SÉCURITÉ INFORMATIQUE ET SYSTÈME DE DÉTECTION D'INTRUSIONS

1.1 Introduction

En raison de plusieurs facteurs notamment l'ouverture des systèmes d'information sur Internet, l'évolution de la technologie et des moyens de communication ainsi que la transmission de données à travers les réseaux, des risques d'accès et de manipulation des données par des personnes non autorisées d'une façon accidentelle ou bien intentionnelle sont apparus. Donc la mise en place d'une politique de sécurité autour de ces systèmes est devenu une nécessité incontournable.

Le système de détection d'intrusion est l'une des techniques utilisées pour garantir un contrôle permanent des attaques ainsi que la détection de toute violation de cette politique, c'est-à-dire toute intrusion.

Dans ce premier chapitre nous introduisons les principales notions de base de la sécurité informatique y compris sa définition, ses objectifs, les problèmes et les attaques informatiques et aussi les mécanismes permettant d'améliorer la sécurité. Ensuite, nous présentons les systèmes de détection d'intrusions, leur définition, architecture, classification...etc, et nous terminons par les limites des systèmes de détection d'intrusions actuels.

1.2 Sécurité informatique

1.2.1 Définition

La sécurité informatique est définie par la protection assurée aux systèmes informatiques ainsi qu'aux données stockées, transférées ou manipulées. Cette protection doit réaliser trois principaux objectifs : l'intégrité, la disponibilité et la confidentialité des ressources du système informatique hôte.[1]

1.2.2 Objectifs de la sécurité

Dans la littérature on trouve plusieurs définitions pour les objectifs de la sécurité, mais les standards (The Federal Information Processing Standards , 2004) citent trois principaux objectifs appelés la triade CIA.

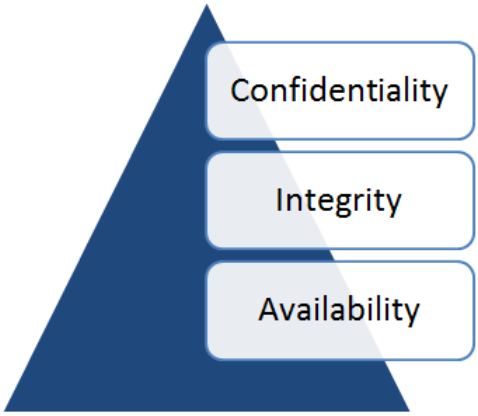

Figure 1.1 – Triade CIA

a) La confidentialité (confidentiality)

Permet d'assurer que les informations sauvegardées ou transmises sur le réseau ne soient pas dévoilées à des personnes, entités ou processus non autorisés, c'est-à-dire seules les personnes autorisées doivent pouvoir accéder aux données ou informations ainsi protégées.

b) L'intégrité (integrity)

Permet d'assurer que les données n'ont pas été altérées ou détruites de façon non autorisée, soit de manière accidentelle ou bien intentionnelle.

c) La disponibilité (availability)

Cet objectif vise à assurer l'accès aux ressources du système d'information conformément aux spécifications en terme de performances. Ceci implique que le temps d'attente et le temps de service sont tout les deux relativement raisonnables.

1.2.3 Soucis de la sécurité informatique

il existe trois problèmes qui affectent la sécurité informatique : les vulnérabilités, les menaces et les attaques.

a) Les vulnérabilités

Ce sont des failles ou des faiblesses dans la spécification, conception, implémentation ou bien configuration des systèmes informatiques dont l'exploitation peut créer une intrusion.

b) Les menaces

Une menace c'est la possibilité d'une violation d'une propriété de la sécurité en exploitant une ou plusieurs vulnérabilités d'une façon intentionnelle ou accidentelle.

c) Les attaques

Une attaque c'est une action malveillante qui tente d'exploiter une faiblesse dans le système et de violer un ou plusieurs besoins de sécurité.

1.2.4 Classification des attaques informatiques

Une attaque peuvent être classée selon son objectif, son point d'initiation ou la façon d'adresser la victime désirée.

a) Selon l'objectif d'attaque

On trouve deux types d'attaques principaux : passives et actives .

- **Les attaques passives :** ce type d'attaque ne provoque pas d'altération aux ressources du système ciblé ce qui le rend généralement indétectable (récupération du contenu d'un message ou bien l'observation du trafic).
- **Les attaques actives :** consistent à effectuer des modifications ou bien une destruction des ressources d'un système d'une manière non autorisée. Ce type d'attaque est plus dangereux que le premier et peut causer des dégâts (usurpation de l'identité, modification, replay, déni de service...etc).

b) Selon le point d'initiation

On distingue deux types d'attaques pour ce critère de classification : attaques de l'intérieur et attaques de l'extérieur.

- **Les attaques de l'intérieur :** provenant des utilisateurs légitimes d'un système lorsqu'ils se comportent de façon non autorisée.
- **Les attaques de l'extérieur :** venant de l'extérieur, souvent via Internet, en utilisant des techniques comme l'usurpation d'identité.

c) Selon la façon d'adresser la victime

Il existe deux façons pour adresser la victime soit d'une manière directe ou bien indirecte.

- **Les attaques directes :** dans ce type d'attaque, l'intrus adresse ses paquets directement à la victime sans passer par un intermédiaire.
- **Les attaques indirectes :** dans ce type d'attaque, l'adversaire envoie ses paquets vers une entité intermédiaire qui à son tour les retransmet vers la victime.

1.2.5 Buts des attaques

Il existe plusieurs objectifs pour les attaques informatiques :

- **Interruption :** vise la disponibilité des informations (DoS).
- **Interception :** vise la confidentialité des informations (sniffing, analyse de trafic,...etc).
- **Modification :** vise l'intégrité des informations.
- **Fabrication :** vise l'authenticité des Informations.

