République Algérienne
Démocratique & Populaire
Ministère de l'Enseignement Supérieur & de la Recherche Scientifique

UNIVERSITE SAIDA - Dr. MOULAY Tahar

Faculté : Technologie **Département : Informatique**

Mémoire de fin d'étude
Pour l'obtention du diplôme de Master

Option : Sécurité Informatique et Cryptographie

Thème

Cryptanalyse du chiffrement de Vigenère par une technique heuristique : (Algorithmes Génétiques)

Présenté par :

AMER Nabil
ABOUDA Benyasaad

Encadré par :

Ahmed Chaouki LOKBANI

Promotion : 2019/2020

Remerciement

Nous remercions tout d'abord Dieu, le miséricordieux de nous avoir aidé et donné la force et le courage.

Qu'il nous soit permis ici de dire nos gratitudes

Nous tenons à remercier notre encadreur Monsieur ***Ahmed ChaoukiLOKBANI*** *qui s'est toujours montré à l'écoute et très disponible tout au long de la réalisation de ce mémoire, ainsi pour l'inspiration, l'aide et le temps qu'il a bien voulu nous consacrer et sans qui ce mémoire n'aurait jamais vu le jour.*

Nous tenons aussi à exprimer nos remerciements aux membres du jury qui ont accepté d'évaluer notre travail.

À remercier sincèrement Monsieur ***Mehdi FEZZA*** *pour le soutien et l'aide qu'il n'a jamais manqué de nous apporter, aussi pour ces conseils et ces orientations durant l'élaboration de ce travail.*

Enfin, nous adressons nos plus sincères remerciements à tous ceux qui de prés ou de loin on apporté un effort pour l'élaboration et la mise en forme de ce modeste travail.

Merci à tous.

DEDICACES

Je dédie ce modeste travail
A ma très chère mère
A toi mon père
A, mes sœurs, mes frères et mes chers amis.

Abouda Benyasaad

DEDICACES

JE DEDIE CE MEMOIRE A …

Mon père

Aucune dédicace ne saurait exprimer l'amour, l'estime, le dévouement et le respect que j'ai toujours eu pour vous.

Rien au monde ne vaut les efforts fournis jour et nuit pour mon éducation et mon bien être.

Ma très chère mère

Je te dédie ce travail en témoignage de mon profond amour. Puisse dieu, le tout puissant, te préserver et t'accorder sante, longue vie et bonheur.

les fleurs de notre maison : Rahaf, Yassemin, Hadjer

Je vous dédie ce travail avec tous mes vœux de bonheur, de sante et de réussite.

A tous les membres de ma famille, petits et grands

Veuillez trouver dans ce modeste travail l'expression de mon affection

A mes chère professeures

Un remerciement particulier et sincère pour tous vos efforts fournis. Vous avez toujours été présents.

Que ce travail soit un témoignage de ma gratitude et mon profond respect.

Merci…Nabil

Table des matières

Table des matières

Chapitre III: Les méthodes évolutionnaires

Liste des tables

Introduction générale

L'échange de données (texte) pour l'homme est une nécessité. La sécurité de cette opération devient parfois plus qu'une exigence. Ainsi depuis César à l'ère de l'informatique, le chiffrement de certains messages a toujours été un besoin afin de les cacher à tout intrus non autorisé de façon à s'abriter d'un éventuel usage malveillant.

De nos jours, l'ensemble de ces méthodes a été regroupé dans une branche appelée la cryptographie. Parallèlement, une autre branche ennemie à la cryptographie appelée la cryptanalyse a été développée, qui est l'art de révéler le texte en clair d'un texte chiffré sans connaître la clé de déchiffrement.

.La cryptologie comporte le domaine de la cryptographie (qui est le domaine où on cherche à protéger le secret) et de la cryptanalyse (qui est le domaine où on cherche à retrouver le message d'origine sans connaitre exactement tout le procédé).

Ces deux domaines étant fortement liés, il n'est pas rare de voir des Cryptologues et des Cryptanalystes travailler ensemble.

Pour cela, depuis des siècles, de nombreux mécanismes ont été inventés. Nous remontons à la fin du XVIe siècle pour étudier le chiffrement de Vigenère, qui n'est qu'une extension du chiffrement de César. Cette extension semble beaucoup plus robuste du fait qu'une clé de 20 caractèresprésente$2 \times 10^{26} = 2^{94}$possibilités, ce qui nécessiterait plus d'un million d'années pour toutes les essayer sur un parc d'un million de machines ! Et pourtant, elle ne résiste pas aux cryptanalyses utilisant les redondances de la langue française.

Nous allons utiliser un des algorithmes heuristiques ou évolutionnistes en l'occurrence (AE : Algorithmes génétiques), pour décrypter un système basé sur le chiffrement de Vigenère.

Les algorithmes évolutionnistes font partie de la classe des algorithmes basés sur une population et cet élément est donc au cœur de leur fonctionnement. Toutefois avoir une population n'est pas suffisant.

En effet, pour utiliser un algorithme évolutionniste il faut: Un problème d'optimisation, une population, une représentation des individus de la population, une méthode d'évaluation des individus ("fitness") des méthodes d' évolution (croisement, mutation) et de sélection des individus, un critère de terminaison (nombre d'itérations / performance)

Ce projet se compose de quatre chapitres, dont le premier chapitre donne une introduction générale à la sécurité informatique.

Le deuxième chapitre est consacrée à la recherche théorique de la cryptographie et autres Surtout pour le problème de Vigenère, qui a été décrit plus en détail.

Dans le troisième chapitre de notre travail, nous nous sommes d'abord concentrés sur l'algorithme évolutif, dans lequel nous spécifions Les paramètres qui sont utilisés pour résoudre le problème de Vigenère.

Enfin, dans la quatrième partie, nous avons mis en place un algorithme génétique qui résout le problème de Vigenère.

L’essentiel de notre travail est de programmer le chiffrement d’un texte clair par la méthode de Vigenère (cryptage de texte avec une clé secrète) et de le décrypter en s’inspirant d’ un algorithme évolutionniste (AE) : Algorithme Génétique comme outil de base pour retrouver tout ou une partie du texte crypté.

Introduction :

Les systèmes d'information prennent de plus en plus une place stratégique au sein des entreprises. Ainsi la notion du risque lié à ces derniers devient une source d'inquiétude et une donnée importante à prendre en compte, ceci en partant de la phase de conception d'un système d'information jusqu'à son implémentation et le suivi de son fonctionnement.

Les pratiques associées à la sécurité des systèmes d'information constituent un point à l'importance croissante dans l'écosystème informatique qui devient ouvert et accessible par utilisateurs, partenaires et fournisseurs de services de l'entreprise. Il devient essentiel pour les entreprises de connaître leurs ressources en matière de système d'information et de définir les périmètres sensibles à protéger afin de garantir une exploitation maîtrisée et raisonnée de ces ressources.

Par ailleurs, les nouvelles tendances de nomadisme et de l'informatique « in the Cloud » permettent, non seulement, aux utilisateurs d'avoir accès aux ressources mais aussi de transporter une partie du système d'information en dehors de l'infrastructure sécurisée de l'entreprise. D'où la nécessité de mettre en place des démarches et des mesures pour évaluer les risques et définir les objectifs de sécurité à atteindre.

Ainsi la sécurité informatique est un ensemble de moyens techniques, organisationnels, juridiques et humains nécessaires pour conserver, rétablir et garantir la sécurité du système d'information

On peut déduire de ces constats que la démarche de sécurité informatique est une activité managériale des systèmes d'information et qu'il convient aussi d'établir un tableau de bord de pilotage associé à une politique de sécurité comprenant les organes vitaux constituant une entreprise.

1) Définition :

La sécurité informatique c'est l'ensemble des moyens mis en œuvre pour réduire la Vulnérabilité d'un système contre les menaces accidentelles ou intentionnelles. Il convient d'identifier les exigences fondamentales en sécurité informatique, qui caractérisent ce à quoi s'attendent les utilisateurs des systèmes informatiques en regard de la sécurité :**[01]**

1. **Disponibilité :** demande que l'information sur le système soit *disponible* aux personnes autorisées.
2. **Confidentialité :** demande que l'information sur le système ne puisse être *lue* que par les personnes autorisées.
3. **Intégrité :** demande que l'information sur le système ne puisse être *modifiée* que par les personnes autorisées.**[01]**

Pour considérer efficacementles besoins de sécurité d'une organisation et évaluer et choisir les nombreux produits et politiques de sécurité, le responsable de la sécurité abesoin de moyens systématiques de définition des exigences de sécurité et de caractérisation des approches qui satisfont ces exigences. Une approche possible est de considérer trois aspects de la sécurité de l'information :

- Services de sécurité : un service qui améliore la sécurité des systèmes informatiques et des transferts d'information d'une organisation. Les services sont conçus pour contrer les attaques de sécurité, et ils utilisent un ou plusieurs mécanismes de sécurité ;
- Mécanismes de sécurité : un mécanisme est conçu pour détecter, prévenir ou rattraper une attaque de sécurité.
- Attaque de sécurité : une action qui compromet la sécurité de l'information possédée par une organisation.

Du point de vue de la sécurité informatique, une menace est une violation potentielle de la sécurité. Cette menace peut-être accidentelle, intentionnelle (attaque), active ou passive. **[02]**

2) Les Attaques informatiques :

La sécurité de l'information traite de la prévention de la fraude, ou, à défaut, de sadétection dans des systèmes d'information à l'intérieur desquels l'information elle-même n'apas d'existence physique significative. On verra dans les transparents suivants une listed'exemples évidents de tricherie, qui se sont produits dans des cas réels. Ce sont desexemples d'attaques spécifiques qu'une organisation ou un individu peut avoir à affronter. Lanature de l'attaque varie considérablement selon les circonstances. Heureusement, il estpossible d'approcher le problème en examinant les types génériques d'attaques pouvant êtrerencontrées.Une attaque peut être définit comme toute action ou ensemble d'actions qui peut porter atteinte la sécurité des informations d'un système ou d'un réseau informatique. **[09]**

3) Typologie des attaques informatique :

3.1Virus

Un virus informatique est un programme, généralement de petite ou très petite taille, doté des propriétés suivantes : infection ; multiplication ; possession d'une fonction nocive (payload). La fonction d'infection permet au virus de s'introduire dans des fichiers de programme, dans des fichiers de données utilisant un langage de script, ou dans une partie de la disquette ou du disque dur contenant un petit programme (secteur de démarrage). Lors de l'accès à ces programmes ou secteur, le code du virus s'exécutera de façon d'abord silencieuse (phase de multiplication pendant laquelle il infectera d'autres fichiers) puis visible (activation de la fonction nocive).**[01]**

3.2Les Vers

Un verest un programme autonome qui se reproduit et se propage àl'insu des utilisateurs. Contrairement aux virus, un ver n'a pas besoind'un logiciel hôte pour se dupliquer. Le ver a habituellement unobjectif malicieux, par exemple :

- Espionner l'ordinateur dans lequel il réside ;
- Offrir une porte dérobée à des pirates informatiques ;
- Détruire des données sur l'ordinateur infecté ;
- Envoyer de multiples requêtes vers un serveur internet dans le but de le saturer.

Le ver Blaster avait pour but de lancer une attaque par déni de servicesur le serveur de mises à jour de Microsoft.Un ver est donc un virus réseau.**[17]**

3.3 Cheval de Troie

Un Cheval de Troie (trojan en anglais) est un programme effectuantune fonction illicite tout en donnant l'apparence d'effectuer unefonction légitime. La fonction illicite peut consister enla divulgationou l'altération d'informations. Trojan.ByteVerify est un cheval deTroie sousforme d'une applet java. Ce cheval de Troie exploite unevulnérabilité de la machine virtuellejava de Microsoft permettant àun pirate d'exécuter du code arbitraire sur la machine infectée.

Parexemple, Trojan.ByteVerify peut modifier la page d'accueil d'InternetExplorer.

C'est un programme qui se cache lui-même dans un autre programme apparemment au-dessus de tout soupçon. Les plus célèbres sont Back Orifice, Net bus et SubSeven.Ces programmes sont composés de deux parties :

Un module serveur : l'équivalent des soldats grecs cachés dans le célèbre cheval. En informatique, ce module doit sembler, lui aussi, anodin. Il peut se présenter sous la forme d'un jeu ou d'une pièce jointe à un courriel.**[09]**

Un module client : l'équivalent de l'armée grecque entrant dans la ville une fois les portes ouvertes. Concernant un ordinateur, il s'agit d'un code malveillant. Une fois mis en place, l'infection peut consister à récupérer des données, détruire des fichiers ou encore « suivre » l'enregistrement des touches du clavier, etc. **[09]**

3.4 Porte dérobée

Une porte dérobée (ou backdoor en anglais) est un moyen decontourner les écanismesde contrôle d'accès. Il s'agit d'une failledu système de sécurité due à une faute de conceptionaccidentelle ouintentionnelle (cheval de Troie en particulier). C'est donc unefonctionnalitéinconnue de l'utilisateur légitime qui donne un accèssecret au logiciel. Une porte dérobée aété découverte dans le SGBDinterbase de Borland au début des années 2000. Il suffisaitd'entrer lenom d'utilisateur " politically " et le mot de passe " correct " pour seconnecter à labase de données avec les droits d'administrateur.**[09]**

3.5Bombe logique

Une Bombe logique est une partie d'un programme malveillant (virus,cheval de Troie, etc.)qui reste dormante dans le système hôte jusqu'àce qu'un instant ou un événement survienne,ou encore que certainesconditions soient réunies, pour déclencher des effets dévastateurs enson sein. Le virus Tchernobyl, qui fut l'un des virus les plusdestructeurs, avait une bombelogique qui s'est activée le 26 avril1999, jour du treizième anniversaire de la catastrophenucléaire deTchernobyl.**[01], [09]**

3.6 Intrusion

L'intrusion dans un système informatique a généralement pour but la réalisation d'une menace et est donc une attaque. Les conséquences peuvent être catastrophiques : vol, fraude, incident diplomatique, chantage…

Le principal moyen pour prévenir les intrusions est le coupe-feu ("firewall"). Il est efficace contre les fréquentes attaques de pirates amateurs, mais d'une efficacité toute relative contre des pirates expérimentés et bien informés. Une politique de gestion efficace des accès, des mots de passe et l'étude des fichiers « log » (traces) est complémentaire.**[19]**

3.7Social engineering

C'est une technique qui a pour but d'extirper des informations à des personnes sans qu'elles ne s'en rendent compte. Contrairement aux autres attaques, elle ne nécessite pas de logiciel. La seule force de persuasion est la clé de voûte de cette attaque. Il y a quatre grandes méthodes de social engineering : par téléphone, par lettre, par internet et par contact direct. **[21]**

