République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université Chadli Bendjedid
Faculté des Sciences et de la Technologie
Département d'Informatique

الجمهورية الجزائرية الديمقراطية الشعبية
وزارة التعليم العالي والبحث العلمي
جامعة الشاذلي بن جديد
كلية العلوم و التكنولوجيا
قسم الاعلام الالي

MEMOIRE

Présentée par

Hanine Abdalia

Pour l'obtention de diplôme de

MASTER

Filière : Informatique

Spécialité : Systèmes Informatiques Intelligents

Thème

Un réseau de neurones profond optimisé via Keras tuner pour la détection des DDoS

Soutenu le : 10 / 09 / 2023

Devant le Jury composé de :

Qualité	Nom et Prénom	Grade	Université
Président	Mr. Chemam. Chaouki	MCA	Chadli Bendjedid El-Tarf
Rapporteur	**Mr.Betouil Abd el latif**	**MCB**	**Chadli Bendjedid El-Tarf**
	Mr. Ahmed ahmim	**MCB**	
Examinateur	Mme. Gasmi Ibtissem	MCA	Chadli Bendjedid El-Tarf

Année Universitaire : 2022/2023

Remerciements

Je remercie Allah le tout puissant de m'avoir donné le courage jusqu'à l'achèvement de ce mémoire. Au terme de ce travail, J'adresse ma profonde gratitude à **Dr Ahmed Ahmim** Et **Dr Betouil Ali Abdelatif**. Je les remercie pour l'aide inestimable, la disponibilité, la compréhension, la gentillesse, les conseils et les encouragements avec lesquelles ils ont bien voulurent diriger ce travail, il étaient vraiment responsables, compétant et patients. J'ai eu le grand plaisir de travailler sous ses directions. J'espère être digne de la confiance qu'ils ont placée en moi.

Je tiens à exprimer aussi mes sincères remerciements à tous les membres du jury qui ont pris le temps d'évaluer mon travail. Je suis reconnaissante pour le nombre de personnes qui ont participé à ce processus et qui ont apporté leur expertise et leur expérience pour évaluer mon travail de manière juste et impartiale

Je désire aussi remercier les professeurs de l'université de Chadli ben jdid El Taref qui m'ont fourni les outils nécessaires à la réussite de mes études universitaires.

Dédicace

A tous ceux que j'aime, a tous ceux qui m'aiment..

A ma chère maman qui n'est plus parmi nous, je voudrais dédier ma réussite à ce jour. Tu as toujours été ma plus grande source d'inspiration et de motivation, et je sais que ta présence bienveillante m'a accompagné tout au long de ce parcours. Même si tu n'es plus là, je sais que tu es fière de moi. Je t'aime et je te dédie cette réussite et aussi la femme qui m'a éduqué et qui n'est plus parmi nous, je voudrais dédier ma réussite à ce jour. Tu as été une figure importante dans ma vie, tu m'as appris tant de choses et tu m'as encouragé à poursuivre mes rêves. Je suis reconnaissante pour tout ce que tu as fait pour moi.

A mon cher papa, je voudrais te dire merci pour tout ton soutien et ton encouragement tout au long de mes études. Ta présence et tes encouragements m'ont permis d'arriver à ce moment important de ma vie. Cette soutenance est autant la mienne que la tienne, car c'est grâce à toi que j'ai pu atteindre cet objectif Merci pour tout ce que tu as fait pour moi. Je t'adore.

A ma chère tante « NACIRA » , je voudrais dédier ma réussite à ce jour. Votre présence bienveillante, votre soutien constant et votre amour inconditionnel ont été essentiels pour moi tout au long de ce parcours. Vous avez su me guider, m'encourager et me soutenir dans les moments difficiles, et je ne pourrai jamais assez-vous remercier pour cela. Votre influence positive sur ma vie restera inoubliable et je suis fière de partager ce moment avec vous. Merci d'être toujours là pour moi et de m'avoir aidé à réaliser mes rêves.

A mes deux chères sœurs « NOURHENE & IKRAM » je voudrais vous dire combien vous êtes importantes dans ma vie. Vous êtes mes meilleures amies, mes confidentes Vous m'avez soutenu, encouragé et guidée, je voudrais aussi dédier cette réussite à vous deux qui avez toujours été là pour moi, dans les bons comme dans les mauvais moments. Votre soutien, vos encouragements et votre amour inconditionnel m'ont aidé à réaliser mes rêves c'est grâce à vous que j'ai pu arriver jusqu'ici. Merci d'être des sœurs extraordinaires, des amies fidèles. Je vous aime toutes les deux de tout mon cœur sans oubliée mon très cher frère « AKRAM »

A mes chères amies et collègues, je voudrais dédier cette réussite à vous toute qui avez toujours été là pour moi, Votre amitié, votre soutien et votre encouragement. Merci d'être toujours là pour moi, de me motiver, de m'encourager et de m'aimer pour qui je suis.

Et la fin je voulais dire à une chère personne à moi merci pour ton encouragement et pour votre soutien merci de toujours vouloir me voir à la hauteur.

Table des matières

Liste des figures

Liste des tableaux

Liste des acronymes

DOS	Denial of Service
DDoS	Distributed Denial of Service
IA	Artificial Intelligence
URL	Uniform Resource Locator
SQL	Structured Query Language
TCP	Transmission Control Protocol
IDS	Intrusion Detection System
IPS	Intrusion Prenvention System
HIDS	Host-based Intrusion Detection System
NIDS	Network-based Intrusion Detection System
PIDS	Protocol-based Intrusion Detection System
APIDS	Intrusion Detection System based on application protocol
ML	Machine learning
DL	Deep learning
DNN	Deep Neural Network
CNN	Convolutional Neural Network
RNN	Recurrent Neural Network

RÉSUMÉ

La détection des attaques DDoS est un défi majeur en matière de sécurité informatique. Les réseaux de neurones profonds sont des modèles puissants capables de reconnaître les comportements malveillants associés à ces attaques en extrayant des caractéristiques complexes des données.

Ce mémoire se concentre sur l'optimisation d'un réseau de neurones profond pour la détection des attaques DDoS en utilisant Keras Tuner. Cet outil permet d'explorer automatiquement et de manière itérative les hyperparamètres du modèle afin de trouver la meilleure configuration.

En utilisant un large ensemble de données d'attaques DDoS, le réseau de neurones est entraîné à reconnaître les schémas caractéristiques de ces attaques. Ensuite, grâce à Keras Tuner, les hyperparamètres tels que le nombre de couches, les fonctions d'activation et les taux d'apprentissage sont ajustés pour maximiser la précision de la détection.

Les résultats obtenus démontrent l'efficacité du modèle optimisé dans la détection des attaques DDoS. En exploitant les capacités d'apprentissage profond du réseau de neurones et en optimisant les hyperparamètres, le modèle atteint une précision élevée et une meilleure capacité à distinguer les attaques DDoS du trafic normal.

Et à la fin , l'utilisation d'un réseau de neurones profond optimisé via Keras Tuner offre une approche prometteuse pour la détection des attaques DDoS. Ce modèle optimisé permet une détection plus précise et fiable des attaques, renforçant ainsi la sécurité des systèmes informatiques face à cette menace croissante.

Mots clés : IDS,DOS, DDoS ,Keras Tuner .

ABSTARCT

Detecting DDoS attacks is a major challenge in computer security. Deep neural networks are powerful models capable of recognizing malicious behaviors associated with these attacks by extracting complex features from the data.

This study focuses on optimizing a deep neural network for DDoS attack detection using Keras Tuner. This tool allows for automatic and iterative exploration of the model's hyperparameters to find the best configuration.

By using a large dataset of DDoS attacks, the neural network is trained to recognize characteristic patterns of these attacks. Then, using Keras Tuner, hyperparameters such as the number of layers, activation functions, and learning rates are adjusted to maximize detection accuracy.

The results demonstrate the effectiveness of the optimized model in detecting DDoS attacks. By leveraging the deep learning capabilities of the neural network and optimizing its hyperparameters, the model achieves high accuracy and improved ability to distinguish DDoS attacks from normal traffic.

In conclusion, the use of an optimized deep neural network through Keras Tuner offers a promising approach for DDoS attack detection. This optimized model enables more precise and reliable detection of attacks, thereby enhancing the security of computer systems against this growing threat.

Kyword :IDS ,DOS, DDoS ,Keras Tuner .

ملخص

يعد اكتشاف هجمات DDoS تحديًا كبيرًا لأمن تكنولوجيا المعلومات. تعد الشبكات العصبية العميقة نماذج قوية قادرة على التعرف على السلوكيات الضارة المرتبطة بهذه الهجمات عن طريق استخراج ميزات معقدة من البيانات.

تركز هذه الأطروحة على تحسين الشبكة العصبية العميقة للكشف عن هجمات DDoS باستخدام. Keras Tuner تتيح لك هذه الأداة استكشاف المعلمات الفائقة للنموذج تلقائيًا وبشكل متكرر من أجل العثور على أفضل تكوين.

باستخدام مجموعة كبيرة من بيانات هجماتDDoS ، يتم تدريب الشبكة العصبية على التعرف على الأنماط المميزة لهذه الهجمات. بعد ذلك، باستخدامKeras Tuner ، يتم ضبط المعلمات الفائقة مثل عدد الطبقات ووظائف التنشيط ومعدلات التعلم لزيادة دقة الاكتشاف إلى أقصى حد.

وتظهر النتائج التي تم الحصول عليها فعالية النموذج الأمثل في الكشف عن هجمات .DDoS من خلال الاستفادة من قدرات التعلم العميق للشبكة العصبية وتحسين المعلمات الفائقة، يحقق النموذج دقة عالية وقدرة أفضل على التمييز بين هجمات DDoSوحركة المرور العادية.

وفي النهاية، فإن استخدام شبكة عصبية عميقة تم تحسينها من خلال Keras Tuner يوفر طريقة واعدة للكشف عن هجمات .DDoS يسمح هذا النموذج الأمثل باكتشاف الهجمات بشكل أكثر دقة وموثوقية، وبالتالي تعزيز أمن أنظمة تكنولوجيا المعلومات في مواجهة هذا التهديد المتزايد

الكلمات المفتاحية: Keras Tuner ,DDoS,DOS,IDS

Introduction Générale

La détection des attaques DDoS (Distributed Denial of Service) est une préoccupation majeure dans le domaine de la sécurité des systèmes informatiques. Les attaques DDoS peuvent paralyser les services en ligne en inondant les serveurs cibles de trafic malveillant, entraînant ainsi une interruption des services et des conséquences financières et opérationnelles importantes. Dans ce contexte, l'utilisation de techniques avancées d'apprentissage automatique, telles que les réseaux de neurones profonds, offre de nouvelles perspectives pour détecter et contrer ces attaques.

Un réseau de neurones profonds est un modèle d'apprentissage automatique qui imite les processus de traitement de l'information dans le cerveau humain. Il est composé de multiples couches de neurones interconnectés qui permettent d'extraire des caractéristiques complexes à partir des données d'entrée. Cette capacité d'apprentissage et de généralisation fait des réseaux de neurones profonds des outils puissants pour la détection d'attaques DDoS.

Cependant, la tâche de configuration et d'optimisation des réseaux de neurones profonds peut être complexe et nécessite une exploration approfondie de l'espace des hyperparamètres. C'est là qu'intervient Keras Tuner, une bibliothèque qui facilite la recherche des meilleurs hyperparamètres pour un modèle de réseau de neurones donné. En utilisant Keras Tuner, il est possible d'automatiser le processus d'optimisation et de trouver les hyperparamètres qui maximisent les performances du modèle dans la détection des attaques DDoS.

L'objectif est de présenter une approche basée sur l'utilisation d'un réseau de neurones profonds optimisé via Keras Tuner pour la détection des attaques DDoS. Nous explorerons les étapes nécessaires à la mise en place de ce modèle, y compris la collecte et le prétraitement des données de trafic réseau, la conception et l'entraînement du réseau de neurones profonds, ainsi que l'optimisation des hyperparamètres à l'aide de Keras Tuner. Enfin, nous évaluerons les performances du modèle obtenu et discuterons de son efficacité et de son applicabilité dans le domaine de la détection des attaques DDoS.

Ce mémoire est organisé en 5 chapitres :

- Dans le premier chapitre, ce dernier est composé de trois parties : les attaques de sécurité, les systèmes d'information (SI) et les méthodes de prévention.
- Dans le deuxième chapitre, nous présentons les IDS (Systèmes de Détection d'Intrusion), le modèle général et la classification des IDS.

- Dans le troisième chapitre, nous approfondissons les concepts du Machine Learning, du Deep Learning et des types de Réseaux de Neurones (RN).
- Dans le quatrième chapitre, nous nous appuyons sur les travaux des algorithmes de la détection d'intrusion.
- Dans le dernier chapitre, nous commençons par la création du modèle, puis nous décrivons les étapes de sa mise en œuvre. Enfin, nous présentons les résultats obtenus.

Comment optimiser l'utilisation d'un réseau de neurones profonds via Keras Tuner pour améliorer la détection des attaques DDoS ?

Chapitre 1 : Sécurité Informatique

1. Introduction

Internet a transformé nos vies à bien des égards. Malheureusement, ce vaste réseau et ses technologies associées ont également entraîné dans leur sillage le nombre croissant de menaces de sécurité. Le moyen le plus efficace de vous protéger contre ces menaces et attaques est de connaître les pratiques de cyber sécurité standard. Ce chapitre sur « Qu'est-ce que la sécurité informatique ? » Présente une introduction à la sécurité informatique et ses concepts clés.

2. Définition

La sécurité informatique est la sécurité appliquée aux dispositifs informatiques tels que ordinateurs et smartphones, ainsi que les réseaux informatiques tels que privés et publics réseaux, y compris l'ensemble de l'Internet. Le domaine couvre tous les processus et mécanismes par quels équipements, informations et services numériques sont protégés contre les l'accès, la modification ou la destruction non autorisés, et revêtent une importance croissante compte tenu de la dépendance croissante vis-à-vis des systèmes informatiques de la plupart des sociétés du monde. Il comprend physique la sécurité pour empêcher le vol d'équipements et la sécurité des informations pour protéger les données qui s'y trouvent. Équipement. Elle est parfois appelée « cybersécurité » ou « sécurité informatique », bien que ces termes ne font généralement pas référence à la sécurité physique (serrures et autres).

3. Les attaques

3.1 Qu'est-ce qu'un cyber attaque ?

Un cyber attaque désigne une action visant à cibler un ordinateur ou tout élément d'un système d'information informatisé pour modifier, détruire ou voler des données, ainsi qu'exploiter ou nuire à un réseau. Les cybers attaques se multiplient, en phase avec la numérisation des entreprises qui est devenue de plus en plus populaire ces dernières années.

3.2 Types d'attaques

Les types d'attaques de cyber sécurité les plus courants :

3.2.1 Attaques DoS et DDoS

Une attaque par déni de service (DoS) est conçue pour submerger les ressources d'un système au point où il est incapable de répondre aux demandes de service légitimes. Une attaque par déni de service distribué (DDoS) est similaire en ce sens qu'elle cherche également à drainer les ressources d'un système. Une attaque DDoS est initiée par un vaste éventail de machines hôtes infectées par des logiciels malveillants contrôlées par l'attaquant. Celles-ci sont appelées attaques de « déni de service » car le site victime est incapable de fournir un service à ceux qui souhaitent y accéder.