Les quatre objectifs sont illustrés dans la figure suivante :

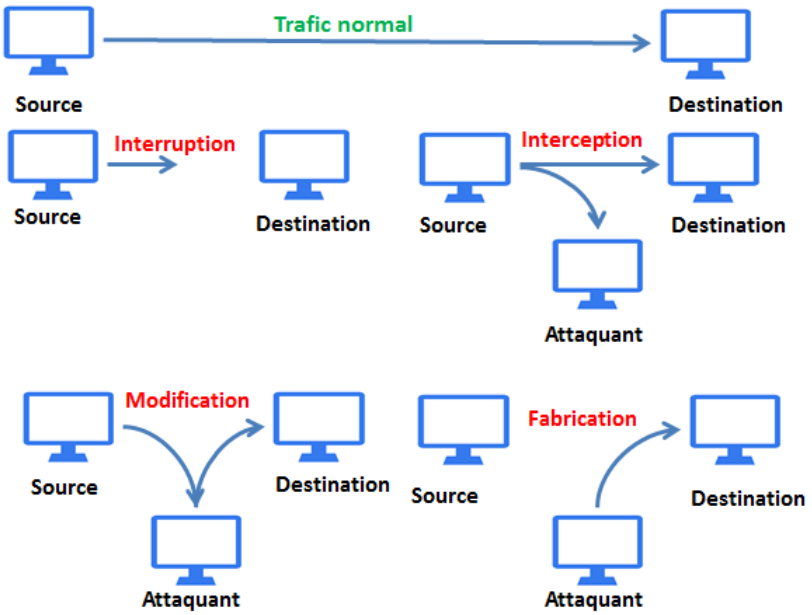

Figure 1.2 – Les objectifs des attaques informatiques
[2]

1.2.6 Motivations des attaques informatiques

Il existe plusieurs motivations d'un attaquant à vouloir exploiter une vulnérabilité et effectuer une attaque, parmi elles on peut citer les suivantes :[3]

- **motivations financières :** lorsqu'il s'agit de prendre possession des données pour rançonner l'entreprise ou le particulier.
- **motivations économique et concurrentielle :** on espionne ou on commet des actes de malveillances envers un concurrent dans le but d'acquérir un avantage commercial.
- **motivations politique ou idéologique :** comme semble l'indiquer la cyberattaque Not-Petya [1] ou le piratage de l'entreprise Ashley Madison (Mansfield-Devine, 2015) [2].

1.2.7 Exemples d'attaques informatiques

Il existe un nombre énorme d'attaques qui menacent les systèmes informatique à travers le monde entier, les plus connues aujourd'hui sont :

a) IP Spoofing

Le principe de l'attaque IP Spoofing est relativement ancien (aux alentours de 1985) alors que sa première application dans une vraie attaque ne remonte qu'à 1995. Kevin Mitnick, un célèbre hacker, l'utilise afin de s'infiltrer dans le réseau d'un expert en sécurité informatique, Tsutomo SHimomura. [4]

1. **NotPetya** est un logiciel malveillant de type wiper (il détruit les données), mais apparait sous la forme d'un rançongiciel (appelé aussi ransomware en anglais) en affichant sur l'écran de l'ordinateur infecté une demande de rançon. Son mécanisme de propagation permet de le classer comme ver informatique.
2. Les données de 32 millions de comptes sur Ashley Madison, le principal site de rencontres adultères aux Etats-Unis, ont été mises en ligne

Cette attaque consiste à usurper l'adresse IP d'une machine pour cacher la source d'attaque ou bien profiter d'une relation de confiance entre deux machines. Il existe des variantes car on peut spoofer aussi des adresses e-mail, des serveurs DNS ...etc.

b) Le dénis de service

Cette attaque consiste à envoyer des milliers des messages depuis des dizaines d'ordinateurs afin de saturer le système et donc le rendre indisponible. Ce type d'attaque est très facile à mettre en place mais très difficile à empêcher.
Un attaquant peut utiliser les DOS pour les raisons suivants :[4]

- obtenir le contrôle sur une machine cible ou sur un réseau. C'est le cas par exemple d'une attaque de type « SYN Flooding »qui est souvent utilisée de paire avec une tentative de spoofing.
- masquer les traces en détruisant les stations qui auraient pu contenir des traces d'un attaquant.
- se venger contre une personne, un administrateur ou bien encore une entreprise...etc.

Il existe plusieurs types d'attaques DOS comme il est montré dans la figure suivante :

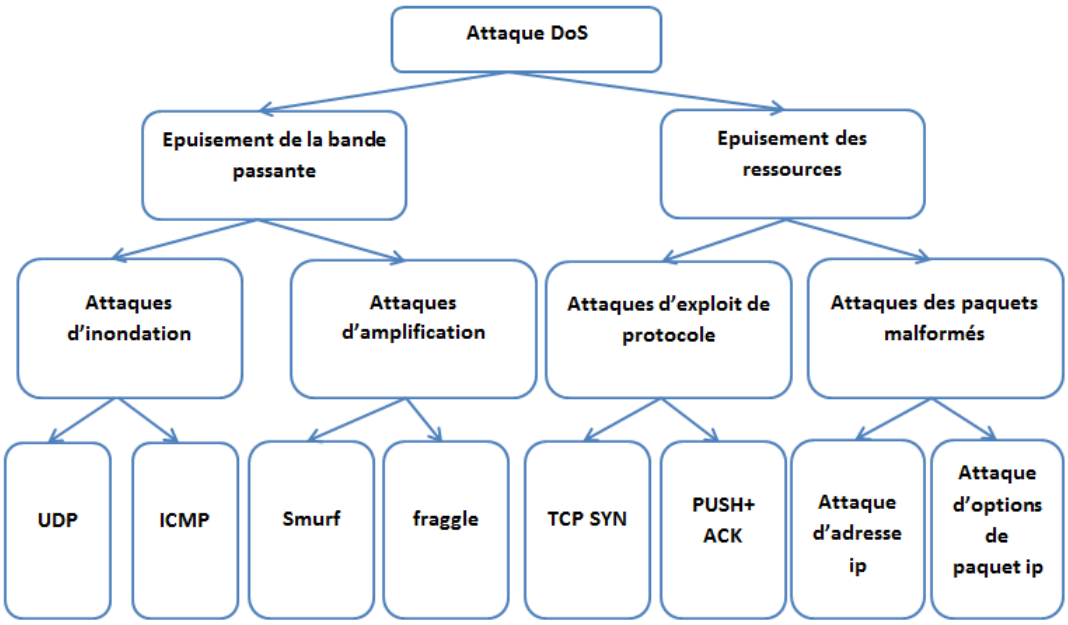

Figure 1.3 – Taxonomie des attaques dos
[5]

c) Probing (Sondage)

Le sondage est une attaque dans laquelle le pirate analyse une machine ou un réseau pour déterminer les faiblesses ou les vulnérabilités qui pourraient être exploitées plus tard afin de compromettre le système. Cette technique est couramment utilisée dans l'exploration de données, par exemple saint, portsweep, mscan, nmap...etc.[6]

Cette classe d'attaque est la plus étendue et qu'elle requit une expertise technique minime, donc il est très important de protéger le système de telles intrusions car elles sont à la base d'autres attaques comme R2L (Remote-to-Local), U2R (User-to-Root)...etc.

d) User to Root

Ces attaques sont des exploitations dans lesquelles le pirate démarre sur le système avec un compte d'utilisateur normal et tente d'abuser des vulnérabilités du système afin d'obtenir des privilèges de super utilisateur.[7]

En d'autre terme, l'objectif de cette attaque est d'obtenir les privilèges de l'administrateur système (Root) en allant d'un simple compte utilisateur (User) et cela en exploitant des failles dans le système comme le débordement de tampon, les erreurs de programmation..etc.

