SYSTÈMES D'AUTHENTIFICATION BASÉS SUR LES CODES CORRECTEURS D'ERREURS

Vincent HERBERT

Université de Grenoble

10 septembre 2008

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

INRIA

centre de recherche PARIS - ROCQUENCOURT

Plan

RFID = Identification par radiofréquences

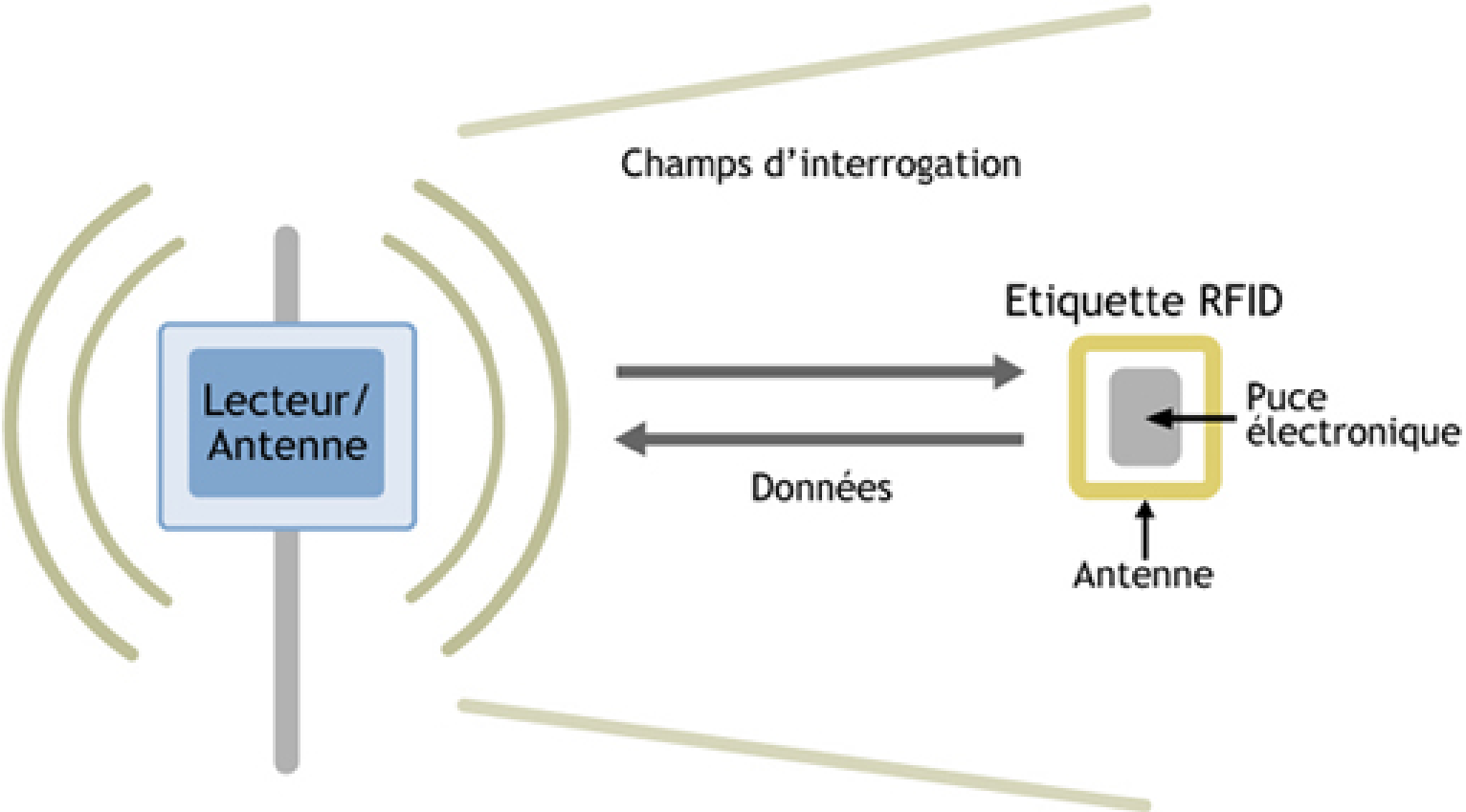


Schéma de principe de la RFID

Caractéristiques des RFID à bas-coût

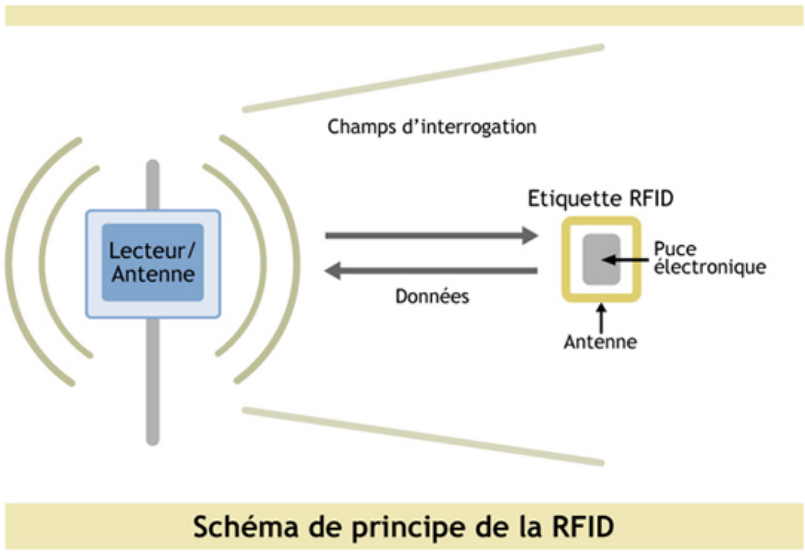


Schéma de principe de la RFID

- Énergie : reçue du lecteur (tags passifs).
- Communication : quelques centimètres à quelques mètres.
- Calculs : simple mémoire jusqu'à microprocesseur.
- Mémoire : centaine de bits jusqu'à plusieurs kilo-octets.
- Prix : dizaine de centimes jusqu'à quelques euros.

Quelques applications de la RFID

Identification : Obtenir l'identité du tag.

- suivi d'objets (le RFID va remplacer les codes barres) ;
- tatouage animal ;
- marquage du bétail ;
- bibliothèque ;
- inventaire.

Authentification : Obtenir une preuve de l'identité du tag.

- passeport ;
- abonnement aux transports publics ;
- badge d'accès ;
- télépéage ;
