Initiation à la cryptographie

Correction TD de cryptographie n°3

—TELECOM Nancy 2A ISS—

1 Modes opératoires

▸ **Exercice 1.** Modes opératoires et authentification

1. Un message chiffré avec un chiffrement par blocs et un mode opératoire quelconque garantit-il l'authentification du message reçu ?
2. Par exemple, soit un chiffré AES-CBC : $(\mathrm{IV}, c_0, c_1, \ldots)$ intercepté par un attaquant. Que peut-il retransmettre au destinataire pour se faire passer pour l'émetteur officiel ?
3. Que faut-il ajouter au système pour obtenir une garantie d'authenticité du message reçu.

Correction :

1. *Non.*
2. *S'il coupe le message chiffré à n'importe quel bloc, le destinataire ne verra pas la différence.*
3. *On ajoute un MAC.*

▸ **Exercice 2.** Electronic Code Book

Le mode de chiffrement ECB (*Electronic Code Book* ou *Dictionnaire de code*) est le mode de chiffrement le plus simple que l'on puisse imaginer : chaque bloc de données est chiffré indépendamment par la fonction de chiffrement.

1. Ce mode de chiffrement n'est pas sûr, expliquer pourquoi.
2. Jack, qui gagne 105000€ par an [1], a retrouvé l'entrée chiffrée qui lui correspond dans la base de donnée des salaires de son entreprise :

   `Q92DFPVXC9IO`

   Sachant que la fonction de chiffrement utilisé a des blocs de deux caractères et que le service informatique de son entreprise ne comprend aucun expert en cryptographie (entendre par là, utilise le mode ECB !), retrouver le salaire de Jane la patronne de Jack parmi le reste de la base de donnée :

   `TOAV6RFPY5VXC9, YPFGFPDFDFIO, Q9AXFPC9IOIO, ACED4TFPVXIOIO, UTJSDGFPRTAVIO.`
3. Exemple 2. Imaginer à quel point ce mode chiffrement est déplorable pour les photographies.

1. Exemple de `https://fr.wikipedia.org/wiki/Mode_d\%27op\%C3\%A9ration_(cryptographie)`.

Correction :

1. *Deux blocs identiques auront le même chiffré, ainsi de l'information peut fuir.*
2. *On peut supposer que l'entrée de Jack donne la correspondance suivante :*

   ```
   Ja|ck|??|10|50|00
   Q9|2D|FP|VX|C9|IO
   ```

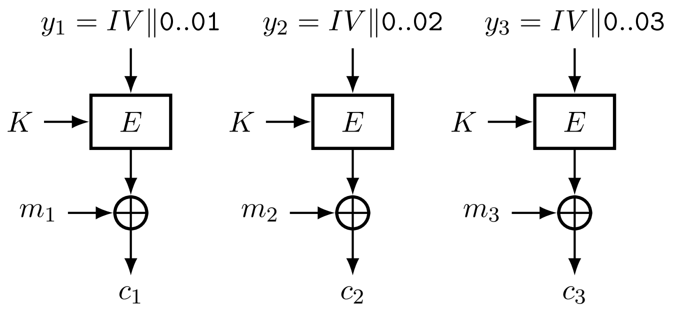
   `FP` *doit correspondre au séparateur de champ. Jane a aussi un prénom de 4 lettres qui commence par ≪Ja≫ donc son entrée chiffrée commence par* `Q9??FP`*, c'est* `Q9AXFPC9IOIO`*. Son salaire est ainsi* `C9IOIO`*, soit 500000€ par an.*
3. *Cela correspond à remplacer les couleurs si la taille du bloc correspond à un pixel, ou des blocs d'image par d'autre, on reconnaitra la forme de l'image : par ex les pixels noir deviendront tous rouges.*

▸ **Exercice 3.** Modes et vecteur d'initialisation

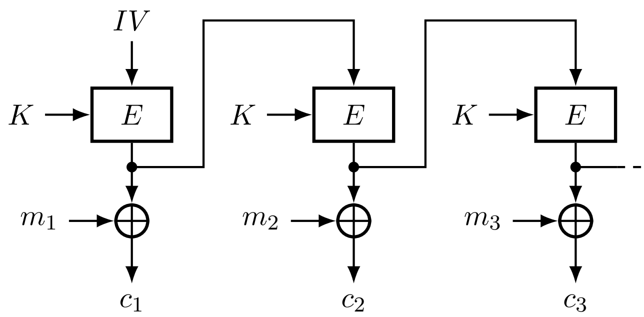
Nous considérons dans cet exercice un chiffrement par bloc E paramétré par une clé secrète K. Notons n la taille (en bits) des blocs en question.

