REPUBLIQUE DU SENEGAL

Un Peuple - Un But - une Foi

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR
DE LA RECHERCHE, ET DE L'INNOVATION

UNIVERSITE GASTON BERGER DE SAINT-LOUIS

UNIVERSITE GASTON BERGER

L'excellence au service du développement

UNITE DE FORMATION ET DE RECHERCHE DES SCIENCES APPLIQUEES ET DE TECHNOLOGIE

Département de Physique Appliquée

Mémoire de fin d'étude pour l'obtention du diplôme de Master en Ingénierie Electronique et Télécommunications

ALLOCATION INTELLIGENTE DE RESSOURCES DANS LES RESEAUX 5G

Présenté par :

Souleymane Bore

Jury :

Pr. Papa Alioune FALL (Président)

Dr Ibrahima GUEYE (Examinateur)

Dr Abdoul Karim MBODJI (Examinateur)

Pr Cherif Diallo (Encadrant)

Année Académique 2023-2024

DEDICACES :

Louange à Dieu, le Tout-Miséricordieux. Sans Lui, rien de tout cela n'aurait été possible.

Je dédie ce travail à celles et ceux qui ont été des piliers dans ma vie, par leur présence, leur amour, leurs sacrifices ou leur souvenir inoubliable. À mon **grand-père Konaté Ndiaye**, qui m'a élevé avec rigueur, amour et dignité depuis mon enfance. Que Dieu lui accorde Sa miséricorde infinie. Son exemple et ses valeurs continuent de guider mes pas. À **ma mère**, une femme exceptionnelle, courageuse, patiente et toujours présente. Merci pour ta force, ton amour et ton soutien inébranlable. À **mon père**, paix à son âme, dont le souvenir me pousse à persévérer chaque jour.

À mon grand frère **Amady Boré**, pour ton soutien constant, tes conseils avisés, et surtout les sacrifices que tu as faits pour me permettre d'avancer. Tu as toujours cru en moi, même dans le silence. À mes grandes sœurs, **Adja et Wédji**, merci pour votre bienveillance, vos encouragements et votre présence à chaque étape. Je pense également à toute la famille Bore, à qui je dédie ce travail avec beaucoup d'émotion : **Hamidou, Idy, Bouba, Mady, Ibrahima** et tous les autres. Merci pour votre soutien, vos conseils et votre amour durant toutes ces années. À mon oncle **Malick Ndiaye,** dont le rôle dans ma vie a été fondamental. Que Dieu te récompense ainsi que toute **la famille Ndiayénne** pour tout ce que vous avez fait pour moi.

À mes promotionnaires et cochambreurs, **Khadim et Bachir**, avec qui j'ai partagé non seulement un espace mais aussi une belle aventure humaine. Merci pour la solidarité, les rires et les moments de complicité. Je penses également à toute la famille 4G3, ainsi qu'à l'ensemble des résidents du G3E.

Je rends aussi un hommage tout particulier à **Ibrahima Kontéye**, un promotionnaire, un frère, un brillant esprit passionné de mathématiques, arraché trop tôt à la vie. Ton souvenir demeure vivant dans mon cœur. Ce travail est aussi le tien.

Enfin, j'adresse une pensée pleine d'émotion à ma sœur **Mame Diarra Thiam**, Que son âme repose en paix. De l'école primaire j'jusqu'au lycée, nous avons partagé tant de souvenir inoubliables. Tu restes a jamais gravée dans mon cœur.

Je n'oublie pas non plus mes enseignants, mes camarades de classe, mes amis, et toutes celles et ceux qui, de près ou de loin, ont marqué ce parcours. Merci pour votre impact dans ma vie.

REMERCIMENTS :

Je rends grâce à Allah, le Tout-Puissant, le Clément et le Miséricordieux, pour Ses bienfaits innombrables et pour la lumière qu'Il m'a accordée tout au long de ce parcours. Sa présence m'a guidé dans les moments de doute et m'a donné la force d'avancer avec foi et persévérance. Ma gratitude la plus profonde va à ma mère, dont le courage, l'amour sans limites et le soutien indéfectible ont été essentiels à chaque étape de ma vie

Je remercie chaleureusement mon encadrant, **Pr. Chérif Diallo**, Professeur titulaire à l'Université Gaston Berger, pour son accompagnement attentif, sa rigueur scientifique et ses précieux conseils qui ont largement contribué à la réussite de ce mémoire. Ma reconnaissance s'étend à tous **les professeurs de l'UFR SAT**, dont l'engagement et la transmission du savoir m'ont permis de grandir tant sur le plan intellectuel que personnel.

Je tiens à remercier ma famille dans son ensemble mes proches, parents, tantes, frères et sœurs pour leur affection, leurs prières et leurs encouragements constants. Ce mémoire vous est aussi dédié. Mes remerciements vont également à mes compagnons de travail, **Moussa Ndiaye, Fatou Diop et Ibrahima Diaby…**, pour leur collaboration et l'esprit d'équipe qui nous a animés tout au long de notre parcours.

À mes camarades de la **promotion 28 (MPI et MASS)** et aux étudiants de la classe **MIETEL**, merci pour votre solidarité, vos partages et tous les moments enrichissants vécus ensemble.

Enfin, je tiens aussi à exprimer ma profonde gratitude au Cercle des Étudiants de Pout à Saint-Louis (**CEPS**). Grâce à votre accueil, j'ai pu m'intégrer, trouver un logement et poursuivre sereinement mes études. C'est grâce à vous j'ai su que la famille n'est pas forcement biologique. À tous mes aînés, ainsi qu'aux bleus et bleuses, merci pour votre solidarité et votre soutien indéfectible.

RÉSUMÉ :

Ce mémoire explore l'utilisation de l'apprentissage automatique pour améliorer l'allocation des ressources dans les réseaux 5G. Après un rappel de l'évolution des réseaux mobiles de la 1G à la 5G, il met en évidence les limites de la 4G face aux exigences actuelles : faible latence, forte densité de connexions et réduction de la consommation énergétique. La 5G, conçue pour répondre à ces besoins, s'appuie sur des technologies comme le SDN et la NFV pour rendre les réseaux plus flexibles et intelligents.

L'étude montre que le machine learning permet d'optimiser la gestion dynamique des ressources, en anticipant la charge, adaptant l'allocation et répondant aux besoins des différents services 5G (eMBB, URLLC, mMTC). Trois modèles sont testés : régression linéaire, régression polynomiale et forêt aléatoire. Leur entraînement et évaluation sont réalisés sur un jeu de données réel. L'analyse des résultats confirme l'intérêt de ces approches pour améliorer la performance réseau.

Ce travail met en avant l'intégration prometteuse du machine learning dans la gestion automatisée des réseaux 5G, et ouvre la voie à l'utilisation de techniques plus avancées comme le deep learning dans les futurs développements.

Mots clés : Réseau 5G, apprentissage automatique, gestion des ressources, IoT, SDN, NFV, l'intelligence artificielle, network slicing, modèles, regression linéaire, régression polynomiale et le forêt aléatoire, deep learning.

ABSTRACT:

This thesis explores the use of machine learning to enhance resource allocation in 5G networks. After reviewing the evolution of mobile networks from 1G to 5G, it highlights the limitations of 4G in meeting modern demands such as low latency, high connection density, and reduced energy consumption. 5G, designed to address these challenges, leverages technologies like SDN and NFV to create more flexible and intelligent networks.

The study demonstrates that machine learning can optimize dynamic resource management by predicting network load, adjusting allocation, and meeting the requirements of various 5G services (eMBB, URLLC, mMTC). Three models are tested: linear regression, polynomial regression, and random forest. These models are trained and evaluated on a real-world dataset. The analysis of results confirms the effectiveness of these approaches in improving network performance.

This work emphasizes the promising integration of machine learning in automated 5G network management and paves the way for the application of more advanced techniques such as deep learning in future developments.

Keywords: 5G network, machine learning, resource management, IoT, SDN, NFV, artificial intelligence, network slicing, models, linear regression, polynomial regression, random forest, deep learning.

TABLES DES MATIÈRES :

DEDICACES : i
REMERCIMENTS : ii
RÉSUMÉ : iii
ABSTRACT: iv
TABLES DES MATIÈRES : v
LISTE DES FIGURES ET DES TABLEAUX : vii
LISTE DES ABREVIATIONS: ix
INTRODUCTION GÉNÉRALE : 2
1. LES RESEAUX MOBILES 3
1.1 Introduction : 3
1.2 Évolution des réseaux mobiles 3
1.2.1 Les technologies de première génération (1G) et de deuxième génération (2G) 4
1.2.2 Les technologies de la troisième génération (3G) 5
1.2.3 La quatrièmes génération (4G) 6
1.3 Limites des réseaux (4G) 8
1.4 La cinquième génération (5G) 9
1.4.1 Exigence de la 5G 9
1.4.2 Architecture de la 5G 11
1.4.3 Les caractéristiques de la 5G 15
1.5 Conclusion 18
2.PROGRAMATION DES RESEAUX (SDN/NFV) 19
2.1 Introduction 19
2.2. Systèmes distribués traditionnels 19
2.3 Réseaux SDN 21
2.3.1 Historique et définition des réseaux SDN 21
2.3.2 Architecture détaillée des réseaux SDN 22
2.4 La virtualisation des fonctions réseau (NFV) 29
2.4.1 Définition de NFV 29

2.4.2 Architecture NFV 30
2.4.3 Réduction du CAPEX /OPEX 33
2.4.4 Chaine de fonctions services réseaux (SFC) 34
2.5 Les outils d'implémentation de SDN et NFV 36
2.6 Synergie entre SDN, NFV et le Machine Learning dans les réseaux 5G 37
2.7 Conclusion 38
3. APPLICATION DU MACHINE LEARNING POUR 39
L'ALLOCATION DE RESSOURCES DANS LES RESEAUX 5G 39
3.1 Introduction 39
3.2 Problématique de l'allocation des ressources dans les réseaux 5G 39
3.2.1 Définition de l'allocation des ressources 39
3.2.2 Méthodes traditionnelles et leurs limites 40
3.2.3 Exemple de paramètres influençant allocation des ressources 40
3.3 Généralités sur le machine Learning 41
3.3.1 Définition 41
3.3.2 Types d'apprentissage automatique 42
3.3.3 Étapes d'un projet machine learning 43
3.4 Étude pratique : Modèle de Machine Learning pour l'allocation 44
3.4.1 Equations 44
3.4.2 Description du jeu de données 46
3.4.3 Prétraitement et visualisations des données 47
3.5 Construction et entrainement des modèles 53
3.5.1 Sélection des modèles 54
3.5.2 Entrainement des modèles 54
3.5.3 Résultats et analyse 57
3.6 Conclusion 60
CONCLUSION GÉNÉRALE : 61
Références Bibliographiques : 62

LISTE DES FIGURES ET DES TABLEAUX :

Figure 1.1 Évolution des réseaux mobiles. ……4
Figure 1.2 Architecture 2G …….5
Figure 1.3 Architecture 3G …….6
Figure 1.4 Architecture 4G …….7
Figure 1.5 Comparaison 4G/5G …….....9
Figure 1.6 Les exigences de capacité de la 5G ……10
Figure 1.7 Architecture 5G en NSA et SA …….....11
Figure 1.8 Architecture 5G ……...12
Figure 1.9 Bandes de fréquences 5G…….....15
Figure 1.10 MIMO Massif …….....17
Figure 1.11 Technique de MIMO Massif……..17
Figure 1.12 Full duplex …….....18
Figure 2.1 Encapsulations desencapsulations du modèle OSI …….20
Figure 2.2 Comparaison entre réseau traditionnel et Réseau SDN ……..……...22
Figure 2.3 Architecture détaillée d'un réseau SDN……..22
Figure2.4 Architecture d'un Commutateur Openflow…….....25
Figure 2.5 Format de table de flux open flow …….……27
Figure 2.6 Organigramme de traitement d'un paquet dans un …….27
Commutateur openflow…….…….....27
Figure 2.7 Architecture de NFV…….....30
Figure 2.8 Les interfaces de l'architecture NFV…….……..33
Figure 2.9 Chaine de fonctions service…….....34
Figure 2.10: Architecture de chaînage des fonctions services……..35
Figure 3.1 IA et Machine learning……41
Figure 3.2 Type d'apprentissage automatique …….42
Figure 3.3 Donnée extrait du dataset……47
Figure 3.4 Traitement et conversion de formats…….....48
Figure 3.5 Nombre d'occurrences par type d'application…...…….49
Figure 3.6 Répartition des types d'application…….49

Figure 3.7 visualisation croisées entre variable……51
Figure 3.8 Matrice de corrélation……..52
Figure 3.9 Programme du modèle régression linéaire ……55
Figure 3.10 Programme du modèle Régression polynomiale……...55
Figure 3.11 Programme du modèle forêt aléatoire…….....56
Figure 3.12 Prédiction à l'aide de régression linéaire……...57
Figure 3.13 Prédiction à l'aide de régression polynomiale…….....58
Figure 3.14 Prédiction à l'aide de forêt aléatoire……..58
Tableau 1.1 Comparaison de la 1G à la 4G……..…….....7
Tableau 1.2 Les fonctions des modules de 5G NGC…….....13
Tableau 1.3 Les fonctions des modules communs ……..…….14
de EPC et 5G NGC …….....14
Tableau 2.1 Quelques types de contrôleurs et caractéristiques …….....24
Tableau 2.2 Table comparative des contrôleurs SDN ……..37
Tableau 3.1 Méthodes traditionnelles d'allocation ……..40
de ressources réseau …….40
Tableau 3.2 Quelques paramètres de ressources……...41
Tableau 3.3 Description des caractéristiques des données ……...46
Tableau 3.4 Comparaison des performances des modèles…….....60

LISTE DES ABREVIATIONS:

1G: first generation, 2G: second generation ,3G: third generation, 4G fourth generation, 5G: five generation

IoT: Internet of Things

SDN: Software Defined Networking

NFV: Network Function Virtualization

QoS : qualité de service

GSM : Generation Système mobile

WAP : Wireless Application Protocol

LTE: Long Term Evolution (technologie 4G)

3GPP :3rd Generation Partnership Project

4G: Fourth Génération (réseau mobile)

5G: Fifth Generation (réseau mobile)

AI: Artificial Intelligence (intelligence artificielle)

AF: Application Function

AMF: Authentication Management Function

ANIPS: Advanced Mobile Phone Service

AMQP: Advanced Message Queuing Protocol

API: Application programming Interface

AUSF: Authentication Server Function

BLE: Bluetooth Low Energy

CAPEX: Capital Expenditure

CDMA: Code Division Multiple Access CoAP: Constrained Application Protocol

DC-HSPA+: Dual-Cell High Speed Packet Access+

DNS: Domain name services

DPI: Deep Packet Inspection

EDGE: Enhanced Data Rates for GSM Evolution

eMBB: Enhanced Mobile Broadband

ETSI: European Telecommunications Standards Institute

GB: Gigabyte

Gbps: Gigabit per second

GNB : gNodeB (station de base en 5G)

IoT: Internet of Things

KPI: Key Performance Indicator

MAE: Mean Absolute Error

MEC: Multi-access Edge Computing

ML Machine Learning

MIMO: Multiple Input Multiple Output

mMTC: massive Machine-Type Communication

MSE: Mean Squared Error

NF: Network Function

NFV: Network Function Virtualization

O-RAN: Open Radio Access Network

QoE: Quality of Experience

SINR: Signal-to-Interference-plus-Noise Ratio

INTRODUCTION GÉNÉRALE :

Les progrès rapides des télécommunications ont radicalement transformé notre manière de communiquer, de travailler et d'interagir avec le monde. Depuis les premières générations de réseaux mobiles (1G, 2G) jusqu'à la 4G, les systèmes ont évolué pour satisfaire des exigences croissantes en termes de débit, de mobilité et de qualité de service (QoS). Toutefois, face à l'émergence de nouveaux usages comme l'Internet des objets (IoT), les véhicules autonomes et la réalité virtuelle, la 4G montre ses limites, notamment en matière de latence, de densité de connexions et d'efficacité énergétique. La 5G a été conçue pour relever ces défis. Elle promet des performances accrues : débits très élevés, latence réduite, forte densité de connexions et meilleure efficacité énergétique. Cependant, la complexité croissante des réseaux 5G rend la gestion des ressources plus délicate. Les besoins variables et dynamiques des utilisateurs exigent des mécanismes intelligents d'allocation des ressources. L'arrivée des technologies SDN (Software Defined Networking) et NFV (Network Function Virtualization), qui séparent le plan de contrôle du plan de données et permettent une plus grande programmabilité et flexibilité des réseaux. Ces innovations préparent le terrain à l'intégration de l'intelligence artificielle, et plus particulièrement du machine learning (apprentissage automatique), dans la gestion des réseaux. Le machine learning permet en effet de prendre des décisions adaptatives en temps réel, d'anticiper la charge réseau et d'optimiser dynamiquement l'allocation des ressources. En exploitant les données collectées, ces techniques offrent des capacités de prédiction et d'automatisation très prometteuses pour les réseaux mobiles de nouvelle génération.

Ce document est structuré en trois chapitres : Le premier chapitre présente l'évolution des réseaux mobiles, des premières générations à la 5G, en mettant en lumière leurs caractéristiques techniques et leurs limites. Ensuite le deuxième chapitre est consacré à la programmation des réseaux via les approches SDN et NFV, qui constituent les fondements des réseaux intelligents. Enfin le troisième chapitre aborde l'application du machine learning à l'allocation des ressources dans les réseaux 5G, avec une étude pratique menée sur un jeu de données réel. Ce travail a pour ambition de démontrer, par une approche concrète, la pertinence de l'apprentissage automatique pour l'optimisation de la gestion des réseaux 5G, et d'ouvrir la voie à l'intégration de méthodes plus avancées comme le deep learning dans les développements futurs.

CHAPITRE 1 : LES RESEAUX MOBILES

1.1 Introduction :

Au cours des deux dernières décennies, le secteur des communications mobiles a connu un essor impressionnant, stimulé par un besoin grandissant des consommateurs pour des services plus performants, plus sûrs et plus faciles d'accès. Cette progression a engendré une série de générations de réseaux mobiles, partant de la toute première (1G), axée sur le transfert analogique de la voix, pour arriver à la cinquième (5G), qui se distingue par des capacités considérablement améliorées. Chaque nouvelle génération a apporté son lot d'améliorations significatives : la 2G a inauguré la communication numérique, la 3G a rendu possible l'accès à Internet depuis un appareil mobile, et la 4G a accru la vitesse de transmission des données ainsi que la qualité du streaming. En dépit de ces avancées, les réseaux 4G montrent des faiblesses, notamment en matière de temps de réponse, de capacité et de consommation d'énergie. La 5G se propose donc de relever ces défis avec de grandes ambitions technologiques, en intégrant des applications novatrices comme l'Internet des objets (IoT), les voitures autonomes, ou encore la réalité virtuelle.

L'objectif de ce chapitre est de revenir sur cette transformation technologique, à mettre en évidence les avantages et les inconvénients de chaque génération, et à présenter les enjeux liés à la 5G, afin de mieux saisir le contexte dans lequel s'inscrit notre problématique d'amélioration intelligente de la distribution des ressources.

1.2 Évolution des réseaux mobiles

Un réseau mobile est un système de réseau cellulaire qui permet l’utilisation simultanée de téléphone en mouvement ou immobile, y compris lors de déplacement à grande vitesse et sur de grandes distances. Il est conçu suivant une architecture de cellules qui permet la réutilisation des fréquences. Il permet aussi aux utilisateurs en mouvement de changer de cellule sans interruption de la communication en cours (handover). Chacune de ces Cellules desservies par au moins un émetteur-récepteur a localisation fixe, connu par la station de base. La station de base assure la connexion entre les téléphones mobiles et le réseau téléphonique.[1] Leur évolution suit une logique d'avancées technologique incessantes, alimentée par une demande grandissante pour la connexion, la capacité à se déplacer et des prestations à haute valeur ajoutée. Depuis l'aube des échanges sans fil, les réseaux mobiles ont connu différentes générations, chacune représentant un saut technologique important par rapport à la précédente. La première génération, souvent désignés par l'appellation 1G, ont fait leur apparition en 1980 dans le paysage des

télécommunications. Toutefois, ces infrastructures ont été progressivement délaissées au profit de la seconde génération, ou 2G, lancée en 1991. Il est intéressant de noter que l'on peut identifier deux évolutions distinctes au sein même de cette 2G : les versions 2.5 et 2.75. Le standard GSM (Global System for Mobile Communications) est l'un des principaux protocoles utilisés par la 2G. Contrairement à la 1G, cette deuxième génération de normes offre un éventail de services plus large, comme l'accès à Internet via le protocole WAP. Quant à la troisième génération, la 3G, elle se distingue par sa capacité à offrir la navigation Internet et le transfert de données. La 4G ou LTE (Long Term Evolution), el offre un débit élevé. Enfin, Pour ce qui est de la toute dernière génération la 5G, les technologies IMT-2020 (International Mobile Telecommunications-2020) offre un débit très élevé et recèlent un potentiel de transformation considérable pour améliorer notre quotidien.