- clef de démarrage des voitures.

Problématiques de sécurité liées à la RFID

Quels sont les objectifs du cryptologue ?

1. L'Authenticité (l'origine ou la personne doivent être reconnues).
2. La Confidentialité (les informations ne doivent pas être divulguées).
3. L'Intégrité (les informations doivent rester intactes).

Quels sont les objectifs de l'adversaire ?

1. Usurpation d'identité ;
2. Fuite / Vol d'information ;
3. Déni de service (disponibilité) ;
4. Traçabilité malveillante.

Les variantes de HB

- HB : Hopper, Blum (Asiacrypt 2001) ;
- HB+ : Juels, Weis (Crypto 2005) ;
- HB++ : Bringer, Chabanne (2006) ;
- variante de Piramathu : (2006) ;
- HB* : Duc, Kim (2007) ;
- HP MP, HB MP' : Munilla, Peinado (2007) ;
- HB#, Random HB# : Gilbert, Robshaw, Seurin (Eurocrypt 2008) ;
- Trusted HB : Bringer, Chabanne (2008) ;
- Improved HP MP, HB MP+ : Leng, Mayes, Markantonakis (2008).
- PUF-HB : Ghait Hammouri, Berk Sunar (2008).

Le protocole HB : Hopper & Blum (Asiacrypt 2001)

Secret partagé : $x \in \{0,1\}^k$.

Données publiques : k, r, η, $u \in]\eta, \frac{1}{2}[$.

