Eléments de cryptographie
- Mode d'opération -

Plus rappels sur le hachage et le « salage ».

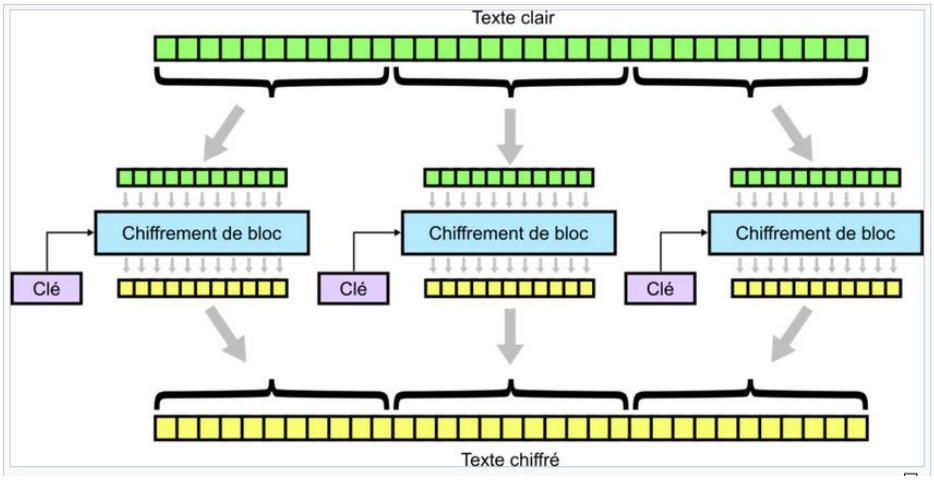
En cryptographie, un mode d'opération est la manière de traiter les blocs de texte clairs et chiffrés au sein d'un algorithme de chiffrement par bloc. Plusieurs modes existent, certains sont plus vulnérables que d'autres :

Dictionnaire de codes (« Electronic codebook » ou ECB)

Il s'agit du mode le plus simple. Depuis 2000 et l'algorithme AES le standard est de blocs de 128 bits.

- **Le message à chiffrer est subdivise en plusieurs blocs qui sont chiffres séparément les uns après les autres.**
- **Le défaut de cette méthode est que deux blocs avec le même contenu seront chiffrés de la même manière** : on peut donc déduire des informations à partir du texte chiffré en cherchant les séquences identiques.

On obtient dès lors un « dictionnaire de codes » avec les correspondances entre le clair et le chiffré d'où le terme « codebook ». Ce mode est pour ces raisons fortement déconseillé dans toute application cryptographique. Le seul avantage qu'il peut procurer est un accès rapide à une zone quelconque du texte chiffré et la possibilité de déchiffrer une partie seulement des données.



Le texte en clair est découpé en blocs et chaque bloc est chiffré, indépendamment des autres, avec la clé de chiffrement.

Exemple :

On chiffre les deux messages suivants avec un mode ECB et un algorithme de chiffrement par bloc qui travaille avec un bloc de deux caractères à la fois. Ce type de fichier pourrait correspondre à une liste de salaires.

```
JOHN__105000
JACK__500000
```

Le chiffrement sur le premier message donne ceci :

```
JO|HN|__|10|50|00
Q9|2D|FP|VX|C9|IO
```

Et sur le deuxième message, on obtient :

```
JA|CK|__|50|00|00
LD|AS|FP|C9|IO|IO
```

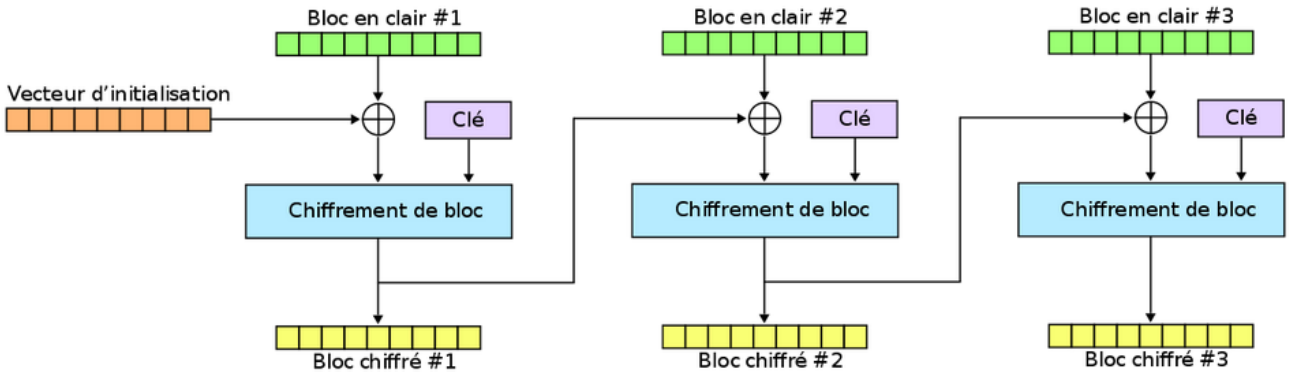
On constate que des paires de caractères apparaissent dans les deux messages chiffrés, il en va de même dans les messages en clair :

```
Q9|2D|FP|VX|C9|IO
LD|AS|FP|C9|IO|IO
```