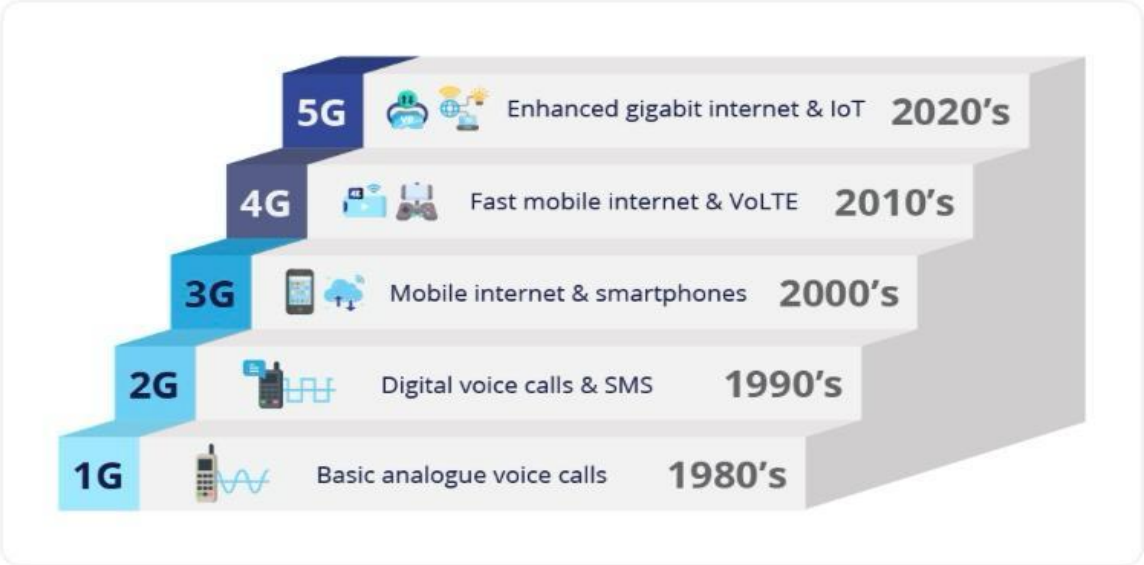


Figure 1.1 : Evolution des réseaux mobiles[4]

1.2.1 Les technologies de première génération (1G) et de deuxième génération (2G)

La toute première génération (1G) s'appuyait sur un système de communication mobile analogique, avant tout imaginée pour offrir un service de voix. Sa manière de moduler la radio ressemblait beaucoup à ce que les stations FM utilisaient. Pour ce qui est de la couverture, chaque cellule s'étendait sur environ de 2 à 20 km [2]. Cette 1G, matériel plutôt imposant, présentait des inconvénients assez clairs : qualité sonore parfois médiocre, un nombre limité d'utilisateurs possibles, une batterie qui tenait difficilement,

un coût assez élevé, et une sécurité laissant à désirer car les signaux analogiques ne se prêtaient pas vraiment à des techniques de cryptage sophistiquées, sans parler d'un réseau souvent saturé. Elle fonctionnait aussi avec des normes comme AMPS, NMT, TACS et Radiocom2000.[1]

Le réseau mobile 2G, ou GSM, représente la toute première incursion dans le monde du numérique. Cette technologie prend en charge diverses tailles de cellules, allant jusqu'a 35 kilomètres, en exploitant les macro, micro, pico et femto cellules. Elle améliore considérablement la qualité audio, assure la confidentialité des échanges et offre une capacité accrue tout en restant abordable pour l'utilisateur [1]. De plus, la 2G est apte a fournir des SMS, ou Short Messaging Service, et des photos avec MMS, ce qui permet de pallier certaines des lacunes de la 1G [2]. On constate aussi une évolution de cette génération avec la 2.25G (GPRS) offrant un débit de 171,2 kbps, ainsi que la 2.75G (EDGE) avec un débit de 384 kbps [3].

La figure 1.2 illustre l'architecture de la 2G, qui comprend une partie radio composée du sous-système de station de base (BTS, BSC). Ce dernier assure la connexion des équipements des utilisateurs a la partie commutation du réseau, incluant les éléments tels que MSC, VLR, HLR et AuC.

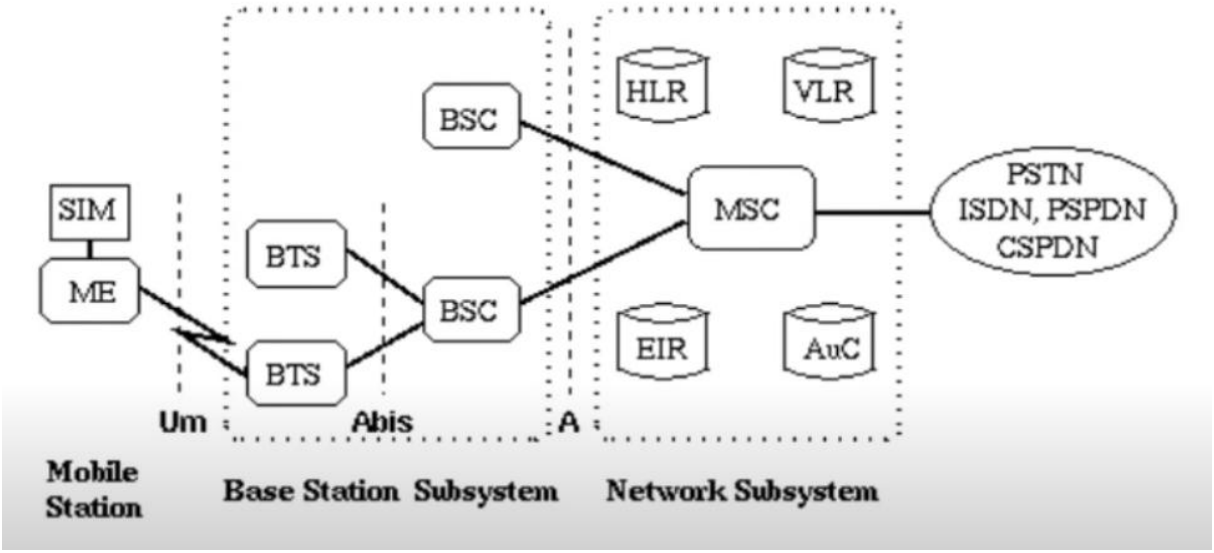


Figure 1.2 : Architecture 2G [4]

1.2.2 Les technologies de la troisième génération (3G)

La 3G offre une communication accélérée, qu'il s'agisse de la voix, de la télécopie ou d'internet, où que vous soyez et à toute heure [1]. La 3G, en Europe, on l'a plutôt nommée UMTS, tandis qu'en Amérique, c'était CDMA2000 [4]. Il y a eu aussi des

améliorations de la 3G : la 3.5G (HSPA) avec un débit théorique maximal de 14,4 Mbit/s, mais en pratique plutôt autour de 3,6 Mbit/s, la 3.75G (HSPA+) avec un débit théorique maximal de 21 Mbit/s, et environ 5 Mbit/s en conditions réelles, et enfin, le DC-HSPA+, remarquable pour un débit théorique maximal de 42 Mbit/s, et concrètement plus de 10 Mbit/s [1].

Concernant l'organisation de la 3G, la section radio est constituée des NodeB et des RNC. Ces éléments connectent les appareils des utilisateurs au réseau central, comme on peut le voir sur le schéma 1.3.

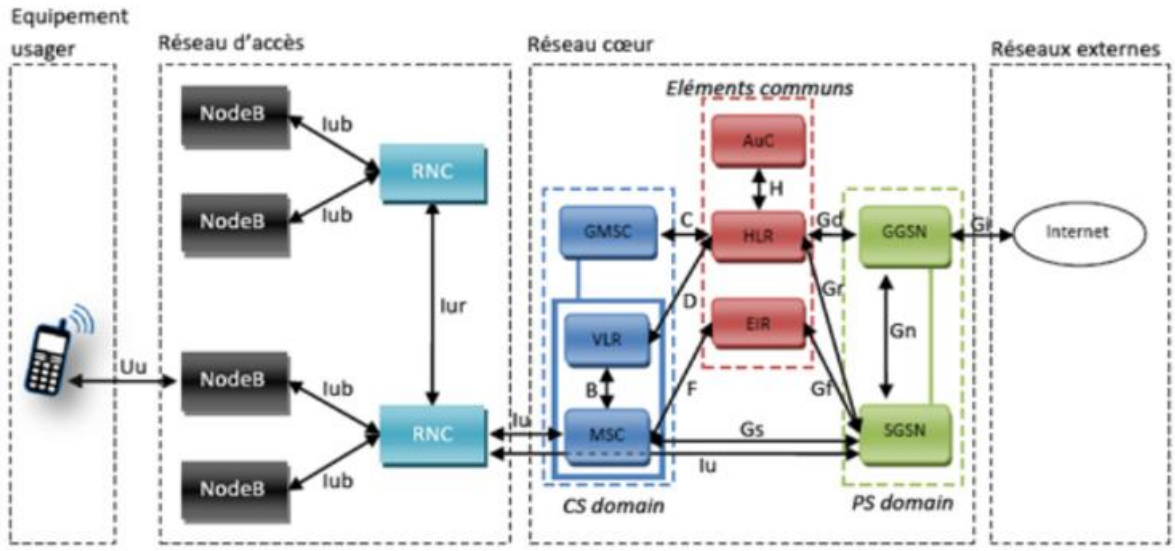


Figure 1.3 : Architecture 3G [4]

1.2.3 La quatrièmes génération (4G)

La 4G, ou LTE, représente une avancée considérable si on la compare à la 2G et à la 3G. Avec la 4G, on constate une nette amélioration de la vitesse de transmission (débit) et un temps de réaction fortement (latence) diminué. La 4G propose les mêmes possibilités que la 3G, tout en ajoutant des prestations en plus. De plus, elle assure un accès universel à divers services, quels que soient l'instant et le lieu, en fonction des besoins de l'usager. Toutefois, il s'agit d'un système de réseau qui repose sur le protocole IP [3].

Comme on peut le voir sur le schéma 1.4, contrairement aux deux configurations antérieures, la section radio, qui comprend les eNodeBs, est directement reliée au réseau central.

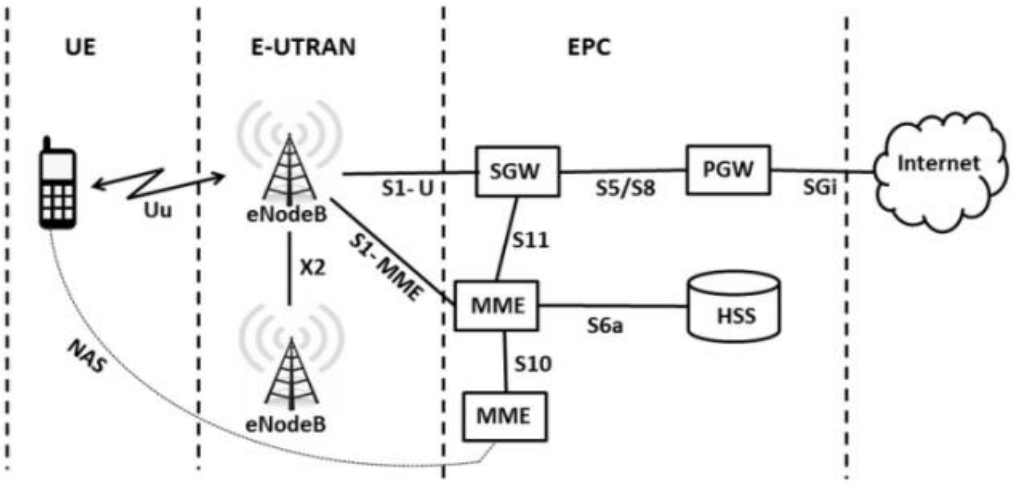


Figure 1.4 : Architecture 4G [4]

Le tableau 1.1 met en lumière certaines particularités propres à chaque génération de réseau mobile, allant de la 1G à la 4G, telles que les largeurs de bande, les débits de données, les fréquences utilisées et les différentes méthodes de commutation.

Tableau 1.1 : Comparaison de la 1G à la 4G

Caractéristiques	1G	2G	3G	4G
Année d'utilisation	1980-1990	1990-2000	2010-2010	2010-2020
Largeur bande	30KHZ	1,8GHZ	1,6-2GHZ	2-8GHZ
Débit	2,4Kb/s	9,6Kb/s 64Kb/s	144Kb/s 2Mb/s	100Mb/s 1Gb/s
Fréquence	869-894 MHz 824-849 MHz	900MHz 1800 MHz	900 MHz 2100 MHz	1800 MHZ 2600 MHz
Mode commutation	Circuit	Circuit	Circuit-paquets	Paquets

1.3 Limites des réseaux (4G)

Bien que la quatrième génération (4G) ait été un bond en avant notable pour les télécommunications mobiles, elle révèle certaines faiblesses techniques et pratiques face aux besoins grandissants des utilisateurs et des nouvelles utilisations.

Tout d'abord, l'un des principaux obstacles de la 4G est qu'elle peine à gérer un nombre considérable d'appareils connectés en même temps. Avec l'essor de l'Internet des Objets (IoT), qui inclut des millions d'appareils connectés dans un espace limité (capteurs, montres, voitures, maisons connectées, etc.), la 4G atteint vite ses limites en termes de densité de connexion par kilomètre carré.

Deuxièmement, la 4G affiche des performances qui ne suffisent pas pour des utilisations qui nécessitent un temps de réponse très court, comme la chirurgie à distance, la réalité virtuelle (VR), la réalité augmentée (AR) ou encore les voitures autonomes. En effet, le temps de réponse moyen de la 4G est d'environ 30 à 50 millisecondes, alors que certaines utilisations cruciales demandent des temps de réponse inférieurs à 1 milliseconde.

Troisièmement, malgré des débits élevés (jusqu'à 100 Mbit/s dans des conditions idéales), la 4G n'arrive pas toujours à assurer un débit stable et une qualité de service (QoS) optimale dans des endroits très fréquentés ou en mouvement (gares, stades, villes, etc.). Cela entraîne une baisse de la qualité pour l'utilisateur, surtout pour les services de diffusion en continu (streaming, jeux en ligne, visioconférences).

Quatrièmement, la 4G utilise plus d'énergie sur les appareils mobiles que les normes plus récentes, ce qui a un effet négatif sur l'autonomie des batteries, surtout pour les appareils IoT à faible consommation.

Enfin, la 4G ne répond pas naturellement aux nouveaux défis de la virtualisation, de la flexibilité du réseau et de l'intelligence artificielle appliquée aux infrastructures. Les réseaux 4G sont peu compatibles avec les concepts modernes comme le network slicing, la programmation du réseau (SDN) ou encore la virtualisation des fonctions réseau (NFV), qui sont essentiels pour les réseaux de nouvelle génération.

Ces limites ont donc encouragé le développement et le déploiement de la cinquième génération (5G), conçue pour répondre aux futures exigences en matière de connectivité, de performance et d'intelligence du réseau.

1.4 La cinquième génération (5G)

La 5G succède à la 4G en tant que nouveau réseau mobile. Elle a été officiellement baptisée IMT-2020 par l'UIT en 2015. La 5G est sur le point d'offrir la vitesse et la sureté des transmissions sans fil, jusqu'ici confinées aux infrastructures de fibre optique. L'application IMT-2020 dépassera les limites de la voix et du son pour s'étendre a des secteurs diversifies tels que la sante, l'automatisation industrielle, la réalité virtuelle, la conduite autonome et la robotique [5].

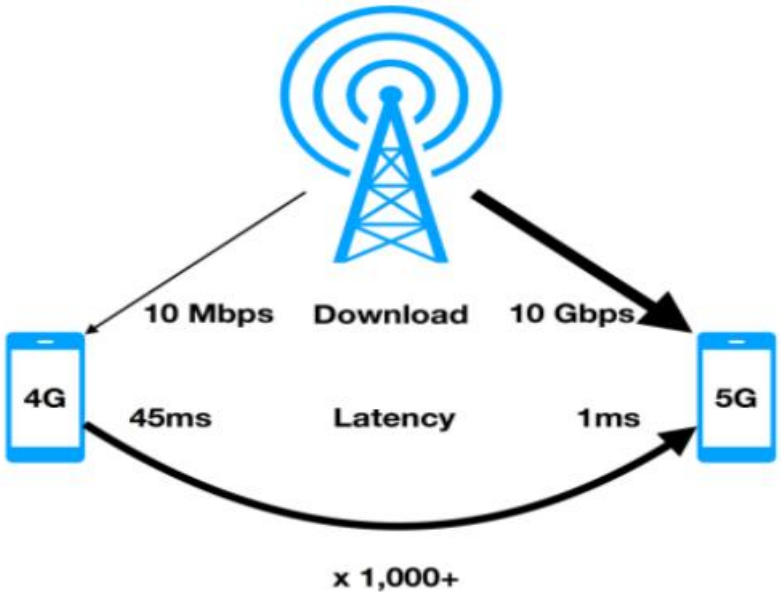


Figure 1.5 : Comparaison 4G/5G [5]

1.4.1 Exigence de la 5G

L'IUT-R a identifié trois cas d'usage principaux pour la 5G :

— emBB : On parle ici de communication à très haut débit. Dans la lignée des technologies précédentes, cette option permet de gérer l'explosion de l'utilisation des données mobiles, avec une vitesse dix fois supérieure à celle de la 4G (10Gbps). Les exemples typiques dans ce domaine sont les flux vidéo de plus en plus volumineux, ainsi que la réalité virtuelle, augmentée ou mixte.

— mMTC : Il s'agit de la connexion d'un très grand nombre d'appareils. Le nombre de connexions dépasse les 10 millions par MHz et par cellule, et excède 1 million par km2. L'internet des objets et les villes intelligentes, par exemple, sont des illustrations parfaites de cette catégorie.

— URLLC : C'est le volet où les communications sont qualifiées de critiques. Ici, les communications doivent être fiables et avec une latence très faible (1ms). Les applications les plus courantes sont les voitures autonomes, les services médicaux et les services de sécurité et d'urgence. Une panne ou une interruption de la communication peut avoir des conséquences désastreuses, et la transmission doit être aussi rapide que possible [5].

Les besoins en matière de performance de la 5G sont structurés autour de huit aspects clés, incluant l'augmentation du débit de données, l'optimisation de la consommation d'énergie, la prise en charge de la mobilité, la capacité de gestion du trafic par zone géographique, le débit moyen offert à chaque utilisateur, une forte densité de connexions simultanées, l'utilisation efficace du spectre radioélectrique et la réduction du temps de latence, comme le montre la figure 1.5 [6,7].

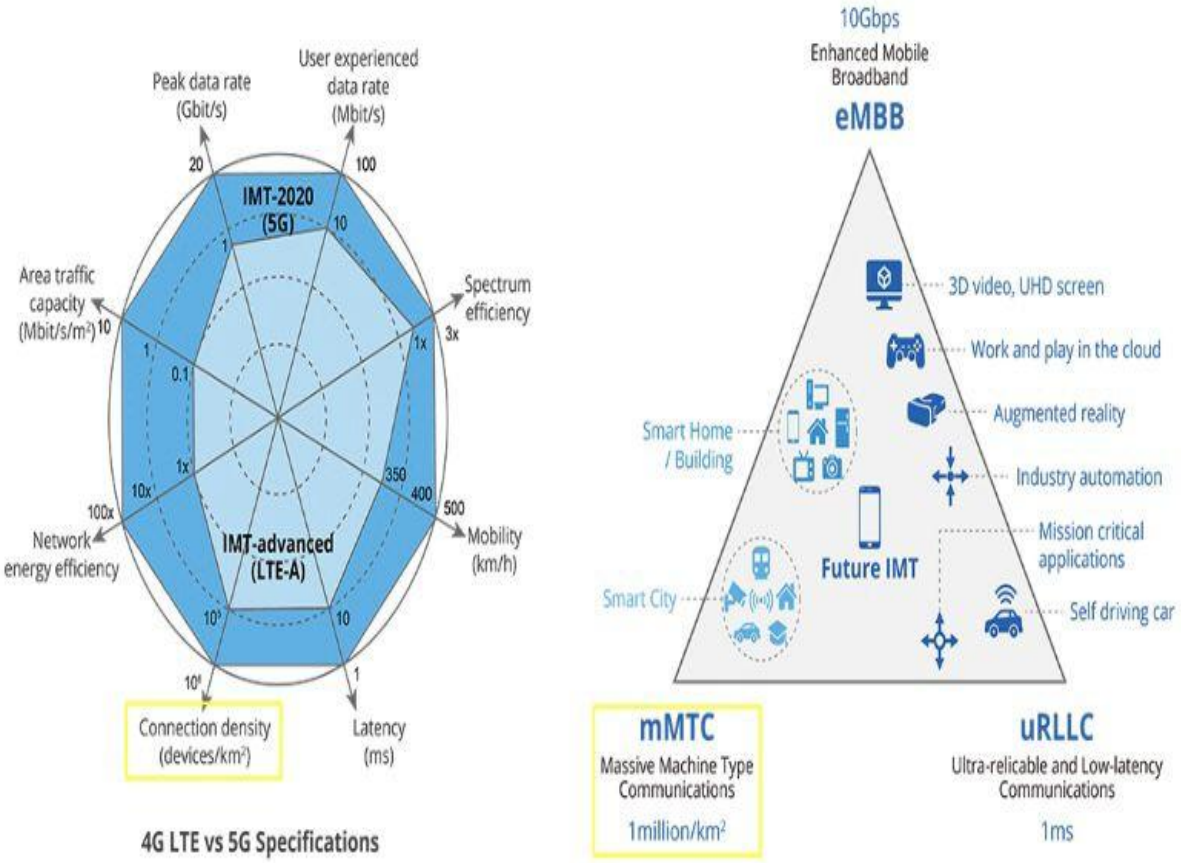


Figure 1.6 : Les exigences de capacite de la 5G [6 ,7]

Un aspect essentiel de la 5G réside dans la segmentation du réseau, souvent désignée par le terme anglais « Network slicing ». Ce découpage permet aux opérateurs de diviser le réseau en plusieurs réseaux virtuels distincts au sein de la même infrastructure physique. Cela offre la possibilité d'allouer dynamiquement des ressources en fonction des exigences spécifiques des utilisateurs. En outre, face à la nécessité de proposer de nouveaux services, il devient superflu de déployer un réseau entièrement nouveau. Il suffit simplement de virtualiser une nouvelle tranche pour déployer rapidement les services demandés [7].

1.4.2 Architecture de la 5G

En matière d'architecture, la façon dont le réseau 5G est pensé – selon les versions R14, R15, R16 – est de garder le cœur du réseau 4G LTE, que ce soit en mode NSA ou SA. En mode NSA, le eNodeB et le gNodeB fonctionnent ensemble dans le RAN, utilisant donc EPC et 5GC comme cœur de réseau. Par contre, en mode SA, le gNodeB fonctionne seul sur le RAN, et c'est le 5GC qui devient le cœur du réseau, comme on le voit sur la figure 1.7 [7]. L'idée du mode NSA, c'est de continuer à se servir du cœur du réseau 4G tout en ajoutant petit à petit les technologies 5G pour exploiter les hautes fréquences en 5G. Mais avec le mode SA, le cœur du réseau passe complètement à la 5G. Du coup, l'architecture ne dépend plus du tout de la 4G LTE. Dans ce cas, on peut utiliser les fréquences basses et hautes.