Avec une attaque DoS, le site cible est inondé de requêtes illégitimes. Comme le site doit répondre à chaque requête, ses ressources sont consommées par toutes les réponses. Cela rend impossible pour le site de servir les utilisateurs comme il le fait normalement et entraîne souvent un arrêt complet du site.

Les attaques DoS et DDoS sont différentes des autres types de cyberattaques qui permettent au pirate soit d'obtenir l'accès à un système, soit d'augmenter l'accès dont il dispose actuellement. Avec ces types d'attaques, l'attaquant bénéficie directement de ses efforts. Avec les attaques réseau DoS et DDoS, en revanche, l'objectif est simplement d'interrompre l'efficacité du service de la cible. Si l'agresseur est embauché par un concurrent commercial, il peut bénéficier financièrement de ses efforts.

Une attaque DoS peut également être utilisée pour créer une vulnérabilité pour un autre type d'attaque. Avec une attaque DoS ou DDoS réussie, le système doit souvent se déconnecter, ce qui peut le rendre vulnérable à d'autres types d'attaques. Un moyen courant de prévenir les attaques DoS consiste à utiliser un pare-feu qui détecte si les demandes envoyées à votre site sont légitimes. Les demandes d'imposteurs peuvent alors être rejetées, permettant au trafic normal de circuler sans interruption. Un exemple d'attaque Internet majeure de ce type s'est produit en février 2020 contre Amazon Web Services (AWS).

3.2.2 Attaques MITM

Les cyberattaques de type Man-in-the-middle (MITM) font référence à des failles dans la cybersécurité qui permettent à un attaquant d'écouter les données échangées entre deux personnes, réseaux ou ordinateurs. C'est ce qu'on appelle une attaque "man in the middle" parce que l'attaquant se positionne au "milieu" ou entre les deux parties essayant de communiquer. En effet, l'attaquant espionne l'interaction entre les deux parties.

Dans une attaque MITM, les deux parties impliquées ont l'impression de communiquer comme elles le font normalement. Ce qu'ils ne savent pas, c'est que la personne qui envoie réellement le message modifie ou accède illicitement au message avant qu'il n'atteigne sa destination. Certains moyens de vous protéger, vous et votre organisation, contre les attaques MITM consistent à utiliser un cryptage fort sur les points d'accès ou à utiliser un réseau privé virtuel (VPN).

3.2.3 Attaques de phishing

L'hameçonnage est une attaque d'ingénierie sociale qui vise à exploiter les vulnérabilités du système opérations causées par les utilisateurs du système . Les attaques de phishing commencent par envoyer le escroc par hameçonnage un e-mail qui semble provenir d'une organisation réelle, contenant des liens qui cliquer dessus pourrait soit diriger la victime vers de fausses pages Web où l'utilisateur est invité à fournir leurs informations d'identification ou d'installer des logiciels espions sur l'appareil. Le mobile de l'agresseur derrière ces escroqueries peuvent être le vol d'identité, le gain financier ou la notoriété illustré le cycle de l'hameçonnage.

Plusieurs approches heuristiques telles que les améliorations technologiques, le processus

L'ingénierie et la formation des utilisateurs sont suivies pour résoudre le problème du phishing .

3.2.4 Attaques de hameçonnage de baleines

Une attaque de hameçonnage à la baleine est ainsi nommée parce qu'elle s'attaque aux «gros poissons» ou aux baleines d'une organisation, qui comprennent généralement les membres de la suite C ou d'autres personnes en charge de l'organisation. Ces personnes sont susceptibles de posséder des informations qui peuvent être précieuses pour les attaquants, telles que des informations exclusives sur l'entreprise ou ses opérations.

Si une « baleine » ciblée télécharge un rançongiciel, elle est plus susceptible de payer la rançon pour empêcher que la nouvelle de l'attaque réussie ne soit diffusée et ne nuise à sa réputation ou à celle de l'organisation. Les attaques de Whale-phishing peuvent être évitées en prenant les mêmes types de précautions pour éviter les attaques de phishing, telles que l'examen attentif des e-mails, des pièces jointes et des liens qui les accompagnent, en gardant un œil sur les destinations ou les paramètres suspects.

3.2.5 Attaques de harponnage

Le spear phishing fait référence à un type spécifique d'attaque de phishing ciblée. L'attaquant prend le temps de rechercher les cibles qu'il vise, puis d'écrire des messages que la cible est susceptible de trouver personnellement pertinents. Ces types d'attaques sont appelés à juste titre hameçonnage « harpon » en raison de la façon dont l'attaquant se concentre sur une cible spécifique. Le message semblera légitime, c'est pourquoi il peut être difficile de repérer une attaque de harponnage.

Souvent, une attaque de harponnage utilise l'usurpation d'e-mail, où les informations contenues dans la partie "De" de l'e-mail sont falsifiées, ce qui donne l'impression que l'e-mail provient d'un expéditeur différent. Il peut s'agir d'une personne en qui la cible a confiance, comme une personne de son réseau social, un ami proche ou un partenaire commercial. Les attaquants peuvent également utiliser le clonage de sites Web pour donner l'impression que la communication est légitime. Avec le clonage de site Web, l'attaquant copie un site Web légitime pour endormir la victime dans un sentiment de confort. La cible, pensant que le site Web est réel, se sent alors à l'aise de saisir ses informations privées.

Semblables aux attaques de phishing classiques, les attaques de spear-phishing peuvent être évitées en vérifiant soigneusement les détails dans tous les champs d'un e-mail et en s'assurant que les utilisateurs ne cliquent sur aucun lien dont la destination ne peut pas être vérifiée comme légitime.

3.2.6 Rançongiciels

Avec les rançongiciels, le système de la victime est retenu en otage jusqu'à ce qu'elle accepte de payer une rançon à l'attaquant. Une fois le paiement envoyé, l'attaquant fournit alors des instructions sur la façon dont la cible peut reprendre le contrôle de son ordinateur. Le nom "ransomware" est approprié car le malware demande une rançon à la victime.

Lors d'une attaque par rançongiciel, la cible télécharge un rançongiciel, soit à partir d'un site Web, soit à partir d'une pièce jointe à un e-mail. Le logiciel malveillant est écrit pour exploiter des vulnérabilités qui n'ont pas été résolues par le fabricant du système ou l'équipe informatique. Le rançongiciel crypte ensuite le poste de travail de la cible. Parfois, les rançongiciels peuvent être utilisés pour attaquer plusieurs parties en refusant l'accès à plusieurs ordinateurs ou à un serveur central essentiel aux opérations commerciales.

Affecter plusieurs ordinateurs est souvent accompli en n'initiant la captation des systèmes que des jours, voire des semaines après la pénétration initiale du logiciel malveillant. Le logiciel malveillant peut envoyer des fichiers AUTORUN qui vont d'un système à un autre via le réseau interne ou des lecteurs Universal Serial Bus (USB) qui se connectent à plusieurs ordinateurs. Ensuite, lorsque l'attaquant lance le cryptage, il fonctionne simultanément sur tous les systèmes infectés.

Dans certains cas, les auteurs de rançongiciels conçoivent le code pour échapper aux logiciels antivirus traditionnels. Il est donc important que les utilisateurs restent vigilants quant aux sites qu'ils visitent et aux liens sur lesquels ils cliquent. Vous pouvez également empêcher de nombreuses attaques de ransomwares en utilisant un pare-feu de nouvelle génération (NGFW) qui peut effectuer des inspections approfondies des paquets de données à l'aide de l'intelligence artificielle (IA) qui recherche les caractéristiques des ransomwares.

3.2.7 Attaque par mot de passe

L'authentification des utilisateurs avec des mots de passe est une méthode bien connue pour accéder à leurs informations sur les sites Web

De nombreuses personnes utilisent un dictionnaire de mots pour créer des mots de passe qui seront stockés sur bases de données sous forme de valeurs de hachage au lieu de texte brut. Certains utilisateurs choisissent un mot de passe court pour

s'en souvenir facilement. En conséquence, les pirates peuvent deviner ces mots de passe en utilisant la force brute, la table arc-en-ciel et attaques par dictionnaire

3.2.8 Attaque par injection SQ L

L'attaque par injection SQL (SQLIA) est l'une des attaques de base de données les plus courantes donne un accès illimité aux attaquants pour extraire les données sensibles stockées dans la base de données . Les attaquants utilisent des commandes SQL simples pour restructurer le code SQL et exécuter des code dans l'application Web . De nombreuses technologies sont vulnérables aux attaques SQL telles que PHP, JSP, ASP, ASP.net. Dans

De plus, les formulaires Web sont également vulnérables à ces types d'attaques telles que les formulaires de connexion, commentaires, support client, paniers d'achatL

3.2.9 Interprétation des URL

Avec l'interprétation d'URL, les attaquants modifient et fabriquent certaines adresses URL et les utilisent pour accéder aux données personnelles et professionnelles de la cible. Ce type d'attaque est également appelé empoisonnement d'URL. Le nom "interprétation d'URL" vient du fait que l'attaquant connaît l'ordre dans lequel les informations d'URL d'une page Web doivent être saisies. L'attaquant "interprète" ensuite cette syntaxe, l'utilisant pour comprendre comment accéder à des zones auxquelles il n'a pas accès.

Pour exécuter une attaque par interprétation d'URL, un pirate peut deviner les URL qu'il peut utiliser pour obtenir des privilèges d'administrateur sur un site ou pour accéder au back-end du site afin d'accéder au compte d'un utilisateur. Une fois sur la page qu'ils souhaitent, ils peuvent manipuler le site lui-même ou accéder à des informations sensibles sur les personnes qui l'utilisent.

Par exemple, si un pirate tente d'accéder à la section d'administration d'un site appelé GetYourKnowledgeOn.com, il peut taper http://getyourknowledgeon.com/admin, ce qui l'amènera à une page de connexion d'administrateur. Dans certains cas, le nom d'utilisateur et le mot de passe de l'administrateur peuvent être "admin" et "admin" par défaut ou très faciles à deviner. Un attaquant peut également avoir déjà trouvé le mot de passe de l'administrateur ou l'avoir réduit à quelques possibilités. L'attaquant essaie ensuite chacun d'entre eux, y accède et peut manipuler, voler ou supprimer des données à volonté.

Pour empêcher les attaques par interprétation d'URL de réussir, utilisez des méthodes d'authentification sécurisées pour toutes les zones sensibles de votre site. Cela peut nécessiter une authentification multi facteur (MFA) ou des mots de passe sécurisés composés de caractères apparemment aléatoires.

3.2.10 Usurpation DNS

Avec l'usurpation du système de noms de domaine (DNS), un pirate modifie les enregistrements DNS pour envoyer du trafic vers un site Web faux ou "usurpé". Une fois sur le site frauduleux, la

victime peut saisir des informations sensibles pouvant être utilisées ou revendues par le pirate. Le pirate informatique peut également construire un site de mauvaise qualité avec un contenu désobligeant ou incendiaire pour donner une mauvaise image d'une entreprise concurrente.

Dans une attaque d'usurpation de DNS, l'attaquant profite du fait que l'utilisateur pense que le site qu'il visite est légitime. Cela donne à l'attaquant la possibilité de commettre des crimes au nom d'une société innocente, du moins du point de vue du visiteur.

Pour éviter l'usurpation de DNS, assurez-vous que vos serveurs DNS sont à jour. Les attaquants visent à exploiter les vulnérabilités des serveurs DNS, et les versions logicielles les plus récentes contiennent souvent des correctifs qui corrigent les vulnérabilités connues.

3.2.11 Piratage de session

Le détournement de session est l'un des nombreux types d'attaques MITM. L'attaquant prend le contrôle d'une session entre un client et le serveur. L'ordinateur utilisé dans l'attaque substitue son adresse IP (Internet Protocol) à celle de l'ordinateur client, et le serveur poursuit la session sans se douter qu'il communique avec l'attaquant au lieu du client. Ce type d'attaque est efficace car le serveur utilise l'adresse IP du client pour vérifier son identité. Si l'adresse IP de l'attaquant est insérée au cours de la session, le serveur peut ne pas suspecter une violation car il est déjà engagé dans une connexion sécurisée.

Pour empêcher le piratage de session, utilisez un VPN pour accéder aux serveurs critiques de l'entreprise. De cette façon, toutes les communications sont cryptées et un attaquant ne peut pas accéder au tunnel sécurisé créé par le VPN.

3.2.12 Attaque par force brute

Une attaque par force brute tire son nom de la méthodologie « brutale » ou simple employée par l'attaque. L'attaquant essaie simplement de deviner les identifiants de connexion d'une personne ayant accès au système cible. Une fois qu'ils ont bien compris, ils sont dedans.

3.2.13 Attaques Web

Les attaques Web font référence aux menaces qui ciblent les vulnérabilités des applications Web. Chaque fois que vous entrez des informations dans une application Web, vous lancez une commande qui génère une réponse. Par exemple, si vous envoyez de l'argent à quelqu'un à l'aide d'une application bancaire en ligne, les données que vous saisissez indiquent à l'application d'accéder à votre compte, de retirer de l'argent et de l'envoyer sur le compte de quelqu'un d'autre. Les attaquants travaillent dans le cadre de ce type de requêtes et les utilisent à leur avantage.

Certaines attaques Web courantes incluent l'injection SQL et les scripts intersites (XSS), qui seront abordés plus loin dans cet article. Les pirates utilisent également des attaques de falsification de requête intersite (CSRF) et la falsification de paramètres. Dans une attaque CSRF, la victime est amenée à effectuer une action qui profite à l'attaquant. Par exemple, ils peuvent cliquer sur quelque chose qui lance un script conçu pour modifier les identifiants de connexion pour accéder à une application Web. Le pirate, armé des nouveaux identifiants de connexion, peut alors se connecter comme s'il était l'utilisateur légitime.

La falsification des paramètres consiste à ajuster les paramètres que les programmeurs implémentent en tant que mesures de sécurité conçues pour protéger des opérations spécifiques. L'exécution de l'opération dépend de ce qui est entré dans le paramètre. L'attaquant modifie simplement les paramètres, ce qui lui permet de contourner les mesures de sécurité qui dépendaient de ces paramètres.

Pour éviter les attaques Web, inspectez vos applications Web pour rechercher et corriger les vulnérabilités. Une façon de corriger les vulnérabilités sans affecter les performances de l'application Web consiste à utiliser des jetons anti-CSRF. Un jeton est échangé entre le navigateur de l'utilisateur et l'application Web. Avant l'exécution d'une commande, la validité du jeton est vérifiée. S'il vérifie, la commande passe - sinon, elle est bloquée. Vous pouvez également utiliser les drapeaux SameSite, qui autorisent uniquement le traitement des demandes provenant du même site, rendant tout site construit par l'attaquant impuissant

3.2.14 Menaces internes

Parfois, les acteurs les plus dangereux viennent de l'intérieur d'une organisation. Les personnes se trouvant à l'intérieur d'une entreprise représentent un danger particulier car elles ont généralement accès à une variété de systèmes et, dans certains cas, à des privilèges d'administrateur qui leur permettent d'apporter des modifications critiques au système ou à ses politiques de sécurité.

De plus, les personnes au sein de l'organisation ont souvent une compréhension approfondie de son architecture de cybersécurité, ainsi que de la façon dont l'entreprise réagit aux menaces. Ces connaissances peuvent être utilisées pour accéder à des zones restreintes, modifier les paramètres de sécurité ou déduire le meilleur moment possible pour mener une attaque.

