République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université Dr Moulay Tahar de Saida
Faculté de Technologie
Département d'Electronique
Spécialité : Télécommunication

Mémoire de fin d'études pour l'obtention du diplôme de master en Télécommunication

Option : Réseaux et télécommunication

Thème
La cryptographie à base d'ADN : application dans un environnement IoT

Présenté par :

MOUAZER saadia

FEZZA faiza

Soutenu le 22/09/ 2020, devant le jury composé de:

Dr.M Damou	président
Dr.M Bouyeddou	examinateur
Dr.k Benyahia	encadreur

Année universitaire 2019/2020

Remerciements

« Aucun travail ne s'accomplit de la solitude »

Michel. Beaud

Nous remercions, tout d'abord Dieu le tout puissant de nous avoir donné la force et d'avoir illuminé notre parcours et guidé nos pas vers le chemin du savoir

Nous adressons nos vifs et sincères remerciements à notre directeur de recherche Monsieur Kadda benyahia de nous avoir prodigué conseils aussi pour sa gentillesse et surtout sa confiance qu'il nous a accordée.

Nos remerciements s'adressent également aux membres de jury qui ont accepté d'évaluer ce mémoire et de nous faire part de leurs critiques constructives qui contribueront surement au perfectionnement de ce travail

Nous ne pouvons omettre d'adresser un grand merci à tous nos enseignants qui nous ont suivis et encouragés tout au long de nos cursus et dans la réalisation de ce mémoire.

Ce mémoire n'aurait jamais vu le jour sans la coopération des enseignants de labo

Ainsi nous sommes très reconnaissantes à tous ceux qui ont contribué de près ou de loin à la réalisation de ce modeste travail.

Dédicace

Je dédie ce modeste travail

** A celui qui m'a donné le gout de la recherche et des études qui est mon cher papa*

** A mes chers parents pour leurs conseils, leurs sacrifices et leurs prières*

** A mon petit frère Zinou et ma sœur Karima*

** A toute ma grande famille*

** A mes camarades de la promo 2019/2020*

** A tous ceux et celles qui me connaissent et qui m'aiment*

Mouazer Saadia

Dédicace

Je dédie ce modeste travail

** A celui qui m'a donné le gout de la recherche et des études qui est mon cher papa*

** A mes chers parents pour leurs conseils, leurs sacrifices et leurs prières*

** A mes deux frères et ma sœur*

** A toute ma grande famille*

** A mes camarades de la promo 2019/2020*

** A tous ceux et celles qui me connaissent et qui m'aiment*

Fezza Faiza

Table des figures

Table des tableaux

Sommaire

CHAPITRE 4 : Implémentations et Résultats

Introduction générale

Etant donné que le marché mondial est concurrentiel et la croissance des télécommunications et leurs sécurité a exponentiellement changé au cours de ces dernières années. En pleine mutation la majorité des secteurs se sont digitalisés pour donner un dynamisme et jouer un rôle dans le développement d'un pays.

Depuis l'antiquité, on cherche à envoyer des données afin qu'aucune personne autre que le destinataire ne puisse les lire, durant cette époque l'une des utilisations les plus connues est le chiffre de César, nommé en référence à Jules César qui l'utilisait pour ses communications secrètes. Mais la cryptographie est bien antérieure à cela : le plus ancien document chiffré est une recette secrète de poterie qui date du 16eme siècle av. J.-C., qui a été découverte dans l'actuelle Irak.

La cryptographie est un domaine clé des systèmes électroniques permettant de protéger les messages tout on assurant leur confidentialité, authenticité et intégrité. Cette discipline consiste à sécuriser les données par une méthode de chiffrement qui utilise la clé pour rendre le message incompréhensible aux personnes extérieures, Par contre le cryptage sert à chiffrer les messages sans utiliser la clé, la fonction permettant de retrouver le texte clair à partir du texte chiffré porte le nom de déchiffrement, lorsqu'on parle de décryptage c'est-à-dire retrouver le texte clair sans connaître la clef de déchiffrement

Plusieurs études ont été menées sur les deux principales catégories de cryptographie classée selon le type de clés de sécurité utilisé pour chiffrer et déchiffrer les données : la cryptographie symétrique ou à clés secrètes (par blocs et par flots) comme AES DES, la cryptographie asymétrique ou à clés publiques comme RSA

Dans un temps passé, il n y'avait pas de relation entre la cryptographie et la biologie moléculaire, après une documentation approfondie sur la recherche en biotechnologie moderne et en informatique de l'ADN ou les chercheurs sont arrivés à trouver un lien entre elles .nous voulons orienter notre travail sur La cryptographie par ADN qui est un nouveau domaine et qui apporte de nouvelles informations sur la sécurisation des informations en utilisant la génétique et le calcul biomoléculaire.

Notre choix est justifié par trois raisons :

La première, nous savons que le matériel génétique tel que l'ADN peut être utilisé comme un vaste espace de stockage. Car, l'ADN est un vecteur naturel d'informations codées par un alphabet à 4 lettres: A, C, G et T appelé nucléotide.

La seconde, nous avons trouvé utile que L'information secrète placée dans une molécule de l'ADN, peut être cachée dans d'autres. Parmi les avantages de l'utilisation de l'ADN en cryptographie, il y a ; l'espace de stockage, la puissance et la complexité de calcul et la génération des clés cryptographiques à partir de ses longues séquences.

Et la troisième, lorsque nous avons parcouru la documentation sur L'Internet of Things (IoT) qui consiste principalement à connecter des objets physiques à l'Internet et avec le temps et que le terme a évolué pour englober maintenant tout l'écosystème des objets connectés, des fabricants de capteurs, des éditeurs de logiciels, des intégrateurs…

Vu cet éclectisme qui fait la richesse du sujet, nous avons décidé de se pencher sur cette question et d'entamer une recherche qui tenterait de mettre en exergue la place de l'ADN dans la cryptographie.

Dans notre modeste travail, nous allons traiter notre pratique sur une machine qui est le « Raspberry pi » et faire la cryptographie par ADN c'est-à-dire, on va essayer de chiffrer des messages de différentes tailles et calculer leurs temps de chiffrement et déchiffrement.

Cette situation nous a ramenées à se poser le questionnement suivant :

- Comment et dans quelle mesure la cryptographie par ADN offre une protection des données dans une communication ?

Pour répondre à cette question, nous avons prévu de mettre en place un dispositif qui consiste à la cryptographie par ADN dans un environnement IoT, en particulier au Raspberry pi en proposant de chiffrer des données de différentes tailles.

-Nous supposons que cela offre une bonne protection en tenant compte des caractéristiques limitées de raspberry.

À travers notre proposition basée sur la cryptographie symétrique, nous voulons atteindre l'objectif principal qui est d'assurer la protection des informations.

Pour mettre en exergue les théories appropriées de notre mémoire ainsi que notre analyse, notre travail s'articule en deux parties :

Une partie théorique, dans laquelle nous allons présenter dans le premier chapitre l'historique et quelques définitions des notions descriptives, sur la cryptographie. Dans le $2^{ème}$ chapitre, on a esquissé tout ce qui concerne la molécule d'ADN, ses fonctions et sa relation avec la cryptographie et enfin, dans le $3^{ème}$ chapitre, on a abordé les composants, les technologies et les domaines d'applications de l'IOT en général.

Dans la partie pratique, nous allons faire un programme de cryptographie par ADN en langage PYTHON et l'exécuter dans un raspberry pi en calculant le temps de chiffrement/déchiffrement en fonction de la taille des données.

Pour finir, nous ferons une apostrophe en conclusion générale de notre démarche de recherche, nos analyses, nos résultats ainsi que nos perspectives dans les études ultérieures de la cryptographie par ADN dans un environnement IoT.

Chapitre 01

Etat de l'art autour de la cryptographie

Introduction

La sécurité de l'information demeure parmi les sciences les plus intéressantes durant l'histoire jusqu'à nos jours. Les méthodes d'enregistrer l'information n'ont pas trop changé typiquement sur des supports en papier ou bien magnétique envoyée par la suite via des systèmes de télécommunications. Ce qui a par contre beaucoup évolué, c'est la copie et la modification illégale de ces données, d'où la nécessité de les protéger contre les menaces accidentelles ou intentionnelles. Cette protection peut être assurée par la cryptographie [1]

Au-delà, nous voyons se dégager un certain nombre de fonctionnalités principales indispensables :

Confidentialité

L'intégrité des données

L'authentification, l'identification

Signature [2]

Le besoins croissants de sécurité au sein de l'environnement matériel et technologique de la cryptographie comme (les ordinateurs, les réseaux ordinateurs les moyens modernes de télécommunication, les cartes à puce) a motivé un développement rapide de la cryptographie, développement impliquant de manière imbriquée les mathématiques, informatique ainsi que la physique.

A côté du terme ancien de cryptographie, limité à l'origine au chiffrement des messages on a introduit récemment, de manière à préciser la terminologie et rendre compte de l'évolution récente du domaine, la dénomination générale de cryptologie pour designer la partie de la sécurité des systèmes d'information qui s'occupe d'assurer, ou au contraire de compromettre les grandes fonctions de sécurité.

Ainsi la cryptologie a un double aspect : d'une part la cryptographie qui consiste à construire des outils propre à assurer des fonctionnalités de sécurité, d'autre part la cryptanalyse qui consiste à s'attaquer aux outils en question pour en trouver les faiblesses. [2]

La cryptographie est divisée en deux grandes catégories toutes deux indispensables, car employées dans des usages différents : le chiffrement symétrique et le chiffrement asymétrique. Dans le système à clé secrète ou symétrique, un expéditeur et un destinataire partagent une même clé secrète pour le chiffrement et déchiffrement et doit rester secrète de tous les ennemis [3]. Par contre dans le système à clé publique ou asymétrique la clé de chiffrement soit différente de la clé de déchiffrement. C'est-à-dire, que n'importe qui peut utiliser la clé de chiffrement pour chiffrer un message appelée (clé publique), mais seul celui qui possède la clé du déchiffrement peut déchiffrer le message chiffré résultant appelée (clé privée). [4]

1.1 Principes de base

1.1.1 Terminologies

Dans ce chapitre on va définir précisément certaines notions couramment utilisées en cryptologie

1.1.1.1 Expéditeur et destinataire

Supposons qu'un expéditeur veut envoyer un message à un destinataire. Cet expéditeur veut envoyer le message de manière sure : il veut s'assurer qu'aucune oreille indiscrète ne puisse s'informer du message

1.1.1.2 La cryptologie

La cryptologie est la science des messages secrets et ses pratiquants sont appelés cryptologues, cette discipline se décompose en cryptographie et cryptanalyse [5] Le mot cryptologie est souvent utilisé comme synonyme de cryptographie.

1.1.1.3 La cryptographie

C'est l'étude des systèmes mathématiques informatiques et électroniques propres à résoudre les problèmes de confidentialité et l'authentification [2] .Elle est aussi l'art et la science de garder le secret de message, pratiquée par des cryptographes [4].

La cryptographie est l'étude des méthodes permettant de convertir des informations "en clair" en informations codées, c'est à dire non compréhensibles c'est ce qu'on appelle « chiffrement », puis, à partir de ces informations codées, de restituer les informations originales appelé « déchiffrement » [6]

La fonction mathématique utilisée pour le chiffrement et le déchiffrement est l'algorithme cryptographique, qui dépend d'un paramètre appelé clef [4].

Le texte en clair est noté M cela peut être une suite de bits, un fichier de texte, un enregistrement de voix numérisé, ou une image vidéo numérique. Du point de vue de l'ordinateur M n'est rien d'autres que de l'information binaire. Le texte en clair peut être transmis ou stocké. Dans tous les cas, M est le message à chiffrer.

Le texte chiffré est noté C. c'est aussi de l'information binaire, parfois de la même taille que M, parfois plus grand. La fonction de chiffrement, notée E, transforme M en C. ce qui en notation mathématique s'écrit :

$$E(M)=C$$

La fonction inverse, notée D, de déchiffrement transforme C en M :

$$D(C)=M$$

Comme le but de toutes ces opérations n'est rien d'autre que de retrouver le message en clair à partir de la version chiffré de ce même message, l'identité suivante doit être vérifiée :

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

Les opérations de chiffrement et de déchiffrement utilisent toutes les deux une clef notée k. cette clef peut prendre une des valeurs parmi un grand nombre de valeurs possibles, aussi les fonctions s'écrivent de la manière suivante :

$$Ek(M)=C$$

$$Dk(C)=M$$

Ces fonctions vérifient la propriété suivante :

$$Dk\ (Ek\ (M))\ =M$$

Figure1.1 : Chiffrement et déchiffrement avec une clef

Certains algorithmes utilisent des clefs différentes pour le chiffrement et le déchiffrement. Dans ce cas, la clef de chiffrement, notée k1, est différente de la clef de déchiffrement, notée k2 .les relations suivantes décrivent un tel cryptosystème :

$$Ek1\ (M)\ =C$$

$$Dk2(C)\ =M$$

$$Dk2\ (Ek1\ (M))\ =M$$

Avec ces algorithmes, toute la sécurité réside dans la (ou les) clef (s), et non dans les détails de l'algorithme. Ceci implique que l'algorithme peut être publié et analysé [4]

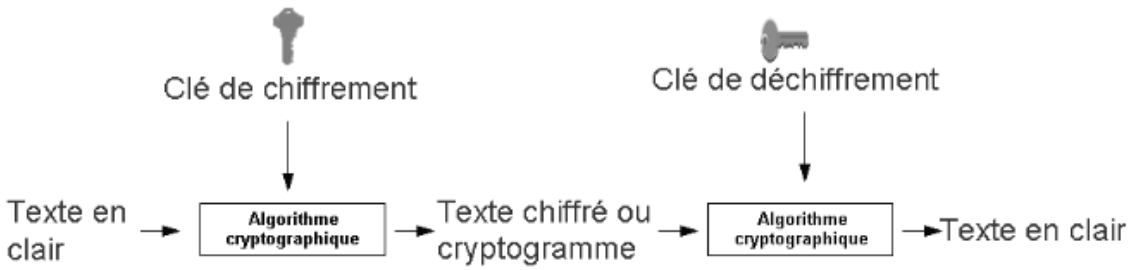


Figure 1.2: Chiffrement et déchiffrement avec deux clefs

1.1.1.4 Chiffrement et déchiffrement

Le processus de transformation d'un message clair de telle manière à le rendre incompréhensible est appelé chiffrement(ou encryption) .le résultat de ce processus de chiffrement est appelé texte chiffré (ou encore cryptogramme). Le processus de reconstruction du texte en clair à partir du texte chiffré est appelé déchiffrement (ou décryptage). [4]

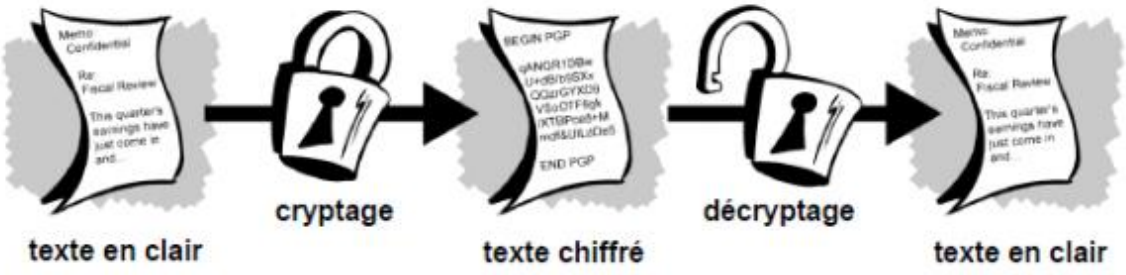


Figure1.3 : Protocole de chiffrement et déchiffrement

1.1.1.5 Clair (ou message clair)

Version intelligible d'un message tel qu'il se présentait avant tout chiffrement [7]

1.1.1.6 Texte chiffré (cryptogramme)

Message écrit à l'aide d'un système de chiffrement [7]

1.1.1.7 Clef

Dans un système de chiffrement, elle correspond à un nombre, un mot, une phrase, etc., qui permet, grâce à l'algorithme de chiffrement, de chiffrer ou de déchiffrer un message [7].

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, On distingue généralement deux types de clefs:

- Les clés symétriques : il s'agit de clés utilisées pour le chiffrement ainsi que pour le déchiffrement. On parle alors de chiffrement symétrique.

- Les clés asymétriques : il s'agit de clés utilisées dans le cas du chiffrement asymétrique (aussi appelé chiffrement à clé publique). Dans ce cas, une clé différente est utilisée pour le chiffrement et pour le déchiffrement.

1.1.1.8 La cryptanalyse

Le but principale de la cryptographie est de préserver le texte en clair de l'indiscrétion des espions (aussi appelé adversaires, attaquants, cryptanalystes ou tout simplement ennemis).

La cryptanalyse est la science de la reconstitution du texte en clair sans connaitre la clef.

On peut aussi appeler cette action le « cassage ».une cryptanalyse réussie peut fournir soit le texte en clair, soit la clef. Une tentative de cryptanalyse est appelée attaque.

Il ya quatre types génériques d'attaques cryptanalytiques

1/ l'attaque à texte chiffré seulement : consiste de retrouver le texte en clair du plus grand nombre de message possible ou mieux encore de trouver la ou les clefs qui on été utilisées pour chiffré le message ce qui permettrait de déchiffrer d'autres messages chiffrés avec ces mêmes clefs

2/L'attaque à texte en clair connue : le cryptanalyste a non seulement accès aux textes chiffrés de plusieurs messages mais aussi aux textes en clair correspondants. La tâche est de retrouver la ou les clef(s) utilisées pour chiffrer ces messages ou un algorithme qui permet de déchiffrer n'importe quel nouveau message chiffré avec la même clef.

3/ l'attaque à texte en claire choisi : cette attaque est plus efficace que l'attaque à texte en clair connu car le cryptanalyste peut choisir des textes en clair spécifiques qui donneront plus d'informations sur la clef.

4/L'attaque à texte chiffré choisi : le cryptanalyste peut choisir différents textes chiffrés à déchiffrer les textes déchiffrés lui sont alors fournis, sa tâche est de retrouver la clef [4]

1.1.1.9 Décryptement

Opération qui consiste à retrouver le clair sans disposer des clefs théoriquement nécessaires. Il ne faut pas confondre déchiffrement et décryptement [7] c'est l'action faite par les cryptanalystes.

1.1.1.10 Cryptosystème

Il est défini comme l'ensemble des clés possibles (espace de clés), des textes clairs et chiffrés possibles associés à un algorithme donné [4]

L'algorithme est en réalité un triplet d'algorithmes :

– l'un générant les clés K

– un autre pour chiffrer M

– un troisième pour déchiffrer C

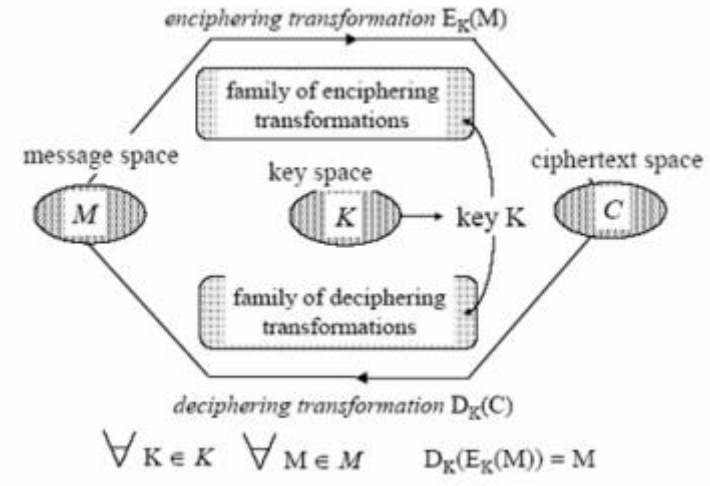


Figure 1.4 : Schéma d'un cryptosystème

1.1.1.11 Stéganographie

Branche particulière de la cryptographie qui consiste non pas à rendre le message inintelligible, mais à le camoufler dans un support (un texte, une image, les mailles d'un tricot, etc.) de manière à masquer sa présence [7]

1.1.1.12 Chiffre

Manière secrète d'écrire un message à transmettre, au moyen de caractères et de signes disposés selon une convention convenue au préalable. Les deux grandes familles de chiffres sont les substitutions et les transpositions [7].