Il existe plusieurs attaques de ce type comme Eject,Ffbconfig, Fdformat, Load module, Perl, Xterm...etc.

e) Remote to Local

C'est une attaque dans laquelle l'attaquant exploite les vulnérabilités d'une machine distante comme les bugs des applications, les mauvaises configuration des systèmes d'exploitation et d'autres afin d'obtenir un accès illégal à celle-ci en exploitant les privilèges d'un utilisateur local.

Plusieurs attaques se trouvent dans cette catégorie parmi elles on cite : xlock, guest, xnsnoop, phf, Dict, warezmaster, spy, warezclient...etc.

1.2.8 Techniques et mécanismes de sécurisation

Pour réaliser les objectifs de sécurité cités dans la section 1.2.2, on doit prévoir un ensemble de mécanismes permettant de détecter toute attaque possible sur le système et même dans certains cas de prévenir ces attaques si cela est possible pour garantir un niveau élevé de protection du réseau et du système d'information. Ces mécanismes peuvent être implémentés à différents niveaux de l'architecture réseau en couche.

a) Le chiffrement

C'est un algorithme généralement basé sur des clefs pour transformer les données. Sa sécurité est dépendante du niveau de sécurité des clefs.[2]

Autrement dit, le chiffrement consiste à utiliser des algorithmes permettant de coder les données en une forme non intelligible afin de les protéger contre toute divulgation non autorisée. Cette technique permet d'assurer la confidentialité des données.

b) La signature numérique

Cette technique consiste à calculer une valeur à l'aide d'un algorithme de chiffrement, cette valeur sera ajoutée à une donnée d'une façon que tout récepteur de cette donnée puisse vérifier son origine.

La signature remplit deux fonctions juridiques principales [8] :

- L'identification de l'auteur et la manifestation de son consentement. La signature numérique est le pendant électronique à la signature manuscrite, mais la signature digitale est liée au document signé,elle n'est pas comparée à une signature témoin mais elle est vérifiée algorithmiquement alors elle est universellement vérifiable.
- Une signature numérique apporte la non-répudiation à l'origine, c'est-à-dire l'auteur d'une action ne peut dénier l'avoir effectué.

c) Le bourrage

Données ajoutées pour assurer la confidentialité, notamment au niveau du volume du trafic.[2]
Les mécanismes de bourrage servent à modifier les caractéristiques du trafic pour assurer différents niveaux de protection contre l'analyse du celui-ci.

d) Le contrôle d'accès

Ce mécanisme consiste à vérifier les droits d'accès aux données en laissant passer que les personnes autorisées et cela pour empêcher toute exploitation de vulnérabilités venant de l'extérieur.

e) La notarisation

Le mécanisme de notarisation consiste à reposer sur un tiers de confiance (notaire) qui détient les informations nécessaires pour assurer certains services de sécurité comme la non-répudiation.

f) Le pare-feu

Un pare-feu, ou coupe-feu ou encore firewall est un équipement ou des systèmes qui contrôlent le flux de trafic entre les différentes zones d'un réseau[9]. Donc, il assure un périmètre de protection entre le réseau interne à l'entreprise et le monde extérieur.
Voici une figure qui montre l'emplacement du pare-feu au seine d'une entreprise à fin de protéger le réseau local et les serveurs sensibles de l'entreprise (DMZ [3]).

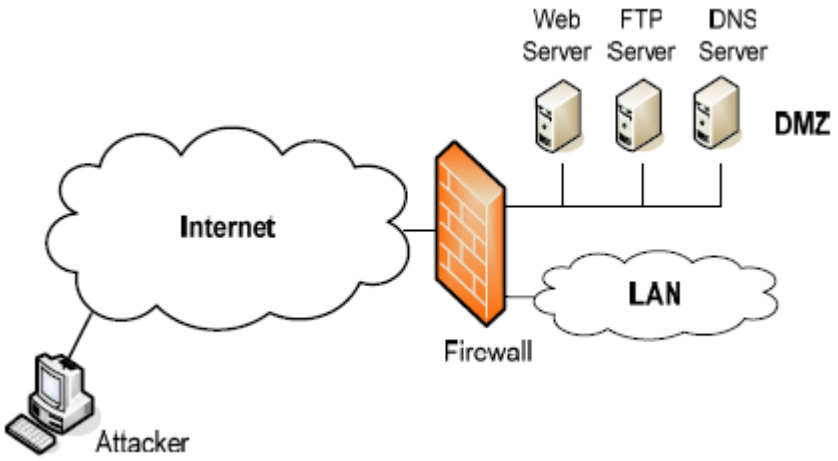

Figure 1.4 – Déploiement tri-hébergé d'un pare-feu de réseau d'entreprise
[10]

Un pare-feu peut assurer les tâches suivantes :

- bloquer l'accès à des services non autorisés.
- protéger en temps réel contre les menaces embarquées dans les applications.
- protéger contre les attaques de type DoS (Deni de service).
- intégrer des techniques de détection d'intrusions et envoyer des alertes afin de prévenir les équipes de surveillance technique.

3. **DMZ** :désigne (zone démilitarisée), est un sous-réseau séparé du réseau local et isolé de celui-ci et d'Internet (ou d'un autre réseau) par un pare-feu. Ce sous-réseau contient les serveurs sensibles de l'entreprise

- gérer les connexions sortantes à partir du réseau local.
- protéger le réseau interne des intrusions venant de l'extérieur.
- identifier et contrôler les applications partageant une même connexion.
- ...etc.

g) L'antivirus

C'est un logiciel permettant de préserver le système de tout type de maliciels (virus, vers, chevaux de troie...etc). Son principe de fonctionnement peut suivre l'une des trois approches :

- comparer la signature virale du virus aux codes vérifiés.
- utiliser les méta-heuristiques pour détecter les codes malveillants.
- utiliser le filtrage basé sur les règles.

h) La détection d'intrusions

La détection des intrusions est un mécanisme de cybersécurité courant dont la tâche est de détecter les activités malveillantes dans des environnements hôte et / ou réseau. La détection des activités malveillantes permettent de réagir en temps opportun, par exemple pour arrêter une attaque en cours. Vu l'importance de détection des intrusions, les milieux de la recherche et de l'industrie ont conçu et développé une variété de systèmes de détection d'intrusion (IDS).[11]

