

République Algérienne Démocratique et Populaire

Ministère de l'Enseignement Supérieure et de la Recherche Scientifique

Université Abderrahmane Mira

Faculté de Technologie



Département d'Automatique, Télécommunication et d'Electronique

Projet de Fin d'étude

Pour l'obtention du diplôme de Master

Filière : Télécommunications

Spécialité : Système des Télécommunications

Thème

Allocation de ressources dans un réseau 5G het-net en utilisant l'algorithme « Decentralized Q-Learning »

Préparé par :

- ABDELFETTAH Yasmine.
- ISSAD Yasmina.

Dirigé par :

- Dr. M. A. OUAMRI.
- Dr. M. AZNI.

Examiner par :

- Mme. MAMMERI.
- Mme. GAGAOUA.

Année universitaire : 2021/2022

Remerciements

Il nous est agréable et important de réserver cette page comme le témoignage de notre reconnaissance à toutes les personnes qui nous ont soutenues pour la réalisation de ce travail.

Nous remercions Dieu le tout-puissant pour sa grâce qu'il nous accorde tous les jours.

Nous remercions nos encadrants Mr OUAMRI.M.A et Mr AZNI.M, pour leurs encadrements qui nous a été précieux afin de mener notre travail à terme.

Nous remercions également toute les personnes qui n'ont pas aidé pour la réalisation de ce projet plus précisément Ben Madani Housseem, Abdelli Lyes et Bouhadad Maria.

Nous remercions également les membres du jury d'avoir accepté d'évaluer ce travail.

Nous remercions particulièrement nos chères familles respectives qui n'ont pas cessé de nous encourager et de nous motiver durant tout notre parcours éducatif.

Nous remercions nos chers ami(e)s de nous avoir soutenus durant toute la phase de préparation.

Table des matières

Liste des figures	iv
Liste des tableaux	vi
Liste des abréviations.....	vii
Introduction générale.....	1
CHAPITRE I Réseaux mobiles de nouvelle génération	
Introduction	4
1. Chronologie des réseaux cellulaires	4
2. Techniques de transmission.....	6
2.1. Propagations LOS et NLOS	6
2.2. Multi-Input Multi Output (MIMO) et Massive MIMO	7
2.3. Beamforming	9
3. Techniques d'accès au canal.....	10
3.1. Modulations OFDM et OFDMA.....	10
3.2. Techniques d'accès OMA et NOMA	11
4. Type de canaux	14
4.1. Canal de Rayleigh	14
4.2. Canal de Rice	15
4.3. Canal de Nakagami	15
5. Interférences cellulaires	15
5.1. Interférences intercellulaires.....	16
5.2. Interférences intracellulaires	16
5.3. Coordinations d'interférences (FFR, SFR).....	16
6. Réseaux hétérogènes.....	17
Conclusion.....	18

CHAPITRE II Allocation de ressources dans un réseau 5G

Introduction	20
1. Principe d'allocation de ressources	20
1.1. Contraintes d'allocation de ressources	21
1.2. Les ordonnanceurs d'allocation de ressources et leurs principes	23
1.3. Les objectifs d'allocation de ressource	24
1.4. Allocation de puissance	24
2. Efficacité spectrale	25
2.1. Spectre de fréquence 5G	25
2.2. Amélioration de l'efficacité spectrale	26
3. Efficacité énergétique	27
3.1. Efficacité énergétique dans les réseaux 5G	28
3.2. L'enjeu de l'efficacité énergétique de la 5G	28
3.3. Sources de consommation d'énergie	28
4. Délai de tolérance	29
4.1. Service 5G et délai de tolérance	29
Conclusion	29

CHAPITRE III Apprentissage par renforcement

Introduction	31
1. L'apprentissage par renforcement et son principe	31
1.1. Méthode de résolution dans AR	32
1.2. Définition des paramètres de l'apprentissage par renforcement	32
2. Processus de décision de Markov	33
3. Politique d'apprentissage par renforcement	34
3.1. Les différentes politiques	34
4. Fonction de valeur	35

4.1. Différentes fonctions de valeurs	35
4.2. Diagramme de sauvegarde des fonctions de valeur	35
5. Fonction de BELLMAN.....	36
6. Algorithme de renforcement Learning.....	36
6.1. Q-Learning décentralisé	37
6.2. Deep Q-network	38
Conclusion.....	38
CHAPITRE IV Simulations et interprétations des résultats	
Introduction	40
1. Modèle et paramètre de simulation	40
1.1. Modèle de perte de trajet dans le canal.....	40
Partie 01	41
1.1. Calcul du débit	41
1.2. Fonction de récompense	41
Partie 02	46
2.1. Calcul de l'efficacité énergétique	46
2.2. Résultats obtenus.....	46
Conclusion.....	49
Conclusion générale	50
Références	51

Liste des figures

Figure I.1 : L'évolution des réseaux mobiles.....	4
Figure I.2 : Propagation Line-of-Sight.....	7
Figure I.3 : Propagation NLOS.....	7
Figure I.4 : Système MIMO.....	8
Figure I.5 : Beamforming.....	9
Figure I.6 : Principe des techniques OFDM et OFDMA.....	11
Figure I.7 : Principe des techniques OMA et NOMA.....	12
Figure I.8 : Interférences intercellulaires.....	16
Figure I.9 : Interférences intracellulaires.....	16
Figure I.10 : Comparaison entre SFR et FRR.....	17
Figure I.11 : Dépoilement hétérogène.....	17
Figure II.1 : Illustration des réseaux sans fil de prochaine génération caractérisés par leur hétérogénéité massive.....	20
Figure II.2 : Modèle de propagation par perte de chemin dans les communications mobiles.....	22
Figure II.3 : Comparaison entre une transmission du signal direct et une transmission avec obstacles.....	22
Figure II.4 : Modèle de propagation d'un signal subissant un évanouissement.....	23
Figure II.5 : Les mécanismes de propagation d'un signal radio-mobile.....	23
Figure II.6 : Fonctionnement de l'ordonnanceur pour les réseaux radios mobiles.....	24
Figure II.7 : Spectre de fréquence allouée à la 5G.....	26
Figure III.1 : Système d'apprentissage par renforcement.....	31
Figure III.2 : Découpage des méthodes en apprentissage par renforcement.....	32
Figure III.3 : Processus de la décision de Markov.....	33
Figure III.4 : Paradigme général de l'apprentissage par renforcement.....	34
Figure III.5 : Diagrammes de sauvegarde pour $V(\pi)$ et $Q(\pi)$	36
Figure III.6 : Principe de Bellman.....	36
Figure III.7 : Algorithme Q-Learning.....	37
Figure III.8 : Deep Q-network.....	38

Figure IV. 1 : Carte du réseau.	40
Figure IV.2 : Débit global optimal.	43
Figure IV.3 : Evaluation de α et γ	43
Figure IV.4 : Débit individuel expérimenté par chaque WN au cours d'un seul cycle de simulation pour différents α , γ et ε_0	44
Figure IV.5 : Débit moyen expérimenté par WN pendant les 5000 dernières itérations sur un total de 10000 itérations pour différents α , γ , et ε_0	44
Figure IV.6 : Probabilité de choisir les différentes actions à chaque WN pour une simulation unique (10000 itérations) pour différents α , γ et ε_0	45
Figure IV.7 : Efficacité énergétique globale.	46
Figure IV.8 : Efficacité énergétique individuel expérimenté par chaque WN au cours d'un seul cycle de simulation pour différents pour différents α , γ et ε_0	47
Figure IV.9 : Efficacité énergétique moyenne expérimenté par WN pendant les 5000 dernières itérations sur un total de 10000 itérations pour différents α , γ et ε_0	47
Figure IV.10 : Probabilité de choisir les différentes actions à chaque WN pour une simulation unique (10000 itérations) pour différents α , γ et ε_0	48

Liste des tableaux

Tableau I.1: Comparaison entre les différentes générations mobiles.	6
Tableau I.2: Comparaison entre MIMO, MISO, SIMO et SISO.	9
Tableau IV.1 : Paramètres de simulation.	42

Liste des abréviations

2G/3G/4G/5G : 2^{ème} /3^{ème} /4^{ème} /5^{ème} Génération.

3GPP : 3rd Generation Partnership Project.

AMPS : Advanced mobile Phone System.

AGRR : Allocation et Gestion des Ressources Radio

AP : Access Point.

AR: Apprentissage par Renforcement.

ARMA: Apprentissage par Renforcement Multi-Agent.

AMRC : Accès Multiple par Répartition en Code.

AWGN : Additive White Gaussian Noise.

BS/BTS : Base Station.

CD : Code Domaine.

CDMA : Code Division Multiple Access.

CNA : Convertisseur Numérique-Analogique.

CQI : Channel Quality Indicator.

D2D : Device To Device.

EE : Efficient Energy.

E-NB: Evolved-Node-Base.

E-MBB : Enhanced Mobile Broadband.

FFR : Fractional Frequency Reuse.

FSPL : Frequency Signal Path-Loss

GSM : Global System for Mobile Communications.

HET-NET : Heterogenic Network.

ICIC : InterCell Interference Coordination.

IDFT : Inverse Descrete Fourier Transform.

IOT: Internet Of Things.

LTE : Long Term Evolution.

LTE-A: Long Term Evolution-Advanced.

LOS: Line-of-Sight.

M2M : Machine to Machine.

MDP : Marcov Decision Process.

MIMO : Multi-Input Multi-Output.

MU-MIMO : Multi-User Multi-Input Multi-Output.

MISO : Multi-Input Single-Output.
MMTC: Massive Machine Type Communication.
NLOS : Non-Line-of-Sight.
NOMA : Non Orthogonal Multiple Access.
NR: New Radio.
OFDM : Orthogonal Frequency Division Multiple.
OFDMA : Orthogonal Frequency-Division Multiple Access.
OMA : Orthogonal Multiple Access.
PD : Puissance Domaine.
PDM : Processus De Marcov.
PLE : Path Loss Exponent.
PSK : Phase-Shift Keying.
QAM : Quadrature Amplitude Modulation.
RF: Radio Frequency.
SE : Spectrum Frequency.
SFR : Soft Frequency Reuse.
SIC : Successive Interference Cancellation.
SMS : Short Message Service.
SNR: Signal and Noise Ratio.
SINR : Signal To Interference and Noise Ratio.
SIMO: Multiple Input Single Output.
SISO : Single Input Single Output.
SU-MIMO : Single User- Multi-Input Multi-Output.
TDMA : Temporal Division.
UE : Utilisation Equipment.
UMTS : Universal Mobile Telecommunication System.
URLLC : Ultra-Reliable Low Latency Communication.

Introduction générale

La 5G est une technologie de téléphonie mobile qui suit celle de la 4G elle ne vient pas pour remplacer les réseaux déjà en place (3G, 4G) mais plutôt une extension et un parfait complément. Les fréquences de la 5G viendront donc s'ajouter aux autres déjà en place. Cette dernière aura un débit plus performants que la 4G et un temps de latence plus bas. Elle permettra de connecter un nombre impressionnant d'objets, d'équipements ou de voitures, tout cela simultanément en dépensant moins d'énergie à cause de sa grande capacité [1].

Le réseau 5G peut fournir une vitesse et un débit élevés, une grande fiabilité, une faible latence, une capacité, une disponibilité et une connectivité accrue, ainsi qu'une allocation dynamique de la bande passante. Ce sont les attributs de qualité de service qui seront consommés par les services 5G qu'ils desservent et qui les caractérisent comme : eMBB (*enhanced Mobile BroadBand*), uRLLC (*ultra-reliable and low-latency communications*), Mmtc (*Massive machine type communications*).

Dans ce cadre, des technologies tels que les réseaux hétérogènes, les communications de machine à machine, les réseaux d'appareil à appareil (D2D) et l'Internet des objets (IoT) ont été proposées pour faire l'objet d'une normalisation sous l'égide du développement et l'optimisation des réseaux 5G qui peuvent être réalisées en utilisant une technique d'intelligence artificielle appelée apprentissage par renforcement (*reinforcement learning*), qui fournit un résultat très satisfaisant en atteignant des performances maximales théoriques [1], [2].

Le but ultime de notre travail est l'étude et l'analyse d'allocation de ressource dans un réseau 5G hétérogène avec l'apprentissage par renforcement pour trouver le débit et l'efficacité énergétique optimale. Donc nous proposons une variation sans état de l'apprentissage Q, que nous appliquons pour exploiter la réutilisation spatiale dans un réseau sans fil. En particulier, nous permettons aux réseaux de modifier à la fois leur puissance de transmission et le canal utilisé uniquement sur la base du débit expérimenté.

Dans le premier chapitre nous allons voir brièvement la notion des réseaux cellulaires de nouvelle génération ainsi que les techniques de transmissions, les types de canaux utilisés dans les réseaux cellulaires et les aspects d'accès à ces derniers.

Le second chapitre est consacré à l'allocation de ressource dans un réseau de 5^{ème} génération ses contraintes et ses objectifs puis on présentera un aperçu sur l'efficacité spectrale et énergétique.

Le troisième chapitre traite l'ensemble des concepts de base qu'on va utiliser dans notre simulation à savoir l'apprentissage par renforcement en introduisant le modèle mathématique de notre cas et ses différentes caractéristiques ayant comme but le calcul du débit et de l'efficacité énergétique.

Enfin, le quatrième chapitre est dédié à la simulation et à la discussion des différents résultats obtenus concernant la variation de la probabilité, le débit moyen, l'écart-type, l'efficacité énergétique et en prenant en considération quelques paramètres pouvant affecter cette dernière. Et nous terminerons par une conclusion générale.

CHAPITRE I

Réseaux mobiles de nouvelle génération

Introduction

Actuellement, les réseaux cellulaires constituent un moyen de communication intéressant pour offrir à plusieurs abonnés une qualité de service satisfaisante. Généralement, les réseaux cellulaires sont déployés avec la plus part des téléphones cellulaires, smartphone ainsi que des appareils à accès commuté. Ces appareils sont associés à des stations de base composant des cellules et peuvent se déplacer d'une zone de couverture à une autre toute en maintenant la communication. Cependant, les stations de base, ainsi que les contrôleurs radio sont équipées d'une capacité informatique dont l'objectif est de suivre l'utilisateur en temps réel [4].

Le déploiement d'une telle infrastructure nécessite une planification ainsi qu'un dimensionnement du réseau. En effet, les stations de base doivent être installées d'une manière à couvrir tous les utilisateurs dans le réseau, avec une contrainte de capacité c'est-à-dire que le trafic véhiculant dans le canal doit satisfaire l'équation de Shannon [4].

Dans ce chapitre, nous allons introduire la notion des réseaux cellulaires ainsi que les techniques de transmission. Ensuite, nous introduisons les aspects d'accès au canal et les types de canaux utilisés dans les réseaux cellulaires. De plus, nous illustrons les schémas d'interférences inter et intra cellulaires. En fin, une initiation au réseau de (5G) est investie.

1. Chronologie des réseaux cellulaires

Alors que les réseaux mobiles continuent d'évoluer rapidement, il est utile de prendre du recul et de revoir leurs évolutions. Tout d'abord, les laboratoires Bell aux États-Unis ont réussi à mettre au point le premier système de téléphonie cellulaire mobile au monde (AMPS) (*Advanced Mobile Phone System*). Cette approche, est mise en service commercial à Chicago, puis rapidement adoptée dans tous les États-Unis et connaît un grand succès. Au cours de la même période, d'autres régions, comme l'Europe, ont également mis en place leurs propres systèmes de communication mobile de première génération, qui reposaient sur la technologie analogique. Néanmoins, La capacité des systèmes était limitée, de sorte que les réseaux ne pouvaient fournir que des services vocaux de base. D'autre part, trouver des moyens de gérer la sécurité et les interférences constituait un problème majeur pour les opérateurs mobiles [1], [4].

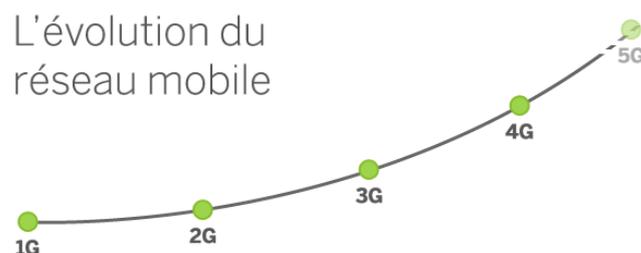


Figure I.1 : L'évolution des réseaux mobiles.

Alors que l'application des circuits intégrés, des microprocesseurs et des signaux numériques gagnaient en maturité à grande échelle, la technologie des communications mobiles a évolué pour devenir une technologie numérique à partir de la 2^{ème} génération. Les principaux systèmes de communication mobile de la 2^{ème} génération comprenaient le GSM (*Global System for Mobile Communications*) en Europe et d'autres systèmes de communication tels que le CDMA (*Code Division Multiple Access*) aux États-Unis et ailleurs. Les réseaux de communication mobile 2G fournissaient des services vocaux, des données à faible débit et de courts messages (SMS) (*Short Message Service*).

Avec la progression des services de données mobiles, la 3^{ème} génération appelé aussi UMTS (*Universal Mobile Telecommunication System*) a été développée au début du 21^{ème} siècle. Par rapport aux réseaux 2G, les réseaux 3G permettent de fournir des services de données à haut débit, ils utilisent la technologie d'accès multiple par répartition en code (AMRC) ou (CDMA) en anglais, qui a permis d'augmenter la capacité des réseaux pour mieux supporter les services de données. La 3G permet la fourniture de services d'information multimédia tels que des images, de la musique et des vidéos. Elle permet aux utilisateurs de naviguer facilement sur des sites web, de participer à des conférences téléphoniques et de faire du commerce électronique sur leur téléphone portable [2].

Le débit de données sur les réseaux 3G peut atteindre plus de 1Mbps, ce qui a favorisé le développement des réseaux 3G, augmentant rapidement le nombre d'utilisateurs et rendant les services de données accessibles à un plus grand nombre d'abonnés. En raison de l'existence de différentes technologies 3G, la compatibilité des terminaux et des réseaux a été une préoccupation pour les opérateurs et les consommateurs. Par conséquent, la quatrième génération est apparue comme une technologie révolutionnaire. Lors de son apparition, la norme technologique LTE, qui a été introduite par le projet de partenariat de 3^{ème} génération (3GPP), est devenue la norme technique mondiale unifiée pour les communications mobiles.

LTE est l'abréviation de "*Long Term Evolution*" (évolution à long terme) [1]. Cependant, par rapport aux générations précédentes, LTE adopte une nouvelle architecture de réseau et une technologie OFDMA (*Orthogonal Frequency Division Multiple Access*) permettant d'améliorer considérablement les capacités du réseau et permet un débit de données théorique atteignant les 100 Mbps. Aujourd'hui Les réseaux 4G sont désormais utilisés pour apporter plus de commodité et de vitalité à la vie quotidienne [1].

De nos jours, Les réseaux de 5^{ème} génération offrent de nouvelles fonctionnalités et de nouveaux services. L'objectif majeur de la 5G sera l'augmentation continue du débit de données, par rapport au réseau.

Le principe de la 5^{ème} génération diffère totalement de celui de la 4G, il inclut des fréquences porteuses très élevées avec des largeurs de bande massives, des stations de base à densité intense et un nombre inédit d'antennes.

Afin que cette nouvelle génération fournisse une couverture à haut débit, on associe le spectre 5G au LTE tout en améliorant la réglementation de ce spectre. De nombreux facteurs rendent la 5G intéressante tel que la demande d'une capacité de réseau plus grande et l'alimentation du débit de connexion [2].

Génération	Utilisation	Nom de technologie	Type d'accès sur voix radio
1 ^{ère} génération	Téléphonie	R2000, NMT	Analogique, FDMA
2 ^{ème} génération	Téléphonie et SMS	GSM	TDMA
3 ^{ème} génération	Téléphonie, SMS, Accès IP à 1Mbps	UMTS	CDMA
4 ^{ème} génération	Téléphonie, SMS, Accès IP à 100Mbps avec faible latence.	LTE, LTE-A	OFDM, OFDMA
5 ^{ème} génération	Téléphonie, SMS, Accès IP à 1Gbps avec faible latence	LTE-A, IMT-2020	OFDMA, OMA, NOMA

Tableau I.1: Comparaison entre les différentes générations mobiles.

