

République Algérienne Démocratique et Populaire

Ministère de l'Enseignement Supérieure et de la Recherche Scientifique

Université Abderrahmane Mira

Faculté de la Technologie



Département d'Automatique, Télécommunication et d'Électronique

## Projet de Fin d'Études

Pour l'obtention du diplôme de Master

**Filière :** Télécommunications

**Spécialité :** Réseaux et Télécommunications

### Thème

**PRÉDICTION DE PERFORMANCES DE RÉSEAUX 5G**

**Préparé par :**

*BOULEGHALEGH Aya - GUERSES Bouthaina*

**Dirigé par :**

Dr M.TOUNSI

Mr Y.OUAZIZ

**Examiné par :**

Dr N.BENAMIROUCHE

Dr N.MEZHOUD

Année universitaire : 2020/202

## ***REMERCIEMENTS***

Au terme de ce travail, nous remercions Dieu le tout puissant de nous avoir donné la force et le courage, la volonté et la patience pour achever ce travail.

Nous avons l'honneur d'exprimer nos remerciements avec un grand plaisir et un grand respect à nos promoteurs Mr. TOUNSI et Mr. OUZZIZ pour leurs conseils, leurs disponibilités et leurs encouragements.

Nous remercions également les membres du jury d'avoir accepté de juger et d'évaluer notre mémoire, et de nous faire part de leurs précieuses remarques.

Enfin, nous tenons à remercier toute personne qui nous a aidés de près ou de loin durant notre travail.

## *Dédicaces*

Ce mémoire est dédié à

Nos parents,

Nos familles,

Nos ami(e)s,

Tous ceux et celles qui nous aiment et que nous aimons.

*GUERSES Bouthaina , BOULEGHALEGH Aya*

## ABRÉVIATIONS

<b>AMPS</b>	Advanced Mobile Phone System
<b>AUC</b>	Authentication Center
<b>BSS</b>	Base Station Sub-system
<b>BPSK</b>	Binary Phase Shift Keying
<b>BTS</b>	La station de base
<b>BSC</b>	Base Station Controller
<b>CDMA</b>	Code division Multiple Access
<b>CN</b>	Core Network
<b>CS</b>	Circuit Switching
<b>CP</b>	Cyclic prefix
<b>DFTS-OFDMA</b>	Direct Fourier Transform Spread-OFDMA
<b>ETACS</b>	Extended Total Access Communication System
<b>EIR</b>	Equipment Identification Register
<b>EDGE</b>	Enhanced Data Rates for GSM Evolution
<b>EPS</b>	Evolved Packet System
<b>E-UTRAN</b>	Evolved UTRAN
<b>EPC</b>	Evolved Packet Core
<b>FDD</b>	Frequency Division Duplex
<b>FDMA</b>	Frequency-Division Multiple Access
<b>HLR</b>	Home Location Register
<b>HSPA</b>	High Speed Packet Access
<b>HSDPA</b>	High Speed Uplink Packet Access
<b>HSUPA</b>	High Speed Downlink Packet Access
<b>HSPA+</b>	High Speed Packet Access +
<b>HSS</b>	Home Subscriber Server
<b>I-HSPA</b>	Internet-HSPA
<b>IMEI</b>	International Mobile Equipment Identity
<b>IMSI</b>	International Mobile subscriber Identity
<b>IMS</b>	IP Multimedia Subsystem
<b>GSN</b>	Gateway Serving GPRS Support Node
<b>G-MSC</b>	Mobile services Switching Center Mobile Equipment
<b>GSM</b>	Gateway Global System for Mobile Communication
<b>GPRS</b>	General Packet Radio Service

<b>ME</b>	Mobile Equipment
<b>MIMO</b>	Multiple Input Multiple Output
<b>MSC</b>	Mobile services Switching Center Mobile Equipment
<b>MME</b>	Mobility Management Entity
<b>MISO</b>	Multiple-input and single-output
<b>NSS</b>	Network Switching Subsystem
<b>NMC</b>	network management center
<b>OSS</b>	Operation SubSystem
<b>OMC</b>	Operation and maintenance center
<b>OFDM</b>	Orthogonal Frequency-division Multiplexing
<b>OFDMA</b>	Orthogonal Frequency Division Multiple Access
<b>PS</b>	Packet Switching
<b>P-GW</b>	Packet Data Network Gateway
<b>PRB</b>	Physical Resource Block
<b>PAPR</b>	Peak to Average Power Ratio
<b>QPSK</b>	Quadrature Phase Shift Keying
<b>RNC</b>	Radio Network Controller
<b>RE</b>	Ressource Element
<b>RB</b>	Ressource Bloc
<b>SGSN</b>	Serving GPRS Support Node
<b>SIM</b>	subscriber identity/identification module
<b>SAE</b>	System Architecture Evolution
<b>Serving</b>	GW (Serving Gateway
<b>SC-FDMA</b>	Single-Carrier
<b>SIMO</b>	single-input and Multiple-output
<b>TDD</b>	Time Division Multiple Access
<b>TD-CDMA</b>	Time Division-Code division Multiple Access
<b>TDMA</b>	Time Division Multiple Access
<b>UMTS</b>	Universal Mobile Telecommunications System
<b>USIM</b>	Universal Subscriber Identity Module
<b>UTRAN</b>	Universal Terrestrial Radio Access Network
<b>VLR</b>	Visitor Location Register
<b>W-CDMA</b>	Wideband-Code division Multiple Access

## ***TABLE DES MATIÈRES***

<b>Introduction Générale</b>	<b>1</b>
<b>Chapitre I</b>	
I.1. Introduction.....	3
I.2. Réseaux cellulaire de première générations.....	3
I.2.1 AMPS .....	3
I.2.2 ETACS .....	3
I.2.3 Radicom2000 .....	3
I.2.4 NMT .....	3
I.3. Réseaux cellulaire de deuxième générations .....	4
I.3.1 Concept cellulaire.....	4
I.3.2 Architecture globale du réseau GSM .....	5
I.3.2.1 ME (Mobile Equipment) La station mobile.....	5
I.3.2.2 BSS (Base Station Sub_system) Le sous-système radio .....	5
I.3.2.3 NSS (Network Switching Sub_system).....	5
I.3.3 FDD.....	7
I.3.4 Technique d'accès .....	7
I.3.4.1 Le partage en fréquence FDMA (Frequency Division Multiple Access).....	7
I.3.4.2 Le partage en Temps TDMA (Time Division Multiple Access).....	7
<b>I.3.5</b> Evolution du réseau GSM .....	7
I.4 Réseaux cellulaire de troisième générations .....	8
I.4.1 Architecture globale du réseau UMTS.....	8
I.4.1.1 Réseau cœur CN (Core Network).....	8
I.4.1.2 Réseau d'accès radio (UTRAN).....	9
I.4.1.3 Station mobile (MS) .....	10
I.4.2 Evolution de l'UMTS.....	11
I.5 Réseau cellulaire de La quatrième génération .....	11
I.5.1 Architecture de réseau LTE.....	11
I.5.1.1 E-UTRAN (Evolved UTRAN) .....	12
I.5.1.2 Réseau cœur EPC (Evolved Packet Core).....	12
I.5.2 La capacité.....	12
I.5.3 La dimension fréquentielle en LTE.....	13
I.5.4 Techniques d'accès multiple.....	14

I.5.4.1 Multiplexage par répartition orthogonale de la fréquence (OFDM).....	15
I.5.4.2 Accès multiple par répartition en fréquence orthogonale (OFDMA).....	15
I.5.4.3 Single-Carrier FDMA (SC-FDMA) .....	15
I.5.5. La latence .....	16
I.5.5.1. Latence du plan de contrôle.....	16
I.5.5.2. Latence du plan usager .....	16
I.5.6. Les Service IP.....	16
I.5.7. L’interface Radio LTE .....	16
I.5.7.1. Interférence de canal.....	17
I.5.7.2. Path loss (les pertes de propagation) .....	17
I.5.7.3. Les évanouissements rapides (fast fading) .....	17
I.5.7.4. MIMO .....	18
I.5.7.5. Préfixe cyclique (CP).....	19
I.6.Conclusion .....	19
<b>Chapitre II</b>	
II.1 Introduction .....	20
<b>II.2.</b> Réseaux cellulaire 5G.....	20
<b>II.3.</b> Fréquences attribuées à la 5G.....	20
<b>II.3.1.</b> Les bandes de fréquences 5G Sub-6 .....	21
<b>II.3.2.</b> Les ondes millimétriques mmWave.....	21
<b>II.4.</b> Architecture du réseau 5G.....	21
II.4.1. Le mobile UE (User Equipement) .....	22
II.4.2. L’accès radio 5G (NG-RAN).....	23
II.4.3. Réseau cœur 5g (NGCN ou 5GC).....	23
II.5. Tranchage (Slicing) de réseau 5G.....	23
II.6. Les modes de Déploiement du réseau 5G .....	24
<b>II.6.1.</b> Mode NAS (Non-standalone access) .....	25
II.6.2. Mode AS (standalone access) .....	25
<b>II.7.</b> Les technologies envisagées par NR .....	25
II.7.1. Massive MIMO .....	26
II.7.2. Accès multiple non-orthogonal NOMA.....	27
II.7.3. Beamforming .....	27
II.7.4. Antenne intelligente .....	28
II.7.5. Accès multiple a division spatiale.....	28

II.7.6. Antenne adaptative.....	29
<b>II.7.7.</b> Algorithme adaptative.....	30
II.8. Conclusion.....	30
<b>Chapitre III</b>	
III.1 Introduction.....	31
III.2 L'apprentissage automatique (Machine learning).....	31
III.3 Apprentissage supervisé et apprentissage non supervisé.....	31
III.3.1 Apprentissage supervisé.....	31
<b>III.3.2</b> Apprentissage non supervisé.....	32
III.4 Fonction d'hypothèse.....	32
III.5 Fonction de Cout (Cost Function).....	32
III.6 La Descente de gradient.....	33
<b>III.7</b> Régression linière avec variable multiple.....	34
<b>III.8</b> Mise à l'échelle des fonctionnalités.....	35
<b>III.9</b> Les différentes techniques de mise à l'échelle (Feature Scaling).....	35
<b>III.9.1</b> Normalisation.....	35
<b>III.9.2</b> Standardisation.....	36
<b>III.10</b> Régression polynomial.....	36
III.11 Équation normale.....	36
III.11.1 Non-inversibilité de l'équation normale.....	36
III.12 Régression logistique.....	37
<b>III.12.1</b> Principe.....	38
III.12.2 La décision.....	40
III.13 Fonction du cout.....	40
III.14 Classification multi classe.....	40
III.15 Le problème du sur apprentissage (Overfitting).....	40
III.16 fonction de cout.....	41
III.17 La régularisation.....	41
III.17.1 Régression linéaire régularisée.....	42
III.17.1.1 Descente graduelle.....	42
III.17.1.2 Équation normale.....	42
III.17.2 Régression Logistique Régularisée.....	43
III.17.2.1 Cost function.....	43
III.17.2.2 La descente de gradient.....	43

III.18 Conclusion.....	43
<b>Chapitre IV</b>	
<b>IV.1.</b> Introduction.....	44
IV.2. Génération des données .....	44
IV.2.1. Émetteurs et récepteurs .....	44
IV.2.2. Calcul des puissances reçues.....	45
IV.3. Régression linéaire à variable unique .....	46
IV.3.1. Représentation des données.....	46
IV.3.2. Fonction Cout.....	46
IV.3.3. Optimisation par descente de gradient .....	47
IV.3.4. Graphe de la fonction d’hypothèse.....	47
<b>IV.4.</b> Régression linéaire à multiple variables.....	48
<b>IV.4.1.</b> Équation normale .....	48
IV.4.2. Graph de la fonction d’hypothèse.....	49
IV.5. Conclusion .....	50
<b>Conclusion Générale .....</b>	<b>51</b>
<b>Bibliographie.....</b>	<b>52</b>

---

# INTRODUCTION GÉNÉRALE

---

Les réseaux mobiles occupent de nos jours, une place plus importante en terme de recherche et d'investissement. Le téléphone mobile est devenu un outil incontournable dans le quotidien de nos sociétés.

Avec la demande croissante et l'augmentation du nombre d'abonnés, la quatrième génération communément appelée Long Term Evolution (LTE) n'as pas permis de répondre aux besoins des utilisateurs. Il fallait donc pousser le développement vers d'autres améliorations afin d'augmenter la capacité des réseaux et d'assister à une révolution intelligente, ce qui a donné naissance à la cinquième génération 5G.

Dans ce mémoire de fin d'étude, nous allons aborder une étude introductive aux techniques de prédiction de performances de réseaux mobiles 5G. Notre but est de nous aider des techniques de régression qui sont des algorithmes d'apprentissage automatique pour prédire des performances d'un réseau 5G avant ou en cours d'exploitation avant d'adapter la configuration du réseau. De plus en plus de fonctions d'apprentissage automatique sont en effet en cours d'intégration dans les émetteurs et récepteurs RF afin d'arriver à des réseaux cognitifs.

Nos investigations sont synthétisées dans ce présent mémoire structuré en quatre chapitres :

- Dans le premier chapitre, nous dressons une revue des différentes évolutions des réseaux mobiles de la première à la quatrième génération ;
- Dans le second chapitre, nous présentons les caractéristiques du nouveau standard 5G, notamment son interface radio NR (new Radio), et des différentes technologies utilisées pour déployer cette nouvelle génération radio-mobile ;
- Le troisième chapitre traite de techniques d'apprentissage automatique (Machine Learning ) notamment pour un objectif de régression linéaire à variable unique et à variables multiples. Nous y présentons le formalisme mathématique des algorithmes utilisés ;

- Le dernier chapitre reporte les résultats des différentes simulations de prédiction de performances de réseaux 5G selon différentes contraintes.

Nous terminons par une conclusion générale suivie des références bibliographiques que nous avons utilisées.

---

*CHAPITRE I*

*ÉVOLUTION DES RÉSEAUX MOBILES*

---

## **I.1. Introduction**

De nos jours, les systèmes de communication sans fils sont devenus un outil incontournable dans le quotidien de nos sociétés. Les communications sans fil ont connu une évolution et une progression constante, définie par quatre générations allant de la téléphonie mobile (1G, GSM) aux communications multimédia (3G, LTE) [1]. Ces générations diffèrent principalement par les techniques d'accès à la ressource radio.

Dans ce chapitre, nous allons voir la composition et les caractéristiques des différentes générations.

## **I.2. Réseaux mobiles de première génération**

La première génération (1G) des systèmes cellulaires reposait sur un système de communications mobiles analogiques et ne propose aucun service autre que la voix. Cette génération a bénéficié de deux inventions techniques majeures des années 1970 : le microprocesseur et le transport numérique des données entre les téléphones mobiles et la station de base. Les appareils utilisés étaient particulièrement volumineux. D'autre part, cette génération utilise les standards suivants :

### **I.2.1 AMPS**

L'AMPS (Advanced Mobile Phone System) était le premier standard de réseaux cellulaires. Il possédait de faibles mécanismes de sécurisation rendant possible le piratage des lignes téléphoniques. Il était Basé sur un codage et une modulation de type analogique et il utilisait la technique d'accès multiple FDMA sur la fréquence 800 MHz. Ce standard dispose d'un nombre total de 832 canaux de largeur 30 KHz chacun.

### **I.2.2 ETACS**

C'est une version améliorée du standard TACS, il été basé sur un codage et une modulation de type analogique. Il utilisait la technique d'accès multiple FDMA sur la fréquence 900 MHz, et disposait de 1000 canaux de largeur 25 KHz chacun.

### **I.2.3 Radicom2000**

Apparu en 1985, il venait remplacer progressivement le réseau analogique. Ce réseau utilisait la technologie numérique pour la signalisation et la modulation analogique pour la voix. Les fréquences étaient attribuées dynamiquement en fonction des besoins.

### **I.2.4 NMT**

Le système est développé par les pays nordiques, pour répondre aux problèmes de congestion des réseaux de téléphonie existants à cette époque. La première version, le NMT-450 était introduit en 1981, fonctionne à 450 MHz sur la bande (450-470 MHz). Ensuite la seconde

version, NMT-900, a été introduite en 1986, fonctionne sur une fréquence de 900 MHz, qui est utilisée pour le GSM. [2]

La première génération recense certains problèmes ce qui reconduit les chercheurs à améliorer la qualité de services. La numérisation et la miniaturisation des équipements ont rendu la communication mobile accessible et beaucoup plus flexible, ce qui a permis l'introduction de la deuxième génération de téléphonie mobile, notamment le GSM.

### I.3. Réseaux cellulaires de deuxième générations

Le GSM est la première norme de téléphonie mobile numérique. Le GSM a connu un grand succès, grâce notamment à équipements terminaux plus petits, plus gérable et plus facile à transporter, avec plus d'autonomie et à faible coût. Il offre plusieurs services à débits différents comme transmission de données et de messages multimédia.

#### I.3.1 Concept cellulaire

Le principe est de diviser le territoire en de petites zones, appelées cellules sous forme d'hexagone, et de partager les fréquences radio entre celles-ci. Chaque cellule est constituée d'une station de base comportant une antenne généralement omnidirectionnelle tri-sectorielle reliée au Réseau Téléphonique Commuté RTCP à laquelle on associe un certain nombre de canaux de fréquences à bande étroite. Ces fréquences ne peuvent pas être utilisées dans les cellules adjacentes afin d'éviter les interférences. En pratique, une distance minimale sépare deux cellules utilisant la même fréquence suivant l'équation I.1. Ainsi, on définit des *motifs*, ou *clusters*, constitués de plusieurs cellules, dans lesquelles chaque fréquence est utilisée une seule fois. [3]

$$D = \sqrt{3 \cdot N} \cdot a \quad (I.1)$$

où  $a$  est le rayon d'une cellule,  $N$  représente la taille du motif de réutilisation. La figure suivante représente la distance de réutilisation des fréquences.