4) ESPIONNAGE :

4.1L'homme du milieu :

Une attaque qui a pour but d'intercepter les communications entre deux parties, sans que ni l'une ni l'autre ne puissent se douter que le canal de communication entre elles a été compromis. L'attaquant doit d'abord être capable d'observer et d'intercepter les messages d'une victime à l'autre. **[15]**

4.2 Le Spyware :

Un spyware est un logiciel espion qui collecte des données personnelles afin de les envoyer à un tiers. Ce type de programme malveillant est la plupart du temps caché dans des logiciels gratuits ou des partagiciels (*shareware*), mais il peut aussi se propager depuis une page Internet infectée. On en trouve notamment sur les sites Web proposant des contenus illégaux. **[05]**

4.3 Cookies :

Un cookie est un petit fichier très simple, en fait un texte, enregistré sur le disque dur de l'ordinateur d'un internaute à la demande du serveur gérant le site Web visité. Il contient des informations sur la navigation effectuée sur les pages de ce site. **[10]**

5) Absence de politique de sécurité :

Une stratégie de sécurité est le point d'entrée de tout bon plan de sécurité. À l'instar des règles de travail, la politique de sécurité ne prend que quelques pages pour refléter l'attitude de sécurité de l'entreprise. Les finalités de sécurité s'articulent auteur de 5 axes :

- ❖ Améliorer la prise de conscience des risques du système d'information et des méthodes de prévention des risques.
- ❖ Créer une structure chargée de formuler et de mettre en œuvre des instructions et des processus cohérents pour assurer la sécurité des systèmes informatiques.
- ❖ Promouvoir la coopération entre les différents départements et départements de l'entreprise pour développer et mettre en œuvre les normes et procédures définies.
- ❖ Renforcez la confiance dans le système d'information de l'organisation.
- ❖ Promouvoir le développement et l'utilisation de systèmes d'information pour tous les utilisateurs autorisés de l'entreprise. **[13]**

6) PROTECTIONS :

6.1 L'Authentification:

L'authentification est une procédure, par laquelle un système informatique certifie l'identité d'une personne ou d'un ordinateur. Le but de cette procédure étant d'autoriser la personne à accéder à certaines ressources sécurisées. Il va comparer les informations des utilisateurs autorisés stockées dans une base de données (en local ou sur un serveur d'authentification) à celles fournies. L'accès sera autorisé seulement si les informations sont identiques. C'est l'administrateur du système d'information qui octroie les droits et paramètre l'accès. L'utilisateur possédant un compte d'accès (identifiant + mot de passe) n'aura accès qu'aux ressources dont il est autorisé à voir.

Il existe 4 facteurs d'authentifications qui peuvent être utilisés dans le processus d'autorisation d'accès à des ressources bloquées et sécurisées :

- **Ce que l'on connait (facteur mémoriel)** : une information que l'utilisateur a mémorisée et que lui seul connait (exemple : un mot de passe, un nom)
- **Ce que l'on possède (facteur matériel)** : une information que seul l'utilisateur possède et enregistrée dans un support (exemple : une clé USB).

- **Ce que l'on est (facteur corporel)** : une information qui caractérise l'utilisateur avec une empreinte qui lui est propre (exemple : voix, pupille, empreinte digitale)
- **Ce que l'on sait faire (facteur réactionnel) :** une information ou un geste que seul l'utilisateur peut produire (exemple : une signature). **[18]**

6.2 Antivirus :

Les antivirus sont des logiciels informatiques ayant pour objectif de détecter et supprimer les virus et malwares du poste ou du flux analysé. Il existe plus de 50 antivirus commerciaux actuellement et quelques antivirus open-source. La qualité d'un antivirus dépend en grande partie de sa rapidité à identifier les nouveaux virus et à mettre à jour sa base de signatures antivirus. Il est donc souvent intéressant d'utiliser des antivirus comportant plusieurs bases de signatures de virus et/ou différents antivirus.**[14]**

6.3Pare-feu :

Un pare-feu, aussi appelé firewall, désigne un dispositif de sécurité chargé de protéger un ordinateur contre les tentatives d'intrusion malveillantes provenant d'un réseau auquel il est connecté. Autrement dit, il s'agit d'un logiciel ou d'un matériel dont la principale fonction est de protéger votre ordinateur contre les virus et contre le piratage lorsque votre ordinateur est relié à n'importe quel réseau, dont Internet.

Le pare-feu analyse pour cela les données entrant sur un réseau (les paquets IP) et les données échangées. Il détecte, dans ces données, la présence d'éléments malveillants pour pouvoir les bloquer. Le pare-feu agit en quelques sortes comme un mur de protection contre tout ce qui pourrait endommager votre ordinateur. **[14]**

6.4 Le cryptage :

Le cryptage consiste à transformer un texte normal en charabia inintelligible appelé texte chiffré. Cette opération permet de s'assurer que seules les personnes auxquelles les informations sont destinées pourront y accéder. **[20]**

6.5 Traduction d'adresse :

Elle consiste à modifier l'adresse IP source ou destination, dans l'en-tête d'un datagramme IP lorsque le paquet transite dans le Pare-feu (Proxy) en fonction de l'adresse source ou destination et du port source ou destination. Lors de cette opération, le Pare-feu garde en mémoire l'information lui permettant d'appliquer la transformation inverse sur le paquet de retour. La traduction d'adresse permet de masquer le plan d'adressage interne (non routable) à l'entreprise par une ou plusieurs adresses routables sur le réseau externe ou sur Internet. Cette technologie permet donc de cacher le schéma d'adressage réseau présent dans une entreprise derrière un environnement protégé. **[3]**

7) Conclusion :

La sécurité à 100% reste un idéal à atteindre, surtout devant le large éventail des menaces qui mettent en danger l'exploitation d'un système d'information.

INTRODUCTION :

La cryptologie englobe deux domaines à savoir la cryptographie (qui est le domaine où on cherche à protéger le secret) et la cryptanalyse (qui est le domaine où on cherche à retrouver le message d'origine sans connaitre exactement tout le procédé).

Ces deux domaines étant fortement liés, il n'est pas rare de voir des cryptologies et des cryptanalyses travailler ensemble.

La cryptologie est en même temps une technique, un art et une science dons le domaine d'application est maintenant souvent plus économique que militaire.

Les cryptologies, autrefois vus comme des gens louches, espions au mieux, sorciers au pire, sont aujourd'hui au centre de notre société. Rendu indispensable par la diversification du transport de l'information (via Internet par exemple), les cryptologies cherchent sans relâche de nouveaux algorithmes de cryptage pour rendre impossible la cryptanalyse.

Nous essaierons au travers de la présente ce chapitre de comprendre les grands principes et catégories de chiffrement, et la classification des algorithmes de cryptage selon la clé de cryptage et selon la structure de cryptage et on parle sur les deux types symétrique et asymétrique, on parle aussisur Cryptographie classique et moderne, Enfin, avant la conclusion, nous dressons un schéma général de cryptologie Pour mieux comprendre le chapitre.

1) Vocabulaire

- Chiffrer : transformer à l'aide d'une convention secrète, appelée clé, des informations claires en informations inintelligibles pour des tiers n'ayant pas la connaissance du secret.
- Déchiffrer : retrouver les informations claires, à partir des informations chiffrées en utilisant la convention secrète de chiffrement.
- Décrypter : retrouver l'information intelligible, à partir de l'information chiffrée sans utiliser la convention sécrète de chiffrement.
- Clé : Une clé est un paramètre utilisé en entrée d'une opération cryptographique (chiffrement, déchiffrement, ...)
- Message clair : chaine de caractères composée de lettres de l'alphabet et dont on veut en général conserver la confidentialité´

2) Historique :

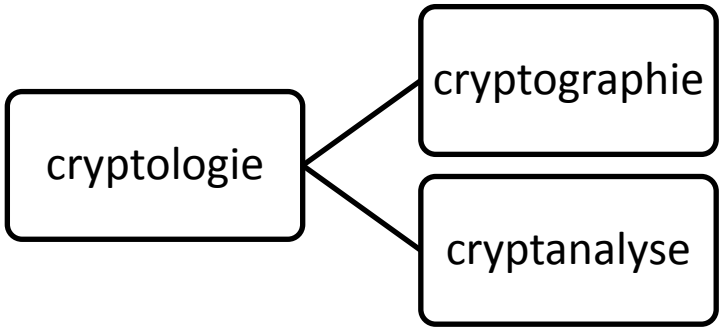
Avant de tentent de donner une définition de la cryptologie nous parlerons de l'histoire de la cryptologie.

L'origine de la cryptologie remonte Sans aucun doute aux origines de l'homme, Cela est dû à cela le secret à toujours tenu une place importante dans la vie de l'homme Surtout sur les plans militaires et informatiques.

- L'origine de mot cryptologie et de grec crypto qui signifie secret et logie que signifie science.
- Le premier ordinateur a été inventé pour cryptanaliser un message.
- Des batailles ont été gagnées grâce à la cryptanalyse.
- On considère que la seconde guerre mondiale a été raccourcie d'au moins un an grâce à la connaissance des communications allemandes par les alliés.

3) Généralités sur la Cryptologie :

La cryptologie est l'ensemble formé de la cryptographie et de la cryptanalyse.



La cryptologie fait partie d'un ensemble de théories et de techniques liées `a la transmission de l'information (théorie des ondes électromagnétiques, théorie du signal, théorie des codes correcteur d'erreurs, théorie de l'information, théorie de la complexité,...). **[04]**

3.1 Cryptographie :

La cryptographie est l'art de cacher l'information, de la rendre accessible uniquement à un nombre restreint de personnes, Pour se faire on transforme le message pour le rendre illisible mais de manière à pouvoir réobtenir le message d'origine. **[04]**

3.2 Cryptanalyse :

Jusqu'à une époque assez récente, la sécurité d'un chiffrement reposait tout autant sur le non-divulgation de l'algorithme utilisé que sur la clé. Ce n'est plus le cas aujourd'hui et on attend d'un algorithme, conformément au principe de Kerckhoffs, qu'il reste solide lorsque l'attaquant en connaît la spécification et ignore seulement la clé utilisée. La cryptanalyse d'un système peut être alors soit partielle (l'attaquant découvre le texte clair correspondant à un ou plusieurs messages chiffrés interceptés), soit totale (l'attaquant peut déchiffrer tous les messages, par exemple en trouvant la clé). Il existe plusieurs types d'attaques selon les moyens dont dispose l'attaquant :

- Attaques à chiffré connu : l'attaquant a seulement accès à des messages chiffrés.
- Attaques à clair connu : l'attaquant dispose d'un ou plusieurs messages clairs et les chiffrés correspondants.
- Attaques à clair choisi : l'attaquant choisit des clairs et peut obtenir les chiffréscorrespondants.
- Attaques à chiffré choisi : l'attaquant peut déchiffrer les messages de son choix. **[07]**

L'attaque la plus simple et la plus brutale est la recherche exhaustive. L'attaquant teste l'ensemble des clés possibles sur un cryptogramme donné dont il est supposé connaître au moins partiellement le clair ; il a découvert la bonne clé lorsque le déchiffrement redonne le clair attendu. La complexité d'une recherche exhaustive sur un algorithme de chiffrement par blocs dont la taille des clés est n bits est de l'ordre de 2^n chiffrements. Il est donc nécessaire lorsque l'on souhaite créer un algorithme de chiffrement de prendre des clés suffisamment longues pour se prémunir contre ce type d'attaque. Malheureusement, ce n'est pas la seule condition à vérifier. C'est pourquoi, la construction de tels algorithmes doit s'appuyer sur des problèmes mathématiques difficiles.

Dans notre travail, pour casser ce système de chiffrement, nous avons opté pour l'utilisation des algorithmes évolutionnistes (AE). **[07]**

4) Domaine de la Cryptologie :

La cryptologie est aujourd'hui présente partout dans notre quotidien.« L'homme moderne porte sur lui, sans forcément le savoir, un ou plusieurs processeurs cryptographiques » (Jacques Stern) Téléphone mobile, carte bancaire, passeport biométrique, carte d'assuré social, carte monéo, passe navigo, carte de TV à péage, clé de démarrage de véhicule, badge d'accès Vigik, carte vitale.

5) Principes d'un système cryptographique :

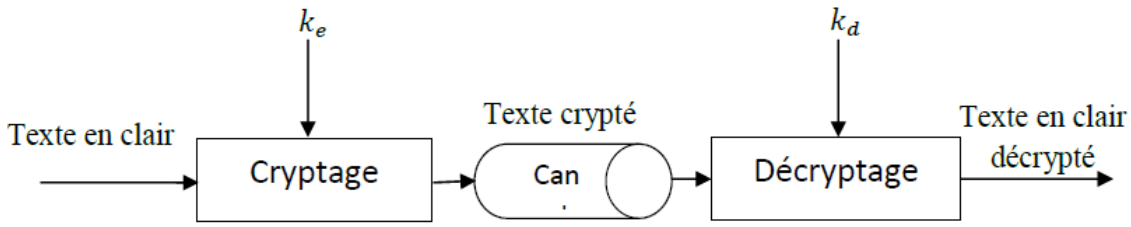


Figure 1:Schéma d'un système cryptographique.

Dans le système cryptographique de la figure 1.1, le résultat de cryptage d'un message appelé texte en clair (plain texte) et noté P est un texte crypté (ciphertext) noté C, la fonction de cryptage notée E_{k_e} transforme P en C selon la formule suivante

$$C = E_{k_e}(P), \qquad (1.1)$$

oùk_e est la clé de cryptage. La fonction de décryptage, notée D_{K_d} transforme C en P selon la formule suivante :

$$P = D_{K_d}(C), \qquad (1.2)$$

où$K_d$est la clé de décryptage

Le type de relation qui unit les clefs K_e et $K_d$utilisées dans le cryptage et le décryptage permet de définir deux grandes catégories de systèmes cryptographiques (Schneier, janvier 2001) :

- Les systèmes à clé secrète: la clé est un secret partagé entre l'émetteur et le destinataire(K_e= K_d) .
- Les systèmes à clé publique: aucune information secrète n'est partagée entre l'émetteur et le destinataire ($K_e \neq K_d$) .

