Chapitre 1 Introduction

Le matériel informatique se répand de plus en plus. En effet, d'une part le matériel est accessible à un prix très abordable, et d'autre part, les logiciels tendent à se simplifier (au niveau de l'utilisation) et permettent une prise en main rapide. D'un autre cote, les entreprises, elles aussi informatisées, nécessitent un réseau sécurise pour le transfert des données, que ce soit entre les machines de cette entreprise, ou avec des machines externes, distantes de plusieurs milliers de kilomètres.
Avec cette popularité grandissante des réseaux, de nombreuses menaces les accompagnent. Parmi celles-ci, on trouve diverses catégories :

- Les menaces accidentelles
- Les menaces intentionnelles :
    - passives
    - actives

Les menaces accidentelles ne supposent aucune préméditation. Dans cette catégorie, sont repris les bugs logiciels, les pannes matérielles, et autres défaillances "incontrôlables".
Les menaces intentionnelles quant a elles, reposent sur l'action d'un tiers désirant s'introduire et relever des informations. Dans le cas d'une attaque passive, l'intrus va tenter de dérober les informations par audit, ce qui rend sa détection relativement difficile. En effet, cet audit ne modifie pas les fichiers, ni n'altère les systèmes. Dans le cas d'une attaque active, la détection est facilitée, mais il peut être déjà trop tard lorsque celle-ci a lieu. Ici, l'intrus aura volontairement modifié les fichiers ou le système en place pour s'en emparer.
Les menaces actives appartiennent principalement à quatre catégories (illustrées à la figure 1.1) :

- Interruption = problème lié a la disponibilité des données
- Interception = problème lié a la confidentialité des données
- Modification = problème lié à l'intégrité des données
- Fabrication = problème lie à l'authenticité des données

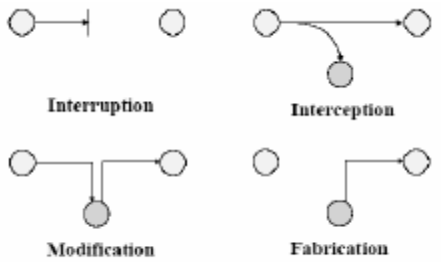


FIG. 1.1 – Types de menaces actives

Les auteurs de ces attaques sont notamment les hackers (agissant souvent par défi personnel), les concurrents industriels (vol d'informations concernant la stratégie de l'entreprise ou la conception de projets), les espions.
Pour empêcher ces attaques, la sécurité doit être omniprésente. Aussi bien lorsque les données ne sont pas utilisées que lorsqu'elles transitent sur le réseau. L'accès aux données doit donc être contrôlé par un système de vérification, axe sur l'audit, l'identification de l'utilisateur (par mot de passe, cartes magnétiques, empreintes digitales, ...) et ses droits de modification (selon les autorisations qui lui sont accordées).

Lorsque les données sont en transit, il faudra à la fois permettre la confidentialité (par l'intermédiaire d'un chiffrement), l'intégrité (fonctions de hash), l'authentification des parties, et la non-répudiation des parties (une des deux parties (ou les deux) ne peu(ven)t renier son (leur) message).

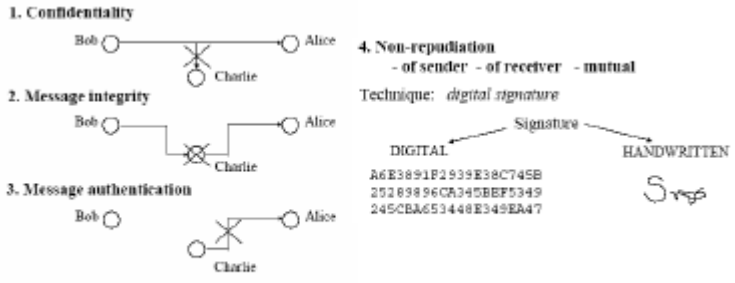


FIG. 1.2 – Représentation des propriétés de sécurité

1.2. VOCABULAIRE DE BASE

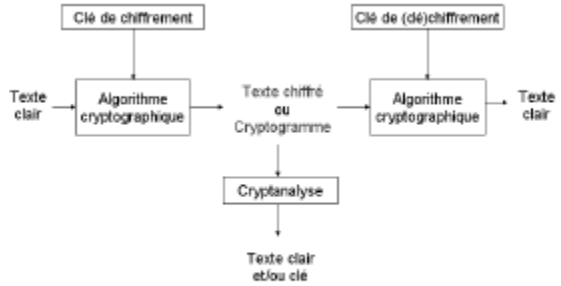


FIG. 1.3 – Protocole de chiffrement

Cryptologie : Il s'agit d'une science mathématique comportant deux branches : la cryptographie et la cryptanalyse.