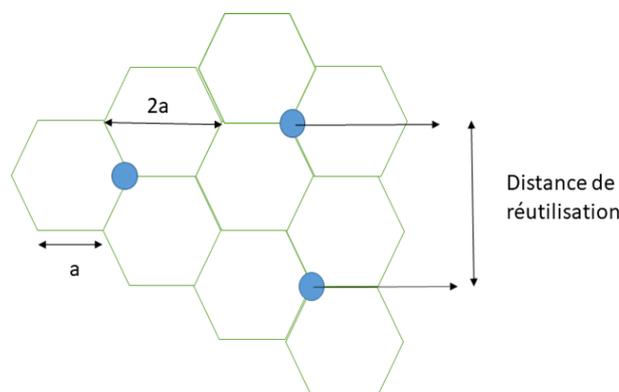


Figure I.1 : Distance de réutilisation de fréquence

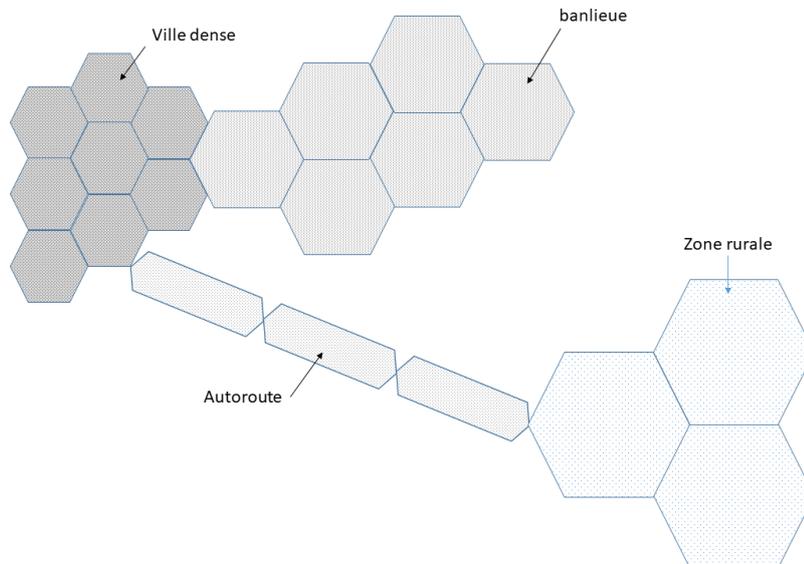


Figure I.2 : Distribution des cellules selon la région d'étude

### I.3.2 Architecture globale du réseau GSM

L'architecture d'un réseau GSM peut être divisée en trois sous-systèmes (Cf. figure I.3) [3] :

#### I.3.2.1 ME (Mobile Equipment) La station mobile :

Représente la partie visible d'un système radio mobile, permet à l'abonnés d'accéder au service GSM au travers du système cellulaire.

La station mobile est composée d'un combiné téléphonique identifié par un numéro unique l'IMEI (International Mobile Equipment Identity) et d'une carte SIM qui contient le numéro d'identification de l'abonné IMSI (International Mobile subscriber Identity) et des algorithmes de chiffrement.

#### I.3.2.2 BSS (Base Station Sub\_system) Le sous-système radio

C'est un sous-système qui assure les transmissions radioélectriques et gère la ressource radio. Il est composé de :

- **BTS (La station de base) :** c'est un groupe d'émetteurs et de récepteurs TRX qui assurent le lien radioélectrique avec les abonnés mobiles.
- **BSC (Le contrôleur de station de base) :** il gère une ou plusieurs stations de base. Cet organe assure l'allocation de canaux, la gestion du saut de fréquence, le transfert intercellulaire des communications et la gestion de la signalisation sur voie radio.

#### I.3.2.3 NSS (Network Switching Sub\_system)

Regroupe les sous-systèmes qui assurent des fonctions du niveau réseau (commutation, routage, interconnexion). Il est constitué de :

- **HLR (Home Location Register)** : est la base de données centrale contenant toutes les informations de localisation et de caractéristiques des abonnés, il mémorise pour chaque abonné le numéro du VLR où il est enregistré.
- **VLR (Visitor Location Register)** : est une base de données associée à un commutateur MSC qui enregistre temporairement les informations relatives aux terminaux présents et actifs (au moins en veille) dans son secteur de couverture.
- **MSC (Mobile services Switching Center)** : centre de commutation généralement associés aux bases de données VLR. Il assure une interconnexion entre le réseau mobile et le réseau téléphonique public RTCP

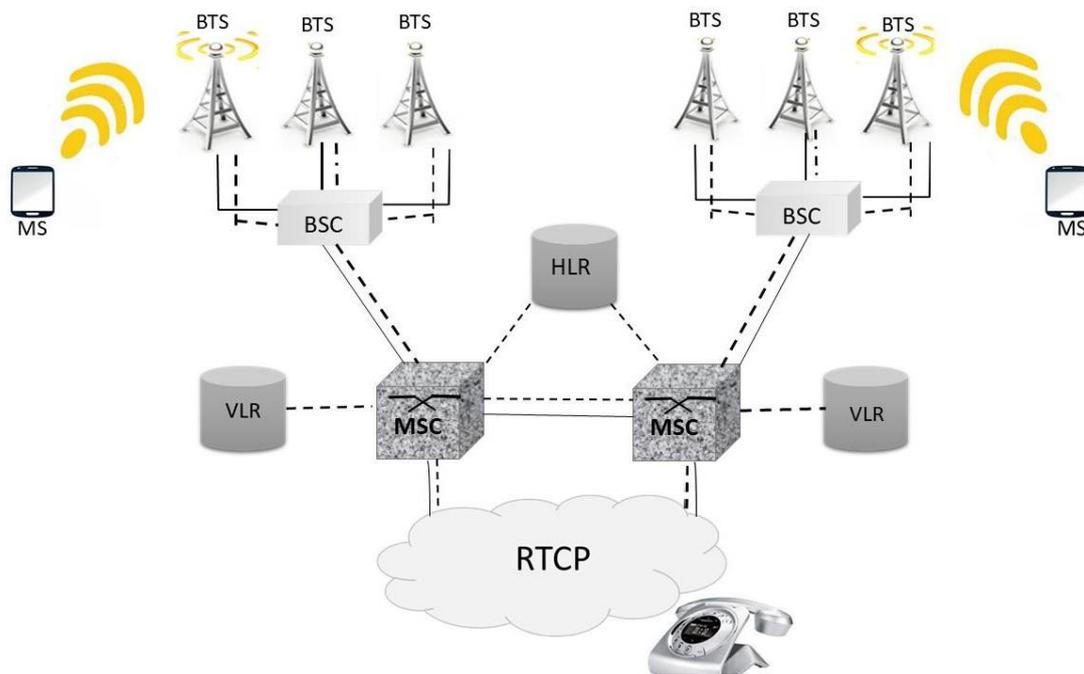


Figure I.3 : Infrastructure d'un réseau GSM

#### I.3.2.4 OSS (Operation Sub\_System) sous système d'exploitation et de maintenance :

Cette partie assure la gestion et la supervision du réseau elle est constituée de :

- **AUC (Authentication Center) centre d'authentification** : mémorise pour chaque abonné une clé secrète qui est utilisée pour authentifier les demandes de services et pour chiffrer les communications.
- **L'EIR (Equipment Identification Register)** : registre d'identités des équipements. C'est une base de données contenant les identités des terminaux. Ce sous système peut être consulté lors des demandes de services d'un abonné pour vérifier si un ME (est autorisé a accéder au système) a ou il n'a pas le droit d'accès au système.

- **Le NMC (network management center) :** centre d'administration d'un réseau permet l'administration générale de l'ensemble du réseau par un contrôle centralisé.
- **OMC (Operation and maintenance center) :** centre d'opération et de maintenance permettent une supervision locale des équipements.

### I.3.3 FDD

Le réseau GSM utilise le réseaux FDD. C'est à dire que la communication est effectuée sur deux fréquences indépendantes et que les données d'émission et de réception sont simultanées. C'est un système de communication Duplex typique.

### I.3.4 Technique d'accès

Le spectre de fréquence, est partagé selon un plan temps/fréquence afin d'obtenir des canaux physiques capables de supporter une communication téléphonique.

#### I.3.4.1 Le partage en fréquence FDMA (Frequency Division Multiple Access)

Cette technique consiste sur le découpage de chaque bande dédiée au GSM en 124 canaux fréquentiels d'une largeur de 200khz. Sur une bande de fréquence sont émis des signaux modulés autour d'une fréquence porteuse qui siège au centre de la bande.

#### I.3.4.2 Le partage en Temps TDMA (Time Division Multiple Access)

C'est le partage d'une porteuse sur plusieurs utilisateurs. la porteuse est divisé sur 8 time slot, chaque slot a une durée de 0.577 ms. Il permet de véhiculer des éléments d'information appelés « burst » a une durée de 0.546 ms. Une trame TDMA sa durée est de  $8 \times 0.577 = 4.6152$  ms comme le montre la figure ci-dessous :

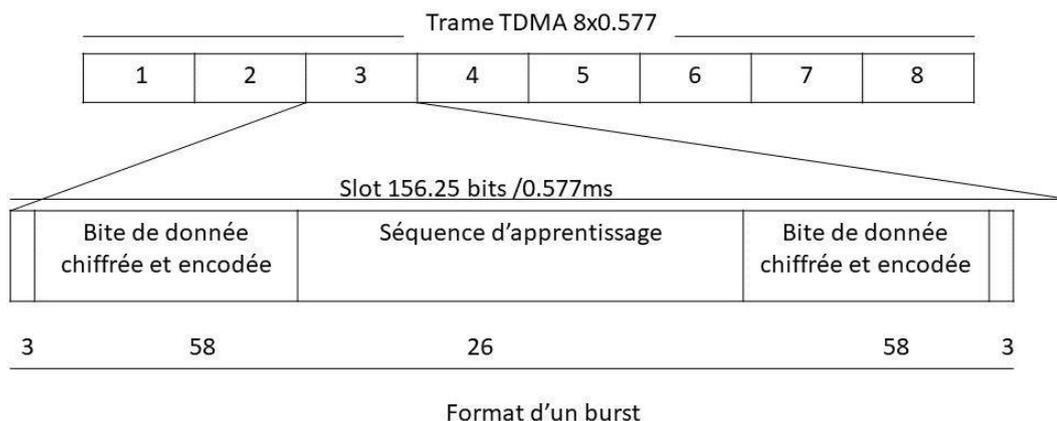


Figure I.4 : format d'une trame TDMA

### I.3.5 Évolution du réseau GSM

Avec l'augmentation du nombre d'utilisateurs, il est nécessaire d'attribuer une bande supplémentaire aux alentours de 1800 MHz, d'où la nécessité de porter la technologie GSM 900

vers une bande à plus haute de fréquence. C'est le système GSM 1800 MHz connu aussi sous le nom de DCS-1800 dont les caractéristiques sont identiques au GSM en termes de protocoles et de service. [3]

Le GPRS peut être considéré comme une évolution des réseaux GSM (2,5G ou la 2G+) avant leur passage aux systèmes de troisième génération en termes de service et de débits (environ 114 kbit/s). Un réseau GPRS est un réseau IP constitué de routeurs IP.

Le GPRS considéré comme mise à jour logicielle et matérielle des éléments de base du réseau GSM. Ainsi l'introduction de nouveaux équipements permettant l'accès au réseau de données (ex. Internet). [4]

Le standard EDGE (Enhanced Data Rates for GSM Evolution, 2.75G) défini par le 3GPP, est une évolution de la norme GSM en modifiant le type de modulation. Tout comme la norme GPRS, le standard EDGE est une extension du réseau GPRS, Seule le sous-système radio est sensiblement modifiée, Il a introduit une nouvelle modulation nommée 8-PSK, autorisant ainsi des débits théoriques maximums de 384 Kbit/s et un débit réel de 177 Kbit/s, ce qui permet de dépasser les limites connues auparavant, ouvrant alors la porte aux applications multimédias (la troisième génération de téléphonie mobile 3G). [5]

#### **I.4 Réseaux cellulaires de troisième génération**

La 3G est apparue dans les années 2000. La standardisation d'un système 3G est nommée UMTS, le réseau 3G est liée aux réseaux GSM et GPRS avec leurs fonctionnalités respectives de la voix et data. Cette génération fait évoluer les usages : accès haut débit à l'internet sans fil, jeux en ligne, visiophonie, messages vidéo, la réception de la télévision sur le téléphone, etc. La particularité de la technologie 3G est d'avoir un réseau cœur IP. [36] [37]

Le système UMTS exploite deux techniques d'accès : W-CDMA (basée sur le mode FDD) qu'est utilisé dans la division duplex « apparées » et TD-CDMA (basée sur le mode TDD) utilisé dans la division duplex dits « non apparées ».

Le mode FDD est un multiplexage fréquentiel à la manière de FDMA et TDD qui est un multiplexage temporel à la manière de TDMA. Les deux modes utilisent un accès CDMA à 3.84 Mbit/sec jusqu'à 2Mbit/sec sur des bandes de 5 MHz. [36] [37]

Les mobiles de la même cellule se partagent un canal radio par des techniques d'étalement de spectre, le système alloue un code unique à chaque client.

La modulation utilisé est BPSK (Uplink) QPSK (downlink). Il est compatible avec tous les réseaux du monde du fait de la possibilité de roaming au niveau mondial (roaming permet aux utilisateurs

de pouvoir appeler et être appelés, envoyer et recevoir des SMS, accéder à leurs services Internet/Intranet en étant pris en charge par un autre opérateur). [36] [37]

La figure 1.5 illustre la bande de fréquence utilisée en UMTS :

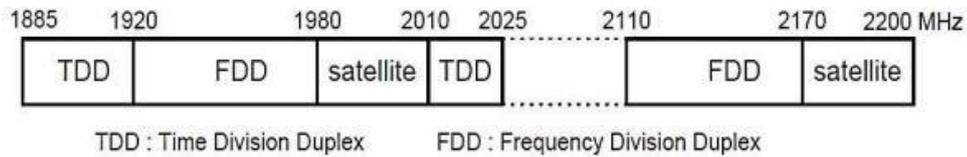


Figure I.5 : division de bande de fréquence. [36]

### I.4.1 Architecture globale du réseau UMTS

La figure suivante illustre Architecture globale du réseau UMTS :

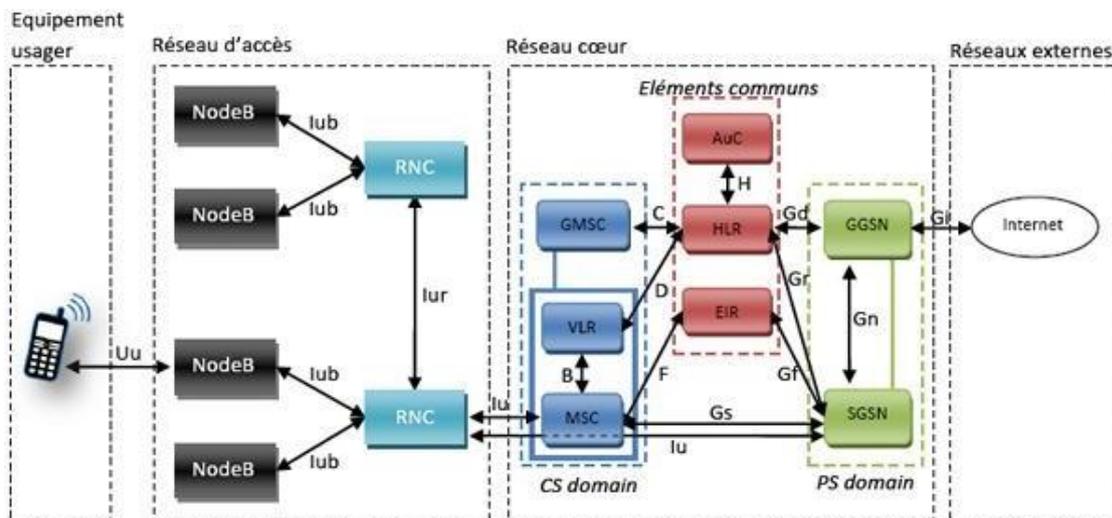


Figure I.6 : Architecture globale du réseau UMTS.

#### I.4.1.1 Réseau cœur CN (Core Network)

Le réseau cœur qui est la partie du système chargée de la gestion des appels. Il permet aux abonnés de communiquer à l'intérieur d'un même réseau de téléphonie mobile et assure l'interconnexion de ce dernier avec des réseaux externes, fixes ou mobiles, analogiques ou numériques. [36]. Le réseau cœur est composé de trois domaines :

##### a. Domaine à commutation de circuits (CS Circuit Switching)

Le mode circuit permettra de gérer les services temps réels dédiés aux conversations téléphoniques (vidéo-téléphonie, jeux vidéo, application multimédia). Ces applications nécessitent un temps de transfert rapide. [36]. L'infrastructure s'appuiera alors sur les principaux éléments du réseau GSM :

- MSC/VLR (bases données existantes)
- G-MSC afin d'avoir une connexion directe vers le réseau externe.

### b. Domaine à commutation de paquet (PS Packet Switching)

Le mode paquet permet de gérer les services non on temps réels. Il s'agit principalement de la navigation sur l'internet, et de l'accès/utilisation des e-mails. Ces applications sont moins sensibles au temps de transfert, c'est la raison pour laquelle les données transiteront mode paquet. [36]

L'infrastructure s'appuiera alors sur les principaux éléments du réseau GPRS[36] :

- **SGSN (Serving GPRS Support Node)** : il envoie les paquets entrants et sortants de/vers un MS. Il fournit aussi le chiffrement et l'authentification et la gestion des sessions. Il est connecté au HLR, MSC
- **GGSN (Gateway Serving GPRS Support Node)** : passerelle s'interfaçant avec les autres réseaux de données et les autres réseaux publics ou privés, il est notamment chargé de fournir une adresse IP aux terminaux mobiles pendant toute la durée de la connexion.

### c. Éléments communs

Sont des éléments partagés par le domaine de commutation de paquet et le domaine de commutation de circuit, ces éléments sont : HLR, EIR et AuC. [37]

#### I.4.1.2 Réseau d'accès radio (UTRAN)

Son rôle principal acheminer les informations depuis l'utilisateur jusqu'au réseau cœur, il est chargé d'autres fonctions telles que : Sécurité, Mobilité, Gestion des ressources radio et Le contrôle du Handover. [36]. Le réseau d'accès UTRAN est composé de [36] :

- **Node B** : Un équipement équivalent à la BTS des réseaux GSM. Son rôle principal est d'assurer les fonctions de réception et de transmission radio pour une ou plusieurs cellules de l'UTRAN.
- **Radio Network Controller (RNC)** : Un équipement équivalent au BSC des réseaux GSM, il s'occupe principalement du routage des communications entre le NodeB et le réseau cœur d'une part, possède et contrôle les ressources radio des Node B auxquels il est connecté d'autre part.

#### I.4.1.3 Station mobile (MS)

- **Équipement mobile (ME)** : L'équipement mobile est chargé de la transmission de l'information vers le réseau d'accès à travers l'interface radio, ou les traiter en réception.
- **Universal Subscriber Identity Module (USIM)** : La carte USIM assure la sécurité du terminal et la confidentialité des communications, l'authentification et le chiffrement de la souscription en utilisant Des algorithmes de cryptages à clés publiques. [37]

### **I.4.2 Évolutions de l'UMTS**

Le HSPA (High Speed Packet Access) a été introduit par le 3GPP dans les Releases 5 en 2002 et 6 en 2005 et évalué à 3.5G (nommé aussi 3G+), regroupe :

- HSDPA (High Speed Uplink Packet Access) pour la voie descendante.
- HSUPA (High Speed Downlink Packet Access) pour la voie montante.

Techniquement, le HSPA utilise la capacité non utilisée par l'UMTS. En d'autres termes, il suffit de rehausser le logiciel sur les contrôleurs afin de permettre aux relais de supporter le Nouveau protocole de transmission de paquets ou la station de base décide dynamiquement du partage des ressources entre les UE actifs. La mise à jour de cette évolution ne nécessite pas d'intervention sur les équipements ni sur l'infrastructure de l'UMTS.

I-HSPA (internet-HSPA) Nommé aussi HSPA+, parfois H+ ou encore 3G++, normalisé par le 3GPP au cours de l'année 2007 dans les Releases 7 et 8. Il permet l'amélioration des débits et de la capacité en introduisant de nouvelles techniques comme MIMO (Multiple Input Multiple Output), opérant dans les bandes de fréquences 900/2100 MHz.