L'un des meilleurs moyens de prévenir les menaces internes dans les organisations consiste à limiter l'accès des employés aux systèmes sensibles à ceux qui en ont besoin pour accomplir leurs tâches. De plus, pour les quelques privilégiés qui ont besoin d'un accès, utilisez MFA, ce qui les obligera à utiliser au moins une chose qu'ils connaissent en conjonction avec un élément physique dont ils disposent pour accéder à un système sensible. Par exemple, l'utilisateur peut avoir à entrer

un mot de passe et insérer un périphérique USB. Dans d'autres configurations, un numéro d'accès est généré sur un appareil portable auquel l'utilisateur doit se connecter. L'utilisateur ne peut accéder à la zone sécurisée que si le mot de passe et le numéro sont corrects.

Bien que la MFA n'empêche pas toutes les attaques à elle seule, elle permet de déterminer plus facilement qui est à l'origine d'une attaque - ou d'une tentative - en particulier parce que relativement peu de personnes ont accès aux zones sensibles en premier lieu. Par conséquent, cette stratégie d'accès limité peut avoir un effet dissuasif. Les cybercriminels au sein de votre organisation sauront qu'il est facile d'identifier l'auteur en raison du nombre relativement restreint de suspects potentiels.

3.2.15 Chevaux de Troie

Une attaque de cheval de Troie utilise un programme malveillant qui est caché à l'intérieur d'un programme apparemment légitime. Lorsque l'utilisateur exécute le programme présumé innocent, le logiciel malveillant à l'intérieur du cheval de Troie peut être utilisé pour ouvrir une porte dérobée dans le système par laquelle les pirates peuvent pénétrer dans l'ordinateur ou le réseau. Cette menace tire son nom de l'histoire des soldats grecs qui se sont cachés à l'intérieur d'un cheval pour infiltrer la ville de Troie et gagner la guerre. Une fois que le "cadeau" a été accepté et amené aux portes de Troie, les soldats grecs ont sauté et ont attaqué. De la même manière, un utilisateur peu méfiant peut accueillir une application d'apparence innocente dans son système uniquement pour inaugurer une menace cachée.

Pour prévenir les attaques de chevaux de Troie, les utilisateurs doivent être informés de ne pas télécharger ou installer quoi que ce soit à moins que sa source ne puisse être vérifiée. En outre, les NGFW peuvent être utilisés pour examiner les paquets de données à la recherche de menaces potentielles de chevaux de Troie.

3.2.16 Attaques au volant

Lors d'une attaque au volant, un pirate informatique intègre un code malveillant dans un site Web non sécurisé. Lorsqu'un utilisateur visite le site, le script est automatiquement exécuté sur son ordinateur, l'infectant. L'appellation "drive by" vient du fait que la victime n'a qu'à "passer" le site en le visitant pour être infectée. Il n'est pas nécessaire de cliquer sur quoi que ce soit sur le site ou de saisir des informations.

Pour se protéger contre les attaques au volant, les utilisateurs doivent s'assurer qu'ils exécutent le logiciel le plus récent sur tous leurs ordinateurs, y compris des applications telles qu'Adobe Acrobat et Flash, qui peuvent être utilisées lors de la navigation sur Internet. En outre, vous pouvez

utiliser un logiciel de filtrage Web, qui peut détecter si un site est dangereux avant qu'un utilisateur ne le visite.

3.2.17 Attaques d'écoute clandestine

Les attaques d'écoute clandestine impliquent que le mauvais acteur intercepte le trafic lorsqu'il est envoyé sur le réseau. De cette manière, un attaquant peut collecter des noms d'utilisateur, des mots de passe et d'autres informations confidentielles telles que des cartes de crédit. L'écoute peut être active ou passive.

Avec l'écoute clandestine active, le pirate insère un logiciel dans le chemin du trafic réseau pour collecter des informations que le pirate analyse pour des données utiles. Les attaques d'écoute passive sont différentes en ce sens que le pirate « écoute » ou écoute clandestinement les transmissions, à la recherche de données utiles qu'il peut voler.

Les écoutes clandestines actives et passives sont des types d'attaques MITM. L'un des meilleurs moyens de les prévenir consiste à crypter vos données, ce qui empêche qu'elles ne soient utilisées par un pirate, qu'il utilise une écoute active ou passive.

3.2.18 Attaque d'anniversaire

Dans une attaque d'anniversaire, un attaquant abuse d'une fonctionnalité de sécurité : les algorithmes de hachage, qui sont utilisés pour vérifier l'authenticité des messages. L'algorithme de hachage est une signature numérique et le destinataire du message la vérifie avant d'accepter le message comme authentique. Si un pirate peut créer un hachage identique à celui que l'expéditeur a ajouté à son message, le pirate peut simplement remplacer le message de l'expéditeur par le sien. L'appareil récepteur l'acceptera car il a le bon hachage.

Le nom « birthday attack » fait référence au paradoxe de l'anniversaire, qui repose sur le fait que dans une pièce de 23 personnes, il y a plus de 50 % de chances que deux d'entre elles aient le même anniversaire. Par conséquent, alors que les gens pensent que leurs anniversaires, comme les hachages, sont uniques, ils ne sont pas aussi uniques que beaucoup le pensent.

Pour éviter les attaques d'anniversaire, utilisez des hachages plus longs pour la vérification. Avec chaque chiffre supplémentaire ajouté au hachage, les chances d'en créer un correspondant diminuent considérablement.

3.2.19 Attaque de logiciels malveillants

Malware est un terme général pour les logiciels malveillants, d'où le "mal" au début du mot. Les logiciels malveillants infectent un ordinateur et modifient son fonctionnement, détruisent des

données ou espionnent l'utilisateur ou le trafic réseau lors de son passage. Les logiciels malveillants peuvent soit se propager d'un appareil à un autre, soit rester en place, n'affectant que son appareil hôte.

Plusieurs des méthodes d'attaque décrites ci-dessus peuvent impliquer des formes de logiciels malveillants, notamment les attaques MITM, le phishing, les ransomwares, l'injection SQL, les chevaux de Troie, les attaques au volant et les attaques XSS.

Lors d'une attaque de logiciel malveillant, le logiciel doit être installé sur l'appareil cible. Cela nécessite une action de la part de l'utilisateur. Par conséquent, en plus d'utiliser des pare-feu capables de détecter les logiciels malveillants, les utilisateurs doivent être informés des types de logiciels à éviter, des types de liens qu'ils doivent vérifier avant de cliquer, ainsi que des e-mails et des pièces jointes avec lesquels ils ne doivent pas interagir. [1]

4. Sécurité Informatique

4.1 Définition :

La sécurité informatique est la sécurité appliquée aux dispositifs informatiques tels que ordinateurs et smartphones, ainsi que les réseaux informatiques tels que privés et publics réseaux, y compris l'ensemble de l'Internet. Le domaine couvre tous les processus et mécanismes par quels équipements, informations et services numériques sont protégés contre les l'accès, la modification ou la destruction non autorisés, et revêtent une importance croissante compte tenu de la dépendance croissante vis-à-vis des systèmes informatiques de la plupart des sociétés du monde. Il comprend physique la sécurité pour empêcher le vol d'équipements et la sécurité des informations pour protéger les données qui s'y trouvent. équipement. Elle est parfois appelée « cybersécurité » ou « sécurité informatique », bien que ces termes ne font généralement pas référence à la sécurité physique (serrures et autres). [2]

4.2 Les bonnes pratiques pour assurer la sécurité :

*Mettre constamment à jour votre système d'exploitation et vos applications.

Utilisez un antivirus.

*Utilisez des mots de passe sécurisés (forts).

*Ne pas ouvrir les pièces jointes aux e-mails d'expéditeurs inconnus

*Cliquer sur des liens dans des e-mails provenant de sites Web inconnus ou d'expéditeurs inconnus n'est pas recommandé.

*N'utilisez pas de réseaux Wi-Fi non sécurisés dans des lieux publics.

5 . Les méthodes de la prévention

Il y a un avis séculaire qui dit: « Il est trop tard pour aiguiser votre épée quand le tambour bat pour la bataille ». Ne vous méprenez pas, nous sommes en guerre et nous devons nous préparer à les cyber-batailles en aiguisant nos compétences. Les professionnels de la sécurité de l'information doivent perfectionner continuellement leurs capacités en travaillant plus intelligemment et non plus durement. C'est toujours mieux prévenir, puis poursuivre et poursuivre. La prévention d'un incident nécessite une attention particulière analyse et planification.

L'information est un bien qui doit être protégé à la mesure de sa valeur.

Des mesures de sécurité doivent être prises pour protéger les informations contre la modification, destruction ou divulgation accidentelle ou intentionnelle. Pendant la phase de prévention, les politiques, contrôles et processus de sécurité doivent être conçus et mis en œuvre. Politiques de sécurité, programmes de sensibilisation à la sécurité et contrôle d'accès procédures, sont toutes interdépendantes et doivent être élaborées dès le début. L'information la politique de sécurité est la pierre angulaire à partir de laquelle tout le reste est construit. [3]

5.1 Serveur proxy :

Le serveur proxy est un serveur intermédiaire entre le client et l'interne. Les serveurs proxy offrent la fonctionnalités de base suivantes :

- Pare-feu et filtrage des données réseau.
- Partage de connexion réseau
- Mise en cache des données

5.1.1 Objectif des serveurs proxy :

Voici les raisons d'utiliser des serveurs proxy :

- **Surveillance et filtrage**

Les serveurs proxy nous permettent de faire plusieurs types de filtrage tels que :

Filtrage du contenu

Filtrage des données chiffrées

Filtres de contournement

Journalisation et écoute clandestine

- **Amélioration des performances**

Il fixe le service en récupérant le contenu du cache qui a été enregistré lorsque

demande précédente a été faite par le client.

- **Traduction**

Il permet de personnaliser le site source pour les utilisateurs locaux en excluant le contenu source ou en remplaçant contenu source avec un contenu local original. Dans ce cas, le trafic des utilisateurs globaux est acheminé vers le

site Web source via le proxy de traduction.

- **Accéder aux services de manière anonyme**

En cela, le serveur de destination reçoit la demande du serveur proxy anonymisation et donc ne reçoit pas d'informations sur l'utilisateur final.

- **La Sécurité**

Étant donné que le serveur proxy cache l'identité de l'utilisateur, il protège donc du spam et du pirate attaques. [3]

5.1.2 Les Type de proxy :

- **Proxy de transfert : Forward Proxies**

Dans ce cas, le client demande à son serveur de réseau interne de le transmettre à Internet.

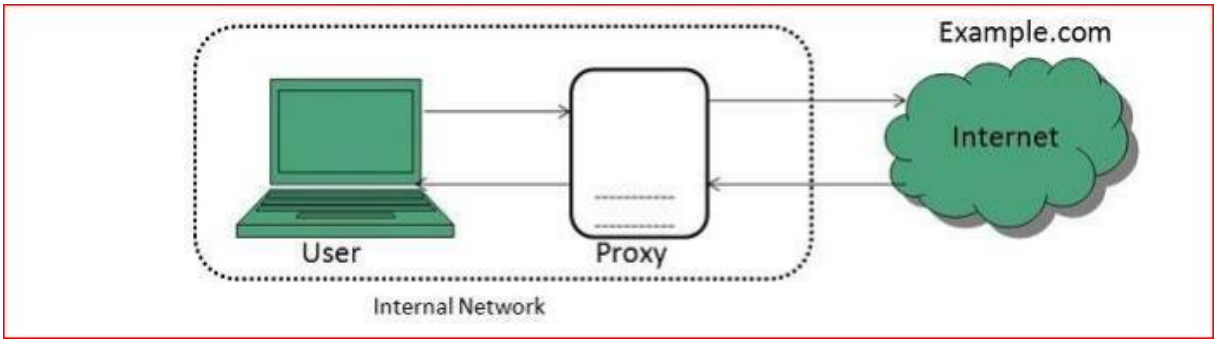

Figure 1 : Proxy de transfert [3]

- **Procurations ouvertes : open proxy**

Open Proxies aide les clients à dissimuler leur adresse IP tout en naviguant sur le Web

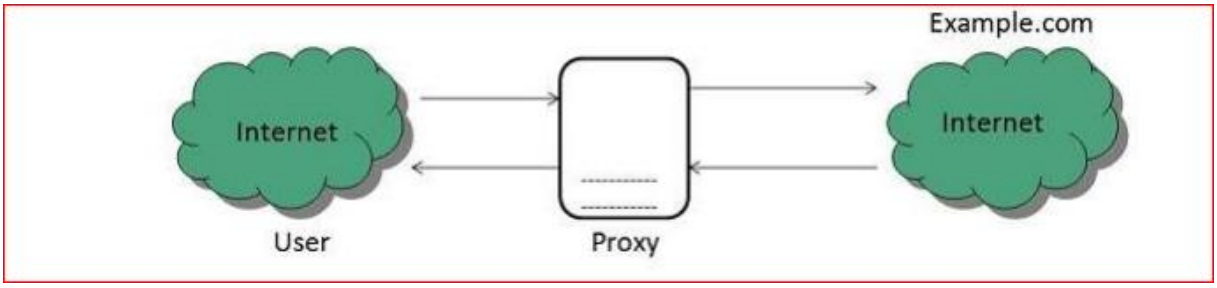

Figure 2 : Procurations ouvertes [3]

- **Proxy inverses**

Dans ce cas, les demandes sont transmises à un ou plusieurs serveurs proxy et la réponse du proxy serveur est récupéré comme s'il provenait directement du serveur d'origine.