Cryptographie : La cryptographie est l'étude des méthodes donnant la possibilité d'envoyer des données de manière confidentielle sur un support donné.

Chiffrement : Le chiffrement consiste à transformer une donnée (texte, message, ...) afin de la rendre incompréhensible par une personne autre que celui qui a crée le message et celui qui en est le destinataire. La fonction permettant de retrouver le texte clair à partir du texte chiffre porte le nom de déchiffrement.

Texte chiffré : Appelé également cryptogramme, le texte chiffré est le résultat de l'application d'un chiffrement à un texte clair.

Clef : Il s'agit du paramètre impliqué et autorisant des opérations de chiffrement et/ou déchiffrement. Dans le cas d'un algorithme symétrique, la clef est identique lors des deux opérations. Dans le cas d'algorithmes asymétriques, elle diffère pour les deux opérations.

Cryptanalyse : Opposée à la cryptographie, elle a pour but de retrouver le texte clair à partir de textes chiffres en déterminant les failles des algorithmes utilisés.

Cryptosystème : Il est défini comme l'ensemble des clés possibles (espace de clés), des textes clairs et chiffres possibles associes à un algorithme donné.

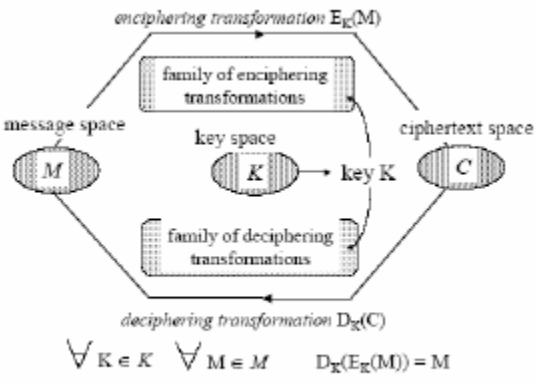


FIG. 1.4 – Schéma d'un cryptosystème

1.3 Notations

En cryptographie, la propriété de base est que

$$M = D(E(M))$$

où

- M représente le texte clair,
- C est le texte chiffre,
- K est la clé (dans le cas d'un algorithme à clé symétrique), E_k et D_k dans le cas d'algorithmes asymétriques,
- E(x) est la fonction de chiffrement, et
- D(x) est la fonction de déchiffrement.

Ainsi, avec un algorithme à clef symétrique,

$$M = D(C) \text{ si } C = E(M)$$

1.4 Principe de Kerckhoff

La sécurité du chiffre ne doit pas dépendre de ce qui ne peut pas être facilement changé.
En d'autres termes, aucun secret ne doit résider dans l'algorithme mais plutôt dans la clé. Sans celle-ci, il doit être impossible de retrouver le texte clair à partir du texte chiffré. Par contre, si on connait K, le déchiffrement est immédiat.

Remarque : il faut distinguer les termes "Secret" et "Robustesse" d'un algorithme. Le secret de l'algorithme revient à cacher les concepts de celui-ci, ainsi que les méthodes utilisées (fonctions mathématiques).
La robustesse quant à elle désigne la résistance de l'algorithme à diverses attaques.

1.5 Algorithme public et algorithme secret

Selon l'endroit ou réside le secret, on peut parler d'algorithme secret ou d'algorithme publié. Chacun possède ses atouts et inconvénients.

1.5.1 Algorithme secret

La cryptanalyse, souvent basée sur le secret de la clé, doit ici en plus retrouver l'entièreté de l'algorithme (mécanisme de récupération). Souvent, de tels algorithmes sont utilisés par un plus petit nombre d'utilisateurs. Et comme souvent dans ce cas, moins il y a de monde l'utilisant, moins il y a d'intérêts à le casser. De tels algorithmes sont rarement distribués par delà les frontières, afin de garder un nombre d'utilisateurs restreint.