La 3G a donc marqué l'avènement du haut débit mobile en proposant sa première application multimédia en mobilité. Un nouvel objectif est de pouvoir aller bien au-delà de cette proposition initiale, d'où l'apparition d'une nouvelle génération. [36]

## **I.5 Réseaux cellulaires de quatrième génération**

La technologie 4G LTE (en réalité est une norme dite 3.9G), représente une évolution dans le domaine de réseaux mobiles, basé sur GSM et l'UMTS/HSPA (3G/3G+) pour différents objectifs : améliorer le support des services de données via une capacité accrue, une augmentation des débits et une réduction de la latence. En complément de ces exigences de performance, le 3GPP a aussi défini des prérequis fonctionnels tels que la flexibilité spectrale et la mobilité avec les autres technologies. [6]

LTE utilise des bandes de fréquence d'une largeur pouvant varier de 1.4MHz à 20 MHz selon les pays avec : une structure tout IP (Internet Protocol), des modes de duplexage FDD TDD et un débit de données élevés (100Mbit/s descendant et 50 Mbit/s montant). [34]

### **I.5.1 Architecture de réseau LTE**

L'EPS (Evolved Packet System) est l'architecture de LTE/4G se divise en trois parties essentielles à savoir la partie UE, partie radio appelée E-UTRAN et un réseau cœur qui assure le routage comme illustrée par la figure suivante [34].

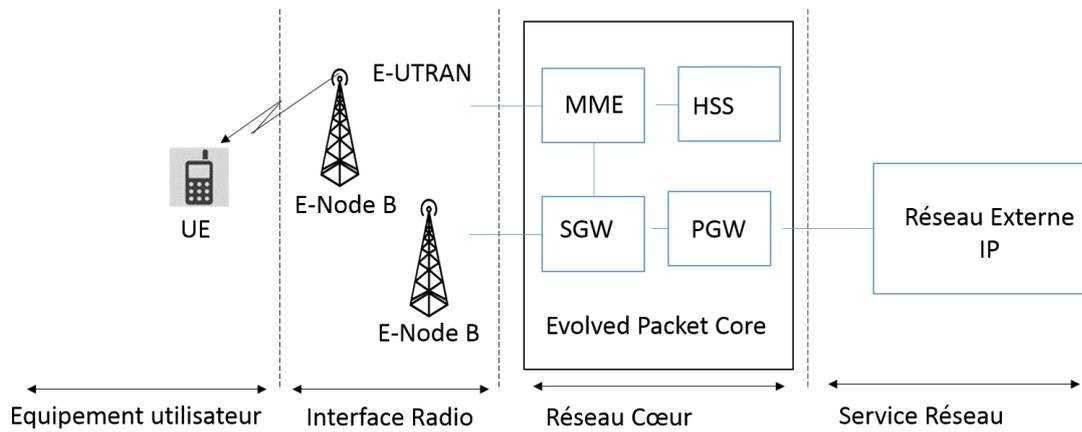


Figure I.7 : Architecture Globale d'un réseau LTE.

### I.5.1.1 E-UTRAN (Evolved UTRAN)

Cette partie contient l'E-Node B qui est responsable de la transmission et de la réception radio entre l'UE et EPC, IL regroupe les entités NodeB et RNC de l'UTRAN 3G. Les fonctions supportées par le RNC ont été réparties entre E-Node B et les entités du réseau cœur MME/SGW. [34]

### I.5.1.2 Réseau cœur EPC (Evolved Packet Core)

Connu aussi sous le nom de SAE (System Architecture Evolution). Ce réseau cœur opère avec un seul domaine de paquet, qui offre tous ces services en IP. Il est composé de [34]:

- **MME (Mobility Management Entity)** : cette entité est responsable de la gestion de la mobilité, l'authentification des utilisateurs, et aussi de la diffusion des messages de paging (quand l'UE est dans l'incapacité de recevoir les paquets qui lui sont destinés). Prendre en charge la mise à jours des paramètres de localisation de l'UE et de gérer leurs sessions (établissement/libération de session de données).
- **Serving GW (Serving Gateway)** : est définie pour gérer les données utilisateurs et il est impliquée dans le routage et la transmission de paquets de données entre les E-UTRAN et le réseau cœur. Il achemine les paquets entre les systèmes 2G/3G et le PGW, assure aussi le Handover (Lors d'une mobilité entre LTE et les réseaux 2G/3G, le SGSN du réseau 2G/3G s'interface avec le Serving GW pour la continuité du service de données sans coupure).
- **La passerelle P-GW (Packet Data Network Gateway)** : elle est définie comme l'interface vers les réseaux externes, fournit à l'UE l'adresse IP, par la suite lorsque le réseau est établi, il l'offre un « default bearer » permanent. [7] [34]
- **Le serveur HSS (Home Subscriber Server)** : HSS se présente comme une version évoluée du registre HLR des réseaux antérieurs, est une base de données qui peut être utilisée simultanément par les réseaux 2G et 3G, LTE et IMS. [34] [7]

### I.5.2 La capacité

La capacité d'une cellule correspond au trafic total maximal qu'elle peut gérer sous une forte charge pendant période précis. Les techniques employés par l'LTE impliquent le support d'un large nombre d'utilisateurs simultanément par cellule. Le débit moyen par utilisateur, en situation de forte charge est approché par la capacité divisée par le nombre d'UE actifs dans la cellule. La capacité d'une cellule est conditionnée par l'efficacité spectrale du système et la ressource spectrale disponible, d'où l'accroissement de cette dernière (en bit/s/ Hz/cellule) En sens descendant, l'efficacité spectrale doit être trois à quatre fois supérieure à celle offerte par le HSPA au sein d'un réseau chargé, et deux à trois fois supérieure en sens montant. [6]

La capacité de canal est donnée par l'équation suivante [8] :

$$C \leq BW \cdot \log_2 \left( 1 + \frac{S}{N} \right) = BW \cdot \log_2 \left( 1 + \frac{E_b \times R}{N_0 \times BW} \right) \quad (I.2)$$

$C$  : est la capacité.

$BW$  : est la bande passante.

$S$  : désigne la puissance du signal reçu et  $N$  désigne la puissance de bruit blanc.

$R$  : Débit de l'information.

$E_b$  : Energie bit.

$N_0$  : Bruit constant.

Le débit d'information ne peut pas dépasser la capacité.

### 1.5.3 La dimension fréquentielle en LTE

L'allocation de ressources en LTE s'effectue à la fois dans les dimensions temporelle et fréquentielle. Les ressources radio en voie descendante et voie montante se présentent ainsi sous la forme d'une grille temps-fréquence représentée dans la figure suivante [35] .

La bande de fréquence totale est découpée :

- Une sous-trame radio complète dure 1ms.
- Chaque sous trame est composée de 2 slots dits PRB (Physical Resource Block) de 0,5 ms.
- Chaque PRB contient 6 ou 7 symboles OFDMA/SC-FDMA dépend de la valeur du préfixe cyclique (CP), soit 7 symbole de 66,7µs considérons un CP normal.
- La largeur de bande d'un PRB est de 12 sous-porteuses de 15 kHz, soit 180 KHz.
- L'élément de base est le RE (Ressource Element) ses dimensions sont de 12 KHz/66.7 µs. A chaque RE correspond un symbole OFDMA/SC-FDMA qui sera encodé en utilisant la sous-porteuse correspondante.
- Un RB est donc composé de 12x7 ou 12x6 REs en considérant un CP normal. [6]

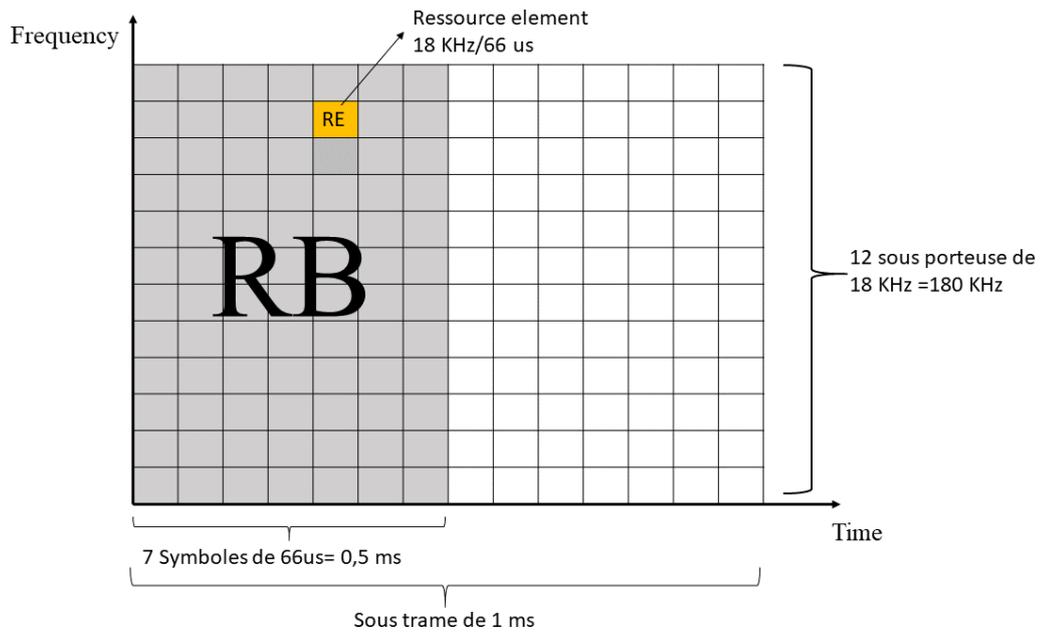


Figure I.8 : La dimension fréquentielle en LTE.

### I.5.4 Techniques d'accès multiple

La modulation du LTE est basée essentiellement sur l'utilisation de la technologie OFDM et des technologies d'accès multiples associés, OFDMA et SC-FDMA.

#### I.5.4.1 Multiplexage par répartition orthogonale de la fréquence (OFDM)

C'est une technique de modulation multi-porteuses à base de transformée de Fourier rapide FFT, qui permet de fragmenter le flux de données de haut débit à transmettre en N sous flux de données parallèles à bas débit, qui seront transmis sur N sous-porteuse orthogonales, différentes et très proche l'une de l'autre. La figure suivante présente la technique OFDM [35] :

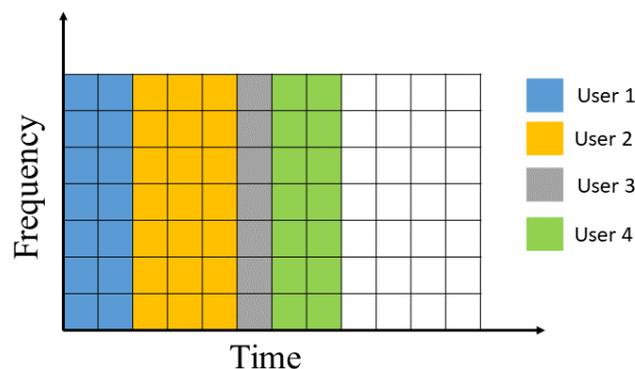


Figure I.9 : Allocation de sous-porteuses en OFDM

Cette subdivision permet d'offrir une grande efficacité au niveau de l'utilisation du spectre et de la puissance, ou Chaque utilisateur a un intervalle de temps durant lequel toutes les sous-porteuses lui sont exclusivement allouées.

L'avantage de ce type d'accès multiple est la réduction de la consommation d'énergie au niveau de récepteur qui ne fonctionne qu'à des instants bien déterminés. [34]

#### I.5.4.2 Accès multiple par répartition en fréquence orthogonale (OFDMA)

L'OFDMA (Orthogonal Frequency Division Multiple Access) basée sur la division en fréquence, assure un partage des ressources fréquentielles entre un nombre variable d'utilisateurs bénéficiant de débits divers pour un time-slot prédéterminé, appelées PRB (Physical Resource Block). Cette technique permet de changer de porteuse pour un usager après une certaine durée OFDMA fournit une facilité et extensibilité de la bande passante par la configuration du nombre de sous-porteuses. Ceci permet à la station de base d'ajuster dynamiquement l'utilisation de la bande passante en fonction des exigences du système. Cette technique utilisée exceptionnellement en LTE DOWNLINK. [35]

La figure suivante présente la technique OFDMA :

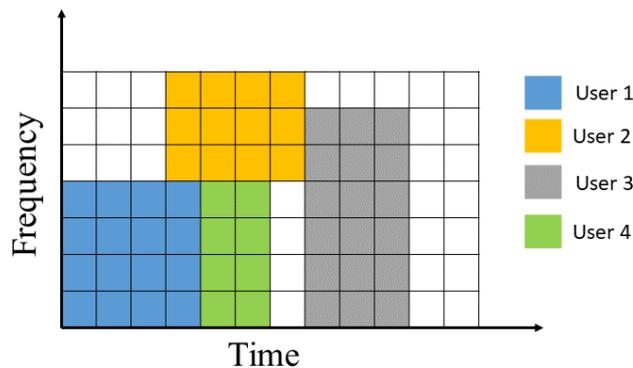


Figure I.10 : Allocation de sous-porteuses en OFDMA

#### I.5.4.3 Single-Carrier FDMA (SC-FDMA)

La SC-FDMA Accès multiple à répartition en fréquences avec porteuse unique aussi appelé DFTS-OFDMA (Direct Fourier Transform Spread-OFDMA), c'est un concept différent utilisé en LTE UPLINK pour les deux modes FDD et TDD, est une technique mono-porteuse consiste à répartir sur un grand nombre de sous-porteuses non pas directement les symboles sources comme en OFDMA, chaque symbole est étalé (pré-coder en utilisant DFT (Direct Fourier Transform)) sur l'ensemble des sous-porteuses allouées.

SC-FDMA est caractérisé par un PAPR (Peak to Average Power Ratio qui est le rapport entre la puissance crête et la puissance moyenne, est une valeur sans dimension exprimer en dB ) plus faible que celui de l'OFDMA car c'est la transmission en parallèle de plusieurs symboles qui le cause, tandis qu'avec la SC-FDMA la transmission est série qui est certes plus lente mais assure

un faible PAPR qui permet de minimiser la consommation énergétique, en diminuant la puissance crête d'émission, aussi de diminuer la consommation électrique. [35]

### **I.5.5. La latence**

La latence d'un système se traduit par sa capacité à réagir rapidement à des demandes de services ou d'utilisateur. Elle est divisée en latence du plan de contrôle et latence du plan usager

#### **I.5.5.1. Latence du plan de contrôle**

La latence du plan de contrôle est définie par un temps de transition entre un état de veille et un état actif de l'UE 80ms en FDD et 85ms en TDD. [6]

#### **I.5.5.2. Latence du plan usager**

La latence du plan usager correspond au délai de transmission d'un paquet IP au sein du réseau d'accès (entre l'UE et un nœud du réseau d'accès). La latence du plan usager doit être inférieure à 5 ms en FDD, en TDD elle dépend de la configuration voie montante/voie descendante, pour certaines configurations la latence est inférieure à 6 ms en voie montante et en voie descendante et pour autres configurations peuvent être inférieure ou égale à 6,2 ms en voie descendante et inférieure ou égale à 9,5 ms en voie montante. [6]

La latence réduite garantit une réponse quasi-instantanée aux requêtes, par exemple pour lancer le chargement d'une page web ou d'une vidéo, elle permet aussi d'accompagner sur les réseaux mobiles des évolutions de l'Internet fixe, comme le développement des services de type cloud computing ou SaaS (Software as à Service). [35]

### **I.5.6. Les Service IP**

Grace à l'architecture de service normalisée IMS (IP Multimedia Sub\_System), plusieurs services sont révélés : services multimédia IPTV, téléphonie sur IP, présence, conférence, messagerie, la vidéo haute définition en 3D...etc.

La VoLTE pour Voice over LTE ou voix sur IP, est un service permettent d'émettre et de recevoir les appels en 4G, sans la VoLTE l'UE passe automatiquement vers le réseaux GSM ou 3G pendant les appels. [34] [6]

### **I.5.7. L'interface Radio LTE**

Dans le cas des communications radio mobiles, le signal est porté par une onde électromagnétique qui se propage dans l'air par un canal hertzien La puissance reçue au récepteur subit plusieurs transformations [35] [6].

#### **I.5.7.1. Interférence**

**de**

**canal**

Le canal intègre des perturbations externes affectant la transmission : l'interférence (créée par d'autres transmissions radio que celle attendue par le récepteur inter-cellule et intra-cellule) et le bruit thermique (provoqué par les équipements électroniques du récepteur). [6]

### I.5.7.2. Path loss (les pertes de propagation)

Traduisent l'atténuation du signal en fonction de la distance entre l'émetteur et le récepteur dans un espace libre sans obstacle selon une loi (1. ), et de l'environnement de propagation (cas de présence d'obstacle), qui engendrent des réflexions, diffractions, diffusions et absorptions de l'onde. [6]

A et B sont des constantes dépendant de l'environnement :

$$P(d) = A + B \cdot \log_{10}(d) \text{ (en dB)} \quad (I.3)$$

Le path loss est caractérisé par le rapport entre la puissance du signal émis sur celle du signal reçu, qui vaut :

$$P_r(d) = \frac{P_t G_r G_t}{PL}$$

Supposons un nœud A ayant les caractéristiques suivantes : gain d'antenne  $G_t$  et puissance émise  $P_t$  désirant émettre une information vers un nœud B distant de  $d$  avec les propriétés suivantes : gain d'antenne  $G_r$ . Le signal émis par A sera reçu par B avec une puissance  $P_r(d)$ .

PL désigne le Path Loss ou l'affaiblissement encouru par la propagation de l'onde.

### I.5.7.3. Les évanouissements rapides (fast fading) :

L'évanouissements suscités par l'existence de trajets multiples due aux obstacle, le signal reçu est une combinaison de plusieurs trajets dont les déphasages, les amplitudes, les décalages Doppler et les retards sont différents.

Au niveau de récepteur les signaux reçus peuvent s'additionner en phase pour donner des évanouissements constructifs, ou en opposition de phase pour donner alors des évanouissements destructifs.

Les retards différents des trajets créent les phénomènes d'interférence entre symboles (IES).

Pour lutter contre l'évanouissement, une technique couramment utilisée est la diversité :

- La diversité spatiale obtenue via plusieurs antennes d'émission et/ou de réception.
- La diversité dans le domaine temporel, est obtenue en recevant le signal via des trajets du canal affecté par les évanouissements rapides indépendants, ou par des retransmissions. [6]
- La diversité en fréquence consiste à transmettre les bits sur différentes porteuses [6]

### I.5.7.4. MIMO

Multiple Input multiple Output MIMO est une combinaison entre deux technique SIMO (single-input and Multiple-output) et MISO (Multiple-input and single-output). MIMO est une ensemble d'antennes émettrices ou bien réceptrices. Il a le potentiel d'augmenter le débit de données et la capacité. Le fonctionnement de cette technique de multiplexage se définit comme suit : le nombre de flux de données doit être inférieur au égal a le nombre des émetteurs, sinon les données ne peuvent pas être entièrement décodées, aussi chaque antenne doivent être identifiée par son signal pilote d'une manière unique pour que le récepteur puisse déterminer quelle combinaison de transmission a été reçue, cette identification utilise des configurations orthogonales pour chaque antenne [9].

