REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE
MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET LA RECHERCHE SCIENTIFIQUE

UNIVERSITE MENTOURI DE CONSTANTINE
FACULTE DES SCIENCES DE L'INGENIEUR

DEPARTEMENT D'INFORMATIQUE
LABORATOIRE LIRE

N° 144/ Mag/ 2006
Série 010 / Inf /2006

MEMOIRE
EN VUE DE L'OBTENTION DU DIPLOME DE MAGISTER EN INFORMATIQUE

OPTION : INFORMATION & COMPUTATION

Thème

Proposition d'un système immunitaire artificiel pour la détection d'intrusions

Présenté par :
M^me^ LABED Ines

Encadré par :
M^r^ M-K Kholladi

Devant le jury composé de :

Pr M.Batouche	**Président**	**Université de Constantine**
Dr M-K Kholladi	**Rapporteur**	**Université de Constantine**
Dr S.Chikhi	**Examinateur**	**Université de Constantine**
Dr N.Zarour	**Examinateur**	**Université de Constantine**

Année universitaire 2005-2006

DÉDICACES

Je dédie ce présent mémoire :

A celle qui s'est toujours dévouée et s'est sacrifiée pour moi, celle qui m'a aidé du mieux qu'elle pouvait pour réussir, celle qui a toujours été là dans mes moments de détresse, ma très chère mère.

A mon père qui m'a toujours encouragé et soutenu moralement.

A mes très chères sœurs Malika, Rayane, mes très chers frères Mohamed et Idris qui m'ont énormément aidé et à qui je témoigne mon affection.

A mon grand père et ma grande mère

A ma belle mère et mon beau père ainsi mes beaux frères et mes belles sœurs.

A celle qui m'a toujours aidé, écouté, soutenu et encouragé tout au long de mon parcours, celle qui est toujours à mes côtés, ma très chère tante Souad.

A mes oncles et mes tantes.

A celui qui a su me supporter, m'aider et m'encourager, mon très cher mari Karim, en témoignage de sa gentillesse et de son affection.

Sans oublier mon trésor, mon fils Mohamed Akram.

Ines

REMERCIEMENTS

Ma reconnaissance va à tous ceux qui m'ont aidé à conduire ce travail à son terme : tout particulièrement, j'aimerais remercier vivement, mon encadreur de thèse, Monsieur Kholladi Mohamed Khireddine, de l'attention et du soutien qu'il a porté à mon travail.

Ainsi je tiens à remercier vivement l'enseignante Madame Meshoul Sihem qui m'a toujours encouragé et soutenu depuis le début de ma thèse, qui m'a donné le goût de la recherche et n'a cessé de m'encourager.

Je remercie tous les membres du jury le Professeur Mohamed Batouche et le Docteur Chikhi Salim ainsi que le Docteur Zarour Nacereddine de me faire l'honneur d'assister à ma soutenance.

Résumé

Le challenge dans le domaine de la sécurité informatique est de pouvoir déterminer la différence entre un fonctionnement normal et un fonctionnement avec intrus. Durant, la dernière décennie, la protection des systèmes s'est faite moyennant des règles de codage qui identifient et bloquent des événements spécifiques. Cependant, la complexité accrue des systèmes et des réseaux qui sont devenus de plus en plus larges ainsi que la nature des intrusions courantes et futures nous incite à développer des outils de défense automatiques et surtout adaptatifs. Une solution prometteuse est d'utiliser les systèmes immunitaires artificiels qui s'inspirent des systèmes immunitaires humains lesquels sont dotés de capacités de détection et de défense d'intrus.

Dans ce travail, nous avons exploité certains modèles immunitaires artificiels pour la détection d'intrusions. Ainsi, nous avons intégré quelques concepts supplémentaires proposés par la nouvelle théorie de danger, afin de surmonter les problèmes liés à l'adoption du modèle de soi et de non soi. Le système proposé permet la détection des intrusions réelles qui sont générées, non seulement par les éléments externes du système, mais aussi par ses membres internes. Cela est garantit par l'incorporation de la notion de danger, qui est le principe de base de la théorie de danger, permettant de s'apercevoir et de faire remarquer les éléments de soi mais nuisibles et les éléments de non soi mais inoffensifs.

Mots clés : sécurité informatique, théorie de danger, systèmes immunitaires artificiels, détection d'intrusions.

Abstract

The central challenge with computer security is determining the difference between normal and potentially harmful activity. During the last decade, developers have protected their systems by coding rules that identify and bloc specific events. However, the increased complexity of the systems and network, which became ever larger in conjunction with the nature of current and future threats require the development of automated and adaptive tools. A promising solution is to use the artificial immune systems (AIS) inspired from the human immune system, which can detect and defend against harmful and previously unseen invaders.

In this work, we use a number of patterns of artificial immune systems for intrusion detection. In addition, we integrate some complementary components, proposed by the new danger theory to overcome self-nonself discrimination problems.

The proposed system allows the detection of the real threats, which are generating, not only by the external elements of the system, but also by its internal members. This is the result of the incorporation of danger, which is the basic concept of the danger theory, which allows to take care of nonself but harmless and the self but harmful invaders into the system.

Keywords: computer security, artificial immune systems, danger theory, intrusion detection.

ملخص

يكمن التحدي في ميدان حماية شبكة المعلومات في إمكانية التفريق بين العمل العادي و العمل الدخيل ; من أجل حماية الأنظمة استعملت خلال السنوات الأخيرة مجموعة من قواعد التشفير وهذا لمعرفة و إيقاف الأحداث الخاصة. لكن لازدياد الصعوبات في الأنظمة و الشبكات التي أصبحت أكثر اتساعا من جهة و طبيعة التدخلات الحالية و المستقبلية من جهة أخرى ,حتم علينا إنشاء مجموعة من الأدوات الأوتوماتيكية ,من بين الحلول الموجودة, حل مهم يتمثل في استعمال أنظمة المناعة الاصطناعية المقتبسة من أنظمة المناعة عند الإنسان التي لديها إمكانيات الاكتشاف ثم الدفاع ضد الدخلاء.

في هذا العمل، قمنا باستعمال مجموعة من نماذج نظام المناعة الاصطناعية من أجل اكتشاف الدخلاء و أيضا عملنا على دمج مجموعة من الوظائف المقترحة من خلال نظرية جديدة (نظرية الخطر) من أجل التغلب على المشاكل الناتجة عن اعتماد نموذج الذات و اللاذات.

النظام المقترح يسمح باكتشاف التدخلات الحقيقية التي تنتج من طرف العناصر الخارجية و العناصر الداخلية للنظام. كل هذا يتحقق من خلال إدماج مفهوم الخطر الذي يعتبر القاعدة الأساسية في نظرية الخطر. هذا الأخير يسمح بملاحظة و استنتاج عناصر الذات الضارة و عناصر اللاذات المسالمة الغير ضارة.

كلمات مفتاحية: حماية شبكة المعلومات، نظرية الخطر، نظام المناعة الاصطناعي، اكتشاف التدخلات.

.

TABLE DES MATIERES

TABLE DES FIGURES

CHAPITRE I

INTRODUCTION GENERALE

I.1 INTRODUCTION

Les systèmes informatiques et les réseaux sont devenus des outils indispensables pour la société actuelle. Ils sont aujourd'hui déployés dans tous les secteurs professionnels. Initialement, isolés les uns des autres, ces systèmes informatiques sont devenus interconnectés et le nombre de points d'accès ne cesse de croître. Ce développement phénoménal est accompagné également par la croissance du nombre d'utilisateurs, qui ne sont pas forcément pleins de bonnes intentions vis-à-vis de ces systèmes informatiques. Ils peuvent exploiter les vulnérabilités des réseaux et les systèmes pour essayer d'accéder à des informations sensibles dans le but de les lire, les modifier ou les détruire, portant atteinte au bon fonctionnement du système.

I.2 PROBLEMATIQUE

L'importance de sécurité des systèmes informatiques motive les angles divers de la recherche dont le but principal est de fournir de nouvelles solutions prometteuses qui ne pourraient être assurées par des méthodes classiques. Les systèmes de détection d'intrusions sont l'une de ces solutions qui permettent la détection des utilisations non autorisées, les mauvaises utilisations et les abus dans un système informatique par les utilisateurs externes ainsi que ses utilisateurs internes. Le défi dans le domaine de la sécurité informatique et plus précisément dans les systèmes de détection d'intrusions est de pouvoir déterminer la différence entre un fonctionnement normal et un fonctionnement avec intrus. Cependant, les systèmes et les

réseaux à protéger sont devenus de plus en plus complexes et larges ainsi que la nature des intrusions courantes et futures nous incite à développer des outils de défense automatiques et surtout adaptatifs. Une solution prometteuse est d'utiliser les systèmes immunitaires artificiels qui s'inspirent des systèmes immunitaires humains, lesquels sont dotés de capacités de détection et de défense d'intrus. Plusieurs travaux ont été proposés pour la détection d'intrusions qui sont basés sur les systèmes immunitaires artificiels, et qui ont intégré différents modèles immunitaires dont le modèle principal est le modèle de soi et de non soi.

L'objectif principal de ces systèmes consiste à augmenter le taux de vrai positif c'est-à-dire la détection des intrusions réelles et à minimiser le taux de vrai négatif qui reflète le taux d'erreurs du système. Vu que les intrusions sont générées non seulement par les membres externes mais aussi par ses membres internes, alors il est nécessaire d'améliorer les systèmes de détection d'intrusions qui sont basés sur le modèle de soi et non soi afin de permettre la détection des éléments nuisibles et dangereux qui peuvent être de soi ou de non soi ce qui permet l'augmentation du taux vrai positif et la minimisation du taux vrai négatif. Ainsi, pour la réalisation de ce but, certains systèmes de détection d'intrusions proposés, exigent l'intervention continue de l'opérateur de sécurité après chaque détection dont le but principal est l'obtention d'un ensemble de détecteurs permettant la détection des intrusions réelles.

I.3 OBJECTIF DU TRAVAIL

L'objectif de ce travail est de présenter, d'une part, le lien qui peut exister entre les systèmes immunitaires et la sécurité des réseaux et de proposer d'autre part, de nouvelles techniques basées sur les systèmes immunitaires artificiels.

Dans ce travail, nous avons utilisé quelques modèles des systèmes immunitaires artificiels ainsi nous avons intégré quelques nouveaux concepts, en se basant principalement sur la nouvelle théorie proposée dans l'immunologie qui est la ***théorie de danger,*** et qui propose des nouveaux horizons à exploiter afin de surmonter les problèmes liés à l'utilisation du modèle de soi et de non soi. Cette nouvelle théorie signale qu'il existe plusieurs facteurs contribuant dans l'initiation de la réponse immunitaire de telle sorte que ce n'est plus le caractère étranger d'un élément qui déclenchera la réponse immunitaire, mais plutôt le caractère dangereux d'un élément. Dans le système proposé, nous essayons d'assurer la détection des éléments dangereux qui peuvent être des éléments internes dont le système est

initialement tolérant ainsi que la détection des éléments externes du système qui ont initialisé des dommages dans l'environnement. Ainsi, nous essayons d'éviter l'étape de costimulation qui nécessite l'intervention continue d'un officier de sécurité.

I.4 ORGANISATION DU MEMOIRE

L'organisation de ce document reflète la démarche que nous avons adoptée lors de la réalisation de ce travail. Ce travail est composé de six chapitres

Après l'introduction générale, le deuxième chapitre décrit les systèmes de détection d'intrusions (IDS), avec une classification générale des systèmes de détection d'intrusions selon plusieurs critères. Ainsi qu'une étude comparative entre certains types d'IDS sera présentée. Puisque cette étude se focalise sur l'approche comportementale, les différentes approches utilisées par cette approche seront exposées avec une vue d'ensemble sur les différentes architectures d'implémentation d'un IDS. La dernière section de ce chapitre sera consacrée à l'exposition de quelques systèmes de détection d'intrusions existants.

Le troisième chapitre décrit les systèmes immunitaires artificiels (AIS). Afin de comprendre les différents algorithmes proposés dans le domaine des systèmes immunitaires artificiels, ce chapitre commence par une présentation générale du système immunitaire naturel et les différents mécanismes utilisés dans l'identification et la détection des intrus. La partie consacrée aux systèmes immunitaires artificiels expose les différents algorithmes immunitaires disponibles ainsi que les différents domaines d'application des systèmes immunitaires artificiels. Avec l'apparition de la théorie de danger qui défie l'immunologie classique et en particulier le modèle de soi et de non soi, une discussion est consacrée sur la proposition d'extension des systèmes immunitaires artificiels par l'intégration de nouveaux concepts proposés par cette théorie.

Le quatrième chapitre montre l'analogie entre l'objectif des systèmes de détection d'intrusions et celui du système immunitaire humain. Cette démonstration se base sur la présentation des exigences principales d'un système de détection d'intrusions compétent ainsi que les buts de conception d'un IDS nécessaires pour la satisfaction de ces exigences. Enfin,

une analyse prudente des propriétés des systèmes immunitaires humains qui peuvent contribuer à la réalisation de ces buts de conception d'un IDS.

Dans le cinquième chapitre, nous aborderons les différents travaux qui ont exploité les systèmes immunitaires artificiels dans le domaine de la détection d'intrusions. Une étude discutée, nous permet la présentation de l'approche proposée, qui tente d'utiliser plus de composants immunitaires afin de surmonter quelques problèmes liés à l'adoption du modèle de soi et de non soi. Le pseudo code de l'algorithme sera détaillé avec les différents mécanismes immunitaires mis en œuvre, avec la description des différents paramètres utilisés et la présentation de l'ensemble de données employé dans la phase test. Enfin, un certain nombre d'expériences sera faite sur l'algorithme.

Nous terminerons ce mémoire par le chapitre présentant nos conclusions ainsi que les perspectives de ce travail.

CHAPITRE II
LE SYSTEME DE DETECTION D'INTRUSIONS

II.1 INTRODUCTION

Avec le développement des réseaux de communication, l'Internet est devenu l'infrastructure critique pour une société moderne. La croissance explosive d'utilisateurs d'Internet a motivé l'expansion rapide de commerce électronique et d'autres services en ligne. Malheureusement derrière la convenance et l'efficacité de ces services, les risques et les chances d'intrusions malveillantes sont aussi augmentés. La sécurité des systèmes informatiques est devenue un défi majeur dont l'objectif est d'assurer la disponibilité des services, la confidentialité et l'intégrité des données et des échanges.

De nombreux mécanismes ont été développés pour assurer la sécurité des systèmes informatiques, particulièrement pour *prévenir les intrusions* dont le but principal est de construire un système sécurisé en déterminant et éliminant les vulnérabilités de sécurité. Citons par exemple l'authentification qui consiste à prouver l'identité des utilisateurs, le contrôle d'accès qui consiste à définir les droits d'accès accordés aux utilisateurs sur les données et les pare-feux qui filtrent l'accès aux services du système informatique vis-à-vis de l'extérieur.

Cependant, ces mécanismes ne peuvent pas garantir la sécurité des systèmes informatiques. En effet, ces systèmes présentent des failles de conception, d'implémentation et de configuration permettant à des attaquants de contourner les mécanismes de prévention. De plus, un certain nombre de ces systèmes se protègent contre des attaques externes, alors plusieurs études ont montré qu'il existe des attaques issues par leurs utilisateurs internes.

L'augmentation de l'importance de sécurité d'ordinateurs mène à des recherches diverses afin de fournir de nouvelles solutions qui ne pourraient pas être réalisables par des approches de sécurité classiques. Les systèmes de détection d'intrusions sont l'une de ces solutions qui permettent de garantir la sécurité.

Pour cette raison, ce chapitre sera consacré à représenter cette nouvelle solution destinée à repérer des activités anormales ou suspectes sur la cible analysée qui peut être un hôte ou un réseau par exemple. Ce chapitre sera organisé comme suit : nous commencerons d'abords par la description d'un système de détection d'intrusions (IDS), ainsi nous présentons les critères nécessaires pour assurer l'efficacité de tel système. La section suivante sera consacrée à la classification des systèmes de détection d'intrusions selon plusieurs critères. Puis, nous étudierons quelques types d'IDS d'une manière détaillée vu de leur importance dans le reste de cette étude. Finalement, nous exposerons d'une manière générale les architectures d'implémentation des systèmes de détection d'intrusions ainsi que la présentation de quelques systèmes de détection d'intrusions existants.

II.2 LE SYSTEME DE DETECTION D'INTRUSIONS

La recherche dans le domaine de système de détection d'intrusions a été commencée en 1980 par le travail de James Anderson [1] qui a souhaité l'amélioration des équipements d'audit et les capacités de surveillance des systèmes informatiques. Suivant ce travail, le premier modèle de détection d'intrusions générique a été proposé par Dorothy Denning en 1987 [2]. Ce modèle de détection d'intrusions générique est indépendant de tout système et environnement d'application, les types d'intrusions et les vulnérabilités de système [2]. Il se compose de six éléments :

- ***Les sujets*** : ce sont les initiateurs des actions effectuées sur le système qui peuvent être : les utilisateurs, les groupes d'utilisateurs, ou encore les processus générés par ces utilisateurs.
- ***Les objets*** : ce sont les ressources gérées par le système tels que les fichiers, les programmes, les messages, les périphériques, etc.
- ***Les enregistrements d'audit*** : ils représentent les actions entreprises par un sujet sur un objet et sont générés par le système. Ces enregistrements sont composés de :
  - Le sujet ayant fait l'action.
  - Le type de l'action (par exemple login, logout, lecture, écriture).

- Les ressources utilisées.
- Le résultat de l'action (échec ou succès).
- Des éléments quantitatifs (par exemple, nombre de lignes ou pages imprimées, nombre d'enregistrements lus ou écrits).
- La date et l'heure où s'est déroulée l'action.

- ***Les profils*** : ce sont des structures qui caractérisent le comportement des sujets envers des objets.
- ***Les enregistrements d'anomalie*** : ils sont générés lorsqu'une activité anormale est détectée.
- ***les règles d'activités*** : elles spécifient les actions à entreprendre lorsque certaines conditions sont satisfaites sur les enregistrements d'audit ou les enregistrements d'anomalies générés.

Dans ce modèle, un profil est défini par un ensemble de *variables* et de *modèles statistiques*. Ces variables représentent des mesures quantitatives accumulées durant une période de temps, qui peut être soit un intervalle de temps fixe (heure, jour, etc.) ou entre deux événements d'audits (login et logout, connexion et déconnexion, etc.). Ce modèle de base est un système expert en temps réel. Il stocke les faits des profils décrivant le comportement normal afin de générer les signatures de comportements incorrects. Quand un sujet agit sur un objet spécifique alors l'événement généré change la valeur des variables. Une base de connaissances contient les règles d'activité dont le rôle de ces règles consiste à mettre à jour les profils, détecter les comportements anormaux, produire des rapports ou encore alerter l'officier de sécurité. Le moteur d'inférence essaye de déclencher les règles qui correspondent aux faits des profils. En ce qui concerne les modèles statistiques proposés par Denning, ils seront décrits dans les sections suivantes.

II.2.1 Définition d'un système de détection d'intrusions

Il existe plusieurs définitions du système de détection d'intrusions. Cependant, nous allons seulement citer quelques unes.

II.2.1.1 Définition 1

Heady et d'autres [4] ont défini un système de détection d'intrusions comme suit :

Définition 1 : Intrusion

Une intrusion est n'importe quel ensemble des actions qui essayent de compromettre l'intégrité, la confidentialité ou la disponibilité d'une ressource d'ordinateur.

Définition 2 : Détection d'intrusions

C'est le problème d'identification des actions qui essayent de compromettre l'intégrité, la confidentialité ou la disponibilité d'une ressource d'ordinateur.

Définition 3 : Système de détection d'intrusions

C'est un système d'ordinateur (une combinaison de logiciel et de matériel) qui essaye d'exécuter la détection d'intrusions.

II.2.1.2 Définition 2

Tandis que Kim [43] a défini un système de détection d'intrusions comme un système automatisé dont le rôle est la détection des intrusions dans un système informatique tout en examinant les audits de sécurité fournis par le système d'exploitation ou bien les outils de contrôle du réseau. Son but principal est la détection des utilisations non autorisées, les mauvaises utilisations et les abus dans un système informatique par les utilisateurs internes et externes.

II.2.2 Terminologie des attaques

Une attaque peut être définie comme toute action ou ensemble d'action qui peut porter atteinte à la sécurité des informations d'un système ou d'un réseau informatique.

II.2.2.1 Classification des attaques

Vu le nombre important des attaques possibles, elles peuvent être classées selon différentes classifications [5] :

II.2.2.1.1La première classification

- ***Les attaques passives :*** ce type d'attaque vise à l'obtention d'accès pour pénétrer dans le système sans compromettre ces ressources.
- ***Les attaques actives :*** dont le résultat de cette attaque est un changement non autorisé d'état des ressources de système.

II.2.2.1.2 La deuxième classification

- ***Les attaques internes*** : ce type d'attaque est causé :
  1. Soit par les utilisateurs autorisés du système qui essayent d'utiliser des privilèges complémentaires dont ils n'ont pas le droit.
  2. Soit par les utilisateurs autorisés qui emploient improprement les privilèges dont ils ont le droit.
- ***Les attaques externes* :** ce type d'attaque est causé par des utilisateurs externes qui essayent d'accéder à des informations ou des ressources d'une manière illégitime et non autorisée.

II.2.2.1.3 La troisième classification

Selon cette classification, les attaques de cette catégorie peuvent porter atteinte à :

- *La confidentialité* des informations en brisant les règles privées.
- *L'intégrité* en altérant les données.
- *La disponibilité* en rendant un système ou un réseau informatique indisponible. Ces attaques sont connues sous le nom des attaques de déni de service.
- *L'authenticité* des informations.

II.2.2.2 Description des attaques

Il existe différentes attaques que les systèmes de détection d'intrusions essayent de les repérées. Parmi ces attaques, nous citons par exemple [5] :

- ***Craquage des mots de passe***.
- ***Cheval de Troie « Trojan Horse »*** qui est un programme qui se cache lui-même dans un autre programme au-dessus de tout soupçon. Quand la victime lance ce programme, elle lance aussi le cheval de Troie caché. Un cheval de Troie peut par exemple, lorsqu'il est exécuté, ouvrir l'accès au système à des personnes particulières ou même à tout le monde.
- ***IP spoofing*** : dans ce cas l'attaquant change son adresse IP par une autre adresse de confiance afin d'obtenir des droits d'accès.
- ***Les scans*** : qui servent principalement à obtenir des informations sur un hôte, un réseau (pour préparer une attaque plus élaborée). Les informations qu'un attaquant peut obtenir sur le réseau sont par exemple : le type de système d'exploitation de la machine, les ports ouverts, etc.

- ***Sniffing:*** elle se caractérise par l'observation et l'analyse du trafic réseau afin d'obtenir des informations pertinentes pour les attaques suivantes.
- ***Les attaques de déni de service (denial of service)*** ont pour but de paralyser le serveur cible pour qu'il devienne inaccessible, au moins pour une durée de temps. De très nombreuses techniques existent pour épuiser les ressources d'un hôte cible, par exemple : ICMP Flooding, smurf, SYN flood, etc.
- Etc.

II.2.3 Concepts de base

Nous désirons dans cette section d'éclairer quelques notions qui seront utilisées dans le reste de ce travail.

- **Système** : dénote un système d'information contrôlé par un système de détection d'intrusions. Cela peut être un poste de travail, un élément du réseau, une unité centrale, un pare-feu, un serveur Web, un réseau d'entreprise, etc.
- **Alarme** *:* c'est la réponse générée par le système de détection d'intrusions lors de la détection d'une intrusion. Cependant les erreurs de détection peuvent être classées selon deux types :
  - ***Le positif faux :*** signifie qu'un système de détection d'intrusions détecte une intrusion là où aucune intrusion réelle n'a été commise.
  - ***Le négatif faux :*** A l'inverse de «positif faux », «négatif faux » signifie que le système de détection d'intrusions n'a pas détecté une intrusion ayant réussi.

II.2.4 Description du système de détection d'intrusions

Un système de détection d'intrusions à un niveau très macroscopique [6] peut être décrit comme un *détecteur*. Ce détecteur est un moteur d'analyse qui reçoit des données de trois sortes de ressources (Figure 2.1). L'analyse de ces données génère une décision d'évaluation de la probabilité que ces actions peuvent être considérées comme des symptômes d'intrusions. Ces données sont :

- Des informations de configuration relatives à l'état actuel du système.
- Des informations à long terme relatives à la technique utilisée pour détecter les intrusions par exemple une base de connaissances d'attaques.

Ø Des informations venant du système à protéger qui sont les informations d'audit décrivant les événements qui apparaissent dans le système.

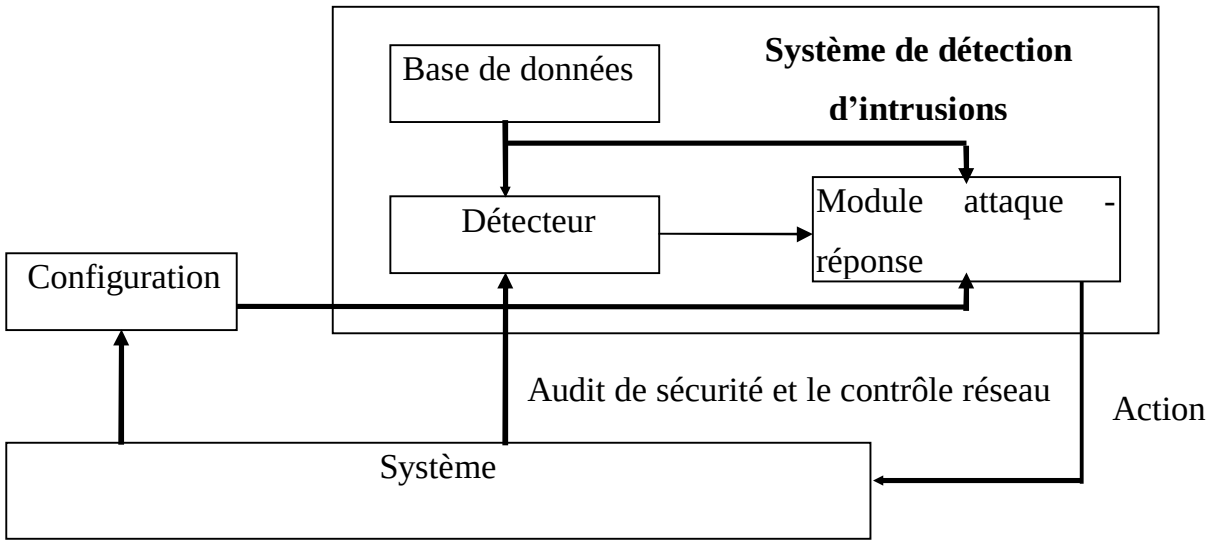


Figure 2.1 : Description d'un système de détection d'intrusions.

II.3 EFFICACITE DES SYSTEMES DE DETECTION D'INTRUSIONS

L'efficacité d'un système de détection d'intrusions est déterminée par les mesures suivantes [6] :

Ø **Exactitude** : Le système de détection d'intrusions n'est pas exact s'il considère les actions légitimes des utilisateurs comme atypiques ou intrusives.

Ø **Performance :** La performance d'un système de détection d'intrusions est mesurée par le taux de traitement des traces d'audits. Si la performance du système de détection d'intrusions est pauvre, donc la détection en temps réel n'est pas possible.

Ø **Perfection :** Un système de détection d'intrusions est imparfait s'il n'arrive pas à détecter une attaque.

Ø **Tolérance aux pannes :** Un système de détection d'intrusions doit être résistant aux attaques, en particulier dans le cas des attaques de déni de service.

Ø **Opportunité :** Un système de détection d'intrusions doit exécuter et propager son analyse d'une manière prompte pour permettre une réaction rapide dans le cas d'existence d'une attaque.

II.4 CLASSIFICATION DES SYSTEMES DE DETECTION D'INTRUSIONS

II.4.1 Une vue d'ensemble du système IDS

Les différents systèmes de détection d'intrusions disponibles peuvent être classés [6,8] selon plusieurs critères (Figure 2.2), qui sont :

- La méthode de détection.
- Le comportement du système après la détection.
- La source des données.
- La fréquence d'utilisation.

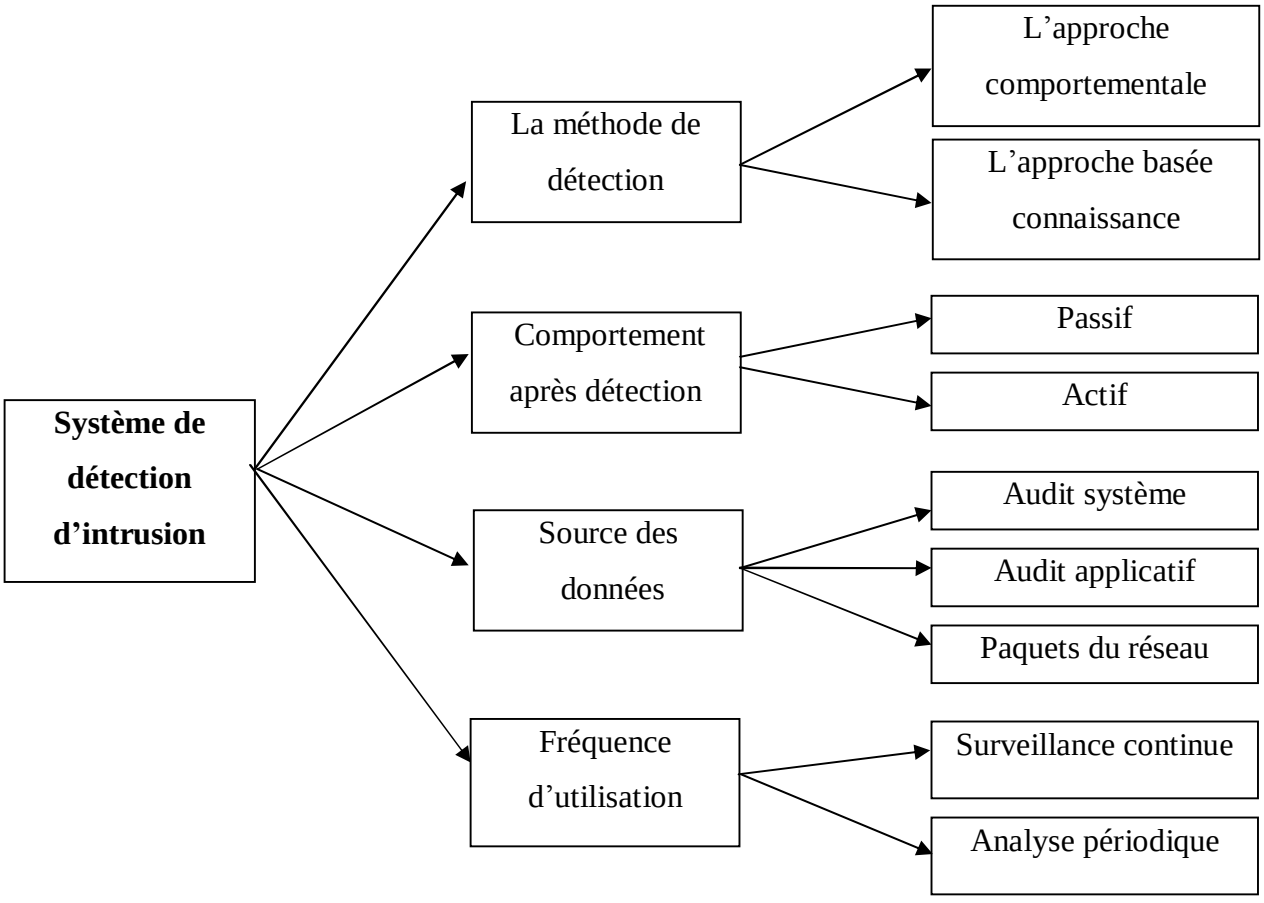


Figure 2.2 : Classification d'un système de détection d'intrusions

II.4.2 La méthode de détection

La détection d'intrusions repose sur deux approches de base :

- L'approche comportementale.
- L'approche basée connaissance.

II.4.2.1 L'approche comportementale

Cette approche est connue aussi par *l'**approche de détection d'anomalies.*** Elle consiste à définir un profil de l'activité normale d'un utilisateur et à considérer les déviations significatives de l'activité d'utilisateur courante par rapport aux profils de comportement normaux comme anomalie.

II.4.2.2 L'approche basée connaissance

Cette approche définit des signatures soupçonneuses basées sur les vulnérabilités connues de système et la politique de sécurité. Une intrusion est signalée lorsque la trace d'une attaque connue est présente dans les traces d'audit.

Ces deux méthodes d'analyse constituent la partie importante des systèmes de détection d'intrusions. Pour cette raison, elles seront détaillées dans les sections suivantes.