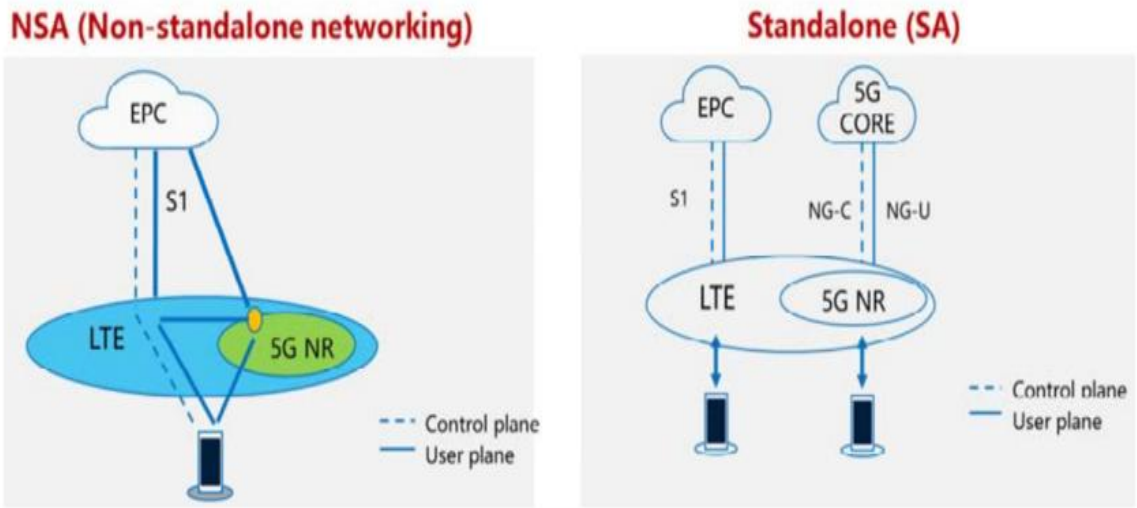


Figure 1.7 : Architecture 5G en mode NSA et SA [7]

La 5G s'articule autour d'un réseau d'accès radio nouvelle génération, le NG-RAN, ainsi que du cœur de réseau NGC. Les liaisons sans fil du réseau 5G incluent les interfaces Xn, qui assurent la connexion entre les gNB, et les NG-C et NG-U, qui relient les gNB au NGC, comme on peut le voir sur la figure 1.8.

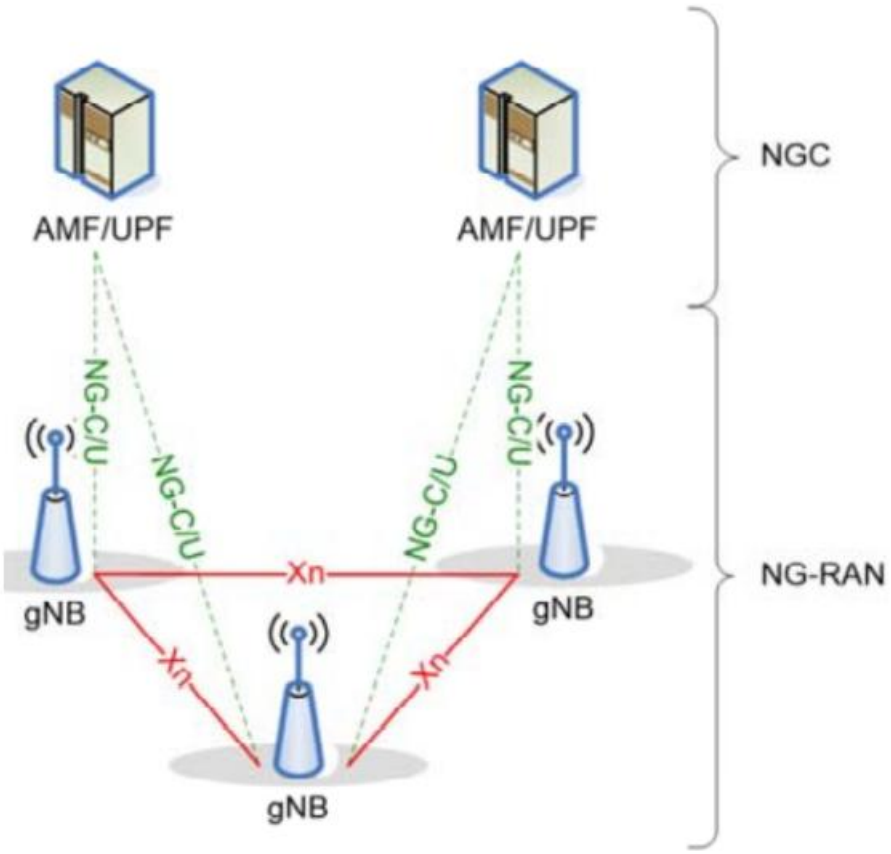


Figure 1.8 : Architecture 5G [8]

Dans le tableau 1.2 [8], on retrouve les rôles divers des éléments du réseau cœur 5G, notamment AMF, AUSF, AF, SMF, UPF, PCF, SSF, UDM, NRF, NEF. On constate aussi que les fonctions réseau de la 5G découlent en réalité des fonctions réseau cœur de la 4G. Ainsi, une part importante des fonctions du NGC 5G a une équivalence dans l'EPC, comme illustré dans le tableau 1.3 [7].

Tableau 1.2 : Les fonctions des modules de 5G NGC [8]

Modules	Fonctions
AMF	Authentification, Autorisation, Compatibilité basées sur UE, Contrôle d'accès et gestion de la mobilité, Orchestration de la sélection des tranches réseaux, Contrôle de SMF et indépendance de la technologie d'accès, Terminaison et s´sécurité de la signalisation NAS, Contrôle de sécurité de la couche d'accès, Signalisation de nœuds inter réseau cœur pour la mobilité entre les réseaux d'accès
AUSF	Stockage des données pour l'authentification UE
AF	Fourniture d'information de flux de paquets pour PCF
SMF	Gestion des sessions basée sur des stratégies réseau (configuration, maintenance. . .), Attribution d'adresse IP aux UEs, Sélection et contrôle de l'UPF pour le transport des données, Notification des données DL, Configuration de la direction du trafic de UPF pour acheminer le trafic vers la destination appropriée
UFP	Gestion de l'emplacement et de la configuration basée sur un type de service Routage et transfert de paquets Inspection de paquets du trafic UP Rapport sur l'utilisation du trafic. . .
PCF	Fourniture d'un cadre stratégique pour le roaming, network slicing, gestion de la mobilité, gestion de la session et prise en charge de QoS
	Sélection de la tranche réseau

SSF	
UDM	Stockage des données d'abonné et des fichiers de l'UE
NRF	Assistance au nœud de d´découverte dans les services de réseau de découverte
NEF	Exposition et publication des données réseau

Tableau 1.3 : Les fonctions des modules communs de EPC et 5G NGC [7]

Modules EPC	Fonctions	Module 5G NGC
MME	Gestion de la mobilité Gestion de l'authentification Gestion de la session PDN	AMF AUSF SMF
PDN-GW	Gestion de la session PDN Transfert des données du plan utilisateur	SMF UPF
SGW	Transfert des données du plan utilisateur	UPF
PCRF	Contrôle de la charge et des politiques	PCF
HSS	Base de données des abonnés	UDM

1.4.3 Les caractéristiques de la 5G

• Les bandes de fréquences employées

Opter pour une certaine bande de fréquences implique forcément un équilibre à trouver entre la portée et les capacités. Par conséquent, pour assurer une couverture étendue, il est préférable de se tourner vers des fréquences basses, tandis que pour optimiser les capacités, on privilégiera des fréquences plus hautes, étant donné que les largeurs de bandes y sont plus conséquentes. Néanmoins, les conditions de propagation s'avèrent plus ardues dans les hautes fréquences en raison de l'environnement. Ainsi, les réseaux mobiles doivent trouver le juste milieu afin d'être en mesure de fournir une couverture adéquate.

La 5G dispose d'une gamme de fréquences :

-Les fréquences bases (inferieures à 1GHz)

-Les frequences moyennes (sub-6GHz, entre 2,4 GHz et 6GHz)

-Et les ondes millimétriques Mmwave (supérieures à 24GHz)

[7,9], ce qui est illustre sur la figure 1.9

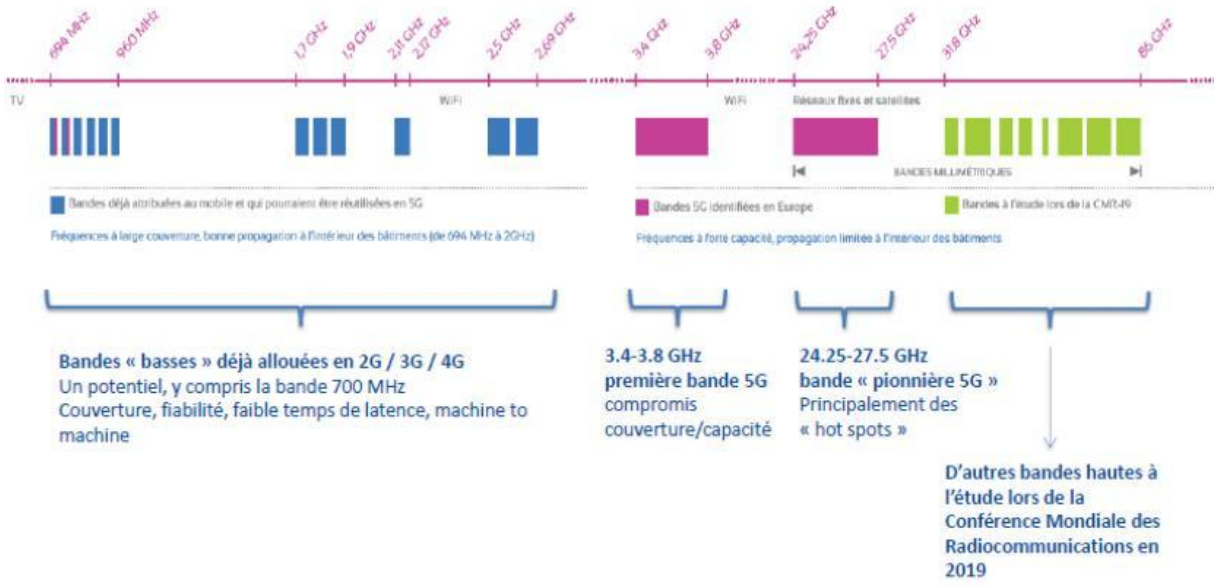


Figure 1.9 : Bandes de frequences 5G [9]

• Méthodes de modulation

En 5G, on s'appuie sur la modulation QAM, autant pour envoyer que pour recevoir des données (liaison montante et descendante). C'est une façon d'accélérer le débit. Les types de modulation QAM employés en 5G sont le 16QAM, le 64QAM et le 256QAM. Si l'on cherche à avoir des débits rapides, on privilégie les modulations d'ordre élevé. Toutefois, on se heurte fréquemment à une certaine sensibilité au bruit. C'est pourquoi le 256QAM est privilégié quand la connexion est bonne ; sinon, il serait plus judicieux de se rabattre sur les 64QAM ou 16QAM. Comme la 4G, la 5G emploiera l'OFDMA, mais avec une flexibilité et une souplesse bien plus grande dans la répartition temporelle et fréquentielle, afin de pouvoir répondre aux besoins d'utilisateurs ayant des contraintes très diverses. La 5G fait aussi appel au GFDM, une technique de multiplexage non orthogonal qui pallie les défauts de l'OFDM tout en conservant ses atouts, comme la suppression de l'interférence inter-symbole [10].

• Le MIMO Massif

Le MIMO Massif, c'est un réseau d'antennes à grande échelle, en quelque sorte. Ce qui rend cette technologie particulière, c'est l'utilisation d'un grand nombre de micro-antennes "intelligentes" regroupées sur un seul équipement. La technique de pondération des antennes, qui sert à gérer les réseaux d'antennes, améliore la couverture sans fil de manière significative. Pour la 4G, le MIMO Massif était une option, mais pour la 5G, c'est devenu une nécessité incontournable. Le MIMO Massif repose sur trois idées fondamentales : la diversité spatiale, le multiplexage spatial et la formation de faisceaux. Avec la diversité spatiale, les mêmes données sont envoyées à travers différents chemins. Alors que la diversité spatiale renforce la fiabilité de la liaison radio, le multiplexage spatial, lui, augmente la capacité de cette liaison, car il permet d'envoyer plusieurs flux de données entre l'émetteur et le récepteur, ce qui booste considérablement le débit, comme le montre la figure 1.11.

La formation de faisceau est une autre astuce qui travaille main dans la main avec le MIMO massif pour doper le débit et la capacité. Elle s'appuie sur des technologies d'antennes pointues pour concentrer le signal sans fil dans une direction précise, au lieu de le diffuser à tout vent. Cette méthode réduit les interférences entre les faisceaux orientés dans des directions différentes, ce qui ouvre la porte au déploiement de réseaux d'antennes plus vastes [11,12], comme on peut le voir sur la figure 1.10.

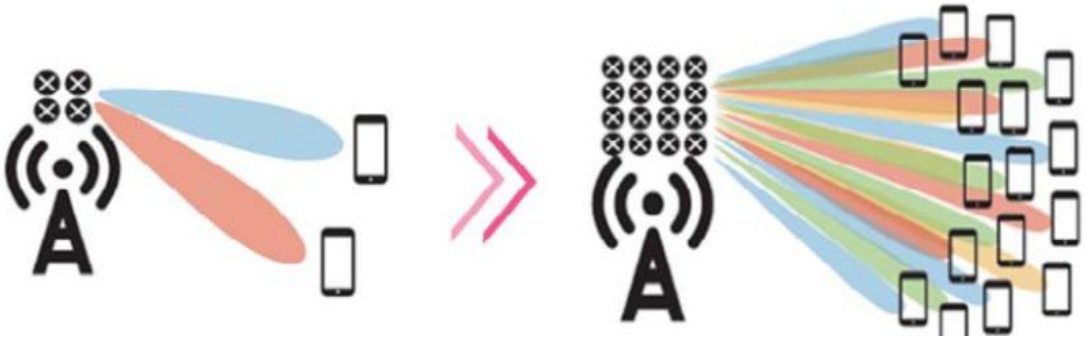


Figure 1.10 : MIMO Massif [12]

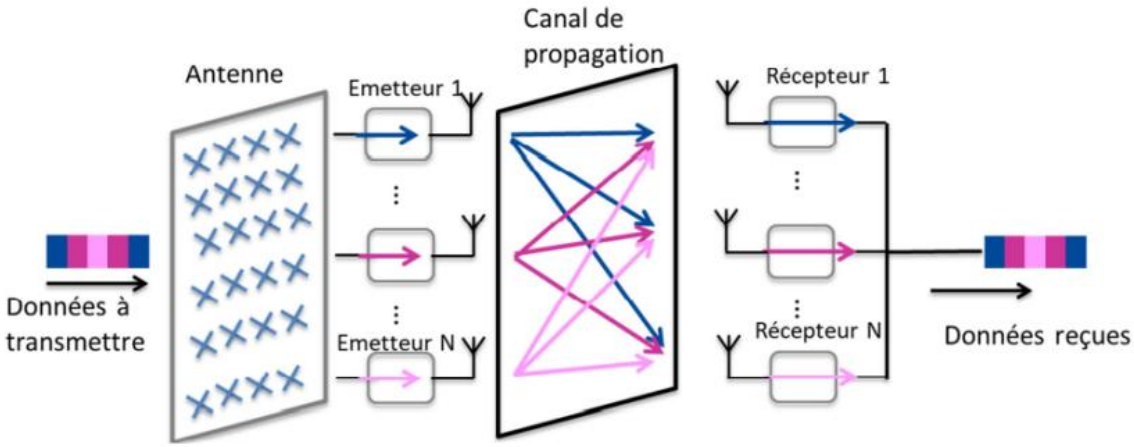


Figure 1.11 : Technique de MIMO Massif [11]

• Communication full-duplex

La transmission sans fil en mode full-duplex est actuellement un point central des discussions portant sur la définition de la 5G. Cette technologie autorise l'envoi et la réception simultanés, utilisant le même laps de temps et la même bande de fréquences, ce qui pourrait théoriquement multiplier le débit par deux. Ceci contraste avec le système half-duplex, où l'émission et la réception des signaux se déroulent dans deux périodes distinctes ou sur deux bandes de fréquences différentes. De plus, l'émetteur-récepteur ne peut pas émettre et capter un signal au même moment sur la même fréquence [12], comme illustré à la figure 1.12.

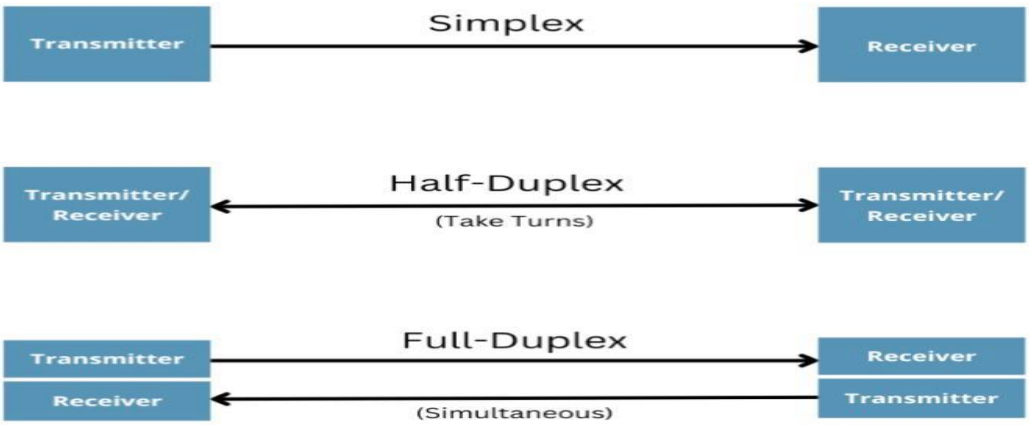


Figure 1.12 : Full-duplex [12]

• Méthodes de codage

Pour le codage du canal, le canal de données emploie le code LDPC, tandis que le canal de contrôle utilise le code polaire. Les codes LDPC, grâce à leurs performances supérieures et à un délai de décodage réduit, sont particulièrement adaptés aux transmissions de données impliquant de grands volumes et nécessitant de faibles latences. Parallèlement, le codage polaire est privilégié comme méthode de codage pour le canal de contrôle, en raison de ses excellentes performances lors de la transmission de petits paquets [13].

1.5 Conclusion

Le passage des réseaux mobiles, de la première à la quatrième génération, a permis une amélioration graduelle de la qualité des services fournies aux usagers. Néanmoins, la 4G montre ses limites face aux nouvelles attentes, surtout en matière de temps de latence réduit, de quantité grandissante d'objets connectés, ainsi que de souplesse. Ces bornes ont poussé à la création de la 5G, une technologie élaborée pour satisfaire ces impératifs et autoriser l'insertion de solutions futées. La 5G ne se limite pas à une simple amélioration des débits ; elle introduit de nouvelles architectures, des paradigmes de gestion dynamique des ressources telles que l'apprentissage automatique. Dans le prochain chapitre, nous allons examiner de près les toutes dernières méthodes de conception des réseaux, en nous penchant sur des technologies telles que le SDN (réseaux définis par logiciel) et la NFV (virtualisation des fonctions réseau), qui sont absolument essentielles à la mise en place des réseaux 5G.

CHAPITRE 2 : PROGRAMATION DES RESEAUX (SDN/NFV)

2.1 Introduction

Avec la connectivité en constante évolution, la variété des services et les impératifs de performance qu'impose la 5G, les architectures réseau classiques, souvent inflexibles et lentes à réagir, montrent leurs limites. Pour surmonter cette complexité grandissante, des solutions novatrices axées sur la programmabilité des réseaux ont vu le jour, où le Software Defined Networking (SDN) et la Network Function Virtualization (NFV) jouent un rôle clé. Ces technologies favorisent une administration plus agile, dynamique et automatisée des infrastructures réseau, en séparant le plan de contrôle du plan de données (côté SDN) et en virtualisant les fonctions réseau traditionnellement ancrées dans des équipements physiques (via la NFV). C'est un véritable tournant dans la conception et l'exploitation des réseaux actuels. Ces deux approches ouvrent la porte à une gestion affinée des ressources, surtout lorsqu'elles sont associées à des méthodes d'intelligence artificielle comme le machine Learning. Ainsi, le SDN et la NFV ne sont pas seulement des composantes essentielles de la 5G : elles sont aussi des moteurs d'une allocation des ressources adaptable, prédictive et performante, prête à relever les défis de la connectivité future.

2.2. Systèmes distribués traditionnels

Les réseaux traditionnels distribués ou systèmes distribués traditionnels, sont des réseaux informatiques où les composantes du programme et les données sont reparties sur plusieurs ordinateurs (nœuds) qui communique entre eux pour accomplir une tâche.

La figure 2.1 illustre la manière dont les données transitent de l'expéditeur au destinataire, en passant par sept couches distinctes, chacune ayant un rôle précis à jouer dans le traitement des données. Un réseau classique fonctionne sur le principe d'un réseau distribué articulé autour de trois plans essentiels :

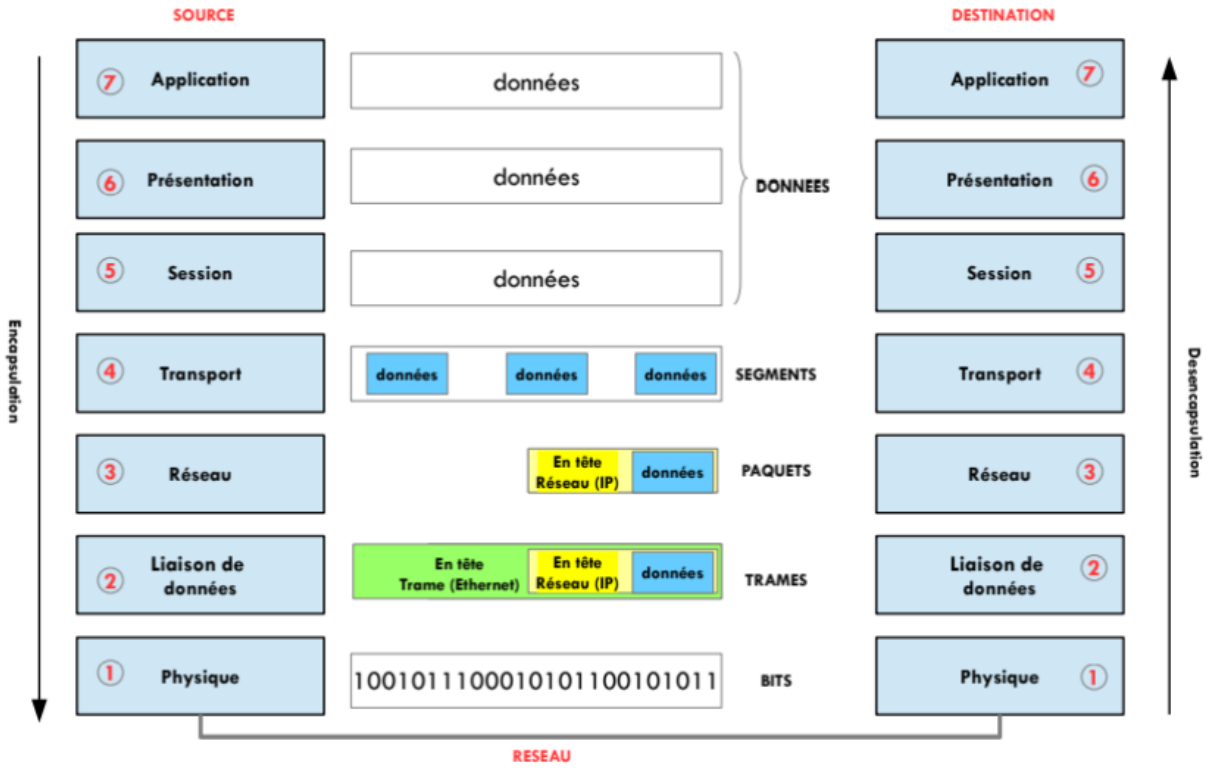


Figure 2.1 : Encapsulations désencapsulassions du modèle OSI [22]

Le plan de gestion, le plan de contrôle et le plan de données. Habituellement, le plan de contrôle et le plan de données sont intégrés au matériel physique. L'architecture interne des équipements réseau s'organise autour de ces trois plans d'opération :

—Plan de gestion : il s'occupe des interactions avec les utilisateurs externes et des tâches administratives, comme l'authentification et la configuration via diverses interfaces.