Les multi-flux des données sont transmis avec la même fréquence simultanément en prenant des chemins différents dans le canal radio. Chaque antenne de réception devrait recevoir les flux de données transmis par toutes les antennes.

Le canal peut ainsi être décrit par la matrice de canal  $H$  suivante :

$$H = \begin{bmatrix} h_{11} & h_{12} & \square & h_{1N_r} \\ h_{21} & h_{22} & \square & h_{2N_r} \\ \vdots & \vdots & \ddots & \vdots \\ h & h & \square & h \\ \text{---} & \text{---} & \text{---} & \text{---} \\ N_r1 & N_r2 & & N_rN_t \end{bmatrix} \quad (\text{I.4})$$

Où :  $N_t$  Est le nombre d'antenne de transmission.

$N_r$  Le nombre antennes de réceptions.

$h_{ij}$  Le coefficient de canal de l'antenne de transmission  $j$  vers l'antenne de réception  $i$ .

La capacité d'un système MIMO :

$$C = W \log_2 \left( I_N + \frac{P}{M} + H H^* \right) \text{ bit/s/Hz} \quad (\text{I.5})$$

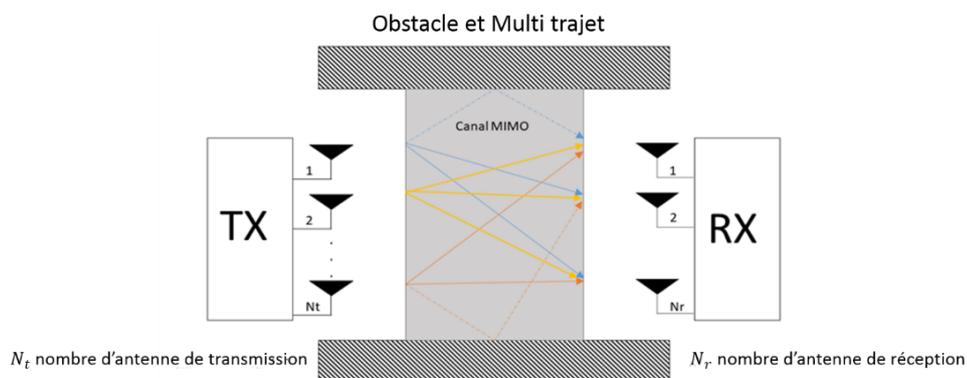


Figure I.11 : Illustration du canal MIMO.

### I.5.7.5. Préfixe cyclique (CP)

Le préfixe cyclique est utilisé dans le cas où l'utilisateur n'est pas en vue directe avec l'antenne, il permet une protection contre les interférences inter-symboles. En effet, le symbole OFDM est allongé avec ce préfixe qui doit être plus grand que le maximum des retards qui apparaissent dans le canal. Si un symbole d'une transmission précédente arrive en retard à cause de la propagation multi-trajets, il entre en collision avec le CP du symbole actuel, et comme la taille du CP est suffisante, cette collision ne peut pas affecter le reste du symbole, où il y a les informations utiles. [8] [34]

La figure suivante présente le principe d'insertion du préfixe cyclique :

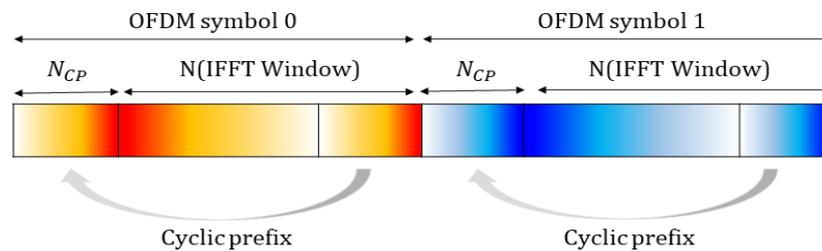


Figure I.12 : Principe d'insertion du préfixe cyclique.

## I.6.Conclusion

Dans ce chapitre, nous avons présenté les concepts des différentes générations de réseaux mobiles qui ont commencé par la première génération analogique et passant au numérique avec la deuxième et la troisième génération pour arriver à la quatrième génération avec leurs architecture [35], technologie et leurs caractéristiques qui s'améliorent afin d'atteindre les besoins des abonnés. Dans le prochain chapitre, nous allons découvrir la nouvelle génération 5G et leur technologie.

---

***CHAPITRE II***

***NEW RADIO 5G***

---

## II.1 Introduction :

L'idée qui commence à se forger derrière la notion de 5G est que celle-ci ne correspondrait pas à une simple augmentation des débits, comme cela a été le cas pour les précédentes générations. Et ce n'est pas seulement la connectivité des gens, mais aussi une connexion pour tous les appareils et toutes sortes d'application, cette connectivité disponible partout et tout le temps pour tout le monde.

Dans le deuxième chapitre, après avoir donné un aperçu d'informations sur les différentes générations, nous nous basons sur 5G qui présente un saut vers les ondes millimétriques et vers une nouvelle architecture pour une meilleure couverture définissant les nouvelles technologies envisagées qui ont une tâche importante dans l'efficacité de transmission de l'information dans ce réseau et qui nous permettrons de résoudre un bon nombre de problème de notre vie moderne.

Dans ce chapitre, nous allons pouvoir vous introduire en premier point le principe de la 5g, découvrir son architecture et ses technologies envisagées ..... etc.

## II.2. Réseaux cellulaire 5G

La 5G est une technologie qui utilise des fréquences plus élevées du spectre par rapport à celle affecté en LTE en utilisant un spectre élargi de 30 à 300 GHz [3] qui représente les ondes millimétriques permettant un débits plus importants, le débit maximum devrait se situer entre 1 et 10 Gbit/s soit 100 à 1000 fois plus rapide que celui de la 4G et une portée plus courte qui varié proportionnellement avec la variation de la fréquence, donné par la relation suivante :

$$\lambda = \frac{c}{f} \quad (\text{II.1})$$

Avec :  $\lambda$  : longueur d'onde,  $c$ : la vitesse de la lumière dans le vide,  $f$  : la fréquence.

La norme 5G a pour le but d'améliorer les débits, la latence (1 ms) et la densité en utilisant une architecture logicielle (Algorithme Adaptatif) pour s'adapter aux différents besoins d'un appareil selon ces trois paramètres (débit, latence et densité). [3]

## II.3. Fréquences attribuées à la 5G

La 5G reprend les technologies déjà utilisées dans la 4G LTE, mais se différencie sur plusieurs points, ce qu'il s'agira d'un nouveau réseau.

En premier lieu, la 5G est une évolution de la 4G LTE, elle peut réutiliser les mêmes bandes de fréquences que cette dernière grâce à la technologie DSS (dynamic Spectrum Switching), cette dernière permet de basculer directement depuis l'antenne entre la 4G et la 5G en temps réel et selon la demande. Grâce à cette mise à jour, un utilisateur 5G peut bénéficier d'un meilleur débit qu'un

utilisateur 4G pourtant connecté à la même antenne, en utilisant la même fréquence ainsi que les mêmes équipements. [13]

La 5G fait également appel à de nouvelles bandes de fréquences divisées en deux groupes, permettant d'augmenter sensiblement le débit au détriment de la portée :

### II.3.1. Les bandes de fréquences 5G Sub-6

Désigne les fréquences utilisées en 5G sous la barre des 6 GHz, soit entre 3,4 et 3,8 GHz en Europe pour la 5G. Ce sont des fréquences permettant une meilleure portée et une meilleure couverture que les ondes millimétriques, mais un débit maximal plus faible. Elles regroupent à la fois les bandes moyennes, qui sont nouvelles, et les bandes basses fréquences, qui réutilisent des fréquences utilisées par la 4G LTE. [13]

### II.3.2. Les ondes millimétriques mmWave

Les ondes radioélectriques allant de 28 GHz jusqu'à 300 GHz avec une longueur d'onde de l'ordre du millimètre. Elles permettent un très haut débit mais une portée limitée (<1Km), une grande sensibilité aux perturbations et une faible capacité à traverser les murs. [13]

La figure suivante présente le spectre de fréquences utilisées :

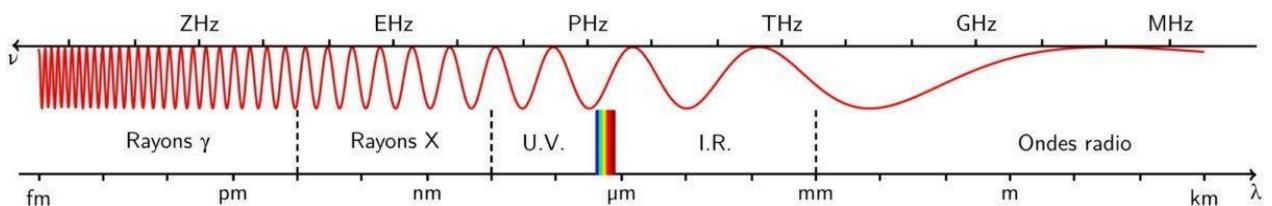


Figure II.1 : spectre de fréquences.

La figure suivante présente Les trois groupes de fréquences utilisés par la 5G NR :



Figure II.2: Les trois groupes de fréquences utilisés par la 5G NR.

## II.4. Architecture du réseau 5G

Un réseau 5G est composé d'un nouveau réseau d'accès radio, et d'un nouveau réseau cœur appelé NGCN ou 5GC (Next Génération Core).

Le réseau d'accès est composé par réseau d'accès radio NG-RAN, (Next Génération RAN) qui utilise une nouvelle interface radio NR et /ou AN non-3GPP (WIFI, XDSL...etc.). [3, 12]

### II.4.1. Le mobile UE (User Equipement)

Le mobile UE (User Equipement) du réseau 5G a la même définition qu'en 3G et 4G. Il est composé d'un USIM, qui contient l'identité de l'abonnée, les fonctions de sécurités et des données d'abonnement. L'abonné est identifié par SUPI (SUBscription Permanent Identifier) qui est une extension de l'IMSI de la 4G, il est aussi identifié à l'aide d'un identifiant temporaire GUTI (Globally Unique Temporary user equipment Identity) attribué localement par le réseau visité et qui est rien d'autre que TMSI dans la 3G. [12]

L'UE ne peut dialoguer avec le 5GC que si 'il supporte le protocole NAS « Non-Access Stratum » « non-standalone Access. L'UE peut établir deux support radio avec deux stations de base c'est la 5G Dual mode, c'est-à-dire que l'appareil est compatible à la fois avec la 5G SA et la 5G NSA. [12]

### II.4.2. L'accès radio 5G (NG-RAN)

Cette partie est constituer de station de base de nouvelle génération. Les UEs communiquent avec la station de base soit par un lien radio 5G, soit par un lien radio 4G.

Si la communication est en 5G, la station de base se nomme gNB (next generation node base station). Si la communication en 4G, la station de base se nomme eNB évoluée pour s'interconnecter avec le réseau cœur 5G nommé ng-eNB (next generation eNB). [14]

Les fonctions de la gNB sont assez similaires avec l'entité eNB. Cependant, les différences concernent la gestion de la qualité de service par flux et non par support et la gestion des tranches de réseau (Slices) sur l'interface radio. Un slice est composé d'instances logique du réseau mobile permettant de garantir la sécurité, les performances et la qualité de service propres à chaque usage. [14] [12]

L'architecture de la gNB est devisé selon des fonctionnalités, elle est formée par une unité centrale (gNB-CU central unit) qui nécessite une ou plusieurs unités distribuées (gNB-DU distributed units) connecter à une carte radio distante (RRH remote radio head). L'unité centrale (gNB-CU) est encore divisée en deux parties, une pour les fonctions de plan de contrôle (gNB-CU-CP control plane functions) et l'autre pour les fonctions de plan d'utilisateur (gNB-CU-UP user plane functions), comme la figure montre ci-dessous : [12]

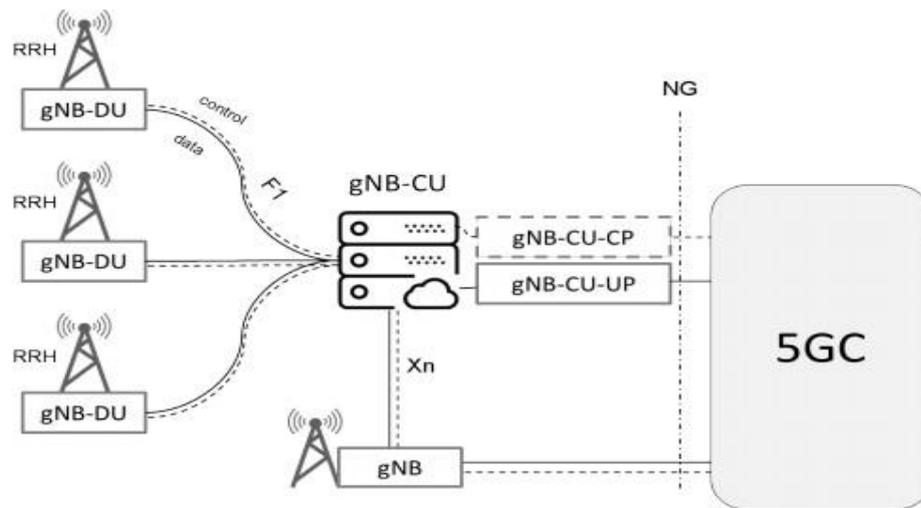


Figure II.3 : Accès radio 5G (NG-RAN)

### II.4.3. Réseau cœur 5g (NGCN ou 5GC)

Le réseau cœur 5GC est adapté pour la virtualisation des fonction réseau aidé par l'utilisation des protocoles Http entre les entêtées fonctionnelles du réseau cœur 5G et le serveur. Il s'appuie sur le découpage du plan de contrôle (Control Plane) et du plan utilisateur (User Plane). Les RANs (Radio Access Networks) dialogue avec le réseau cœur 5G (NGCN ou 5GC) a l'aide des interfaces.

L'architecture 5G est complètement définie en termes de fonctions réseau NF (Network Function) qui sont exposés en tant que services. Dans le plan utilisateur, nous avons une ou plusieurs fonctions de plan utilisateur (UPF), qui portent principalement Transfert de paquets entre les différents tunnels NG-U. Toutes les autres fonctions réseau appartiennent au plan de contrôle. [12]

Les seuls radios autorisés à s'interfacer au 5GC sont LTE et ses évolutions (LTE-Advanced, LTE-Advanced Pro), aussi 5GC donne accès à des réseaux externes IP (ex : Internet, IMS...etc.) appelées Data Networks. [11]

La figure suivante présente l'architecture d'un réseau 5G [12] :

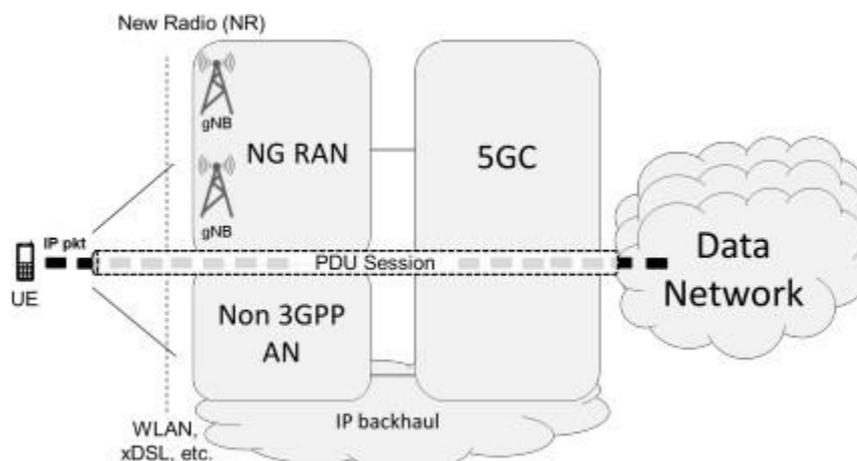


Figure II.4 : architecture d'un réseau 5G.

## II.5. Tranchage (Slicing) de réseau 5G

Le réseau cellulaire de la 5G, est divisé en trois grande catégorie (tranche/slices) qui permettraient de répondre aux besoins métier : [11], [12]

- **mMTC ( Massive Machine Type Communications)** : regroupe tous les usages lié à l'IoT ( Internet of Things) avec des besoins de qualité de service variés. L'objectif de cette catégorie est de répondre à l'augmentation exponentielle de la densité d'objets connectés. [11]
- **eMBB (Enhanced Mobile Broadband)** : concerne tous les applications et services qui nécessite une connexion en ultra haut débit en outdoor, indoor et en bordure de cellule avec une uniformité de la qualité de service (par exemple visionner des vidéos en ultra haute définition (8K)). [11]
- **uRLLC (Ultra-reliable and Low Latency Communications)** : les communications ultra-fiables pour les besoins critiques avec une très faible latence, pour une réactivité accrue ainsi qu'une garantie très forte ( par exemple la téléchirurgie ).[11]

Pour répondre aux besoins et mettre en œuvre ces trois types d'usage, L'UIT à fonder des indicateurs de performance (KPI – Key performance indicator) pour préciser, quantifier et mesurer les caractéristiques de systèmes 5G [11]:

- Débit crête par utilisateur (Gbit/s) : Les exigences minimales de 5G pour un débit de données maximal sont de 20 Gbit/s dans le sens descendant DL et 10 Gbit/s dans le sens montant UL ;
- Débit moyen perçu par l'utilisateur (Mbit/s) : fixé à 100 Mbit/s en DL et 50 Mbit/s en UL.
- Efficacité spectrale (bit/Hz) : Définie comme le débit de données maximal dans des conditions idéales normalisées par la bande passante du canal, en bit/s/Hz, fixé à 30bit/s/Hz en DL et 15bit/s/Hz en UL.
- Vitesse maximale des terminaux (km/h) : C'est la vitesse maximum à laquelle les terminaux peuvent se déplacer en gardant toujours Une connexion stable même en mobilité, cette dernière est classée selon la vitesse considérée par exemple : à pied 0km/h à 10km/h et en véhicule 10km/h à 120km/h.
- Latence (ms) :la latence unidirectionnelle de bout en bout (E2E) est à 4 ms pour les services eMBB et à 1 ms pour les services URLLC.
- Nombre d'objets connectés sur une zone (quantité d'objets/km<sup>2</sup>) : L'UIT-R a spécifié un objectif de 10<sup>6</sup> de dispositifs par km<sup>2</sup> pour des communications de type machine.
- Efficacité énergétique du réseau : C'est la quantité de bits d'information transmise ou reçue des utilisateurs, par unité de consommation d'énergie du RAN, et du côté des appareils c'est la quantité de bits d'information par unité de consommation d'énergie du module de communication.