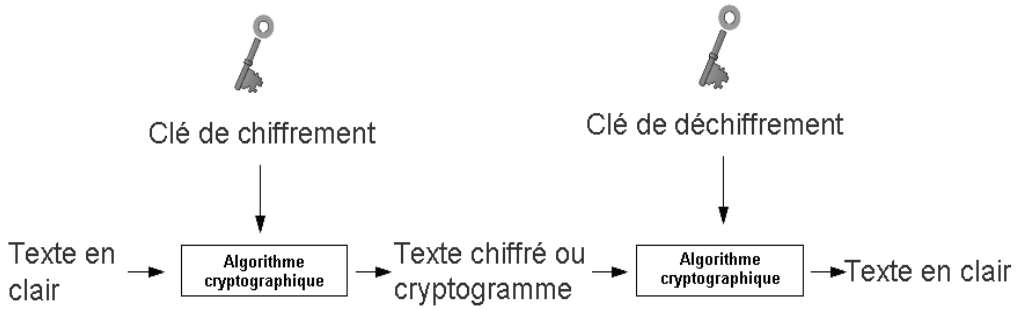
1.5.2 Algorithme public

Puisque l'algorithme est public, tout le monde a le droit de l'explorer. Ainsi, les failles (laissées intentionnellement ou non par les concepteurs) peuvent être plus facilement découvertes. La sécurité en est donc améliorée. Comme la publication est autorisée, il n'est pas nécessaire de chercher à protéger le code contre le reverse-engineering. Cette publication permet d'étendre les travaux sur l'algorithme au niveau mondial. Toute une série d'implémentations logicielles peuvent donc être réalisées. Tout le monde utilise la même version publique ce qui permet une standardisation générale. En conséquence, on préférera les algorithmes publiés, souvent plus surs pour les raisons explicitées ci-dessus. Si un bon nombre de gens futés n'ont pas résolu un problème, alors il ne sera probablement pas résolu de si tôt.

1.6 Les principaux concepts cryptographiques

1.6.1 Crypto système à clé symétrique

Caractéristiques :



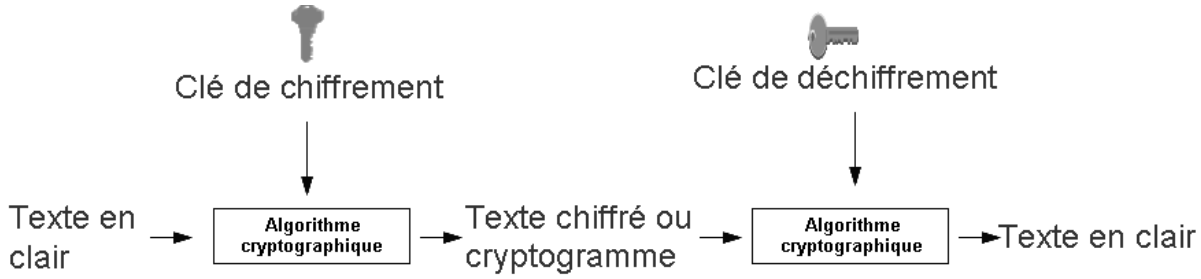
Chiffrement symétrique

- Les clés sont identiques : $K_E = K_D = K$,
- La clé doit rester secrète.
- Les algorithmes les plus répandus sont le DES, AES, 3DES, ...
- Au niveau de la génération des clés, elle est choisie aléatoirement dans l'espace des clés, et donc, dans lequel le secret réside dans la clé.
- Ces algorithmes sont bases sur des opérations de transposition et de substitution des bits du texte clair en fonction de la clé,
- La taille des clés est souvent de l'ordre de 128 bits. Le DES en utilise 56, mais l'AES peut aller jusque 256,
- L'avantage principal de ce mode de chiffrement est sa rapidité,
- Le principal désavantage réside dans la distribution des clés : pour une meilleure sécurité, on pratiquera à l'échange de manière manuelle. Malheureusement, pour de grands systèmes, le nombre de clés peut devenir conséquent. C'est pourquoi on utilisera souvent des échanges sécurisés pour transmettre les clés. En effet, pour un système à N utilisateurs, il y aura N*(N-1)/2 paires de clés.

