de l'oracle aléatoire (rappel informel : $\simeq$ fonctions de hachage idéales), la **sécurité sémantique** (CPA) de l'IBE de Boneh-Franklin se réduit au problème de Diffie-Hellman bilinéaire.

Problème de Diffie-Hellman bilinéaire (BDH). Soient $\mathbb{G}_1$ et $\mathbb{G}_2$ deux groupes finis d'ordre q, notés additivement, et soit $e : \mathbb{G}_1 \times \mathbb{G}_2 \to \mathbb{G}_3$ un couplage bilinéaire où $\mathbb{G}_3$ est noté multiplicativement.

Instance. un quadruplet (P, Q, aQ, bQ), où $P \in \mathbb{G}_1$, $Q \in \mathbb{G}_2$ d'ordre q, et $a, b \in (\mathbb{Z}/q\mathbb{Z})^\times$.

Question. Trouver Z tel que $Z = e(P, Q)^{ab}$.

Ce problème est supposé **difficile** pour des groupes génériques $\mathbb{G}_i$.

Exercice : Montrer que si BDH est difficile, alors DL est également difficile dans $\mathbb{G}_1$ et $\mathbb{G}_2$.

Fin du cours

Questions ?