II.4.3 Le comportement après la détection d'intrusions

Le comportement d'un IDS après la détection d'une intrusion est l'ensemble des actions prises par le système lorsqu'il détecte une attaque. Ces réponses peuvent être *actives* ou bien *passives.*

II.4.3.1 Réponse active

La réponse active implique des actions automatisées prises par un IDS quand le système détecte une intrusion. Par exemple interrompre le progrès d'une attaque pour bloquer ensuite l'accès suivant de l'attaquant.

II.4.3.2 Réponse passive

Dans ce cas, quand une attaque est détectée, le système de détection d'intrusions ne prend aucune action. Il génère seulement une alarme pour notifier l'administrateur de système qui va prendre des mesures en se basant sur les rapports générés par le système de détection d'intrusions.

II.4.4 La nature des données analysées

Les systèmes de détection d'intrusions sont classés en fonction de l'origine des données qui seront exploitées pour détecter des actions intrusives. La source de données utilisée est une

caractéristique essentielle pour classer les systèmes de détection d'intrusions. On distingue trois catégories de sources d'informations :

- Ø Les audits systèmes.
- Ø Les audits applicatifs.
- Ø Le trafic réseau.

II.4.4.1 Les audits systèmes

Les audits systèmes sont produits par le système d'exploitation d'un hôte. Ces données permettent à un IDS de contrôler les activités d'un utilisateur sur un hôte. Elles peuvent être également de plusieurs types, par exemple :

- Ø ***Historique des commandes systèmes :*** tous les systèmes d'exploitation possèdent des commandes pour obtenir *des informations instantanées* sur les processus actifs courants dans un ordinateur. Grâce à ces commandes, l'IDS peut avoir des informations précises sur les événements systèmes.
- Ø ***Accounting :*** l'accounting fournit des informations sur l'usage des ressources partagées par les utilisateurs. Ces ressources sont par exemple : le temps processeur, la mémoire, l'espace disque, les applications lancées, etc.
- Ø ***Systèmes d'audit de sécurité :*** les systèmes d'exploitation sont dotés par ce service pour définir des événements, les associer à des utilisateurs et assurer leurs collectes dans un fichier d'audit. L'IDS possède potentiellement des informations sur toutes les actions effectuées par un utilisateur.

L'avantage de ces données systèmes réside dans leur fiabilité et leur granularité fine, qui permettent un diagnostic précis des actions effectuées sur un hôte par un attaquant. Cependant, le volume d'événements généré par les audits systèmes est très volumineux ce qui implique un impact très important sur les performances de la machine surveillée.

Les IDS qui se basent sur cette catégorie des sources de données sont appelés : **Les IDS basés hôte « Host Based Intrusion Detection System ».**

II.4.4.2 Les sources d'informations réseau

Ce sont des données du trafic réseau. Cette source d'informations est prometteuse car elle permet de collecter et analyser les paquets de données circulant sur le réseau.

Les IDS qui exploitent ces sources de données sont appelés : **Les IDS basés réseau « Network Based Intrusion Detection System ».**

II.4.4.3 Les audits applicatifs

La troisième catégorie de source de données est constituée des audits applicatifs. Les données à analyser sont produites directement par une application, par exemple des fichiers logs générés par les serveurs ftp et les serveurs Web. L'avantage de cette catégorie est que les données produites sont très synthétiques, elles sont sémantiquement riches et leur volume est modéré. On note que ces types d'informations sont généralement intégrés dans les IDS basés hôte.

Vu de l'importance des IDS basés hôte et basés réseau, une étude détaillée de ces deux types d'IDS sera exhibée dans les prochaines sections.

II.4.5 La fréquence d'utilisation

La fréquence d'utilisation d'un système de détection d'intrusions peut exister selon deux formes :

II.4.5.1 Surveillance périodique

Ce type de système de détection d'intrusions analyse périodiquement les différentes sources de données à la recherche d'une éventuelle intrusion ou une anomalie passée.

II.4.5.2 Surveillance en temps réel

Les systèmes de détection d'intrusions en temps réel fonctionnent sur le traitement et l'analyse continue des informations produites par les différentes sources de données. La détection d'intrusions en temps réel permet de limiter les dégâts produits par une attaque car elle permet de prendre des mesures qui réduisent le progrès de l'attaque détectée.

II.5 L'ANALYSE BASEE CONNAISSANCE VERSUS L'ANALYSE COMPORTEMENTALE

Comme nous l'avons vu dans la section précédente, il existe deux approches pour détecter les intrusions dans les systèmes informatiques [6, 8,10]. L'approche *basée connaissance* qui se base sur la définition d'un modèle constitué des actions interdites dans les systèmes

informatiques et *l'approche comportementale* qui est basée sur la définition d'un modèle constitué des actions autorisées.

II.5.1 L'analyse basée connaissance

Cette approche de détection est désignée en anglais par le terme « ***Misuse Detection*** », qui signifie dans la littérature la détection d'une mauvaise utilisation, et il existe plusieurs traductions françaises adoptées pour cette approche, par exemple l'approche par signatures ou par scénarios.

Elle est caractérisée par l'existence d'une base de connaissances qui comporte des modèles d'attaque connus a priori qui sont appelés *les signatures*. Elle examine les activités du système et du réseau en cherchant des événements ou l'ensemble des événements qui décrivent une attaque connue. Ainsi, dans cette approche, tous ce qui n'est pas explicitement interdit est autorisé. Cette approche possède un certain nombre d'avantages et d'inconvénients [8, 9].

II.5.1.1 Les avantages de l'analyse basée connaissance

- L'analyse basée connaissance est très efficace pour la détection d'attaque avec un taux très bas des alarmes de type positif faux.
- Les alarmes générées sont significatives.

II.5.1.2 Les inconvénients de l'analyse basée connaissance

- Cette analyse basée connaissance permet seulement la détection des attaques qui sont connues au préalable. Donc, la base de connaissances doit être constamment mise à jour avec les signatures de nouvelles attaques.
- Le risque que l'attaquant peut influencer sur la détection après la reconnaissance des signatures.

II.5.2 L'analyse comportementale

Dans l'analyse comportementale, un modèle de comportement normal du système surveillé est préalablement construit. Ce modèle est appelé *profil de comportement normal qui* sera utilisé comme une référence dans la détection. Au cours de la surveillance du système, toute déviation significative du comportement courant de système contrôlé par rapport au

comportement normal de référence donne lieu à une attaque. Cette approche possède aussi un certain nombre d'avantages et d'inconvénients [8, 9].

II.5.2.1 Les avantages de l'analyse comportementale

- Ø L'analyse comportementale n'exige pas des connaissances préalables sur les attaques.
- Ø Elle permet la détection de la mauvaise utilisation des privilèges.
- Ø Elle permet de produire des informations qui peuvent être employées pour définir des signatures pour l'analyse basée connaissance.

II.5.2.2 Les inconvénients de l'analyse comportementale

- Ø Les approches comportementales produisent un taux élevé des alarmes de type positif faux en raison des comportements imprévisibles d'utilisateurs et des réseaux.
- Ø Ces approches nécessitent des phases d'apprentissage pour caractériser les profils de comportement normaux.
- Ø Les alarmes générées par cette approche ne sont pas significatives.

Cette étude comparative entre les deux approches d'analyses utilisées par les systèmes de détection d'intrusions montre l'existence d'une complétude entre ces deux méthodes. Cette complétude qui permettra de surmonter les inconvénients relatifs à chaque méthode d'analyse. Pour cette raison, il est préférable d'adopter les deux techniques d'une manière parallèle pour obtenir un système de détection d'intrusions efficace [43]. Cependant, les systèmes de détection d'intrusions commerciaux disponibles emploient seulement la technique basée signature, ce qui motive les efforts de recherche croissants pour construire des détecteurs d'anomalies efficaces pour des buts de détection d'intrusions. L'effort principal de cette recherche est concentré sur les systèmes de détection d'intrusions qui sont basés sur la technique comportementale. Pour cette raison, nous présenterons dans la section suivante les différentes approches utilisées dans la méthode de détection comportementale.

II.6 LES TECHNIQUES UTILISEES DANS L'APPROCHE COMPORTEMENTALE

Un système de détection d'intrusions basé sur la détection d'anomalies contrôle les activités du système afin de les classer comme normales ou anomalies. Il procède à construire des

profils d'un comportement normal pour les activités des utilisateurs et à observer les déviations significatives de l'activité de l'utilisateur courante par rapport à la forme normale établie. D'une façon générale, la détection d'anomalies est composée de deux phases :

- **Une phase d'apprentissage** : le système apprend le comportement normal d'un utilisateur ou un système. Il crée ainsi ***« le profil normal »*** d'un utilisateur ou d'un système à partir des données collectées.
- **Une phase de détection** : le système compare les traces d'audit courantes ou le trafic réseau aux profils pour vérifier s'il n'y a pas une activité intrusive. Si la différence entre le profil et les traces d'audit est significative, une alarme est déclenchée.

Pour pouvoir formaliser le comportement normal d'un système, des approches diverses ont été utilisées [6, 10,43]. Cette section sera consacrée à une présentation générale de ces différentes approches.

II.6.1 L'approche statistique

L'approche statistique est utilisée pour la génération d'un modèle de comportement normal d'un système. Elle consiste à générer le profil de comportement normal à partir d'un ensemble de variables aléatoires, échantillonnées à des intervalles réguliers dans le temps, ces variables peuvent être par exemple :

- Le temps CPU utilisé.
- Le nombre de connexions établi durant une période de temps.
- Les fichiers les plus fréquemment utilisés.
- Les entrées/sorties effectuées.
- Etc.

Dans cette approche, Denning [2] a proposé un ensemble de modèles statistiques, leur but est de définir à partir de n observations x_1, x_2,…x_n sur une variable donnée x, si la valeur x_{n+1} de l'observation n+1 est anormale. Parmi ces modèles, on peut citer les modèles suivants :

- ***Le modèle opérationnel*** : ce modèle est très simple, une anomalie est détectée par la comparaison de la valeur d'une nouvelle observation avec un seuil fixe qui est défini d'une manière intuitive en se basant sur les données historiques.
- ***Le modèle de déviation standard et moyen*** : Ce modèle définit un seuil d'anomalie par l'estimation d'un intervalle de confiance. L'intervalle de confiance est la moyenne et

l'écart type des n observations qui peuvent être considérées normales. Si la valeur d'une nouvelle observation est en dehors de cet intervalle alors elle est considérée anormale.

- ***Le modèle de covariances*** : Il est similaire au modèle précédent mais il se base sur la corrélation de plusieurs variables pour tirer des conclusions.

Ces approches ont été adoptées dans le développement de plusieurs systèmes de détection d'intrusions, on peut citer par exemple :

- MIDAS « Multics Intrusion Detection and Alerting System » [3].
- NIDES « Next Generation Real time Intrusion Detection Expert System » [15].

II.6.2 L'approche de la machine learning

Le but principal de l'utilisation de la machine learning est l'extraction automatique des caractéristiques des activités normales qui sont critiques pour la détection d'anomalies. A partir des données d'audit, le modèle de la machine learning essaye d'identifier des règles pour définir les comportements normaux. Ces règles seront employées pour déterminer si des événements nouvellement observés sont anormaux ou non.

Parmi les travaux qui sont basés sur cette approche, le système de détection d'intrusions basé règle TIM « Time based Inductive Machine » [16] proposé par Teng et son groupe. TIM génère des règles qui essayent de prédire les événements futurs en se basant sur des événements qui se sont déjà produits dans le passé.

Durant la phase d'apprentissage, cette approche détermine des règles temporelles qui caractérisent le comportement normal des utilisateurs. Les règles ont par exemple la forme suivante :

E1 **à** E2 **à** E3 **è** (E4 = 95%, E5 = 5%) où E1, E2,… E5 sont des événements. Cette règle signifie que la séquence d'évènements observée E1 **à** E2 **à** E3 implique ensuite l'occurrence de l'évènement E4 avec une probabilité de 95% ou l'occurrence de l'évènement E5 avec une probabilité de 5%.

Durant la phase de détection, les règles possédant des parties gauches qui correspondent à la séquence d'événements observée seront sélectionnées et l'évènement prédit de cette règle sera comparé avec le dernier évènement qui apparaît dans la séquence d'évènements observée. Si cet évènement dévie d'une manière significative de ceux prédits dans la règle, alors TIM alerte l'officier de sécurité.

II.6.3 L'approche de réseaux de neurones

Les réseaux de neurones sont utilisés dans la détection d'anomalies afin d'exploiter leurs capacités d'apprentissage. L'idée de base est d'utiliser les mécanismes d'apprentissage des réseaux de neurones pour apprendre les profils de comportements normaux des utilisateurs ou d'un système.

Plusieurs travaux ont été élaborés qui ont essayé d'abord d'apprendre à un réseau de neurones le comportement normal d'un système pour qu'il puisse par la suite de décider si un ensemble d'action est normal ou suspect. Parmi ces travaux, nous citons le travail de Debar [7] qui a proposé l'utilisation des réseaux de neurones pour construire un modèle du comportement des utilisateurs du système informatique. Le travail proposé s'intéresse à l'aspect dynamique du comportement et à sa présentation sous des séries d'actions temporelles.

II.6.4 L'approche de data mining

Le but de cette approche est l'exploitation des techniques de data mining pour extraire des anomalies à partir des grandes quantités de données du trafic réseau. Parmi les travaux existants, on peut citer ADAM « Audit Data Analysis and Mining » [17] qui est un système de détection d'intrusions qui exploite des techniques de data mining pour construire des profils du trafic réseau normaux.

ADAM utilise les règles d'association pour construire des profils du trafic de réseau normaux qui seront employées par la suite pour détecter les comportements incorrects de trafic de réseau. Pour détecter des anomalies, ADAM extrait les règles d'association à partir des données du trafic réseau et qui seront comparées aux profils du réseau. Si n'importe quelle règle d'association produite à partir des données de trafic de réseau rassemblées n'est pas incluse dans les profils, alors cette règle est considérée comme une indication d'un comportement incorrect.

II.6.5 L'approche immunologique

Vu que la détection d'anomalies est une application directe de la métaphore immunitaire, plusieurs travaux tentent de calquer la manière dont le système immunitaire naturel procède pour la distinction entre le comportement normal et le comportement suspect afin de construire des systèmes de détection d'intrusions efficaces. Parmi ces travaux nous citons le système LYSIS [35,36] qui a intégré des différentes propriétés et mécanismes inspirés par le

système immunitaire humain. Il se base principalement sur l'algorithme de la sélection négative proposé par Forrest [31]. Dans ce travail, une population de détecteurs est générée aléatoirement, puis en se basant sur les modèles de comportement normaux des utilisateurs, les détecteurs qui identifient ces modèles seront éliminés, en d'autre terme élimination des détecteurs qui détectent le soi. La population des détecteurs restants procède à contrôler les opérations effectuées dans le système de telle sorte que s'il y a une correspondance entre un détecteur et l'opération courante dans le système alors cette opération est considérée litigieux ou anormal.

II.7 LES IDS BASES HOTES VERSUS LES IDS BASES RESEAU

En raison des multiples possibilités d'attaques des systèmes informatiques et les réseaux. Il existe différents types de systèmes de détection d'intrusions [6, 8, 46,11] qui varient selon l'endroit qu'ils surveillent et ce qu'ils contrôlent (les sources d'information).

II.7.1 L'IDS basé hôte (host- based IDS)

L'IDS basé hôte contrôle un seul hôte. Il analyse des informations rassemblées d'un système d'ordinateur individuel, ce qui permet à l'IDS basé hôte d'analyser des activités avec une grande fiabilité et précision en déterminant exactement les processus et les utilisateurs impliqués dans une attaque particulière. Les avantages et les inconvénients [8, 9] de l'IDS basé hôte sont :

II.7.1.1 Les avantages d'un IDS basé hôte

- La capacité de contrôler les activités locales des utilisateurs avec précision.
- Capable de déterminer si une tentative d'attaque est couronnée de succès.
- La capacité de fonctionnement dans des environnements cryptés.
- L'IDS basé hôte fonctionne sur les traces d'audit des systèmes d'exploitation ce qui lui permet de détecter certains types d'attaques (ex : Cheval de Troie).

II.7.1.2 Les inconvénients d'un IDS basé hôte

- La vulnérabilité aux attaques du type déni de service puisque l'IDS peut résider dans l'hôte cible par les attaques.
- La difficulté de déploiement et de gestion, surtout lorsque le nombre d'hôtes qui ont besoin de protection est large.

- Ces systèmes sont incapables de détecter des attaques contre de multiples cibles dans le réseau.

II.7.2 L'IDS basé réseau (Network- based IDS)

Bien que le système de détection d'intrusions basé hôte a montré des résultats encourageants mais son problème majeur est la détection des intrusions essayées à travers le réseau. Pour détecter cette sorte d'intrusion, l'IDS a besoin de contrôler des événements multiples produits sur plusieurs hôtes. En effet, une proportion large d'intrusions est réalisée via les réseaux et en conséquence l'utilisation des informations sur le trafic réseau rend l'IDS plus efficace. Ce problème motive l'évolution des IDS basés hôte vers l'IDS basé réseau. Le système de détection d'intrusions basé réseau détecte des attaques en capturant et analysant des paquets du réseau. Les avantages et les inconvénients [8,9] de ce type d'IDS sont :

II.7.2.1 Les avantages d'un IDS basé réseau

- L'IDS basé réseau est capable de contrôler un grand nombre d'hôte avec un petit coût de déploiement.
- Il n'influence pas sur les performances des entités surveillées.
- L'IDS basé réseau est capable d'identifier les attaques de /à multiples hôtes.
- L'IDS basé réseau assure une grande sécurité contre les attaques parce qu'il est invisible aux attaquants.

II.7.2.2 Les inconvénients d'un IDS basé réseau

- L'IDS basé réseau ne peut pas fonctionner dans des environnements cryptés.
- Ce type d'IDS ne permet pas d'assurer si une tentative d'attaque est couronnée de succès.

II.8 LES ARCHITECTURES D'IMPLEMENTATION DES IDS

L'architecture d'implémentation d'un système de détection d'intrusions qui est considérée comme une stratégie de contrôle décrit la manière de contrôle effectuée par les éléments d'un système de détection d'intrusions. Nous distinguons trois approches d'implémentation [8, 13, 14,43] : ***Monolithique, hiérarchique*** et ***coopérative***.

II.8.1 L'approche monolithique (centralisée)

Les premières mises en œuvre des systèmes de détection d'intrusions ont employé une architecture monolithique sous laquelle les données rassemblées seront analysées à un point central. Puisque le contrôle de l'activité des utilisateurs d'un seul hôte ne révèle pas les attaques impliquant des hôtes multiples. L'IDS basé réseau a été développé, qui analyse le trafic de réseau pour déduire les anomalies venant du réseau.

Bien qu'un IDS basé réseau avec un serveur central a montré des résultats prometteurs pour des réseaux à petite échelle. Cependant, cette approche ne peut pas supporter un grand réseau à cause de la quantité énorme des données des différents hôtes qui doivent être analysée par le serveur central, ce qui engendre une dégradation sévère des performances de réseau. Un exemple d'un système de détection d'intrusions qui se base sur l'approche monolithique est le système NADIR [3, 76,79].

II.8.2 L'approche hiérarchique

Cette approche a été proposée pour surmonter les problèmes de l'approche monolithique. Elle est caractérisée par l'existence des secteurs de contrôle hiérarchiques. Chaque IDS contrôle un secteur avec l'élimination du transfert des données d'audit rassemblées par les hôtes locaux à un point central. Chaque IDS à n'importe quel niveau de contrôle exécute une analyse locale et envoie ses résultats d'analyse au niveau suivant dans la hiérarchie.

L'approche hiérarchique montre la meilleure incrémentabilité « scalability » en permettant des analyses locales aux secteurs de contrôle distribués. Cependant, les problèmes vus précédemment demeurent toujours. En plus, le changement de la topologie du réseau cause un changement aussi bien dans la hiérarchie de réseau et dans les mécanismes de rassemblement des rapports d'analyse locaux. Ainsi, la difficulté de détecter les attaques qui visent le niveau le plus haut de la hiérarchie. Un exemple de système de détection d'intrusions hiérarchique : GrIDS [77,78], EMERALD [74].

II.8.3 L'approche coopérative (distribuée)

Cette approche a été suggérée pour résoudre les problèmes de l'approche précédente. Elle essaye de distribuer les responsabilités d'un serveur central à un nombre de systèmes de détection d'intrusions coopératifs. La différence de cette approche avec l'approche hiérarchique est qu'il n'y a aucune hiérarchie entre les IDS distribués ce qui signifie que

l'échec de n'importe quel IDS n'empêche pas la détection d'attaques coordonnées. Parmi les systèmes de détection d'intrusions coopératifs, nous pouvons citer par exemple le système CSM [76,79] et le système AAFID [80]

II.9 UNE VUE GENERALE DE QUELQUES SYSTEMES DE DETECTION D'INTRUSIONS EXISTANTS

Il existe plusieurs systèmes de détection d'intrusions qui ont été développés. Dans cette section, nous présenterons quelques systèmes de détection d'intrusions existants.

II.9.1 IDES

IDES (Intrusion-Detection Expert System) a été développé par SRI International. Il représente le modèle de référence pour un grand nombre de systèmes de détection d'intrusions. Il a été conçu pour surveiller un seul hôte et il traite uniquement les données d'audit. Ce système de détection d'intrusions est indépendant du système surveillé, il fonctionne sur une machine dédiée, reliée au système par un réseau. Afin de détecter les violations de sécurité en temps réel, IDES s'appuie aussi bien sur une approche statistique que sur un système expert [3, 76]. Ainsi, il est constitué de deux éléments importants :

- ***Le détecteur d'anomalie*** : qui est responsable de la détection des comportements atypiques, en utilisant des méthodes statistiques du modèle de Denning [2].
- ***Le système expert*** : qui est chargé de détecter les attaques suspectes en s'appuyant sur une base de connaissances de scénarios d'attaques connus [3].

II.9.2 NIDES

NIDES (Next- Generation IDES) [3,76] est une version améliorée du système de détection d'intrusions IDES. Il assure la détection d'intrusions sur plusieurs hôtes (distribuées) en se basant toujours sur les données d'audit. Il n'y a aucune analyse du trafic réseau. Il utilise les mêmes algorithmes qu'IDES.

II.9.3 NADIR

NADIR (Network Anomaly Detection and Intrusion Reporter) [3, 76,79] est un système expert qui a été conçu pour le réseau ICN (Integrated Computing Network) du Laboratoire National Los Alamos. Son but est d'analyser les activités réseaux des utilisateurs et d'ICN en

se basant sur les règles du système expert qui définissent la politique de sécurité et les comportements suspects. L'inconvénient majeur de ce système est qu'il ne peut être porté sur d'autres réseaux, étant donné que les protocoles réseaux d'ICN ne sont pas standards.

II.9.4 DIDS

DIDS (Distributed Intrusion Detection System) [3, 76,79] est un système de détection d'intrusions basé réseau qui se base sur l'approche hiérarchique. Afin d'éviter la dégradation des performances de système, DIDS délègue certaines analyses locales aux hôtes locaux. Son architecture se compose de trois entités :

- Le « Host Monitor » : Il en existe un par hôte. Il collecte les données de l'hôte surveillé, fait une première analyse simple sur ces données puis transmet les événements pertinents au « DIDS Director ».
- Le « LAN Monitor » : Il en existe un pour chaque segment LAN. Il surveille le trafic sur le LAN, collecte les informations réseaux et reporte au « DIDS Director » les activités suspectes et non autorisées qui se sont produites sur le réseau.
- Le « DIDS Director » : Il analyse les rapports reçus du « LAN Monitor » et des « Host Monitor » afin de détecter les attaques potentielles.

II.9.5 GrIDS

GrIDS (Graph-Based Intrusion Detection System) [77, 78] a été conçu pour détecter des attaques à grande échelle. GrIDS considère les réseaux larges comme une agrégation de sous réseaux. Les données concernant l'activité des hôtes et le trafic réseau entre ces hôtes sont rassemblées dans des graphes d'activité qui révèlent la structure causale de l'activité réseau. Les nœuds d'un graphe d'activité correspondent aux hôtes constituant le réseau alors que les arêtes représentent l'activité réseau entre les différents hôtes.

Durant la phase de détection, GrIDS analyse les caractéristiques des graphes d'activité et compare ces graphes à des formes intrusives connues. S'il y a des similitudes entre ces graphes et des attaques connues, il en informe l'officier de sécurité.

II.9.6 CSM

CSM (Cooperating Security Manager) [76,79] est un système de détection d'intrusions qui peut être utilisé dans un environnement de réseau distribué. Son principal objectif est de

détecter les activités intrusives de façon non centralisée car utiliser un directeur central qui coordonnerait toutes les activités limiterait la taille du réseau « le problème d'incrémentabilité ». Pour cela, CSM doit s'exécuter sur chaque hôte connecté au réseau. Ainsi, au lieu de reporter les activités anormales à un directeur central, les CSM communiquent entre eux pour détecter d'une manière coopérative les intrusions réseaux. Les composants principaux de ce système de détection d'intrusions sont :

- Un système de détection d'intrusions local (IDS) : qui assure la détection d'intrusions pour un hôte local.
- Un gestionnaire de sécurité : qui coordonne la détection d'intrusions distribuée entre les CSM.
- Un gestionnaire d'intrus (IH : intruder handling component) : dont le rôle est d'entreprendre les actions nécessaires lorsqu'une intrusion est détectée.

II.9.7 AAFID

Le système AAFID (Autonomous Agent for Intrusion Detection) [62, 63,80] est la première tentative d'utilisation des agents autonomes pour les systèmes de détection d'intrusions basés réseau où plusieurs agents indépendants opèrent de manière coopérative pour assurer la surveillance du système cible. La décision finale du système est le résultat de coopération entre ces différents processus.

II.10 DISCUSSION

Les recherches actuelles visent à améliorer les systèmes de détection d'intrusions en raison de la complexité croissante des environnements à protéger, qui sont de plus en plus larges et dynamiques. Ainsi, la nature des intrusions actuelles et futures nous incite à développer des outils adaptatifs et automatiques. Une solution prometteuse consiste à s'inspirer à partir des métaphores biologiques pour résoudre ces problèmes. Cela est réalisé via l'exploitation des concepts et des méthodes d'identification et de détection du système immunitaire humain, qui est capable d'assurer la protection du corps contre les différents intrus d'une manière robuste, autonome, distribuée et adaptative. Alors, pourquoi de ne pas concevoir des systèmes immunitaires afin de protéger les systèmes et les réseaux informatiques.

Le système immunitaire constitue un intérêt croissant des recherches vu de sa capacité de traitement des informations. En particulier, il assure des calculs d'une manière distribuée et

parallèle. Il peut apprendre des nouvelles informations et identifier les différents modèles d'une manière décentralisée. Il détecte et répond aux envahisseurs étrangers d'une façon distribuée. L'approche immunologique est une solution prometteuse pour la détection d'anomalies vue l'analogie puissante qui existe entre l'objectif du système immunitaire humain et celui du système de détection d'intrusions ainsi que la capacité du système immunitaire humain à protéger le corps contre les intrus. Ce système présente un intérêt croissant des différents travaux existants pour exploiter ces méthodes d'identification et de détection dans des systèmes de détection d'intrusions. Pour cette raison, dans ce travail nous nous intéresserons par ce domaine de recherche.

II.11 CONCLUSION

Dans ce chapitre, nous avons présenté le système de détection d'intrusions et nous avons également étudié d'une manière détaillée les différents types d'IDS selon différents critères de classification avec la présentation générale des différentes techniques utilisées pour la détection d'intrusions.

Afin d'obtenir un système de détection d'intrusions compétent et efficace, il est souhaitable d'utiliser les deux techniques de détection comportementale et basée connaissances en parallèle pour surmonter les problèmes liés à chacune de ces deux techniques de détection. Cependant, les systèmes de détection d'intrusions commercialisés emploient seulement la technique de détection basée connaissance, ce qui motive les différents efforts de recherche dans le domaine de la détection d'anomalies.

Pour cette raison, différentes approches sont utilisées pour implémenter la technique de la détection d'anomalies. Parmi ces approches diverses, nous intéresserons à l'approche immunologique qui constitue un intérêt accru des recherches actuelles vu de l'analogie qui existe entre le système de détection d'intrusions et le système immunitaire humain. Cette approche qui présente beaucoup d'aspects intéressants pour le développement d'un système de détection d'intrusions efficace.

Le deuxième chapitre sera consacré à étudier d'une manière détaillée les systèmes immunitaires artificiels. Cette approche qui s'inspire par le mécanisme de défense humain et qui présente des capacités intéressantes d'apprentissage, d'adaptation et d'évolution pour détecter les anomalies afin d'entretenir les réseaux informatiques sereins.

CHAPITRE III
LE SYSTEME IMMUNITAIRE ARTIFICIEL

III.1 INTRODUCTION

Un intérêt croissant d'utiliser la biologie comme source d'inspiration pour résoudre différents problèmes. Ce domaine de recherche se base principalement sur l'extraction des métaphores utiles à partir des systèmes biologiques afin de créer des solutions informatiques efficaces aux problèmes complexes. Les développements les plus appréciables ont été les réseaux de neurones inspirés par le fonctionnement du cerveau, et les algorithmes évolutionnaires inspirés par la théorie de l'évolution darwinienne.

Cependant, plus récemment, un intérêt croissant pour l'utilisation d'un autre système biologique qui est le système immunitaire comme source d'inspiration pour résoudre des problèmes complexes. Le système immunitaire biologique est doté par des capacités de traitement de l'information y compris l'identification du modèle, l'apprentissage, la mémorisation et le traitement parallèle distribué. Pour ces dernières et d'autres raisons, le système immunitaire a suscité un intérêt significatif pour l'employer comme une métaphore

d'inspiration dans le calcul. Ce domaine de recherche est connu sous l'appellation des ***systèmes immunitaires artificiels***.

Ce chapitre sera composé de deux parties principales dont la première partie sera consacrée à la présentation du système immunitaire biologique, en exhibant les différents composants immunitaires et les différents mécanismes utilisés par ce système. Enfin, nous récapitulerons les propriétés intéressantes du système immunitaire qui constituent d'un point de vue informatique une source d'inspiration très riche. Tandis que la deuxième partie sera consacrée à définir le système immunitaire artificiel (AIS) et le processus de conception d'un AIS. Ainsi, nous intéresserons à présenter les différents algorithmes et les modèles immunitaires. La dernière section de cette partie expose les tentatives d'extension du système immunitaire artificiel en intégrant une nouvelle théorie immunologique qui est la théorie de danger.

III.2 LE SYSTEME IMMUNITAIRE NATUREL

III.2.1 Introduction

Toutes les créatures vivantes sont dotées par un système immunitaire, par exemple quelques plantes ont des épines protectrices pour fournir la protection de prédateurs qui les attaquent. Les animaux contiennent des os (des vertébrés) qui ont développé un système immunitaire fortement efficace et complexe.

Notre étude sera focalisée sur le système immunitaire de vertébrés, plus spécifiquement le système immunitaire humain. C'est dû aux caractéristiques intéressantes d'une perspective biologique et informatique et la compréhension de son fonctionnement.

III.2.2 Le système immunitaire

Le système immunitaire est une collection de cellules, des molécules et des organes. Il représente un mécanisme d'identification capable de percevoir et de combattre le dysfonctionnement de ses propres cellules et les micro-organismes exogènes infectieux qui envahissent le corps [20].

III.2.3 L'architecture du système immunitaire

Le corps humain est doté par plusieurs mécanismes de défense qui s'étendent à plusieurs niveaux. La première ligne de défense est composée de barrières physiques qui sont : la peau,

l'urine, les membranes muqueuses, etc. Si cette première ligne échoue d'éliminer un intrus alors le système immunitaire utilise d'autres mécanismes de défense. Le système immunitaire possède une architecture multicouche [20, 25,43] qui est constituée de deux couches inter-liées qui sont *le système immunitaire inné* et *le système immunitaire adaptatif ou acquis* (Figure 3.1).