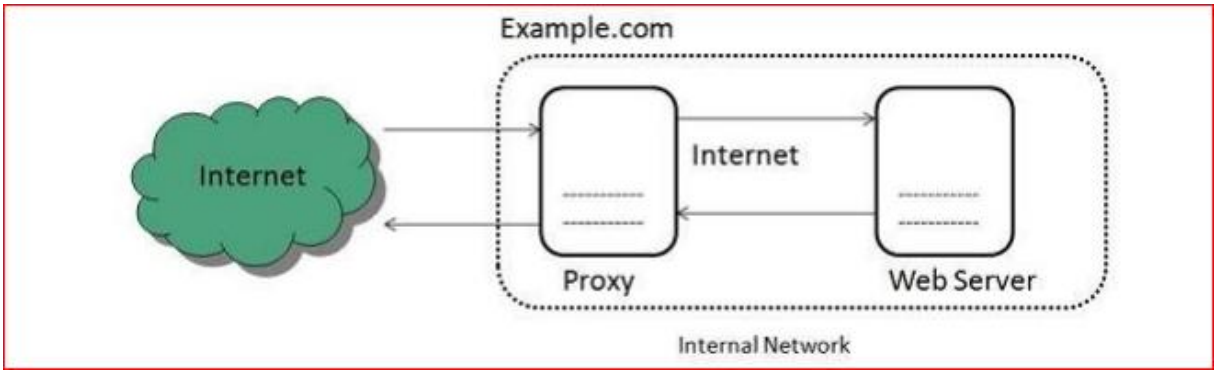

Figure 3 : Proxy inverses [3]

5.1.3 L'Architecture

L'architecture du serveur proxy est divisée en plusieurs modules comme illustré dans le schéma suivant :

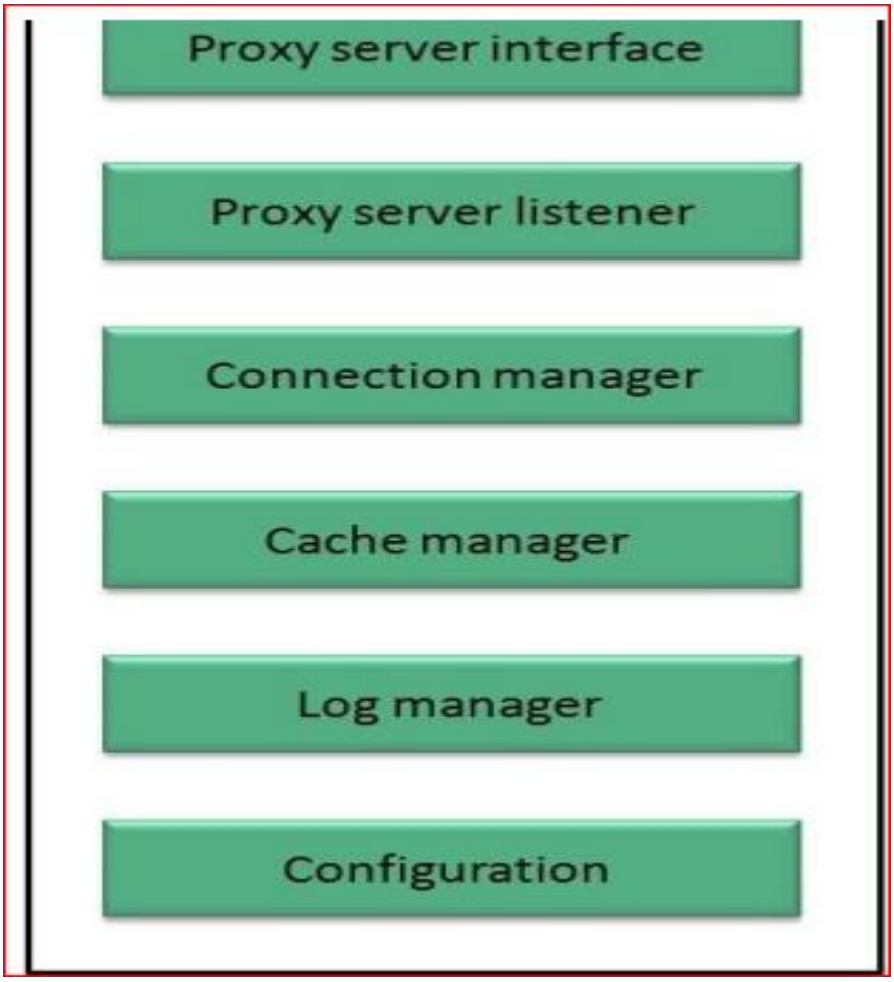

Figure 4 : Architecture de serveur proxy [3]

5.2 Par feu :

5.2.1 Définition :

C'est un ensemble de différents composants matériels (physique) et logiciels (logique) qui contrôlent le trafic intérieur/extérieur selon une politique de sécurité. Un système pare-feu fonctionne la plupart du temps grâce à des règles de filtrage indiquant les adresses IP autorisées à communiquer avec les machines aux réseaux, il s'agit ainsi d'une passerelle filtrante. Il permet d'une part de bloquer des attaques ou connexions suspectes d'accéder au réseau interne. D'un autre côté, un firewall sert dans de nombreux cas également à éviter la fuite non contrôlée d'informations vers l'extérieur. Il propose un véritable contrôle sur le trafic réseau de l'entreprise, Il permet donc d'analyser, de sécuriser et de gérer le trafic réseau [4]

Figure 5 : le fonctionnement d'un pare-feu. [5]

5.2.2 Les deux types de pare-feu

Deux types de firewall existent : le pare-feu matériel et le pare-feu logiciel. En fonction de la situation, il est possible d'installer l'un ou l'autre, ou de cumuler les deux pour accroître la sécurité du réseau.

- Le pare-feu matériel. Ce type de firewall est installé à l'entrée et à la sortie du réseau local. Son installation est généralement plus coûteuse que le firewall logiciel, mais il garantit davantage de protection en termes de sécurité. Cette solution est notamment privilégiée pour les réseaux comportant plusieurs ordinateurs, par exemple dans le cadre de sociétés privées (le pare-feu matériel se révèle alors moins onéreux qu'un pare-feu logiciel, et il assure une plus grande protection pour le réseau).
- Le pare-feu logiciel. Installé directement sur l'ordinateur, le pare-feu logiciel joue un rôle similaire au pare-feu matériel mais de façon locale. Il contrôle les paquets de données entrants et sortants et peut les bloquer si nécessaire. Son prix est moins élevé qu'un pare-feu matériel, et son utilisation est privilégiée lorsqu'il s'agit de protéger uniquement un ordinateur.

5.2.3 Les différents types de filtrage

Il existe aujourd'hui différents types de filtrage, ayant chacun un rôle différent :

- Le pare-feu à états. Il vérifie que chaque paquet est conforme à une connexion en cours. Il s'assure donc que le paquet est bien la suite d'un précédent paquet, et la réponse à un paquet dans le sens inverse.
- Le pare-feu sans état. Il contrôle séparément chaque paquet en vérifiant qu'il répond aux règles définies.
- Le pare-feu applicatif. Aussi appelé proxy, le filtrage applicatif joue le rôle de filtre au niveau applicatif. Ce firewall contrôle la conformité complète du paquet à un protocole attendu.

- Le pare-feu identifiant. Il associe les utilisateurs et l'IP, pour suivre l'activité réseau de chaque utilisateur.
- Le pare-feu personnel. Ce firewall est installé directement sur un ordinateur et agit comme un pare-feu à états pour lutter contre les virus informatiques. [6]

5.3 Antivirus :

5.3.1 Définition :

Antivirus Il s'agit d'un logiciel capable de détecter et de détruire les virus contenus sur un disque. Le logiciel a pour charge de surveiller la présence de virus et éventuellement de nettoyer, supprimer ou mettre en quarantaine le ou les fichiers infectés. Ils surveillent tous les espaces dans lesquels un virus peut se loger, c'est à dire la mémoire et les unités de stockage qui peuvent être locales ou réseau. [7]

Figure 6 : Antivirus fonctionnement [7]

5.3.2 Fonctionnement d'un antivirus

Le programme est composé de 3 parties ayant chacune un rôle essentiel :

- Un " moteur " qui a pour rôle la détection des virus.

- Une base de données contenant des informations sur les virus connus. C'est cette base de données qu'il faut maintenir à jour le plus régulièrement possible, afin de permettre à l'antivirus de connaître les virus les plus récents.

- Un module de nettoyage qui a pour but de traiter le fichier infecté. A chaque fichier testé, si le programme pense voir un virus, il regarde dans sa base de données si le virus est connu (chaque virus ainsi que ses variantes a une signature particulière, et c'est cette signature qui est comparée avec la base). Si le virus est connu, il y a de fortes chances qu'un antidote soit connu.

• Si le virus n'est pas connu, le logiciel emploie une méthode heuristique (Technique consistant à apprendre petit à petit, en tenant compte de ce que l'on a fait précédemment pour tendre vers la solution d'un problème. L'heuristique ne garantit pas du tout que l'on arrive à une solution satisfaisante. Opposé à l'algorithmique, l'heuristique est essentiellement utilisée dans les antivirus, pour détecter des virus en les reconnaissant selon ce qu'ils sont capables de faire plutôt que selon leur signature) qui recherche une activité anormale ressemblant à celle d'un virus. Si tel est le cas, il met le programme infecté en quarantaine et affiche un message. Si le virus n'apparaît plus (parce qu'il est boggué et qu'il se réplique mal ou qu'il se détériore), les éditeurs d'antivirus le cataloguent comme «dormant » [8]

5.3.3 Les techniques de détection utilisées par un antivirus :

Les antivirus utilisent principalement cinq méthodes pour détecter les virus :

- Recherche par signature
- Recherche heuristique
- Analyse spectrale
- Contrôleur d'intégrité
- Moniteur de comportement [8]

5.4 IPSEC

5.4.1 Définition

IPsec (Internet Protocol Security) est un ensemble de protocoles de sécurité utilisé pour sécuriser les communications sur un réseau IP. Il fournit des fonctionnalités de confidentialité, d'intégrité et d'authentification des paquets IP.

IPsec est principalement utilisé pour sécuriser les communications sur des réseaux publics tels qu'Internet. Il permet de créer des tunnels virtuels sécurisés entre des points finaux, tels que des routeurs, des passerelles ou des hôtes, afin de protéger les données qui transitent entre eux [9]

5.4.2 Comment les utilisateurs se connectent-ils à un VPN IPsec ?

Les utilisateurs peuvent accéder à un VPN IPsec en se connectant à une application VPN, ou au client "." Cela nécessite généralement que l'utilisateur ait installé l'application sur son appareil.

Les connexions à un VPN sont généralement basées sur un mot de passe. Bien que les données envoyées par un VPN soient cryptées, si les mots de passe des utilisateurs sont compromis, les attaquants peuvent se connecter au VPN et voler ces données cryptées. L'utilisation de

l'authentification à deux facteurs (2FA) peut renforcer la sécurité des VPN IPsec, puisque le vol d'un mot de passe seul ne donnera plus accès à un attaquant. [9]

5.4.3 Comment fonctionne l'IPsec ?

Les connexions IPsec comprennent les étapes suivantes :

- **Échange de clés :** Les clés sont nécessaires au cryptage ; une clé est une chaîne de caractères aléatoires qui peut être utilisée pour "verrouiller" (crypter) et "déverrouiller" (décrypter) les messages. IPsec établit des clés avec un échange de clés entre les appareils connectés, afin que chaque appareil puisse décrypter les messages de l'autre.
- **En-têtes et remorques de paquets :** Toutes les données envoyées sur un réseau sont décomposées en petits morceaux appelés paquets. Les paquets contiennent à la fois une charge utile, ou les données réelles envoyées, et des en-têtes, ou des informations sur ces

  données afin que les ordinateurs qui reçoivent les paquets sachent quoi en faire. L'IPsec ajoute aux paquets de données plusieurs en-têtes contenant des informations d'authentification et de cryptage. IPsec ajoute également des trailers, qui sont placés après la charge utile de chaque paquet au lieu d'être placés avant.

- **Authentification :** IPsec fournit une authentification pour chaque paquet, comme un cachet d'authenticité sur un objet de collection. Cela garantit que les paquets proviennent d'une source de confiance et non d'un attaquant.
- **Cryptage :** IPsec crypte les charges utiles de chaque paquet et l'en-tête IP de chaque paquet (sauf si le mode transport est utilisé au lieu du mode tunnel - voir ci-dessous). Les données envoyées par IPsec restent ainsi sécurisées et privées.
- **Transmission :** Les paquets IPsec chiffrés traversent un ou plusieurs réseaux jusqu'à leur destination en utilisant un protocole de transport. À ce stade, le trafic IPsec diffère du trafic IP ordinaire en ce qu'il utilise le plus souvent UDP comme protocole de transport, plutôt que TCP. TCP, le protocole de contrôle de transmission, établit des connexions dédiées entre les périphériques et garantit l'arrivée de tous les paquets. Le protocole UDP, User Datagram Protocol, n'établit pas ces connexions dédiées. IPsec utilise UDP car cela permet aux paquets IPsec de traverser les pare-feu.
- **Décryptage :** À l'autre bout de la communication, les paquets sont décryptés et les applications (par exemple, un navigateur) peuvent maintenant utiliser les données fournies. [9]

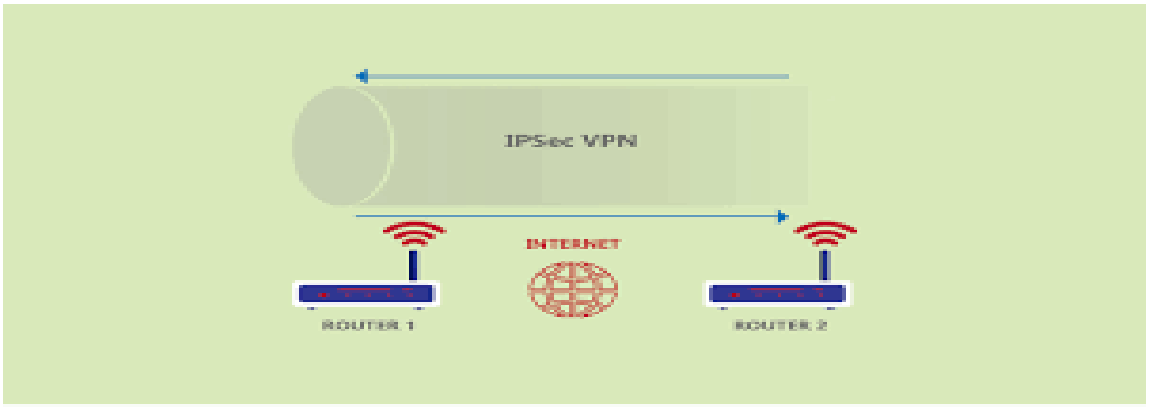

Figure 7 : le fonctionnement de VPN IPsec [10]

5.5 La cryptographie :

5.5.1 Définition :

La cryptographie est la science qui utilise les mathématiques pour chiffrer et déchiffrer des données. La cryptographie vous permet de stocker des informations sensibles ou de les transmettre sur des réseaux non sécurisés (comme Internet) afin qu'elles ne puissent être lues par personne d'autre que le destinataire.

Alors que la cryptographie est la science de la sécurisation des données, la cryptanalyse est la science de l'analyse et de la rupture des communications sécurisées. La cryptanalyse classique implique une combinaison intéressante de raisonnement analytique, d'application d'outils mathématiques, de modèles découverte, patience, détermination et chance. Les cryptanalystes sont aussi appelés attaquants.

La cryptologie englobe à la fois la cryptographie et la cryptanalyse. [11]

5.5.2 Comment fonctionne la cryptographie ?

Un algorithme cryptographique, ou chiffrement, est une fonction mathématique utilisée dans le processus de chiffrement et de déchiffrement. Un algorithme cryptographique fonctionne en combinaison avec un key (un mot, un nombre ou une phrase) pour chiffrer le texte en clair. Le même texte clair crypte un texte chiffré différent avec des clés différentes. La sécurité des données cryptées et dépend entièrement de deux choses : la force de l'algorithme cryptographique et le secret de la clé.