6) Principe de KERCKHOFF en cryptographie :

Un système cryptographique dont les mécanismes internes sont librement diffusés et qui résiste aux attaques continues de tous les cryptanalystes pourra être considéré comme sûr. Le premier à avoir formalisé ce principe est Auguste Kerckhoffs en 1883 dans l'article 'La cryptographie militaire paru dans le Journal des Sciences Militaires. Son article comporte en réalité six principes, connus depuis, sous le nom de Principes de Kerckhoffs. On en résumera ici que trois, les plus utiles aujourd'hui :

1. La sécurité repose sur le secret de la clé et non sur le secret de l'algorithme. Ce principe est notamment utilisé au niveau des cartes bleues et dans le chiffrement des images et du son sur Canal+ ;
2. Le déchiffrement sans la clé doit être impossible (en temps raisonnable) ;
3. Trouver la clé à partir du clair et du chiffré est impossible (en temps raisonnable).

7) Classification des algorithmes de cryptage :

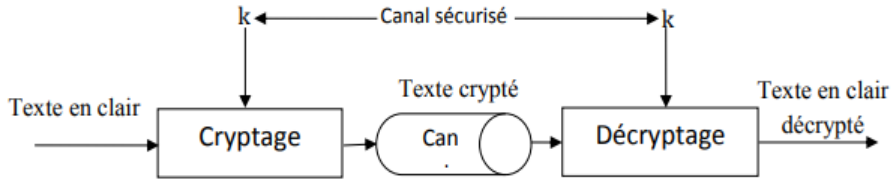
Les algorithmes de cryptage peuvent être classés de différentes manières: selon les clés, selon la structure du cryptage ou selon le domaine de travail **[08]**

7.1 Classification selon la clé de cryptage :

Selon les clés, il existe deux types de chiffrements suivant la relation entre les clés Ke et Kd .

7.1.1 Cryptographie symétrique :

Lorsque ($K_e = K_d =$ k), le chiffrement est appelé un chiffrement à clé privée ou un chiffrement symétrique. Pour les chiffrements par clé privée, la clé de chiffrement / déchiffrement doit être transmise de l'expéditeur au destinataire via un canal sécurisé distinct. Comme illustré à la figure 1.2, en cryptage symétrique, les clés de cryptage et de décryptage sont identiques. Un exemple de cryptage symétrique est le fameux standard de cryptage AES (Advanced Encryption Standard).



.Figure2:Schéma de cryptage symétrique.

7.1.2 Cryptographie asymétrique :

Lorsque ($K_e = K_d$), le chiffrement est appelé un chiffrement à clé publique ou un chiffrement asymétrique. Pour les chiffrements par clé publique, la clé de chiffrement K_e est publiée et la clé de déchiffrement K_d est gardée privée, pour laquelle aucun canal secret supplémentaire n'est nécessaire pour le transfert de la clé. **[08]**

Le RSA fait partie des algorithmes de cryptage asymétriques.

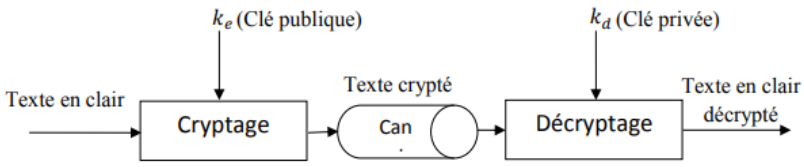


Figure 3:Schéma de cryptage asymétrique

.

7.1.3 Cryptographie classique :

La cryptographie classique décrit la période avant les ordinateurs durant laquelle, les principaux outils utilisés consistent à remplacer des caractères par d'autres et les transposer dans des ordres différents tout en gardant secrètes les procédures de chiffrement ou de déchiffrement. Sans cela le système est complètement inefficace, puisque n'importe qui peut déchiffrer le message codé. On appelle généralement cette classe de méthodes : le chiffrement à usage restreint.

7.1.3.1 Les Transformation s par substituions :

Dans ce mode de cryptage, les lettres du message en clair sont remplacées par d'autres lettres, des chiffres ou d'autres symboles. Selon la façon de substituer, on distingue la substitution mono-alphabétique, la substitution homophonique et la substitution poly alphabétique.

7.1.3.1.1 Les Substitutions Multiples:

7.1.3.1.1.1Chiffrement polygraphique (Chiffre Hill) :

Ce crypto système généralise celui de Vigénére Il été publié par L. S. Hill en 1929.

- On choisit un alphabet de n lettres (on prendra dans nos exemples n = 26) et une taille m pour les blocs, par exemple m = 2.
  Alors P = ε= $(\mathbb{Z}/26\mathbb{Z})^2$, (en général $\mathbb{Z}$ /n $\mathbb{Z}$).
- La clef de codage est une matrice inversible K ∈GLm($\mathbb{Z}$ /26 $\mathbb{Z}$), si m = 2

  K=$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL_2(\mathbb{Z}$ /26 $\mathbb{Z})$

  Si (x1, x2) $\in(\mathbb{Z}/26\mathbb{Z})^2$2 est le message clair alors le message codé sera :

  (y1, y2) = e_K((x1, x2)) = (x1, x2) $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$= (ax1 + cx2, bx1 + dx2)

  La clé de déchiffrage est la matrice inverse de K dan GLm($\mathbb{Z}$ /26 $\mathbb{Z}$).

  Par exemple avec m = 2 et K=$\begin{pmatrix} 11 & 8 \\ 3 & d7 \end{pmatrix}$alors k^{-1} = $\begin{pmatrix} 7 & 8 \\ 23 & 11 \end{pmatrix}$

  Le crypto système de Hill succombe facilement aux attaques à texte clair chois. **[04]**

7.1.3.1.1.2 Chiffrement Poly alphabétique (Chiffre de Vigenère) :

Le Chiffre de Vigenère est un système de chiffrement, élaboré par Blaise de Vigenère (1523-1596), diplomate français du XVIe siècle. C'est un système de substitution poly alphabétique. Cela signifie qu'il permet de remplacer une lettre par une autre qui n'est pas toujours la même, contrairement aux autres chiffres qui se contentaient d'utiliser la même lettre de substitution. C'est donc un système relativement plus « solide ».

1-La table de Vigenère :

L'outil indispensable du chiffrement de Vigenère est : « La table de Vigenère » ou « Carré de Vigenère ». On l'obtient en écrivant 26 fois l'alphabet, et en décalant chaque ligne d'unelettre.

Pour la 1ère ligne : ABCDE ... XYZ

Pour la 2eme : BCDEF... YZA

La 3eme : CDEFG...ZAB, Etc……**[02]**

2-Chiffrement :

Pour chaque lettre en clair, on sélectionne la colonne correspondante et pour une lettre de la clé on sélectionne la ligne adéquate, puis au croisement de la ligne et de la colonne on trouve la lettre codée. La lettre de la clé est à prendre dans l'ordre dans laquelle elle se présente et on répète la clé en boucle autant que nécessaire. **[02]**

3-Exemple

Clé : Musique

Texte : j'adore écouter la radio toute la journée

On répète la clé jusqu'à ce qu'elle soit aussi longue que le texte à chiffrer :

Clé : MUSIQU EMUSIQU EM USIQU EMUSI QU EMUSIQU

Texte : JADORE ECOUTER LA RADIO TOUTE LA JOURNEE

Pour la 1^{ere}lettre:On prend dans la table, la colonne de la clé (M) et la ligne de la lettre (J) :V

Pour la 2^{eme}lettre : On prend la colonne de la clé (U) et la ligne de la lettre (A) : U

Pour la 3^{eme} lettre : On prend la colonne de la clé (S) et la ligne de la lettre (D) : V

Pour la 4^{eme} lettre : On prend la colonne de la clé (I) et la ligne de la lettre (O) : W

Etc

Le texte chiffré est alors :

VUVWHY IOIMBUL PM LSLYI XAOLM BU NAOJVUY.

4- Déchiffrement :

Si on veut déchiffrer ce texte, on regarde pour chaque lettre de la clé répétée la ligne correspondante, et on y cherche la lettre codée. La première lettre de la colonne que l'on trouve ainsi est la lettre décodée.

Texte codé : VUVWHY IOIMBUL PM LSLYI XAOLM BU NAOJVUY

Clé répétée : MUSIQU EMUSIQU EM USIQU EMUSI QU EMUSIQU

|||Ligne I, on cherche W: on trouve la colonne O.

||Ligne S, on cherche V: on trouve la colonne D.

|Ligne U, on cherche U: on trouve la colonne A.

Ligne M, on cherche V: on trouve la colonne J.

7.1.3.1.2 Substitution Unique (Chiffrement Mono alphabétique):

Le chiffrement mono alphabétique ou chiffrement par substitution est une des plus anciennes méthodes de chiffrement. Elle consiste à remplacer chaque lettre d'un texte par un symbole donné (ce symbole peut être une autre lettre de l'alphabet). Sachant que deux lettres distinctes doivent être chiffrées en deux signes distincts pour permettre un déchiffrement du message sans ambiguïté. **[12]**

7.1.3.1.2.1 Alphabet désordonné :

L'alphabet désordonné est la manière la plus classique de chiffrer des messages, il consiste à remplacer une lettre par une autre. Cela donne par exemple la grille de chiffrement ci-dessous : **[12]**

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

7.1.3.1.2 .2 Code de César :

Le chiffre de César (ou code de César ou alphabet décalé) est un type d'alphabet désordonné qui consiste à décaler les lettres de l'alphabet de quelques crans vers la droite ou la gauche. Par exemple, Jules César (qui a donné son nom à ce code) décalait les lettres de 3 rangs vers la gauche : **[12]**

Original	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
Chiffrement	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

L'avantage de ce chiffrement est que la clef de chiffrement est un nombre et qu'il est possible (plus simple) de déchiffrer le message sans avoir de grille de déchiffrement à côté. L'inconvénient majeur est qu'il n'existe que 25 permutations possibles. Il suffit donc d'essayer tous les décalages pour trouver le bon. Cette technique s'appelle la recherche exhaustive des clefs.

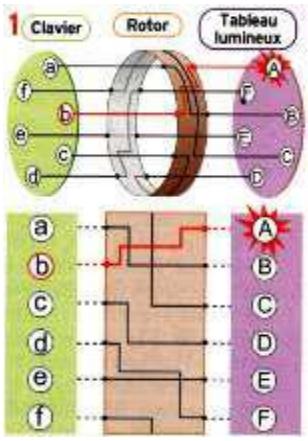
7.1.3.2 Chiffrement par transposition :

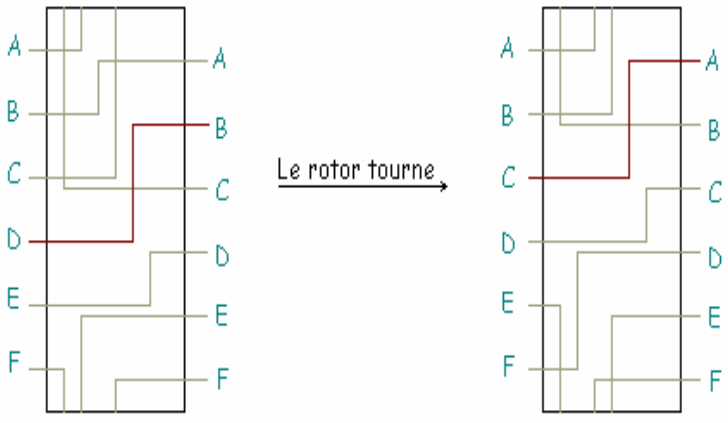
Un chiffrement par transposition (ou chiffrement par permutation) est un chiffrement qui consiste à changer l'ordre des lettres, donc à construire des anagrammes. Cette méthode est connue depuis l'Antiquité, puisque les Spartiates utilisaient déjà une scytale.

Le chiffrement par transposition demande de découper le texte clair en blocs de taille identique. La même permutation est alors utilisée sur chacun des blocs. Le texte doit éventuellement être complété (procédé de bourrage) pour permettre ce découpage. La clef de chiffrement est la permutation elle-même. **[11]**

Machines à rotors :

Très vite après la première guerre, on s'est rendu compte que si l'on souhaitait diffuser beaucoup de documents chiffrés rapidement, et pouvoir changer de clef de chiffrement facilement, il fallait fabriquer des machines à chiffrer et à déchiffrer. Les machines utilisées à ces fins sont les machines à rotors (dont la plus célèbre était la machine ENIGMA à 3 rotors inventée dans les années 30). Cette machine électrique est composée d'un clavier alphabétique, d'un écran lumineux et de trois rotors. Le système est simple : l'utilisateur tape une lettre sur le clavier et le texte chiffré apparaît alors sur l'écran. A chaque frappe sur le clavier, le premier rotor tournait d'une unité puis à la fin d'un tour complet décalait le deuxième rotor d'une unité et ainsi de suite. On positionnait initialement les rotors comme on voulait, ce qui définissait ainsi la clef (FAC par exemple). **[06]**





Grâce à cette machine un texte pouvait être codé d'un million de façons différentes. La frappe au clavier d'une lettre en allumait une autre sur l'écran de manière symétrique (si A donnait C alors C donnait A). Ainsi pour chiffrer un message, une fois la clef fixée, il suffisait de le taper sur la machine et pour le déchiffrer de mettre les rotors dans la même position initiale et de taper le message chiffré. Chaque message commençait par la donnée de la clef choisie par l'opérateur, qu'il cryptait elle aussi selon une liste de clef changeant tous les jours. La machine ENIGMA a été utilisée pendant toute la seconde guerre mondiale par l'armée allemande qui croyait en son inviolabilité. Une équipe de mathématiciens (spécialisée en cryptanalyse, art de déchiffrer des messages) anglais dirigée par A.Turing finit par la décrypter.

On voit donc ici que la cryptographie est une technique de guerre à part entière et qu'elle joue un vrai rôle dans l'ère moderne. En effet on se rend compte qu'en l'espace de quelques années, la cryptographie et la cryptanalyse sont passées de simples techniques désuètes, à véritables sciences. Cette progression des techniques et algorithmes de cryptage ne s'est pas faite toute seule et c'est totalement à cause des attaques incessantes, visant à «casser» les techniques adverses, que l'on a pu assister à un tel bond.

7.2 La cryptographie moderne :

Les techniques quand a vu précédemment n'ont presque plus aujourd'hui qu'une importance historique, les améliorations qu'on leur apporte a permis à d'autres techniques de chiffrement plus sophistiquées de voir le jour, parmi elles, celles qui sont même considérés comme des standards. En fait, nous faisons ici allusion aux crypto systèmes à clé secrète et à clé publique, nous essayerons de les détailler plus dans ce qui suit: **[22]**

7.2.1 Algorithmes de cryptographie symétrique (Chiffrement par blocs):

Les crypto-systèmes symétriques, ou aussi appelés crypto-systèmes à clé secrète, sont donc les crypto-systèmes où les deux opérations de chiffrement et de déchiffrement utilisent la même et unique clé, qu'ils sont utilisés depuis des siècles certes, et même de nos jours vu la simplicité des opérations et donc la rapidité et la facilité d'implémentation mais le seul inconvenant c'est que sur la sécurité de cette clé que repose toute la sécurité du système de chiffrement.