## II.6. Les modes de Déploiement du réseau 5G

### II.6.1. Mode NAS (Non-standalone access)

Non-Access Stratum est un protocole qui permet l'échange d'information de contrôle ou de données quel que soit l'accès radio et il s'appuie sur les installations existantes du réseau 4G LTE de l'opérateur tout en ajoutant petit à petit des antennes 5G, et permettre notamment l'utilisation de hautes fréquences en 5G NR afin de monter graduellement en puissance pour proposer des débits plus importants. Utilisant le DSS pour basculer entre les deux réseaux. [15]

### II.6.2. Mode AS (standalone access)

La 5G SA (Access Stratum) est un réseau complet 5G qui dispose de ses propres installations et qu'il comporte d'importantes améliorations en termes de débit, latence et propose de nouveaux services pour les utilisateurs. SA est totalement indépendante du réseau 4G, cela demande des investissements bien plus conséquents, et ne sera donc disponible qu'à long terme. [15]

La figure suivante présente les deux modes de déploiements NAS et AS [12] :

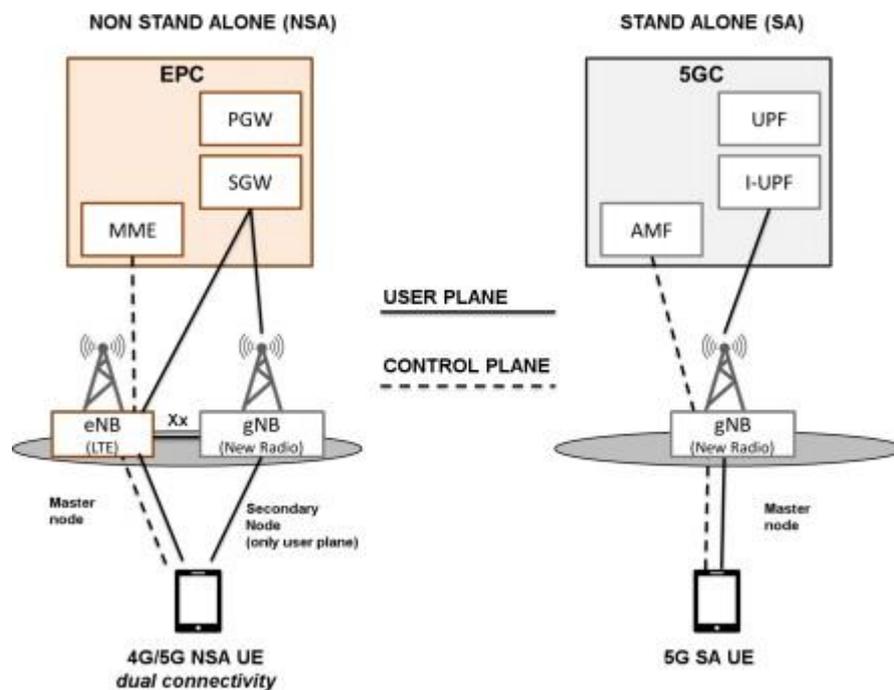


Figure II.5 : les modes de déploiements NAS et AS.

## II.7. Les technologies envisagées par NR

Les bandes millimétriques n'ont jamais encore été prises en compte pour le déploiement des réseaux mobiles, car leur utilisation impose le développement de toutes les technologies nécessaires, à bas coût et avec une consommation énergétique compatible avec des terminaux portables. En particulier, à cause de la faible qualité de propagation des ondes millimétriques,

chaque cellule aura une couverture réduite, ce qui nécessitera la mise en place de nouvelles techniques pour mieux focaliser l'énergie transmise par les antennes [11] :

### II.7.1. Massive MIMO

Massive MIMO (mMIMO) est un mot largement utilisés dans le secteur des télécommunications pour faire référence à la 5G et aux dernières avancées de la technologie LTE. Cette technologie se caractérise par l'utilisation d'un nombre élevé comprendront 100 antennes, de telle façon que le nombre de transmetteurs dépasse le nombre d'utilisateur.

Les systèmes MIMO massif sont extrêmement complexes, mais il offre des avantage importante : propose une efficacité spectrale 10 fois supérieure à celles des systèmes MIMO utilisés pour la 4G [10], il permet de focaliser l'énergie sur un terminal pour améliorer son bilan de liaison (Beamforming),il fournit une augmentation considérable des débits et la qualité des liens de transmission grâce au multiplexage spatiotemporel.[11]

La formation de faisceau utilise plusieurs antennes pour contrôler la direction d'un front d'onde en pesant de manière appropriée l'amplitude et la phase des signaux d'antenne individuels dans un réseau d'antennes multiples. En d'autres termes, le même signal est envoyé par plusieurs antennes disposant d'un espace suffisant entre elles (au moins  $\frac{1}{2}$  longueur d'onde). Quel que soit l'endroit, le récepteur recevra plusieurs copies du même signal.

### II.7.2. Accès multiple non-orthogonal NOMA

Les technologies LTE et 4G (LTE-Advanced) actuelles sont basées sur un schéma OFDM, il est bien connu que cette technique souffre d'un PAPR élevé, en raison de la perte de l'orthogonalité des sous-porteuses et également aux autres déficiences. Ainsi, Pour améliorer l'efficacité spectrale, un schéma non orthogonal, à savoir NOMA (Non-Orthogonal Multiple Access) a été envisagé pour la 5G, elle apporte un attribut supplémentaire qui n'a pas été prise en compte pour différencier les utilisateurs par toute les méthodes d'accès multiple déployer actuellement.

Dans NOMA, les utilisateurs sont multiplexés dans le domaine de puissance formant un codage de superposition, dont la quelle il peuvent transmettre en même temps en utilisant le même code et fréquence mais avec des niveaux de puissance différents. Ce partage de puissance réduit la quantité d'énergie allouée à chaque utilisateur, donc les utilisateurs avec les gains de canal élevés reçoivent moins de puissance par rapport aux ceux des gains de canal inférieurs.

NOMA superpose plusieurs utilisateurs dans le domaine de la puissance formant un codage de superposition au niveau de l'émetteur, tout en permettant la séparation des utilisateurs à l'extrémité de réception via l'annulation des interférences successives SIC (Successive Interference Cancellation). L'UE avec un gain de canal plus élevé soustrait le signal de l'UE avec un gain

inferieur via SIC puis il décode son propre signal, et l'UE avec un gain inferieur considère le signal de l'UE avec un gain élevé comme du bruit et détecte directement son propre signal avec plus des interférences et en lui attribuent plus de puissance. [16]

### II.7.3. Beamforming

Beamforming (aussi appelé filtrage spatial, formation de faisceaux ou formation de voies) est la forme la plus avancée de techniques d'antennes multiples, il peut s'utiliser avec les ondes radio ou sonores et il a de nombreuses applications dans les techniques de radar, transmission sans fil ... etc. Il utilise les mêmes techniques d'antennes et de traitement du signal que MIMO avec une corrélation de sorte que le diagramme de rayonnement de l'émetteur soit orienté intelligemment vers le récepteur. La formation de faisceau utilise plusieurs antennes pour contrôler la direction d'un front d'onde en pondérant de manière appropriée l'amplitude et la phase des signaux d'antenne individuels dans un réseau d'antennes multiples. Ainsi, quel que soit le lieu, le récepteur recevra plusieurs copies du même signal. [17]

L'efficacité de Beamforming est variable selon le nombre d'antenne, l'utilisation de cette application avec le massive MIMO offre des gains et efficacité très importante [12].

Les techniques de Beamforming permettent à la fois : d'étendre une couverture radio (d'une station de base ou d'un point d'accès par exemple), limiter les interférences entre utilisateurs, réduire la pollution électromagnétique environnante (en ciblant le récepteur visé) et la réduction de la latence. Cette technique ne permet pas l'émission en permanence mais seulement quand c'est nécessaire, cela représente donc un gain d'énergie considérable. [18]

### II.7.4. Antenne intelligente

La figure suivante présente une antenne intelligente :

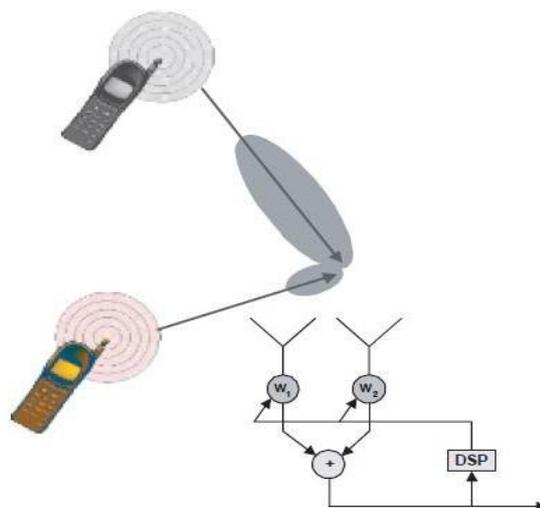


Figure II.6 : Une antenne intelligente électrique à deux éléments.

Un système d'antenne intelligente est généralement constitué par un réseau d'éléments d'antenne équidistant (dont la phase et l'amplitude sont contrôlable) à faible gain dotés d'une capacité de traitement du signal afin d'optimiser automatiquement son rayonnement et / ou son diagramme de réception en fonction de l'environnement de signal. Il peut être visualisé comme une antenne focalisant uniquement vers l'utilisateur de la communication. [8]

Les systèmes d'antenne intelligents, permettant d'améliorer les performances des systèmes en augmentant l'efficacité du spectre en élargissant la zone de couverture, adaptant la mise en forme du faisceau et en orientant plusieurs faisceaux. Le système d'antenne intelligent augmente la capacité de canal à long terme grâce au schéma d'accès multiple par répartition spatiale [6]. En outre, il réduit également les évanouissements par trajets multiples, les interférences entre canaux, le coût de configuration initiale et le taux d'erreur sur les bits (BER).

### II.7.5. Accès multiple à division spatiale

Un concept complètement différent des schémas d'accès multiple décrits précédemment est le schéma d'accès multiple par division spatiale (SDMA). Les systèmes SDMA reposent sur le partage de l'espace en plusieurs régions (par techniques de traitement de l'antenne) permettant d'obtenir un partage spatial du spectre [19]. Cette technique découle directement du concept d'antenne intelligente, son traitement spatial avancé et sa capacité lui permet de localiser de nombreux utilisateurs, créant ainsi des faisceaux différents pour chaque utilisateur, comme indiqué dans la figure suivante.

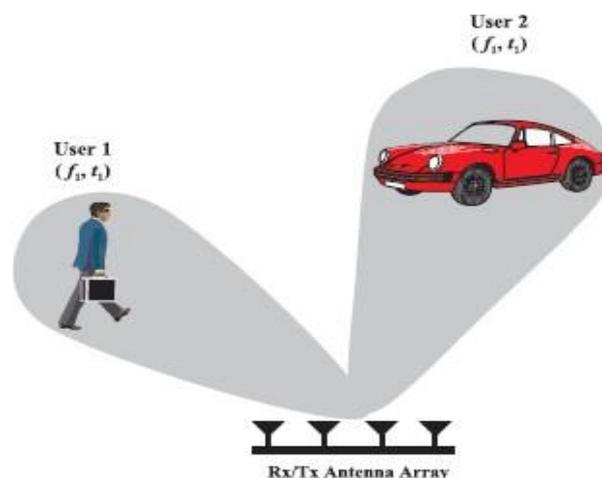


Figure II.7 : Concept SDMA

SDMA est généralement utilisé conjointement avec FDMA, TDMA ou CDMA afin de donner à ce dernier la possibilité supplémentaire d'explorer les propriétés spatiales des signaux. Ce qui permet l'introduction de SDMA dans les systèmes c'est que les modifications nécessaires sont limitées aux stations de base et ne concernent pas les mobiles [19]

La SDMA adaptatif (A-SDMA pour adaptative-SDMA), est une technique dont laquelle les diagrammes des antennes doivent être adaptés afin que le faisceau pointant vers un mobile puisse suivre ses mouvements, cette approche comporte comme avantages : la réutilisation des porteuse (qui implique la bande de fréquence a l'intérieure de la zone de couverture ; plusieurs utilisateurs peuvent donc être desservis par la même bande de fréquences), la réduction de la propagation par multi-trajet et des interférences entre canaux et assure la confidentialité des communications.[19]

### II.7.6. Antenne adaptative

Les antennes adaptatives sont des antennes-réseaux dont les sorties sont combinées pour assurer une commande adaptative de l'émission et de la réception du signal. Ils abordent la communication entre un utilisateur et une station de base d'une manière différente en ajoutant la dimension d'espace selon des stratégies prédéfinies à l'environnement RF, au fur et à mesure que cet environnement évolue, avec traitement du signal qui peuvent modifier leur diagramme de rayonnement de façon dynamique pour s'adapter au bruit, au brouillage et à la propagation par trajets multiples [20]. Les techniques utilisé dans système d'antenne adaptative permettent d'améliorer le rapport signal/bruit + brouillage (SINR) à la réception.

Les systèmes à réseaux adaptatifs [20] offrent davantage de degrés de liberté car ils sont capables d'adapter en temps réel le diagramme de rayonnement à l'environnement du signal RF; En d'autres termes, ils peuvent diriger le faisceau principal vers le signal pilote tout en supprimant le diagramme d'antenne en direction des interféreurs. Pour le dire simplement, les systèmes matriciels adaptatifs peuvent personnaliser un diagramme de rayonnement approprié pour chaque utilisateur. La Figure suivante illustre l'idée générale d'un système d'antenne adaptatif.

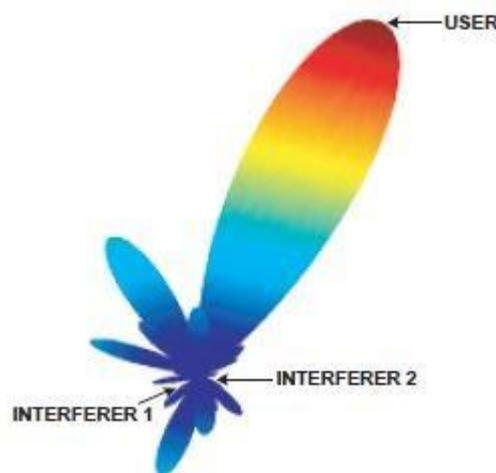


Figure II.8 : Couverture adaptative en réseau

La figure précédente représente un lobe principal s'étendant vers un utilisateur avec des valeurs nulles dirigées vers deux interféreurs dans le même canal.

### II.7.7. Algorithme adaptatif

Un algorithme adaptatif est un algorithme qui est capable de changer automatiquement son comportement en fonction de son contexte d'exécution pour atteindre des performances optimales. Les changements peuvent être sur les données manipulées par l'algorithme, des paramètres de configuration de l'environnement d'exécution et de l'occupation des ressources.

Le terme adaptatif est utilisé pour souligner le fait que le comportement de l'algorithme peut varier de façon importante selon l'environnement.

Pour l'environnement de propagation du signal variant dans le temps, une mise à jour récursive du vecteur de pondération est nécessaire pour suivre le mobile en mouvement, de sorte que le faisceau de filtrage spatial se dirige de manière adaptative vers la direction d'arrivée variable dans le temps du mobile cible, permettant ainsi une transmission / réception optimale du signal [21].

L'algorithme adaptatif calcule les poids complexes appropriés pour diriger le rayonnement maximal du diagramme d'antenne vers le UE et place des valeurs nulles vers les UEI.

Le choix de l'algorithme qui permet d'atteindre la solution optimale est une étape cruciale [9,22], ils sont généralement caractérisés en termes de propriétés de convergence et de complexité de calcul. L'algorithme simple est l'algorithme DMI (Digital multiplexed interface) ; inversion direct de la matrice de convergence [17]. La précision de l'estimation de cette matrice augmente en fonction du nombre d'échantillons de données reçus, ce qui permet de calculer des poids plus précis.

## II.8. Conclusion

Dans ce chapitre nous avons présenté un aperçu sur la cinquième génération des réseaux cellulaires, ces caractéristiques qui s'améliorent afin d'atteindre les besoins des abonnés, sécuriser et assurer une bonne transmission de l'information, augmentation du débit, le nombre d'utilisateurs tel que les techniques d'accès qui diffèrent et se développent et le concept cellulaire parle-on sur l'architecture de base, sa couverture et sa capacité.

Dans le chapitre suivant nous allons définir et déterminer l'apprentissage automatique (machine Learning) qui va nous aider à résoudre un nombre important de problèmes imposés dans la cinquième génération, et qu'il va nous idée à améliorer cette génération encore plus.

---

*CHAPITRE III*

*APPRENTISSAGE AUTOMATIQUE*

---

## III.1 Introduction

Le Machine Learning ou bien l'apprentissage automatique est une branche de l'intelligence artificielle (IA) qui a pour but de donner la possibilité aux ordinateurs d'apprendre de leur environnement. Le but est d'atteindre une conception de machines capables d'effectuer des tâches humaines avec une grande précision.

Ce chapitre présente des concepts de Machine Learning, particulièrement la régression linéaire avec une seule variable, puis avec des variables multiples ainsi que la régression logistique binaire et multi-classes.

## III.2 L'apprentissage automatique (Machine learning)

Deux définitions du Machine Learning sont proposées :

Arthur Samuel l'a décrit comme : "le domaine d'étude qui donne aux ordinateurs la capacité d'apprendre sans être explicitement programmés." Il s'agit d'une définition plus ancienne et informelle. [23]

Tom Mitchell donne une définition plus moderne : « On dit qu'un programme informatique apprend de l'expérience E en ce qui concerne une certaine classe de tâches T et une mesure de performance P, si ses performances aux tâches en T, telles que mesurées par P, s'améliorent avec l'expérience E. [23]

## III.3 Apprentissage supervisé et apprentissage non supervisé

En général, tout problème d'apprentissage automatique peut être attribué à l'une des deux grandes classifications, Apprentissage supervisé et apprentissage non supervisé. [24]

### III.3.1 Apprentissage supervisé

L'apprentissage supervisé est le concept derrière plusieurs applications de nos jours (reconnaissance faciale, filtres anti-spam des emails, etc.), le principe est de recevoir un ensemble de données et savoirs déjà à quoi devrait ressembler notre sortie correcte, ayant l'idée qu'il existe une relation entre l'entrée et la sortie. [24]

Le concept d'apprentissage supervisé est basé sur un algorithme tiré de l'ensemble de données de formation (training set), qui peut être considéré comme un enseignant supervisant le processus d'apprentissage. Connaissant les réponses correctes, l'algorithme effectue des prédictions itératives sur les données d'apprentissage et est corrigé par l'enseignant. L'apprentissage s'arrête lorsque l'algorithme atteint un niveau de performance acceptable [26],

Par exemple elle peut apprendre à reconnaître une photo de chien après qu'on lui ait montré des millions de photos de chiens [25].

Il existe deux problèmes d'apprentissage supervisé classés en problèmes de :

- **Régression**

Un problème de régression, se pose lorsque les résultats dans une sortie continue (une valeur réelle), ce qui signifie que nous essayons de mapper des variables d'entrée à une fonction continue. [26]

- **Classification**

Dans un problème de classification, nous essayons plutôt de prédire les résultats dans une sortie discrète, donc nous essayons de mapper les variables d'entrée en catégories discrètes. [26]

### III.3.2 Apprentissage non supervisé

L'apprentissage non supervisé nous permet d'aborder les problèmes avec peu ou aucune idée à quoi devraient ressembler nos résultats. [26]

Nous pouvons dériver une structure à partir de données dont nous ne connaissons pas nécessairement l'effet des variables, ou les algorithmes sont laissés à leurs propres mécanismes pour découvrir et présenter la structure intéressante des données. [26]

### III.4 Fonction d'hypothèse

On considère un ensemble de test avec  $m$  est le nombre d'occurrences,  $X$  l'espace des valeurs d'entrées et  $Y$  l'espace des valeurs des sorties avec  $X=Y=R$ .