1. Rappelez les valeurs typiques de n.

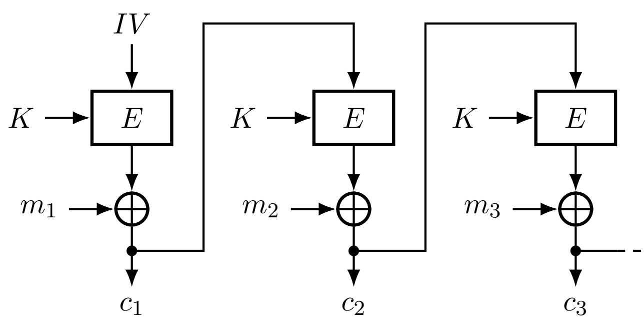
Nous nous intéressons tout d'abord au cas du mode CTR (*Counter*), dont nous rappelons ici la définition. Étant donné un vecteur d'initialisation IV de $n - 64$ bits, nous notons $y_i = IV \| i$, pour tout $0 < i < 2^{64}$, la concaténation de cet IV avec l'entier i, représenté sur 64 bits. Le chiffrement du $i^{\text{ème}}$ bloc en clair m_i est alors donné par la formule $c_i = m_i \oplus E_K(y_i)$, pour tout $0 < i < 2^{64}$, comme représenté sur le schéma suivant :



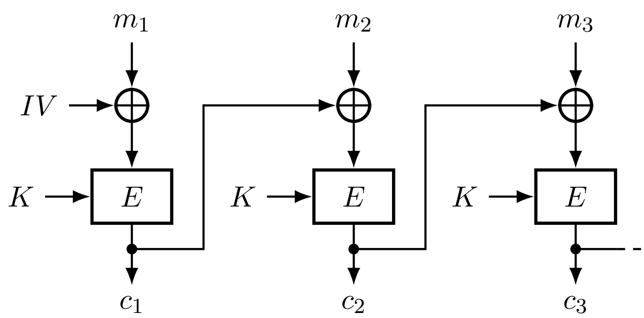
1. Donnez le schéma de déchiffrement, ainsi que la formule correspondante pour calculer chaque bloc m_i en fonction de la clé K, du vecteur d'initialisation IV et du bloc chiffré c_i.
2. Supposons alors qu'un utilisateur décide d'utiliser toujours le même vecteur d'initialisation IV pour chiffrer plusieurs messages. Supposons de surcroît que vous, l'attaquant, disposiez d'un couple clair / chiffré (M, C). Quelle est le nom de ce type d'attaque ?
3. Pouvez-vous utiliser la connaissance de M et C afin de décrypter d'autres messages chiffrés avec la même clé K et le même vecteur d'initialisation IV ? Si oui, comment faites-vous ?
4. Quel intérêt voyez-vous à ce mode de chiffrement quant à son implémentation ?
5. Considérons alors le mode OFB (*Output Feedback*) suivant :



6. Donnez les formules de chiffrement et de déchiffrement de ce mode.
7. Est-ce qu'une attaque à clair connu est possible sur le mode OFB si un même vecteur d'initialisation IV est utilisé pour tous les messages ?
8. Mêmes questions pour le mode CFB (*Cipher Feedback*) suivant :



9. Mêmes questions pour le mode CBC (*Cipher Block Chaining*) suivant :



10. À quoi sert le vecteur d'initialisation (IV) ? Doit-il rester secret ?
11. Que se passe-t-il lors du déchiffrement si l'un des blocs chiffrés a été altéré ?

✂ -

Correction :

1. *DES : $n = 64$ bits (cassé) ; AES : $n = 128$ bits.*
2. *Déchiffrement : $m_i = c_i \oplus E_K(y_i) = c_i \oplus E_K(IV\|i)$, pour tout $i > 0$.*
3. *On parle alors d'attaque à* clair connu, *ou KPA, avec un message.*
4. *Quel que soit le message envoyé, la suite $(y_i)_{i>0}$ est toujours la même. De même, la suite chiffrante utilisée pour masquer le message, $(z_i)_{i>0}$ avec $z_i = E_K(y_i)$, ne change pas. Le couple clair / chiffré connu (M, C) peut donc nous permettre de retrouver la valeur des premiers blocs de cette suite chiffrante : $z_i = m_i \oplus c_i$, pour tout $0 < i \leq \ell$, où ℓ désigne le nombre de blocs de M.*

   Ainsi, pour chaque message chiffré intercepté C', on peut retrouver les ℓ premiers blocs du message en clair M' en calculant $m'_i = c'_i \oplus z_i = c'_i \oplus m_i \oplus c_i$, pour tout $0 < i \leq \ell$.
5. *Ce mode est facilement parallélisable. De plus il n'y a pas besoin de fonction de déchiffrement C'est avantageux dans certain cas : pour AES le déchiffrement est plus cher que le chiffrement.*
6. *Chiffrement : $c_i = m_i \oplus z_i$, avec $z_i = E_K(z_{i-1})$ pour tout $i > 0$ et $z_0 = IV$.*

   Déchiffrement : $m_i = c_i \oplus z_i$.