1.6.2 Crypto système à clé publique

Caractéristiques :

- Une clé publique P_K (symbolisée par la clé verticale),
- Une clé privée secrète S_K (symbolisée par la clé horizontale),
- Propriété : La connaissance de P_K ne permet pas de déduire S_K,
- $D_{SK}(E_{PK}(M)) = M$,



Chiffrement asymétrique

- L'algorithme de cryptographie asymétrique le plus connu est le RSA,
- Le principe de ce genre d'algorithme est qu'il s'agit d'une fonction unidirectionnelle à trappe. Une telle fonction a la particularité d'être facile à calculer dans un sens, mais difficile voire impossible dans le sens inverse. La seule manière de pouvoir réaliser le calcul inverse est de connaitre une trappe. Une trappe peut par exemple être une faille dans le générateur de clés. Cette faille peut être soit accidentelle ou intentionnelle de la part du concepteur.
- Les algorithmes se basent sur des concepts mathématiques tels que l'exponentiation de grands nombres premiers (RSA), le problème des logarithmes discrets (ElGamal), ou encore le problème du sac à dos (Merkle-Hellman).
- La taille des clés s'étend de 512 bits à 2048 bits en standard. Dans le cas du RSA, une clé de 512 bits n'est plus sure au sens "militaire" du terme, mais est toujours utilisable de particulier à particulier.
- Au niveau des performances, le chiffrement par voie asymétrique est environ 1000 fois plus lent que le chiffrement symétrique.
- Cependant, à l'inverse du chiffrement symétrique ou le nombre de clés est le problème majeur, ici, seules n paires sont nécessaires. En effet, chaque utilisateur possède une paire (S_K, P_K) et tous les transferts de message ont lieu avec ces clés.
- La distribution des clés est grandement facilitée car l'échange de clés secrètes n'est plus nécessaire. Chaque utilisateur conserve sa clé secrète sans jamais la divulguer. Seule la clé publique devra être distribuée.

1.6.3 Fonction de hachage

Il s'agit de la troisième grande famille d'algorithmes utilises en cryptographie. Le principe est qu'un message clair de longueur quelconque doit être transformé en un message de longueur fixe inférieure à celle de départ. Le message réduit portera le nom de "Haché" ou de "Condensé". L'intérêt est d'utiliser ce condensé comme empreinte digitale du message original afin que ce dernier soit identifie de manière univoque. Deux caractéristiques (théoriques) importantes sont les suivantes :

1. Ce sont des fonctions unidirectionnelles :
   A partir de H(M) il est impossible de retrouver M.
2. Ce sont des fonctions sans collisions :
   A partir de H(M) et M il est impossible de trouver M≠M tel que H(M)=H(M).

1.6.4 Protocoles cryptographiques

Des que plusieurs entités sont impliquées dans un échange de messages sécurisés, des règles doivent déterminer l'ensemble des opérations cryptographiques à réaliser, leur séquence, afin de sécuriser la communication. C'est ce que l'on appelle les protocoles cryptographiques.

Il y a trois propriétés fondamentales :

- Confidentialité
- Intégrité
- Authentification

1.6.4.1 Confidentialité

Elle est amenée par le chiffrement du message. Dans le cas de systèmes à clés symétriques, la même clé est utilisée pour $E_K(M)$ et $D_K(C)$. Ce type de chiffrement nécessite un échange sur préalable de la clé K entre les entités A et B.
Comme dit précédemment, à l'aide d'un crypto système asymétrique, cet échange préalable n'est pas nécessaire. Chaque entité possède sa propre paire de clés. On aura donc la paire P_{KA}, S_{KA} pour l'entité A et la paire P_{KB}, S_{KB} pour l'entité B.
En marge de ces deux systèmes, existe également un système appelé "hybride", reposant comme son nom l'indique sur les deux systèmes précédents. Par l'intermédiaire du système à clé publique, on sécurise l'échange de la clé K. Ensuite, les deux parties ayant acquis de manière sécurisée cette clé de chiffrement K , on utilisera le système à clé symétrique pour chiffrer le message.