2. Techniques de transmission

La transmission de donnée à longue distance est plus efficace grâce aux réseaux sans fil, mais les obstacles géographiques et la courbure de la terre peuvent causer certains problèmes. Pour assurer le bon acheminement et la transmission de donnée et garantir une bonne communication entre deux points plusieurs techniques sont utilisées par exemple le MIMO (*Multi-Input Multi-Output*). Il est bien important d'introduire la manière dont l'onde est véhiculée dans le réseau. En effet, deux types de propagation sont envisagés à savoir Line-of-Sight (LOS) et Non-line-of-Sight (NLOS) [3].

2.1. Propagations LOS et NLOS

En premier lieu la transmission LOS permet un acheminement des données entre deux stations de base, l'une est émettrice et l'autre réceptrice qui est en contact visuel direct l'une avec l'autre ce qui signifie que les ondes se déplacent sur un trajet direct de la source au récepteur sans aucun obstacle entre elles comme illustré dans la figure ci-dessous [3].

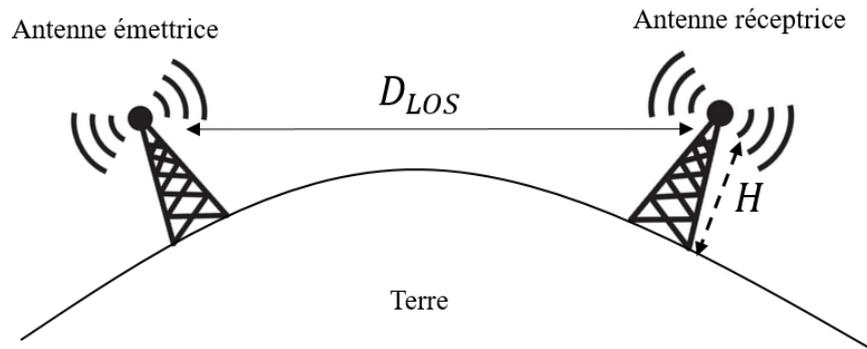


Figure I.2 : Propagation Line-of-Sight.

D'où on donne :

$$D_{LOS} = \sqrt{2HR} \quad (I.1)$$

Avec D_{LOS} est la distance en visibilité directe entre deux antennes terrestres, H est la hauteur de la station de base et R est le rayon de la terre $R = 6,38 \times 10^6$ m [3].

Contrairement au LOS, le NLOS signifie que l'émetteur et le récepteur ne sont pas en visibilité directe, en d'autre terme le chemin de propagation est obscurci par des obstacles dont la longueur d'onde est plus grande tel que des arbres, des bâtiments ou un conducteur électrique, créant ainsi des difficultés de passage du signal. Les propagations NLOS limitent la capacité de transmission des signaux et réduit la puissance reçue au niveau des utilisateurs [3].

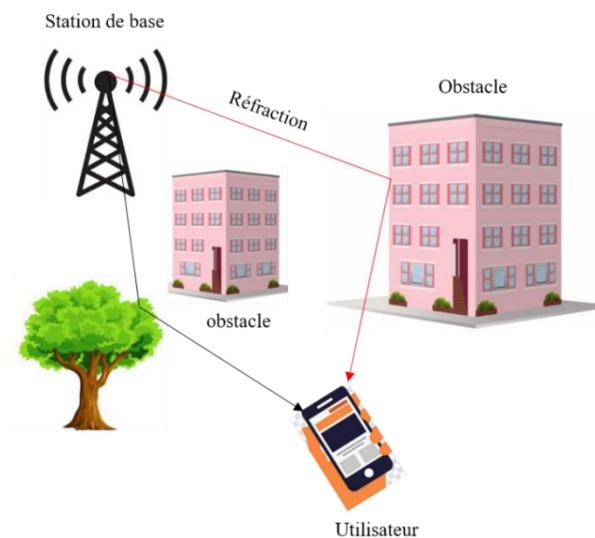


Figure I.3 : Propagation NLOS.

2.2. Multi-Input Multi Output (MIMO) et Massive MIMO

MIMO est une Technique de multiplexage utilisé dans les réseaux mobile et sans fil qui permet de transférer des données sur de longues distances à un débit de transmission élevé et d'améliorer la qualité du rapport signal sur bruit. Elle est aussi utilisée pour réduire les évanouissements du signal provoqués par l'environnement de propagation des trajets multiples. Cette technique consiste

à placer plusieurs antennes en émission (T_x) et en réception (R_x). Lors de la transmission les données sont divisées en plusieurs flux de données puis recombinaées une fois à la réception par une autre radio MIMO configurée avec le même nombre d'antennes [4]. Il existe trois catégories principales de la technique MIMO on cite : la diversité spatiale MIMO ; le multiplexage spatiale et le MIMO filtrage spatial (Beamforming).

- **Diversité spatiale MIMO** : consiste à transmettre le même message sur plusieurs différentes antennes de transmission, une fois à la réception les antennes réceptrices remettent le message en phase et qui serait sommé d'une façon ordonnée ; ensuite l'antenne qui reçoit le meilleur signal c'est celle qui sera utilisée [4].
- **Multiplexage spatial MIMO** : le multiplexage MIMO permet d'avoir un débit de transmission plus élevé ; le principe consiste à découper un message en plusieurs sous messages ensuite le transmettre simultanément sur chacune des antennes émettrices ; une fois à la réception les signaux sont réassemblés afin d'avoir le message original [4], [5].
- **MIMO Beamforming** : a pour avantage d'éliminer les interférences entre les utilisateurs et accroître la couverture radio ; elle est utilisée pour orienter et centrer le faisceau d'onde radio sur une zone donnée [4].

On considère un système MIMO composé de N_t antennes à l'émission et N_r d'antennes à la réception. L'équation de la capacité s'écrit donc comme suit :

$$C = \log_2(\det[IN_r + \frac{\rho}{N_t}HH^*]) \quad (I.2)$$

Avec IN_r est la matrice d'identité, ρ est le rapport signal sur bruit et H la matrice du canal MIMO $N_t \times N_r$.

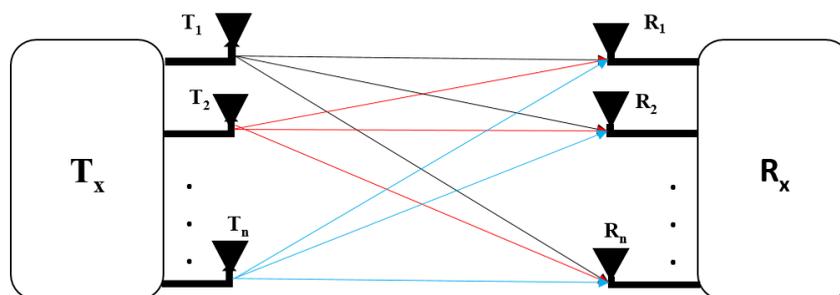


Figure I.4 : Système MIMO.

Massive MIMO est une technologie convaincante pour les réseaux sans fil son principe consiste à utiliser un grand nombre de réseaux d'antennes dans des stations de base. Elle est déployée lors de l'arrivée de la 5^{ème} génération ; dans lequel chaque station de base est équipée d'un réseau avec multiple émetteurs rassemblés dans un même équipement qui sera assisté par un logiciel, qui permet de centrer le faisceau d'onde radio sur une zone donnée [4], [5].

Comparaison entre MIMO, Massive MIMO et SISO, MISO, SIMO :

Technologie	Antenne émettrice	Antenne réceptrice
SISO	$T_X=1$	$R_X=1$
SIMO	$T_X=1$	$R_X=N$
MISO	$T_X=N$	$R_X=1$
MIMO	$T_X=N$	$R_X=N$

Tableau I.2 : Comparaison entre MIMO, MISO, SIMO et SISO.

SISO (*Single Input Single Output*) est le système le plus simple qui possède une seule antenne émettrice et une seule antenne réceptrice puis on a le système SIMO (*Single Input Multi Output*) qui son principe est d'avoir une seule antenne émettrice et plusieurs antennes réceptrices [5].

Contrairement à SIMO, MISO (*Multiple Input Single Output*) possède plusieurs antennes émettrices et une seule antenne réceptrice. En fin on a le système MIMO (*Multi Input Multi Output*) dont on possède plusieurs antennes à l'émission et à la réception.

2.3. Beamforming

La technique du Beamforming est utilisée dans les antennes intelligentes pour transmettre et recevoir des signaux dans des systèmes massifs MIMO. Utilisée dans le but d'augmenter les performances du système sans fil, d'améliorer la qualité de la transmission d'une façon à couvrir une plus grande portée ainsi que le débit et d'annuler les interférences du bruit et renforcer le signal d'intérêt [6]. Cette technique est essentiel pour la technique MU-MIMO (ou MIMO multi-utilisateurs) où plusieurs utilisateurs peuvent communiquer chacun avec plusieurs antennes du routeur et l'utilisation du Beamforming assure que la communication du routeur cible efficacement chaque client connecté [8]. Différentes techniques de formation de faisceau ont été étudiées, telles que la formation de faisceau analogique, numérique et hybride [6].

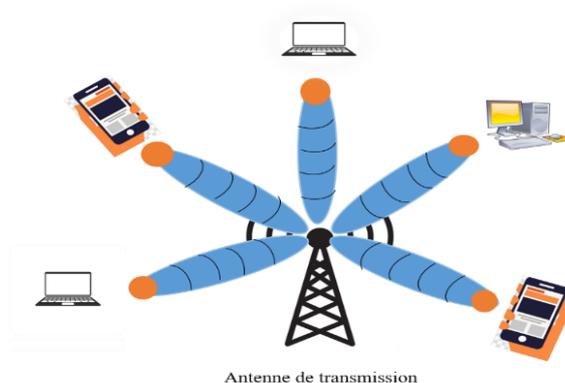


Figure I.5 : Beamforming.

3. Techniques d'accès au canal

Les réseaux cellulaires mettent en œuvre un ensemble de différentes techniques d'accès et Avec l'arrivée de la 4G, l'accès multiple par répartition orthogonale de la fréquence OFDMA a fait son apparition, et afin de satisfaire les exigences élevées de la nouvelle génération 5G de nouvelles techniques telles que NOMA et OMA seront adoptées.

3.1. Modulations OFDM et OFDMA

La modulation OFDM (*Orthogonal Frequency Division Multiple*) est une modulation multi-porteuse orthogonale et une technique qui s'avère être efficace pour lutter contre les évanouissements par trajets multiples dans les canaux sans fil. Son principe est de transmettre un bloc de donnée binaire sur un grand nombre de porteuses en même temps, Pour que les fréquences des porteuses soient les plus proches possibles et ainsi transmettre le maximum d'informations sur une portion de fréquences données [4]. L'OFDM est intéressante car elle offre des solutions relativement faciles à certains problèmes difficiles rencontrés lors de l'utilisation de schémas de modulation à porteuse unique sur des canaux sans fil. Le plus grand avantage de l'utilisation de l'OFDM est que la modulation de sous-porteuses orthogonales très rapprochées divise la largeur de bande disponible en une série de sous-bandes étroites. Mais les spectres des signaux correspondant aux différentes sous-porteuses se chevauchent en fréquence. Le chevauchement spectral donne une forme d'onde qui utilise la largeur de bande disponible avec une très grande efficacité [9]. Un modulateur OFDM peut être mis en œuvre comme une transformée de Fourier discrète inverse (IDFT) à N points sur un bloc de N symboles d'information, suivie d'un convertisseur numérique-analogique (CNA) sur les échantillons IDFT.

L'IDFT du bloc de données est :

$$S_n = \sum_{k=0}^{N-1} s_k \exp\left\{\frac{j2\pi nk}{N}\right\} \quad n=0,1,\dots, N-1 \quad (\text{I.3})$$

Dont $\{S_k, k = 1, \dots, N\}$: représente un bloc de N symboles de données complexes choisis dans une constellation de signaux appropriée telle que la modulation d'amplitude en quadrature (QAM) ou la modulation par déplacement de phase (PSK) [9]. Le signal OFDM (modulé) peut être exprimé comme suit :

$$S(t) = \sum_{k=0}^{N-1} S_k e^{j2\pi f_k t} = \sum_{k=0}^{N-1} s_k \varphi_k(t), \quad \text{pour } 0 \leq t \leq T_s \quad (\text{I.4})$$

$$\text{Avec} \quad f_k = f_0 + k\Delta f \quad (\text{I.5})$$

S_k : Les symboles complexes à transmettre par OFDM.

Et :

$$\varphi_k(t) = \begin{cases} e^{j2\pi f_k t} & \text{si } 0 \leq t \leq T_s, \text{ avec } k = 0, 1, \dots, N-1 \\ 0 & \text{sinon} \end{cases} \quad (\text{I.6})$$

Dont T_s est durée d'un symbole, et Δf l'espace de sous canal de l'OFDM.

Pour que le récepteur puisse démoduler le signal OFDM, la durée du symbole doit être suffisamment longue pour que $T_s \Delta f = 1$ qui est également appelée condition d'orthogonalité.

OFDMA (*orthogonal frequency-division multiple access*) est une technique de multiplexage d'accès orthogonale et de codage de données, qui est une extension simple de l'OFDM dans un environnement multi-utilisateurs [10] et un moyen efficace de résoudre le problème de l'interférence inter-symbole causée par les canaux à évanouissement en fréquence. Utilisée dans les réseaux de téléphonie mobile 4G et 5G, son principe consiste à répartir sur un grand nombre de sous-porteuses le signal numérique que l'on veut transmettre, permettant d'avoir sur chaque canal un débit plus faible pour un même débit global et donc un temps d'émission de chaque symbole plus long (66.7 us pour le LTE). L'OFDMA divise la largeur de bande disponible en un certain nombre de sous-porteuses également espacées et transporte une partie des informations d'un utilisateur sur chaque sous-porteuse.

Toutefois, l'OFDMA possède une propriété importante, à savoir que chaque sous-porteuse est orthogonale à toutes les autres sous-porteuses. Les signaux des différentes sous-porteuses se chevauchent mais grâce à l'orthogonalité ils n'interfèrent pas entre eux. [4] l'avantage attirant de l'OFDMA est sa robustesse aux canaux à trajets multiples et l'égalisation simplifiée [10].

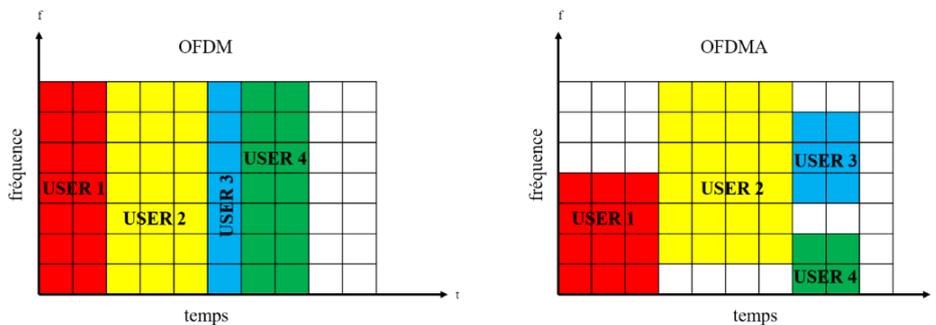


Figure I.6 : Principe des techniques OFDM et OFDMA.

3.2. Techniques d'accès OMA et NOMA

OMA (*Orthogonal Multiple Access*) est un système classique d'accès multiple orthogonal dont on attribue à des utilisateurs des ressources radios qui sont orthogonales dans le domaine du temps, des fréquences ou des codes. Dans l'OMA en raison de l'attribution orthogonale des ressources il n'y a pas d'interférences entre les utilisateurs multiples de sorte qu'une simple détection à utilisateur unique peut être utilisée pour séparer les signaux des différents utilisateurs. L'OMA à répartition orthogonale de la fréquence, combine les domaines temporel et fréquentiel et a été intégrée à la LTE (*Long Term Evolution*) [6].

Pour la 5G New Radio, l'adoption d'une OMA basée sur le multiplexage par répartition en fréquence orthogonale (OFDM) implique des différences significatives liées, entre autres, au fait que la NR prend en charge différentes numérations en fonction de l'espacement des sous-porteuses sélectionnées. Récemment, le NOMA (*Non Orthogonal Multiple Access*) a été étudié pour traiter les problèmes de l'OMA. NOMA est une technique d'accès non orthogonale qui est basée sur le multiplexage en puissance, son principe se résume à servir plusieurs utilisateurs au sein de la même ressource (temps, fréquence, espace ou code) et qui permet de contrôler les interférences et l'amélioration de l'efficacité spectrale [6].

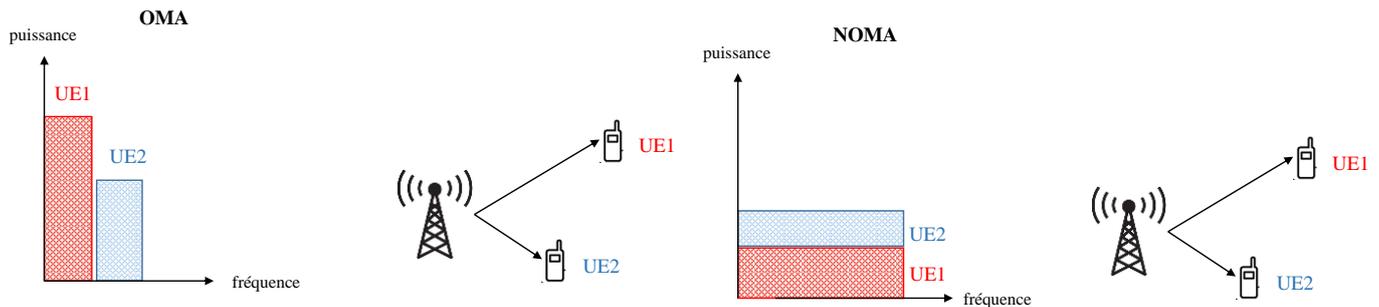


Figure I.7 : Principe des techniques OMA et NOMA.

On peut classer les systèmes NOMA en deux catégories :

- **Domaine de puissance NOMA (PD-NOMA)** : ou NOMA réalise un multiplexage basé sur différents niveaux de puissance.
- **Domaine de code NOMA (CD-NOMA)** : ou NOMA réalise un multiplexage basé sur différents codes.

• NOMA pour la liaison descendante

Dans la liaison descendante de NOMA, la station de base (BTS) superpose les formes d'onde d'information pour ses utilisateurs entretenus. Chaque équipement utilisateur (UE) utilise le SIC pour détecter leurs propres signaux. Dans cette liaison on alloue plus de puissance à l'UE situé plus loin de la BTS et la plus faible puissance à l'UE le plus proche de la BTS. Tous les UE reçoivent le même signal qui contient les informations pour tous les utilisateurs, chaque UE décode d'abord le signal le plus fort, puis soustrait le signal décodé du signal reçu. Le récepteur SIC répète la soustraction jusqu'à ce qu'il trouve son propre signal. Le signal transmis par la BTS peut être écrit comme :

$$x(t) = \sum_{k=1}^K \sqrt{a_k} P_{TX_K} \quad (I.7)$$

Avec $X_K(t)$ représente l'information d'utilisateur K, a_k le coefficient d'allocation de puissance pour l' UE_K et P_T la puissance totale disponible à la BS. La puissance allouée à chaque UE_K devient alors :

$$P_K = P_K \times P_t \quad (I.8)$$

Le signal reçu au niveau de l' UE_K est :

$$Y_K(t) = X(t)g_k + W_g(t) \quad (I.9)$$

Dont g_k est le facteur d'atténuation de canal pour la liaison entre la BS et l'UE $_K$, $W_K(t)$ le bruit gaussien blanc additif au niveau de l'UE $_K$ avec une moyenne de zéro et une densité : $N_0(W / Hz)$.

Le rapport signal sur bruit (SNR) pour UE $_K$ peut-être écrit comme :

$$SNR_K = \frac{P g_k^2}{N_0 W + \sum_{i=1}^{K-1} P_i g_i^2} \quad (I.10)$$

Où : W est la bande passante de transmission.