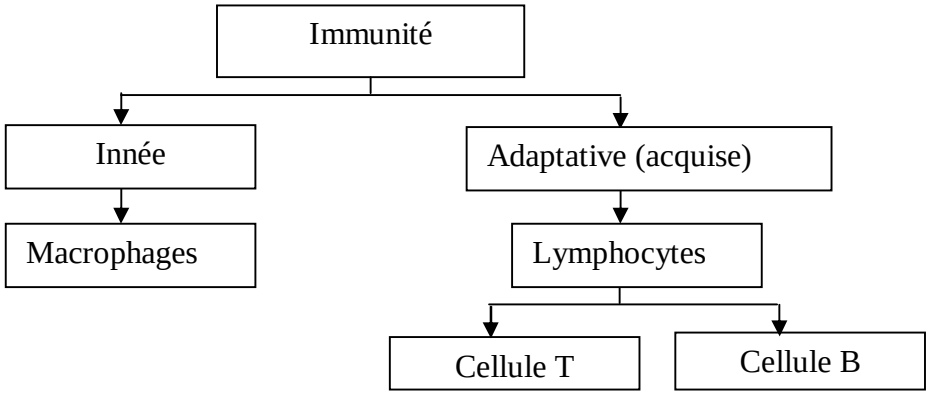


Figure 3.1 : Architecture du système immunitaire

III.2.3.1 Le système immunitaire inné

Le système immunitaire inné est composé d'un ensemble de cellules spécialisées dont le rôle principal est la liaison avec des modèles moléculaires trouvés dans des micro-organismes. Cependant, ce système ne peut pas assurer la protection complète du corps. Il est caractérisé par [25,43] :

- Les mécanismes de détection des organismes étrangers sont constants, aussi bien pour les infections répétées.
- La réponse du système immunitaire inné est non spécifique à un type particulier d'intrus mais elle est identique contre tous les pathogènes qui envahissent le corps.
- Il joue un rôle vital pour l'initialisation et la régularisation de la réponse immunitaire adaptative.

III.2.3.2 Le système immunitaire adaptatif

Le système immunitaire adaptatif est constitué de types différents de cellules dont chacun joue un rôle important. Le rôle central est assuré par les lymphocytes qui sont composés de deux types de cellules : ***cellule B*** et ***cellule T***. Le système immunitaire adaptatif est caractérisé par [25,43] :

- Ø Le système immunitaire adaptatif s'occupe avec les intrus qui ne sont pas détectés par le système immunitaire inné.
- Ø Le système immunitaire adaptatif est généré dynamiquement contre les organismes étrangers pendant sa durée de vie. Il fournit des mécanismes plus efficaces qui seront adaptés aux changements antigéniques.
- Ø Le système adaptatif est adressé à des intrus spécifiques.
- Ø La présence d'une mémoire immunologique qui permet aux cellules de se souvenir des intrus déjà rencontrés lors des prochaines rencontres.

III.2.4 La physiologie du système immunitaire

Le processus de génération et de développement de cellules immunitaires est assuré principalement par deux organes [20] : ***la moelle osseuse*** et ***le thymus***. La moelle osseuse est l'endroit de production de toutes les cellules sanguines et où certaines classes de cellules se développent. Le thymus est l'organe où une autre classe de cellules immunitaires migre pour passer l'étape de maturation.

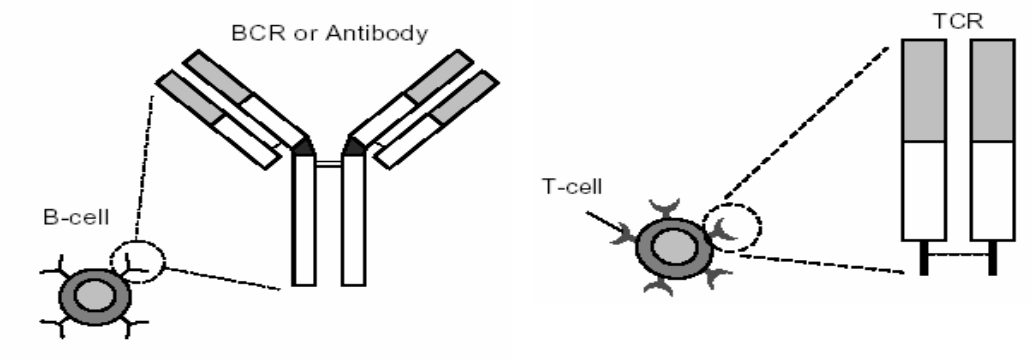
Le système immunitaire possède d'autres types de cellules immunitaires, mais cette étude sera focalisée principalement sur les ***lymphocytes.*** Les lymphocytes sont des globules blancs produits dans la moelle osseuse spécialisés dans l'identification de pathogènes qui sont composés de deux types [18,20] : les cellules B et les cellules T. Les lymphocytes qui se développent dans la moelle osseuse sont nommés des ***cellules B*** et ceux qui migrent et se développent dans le thymus sont nommés des ***cellules T***. Ces lymphocytes possèdent des récepteurs qui sont localisés sur leur surface responsable de la reconnaissance des modèles antigéniques.

III.2.4.1 Les cellules B

La fonction principale des cellules B consiste à produire et à sécréter des molécules appelées ***anticorps*** comme une réponse aux corps étrangers [20]. Chaque cellule B produit un anticorps spécifique. Les anticorps sont des protéines spécifiques qui reconnaissent et lient avec d'autre protéine particulière. Le récepteur de cellule B est appelé ***BCR*** ou bien ***anticorps*** (Ab) (Figure 3.2 (a)).

III.2.4.2 Les cellules T

Les cellules T peuvent être classées selon deux types : les cellules T d'aide (T helper) et les cellules T cytotoxiques (T killer). Les cellules T d'aide assurent des fonctions essentielles pour la régularisation de la réponse immunitaire par exemple l'activation ou la suppression du développement de certain type de réponse immunitaire. Par contre, les cellules T cytotoxiques assurent des fonctions de suppression des envahisseurs microbiens, des virus ou les cellules cancéreuses. Ainsi, les cellules T présentent des récepteurs sur leur surface (Figure 3.2 (b)).



(a) **(b)**

Figure 3.2 : Présentation d'une cellule B et une cellule T.

III.2.5 Comment le système immunitaire assure t-il la protection du corps humain ?

Notre corps est protégé par une collection diverse des cellules et des molécules qui collaborent contre n'importe quelle molécule étrangère comme les bactéries ou d'autres envahisseurs. La figure ci-dessous présente une version simplifiée des mécanismes de base de défense immunitaire (figure 3.3), et qui peuvent être résumés par les étapes suivantes :

1. Quand un intrus envahit le corps, les cellules de présentation antigénique (APC[1]) comme les macrophages procèdent à l'ingestion et la digestion de l'antigène rencontré pour le présenter comme des fragments de peptides antigéniques.

[1] **APC** est un type particulier de globine blanc du sang dont le rôle principal est la digestion des intrus cachés à l'intérieur des cellules pour les présenter aux lymphocytes.

2. Ces peptides seront liées avec les molécules MHC pour permettre leurs liaisons avec les cellules T qui ont la capacité de reconnaître la combinaison de peptide / MHC.
3. Les cellules T activées par cette identification produisent et sécrètent des lymphokines ou des signaux chimiques pour mobiliser d'autres composants du système immunitaire.
4. Les cellules B qui ont aussi des molécules de récepteur complémentaires répondent à ces signaux. À la différence des récepteurs de cellules T, ceux de cellules B peuvent reconnaître les parties d'antigènes libres sans les molécules MHC.
5. Après cette activation, les cellules B prolifèrent et se différencient et sécrètent des protéines d'anticorps.
6. La liaison entre les anticorps et les antigènes disponibles mènent à la destruction et la suppression des antigènes.
7. Un nombre de cellules B et T deviennent des cellules mémoires qui ont une durée de vie illimitée, en permettant l'élimination rapide de l'antigène s'il se présente une autre fois dans l'avenir.

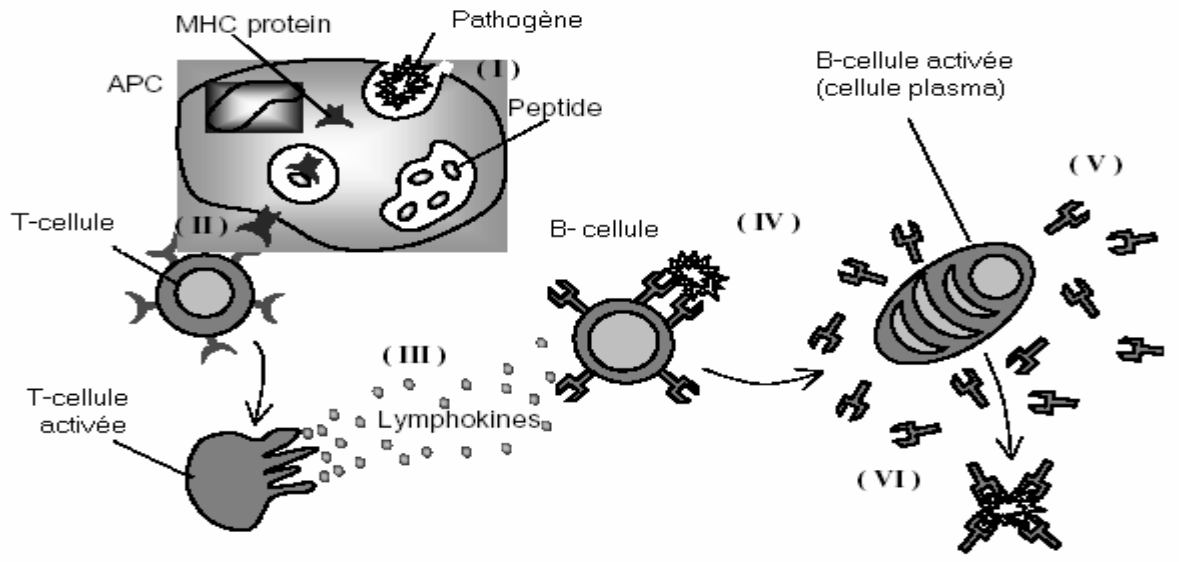


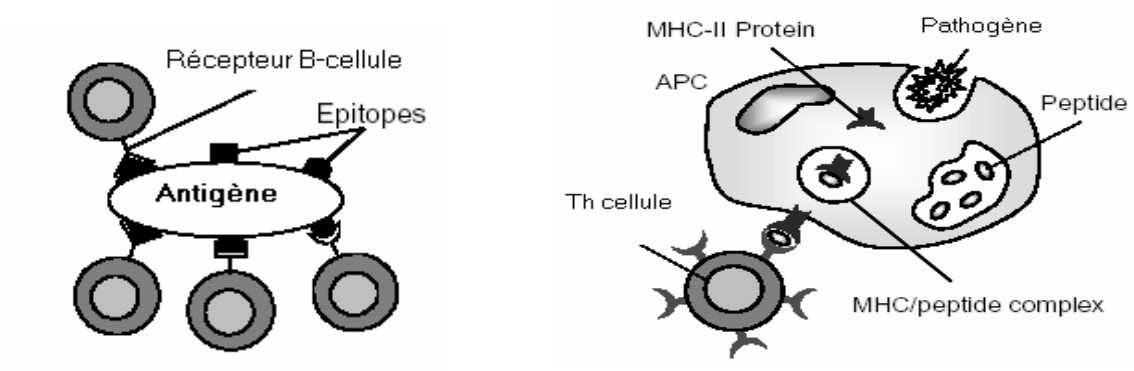
Figure 3.3 : Le processus de base de défense immunitaire

III.2.6 Les processus de base d'un système immunitaire

III.2.6.1 L'identification dans le système immunitaire naturel

La reconnaissance d'un antigène est assurée par les lymphocytes. La réaction de chaque lymphocyte est limitée au nombre de cellules étrangères connues comme *antigènes*[2]. En effet, chaque lymphocyte (cellules B et T) possède un ensemble de récepteur spécifique sur sa surface et ces récepteurs ont une forme complémentaire aux déterminants spécifiques connus comme ***épitopes*** présents sur la surface des antigènes. Un antigène est identifié s'il y'a une correspondance entre les récepteurs des cellules et l'épitope de l'antigène [18, 19].

Les cellules B et T ont une structure semblable mais elles ont une manière de reconnaissance différente [19, 20]. Les cellules B sont capables de reconnaître les antigènes libres (Figure 3.4 (a)), tandis que les cellules T ont la possibilité de reconnaître l'antigène qui est présenté par les molécules MHC[3] (Figure 3.4 (b)).



(a) **(b)**

Figure 3.4 : L'identification dans le système immunitaire

III.2.6.2 L'activation

L'identification de l'antigène est l'étape préalable pour déclencher une réponse immunitaire pour détruire l'antigène reconnu [19]. Le système immunitaire humain emploie

[2] Pathogènes

[3] Le rôle de la molécule MHC consiste à rassembler les fragments de protéines cachés à l'intérieur des cellules pour les présenter sur la surface de cette cellule infectée afin de permettre leur identification par les cellules T.

l'appariement approximatif pour déclencher la réponse immunitaire. Grâce à l'appariement approximatif, le système immunitaire est capable de détecter de nombreux antigènes. L'appariement entre un récepteur d'un lymphocyte et un épitope de l'antigène détermine ***l'affinité***[4] entre un lymphocyte et un antigène [27]. Si l'appariement entre un récepteur et l'épitope est fort alors l'affinité est grande sinon elle est petite.

Les anticorps de cellules B matures seront activés d'une manière directe ou indirecte, le type d'activation est déterminé en fonction d'un *seuil d'affinité*. Quand une cellule B correspond à un antigène avec une affinité forte au-dessus du seuil d'affinité, alors elle sera activée directement pour se développer et se différencier. Sinon, si une cellule B correspond à l'antigène avec une affinité faible au-dessous du seuil d'affinité, elle a besoin de l'aide de cellule T d'aide (T helper) pour qu'elle puisse être activée (l'activation indirecte). Pour le cas des cellules T, l'activation aura lieu s'il y a un appariement entre une cellule T et une molécule MHC qui contient un fragment de l'antigène [20].

Après l'apparition d'une réponse immunitaire suite à un stimulus antigénique. Les cellules ayant identifiées l'antigène prolifèrent et différencient, ainsi elles constitueront des cellules mémoires. Pour cette raison, la section suivante sera consacrée à la présentation de ce principe de base.

III.2.7 Le principe du mécanisme de la sélection clonale

En effet, les modèles antigéniques changent constamment, l'efficacité de la détection est maintenue par l'apprentissage dynamique de ces changements qui est assuré par la sélection clonale [20]. La théorie de la sélection clonale décrit les conséquences de la réponse immunitaire suite à un stimulus antigénique en assurant que seules les cellules qui reconnaissent l'antigène subissent aux proliférations et différentiations.

Quand un antigène envahit le corps, des cellules immunitaires reconnaissent cet antigène avec des degrés d'affinité différents. La réponse des cellules B est la production des anticorps dont chaque cellule sécrète un seul type d'anticorps qui est relativement spécifique à l'antigène. L'appariement fort entre les récepteurs des anticorps et l'antigène produit la stimulation des cellules B c'est-à-dire la prolifération (clone) et la maturation en des cellules

[4] L'affinité est le degré de liaison entre le récepteur d'une cellule et l'antigène.

de plasma. Le taux de prolifération d'une cellule est directement proportionnel à son affinité avec l'antigène [18, 20], les cellules qui ont les plus grandes affinités seront les plus proliférées et réciproquement. En plus, les lymphocytes qui ont une forte affinité peuvent se différencier en des ***cellules mémoires***. (Figure 3.5).

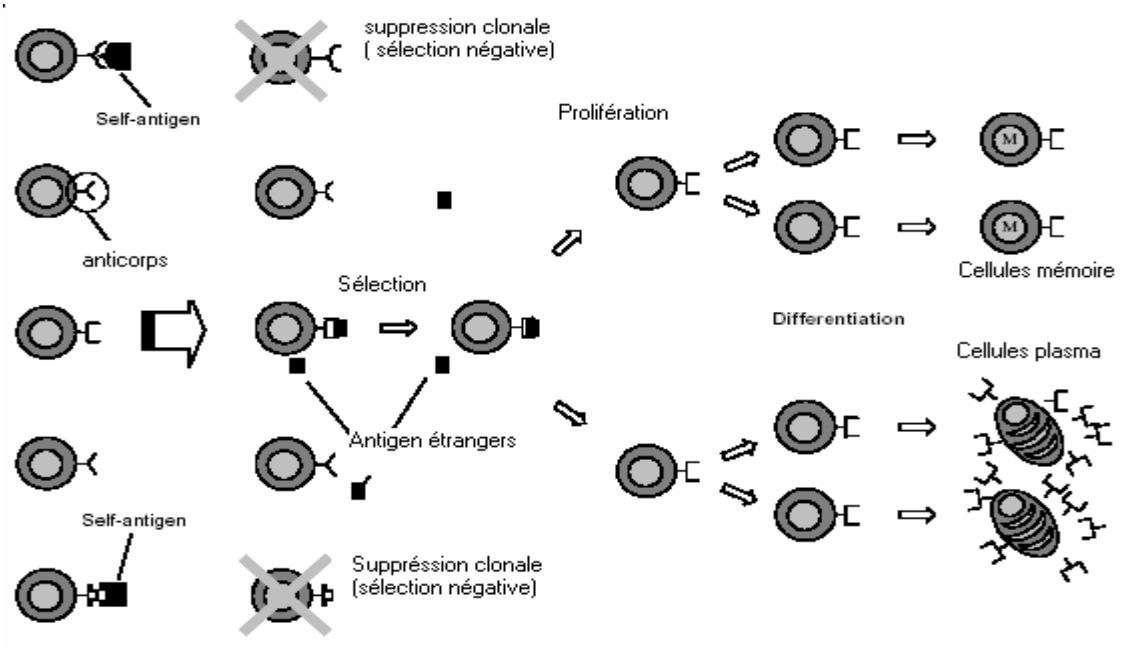


Figure 3.5 : Le principe de la sélection clonale

III.2.7.1 L'hypermutation somatique

Le résultat du processus de la sélection clonale est la reproduction de nouvelles cellules qui sont des sosies de leurs parents. Ces clones seront soumis à un mécanisme de mutation avec des taux très élevés (plus haut que des taux de mutation de cellules ordinaires). Ce mécanisme est appelé ***l'hypermutation somatique*** [18, 20]. Le résultat est des filles de la cellule B initiale qui ont des récepteurs différents du parent et par conséquence des affinités différentes aux pathogènes. L'hypermutation somatique est inversement proportionnelle à l'affinité d'une cellule c'est-à-dire les cellules qui ont les plus hautes affinités seront les moins mutées et réciproquement. Le mécanisme de l'hypermutation somatique permet au système immunitaire d'augmenter la capacité d'identification des anticorps par rapport à un antigène sélectif.

III.2.7.2 La mémoire immunitaire

L'identification de l'antigène n'est pas suffisante, le système immunitaire humain possède des capacités supplémentaires pour avoir une réponse efficace contre les pathogènes [20]. Le système immunitaire humain possède plusieurs types de réponses immunitaires (Figure 3.6), qui sont :

- La réponse primaire.
- La réponse secondaire.
- La réponse réactive croisée.

III.2.7.2.1 La réponse primaire

Cette réponse est le résultat de la première exposition à un antigène qui stimule une réponse immunitaire adaptative. Cette réponse est traitée par un petit nombre de cellules B dont chacune produisent des anticorps à des affinités différentes. Elle est caractérisée par un temps de latence grand et un petit nombre d'anticorps.

III.2.7.2.2 La réponse secondaire

Les prochaines expositions à des antigènes rencontrés précédemment stimulent des réponses secondaires. Ce processus est dû à une rétroaction d'événements passés aidant le système à apprendre. L'efficacité de cette réponse est augmentée par l'existence des ***cellules mémoires*** qui correspondent à un nombre de clone à forte affinité. Ces cellules mémoires sont produites lors de la première exposition à l'antigène.

Cette stratégie assure que la vitesse et l'exactitude de la réponse immunitaire deviennent successivement plus élevées après chaque intrusion où le système améliore continuellement ces capacités d'exécution des tâches. Ainsi, on peut noter que la réponse secondaire est caractérisée par un temps de latence très court, un taux d'anticorps plus élevé et une plus longue persistance des anticorps.

III.2.7.2.3 La réponse réactive croisée

Une réponse secondaire n'est pas seulement déclenchée par la réintroduction de même pathogène dans le futur. En effet, Il est possible qu'un antigène puisse se présenter avec des formes différentes qui sont des variations légères de l'antigène initial. Une caractéristique importante de la mémoire immunitaire est qu'elle est ***associative*** [20]. Cette particularité permet aux cellules B adaptées à un certain type d'antigène, par exemple l'antigène (A) de

présenter également une réponse secondaire efficace et rapide aux antigènes qui sont semblables à l'antigène (A). Cette réponse est connue sous l'appellation : **la *réponse réactive croisée*** ou bien ***la réaction immunologique croisée.***

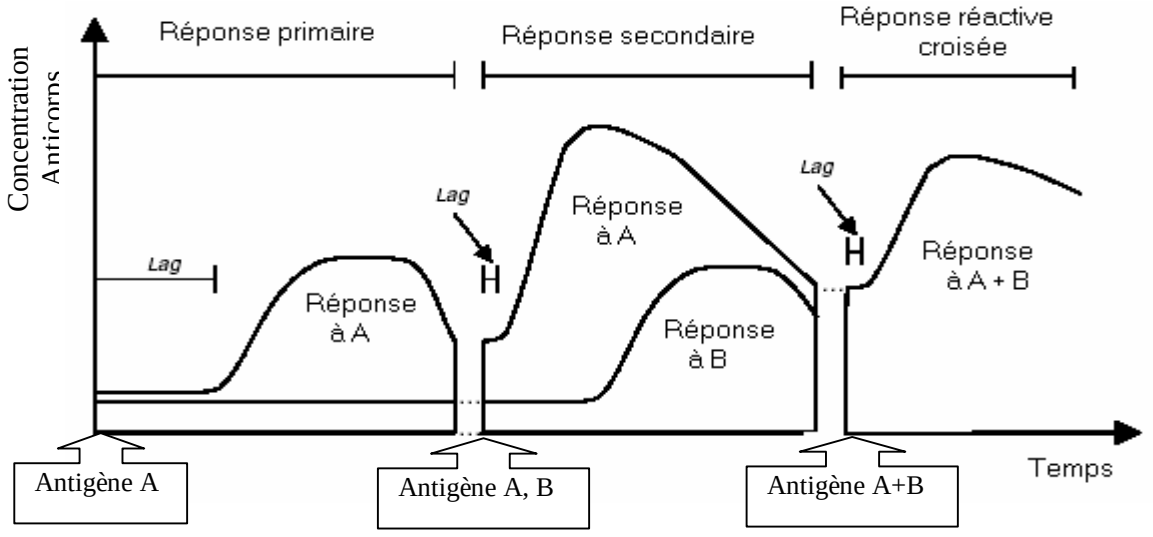


Figure 3.6 : Les différents types de réponses immunitaires

III.2.7.3 La maturation d'affinité

La maturation d'affinité est le processus qui garantit que le système immunitaire possède de plus en plus des cellules immunitaires spécialisées pour la reconnaissance des modèles antigéniques [18]. Ce processus est le résultat du mécanisme de l'hypermutation somatique suivi par une sélection. La mutation qui affecte les parties des récepteurs qui lient avec l'antigène suivi par une sélection qui garantit la préservation des solutions candidates de hautes qualités. Le récepteur qui possède la plus haute affinité permet d'avoir le plus fort appariement et ainsi la meilleure identification, ce qui permet d'avoir une réponse immunitaire exacte et efficace.

La réponse immunitaire est adaptative parce que l'opération de mutation suivie par une sélection permet aux récepteurs de cellules de s'y adapter à l'antigène. Cela garantit que les rencontres suivantes avec un certain type d'antigène mènent aux réponses plus puissantes.

III.2.8 Le répertoire cellulaire

La capacité du système immunitaire à identifier les antigènes est *complète [20].* Les récepteurs des différentes cellules immunitaires peuvent identifier les intrus externes et même les cellules du soi (la théorie du réseau immunitaire idiotypique qui sera détaillée dans les

sections suivantes). La diversité des récepteurs est assurée d'une part pendant la reproduction de molécules de récepteur par la recombinaison des segments de gène à partir de la bibliothèque de gènes. D'autre part par le mécanisme de l'hypermutation somatique qui permet la génération continuelle de nouveaux récepteurs [25].

III.2.9 La discrimination entre soi / non soi.

Si le système immunitaire est capable de reconnaître n'importe quel modèle antigénique qui est le complément des récepteurs de cellule immunitaire. Comment le système immunitaire se comporte quand il est confronté avec un antigène de soi[5] ?

La capacité du répertoire du système immunitaire pour reconnaître les antigènes est complète. Cependant, cette propriété représente un paradoxe fondamental parce que toutes les molécules qui peuvent être reconnues incluant les cellules du corps seront considérées comme ***antigènes*** ou ***antigènes de soi*** [20].

Pour que le système immunitaire fonctionne correctement, il doit être capable de distinguer entre les cellules de soi et les cellules étrangères (cellules de non soi), cette capacité est appelée la ***tolérance de soi***[6]. Ce problème est reconnu sous le nom problème de discrimination entre soi / non soi [18,20]. Donc, il doit y avoir quelque forme de sélection négative qui empêche les cellules immunitaires de devenir auto réactives.

III.2.9.1 La sélection négative pour les cellules T

Après la production des cellules T naïves dans la moelle osseuse, elles migrent vers le thymus. Les cellules T *immatures* ou *naïves* subiront alors un processus de sélection négative [18,20] dans le thymus[7]. Le processus de la sélection négative permet l'élimination des cellules T naïves qui peuvent reconnaître un antigène de soi. Les cellules T naïves qui ne reconnaissent aucun antigène du soi dans le thymus seront libérées pour la recherche éventuelle des cellules de non soi.

---

[5] Antigène de soi est une autre appellation pour les propres cellules du corps humain.

[6] Si le système immunitaire n'est pas tolérant au soi, donc une réponse immunitaire sera déclenchée contre les cellules de soi causant la maladie de l'auto-immunité.

[7] Le thymus est un organe qui est doté par une barrière thymique de sang pour éviter l'assistance des antigènes du non soi.

III.2.9.2 La sélection négative pour les cellules B

La sélection négative est appliquée aussi sur les cellules B dans la moelle osseuse, quand les cellules B immatures identifient les cellules du soi, elles seront éliminées. Ce mécanisme est appliqué seulement sur les cellules B immatures dans la moelle osseuse. La tolérance au soi des cellules nouvellement générées après le processus de la sélection clonale et l'hypermutation somatique, sera assurée par l'assistance des cellules T d'aide [43].

III.2.10 La théorie du réseau immunitaire

La sélection clonale est la théorie qui explique comment le système immunitaire répond à un antigène de non soi. Tandis que la sélection négative est employée pour éliminer les cellules auto réactives. Une autre question cruciale à être répondue est comment les cellules du système immunitaire interagissent avec d'autres cellules immunitaires ? Intéressé par ce problème, Jerne a proposé la théorie du réseau immunitaire [37] qui suggère que les interactions au sein du système immunitaire ne se limitent pas entre anticorps et antigènes, mais aussi entre les anticorps même en absence d'un stimulus antigénique. Cette interaction est assurée par des récepteurs spécialisés présents sur la surface des anticorps appelés : ***idiotope***. Alors, le système immunitaire est formellement défini par un réseau énorme et complexe de paratopes qui reconnaissent un ensemble d'idiotopes et d'idiotopes reconnus par un ensemble de paratopes. Ainsi, chaque élément pourrait reconnaître aussi bien qu'être reconnu. Ce réseau est appelé ***le réseau immunitaire idiotypique*** [18,20].

Cette théorie synthèse la propriété de la détection distribuée des systèmes immunitaires, elle montre l'état dynamique des interactions internes des lymphocytes, des anticorps et antigènes. Les lymphocytes stimulés peuvent répondre positivement ou négativement à un signal d'identification (Figure 3.7). La réponse positive est le résultat d'une liaison entre un anticorps et un antigène qui provoque l'activation et la prolifération des cellules ainsi la sécrétion d'anticorps. Par contre la réponse négative est le résultat d'une liaison entre un anticorps et un anticorps qui entraîne une suppression [20]. La chaîne continue de différentiation par antigène et la suppression par anticorps forme un réseau. Ce réseau peut atteindre le statut d'équilibre entre la suppression et la stimulation pour déterminer le système immunitaire complet [43]. Cette théorie en particulier reflète les propriétés parallèles et distribuées de système immunitaire parce que différentes réponses locales entre un anticorps et un antigène ou un anticorps et un anticorps arrivent en parallèle et à des endroits dispersés.

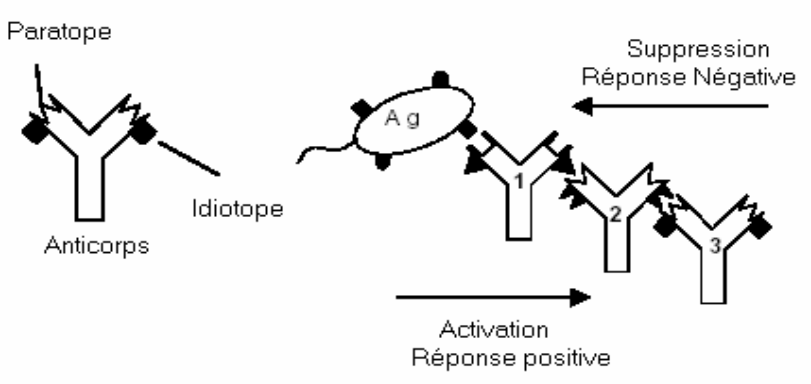


Figure 3.7 : La représentation du réseau immunitaire idiotypique

III.3 CARACTERISTIQUE DU SYSTEME IMMUNITAIRE

Dans les sections précédentes, nous avons essayé de présenter le système immunitaire humain, son processus d'identification et d'activation. Ainsi que certains processus immunitaires de base. Cette partie sera consacrée à une récapitulation des propriétés intéressantes du système immunitaire qui constituent d'un point de vue informatique une source d'inspiration très riche, qui sont :

- ***Multicouche*** **:** Le système immunitaire possède une architecture multicouche qui consiste en deux sous systèmes inter-liés qui sont le système immunitaire inné et le système immunitaire adaptatif. Ces deux systèmes combinent leurs tâches et responsabilités pour assurer la protection et la sécurité globale.
- ***Unicité*** **:** Chaque élément dans le système immunitaire assume des responsabilités particulières.
- ***Autonomie*** **:** Le système immunitaire humain ne possède aucun contrôle central ou un gestionnaire particulier. Il possède une autonomie globale dans la détection et l'élimination des intrus.
- ***Distribution*** **:** Les cellules immunitaires et les molécules sont distribuées dans le corps humain pour assurer sa protection. Il n'existe pas un point de contrôle centralisé.
- ***Parallélisme*** **:** Le système immunitaire est capable de produire plusieurs réponses immunitaires en même temps à des endroits dispersés.
- ***Tolérance au soi*** : Le système immunitaire humain peut différencier entre les cellules de soi et les cellules de non soi.

- ***Apprentissage* :** Le système immunitaire augmente la capacité d'identification des anticorps à un antigène sélectif (les réponses primaire et secondaire). Il apprend continuellement les structures de pathogènes.
- ***Adaptabilité* :** Le système immunitaire humain permet la production des cellules de plus en plus spécialisées pour l'identification des antigènes. Cela est garantit par la théorie de la sélection clonale suivie par le mécanisme de l'hypermutation somatique.
- ***Dynamique* :** Le système immunitaire change continuellement par la création de nouvelles cellules et molécules, l'élimination des cellules vieilles ou endommagées. Un bon exemple de la dynamique du système immunitaire est la théorie du réseau idiotypique.
- ***Diversité* :** Le système immunitaire naturel est capable d'identifier n'importe quels intrus qui envahissent le corps. Le répertoire cellulaire est complet. Cette particularité est due à plusieurs mécanismes qui sont par exemple : l'hypermutation somatique, la reproduction de récepteur, l'identification approximative des intrus, etc.
- ***Mémorisation* :** Après une réponse immunitaire à un antigène donné, un ensemble de cellules constituent l'ensemble de cellules mémoires qui seront dotées par une durée de vie longue afin de fournir des réponses immunitaires plus rapides et plus puissantes aux rencontres suivantes d'un même antigène.
- ***Coopération* :** Les cellules immunitaires coopèrent leurs capacités pour assurer une meilleure détection et également une protection puissante par exemple les cellules T d'aide, les molécules MHC, etc.
- ***Détection* :** Le système immunitaire est capable d'identifier et détecter les intrus dans le corps sans aucune connaissance antérieure de la structure de ces intrus.

III.4 LE SYSTEME IMMUNITAIRE ARTIFICIEL

Le système immunitaire biologique possède la capacité pour protéger le corps humain contre une variété énorme de pathogènes étrangers. Dans les dernières années, un nombre de chercheurs ont étudié le succès et la compétence de ce système naturel et ont proposé ***le modèle immunitaire artificiel*** pour la résolution de divers problèmes. Des approches diverses ont été proposées pour mettre en œuvre les mécanismes de base du système immunitaire humain [30]. Cette section sera consacrée à introduire le système immunitaire artificiel avec une présentation des différents modèles qui ont été mis en œuvre.

III.4.1 Définitions

III.4.1.1 Définition 1

Selon Timmis [27] : « Un système immunitaire artificiel est un système informatique basé sur les métaphores du système immunitaire naturel ».