1.3 Systèmes de détection d'intrusions

1.3.1 Définitions

a) Intrusion

C'est toute utilisation d'un système informatique à des fins autres que celles prévues.[12] Autrement dit, c'est toute action malveillante qui vise l'un des objectifs de sécurité : La confidentialité, l'intégrité ou la disponibilité.

b) Détection d'intrusions

Consiste à analyser les informations collectées par les mécanismes d'audit de sécurité, à la recherche d'éventuelles attaques.[12]

c) Audit de sécurité

C'est un examen méthodique d'une organisation ou d'un site visant à identifier ses risques, ses vulnérabilités et les faiblesses de ses protections existantes ainsi qu'à statuer sur son niveau de sécurité et à recommander des solutions aux problèmes identifiés.[13]

d) Système de détection d'intrusions

Le système de détection d'intrusions inclure tous les systèmes logiciels et matériels permettant d'automatiser les processus de surveillance et d'analyse des évènements au sein d'un système informatique afin de détecter toute activité pouvant conduire à une défaillance de sécurité. Il peut être déployé sur une hôte, on parle alors de Host-Based Intrusion Detection System

(HIDS), ou bien sur un réseau, on parle alors de Network-Based Intrusion Detection System (NIDS).

1.3.2 Architecture d'un IDS

Plusieurs architectures ont été proposées pour décrire les différents éléments constituant un système de détection d'intrusions. L'architecture la plus simple est composée de trois modules :le capteur, l'analyseur et le manager. Cette architecture est montrée dans la figure suivante :

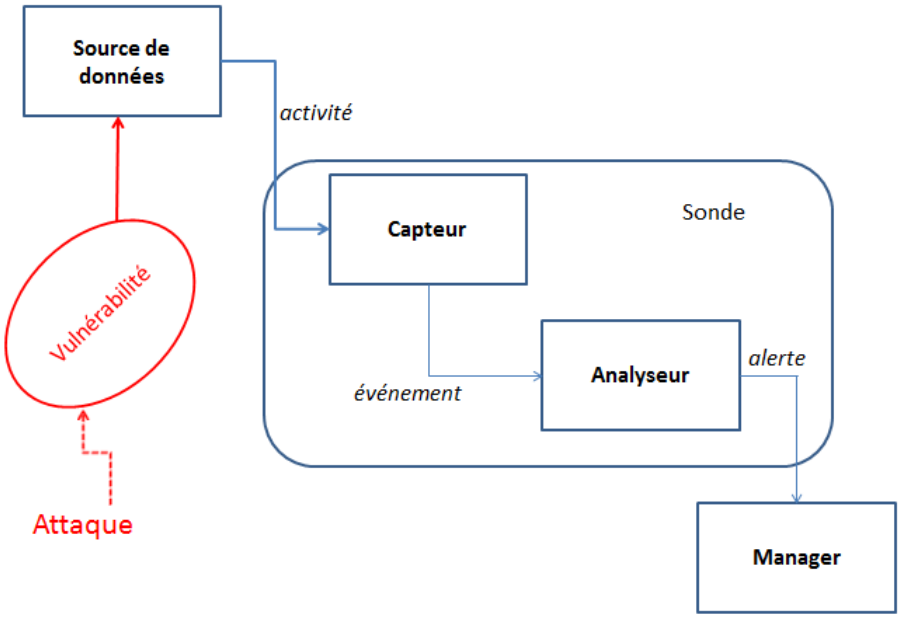

Figure 1.5 – Architecture classique d'un IDS
[14]

a) Le capteur

Chargé de collecter, filtrer et formater les informations brutes envoyées par la source de données concernant l'évolution de l'état du système. Le résultat de traitement est un message formaté appelé événement.

b) L'analyseur

Permet d'analyser les événements générés par le capteur en détectant toute activité malveillante qui peut se produire à partir d'un sous-ensemble de ces événements, et donc envoyer une alerte qui sera stockée dans les journaux du système ou bien utilisée pour lutter contre les attaques selon le type d'IDS.

c) Le manager

Permet de collecter et notifier les alertes envoyées par l'analyseur. Éventuellement, le manager est chargé de la réaction à adopter qui peut être :

- Isolement de l'attaque pour réduire les dégâts.
- Suppression d'attaque.

- Restauration du système dans un état sain.
- Identification du problème qui a engendré cette attaque.

1.3.3 Principe de fonctionnement d'un IDS

Le fonctionnement d'un IDS et le processus de détection d'intrusions sont illustrés dans la figure suivante :

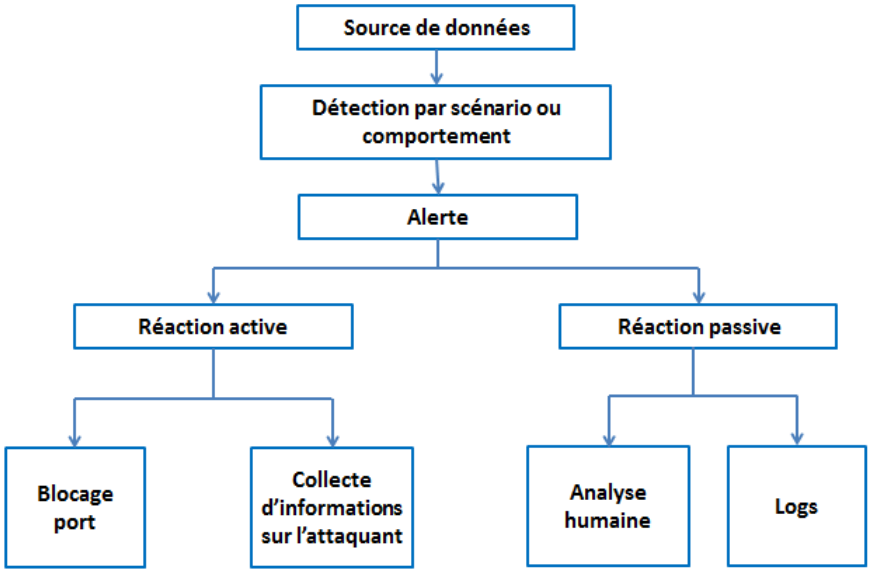

Figure 1.6 – Fonctionnement d'un IDS
[15]

1.3.4 Emplacement des IDS

Il existe plusieurs endroits stratégiques où il convient de placer un IDS pour atteindre le niveau de protection attendu selon la politique de sécurité choisie.