En supposant une annulation parfaite, le SNR pour UE le plus proche devient :

$$SNR_1 = \frac{P g_1^2}{N_0 W} \quad (I.11)$$

Le débit pour chaque UE s'écrit :

$$D_K = W \log \left(1 + \frac{P_K g_K^2}{N + \sum_{i=1}^{K-1} P_i g_i^2} \right) \quad (I.12)$$

En OFDMA, le débit s'écrit :

$$D_K = W_K \log_2 \left(1 + \frac{P_K g_K^2}{N_K} \right) \quad (I.13)$$

Avec $W_K = \frac{W}{K}$ (I.14)

Et $N_k = N_0 W_K$ (I.15)

La somme des capacités OFDMA et NOMA est :

$$D_T = \sum_{K=1}^K D_K \quad (I.16)$$

• NOMA pour la liaison montante

Dans la liaison montante, chaque utilisateur transmet son signal à la BTS d'où des itérations SIC sont effectuées afin de séparer les signaux des utilisateurs.

Le signal reçu à la BS comprend tous les signaux des utilisateurs et s'écrit comme suit :

$$y(t) = \sum_{k=1}^k x_k(t) g_k + w(t) \quad (I.17)$$

Avec g_k qui est le coefficient de canal entre la BTS et l'UE $_k$, $x_k(t)$ l'information de l'utilisateur k et $w(t)$ le bruit AWGN à la BS avec une moyenne zéro et une densité spectrale de puissance $N_0(W/Hz)$.

Le premier signal à détecter provient de l'utilisateur le plus proche de la BTS son rapport signal sur bruit sera donc :

$$SNR_1 = \frac{P g_1^2}{N + \sum_{i=2}^k P g_i^2} \quad (I.18)$$

Où : P est la puissance d'émission des UE.

$$N = N_0 W \quad (I.19)$$

Le dernier signal à détecter provient de l'utilisateur le plus éloigné de la BS le rapport signal sur bruit sera donc :

$$SNR_k = \frac{P_k g_k^2}{N} \quad (I.20)$$

Le rapport signal sur bruit pour tout utilisateur k vaut :

$$SNR_k = \frac{P_k g_k^2}{N + \sum_{i=k+1}^K P_i g_i^2} \quad (I.21)$$

Le débit de transmission pour chaque utilisateur dans un système NOMA est défini comme suit :

$$D_K = W \log_2 \left(1 + \frac{P_K g_K^2}{N + \sum_{i=K+1}^K P_i g_i^2} \right) \quad (I.22)$$

Dans un système OFDMA, la bande passante et la puissance sont partagées de manière égale. Le débit de transmission est défini comme suit :

$$D_K = W_K \log_2 \left(1 + \frac{P_K g_K^2}{N_K} \right) \quad (I.23)$$

La relation de la capacité totale pour NOMA et OFDMA s'écrit comme suit :

$$D_T = \sum_{k=1}^K D_K \quad (I.24)$$

4. Type de canaux

Les modèles de canaux sans fil sont des modèles qui décrivent l'effet de propagation par trajets multiples dont on cite le modèle de Rayleigh, le modèle de Rice et le modèle Nakagami. Ces modèles décrivent les fonctions de distribution de probabilité de la puissance du signal dans le site récepteur. Le modèle de Rayleigh suppose de nombreux chemins aléatoires sans ligne de visée, tandis que le modèle de Rice suppose en plus un chemin en ligne de visée. Le modèle de Nakagami est un modèle empirique couvrant une large gamme de modèles d'évanouissement par trajet multiple grâce à l'introduction du paramètre 'm'. Il peut donc donner des résultats plus précis pour décrire le canal lorsqu'il est utilisé pour modéliser les systèmes de communication [16].

4.1. Canal de Rayleigh

La loi de Rayleigh est une loi de probabilité de densité baptisée d'après Lord Rayleigh. Le modèle d'évanouissement de Rayleigh part du principe que l'amplitude d'un signal qui a traversé un tel canal de transmission varie de manière aléatoire, ou s'évanouit, selon une distribution de Rayleigh la composante radiale de la somme de deux variables aléatoires gaussiennes non corrélées [13]. La loi de Rayleigh est donnée par :

$$f(x, \sigma^2) = \frac{x}{\sigma^2} \exp\left(\frac{-x^2}{2\sigma^2}\right) ; \text{ Pour } x \in [0, \infty[\quad (I.25)$$

σ^2 : Paramètre de Rayleigh.

4.2. Canal de Rice

La distribution de Rice peut aussi être appelée « *distribution Nakagami-n* » ou « *Nakagami-Rice* », utilisée pour la modélisation d'un canal avec un chemin direct (LOS). Le canal sera composé d'une ligne de visée fixe entre l'émetteur et le récepteur, ainsi que de nombreux composants aléatoires très faibles [14]. Sa fonction de densité de probabilité est définie par :

$$P_{\alpha}(\alpha) = \frac{2(1+n^2)e^{-n^2}}{\Omega} \exp\left[-\frac{(1+n)^2\alpha^2}{\Omega}\right] I_0\left(2n\alpha\sqrt{\frac{1+n^2}{\Omega}}\right), \quad \alpha \geq 0. \quad (I.26)$$

Avec : I_0 qui est La fonction de Bessel, n le paramètre d'évanouissement de Rice qui prend une valeur comprise entre 0 et $+\infty$ et n^2 le paramètre de Rice également noté K , il correspond au rapport entre la puissance de la composante du trajet directe x_d (LOS) avec la puissance de la composante des trajets multiples $2\sigma^2$ (NLOS), et K est défini par [15] :

$$K = n^2 = \frac{x_d^2}{2\sigma^2} \quad (I.27)$$

4.3. Canal de Nakagami

La distribution m-Nakagami fait partie des distributions qui représentent l'évanouissement dans un canal de communication sans fil. [16] ; C'est une distribution de probabilité qui conçoit généralement des évanouissements à petite échelle (un phénomène qui affecte le signal transmis pendant un petit intervalle de temps) [17]. Le modèle du canal Nakagami ou m-Nakagami dont 'm' est défini comme un facteur d'évanouissement de Nakagami est utilisé pour les évanouissements modélisés par multi-trajets. Il s'agit d'un modèle d'évanouissement général qui peut être adapté aux différents modèles de canaux et qui permet de fournir des résultats plus précis pour décrire le canal lors de l'utilisation des systèmes de communication [18]. La loi de distribution m-Nakagami est une loi de distribution de probabilité qui utilise les deux paramètres m et Ω . Sa fonction de densité de probabilité de l'amplitude α par symbole est définie par [17] :

$$P_{\alpha}(\alpha) = \frac{2m^m\alpha^{2m-1}}{\Omega^m\Gamma(m)} \exp\left[-\frac{m\alpha^2}{\Omega}\right] \quad (I.28)$$

Dont $\Gamma(m)$ est la fonction gamma, Ω est la valeur quadratique moyenne et m le paramètre d'évanouissement dont sa valeur est comprise entre 1 et $+\infty$, il est défini par :

$$m = \frac{\Omega^2}{E[(\alpha^2 - \Omega)^2]} \quad (I.29)$$

5. Interférences cellulaires

Les interférences dans les réseaux de communication constituent une nuisance indésirable car elles modifient le signal utile d'une manière perturbatrice lorsqu'il se déplace sur un canal de communication entre la source et le récepteur en ajoutant des signaux non désirés. Plus

précisément, il existe deux types d'interférences : les interférences intracellulaires et les interférences intercellulaires.

5.1. Interférences intercellulaires

Dans ce cas le terminal reçoit un signal perturbé par les interférences dues à la portion des cellules voisines lorsque chacune des stations de base des cellules adjacentes émettent des signaux simultanément [7].

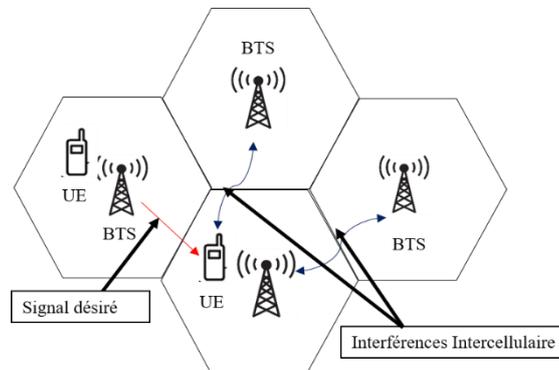


Figure I.8 : Interférences intercellulaires.

5.2. Interférences intracellulaires

Les interférences sont produites dans une même cellule entre deux signaux reçus ou transmis par une même station de base lorsque deux utilisateurs émettent un signal en même temps vers la BTS et qui va recevoir deux signaux perturbés qui correspondent aux signaux émis [7].

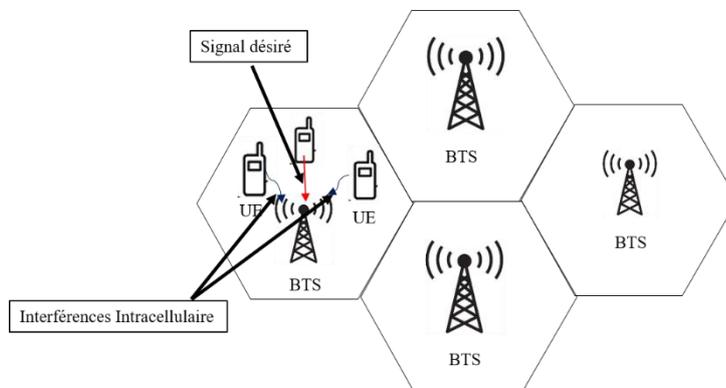


Figure I.9 : Interférences intracellulaires.

5.3. Coordinations d'interférences (FFR, SFR)

Pour éviter l'impact négatif des interférences sur les performances d'un système cellulaire, des techniques de coordination des interférences intercellulaires (ICIC) (*InterCell Interference Coordination*) sont proposées. Elles établissent des restrictions sur l'utilisation des ressources, comme la réutilisation fractionnée des fréquences (FFR), et sur l'allocation de puissance, comme la réutilisation souple des fréquences (SFR) [7]. La réutilisation fractionnée des fréquences FFR (*Fractional Frequency Reuse*) est une technique efficace d'atténuation des interférences pour les réseaux cellulaires modernes en raison de sa faible complexité, de ses faibles exigences de coordination et de sa flexibilité en matière d'allocation des ressources [11]. Dans le mode de

déploiement FFR chaque cellule est divisée en deux zones : le centre qui contient les utilisateurs les plus proches à la station de base, et la bordure qui est la zone qui intègre les UE qui se situent à l'extrémité des cellules dont l'objectifs du FFR est de protéger les ressources radios attribuer à ces utilisateurs et cela en modifiant la distribution des ressources fréquentielles entre les différentes zones des cellules [12]. La réutilisation souple des ressources SFR (*soft frequency reuse*) est une solution ICIC qui est conçue pour atteindre des niveaux de performance comparable à ceux de FFR et contrairement à cette dernière le SFR permet l'utilisation de tout le spectre dans la cellule [12].

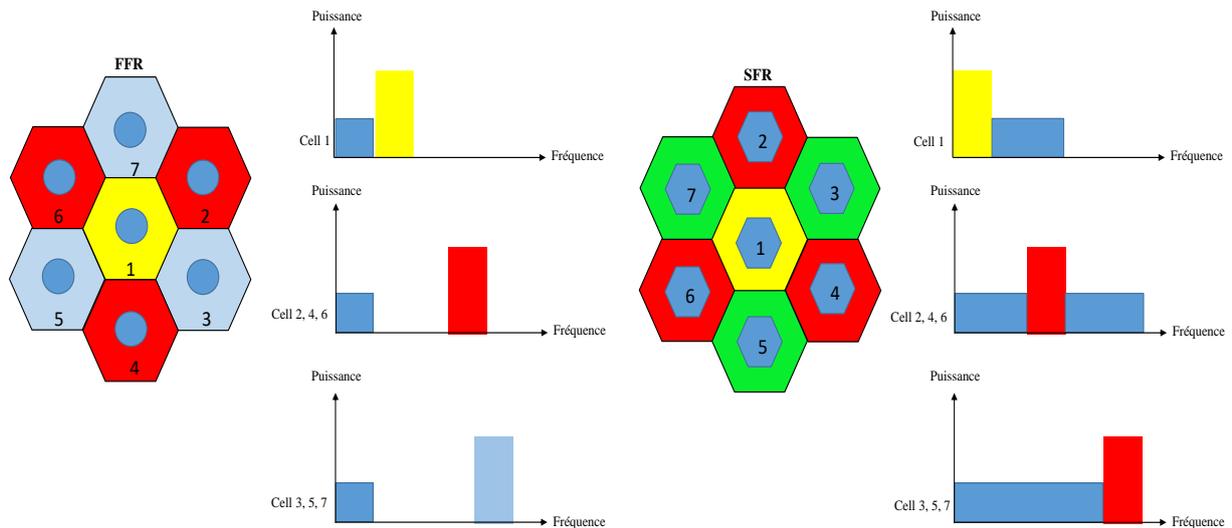


Figure I.10 : Comparaison entre SFR et FFR.

6. Réseaux hétérogènes

Afin de fournir une capacité élevée il est nécessaire de densifier l'accès radio en augmentant le nombre de nœuds. Le déploiement hétérogène consiste à superposer aux macro-cellules (réseau *overlaid*) un réseau de petites cellules (pico-cellules, microcellules) (réseau *underlaid*) comme le montre la figure suivante :

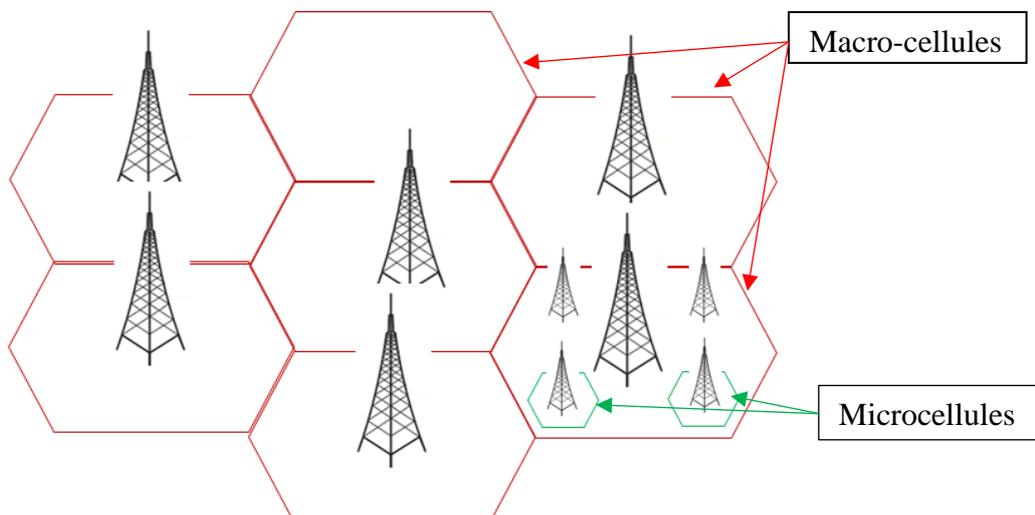


Figure I.11 : Déploiement hétérogène.

Conclusion

Dans ce chapitre nous avons pu comprendre diverses choses sur les réseaux cellulaires mobiles principalement sur les techniques de transmission 'LOS et NLOS' en terme de propagation des données qui connaissent dernièrement une grande révolution et mis en valeur par rapport aux avantages qu'elles apportent. Ainsi que sur les techniques d'accès, OFDM et OFDMA qui a pour but de limiter les problèmes d'interférences et de fading, OMA et NOMA qui sont là pour améliorer les performances du réseau en termes d'efficacité spectrale en expliquant le principe de chacune, de plus on a pu avoir une vue sur certains types de canaux.

CHAPITRE II

Allocation de ressources dans un réseau 5G

Introduction

Les réseaux sans fil de prochaine génération vont être extrêmement complexes en raison de leur hétérogénéité en termes de types d'architectures de réseau qu'ils intègrent, de types et de nombres de dispositifs qu'ils desservent et de types d'applications émergentes qu'ils prennent en charge. Dans ces réseaux hétérogènes et à grande échelle (Het-Nets), l'allocation et la gestion des ressources radio (AGRR) deviennent l'un des principaux défis rencontrés lors de la conception et du déploiement du système. Dans ce contexte, les techniques émergentes d'apprentissage par renforcement profond devraient être l'une des principales technologies habilitantes pour traiter l'AGRR dans les futurs Het-Nets sans fil.

L'objectif de l'allocation de ressources est de garantir une meilleure qualité de service qui est évalué sous trois propriétés, assurer une distribution équitable de ressources, maximiser le débit et minimiser le délai, ainsi, satisfaire tous les utilisateurs.

Dans ce chapitre on va introduire le principe d'allocation de ressource et de puissance ensuite nous parlerons de l'efficacité spectrale et énergétique et enfin nous présenterons le délai de tolérance dans le réseau 5G.

1. Principe d'allocation de ressources

L'allocation et la gestion des ressources radio (AGRR) sont considérées comme l'un des défis essentiels rencontrés dans les réseaux de communication sans fil modernes. L'allocation se fait dans l'eNB par les BTS, cette tâche est très complexe, car elle nécessite de prendre en considération plusieurs facteurs en même temps, en plus elle doit être immédiate (en temps réel). Les algorithmes d'allocation de ressources radio ont pour objectif d'améliorer les performances du système en augmentant l'efficacité spectral et l'équité dans le réseau. Il est donc essentiel de trouver un compromis entre l'efficacité (augmentation en débit) et l'équité entre les utilisateurs [33].

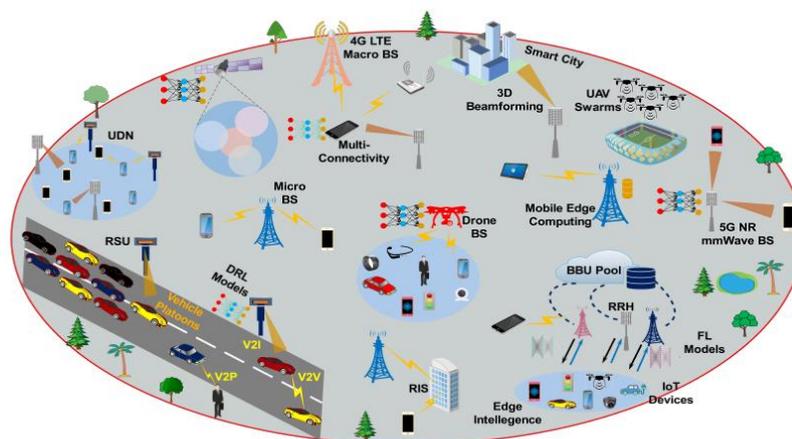


Figure II.1 : Illustration des réseaux sans fil de prochaine génération caractérisés par leur hétérogénéité massive.

1.1. Contraintes d'allocation de ressources

Dans les systèmes de communication sans fil, les signaux sont transmis sous forme d'ondes électromagnétiques entre l'émetteur et le récepteur. Ces transmissions dépendent fortement de changement dans les canaux de communication. En raison de certains phénomènes produits pas ce dernier (le pathloss, le shadowing et le multipath fading), la qualité du signal est fortement dégradée. Cela diminue la capacité de transmission et rend l'allocation de ressource plus compliqué.

- **Les pertes de trajet (pathloss)**

Les pertes de propagation sont généralement exprimées en décibels, elles caractérisent l'atténuation du signal due à la séparation des trajets émetteur récepteur et l'environnement de propagation.

En effet le signal transmis est souvent soumis à plusieurs phénomènes liés à l'environnement de propagation (Réflexion, diffraction, diffusion) [34].