— Plan de contrôle : il assure le bon fonctionnement interne des appareils, fournit les directives pour acheminer les paquets, exécute les protocoles de routage et de commutation, et transmet les données opérationnelles au plan de gestion.

— Plan de données : c'est en quelque sorte la salle des machines, chargée de déplacer les paquets à travers le périphérique, en se basant sur la table de transfert fournie par le plan de contrôle pour déterminer le port de sortie [22].

En cas de modifications sur le réseau, qu'il s'agisse d'un changement d'état ou d'une mise à jour, l'information est automatiquement diffusée à l'ensemble du réseau. De plus, cette approche de gestion implique que chaque équipement est configuré de manière

indépendante et effectue ses calculs sans coordination avec les autres. Cela peut rendre le réseau plus lourd et plus difficile à gérer. Avec le temps, ces réseaux traditionnels sont devenus de plus en plus complexes, notamment en raison de la nécessité d'intégrer de nouveaux services et d'ajouter des équipements pour maintenir la qualité du service.

Les réseaux traditionnels présentent donc des limitations qu'il est impératif de surmonter pour garantir une qualité de service optimale. On peut identifier trois principales faiblesses : leur manque de flexibilité, leur complexité, les difficultés de provisionnement et la lenteur de la diffusion des informations [22].

2.3 Réseaux SDN

2.3.1 Historique et définition des réseaux SDN

Issu d'une tendance intellectuelle visant à programmer les dispositifs réseau et de la nécessité de surmonter certains défis, notamment la complexité de piloter et de faire progresser le réseau. L'appellation SDN fut initialement lancée par Martin Casado de l'université de Stanford en Californie, et des figures telles que Nick McKeown ou Scott Shenker ont aussi contribué au SDN. L'aboutissement de ces efforts avait pour objectif de rendre les réseaux adaptables, souples, progressifs et novateurs [23].

Le SDN se conçoit comme une structure séparant la strate (plan) de commande (contrôle) de la strate de données, dans le but de disposer d'une assise matérielle libérée de tout service réseau. Cette structure intègre une instance intelligente nommée contrôleur, qui perçoit le réseau dans son ensemble et permet d'intégrer les directives dans les divers équipements du réseau. De plus, ce contrôleur administre tous les appareils du réseau de manière centralisée [23,24].

La figure 2.2 illustre la distinction entre le réseau classique, où la strate de commande et la strate de données sont incluses dans le matériel physique, tandis que dans le réseau SDN, la strate de commande est séparée de la strate de données.

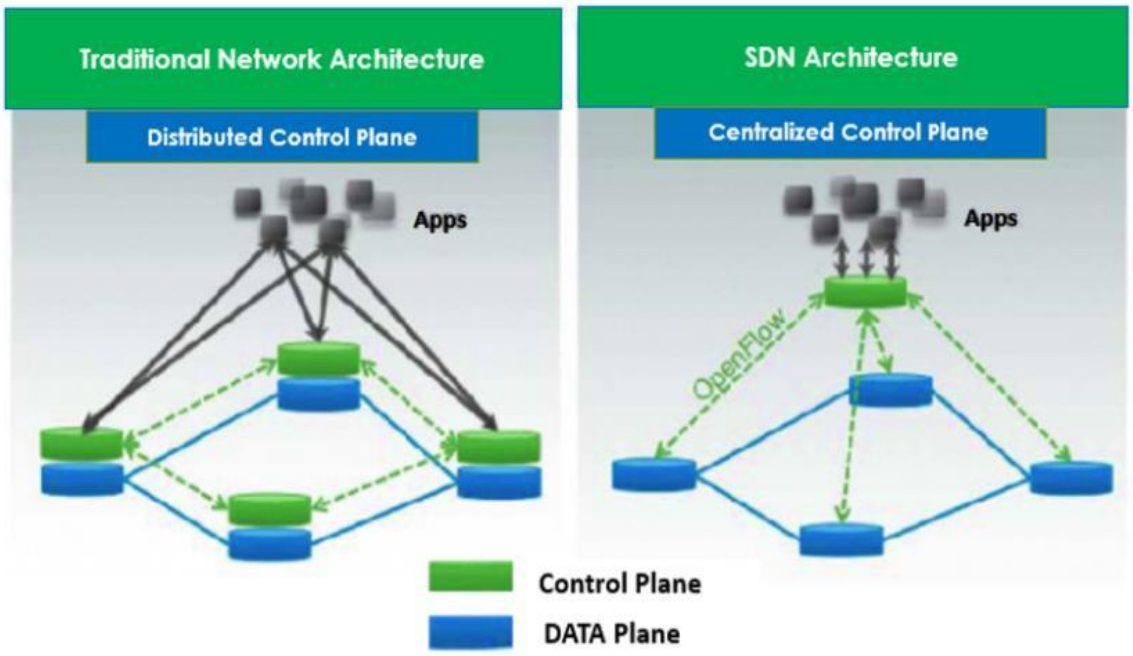


Figure 2.2 : Comparaison entre réseau traditionnel et réseau SDN [24]

2.3.2 Architecture détaillée des réseaux SDN

L'organisation d'un réseau reposant sur la technologie SDN se présente sous trois couches distinctes : la couche des applications, la contrôle et la couche infrastructure.

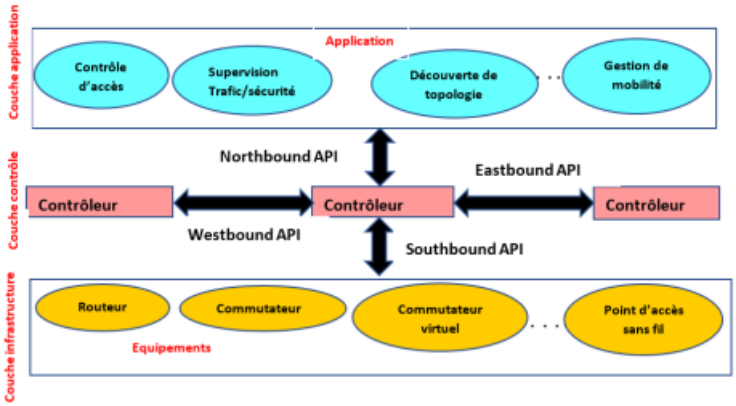


Figure 2.3 : Architecture détaillée d'un réseau SDN [23]

Chacune de ces couches possède une fonction propre, une délimitation claire et des interfaces programmable d'application (API) orientées nord, sud, est et ouest, qui facilitent l'´échange d'informations entre les différentes couches voisines, comme le montre la figure 2.3 [23].

2.3.2.1 La couche infrastructures ou plan de données

Cette couche s'articule autour d'infrastructures réseau comme des commutateurs, des routeurs, des entités virtuelles, et autres [22]. Son objectif principal est d'acheminer les échanges réseau selon des directives définies par le plan de contrôle. La liaison entre le plan de contrôle et le plan de données est assurée par des interfaces dites sud, désignées sous le terme d'API Southbound [24]. De plus, le plan de données est géré par un commutateur Openflow. Ce dernier est conçu avec une interface de programmation ouverte et standard, ce qui lui permet de garantir sa configuration et son aptitude à fonctionner avec le contrôleur SDN et les autres commutateurs SDN. Le commutateur Openflow et le contrôleur SDN dialoguent grâce au protocole Openflow. On distingue deux catégories de commutateurs Openflow : le commutateur Openflow only, qui ne prend en charge que le protocole Openflow, et le commutateur Openflow enable, qui cumule les fonctions d'un commutateur classique et celles d'un commutateur Openflow [23].

2.3.2.2 La couche de contrôle ou plan de contrôle

Il orchestre les décisions relatives au flux de données sur le réseau, en s'appuyant sur les impératifs des applications, les besoins des utilisateurs et les règles de communication réseau issues du plan de données. Au cœur de ce dispositif se trouve un élément clé : le contrôleur SDN. Ce dernier, doté d'une vision d'ensemble du réseau, en assure la gestion de manière centralisée. Le contrôleur SDN interprète les demandes des applications et les objectifs commerciaux, puis relaie ces informations aux composantes du plan de données. Via un canal sécurisé qui le connecte aux commutateurs OpenFlow, il échange des paquets, aussi bien en réception qu'en émission. Il est à noter que le nombre de contrôleurs peut varier en fonction de l'envergure du réseau [24].

Il existe une diversité de contrôleurs SDN, qui se différencient par leurs langages de programmation et les protocoles pris en charge. Parmi les plus répandus, on retrouve ONOS, Opendaylight, POX, RYU, Beacon et Floodlight (comme illustré dans le tableau 2.1). De plus, certains contrôleurs sont distribués, tels qu'ONOS et Opendaylight, tandis que d'autres sont centralisés, à l'image de Beacon et Floodlight [23]. Les contrôleurs SDN ont la capacité de communiquer entre eux grâce aux interfaces de programmation Westbound et Eastbound.

Tableau 2.1 : Quelques types de contrôleurs et caractéristiques [23]

Contrôleur	Langages de programmation	Développé par	Open source
NOX	C++	Nacira	Oui
POX	Python	Nacira	Oui
Trema	Ruby, C	Nec	Oui
Maestro	Java	RICE University	Oui
Beacon	Java	Stanford University	Oui
Floodlight	Java	Big switch	Oui
RYU	Python	ON. Lab	Oui
ONOS	Java	NTT	Nom
Helio	C	ONF	Oui
Opendaylight	Java Python	DreamersLab	Oui
HP VAN	Controller Java, REST	HP	Nom

2.3.2.3 La couche application ou plan d'application

C'est le niveau supérieur de l'édifice SDN, là où l'on conçoit et met en œuvre les directives pour gérer le réseau. Grâce au contrôleur SDN et à ses interfaces de programmation Northbound [22], elle offre la possibilité d'agir et de modifier la manière dont les équipements du réseau se comportent. Cette strate applicative englobe diverses applications conçues pour satisfaire les multiples exigences des utilisateurs, telles que l'automatisation, la souplesse et la capacité de programmation du réseau. Parmi les champs d'application SDN, on trouve l'optimisation du trafic, la virtualisation du réseau, la supervision et l'analyse du réseau, l'identification des services réseau, le contrôle d'accès, et bien d'autres. La logique de contrôle pour chaque application peut s'exécuter

comme un processus distinct, directement sur le matériel du contrôleur, au sein de chaque zone [22].

2.3.2.4 Les protocoles d'interfaces SDN

- **Le protocole OpenFlow**

Le protocole OpenFlow se présente comme le tout premier protocole d'interface Southbound standardise au sein de l'architecture de réseau SDN, et il est largement reconnu comme un composant essentiel de ce type de réseau. On constate que l'architecture OpenFlow introduit deux rôles matériels distincts : le contrôleur OpenFlow et le commutateur OpenFlow [22]. Le contrôleur OpenFlow assume la responsabilité du calcul des chemins et de la distribution des entrées de flux au niveau du plan de données, qui gère la transmission des paquets de données en fonction des flux d'entrée reçus. Quant au commutateur OpenFlow, il assure le transfert physique des paquets de données [23], en se basant sur les informations d'entrée de flux générées et transmises par le contrôleur.

Le commutateur OpenFlow est constitué de tables de flux, d'une table de groupe et d'un canal s´sécurisé, comme le montre la figure 2.4, facilitant ainsi les ´échanges avec le contrôleur OpenFlow. Une table de flux est essentiellement une liste d'entrées de flux. Un commutateur OpenFlow peut intégrer plusieurs de ces tables, qui sont ensuite traitées via un processus de pipeline [23].

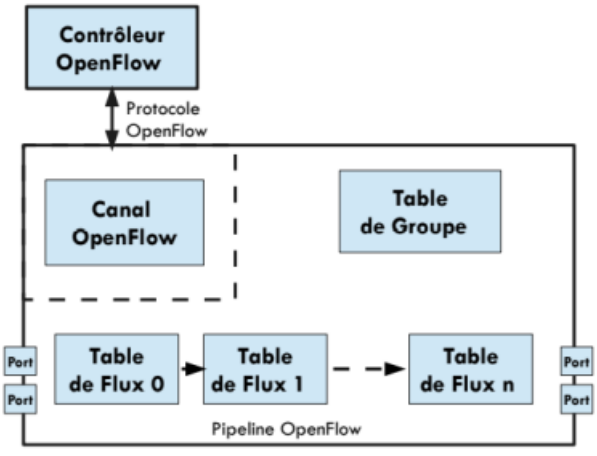


Figure 2.4 : Architecture d'un Commutateur Openflow [23]

On trouve aussi une table de groupes, qui, elle, répertorie différentes entrées de groupe. Chacune de ces entrées est liée à un ensemble d'actions spécifiques. Ainsi, si un paquet correspond à un critère prédéfini, ces actions sont d´enclenchées pour ce paquet [23].

Un commutateur Openflow gère trois genres de ports : le port physique, directement lié au matériel concret ; le port logique, un port de niveau supérieur sans lien avec les interfaces physiques ; et le port réservé, servant à préciser des actions de transfert comme l'envoi à un contrôleur, l'inondation ou le transfert direct des paquets [22].

On distingue trois types de tables : la table des flux, la table de groupe et celle de mesure.

Une table de flux se compose de :

— **Champ de correspondance** : il sert à apparier les paquets et à contrôler le port d'entrée ou l'en-tête du paquet pour appliquer des actions ;

— Priorité : établie par la table de flux précédente, elle permet de faire correspondre d'abord une entrée de flux à priorité élevée ;

— Compteur : il donne une vue du nombre de paquets traités ;

— Instructions : exécute des actions sur les paquets ;

— Délais d'expiration : c'est la durée maximale pour définir une expiration d'une entrée de flux ;

— Cookie : un identifiant d'entrée de flux choisi par le contrôleur, utile pour filtrer, modifier ou supprimer le flux ;

— Indicateur (flags) : modifie le mode de gestion des entrées de flux, comme illustré figure 2.5. Une table de flux débute par la table 0. S'il y a plus d'une table, un traitement pipeline s'impose. Quand un paquet arrive dans la première table, l'entrée de flux prioritaire est appariée et les instructions le dirigent vers la table suivante, et ainsi de suite. Si le paquet n'est pas dirigé vers la table suivante, le pipeline s'arrête et les

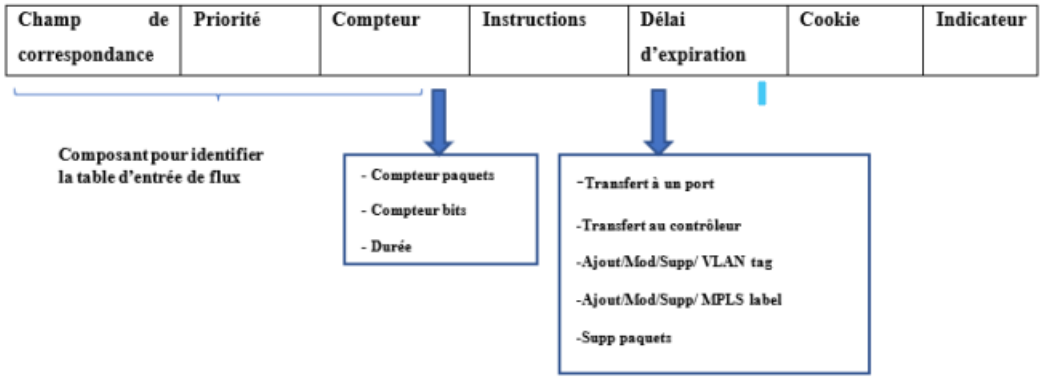


Figure 2.5 : Format de table de flux Openflow [23]

actions seront exécutées. Dans le cas où si un paquet ne correspond à aucune entrée de flux dans la table de flux, il est appelé table miss. Si cette dernière existe alors des actions sont exécutées dans le cas contraire, le paquet sera abandonné. Ce traitement est illustré sur la figure 2.6 sous forme d'organigramme.

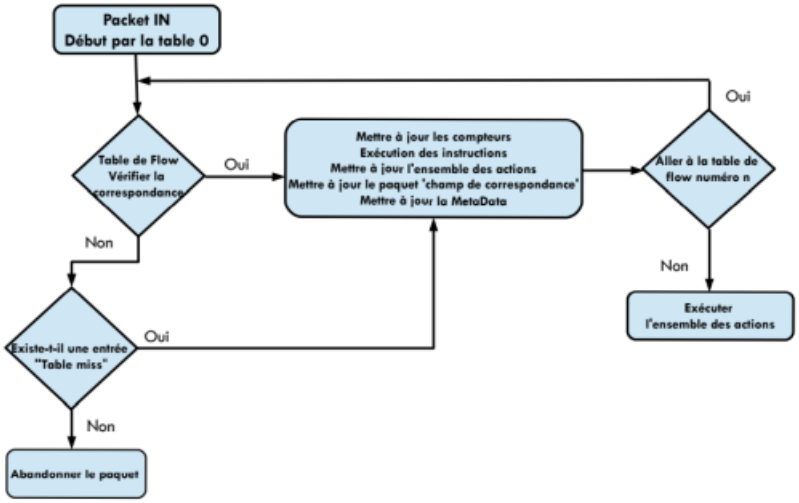


Figure 2.6 : Organigramme de traitement d'un paquet dans un commutateur openflow [23]

Pour la gestion des paquets, le commutateur s'appuie sur une TCAM, ou mémoire ternaire adressable par contenu, et une RAM, mémoire vive a accès aléatoire [23]. Le protocole OpenFlow prend en charge trois catégories de messages distinctes :

— **Les échanges entre le contrôleur et le commutateur :** ce type de communication, généralement initiée par le contrôleur, vise à obtenir des informations d'identification ou des fonctionnalités essentielles du commutateur. Ces données peuvent s'avérer indispensables lors de la configuration du canal OpenFlow. On y trouve, par exemple, les messages Modify-state, utilises pour l'ajout, la suppression ou la modification des

entrées de flux dans une table de flux, les messages Read-state, qui permettent la collecte d'informations telles que les paramètres de configuration du commutateur, ou encore les messages packet-out, qui assurent le transfert des paquets reçus via les messages packet-in. D'autres messages, comme ceux relatifs à la configuration du commutateur, les commandes envoyées par le contrôleur, les statistiques, la configuration des files d'attente et les barrières, sont également inclus ;

— **Les messages asynchrones** : ce type de message, transmis par le commutateur au contrôleur, signale l'arrivée de paquets ou notifie tout changement d'état du commutateur. Parmi les exemples notables, on distingue les messages packet-in, servant à transférer un paquet au contrôleur, les messages Flowremoved, qui signalent la suppression d'entrées de flux dans la table de flux, et les messages port-status, informant le contrôleur de toute modification affectant un port ;

— **Les messages symétriques** : ce type de message peut être initie indifféremment par le commutateur ou le contrôleur. Il comprend les messages Hello, utilises lors du démarrage de la connexion, les messages écho request/reply, servant à vérifier l'état de la connexion, et les messages error, informant l'entité homologue de problèmes éventuels [22, 23].

- **Le protocole Netconf**

Netconf, c'est en quelque sorte un chef d'orchestre réseau. Il s'appuie sur XML et les RPC pour bien configurer les équipements du réseau. Netconf utilise le protocole SSH pour le transport, ce qui simplifie la gestion des données et assure une bonne entente entre les équipements de divers fournisseurs, en utilisant XML pour structurer les messages et les RPC pour ajuster la configuration. Cela aide à limiter les pannes dues aux erreurs de configuration manuelle, à booster l'efficacité des mises à jour logicielles via des outils de configuration, et à renforcer la sécurité des données avec des mécanismes d'authentification et d'autorisation [22].

L'architecture de Netconf repose sur trois piliers :

— **Le gestionnaire Netconf**, un peu comme un client, supervise les équipements. Il peut envoyer des requêtes RPC à l'agent Netconf d'un équipement pour obtenir des infos ou

modifier sa configuration. Il peut aussi suivre l'état d'un équipement et les alertes envoyées par l'agent Netconf ;

— L'agent Netconf, agissant comme un serveur, gère la configuration d'un équipement, répond aux requêtes RPC et informe le gestionnaire Netconf. Dès qu'il reçoit une requête RPC, il l'examine, la traite en utilisant un cadre de gestion de configuration, et renvoie une réponse "RPC-reply" au gestionnaire Netconf. Il lui signale aussi les alarmes ou événements, pour que ce dernier soit au courant de l'état de gestion des équipements ;

— Le fichier Schéma, c'est un peu le manuel des données, avec toutes les règles pour rédiger un document XML. Il précise tous les éléments de gestion sur l'appareil, les contraintes, les liens entre ces éléments, et qui a le droit de lire ou d'écrire [22].

Les infos que l'on peut récupérer d'un équipement en marche, ce sont les données de configuration, c'est-à-dire toutes les données nécessaires pour passer de l'état initial à l'état actuel de l'équipement, et les données d'état, qui sont des infos supplémentaires hors configuration.

Il existe d'autres protocoles en plus de ceux que nous avons mentionnés. On peut citer par exemple ForCes, Opflex, Netstream, SNMP, etc [23].