Un algorithme cryptographique, plus toutes les clés possibles et tous les protocoles qui le composent travail, comprennent un crypto système. PGP est un crypto système

Dans la cryptographie conventionnelle, également appelée chiffrement à clé secrète ou à clé symétrique, La clé est utilisée à la fois pour le chiffrement et le déchiffrement. La norme de chiffrement des données(DES) est un exemple de crypto système conventionnel largement utilisé par les Gouvernement des États-Unis. [11]

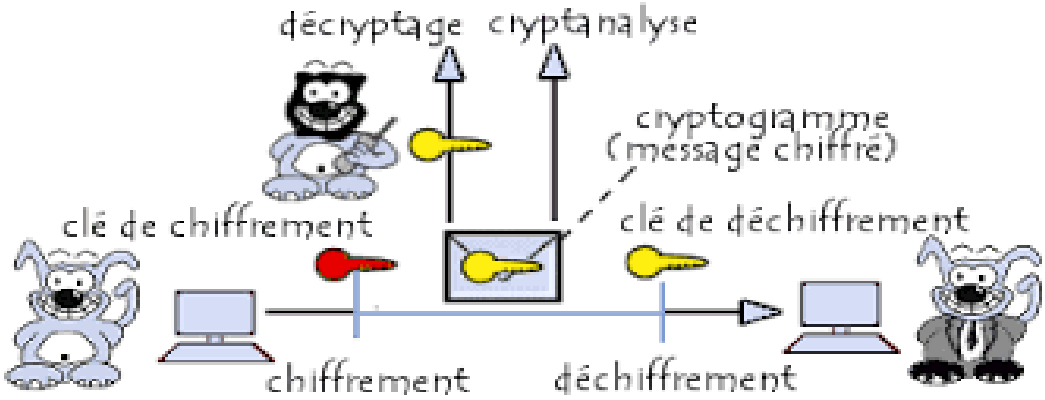

Figure 8 : Cryptographie [12]

6. Conclusion :

Internet a rendu notre monde plus petit à bien des égards, mais il a également ouvert

Jusqu'à un impact qui n'a jamais été aussi divers et aussi difficile. Avec sécurité

Croissance rapide, le monde du piratage a grandi encore plus vite.

Les méthodes de la prévention ne sont pas suffisantes et il faut utiliser une seconde ligne de défense qui est L'IDS

Dans ce chapitre, nous avons présenté les différentes attaques informatiques, où nous avons

Présenté un aperçu général de ces attaques.

Dans le chapitre suivant, nous nous concentrerons sur le Système de Détection d'Intrusion, et ses types, son emplacement et son modèle général.

Chapitre02 : Système de détection d'intrusion :

1. Introduction

La détection d'intrusion est un élément crucial d'un système de protection et de sécurité. Parce que nouveau des cyberattaques apparaissent chaque jour, les systèmes de détection d'intrusion (IDS) jouent un rôle majeur à reconnaître d'éventuelles cyberattaques sur un réseau ou un système et à fournir réponses. Les IDS doivent s'acclimater à ces nouvelles menaces (cyberattaques) et à leurs stratégies, ainsi que continuer à évoluer. De nombreux outils sur le marché offrent désormais plusieurs niveaux de détection d'intrusion. Quelques utilisent des signatures pour surveiller les attaques connues. Certaines plateformes proposent surveillance de réseaux; d'autres sont des systèmes basé1s sur l'hôte. Certaines solutions répondent à des avertissements en fermant les services. Nous devons choisir avec soin les tactiques de détection d'intrusion pour garantir la sûreté et la sécurité des ressources de notre réseau contre les intrus indésirables.

Dans ce chapitre, nous allons présenter le Système de Détection d'Intrusion, son emplacement et son modèle général. Nous décrirons ensuite la classification de IDS.

2. Définition

2.1 Intrusion :

Les méthodes traditionnelles de protection contre les intrusions, telles que les pare-feu, la protection d'accès et le chiffrement ont eu du mal à sécuriser complètement les réseaux et les serveurs de plus en plus menaces avancées et logiciels malveillants. En conséquence, les systèmes de détection d'intrusion (IDS) sont devenus un composant essentiel de l'infrastructure de défense pour détecter ces attaques avant qu'elles ne le fassent

préjudice généralisé .

2.2 Un système de détection d'intrusion (IDS) :

Est un système logiciel ou matériel qui détecte accès illégal à un système ou à un réseau informatique et qui surveille le trafic réseau pour détecter toute activité suspecte et alerte lorsqu'une telle activité est découverte.

Bien que la détection et le signalement des anomalies soient les principales fonctions d'un IDS, certains systèmes de détection d'intrusion sont capables de prendre des mesures lorsqu'une activité

malveillante ou un trafic anormal est détecté, notamment en bloquant le trafic envoyé à partir d'adresses IP (Internet Protocol) suspectes.

Un IDS peut être mis en contraste avec un système de prévention des intrusions (IPS), qui surveille les paquets réseau pour le trafic réseau potentiellement dommageable, comme un IDS, mais a pour objectif principal de prévenir les menaces une fois détectées, par opposition à la détection et à l'enregistrement des menaces. [13]

Les systèmes de détection d'intrusion , composés de capteurs placés à différents emplacements dans un réseau, renforcent la sécurité de l'environnement. Leur objectif principal est de détecter les signatures d'attaques connues ainsi que les comportements anormaux des paquets ou des flux de données sur les réseaux informatiques. Cela permet aux entreprises de surveiller les activités sur leurs réseaux, en particulier les interfaces de communication avec Internet. Le placement stratégique des capteurs est essentiel pour assurer une efficacité optimale du système. Il existe différents modèles et formats de déploiement d'IDS/IPS qui doivent être bien compris pour garantir une protection adéquate dans l'environnement de l'entreprise. [14]

3.Shéma générale

L'IDWG (Intrusion Detection Working Group) de l'IETF a défini une modèle de détection d'intrusion qui représente la fonctionnalité commune à tous les IDS (qu'ils utiliser l'approche comportementale ou l'approche par scénarios).

La figure09 : proposée par l'IDWG (Intrusion Detection Working Group) montre processus générique des systèmes de détection d'intrusion

L'administrateur configure les différents composants (capteur(s), analyseur(s), gestionnaire (s)). Les capteurs accèdent aux données brutes, les filtrent et les formatent pour ne renvoyer que les données intéressants événements à un analyseur. Les analyseurs utilisent ces événements pour décider si une intrusion est ou non présent et, si nécessaire, envoyer une alerte au gestionnaire (qui prévient l'opérateur humain). UNE réaction éventuelle peut être effectuée automatiquement par le gestionnaire ou manuellement par le opérateur. [15]

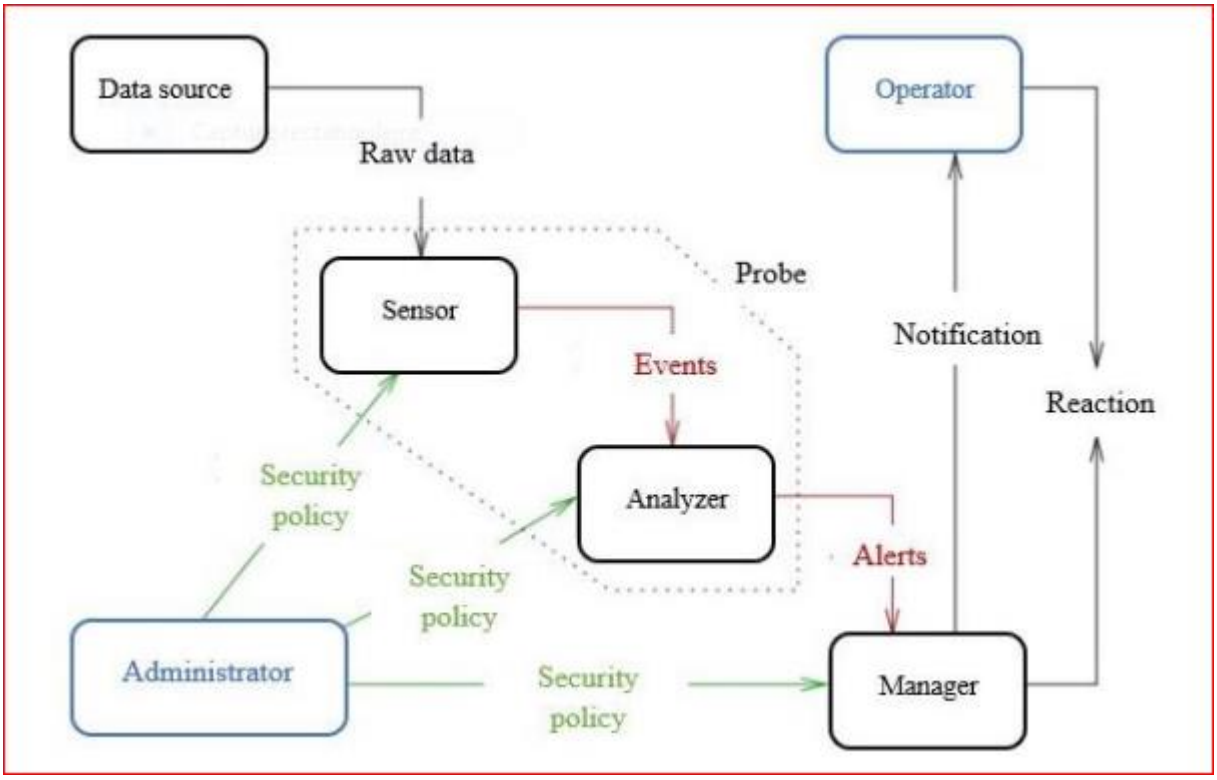

Figure 9 : Modèle générique de détection d'intrusion [15]

4.Classification IDS

Les IDS sont classées en 5 types :

5.1 HIDS

IDS basé sur l'hôte , HIDS est le premier type de détection d'intrusion qui a été développé. Le HIDS est surveiller et analyser les activités au niveau du système de l'hôte unique ou de l'informatique interne système tel que la configuration du système. [16]

5.2 NIDS

Le NIDS (Network Intrusion Detection System) est un système de détection d'intrusion réseau qui surveille le trafic en temps réel pour détecter les activités suspectes. Il analyse les paquets, les protocoles et les signatures afin d'identifier les comportements malveillants. Lorsqu'une activité suspecte est détectée, des alertes sont générées et des mesures préventives peuvent être prises. Les NIDS sont essentiels pour renforcer la sécurité des réseaux en détectant et en prévenant les intrusions et les attaques. [16]

5.3 PIDS

Système de détection d'intrusion basé sur protocole (PIDS) composé d'un système ou d'un agent qui résiderait systématiquement à l'avant d'un serveur, contrôlant et interprétant le protocole entre

un utilisateur/appareil et le serveur. Il essaie de sécuriser le serveur Web en surveillant régulièrement le flux de protocole HTTPS et en acceptant le protocole HTTP associé. Comme HTTPS n'est pas crypté et avant d'entrer instantanément dans sa couche de présentation Web, ce système devrait résider dans cette interface, entre pour utiliser le HTTPS. [16]

5.4 APIDS

Un système de détection d'intrusion basé sur le protocole d'application (APIDS) est un système ou un agent qui réside généralement dans un groupe de serveurs. Il identifie les intrusions en surveillant et en interprétant la communication sur des protocoles spécifiques à l'application. Par exemple, cela surveillerait explicitement le protocole SQL vers le middleware lors de ses transactions avec la base de données sur le serveur Web. [16]

5.5 Système de détection d'intrusion hybride

Le système de détection d'intrusion hybride est constitué par la combinaison de deux ou plusieurs approches du système de détection d'intrusion. Dans le système de détection d'intrusion hybride, l'agent hôte ou les données du système sont combinés avec des informations de réseau pour développer une vue complète du système de réseau. Le système de détection d'intrusion hybride est plus efficace par rapport à l'autre système de détection d'intrusion. Prélude est un exemple d'IDS hybride. [16]

5. Conclusion :

La sécurité du réseau est devenue un problème sérieux car plusieurs attaquants tentent d'attaquer réseaux pour atteindre un but, comme l'économie. De nombreuses méthodes de protection de réseau ont été proposés, tels que les systèmes de détection d'intrusion, la cryptographie, les pare feu, etc.

Ces outils de sécurité, la détection d'intrusion est généralement considérée comme l'une des plus encourageantes méthodes pour nous protéger contre les cyberattaques nouvelles, dynamiques et complexes.

Dans ce chapitre, nous avons d'abord présenté les Systèmes de Détection d'Intrusion et ses emplacements et leur modèle général. Nous avons ensuite discuté des 5 types de systèmes de détection d'intrusion :

HIDS ,NIDS, PIDS, APIDS, IDS hybride, Dans le chapitre suivant, nous présenterons l'apprentissage automatique, l'apprentissage profond et les réseaux de neurones.

Chapitre 3 : Machine Learning et Deep Learning

1. Introduction

L'intelligence artificielle (IA) fait référence à la production de machines intelligentes comme l'humain esprit. L'intelligence artificielle concerne la conception et la mise en œuvre de systèmes de machines.

Ces machines peuvent résoudre des problèmes qui nécessitent habituellement la capacité des humains.

Depuis l'article fondateur d'Alan Turing en 1950 sur la faisabilité de la programmation d'un machine électronique pour agir intelligemment, l'intelligence artificielle au cours des dernières décennies a connu une croissance rapide de la recherche et du développement.

Au XXIe siècle, l'intelligence artificielle (IA) est devenue un sujet essentiel de la recherche dans tous les domaines tels que la science, le droit, la médecine, l'éducation, la comptabilité, les affaires, marketing, finance, économie et ingénierie.

L'apprentissage automatique est l'un des domaines du monde informatique actuel. Beaucoup de recherches avait fait pour rendre les machines intelligentes. L'apprentissage est une activité humaine naturelle qui a également devenir un aspect crucial des ordinateurs. Les chercheurs ont déployé beaucoup d'efforts pour améliorer la précision des algorithmes d'apprentissage automatique.

L'apprentissage en profondeur (DL) devient de plus en plus crucial dans notre vie quotidienne. Il a déjà a eu une influence significative dans plusieurs domaines, tels que les voitures autonomes et la reconnaissance vocale.

Dans ce chapitre, nous présenterons l'apprentissage automatique avant de discuter de la façon dont il se transforme à l'apprentissage en profondeur avec ses types , Ensuite nous présentons les réseaux de neurones .

2. Machine Learning

2.1 Définition

L'apprentissage automatique (ML) est un sous-domaine de l'intelligence artificielle. Il s'agit de étudier les algorithmes informatiques qui permettent aux systèmes d'apprendre et de s'améliorer automatiquement à partir de chacun expérience.

Selon l'approche d'apprentissage, le type de données d'entrée et de sortie, et le type de problème résolu, les algorithmes d'apprentissage automatique ont été divisés en plusieurs types :

Apprentissage supervisé, non supervisé, semi-supervisé et apprentissage renforcé. [17]

2.2. Les types d'apprentissage automatique

2.2.1 Apprentissage supervisé

On utilise ce type d'apprentissage lorsque les données sont sous l'une des deux formes, variables d'entrée, et les valeurs cibles de sortie. L'algorithme commence à apprendre la fonction de mappage à partir de l'entrée variable aux valeurs cibles de sortie.

Il existe deux types de ce modèle : la classification (où la variable de sortie est discrète), régression (où la variable de sortie est continue). La figure 10 montre un exemple de

Enseignement supervisé. [17]

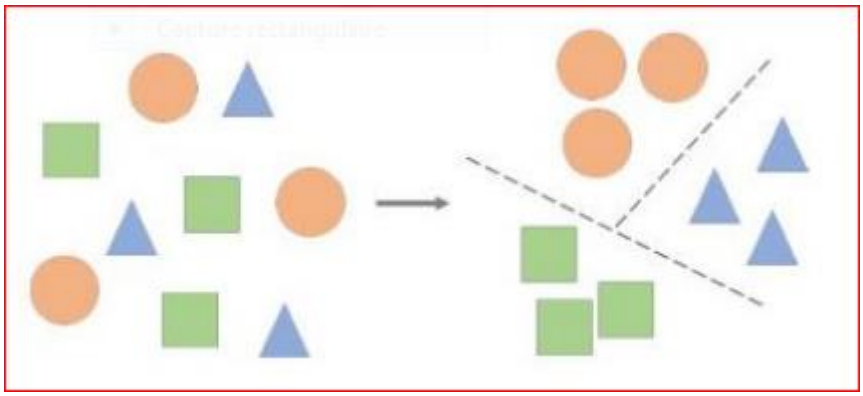

Figure 10 :Aperçu de l'apprentissage supervisé [17]

2.2.2 Apprentissage non supervisé

L'apprentissage non supervisé est utilisé lorsque les données ne sont utilisées que comme entrée et qu'il y a aucune variable de sortie correspondant à ces données. Pour en savoir plus sur les caractéristiques des données, telles que un algorithme modélise les modèles sous-jacents de ces données.