7.2.1.1 DES :

Le 15 mai 1973 le **NBS** (National Bureau of Standards, aujourd'hui appelé NIST - National Institute of Standards and Technology) a lancé un appel dans le FederalRegister (l'équivalent aux Etats-Unix du Journal Officiel en France) pour la création d'un algorithme de chiffrement répondant aux critères suivants :**[16]**

- Posséder un haut niveau de sécurité lié à une clé de petite taille servant au chiffrement et au déchiffrement
- Être compréhensible
- Ne pas dépendre de la confidentialité de l'algorithme
- Être adaptable et économique
- Être efficace et exportable

Principe du DES :

Il s'agit d'un système de chiffrement symétrique par blocs de 64 bits, dont 8 bits (un octet) servent de test de parité (pour vérifier l'intégrité de la clé). Chaque bit de parité de la clé (1 tous les 8 bits) sert à tester un des octets de la clé par parité impaire, c'est-à-dire que chacun des bits de parité est ajusté de façon à avoir un nombre impair de '1' dans l'octet à qui il appartient. La clé possède donc une longueur « utile » de 56 bits, ce qui signifie que seuls 56 bits servent réellement dans l'algorithme.

L'algorithme consiste à effectuer des combinaisons, des substitutions et des permutations entre le texte à chiffrer et la clé, en faisant en sorte que les opérations puissent se faire dans les deux sens (pour le déchiffrement). La combinaison entre substitutions et permutations est appelée code produit.

La clé est codée sur 64 bits et formée de 16 blocs de 4 bits, généralement notés *k1* à *k16*. Etant donné que « seuls » 56 bits servent effectivement à chiffrer, il peut exister 256 (soit 7.2*1016) clés différentes ! **[16]**

L'algorithme du DES :

Les grandes lignes de l'algorithme sont les suivantes :

- Fractionnement du texte en blocs de 64 bits (8 octets) ;
- Permutation initiale des blocs ;

- Découpage des blocs en deux parties : gauche et droite, nommées *G* et *D* ;
- Etapes de permutation et de substitution répétées 16 fois (appelées **rondes**) ;
- Recollement des parties gauche et droite puis permutation initiale inverse.

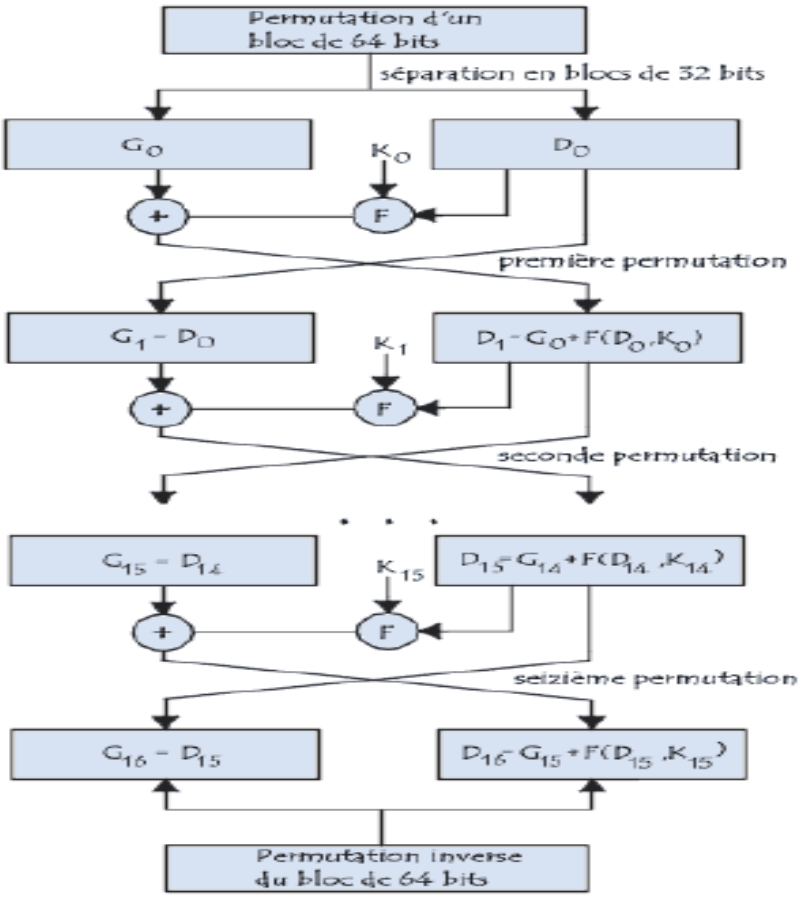


Figure 4:cryptage par DES

7.2.1.2 AES :

En Janvier 1997, la NIST (National Institute of Standards and Technology) lance un appel d'offre international pour remplacer le vieillissant DES : il en résulte 15 propositions.

Parmi ces 15 algorithmes, 5 furent choisis pour une évaluation plus avancée en avril 1999 : MARS, RC6,Rijndael, Serpent et Twofish. Finalement, en octobre 2000 la NIST élit Rijndael comme nouveau standard qu'on nomme aussi AES (Advanced Encryption Standard). Rijndael, du nom condensé de ses concepteurs Rijmen et Daemen, est un algorithme de chiffrement par blocs à plusieurs tours similaires à DES mais avec une taille de blocs et de clefs supérieures et variables, choisis entre 128, 196 et 256 bits. [**06]**

Structure d'état dans l'AES :

On appelle état un bloc vu comme un tableau de 4 x Nb octets où Nb est égal à Taille du bloc / 32. On représente la clef de la même façon, le nombre de colonnes étant Nk = longueur de la clef / 32.**[06]**

$a_{0,0}$	$a_{0,1}$	$a_{0,2}$	$a_{0,3}$
$a_{1,0}$	$a_{1,1}$	$a_{1,2}$	$a_{1,3}$
$a_{2,0}$	$a_{2,1}$	$a_{2,2}$	$a_{2,3}$
$a_{3,0}$	$a_{3,1}$	$a_{3,2}$	$a_{3,3}$

$k_{0,0}$	$k_{0,1}$	$k_{0,2}$	$k_{0,3}$
$k_{1,0}$	$k_{1,1}$	$k_{1,2}$	$k_{1,3}$
$k_{2,0}$	$k_{2,1}$	$k_{2,2}$	$k_{2,3}$
$k_{3,0}$	$k_{3,1}$	$k_{3,2}$	$k_{3,3}$

Exemple d'état (avec des blocs de 128 bits, Nb = 4) et de clef (de longueur 128 bits, Nk = 4)

Nombre de tours :

Le nombre de tours dans l'AES dépend à la fois de la taille des blocs et de la clef. Le nombre r de tours est donné par le tableau :

N_r	$N_b = 4$ (128 bits)	$N_b = 6$ (192 bits)	$N_b = 8$ (256 bits)
$N_k = 4$ (128 bits)	10	12	14
$N_k = 6$ (192 bits)	12	12	14
$N_k = 8$ (256 bits)	14	14	14

Chaque tour utilise une sous-clefki différente et est composée de quatre étapes : ByteSub, ShiftRow, MixColumn et AddRoundKey.**[06]**

ByteSub :

ByteSub est une substitution qui agit isolément sur tous les octets $a_{i,j}$d'un état en 2 étapes :

1. on regarde $a_{i,j}$comme polynôme dans GF(2_8), et on prend son inverse $a^{-1}{}_{i,j}$.
2. on calcule l'image du résultat par la fonction y = f(x) suivante : **[06]**

$$\begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \\ x_7 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}$$

ShiftRow :

ShiftRow effectue un décalage des lignes de l'état courant. La ligne 0 n'est pas décalée, la ligne 1 l'est de **C1** octets, la 2 de **C2** octets et la ligne 3 de **C3** octets. Les valeurs de C1, C2 et C3 dépendant de la taille du bloc, selon la table suivante : **[06]**

N_b	C_1	C_2	C_3
4	1	2	3
6	1	2	3
8	1	3	4

MixColumn :

La transformation MixColumn consiste à prendre chaque colonne de l'état et à la multiplier par la matrice suivante : **[06]**

$$\begin{pmatrix} b_{0,x} \\ b_{1,x} \\ b_{2,x} \\ b_{3,x} \end{pmatrix} = \begin{pmatrix} 02 & 03 & 01 & 01 \\ 01 & 02 & 03 & 01 \\ 01 & 01 & 02 & 03 \\ 03 & 01 & 01 & 02 \end{pmatrix} = \begin{pmatrix} a_{0,x} \\ a_{1,x} \\ a_{2,x} \\ a_{3,x} \end{pmatrix}$$

AddRoundKey :

AddRoundKey consiste en un OU exclusif de l'état courant et de la clef du tour. **[06]**

7.2.2 Cryptographie à clefs publiques :

Principe

Tous les algorithmes évoqués jusqu'à présent sont symétriques en ce sens que la même clef est utilisée pour le chiffrement et le déchiffrement. Le problème essentiel de la cryptographie symétrique est la distribution des clefs : pour que n personnes puissent communiquer de manière confidentielle il faut n (n-1)/2 clefs.

L'idée de base des cryptosystèmes à clefs publiques a été proposée dans un article fondamental de Diffie et Hellman en 1976. Le principe fondamental est d'utiliser des clefs de chiffrement et déchiffrement différentes, non reconstructibles l'une à partir de l'autre :

- Une clef publique pour le chiffrement
- Une clef secrète pour le déchiffrement

Ce système est basé sur une fonction à sens unique, soit une fonction facile à calculer dans un sens mais très difficile à inverser sans la clef privée.

Pour faire une explication imagée, la clef publique joue le rôle d'un cadenas. Imaginons que seul Bob possède la clef (clef secrète), Alice enferme son message dans une boîte à l'aide du cadenas et l'envoie à Bob. Personne n'est en mesure de lire le message puisque seul Bob possède la clef du cadenas.

Le gros avantage de ce système est qu'il n'y ait pas besoin d'avoir partagé un secret au préalable pour s'échanger des messages cryptés. En revanche les implémentations de tels systèmes (RSA, ElGamal,…) ont un inconvénient majeur : leur lenteur par rapport à leurs homologues à clefs secrètes qui tournent eux jusqu'à près de mille fois plus vite. **[06]**

7.2.2.1 RSA :

L'algorithme le plus célèbre d'algorithme à clef publique a été inventé en 1977 par Ron Rivest, Adi Shamir et Len Adleman, à la suite de la publication de l'idée d'une cryptographie à clef publique par Diffie et Hellman. Il fut appelé RSA, des initiales de ces inventeurs. RSA est basé sur la difficulté de factoriser un grand nombre en produit de deux grands facteurs premiers. L'algorithme fonctionne de la manière suivante :
Imaginons que Bob souhaite recevoir d'Alice des messages en utilisant RSA. **[06]**

1- génération des clefs :

a) p et q, deux grands nombres premiers sont générés au hasard grâce à un algorithme de test de primalité probabiliste, avec n = pq.
b) Un nombre entier e premier avec (p-1)(q-1) est choisi. Deux nombres sont premiers entre eux s'ils n'ont pas d'autre facteur commun que 1.
c) L'entier d est l'entier de l'intervalle [2, (p-1)(q-1)[tel que ed soit congrue à 1 modulo (p-1)(q-1), c'est-à-dire tel que ed-1 soit un multiple de (p-1)(q-1).

2- distributions des clefs : le couple (n, e) constitue la clef publique de Bob. Il la rend disponible à Alice en lui envoyant ou en la mettant dans un annuaire. Le couple (n, d) constitue quand à lui sa clef privée.

3- chiffrement du message : Pour crypter le message Alice représente le message sous la forme d'un ou plusieurs entiers M compris entre 0 et n-1. Elle calcule $C= M^e \bmod n$ grâce à la clef publique (n, e) de Bob et envoie C à Bob.

4- déchiffrement du message : Bob reçoit C et calcule grâce à sa clef privée $C^d \bmod n$. Il obtient ainsi le message initial M.

Exemple :
Bob choisit p = 17 et q = 19, n = p x q = 323 et e = 5.
Sa clef privée est alors d=173 car 173 x 5 = 1 (mod (16 x 18))
Supposons qu'Alice veuille lui envoyer le message « BONJOUR » en se servant du tableau suivant pour transformer les lettres en nombre :

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
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26

Cela donne :

B	O	N	J	O	U	R
2	15	14	10	15	21	18

Après avoir chiffré en remplaçant chaque nombre b par (b^e mod n) on obtient :

32	2	29	193	2	89	18

Qu'Alice envoie à Bob.

Bob réalise pour chaque nombre b du message b^d mod n pour trouver :

2	15	14	10	15	21	18
B	O	N	J	O	U	R

qui est bien le message initial.

RSA est basé sur la difficulté de factoriser n. En effet celui qui arrive à factoriser n peut retrouver facilement la clef secrète de Bob connaissant seulement sa clef publique. C'est pourquoi dans la pratique la taille des clefs est au minimum de 512 bits.

Conclusion :

Dans ce chapitre, nous présentons en détail les différentes méthodes de cryptage, ainsi que la méthode de cryptage Vigenère. Nous utiliserons AE pour exécuter le système de cryptanalyse sur ces méthodes, qui sera présenté en détail dans le chapitre suivant.

Introduction

Les phénomènes physiques ou biologiques ont été la source d'inspiration de nombreux algorithmes. Ainsi les réseaux de neurones artificiels s'inspirent du fonctionnement du cerveau humain, l'algorithme de recuit simulé de la thermodynamique, et les algorithmes évolutionnaires (AEs) de l'évolution darwinienne des populations biologiques qui leur permis d'évoluer au cours du temps en créant des systèmes biologiques très complexes adaptés à de nombreuses conditions. Cette dernière branche a été exposée pour la première fois en 1859 par Charles Darwin en publiant « L'Origine des espèces »

Le principe de l'évolution darwinienne repose sur les observations suivantes :

- Il existe au sein de chaque espèce de nombreuses variations, ainsi, chaque individu étant différent ;
- les ressources naturelles étant finies, d'autre part, il naît rapidement plus d'être vivants que la nature ne peut nourrir ; il en résulte une lutte pour l'existence entre chaque organisme ;
- les individus survivants possèdent des caractéristiques qui les rendent plus aptes à survivre. *Darwin* baptise ce concept sélection naturelle ;
- les organismes survivants transmettent leurs avantages à leur descendance qui peuvent être encore meilleurs que leurs parents. L'accumulation au cours des générations des petites différences entre chaque branche généalogique crée de nouvelles espèces de plus en plus aptes à survivre.