Le schéma suivant illustre un réseau local ainsi que les trois positions que peut y prendre un IDS :

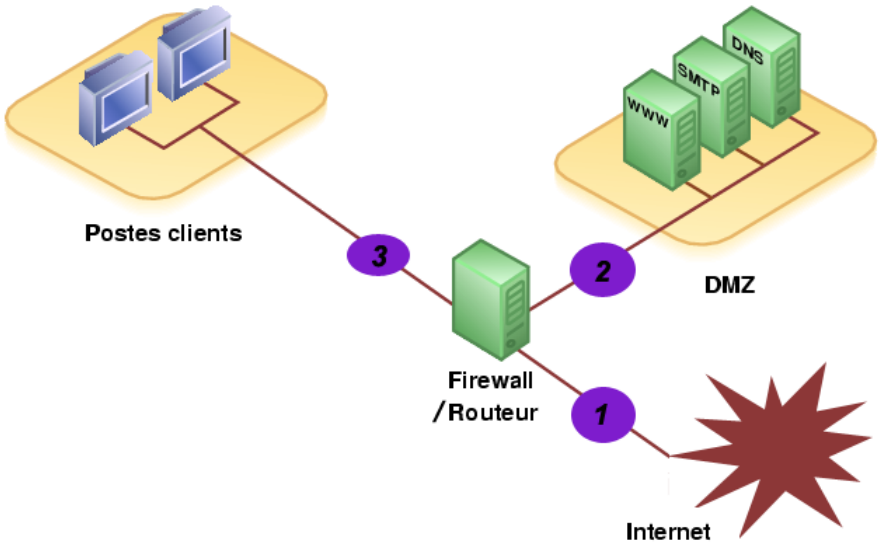

Figure 1.7 – Emplacement d'un IDS
[16]

Position 1 : Lorsque l'IDS prend cette position, son rôle sera de détecter l'ensemble des attaques frontales, provenant de l'extérieur, vers le pare-feu. Donc, plusieurs alertes seront remontées ce qui rendra les logs difficilement consultables.

Position 2 : Si l'IDS est placé sur la DMZ, il détectera les attaques qui n'ont pas été filtrées par le pare-feu et qui relèvent d'un certain niveau de compétence. Les logs seront ici plus clairs à consulter puisque les attaques bénines ne seront pas recensées.

Position 3 : L'IDS dans cette position a pour objectif de rendre compte des attaques internes, provenant du réseau local de l'entreprise. Il peut être judicieux d'en placer un à cet endroit étant donné le fait que 80% des attaques proviennent de l'intérieur. De plus, si des trojans ont contaminé le parc informatique (navigation peu méfiante sur internet) ils pourront être ici facilement identifiés pour être ensuite éradiqués.

1.3.5 Approches pour la détection d'intrusion

Il existe deux approches pour la détection d'intrusions. La première consiste à rechercher des signatures connues d'attaques tandis que la seconde consiste à définir un comportement normal du système et à rechercher ce qui ne rentre pas dans ce comportement.

Un système de détection d'intrusions par recherche de signatures connaît ce qui est mal, alors qu'un système de détection d'intrusions par analyse de comportement connaît ce qui est bien. On parle de détection de malveillances et de détection d'anomalies.[12]

a) Détection de malveillances

Cette approche est plutôt ancienne, remontant aux années 1990,et s'avère très efficace pour trouver des menaces connues. Elle consiste à rechercher des activités abusives par comparaison avec des descriptions abstraites de ce qui est considéré comme malveillant. Cette approche tente de mettre en forme des règles qui décrivent les usages non désirés, en s'appuyant sur des intrusions passées ou des faiblesses théoriques connues. En cas de détection d'une menace, une alerte est émise et le processus de remédiation est enclenché.
L'implémentation de cette approche peut être réalisée en utilisant plusieurs méthodes :

- **Systèmes experts :** Ils peuvent être utilisés pour coder les signatures de malveillance avec des règles d'implication si . . .alors. Les signatures décrivent un aspect d'une attaque ou d'une classe d'attaque. Il est possible d'ajouter des nouvelles règles pour les nouvelles attaques.
- **Analyse des transitions d'états :** On crée un modèle tel que le système au début ne soit pas compromis. L'intrus accède au système. Il exécute une série d'actions qui provoquent des transitions sur les états du modèle, qui peuvent être des états où on considère que le système soit compromis. Cette approche de haut niveau peut reconnaître des variations d'attaques qui passeraient inaperçues avec des approches de plus bas niveau.
- **Réseaux de neurones :** La flexibilité apportée par les réseaux neuronaux permet d'analyser des données même si elles sont incomplètes ou déformées. Ils peuvent de plus permettre une analyse non-linéaire de ces données. Leur rapidité permet l'analyse d'importants flux d'audit en temps réel. On peut utiliser les réseaux neuronaux pour filtrer et sélectionner les informations suspectes pour permettre une analyse détaillée par un système expert. On peut aussi les utiliser directement pour la détection de malveillances. Mais leur apprentissage est extrêmement délicat, et il est difficile de savoir quand un réseau est prêt pour l'utilisation. On peut également lui reprocher son côté boite noire (on ne peut pas interpréter les coefficients).
- **Raisonnement sur des modèles :** On essaye de modéliser les malveillances à un niveau élevé et intuitif d'abstraction en termes de séquences d'événements qui définissent l'intrusion. Cette technique peut être utile pour l'identification d'intrusions qui sont proches mais différentes. Elle permet aussi de cibler les données sur lesquelles une analyse approfondie doit être faite.
- **Algorithmes génétiques :** On définit chaque scénario d'attaque comme un ensemble pas forcément ordonné d'événements. Lorsqu'on veut tenir compte de tous les entremêlements possibles entre ces ensembles, l'explosion combinatoire qui en résulte interdit l'usage d'algorithmes de recherche traditionnels, et les algorithmes génétiques sont d'un grand secours.

On peut rapprocher les méthodes utilisées à cette approche à ceux qu'on peut les rencontrer au domaine des antivirus ou encore dans le domaine de la génomique où l'on recherche une séquence d'ADN dans un brin

La détection de malveillance a deux inconvénients principales :

- La difficulté de construction des bases de signatures.
- La non détection des attaques non connues.

b) Détection d'anomalies

Cette approche se base sur l'hypothèse que l'exploitation d'une faille du système nécessite une utilisation anormale de ce système, et donc un comportement inhabituel de l'utilisateur. Elle cherche donc à répondre à la question « *le comportement actuel de l'utilisateur ou du système est-il cohérent avec son comportement passé ?* ».