L'un des algorithmes non supervisés les plus courants est le clustering. Cette technique découvre des groupes inhérents dans les données, puis les utilise pour la prédiction de la sortie pour l'invisible contributions. [17]

La figure 11 montre un exemple d'apprentissage non supervisé

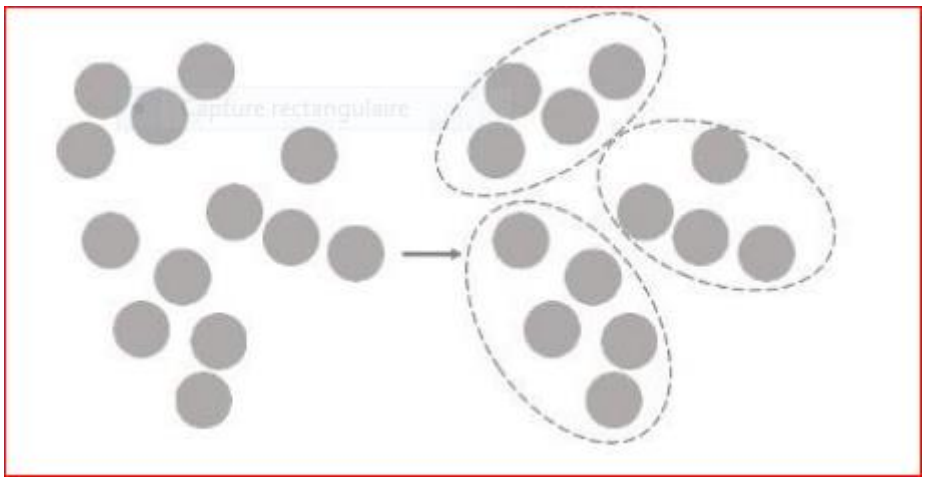

Figure 11 : Aperçu de l'apprentissage non supervisé. [17]

Les échantillons d'entrée sont regroupés en grappes sur la base des modèles sous-jacents.

2.2.3 Apprentissage semi-supervisé

Ce type d'algorithme est un intermédiaire entre l'apprentissage non supervisé et l'apprentissage supervisé techniques. Cet algorithme s'entraîne en utilisant une combinaison de non étiqueté (petite quantité) et données étiquetées (petite quantité). L'algorithme d'apprentissage non supervisé regroupe les premières données, puis il étiquette les données non étiquetées au repos en utilisant les données étiquetées existantes. [17]

Ont montré un exemple d'apprentissage semi-supervisé .

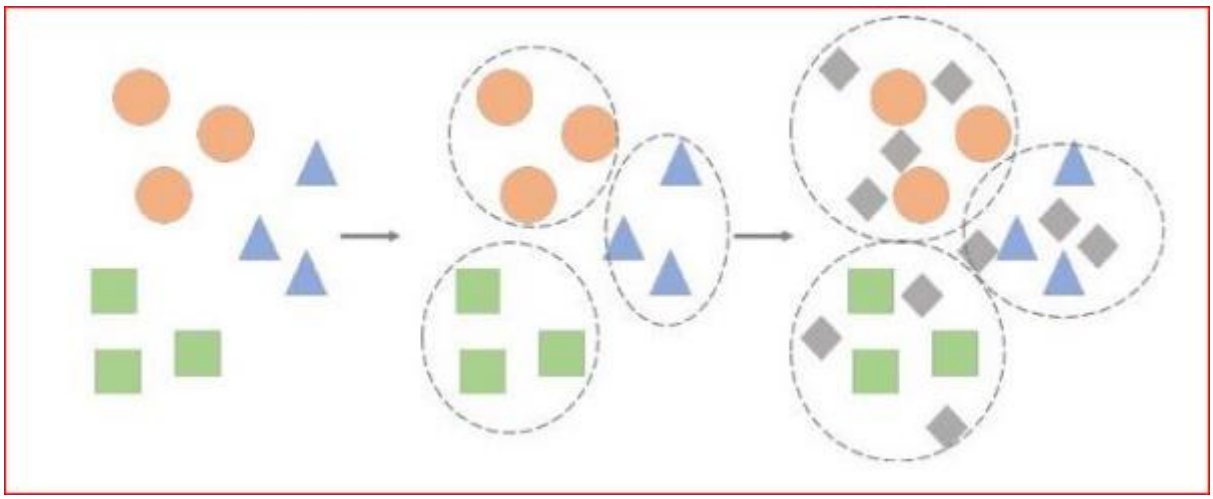

Figure 12: Aperçu de l'apprentissage semi-supervisé. [17]

Les grappes formées par une grande quantité de les données non étiquetées sont utilisées pour classer une quantité limitée de données étiquetées.

2.2.4 Apprentissage par renforcement

L'apprentissage par renforcement est utilisé lorsque l'objectif est de prendre une série de décisions qui mènent à une récompense finale. Un agent artificiel reçoit soit des pénalités, soit des récompenses selon sur les actions qu'il fait pendant le processus d'apprentissage. L'apprentissage par renforcement vise à maximiser la récompense totale. La figure 3.4 montre un exemple d'apprentissage par renforcement. [17]

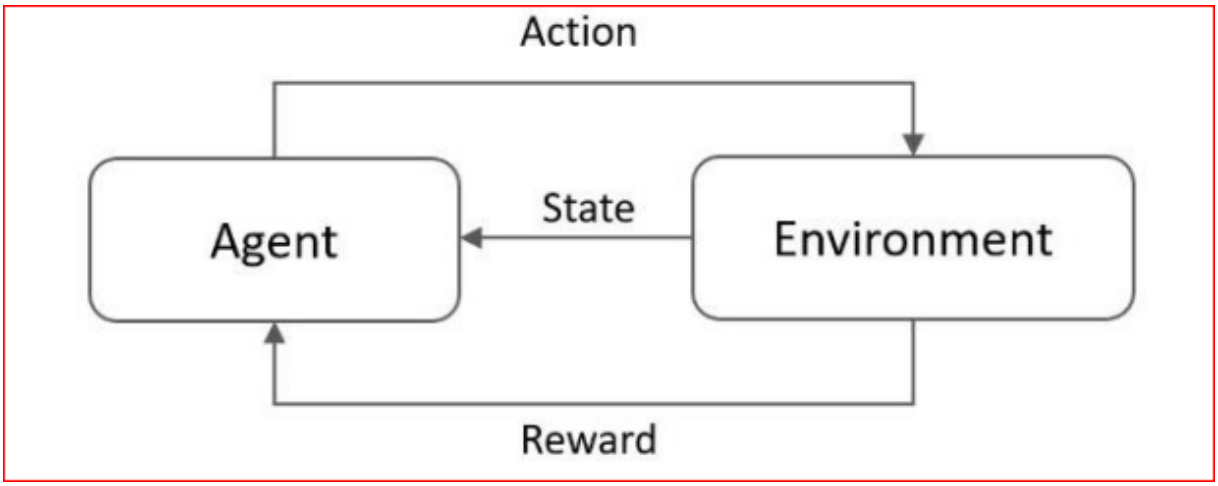

Figure 13 :Vue d'ensemble de l'apprentissage par renforcement. [17]

Un agent observe l'état de l'environnement et effectue des actions pour maximiser une récompense globale .

3.Deep Learning

3.1 Définition :

L'apprentissage en profondeur est un sous-domaine de l'apprentissage automatique qui utilise des réseaux de neurones artificiels qui contiennent deux couches cachées ou plus pour approximer une fonction f∗. ce f∗ peut être utilisé pour faire des prédictions ou mapper les données d'entrée à de nouvelles représentations [18]

3.2 Pourquoi et quand appliquer DL

DL est utilisé dans de nombreux cas où l'intelligence artificielle serait bénéfique.

-Absence d'un expert humain (navigation sur Mars)

-Les humains sont incapables d'exprimer leur expérience (compréhension du langage,

vision et reconnaissance vocale)

-La solution du problème évolue avec le temps (prévision de prix, stock, préférence,

prévision météorologique, suivi)

-Les solutions nécessitent d'être adaptées à certains cas (personnalisation,

biométrie).

-La taille du problème est trop énorme pour nos capacités logiques limitées (sentiment

analyse, mise en correspondance des publicités avec Facebook, classement des pages Web de calcul).[19]

4. De l 'apprentissage automatique à l'apprentissage profond

Pendant longtemps, l'apprentissage automatique a été l'outil principal. Cependant, avec le

Apparition des approches de machine learning Big Data se sont tournées vers la notion de deep l'apprentissage, où de nouvelles architectures ont été créées, plus puissantes et efficaces, pour faire face aux Notion de données volumineuses. Depuis 2006, le deep learning est devenu un outil très courant dans de nombreux recherche [20].

5.Les réseaux de nouerons

5.1Réseau neuronal profond (DNN) :

L'apprentissage en profondeur est un type d' apprentissage automatique et d'intelligence artificielle (IA) qui imite la façon dont les humains acquièrent certains types de connaissances. L'apprentissage en profondeur est un élément important de la science des données, qui comprend les statistiques et la modélisation prédictive . Il est extrêmement bénéfique pour les scientifiques des données qui sont chargés de collecter, d'analyser et d'interpréter de grandes quantités de données ; l'apprentissage en profondeur rend ce processus plus rapide et plus facile [21]

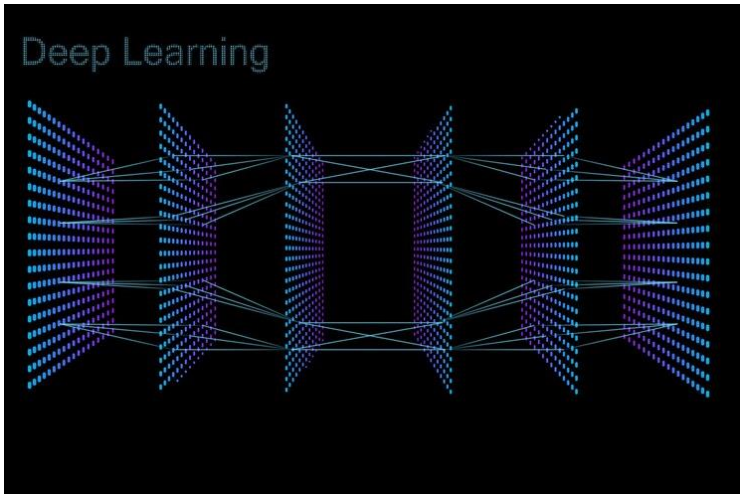

Figure 14 :Réseau neuronal profond (DNN) [22]

5.2 Réseau neuronal convolutif (CNN)

Le réseau de neurones convolutifs (CNN) est un type de réseau de neurones à réaction, il comprend des couches de convolution et des opérations de mise en commun. Il est capable d'extraire à la fois local et fonctionnalités globales.

Le principal avantage des réseaux de neurones convolutifs (CNN) est qu'ils reconnaît automatiquement les caractéristiques essentielles sans aucune intervention humaine. Pour

exemple, étant donné un grand nombre d'images de chiens et de chats, il peut apprendre les principales caractéristiques de chaque classe en soi [23]

5.3 Réseau neuronal récurrent (RNN)

Le réseau neuronal récurrent (RNN) est un modèle d'apprentissage séquentiel courant. RNN apprend les caractéristiques des données en série à l'aide d'une mémoire des entrées précédentes stockées dans le neurone l'état interne du réseau. Comme le montre la Figure 15, un cycle dirigé est utilisé pour établir les connexions entre les neurones [24].

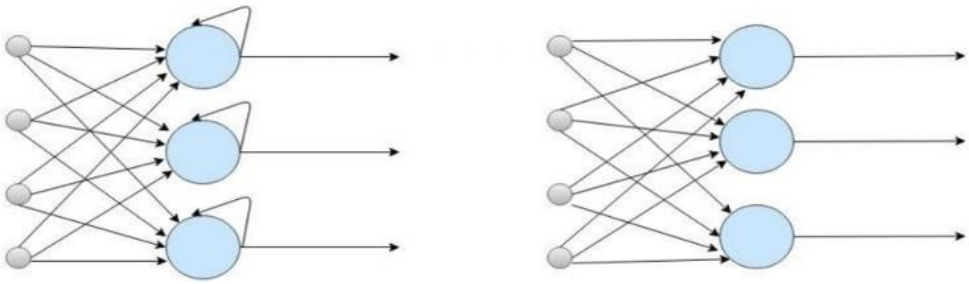

Figure 15 : Le réseau neuronal récurrent [25]

6. Conclusion

L'apprentissage profond est un nouveau sujet de recherche en apprentissage automatique, qui apporte de nouvelles architectures basées principalement sur l'ANN classique, mais avec plusieurs améliorations pour faire face aux défis du Big Data.

L'apprentissage automatique et l'apprentissage en profondeur ont été introduits dans ce troisième chapitre. Nous avons commencé avec la définition de l'apprentissage automatique comme un sous-domaine de l'intelligence artificielle avant discuter de ses différents types, y compris l'apprentissage supervisé, non supervisé, semi-supervisé et un apprentissage renforcé. Puis nous avons présenté le Deep Learning avant de terminer le chapitre sur la différente architecture de réseaux de neurones dans la dernière section.

Dans le chapitre suivant, nous présenterons les travaux de quelques algorithmes dans la détection d'intrusion en machine Learning .

Chapitre 4 :La détection d'intrusion basé sur l'apprentissage automatique

1. Introduction

La détection d'intrusion basée sur machine Learning utilise une variété de méthodes et d'algorithmes pour identifier les activités malveillantes et les tentatives d'intrusion dans les systèmes informatiques. Ces méthodes exploitent les capacités d'apprentissage automatique des algorithmes pour analyser les données et détecter les schémas anormaux ou suspects.

Dans ce chapitre, nous aborderons les méthodes de quelques algorithmes couramment utilisés dans la détection d'intrusion basée sur machine learning :

2. Les travaux dans la détection d'intrusion

Algorithme	Auteur	Jeux de donnée	Résultat	Taux d'exactitude 'accuracy'
k-means	S.Sbhatia A.K sharma [26]	Jeux de donnée k.DDcup 1999	Efficace	98.9
KNN	SvShinde RP Borole RN PHursule [27]	Jeux de donnée NSL.KDD		97.5
SVM	RS Thakur et K THakur [28]	Jeu de donnée NSL.KDD		99.6
R N	S Sharma .Sk Sharma [29]	NSL.KDD		98.9
Régression logistique	SG Hakeem MM Hasen [30]	NSL.KDD		98.5
Random forest	MS AL hashemi Al- Jumail [31]	NSL.KDD		99.1
Arbre de décision	SR Choudhary KK Jupta [32]	NSL.KDD		92.3

Tableau 1: les méthodes des algorithmes dans la détection intrusion

3. Conclusion

Dans ce chapitre, nous résumerons dans un tableau quelques algorithmes couramment utilisés dans la détection d'intrusion, ainsi que les chercheurs associés à leurs travaux et les précisions obtenues , dans le chapitre suivant présenterons la création d'un modèle de classification des attaques DDoS .et l'importance de la classification des attaques et ses applications, ainsi que quelques exemples de modèles utilisés , nous décrivons brièvement les étapes nécessaires à la mise en œuvre du modèle, Enfin, nous parlerons des résultats obtenus.