On dispose de deux modèles d'affaiblissement ABG et CI qui sont des modèles génériques de toutes fréquences qui décrivent l'affaiblissement de propagation à grande échelle à toutes les fréquences pertinentes dans un certain scénario. Le modèle CI est facilement mis en œuvre dans les modèles 3GPP existants en apportant une modification très subtile en remplaçant une constante flottante non physique par une constante dépendant de la fréquence qui représente l'affaiblissement sur le trajet en espace libre dans le premier mètre de propagation. L'équation du modèle ABG est donnée par :

$$P_L^{ABG}(f, d)[dB] = 10\alpha \log_{10} \left(\frac{d}{1m} \right) + \beta + 10\gamma \log_{10} \left(\frac{f}{1GHz} \right) + x_\sigma \quad (II.1)$$

Où α et γ sont respectivement des coefficients montrant la dépendance de l'affaiblissement de trajet en fonction de la distance et de la fréquence, β est la valeur de décalage optimisée pour l'affaiblissement de parcours en dB, d est la distance de séparation émetteur-récepteur, f est la fréquence porteuse en GHz et x_σ est l'écart type décrivant les fluctuations du signal à grande échelle par rapport à l'affaiblissement de parcours moyen sur la distance. Et l'équation du model CI est donnée par :

$$P_L^{CI}(f, d)[dB] = FSPL(f, 1m) + 10n \log_{10} \left(\frac{d}{1m} \right) \quad (II.2)$$

D'où n désigne le paramètre unique du modèle, l'exposant d'affaiblissement sur le trajet (PLE), avec $10n$ décrivant l'affaiblissement sur le trajet en dB en termes de décennies de distances commençant à 1m ce qui rend très facile le calcul de la puissance sur la distance d est la distance de séparation Emetteur-Recepteur, et $FSPL(f, 1m)$ désigne l'affaiblissement sur le trajet en espace libre en dB à une distance de séparation Emetteur-Recepteur de 1m à la fréquence porteuse d'où :

$$FSPL(f, 1m)[db] = 20 \log_{10} \left(\frac{4\pi}{c} \right) \quad (II.3)$$

Avec c : la vitesse de la lumière dans le vide.

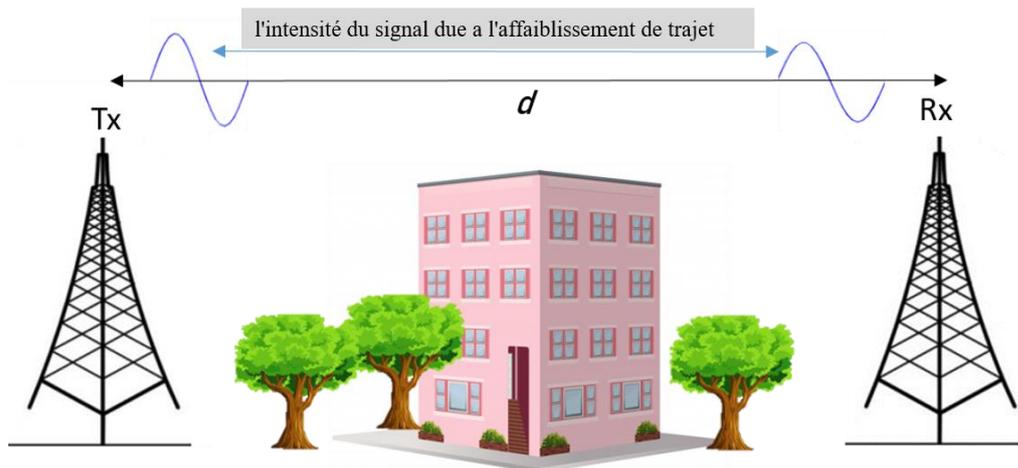


Figure II.2 : Modèle de propagation par perte de chemin dans les communications mobiles.

- **L'effet de masque (shadowing)**

L'effet de masquage ou zone d'ombre (*shadowing*) est responsable des variations à long terme de la puissance moyenne au récepteur. Il est dû à la présence d'obstacles importants entre l'émetteur et le récepteur (immeubles ou forêts). L'onde électromagnétique traverse ces obstacles, mais subit une atténuation de puissance en fonction de la nature du matériau Traversé [35].

L'ampleur de l'évanouissement dû à l'ombre est analysée et modélisée en fonction de la distance de séparation Emetteur-Récepteur à partir des mesures en utilisant les modèles d'affaiblissement sur le trajet à pente unique CI pour l'affaiblissement sur le trajet à une seule pente. La relation entre l'ampleur de l'évanouissement dû aux ombres et la distance de séparation Emetteur-Récepteur est modélisée comme suit :

$$SF[dB] = A \times d + B \quad (II.4)$$

D'où SF représente l'ampleur de l'évanouissement de l'ombre, A Reflète le taux de variation de SF sur la distance, d est la distance de séparation 3D T-R en mètres et B l'intercepte déterminé par ajustement linéaire sur SF.

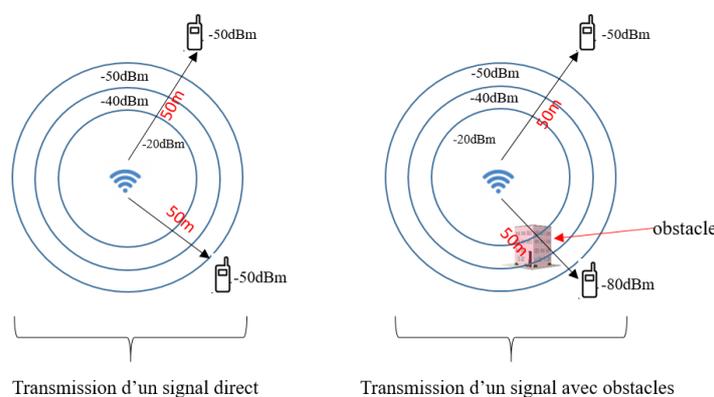


Figure II.3 : Comparaison entre une transmission du signal direct et une transmission avec obstacles.

- **Evanouissement du signal (fading)**

Le signal reçu par une antenne est déformé par le canal de propagation, c'est-à-dire par l'environnement qui sépare l'antenne d'émission à l'antenne de réception. D'une part, le canal atténue la puissance du signal émis en fonction de la distance, mais d'autre part, les transmissions sur le canal mobile dans des environnements plutôt urbains (présence de nombreux bâtiments) ou intérieurs (murs, meubles) provoquent des multi-trajets ce qui génère en réception un évanouissement du signal. Cet évanouissement (*fading*) est un effet aléatoire et souvent résultant de trajets multiples. C'est un phénomène qui perturbe fortement des données d'un système de communication [35].

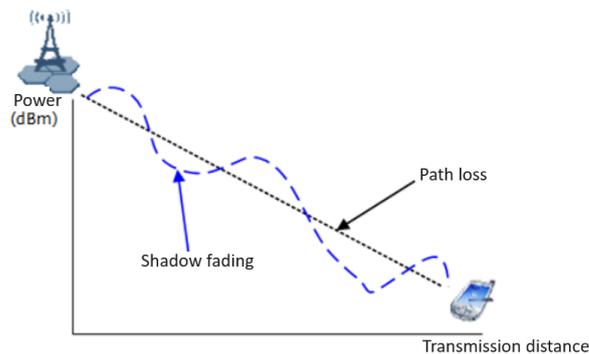


Figure II.4 : Modèle de propagation d'un signal subissant un évanouissement.

On distingue plusieurs types d'évanouissement, les évanouissements à moyenne échelle, dont l'origine est la présence de zones d'ombre, influent sur la distribution de la puissance moyenne reçue. Les évanouissements à petites échelles provoquent une variation rapide du signal reçu. Cette variation est importante sur des faibles distances, sur des faibles temps et sur des faibles bandes de fréquence. Les principales sources d'évanouissement sont les perturbateurs entre l'émetteur et le récepteur créant différentes interactions sur l'onde comme la réflexion, la réfraction, la diffraction et la diffusion.

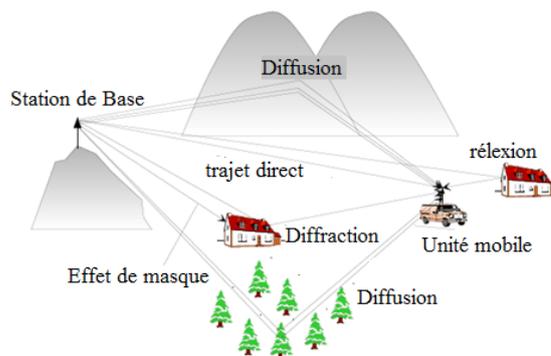


Figure II.5 : Les mécanismes de propagation d'un signal radio-mobile.

1.2. Les ordonnanceurs d'allocation de ressources et leurs principes

Pour répondre à la demande croissante des réseaux sans fil, les technologies du multiplexage OFDMA et NOMA seules ne suffisent pas. En fait elles doivent être combinées avec des

algorithmes d'allocation de ressources efficaces appelées ordonnanceurs qui vise à améliorer les performances du système. Il doit être implémenté au niveau de la station de base qui est responsable de l'établissement du lien entre les buffers et l'interface radio et d'allocation de ressource radio. Son but est de trouver une allocation de sous-porteuses, qui permet à chaque utilisateur de satisfaire ses exigences de débit, sans compromettre l'équité. Il utilise d'une manière intelligente les informations sur l'état de canal CQI (*Channel Quality Indicator*) afin d'adapter le débit de chaque transmission à la variation de l'état du canal radio. Ces informations sont supposées être connues à la fois dans la station mobile et dans la station de base. L'attribution des sous-porteuses est effectuée dans la BS et les utilisateurs sont informés des sous-porteuses choisies pour eux [36].

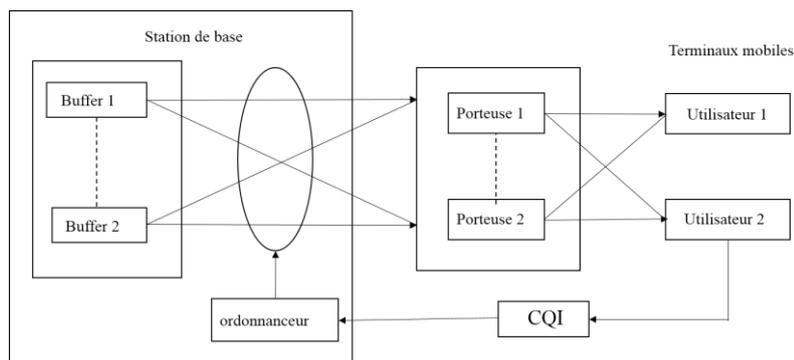


Figure II.6 : Fonctionnement de l'ordonnanceur pour les réseaux radios mobiles.

1.3. Les objectifs d'allocation de ressource

L'objectif principal de toutes les stratégies d'allocation de ressources est d'optimiser l'utilisation du spectre fréquentiel qui est à la fois rare et parfois encombré, les objectifs attendus de cette étude sont les suivant :

- **Maximiser le débit** : il est nécessaire de garantir un débit maximum à chaque utilisateur du réseau néanmoins aujourd'hui ce paramètre est largement limité par le réseau d'accès de type radio et couteuse.

- **Optimiser l'équité** : L'équité d'un système, ne s'exprime pas toujours en termes de valeur quantitative. Il existe plusieurs types d'équité, l'équité en bande et l'équité en débit. l'équité optimale ne garantit pas à tous les utilisateurs d'avoir le même nombre de porteuse et le même débit mais plutôt de leur assurer le même degré de satisfaction.

- **Minimiser le délai** : parmi les attentes de la 5G, fournir un débit élevé avec un très faible délai pour chaque utilisateur et chaque service, afin de maintenir une meilleure qualité de service.

1.4. Allocation de puissance

Dans les réseaux sans fil, l'allocation de puissance est une technique efficace pour prolonger la durée de vie des terminaux du réseau. En général, l'allocation optimale de la puissance améliore

l'efficacité des systèmes sans fil. Lorsque l'allocation de puissance est correctement effectuée, les informations de la source peuvent atteindre la destination efficacement [32].

2. Efficacité spectrale

L'efficacité spectrale décrit la quantité de données transmises sur un spectre ou une bande passante donnée avec un minimum d'erreurs de transmission. Également connue sous le nom d'efficacité de la bande passante, l'efficacité spectrale d'un réseau cellulaire équivaut au nombre maximal de bits de données qui peuvent être transmis à un nombre donné d'utilisateurs par seconde tout en maintenant une qualité de service acceptable [24].

L'efficacité spectrale dans les vitesses de communication sans fil est affectée par le nombre d'utilisateurs accédant simultanément au réseau. Le taux de transfert des données dépend de la bande passante du dispositif de transmission et du signal transmis ou du rapport signal/bruit. Lorsque le rapport signal/bruit est amélioré, cela augmente également l'efficacité spectrale et la capacité du canal. En d'autres termes, il faut envoyer davantage de données sur le spectre disponible pour l'utiliser efficacement [24]. Elle est exprimée par le rapport entre le débit du réseau et la bande passante du signal par : [37]

$$ES = \log_2 \left(1 + \frac{hP}{N_0 + \sum_K h_k P_k} \right) \quad (\text{bit/s/Hz}) \quad (\text{II.5})$$

Avec P la puissance de transmission, N_0 la densité spectrale du bruit, K le nombre d'utilisateurs et H est une matrice donnée par :

$$h = |g|^2 G = |g|^2 \beta d^{-\alpha} \quad (\text{II.6})$$

Dont g est l'évanouissement par trajet multiple entre l'émetteur et le récepteur et G le gain du canal lié à cette distance.

α et β Sont respectivement le coefficient de perte, et exposant de perte sur le chemin et d est la distance entre l'émetteur et le récepteur.

2.1. Spectre de fréquence 5G

Il est prévu que la 5^{ème} génération supporte des vitesses de connexion à un très haut débit largement plus rapide et une utilisation des données nettement plus intensive que les générations précédentes et cela dépendra de la bonne quantité et le bon choix des fréquences, sous les bonnes conditions [26]. Pour répondre à ces exigences la 5G doit avoir accès à des fréquences "élevées", "moyennes" et "basses", en exploitant les caractéristiques spécifiques des différentes parties du spectre de fréquences comprises entre 1 et 6 GHz en combinaison avec des fréquences inférieures à 1 GHz et supérieures à 6 GHz [27]. Les fréquences en-dessous de 1 GHz permettront une couverture large dans les zones urbaines, périurbaines et rurales, et la mise en place de services d'Internet des Objets (IoT), Les fréquences de 1 à 6 GHz proposent un bon compromis entre la

couverture et la capacité. Ceci inclut les fréquences dans la plage 3,3-3,8 GHz qui sont visées pour accueillir nombre des premiers services 5G. Les fréquences au-dessus de 6 GHz sont nécessaires pour atteindre les grandes vitesses très hautes débit envisagées pour la 5G. [26] Les basses fréquences (inférieures à 2 GHz) resteront essentielles pour étendre l'expérience du haut débit mobile 5G à de vastes zones et à des environnements intérieurs profonds ; les scénarios d'utilisation mMTC et URLLC bénéficieront aussi largement de la couverture étendue des basses fréquences [27].

Les bandes de basses fréquences disponibles (700, 800, 900, 1800 et 2100 MHz) peuvent être exploitées pour le partage du spectre des liaisons montantes LTE/NR en combinaison avec NR afin de permettre aux opérateurs d'assurer un déploiement plus rapide et rentable. Les hautes fréquences (au-dessus de 6 GHz) s'avéreront indispensables pour fournir une capacité supplémentaire et offrir les débits de données extrêmement élevés requis par certaines applications 5G eMBB. L'attribution d'une large bande passante de spectre contiguë dans chaque couche réduit la complexité du système associée à l'agrégation de porteuses, ce qui améliorera l'efficacité énergétique [27].

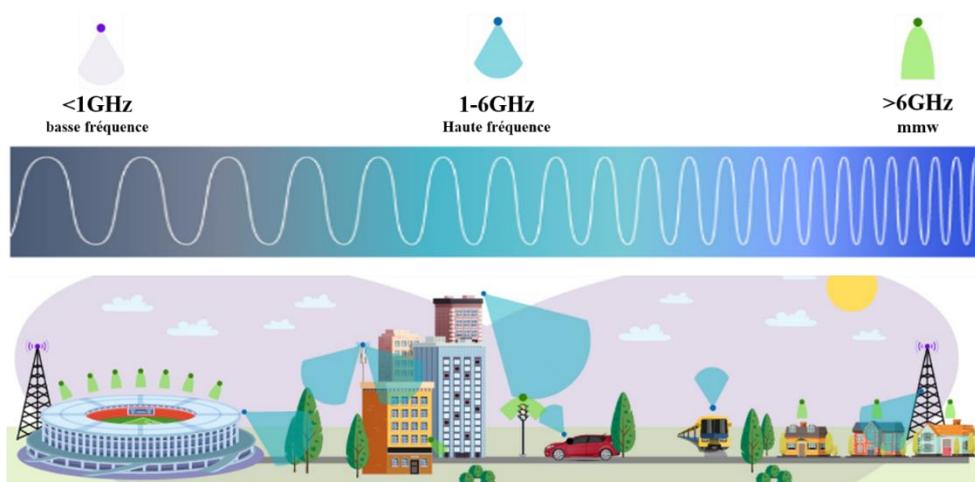


Figure II.7 : Spectre de fréquences allouées à la 5G.

2.2. Amélioration de l'efficacité spectrale

Il existe de nombreuses façons d'améliorer l'efficacité spectrale. Cependant, il ne suffit pas d'ajouter des antennes supplémentaires, car cela peut potentiellement augmenter les interférences. Pour éviter toute interférences potentielles, des systèmes massifs à entrées et sorties multiples (MIMO) utilisant des techniques de formation de faisceau peuvent être utilisés pour améliorer l'efficacité des transmissions. Qui est une technique de gestion des fréquences radio (RF). Ces techniques de formation de faisceaux sont également essentielles aux réseaux 5G car elles permettent de diriger et d'ajuster les ondes radio pour les rendre plus fortes et plus ciblées. Elles peuvent également étendre la portée des transmissions RF et transmettre des données directement

à l'utilisateur final [24]. L'efficacité du spectre dans la 5G crée des opportunités pour les opérateurs de réseaux qui utilisent des petites cellules pour densifier leurs réseaux et réutiliser le spectre plus fréquemment. Cette approche est de plus en plus populaire car elle offre des améliorations de capacité et permet une transition transparente de la 4G à la 5G. Plus la taille du canal augmente, l'efficacité spectrale augmente aussi. Une large canalisation permet également d'obtenir des attributs cruciaux de la 5G, comme une latence à un chiffre et des vitesses jusqu'à 100 fois plus rapides que celles des réseaux 4G [24].

Trois approches clés permettent d'atteindre les objectifs de conception de la 5G. Il s'agit des réseaux ultra-denses, de la bande mmWave à large bande passante et du MIMO massif qui permet une grande efficacité spectrale [25]. La densification du réseau peut être obtenue par un déploiement massif des macro-cellules, des microcellules et des femto-cellules. Ces cellules améliorent la capacité du réseau, l'efficacité énergétique et les performances de couverture [24].

Pour améliorer encore plus les performances du réseau, on peut utiliser des Het-Nets sans fil comprenant des têtes radio distantes et des relais sans fil. Cette approche de relais et de communication multi-sauts promet d'être un élément essentiel de l'architecture sans fil de la 5G [25].

3. Efficacité énergétique

La consommation d'énergie est devenue une préoccupation majeure dans la conception et l'exploitation des systèmes de communication sans fil. Les réseaux de communication sont principalement conçus dans le but d'optimiser les mesures de performance telles que le débit, la latence et l'efficacité énergétique qui est apparue comme une nouvelle figure de proue [33].

L'efficacité énergétique est l'un des principaux critères de conception des réseaux sans fil et elle est traitée comme l'une des solutions proposées pour l'amélioration des communications mobiles sans fils future tel que la 5G. Du côté des opérateurs, l'utilisation de la technologie Massive MIMO, une nouvelle organisation d'accès au réseau mobile ainsi qu'une concentration des équipements vont permettre une meilleure efficacité énergétique [19]. On peut définir l'efficacité énergétique comme suit [37] :

$$EE = \frac{ES}{(P_i + P_c)} \quad (bits/J) \quad (II.7)$$

Avec, P_c et P_i sont respectivement la puissance consommée par le circuit transmetteur et la puissance de transmission et ES l'efficacité spectrale.