1.1.1.13 Scytale

Les Grecs emploient un dispositif appelé la scytale - un bâton autour du quel une bande de cuir longue et mince était enveloppée et sur laquelle on écrivait le message. Le cuir était ensuite porté comme une ceinture par le messager. Le destinataire avait un bâton identique permettant d'enrouler le cuir afin de déchiffrer le message [7]

Figure 1.5 : Scytale

1.2 Les objectifs de la cryptographie

Le but de la cryptographie c'est assurer la sécurité des communications transmises sur un canal public en présence d'adversaires

- Adversaire passif : Écoute les communications

- Adversaire actif : capable d'écrire, modifier et effacer des informations passant sur le canal de communication

Les objectifs de la cryptographie moderne sont plus complexes et plus nombreux, mais on peut en distinguer quatre principaux :

1.2.1 Confidentialité, secret, chiffrement

Les données doivent rester inintelligibles à toute personne non autorisée. Plus précisément la confidentialité est la propriété qu'une information n'est ni disponible ni divulguée aux personnes, entités ou processus non autorisés [2]

1.2.2 Intégrité des données

On doit éviter que les données transmises soient modifiées ou forgées par un adversaire. Plus précisément l'intégrité est la prévention d'une modification non autorisée de l'information. L'intégrité du système et de l'information garantit que ceux-ci ne sont modifiés que par une action volontaire et légitime .les attaques contre l'intégrité sont appelés substitutions [2].

1.2.3 Authentification

Consiste à vérifier l'identité des différents éléments impliqués dans un dialogue. L'émetteur est sur de l'identité du destinataire c'est à dire que seul le destinataire pourra prendre connaissance du message car il est le seul à disposer de la clef de déchiffrement, le receveur est sur de l'identité de l'émetteur. Les attaques conte l'authentification sont appelées mascarades [2].

1.2.4 Non-répudiation (signature)

C'est un mécanisme qui empêche de nier un contrat. La non répudiation consiste à prouver par exemple qu'un message a bien été émis par son expéditeur ou a bien été reçu par son destinataire.

L'auteur ne peut pas nier, d'avoir écrit ou transmis un message [2].

Non répudiation se décompose en trois:

1- non-répudiation d'origine l'émetteur ne peut nier avoir écrit le message et il peut prouver qu'il ne l'a pas fait si c'est effectivement le cas.

2- non-répudiation de réception le receveur ne peut nier avoir reçu le message et il peut prouver qu'il ne l'a pas reçu si c'est effectivement le cas.

3- non-répudiation de transmission l'émetteur du message ne peut nier avoir envoyé le message et il peut prouver qu'il ne l'a pas fait si c'est effectivement le cas.

On peut regarder ces quatre qualités du point de vue de l'émetteur. Alice veut être certaine

- qu'une personne non-autorisée (Eve) ne peut pas prendre connaissance des messages qu'elle envoie, confidentialité.

- que ses messages ne seront pas falsifiés par un attaquant malveillant (Martin), intégrité.

- que le destinataire (Bob) a bien pris connaissance de ses messages et ne pourra pas nier l'avoir reçu, non-répudiation.

De plus elle veut être certaine que son message ne sera pas brouillé par les imperfections du canal de transmission (cette exigence ne relève pas du cryptage mais de la correction d'erreur).

Bob veut être certain

- que personne d'autre que lui (et Alice bien sûr) n'a accès au contenu du message, confidentialité.

- que le message reçu vient bien d'Alice authentification, par exemple qu'un attaquant malveillant (Oscar) ne puisse pas se faire passer pour Alice, mascarade ou usurpation d'identité

- que le message n'a pas été falsifié par un attaquant malveillant (Martin), intégrité des données

- que l'expéditeur (Alice) ne pourra pas nier avoir envoyé le message, non-répudiation [3]

1.3 Les types de la cryptographie

Il existe deux grandes familles d'algorithmes cryptographiques à base de clefs, à savoir

1/ La cryptographie à clef secrète ou cryptographie symétrique.

2/ La cryptographie à clef publique ou cryptographie asymétrique

1.3.1 La cryptographie symétrique

Si la clé est unique, elle sert à chiffrer et à déchiffrer le message .le chiffrement à clé secrète a des origines lointaines, et a toujours été associé à des applications militaires. Dans un système à clé secrète ou symétrique un expéditeur et un destinataire partagent une même clé secrète .cette clé est utilisée à la fois pour le chiffrement et pour le déchiffrement et doit rester secrète de tout observateur ennemi [2]. L'émetteur et destinataire doivent se mettre d'accord sur une clef à utiliser avant d'échanger des messages .la sécurité d'un algorithme à clef secrète repose sur la clef : si celle-ci est dévoilée, alors n'importe qui peut chiffrer ou déchiffrer des messages dans ce cryptosystème [4]

Les cryptosystèmes symétriques se répartissent en deux familles :

- le chiffrement à flot : une méthode de chiffrement à flot opère individuellement sur chaque bit de texte clair en utilisant une transformation qui varie en fonction de la place du bit d'entrée comme : xor, RC4, A5 ……

-chiffrement par bloc : Un système de chiffrement par bloc opère avec une transformation fixe qui s'applique sur des blocs de texte clair de taille fixe comme : DES, AES, IDEA …. [2]

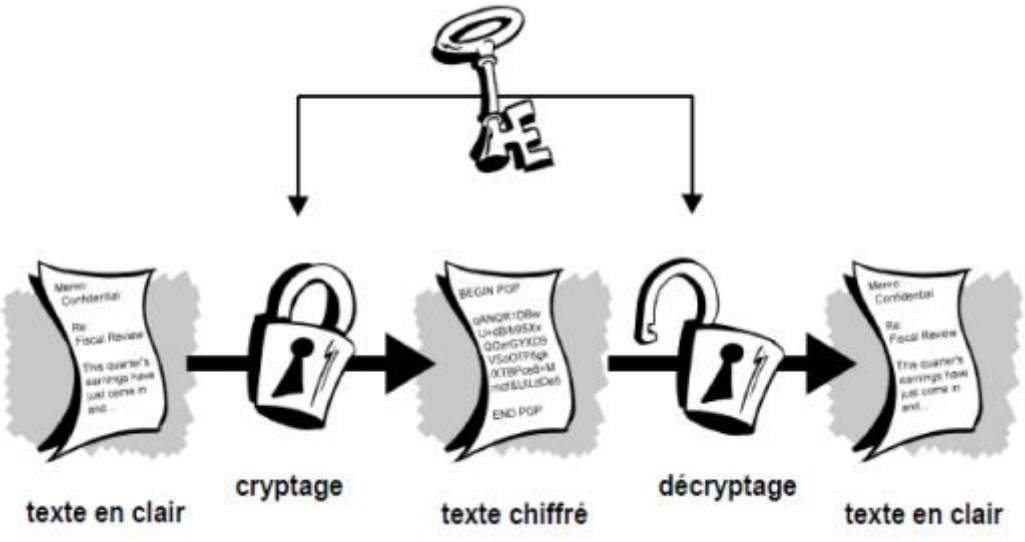


Figure1.6: Le schéma général de la cryptographie symétrique

1.3.1.1 Exemples des méthodes symétriques anciennes

A. Le chiffre de césar

Le chiffre de César est la méthode de cryptographie la plus ancienne communément admise par l'histoire. Il consiste en une substitution mono-alphabétique : chaque lettre est remplacée ("substitution") par une seule autre ("mono-alphabétique"), selon un certain décalage dans l'alphabet ou de façon arbitraire. D'après Suétone, César avait coutume d'utiliser un décalage de 3 lettres : A devient D, B devient E, C devient F, etc. Il écrivait donc son message normalement, puis remplaçait chaque lettre par celle qui lui correspondait [8]

Clair	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Chiffré	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C

Tableau1.1: Chiffrement de César

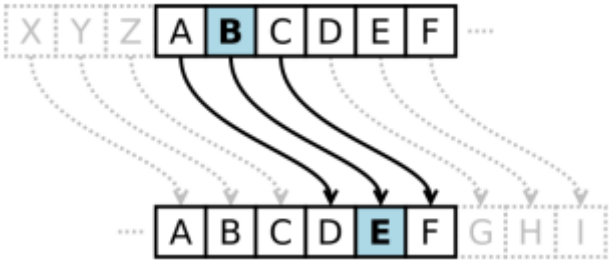


Figure 1 .7 : chiffre de César

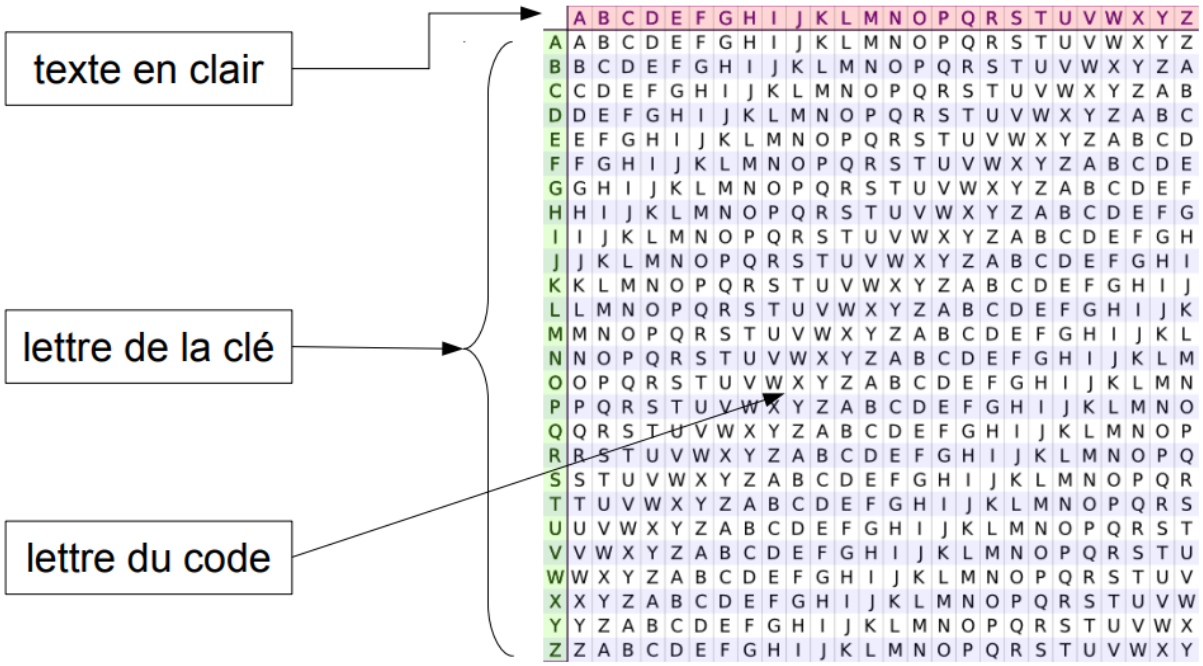
B. Chiffre à substitution simple (monoalphabétique)

C'est un chiffre dans lequel chaque caractère du texte en clair est remplacé par un caractère correspondant dans le texte chiffré. Les cryptogrammes publiés dans les journaux sont des exemples de chiffres à substitution simple [4]

C. Méthode de Vigenère (substitution polyalphabétique)

Un chiffre à substitution polyalphabétique est composé à partir de plusieurs chiffres à substitution simple [4]. On remplace les lettres par d'autres, ce remplacement se fait en fonction de la lettre ainsi que de sa position dans le mot. Le chiffrement de Vigenère ressemble beaucoup au chiffrement de César, à la différence près qu'il utilise une clef plus longue afin de pallier le principal problème du chiffrement de César

Considérons la table de chiffrement de Vigenère suivante (La première ligne correspond aux lettres du texte en clair à crypter et la 1ere colonne correspond à la clé utilisée) [9]



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
B	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A
C	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B
D	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C
E	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D
F	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E
G	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F
H	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G
I	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H
J	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I
K	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J
L	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K
M	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L
N	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M
O	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N
P	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
Q	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
R	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
S	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
T	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
U	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
V	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
W	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
X	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
Y	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
Z	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y

Tableau1.2 : Chiffrement de vigenére

1.3.1.2 Méthodes symétriques modernes

A. DES (Data Encryption Standard)

C'est le standard ANSI X3.92, proposé en 1974, publié par le NIST (National Institute of Standards and Technology) dans le federal Register en 1975, adopté comme standard en 1977 (FIPS-46) [2]

Le système DES est certainement le plus célèbre des systèmes de chiffrement à clé secrète par bloc, il chiffre les données par bloc de 64 bits. Un bloc de 64 bits du texte en clair entre par un coté de l'algorithme et un bloc de 64 bits du texte chiffré sort de l'autre côté.

Le chiffrement et le déchiffrement utilisent tous deux le même algorithme (avec des différences uniquement dans le plan de génération des clefs).

La longueur de la clef est des 56 bits. Généralement, la clef est exprimée comme un nombre de 64 bits mais un bit sur huit utilisé comme bit de contrôle de parité est ignoré. La clef peut être n'importe quel nombre de 56 bits et peut être changée à tout moment.

Au niveau le plus simple, l'algorithme n'est rien d'autre que la combinaison de deux techniques de base de chiffrement : confusion et diffusion. L'élément constitutif de DES est une seule combinaison de ces techniques (une substitution suivie d'une permutation) appliquée au texte, basée sur la clef. On parlera alors de ronde. Le DES a 16 rondes, c'est-à-dire qu'il applique 16 fois la même combinaison de techniques au bloc de texte en clair [4].

Ce système est maintenant considéré comme trop faible en raison de la taille trop petite des clés [2].

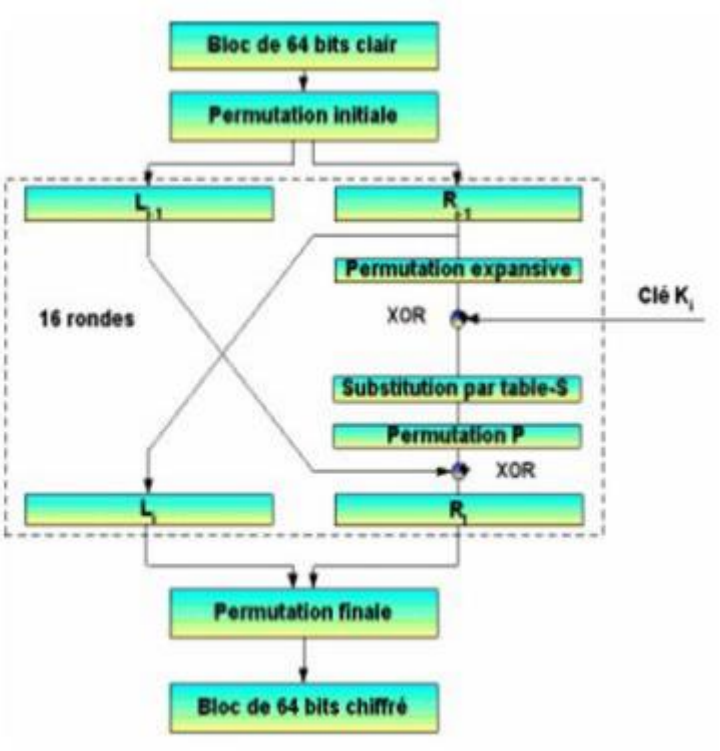


Figure 1 .8 : algorithme principal du DES

B. AES (Advanced Encryption Standard)

C'est un système de chiffrement basé sur le système Rijndael, élaboré par Joan Daemen et Vincent Rijmen en réponse à un appel d'offre du NIST lancé en 1977 pour remplacer le DES.

Tout comme le DES ,l'AES est un standard de chiffrement pour les applications « non classifiées « .pour l'AES les blocs de données en entrée et en sortie sont des blocs de 128 bits c'est -à-dire de 16 octets , 192 ou 256 bits, les clés secrètes ont au choix , suivant la version du système : 128 bits (16 octets) 192 bits (24 octets) ou 256 bits (32 octets),Le choix de la taille de la clef et de la taille des blocs sont indépendants .

Ce chiffrement est constitué de substitutions, de décalages, de «ou exclusif » et De multiplications dans un corps fini de polynômes fixes ; ces opérations sont élémentaires, Simples et rapides à calculer [2].

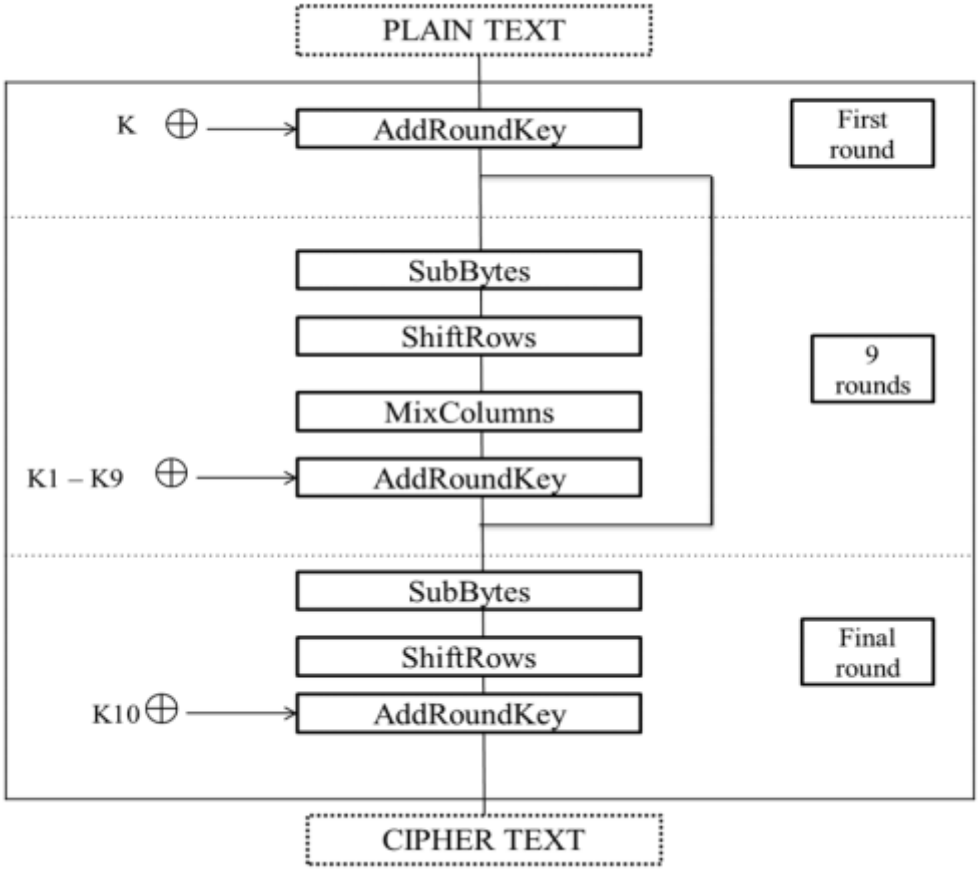


Figure1.9 : Le schéma de fonctionnement de l'AES.