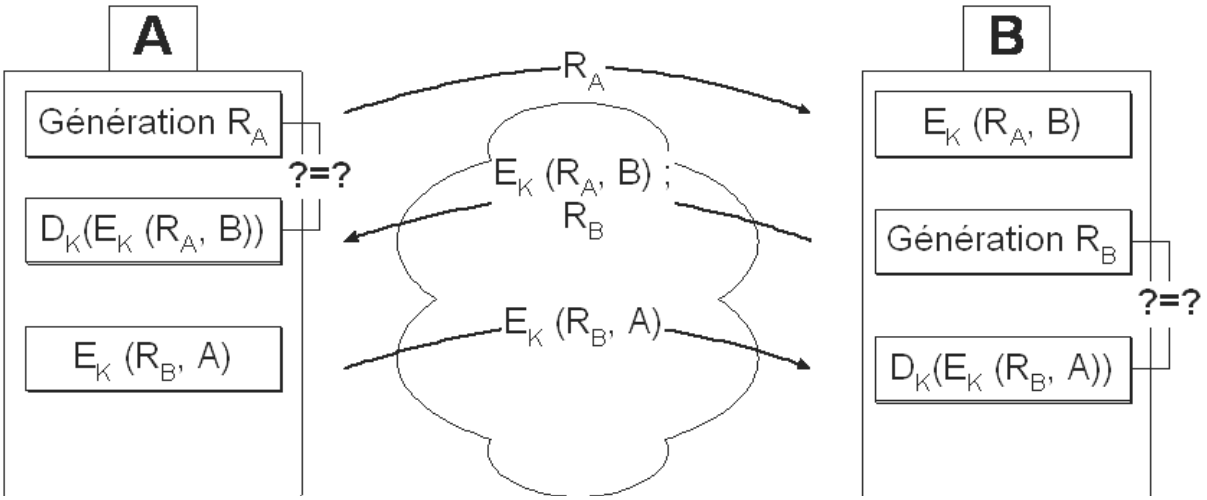
1.6.4.2 Intégrité

Il faut ici vérifier si le message n'a pas subi de modification durant la communication. C'est ici qu'interviennent les fonctions de hachage.

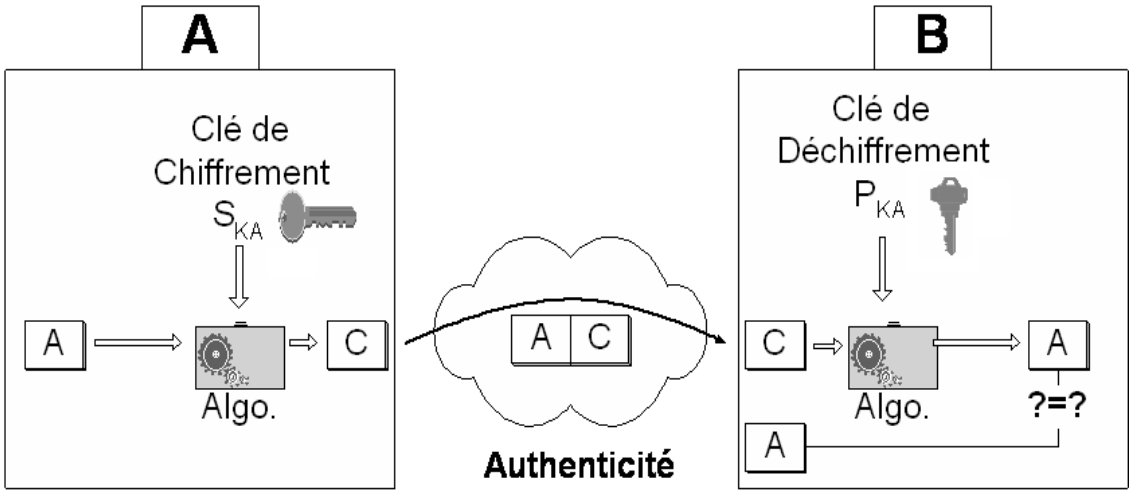
1.6.4.3 Authentification

Elle a lieu à plusieurs niveaux.

- Au niveau des parties communicantes, dans le cas d'un système symétrique ou asymétrique. Ra est une nonce (p. ex. nombre aléatoire), propre à l'utilisateur A. Les lettres A et B représentent des identificateurs personnels.



A la seconde figure, la clé de chiffrement utilisée est bien la clé privée. Comme le propriétaire de cette clé est le seul à la connaitre, cela prouve qu'il est bien la personne ayant chiffré le message. Attention, dans cet exemple, seule l'authentification est souhaitée. Le message envoyé pourra être lu par toute personne possédant la clé publique, c'est-à-dire, n'importe qui. La confidentialité est ici nulle.



- Au niveau du message
    - Par l'utilisation d'un MAC (Message Authentication Code) généré à l'aide d'un crypto système à clé symétrique ou le MAC est constitué des derniers digits de C, ou généré à l'aide d'une fonction de hachage, la clé secrète K utilisée étant partagée par les deux entités A et B. Dans les deux cas, l'authentification repose sur l'utilisation de la clé K.

A noter que par l'intermédiaire du MAC, il y a aussi une vérification de l'intégrité du message.

    - Par l'utilisation d'une signature digitale. Parmi les propriétés remarquables de ces signatures, on peut dire qu'elles doivent être authentiques, infalsifiables, non-réutilisables, non-répudiables, et inaltérables.