III.4.1.2 Définition 2

Dasgupta a défini le système immunitaire artificiel comme suit [29] : « Le système immunitaire artificiel est la composition de méthodologies intelligentes inspirées par le système immunitaire naturel afin de résoudre des problèmes du monde réel ».

III.4.1.3 Définition 3

Tandis que Timmis et De Castro [26] ont donné la définition suivante : « Les systèmes immunitaires artificiels sont des systèmes adaptatifs inspirés par des théories immunologiques et des observations de fonctions immunitaires, des principes et des modèles, qui seront appliqués à la résolution des problèmes ».

III.4.2 Le processus de conception d'un AIS

Un schéma pour concevoir un algorithme de point de vue quantitatif exige au moins les éléments de base suivants [18,25] :

- Une représentation pour les composants du système.
- Un ensemble de mécanismes pour évaluer l'interaction des individus avec l'environnement. Les environnements sont simulés par un ensemble de stimulus d'entrée, une ou plusieurs fonctions d'évaluation.
- La procédure d'adaptation qui dirige la dynamique du système, c'est-à-dire comment son comportement varie dans le temps.

Ce schéma est adopté par Timmis & de Castro [26] qui ont proposé un processus de conception d'un AIS. Leur principe est :

- Une représentation pour créer les modèles abstraits des cellules et d'organes immunitaires.
- Un ensemble de fonction nommée *fonction d'affinité* pour évaluer les interactions entre ces éléments artificiels d'une manière quantitative.
- Un ensemble d'algorithmes pour diriger la dynamique du système immunitaire artificiel.

La figure 3.8 récapitule les éléments impliqués dans la structure de conception d'un système immunitaire artificiel (AIS).

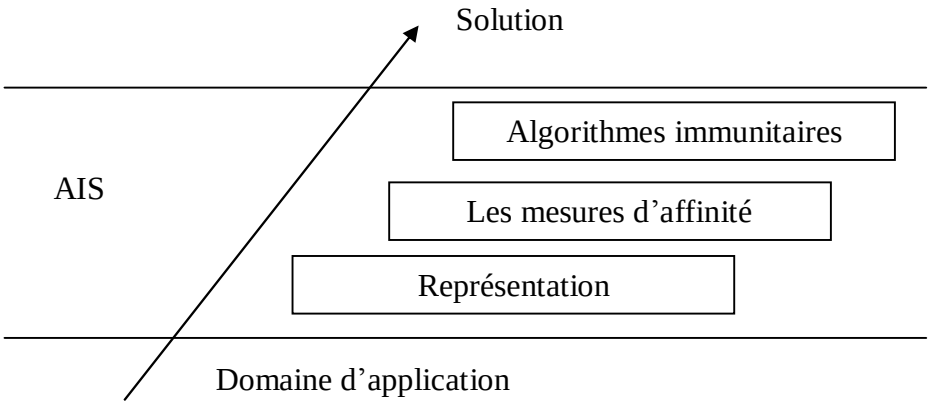


Figure 3.8 : La structure de conception d'un AIS

Pour cette raison, la section suivante sera consacrée à présenter la manière de représentation des cellules immunitaires ainsi que les mesures d'affinité existantes. Les différents algorithmes et les modèles immunitaires seront exposés d'une manière détaillée dans les sections suivantes.

III.4.2.1 La représentation

Dans les sections précédentes, les cellules B et T ont été décrites comme les cellules les plus importantes dans le système immunitaire. Elles présentent des récepteurs superficiels utiles pour la reconnaissance des intrus dont les formes de ces récepteurs sont complémentaires à la forme d'antigène. Les cellules et les molécules immunitaires sont alors les éléments qui doivent être modélisés et employés dans les modèles proposés par le système immunitaire artificiel [18].

III.4.2.1.1 Le modèle Shape - Space

Le modèle Shape- Space (Forme - Espace) a été proposé par Perelson et Oster [38] en 1979. Ce modèle permet une description quantitative des interactions de molécules de récepteur et les antigènes.

Dans le système immunitaire biologique, le concept Forme - Espace S est le degré de liaison (le degré de correspondance ou l'affinité) entre le récepteur d'anticorps (Ab ou TCR) et un

antigène (Ag). Ce degré de liaison est mesuré via les *régions de complémentarité* entre les deux éléments (Figure 3.9).

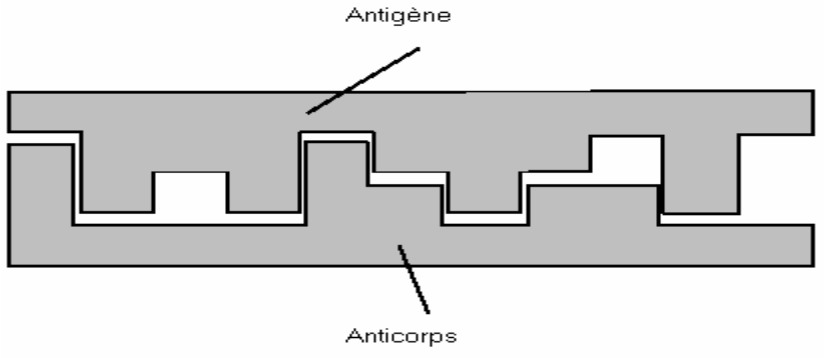


Figure 3.9 : La représentation du modèle Shape-Space.

III.4.2.1.2 Les concepts de base du modèle

Ø La forme généralisée

L'ensemble de caractéristique qui décrit les propriétés relatives à une molécule d'une perspective d'identification est nommé sa *forme généralisée*.

Mathématiquement, la forme généralisée d'une molécule (m), peut être représentée par un ensemble de coordonnées $m = < m_1, m_2, \ldots m_L >$, $m \in S^L \subseteq R^L$ tel que m est un point dans un espace L- dimensionnel, où S représente le modèle Shape-Space.

Ø L'identification via les régions de complémentarité

Une population ou bien le répertoire de N individus (les récepteurs de cellule) correspond au Shape-Space avec un volume V fini contenant N points. Comme les interactions antigène anticorps sont mesurées via *les régions de complémentarité*, les déterminants antigéniques sont aussi caractérisés par des formes généralisées.

Ø Le seuil d'affinité

Il est assumé que chaque anticorps agit spécifiquement avec tous les antigènes dont les compléments existent dans une petite région d'encerclement. Cette région est caractérisée par un paramètre « e » appelé le *seuil d'affinité*.

Ø La région d'identification

Le résultat de la définition du seuil d'affinité est le volume Vε qui est appelé la *région d'identification.*

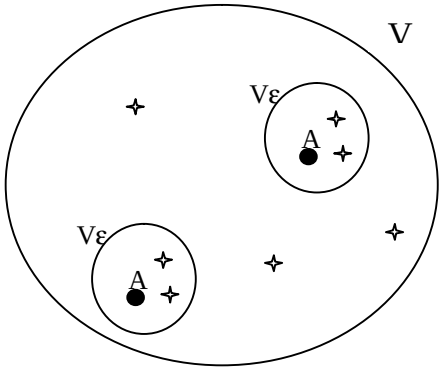


Figure 3.10 : Une représentation schématique du modèle Shape-Space

III.4.2.2 Les mesures d'affinités

L'affinité entre un anticorps et un antigène est relative à leur distance [18]. Elle peut être estimée via n'importe quelle mesure de distance entre deux chaînes (ou vecteurs) par exemple par l'utilisation de la distance *Euclidienne*, la distance de *Manhattan* ou la distance de *Hamming* [18]. Si on considère un anticorps Ab = <Ab_1, Ab_2, …Ab_L> et un antigène Ag = <Ag_1, Ag_2, …Ag_L>, alors la distance D peut être calculée selon l'une des distances précédentes qui seront présentées respectivement dans la figure suivante :

$$D = \sqrt{\sum_{i=1}^{L} (Ab_i - Ag_i)^2}$$

$$D = \sum_{i=1}^{L} \left| Ab_i - Ag_i \right|$$

$$D = \sum_{i=1}^{L} d_i \text{ où } d = \begin{cases} 1 & si \quad Ab_i \neq Ag_i \\ 0 & \text{sinon} \end{cases}$$

Figure 3.11 : Les différentes équations pour calculer l'affinité entre un antigène et un anticorps.

III.4.3 Les algorithmes du système immunitaire artificiel

III.4.3.1 L'algorithme de la sélection négative

Le système immunitaire humain utilise la sélection négative pour éliminer les cellules immunitaires immatures qui se lient avec les cellules du soi. Seulement les cellules nouvellement générées qui n'apparient à aucune cellule du soi seront libérées du thymus et la

moelle osseuse. Ensuite elles sont distribuées dans le corps humain afin d'assurer son contrôle contre les organismes étrangers. Forrest et al [31] ont proposé l'algorithme de la sélection négative qui reflète ce principe. Ils ont considéré l'algorithme de la sélection négative comme un processus de détection d'anomalies composé de trois phases principales :

- La définition du soi.
- La génération des détecteurs et
- Le contrôle d'occurrence des anomalies.

L'algorithme de la sélection négative déroulera comme suit : Étant donné l'ensemble des modèles de soi à être protégé **(P)**, générer un ensemble **(M)** de *détecteurs* qui n'identifie aucun élément appartenant à l'ensemble P. Le processus itératif pour produire l'ensemble des détecteurs **(M)** est décrit comme suit (Figure 3.12 (a)) :

3. Générer des éléments candidats **(C)** de type chaîne d'une façon aléatoire.
4. Déterminer l'affinité entre chaque élément en (**C)** avec tous les éléments de l'ensemble de soi (P).
5. **Si** l'affinité d'une chaîne dans (C) avec au moins une chaîne dans (P) est plus grande ou égale à un seuil d'affinité prédéfini **Alors**

   Cette chaîne reconnaît l'ensemble de soi, ce qui implique qu'elle doit être éliminée.

   Sinon la chaîne est ajoutée à l'ensemble de détecteurs (M).

Une fois que l'ensemble de détecteurs est produit, l'étape suivante de l'algorithme consiste à contrôler le système contre la présence des modèles de non soi tel que chaque élément détecté par les détecteurs générés est considéré comme un élément de non soi (Figure 3.12 (b)).

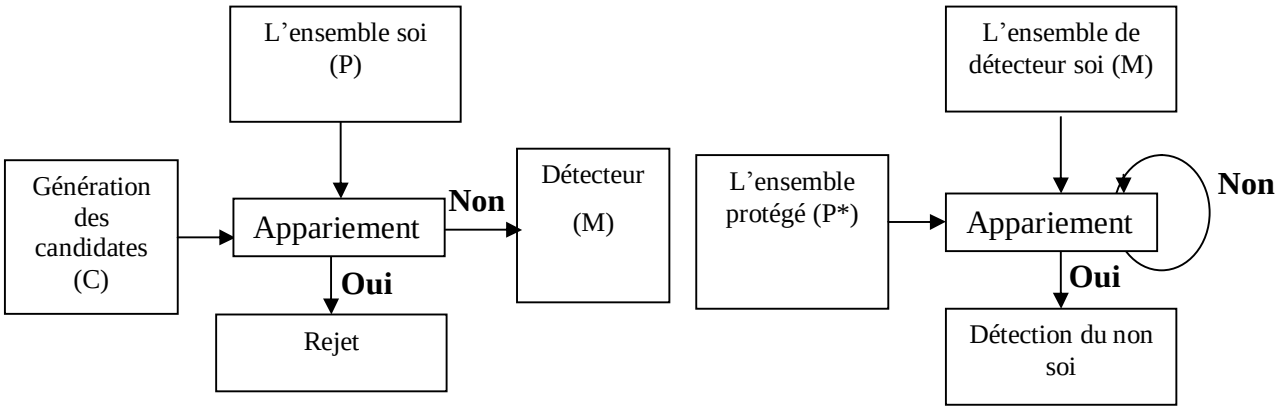


Figure 3.12 : La structure générale de l'algorithme de la sélection négative.

III.4.3.2 L'algorithme de la sélection positive

L'algorithme de la sélection positive est proposé par Forrest et al [83] et Somayaji et Forrest [84]. Cet algorithme est la solution alternative de l'algorithme de la sélection négative. La différence principale est la génération des détecteurs qui détectent des éléments de soi au lieu la génération des détecteurs qui détectent des éléments de non soi. Selon cet algorithme, un élément de non soi suspect doit être comparé avec tout l'ensemble des détecteurs de soi, s'il n'est pas détecté alors il est considéré comme un élément de non soi.

III.4.3.3 L'algorithme de la sélection clonale

De Castro & Von Zuben [23] ont proposé l'algorithme de la sélection clonale nommé CLONALG qui accompli les tâches de base impliquées dans le processus de la sélection clonale dans le système immunitaire humain. Les étapes de base de l'algorithme CLONALG sont résumées comme suit :

1. Générer une population (M) de solution candidate.
2. Déterminer l'affinité de chaque solution candidate avec un ensemble d'antigènes.
3. Choisir les n_1 meilleurs éléments de (M) dont ils ont les plus hautes affinités et produire des copies de ces individus proportionnellement à leur affinité avec l'antigène : l'élément qui possède la plus haute affinité aura le plus haut nombre de clones et réciproquement.
4. Muter toutes ces copies avec un taux inversement proportionnel à leur affinité avec l'antigène : l'élément qui possède la plus haute affinité aura un taux de mutation faible et réciproquement.
5. Sélectionner n_2 éléments à partir des clones mutés dont ils ont la plus haute affinité pour remplacer les n_1 anticorps à l'origine choisis par ces mutants.
6. Remplacer les cellules de faible affinité par des nouvelles cellules aléatoires.
7. Répéter les pas 2 à 5 tant que certains critères sont vérifiés.

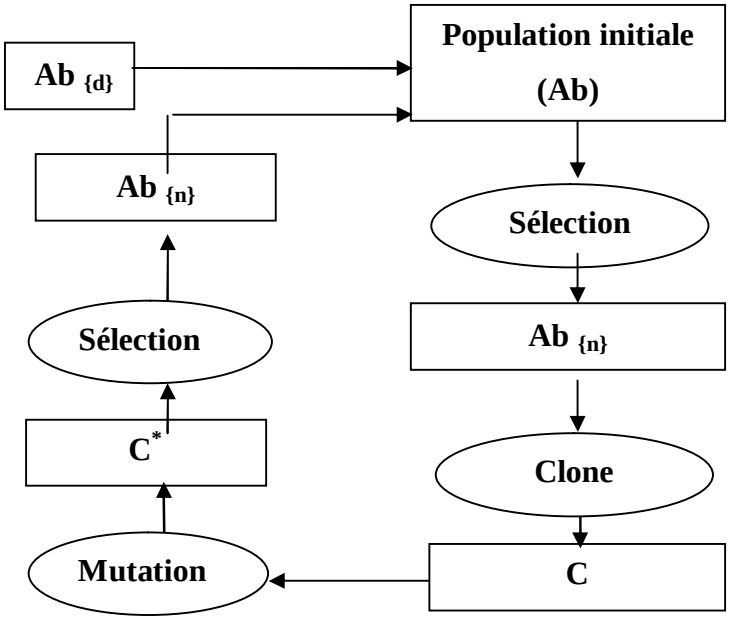


Figure 3.13 : Une représentation de l'algorithme de la sélection clonale

Smith et al [57] ont proposé une autre approche qui utilise les algorithmes génétiques et une technique de niche pour représenter le processus de la sélection clonale dans le système immunitaire humain dont le but principal est de maintenir la généralité et la diversité d'une population d'anticorps. Dans le modèle proposé, ils ont employé les algorithmes génétiques pour générer une population d'anticorps par l'utilisation d'une population d'antigène constante. En reflétant la stratégie de niche du système immunitaire humain, pour chaque génération, leur algorithme choisit un échantillon aléatoire de la population d'anticorps et un antigène aléatoire de la population d'antigènes. Ensuite, l'algorithme vérifie la correspondance entre chaque anticorps de l'échantillon avec l'antigène dont chaque correspondance implique l'augmentation du nombre d'appariements de l'anticorps correspondant. Puis, la valeur de fitness de l'anticorps, qui possède le plus grand nombre de correspondances sera incrémentée par cette valeur, par contre les valeurs de fitness des autres anticorps restent inchangeables. Dans le cas où plusieurs anticorps ont la valeur de correspondances la plus grande alors la valeur de fitness de chacun de ces anticorps sera incrémentée par le résultat de division de la valeur de correspondances la plus grande sur le nombre de ces anticorps et la valeur de fitness de ces anticorps sera incrémentée par le résultat de cette division.

III.4.3.4 L'algorithme du réseau immunitaire

La théorie du réseau immunitaire a suggéré un système immunitaire avec un comportement dynamique même en absence d'un antigène de non soi. Il existe plusieurs modèles du réseau immunitaire [18,30] à titre d'exemple on peut citer le modèle proposé par Timmis & Neal [17]. Cependant, Nous limiterons notre discussion à la description d'un seul modèle du réseau immunitaire. Ce modèle est nommé **aiNet « artificial immune NETwork »** qui a été proposé par De Castro & Von Zuben [24].

Dans ce modèle, le réseau est initialisé avec un petit nombre d'éléments aléatoirement produits. Chaque élément correspond à une molécule d'anticorps modélisée par une chaîne dans le modèle Forme-Espace. L'étape suivante est la présentation des modèles d'antigène. Chaque modèle d'antigène est présenté à chaque élément du réseau pour déterminer leur affinité selon l'équation euclidienne**.** Quelques anticorps de forte affinité sont choisis et reproduits (l'expansion clonale) proportionnellement à leur affinité. Les clones produits subissent à l'hypermutation somatique qui est inversement proportionnelle à leur affinité antigénique. Quelques clones de hautes affinités sont choisis d'être maintenus dans le réseau pour constituer la mémoire clonale. L'affinité entre tous les anticorps restants est déterminée de telle sorte que les anticorps dont l'affinité est inférieure d'un seuil donné seront éliminés du réseau (la suppression clonale). Tous les anticorps dont l'affinité avec l'antigène est inférieure d'un seuil donné sont aussi éliminés du réseau. En addition, les nouveaux anticorps aléatoirement produits sont incorporés dans le réseau (méta- dynamique). Les anticorps restants sont incorporés dans le réseau et leur affinité avec les anticorps existants est déterminée avec l'élimination des anticorps dont l'affinité est inférieure ou en dessous d'un seuil donné. L'algorithme aiNet peut être récapitulé comme suit :

1. Initialisation : créer une population initiale d'anticorps d'une façon aléatoire.

2. Présentation antigénique : pour chaque modèle antigénique faire :

2.1. Sélection clonale et expansion : pour chaque élément de réseau, déterminer son affinité avec l'antigène présenté. Sélectionner les éléments de haute affinité et les reproduire proportionnellement à leur affinité.

2.2. Maturation d'affinité : chaque clone est muté inversement proportionnel à son affinité. Sélectionner quelques clones de plus haute affinité pour constituer l'ensemble mémoire.

2.3. Interactions clonales : déterminer l'interaction réseau (affinité) de tous les éléments de l'ensemble mémoire.

2.4. Suppression clonale : éliminer ces clones mémoires dont l'affinité est moins d'un seuil prédéfini.

2.5. Méta dynamique : éliminer tous les clones mémoires dont l'affinité avec l'antigène est moins d'un seuil prédéterminé.

2.6. Construction de réseau : incorporer les clones restants de l'ensemble mémoire avec les anticorps du réseau.

2.7. Interactions de réseau : déterminer la similitude entre chaque paire d'anticorps du réseau.

2.8. Suppression de réseau : éliminer tous les anticorps du réseau dont l'affinité est moins d'un seuil prédéterminé.

3. Cycle : répéter ces pas un certain nombre d'itérations.

III.4.3.5 La mémoire immunitaire

La mémoire immunitaire indique la capacité du système immunitaire humain à protéger le corps humain contre les attaques des pathogènes qu'il a déjà rencontrés dans le passé. Les cellules B activées produisent des cellules mémoires pour assurer la détection rapide et efficace de l'antigène pour les prochaines occurrences. Il existe plusieurs manières pour mettre en œuvre la mémoire immunitaire, nous citerons seulement deux exemples parmi les travaux existants.

Le premier exemple est celui de Timmis [27] où la mémoire immunitaire a été assurée via l'exploitation de la théorie du réseau immunitaire. Les études de ce mécanisme ont montré que l'identification immunitaire forme une structure du réseau, qui décrit la mémoire du système immunitaire. Le nouveau réseau immunitaire construit par les cellules B qui survivent fournit une nouvelle solution pour les nouveaux environnements apparus sans perdre les solutions de l'environnement précédent. Cela est possible parce que la sélection des cellules B qui survivent dans le réseau n'est pas seulement basée sur leur niveau de stimulation d'antigène mais aussi par le niveau de suppression d'anticorps. Si les anticorps secrétés par des cellules B ne reçoivent pas un degré suffisant de stimulation par les nouveaux antigènes, ils ne seront pas supprimés tant qu'ils n'ont pas reçu un degré suffisant

de suppression par d'autres anticorps. Les cellules B secrétant ces anticorps restent et agissent comme des cellules mémoires dans le système.

Un autre exemple de la mise en œuvre de la mémoire immunitaire est le système immunitaire artificiel proposé par Hofmeyr [35]. Ce travail propose l'utilisation d'une population de cellule mémoire séparée au contraire des différents travaux proposés afin d'assurer une mémoire immunitaire semblable à la mémoire immunitaire du système immunitaire humain. Il a défini un cycle de vie pour les cellules générées par le système avec une représentation explicite de la population des cellules mémoires qui ont une durée de vie plus longue. Vu que cette méthode sera utilisée, elle sera détaillée dans les chapitres suivants.

III.4.4 Etude comparative entre les différents systèmes inspirés de la biologie

La figure suivante présente un tableau récapitulatif qui compare entre les différents systèmes inspirés de la biologie qui sont : les systèmes immunitaires artificiels qui sont inspirés du système immunitaire humain, les réseaux de neurones qui sont inspirés du fonctionnement du cerveau et les algorithmes évolutionnaires inspirés par la théorie de l'évolution darwinienne [18,21].

Systèmes / Caractéristiques	**Systèmes immunitaires artificiels (AIS)**	**Réseaux de neurones artificiels (RNA)**	**Algorithmes génétiques (AG)**
Composants	Chaîne d'attribut	Neurones artificiels	Chaînes de chromosomes
Endroits des composants	Endroits dynamiques	Endroits prédéfinis/ dynamiques	Endroits dynamiques
Structure	Ensemble d'éléments discrets ou gérés en réseau	Neurones gérés en réseau	Elément discret
Stockage de la connaissance	Chaînes d'attributs / connexion réseau	Poids de connexion	Chaînes Chromosomiques
Dynamique	Apprentissage / Evolution	Apprentissage	Evolution
Méta dynamique	Elimination / recrutement des composants	Algorithme constructif	Elimination / recrutement des composants
Interactions avec d'autres composants	Par l'identification des chaînes d'attribut ou des connexions réseau	Par des connexions du réseau	Par des opérateurs de recombinaison et/ou la fonction d'évaluation
Interaction avec l'environnement	Identification d'un modèle en entrée ou d'une évaluation d'une fonction objective	Les unités d'entrée reçoivent les stimuli environnementaux	Evaluation d'une fonction objective
Seuil	Influence l'affinité des éléments	Influence l'activation de neurone	Influence les variations génétiques
Robustesse	Population / réseaux d'individus	Réseau d'individu	Population d'individu
Etat	Concentration et affinité	Niveau d'activation des neurones de sortie	L'information génétique dans les chromosomes
Contrôle	Principe, théorie ou processus immunitaire	Algorithme d'apprentissage	Algorithme évolutionnaire
Possibilités de généralisation	Réaction croisée	Extrapolation du réseau	Détection des schémas communs
Non - linéarité	Fonction d'activation par attachement	Fonction d'activation neuronale	Non explicite

Figure 3.14 : Un tableau comparatif entre les caractéristiques des différents systèmes inspirés de la biologie (AIS, RNA, AG)

III.4.5 Les domaines d'application des AIS

Comme vu dans les sections précédentes, le système immunitaire artificiel possède une variété de modèles de telle sorte que chaque modèle est basé sur une partie particulière de fonctionnement du système immunitaire humain. Cette diversité permet d'utiliser le système immunitaire artificiel à plusieurs secteurs d'application pour des buts différents. D'une manière générale, parmi ces secteurs on peut citer [30, 25,43] :

III.4.5.1 La sécurité des ordinateurs

La sécurité des ordinateurs est une application directe de la métaphore du système immunitaire humain. Plusieurs travaux intéressants ont été proposés afin d'exploiter les principes de base de la détection et l'élimination, employés par le système immunitaire humain dans la sécurité des systèmes informatiques. Un travail très intéressant, et qui est considéré parmi les premières tentatives dans ce secteur de recherche et celui de Stéphanie Forrest et son groupe [31]. Dans ce travail, le problème de la protection des systèmes informatiques est vu comme une instance d'un problème de la discrimination entre le soi et le non soi ce qui signifie la capacité de distinction entre les utilisateurs légitimes et les données non infectées qui constituent le soi et les virus et les utilisateurs non autorisés qui constituent le non soi.

III.4.5.2 La détection et l'élimination des virus informatiques

Okamoto et Ishida [73] ont proposé un système multi agent basé AIS. Ce système de détection de virus opère dans un environnement distribué et hétérogène. L'algorithme de la sélection négative a été utilisé comme une méthode d'authentification de fichier.

La détection des virus est réalisée via l'appariement entre les informations propres d'un fichier tel que les premiers bits de l'entête du fichier, sa taille, le chemin d'accès et le fichier de l'hôte. La neutralisation des virus est faite par la réécriture des informations initiales sur le fichier infecté. Le système est composé de quatre types d'agents qui sont :

- Les agents anticorps qui détectent les virus sur les hôtes locaux.
- Les agents tueurs qui neutralisent les virus par les réécritures des informations initiales sur les fichiers infectés.
- Les agents de copie qui copient les fichiers non infectés qui sont équivalents aux fichiers infectés à partir des différents hôtes.

- Les agents de contrôle qui aident la communication entre les différents agents.

III.4.5.3 Optimisation

Le problème d'optimisation consiste à trouver l'ensemble absolu des meilleures conditions admissibles pour atteindre un certain objectif. Les problèmes d'optimisation apparaissent dans plusieurs secteurs d'application. Pour cette raison, ce problème est caractérisé par l'existence de plusieurs travaux qui ont exploité le système immunitaire artificiel afin de résoudre les différents problèmes d'optimisation. Par exemple le travail de De Castro & Von Zuben [23] dont le but principal est le développement d'un algorithme approprié pour le problème d'optimisation, la reconnaissance de forme. Ce travail se focalise sur le principe de la sélection clonale et la maturation d'affinité lors d'une réponse immunitaire adaptative afin de résoudre des problèmes complexes tels que l'optimisation combinatoire et l'optimisation multi modale.

III.4.5.4 Robotique

Plusieurs tentatives ont été faites pour appliquer le principe du réseau immunitaire pour contrôler les grandes populations de robots dont le but principal est d'obtenir les propriétés de base du système immunitaire qui sont l'auto organisation et le comportement de groupe. Le travail de Mitsumoto et al [70] est parmi les premiers travaux, ils ont essayé de créer un groupe de robots qui se comportent d'une façon autonome pour chercher l'alimentation sans aucun mécanisme de contrôle global.

L'idée principale dans ce travail est l'interaction entre les robots au niveau local. Les auteurs emploient trois métaphores immunologiques principales. La première métaphore est les cellules B, où un robot représente une cellule B dont chaque robot possède une stratégie particulière pour trouver l'alimentation. La deuxième est le réseau immunitaire pour garantir l'interaction entre ces robots. La troisième est le calcul de stimulation des cellules B, où le robot qui est le plus stimulé alors sa stratégie est la meilleure pour être prise en considération. Suite à ce travail, plusieurs travaux ont été proposés dans ce domaine de recherche.

III.4.5.5 Autres domaines d'utilisation

Le système immunitaire artificiel est utilisé dans plusieurs secteurs de recherche. Voici une liste non exhaustive d'autres domaines d'application :

- La maintenance des systèmes d'ordinateurs [59, 60,61].

- La reconnaissance de formes [57,69].
- Apprentissage [66, 67, 72,75].
- La classification des données [24,27]
- La planification [64, 65,71]
- Etc.

III.5 EXTENTION DU SYSTEME IMMUNITAIRE ARTIFICIEL

III.5.1 Le modèle de soi / non soi

Le système immunitaire adaptatif et en particulier les cellules B secrètent des anticorps spécifiques pour reconnaître et réagir au stimulus. La correspondance entre l'antigène et l'anticorps est l'élément de base dans la plupart des implémentations des AIS.

La caractéristique principale du système immunitaire est sa capacité à répondre aux envahisseurs étrangers sans réagir aux molécules de soi. Afin d'assurer ce rôle, le système immunitaire a besoin de différencier entre les cellules de soi et entre les cellules étrangères ou les pathogènes.

Cette discrimination est apprise tôt dans la vie grâce aux différents processus immunitaires qui jouent un rôle important pour réaliser la tolérance au soi [45]. Le modèle de soi / non soi se base sur ce principe de telle sorte que la réponse immunitaire soit déclenchée quand le corps rencontre quelque chose de non soi ou étrangère.

III.5.2 Le modèle de la théorie de danger

Au cours de la dernière décennie, une nouvelle théorie appelée la ***théorie de danger*** est devenue populaire parmi les immunologistes. Cette théorie qui est proposée en 1994 par Matzinger [47, 48,49] et qui propose des nouvelles conditions pour le déclenchement de la réponse immunitaire. Elle propose que la réponse immunitaire soit déclenchée suite à l'existence de danger et non suite à l'existence d'un élément étranger.

III.5.2.1 Critique du modèle de soi / non soi

La théorie de danger défie le point de vue du modèle de soi / non soi, elle signale qu'il existe des exemples de discrimination apparaissant après la distinction de soi / non soi [44, 45]. Par exemple :

- Aucune réaction immunitaire aux bactéries étrangères dans l'intestin ou à l'alimentation que nous mangeons bien que tous les deux soient des entités étrangères.
- L'utilité de quelques processus auto réactifs par exemple contre les molécules de soi exprimées par les cellules de stresses.
- La définition de soi est limitée au sous-ensemble vu par les lymphocytes pendant la période de maturation.
- Le corps humain change pendant sa durée de vie et par conséquence le soi change aussi. Donc, les défenses apprises tôt dans la vie contre le non soi pourraient être auto réactives plus tard.
- L'existence des situations où le système immunitaire peut réagir au soi par exemple le cas des maladies auto- immunitaires et des tumeurs. Ainsi, le cas où il n'existe aucune attaque contre le non soi dans le cas des greffes.

III.5.2.2 Le principe de base de la théorie de danger

La théorie de danger propose que le système immunitaire ne réponde pas aux éléments de ***« non soi »*** mais aux éléments qui déclenchent le ***« danger »*** dans le corps. Dans cette théorie, les envahisseurs étrangers qui sont dangereux encouragent la génération des signaux de danger. Le danger est mesuré par des dégâts sur les cellules indiquées par les signaux de détresses qui sont envoyés quand les cellules meurent d'une façon inhabituelle (nécrose) par opposition à la mort de cellule programmée (apoptose). Ces signaux de danger sont reconnus par les cellules de présentation d'antigène (APC) qui sont des cellules critiques pour l'initialisation de la réponse immunitaire.

La figure 3.14 décrit la manière de déclenchement d'une réponse immunitaire selon la théorie de danger [44]. Une cellule qui est dans la détresse déclenche un signal d'alarme et par conséquence les antigènes dans le voisinage seront capturés par les cellules de présentation d'antigène (APC) qui déplacent au nœud de lymphe local afin de présenter l'antigène aux lymphocytes. Le signal d'alarme engendre une zone dangereuse autour de lui pour permettre la stimulation des cellules B qui correspondent à l'antigène dans la zone dangereuse et qui subissent ainsi le processus d'expansion clonale, par contre celles qui ne correspondent pas ou sont trop loin ne seront pas stimulées.