1.3.2 La cryptographie asymétrique

En 1976, Diffie et Hellman introduisent la notion de couple de clés : l'une servant au chiffrement et l'autre au déchiffrement. C'est le début de la première cryptographie à clé publique et du règne du système RSA développé en 1977 par Rivest, Shamir et Adleman. Aujourd'hui il existe divers systèmes à clé publique.

Dans un cryptosystème à clé publique, chaque utilisateur A dispose d'une paire de clés : une clé privée dA et une clé publique eA. La clé privée de A n'est connue que de A .la clé publique est publié et connu de tous .il doit, bien entendu, être impossible en pratique de calculer dA à partir de eA. On dispose en outre d'une fonction publique de chiffrement ξ qui a une clé eA et un texte clair x fait correspondre $y=\xi$ (eA, x), le chiffré de x à destination de A. On dispose également d'une fonction publique de déchiffrement D qui à la clé privé dA de A et à un chiffré y à destination de A fait correspondre x =D (dA, y), le texte clair associé à y. remarquant que seule la clé privée est secrète ; les fonctions ξ et D sont publiques .on notera

EA la fonction de chiffrement à destination de A, c'est à dire la fonction définie par

EA(x)= ξ (eA, x) ; de même en désignera par DA la fonction de déchiffrement de A, c'est à dire la fonction définie par DA(y)=D (dA, y). Pour tout utilisateur A de système on a donc :

DA oEA=identité

Pour résumer la situation disons que si l'expéditeur B veut communique le texte clair m à A, il calcule le texte chiffré c =EA(m) on utilisant la clé publique de A et il envoie c à A. le destinataire A retrouve le texte clair en calculant m=DA(c) grâce à sa clé privé .les systèmes à clé publique reposent sur la difficulté d'effectuer en pratique certains calculs. [2]

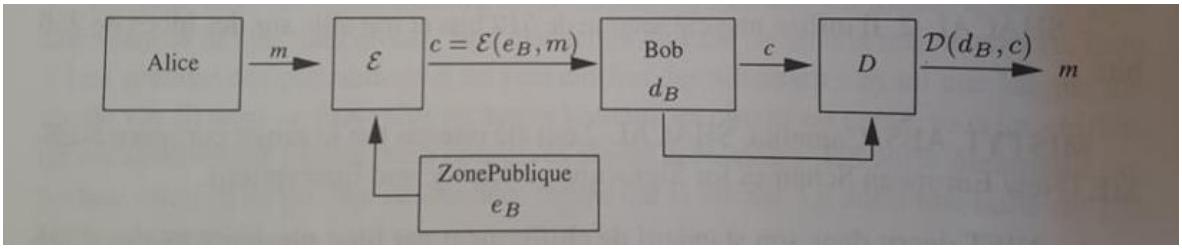


Figure1. 10: Schéma classique d'un système de chiffrement à clé publique

Dans un système à clé publique il n'y a pas de clé secrète à échanger : la clé privée ne sort pas de chez A et la clé publique est connue de tous.

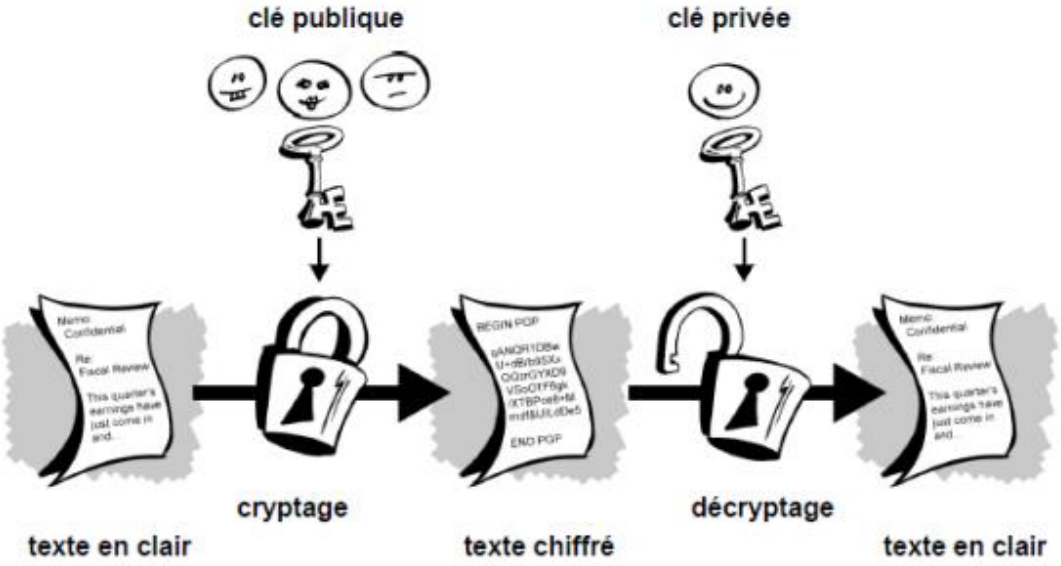


Figure 1 .11 : la cryptographie à clef publique

1.3.2.1 Quelque Exemples de Système de chiffrement à clé publique (asymétrique)

Citons quelques systèmes à clé publique comme

A. Diffie –Hellman

Ce système qui permet d'échanger des clés a été le premier protocole assimilable à un chiffrement utilisant le principe de la cryptographie à clé publique. il utilise la difficulté du problème du logarithme discret sur certains groupes et utilise des paires de clés publiques, privées et éphémères.

B. RSA (Rivest-shamir-Adleman)

Ce système est basé sur la difficulté de factorisé un produit de deux nombre premiers. Associé à l'encodage KEM (Key Encapsulation Mechanism) il a été retenu par le projet NESSIE.

C. ELGamal

Ce système est basé sur la difficulté de probléme du logarithme discret dans certains groupes [2]

1.4 La cryptographie hybride

La cryptographie hybride a été introduite pour profiter des avantages des deux types de cryptographie précédents, en fait cette technique bénéficie de la cryptographie symétrique par sa rapidité de traitement des données et de la cryptographie asymétrique par sa puissance de chiffrement.

Le principe est assez simple, l'échange des clés pour un chiffrement symétrique est effectué grâce à la cryptographie à clé publique, et les données à échanger sont chiffrées en utilisant un algorithme de chiffrement symétrique, cela rend la communication assez rapide [10]

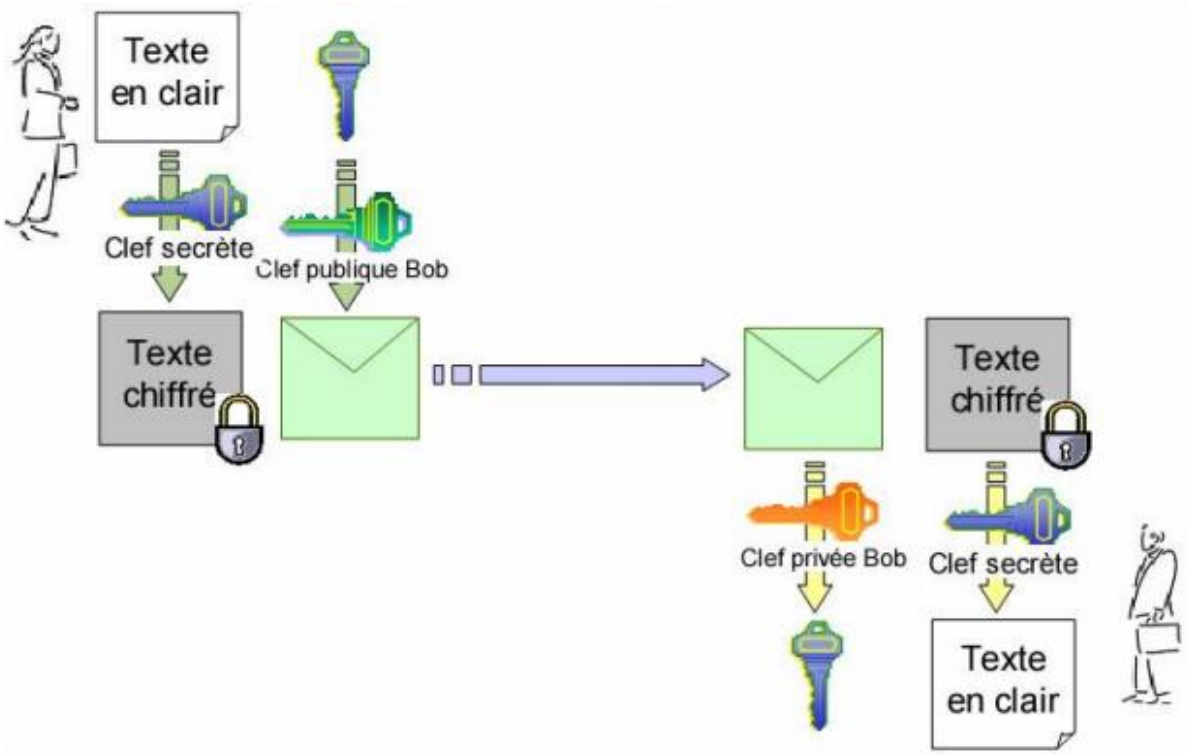


Figure1.12: cryptage hybride

1.5 La signature numérique

L'un des principaux avantages de la cryptographie de clé publique est qu'elle offre une méthode d'utilisation des signatures numériques. Celles-ci permettent au destinataire de vérifier leur authenticité, leur origine, mais également de s'assurer qu'elles sont intactes.

Ainsi, les signatures numériques de clé publique garantissent l'authentification et l'intégrité des données. Elles fournissent également une fonctionnalité de non répudiation, afin d'éviter que l'expéditeur ne prétende qu'il n'a pas envoyé les informations. Ces fonctions jouent un rôle tout aussi important pour la cryptographie que la confidentialité, sinon plus [11]

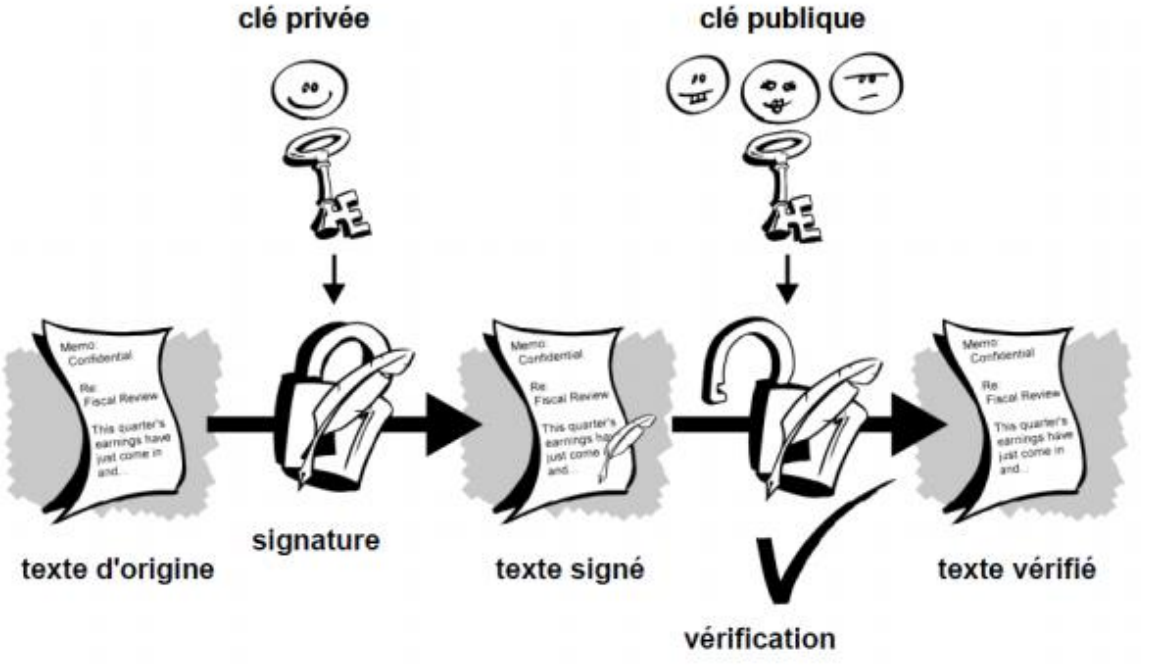


Figure1.13 : Signature numérique

Comme montre la figure ci-dessus, pour signer un message, il suffit de le chiffrer par la clef privée. Et pour vérifier la signature numérique d'un message, il suffit de le déchiffrer à l'aide de la clef publique de l'expéditeur, on ne signe pas des messages directement mais on signe leurs résumés pour gagner le temps.

1.6 Fonctions de hachage

Une fonction de hachage transforme un long message en un bloc court, de taille fixe. L'image d'un message par une fonction de hachage s'appelle, L'empreinte du message, le résumé du message ou encore le message haché. Une fonction de hachage est une application qui vérifie un certain nombre de propriétés :

-Résistance à la détermination d'une préimage, ce qui signifie qu'il doit être impossible en pratique, à partir d'un résumé m, de trouver un message M ayant ce résumé, c'est-à-dire tel que m=h(M), l'impossibilité en pratique dont nous parlons ici, fait référence à la difficulté calculatoire de l'opération en question. Un peu plus précisément, les algorithmes connus qui peuvent mener à bien l'opération impossible s'exécutent en un temps bien trop long pour être en pratique mènes à leur terme.

-Résistance à la détermination d'un deuxième pré image, ce qui signifie que si on se donne un message M1 ainsi que son haché h(M1) il est impossible en pratique de trouver un message M2 distinct de M1 tel que h(M1) =h(M2)

-Résistance aux collisions, ce qui signifie qu'il est impossible en pratique de construire deux messages M1 et M2 ayant le même résumé : h(M1)=h(M2).

Bien entendu, comme une fonction de hachage h transforme un long message en un message court, la fonction h n'est certainement pas injective. Donc les deux propriétés demandées aux fonctions de hachage ne peuvent être réalisable qu'à cause de l'inextricabilité des calculs qui permettraient de revenir aux préimages, ou de provoquer des collisions. Cependant l'inextricabilité des calculs n'est assurée que si la taille fixe des images de la fonction h est suffisamment grande. [2]

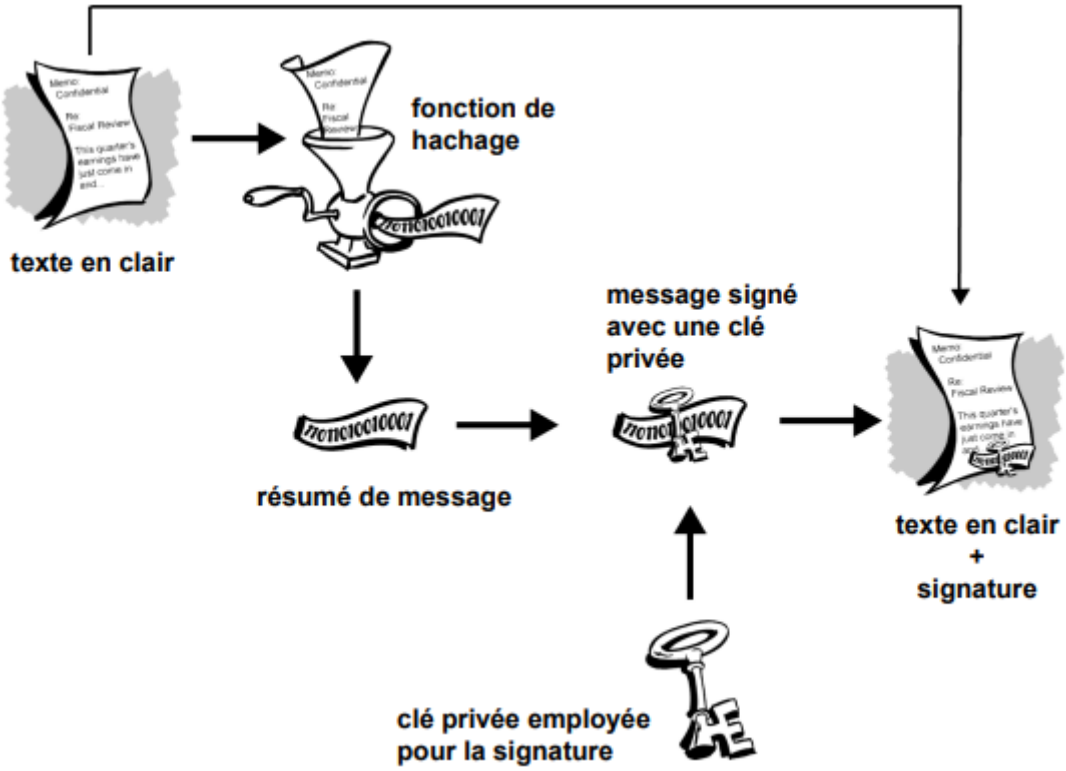


Figure1. 14: Signatures numériques sécurisées

1.7 Certificats numériques

Les certificats numériques ou certificats simplifient la tâche qui consiste à déterminer si une clé publique appartient réellement à son détenteur supposé.

Un certificat correspond à une référence. Il peut s'agir par exemple de votre permis de conduire, de votre carte de sécurité sociale ou de votre certificat de naissance. Chacun de ces éléments contient des informations vous identifiant et déclarant qu'une autre personne a confirmé votre identité. Certains certificats, tels que votre passeport, représentent une confirmation de votre identité suffisamment importante pour ne pas les perdre, de crainte qu'une autre personne ne les utilise pour usurper votre identité.

Un certificat numérique contient des données similaires à celles d'un certificat physique. Il contient des informations associées à la clé publique d'une personne, aidant d'autres personnes à vérifier qu'une clé est authentique ou valide. Les certificats numériques permettent de contrecarrer les tentatives de substitution de la clé d'une personne par une autre. Un certificat numérique se compose de trois éléments :

• Une clé publique.

• Des informations sur le certificat. (Informations sur l'« identité » de l'utilisateur, telles que son nom, son ID utilisateur, etc.)

• Une ou plusieurs signatures numériques [11]

1.8 Les modes de chiffrement

Les cryptosystèmes symétriques se répartissent en deux familles :le chiffrement continu et le chiffrement par bloc.

1.8.1 Le mode de chiffrement en continu

Une méthode de chiffrement continu ou à flot opère individuellement sur chaque bit de texte clair en utilisant une transformation qui varie en fonction de la place du bit d'entrée. le cryptosystème de Vernam appelé aussi one-time-pad ou encore masque jetable est le prototype de ces systèmes .il utilise une clé secrète très longue qui devrait de manière idéale représenter une suite aléatoire de bits.si on a un message m de n bits à chiffrer, on considère les n premiers bits de la clé qui constituent un mot k et on calcule le « ou exclusif bit à bit » entre le message et cette partie de la clé [2]

1.8.2 Le mode de chiffrement par bloc

Un système de chiffrement par bloc opère avec une transformation fixe qui s'applique sur des blocs de texte clair de taille fixe. Ce mode permet de découper le message en bloc et chiffrer chaque bloc.