2.4 La virtualisation des fonctions réseau (NFV)

L’essor fulgurant des appareils connectés à Internet, constaté ces dernières années, a engendré une demande accrue de la part des usagers. Afin de satisfaire ces exigences toujours plus grandes et variées, l’intégration de nouveaux dispositifs au réseau s’avère indispensable pour enrichir l’offre de services. Néanmoins, cette augmentation continue du nombre d’appareils entraîne des coûts importants, une complexité accrue et une maintenance ardue du réseau. Au vu de ces multiples difficultés rencontrées par les opérateurs, la virtualisation des fonctions réseau apparaît comme une alternative.

2.4.1 Définition de NFV

La virtualisation des fonctions réseau s'appuie sur une idée simple : dissocier les fonctions réseau de l'infrastructure physique. Les fonctions réseau virtualisées sont alors mises en œuvre par le biais de machines virtuelles. Tout cela a été rendu possible grâce aux avancées des technologies de virtualisation. La virtualisation des fonctions réseau introduit une approche novatrice de l'architecture réseau, qui consiste à séparer les

fonctions réseau, comme le NAT, le pare-feu, la détection d'intrusion ou le DNS, de leurs équipements dédiés pour qu'elles tournent directement dans des logiciels [22]. L'objectif de la virtualisation des fonctions réseau est d'exploiter les technologies issues de la virtualisation informatique pour virtualiser des ensembles complets de fonctions de nœuds réseau en éléments de base pouvant être assemblés ou mis bout à bout afin de concevoir des services de communication. Cela permettra de réaliser des économies, tant en termes d'espace que de consommation d'électricité, tout en contribuant à diminuer les dépenses d'investissement (CAPEX) et les coûts opérationnels (OPEX) [25].

2.4.2 Architecture NFV

L'architecture NFV s'articule autour de trois axes majeurs : les fonctions réseau virtualisées (VNF), l'orchestration et la gestion (MANO) et l'infrastructure NFV (NFVI).

La MANO se compose d'un orchestrateur NFV, d'un gestionnaire VNF et d'un gestionnaire d'infrastructure virtualisée, tandis que l'infrastructure NFV englobe les ressources virtualisées et matérielles ainsi que la couche de virtualisation, (voir figure 2.7) [26].

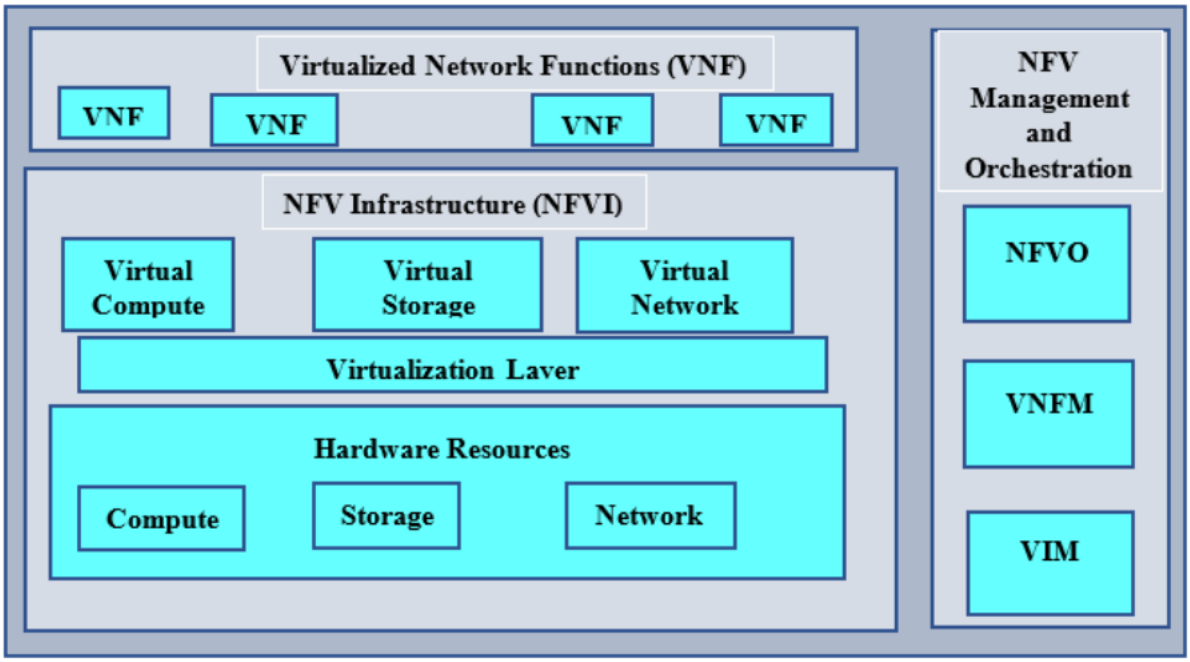


Figure 2.7 : Architecture de NFV[24]

2.4.2.1 Les fonction réseaux virtualisées (NFV)

La virtualisation des fonctions réseau, ou VNF, représente une approche où cette strate VNF est perçue comme le niveau applicatif. Ici, diverses fonctions réseau virtualisées résident, étant considérées comme des applications à part entière. Ces VNF peuvent être déployées sur plusieurs machines virtuelles, avec la possibilité pour chaque machine d'héberger une ou plusieurs VNF [26]. L'élément de gestion (EM) assume la responsabilité de gérer les aspects Fault, Configuration, Accounting, Performance et Security (FCAPS) d'une fonction réseau virtuelle (VNF) en cours d'exécution [25].

2.4.2.2 Infrastructures NFV (NFVI)

L'infrastructure NFVI englobe tout : le matériel, les logiciels, et cette couche de virtualisation cruciale. C'est elle qui façonne l'environnement où les fonctions réseau virtualisées prennent vie, sont administrées et fonctionnent [25]. L'infrastructure physique, elle, s'appuie sur des ressources bien concrètes : puissance de calcul, espace de stockage et connectivité réseau. En parallèle, l'infrastructure virtuelle est faite de ressources virtualisées comme le calcul, le stockage et la connectivité également. C'est là que la couche de virtualisation intervient, faisant office de pont. Elle sépare les ressources physiques grâce à un hyperviseur.

2.4.2.3 Gestion et Orchestration de NFV (MANO)

Le rôle premier de MANO consiste à piloter l'ensemble des fonctions, incluant par exemple un système de virtualisation, une organisation des ressources physiques, un suivi du cycle de vie des instances VNF, etc. [22].

La gestion des infrastructures virtualisées (VIM) assure le contrôle et le pilotage des NFVI, de même que la liaison entre un VNF et les ressources de calcul, de stockage et de réseau. L'orchestrateur NFV (NFVO), lui, assure l'organisation et le pilotage des infrastructures NFV et des ressources logicielles. Quant à NFVM, il prend en charge le suivi du cycle de vie des VNF. Un NFVM peut donc gérer un ou plusieurs VNF [25,26].

2.4.2.3 Les interfaces de l'architecture NFV

On note l'existence de plusieurs interfaces qui permettent aux différentes couches de l'architecture NFV de communiquer.

Parmi ces interfaces, on peut citer :

— Virtualizations Layer - Hardware Resources -**(Vl-Ha)** : elle se trouve entre la couche de virtualisation et les ressources materielles. Cette interface est responsable de la creation de l'environnement d'execution et de la collecte des informations sur l'etat des ressources materielles pour les VNFs ;

— VNF - NFV Infrastructure **(Vn-Nf)** : elle relie l'infrastructure NFV et les VNFs. Cette interface garantit que les machines virtuelles peuvent être déployées sur le NFVI pour répondre à la fiabilité et aux exigences d'évolutivité. Il représente l'environnement d'execution fourni par le NFVI aux VNFs afin de garantir, un cycle de vie indépendant du matériel, des performances et des exigences de portabilité du VNF ;

— NFV Orchestrator - VNF Manager **(Or-Vnfm)** : il est l'interface liant NFVO et VNFM. Il envoie les informations de configuration au VNFM. Il collecte également les informations sur l'etat des VNFs pour la gestion du cycle de vie des services réseau ;

— Virtualized Infrastructure Manager - VNF Manager **(Vi-Vnfm)** : elle se trouve entre le VIM et VNFM. Elle permet les demandes d'allocation de ressources par le VNFM et gère aussi la configuration des ressources materielles virtualisées et l'′échange d'informations d'etat ;

— NFV Orchestrator - Virtualized Infrastructure Manager **(Or-Vi)** : elle lie NFVO et VIM. Elle permet l'allocation et la réservation des ressources par le VNFM. Elle est responsable aussi de la configuration des ressources virtualisées et l'′échange d'informations d'′etat ;

— NFVI - Virtualised Infrastructure Manager **(Nf-Vi)** : cette interface lie le VIM et NFVI. Elle gère les machines virtuelles, le stockage et les systèmes réseau de NFVI, mais aussi la configuration des ressources materielles et l'échange d'informations d'etat ;

— OSS/BSS - NFV Management and Orchestration **(Os-Ma)** : cette interface relie OSS/BSS et MANO. Elle est responsable de la gestion des cycles de vie des services

réseau et des VNFs. Elle gère également l'échange d'analyse de données et l'échange de gestion de stratégies ;

— VNF/EM - VNF Manager **(Ve-Vnfm)** : Elle se trouve entre VNFM et VNF. Elle permet la gestion, du cycle de vie, de la configuration et des performances des VNFs. Elle est responsable aussi de la gestion des pannes, (voir figure 2.8)

[26,27]

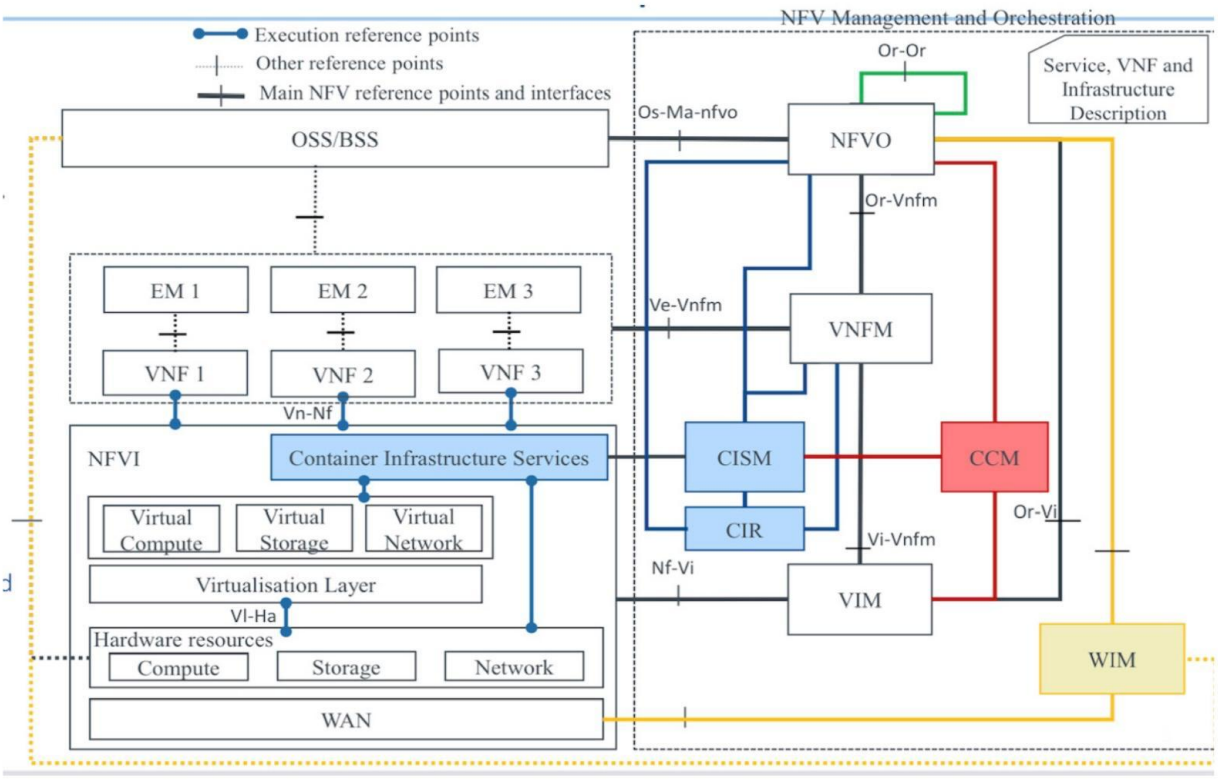


Figure 2.8 : Les interfaces de l'architecture NFV [24]

2.4.3 Réduction du CAPEX /OPEX

La virtualisation est l'un des moyens de diminuer des dépenses initiales et courantes, ce qui représente un coût important en investissement et en fonctionnement. Cette approche entraîne une baisse du matériel requis, de l'énergie consommée, des besoins en refroidissement et de l'espace utilisé. De plus, elle aide à optimiser, à restaurer et à sauvegarder les données, tout en renforçant la sécurité. [28]

2.4.4 Chaine de fonctions services réseaux (SFC)

Une suite ordonnée de fonctions de service constitue une chaîne de fonctions de service réseau. En d'autres termes, c'est une technologie permettant l'utilisation successive de divers services. La translation d'adresse réseau (NAT), l'inspection approfondie des paquets (DPI) et le pare-feu, traditionnellement déployés dans des composants matériels, en sont des exemples typiques. Or, grâce à la virtualisation des fonctions réseau, ces fonctions de service peuvent être virtualisées [29].

La figure 2.9 illustre une composition de multiples chaînes de fonctions de service, chaque bloc représentant un nœud de fonction de service. On remarque que chaque nœud peut appartenir à une ou plusieurs chaînes. De plus, un même nœud peut apparaître plusieurs fois au sein d'une chaîne. Il est également à noter que les chaînes de fonctions de service peuvent être soit unidirectionnelles, soit bidirectionnelles [29].

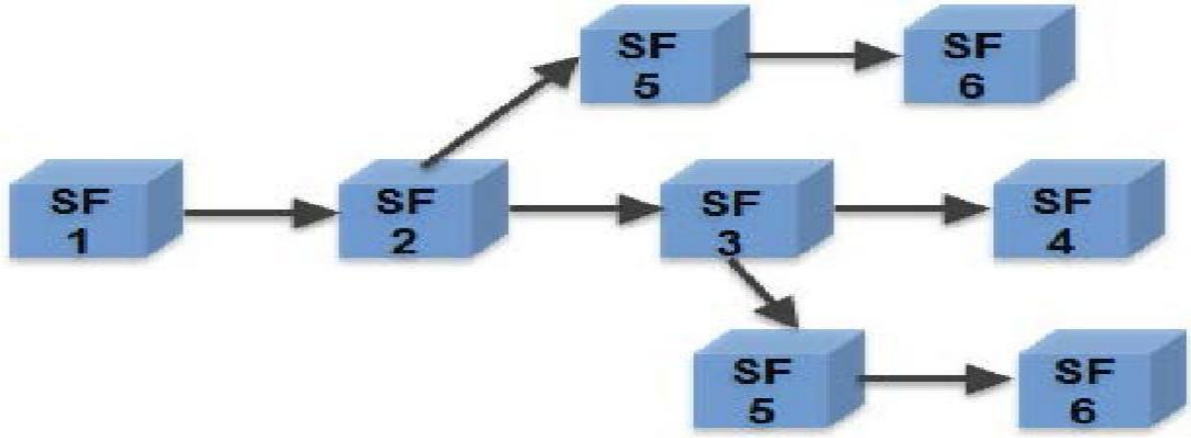


Figure 2.9 : Chaine de fonctions services [29]

Tel qu'illustré à la figure 2.10, l'architecture de SFC englobe un classificateur, des commutateurs, de proxy et des fonctions de service. Entre le réseau et le classificateur, un lien en pointillés indique un flux non encapsulé, tandis qu'entre le classificateur et le commutateur, un lien continu signale un flux encapsulé.

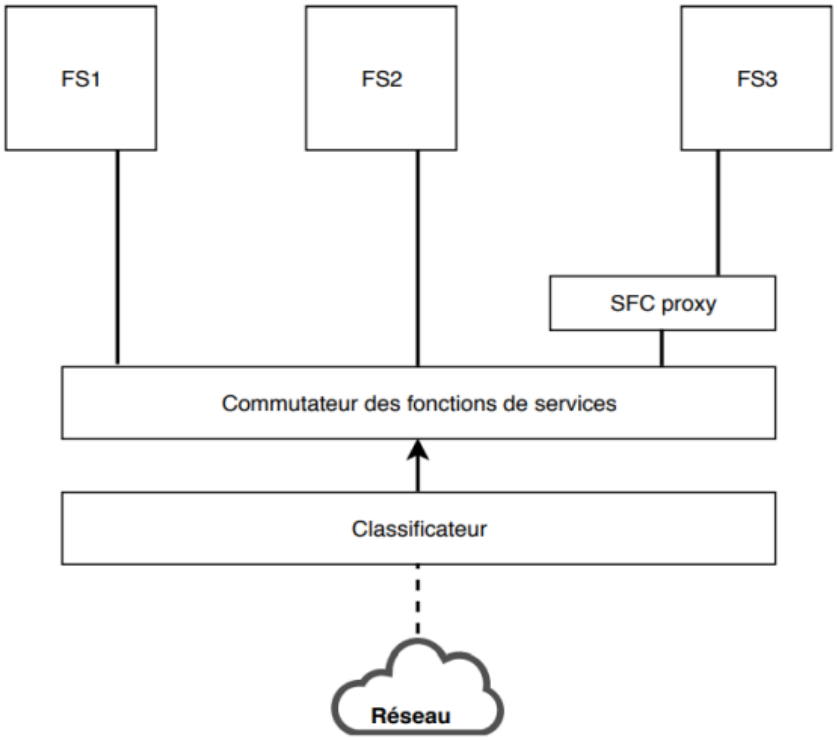


Figure 2.10 : Architecture de chainage des fonctions services [29]

— **Le classificateur** est responsable de l'ordre dans lequel les fonctions sont visitées, il gère également l'encapsulation des informations spécifiques à la chaine ;

— **Le commutateur** prend en charge l'acheminement des paquets encapsulé par le classificateur et les dirige vers la fonction de service initiale. Le flux est ensuite retourné au commutateur, qui se charge de l'envoyer à la fonction suivante, et ainsi de suite ;

— **Le proxy** sert à extraire les données afin que la fonction de service puisse comprendre et traiter les informations, comme illustré par FS3 dans la figure 2.9. Une fois le traitement terminé, les données sont réencapsulées. Cependant, d'autres fonctions de service peuvent interpréter directement les informations sans nécessiter l'intervention d'un proxy [29,30].

2.5 Les outils d’implémentation de SDN et NFV

On trouve divers contrôleurs SDN, tels que ODL, ONOS, Beacon, NOX, POX, RYU, etc. La performance d'un contrôleur s'évalue souvent en débit, mais aussi en latence. De plus, lors du choix d'un contrôleur, divers éléments entrent en jeu, comme le langage de programmation, la compatibilité avec Openflow, la capacité à programmer le réseau, les interfaces northbound et southbound, ainsi que l'efficacité, englobant performance, fiabilité, évolutivité et sécurité [31].

Un facteur essentiel lors du choix est le langage de programmation, car il influe sur les performances du contrôleur et la rapidité de développement. Python, Java et C++ sont les plus utilisés dans la programmation des contrôleurs SDN.

De plus, on peut dire que les contrôleurs codés en Java (ONOS, RYU, Beacon, Maestro, Floodlight) sont supérieurs, car ils fonctionnent sur diverses plateformes. Parmi eux, ONOS et ODL se distinguent par leur bonne modularité et une très bonne documentation [25].

La compatibilité avec diverses plateformes est un critère important. ONOS, ODL, POX, Beacon, Maestro et Floodlight se démarquent, car ils sont compatibles avec Linux, Mac OS et Windows. POX et RYU sont encore mieux lotis, car ils offrent également une productivité élevée [31].

Il faut aussi considérer la prise en charge d'Openflow. Ce dernier permet une manipulation directe des commutateurs Openflow. La connaissance des fonctionnalités Openflow prises en charge par les contrôleurs est cruciale, car les versions Openflow diffèrent par leurs caractéristiques. Dans ce cas, ONOS, ODL, RYU, Floodlight et MUL sont préférables, car ils supportent les versions récentes d'Openflow (1.3, 1.4) [25].

Le multithreading, qui permet de simuler plusieurs tâches simultanées, est un critère important pour le choix du contrôleur, car il améliore ses performances. Donc, ONOS, ODL, RYU, Beacon, Maestro, Floodlight et MUL sont à privilégier. De plus, en tenant compte de la prise en charge d'Openstack, ONOS, ODL, RYU et MUL sont supérieurs, car ils permettent de créer un environnement Cloud pour assurer le stockage, la sécurité et l'orchestration des NFV, garantissant ainsi une bonne gestion des ressources virtuelles via le module neutron, comme illustré dans le tableau 2.2 [31].

Tableau 2.2 : Table comparative des contrôleurs SDN [31]

	ONOS	ODL	NOX	POX	RYU	Beacon	Maestro	Flood-light
Langage de programmation	Java	Python, Java	C++	Python	Java	Java	Java	Java
GUI	Interface web	Interface web	Python + QT4	Python + QT4	Interface web	Interface web	Interface web	Interface web
Modularité	Elevé	Elevé	Faible	Faible	Moyen	Moyen	Moyen	Moyen
Distribué/ Centralisé	D	D	C	C	C	C	C	C
Plateforme	Linux MAC OS Win-dows	Linux MAC OS Win-dows	Supporté sur Li-nux	Linux MAC OS Win-dows	Supporté sur Li-nux	Linux MAC OS Win-dows	Linux MAC OS Win-dows	Linux MAC OS Win-dows
Productivité	Moyen	Moyen	Moyen	Elevé	Elevé	Moyen	Moyen	Moyen
API south-bound	OF1.0, 1.3, NETCO NF	OF1.0, 1.3, 1.4, OVSDB, SNMP,etc	OF 1.0	OF 1.0	OF 1.0, 1.2, 1.3, 1.4, NETCO NF, etc	OF 1.0	OF 1.0	OF 1.0, 1.3
API north-bound	REST API	REST API	REST API	REST API	REST API	REST API	REST API	REST API
Multithreading	Oui	Oui	NOX MT	Non	Oui	Oui	Oui	Oui
Openstack	Oui	Oui	Non	Non	Oui	Non	Non	Non

2.6 Synergie entre SDN, NFV et le Machine Learning dans les réseaux 5G

Les innovations en SDN (réseautage défini par logiciel) et NFV (virtualisation des fonctions réseau) ont profondément modifié l'architecture et la gestion des réseaux, en apportant une flexibilité, une programmabilité et une automatisation accrue. Ces avancées sont particulièrement cruciales dans le cadre des réseaux 5G, où la complexité et les exigences relatives aux performances et à la qualité de service (QoS) sont considérables.