Le principe de base de l'apprentissage automatique est de créer des algorithmes capables de recevoir des données d'entrée et d'utiliser une analyse statistique pour prédire une sortie tout en les mettant à jour à mesure que de nouvelles données deviennent disponibles. la fonction hypothèse  $h_{\theta}$  est de la forme :

$$h_{\theta}(x^{(i)}) = \theta_0 + \theta_1 x^{(i)} \quad \theta_i \text{ comme paramètres} \quad (\text{III. 1})$$

### III.5 Fonction de Coût (Cost Function)

Nous pouvons mesurer la précision de notre fonction d'hypothèse en utilisant une fonction de Coût (Cost)  $J$ , dans le but d'obtenir un meilleur choix pour adapter une ligne droite à nos données : [25]

$$J(\theta_0, \theta_1) = \frac{1}{2M} \sum_{i=1}^M (h_{\theta}(x^{(i)}) - y^{(i)})^2 \quad (\text{III. 2})$$

Il s'agira de choisir  $\theta_0, \theta_1$  de telle manière que  $h_{\theta}(x)$  est proche de  $y$  dans l'ensembles de données  $(X, Y)$ . En synthèse cela revient à Minimiser la différence des carrés entre la sortie de l'hypothèse et la valeur réelle  $(h_{\theta}(x) - Y)^2$  avec le choix  $\theta_1, \theta_2$ . [33]

### Hypothèse 1 : $h_{\theta}(x) = \theta_1 x$

Dans ce cas  $\theta_0 = 0$  donc toutes les fonctions passent par l'origine  $O(0,0)$

- paramètres :  $\theta_1$
- fonction de cout :  $J(\theta_1) = \frac{1}{2M} \sum_{i=1}^M (h_{\theta}(x^{(i)}) - y^{(i)})^2$
- Objectif : minimiser  $\theta_1, J(\theta_1)$

En termes visuels, notre ensemble de données d'entraînement est dispersé sur le plan x-y. Nous essayons de faire une ligne droite définie par  $h_{\theta}(x)$  qui passe par ces points de données dispersés. Idéalement, la ligne doit passer par tous les points de notre ensemble de données d'entraînement. Dans un tel cas, la valeur de  $J(\theta_0, \theta_1)$  sera égale à 0. [33]

### Hypothèse 1 : $h_{\theta}(x) = \theta_0 + \theta_1 x$

- paramètres :  $\theta_0, \theta_1$
- fonction de cout :  $J(\theta_0, \theta_1) = \frac{1}{2M} \sum_{i=1}^M (h_{\theta}(x^{(i)}) - y^{(i)})^2$
- Objectif : minimiser  $\theta_0, \theta_1, J(\theta_0, \theta_1)$

Un tracé de contour est un graphique qui contient de nombreuses lignes de contour. Une ligne de contour d'une fonction à deux variables à une valeur constante en tous les points de la même ligne. [33]

## III.6 La Descente de gradient

L'un des algorithmes les plus importants de tout le Machine Learning c'est la descente de Gradient (Gradient Descent). Il s'agit d'un algorithme d'optimisation puissant qui permet d'entraîner les modèles de régression linéaire, régression logistiques ou encore les réseaux de neurones. Il permet aussi de trouver le minimum de n'importe quelle fonction convexe en convergeant progressivement vers celui-ci [25]. La formule de la descente de Gradient est la suivante :

$$\theta_j = \theta_j - \sigma \frac{\partial J(\theta_j)}{\partial \theta_j} \quad \text{avec : } \sigma \in (0, 1) \quad (\text{III.3})$$

On utilise l'algorithme de la descente de gradient dans les problèmes d'apprentissage supervisé pour minimiser la fonction de cout.

Lorsqu'elle est spécifiquement appliquée au cas de la régression linéaire, une nouvelle forme de l'équation de descente de gradient peut être dérivée. Nous pouvons remplacer notre fonction de coût réel et notre fonction d'hypothèse réelle et modifier l'équation pour :

$$\theta_0 \rightarrow \theta_0 - \alpha \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)}) \quad (\text{III. 4})$$

$$\theta_1 \rightarrow \theta_1 - \alpha \frac{1}{m} \sum_{i=1}^m ((h_{\theta}(x^{(i)}) - y^{(i)}) x^{(i)}) \quad (\text{III. 5})$$

$m$  : est la taille de l'ensemble d'apprentissage.

$\theta_0$  : constante qui changera simultanément avec  $\theta_1$ .

$x^{(i)}, y^{(i)}$  : sont les valeurs de l'ensemble d'apprentissage donné.

$\alpha$  : Taux d'apprentissage (Learning rate).

Si  $\alpha$  est trop grand, alors on aura de trop grands pas dans la descente de gradient. Cela a l'avantage de descendre rapidement vers le minimum de la fonction coût, mais on risque de louper ce minimum en oscillant autour à l'infini. S'il est trop faible, alors on risque de mettre un temps infini avant de converger vers le minimum de la fonction coût. [25]

### III.7 Régression linéaire à variables multiples

À la suite de la régression linéaire simple, on introduit le modèle linéaire multidimensionnel, qu'est l'outil statistique le plus habituellement mis en œuvre pour l'étude de données dans lequel une variable quantitative  $Y$  est expliquée, modélisée, par plusieurs variables quantitatives  $X_i (i = 1, \dots, n)$ , [29]. Elle sert à estimer une fonction linéaire entre la sortie (avec des valeurs continues, numériques) et les entrées.

La forme multivariée de la fonction d'hypothèse tenant compte de ces multiples caractéristiques est :

$$h_{\theta}(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 + \dots + \theta_n x_n \quad (\text{III. 6})$$

$x_j^{(i)}$  : Valeur de la caractéristique  $j$  dans le  $i$ ème exemple.

$x^{(i)}$  : L'entrée de caractéristique pour  $i$ ème exemple.

$m$  : Nombre des exemples.

$n$  : Nombre de caractéristique.

L'équation de descente de gradient elle-même est généralement de la même forme ; il suffit de le répéter pour nos fonctionnalités ' $n$ ' :

$$\theta_j \rightarrow \theta_j - \alpha \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)}) \cdot x_j^{(i)} \quad (\text{III. 7})$$

La régression linéaire est utile pour l'analyse des données, mais elle donne des mauvaises performances s'il n'y a pas une relation linéaire. La plupart des phénomènes réelles ne correspondent pas la supposition du modèle linéaire. Elle est sensible aux valeurs aberrantes (absurde). [30]

### III.8 Mise à l'échelle des fonctionnalités

De plus, à mesure que les valeurs d'entrées deviennent de plus en plus grandes, les points de données seront tracés de plus en plus loin. Cela augmente non seulement la distance entre les points mais aussi en raison des grandes valeurs, le temps de calcul de l'algorithme est plus grand. [27]

Lorsque les valeurs d'entrée sont dans la même plage ; nous pouvons accélérer la descente de gradient d'où  $\theta$  descendra rapidement sur de petites plages, lentement sur de grandes plages et oscillera de manière inefficace jusqu'à l'optimum lorsque les variables sont très inégales. [28]

Nous pouvons accélérer la descente de gradient en ayant chacune de nos valeurs d'entrée à peu près dans la même plage. Ceci est dû au fait que  $\theta$  descendra rapidement sur de petites plages et lentement sur de grandes plages, et oscillera donc de manière inefficace jusqu'à l'optimum lorsque les variables sont très inégales. [28]

### III.9 Les différentes techniques de mise à l'échelle (Feature Scaling)

Les deux méthodes les plus courantes de mise à l'échelle des fonctionnalités sont :

#### III.9.1 Normalisation

Transforme les entités avec différentes échelles en une échelle fixe de 0 à 1. Cela garantit qu'aucune caractéristique spécifique ne domine l'autre.

Le but d'avoir un tel intervalle restreint est de réduire l'espace de variation des valeurs. [27]

$$X_{normalise} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (\text{III. 8})$$

$X_{max}$ : la plus grande valeur observée pour la feature X

$X_{min}$ : la plus petite valeur observée pour la feature X

X: la valeur de la feature qu'on cherche à normaliser

Les valeurs min et max ne sont apprises qu'à partir des données d'entraînement, donc un problème survient lorsqu'une nouvelle donnée a une valeur de  $x$  qui est en dehors des limites des valeurs min et max, la valeur  $X$  résultante ne sera pas dans la plage de 0 et 1. [27]

### III.9.2 Standardisation

Transforme les entités de telle sorte que leur moyenne ( $\mu$ ) est égale à 0 et l'écart type ( $\sigma$ ) égal à 1. La plage des nouvelles valeurs min et max est déterminée par l'écart type de l'entité non normalisée initiale. [27]

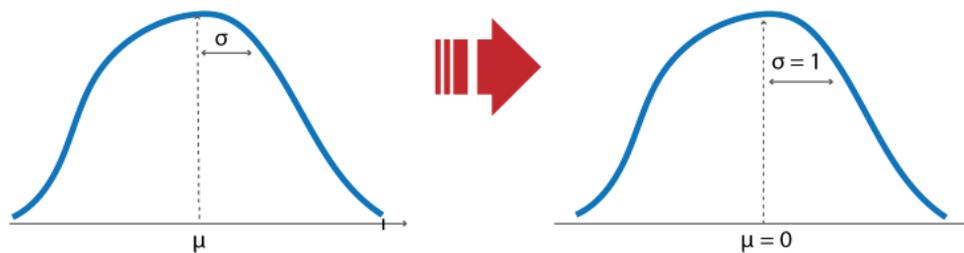


Figure (III.1) : Avant et après standardisation.

$\mu = 0$  la moyenne de la loi de distribution

$\delta = 1$  est l'écart – type (Standard Deviation)

$$x_i = \frac{x_i - \mu_i}{s_i} \quad (\text{III. 9})$$

Où :  $x$  est La valeur qu'on veut standardiser (input variable)

$\mu$ : La moyenne des observation pour cette feature

$\delta$  : L'écart – type (Standard Deviation)des observations pour cette feature

Peut-elle être appliquée quand les caractéristiques d'entrée ( Input Features ) répondent à des distributions normales (Distributions Gaussiennes) avec des moyennes et des écart-types différents. Par conséquent, cette transformation aura pour impact d'avoir toutes nos caractéristiques (Features) répondant à la même loi normale. La standardisation peut également être appliquée quand les caractéristiques (Features) ont des unités différentes. [24].

### III.10 Régression polynomiale

La régression polynomiale est un cas spécial de la régression linéaire. Nous pouvons combiner plusieurs fonctionnalités en une seule. Par exemple, nous pouvons combiner  $x_1$  et  $x_2$  dans une nouvelle fonctionnalité  $x_3$  en prenant  $(x_1.x_2)$ .

La fonction d'hypothèse n'a pas besoin d'être linéaire, on peut changer le comportement ou la courbe de notre fonction en en faisant une fonction quadratique, cubique ou racine carrée (ou toute autre forme).

Par exemple, notre fonction d'hypothèse est :

$$h_{\theta}(x) = \theta_0 + \theta_1 x_1 + \theta_2 x_1^2 + \theta_3 x_1^3 \quad (\text{III. 10})$$

$$x_2 = x_1^2, x_3 = x_1^3.$$

La régression polynomiale fournit une bonne approximation de la relation entre la variable dépendante  $y$  et la variable indépendante  $x$ . [30]

### III.11 Équation normale

Une deuxième façon de faire minimiser  $J$ , cette fois en effectuant la minimisation explicitement en prenant ses dérivées par rapport aux  $\theta_j$  et en les mettant à zéro. Cela nous permet de trouver le  $\theta$  optimal sans itération. La formule de l'équation normale est donnée ci-dessous :

$$\theta = (X^T X)^{-1} X^T y \quad (\text{III. 11})$$

Avec l'équation normale, le calcul de l'inversion a une complexité, Donc, si nous avons un très grand nombre de fonctionnalités, l'équation normale sera lente.

#### III.11.1 Non-inversibilité de l'équation normale :

Si  $X^T X$  est non inversible, les causes courantes pourraient être:

- Fonctions redondantes : où deux caractéristiques sont très étroitement liées (c'est-à-dire qu'elles sont linéairement dépendantes) ex : écrire  $x_1$  en fonction de  $x_2$ .
- TROP de fonctionnalités.

Les solutions aux problèmes ci-dessus incluent la suppression d'une fonctionnalité qui dépend linéairement d'une autre, ou la suppression d'une ou plusieurs fonctionnalités lorsqu'il y a trop de fonctionnalités.

### III.12 Régression logistique

La régression logistique est l'un des modèles de classification et d'analyse multivariée les plus couramment utilisés. Elle est considérée comme une méthode de régression puisqu'elle sert à estimer la probabilité d'appartenir à une classe. La régression logistique permet d'étudier les relations entre un ensemble de variables qualitatives  $X_i$  et une variable qualitative  $Y$ .

Il y a trois types de régression logistique :

- **Régression logistique binaire** : le but de la classification est d'identifier si un échantillon appartient à une classe ou non.

- **Régression logistique multinomiale** : le but de la classification est d'identifier à quelle classe appartient-t-il un échantillon parmi plusieurs classes.
- **Régression logistique ordinale** : le but de la classification est de chercher la classe d'un échantillon parmi des classes ordonnées. Un exemple de classes : non satisfait, satisfait, très satisfait.

La classification binaire prend un petit nombre de valeurs discrètes dans lequel  $y$  ne peut prendre que deux valeurs, 0 et 1.

### III.12.1 Principe

Considérons une entrée  $X = x_1 \ x_2 \ x_3 \ \dots \ x_n$ , la régression logistique a pour objectif de trouver une fonction  $h$  telle que nous puissions calculer :

$$y = \begin{cases} 1 & \text{si } h_{\theta}(x) \geq \text{seuil} \\ 0 & \text{si } h_{\theta}(x) < \text{seuil} \end{cases} \quad (\text{III. 12})$$

$$\text{on a :} \quad h_{\theta}(x) = g(\theta^T x) \quad (\text{III. 13})$$

$$z = \theta^T x \quad (\text{III. 14})$$

$$g(z) = \frac{1}{1+e^{-z}} \quad (\text{III. 15})$$

L'image suivante nous présente la fonction sigmoïde

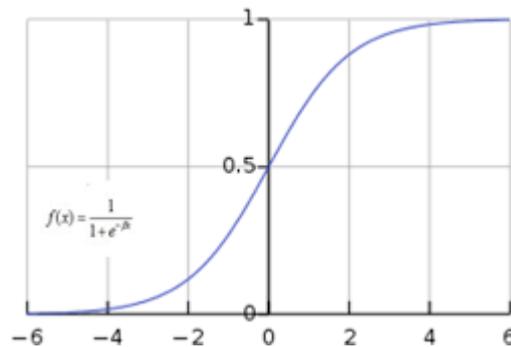


Figure (III.2) : Fonction sigmoïde.

Cette valeur est transformée à une probabilité en utilisant la fonction logistique. Donc, la probabilité qu'un échantillon avec les caractéristiques  $x_1, \dots, x_n$  appartienne à une classe  $y$  est calculée tq :

$$h_{\theta}(x) = p(y = 1 | x; \theta) = 1 - p(y = 0 | x; \theta) \quad (\text{III. 16})$$

$h_{\theta}(x)$  Donne la probabilité que la sortie soit 1

### III.12.2 La décision

On calcule la probabilité en utilisant l'équation et on utilise un seuil sur cette probabilité pour décider et prédire si un échantillon  $x$  appartient à une classe donnée (classe positive)  $y=1$ .

$$h_{\theta}(x) \geq 0.5, y = 1 \quad (\text{III. 17})$$

$$h_{\theta}(x) < 0.5, y = 0 \quad (\text{III. 18})$$

$$g(z) \geq 0.5 \text{ lorsque } z \geq 0 \quad (\text{III. 19})$$

La limite de décision est la ligne qui sépare la zone où  $y=0$   $y=1$ . Elle est créée par notre fonction d'hypothèse.

En cas de plusieurs classes, on utilise une stratégie d'un-contre-le-reste. On entraîne plusieurs classifieurs, chacun pour une classe. Pour décider quelle est la classe d'un échantillon, on prend celle avec la probabilité la plus élevée.

### III.13 Fonction du coût

On n'utilise pas la même fonction de coût comme celle utilisée dans la régression car on a la fonction de prédiction est non linéaire. La fonction du coût va être non-convexe avec plusieurs minimums locaux. Lors de la minimisation, on peut tomber sur un minimum local et l'algorithme du gradient va s'arrêter sans converger vers la solution optimale.

Dans ce cas, on utilise l'entropie croisée. Étant donnée un ensemble de données avec  $N$  échantillons, où le résultat  $y$  est soit 1 ou 0. La fonction du coût est calculée comme suit, où  $(i)$  réfère au  $i$ -ème échantillon dans les données d'entraînement:

$$j(\theta) = \frac{1}{m} \sum_{i=0}^m \text{cost}(h_{\theta}(x^i), y^i) \quad (\text{III. 22})$$

$$\text{cost}(h_{\theta}(x), y) = -\log h_{\theta}(x) / , \text{ si } y = 1 \quad (\text{III. 21})$$

$$\text{cost}(h_{\theta}(x), y) = -\log(1 - h_{\theta}(x)) / , \text{ si } y = 0 \quad (\text{III. 22})$$

On peut regrouper les deux cas de notre fonction de coût dans un seul cas :

$$\text{cost}(h_{\theta}(x), y) = -y \log h_{\theta}(x) / - (1 - y) \log(1 - h_{\theta}(x)) / \quad (\text{III. 23})$$

$$j(\theta) = -\frac{1}{m} \sum_{i=0}^m y^{(i)} \log h_{\theta}(x^{(i)}) / + (1 - x^{(i)}) \log(1 - h_{\theta}(x^{(i)})) / \quad (\text{III. 24})$$

La fonction de gradient descente est comme suivant :

$$\theta_j \rightarrow \theta_j - \alpha \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x_j) - y_j) x_j^{(i)} \quad (\text{III. 25})$$

Le gradient de chaque poids  $\theta_j$  est calculé en utilisant le dérivé partiel de la fonction du coût par rapport à ce poids. Donc, le gradient d'un poids  $\theta_j$  est calculé comme suit :

$$\frac{\partial j}{\partial \theta_j} = \frac{1}{m} \sum_{i=1}^m (h_{\theta_j}(x_j) - y_j) x_j^{(i)} \quad (\text{III. 26})$$

- Elle est utilisée pour la classification, elle donne des probabilités pour les sorties.
- Les modèles logistiques peuvent être mis à jour facilement

Donne une mauvaise performance lorsqu'il existe plusieurs limites de décision ou des limites de décision non linéaires

### III.14 Classification multi classe

Nous allons maintenant aborder la classification des données lorsque nous avons plus de deux catégories. Au lieu de  $y = \{0,1\}$ , nous allons étendre notre définition de sorte que  $y = \{0,1\dots n\}$ , Cette méthode est appelée one-versus-rest OVR « une contre le reste », ou « une contre toutes ». [31]

Puisque  $y = \{0,1\dots n\}$ , nous divisons notre problème en  $n+1$  problèmes de classification binaire ; dans chacun, nous prédisons la probabilité que 'y' soit membre de l'une de nos classes.

$$y \in \{0, 1, \dots, n\}$$

$$h_{\theta}^{(0)}(x) = P(y = 0|x; \theta) \quad (\text{III. 27})$$

$$h_{\theta}^{(1)}(x) = P(y = 1|x; \theta) \quad (\text{III. 28})$$

...

$$h_{\theta}^{(n)}(x) = P(y = n|x; \theta) \quad (\text{III. 29})$$

$$\text{prediction} = \max_i (h_{\theta}^{(i)}(x)) \quad (\text{III. 30})$$

Nous choisissons essentiellement une classe, puis nous regroupons toutes les autres dans une seule seconde classe. Nous faisons cela à plusieurs reprises, en appliquant une régression logistique binaire à chaque cas, puis en utilisant l'hypothèse qui a renvoyé la valeur la plus élevée comme prédiction. [31]

- Former un classificateur de régression logistique  $h_{\theta}(x)$  pour chaque classe pour prédire la probabilité que  $y = i$ .
- Pour faire une prédiction sur un nouveau  $x$ , nous allons donc simplement prédire la classe pour laquelle la fonction de décision retourne la valeur la plus élevée, donc la classe qui maximise  $h_{\theta}(x)$ .