Les AEs font partie du champ de l'Intelligence Artificielle (IA), inspirée de « l'intelligence » de la nature. Intelligence que Fogel avaitdéfini de la façonsuivante:
« The capability of a system to adapt its behavior to meet its goals in a range of environments». Ces algorithmes fournissent des solutions aux problèmes difficiles pour lesquels aucune méthode de résolution n'est envisageable et où les solutions sont représentées comme un ensemble de paramètres. **[23]**

Ils constituent une méthode d'exploration automatique d'un espace de recherche potentiellement très vaste. Contrairement à d'autres algorithmes partant d'une solution singulière et cherchant à remonter un gradient de performances, les Aes utilisent un ensemble

de solutions dont seule la performance ponctuelle est utilisée et aucune autre propriété mathématique n'est nécessaire. De ce fait, et par rapport aux méthodes habituelles que l'on peut qualifier d'*analytiques*, les algorithmes évolutionnaires peuvent être qualifiés de *synthétiques* puisqu'ils peuvent parfois synthétiser des solutions nouvelles et originales à des problèmes connus car ils expérimentent sans idées préconçues, si ce n'est les paramètres et l'espace de recherche qu'on leur impose. **[23]**

Par conséquence, le champ des applications potentielles est très vaste à cause de la pauvreté des informations nécessaires et de la généricité des principes exploités. Ainsi, ces algorithmes permettent de résoudre non seulement

- des problèmes purement théoriques en combinatoire,
- en économie,
- en apprentissage,
- dans la théorie des jeux,

Mais aussi des problèmes liés à des applications réelles complexes,

- telles que l'analyse des sondages de sous-sol
- et la détection des champs pétrolifères,
- la fabrication des emplois du temps,
- prévision des cours de la bourse,
- le contrôle des pipe-lines de gaz,
- la conception des automobiles,
- l'optimisation des ailes d'avion,
- les manœuvres des avions de combat,
- les allocations de routes aériennes,
- les allocations dynamiques de fréquences en téléphonie mobile (meilleur résultat actuel),
- le positionnement d'antennes,
- le routage dans les réseaux, …

De plus, ils peuvent aussi être utilisés pour contrôler un système évoluant dans le temps (chaîne de production, centrale nucléaire…) car la population peut s'adapter à des conditions changeantes.

En fait selon les principes Darwiniens, le terme algorithmes évolutionnaires couvre un ensemble de techniques, nommées algorithmes génétiques, stratégies d'évolution, programmation évolutionnaire, et programmation génétique.

1) Historique :

C'est en 1860 que *Charles Darwin* publie son livre intitulé « *L'origine des espèces au moyen de la sélection naturelle ou la lutte pour l'existence dans la nature* », dans lequel ilrejette l'existence «de systèmes naturels figés», adaptés pour toujours à toutes les conditionsextérieures, et il expose sa théorie de l'évolution des espèces: « sous l'influence descontraintes extérieurs, les êtres vivants se sont graduellement adaptés à leur milieu naturel autravers de processus de reproductions ». **[24]**

Presque simultanément, en 1866, *Mendel* publie l'article retraçant dix années d'expériences d'hybridation chez les végétaux en termes de recombinaison de leurs gênes, et l'adresse aux sociétés scientifiques à travers le monde, mais les réactions n'étaient pas du tout encourageant. Ce n'est qu'en 1900, que des résultats similaires à ceux de *Mendel* ont fait l'objet de trois nouveaux articles signés *Hugo de Vries*, *Carl Correns* et *Erich von Tschermark*.

C'est alors à partir du 20ième siècle que la mutation génétique a été mise en évidence par des chercheurs en informatique qui essayent de développer des méthodes permettant aux systèmes d'évoluer de manière normale et efficace face à de nouvelles conditions d'environnement inconnues, variables ou évolutives. Ainsi, les problèmes de traitement de l'information ne seront plus résolus de manières figées, car il ne sera plus indispensable lors de la phase de conception d'un système, d'énumérer toutes les caractéristiques nécessaires pour les conditions d'exploitations connues au moment de la conception.

Dans les années 1960, *John Holland*, ses collègues et ses étudiants ont mené des recherches, à l'université de Michigan, poussés par deux objectifs principaux : premièrement, mettre en évidence et expliquer rigoureusement les processus d'adaptation des systèmes naturels, et deuxièmement, concevoir des systèmes artificiels qui possèdent les propriétés importantes des systèmes naturels. Toutefois, c'était *Bagley*qui a mentionné, en premier lieu, l'expression « Algorithme Génétique ». C'était en 1967. Et en 1975 *Holland* a introduit, dans son livre *« Adaptation in Natural and ArtificialSystems*», le premier modèle formel des algorithmes génétiques : « *the Canonical GeneticAlgorithm CGA* ». **[24]**

2) Les quatre grandes familles des algorithmes évolutionnaires :

2.1 Programmation évolutionnaire :

La programmation évolutionnaire développée par *L.J. Fogel*, se base sur l'évolution d'une population d'automates à états finis (Figure II.1) pour résoudre des problèmes de prédiction. Ce modèle évolutionniste accentue l'utilisation de la mutation et n'utilise pas dans sa version originale la recombinaison des individus par croisement. **[24]**

2.2 Stratégies d'évolution :

Les stratégies d'évolution sont dédiées à l'optimisation des problèmes continus dans l'espace de vecteurs des réels. Les premiers efforts pour la mise en place des stratégies d'évolution ont eu lieu en 1973 à l'université de Berlin au cours de la résolution d'un problème aérodynamique. C'est avec ces méthodes évolutionnistes que la notion d'auto-adaptabilité pour la mutation permettant de contrôler cette fonction de mutation, a été apparue. **[24]**

2.3 Algorithmes génétiques :

Les algorithmes génétiques (AGs) sont probablement les algorithmes les plus connus et les plus utilisés dans le calcul évolutionnaire. Ils ont été développé dans les années soixante par*Holland* qui les a appliqué à l'optimisation paramétrique pour la première fois en 1975, en posant, ainsi, les fondements de cette technique d'application. Cependant, cette technique n'a pas été appliquée sur des problèmes réels de grande taille, à cause des machines calculatoires, de l'époque, et qui n'ont pas été suffisamment puissantes. Ce n'est que vers la fin des années quatre-vingt, précisément, avec l'apparition de l'ouvrage de référence écrit par *Goldberg*, que les algorithmes génétiques ont été connu dans la communauté scientifique. Leur particularité est qu'ils sont fondés sur le *Néo-Darwinisme*, c'est-à-dire l'union de la théorie de l'évolution et de la génétique moderne. Ainsi, les variables sont généralement codées en binaire, par Analogie avec les quatre lettres de l'alphabet génétique d'ADN, sous forme de gènes dans un Chromosome. Ensuite, des opérateurs génétiques, à savoir le croisement et la mutation, sont appliqués à ces chromosomes.

Les AGs sont utilisés pour retrouver une solution résolvant un problème donné, et ce sans avoir de connaissance à priori sur l'espace de recherche. Seul un critère de qualité est nécessaire pour évaluer les différentes solutions en quantifiant, ainsi, leur capacité à résoudre le problème donné. Donc, le but scientifique et technologique visé par ces algorithmes est de pouvoir traiter des problèmes d'optimisation globaux, grâce à la *généralité* avec laquelle on représente l'espace de recherche qui peut contenir des booléens (système actif ou non), des entiers (nombre de composants à optimiser), des réels (intensités associées aux composants réglables), ou des

fonctions discrétisées (optimisation de forme), et grâce aussi à la *robustesse* de la convergence. **[25]**

Il convient de noter, que cette robustesse observée en pratique dans de nombreux domaines applicatifs de l'ingénierie, n'a pas encore selon *Simon Baudot-Roux* une assise théorique très utilisable pour guider les choix algorithmiques, qui sont encore aujourd'hui surtout empiriques. Mais ça n'empêche de noter qu'il existe une école théoricienne qui s'appuie sur les processus stochastiques, mais dont l'impact méthodologique n'est pas encore tout à fait clair.

2.4 Programmation génétique :

L'idée de faire évoluer des programmes date des années cinquante où *Friedberg*, en 1958, a fait plusieurs tentatives pour avoir des ordinateurs auto-programmables en utilisant ce qui est de la mutation actuellement. Donc, et à partir d'une population constituée de programmes aléatoires dont il modifie leurs contenus stochastiquement, il essaye de les améliorer pour aboutir à des résultats satisfaisants. Plus tard, *Smith* (1980) travaillant sur les systèmes classifieurs d'apprentissage, a introduit de petits programmes dans les règles qu'il cherche à les faire évoluer. **[24]**

3) Vocabulaire et Principe :

3.1 Vocabulaire :

Les mécanismes utilisés par les AEs reposent sur le principe de compétition entre lesindividus, où les mieux adaptés aux conditions survivent et peuvent laisser une descendancequi répandra leurs gènes. De ce fait, le vocabulaire employé est directement calqué sur celuide la théorie de l'évolution et de la génétique. Ainsi, il regroupe les termes suivants :

- **Individu :** Un élément de l'espace de recherche. C'est aussi une solution potentielledu problème.
- **Population :** Un ensemble fini (de taille N) d'individus.
- **Génération :** Correspond à l'itération, c.-à-d. repère le moment de l'évolution. Maisparfois ce terme signifie la population en une certaine itération.
- **Evolution :**Un processus itératif de recherche d'un, ou de plusieurs, individu optimal.
- **Performance :**La mesure de la qualité des individus basé sur l'objectif del'optimisation et permettant de comparer les individus entre eux afin d'en déterminerles plus aptes.

- **Evaluation d'un individu :**Le calcul de sa performance.
- **Croisement :** L'opérateur de reproduction appliqué avec une probabilité Pc etcorrespondant à un mélange d'information des individus de la population entre elles. Ilconsiste à échanger des parties composantes (gènes) entre un ou plusieurs individus.
- **Mutation :** L'opérateur de modification d'un ou plusieurs gènes, appliqué avec uneprobabilité P_m dans le but de produire une nouvelle diversité dans la population.
- **Sélection :** Processus du choix des individus pour la reproduction en se basant sur leurperformance.
- **Remplacement :** Processus de formation d'une nouvelle population à partir desensembles de parents et d'enfants choisis suivant leur performanceCe vocabulaire permet de comprendre le principe de fonctionnement des AEs, qui sera détaillé par la suite.

3.2 Principe :

Pour qu'ils puissent surpasser d'autres méthodes plus classiques dans la quête de la robustesse, les AEs doivent être fondamentalement différents. Ils le sont en fait selon quatre axes principaux **[26]**:

1) Les AEs utilisent un codage des éléments de l'espace de recherche et non pas les éléments eux-mêmes.

2) Les AEs recherchent une solution à partir d'une population de points et non pas à partir d'un seul point.

3) Les AEs n'imposent aucune régularité sur la fonction étudiée (continuité, dérivabilité, convexité…). C'est l'un des gros atouts de ces algorithmes.

4) Les AEs ne sont pas déterministes, ils utilisent des règles de transition probabilistes.

La question qui se pose maintenant, est : Comment crée-t-on cette évolution ? Comment faiton évoluer ces gènes ? En fait, la recette est assez simple, et tourne autour des huit points fondamentaux suivants :

1) Tout d'abord, une population initiale d'individus est créé en choisissant des valeurs aléatoires pour les gènes constituant ces individus-là. Le résultat peut contenir des individus de très mauvaises performances.

2) Ces individus sont ensuite évalués. Cela revient à déterminer leurs performances.

3) Maintenant, c'est là que le principe Darwinien intervient. Il s'agit de sélectionner, parmi les individus formant la population, un certain nombre de « géniteurs » qui sont, par exemple, les individus ayant donné les meilleurs résultats lors de l'évaluation.

4) De nouveaux individus, appelés enfants, sont à ce niveau créés en reproduisant les individus sélectionnés, appelés parents, entre eux en imitant la nature, c'est à dire en recombinant leurs gènes (croisement) et/ou en les mutant ; pour obtenir une nouvelle population constituée de la population initiale plus les enfants qui viennent d'être créés,

5) Evaluer les enfants,

6) Supprimez de la population regroupant les parents et les enfants, certains individus pour revenir à la taille de la population initiale. Par exemple, les moins performants des individus sont à supprimer.

7) Définir un critère d'arrêt de l'algorithme. Il peut être un nombre prédéterminé de générations à atteindre, ou la génération d'un enfant satisfaisant les conditions requises. Dans ce cas, l'algorithme s'arrête et donne comme résultat le meilleur individu,

8) Retourner à la phase 3, et recommencer.

Les applications efficaces à base d'AEs sont évidemment plus complexes. Le problème essentiel étant d'adapter, ou même de créer, les opérateurs répondant aux spécificités du problème. De même qu'une recette d'un livre de cuisine nécessite d'être adaptée avec finesse au matériel et ingrédients disponibles, aux goûts des convives, pour être vraiment appréciée**[27]** Ce procédé peut être résumé ou schématisé comme suit :

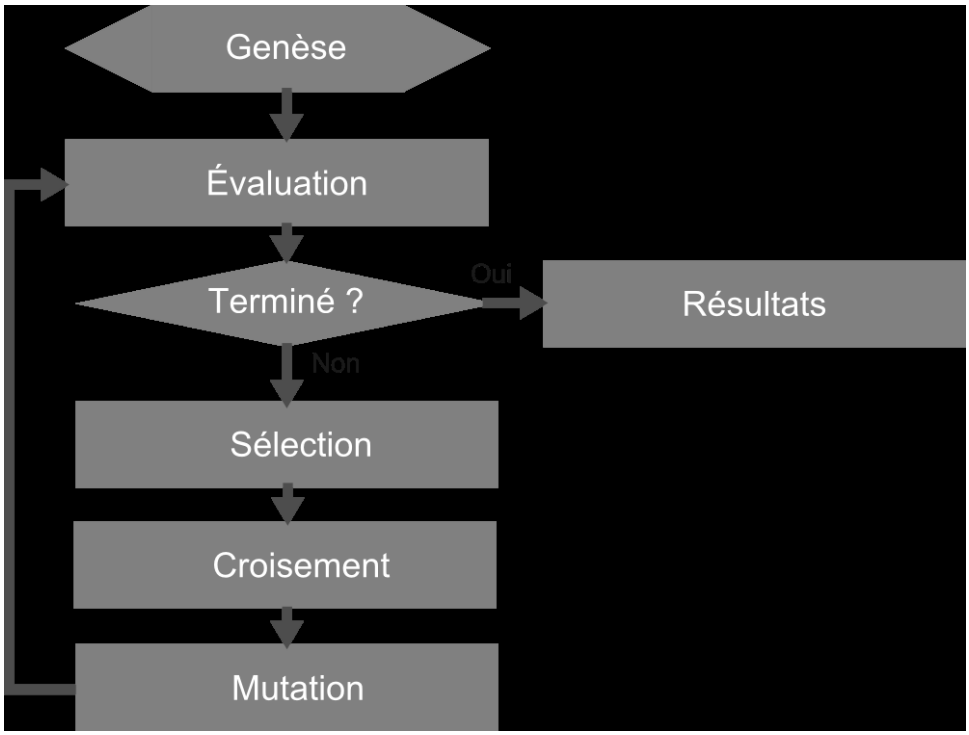


Figure 5:Un algorithme génétique

Détaillons maintenant ses principales phases.