Il existe plusieurs méthodes pour la mise en œuvre de cette approche, parmi elles on peut citer :

- **Observation de seuils :** On fixe le comportement normal d'un utilisateur à certaine valeur (seuil), par exemple le nombre maximum de mots de passe erronés, mais Il est très difficile de caractériser un comportement intrusif en termes de seuils. En effet, on peut avoir

beaucoup de fausses alertes ou d'intrusions non détectées dans une population d'usagers non uniforme par exemple.

- **Profilage d'utilisateurs :** On crée et on maintiens des profils individuels du travail des utilisateurs, auxquels ils sont censés adhérer ensuite. Au fur et à mesure que l'usager change ses activités, son profil de travail attendu se met a jour. Certains systèmes tentent de concilier l'utilisation de profils à court terme et de profils à long terme. Il reste cependant difficile de profiler un utilisateur irrégulier ou très dynamique. De plus, un utilisateur peut arriver à habituer lentement le système à un comportement intrusif.
- **Profilage de programmes exécutables :** On observe l'utilisation des ressources du système par les programmes exécutables. Les virus, chevaux de Troie, vers, bombes logiques et autres programmes du même goût se voient démasqués en profilant la façon dont les objets du système comme les fichiers ou les imprimantes sont utilisés. Le profilage peut se faire par type d'exécutable.
- **Profilage adaptatif à base de règles :** Contrairement à la détection de malveillances à base des règles, là on n'a pas besoin des connaissances d'un expert car ces règles sont générées automatiquement lors de la phase d'apprentissage. Donc, l'efficacité de cette méthode nécessite la génération de beaucoup de règles ce qui engendre des problèmes de performance.
- **Réseaux de neurones :** Les réseaux neuronaux offrent une alternative à la maintenance d'un modèle de comportement normal d'un utilisateur. Ils peuvent offrir un modèle plus efficace et moins complexe que les moyennes et les déviations standards.

Cette approche a aussi beaucoup d'inconvénients comme :

- La difficulté à dire si les observations faites pour un utilisateur particulier correspondent à des activités que l'on voudrait prohiber.
- Pour un utilisateur au comportement erratique, toute activité est normale.
- Pas de prise en compte des tentatives de collusion entre utilisateurs.
- Choix délicat des différents paramètres du modèle statistique...etc.

c) Systèmes hybrides

Pour compenser les lacunes des méthodes précédentes, certains systèmes utilisent une combinaison de la détection d'anomalies et la détection de malveillances. Par exemple, un administrateur peut avoir un profil qui lui permet d'accéder à certains fichiers sensibles, mais on doit vérifier que les attaques connues ne soient pas utilisées contre ces fichiers. À l'inverse,l'utilisation des fichiers comportant le mot « nucléaire »ne caractérise aucune signature d'attaque, mais cela est possible si ce n'était pas dans les habitudes de l'utilisateur.

1.3.6 Efficacité des IDS

L'efficacité d'un système de détection d'intrusions est déterminée par les mesures suivantes : [17][18][19]

- **Exactitude :** Un système de détection d'intrusions n'est pas exact s'il déclare comme malicieux une activité légale. Ce critère correspond au faux positif.
- **Performance :** La performance de système de détection d'intrusions est le taux de traitement des événements. Si ce taux est faible, la détection en temps réel est donc impossible.

- **Perfection :** Un système de détection d'intrusions est imparfait s'il n'arrive pas à détecter une attaque.
- **Tolérance aux pannes :** Le système de détection d'intrusions doit lui-même résister aux attaques, en particulier dans le cas des attaques de déni de service. Ceci est important car plusieurs systèmes de détection d'intrusions s'exécutent sur des matériels ou logiciels connus vulnérables aux attaques.
- **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.

1.3.7 Limites des IDS

La plupart des systèmes de détection d'intrusions existants souffrent d'au moins deux des problèmes suivants :[20]

- Tout d'abord, les informations utilisées par le système de détection d'intrusions sont obtenues à partir d'un audit des chemins ou des paquets sur un réseau. Les données doivent parcourir un long chemin à partir de leur origine à l'IDS, donc elles peuvent potentiellement être détruites ou modifiées par un attaquant. En outre, le système de détection d'intrusions doit déduire le comportement du système à partir des données collectées, ce qui peut entraîner des interprétations erronées ou des événements manqués. Cela est appelé problème de fidélité.
- Deuxièmement, le système de détection d'intrusions utilise en permanence des ressources système supplémentaires, il surveille même lorsqu'il n'y a pas d'intrusions, car les composants du système de détection d'intrusions doivent fonctionner en permanence. C'est le problème de consommation des ressources.
- Troisièmement, parce que les composants du système de détection d'intrusions sont mis en œuvre comme programmes distincts, ils sont susceptibles d'être altérés. Un intrus peut potentiellement désactiver ou modifier les programmes exécutés sur un système, rendant la détection d'intrusions inutile ou peu fiable. C'est le problème de fiabilité.

1.4 Conclusion

Dans un monde où le progrès technologique avance à grande vitesse, où les gens, les entreprises, les organismes, les pays et même les objets sont de plus en plus connectés, les attaques informatiques sont de plus en plus fréquentes. La question de la cybersécurité se pose à tous les niveaux et tend à devenir un enjeu essentiel ces prochaines années.

Dans ce chapitre, nous avons abordé différentes notions concernant la sécurité informatique, où nous avons présenté les différents types d'attaques, leur classification, leurs objectifs et motivations et les techniques utilisées pour protéger le système contre ces attaques. Parmi ces mécanismes, nous avons détaillé les systèmes de détection d'intrusions, vu que c'est notre objectif dans ce mémoire, qui continuent d'évoluer pour répondre aux exigences et offre un éventail de fonctionnalités capables de satisfaire les besoins de tous les types d'utilisateurs.

Donc, nous avons détaillé l'architecture des systèmes de détection d'intrusions, leur principe de fonctionnement et les différentes approches pour la détection d'intrusions où on a divisé les IDS en deux grandes catégories , les IDS comportementaux et les IDS à base de signatures. Ces derniers consistent à détecter les attaques en se basant sur leurs signatures ce qui demande

une mise à jour périodique de la base des signatures et rend la détection des nouvelles attaques impossible. C'est pour cela qu'on a basé dans ce travail sur l'approche comportementale qui offre la possibilité de détecter les attaques inconnues en s'appuyant sur les réseaux de neurones et les méta-heuristiques qui seront détaillés dans les prochaines chapitres.