Cependant, cette complexité croissante rend la gestion manuelle des ressources inefficace et difficile à adapter. C'est ici que le machine Learning (ML) se présente comme un atout stratégique pour les approches SDN/NFV. En tirant parti des données issues du contrôle centralisé (via SDN) et des fonctions virtualisées (via NFV), les modèles d'apprentissage automatique permettent :

D'automatiser l'allocation des ressources en temps réel,

De prévoir les congestions ou défaillances réseau,

D'optimiser les chemins de routage de manière dynamique,

Ou encore de suggérer des politiques de gestion fondées sur les comportements constatés.

L'intégration de ces trois aspects — SDN, NFV et ML — constitue la base d'une gestion intelligente et autonome des réseaux 5G. Cette synergie est au centre des recherches contemporaines visant à concevoir des réseaux plus efficaces, adaptables et autonomes, tout en anticipant les besoins futurs de la 6G. Cette connexion logique entre les technologies de réseautage programmables et l'intelligence artificielle nous conduit naturellement au chapitre suivant, qui traite de l'application du machine Learning pour optimiser l'allocation des ressources dans les réseaux 5G.

2.7 Conclusion

Ce chapitre a permis d'examiner les principes de la programmation des réseaux à travers les technologies SDN et NFV. Nous avons mis en évidence les transformations significatives qu'elles apportent concernant la gestion, l'orchestration et la virtualisation des services réseaux. Grâce à leur complémentarité, SDN et NFV constituent une base technologique essentielle pour la 5G, offrant une infrastructure plus flexible, réactive et automatisée.

La dernière section a présenté un point de vue important : l'incorporation du machine Learning comme un levier supplémentaire face aux défis de complexité et d'optimisation. C'est précisément cet aspect que nous allons explorer plus en détail dans le prochain chapitre.

CHAPITRE 3 : APPLICATION DU MACHINE LEARNING POUR L'ALLOCATION DE RESSOURCES DANS LES RESEAUX 5G

3.1 Introduction

Les réseaux 5G, en raison de leur complexité croissante et de leurs exigences rigoureuses en matière de qualité de service (QoS), requièrent des mécanismes d'allocation des ressources qui soient à la fois dynamiques, précis et intelligents. Comme mentionné précédemment, les architectures programmables reposant sur SDN et NFV facilitent une centralisation des données réseau, ce qui ouvre la voie à l'intégration de l'intelligence artificielle pour automatiser la gestion des ressources.

Dans ce chapitre nous allons donc voire d'abord la problématique de l'allocation des ressources dans les réseaux mobiles ensuite les généralités sur la machine Learning, concevoir aussi plusieurs modèles de machine Learning capables de prévoir ou d'optimiser l'allocation des ressources, grâce à un ensemble de jeu de donnée contenant des caractéristiques liées à l'allocation des ressources. Et enfin, faire une comparaison des résultats afin de déterminer le modèle le plus approprié pour cette tâche spécifique. [32]

3.2 Problématique de l'allocation des ressources dans les réseaux 5G

3.2.1 Définition de l'allocation des ressources

L'allocation des ressources dans les réseaux mobiles fait référence à l'ensemble des méthodes utilisées pour distribuer efficacement les ressources radioélectriques disponibles comme la bande passante, le temps, la puissance d'émission ou encore les blocs de ressources entre les divers utilisateurs, applications ou services. Son objectif est d'optimiser les performances globales du réseau tout en maintenant le respect des exigences de qualité de service (QoS). Dans un réseau 5G, cette fonction devient beaucoup plus complexe que dans les générations antérieures en raison de la diversité des cas d'utilisation (eMBB, URLLC, mMTC), de la variété des appareils connectés, de l'augmentation de la densité des cellules, ainsi que de l'adoption massive de technologies avancées telles que le MIMO massif et le slicing de réseau.

3.2.2 Méthodes traditionnelles et leurs limites

Les méthodes traditionnelles d'allocation des ressources dans les réseaux mobiles reposent principalement sur des algorithmes simples et équitables, mais qui présentent des limites en termes d'efficacité et de qualité de service. Parmi les plus connus, on peut citer :

Tableau 3.1 : Méthodes traditionnelles d'allocation de ressources réseau [31]

Méthode	Principe de fonctionnement	Avantages	Limites
Round Robin (RR)	Attribution cyclique et égale des ressources entre tous les utilisateurs	Équité en nombre de ressources, simplicité de mise en œuvre	Ignore la qualité du lien et les besoins spécifiques ; débit variable selon les conditions radio
Accès Aléatoire	Distribution aléatoire des ressources avec probabilité uniforme	Simple à implémenter, distribution homogène à long terme	Absence de différenciation de QoS ; inefficace dans les environnements exigeants
Max-Min Fair (MMF)	Priorité aux utilisateurs à faibles besoins, allocation itérative jusqu'à saturation	Garantie de service minimum pour les utilisateurs faibles	Ne satisfait pas les utilisateurs à haut débit ; approche sous-optimale en cas de forte demande
MaxSNR	Priorité à l'utilisateur ayant le meilleur rapport signal/bruit	Optimise le débit global ; efficace en bonnes conditions	Pénalise les utilisateurs éloignés ; absence totale d'équité ou de QoS
Proportional Fair (PF)	Compromis entre efficacité (débit) et équité, selon le débit moyen de chaque utilisateur	Équilibre acceptable entre performance et équité	Ne prend pas en compte les besoins QoS spécifiques ; complexité de paramétrage

Avec l'explosion du trafic et des exigences de performance dans les réseaux 5G, ces approches montrent leurs limites : elles manquent de souplesse, n'anticipent pas les variations de charge, et ne peuvent pas exploiter efficacement les données générées par le réseau.

3.2.3 Exemple de paramètres influençant allocation des ressources

Pour prendre des décisions intelligentes, les algorithmes d'allocation doivent considérer plusieurs variables d'entrée. En voici quelques-unes présentées dans le tableau ci-dessous :

Tableau 3.2 : Quelques paramètres de ressources [31]

Paramètre	Description	Impact sur l'allocation
RSRP	Puissance du signal reçu	Permet de prioriser les utilisateurs bien couverts
SINR	Rapport signal sur interférence + bruit	Indique la qualité du lien radio
Bande passante requise	Dépend du type de service (vidéo, IoT...)	Conditionne le nombre de ressources à allouer
Latence tolérée	Temps maximal admissible pour la transmission	Influence le choix des ressources à faible délai
Type de service	eMBB, URLLC, mMTC	Détermine les contraintes spécifiques
Charge du réseau	Nombre total d'utilisateurs connectés dans la cellule	Peut déclencher des stratégies de réallocation dynamique

Ces données, souvent accessibles via les équipements de réseau, servent de base aux algorithmes d'apprentissage.

3.3 Généralités sur le machine Learning

3.3.1 Définition

L'apprentissage automatique, également connu sous le nom de machine Learning, est une discipline de l'intelligence artificielle qui vise à créer des systèmes capables d'apprendre à partir d'informations, sans avoir à être spécifiquement programmés pour chaque tâche. Au lieu de suivre des directives strictes, un algorithme d'apprentissage automatique repère des schémas, élabore des modèles et prend des décisions sur la base de l'expérience accumulée grâce aux données.

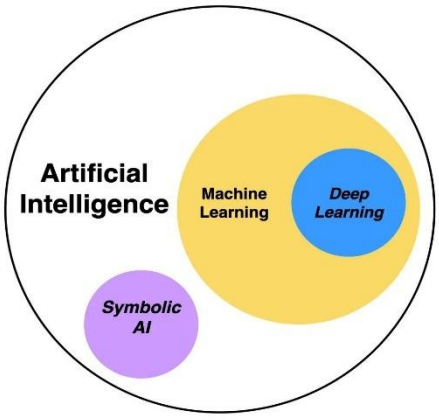


Figure 3.1 : IA et Machine Learning [32]

Le concept fondamental de l'apprentissage automatique repose sur la notion qu'un système peut perfectionner ses performances au fil du temps en analysant de nouvelles informations. Cela le rend particulièrement adapté à des environnements complexes et en

constante évolution, tels que les réseaux 5G, où les conditions réseau (trafic, latence, bande passante) évoluent continuellement et nécessitent des ajustements rapides. Le ML est donc considéré comme un outil central pour l'automatisation intelligente des réseaux [32].

3.3.2 Types d'apprentissage automatique

Le machine learning peut être classé en plusieurs grandes catégories selon la manière dont le modèle apprend :

Apprentissage supervisé

Apprentissage non supervisé

Apprentissage par renforcement

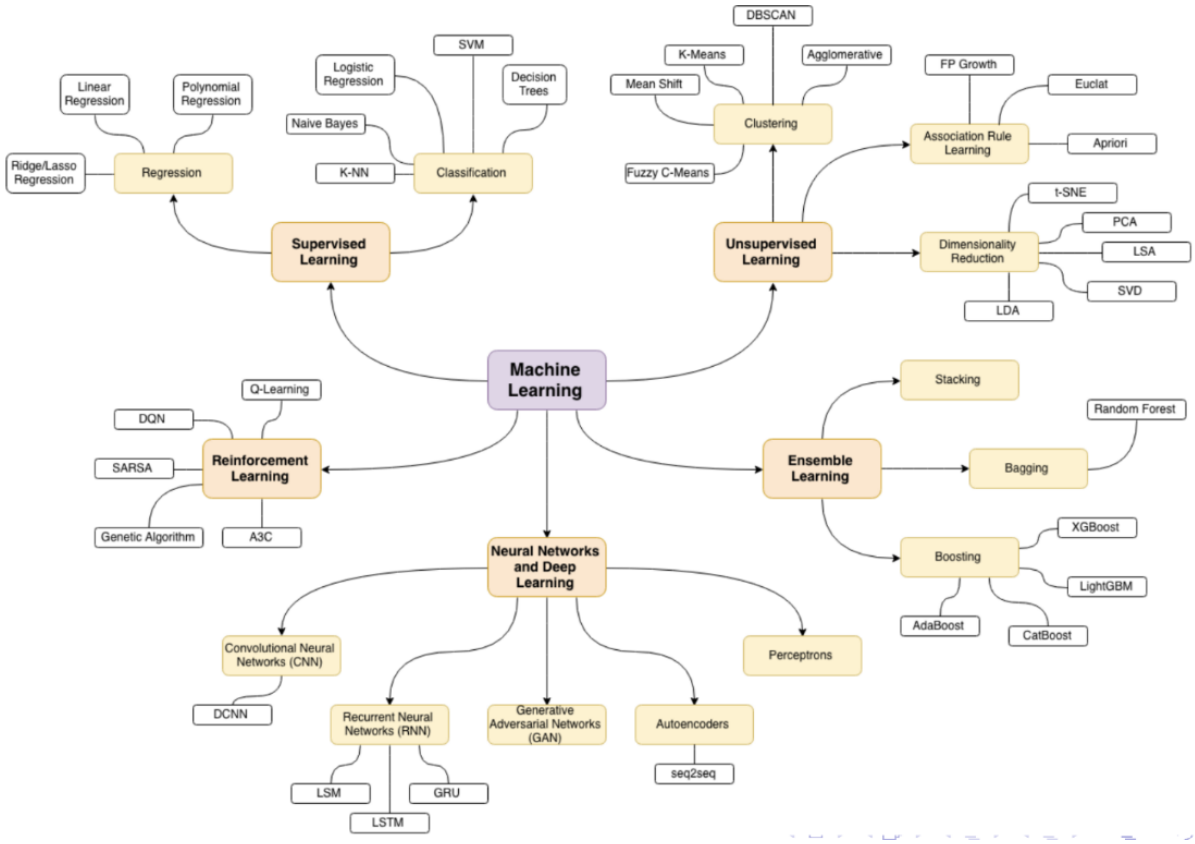


Figure 3.2 : Type d'apprentissage automatique [33]

Apprentissage supervisé : le modèle apprend à partir de données étiquetées. Par exemple, il apprend à prédire une quantité (régression) ou une classe (classification) en se basant sur des observations passées. C'est le type d'apprentissage utilisé dans ce mémoire.

Apprentissage non supervisé : les données ne sont pas étiquetées. Le modèle cherche à découvrir des structures ou des régularités cachées (ex. : clustering, réduction de dimension).

Apprentissage par renforcement : le modèle apprend à prendre des décisions séquentielles en interagissant avec un environnement, recevant des récompenses selon ses actions.

3.3.3 Étapes d'un projet machine learning

La mise en œuvre d'un projet de machine learning suit généralement un processus en plusieurs étapes, dont :

— **Collecte et nettoyage des données :** Il s'agit ici d'obtenir un ensemble de données fiables, représentatives du problème étudié, gérer le nettoyage des données, gestion des valeurs manquantes, normalisation, encodage des variables catégorielles, etc. Il faut disposer d'un grand nombre de données pour une construction plus précise du modèle et la prise de bonne décision par ce dernier. Apres la collecte d'une quantité importante de données, il faut les traiter par transformation des données brutes a des données utiles afin de rendre la modélisation simple, rapide et aussi optimiser le temps d'exécution et d'apprentissage. Ainsi l'extraction de caractéristiques dans ces données utiles entrainera une précision du modèle ;

— **Séparation des données :** division du jeu de données en ensemble d'entraînement (training set) et ensemble de test (test set), parfois avec une validation croisée (cross-validation).

— **Choix et entraînement du modèle :** sélection de l'algorithme approprié et entraînement sur les données d'apprentissage. Le choix de l'algorithme dépend également d'autre facteurs comme par exemple la quantité de données, le nombre de caractéristiques ou la sortie qui peut donner une valeur continue ou une valeur discrète appelée respectivement la régression ou la classification ;

— **Phase de validation** : Une fois les différents modèles de machine learning entraînés, il est essentiel d'évaluer leur capacité à généraliser à de nouvelles données, c'est-à-dire à fournir des prédictions fiables sur des cas qu'ils n'ont jamais vus auparavant. Cette

évaluation se fait à travers une phase de validation rigoureuse, qui constitue une étape fondamentale dans tout processus d'apprentissage automatique.

— **Évaluation du modèle** : utilisation de métriques comme la Root Mean Square Error (RMSE), le Mean Absolute Error (MAE) ou le coefficient de détermination (R^2) pour mesurer la performance.

— **Optimisation** : réglage des hyperparamètres et amélioration du modèle via des techniques comme la validation croisée, grid search, etc. [34]

3.4 Étude pratique : Modèle de Machine Learning pour l'allocation

Dans ce mémoire de recherche, nous tirons parti de l'apprentissage automatique pour optimiser l'allocation des ressources, comme la bande passante, afin de garantir une qualité de service (QoS) élevée et l'efficacité du réseau face aux diverses demandes des utilisateurs. Cette recherche vise à comparer et à évaluer les performances de divers modèles d'apprentissage automatique, dans l'allocation dynamique de la bande passante au sein des réseaux 5G, en identifiant le modèle le plus approprié pour des scénarios spécifiques sur la base de métriques clés.

3.4.1 Equations

Des expressions mathématiques sont utilisées pour modéliser le processus d'allocation des ressources dans la 5G. L'approche adoptée repose sur une modélisation en cinq étapes, chacune décrite par une équation clé :

Fonction Générale d'allocation

$$\mathbf{F(allocation) = \sum_{i=1}^{NU} U_i\ (Resource\ Allocation_i)} \qquad \mathbf{(1)}$$

Où :

- Nu est le nombre total d'utilisateurs ou d'unités à desservir,
- U_i est la fonction d'utilité de l'utilisateur i (bénéfice obtenu par l'allocation).

Modèle de régression pour la prédiction

$$\mathbf{Allocation = \beta_0 + \beta_1 \cdot x_1 + \beta_2 \cdot x_2 + ... + \beta_n \cdot x_n} \qquad \mathbf{(2)}$$

Où

- $\beta 0$ est l'ordonnée à l'origine (biase)

- β_i sont les poids ou coefficients associés aux variables explicatives x_i(paramètre) (ex : RSRP, SINR, latence, bande passante requise, etc.).

Fonction de perte

$$\min_{\theta} L(X, y ; \theta) \quad (3)$$

où :

- X représente les données d'entrée,

- y les valeurs observées (ressources réellement allouées),

- θ les paramètres du modèle (les β dans le cas d'une régression).

Mise a jour des poids par descente de gradient

$$\theta^{(t+1)} = \theta^{(t)} - \alpha \cdot \nabla_{\theta} L(X, y ; \theta^{(t)}) \quad (4)$$

Où

- α est le taux d'apprentissage (learning rate),

- $\nabla_{\theta} L$ est le gradient de la fonction de perte par rapport aux paramètres.

Prédiction finale du modèle

$$\hat{y} = f(X ; \theta) \quad (5)$$

Où

- $\hat{y}$ est la valeur prédite (quantité de ressource a prédite)

- f est la fonction d'apprentissage défini par le modèle appris

Enfin, nous évaluons les performances des modèles formé sur un ensemble de données de validation ou de test distinct afin d'évaluer leurs capacités de généralisation et de précision prédictive. Nous mesurons des paramètres tels que l'erreur quadratique moyenne et le coefficient de détermination pour quantifier les performances du modèle.[34]

3.4.2 Description du jeu de données

Il faut noter que l'outil ou l'environnement de développement utilisé est **Jupyter Notebook**, tandis que le langage de programmation adopté est **python**. Le jeu de données utilisé dans cette étude a été télécharger sur le net et est fourni par Nokia et concerne un sous-ensemble de données réelles collectées dans le cadre de l'optimisation des réseaux mobiles 5G. Il contient environ 400 lignes d'observations, liées à l'allocation des ressources du réseau 5G, telles que la force du signal, la bande passante requise, la latence et les ressources allouées. Ce jeu de données est utilisé comme base d'apprentissage supervisé pour prédire la quantité optimale de ressources à allouer selon les caractéristiques d'usage et les conditions réseau observées. On a effectué aussi une augmentation des données (800 lignes) en appliquant des techniques d'augmentations adaptées aux caractéristiques spécifiques [35]

Dans le tableau 3.3 suivant nous avons les différentes caractéristiques et leur description.

Tableau 3.3 : Description des caractéristiques

Nom de la variable	Type	Description
`Timestamp`	Temporelle	Date et heure de l'observation. Permet de suivre l'évolution temporelle du trafic.
`User_ID`	Identifiant	Identifiant unique de l'utilisateur. Utilisé pour distinguer les sessions.
`Application_Type`	Catégorique	Type de service utilisé : streaming, emergency service, appel, IoT, gaming etc.
`Signal_Strenghth`	Numérique (dBm)	Puissance du signal reçu par l'utilisateur. Indique la qualité du lien radio.
`Latency`	Numérique (ms)	Temps de réponse requis par l'application. Mesure la tolérance au délai.
`Required_Bandwidth`	Numérique (MHz)	Quantité de bande passante demandée par l'utilisateur selon l'application.
`Allocated_Bandwidth`	Numérique (MHz)	Bande passante effectivement allouée par le réseau. Sert à mesurer la performance.
`Resource_Allocation`	Numérique	Quantité totale de ressources réseau (ex : blocs de ressources, PRB) allouées.

La figure 3.3 montre les données contenues dans les 5 premières lignes ainsi que leurs caractéristiques contenues dans le dataset.

	Timestamp	User_ID	Application_Type	Signal_Strength	Latency	Required_Bandwidth	Allocated_Bandwidth	Resource_
0	9/3/2023 10:00	User_1	Video_Call	-75 dBm	30 ms	10 Mbps	15 Mbps	
1	9/3/2023 10:00	User_2	Voice_Call	-80 dBm	20 ms	100 Kbps	120 Kbps	
2	9/3/2023 10:00	User_3	Streaming	-85 dBm	40 ms	5 Mbps	6 Mbps	
3	9/3/2023 10:00	User_4	Emergency_Service	-70 dBm	10 ms	1 Mbps	1.5 Mbps	
4	9/3/2023 10:00	User_5	Online_Gaming	-78 dBm	25 ms	2 Mbps	3 Mbps	

Figure 3.3 : Donnée extrait du dataset

3.4.3 Prétraitement et visualisations des données

Dans cette étape qui est essentielle nous avons préparé les données pour garantir leur qualité, leur cohérence et leur compatibilité avec les algorithmes utilisés. Cette phase inclut le diagnostic des anomalies, la gestion des types de données, le nettoyage des unités, ainsi que l'analyse exploratoire des variables.

Dans un premier temps, une analyse exploratoire a été réalisée pour identifier :

Les valeurs manquantes **(df.isnull().sum()),**

Les doublons **(df.duplicated().sum()),**

Les types de données **(df.info()),**

Et la distribution des valeurs **(df.describe()).**

Ensuite, plusieurs corrections ont été apportées :

Les colonnes contenant des unités (comme **Signal_Strength**, **Latency, Required_Bandwidth)** ont été nettoyées à l'aide d'expressions régulières pour extraire les valeurs numériques.

La variable Timestamp a été convertie en format temporel avec **pd.to_datetime()** pour permettre une analyse chronologique.

Les colonnes ont été classées comme numériques ou catégoriques afin d'adapter les traitements futurs (normalisation, encodage, etc.).

Enfin, une visualisation simple de la variable **Application_Type** a permis de comprendre la répartition des types de services présents dans le dataset (streaming, IoT, etc.), éléments importants pour les besoins en bande passante. Cette phase de préparation garantit que les données sont propres, cohérentes et prêtes pour l'entraînement des modèles prédictifs.

```python
# Supprimer les unités et convertir les colonnes en types numériques
df['Signal_Strength'] = df['Signal_Strength'].str.replace(r'[^\d.-]', '', regex=True).astype(fl
df['Latency'] = df['Latency'].str.replace(r'[^\d.-]', '', regex=True).astype(float)
df['Required_Bandwidth'] = df['Required_Bandwidth'].str.replace(r'[^\d.-]', '', regex=True).ast
df['Allocated_Bandwidth'] = df['Allocated_Bandwidth'].str.replace(r'[^\d.-]', '', regex=True).a
df['Resource_Allocation'] = df['Resource_Allocation'].str.replace(r'[^\d.-]', '', regex=True).a

# Obtenir une liste des colonnes avec des types de données objet (catégorielles)
object_columns = df.select_dtypes(include='object').columns.tolist()

# Obtenir une liste des colonnes avec des types de données numériques (entier ou flottant).
numerical_columns = df.select_dtypes(include=['int', 'float']).columns.tolist()
# afficher les listes des colonnes
print("Object columns:", object_columns)
print("Numerical columns:", numerical_columns)

```

```
Object columns: ['User_ID', 'Application_Type']
Numerical columns: ['Signal_Strength', 'Latency', 'Required_Bandwidth', 'Allocated_Bandwidth', 'Re
source_Allocation']
```

Figure 3.4: Traitement et conversion des formats

```
1 df['Application_Type'].value_counts()
```

```
Application_Type
Video_Call              58
Web_Browsing            48
Streaming               47
Emergency_Service       47
Background_Download     47
Video_Streaming         47
VoIP_Call               46
Online_Gaming           45
IoT_Temperature         13
Voice_Call               1
File_Download            1
Name: count, dtype: int64
```

Figure 3.5: Nombre d'occurrences par type d'application

On observe que les types d'applications les plus représentés sont les appels vidéo, la navigation web, les services d'urgence et le streaming, avec environ 45 à 58 occurrences chacun. Les services liés à l'Internet des objets (IoT) et aux appels vocaux classiques sont beaucoup moins fréquents, ce qui reflète leur usage limité dans cet échantillon.