Plusieurs modes opératoires peuvent être employées : ECB, CBC, CFB, OFB, CTR [12]

Conclusion

Dans notre premier chapitre on a présenté la cryptologie qui se décompose de la cryptographie et la cryptanalyse. Au premier lieu on a défini les notions de base de la cryptographie avec ses principaux objectifs, ensuite nous avons expliqué cette discipline avec ses deux types : la cryptographie symétrique et asymétrique tout en illustrant avec des exemples d'algorithmes cryptographiques anciens et modernes. Enfin on a étudié les deux modes de chiffrement symétrique : en bloc et continu.

Dans le deuxième chapitre on va parler d'une nouvelle technique de chiffrement qui est la cryptographie par ADN.

Chapitre 02

La cryptographie à base d'ADN

Introduction

Le calcul à l'ADN est un nouveau paradigme de calcul, apparu en 1994 après l'expérience de Leonard Adleman .La cryptographie à l'ADN (DNA cryptography) est un nouvel axe de recherche en cryptographie les molécules d'ADN, ayant la capacité de stocker, traiter et transmettre des informations, inspirent l'idée de la cryptographie ADN. Ce sont les techniques non conventionnelles qui émergent rapidement qui combine les caractéristiques chimiques des séquences d'ADN biologique avec la cryptographie classique assurer une transmission non vulnérable des données. Cette méthode innovante est basée sur la notion de Calcul ADN. [13]

En effet, il est possible de bénéficier des avantages des systèmes cryptographiques classiques et de les rendre plus efficaces sur certaines méthodes grâce à l'utilisation de l'ADN. Il y a différentes façons d'utiliser l'ADN pour sécuriser le contenu de l'information Il y'a deux solutions différentes pour utiliser l'ADN dans la cryptographie : sous sa forme biologique ou alors sous forme numérique. D 'une part, l'ADN biologique peut être utilisé pour le stockage et pour cacher des données à l'intérieur de celui-ci. L'information secrète est placée dans une molécule de l'ADN et caché parmi d'autres molécules d'ADN. D'autre part, les nombres aléatoires peuvent être générés à partir de séquences numériques d'ADN. Enfin, la sécurité et la compression sont très importantes lors de la transmission et du stockage des données informatiques. Cependant, la plupart des systèmes de cryptage peuvent augmenter la taille des données, ou encore augmenter la complexité de calcul. [14]

Dans ce chapitre nous allons définir la molécule d'ADN et ses notions de base en présentant les structures physiques et chimiques, ainsi que leurs approches et fonctions.

2.1 La naissance de la biologie moléculaire

Le XXe siècle coïncide avec la naissance génétique : débutant avec la redécouverte des travaux de Mendel précisément en 1900, se poursuivant par l'élaboration de la théorie chromosomique de l'hérédité au début du siècle, la découverte de l'ADN comme support biochimique de l'information génétique, l'élucidation de sa structure, et l'explosion de la biologie moléculaire à partir des années 70. [15]

Une cellule est l'unité fonctionnelle fondamentale des organismes biologiques. La plupart des cellules contiennent un noyau et des chromosomes à l'intérieur de celui-ci. Information génétique (ADN) qui contrôle la fonctionnalité des cellules sont divisée en chromosomes.

Chaque chromosome est composé d'une molécule d'ADN unique qui porte des gènes

(Fig. 2.1).

Chaque cellule tient dans son noyau la même copie des chromosomes, mais en fonction du type de cellule, il active uniquement une spécifique partie de l'ensemble du matériel génétique

(expression des gènes). La cellule a la capacité de stocker, récupérer et traduire des instructions génétiques donnant vie à l'organisme [16]

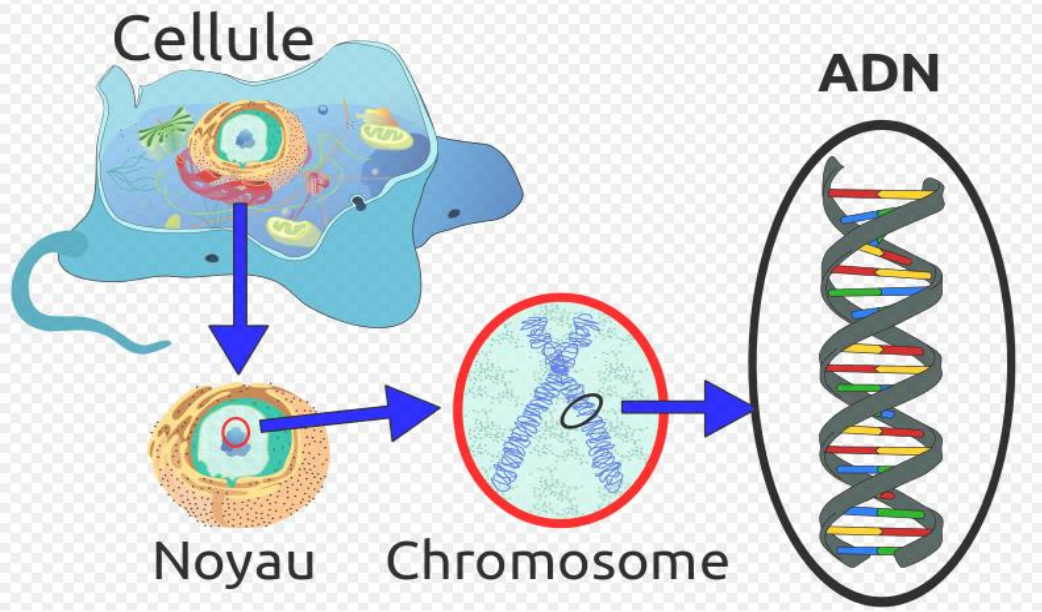


Figure 2.1: l'ADN dans le noyau cellulaire.

2.2 Comprendre L'ADN

2.2.1 Que signifie ADN ?

L'acide désoxyribonucléique ou ADN est un composé organique dont les molécules contiennent les instructions génétiques qui coordonnent le développement et le fonctionnement de tous les êtres vivants et quelques virus, et qui transmettent les caractéristiques héréditaires de chaque être vivant.
ADN est le sigle de acide désoxyribonucléique, un acide nucléique composé de désoxyribose, de phosphate, d'adénine, de cytosine, de guanine et de thymine. L'ADN contient les instructions génétiques utilisées dans le développement et le fonctionnement de tous les organismes vivants et de certains virus, et qui est responsable de sa transmission héréditaire.

Cette macromolécule constitue le support des informations génétiques de tous les êtres vivants exceptés les virus à ARN. Elle est formée d'une double chaîne hélicoïdale de désoxyribonucléotides, chaque chaîne ou brin étant complémentaire de l'autre. [17]

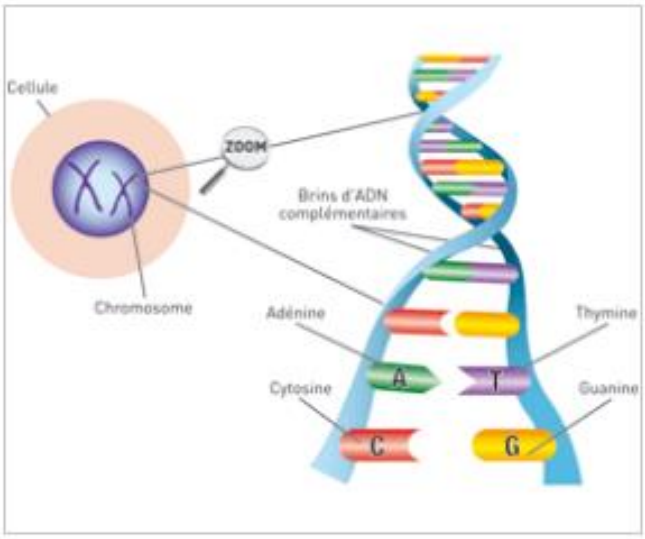


Figure 2.2 : vue globale de l'ADN

2.2.2 Structure de l'ADN

L'ADN est le principal constituant des chromosomes présents dans nos noyaux cellulaires. L'information génétique est contenue dans notre ADN, on peut donc dire que notre ADN est le "support de l'information génétique". Cette information génétique permet la fabrication des protéines qui gèrent la quasi-totalité de nos fonctions biologiques. [18]

L'acide nucléique désoxyribose (ADN) a une forme hélicoïdale, composée de deux longs brins de nucléotides. Un nucléotide a l'une des 4 bases: A - adénine, G - guanine, C - cytosine ou T - thymine, un sucre désoxyribose et un groupe phosphate (Fig2.3)

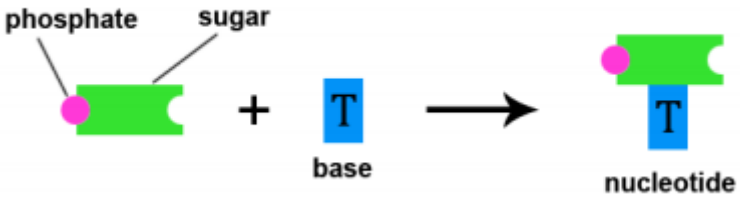
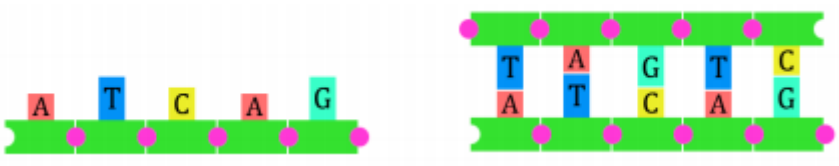


Figure 2.3: structure de nucléotide

Les sucres et les phosphates permettent aux nucléotides de se lier dans un seul brin d'ADN. Les liaisons hydrogène maintiennent ensemble 2 brins et créent un ADN double brin

(Fig. 2.3). Les liaisons hydrogène ne durent qu'entre paires complémentaires: A-T et C-G



Brin d'ADN double brins d'ADN

Figure 2.4: Brins d'ADN formés par des liaisons nucléotidiques et hydrogène

La structure tridimensionnelle de l'ADN a été découverte en 1953 par Watson et Crick. Les brins d'ADN se tordent les uns aux autres pour former une hélice [17], Les bases sont maintenues par une armature sucre-phosphate, les phosphates sont attachés par des liaisons esters aux groupes hydroxyle en 3' et 5'.

Par convention le groupe phosphorique représente le bout 5', le groupe OH le bout 3', c'est également dans ce sens que se lit une molécule d'ADN [16]

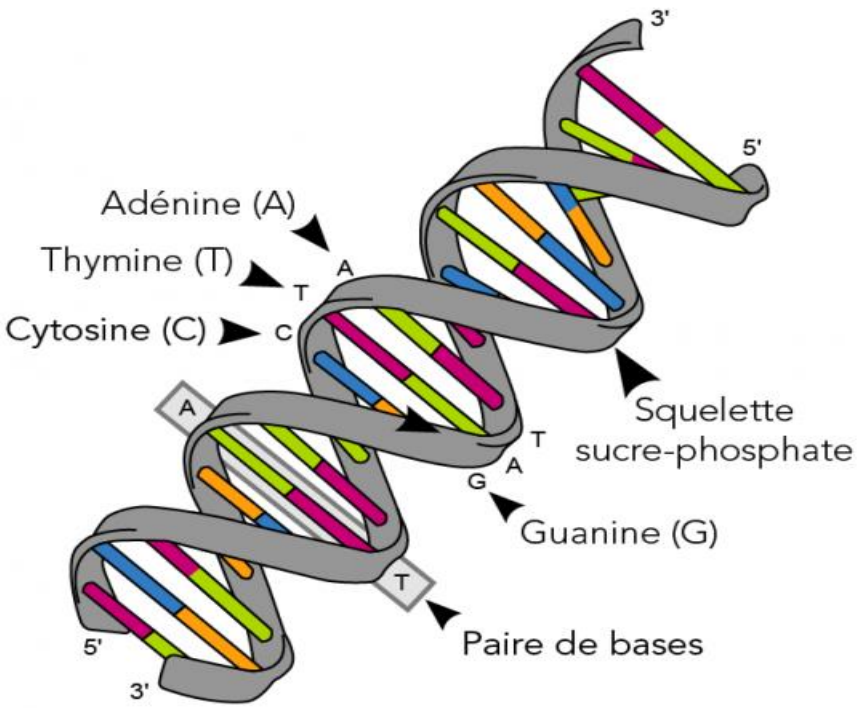


Figure 2 .5 : structure de l'ADN

2.2.3 Les Fonctions de l'ADN

L'ADN à un double rôle : d'une part il contrôle l'information génétique et assure sa permanence au cours de la division cellulaire, d'autre part la biosynthèse protéique.

- Au niveau de l'ADN on distingue différentes régions ayant des fonctions différentes :

• Gène de structure → ARNm → protéine.

• Région codant les ARN ribosomiques et ARNt.

• Région permettant la régulation de la synthèse protéique [19]

Cette molécule assure également un certain nombre de fonctions au sein de la cellule :

-La transcription de l'ADN à l'ARN : L'ADN ne quitte pas le noyau de la cellule. La double hélice s'ouvre afin que le gène soit copié en ARN messager.

- la traduction de l'ARN à la protéine : Les ribosomes traduisent selon le code, l'enchaînement des bases nucléotidiques de l'ARN en une séquence d'acides aminés

- la réplication : Quand les cellules se multiplient, l'ADN est recopié puis enroulé en pelotes : les chromosomes. L'ADN de la cellule mère est reproduit à l'identique pour former celle de la cellule fille [20]

2. 3 Définition de quelques notions

2.3.1 Chromosome

Un chromosome est une structure cellulaire microscopique représentant le support physique des gènes et de l'information génétique, toujours constituée d'ADN, et souvent de protéines. Les chromosomes existent dans les cellules de tous les êtres vivants, en nombre variable, spécifique à chaque espèce.[21]

Fonction

Les chromosomes constituent le matériel héréditaire des cellules. Supports de l'information génétique, ils portent les gènes qui sont transmis de génération en génération. Chaque gène occupe un emplacement précis sur un chromosome donné : c'est son locus. Un même gène sera toujours situé sur un même locus pour tous les individus d'une espèce donnée [21]

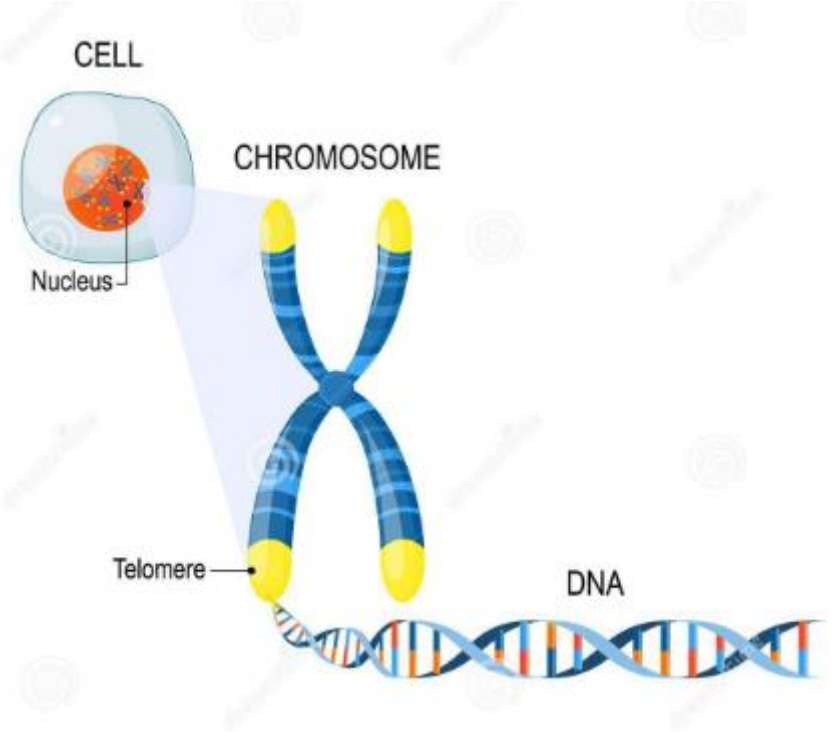


Figure 2.6 : Structure d'un chromosome

2.3.2 Les bases azotées

Une base azotée est aussi appelée base nucléique ou nucléobases. Une base azotée qualifie un corps hétérocyclique azoté, un composé organique possédant des propriétés basiques. Dans les cellules, les bases azotées sont constitutives des acides nucléiques. Les nucléobases sont un constituant des nucléosides et des nucléotides et donc des éléments constitutifs des acides nucléiques, dans l'ARN tel que l'ADN.En effet, il n'existe que deux types complémentaires de bases : une pyrimidine sera toujours en face d'une purine. La thymine (T) et la cytosine (C) sont de la famille des pyrimidines. L'adénine (A) et la guanine (G) sont de la famille des purines.

La structure chimique de l'ADN montre quatre paires de nucléobases (bases azotées, bases nucléiques) produites par huit nucléotides: l'adénine (A) est liée à la thymine (T) et la guanine (G) est liée à la cytosine (C).Si l'on assimile la molécule d'ADN à une échelle, les bases azotées en constituent les barreaux. Les liaisons complémentaires sont réalisées par un nombre différent de liaisons hydrogène: 2 entre adénine et thymine, 3 entre cytosine et guanine, ce qui détermine une différence dans la force de ces liaisons [22]

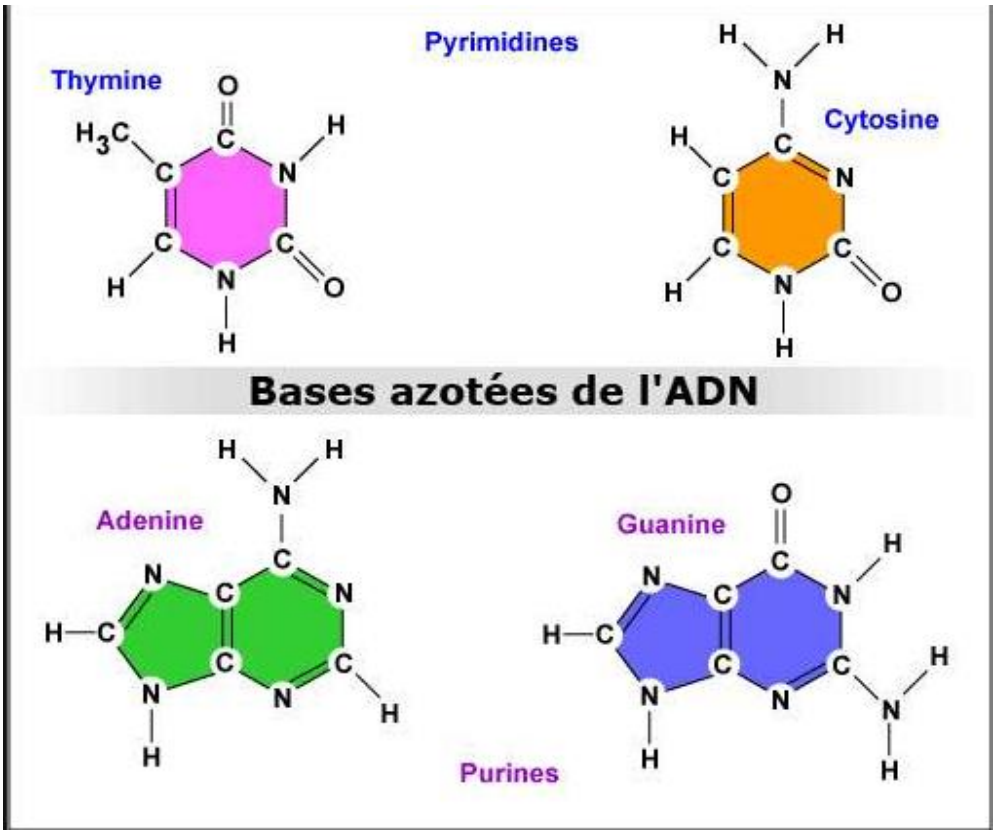


Figure 2.7 : Structure Chimique des bases azotées

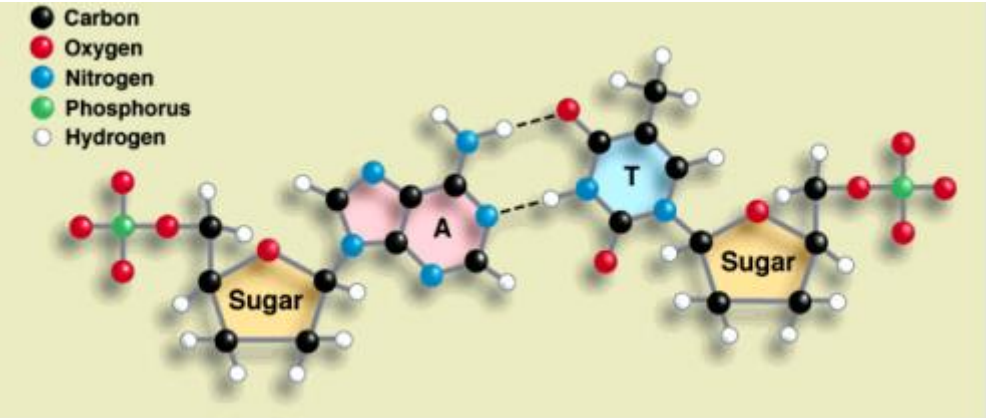


Figure 2.8 : A avec T; deux liaisons hydrogène (liaisons faibles).