Chapitre 5 : Évaluation et Discussion

1. Introduction

La classification des attaques DDoS est cruciale pour la sécurité des systèmes informatiques. Différents modèles ont été développés pour cette classification, notamment le modèle de syntonisation (tuner model) qui joue un rôle important. Dans ce chapitre nous présentons la création d'un modèle de classification des attaques DDoS. Nous abordons l'importance de la classification des attaques et ses applications, ainsi que quelques exemples de modèles utilisés. Nous mettons l'accent sur le modèle de syntonisation. Ensuite, nous décrivons brièvement les étapes nécessaires à la mise en œuvre du modèle, en mentionnant que les détails seront expliqués ultérieurement. Enfin, nous discutons des résultats obtenus.

2. Proposition (optimisation du resnet via kerastuner)

2.1 ResNet

ResNet est une abréviation de "Residual Network" (réseau résiduel en français). Il s'agit d'une architecture de réseau de neurones profonds qui a été introduite par Kaiming He et al. dans leur article de recherche de 2015 intitulé "Deep Residual Learning for Image Recognition" (Apprentissage résiduel profond pour la reconnaissance d'images).

L'idée principale derrière ResNet est de résoudre le problème des gradients qui disparaissent lors de l'apprentissage de réseaux de neurones très profonds. Lorsque le nombre de couches augmente, il devient difficile pour le réseau d'apprendre et de propager efficacement les gradients, ce qui entraîne une performance dégradée. ResNet aborde ce problème en introduisant une connexion résiduelle ou une connexion sautée.

Dans un ResNet, l'entrée d'une couche est combinée avec la sortie d'une couche précédente grâce à une connexion sautée. Cela signifie que le réseau apprend à modéliser le résidu ou la différence entre la sortie souhaitée et la sortie actuelle. En utilisant cette approche d'apprentissage résiduel, ResNet permet l'entraînement de réseaux très profonds (par exemple, des centaines de couches) tout en maintenant une bonne précision.

Les architectures ResNet existent en différentes variantes, telles que ResNet-50, ResNet-101 et ResNet-152, qui diffèrent par le nombre de couches et de paramètres. Ces architectures ont atteint des performances de pointe dans diverses tâches de vision par ordinateur, notamment la classification d'images, la détection d'objets et la segmentation d'images.

ResNet a eu un impact significatif dans le domaine et a inspiré le développement d'autres architectures basées sur les résidus. Il est devenu un bloc de construction fondamental dans de nombreuses applications d'apprentissage profond, en particulier dans les tâches de vision par ordinateur qui impliquent le traitement et la compréhension de données visuelles. [33]

2.2 Hyper Resnet

Le terme "Hyper ResNet" n'est pas couramment utilisé et il n'y a pas de définition standard pour cette terminologie. Cependant, si nous interprétons "Hyper ResNet" comme une combinaison des concepts de ResNet et d'hyperparamètres, cela pourrait faire référence à une adaptation ou une personnalisation spécifique de l'architecture ResNet en utilisant des techniques d'optimisation des hyperparamètres..

L'optimisation des hyperparamètres concerne le processus de recherche des meilleurs hyperparamètres (comme la taille du lot, le taux d'apprentissage, le nombre de couches, etc.) pour un modèle d'apprentissage automatique donné. En utilisant des techniques telles que la recherche par grille, la recherche aléatoire ou l'optimisation bayésienne, on peut ajuster les hyperparamètres de l'architecture ResNet pour obtenir des performances optimales sur une tâche spécifique.

En résumé, ResNet est une architecture de réseau de neurones profonds avec des connexions résiduelles, tandis que "Hyper ResNet" peut faire référence à une adaptation spécifique de ResNet en utilisant des techniques d'optimisation des hyperparamètres pour une meilleure performance. [34]

2.3 Kerastuner

Le Keras Tuner est une bibliothèque qui nous aide à choisir l'ensemble optimal d'hyperparamètres pour le programme TensorFlow. Le processus de sélection du bon ensemble d'hyperparamètres pour votre application d'apprentissage automatique (ML) est appelé réglage d'hyperparamètres ou hyperréglage.

Keras Tuner simplifie le processus de recherche des meilleurs hyperparamètres en automatisant la recherche et l'évaluation de différentes combinaisons d'hyperparamètres. Il fournit une interface conviviale pour définir l'espace des hyperparamètres à explorer, choisir une stratégie de recherche (par exemple, la recherche aléatoire, la recherche par grille ou l'optimisation bayésienne), et effectuer la recherche elle-même.

Voici quelques fonctionnalités clés de Keras Tuner

- Définition de l'espace des hyperparamètres : Keras Tuner permet de spécifier l'espace des hyperparamètres à explorer, tels que le taux d'apprentissage, la taille du lot, le nombre de couches, les fonctions d'activation, etc. Vous pouvez définir des distributions continues ou discrètes pour chaque hyperparamètre.
- Stratégies de recherche : Keras Tuner propose plusieurs stratégies de recherche, y compris la recherche aléatoire, la recherche par grille et l'optimisation bayésienne. Chaque stratégie explore l'espace des hyperparamètres de manière différente, en évaluant plusieurs configurations de modèles.
- Évaluation des modèles : Keras Tuner fournit des mécanismes intégrés pour évaluer et comparer les performances des modèles avec différentes configurations d'hyperparamètres. Il gère automatiquement l'entraînement des modèles avec chaque configuration et enregistre les métriques de performance pour faciliter la comparaison et la sélection des meilleures configurations.

- Intégration avec Keras : Keras Tuner est spécialement conçu pour fonctionner de manière transparente avec Keras, ce qui facilite l'optimisation des hyperparamètres pour les modèles Keras. Il est également compatible avec TensorFlow 2.0+.
- En utilisant Keras Tuner, vous pouvez accélérer le processus d'optimisation des hyperparamètres, trouver plus efficacement les configurations optimales pour vos modèles, et améliorer les performances de votre apprentissage automatique. [35]

2.4 kaggle

Kaggle est une plateforme communautaire en ligne pour les scientifiques des données et les passionnés d'apprentissage automatique. Kaggle permet aux utilisateurs de collaborer avec d'autres utilisateurs, de rechercher et de publier des ensembles de données, d'utiliser des blocs-notes intégrés au GPU et de rivaliser avec d'autres scientifiques des données pour résoudre les défis de la science des données. L'objectif de cette plateforme en ligne (fondée en 2010 par Anthony Goldbloom et Jeremy Howard et acquise par Google en 2017) est d'aider les professionnels et les apprenants à atteindre leurs objectifs dans leur parcours en science des données grâce aux puissants outils et ressources qu'elle fournit. À ce jour (2021), il y a plus de 8 millions d'utilisateurs enregistrés sur Kaggle. [36]

et voici quelques-unes des principales fonctionnalités de Kaggle :

Jeux de données : Kaggle propose une vaste collection de jeux de données provenant de divers domaines, tels que l'image, le texte, le son, les séries chronologiques, etc. Les utilisateurs peuvent explorer et télécharger ces jeux de données pour les utiliser dans leurs projets.

Compétitions : Kaggle est bien connu pour ses compétitions de data science. Les compétitions permettent aux participants de mettre en pratique leurs compétences en résolvant des problèmes réels et de concourir pour des prix en soumettant leurs modèles prédictifs. Les compétitions peuvent varier en termes de complexité et de domaines d'application.

Notebooks : Les notebooks Kaggle sont des environnements basés sur le cloud qui permettent aux utilisateurs d'écrire, d'exécuter et de partager du code en Python ou R. Les notebooks intègrent des fonctionnalités telles que le code, les visualisations, les commentaires et la documentation dans un seul endroit, ce qui facilite la collaboration et le partage des résultats.

Kernels : Les kernels Kaggle sont des environnements d'exécution de code partagés qui permettent aux utilisateurs de créer et d'exécuter des scripts et des analyses. Les kernels prennent en charge plusieurs langages de programmation, notamment Python et R, et permettent aux utilisateurs de collaborer, de commenter et de partager leurs travaux.

Discussions : Kaggle offre une plateforme de discussion où les utilisateurs peuvent poser des questions, discuter de projets, partager des connaissances et interagir avec une communauté de data scientists du monde entier. Cela favorise l'apprentissage collaboratif et permet aux utilisateurs de bénéficier de l'expérience et de l'expertise des autres membres de la communauté.

Cours et didacticiels : Kaggle propose également des cours et des didacticiels en ligne gratuits pour apprendre les bases de l'apprentissage automatique, de l'analyse de données et d'autres sujets connexes. Ces ressources aident les débutants à se familiariser avec les concepts et les outils de data science.

Profils et portefeuilles : Les utilisateurs peuvent créer des profils sur Kaggle pour présenter leurs compétences, leurs projets et leurs réalisations dans le domaine de l'apprentissage automatique. Cela permet aux utilisateurs de se connecter avec d'autres professionnels du domaine, de recevoir des commentaires et d'établir leur réputation en tant que data scientistes

En résumé, Kaggle offre une gamme de fonctionnalités allant des jeux de données et des compétitions à l'hébergement de notebooks et de kernels, en passant par les discussions et les cours en ligne. Cette plateforme communautaire permet aux data scientistes d'apprendre, de collaborer, de partager et de concourir dans le domaine de l'apprentissage automatique. [37]

2.5 TensorFlow

TensorFlow Keras Tuner est une bibliothèque open-source qui fournit une interface de haut niveau pour l'optimisation des hyperparamètres dans les modèles d'apprentissage automatique construits avec TensorFlow et Keras. Elle offre une manière simplifiée et efficace de rechercher les hyperparamètres optimaux pour une architecture de modèle donnée, permettant aux développeurs d'améliorer les performances des modèles et de réduire les essais manuels et les erreurs.

- Algorithmes de recherche : TensorFlow Keras Tuner propose différents algorithmes de recherche, tels que la recherche aléatoire, Hyperband et l'optimisation bayésienne. Ces algorithmes explorent différentes combinaisons d'hyperparamètres dans un espace de recherche défini pour trouver la meilleure configuration.
- Espaces d'hyperparamètres : Il permet aux développeurs de définir l'espace de recherche des hyperparamètres en spécifiant les plages ou distributions de valeurs. Cette flexibilité permet d'explorer différentes configurations d'hyperparamètres lors du processus d'optimisation.
- Intégration avec TensorFlow Keras : TensorFlow Keras Tuner s'intègre parfaitement avec l'API TensorFlow Keras, facilitant la définition et l'optimisation des hyperparamètres pour les modèles Keras. Il prend en charge une large gamme d'architectures de modèles, notamment les modèles séquentiels, fonctionnels et sous-classe.
- Configuration facile : Avec TensorFlow Keras Tuner, les développeurs peuvent définir les hyperparamètres directement dans l'architecture du modèle, ce qui facilite la spécification des paramètres à optimiser et de leurs espaces de recherche correspondants.
- Entraînement et validation automatisés : Il automatise le processus d'entraînement et de validation des modèles avec différentes configurations d'hyperparamètres, permettant une expérimentation efficace. Il propose des rappels (callbacks) et des mécanismes d'arrêt prématuré pour contrôler le processus d'entraînement.
- Métriques de performance et visualisation : TensorFlow Keras Tuner permet aux développeurs de définir et de suivre plusieurs métriques d'évaluation lors de la recherche

des hyperparamètres. Il fournit des visualisations et des rapports pour analyser les résultats et comparer les différentes configurations d'hyperparamètres.

- Exécution distribuée et parallèle : Il prend en charge l'optimisation distribuée et l'exécution parallèle, permettant une exploration plus rapide de l'espace de recherche des hyperparamètres en exploitant plusieurs ressources informatiques.

TensorFlow Keras Tuner simplifie le processus d'optimisation des hyperparamètres en fournissant une interface intuitive et flexible, ce qui permet aux développeurs d'optimiser plus facilement les performances de leurs modèles. En automatisant la recherche des meilleurs hyperparamètres, il permet de gagner du temps et des efforts, ce qui facilite le développement et le déploiement rapides de modèles d'apprentissage automatique performants. [38]

3. Implémentation

3.1 Dataset CICIDS 2017

3.1.1 Description

Le jeu de données CICIDS2017 est un ensemble de données de détection d'intrusion réseau disponible publiquement. Il contient des données de trafic réseau capturées lors d'un scénario de cyberattaque simulée. L'objectif de ce jeu de données est d'évaluer les performances des systèmes de détection d'intrusion (IDS) et des algorithmes d'apprentissage automatique dans la détection et la classification de différents types d'attaques réseau

Voici une description succincte du jeu de données CICIDS2017 :

Origine du jeu de données : Le jeu de données a été généré en simulant diverses attaques réseau dans un environnement contrôlé à l'aide du laboratoire Cyber Range de l'Institut canadien de cybersécurité (CIC). Les attaques ont été conçues pour imiter des scénarios d'attaque réels, comprenant différents types d'anomalies de trafic réseau et d'activités malveillantes.

Collecte des données : Les données de trafic réseau ont été collectées à l'aide de CICFlowMeter, un outil spécialement développé pour la capture et l'analyse des flux de données réseau. Le jeu de données comprend à la fois du trafic bénin et du trafic associé à différents types d'attaques.

Caractéristiques : Le jeu de données fournit une large gamme de caractéristiques extraites des flux de trafic réseau. Ces caractéristiques comprennent la durée du flux, les adresses IP source et de destination, les ports source et de destination, les types de protocoles, les compteurs de paquets et d'octets, les indicateurs TCP, ainsi que diverses caractéristiques statistiques calculées à partir des données de flux.

Catégories d'attaques : Le jeu de données couvre plusieurs catégories d'attaques, notamment les attaques par déni de service (DoS), les attaques par déni de service distribué (DDoS), les attaques

par force brute, les botnets, les attaques web, les infiltrations, et bien d'autres. Chaque instance d'attaque est étiquetée avec la catégorie d'attaque correspondante.

Format des données : Le jeu de données est disponible au format CSV (valeurs séparées par des virgules), ce qui facilite son importation et son traitement à l'aide de différents outils d'analyse de données et d'apprentissage automatique

Le jeu de données CICIDS2017 a été largement utilisé dans la recherche et le développement de systèmes de détection d'intrusion, de sécurité réseau et d'algorithmes d'apprentissage automatique. Il offre une représentation réaliste du trafic réseau avec à la fois des activités bénignes et malveillantes, permettant aux chercheurs et aux professionnels d'évaluer l'efficacité de différentes approches de détection et de classification dans l'identification des attaques réseau. [39]

3.1.2 Prétraitement

Le but du prétraitement des données est de préparer les données pour le traitement afin qu'elles soient compatibles avec le modèle. Ensuite, étant donné que les données sont initialement réparties dans plusieurs fichiers, nous les avons fusionnées en un seul fichier pour faciliter le traitement. Pendant cette fusion, nous avons supprimé toutes les informations non pertinentes telles que les valeurs inconnues, les valeurs infinies et les cases vides.