### III.15 Le problème du sur apprentissage (Overfitting)

L'Overfitting, ou variance très élevée, intervient lorsque l'algorithme sur-apprend (overfit). On observe généralement de l'Overfitting lorsque la fonction d'hypothèse s'ajuste aux données disponibles mais ne se généralise pas bien pour prédire de nouvelles données. Cela est généralement causé par une fonction compliquée qui crée de nombreuses courbes, schémas, structures et angles inutiles sans rapport avec les données pour résoudre des problèmes simples mais bruités. [32]

Cette terminologie est appliquée à la fois à la régression linéaire et à la régression logistique. Il existe deux options principales pour résoudre le problème du surapprentissage :

- Réduire le nombre de fonctionnalités

Sélectionnez manuellement les fonctionnalités à conserver.

Utiliser un algorithme de sélection de modèle.

- Régularisation

Conserver toutes les fonctionnalités, mais réduire l'amplitude des paramètres  $\theta_j$ .

La régularisation fonctionne bien lorsque nous avons beaucoup de fonctionnalités légèrement utiles.

### III.16 fonction de coût

Si nous avons un Overfitting de notre fonction d'hypothèse, nous pouvons réduire le poids que certains termes de notre fonction en augmentant leur coût. Sans réellement changer la forme de notre hypothèse, nous pouvons à la place modifier notre fonction de coût en régulariser tous nos paramètres  $\theta$  en une seule sommation comme :

$$\min_{\theta} \frac{1}{2m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)})^2 + \frac{\lambda}{2} \sum_{j=1}^n \theta_j^2 \quad (\text{III. 31})$$

Le  $\lambda$ , ou lambda, est le paramètre de régularisation. Il détermine de combien les coûts de nos paramètres  $\theta$  sont gonflés.

En utilisant la fonction de coût ci-dessus avec la sommation supplémentaire, nous pouvons lisser la sortie de notre fonction d'hypothèse pour réduire le surajustement.

### III.17 La régularisation

Nous pouvons appliquer la régularisation à la fois à la régression linéaire et à la régression logistique. Nous aborderons d'abord la régression linéaire.

#### III.17.1 Régression linéaire régularisée

##### III.17.1.1 Descente graduelle

Nous allons modifier notre fonction de descente de gradient pour séparer  $\theta_0$  du reste des paramètres car nous ne voulons pas pénaliser  $\theta_0$ .

$$\theta_0 = \theta_0 - \alpha \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)}) x_0^{(i)} \quad (\text{III. 32})$$

$$\theta_j = \theta_j - \alpha \left( \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)}) x_j^{(i)} + \frac{\lambda}{m} \theta_j \right) \quad j = 1, 2, \dots, n \quad (\text{III. 33})$$

Le terme  $\frac{\lambda}{m} \theta_j$  effectue notre régularisation. Avec quelques manipulations, notre règle de mise à jour peut également être représentée par :

$$\theta_j := \theta_j \left(1 - \alpha \frac{\lambda}{m}\right) - \alpha \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)}) x_j \quad (\text{III. 34})$$

Le premier terme de l'équation ci-dessus,  $1 - \alpha \frac{\lambda}{m}$  sera toujours inférieur à 1. Intuitivement, vous pouvez le voir comme réduisant la valeur de  $\theta_j$  d'un certain montant à chaque mise à jour. Notez que le deuxième terme est maintenant exactement le même qu'avant.

### III.17.1.2 Équation normale

Pour ajouter en régularisation, l'équation est la même que notre originale, sauf que nous ajoutons un autre terme à l'intérieur des parenthèses :

$$\theta = (X^T X + \lambda \cdot L)^{-1} X^T y \quad (\text{III. 34})$$

where  $L = \begin{bmatrix} 0 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & \dots & \\ & & & & 1 \end{bmatrix}$

L est une matrice avec 0 en haut à gauche et 1 en diagonale, avec des 0 partout ailleurs. Il doit avoir la dimension  $(n+1) \times (n+1)$ . Intuitivement, c'est la matrice d'identité (bien que nous n'incluons pas  $x_0$ ), multiplié par un seul nombre réel  $\lambda$ . Rappelons que si  $m < n$ , alors  $X^T X$  est inversible. Cependant, lorsque nous ajoutons le terme  $\lambda \cdot L$ , alors  $X^T X + \lambda \cdot L$  devient inversible.

### III.17.2 Régression Logistique Régularisée

Nous pouvons régulariser la régression logistique de la même manière que nous régularisons la régression linéaire. En conséquence, nous pouvons éviter l'Overfitting.

#### III.17.2.1 Cost function

On peut régulariser cette équation en ajoutant un terme à la fin :

$$J(\theta) = - \frac{1}{m} \sum_{i=1}^m [y^{(i)} \log(h_{\theta}(x^{(i)})) + (1 - y^{(i)}) \log(1 - h_{\theta}(x^{(i)}))] + \frac{\lambda}{2m} \sum_{j=1}^n \theta_j^2 \quad (\text{III. 35})$$

La deuxième somme  $\sum_{j=1}^n \theta_j^2$ , signifie exclure explicitement le terme de biais,  $\theta_0$ . C'est-à-dire le vecteur  $\theta$  est indexé de 0 à n (contenant n+1 valeurs), et cette somme saute explicitement  $\theta_0$ , en

passant de 1 à n, en sautant 0. Ainsi, lors du calcul de l'équation, nous devons continuellement mettre à jour les deux équations du gradient descent.

### III.17.2.2 La descente de gradient

$$\theta_0 = \theta_0 - \alpha \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)}) x_0^{(i)} \quad (\text{III. 36})$$

$$\theta_j = \theta_j - \alpha \frac{1}{m} \sum_{i=1}^m (h_{\theta}(x^{(i)}) - y^{(i)}) x_j^{(i)} \quad j = 1, 2, \dots, n \quad (\text{III. 3})$$

## III.18 Conclusion

Dans ce chapitre nous avons définies les notions de bases et les principes de l'apprentissage automatique (Machine Learning) qui nous permettons d'effectuer de multiples actions tel que classifications des données, faire apprendre à un programme à partir d'expérimentations ou encore de créer un programme évolutionnaire qui s'améliore sans cesse. Dans le chapitre suivant nous allons utiliser ces notions pour améliorer la

---

*CHAPITRE IV*

*SIMULATIONS ET RÉSULTATS*

---

## N1. Introduction

Au cours de ce chapitre, nous allons tâcher d'appliquer la régression linéaire à variable unique puis la régression linéaire à multiples variables à un ensemble de données préalablement générées.

Cela se fera sous environnement Matlab, à travers la 5G toolbox, qui nous permettra de simuler un réseau 5G à base d'une station de base et de plusieurs récepteurs.

## N2 Génération des données

### IV.2.1. Émetteurs et récepteurs

Pour la génération des données, nous avons implémenté une station de base 5G au centre-ville de Bejaia, émettant à une fréquence de 3.5GHz, avec une puissance de 1dB.

Nous avons ensuite placé une multitude de récepteurs, à des distances variant de 5 à 500m avec un pas de 15m, et formant des angles de 45 à 360° avec un pas de 45°.

Le réseau simulé est illustré par la figure IV.1.



Fig. IV.1 : Réseau simulé pour la génération des données.

 : Émetteur  : Récepteur

## IV.2.2. Calcul des puissances reçues

Pour le calcul des puissances parvenant aux récepteurs, et dans un souci de réalisme, nous avons ajouté plusieurs contraintes de transmissions à notre réseau.

Nous avons d'abord dû insérer des antennes émettrices à notre station de base (diagramme de rayonnement en fig. IV.2), puis nous avons simulé pour plusieurs récepteurs, les puissances reçues en présence de pluie ou de feuillage, qui sont des éléments atténuants pour un signal.

Les données générées par notre algorithme, et qui seront stockées pour être utilisées dans la régression linéaire sont les suivantes :

- Puissances reçues par les utilisateurs (272 au total).
- Distance séparant l'émetteur du récepteur.
- Elévation du récepteur par rapport à l'émetteur.
- Angle formé par le récepteur avec l'émetteur.
- Présence ou absence de pluie lors de la mesure (binaire).
- Présence ou absence de feuillage entre l'émetteur et le récepteur (binaire).

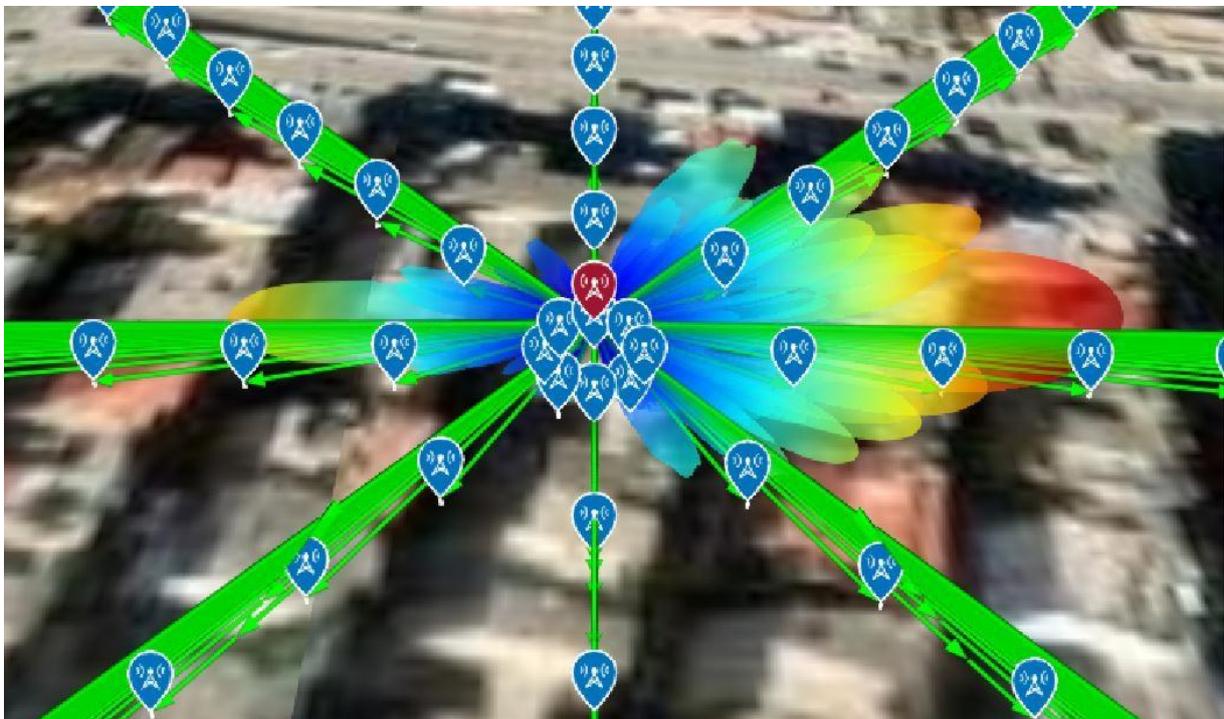


Fig. IV.2 : Diagramme de rayonnement de la station de base.

## IV.3. Régression linéaire à variable unique

### IV.3.1. Représentation des données

Pour la régression linéaire à variable unique, nous avons pris en considération la puissance reçue par les utilisateurs uniquement en fonction de la distance qui les sépare de l'émetteur.

La figure IV.3 en est une représentation :

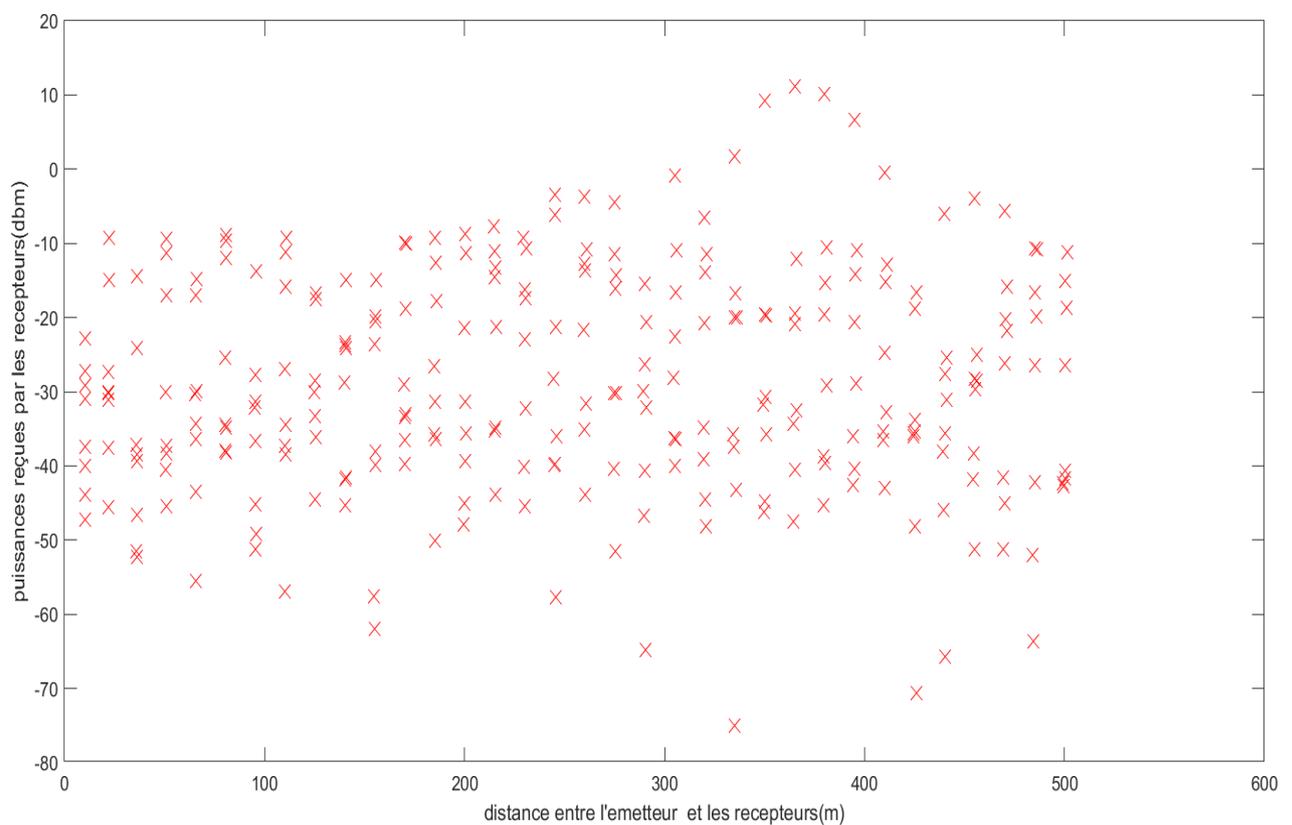


Fig.IV.3 : Puissance reçue en fonction de la distance.

### IV.3.2. Fonction Cout

La fonction coût ou « cost function », qui résulte de notre ensemble de données est représentée en Fig. IV.4, notre but sera de trouver des valeurs de  $\theta_0$  et  $\theta_1$  la minimisant.

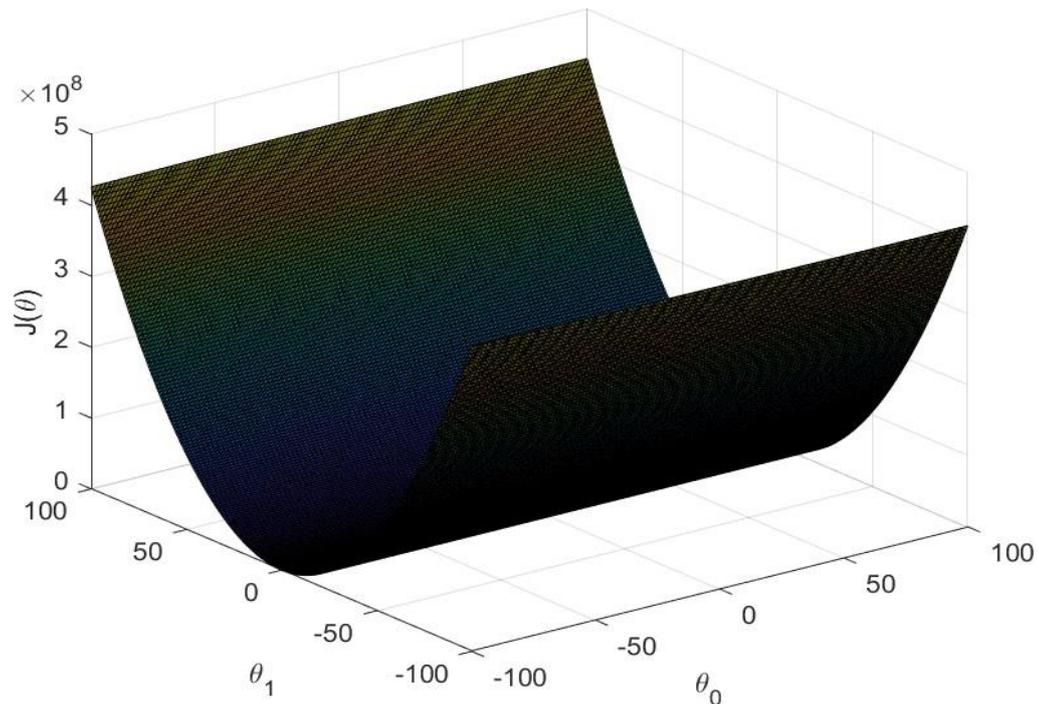


Fig. IV.4 : Cost Function

### IV.2.3. Optimisation par descente de gradient

L'implémentation de l'algorithme de descente de gradient (gradient descent) se fera en choisissant rigoureusement la valeur du taux d'apprentissage  $\alpha$ . En effet, les puissances du signal sont de l'ordre de -20 dbm, tandis que les valeurs des distances vont jusqu'à 500m. La valeur de l'écart entre la fonction d'hypothèse et les puissances mesurées sera donc élevée, il faudra donc choisir un taux d'apprentissage bas pour éviter que l'algorithme ne converge trop vite, risquant ainsi de rater la valeur minimale de la fonction coût.