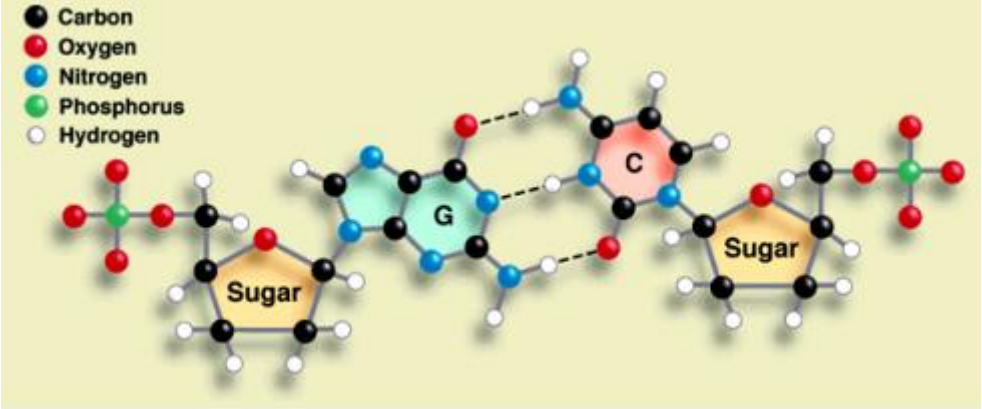


Figure 2.9 : C avec G: trois liaisons hydrogène

2.3.3 Le nucléotide

Molécule composée d'un sucre, d'un phosphate et d'une molécule azotée (base purique ou pyrimidique).Ce sont les unités de base de l'ADN et de l'ARN. [23]

Le sucre présent dans l'ADN est le Désoxyribose, ce sucre est relié à l'une des bases azotée (A, T, C, G). Si on enlève le groupement de phosphate au nucléotide, il devient un nucléoside

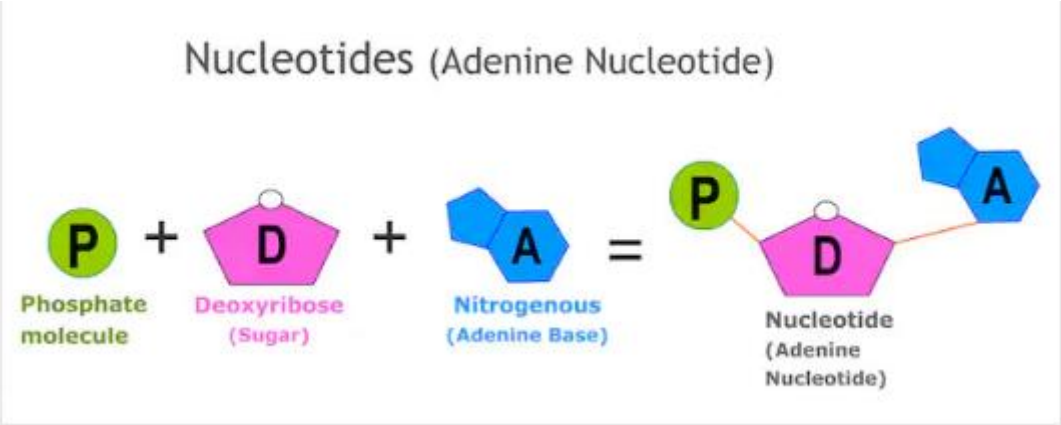


Figure 2.10 : structure de nucléotide

2.3.4 Le nucléoside

Un nucléoside identifie un composé formé d'une base, une base azotée, et d'un pentose: le ribose (ribonucléoside) ou le désoxyribose (désoxyribonucléoside).il ne contient pas de radicauxphosphate. D'un autre terme c'est un nucléotide sans groupe phosphate [24]

{« Sucre » + « Base Azoté »}= Nucléoside

{« Sucre » + « Base Azoté » + « groupement de phosphate »}= Nucléotide

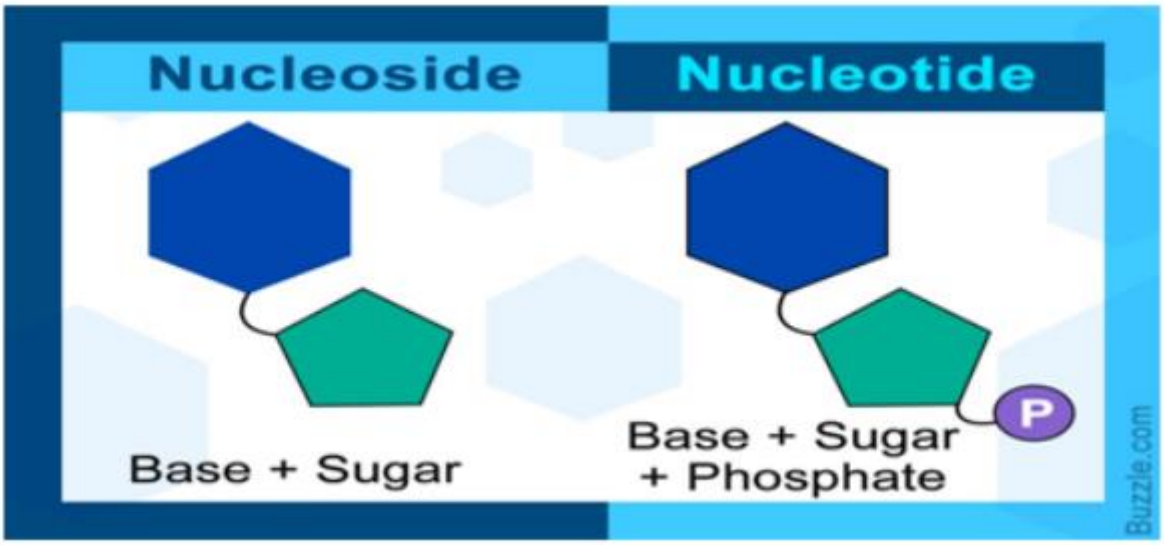


Figure 2.11 : Nucléoside et Nucléotide.

2.4 Structure et fonction de l'ARN

Les gènes portés par l'ADN vont être codés sous une autre forme : en ARN messager, au cours d'un processus nommé « transcription ».

2.4.1 La transcription de l'ADN en ARN

Les molécules d'ADN et d'ARN sont chimiquement très proches, mais le second possède un oxygène supplémentaire sur les sucres (riboses) qui composent ses nucléotides (l'ADN contient en réalité du désoxyribose).

La transcription est un mécanisme biologique permettant la synthèse d'une molécule d'ARN à partir d'une molécule d'ADN complémentaire. C'est la première étape du processus qui permet de passer de l'ADN à la protéine. La transcription est catalysée par une enzyme : l'ARN polymérase

En outre, la thymine (T) de l'ADN est remplacée par l'uracile (U) dans l'ARN [25]

A une séquence d'ADN (constituée des quatre nucléotides désignés par les lettres A, T, C et G) correspond à une unique séquence d'ARN grâce aux règles de complémentarité (G et C, A et U, le U remplaçant le T). Par exemple, la séquence d'ADN AATCGA est transcrite en UUAGCU [26]

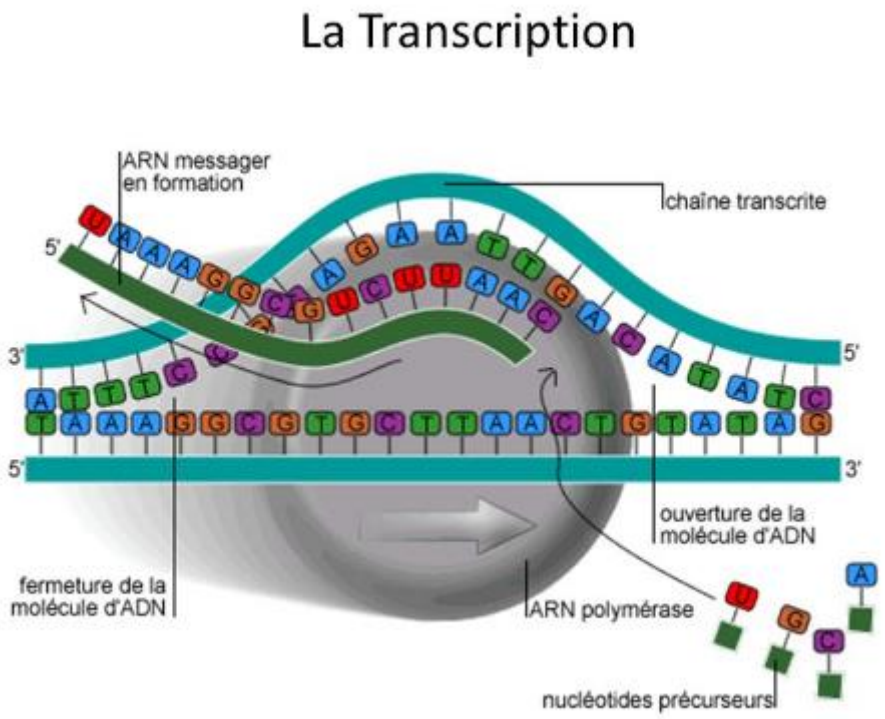


Figure 2.12 : principe de transcription de l'ADN vers l'ARN

Le ribosome, un énorme complexe protéique, réalise la traduction de l'ARN messager en protéine, dernière étape phare de la conversion de l'information génétique en outil fonctionnel

2.4.2 La traduction du gène en protéine

La lecture du message porté par l'ARNm et sa traduction sous forme de protéine sont assurées par le ribosome, une des plus complexes machines cellulaires. Le ribosome est constitué de trois molécules d'ARN et de plus d'une cinquantaine de protéines. Son rôle est le décodage d'une information écrite avec quatre lettres - A, C, G, T de l'ADN, puis A, C, G, U de l'ARN - vers un alphabet à vingt lettres : les vingt acides aminés composant les protéines. Il réalise cette opération en lisant la séquence de l'ARN par groupes de trois bases, chaque triplet ou codon indiquant un acide aminé de la séquence de la protéine en cours de fabrication [25].

Le ribosome parvient sur un des trois **codons « stop »** ou « non-sens », codon auquel ne correspond aucun acide aminé. La dissociation entre l'ARN messager et la chaîne polypeptidique terminée s'effectue alors. [27]

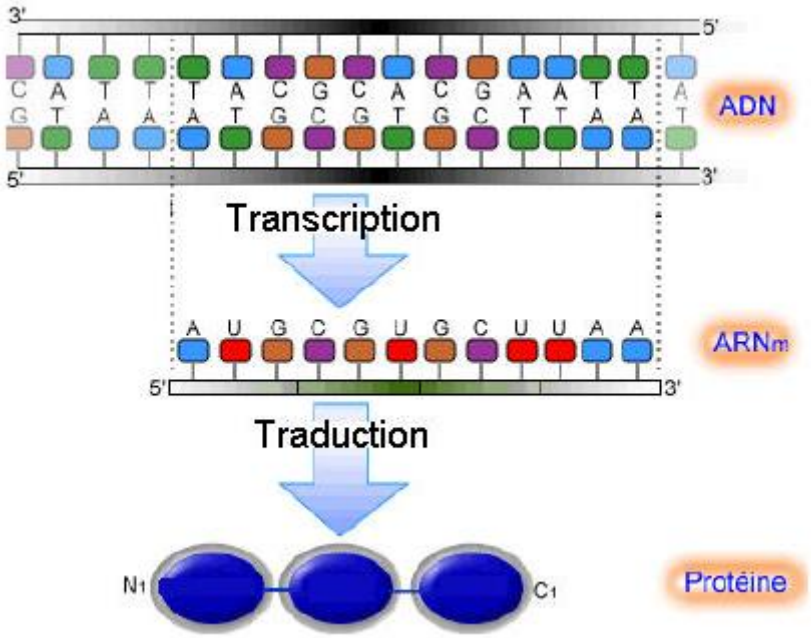


Figure 2.13 : principe de traduction de l'ARNm vers protéine

2.5 Réplication de l'ADN

La réplication se fait par l'intermédiaire d'une enzyme, l'ADN polymérase qui, à partir d'une amorce, va synthétiser un second brin en prenant le premier comme complexe. Les ADN polymérase travaillent de façon unidirectionnelle allant du 5' vers le 3', rendant difficile la synthèse du second brin qui est antiparallèle ; la cellule utilise donc une technique différente qui consiste en la synthèse « à reculons » de petite brins d'ADN, ces petits brins sont ensuite attachés les uns des autres par une ligase [16]

Une réplication semi-conservative veut dire que sur les deux brins toute molécule d'ADN, il y a toujours :

- un brin d'ADN parental (brin ancien)
- un brin d'ADN fils (brin nouvellement formé) [28]

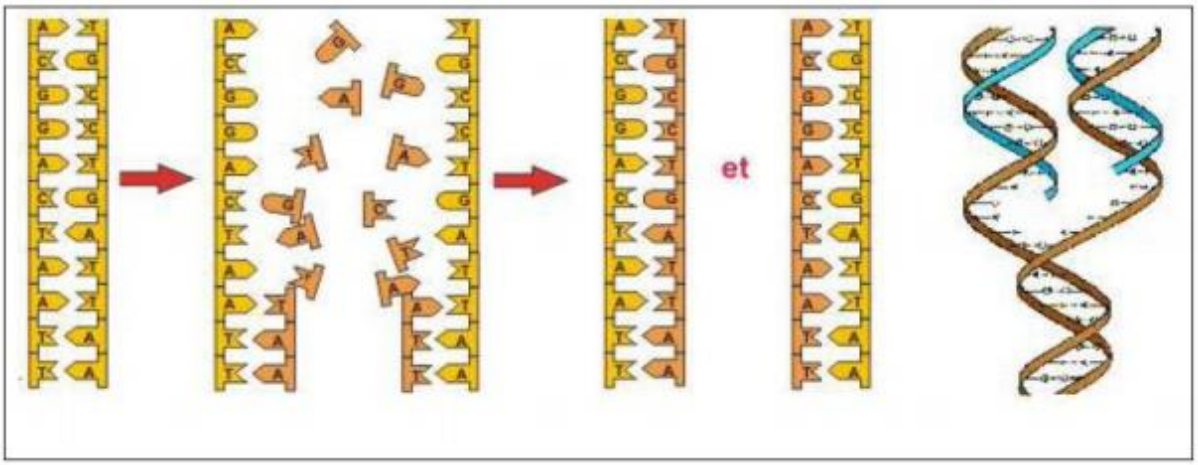


Figure 2 .14 : Réplication semi-conservative de l'ADN

2.6 ADN informatique

Depuis la première démonstration par Leonard Adleman, en 1994, de la résolution d'un problème mathématique grâce à la programmation d'une molécule d'ADN, un nombre croissant de scientifiques se penchent sur l'ADN pour développer de nouveaux systèmes informatiques moléculaires capables de performances inouïes.

La cellule vivante est la technologie la plus complexe que nous connaissions actuellement. » John Reif n'est pas biologiste moléculaire, mais professeur de sciences informatiques à Duke University. Il est aussi directeur du Consortium d'informatique biomoléculaire. Inspiré par les travaux de Leonard Adleman, qui a résolu en 1994 un problème mathématique en programmant des brins d'ADN la molécule qui est au cœur de la cellule et possède le « code » sans lequel celle-ci ne pourrait ni exister ni fonctionner, John Reif fait partie des chercheurs qui ont entrepris de défricher le terrain vierge de l'informatique ADN.

Une opération informatique exige deux conditions fondamentales : la possibilité de stocker des données et la possibilité de les manipuler. L'ADN remplit ces deux conditions : l'enchaînement des nucléotides A, T, C, G peut être programmé pour stocker l'information (de la même manière qu'un ordinateur classique stocke l'information sous forme de 0 et de 1), et les réactions biochimiques décrites plus haut peuvent être utilisées pour manipuler et exploiter cette information. [29]

2.7 ADN Cryptographie

La cryptographie à ADN est une nouvelle large branche scientifique, qui comprend une variété de domaines scientifiques: sécurité de l'information (cryptographie, Stéganographie, gestion des clés), biologie moléculaire, bioinformatique, calcul biomoléculaire. C'est un nouveau et prometteur domaine de la sécurité de l'information. Il combine les solutions classiques de cryptographie avec le force du matériel génétique. L'ADN biologique peut être utilisé en Stéganographie et la cryptographie comme matériel de stockage. Les calculs moléculaires peuvent être effectués avec 25 structures d'ADN biologique et ensuite appliqué sur les chiffres classiques. Plusieurs projets en la séquençage du génome offre la possibilité d'exploiter des bases de données numériques d'ADN fins cryptographiques. [30]

2.7.1 Substitution d'ADN et one time pad

Gehani et al ont d'abord proposé la possibilité de concevoir cryptosystème utilisant des molécules d'ADN en 1999. Ils ont développé un chiffrement ponctuel méthodologie utilisant deux techniques différentes; le premier est par la méthode de substitution d'ADN en utilisant bibliothèques de tampons distincts générés aléatoirement représentés par des brins d'ADN et le second est en utilisant le schéma XOR bit par bit en utilisant le calcul moléculaire [30]

2.7.1.1 Schéma d'one time pad par substitution d'ADN

Le cryptosystème conçu par Gehani et Al. Utilisant une méthodologie de bloc unique contiennent les éléments suivants:

1. Message binaire en texte clair qui est représenté par des brins d'ADN de longueur n et crachés en mots de longueur fixe.