3.2.1 La population initiale :

Comme les AEs agissent sur une population d'individus, et non pas sur un individu isolé, le premier point dans l'implantation des AEs est la création d'une population initiale d'individus. Par analogie avec la biologie, chaque individu de la population est codé par un *chromosome* ou *génotype*. Une population est donc un ensemble de chromosomes, où chaque chromosome code un point de l'espace de recherche**.[23]**

3.2.2 L'évaluation :

L'objectif des AEs est de maximiser une *fonction d'adaptation* (*fitness* en anglais) ou encore dite *fonction d'évaluation*. Intuitivement, cette fonction peut être envisagée comme une mesure de profit, d'utilité ou de qualité que l'on souhaite maximiser. Autrement dit, elle est représentative de l'efficacité des solutions générées sur un problème posé. Elle est le principal biais introduit par le concepteur pour guider l'évolution vers les solutions recherchées. La nature n'est pas aussi directive, c'est pourquoi l'analogie ne doit pas être prise au pied de la lettre **[28]**: l'objectif n'est pas de reproduire, ni même de comprendre le fonctionnement de la

sélection naturelle, le but est de disposer d'une méthode efficace de conception et d'optimisation de structures pour répondre à des problèmes particuliers, mêmesi des interactions sont possibles avec des biologistes.

On définit au préalable la fonction d'évaluation qui associe à chaque phénotype, qui est un vecteur de l'espace de recherche, donc c'est la solution correspondante à un génotype représentant le codage d'un individu ; il lui associe une valeur dite d'*évaluation*. C'est la note de l'individu, plus elle est élevée, plus la solution donnée par ce phénotype est meilleure. Donc, cette fonction sert à calculer le coût d'un point de l'espace de recherche, sachons que l'évaluation d'un individu ne dépend pas de celle des autres individus, et le résultat fournit par la fonction d'évaluation va permettre de sélectionner ou de refuser un individu pour ne garder que les individus ayant le meilleur coût en fonction de la population courante : c'est le rôle de la fonction *fitness*. Autrement dit, cette méthode permet de s'assurer que les individus performants seront conservés, alors que les individus peu adaptés seront progressivement éliminés de la population.

La construction de cette fonction doit être particulièrement optimisée car elle sera exécutée un grand nombre de fois, qui peut augmenter jusqu'à N fois à toutes les générations, ce qui fait que la rapidité de l'algorithme dépend essentiellement d'elle. Elle doit aussi pouvoir tenir compte des phénotypes invalides si le problème doit satisfaire des contraintes que les opérateurs de mutation et croisement ne respectent pas.

3.2.3 La Sélection :

La partie darwinienne de l'algorithme comprend deux étapes totalement indépendantes de l'espace de recherche, qui sont la **sélection** et le **remplacement**. La première, qui est celle de sélection, sert à déterminer les individus d'une population ayant le droit de participer à l'élaboration de la population descendante. Elle ne crée aucune nouveauté, elle se contente de choisir quels individus seront ou ne pas en mesure de contribuer à la création de la population descendante, suivant une stratégie particulière. Cette opération est bien entendu une version artificielle de la sélection naturelle : la survie darwinienne des chaînes les plus adaptées. Sachons que dans les populations naturelles, l'adaptation est déterminée par la capacité d'une créature à survivre aux prédateurs, aux maladies, et aux autres obstacles à franchir pour atteindre l'Age adulte et la période de reproduction**.[26]** Dans notre environnement indéniablement artificiel, une telle décision, qui permet soit de garder en vie ou de tuer un certain individu, revient à la fonction à optimiser.

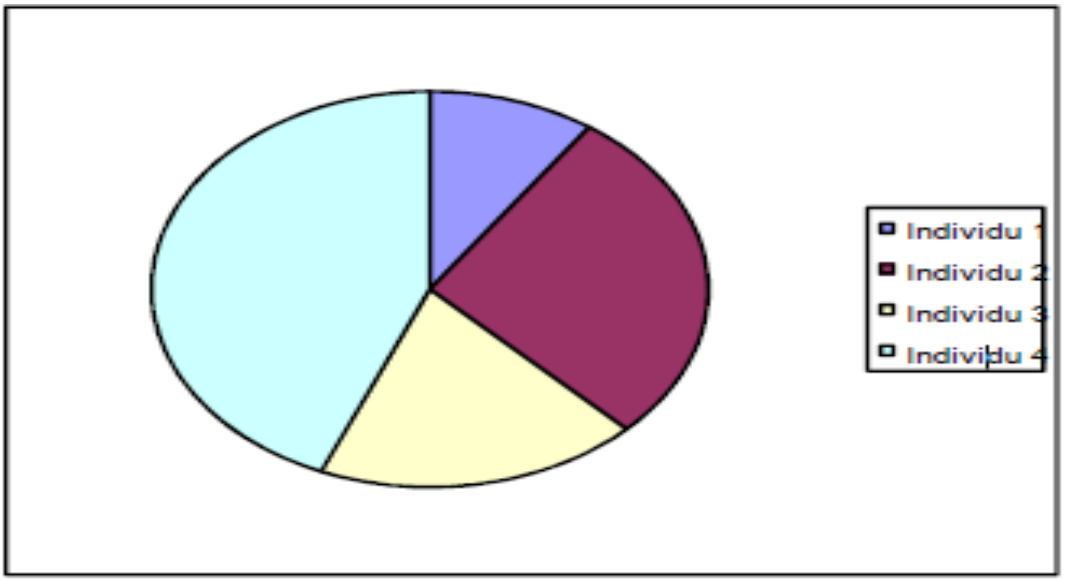
L'algorithme de sélection mis en œuvre doit assurer la convergence vers une solution efficace tout en maintenant la diversité de la population. Ceci est l'aspect caractéristique des AEs par rapport aux autres algorithmes d'apprentissage ou d'optimisation. La diversité dans la population diminue les risques de convergence prématurée vers un extremum local.

3.2.3.1 Les techniques de sélection :

Une panoplie de techniques de sélection est présentée à travers les points suivants :

3.2.3.1.1 La sélection par roulette (Wheel)

Avec cette technique, encore dite **sélection proportionnelle**, les parents sont sélectionnés en fonction de leur performance, où les chromosomes codant les meilleurs résultats, auront plus de chances d'être sélectionnés. Son principe ressemble à celui de la roulette de casino, du fait qu'il revient à imaginer une sorte de roulette de casino sur laquelle sont placés tous les chromosomes de la population suivant leurs valeurs d'adaptation. Un exemple de cette roulette est représenté par la figure.



Ensuite, la bille est lancée et s'arrête sur un chromosome. Cela peut être simulé par la succession des points suivants **[29]**:

1) On calcule la somme S1 de toutes les valeurs d'évaluation des individus d'une population.

2) On génère un nombre r entre 0 et S1.

3) On calcule ensuite une somme S2 des évaluations en s'arrêtant dès que r est dépassé.

4) Le dernier chromosome dont la fonction d'évaluation vient d'être ajoutée est sélectionné.

Ainsi, le meilleur individu a de grandes chances d'être conservé, mais des individus de performance plus faible peuvent également rester d'une génération à l'autre, ce qui peut aider à conserver la diversité de la population. Cependant, cette technique semble inéquitable lorsque les valeurs d'adaptation des chromosomes varient énormément. Si par exemple, la meilleure valeur d'adaptation d'un chromosome représente 90% de la roulette, les autres chromosomes auront très peu de chance d'être sélectionnés et on arrive à une stagnation de l'évolution.

3.2.3.1.2 La sélection par rang :

La sélection par rang trie d'abord la population par fitness. Ensuite, elle associé à chaque chromosome un rang suivant sa position. De ce fait, le plus mauvais chromosome aura le rang 1, le suivant 2, et ainsi de suite jusqu'au meilleur chromosome qui aura le plus grand rang et qui égale à la taille de la population. Le principe de cette méthode de sélection est le même que par roulette, mais les proportions sont en relation avec le rang plutôt qu'avec la valeur d'évaluation.

.

Chromosomes	1	2	3	4	5	Total
Probabilités initiales	24 %	16 %	7 %	2 %	51 %	100 %
Rang	4	3	2	1	5	15
Probabilités finales	27 %	20 %	13 %	7 %	33 %	20 %

Tableau 1:Un exemple de sélection par rang

Dans cet exemple et après l'application de la sélection par rang, c'est l'individu ayant comme rang 3 et comme probabilité finale correspondante 20 % qui est sélectionné.

L'avantage majeur de cette méthode de sélection est que, tous les chromosomes ont une chance d'être sélectionnés. Cependant, elle conduit à une convergence plus lente vers la bonne solution, du fait que les meilleurs chromosomes ne diffèrent pas énormément des plus mauvais.

3.2.3.1.3 La sélection steady-state

Avec cette méthode, une grande partie de la population peut survivre à la prochaine génération. Tout d'abord, un premier ensemble de chromosomes parents choisis parmi ceux qui ont le meilleur coût, sera utilisé pour créer des chromosomes fils. Ces derniers vont remplacer les chromosomes les plus mauvais pour construire, avec les chromosomes parents, la nouvelle population.

3.2.3.1.4 La sélection par tournoi :

La sélection par tournoi consiste à choisir aléatoirement un certain nombre d'individus, parmi lesquels celui qui a la plus grande valeur d'adaptation sera sélectionné. Cette étape est répétée autant de fois qu'il y a d'individus à remplacer dans la génération, sachons que, les individus qui participent à un tournoi restent dans la population et sont de nouveau disponibles pour les tournois ultérieurs.

3.2.3.1.5 L'élitisme :

Cette méthode est utilisée pour éviter, lors de la création d'une nouvelle population, la perte de meilleurs chromosomes suite à l'application des opérations d'hybridation et de mutation. Ainsi, un ou plusieurs des meilleurs chromosomes sont copiés dans la nouvelle génération. Ensuite, le reste de la population sera généré selon l'algorithme de reproduction. Cette méthode améliore considérablement les AEs, car elle permet de ne pas perdre les meilleures solutions.

3.2.3.1.6 La sélection uniforme :

C'est une technique très simple qui consiste à sélectionner un individu aléatoirement de la population. La probabilité P_i pour qu'un individu soit sélectionner est définie par :

P_i = 1 / taille-pop

3.2.4 La génération de nouveaux individus :

Pour assurer la reproduction de nouveaux individus, on fait appel à des opérateurs génétiques.

3.2.4.1 Le croisement (cross-over) :

Le croisement est vu comme l'opérateur d'exploitation essentiel des algorithmes évolutionnaires. L'idée générale du croisement est l'échange de matériel génétique entre les parents. Dans ce cas, si les deux parents sont plus performants que la moyenne, on peut espérer que cela est dû à certaines parties de leur génotype, et que certains des enfants recevant les « bonnes » parties de leurs deux parents, n'en seront que plus performants **[30]**. Tout simplement, cet opérateur consiste à combiner deux génotypes de deux individus pour en obtenir deux nouveaux, en échangeant un ou plusieurs fragments des deux génotypes. On distingue plusieurs croisements possibles, dont les plus utilisés sont :

3.2.4.2 Le croisement en un point:

Pour obtenir les deux nouveaux génotypes, on choisit au hasard un même point de coupure sur ces deux génotypes et on échange les fragments situés après le point de coupure.

3.2.4.3 Le croisement en 2 points:

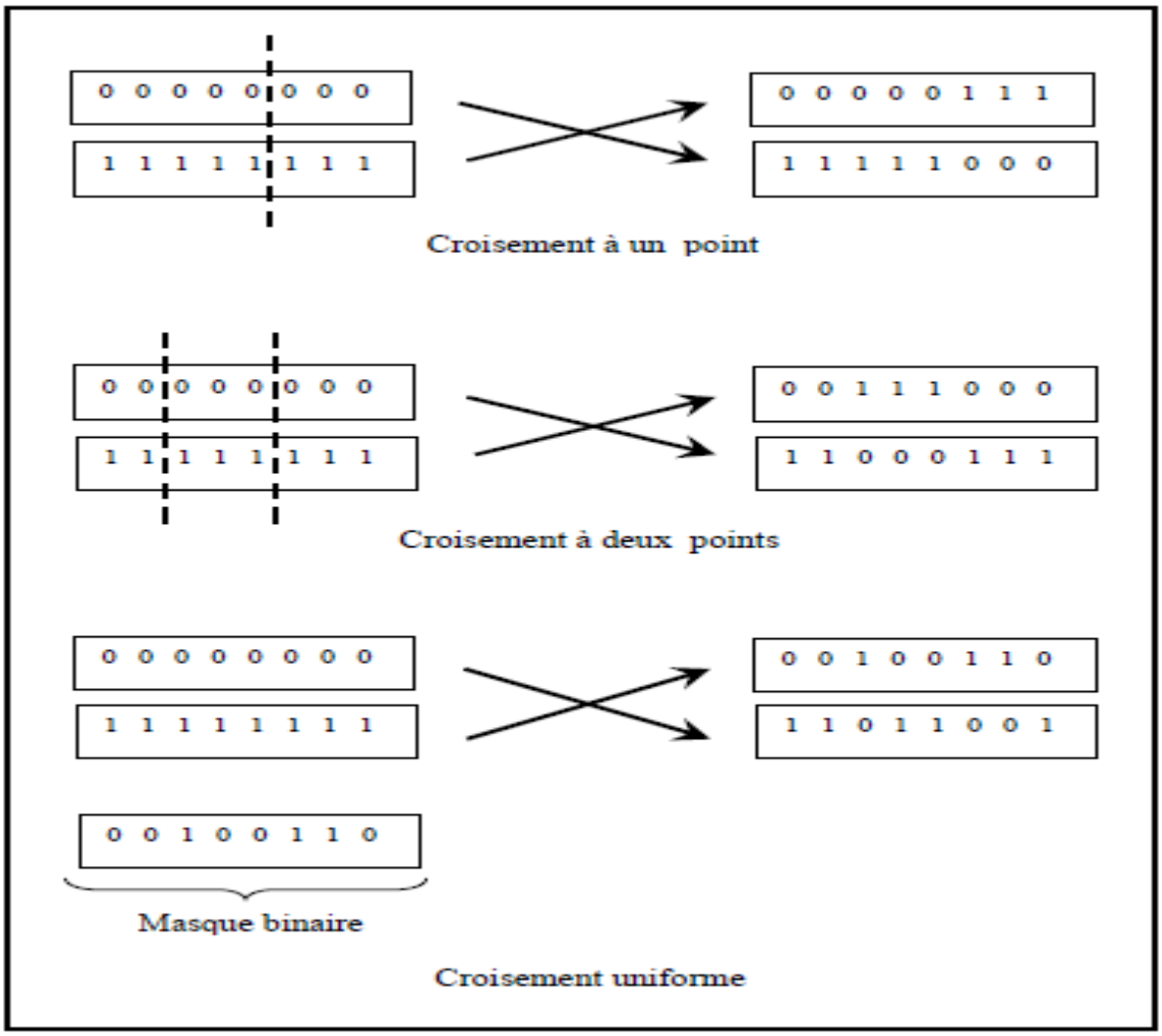
Dans ce cas, on choisit au hasard deux points de croisement et on échange les fragments situés entre ces deux points.