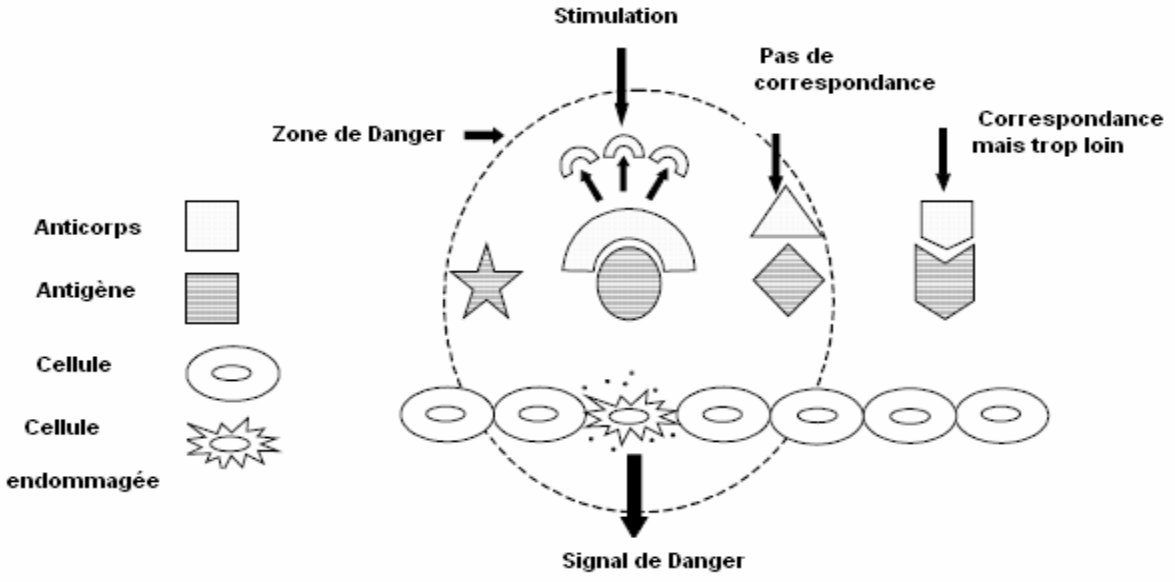


Figure 3.15 : Le modèle de la théorie de danger

III.5.2.3 La mort cellulaire

Il existe deux types majeurs de mort cellulaire soit la nécrose et l'apoptose [45].

III.5.2.3.1 La nécrose

La nécrose est une mort cellulaire dite « accidentelle » qui survient lors d'un dommage tissulaire et qui provoque l'inflammation dans le tissu.

III.5.2.3.2 L'apoptose

L'apoptose ou la mort cellulaire programmée est une forme physiologique de mort cellulaire hautement régulée nécessaire à la survie des organismes multicellulaires. Les cellules d'organismes multicellulaires s'autodétruisent lorsqu'elles sont endommagées ou lorsqu'elles présentent des dysfonctionnements. L'apoptose ne provoque pas l'inflammation dans le tissu.

III.5.3 L'emploi de la théorie de danger dans un AIS

La théorie de danger fournit des nouvelles idées intéressantes pour la représentation et le traitement des données dans les systèmes immunitaires artificiels. Elle se concentre sur ce qui est dangereux et le traitement des signaux de danger. Elle propose une nouvelle condition pour déclencher la réponse immunitaire [44, 45] et cela par l'exploitation de la mort cellulaire

nécrose qui signifie qu'il existe quelque chose anormale dans le système au contraire à la mort cellulaire apoptose.

III.6 CONCLUSION

Le système immunitaire humain protège le corps contre des dégâts d'un grand nombre de bactéries et des virus nommés pathogènes. Il assume cette tâche sans aucune connaissance antérieure de la structure de ces pathogènes. En plus, le système immunitaire humain présente plusieurs caractéristiques très intéressantes qui sont décrites dans ce chapitre, qui le rend au centre d'intérêt accru pour résoudre divers problèmes. Quelques systèmes immunitaires artificiels ont été construits pour plusieurs domaines d'applications incluant la classification, la reconnaissance de forme et la robotique, etc.

La sécurité des ordinateurs est l'un des secteurs d'application auxquels le système immunitaire artificiel est plus fréquemment adressé. Ce domaine présente une application directe de la métaphore immunitaire. Cette similitude entre les deux systèmes mène à l'apparition de plusieurs travaux dans ce domaine.

Vu d'une telle perspective, le système immunitaire humain peut être vu comme une forme de détecteur d'anomalies avec des taux très bas de faux positifs et faux négatifs. Un nombre croissant de travaux est apparu pour comprendre et extraire les mécanismes clefs par lesquels le système immunitaire humain est capable de réaliser sa détection et ses capacités de protection et d'apprentissage.

Comme le but de ce travail est lié au domaine de la sécurité des ordinateurs et plus précisément à la détection d'intrusions. Une fois que nous avons présenté les deux systèmes, le chapitre suivant sera consacré à étudier le lien étroit entre ces deux systèmes (AIS & IDS).

CHAPITRE IV
LE LIEN ENTRE UN AIS & UN IDS

IV.1 INTRODUCTION

Les approches de la sécurité des ordinateurs inspirées de la biologie sont devenues intéressantes par rapport à d'autres approches pour deux raisons à savoir :

- Ø Les systèmes informatiques et les espèces biologiques sont souvent attaqués.
- Ø Les systèmes informatiques deviennent de plus en plus complexes et les approches traditionnelles de la sécurité ne peuvent pas assumer le rôle de protection d'une manière parfaite, par contre les métaphores biologiques deviennent de plus en plus très puissantes.

Comme vu dans le chapitre précédent, le système immunitaire artificiel a été appliqué aux différents domaines de recherche. Parmi ces domaines, la sécurité et la détection d'intrusions sont le secteur d'application qui est le plus étroitement lié avec le système immunitaire humain, puisque les deux systèmes ont un but commun qui consiste à assurer la protection contre des agents étrangers.

Afin de détailler ce lien, nous allons consacrer ce chapitre pour la présentation de l'analogie entre le système immunitaire humain et le système de détection d'intrusions.

IV.2 L'IMMUNOLOGIE ET LA SECURITE DES SYSTEMES INFORMATIQUES

IV.2.1 L'immunologie

Le corps humain est constamment sous l'attaque par des micro-organismes hostiles qui sont la source de beaucoup de maladies. Le but du système immunitaire est la protection du corps contre ces pathogènes, il est face à deux aspects de problème qui sont [35] : l'identification ou la détection des pathogènes et l'élimination efficace de ces pathogènes en réduisant au minimum les dégâts causés.

IV.2.2 La sécurité des systèmes informatiques

Le problème qui touche le système immunitaire est semblable à celui de système de sécurité des systèmes informatiques : le système immunitaire protège le corps contre les pathogènes et analogiquement le système de sécurité d'ordinateur doit protéger les systèmes informatiques contre les différentes intrusions. Cette analogie peut être bien définie en exposant les problèmes confrontés par les systèmes de sécurité des systèmes informatiques [35].

- ***Confidentialité*** : le système de sécurité doit assurer la protection contre les accès non autorisés aux systèmes et aux informations.
- ***Intégrité*** : il doit protéger les données contre les opérations non autorisées telles que : la modification, la suppression, etc.
- ***Disponibilité*** : la protection des utilisateurs légitimes contre l'indisponibilité des ressources.
- ***Responsabilité*** : si le compromis d'un système d'ordinateur a été détecté, le système de sécurité d'ordinateur doit préserver l'information suffisante pour identifier ces intrus.
- ***Justesse*** : les alarmes fausses de la classification incorrecte d'événements doivent être réduites au minimum.

La similitude entre le problème de sécurité et le problème de système immunitaire peut être montrée en traduisant la langue d'immunologie dans des termes de sécurité d'ordinateur [35] : le système immunitaire détecte les abus d'une politique de sécurité implicitement indiquée par la sélection naturelle et répond à ces abus par des contre attaques de la source de l'abus. La disponibilité permet au corps de continuer son fonctionnement même dans le cas d'existence des attaques de pathogènes. La justesse signifie que le système immunitaire ne doit pas

attaquer le corps. L'intégrité signifie l'assurance que les gènes de cellule ne soient pas infectés par les pathogènes et la responsabilité signifie la recherche et l'élimination des pathogènes responsables de la maladie. Un aspect de sécurité qui n'est pas important pour le système immunitaire est la confidentialité parce qu'il n'existe aucune notion de données secrètes dans le corps qui doit être protégé à tout prix.

IV.3 L'ANALOGIE ENTRE LE SYSTEME IMMUNITAIRE ET UN SYSTEME DE DETECTION D'INTRUSIONS

Dans cette section, l'étude de l'analogie entre le système immunitaire et le système de détection d'intrusions est basé essentiellement sur le travail établit par Kim [39,43] dans lequel la démonstration de cette analogie est composée de trois étapes essentielles. La première étape présente les exigences principales d'un IDS basé réseau compétent, la deuxième étape introduit les buts de conception d'un IDS pour satisfaire les exigences de la première étape. Enfin, la dernière étape analyse les propriétés significatives du système immunitaire par une comparaison avec les buts de conception d'un IDS basé réseau. Ainsi, cette démonstration est basée d'une manière générale sur un IDS basé réseau pour deux raisons principales :

- Un IDS basé hôte peut être considéré comme l'un des composants d'un IDS basé réseau.
- Un IDS basé réseau possède la possibilité de contrôler des hôtes multiples d'une manière distribuée de la même façon que le système immunitaire.

IV.3.1 Les exigences d'un IDS basé réseau

La conception d'un IDS basé réseau compétent doit prendre en considération les fonctions suivantes :

1. ***Robustesse :***

Le système de détection d'intrusions doit être doté par des points de détection multiples pour qu'il soit assez robuste contre les attaques et les fautes de système.

2. ***Configurabilité :***

La configuration d'un IDS doit être facile aux exigences locales de chaque hôte et aux composants du réseau.

3. ***Extensibilité :***

La facilité d'étendre la portée du contrôle d'un IDS par l'ajout de nouveaux hôtes d'une manière simple indépendamment des systèmes d'exploitation.

4. ***Incrémentabilité « Scalability » :***

Il est nécessaire de réaliser l'incrémentabilité fiable pour réunir et analyser correctement le grand volume de données d'audit à partir des hôtes distribués. Dans le cas d'un IDS centralisé, la procédure de collection des données d'audit est distribuée alors que son analyse est centralisée. Cependant, il est difficile d'analyser toutes les données sur un seul IDS sans aucune perte des données.

5. ***Adaptabilité :***

Les environnements de système informatique ne sont pas statiques, les utilisateurs et les administrateurs de système changent constamment et par conséquence les intrusions changent. Un IDS doit être capable de s'adapter aux changements dynamiques afin de détecter les différentes intrusions.

6. ***Analyse Globale :***

Afin de détecter les intrusions issues du réseau, il est nécessaire de contrôler la corrélation entre les différents événements produits sur les différents hôtes car l'analyse établie par un seul hôte peut donner juste une erreur normale.

7. ***Efficacité :***

Le système de détection d'intrusions doit être simple et assez souple pour ne pas influencer sur les activités des hôtes et le réseau ce qui peut engendrer la dégradation de performance du réseau.

IV.3.2 Les buts de conception d'un IDS basé réseau

L'analyse des exigences identifiées ci-dessus peut être employée pour tirer trois buts de conception principaux d'un IDS basé réseau [39,43]. Ces buts sont la distribution, l'auto organisation et la souplesse « lightweight ».

IV.3.2.1 La distribution

Un système de détection d'intrusions basé réseau distribué délègue ses responsabilités à un nombre de composants distribués dont chacun contrôle un sous espace du système complet

d'une manière concurrente et coopérative. Un IDS basé réseau distribué satisfera les exigences suivantes :

- ***Robustesse*** : pour un IDS basé réseau distribué, l'échec d'un composant de détection d'intrusions local n'endommage pas l'IDS complet bien qu'il cause la dégradation minimale de l'exactitude de la détection complète.
- ***Configurabilité*** : la facilité de configuration d'un processus de détection d'intrusions aux exigences locales d'un hôte spécifique sans considération des exigences d'autres hôtes.
- ***Extensibilité*** : si un nouvel hôte exécutant un système d'exploitation différent est ajouté à un réseau, il est facile d'ajouter des nouveaux processus de détection d'intrusions sur cet hôte, parce que les processus de détection d'intrusions sont indépendants et ne seront pas modifiés quand un nouveau processus est ajouté.
- ***Incrémentabilité « scalability »*** : puisque la collecte et l'analyse des données d'audit seront effectuées dans le même endroit dans un hôte contrôlé localement, le grand volume de données d'audit est distribué sur plusieurs hôtes locaux et par conséquence l'IDS distribué permet plus d'incrémentabilité que l'IDS basé sur un serveur local.

IV.3.2.2 L'auto organisation

Un système de détection d'intrusions basé réseau auto organisé apprend les signatures d'intrusions qui sont inconnues et/ou distribuées sans aucune information prédéfinie. Un IDS basé réseau auto organisé satisfera les exigences suivantes :

- ***Adaptabilité*** : il est adaptatif parce qu'il n'y a aucun besoin de la mise à jour manuelle de ses signatures d'intrusions.
- ***Analyse globale*** : le système de détection d'intrusions complet fournit l'analyse globale parce qu'il est auto organisé à partir des interactions entre les différents processus de détection d'intrusions.

IV.3.2.3 La souplesse « lightweight »

Un IDS basé réseau est souple parce qu'il n'influence pas sur les performances du système. Un IDS basé réseau souple satisfera la dernière exigence.

- ***Efficacité*** : quand chaque composant d'un IDS assure une partie minimale du contrôle, les activités principales qui doivent être exécutées par les hôtes locaux et le réseau ne sont pas défavorablement affectées par le contrôle.

IV.3.3 Une analyse des caractéristiques du SIH

Une analyse prudente des capacités complexes du système immunitaire humain permet l'identification de plusieurs propriétés significatives pour la détection d'intrusions basée réseau. Ces propriétés spécifiques du système immunitaire peuvent agir ensemble pour satisfaire chacun des trois buts de conception d'un IDS basé réseau compétent.

IV.3.3.1 Un modèle distribué

La détection d'un antigène dans le système immunitaire humain est distribuée. Cette caractéristique est assurée par l'ensemble des mécanismes suivants :

- ***Réseau immunitaire idiotypique*** : le système immunitaire humain est mis en œuvre par les interactions entre différents types de cellules. Au lieu de l'utilisation d'un coordinateur central, le système immunitaire humain assure des réponses immunitaires appropriées par maintenir le statut d'équilibre du réseau immunitaire entre la suppression par anticorps et l'activation par antigène.
- ***Ensemble d'anticorps unique*** : le système immunitaire humain produit des groupes divers d'anticorps pour détecter les différents antigènes. La diversité des anticorps est maintenue par la sélection naturelle dans la bibliothèque de gènes et la sélection clonale de telle sorte que chaque ensemble d'anticorps est unique et indépendant Ces propriétés n'exigent pas un coordinateur central et permettent la détection des intrus au niveau local des anticorps.

IV.3.3.2 L'auto organisation

La réponse immunitaire est composée de trois étapes évolutionnaires qui sont : l'évolution de la bibliothèque de gènes, la sélection négative et la sélection clonale. Ces trois étapes sont auto organisées plutôt que la direction par un organe central ou obtenir une information prédéfinie.

- ***L'évolution de la bibliothèque de gènes*** : La production des anticorps compétents nécessite une certaine connaissance de propriétés antigéniques. Le système immunitaire apprend ces connaissances par l'évolution de la bibliothèque de gènes. Puisque ce processus d'évolution est auto organisé, il permet aux bibliothèques de gènes d'agir comme une archive d'information afin de détecter les antigènes observés.

- ***La sélection négative*** : la sélection négative élimine les anticorps immatures qui correspondent avec les cellules du soi. Le système ne possède aucune information globale sur les cellules de soi, la satisfaction de cette contrainte est assurée dans le thymus et la moelle osseuse par la présentation des cellules de soi afin de supprimer les anticorps qui attaquent ces cellules.
- ***La sélection clonale*** : Ce processus permet la prolifération des meilleurs anticorps alors que les anticorps de faible affinité meurent après une durée de vie. Ainsi, selon les antigènes existants seulement les anticorps les plus convenables survivent. De la même façon au lieu de l'obtention de l'information prédéterminée sur les antigènes spécifiques, le système immunitaire est capable de sélectionner d'une manière autonome les anticorps les plus convenables en agissant avec les antigènes existants.

IV.3.3.3 La souplesse

Le système immunitaire humain est souple. Les mécanismes décrits au-dessous permettent au système immunitaire d'être souple et sont concentrés sur trois idées :

i) La détection d'un ensemble énorme d'antigènes avec un ensemble plus petit d'anticorps.
ii) La réutilisation de l'information antigénique connue préalablement d'une manière efficace pour les prochaines détections.
iii) La production de nombreux anticorps avec un nombre limité de gènes.

- ***La liaison approximative*** : la réponse immunitaire est déclenchée quand l'affinité de la correspondance entre un anticorps et un antigène dépasse un certain seuil. Cette liaison approximative permet qu'un seul anticorps puisse détecter n'importe quel nombre d'antigènes tant que leur affinité est au-dessus de seuil d'affinité. Cette liaison approximative contribue pour augmenter la généralité de système immunitaire.
- ***Les cellules mémoires*** : les cellules mémoires stockent l'information génétique d'un antigène précédemment détecté afin de répondre d'une façon rapide et efficace à ces prochaines rencontres dans l'avenir.
- ***Expression de gène*** : le système immunitaire maintient la diversité d'anticorps pour assurer la détection efficace d'un grand nombre d'antigènes. Le processus de développement d'anticorps « expression de gène » emploie plusieurs mécanismes génétiques pour produire différents anticorps.

- ***Hypermutation somatique*** : le système immunitaire apprend dynamiquement le changement antigénique via la sélection clonale. Le résultat de la sélection clonale suivie par l'hypermutation somatique est un ensemble d'anticorps avec des variations plus larges de leurs gènes de correspondance avec les antigènes.

IV.3.4 Discussion

Le système immunitaire humain est distribué par son réseau immunitaire et les ensembles d'anticorps uniques. Ainsi, il est auto organisé en conséquence de trois processus évolutionnaires qui sont l'évolution de la bibliothèque de gènes, la sélection négative et la sélection clonale. Il est souple par la généralité de la liaison approximative, l'expression de gène, l'hypermutation somatique et l'efficacité des cellules mémoires. Ces propriétés significatives montrent le lien étroit entre le système immunitaire humain et le système de détection d'intrusions. Elles montrent que la réalisation des exigences principales pour la conception d'un système de détection d'intrusions basé réseau est envisageable par l'utilisation d'un système immunitaire artificiel, ce qui motive les différentes recherches exploitant les systèmes immunitaires artificiels dans le domaine de sécurité.

IV.4 CONCLUSION

Dans ce chapitre, nous avons exposé le lien entre l'objectif d'un système de détection d'intrusions et celui du système immunitaire. A cette démonstration, nous pouvons ajouter un autre exemple simple qui illustre encore cette analogie. Le système immunitaire possède une architecture multicouche comme nous l'avons décrit dans le chapitre précédent qui est composée principalement de deux couches qui sont le système immunitaire inné et le système immunitaire adaptatif. Le système immunitaire inné est semblable aux détecteurs basés signature d'un IDS car les deux systèmes ont les connaissances antérieures des attaques. De la même façon le système immunitaire adaptatif est semblable au détecteur d'anomalies d'un IDS parce que les deux systèmes produisent de nouveaux détecteurs d'une manière adaptative afin de détecter les attaques inconnues.

A la fin de ce chapitre, nous pouvons conclure que le système immunitaire est la meilleure solution qui pourrait être utilisée pour concevoir un système de détection d'intrusions compétent et efficace.

CHAPITRE V
L'APPROCHE PROPOSEE POUR LA DETECTION D'INTRUSION PAR SIA

V.1 INTRODUCTION

Suite à la proposition de l'algorithme de la sélection négative par Stéphanie Forrest et son groupe qui ont essayé d'appliquer des métaphores du système immunitaire pour la détection des virus dans un système d'ordinateur. Ainsi, vu l'analogie entre l'objectif du système de détection d'intrusions et celui du système immunitaire humain, plusieurs travaux sont apparus dont le but principal est d'exploiter et intégrer les différents mécanismes utilisés par le système immunitaire pour la détection des intrus.

Dans ce chapitre nous exposerons les différents travaux qui ont exploité les systèmes immunitaires artificiels dans le domaine de détection d'intrusions. Ensuite, une discussion sera établie sur ces différents travaux afin de montrer certains problèmes existants dont le but est la présentation des motivations de l'approche proposée. La section suivante sera consacrée à la présentation détaillée de cette proposition et cela par la description des caractéristiques de l'algorithme ainsi le pseudo code de l'algorithme sera exposé avec la

exposition des différents mécanismes mis en œuvre. La dernière section de ce chapitre sera consacrée à l'exécution d'un certain nombre d'expérience afin d'effectuer une vue générale sur les résultats donnés par cet algorithme

V.2 LE SYSTEME IMMUNITAIRE ARTIFICIEL POUR LA DETECTION D'INTRUSIONS

Le travail de ***Somayaji et al [32]*** est la première tentative qui intègre l'immunologie dans un système de détection d'intrusions dont le but était la conception et la vérification d'un système de détection d'intrusions basé sur la notion de soi. Ce travail s'inspire principalement du travail de Forrest et al [31]. Le système proposé est basé hôte, il contrôle principalement les processus privilégiés[8]. Le système collecte les informations pour définir le soi pendant la période d'apprentissage, ces informations sont sous la forme des séquences de commande de sendmail (un agent de transport des emails dans le système UNIX) dont le résultat est une base de données de séquences de commandes. Ensuite, pendant la phase de test, le système vérifie l'occurrence des nouvelles séquences qui n'existent pas dans la base de données du programme en exécution. Chaque séquence qui n'existe pas dans la base de données est considérée comme une erreur. Une anomalie est déclenchée si le nombre d'erreurs atteint un seuil prédéfini.

Le travail proposé par ***Hofmeyr et al [33]*** espère améliorer les systèmes de détection d'intrusions basés sur la détection d'anomalies. Le principe de ce travail est semblable au travail précédent de Somayaji et al [32] mais avec quelques améliorations. Les séquences d'appels système sont représentées dans des fenêtres d'appels système qui seront confrontées à la base de données de comportements normaux. L'évaluation de la similarité entre deux séquences est établie via la distance de Hamming. Une erreur est déclenchée s'il y'a une déviation du comportement normal et si le nombre d'erreurs dépasse un certain seuil, une alerte est générée. Le point intéressant de ces travaux c'est la démonstration que les

[8] Ce sont des programmes en exécution qui performent des services (comme l'envoi ou la réception du courrier), qui exige l'accès aux ressources de système qui sont inaccessibles à l'utilisateur ordinaire.

séquences d'appels système des processus privilégiés sont appropriées pour définir l'ensemble de soi dans un hôte contrôlé. Cependant, ces travaux pour la détection d'intrusions inspirés d'immunologie n'ont pas incorporés les différentes propriétés du système immunitaire, ils ont essayé seulement d'utiliser le concept de base du système immunitaire humain qui est la détection des entités du non soi.

Afin de construire un système robuste, distribué, tolérant aux erreurs et adaptatif et qui intègre les différentes propriétés du système immunitaire, ***Hofmeyr [35] et Forrest & Hofmeyr [34,36]*** ont développé un système de détection d'intrusions basé AIS nommé ***LYSIS « Lightweight Intrusion detection SYStem »***, qui est désigné à protéger un réseau local (LAN) contre les attaques arrivants du réseau par l'exploitation de plusieurs mécanismes inspirés du système immunitaire humain.

Le système de détection d'intrusions proposé est basé réseau et il examine les connexions TCP dont les connexions normales définissent l'ensemble de soi. Ces connexions TCP sont sous la forme des triplets de chemin de données qui contient les attributs suivants : l'adresse IP source, l'adresse IP destination, le port ou le service.

Les détecteurs sont représentés par des chaînes binaires, générés par l'algorithme de la sélection négative. La règle de correspondance utilisée entre un détecteur et un élément de soi ou de non soi est une règle de correspondance partielle qui est *la règle de r bits contigus* qui indique l'existence d'une correspondance entre un détecteur et une chaîne quelconque s'ils ont au moins r bits contigus en commun.

Si le nombre de correspondances d'un détecteur avec les chaînes de non soi dépasse le seuil d'activation alors une alarme est générée. Les détecteurs qui produisent beaucoup d'alarmes seront choisis pour constituer l'ensemble mémoire avec un seuil d'activation minimal. La costimulation est fournie par l'utilisateur afin d'augmenter le taux vrai positif. Le cycle de vie d'un détecteur peut être résumé par le schéma suivant :

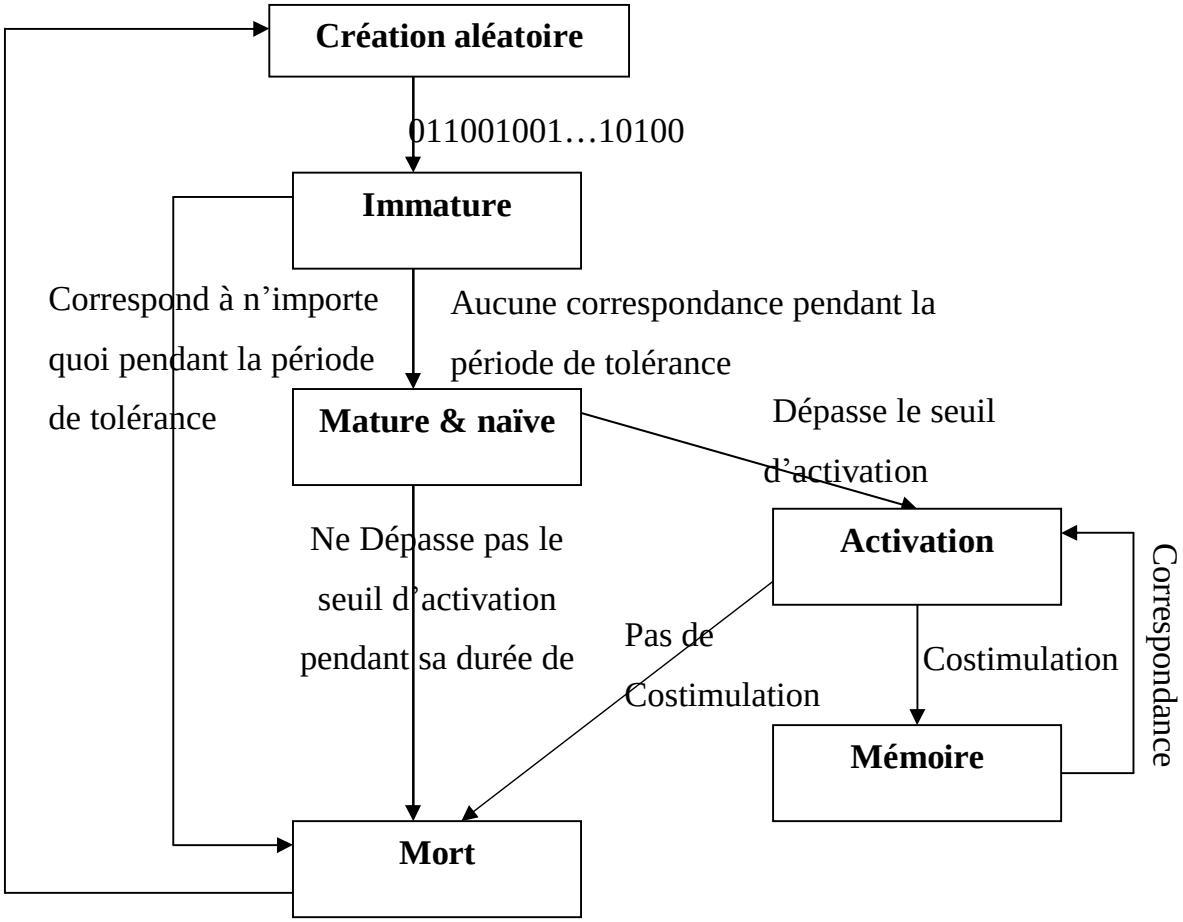


Figure 5.1 : Le cycle de vie d'un détecteur

Dasgupta et al [81] propose une structure générale pour un système de détection d'intrusions basé AIS. Cette structure utilise une architecture multi agent pour supporter le modèle du système immunitaire pour la détection d'intrusions. Cette structure du système de détection d'intrusions basé AIS suit simplement la caractéristique de détection multi niveau du système immunitaire humain plutôt que l'emploi des propriétés de système immunitaire humain, comme la détection distribuée, la sélection clonale et la sélection négative, etc.

Les agents dans ce modèle contrôlent le système aux niveaux divers d'une façon hiérarchique, activent les signaux d'avertissement, communiquent leurs signaux d'avertissement locaux et prennent des décisions basées sur les signaux d'avertissement rassemblés localement. Pour réduire les avertissements faux dans l'IDS proposé, ce modèle se base sur les décisions collectives. Il contrôle en même temps les anomalies dans les modèles d'utilisation de système par les utilisateurs, les modèles d'utilisation de ressources de système et les modèles d'activité de processus et les modèles du trafic réseau et produit les

signaux d'avertissements finaux si la corrélation des avertissements locaux constituent une intrusion.

Kim & Bentley [41] ont évalué les performances de l'utilisation de l'algorithme de la sélection négative comme un détecteur d'anomalies pour la détection d'intrusions de réseau. Ce travail se base sur le système LISYS avec la définition d'un ensemble large de soi afin de tester les performances de cet algorithme dans un environnement du réseau réel. Deux changements principaux ont été faits pour s'adapter à des ensembles de soi plus compliqués à savoir : l'adoption de plus grande cardinalité de génotypes et l'application d'une fonction de correspondance sur phénotypes plutôt que génotypes.
Les résultats de ce travail ont montré une performance pauvre de l'algorithme de la sélection négative lorsqu'il est appliqué aux problèmes du monde réel. Comme le système à protéger devient de plus en plus large, il est difficile de trouver un ensemble de détecteur adéquat pour fournir la couverture nécessaire.
Kim et Bentely ont suggéré de redéfinir le rôle de l'algorithme de la sélection négative dans le système de détection d'intrusions réseau afin de surmonter les problèmes d'incrémentabilité.

Kim [42] a proposé un nouvel algorithme ***DynamiCS*** qui intègre les principaux concepts utilisés dans le système LISYS proposé par Hofmeyr [35] comme par exemple : la mémoire immunitaire, la période de tolérance des détecteurs, le seuil d'activation, la costimulation, etc. Ce travail tente de traiter le comportement dynamique de l'environnement contrôlé par un système de détection d'intrusions qui change d'une manière constante dans lequel le comportement normal à un moment peut devenir anormal plus tard, Kim a essayé de vérifier les deux propriétés suivantes :

- L'apprentissage progressif du comportement normal du système en se basant sur un petit sous-ensemble de soi à la fois.
- Les détecteurs doivent être remplacés si les comportements normaux déjà appris ne représentent plus des comportements normaux puisqu'ils sont soudainement altérés par une opération de changement de soi légale.

L'algorithme intègre trois stages évolutionnaires inspirés du système immunitaire humain qui sont : l'algorithme de la sélection négative, l'algorithme de la sélection clonale et l'évolution

de la bibliothèque de gènes. Le stage de l'évolution de la bibliothèque de gènes consiste à réintégrer les détecteurs mémoires qui ne détectent plus des intrusions dans la population des détecteurs immatures après une opération de mutation effectuée sur ces détecteurs dont le but principal est de sauvegarder l'information génétique utile de ces détecteurs.
Les détecteurs sont représentés sous la forme des règles conjonctives et la règle de correspondance est effectuée au niveau phénotype. Cet algorithme présente des taux bas de vrai négatif avec des taux élevés de vrai positif. Cependant, il requiert un nombre élevé de costimulation qui nécessite l'intervention d'un opérateur humain pour confirmer si un détecteur mémoire à réellement détecter un antigène de non soi afin de réduire le taux d'erreurs de vrai négatif.

Depuis la proposition de l'algorithme de la sélection négative par Forrest et al comme un algorithme pour la détection d'anomalies, plusieurs travaux ont été proposés en se basant sur cet algorithme. ***Dasgupta & al [52]*** ont examiné l'idée d'utiliser la sélection négative pour le problème de la détection d'intrusions réseau. Ils ont comparé les performances de l'algorithme de la sélection négative et l'algorithme de la sélection positive, qui est l'un des algorithmes de la détection d'anomalies conventionnelles.
L'algorithme de la sélection positive proposé génère les échantillons positifs à partir des données d'apprentissage et utilise les arbres k dimensionnels afin d'obtenir une recherche rapide des k plus proches voisins. Afin de démontrer l'effet de la corrélation des paramètres dans la détection d'intrusions, ils ont utilisé un seul paramètre puis la combinaison des trois paramètres. Une alerte est générée si la distance entre un élément et l'ensemble de soi dépasse un certain seuil. Un sous ensemble de données est employé pour tester le système avec l'utilisation de cinq attaques différentes. Le test a déroulé en deux semaines dont la première est consacrée à la phase d'apprentissage et la deuxième pour la phase de vérification. L'utilisation d'un seul paramètre a produit la détection de trois attaques parmi les cinq attaques simulées tandis que les cinq attaques ont été détectées par la combinaison des trois paramètres. Les résultats expérimentaux montrent que l'algorithme de la sélection positive assure des taux élevés de détection mais il nécessite un espace mémoire très grand pour stocker tous les échantillons qui constituent l'espace de soi, et comme la quantité des données générée par le trafic réseau est très large, cette méthode devient inacceptable.