Les valeurs de  $\theta_0$  et  $\theta_1$  minimisant la fonction coût, calculées par la descente de gradient sont :

$$[-0.001019, -0.084287]$$

### IV.3.3. Graphe de la fonction d'hypothèse

La fonction d'hypothèse, résultant de la régression linéaire à variable unique, est superposée aux données réelles dans la figure IV.5.

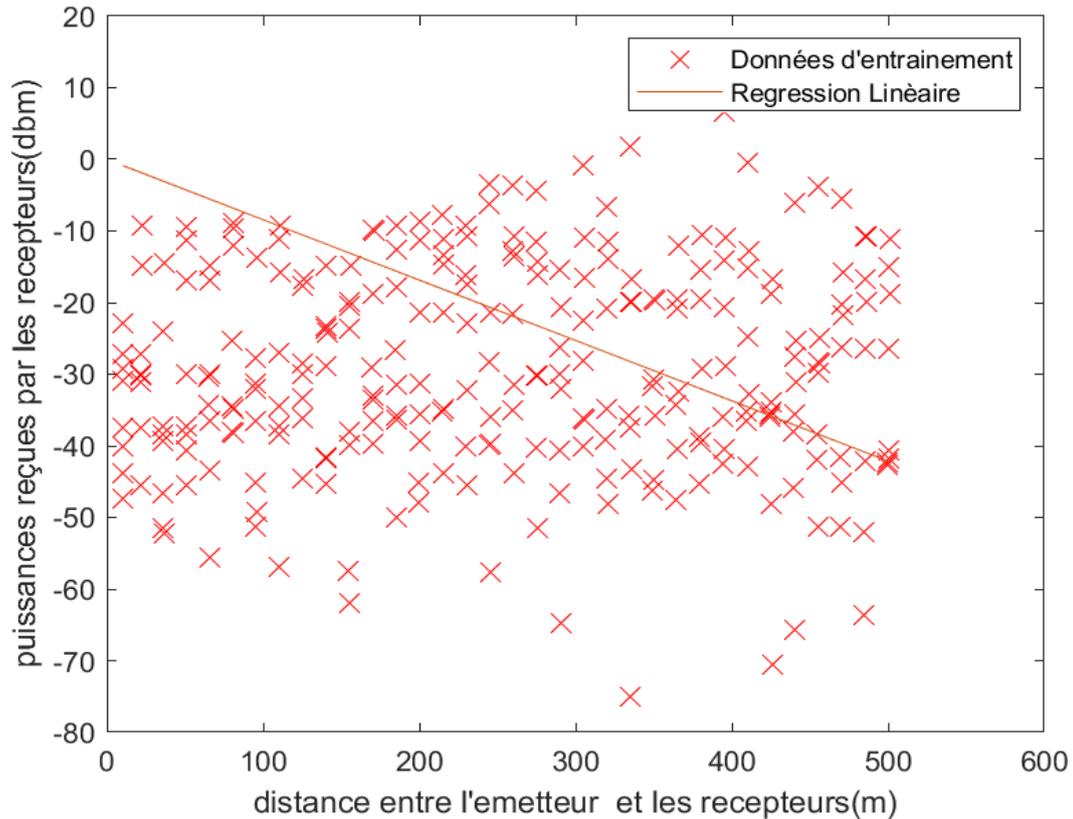


Fig. IV.5 : Régression linéaire à variable unique.

Comme illustré dans le graphe ci-dessus, ainsi que dans la figure IV.6, la prédiction de la régression linéaire à variable unique ne peut pas réellement être considérée comme efficace. Nous allons donc intégrer d'autres paramètres, ayant une incidence directe sur la puissance de signal reçue, dans ce qui suit.

pour un utilisateur à 500m, on prédit une puissance reçue de -42.144716 dbm  
 pour un utilisateur à 200m, on prédit une puissance reçue de -16.858498 dbm

Fig.IV.6 : Prédiction pour des utilisateurs donnés.

## IV.4 Régression linéaire à multiple variables

Comme mentionné dans l'introduction, les paramètres pris en considération pour l'application de la régression linéaire à multiple variables sont les suivants :

- Distance séparant l'émetteur du récepteur.

- Elévation du récepteur par rapport à l'émetteur.
- Angle formé par le récepteur avec l'émetteur.
- Présence ou absence de pluie lors de la mesure(binaire).
- Présence ou absence de feuillage entre l'émetteur et le récepteur(binaire).

Nous aurons donc 6 valeurs de  $\theta$  à trouver. (Nombre de variables+1)

#### IV.4.1. Équation normale

Pour calculer les valeurs de  $\theta$  nous allons, pour ce cas, appliquer l'équation normale.

Les valeurs résultantes, minimisant la nouvelle fonction cout, sont :

**[-21.2023 0.0017 0.1057 0.0180 -20.6673 4.4403]**

#### IV.4.2. Graph de la fonction d'hypothèse :

La fonction d'hypothèse, résultant de la régression linéaire à multiples variables, est superposée aux données réelles dans la figure IV.7 :

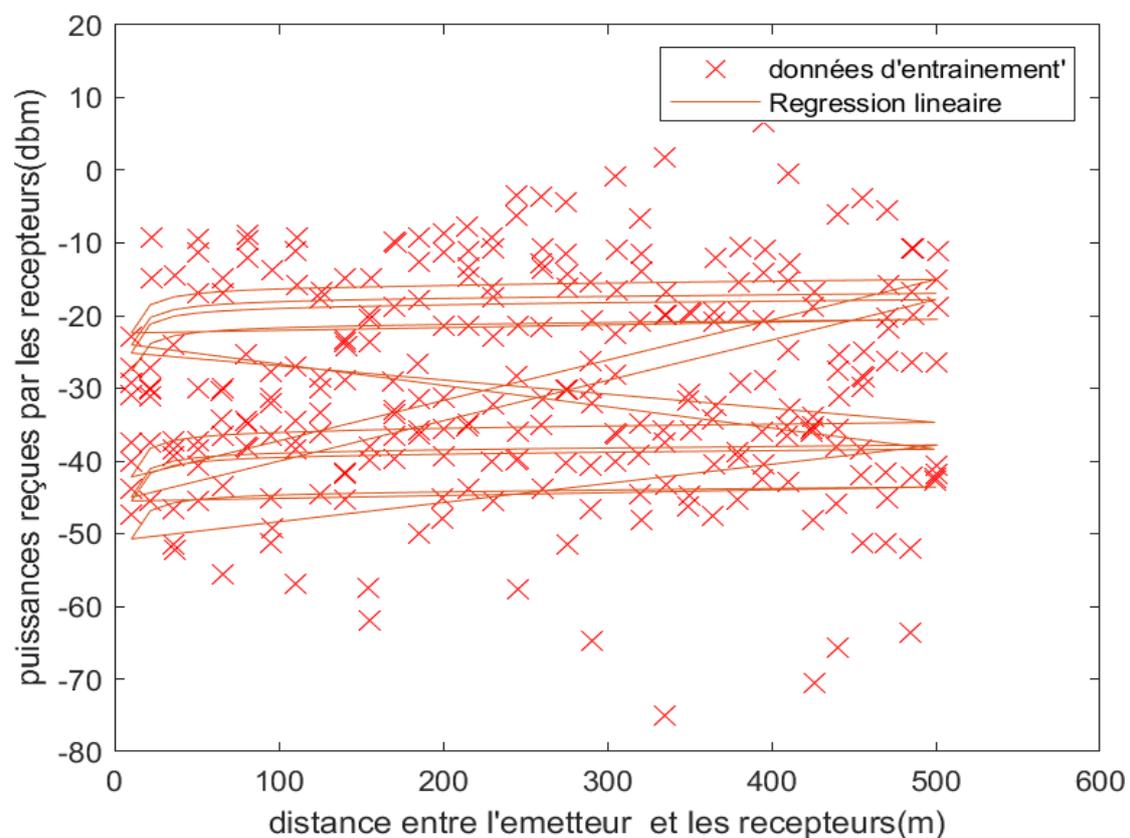


Fig. IV.7 : Régression linéaire à multiples variables

Nous voyons, en observant la figure IV.7, que la fonction d'hypothèse est bien plus adaptée que dans le cas de la régression linéaire à variable unique, nous allons maintenant l'appliquer à plusieurs utilisateurs, absents des données d'entraînement, les résultats sont illustrés dans la figure II.8 :

un utilisateur à 223m de l'émetteur, une élévation de -10m, un azimut de  $25^\circ$  sans feuillage et en temps pluvieux, on prédit une puissance reçue de -17.00 dbm  
un utilisateur à 382m de l'émetteur, une élévation de -2m, un azimut de  $68^\circ$  avec feuillage et en temps pluvieux, on prédit une puissance reçue de -35.78 dbm  
un utilisateur à 244m de l'émetteur, au même niveau d'élévation, un azimut de  $25^\circ$  sans feuillage et sans pluie, on prédit une puissance reçue de -20.34 dbm  
un utilisateur à 12m de l'émetteur, une élévation de -6m, un azimut de  $9^\circ$  sans feuillage et en temps pluvieux, on prédit une puissance reçue de -17.21 dbm  
un utilisateur à 354m de l'émetteur, une élévation de -18m, un azimut de  $99^\circ$  avec feuillage et en temps pluvieux, on prédit une puissance reçue de -36.96 dbm

Fig. IV.8 : Prédiction pour divers utilisateurs.

## V5. Conclusion

La régression linéaire, peut être considérée comme un excellent moyen de prédire le niveau de performance d'un réseau avant son déploiement, de par sa capacité à prédire des résultats en se basant sur des données d'entraînement très réduites, elle permet de connaître, approximativement, la qualité de service dont pourront jouir les utilisateurs.

---

## **CONCLUSION GÉNÉRALE**

---

Dans ce mémoire on a abordé d'abord l'évolution des réseaux mobiles de la première génération jusqu'à la cinquième génération en présentant leurs architectures, leurs différentes techniques d'accès multiple et leurs spécificités.

Ensuite, nous nous sommes attelées à présenter les différentes exigences du nouveau standard 5G, suivies des concepts de l'apprentissage automatique qui sont de plus en plus intégrés dans la 5G.

Notre objectif est d'explorer les techniques de prédiction de performances d'un réseau avant son déploiement, en utilisant la régression linéaire à variable unique dans une première étape et à variables multiples dans une seconde étape.

Au terme de nos investigations, nous avons réalisé que l'apprentissage automatique et la 5G ont une grande marge de progression ensemble [38]. En effet, l'intégration des fonctions d'intelligence artificielle vise à développer des réseaux cognitifs capables d'apprendre de leur environnement RF.

Ce travail nous a permises d'approfondir nos connaissances en réseaux de communications modernes, en simulation sous matlab de tels réseaux, et en techniques de Machine Learning. Nous espérons que ce travail puisse servir de futures études dans la même thématique.

## *Références Bibliographiques*

- [1] A. Cavoukian, Ph.D, "Bureau du commissaire à l' information et à la protection de la vie privée", ISSN 1188-3006, No. 14, août 2007
- [2] Guy Pujolles, Les Réseaux édition 2014, Paris : Edition Eyrolles, 27 février 2014
- [3] Jörg Eberspächer, Hans-Joerg Voegel, christian bettstetter, GSM Switching, Services and Protocols, Second Edition, 3eme Edition, WILEY, mai 2001.
- [4] J. tisal, le réseau GSM évolutions GPRS, I-MODE et WAP, DUNOD, 4ème édition, juillet 2003
- [5] E. Meurisse, " L'UMTS et le haut-débit mobile", février 2007
- [6] Yannick Bouguen, Éric Hardouin, François-Xavier Wolff, LTE et les réseaux 4G, Préface d'Alain Maloberti, Groupe Eyrolles, 2012, ISBN : 978-2-212-12990-8.
- [7] efort. 2009, site d'études et formations en Télécommunications Service et Réseaux de Télécommunication. [en ligne]. [www.efort.com](http://www.efort.com)
- [8] Erik Dahlman, Stefan Parkvall, and Johan Sköld , 4G LTE/LTE-Advanced for Mobile Broadband. First published . The Boulevard, Langford Lane, Kidlington, Oxford, OX5 1GB, UK 30 Corporate Drive, Suite 400, Burlington, MA 01803, USA, 2011, page 15
- [9] Aliou DIALLO, SYTEMES MULTI-ANTENNES POUR DIVERSITE ET MIMO, THESE pour obtenir le titre de Docteur en Sciences spécialité Electronique, NICE, 1<sup>re</sup> UNIVERSITE de Nice-Sophia Antipolis, 2007, 278p.
- [10] altium. 24 aout 2017 les avantages du MIMO massif pour infrastructure cellulaire 5G. [en ligne]. <https://resources.altium.com>
- [11] EFORT. (2019) Introduction à la 5G. [en ligne]. <http://www.efort.com/>
- [12] Marco Ajmone , Marsan Nicola, Blefari Melazzi , Stefano Buzzi, Sergio Palazzo. The 5G Italy Book 2019: a Multiperspective View of 5G. 43124 Parma (PR) – ITALY . Edited by: Raffaele Bolla University of Genova and CNIT Giovanni Schembra University of Catania, 2019. 572pages : 100,226
- [13] frandroid. 28 mars 2021. [En ligne]. <https://www.frandroid.com/>
- [14] DEHEMCHI ? MARWA, et ASMA BOUHAFER. Etude et analyse de méthodes de Beamforming pour des système multi-utilisateur. Mémoire de fin d'étude, Guelma : université 8 mai 1945, juillet 2019.
- [15] Ali Zaidi Fredrik Athley Jonas Medbo Ulf Gustavsson Giuseppe Durisi Xiaoming Chen, 5G Physical Layer Principles, Models and Technology Components. ACADEMIC PRESS An imprint of Elsevier , 2018, 294p. page 14
- [16] Saad Z. Asif, Livre\_ 5G Mobile Communications Concepts and Technologies\_2019. CRC Press Taylor & Francis Group 6000 Broken Sound Parkway NW, Suite 300 Boca Raton, FL 33487-2742, 2019, 312 page.
- [17] Deep learning based massive MIMO beamforming for 5G, 19 jan 2019. [en ligne].

## *Références Bibliographiques*

- [https://www.researchgate.net/publication/328836812\\_Deep\\_Learning\\_Based\\_Massive\\_MIMO\\_Beamforming\\_for\\_5G\\_Mobile\\_Network](https://www.researchgate.net/publication/328836812_Deep_Learning_Based_Massive_MIMO_Beamforming_for_5G_Mobile_Network)
- [18] BENDIMERAD FETHI TARIK, « MIMO ET ACCÈS MULTIPLES AVANCÉS POUR RÉSEAUX SANS FIL ». Projet national de recherche : rapport final, Direction Générale de la Recherche Scientifique et du Développement Technologique, Tlemcen : Université Abou Bekr Belkaid.
- [19] BENYAROU, MOURAD. (7-avr-2014) Détection Multi-Utilisateurs par réseau d'antennes appliquée à un signal CDMA à porteuses multiples. [en ligne]. <http://dspace.univ-tlemcen.dz/handle/112/4635>
- [20] RAPPORT UIT-R M2040 – Antennes adaptatives et principales caractéristique URL : [https://www.itu.int/dms\\_pub/itu-r/opb/rep/R-REP-M.2040-2004-PDF-F.pdf](https://www.itu.int/dms_pub/itu-r/opb/rep/R-REP-M.2040-2004-PDF-F.pdf)
- [21] : <https://www.researchgate.net/publication/48418277>
- [22] Guillermo Andrés, Pocovi Gerardino, Sonia Barcos Sánchez, master's Thesis „LTE-Advanced HetNet Investigations Under Realistic Conditions“, AALBORG UNIVERSITY DENKARK, September 2013 – June 2014, Pages :18-19.
- [23] Manuel Eugenio Morocho Cayamcela\* , Wansu Lim, « Artificial Intelligence in 5G Technology: A Survey », IEEE, 2018, 6 pages.
- [24] Akansha Gupta\*, Kamal Ghanshala, R. C. Joshi, « Machine Learning Classifier Approach with Gaussian Process, Ensemble boosted Trees, SVM, and Linear Regression for 5G Signal Coverage Mapping », International Journal of Interactive Multimedia and Artificial Intelligence, Vol. 6, N°6, 30 Mars 2021, 8 pages.
- [25] MACHINE LEARNIA, Guillaume Saint-Cirgue. Juillet 2, 2019. [En ligne]. <https://machinelearnia.com/apprentissage-supervise-4-etapes/>
- [26] ANALYTICS & INSIGHTS. [En ligne]. <https://analyticsinsights.io/apprentissage-supervise-vs-non-supervise/>
- [27] ICHI.PRO. [En ligne]. <https://ichi.pro/fr>
- [28] le blog Business & Decision. [En ligne]. <https://fr.blog.businessdecision.com/tutoriel-regression-lineaire-et-descente-de-gradient-en-machine-learning/>
- [29] WikiStat, Introduction à la régression multiple. [En ligne]. <https://www.math.univ-toulouse.fr/~besse/Wikistat/pdf/st-l-inf-intRegmult.pdf>
- [30] introduction a l'apprentissage automatiques. [En ligne]. [https://proeduc.github.io/intro\\_apprentissage\\_automatique/regression.html#v-3-r%C3%A9gression-polynomiale](https://proeduc.github.io/intro_apprentissage_automatique/regression.html#v-3-r%C3%A9gression-polynomiale)
- [31] machine Learning. [En ligne]. <https://developers.google.com/machine-learning/crash-course?hl=fr>
- [32] DATA ANALYTICS POST. [En ligne]. <https://dataanalyticspost.com/Lexique/overfitting/>

## *Références Bibliographiques*

- [33] FONCTIONS DE COÛT ET MACHINE LEARNING. [EN LIGNE].  
<https://franckybox.com/fonctions-de-cout-et-machine-learning/>
- [34] André pérez, LTE et LTE Advanced (les interfaces radioélectriques du reseau mobile4G), British Libaray Cataloguing-in-Publication Data, Edition ETSE LTD 2016.
- [35] André pérez,Frédéric Launay, LTE-Advanced Pro (une étape vers le réseau de mobiles5G), British Libaray Cataloguing-in-Publication Data, Edition ETSE LTD 2019.
- [36] Pierre Lescuyer, UMTS (les origines, l'architecture, la norme), 2eme Edition, Dunaud, paris, 2002.
- [37] Pierre Lescuyer, réseaux 3G (principe, architecture et services de l'UMTS), 2eme Edition, Dunaud, paris, 2006.
- [38] MANUEL EUGENIO, MOROCHO-CAYAMCELA, HAEYOUNG LEE et WANSU LIM. Machine Learning for 5G/B5G mobile and wireless communication : potential, limitations, and future directions. IEEE ACCESS (19 septembre 2019).