L'efficacité spectrale (ES) et énergétique (EE) seront également un enjeu clé dans les futurs réseaux cellulaires 5G. Les efforts les plus importants pour augmenter l'efficacité énergétique d'un

réseau sans fil se concentrent sur le réseau d'accès, car il consomme la plus grande partie de l'énergie.

3.1. Efficacité énergétique dans les réseaux 5G

Les réseaux de communication 5G se caractérisent par un certain nombre de nouvelles fonctionnalités, telles que les fréquences d'ondes millimétriques, les réseaux d'antennes massives, la formation des faisceaux et les cellules denses. Avec la formation de faisceau le réseau 5G consomme 4 fois moins d'énergie que le réseau de référence 4G, tout en offrant une capacité plus élevée et une couverture identique. En revanche, il nécessite le déploiement de 15% de stations de base supplémentaires. [21]

Le secteur des télécommunications a pris les devants et a joué un rôle exemplaire, en fixant des exigences strictes en matière d'efficacité énergétique et de consommation de ses réseaux, ainsi qu'en matière de sécurité. Dont la nouvelle technologie 5G permet un nouveau type de mise en réseau capable de connecter pour la première fois tout le monde et tout ce qui existe, y compris les machines, les objets et les appareils, grâce à une capacité plus élevée, une latence plus faible, une fiabilité améliorée et un plus grand nombre de connexions prises en charge. Ces nouvelles capacités de communication 5G aident les industries actuelles et les nouvelles formes d'entreprises à mettre en œuvre de nouveaux processus avec une efficacité plus importante, en favorisant une utilisation plus souple, plus adaptée et plus efficace des ressources [20].

3.2. L'enjeu de l'efficacité énergétique de la 5G

Dans le secteur de télécommunication l'effet de la 5G s'expose à de nombreux défis car accroître l'efficacité énergétique et réduire la consommation via un réseau de télécommunication entraîne une croissance importante de l'utilisation des données dans certains scénarios, et une nécessité d'une mise en réseau plus sophistiquée pour répondre aux exigences de faible latence, de haute fiabilité et de grand volume de connexions de données dans certains autres [20].

3.3. Sources de consommation d'énergie

Les sources de consommation d'énergie, en ce qui concerne le fonctionnement des réseaux, peuvent être classées en deux types : celles liées à la communication et celles liées au calcul. La communication implique l'utilisation de l'émetteur-récepteur au niveau des nœuds source et destination. L'émetteur est utilisé pour envoyer des commandes, des demandes de route et des réponses, ainsi que des paquets de données provenant de l'émetteur ou acheminés par celui-ci. Le récepteur est utilisé pour recevoir les paquets de données et de contrôle. La compréhension des caractéristiques de puissance de la radio mobile utilisée dans les dispositifs sans fil est importante pour la conception efficace des protocoles de communication. La radio mobile typique peut exister en trois modes : transmission, réception et veille. La consommation d'énergie est maximale en mode de transmission et minimale en mode de veille. [23]

4. Délai de tolérance

Un service tolérant aux retards est un service dont la longueur d'un bloc de code du canal être est suffisamment longue, où le débit moyen de chaque utilisateur s'approche de la capacité de Shannon, [28] on a :

$$\bar{R} = NEg \left[w \ln \left(1 + \frac{\alpha g P}{N_0 N_T N w} \right) \right] \quad (bit/s) \quad (II.8)$$

Dont α et g sont respectivement le gain du canal à grande échelle et le gain du canal à petite échelle sur le $n^{i\text{ème}}$ sous-canal, P est la puissance d'émission, N_0 est la densité spectrale du bruit unilatéral, en fin N_T et N sont respectivement le nombre d'antennes de la station de base et le nombre de sous-porteuses attribuées au $k^{i\text{ème}}$ utilisateur tolérant au retard.

4.1. Service 5G et délai de tolérance

Ce service concerne des procédés et des appareils pour un système 5G qui déploie des services apportant des optimisations à valeur ajoutée basés sur des échanges de données tolérantes au retard (DTD) entre des dispositifs d'extrémités et serveurs de machine à machine (M2M). Selon un mode de réalisation, un appareil peut recevoir, en provenance d'un deuxième appareil via le réseau, des premières informations comportant une configuration de communication tolérante au retard comportant un identifiant et une caractéristique tolérante au retard. L'appareil peut recevoir, en provenance d'un troisième appareil, des données dans une pluralité de lots de données dont chacun comporte l'identifiant et qui peut traiter et acheminer sur la base des premières informations, la pluralité de lots de données et peut comporter un équipement d'utilisateur (EU) [29].

Conclusion

Dans ce chapitre, nous avons présenté le principe d'allocation de ressource dans les réseaux 5G. On a également présenté les ordonnanceurs d'allocation de réseau et leur principe, ensuite on a donné un aperçu sur les objectifs de l'allocation de ressource. On a donné une définition de l'allocation de puissance. Ensuite on a présenté l'efficacité spectrale et énergétique qui sont un moyen efficace pour une meilleure qualité d'allocation de ressource, et on a conclu avec une vue globale sur les délais de tolérance.

CHAPITRE III

Apprentissage par renforcement

Introduction

Le ‘*Reinforcement Learning*’ ou apprentissage par renforcement est une méthode de Machine Learning de plus en plus utilisée. Elle consiste à entraîner des modèles d’intelligence artificielle d’une manière bien spécifique. Il est basé sur l’hypothèse de récompense : l’objectif peut être décrit comme la maximisation des récompenses cumulatives attendues. Dans l’Apprentissage par Renforcement, il n’y a aucun superviseur, mais uniquement un système de récompense. A chaque action effectuée, l’agent reçoit une récompense, qui va affecter ses décisions suivantes. Il est important de comprendre qu’il n’y a pas de données indépendantes et identiquement distribuées comme est le cas pour l’Apprentissage Supervisé. La récompense peut être différée dans le temps, la dimension temporelle a alors un réel impact.

Dans ce chapitre on va définir en premier lieu l’apprentissage par renforcement son principe et sa politique ensuite on introduira la décision de Markov et on aura une vue sur les fonctions de valeur et la fonction de Bellman. En fin on donnera les types du renforcement Learning.

1. L’apprentissage par renforcement et son principe

L’apprentissage par renforcement est une approche informatique dans laquelle un agent essaie de maximiser le montant totale de récompense qu’il reçoit en interagissant avec un environnement complexe et incertain et qui permet de trouver par un processus d’essais et d’erreurs, il utilise cette évaluation pour améliorer les performances du système. Son principe consiste à ce que l’agent peut interagir avec l’environnement par le biais d’actions et l’environnement lui rend une récompense (positive, négative ou nulle) en fonction de celles-ci. Plus formellement, l’apprentissage par renforcement fait référence à une classe de problèmes d’apprentissage automatique, dont le but est d’apprendre, à partir d’expériences, ce qu’il convient de faire en différentes situations, de façon à optimiser une récompense numérique au cours du temps. L’agent va donc chercher à maximiser les récompenses obtenues en effectuant différentes expériences et types d’expériences (exploration, suivi d’une politique, ...) au sein de l’environnement. La nature et le choix des expériences à effectuer à un moment précis de l’apprentissage varient en fonction de l’algorithme d’apprentissage choisi [38].

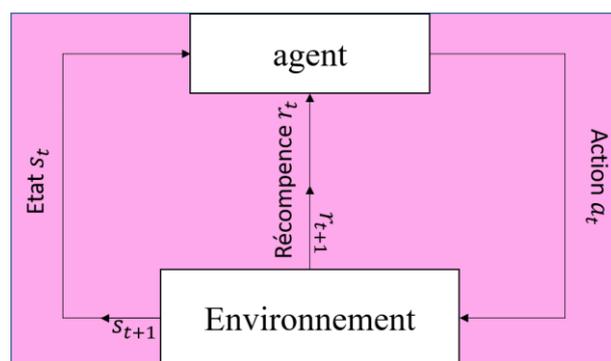


Figure III.1 : Système d’apprentissage par renforcement.

1.1. Méthode de résolution dans AR

Il existe différentes méthodes de résolution selon les hypothèses faites au départ. La programmation dynamique (Bellman) qui s'intéresse à la résolution de la décision de Markov lorsque les modèles de l'environnement 'T' (fonction de transition markovienne) et 'R' (fonction des renforcements Markovienne) sont connus et discrets.

Dans le cas où les fonctions 'T' et 'R' ne sont pas connues. L'agent apprend à travers son interaction avec l'environnement en échantillonnant ces fonctions en des points précis et dans ce cas on aurait deux approches de résolution : la première basées sur des modèles (*model based*) qui cherchent à apprendre le modèle de l'environnement (les fonctions T et R) puis à utiliser, entre autres, la programmation dynamique sur ces approximations. Et celles sans modèles (*model-free*) qui cherchent une politique optimale sans représenter explicitement le modèle de l'environnement [39].

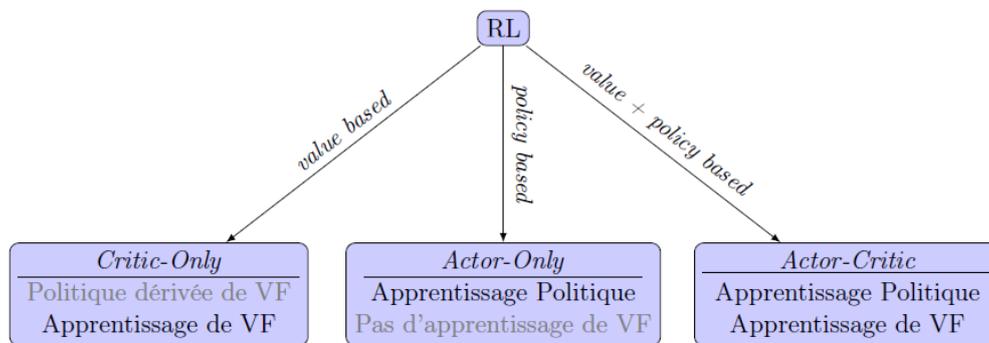


Figure III.2 : Découpage des méthodes en apprentissage par renforcement.

L'acteur est le composant de l'agent qui agit sur l'environnement : il est défini par la politique. Le critique sert à critiquer les actions faites par l'agent ; il s'appuie sur les fonctions de valeurs. Les méthodes (*actor-only*) cherchent à apprendre directement la meilleure politique en estimant des fonctions de valeur mais sans les apprendre, tandis que les méthodes (*critic-only*) apprennent des fonctions de valeurs pour seulement ensuite en déduire des politiques. Enfin, les méthodes (*actor-critic*) apprennent à la fois les fonctions de valeur et les politiques [39].

1.2. Définition des paramètres de l'apprentissage par renforcement

L'algorithme d'AR se déroule en fonction de certains paramètres plus essentiellement :

α qui est le taux d'apprentissage utilisé comme outil pour déterminer la quantité d'informations conserver sur la valeur Q précédemment calculée pour la paire état-action donnée par rapport à la nouvelle valeur Q calculée pour la même paire état-action à un pas de temps ultérieur.

γ qui est le taux d'actualisation qui est utilisé pour la convergence et qui démontre que les valeurs des choses dans le présent sont supérieures à leurs valeurs dans le futur.

ϵ_0 ou ϵ -greedy qui est utilisé Pour déterminer si l'agent choisira l'exploration ou l'exploitation à chaque pas de temps, d'où un nombre aléatoire entre 0 et 1 est généré. Si ce nombre est supérieur

à epsilon, alors l'agent choisira sa prochaine action via l'exploitation, c'est-à-dire qu'il choisira l'action avec la valeur Q la plus élevée pour son état actuel dans la Q-table. Sinon, sa prochaine action sera choisie par exploration, c'est-à-dire qu'il choisira son action au hasard et explorera ce qui se passe dans l'environnement [40].

2. Processus de décision de Markov

Le processus de décision de Markov qui pourra être noté PDM ou Markov Decision Process (MDP) est un modèle stochastique peut être considéré comme une chaîne de Markov contrôlée, c'est à dire à laquelle on ajoute une composante décisionnelle. C'est un outil mathématique qui permet de formaliser l'apprentissage par renforcement. Un MDP vérifie la propriété de Markov qui peut être résumée ainsi : prédire le futur à partir du présent est tout aussi efficace que de le prédire en possédant des informations concernant le passé [38]. Plus formellement la propriété de Markov peut être définie de la manière suivante :

Soient $X_0 \dots X_n$ des variables aléatoires réelles, nous avons donc :

$$\forall n \geq 0, \forall (i_0, \dots, i_{n-1}, i, j) \in E^{n+2}$$

$$P(X_{n+1} = j | X_0 = i_0, \dots, X_{n-1} = i_{n-1}, X_n = i) = P(X_{n+1} = j | X_n = i) \quad (III.1)$$

Un processus de décision de Markov est un quadruplet $\langle S, A, T, R \rangle$ qui peut être défini de la façon suivante : S est l'espace d'états, A est l'espace d'actions et T est une fonction de transition markovienne qui est défini par :

$S \times A \times S' \rightarrow [0,1]$, T (s, a, s') représente la probabilité d'aller de l'état « s » à l'état « s' » en effectuant l'action 'a'.

R : c'est une fonction des renforcements Markovienne qui est défini par :

$$S \times A \times S' \rightarrow R \quad (III.2)$$

R (s, a, s') représente le renforcement obtenu en effectuant l'action « a » et en se retrouvant dans l'état « s ».

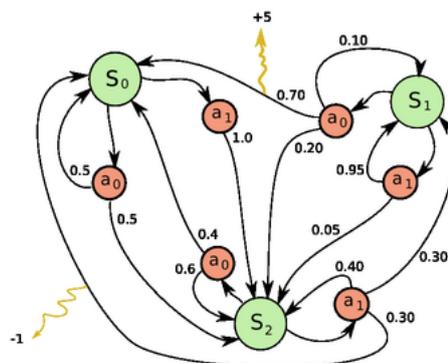


Figure III.3 : Processus de la décision de Markov.

3. Politique d'apprentissage par renforcement

La politique de l'apprentissage par renforcement est définie comme étant une fonction qui permet de choisir une action dans un état donné elle fait partie de l'agent et son objectif est de contrôler l'environnement modélisé comme un processus de décision de Markov $\{S, A, T, R\}$ dont on définit une politique comme étant une fonction qui est générée pour chaque état $s \in S$ et action $a \in A$ et on définit la politique π par la fonction $\pi : S \rightarrow A \rightarrow [0, 1]$. L'application de cette politique se fait tout d'abord en générant un état initial s_0 à partir de la distribution d'état initiale I puis la politique π va suggérer une action $a_0 = \pi(s_0)$ qui va être effectuée en se basant sur T qui est la fonction de transition et R qui est la récompense, une transition est effectuée vers l'état s_1 , avec une probabilité $T(s_0, a_0, s_1)$ et une récompense $r_0 = R(s_0, a_0, s_1)$ est reçue. Ce processus continue jusqu'à ce qu'on atteigne notre objectif [38].

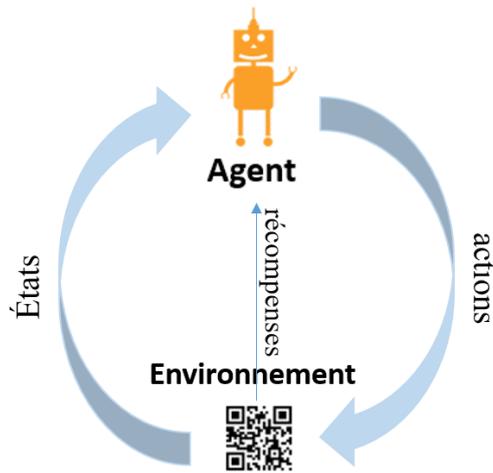


Figure III.4 : Paradigme général de l'apprentissage par renforcement.

3.1. Les différentes politiques

Il existe plusieurs différents types de politique dont on peut citer par exemple [38] :

- **Politique stochastique** : cette politique π associe une densité de probabilité sur l'espace des actions A à un état $s \in S$:

$$\pi : S \times A \rightarrow R^+ \text{ Avec : } \int \pi(a|s) da = 1 \text{ et } s, a \rightarrow \pi(a|s) \quad (\text{III.3})$$

- **Politique déterministe** : Dans ce cas, l'action choisie est toujours la même pour un état s . Une politique déterministe π associe un état $s \in S$ à une action $a \in A$:

$$\pi : S \rightarrow A \text{ Et } s \rightarrow \pi(s) \quad (\text{III.4})$$

- **Politique stationnaire ou markovienne** : Une politique π est dite stationnaire ou markovienne si elle ne dépend pas du temps :

$$\forall t \in \mathbb{N}, \pi_t = \pi \quad (\text{III.5})$$

4. Fonction de valeur

Les algorithmes d'apprentissage par renforcement sont basés sur l'estimation de fonctions de valeur qui sont l'espérance du gain futur dans un état 't' et l'action 'a' qui va être la valeur d'action qui ne dépend pas du temps. Ces valeurs sont un moyen de lier les critères d'optimalité aux politiques. La plupart des algorithmes d'apprentissage pour les processus de décision de Markov calculent des politiques optimales en apprenant une fonction de valeur. Cette dernière représente une estimation de la qualité de l'agent dans un certain état ou de la qualité de l'exécution d'une certaine action dans cet état [38].

4.1. Différentes fonctions de valeurs

Réussir à associer une valeur à un état 's' qui permettrait de quantifier l'intérêt d'être dans cet état est une façon de résoudre des processus de Markov [39].

• **Fonction de valeur d'état** : Étant donnée une politique π , pour tout état $s \in S$, la fonction de valeur d'état V_π est définie telle que [39] :

$$V_\pi(s) = \lim_{N \rightarrow \infty} E_\pi \{ \sum_{t=1}^{\infty} r_t^\pi | s_t = s \} \quad (\text{III.6})$$

Avec E est l'espérance sous la politique π et r_t^π est la récompense obtenue à l'itération 't' après être parti de l'état 's' et en suivant la politique π .

• **Fonction action valeur** : cette fonction notée Q_π , n'évalue plus seulement la pertinence d'un état, mais d'un couple état-action, Étant donnée une politique π , pour tout couple $(s, a) \in S * A$, la fonction action-valeur $Q_\pi(s, a)$ est définie telle que [39] :

$$Q_\pi(s, a) = E_\pi \{ \sum_{k=0}^{\infty} \gamma^k r_{t+k+1} | s_t = s, a_t = a \} \quad (\text{III.7})$$

Dont E est l'espérance sous la politique π , γ est le paramètre du facteur d'actualisation et t est un pas de temps quelconque. La relation entre ces deux fonctions est comme suit :

$$V_\pi(s) = \max Q_\pi(s, \pi(s)) \quad (\text{III.8})$$

Les valeurs d'états et d'actions changent selon la façon dont on choisit nos états et nos actions, et la meilleure action est celle qui va maximiser notre valeur d'état.

4.2. Diagramme de sauvegarde des fonctions de valeur

Les diagrammes de sauvegarde représentent les relations qui constituent la base des opérations de mise à jour ou de sauvegarde qui sont au cœur des méthodes d'apprentissage par renforcement. Ces opérations transfèrent des informations de valeur à un état à partir de ses états successeurs. On utilise des diagrammes de sauvegarde pour fournir des résumés graphiques des algorithmes. Contrairement aux graphes de transition, les nœuds d'état des diagrammes de sauvegarde ne

représentent pas nécessairement des états distincts, par exemple, un état peut être son propre successeur [40].

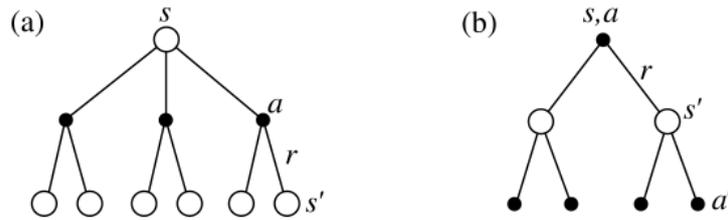


Figure III.5 : Diagrammes de sauvegarde pour $V(\pi)$ et $Q(\pi)$.