Pour l'approche de la sélection négative, ils ont utilisé une représentation basée sur des valeurs réelles pour caractériser l'espace de soi / non soi. Un algorithme génétique est utilisé pour générer des règles afin de couvrir l'espace de non soi. Ils ont exploité une technique de niche pour générer un ensemble de solution pour la couverture de la région de non soi.
Egalement, ils ont essayé d'étendre l'approche de Forrest de deux classes (soit le soi et le non soi) en une approche multi classes ce qui signifie la définition des degrés de normalité dans l'espace de soi (ou bien dans l'espace de non soi). Pour cela, ils ont défini un paramètre de variabilité v comme la distance de soi qui est considérée toujours comme normale. Les mêmes données utilisées dans le test de l'algorithme de la sélection positive seront employées pour vérifier l'algorithme de la sélection négative. Les résultats ont montré que l'algorithme de la sélection négative présente un taux de détection moins que celui présenté par l'algorithme de la sélection positive. Cependant, l'algorithme de la sélection négative utilise 1% de l'espace mémoire utilisé par l'algorithme de la sélection positive. Dasgupta et son groupe concluent la possibilité d'utiliser l'algorithme de la sélection négative pour construire un IDS. Comme un futur travail, ils espèrent d'utiliser plus de données pour tester le système.

Suivant l'étude établie par Kim et Bentley [41] sur les problèmes d'incrémentabilité et le taux de faux positif de l'algorithme de la sélection négative. ***Balthrop et al [51]*** fournissent une étude détaillée du système de détection d'intrusions LYSIS.
Ils ont proposé une forme simplifiée de LYSIS qui explore les différents mécanismes utilisés pour contrôler le taux de faux positif. Cette forme simplifiée contient seulement quelques mécanismes existant dans le système LYSIS, qui sont :

- Les détecteurs utilisés sont modélisés par des chaînes binaires de 49 bits.
- L'utilisation de l'algorithme de la sélection négative pour garder seulement les détecteurs valides pendant une période de tolérance.
- L'utilisation de la règle de correspondance partielle r bits contigus.
- L'utilisation d'un seuil d'activation pour définir le nombre de liaison qu'un détecteur doit atteindre pour qu'il soit activé. Une fois le détecteur atteint le seuil d'activation, il peut générer une alarme.

En première étape, ils ont essayé de définir les meilleurs paramètres qui contribuent à la minimisation de taux de faux positif. D'après l'analyse de l'effet des différents paramètres, Ils ont constaté les remarques suivantes :

- Ø Si le paramètre r possède la valeur maximale alors le nombre de détecteur atteint un point de saturation en assurant une meilleure couverture.
- Ø Une longue période de tolérance réduit le taux de faux positif.
- Ø Et que l'incrémentation du seuil d'activation réduit le nombre de faux positif.

Puis, en se basant sur les résultats de la dernière expérience, les meilleurs paramètres ont été affectés pour exécuter différentes attaques sur le système.

La version simplifiée de LYSIS décrit dans le travail précédent a été réutilisée dans le travail proposé par ***Balthrop et al [50]*** dont l'objectif principal de cette recherche se focalise sur l'amélioration de la représentation des détecteurs en explorant une nouvelle représentation par l'amélioration de la règle de r bits contigus par le schéma de r-chunk. La différence de cette représentation avec la règle de r bits contigus est que seulement les r bits de détecteur sont spécifiés afin de réduire le nombre de holes[9] qui peuvent être présentés dans la couverture des détecteurs. Ainsi, ils ont examiné l'effet des masques de permutation[10] sur le système et qui permettent l'incrémentation de la généralisation de la couverture des détecteurs. L'exécution du système avec le schéma de r-chunk a montré des bons résultats, bien que les auteurs aient constaté par la suite que l'augmentation de la performance est relative à la configuration du réseau d'essai. En plus, ils ont observé que les variations de r engendrent des petits effets au contraire avec des détecteurs de longueur complète.
Balthrop et al concluent que cette nouvelle règle de correspondance r-chunk est appropriée et que l'ajout de masque de permutation permet la réduction du taux de faux positif.

L'AIS décrit par **Kephart [54]** est parmi les premières tentatives d'utiliser les mécanismes du système immunitaire humain pour la détection d'intrusions. Le travail se focalise sur la définition d'un système qui est capable d'effectuer la détection et la réponse automatique à des virus informatiques.

---

[9] Hole se sont des chaînes dans l'ensemble de non soi qui ne seront pas couverts par les détecteurs de non soi valides.

[10] Le masque de permutation permet d'effectuer des changements sur les valeurs attribuées à un détecteur.

Le système proposé détecte les virus soit par l'utilisation de la correspondance floue avec les signatures préexistantes des virus ou par le contrôle de changement dans le système et les fichiers de données. Afin de réduire le taux de faux positif dans le système, si un virus détecté est suspect alors il est séduit par le système pour infecter un ensemble de programme de decoy[11]. Si ces programmes deviennent infectés alors le système est certain que le programme détecté est un virus. Dans ce cas un algorithme est utilisé pour extraire la signature de ce programme. Afin de réduire la propagation rapide des virus dans les réseaux, le système infecté contacte les systèmes voisins pour communiquer la base de données de signatures.

Gonzalez et al [53] propose une nouvelle technique pour la détection d'anomalies qui utilise au départ les échantillons positifs pour générer des échantillons anormaux. Ensuite, les échantillons positifs et négatifs seront utilisés par un algorithme de classification. Cette approche hybride est ensuite comparée avec une technique de la détection d'anomalies qui utilise la carte auto organisatrice[12] (SOM) pour classifier un ensemble de données. Ce travail explore les problèmes liés à l'incrémentabilité, les détecteurs de valeurs réelles contre les détecteurs de valeurs binaires et la distinction floue entre le soi et le non soi.

L'approche proposée utilise l'algorithme de la sélection négative basé sur des valeurs réelles dont ces caractéristiques principales sont :

- L'espace de soi /non soi correspond à un sous-ensemble R^n.
- Un détecteur est défini par un vecteur n dimensionnel qui peut être vu comme une boule dans l'espace R^n avec un rayon r.
- L'entrée de l'algorithme est un ensemble d'échantillon de soi modélisé par des points n dimensionnel afin de produire un ensemble de détecteur pour couvrir l'ensemble de non soi.

[11] Les programmes de decoy se sont des programmes dont la seule fonction est de devenir infectés.

[12] Self-Organizing Map est un type de réseau de neurones qui utilise l'apprentissage compétitif, il fournit la capacité de représenter des données multidimensionnelles dans un espace de un ou deux dimensions.

- Ø Un détecteur (d) correspond au soi si la distance médiane des k plus proches voisins de détecteur d dans l'ensemble de soi est inférieure de r.
- Ø Un détecteur sera remplacé par un nouveau détecteur généré aléatoirement si l'âge de ce détecteur atteint une certaine valeur.

Ces détecteurs seront utilisés pour produire des échantillons anormaux (de non soi). Ensuite, ils ont utilisé un perceptron multicouche avec des propagations en arrière pour apprendre à distinguer entre le soi et le non soi. Ce dernier correspond à la fonction de détection d'anomalies qui sera utilisée pendant la phase de test pour classifier des nouveaux échantillons comme soi ou non soi. A la fin de ce travail, Gonzalez propose l'utilisation des réseaux immunitaires à la place des réseaux de neurones comme un futur travail.

Aickelin et al [45] propose l'incorporation de la théorie de danger dans les techniques de détection d'intrusions pour construire un système de détection d'intrusions capable de répondre d'une manière efficace aux menaces connues ainsi qu' aux nouvelles attaques avec des taux réduits de faux positif. Ils souhaitent la construction d'un modèle de calcul pour la théorie de danger avec l'intégration des signaux d'alarme.

Cette théorie qui montre la capacité du système immunitaire humain à repérer les signaux de danger pour répondre aux menaces en se basant sur la corrélation de ces signaux. Aickelin et al propose d'utiliser le même concept pour construire un IDS basé théorie de danger pour traiter le problème de corrélation d'alerte dont les signaux seront rassemblés à partir des hôtes et de réseaux. Ces signaux sont corrélés avec des alertes qui peuvent être de deux types : apoptose ou nécrose qui sont définies en parallèle de la mort cellulaire de type apoptose et nécrose respectivement.

Les alertes apoptoses correspondent aux alertes nuisibles de bas niveau, ces alertes seules ne signifient aucun mauvais comportement, mais elles sont souvent les prés-requis d'une attaque. Tandis que les alertes nécroses sont produites lors des attaques sérieuses où ils existent réellement des dommages dans le système. Aickelin et al espèrent la corrélation de ces alertes pour trouver les scénarios d'intrusions.

Greensmith et al [56] ont proposé un nouvel algorithme inspiré d'immunologie pour la détection d'intrusions qui se base sur l'abstraction du fonctionnement des cellules dendritiques. Ce travail est apparu après le projet souligné par Aickelin et al [45] qui propose

l'application de la théorie de danger aux systèmes de détection d'intrusions. Ce travail utilise les cellules dendritiques qui sont des cellules de présentation d'antigène qui ont la capacité d'agir comme un détecteur d'anomalies.

L'algorithme proposé est une abstraction de fonctionnement de base et des chemins de différenciation des cellules dendritiques qui sont responsables de la traduction de l'information sur la santé du tissu au système immunitaire adaptatif. Cela est réalisé par la combinaison des différents signaux qui apparaissent dans le système. Initialement, les cellules dendritiques se trouvent dans un état immature dont le rôle principal est la collection des antigènes, puis selon la combinaison des différents signaux de danger présents dans le système, ces cellules deviennent soient des cellules matures ou bien des cellules semi matures. Par exemple, en présence de signaux d'alarme dus aux dommages de l'hôte, les cellules dendritiques changent leur état en des cellules prétendues matures et présentent les protéines rassemblées du tissu endommagé avec une information de contexte indiquant que le tissu est endommagé.

De la même façon, Greensmith et al ont défini les signaux qui peuvent initier des changements dans les cellules dendritiques artificielles. Ensuite, ils ont appliqué ce principe à un ensemble de données, ils ont comparé le nombre de cellules dendritiques matures contre le nombre de cellules dendritiques semi matures. Un grand nombre de cellules dendritiques matures indique l'existence d'une anomalie dans le système. Les auteurs de ce travail proposent l'utilisation de l'algorithme comme un composant de base dans les systèmes immunitaires artificiels. A la fin de ce travail, les auteurs espèrent l'amélioration de l'algorithme par l'intégration des cytokines inflammatoires qui ont des effets différents sur les signaux d'alarmes ainsi l'incorporation d'autres composants immunitaires comme les cellules T dont le but principal est l'amélioration des performances de l'algorithme.

V.3 L'APPROCHE PROPOSEE

V.3.1 Discussion

Comme nous avons vu dans la section précédente, les différents travaux dans le domaine de la détection d'intrusions ont exploité les différents modèles du système immunitaire artificiel, le tableau ci-dessus (Figure 5.2) récapitule les différents travaux avec les différents modèles immunitaires utilisés qui sont :

- *La bibliothèque de gènes* signifie que le système implémenté utilise une méthode évolutionnaire pour initialiser les génotypes des détecteurs et non pas d'une façon aléatoire.
- *La sélection clonale* qui correspond au processus de prolifération et différenciation des cellules B pour incrémenter la généralité et la couverture des détecteurs par le processus d'hypermutation.
- *La mémoire immunitaire* permet l'apparition de la réponse secondaire afin d'obtenir une réponse plus rapide et plus efficace à une attaque déjà connue.
- *le réseau idiotypique* correspond à l'implémentation de la théorie du réseau idiotypique qui propose qu'il existe une interaction entre les différents composants immunitaires.
- *Le modèle de soi et non soi* qui permet au système de reconnaître ce qui est normal ou ce qui lui appartient afin de détecter le non soi.

La réponse dans ce contexte ne signifie pas tout simplement la génération d'une alerte, mais l'implémentation d'un changement dans le système en fonction du résultat de détection.

Modèles / **Travaux**	Soi / non-soi	Bibliothè-que de gènes	Sélection négative	Sélection clonale	Mémoire immunitaire	Réseau idiotypique	Théorie de danger	**Réponses**
Kephart [54]	×							×
Forrest [31]	×							
Hofmeyr [32]	×							
Hofmeyer [35]	×		×		×			
Balthrop [51]	×		×					
Gonzalez [53]	×		×					
Dasgupta [52]	×		×	×				
Greensmith [56]							*	

Figure 5.2 : Un récapitulatif des différents travaux dans le domaine des IDS

L'analyse prudente de ces travaux permet de tirer les remarques suivantes [46] :

- Ø Le modèle de soi / non soi est le modèle dominant parce qu'il est adopté par les différents travaux proposés.
- Ø Les systèmes proposés n'ont pas encore exploité les modèles suivants :
  1. Le réseau immunitaire idiotypique.
  2. La bibliothèque de gènes.
- Ø L'utilisation de l'algorithme de la sélection négative défini par Stéphanie Forrest et son groupe comme un algorithme de base pour générer les détecteurs de non soi.
- Ø La proposition d'un nouveau modèle immunitaire qui se base sur la théorie de danger par Greensmith et al [56]. Ce modèle se base sur l'abstraction du fonctionnement des cellules dendritiques et qui intègre la notion de danger.

A la fin de cette section, nous pouvons constater que les systèmes de détection d'intrusions basés sur les systèmes immunitaires artificiels ont encore des points à explorer. Ils peuvent par exemple adopter des concepts et des aspects plus larges inspirés d'immunologie comme par exemple : la théorie de danger, la bibliothèque de gènes et le réseau idiotypique.

V.3.2 Problématique

La réalisation de la sécurité des systèmes et des réseaux reste une grande épreuve, de nombreuses solutions ont été proposées pour prévenir les différents types d'intrusions. Ces intrusions qui peuvent être non seulement produites par des éléments externes du système mais aussi par ses utilisateurs internes qui essayent d'exploiter des droits d'accès autorisés pour effectuer des opérations non autorisées ou bien d'obtenir des droits supplémentaires non autorisés. Cependant, les différents algorithmes proposés dans le domaine de détection d'intrusions qui sont basés sur les systèmes immunitaires artificiels se basent principalement sur le modèle de soi et de non soi qui est le modèle le plus populaire.

L'objectif de ces différents algorithmes consiste à minimiser le taux de vrai négatif et maximiser le taux de vrai positif. Le taux de vrai négatif reflète le taux d'erreurs c'est-à-dire le cas où le système détecte un élément de soi comme un intrus par contre le taux de vrai positif présente la détection réelle d'un élément de non soi. Cependant, dans un environnement réel, une intrusion n'est pas seulement générée par des éléments externes du système (des éléments de non soi) mais aussi elle peut être produite par des éléments internes du système (les éléments de soi).

En plus, l'environnement du réseau réel change constamment, il est difficile de caractériser d'une manière complète les comportements normaux dans le système. Pour cette raison, les algorithmes proposés optent l'étape de la costimulation afin de permettre la détection des intrusions réelles. Le signal de costimulation est déclenché quand un détecteur détecte un élément. Elle nécessite l'intervention d'un opérateur humain pour indiquer si l'élément détecté désigne réellement une intrusion. Dans cette situation, l'opérateur de sécurité envoie un signal de confirmation et par conséquence le détecteur correspondant sera activé afin de supprimer cet intrus, par contre si l'élément détecté ne constitue aucune intrusion alors l'officier de sécurité n'envoie aucun signal et par conséquence le détecteur correspondant sera supprimé. Cette propriété permet la satisfaction de l'objectif des algorithmes de détection d'intrusions qui consiste à maximiser le taux de vrai positif et minimiser le taux de vrai négatif. Cependant, la satisfaction d'un taux bas de vrai négatif requiert un nombre élevé de costimulation qui signifie que le système ne peut pas s'adapter seul pour assurer les meilleurs taux de détection.

V.3.3 Proposition

Avec l'apparition de la théorie de danger qui propose des nouvelles idées intéressantes qui peuvent surmonter les problèmes liés à l'adoption du modèle de soi / non soi classique. Pour Cette théorie, ce n'est plus le caractère étranger d'un élément qui déclenchera une réponse immunitaire mais c'est le caractère dangereux de l'élément. Bien que la théorie de danger défie le modèle de soi / non soi mais elle ne nie pas l'existence de la discrimination entre le soi et le non soi, elle signale qu'il y a d'autres facteurs contribuant impliqués dans l'initiation d'une réponse immunitaire [44]. Cette théorie qui permet de contrôler les envahisseurs dans le système qui peuvent être de « non soi mais inoffensifs » et de « soi mais nuisibles ». Nous désirons dans ce travail d'intégrer des idées inspirées de cette nouvelle théorie afin de surmonter les problèmes liés au modèle de soi / non soi.

Ainsi, le travail proposé par Greensmith est al [56] qui a introduit les cellules dendritiques comme un détecteur d'anomalies ouvre des nouveaux horizons à exploiter. Ce travail qui va nous permettre d'établir la liaison entre le système immunitaire adaptatif et le système immunitaire inné. Ces cellules qui ont des effets différents sur la réponse immunitaire des cellules T selon l'information de contexte présentée avec l'antigène, qui reflète l'état du tissu cellulaire. Cette information qui peut indiquer soit que l'élément détecté est dangereux parce

qu'il a occasionné des dégâts dans le tissu ce qui implique l'activation des cellules T, ou bien que l'élément détecté n'a rien produit dans le tissu ce qui implique la tolérance des cellules T à l'antigène présenté.

Nous voulons dans ce travail d'exploiter ce fonctionnement de base du système immunitaire naturel pour la détection d'intrusions afin de résoudre les problèmes décrits dans la section précédente. Et cela par l'ajout de quelques améliorations sur l'algorithme de la sélection négative qui se base sur le modèle de soi et de non soi. Cette amélioration consiste à intégrer la notion de danger pour permettre la détection des intrusions réelles causées par des utilisateurs internes ou externes qui endommagent le système. D'une manière générale, l'approche proposée tente de combiner entre deux éléments de base dans la détection des intrus qui sont les cellules T et les cellules dendritiques. Cette combinaison est tirée à partir du modèle de la théorie de danger proposé par Matzinger [44,56] et qui décrit l'interaction entre les cellules dendritiques et les cellules T.

Dans le contexte d'un IDS, nous supposons que les signaux de danger apparaissent après une attaque limitée afin de minimiser les dommages qui peuvent être apparus dans le système. Le traitement de ces signaux est effectué par les cellules dendritiques. Ensuite, ces cellules présentent les éléments détectés aux cellules T avec l'information de contexte qui indique la nature de l'élément détecté qui peut être dangereux ou non dangereux dont le but est la détection et l'élimination des éléments nuisibles.

V.4 DESCRIPTION DE L'ALGORITHME

La solution proposée combine entre deux éléments de base pour la détection des intrus qui sont les cellules T et les cellules dendritiques. Cela est réalisé via l'exploitation de la notion de danger liée à chaque élément présenté par les cellules dendritiques dont le but principal est de rendre l'algorithme de la sélection négative apte à détecter des intrusions réelles initialisées par les membres externes ainsi par les éléments internes. Avant de décrire les étapes essentielles de l'algorithme, nous souhaitons d'abord éclaircir ce fonctionnement dans le cas du système immunitaire humain.

V.4.1 Le système immunitaire humain

Comme nous avons présenté dans le chapitre qui expose le système immunitaire artificiel, le système immunitaire humain possède plusieurs éléments qui travaillent en collaboration pour

assurer la protection du corps humain. Un élément primordial est les cellules dendritiques qui sont des cellules de présentation d'antigènes, spécialisées dans la présentation des protéines rassemblées pendant leurs phases d'immaturation [56]. Chaque élément rassemblé sera présenté avec une information de contexte qui indique s'il est collecté dans un environnement serein ou dangereux. Quand les cellules dendritiques achèvent l'étape de maturation, elles peuvent présenter ces éléments au système immunitaire adaptatif et en particulier aux cellules T avec des signaux de contexte dont le but principal est l'activation des cellules T naïves. Les cellules T qui ont un récepteur complémentaire à l'antigène seront activées si le contexte de présentation est dangereux ou nécrotique. Cependant, si le contexte de présentation est apoptose, alors cela engendre la tolérance des cellules T qui peuvent détecter l'antigène.

V.4.2 Une vue générale de l'algorithme

Avec la même démarche dans le système immunitaire humain, l'algorithme proposé dans ce travail se base sur l'incorporation des interactions existantes entre les deux systèmes immunitaires inné et adaptatif. Ces interactions qui peuvent engendrer une activation des détecteurs si l'élément présenté par les cellules dendritiques est un élément dangereux ou bien la tolérance des détecteurs dans le cas contraire. En plus, ces interactions génèrent des changements considérables au niveau des éléments mémoires dont le but est l'obtention d'une mémoire immunitaire qui permet uniquement la détection des éléments dangereux.

V.5 PSEUDO CODE DE L'ALGORITHME

Le système proposé intègre les concepts de base du système de détection d'intrusions LYSIS basé sur les systèmes immunitaires artificiels proposé par Hofmeyr [35,36] qui a amélioré l'algorithme de la sélection négative par l'ajout des propriétés immunitaires supplémentaires pour lui permettre de s'adapter aux changements de l'environnement. Ainsi, l'utilisation de l'algorithme des cellules dendritiques proposé par Greensmith et son groupe [56].

Dans ce pseudo code (figure 5.3) , on suppose que ***détectable_mature (élément)*** et ***détectable_mémoire(élément)*** sont des fonctions prédéfinies qui retournent la valeur vraie si l'élément en paramètre est détectable par la population des détecteurs matures ou bien par la population des détecteurs mémoires respectivement. L'ensemble d'antigènes qui sera contrôlé par la population des détecteurs mémoires est un ensemble de chaînes binaires de longueurs 36, et qui sera présenté dans les sections suivantes.

```
Début

// Initialisation des paramètres
  Nombre de génération = n ;
 Période de tolérance = T ;
 Seuil d'activation = A ;
 Durée de vie = L ;
taille_population non mémoire = M ;
 Génération de la population des détecteurs immatures initiaux dont l'âge des
 détecteurs est initialisé à zéro ;
  N = 1 ;
// Répéter cette boucle un nombre de génération
Tant que (N <= Nombre de génération) Faire
 Début
    Représentation de P éléments par les cellules dendritiques accompagnés avec
    leur information de contexte ;
    Pour chaque (élément présenté) Faire
     Début
       Si (contexte (élément) = "dangereux") Alors
        Début
       //Vérifier si l'élément n'est pas détectable par la population des détecteurs
         mémoires ;
           Si (détectable_mémoire (élément)) Alors
              Supprimer cette protéine de la population des éléments présentés par
             les cellules dendritiques ;
           Sinon
              //Vérifier s'il est détectable par la population des détecteurs matures ;
            Si (détectable_mature (élément)) Alors
              Début
                 Le détecteur correspondant deviendra un détecteur mémoire ;
                Supprimer cette protéine de la population des éléments présentés par les
                 cellules dendritiques ;
               Fin
            Sinon
              Garder cette protéine pour une nouvelle vérification avec de nouveaux
              détecteurs ;
            Fin si
         Fin si
       Fin
      Sinon /* contexte (élément) = " non dangereux"*/

       Début
          //Vérifier si l'élément est détectable par la population des détecteurs mémoires
         Si (détectable_mémoire (élément))) Alors
              Supprimer le détecteur mémoire correspondant ;
       Fin
      Fin si
    Fin pour chaque
```

```
// Contrôler l'ensemble d'antigènes par la population des détecteurs mémoires
    Pour chaque (détecteur mémoire) Faire
     Début
        Vérifier si le détecteur mémoire détecte un élément dangereux ;
        Vérifier si le détecteur mémoire détecte un élément non dangereux ;
        Si (l'antigène est détecté) Alors Supprimer l'antigène détecté de l'ensemble
          d'antigène ;
    Fin pour chaque
 // Contrôler l'ensemble d'antigènes par la population des détecteurs matures
    Pour chaque (détecteur mature) Faire
     Début
        Si le détecteur mature détecte un élément de non soi alors augmenter le nombre de
        correspondances du détecteur ;
         // Génération de nouveaux détecteurs mémoires
        Si le nombre de correspondances du détecteur est supérieur au Seuil d'activation
        alors ce détecteur mature devient un détecteur mémoire ;
        Si l'âge du détecteur mature dépasse la Durée de vie alors supprimer ce détecteur
        mature ;
    Fin pour chaque
 // Contrôler l'ensemble d'antigènes de soi par la population des détecteurs immatures
    Pour chaque (détecteur immature) Faire
     Début
       Vérifier si le détecteur immature détecte un antigène de soi ;
        Supprimer le détecteur immature qui détecte l'antigène de soi ;
        // Génération de nouveaux détecteurs matures
        Si ce détecteur immature dépasse la Période de tolérance alors ce détecteur
        immature devient un détecteur mature ;
    Fin pour chaque

// Génération de nouveaux détecteurs immatures
  Si ((taille (population cellule mature) + taille (population cellule immature)) <
      taille_population non mémoire) Alors
    Répéter
       Début
          Générer des nouveaux détecteurs immatures ;
          Ajouter ces détecteurs à la population des détecteurs immatures ;
       Fin
    Jusqu'à ((taille (population cellule mature) + taille (population cellule immature))
            < taille_population non mémoire)

   // Augmenter les paramètres
        N = N +1 ;
        Augmenter l'âge des détecteurs immatures ;
       Augmenter l'âge des détecteurs matures ;

  Fin tant que
Fin
```

Figure 5.3 : Pseudo code de l'algorithme

Le pseudo code de l'algorithme présenté au dessus est décrit de la manière suivante :

A chaque génération, les cellules dendritiques présentent un ensemble d'élément de taille prédéfinie choisi aléatoirement à partir de l'ensemble d'antigènes. En fonction du contexte de l'élément présenté un nombre d'opération sera établi afin de permettre à la population des détecteurs mémoires la détection des éléments intrusifs. Si le contexte de l'élément est dangereux alors l'algorithme vérifie si cet élément est détectable par la population des détecteurs mémoires et par conséquence seulement cette protéine sera supprimée de l'ensemble d'éléments présenté. Mais, si l'élément dangereux n'est pas détectable par la population des détecteurs mémoires alors l'algorithme vérifie s'il y'a un détecteur qui peut détecter cet élément dans la population des détecteurs matures. Si tel détecteur existe alors il sera ajouté à la population des détecteurs mémoires et la protéine correspondante sera éventuellement supprimée de l'ensemble des éléments présentés. Cependant, s'il n'existe aucun détecteur mature qui permet l'identification de cette protéine alors elle sera sauvegardée pour une nouvelle recherche dans les générations suivantes. Dans le cas où l'élément présenté est un élément non dangereux alors l'algorithme vérifie s'il est détectable par la population des détecteurs mémoires afin de supprimer le détecteur correspondant.

Pendant la phase de contrôle effectuée par la population des détecteurs mémoires, chaque élément détecté est considéré comme une intrusion. En plus, deux vérifications sur la nature de détection seront établies et qui seront utilisées par la suite pour vérifier les performances de l'algorithme, qui sont :

- Vérifier si l'élément détecté est un élément dangereux afin de compter le taux de détection des éléments dangereux pour définir le ***taux de détection vrai positif***.
- Vérifier si l'élément détecté est un élément non dangereux afin de recenser le taux d'erreurs effectué par l'algorithme pour définir le ***taux de détection vrai négatif***.

Les éléments de l'ensemble d'antigènes restant qui ne sont pas détectés par la population des détecteurs mémoires seront utilisés dans la phase de contrôle effectuée par la population des détecteurs matures dont le but principal est la génération de nouveaux détecteurs mémoires capables de détecter ces éléments dans le futur. Ainsi, les éléments de l'ensemble de soi seront exploités pour générer des nouveaux détecteurs immatures. Pour éclaircir ces différentes étapes de l'algorithme, la section suivante sera consacrée à présenter d'une manière détaillée les différents mécanismes utilisés ainsi le processus de génération des différentes populations des détecteurs.

V.6 LES DIFFERENTS MECANISMES MIS EN ŒUVRE

L'algorithme présenté intègre plusieurs modèles du système immunitaire artificiel, qui sont :

- L'algorithme de la sélection négative qui est utilisé afin de générer des détecteurs tolérants au soi.
- La sélection clonale qui est intégrée pendant le processus de génération des détecteurs initiaux de telle sorte que les détecteurs qui assurent le plus grand nombre de détections seront choisis pour produire des nouveaux détecteurs.
- La mémoire immunitaire qui permet d'avoir une réponse secondaire rapide.
- Le modèle des cellules dendritiques qui permet de présenter les protéines rassemblées avec une information de contexte afin de générer des changements constants sur la population des détecteurs mémoires selon le contexte de ces éléments.

Afin d'illustrer la manière d'utilisation de ces différents modèles immunitaires dans l'algorithme proposé, la section suivante sera consacrée à présenter la manière d'utilisation de ces modèles.

V.6.1 La génération de la population des détecteurs initiaux

Le système immunitaire est capable de reconnaître un nombre important de pathogènes par un petit nombre d'anticorps. Nous avons utilisé le modèle proposé par Smith et al [57] qui se base sur le processus de la sélection clonale du système immunitaire humain et qui intègre une technique de niche dont le but principal est la génération des solutions diverses et coopératives. Ce modèle utilise les algorithmes génétiques ainsi une stratégie de niche pour garantir la possibilité de détecter les modèles communs d'antigènes présentés d'une façon aléatoire ainsi maintenir et discerner la population d'antigènes.

V.6.1.1 Calcul de la fonction d'évaluation

Ce modèle propose un ensemble d'étapes pour calculer les valeurs de fitness des différents détecteurs générés initialement d'une façon aléatoire, qui sont résumées par les étapes suivantes :

1. Les valeurs de fitness des différents détecteurs sont initialisées par zéro.
2. Un échantillon de détecteur de taille D est sélectionné d'une façon aléatoire à partir de la population de détecteurs.

3. un échantillon d'antigènes est sélectionné aussi d'une façon aléatoire à partir de la population des antigènes (les éléments de non soi).
4. Le calcul du nombre de correspondances entre les éléments de l'échantillon des détecteurs et chaque antigène de l'échantillon des antigènes via la règle de correspondance r bits contigus.
5. L'incrémentation de la valeur de fitness du détecteur qui possède le plus grand nombre de détections par cette valeur, tandis que les valeurs de fitness des autres détecteurs restent inchangeables. Si plusieurs détecteurs possèdent la valeur la plus grande du nombre de correspondances, alors cette valeur est divisée par le nombre de ces détecteurs et les valeurs de fitness de ces détecteurs seront incrémentées par le résultat de cette division.
6. Répéter les étapes 2 à 5 trois fois le nombre de détecteur.

V.6.1.2 Paramètres de génération

Initialement, nous avons généré d'une manière aléatoire une population des détecteurs immatures de taille 100, qui sont sous la forme des chaînes binaires de longueur 36. La longueur des détecteurs est choisie en fonction de la taille des éléments utilisés dans la phase de test. Le modèle décrit à la section précédente est appliqué pour affecter une valeur d'évaluation aux différents détecteurs de la population par la sélection aléatoire d'un échantillon de détecteurs de taille 10 à chaque itération ainsi la sélection aléatoire d'un seul antigène à partir de l'ensemble de non soi. Ce processus d'évaluation est répété un nombre d'itérations. La règle d'appariement entre un détecteur et un antigène est la règle de r bits contigus avec un r égal à 12. Ensuite, les différentes étapes des algorithmes génétiques seront appliquées sur la population de détecteurs initiaux. La sélection d'un meilleur individu est établie via la sélection par la roulette de wheel. En ce qui concerne l'opérateur de croisement, deux individus sont sélectionnés d'une façon aléatoire et un croisement simple est effectué sur un seul point aléatoirement choisi. Les deux fils générés subissent à l'opérateur de mutation avec un taux de mutation égal à 0,1. Ainsi ce processus de sélection et application des opérateurs génétiques est répété un nombre d'itérations.

La nouvelle population générée est jugée par le même principe d'évaluation de fitness décrit précédemment afin de remplacer 20% des éléments mauvais de la population initiale par les meilleurs éléments de la population obtenue. Ce processus de génération et de remplacement

des détecteurs est répété un nombre d'itérations pour obtenir une population des détecteurs immatures initiaux qui sera utilisée dans l'algorithme principal.