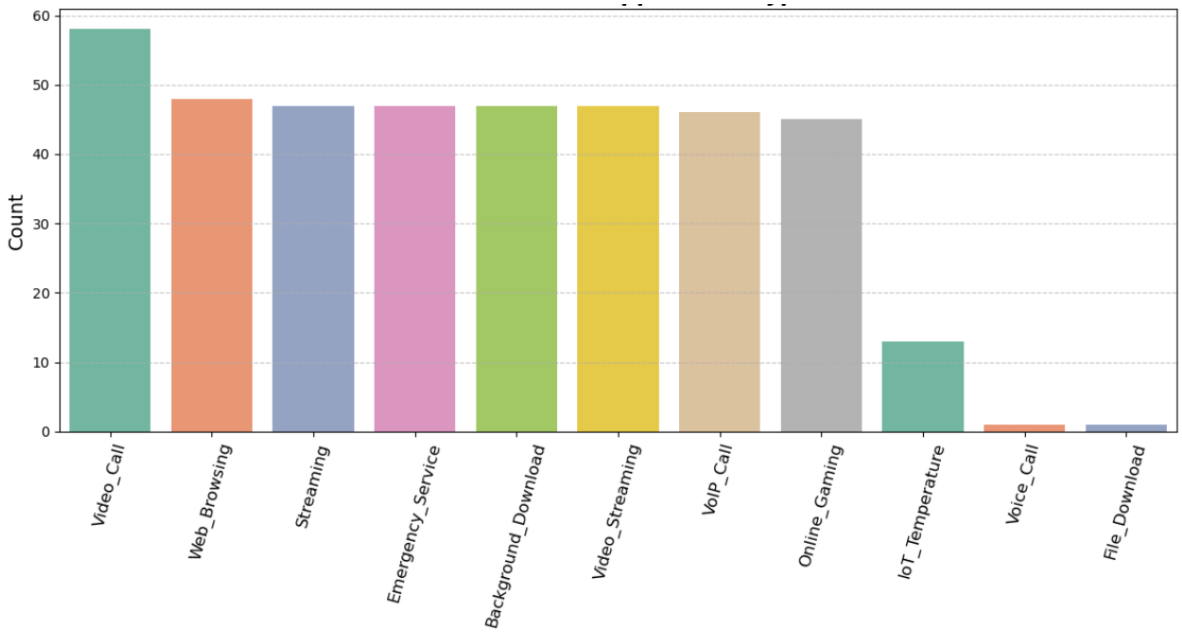
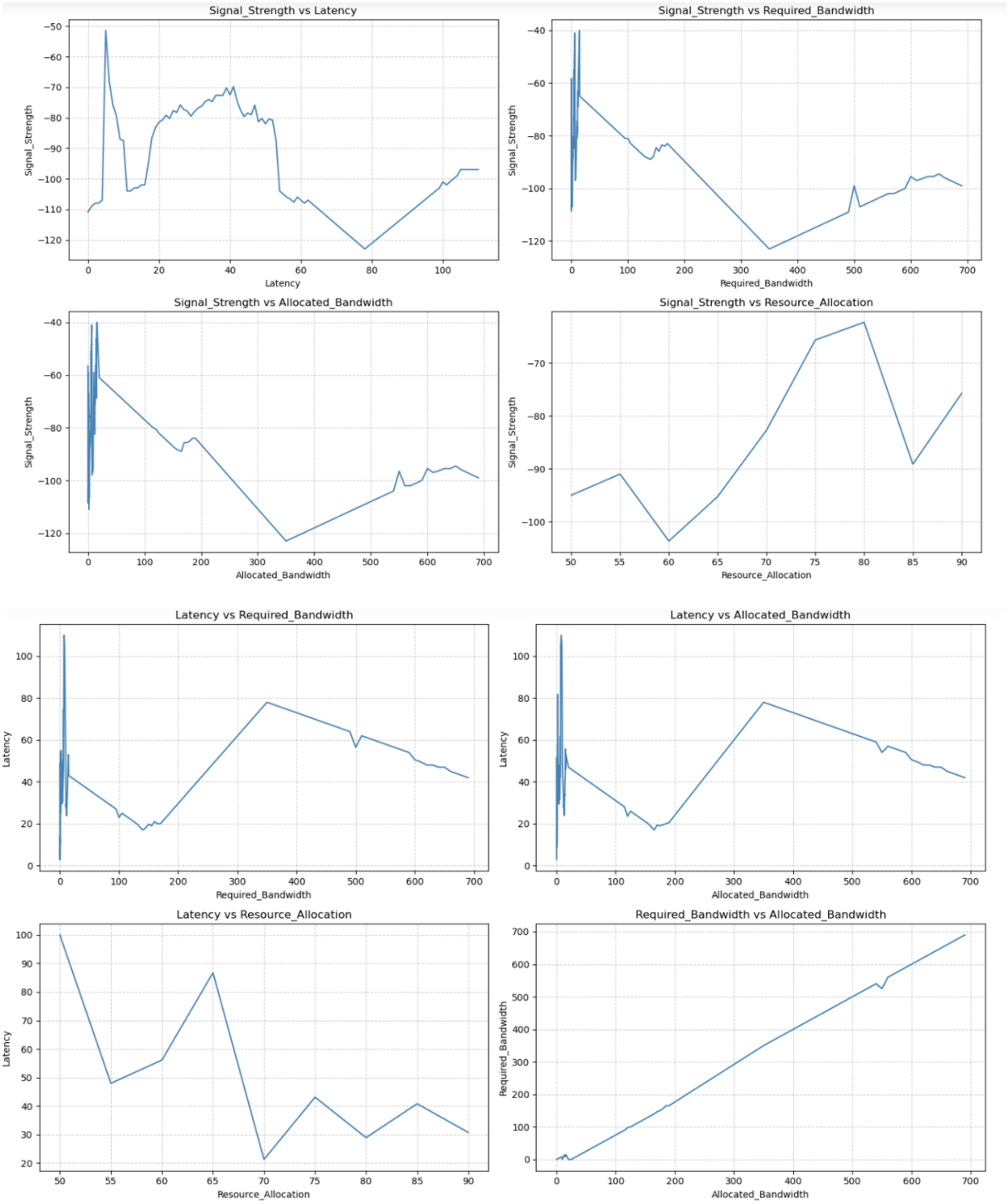


Figure 3.6: Répartition des types d'application

Cette distribution qui montre encore la prédominance des usages à forte demande de bande passante comme les appels vidéo, le streaming. Cela indique que les besoins en ressources varient fortement selon le type de service, ce qui justifie l'utilisation de modèles intelligents d'allocation.



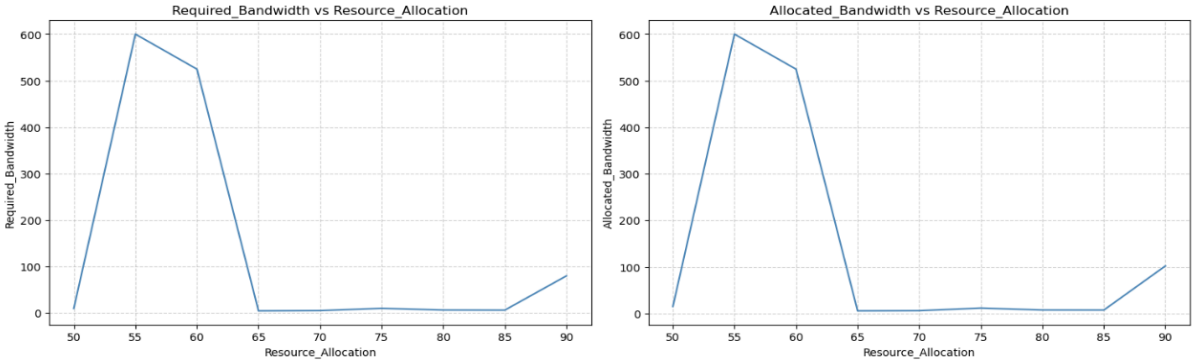


Figure 3.7: Visualisation croisée entre variables

Ces courbes de tendance entre les principales variables numériques du réseau. Plusieurs observations clés s'en dégagent :

Signal fort = meilleure allocation : Les utilisateurs ayant une meilleure puissance de signal reçoivent généralement plus de bande passante et de ressources, avec une latence plus faible.

Latence et compensation : Une hausse de la latence est souvent accompagnée d'une augmentation modérée des ressources allouées, suggérant un ajustement pour maintenir la qualité de service.

Besoins vs allocations : Les variables Required_Bandwidth et Allocated_Bandwidth montrent une relation croissante, ce qui indique que les demandes des utilisateurs sont globalement prises en compte.

Corrélation forte entre Allocated_Bandwidth et Resource_Allocation, confirmant la cohérence du système d'allocation.

```python
import matplotlib.pyplot as plt
import seaborn as sns

# Calculate the correlation matrix for numerical columns in your DataFrame
correlation_matrix = df[numerical_columns].corr()

# Plotting the heatmap
plt.figure(figsize=(10, 6))
sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm', linewidths=0.5)

# Adding title
plt.title('Correlation Heatmap')

# Show the plot
plt.show()

```

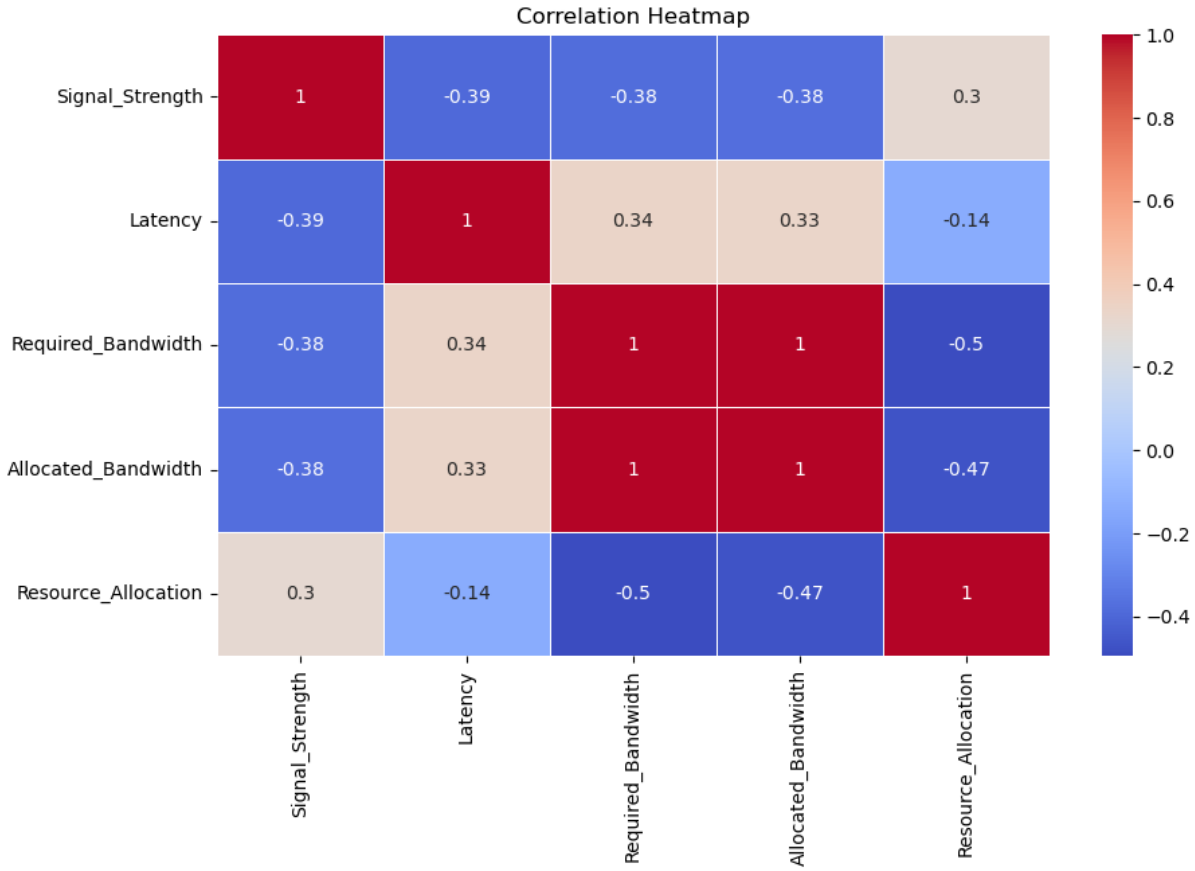


Figure 3.8: Matrice de corrélation

Cette matrice qui a des valeurs comprissent entre - 1 et 1 permet d'identifier les relations linéaires ou les dépendances entre les variables réseau (signal, latence, bande passante, ressources). Elle aide à sélectionner les variables les plus pertinentes pour la prédiction de l'allocation, en évitant les redondances ou dépendances forte.

Signal_Strength : Corrélée négativement avec Latency ,allocated_bandwith (= -0,38) ;cela signifie que plus le signal est fort, moins la latence est élevée, et moins les utilisateurs demandent ou reçoivent de bande pansante.Cela peut refléter une meilleure efficacité pour les utilisateurs bien couverts.

Corrélation modérée positive avec Resource_Allocation (0.30) : Les utilisateurs avec un bon signal reçoivent plus de ressources physiques, probablement car ils peuvent mieux les exploiter.

Latency : Corrélation faible à modérée avec les variables de bande passante : Une latence plus élevée est associée à une demande accrue de bande passante, ce qui peut traduire un effort pour compenser une mauvaise qualité de service.

Corrélation négative avec Resource_Allocation (-0.14) : Les utilisateurs avec plus de latence reçoivent légèrement moins de ressources, ce qui peut indiquer une congestion réseau ou une priorisation non optimisée.

Required_Bandwidth et Allocated_Bandwidth : Corrélation parfaite entre elles (1.00) : Cela suggère que l'allocation suit strictement la demande, sans ajustement dynamique. C'est à noter, car un réseau adaptatif chercherait à gérer l'allocation selon d'autres facteurs (signal, congestion...).

Corrélation négative avec Resource_Allocation : -0.50 pour la bande requise -0.47 pour la bande allouée. Cela peut sembler contre-intuitif : plus on alloue de bande passante, moins on attribue de "ressources" au sens physique. Cela s'explique probablement par une efficacité spectrale : les utilisateurs avec bon signal peuvent transmettre plus avec moins de blocs physiques.

Resource_Allocation : Corrélée positivement avec Signal_Strength (0.30)Corrélée négativement avec toutes les autres, surtout la bande passante demandée/allouée Cela montre que l'attribution physique de ressources n'est pas proportionnelle à la quantité de bande passante, mais influencée par la qualité du lien.

L'analyse de la matrice de corrélation révèle des relations clés entre la qualité du signal, la latence et l'allocation de ressources. Notamment, les utilisateurs ayant un bon signal reçoivent plus de ressources avec moins de latence, tandis que ceux avec une latence élevée ou une bande passante importante n'obtiennent pas nécessairement plus de blocs physiques.

3.5 Construction et entrainement des modèles

Dans cette section nous allons construire des modèles de machine learning capables de prédire l'allocation des ressources dans un réseau 5G en fonction de plusieurs paramètres (force du signal, latence, bande passante, etc.). Pour cela, nous avons sélectionné plusieurs algorithmes supervisés de régression, entraînés et évalués à partir du jeu de données préalablement nettoyé.

3.5.1 Sélection des modèles

Trois modèles de régression ont été retenus pour cette étude :

Régression linéaire : un modèle simple servant de baseline pour évaluer les performances minimales. **Régression polynomiale :** extension de la régression linéaire permettant de modéliser des relations non linéaires entre les variables. Et la **Forêt aléatoire (Random Forest)** : ensemble d'arbres de décision, robuste aux sur-apprentissages et généralement performant sur des données tabulaires. Ces modèles ont été choisis pour leur complémentarité, leur simplicité de mise en œuvre, et leur capacité à gérer efficacement des données numériques et des relations non linéaires.

Avant l'entraînement, les variables ont été séparées en deux ensembles : **Variables explicatives (X)** *signal_Strength,Required_Bandwidth,* **Variable cible (y) :***Resource_Allocation*

Notre ensemble de donnée sera divise en deux parties : données d'entrainement avec 80% des échantillons et données de test avec 20%.

3.5.2 Entrainement des modèles

Chaque modèle a été entraîné sur le jeu d'entraînement X_train, y_train en utilisant la bibliothèque Scikit-learn :

```
from sklearn.model_selection import train_test_split
from sklearn.linear_model import LinearRegression
from sklearn.metrics import mean_squared_error, r2_score
```

```
X = df[['Signal_Strength', 'Required_Bandwidth']]
y = df['Resource_Allocation']
```

```
#  separation en donnee d'entraînement et test
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

```

```
model = LinearRegression()
model.fit(X_train, y_train)
```

▾ LinearRegression

LinearRegression()

```
y_pred = model.predict(X_test)
```

```
mse = mean_squared_error(y_test, y_pred)
r2 = r2_score(y_test, y_pred)
print(f'Mean Squared Error: {mse}')
print(f'R-squared: {r2}')
```

```
# Nuage de points pour les valeurs réelles par rapport aux valeurs prédites
plt.figure(figsize=(10, 6))
plt.scatter(X_test['Signal_Strength'], y_test, color='blue', label='Actual')
plt.scatter(X_test['Signal_Strength'], y_pred, color='red', label='Predicted')

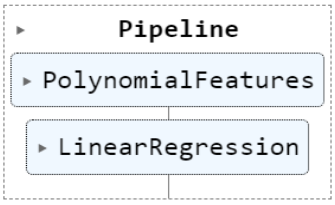
# Labeling the axes and adding the title
plt.xlabel('Signal Strength (dBm)')
plt.ylabel('Resource Allocation')
plt.legend()
plt.title('Linear Regression: Resource Allocation vs. Signal Strength and Predicted Values')

# Show the plot
plt.show()
```

Figure 3.9: Programme du modèle régression linéaire

```
from sklearn.preprocessing import PolynomialFeatures
from sklearn.pipeline import make_pipeline
```

```
degree = 5
polyreg = make_pipeline(PolynomialFeatures(degree), LinearRegression())
polyreg.fit(X_train, y_train)
```




```
y_pred_poly = polyreg.predict(X_test)
```

```
mse_poly = mean_squared_error(y_test, y_pred_poly)
r2_poly = r2_score(y_test, y_pred_poly)
print(f'Polynomial Regression - Mean Squared Error: {mse_poly}')
print(f'Polynomial Regression - R-squared: {r2_poly}')
```

```
Polynomial Regression - Mean Squared Error: 47.632824120609875
Polynomial Regression - R-squared: 0.4563772023148265
```

```
# Scatter plot for actual vs predicted values
plt.figure(figsize=(10, 6))
plt.scatter(X_test['Signal_Strength'], y_test, color='blue', label='Actual')
plt.scatter(X_test['Signal_Strength'], y_pred_poly, color='red', label='Predicted')

# Labeling the axes and adding the title
plt.xlabel('Signal Strength (dBm)')
plt.ylabel('Resource Allocation')
plt.legend()
plt.title('Polynomial Regression: Resource Allocation vs. Signal Strength and Predicted Values

# Show the plot
plt.show()
```

Figure 3.10: Programme du modèle régression polynômiale

```python
# Import necessary libraries
from sklearn.ensemble import RandomForestRegressor
from sklearn.model_selection import train_test_split, GridSearchCV
from sklearn.metrics import mean_squared_error, r2_score
```

```python
# Feature selection: Using Signal_Strength and Required_Bandwidth as predictors
X = df[['Signal_Strength', 'Required_Bandwidth']]
y = df['Resource_Allocation']

# Splitting the data into training and testing sets
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)

```

```python
# Setting up the Random Forest Regressor
rf_model = RandomForestRegressor(random_state=42)

# Parameter grid for tuning
param_grid = {
    'n_estimators': [100, 200, 300],  #  Nombre d'arbres dans la forêt
    'max_depth': [4, 6, 8],           # Profondeur maximale des arbres
    'min_samples_split': [2, 5, 10],  # Nombre minimum d'échantillons requis pour diviser un no
    'min_samples_leaf': [1, 2, 4],    # Nombre minimum d'échantillons requis pour être à un nœu
}

# Grid search for hyperparameter tuning
grid_search = GridSearchCV(estimator=rf_model, param_grid=param_grid, cv=3, scoring='r2', verbo
grid_search.fit(X_train, y_train)

# Best model after tuning
best_rf = grid_search.best_estimator_
```

```python
# Faire des prédictions
y_pred = best_rf.predict(X_test)

# Evaluate model performance
mse = mean_squared_error(y_test, y_pred)
r2 = r2_score(y_test, y_pred)
print(f"Best Model Parameters: {grid_search.best_params_}")
print(f"Mean Squared Error: {mse}")
print(f"R-squared (Accuracy): {r2 * 100:.2f}%")

```

```python
# Diagramme de dispersion des valeurs réelles par rapport aux valeurs prévues avec la puissance
plt.figure(figsize=(10, 6))
plt.scatter(X_test['Signal_Strength'], y_test, color='blue', label='Actual', alpha=0.6)
plt.scatter(X_test['Signal_Strength'], y_pred, color='red', label='Predicted', alpha=0.6)
plt.xlabel('Signal Strength (dBm)', fontsize=14)
plt.ylabel('Resource Allocation', fontsize=14)
plt.title('Resource Allocation vs. Signal Strength and Predicted Values (Random Forest)', fonts
plt.legend()
plt.grid(True, linestyle='--', alpha=0.7)
plt.tight_layout()
plt.show()
```

Figure 3.11: Programme du modèle foret aléatoire

3.5.3 Résultats et analyse

Après l'entraînement des différents modèles sur les données historiques, nous les avons testés sur un jeu de données de validation afin de mesurer leur efficacité prédictive. Les graphs et les résultats suivants mettent en évidence les différences de performances entre les approches simples (comme la régression linéaire) et les modèles plus complexes (comme les forêts aléatoires), à travers des indicateurs standards tels que le RMSE, le MAE ou le coefficient de détermination R^2.