3.2.4.4 Le croisement en k points:

Cette façon d'hybrider est la généralisation à k points de coupure des méthodes précédentes. Elle est aussi connue par le croisement multi points.

3.2.4.5 Le croisement uniforme:

Cette méthode peut être considérée comme un cas particulier du croisement multi points, puisqu'elle consiste à échanger les bits à chaque position indépendamment avec une probabilité de 0.5 ; donc, le nombre de coupures n'est pas connu à priori mais il est déterminé aléatoirement au cours de l'opération. En pratique, pour chaque couple de génotypes, un masque binaire généré aléatoirement et de même longueur que ces génotypes est utilisé. Il permet de conserver les gènes dans les positions correspondantes aux positions contenant des zéros dans le masque. Les autres gènes seront échangés.



3.2.5 La mutation

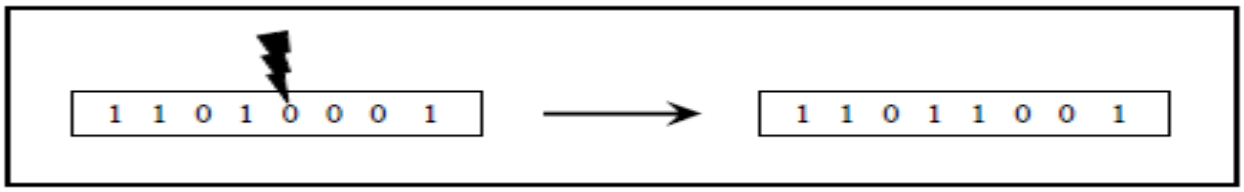
L'operateur de mutation est applique, avec une certaine probabilité Pmut, aux individus issus du croisement. La mutation la plus classique consiste à sélectionner aléatoirement un gène du chromosome d'un individu et a modifier sa valeur.**[30]**

Les mutations les plus utilisées sont :

3.2.5.1 La mutation stochastique (Bit Flip) :

C'est la mutation la plus employée dans la représentation binaire. Elle consiste à inverser chaque bit indépendamment avec une probabilité (c / n), tel que : $0 < c \leq n$ et n est la taille du vecteur.

La mutation 1 bit :Inverse le symbole d'un bit choisi au hasard avec une probabilité Pm.



3.2.6 Le critère d'arrêt :

Le critère d'arrêt peut être arbitraire, par exemple le nombre maximal de générations, ou base sur le critère de convergence décrit ci-dessous.

4. Conclusion :

Les algorithmes génétiques sont des méthodes d'optimisation très utiles dans le cas non linéaire (évidemment, cette méthode fonctionne également pour les cas linéaires, mais dans ce cas, est inutilement puisque elle est lourde en temps et en calcul).

Cette technique part du principe évolutif de la sélection naturelle de Darwin. Celle-ci énonçait que les individus les plus aptes à survivre (les meilleurs) se reproduiront plus souvent et auront plus de descendants. Ainsi, la qualité d'enfant génétique de la population sera augmentée, les gènes plus efficaces deviendront plus fréquent; la population s'améliore. Selon le même principe, un algorithme génétique part d'une population de solutions initiales, les fait se reproduire (les meilleures solutions ont plus de chances de se reproduire), créant ainsi la nouvelle génération de solutions. En répétant ce cycle plusieurs fois, on obtient une population composée de solutions meilleures, l'application de cette technique sur la commande prédictive non linéaire peut porte une solution prometteuse.

Introduction :

Dans ce dernier chapitre, on va voir les différentes étapes suivies durant la réalisation de notre application, Nous appliquerons les algorithmes dérivées de la génétique et des mécanismes d'évolution naturels: croisement, mutation, sélection présenté dans le chapitre 3 pour déchiffrez un texte chiffré par la méthode Vigenère.

1) Les outils utilisés dans l'implémentation :

Dans cette section on va présenter les différents outils utilisés pour la réalisation des cryptanalyse de chiffrement Vigenère.

1.1 Langage JAVA :

Java est un langage de programmation et une plate-forme informatique qui ont été créés par Sun Microsystems en 1995. Beaucoup d'applications et de sites Web ne fonctionnent pas si Java n'est pas installé et leur nombre ne cesse de croître chaque jour.

1.2 Eclipse:

Eclipse est un environnement de développement intégré dont le but est de fournir une plate-forme modulaire pour permettre de réaliser des développements informatiques. Eclipse, de par le fort soutien de la communauté open source et à la base crée sur une initiative d'IBM, s'est imposé comme un environnement de développement multi-langages de premier ordre. Sa capacité à être étendu à travers des plugins en fait un outil de choix lors d'un développement demandant l'utilisation et l'intégration de composants logiciels hétérogènes.

1.3 WordNet :

Est une base de données lexicale développée par des linguistes du laboratoire des sciences cognitives de l'université de Princeton depuis une vingtaine d'année. Son but est de répertorier, classifier et mettre en relation de diverses manières le contenu sémantique et lexical de la langue anglaise. Des versions de WordNet pour d'autres langues existent, mais la version anglaise est cependant la plus complète à ce jour. La base de données ainsi que des outils sont disponibles pour téléchargement gratuit. Par rapport aux outils fournis, un développeur peut aussi accéder la base de données à partir des interfaces disponibles pour de nombreux langages de programmation. WordNet est distribué sous une licence libre, permettant de l'utiliser commercialement ou à des fins de recherche. La dernière version distribuée en avril 2013 est la 2.1. lnk.

2) L'implémentation :

L'implémentation des algorithmes génétiques passe par ces étapes suivantes :

2.1 La population initiale :

Il s'agit d'une population de base générée aléatoirement, elle génère des textes de la même taille que le texte chiffré

2.2 L'évaluation :

C'est l'étape de la comparaison des individus, grâce à laquelle nous pourrons évaluer les textes les plus appropriés lors de la sélection.

2.3 La sélection :

Seuls les meilleurs individus (textes) sont conservés. Dans nos applications, les meilleurs textes sont sélectionnés en fonction du seuil de sélection.

2.4 L'application des opérateurs génétiques:

Cette phase se compose de deux parties :

2.4.1 Le croisement :

Cela comprend la combinaison de deux individus (parents) pour construire deux autres individus (enfants). Par conséquent, nous avons coupé deux individus (parents) à un moment donné (les deux parents sont dans la même position) et avons échangé les pièces entre ces points.

2.4.1.1 Le taux de croisement :

- L'opérateur croisement est appliqué avec la probabilité Pc.
- Plus le ratio est élevé, plus de nouveauxindividus sont introduites.
- En général, Pc varie entre 0.25 et 0.70.

2.4.2 La mutation :

Consiste à modifier un gène sélectionné dans un chromosome et à leremplacer par une valeur aléatoire.

2.4.2.1 Le taux de mutation:

- L'opérateur de mutation est appliqué avec une probabilité Pm.
- En général, Pm varie entre 0.01 et 0.1.
- • Si ce taux est élevé, la recherche deviendra purement aléatoire.
- S'il est petit, la population est moins diversifiée et il y a un risque de stagnation.

3) L'interface :

Notre application se compose de trois étapes principales

1. Le chiffrement d'un texte par la méthode de Vigenère
2. L'application de l'algorithme génétique pour le décryptage
3. Donne les résultats les plus approchés au texte clair.

   Dans les figures qui viennent ci-dessous on va illustrer notre application avec la description de chaque onglet :

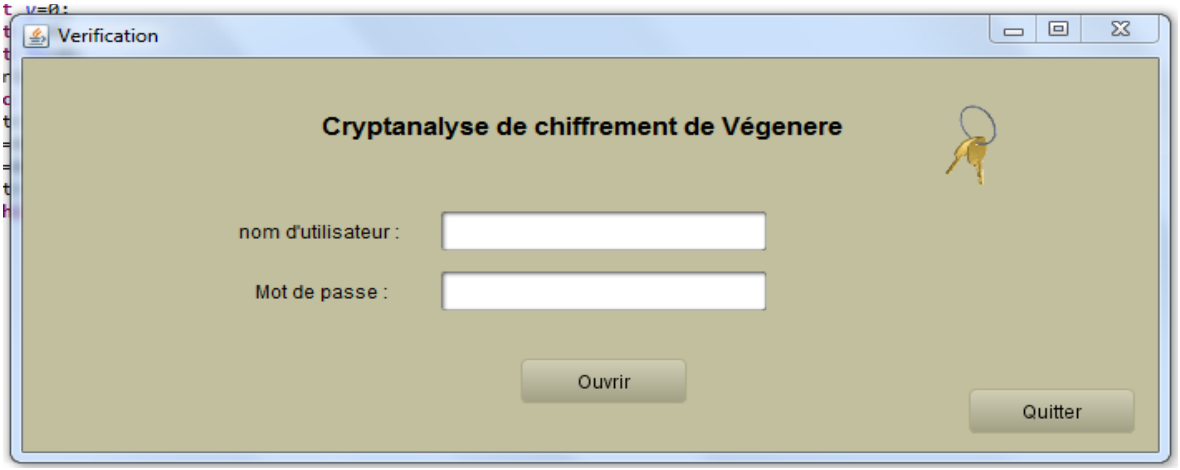


Figure6 : Interface d'entrée.

Cet onglet a pour rôle d'identification de l'utilisateur.

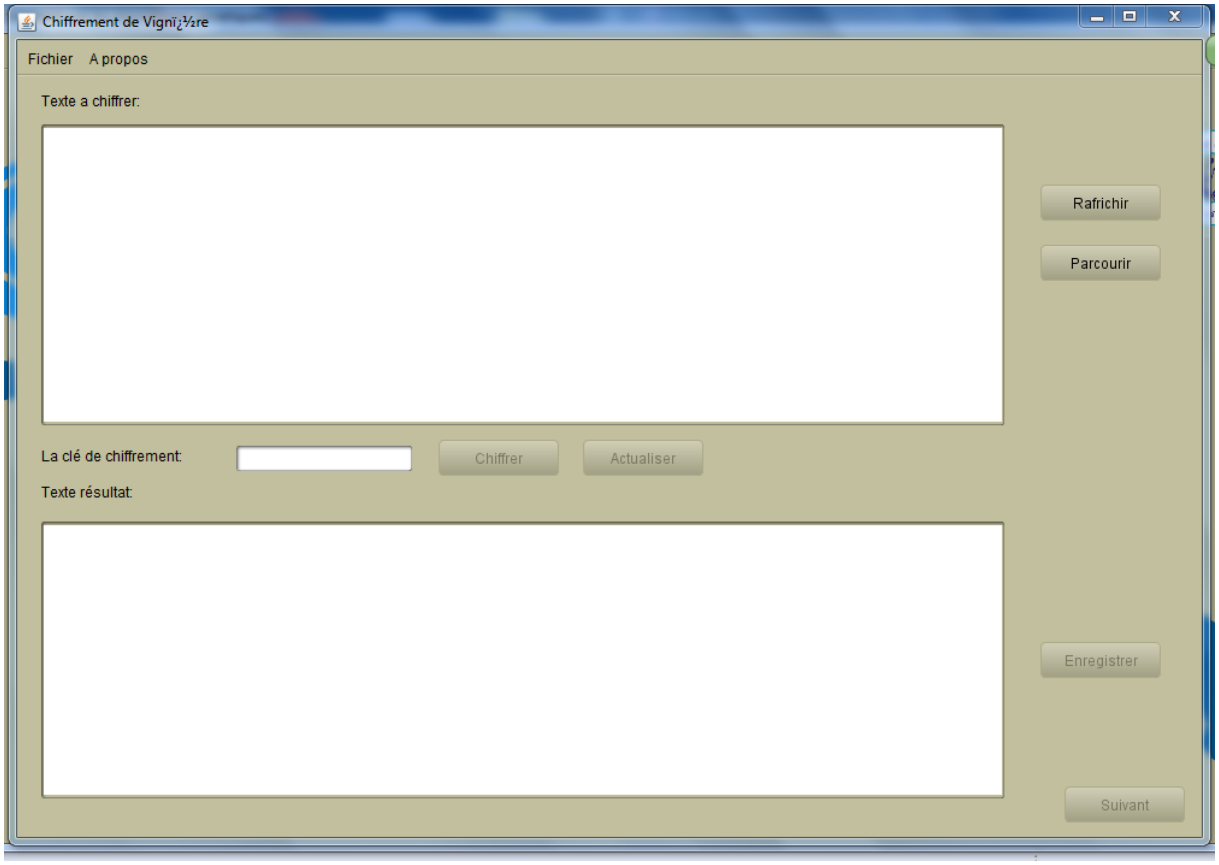


Figure 7: Chiffrement de Vigenère

- Le bouton **Parcourir**pour ajouter un texte.
- Le JTextArea **en haut** :pour afficher le texte ajouté
- Le JTextField **La clé de chiffrement** le champ où on va mettre la clé de chiffrement.
- Le JTextArea **en bas** : le champ où il va apparaitre le texte chiffré.
- Le bouton **Chiffrer** fait le chiffrement de texte puis son apparition dans le JTextArea.
- Le bouton **Enregistrer** fait l'enregistrement le texte crypté
- Le bouton **Initialiser** fait l'initialisation de tous les paramètres.
- Le bouton **Suivant** le passage à l'étape suivante qu'on va l'expliquer.