7. *La suite chiffrante est là aussi complètement déterminée par la clé et l'IV : $z_i = E_K(z_{i-1})$ pour tout $i > 0$, avec $z_0 = IV$.*
   Comme pour le mode CTR, on a donc $z_i = m_i \oplus c_i$ et ainsi $m'_i = c'_i \oplus m_i \oplus c_i$, pour tout $0 < i \leq \ell$.
8. *Chiffrement : $c_i = m_i \oplus z_i$, avec $z_i = E_K(c_{i-1})$ pour tout $i > 0$ et $c_0 = IV$.*
   Déchiffrement : $m_i = c_i \oplus z_i$.
   Dans le cas de CFB, seul le premier bloc z_1 de la suite chiffrante ne dépend que de K et de IV : les suivants dépendent aussi des blocs du message en clair.
   L'attaque à clair connu ne permet donc de retrouver que le premier bloc du message en clair : $m'_1 = c'_1 \oplus z_1 = c'_1 \oplus m_1 \oplus c_1$.
   En fait, si les k premiers blocs de M' sont identiques aux k premiers blocs du clair connu M, alors on pourra retrouver $k+1$ blocs de la suite chiffrante, et donc décrypter le bloc suivant de C'.
9. *Chiffrement : $c_i = E_K(m_i \oplus c_{i-1})$, pour tout $i > 0$, avec $c_0 = IV$.*
   Déchiffrement : $m_i = D_K(c_i) \oplus c_{i-1}$.
   Ici, il n'est plus question de suite chiffrante. Les attaques précédentes ne fonctionneront donc pas.
   Néanmoins, le rôle du vecteur d'initialisation est aussi de garantir que deux messages en clair identiques donneront des chiffrés différents. Cela n'est plus le cas lorsque l'IV est fixé.
10. *Si il n'y avait pas d'IV deux fichiers identiques auraient les mêmes chiffrés.*
11. *Seulement deux blocs sont altérés.*

▸ **Exercice 4.** Attaque par insertion

On considère un chiffrement par blocs utilisant un mode opératoire OFB ou CTR. Un attaquant parvient à intercepter un chiffré $C = (c_0, c_1, \ldots)$, correspondant à un message $M = (m_0, m_1, \ldots)$. L'attaquant connaît uniquement C, mais pas M, ni bien sûr la clé K ou encore la valeur IV (pour OFB) ou la *nonce* (pour CTR).

On suppose que l'attaquant parvient à forcer la personne qui chiffre à re-chiffrer un message M' quasiment identique à M, mais avec uniquement un bloc de zéros *inséré* parmi les autres blocs. On suppose en outre que l'attaquant parvient à forcer ce deuxième chiffrement à être réalisé avec la même IV (pour OFB) ou *nonce* (pour CTR). L'attaquant obtient donc un nouveau chiffré C'.

1. Comment l'attaquant peut-il déterminer le bloc à partir duquel M et M' diffèrent ?
2. Supposons que ce premier bloc différent ait pour indice i. Que vaut alors c'_i. Comment l'attaquant peut-il en déduire m_i ?
3. Montrer comment l'attaquant peut alors déduire toute la suite du message.
4. Que doit-on en conclure comme précaution sur l'utilisation de OFB ou CTR ?

Correction :

1. *Si l'insertion est au bloc i (les blocs étant numérotés à partir de 1), alors les blocs de 1 à $i-1$ des chiffrés sont identiques. On repère donc i comme étant le premier indice où les chiffrés diffèrent.*
2. *Mode CTR : Le couple (M'/C') nous donne d'abord la valeur de $E_K(\text{nonce} + i)$. En effet le bloc c'_i vaut $0 \oplus E_K(\text{nonce} + i)$. Ensuite, le couple (M/C) nous indique lui la valeur de m_i car $m_i = c_i \oplus E_K(\text{nonce} + i) = c_i \oplus c'_i$ qui sont tous les deux connus.*
   Mode OFB : Le couple (M'/C') nous donne d'abord la valeur de z_i, avec la suite z définie par $z_i = E_K(z_{i-1}$ et $z_0 = \text{IV}$. En effet le bloc c'_i vaut $0 \oplus z_i$. Ensuite, le couple (M/C) nous indique lui la valeur de m_i car de $m_i = c_i \oplus z_i = c_i \oplus c'_i$ qui sont tous les deux connus.
3. *On veut retrouver toute la suite du message clair, c'est-à-dire les $m_{i+1}, \cdots m_\ell$. On va tirer partie du fait que m_i est chiffrée dans le couple (M'/C') par la même valeur que m_{i+1} l'est dans le couple (M/C). Dans les deux cas (CTR ou OFB), le couple (M'/C') nous donne d'abord la valeur chiffrante f_{i+1} (z_{i+1} pour OFB, et $E_K(\text{nonce} + i + 1)$ pour CTR). En effet le bloc c'_{i+1} vaut $m_i \oplus f_{i+1}$. Comme m_i est connu on peut calculer $f_{i+1} = c'_{i+1} \oplus m_i$. Ensuite, le couple (M/C) nous indique lui la valeur de m_{i+1} car $m_{i+1} = c_{i+1} \oplus f_{i+1} = c_{i+1} \oplus c'_{i+1} \oplus m_i$ qui sont tous les trois connus. Ainsi de suite jusqu'à la fin du message.*

4. *Précaution à prendre : il ne faut jamais réutiliser la même* IV.