2. Bibliothèque de codes contenant un grand nombre de longues plaquettes d'ADN; chacun représentant unique et mappage aléatoire du mot en texte clair au mot de chiffrement. Il sert

de clé au schéma proposé. Le one time pad codebook se compose de l'unité répétitive (i) qui a trois domaines;

• Le premier domaine représente l'ensemble des mots

• Le domaine suivant représente les mots en clair correspondants

• Le dernier domaine est la séquence d'arrêts qui sert de ponctuation entre les unités répétitives. [31]

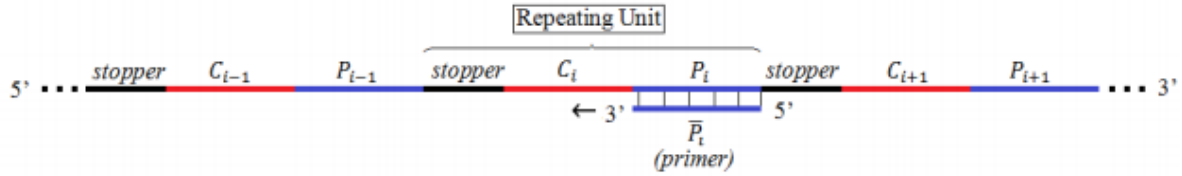


Figure 2 .15 : unité répétitive de l'ADN codebook

2.7.1.2 Schéma de one time pad par Bit -wise XOR

Une autre technique d'one time pad proposée par Gehani et al est basée sur le principe fondamental du chiffrement de Vernam [30]

On utilise le tableau du xor suivant :

P_i	K_i	C_i
0	0	0
0	1	1
1	0	1
1	1	0

Tableau 2.1: Table de vérité XOR

2.7.1.3 Stéganographie ADN

Gehani et al proposent également des méthodes de Stéganographie utilisant la séquence d'ADN dans leur papier.

Étape 1: Le texte en clair est représenté sous la forme de brins d'ADN d'entrée.

Étape 2: Les brins d'entrée sont étiquetés avec une clé secrète qui se présente également sous la forme d'ADN séquence.

Étape 3: Ces brins sont mélangés avec des brins d'ADN aléatoires qui sont spécifiés comme distracteur.

Étape 4: Si la clé secrète est connue du récepteur, les brins représentant le texte en clair peuvent être extraits de la solution mélangée suivant le protocole de purification par affinité.

La séquence échouée utilisée dans l'expérience est la séquence complémentaire de la clé secrète. [30]

2.7.2 Protection de l'information chez l'hôte vivant

La méthode proposée pour Stéganographie par Clelland et al où les auteurs ont recommandé que pour les séquences d'ADN de stockage d'informations peuvent être aussi fiables qu'un morceau de papier. Où les auteurs ont chiffré informations dans des brins d'ADN synthétiques et stockées en permanence les informations dans l'hôte vivant permettant à l'organisme de croître et de se multiplier en toute sécurité. Ils ont assuré la protection des séquences cryptées d'ADN provenant de circonstances défavorables, telles que la rupture fatale d'ADN double brin causée par une température extrême et la dessiccation ou la réhydratation; présence d'ADN nucléases; rayon ultraviolet, rayonnement ionisant; attaque intentionnelle de toute personne, etc. [30]

2.7.3 Cryptographie d'ADN utilisant des brins binaires

Leier et al. Utilisés des brins binaires d'ADN pour effectuer la cryptographie dans leur article. Ils ont montré que si l'adversaire a les mêmes potentiels techniques que l'expéditeur et le destinataire du message secret; même alors, le cryptosystème proposé fonctionne efficacement. Ils aussi projeté une autre technique de cryptographie basée sur la soustraction graphique d'images de gel binaire. [30]

2.7.4 Cryptosystème à ADN à clé symétrique

MingXin et al. Proposé un cryptosystème ADN basé sur la technique de la clé symétrique dans leur papier. Ils ont incorporé la technologie des puces à ADN dans l'ADN

La cryptographie pour concevoir un cryptosystème sécurisé qui n'est pas affecté même par un système hautement efficace ordinateur quantique. Les chercheurs ont exploité le parallélisme massif du calcul de l'ADN et étonnante capacité de stockage d'informations des molécules d'ADN pour développer la clé symétrique d'ADN Cryptosystème (DNASC). Le chiffrement à clé symétrique utilise la même clé pour chiffrer et déchiffrer la secrète information. [30]

2.7.5 Cryptosystème à ADN à clé asymétrique et méthode de signature

Lai et al. Proposé un cryptosystème à ADN à clé asymétrique et une méthode de signature

Papier. Les auteurs ont conçu DNA-PKC qui peut effectuer chiffrement de clé asymétrique et méthode de signature numérique imitant la technologie à l'aide d'un micro réseau d'ADN.

Dans cette méthodologie, un ensemble de sondes ADN représente la clé publique (Ek) et l'autre ensemble des sondes ADN représente la clé privée (Dk). Avec la norme sélective des sondes la condition d'hybridation peut également être incluse comme clé de déchiffrement. On Utilisant la clé publique tout le monde peut crypter un message secret sur une puce à ADN et le transmettre physiquement au récepteur.

Seul le destinataire prévu qui possède la clé privée peut déchiffrer le texte chiffré. L'ensemble des clés privées est conçue de manière à avoir un certain lien avec la clé publique. Dans ce modèle les sondes de clé privée sont complémentaires de la clé publique à des fins d'hybridation. Sur la Puce d'ADN si les sondes s'hybrident selon une norme prédéterminée, cela produira signal à intensité variable. Si l'intensité du signal est supérieure à une valeur seuil prédéfinie, il représente le chiffre binaire "1" et les sondes dont l'intensité du signal est inférieure à la valeur seuil représente le chiffre binaire "0". Pour le déchiffrement, le signal d'hybridation est traduit à nouveau texte original.

Dans la méthodologie de signature, le détenteur de la clé de signature privée peut générer une signature et seuls les détenteurs de clés de vérification publiques peuvent authentifier la signature. [30]

2.7.6 Cryptographie d'ADN à trois étapes

Soni et al ont proposé un nouvel algorithme de cryptographie ADN basé sur le concept de machine Moore dans la théorie des automates. Les auteurs affirment que la conception du cryptosystème est plus fiable et la sécurité de cette technique repose sur les trois cryptages

Les étapes qui utilisent une clé secrète, une machine Moore et un mot de passe générés automatiquement.

La théorie des automates peut être définie comme l'étude de dispositifs automoteurs abstraits qui suivre une séquence d'opérations prédéterminée pour résoudre automatiquement le problème de calcul.

La machine à état final est un type d'automate qui peut être défini comme une machine à nombre d'états et l'automate peut être dans un seul état à tout moment. Machine Moore est un FSM qui produit une sortie en fonction uniquement de l'état actuel. [30]

Conclusion

Dans ce chapitre, nous avons commencé d'abord par la naissance de la biologie moléculaire ainsi que la découverte de l'ADN comme support de l'information génétique. Ensuite on a expliqué la structure de la molécule d'ADN avec ses différents composants et fonctions, comme on a défini quelques notions liées à cette molécule. Cependant la structure et fonction de l'ARN qui est représentée par la transcription de l'ADN en ARN, la traduction en protéine et la réplication de l'ADN. La naissance de la cryptographie à base d'ADN est donnée par le développement remarquable du calcul à l'ADN et ceci en exploitant le parallélisme ainsi la capacité importante du stockage qu'offre cette molécule, enfin nous avons présenté quelques approches qui sont faites par plusieurs chercheurs pour résoudre des problèmes complexes.

Dans le chapitre suivant, nous allons présenter un nouvel axe de recherche de l'IOT

Chapitre 03

Internet des objets (IoT)

Introduction

Depuis la fin des années 1980, Internet a évolué de manière spectaculaire. La dernière étape est l'utilisation de ce réseau mondial pour la communication avec des objets ou entre objets, évolution nommée Internet des Objets (IoT pour *Internet of Things*). L'évolution de l'IoT est rapide : depuis 2014, le nombre d'objets connectés est supérieur au nombre d'humains connectés et il est prévu que 50 milliards d'objets seront connectés en 2020. [32]

Tous les appareils présents à différents endroits autour de nous, comme les maisons, les bâtiments, les villes et même dans notre point de vue des données, les organismes peuvent détecter ou générer des données pour diverses applications de notre vie quotidienne telles que soins de santé, surveillance de l'environnement, militaire et industrie. Lorsque ces appareils communiquent et partagent les informations entre eux sur un espace distribué via internet, ils constituent l'internet des objets (IoT). Par conséquent, un appareil IoT a la capacité de communiquer, de télécharger des informations via Internet sans intervention humaine. En d'autres termes, les appareils sont capables de penser et prendre une décision. Parallèlement au développement rapide de l'application IoT, la sécurité de l'IoT est une question cruciale qui comprend des menaces visant à exploiter d'éventuelles faiblesses. Dans l'IoT, la sécurité est divisée en deux parties

Tout d'abord, un mécanisme d'authentification et d'autorisation est nécessaire pour assurer la sécurité de la communication réseau qui protège le réseau de tout périphérique intrus, qui peut envoyer ou recevoir des informations dans le réseau. Deuxièmement, les informations elles-mêmes devraient également être sécurisées au moyen de techniques de cryptage. Bientôt la base de différents algorithmes de cryptographie, la sécurisation du dispositif de données est possible. La cryptographie est principalement utilisée pour sécuriser les informations en partageant la clé secrète sur différents appareils. Deux types de clés sont disponibles symétriques et clé asymétrique.

En symétrique, les touches sont utilisées des deux côtés émetteur et récepteur tandis qu'en asymétrique deux différentes les clés sont utilisées. L'IoT traite des données en temps réel telles que le point critique, la taille des données est également une mesure importante.

Pour certaines applications telles que la surveillance de l'environnement, le temps d'échantillonnage n'est pas très critique car les données peuvent être collectées toutes les minutes ou heures lors de la surveillance du trafic ou des soins de santé. Lors du téléchargement ou du téléchargement petite quantité de données, il ne nécessitera pas une bande passante Internet très élevée et vice versa. La cryptographie peut modifier les données de type ou de taille en fonction de l'algorithme utilisé de sorte que l'intrus ne puisse pas identifier les données d'origine. Par conséquent, l'algorithme utilisé pour le chiffrement des données dans l'IoT doit être choisi avec soin qu'il ne surchargerait pas la bande passante ou n'affecterait pas l'application en temps réel, ce qui pourrait conduire à une mauvaise performance de l'appareil. La sécurité typique du système IoT peut être classée selon le terme

Suivant: accès contrôle, authentification, protection de la vie privée, sécurité des communications, intégrité et confidentialité des données, et disponibilité. [32]

3.1 Qu'est-ce que l'IoT ?

L'IoT est l'acronyme d'Internet Of Things (Internet des Objets en français). Le terme IoT est apparu la première fois en 1999 dans un discours de Kevin ASHTON, un ingénieur britannique. Il servait à désigner un système où les objets physiques sont connectés à Internet. Il s'agit également de systèmes capables de créer et transmettre des données afin de créer de la valeur pour ses utilisateurs à travers divers services (agrégation, analytique…).

Selon l'UIT (Union Internationale des Télécommunications), l'Internet des Objets est défini comme « une infrastructure mondiale pour la société de l'information, qui permet de disposer de services évolués en interconnectant des objets (physique ou virtuels) grâce aux technologies de l'information et de la communication interopérables existantes ou en évolution ».

Au fil du temps, le terme a évolué et il englobe maintenant tout l'écosystème des objets connectés. Cet écosystème englobe, des fabricants de capteurs, des éditeurs de logiciels, des opérateurs historiques ou nouveaux sur le marché, des intégrateurs… Cet éclectisme en fait sa richesse.

3.1.1 Les 5 composantes de l'IoT

Une solution d'IoT s'articule autour de 5 composants essentiels que sont :

- Les objets (capteurs)
- Le réseau (connectivité)
- Les données
- Les informations
- Les applications d'exploitation [33]

Les composants matériels et leur environnement de développement sont illustrés. Ils comprennent :

- Les identificateurs, capteurs, afficheurs, actuateurs au niveau des objets ;

- Les microcontrôleurs ou processeurs et cartes bas coût sans ou avec OS léger pour les objets simples, des processeurs avec OS comme Linux, Android, IOS et les ressources du nuage pour les objets complexes.

- Les réseaux, qui peuvent être « courte distance », « longue distance » ou « cellulaires ». Les différentes technologies de réseau utilisables sont détaillées en fonction des contraintes : portée, débit, coût, sécurité, etc.

La gestion d'une application comprenant des centaines, des milliers voire plus, d'objets implique de disposer d'un support de développement performant (plateforme). Des plateformes « solutions propriétaires » et « libres » sont disponibles. Elles doivent satisfaire à un certain nombre de besoins :

- gérer les composants et le support d'intégration ;
- assurer la sécurité de l'information ;
- définir le protocole de recueil des données ;
- permettre l'analyse des données pour obtenir l'information pertinente (Big Data) [31]

3.1.2 Les technologies qui ont rendu l'IoT possible

Bien que l'idée de l'Internet of Things soit loin d'être nouvelle, c'est un ensemble de progrès récents de différentes technologies qui a permis de le concrétiser.

- **Accès à une technologie de capteurs à coût réduit et faible consommation**

Des capteurs fiables et abordables rendent possible la technologie IoT pour un plus grand nombre d'industriels.

- **Connectivité**

La prolifération des protocoles réseau pour Internet a facilité la connexion des capteurs au cloud et à d'autres "objets" pour un grain d'efficacité des transferts de données.

- **Plates-formes de cloud computing**

La disponibilité accrue des plates-formes de cloud permet aux entreprises et aux consommateurs d'accéder à l'infrastructure dont ils ont besoin pour évoluer, sans pour autant avoir à s'occuper de sa gestion.

- **Machine Learning et analyses**

Grâce aux progrès effectués dans les domaines du Machine Learning et des analyses, et avec l'accès à de vastes quantités de données diversifiées stockées dans le cloud, les entreprises obtiennent des informations plus rapidement et plus facilement. L'émergence de ces

technologies associées continue à repousser les limites de l'IoT, et les données produites par l'IoT viennent à leur tour renforcer ces technologies.

- **Intelligence artificielle (IA) conversationnelle**

Les progrès effectués en matière de réseaux neuronaux ont permis aux appareils IoT de gérer le traitement des langues naturelles (avec notamment les assistants digitaux personnels tels qu'Alexa, Cortana et Siri), et les ont rendus attrayants, abordables et viables pour une utilisation domestique. [34]

3.1.3 Domaines d'applications de l'IoT

Plusieurs domaines d'application sont touchés par l'IoT, Parmi ces principaux domaines nous citons: le domaine de la sécurité, le domaine du transport, l'environnement et l'infrastructure et les services publics….etc.

Quelques exemples courants sont présentés ci-dessous :

A- Les transports

Depuis la création de l'IoT en 1999, le nombre des véhicules intelligents sont en croissance, presque Tous les véhicules vendus aujourd'hui dans le monde renferment déjà des capteurs et de moyens de communication pour traiter la congestion du trafic, la sécurité, la pollution et le transport efficace des marchandises, etc.

L'objectif est qu'une voiture soit capable de communiquer de façon autonome avec d'autres véhicules ou une centrale de surveillance pour prévenir les accidents et réduire les coûts d'assurance.

B- La santé

Le secteur de la santé a connu un très grand nombre d'applications permettant à un patient et à son docteur de recevoir des informations, parfois même en temps réels, qu'il aurait été impossible de connaître avant l'apparition d'IoT.

Par exemple, (Porteuse Digital Health) qui est le premier médicament connecté sur le marché grâce à un capteur directement intégré dans l'être humain qui permet après ça le suivi des patients à distance.

C- La domotique

La domotique regroupe l'ensemble des technologies permettant l'automatisation des équipements d'un habitat. Elle vise à apporter des fonctions de confort : commandes à distance, gestion d'énergie (optimisation de l'éclairage et du chauffage… etc.), sécurité

(Comme les alarmes) et de communication (contacts et discussion avec des personnes extérieures)

D- Agriculture

L'agriculture intelligente a pour objet de renforcer la capacité des systèmes agricoles, de contribuer à la sécurité alimentaire en intégrant le besoin d'adaptation et le potentiel d'atténuation dans les stratégies de développement de l'agriculture durable.

Cet objectif a été atteint enfin par l'utilisation des nouvelles technologies, telles que l'imagerie satellitaire et l'informatique, les systèmes de positionnement par satellite comme GPS, aussi par l'utilisation des capteurs qui vont s'occuper de récolter les informations utiles sur l'état du sol, taux d'humidité, taux des sels minéraux, etc. et envoyer ces informations au fermier pour prendre les mesures nécessaires garantissant la bonne production. [35]

Figure 3.1 : Domaines d'applications de l'IoT

3.1.4 Architecture d'Internet des Objets

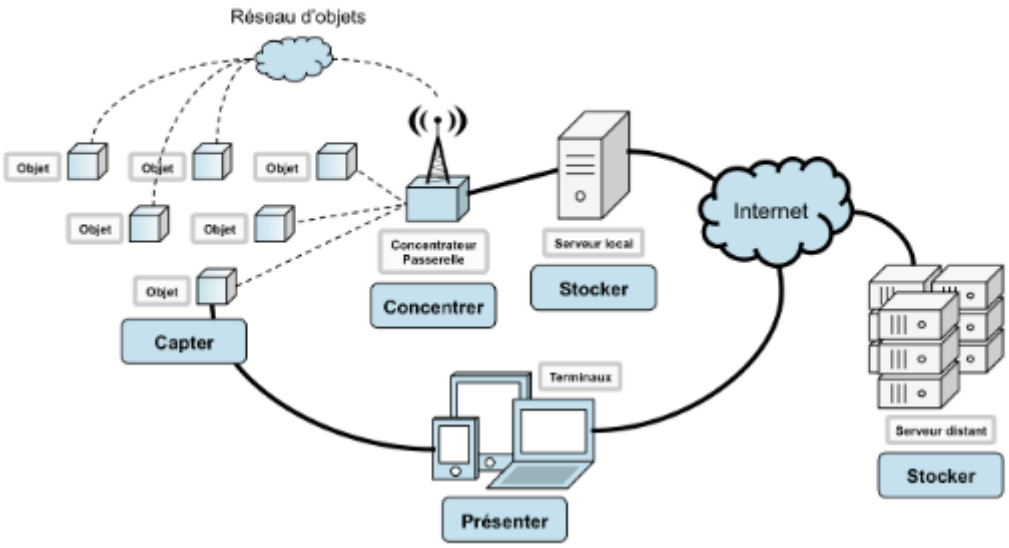


Figure 3.2 : Architecture d'un environnement IoT

Précisons le rôle des différents processus présentés sur ce schéma :

- **Capter** désigne l'action de transformer une grandeur physique analogique en un signal numérique.
- **Concentrer** permet d'interfacer un réseau spécialisé d'objet à un réseau IP standard (e.g. Wifi) ou des dispositifs grand public.
- **Stocker** qualifie le fait d'agréger des données brutes, produites en temps réel, méta taguées, arrivant de façon non prédictible.
- Enfin, **présenter** indique la capacité de restituer les informations de façon compréhensible par l'Homme, tout en lui offrant un moyen d'agir et/ou d'interagir [36]

Conclusion

Dans ce chapitre nous avons parlé d'abord de l'évolution rapide de l'IoT qui est l'interconnexion entre l'internet et les objets, des lieux et des environnements physiques, ensuite on a cité ses différents composants. Après avoir donné un aperçu sur les technologies qui ont rendu l'IoT possible on a présenté quelques domaines d'applications de l'IoT. Enfin on a déterminé un exemple d'architecture d'Internet des Objets.