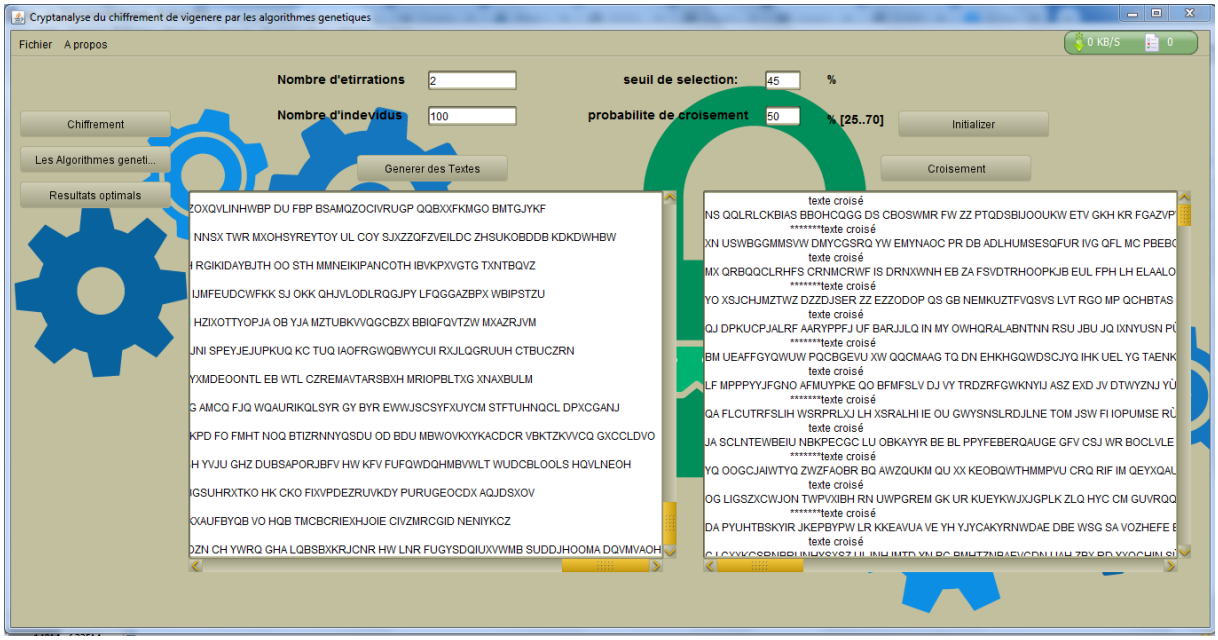


Figure 8: L'utilisation des AG pour le déchiffrement

Cet onglet (Fig 8) permet de traiter le texte crypté pour le but de déchiffrer le texte à l'aide de l'algorithme génétique. Alors il est nécessaire d'introduire les paramètres suivants :

- ☐ Nombre d'individus (le nombre des textes à générer)
- ☐ Nombre d'itérations.
- ☐ Seuil de sélection (le nombre des mots significatifs sur le nombre de tous les♣ mots d'un texte).
- ☐ Probabilité de croisement.
- ☐ Probabilité de mutation.

➢ Le bouton **Générer** les textes fait la génération des textes d'une façon aléatoire de taille similaire au texte chiffré et les afficher dans le JTextArea à gauche.

➢ Le bouton **Croisement** fait la combinaison entre deux textes quelconques (parents) pour construire deux autres textes (enfants).

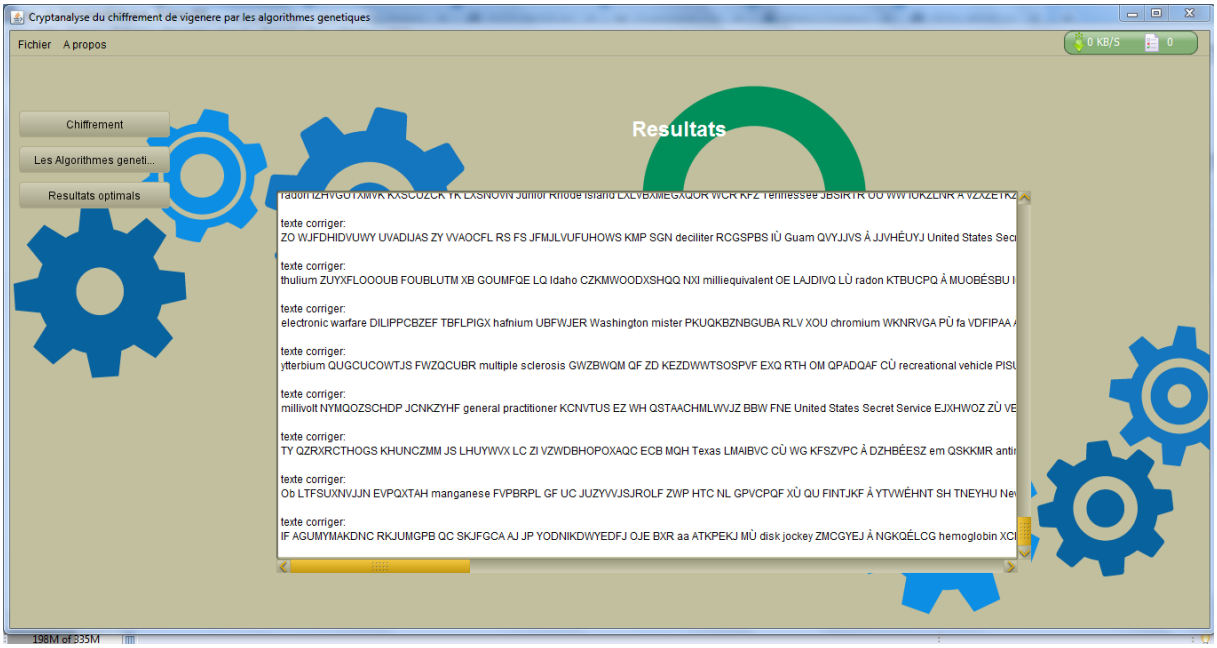


Figure 9: Résultat optimal : «Affichage du texte décrypté optimal par rapport au texte clair»

4) Les tests et résultats de l'exécution de l'application:

<u>Exemples</u>:

- ☐ Le chiffrement d'un texte par la méthode de Vigenère

Texte **: je vais manger du pain**

Clé : nabil

Le texte chiffré : WE WITF MBVRRR EC ANIO

- ☐ L'application de l'algorithme génétique pour le décryptage

a) Teste_01

- Fixer le nombre d'itération à 50 est varier le seuil de sélection de 10 à 50 cela nous a donné le Tableau ci-dessus :

Population initiale (Texte)	Nombre d'itération	Seuil de Sélection (%)	Texte Trouvés (Sur 500)	Significatif	Non-Significatif	Temps (minute)
500	50	10	328	0	328	15m.22s
500	50	30	25	0	25	27m.03s
500	50	40	8	0	8	48m.09s
500	50	50	0	0	0	1h.3m

Tableau 2:Nombre de textes selectionés lorsque nombre d'itération est égale à 50

Remarque : Seuil de Sélection =50 → Texte Trouvés=0

b) Teste_02.

- Fixer le nombre d'itération à 200 et varier le seuil de sélection de 10 à 50 cela nous a donné le Tableau ci-dessus :

Population initiale (Texte)	Nombre d'itération	Seuil de Sélection (%)	Texte Trouvé (Sur1000)	Significatif	Non-Significatif	Tempe (minute)
500	200	10	303	0	303	1h.15m
500	200	30	175	0	175	1h.38m
500	200	40	60	0	60	1h.55m
500	200	50	1	0	1	2h.3m

Tableau 3: Nombre de textes sélectionnés lorsque Le nombre d'itération est égal à 200

Remarque : Seuil de Sélection =50 → Texte Trouvé=1 ; Mais ce dernier n'était pas significatif.

c) Teste_03.

- Fixer le nombre d'itération à 300 est varier le seuil de sélection de 10 à 50 cela nous a donné le Tableau ci-dessus :

Population initiale (Texte)	Nombre d'itération	Seuil de Sélection (%)	Texte Trouvés (Sur 500)	Significatif	Non-Significatif	Tempe (minute)
500	300	10	380	0	380	2h.06m
500	300	30	250	0	250	2h.19m
500	300	40	106	0	106	2h.25m
500	300	50	69	5	69	3h

Tableau 4:Nombre de texte sélectionnés lorsque Le nombre d'itération est égal à 300

Remarque : Seuil de Sélection =50→ Texte Trouvés=69 → Texte significatif=5.

d) Teste 04.

- Fixer le nombre d'itération à 500 est varier le seuil de sélection de 10 à 50 cela nous a donné le Tableau ci-dessus :

Population initiale (Texte)	Nombre d'itération	Seuil de Sélection (%)	Texte Trouvés (Sur 1000)	Significatif	Non-Significatif	Tempe (minute)
1000	500	10	726	0	726	4h.06m
1000	500	30	504	0	504	4h.19m
1000	500	40	226	0	226	4h.25m
1000	500	50	69	5	69	5h

Tableau 5:Nombre de texte sélectionnés lorsque le nombre d'itération est égal à 500

Remarque : Seuil de Sélection =50→ Texte Trouvés=69 → Texte significatif=5.

Conclusions :

Selon les résultats trouvés par les différents tests, nous remarquons que lorsque le nombre de seuil de la population est grand, la possibilité de trouver un texte claire est élevée. Ce qui montre que notre application donne une possibilité de trouver un texte claire avec un taux égal à 50%.

Dans ce chapitre nous avons trouvé que l'utilisation de l'algorithme génétique pour cassée le chiffrement de Vigenère c'est une méthode efficace et ça ne prend pas de temps par rapport à d'autres méthodes, nous avons démontré ceci par implémenté quelques exemples. Nous espérons dans le futur, développer cet algorithme pour obtenir de meilleurs résultats.

Dans ce cas, nous sommes en mesure de décrypter le système de cryptage Vigenère. Notre outil de base est l'algorithme évolutif.

Le premier obstacle à ce travail est la formalisation du problème de décryptage afin de le restituer au problème d'optimisation combinatoire.

Le deuxième obstacle à résoudre est d'établir les éléments de base des algorithmes évolutifs, à savoir: le codage chromosomique, la définition des fonctions d'évaluation et la sélection des opérateurs génétiques. Enfin, la génération de texte grâce à des algorithmes évolutifs est définitive.

Par conséquent, Les domaines d'application des algorithmes évolutionnistes s'étendent de plus en plus et l'avenir est à eux. Mais en dépit de leur simplicité, les algorithmes évolutionnistes ne sont pas évidents dans leur application. En fait, concevoir et réaliser un bon algorithme évolutionniste demande une bonne connaissance du problème, une bonne compréhension des mécanismes évolutionnistes et beaucoup de créativité

Nous espérant dans le futur développer un peu plus cet algorithme pour obtenir des résultats beaucoup plus meilleurs que celle trouvées par notre application en poussant d’avantage les calculs en jouant sur les paramètres et méthodes utilisées.

Bibliographie

1. (Julien,2010). *Sécurité Informatique et Piratage.*
2. Pdf.initiation à la cryptologie Mars 2009.
3. Récupéré sur https://www.memoireonline.com/11/19/11308/m_etude-pour-la-securisation-dun-reseau-par-la-mise-en-place-dun-pare-feu-open-source-cas-de27.html.
4. Barsky, D. (f´evrier 2006). *Cours de Cryptographie (version pr´eliminaire 2005/2006)* .
5. Récupéré sur https://www.futura-sciences.com/tech/definitions/internet-spvwar-1956.
6. BLANC, J. (2003). *technique de cyptographie.*
7. Récupéré sur https://www.simplilearn.com/cryptography-a-detailied-insight-rar217-article.
8. El-Samie, A. (2013.). *Image encryption : a communication perspective.*
9. laila, s. (2015). *Attaques Informatique.*
10. Récupéré sur https://www.futura-sciences.com/tech/definitions/internet-cookies-469/
11. Récupéré sur https://fr.wikipedia.org/wiki/Chiffrement_par_transposition
12. Récupéré sur https://tmonseigne.github.io/Chiffre_Mono/
13. Didier. Godart, 2002. Sécurité informatique risques stratégies et solutions. 2 Edition. ENI.
14. Benoit 2016, article sur le par-feu.
15. Récupéré sur https://tpesecuriteinformatique.wordpress.com/les-differentes-attaques-informatiques/
16. Récupéré sur https://web.maths.unsw.edu.au/~lafaye/CCM/crypto/des.htm
17. Récupéré sur https://web.maths.unsw.edu.au/~lafaye/CCM/virus/worms.htm
18. https://www.syloe.com/glossaire/authentification/Alexandre Viardin, 2004. Un petit guide pour la sécurité.
19. Récupéré sur https://www.institut-numerique.org/iii4-definitions-525681a39aab7
20. http://laurent.flaum.free.fr/pgpintrofr.htm
21. Récupéré sur https://www.securiteinfo.com/attaques/divers/social.shtml
22. Rima, D. (2009). *CRYPTOGRAPHIE : APPROCHE QUANTIQUE.*

23. Vincent Magnin, « Optimisation et algorithmes génétiques », Cyber-cour gratuit de l'EUDIL, 2001.

24. Pierre Collet, Jean-Philippe Rennard, « Handbook of Research on Nature Inspired Computing for Economics and Management: Stochastic Optimization Algorithms »,Edition Rennard, J.P, Hershey, IGR, 2006.

25. Simon Baudot-Roux, «OPALE : OPtimisation et contrôle, ALgorithmes numériques et intégration de systèmes complexes multidisciplinaires régis par des E.d.p », Projet commun CNRS-INRIA-UNSA bilocalisé Sophia-Antipolis / Rhône-Alpes.

26. David E. Goldberg, « Algorithmes génétiques : exploration, optimisation et apprentissage automatique », Editions Addison-Wesley France, SA.

27. Evelyne Lutton, « Darwinisme artificiel : une vue d'ensemble», Revue Technique et Science Informatique (TSI), numéro spécial "Méthodologie de la gestion intelligente des senseurs", 2005.

28. Stéphane Doncieux, « Algorithmes évolutionnistes: de l'optimisation de paramètres à la conception complète d'un système de contrôle», Journées MicroDrones, Toulouse,2002.

29. Yves Coueque, Julien Ohler, Tollari Sabrina, « Algorithmes génétiques pour résoudre le problème du commis voyageur », Avril 2002.

30. Éric Goubault, Frédéric Nataf, Marc Schoenauer, « calcul parallèle : algorithmes évolutionnaires », École Polytechnique.