5. Fonction de BELLMAN

Une équation de Bellman est une condition nécessaire à l'optimalité associée à la méthode d'optimisation mathématique, elle définit la valeur d'un problème de décision à un moment donné en termes de gain résultant de certains choix initiaux et de la valeur du problème de décision restant qui résulte de ces choix initiaux. L'équation de Bellman est donnée par :

$$Q^*(s, a) = \mathbb{E}\{r_{t+1} + \gamma \max_{a'} Q^*(s_{t+1}, a') | s_t = s, a_t = a\} \quad (\text{III.9})$$

Cette équation déclare que la valeur d'un état sous une politique optimale doit être égale au rendement attendu pour la meilleure action dans cet état [38]. Et indique qu'au temps 't', pour toute paire état-action (s, a) le rendement attendu en commençant par l'état 's', en prenant l'action 'a' et avec la politique optimale par la suite sera égal à la récompense attendue 'R_{t+1}' en plus du rendement maximal qui est réalisable de tout (s', a') où (s', a') est une paire état-action suivante potentielle. En considérant que l'agent suit une politique optimale, le dernier état (s') sera l'état à partir duquel nous pouvons prendre la meilleure action suivante possible (a') au temps 't+1'. L'essentiel est que cette équation peut être utilisée pour trouver le Q* optimal afin de trouver la politique optimale π et ainsi un algorithme d'apprentissage par renforcement peut trouver l'action a qui maximise le $Q^*(s, a)$. C'est pourquoi cette équation a son importance [38].

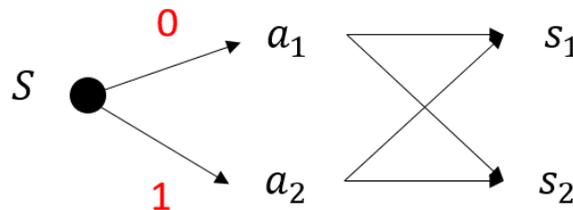


Figure III.6 : Principe de Bellman.

6. Algorithme de renforcement Learning

L'apprentissage par renforcement correspond au cas où l'algorithme apprend un comportement étant donné une observation. L'action de l'algorithme sur l'environnement produit une valeur de retour qui guide l'algorithme d'apprentissage. On dispose de plusieurs algorithmes qui ont été

conçus pour résoudre des processus de décision de Markov on cite par exemple le Q-Learning décentralisé et le Deep Q-network.

6.1. Q-Learning décentralisé

L'apprentissage Q (*Q-learning*) est un algorithme d'apprentissage par renforcement hors politique qui cherche à trouver la meilleure action à entreprendre compte tenu de l'état actuel. Il est considéré comme non conforme à la politique, car la fonction de Q-Learning apprend des actions qui sont en dehors de la politique actuelle, comme prendre des mesures aléatoires, et par conséquent, une politique n'est pas nécessaire. Plus précisément, le Q-Learning cherche à apprendre une politique qui maximise la récompense totale [38]. L'idée de base du Q-Learning est d'estimer de manière à incrémenter les fonctions de valeurs Q pour les actions, en fonction des récompenses et de la fonction de valeur Q de l'agent. On utilise des valeurs Q et un opérateur max intégré sur les valeurs Q de l'état suivant afin de mettre à jour Q_t en Q_{t+1} :

$$Q_{k+1}(s_t, a_t) = Q_k(s_t, a_t) + \alpha(r_t + \gamma \max_a Q_k(s_{t+1}, a_t) - Q_k(s_t, a_t)) \quad (\text{III.9})$$

Avec $\alpha \in [0, 1]$ qui est défini comme le degré d'acceptation de la nouvelle valeur par rapport à l'ancienne.

L'agent fait un pas dans l'environnement de l'état s_t à s_{t+1} en utilisant l'action a_t tout en recevant la récompense r_t . La mise à jour a lieu sur la valeur Q de l'action a_t dans l'état s_t à partir duquel cette action a été exécutée. Pour cela, Q-Table est la structure de données utilisée pour calculer les récompenses futures maximales attendues pour l'action à chaque état, ce tableau nous guidera vers la meilleure action dans chaque état [38].

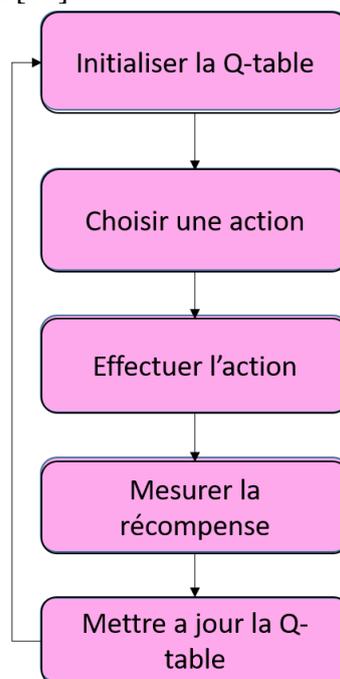


Figure III.7 : Algorithme Q-Learning.

6.2. Deep Q-network

Le deep Q-Learning est une extension du Q-Learning et qui est utilisée lorsque la taille de l'espace des états et actions est très grande.

Les étapes de l'apprentissage par renforcement à l'aide de l'algorithme deep Q-network sont :

- Toute l'expérience passée est stockée dans une mémoire qu'on appelle *Replay Memory*.
- L'action suivante est déterminée par la sortie maximale du réseau Q.
- La fonction de perte est ici l'erreur quadratique moyenne de la valeur Q prédite et de la valeur Q cible qu'on a représentée par Q^* [38].

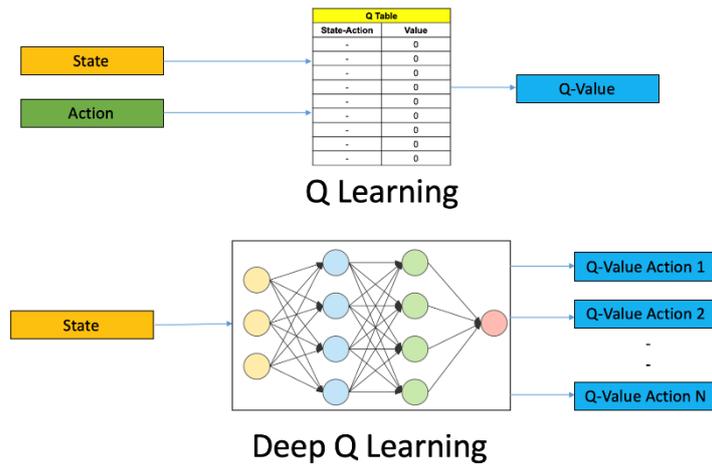


Figure III.8 : Deep Q-network.

Conclusion

Dans ce chapitre, on a donné un aperçu général de l'apprentissage par renforcement, son principe et ses différentes méthodes, ensuite on s'est concentré sur le processus de décision de Markov, et on a présenté les différentes fonctions de valeur et la fonction de Bellman, en fin quelques algorithmes utilisés en AR.

CHAPITRE IV

Simulations et interprétations des résultats

Introduction

L'hétérogénéité est au cœur des réseaux 5G à venir. L'une des principales conséquences de ces déploiements hétérogènes est l'émergence de différents types de réseaux spatiaux, qui nécessitent une analyse continue de leurs paramètres afin de satisfaire les utilisateurs en termes de qualité de service. Cependant, l'apprentissage par renforcement est un outil d'intelligence artificielle intéressant pour l'analyse des réseaux cellulaires 5G en termes de couverture, de blocage et de débit en raison de sa capacité à apprendre des configurations performantes uniquement à partir des résultats observés. Après avoir développé nos équations de la probabilité d'association et de couverture, dans ce chapitre nous allons discuter des différents résultats de simulation obtenus. En premier lieu, nous étudions la variation du débit pour différentes valeurs des paramètres du Q-Learning et nous identifierons la cause de la grande variabilité dans le débit expérimenté par les réseaux individuels. Nous évaluerons également l'effet des paramètres d'apprentissage intrinsèques de l'algorithme sur cette variabilité en utilisant MATLAB comme logiciel de simulation.

1. Modèle et paramètre de simulation

On considère un scénario dans lequel on dispose de quatre réseaux sans fils différents (WN1, WN2, WN3, WN4) placés dans une carte avec des paramètres qui seront décrits dans le tableau IV.1. Chaque réseau est formé d'un point d'accès (PA) qui transmet à une seule station en liaison descendante.

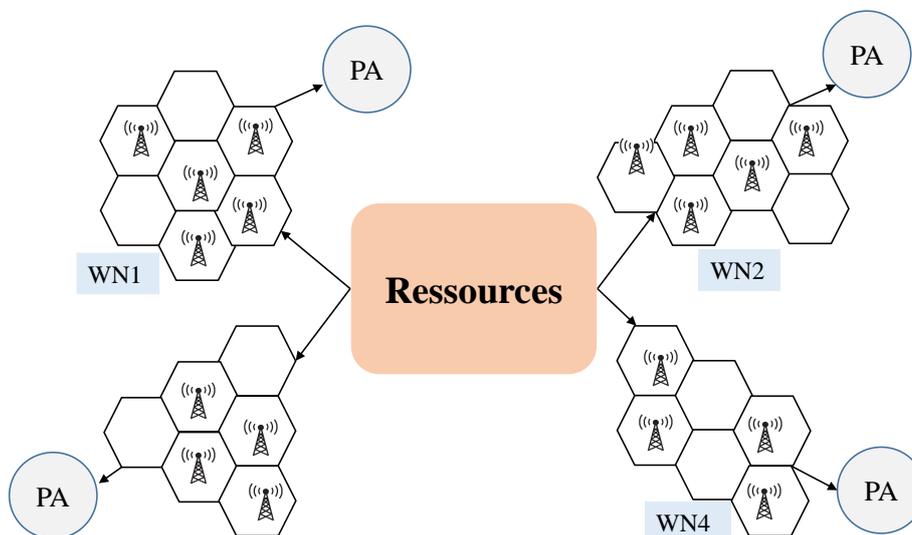


Figure IV. 1 : Carte du réseau.

1.1. Modèle de perte de trajet dans le canal

Les effets d'affaiblissement du chemin et d'ombrage sont modélisés à l'aide du modèle de distance logarithmique pour les communications intérieures. Dans notre modèle la perte de trajet (PL) Alpha Beta Gamma entre les WN i et j est donné par :

$$PL^{ABG}(f, d)[dB] = 10\alpha_{PL} \log_{10}(d) + \beta + 10\gamma_{PL} \log_{10} + X^{ABG} \quad (IV.1)$$

Ou α_{PL} et γ_{PL} sont respectivement des coefficients montrant la dépendance de l'affaiblissement de trajet en fonction de la distance et de la fréquence, β est la valeur de décalage optimisée pour l'affaiblissement de parcours en dB, d est la distance de séparation émetteur-récepteur, f est la fréquence porteuse en GHz et x_σ est l'écart type décrivant les fluctuations du signal à grande échelle par rapport à l'affaiblissement de parcours moyen sur la distance.

Partie 01

1.1. Calcul du débit

Dans cette étude on utilise la puissance reçue et l'interférence pour calculer le débit maximal théorique de chaque réseau sans fil 'i' en fonction du temps 't' en utilisant la capacité de Shannon.

$$\Gamma_{i,t} = B \log_2(1 + SINR_{i,t}) \quad (IV.2)$$

Ou B est la bande passante du canal, et le rapport signal expérimenté sur interférence plus bruit est donné par [40] :

$$SINR_{i,t} = \frac{P_{i,t}}{I_{i,t} + N} \quad (IV.3)$$

D'où $P_{i,t}$ et $I_{i,t}$ sont la puissance reçue et la somme des interférences dans le réseau sans fil 'i' en fonction du temps 't' respectivement, et N est la puissance du bruit de fond.

Pour chaque BTS dans ce réseau l'interférence est considérée comme étant la puissance totale reçue de tous les PA des autres réseaux étendus coexistant et nous considérons que la puissance transmise qui fuit vers les canaux adjacents est inférieure de 20 dBm pour chaque séparation de canaux.

1.2. Fonction de récompense

En se concentrant sur un scénario complètement décentralisé où aucune information sur les autres nœuds n'est disponible, le système peut alors être entièrement décrit par l'ensemble des actions et des récompenses. Ainsi, nous proposons d'utiliser une variante apatriote de l'algorithme d'apprentissage-Q. Pour mettre en œuvre l'apprentissage décentralisé du problème d'allocation des ressources, nous considérons chaque WN comme étant un agent exécutant le Q-Learning sans état par le biais d'une stratégie de sélection d'actions « ϵ -greedy », de sorte que les actions 'a \in A' correspondent à toutes les configurations possibles qui peuvent être choisies en ce qui concerne le canal et la puissance d'émission. Pendant le processus d'apprentissage, nous supposons que les réseaux sélectionnent les actions de manière séquentielle, de sorte qu'à chaque itération d'apprentissage, chaque agent prend une action de manière ordonnée. L'ordre dans lequel les réseaux choisissent une action à chaque itération est choisi aléatoirement au début de celle-ci. La récompense après le choix d'une action est définie comme suit [40] :

$$r_{i,t} = \frac{\Gamma_{i,t}}{\Gamma_i^*} \quad (IV.4)$$

D'où $\Gamma_{i,t}$ est le débit expérimenté au temps t par WN, $i \in \{1, \dots, n\}$, 'n'étant le nombre de WN dans notre scénario, et $\Gamma_i^* = B \log_2(1 + SNR_i)$ est le débit maximal réalisable par le WN.

1.3. Paramètres de simulation

On considère un scénario qui contient 4 réseaux sans fil (WN1, WN2, WN3, WN4) qui forment une topologie dans laquelle les BTS sont placés à la distance maximale possible des autres réseaux (figure IV.1). Ce scénario nous permet d'étudier les performances de l'apprentissage-Q sans état dans un environnement contrôlé, ce qui est utile pour vérifier l'applicabilité de l'apprentissage par renforcement dans les réseaux en utilisant uniquement des informations locales. Et on considère que le nombre de canaux est égal à la moitié du nombre de WN coexistant c'est-à-dire deux canaux, ce qui nous permet d'étudier une situation difficile concernant la réutilisation spatiale. Le tableau IV.1 détaille les paramètres utilisés.

Paramètres	Valeurs
Nombre de réseaux coexistant	4
AP/BTS par WN	1/1
Distance AP-BTS	$\sqrt{2}$
Nombre de canaux	2
Bande de fréquence	30GHz
Modèle initial de sélection des canaux	Distribué uniformément
Valeurs de la puissance transmise (dBm)	{5, 10, 15, 20}
PL_0 (dB)	5
α_{PL}	4.4
D_{obs} (m) (la distance entre deux obstacles)	5
γ_{PL}	0.2
β_{PL}	5
X^{ABG}	5
niveau du bruit (dBm)	100
modèle de trafic	tampon plein (liaison descendante)

Tableau IV.1 : Paramètres de simulation.

1.4. Solution optimale

On identifie d'abord les solutions optimales qui maximisent le débit global, et l'équité proportionnelle, qui est calculée comme la somme logarithmique du débit de chaque WN c'est-à-dire :

$$EP = \max \sum_i \log(\Gamma_{i,k}) \quad (IV.5).$$

1.5. Analyse des paramètres d'entrée

Nous analysons d'abord les effets de la modification des paramètres α (le taux d'apprentissage), γ (le taux d'actualisation) et ϵ_0 (coefficient d'exploration initiale) par rapport au débit du réseau obtenu. On effectue des simulations de 10000 itérations et on capture les résultats des 5000 dernières itérations pour nous assurer que la phase transitoire initiale est terminée. Chaque simulation est répétée 100 fois à des fins de calcul de la moyenne. La figure IV.2 illustre le débit moyen agrégé obtenu pour $\alpha=1$, $\gamma=0.95$ et $\epsilon_0=1$. On peut observer que les meilleurs résultats en ce qui concerne le débit global est de 1440.97 Mbps

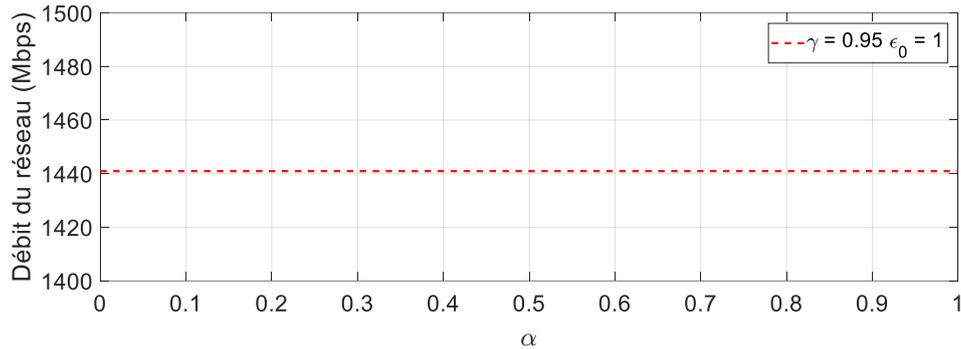
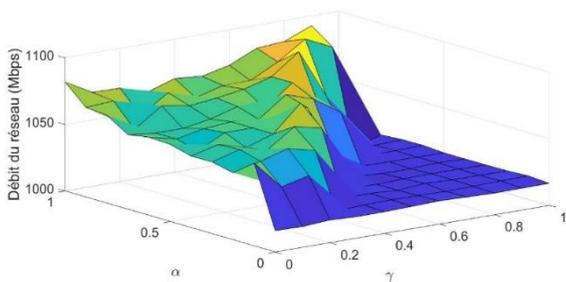
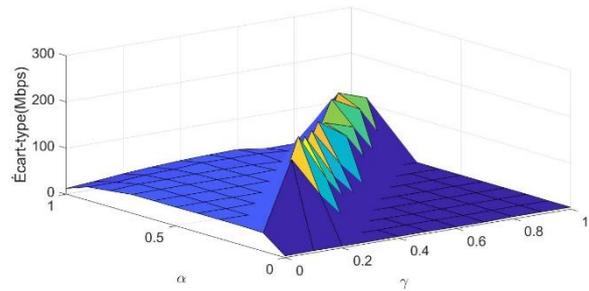


Figure IV.2 : Débit global optimal.

On évalue également la relation entre la valeur α et de γ dans le débit agrégé moyen et l'écart type (illustré à la figure 2) et on remarque un débit agrégé plus grand lorsque $\alpha > \gamma$, et un pic dans l'écart-type lorsque $\gamma \approx \alpha$ et $\gamma > \alpha$.



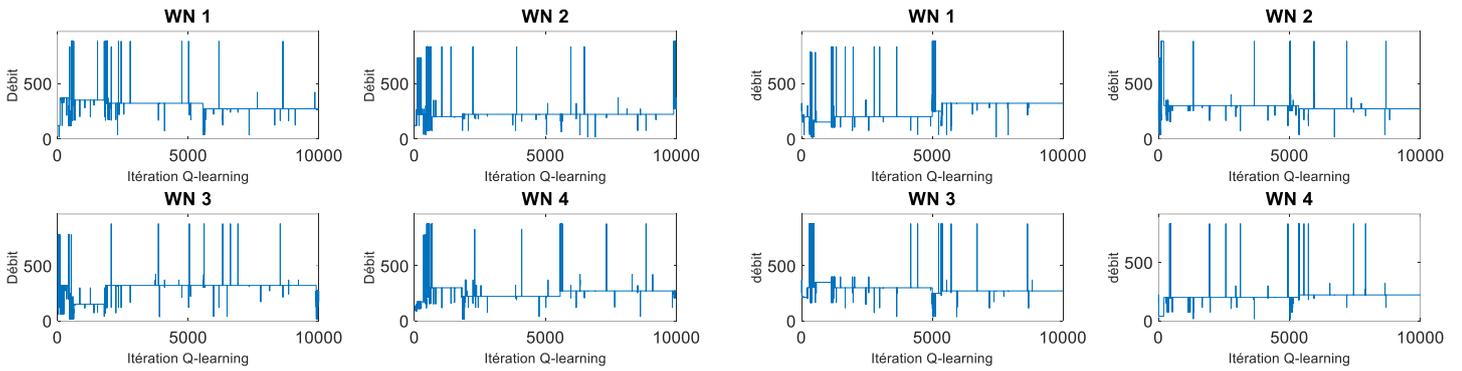
(a) Débit moyen agrégé.