CHAPITRE 2

CLASSIFICATION ET RÉSEAUX DE NEURONES

2.1 Introduction

La classification de données est un problème délicat qui a apparu dans de nombreux domaines tels que l'analyse d'image, le diagnostic médical, l'apprentissage automatique...etc, et cela à cause de l'explosion de la quantité de données traitée par les systèmes informatiques lors de ces dernières années.

Dans le domaine de la sécurité informatique, la classification sert à classifier le trafic réseau ou les données du système surveillé en deux classes principale (normale/attaque,légal/illégal,...etc).En effet, les méthodes de classification sont implémentées dans les systèmes de détection d'intrusions pour les aider à prendre des décisions et générer des alertes en cas d'attaques avec le moins de fausses alarmes possibles. Pour ce faire, plusieurs méthodes de l'intelligence artificielle peuvent être utilisées, telle que les réseaux de neurones.

Dans ce chapitre, nous allons présenter la classification des données vu que c'est une phase principale dans le processus de génération de notre modèle de détection d'intrusions. Nous commençons tout d'abord par sa définition, ses catégories et les méthodes de classement les plus utilisées dans chacune. Ensuite, on focalise sur les réseaux de neurones en exposant leur historique, leur définition, leurs techniques d'apprentissage et leur architecture avant de passer aux détails de perceptron multicouches, vu que c'est l'outil de notre travail, et nous terminons par les avantages et les inconvénients des réseaux de neurones et leurs domaines d'applications.

2.2 Classification

2.2.1 Définition

La classification est l'opération statistique qui consiste à regrouper des objets(individus ou variables) en un nombre limité de classes de sort qu'on doit satisfaire deux conditions principales :[21]

- Ces classes ne sont pas prédéfinies par l'analyste mais découvertes au cours de l'opération.
- Une homogénéité interne et hétérogénéité externe, c'est-à-dire les classes de la classification regroupent les objets ayant des caractéristiques similaires et séparent les objets ayant des caractéristiques différentes.

2.2.2 Architecture typique d'une application basée sur la classification

La classification est une tâche très importante dans le data mining [1], mais le développement d'un outil de classification dans n'importe quel domaine doit être réalisé en appliquant d'autre phases d'extraction et de l'analyse d'informations qui sont montrées dans la figure suivante :

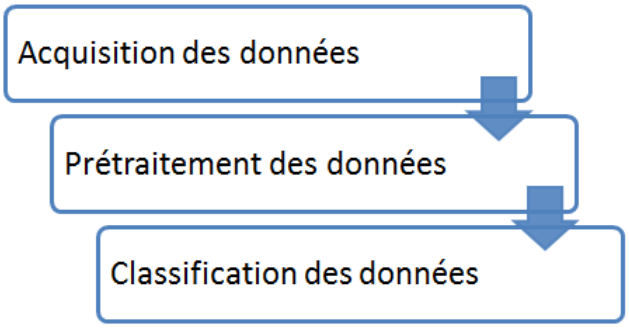

Figure 2.1 – Processus du data mining

- **Acquisition des données :** D'une manière générale, il s'agit de mettre en place l'ensemble d'instrumentation (capteurs, matériel d'acquisition, etc.) de façon à reproduire le phénomène observé le plus fidèlement possible. [22]. Dans notre cas, il s'agit de placer les sondes (IDS) pour écouter le trafic réseau.
- **Pré-traitement de données :** Cette phase correspond au filtrage des informations en ne conservant que ce qui est pertinent dans le contexte d'étude puisque ces données peuvent contenir plusieurs types d'anomalies (elles peuvent être omises à cause des erreurs de frappe ou à cause des erreurs dues au système lui-même, elles peuvent être incohérentes donc on doit les écarter ou les normaliser...etc). Parfois on est obligé à faire des transformations sur les données pour unifier leur poids.
- **Classification des données :** Dans cette étape, on doit choisir la bonne technique pour extraire les connaissances des données (les réseaux de neurones, les arbres de décision, les réseaux bayésiens...etc). Dans notre cas, la classification des connexions TCP/IP, on se base sur les réseaux de neurones.

2.2.3 Catégories de classification

Il existe un grand nombre des méthodes pour la résolution des problèmes de classification. Cependant, il est possible de les regrouper sous forme d'une hiérarchie de méthodes puisque certaines de ces approches partagent des caractéristiques communes, soit dans le type d'apprentissage utilisé (apprentissage supervisé ou non), ou bien dans la sortie réalisée (groupes disjoints ou classification floue).

La figure suivante montre une taxonomie des méthodes de classification dérivée de celle de Jain et Dubes (1988) :

1. Le **data mining** est un procédé d'exploration et d'analyse de grands volumes de données en vue d'une part de les rendre plus compréhensibles et d'autre part de découvrir des corrélations significatives

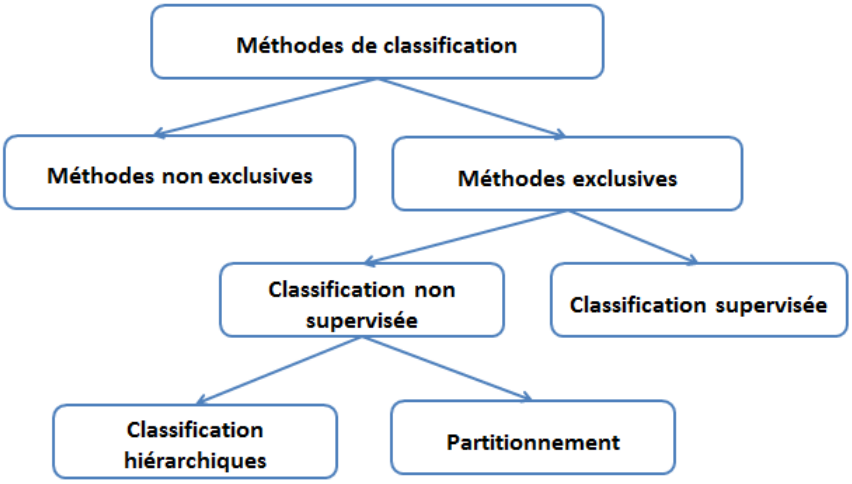

Figure 2.2 – Les méthodes de classification
[23]

a) Classification non exclusive

Une méthode est dite non exclusive si chaque objet est associé à une densité de probabilité qui indique pour chacune des classes la probabilité que l'objet considéré y appartienne (une classification floue).

b) Classification exclusive

Une méthode est dite exclusive si un objet ne peut être affecté qu'à une classe et une seule. Elle peut être divisée en deux autres classes :

I. **Classification non supervisée :** il s'agit d'extraire à partir d'une population des classes ou groupes d'individus présentant des caractéristiques communes. Le nombre et la définition des classes n'étant pas donnés a priori[24]. Le clustering regroupe un ensemble de techniques qui visent à regrouper les enregistrements d'une base de données en des groupes selon leur rapprochement les uns des autres en ne se basant sur aucune information antérieure.

il existe deux approches pour le clustering : le partitionnement et les méthodes hiérarchique.