En partant du principe que John connait son salaire, il pourrait deviner le salaire de Jack car la séquence "C9" correspond à "50" et "IO" à "00". John en déduit que le salaire de Jack, chiffré en « C9IOIO » correspond à « 500000 ».

Enchaînement des blocs (« Cipher Block Chaining » ou CBC)

Dans ce mode, on applique sur chaque bloc un « OU EXCLUSIF » avec le chiffrement du bloc précédent avant qu'il soit lui-même chiffré. De plus, afin de rendre chaque message unique, un VECTEUR D'INITIALISATION (IV) est utilisé.



Que signifie le vecteur d'initialisation (IV)?

- Un IV est un NOMBRE ALEATOIRE qui est utilisé en combinaison avec une clé secrète comme moyen de crypter des données. Un programme de chiffrement ne l'utilise qu'UNE SEULE fois par session.
- Un IV est utilisé pour EVITER LA REPETITION pendant le cryptage des données : son l'utilisation permet à une même lettre d'être représentée par une séquence particulière dans le premier cas, puis représentée par une séquence binaire complètement différente dans le second cas, et ainsi de suite…

L'IV permet donc d'ajouter un peu plus de confusion en offrant la possibilité d'obtenir deux messages chiffrés différents correspondant à un seul et unique message. Ainsi, lors d'un échange, Alice et Bob doivent partager :

- Un algorithme commun : DES, Triple-DES, AES, RC4, RC5…
- La clé secrète.
- Le vecteur d'initialisation (IV).

REMARQUE : ne pas confondre avec le HACHAGE !

Une FONCTION DE HACHAGE (MD5, SHA-XXX…) convertit un grand ensemble en un plus petit ensemble : l'EMPREINTE. Il est impossible de la déchiffrer pour revenir à l'ensemble d'origine ; ce n'est donc pas une technique de chiffrement. L'empreinte d'un message ne dépasse généralement pas 256 bits (maximum 512 bits pour SHA-512).

Pour renforcer la sécurité, la technique du « salage » consiste à ajouter une chaîne de caractères à l'information avant le hachage. EXEMPLE : soit le hash (SHA256) d'un mot de passe et enregistré dans une BD :

3d6155e85ccd2c725d39827ad97d23627265a4bd1b84561b5b145dea5d4e5108

Si un hacker réussit à récupérer le haché du mot de passe, il va essayer de retrouver sa valeur initiale (en clair), à condition de connaître la fonction de hachage utilisée et en utilisant des techniques telles que :

- **L'attaque par dictionnaire** : il va hacher pleins de mots d'un « dictionnaire" de mots de passe » couramment utilisés.
- **L'attaque par force brute** : il va hacher toutes les combinaisons possibles de lettres/chiffres/symboles : c'est très long mais comme la plupart des fonctions de hachage sont très rapides, cela fonctionne rapidement sur les mots de passe courts.
- **Les tables arc-en-ciel** : elles contiennent toutes les combinaisons possibles des mots de passe en fonction des paramètres choisis à la génération et disposent de correspondances directes mot de passe <-> hash correspondant.

Cependant, si l'on a utilisé un « grain de sel », donnée informatique différente pour chaque utilisateur, soit connue par l'administrateur, mais généralement aléatoire, celle-ci sera être ajoutée au mot de passe avant hachage. Exemple ;

SEL = dac6595c04dda81 + mot de passe = f1c829b4039db06ef077637b8c5c25544810c557b82d40c1e22c5f2cc2889b5e

- Cette fois-ci le hash et le sel seront sauvegardés (pas au même endroit si possible).
- Lors de la connexion au site, le serveur récupère le mot de passe saisi, trouve le sel correspondant à l'utilisateur et applique sa fonction de hachage avec sel. Si le haché ainsi calculé est identique à celui stocké sur le serveur, le mot de passe saisi est correct.
- Si le hacker n'a pas accès au sel mais uniquement aux empreintes de mot de passe, il lui devient impossible d'attaquer par dictionnaire, brute-force ou en encore table arc-en-ciel.