Tag		Lecteur
$\nu \hookrightarrow Ber(\eta)$		
	$\xleftarrow[Challenge]{a \in_R \{0,1\}^k}$	
$z = a \cdot x \oplus \nu$	$\xrightarrow[Réponse]{}$	$z == a \cdot x$

Protocole de type Challenge-Réponse en r rondes.
Le tag est authentifié si le nombre d'erreurs est $\leq t = u \times r$.
La sécurité de HB repose sur la difficulté du problème LPN.

Le problème LPN

Learning from Parity with Noise (LPN)

Entrées :

- A : matrice binaire $k \times n$.
- z : vecteur binaire de longueur n
- η : nombre réel dans $[0, \frac{1}{2}[$

Problème : Existe-t-il $x \in \{0,1\}^k$ tel que $d_H(z, x * A) \leq \eta \times n$.

- LPN est un problème NP-complet (pire des cas).
- LPN est conjecturé difficile (cas moyen).

Il existe plusieurs LPN-Solvers : BKW (2003), Fossorier et Mihaljevic (2006), Levieil et Fouque (2006).

Quelques problèmes difficiles liés aux codes

Syndrome Decoding (SD)

Entrées :

- H : matrice binaire $n \times r$.
- s : vecteur binaire de longueur r
- w : entier positif

Problème : Existe-t-il un vecteur binaire de longueur n et de poids inférieur ou égal à w tel que $e * H = s$.

Ce problème est *NP-Complet* [Berlekamp, McEliece, van Tilborg 1978].

Problème du décodage borné d'un code linéaire

Entrées :

- G : matrice binaire $k \times n$ de rang k
- x : vecteur binaire de longueur n

Problème : Existe-t-il un vecteur binaire de longueur k tel que $d_H(x, m * G) \leq w(n, k)$.

Ce problème est *NP-Complet* [Finiasz 2004].

Attaque sur HB

1. HB résiste aux attaques passives.
2. HB sensible aux attaques actives.

- On envoie un challenge **a** plusieurs fois au tag
- On en déduit la valeur de **a**.**x** sachant que $\eta < \frac{1}{2}$.
- On retrouve **x** par élimination de Gauss en retrouvant k équations avec des **a** linéairement indépendants.
- La complexité algorithmique asymptotique de l'élimination de Gauss (méthode naïve) est $O(k^3)$
- On peut descendre à $O(k^{2.38})$ avec les dernières techniques connues. Intérêt limité ici puisque k est petit (quelques centaines de bits).

De HB à HB+

Il faut savoir que HB n'a pas été pensé par Hopper et Blum pour des applications RFID. En 2001, HB est considéré comme étant un protocole HumanAut, i.e. un protocole qui permet aux humains de s'authentifier sur un ordinateur.

C'est en 2005 que Juels et Weis eurent l'idée de l'adapter à la RFID en constatant les similarités entre humains et tags du point de vue puissance de calcul et mémoire.

Le protocole HB+ : Juels & Weis (Crypto 2005)

HB+ se déroule en r rondes de 3 passes.
Il suit le schéma classique :

1. Engagement* ;
2. Challenge ;
3. Réponse.

Ce schéma de communication fut conservé par les variantes qui suivirent.

*On utilise aussi les termes de *commitment* et de *promesse*.

Le protocole HB+ : Juels & Weis (Crypto 2005)

Secrets partagés : $x \in \{0,1\}^{k_x}$, $y \in \{0,1\}^{k_y}$.

Données publiques : k_x, k_y, r, η, $u \in]\eta, \frac{1}{2}[$.

Tag		Lecteur
$\nu \hookrightarrow Ber(\eta)$		
	$\xrightarrow[\text{Blinding Factor}]{b \in_R \{0,1\}^{k_y}}$	
	$\xleftarrow[\text{Challenge}]{a \in_R \{0,1\}^{k_x}}$	
$z = a \cdot x \oplus b \cdot y \oplus \nu$	$\xrightarrow[\text{Réponse}]{}$	$z == a \cdot x \oplus b \cdot y$

Protocole en r rondes.

Le tag est authentifié si le nombre d'erreurs est $\leq t = u \times r$.

Pourquoi HB+ suscite-t-il tant d'intérêt ?