- **Partitionnement :** consiste à construire plusieurs partitions puis les évaluer selon certains critères. Parmi les méthodes de partitionnement les plus connues on trouve **la méthode des k-moyennes(K-Means)**.
  Le principe de l'algorithme *k-means* est le suivant :
  - Choisir k objets formant ainsi k clusters.
  - (Ré)affecter chaque objet O au cluster Ci de centre Mi tel que distance(O,Mi) est minimale.
  - Recalculer Mi de chaque cluster.
  - Aller à la deuxième étape si on vient de faire une affectation.

Figure 2.3 – Exemple de partition obtenue en utilisant le K-means

- **Classification ascendante hiérarchique :** consiste à créer une décomposition hiérarchique des objets selon certains critères. Elle a pour objectif de construire une suite de partitions emboitées des données en n classes, n-1 classes,...,1 classe. Ces méthodes peuvent être vues comme la traduction algorithmique de l'adage « *qui se ressemble s'assemble* ».
  Le principe de l'algorithme CAH peut être exprimé comme suit :
  — A l'étape initiale, les n individus constituent des classes à eux seuls.
  — On calcule les distances deux à deux entre les individus, et les deux individus les plus proches sont réunis en une classe.
  — La distance entre cette nouvelle classe et les n-2 individus restants est ensuite calculée, et à nouveau les deux éléments les plus proches sont réunis.
  — Ce processus est réitère jusqu'à ce qu'il ne reste plus qu'une unique classe constituée de tous les individus.

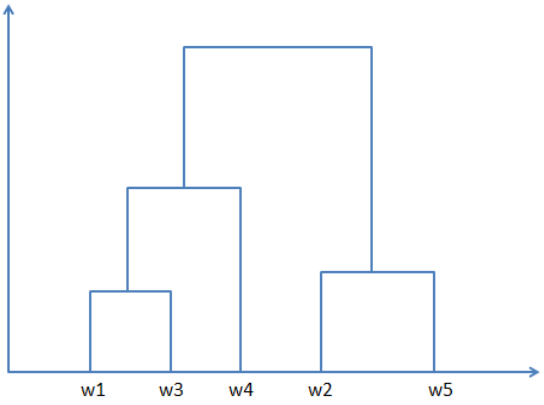

Figure 2.4 – Exemple de dendogramme (arbre hiérarchique)

II. **Classification supervisée :** elle consiste à inférer à partir d'un échantillon d'exemples classés une procédure de classification. Le problème est alors d'être capable d'associer à tout nouvel objet sa classe la plus adaptée, en se servant des exemples déjà étiquetés. [25]
Il existe plusieurs méthodes de classification supervisée, les plus connues sont : la méthode de k plus proche voisins et la méthode d'arbre de décision.

- **K plus proche voisins :** l'algorithme des k-plus proches voisins est un des algorithmes de classification les plus simples. Le seul outil dont on a besoin est une distance entre les éléments que l'on veut classifier.
  On considère que l'on dispose d'une base d'éléments dont on connait la classe, on parle de la base d'apprentissage, bien que cela soit de l'apprentissage simplifié. Dès que l'on reçoit un nouvel élément que l'on souhaite classifier, on calcule sa distance à tous les éléments de la base. Si cette base comporte 50 éléments, alors on calcule 50 distances et on obtient donc 50 nombres réels. Si k=5 par exemple, on cherche alors les 5 plus petits nombres parmi ces 50 nombres. Ces 5 nombres correspondent donc aux 5 éléments de la base qui sont les plus proches de l'élément que l'on souhaite classifier. On décide d'attribuer à l'élément à classifier la classe majoritaire parmi ces 5 éléments.
- **Arbres de décision :** les arbres de décision représentent l'une des techniques les plus connues et les plus utilisées en classification. Leur succès est notamment dû à leur aptitude à traiter des problèmes complexes de classification. En effet, ils offrent une représentation facile à comprendre et à interpréter, ainsi qu'une capacité à produire des règles logiques de classification.[26]
  Un arbre de décision est caractérisé par :
  — chaque nœud correspond à un test sur la valeur d'un ou plusieurs attributs.
  — chaque branche partant d'un nœud correspond à une ou plusieurs valeurs de ce test.
  — à chaque feuille, une valeur de l'attribut cible est associée.

  L'utilisation des arbres de décision dans les problèmes de classification se fait en deux étapes principales :
  — la construction d'un arbre de décision à partir d'une base d'apprentissage.
  — la classification ou l'inférence consistant à classer une nouvelle instance à partir de l'arbre de décision construit dans la première étape.

  Voici un exemple d'arbre de décision et la partition qu'il implique :

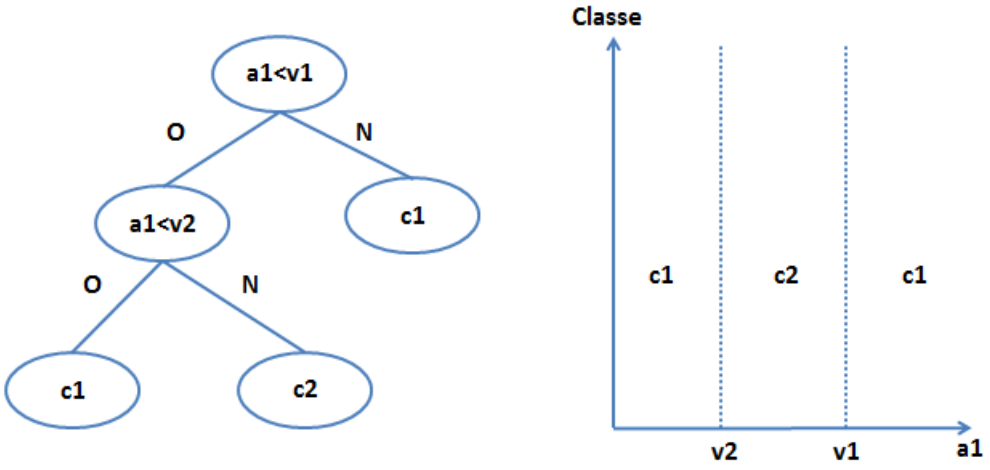

Figure 2.5 – Exemple d'une classification par arbre de décision
[27]