Toute l'analyse faite dans les chapitres précédents reste qu'une théorie, il est nécessaire pour nous de la confronter à la pratique, c'est ce que va développer notre cadre expérimentale.

Chapitre 04

Implémentations et résultats

Introduction

Nous avons proposé une méthode de cryptage symétrique qui a été implémenté en python sur un Raspberry.

Dans ce chapitre, nous détaillons l'algorithme proposé et nous présentons les résultats obtenus.

4.1 Présentation de l'algorithme

La tâche de notre système est de chiffrer / déchiffrer un texte en le découpant en blocs de 16 caractères, la clé utilisée pour cette tâche est une clé symétrique extraite à partir d'un chromosome humain.

Notre approche est illustrée sur la figure **4.2**.

4.2 Le chiffrement

4.2.1 Extraction des blocs et codage

Dans cette phase, le texte original passe par l'étape de découpage en lettres On calcule la fréquence de chaque lettre et on calcule aussi l'indicateur. Pour chaque lettre en texte brut, calculer le caractère chiffré à l'aide de l'équation eq (1). Le texte résultat, sera ensuite codé en blocs de 16 caractères.

Caractère codé = Fréquence XOR Limit Indicator XOR Code ASCII …….. 1

Chaque lettre du texte résultat est converti en code ASCII puis en binaire.

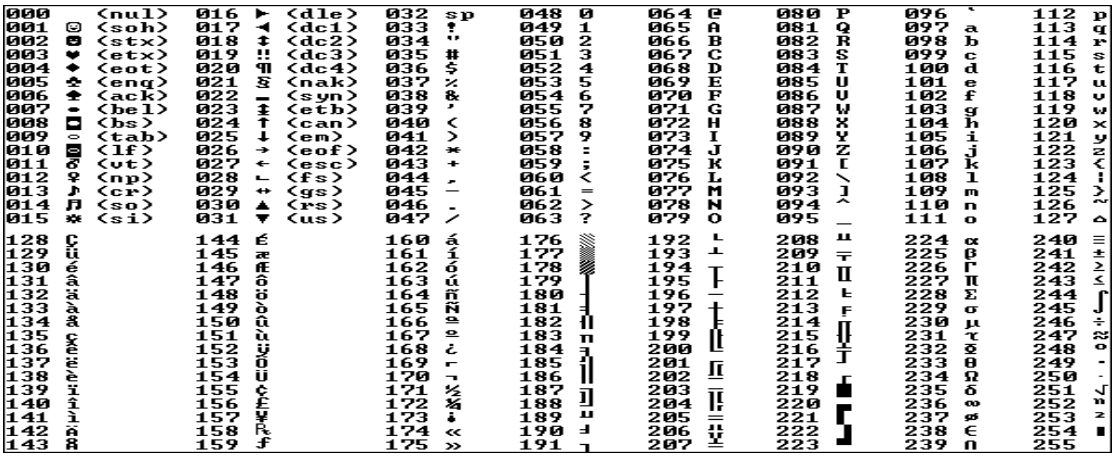


Figure 4.1 : code ASCII

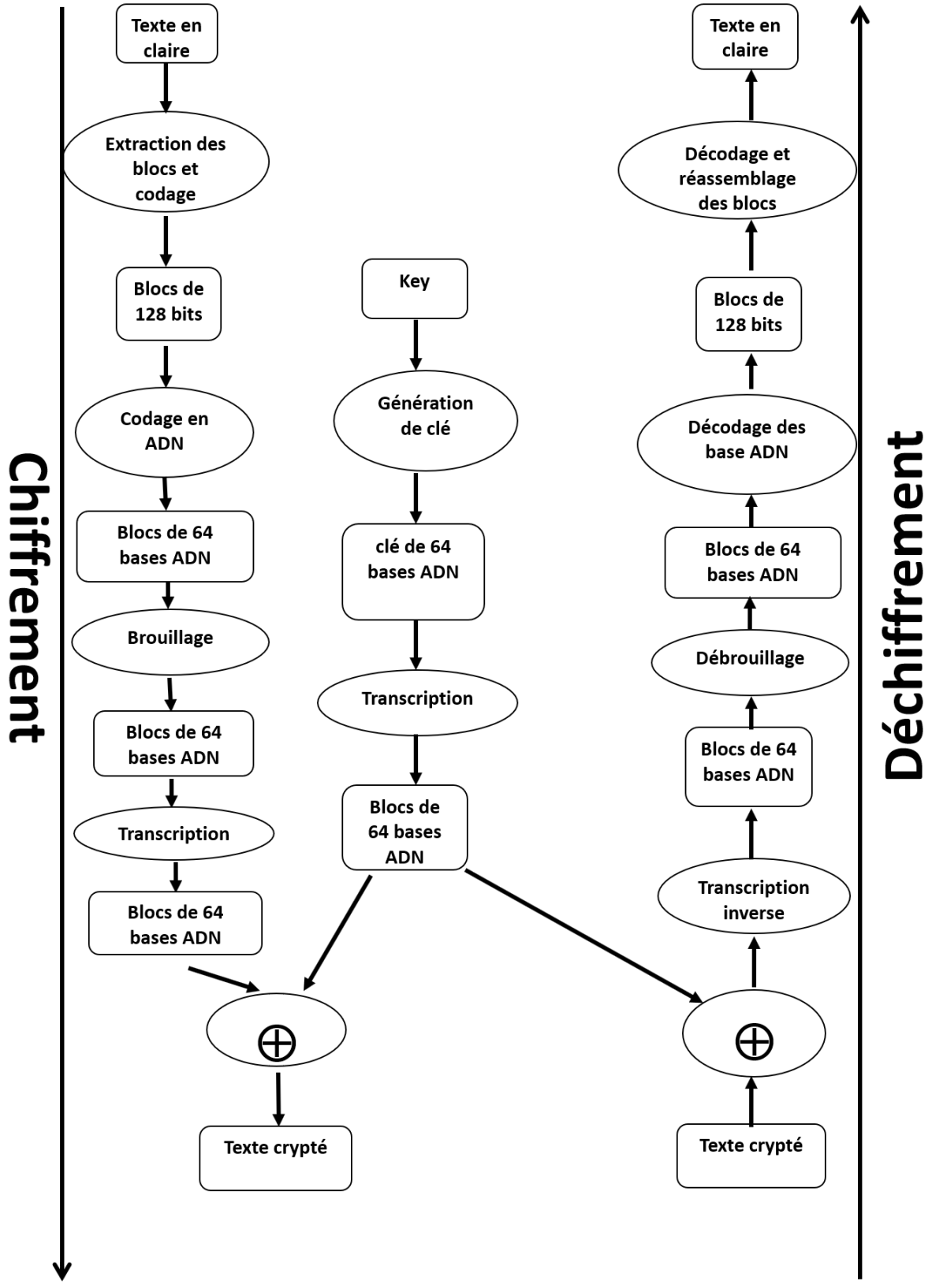


Figure 4.2: schéma fonctionnel chiffrement/déchiffrement

4.2.2 Codage en ADN

Le code génétique est l'ensemble des règles permettant de traduire les informations contenues dans le génome des cellules vivantes afin de synthétiser les protéines du coté scientifique. Du coté cryptographie ADN, les données binaires codées numériquement sont codées par une combinaison de deux états 0 ou 1.Quatre types de bases peuvent être utilisés pour coder l'ADN telles que l'adénine (A), Thymine (T), cytosine (C) et guanine (G).

Base ADN	Valeur binaire
A	00
C	01
G	10
T	11

Tableau 4.1: Codage en base nucléotide

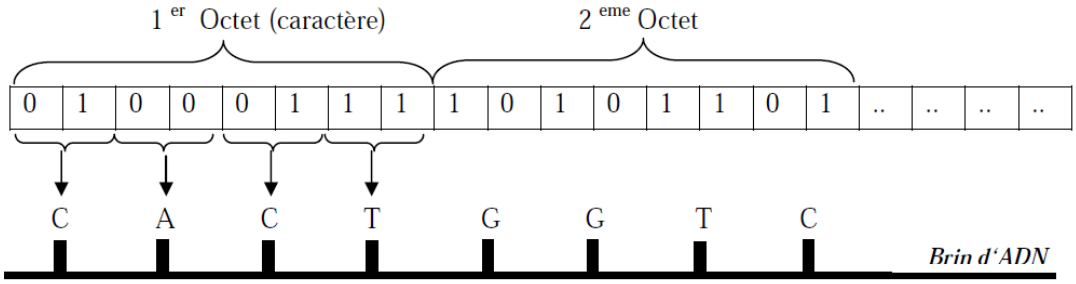


Figure 4.3 : Transformation Bits Base

4.2.3 Brouillage

Le tableau se compose de 256 caractères ADN.

0	**1**	**2**	**3**	**……**	**254**	**255**
AAAA	AAAC	AAAG	AAAT	……	TTTG	TTTT

Tableau 4.2 : de séquence d'ADN

En découpant le bloc en séquences de 4 nucléotides, à partir du tableau de séquence on extrait la position de la séquence, en appliquant la suite Fibonacci (2) pour faire un brouillage dans l'ordre des séquences dans le bloc (d'une autre façon décalage des séquences vers le bas).

La suite de Fibonacci est définie comme suit :

$$\begin{cases} f_{0=0} \\ f_{0=}1 \\ f_{n+2=}f_{n+1} + f_n \end{cases} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(2)$$

La formule de brouillage est :

R= position de la séquence ADN + fib (séquence)%256 ; …… (3)

4.2.4 La transcription

L'information contenue dans les gènes va servir à la fabrication de milliers de protéines qui interviennent dans le fonctionnement de la cellule. La première étape de l'expression d'un gène consiste à recopier son information sous la forme d'une molécule très proche de l'ADN, l'acide ribonucléique ou ARN. Dans le cas de notre code c'est transformer les nouvelles séquences d'ADN a des ARN selon le tableau suivant :

Base ADN	**Base ARN_m**
A	**U**
C	**G**
G	**C**
T	**A**

Tableau 4.3: conversion d'ADN en ARN_m

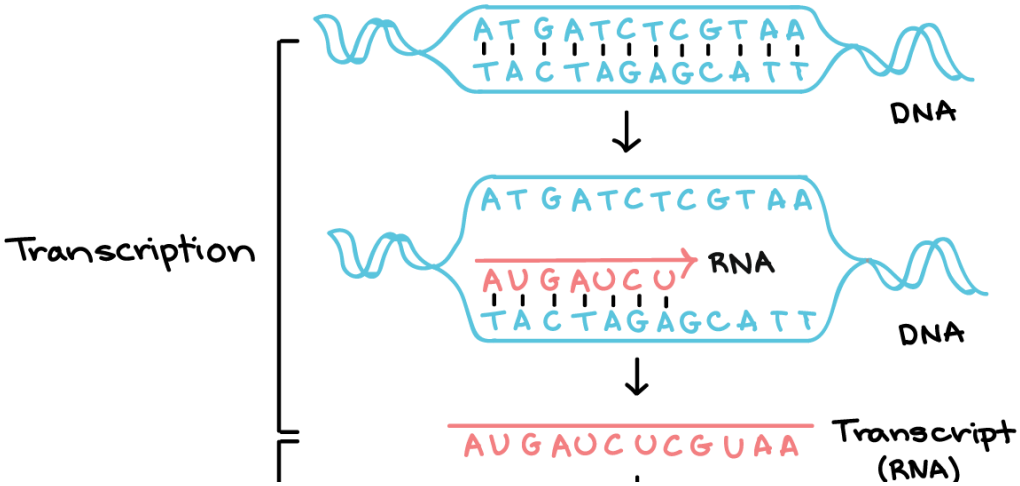


Figure 4.4 : la Transcription

4.2.5 Génération des clés

Notre algorithme utilise une clé composée de deux parties, sa première partie est un entier long **P** qui permet de calculer la position dans la séquence à partir de laquelle, on commence la génération des sous clés. **P** sera envoyé en clair par voie publique à l'émetteur.

Pos= **P** mod longueur (séquence)

Où :

P : Entier générée par l'émetteur

Séquence : est la séquence ADN à partir de laquelle on fait l'extraction des sous clés qui sont généralement un chromosome du gène humain.

Pos **:** c'est la position de départ dans la séquence ou l'extraction se commence.

La deuxième partie est le numéro du chromosome du gène humain, c'est l'identificateur de la séquence.

Les sous clés sont générées à partir de la position pos dans la séquence définie par numéro chromosome de taille 64 (128 bits), le nombre des sous clés générées égale au nombre des blocs crées dans la phase de découpage du texte chiffré.

Transcription de la clé

En procédant avec la même démarche utilisée dans la section **(4.3.4)**, on applique une transcription sur chaque sous clé.

4.2.6 Le XOR

Le bloc et la clé qui sont en ARNm, seront convertis de nouveau en binaire selon le tableau suivant :

Base ARN	Valeur binaire
U	00
G	01
C	10
A	11

Tableau 4.4: conversion d'ARN_m en binaire

La dernière phase est le xor entre le bloc et la clé

A	**B**	**A ⊕ B**
0	0	0
0	1	1
1	0	1
1	1	0

Tableau 4.5: XOR binaire

Le résultat de cette phase c'est le texte chiffré.

4.3 Le déchiffrement

L'algorithme de déchiffrement est utilisé pour déchiffrer les blocs numériques qui ont été chiffrés par l'algorithme de chiffrement. Les clés de chiffrement et de déchiffrement sont identiques.

L'algorithme de déchiffrement est complètement opposé à l'algorithme de chiffrement, et tous les modules présents dans le chiffrement sont inversés dans le déchiffrement selon les étapes suivantes :

1. Génération de la clé + transcription
2. Le texte chiffre XOR la clé utilise dans le chiffrement.
3. La transcription inverse
4. Débrouillage
5. Décodages des bases nucléotide.
6. Décodage et réassemblage des blocs
7. Texte en claire

4.4 Expérimentations et résultats

Cette partie a pour objectif d'évaluer et de tester les performances de notre algorithme. Nous souhaitons, en utilisant un ensemble de documents textuels, illustrer les avantages d'utilisation de notre système de chiffrement et déchiffrement dans l'environnement IoT.

Les tests des méthodes cryptographiques s'effectuent sur deux axes importants :

- Temps d'exécution : Pour évaluer les performances de l'algorithme de point de vue temps de chiffrement/déchiffrement.
- Sécurité : Portant sur la robustesse de la clé proposée aux différentes attaques

4.4.1 Environnement de développement

Au cours de notre phase d'expérimentation, nous avons travaillé sur une machine appelée Raspberry pi 3 model B qui contient les caractéristiques suivantes :

RAM	Processeur intégré	Ports
1 Go	Quad-core ARM Cortex-A53 1.2 GHz	HDMI, 4x USB, RJ45, jack3.5 mm, connecteurs pour APN et écran tactile

Tableau 4 .6 : les caractéristiques du raspberry pi 3 model B

4 .4.2 Langage de programmation

Nous avons implémenté notre algorithme en langage python.

Python est un langage qui peut être utilisé dans de nombreux contextes et adapté à tout type d'utilisation grâce à des bibliothèques spécialisées. Il est cependant particulièrement utilisé comme langage de script pour automatiser des tâches simples mais fastidieuses, comme un script qui récupérerait la météo sur Internet ou qui s'intégrerait dans un logiciel de conception assistée par ordinateur afin d'automatiser certains enchaînements d'actions répétitives. On l'utilise également comme langage de développement de prototype lorsqu'on a besoin d'une application fonctionnelle avant de l'optimiser avec un langage de plus bas niveau. Il est particulièrement répandu dans le monde scientifique, et possède de nombreuses bibliothèques optimisées destinées au calcul numérique. Python a été conçu pour être un langage lisible. Il vise à être visuellement épuré. Par exemple, il possède moins de constructions syntaxiques que de nombreux langages structurés tels que C, Perl, ou Pascal. Les commentaires sont indiqués par le caractère croisillon (#)

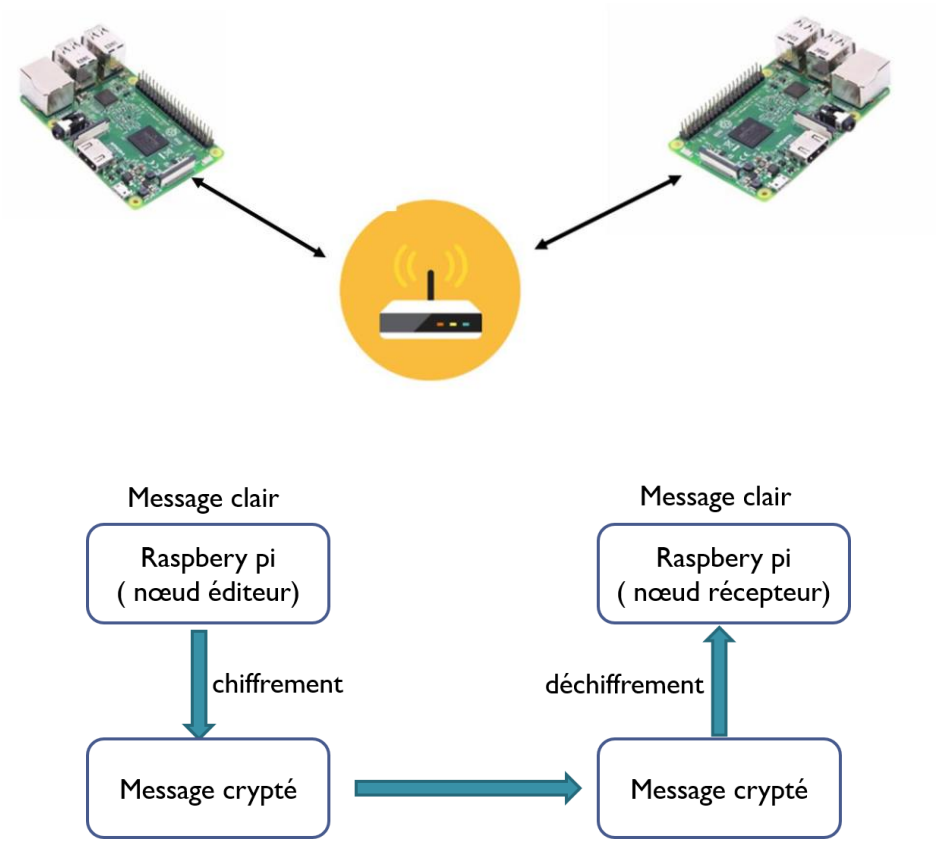


Figure 4.5 : Schéma fonctionnel de la phase des tests

4.4.3 Corpus de tests

Les textes

Nous avons utilisé un ensemble de documents textuels, de différentes tailles

Les séquences

Nous avons utilisé les chromosomes humains qui sont disponible sur le site NCBI

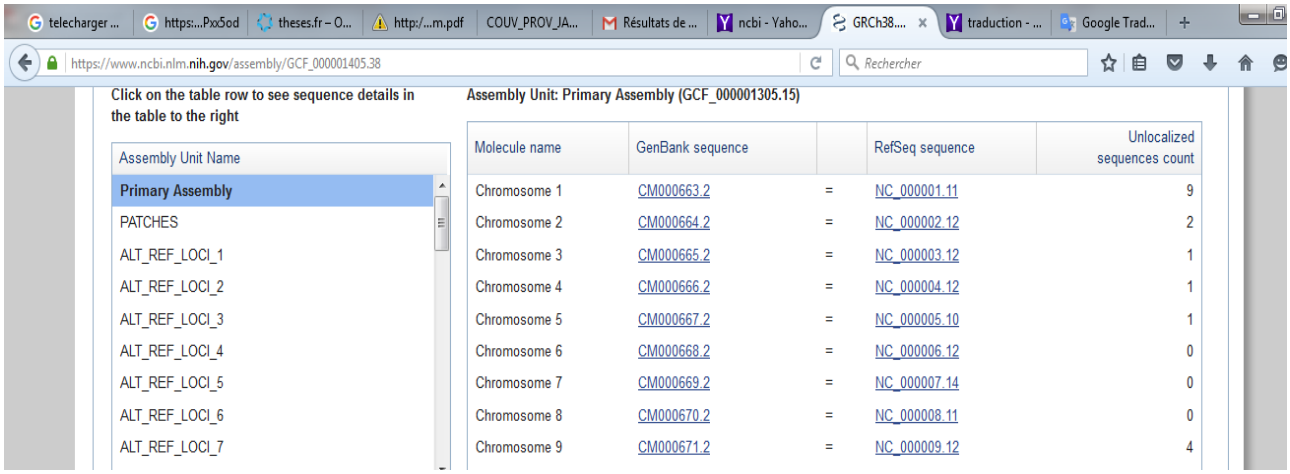


Figure 4.6 : Extrait- Les chromosome Humain (NCBI web site)

4.4.4 Evaluation du temps d'exécution

Nous avons effectué 3 tests pour différents fichiers, puis nous avons évalué le temps moyen de chiffrement/déchiffrement.