- HB+ est simple, élégant, facile à implémenter. Il requiert uniquement des opérations logiques XOR et ET bit à bit ainsi que des bits de bruit.
- La sécurité de HB+ a été étudié (travaux de Katz...). On dispose de preuves de sécurité ce qui est très appréciable.

Les limitations de HB

La famille des protocoles HB n'est pas une solution de sécurité complète, elle ne traite que le problème de l'authentification des tags.

Elle néglige d'autres aspects importants comme l'authentification du lecteur, le problème du pistage, le problème de l'anonymat ainsi que celui de la confidentialité de l'identification du tag.

Sécurité de HB+

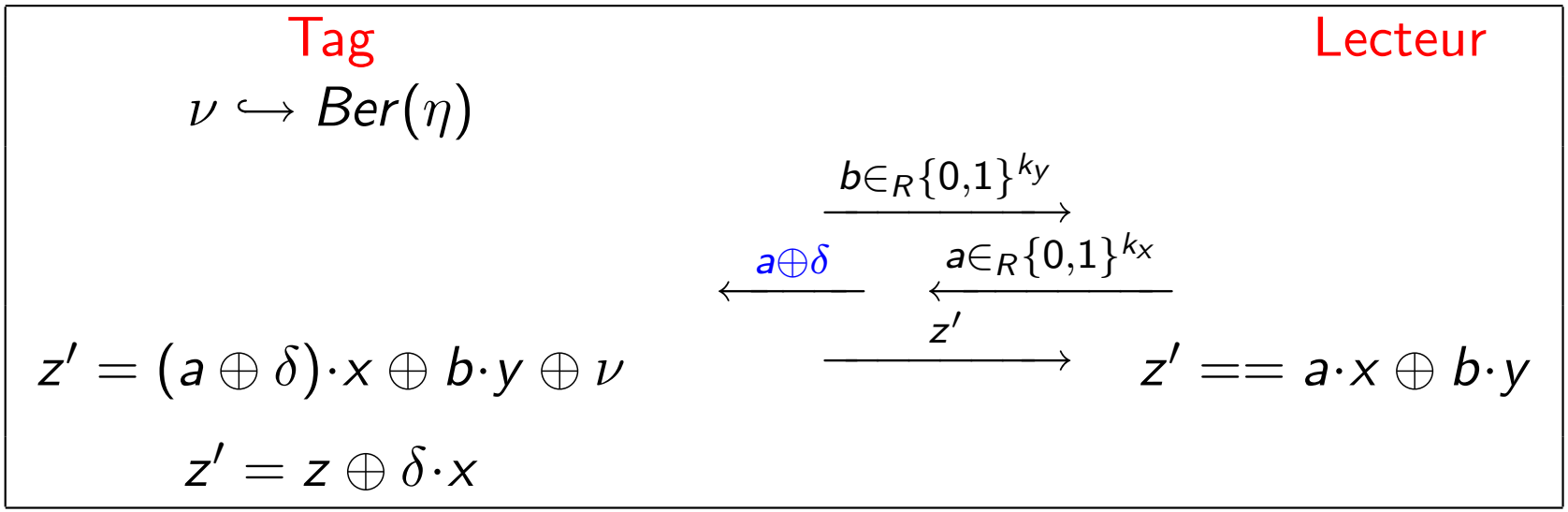
1. HB+ résiste aux attaques actives (modèle de détection).
2. HB+ sensible aux attaques Man in the Middle (modèle GRS-MiM).

La sécurité de HB+ a été prouvée dans le modèle de détection par le biais d'une réduction de LPN. Une attaque contre HB+ a été découverte dès 2005 par Gilbert, Robshaw et Seurin (GRS) dans le modèle GRS-MiM...

Attaque sur HB+ : Gilbert, Robshaw & Seurin

Secrets partagés : $x \in \{0,1\}^{k_x}$, $y \in \{0,1\}^{k_y}$.

Données publiques : k_x, k_y, r, η, $u \in]\eta, \frac{1}{2}[$.



Protocole inchangé (resp. inversé) si $\delta \cdot x = 0$ (resp. $\delta \cdot x = 1$).

$\Rightarrow$ 1 bit d'information à chaque exécution du protocole.