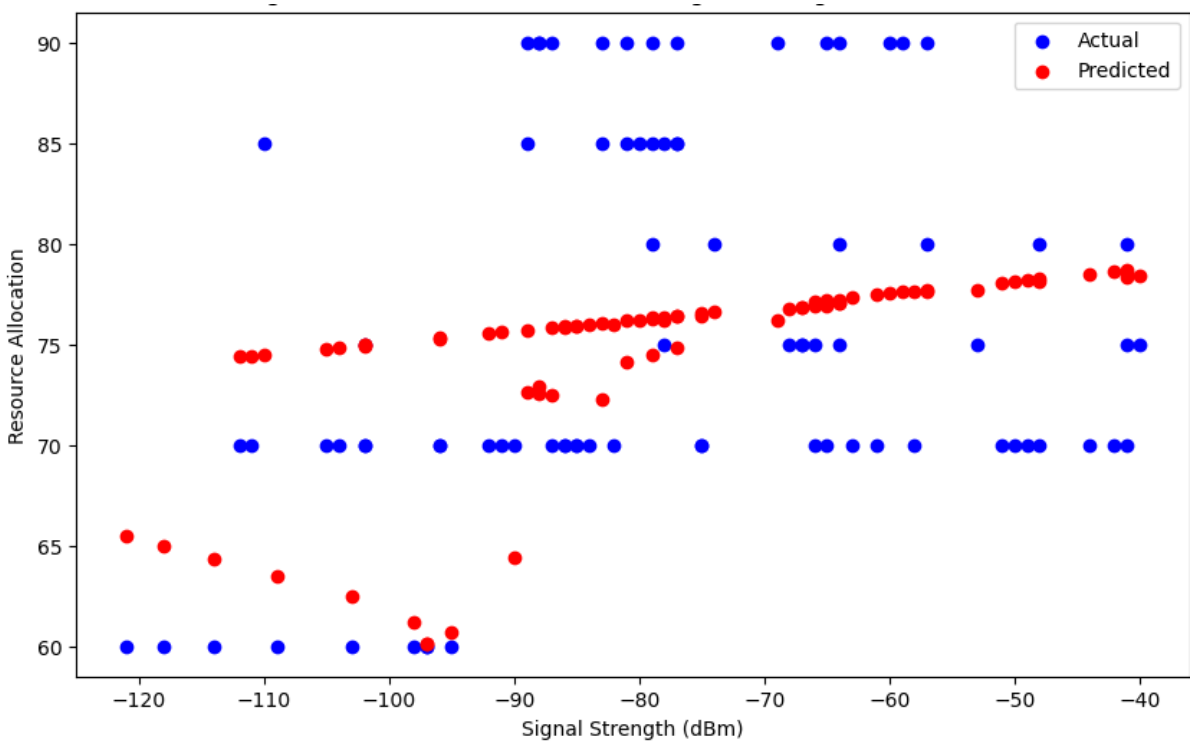


Figure 3.12: Prédiction à l'aide de la régression linéaire

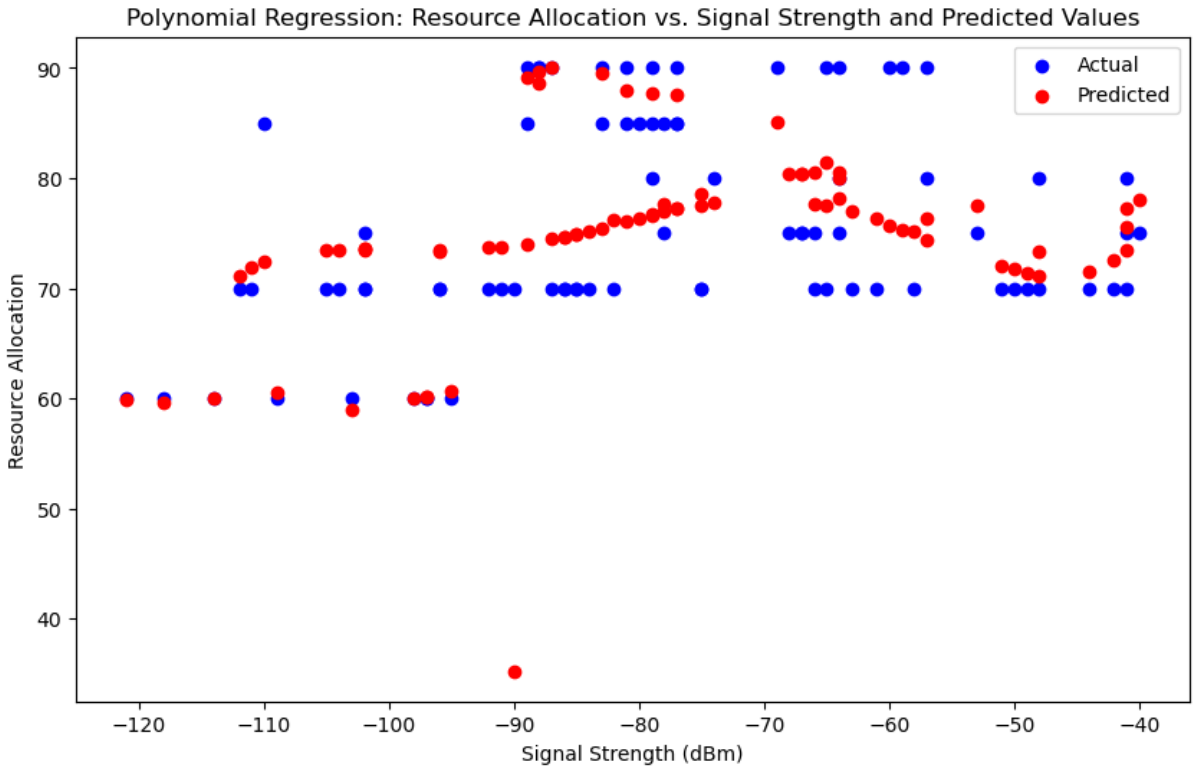


Figure 3.13: Prédiction à l'aide de la régression polynomiale

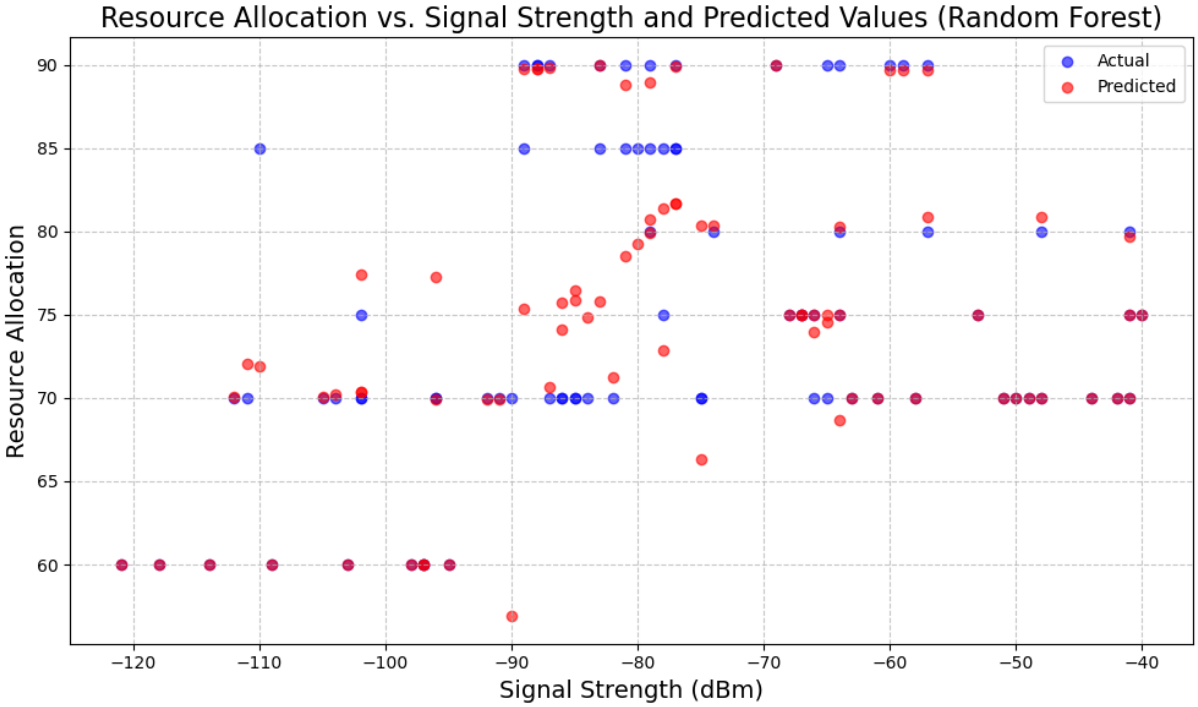


Figure 3.14: Prédiction à l'aide de la forêt aléatoire

La robustesse de l'allocation des ressources est testée sur un ensemble de données 5G. Cela comprend l'acquisition des données, le prétraitement, l'ingénierie des caractéristiques, la sélection des modèles qui impliquent les étapes de formation du modèle et le test avec plusieurs modèles ML afin de minimiser la perte et d'améliorer la précision de la prédiction.

Le graphique de la figure 3.12 illustre l'allocation de ressources prévue en fonction de l'intensité du signal. Le point rouge indique les valeurs prédites. Pour une puissance de signal de 75, l'algorithme d'apprentissage automatique de la forêt aléatoire montre généralement un pourcentage d'allocation des ressources de 82,4, ce qui constitue une bonne prédiction pour la suite de l'optimisation. Le graphique de régression linéaire est illustré à la figure 3.10. L'analyse est effectuée pour quantifier la relation entre l'allocation des ressources et la force du signal afin de comprendre comment les changements dans l'allocation des ressources affectent la force du signal. Chaque point du graphique représente une paire de valeurs réelles et prédites de l'intensité du signal. La ligne de régression est la ligne de meilleur ajustement qui minimise les différences (erreurs) entre les valeurs prédites et les valeurs réelles. Le graphique déduit le meilleur ajustement lorsque les valeurs prédites prennent la ligne de régression. De la même manière, la figure 3.11 illustre les valeurs prédites par la régression polynomiale qui s'adaptent à la ligne de régression.

Dans notre étude, nous avons amélioré, prédit d'autres idées en appliquant divers algorithmes complexes d'apprentissage automatique pour l'optimisation. Il s'avère les valeurs MSE et R^2 sont des mesures précieuses pour évaluer les performances des modèles utilisés dans l'allocation des ressources pour les réseaux 5G. L'ensemble de données a donc été testé sur les modèles d'apprentissage automatique afin de prédire l'adéquation globale. Le tableau 3.4 présente les résultats obtenus par ligne. Les meilleures valeurs MSE sont obtenues avec le modèle Random Forest qui prédit une allocation efficace puisqu'il donne le MSE le plus bas (21,35). Cela indique qu'il fournit la meilleure adéquation globale à la ressource parmi les autres modèles.

Tableau 3.4 : Comparaison des performances des modèles

Modèle	**Erreur quadratique moyenne MSE**	**Valeur de R^2**
Régression Linéaire	69,07	0,21
Régression polynômiale	47,63	0,46
Forêt aléatoire	21,35	0,76

Ces résultats montrent que plus les modèles sont capables de gérer la complexité des relations entre variables (non-linéarité, interactions), plus ils produisent des prédictions fiables. La forêt aléatoire se révèle être le modèle le plus performant dans notre cas, avec une erreur significativement plus faible et une meilleure capacité de généralisation. La régression linéaire, bien qu'interprétable, est trop limitée pour ce type de problématique. Le modèle de la forêt aléatoire permet donc de prévoir une meilleure allocation des ressources. Une solution est donc proposée aux fournisseurs de services pour le déploiement des stations de base.

3.6 Conclusion

Dans ce chapitre, nous avons exploré les complexités et les défis associés à l'allocation des ressources dans les réseaux 5G et discuté des différents facteurs qui influencent cet aspect critique de la gestion du réseau. En analysant l'intensité du signal, la demande des utilisateurs, la capacité du réseau, les exigences des applications et d'autres variables pertinentes, nous avons cherché à donner un aperçu de l'optimisation des stratégies d'allocation des ressources afin d'améliorer les performances du réseau et l'expérience de l'utilisateur. En donnant la priorité à l'amélioration de l'intensité du signal, on peut réduire le temps de latence, améliorer la fiabilité et optimiser l'utilisation de la bande passante, ce qui, en fin de compte, améliore la qualité de service pour les utilisateurs. Des mesures d'évaluation telles que le R^2 et l'erreur quadratique moyenne (MSE) ont été utilisées pour évaluer les performances des modèles. Dans les travaux futurs, l'allocation des ressources basée sur le ML peut être davantage axée sur l'étude des réseaux neuronaux profonds afin de trouver la meilleure précision, ce qui donne une justification supplémentaire à l'allocation des ressources.

CONCLUSION GÉNÉRALE :

Dans ce mémoire de recherche, nous pouvons dire que notre objectif, qui était de concevoir, entraîner et évaluer des modèles de machine learning pour la prédiction de l'allocation des ressources dans les réseaux 5G, a été atteint.Dans un premier temps, nous avons étudié la cinquième génération de réseaux mobiles (5G), qui se distingue par sa capacité à offrir un haut débit, une latence ultra-faible et une connectivité massive, ouvrant la voie à de nouveaux usages critiques tels que l'Internet des objets (IoT), la télémédecine ou les véhicules autonomes. Pour répondre à ces exigences, la 5G repose également sur une évolution de l'architecture réseau à travers les paradigmes SDN (Software Defined Networking) et NFV (Network Function Virtualization). Ces technologies apportent souplesse, programmabilité et virtualisation des fonctions réseau, facilitant ainsi une gestion fine, dynamique et centralisée des ressources.

Dans ce contexte, nous avons exploré l'apport du machine learning pour améliorer l'allocation des ressources dans ces environnements complexes. À partir d'un jeu de données réelles, nous avons appliqué plusieurs étapes de prétraitement, d'analyse exploratoire, puis construit différents modèles supervisés (régression linéaire, régression polynomiale et forêt aléatoire). Chaque modèle a été entraîné sur des paramètres clés du réseau (signal, latence, bande passante) pour prédire la ressource allouée.Les résultats montrent que la forêt aléatoire est la plus performante, atteignant une variance expliquée de plus de 75 %, confirmant que les modèles d'ensemble sont bien adaptés à ce type de problématique. Cette approche ouvre la voie à une allocation prédictive, optimisée et intégrée dans un contrôleur SDN pour un réseau plus intelligent.

En somme, cette étude a démontré la faisabilité et la pertinence d'une solution hybride combinant 5G, SDN/NFV et apprentissage automatique pour anticiper et optimiser l'allocation des ressources réseau. Pour aller plus loin, il serait intéressant d'intégrer des modèles de réseaux de neurones, d'étendre l'étude à des environnements IoT ou de tester l'approche en temps réel dans un environnement SDN simulé (ex : avec Mininet ou GNS3).

Références Bibliographiques :

[1] Mémoire, BEDDIAF Safia, Y.S. 2019.“Technologie radio sur fibre pour les applications sans fil(5G)” 127 pages.

[2] Article, L.J.Vora, 2015.“EVOLUTION OF MOBILE GENERATION TECHNOLOGY : 1G TO 5G AND REVIEW OF UPCOMING WIRELESS TECHNOLOGY 5G”, 11 pages.

[3] Article, Nikhil Bhandari, S.D.K.S 2017.“Evolution of Cellular Network : From 1Gto 5G”, International Journal of Engineering and Techniques-Volume 3 Issue 5,Sep-Oct 2017, 9 pages

[4] Article, P. Sharma, 8 August 2013.“Evolution of Mobile Wireless CommunicationNetwork-1G to 5G as well as Future Prospective of Next Generation Communication Network”,International Journal of Computer Science and Mobile Computing Vol.2 Issue, pages 47-53, 7 pages

[5] Cour, Huawei. 2019.“5G Motivation and Industry Progress”, 296 pages

[6] Thèse, A.Ngom, 2019.“Conception de petits réseaux d’antennes reconfigurable ou”Small-Cells” pour le standard 5G.”, 165 pages

[7] Cour, Huawei, 2021.“5G Development and Evolution.”, 464 pages

[8] Article, N.U.M.O.C.O.a.A.J.A.Opeoluwa Tosin Eluwole,28 August 2018.“”From 1G to 5G, What Next ?”.”IAENG International Journal of Computer Science, 45 :3,IJCS306, 22pages

[9] Cour, Agence National des fréquence (ANFR), 2019. “Evaluation de l’exposition du public aux ondes ´électromagnétique.”, Volet 1 : Présentation générale de la 5G,17 pages

[10] Mémoire, R.Arthure. 2018.“DIMENSIONNEMENT DU RESEAU 5G.”, 82 pages

[11] Cour, Abacus AVNET. 2021.“Comprendre la technologie Massive MIMO.”, 5 pages

[12] Mémoire, B.A.M.e.G.Rania, 2018 .“Etude de méthode de beamforming `a maximisation du rapport signal sur bruit pour des systèmes FBMC.”, 91 pages

[13] Mémoire, D.A.A. Bachir Sadou, 2020. “Etude des techniques de modulation pour les réseaux mobiles 5G et 6G nouvelles générations.”, 94 pages

[14] M. Giordani, M. Polese, M. Mezzavilla, et al., 2020. “Toward 6G Networks: Use Cases and Technologies.” IEEE Communications Magazine, vol. 58, no. 3, pp. 55–61. (pour une perspective sur le ML dans les réseaux futurs)

[15] Article, T.Salman, 30 Novembre 2015.“Networking Protocols and Standards for Internet of Thing.”, 28 pages

[16] Cour, Huawei, 2018. “Huawei, SDN Concepts and values, 2018..”, 49 pages

[17] ETSI, 2014. “Network Functions Virtualisation (NFV) ; Architectural Framework.” ETSI GS NFV 002 V1.2.1, 21 pages.

[18] N. McKeown, T. Anderson, H. Balakrishnan et al., 2008. “OpenFlow: Enabling Innovation in Campus Networks.” ACM SIGCOMM Computer Communication Review, vol. 38, no. 2, pp. 69–74.

[19] B. A. A. Nunes, M. Mendonca, X. N. Nguyen, K. Obraczka and T. Turletti, 2014. “A Survey of Software-Defined Networking: Past, Present, and Future of Programmable Networks.” IEEE Communications Surveys & Tutorials, vol. 16, no. 3, pp. 1617–1634.

[20] M. Chiosi et al., 2012. “Network Functions Virtualisation – Introductory White Paper.” ETSI NFV ISG, 17 pages.

[21] X. Foukas, G. Patounas, A. Elmokashfi and M. K. Marina, 2017. “Network slicing in 5G: Survey and challenges.” IEEE Communications Magazine, vol. 55, no. 5, pp. 94–100.

[22] A. Lara, A. Kolasani and B. Ramamurthy, 2014. “Network Innovation using OpenFlow: A Survey.” IEEE Communications Surveys & Tutorials, vol. 16, no. 1, pp. 493–512.

[23] A. Gember-Jacobson, R. Viswanathan, C. Prakash, et al., 2014. “OpenNF: Enabling innovation in network function control.” ACM SIGCOMM, pp. 163–174.

[24] R. Mijumbi, J. Serrat, J.-L. Gorricho, et al., 2016. “Network Function Virtualization: State-of-the-Art and Research Challenges.” IEEE Communications Surveys & Tutorials, vol. 18, no. 1, pp. 236–262.

[25] D. Kreutz, F. M. V. Ramos, P. Verissimo, et al., 2015. “Software-Defined Networking: A Comprehensive Survey.” Proceedings of the IEEE, vol. 103, no. 1, pp. 14–76.

[26] Article, ETSI, 2014. “Network Functions Virtualisation (NFV) ; Architectural Framework.”, 21 pages.

[27] Y. Li and M. Chen, 2015. “Software-Defined Network Function Virtualization: A Survey.” IEEE Access, vol. 3, pp. 2542–2553.

[28] H. Hawilo, A. Shami, M. Mirahmadi and R. Asal, 2014. “NFV: State of the Art, Challenges, and Implementation in Next Generation Mobile Networks (NGMN).” IEEE Network, vol. 28, no. 6, pp. 18–26.

[29] Y. Zhang, M. Wang, and L. Wang, 2020. “Machine Learning in 5G and Beyond: A Survey.” IEEE Access, vol. 8, pp. 89–101.

[30] R. Li, Z. Zhao, X. Zhou, et al., 2018. “Intelligent 5G: When Cellular Networks Meet Artificial Intelligence.” IEEE Wireless Communications, vol. 24, no. 5, pp. 175–183.

[31] M. Chen, U. Challita, W. Saad, et al., 2019. “Artificial Intelligence in Wireless Communications: A Comprehensive Survey.” IEEE Communications Surveys & Tutorials, vol. 21, no. 4, pp. 3039–3071.

[32] S. M. Zafaruddin, P. S. S. Chowdhury, et al., 2020. “Machine Learning for Resource Allocation in 5G HetNets: A Survey.” IEEE Access, vol. 8, pp. 188675–188709.

[33] M. Giordani, M. Polese, M. Mezzavilla, et al., 2020. “Toward 6G Networks: Use Cases and Technologies.” IEEE Communications Magazine, vol. 58, no. 3, pp. 55–61. (pour une perspective sur le ML dans les réseaux futurs)

[34] X. Foukas, G. Patounas, A. Elmokashfi and M. K. Marina, 2017. “Network Slicing in 5G: Survey and Challenges.” IEEE Communications Magazine, vol. 55, no. 5, pp. 94–100.

[35] T. J. Oechtering, F. Gabry, and E. Ekici, 2021. “Machine Learning for Resource Allocation in Wireless Networks: Potentials and Challenges.” EURASIP Journal on Wireless Communications and Networking.

[36] ETSI, 2021. “Zero-touch Network and Service Management (ZSM); Requirements.” ETSI GS ZSM 001 V1.1.1.

[37] A. Gupta and R. K. Jha, 2015. “A Survey of 5G Network: Architecture and Emerging Technologies.” IEEE Access, vol. 3, pp. 1206–1232.