(b) Ecart type.

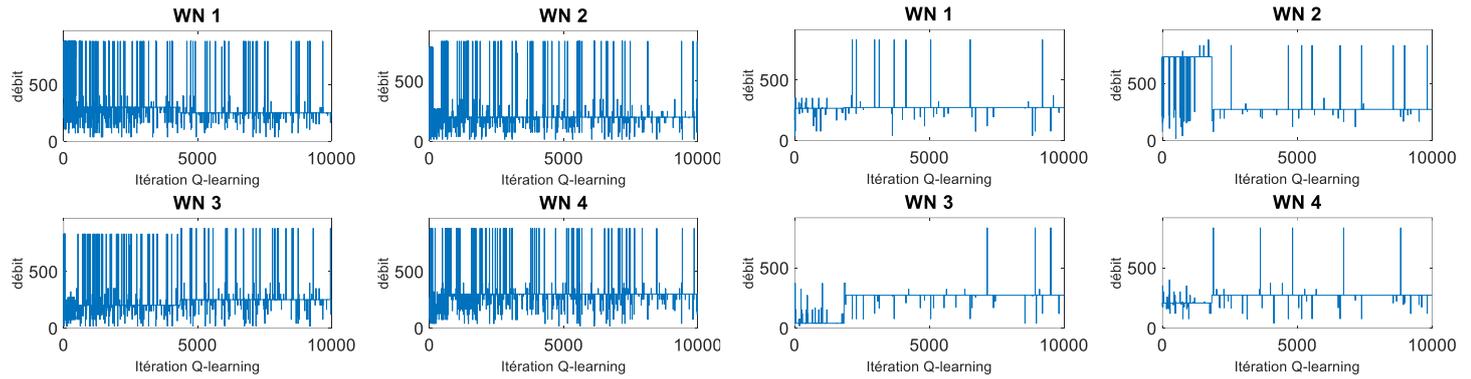
Figure IV.3 : Evaluation de α et γ .

Pour mieux comprendre les effets de la modification de chacun des paramètres susmentionnés, nous montrons le débit individuel expérimenté par chaque WN pour différents α , γ et ϵ_0 pendant les 10000 itérations totales d'une seule exécution de simulation (Figure IV.4), le débit moyen expérimenté par chaque WN pour les 5000 dernières itérations, également pour une seule exécution de simulation (Figure IV.5), et la probabilité de choisir chaque action à chaque WN (figure IV.6).



(a) $\epsilon_0 = 1, \alpha = 1, \gamma = 0.95$

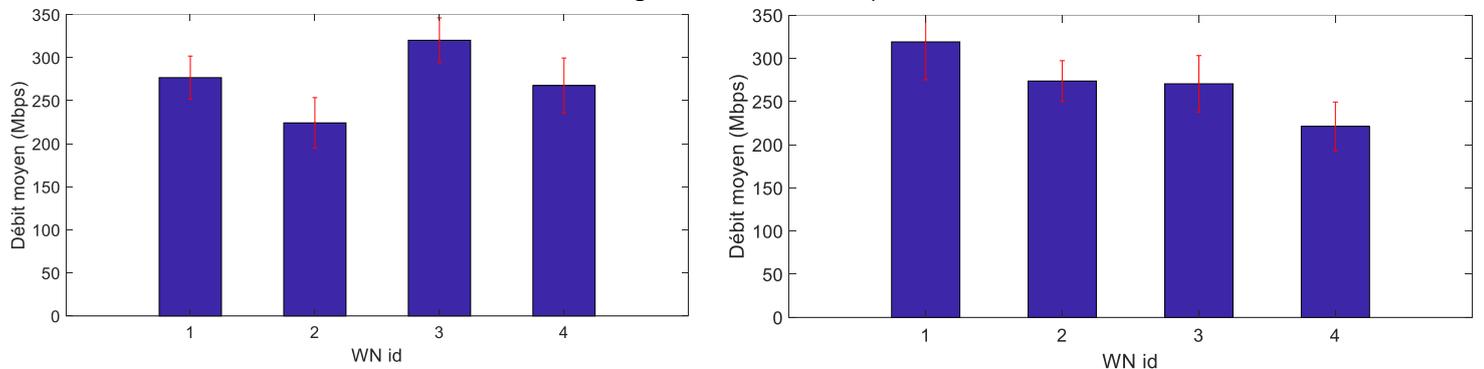
(b) $\epsilon_0 = 0.1, \alpha = 1, \gamma = 0.95$



(c) $\epsilon_0 = 1, \alpha = 0.1, \gamma = 0.05$

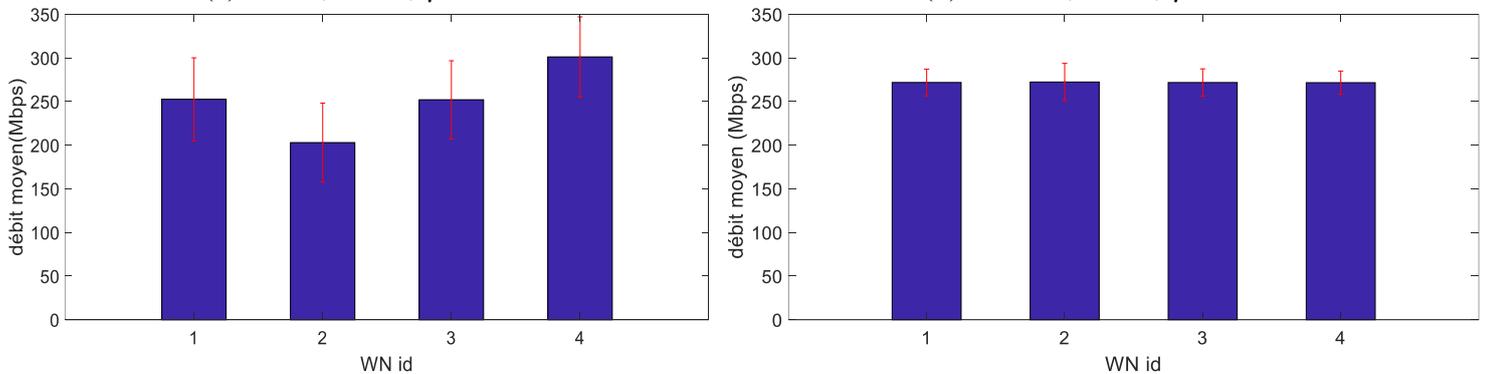
(d) $\epsilon_0 = 0.1, \alpha = 0.1, \gamma = 0.05$

Figure IV.4 : Débit individuel expérimenté par chaque WN au cours d'un seul cycle de simulation pour différents α, γ et ϵ_0 .



(a) $\epsilon_0 = 1, \alpha = 1, \gamma = 0.95$

(b) $\epsilon_0 = 0.1, \alpha = 1, \gamma = 0.95$



(c) $\epsilon_0 = 1, \alpha = 0.1, \gamma = 0.05$

(d) $\epsilon_0 = 0.1, \alpha = 0.1, \gamma = 0.05$

Figure IV.5 : Débit moyen expérimenté par WN pendant les 5000 dernières itérations sur un total de 10000 itérations pour différents α, γ , et ϵ_0 .

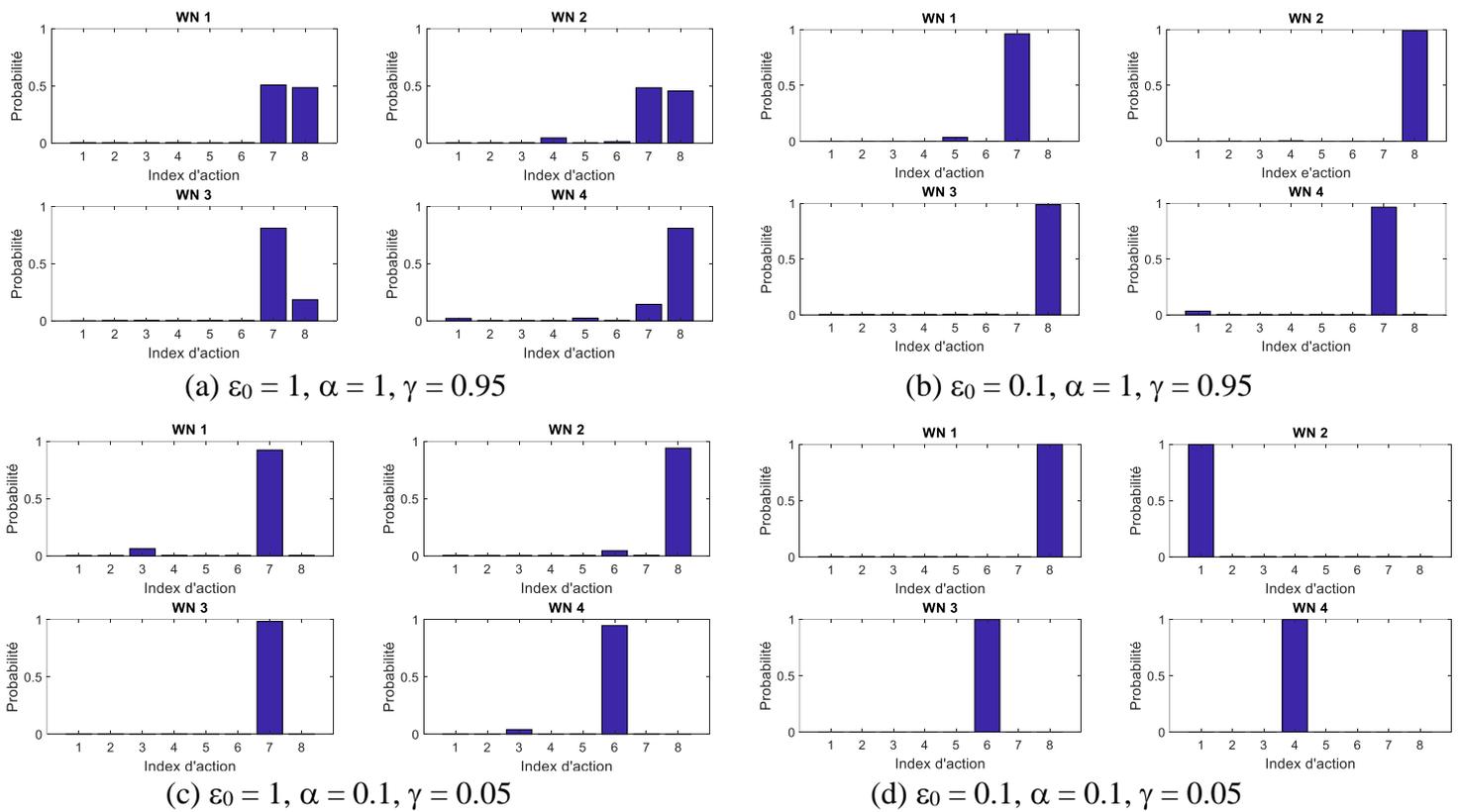


Figure IV.6 : Probabilité de choisir les différentes actions à chaque WN pour une simulation unique (10000 itérations) pour différents α , γ et ϵ_0 .

On peut remarquer par ces résultats les aspects suivants : Dans la figure IV.4, on peut observer une grande variabilité du débit expérimenté par chaque WN, surtout si ϵ_0 est élevé (comme dans la figure IV.4(c)).

Un haut degré d'exploration permet aux WN de découvrir des changements dans la performance résultante de leurs actions en raison de l'activité des autres nœuds, ce qui génère en même temps plus de variabilité c'est-à-dire les WN s'adaptent aux changements en dépit de la variabilité générée et nous obtenons des résultats plus justes pour des ϵ_0 élevés (figure IV.5). Dès lors, il existe une relation entre la variabilité générée et l'équité du débit moyen.

Enfin, dans les figures IV.6(a), IV.6(c) et IV.6(d), nous observons que pour la première, il y a deux actions favorites qui sont jouées le plus, mais pour les deux dernières, il n'y a qu'une seule action favorite donc plus le taux d'apprentissage (α) le facteur d'actualisation (γ), sont faible, plus la probabilité de choisir une action unique, qui s'avère être celle qui a fourni la meilleure performance dans le passé, est élevée. Le contraire se produit pour des valeurs α et γ plus élevées, puisque le fait de donner plus d'importance à la récompense immédiate permet de ne réagir qu'aux actions récemment jouées par les nœuds voisins.

Partie 02

2.1. Calcul de l'efficacité énergétique

Dans cette partie on fait en sorte que la récompense obtenue par ce programme d'apprentissage par renforcement est l'efficacité énergétique des réseaux avec les mêmes valeurs des paramètres de simulation et on observe les résultats obtenue, d'où notre équation de récompense est donnée comme suit :

$$EE = \frac{\Gamma_{i,t}}{(P_E + P_C)} \quad (\text{bits/J}) \quad (\text{IV.6})$$

Avec $\Gamma_{i,t}$ le débit maximal théorique de chaque réseau sans fil 'i' en fonction du temps 't' calculé dans la partie 01, P_E P_C sont respectivement la puissance émise et la puissance du circuit.

2.2. Résultats obtenus

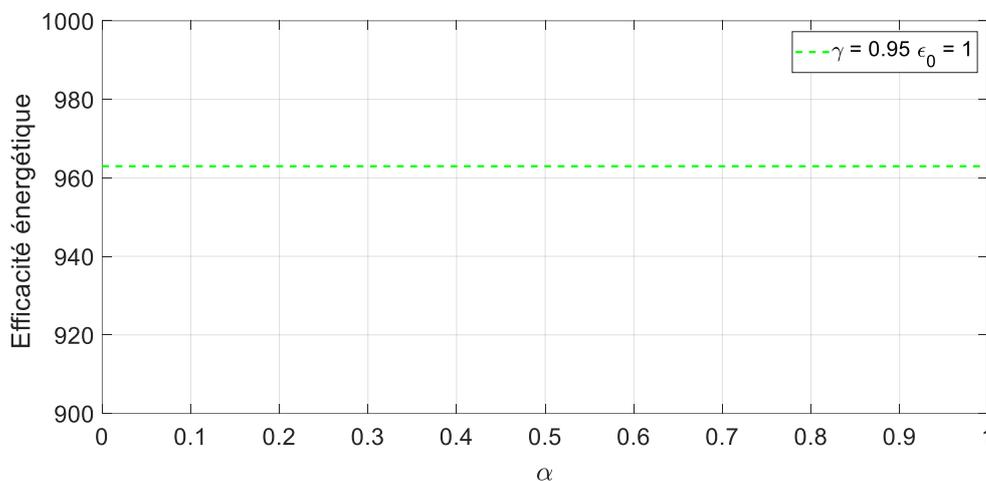


Figure IV.7 : Efficacité énergétique globale.

Dans la figure IV.7, nous observons la variation de l'efficacité énergétique en fonction du taux d'apprentissage pour une valeur d'un taux d'actualisation et le coefficient d'exploration initiale égale à 0.95 et 1 respectivement. D'après la Figure IV.7, l'efficacité énergétique est stable pour une variation de α . Cela s'explique par le fait que la meilleure valeur de l'efficacité énergétique a été atteinte pour un $\gamma=0.95$ et $\epsilon_0=1$.

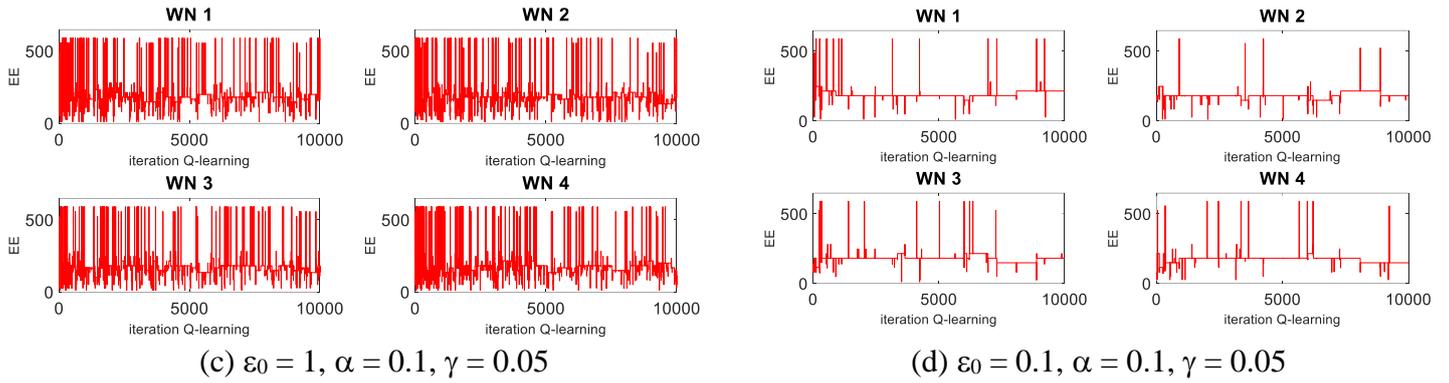
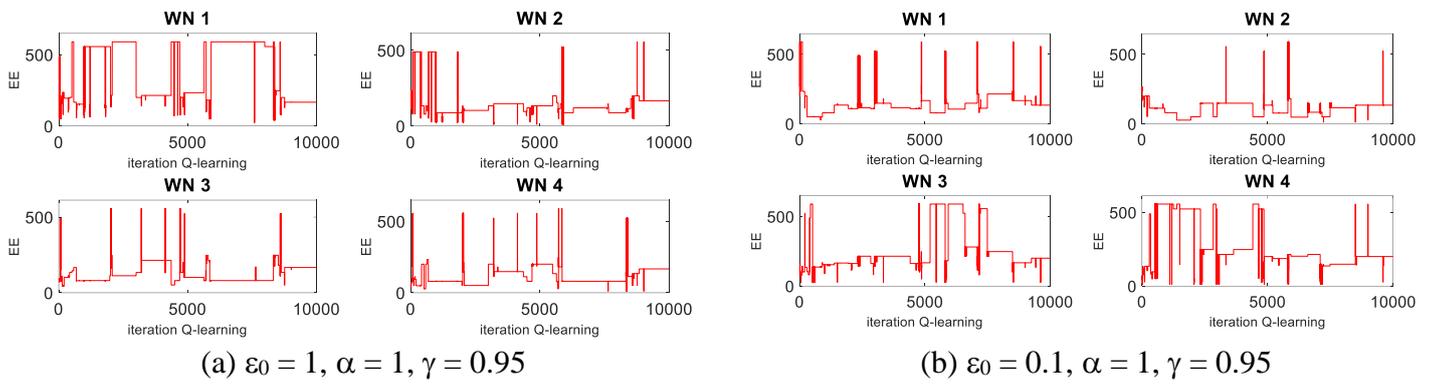


Figure IV.8 : Efficacité énergétique individuel expérimenté par chaque WN au cours d'un seul cycle de simulation pour différents pour différents α, γ et ε_0 .