Le protocole Random HB# : Gilbert, Robshaw & Seurin (Eurocrypt 2008)

Secrets partagés : $X \in \{0,1\}^{k_X \times m}$, $Y \in \{0,1\}^{k_Y \times m}$.

Données publiques : k_X, k_Y, m, η, $u \in]\eta, \frac{1}{2}[$.

Tag		Lecteur
$\nu = (\nu_0, \nu_1, .., \nu_m)$ $\nu_i \hookrightarrow Ber(\eta)$ i.i.d.		
	$\xrightarrow[\text{Blinding Factor}]{b \in_R \{0,1\}^{k_Y}}$	
	$\xleftarrow[\text{Challenge}]{a \in_R \{0,1\}^{k_X}}$	
$z = a * X \oplus b * Y \oplus \nu$	$\xrightarrow[\text{Réponse}]{}$	$\nu = z \oplus a * X \oplus b * Y$

Protocole en 1 ronde.

Le tag est authentifié si $w_H(\nu) \leq t = u \times m$.

Le protocole HB# : Gilbert, Robshaw & Seurin (Eurocrypt 2008)

X et Y ne sont plus aléatoires, ce sont des matrices de Toeplitz. Ces matrices sont entièrement déterminées par leur première ligne et leur première colonne

Toeplitz = diagonales constantes

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 6 & 1 & 2 & 3 & 4 \\ 7 & 6 & 1 & 2 & 3 \\ 8 & 7 & 6 & 1 & 2 \end{pmatrix}$$

Problème du coût de stockage des données :

- Random HB# : bien en théorie mais impraticable.
  coût de stockage : $k \times m$ bits.
- HB# : bien pour la pratique mais moins sécurisé.
  coût de stockage : $k + m - 1$ bits.

Attaque sur HB# : Ouafi, Overbeck & Vaudenay (2008)

1. On obtient par écoute passive un triplet HB valide $(\overline{a},\overline{b},\overline{z})$. Notre but est de trouver $\overline{w}$, le poids de l'erreur $\overline{\varepsilon}$;
2. On construit un oracle qui donne $w(\overline{\varepsilon})$;

Tag		Lecteur
	$\xrightarrow{b\oplus\overline{b}}$	
	$\xleftarrow{a\oplus\overline{a}}$	
$z = (a \oplus \overline{a}) * X \oplus b * Y \oplus \varepsilon$	$\xrightarrow{z\oplus\overline{z}}$	$w(\overline{\varepsilon} \oplus \varepsilon) \leq t$

$\nu = z \oplus \overline{z} \oplus a * X \oplus (b \oplus \overline{b}) * Y = \overline{\varepsilon} \oplus \varepsilon$

3. Pour chaque i, on calcule $w(\overline{\varepsilon} \oplus 2^i)$;
4. On en déduit $\overline{\varepsilon}$ puis $\overline{a} * X \oplus \overline{b} * Y$;
5. On itère ce procédé pour d'autres triplets $(\overline{a},\overline{b},\overline{z})$ et on trouve X et Y.

Proposition de schéma d'authentification low-cost basé sur LPN

Secrets partagés : $x \in_R \{0,1\}^{k_x}$, $y \in_R \{0,1\}^{k_y}$

Données publiques : k_x, k_y, m, η, $u \in]\eta, \frac{1}{2}[$

Tag		Lecteur
$\nu = (\nu_0, \nu_1, .., \nu_m)$		
$\nu_i \hookrightarrow Ber(\eta)$ i.i.d.		
	$\xrightarrow[\text{Blinding Factor}]{B \in Toeplitz(k_y,m)}$	
	$\xleftarrow[\text{Challenge}]{A \in Toeplitz(k_x,m)}$	
$z = x * A \oplus y * B \oplus \nu$	$\xrightarrow[\text{Réponse}]{}$	$\nu = z \oplus x * A \oplus y * B$

En pratique

k_x	k_y	m	η	t	$P[FR]$	$P[FA]$
80	512	1164	0.25	405	2^{-45}	2^{-83}
80	512	441	0.125	113	2^{-45}	2^{-83}

NB : k_x et k_y jouent deux rôles distincts. La sécurité de y repose sur la difficulté de LPN. 512 bits sont requis pour obtenir 80 bits de sécurité [Levieil, Fouque (2006)].