De plus, la plupart des données de l'ensemble de données se trouvent dans des plages différentes. Par conséquent, il est nécessaire de normaliser ces données pour les ramener dans une plage unique, qui va de 0 à 1. Le processus de normalisation peut être réalisé en utilisant l'équation suivante :

df_norm = (df_num - df_num.min()) / (df_num.max() - df_num.min())

Le dataset a été fractionné en deux parties, une pour l'entraînement et une pour les tests, avec 70% réservés à l'entraînement et le reste pour les tests .

Le modèle sera entraîné en utilisant des données avec des dimensions de (9, 9, 1). Si vous souhaitez obtenir des explications plus détaillées sur l'étape de prétraitement, vous pouvez les trouver ici.

Nom de fichier	Taille de fichier
Friday-WorkingHours-Afternoon-DDos.pcap_ISCX.csv	77.12 MB
Friday-WorkingHours-Afternoon-PortScan.pcap_ISCX.csv	76.91 MB

Friday-WorkingHours-Morning.pcap_ISCX.csv	58.32 MB
Monday-WorkingHours.pcap_ISCX.csv	176.39 MB
Thursday-WorkingHours-Afternoon-Infilteration.pcap_ISCX.csv	83.1 MB
Thursday-WorkingHours-Afternoon-Infilteration.pcap_ISCX.csv	52.02 MB
Tuesday-WorkingHours.pcap_ISCX.csv	135.08 MB
Wednesday-workingHours.pcap_ISCX.csv	225.17 MB

Tableau 2 : nom et taille des fichiers

L'étape de prétraitement des données revêt une importance cruciale dans le domaine de l'apprentissage automatique. Elle vise à préparer les données en amont de l'entraînement du modèle, afin d'optimiser sa performance et sa capacité à généraliser sur de nouvelles données.

L'une des étapes clés du prétraitement est la normalisation des données. Lorsque les données présentent des échelles différentes, il devient difficile pour le modèle d'apprentissage de les interpréter correctement. La normalisation permet de ramener toutes les données à une même échelle, souvent entre 0 et 1. Cela facilite la comparaison des variables et permet au modèle de mieux appréhender les relations entre les caractéristiques.

La normalisation aide également à éviter les problèmes liés aux valeurs extrêmes. Les valeurs aberrantes peuvent avoir un impact disproportionné sur les résultats du modèle, faussant ainsi les prédictions. En normalisant les données, on atténue l'effet de ces valeurs extrêmes, ce qui permet au modèle de mieux s'adapter aux données globales.

De plus, la normalisation des données contribue à accélérer la convergence de l'algorithme d'apprentissage. En ramenant les données dans une plage spécifique, on évite les oscillations et les

instabilités lors de la mise à jour des poids du modèle. Cela permet au modèle de converger plus rapidement vers une solution optimale.

On a permet de réaliser un échantillonnage d'un ensemble de données en fonction des étiquettes de classe.

Etiquette	Nombre d'échantillon
BENIGN	50000
DoS Hulk	2000
PortScan	2000
DDoS	2000
DoS GoldenEye	2000
FTP-Patator	2000
SSH-Patator	2000
DoS slowloris	2000
DoS Slowhttptest	2000
Bot	1966
Infiltration	36
Heartbleed	11

Tableau 3: étiquette et nombre d'echantillon

3.2 Paramètre utilisé

Le modèle était très complexe, avec un nombre de paramètres de 42.682.639 Le temps de traitement était considérablement long, et voici les résultats de la configuration des paramètres Vous trouverez ci-dessous les paramètres utilisés dans les modéle, y compris les paramètres fixes et les paramètres pouvant être entraînés.

```
Total params: 42,682,639
Trainable params: 42,577,295
Non-trainable params: 105,344
```

Figure 16 : paramètre utilisé

3.3 Modèle obtenu

Le modèle proposé dans ce travail est basé sur l'hypermodel avec la bibliothèque de réglage automatique. Le modèle est construit à partir de couches de convolution et de regroupement.

Le modèle était trop grand, donc nous ne pouvions pas le représenter dans une figure. Nous l'avons donc résumé dans un tableau.

L'architecture du modèle est présentée dans le tableau 04 Le nombre total de paramètres d'apprentissage est de 42.682.639 , Ce modèle est considéré comme quelque peu complexe, mais il a donné de bons résultats, comme cela sera précisé ultérieurement.

Modèle : resnet

Type	Layer
Zero padding 2D	02
Conv 2D	112
BatchNormalization	113
Activation	109
MaxPooling 2D	01
Add	35

Tableau 4 : Architecture du modèle

Et les paramètres :

```
Total params: 42,682,639
Trainable params: 42,577,295
Non-trainable params: 105,344
```

La prédiction de la partie de test de l'ensemble de données, qui est considérée comme de nouvelles données pour tester l'efficacité du modèle. Le modèle a réussi à atteindre une précision ‘ accuracy’ de 0.96452 avec une grande précision ’accuracy’ de 0.96588 et taux de rappel 0.96588

4. Résultats et discussion

Les informations ont été divisées en deux parties distinctes : la première partie a été utilisée pour le traitement du modèle, tandis que la seconde partie a été réservée pour les tests. Lors des tests sur ces nouvelles données, le modèle a démontré une bonne précision, mesurée par un score d'accuracy de 0.96424, accompagné d'une perte (loss) très faible. Ces résultats témoignent de la performance du modèle et suggèrent qu'il est capable de bien prédire les valeurs cibles.

De plus, il est important de noter que les résultats de validation obtenus lors des tests sont similaires à ceux obtenus lors de l'entraînement du modèle. Cette cohérence entre les deux jeux de données indique que le modèle a la capacité de généraliser et d'effectuer des prédictions précises sur de nouvelles données. et la validation présente également une apparence similaire de traitement.

En résumé, après avoir traité le modèle, il a été testé sur de nouvelles données pour évaluer sa précision. Les résultats obtenus sont satisfaisants, ce qui confirme que le modèle est performant et capable de traiter efficacement de nouvelles données, se présentent comme suit :

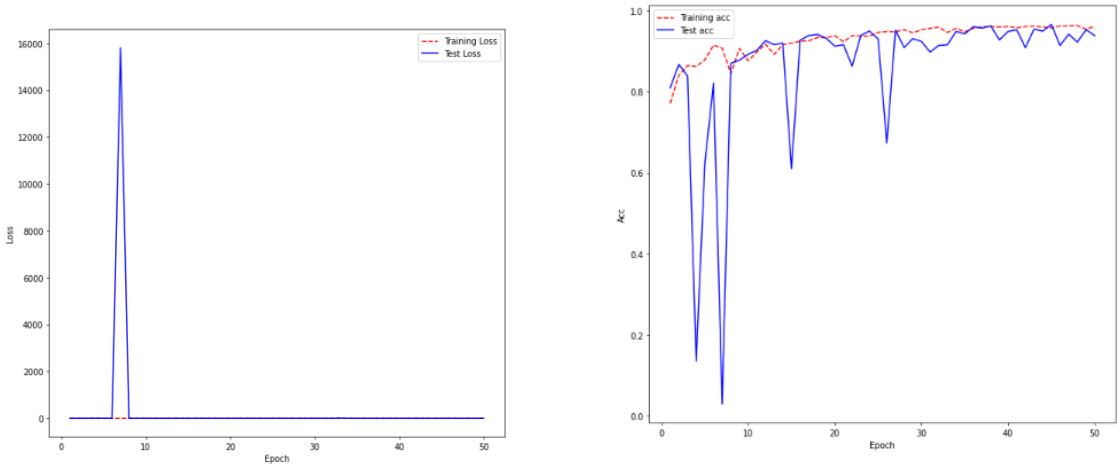

Figure 17 : Resultats

Par rapport au Loss : Le résultat obtenu est bon , cependant, il n'est pas clair .

```
638/638 [==============================] - 7s 11ms/step
Accuracy: 0.96424
Recall: 0.96452
Precision: 0.96723
F1 Score: 0.96588
```

Figure 18 : Evaluation

5. Conclusion

En conclusion, la création d'un modèle de classification des attaques DDoS optimisé par l'utilisation de tuner modèles est essentielle pour renforcer la sécurité des systèmes et des réseaux. Ces modèles permettent une détection précise des attaques DDoS en optimisant les hyperparamètres du modèle. Des exemples de modèles couramment utilisés tels que les réseaux de neurones convolutifs et les réseaux de neurones récurrents ont été mentionnés. Les étapes d'implémentation, comprenant la collecte
des données, la construction et l'entraînement du modèle, ont été brièvement expliquées. À travers cette approche, des résultats concrets et des améliorations significatives dans la précision de la classification des attaques DDoS peuvent être obtenus. En continuant d'explorer et d'améliorer ces modèles, nous pourrons mieux protéger les systèmes contre les attaques DDoS, assurant ainsi la disponibilité et la sécurité des services en ligne.

Conclusion et Perspectives

La détection des attaques DDoS est un défi majeur en matière de sécurité informatique. Les réseaux de neurones profonds sont des modèles puissants capables de reconnaître les comportements malveillants associés à ces attaques en extrayant des caractéristiques complexes des données.

Ce mémoire se concentre sur l'optimisation d'un réseau de neurones profond pour la détection des attaques DDoS en utilisant Keras Tuner. Cet outil permet d'explorer automatiquement et de manière itérative les hyperparamètres du modèle afin de trouver la meilleure configuration.

En utilisant un large ensemble de données d'attaques DDoS, le réseau de neurones est entraîné à reconnaître les schémas caractéristiques de ces attaques. Ensuite, grâce à Keras Tuner, les hyperparamètres tels que le nombre de couches, les fonctions d'activation et les taux d'apprentissage sont ajustés pour maximiser la précision de la détection.

Les résultats obtenus démontrent l'efficacité du modèle optimisé dans la détection des attaques DDoS. En exploitant les capacités d'apprentissage profond du réseau de neurones et en optimisant les hyperparamètres, le modèle atteint une précision élevée et une meilleure capacité à distinguer les attaques DDoS du trafic normal.

Et à la fin , l'utilisation d'un réseau de neurones profond optimisé via Keras Tuner offre une approche prometteuse pour la détection des attaques DDoS. Ce modèle optimisé permet une détection plus précise et fiable des attaques, renforçant ainsi la sécurité des systèmes informatiques face à cette menace croissante.

Références

A. Références Bibliographiques

[2] Michael E. Whitman, Herbert J. Mattord, "Principles of Information Security", Cengage Learning ,2018.

[4] Z.BENDELLA, «Gestion de la sécurité d'une application Web à l'aide d'un IDS comportemental optimisé par l'algorithme des K-means », Thèse de Master, Université de Tlemcen, 2013

[7] G.CHARPENTIER, O.MONTIGNY, M.ROUSSEAU, « Virus / antivirus », janvier 2004.

[8] J. LEGRAND, « Virus et Antivirus », STS Informatique de Gestion

[13] Richard Bejtlich ,"intrusion Detection Systems", No Starch Press,2013

[14] **IDS: history, concept and terminologys**

[16] S. M. Othman, F. Mutaher Ba-Alwi, N. T. Alsohybe, and A. T. Zahary, "Survey on Intrusion Detection System Types," Int. J. Cyber-Security Digit. Forensics, vol. 7, no. 4, pp. 444–462, 2018, [Online]. Available: https://www.researchgate.net/publication/329363322.

[18] J. M. Johnson and T. M. Khoshgoftaar, "Survey on deep learning with class imbalance," J. Big Data, vol. 6, no. 1, 2019, doi: 10.1186/s40537-019- 0192-5.

[19] M. Z. Alom et al., "A state-of-the-art survey on deep learning theory and architectures," Electron., vol. 8, no. 3, pp. 1–66, 2019, doi: 10.3390/electronics8030292.

[20] **S. Sah, "Machine Learning: A Review of Learning Types," 2020.**

https://www.researchgate.net/publication/342890321_Machine_Learning_

[21] S. Zhang, L. Yao, A. Sun, and Y. Tay, "Deep Learning Based Recommender Systems," ACM Comput. Surv., vol. 52, no. 1, pp. 1–38, 2019, doi: 10.1145/3285029 1.

[23] L. Yang and A. Shami, "On Hyperparameter Optimization of Machine Learning Algorithms: Theory and Practice," 2020.

https://www.researchgate.net/publication/343390531_On_Hyperparameter_

Optimization_of_Machine_Learning_Algorithms_Theory_and_Practice.

[24] Q. Zhang, L. T. Yang, Z. Chen, and P. Li, "A survey on deep learning forbig data," Inf. Fusion, vol. 42, pp. 146–157, 2018, doi: 10.1016/j.inffus.2017.10.006.

[26] MacQueen, J. (1967). "Some Methods for Classification and Analysis of Multivariate Observations." Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability.

[27] Cover, T., Hart, P. (1967). "Nearest neighbor pattern classification." IEEE Transactions on Information Theory

[28] Cortes, C., Vapnik, V. (1995). "Support-vector networks." Machine Learning

[29] Goodfellow, I., Bengio, Y., Courville, A. (2016). "Deep Learning." MIT Press.

[30] Cox, D. (1958). "The Regression Analysis of Binary Sequences." Journal of the Royal Statistical Society. Series B (Methodological).

[31] reiman, L. (2001). "Random Forests." Machine Learning.

[32] Quinlan, J.R. (1986). "Induction of Decision Trees." Machine Learning

[33] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, "Deep Residual Learning for Image Recognition",2015

[34] "deep Residual Learning for Image Recognition "

https://www.mygreatlearning.com/blog/resnet/

B. Références Web (Techniques)

[1] https://www.fortinet.com/resources/cyberglossary/types-of-cyber-attacks#:~:text=A%20cyber%20attack%20refers%20to,exploit%20or%20harm%20a%20network

[3] file:///C:/Users/tsi/Downloads/proxy_servers.pdf

[5] https://www.weodeo.com/la-securite-informatique/quest-ce-quun-pare-feu/

[6] https://blogs.oracle.com/oracle-france/post/quest-ce-quun-pare-feu

[9] https://www.cloudflare.com/fr-fr/learning/network-layer/what-is-ipsec/

[10] https://www.privacyaffairs.com/fr/vpn-ipsec/

[11] https://www.cs.unibo.it/babaoglu/courses/security/resources/documents/intro-to-crypto.pdf

[12] https://web.maths.unsw.edu.au/~lafaye/CCM/crypto/crypto.htm

[15] P. Illy, “Les systèmes de détection d ’ intrusion (IDS),” 2018.

https://www.researchgate.net/publication/335639245_Les_systemes_de_detection_d’intrusion_IDS (accessed May 27, 2021).

[17] S. Sah, “Machine Learning: A Review of Learning Types,” 2020.https://www.researchgate.net/publication/342890321_Machine_Learning_A_Review_of_Learning_Types (accessed May 30, 2021).

[22] https://fr.freepik.com/vecteurs-premium/modele-reseau-neurones-apprentissage-profond-modele-six-couches-points-neon-fond-noir_22595731.htm

[25] https://datavalue-consulting.com/deep-learning-reseaux-neurones-recurrents-rnn/

[35] https://keras-team.github.io/keras-tuner/

"Keras Tuner - An Easy-to-Use Hyperparameter Optimization Library" - Aashay Sanghvi, Eibe Frank (2020)

[36] https://www.datacamp.com/blog/what-is-kaggle

[37] https://www.kaggle.com/

[38] https://keras-team.github.io/keras-tuner/

[39] https://www.unb.ca/cic/datasets/ids-2017.html