Master M2 Cryptis - Université de Limoges

Examen Codes et cryptographie - 12 fevrier 2016

Documents autorisés - durée 3h

Questions de cours:

a. Soit un code de Reed-Solomon $[10, 2, 9]$ sur le corps $GF(11)$ jusqu''a quelle distance peut on décoder avec un décodage classique ? Et avec l'algorithme de Sudan (justifier le calcul)?

b. Soit un code de Goppa $[2^m, 2^m - mt, 2t + 1]$ qui peut décoder t erreurs, quelle est la densité de mots de l'espace qu'il peut décoder (ie le rapport des mots decodables de l'espace par le nombre de mots total de l'espace)? (justifier le calcul).

c. Soit un code de Reed-Solomon $[2^m - 1, k]$ sur $GF(2^m)$, pour $k = n - 2t$ et $n >> t$ calculer de maniere approchée la densité des mots décodés par le code. Comparer à la densité obtenue dans la question précédente.

d. Soit le corps $K = GF(q^m)$, on considère un code en metrique rang de longueur n et de dimension k. Rappeler le principe de la metrique rang, donner une borne inferieure equivalente à la borne de Singleton poru la metrique rang (attention la borne doit dependre de m,n et k, pas simplement de n et k).

e) Des codes codes linéaires binaires $[19, 10, 14]$ et $[16, 8, 6]$ peuvent-ils exister ? (justifier)

Partie I (Authentification par les codes : schéma de Veron):

Les questions sont indépendantes, les question 1,2,5,6 sont faciles, la 3 un peu moins et la 4 encore moins

On considère le schéma d'authentification de Veron, qui est une variation sur le schéma d'authentification de Stern vu en cours. Dans le cas de l'algo de Stern, la clé publique est un syndrome et la clé secrete un mot de petit poids associé, dans le cas de Veron, la clé publique x est un mot du code mG (G une matrice generatrice d'un code aleatoire $[n, k]$) bruité par une erreur e de poids w. Plus précisément la clé publique est le triplet (G, x, w), avec G une matrice aleatoire $k \times n$, $x = mG + e$ et w le poids de e. La clé secrete est le couple décodé (m, e) (qu'on supposera unique pour un x fixé). On supposera dans la suite que C est un code de parametre $[n, n/2]$ (typiquement $n = 700$). Le protocole a pour but de montrer que le prouveur P connait le decodage (m, e) du mot bruité $mG + e$ au verifieur V. Dans la suite h est une fonction de hachage, on considere de

1. [Engagement] P choisit au hasard $u \in GF(2)^k$ et une permutation σ de $\{1, 2, \ldots, n\}$. P envoie à V les engagements c_1, c_2 and c_3 tels que:

$$c_1 = h(\sigma);\ c_2 = h(\sigma((u+m)G));\ c_3 = h(\sigma(uG + x));$$

2. [Défi] V envoie $b \in \{0, 1, 2\}$ à P.

3. [Réponse] 3 cas :
   - si $b = 0$: P revele $(u + m)$ et σ.
   - si $b = 1$: P revele $\sigma((u + m)G)$ et $\sigma(e)$.
   - si $b = 2$: P revele u and σ.

4. [Verification Step] 3 cas :
   - si $b = 0$: V verifie que c_1, c_2 ont été calculé honnetement (ie qu'il est capable de reconstruire c_1 et c_2 et que cela correspond aux valeur de l'engagement).
   - si $b = 1$: V verifie que c_2, c_3 ont été calculé honnetement, et que le poids de $(\sigma(e)) = w$.
   - si $b = 2$: V verifie que c_1, c_3 ont été calculé honnetement.

Figure 1: Protocol of Veron

plus que la description la permutation σ équivaut à donner une 'graine' de 80 bits qui permet de reconstruire σ.

1) Montrer que le protocole fonctionne (montrer que si tout se passe normalement le verifieur peut effectivement verifier tous les cas), quand se sert-on de la clé publique ?

2) Montrer qu'un tricheur peut facilement anticiper n'importe quel choix de b pour le défi (ie choisir un engagement adequat qui lui permet de se faire passer pour P)

3) Montrer qu'un tricheur peut facilement anticiper 2 choix sur 3 de b (ie soit $b = 0$ ou 1, soit $b = 1$ ou 2, soit $b = 0$ ou 2). (Montrer au moins un des trois cas au choix $(0, 1), (1, 2)$ ou $(0, 2)$). En déduire que la proba de triche est au moins 2/3.

4) (**question difficile**) Montrer que si un tricheur peut anticiper les 3 possibilités pour b, alors soit il est capable de trouver une collision pour la fonction de hachage h, soit il connait le secret m. (indice: l'idée est de dire que si un tricheur peut repondre à tout b, alors il est capable de construire des c_i de manière differente, et donc ou bien ces valeurs sont egales auquel cas on montre

qu'on connait le secret, ou bien on a trouvé une collision pour h). On en déduit que la proba de triche est exactement 2/3.