Nb de bits stockés	Nb de bits transmis
592 (1756)	4082 (2918)
592 (1033)	1913 (1472)

Coût de stockage : $k_x + k_y$ bits.

Coût de transmission : $k_x + k_y + 3 \times m - 2$ bits.

- Avantages : Simplicité et coût de stockage minimisé.
- Inconvénients : Coût de communication plus important, sensible à l'attaque de Overbeck et Vaudenay.

Autres résultats

- LPN et Syndrome Decoding (ou décodage à maximum de vraisemblance) sont équivalents. (preuve réductionniste forte dans le rapport)
- Résultat négatif : Les algorithmes de décodage génériques (Canteaut-Chabaud, Stern) sont moins performants que les meilleurs LPN-Solvers avec les jeux de paramètres recommandés pour HB (implémentation, tests).

Un problème de décodage particulier

1. Il s'agit d'un code linéaire aléatoire (sans structure).
2. Son rendement k/n peut-être rendu aussi proche de 0 que l'adversaire le veut (selon sa puissance de calcul).
3. La capacité de correction (théorique) t/n du code est nettement supérieure au taux d'erreur η injecté dans les réponses. (possibilité de gain ?)

Remarque : Pour LF2, le meilleur LPN solver, $n = 10000$ (10 authentifications) est suffisant.

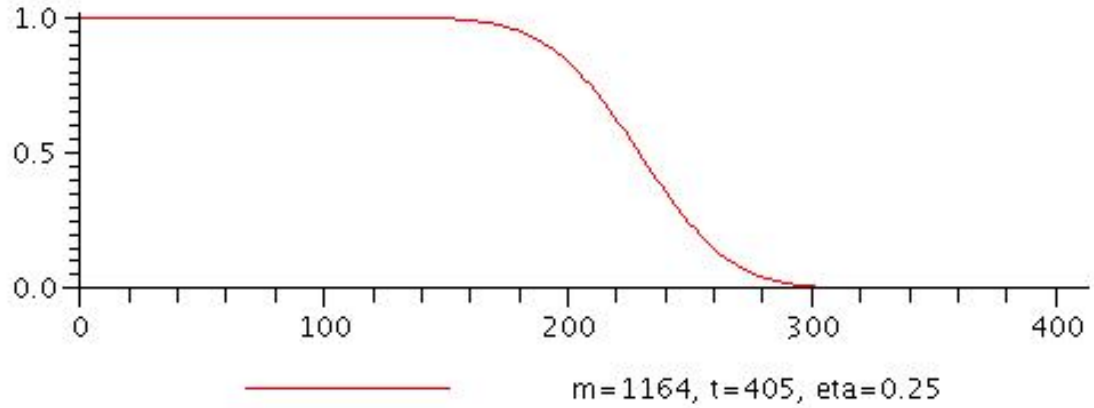
Bilan

- Il existe des solutions cryptographiques basées sur les codes correcteurs pour la RFID.
- "De bonnes variantes de HB+ sont très difficiles à trouver" (GRS)
- Besoins de sécurité toujours importants avec des contraintes lourdes.
- Quelle que soit la variante, une attaque sur la primitive LPN est toujours possible.
- HB est un problème de décodage particulier et encore mal étudié en tant que tel.

Merci de votre attention ! ! !

Comportement de

$$f_{m,t,\eta}(\overline{w}) = Pr_{\varepsilon \hookrightarrow Bin(m,\eta)}(w(\overline{\varepsilon} \oplus \varepsilon) \leq t)$$



$$Pr(w(\overline{\varepsilon} \oplus \varepsilon) \leq t) = \sum_{i=0}^{\overline{w}} \sum_{j=0}^{i+t-\overline{w}} \binom{\overline{w}}{i} \times \binom{m-\overline{w}}{j} \times \eta^{i+j} \times (1-\eta)^{m-i-j}$$