Exemple :

Test 1 : Un ***fichier texte*** avec les caractéristiques suivantes :

- Type de fichier : format texte.
- Taille est : 1.47 kilo-octets (exactement 1506 octets)
- Nombre de bloc : 95blocs (bloc de 128 bits)

Format de fichier	Temps de chiffrement	Temps de déchiffrement
Fichier texte	8ms	7ms
	9ms	10ms
	10 ms	9 ms

Tableau 4 .7 : Temps d'exécution chiffrement/déchiffrement du fichier texte.

- Temps moyen de chiffrement : 9 ms
- Temps moyen de déchiffrement : 8,7ms

Nous avons suivi la variation du temps de chiffrement/déchiffrement en augmentant la taille de fichier en clair, le tableau suivant regroupe l'ensemble des résultats obtenus :

Fichier	**Taille**	**Temps de chiffrement** (ms)	**Temps de déchiffrement** (ms)
Fichier 1	1.20 ko	6	7
Fichier 2	8.31 Ko	10	10
Fichier 3	102 Ko	27	28
Fichier 4	385Ko	75	74

Tableau 4.8 : Temps d'exécution avec la variation de la taille de fichier.

Les deux graphes suivants représentent la relation entre la taille de texte en clair (respectivement texte chiffré) et le temps de chiffrement (respectivement temps de déchiffrement).

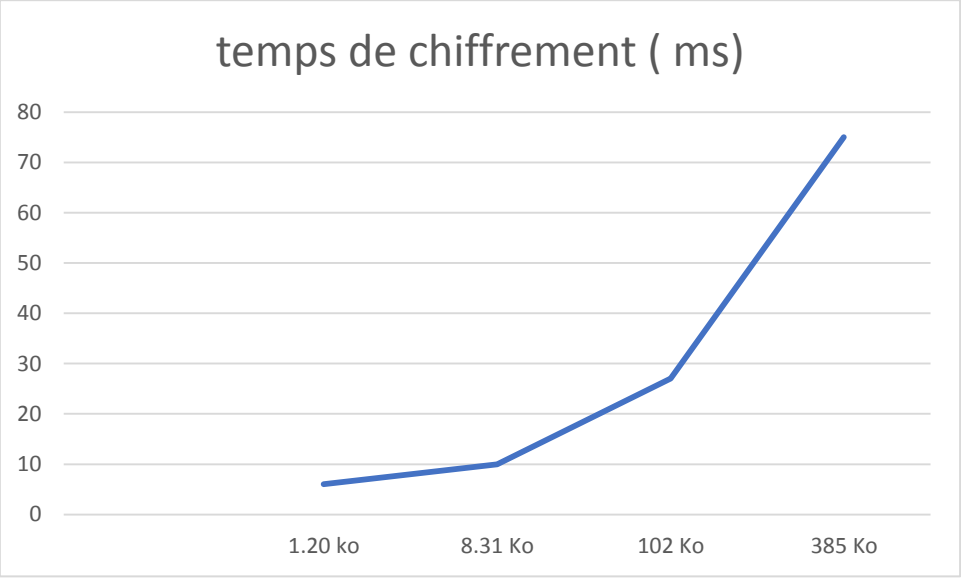


Figure 4.7: Graphe de la variation de taille de fichier pour le chiffrement

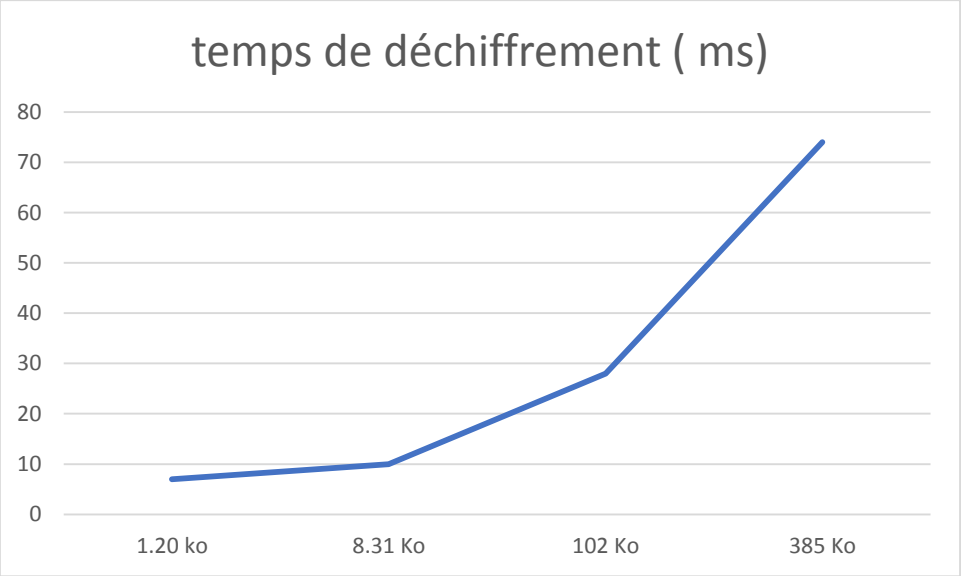


Figure 4.8 : Graphe de la variation de taille de fichier pour le déchiffrement

On remarque que le temps de chiffrement (respectivement déchiffrement) augmente d'une façon linéaire avec l'augmentation de la taille du texte clair. Cela s'explique qu'à chaque fois que le texte clair augmente de taille, le temps de chiffrement (respectivement déchiffrement) augmente autant de fois qu'il y a de blocs de 128 bits dans le texte clair. Le temps est jugé acceptable.

4.4.5 Evaluation de l'espace de la clé

L'espace clé d'un algorithme de chiffrement est un facteur très important dans la phase de test d'un algorithme de chiffrement. Il doit résister à l'attaque de force brute.

Notre clé est extraite à partir d'un chromosome humain ; nous prenons par exemple le chromosome qui contient environ 135 440 924 enregistrements de séquence d'ADN. Essayer chacune de ces séquences dans la construction de la clé sera équivalent à 2^{27} positions de départ.

Et donc la génération des sous clés avec les combinaisons possibles donne un nombre qui tend vers l'infini.

Conclusion

Dans ce chapitre, nous avons présenté une approche de chiffrement /déchiffrement d'un message texte transmis dans un environnement IoT et les expérimentations que nous avons mené.
Nous avons présenté deux phases d'évaluation, la première concerne le calcul du temps de chiffrement par rapport à la taille des textes, la deuxième concerne l'évaluation de la structure de clé proposée. Les résultats obtenus sont encourageants, surtout de point de vue robustesse de la clé proposée en tenant compte du niveau garanti de sécurité.

Conclusion générale

Conclusion générale

A la fin de ce mémoire, nous pourrons avancer que la sécurisation des informations dans une communication quelconque est un domaine complexe qui nécessite beaucoup d'attention de la part des chercheurs en cryptographie.

Dans ce sens, nous avons jugé utile d'évoquer au cours de cette recherche, le cadre théorique. Nous avons d'abord, revu l'état de l'art de la cryptographie .Ensuite, présenté un panorama sur la cryptographie par ADN dans lequel s'inscrit notre travail. Enfin, nous avons mis en exergue les caractéristiques de l'IoT

Pour répondre à notre questionnement, nous avons opté pour un programme en PYTHON pour faire une cryptographie par ADN dans un environnement IoT. C'est une nouvelle méthode qui permet la sécurisation des informations.

Dans cette méthode on a proposé un algorithme de cryptage symétrique qui a été réalisé en python sur un Raspberry pi. Le but de notre système est de chiffrer / déchiffrer un texte en les découpant en blocs de 16 caractères, la clé utilisée pour cette tâche est une clé symétrique extraite à partir d'un chromosome humain. Dans cet algorithme le chiffrement d'un texte passe par six étapes.

En commençant par l'extraction des blocs et codage qui permet de découper le texte original et calculer la fréquence et l'indicateur de chaque lettre pour calculer le caractère codé après le texte résultat sera converti en code ASCII puis en binaire, ensuite l'étape de codage en ADN qui consiste à transformer les données binaires en base ADN telles que l'adénine (A), Thymine (T), cytosine (C) et guanine (G). Dans l'étape suivante qui est le brouillage on a utilisé la fonction de Fibonacci qui sert à décaler les séquences vers le bas, l'étape de transcription comprend la conversion d'ADN en ARN_m .Cet algorithme est basé sur l'étape de la génération des clés qui est composé de deux parties : La première c'est de calculer la position de départ dans la séquence où l'extraction commence, la deuxième partie est le numéro du chromosome du gène humain puis on applique une transcription sur chaque sous clé. Pour finir la phase de chiffrement et avoir un texte chiffré on a reconverti en binaire le bloc et la clé pour faire un xor entre eux.

L'algorithme de déchiffrement passe aussi par des étapes qui sont complètement opposées aux étapes de l'algorithme de chiffrement

On a estimé que le temps de chiffrement/déchiffrement où nous avons utilisé un ensemble de documents textuels de différentes tailles et on a établi une vue graphique qui représente la variation du temps en fonction de la taille. On a remarqué que le temps de chiffrement /déchiffrement augmente d'une façon linéaire avec l'augmentation de la taille du texte. De là, nous dirons que notre modeste analyse pratique nous permet de confirmer notre hypothèse dans laquelle on a prédit que la cryptographie par ADN offre une bonne protection en tenant compte des caractéristiques limitées de raspberry.

Conclusion générale

Les résultats obtenus dans ce présent mémoire nous paraissent encourageants, surtout du point de vue de temps d'exécution et la robustesse de la clé proposée.

Nous croyons que ces résultats pourraient servir d'assise descriptive pour les recherches ultérieures dans le domaine de l'évolution de la sécurité par ADN qui demeure encore un champ de recherche vierge à exploiter où d'autres problématiques peuvent être soulevées.

Références bibliographies

Références bibliographiques

THESE DE DOCTORAT Présentée Par ZAIBI Ghada le thème : « Sécurisation par dynamiques chaotiques des réseaux locaux sans fil au niveau de la couche MAC » la Discipline : génie électrique ville : Toulouse Etablissement de soutenance : université de Sfax Ecole Nationale d'ingénieurs de Sfax Année : 2012 [1]

(Pierre Barthélemy, mai 2012) Livre de Cryptographie principe et mises en œuvre 2em édition revue et augmentée [2]

(Dartois, 2010)Cryptographie Paris 13 (version 2010/2011) d'après un cours de Daniel Barsky & Ghislain Dartois 1 octobre 2010 [3]

(Schneier, 2001)Livre de cryptographie appliquée Algorithmes, protocoles et codes source en C 2em édition [4]

https://www.supinfo.com/cours/1ARI/chapitres/01-introduction-cryptologie consulté le 5/02/2020 [5]

https://www.securiteinfo.com/cryptographie/cryptographie.shtml consulté le 5/02/2020 [6]

(Muller, 2018)Livre de Les codes secrets décryptés 3ème édition corrigée et augmentée [7]

http://www.cryptage.org/chiffre-cesar.html consulté le 6/02/2020 [8]

https://www.tala-informatique.fr/wiki/images/e/eb/Cryptographie.pdfconsulté le

07/02/2020 [9]

THESE DE DOCTORAT Présentée par KADDOURI Zakaria le thème : « MISE EN ŒUVRE DE NOUVELLES TECHNIQUES POUR LA SECURITE INFORMATIQUE BASEES SUR LES ALGORITHMES EVOLUTIONNISTES ET LES FONCTIONS DE HACHAGE » Discipline : Informatique ville : Maroc Rabat Etablissement de soutenance : UNIVERSITE MOHAMMED V FACULTE DES SCIENCES Rabat Année : 2014 [10]

https://www.apprendre-en-ligne.net/crypto/bibliotheque/PDF/IntroToCrypto.pdf consulté le 12 /02/2020 [11]

https://www.ibisc.univ-evry.fr/~petit/Enseignement/Chiffrement-compression/crypto-2009-2010/crypto-2009-2010-cours2.pdf consulté le 13/02/2020 [12]

https://arxiv.org/ftp/arxiv/papers/1904/1904.05528.pdf consulté le 19/03/2020 [13]

https://tel.archives-ouvertes.fr/tel-00942608 consulté le 15/02/2020 [14]

http://www.genoscope.cns.fr/externe/HistoireBM/ consulté le 16/02/2020 [15]

https://www.academia.edu/19849085/INTRODUCTION_A_LA_GENETIQUE_MOLECULAIRE consulté le 17/02/2020 [16]

Références bibliographiques

https://www.aquaportail.com/definition-530-adn.html consulté le 18/02/2020 [17]

https://www.police-scientifique.com/adn/structure-et-principe-de-base consulté le

18/02/2020 [18]

https://natyinfirmiere.files.wordpress.com/2010/10/linformation-genetique-adn-transfert-et-conservation.pdf consulté le 21/02/2020 [19]

https://studylibfr.com/doc/932987/sert-l-adn consulté le 23/03/2020 [20]

http://galactosemie.free.fr/telechargements/chromosome.pdfconsulté le 23/03/2020 [21]

https://www.aquaportail.com/definition-4529-base-azotee.html consulté le 27/02/2020 [22]

https://www.doctissimo.fr/sante/dictionnaire-medical/nucleotide consulté le 27/02/2020 [23]

https://www.aquaportail.com/definition-8498-nucleoside.html consulté le 28/02/2020 [24]

https://www.futura-sciences.com/sante/dossiers/genetique-gene-adn-proteines-1130/page/5/ consulté le 3/03/2020 [25]

https://www.pourlascience.fr/sd/genetique/adn-arn-une-copie-infidele-10969.php consulté le 10/03/2020 [26]

https://www.maxicours.com/se/cours/transcription-et-traduction-synthese-des-proteines/ consulté le 13/03/2020 [27]

https://www.ebiologie.fr/cours/s/35/la-replication-de-l-adn consulté le 15/03/2020 [28]

https://www.lesechos.fr/2002/11/ladn-molecule-informatique-703535 consulté le

16/03/2020 [29]

https://arxiv.org/ftp/arxiv/papers/1904/1904.05528.pdf consulté le 19/03/2020 [30]

https://www.techniques-ingenieur.fr/base-documentaire/technologies-de-l-information-th9/systemes-embarques-42588210/introduction-a-l-internet-des-objets-h8050/ consulté le 20/03/2020 [31]

(Noor A. Hussein, 2020)DNA computing based stream cipher for internet of things using MQTT protocol Noor A. Hussein, Mohamed Ibrahim Shujaa Department of Computer Engineering Techniques, Electrical Engineering Technical College, Middle Technical University (MTU), Iraq [32]

https://www.digora.com/fr/blog/definition-iot-et-strategie-iot consulté le 21/03/2020 [33]

https://www.oracle.com/dz/internet-of-things/what-is-iot.html consulté le 22/03/2020 [34]

Références bibliographiques

https://wikimemoires.net/2019/09/domaines-d-applications-de-l-iot/ consulté le 22/03/2020 [35]

https://blog.octo.com/modeles-architectures-internet-des-objets/ consulté le 05/04/2020 [36]

Table des matières

Table des matières

Table des matières

Table des matières

Résumé

A travers cette recherche nous avons essayé d'étudier le sujet de la cryptographie par ADN dans un environnement IoT qui est un nouveau domaine pour la sécurisation des informations. Cette méthode sert à combiner les avantages du matériel génétique avec les solutions classiques de la cryptographie ou l'information est stockée à l'intérieur de la molécule d'ADN et cachée parmi d'autres. Afin de vérifier notre hypothèse on s'est basé sur la méthode de cryptage symétrique ou on a réalisé un programme implémenté en python Sur un Raspberry pi en proposant de chiffrer des textes de différentes tailles basés sur une clé extraite à partir d'un chromosome humain. Après l'analyse des résultats nous sommes arrivés à montrer que cette méthode de la cryptographie par ADN s'articule autour deux axes importants, le temps d'exécution Pour évaluer les performances de l'algorithme d'une part et la robustesse de la clé d'autre part, tout cela pour assurer une bonne sécurité d'informations.

Mots clés: cryptographie, ADN, sécurisation, cryptage, informations,

Abstract

Through this research we have tried to study the subject of DNA cryptography in a sensor network with which is a new field for securing information. This method is used to combine the advantages of genetic material with the classical solutions of cryptography where information is stored inside the DNA molecule and hidden among others. In order to verify our hypothesis we used the symmetrical encryption method where we realized a program implemented in python on a Raspberry pi proposing to encrypt texts of different sizes based on a key extracted from a human chromosome. After analyzing the results we have been able to show that this method of DNA cryptography has two important axes, the execution time to evaluate the performance of the algorithm on the one hand and the robustness of the key on the other hand, all this to ensure good information security.

Key words: cryptography, DNA, security, encryption, information

الملخص

من خلال هذا البحث حاولنا دراسة موضوع تشفير الحمض النووي في بيئة إنترنت الأشياء والتي تعد مجالًا جديدًا لأمن المعلومات. تستخدم هذه الطريقة للجمع بين مزايا المادة الجينية مع الحلول التقليدية للتشفير حيث يتم تخزين المعلومات داخل جزيء الحمض النووي وإخفائها من بين أمور أخرى. من أجل التحقق من فرضيتنا ، استخدمنا طريقة التشفير المتماثل أو أنتجنا برنامجًا تم تنفيذه في برنامج بايثون في جهاز راسبيري بي من خلال اقتراح تشفير نصوص بأحجام مختلفة بناءً على مفتاح مستخرج من كروموسوم بشري. بعد تحليل النتائج، تمكنا من إظهار أن طريقة تشفير الحمض النووي هذه تدور حول محورين مهمين، وقت التنفيذ لتقييم أداء الخوارزمية من جهة وقوة المفتاح من ناحية أخرى، كل هذا لضمان أمن المعلومات الجيد

الكلمات الأساسية: التشفير، الحمض النووي، الأمن، التشفير، المعلومات