5) Calculer le cout moyen pour les communications (nbre de bits envoyés lors du protocole) pour l'execution de 1 tour dans le cas $n = 700, k = n/2$ et la taille du haché 160 bits. Si on veut une authentification avec proba de triche de 2^{-32}, combien de fois faut-il executer le protocole ? Quelle est alors le cout moyen des communications pour une telle authentification ?

6) Ecrire le cout des communications pour le protocole de Veron en fonction de $n, k = n/2$ pour un haché de taille 160. Faire pareil pour Stern, montrer que le protocole de Veron permet de gagner un peu sur le cout des communications.

Partie II: Schémas basés sur l'identité

1) Rappeler en quelques lignes la notion de Schéma basé sur l'identité et le principe de la signature Courtois-Finisaz-Sendrier (CFS) pour les codes.

2) Proposer un schéma d'authentification basé sur l'identité obtenu en utilisant d'abord le schéma de signature de CFS sur les codes puis le schéma d'authentification de Stern sur les codes. Peut-on en déduire un schéma de signature basée sur l'identité ?

3) Montrer qu'on peut construire de manière générique un schéma de signature basée sur l'identité à partir de deux schémas de signature quelconques (voire 2 fois le même schéma).

Partie III Attaques sur le schéma de McEliece

1) a) Rappeler le principe du cryptosystème (ou schéma) de McEliece. Et présenter ses intérêts par rapport aux systèmes classiques basés sur la théorie des nombres.

b) Un des inconvénients de ce système est la taille de la clé publique. Est-il possible de donner la matrice masquée utilisée pour le chiffrement sous forme systématique sans compromettre la sécurité du message ? Justifier votre réponse. Dans le cas négatif voyez-vous une manière de modifier légèrement le schéma en utilisant l'erreur dans le message chiffré pour donner la matrice masquée sous forme systématique sans compromettre la sécurité du message.

2) a) Le syst'eme de McEliece est-il un système déterministe ? Rappeler comment on peut retrouver un message m si on connait deux chiffrés différents c_1 et c_2 de ce même message.

b) Montrer qu'en modifiant légèrement le schéma de McEliece en utilisant une fonction de hachage qui relie le message encodé et l'erreur on peut très facilement contrer cette attaque.

3) "Attaque par réaction". Une attaque est dite par réaction lorsqu'on envoie un message chiffré légèrement modifié à un déchiffreur et qu'on observe sa réaction. On suppose que le déchiffreur quand il reçoit le message réagit de

la façon suivante: s'il a pu quand même déchiffrer le message il ne fait rien et on déduit qu'alors la modification n'a pas gêné le déchiffrement, et s'il n'a pas pu déchiffrer le message alors il redemande un nouvel envoi du chiffré et on déduit que la modification a empêché le déchiffrement. L'idée de l'attaque est alors d'utiliser l'information donnée par le fait que la petite modification faite sur le chiffré a empêché le déchiffrement ou non.

a) On suppose qu'on intercepte un message chiffré, sur lequel on va essayer l'attaque précédente. Proposer une attaque par réaction sur le schéma de McEliece qui permet, en observant la réaction à l'envoi de l'ordre de n (pour n la longueur ducode utilisé) messages chiffrés modifiés d'1 bit, de retrouver le message clair?

b) Quelles sont les limites de cette attaque.

Partie IV

Soit le corps à 4 éléments $F_4 = \{0, 1, \omega, \omega^2\}$ tel que $\omega^3 = 1$ et $1+\omega+\omega^2 = 0$. On appelle hexacode le code H_6 de matrice génératrice

$$\begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & \omega & \omega^2 \\ 0 & 0 & 1 & 1 & \omega^2 & \omega \end{bmatrix}.$$

1) Rappeler la définition du produit scalaire sur F_2 entre 2 mots $x(x_1, ...x_n)$ et $y(y_1, ..., y_n)$. Si G est une matrice generatrice d'un code binaire sous forme systematique $[I_k|A]$ (pour I_k la matrice identité $k \times k$, et A une matrice $k \times (n-k)$) , rappeler la forme d'une matrice generatrice du code dual associé.

2) On considére le produit scalaire hermitien sur F_4 défini par $x.y = \sum_{i=1}^{n} x_i.y_i^2$. Si G est une matrice generatrice d'un code sur F_4 sous forme systematique $[I_k|A]$ (pour I_k la matrice identité $k \times k$, et A une matrice $k \times (n-k)$) rappeler la forme d'une matrice generatrice du code dual associé.

3) Montrer que si un code C sur F_4 est tel que $C \subset C^{\perp}$ (ou $C^{\perp}$ est le code dual de C pour le produit scalaire hermitien) alors le poids de Hamming de tous les mots de C est pair.

4) Montrer que l'hexacode est un code auto-dual sur F_4 pour le produit sclaire hermitien.

5) En utilisant des idées analogues à celles utilisées pour montrer que le poids minimum du code de Golay étendu est 8, montrer que le poids miminum de l'hexacode est 4. Montrer que ce code alors vérifie la borne de Singleton et donc qu'il est MDS.