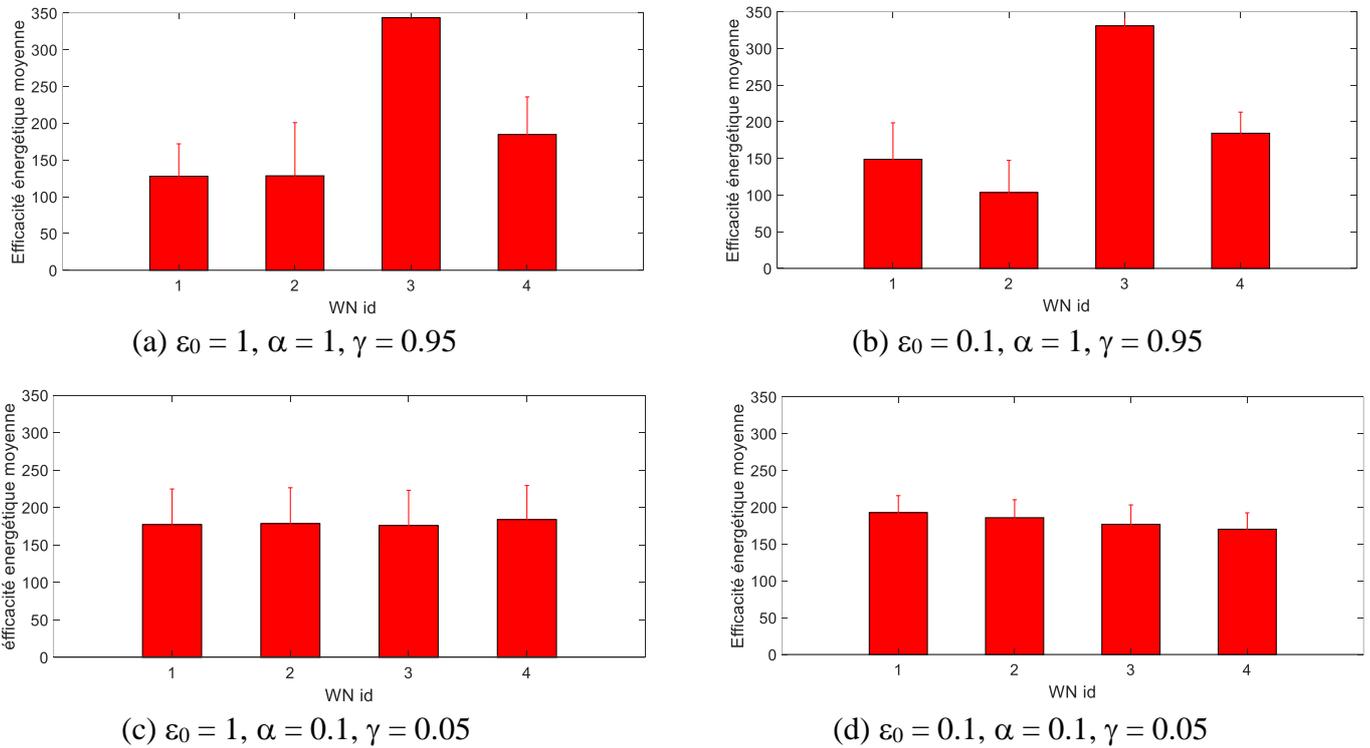


Figure IV.9 : Efficacité énergétique moyenne expérimenté par WN pendant les 5000 dernières itérations sur un total de 10000 itérations pour différents α, γ et ε_0 .

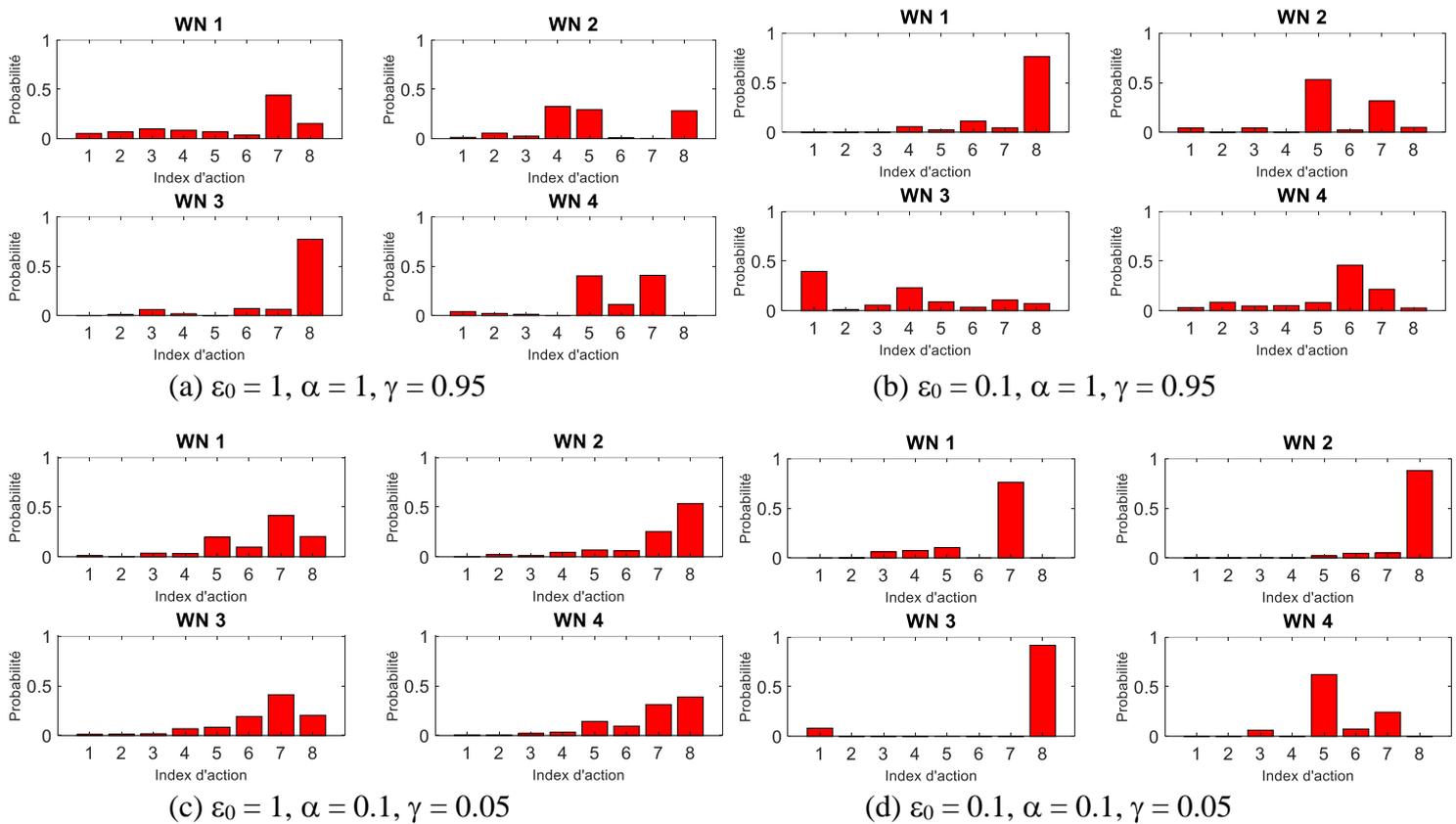


Figure IV.10 : Probabilité de choisir les différentes actions à chaque WN pour une simulation unique (10000 itérations) pour différents α , γ et ϵ_0 .

La Figure IV.8 illustre la variation d'efficacité énergétique en fonction des itérations de l'algorithme proposé dans chaque WN au cours d'un seul cycle de simulation. L'observation courante de cette figure est que les paramètres ϵ_0 et γ ont une influence directe sur l'efficacité énergétique. Cela peut être expliqué par le fait que l'agent s'adapte aux activités des nœuds, il apprend de l'environnement étudié en attribuant des puissances convenables pour maximiser le débit. Par conséquent, l'efficacité énergétique est améliorée.

Par contre on remarque une variation dans la figure IV.9 lorsque les valeurs de α et γ sont élevées et des résultats plus ou moins stable pour des valeurs plus faibles. Ce qui explique qu'il existe une relation entre la variabilité et l'équité de l'efficacité énergétique.

D'autre part, chaque action a une probabilité de réussite différente comme le montre la figure IV.10. En effet, l'obtention d'une efficacité énergétique imperméable peut se faire dans les premières actions de l'agent, la probabilité de choisir une action particulière, qui fournit la meilleure performance dans le passé, est élevée. Néanmoins, d'autres actions peuvent être à faible probabilité après l'obtention d'un meilleur résultat. Cela est dû principalement au changement dans l'environnement du réseau qui reste dynamique surtout en matière de propagation.

Conclusion

Dans ce chapitre on a présenté les différentes étapes de simulation de notre modèle de propagation dans un environnement complètement décentralisé, et on a montré nos résultats tels que le débit optimal obtenu et l'efficacité énergétique. Ainsi, nous avons également identifié une grande variabilité dans le débit individuel expérimenté et l'EE en raison des changements constants des actions jouées.

Nous avons évalué l'impact des paramètres de l'algorithme d'apprentissage α γ et ϵ_0 . Sur cette variabilité montrant qu'elle peut être réduite en diminuant le degré d'exploration et le taux d'apprentissage.

Nos résultats montrent que, bien que l'algorithme soit capable de trouver les actions les plus performantes pour améliorer le débit global, et l'efficacité énergétique il existe une grande variabilité dans les résultats expérimenté par les réseaux individuels.

Conclusion générale

La croissance explosive du nombre d'appareils mobiles et des applications Internet mobiles apporte un grand confort à la société. Il crée également une énorme demande de trafic qui remet en question la conception des systèmes de communication mobile de prochaine génération.

Ce projet de fin d'étude aborde la nouvelle génération des réseaux mobiles 5G. C'est une étude en perspective générale de l'allocation de ressource dans un réseau 5G hétérogène en utilisant l'apprentissage par renforcement avec l'algorithme « decentralized Q-Learning » qui est une méthode d'amélioration de machine Learning consistant à résoudre des problèmes tel que l'association des points d'accès, la sélection des canaux et l'ajustement de la puissance d'émission et qui permet de calculer le gain potentiel. La nécessité de l'amélioration de taux de transmission de donnée (le débit) nous a poussés à chercher à améliorer les performances de ce système de communication et de faire des études approfondies pour des solutions meilleures dans les futures normes.

Notre travail est divisé en quatre chapitres dont le premier est une présentation des réseaux mobiles de nouvelle génération ainsi que leurs techniques de transmission telle que la propagation LOS et NLOS, la technique MIMO et Massive MIMO et le Beamforming. On a aussi abordé les types de canaux tels que Rayleigh, Rice et Nakagami et les techniques d'accès à ces derniers qui sont L'OFDM, OFDMA, NOMA et OMA ayant pour but de partager les ressources spectrales entre plusieurs utilisateurs d'un système radio mobile. Et en dernier on a abordé le problème d'interférences inter et intracellulaires et comment essayer de les diminuer par les coordinations FFR et SFR.

Dans le second chapitre, on a expliqué l'allocation de ressource dans un réseau 5G, son principe et ses objectifs. Ainsi que les ordonnanceurs d'allocation de ressource. On a présenté aussi l'efficacité spectrale et comment l'améliorer et l'efficacité énergétique et ses enjeux. De plus, on a parlé sur le délai de tolérance dans un réseau 5G.

Dans le troisième chapitre, on s'est basé sur les concepts de base de notre simulation, à savoir l'importance de l'apprentissage par renforcement dans les communications sans fil permettant de faire de bons choix de conception futurs. On s'est intéressé sur un réseau hétérogène précisément avec un modèle décentralisé tout en introduisant notre modèle mathématique afin de calculer la probabilité de couverture de notre réseau, améliorer notre débit et notre efficacité énergétique.

Dans le dernier chapitre, on a analysé les différents résultats obtenus lors de notre simulation qui s'est porté sur l'étude de la variation des paramètres du Q-Learning pour pouvoir trouver la meilleure politique pour notre étude.

Références

- [1] SAUTER, Martin. From GSM to LTE-advanced Pro and 5G : An introduction to mobile networks and mobile broadband. John Wiley & Sons, 2017.
- [2] ANDREWS, Jeffrey G., BUZZI, Stefano, CHOI, Wan, et al. What will 5G be?. IEEE Journal on selected areas in communications, 2014, vol. 32, no 6, p. 1065-1082.
- [3] ARNAU, Jesús, ATZENI, Italo, et KOUNTOURIS, Marios. Impact of LOS/NLOS propagation and path loss in ultra-dense cellular networks. In : 2016 IEEE international conference on communications (ICC). IEEE, 2016. p. 1-6.
- [4] OSSEIRAN, Afif, MONSERRAT, Jose F., et MARSCH, Patrick (Ed). 5G mobile and wireless communications technology. Cambridge University Press, 2016.
- [5] SIBILLE, Alain, OESTGES, Claude, et ZANELLA, Alberto. MIMO: from theory to implementation. Academic Press, 2010.
- [6] VAN VEEN, Barry D. et BUCKLEY, Kevin M. Beamforming : A versatile approach to spatial filtering. IEEE assp magazine, 1988, vol. 5, no 2, p. 4-24.
- [7] MAHMUD, Azwan, LIN, Zhangxian, et HAMDI, Khairi Ashour. On the energy efficiency of fractional frequency reuse techniques. In : 2014 IEEE Wireless Communications and Networking Conference (WCNC). IEEE, 2014. p. 2348-2353.
- [8] CAMPO, C., STEFER, M., BOEGLIN, H., et al. Système de contrôle d'antennes pour un réseau linéaire uniforme basé sur la radio logicielle. In : XXèmes Journées Nationales Microondes. 2017.
- [9] LI, Ye Geoffrey et STUBER, Gordon L. (ed.). Orthogonal frequency division multiplexing for wireless communications. Springer Science & Business Media, 2006.
- [10] JIANG, Tao, SONG, Lingyang, et ZHANG, Yan (ed.). Orthogonal frequency division multiple access fundamentals and applications. CRC Press, 2010.
- [11] HASHIMA, Sherief, MUTA, Osamu, ELNOUBI, Said, et al. Analysis and comparison of Fractional Frequency Reuse schemes based on worst case Signal to Interference power Ratio in OFDMA uplink. In : WTS. 2014. p. 1-6.
- [12] EZZAOUIA, Mahdi. Allocation de ressource opportuniste dans les réseaux sans fil multicellulaires. 2018. Thèse de doctorat. Ecole nationale supérieure Mines-Télécom Atlantique Bretagne Pays de la Loire.
- [13] PROAKIS, John G. et SALEHI, Masoud. *Digital communications*. New York: McGraw-hill, 2001.
- [14] SABRA, Mme BENKRINAH et OUARGLA, UKM. Performances d'un système de communication sans fil à base de la technique FBMC pour la 5G.

- [15] Z. Kadri, S. Dib, “Techniques de positionnement dans le réseau 5G”, université A. Mira-Bejaia, 2019/2020.
- [16] “What is the difference between Nakagami-m and general fading?”, November 2020.
- [17] ANDRIAMANALINA, A. Modélisation des différents canaux des systèmes de communication numérique. 2010.
- [18] “Why is Nakagami-m fading channel considered? What does m indicate? Where it is we can consider?”, December 2015.
- [19] LIM, Yujin, FERRARI, Gianluigi, TAKAHASHI, Hideyuki, et al. Energy efficient wireless networks. *Wireless Communications and Mobile Computing*, 2019, vol. 2019.
- [20] LÓPEZ-PÉREZ, David, DE DOMENICO, Antonio, PIOVESAN, Nicola, *et al.* A Survey on 5G Radio Access Network Energy Efficiency: Massive MIMO, Lean Carrier Design, Sleep Modes, and Machine Learning. *IEEE Communications Surveys & Tutorials*, 2022.
- [21] DEMESTICHAS, Konstantinos, ADAMOPOULOU, Evgenia, et CHORAŚ, Michał. « 5G communications : Energy efficiency ». *Mobile Information Systems*, 2017, vol. 2017.
- [22] JAOUADI, Randa. *Compromis efficacité énergétique et efficacité spectrale pour les objets communicants autonomes*. 2017. Thèse de doctorat. Nantes.
- [23] JONES, Christine E., SIVALINGAM, Krishna M., AGRAWAL, Prathima, et al. A survey of energy efficient network protocols for wireless networks. *Wireless networks*, 2001, vol. 7, no 4, p. 343-358.
- [24] OLANREWAJU, B. S. Et OSUNADE, O. Survey of Related Technologies for Improved Spectrum Efficiency.
- [25] LEE, William CY. Spectrum efficiency in cellular (radio). *IEEE Transactions on Vehicular Technology*, 1989, vol. 38, no 2, p. 69-75.
- [26] WANG, Tan, LI, Gen, DING, Jiaxin, *et al.* 5G spectrum: Is China ready?. *IEEE Communications Magazine*, 2015, vol. 53, no 7, p. 58-65.
- [27] SGORA, Aggeliki. 5G spectrum and regulatory policy in Europe: an overview. In: *2018 Global Information Infrastructure and Networking Symposium (GIIS)*. IEEE, 2018. p. 1-5.
- [28] DONG, Rui, SHE, Changyang, HARDJAWANA, Wibowo, et al. Deep learning for radio resource allocation with diverse quality-of-service requirements in 5G. *IEEE Transactions on Wireless Communications*, 2020, vol. 20, no 4, p. 2309-2324.
- [29] CHAOUCH, S. Gestion Des Ressources Des Réseaux Cloud RAN Dans Un Contexte 5G. 2017. Thèse de doctorat. Thèse de doctorat, Télécom SudParis.
- [30] ROY, Animesh, ACHARYA, Tamaghna, et DASBIT, Sipra. Quality of service in delay tolerant networks: A survey. *Computer Networks*, 2018, vol. 130, p. 121-133.

- [31] VASILAKOS, Athanasios, ZHANG, Yan, et SPYROPOULOS, Thrasyvoulos. Delay Tolerant Networks. CRC press, 2016.
- [32] MANGAYARKARASI, P., RAMYA, M., et JAYASHRI, S. Analysis of various power allocation algorithms for wireless networks. In: 2012 International Conference on Communication and Signal Processing. IEEE, 2012. p. 133-136.
- [33] BUZZI, Stefano, CHIH-LIN, I., KLEIN, Thierry E., *et al.* A survey of energy-efficient techniques for 5G networks and challenges ahead. *IEEE Journal on Selected Areas in Communications*, 2016, vol. 34, no 4, p. 697-709.
- [34] S. Zenadji, “Allocation de ressources dans une cellule 5G tri-sectorielle ”, Thèse de doctorat, Université A. MIRA Bejaia, 2020.
- [35] Benghabrit, “Allocation de ressources dans les réseaux radio cognitifs”, Thèse de doctorat, Université Oran1 Ahmed Ben Bella, Mars 2019.
- [36] EZZAOUIA, Mahdi. *Allocation de ressource opportuniste dans les réseaux sans fil multicellulaires*. 2018. Thèse de doctorat. Ecole nationale supérieure Mines-Télécom Atlantique Bretagne Pays de la Loire.
- [37] LEE, Woongsup, KIM, Minhoe, et CHO, Dong-Ho. Deep power control: Transmit power control scheme based on convolutional neural network. *IEEE Communications Letters*, 2018, vol. 22, no 6, p. 1276-1279.
- [38] MOUSSAID, Achraf. Appariement intelligent des appareils D2D sous-jacents aux réseaux cellulaires et activés pour le cache. 2020.
- [39] ZIMMER, Matthieu. Apprentissage par renforcement développemental. 2018. Thèse de doctorat. Université de Lorraine.
- [40] WILHELMI, Francesc, BELLALTA, Boris, CANO, Cristina, *et al.* Implications of decentralized Q-learning resource allocation in wireless networks. In: *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*. IEEE, 2017. p. 1-5.

Résumé

Le réseau 5G est actuellement considéré comme une technologie émergente afin d'optimiser la capacité et accroître le débit. Cependant, avec l'augmentation des utilisateurs, les performances du réseau doivent être améliorées. Ce travail présente une étude de l'allocation de ressources dans un réseau 5G hétérogène en appliquant l'apprentissage par renforcement pour objectif d'améliorer le débit et l'efficacité énergétique et de l'évaluer en termes de probabilité. Le logiciel MATLAB nous a permis de réaliser cette analyse et de comprendre certains aspects liés à notre étude.

Mots clés : 5G, OFDMA, NOMA, OMA, LOS, NLOS, Beamforming, SINR, Apprentissage par renforcement, Q-Learning, Efficacité énergétique, MARKOV, BELLMAN.

Abstract

The 5G network is currently considered as an emerging technology to optimize capacity and increase throughput. However, with the increase of users, the network performance needs to be improved. This work presents a study of resource allocation in a heterogeneous 5G network by applying reinforcement learning with the objective of improving throughput and energy efficiency and evaluating it in terms of probability. The MATLAB software allowed us to perform this analysis and to understand some aspects related to our study.

Keywords: 5G, OFDMA, NOMA, OMA, LOS, NLOS, Beamforming, SINR, Reinforcement learning, Q Learning, Energy efficiency, MARKOV, BELLMAN.