V.6.2 La sélection négative et la mémoire immunitaire

Le processus de génération de la population des détecteurs se base sur l'algorithme de la sélection négative qui permet la génération des détecteurs tolérants au soi. Ce travail se base sur le système de détection d'intrusions LYSIS basé sur les systèmes immunitaires artificiels proposé par Hofmeyr [35,36] par l'exploitation de certains concepts de base de ce système tel que : la période de tolérance, le seuil d'activation, etc. Le processus de génération des détecteurs suit le cycle de vie des détecteurs proposé par Hofmeyr où la population des détecteurs est divisée selon trois sous populations, qui sont :

- La population des détecteurs immatures qui sont générés initialement d'une façon aléatoire, ces détecteurs doivent subir à la phase de la sélection négative pour éliminer les détecteurs qui détectent le soi.
- La population des détecteurs matures qui doivent expérimenter l'ensemble d'antigènes présenté dans chaque génération jusqu'à l'obtention d'un certain degré d'affinité.
- La population des détecteurs mémoires qui présente la mémoire immunitaire. Ces détecteurs mémoires effectuent le contrôle du système dont chaque détection est considérée comme une intrusion.

Ce processus est caractérisé par l'intégration de certains nombres de paramètres afin d'assurer la meilleure tolérance aux soi, ainsi l'adaptation aux changements constants de l'environnement, à savoir :

- ***La période de tolérance*** qui est la période pendant laquelle le détecteur immature généré subit à l'algorithme de la sélection négative.
- ***Le seuil d'activation*** qui détermine le nombre de correspondances qu'un détecteur doit réaliser pour qu'il soit activé.
- ***La durée de vie*** qui indique l'âge maximal des détecteurs matures afin de rafraîchir la tolérance au soi obtenue. Si l'âge du détecteur mature dépasse cette durée de vie, il sera éliminé et remplacé par un nouveau détecteur généré aléatoirement.

Le processus de génération des différentes populations des détecteurs est accompli de la façon suivante : les détecteurs immatures initiaux générés en fonction de l'ensemble d'antigènes selon l'étape décrite dans la section précédente subissent à la phase de la

sélection négative contre l'ensemble de soi pour générer un ensemble de détecteurs immatures tolérants au soi de telle sorte que les détecteurs qui détectent un élément de soi seront éliminés. Les détecteurs qui survivent et dépassent *la période de tolérance* prédéfinie deviennent des détecteurs matures. Ces derniers contrôlent et expérimentent l'ensemble d'antigènes présenté dont chaque appariement implique l'incrémentation du nombre de correspondances du détecteur correspondant.

Les détecteurs matures qui ont un nombre de correspondances supérieur au *seuil d'activation* évoluent vers des détecteurs mémoires, par contre, les autres détecteurs restent dans l'état mature jusqu'à l'obtention d'une certaine affinité qui est définie par le seuil d'activation. On note que les détecteurs matures et immatures ont un âge qui sera incrémenté de un après chaque génération, si l'âge des détecteurs dépasse la *durée de vie* prédéfinie alors le détecteur correspondant sera éliminé. La population des détecteurs mémoires générée continuellement en fonction de l'ensemble des antigènes effectue le contrôle dont chaque détection signalée est considérée comme une intrusion.

Cependant, dans ce travail nous n'avons pas intégré l'étape de costimulation qui sert à la construction d'une population des détecteurs mémoires qui ne détectent que des intrusions. Cette étape qui est réalisée d'une manière explicite dans le cycle de vie proposé par Hofmeyr est assurée d'une manière implicite dans cet algorithme soit par l'opération de suppression réalisée sur la population des détecteurs mémoires après la présentation d'un élément non dangereux détectable par les détecteurs mémoires, ou bien par l'opération d'activation qui permet l'ajout d'un nouveau détecteur mémoire permettant la détection des éléments dangereux.

La figure suivante résume les différentes étapes suivies durant le processus de génération des différents détecteurs.

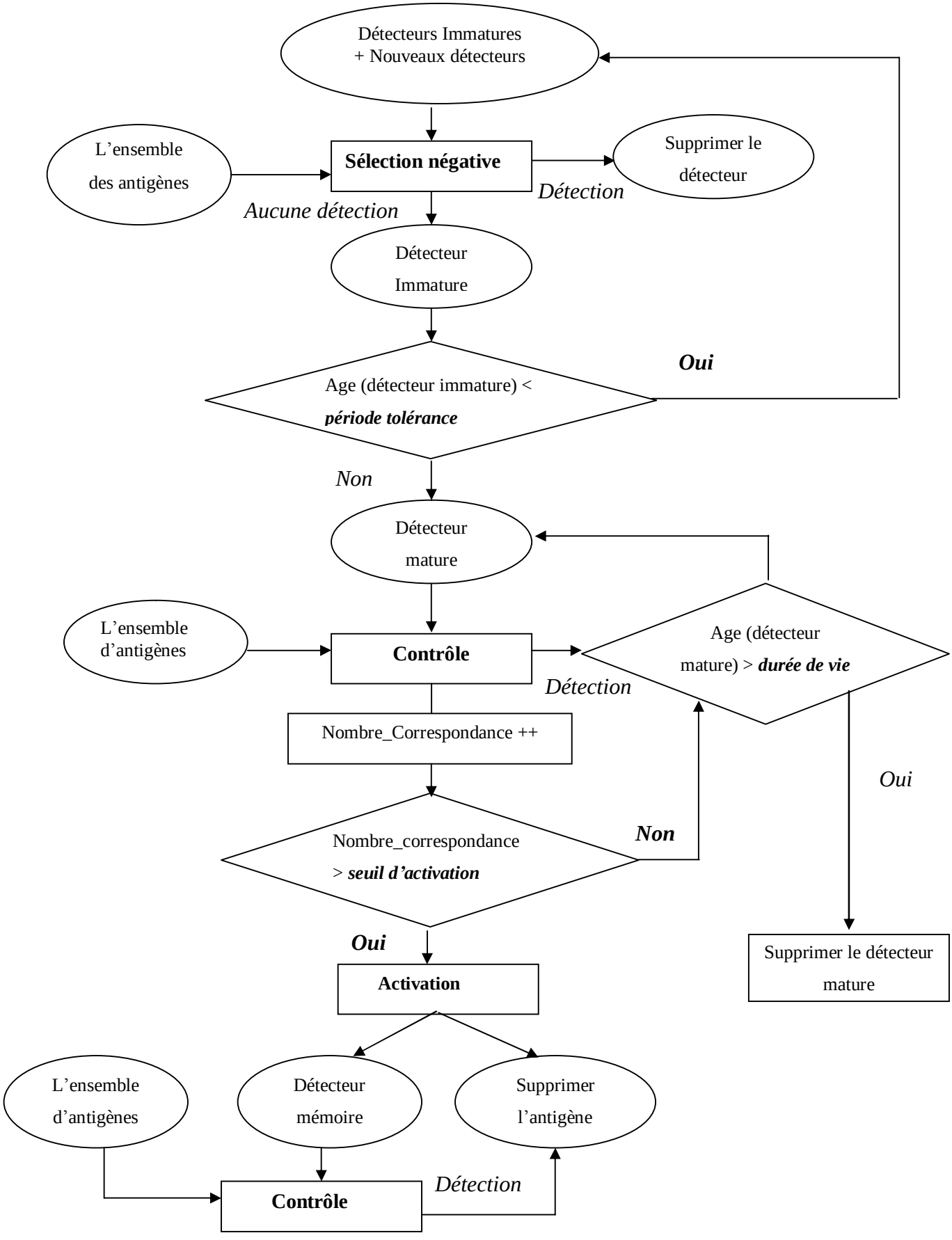


Figure 5.4 : Le processus de génération des détecteurs

V.6.3 Le modèle des cellules dendritiques

Les cellules dendritiques sont des cellules spécialisées dans la présentation des protéines (antigènes) collectées avec leur contexte environnemental. Cette information est présentée aux cellules T et génère l'identification et la suppression des pathogènes. L'information de contexte est traduite par l'utilisation des différents états de maturation des cellules dendritiques qui peuvent exister selon trois états : immatures, matures ou bien semi matures. Initialement, les cellules dendritiques se trouvent dans un état immature dont le rôle principal est la collection des antigènes afin de préparer leur présentation aux cellules T naïves.

Cependant, si le tissu est endommagé suite à des infections engendrées par des antigènes ou l'existence des cellules de stresse alors les signaux de danger apparaissent dans le tissu. L'exposition de cellules dendritiques aux signaux de danger implique la maturation complète des cellules dendritiques immatures. Dans le cas contraire, si le tissu est sein et il n'existe pas des cellules de stresse, suite à la mort cellulaire apoptose qui implique la semi maturation des cellules dendritiques immatures (voir Annexe A).

D'une manière générale, les cellules dendritiques immatures combinent entre les différents signaux de danger (les signaux apoptoses, les signaux de danger, etc.) afin de présenter les antigènes rassemblés avec une information de contexte qui reflète l'état du tissu cellulaire.

Dans le contexte d'un IDS, les signaux apoptoses indiquent l'état normal du système surveillé tandis que les signaux nécroses ou de danger indiquent des changements importants dans l'environnement par exemple : l'utilisation du CPU ou la taille mémoire utilisée, etc.

Afin de vérifier les conséquences des effets d'activation et de suppression engendrés par les cellules dendritiques sur la population des détecteurs mémoires. Dans ce travail, nous avons supposé qu'à chaque itération de l'algorithme, les cellules dendritiques présentent un échantillon d'éléments choisi d'une manière aléatoire de taille prédéfinie, dont chaque élément est présenté avec une information de contexte.

V.7 PARAMETRES ET ENSEMBLE DE DONNEES

V.7.1 Description de l'ensemble de données

Nous avons utilisé le standard UCI qui comporte un ensemble de données sur des cellules cancéreuses « Wisconsin Breast cancer data set »[13] [58]. Cette base de données contient 699 exemples qui sont divisés selon 2 classes : « maligne » et « bénigne » de la manière suivante : 241 exemples appartient à la classe « maligne » et le reste c'est-à-dire 458 appartient à la classe « bénigne ».

Chaque élément de cet ensemble de données possède 9 attributs continus pour représenter les différentes caractéristiques des cellules cancéreuses. Ainsi, les éléments de cette base de données ont un dixième attribut qui désigne l'appartenance de l'élément à la classe des cellules bénignes ou à la classe des cellules malignes. Cet ensemble de données contient 16 éléments dont la valeur d'un attribut est inconnue.

V.7.1.1 Hypothèse

Afin d'exploiter cette base de données dans le test de cette application, nous considérons les hypothèses suivantes :

- Ø Les éléments de la classe bénigne sont considérés comme des éléments de soi tandis que les éléments de la classe maligne sont considérés comme des éléments de non soi.
- Ø Les valeurs manquantes des attributs de cette base de données sont remplies par des valeurs générées aléatoirement.

Nous avons ajouté une autre étiquette à chaque attribut de l'ensemble de données pour différencier entre les éléments qui sont « dangereux » et les éléments « non dangereux » de telle sorte que nous avons divisé les éléments de soi et de non soi en deux sous-ensembles pour permettre la présentation de ces éléments avec une information de contexte s'ils sont sélectionnés par les cellules dendritiques.

[13] ftp://ftp.ics.uci.edu/pub/machine-learning-databases.

V.7.1.2 La codification de données

Les éléments de cet ensemble de données sont composés de 9 attributs dont chacun peut avoir une valeur entre le 1 et le 10. Les attributs de chaque élément de cet ensemble de données sont représentés par une valeur binaire de 4 bits de telle sorte que chaque élément de l'ensemble de données sera représenté par des chaînes binaires de longueur 36. Ces chaînes binaires obtenues définissent l'ensemble d'antigènes utilisé dans l'algorithme.

V.7.2 La règle de correspondance

La règle de correspondance utilisée pour vérifier l'appariement entre un élément quelconque et un élément de la population des détecteurs est la règle de r bits contigus. Un détecteur correspond à un antigène selon la règle de r bits contigus s'ils ont au moins r bits contigus en commun. Selon la figure 5.5, dans le cas où r = 4, il y'a une correspondance entre les deux chaînes dans le premier cas seulement.

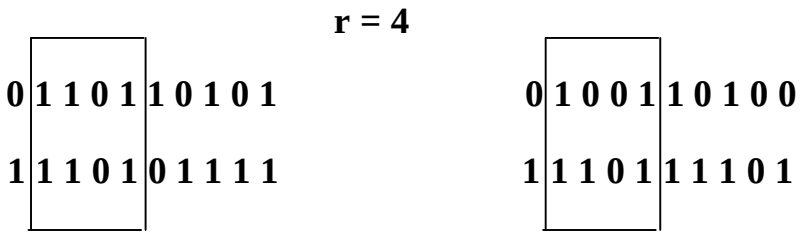


Figure 5.5 : L'appariement selon la règle de r bits contigus

Le choix de la valeur de r est très important puisque il a une grande influence sur la qualité et le nombre des détecteurs générés. Une valeur très grande mène à obtenir des détecteurs spécialisés qui peuvent détecter un nombre limité d'éléments et par conséquence la génération d'un nombre élevé des détecteurs pour garantir la détection de tous les éléments. Par contre si la valeur de r est petite, cela permet d'avoir des détecteurs généralisés en détectant plus d'éléments et le résultat est un ensemble de détecteurs de taille inférieure.

V.8 RESULTATS EXPERIMENTAUX ET DISCUSSIONS

Afin de montrer les performances de cet algorithme et la possibilité de détecter les éléments dangereux ainsi la tolérance aux éléments non dangereux qui ne représentent pas des

intrusions, l'algorithme décrit auparavant sera exécuté selon plusieurs expériences dont le but principal est la vérification des taux de détection vrai positif et vrai négatif produits par la population des détecteurs mémoires durant leur phase de contrôle.

Au contraire des autres travaux qui considèrent le taux de détection vrai positif comme le taux de détection des éléments de non soi et le taux de détection vrai négatif comme le taux de détection des éléments de soi. Dans ces expériences, nous avons considéré que *le taux de détection vrai positif* représente la détection des éléments *dangereux* représentant des intrusions réelles et qui peuvent être engendrées par des éléments de soi ou de non soi. Tandis que *le taux de détection vrai négatif* présente le cas de détection des éléments *non dangereux* et qui peuvent ainsi être des éléments de soi ou de non soi.

V.8.1 Expérience 1

V.8.1.1 Description

Une fois que l'ensemble des détecteurs est généré selon le processus décrit précédemment, l'algorithme proposé est exécuté 1000 itérations. Pendant chaque génération, l'ensemble de données décrit précédemment qui est constitué de chaînes binaires de longueur 36 sera utilisé pour définir l'ensemble des antigènes ainsi les cellules dendritiques présentent 20 éléments choisis d'une manière aléatoire accompagnés avec leur information de contexte.

Puisque l'ensemble de données utilisé dans le test est un ensemble statique, la période de tolérance des détecteurs est égale à 1, un détecteur mature sera activé s'il assure la détection de 4 éléments dont l'âge maximal ne dépasse pas 2.

L'importance de l'utilisation de ces paramètres n'est pas illustrée dans ces expériences car l'ensemble de données utilisé est statique. Le test de ce programme avec un ensemble de données dynamique est souhaitable comme un futur travail.

V.8.1.2 Discussion

Les résultats expérimentaux obtenus sont illustrés dans le graphe présenté dans la figure 5.6 dont l'axe des abscisses représente le nombre de génération et l'axe des ordonnées représente le taux de détection. Ce graphe possède deux lignes dont l'une pour représenter le taux de détection vrai positif (Taux TP) et l'autre pour représenter le taux de détection vrai négatif (Taux TN).

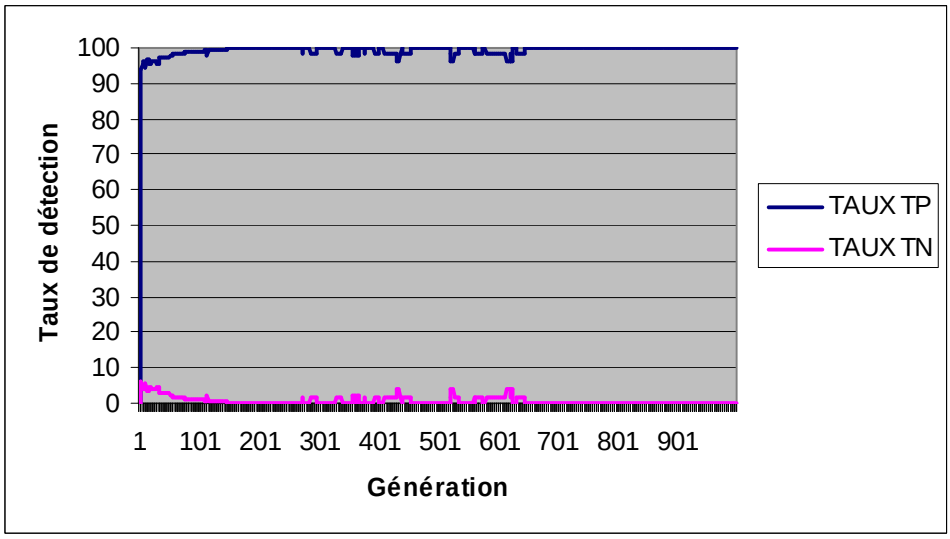


Figure 5.6 : Les taux de détection vrai positif (TP) et vrai négatif (TN)

Cette figure montre que le taux de détection vrai positif est élevé par rapport au taux de détection vrai négatif. Cependant, pendant les premières générations le taux de détection vrai négatif (Taux TN) qui représente la détection des éléments présent dans un contexte non dangereux est un petit peu élevé, ce taux peut atteindre jusqu'au 6,29% du total du taux de détection. Ensuite, les taux de détection se balancent entre une augmentation et une diminution jusqu'à l'arrivée à une stabilité avec un taux de détection vrai positif maximal et un taux de détection vrai négatif égal à 0%. Cette perturbation dans les taux de détection est la conséquence des effets d'activation et de suppression engendrés par les éléments présentés par les cellules dendritiques qui permettent de supprimer les détecteurs mémoires qui peuvent détecter des éléments non dangereux et qui ne représentent pas des intrusions réelles ce qui implique l'augmentation du taux vrai positif (Taux TP).

Afin de détailler les résultats obtenus en termes de la possibilité de détecter les éléments dangereux et la tolérance aux éléments non dangereux qui ne représentent pas des intrusions, les taux de détection vrai positif et vrai négatif seront détaillés en fonction de la nature des éléments détectés (soi ou de non soi). A partir de cette étape nous considérons les taux suivants :

- *Le taux vrai positif non soi* (Taux TP Non Soi) qui représente le taux de détection des éléments de non soi dangereux.
- *Le taux vrai positif soi* (Taux TP Soi) qui représente le taux de détection des éléments de soi dangereux.

- *Le taux vrai négatif non soi* (Taux TN Non Soi) qui représente le taux de détection des éléments de non soi non dangereux.
- *Le taux vrai négatif soi* (Taux TN Soi) qui représente le taux de détection des éléments de soi non dangereux.

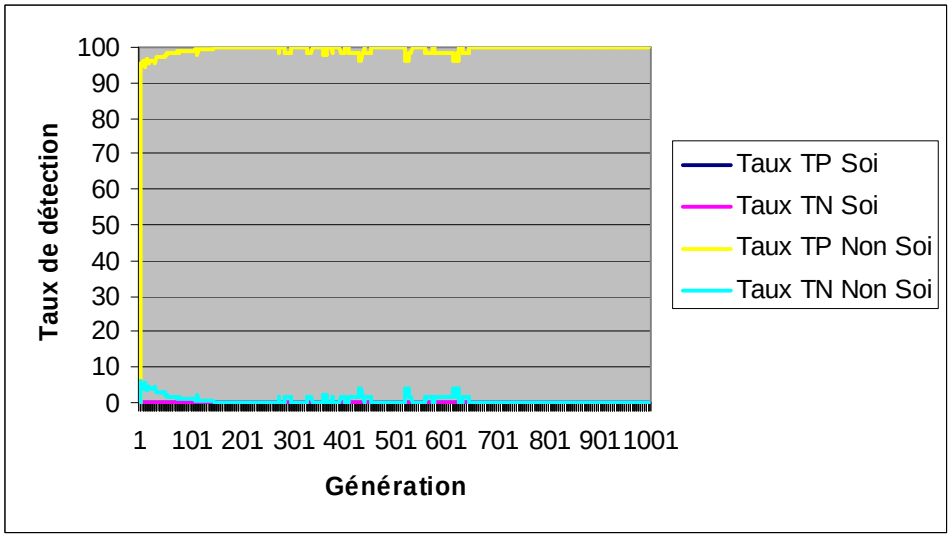


Figure 5.7 : Les taux de détection vrai positif et vrai négatif détaillés

La figure 5.7 illustre que le taux de détection vrai positif (Taux TP) obtenu dans la figure précédente représente seulement la détection des éléments de non soi dangereux (Taux TP Non Soi) et le taux de détection vrai négatif (Taux TN) représente aussi la détection des éléments de non soi dangereux (Taux TN Non Soi).

Ces résultats montrent la détection des éléments de non soi présentés dans un contexte dangereux avec la tolérance graduelle aux éléments de non soi non dangereux. Cependant, les éléments de soi dangereux ne sont jamais détectés par les détecteurs mémoires.

En effet, ces éléments de soi dangereux ne seront jamais détectés parce que les détecteurs générés par cet algorithme sont tolérants à ces éléments de soi (ils ont subi à la phase de la sélection négative contre cet ensemble de soi) alors il est impossible de trouver au fil des générations un détecteur mature qui sera activé pour localiser un élément de soi dangereux.

V.8.2 Expérience 2

V.8.2.1 Description

Afin de permettre la détection des éléments de soi présent dans un contexte dangereux, nous avons modifié l'algorithme précédent de telle sorte que si l'élément présenté par les cellules

dendritiques est dangereux et il est aussi un élément de soi alors s'il existe un détecteur immature qui correspond à cet antigène il sera activé. Cette activation est tolérée s'il n'existe aucun détecteur mémoire qui peut détecter cet élément. Les modifications apportées sur l'algorithme initial seront représentées dans la figure 5.8. On note que ***détectable_immature (élément)*** est une fonction prédéfinie qui retourne une valeur booléenne, qui indique si l'élément en paramètre est détectable ou non par la population des détecteurs immatures. Ainsi que la fonction ***élément_non soi (élément)*** est une fonction prédéfinie qui retourne la valeur vraie si l'élément présenté en paramètre n'est pas un élément de soi.

Les mêmes paramètres de l'expérience précédente seront utilisés pour tester les résultats obtenus après cette modification.

```
Représentation de P éléments par les cellules dendritiques accompagnés avec leur information
     de contexte ;
Pour chaque (élément présenté) Faire
 Début
   Si (contexte (élément) = "dangereux") Alors
    Début
        //Vérifier si l'élément n'est pas détectable par la population des détecteurs
          mémoires ;
         Si (détectable_mémoire (élément)) Alors
           Supprimer cette protéine de la population des éléments présentés par les
           cellules dendritiques ;
         Sinon // l'élément n'est pas détectable par la population des  détecteurs mémoires
          Début
             //Vérifier s'il est détectable par la population des détecteurs matures ;
            Si (détectable_mature (élément)) Alors
              Début
               Le détecteur correspondant deviendra un détecteur mémoire ;
               Supprimer cette protéine de la population des éléments présentés par les
               cellules dendritiques ;
              Fin
            Sinon
               Si (élément_non soi (élément)) Alors
                  Garder cette protéine pour une nouvelle vérification avec de nouveaux
                 détecteurs ;
                Sinon // l'élément présenté est un élément de soi dangereux
                    Si (détectable_immature (élément)) Alors
                       Début
                          Le détecteur correspondant deviendra un détecteur mémoire ;
                          Supprimer cette protéine de la population des éléments présentés
                          par  les cellules dendritiques ;
                       Fin si
                Fin si
            Fin si
         Fin si
    Fin
   Sinon /* contexte (élément) = " non dangereux" */

    Début
      //Vérifier si l'élément est détectable par la population des détecteurs mémoires
       Si (détectable_mémoire (élément))) Alors
          Supprimer le détecteur mémoire correspondant ;
       Fin
   Fin si
Fin pour chaque
```

Figure 5.8 : La modification apportée sur l'algorithme

V.8.2.2 Discussion

Après les modifications établies sur l'algorithme, les résultats de cette expérience montrent la possibilité de détecter des éléments de soi qui sont présentés dans un contexte dangereux avec un taux de détection qui peut atteindre 52% du nombre des éléments de soi dangereux (Figure 5.9) avec un taux d'erreurs minimal égal à 0%.

Cependant, l'activation des détecteurs immatures engendre la détection des éléments de soi non dangereux, ce qui implique des augmentations potentielles du taux de vrai négatif soi (Taux TN Soi) car le détecteur mémoire n'est pas parfaitement tolérant au soi c'est-à-dire il peut détecter d'autres éléments de soi non dangereux. Mais, ce problème est résolu par le fait que ces détecteurs immatures activés peuvent être supprimés, si les éléments de soi qui sont détectés par les détecteurs mémoires seront présentés par les cellules dendritiques dans un contexte non dangereux ce qui implique la suppression de ces détecteurs c'est-à-dire assurer seulement la tolérance aux éléments de soi non dangereux.

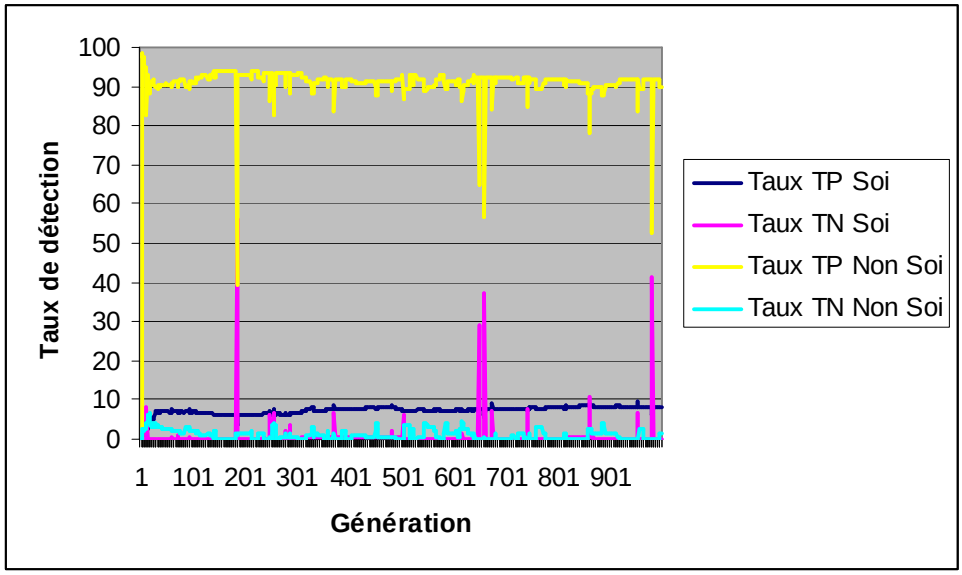


Figure 5.9 : Les taux de détection de l'expérience 2 avec r = 12

Ces courbes montrent la bonne adaptabilité de l'algorithme face à l'ensemble des antigènes parce qu'il peut rattraper toujours d'une manière autonome ces erreurs de détection par les modifications constantes établies sur l'ensemble des détecteurs mémoires.

L'augmentation élevée du taux de vrai négatif soi (Taux TN Soi) potentielle est liée ainsi à la valeur attribuée au nombre de bits de la règle de correspondance utilisée. En effet, il est nécessaire d'augmenter la valeur de r afin d'obtenir un détecteur optimal qui peut détecter les

éléments de soi dangereux tout en évitant la détection d'un grand nombre d'éléments de soi non dangereux. Puisque, la règle r bits contigus permet la liaison approximative entre un détecteur et un élément quelconque, pour tester l'influence de la valeur attribuée à la règle de r bits contigus sur les résultats d'exécution de l'algorithme, cette expérience sera rétablie avec une valeur de r égale à 16.

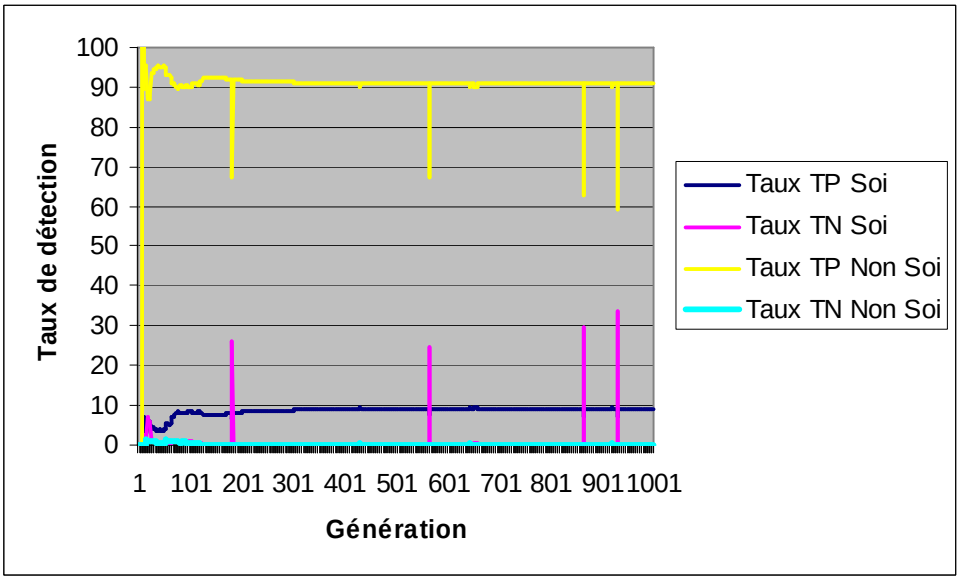


Figure 5.10 : Les taux de détection de l'expérience 2 avec r = 16.

Les résultats de la même expérience avec la modification du nombre de bits à 16 (Figure 5.10) montrent des résultats prometteurs par rapport aux résultats obtenus avec un nombre de bits égal à 12 pour la règle de correspondance, même s'il y'a quelques augmentations au niveau du taux vrai négatif soi (Taux TN Soi) mais ces taux sont faibles par rapport à ceux obtenus dans le cas précédent.

En plus, cette expérimentation montre mieux la tolérance graduelle aux éléments de non soi non dangereux avec l'augmentation de la détection des éléments de soi dangereux avec une moyenne de détection qui peut atteindre 55% du nombre total des éléments de soi dangereux et 99% du nombre total des éléments de non soi dangereux.

V.8.3 Expérience 3

V.8.3.1 Description

Le nombre d'éléments présenté par les cellules dendritiques influence sur les résultats de détection, un grand nombre d'éléments permet d'exploiter mieux les différents détecteurs

générés à chaque itérations qui sont soient matures ou immatures, ce qui augmente la probabilité pour trouver un détecteur adéquat qui correspond à l'élément présenté dans un contexte dangereux et qui n'est pas encore détectable par aucun détecteur mémoire.

Afin de montrer l'importance de ce paramètre, nous allons exécuter le programme avec les mêmes paramètres des expériences précédentes avec un nombre d'éléments présenté par les cellules dendritiques égal à 50.

V.8.3.2 Discussion

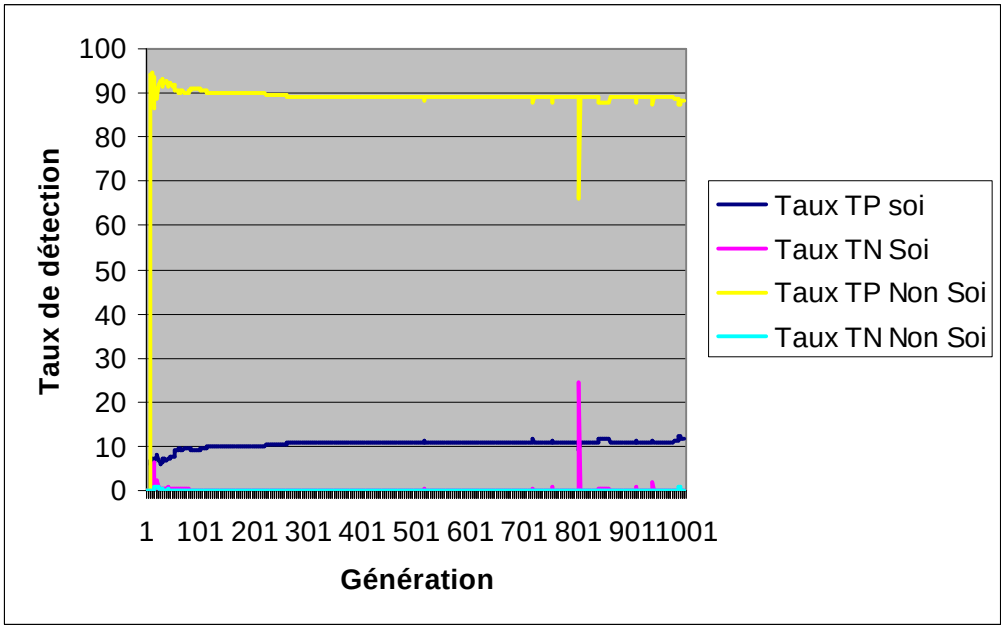


Figure 5.11 : Les taux de détection de l'expérience 3 avec un échantillon = 50.

Selon cette expérience, les résultats obtenus illustrés dans la figure 5.11 montrent qu'avec l'augmentation du nombre d'éléments présenté par les cellules dendritiques, le taux de détection des éléments dangereux de soi (Taux TP Soi) a augmenté pour atteindre environ 80 % du nombre total des éléments de soi dangereux avec un taux d'erreurs (Taux TN Soi) minimal égal à 0%. Cela est dû à la meilleure exploitation de l'ensemble des détecteurs générés pendant chaque itération.

En plus, ces résultats expérimentaux présentent des meilleurs taux de détection, c'est-à-dire un taux de détection vrai positif élevé et un taux de détection vrai négatif faible. Cela est dû au taux de modifications de la population des détecteurs mémoires qui est très élevé, par conséquent, il existe plus de chance pour supprimer les détecteurs détectant les éléments non

dangereux, ce qui implique la minimisation du taux vrai négatif, ainsi l'activation des détecteurs adéquats pour la détection des éléments dangereux qui ne sont pas encore détectés.

V.9 DISCUSSION

Afin de réaliser le but principal des systèmes de détection d'intrusions, qui consiste à augmenter le taux de vrai positif et réduire le taux de vrai négatif, le système doit détecter des intrusions réelles qui peuvent être produites par des éléments de soi ou de non soi. L'algorithme proposé effectue quelques modifications sur l'algorithme de la sélection négative par l'intégration de la notion de danger. L'incorporation de ce nouveau concept est assurée via la combinaison entre deux systèmes immunitaires inter-liés qui sont le système immunitaire inné et le système immunitaire adaptatif

Les résultats expérimentaux effectués sur cet algorithme montrent des résultats prometteurs car il est possible de détecter les éléments nuisibles qui peuvent être des éléments de soi ou de non soi avec la tolérance progressive aux éléments non dangereux avec des taux de détection de vrai positif élevés et des taux de détection de vrai négatif faibles.

Cependant, l'utilisation de la règle r bits contigus qui permet une liaison approximative entre un antigène et un détecteur influence sur les taux de détection après l'activation d'un détecteur immature qui peut détecter un élément de soi dangereux. Cette influence est marquée par des augmentations potentielles du taux de détection vrai négatif soi (Taux TN Soi) mais qui sera par la suite minimisé lors de la présentation de ces éléments de soi par les cellules dendritiques dans un contexte non dangereux, ce qui engendre la tolérance de la population des détecteurs mémoires aux éléments de soi non dangereux. En plus, le système arrive à détecter environ 55% d'éléments de soi dangereux avec un taux d'erreurs vrai négatif de soi minimal égal à 0%.

Ainsi, les résultats expérimentaux montrent l'influence du nombre d'éléments présenté par les cellules dendritiques sur les performances de l'algorithme. Un grand nombre d'éléments permet d'augmenter la probabilité d'obtenir des détecteurs adéquats pour la détection des éléments dangereux ainsi la tolérance aux éléments non dangereux par la meilleure exploitation des détecteurs générés.

Cet algorithme n'intègre pas l'étape de costimulation qui est une étape supplémentaire qui nécessite à chaque détection l'intervention de l'opérateur humain afin d'assurer que l'élément détecté constitue réellement une intrusion, mais il se base sur les changements constants

produits sur la population des détecteurs mémoires selon le contexte dés éléments présentés. Cette étape qui est assurée d'une façon implicite par les effets de suppression et d'activation sur la population des cellules mémoires suite à la présentation des éléments avec une information de contexte par les cellules dendritiques, sans aucune intervention de l'officier de sécurité. Bien que l'algorithme présente quelques augmentations au niveau du taux vrai négatif en conséquence de l'opération d'activation des détecteurs qui ne sont pas parfaitement tolérants au soi. Cependant, le système est capable de diminuer d'une manière autonome le taux d'erreurs par les modifications constantes établies sur la population des détecteurs mémoires.

V.10 CONCLUSION

Dans ce chapitre, nous avons présenté les différents travaux qui ont utilisé les systèmes immunitaires artificiels dans la détection d'intrusions. Egalement, nous avons constaté que le modèle de soi et de non soi est le modèle dominant parce qu'il est adopté par les différents travaux proposés. Cependant, ce modèle qui est basé sur la distinction entre le soi et non soi présente quelques problèmes. Avec l'apparition de la théorie de danger qui défie ce modèle et qui présente des nouvelles idées intéressantes, nous avons exploité quelques concepts proposés par cette nouvelle théorie dans la détection d'intrusions afin de surmonter les problèmes liés au modèle de soi et de non soi.

L'algorithme proposé dans ce travail essaye de surmonter les problèmes liés au modèle de soi et de non soi par l'amélioration de l'algorithme de la sélection négative qui se base principalement sur la discrimination entre le soi et le non soi, et qui ne permet pas la détection des éléments dangereux qui constituent réellement des intrusions, ces éléments qui peuvent être des éléments de soi ou de non soi, dont le but est de réaliser les meilleurs taux de détection. Cette solution intègre plusieurs concepts du système de détection d'intrusions LYSIS basé AIS, elle intègre aussi quelques concepts de base de la théorie de danger par l'exploitation de la notion de danger initiée par les éléments nuisibles dans le système ainsi que le modèle de la sélection clonale et la mémoire immunitaire. Dans cet algorithme, la détection des intrus est liée aux dommages qui peuvent apparaître dans le système, impliqués par des éléments internes ou bien par des éléments extérieurs. Cette détection est envisageable par l'exploitation des éléments présentés par les cellules dendritiques accompagnés avec une information de contexte indiquant l'état de l'environnement.

Les résultats expérimentaux appliqués sur cet algorithme illustrent l'adaptation graduelle de détection selon le contexte des éléments par des changements constants effectués sur la population des détecteurs mémoires afin de préserver que les détecteurs qui peuvent discerner les intrus. Ces changements qui sont les résultats de suppression des détecteurs mémoires et d'activation de nouveaux détecteurs selon le contexte de présentation des éléments par les cellules dendritiques. Ainsi, ces résultats exhibent la possibilité de surmonter les problèmes liés aux modèles de soi non soi par le fait qu'il est possible de détecter des éléments de soi qui sont dangereux dont le système est initialement tolérants. Ainsi, le système détecte les éléments de non soi qui constituent réellement des intrusions, qui ont provoqué des dégâts dans le système. En plus, cette solution n'intègre pas l'étape de costimulation qui nécessite à chaque détection l'intervention de l'opérateur humain afin d'assurer que l'élément détecté constitue réellement une intrusion, mais il se base sur les changements constants produits sur la population des détecteurs mémoires. Cette étape est assurée d'une manière indirecte par les effets de suppression et d'activation sur la population des cellules mémoires suite à la présentation des éléments avec une information de contexte, sans aucune intervention de l'officier de sécurité. Bien que l'algorithme présente quelques augmentations au niveau du taux vrai négatif qui sont dues à l'opération d'activation des détecteurs qui ne sont pas parfaitement tolérants au soi. Cependant, le système est capable de diminuer d'une manière autonome le taux d'erreurs par les modifications constantes établies sur la population des détecteurs mémoires.

CHAPITRE VI
CONCLUSION GENERALE

VI.1 INTRODUCTION

Avec la complexité croissante des réseaux et les systèmes informatiques, les solutions inspirées de la biologie constituent une source d'inspiration intéressante. Ces approches inspirées de la biologie restent intéressantes par rapport à d'autres approches pour deux raisons principales. D'une part, les systèmes informatiques et les espèces biologiques sont souvent attaqués, et d'autre part, les systèmes informatiques deviennent de plus en plus complexes et les approches traditionnelles de la sécurité ne peuvent pas assumer le rôle de protection d'une manière parfaite, par contre les métaphores biologiques assurent la protection du corps contre les intrus d'une manière très puissante.

Dans ce travail nous nous sommes focalisés sur l'utilisation des systèmes immunitaires artificiels dans la sécurité des systèmes informatiques et plus précisément les systèmes de détection d'intrusions. Pour cette raison, nous avons exposé les systèmes de détection d'intrusions, puis nous avons étudié les systèmes immunitaires artificiels afin de pouvoir établir l'analogie entre les systèmes de détection d'intrusions et les systèmes immunitaires artificiels. Ensuite, nous avons exposé les différents travaux proposés pour la détection d'intrusions qui ont exploité les systèmes immunitaires artificiels.

Vu l'émergence de la théorie de danger dans l'immunologie et qui défie l'immunologie classique qui se focalise sur la discrimination entre le soi et le non soi. Nous avons essayé d'exploiter la notion de danger stipulée par cette théorie afin de surmonter les problèmes liés

au modèle de soi et de non soi qui est le modèle le plus populaire puisqu'il est adopté dans les différents travaux proposés dans ce domaine, dont l'objectif principal est de pouvoir détecter les éléments dangereux qui ont initié des dommages dans le système.

VI.2 CONTRIBUTION

Dans le cadre de ce travail, nous avons essayé de combiner entre certains modèles du système immunitaire artificiel et la nouvelle théorie de danger qui s'intéresse au caractère dangereux d'un élément au contraire du modèle de soi et non soi qui se base principalement sur la distinction entre les éléments de soi et les éléments de non soi.

L'algorithme proposé intègre les principaux concepts du système de détection d'intrusions LYSIS proposé par Hofmeyr qui a amélioré l'algorithme de la sélection négative par l'intégration de certains paramètres supplémentaires tels que : la période de tolérance, le seuil d'activation, etc. Ainsi que l'utilisation d'autres modèles immunitaires tels que : la sélection clonale et la mémoire immunitaire ainsi le modèle des cellules dendritiques par l'exploitation de la notion de danger dont le but principal est de permettre la détection des éléments qui constituent des intrusions réelles et cela par l'ajout de quelques fonctionnalités supplémentaires qui reflètent les interactions existantes entre le système immunitaire inné et le système immunitaire adaptatif.

Nous avons utilisé la notion de danger et les interactions immunitaires existantes entre le système immunitaire inné et le système immunitaire adaptatif pour améliorer l'algorithme de la sélection négative qui est fondé sur le modèle de soi et de non soi afin d'obtenir la possibilité de détecter les éléments dangereux dans le système, ces intrus qui peuvent être soient des éléments internes ou externes du système. Ainsi, dans ce travail nous n'avons pas intégré l'étape de costimulation adoptée par les différents algorithmes dont le but principal de son utilisation est l'obtention d'une population des détecteurs mémoires qui détectent des intrusions réelles. Cependant, nous avons exploité les effets de suppression et d'activation engendrés par les éléments présentés par les cellules dendritiques qui impliquent des changements constants sur la population des détecteurs mémoires. Bien que l'effet d'activation des détecteurs puisse engendrer quelques augmentations du taux d'erreurs, le système peut rattraper ces erreurs d'une manière autonome dont le but principal est la détection des éléments nuisibles de soi ou de non soi et par conséquence l'augmentation du taux vrai positif et la réduction du taux vrai négatif.

VI.3 PERSPECTIVES

Vu l'analogie entre l'objectif du système immunitaire naturel et l'objectif des systèmes de détection d'intrusions. Ainsi, la capacité puissante du système immunitaire humain à assurer la protection du corps contre les différents intrus qui envahissent le corps. Les systèmes immunitaires artificiels présentent des solutions prometteuses pour assurer la protection des systèmes informatiques. Ce domaine de recherche constitue toujours le centre d'intérêt des différentes recherches afin d'exploiter tous les concepts et les mécanismes d'identification et de détection utilisés par le système immunitaire humain.

Les résultats expérimentaux effectués sur cet algorithme montrent la possibilité de détecter les éléments de soi qui peuvent provoquer des dégâts dans le système dont le système est initialement tolérant. Ainsi, l'algorithme détecte les éléments de non soi dangereux qui ont établi des intrusions réelles.

Nous espérons dans l'avenir d'intégrer d'autres concepts inspirés du système immunitaire humain et en particulier les fonctionnalités proposées par la théorie de danger comme par exemple la zone de danger qui est établie au tour de signaux d'alarme déclenchés, cette zone qui peut être intégrée dans le système immunitaire artificiel en terme temporelle. Ainsi, l'intégration des molécules de costimulation produisent par les cellules dendritiques qui ont un effet amplifiant sur les signaux d'alarme produits dans le système.

A la fin de ce travail, le système immunitaire naturel constitue toujours une source d'inspiration très riche dont le but principal des différentes recherches est la compréhension et l'extraction des mécanismes clefs utilisés par ce système dans l'identification, la détection et l'élimination des intrus afin de construire des systèmes immunitaires pour protéger les systèmes et les réseaux d'une manière efficace.

Annexe A

Les cellules dendritiques et le signal de danger.

Les cellules dendritiques sont des cellules spécialisées dans la représentation des protéines (antigènes) collectées avec leur contexte environnemental. Cette information est présentée aux cellules T et génère l'identification et la suppression des pathogènes.

Suite à la migration des cellules dendritiques, l'antigène est présenté aux cellules T avec les signaux de contexte dont le but est l'activation des cellules T naïves. Les cellules T qui ont un récepteur complémentaire à l'antigène présenté seront activées si l'antigène est présenté dans un contexte dangereux ou nécrose. Cependant, le contexte apoptose ou non dangereux engendre la tolérance des cellules T qui peuvent correspondre à l'antigène présenté.

L'information de contexte est traduite par l'utilisation des différents états de maturation des cellules dendritiques qui peuvent exister selon trois états de maturation : immature, mature ou bien semi mature.

Initialement, les cellules se trouvent dans un état immature dont la fonction principale est la collection des débris cellulaires du tissu via l'ingestion. Puis l'extraction des protéines existantes dans les débris et les stockées afin de préparer leur présentation aux cellules T. Cependant, la présentation des antigènes par les cellules dendritiques immatures ne permet pas l'activation des cellules T parce que la quantité des molécules de costimulation nécessaires ou les cytokines inflammatoires (des messagers locaux chimiques) est insuffisante, et qui sont nécessaire pour l'activation des cellules T.

Cependant, si le tissu est endommagé suite à une infection engendrée par des pathogènes ou d'autres cellules de stresse, alors ***les signaux d'alarme*** apparaissent dans le tissu. En plus, les pathogènes comme des bactéries expriment des protéines qui peuvent être reconnues par des récepteurs spécifiques sur les cellules dendritiques. Ces protéines sont connues comme des PAMP (Pathogen associated molecular patterns)[14]. L'exposition des cellules dendritiques à des signaux exogènes (PAMP) ou à des signaux de danger ou aux deux types de signaux au même temps implique la maturation complète des cellules dendritiques immatures et la migration des cellules dendritiques du tissu au nœud de lymphe. Avec la rupture des membranes, une cellule subit des lésions et perte tout son contenu dans le tissu. Quelques molécules trouvées seulement à l'intérieur de cellules apparaissent dans le liquide interstitiel. Les signaux sont un indicateur de cellule de stresse, impliquant la présence de danger dans ce tissu particulier. Les cytokines inflammatoires produisent par d'autres cellules dendritiques matures dans le secteur peuvent avoir un effet amplifiant sur les PAMP et les signaux de danger. Les effets cellulaires d'exposition à PAMP et des signaux de danger aboutissent à la production accrue des molécules de costimulation nécessaires pour la correspondance de cellule T et l'expression de cytokines qui active les cellules T naïves pour avoir une réponse immunitaire adaptative.

Au contraire, si le tissu est sain et les cellules ne sont pas sous le stresse, l'apoptose est la mort de cellule dominante par la destruction réglée des cellules. Cela assure que le contenu de cellule ne sort pas au liquide interstitiel. Les cytokines apparaissant suite à l'apoptose se lient avec les différents récepteurs des cellules dendritiques pour modifier la production des cytokines exprimées avec la production accrue des molécules de costimulation différentes (anti-inflammatoires). Les cytokines produisent par les cellules dendritiques semi matures impliquent la tolérance des cellules T qui peuvent détecter l'antigène présenté.

En résume, les cellules dendritiques ont la capacité d'agir comme un détecteur d'anomalies. Elles combinent entre les différents signaux d'entrée (les signaux de danger, les signaux apoptoses, les cytokines inflammatoires et les PAMP) afin de fournir les cellules T par l'information de contexte qui décrit l'état du tissu.

En se basant sur ce fonctionnement, Greensmith et al [56] ont proposé l'algorithme suivant :

[14] Les PAMP sont produits par tous les micro-organismes pathogéniques.

Créer une population de cellule dendritique taille 100
Pour chaque article de données
Choisir 10 cellules dendritiques dans la population
Pour chaque cellule dendritique
Ajouter l'antigène à la liste des antigènes rassemblés.
Mettre à jour les concentrations de signal d'entrée.
Calculer les concentrations des cytokines de sortie.
Mettre à jour le total cumulé de chaque cytokine de sortie.
Si (la concentration de molécule de costimulation> seuil flou)
Supprimer la cellule dendritique de la population
Migrer la cellule dendritique.
Créer une nouvelle cellule dendritique.

Pour chaque cellule dendritique qui migre
Si concentration de semi > mature
Contexte antigène = semi.
Sinon
Contexte antigène = mature.

Pour chaque antigène entrant dans le système
Calculer le nombre de fois que l'antigène est présenté comme mature ou semi mature
Si (semi > mature) **Alors** Antigène = normale
Sinon Antigène = anormale

Références bibliographiques

[1] Anderson. J. P « Computer Security Threat Monitoring and Surveillance », Technical Report James P Anderson Co., Fort Washington, PA, April 15, 1980.

[2] D. Denning «An intrusion detection models », IEEE, transaction on software engineering 13(2): 222-232, 1987.

[3] Mykerjee. B & Heberlein. L.T & Levitt .K.N « Network Intrusion Detection », IEEE Network, Vol 8, No 3, pp .26-41, 1994.

[4] R. Heady & G. Luger & A. Maccabe & M. Servilla « The Architecture of a Network Level Intrusion Detection System ». Technical Report CS90-20, University of New Mexico, Department of Computer Science, August 1990.

[**5**] **Przemysiam Kazienko & Piotr Dorosz** « Intrusion Detection Systems (IDS) Part I - (network intrusion; attack symptoms; IDS tasks; and IDS architecture) », 2004. http://www.windowsecurity.com/pages/article_p.asp?id=1147

[**6**] **H. Debar & M. Dacier & A. Wespi** « A revised taxonomy for Intrusion Detection Systems », Computer science, 1999

[7] Hervé Debar, « Application des réseaux de neurones à la détection d'intrusions sur les systèmes informatiques », thèse de doctorat, Université Paris 6, 1993.

[**8**] **NIST,** « Intrusion detection Systems ». NIST Computer Science Special reports SP 800-31, November 2001

[9] D. Zamboni « Intrusion Detection - Basic concept and current research at IBM », IBM Global security Analysis Lab, Information Technology Security Spring School, 2005

[10] A. Sundaram, « An introduction to Intrusion Detection », 1996 http://www.acm.org/crossroads/xrds2-4/intrus.html

[11] Axelsson. S.: « Intrusion Detection Systems: A Taxonomy and Survey ». Technical Report No 99-15, Dept. of Computer Engineering, Chalmers University of Technology, Sweden, 2000,

[12] Allen. J & al. « State of the practice of Intrusion Detection Technologies », Technical Report (No. CcMmU/SEI-99-TR-028)

[13] W. Jansen & P. Mell, T.Karygiannis & D.Marks «Mobile Agents in Intrusion Detection And Response », 2000

[14] C. Kruegel & T. Toth, « Applying Mobile Agent Technology to Intrusion Detection », Technical University of Vienna, Distributed Systems group, 2002.

[15] Jackson. K & DuBois. D & Stallings. C « The NIDES Statistical Component Description and Justification » Technical Report, Computer Science Laboratory, SRI International, Menlo Park, CA, March, 1994.

[16] Teng. H & Chen . K & Lu. S « Adaptive Real-Time Anomaly Detection Using Inductively Generated Sequential Patterns », Proceeding of the 1990 Symposium on Security and Privacy, Oakland, CA, May 7-9, pp 278-284, 1990.

[17] Li. Y& Wu. N & Jajosia. S & Sean Wang. X « Enhancing Profiles for Anomaly Detection Using Time Granularities » Proc. 1st ACM workshop on Intrusion Detection Systems, Athens, Greece, Nov. 2000.

[**18**] **J.Timmis & De Castro.L.N** « Artificial Immune System as a novel Soft Computing Paradigm ». To appear in the Soft Computing Journal, vol7, Issue 7, July 2003.

[19] J.Timmis & De Castro.L.N «Artificial Immune Systems : A novel paradigm to pattern recognition », In artificial Neural Networks in pattern recognition, Soco 2002, university of Paisely, UK, pp. 67-84,2002.

[20] De Castro .L.N & Von Zuben .F.J «Artificial Immune Systems: Part I - Basic theory and applications », Technical report, TR-DCA-01/99, December 99.

[21] De Castro .L.N & Von Zuben .F.J «Artificial Immune Systems: Part II - A survey of application », Technical report- DCA RT 02 /00, 2000.

[22] De Castro.L.N « An introduction to the Artificial Immune Systems », State University of Compinas- UNICAMP/ BRAZILE.2001. Http://www.dca.fee.unicamp.br/~Inunes

[23] De Castro.L.N & Von Zuben.F.J «The clonal selection algorithm with engineering application» Proc of GECCO00, Workshop Proceeding, 36-37, 2000

[24] Von Zuben.F.J & De Castro.L.N « aiNet: An artificial Immune Network for Data analysis », In Data Mining: A Heuristic Approach, H.A. Abbass, R.A.Sarker, C.S. Newlton (Eds.), Idea Group publishing, USA, Chapter XII, pp.213-259, 2001.

[25] J. Timmis & T. Knight & L.N. De Castro & E.Hart, «An overview of Artificial immune Systems », Natural computation series, pages 51-86, Springer, 2004.

[**26**] **J. Timmis & L.N. De Castro** «Artificial immune Systems: A New computational Intelligence Approach », Springer-Verglas, London, 2002.

[27] J. Timmis « Artificial Immune Systems: A novel data analysis technique inspired by the immune network theory », PhD Thesis, University of Wales, 2001.

[28] J. Timmis & M.Neal «A ressource limited artificial immune system », Knowledge based systems, 2001

[29] D. Dasgupta «Artificial Immune Systems and Their Application », Springer-Verglas, 1999.

[30] D. Dasgupta & Z.Ji & F.Gonzalez « Artificial Immune System (AIS) Research in the Last Five Years », IEEE.2003.

[31] Forrest .S & Perelson.A & Allen.L & Cherukuri. R «Self-Nonself Discrimination in a computer», Proc. Of the IEEE Symposium on research in Security and Privacy, pp. 2002-212, 1994.

[32] A .Somayaji & S.Forrest & S .Hofmeyr & T. Longstaff. « A sense of self for Unix processes ». IEEE Symposium on Security and Privacy, pages 120–128, 1996.

[33] S. Hofmeyr & S. Forrest. « Intrusion detection using sequences of system calls ». Journal of Computer Security, 6:151–180, 1998.

[34] S. Hofmeyr & S. Forrest. « Immunity by design ». Proceedings of GECCO, pages 1289–1296, 1999.

[35] Steven Hofmeyr « An immunological model of distributed detection and its application to computer security ». PhD thesis, University Of New Mexico, 1999.

[36] S. Hofmeyr & S. Forrest « Architecture for an artificial immune system », Evol. Comput, vol. 8, no. 4, pp 443 – 473, 2000.

[37] Jerne. N. K «Towards a Network theory of the immune system », Ann. Immunol. (Inst. Pasteur) 125C, pp. 373-389, 1974.

[38] Perelson .A.S & Oster .G.F « Theoretical studies of clonal selection minimal antibody repertoire size and reliability of Self-Nonself discrimination » J. theor.Biol, 81, pp 645-670, 1979.

[39] J. Kim & P. Bentely «The Human Immune System and Network Intrusion Detection », the proceeding of 7th European congress on intelligent Techniques and Soft computing (EUFIT 99), Aachen, Germany, 1999.

[40] J.Kim & P.Bentely « The artificial Immune model for Network Intrusion Detection », the proceeding of 7th European congress on intelligent techniques and soft computing (EUFIT 99), Aechan, Germany, 1999.

[41] J. Kim & P. Bentley. « Evaluating negative selection in an artificial immune system for network intrusion detection ». Proceedings of GECCO, pages 1330 – 1337, July 2001.

[42] J. Kim & P. Bentley « Towards an artificial immune system for network intrusion detection: An investigation of dynamic clonal selection ». In the Congress on Evolutionary Computation (CEC-2001), Seoul, Korea, pages 1244–1252, 2001

[43] J. Kim « Integrating Artificial Immune Algorithms for Intrusion Detection », PhD Thesis, University College London, 2002.

[44] Aickelin. U & Cayzer. S « The Danger Theory and Its Application to AIS», 1st International Conference on AIS, pp 141- 148, 2002.

[45] Aickelin. U & Bentley . P & Cayzer. S & Kim . J & McLeod. J « Danger Theory: The Link between AIS and IDS? », in Proceedings ICARIS – 2003, 2nd International Conference on Artificial Immune Systems, 147 – 155, 2003.

[**46**] **U. Aickelin & J. Greensmith & J. Twycross** « Immune System Approaches to Intrusion Detection - A review », School of Computer Science, University of Nottingham, 2004

[47] Matzinger. P « Tolerance, danger and the extended family », Annual reviews in Immunology, 12: 991- 1045, 1994.

[48] P. Matzinger. « An innate sense of danger». Seminars in Immunology, 10:399-415, 1998.

[49] Matzinger. P « The Danger Model: A renewed Sense of Self », Science 296: 301-305, 2002.

[50] J.Balthrop & F Esponda & S Forrest & M. Glickman. « Coverage and generalization in an artificial immune system ». Proceedings of GECCO, pages 3–10, 2002.

[51] J. Balthrop & S. Forrest & M. Glickman. « Revisiting lysis: Parameters and normal behaviour ». Proceedings of the Congress on Evolutionary Computation, pages 1045– 1050, 2002.

[52] D. Dasgupta & F. Gonzalez. « An immunity-based technique to characterize intrusions in computer networks ». IEEE Transactions on Evolutionary Computation, 6(3):281–291, 2002.

[53] F. Gonzalez & D. Dasgupta. « Anomaly detection using real-valued negative selection ». Journal of Genetic Programming and Evolvable Machines, 4:383–403, 2003.

[54] J. Kephart. « A biologically inspired immune system for computers ». In Proceedings of the Fourth International Workshop on Synthesis and Simulation of Living Systems, Artificial Life IV, pages 130–139, 1994.

[55] Twycross. j « Immune System, Danger Theory and Intrusion Detection », to be presented at the AISB 2004 Symposium on Immune System and Cognition (ImmCog-04), Leeds, U.K, 2004.

[56] Greensmith. J & Aickelin. U & Cayzer. S « Introducing Dendritic Cells as a Novel Immune-Inspired Algorithm for Anomaly Detection », To appear in the Proceeding of the fourth International Conference on Artificial Immune Systems (ICARIS-05), 2005.

[57] Robert E. Smith & Stephanie Forrest & Alan S. Perelson. « Searching for diverse, cooperative populations with genetic algorithms ». Evolutionary Computation, 1(2):127–149, 1993.

[58] C.L.Blacke & S.Hettich & C.J.Merz « UCI repository of machine learning databases », 1998. ftp://ftp.ics.uci.edu/pub/machine-learning-databases.

[59] Burgess, M. « Computer Immunology », *Proceedings of the 12th System Administration Conference (USENIX/LISA)* 1998. http://www.iu.hio.no/cfengine/papers.html

[60] Burgess, M. « Cfengine as a Component of Computer Immune-Systems », Proceedings of the Norwegian Informatics Conference 1998. http://www.iu.hio.no/cfengine/papers.html

[61] Burgess, M. « Evaluating Cfengine's Immunity Model of Site Maintenance », Proceedings of the 2nd SANE system administration conference (USENIX/NLUUG), 2000. http://www.iu.hio.no/cfengine/papers.html

[62] M. Crosbie & E. H. Spafford. « Active Defense of a Computer System using Autonomous Agents ». Technical Report CSD-TR-95-008, Purdue University, 1995.

[63] M. Crosbie & E. H. Spafford. « Defending a Computer System using Autonomous Agents ». Technical Report CSD-TR-95-022, Computer Sciences Department, Purdue University, 1995.

[64] Hart. E & Ross. P & Nelson. J « Producing Robust Schedules Via an Artificial Immune System », Proceeding of IEEE International Conference on Evolutionary Computing, 1998. http://www.dai.ed.ac.uk/daidb/people/homes/emmah/research.html

[65] Hart. E. & Ross. P «An Immune System Approach to Scheduling in Changing Environments », Proceeding of Genetic and Evolutionary Computation Conference (GECCO'99), pp. 1559 -1566, 1999.

[66] Hunt. J & Cooke. D., «Learning using an Artificial Immune System », Journal of Network and Computer Applications: Special Issue on Intelligent Systems Design and Application, Vol. 19, pp.189-212, 1996.

[67] Hunt. J & Timmis. J & Cooke. D & Neal. M & King. C. «Jisys: Development of Artificial Immune Systems for Real World Applications », Artificial Immune Systems and Their Applications, (Ed) Dasgupta, D., Springer-Verlag, Berlin, pp.157-186, 1998.

[69] McCoy. D. F & Devarajan. V « Artificial Immune Systems and Aerial Image Segmentation », Proceeding of IEEE Systems, Man and Cybernetics, pp.867- 873, 1997.

[70] Mitsumoto. N & Fukuda.T & Idogaki. T « Self-Organising Multiple Robotic System ». Proceedings of IEEE International Conference on Robotics and Automation. Pp. 1614-1619. Minneapolis, USA. IEEE.1996

[71] Mori. K & Tsukiyama. M & Fukuda. T «Adaptive Scheduling System Inspired by Immune System », Proceeding of IEEE Systems, Man and Cybernetics, pp.3833- 3837, 1998.

[72] Nikolaev. N & Iba. H & Slavov. V «Inductive Genetic Programming with Immune Network Dynamics », Advances in Genetic Programming 3, MIT Press, Chapter 15, pp. 335-376, 1999

[73] Okamoto, T & Ishida, Y « A Distributed Approach to Computer Virus Detection and Neutralization by Autonomous and Heterogeneous Agents », the Proceeding of the ISADS'99, pp.328-331. 1999.

[74] Porras. P. A & Neumann. P. G. « EMERALD: Event Monitoring Enabling Responses to Anomalous Live Disturbances », Proceeding of 20th National Information System Security Conference, 1998.

[75] Potter. M. A. & De Jong, K.A. «The Coevolution of Antibodies for Concept Learning », Proceeding of the fifth Intl. Conference on Parallel Problem Solving From Nature, pp.530-539, 1998.

[76] K. Price. « *Intrusion Detection Pages ».* Purdue University, 1998. http://www.cs.purdue.edu/coast/intrusion-detection/ids.html.

[77] S. Staniford-Chen & S. Cheung & R. Crawford & M. Dilger & J. Frank & J. Hoagland & S. Templeton & K. Levitt & S. Walnum & C. Wee, & R. Yip. « GrIDS-A Graph-Based Intrusion Detection System for Large Network »s. Proc of the 19th National Information Systems Security Conference, 1996.

[78] S. Staniford-Chen , « GrIDS Outline Design Document ». GrIDS Project Home Page at UC Davis'sComputer Science Department, 1997. http://olympus.cs.ucdavis.edu/arpa/grids/design.html,

[79] B. White & E. A. Fisch, & U. W. Pooch. « Cooperating Security Managers: A Peer-Based Intrusion Detection System ». IEEE Network Journal, pp. 20-23, January/February 1996.

[80] J. S. Balasubramaniyan & J. O. Garcia-Fernadez & D. Isacoff & E. H. Spafford & D. Zamboni. « An Architecture for Intrusion Detection using Autonomous Agents ». Technical Report Coast-TR-98-05, Computer Sciences Department, Purdue University, 1998.

[81] Dasgupta. D. « Immunity-Based Intrusion Detection Systems: A General Framework », the proceedings of the 22nd National Information Systems Security Conference (NISSC), October 18-21, 1999.

[82] J.Kim & J.Greensmith & J.Twycross & U.Aickelin « Malicious Code Execution Detection and Response Immune System inspired by the Danger Theory » , Adaptive and Resilient Computing Security Workshop (ARCS-05) , Nov 02-03, 2005.

[83] Forrest .S & Hofmeyr. S & Somayaji .A & Longstaff .T « A sense of self for Unix processes », Proceedings of the 1996 IEEE Symposium on Research in Security and Privacy, 120-128, 1996.

[84] Somayaji. A & Forrest. S « Automated response using system-call delays », Proceedings of the ninth USENIX Security Symposium, 185-197, 2000.